Peptide fingerprinting and predictive modelling of fermented milk

A thesis presented in partial fulfilment of the requirements for the degree of

Doctor of Philosophy in Food Technology

at Massey University, Palmerston North Campus, New Zealand.

Fionnuala Murphy

BSc Environmental Sciences, MSc Bioinformatics

December 2021

Acknowledgements

I would like to thank my supervisors Dr Stefan Clerens, Dr Julie Dalziel, Dr Esther Meenken and Prof Joanne Hort for their incredible support, encouragement, guidance, and patience throughout the last few years. I feel very privileged to have had such a supportive and approachable supervisory team.

In particular, I would like to express my gratitude to Prof Joanne Hort. There have been many times over the last few years that I have doubted myself; I never once doubted Joanne, and my absolute faith in her as a supervisor provided me with the confidence and certainty to finish this PhD.

I cannot thank Dr Esther Meenken enough for her friendship and mentorship these past few years. I have so much admiration and respect for Esther as a scientist, and I can only hope that someday I will be half the scientist that she is. Thank you for all the wonderful, and sometimes challenging, conversations.

A big thank you to everyone at AgResearch who were always so generous with their time and with sharing their expertise – in particular, Dr Evelyne Maes, Dr Jess Gathercole, Erin Lee and everyone on the Protein and Metabolites Team. Without their help, I would not have made any progress on this project.

Thank you to all of my FEAST colleagues, who were so supportive and encouraging from afar. Especially thanks to Dr Amanda Dupas and Rebecca Orr for their help running the consumer trial.

Thank you to the Fermented Foods team for providing me with this opportunity. I was grateful to be part of such an interesting project, and to learn from the incredible researchers that made up the team.

I am eternally grateful to my friends who have provided incredible motivation and support – Dr Sarah Fitzpatrick, Esther Onguta and Jorie Knook, and to my partner Patrick who was subjected to hearing every struggle I had in great detail.

Publications and presentations

Publications

Murphy, F., Gathercole, J., Lee, E., Homewood, I., Ross, A.B., Clerens, S. and Maes, E., 2021. Discrimination of milk fermented with different starter cultures by MALDI-TOF MS and REIMS fingerprinting. *International Dairy Journal*, *122*, p.105143.

Oral presentations

Murphy, F., Meenken, E., Clerens, S., Dalziel, J., Hort, J., 2021. Exploring the changes in sensory attributes and peptide fingerprints throughout the fermentation of milk. Oral presentation for the 2021 New Zealand Institute of Food Science and Technology (NZIFST) conference, Palmerston North, New Zealand.

Murphy, F., Meenken, E., Clerens, S., Dalziel, J., Hort, J., 2021. A paired comparison study exploring the changes in flavour intensity and bitter taste during the fermentation of milk. Oral presentation for the 13th NZOZ Sensory Symposium, virtual conference, 2021. Awarded best student presentation.

Murphy, F., Meenken, E., Clerens, S., Dalziel, J., Hort, J., 2021. Exploring the sensory and peptidomic changes during the fermentation of milk. Oral presentation prepared for the 2021 Riddet Institute Student Colloquium, Wellington, New Zealand.

Murphy, F., Meenken, E., Clerens, S., Dalziel, J., Hort, J., 2019. Peptide profiling of fermented foods. Oral presentation prepared for the 8th Canterbury Omics Symposium, Christchurch, New Zealand.

Poster presentations

Murphy, F., Meenken, E., Clerens, S., Dalziel, J., Hort, J., 2020. Peptide profiling and predictive modelling of dairy products throughout fermentation. Poster presentation prepared for the 68th American Society of Mass Spectrometry (ASMS), Reboot Virtual Conference.

Murphy, F., Meenken, E., Clerens, S., Dalziel, J., Hort, J., 2019. Peptide profiling of fermented foods. Poster prepared for the 5th International Conference on Food Structures, Digestion and Health (FSDH), Rotorua, New Zealand.

Abstract

Fermented milk products are valued by consumers and the food industry for their nutritional properties, pleasant taste, and texture. Consumer demands and expectations for such products are constantly changing. Understanding how consumers perceive the sensory characteristics of food and the relationship these characteristics have with the chemical components of food can provide insight that can enable food researchers and manufacturers to develop food products that are tailored to provide enhanced sensory qualities. Establishing techniques that allow for in-silico prediction or correlation of sensory qualities can enable a more rapid approach that would aim to enable researchers to meet the demands of consumers.

This research firstly explored mass spectrometric techniques for the rapid fingerprinting of milk and fermented milk products, using Matrix-Assisted Laser Desorption Ionisation - Time-of-Flight Mass Spectrometry (MALDI-TOF MS) and Rapid Evaporative Ionisation Mass Spectrometry (REIMS), two technologies that require minimal sample preparation and can rapidly generate a fingerprint of a food's chemical components. Peptide fingerprints obtained by MALDI-TOF MS and analysed by principal component analysis were effective at discriminating the two fermented milk and milk samples. Supervised discrimination of low molecular weight fingerprints obtained via REIMS and MALDI-TOF MS proved less effective but demonstrated some potential and could be used alongside other analyses in future studies. These techniques were explored with a view to establishing a technique that could provide rapid insights into a food's chemical composition, and which could also effectively discriminate the chemical components of the product. Such techniques could be used for rapid screening of products and can provide insight into the chemical components that are driving the variation in different products, which may be reflective of the differences in sensory characteristics.

Next, peptide fingerprinting and predictive modelling were investigated in milk fermented with various bacterial combinations, including probiotic cultures. Fingerprinting was performed on samples collected at each hour of fermentation. Predictive modelling techniques, using both

regression and classification approaches, were trialled in order to predict the change in signal intensity throughout fermentation. This aimed to understand if peptides could be predicted throughout fermentation, with a view to enable the targeted prediction of desirable peptides, or other relevant components, which may impart favourable sensory qualities in the final product. Regression techniques were somewhat effective for predicting the signal intensity of individual *m/z* ions throughout fermentation. Most of the ions did not follow a linear relationship, and, as such, a multiple linear regression model was unable to model most of the ions. Using a generalised additive model, a non-linear approach, improved the performance in most cases and could predict the signal intensity of individual ions throughout fermentation. However, the model was unable to correctly predict all cases. Classification techniques were effective for predicting the general direction of the signal intensity between start and end fermentation times. Five classification techniques were trialled, with each model providing accurate predictions for the increase or decrease of signal intensity between early and late fermentation times.

Lastly, consumer panellists were recruited to evaluate the change in important sensory characteristics throughout the fermentation of milk prepared using two different starter cultures. This aimed to understand if consumer responses to such products could be correlated with instrumental analysis, in order to predict the consumer responses from instrumental data. Consumers perceived significant differences in bitterness and flavour intensity between fermented milk samples at different fermentation time points. There were significant correlations between peptide fingerprints and the consumer rankings for the sensory attributes in each fermented milk product. XGBoost regression could predict consumer responses with reasonable accuracy.

This thesis explored the fermentation of milk using specific bacteria and fermentation processes. To validate this work, further products could be explored, in addition to different processing parameters. Furthermore, a more in-depth analysis of the chemical components of the products could be investigated and analysed with additional sensory evaluation to further explore and confirm the findings.

List of abbreviations

2-AFC	2-Alternative Forced Choice Test
AIC	Akaike Information Criterion
ANOVA	Analysis Of Variance
AUC	Area Under the Curve
AUCPR	Area Under the Precision-Recall Curve
BB12	Bifidobacterium
CaSR	Calcium-Sensing Receptor
CNN	Convolutional Neural Networks
DHB	2,5-Dihydroxybenzoic Acid
DHPT	2,3,4,5-Tetrakis(3′,4′-dihydroxylphenyl)thiophene
DL	Deep Learning
DNN	Deep Neural Networks
ENaC	Epithelial-Sodium Channel
FDR	False Discovery Rate
FNN	Feedforward Neural Networks
GAM	Generalised Additive Model
GBM	Gradient Boosting Machines
GC	Gas Chromatography
GLM	Generalised Linear Model
HCCA	A-Cyano-4-Hydroxycinnamic Acid
HILIC	Hydrophilic Interaction Liquid Chromatography
HPLC	High-Performance Liquid Chromatography
HSD	Honestly Significant Difference
LA5	*Lactobacillus acidophilus*
LC	*Lactobacillus paracasei* subsp. *paracasei*
LC	Liquid Chromatography
MAE	Mean Absolute Error
MALDI-TOF	Matrix-Assisted Laser Desorption Ionisation - Time-Of-Flight
MCC	Matthew's Correlation Coefficient
MINT-PLS-DA	Multivariate Integrative Partial Least Squares Discriminant Analysis
MS	Mass Spectrometry
MSE	Mean Square Error
PCA	Principal Component Analysis
PERMANOVA	Permutational Multivariate Analysis of Variance
PLS-DA	Partial Least Squares -Discriminant Analysis
PROP	6-N-Propylthiouracil
REIMS	Rapid Evaporative Ionisation Mass Spectrometry
RF	Random Forest
RMSE	Root Mean Squared Error
RNN	Recurrent Neural Networks
SWATH-MS	Sequential Window Acquisition of All Theoretical Mass Spectra
TFA	Trifluoroacetic Acid
TOF	Time Of Flight
TRP	Transient Receptor Potential
WSE	Water-Soluble Extract
XGBoost	Extreme Gradient Boosting

Table of Contents

Chapter 1 Introduction 1

1.1 Background 1

1.2 Rationale and importance 2

1.3 Hypothesis and objectives 3

1.4 Thesis outline 5

Chapter 2 Literature review 7

2.1 Peptides as tastant compounds 7

2.1.1 Bitter peptides 8

2.1.2 Umami peptides 11

2.1.3 Kokumi peptides 13

2.1.4 Salty, sour, sweet peptides 15

2.2 Mass spectrometry techniques to detect tastant compounds 16

2.2.1 Preparation of food samples for MS analysis 17

2.2.2 Rapid techniques for obtaining chemical fingerprints 20

2.2.3 Data processing 24

2.3 Sensory techniques to evaluate tastant compounds 25

2.4 Predictive modelling overview 29

2.4.1 Regression techniques 29

2.4.2 Machine learning 32

2.4.3 Exploratory techniques 36

2.5 Data-driven approaches to evaluate tastant peptides 37
2.6 Challenges in inferring the role of compounds in taste 39
2.7 Fermented milk 41
2.7.1 Manufacture of fermented milk 41
2.7.2 Characteristic taste and consumer preferences for fermented milk 43
2.7.3 Changes in peptide composition throughout fermentation of milk 45
2.8 Literature review conclusions and perspectives 47
2.9 Rationale of the techniques selected for use in this thesis 48
Chapter 3 Discrimination of milk and fermented milk by MALDI-TOF MS and REIMS fingerprinting 50
3.1 Introduction 50
3.2 Materials and methods 52
3.2.1 Samples and reagents 52
3.2.2 Fermented milk preparation 52
3.2.3 MALDI-TOF MS 53
3.2.4 REIMS 55
3.2.5 Data analysis 56
3.3 Results and discussion 57
3.3.1 MALDI-TOF 57
3.3.2 REIMS 65
3.4 Conclusions 69
Chapter 4 Peptide fingerprinting and predictive modelling of fermented milk 71

4.1 Introduction 71
4.2 Materials and methods 74
4.2.1 Chemicals and reagents 74
4.2.2 Experimental design 75
4.2.3 Preparation of fermented milk 76
4.2.4 MALDI-TOF MS sample preparation 79
4.2.5 MALDI-TOF analysis 80
4.2.6 Data pre-processing 80
4.2.7 Data analysis and modelling 80
4.3 Results and discussion 85
4.3.1 Exploration of samples and peptide fingerprints through multivariate analysis 85
4.3.2 Exploration of individual peaks during fermentation using regression techniques 89
4.3.3 Classification exercises to predict change in signal intensity 99
4.4 Conclusions 104
Chapter 5 Consumer evaluation and peptide fingerprinting of milk throughout fermentation ... 106
5.1 Introduction 106
5.2 Materials and methods 108
5.2.1 Consumer study 108
5.2.2 Matrix-assisted laser desorption/ionisation – time of flight Mass Spectrometry (MALDI – TOF MS) 115
5.3 Results and discussion 118
5.3.1 Consumer evaluation throughout fermentation of milk 120

5.3.2 Peptide fingerprinting of samples 126

5.3.3 Correlating consumer responses with peptide fingerprints 131

5.3.4 Prediction of consumer responses from peptide fingerprint data 135

5.4 Conclusions 139

Chapter 6 Discussion and conclusion 140

References 151

Appendices 166

List of Figures

Figure 1.1 Summary of the primary project objectives. This figure was generated using BioRender.com. 4

Figure 2.1 Typical workflow for determining taste-activity in samples using mass spectrometry and sensory analysis. This figure was created using BioRender. 19

Figure 3.1 MALDI MS target plate. Sample is mixed with a matrix and spotted on the target plate prior to being inserted into the instrument. 54

Figure 3.2 To analyse samples using REIMS, milk/fermented milk was placed in an aluminium foil cup; a surgical knife was applied to the surface of the milk. This cuts directly through the sample surface, vaporising it and injecting it into a mass spectrometer via a pump, generating a fingerprint of the metabolites. 55

Figure 3.3 Workflow for sample preparation and data acquisition using MALDI-TOF MS. Samples are pH adjusted, followed by centrifugation to obtain a supernatant. Samples are diluted 1:25 ul TFA and spotted on target plate. Data acquisition can be performed manually or automated to obtain a spectrum **(a).** Workflow for data acquisition using REIMS. Samples require no sample preparation prior to MS analysis. Spectra are obtained in near-real time **(b).** This figure was generated using BioRender.com. 57

Figure 3.4 Raw, unprocessed peptide spectra for milk and fermented milk prepared from cultures CH-1 and YF-L811. Peptide fingerprints obtained by MALDI-TOF MS between m/z 700-3,500. 58

Figure 3.5 PCA plot of peptide fingerprints obtained using MALDI-TOF MS on milk (blue) and milk fermented using CH-1 (red) and YF-L811 (green), with 95% confidence ellipses. N = 5 replicates for each sample type, MALDI-TOF MS replicates (n = 2) were averaged. A PERMANOVA of the model was statistically significant ($P \leq 0.001$). 59

Figure 3.6 Small molecule fingerprints obtained in negative mode by MALDI-TOF MS, between m/z 50-750. Only m/z 50-300 is displayed **(a).** Small molecule fingerprints obtained by MALDI-TOF MS in positive mode, between m/z 50-750 **(b).** Graphene oxide was used as matrix. 61

Figure 3.7 PCA plot of small molecule fingerprints obtained using MALDI-TOF on milk (blue) and milk fermented using CH-1 (red) and YF-L811 (green) in negative mode **(a)** and in positive mode **(b)**, with 95% confidence ellipses. N = 5 replicates for each sample type, MALDI-TOF replicates (n = 2) were averaged. A PERMANOVA of each model was statistically significant (negative mode $p \leq 0.001$, positive mode $p = 0.004$)..... 63

Figure 3.8 Milk (blue) and fermented milk samples prepared from starter cultures CH-1 (red) and YF-L811 (green). PLS-DA performed on small molecules obtained by MALDI-TOF MS in negative mode **(a)** and in positive mode **(b)**, with 95% confidence ellipses. 64

Figure 3.9 Negative mode profiles obtained by REIMS, between m/z 0 – 900 **(a).** Positive mode profiles obtained by REIMS, between m/z 0 - 900 **(b).**..... 66

Figure 3.10 Milk (blue) and fermented milk samples prepared from starter cultures CH-1 (red) and YF-L811 (green). PCA plot of metabolites obtained by REIMS in negative mode **(a)** and positive mode **(b)**, with 95% confidence ellipses. A PERMANOVA of the models was statistically significant (negative mode $p \leq 0.001$; positive mode $p \leq 0.001$)**.** N = 5 replicates for each sample type, REIMS replicates (n = 3) were averaged. 67

Figure 3.11 PLS-DA of metabolites from REIMS in negative mode **(a).** PLS-DA of metabolites from REIMS in positive mode **(b)**, with 95% confidence ellipses. 68

Figure 4.1 Workflow to demonstrate the preparation of fermented milk. Skimmed milk powder was reconstituted, followed by a heat treatment, and a 1:1 inoculation of starter: probiotic. Samples were fermented at 43 °C for up to 5 hours. Samples were removed each hour and placed on ice. This figure was generated using BioRender.com. 77

Figure 4.2 Placement of tubes/fermented milk in the water bath. The position of each tube in the water bath was randomised according to the experimental design prepared in CycDesignN. Each tube corresponds to a different starter/probiotic/fermentation time point. 79
Figure 4.3 PCA plot of the peptide fingerprints, for all replicates, coloured by fermentation time **(a)**, starter culture **(b)**, probiotic culture **(c).** A PERMANOVA of each model indicated significant differences between the treatments (P<0.05). 87
Figure 4.4 Multiple linear regression model of m/z value 1860.65. Fermentation hour, starter culture and probiotic terms were used as predictor terms to predict the response of signal intensity. The signal intensity is along the y-axis, fermentation time is along the x-axis. The plot is faceted by starter culture and coloured by probiotic culture. Legend indicates the corresponding probiotic culture. Formula: lm(Signal_Intensity ~ Time * Starter * Probiotic). 91
Figure 4.5 Model diagnostics for multiple linear regression of Signal Intensity ~ Time * Starter * Probiotic. Model diagnostic plots: Raw data plotted by signal intensity ~ time **(a)**, a scatter plot of residuals vs fitted values **(b)**, a histogram of residuals **(c)**, QQ plot of residuals **(d)**. 92
Figure 4.6 Multiple linear regression model of m/z value 1189.79, fermentation hour, starter culture and probiotic terms were used as predictor terms to predict the response of signal intensity. The signal intensity is along the y-axis, fermentation time is along the x-axis. The plot is faceted by starter culture and coloured by probiotic culture. Legend indicates the corresponding probiotic culture. Formula: lm(Signal_Intensity ~ Time * Starter * Probiotic). 93
Figure 4.7 Multiple linear regression model of m/z value 1189.79, fermentation time, starter culture and probiotic terms were used as predictor terms to predict the response of signal intensity. The signal intensity is along the y-axis, fermentation time is along the x-axis. The plot is faceted by starter culture and coloured by probiotic culture. 94
Figure 4.8 GAM model of m/z 1189.79, fermentation time, starter culture and probiotic terms were used as predictor terms to predict the response of signal intensity. The signal intensity is along the y-axis, fermentation time is along the x-axis. The plot is faceted by starter culture and coloured by

probiotic culture. Confidence intervals are indicated by the shaded area around the dotted line. Legend indicates the corresponding probiotic culture. GAM formula: bam(Signal_Intensity ~ Starter + Probiotic +s(Time, k = 6) + s(Time, by = interaction(Starter, Probiotic), k = 6)). 97
Figure 4.9 Diagnostic plot of GAM model on m/z 1189.79. Diagnostic plots are a QQ plot comparing model residuals to a normal distribution **(a)**, a plot of the residuals **(b)**, a histogram of residuals **(c)**, and a plot of response vs fitted values **(d)**. 98
Figure 4.10 Top 40 ranking peaks selected by discriminant analysis. To identify the most important class discriminating peaks, standard t-scores are used. The peaks (rows) in the feature matrix are ranked by their t-score. Fermentation times are along the columns. 100
Figure 4.11 Boxplot of signal intensity for each m/z value selected through discriminant analysis; comparison between Time 0 (pink) and Time 4 (blue). 25 out of the 40 m/z values were decreasing between start and later fermentation times. Vertical lines represent minimum and maximum, horizontal line represents the median. Data below the line is in quartile 1, data above the median is quartile 3. The black dots indicate outliers. 101
Figure 5.1 Fermented milk for the consumer study prepared at the Massey University Food Pilot Plant. Milk was reconstituted by mixing skim milk powder and water using an overhead stirrer (a). The milk was heat-treated using a steam kettle to rapidly heat the milk to 85 °C (b). The milk was then held in a water-jacketed kettle, set to 85 °C for 30 mins. Fermentation took place in the water-jacketed kettle (c). 108
Figure 5.2 The sessions took place in individual partitioned sensory booths. Participants were presented with pairs of samples on trays, with random codes assigned to each sample. The presentation of the samples was randomised for each participant. Participants were presented with a screen on an iPad with instructions to consume samples and evaluate the sample for bitterness, or flavour intensity. 113
Figure 5.3 Samples collected at hour 3 for YF-L811 had whey separation. This defect was found only in this sample. 118

Figure 5.4 mean pH (n=3) recorded on day of preparation and after 4 days in storage for samples prepared using the starter culture YF-L811 (a), and pH recorded on day of preparation and after 4 days in storage for samples prepared using the starter culture YC380 (b). Post-acidification occurred while the samples were in storage. Average of three measurements for each sample.. 119

Figure 5.5 Rank sum scale of bitterness for YF-L811 (n=41 participants) and YC380 (n=43 participants). The arbitrary numbers correspond to the samples rank position within the sample set. Different lettering above the samples indicates a significant increase in bitterness (at α=0.05), determined by Tukey's HSD test. The samples prepared for each starter culture increase in perceived bitterness, with an increased fermentation time. 121

Figure 5.6 d` values for bitterness in YF-L811 (a) and YC380 (b). The 15 pairs of samples are organised on the x-axis, in order of increasing fermentation time. The d` denotes the sensory distance between the more fermented vs less fermented samples in each pairing..... 122

Figure 5.7 Rank sum scale of flavour intensity for YF-L811(n=41 participants) and YC380 (n=43 participants). The arbitrary numbers correspond to the samples rank position within the sample set. Different lettering above the samples indicates a significant increase in bitterness (α=0.05), determined by Tukey's HSD test. The samples prepared for each starter culture increase in perceived flavour intensity, with an increased fermentation time. 124

Figure 5.8 d` values for flavour intensity in YF-L811 (a) and YC380 (b). The 15 pairs of samples are organised on the x-axis, in order of increasing fermentation time. The d` denotes the sensory distance between the more fermented vs less fermented samples in each pairing..... 125

Figure 5.9 pH profiles of replicate samples prepared in Lincoln (labelled Lin) and in Palmerston North (PN) for YF-L811 (a) and YC380 (b). Average of three measurements for each sample. A two-way ANOVA (pH ~ location * time) for each starter type was not significant (p = 0.679 for YF-L811, p = 0.350 for YC380). 127

Figure 5.10 A Sparse MINT-PLSDA on peptide fingerprints from samples prepared using YF-L811. Samples are colour-coded by fermentation time point. The different shapes correspond to

location of sample preparation: Lin = Lincoln, PN = Palmerston North. Five replicates were prepared for each sample in Lincoln. Three measurements were taken using MALDI-TOF MS for each sample type. 129
Figure 5.11 A Sparse MINT-PLSDA on peptide fingerprints from samples prepared using YC380. Samples are colour-coded by fermentation time point. The different shapes correspond to location of sample preparation: Lin = Lincoln, PN = Palmerston North. Five replicates were prepared for each sample in Lincoln. Three measurements were taken using MALDI-TOF MS for each sample type..... 130
Figure 5.12 Principal components analysis of peptide fingerprints for samples prepared using YF-L811. Data is colour-coded by fermentation time point. Large dots indicate the average for that sample type. Five replicates were prepared for each sample type. MALDI-TOF technical repeats were averaged. Samples measured were prepared in Lincoln. 131
Figure 5.13 Principal components analysis of peptide fingerprints for samples prepared using YC380. Data is colour-coded by fermentation time point. Large dots indicate the average for that sample type. Five replicates were prepared for each sample type. MALDI-TOF technical repeats were averaged. Samples measured were prepared in Lincoln. 132
Figure 5.14 The Spearman correlation was performed by extracting the top four principal components from PCA of peptide fingerprints and performing a correlation with the consumer scores obtained for bitter intensity in YF-L811 (a) and YC380 (b). Consumer rankings for bitterness are denoted along the y-axis; the x-axis represents the extracted principal component scores. .. 133
Figure 5.15 The Spearman correlation was performed by extracting the top four principal components and performing a correlation with the consumer scores obtained for flavour intensity in YF-L811 (a) and YC380 (b). Consumer rankings for flavour intensity are denoted along the y-axis; the x-axis represents the extracted principal component scores. 134
Figure 5.16 Predicted vs Observed values for XGBoost model of flavour intensity on the principal components of peptide fingerprints obtained from YF-L811 samples. 135

Figure 5.17 Predicted vs Observed values for XGBoost model of flavour intensity on the principal components of peptide fingerprints obtained from YC380 samples..... 136
Figure 5.18 Predicted vs Observed values for XGBoost model of bitter intensity on the principal components of peptide fingerprints obtained from YF-L811 samples. 137
Figure 5.19 Predicted vs Observed values for XGBoost model of bitter intensity on the principal components of peptide fingerprints obtained from YC380 samples..... 138
Figure 6.1 A summary of the primary project objectives and their potential applications. This thesis was divided in to three chapters to address the thesis objectives: Clustering and discrimination of fermented milk prepared from different bacterial cultures; Prediction of signal intensity throughout fermentation of milk, and evaluation and prediction of consumer responses throughout fermentation of milk. This figure was generated using BioRender.com. 142
Figure 6.2 Opportunities for future work to further the work explored in this thesis..... 149

List of Tables

Table 1 Overview of sensory tests performed to identify taste activity in samples originating from fermented dairy foods 27

Table 2 Factor levels used in experimental design. 75

Table 3 An example of the placement of samples in the water bath according to experimental design generated in CycDesignN. The numbers correspond to the Starter, Probiotic, Fermentation Time Point, where e.g., (2,3,5) corresponds to Starter YC380, Probiotic LC and fermentation time point 5 (e.g., 5 hours). 76

Table 4 Comparing added terms for each peak (n = 143). Performance and significance of models with added complexity were assessed by AIC and ANOVA. m/z signal intensity was used as the response variable. 89

Table 5 Adjusted R^2 and p-values of model with full interaction terms. This is a subset of peaks with the highest and lowest R^2 values. P-values were adjusted using Benjamini-Hochberg correction. 90

Table 6 Multiple linear regression model vs general additive model across all peaks. Models were compared by AIC, ANOVA and R^2. 95

Table 7 Comparison of adjusted R^2 values outputted by the multiple linear regression and GAM models. These are organised by lowest/highest based on R^2 from the linear regression model. ... 96

Table 8 Performance metrics for six models on training data: XGBoost, Gradient Boost Machine, Deep Learning, C4.5 and Random Forest. Each model was trained using a 70% subset of the data. 102

Table 9 Performance metrics for six models on test data: Deep Learning, Gradient Boost Machine, XGBoost, C4.5 and Random Forest. 103

Table 10 Confusion matrix showing correctly and incorrectly predicted labels for deep learning model on data from hour 0 vs hour 4. For the label ‘Decrease’, 195 were correctly labelled; 9 were

incorrectly labelled as ‘Increase’. For the label ‘Increase’, 109 were correctly labelled; 28 were incorrectly labelled as ‘Decrease’. 103
Table 11 Random, blind codes assigned to each sample type and the actual sample it corresponds to. The positioning of each sample is partially balanced and randomised per participant. 112
Table 12 Experimental design; position one in the three-digit code represents the starter culture, followed by fermentation time and replicate. 115
Table 13 Spearman correlation of top four components from a PCA of peptide fingerprints for YF-L811 and YC380 on consumer responses to bitter intensity. * Indicates a significant correlation ($p \leq 0.05$). 133
Table 14 Spearman correlation of top four components from a PCA of peptide fingerprints for YF-L811 and YC380 on consumer responses to flavour intensity. Significant correlations are denoted by “*”. 134
Table 15 XGBoost model metrics for predictions of consumer responses to flavour intensity for YF-L811 and YC380. 135
Table 16 XGBoost model metrics for predictions of consumer responses to bitter intensity for YF-L811 and YC380. 137

Chapter 1 Introduction

1.1 Background

The fermentation of milk transforms it into various products with an extended shelf-life and is valued by consumers for its enhanced sensory and nutritional properties (Tamime and Robinson, 1999). Milk fermentation involves the inoculation of a starter culture, which drives the fermentation process. The activity of this culture results in the generation of the various compounds that impart the typical flavour and taste of fermented milk (Tamime and Robinson, 1999). Different starter cultures can create products with variations in these characteristics; adjunct cultures can also be added to enhance sensory and nutritional properties (Routray and Mishra, 2011). The Codex standard defines fermented milk as a product obtained by fermentation of milk and includes yoghurt, Kefir, and Acidophilus milk, amongst others (WHO/FAO, 2003).

The flavour profile of fermented milk is widely known to result from a combination of compounds such as lactic acid and acetaldehyde (Tamime and Robinson, 1999). However, the taste and flavour may also be affected by the proteolytic activity during fermentation; peptides and amino acids generated via proteolysis can have a direct and indirect impact on the taste and flavour by contributing to bitter, umami and kokumi sensations (Zhao et al., 2016) and by acting as precursors of flavour-producing reactions (Kilara and Panyam, 2003). Additionally, bioactive peptides may be generated which can contribute to the nutritional properties, adding to the commercial value of the product (Fijan, 2014).

There has been an increased focus on understanding how tastant compounds develop in various fermented dairy foods in recent years (Toelstede and Hofmann, 2008a, Toelstede and Hofmann, 2008b, Sebald et al., 2018, Sebald et al., 2019, Murray et al., 2018, Schäfer et al., 2019, Alim et al., 2020). These approaches have enabled researchers to explore the potential role of peptides, and other relevant compounds, in generating taste in fermented dairy products, with a view to improving the sensorial and nutritional quality of the product to make it more desirable to

consumers. Furthermore, the application and advancement of machine learning and predictive techniques have aided researchers in understanding and predicting fermentation processes and sensory properties, which can be utilised to improve and monitor product development to ensure product quality (Rocchetti et al., 2018, Dalabasmaz et al., 2019, Li et al., 2020b, Piras et al., 2021). The ability to rapidly screen different products and provide insight into the fermentation process as well as the generation of important compounds and sensory properties throughout the fermentation of milk could permit the development of new, improved products that are appreciated by consumers.

1.2 Rationale and importance

There is a growing demand globally for quality fermented dairy products, and the expectations and demands of consumers are constantly changing (Conti-Silva and Souza-Borges, 2019, Martins et al., 2018, Aryana and Olson, 2017, Chollet et al., 2013). In 2018, the global fermented milk market was valued at US$264.7 billion; this is expected to grow to $396.9 by 2026 (Allied Market Research, 2020). New Zealand is reputed worldwide for producing high-quality food products, in particular dairy foods (Tait et al., 2018).

This PhD project was part of a wider programme (Fermented Foods) which aimed to accelerate the evolution of bacterial cultures to select for new cultures which could impart interesting flavours to fermented products. The project presented here explored a strategy to enable a more rapid and high-throughput screening of products and to provide an instrumental approach to understand consumer responses to such products. This approach was with a view to providing New Zealand dairy manufacturers, as well as the dairy industry as a whole, with valuable insights that may enable the speedy development of new and interesting dairy products with targeted desirable properties. A suitable dairy food was required to investigate the feasibility of the approach taken for this project. Fermented milk was selected for this project as it is a product with a growing popularity worldwide and an increasing market share in the food and dairy industry (Aryana and Olson, 2017, Rul, 2017). Fermented milk also has a short fermentation time, lending itself to quick analysis.

1.3 Hypothesis and objectives

The overall objective of this thesis was to evaluate rapid methods to provide insight into the molecular composition of fermented milk, and the sensory characteristics which may be affected by these molecules. The working hypothesis for this thesis was that the chemical fingerprint, which represents the unique pattern of molecular components in the sample, could provide us with insight that can be used for developing and screening new products. More specifically, that the peptide fingerprint of a fermented milk product changes throughout fermentation in a way that can be predicted and linked with changing tastant characteristics, as perceived by the consumer. This thesis focussed on how a fingerprint can be utilised to extract knowledge and insight into fermented milk products, without basing the assumptions and driving factors on the biological significance of such components, i.e., the detected compounds were assumed to be significant (with flavour/tastant properties) and following this their relationship and biological significance could be established, although the latter part of this was not investigated in this thesis.

This thesis sought to answer the following research questions:

1. Can milk fermented from different starter cultures be discriminated using rapid fingerprinting technologies coupled with multivariate modelling techniques?
2. Is the peptide fingerprint different across fermentation times and for different starter and probiotic culture combinations?
3. Can the peptide fingerprint be predicted across fermentation time points for different bacterial cultures?
4. Can consumers perceive differences in important sensory attributes throughout fermentation?
5. Can the peptide fingerprint be correlated with consumer responses and be used to predict such responses?

With these questions in mind, the following objectives, as summarised in Figure 1.1, were set:

- Investigate appropriate techniques to rapidly discriminate fermented milk samples prepared from different culture combinations.
- Develop an approach for predicting the changes in the peptide fingerprint throughout fermentation.
- Evaluate the changes in key sensory attributes throughout fermentation using consumer panellists.
- Integrate consumer and peptide fingerprints to correlate these two measures with a view to understanding the change in consumer response for different treatments.

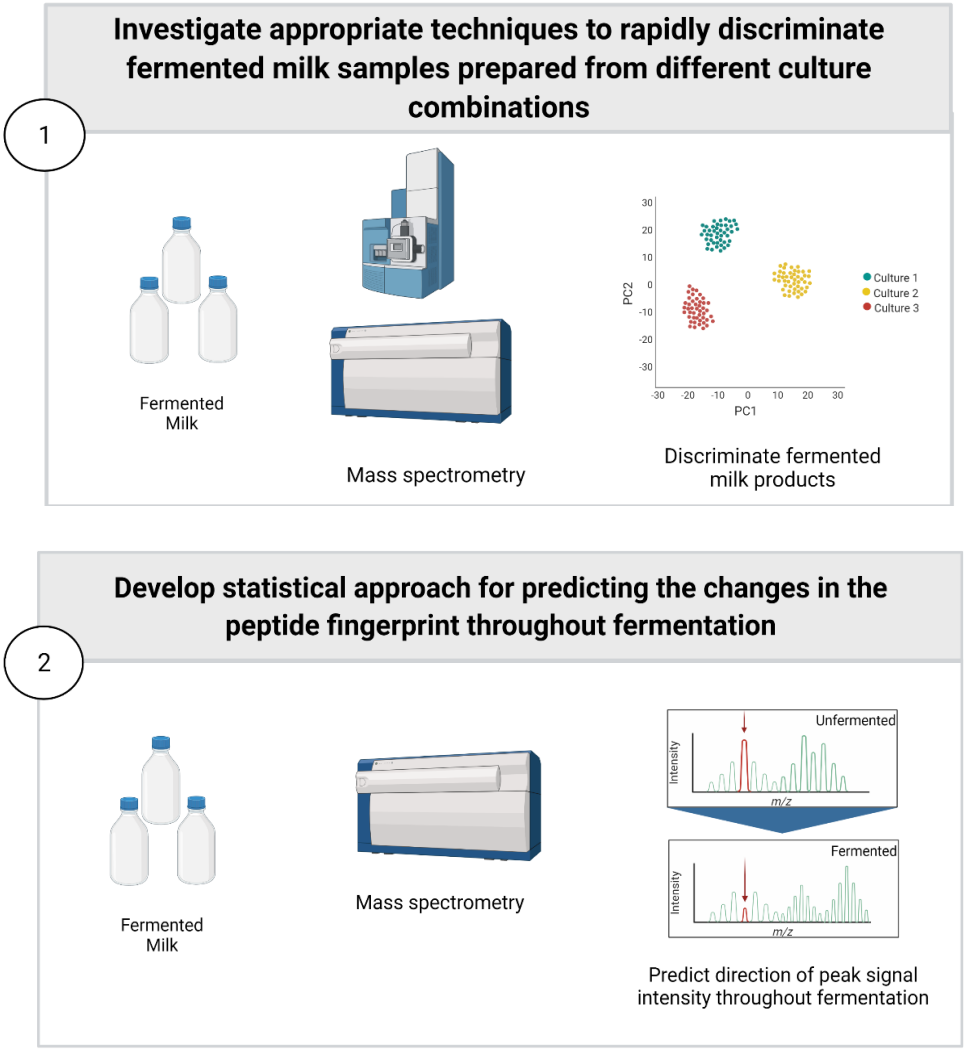
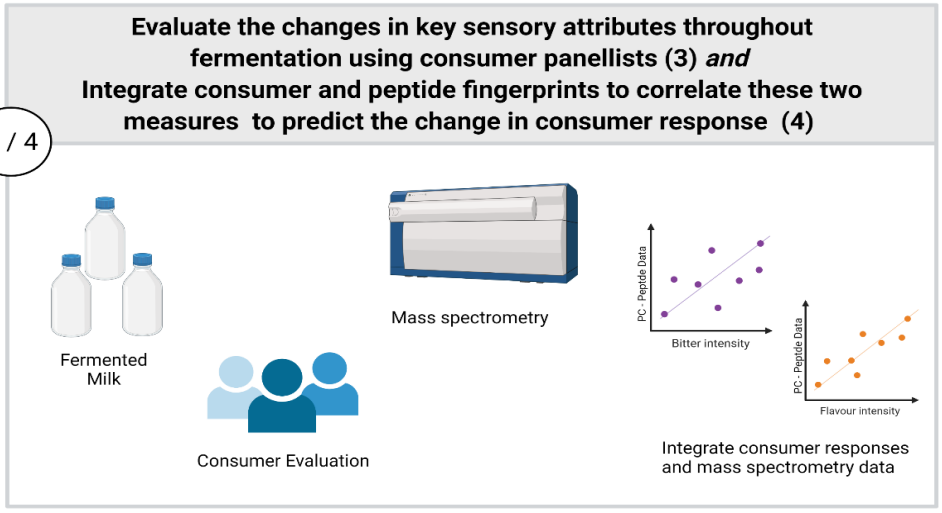


Figure 1.1 Summary of the primary project objectives. This figure was generated using BioRender.com.

Figure 1.1 continued



1.4 Thesis outline

To achieve the project objectives, the study was conducted in four stages, organised into the following chapters.

Chapter 2 Literature review

The literature review provides context and justification for some of the techniques used in this thesis, in addition to relevant literature that aims to summarise the techniques and tools used throughout. The contribution of peptides to taste in fermented dairy foods is described, as well as mass spectrometric, data and sensory analysis techniques used to investigate tastant peptides. Additionally, the mass spectrometry approach and modelling techniques used in this thesis were summarised. Finally, fermented milk is described in relation to its production, taste, and flavour.

Chapter 3 Discrimination of milk and fermented milk by MALDI-TOF MS and REIMS fingerprinting

A study (*research question 1*) was conducted to evaluate the potential for rapid fingerprinting technologies to discriminate the peptides and small molecules of fermented milk. Two high-throughput techniques requiring minimal sample preparation, MALDI-TOF MS (Matrix-Assisted Laser Desorption Ionisation - Time-of-Flight Mass Spectrometry) and REIMS (Rapid Evaporative

Ionisation Mass Spectrometry), were employed to fingerprint and discriminate fermented milk prepared from different starter cultures. This study provided a means for rapid screening of fermented milk samples. The fingerprinting method and compound type that was most effective for discriminating these samples was employed in Chapter 4. This work was published in part as a manuscript entitled 'Discrimination of milk fermented with different starter cultures by MALDI-TOF MS and REIMS fingerprinting'.

Chapter 4 Peptide fingerprinting & predictive modelling

Chapter 4 (*research questions 2 & 3*) explores the feasibility of predicting the peptide fingerprint throughout the fermentation of milk prepared from different combinations of starter and probiotic cultures. This study provided an understanding of how the peptide fingerprint can change throughout fermentation in different culture combinations, and how it can be exploited to provide a means for targeted tracking of peptides throughout fermentation. Findings from this chapter, in terms of the most discriminating factors in the analysed treatments, were used for the basis of Chapter 5.

Chapter 5 Consumer evaluation and peptide fingerprinting of milk throughout fermentation

This chapter (*research questions 4 & 5*) describes a consumer study to evaluate the changes in key sensory attributes throughout fermentation, in order to provide insights on how these attributes develop and continue to change during fermentation. In parallel, samples were fingerprinted for peptides to establish if a correlation existed between the peptide fingerprints and the consumer responses to these products.

Chapter 6 Discussion and conclusion

The final chapter provides a summary of the main findings and applications of this work, as well as suggestions for future research.

Chapter 2 Literature review

Summary

This chapter presents a summary of literature that is considered relevant to this thesis. It describes relevant literature concerning tastant peptides in fermented dairy foods, and the mass spectrometric, sensory and data analysis techniques that may be used to detect and evaluate their tastant properties. Lastly this review describes fermented milk production and important properties of fermented milk. Much of the literature described in this chapter pertains to peptides in cheese, this is due to the majority of published literature on tastant peptides focussing on cheese, rather than other fermented dairy products.

2.1 Peptides as tastant compounds

The basic taste sensations are made up of sweet, sour, salty, bitter and umami and are evoked through specialised taste receptor cells in the mouth (Reineccius, 2006). The kokumi sensation was discovered in recent years and is associated with "mouthfulness", "continuity" and "thickness". Although they are not reported to be tasty on their own, kokumi-active compounds can enhance other tastes (Ueda et al., 1997). Peptides are said to impart taste across all taste sensations, although it is understood that in most cases, they contribute to background tastes and are not responsible for the primary taste of the food (Temussi, 2012). In a variety of fermented and aged foods, however, peptides can be potent and contribute primarily to bitter, umami and kokumi sensations (Zhao et al., 2016). The taste of peptides, and amino acids, is complex, and it can be difficult to assign and verify their taste. These compounds have the potential to elicit multiple taste qualities and are affected by their interactions with other compounds, as is the case with bitter peptides in fermented fish sauce, which were reported to have sweet or umami tastes in the presence of NaCl (Park et al., 2002). It is a similar case with amino acids; L-Thr, for example, has been associated with both sour and sweet tastes in cheese (Fox et al., 2017). Aside from the direct contribution that peptides can have on taste, they also act as precursors for flavour production. In fermented dairy products, many

important flavour compounds are derived from casein (Liu et al., 2008, Kilara and Panyam, 2003). Cadwallader and Singh (2009) noted the absence of research on tastant compounds in dairy products; however, recently, there has been a greater focus on understanding tastant compounds in various dairy foods (Toelstede et al., 2008a, Toelstede et al., 2008b, Sebald et al., 2018, Sebald et al., 2019, Murray et al., 2018, Schäfer et al., 2019, Alim et al., 2020).

2.1.1 Bitter peptides

Bitter peptides are found in a variety of fermented dairy foods and, along with amino acids, can activate at least five of the 25 human bitter taste receptors, known as the TAS2Rs (Kohl et al., 2013). Although they rarely have a significant impact on the bitter taste, some bitter peptides are known to be very potent (Fox et al., 2017) and are present in various types of cheeses where bitterness is usually viewed as a major defect. Small bitter peptides have also been reported in various aged meats (Sforza et al., 2001, Sentandreu et al., 2003) and fermented fish sauces (Park et al., 2002, Schindler et al., 2011). Humans are sensitive to bitterness and can detect it in micromolar amounts (Drewnowski and Gomez-Carneros, 2000); this, coupled with the negative perception consumers often have of bitterness in dairy foods, can be problematic to the development and quality of these products and, as a result, a great deal of research has been carried out on the mechanism of bitter peptides, as well as the potential for their debittering.

2.1.1.1 Mechanisms for bitterness evoked by peptides

The intensity of bitterness elicited by bitter peptides has been related to the number of hydrophobic amino acids, the sequence of the amino acids, and the size of the peptide (Fox et al., 2017). In general, the bitterness of a peptide can increase with an increasing number of amino acids, but only up to eight amino acids; the potency of the peptide does not substantially increase with more than seven amino acids (Tamura et al., 1990). The position and abundance of certain amino acids in the peptide can impact the bitterness (Ishibashi et al., 1987a, Ishibashi et al., 1987b, Ishibashi et al., 1988a, Ishibashi et al., 1988b). A key determinant for bitterness in di- and tripeptides is a bulky hydrophobic amino acid at the C-terminal and a bulky amino acid at the N-terminal (Kim and Li-

Chan, 2006, Xu et al., 2019). In larger peptides (i.e., four or more), it was recently shown that hydrophobic amino acids at the C-terminal were a less important determinant for bitterness than was previously thought (Xu and Chung, 2019).

Bitter peptides in fermented dairy foods are known to be mostly casein-derived, particularly bovine β-casein, which is highly hydrophobic. Casein itself does not have a bitter taste; however, when hydrolysed, it can release bitter-tasting peptides (Kilara and Panyam, 2003). The peptide β-CN f193-209, for example, has a potent bitter taste, and it can accumulate in a variety of cheese types (Appendix 1). Numerous α-CN-derived peptides are also reported to have a bitter taste, many of which are products of the breakdown of αs1-CN f1-23. Several bitter peptides are also reported to have bioactive properties, such as β-CN f193-209 (Ong and Shah, 2008), αs1-CN fl-13 (Saito et al., 2000), and αs1-CN f1-9 (Saito et al., 2000, Ong and Shah, 2008) which can exhibit antihypertensive effects, as well as playing a role in immune defence and nervous system activity. κ-casein was not previously thought to be a source of bitterness, but several bitter peptides derived from κ-casein were recently identified in cheese (Sebald et al., 2018, 2019). Casein-derived peptides are typically abundant in Pro, which contributes to their bitter taste as Pro is a source of bitterness in peptides. The role of Pro in bitterness is unrelated to its hydrophobicity, rather it is associated with the conformational alteration of the peptide skeleton, as well as the position of the amino acid in the peptide sequence; the bitter taste is reported to be more intense when Pro is situated at the C-terminus (Ishibashi et al. 1988b). The structure of peptides containing Pro also allow them to easily bind to the bitter taste receptor (Tamura et al., 1990). It was suggested that the distribution of Pro in the peptide can make it less susceptible to degradation, thereby hindering the breakdown of the peptide into smaller, non-bitter products (Lee et al., 1996).

The association of bitterness and hydrophobicity can be used as a reliable indicator of the peptide's potential to be bitter-tasting. Ney (1971) derived a rule based on this association: Ney proposed that a peptide must have a hydrophobicity, or Q-value, greater than 1,400, as well as a molecular weight < 6000 Dalton, to be bitter tasting. The Q value is defined as the average amount of free

energy required for the transfer of amino acid side chains from ethanol to water (Ney, 1971). This rule is widely used to predict the bitterness of peptides (Nielsen et al., 2017, Murray et al., 2018, Sebald et al., 2019). Some inconsistencies with the Q-rule have been reported, however (Toelstede and Hofmann, 2008b), and it may not be reliable in all cases as it does not consider the position of amino acids in the peptide, which is known to affect bitterness (Murray et al., 2018).

2.1.1.2 Development of bitter peptides in fermented dairy foods

In fermented dairy foods, the generation of bitter peptides (Appendix 1) is related to the activity of the lactic acid bacteria. Certain starter cultures may have a propensity for generating fewer bitter peptides or may contain enzymes that target bitter peptides for hydrolysis, thus reducing their bitter taste (Broadbent et al., 2002). In yoghurt, the proteolytic activity of *Lactobacillus bulgaricus* is mostly responsible for the production of bitter peptides, and an increased inoculation rate of this bacterium can lead to a bitter taste (Tamime and Robinson, 1999). The addition of adjunct cultures can aid in reducing bitterness by targeting bitter peptides; when *Lactobacillus helveticus* was used as an adjunct culture in Cheddar cheese production it resulted in the reduced concentration of the bitter peptide, β-CN f193-209 (Soeryapranata et al., 2002).

During the fermentation process, bitter peptides may be liberated from the native protein via hydrolysis. Subsequent hydrolysis can reduce these products to amino acids and derivatives, thereby eliminating the bitter tasting peptides (Newman et al., 2014). Unequal rates of proteolysis and peptide hydrolysis, however, can cause bitter peptides to accumulate; excessive degradation of casein occurs via primary proteolysis, releasing bitter peptides. This was the case with Ragusano cheese where there were significantly greater rates of primary proteolysis than peptide hydrolysis leading to an increased concentration of bitter peptides (Fallico et al., 2005). This imbalance of hydrolysis was related to a combination of high salt-in-moisture and low moisture levels. In Prato cheese, the bitter peptide *α*s1-CN f1-13 accumulated more in cheeses with salt-in-moisture concentrations greater than 5%, compared to cheeses prepared at lower concentrations. It was speculated that the high salt content may inhibit bacterial activity, reducing the rate of hydrolysis

(Baptista et al., 2017). Conversely, in milk containing hydrolysed dairy proteins, excessive hydrolysis resulted in an accumulation of bitter peptides causing an increase in bitterness (in addition to higher salty and umami tastes), compared to samples that were only partially hydrolysed (Alim et al., 2020). In Parmesan cheese, an increasing salt concentration during ageing was found to inhibit the breakdown of the peptide α_{s1}-CN (1–23) by interfering with endoproteases (Sforza et al., 2012). The percentage of fat may also suppress bitterness; bitter peptides may reside in the fat layers of high-fat dairy products, masking the bitter taste (Singh et al., 2005). Other processing parameters can also impact the generation of bitter peptides. Rocha et al. (2020) related a lower bitter taste to a decreased production of bitter peptides in cheese samples that were prepared from milk heat-treated with ohmic heating, compared to conventionally heat-treated milk. Yoghurts fermented at 38°C were more likely to have a bitter taste, attributed to bitter peptides, compared to yoghurts fermented at 44°C (Tamime and Robinson, 1999). Bitter peptides may also accumulate during storage. In hydrolysed-lactose milk samples, for instance, the bitter taste increased significantly during storage and was attributed to an increase in bitter peptides that formed via residual proteolytic activity (Nielsen et al., 2017). Schäfer et al. (2019) found a decreased bitterness intensity and bitter peptide production in cheese samples that were prepared from fermented skim milk retentates. The increased bitterness was associated with a decreased content of calcium; however, the bitterness and concentration of bitter peptides increased during storage.

2.1.2 Umami peptides

Umami taste can be found naturally in many foods and is described as a “pleasantly savoury” sensation (van den Oord and van Wassenaar, 1997). The principal umami taste receptor is the TAS1R1/TAS1R3 receptor, which is stimulated by L-Glu, the primary compound responsible for imparting an umami taste (Nelson et al., 2002). This receptor is also referred to as the “amino acid taste receptor” as it responds to numerous amino acids (Nelson et al., 2002). Umami taste is a desirable property in foods, and a better understanding of its formation and the compounds that generate its taste can aid to enhance the taste of food and consumer acceptance.

The role of peptides in forming an umami taste has been controversial (Yamasaki and Maekawa, 1978, van den Oord and van Wassenaar, 1997, Maehashi et al., 1999). There have been at least 52 reports of peptides with an umami taste in various foods; however, at least 20 of these are in dispute (Zhang et al., 2017). Umami peptides have been identified in various foods, such as soy sauce (Kaneko et al., 2011, Zhuang et al., 2016), aged meats (Dang et al., 2015), and a variety of food-derived hydrolysates (Arai et al., 1972, Noguchi et al., 1975, Tamura et al., 1989, Maehashi et al., 1999, Schlichtherle-Cerny and Amadò, 2002). In fermented dairy foods, however, it is suggested that in most cases, peptides likely do not have a significant impact on the umami taste (Engel et al., 2000, Toelstede et al., 2009). Peptides with a reported umami taste are highly hydrophilic (Ardö and Varming, 2010), and will usually contain Glu in their sequence. Dipeptides containing Glu were first reported in Comté cheeses by Roudot-Algaron et al. (1994), some of which were said to have an umami taste, as well as having sour, bitter and salty notes. The authors were apprehensive about conclusively assigning an umami taste to these peptides but did conclude that the peptides contributed to the overall complexity of flavour in the cheese. As a free amino acid, Glu is thought to contribute to umami taste in many cheese varieties. In Swiss and Cheddar cheeses, Glu was primarily responsible for generating umami taste, though organic acids were also suggested to play a role in Swiss cheese (Drake et al., 2007). In Manchego cheese fractions, umami taste was related to Glu, as well as Asp (Taborda et al., 2008). Although in some umami-tasting fractions, the taste was unrelated to the concentration of Glu and Asp and was suggested to be associated with the presence of small hydrophilic peptides, which contained Glu-Glu in their sequences (Taborda et al., 2008). Gomez-Ruiz et al. (2007) also found several known umami peptides abundant in Glu in umami-tasting fractions of Manchego cheese. However, the peptides were undetected in some umami-tasting fractions, leaving uncertainty as to the contribution of peptides to umami taste in this cheese. Di- and tripeptides containing Glu have been linked with an umami taste in Cheddar cheese (Andersen et al., 2010, Andersen et al., 2008). Li et al. (2020) reported the presence of several small peptides, with known umami taste, namely Asp-Trp-Asp-Ser, Glu-Glu, Asp-Leu, and

Glu-Leu, in fermented brown milk. Umami-tasting amino acids and peptides may also play a role in suppressing salty and bitter tastes; the omission of Glu led to an increased salty taste in fractions of Cheddar cheese (Andersen et al., 2010), and umami-tasting peptides in protein hydrolysates blocked the bitter taste receptor (Schlichtherle-Cerny and Amadò, 2002, Kim et al., 2015).

Though it is clear that the amino acid, Glu, does contribute to umami taste in various cheeses, the role of peptides remains unclear. In some cheese varieties, the umami taste has been attributed to the presence of organic acids and mineral salts (Drake et al., 2007, Andersen et al., 2010), and so it is likely that other compounds, and not peptides, are primarily forming an umami taste in fermented dairy foods.

2.1.3 Kokumi peptides

The kokumi sensation is associated with "mouthfulness", "continuity", and "thickness" and serves to complement other taste attributes. The first compound to be described as kokumi-active is the tripeptide (γ-L-glutamyl-L-cysteinylglycine - glutathione) (Ueda et al., 1997). The calcium-sensing receptor (CaSR) has been shown to play a role in the detection of the kokumi sensation (Ohsu et al., 2010). On their own, kokumi-active compounds are incapable of exhibiting taste-activity. In a water solution, kokumi-active peptides are tasteless or display other taste qualities such as sour, bitter, and salty tastes (Toelstede et al., 2009, Yang et al., 2019). In turn, the basic tastes are unable to generate kokumi activity in the absence of kokumi-active compounds, and it is only in the presence of a mixture of different tastes that the kokumi activity can be detected (Yang et al., 2019).

Kokumi activity is mostly thought to be generated by small γ-L-glutamyl peptides (Appendix 2). The peptide γ-Glu-Val-Gly, in particular, has been reported to be a potent kokumi peptide and is said to enhance the sensorial quality of fermented foods (Kuroda et al., 2020). This peptide is suggested to be as much as 12.8 times more potent than glutathione (Ohsu et al., 2010). γ-Glu-Val-Gly has not been detected in the raw material of these products, indicating that it is formed during the fermentation process (Kuroda et al, 2020). Kokumi-active peptides have been reported in

various cheeses and have also been found in a variety of other foods such as fermented cocoa beans (Salger et al., 2019), yeast extracts (Liu et al., 2015), fish sauce (Kuroda et al., 2012) and soy sauce (Kuroda et al., 2013).

In Parmesan cheese, fifteen kokumi-enhancing γ-glutamyl di-peptides were recently identified and were present in high concentrations, although most of these peptides had a low dose-over-threshold (DoT) factor, which is the ratio of concentration and taste threshold (Hillmann and Hofmann, 2016). However, the threshold for γ-glutamyl peptides enhancing kokumi-activity is said to be lower, relative to other peptides (Hillmann and Hofmann, 2016). Six of the identified γ-glutamyl dipeptides were also found to be kokumi-active in Gouda cheese (Toelstede et al. 2009). A number of the kokumi-active peptides identified by Hillmann and Hofmann (2016) had previously been proposed as umami-tasting peptides by Roudot-Algaron et al. (1994) in Comté cheese. Toelstede and Hofmann (2009) investigated a variety of cheeses for γ-glutamyl peptides, including blue, Gouda, Camembert, goat's cheese, and others. Of these, the blue cheese was found to have the highest concentration of γ-glutamyl peptides. In most samples, the peptide γ -Glu-Glu was the most dominant of the γ-glutamyl peptides. Interestingly, Kuroda et al. (2020) recently investigated the potent kokumi peptide γ-Glu-Val-Gly in cheeses prepared from cow's and ewe's milk, and only found the peptide present in cheeses produced from ewe's milk, which they related to the different proteins present in the raw milk. α-glutamyl peptides have also been investigated for their kokumi-activity but were not found to be kokumi-active in either Gouda (Toelstede et al., 2009) or Parmesan cheeses (Hillmann and Hofmann, 2016) and were found at concentrations below their taste threshold. A number of kokumi peptides were also reported to elicit bitter notes in aged Gouda, namely γ-Glu-Tyr, α -Glu-Tyr, and α -Glu-Trp (Toelstede et al., 2009). γ- and α-glutamyl may also play a role in masking bitterness; reconstituted extracts of cheese containing γ- and α-glutamyl peptides were said to have a less harsh bitter taste than extracts without these peptides (Toelstede et al., 2009).

2.1.4 Salty, sour, sweet peptides

Peptides are not known for having salty, sour, or sweet tastes, although there have been some peptides and related compounds reported in the literature for their association to these tastes.

Salty taste is primarily evoked by sodium ions (Briand and Salles, 2016) and is detected by the epithelial-sodium channel (ENaC) receptor (DeSimone and Lyall, 2006). There are few references to salty peptides or amino acids in fermented dairy foods, although there are some suggestions of these compounds enhancing the salty taste. Omission tests with Gouda cheese, for instance, indicated that the salty taste was significantly enhanced by the concentration of the amino acid, L-Arg (Toelstede and Hofmann, 2008a). In various other fermented foods, peptides have been reported to enhance salty taste, such as in soy and fish sauces (Schindler et al., 2011, Yamamoto et al., 2014) and aged meats (Sforza et al., 2001).

A sour taste can be found in a variety of foods and is particularly characteristic of some fermented dairy foods. This sour taste can be mostly attributed to lactic acid and is generated through fermentation (Tamime and Robinson, 1999). Transient receptor potential (TRP) channel PKD2L1 and PKLD13 were previously reported as promising candidates for sour taste transduction (Challis and Ma, 2016). More recently, Tu et al. (2018) described Otopetrin1 (OTOP1), a protein involved in gravity-sensing otoconia in the vestibular system, as a potential candidate for sour taste reception. Peptides containing at least one Glu or Asp in their sequence have been associated with sour tastes (Linden and Lorient, 1999). In various fermented and aged foods, amino acids and peptides have been suggested to play a role in sourness (Maehashi et al., 1999, Park et al., 2002, Sforza et al., 2006, Kęska and Stadnik, 2017). In cheese, the L-configuration of amino acids His, Asp, Ser and Thr have been associated with a sour taste (Fox et al., 2017). Andersen et al. (2010) reported that a high concentration of free amino acids contributed to a sour taste, as well as contributing to the overall complexity and intensity, in mature Cheddar cheese.

