



## Prediction of Tenderness, Juiciness, and Flavor Profile of 2 Beef Muscles with Different Aging Times Using Rapid Evaporative Ionization Mass Spectrometry (REIMS)

Michael J. Hernandez-Sintharakao<sup>1</sup>, Chandler J. Sarchet<sup>2</sup>, Jessica E. Prenni<sup>3</sup>, Dale R. Woerner<sup>2</sup>, Chaoyu Zhai<sup>4</sup>, and Mahesh N. Nair<sup>1\*</sup>

<sup>1</sup>Center for Meat Safety & Quality, Department of Animal Sciences, Colorado State University, Fort Collins, CO 80523, USA

<sup>2</sup>Department of Animal and Food Sciences, Texas Tech University, Lubbock, TX 79409, USA

<sup>3</sup>Department of Horticulture & Landscape Architecture, Colorado State University, Fort Collins, CO 80521, USA

<sup>4</sup>Department of Animal Science, University of Connecticut, Storrs, CT 06269, USA

\*Corresponding author. Email: [mahesh.narayanan\\_nair@colostate.edu](mailto:mahesh.narayanan_nair@colostate.edu) (Mahesh N. Nair)

**Abstract:** Rapid evaporative ionization mass spectrometry (REIMS) is a novel technique that provides rapid chemical information on biological tissues and has the potential to predict beef quality attributes in real time. This study aims to assess the ability of analysis by REIMS coupled with chemometric modeling to predict the quality attributes of wet-aged beef at the grading time. USDA Select and upper two-thirds Choice ( $n = 42$ ,  $N = 84$ ) striploins (longissimus lumborum [LL]) and tenderloins (psoas major [PM]) were collected 36 h postmortem from a commercial beef abattoir. The LL and PM were cut into portions and aged for 3, 14, and 28 d. Aged samples were analyzed for slice shear force, Warner-Bratzler shear force (WBS), and trained sensory panels (tenderness, juiciness, and flavor attributes), and results were used to categorize both LL and PM into binary classes. Additionally, slivers of the longissimus dorsi muscle between the 12th and 13th rib were collected during grading (36 h postmortem) and analyzed using REIMS. The REIMS data were used to build predictive models for tenderness, juiciness, and flavor classes for the 3 aging periods and 2 muscles. Prediction accuracies of all models were higher than classifying the samples by chance ( $P < 0.05$ ), except WBS of 3 d aging model ( $P > 0.05$ ). However, model accuracies were not too high, which could be due to overlaps between classes, small sample sizes, and unbalanced data, which could negatively affect predictive models. Results demonstrated that the chemical fingerprints obtained with REIMS could potentially sort carcasses by flavor, juiciness, and tenderness in real time. However, the full realization of this approach will require an increased sample size and the development of a sampling method that allows improved separation between sensory classes.

**Key words:** rapid evaporative ionization mass spectrometry, beef, tenderness, flavor, juiciness

*Meat and Muscle Biology* 7(1): 16120, 1–14 (2023)

doi:10.22175/mmb.16120

Submitted 26 February 2023

Accepted 12 May 2023

## Introduction

Beef palatability depends on tenderness, juiciness, and flavor. Tenderness has been previously defined as the most important quality trait for consumer satisfaction when consuming beef (Wheeler et al., 1990; Miller et al., 1995). However, when tenderness is acceptable, beef flavor plays a major role in the

overall eating experience (Goodson et al., 2002; Behrends et al., 2005; Legako et al., 2015; Liu et al., 2020). In addition, juiciness contributes 7.4% to the overall palatability of beef (O'Quinn et al., 2018). The USDA quality grade attempts to predict palatability attributes of beef based on carcass traits, including sex, lean texture/firmness, marbling score, and animal maturity, although the marbling score is the main determinant. The USDA quality grade segregates

carcasses by the probability of having a positive eating experience (O'Quinn et al., 2018). However, the instrumental marbling score explains less than 45% variability in juiciness, 40% variability in tenderness, and 61% overall palatability of the longissimus muscle (Smith et al., 1985; Emerson et al., 2013). Therefore, the current grading system does not account for other significant sources of variation in tenderness and juiciness.

Real-time assessment of beef quality traits (flavor, tenderness, and juiciness) at production speeds is challenging, and implementation of techniques such as sensory analysis or shear force measurement is impractical. Large commercial beef facilities process approximately 4,000 to 6,000 heads of beef cattle per day (MacDonald et al., 2000; DeOtte et al., 2010). Current methods to assess beef quality attributes are slow compared with the production rate of commercial facilities. Mechanical methods for measuring beef tenderness, such as slice shear force (SSF) or Warner-Bratzler shear force (WBS), have been suggested to be implemented in line (Shackelford et al., 1999). However, mechanical methods are not popular among producers because a 1-inch ribeye steak is wasted during the analysis (Wheeler et al., 2002).

Nondestructive techniques that can predict beef palatability traits (flavor, tenderness, and juiciness) in real time could improve the current grading system. Rapid evaporative ionization mass spectrometry (REIMS) is an ambient mass spectrometry technique that enables rapid analysis of biological tissues in situ, without sample preparation (Balog et al., 2016). Because this method does not require sample preparation, data collection takes only a few seconds, making REIMS a good alternative for assessing quality attributes in production lines. Previously, REIMS analysis coupled with chemometric modeling was used to classify beef striploins samples aged 14 d into “tough” and “tender” classes defined by a cutoff of 20 kg of SSF for tenderness. Results demonstrated that this technique could classify beef samples into the 2 tenderness categories with 91% accuracy (Gredell et al., 2019). However, the analysis for this study was performed on steaks aged 14 d, so the potential of REIMS data to be used to predict quality during grading is still unknown. Therefore, this study aimed to evaluate the ability of REIMS analysis coupled with chemometric modeling as a real-time method to predict meat tenderness, juiciness, and flavor of aged meat using data collected at the grading time.

## Materials and Methods

Product collection, aging, and measurement of sensory attributes were performed at Texas Tech

University. REIMS data acquisition and data analysis were performed at Colorado State University (CSU).

### *Product collection and aging*

Forty-two ( $n = 42$ ) USDA Select and 42 USDA upper 2/3 of the Choice grade (High and Average Choice) were selected from a commercial beef processing facility over 3 production days. Slivers of the longissimus lumborum (LL) muscle were collected between the 12th and 13th ribs from both sides of the carcasses at the time of grading (GR; around 36 h postmortem), frozen with liquid nitrogen, transported to CSU in coolers with dry ice, and stored at  $-80^{\circ}\text{C}$  until further analysis. After grading, whole striploins (LL) and tenderloins (psoas major [PM]) were collected from both sides of the carcasses and transported in coolers ( $2^{\circ}\text{C}$  to  $4^{\circ}\text{C}$ ) to the Texas Tech University meat laboratory. Upon arrival, LL (from both sides of the carcasses) was fabricated into 6-cm sections. The gluteus medius from the posterior part of the LL was excluded. Similarly, heavy connective tissue was removed from the PM from both sides and fabricated into 9-cm sections. Each section per muscle was vacuum sealed and randomly assigned to 1 of the 3 aging periods (3, 14, and 28 d). Portions were aged at  $0^{\circ}\text{C}$  to  $2^{\circ}\text{C}$  for the corresponding aging period and were stored at  $-80^{\circ}\text{C}$  until further analysis. Frozen LL sections were fabricated into two 2.54-cm steaks using a band saw, and each steak was randomly assigned to shear force and sensory panels. Likewise, frozen PM sections were fabricated into three 2.54-cm steaks, and two steaks were randomly assigned to sensory panels and the remaining steak to shear force.

### *Cooking procedure*

Steaks for sensory panels and shear force evaluation were cooked using the same procedure. Frozen steaks were thawed at  $0^{\circ}\text{C}$  to  $2^{\circ}\text{C}$  for 24 to 48 h to attain raw internal temperatures of  $0^{\circ}\text{C}$  to  $2^{\circ}\text{C}$  before cooking. Steaks were cooked in an oven (Model SCC WE 61 E; Rational, Landsberg am Lech, Germany) at  $204^{\circ}\text{C}$  and 0% relative humidity to a peak internal temperature of  $71^{\circ}\text{C}$ . Peak internal temperatures were recorded in the geometric center of the steaks using a thermometer (AccuTuff 34032, Cooper-Atkins Corporation, Middlefield, CT).