Natural peptides are not typically associated with sweetness (Temussi, 2012). There are several widely found artificial, calorie-free peptides, such as aspartame, which is frequently used as a potent alternative to sugar in foods and is 100-200 times sweeter, on a weight basis, than sucrose itself (Mazur et al., 1969, Briand and Salles, 2016). There are also several large sweet-tasting proteins derived from plants, for example, thaumatin, monellin and brazzein, which are several hundred times sweeter than sucrose, on a weight basis (Temussi, 2012, Briand and Salles, 2016). The TAS1R1/TAS1R3 dimer is responsible for the perception of sweet taste and has been reported to be stimulated by various amino acids, such as Gly and D-Trp (Zhang et al., 2017). Nelson et al. (2002) also reported the response of the umami taste receptor (TAS1R1/TAS1R3) to the L-configuration of sweet amino acids. The receptor was unresponsive to the D-configuration of these amino acids, as well as to other artificial and natural sweeteners. In aged pork meats and soy sauce, small peptides were correlated with a sweet taste (Kęska and Stadnik, 2017, Yamamoto et al., 2014). There have been some reports of sweet-tasting amino acids in fermented foods. In cheese, the L configuration of Glu, Lys, Leu, Val, Ala, Gly, Pro, Met, Gln, Thr and Ser were all suggested to have some sweet notes (Fox et al., 2017). Lawlor et al. (2002) investigated eight different cheese types and reported that the sweetness was positively correlated with the concentrations of free amino acids, including Thr, Ser and Pro. In Swiss cheese, Pro was also linked to sweetness (Fox et al., 2017). In Gouda cheese, the amino acids Gly, L-Met, L-Ser, L-Ala, L-Thr and L-Pro were reported to have a sweet taste (Toelstede and Hofmann, 2008a).

2.2 Mass spectrometry techniques to detect tastant compounds

Mass spectrometry (MS) and chromatographic techniques are frequently used together to separate and identify compounds in complex food matrices and can be used to perform semi-quantitation on the compounds of interest (Desfontaine et al., 2018). Chromatographic methods, such as high-performance liquid chromatography (HPLC), are commonly used for the separation of a wide variety of different compounds and are used for the analysis of various non-volatile and volatile compounds in fermented foods for their contribution to taste and flavour. Its ability to separate low

molecular weight compounds make it an attractive tool (Manso et al., 2005), although conventional HPLC systems may struggle to separate several different components in parallel (Mozzi et al., 2013). HPLC is not well-suited for the detection of small hydrophilic peptides (i.e., umami peptides) as they have been found to elute in the void volume. Hydrophilic interaction liquid chromatography (HILIC) can be performed for the analysis of such peptides (Andersen et al., 2008). Chromatographic systems coupled to a mass spectrometer can be used for the simultaneous separation and identification of sensory-active peptides and amino acids (Salles et al., 2002, Toelstede and Hofmann, 2008a, Toelstede and Hofmann, 2008b, Hillmann and Hofmann, 2016, Sebald et al., 2018, Sebald et al., 2019). Identification of compounds can be achieved through MS/MS (or tandem MS), where an initial mass spectrum is obtained, which undergoes a second MS, producing a spectrum for each precursor ion selected by the first round of MS. This can be used to identify the compound, by comparing experimental spectra to theoretical spectra, through a database search (McLafferty, 1983).

2.2.1 Preparation of food samples for MS analysis

Dairy products are a complex food matrix containing a wide variety of different compound classes, most of which are not of interest in the study of tastant peptides and can interfere or suppress compounds during MS analysis (de Hoffmann and Stroobant, 2007). As such, before mass spectrometry or chromatographic analysis, food samples must be appropriately prepared to isolate the fraction or individual compounds of interest. Fractionation may be performed based on the physicochemical and structural properties of the compound, such as molecular weight and hydrophobicity, amongst others (Martínez-Maqueda et al., 2013). Tastant peptides are contained in the water-soluble extract (WSE), and so typically, separation in fermented dairy foods will first involve the isolation of this extract, via homogenisation of the sample (in the case of cheese), followed by centrifugation and ultrafiltration to separate fat, casein and larger peptides and proteins. The ultrafiltrate can then be analysed using mass spectrometry and used for the identification of peptides (Soeryapranata et al., 2002, Toelstede et al., 2008b), although often

additional and repeated separation steps may be performed to further fractionate the sample or to isolate individual compounds. For instance, where Toelstede et al. (2008a) performed sensomics mapping of bitter-tasting peptides, utilising gel permeation chromatography, solid-phase extraction, and HPLC to obtain fractions of cheese samples which then underwent additional HPLC and HILIC steps to isolate individual compounds. This was in parallel with the evaluation of different fractions for taste activity via sensory testing. Liquid Chromatography-Time of Flight (LC-TOF-MS) was then performed to identify the peptides. The identified peptides can be compared against databases of known tastant peptides to infer taste activity. This approach is referred to as sensory-guided fractionation, whereby a sample is repeatedly fractionated and evaluated by a sensory panel to determine the most taste-active fraction. The most intensely tasting fraction, as determined by the panel, may then be searched for known taste-active peptides. These peptides can be isolated and further assessed by the taste panel to determine their taste activity, although synthetic peptides are often used for evaluation by the panel (Toelstede et al., 2008a). Where a specific peptide is targeted, MS analysis can be performed directly on the WSE of the sample, as was the case with quantification of the bitter peptide, β-CN 193-209 in cheese, where the WSE was analysed directly using MALDI-TOF MS (Soeryapranata et al., 2002).

The mass spectrometry techniques used will depend on the objective of the study and the sensory analysis being performed. When being assessed for taste activity via sensory evaluation, either the whole sample, fractions of interest only (i.e., a fraction that contains the greatest number of known tastant peptides), or individual peptides of interest may be assessed. In the case of evaluation of specific fractions or peptides, additional separation and purification steps need to be carried out as mentioned above, and it is necessary to bear in mind the food grade standard of any chemicals used. Martínez-Maqueda et al. (2013) have reviewed the various steps employed by researchers to extract and fractionate peptides and proteins in a variety of foods. A typical workflow for preparing and analysing samples using MS is summarised in Figure 2.1.

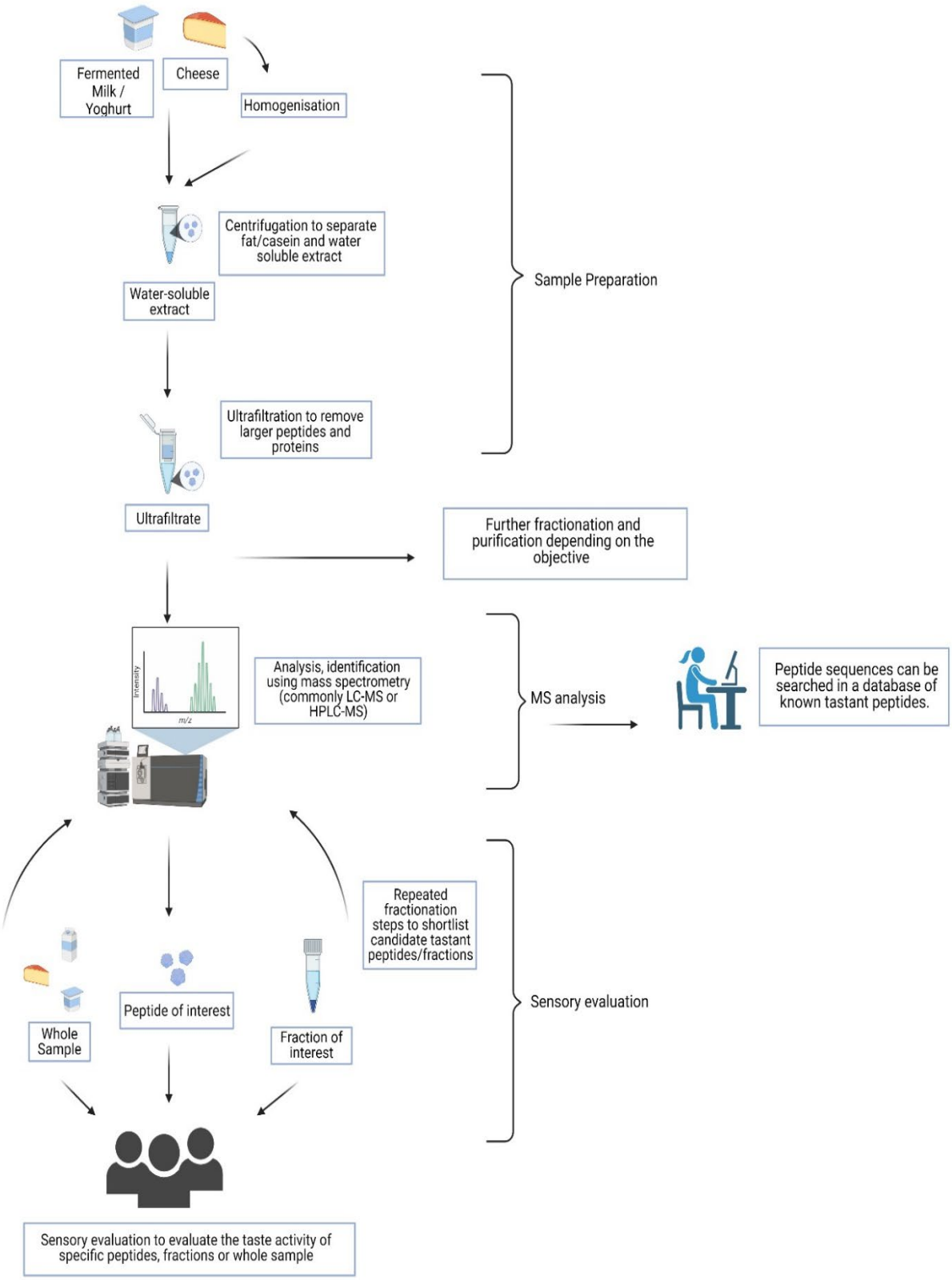


Figure 2.1 Typical workflow for determining taste-activity in samples using mass spectrometry and sensory analysis. This figure was created using BioRender.

2.2.2 Rapid techniques for obtaining chemical fingerprints

Fingerprinting produces a unique pattern that corresponds to the chemical components in the sample (Huang et al., 2019). In fingerprinting studies, identification of the compound is not a requirement, and the analysis can be based on MS1 only (i.e., the unique pattern). Fingerprinting technologies have a wide range of applications that can provide invaluable and rapid insights for the food industry.

2.2.2.1 Matrix-assisted laser desorption/ionisation – time-of-flight

Matrix-assisted laser desorption/ionisation (MALDI) is a popular method for direct mass spectrometry analysis, and it is routinely used for the analysis of milk-derived compounds. MALDI MS has been used for untargeted fingerprinting in dairy samples to monitor milk during storage (Dalabasmaz et al., 2019), to determine the authenticity of cheese (Rocchetti et al., 2018) and yoghurt (Liu et al., 2010), for the discrimination of different types of milk and milk powders (Garcia et al., 2012, Calvano et al., 2013a, Sassi et al., 2015, England et al., 2020, Piras et al., 2021), and to demonstrate the effects of fermentation on the peptide profile of kefir (Amorim et al., 2019).

MALDI MS has become a widely used technique in the field of proteomics since it was first introduced by Karas and Hillencamp in the 1980s (Karas and Hillenkamp, 1988), due to its good mass accuracy, good resolution, and its rapid and high-throughput capability (Kempka, 2005). It has become a popular choice for the analysis of a broad range of analytes, from small to large, non-volatile, and thermally labile molecules, including proteins, oligonucleotides, synthetic polymers, and inorganic compounds, amongst others (de Hoffmann and Stroobant, 2007). MALDI is achieved in two steps: firstly, the compound of interest, the analyte, is mixed with the matrix. The mixture of analyte and matrix is then deposited and co-crystallised on an electrically conducting target, e.g., a stainless-steel plate (Kempka, 2005). The target plate is then placed in a vacuum, inside the source of the mass spectrometer. The analyte and matrix solution is ablated by short bursts of intense laser pulses (de Hoffmann and Stroobant, 2007). The matrix will absorb the photon energy and rapid heating occurs. The analyte enters the gas phase, and the analyte is ionised (Kempka, 2005). When

Time of Flight (TOF) is utilised as the mass analyser, the technique is referred to as MALDI-TOF. Other mass analysers include orbitrap, ion traps, and ion cyclotron resonance, amongst others (de Hoffmann and Stroobant*,* 2007*)*. These ions are accelerated and led through a flight path, at the end of which is an ion detector that records the time-of-flight and the intensity of the individual ions that meet the ion detector. The time-of-flight for each ion is converted to a mass-to-charge ratio *(m/z)* (Hosseini and Martinez-Chapa, 2016).

2.2.2.2 MALDI-TOF sample preparation

MALDI-TOF MS can be used to directly analyse samples and does not require extensive sample preparation, making it a more rapid technique compared to GC or LC-MS. MALDI-TOF still requires some sample preparation, however, in order to mix the matrix and analyte.

The matrix selection and optimisation of the sample preparation protocol are the most important steps in the analysis and can determine the quality of the results (de Hoffmann and Stroobant, 2007). The selection of the matrix is dependent on the class and size of the analyte, e.g., lipid, peptide, or protein. The matrix should meet a number of criteria – it should have strong absorbance at the laser wavelength, a low enough mass to avoid interference (in the case of larger molecules), vacuum stability, ability to promote analyte ionisation, solubility in a solvent compatible with the analyte and an absence of chemical reactivity (de Hoffmann and Stroobant, 2007). A number of sample deposition methods have been described. The dried-droplet is the most widely used procedure. It is a simple procedure requiring the matrix and analyte to be pre-mixed, and a droplet of this mixture to be applied to the target (de Hoffmann and Stroobant, 2007). A variety of other methods have been reported, such as the double layer, thin layer, and sandwich method. The choice of layering technique may depend on the choice of matrix and the analyte.

There are several established and routinely used methods for the analysis of peptides and proteins using MALDI MS (Bruker-Daltonics, 2015). However, the quality of the results from MALDI-TOF MS may vary considerably depending on the choice of sample preparation and acquisition

mode (Dave et al., 2011). A number of factors can influence the absolute ion intensity between different shots within a sample, these include the sample crystal size and homogeneity, the analyte distribution within the crystals and the thickness of the crystal layer (Parker et al., 2008). "Hot-spots", which is a region of high sample density resulting from uneven co-crystallisation of the matrix and analyte, may be introduced during the sample preparation steps (Duncan et al., 2008). Quantitation can be difficult to achieve when hot-spots are present; the spots may contain high levels of the sample, causing ion suppression, while other areas of the spot may not contain enough sample to produce a good signal and spectrum (O'Rourke et al., 2018). To improve upon the reproducibility of MALDI-TOF data, technical and experimental replicates are used, and averaging is performed across replicates which can increase the confidence in the profiles for comparison (Lancashire et al., 2009).

Automating data acquisition can also reduce the bias toward collecting data from "hot-spot" regions, which would likely occur during manual acquisition (Duncan et al., 2008). Automated acquisition also enables more high-throughput data acquisition. Spectra acquired manually have been shown to have a higher reproducibility (Zhang et al., 2014); however, blind, and randomised data acquisition is an important aspect of the experimental design in proteomic mass spectrometry experiments (Hu et al., 2005a), and so it is important to assess the performance of the protocols when data is acquired automatically.

Smaller compounds (m/z < 700) are difficult to detect using MALDI-TOF, largely due to the interference of the matrix, which typically has a low mass (Kang et al., 2001). Several matrices for the analysis of low molecular weight compounds using MALDI-TOF have been reported in the literature. High molecular weight matrices, for example, can minimise the interference in the low mass region, e.g., 2,3,4,5-tetrakis(3′,4′-dihydroxylphenyl)thiophene (DHPT) which has been used for the analysis of various small compounds (Chen et al., 2012). Ionic liquids, which are used in combination with conventional matrices, can enhance the signal-to-noise ratio and have been reported to produce fewer interfering ions in the low mass region (Bronzel et al., 2017). The

performance of the ionic liquid matrix, however, requires the correct pairing of base with matrix, and several combinations have been reported in the literature (Li and Gross, 2004, Zabet-Moghaddam et al., 2004, Abdelhamid et al., 2014). A range of carbon-based materials have been reported for use as MALDI MS matrices in recent years, including graphite particles (Zumbühl et al., 1998), carbon nanotubes (Jing et al., 2005, Hu et al., 2005b, Pan et al., 2005, Wang et al., 2007), various forms of graphene (Dong et al., 2010, Liu et al., 2011, Liu et al., 2012), and pencil lead (Black et al., 2006, Langley et al., 2007).

2.2.2.3 Rapid Evaporative Ionisation Mass Spectrometry

Rapid Evaporative Ionisation Mass Spectrometry (REIMS) is an emerging technique that has been demonstrated to successfully fingerprint metabolites, including lipids, in tissues and micro-organisms (Balog et al., 2016, Phelps et al., 2018, Shen et al., 2020, Barlow et al., 2021). REIMS can rapidly characterise samples in near real-time using a tool for volatilisation. For example, a handheld surgical knife (an iKnife) can create a metabolite-rich vapour by point heating the surface of a sample, allowing for direct analysis of intact samples, and eliminating the need for intensive sample preparation that is typically performed in MS analysis (Barlow et al., 2021). A Venturi pump pulls the volatilised metabolites into the mass spectrometer, resulting in the detection of a fingerprint of the metabolites (Balog et al., 2010). REIMS was initially developed as a tool to detect cancerous tissues (Balog et al., 2010, Phelps et al., 2018), and has recently been applied in various food-related studies to authenticate meat and fish species (Balog et al., 2016, Ross et al., 2020, Shen et al., 2020), to characterise fruits (Arena et al., 2020), and to detect fraud in honey samples (Wang et al., 2019). REIMS is a powerful and rapid technique that has shown promise in the food industry in recent years. The greatest advantage to REIMS is that it requires no sample preparation, and a fingerprint can be obtained in near real-time. Milk and fermented milk contain a variety of low molecular weight compounds, such as amino acids, small peptides, lipids, and more than 100 different volatile compounds (Cheng, 2010). REIMS can detect a wide range of small molecules and lipids, with polar compounds, such as small phenolic compounds, fatty acids and phospholipids

readily detected in negative mode, and non-polar compounds, (e.g., acylglycerols) and some phenol derivatives readily detected in positive mode (Arena et al., 2020). The ionisation mode used for routine measurement varies depending on the matrix, with negative mode favoured for meat measurements (Balog et al., 2016) and honey measurements (Wang et al., 2019). Tentative compound identification is possible with REIMS based on high-resolution mass, though current instrumental configuration prevents automated MS/MS fragmentation to aid identification and identification of compounds ideally also requires verification using other analytical techniques (Ross et al., 2020).

2.2.3 Data processing

In order to extract meaningful information from mass spectrometry data, it must first be carefully processed. MS data is characterised by high-dimensionality, whereby the number of samples is significantly lower than the number of data points (Plechawska-Wojcik, 2012). As dimensionality increases, it becomes increasingly more difficult to identify the important data points amongst the noisy or irrelevant data, which can be problematic when performing data analysis and modelling (Lancashire et al., 2009). High-dimensional data should be appropriately processed (Plechawska-Wojcik, 2012), and the use of feature extraction or selection may be performed to reduce the dimensionality of the data (Lancashire et al., 2009). MS data suffers from poor reproducibility, resulting from variability in sample preparation, such as ‘hot-spots’, as well as instrumental variability, amongst others. Issues with reproducibility, alongside high-dimensionality, can cause further issues by making the important features within the data even sparser with respect to the level of noise. Random and noisy features may be introduced to the data through the instrument settings, sample preparation, chemical noise, instrument temperature, sample runs, and storage (Lancashire et al., 2009, Tong et al., 2011). As a consequence, the pre-processing of data is especially important before conducting an analysis of the data to remove these redundant features and to extract meaningful information from the dataset.

Typical processing steps may include some or all of the following:

- Transformation to simplify visualisation and to reduce the potential dependency of the variance from the mean.
- Intensity smoothing to suppress high-frequency noise.
- Baseline removal to eliminate background noise, i.e., chemical noise.
- Equalising the intensities across multiple spectra.
- Calibration to align spectra corresponding to different technical repeats.
- Peak detection to reduce the number of features and identify true signals in the data.
- Binning to make similar peak mass values identical.
- Peak filtering to removal false positive peaks.
- Feature matrix generated from the list of mass peaks and intensities.

The generated feature matrix can then be used for downstream analysis. There are several well-maintained packages that can be implemented using the statistical software, R, to process MALDI-TOF MS data. There are a number of open-source software programs that can also be used, for instance Mass-up (López-Fernández et al., 2015) and MASSyPup (Winkler, 2014). REIMS data can be processed using ProGenesis software (Waters Corp. Wilmslow, UK), which is supplied by the instrument vendor.

2.3 Sensory techniques to evaluate tastant compounds

Through identification and database searching, it may be inferred that a peptide has tastant properties. Using sensory evaluation techniques, these properties can be explored and verified. Sensory evaluation is used to "measure, analyse and interpret" the sensory properties of food (Anonymous, 1975). Sensory evaluation can be carried out using either trained or naïve, consumer panellists, the choice of which very much depends on the type of sensory test being performed, the objective of the test, as well the resources and budget available. Typically, training a panel will involve screening prospective panellists, the most discriminating of which are then selected and trained (Kemp et al., 2009). This can be an expensive and time-consuming process. Where naïve

consumers are used, larger sample sizes are needed due to the variability of taste preferences and sensitivity between individuals (Civille and Oftedal, 2012).

When evaluating the taste contribution of peptides in fermented dairy foods, several techniques are commonly used (Table 1). These can be categorised as either discriminative or descriptive tests. A discriminative test is used to determine if there is a perceivable difference or similarity between two or more samples. This can be attribute-specific, or simply that a difference or similarity exists between the samples. Inferring similarity between products is more challenging than inferring a difference and requires a larger number of panellists for statistical robustness (Kemp et al., 2009). Discriminative tests do not quantify the difference or similarity between the tested products, but simply indicate whether a perceivable difference or similarity does exist. Discriminative tests exist in many forms and have been utilised in a variety of studies aimed at establishing the role of peptides in the taste of fermented dairy foods. Commonly used techniques include the triangle test, duo-trio, and multiple paired comparisons (alternative forced-choice). In a triangle test, panellists are presented with three samples, two of which are identical and one which is different. The panellists are asked to select the 'odd-one-out'. In a paired-comparison, or alternative-forced choice, test, panellists evaluate all samples in all possible pairings, and for each pairing are asked to select the sample with the greatest intensity of a specified attribute (Kemp et al., 2009). Tests such as these can provide insights into the role of whole tastant groups and individual taste compounds within the samples, for instance where Toelstede and Hofmann (2008a; b) performed omission tests, presented to panellists as a triangle test, to determine the taste recognition threshold concentration of bitter peptides and fractions of cheese samples. Discriminative tests can be used alongside analytical techniques to determine taste activity in fermented dairy samples. As previously discussed, this is referred to as sensory-guided fractionation, a systematic approach to identify potential tastant compounds, where fractionated samples are assessed for taste activity, thus guiding the researchers to search for tastant peptides only in the most taste-intensive fractions (Toelstede et al., 2008a, Murray et al., 2018, Sebald et al., 2018; 2019, Alim et al., 2020).

Table 1 Overview of sensory tests performed to identify taste activity in samples originating from fermented dairy foods

Class of Test	Sensory Test	Sample	Description of Test	Reference
Discriminative	Triangle test	Tastant groups; individual peptides	Omission tests performed by omitting whole tastant groups or individual compounds. Samples presented to panellists as a triangle test to determine compounds or tastant groups that are different in taste to control stimulus.	Toelstede and Hofmann, 2008b; Toelstede et al., 2009; Hillmann & Hoffmann, 2016
		Individual peptides	Individual peptides presented to panellists as a triangle test to determine the taste recognition threshold concentration or to perform taste dilution analysis.	Toelstede and Hofmann, 2008a; Toelstede et al., 2009; Hillmann & Hoffmann, 2016; Alim et al. 2020
	Alternative-forced choice	Fractionated samples	Taste dilution analysis performed on fractionated samples by presenting panellists with pairs of samples, as a 2-AFC test, to determine at which dilution step there was a perceivable difference between the dilution sample and the blank stimulus.	Sebald et al., 2018
		Individual peptides	Taste recognition threshold concentration performed on purified synthetic peptides. Samples presented as a 3-AFC using incremental concentrations of the stimulus.	Sebald et al., 2018; Sebald et al., 2019
Descriptive	Attribute rating test	Whole sample; fractionated samples	Peptide fractions or whole samples presented to panellists and rated for their bitterness intensity using a scale (i.e., rated using descriptor terms, or 9-15-point scale)	Lee et al., 1996; Soeryapranata et al., 2002; Broadbent et al., 2002; Singh et al., 2005; Schäfer et al., 2019
	Taste profile analysis	Whole sample; fractionated samples	Taste profile analysis performed by presenting peptide fractions or whole samples and rating the intensity of each taste sensation. Texture and flavour descriptors also used (rated using 5-9-point scales).	Salles et al., 2000; Toelstede and Hofmann, 2008a; Toelstede et al., 2009; Hillmann & Hoffmann, 2016; Sebald et al., 2018; Alim et al., 2020; Xia et al., 2020

Descriptive analysis is another class of sensory evaluation, which typically uses trained individuals to identify the characteristics of a food. Descriptive techniques can be used to quantitatively measure the intensity of specified characteristics and include tests such as profiling and quantitative descriptive analysis. In recent years, rapid techniques using naïve panellists have been introduced and include sorting and projective mapping. Though descriptive tests can be more complex than discriminative tests, they can provide a greater depth of information about the characteristics of food that cannot be obtained solely through discriminative analysis (Kemp et al., 2009). In investigations of tastant peptides, commonly used descriptive techniques include taste profile analysis and attribute rating. The data from these descriptive tests may be correlated with instrumental analyses to gain insights into the physical and chemical composition of the food (Kemp et al., 2009). Studies such as these can reveal insights into certain taste or flavour characteristics that are appealing or unappealing to consumers, for instance where off-flavours and bitterness can decrease consumer acceptability.

Sensory analysis can be used in product development and to monitor the quality of products. Xia et al. (2020) demonstrated this recently by using a trained panel to measure the intensity of 14 cheese descriptors to determine the suitability of using alternative starter cultures in blue cheese production. By measuring the sensory properties, alongside characteristics of proteolysis, lipolysis, and texture, the researchers could determine the impact of using new starter cultures on these properties, compared to cheese prepared from conventional starter cultures. Hillman and Hoffman (2016) used a combination of discriminative and descriptive tests to identify individual tastant peptides in Parmesan cheese samples, as well as the role of compounds in reconstituted mixtures. This can enable a more targeted approach in developing new cheeses with desirable sensory properties. Recently, Alim et al. (2020) performed taste profile analysis and taste dilution analysis to identify bitter peptides in milk with hydrolysed milk proteins. By doing so, they could determine the level of hydrolysis that generated the most bitter peptides which would enable them to control the processing conditions to generate a product with optimal flavour.

2.4 Predictive modelling overview

Predictive modelling techniques are implemented to predict future outcomes as well as to provide insights on the association between explanatory variables and the response variable of interest (Wasserman, 2004). When the outcome is a numerical variable, regression techniques can be applied, when the outcome is categorical, classification techniques are used to generate the predictive model. Many of these techniques are interchangeable to solve either regression or classification problems.

2.4.1 Regression techniques

A simple linear regression models the relationship between a continuous response variable, y, and a single explanatory variable, x. Multiple linear regression builds on this by allowing additional explanatory variables. Linear regression assumes that there is a linear relationship between the outcome and explanatory variables, that the residuals are homoscedastic (constant variance), and normally distributed, and that the errors and observations are independent (Wasserman, 2004). A linear regression model assumes a straight-line relationship between the response and the predictor variables. If the true relationship is far from linear, then almost all of the conclusions derived from the fit are questionable, as well as potentially reducing the prediction accuracy significantly (James et al., 2013).

An important aspect of model-building is assessing the model's performance. A violation of the model assumptions does not invalidate the model; the limitations of the model can be acknowledged, or the interpretation of the model can be adjusted. Transformation of the data may also be performed which can mitigate the effects of the violations. For instance, squaring one of the terms to create a quadratic model may improve the model fit. However, there are drawbacks to higher-order polynomial terms as they can have a large influence on extreme values of the predictor variable (James et al., 2013, Kuhn and Johnson, 2013).

There are several criteria that can be used to assess the performance of a regression model, and it is recommended to use more than one to understand the model's performance and fit (Wasserman, 2004, James et al., 2013, Kuhn and Johnson, 2013). Visualising the model fit, in addition to visualising the residual plots can provide a good basis for assessing the performance. The residuals are defined as the observed value minus the predicted value. The mean squared error (MSE) can be calculated by squaring the residuals, summing them, and dividing by the number of observations. The root mean square error (RMSE) is the square root of MSE and is a metric that is frequently used to measure the model performance. RMSE provides a measure of the distance between the observed values and the predicted values. Because RMSE is the square root of MSE, the unit is the same as the original data, allowing it to be easily interpreted. R^2 is another popular metric to evaluate model performance and is defined as the proportion of variation explained by a component (Blasco et al., 2015). This provides a measure of the amount of variance explained by the model (Wasserman, 2004, James et al., 2013, Kuhn and Johnson, 2013). R^2 will usually return a number between 0 and 1, where a value closer to 1 indicates a model is explaining a good portion of the variation in the data. A value closer to 0 indicates a poor model fit and is closer to a horizontal line, indicating no relationship between predictor and response variable. R^2 also accounts for how scattered the data is (James et al., 2013).

When fitting a linear regression model, other potential problems may occur such as collinearity, whereby two or more predictors are highly correlated with one another, outliers, which are values that are far from the predicted value, and high-leverage data points, where a point has an unusual or extreme predictor value (James et al., 2013). Collinearity can be problematic in determining the effects of related predictor variables on the response variable, can reduce the accuracy of the estimates of the model coefficients, and can cause an increase in the standard error. Visualising correlation plots or calculating the variable inflation factor can be used to assess cases of (multi)collinearity. Outliers can have implications for the interpretation of the fit. Plotting of the studentised residuals can reveal outliers; typically, values that are not between -3 and 3, can be

considered as possible outliers. Outliers may be introduced erroneously through data collection or recording, but may also reflect model deficiencies, e.g., missing predictor variables, and so caution should be taken in removing such points (James et al., 2013). Studentised residuals are deleted residuals that have been standardised. High-leverage can be problematic where the least squares line is influenced by a small number of observations. A leverage statistic can be calculated; large values suggest high-leverage points (James et al., 2013).

Parametric methods are typically easy to fit and interpret; however, they make strong assumptions about the data: data should be normally distributed, have approximately equal variance, and should not contain extreme outliers. If the data do not meet these criteria, then a parametric method may perform poorly, i.e., when the true relationship of the predictor and response variable is far from linear, then the model will be a poor fit of the data and the conclusions drawn from the model cannot be relied on (James et al., 2013). In practice, a true linear relationship rarely exists. Non-parametric techniques can provide a more flexible approach.

2.4.1.1 Generalised additive models

Generalised additive models (GAM) are a non-parametric extension of the generalised linear model (GLM) (Zaniewski et al., 2002). In GLM, the response variable can follow a non-normal distribution, e.g., binomial, Poisson, or Gamma. A link function is used to fit the linear model in GLM (Ravindra et al., 2019). An advantage to GAMs is that they make no assumptions about the distribution of the data and can fit nonlinear relationships without the need to trial different data transformations (Ravindra et al., 2019). The model captures nonlinear relationships by fitting smooth functions, rather than linear or quadratic functions, which can give more flexible regression lines (Zaniewski et al., 2002). Similar to the linear regression model, GAM can be evaluated through inspection of the model residuals. The key to a good-fitting GAM is a trade-off between likelihood and the model's complexity. The smoothing parameter, lambda, controls the balance between the complexity and likelihood; a high lambda that smooths too much can effectively create a straight line, with no curvature, whereas a lambda that is too low can lead to fitting noise rather

than fitting the actual trend of the data (Ross, 2019). The optimal smooth should balance over-and-under-fitting (Pedersen et al., 2019). Over-fitting is the concept of a model fitting too closely with the training dataset and learning noise in the data, whereas under-fitting fails to learn the training data or make adequate predictions on new data. This can impact model accuracy and the model may fail to generalise over unseen data (Kuhn and Johnson, 2013).

A GAM model can be evaluated by inspecting plots of residuals. Additional methods for assessing the suitability of a GAM model include inspecting concurvity, which is essentially a nonparametric extension of collinearity (Pedersen et al., 2019). GAM has been used successfully for a variety of applications and is particularly popular in ecological studies (Zaniewski et al., 2002, Ravindra et al., 2019). GAM has also been successfully applied to predict peptide retention times from chromatographic data (Stanstrup et al., 2015).

2.4.2 Machine learning

2.4.2.1 Gradient boosting machine

Gradient boosting machine (GBM) is a popular machine learning technique, which sequentially learns new models. At each iteration a new model is trained with respect to the error of the previous model, generating a more accurate estimate of the response variable (Natekin and Knoll, 2013). The GBM algorithm is implemented in three steps: firstly, the loss function is optimised, next the weaker learner is predicted, and finally, an adaptive model is generated by adding trees to the weaker learner, reducing the loss function (Sandhu and Batth, 2020). The loss function used will depend on the response data, i.e., for categorical data (classification problems) a binomial or adaboost loss function may be used (Natekin and Knoll, 2013). GBM is reported to have improved performance for prediction of retention times from HPLC, compared to other machine learning techniques such as neural networks, random forests, and adaptive boosting amongst others for regression problems (Bouwmeester et al., 2019). In proteomics research, GBM models have been used effectively for estimating the quality score of peptide feature matches based on MS1 spectra (Ivanov et al., 2020).

2.4.2.2 eXtreme gradient boosting

eXtreme gradient boosting (XGBoost) has become a popular technique for a variety of data science problems and is reported to outperform other techniques. XGBoost is implemented under the gradient boosting framework, combining many weak classifiers to form a strong classifier (Chen and Guestrin, 2016). XGBoost is known to perform well on small-medium sized datasets (Margulis et al., 2021), but is robust to scaling, and can be extended to handle very large datasets, without compromising performance or computational speed (Chen and Guestrin, 2016). Several parameters can be adjusted in an XGBoost to improve the model performance. Maximum tree depth and nrounds (similar to number of trees) have both been used as criteria to prevent model complexity and over-fitting (Nguyen et al., 2019). Parameter tuning can be introduced for complex or large datasets, but default parameters have been found to frequently produce good results. Max depth can cause over-fitting if the depth is too great; the larger the depth, the more complex the model is and thus there is a risk of over-fitting. Larger datasets require higher depth to learn the data. One approach to tune parameters is to fix certain settings, and iteratively adjust others until optimal parameters are found through improved performance.

Recently, XGBoost has been implemented to predict antihypertensive properties in food-derived peptides (Wang et al., 2020), to predict milk source and quality using electronic nose data (Mu et al., 2020) and to predict the bitterness intensity of molecules (Margulis et al., 2021), amongst many others. In all applications, XGBoost performed particularly well.

2.4.2.3 Random forest

Random forest is a bagging type ensemble learning algorithm, composed of many decision trees. Random forest models are built by extracting m data points from the training data to form a new training subset. A classification decision tree or regression model is built for each training subset, by randomly selecting k features among all features as split nodes. The output is a category, for classification models, with the highest number of votes of each decision tree (Mu et al. 2020). Parameters such as number of trees, ntree, and mtry can be adjusted to prevent complexity and

computational time (Nguyen et al., 2019). Random forests have been used for identifying milk source and quality (Mu et al., 2020) amongst others. Random forest has been applied to various research predicting bioactive peptides, although it has been criticised for dealing poorly with data imbalances and peptide characteristics with high-dimensionality (Wang et al., 2020).

2.4.2.4 Deep learning

Deep learning is a machine learning algorithm that utilises the concept of neural networks in the human brain to learn from data. The model is built using deep layers that allow continuous learning to improve the model performance. A network model, a loss function and an optimisation method are used as the building blocks of the model. The model is constructed by mapping the input data to predictions, via the network model which connects multiple layers. Next, a loss function is calculated which evaluates the performance of the network by comparing the model predictions with the target. The optimisation methods then use the loss value to optimise the model. This is achieved through backpropagation, a procedure which uses the loss value as a feedback signal which sequentially adjusts the weights of the network connection. Deep neural networks (DNNs), convolutional neural networks (CNN), feedforward neural networks (FNN) or recurrent neural networks (RNN) are different network architectures that can be implemented. These vary in the number of neurons in each layer, the number of layers and the type of connection between the layers. The network in a DNN, for instance, is made up of an input layer, multiple hidden layers, and an output layer (Wen et al., 2020). Deep learning has been applied to solve various problems in proteomics, for instance, to predict the retention time for peptides using LC-MS data (Wen et al., 2020, Ma et al., 2018), to predict MS/MS spectra (Lin et al., 2019, Zhou et al., 2017) and to predict de novo peptide sequences based on MS/MS spectra (Tran et al., 2017). Deep learning techniques have also been applied to MALDI-TOF data to identify bacterial species and genera (using convolutional network: Zielinski et al. (2017), Papagiannopoulou et al. (2020)).

2.4.2.5 C4.5 Decision tree

Decision trees are frequently used for classification model building. A decision tree is a directed tree comprised of a root node, with no incoming edges, and decision nodes, which are all the remaining nodes that each have an incoming edge. To optimise the model performance, the internal nodes partition the instance space into two or more parts during training. Each path from the root node to leaf node then forms a decision rule to classify each new instance (Dai and Ji, 2014). C4.5 is an extension of the ID3 algorithm. It uses normalised information gain (i.e., the difference in entropy (i.e., the expected information needed to classify a record in a given node)) to select splitting attributes. This considers the number of outcomes that a specific attribute might produce. The attribute with the greatest normalised information gain is selected to make the decision on the data split (Quinlan, 1993, Hssina et al., 2014). C4.5 also uses tree-pruning methods, which work well for noisy datasets and minimises the risk of over-fitting (Hssina et al., 2014).

2.4.2.6 Evaluating classification model performance

There are several approaches to evaluate the performance of a model implemented in classification problems, but as with regression techniques, it is preferable to assess multiple metrics before deciding on a model. The appropriate metric can vary depending on the data, and the objective of the test. In binary classifications, accuracy and F-measure are among the most popular metrics used to evaluate performance. However, these are not without their faults. Accuracy is reported as the number of correct predictions as a ratio of all predictions made. Accuracy can be very misleading in cases where there are imbalances in the data and should not be used as the sole criteria for assessing model performance (Gay et al., 2002). The F-measure combines two other metrics, precision, and recall, where precision is the accuracy of the positive class and recall is the number of true positives divided by the number of all possible positive samples (Gay et al., 2002). F-measures have been criticised for their over-optimistic score, particularly for imbalanced data (Chicco and Jurman, 2020). Matthew's correlation coefficient (MCC) is a metric that is considered to be more robust and informative for binary classification than the F-measure, and also accounts

for imbalances in the data. MCC returns a value between -1 and 1, where -1 indicates a perfect misclassification, and +1 indicates a perfect correct classification. It returns a high score only if the majority of instances of both classes are correctly predicted. Recently, Chicco and Jurman, (2020) proposed that assessing binary classification problems is better done using MCC to evaluate performance. The area under the curve (AUC) evaluates how well a binary classification can distinguish between true positives and false positives, where a value closer to 1 is a perfect classifier. The area under the precision-recall curve (AUCPR) is more sensitive than AUC and has been hailed as a more informative metric than the AUC, particularly for data that is imbalanced (Chicco and Jurman, 2020).

2.4.3 Exploratory techniques

In the field of proteomics and metabolomics, exploratory tools, such as principal component analysis partial least squares -discriminant analysis, are frequently used as a first step in the data analysis pipeline to visualise patterns in the sample set and can provide an indication of the separation of classes by showing clustering or overlapping of samples.

2.4.3.1 Principal component analysis

Principal component analysis (PCA) is frequently used for data reduction and as an exploratory tool. PCA is an unsupervised technique, which does not consider the response variable during modelling. PCA functions by finding linear combinations of predictors, which are the principal components (PCs), and capturing the greatest variance in the data. PCA summarises the data by essentially creating new, uncorrelated variables (Kuhn and Johnson, 2013). PCA also has the advantage of being able to handle high-dimensionality in datasets, which is a characteristic of proteomics and metabolomics studies (Lancashire et al., 2009).

2.4.3.2 Partial least squares-discriminant analysis

Partial least squares-discriminant analysis (PLS-DA) can essentially be described as a supervised version of PCA, whereby it considers the class of the response variable during modelling. Similar

to PCA, PLS finds linear combinations of the predictor, which are referred to as the components or latent variables. In PLS, the model finds components that maximally summarise the variation of the predictor, while requiring the components to have a maximum correlation with the response (Kuhn and Johnson, 2013). In the context of supervised classification, PLS-DA is performed to differentiate and discriminate between sample classes. The concept was originally applied to regression problems but was extended to classification; in this context, the sample class is converted to a dummy variable, which is predicted using a regression (Rohart et al., 2017). In metabolomic fingerprinting studies, clustering of sample classes typically depends on the use of discriminant analysis techniques, such as PLS-DA, over PCA (Wang et al., 2019, Paxton, 2020). Recently, Multivariate INTegrative PLS-DA (MINT-PLS-DA; Rohart et al., 2017) was introduced as a technique to integrate omics datasets originating from independent studies. This is a novel approach that accounts for variation originating from different sample and data collection protocols, in addition to identifying molecular signatures. A sparse MINT can also be applied which allows for variable selection by selecting the optimal number of variables to retain for each component. This reduces the influence of noisy features and can improve the predictive performance of the model (Rohart et al., 2017). The MINT-PLS-DA approach offers a solution to overcome problems that are typical of omics studies, such as poor reproducibility and consistency between sample sets (Rohart et al., 2017).

2.5 Data-driven approaches to evaluate tastant peptides

Integrating sensory, analytical and data analysis techniques can provide rapid insights into tastant peptides in food. Recently, Sebald et al. (2018) coined the term ‘sensoproteomics’, a process that uses both targeted proteomics and sensory techniques to identify taste active peptides in food. This approach involved a targeted analysis by first compiling a database of known peptides from fermented foods. Cheese samples were fractionated and then evaluated for their sensory properties. In parallel with this, the cheese fractions were analysed using targeted proteomics with proteomics software and LC-MS/MS. The combined methods allowed the researchers to shortlist 17 bitter

peptide candidates from their original database of 1600 peptides. Sebald et al. (2019) later developed an alternative approach to shortlist potential bitter peptides in cheese samples, that required no prior literature knowledge of tastant peptides. Their approach involved systematically shortlisting candidate bitter peptides by first conducting a database search of MS/MS spectra obtained, further refining this list using SWATH-MS (Sequential Window Acquisition of All Theoretical Mass Spectra), followed by using targeted proteomics software. The list of candidate peptides was then further shortlisted by filtering the peptide list and selecting only peptides that were present in the most bitter-tasting fractions. Using this approach, they shortlisted their list of candidate bitter peptides from nearly 1000 to 42, which were then evaluated for their taste activity. The authors concluded that this technique holds promise for selecting candidate bitter peptides and was less time and labour-intensive than alternative methods.

The application of predictive modelling and machine learning can aid in better understanding the role of different taste compounds. Newman et al. (2014) assessed various methods of evaluating the bitterness intensity in dairy protein hydrolysates. They integrated HPLC with trained taste panels and compared this to an electronic tongue. Using a PLS (Partial Least Squares) regression, they were able to build a predictive model using the results of the electronic tongue to predict the bitterness intensity of their samples, although they did conclude that ideally the HPLC and electronic tongue data should both be used in combination to get more accurate and reliable results. A similar approach was adopted to evaluate the correlation between electronic tongue and trained panel measurements in fermented milk (Hruskar et al., 2010), using PLS-regression and artificial neural networks. Recently, Rocha et al. (2020) noted the advantages of using machine learning techniques with sensory evaluation. By using machine learning techniques, the authors could determine the sensory drivers for cheese samples prepared using different processing parameters. They found these techniques to be more rapid, efficient, and accurate for analysing sensory data compared to conventional statistical modelling techniques. In soy sauces, Yamamoto et al. (2014) correlated the taste differences in soy sauces with dipeptides. They constructed a PLS regression

model using peptidomic data obtained from both GC-MS and LC-MS, and sensory data, obtained via quantitative descriptive analysis. Using the PLS model, they found a number of dipeptides correlated with the basic tastes. Recently, Daher et al. (2020) developed a PCA model using peptidomic data and descriptive sensory data of dairy protein hydrolysates to demonstrate the similarities in instrumental and sensory data. The authors inferred the relevance of such a model in evaluating the bitterness of protein hydrolysates.

In-silico approaches have also been developed to predict the potential taste activity of peptides, many of which draw on the wealth of information in publicly available databases. Using the amino acid sequence of known porcine-derived proteins, Kęska and Stadnik (2017) compared their sequences to the BIOPEP database and identified the number of taste active peptides and amino acids in the protein sequences. This study aimed to predict the role of a protein in taste activity based on the protein's ability to produce peptides and amino acids, which were known for their taste activity. BitterDB is a curated database composed of over 1000 bitter molecules cited in the literature (Wiener et al., 2012) and has been used to build predictive tools, such as BitterIntense. This tool was developed using machine learning techniques to predict the bitter intensity of a compound, utilising available information from BitterDB. The developed models could classify molecules as either “very bitter” or “not very bitter”, with about 80% accuracy (Margulis et al., 2021). BitterPredict is a similar predictive tool, based on BitterDB, which can classify with 80% accuracy a compound as either “bitter” or “non-bitter”, based on the chemical structure of the compound (Dagan-Wiener et al., 2017).

2.6 Challenges in inferring the role of compounds in taste

Understanding the role of tastant compounds within food can be problematic, and there have been some questionable results presented. Some authors have suggested anomalies in results are due to the purity of the peptides studied. This was the case with the reported “delicious” umami peptide (Yamasaki and Maekawa, 1978), which was later suggested to have bitter peptides interfering with the results (van den Oord and van Wassenaar, 1997). Similarly, a proposed salty peptide had

artefacts of NaCl introduced during peptide synthesis and did not impart a salty taste, as previously suggested (van den Oord and van Wassenaar, 1997). Additionally, the interactions the peptide may have with other tastant compounds may affect the taste elicited by the peptide (Zhao et al., 2016). For instance, the presence or absence of 5′-ribonucleotides can affect the umami taste of peptides (Zhuang et al., 2016), and the presence of NaCl in low concentrations caused sour-tasting peptides to elicit sweet and umami tastes in fish sauce (Park et al., 2002).

Whilst it is plausible that discrepancies are due to impurities, it is also possible that the difference in sensitivity among participants during sensory evaluations could lead to varying results. This is particularly notable for studies on umami peptides. The umami taste has long been recognised in Asian cultures and was first described by Japanese researchers at the beginning of the 20th century (Ikeda, 1909). However, it was refuted by Western researchers that the umami taste existed (Behrens et al., 2011), and numerous studies have disputed the existence of umami peptides and described their taste as mostly bitter or tasteless (van den Oord and van Wassenaar, 1997, Zhang et al., 2017). One potential explanation for this could be the existence of "supertasters" and "non-tasters". A supertaster experiences an increased intensity of a taste and was reported by Bartoshuk et al. (1994) in relation to the compound 6-n-propylthiouracil (PROP), which was perceived as intensely bitter by supertasters. The concept of the supertaster is complicated, however, insofar as someone may have a heightened response to some bitter compounds but not others (Hayes and Keast, 2011). Since the discovery of the bitter supertaster, further studies have found evidence to support heightened responses to all taste modalities (Hayes and Keast, 2011). Chen et al. (2009) reported variations by individuals in response to different concentrations of umami-tasting compounds, which was related to variations in the gene, TAS1R3. Reports have found that there is a higher prevalence of PROP-supertasters and low sweet likers in Asian populations compared to Caucasians (Yang et al., 2020), which could offer an explanation for the differences in findings between Western and Asian sensory studies. Aside from genetics and ethnicity, differences in taste perception may also be affected by differences in gender and age. For instance, females are more

sensitive than males to the bitter and sweet taste (Michon et al., 2009). Monteleone et al. (2017) reported some similar findings, and also found that taste perception was significantly decreased in older cohorts.

2.7 Fermented milk

The fermentation of milk transforms it into various products with an extended shelf-life and enhanced sensory and nutritional properties (Tamime and Robinson, 1999). Fermented milk is generally produced by inoculating milk with lactic acid bacterial culture, which acidifies the milk. Fermented milk products are valued by consumers and the food industry for their perceived health and nutritional properties, pleasant taste, and texture. Demand for fermented dairy products has surged worldwide in recent years, particularly in North and South America, as well as many parts of Asia, where fermented dairy products are not traditionally eaten regularly (Rul, 2017).

2.7.1 Manufacture of fermented milk

The raw materials used in the production of fermented milk can determine the taste, texture, and quality of the final product. Fermented milk may be produced using the milk of a variety of animals, although, in industrial yoghurt-making, cow's milk is the most commonly used milk base (Chandan, 2014). The origin of the milk may affect the sensory characteristics in the final product, which is related to the protein content of the milk base, in addition to the fat content (Tamime and Robinson, 1999). For instance, in Manchego-type cheeses produced from both cow's and ewe's milk, the cheese produced from the cow's milk was found to have a stronger bitter taste that was not observed in cheese produced from ewe's milk (Fernandez-Garcia et al., 1990). Fermented milk produced from sheep and buffalo milk (which are high-fat) resulted in a much richer and creamier product, with a more pleasant "mouth-feel" than that produced from lower fat milk (e.g., horse or cow milk), or from milk that has had the fat content removed, such as skimmed milk (Tamime and Robinson, 1999).

Fermented milk is manufactured by heat-treating a milk mix, inoculating with a starter culture, and fermenting until a pH of 4.5 is reached (Nguyen et al., 2018). A final pH of 4.2-4.5 is typical in commercial fermentations; a pH ≤4.6 is the isoelectric point of casein and causes precipitation of the protein (Tamime and Robinson, 1999). In industrial manufacturing, additional processing steps such as homogenisation and the addition of stabilisers or sweeteners might take place.

The heat treatment of the milk serves to eliminate any existing micro-organisms in the milk, reducing the competition for the starter culture and ensuring the product is safe for consumption. The heat treatment also results in thermal breakdown of milk constituents, which releases compounds that are utilised for bacterial growth (Chandan, 2014). The temperature and duration of the heat treatment can impact the sensorial properties of the product; increased temperatures and treatment times can increase the denaturation of the milk proteins. During denaturation, the whey proteins interact with κ-casein on the casein micelle surface, leading to increased gel firmness and viscosity (Lucey and Singh, 1999). Lucey et al. (1998) reported that milk treated at high temperatures (greater than 80°C) resulted in a final product that had acid gels with a rough surface, cracks, reduced thickness and some whey separation. In contrast, Yadav et al. (2018) reported that fermented milk produced from milk treated at high temperatures (85°C for 30 minutes) was found to have significantly improved body, texture and overall acceptability compared to a product made at lower temperatures for the same duration (e.g., 80°C, and 75°C).

Following the heat-treatment and subsequent cooling of the milk to fermentation temperature, the starter culture is added. The mix is incubated, and typically maintained at a temperature of 43°C, as the fermentation process takes place (Chandan, 2014). The fermentation temperature is also reported to impact the taste in fermented milk; milk fermented at 44°C is reported to yield a product that is less likely to be bitter than one fermented at 38°C (Tamime and Robinson, 1999).

Starter cultures refer to a microbial composition of a large number of cells containing at least one micro-organism that are added to a raw material, such as milk, and which drive the fermentation

process to produce a fermented product (Leroy and De Vuyst, 2004). Depending on its primary function, added micro-organisms may be considered a starter or primary culture (if it contributes to the acidification process), and adjunct, maturing or secondary cultures (if it contributes to flavour, aroma, probiotic or maturing properties). The starter culture used in fermented milk production is comprised of *Streptococcus thermophilus* and *Lactobacillus* subsp. *bulgaricus*, which drive the fermentation process and are responsible for creating the flavour, body and texture that is typical of fermented milk (Tamime and Robinson, 1999). Often, other adjunct cultures will be added which do not contribute to the acidification but are added for their contribution to sensory or health-promoting properties, for example *Lactobacillus acidophilus, Bifidobacterium lactis,* and *Lactobacillus casei* which are commonly added for their probiotic potential (Routray and Mishra, 2011, Fijan, 2014). The choice of starter strain can play a role in the formation of flavour defects, such as bitterness (Lemieux and Simard, 1991). A high inoculation rate of the starter bacteria can also lead to a product of inferior quality and flavour defects, particularly *Lb. bulgaricus* which is mostly responsible for the production of bitter peptides in fermented milk products (Tamime and Robinson, 1999).

2.7.2 Characteristic taste and consumer preferences for fermented milk

Fermented milk is characterised by a distinctive sharp acidic and green apple flavour, attributed to lactic acid (Cheng, 2010). The concentration of lactic acid in the final product can determine its acceptability; either too low or too high can have a negative impact on the typical mild-flavour (Ravyts *et al*., 2012). The activity of the starter culture produces a variety of compounds that form the distinctive taste and flavour of the product (Beshkova et al., 1998). These compounds are mostly derived from fat, protein or carbohydrates present in the milk (Cheng, 2010) and may be divided into volatile and non-volatile acids, carbonyl compounds and miscellaneous compounds, such as products generated through the degradation of protein, fat, or lactose (Tamime and Robinson, 1999).

More than 100 different aroma compounds have been identified in fermented milk products, most of which are present at very low concentrations and are not known to make a significant contribution to flavour formation (Cheng 2010). The typical aroma of fermented milk is a result of acetaldehyde, imparting a fresh and fruity aroma of green apple or nuts (Chen et al., 2017, Cheng, 2010, Rul, 2017). The concentration of acetaldehyde depends on the bacterial strains and the processes used during fermentation (Cheng, 2010). Acetaldehyde can be formed either through pyruvate decarboxylation or amino acid conversion, via threonine aldolase (Chaves et al., 2002). Diketones such as 2,3- butanedione (diacetyl), acetoin, acetone, and 2–3 pentanedione are also major compounds found in fermented milk products and impart a butter-like flavour (Imhof et al. 1995, Ott et al. 1997; Rul, 2017). The balance of these flavour compounds is important in determining the sensory characteristics of dairy products and can often determine a consumer's liking of the product (Smit et al., 2005).

The acceptability for fermented milk products can vary considerably among consumers. Unsurprisingly, many studies have concluded that the preference and acceptability of fermented milk products is strongly influenced by its sweetness (Harper et al., 1991, Barnes et al., 1991a, Barnes et al., 1991b, Allgeyer et al., 2010, Bayarri et al., 2011). Consumer studies investigating the liking of six fermented milk products demonstrated that sweetness was positively correlated with consumer liking for a large proportion of the surveyed group (52%) (Bayarri et al., 2011). Fermented milk, however, is typically regarded as a healthy food, and the high levels of added sugar and the addition of artificial sugars may damage this perception (Chollet et al., 2013). Consumer studies indicated that sweet taste alone is not the sole determinant for acceptability in fermented milk products. Pohjanheimo and Sandell (2009) investigated the motivation for food choice and preferences in drinking yoghurt and concluded that consumers who generally appreciate more natural ingredients and tend to avoid additives are more inclined to eat yoghurt products that are less sweet tasting. Jaworska et al. (2005) reported that negative sensory attributes, such as off-flavours and bitterness, were of critical importance for consumer acceptance of natural fermented

milk products. Although this is not always the case; following a consumer study, Bayarri et al. (2011) identified a small subset (about 10% of 120 participants) of those surveyed whose liking for yoghurt and fermented drink products was driven by attributes such as bitterness, saltiness, acidity and astringency, and a disliking for products that was driven by sweetness.

2.7.3 Changes in peptide composition throughout fermentation of milk

Typical starter cultures, *St. thermophilus* and *Lb. bulgaricus* can ferment milk individually; however, when grown together, these bacteria are known to have a mutually beneficial interaction, referred to as 'proto-cooperation' (Tamime and Robinson, 1999). The mutual stimulation of the two bacteria is said to positively contribute to a number of desirable characteristics within the final product (Ott et al., 2000, Sodini et al., 2004, Sieuwerts et al., 2008). Because *St. thermophilus* is non-proteolytic, it depends on the supply of amino acids and peptides that are generated through the extracellular protease activity of *Lb. bulgaricus* on milk casein (Radke-Mitchell and Sandine, 1984). Briefly, the process of proto-cooperation in fermented milk is as follows: following inoculation, *St. thermophilus* grow exponentially by utilising the free amino acids, dipeptides, tripeptides, and oligopeptides in the milk (Sieuwerts et al., 2008). Initially, *Lb. bulgaricus* benefits *St. thermophilus* by releasing amino acids (e.g., valine, leucine, histidine, and methionine) from the milk which are utilised by the growing *St. thermophilus*. The high pH at this point is also optimal for the growth of *St. thermophilus.* During this exponential growth period, almost no growth is observed in *Lb. bulgaricus*. The proto-cooperation between the two cultures during this phase results in increased lactic acid and aromatic compound formation (compared to when the two cultures are individually grown) (Routray and Mishra, 2011). During the second growth phase, the concentration of essential amino acids (e.g., glutamic acid and methionine) in the milk are insufficient for the requirements of *St. thermophilus*, and the available amino acids soon become limited. *St. thermophilus* subsequently enter into a non-exponential growth phase (Juillard et al., 1995, Letort et al., 2002). *St. thermophilus* stimulates the growth of *Lb. bulgaricus* by releasing small amounts of formic acid, folic acid, and pyruvic acid (Tamime and Robinson, 1999). As the

levels of acid in the milk increase, the pH of the milk lowers and favours the growth rate of *Lb. bulgaricus*, which enters an exponential growth phase and proteases expression is initiated (Routray and Mishra, 2011). *Lb. bulgaricus* continue to grow in a third growth phase. *St. thermophilus* proceeds to enter a second growth phase. During this growth phase, *St. thermophilus* use milk casein as a source of amino acids (Letort et al., 2002).

Products generated through proteolytic processes may be subsequently utilised in other processes. Amino acids can be converted in many different ways by enzymes such as deaminases, decarboxylases, transaminases (aminotransferases), and lyases (van Kranenburg et al., 2002). Aromatic amino acids (e.g., Phe, Tyr, Trp), branched-chain amino acids (e.g., Val, Leu, Ile), and sulfuric amino acids (e.g., Cys, Met) are reported to be important substrates for flavour development (van Kranenburg et al., 2002). Most amino acids may be converted by aminotransferases to α-keto acids. From here, they may proceed to be converted to hydroxy acids, aldehydes, organic acids, alcohols, and esters (Smit et al., 2005).

The concentration of peptides is reported to increase in fermented milk products, compared to unfermented milk. Li et al., (2020) reported an increase in the content of small peptides in fermented milk products compared to unfermented products; 19 small peptides were differentially abundant in the fermented product. Amorim et al. (2019) also noted an increase in the number of ions detected pre- and post-fermentation when analysing their samples: they identified 51 *m/z* ions, via MALDI-TOF MS, that were unique to unfermented Kefir samples and 183 unique to fermented samples. The increased rate of peptides in the fermented products could be attributed to the bacterial utilisation of protein during fermentation; proteins are degraded during fermentation, which results in the accumulation of smaller products. In turn these products may be used directly for bacterial growth as nutrients or are further degraded into smaller products. The resulting products may also be used for bacterial growth, depending on the stage of bacterial growth, or are transformed into aroma or flavour compounds via enzymatic reactions as mentioned above.

2.8 Literature review conclusions and perspectives

Methods of fermentation have been used by humans for thousands of years to extend the shelf-life of food and impart favourable sensory qualities (Tamime and Robinson, 1999). Due to perceived health benefits and the positive view consumers have of many fermented products, their consumption worldwide has increased considerably since their commercialisation (Aryana and Olson, 2017). To meet such demands and to ensure the quality of these products, it is imperative to understand the role of tastant compounds in these foods.

Peptides and amino acids are widely regarded to be important tastants in a variety of fermented and aged foods, and it has been shown that they can contribute or interact with taste across all taste modalities (Park et al., 2002, Lawlor et al., 2002, Toelstede and Hofmann, 2008a, Maehashi and Huang, 2009, Andersen et al., 2010, Kohl et al., 2012, Zhuang et al., 2016, Kęska and Stadnik, 2017, Sebald et al., 2018, 2019). A great deal of the work conducted on tastants in fermented foods relates to the bitter taste. There are a few reasons for the extensive body of work conducted on the bitter taste – firstly, it has been demonstrated that the bitter taste, and the role of peptides and amino acids in generating this taste, is both complex and interesting (Ishibashi et al., 1987a, Ishibashi et al. 1987b, Ishibashi et al., 1988a, Ishibashi et al., 1998b, Tamura et al., 1990, Maehashi and Huang, 2009, Kohl et al., 2013). Secondly, bitterness in many foods is considered to be undesirable and can compromise the quality of the product, and so understanding the bitter taste and its prevention and/or masking is in the interest of food manufacturers. Further to this, many bitter peptides also possess bioactive properties. This can be problematic as efforts to debitter such peptides have been found to negatively impact the peptide's bioactive properties (Murray et al., 2018), and when bioactive peptides are added to products, such as fermented milk, they have been found to significantly increase the bitterness intensity (Ahtesh et al., 2018). While the bitter taste has been well documented in the literature, the mechanism for taste in other taste types is somewhat neglected in the literature, by comparison, particularly pertaining to peptides and fermented dairy foods.

There are many new and exciting approaches proposed to explore tastant peptides in foods. Sensory evaluation using human panels is ultimately the gold standard; however, these can be costly and time-consuming, and the advent of more data-driven and integrative approaches could prove invaluable to provide quick and less-costly insights into taste-formation and activity in a variety of foods. Such insights can enable food manufacturers to develop products with targeted, desirable characteristics, enhancing the sensory properties of the food.

2.9 Rationale of the techniques selected for use in this thesis

- As demonstrated, there are numerous mass spectrometric, data analysis and sensory analysis approaches which can be taken to investigate tastant properties in food. This thesis sought to build on this by exploring rapid techniques that can provide insights into the early stages of product development. The work carried out in this thesis is not conclusive and provides a potential preliminary approach for short-listing potential compounds which could be further investigated using some of the more in-depth approaches outlined throughout the literature review.
- Conventional mass spectrometry techniques, such as GC- and LC-MS, can provide powerful insights into the tastant properties of food. These protocols typically will include derivatisation and time-consuming separation techniques, which are not very conducive to rapid, high-throughput screening. Fingerprinting technologies may offer an alternative approach to provide quick insights, that may be particularly useful to manufacturers during the early stages of product development. This thesis wished to explore alternative techniques which may not provide as much information and insight as LC-MS but offer quick insights that would lend themselves to more to high-throughput screening.
- It is widely understood that the characteristic taste and flavour in fermented milk results from lactic acid, acetaldehyde, diacetyl, and acetoin, amongst others. Peptides, although not widely reported to contribute to taste in fermented milk, are widely reported in cheese.

As mentioned in this literature review, tastant peptides in fermented dairy foods are known to contribute directly or indirectly to taste and flavour. As such, peptides may potentially play a role in taste and flavour in fermented milk. This thesis sought to investigate if such a correlation can be made.

- There is a wide variety of sensory evaluation techniques that can be employed to characterise or measure various attributes of a food product. Multiple paired comparison tests are often performed for evaluating taste perception in various foods and can provide an initial insight into the differences that exist between products. Although they do not provide a quantitative measure of the attribute, they can provide an indication of the difference in samples which may guide further analysis.

Chapter 3 Discrimination of milk and fermented milk by MALDI-TOF MS and REIMS fingerprinting

Summary

This chapter describes rapid techniques for fingerprinting the peptides and small molecules in milk and fermented milk products with different flavour characteristics. Rapid Evaporative Ionisation Mass Spectrometry (REIMS) and Matrix-Assisted Laser Desorption Ionisation - Time-of-Flight Mass Spectrometry (MALDI-TOF MS) are rapid techniques, with high-throughput potential and require minimal sample preparation. Using these techniques, different fermented milk produced from different bacterial strains were fingerprinted. Supervised and unsupervised multivariate analysis techniques were employed to analyse the data and could distinguish the fermented milk samples at peptide and small molecular level. These techniques can aid in quick and cost-effective fingerprinting to screen different fermented milk products. This is the first study demonstrating the use of REIMS for the measurement of liquid dairy samples.

3.1 Introduction

The first objective of this thesis was to explore a rapid and effective technique to distinguish the compounds generated by different bacteria in fermented milk samples. Many metabolomic untargeted analyses use separation (e.g., liquid (LC) and gas chromatography (GC)) prior to mass spectrometry (MS) analysis to fractionate the sample, which aims to increase the number of compounds observed and improve identification via MS/MS (Moros et al., 2017, Calvano et al., 2013b, Mischak et al., 2013). Sample preparation and data acquisition via LC-MS or GC-MS are both relatively time- and labour-intensive, and as such, they are not well-suited for high-throughput use in the food processing industry for monitoring, quality control and rapid screening of samples during product development. Molecular fingerprinting is a rapid means to differentiate samples

(England et al., 2020) where the fingerprint represents the distinctive pattern of the sample's chemical components (Huang et al., 2019). Fingerprinting studies are typically based on MS1 only, where MS1 is a scan of the ions by their mass-to-charge ratio. MS/MS (or MS2) is the fragmentation of ions selected by MS1 and subsequent identification of compounds through database searching (de Hoffmann and Stroobant, 2007). Fingerprinting technologies allow for a quick analysis time between samples and are considerably quicker than using a separation method, making them more suited for high-throughput use in the food processing industry (Barlow et al., 2021, Cohen and Gusev, 2002, Huang et al., 2019).

This chapter explored if Matrix-Assisted Laser Desorption Ionisation - Time-of-Flight Mass Spectrometry (MALDI-TOF MS) and Rapid Evaporative Ionisation Mass Spectrometry (REIMS), mass spectrometry-based fingerprinting techniques requiring minimal sample preparation, could be used to distinguish fermented milk samples prepared with different starter cultures in a high-throughput manner. The molecular fingerprint obtained by these techniques can provide rapid insights into the chemical composition of different fermented products. Such insights could be used for the screening and development of fermented products and bacterial strains with targeted properties. As an exemplar, cow skim milk and two types of fermented milk produced from the skim milk were analysed. The fermented milk was produced from two commercial bacterial culture combinations, both containing *Lactobacillus delbrueckii* subsp. *bulgaricus* and *Streptococcus thermophilus*. The two commercial cultures are known for producing fermented milk with opposing flavour and viscosity; the product information sheets describe the CH1 culture as producing fermented milk with a strong flavour and low viscosity, whereas the YF-L811 culture produces fermented milk with a very mild flavour and a high viscosity (Hansen, 2011, Hansen, 2016). The performance of both unsupervised and supervised multivariate techniques to discriminate the peptide and small molecular fingerprints of fermented milk was evaluated. By coupling mass spectrometry with multivariate data analysis techniques, an efficient means to discriminate between

different fermented milk products, known to have contrasting flavour characteristics, was established.

3.2 Materials and methods

3.2.1 Samples and reagents

MALDI matrices and peptide standards were purchased from Bruker Daltonics (Bremen, Germany). Acetonitrile and ethanol were obtained from Fluka Analytical (Merck, Kenilworth, NJ, USA) and were LC-MS grade. Isopropanol, trifluoroacetic acid (TFA) and water were purchased from Fisher Scientific (Hampton, NH, USA), and were all LC-MS grade. Graphene oxide was purchased from GRAPHENEA (San Sebastián, Spain). Starter cultures were purchased from Christian Hansen (CHR Hansen A/S, Denmark). Skim milk powder was purchased from Fonterra (Co-operative Group, New Zealand).

3.2.2 Fermented milk preparation

Two types of fermented milk were prepared using commercial starter cultures (CH-1 and YF-L811, CHR Hansen A/S, Denmark). The fermented milk used for the following experiments were prepared in collaboration with the Fermented Foods programme, at AgResearch, Palmerston North.

Each culture contained different strains of *Lactobacillus delbrueckii* subsp. *bulgaricu*s and *Streptococcus thermophilus.* Briefly, the samples were prepared as follows: milk was prepared by reconstituting skim milk powder (Fonterra Co-operative Group, New Zealand) at 9.1% (w/v) and heat-treating at 85°C for 30 minutes in a Grant OLS Aqua Pro Shaking Water Bath (Grant Instruments Cambridge Ltd., Shepreth, UK). The milk was allowed to cool to incubation temperature (43°C) and inoculated with the starter cultures CH-1 and YF-L811. The milk was inoculated at a rate of 110 mg/L of bacteria/skim milk. The pH of the milk was monitored using a pH meter (Hanna Instruments, USA), calibrated using buffers of 4.01 and 7.01 The inoculated milk was incubated until pH 4.5 was reached. Reconstituted milk from the same batch of milk powder

was used as a control sample. Five replicates were prepared for each of the three sample types. The milk and fermented milk samples were then stored at −20 °C until MS analysis.

3.2.3 MALDI-TOF MS

3.2.3.1 Sample preparation

The water-soluble extract (WSE) for each sample was obtained according to the protocol outlined in De Noni and Cattaneo (2010), with adjustments. The milk and fermented milk samples were defrosted and brought to room temperature prior to sample preparation.

Milk-like samples: For milk-like samples (e.g., those with a pH > 4.5), 1 mL milk was acidified to pH 4.5 by slowly adding small drops of 1 M HCl. The samples were left for approximately 15 minutes prior to pH measurement. The acidified milk was then centrifuged at 5000xg for 20 minutes at 4 °C (Kubota, Tokyo, Japan). The supernatant was transferred to a fresh Eppendorf and centrifuged again at 5000xg for 20 minutes at 4 °C, to remove any additional casein and fat.

Fermented samples: For fermented samples (e.g., those with a pH < 4.5), 1 mL of fermented milk was placed in an Eppendorf tube. Samples were centrifuged at 5000xg for 20 min at 4 °C. The supernatant was collected, and the pH adjusted to 4.50 (+/- 0.01) using 1 M NaOH. The pH adjusted WSEs were re-centrifuged at 5000xg for 20 minutes at 4 °C.

The dilution resulting from the addition of HCl or NaOH was kept constant across all samples by adding Milli-Q water, where needed. Once the WSEs were obtained for all samples, MS analysis was performed as described in the following sections.

3.2.3.2 Peptide analysis

In the first stage of MALDI-TOF MS fingerprinting, the peptide fingerprint of the three samples was acquired. As peptides are typically ionised in positive mode, only this mode was used at the peptide level. MS/MS-based identification of molecules was not performed as the goal was to determine if fermented milk could be differentiated rapidly based on MS1 only. To perform peptide fingerprinting, a saturated solution of α-cyano-4-hydroxycinnamic acid (HCCA) was prepared in a

30:70 solution of acetonitrile and 0.1% TFA. A dried droplet deposition was done in duplicate by pre-mixing 1 µL of sample and 1 µL matrix and spotting 1 µL of the mixed solution onto the target plate (Anchor Chip plate, Bruker Daltonics, Germany) for each sample (Figure 3.1).

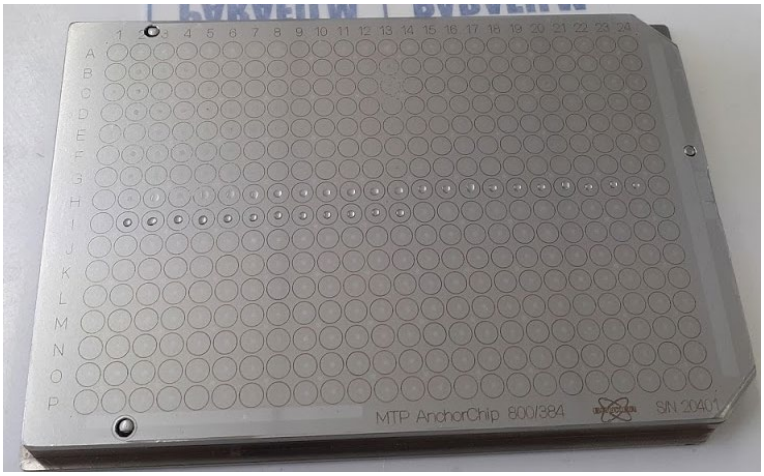


Figure 3.1 MALDI MS target plate. Sample is mixed with a matrix and spotted on the target plate prior to being inserted into the instrument.

3.2.3.3 Metabolite analysis

Graphene oxide matrix was prepared according to Liu et al. (2012). For each sample in duplicate, 1 mg of matrix was mixed in 1 mL solution of equal parts ethanol and 0.1% TFA and water. The mixture was sonicated for 3 minutes. A thin layer of the matrix was formed by pipetting 1 µL of the matrix onto the MALDI target plate. Once dried, 1 µL of the sample solution was pipetted onto the layer of dried matrix.

3.2.3.4 Instrument settings

All samples were fingerprinted using an Ultraflex III MALDI-TOF tandem mass spectrometer, containing a smartbeam™-I laser, with 355 nm wavelength (Bruker Daltonics, Bremen, Germany). For peptide analysis, the instrument was calibrated using peptide standard II (Bruker Daltonics, Bremen, Germany) as an external calibrant. The mass spectra were acquired in the range of *m/z* 700 to 3,500 in positive ion mode. 1,500 shots were accumulated for each spot in AutoXecute mode, with a frequency of 100 Hz, and the laser movement set to random walk.

For metabolites, the instrument was calibrated using a 1:9 mixture of 2,5-dihydroxybenzoic acid (DHB) and HCCA. The mass spectra were manually acquired in the range of *m/z* 50 to 750, in both positive and negative ion mode. 1,500 shots were accumulated for each spot, with a frequency of 200 Hz, and the laser movement set to random walk. The instrument was optimised for small compounds (i.e., the matrix suppression was turned off).

3.2.4 REIMS

REIMS analysis was performed directly on each sample as follows: 2 mL of sample was pipetted into an aluminium foil cup and a section of the surface vaporised using an electronic monopolar surgical knife (Electrosurgical pencil and Erbe VIO 50C generator, Erbe Medical UK Ltd, UK) at 50 V in cutting mode (Figure 3.2). The vaporisation time was 2 s, with 10 s between 'burns' to allow the MS signal to return to baseline. Samples were measured in triplicate. The MS data were acquired in both positive and negative ionisation mode using a Waters Xevo® G2 Q-TOF MS with REIMS interface (Waters Corp. Wilmslow, UK), in sensitivity mode at 100 µL/min isopropanol flow rate, with a scan time of 0.5 s and mass range *m/z* 50-1,200. The instrument was calibrated using 5 mM sodium formate in isopropanol before analysis. REIMS data collection was carried out in collaboration with the wider Fermented Foods team, at AgResearch, Lincoln.

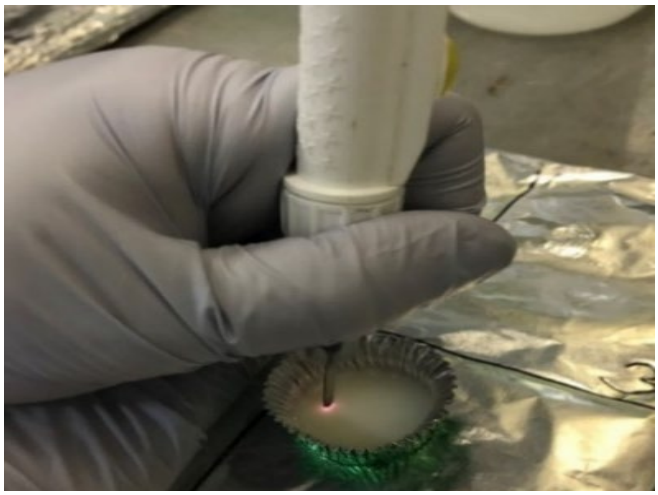


Figure 3.2 To analyse samples using REIMS, milk/fermented milk was placed in an aluminium foil cup; a surgical knife was applied to the surface of the milk. This cuts directly through the sample surface, vaporising it and injecting it into a mass spectrometer via a pump, generating a fingerprint of the metabolites.

3.2.5 Data analysis

Prior to data analysis, the MALDI-MS fid files (Bruker Daltonics, Bremen, Germany) were converted to mzXML format using CompassXport (Version 3.0.13.1, Bruker Daltonics, Bremen, Germany). The mzXML files were processed using the R packages 'MALDIquant' (Gibb and Strimmer, 2012), and 'MALDIquantForeign' (Gibb, 2018), with R version 4.0.2 (R Core Team., 2020). For both peptide and small molecules, processing steps included Savitzky-Golay smoothing, baseline correction by Top Hat method, intensity calibration by total ion current, spectra alignment and averaging across technical repeats. For peptide spectra, a square root transformation was performed before smoothing. Peak detection was performed using the Super Smoother method and a signal-to-noise ratio of 2. A feature matrix was then generated with detected *m/z* ions (peptides and metabolites) and corresponding intensity values. Raw spectra were visualised using FlexAnalysis, version 3.4 (Bruker Daltonics, Bremen, Germany).

The REIMS measurements were split into individual replicates, baseline removed, and mass-aligned to *m/z* 325.240 in negative mode and *m/z* 91.059 in positive mode using ProGenesis Bridge (Waters Corp. Wilmslow, UK). Further data processing in ProGenesis QI (Waters Corp. Wilmslow, UK) removed noise, grouped mass adducts and normalised the data. The normalised abundance of the detected ions from the technical replicates were averaged for each sample. The processed data were exported into Excel and imported into R for further analysis.

Principal components analysis (PCA) was conducted on the processed feature matrices for each measurement and visualised using the R packages 'FactoMinerR' (Le et al., 2008), and 'factoextra' (Kassambara and Mundt 2020). A permutational multivariate analysis of variance (PERMANOVA) was calculated to determine the PCA model significance using the 'vegan' package (Oksanen et al., 2013). A partial least squares discriminant analysis (PLS-DA) was performed on the processed data using the R package 'mixOmics' (Rohart et al., 2017). The R package 'ropls' was used to calculate performance metrics of the PLS-DA models (Thévenot et al., 2015). In the context of supervised classification, PLS-DA is performed to differentiate and

discriminate between sample classes. The concept was originally applied to regression problems but was extended to classification; in this context, the sample class is converted to a dummy variable, which is predicted using a regression (Rohart et al., 2017). The package 'pcaMethods' was used to calculate metrics for the PCA models (Stacklies et al., 2007).

Both MALDI-TOF and REIMS are rapid fingerprinting techniques with minimal sample preparation, allowing for a quick analysis time (Figure 3.3).

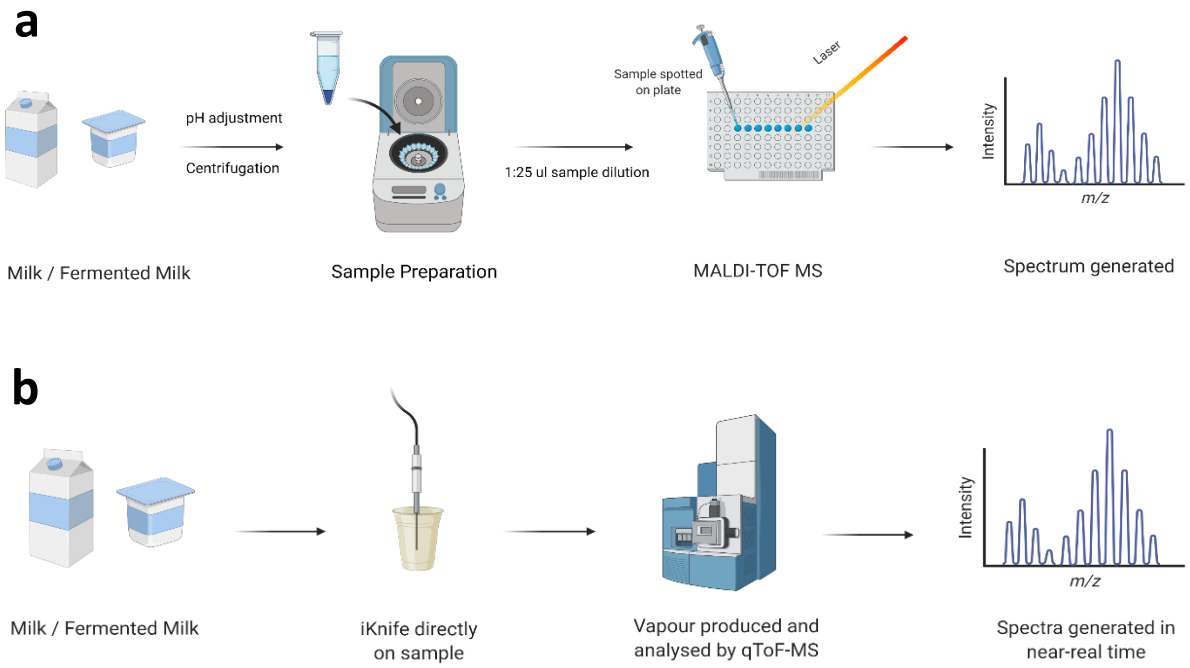


Figure 3.3 Workflow for sample preparation and data acquisition using MALDI-TOF MS. Samples are pH adjusted, followed by centrifugation to obtain a supernatant. Samples are diluted 1:25 ul TFA and spotted on target plate. Data acquisition can be performed manually or automated to obtain a spectrum ***(a).*** *Workflow for data acquisition using REIMS. Samples require no sample preparation prior to MS analysis. Spectra are obtained in near-real time* ***(b).*** *This figure was generated using BioRender.com.*

3.3 Results and discussion

3.3.1 MALDI-TOF

The raw mass spectra acquired from the MALDI-TOF MS peptide analysis showed distinct differences between the sample types (Figure 3.4) and resembled similar spectra observed previously for endogenous peptide profiles of milk (England et al., 2020, Ebner et al., 2015). Both

fermented milk products contained unique ions, i.e., *m/z* 2824.38 is present in YF-L811 only. The fermented products also had many compounds in common, i.e., *m/z* 1329.65 and 1991.09, as well as in common with the unfermented milk, i.e., *m/z* 2460.43.

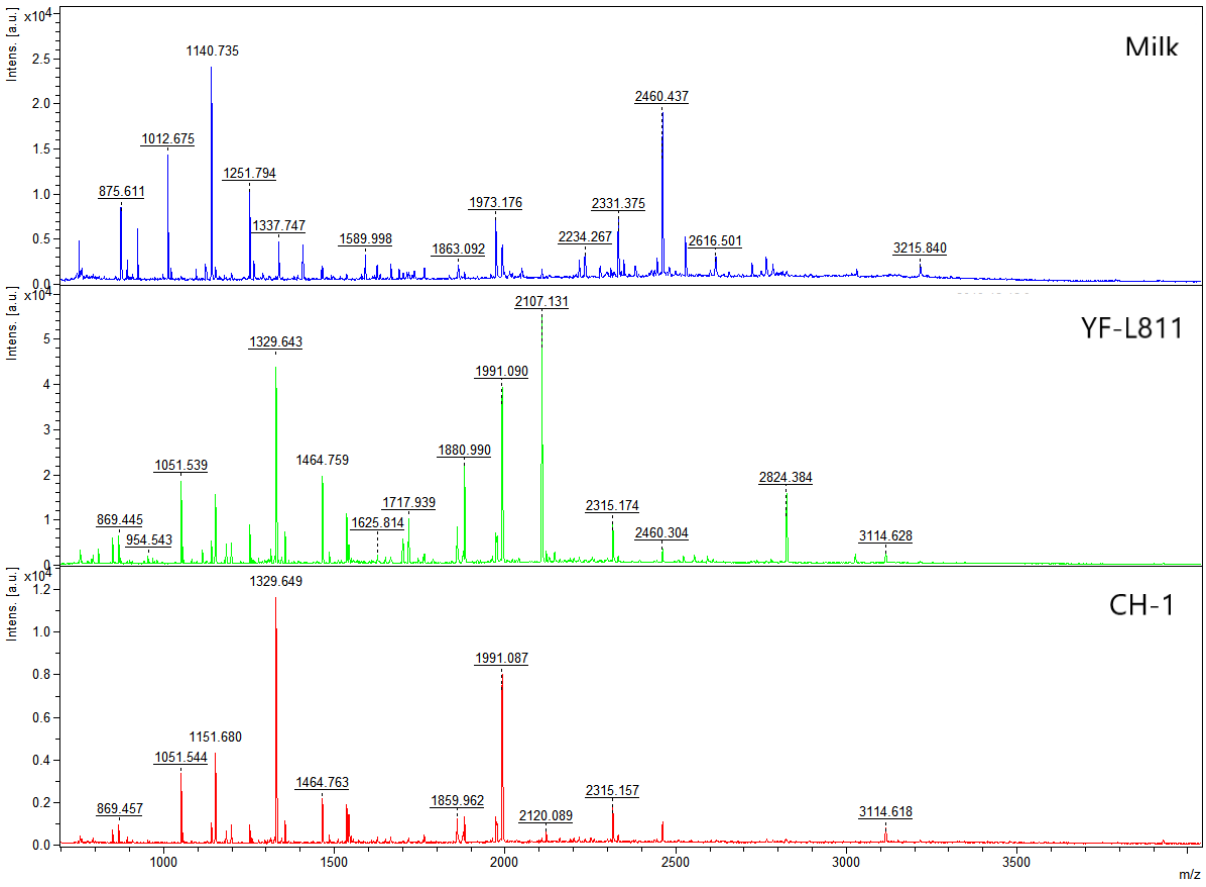


Figure 3.4 Raw, unprocessed peptide spectra for milk and fermented milk prepared from cultures CH-1 and YF-L811. Peptide fingerprints obtained by MALDI-TOF MS between m/z 700-3,500.

Unsupervised multivariate analysis of the peptide fingerprints shows three distinct groupings, corresponding to the three different products (Figure 3.5). The model metrics were: $R^2 = 0.69$ and $Q^2 = 0.58$. R^2 is the explained variance or goodness of fit of the model, and Q^2 is a measure of the goodness of prediction of the model (Blasco et al., 2015). An R^2 value > 0.5 has been suggested as a good explanation of biological models (Blasco et al., 2015). There is no critical value of Q^2 for statistical significance, although a value greater than 0.4 is considered adequate (Blasco et al., 2015, Worley and Powers, 2013). A PERMANOVA of the model was statistically significant ($P \leq 0.001$). There is little variation between the replicates of each sample type, as they cluster closely together

in the PCA space. The PCA demonstrates that MALDI-TOF MS can effectively group these samples based on their peptide fingerprints and that different peptides were present in the milk before and post-fermentation.

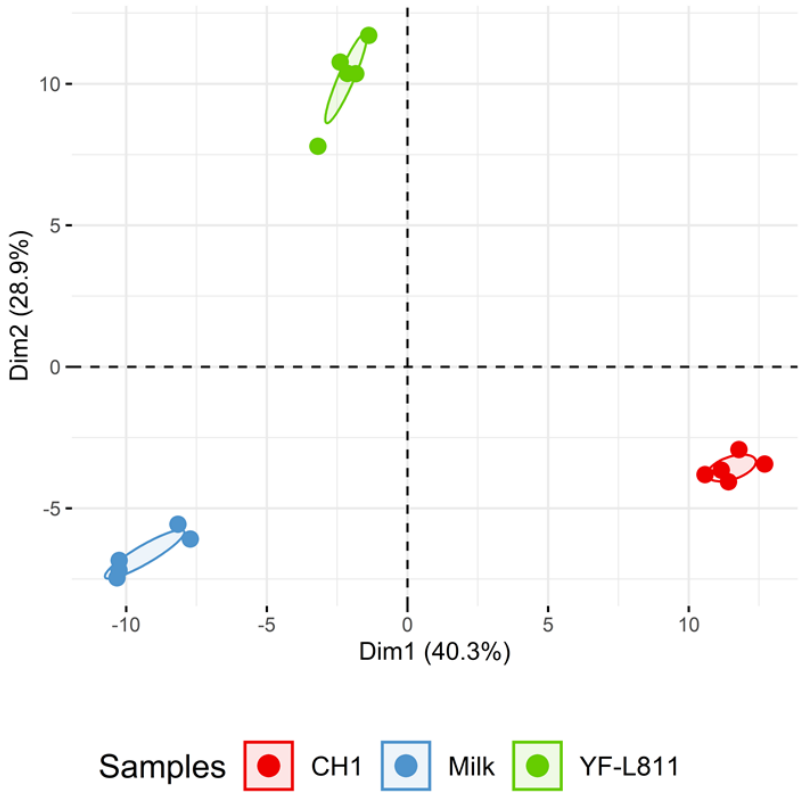


Figure 3.5 PCA plot of peptide fingerprints obtained using MALDI-TOF MS on milk (blue) and milk fermented using CH-1 (red) and YF-L811 (green), with 95% confidence ellipses. N = 5 replicates for each sample type, MALDI-TOF MS replicates (n = 2) were averaged. A PERMANOVA of the model was statistically significant ($P \leq 0.001$).

Prior to sample preparation, the performance of the matrix (HCCA) was compared to other commonly used MALDI-TOF MS matrices. HCCA performed better than DHB: the number of peaks and signal intensity of samples fingerprinted with HCCA was greater than those prepared using DHB, and HCCA samples were also less similar in a PCA plot, compared to DHB samples (Appendix 3, 4). HCCA was selected as a matrix for use in these experiments, as a result. Filtration devices are frequently used in the sample preparation of fermented milk samples prior to MS analysis (Yildiz et al., 2011, Nguyen et al., 2014, Theodorou and Politis, 2016). This removes large proteins and peptides which can interfere or suppress smaller ions. Ultrafiltration devices (3 kDa

and 10 kDa filter) were trialled prior to this experiment and were found to remove a large number of ions in the ultrafiltrate that were present in the unfiltered samples, as well as introducing some suppressing and interfering ions (Appendix 5). As such, no filtration devices were used for these experiments.

In the second stage of fingerprinting with MALDI-TOF MS, small molecules present between *m/z* 50 – 750 were analysed in positive and negative ionisation mode. The spectra obtained for these small molecules had substantially more background noise compared to the peptide spectra; this is expected due to the matrix interference at the low mass region. For both modes, the raw spectra contained many ions in common between the samples, though there were some clear differences in the compounds of the fermented versus unfermented samples (Figure 3.6a; b). The spectra acquired in positive ionisation mode showed more distinguishable differences between the sample types than those observed in the negative ion mode spectra. In the negative mode spectra, there were no ions present in the region of ≥ *m/z* 300. The positive mode spectra, on the other hand, contained numerous ions between *m/z* 500-750, that were present only in the fermented samples. This could be a result of the amalgamation of smaller metabolites or the breakdown of larger ones. Notably, there was an increase in the relative intensity of some compounds (e.g., *m/z* 381.1 in positive mode) in the fermented product compared to the milk. This was also notable between the *m/z* 150 – 250 region. A blank matrix spectrum was also collected to assess the level of interference introduced by the matrix. Conventional matrices, such as HCCA, are not well-suited for the analysis of molecules less than *m/z* 500, due to interference from the matrix peaks in the low mass region and its inefficiency in ionising low molecular weight compounds (Kang et al., 2001). To allow efficient ionisation of small molecules, graphene oxide was used as a matrix to fingerprint the smaller peptides and metabolites in this experiment. Graphene oxide is not a well-established matrix and because MALDI-TOF MS is not routinely used for the analysis of such compounds, various matrices were first trialled prior to the experiment (Appendix 7, 8).

a

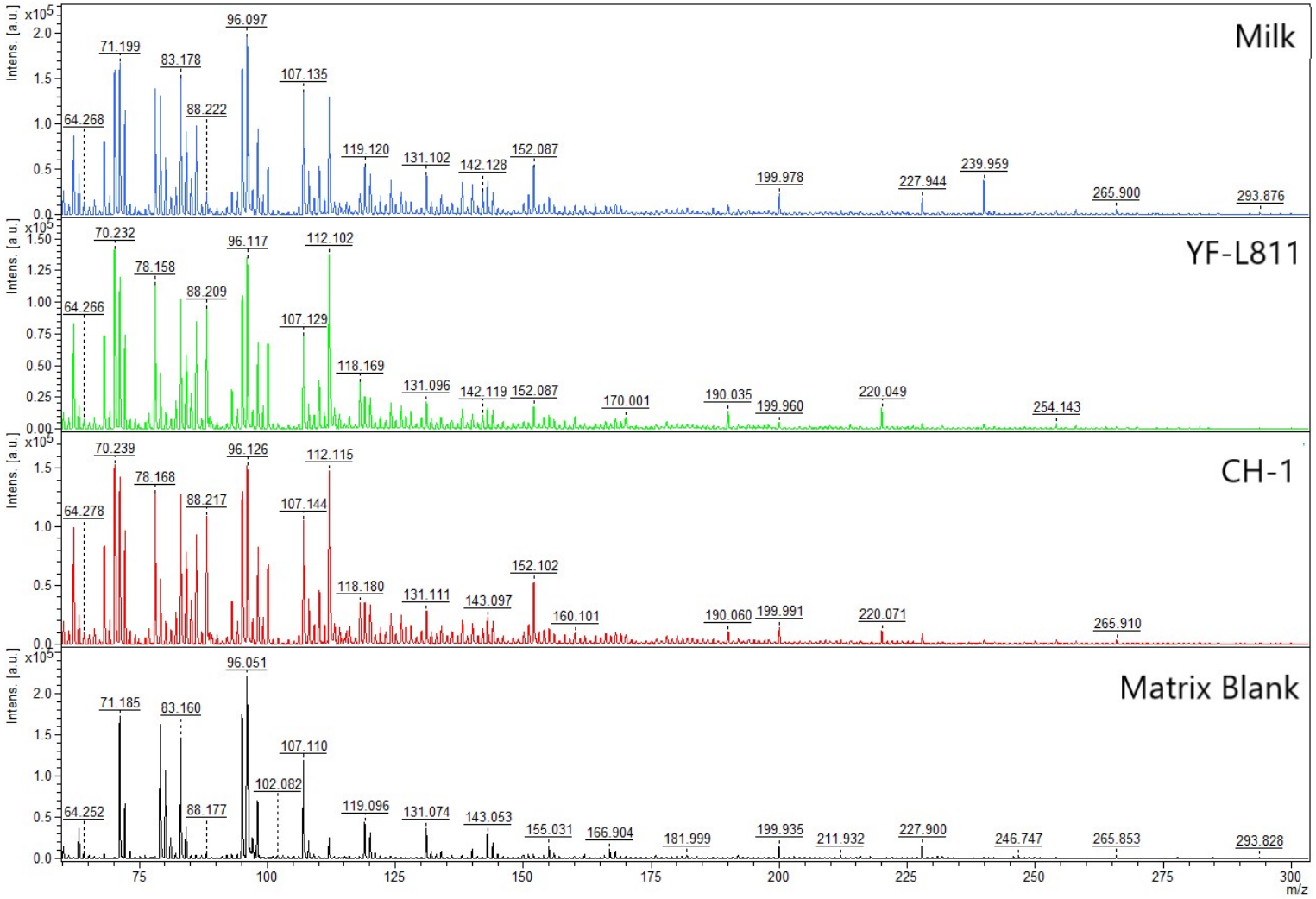
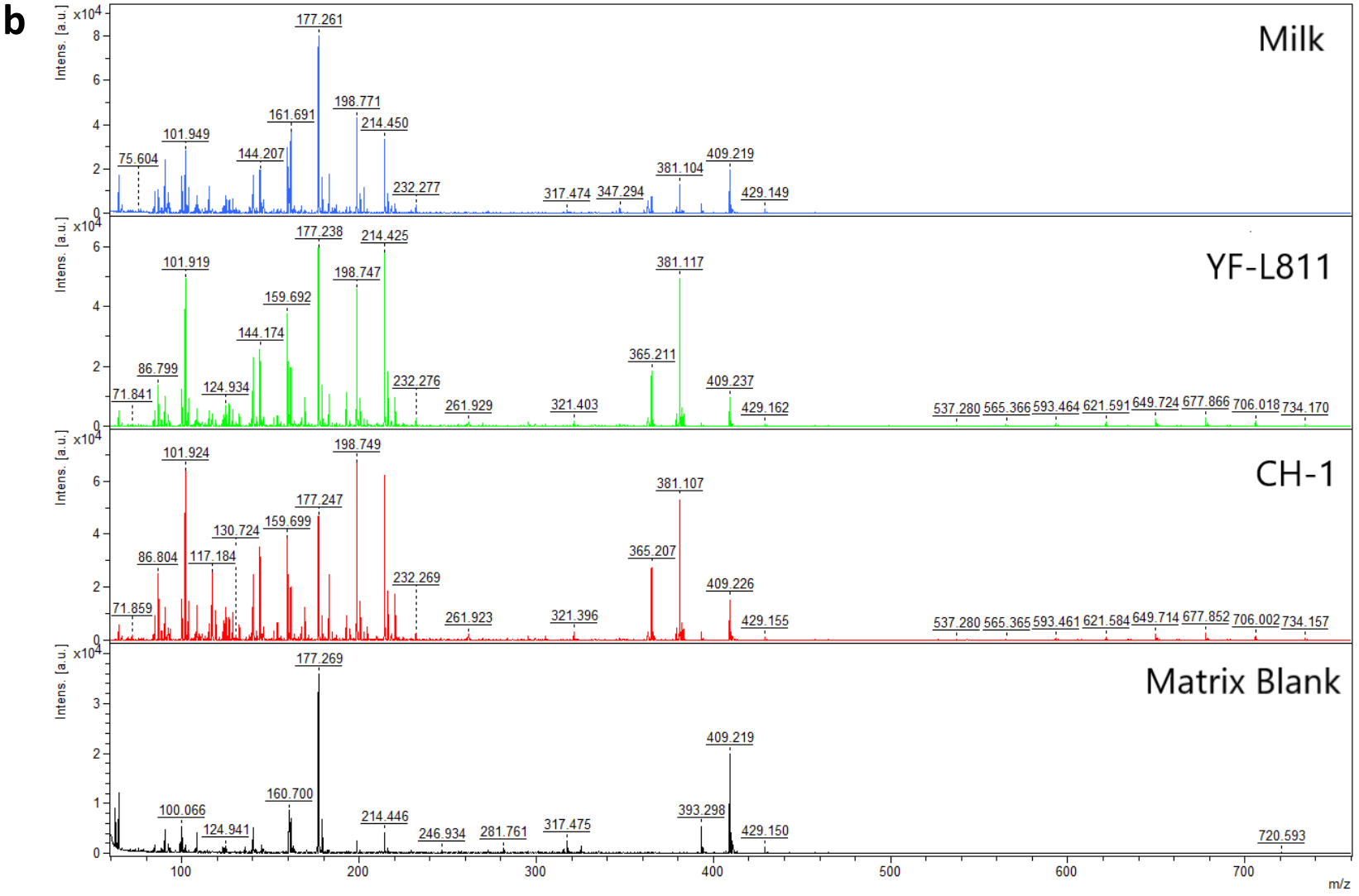


Figure 3.6 Small molecule fingerprints obtained in negative mode by MALDI-TOF MS, between m/z 50-750. Only m/z 50-300 is displayed ***(a).*** *Small molecule fingerprints obtained by MALDI-TOF MS in positive mode, between m/z 50-750* ***(b).*** *Graphene oxide was used as matrix.*

Figure 3.6 continued



A PCA of the small molecule fingerprints indicated some clustering according to sample type, although there was no clear grouping between the different sample types (Figure 3.7a). The model metrics were: R^2 = 0.49 and Q^2 = 0.08, indicating overfitting. Overfitting is the concept of overly optimising results, where the model fails to generalise over new, unseen data (Kuhn and Johnson, 2013). Large variations between R^2 and Q^2 (i.e., $R^2 >> Q^2$), in addition to a low Q^2, are an indication of overfitting (Worley and Powers, 2013). In positive mode, the unfermented milk formed a different cluster from the fermented milk, although there was an overlap of the YF-L811 and CH-1 samples (Figure 3.7b). Compared to the peptide fingerprints, there was more variation within the replicates for each sample, particularly for the milk samples which did not cluster closely in the PCA space for either negative or positive mode. The variation seen in the PCA plot could be due to poor ionisation of the low molecular weight compounds, as well as interference from the matrix. The model metrics were: R^2 = 0.51 and Q^2 = 0.10, again indicating overfitting. Both models were statistically significant, according to a PERMANOVA ($p \leq 0.05$).

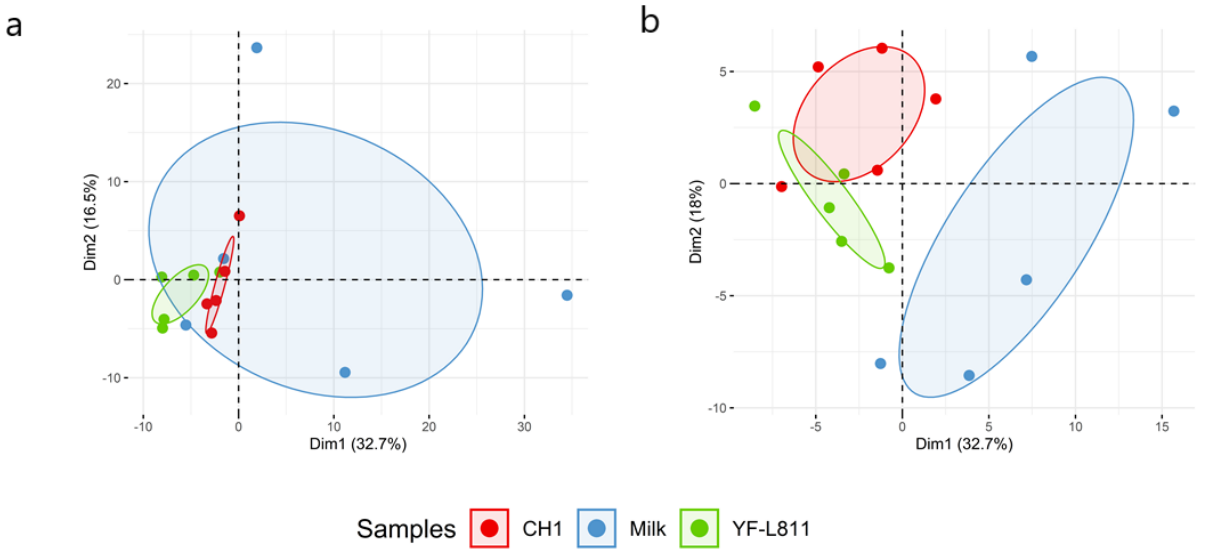


Figure 3.7 PCA plot of small molecule fingerprints obtained using MALDI-TOF on milk (blue) and milk fermented using CH-1 (red) and YF-L811 (green) in negative mode ***(a)*** *and in positive mode* ***(b)****, with 95% confidence ellipses. N = 5 replicates for each sample type, MALDI-TOF replicates (n = 2) were averaged. A PERMANOVA of each model was statistically significant (negative mode p ≤ 0.001, positive mode p = 0.004).*

Next, a supervised analysis was performed to further explore the differences in the fingerprints of the small molecules obtained by MALDI-TOF using a PLS-DA. Using a PLS-DA on the negative mode fingerprints, the three groups were separated using components 1, 3 and 4 (Figure 3.8a). A PLS-DA using component 2 was unable to group the samples according to their respective classes. The model fit metrics were: $R^2X = 0.56$, $R^2Y = 0.96$ and a $Q^2 = 0.6$. R^2X is the X variance explained by the model (Blasco et al., 2015); R^2Y is the sum of variation in Y (sample class) explained by the model. This model indicates good class separation (R^2Y), and good predictive ability (Q^2), but the disparity between R^2Y and Q^2 indicate it may be overfit. A PLS-DA applied to the positive mode samples was able to separate the two fermented milk samples using components 1, 2 and 4, although the YF-L811 samples were not fully separated from the unfermented milk (Figure 3.8b). The model fit metrics were: $R^2X = 0.69$, $R^2Y = 0.87$ and a $Q^2 = 0.28$. This model indicates good class separation, but a low predictive ability and the variability between R^2 and Q^2 suggest it is overfit.

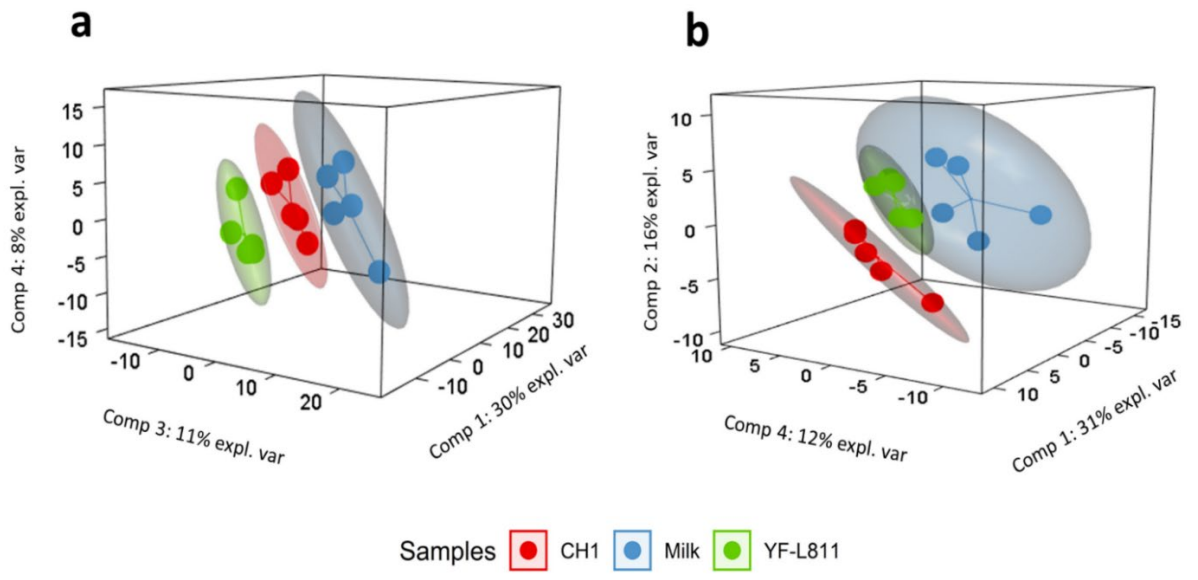


Figure 3.8 Milk (blue) and fermented milk samples prepared from starter cultures CH-1 (red) and YF-L811 (green). PLS-DA performed on small molecules obtained by MALDI-TOF MS in negative mode ***(a)*** *and in positive mode* ***(b)****, with 95% confidence ellipses.*

Unsupervised analysis of the small molecules detected by MALDI-TOF MS was able to cluster the fermented milk to some extent, although, the peptide fingerprints demonstrated a greater ability to distinguish the different milk products. The peptide profile of other fermented milk products has previously been shown to change during fermentation; the number of peptides increased in the fermented product and there were few peptides reported to be in common between fermented and unfermented samples (Ebner et al., 2015). This is consistent with the results of the present experiment: the *m/z* ions in the fermented samples were more abundant in the raw spectra of the fermented samples, and multivariate analysis demonstrates there are few similarities between the peptide fingerprints of these products. The profile of small compounds (including amino acids, peptides, and fatty acids), obtained via UPLC-QTOF-MS, in fermented milk was recently shown to be different to that of unfermented milk and has been demonstrated using multivariate analysis techniques (Li et al., 2020b). In this study, the low molecular weight compounds proved less informative than the peptides.

3.3.2 REIMS

The raw spectra obtained by REIMS for the fermented milk samples contained many of the same *m/z* ions, for both positive and negative mode, though there were several compounds present in the CH1 samples that are not observed in the YF-L811 samples (Figure 3.9a; b). There were few ions present between *m/z* 900-1200, and so this region was excluded when visualising the raw spectra. Both the fermented milk samples and unfermented milk showed compounds with the same *m/z* in the negative ion mode. Some compounds (e.g., *m/z* 311.17, 325.18 and 329.20) were observed in the negative mode in the milk as well as both of the fermented milk. The unfermented milk contained a number of ions > *m/z* 600, that were not present in either fermented milk. The positive spectra for the fermented milk samples were also very similar and contained many of the same peaks between *m/z* 50 – 300. The CH1 samples, however, showed more compounds than YF-L811, particularly between *m/z* 400 to 900.

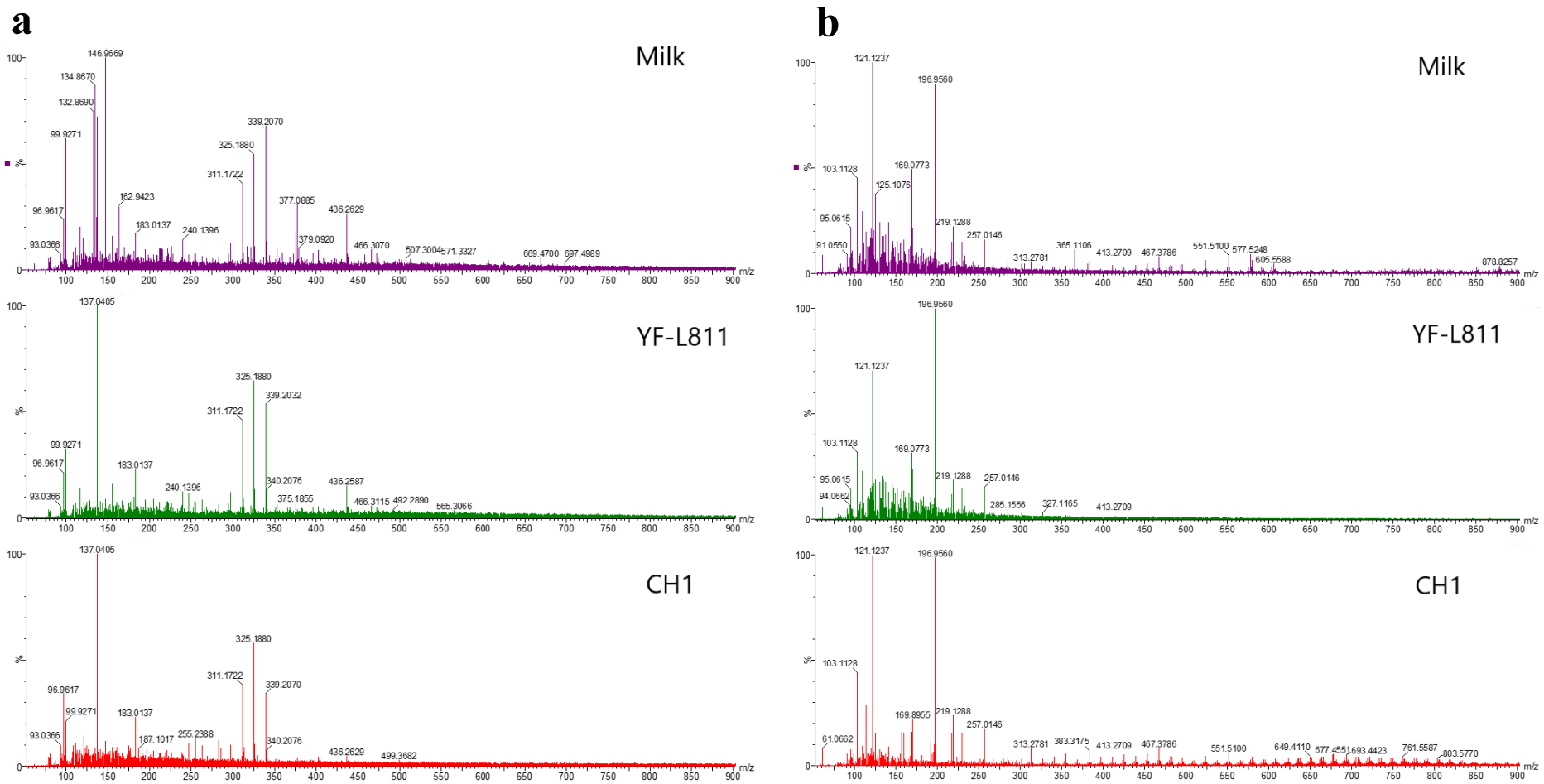


Figure 3.9 Negative mode profiles obtained by REIMS, between m/z 0 – 900 ***(a).*** *Positive mode profiles obtained by REIMS, between m/z 0 - 900* ***(b).***

Unsupervised analysis of the REIMS fingerprints in negative mode was able to cluster the unfermented milk separately from the two fermented milk samples; however, there was an overlap between the fermented milk (Figure 3.10a). The model metrics were: $R^2 = 0.46$ and $Q^2 = 0.27$. The low Q^2 suggest poor predictive ability. Post-processing, REIMS detected 3,320 *m/z* ions in negative mode. A PCA of the positive mode samples showed the milk samples forming a different cluster to the fermented milk (Figure 3.10b). There was some overlap between the two fermented milk samples, with one YF-L811 sample clustering closely with the CH-1 samples. The model metrics were: $R^2 = 0.42$ and $Q^2 = 0.10$. REIMS detected 6,092 *m/z* ions in positive mode. A PERMANOVA of both models was statistically significant ($p \leq 0.001$). In this analysis, distinct clustering could be seen in the fermented versus unfermented samples, suggesting that the small molecules in the fermented samples undergo a compositional change during fermentation that is detectable by REIMS. Although the clusters corresponding to the fermented samples were overlapping, indicating that the molecular fingerprint is similar in the two fermented products.

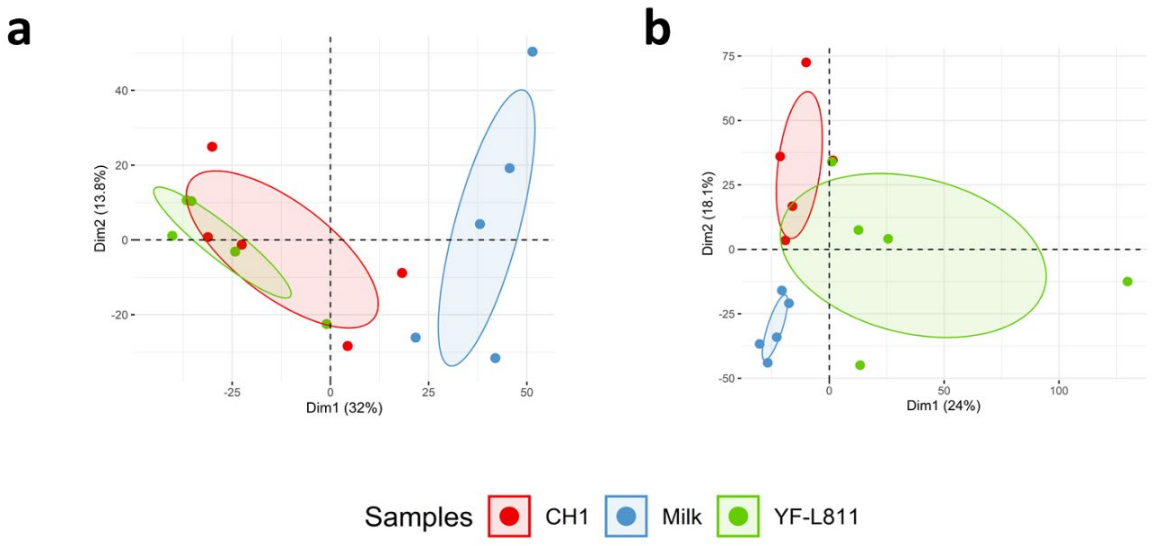


Figure 3.10 Milk (blue) and fermented milk samples prepared from starter cultures CH-1 (red) and YF-L811 (green). PCA plot of metabolites obtained by REIMS in negative mode ***(a)*** *and positive mode* ***(b)****, with 95% confidence ellipses. A PERMANOVA of the models was statistically significant (negative mode $p \leq 0.001$; positive mode $p \leq 0.001$)****.*** *N = 5 replicates for each sample type, REIMS replicates (n = 3) were averaged.*

Supervised analysis of the negative mode fingerprints, using components 1, 2, 3 was able to separate the three sample groups (Figure 3.11a). The model fit metrics were: $R^2X = 0.47$, $R^2Y = 0.88$ and a $Q^2 = 0.38$. A PLS-DA on the positive mode fingerprints using components 1, 2, 3 was also able to discriminate the three sample groups (Figure 3.11b). The model fit metrics were: $R^2X = 0.49$, $R^2Y = 0.88$ and a $Q^2 = 0.55$. Both models indicate good class separation (> 0.5), but a low predictive ability and the variability between R^2 and Q^2 suggest they are overfit, particularly in the case of negative mode.