### *Trained sensory panel*

The LL and PM sensory panels were performed similarly but separately. Panels of the LL were performed in 42 sessions and the PM in 44 sessions, with

a maximum of 2 sessions/day, a resting time of at least 7 h between sessions, a maximum of 12 samples/session, and at least 8 trained panelists/session. All samples ( $n = 504$  per USDA grade) were randomly assigned to each session, ensuring that both USDA quality grades and all aging treatments were present. Immediately after cooking, external fat and connective tissue were removed, and the steaks were cut into small cubes. Two to three cubes were served to the panelists

**Table 1.** Definition and reference standards for beef sensory attributes and their intensities based on Adhikari et al. (2011), where 0 = “not present,” “extremely dry,” or “extremely tough,” and 100 = “extremely intense,” “extremely juicy,” or “extremely tender”

Attribute	Definition	Reference
<b>Tenderness</b>	The overall tenderness of the sample	Strip steak to 71°C = 60 Tenderloin to 71°C = 95
<b>Juiciness</b>	The amount of perceived juice that is released from the product during mastication	Carrot = 55 Strip steak cooked to 79°C = 60 Strip steak cooked to 57°C = 75 Watermelon = 95
<b>Beef Flavor ID</b>	Amount of beef flavor identity in the sample	Swanson Beef Broth = 35 Beef brisket (71°C) = 80
<b>Bitter</b>	The fundamental taste factor associated with a caffeine solution	0.01% caffeine solution = 15 0.02% caffeine solution = 25
<b>Browned</b>	Aromatic associated with the outside of grilled or broiled meat; seared but not blackened or burnt	Beef suet (broiled) = 55
<b>Buttery</b>	Sweet, dairy-like aromatic associated with natural butter	Land O'Lakes unsalted butter = 45
<b>Fat-Like</b>	The aromatics associated with cooked animal fat	Hillshire Farm Beef Lit'l Smokies = 45 Beef suet = 80
<b>Liver-Like</b>	The aromatics associated with cooked organ meat/liver	Beef liver = 50
<b>Metallic</b>	The impression of slightly oxidized metal, such as iron, copper, and silver spoons	0.10% potassium chloride solution = 10 Select striploin steak (60°C internal) = 25 Dole canned pineapple juice = 40
<b>Roasted</b>	Aromatic associated with roasted meat	80% lean ground chuck = 65
<b>Sour</b>	The fundamental taste factor associated with citric acid	0.015% citric acid solution = 10 0.050% citric acid solution = 25
<b>Umami</b>	Flat, salty, somewhat brothy. The taste of glutamate, salts of amino acids, and other molecules called nucleotides	0.035% Accent Flavor Enhancer solution = 50

in individual booths equipped with a red incandescent light. Panelists evaluated tenderness, juiciness, and 12 flavor descriptors from the beef lexicon (Adhikari et al., 2011) described in Table 1 on a 100-mm unstructured line scale. For all flavor attributes, the left end of the line scale was anchored with “not present,” and the right end represented “extremely intense.” For juiciness and tenderness, the left end was fixed with “extremely dry” and “extremely tough,” and the right one with “extremely juicy” and “extremely tender,” respectively. A warm-up sample (USDA Low Choice strip steak) was served at the beginning of each session to calibrate panelists. Consensus on the attributes of the warm-up sample was reached before moving to the study samples.

### Shear force

The tenderness of both muscles was evaluated using WBS and SSF as described by Shackelford et al. (1999) and Lorenzen et al. (2010). Samples were cooked, and internal peak temperatures were recorded following the protocol described. Within 2 to 3 min after recording peak temperature, the lateral ends of the steak were removed, and a 1 × 5 cm slice was cut parallel to the muscle fiber. SSF was measured by shearing the slice perpendicular to muscle fibers with a universal testing machine (Instron Corp., Norwood, MA) equipped with a flat, blunt-end blade (500 mm/min crosshead speed and 100 kg load capacity). The remaining part of the sample was cooled down to room temperature or below, and 2 to 6 cores (1.2 cm diameter) were removed parallel to the muscle fiber orientation. The WBS was measured by shearing each core perpendicular to the fiber with the universal testing machine equipped with a WBS blade (74.2 mm wide × 138.7 mm long × 0.99 mm thick, 200 mm/min crosshead speed, and 100 kg load cell capacity).

### Rapid evaporative ionization mass spectrometry

Analysis of the GR samples was performed using a quadrupole time-of-flight mass spectrometer (Synapt G2-Si Q-ToF, Waters Corporation, Milford, MA) equipped with a REIMS source (Waters Corporation). An electronic probe (Waters Corporation) powered by an electrosurgical generator (Erbe VIO 50C, Erbe Elektromedizin GmbH, Tübingen, Germany) was used as the sampling device. The electrosurgical generator was set to dry cut mode and maximum cutting power of 40 W. Mass spectra from 100 to 1,500  $m/z$  were acquired in negative ion mode, with cone voltage at

40 V and heater bias at 60 V. Samples were thawed at 0°C to 4°C for 16 to 24 h and randomly sorted for REIMS analysis. The sampling device was used to generate at least 5 burns over the surface of individual samples. Burns were made in the 4 corners and the middle of a 2.5 by 2.5 cm square from the surface of the samples, and each burn lasted approximately 1 s. Extra burns were made in a random location equidistant to the last burn when additional burns were required because of low signal intensity. A 40 µg/L of leucine-enkephalin/isopropanol (Sigma-Aldrich, St. Louis, MO) solution was injected directly into the REIMS sources (flow rate: 200 µL/min) for lock mass correction. REIMS data were preprocessed with AMX Recognition software (v. 1.0.2184.0, Waters Corporation). This process included lock mass correction using leucine-enkephalin (554.2615 m/z), background subtraction using standard MassLynx preprocessing algorithms, total ion current normalization, peak binning in 0.5 m/z intervals, selection and averaging 5 peaks per sample, and exclusion of bins in the range of 550 to 600 m/z to remove the internal standard peak.

### ***Tenderness, juiciness, and flavor classification***

Individual samples were classified into 2 classes of juiciness, sensory panel tenderness (SPT), SSF, WBS, and flavor performance. Classification cutoffs for juiciness and tenderness were set close to the mean value of each attribute to achieve relatively balanced classes. LL and PM samples with juiciness values  $\leq 55$  were classified as “dry,” and samples with values  $> 55$  were classified as “juicy.” LL samples with SPT  $\geq 55$  were classified “tender,” and samples with values  $< 55$  were classified as “tough,” whereas PM samples with SPT  $\geq 77$  were classified “tender,” and samples with values  $< 77$  were classified as “tough.” SSF cutoffs were 14.0 and 12.0 kg, and WBS cutoffs were 3.1 and 2.9 kg for LL and PM, respectively. Because the PM is more tender than the LL muscle, the tenderness cutoffs for the PM muscle were set lower than for the LL muscle to achieve balanced classes (McKeith et al., 1985; Legako et al., 2015; Nair et al., 2019). Samples with higher values of SSF and WBS than the cutoffs were classified as “tough,” and lower than the cutoff were classified as “tender.”

Hierarchical clustering was used to group samples into classes with similar flavor performance. Principal component analysis (PCA) of the 13 flavor attributes was performed using the PCA function from the FactoMineR package (Lê et al., 2008). Then, the

HCPC function from the FactoMineR package was used to group the samples with similar flavor performance based on the first 5 principal components. Samples aligned to positive flavor attributes were assigned to the “positive” class and those aligned to the negative ones to the “negative” class. Pairwise comparisons of the proportion of samples within each class for each aging treatment were analyzed with McNemar’s test, a confidence level of 0.05, and Bonferroni adjustment using the pairwise\_mcnemar\_test function from the rstatix package (Kassambara, 2021). Comparisons of proportions were performed using carcass numbers as blocks and individually for each muscle because the classification cutoffs differ by muscle.

### ***Predictive model building***

Classification models were built using REIMS data of GR samples to predict quality attributes of PM and LL samples on days 3, 14, and 28. Before modeling, mass bins were mean-centered and normalized to the variance of all samples. Models were built using combinations of 3 dimension reduction methods and 15 machine-learning algorithms. PCA, feature selection (FS), and combined principal component analysis–feature selection (PCAFS) methods were used to reduce the dimensionality of the REIMS data. The PCA was performed with the PCA function from the FactoMineR package, and the number of principal components was selected with a scree plot. In FS, highly correlated mass bins ( $|r| > 0.9$ ) were identified and removed from the data; then, a recursive feature elimination was performed using the rfe function from the caret package (Kuhn, 2008). The PCAFS was performed with PCA, followed by recursive feature elimination using the aforementioned functions. Fifteen machine-learning algorithms, including 14 (Table 2) from the caret package repository (Kuhn, 2008) and PLS-DA using the plsDA function from the DiscriMiner package (Sanchez, 2012), were used to build predictive models for each combination of muscle, aging period, response variable, and dimensionality reduction method. All models were trained using 10-fold cross-validation using the train function for models from the caret package and manually for the PLS-DA. The PLS-DA plots with 95% confident interval ellipses were produced to visualize data overlap using the function plsda from the mixOmics package (Lê Cao et al., 2016).

The performance of the models was evaluated with the accuracy, sensitivity, and specificity obtained with

**Table 2.** R machine-learning algorithms, functions, and libraries used for building classification models

Algorithm	Function	Libraries
Stochastic gradient boosting	gbm	gbm, plyr
Support vector machine radial kernel	svmRadial	kernelab
Support vector machine linear kernel	svmLinear	kernelab
Support vector machine polynomial kernel	svmPoly	kernelab
Linear discriminant analysis	lda	MASS
eXtreme gradient boosting	xgbTree	xgboost, plyr
Penalized discriminant analysis	pda	mda
Boosted logistic regression	LogitBoost	caTools
Random forest	rf	randomForest
Generalized linear model	glm	stats
Lasso and elastic-net regularized generalized linear models	glmnet	glmnet, Matrix
K-nearest neighbors	knn	class
Recursive partitioning tree	rpart	rpart
Bagged classification tree	treebag	ipred, plyr, e1071

the confusionMatrix function from the caret package. A one-sided binomial test with a 95% confidence interval was used to evaluate if model accuracies were greater than the largest class proportion. Sensitivity and specificity were used to evaluate the performance of the models to classify individual classes. Sensitivity was calculated by dividing the number of samples classified correctly as positives by the total actual positives, whereas specificity was calculated as the fraction of negatives samples classified correctly by the total actual negatives. “Tough,” “dry,” and “positive” samples were treated as positive for sensitivity calculations of tenderness, juiciness, and flavor performance models, respectively. However, “tender,” “juicy,” and “negative” were treated as negative for specificity calculations. The statistical analysis described here was performed in R statistical computing program (v. 4.1.0, 2018).