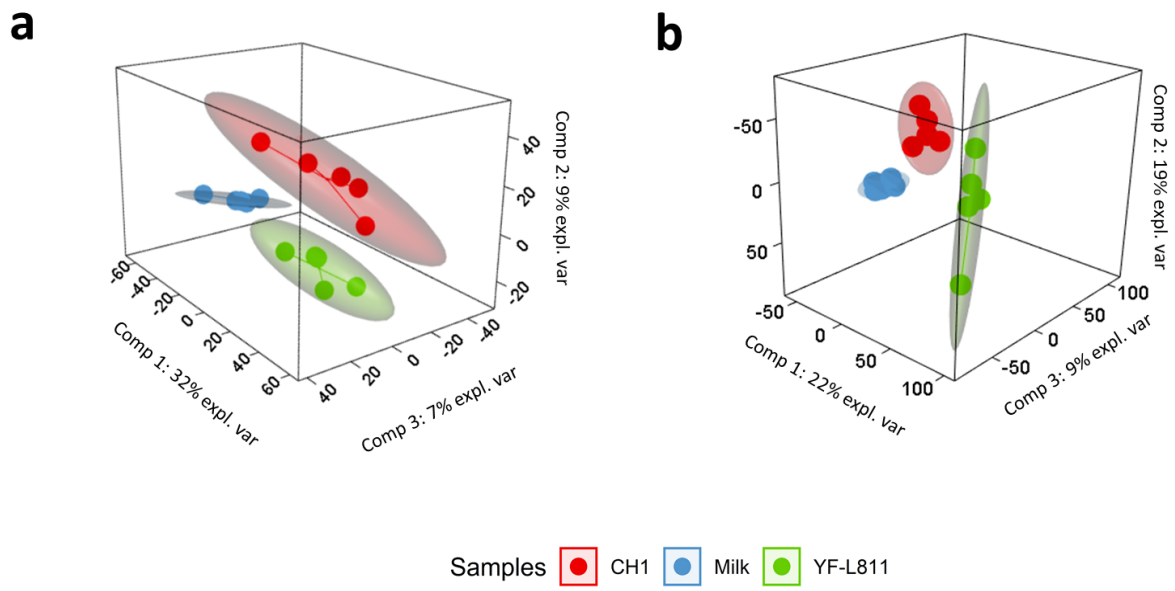


Figure 3.11 PLS-DA of metabolites from REIMS in negative mode ***(a).*** *PLS-DA of metabolites from REIMS in positive mode* ***(b)****, with 95% confidence ellipses.*

The greatest advantage to REIMS is that it requires no sample preparation, and a fingerprint can be obtained in near real-time. REIMS can detect a wide range of small molecules and lipids, with polar compounds, such as small phenolic compounds, fatty acids and phospholipids readily detected in negative mode, and non-polar compounds, (e.g., acylglycerols) and some phenol derivatives readily detected in positive mode (Arena et al., 2020). While PCA can be useful as an exploratory tool to visualise the relationship between samples, in metabolomic fingerprinting studies separation of sample class typically depends on the use of discriminant analysis techniques, such as PLS-DA or orthogonal-PLS-DA (Wang et al., 2019, Paxton, 2020). Classification rates in discriminant

models can also be improved using feature reduction or sparse models (i.e., removal of noisy features that do not contribute to the discrimination). It is possible that the metabolites primarily detected by REIMS, such as lipids and fatty acids, do not undergo major compositional change during fermentation of milk using the analysed bacterial cultures, and thus there is little variation between the samples at this level. It is also possible that REIMS does not readily detect the primary metabolites that are present in milk and fermented milk, though there are no similar studies with which to compare. The absence of published work on non-solid substrates also limits the interpretation of these results. Techniques that will allow REIMS to better analyse liquid samples are currently being investigated by other researchers, which may improve REIMS' capabilities for the analysis of dairy and liquid products (Paxton, 2020).

3.4 Conclusions

This chapter reported the use of different fingerprinting technologies to rapidly discriminate between unfermented milk and milk fermented milk using different starter cultures. Both techniques can detect a broad range of molecule class, and as preliminary screening tools can provide insights into the compounds undergoing compositional change during fermentation. This can direct further studies targeting a specific class of compound during the development of fermented milk products. Both fingerprinting techniques were able to show some clustering according to the sample type using unsupervised multivariate analysis. PCA analysis revealed that the molecules driving variation between the groups is more apparent at the peptide level compared to the small molecule level. The REIMS fingerprint did not result in distinct clusters for the different cultures using non-discriminant analysis (PCA), likely this is due to the majority of molecules detected by REIMS not being impacted by the differences in the two cultures. Discriminant analysis did lead to the separation of the different sample types, reflecting that REIMS could be used for differentiating between the two different cultures. Although model performances were poor in this instance, feature reduction can improve model performance and lead to improved classification rates. Advances in applications of REIMS for dairy and liquid

samples may also show promise for the use of this instrument in the dairy industry in the future. In summary, changes in the peptide fingerprints demonstrated the greatest discrimination between the different products analysed in this study, suggesting that the mode of proteolysis is the main differentiator between the two cultures tested, which could relate to the contrasting flavour and textural properties of the fermented milk. This chapter demonstrated that peptide fingerprinting via MALDI-TOF was an effective and rapid means to discriminate between different fermented milk products and may be used in further applications to discriminate fermented milk products and to monitor product development. This technique was therefore selected as a means for rapid fingerprinting and discrimination of milk fermented with various bacterial cultures in subsequent chapters in this thesis.

Chapter 4 Peptide fingerprinting and predictive modelling of fermented milk

Summary

The objective of this chapter was primarily to model the peptide fingerprint of fermented milk throughout fermentation. Fingerprinting techniques established in Chapter 3 were employed to obtain peptide fingerprints of fermented milk prepared from a variety of starter and probiotic bacterial combinations. The results section is divided into three parts: exploration and visualisation of peptide fingerprints using untargeted multivariate analysis techniques, implementation of regression techniques to predict the change in signal intensity across fermentation time for individual *m/z* ions, and finally implementation of classification techniques to classify signal intensities between unfermented and fermented samples. Exploration of fingerprints using untargeted multivariate analysis techniques revealed that the samples primarily group by starter culture and fermentation time. Regression techniques could successfully model the signal change in some cases. Classification techniques were successful in classifying the direction of signal intensity over fermentation time.

4.1 Introduction

Understanding the changes occurring during fermentation can be useful for quality control and monitoring of certain parameters that might be critical for sensory, safety or health-promoting properties. Monitoring and prediction of fermentation parameters have been applied for various fermented foods using regression and classification techniques, using a variety of techniques to collect data as input for predictive models (Mains et al., 2017, Temizkan et al., 2020, Li et al., 2020a). Such techniques have been applied to predict fermentation characteristics and outputs, such as pH, fat, protein, lactose, and salt content (Subramanian et al., 2009, Cimander et al., 2002, Temizkan et al., 2020), as well as predicting the concentration of important taste-related compounds (Hruskar et al., 2010). These models are beneficial for enabling scheduled fermentation

processes, improving throughput and efficiency, improved quality control, and enhancing sensory characteristics.

Products generated during fermentation, via proteolysis, can contribute to the taste, flavour, and can have bioactive functions in fermented milk. Prediction of MALDI-TOF MS signal intensity has been explored previously, using various physicochemical properties of peptides to predict peak presence and/or intensity (Gay et al., 2002, Aiche et al., 2012), but not with regard to the development of peptides during fermentation of milk. Predicting changes in the peptide fingerprint during fermentation can be beneficial while screening for new bacterial strains to enable the development of new and diverse fermented products, with targeted and desirable characteristics. Such insights can be used as a starting point for monitoring, prediction, and targeted screening of potentially important compounds generated during the fermentation of milk.

Background to chapter: To provide some context and rationale for this chapter, the objectives and methods are discussed in more detail below.

The primary aim of this chapter was to explore and predictively model the peptide fingerprints over fermentation time. Ultimately, this chapter sought to answer the following: can the peptide fingerprint be used to provide a means to discriminate milk fermented with different bacterial combinations, and can this fingerprint be predicted with the expectation that some of the components may have a direct or indirect impact on the sensory profile? Insights from such analyses could then be used as basis for targeted screening, prediction and monitoring of potentially important molecules produced during fermentation.

Probiotic cultures were investigated in this chapter for several reasons. The addition of probiotic cultures can contribute to the commercial value of fermented products by generating peptides that can enhance the nutritional properties (Aryana and Olson, 2017). Probiotic cultures have also been reported to impact the metabolites generated in yoghurt products: the concentration of volatile compounds was increased in yoghurts fermented with *Lactobacillus acidophilus* (Østlie et al.,

2003, Østlie et al., 2005) and Bifidobacterium (Oliveira et al., 2012, Özer and Kirmaci, 2010, Prasanna et al., 2014). Because of the potential value of adding probiotics, as well as the comparatively few studies investigating the influence of probiotics on peptide fingerprint (and related compounds) throughout fermentation, this chapter sought to investigate if differences in the peptide fingerprint of probiotic cultures could be modelled throughout fermentation.

Furthermore, as the intention of this work is to contribute to screening for new bacterial cultures, it was considered important to understand if rapid fingerprinting technologies (MALDI-TOF MS) could extend to discriminate milk fermented with different probiotics. This would provide some understanding of the limitations of this work: can MALDI-TOF MS also be used for screening different probiotic cultures, or is the variation in the peptide fingerprint mostly influenced by the starter culture? To answer some of these questions, principal components analysis was performed to visualise the fingerprints and to establish which factors were driving the variation.

Predictive modelling techniques were performed to explore if the peptide fingerprint could be predicted throughout fermentation. The objective in this section was to initially explore the feasibility of predicting the signal intensity for individual peaks (which are *assumed* to possess relevant biological functions for the purposes of this work) throughout fermentation. The goal for this work was to understand how peptides break down throughout fermentation, components that may be impactful on the final sensory profile, and to be able to predict how such changes to the fingerprint occur throughout the course of fermentation. By providing some insight into the change in these components throughout fermentation, this potentially could reduce the need to carry fermented milk products to full fermentation during initial screening, saving time and cost. As a first step in the modelling process, linear and nonlinear regression models were explored to predict individual *m/z* ions/peaks throughout the course of fermentation. It should be acknowledged that although these techniques have their limitations and may not be the most suitable, they were employed as a first step to explore the data, and to understand if classical statistical techniques could suitably model this data.

Following on from this, classification techniques were explored. The objective was to predict more generally whether a peak would "increase" or "decrease" in signal intensity from early to late fermentation time. For this purpose, only a subset of the peptide fingerprint was used; only the peaks that displayed a relatively large change in intensity between fermentation times were considered here, as it is assumed that they would have a more pronounced effect on the sensory profile. Essentially, it is assumed that peaks that are not changing throughout fermentation are unlikely to have a significant impact on the sensory profile of the final product.

For the purposes of this work, some assumptions were made which could be expanded on in future works. The *m/z* ions, or peaks (which comprise the peptide fingerprint), referred to throughout were not identified and as such it cannot be said for certain that they do correspond to actual peptides. As such, when referring to prediction, the term *m/z* ion or peak is used, rather than referring to them as "peptides". However, the working assumption in this thesis is that these peaks do correspond to real peptides, with biological properties that potentially may be of significance.

4.2 Materials and methods

4.2.1 Chemicals and reagents

MALDI-TOF MS matrices and peptide standards were purchased from Bruker Daltonics (Bremen, Germany). Acetonitrile was liquid chromatography-mass spectrometry (LC-MS) grade from Fluka Analytical (Merck, Kenilworth, NJ, USA). Optima LC-MS grade water and trifluoroacetic acid (TFA) were obtained from Fisher Scientific (Hampton, NH, USA).

Freeze-dried yoghurt starter cultures YF-L811 and YC380, both containing different strains of *Streptococcus thermophilus* and *Lactobacillus delbrueckii* subsp. *bulgaricus* were purchased from Christian Hansen (Hørsholm, Denmark). According to the technical information for the two cultures, YC380 has a strong flavour with medium viscosity, and YF-L811 has a very mild flavour and very high viscosity. The experiments carried out in Chapter 3 were done in collaboration with the wider project team (Fermented Foods Team) and due to the availability of bacterial cultures,

YC380 was used in place of CH-1 for the remaining work in this thesis. CH-1 and YC380 both have similar properties in terms of flavour, and so it was assumed that results from Chapter 3 would be consistent with results in subsequent chapters.

Probiotic cultures of *Lactobacillus acidophilus* (LA5), Bifidobacterium (BB12), and *Lactobacillus paracasei* subsp. *paracasei* (LC) were also purchased from Christian Hansen (Hørsholm, Denmark). The cultures were stored at −20 °C until use. Skimmed bovine milk powder was purchased from Fonterra (Fonterra Co-operative Group, New Zealand) and stored at 4 °C until use.

4.2.2 Experimental design

A randomised block design was generated using CycDesignN (version 6.0) using the following parameters: a resolvable row-column design was generated with 6 rows and 6 columns. The spatial arrangement was intended to account for potential temperature variations in the water bath that might induce slower or faster fermentation rates (Aguirre-Ezkauriatza et al., 2008). Three factors (starter culture, probiotic, fermentation time) with 2 * 3 * 6 levels, for a total of 36 samples (Table 2).

Table 2 Factor levels used in experimental design.

Factor	1	2	3
Levels	2	3	6
	Starter	**Probiotic**	**Fermentation Time**
	YF-L811	LA5	0
	YC380	BB12	1
		LC	2
			3
			4
			5

This experimental design was replicated and randomised across five days for a total of 180 samples. Each time point was prepared in a separate tube and removed from the water bath, accordingly. Samples were prepared in separate tubes according to fermentation time point to limit potential

contamination when removing samples. Furthermore, preparing the samples separately would limit agitation of the mixture which is reported to interfere with the fermentation process (Aguirre-Ezkauriatza et al., 2008). An example layout of the sample placement in the water bath is presented in Table 3.

Table 3 An example of the placement of samples in the water bath according to experimental design generated in CycDesignN. The numbers correspond to the Starter, Probiotic, Fermentation Time Point, where e.g., (2,3,5) corresponds to Starter YC380, Probiotic LC and fermentation time point 5 (e.g., 5 hours).

(2,3,5)	(1,3,3)	(2,1,3)	(2,2,4)	(1,2,5)	(1,1,2)
(1,2,4)	(2,2,1)	(1,3,1)	(2,3,2)	(1,1,3)	(2,1,5)
(1,1,1)	(2,1,2)	(1,2,2)	(1,3,5)	(2,3,1)	(2,2,0)
(2,2,3)	(2,3,4)	(1,1,4)	(2,1,1)	(1,3,0)	(1,2,1)
(1,3,2)	(1,2,0)	(2,2,5)	(1,1,0)	(2,1,4)	(2,3,3)
(2,1,0)	(1,1,5)	(2,3,0)	(1,2,3)	(2,2,2)	(1,3,4)

4.2.3 Preparation of fermented milk

Six different inoculated milk mixes were prepared using a combination of two different starter cultures combined with one of three probiotic cultures. The probiotic cultures were not mixed in combination with one another. Samples were prepared as outlined in Figure 4.1.

Reconstitution of skim milk: Two litres of milk were reconstituted at 9.1% (w/v; 3% protein) by slowly adding milk powder to distilled water while agitating gently using a magnetic stirrer. The mixture was prepared to contain 3% protein; this was in line with the objectives of the wider project team. The solution was mixed for 30 minutes and stored overnight in a fridge at 4 °C to allow the milk to fully rehydrate.

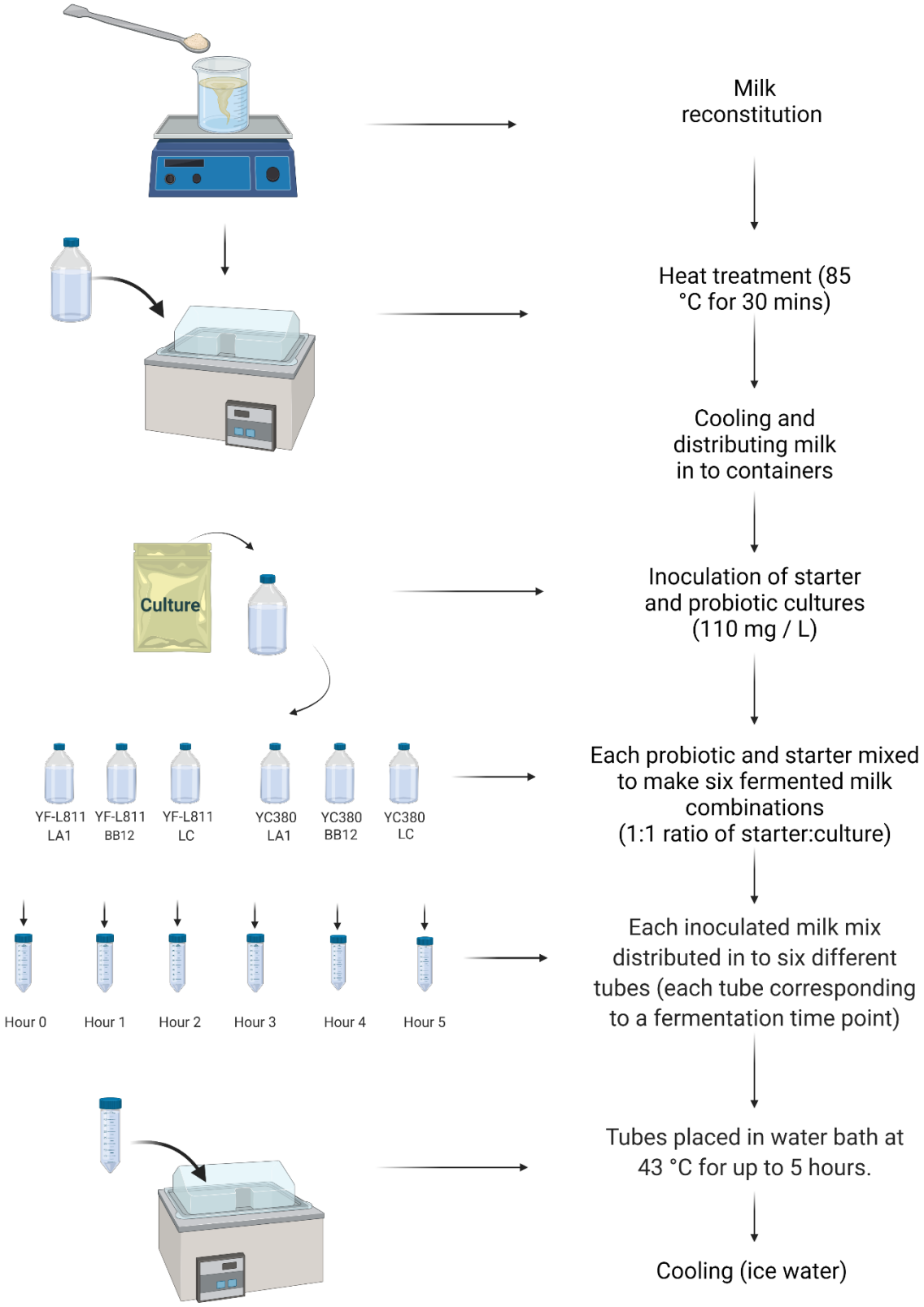


Figure 4.1 Workflow to demonstrate the preparation of fermented milk. Skimmed milk powder was reconstituted, followed by a heat treatment, and a 1:1 inoculation of starter: probiotic. Samples were fermented at 43 °C for up to 5 hours. Samples were removed each hour and placed on ice. This figure was generated using BioRender.com.

Heat treatment: The milk was heat-treated in a Grant OLS Aqua Pro Shaking Water Bath (Grant Instruments Cambridge Ltd., Shepreth, UK). The reconstituted milk was aliquoted in to 500 mL Duran® bottles and placed in the water bath, set to 88 °C at 80 rpm. The temperature of the milk was monitored every 5-10 minutes until the milk reached a temperature of 84 °C (+/- 0.2 °C). A "test" bottle was used to monitor the temperature milk, to reduce potential contamination of the milk. Once the milk reached 84 °C, the water bath was adjusted to 85 °C and timed for 30 minutes. The milk was cooled in ice water to 43 °C. The milk, minus test bottle, was pooled together under a laminar flow hood and then aliquoted in to six 200 mL bottles. The milk was pooled to ensure that there were no variations in the milk being used for each treatment.

Inoculation of starter culture: Each culture was prepared in sterile milk at a concentration of 110 mg/L of skim milk. This was achieved by mixing 500 mg of each culture in 50 mL of sterile milk and gently agitating using a magnetic stirrer for 15 minutes. Each starter culture was mixed with heat-treated milk by adding 2.2 mL of the starter culture in to three different bottles of 200 mL milk. Next, 2.2 mL of each probiotic mixture was added to a bottle containing each of the starter cultures to make six different inoculation milk mixes. The mixtures were gently agitated for 10 mins using a magnetic stirrer to ensure the culture was fully dispersed in the milk.

Fermentation: Approximately 35 mL of each of the six fermented milk treatments was aliquoted in to six different polypropylene tubes (each tube representing a fermentation time point). Each tube was labelled and placed into the water bath according to the randomised design (Figure 4.2), set to 43 °C. Samples corresponding to Hour 0 were immediately withdrawn from the water bath. Samples were collected at five time points/hours during fermentation. Samples were removed and immediately placed on ice water to halt fermentation. Samples were allowed to ferment up to 5 hours to ensure that they had reached a pH of ≤ 4.5. Samples were stored at −20 °C until MS analysis.

Figure 4.2 Placement of tubes/fermented milk in the water bath. The position of each tube in the water bath was randomised according to the experimental design prepared in CycDesignN. Each tube corresponds to a different starter/probiotic/fermentation time point.

pH Measurements: The pH of each sample was recorded in triplicate to track the progress of the fermentation using a pH meter (Hanna Instruments, USA). The pH meter was calibrated using buffers of 4.01 and 7.01 (Hanna Instruments, USA). To measure the pH of each sample, approximately 1 mL of sample at room temperature was placed in a small beaker. The pH electrode was placed directly into the milk sample. The pH reading was allowed to stabilise for a few seconds before the reading was taken. Between measurements of different samples, the pH electrode was rinsed using distilled water. Samples were discarded after the pH reading was taken.

4.2.4 MALDI-TOF MS sample preparation

Samples were first thawed and brought to room temperature. Samples were then prepared as described in Section 3.2.3.1 and 3.2.3.2. Samples were prepared in the same block as the preparation of the fermented milk, i.e., milks fermented on day 1 were prepared for MS analysis at the same time.

4.2.5 MALDI-TOF analysis

Three repeats of MALDI-TOF data acquisition were performed for each sample.

Instrument Settings: All samples were profiled using an Ultraflex III MALDI-TOF tandem mass spectrometer, containing a smartbeam™-I laser, with 355 nm wavelength (Bruker Daltonics, Bremen, Germany). The instrument was calibrated using peptide standard II (Bruker Daltonics, Bremen, Germany) as an external calibrant. The mass spectra were acquired in the range of m/z 700 to 3,500 in positive ion mode. 3,500 shots were accumulated for each spot in AutoXecute mode, with a frequency of 66.7 Hz, 90 shots were collected per raster shot and the laser movement was set to random walk.

4.2.6 Data pre-processing

Processing of the MALDI-TOF spectra was performed as described in Section 3.2.5. Following processing, a feature matrix was generated with detected peaks (or *m/z* ions) and corresponding intensity values. The feature matrix generated was averaged by technical repeats (n=3). This resulted in a feature matrix with 143 peaks and 178 samples. Two samples were removed as they had been mis-labelled during sample preparation. This was observed during exploratory analysis via principal components analysis. The feature matrix was used for subsequent analysis, unless otherwise stated. The terms peak and *m/z* ion may be used interchangeably throughout this chapter. Although peptide fingerprinting was performed in this chapter and it is assumed for the purpose of this work that the detected peaks/ions correspond to peptides, the peaks were not identified and therefore cannot be confirmed or referred to as true peptides.

4.2.7 Data analysis and modelling

4.2.7.1 Exploratory analysis

Principal component analysis (PCA) was performed on the processed feature matrix using the R packages ‘factoextra’ (Kassambara and Mundt 2020) and ‘factominer’ (Le et al., 2008). The PCA allowed visualisation of patterns and trends within the fingerprints. To understand the influence of

each factor (i.e., fermentation time, starter, probiotic), three PCA plots were generated and colour-coded by factor. A permutational multivariate analysis of variance (PERMANOVA) was calculated using the package 'vegan' (Oksanen et al., 2013), to compare groups in the PCA.

4.2.7.2 Predictive modelling

Individual *m/z* values were investigated using regression techniques to establish if these techniques could be used to model individual peaks over fermentation time and to predict the signal intensity recorded by MALDI-TOF. This also served as an exploratory exercise to understand the behaviour of individual peaks across different bacterial combinations throughout fermentation.

Multiple Linear Regression: Multiple linear regression analysis (Wasserman, 2004) was performed on individual peaks obtained from the feature matrix (n=143 peaks) using base R function 'lm' and visualised using the package 'ggplot2' (Wickham, 2011). Diagnostic plots were generated using base R functions and 'ggplot2' (Wickham, 2011). The multiple regression models were generated using the signal intensity of individual peaks as response variable, and the fermentation time (numeric), starter culture (factor) and probiotic culture (factor) as explanatory variables. Note, assumptions were checked after model generation.

First, a null, intercept-only model was generated for each peak (n=143 models) and compared against a model with increasing complexity to understand the significance and performance of each predictor term. This was assessed using two criteria: first, by generating each model with added interaction terms and comparing the resulting Akaike information criterion (AIC) for each model. AIC is a technique frequently used to compare multiple statistical models (James et al., 2013, Pedersen et al., 2019, Portet, 2020). Two or more models may be evaluated, and an AIC value calculated; the lower the AIC value, the more optimal the model, though the values themselves bear no actual meaning. When comparing models using AIC, as a standard rule of thumb, for a model to be considered significantly better, it should be ≥2 AIC units lower than another model (Burnham & Anderson 1998).

Secondly, an Analysis of Variance (ANOVA) was carried out to compare each model to assess the significance of each added term and evaluate whether the final model significantly outperforms null or less complex models (James et al., 2013). The model for each peak was summarised by extracting the adjusted R^2 and model p-value, with adjustment of p-values by Benjamini-Hochberg correction (Benjamini and Hochberg, 1995). An adjusted R^2 accounts for the number of predictors in the model (James et al., 2013).

Multiple linear regression models were visualised using the package 'ggplot2' (Wickham, 2011). Plotting the data can reveal problems with a model that cannot be determined from statistics alone, such as the R^2 (James et al., 2013), i.e., over, or underestimating data points. If the plot resembles data from a bivariate normal population, then R^2 can be considered a meaningful metric; however, R^2 can be viewed as less reliable when data do not follow this population (Weisberg, 2013).

Model evaluation: The suitability and performance of the multiple linear regression models were then assessed by inspecting model diagnostic plots. The diagnostic plots display four plots: (1) the original data to assess for a linear relationship between predictor and response variable. (2) The residuals vs fitted values should show the points sitting randomly and evenly around zero. When there is an obvious pattern in the shape of the data points, the model may be poorly fit and not follow assumptions of a linear model. (3) A histogram of residuals should be normally distributed, with a bell-shaped curve where most of the data points sit around 0. If there is a tail either side, it indicates it does not follow assumptions of the model. (4) A QQ-plot of residuals, the model residuals are compared to a normal distribution. Well-fit residuals should be close to a straight line; data that deviate from this line indicate a poor fit and can be used to indicate potential outliers (Wasserman, 2004, James et al., 2013). Other criteria, such as collinearity or high-leverage could also be used for assessing the assumptions and suitability of a regression model, although these issues are not reported here.

Generalised additive model: Generalised additive models (GAM) were generated using the R package 'mgcv' (Wood, 2015). The models were visualised using the package 'tidymv' (Coretta, 2021). The GAM model was performed to further explore the behaviour of the peaks through time, and to establish whether a non-linear model may be more effective in generalising over these data.

GAM models were generated using the signal intensity of individual peaks as response variable, and using a smooth term for fermentation time, and an interaction smooth between all three factors. To compare the performance of a GAM model over a multiple linear regression model, a model was generated using each technique for all peaks (n=143), using base R functions, 'aov' and 'AIC'. The models were assessed by comparing the two models by AIC, ANOVA and adjusted R^2.

A diagnostic check of the GAM model was performed to test for patterns in the residuals, in a similar manner to the multiple linear regression diagnostics.

Classification and data preparation:

Discriminant analysis: The feature matrix was reduced to select only for peaks that were significantly changing throughout fermentation, eliminating peaks of low intensity and peaks that were changing minimally throughout fermentation. A linear discriminant analysis was performed to rank the peaks changing significantly over time using the R package 'sda' (Ahdesmaki et al., 2015). A subset of the peaks changing most over time was selected; a subset of '40' peaks were retained and were visualised according to their t-score ranking. The t-scores, based on a statistical t-test, are used to identify the most important class discriminating peaks (Bø and Jonassen, 2002, Inza et al., 2004, Mundra and Rajapakse, 2016). The peaks in the feature matrix are then ranked by their t-score. In this case, this allowed for visualising peaks that are typical or characteristic of early/late fermentation times. The assumption being that if a peak is characteristic of early fermentation, then it is less prevalent at later fermentation times. The discriminant analysis uses a shrinkage t-statistic (James-Stein) which makes no assumptions of distributions for data or the model parameters. This approach has been found to lead to consistently high rankings of features

(Opgen-Rhein and Strimmer, 2007). The decision to retain 40 was somewhat arbitrary and not based on any specific threshold or criteria. This has been reported in other studies, where a seemingly arbitrary number of features has been selected for analysis (Blanco et al., 2018, Ahdesmaki et al., 2015, Bø and Jonassen, 2002). Incorporating false discovery rates (FDR) can be used as a threshold to select a subset of features that meet a certain criteria. This was not utilised in this case as the number of samples was deemed to be too low to produce reliable statistics for FDR calculations (Ahdesmaki et al., 2015).

Data preparation: The top 40 peaks were extracted from the original feature matrix and were reformatted for classification using the packages 'reshape2' (Wickham, 2007) and 'tidyverse' (Wickham et al., 2019). The processed feature matrix (now containing 40 peaks (columns) * 178 samples (rows)) was reformatted by converting the data to a long format, i.e., the feature matrix was transformed to have each instance of a peak as a row rather than a column, giving a total of 1120 data points at each fermentation time point. For classification, fermentation hour 4 was selected as an exemplar, with the aim of predicting the changes in peak intensity over time through prediction of the class labels (increase or decrease), from hour 0 to hour 4. Hour 4 was selected as the pH of these samples most closely resembled that of commercial fermented milk (most samples had pH ~ 4.5). The differences in intensity between time points were calculated and each instance of a peak was assigned a label to indicate the direction of the intensity (either increasing or decreasing) over time. Boxplots were generated using 'ggplot2' of the top 40 peaks to visualise the change in signal intensity between early and late fermentation time points. Data were randomly sampled and divided into training and test sets (70% training and 30% test) for model generation.

Model generation: Classification techniques were performed using the R package 'H_2O' (LeDell et al., 2020), 'xgboost' (Chen et al., 2021), and 'RWeka' (Hornik et al., 2007). Five models were generated: deep learning (DL; Kuhn and Johnson, 2013), random forest (RF; Breiman, 2001), gradient boost machine (GBM; Friedman, 2001), eXtreme Gradient Boosting (XGBoost; Chen and Guestrin, 2016), and C4.5 (Quinlan, 1993). Each model was generated using the default model

parameters. The response variable was the class label (increase or decrease), with signal intensity as predictor variable.

Assessing model performance: The following metrics were calculated for each model and used to assess model performances: mean class error, area under the curve (AUC), area under the precision-recall curve (AUCPR), F-measure, Matthew's correlation coefficient (MCC) and predictive accuracy. There are numerous metrics with which to assess model performance in classification models, and it is recommended to use more than one metric for model selection (Hastie et al., 2009, Kuhn and Johnson, 2013, James et al., 2013). The appropriate metric can vary depending on the data, and the objective of the test. In binary classifications, accuracy and F-measure are among the most popular metrics used to evaluate performance (Hastie et al., 2009, Kuhn and Johnson, 2013, James et al., 2013). The F-measure combines two other metrics, precision, and recall, where precision is the accuracy of the positive class and recall is the number of true positives divided by the number of all possible positive samples. MCC is a metric that is said to be more robust and informative for binary classification than the F-measure (Chicco and Jurman, 2020). MCC returns a value between -1 and 1, where -1 indicates a perfect mis-classification, and +1 indicates a perfect correct classification. Other metrics used include accuracy, which is the number of correct predictions as a ratio of all predictions made. The AUC evaluates how well a binary classification can distinguish between true positives and false positives, where a value closer to 1 is a perfect classifier. The AUCPR is more sensitive than AUC and is suggested to be a more informative metric than the AUC (Chicco and Jurman, 2020).

4.3 Results and discussion

4.3.1 Exploration of samples and peptide fingerprints through multivariate analysis

Untargeted multivariate analysis was performed on the processed peptide fingerprints (feature matrix) using principal components analysis (PCA: Figure 4.3). The PCA in Figure 4.3b indicated that the biggest chemical variation in these samples is derived from the difference in starter culture. The samples were separated along the first principal component (PC1), accounting for 33.4% of

the variance, with the exception of a cluster of samples corresponding to the unfermented milk samples at hour 0. Samples were separated along PC2 by fermentation time (Figure 4.3a), accounting for 19.3% of the total variance. In the case of YC380, the unfermented and hour 1 samples were further separated along PC1. No clear grouping could be seen by probiotic culture (Figure 4.3c).

For YF-L811, the earlier ferments cluster separately, with the hour 1 samples clearly splitting from the unfermented samples, but the later ferments clustered closely in the PCA space (hour 3-5). For YC380, samples formed more distinctive groups at later fermentation times, with the exception of hours 4 and 5. The grouping in the PCA plots was consistent with what is seen in the pH profiles (Appendix 10); YC380 samples were slower to ferment and grouped closely at the early fermentation times, whereas YF-L811 was quicker to ferment and grouped separately at early fermentation times. The PCA suggest that samples fermented using YC380 as starter culture were undergoing greater changes in the peptide composition during fermentation than those fermented using YF-L811. This is further supported by the visualisation of peak patterns, where there were distinct differences between fermentation time points and between starter cultures. YC380 is known to generate a product with a strong flavour; the separation of YC380 samples at later fermentation times could be reflective of the intense flavour being generated and may suggest that the peptide fingerprint is a component of the flavour formation of the product, amongst the numerous other compounds known to contribute to flavour in fermented milk products.

Each treatment was also modelled separately by PCA by bacterial combination which indicated a greater separation and more distinctive grouping by time in the PCA space. Hierarchical clustering and PCA modelling by individual time point also indicated some clearer grouping by probiotic type (Appendix 16, 17).

The PCA plots established that MALDI-TOF MS can capture the variation in peptide fingerprint attributed to fermentation time and starter culture, and to a lesser extent the probiotic cultures.

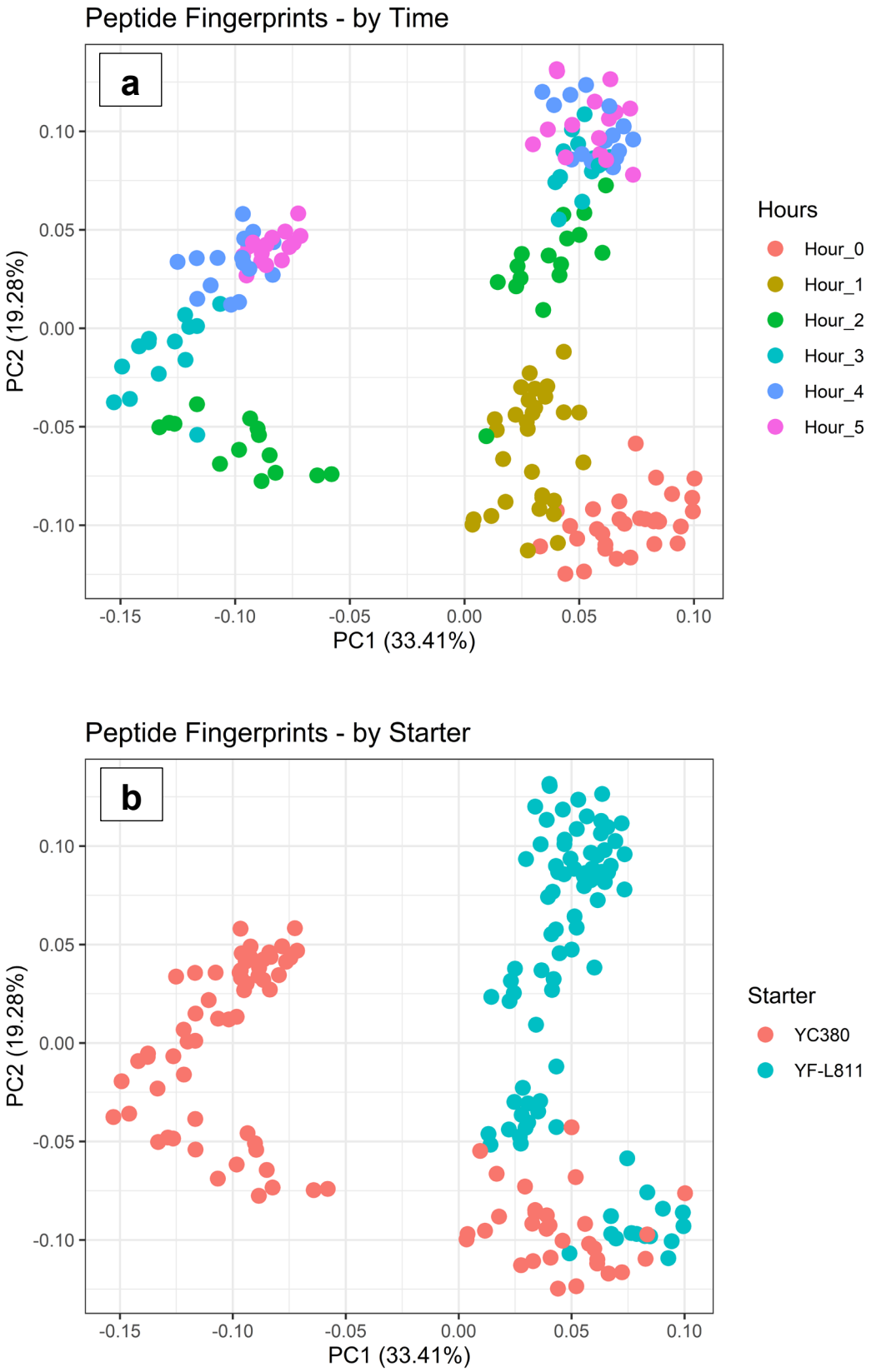
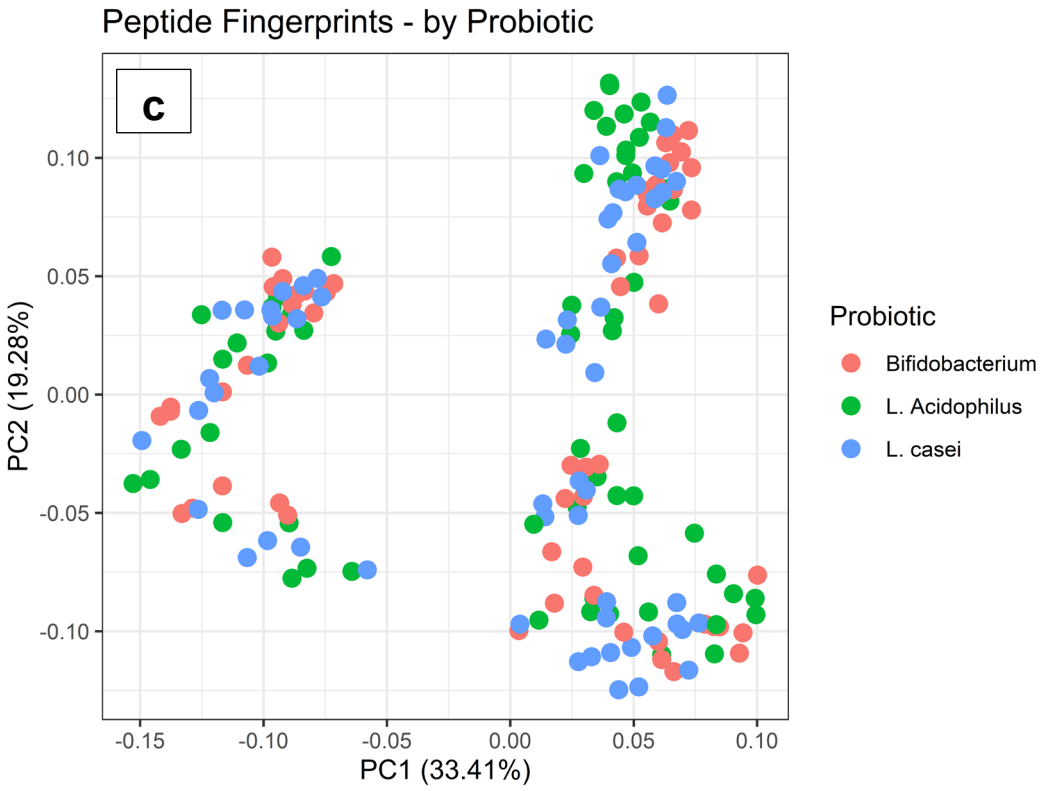


Figure 4.3 PCA plot of the peptide fingerprints, for all replicates, coloured by fermentation time ***(a)****, starter culture* ***(b)****, probiotic culture* ***(c).*** *A PERMANOVA of each model indicated significant differences between the treatments (P<0.05).*

Figure 4.3 continued



Multivariate techniques have previously been demonstrated to be effective at discriminating the small compounds in various types of raw milk (England et al., 2020), to characterise the effects of fermentation on milk (Li et al., 2020), as well as various processing parameters including inoculation rate (Dan et al., 2017) and addition of adjunct cultures (Tian et al., 2017, Ebner et al., 2015). The volatile compounds of yoghurts fermented using various probiotic cultures were analysed while in storage by Tian et al., (2017). Yoghurts fermented using *L. casei* were sampled across multiple storage time points and all formed one group in a hierarchical clustering analysis, whereas all other samples (fermented with *L. acidophilus, L. plantarum*, and *L. rhamnosus*, and starter cultures only), clustered together by storage time rather than by bacterial culture (Tian et al., 2017), indicating that *L. casei* has a pronounced effect on the volatile profile of yoghurts, while in storage.

4.3.2 Exploration of individual peaks during fermentation using regression techniques

4.3.2.1 Linear regression

Next, individual *m/z* values were explored using classical statistical techniques. A multiple linear regression model was assessed for its suitability to predict the change in signal intensity across time points for each peak detected by MALDI-TOF MS (n=143 peaks/models). Predictor terms were assessed using AIC and ANOVA (Table 4). The full interaction model was significant (according to ANOVA) and was more suitable (according to AIC) for about two-thirds of the peaks analysed. As indicated in Table 4, some of the models were not improved by the addition of any term, i.e., only 115 out of 143 models were significant when time was added as a term vs. the null model. Inspection of the AIC values across each model revealed that differences in AIC between some models was minimal (<1) and would suggest that the added term is not always justified based on this criteria alone.

Table 4 Comparing added terms for each peak (n = 143). Performance and significance of models with added complexity were assessed by AIC and ANOVA. m/z signal intensity was used as the response variable.

Intercept-only model vs. added interaction terms		
Model	Number of models with lower AIC	Number of significant models ($p \leq 0.05$)
Null vs Model 1 (time)	123	115
Model 1 vs Model 2 (time * starter)	139	134
Model 2 vs Model 3 (time * starter * probiotic)	91	90

Next, a model was generated for each peak using all interaction terms. Though this model did not significantly improve performance for all peaks, the full model was generally improved compared to a less complex model (Table 5). A summary of the peaks with the highest and lowest adjusted R^2 is presented in Table 5. Of these, 116 models/peaks were significant ($p \leq 0.05$), and 96 had an adjusted R^2 greater than 0.5, 35 had an R^2 value greater than 0.7. The multiple linear regression

model performed poorly in some cases, with adjusted R^2 values < 0 for some (Table 5). A full list of the *m/z* ions and corresponding adjusted R^2 and *p*-values from the multiple linear regression model are reported in Appendix 18. Evaluating and interpreting R^2 very much depends on the application and data being analysed. James et al., (2013) suggest that large residual errors can be expected in typical biological applications, due to various other factors that have not been measured and accounted for. In such cases, an R^2 as low as 0.1 might be expected. In the current study, there are other factors that might be influencing the response variable. This could include anything from variations during the fermentation process, sample preparation and instrumental analysis. These factors, as well as other unmeasured factors, may account for some of the variation in this data.

Table 5 Adjusted R^2 and p-values of model with full interaction terms. This is a subset of peaks with the highest and lowest R^2 values. P-values were adjusted using Benjamini-Hochberg correction.

	m/z	*Adj_R^2*	*p-value*
Lowest	1988.17	-0.01	1.00
	1746.62	0.07	0.13
	1761.65	0.11	0.61
	786.7	0.12	0.63
	1066.13	0.15	0.25
	1189.79	0.16	1.00
	3307.15	0.16	≤0.001
	3499.59	0.18	1.00
	1614.42	0.22	≤0.001
	1633.88	0.22	0.06
Highest	2395.59	0.82	≤0.001
	1976.61	0.82	≤0.001
	2201.87	0.84	0.01
	3346.38	0.85	≤0.001
	2807.34	0.86	≤0.001
	1860.65	0.86	≤0.001
	2120.84	0.87	≤0.001
	2300.61	0.88	≤0.001
	2854.27	0.89	≤0.001
	2316.09	0.89	≤0.001

To assess the performance of the multiple linear regression on these data and to explore how each peak behaved by treatment, the model was visualised and evaluated. As an exemplar, two models

were evaluated: one model which had a low adjusted R^2 and one with a high R^2. A multiple linear regression model of *m/z* value 1860.65 had an adjusted R^2 of 0.86 and was significant ($p \leq 0.05$). The model was visualised, faceting by starter culture and colour-coded according to probiotic culture (Figure 4.4). This model indicated a positive linear relationship of signal intensity over time, the model performed well for each bacterial combination and was suitably describing the variance in the model.

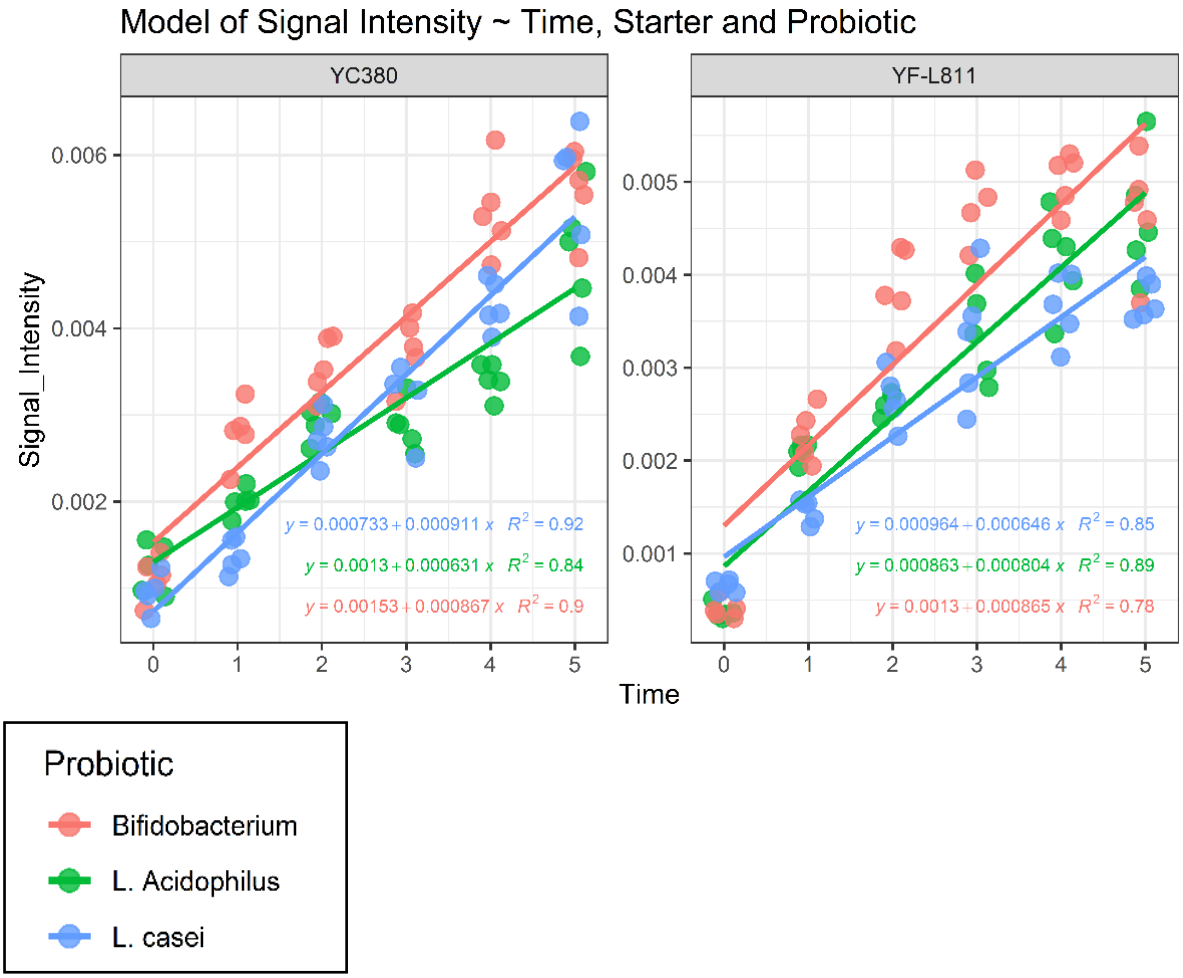


*Figure 4.4 Multiple linear regression model of m/z value 1860.65. Fermentation hour, starter culture and probiotic terms were used as predictor terms to predict the response of signal intensity. The signal intensity is along the y-axis, fermentation time is along the x-axis. The plot is faceted by starter culture and coloured by probiotic culture. Legend indicates the corresponding probiotic culture. Formula: lm(Signal_Intensity ~ Time * Starter * Probiotic).*

The suitability and performance of the regression model were assessed by inspecting the model diagnostic plots (Figure 4.5). A scatter plot of residuals vs fitted values (Figure 4.5b) show most of

the data points scattered around the blue line, with no distinctive pattern. The histogram of residuals (Figure 4.5c) appears to be normal, indicating no obvious skewness. The QQ-plot of residuals (Figure 4.5d) shows most of the data points follow the line, though there are two samples that are selected as potential outliers: samples 58 and 75. Overall, this appears to be a good fit and the model describes the data well, with no major concerns for violations of model assumptions.

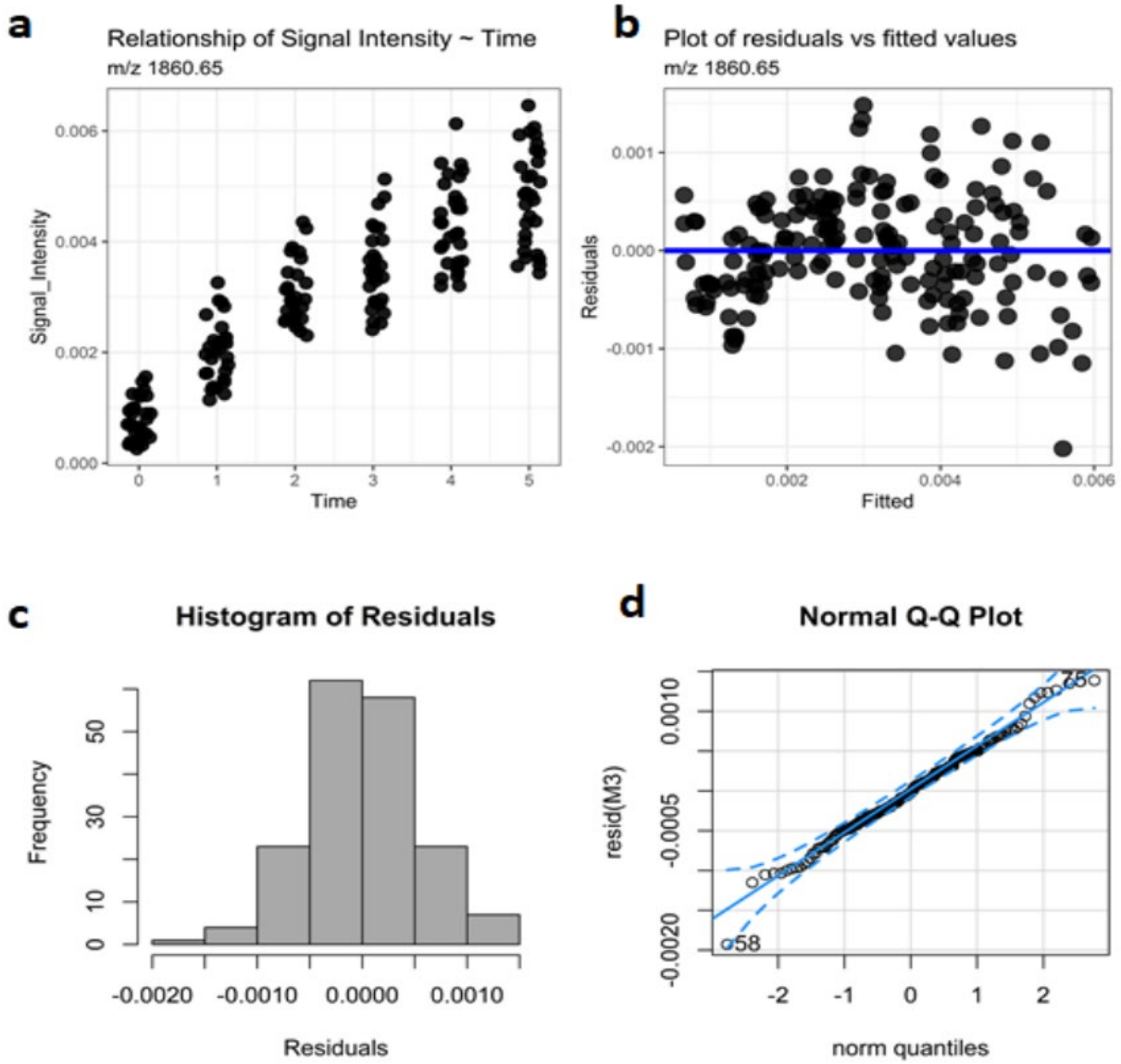


*Figure 4.5 Model diagnostics for multiple linear regression of Signal Intensity ~ Time * Starter * Probiotic. Model diagnostic plots: Raw data plotted by signal intensity ~ time* ***(a)****, a scatter plot of residuals vs fitted values* ***(b)****, a histogram of residuals* ***(c)****, QQ plot of residuals* ***(d)****.*

The model performance was also assessed for a model with low R^2 value. A multiple linear regression model of *m/z* value 1189.79 had an adjusted R^2 of 0.16 and was not significant ($p > 0.05$). The model was visualised, faceting by starter culture and colour-coded according to probiotic culture (Figure 4.6). Unsurprisingly, the line does not run through most of the data; the data modelled here does not follow a linear relationship, so the straight line of best fit is unsurprisingly not very helpful and does not run through most of the data.

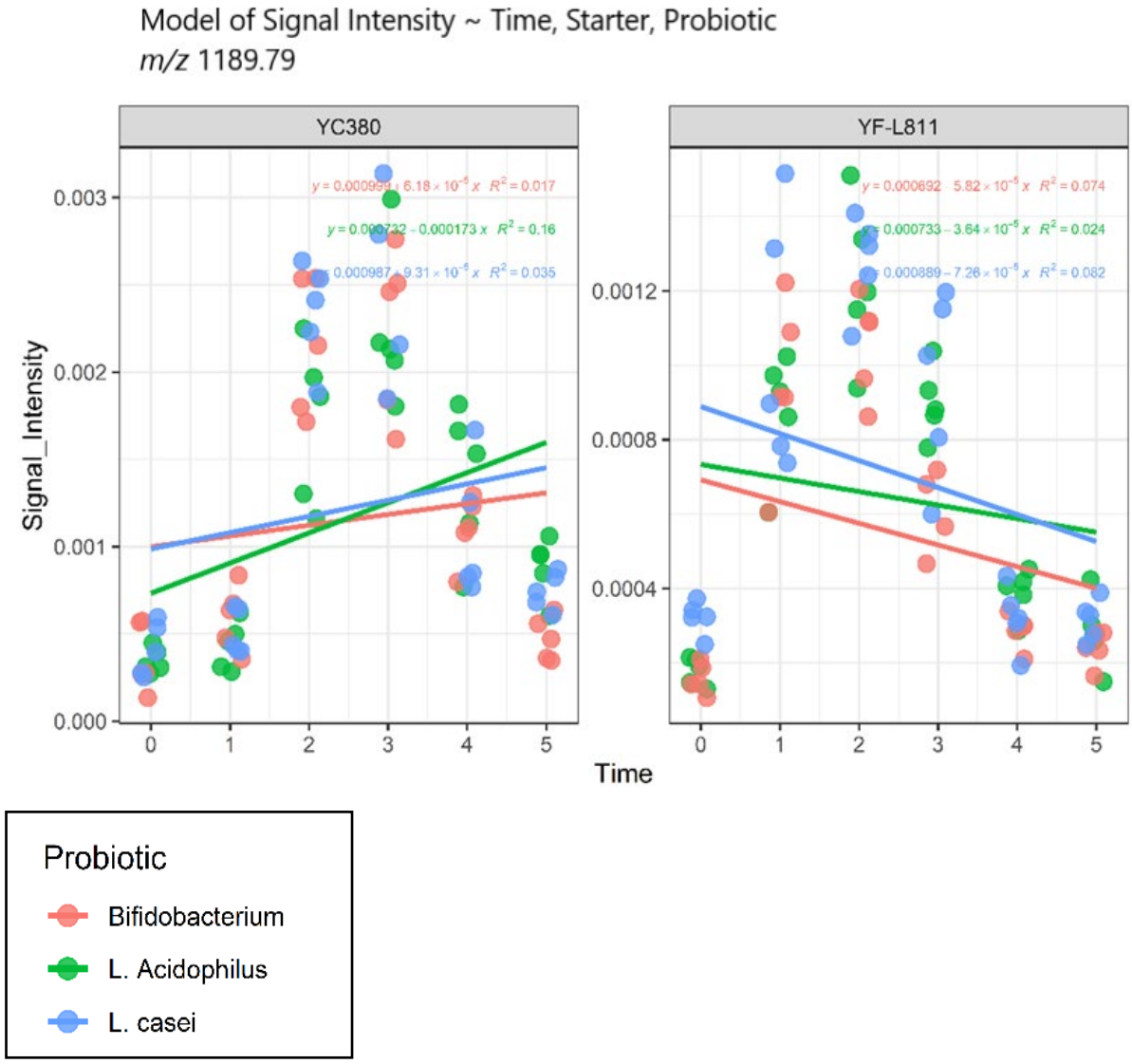


*Figure 4.6 Multiple linear regression model of m/z value 1189.79, fermentation hour, starter culture and probiotic terms were used as predictor terms to predict the response of signal intensity. The signal intensity is along the y-axis, fermentation time is along the x-axis. The plot is faceted by starter culture and coloured by probiotic culture. Legend indicates the corresponding probiotic culture. Formula: lm(Signal_Intensity ~ Time * Starter * Probiotic).*

As expected, the diagnostics, in this case, indicate a poor model performance (Figure 4.7). The data do not indicate linearity (Figure 4.7a). The scatter plot of residuals confirms nonlinearity and shows that the variance of the residuals is increasing with the fitted values (Figure 4.7b); the residual plot shows a discernible pattern, indicating there is a problem with some aspect of the linear model. The residuals show a V-shaped pattern, strongly indicating non-linearity in the data. The histogram of residuals (Figure 4.7c) is skewed to the right (non-normal), and several data points along the QQ plot (Figure 4.7d) are deviating from the reference line indicating it is non-normal.