## Results and Discussion

### Tenderness and juiciness classes

The LL and PM samples were classified into “tough” and “tender” classes for tenderness based on 3 measurement methods (SSF, WBS, and SPT) and into “dry” and “juicy” classes for juiciness. The numbers of samples per class and aging period of LL and PM are presented in Tables 3 and 4, respectively. As expected, the proportion of tough LL samples decreased with the aging time ( $P < 0.05$ , Table 3) for all tenderness measurement methods. Tenderness

**Table 3.** Number of observations (percent) of striploins (LL) per class of SSF, WBS, SPT, flavor, and juiciness

Attribute	Classes <sup>1</sup>	3 d	14 d	28 d
SSF	Tender (SSF ≤ 14.0 kg)	21 (25.3) <sup>a</sup>	44 (55.0) <sup>b</sup>	66 (82.5) <sup>c</sup>
	Tough (SSF > 14.0 kg)	62 (74.7) <sup>a</sup>	36 (45.0) <sup>b</sup>	14 (17.5) <sup>c</sup>
WBS	Tender (WBS ≤ 3.1 kg)	17 (20.5) <sup>a</sup>	44 (55.0) <sup>b</sup>	60 (75.0) <sup>c</sup>
	Tough (WBS > 3.1 kg)	66 (79.5) <sup>a</sup>	36 (45.0) <sup>b</sup>	19 (23.8) <sup>c</sup>
SPT	Tender (SPT ≥ 55)	12 (14.5) <sup>a</sup>	28 (35.4) <sup>b</sup>	57 (71.3) <sup>c</sup>
	Tough (SPT < 55)	71 (85.5) <sup>a</sup>	51 (64.6) <sup>b</sup>	23 (28.8) <sup>c</sup>
Flavor <sup>1</sup>	Positive	53 (63.9) <sup>a</sup>	46 (58.2) <sup>ab</sup>	35 (43.8) <sup>b</sup>
	Negative	30 (36.1) <sup>a</sup>	33 (41.8) <sup>ab</sup>	45 (56.3) <sup>b</sup>
Juiciness	Dry (juiciness ≤ 55)	37 (44.6)	39 (49.4)	40 (50.0)
	Juicy (juiciness > 55)	46 (55.4)	40 (50.6)	40 (50.0)

<sup>a-c</sup>Attributes in the same row lacking common superscript differ ( $P < 0.05$ ).

<sup>1</sup>Flavor classes based on hierarchical clustering of the samples using principal component analysis.

LL = longissimus lumborum; SPT = sensory panel tenderness; SSF = slice shear force; WBS = Warner-Bratzler shear force.

**Table 4.** Number of observations (percent) of tenderloins (PM) per class of SSF, WBS, SPT, flavor, and juiciness

Attribute	Classes <sup>1</sup>	3 d	14 d	28 d
SSF	Tender (SSF ≤ 12.0 kg)	32 (38.1) <sup>a</sup>	47 (57.3) <sup>b</sup>	50 (61) <sup>b</sup>
	Tough (SSF > 12.0 kg)	52 (61.9) <sup>a</sup>	35 (42.7) <sup>b</sup>	32 (39) <sup>b</sup>
WBS	Tender (WBS ≤ 2.9 kg)	34 (40.5) <sup>a</sup>	39 (47.6) <sup>ab</sup>	52 (63.4) <sup>b</sup>
	Tough (WBS > 2.9 kg)	50 (59.5) <sup>a</sup>	43 (52.4) <sup>ab</sup>	30 (36.6) <sup>b</sup>
SPT	Tender (SPT ≥ 77)	54 (64.3)	41 (48.8)	45 (54.2)
	Tough (SPT < 77)	30 (35.7)	43 (51.2)	38 (45.8)
Flavor <sup>1</sup>	Positive	43 (51.2) <sup>a</sup>	43 (51.2) <sup>a</sup>	23 (27.7) <sup>b</sup>
	Negative	41 (48.8) <sup>a</sup>	41 (48.8) <sup>a</sup>	60 (72.3) <sup>b</sup>
Juiciness	Dry (juiciness ≥ 55)	31 (36.9)	39 (46.4)	43 (51.8)
	Juicy (juiciness < 55)	53 (63.1)	45 (53.6)	40 (48.2)

<sup>a,b</sup>Attributes in the same row lacking common superscript differ ( $P < 0.05$ ).

<sup>1</sup>Flavor classes based on hierarchical clustering of the samples using principal component analysis.

PM = psoas major; SPT = sensory panel tenderness; SSF = slice shear force; WBS = Warner-Bratzler shear force.

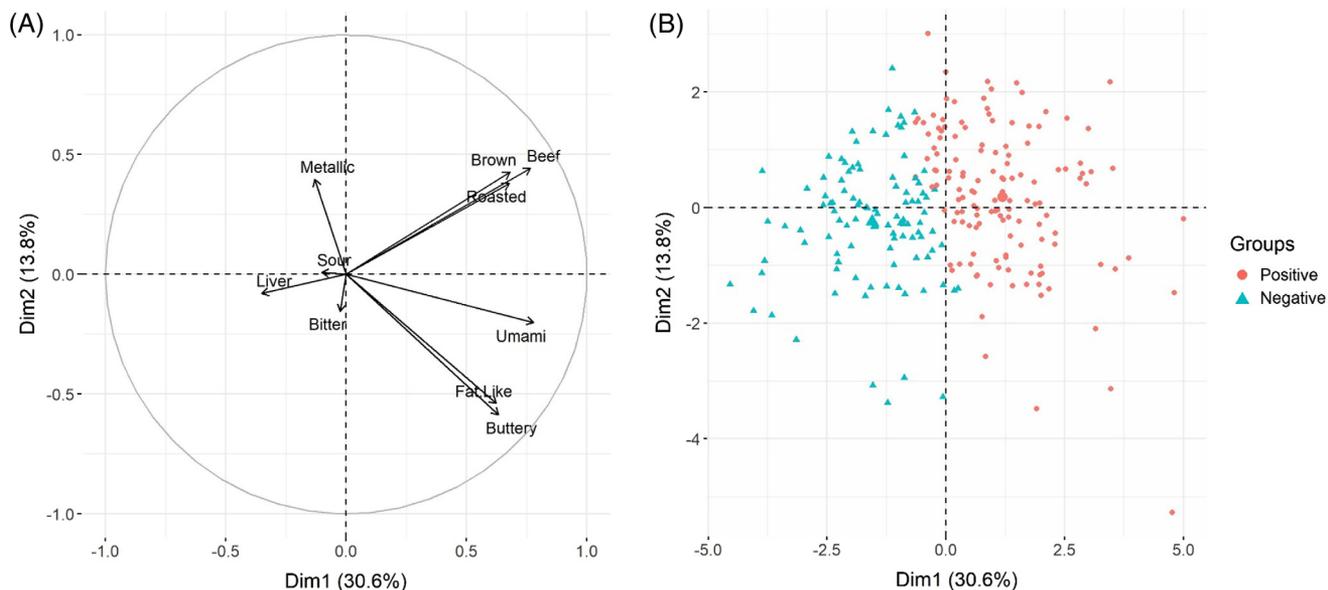
increases with aging because of the proteolysis of structural proteins of the muscles (Koochmarai, 1996). However, there were no differences in the proportions of PM samples classified as tough in SPT ( $P > 0.05$ ) for all aging periods. The proportion of PM samples classified as tough by the SSF method on day 3 was greater than the proportion of tough samples on day 14 and 28 postmortems ( $P < 0.05$ ). Also, the proportion of tough samples by WBS on day 3 was greater than the proportion of tough samples on day 28 ( $P < 0.05$ , Table 4).

The effect of aging on tenderness is muscle-specific, which could explain why the proportions of tough LL samples decreased in all aging periods and a different behavior was observed with the PM samples. Tenderness of the LL muscle can improve during 21 d of aging, whereas tenderness of the PM muscle does not improve after 7 d of aging (Nair et al., 2019). The proportion of dry samples did not change with aging periods in both muscles. These results were expected because the juiciness of LL and PM steaks does not

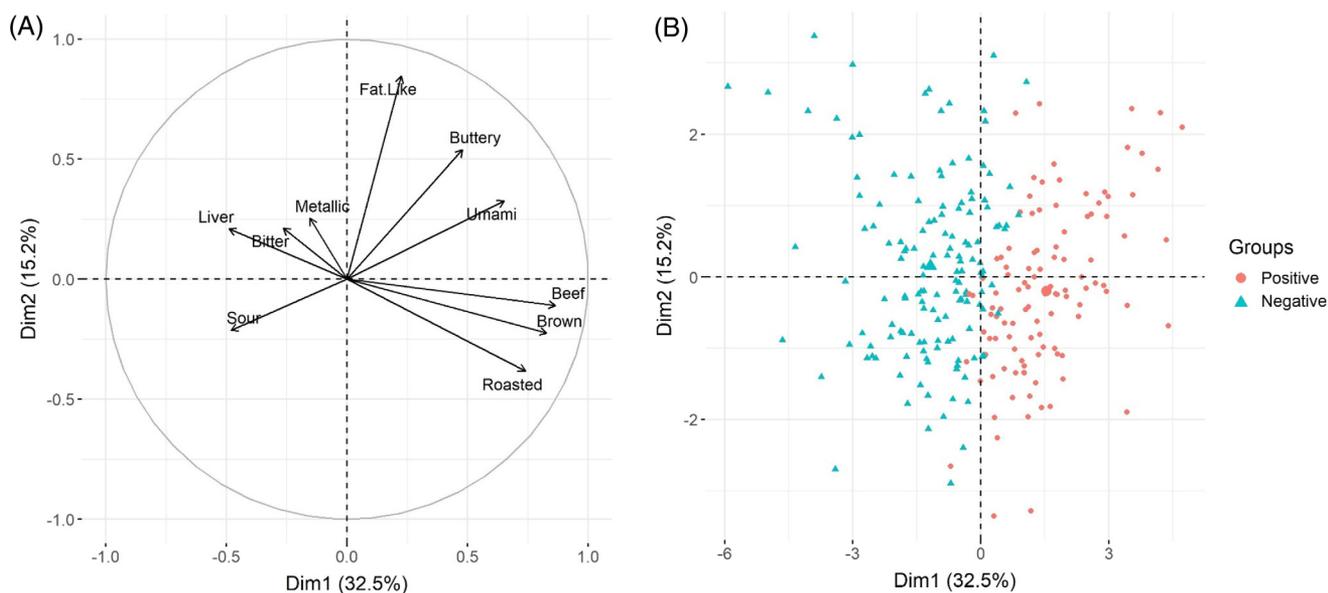
change with wet-aging (Lepper-Blilie et al., 2016; Foraker et al., 2020; Kim et al., 2021).