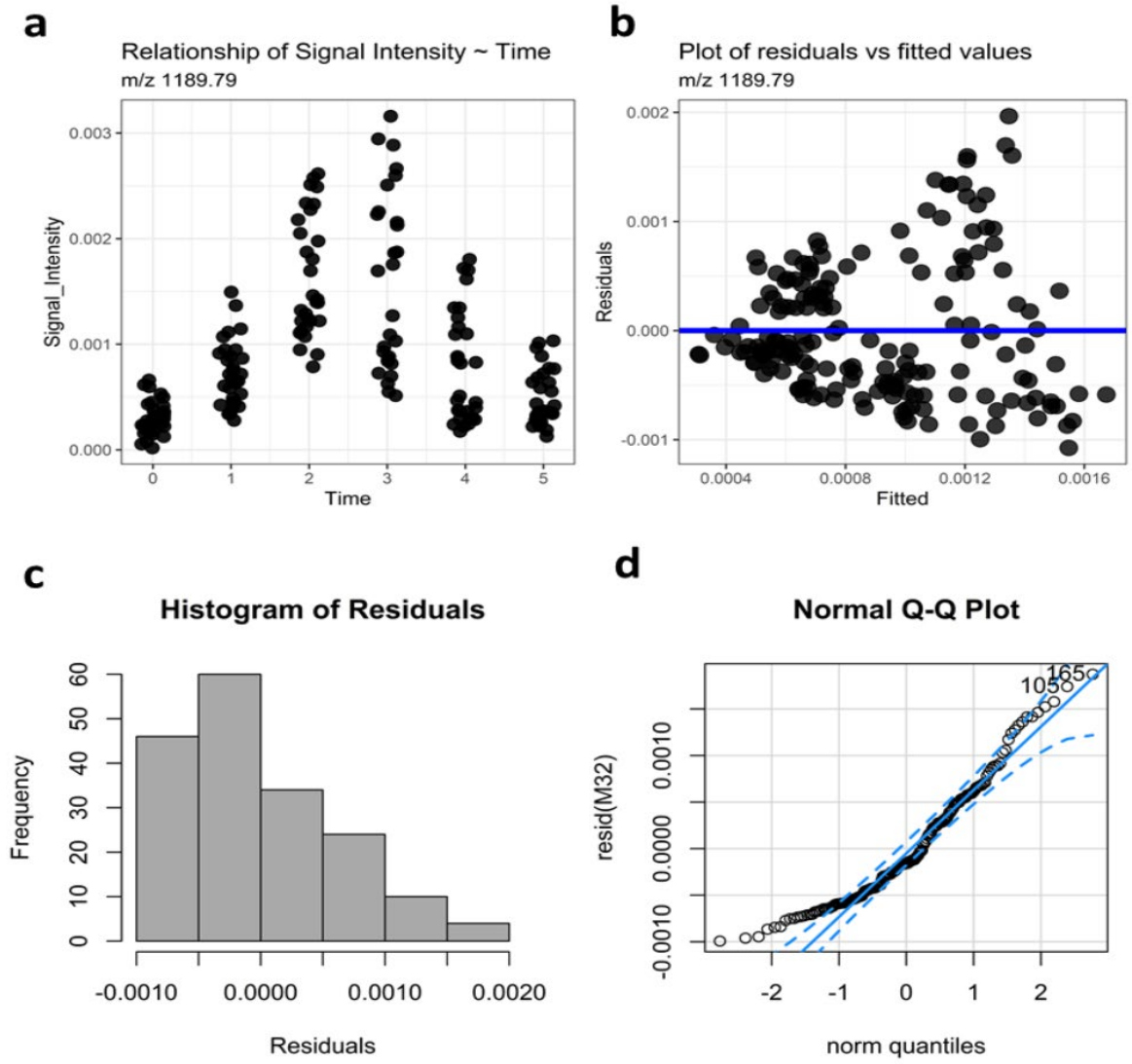


Figure 4.7 Multiple linear regression model of m/z value 1189.79, fermentation time, starter culture and probiotic terms were used as predictor terms to predict the response of signal intensity. The signal intensity is along the y-axis, fermentation time is along the x-axis. The plot is faceted by starter culture and coloured by probiotic culture.

Of note, the signal intensity for this peak (*m/z* 1189.79) increased in the middle of fermentation before decreasing towards the end of fermentation, indicating it is an intermediate peptide. This could be related to bacterial utilisation or the conversion of peptides into other products. The increase in intensity could be explained by the degradation of milk proteins and the subsequent liberation of peptides, causing an increase in the peptides during fermentation. For instance, during the fermentation process, bitter peptides may be liberated from the native protein via hydrolysis. Subsequent hydrolysis can reduce these products to amino acids and derivatives, thereby eliminating the bitter tasting peptides (Newman et al., 2014). As such, this could translate to an increase in the concentration of certain peptides as they are liberated, followed by a decrease as they are used in other processes.

Overall, the multiple linear regression described some of this data well but failed to model all of the peaks adequately. The multiple linear regression was significant in most cases and the R^2 value was reasonable (> 0.5). In many cases, however, the data did not follow a linear trend and inspection of the model diagnostics did not suggest a good fit.

4.3.2.2 Generalised additive modelling

Next, in an attempt to further explore the behaviour of these peaks through time, and to establish whether a non-linear model may be more effective in generalising over this data, a general additive model (GAM) was trialled (James et al., 2013). To compare the performance of a GAM model over a multiple linear regression model, a model was generated using each technique for all peaks (n=143). The models were assessed by comparing the two models by AIC, ANOVA and adjusted R^2 (Table 6).

Table 6 Multiple linear regression model vs general additive model across all peaks. Models were compared by AIC, ANOVA and R^2.

Linear model vs GAM model	
Number of GAM models with lower AIC	138
Number of significant models ($p \leq 0.05$)	136
Number of GAM models with greater Adj R^2	136

Table 6 indicates that the GAM model improved performance for the majority of peaks (138 out of 143), the GAM was significantly different to the linear model in most cases (136 out of 143) and had increased adjusted R^2 for most peaks (136/143).

The adjusted R^2 values were inspected again for the lowest/highest R^2 values as outputted in the linear regression model to understand how the GAM model improved performance (Table 7). In many cases, the R^2 value was increased considerably, for instance in *m/z* 1066.13 which increased from 0.15 to 0.91 in the GAM model. Although, in some cases, it was only marginally improved, i.e., *m/z* 3499.59. In the peaks with the "highest" R^2 value in the linear model, in most cases, the performance remained the same or was improved. For some *m/z* values, the R^2 did marginally decrease in the GAM model, i.e., for *m/z* 2201.87 which decreased from 0.84 to 0.80. A full list of *m/z* ions and corresponding adjusted R^2 for the GAM model are reported in Appendix 18.

Table 7 Comparison of adjusted R^2 values outputted by the multiple linear regression and GAM models. These are organised by lowest/highest based on R^2 from the linear regression model.

Adj R^2	*m/z*	Linear Model	GAM Model
Lowest	1988.17	-0.01	0.59
	1746.62	0.07	0.78
	1761.65	0.11	0.72
	786.7	0.12	0.73
	1066.13	0.15	0.91
	1189.79	0.16	0.88
	3307.15	0.16	0.49
	3499.59	0.18	0.23
	1614.42	0.22	0.89
	1633.88	0.22	0.82
Highest	2395.59	0.82	0.94
	1976.61	0.82	0.88
	2201.87	0.84	0.80
	3346.38	0.85	0.94
	2807.34	0.86	0.94
	1860.65	0.86	0.93
	2120.84	0.87	0.86
	2300.61	0.88	0.88
	2854.27	0.89	0.92
	2316.09	0.89	0.82

The GAM model was visualised for *m/z* 1189.79 to evaluate how well this model improved over the multiple linear regression model (Figure 4.8). For peak *m/z* 1189.79, the R^2 for the GAM model was vastly improved at 0.88 (compared to 0.16 in the linear model). The GAM model appears to fit the trend of the data better than the linear model. In Figure 4.8, it is clear how differently this peak behaves over fermentation for the two different starter cultures. The peak has a higher signal intensity in YC380, compared to YF-L811.

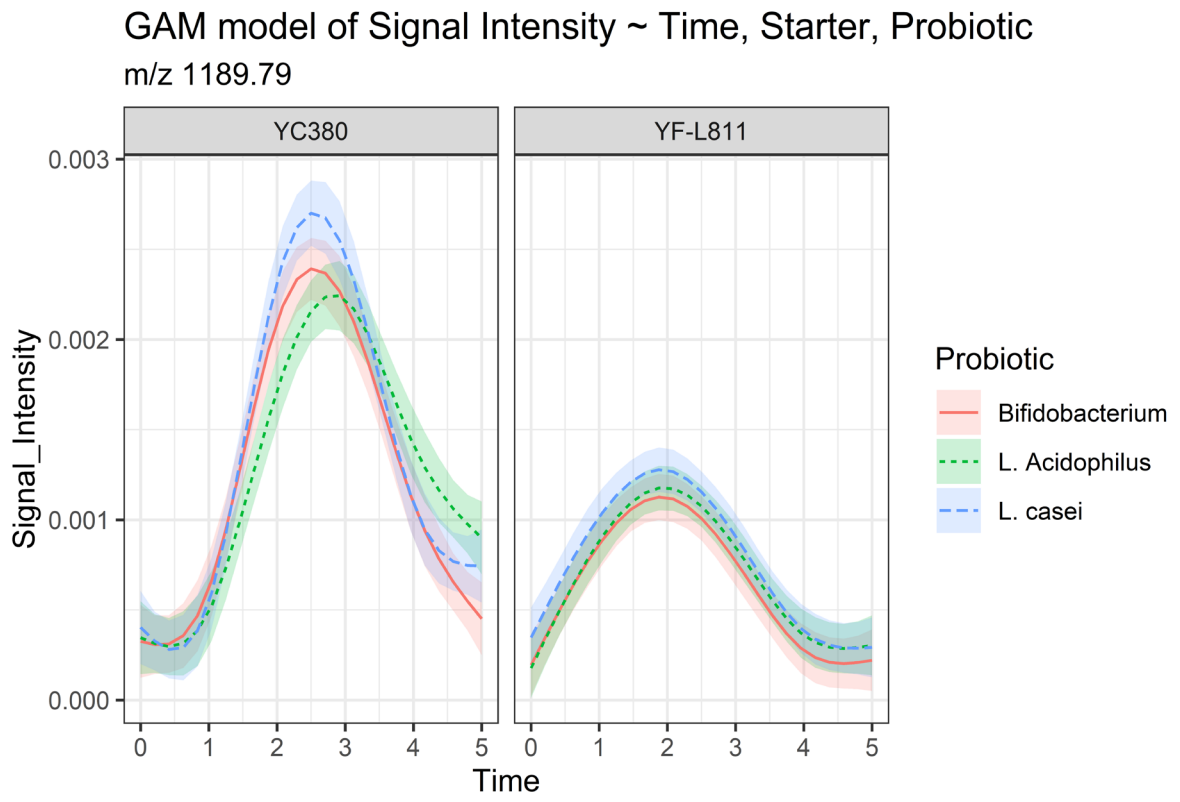


Figure 4.8 GAM model of m/z 1189.79, fermentation time, starter culture and probiotic terms were used as predictor terms to predict the response of signal intensity. The signal intensity is along the y-axis, fermentation time is along the x-axis. The plot is faceted by starter culture and coloured by probiotic culture. Confidence intervals are indicated by the shaded area around the dotted line. Legend indicates the corresponding probiotic culture. GAM formula: bam(Signal_Intensity ~ Starter + Probiotic +s(Time, k = 6) + s(Time, by = interaction(Starter, Probiotic), k = 6)).

A diagnostic check of the GAM model performs a statistical test for patterns in the residuals. The QQ plot in Figure 4.9a shows the points are deviating from the line somewhat for this model. The plot of residuals shows there is an increasing variance in the residuals, which indicate it might not be a great fit (Figure 4.9b). The histogram of residuals looks to be normal (Figure 4.9c). The plot

of response vs fitted values would ideally form a straight line and sit around the 1-to-1 range on the y-axis. The data points in Figure 4.9d trend towards a straight line although some points deviate as the fitted values increase. It is not an ideal fit in all cases but is a vast improvement on the multiple linear regression model.

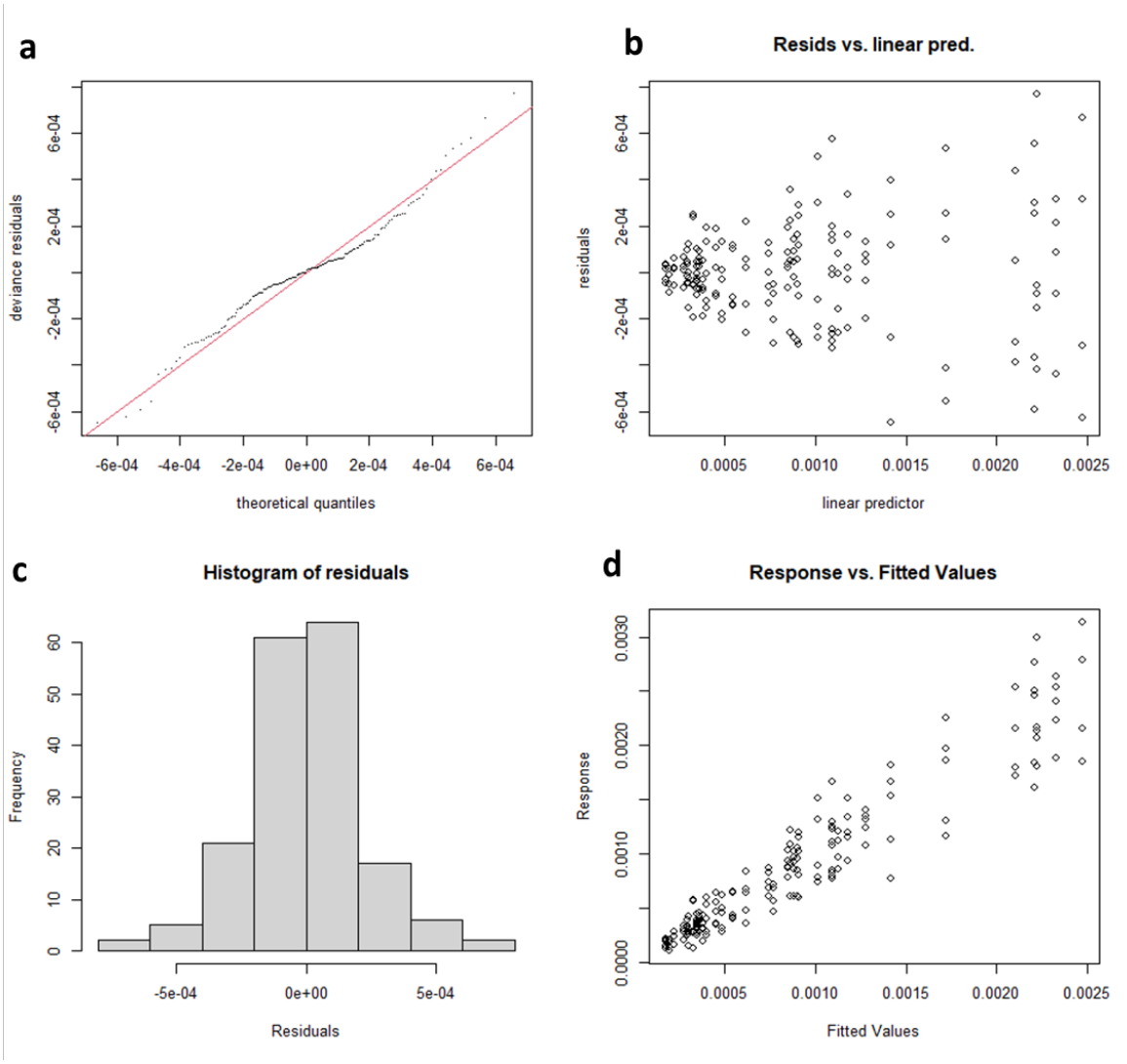


Figure 4.9 Diagnostic plot of GAM model on m/z 1189.79. Diagnostic plots are a QQ plot comparing model residuals to a normal distribution ***(a)****, a plot of the residuals* ***(b)****, a histogram of residuals* ***(c)****, and a plot of response vs fitted values* ***(d)****.*

The GAM models were a considerable improvement over the multiple linear regression models, in most cases. The R^2 was still low for some *m/z* values, i.e., *m/z* 3499.59, which had an R^2 value of 0.18 in the linear model and 0.23 in the GAM model. Inspection of this *m/z* value indicated that the signal intensity was not changing notably over fermentation time, and the signal intensity was also comparatively lower than many of the other peaks detected and was sitting just above zero on the

y-axis (Appendix 22, Appendix 23). It is possible in this case that *m/z* 3499.59 is simply noise and ought to have been eliminated during processing steps, although MS/MS would be required to verify whether it is in fact an actual peptide. During data pre-processing, peak-picking algorithms attempt to find *m/z* values which are likely to be a 'true' peptide, amongst thousands of redundant data points. During this step, there is a compromise in how strict the filtering and peak detection is, for instance increasing the signal-to-noise would eliminate peaks such as this but runs the risk of eliminating peaks that are potentially real peptides of biological interest.

Exploring individual peaks through regression techniques revealed how differently peptides behave throughout fermentation and how this can differ by starter culture. Evidently, this exercise revealed that differences in probiotic culture were impacting the peptide fingerprint less than the starter culture in most instances. Both multiple linear regression and GAM appropriately modelled some of the peaks throughout fermentation time but failed to adequately generalise and fit all of the peaks. Many of the detected peaks inherently will not follow a linear trend as they are likely to correspond to intermediate peptides, whereby the peptide increases during the early stages of fermentation as a result of protein degradation, before being degraded into smaller products. While GAM was a considerable improvement on the linear modelling, it was unable to effectively model all peaks, and would suggest some limitations to this work; the starter cultures cause major differences in the development of peptides during fermentation, resulting in different rates of fermentation and generation of compounds. As such, using a regression model to generalise over the peptide fingerprint obtained from different culture combinations at different fermentation time points may not be feasible.

4.3.3 Classification exercises to predict change in signal intensity

The previous exercise revealed how peptide mass spectra behave during fermentation and how suitable they are for modelling using classical statistical techniques. To further attempt to model the peptide mass spectra over fermentation time, classification techniques were employed, which attempted to predict simply whether a peak would increase or decrease in signal intensity from the

beginning until the end of fermentation. A subset of '40' peaks were retained and were visualised according to their t-score ranking (Figure 4.10). The peaks were organised based on their change in signal intensity through fermentation time, i.e., out of the 143 peaks in the feature matrix, these were the top 40 peaks changing throughout fermentation. Figure 4.10 is organised by the top-ranking peaks along the rows, and subset by fermentation time in the columns. Certain peaks seem to be typical of a particular fermentation time point, i.e., peaks *m/z* 1590.58 and 3031.13 are typical of unfermented milk, whereas 1860.65 is more typical of fermented milk.

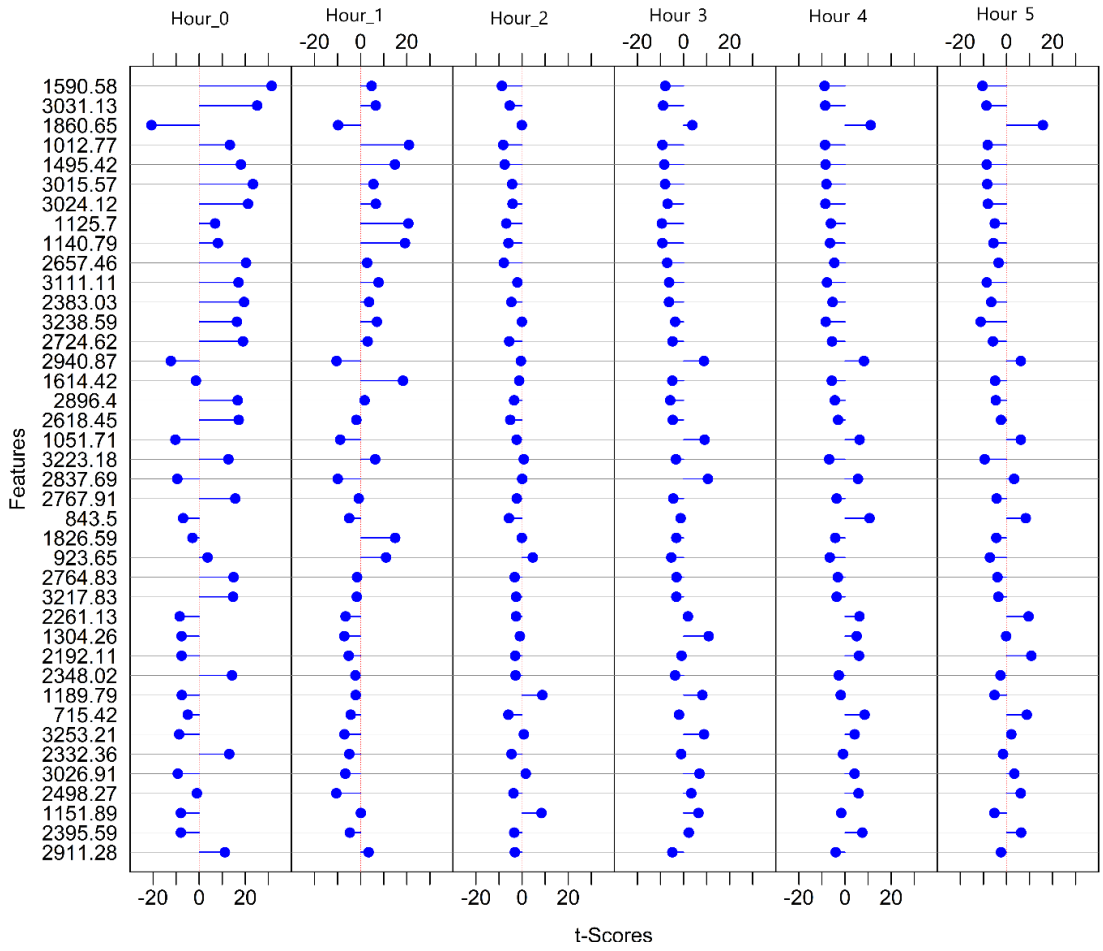


Figure 4.10 Top 40 ranking peaks selected by discriminant analysis. To identify the most important class discriminating peaks, standard t-scores are used. The peaks (rows) in the feature matrix are ranked by their t-score. Fermentation times are along the columns.

The top 40 peaks were then extracted from the original feature matrix, and the peak intensities between early and later ferments were compared and assigned a class label of "increasing" or "decreasing" depending on their change in intensity (Figure 4.11).

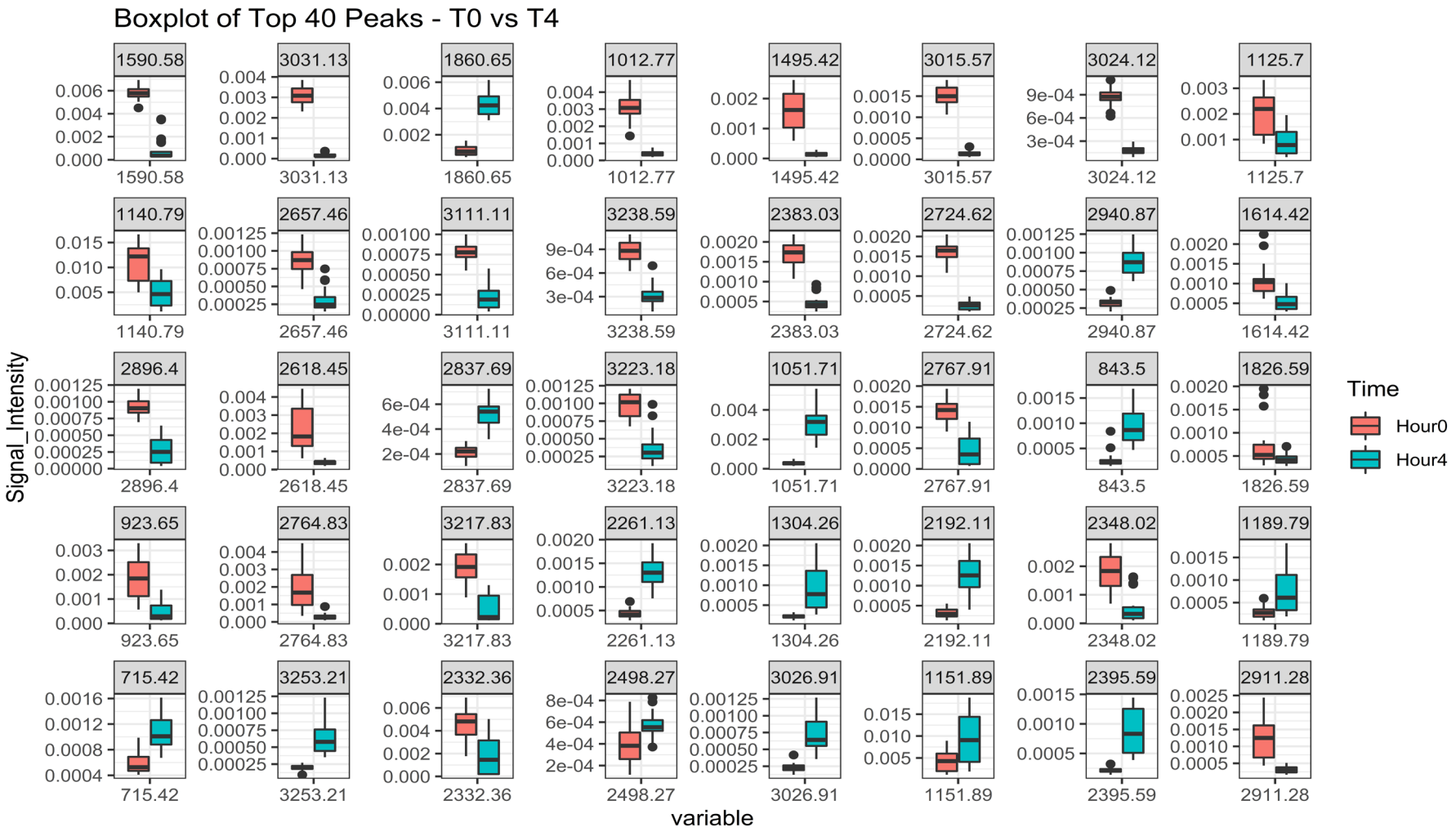


Figure 4.11 Boxplot of signal intensity for each m/z value selected through discriminant analysis; comparison between Time 0 (pink) and Time 4 (blue). 25 out of the 40 m/z values were decreasing between start and later fermentation times. Vertical lines represent minimum and maximum, horizontal line represents the median. Data below the line is in quartile 1, data above the median is quartile 3. The black dots indicate outliers.

Visualising the differences in signal intensity between the time points for each *m/z* value demonstrated that most of the peaks were decreasing in signal intensity over fermentation. As expected, the top-ranked peaks have the biggest changes in intensity between early and later fermentation.

Next, five machine learning techniques were trialled to predict the class label for *m/z* values changing between hour 0 – hour 4. Of the techniques attempted, all models performed reasonably well on both training and test data and were able to predict the class label (increasing or decreasing) with good accuracy (Table 8). For the training data, Matthew's correlation coefficient (MCC) and accuracy are all on-par with one another for most models. Gradient boosting machine (GBM), eXtreme gradient boosting (XGBoost), and deep learning (DL) all had comparable MCC scores, as well as accuracy. Of note is the high MCC, accuracy and F-measure that is returned in the XGBoost and random forest (RF) models, but considerably lower area under the precision-recall (AUCPR) values which reveal a much poorer performance. For this reason, it is worth assessing the models based on more than one criteria, and that a metric more suited to imbalanced data is warranted. The models were ranked by MCC, AUCPR, F-measure then accuracy.

Table 8 Performance metrics for six models on training data: XGBoost, Gradient Boost Machine, Deep Learning, C4.5 and Random Forest. Each model was trained using a 70% subset of the data.

Model	F1	Mean class error	AUC	AUCPR	MCC	Training Accuracy
XGBoost	0.930	0.053	0.989	0.529	0.888	0.947
GBM	0.911	0.073	0.974	0.959	0.859	0.935
DL	0.895	0.089	0.933	0.867	0.840	0.926
C4.5	0.922	0.089	0.911	0.887	0.834	0.930
RF	0.852	0.126	0.916	0.193	0.791	0.901

When applied to test data, each model had a marginally worse performance but still had > 85% accuracy in all cases (Table 9). The GBM and deep learning models all performed well, and

performance metrics were comparatively the same. On test data, the XGBoost model had marginally worse performance based on accuracy (88%) and MCC (0.75). The C4.5 model performed worst based on test data for MCC and accuracy but had a higher AUCPR compared to the random forest and XGBoost.

Table 9 Performance metrics for six models on test data: Deep Learning, Gradient Boost Machine, XGBoost, C4.5 and Random Forest.

Model	F1	Mean class error	AUC	AUCPR	MCC	Test Accuracy
DL	0.855	0.124	0.898	0.906	0.775	0.892
GBM	0.857	0.122	0.915	0.877	0.774	0.892
XGBoost	0.855	0.129	0.898	0.421	0.755	0.883
RF	0.830	0.144	0.892	0.208	0.730	0.871
C4.5	0.856	0.163	0.822	0.809	0.708	0.859

Taking all metrics into account, the deep learning model performed marginally better than other models. This model performed well on the test data and could predict both classes with high accuracy (89%). The prediction of the "Increase" label performed worse with 20% incorrectly labelled, compared to 4.4% for the "Decrease" label (Table 10). This was not unexpected as the number of labels "increasing" was less than those "decreasing".

Table 10 Confusion matrix showing correctly and incorrectly predicted labels for deep learning model on data from hour 0 vs hour 4. For the label 'Decrease', 195 were correctly labelled; 9 were incorrectly labelled as 'Increase'. For the label 'Increase', 109 were correctly labelled; 28 were incorrectly labelled as 'Decrease'.

	Decrease	Increase	Error	Rate	Precision
Decrease	195	9	0.0441	9 / 204	0.87
Increase	28	109	0.2044	28 / 137	0.92
Total	223	118	0.1085	37 / 341	
Recall	0.96	0.8			

Although the deep learning model performed well, these models tend to require large amounts of data for reliable and robust models (Costello and Martin, 2018). The performance of random forest models has also been noted to improve by increasing the size of the dataset (Wang et al., 2020). With that in mind, GBM or XGBoost would be preferable models for this purpose.

Classification techniques were successful in predicting the direction of the signal intensity between start and end fermentation times. These models were implemented using default model parameters only. Although adjustments and fine-tuning of parameters could be made to improve model performances, the model performed adequately using only the default arguments, and so no adjustments were deemed necessary in this case. This exercise indicated that the changes in peptide signal intensity recorded by MALDI-TOF MS could be predicted over fermentation time for these bacterial combinations. This was a simple approach using minimal information from the peptide fingerprint. Modelling of these data could be approached in numerous different ways, using enumerable modelling techniques for both regression and classification which could provide improved or more robust performances. The modelling question could be reframed to predict the "presence" or "absence" of peaks (Timm et al., 2008, Gay et al., 2002). The inclusion of additional physicochemical properties of the peptides might also be useful for building more informative and robust models (Gay et al., 2002). Furthermore, a larger dataset might lead to improved models, as well as more balanced classes.

4.4 Conclusions

This chapter revealed how different starter cultures and probiotics behaved throughout the course of fermentation, and the effects of fermentation on peptide fingerprint. Exploration of peptide fingerprints via multivariate analysis techniques revealed that the primary changes to the peptide fingerprints of the fermented milk occur between early fermented samples vs later ferments, i.e., hour 0-1 vs hour 2 onwards. The starter culture accounted for much of the variation in the peptide fingerprint, indicating that there are different peptides generated by different starter cultures. The

probiotic cultures had less influence on the peptide fingerprints. Probiotic cultures were not used in further analyses because of this.

The primary aim of this chapter was to investigate the potential to predict and model the changes in peptide signal intensity over fermentation time. This was achieved using regression techniques, with some success using both linear and nonlinear techniques. Modelling the peptide fingerprints of fermented milk products can be used as a means to monitor, predict, and screen potentially important compounds during fermentation of milk, which could be used for targeting desirable compounds that impart favourable sensory characteristics. It has been established that certain *m/z* values are typical of a product that is at its end-fermentation time; such *m/z* values could be used as markers to signal fermentation end-points. This is beneficial for the screening of new products using various starter strains which may have variable lengths of fermentation time. Furthermore, classification techniques proved promising for predicting the general direction of the signal intensity over fermentation time. This was a simple model which proved effective at generally predicting whether the signal intensity would increase or decrease throughout fermentation. This could be extended to predict the presence/absence of a peak and could include additional parameters or peptide characteristics to build a more robust and accurate model.

Chapter 5 Consumer evaluation and peptide fingerprinting of milk throughout fermentation

Summary

Consumer perception of important sensory attributes in fermented milk was investigated. Fermented milk was prepared using two different commercial starter cultures, YF-L811 and YC380, and sampled throughout fermentation. Samples were evaluated by consumers in a multiple paired comparison test to understand the changes in bitterness and flavour intensity throughout fermentation for each starter culture. Consumers perceived significant differences in these attributes between pairs of samples fermented for different lengths of time. The intensity of bitterness and flavour increased significantly with fermentation time in fermented milk prepared from the different starter cultures. In parallel with this, peptide fingerprinting with MALDI-TOF MS was performed on replicate samples. Significant correlations were reported between the peptide fingerprints and the consumer rankings for the sensory attributes in each product. Machine learning techniques were employed to predict the consumer responses. XGBoost models performed reasonably well for predicting the consumer responses.

5.1 Introduction

The quality and acceptability of fermented milk depends on consumers perception of sensory characteristics, including the taste, flavour, and texture of the product (Reineccius, 2006, Briand and Salles, 2016). The sensorial quality of fermented dairy products is mostly thought to be derived from lactic acid and aroma compounds, although it is said to be partially dependent on the proteolytic activity, as peptides and amino acids generated via proteolysis can have a direct impact on flavour and taste (Ahtesh et al., 2018). Peptides can directly contribute to taste in fermented foods by producing bitter, umami and kokumi sensations, and are also critical to flavour formation by acting as precursors for flavour-producing enzymatic reactions (Tamime and Robinson, 1999, van Kranenburg et al., 2002, Sanlier et al., 2019). Consumers typically have a negative perception

of bitterness in fermented dairy foods, and it is seen as a major defect by food manufacturers (Zhao et al., 2016). Kokumi peptides are related to continuity[1] and mouthfulness[2] (Ueda et al., 1997), and can contribute to the taste profile of a product by enhancing other taste qualities (Yang et al., 2019). Umami taste is viewed as pleasant and desirable by consumers, and umami peptides have been reported to contribute to the overall complexity of flavour in fermented dairy products (Roudot-Algaron et al., 1994). Understanding how the peptide fingerprint of a product relates to consumer acceptability can contribute to developing strategies to mitigate deleterious flavour characteristics and enhance desirable flavours. Many of the current approaches used in the literature to evaluate flavour and taste in fermented products are costly and time-consuming. The current approach aims to enable a simple, rapid, high-throughput and consumer-led means to correlate important sensory attributes with instrumental analysis. By better understanding the development of desirable and off-flavour characteristics during the fermentation of milk, this information can potentially aid in the selection and screening of fermented milk products with enhanced sensory properties.

The previous two chapters in this thesis explored rapid instrumental techniques to fingerprint and discriminate between different fermented milk products, as well as to couple such techniques with machine learning and predictive modelling. Six bacterial combinations were used for fermented milk preparation in Chapter 4; differences in the peptide fingerprints were found not to be primarily driven by the probiotic cultures used. In the current chapter, fermented milk was prepared using only the starter cultures, YF-L811 and YC380, and were not prepared in combination with the probiotic cultures used in Chapter 4. The objective of this chapter was to explore how consumer perception of key sensory attributes in different fermented milk changes throughout fermentation and to develop an approach to link consumer response with instrumental analysis.

[1] Continuity: the prolonged taste intensity after consumption.
[2] Mouthfulness: reinforcement of the taste sensation throughout the mouth, not limited to just the tongue.

5.2 Materials and methods

5.2.1 Consumer study

5.2.1.1 Fermented milk preparation

Materials: Skim milk powder was obtained from Fonterra (Fonterra Co-operative Group, New Zealand). Commercial freeze-dried starter cultures YF-L811 and YC380, both containing different strains of *Streptococcus thermophilus* and *Lactobacillus delbrueckii* subsp. *bulgaricus*, were purchased from Christian Hansen (Hørsholm, Denmark) and stored at −20 °C until use.

Milk reconstitution: Sample preparation took place at the Massey Food Pilot Plant, Palmerston North over two separate days for each of the two starter cultures used. The evening prior to fermentation, approximately 60 L of skim milk powder (Fonterra Co-operative Group, New Zealand) was reconstituted in water (8.5% w/v; 3% protein). The skim milk powder was slowly added to water and mixed using an overhead stirrer for 30 minutes at 700 rpm (Figure 5.1). The reconstituted milk was stored in a chiller overnight, at 4 °C.

Figure 5.1 Fermented milk for the consumer study prepared at the Massey University Food Pilot Plant. Milk was reconstituted by mixing skim milk powder and water using an overhead stirrer (a). The milk was heat-treated using a steam kettle to rapidly heat the milk to 85 °C (b). The milk was then held in a water-jacketed kettle, set to 85 °C for 30 mins. Fermentation took place in the water-jacketed kettle (c).

Heat treatment: The milk was first heat treated in batches in a steam-jacket pan (Mercer Stainless, New Zealand), to rapidly bring the milk to 85 °C (Figure 5.1). The milk was then placed in a water-jacketed kettle (Mercer Stainless, New Zealand), and held at 85 °C for 30 minutes. Following the

heat-treatment, the milk was cooled to fermentation temperature (43 °C). This was achieved by flushing out the hot water in the kettle and re-filling with cold water.

Inoculation: The starter culture was removed from the freezer 15 minutes prior to mixing. The inoculation mix was prepared by adding 6.6 g of starter culture to sterile milk and mixing by hand for approximately 15 minutes, to ensure the mix was fully dissolved. Once the milk had cooled to fermentation temperature, the starter culture mix was added and stirred gently using an overhead stirrer (400 rpm) for approximately 15 minutes (Figure 5.1).

Fermentation: The inoculated milk was fermented at 43 °C for 5 hours (until pH ≤ 4.5) in the water-jacketed kettle (Figure 5.1). As described in section 4.2.1, fermented milk was withdrawn every hour before fermentation had commenced (inoculated milk), up until after fermentation was completed (pH ≤ 4.5). The fermented milk samples were removed via the tap of the water-jacketed kettle. Approximately 10 L was removed per fermentation time point (n=6) and placed in sterile containers. Samples were removed before (hour 0), during and after fermentation was complete. The containers were immediately covered and placed on ice water to halt the fermentation. At each hour of fermentation (hour 0 – 5), samples were withdrawn as such. Containers containing the samples were placed in a cooler (4 °C) for 3-4 days to enable microbial safety checks.

All equipment used was hand-washed, followed by running in a dishwasher at a high temperature (90 °C), and finally using alcohol to sanitise. The tank used for fermentation was hand washed using detergent and water, followed by sterilisation with alcohol.

pH measurements: pH was recorded in triplicate using a pH meter in order to track the progress of the fermentation (Orion Versa Star Pro Benchtop pH Meter, Thermo Scientific, USA). The pH meter was calibrated using buffers of 4.01 and 7.01 (Thermo Scientific, USA). To measure the pH, approximately 10 mL of sample was placed in a small beaker. The pH electrode was placed directly into the milk sample. The pH reading was allowed to stabilise for a few seconds before the reading was taken. Between measurements of different samples, the pH electrode was rinsed using distilled

water. pH measurements were recorded at room temperature. Samples were discarded after the pH reading was recorded.

Microbial testing: Each sample was tested for coliforms before being deemed safe for consumption. Microbial testing was carried out independently by the team at the Microbiology Laboratory at Massey University.

5.2.1.2 Multiple paired comparison / 2-alternative forced choice test

Participants: Four consumer sessions were carried out to evaluate bitterness and flavour intensity for each of the two starter cultures. The participants were recruited from Massey's Food experience and sensory testing (Feast) Lab's consumer database. The criteria for selection were availability to attend the sessions, no dairy allergies, and a frequent consumer of dairy products. A minimum of 40 participants were recruited for each session based on the chosen statistical parameters of (α risk[1] = 0.05, β risk[2] = 0.3 and P_{max}[3] 70% for 40 assessors (Rogers, 2017)). The same participants returned for each pair of sessions (i.e., the same participants evaluated bitterness and flavour intensity in each starter culture).

Sample preparation: The samples used for the consumer test were held in the chiller, at 4 °C following fermentation; samples evaluated for bitterness were held for three days following production, and samples evaluated for flavour intensity were held for four days. On the morning of each session, each fermented milk sample was mixed using an overhead stirrer (700 rpm for 60 seconds). The sample was then poured in to 1 L sterile Durant bottles. The bottles were gently rotated up and down five times before being poured into small plastic cups to ensure the milk was homogenously mixed. Approximately 15 mL of each sample was poured into each cup. The cups were covered with plastic lids and were kept chilled in a fridge until ready for evaluation.

[1] α risk: probability of concluding a perceivable difference exists when there is none.
[2] β risk: the probability of concluding that there is no perceivable difference when there is an actual difference.
[3] P_{max}: the meaningful difference from equal intensity (50:50 split) that the test should be able to detect, i.e., P_{max} of 70% indicates a 70:30 split as a meaningful difference from equal intensity (Rogers, 2017).

Experimental design: In order to compare each fermentation time-point with all others, a 2-alternative forced-choice (2-AFC), or multiple paired comparison test, was employed. A William's Latin Square, partially balanced design was generated using Compusense (version 20, Compusense Inc., Guelph, Canada). Each sample (n=6 samples) was assigned five random 3-digit codes, for a total of 30 codes. Each sample was assigned a new random code for every pairing it was in to avoid biases or memorisation of codes. The samples were presented as 15 pairs. The codes assigned to each sample type were alternated every second day, i.e., sessions 1 and 3 had the same sample codes, and sessions 2 and 4 had the same sample codes. This was to ensure that the consumers did not memorise codes or associate them with a specific sample.

The cups for each sample type were labelled with their corresponding random code. The position of each sample and pairing was randomised for every participant, according to the experimental design. Table 11 displays an example layout for the samples presented to three different participants, where the random (blind code) and the actual sample (fermentation time) that it corresponds to. The participants received the sample cups labelled with the random, blind code.

Test instructions: Prior to the test, the participants were provided with instructions for carrying out the test. Participants were provided with brief instructions on what to consider when evaluating the products for each characteristic. For flavour intensity, the participants were advised to imagine the contrast between a food full of flavour, such as a blue cheese, compared to a mild-flavoured food, such as a mild cheddar cheese. A participant code was generated for each panellist, and they were allocated a sample set that would appear on screen during the test, in the order determined by the experimental design. The test was created to display the random pairings of samples, as above. For each session, the participants were presented with an iPad instructing them to consume each pair of samples from left to right, and in the order presented on the screen. For each pair of samples, the participants were asked to select the most bitter-tasting sample, and in a separate session, the most intensely flavoured sample. This was repeated for each starter culture. The test was forced-choice, where participants could not select "no difference". The participants were prompted to take at least

a 10-second break between each pairing and to consume water and plain crackers to cleanse the palate and eliminate carryover. The participants evaluated the samples in ten individual partitioned sensory booths. The sessions were conducted under red-lighting to minimise visual differences (Figure 5.2). The participants were presented with three trays in total, each containing five pairs of samples (Figure 5.2). The participants were given approximately 5 minutes break between each tray. The questionnaire design and data collection were conducted in Compusense.

Table 11 Random, blind codes assigned to each sample type and the actual sample it corresponds to. The positioning of each sample is partially balanced and randomised per participant.

	1- Blind Code		1- Actual Sample		2 - Blind Code		2 - Actual Sample		3 - Blind Code		3 - Actual Sample	
Pair 1	671	156	4	0	475	715	2	0	904	623	2	3
Pair 2	315	963	3	1	281	672	3	0	632	143	2	1
Pair 3	735	832	4	5	943	893	5	0	154	591	5	1
Pair 4	632	143	2	1	458	593	1	0	315	963	3	1
Pair 5	823	124	4	2	259	198	1	4	723	982	5	3
Pair 6	904	623	2	3	361	527	2	5	671	156	4	0
Pair 7	204	415	4	3	415	204	3	4	527	361	5	2
Pair 8	154	591	5	1	982	723	3	5	735	832	4	5
Pair 9	198	259	4	1	124	823	2	4	593	458	0	1
Pair 10	723	982	5	3	591	154	1	5	823	124	4	2
Pair 11	893	943	0	5	832	735	5	4	672	281	0	3
Pair 12	527	361	5	2	623	904	3	2	204	415	4	3
Pair 13	715	475	0	2	156	671	0	4	715	475	0	2
Pair 14	593	458	0	1	143	632	1	2	198	259	4	1
Pair 15	672	281	0	3	963	315	1	3	893	943	0	5

Figure 5.2 The sessions took place in individual partitioned sensory booths. Participants were presented with pairs of samples on trays, with random codes assigned to each sample. The presentation of the samples was randomised for each participant. Participants were presented with a screen on an iPad with instructions to consume samples and evaluate the sample for bitterness, or flavour intensity.

The paired comparison test was performed in four sessions across four days. For each session, the testing took place throughout the day. On each day, the first set of consumers began evaluating samples at 10:00am and the final consumers began evaluating samples at 16:00pm.

Session 1: participants evaluated bitterness in pairs of samples prepared using starter culture YF-L811 (n = 41 participants). Session 2: participants evaluated flavour intensity (e.g., umami/kokumi sensation) in pairs of samples prepared using YF-L811 (n = 41 participants). Session 3: participants evaluated bitterness in pairs of samples prepared using YC380 (n = 43 participants). Session 4: participants evaluated flavour intensity in pairs of samples prepared using YC380 (n = 43 participants). Four participants were removed from the overall results as they missed one of the sets of sessions.

5.2.1.3 Data and statistical analysis

The data were analysed using a Friedman analysis (Equation 1) as described in Meilgaard et al. (1999) using Microsoft Excel (version 16.0). This test allows for a global comparison across the samples. Data were first converted to ranks by calculating a ranking score. The ranking score is equal to = (Number of times the product is selected x 2) + (Number of times the product is not selected x 1).

Equation 1 Friedman's T is computed using the following formula, where p=total panellists, t=treatments and R=rank sum:

$$T = \left(\frac{4}{pt}\right)\sum_{i=1}^{t} R_i^2 - 9p(t-1)^2$$

Where the Friedman test was significant, a multiple comparison was performed using Tukey's Honestly Significant Difference (HSD; Equation 2) test to determine the significance of all pairwise comparisons. The significance in the Friedman test is determined by comparing the test statistic, Friedman's *T,* with the critical value of χ^2, at α=0.05, with (t – 1) degrees of freedom, where t is the number of treatments (Meilgaard et al., 1999). The rank sum scale of each attribute was calculated and plotted in R (version 4.0.3), using base R functions.

Equation 2 Tukey's HSD is computed as follows:

$$HSD = q_{\alpha,t,\infty}\sqrt{\frac{pt}{4}}$$

Values of d` (d-prime) were calculated automatically in Compusense. The value of d` is used as a measure of sensory difference between signal and noise, derived from signal detection theory (Lawless and Heyman, 1999). It is an indication of the difference perceived in the samples, where smaller values are small perceptual differences (or noise), and larger values represent a greater difference in the perceptual difference. A d` of 1 can be used as a threshold to determine a noticeable difference for consumers (O' Mahony and Rousseau, 2003). d` was calculated in addition to performing a Friedman analysis in order to confirm the interpretation of results.

5.2.2 Matrix-assisted laser desorption/ionisation – time of flight Mass Spectrometry (MALDI – TOF MS)

5.2.2.1 Chemicals and reagents

MALDI-TOF MS matrices and peptide standards were purchased from Bruker Daltonics (Bremen, Germany). Acetonitrile and ethanol were liquid chromatography-mass spectrometry (LC-MS) grade solvents from Fluka Analytical (Merck, Kenilworth, NJ, USA). Optima LC-MS grade water and trifluoroacetic acid (TFA) were obtained from Fisher Scientific (Hampton, NH, USA). Skim milk powder was obtained from Fonterra (Fonterra Co-operative Group, New Zealand). Freeze-dried DVS yoghurt starter cultures YF-L811 and YC380, both containing different strains of *Streptococcus thermophilus* and *Lactobacillus delbrueckii* subsp. *bulgaricus* were purchased from Christian Hansen (Hørsholm, Denmark). The cultures were stored at −20 °C until use.

5.2.2.2 Experimental design

A randomised block design was generated using CycDesignN (version 6.0) using the following parameters: a resolvable row-column design was generated with 10 rows and 6 columns. Three factors (starter culture, fermentation time, replicate) with 2 * 6 * 5 levels, for a total of 60 samples (Table 12). Each fermented milk time point was prepared in a separate falcon tube and removed from the water bath, accordingly, as outlined in section 4.2.3.

Table 12 Experimental design; position one in the three-digit code represents the starter culture, followed by fermentation time and replicate.

(2,2,4)	(1,5,1)	(1,1,3)	(2,4,3)	(1,3,6)	(2,3,2)
(1,1,5)	(2,4,1)	(2,3,4)	(1,5,3)	(2,1,2)	(1,2,1)
(1,4,2)	(2,2,2)	(2,5,6)	(1,2,5)	(2,3,3)	(1,1,6)
(2,4,5)	(1,3,3)	(2,2,3)	(2,1,6)	(1,4,1)	(1,5,4)
(1,2,3)	(2,3,6)	(2,4,2)	(2,5,4)	(1,1,4)	(1,3,5)
(2,3,1)	(1,1,2)	(1,2,4)	(1,1,1)	(2,4,4)	(2,5,3)
(2,1,3)	(1,2,6)	(1,4,5)	(1,3,2)	(2,5,1)	(2,2,5)
(1,5,6)	(1,4,4)	(2,1,5)	(2,2,1)	(1,2,2)	(2,4,6)
(2,5,2)	(2,1,4)	(1,3,1)	(2,3,5)	(1,5,5)	(1,4,3)
(1,3,4)	(2,5,5)	(1,5,2)	(1,4,6)	(2,2,6)	(2,1,1)

5.2.2.3 Fermented milk preparation

Additional fermented milk replicates were prepared in the laboratory in Lincoln for MS analysis. The fermented milk samples were prepared using the same milk powder and starter cultures as those used for the consumer study. The fermented milk was prepared as outlined in section 4.2.3.

pH measurements: The pH of each sample was recorded at room temperature using a pH meter (Hanna Instruments, USA). The pH meter was calibrated using buffers of 4.01 and 7.01. Three measurements were recorded for each sample, in the same manner as the consumer study.

5.2.2.4 MALDI-TOF MS sample preparation

MALDI-TOF MS analysis was carried out as outlined in section 4.2.4 and section 4.2.5. Pre-processing of MALDI-TOF data was carried out as described in section 4.2.6. MALDI-TOF MS was carried out on both the samples prepared during the consumer trial, and the additional replicates prepared in section 5.2.2.3.

5.2.2.5 Data analysis and modelling

Plots of pH values were created using Excel. To determine how different the peptide fingerprints were for the two sample sets (i.e., samples prepared for consumer and instrumental analysis), a Multivariate INTegrative Partial Least Squares Discriminant Analysis (MINT-PLS-DA) was employed. This is an emerging technique that enables combining and integrating independent studies measuring the same variable (peptides). This allows us to gain some insights into the differences that may arise due to differences in geographical sites or different sampling times (Rohart et al., 2017) in order to determine how different the peptide fingerprints were for the different studies. This analysis was performed separately for each starter culture. A sparse MINT-PLS-DA was performed on the feature matrices using the R package 'mixOmics' (Rohart et al., 2017), including samples prepared during the consumer study. A 10-fold cross validation was performed for each dataset to optimise the tuning parameters. For the YF-L811 model, three

components were retained with 40, 10, and 6 variables on these components, respectively. For the YC380 model, two components were retained with 80 and 20 variables, respectively.

Principal component analysis (PCA) was performed on the processed feature matrices using the R packages 'factoextra' (Kassambara and Mundt, 2020) and 'factominer' (Le et al., 2008) to visualise the similarities in the peptide fingerprints of each fermentation time point, using only data obtained from the instrumental analysis. PCA captures and reports the largest variation between samples and creates new variables that capture this variation in the dataset, allowing a summary of the data using a few relevant vectors. The analysis was performed separately for each starter culture. A Spearman correlation was performed using the package 'GGally' (Schloerke et al., 2021). An XGBoost regression model was applied to predict the consumer responses for each attribute, using the top principal components as predictors. XGBoost is a popular gradient boosting algorithm that has been used to much success to solve a wide range of data science problems (Chen et al., 2021). Extreme Gradient Boosting (XGBoost) regression modelling was carried out using the R package 'H_2O' (LeDell et al., 2020). The ranks calculated from the consumer responses for each attribute were used as response variable and extracted principal components were predictor variables. Default parameters were used to generate each model. Model performance was evaluated using the following: Root Mean Squared Error (RMSE) evaluates how well a model can predict a value. Mean Square Error (MSE) measures the average of the squares of the errors or deviations. For both MSE and RMSE, the smaller the values the better the model performance. RMSE and MSE, however, can be sensitive to outliers; the Mean Absolute Error (MAE) is a more robust metric to outliers. MAE is the average of the absolute errors; the units are the same as the predicted target, which allows for easy interpretation. R^2 is a popular metric to evaluate regression model performance. This provides a measure of the amount of variance explained by the model (Kuhn and Johnson, 2013).

5.3 Results and discussion

Fermented milk, prepared from two different starter cultures, was sampled throughout fermentation, and presented to consumers in a multiple paired comparison test to assess the perceived change in bitterness and flavour in these products. Several defects were evident in the samples prepared for the consumer study. Samples from hour 3 (YF-L811) had considerable whey separation after production (Figure 5.3). There are numerous possibilities for this: high incubation temperatures (Lee and Lucey, 2003) or disruption of the gel network during sample removal (Lee and Lucey, 2010) may have led to the defects in this sample, although they were not observed in other samples. Consumers have a negative perception of whey separation (Lee and Lucey, 2010); however, this defect is visual and is not known to directly affect taste or flavour. Samples collected at hour 2 (YC380) had an unpleasant pungent odour. Neither of these defects appeared to influence the consumer responses. These defects are not unexpected, as most of the samples were semi-fermented making them defective by definition.

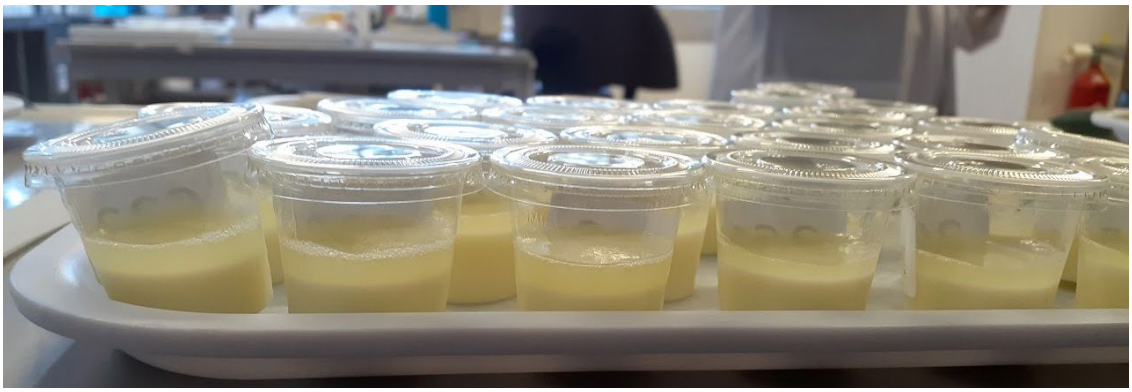


Figure 5.3 Samples collected at hour 3 for YF-L811 had whey separation. This defect was found only in this sample.

The samples were tested for microbial safety prior to the consumer test (Appendix 26, 27). Samples were prepared several days in advance of the consumer sessions to allow for microbial testing. Preliminary trials showed contamination with coliforms and *Escherichia coli* (Appendix 28): this was determined to be due to the tap from which the samples were withdrawn from the tank, as well as a homogeniser for mixing the samples. To address this, additional cleaning and sterilisation steps

were introduced to ensure that all of the equipment was adequately sterilised before sample preparation, during fermentation and during sample withdrawal. The protocol was adjusted so that samples could be prepared in bulk and were only sampled by participants following a clear microbial test.

Post-acidification occurred during storage resulting in a decrease in the pH for all samples (Figure 5.4a, b). Post-acidification can be explained by the continued metabolic activity of the lactic acid bacteria during cooling and refrigerated storage (at 4°C) (Beal et al., 1999). Fermented milk is more stable at a lower pH; high final pH values are known to lead to greater rates of post-acidification (Beal et al., 1999). The rate of post-acidification is consistent with what is reported in the product information for each starter culture and is similar to other studies sampling yoghurts with different pH values throughout storage (Beal et al., 1999, Papadimitriou et al., 2007). *Lb. delbrueckii* subsp. *bulgaricus* is considered to be responsible for post-acidification and for the generation of any bitter peptides during post-acidification (Moller et al., 2007).

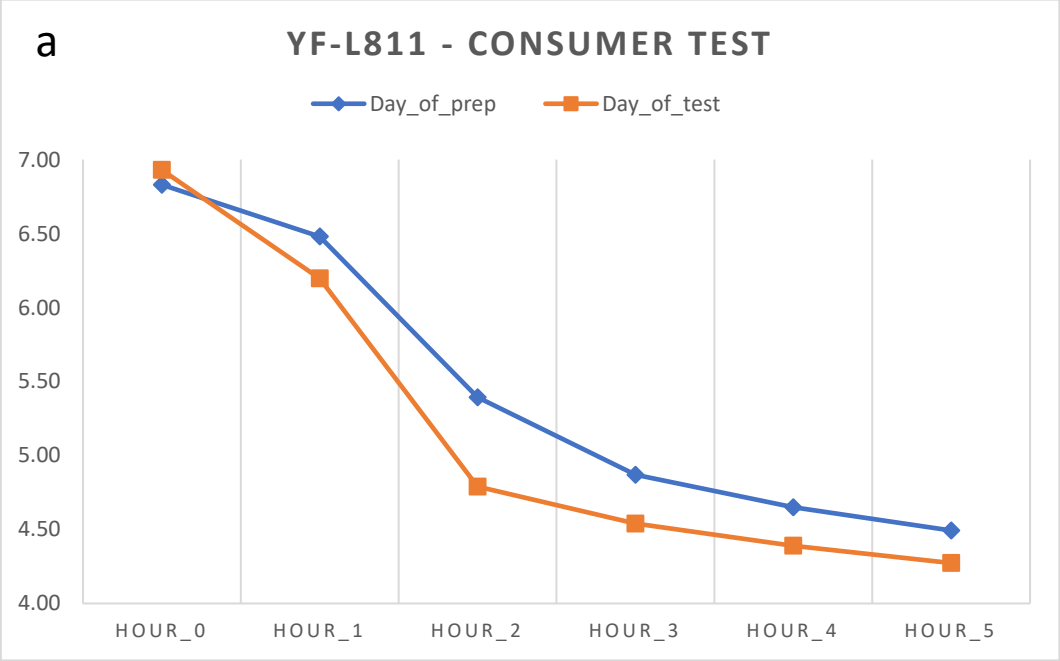
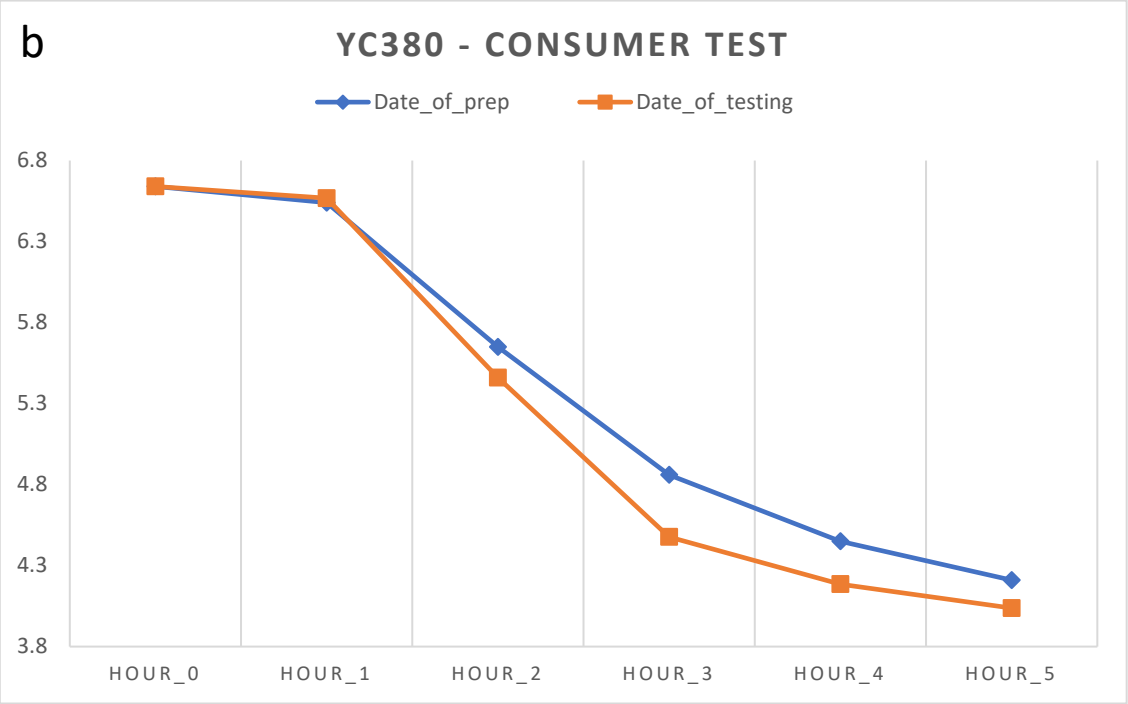


Figure 5.4 mean pH (n=3) recorded on day of preparation and after 4 days in storage for samples prepared using the starter culture YF-L811 (a), and pH recorded on day of preparation and after 4 days in storage for samples prepared using the starter culture YC380 (b). Post-acidification occurred while the samples were in storage. Average of three measurements for each sample.

Figure 5.4 continued



5.3.1 Consumer evaluation throughout fermentation of milk

The bitter taste perceived by consumers increased with an increased fermentation time for both YF-L811 and YC380 (Figure 5.5). The arbitrary numbers on the line scale represent the rank position of the samples within the sample set, and different lettering indicates a significant difference (at α=0.05). In YF-L811, there was an increase in the perceived bitterness after two hours of fermentation, though this did not change significantly for the remainder of the fermentation time.

For YC380, after two hours there was also a significant change in the perceived bitterness. This bitterness continued to increase but did not change significantly after three hours of fermentation. This suggested that the bitter taste did not significantly change in either starter culture after a certain period of fermentation. This may be related to the breakdown of potentially bitter peptides at early stages of fermentation.

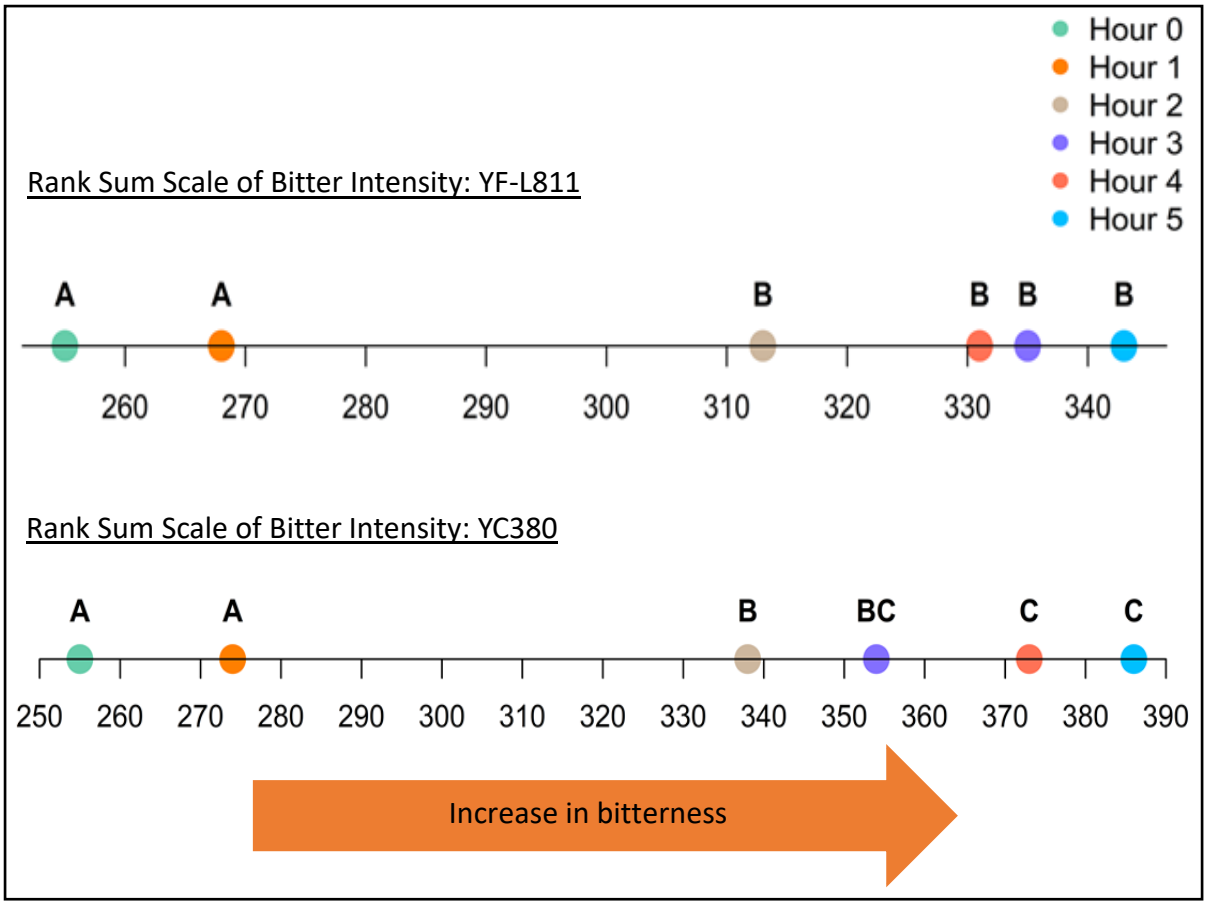


Figure 5.5 Rank sum scale of bitterness for YF-L811 (n=41 participants) and YC380 (n=43 participants). The arbitrary numbers correspond to the samples rank position within the sample set. Different lettering above the samples indicates a significant increase in bitterness (at α=0.05), determined by Tukey's HSD test. The samples prepared for each starter culture increase in perceived bitterness, with an increased fermentation time.

The d` values were obtained for each pair of samples for bitter intensity in each fermented milk type (Figure 5.6a, b). For bitter intensity in samples prepared using YF-L811, the d` ranged from 0.13 (for pairs at hours 5 - 3, and hours 4 - 3) to 1.35 (pair hours 5 - 0). The d` was 0.22 for the pair comparing hours 1 - 0, indicating that there were no differences in perceived bitterness for this pair, and thus there was no perceivable change in the development of bitter taste in the first hour of fermentation. The d` for hours 2-1 increased to 1.09, indicating that the perceived bitterness began to develop and increase after one hour of fermentation. Pairs of samples comparing hours 3 – 1 and 2 – 1 were significantly different (d` = 1.09) but decreased for the pair comparing samples at hours 4 – 1 (0.77) and 5 – 1 (0.98).

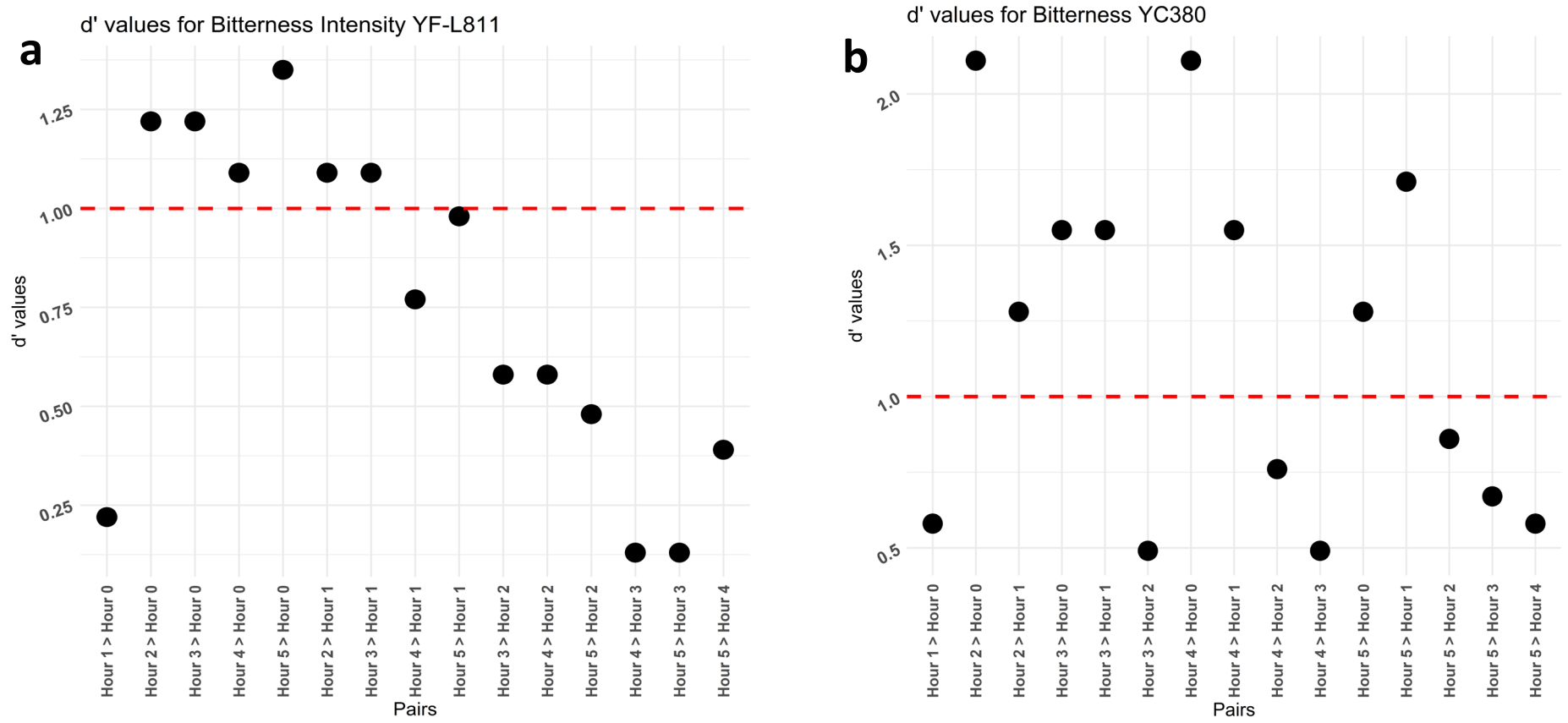


Figure 5.6 d` values for bitterness in YF-L811 (a) and YC380 (b). The 15 pairs of samples are organised on the x-axis, in order of increasing fermentation time. The d` denotes the sensory distance between the more fermented vs less fermented samples in each pairing.

It could be that bitter peptides were developing throughout fermentation, via proteolysis, and were subsequently breaking down as fermentation continued, thereby limiting the development of bitterness.

The d` ranged from 0.49 (pair hour 3 - 2 and 4 - 3) to 2.11 (pair hour 2 - 0 and 4 - 0) for bitterness intensity in YC380 samples, indicating that the difference in bitterness intensity between pairs was more perceivable by consumers. Similar to YF-L811, the d` increased in pair hour 5 – 1 (1.71), compared to hour 4 – 1 (1.55) and 3 – 1 (1.55), which might again suggest that there was some bitterness developing during late fermentation.

Bitterness intensity in YC380 yielded higher d` values, compared to YF-L811. The number of participants assessing samples prepared from YC380 was higher by two than the number assessing YF-L811 samples; a higher sample size can lead to higher d` values (O'Mahony and Rousseau, 2003).

For both starter cultures, the bitterness was not easily distinguished by the panellists in pairs of samples close together in fermentation time (i.e., pair hour 4 - 3). These results are consistent with the results from the Friedman analysis, where the bitterness intensity did not increase significantly after a certain period of fermentation for either starter culture. These results indicate that fermentation time significantly affected the perceived bitterness for these starter cultures.

Flavour intensity increased significantly with fermentation time for both starter cultures (Figure 5.7). The flavour intensity in YF-L811 developed slowly at the beginning of fermentation. There was not a significant increase in the flavour intensity in samples fermented for less than two hours, after this, the flavour increased but did not change significantly again after three hours of fermentation. In YC380, there was a steady increase in the flavour intensity across fermentation. After one hour of fermentation, the flavour intensity continued to increase significantly with fermentation time. This continued throughout the fermentation, suggesting flavour continued to increase with increased fermentation time.

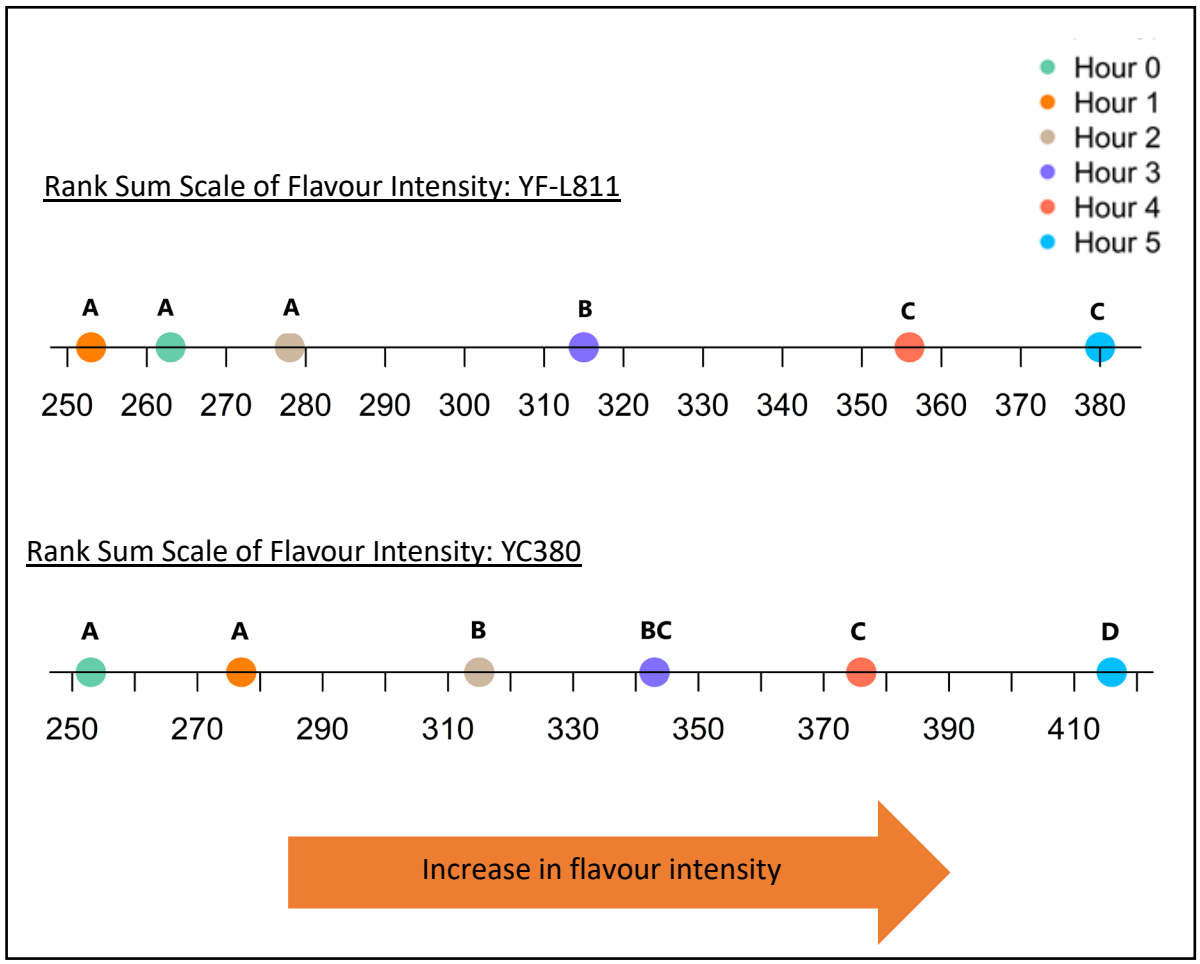


Figure 5.7 Rank sum scale of flavour intensity for YF-L811(n=41 participants) and YC380 (n=43 participants). The arbitrary numbers correspond to the samples rank position within the sample set. Different lettering above the samples indicates a significant increase in bitterness (α=0.05), determined by Tukey's HSD test. The samples prepared for each starter culture increase in perceived flavour intensity, with an increased fermentation time.

The d` for flavour intensity in YF-L811 ranged from -0.22 (pair hour 1-0) to 2.34 (pair hour 4-2) (Figure 5.8a). For the pair at hours 1 - 0, samples at hour 1 were perceived to be less flavoursome than hour 0. The d` increased between pairs of samples, with an increase in fermentation time. Compared to the bitter intensity, there was a gradual increase in perceivable flavour intensity between all pairs with hour 0.

The d` for flavour intensity in YC380 ranged from 0.86 (pair 2 - 1) and 2.11 (pair 3 - 0, 5 - 0, and 5 - 3) (Figure 5.8b).

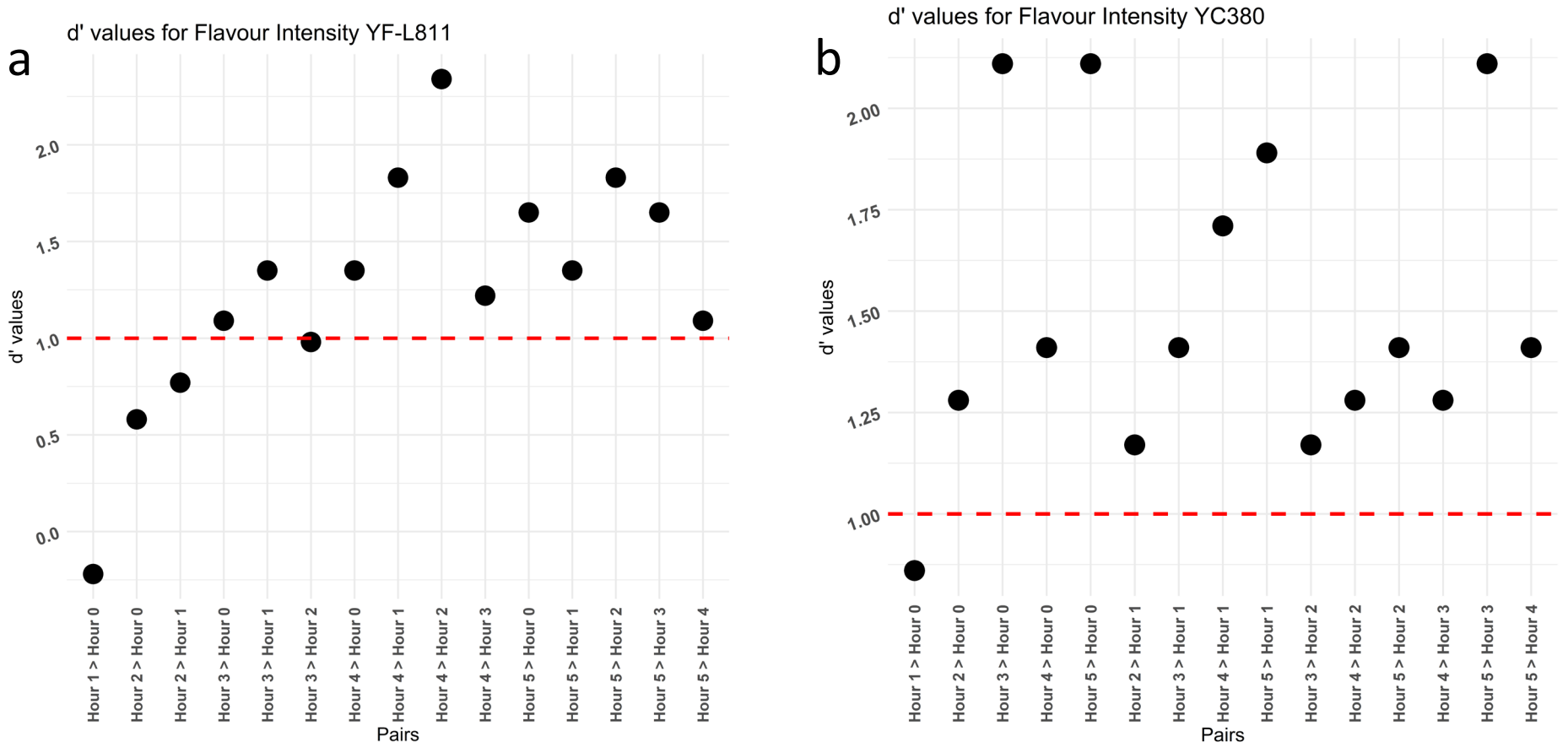


Figure 5.8 d` values for flavour intensity in YF-L811 (a) and YC380 (b). The 15 pairs of samples are organised on the x-axis, in order of increasing fermentation time. The d` denotes the sensory distance between the more fermented vs less fermented samples in each pairing.

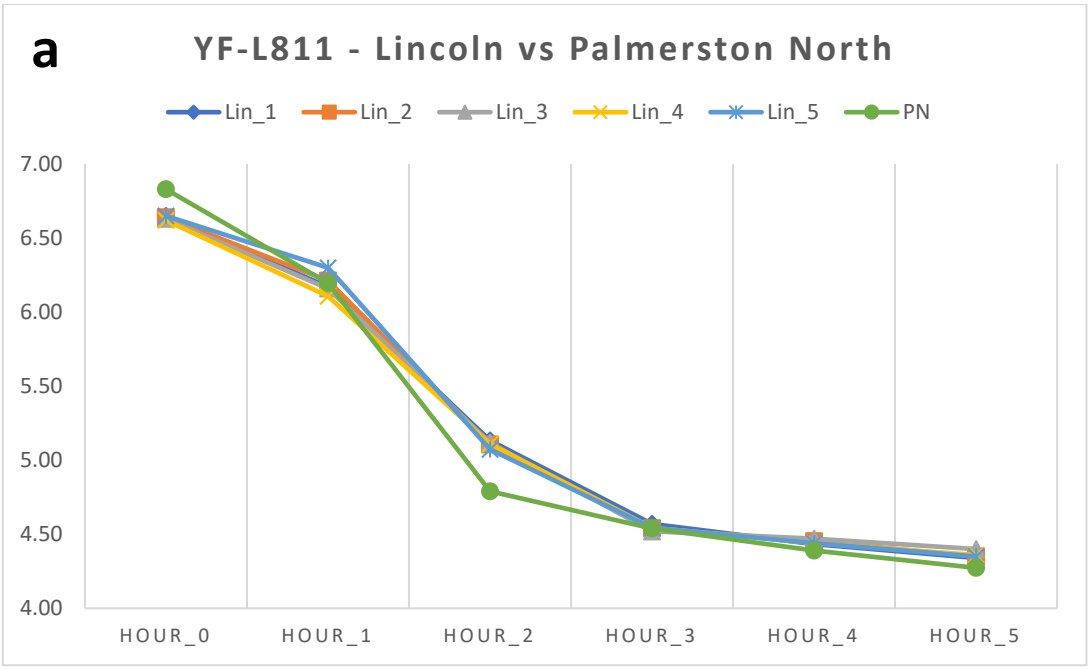
The panellists could discriminate between all pairs of samples, with the exception of pair 2 – 1. Of note was the decrease in d` between pair 4 - 0, and the increase again between pair 5 – 0, which may be reflective of compounds developing with continued fermentation that are affecting the flavour intensity. The d` values were high for all pairings, thus the flavour intensity between all pairs of samples was different and perceivable by consumers. YC380 is known for producing fermented milk with a strong flavour, compared to YF-L811 which produces a milder flavoured product. As a result, it was not unexpected that the consumers were able to easily discriminate between the different samples fermented with YC380.

Compared to bitter intensity, there was a gradual increase in perceivable flavour intensity between all pairs with hour 0. With bitter intensity, after a certain period of fermentation, there was a decrease in the ability to discriminate between samples. The d` decreased in pairs of samples that were close in fermentation time, although in most cases there was still a perceivable difference in the flavour intensity between the samples in each pairing. These results indicate that fermentation time significantly affected perceived flavour intensity for these starter cultures.

5.3.2 Peptide fingerprinting of samples

Replicate samples were prepared on a small scale in the laboratory at AgResearch, Lincoln and, alongside samples used for the consumer trial, were fingerprinted for peptides using MALDI-TOF.

pH profiles: The acidification profile was inconsistent between the samples prepared in the different locations; samples prepared in Lincoln fermented slower than those prepared for the consumer trial in Palmerston North (Figure 5.9a, b). The pH profile was similar for both locations for samples prepared using YF-L811, with the exception of the sample from Palmerston North at hour 2, which had a pH of 4.79 (± 0.01), compared to 5.11 (± 0.02) in samples prepared in Lincoln. The pH profiles for hours 2 onwards were different for samples prepared using YC380.



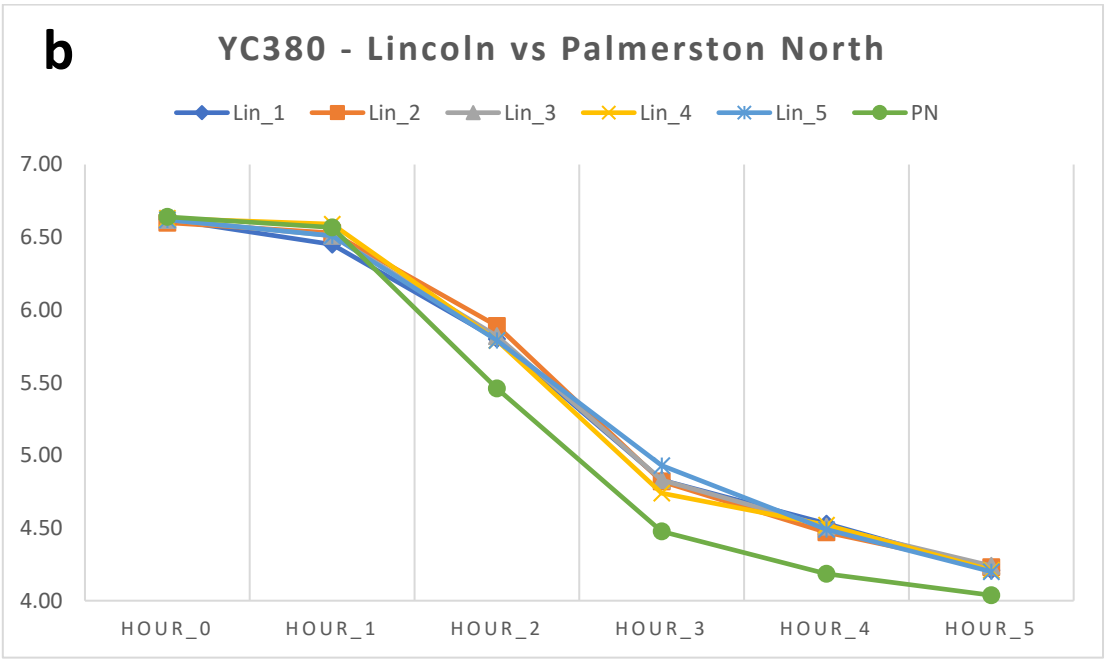


*Figure 5.9 pH profiles of replicate samples prepared in Lincoln (labelled Lin) and in Palmerston North (PN) for YF-L811 (a) and YC380 (b). Average of three measurements for each sample. A two-way ANOVA (pH ~ location * time) for each starter type was not significant ($p = 0.679$ for YF-L811, $p = 0.350$ for YC380).*

There are several potential reasons for the variations in pH change. The samples prepared for the consumer trial were prepared in a large 60 L tank. This may have resulted in an uneven distribution

of heat throughout the milk, and milk closer to the perimeter of the tank is closer to the heat source and as such may have been subjected to higher temperatures than milk at the centre of the tank. Samples prepared in Lincoln were prepared in small 15 mL tubes, where the temperature is much easier to distribute throughout the milk. Temperature variations have been recorded in fermenting tanks in other studies, affecting the pH of the milk during fermentation. Aguirre-Ezkauriatza et al. (2008) observed temperatures in a 5 L tank varying from 39 to 44 °C in radial distance of less than 7 cm, and the temperature of samples collected from the bottom of the tank were recorded as high as 70 to 80 °C after 4 and 5 hours of fermentation. After 5 hours of fermentation, temperatures recorded at the top and bottom of the tank were 30 and 80°C, respectively. Samples collected at the same time point varied in pH by as much as 0.5 units when sampled from different radial locations within the tank (Aguirre-Ezkauriatza et al., 2008). Variations in temperatures throughout the tank may have induced pH gradients resulting in different growth and lactic acid production rates in this study. Moreover, once fermentation begins, heat transfer is slowed due to the increase in viscosity, which can cause wider ranges in temperature (Aguirre-Ezkauriatza et al., 2008). Additionally, samples prepared in Palmerston North were removed in 10 L batches which were distributed into three 3 L containers and placed on ice water in a cold room. However, these samples likely continued to ferment while the mixture was being cooled. The small batches prepared in Lincoln cooled much more rapidly due to their small volume, causing the fermentation to halt faster. That said, it is unexpected that the variations in the pH are inconsistent between the different products.

Comparison of samples prepared for consumer and instrumental analysis

Next, the peptide fingerprints obtained from each set of samples (Palmerston North and Lincoln) were assessed to understand how comparable the samples were. Differences in the samples have implications for scaling up the sample preparation and for making comparisons with experimental samples prepared in a laboratory.

A sparse MINT-PLS-DA for the peptide fingerprints obtained for YF-L811 demonstrated that most of the samples could be grouped according to their fermentation time (Figure 5.10). Samples corresponding to hours 3, 4 and 5 were all grouped close together, though there was a split in the group for hour 4 and 5. The samples prepared in the different locations were grouping together at their respective fermentation time points, although the replicates prepared in Palmerston North at hours 0 and 2 were tending to group away from the Lincoln samples.

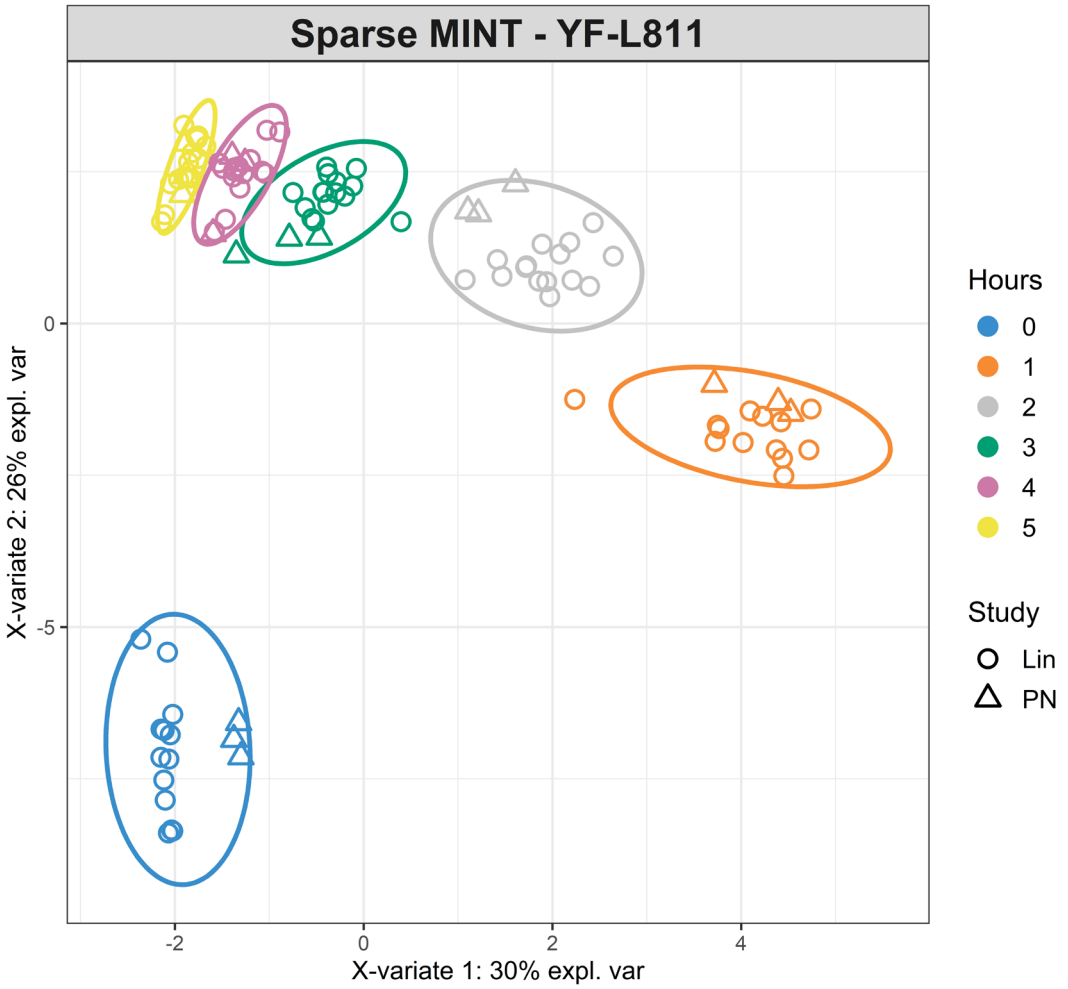


Figure 5.10 A Sparse MINT-PLSDA on peptide fingerprints from samples prepared using YF-L811. Samples are colour-coded by fermentation time point. The different shapes correspond to location of sample preparation: Lin = Lincoln, PN = Palmerston North. Five replicates were prepared for each sample in Lincoln. Three measurements were taken using MALDI-TOF MS for each sample type.

A sparse MINT-PLS DA of YC380 samples was able to group all samples separately by each time point (Figure 5.11). For most of the fermentation time points, the samples prepared in both Lincoln

and Palmerston North were grouping together, with the exception of one replicate corresponding to hour 1, 3 and 5 from Palmerston North. However, these samples were from the same biological replicate and this separation was driven by technical variation.

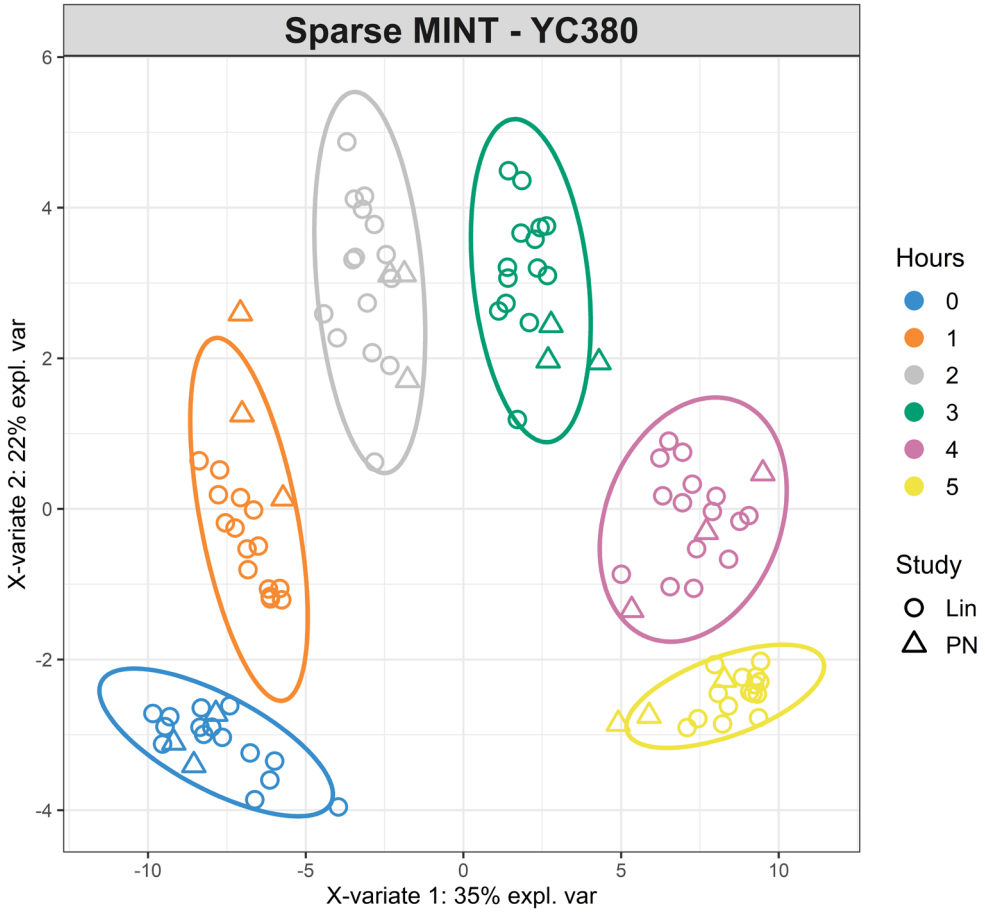


Figure 5.11 A Sparse MINT-PLSDA on peptide fingerprints from samples prepared using YC380. Samples are colour-coded by fermentation time point. The different shapes correspond to location of sample preparation: Lin = Lincoln, PN = Palmerston North. Five replicates were prepared for each sample in Lincoln. Three measurements were taken using MALDI-TOF MS for each sample type.

The sparse MINT-PLS-DA indicated that comparisons can be made between the peptide fingerprints of samples prepared in different locations, at different time points, and in different quantities. This is an important conclusion for future studies when attempting to draw conclusions based on samples prepared on a laboratory scale. However, this should be approached cautiously as the deviations in the pH may have led to variations in other characteristics in the fermented products.

5.3.3 Correlating consumer responses with peptide fingerprints

Next, a means to correlate the consumer response with the peptide fingerprints of the fermented products was explored. First, principal components analysis (PCA) was performed on the samples prepared in Lincoln, averaged by their technical replicate (n=3). PCA of peptide fingerprints for YF-L811 showed a similar grouping as the MINT-PLS-DA (Figure 5.12). Samples at fermentation hours 0 and 1 did not cluster closely in the PCA space, which is in contrast to the bitterness and flavour intensity responses by the consumers, where hours 0 and 1 were not significantly different in either attribute. The samples clustered closely at the end of fermentation, which was similar to what was seen in the consumer responses where bitterness and flavour intensity did not increase significantly after a certain period of fermentation.

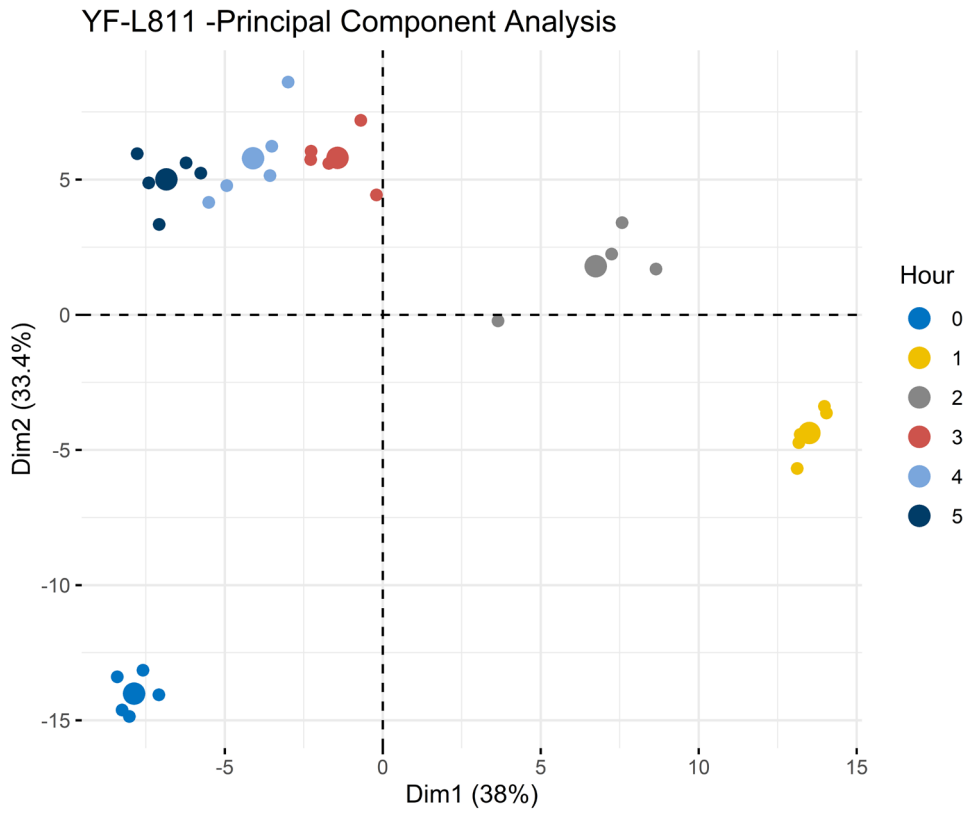


Figure 5.12 Principal components analysis of peptide fingerprints for samples prepared using YF-L811. Data is colour-coded by fermentation time point. Large dots indicate the average for that sample type. Five replicates were prepared for each sample type. MALDI-TOF technical repeats were averaged. Samples measured were prepared in Lincoln.

PCA of peptide fingerprints for YC380 was able to group each fermentation time (Figure 5.13). The more fermented samples (hour 2 onwards) were not grouped closely. This was similar to the consumer responses to flavour intensity for this starter culture; the flavour intensity continued to develop significantly throughout the fermentation. The PCA plot indicated that the peptide components in this starter culture were continuing to change with an increased fermentation time.

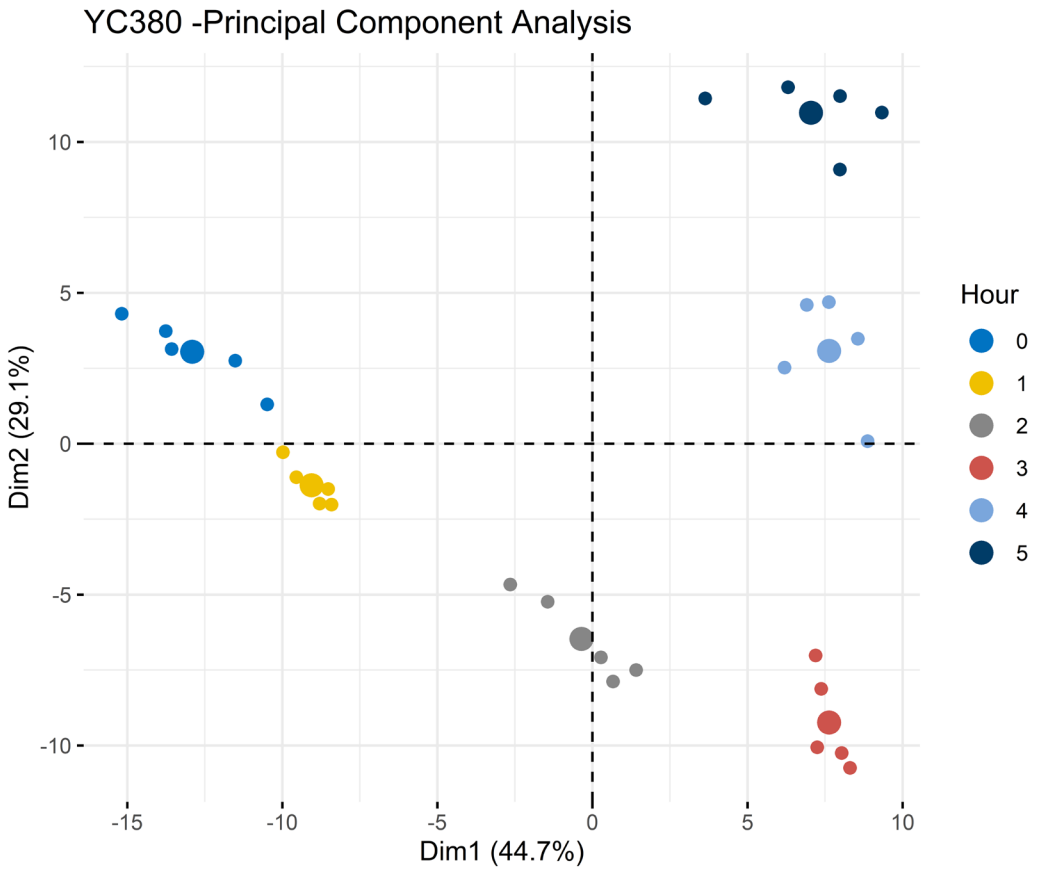


Figure 5.13 Principal components analysis of peptide fingerprints for samples prepared using YC380. Data is colour-coded by fermentation time point. Large dots indicate the average for that sample type. Five replicates were prepared for each sample type. MALDI-TOF technical repeats were averaged. Samples measured were prepared in Lincoln.

For each PCA, the top four components were extracted. A Spearman correlation was then applied to the top components and the consumer responses for each attribute (Figure 5.14a, b; Table 13). There was a significant correlation for bitterness intensity on component 2 for YF-L811 (correlation coefficient = 0.506) and component 1 (0.583), 2 (0.188) and 3 (-0.158) for YC380. The

Spearman correlation returned a correlation coefficient between -1 and 1, where values closer to 1 indicate a perfect correlation and values closer to 0 indicate no correlation. Although the correlation coefficients were significant, the correlation with components 2 and 3 for YC380 was low.

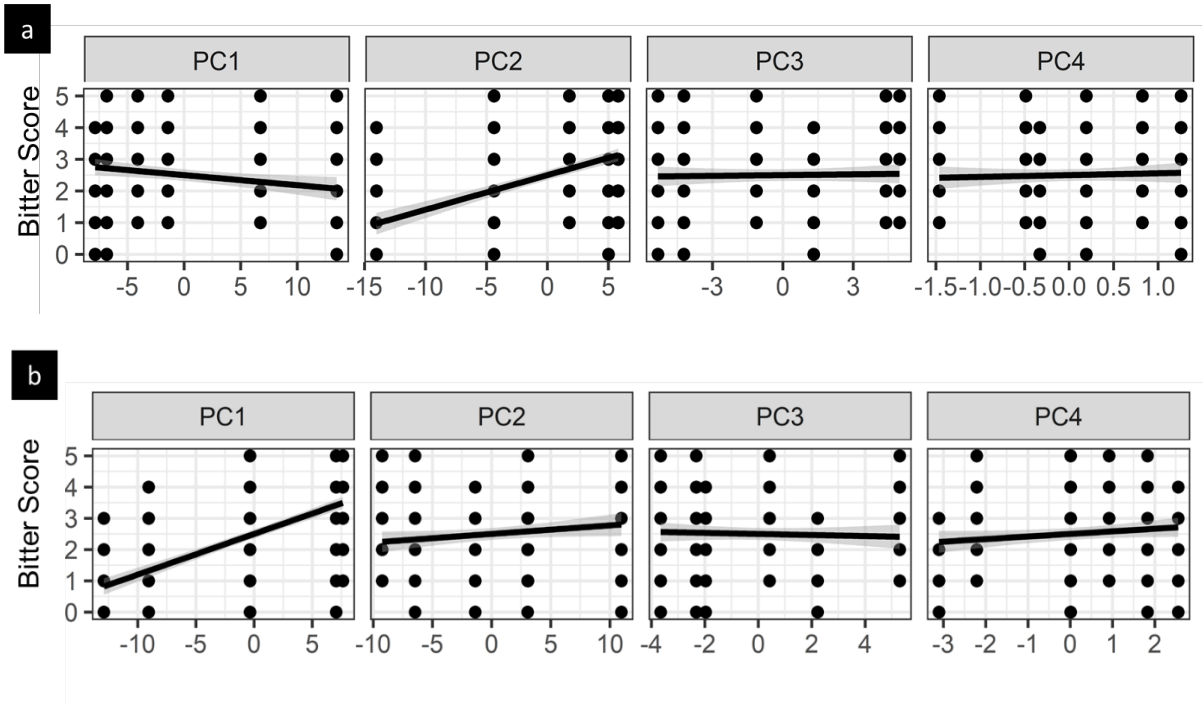


Figure 5.14 The Spearman correlation was performed by extracting the top four principal components from PCA of peptide fingerprints and performing a correlation with the consumer scores obtained for bitter intensity in YF-L811 (a) and YC380 (b). Consumer rankings for bitterness are denoted along the y-axis; the x-axis represents the extracted principal component scores.

*Table 13 Spearman correlation of top four components from a PCA of peptide fingerprints for YF-L811 and YC380 on consumer responses to bitter intensity. * Indicates a significant correlation ($p \leq 0.05$).*

Correlation – Bitter Intensity	Component 1	Component 2	Component 3	Component 4
YF-L811	-0.028	0.506*	-0.02	0.001
YC380	0.583*	0.188*	-0.158*	0.022

The correlations with flavour intensity were higher than bitterness for each starter culture (Figure 5.15a, b; Table 14). There was a significant correlation with the top three components for YF-L811, with correlation coefficients of -0.31, 0.529 and -0.32, respectively. There was also a significant correlation with the top three components for YC380, with correlation coefficients 0.602, 0.361 and -0.207, respectively.

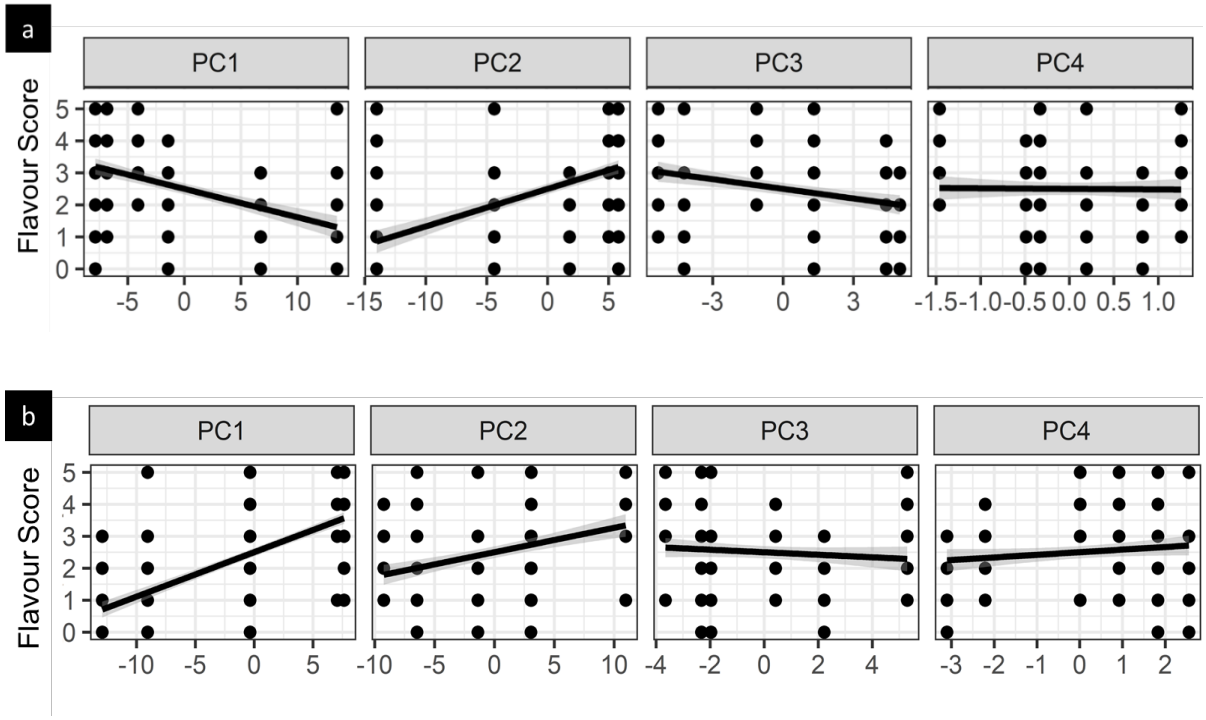


Figure 5.15 The Spearman correlation was performed by extracting the top four principal components and performing a correlation with the consumer scores obtained for flavour intensity in YF-L811 (a) and YC380 (b). Consumer rankings for flavour intensity are denoted along the y-axis; the x-axis represents the extracted principal component scores.

Table 14 Spearman correlation of top four components from a PCA of peptide fingerprints for YF-L811 and YC380 on consumer responses to flavour intensity. Significant correlations are denoted by "".*

Correlation – Flavour Intensity	Component 1	Component 2	Component 3	Component 4
YF-L811	-0.31*	0.529*	-0.32*	-0.015
YC380	0.602*	0.361*	-0.207*	-0.01

5.3.4 Prediction of consumer responses from peptide fingerprint data

XGBoost models were built using components 1, 2 and 3 extracted from a PCA of the peptide fingerprints. The XGBoost model metrics indicated adequate performance for predictions of flavour intensity in both YF-L811 and YC380 (Table 15). For the XGBoost model of flavour intensity in YF-L811, the model metrics were R^2 of 0.53, and 1.088, 1.184 and 0.856, for RMSE, MSE and MAE, respectively. For the XGBoost model of flavour intensity in YC380, the model metrics were R^2 = 0.69, RMSE = 0.896, MSE = 0.802 and MAE = 0.710.

Table 15 XGBoost model metrics for predictions of consumer responses to flavour intensity for YF-L811 and YC380.

Flavour Intensity	RMSE	MSE	MAE	R^2
YF-L811	1.088	1.184	0.856	0.53
YC380	0.896	0.802	0.710	0.69

The XGBoost predictions of flavour intensity in YF-L811 were able to closely predict some of the ranks (Figure 5.16).

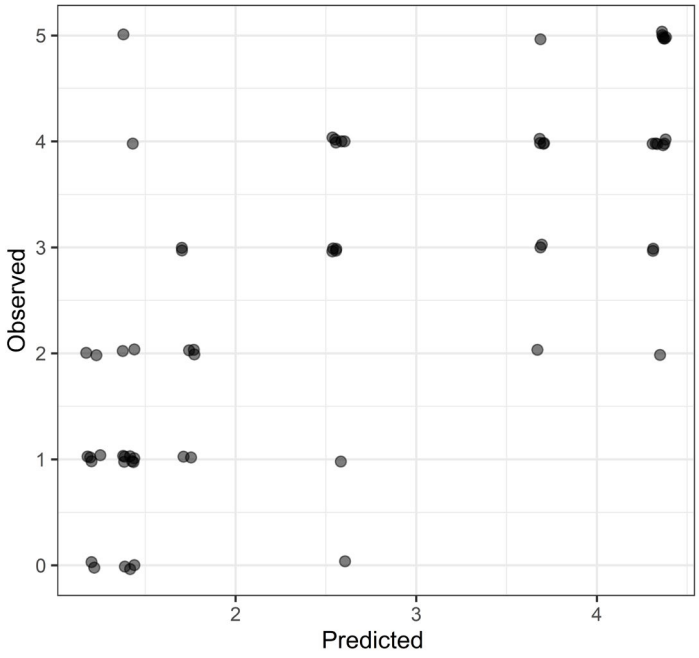


Figure 5.16 Predicted vs Observed values for XGBoost model of flavour intensity on the principal components of peptide fingerprints obtained from YF-L811 samples.

Although, in most cases, the predictions were inaccurate, and the model struggled to correctly predict higher ranking samples. This is not surprising as the consumer results indicated that the panellists were unable to discriminate differences in flavour and bitterness intensity towards the end of fermentation (with the exception of flavour intensity in YC380).

The XGBoost model's predictions of flavour intensity rankings for YC380 demonstrated good accuracy (Figure 5.17). The XGBoost model was able to accurately predict the consumer responses for each ranking, and had a high R^2 (0.69), indicating that consumer responses to flavour intensity in milk fermented with YC380 can be reliably predicted.

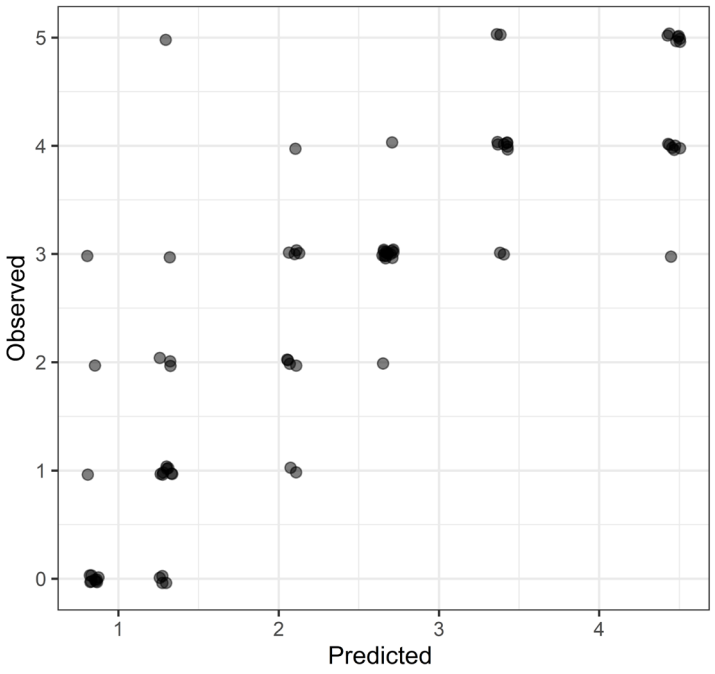


Figure 5.17 Predicted vs Observed values for XGBoost model of flavour intensity on the principal components of peptide fingerprints obtained from YC380 samples.