### Flavor classes

The LL and PM samples aged 3, 14, and 28 d were classified using PCA of the 10 flavor attributes evaluated in the sensory panels (Figures 1 and 2). Loading plots of PCA of both muscles showed that brown, beefy, and roasted flavors were positively correlated,



**Figure 1.** Principal component analysis (PCA) of 10 flavor attributes of beef striploins (longissimus lumborum [LL];  $n = 242$ ) aged 3, 14, and 28 d. (A) Loading plot of PCA and (B) flavor classes based on hierarchical clustering of the samples in the PCA.



**Figure 2.** Principal component analysis (PCA) of 10 flavor attributes of beef tenderloins (psoas major [PM];  $n = 251$ ) aged 3, 14, and 28 d. (A) Loading plot of PCA and (B) flavor classes based on hierarchical clustering of the samples in the PCA.

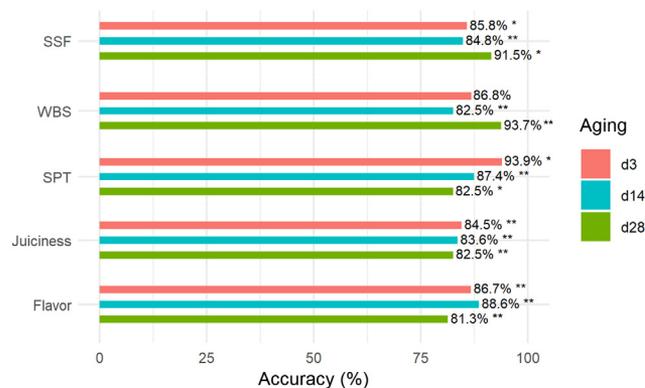
and these attributes were negatively associated with livery flavor. As expected, fat-like and buttery flavors were associated together because the intensity of both of these attributes has been shown to increase with intramuscular fat (Legako et al., 2016; O'Quinn et al., 2016). However, they were more closely related to umami than to other flavors. Foraker et al. (2020) found similar results when performing a discriminant function analysis of aged beef LL sensory attributes. They found that positive flavors such as roasted, beef identity, fat-like, and browned were positively associated with each other but negatively associated with off-flavors (sour, liver-like, oxidized, and metallic). Gredell (2018) also found similar results when performing a PCA of beef samples from different USDA grades, cattle breeds, and production systems. PCA analysis by Gredell (2018) showed association of browned, beef identity, buttery, and fat-like flavors.

Hierarchical clustering of the sensory data resulted in 2 classes that were defined as “positive” and “negative” (Figures 1 and 2). Samples classified as positive were mostly located in the positive quadrant of dimension 1 of the PCA, corresponding to higher intensity of the positive attributes and lower intensity of the negative attributes. Classifications were mostly driven by dimension 1 of the PCA, and the major contributor to this dimension was the beef identity. Foraker et al. (2020) and Gredell (2018) obtained similar results by performing a multivariable analysis of flavor attributes. Both studies concluded that beef flavor identity was the major contributor to separating samples with positive from negative performance. The proportion of LL samples classified as positive after 3 d of aging was greater than the proportion of samples classified as positive after 28 d ( $P < 0.05$ , Table 3). More PM samples were classified as positive on 3 and 14 d of aging compared with 28 d ( $P < 0.05$ , Table 4).

### Classification models of tenderness, juiciness, and flavor of striploins

Three different dimensionality reduction methods (PCA, FS, and PCAFS) combined with different machine-learning algorithms were trained using 10-fold cross-validation. Accuracies of the best models for each attribute and aging time of LL are reported in Figure 3. The sensitivity and specificity of the highest performing models are presented in Table 5.

**Tenderness.** All the highest performance models for tenderness, except for the WBS of 3 d (d3) model, showed accuracies higher than the proportion of the



**Figure 3.** Model accuracies of best performance striploin (longissimus lumborum [LL]) models for slice shear force (SSF), Warner-Bratzler shear force (WBS), sensory panel tenderness (SPT), juiciness, and flavor aged 3, 14, and 28 d. \* $p$  value  $< 0.05$ , \*\* $p$  value  $< 0.01$ .

largest class ( $P < 0.05$ , Figure 3). These results demonstrate that the models' prediction abilities were better than the chance of guessing the tenderness classes and that REIMS can predict the tenderness of LL. However, the models' accuracies were different for the tenderness measurement methods and aging periods (Table 5). The combination of FS and Lasso and elastic-net regularized generalized linear models (GLMNET) performed the best in predicting SSF of d3 with 85.8% accuracy, whereas for 14 d (d14) and 28 d (d28), the PCAFS with extreme gradient boosting (XGBoost) were the best models, with 84.8% and 91.5% accuracies, respectively (Table 5). None of the models were able to predict WBS d3 ( $P > 0.05$ ), but FS with random forest predicted WBS d14 with 82.5% and PCAFS with support vector machine with polynomial kernel (SVM poly) predicted WBS d28 with 93.7% accuracy. The FS and XGBoost models were the best to predict SPT of d3 and d14 with 93.9% and 87.4% accuracy, and FS with PLS-DA displayed 82.5% accuracy in predicting SPT d28 (Table 5). Because the cutoffs to separate samples into tenderness classes were close to the means and tenderness is a numerical attribute, classes overlapped in the PLS-DA plots (Figures 4 to 6). The PLS-DA is a supervised algorithm that projects the data into new hyperplanes that optimize linearity between the data and the response variable (Gareth et al., 2021). Therefore, overlaps between classes suggest that the REIMS profiles of samples located in the overlapped area are similar and induce errors in the classification models. The PLS-DA plot of SSF and WBS of d28 and SPT of d3 showed the smallest overlap between those tenderness classes, which may explain the highest accuracy of those models.

Although SSF d28 and SPT d3 models showed high accuracy, their sensitivity to classify tough

**Table 5.** Top prediction accuracies (based on 10-fold cross-validation, percent) for striploin (LL) tenderness, juiciness, and flavor based on REIMS of GR samples

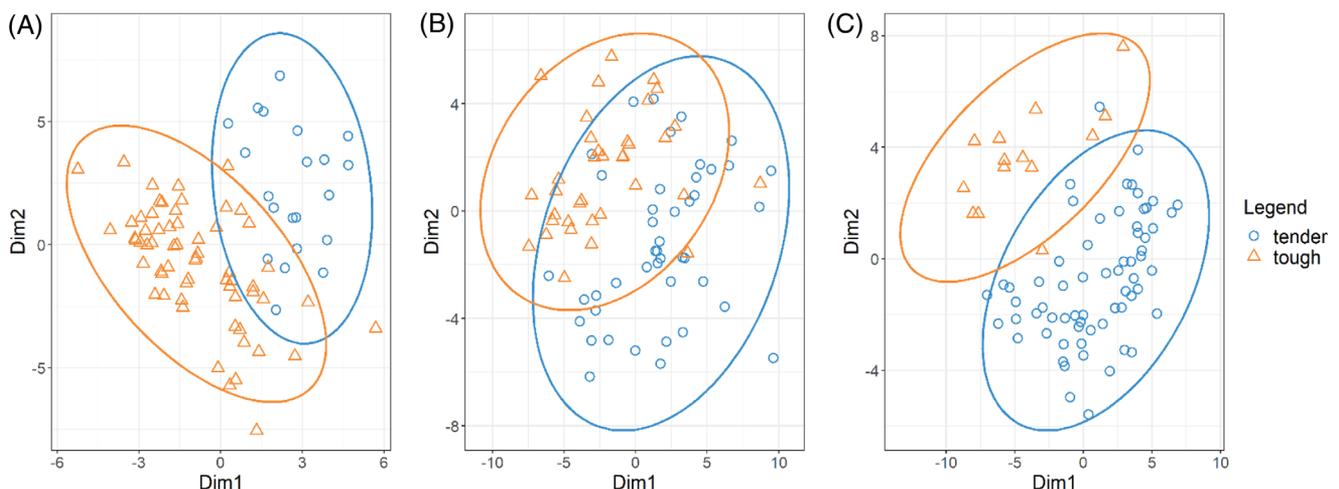
Models <sup>1</sup>	Aging Period <sup>2</sup>	Top Model <sup>3</sup>	Number of Features	Maximum Accuracy	<i>p</i> Value	Sensitivity	Specificity
SSF	d3	FS/GLMNET	155	85.8	1.2E-02	95.2	57.1
	d14	PCAFS/XGBoost	12	84.8	1.7E-10	80.6	88.6
	d28	PCAFS/XGBoost	3	91.5	2.1E-02	50.0	100.0
WBS	d3	FS/XGBoost	41	86.8	6.2E-02	95.5	52.9
	d14	FS/RF	29	82.5	2.0E-07	77.8	86.4
	d28	PCAFS/SVM Poly	21	93.7	3.3E-05	84.2	96.7
SPT	d3	FS/XGBoost	23	93.9	1.4E-02	58.3	100.0
	d14	FS/XGBoost	46	87.4	4.6E-06	78.6	92.2
	d28	FS/PLS-DA	600	82.5	1.5E-02	47.8	96.5
Flavor	d3	FS/GBM	180	86.7	2.8E-06	92.5	76.7
	d14	FS/PLS-DA	265	88.6	1.3E-07	78.3	92.5
	d28	FS/XGBoost	17	81.3	2.2E-06	74.3	86.7
Juiciness	d3	FS/XGBoost	84	84.5	2.1E-08	89.1	78.4
	d14	FS/XGBoost	26	83.6	1.1E-09	84.6	82.5
	d28	FS/XGBoost	61	82.5	1.6E-09	85.0	80.0