The results of these predictive models indicated that consumer responses to flavour intensity can be reasonably predicted for these two starter cultures using data obtained from peptide fingerprinting. The ability to predict the flavour intensity for different extremes of flavour (i.e., mild vs strong) could prove beneficial for screening new products or bacterial cultures that possess different levels of flavour in the final product.

The metrics were worse for models predicting bitter intensity scores. For bitter intensity on YF-L811 samples (Table 16), the R^2 was low: 0.22, accounting for a low proportion of the variation in the model. The RMSE and MSE were high for this model at 1.238 and 1.532 respectively. The MAE is 0.971 for bitterness in YF-L811. The model performance for the XGBoost model predicting bitter intensity ranks in YC380 samples was slightly improved: the R^2 was 0.54. The RMSE, MSE and MAE were slightly lower, compared to YF-L811, at 1.134, 1.285 and 0.910, respectively.

Table 16 XGBoost model metrics for predictions of consumer responses to bitter intensity for YF-L811 and YC380.

Bitter Intensity	RMSE	MSE	MAE	R^2
YF-L811	1.238	1.532	0.971	0.22
YC380	1.134	1.285	0.910	0.54

A plot of the predicted vs observed values for bitterness prediction in YF-L811 samples demonstrated the error in the predictive accuracy (Figure 5.18). The model did not correctly predict any sample with a rank greater than 4.

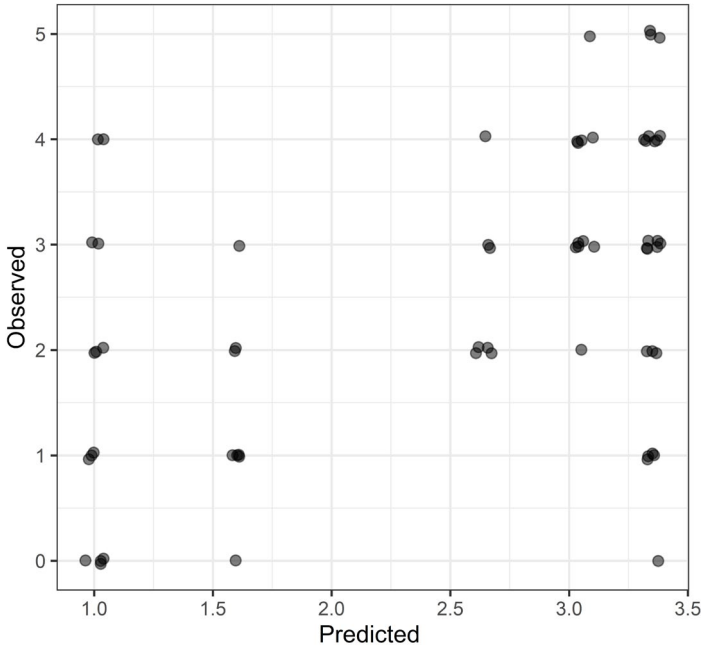


Figure 5.18 Predicted vs Observed values for XGBoost model of bitter intensity on the principal components of peptide fingerprints obtained from YF-L811 samples.

A plot of the predicted vs observed values for bitterness prediction in YC380 samples showed the model was able to predict some of the bitterness rankings with reasonable accuracy (Figure 5.19). However, similar to YF-L811, the model did not correctly predict any samples with a ranking greater than 4.

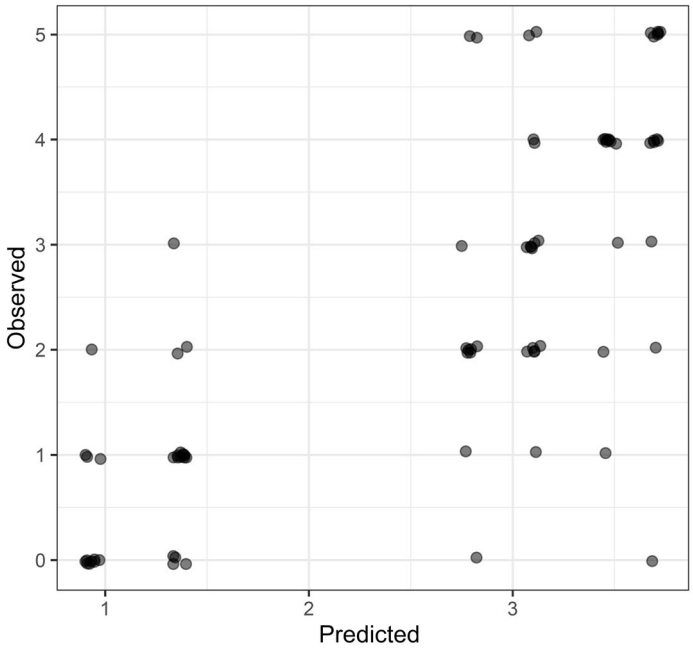


Figure 5.19 Predicted vs Observed values for XGBoost model of bitter intensity on the principal components of peptide fingerprints obtained from YC380 samples.

In conclusion, significant correlations could be made with MALDI-TOF data and consumer responses to important sensory attributes in fermented milk. Predictions for flavour and bitterness intensity were better for fermented milk prepared using YC380 than for YF-L811. This may suggest that predictions of flavour intensity can be made with reasonable accuracy for both starter cultures, and potentially could be used for future studies. Predictions of YF-L811 were poor for bitterness intensity but were improved for flavour intensity. Higher ranking samples, for both bitterness and flavour intensity, were not accurately predicted. This could be related to the consumer's inability to discriminate between more fermented samples (i.e., higher ranking samples). Alternative modelling approaches and techniques, as well as incorporating other data measurements, could be explored to improve predictions. Additionally, increasing the sample size may lead to more accurate and reliable results.

5.4 Conclusions

This chapter explored an approach to correlate and predict responses to key sensory attributes throughout fermentation of milk with the peptide fingerprint of the same products. Consumers perceived significant differences in bitter and flavour intensity throughout the fermentation of milk prepared from different starter cultures. Flavour and bitterness intensity developed at different rates during fermentation for different starter cultures, and through a multiple paired comparison test, the points in fermentation where these attributes began to change in their perceived intensity could be identified. Instrumental analysis, using MALDI-TOF MS, was able to track changes in the peptide fingerprint throughout fermentation when analysed using multivariate analysis. The peptide fingerprints could be significantly correlated to the consumer responses for each attribute. XGBoost regression was employed to predict the consumer responses, with varying success. The model resulted in an adequate performance for the prediction of flavour intensity. Exploring other modelling techniques could improve the performance. By better understanding the development of desirable and off-flavour characteristics during fermentation, and by correlating consumer and instrumental data, this information can potentially aid in the rapid screening and selection of new and interesting bacterial cultures to enhance the sensory perception of fermented milk products.

Chapter 6 Discussion and conclusion

Developing new products that are acceptable to consumers can be a timely and costly task. Insights into the product development process could provide invaluable information to food manufacturers, aiding in monitoring and quality control during development stages. The overall goal of this thesis was to identify a means to aid in the rapid screening of fermented products which may enable the development of new products with desirable properties, and which can also contribute to understanding potentially important characteristics of the end product. The overall project objectives and applications are summarised in Figure 6.1.

This thesis constituted a feasibility study: the feasibility of using fingerprinting technologies to extract knowledge and as a tool for screening fermented milk was investigated (Chapter 3). It was demonstrated that using rapid fingerprinting technologies, MALDI-TOF and REIMS, could be used to screen fermented milk with differing flavour attributes. REIMS has not previously been demonstrated for use on dairy products, and this is the first study showing a side-by-side analysis of REIMS and MALDI-TOF data. As REIMS provides results in near-real-time, this could prove hugely beneficial for the dairy industry during product development. For the starter cultures investigated in this study, peptides fingerprinted using MALDI-TOF MS and analysed using principal component analysis (PCA) revealed the strongest discrimination between samples. Using peptide fingerprints, it was demonstrated that unique peptides are produced from these two bacterial combinations, which create a product with differing flavour and textural attributes. This technique could be used in further applications to discriminate different fermented milk and to monitor product development. Although REIMS did not provide a powerful split between the different samples, it indicated some promise for visualising and clustering samples, and perhaps can be explored as a complementary screening tool. These two techniques could be used in tandem as an indication of the differences in the molecular composition of fermented milk prepared using various starter cultures. These tools require much less sample preparation compared to other instrumentation, as well as having shorter data acquisition and data processing time, making them

more suitable for high-throughput use in the food processing industry for monitoring, quality control and rapid screening of samples during product development (Barlow et al., 2021, Cohen and Gusev, 2002, Huang et al., 2019). The techniques described in the literature, such as 'sensoproteomics' (Sebald et al., 2018), can provide substantial insights into the chemical composition of food products and the peptides that may be impacting the sensory profile. However, these techniques employ sample preparation, instrumentation, and data analysis that could be laborious and costly, and may not lend themselves to rapid screening of products. This thesis explored techniques that could be first used as preliminary screening tools. It is the intention that follow-up techniques, similar to 'sensoproteomics', could then be utilised on candidate products to explore these further.

The ability to reveal strong clustering of different sample types within a short period of time can enable the rapid screening of a large number of samples. Furthermore, when samples are analysed alongside unfermented milk samples, multivariate analysis can indicate which samples undergo the greatest compositional change during fermentation, which may be reflective of end product characteristics, such as flavour, taste, textural or nutritional characteristics. To determine this for certain, the peptides in the samples would need to be identified and their sensory properties determined through database searching or by conducting sensory evaluation. Lastly, this chapter demonstrated the potential predictive power of classifying fermented milk prepared using different bacterial strains. This could prove advantageous and has become a common approach in food studies of late, particularly due to the increased occurrence of food fraud; in the dairy industry this has been associated with authenticity of milk type and origin (Garcia et al., 2012, Calvano et al., 2013a, Sassi et al., 2015, England et al., 2020, Piras et al., 2021) as well as the bacterial composition (Liu et al., 2010, Rocchetti et al., 2018).

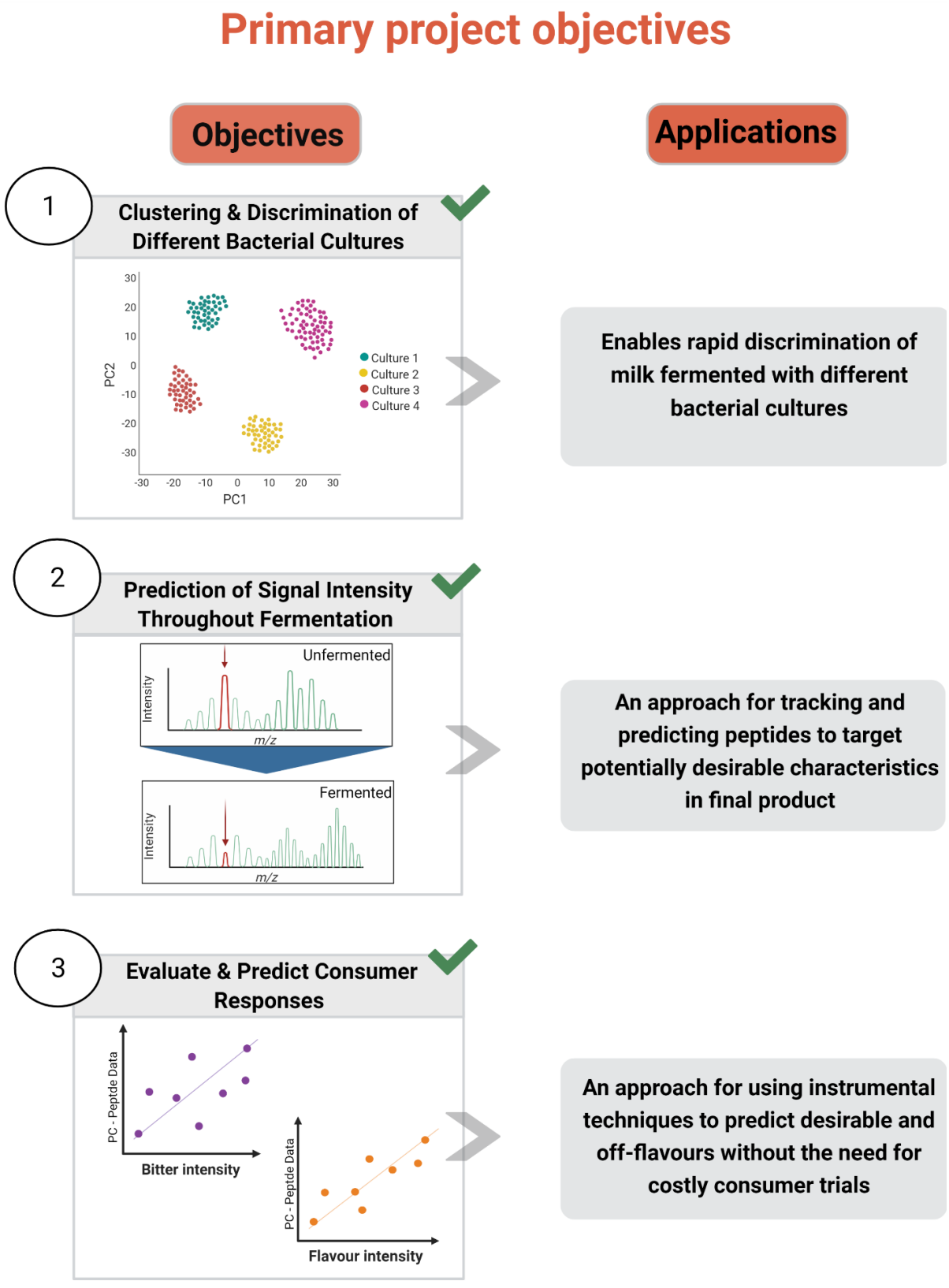


Figure 6.1 A summary of the primary project objectives and their potential applications. This thesis was divided in to three chapters to address the thesis objectives: Clustering and discrimination of fermented milk prepared from different bacterial cultures; Prediction of signal intensity throughout fermentation of milk, and evaluation and prediction of consumer responses throughout fermentation of milk. This figure was generated using BioRender.com.

The peptide fingerprint was further explored in milk fermented with various probiotic cultures, that was sampled throughout fermentation (Chapter 4). This study indicated that the starter culture and fermentation time had a more pronounced effect on the peptide fingerprint, compared to the probiotic bacterial cultures. Studies have shown that there is an increase in certain compounds during fermentation with probiotics, compared to milk fermented without probiotics (Settachaimongkon et al., 2014, Tian et al., 2017), although it is said that the bacteria may not significantly impact the sensory profile (Chen et al., 2017). Probiotic cultures have been reported to generate bioactive peptides in dairy products (Hernández-Ledesma et al., 2014), as such it would be expected that the generation of such peptides might be reflected in the peptide fingerprint by demonstrating some clustering between these cultures.

Fingerprinting of these samples also provided some insight into how the fermented milk with milder flavour clustered more closely with the unfermented milk samples, which may be reflective of the potentially low proteolytic activity occurring in fermentation with this starter culture, and which may also be related to flavour and textural properties. For instance, as indicated by the multivariate analysis of YF-L811 and YC380, samples fermented with YF-L811 formed a similar cluster with unfermented milk inoculated with YC380. This suggests that YC380 is undergoing a greater change of the peptide fingerprint during fermentation. Unique *m/z* ions were also shown to be characteristic of fermented and unfermented samples; this could also be used as a means to monitor the fermentation process and as a predictor of the endpoint of fermentation. Such approaches have been reported by other researchers in milk products (Amorim et al., 2019, Dalabasmaz et al., 2019, Ebner et al., 2016).

Classical statistical and machine learning techniques were employed to explore the feasibility of predicting the peptide fingerprinting throughout fermentation of milk inoculated with various combinations of starter and probiotic culture (Chapter 4). This objective aimed to predict the signal intensity of *m/z* ions detected by MALDI-TOF throughout fermentation. Regression techniques were used to predict the peptide intensity at different stages of fermentation using both linear and

nonlinear techniques. Nonlinear general additive models proved effective for modelling most *m/z* ions detected, and also provided some valuable insight into the behaviour of *m/z* ions throughout fermentation for different bacterial combinations. This behaviour could be reflective of the underlying proteolytic processes, i.e., proteins break down into peptides, peptides break down into smaller peptides and amino acids and derivatives, which may contribute to flavour, taste, textural or bioactive properties (Tamime and Robinson, 1999). Machine learning classification techniques were effective for predicting the general change in direction of these *m/z* ions based on their starting signal intensity. This approach used only a subset of the *m/z* ions detected, which were changing the most throughout fermentation. The models were effective for predicting the direction of the signal intensity from the beginning until the end of fermentation time. This may prove useful for tracking specific peptides with desirable properties and could provide a quick and efficient way to understand the breakdown of a peptide, without the need for complex data analysis.

Finally, important sensory attributes in milk fermented for different lengths of time were investigated using a multiple paired comparison test (2-alternative forced-choice) with a consumer panel (Chapter 5). This indicated significant differences in perceived bitterness and flavour intensity at different fermentation times. Furthermore, this work explored the consistency of samples prepared on a small laboratory scale, compared to samples prepared on a larger scale intended for human consumption. This work indicated that comparisons could be drawn between the peptide fingerprints of samples produced under different conditions and different geographical locations. Finally, the feasibility of integrating consumer data obtained via multiple comparisons with peptide mass spectra was investigated. This work provided insights into how important sensory characteristics develop throughout fermentation, and how they continue to develop after the standard period of fermentation (e.g., pH ~ 4.5). This emphasises how critical the end fermentation time is during fermented milk preparation and development. Lastly, this study demonstrated that significant correlations could be made between peptide mass spectra and consumer responses to fermented milk, particularly in the case of responses to flavour intensity.

Predictive modelling proved somewhat successful in predicting these responses using XGBoost regression models, in particular for the prediction of flavour intensity. The ability to predict key sensory attributes from instrumental data can provide huge advantages to the food industry, providing a less-costly and less-timely alternative to sensory evaluation. Correlating sensory results with instrumental analysis is a valuable methodology that can be applied for improving and maintaining quality during production and development. Sensory evaluation is ultimately the gold standard, however, using instrumental analysis in particular at early stages of product development may aid in steering the decision-making and help to shortlist prospective candidate products that could then be taken to sensory evaluation. Correlating sensory and instrumental analysis has been successfully shown in fermented milk products previously (Gallardo-Escamilla et al., 2005, Güler, 2007) and in various cheese products (Lawlor et al., 2002) with much success. The current study investigated the feasibility of correlating peptide fingerprints using rapid mass spectrometry, MALDI-TOF, with consumer responses to fermented milk, which has not previously been demonstrated in the literature.

The typical flavour of fermented milk results from a combination of lactic acid, aroma compounds and various compounds generated during fermentation, via proteolytic or enzymatic activity (Tamime and Robinson, 1999). Tastant peptides in fermented dairy foods are typically derived from the native protein (Toelstede et al., 2008a, b, Toelstede et al., 2009). The correlation between the peptide fingerprint and sensory responses to flavour in this study indicate that the peptide composition in these samples is impacting the flavour profile. Whether there are peptides (e.g., kokumi or umami) directly impacting the flavour profile would require further investigation through peptide identification and sensory evaluation of such peptides (Toelstede et al., 2008a, b, Toelstede et al., 2009).

Limitations

Investigating chemical composition and its relation to sensory properties in food poses a number of challenges. In fermented dairy products, this is further complicated by the effects of the

fermentation process and storage on the product. Compounds may be degraded when exposed to heat or oxygen, and storage time and temperature can induce changes in chemical composition (Cheng, 2010). Furthermore, sample preparation techniques for MS analysis can result in the loss or degradation of important compounds that may influence the sensory properties (Cheng, 2010). As such, it can be difficult to accurately characterise the product's chemical composition.

Samples prepared for sensory evaluation in this thesis were unavoidably sampled throughout the day by consumers, which raises the issue of whether samples evaluated at the beginning of the day are identical to those evaluated at the end of the day. It is possible that samples have undergone fermentation while in storage during the day which may have affected the sensory characteristics of the product. Due to the logistical challenges of conducting sensory evaluation and the challenging nature of fermented products, such a limitation is difficult to avoid.

As indicated by fingerprinting analysis of fermented milk (Chapter 4), different bacterial combinations result in different proteolytic activity and behaviour throughout fermentation, as such attempting to treat and analyse all of these different samples simultaneously may not be practical. MALDI-TOF is not quantitative and only provides a relative quantitation; however, the signal intensity is correlated with the concentration of a molecule and so the prediction of signal intensity could provide an understanding of the general trend of the molecule. Although, this should be approached with caution as complex mixtures can result in poor correlations with ion intensity (Duncan et al., 2008).

The multiple paired comparison test performed during the consumer trial did not provide a measure of the actual intensity of the attribute (e.g., bitterness or flavour intensity), but rather it provided a measure of the perceivable differences in intensity within that sample set (Kemp et al., 2009). A trained panel could be used to measure the intensity of these attributes, which may provide additional insights and measurements to correlate with instrumental analysis (Toelstede et al. 2008a, b).

The large number of samples may have induced carry-over effects or fatigue in the participants during the multiple paired comparison testing (Rogers, 2017). During each session of the consumer trial, the participants were consuming 30 samples. Although the experimental design and breaks between each sample set should offset such issues to some extent, the volume and number of samples being consumed may have caused some fatigue and carry-over which may have impacted the results. Untrained panellists are also more prone to fatigue and carry-over effects (Rogers, 2007).

Value for dairy industry

This study explored numerous methodologies which may be valuable to the dairy industry. Rapid fingerprinting techniques could be utilised to discriminate different fermented milk products and to monitor product development. Short data acquisition and data processing time make these techniques more suitable for high-throughput use in the food processing industry for monitoring, quality control and rapid screening of samples during product development. Multivariate analysis can reveal which cultures undergo the greatest compositional change during fermentation, which may be reflective of the end product characteristics, such as flavour, taste, textural or nutritional characteristics. Unique *m/z* ions were also shown to be characteristic of fermented and unfermented samples, this could be used as a means to monitor the fermentation process and as an indicator of the endpoint of fermentation. Prediction of peptides is useful for tracking specific peptides with desirable properties and could provide a quick and efficient way to understand the breakdown of a peptide, without the need for complex data analysis. By being able to correlate the changes in important sensory attributes with peptide fingerprints, the aim was to be able to provide an understanding of how the breakdown of products of proteolysis may affect the generation and development of these attributes.

Future prospects

The peptide fingerprint on its own does not provide an abundance of information. This study wished to understand how the peptide fingerprint could be exploited, without the need for time-consuming

peptide identification via MS/MS or other complex data analysis. Of course, this type of analysis can provide important biological insight into peptides and can provide a further understanding of the complex processes occurring during fermentation. However, this study was approached as a feasibility study: can the peptide fingerprint alone be used to provide rapid insights into a fermented product? A follow-up analytical study could be carried out to identify the peptides within the samples, which could then extend this work for tracking and targeting specific peptides with desirable properties. Several future works are proposed in Figure 6.2.

Findings from this research were specific to fermented milk and, for the most part, specific to peptide fingerprinting. There are numerous approaches that could be taken to further the work undertaken in this thesis, including using different food products, instrumental, sensory and data analysis techniques.

Fermented milk prepared under different processing conditions (heat-treatment, fermentation time, different milk base) using additional starter cultures could be examined to investigate whether the results in this study hold true under these conditions. Furthermore, additional measurements of peptide characteristics, using different analytical techniques, could be performed to validate and complement these findings. Further work could investigate other fermented foods with longer fermentation times, such as cheese. The effects of extended storage times on the peptide fingerprint and sensory properties could be examined to understand how these properties change throughout the product's life cycle.

Additional sensory testing, using trained panellists, could be carried out on specific peptides to determine their sensory properties and to quantify the intensity of the sensory attributes. Other sensory properties, such as the texture of fermented milk could also be investigated to understand if these properties may link with the peptide fingerprints.

Importantly, MALDI-TOF *m/z* ions could be identified to determine if these correspond to real peptides with relevant or interesting biological functions which could be utilised for tracking and

targeting during development. Additionally, quantitation on the peptides could be performed to validate whether the trend in the signal intensity recorded in MALDI-TOF is consistent with the actual concentration of peptides in these samples.

Further modelling on the peptide fingerprints could be performed to explore other predictor or response variables, i.e., pH values. Additional mass spectrometry analysis could be performed to measure other compounds (i.e., volatiles) which may be used alongside the current data to extend predictive capabilities and to understand whether the generation of these compounds could be correlated with the breakdown of the peptide profile. The inclusion of physicochemical properties, such as hydrophobicity could also be incorporated which may improve prediction and correlation with bitterness (Newman et al., 2014). Other modelling techniques for the prediction of sensory attributes could also be investigated. MALDI-TOF and REIMS datasets could be integrated to potentially strengthen the predictive power and capabilities of these techniques.

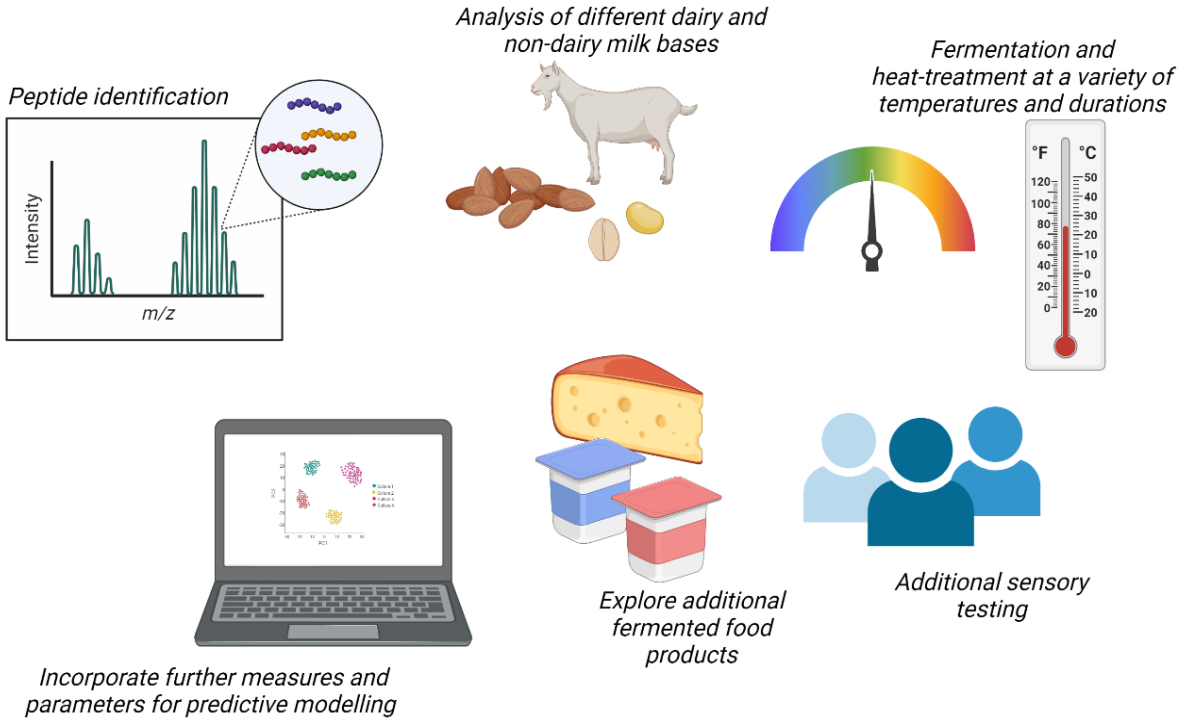


Figure 6.2 Opportunities for future work to further the work explored in this thesis

Concluding remarks

In conclusion, peptide fingerprinting of fermented milk products can provide insights into sensory characteristics, as perceived by consumers, and may be additionally used to predict changes in the fingerprint that could potentially enable the targeting of peptides that infer desirable characteristics.

Both MALDI-TOF and REIMS have promise as rapid screening methods for fermented milk and other dairy foods. These fingerprinting techniques are a rapid and inexpensive means, relative to GC-MS or LC-MS, to discriminate between flavour characteristics of fermented milk. Because of the quick analysis times, these techniques could be used in the food processing industry to rapidly screen samples during the development of new fermented milk products with different flavour characteristics. Molecular fingerprinting could distinguish the differences in proteolytic activity at different stages of fermentation for milk fermented using different bacterial combinations. Predictive modelling techniques could predict changes in the peptide fingerprint throughout fermentation; this was achieved to some degree using regression techniques to predict the intensity of individual *m/z* ions at different times and using classification techniques to generally predict the direction of intensity. This enables an understanding of how the peptide fingerprint changes throughout fermentation and could be utilised in the development of products to target specific molecules with desirable properties, without the need to carry products to full fermentation. Lastly, consumer responses to important sensory characteristics could be correlated with the peptide fingerprints, specifically in the case of flavour intensity. This enables an approach to understand and predict how consumers respond to these products.

This thesis presented novel approaches for screening the molecules in fermented milk products, for prediction of peptide fingerprints throughout fermentation and the correlation of consumer data with MALDI-TOF data, which may be utilised by food researchers and industry.

References

ABDELHAMID, H. N., KHAN, M. S. & WU, H. F. 2014. Design, characterization and applications of new ionic liquid matrices for multifunctional analysis of biomolecules: a novel strategy for pathogenic bacteria biosensing. *Anal Chim Acta,* 823**,** 51-60.

AGUIRRE-EZKAURIATZA, E. J., GALARZA-GONZALEZ, M. G., URIBE-BUJANDA, A. I., RIOS-LICEA, M., LOPEZ-PACHECO, F., HERNANDEZ-BRENES, C. M. & ALVAREZ, M. M. 2008. Effect of mixing during fermentation in yogurt manufacturing. *J Dairy Sci,* 91**,** 4454-65.

AHDESMAKI, M., ZUBER, V., GIBB, S. & STRIMMER, K. 2015. sda: Shrinkage Discriminant Analysis and CAT Score Variable Selection.

AHTESH, F. B., STOJANOVSKA, L. & APOSTOLOPOULOS, V. 2018. Processing and sensory characteristics of a fermented low-fat skim milk drink containing bioactive antihypertensive peptides, a functional milk product. *International Journal of Dairy Technology,* 71**,** 230-239.

AICHE, S., REINERT, K., SCHUTTE, C., HILDEBRAND, D., SCHLUTER, H. & CONRAD, T. O. 2012. Inferring proteolytic processes from mass spectrometry time series data using degradation graphs. *PLoS One,* 7**,** e40656.

ALIM, A., SONG, H., RAZA, A. & HUA, J. 2020. Identification of bitter constituents in milk-based infant formula with hydrolysed milk protein through a sensory-guided technique. *International Dairy Journal,* 110.

ALLGEYER, L. C., MILLER, M. J. & LEE, S. Y. 2010. Drivers of liking for yogurt drinks with prebiotics and probiotics. *J Food Sci,* 75**,** S212-9.

AMORIM, F. G., COITINHO, L. B., DIAS, A. T., FRIQUES, A. G. F., MONTEIRO, B. L., REZENDE, L. C. D., PEREIRA, T. M. C., CAMPAGNARO, B. P., DE PAUW, E., VASQUEZ, E. C. & QUINTON, L. 2019. Identification of new bioactive peptides from Kefir milk through proteopeptidomics: Bioprospection of antihypertensive molecules. *Food Chem,* 282**,** 109-119.

ANDERSEN, L. T., ARDÖ, Y. & BREDIE, W. L. P. 2010. Study of taste-active compounds in the water-soluble extract of mature Cheddar cheese. *International Dairy Journal,* 20**,** 528-536.

ANDERSEN, L. T., SCHLICHTHERLE-CERNY, H. & ARDÖ, Y. 2008. Hydrophilic di- and tripeptides are not a precondition for savoury flavour in mature Cheddar cheese. *Dairy Science and Technology,* 88**,** 467-475.

ANONYMOUS 1975. Minutes of Division Business Meeting. *In:* DIVISION, I. O. F. T. S. E. (ed.). Chicago, IL.: IFT.

ARAI, S., YAMASHITA, M. & FUJIMAKI, M. 1972. Glutamyl oligopeptides as factors responsible for tastes of a proteinase-modified soybean protein. *Agricultural and Biological Chemistry,* 36**,** 1253-1256.

ARDÖ, Y. & VARMING, C. 2010. Bacterial influence on characteristic flavour of cheeses made with mesophilic DL-starter. *Australian journal of dairy technology,* 65**,** 153.

ARENA, K., RIGANO, F., MANGRAVITI, D., CACCIOLA, F., OCCHIUTO, F., DUGO, L., DUGO, P. & MONDELLO, L. 2020. Exploration of Rapid Evaporative-Ionization Mass Spectrometry as a Shotgun Approach for the Comprehensive Characterization of Kigelia Africana (Lam) Benth. Fruit. *Molecules,* 25.

ARYANA, K. J. & OLSON, D. W. 2017. A 100-Year Review: Yogurt and other cultured dairy products. *J Dairy Sci,* 100**,** 9987-10013.

BALOG, J., PERENYI, D., GUALLAR-HOYAS, C., EGRI, A., PRINGLE, S. D., STEAD, S., CHEVALLIER, O. P., ELLIOTT, C. T. & TAKATS, Z. 2016. Identification of the Species of Origin for Meat Products by Rapid Evaporative Ionization Mass Spectrometry. *J Agric Food Chem,* 64**,** 4793-800.

BALOG, J., SZANISZLO, T., SCHAEFER, K. C., DENES, J., LOPATA, A., GODORHAZY, L., SZALAY, D., BALOGH, L., SASI-SZABO, L., TOTH, M. & TAKATS, Z. 2010. Identification of biological tissues by rapid evaporative ionization mass spectrometry. *Analytical chemistry,* 82**,** 7343-7350.

BAPTISTA, D. P., ARAUJO, F. D., EBERLIN, M. N. & GIGANTE, M. L. 2017. A Survey of the Peptide Profile in Prato Cheese as Measured by MALDI-MS and Capillary Electrophoresis. *J Food Sci,* 82**,** 386-393.

BARLOW, R. S., FITZGERALD, A. G., HUGHES, J. M., MCMILLAN, K. E., MOORE, S. C., SIKES, A. L., TOBIN, A. B. & WATKINS, P. J. 2021. Rapid Evaporative Ionization Mass Spectrometry: A Review on Its Application to the Red Meat Industry with an Australian Context. *Metabolites,* 11.

BARNES, D. L., HARPER, S. J., BODYFELT, F. W. & MCDANIEL, M. R. 1991a. Correlation of descriptive and consumer panel flavor ratings for commercial prestirred strawberry and lemon yogurts. *Journal of dairy science,* 74**,** 2089-2099.

BARNES, D. L., HARPER, S. J., BODYFELT, F. W. & MCDANIEL, M. R. 1991b. Prediction of Consumer Acceptability of Yogurt by Sensory and Analytical Measures of Sweetness and Sourness. *Journal of Dairy Science,* 74**,** 3746-3754.

BARTOSHUK, L. M., DUFFY, V. B. & MILLER, I. J. 1994. PTC/PROP tasting: anatomy, psychophysics, and sex effects. *Physiology & behavior,* 56**,** 1165-1171.

BAYARRI, S., CARBONELL, I., BARRIOS, E. X. & COSTELL, E. 2011. Impact of sensory differences on consumer acceptability of yoghurt and yoghurt-like products. *International Dairy Journal,* 21**,** 111-118.

BEAL, C., SKOKANOVA, J., LATRILLE, E., MARTIN, N. & CORRIEU, G. 1999. Combined effects of culture conditions and storage time on acidification and viscosity of stirred yogurt. *Journal of Dairy Science,* 82**,** 673-681.

BEHRENS, M., MEYERHOF, W., HELLFRITSCH, C. & HOFMANN, T. 2011. Sweet and umami taste: natural products, their chemosensory targets, and beyond. *Angew Chem Int Ed Engl,* 50**,** 2220-42.

BENJAMINI, Y. & HOCHBERG, Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B,* 57**,** 289-300.

BESHKOVA, D., SIMOVA, E., FRENGOVA, G. & SIMOV, Z. 1998. Production of flavour compounds by yogurt starter cultures. *Journal of Industrial Microbiology and Biotechnology,* 20**,** 180-186.

BLACK, C., POILE, C., LANGLEY, J. & HERNIMAN, J. 2006. The use of pencil lead as a matrix and calibrant for matrix-assisted laser desorption/ionisation. *Rapid Commun Mass Spectrom,* 20**,** 1053-60.

BLANCO, J. L., PORTO-PAZOS, A. B., PAZOS, A. & FERNANDEZ-LOZANO, C. 2018. Prediction of high anti-angiogenic activity peptides in silico using a generalized linear model and feature selection. *Sci Rep,* 8**,** 15688.

BLASCO, H., BLASZCZYNSKI, J., BILLAUT, J. C., NADAL-DESBARATS, L., PRADAT, P. F., DEVOS, D., MOREAU, C., ANDRES, C. R., EMOND, P., CORCIA, P. & SLOWINSKI, R. 2015. Comparative analysis of targeted metabolomics: dominance-based rough set approach versus orthogonal partial least square-discriminant analysis. *J Biomed Inform,* 53**,** 291-9.

BØ, T. H. & JONASSEN, I. 2002. New feature subset selection procedures for classification of expression profiles. *Genome biology,* 3**,** 1-11.

BOUWMEESTER, R., MARTENS, L. & DEGROEVE, S. 2019. Comprehensive and Empirical Evaluation of Machine Learning Algorithms for Small Molecule LC Retention Time Prediction. *Anal Chem,* 91**,** 3694-3703.

BREIMAN, L. 2001. Random forests. *Maching Learning,* 45**,** 5-32.

BRIAND, L. & SALLES, C. 2016. Taste perception and integration. *Flavor.* Woodhead Publishing.

BROADBENT, J. R., BARNES, M., BRENNAND, C., STRICKLAND, M., HOUCK, K., JOHNSON, M. E. & STEELE, J. L. 2002. Contribution of Lactococcus lactis Cell Envelope Proteinase Specificity to Peptide Accumulation and Bitterness in Reduced-Fat Cheddar Cheese. *Appl Environ Microbiol,* 68**,** 1778-1785.

BRONZEL, J. L., JR., MILAGRE, C. D. F. & MILAGRE, H. M. S. 2017. Analysis of low molecular weight compounds using MALDI- and LDI-TOF-MS: Direct detection of active pharmaceutical ingredients in different formulations. *J Mass Spectrom,* 52**,** 752-758.

BRUKER-DALTONICS. 2015. *Instructions for Use Bruker Guide to MALDI Sample Preparation* [Online]. Available: https://www.bruker.com/fileadmin/user_upload/8-PDF-Docs/Separations_MassSpectrometry/InstructionForUse/8702557_IFU_Bruker_Guide_MALDI_Sample_Preparation_Revision_E.pdf [Accessed].

BURNHAM, K. P. & ANDERSON, D. R. 1998. Practical use of the information-theoretic approach. *In:* SPRINGER (ed.) *Model selection and inference* New York.

CADWALLADER, K. R. & SINGH, T. K. 2009. Flavours and Off-Flavours in Milk and Dairy Products. 631-690.

CALVANO, C. D., DE CEGLIE, C., ARESTA, A., FACCHINI, L. A. & ZAMBONIN, C. G. 2013a. MALDI-TOF mass spectrometric determination of intact phospholipids as markers of illegal bovine milk adulteration of high-quality milk. *Anal Bioanal Chem,* 405**,** 1641-9.

CALVANO, C. D., MONOPOLI, A., LOIZZO, P., FACCIA, M. & ZAMBONIN, C. 2013b. Proteomic approach based on MALDI-TOF MS to detect powdered milk in fresh cow's milk. *J Agric Food Chem,* 61**,** 1609-17.

CHALLIS, R. C. & MA, M. 2016. Sour taste finds closure in a potassium channel. *Proc Natl Acad Sci U S A,* 113**,** 246-7.

CHANDAN, R. C. 2014. Dairy–fermented products. *Food Processing: Principles and Applications.*

CHAVES, A. C. S. D., FERNANDEZ, M., LERAYER, A. L. S., MIERAU, I., KLEEREBEZEM, M. & HUGENHOLTZ, J. 2002. Metabolic Engineering of Acetaldehyde Production by Streptococcus thermophilus. *Appl Environ Microbiol,* 68**,** 5656-5662.

CHEN, C., ZHAO, S., HAO, G., YU, H., TIAN, H. & ZHAO, G. 2017. Role of lactic acid bacteria on the yogurt flavour: A review. *International Journal of Food Properties,* 20**,** S316-S330.

CHEN, Q. Y., ALARCON, S., THARP, A., AHMED, O. M., ESTRELLA, N. L., GREENE, T. A., RUCKER, J. & BRESLIN, P. A. 2009. Perceptual variation in umami taste and polymorphisms in TAS1R taste receptor genes. *Am J Clin Nutr,* 90**,** 770S-779S.

CHEN, S., CHEN, L., WANG, J., HOU, J., HE, Q., LIU, J., XIONG, S., YANG, G. & NIE, Z. 2012. 2,3,4,5-Tetrakis(3',4'-dihydroxylphenyl)thiophene: a new matrix for the selective analysis of low molecular weight amines and direct determination of creatinine in urine by MALDI-TOF MS. *Anal Chem,* 84**,** 10291-7.

CHEN, T. & GUESTRIN, C. 2016. XGBoost. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.*

CHEN, T., HE, T., BENESTY, M., KHOTILOVICH, V., TANG, Y., CHO, H., CHEN, K., MITCHELL, R., IGNACIO C., ZHOU, T., LI M., XIE, J., LIN, M., GENG, Y. & LI, Y. 2021. xgboost: Extreme Gradient Boosting.

CHENG, H. 2010. Volatile flavor compounds in yogurt: a review. *Crit Rev Food Sci Nutr,* 50**,** 938-50.

CHICCO, D. & JURMAN, G. 2020. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC genomics,* 21**,** 1-13.

CHOLLET, M., GILLE, D., SCHMID, A., WALTHER, B. & PICCINALI, P. 2013. Acceptance of sugar reduction in flavored yogurt. *J Dairy Sci,* 96**,** 5501-11.

CIMANDER, C., CARLSSON, M. & MANDENIUS, C. F. 2002. Sensor fusion for on-line monitoring of yoghurt fermentation. *Journal of biotechnology,* 99**,** 237-248.

CIVILLE, G. V. & OFTEDAL, K. N. 2012. Sensory evaluation techniques—Make "good for you" taste "good". *Physiology & behavior,* 107**,** 598-605.

COHEN, L. H. & GUSEV, A. I. 2002. Small molecule analysis by MALDI mass spectrometry. *Anal Bioanal Chem,* 373**,** 571-86.

CONTI-SILVA, A. C. & SOUZA-BORGES, P. K. 2019. Sensory characteristics, brand and probiotic claim on the overall liking of commercial probiotic fermented milks: Which one is more relevant? *Food Res Int,* 116**,** 184-189.

CORETTA, S. 2021. tidymv: Tidy Model Visualisation for Generalised Additive Models.

COSTELLO, Z. & MARTIN, H. G. 2018. A machine learning approach to predict metabolic pathway dynamics from time-series multiomics data. *NPJ systems biology and applications,,* 4**,** 1-14.

DAGAN-WIENER, A., NISSIM, I., BEN ABU, N., BORGONOVO, G., BASSOLI, A. & NIV, M. Y. 2017. Bitter or not? BitterPredict, a tool for predicting taste from chemical structure. *Sci Rep,* 7**,** 12074.

DAHER, D., DERACINOIS, B., BANIEL, A., WATTEZ, E., DANTIN, J., FROIDEVAUX, R., CHOLLET, S. & FLAHAUT, C. 2020. Principal Component Analysis from Mass Spectrometry Data Combined to a Sensory Evaluation as a Suitable Method for Assessing Bitterness of Enzymatic Hydrolysates Produced from Micellar Casein Proteins. *Foods,* 9**,** 1354.

DAI, W. & JI, W. 2014. A MapReduce Implementation of C4.5 Decision Tree Algorithm. *International Journal of Database Theory and Application,* 7**,** 49-60.

DALABASMAZ, S., DITTRICH, D., KELLNER, I., DREWELLO, T. & PISCHETSRIEDER, M. 2019. Identification of peptides reflecting the storage of UHT milk by MALDI-TOF-MS peptide profiling. *J Proteomics,* 207**,** 103444.

DAN, T., WANG, D., JIN, R. L., ZHANG, H. P., ZHOU, T. T. & SUN, T. S. 2017. Characterization of volatile compounds in fermented milk using solid-phase microextraction methods coupled with gas chromatography-mass spectrometry. *J Dairy Sci,* 100**,** 2488-2500.

DANG, Y., GAO, X., MA, F. & WU, X. 2015. Comparison of umami taste peptides in water-soluble extractions of Jinhua and Parma hams. *LWT - Food Science and Technology,* 60**,** 1179-1186.

DAVE, K. A., HEADLAM, M. J., WALLIS, T. P. & GORMAN, J. J. 2011. Preparation and analysis of proteins and peptides using MALDI TOF/TOF mass spectrometry. *Curr Protoc Protein Sci,* Chapter 16**,** Unit 16 13.

DE HOFFMANN, E. & STROOBANT, V. 2007. *Mass Spectrometry: Principles and Applications*, Wiley-Interscience.

DE NONI, I. & CATTANEO, S. 2010. Occurrence of β-casomorphins 5 and 7 in commercial dairy products and in their digests following in vitro simulated gastro-intestinal digestion. *Food Chemistry,* 119**,** 560-566.

DESFONTAINE, V., GOYON, A., VEUTHEY, J. L., CHARVE, J. & GUILLARME, D. 2018. Development of a LC-MS/MS method for the determination of isomeric glutamyl peptides in food ingredients. *J Sep Sci,* 41**,** 847-855.

DESIMONE, J. A. & LYALL, V. 2006. Taste receptors in the gastrointestinal tract III. Salty and sour taste: sensing of sodium and protons by the tongue. *American Journal of Physiology-Gastrointestinal and Liver Physiology,* 291**,** 1005-1010.

DONG, X., CHENG, J., LI, J. & WANG, Y. 2010. Graphene as a novel matrix for the analysis of small molecules by MALDI-TOF MS. *Analytical Chemistry,* 82**,** 6208-14.

DRAKE, S. L., CARUNCHIA WHETSTINE, M. E., DRAKE, M. A., COURTNEY, P., FLIGNER, K., JENKINS, J. & PRUITT, C. 2007. Sources of umami taste in Cheddar and Swiss cheeses. *J Food Sci,* 72**,** S360-6.

DREWNOWSKI, A. & GOMEZ-CARNEROS, C. 2000. Bitter taste, phytonutrients, and the consumer: a review. *The American journal of clinical nutrition,* 72**,** 1424-1435.

DUNCAN, M. W., RODER, H. & HUNSUCKER, S. W. 2008. Quantitative matrix-assisted laser desorption/ionization mass spectrometry. *Brief Funct Genomic Proteomic,* 7**,** 355-70.

EBNER, J., ASCI ARSLAN, A., FEDOROVA, M., HOFFMANN, R., KUCUKCETIN, A. & PISCHETSRIEDER, M. 2015. Peptide profiling of bovine kefir reveals 236 unique peptides released from caseins during its production by starter culture or kefir grains. *J Proteomics,* 117**,** 41-57.

ENGEL, E., NICKLAUS, S., SEPTIER, C., SALLES, C. & LE QUERE, J. L. 2000. Taste active compounds in a goat cheese water-soluble extract. 2. Determination of the relative impact of water-soluble extract components on its taste using omission tests. *Journal of agricultural and food chemistry,* 48**,** 4260-4267.

ENGLAND, P., TANG, W., KOSTRZEWA, M., SHAHREZAEI, V. & LARROUY-MAUMUS, G. 2020. Discrimination of bovine milk from non-dairy milk by lipids fingerprinting using routine matrix-assisted laser desorption ionization mass spectrometry. *Sci Rep,* 10**,** 5160.

FALLICO, V., MCSWEENEY, P. L. H., HORNE, J., PEDILIGGIERI, C., HANNON, J. A., CARPINO, S. & LICITRA, G. 2005. Evaluation of bitterness in Ragusano cheese. *Journal of dairy science,* 88**,** 1288-1300.

FAO/WHO 2018. Standards for Fermented Milk CXS 243-2003. Adopted in 2003. Revised in 2008, 2010, 2018. *In:* ALIMENTARIUS, F. W. C. (ed.).

FERNANDEZ-GARCIA, E., LOPEZ-FANDINO, R., OLANO, A. & RAMOS, M. 1990. Comparative study of the proteolytic activity of a Bacillus subtilis neutral protease preparation during early stages of ripening of cheeses made from cow and ewe milk. . *Milchwissenschaft,* 45**,** 428-431.
FIJAN, S. 2014. Microorganisms with claimed probiotic properties: an overview of recent literature. *Int J Environ Res Public Health,* 11**,** 4745-67.
FOX, P. F., GUINEE, T. P., COGAN, T. M. & MCSWEENEY, P. L. 2017. *Fundamentals Of Cheese Science*, Springer Nature.
FRIEDMAN, J. 2001. Greedy function approximation: a gradient boosting machine. *Annals of Statistics,* 29**,** 1189-1232.
GALLARDO-ESCAMILLA, F. J., KELLY, A. L. & DELAHUNTY, C. M. 2005. Influence of starter culture on flavor and headspace volatile profiles of fermented whey and whey produced from fermented milk. *Journal of dairy science,* 88**,** 3745-3753.
GARCIA, J. S., SANVIDO, G. B., SARAIVA, S. A., ZACCA, J. J., COSSO, R. G. & EBERLIN, M. N. 2012. Bovine milk powder adulteration with vegetable oils or fats revealed by MALDI-QTOF MS. *Food Chemistry,* 131**,** 722-726.
GAY, S., BINZ, P. A., HOCHSTRASSER, D. F. & APPEL, R. D. 2002. Peptide mass fingerprinting peak intensity prediction: extracting knowledge from spectra. *Proteomics,* 2**,** 1374-1391.
GIBB, S. 2018. MALDIquantForeign: Import/Export Routines for 'MALDIquant'. R package version 0.11.5. .
GIBB, S. & STRIMMER, K. 2012. MALDIquant: a versatile R package for the analysis of mass spectrometry data. *Bioinformatics,* 28**,** 2270-2271.
GOMEZ-RUIZ, J. A., TABORDA, G., AMIGO, L., RAMOS, M. & MOLINA, E. 2007. Sensory and mass spectrometric analysis of the peptidic fraction lower than one thousand daltons in Manchego cheese. *J Dairy Sci,* 90**,** 4966-73.
GÜLER, Z. 2007. Changes in salted yoghurt during storage. *International Journal of Food Science & Technology,* 42**,** 235-245.
HANSEN, C. 2011. FD-DVS YF-L811 Yo-Flex® - Product Information. 3 PI-EU-EN 11-24-2011 ed.
HANSEN, C. 2016. CH-1 Product Information. 5 PI EU EN 06-28-2016 ed.
HARPER, S. J., BARNES, D. L., BODYFELT, F. W. & MCDANIEL, M. R. 1991. Sensory Ratings of Commercial Plain Yogurts by Consumer and Descriptive Panels. *Journal of dairy science,* 74**,** 2927-2935.
HASTIE, T., TIBSHIRANI, R. & FRIEDMAN, E. 2009. *The Elements of Statistical Learning*, Springer.
HAYES, J. E. & KEAST, R. S. 2011. Two decades of supertasting: where do we stand? *Physiol Behav,* 104**,** 1072-4.
HERNÁNDEZ-LEDESMA, B., GARCÍA-NEBOT, M. J., FERNÁNDEZ-TOMÉ, S., AMIGO, L. & RECIO, I. 2014. Dairy protein hydrolysates: Peptides for health benefits. *International Dairy Journal,* 38**,** 82-100.
HILLMANN, H. & HOFMANN, T. 2016. Quantitation of Key Tastants and Re-engineering the Taste of Parmesan Cheese. *J Agric Food Chem,* 64**,** 1794-805.
HORNIK, K., KARATZOGLOU, D. M., ZEILEIS, A. & HORNIK, M. K. 2007. The rweka package.
HOSSEINI, S. & MARTINEZ-CHAPA, S. O. 2016. *Fundamentals of MALDI-ToF-MS Analysis: Applications in Bio-diagnosis, Tissue Engineering and Drug Delivery*, Springer.
HRUSKAR, M., MAJOR, N., KRPAN, M. & VAHCIC, N. 2010. Simultaneous determination of fermented milk aroma compounds by a potentiometric sensor array. *Talanta,* 82**,** 1292-7.
HSSINA, B., MERBOUHA, A., EZZIKOURI, H. & ERRITALI, M. 2014. A comparative study of decision tree ID3 and C4. 5. *International Journal of Advanced Computer Science and Applications,* 4**,** 13-19.
HU, J., COOMBES, K. R., MORRIS, J. S. & BAGGERLY, K. A. 2005a. The importance of experimental design in proteomic mass spectrometry experiments: some cautionary tales. *Briefings in Functional Genomics,* 3**,** 322-331.
HU, L., XU, S., PAN, C., YUAN, C., ZOU, H. & JIANG, G. 2005b. Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry with a matrix of carbon nanotubes for the analysis of low-mass compounds in environmental samples. *Environmental science & technology,* 39**,** 8442-8447.

HUANG, J., WONG, K. H., TAY, S. V., HOW, A. & TAM, J. P. 2019. Cysteine-Rich Peptide Fingerprinting as a General Method for Herbal Analysis to Differentiate Radix Astragali and Radix Hedysarum. *Front Plant Sci,* 10**,** 973.

IKEDA, K. 1909. *J. Tokyo Chem. Soc.,* 30**,** 820-836.

IMHOF, R., GLAITTLI, H. & BOSSET, J. O. 1995. Volatile Organic Compounds Produced by Thermophilic and Mesophilic Single Strain Dairy Starter Cultures *Lebensm-Wiss-Technol,* 27**,** 442-449.

INZA, I., LARRANAGA, P., BLANCO, R. & CERROLAZA, A. J. 2004. Filter versus wrapper gene selection approaches in DNA microarray domains. *Artif Intell Med,* 31**,** 91-103.

ISHIBASHI, N., ARITA, Y., KANEHISA, H., KOUGE, K., OKAI, H. & FUKUI, S. 1987a. Bitterness of Leucine-containing Peptides. *Agricultural and Biological Chemistry,* 51**,** 2389-2394.

ISHIBASHI, N., KOUGE, K., SHINODA, I., KANEHISA, H. & OKAI, H. 1988a. A Mechanism for Bitter Taste Sensibility in Peptides. *Agricultural and Biological Chemistry,* 52**,** 819-827.

ISHIBASHI, N., KUBO, T., CHINO, M., FUKUI, H., SHINODA, I., KIKUCHI, E., OKAI, H. & FUKUI, S. 1988b. Taste of proline-containing peptides. *Agricultural and Biological Chemistry,* 52**,** 95-98.

ISHIBASHI, N., SADAMORI, K., YAMAMOTO, O., KANEHISA, H., KOUGE, K., KIKUCHI, E., OKAI, H. & FUKUI, S. 1987b. Bitterness of Phenylalanine- and Tyrosine-containing Peptides. *Agricultural and Biological Chemistry,* 51**,** 3309-3313.

IVANOV, M. V., BUBIS, J. A., GORSHKOV, V., ABDRAKHIMOV, D. A., KJELDSEN, F. & GORSHKOV, M. V. 2020. Boosting the MS1-only proteomics with machine learning allows 2000 protein identifications in 5-minute proteome analysis. *bioRxiv*.

JAMES, G., WITTEN, D., HASTIE, T. & TIBSHIRANI, R. 2013. *An introduction to statistical learning,* New York, Springer.

JAWORSKA, D., WASZKIEWICZ-ROBAK, B. O. Z. E. N. A., KOLANOWSKI, W. & SWIDERSKI, F. 2005. Relative importance of texture properties in the sensory quality and acceptance of natural yoghurts. *International Journal of Dairy Technology,* 58**,** 39-46.

JING, Z., HAO-YANG, W. & YIN-LONG, G. 2005. Amino acids analysis by MALDI mass spectrometry using carbon nanotube as matrix. *Chinese Journal of Chemistry,* 23**,** 185-189.

JUILLARD, V., BARS, D. L., KUNJI, E. R. S., KONINGS, W. N., GRIPON, J.-C. & RICHARD, J. 1995. Oligopeptides are the main source of nitrogen for Lactococcus lactis during growth in milk. *Appl. Environ. Microbiol,* 61**,** 3024-3030.

KADAM, A. & DESHMUKH, R. 2020. *Fermented Milk Market by Type (Cheese, Yogurt, Butter, Sour Cream, and Others) and Distribution Channel (Specialty Stores, Supermarket & Hypermarket, and Online Stores): Global Opportunity Analysis and Industry Forecast 2019–2026* [Online]. Available: https://www.alliedmarketresearch.com/fermented-milk-market-A05952 [Accessed].

KANEKO, S., KUMAZAWA, K. & NISHIMURA, O. 2011. Isolation and identification of the umami enhancing compounds in Japanese soy sauce. *Bioscience, biotechnology, and biochemistry***,** 1105282507-1105282507.

KANG, M. J., THOLEY, A. & HEINZLE, E. 2001. Application of automated matrix-assisted laser desorption ionization time-of-flight mass spectrometry for the measurement of enzyme activities. *Rapid Commun Mass Spectrom,* 15**,** 1327-1333.

KARAS, M. & HILLENKAMP, F. H. 1988. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Analytical chemistry,* 60**,** 2299-2301.

KASSAMBARA, A. & MUNDT , F. 2020. factoextra: Extract and Visualize the Results of Multivariate Data Analyses. R package version 1.0.5.

KEMP, S. E., HOLLOWOOD, T. & HORT, J. 2009. Planning your sensory project. *Sensory evaluation: a practical handbook.* John Wiley & Sons.

KEMPKA, M. 2005. *Improved mass accuracy in MALDI-TOF-MS analysis* PhD, Royal Institute of Technology Stockholm.

KĘSKA, P. & STADNIK, J. 2017. Taste-active peptides and amino acids of pork meat as components of dry-cured meat products: An in-silico study. *Journal of Sensory Studies,* 32**,** e12301.

KILARA, A. & PANYAM, D. 2003. Peptides from milk proteins and their properties. *Crit Rev Food Sci Nutr,* 43**,** 607-33.