<sup>1</sup>SPT = sensory panel tenderness; SSF = slice shear force; WBS = Warner-Bratzler shear force.

<sup>2</sup>d3 = 3 d; d14 = 14 d; d28 = 28 d of aging.

<sup>3</sup>FS = feature selection; GLMNET = Lasso and elastic-net regularized generalized linear models; GR = grading; PCAFS = principal component analysis–feature selection; PLS-DA = partial least square discriminant analysis; REIMS = rapid evaporative ionization mass spectrometry; RF = random forest; SVM = support vector machine; SVM Poly = support vector machine with polynomial kernel; XGBoost = extreme gradient boosting.

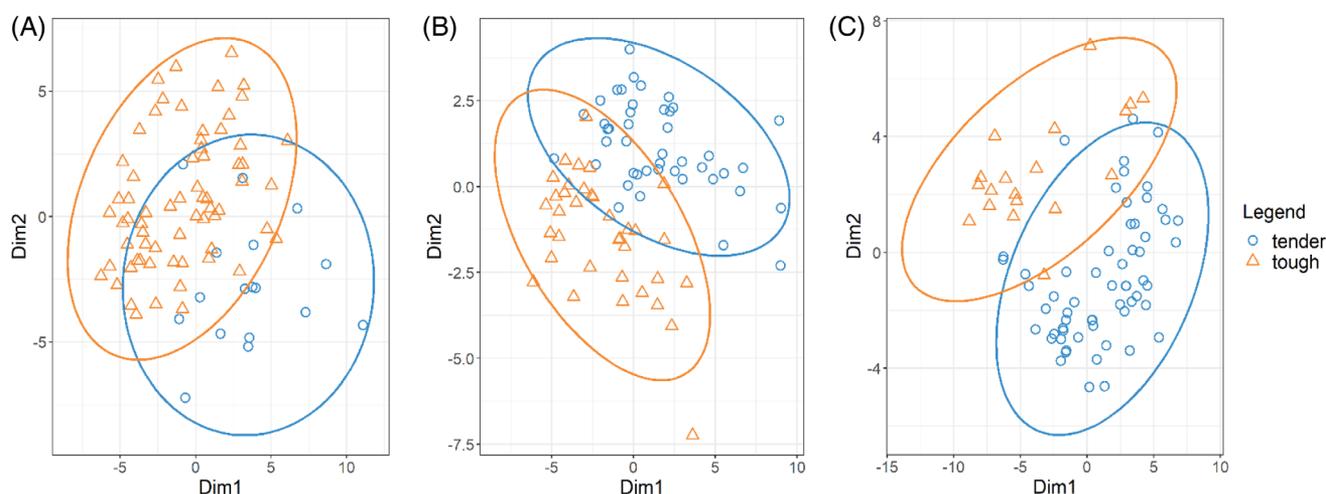
LL = longissimus lumborum.



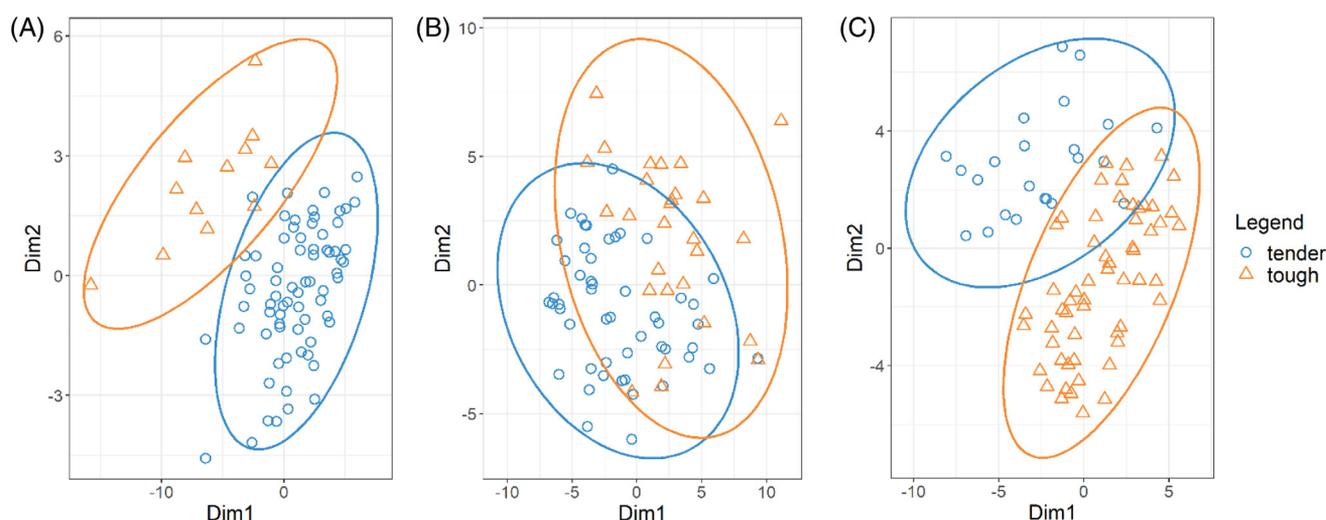
**Figure 4.** Partial least-squares–discriminant analysis (PLS-DA) plot of slice shear force of striploin (longissimus lumborum [LL]) classes corresponding to (A) 3, (B), 14, and (C) 28 d of aging.

samples was low (Table 5). The best performing model for tenderness was the WBS d28, showing an accuracy of 94.8% with 84.2% sensitivity and 96.7% specificity. The specificity of all tenderness models increased with aging, and the sensitivity of all tenderness models decreased with aging, except for the WBS models. The ability of REIMS to classify tough samples decreased with aging time, and the ability to sort tender samples increased with aging time (Table 5). This observation could be attributed to the data structure

used in this study. Data of d3 and d28 were unbalanced because most d3 observations were tough, whereas most d28 were tender. Machine-learning algorithms perform poorly when data are unbalanced because the models favor the larger class (Cieslak and Chawla, 2008). In addition, overlaps between classes and small numbers of observations in the minority groups contribute to the low performance of predictive models (Batista et al., 2004). The results obtained in this study were similar to those reported by Gredell et al. (2019). These authors



**Figure 5.** Partial least-squares–discriminant analysis (PLS-DA) plot of Warner-Bratzler shear force of striploin (longissimus lumborum [LL]) classes corresponding to (A) 3, (B) 14, and (C) 28 d of aging.



**Figure 6.** Partial least-squares–discriminant analysis (PLS-DA) plot of sensory panel tenderness of striploin (longissimus lumborum [LL]) classes corresponding to (A) 3, (B) 14, and (C) 28 d of aging.

reported 90.8% accuracy of an FS/SVM poly model that differentiated tough from tender samples of LL aged 14 d using SSF = 20.0 kg as a cutoff.

The limitation of REIMS to analyze high molecular weight compounds could constrain the use of this technique to predict tenderness with high accuracy. Studies have shown that a large part of the variability in beef tenderness depends on muscle proteins and intramuscular fat (Platter et al., 2003; Emerson et al., 2013; Picard et al., 2014; Gagaoua et al., 2018). During the first few days postmortem, the cross-sectional area of muscle fiber and sarcomere length are large contributors to beef tenderness, whereas background toughness becomes more relevant after 14 d of aging (Dubost et al., 2013). Those factors depend on contractile, cytoskeletal, and stromal proteins found in muscle tissues. Because REIMS

requires that the compounds found in the samples be vaporized, only compounds with molecular weight between 1,000 and 1,500 Da reach the mass spectrometer. Therefore, REIMS can analyze chemical compounds, including fatty acids, sugars, phospholipids, and small peptides, but cannot detect large peptides or proteins (Ross et al., 2021). REIMS could provide information on metabolites identified as possible indicators of beef tenderness (for example, malic acid, glucose, glucose, glucose-6-phosphate) and compounds related to intramuscular fat (King et al., 2019). However, REIMS will likely not detect intact proteins (e.g., collagen, desmin) that also influence tenderness variability.

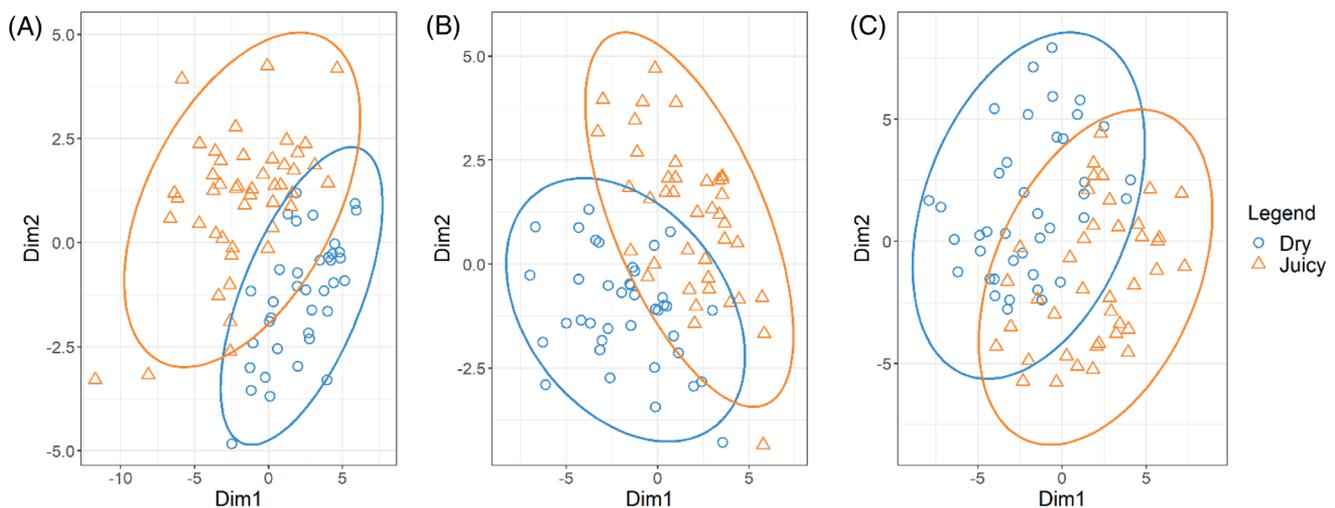
**Flavor.** The results of the flavor models are shown in Figure 3. Accuracies of the flavor models were higher

than the highest-class proportions ( $P < 0.05$ ) in all aging periods, but none of them showed an accuracy higher than 90%. The PCAFS/XGBoost was the best model to predict flavor classes of d3 with 83.3%. The FS/XGBoost of d14 had 73.8%, and PCAFS/SVM poly of d28 had 84.3% accuracy, respectively. The low performance of the flavor models may be due to the complexity of the response variable. Flavor perception results from taste and odor sensations and can be defined by multiple flavor descriptors (e.g., browned, sourness, or beef identity). Specific flavor attributes have been associated with numerous chemical compounds or groups of compounds. Elaborating a composite score that condenses 10 flavor attributes (Figure 1) into 2 classes might oversimplify beef flavor interpretation. However, clustering multiple flavor attributes represent a more difficult problem in building classification models based on the chemical fingerprint. Mathematically, when the number of dimensions (attributes) increases, the distance between observations increases, making it more difficult to find nearby observations in the hyperplanes (Aremu et al., 2020). In addition, samples in one class could be chemically different because the classification results from a combination of chemically unrelated attributes. Because machine-learning algorithms use similarities in the training dataset to predict classes of the test data, more data will be required to cover a broader range of flavor profiles.

Gredell (2018) previously evaluated the ability of REIMS to predict flavor classes defined with multivariate analysis and reported an accuracy of 73.8% for binary classification. However, in that study, beef samples were from different USDA grades, cattle breeds, and production systems, which could increase

variability in the fingerprint across classes and increase error in the model. Additionally, Gredell (2018) only used PLS-DA algorithms to build the model, whereas in the current study, multiple algorithms were evaluated, and the best one was reported. Verplanken et al. (2017) evaluated the ability of REIMS to predict pork flavor. They demonstrated that REIMS could classify pork carcasses with untainted and tainted boar flavor with perfect accuracy using an orthogonal PLS-DA model. Boar taint is an odor occurring in the meat of noncastrated male pigs and is associated with high levels of skatole and androstenone (Verplanken et al., 2017). Therefore, identifying boar taint may be less complex than identifying beef flavor profiles because boar taint is a specific flavor with a physiological origin. In contrast, the beef flavor profiles evaluated in the present study result from a combination of attributes that could come from different sources of variation.

**Juiciness.** Accuracies of all juiciness models were higher than the largest proportion of juiciness classes ( $P < 0.05$ , Figure 3). The FS/XGBoost model predicted juiciness classes of d3 with 84.5%, d14 with 83.6%, and d28 with 82.5% accuracy, respectively (Table 5). The PLS-DA plots corresponding to juiciness classes (Figure 7) showed an overlap between the 2 classes during aging times, which could explain the low performance of the 3 juiciness models. In the PLS-DA plot, the overlap implies similarity in the REIMS data of both classes. These results are not surprising because juiciness classes were defined with a cutoff close to the average juiciness of all the samples. Samples in different classes but with similar values of juiciness are probably more similar than samples in the same class with more



**Figure 7.** Partial least-squares-discriminant analysis (PLS-DA) plot of juiciness of striploin (longissimus lumborum [LL]) classes corresponding to (A) 3, (B) 14, and (C) 28 d of aging.

variation in juiciness values. For example, a sample considered as juicy with a juiciness value of 55.1 is perhaps more similar to a sample with a juiciness value of 54.7 classified as dry than a sample with a juiciness value of 63 classified as juicy. Juiciness is influenced by intramuscular fat and protein composition of muscle tissues (Dubost et al., 2013). Therefore, the inability of REIMS to detect intact proteins could contribute to low prediction accuracy for this attribute.

Another source of error in all models could be the inconsistency in the cooking temperatures of the samples used for sensory evaluation. Even following the same cooking protocol, it is very common to have minor variations in the cooking temperature because of uncontrolled variables during the cooking process (e.g., steak size or location in the oven). The 3 sensory attributes evaluated are highly dependent on the cooking temperature. Steaks cooked with a higher internal temperature are less tender and less juicy than steaks cooked with a lower internal temperature (Cross et al., 1976; Savell et al., 1999; McKillip et al., 2017). Beef flavor is produced by chemical reactions catalyzed by temperature, and cooking temperature plays a major role in developing beef flavor compounds (Kerth and Miller, 2015). Consequently, using REIMS on GR samples does not account for variations in sensory attributes affected by extrinsic factors during storage, transportation, and cooking.

**Table 6.** Top prediction accuracies (based on 10-fold cross-validation, percent) for tenderloin (PM) tenderness, juiciness, and flavor based on REIMS of GR samples

Models <sup>1</sup>	Aging Period <sup>2</sup>	Top Model <sup>3</sup>	Number of Features	Maximum Accuracy	<i>p</i> Value	Sensitivity	Specificity
SSF	d3	PCAFS/SVM Poly	59	82.1	4.90E-05	88.5	71.9
	d14	PCAFS/XGBoost	51	73.4	2.18E-03	68.6	76.6
	d28	FS/XGBoost	38	89.0	1.56E-08	75.0	98.0
WBS	d3	FS/SVM Poly	205	78.6	1.76E-04	88.0	64.7
	d14	FS/RF	4	78.1	1.47E-06	81.4	74.4
	d28	FS/SVM Poly	32	81.5	2.47E-04	76.7	84.6
SPT	d3	PCAFS/XGBoost	34	83.8	9.78E-05	73.3	88.9
	d14	FS/XGBoost	80	81.0	1.46E-08	81.4	80.5
	d28	PCAFS/SVM	46	84.3	6.39E-09	78.9	88.9
Flavor	d3	PCAFS/XGBoost	29	83.3	7.62E-10	90.7	75.6
	d14	FS/XGBoost	4	73.8	1.90E-05	72.1	75.6
	d28	PCAFS/SVM Poly	35	84.3	7.42E-03	56.5	95.0
Juiciness	d3	FS/XGBoost	30	81.0	3.06E-04	90.6	64.5
	d14	PCAFS/SVM Poly	65	79.8	5.30E-07	84.4	74.4
	d28	PCAFS/SVM Poly	59	79.5	1.60E-07	80.0	79.1

<sup>1</sup>SPT = sensory panel tenderness; SSF = slice shear force; WBS = Warner-Bratzler shear force.

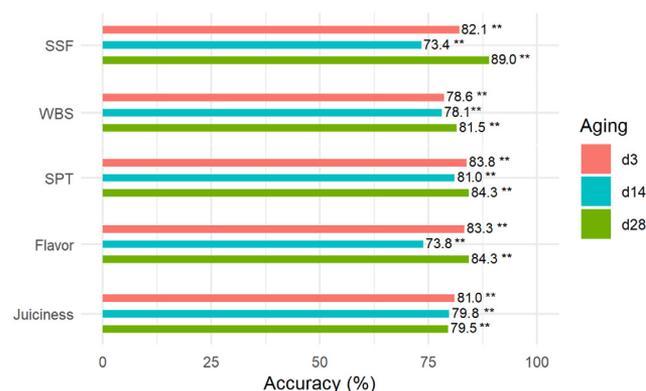
<sup>2</sup>d3 = 3 d; d14 = 14 d; d28 = 28 d of aging.

<sup>3</sup>FS = feature selection; GR = grading; PCAFS = principal component analysis–feature selection; PLS-DA = partial least square discriminant analysis; REIMS = rapid evaporative ionization mass spectrometry; RF = random forest; SVM = support vector machine; SVM Poly = support vector machine with polynomial kernel; XGBoost = extreme gradient boosting.

PM = psoas major.

### Classification models of tenderness, juiciness, and flavor of tenderloins

The accuracy, sensitivity, and specificity of the best models for PM are reported in Figure 8 and Table 6. In all cases, except for the flavor model of d3 and SPT of d28, the accuracies of LL models were numerically higher than PM models. Similar to the LL models, the specificity of SSF and WBS increased, and sensitivity decreased with aging time because of unbalanced data. The number of tender samples increases



**Figure 8.** Model accuracies of best performance tenderloin (psoas major [PM]) models for slice shear force (SSF), Warner-Bratzler shear force (WBS), sensory panel tenderness (SPT), juiciness, and flavor aged 3, 14, and 28 d. \**p* value <0.05, \*\**p* value <0.01.

with aging, enhancing the ability of the model to predict the majority class.

Because REIMS was collected on GR samples from LL muscle, the ability of REIMS to predict sensory attributes of PM muscle was expected to be lower. These 2 muscles are metabolically different, with distinct biochemical properties (Abraham et al., 2017). Although LL is predominantly glycolytic, PM is primarily oxidative (Kim et al., 2021). LL contains more muscle fiber type IIX than type I and IIA, whereas PM comprises more type I fibers than type IIX and IIA fibers (Kim et al., 2021). Several proteins degrade at different rates in both muscles suggesting that proteolysis mechanisms related to quality attributes are muscle-dependent (Gagaoua et al., 2018; Kim et al., 2021). Proteome analysis of LL and PM revealed that sarcoplasmic proteins, including metabolic enzymes, antioxidants, and chaperone proteins, differed in both muscles and during aging (Joseph et al., 2012, Nair et al., 2019). Previous authors have reported that sensory attributes change differently in both muscles because of metabolic differences. Tenderness of the LL improves until 21 d postmortem, whereas in PM, tenderness only improves until day 7, with no additional improvement observed with extended aging (Nair et al., 2019). Therefore, even if REIMS data from a given set of LL samples are found to classify carcasses on tenderness with perfect accuracy, it is expected that, in general, the prediction accuracy for tenderness in other muscles is lower.

## Conclusions

Consumer eating satisfaction of beef depends on the tenderness, juiciness, and flavor. Therefore, nondestructive techniques that allow the prediction of these attributes in real time can be implemented in the current grading system and improve the consumer experience, increasing consumer trust and demand for beef. This study demonstrated that REIMS could potentially be used as a real-time, in situ technique to classify beef carcasses into flavor, juiciness, and tenderness classes at different aging times. However, the ability of REIMS to predict meat juiciness and tenderness may have been limited by its inability to detect intact proteins known to be related to these attributes. Moreover, working with a composite score for flavor classes improves the interpretation of this complex attribute but also increases the model complexity. Developing a predictive model of beef flavor that encompasses a wide range of sensory profiles will require a large sample size to cover a wide

range of flavor profiles. In the current study, overlap in the data and unbalanced class size both affected model accuracies. Future work will require increasing the sample size and overcoming the challenge of collecting samples with greater differentiation in sensory performance between classes.

## Acknowledgements

The authors declare that they have no conflicts of interest. The United States Department of Agriculture (USDA) provided funding for this research. This work was also partially supported by the USDA National Institute of Food and Agriculture, Multi-State Hatch project COL00276B, accession number 7003980. The use of trade names in this publication does not imply endorsement or criticism by Colorado State University of those or similar products not mentioned.

## Literature Cited

- Abraham, A., J. W. Dillwith, G. G. Mafi, D. L. VanOverbeke, and R. Ramanathan. 2017. Metabolite profile differences between beef longissimus and psoas muscles during display. *Meat Muscle Biol.* 1:18–27. <https://doi.org/10.22175/mmb2016.12.0007>
- Adhikari, K., E. Chambers IV, R. Miller, L. Vázquez-Araújo, N. Bhumiratana, and C. Philip. 2011. Development of a lexicon for beef flavor in intact muscle. *J. Sens. Stud.* 26:413–420. <https://doi.org/10.1111/j.1745-459X.2011.00356.x>
- Aremu, O. O., D. Hyland-Wood, and P. R. McAree. 2020. A machine learning approach to circumventing the curse of dimensionality in discontinuous time series machine data. *Reliab. Eng. Syst. Safe.* 195:106706. <https://doi.org/10.1016/j.ress.2019.106706>
- Balog, J., D. Perenyi, C. Guallar-Hoyas, A. Egri, S. D. Pringle, S. Stead, O. P. Chevallier, C. T. Elliott, and Z. Takats. 2016. Identification of the species of origin for meat products by rapid evaporative ionization mass spectrometry. *J. Agr. Food Chem.* 64:4793–4800. <https://doi.org/10.1021/acs.jafc.6b01041>
- Batista, G. E. A. P. A., R. C. Prati, and M. C. Monard. 2004. A study of the behavior of several methods for balancing machine learning training data. *ACM SIGKDD Explorations Newsletter.* 6:20–29. <https://doi.org/10.1145/1007730.1007735>
- Behrends, J. M., K. J. Goodson, M. Koochmarai, S. D. Shackelford, T. L. Wheeler, W. W. Morgan, J. O. Reagan, B. L. Gwartney, J. W. Wise, and J. W. Savell. 2005. Beef customer satisfaction: Factors affecting consumer evaluations of calcium chloride-injected top sirloin steaks when given instructions for preparation. *J. Anim. Sci.* 83:2869–2875. <https://doi.org/10.2527/2005.83122869x>
- Cieslak, D. A., and N. V. Chawla. 2008. Learning decision trees for unbalanced data. In: W. Daelemans, B. Goethals, and K. Morik, editors. *Machine learning and knowledge discovery*

- in databases. Springer, Berlin, Heidelberg. p. 241–256. [https://doi.org/10.1007/978-3-540-87479-9\\_34](https://doi.org/10.1007/978-3-540-87479-9_34)
- Cross, H. R., M. S. Stanfield, and E. J. Koch. 1976. Beef palatability as affected by cooking rate and final internal temperature. *J. Anim. Sci.* 43:114–121. <https://doi.org/10.2527/jas1976.431114x>
- DeOtte, R. E., K. S. Spivey, H. O. Galloway, and T. E. Lawrence. 2010. Conservation of water for washing beef heads at harvest. *Meat Sci.* 84:371–376. <https://doi.org/10.1016/j.meatsci.2009.09.004>
- Dubost, A., D. Micol, B. Picard, C. Lethias, D. Andueza, D. Bauchart, and A. Listrat. 2013. Structural and biochemical characteristics of bovine intramuscular connective tissue and beef quality. *Meat Sci.* 95:555–561. <https://doi.org/10.1016/j.meatsci.2013.05.040>
- Emerson, M. R., D. R. Woerner, K. E. Belk, and J. D. Tatum. 2013. Effectiveness of USDA instrument-based marbling measurements for categorizing beef carcasses according to differences in longissimus muscle sensory attributes. *J. Anim. Sci.* 91:1024–1034. <https://doi.org/10.2527/jas.2012-5514>
- Foraker, B. A., D. A. Gredell, J. F. Legako, R. D. Stevens, J. D. Tatum, K. E. Belk, and D. R. Woerner. 2020. Flavor, tenderness, and related chemical changes of aged beef strip loins. *Meat Muscle Biol.* 4:28. <https://doi.org/10.22175/mmb.11115>
- Gagaoua, M., V. Monteils, and B. Picard. 2018. Data from the farm-gate-to-meat continuum including omics-based biomarkers to better understand the variability of beef tenderness: an integrative approach. *J. Agr. Food Chem.* 66:13552–13563. <https://doi.org/10.1021/acs.jafc.8b05744>
- Gareth, J., D. Witten, T. Hastie, and R. Tibshirani. 2021. Linear model selection and regularization. In: *An introduction to statistical learning with applications in R*. Second ed. NY. p. 225–288. [https://hastie.su.domains/ISLR2/ISLRv2\\_website.pdf](https://hastie.su.domains/ISLR2/ISLRv2_website.pdf)
- Goodson, K. J., W. W. Morgan, J. O. Reagan, B. L. Gwartney, S. M. Courington, J. W. Wise, and J. W. Savell. 2002. Beef customer satisfaction: Factors affecting consumer evaluations of clod steaks. *J. Anim. Sci.* 80:401–408. <https://doi.org/10.2527/2002.802401x>
- Gredell, D. 2018. Assessment of rapid evaporative ionization mass spectrometry (REIMS) to characterize beef quality and the impact of oven temperature and relative humidity on beef. Ph.D. thesis, Colorado State Univ., Fort Collins, CO. (<https://hdl.handle.net/10217/193115>)
- Gredell, D. A., A. R. Schroeder, K. E. Belk, C. D. Broeckling, A. L. Heuberger, S.-Y. Kim, D. A. King, S. D. Shackelford, J. L. Sharp, T. L. Wheeler, D. R. Woerner, and J. E. Prenni. 2019. Comparison of machine learning algorithms for predictive modeling of beef attributes using rapid evaporative ionization mass spectrometry (REIMS) data. *Sci Rep-UK.* 9:5721. <https://doi.org/10.1038/s41598-019-40927-6>
- Joseph, P., S. P. Suman, G. Rentfrow, S. Li, and C. M. Beach. 2012. Proteomics of muscle-specific beef color stability. *J. Agr. Food Chem.* 60:3196–3203. <https://doi.org/10.1021/jf204188v>
- Kassambara, A. 2021. rstatix: Pipe-friendly framework for basic statistical tests. <https://rpkgs.datanovia.com/rstatix/>
- Kerth, C. R., and R. K. Miller. 2015. Beef flavor: A review from chemistry to consumer. *J. Sci. Food Agr.* 95:2783–2798. <https://doi.org/10.1002/jsfa.7204>
- Kim, G.-D., S. Yun Lee, E.-Y. Jung, S. Song, and S. Jin Hur. 2021. Quantitative changes in peptides derived from proteins in beef tenderloin (psoas major muscle) and striploin (longissimus lumborum muscle) during cold storage. *Food Chem.* 338:128029. <https://doi.org/10.1016/j.foodchem.2020.128029>
- King, D. A., S. D. Shackelford, C. D. Broeckling, J. E. Prenni, K. E. Belk, and T. L. Wheeler. 2019. Metabolomic investigation of tenderness and aging response in beef longissimus steaks. *Meat Muscle Biol.* 3:76–89. <https://doi.org/10.22175/mmb2018.09.0027>
- Koohmaraie, M. . 1996. Biochemical factors regulating the toughening and tenderization processes of meat. *Meat Sci.* 43:193–201. [https://doi.org/10.1016/0309-1740\(96\)00065-4](https://doi.org/10.1016/0309-1740(96)00065-4)
- Kuhn, M. 2008. Building predictive models in R using the caret package. *J. Stat. Softw.* 28:1–26. <https://doi.org/10.18637/jss.v028.i05>
- Lê, S., J. Josse, and F. Husson. 2008. FactoMineR: An R package for multivariate analysis. *J. Stat. Softw.* 25:1–18. <https://doi.org/10.18637/jss.v025.i01>
- Lê Cao, K.-A., F. Rohart, I. Gonzalez, S. Dejean, B. Gautier, F. Bartolo, P. Monget, J. Coquery, F. Yao, and B. Lique. 2016. mixOmics: omics data integration project. <https://doi.org/10.18129/B9.bioc.mixOmics>
- Legako, J. F., J. C. Brooks, T. G. O’Quinn, T. D. J. Hagan, R. Polkinghorne, L. J. Farmer, and M. F. Miller. 2015. Consumer palatability scores and volatile beef flavor compounds of five USDA quality grades and four muscles. *Meat Sci.* 100:291–300. <https://doi.org/10.1016/j.meatsci.2014.10.026>
- Legako, J. F., T. T. N. Dinh, M. F. Miller, K. Adhikari, and J. C. Brooks. 2016. Consumer palatability scores, sensory descriptive attributes, and volatile compounds of grilled beef steaks from three USDA Quality Grades. *Meat Sci.* 112:77–85. <https://doi.org/10.1016/j.meatsci.2015.10.018>
- Lepper-Bililie, A. N., E. P. Berg, D. S. Buchanan, and P. T. Berg. 2016. Effects of post-mortem aging time and type of aging on palatability of low marbled beef loins. *Meat Sci.* 112:63–68. <https://doi.org/10.1016/j.meatsci.2015.10.017>
- Liu, J., M.-P. Ellies-Oury, S. Chriki, I. Legrand, G. Pogorzelski, J. Wierzbicki, L. Farmer, D. Troy, R. Polkinghorne, and J.-F. Hocquette. 2020. Contributions of tenderness, juiciness and flavor liking to overall liking of beef in Europe. *Meat Sci.* 168:108190. <https://doi.org/10.1016/j.meatsci.2020.108190>
- Lorenzen, C. L., C. R. Calkins, M. D. Green, R. K. Miller, J. B. Morgan, and B. E. Wasser. 2010. Efficacy of performing Warner-Bratzler and slice shear force on the same beef steak following rapid cooking. *Meat Sci.* 85:792–794. <https://doi.org/10.1016/j.meatsci.2010.03.030>
- MacDonald, J. M., M. E. Ollinger, K. E. Nelson, and C. R. Handy. 2000. Consolidation in U.S. meatpacking. U.S. Department of Agriculture.
- McKeith, F. K., D. L. De Vol, R. S. Miles, P. J. Bechtel, and T. R. Carr. 1985. Chemical and sensory properties of thirteen major beef muscles. *J. Food Sci.* 50:869–872. <https://doi.org/10.1111/j.1365-2621.1985.tb12968.x>
- McKillip, K. V., A. K. Wilfong, J. M. Gonzalez, T. A. Houser, J. A. Unruh, E. A. E. Boyle, and T. G. O’Quinn. 2017. Sensory evaluation of enhanced beef strip loin steaks cooked to 3

- degrees of doneness. *Meat Muscle Biol.* 1:227–241. <https://doi.org/10.22175/mmb2017.06.0033>
- Miller, M. F., L. C. Hoover, K. D. Cook, A. L. Guerra, K. L. Huffman, Ks. Tinney, C. B. Ramsey, H. C. Brittin, and L.M. Huffman. 1995. Consumer acceptability of beef steak tenderness in the home and restaurant. *J. Food Sci.* 60:963–965. <https://doi.org/10.1111/j.1365-2621.1995.tb06271.x>
- Nair, M. N., A. C. V. C. S. Canto, G. Rentfrow, and S. P. Suman. 2019. Muscle-specific effect of aging on beef tenderness. *LWT.* 100:250–252. <https://doi.org/10.1016/j.lwt.2018.10.038>
- O’Quinn, T. G., J. F. Legako, J. C. Brooks, and M. F. Miller. 2018. Evaluation of the contribution of tenderness, juiciness, and flavor to the overall consumer beef eating experience. *Translational Animal Science.* 2:26–36. <https://doi.org/10.1093/tas/txx008>
- O’Quinn T. G., D. R. Woerner, T. E. Engle, P. L. Chapman, J. F. Legako, J. C. Brooks, K. E. Belk, and J. D. Tatum. 2016. Identifying consumer preferences for specific beef flavor characteristics in relation to cattle production and post-mortem processing parameters. *Meat Sci.* 112:90–102. <https://doi.org/10.1016/j.meatsci.2015.11.001>
- Picard, B., M. Gagaoua, D. Micol, I. Cassar-Malek, J.-F. Hocquette, and C. E. M. Terlouw. 2014. Inverse relationships between biomarkers and beef tenderness according to contractile and metabolic properties of the muscle. *J. Agr. Food Chem.* 62:9808–9818. <https://doi.org/10.1021/jf501528s>
- Platter, W. J., J. D. Tatum, K. E. Belk, P. L. Chapman, J. A. Scanga, and G. C. Smith. 2003. Relationships of consumer sensory ratings, marbling score, and shear force value to consumer acceptance of beef strip loin steaks. *J. Anim. Sci.* 81:2741–2750. <https://doi.org/10.2527/2003.81112741x>
- Ross A., C. Brunius, O. Chevallier, G. Dervilly, C. Elliott, Y. Guitton, J. E. Prenni, O. Savolainen, L. Hemeryck, N. H. Vidkjær, N. Scollan, S. L. Stead, R. Zhang, and L. Vanhaecke. 2021. Making complex measurements of meat composition fast: Application of rapid evaporative ionisation mass spectrometry to measuring meat quality and fraud. *Meat Sci.* 181:108333. <https://doi.org/10.1016/j.meatsci.2020.108333>
- Sanchez, G. 2012. DiscrMiner: Tools of the trade for discriminant analysis. <https://rdrr.io/cran/DiscrMiner/>
- Savell, J. W., C. L. Lorenzen, T. R. Neely, R. K. Miller, J. D. Tatum, J. W. Wise, J. F. Taylor, M. J. Buyck, and J. O. Reagan. 1999. Beef customer satisfaction: Cooking method and degree of doneness effects on the top sirloin steak. *J. Anim. Sci.* 77:645–652. <https://doi.org/10.2527/1999.773645x>
- Shackelford, S. D., T. L. Wheeler, and M. Koochmariaie. 1999. Evaluation of slice shear force as an objective method of assessing beef longissimus tenderness. *J. Anim. Sci.* 77:2693–2699. <https://doi.org/10.2527/1999.77102693x>
- Smith, G. C., Z. L. Carpenter, H. R. Cross, C. E. Murphey, H. C. Abraham, J. W. Savell, G. W. Davis, B. W. Berry, and F. C. Parrish Jr. 1985. Relationship of USDA marbling groups to palatability of cooked beef. *J. Food Quality.* 7:289–308.
- Verplanken, K., S. Stead, R. Jandova, C. V. Poucke, J. Claereboudt, J. V. Bussche, S. D. Saeger, Z. Takats, J. Wauters, and L. Vanhaecke. 2017. Rapid evaporative ionization mass spectrometry for high-throughput screening in food analysis: The case of boar taint. *Talanta.* 169:30–36. <https://doi.org/10.1016/j.talanta.2017.03.056>
- Wheeler T. L., R. K. Miller, J. W. Savell, and H. R. Cross. 1990. Palatability of chilled and frozen beef steaks. *J. Food Sci.* 55:301–304. <https://doi.org/10.1111/j.1365-2621.1990.tb06748.x>
- Wheeler, T. L., D. Vote, J. M. Leheska, S. D. Shackelford, K. E. Belk, D. M. Wulf, B. L. Gwartney, and M. Koochmariaie. 2002. The efficacy of three objective systems for identifying beef cuts that can be guaranteed tender. *J. Anim. Sci.* 80:3315–3327. <https://doi.org/10.2527/2002.80123315x>