KIM, H. O. & LI-CHAN, E. C. 2006. Quantitative structure− activity relationship study of bitter peptides. *Journal of agricultural and food chemistry,* 54**,** 10102-10111.
KIM, M. J., SON, H. J., KIM, Y., MISAKA, T. & RHYU, M. R. 2015. Umami–bitter interactions: The suppression of bitterness by umami peptides via human bitter taste receptor. *Biochemical and Biophysical Research Communications,* 456**,** 586-590.
KOHL, S., BEHRENS, M., DUNKEL, A., HOFMANN, T. & MEYERHOF, W. 2013. Amino acids and peptides activate at least five members of the human bitter taste receptor family. *J Agric Food Chem,* 61**,** 53-60.
KUHN, M. & JOHNSON, K. 2013. *Applied predictive modeling,* New York:, Springer.
KURODA, M., KATO, Y., YAMAZAKI, J., KAI, Y., MIZUKOSHI, T., MIYANO, H. & ETO, Y. 2012. Determination and quantification of gamma-glutamyl-valyl-glycine in commercial fish sauces. *J Agric Food Chem,* 60**,** 7291-6.
KURODA, M., KATO, Y., YAMAZAKI, J., KAI, Y., MIZUKOSHI, T., MIYANO, H. & ETO, Y. 2013. Determination and quantification of the kokumi peptide, gamma-glutamyl-valyl-glycine, in commercial soy sauces. *Food Chem,* 141**,** 823-8.
KURODA, M., SASAKI, K., YAMAZAKI, J., KATO, Y. & MIZUKOSHI, T. 2020. Quantification of the kokumi peptide, gamma-glutamyl-valyl-glycine, in cheese: Comparison between cheese made from cow and ewe milk. *J Dairy Sci,* 103**,** 7801-7807.
LANCASHIRE, L. J., LEMETRE, C. & BALL, G. R. 2009. An introduction to artificial neural networks in bioinformatics--application to complex microarray and mass spectrometry datasets in cancer studies. *Brief Bioinform,* 10**,** 315-29.
LANGLEY, G. J., HERNIMAN, J. M. & TOWNELL, M. S. 2007. 2B or not 2B, that is the question: further investigations into the use of pencil as a matrix for matrix-assisted laser desorption/ionisation. *Rapid Commun Mass Spectrom,* 21**,** 180-90.
LAWLOR, J. B., DELAHUNTY, C. M., WILKINSON, M. G. & SHEEHAN, J. 2002. Relationships between the gross, non-volatile and volatile compositions and the sensory attributes of eight hard-type cheeses. *International Dairy Journal,* 12**,** 493–509.
LE, S., JOSSE, J. & HUSSON, F. 2008. FactoMineR: An R Package for Multivariate Analysis. *Journal of Statistical Software,* 25**,** 1-18.
LEDELL, E., GILL, N., AIELLO, S., FU, A., CANDEL, A., CLICK, C., KRALJEVIC, T., NYKODYM, T., ABOYOUN, P., KURKA, M. & MALOHLAVA, M. 2020. h2o: R Interface for the 'H2O' Scalable Machine Learning Platform.
LEE, K. D., LO, C. G. & WARTHESEN, J. J. 1996. Removal of Bitterness from the Bitter Peptides Extracted from Cheddar Cheese with Peptidases from Lactococcus lactis ssp. cremoris SK111. *Journal of Dairy Science,* 79**,** 1521-1528.
LEE, W. J. & LUCEY, J. A. 2003. Rheological properties, whey separation, and microstructure in set-style yogurt: Effects of heating temperature and incubation temperature. *Journal of Texture Studies,* 34**,** 515-536.
LEE, W. J. & LUCEY, J. A. 2010. Formation and Physical Properties of Yogurt. *Asian-Aust. J. Anim. Sci.,* 23**,** 1127-1136.
LEMIEUX, L. & SIMARD, R. E. 1991. Bitter flavour in dairy products. I. A review of the factors likely to influence its development, mainly in cheese manufacture. *Le Lait,* 71**,** 599-636.
LEROY, F. & DE VUYST, L. 2004. Lactic acid bacteria as functional starter cultures for the food fermentation industry. *Trends in Food Science & Technology,* 15**,** 67-78.
LETORT, C., NARDI, M., GARAULT, P., MONNET, V. & JUILLARD, V. 2002. Casein Utilization by Streptococcus thermophilus Results in a Diauxic Growth in Milk. *Appl Environ Microbiol,* 68**,** 3162-3165.
LI, B., LIN, Y., YU, W., WILSON, D. I. & YOUNG, B. R. 2020a. Application of mechanistic modelling and machine learning for cream cheese fermentation pH prediction. *Journal of Chemical Technology & Biotechnology,* 96**,** 125-133.
LI, D., ZHENG, Y., KWOK, L. Y., ZHANG, W. & SUN, T. 2020b. Metabolic footprinting revealed key biochemical changes in a brown fermented milk product using Streptococcus thermophilus. *J Dairy Sci,* 103**,** 2128-2138.
LI, Y. L. & GROSS, M. L. 2004. Ionic-liquid matrices for quantitative analysis by MALDI-TOF mass spectrometry. *J Am Soc Mass Spectrom,* 15**,** 1833-7.

LIN, Y. M., CHEN, C. T. & CHANG, J. M. 2019. MS2CNN: predicting MS/MS spectrum based on protein sequence using deep convolutional neural networks. *BMC Genomics,* 20**,** 906.
LINDEN, G. & LORIENT, D. 1999. *New ingredients in food processing: biochemistry and agriculture*, Woodhead Publishing.
LIU, J., LIU, Y., GAO, M. & ZHANG, X. 2012. High throughput detection of tetracycline residues in milk using graphene or graphene oxide as MALDI-TOF MS matrix. *Journal of the American Society for Mass Spectrometry,* 23**,** 1424-7.
LIU, J., SONG, H., LIU, Y., LI, P., YAO, J. & XIONG, J. 2015. Discovery of kokumi peptide from yeast extract by LC-Q-TOF-MS/MS and sensomics approach. *J Sci Food Agric,* 95**,** 3183-94.
LIU, M., NAUTA, A., FRANCKE, C. & SIEZEN, R. J. 2008. Comparative Genomics of Enzymes in Flavor-Forming Pathways from Amino Acids in Lactic Acid Bacteria. *Applied and Environmental Microbiology,* 74**,** 4590-4600.
LIU, Y., LIU, J., DENG, C. & ZHANG, X. 2011. Graphene and graphene oxide: two ideal choices for the enrichment and ionization of long-chain fatty acids free from matrix-assisted laser desorption/ionization matrix interference. *Rapid Communications in Mass Spectrometry,* 25**,** 3223-34.
LÓPEZ-FERNÁNDEZ, H., SANTOS, H. M., CAPELO, J. L., FDEZ-RIVEROLA, F., GLEZ-PEÑA, D. & REBOIRO-JATO, M. 2015. Mass-Up: an all-in-one open software application for MALDI-TOF mass spectrometry knowledge discovery. *BMC Bioinformatics,* 16**,** 1471-2105.
LUCEY, J. A. & SINGH, H. 1997. Formation and physical properties of acid milk gels: a review. *Food Research International,* 30**,** 529-542.
LUCEY, J. A., TEO, C. T., MUNRO, P. A. & SINGH, H. 1998. Microstructure, permeability and appearance of acid gels made from heated skim milk. *Food Hydrocolloids,* 12**,** 159-165.
MA, C., REN, Y., YANG, J., REN, Z., YANG, H. & LIU, S. 2018. Improved Peptide Retention Time Prediction in Liquid Chromatography through Deep Learning. *Anal Chem,* 90**,** 10881-10888.
MAEHASHI, K., MATSUZAKI, M., YAMAMOTO, Y. & UDAKA, S. 1999. Isolation of Peptides from an Enzymatic Hydrolysate of Food Proteins and Characterization of Their Taste Properties. *Biosci Biotechnol Biochem,* 63**,** 555-559.
MAINS, T. P., PAYNE, F. A. & SAMA, M. P. 2017. Monitoring Yogurt Culture Fermentation and Predicting Fermentation Endpoint with Fluorescence Spectroscopy. *Transactions of the ASABE,* 60**,** 529-536.
MANSO, M. A., LÉONIL, J., JAN, G. & GAGNAIRE, V. 2005. Application of proteomics to the characterisation of milk and dairy products. *International Dairy Journal,* 15**,** 845-855.
MARGULIS, E., DAGAN-WIENER, A., IVES, R. S., JAFFARI, S., SIEMS, K. & NIV, M. Y. 2021. Intense bitterness of molecules: Machine learning for expediting drug discovery. *Comput Struct Biotechnol J,* 19**,** 568-576.
MARTÍNEZ-MAQUEDA, D., HERNÁNDEZ-LEDESMA, B., AMIGO, L., MIRALLES, B. & GÓMEZ-RUIZ, J. Á. 2013. Extraction/Fractionation Techniques for Proteins and Peptides and Protein Digestion. 21-50.
MARTINS, N., OLIVEIRA, M. B. P. P. & FERREIRA, I. C. F. R. 2018. Development of Functional Dairy Foods. *Sweeteners.*
MAZUR, R. H., SCHLATTER, J. M. & GOLDKAMP, A. H. 1969. Structure-taste relationships of some dipeptides. *Journal of the American Chemical Society,* 91**,** 2684-2691.
MCLAFFERTY, F. W. 1983. *Tandem Mass Spectrometry,* New York, JohnWiley & Sons, Inc.
MEILGAARD, M. C., CARR, B. T. & CIVILLE, G. V. 1999. *Sensory evaluation techniques*, CRC press.
MICHON, C., O'SULLIVAN, M. G., DELAHUNTY, C. M. & KERRY, J. P. 2009. The Investigation of Gender-Related Sensitivity Differences in Food Perception. *Journal of Sensory Studies,* 24**,** 922-937.
MISCHAK, H., VLAHOU, A. & IOANNIDIS, J. P. 2013. Technical aspects and inter-laboratory variability in native peptide profiling: the CE-MS experience. *Clin Biochem,* 46**,** 432-43.
MOLLER, C., BOCKELMANN, W., AMMANN, A. & HELLER, K. J. 2007. Production of yoghurt with mild taste by a Lactobacillus delbrueckii subsp. bulgaricus mutant with altered proteolytic properties. *Biotechnol J,* 2**,** 469-79.

MONTELEONE, E., SPINELLI, S., DINNELLA, C., ENDRIZZI, I., LAUREATI, M., PAGLIARINI, E., SINESIO, F., GASPERI, F., TORRI, L., APREA, E. & BAILETTI, L. I. 2017. Exploring influences on food choice in a large population sample: The Italian Taste project. *Food Quality and Preference,* 59**,** 123-140.

MOROS, G., CHATZIIOANNOU, A. C., GIKA, H. G., RAIKOS, N. & THEODORIDIS, G. 2017. Investigation of the derivatization conditions for GC–MS metabolomics of biological samples. *Bioanalysis,* 9**,** 53-65.

MOZZI, F., ORTIZ, M. E., BLECKWEDEL, J., DE VUYST, L. & PESCUMA, M. 2013. Metabolomics as a tool for the comprehensive understanding of fermented and functional foods with lactic acid bacteria. *Food Research International,* 54**,** 1152-1161.

MU, F., GU, Y., ZHANG, J. & ZHANG, L. 2020. Milk Source Identification and Milk Quality Estimation Using an Electronic Nose and Machine Learning Techniques. *Sensors (Basel),* 20.

MUNDRA, P. A. & RAJAPAKSE, J. C. 2016. Gene and sample selection using T-score with sample selection. *J Biomed Inform,* 59**,** 31-41.

MURRAY, N. M., O'RIORDAN, D., JACQUIER, J. C., O'SULLIVAN, M., HOLTON, T. A., WYNNE, K., ROBINSON, R. C., BARILE, D., NIELSEN, S. D. & DALLAS, D. C. 2018. Peptidomic screening of bitter and nonbitter casein hydrolysate fractions for insulinogenic peptides. *J Dairy Sci,* 101**,** 2826-2837.

NATEKIN, A. & KNOLL, A. 2013. Gradient boosting machines, a tutorial. *Front Neurorobot,* 7**,** 21.

NELSON, G., CHANDRASHEKAR, J., HOON, M. A., FENG, L., ZHAO, G., RYBA, N. J. & ZUKER, C. S. 2002. An amino-acid taste receptor. *Nature,* 416**,** 199.

NESTRUD, M. A. & LAWLESS, H. T. 2010. Perceptual Mapping of Apples and Cheeses Using Projective Mapping and Sorting. *Journal of Sensory Studies,* 25**,** 390-405.

NEWMAN, J., EGAN, T., HARBOURNE, N., O'RIORDAN, D., JACQUIER, J. C. & O'SULLIVAN, M. 2014. Correlation of sensory bitterness in dairy protein hydrolysates: Comparison of prediction models built using sensory, chromatographic and electronic tongue data. *Talanta,* 126**,** 46-53.

NEY, K. H. 1971. Prediction of bitterness of peptides from their amino acid composition. *Lebensm Unters Forsch.,* 147**,** 64-71.

NGUYEN, D. D., SOLAH, V. A., JOHNSON, S. K., CHARROIS, J. W. & BUSETTI, F. 2014. Isotope dilution liquid chromatography-tandem mass spectrometry for simultaneous identification and quantification of beta-casomorphin 5 and beta-casomorphin 7 in yoghurt. *Food Chem,* 146**,** 345-52.

NGUYEN, H., BUI, X.-N., BUI, H.-B. & CUONG, D. T. 2019. Developing an XGBoost model to predict blast-induced peak particle velocity in an open-pit mine: a case study. *Acta Geophysica,* 67**,** 477-490.

NGUYEN, H. T. H., AFSAR, S. & DAY, L. 2018. Differences in the microstructure and rheological properties of low-fat yoghurts from goat, sheep and cow milk. *Food Res Int,* 108**,** 423-429.

NIELSEN, S. D., JANSSON, T., LE, T. T., JENSEN, S., EGGERS, N., RAUH, V., SUNDEKILDE, U. K., SØRENSEN, J., ANDERSEN, H. J., BERTRAM, H. C. & LARSEN, L. B. 2017. Correlation between sensory properties and peptides derived from hydrolysed-lactose UHT milk during storage. *International Dairy Journal,* 68**,** 23-31.

NOGUCHI, M., ARAI, S., YAMASHITA, M., KATO, H. & FUJIMAKI, M. 1975. Isolation and identification of acidic oligopeptides occurring in a flavor potentiating fraction from a fish protein hydrolysate. *Journal of agricultural and food chemistry,* 23**,** 49-53.

O'MAHONY, M. & ROUSSEAU, B. 2003. Discrimination testing: a few ideas, old and new. *Food Quality and Preference,* 14**,** 157-164.

O'ROURKE, M. B., DJORDJEVIC, S. P. & PADULA, M. P. 2018. The quest for improved reproducibility in MALDI mass spectrometry. *Mass Spectrom Rev,* 37**,** 217-228.

OHSU, T., AMINO, Y., NAGASAKI, H., YAMANAKA, T., TAKESHITA, S., HATANAKA, T., MARUYAMA, Y., MIYAMURA, N. & ETO, Y. 2010. Involvement of the calcium-sensing receptor in human taste perception. *J Biol Chem,* 285**,** 1016-22.

OKSANEN, J., BLANCHET, F. G., KINDT, R., LEGENDRE, P., MINCHIN, P. R., O'HARA, R. B., SIMPSON, G. L., SOLYMOS, P., STEVENS, M. H. H., WAGNER, H. & OKSANEN, M. J. 2013. Package 'vegan'. Community ecology package, R package version 2.5-6. 2**,** 1-295.

OLIVEIRA, R. P. D. S., PEREGO, P., DE OLIVEIRA, M. N. & CONVERTI, A. 2012. Growth, organic acids profile and sugar metabolism of Bifidobacterium lactis in co-culture with Streptococcus thermophilus: The inulin effect. *Food Research International,* 48**,** 21-27.

ONG, L. & SHAH, N. P. 2008. Release and identification of angiotensin-converting enzyme-inhibitory peptides as influenced by ripening temperatures and probiotic adjuncts in Cheddar cheeses. *LWT - Food Science and Technology,* 41**,** 1555-1566.

OPGEN-RHEIN, R. & STRIMMER, K. 2007. Accurate ranking of differentially expressed genes by a distribution-free shrinkage approach. *Statistical applications in genetics and molecular biology,* 6.

ØSTLIE, H. M., HELLAND, M. H. & NARVHUS, J. A. 2003. Growth and metabolism of selected strains of probiotic bacteria in milk. *International Journal of Food Microbiology,* 87**,** 17-27.

ØSTLIE, H. M., TREIMO, J. & NARVHUS, J. A. 2005. Effect of temperature on growth and metabolism of probiotic bacteria in milk. *International Dairy Journal,* 15**,** 989-997.

OTT, A., FAY, L. B. & CHAINTREAU, A. 1997. Determination and origin of the aroma impact compounds of yogurt flavor. *J. Agric. Food Chem,* 45**,** 850-858.

OTT, A., HUGI, A., BAUMGARTNER, M. & CHAINTREAU, A. 2000. Sensory Investigation of Yogurt Flavor Perception Mutual influene of volatiles and acidity. *48***,** 441-450.

ÖZER, B. H. & KIRMACI, H. A. 2010. Functional milks and dairy beverages. *International Journal of Dairy Technology,* 63**,** 1-15.

PAN, C., XU, S., HU, L., SU, X., OU, J., ZOU, H., GUO, Z., ZHANG, Y. & GUO, B. 2005. Using oxidized carbon nanotubes as matrix for analysis of small molecules by MALDI-TOF MS. *J Am Soc Mass Spectrom,* 16**,** 883-92.

PAPADIMITRIOU, C., VAFOPOULOU-MASTROJIANNAKI, A., SILVA, S., GOMES, A., MALCATA, F. & ALICHANIDIS, E. 2007. Identification of peptides in traditional and probiotic sheep milk yoghurt with angiotensin I-converting enzyme (ACE)-inhibitory activity. *Food Chemistry,* 105**,** 647-656.

PAPAGIANNOPOULOU, C., PARCHEN, R., RUBBENS, P. & WAEGEMAN, W. 2020. Fast Pathogen Identification Using Single-Cell Matrix-Assisted Laser Desorption/Ionization-Aerosol Time-of-Flight Mass Spectrometry Data and Deep Learning Methods. *Anal Chem,* 92**,** 7523-7531.

PARK, J. N., ISHIDA, K., WATANABE, T., ENDOH, K. I., WATANABE, K., MURAKAMI, M. & ABE, H. 2002. Taste effects of oligopeptides in a Vietnamese fish sauce. . *Fisheries science,* 68**,** 921-928.

PARKER, L., ENGEL-HALL, A., DREW, K., STEINHARDT, G., HELSETH, D. L., JR., JABON, D., MCMURRY, T., ANGULO, D. S. & KRON, S. J. 2008. Investigating quantitation of phosphorylation using MALDI-TOF mass spectrometry. *J Mass Spectrom,* 43**,** 518-27.

PAXTON, T. 2020. Rapid evaporative ionization mass spectrometry. *Ambient Ionization Mass Spectrometry in Life Sciences.*

PEDERSEN, E. J., MILLER, D. L., SIMPSON, G. L. & ROSS, N. 2019. Hierarchical generalized additive models in ecology: an introduction with mgcv. *PeerJ,* 7**,** e6876.

PHELPS, D. L., BALOG, J., GILDEA, L. F., BODAI, Z., SAVAGE, A., EL-BAHRAWY, M. A., SPELLER, A. V., ROSINI, F., KUDO, H., MCKENZIE, J. S., BROWN, R., TAKATS, Z. & GHAEM-MAGHAMI, S. 2018. The surgical intelligent knife distinguishes normal, borderline and malignant gynaecological tissues using rapid evaporative ionisation mass spectrometry (REIMS). *Br J Cancer,* 118**,** 1349-1358.

PIRAS, C., HALE, O. J., REYNOLDS, C. K., JONES, A. K., TAYLOR, N., MORRIS, M. & CRAMER, R. 2021. Speciation and milk adulteration analysis by rapid ambient liquid MALDI mass spectrometry profiling using machine learning. *Sci Rep,* 11**,** 3305.

PLECHAWSKA-WOJCIK, M. 2012. A Comprehensive Analysis of MALDI-TOF Spectrometry Data. *Medical Informatics*.

POHJANHEIMO, T. & SANDELL, M. 2009. Explaining the liking for drinking yoghurt: The role of sensory quality, food choice motives, health concern and product information. *International Dairy Journal,* 19**,** 459-466.

PORTET, S. 2020. A primer on model selection using the Akaike Information Criterion. *Infect Dis Model,* 5**,** 111-128.

PRASANNA, P. H. P., GRANDISON, A. S. & CHARALAMPOPOULOS, D. 2014. Bifidobacteria in milk products: An overview of physiological and biochemical properties, exopolysaccharide production, selection criteria of milk products and health benefits. *Food Research International,* 55**,** 247-262.

QUINLAN, R. 1993. *C4.5: Programs for Machine Learning,* San Mateo, Morgan Kaufmann.

RADKE-MITCHELL , L. & SANDINE, W. 1984. Associative Growth and Differential Enumeration of Streptococcus thermophilus and Lactobacillus bulgaricus: A Review. *J. Food Prot,* 47**,** 245–248.

RAMSEY, I. S. & DESIMONE, J. A. 2018. Otopetrin-1: A sour-tasting proton channel. *J Gen Physiol,* 150**,** 379-382.

RAVINDRA, K., RATTAN, P., MOR, S. & AGGARWAL, A. N. 2019. Generalized additive models: Building evidence of air pollution, climate change and human health. *Environ Int,* 132**,** 104987.

REINECCIUS, G. 2006. *An overview of flavor perception. Flavor chemistry and technology*, Boca Raton, Fla., U.S.A.: Taylor and Francis.

ROCCHETTI, G., LUCINI, L., GALLO, A., MASOERO, F., TREVISAN, M. & GIUBERTI, G. 2018. Untargeted metabolomics reveals differences in chemical fingerprints between PDO and non-PDO Grana Padano cheeses. *Food research international,* 113**,** 407-413.

ROCHA, R. S., CALVALCANTI, R. N., SILVA, R., GUIMARÃES, J. T., BALTHAZAR, C. F., PIMENTEL, T. C., ESMERINO, E. A., FREITAS, M. Q., GRANATO, D., COSTA, R. G. B., SILVA, M. C. & CRUZ, A. G. 2020. Consumer acceptance and sensory drivers of liking of Minas Frescal Minas cheese manufactured using milk subjected to ohmic heating: Performance of machine learning methods. *Lwt,* 126.

ROGERS, L. 2017. *Discrimination testing in sensory science: A practical handbook*, Woodhead Publishing.

ROHART, F., GAUTIER, B., SINGH, A. & LE CAO, K. A. 2017. mixOmics: An R package for 'omics feature selection and multiple data integration. *PLoS Comput Biol,* 13**,** e1005752.

ROSS, A., BRUNIUS, C., CHEVALLIER, O., DERVILLY, G., ELLIOTT, C., GUITTON, Y., PRENNI, J. E., SAVOLAINEN, O., HEMERYCK, L., VIDKJAER, N. H., SCOLLAN, N., STEAD, S. L., ZHANG, R. & VANHAECKE, L. 2020. Making complex measurements of meat composition fast: Application of rapid evaporative ionisation mass spectrometry to measuring meat quality and fraud. *Meat Sci***,** 108333.

ROUDOT-ALGARON, F., KERHOAS, L., LE BARS, D., EINHORN, J. & GRIPON, J. C. 1994. Isolation of γ-glutamyl peptides from Comté cheese. *Journal of dairy science,* 77**,** 1161-1166.

ROUTRAY, W. & MISHRA, H. N. 2011. Scientific and Technical Aspects of Yogurt Aroma and Taste: A Review. *Comprehensive Reviews in Food Science and Food Safety,* 10**,** 208-220.

RUL, F. 2017. Yogurt: microbiology, organoleptic properties and probiotic potential. *Fermented Foods, Part II: Technological Interventions***,** 525.

SAITO, T., NAKAMURA, T., KITAZAWA, H., KAWAI, Y. & ITOH, T. 2000. Isolation and structural analysis of antihypertensive peptides that exist naturally in Gouda cheese. *Journal of dairy science,* 83**,** 1434-1440.

SALGER, M., STARK, T. D. & HOFMANN, T. 2019. Taste Modulating Peptides from Overfermented Cocoa Beans. *J Agric Food Chem*.

SALLES, C., SOMMERER, N., SEPTIER, C., ISSANCHOU, S., CHABANET, C., GAREM, A. & QUÉRÉ, J.-L. L. 2002. Goat Cheese Flavor: Sensory Evaluation of Branched-Chain Fatty Acids and Small Peptides. *Journal of Food Science* 67**,** 835-841.

SANDHU, A. K. & BATTH, R. S. 2020. Software reuse analytics using integrated random forest and gradient boosting machine learning algorithm. *Software: Practice and Experience,* 51**,** 735-747.

SANLIER, N., GOKCEN, B. B. & SEZGIN, A. C. 2019. Health benefits of fermented foods. *Crit Rev Food Sci Nutr,* 59**,** 506-527.

SANTOS, B. A., POLLONIO, M. A. R., CRUZ, A. G., MESSIAS, V. C., MONTEIRO, R. A., OLIVEIRA, T. L. C., FARIA, J. A. F., FREITAS, M. Q. & BOLINI, H. M. A. 2013. Ultra-flash profile and projective mapping for describing sensory attributes of prebiotic mortadellas. *Food Research International,* 54**,** 1705-1711.

SASSI, M., ARENA, S. & SCALONI, A. 2015. MALDI-TOF-MS Platform for Integrated Proteomic and Peptidomic Profiling of Milk Samples Allows Rapid Detection of Food Adulterations. *J Agric Food Chem,* 63**,** 6157-71.

SCHÄFER, J., SEBALD, K., DUNKEL, A., HOFMANN, T., ROSENTHAL, I., SCHUSTER, R., ATAMER, Z. & HINRICHS, J. 2019. A feasibility study on the pilot scale manufacture of fresh cheese from skim milk retentates without acid whey production: Effect of calcium content on bitterness and texture. *International Dairy Journal,* 93**,** 72-80.

SCHINDLER, A., DUNKEL, A., STAHLER, F., BACKES, M., LEY, J., MEYERHOF, W. & HOFMANN, T. 2011. Discovery of salt taste enhancing arginyl dipeptides in protein digests and fermented fish sauces by means of a sensomics approach. *J Agric Food Chem,* 59**,** 12578-88.

SCHLICHTHERLE-CERNY, H. & AMADÒ, R. 2002. Analysis of taste-active compounds in an enzymatic hydrolysate of deamidated wheat gluten. *Journal of agricultural and food chemistry,* 50**,** 1515-1522.

SCHLOERKE, B., CROWLEY, J. & COOK, D. 2021. Package 'GGally'.

SEBALD, K., DUNKEL, A. & HOFMANN, T. 2019. Mapping Taste-Relevant Food Peptidomes by Means of Sequential Window Acquisition of All Theoretical Fragment Ion-Mass Spectrometry. *J Agric Food Chem,* 68**,** 10287-10298.

SEBALD, K., DUNKEL, A., SCHAFER, J., HINRICHS, J. & HOFMANN, T. 2018. Sensoproteomics: A New Approach for the Identification of Taste-Active Peptides in Fermented Foods. *J Agric Food Chem,* 66**,** 11092-11104.

SENTANDREU, M. A., STOEVA, S., ARISTOY, M. A., LAIB, K., VOELTER, W. & TOLDRA, E. 2003. Identification of small peptides generated in Spanish dry-cured ham. *Journal of Food Science,* 68**,** 64-69.

SETTACHAIMONGKON, S., NOUT, M. J., ANTUNES FERNANDES, E. C., HETTINGA, K. A., VERVOORT, J. M., VAN HOOIJDONK, T. C., ZWIETERING, M. H., SMID, E. J. & VAN VALENBERG, H. J. 2014. Influence of different proteolytic strains of Streptococcus thermophilus in co-culture with Lactobacillus delbrueckii subsp. bulgaricus on the metabolite profile of set-yoghurt. *Int J Food Microbiol,* 177**,** 29-36.

SFORZA, S., CAVATORTA, V., LAMBERTINI, F., GALAVERNA, G., DOSSENA, A. & MARCHELLI, R. 2012. Cheese peptidomics: a detailed study on the evolution of the oligopeptide fraction in Parmigiano-Reggiano cheese from curd to 24 months of aging. *J Dairy Sci,* 95**,** 3514-26.

SFORZA, S., PIGAZZANI, A., MOTTI, M., PORTA, C., VIRGILI, R., GALAVERNA, G., DOSSENA, A. & MARCHELLI, R. 2001. Oligopeptides and free amino acids in Parma hams of known cathepsin B activity. *Food Chem,* 75**,** 267-273.

SHEN, Q., LI, L., SONG, G., FENG, J., LI, S., WANG, Y., MA, J. & WANG, H. 2020. Development of an intelligent surgical knife rapid evaporative ionization mass spectrometry based method for real-time differentiation of cod from oilfish. *Journal of Food Composition and Analysis,* 86.

SIEUWERTS, S., DE BOK, F. A., HUGENHOLTZ, J. & VAN HYLCKAMA VLIEG, J. E. 2008. Unraveling microbial interactions in food fermentations: from classical to genomics approaches. *Appl Environ Microbiol,* 74**,** 4997-5007.

SINGH, T. K., YOUNG, N. D., DRAKE, M. & CADWALLADER, K. R. 2005. Production and sensory characterization of a bitter peptide from β-casein. *Journal of agricultural and food chemistry,* 53**,** 1185-1189.

SMIT, G., SMIT, B. A. & ENGELS, W. J. 2005. Flavour formation by lactic acid bacteria and biochemical flavour profiling of cheese products. *FEMS Microbiol Rev,* 29**,** 591-610.

SODINI, I., REMEUF, F., HADDAD, S. & CORRIEU, G. 2004. The relative effect of milk base, starter, and process on yogurt texture: a review. *Crit Rev Food Sci Nutr,* 44**,** 113-37.

SOERYAPRANATA, E., POWERS, J. R., FAJARRINI, F., WELLER, K. M., HILL, H. H. & SIEMS, W. F. 2002. Relationship between MALDI-TOF analysis of β-CN f193− 209 concentration and sensory evaluation of bitterness intensity of aged Cheddar cheese. *Journal of agricultural and food chemistry,* 50**,** 4900-4905.

STACKLIES, W., REDESTIG, H., SCHOLZ, M., WALTHER, D. & SELBIG, J. 2007. pcaMethods—a bioconductor package providing PCA methods for incomplete data. *Bioinformatics,* 23**,** 1164-1167.

STANSTRUP, J., NEUMANN, S. & VRHOVSEK, U. 2015. PredRet: prediction of retention time by direct mapping between multiple chromatographic systems. *Anal Chem,* 87**,** 9421-8.

SUBRAMANIAN, A., HARPER, W. J. & RODRIGUEZ-SAONA, L. E. 2009. Rapid prediction of composition and flavor quality of cheddar cheese using ATR-FTIR spectroscopy. *J Food Sci,* 74**,** C292-7.

SZAJLI, E., FEHER, T. & MEDZIHRADSZKY, K. F. 2008. Investigating the quantitative nature of MALDI-TOF MS. *Mol Cell Proteomics,* 7**,** 2410-8.

TABORDA, G., GÓMEZ-RUIZ, J. A., MARTÍNEZ-CASTRO, I., AMIGO, L., RAMOS, M. & MOLINA, E. 2008. Taste and flavor of artisan and industrial Manchego cheese as influenced by the water-soluble extract compounds. *European Food Research and Technology,* 227**,** 323-330.

TAIT, P. R., RUTHERFORD, P., DRIVER, T., LI, X., SAUNDERS, C. M., DALZIEL, P. C. & GUENTHER, M. 2018. Consumer insights and willingness to pay for attributes: New Zealand yogurt products in Shanghai, China. *Lincoln University. AERU.*

TAMIME, A. Y. & ROBINSON, R. K. 1999. *Yoghurt Science and Technology* Cambridge, England, Woodhead Publishing.

TAMURA, M., MIYOSHI, T., MORI, N., KINOMURA, K., KAWAGUCHI, M., ISHIBASHI, N. & OKAI, H. 1990. Mechanism for the Bitter Tasting Potency of Peptides UsingO-Aminoacyl Sugars as Model Compounds+. *Agricultural and Biological Chemistry,* 54**,** 1401-1409.

TAMURA, M., NAKATSUKA, T., TADA, M., KAWASAKI, Y., KIKUCHI, E. & OKAI, H. 1989. The Relationship between Taste and Primary Structure of "Delicious Peptide" (Lys-Gly-Asp-Glu-Glu-Ser-Leu-Ala) from Beef Soup. *Agricultural and Biological Chemistry,* 53**,** 319-325.

TEAM., R. C. 2020. R: A language and environment for statistical computing. *R Foundation for Statistical Computing.* 3.6.2 ed. Vienna, Austria.

TEMIZKAN, R., CAN, A., DOGAN, M. A., MORTAS, M. & AYVAZ, H. 2020. Rapid detection of milk fat adulteration in yoghurts using near and mid-infrared spectroscopy. *International Dairy Journal,* 110.

TEMUSSI, P. A. 2012. The good taste of peptides. *J Pept Sci,* 18**,** 73-82.

THEODOROU, G. & POLITIS, I. 2016. Effects of peptides derived from traditional Greek yoghurt on expression of pro- and anti-inflammatory genes by ovine monocytes and neutrophils. *Food and Agricultural Immunology,* 27**,** 484-495.

THÉVENOT, E. A., ROUX, A., XU, Y., EZAN, E. & JUNOT, C. 2015. Analysis of the human adult urinary metabolome variations with age, body mass index, and gender by implementing a comprehensive workflow for univariate and OPLS statistical analyse. *Journal of proteome research.*

TIAN, H., SHEN, Y., YU, H., HE, Y. & CHEN, C. 2017. Effects of 4 Probiotic Strains in Coculture with Traditional Starters on the Flavor Profile of Yogurt. *Journal of Food Science,* 82**,** 1693-1701.

TIMM, W., SCHERBART, A., BÖCKER, S., KOHLBACHER, O. & NATTKEMPER, T. W. 2008. Peak intensity prediction in MALDI-TOF mass spectrometry: a machine learning study to support quantitative proteomics. *BMC bioinformatics,* 9**,** 1-18.

TOELSTEDE, S., DUNKEL, A. & HOFMANN, T. 2009. A Series of Kokumi Peptides Impart the Long-Lasting Mouthfulness of Matured Gouda Cheese. *J. Agric. Food Chem.,* 57**,** 1440-1448.

TOELSTEDE, S. & HOFMANN, T. 2008a. Quantitative studies and taste re-engineering experiments toward the decoding of the nonvolatile sensometabolome of Gouda cheese. *Journal of agricultural and food chemistry,* 56**,** 5299-5307.

TOELSTEDE, S. & HOFMANN, T. 2008b. Sensomics mapping and identification of the key bitter metabolites in Gouda cheese. *Journal of agricultural and food chemistry,* 56**,** 2795-2804.

TONG, D. L., BOOCOCK, D. J., COVENEY, C., SAIF, J., GOMEZ, S. G., QUEROL, S., REES, R. & BALL, G. R. 2011. A simpler method of preprocessing MALDI-TOF MS data for differential biomarker analysis: stem cell and melanoma cancer studies. *Clin Proteomics,* 8**,** 14.

TRAN, N. H., ZHANG, X., XIN, L., SHAN, B. & LI, M. 2017. De novo peptide sequencing by deep learning. *Proc Natl Acad Sci U S A,* 114**,** 8247-8252.
TU, Y. H., COOPER, A. J., TENG, B., CHANG, R. B., ARTIGA, D. J., TURNER, H. N., MULHALL, E. M., YE, W., SMITH, A. D. & LIMAN, E. R. 2018. An evolutionarily conserved gene family encodes proton-selective ion channels. *Science,* 359**,** 1047-1050.
UEDA, Y., YONEMITSU, M., TSUBUKU, T., SAKAGUCHI, M. & MIYAJIMA, R. 1997. Flavor characteristics of glutathione in raw and cooked foodstuffs. *Biosci Biotechnol Biochem,* 61**,** 1977-80.
VAN DEN OORD, A. H. A. & VAN WASSENAAR, P. D. 1997. Umami peptides: assessment of their alleged taste properties. *Z Lebensm Unters Forsch,* 205**,** 125-130.
VAN KRANENBURG, R., KLEEREBEZEM, M., VAN HYLCKAMA VLIEG, J., URSING, B. M., BOEKHORST, J., SMIT, B. A., AYAD, E. H., SMIT, G. & SIEZEN, R. J. 2002. Flavour formation from amino acids by lactic acid bacteria: predictions from genome sequence analysis. *International Dairy Journal,* 12**,** 111-121.
WANG, C. H., LI, J., YAO, S. J., GUO, Y. L. & XIA, X. H. 2007. High-sensitivity matrix-assisted laser desorption/ionization Fourier transform mass spectrometry analyses of small carbohydrates and amino acids using oxidized carbon nanotubes prepared by chemical vapor deposition as matrix. *Anal Chim Acta,* 604**,** 158-64.
WANG, H., CAO, X., HAN, T., PEI, H., REN, H. & STEAD, S. 2019. A novel methodology for real-time identification of the botanical origins and adulteration of honey by rapid evaporative ionization mass spectrometry. *Food Control,* 106.
WANG, L., NIU, D., WANG, X., SHEN, Q. & XUE, Y. 2020. A Novel Machine Learning Strategy for the Prediction of Antihypertensive Peptides Derived from Food with High Efficiency. *Foods,* 10**,** 550.
WASSERMAN, L. 2004. *All of statistics: a concise course in statistical inference.*, Springer Science & Business Media.
WEISBERG, S. 2013. *Applied linear regression*, John Wiley & Sons.
WEN, B., ZENG, W. F., LIAO, Y., SHI, Z., SAVAGE, S. R., JIANG, W. & ZHANG, B. 2020. Deep Learning in Proteomics. *Proteomics,* 20**,** e1900335.
WICKHAM, H. 2007. Reshaping Data with the reshape Package. *Journal of Statistical Software,* 21**,** 1-20.
WICKHAM, H. 2011. ggplot2. *Wiley Interdisciplinary Reviews: Computational Statistics,* 3**,** 180-185.
WICKHAM, H., AVERICK, M., BRYAN, J., CHANG, W., MCGOWAN, L. D. A., FRANÇOIS, R., GROLEMUND, G., HAYES, A., HENRY, L., HESTER, J. & KUHN, M. 2019. Welcome to the Tidyverse. *Journal of Open Source Software,* 4**,** 1686.
WIENER, A., SHUDLER, M., LEVIT, A. & NIV, M. Y. 2012. BitterDB: a database of bitter compounds. *Nucleic acids research,* 40**,** 413-419.
WINKLER, R. 2014. ASSyPup—an 'Out of the Box'solution for the analysis of mass spectrometry data. *Journal of Mass Spectrometry,* 49**,** 37-42.
WOOD, S. 2015. Package 'mgcv'. R package version, 1.
WORLEY, B. & POWERS, R. 2013. Multivariate Analysis in Metabolomics. *Curr Metabolomics,* 1**,** 92-107.
XIA, Y., YUAN, R., WENG, S., WANG, G., XIONG, Z., ZHANG, H., SONG, X., LIU, W. & AI, L. 2020. Proteolysis, lipolysis, texture and sensory properties of cheese ripened by Monascus fumeus. *Food Res Int,* 137**,** 109657.
XU, Q., SINGH, N., HONG, H., YAN, X., YU, W., JIANG, X., CHELIKANI, P. & WU, J. 2019. Hen protein-derived peptides as the blockers of human bitter taste receptors T2R4, T2R7 and T2R14. *Food chemistry,* 283**,** 621-627.
YADAV, V., GUPTA, V. K. & MEENA, G. S. 2018. Effect of culture levels, ultrafiltered retentate addition, total solid levels and heat treatments on quality improvement of buffalo milk plain set yoghurt. *J Food Sci Technol,* 55**,** 1648-1655.
YAMAMOTO, S., SHIGA, K., KODAMA, Y., IMAMURA, M., UCHIDA, R., OBATA, A., BAMBA, T. & FUKUSAKI, E. 2014. Analysis of the correlation between dipeptides and taste differences among soy sauces by using metabolomics-based component profiling. *J Biosci Bioeng,* 118**,** 56-63.

YAMASAKI, Y. & MAEKAWA, K. 1978. A Peptide with Delicious Taste. *Agricultural and Biological Chemistry,* 42**,** 1761-1765.

YANG, J., BAI, W., ZENG, X. & CUI, C. 2019. Gamma glutamyl peptides: the food source, enzymatic synthesis, kokumi-active and the potential functional properties–a review. *Trends in Food Science & Technology.,* 91**,** 339-346.

YANG, Q., WILLIAMSON, A. M., HASTED, A. & HORT, J. 2020. Exploring the relationships between taste phenotypes, genotypes, ethnicity, gender and taste perception using Chi-square and regression tree analysis. *Food Quality and Preference***,** 103928.

YILDIZ, A., ERDOGAN, S., SAYDUT, A. & HAMAMCI, C. 2011. High-Performance Liquid Chromatography Analysis and Assessment of Benzoic Acid in Yogurt, Ayran, and Cheese in Turkey. *Food Analytical Methods,* 5**,** 591-595.

ZABET-MOGHADDAM, M., HEINZLE, E. & THOLEY, A. 2004. Qualitative and quantitative analysis of low molecular weight compounds by ultraviolet matrix-assisted laser desorption/ionization mass spectrometry using ionic liquid matrices. *Rapid Commun Mass Spectrom,* 18**,** 141-8.

ZANIEWSKI, A. E., LEHMANN, A. & OVERTON, J. M. 2002. Predicting species spatial distributions using presence-only data a case study of native New Zealand ferns. *Ecological modelling,* 157**,** 261-280.

ZHANG, L., BORROR, C. M. & SANDRIN, T. R. 2014. A designed experiments approach to optimization of automated data acquisition during characterization of bacteria with MALDI-TOF mass spectrometry. *PLoS One,* 9**,** e92720.

ZHANG, Y., VENKITASAMY, C., PAN, Z., LIU, W. & ZHAO, L. 2017. Novel Umami Ingredients: Umami Peptides and Their Taste. *J Food Sci,* 82**,** 16-23.

ZHAO, C. J., SCHIEBER, A. & GANZLE, M. G. 2016. Formation of taste-active amino acids, amino acid derivatives and peptides in food fermentations - A review. *Food Res Int,* 89**,** 39-47.

ZHOU, X. X., ZENG, W. F., CHI, H., LUO, C., LIU, C., ZHAN, J., HE, S. M. & ZHANG, Z. 2017. pDeep: Predicting MS/MS Spectra of Peptides with Deep Learning. *Anal Chem,* 89**,** 12690-12697.

ZHUANG, M., LIN, L., ZHAO, M., DONG, Y., SUN-WATERHOUSE, D., CHEN, H., QIU, C. & SU, G. 2016. Sequence, taste and umami-enhancing effect of the peptides separated from soy sauce. *Food Chem,* 206**,** 174-81.

ZIELINSKI, B., PLICHTA, A., MISZTAL, K., SPUREK, P., BRZYCHCZY-WLOCH, M. & OCHONSKA, D. 2017. Deep learning approach to bacterial colony classification. *PLoS One,* 12**,** e0184554.

ZUMBÜHL, S., KNOCHENMUSS, R., WÜLFERT, S., DUBOIS, F., DALE, M. J. & ZENOBI, R. 1998. A graphite-assisted laser desorption/ionization study of light-induced aging in triterpene dammar and mastic varnishes. *Analytical Chemistry,* 70**,** 707-715.

Appendices

Chapter 2: Literature Review

Appendix 1

Sequence and fragment bitter peptides isolated from fermented dairy foods and evaluated for taste activity via sensory techniques

Peptide Sequence	Fragment	Food Type	Reference
ELEEL	β-CN (2−6)	Various cheeses & commercial dairy products	Sebald et al. (2019)
VPGEIVESL	β-CN (8- 16)	Cheddar	Lee et al., (1996)
SLVYPFPGPIPNSL	β-CN (57−70)	Various cheeses & commercial dairy products	Sebald et al. (2019)
LVYPFPGPIHN	β-CN (58–68)	Gouda	Toelstede and Hofmann, (2008a)
VYPFPGPIPN	β-CN-A2 (59–68)		
YPFPGPIHN	β-CN (60-68)	Gouda	Toelstede and Hofmann, (2008a; b)
YPFPGPIPN	β-CN- A2 (60–68)	Gouda	Toelstede and Hofmann, (2008a)
YPFPGPIHNS	β-CN (60-69)	Gouda	Toelstede and Hofmann, (2008a; b)
TQTPVVVPPFLQPE	β-CN (78−91)	Various cheeses & commercial dairy products	Sebald et al. (2018); Sebald et al. (2019)
TQTPVVVPPFL	β-CN (78−88)	Various cheeses & commercial dairy products	Sebald et al. (2019)
VVPPFL	β-CN (84−88)		
VPPFLQPE	β-CN (84−91)	Various cheeses & commercial dairy products	
MAPKHKEMPFPKYPVEPF	β-CN (102−119)	Various cheeses & commercial dairy products	Sebald et al. (2018); Sebald et al. (2019)
VENLHLPLPLL	β-CN (130−140)	Various cheeses & commercial dairy products	Sebald et al. (2019)
VENLHLPLPLLQSW	β-CN (130−143)		
NLHLPLPLLQS	β-CN (132−142)		
LHLPLP	β-CN (133−138)	Various cheeses & commercial dairy products	Sebald et al. (2018); Sebald et al. (2019)
LHLPLPLL	β-CN (133−140)		
LHLPLPLLQS	β-CN (133−142)	Various cheeses & commercial dairy products	Sebald et al. (2019)
HLPLPLLQ	β-CN (134−141)	Various cheeses & commercial dairy products	Sebald et al. (2018); Sebald et al. (2019)
HLPLPLLQS	β-CN (134−142)	Various cheeses & commercial dairy products	Sebald et al. (2019)
LPLPLLQSW	β-CN (135−143)		
WMHQPHQPLPPTVMFPPQ	β-CN (143−160)		
KVLPVPQKAVPYPQ	β-CN (169−182)	Various cheeses & commercial dairy products	Sebald et al. (2018); Sebald et al. (2019)
VLPVPQ	β-CN (170−175)		
VLPVPQKAVPYPQ	β-CN (170−182)	Various cheeses & commercial dairy products	Sebald et al. (2019)
QEPVLGPVRGPFPII	β-CN (194−208)		

Appendix 1 - Continued

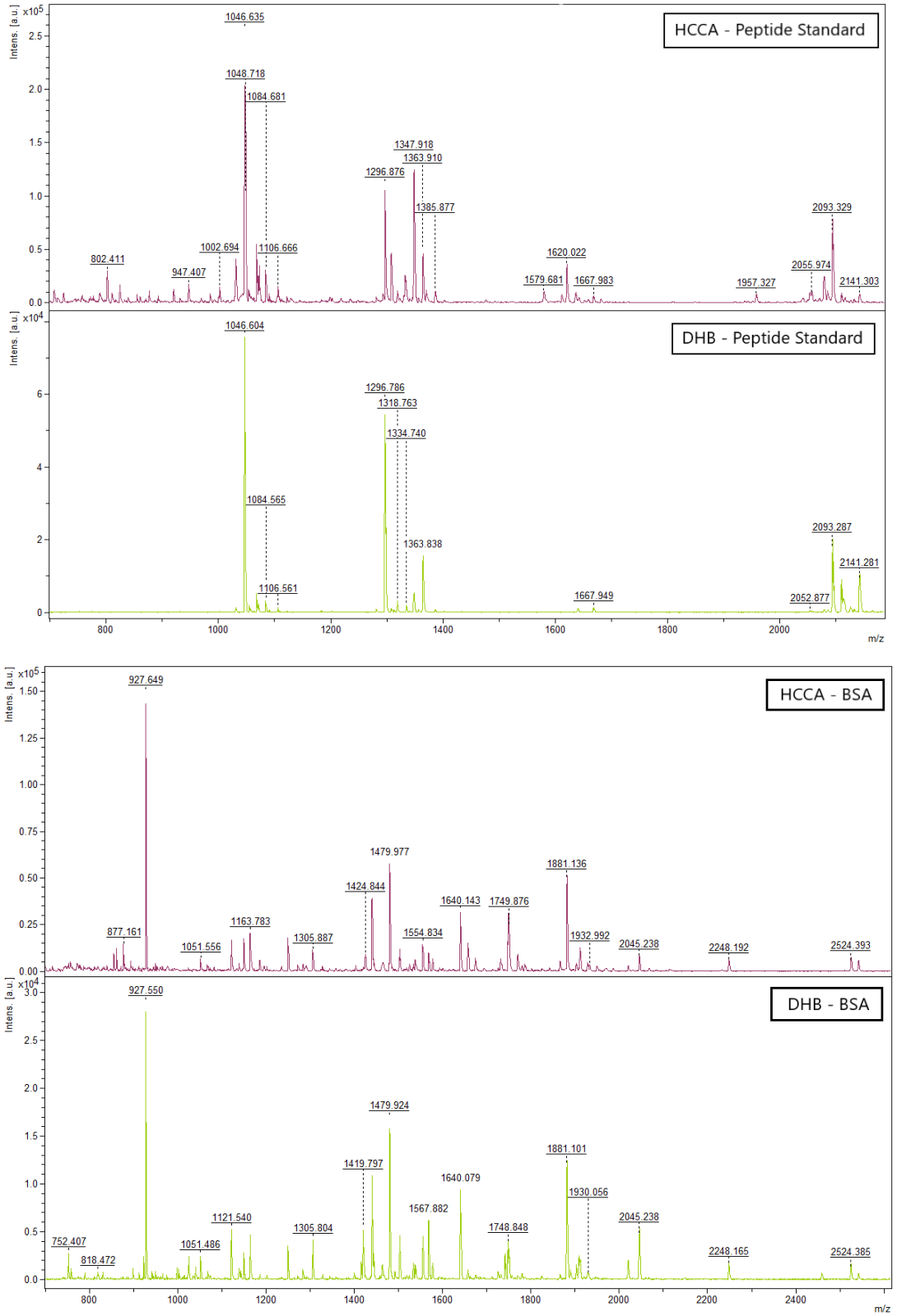
Peptide Sequence	Fragment	Food Type	Reference
YQEPVLGPVRGPF PIIV	β-CN (193-209)	Cheddar; Gouda	Broadbent et al. (1998); Soeryapranata et al. (2002); Broadbent et al. (2002); Singh et al, (2005); Soeryapranata et al. (2008); Toelstede and Hofmann, (2008a)
RPKHPIKHQ	αs1-CN (1-9)	Cheddar	Broadbent et al. (1998); Broadbent et al. (2002)
RPKHPIK	αs1-CN (1-7)	Cheddar	Lee et al. (1996)
RPKHPIKHQGLPQ	αs1-CN (l-13)		
RPKHPIKHQGLPQ EVLNENLLRF	αs1-CN (1-23)	Cheddar	Lee et al. (1996); Soeryapranata et al. (2008)
LPQE	αs1-CN (11-14)	Cheddar; Gouda	Lee et al. (1996); Toelstede and Hofmann, (2008a; b)
NLLRFF	αs1-CN (19−24)	Various cheeses & commercial dairy products	Sebald et al. (2019)
FFVAPFPEVF	αs1-CN (23−32)		
VAPFPEVFGKE	αs1-CN (25−35)	Various cheeses & commercial dairy products	Sebald et al. (2018); Sebald et al. (2019)
VFGKEKVNEL	αs1-CN (31−40)		
DIKQM	αs1-CN (56−60)		
IQKEDVPS	αs1-CN (81−88)		
ERYLGYLEQ	αs1-CN (89−97)	Various cheeses & commercial dairy products	Sebald et al. (2019)
LLRLKK	αs1-CN (98−103)		
KPWIQPK	αs2-CN (f191- 197)	Cheddar	Lee et al. (1996)
FALPQYLKT	αs2-CN (174−182)	Various cheeses & commercial dairy products	Sebald et al. (2019)
EIVPN	αs1-CN (70–74)/ αs1-CN (110–114)	Gouda	Toelstede and Hofmann, (2008a; b)
FFSDKIAK	κ-CN (17−24)	Various cheeses & commercial dairy products	Sebald et al. (2018)
IAKYIPI	κ-CN (22−28)	Various cheeses & commercial dairy products	Sebald et al. (2019)
NYYQQKPVA	κ-CN (41−49)		
YQQKPVAL	κ−CN (43−50)	Various cheeses & commercial dairy products	Sebald et al. (2018); Sebald et al. (2019)
ARHPHPHLSFM	κ-CN (96−106)		
AIPPKKNQDKTEIP TIN	κ-CN (107−123)	Various cheeses & commercial dairy products	Sebald et al. (2019)
AIPPKKNQDKTEIP TINTIASGEPT	κ-CN (107−131)	Various cheeses & commercial dairy products	Sebald et al. (2018); Sebald et al. (2019)

Peptides are abbreviated to one letter code.

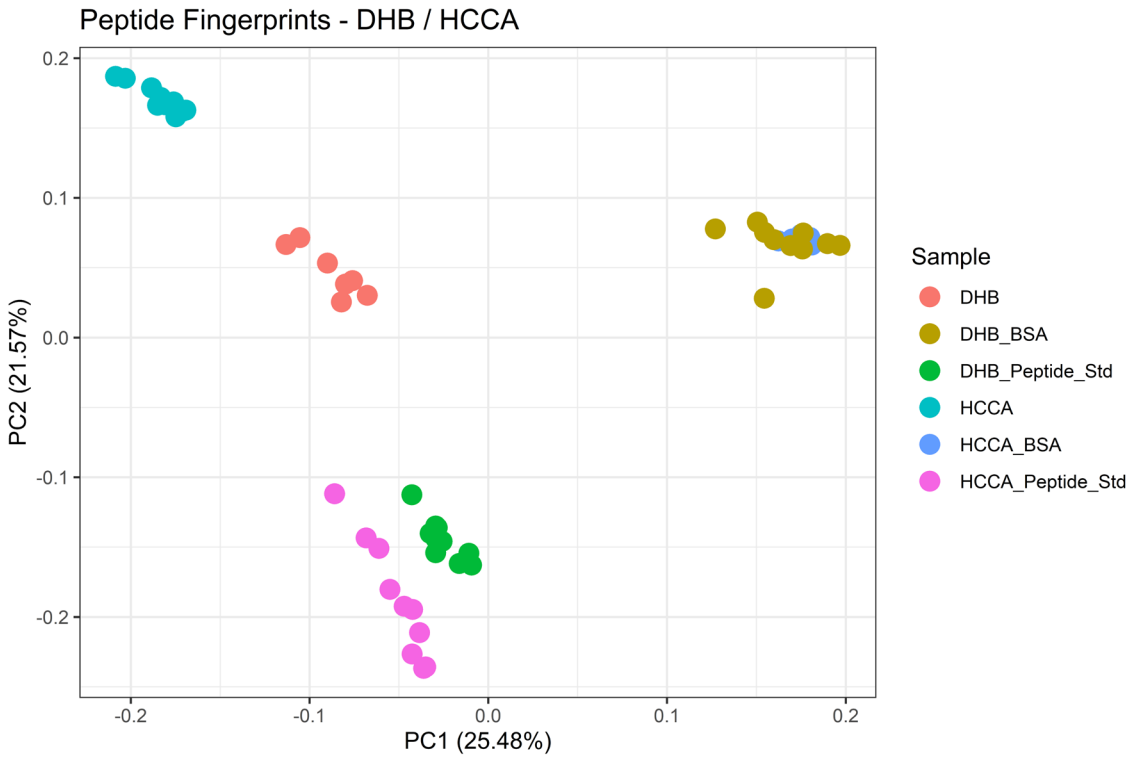
Appendix 2.

Kokumi active peptides identified in cheese and evaluated for taste activity via sensory techniques

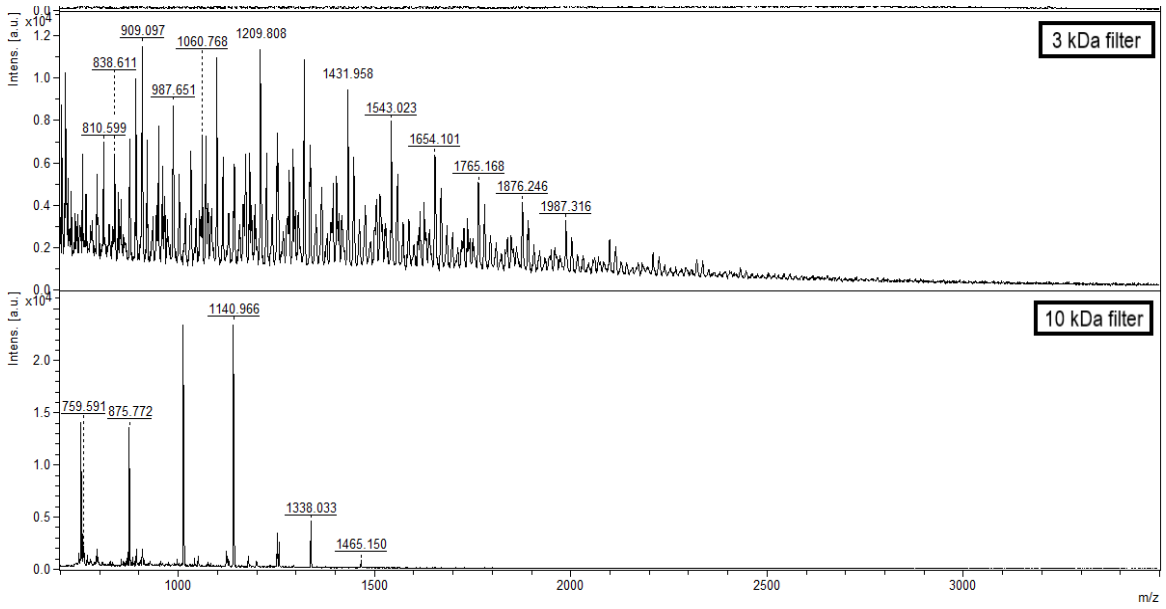
Peptide Sequence	Food Type	Reference
γ-Glu-Glu	Various cheeses	Toelstede et al. (2009); Hillman and Hofmann (2016)
γ-Glu-Gly		
γ-Glu-Met		
γ-Glu-His		
γ-Glu-Gln	Various cheeses	Toelstede et al. (2009)
γ-Glu-Leu	Various cheeses	Toelstede et al. (2009)
γ-Glu-Thr	Parmesan cheese	Hillman and Hofmann (2016)
γ-Glu-Ala		
γ-Glu-Lys		
γ-Glu-Val		
γ-Glu-Tyr		
γ-Glu-Asp		
γ-Glu-Trp		
γ-Glu-Ile		
γ-Glu-Phe		



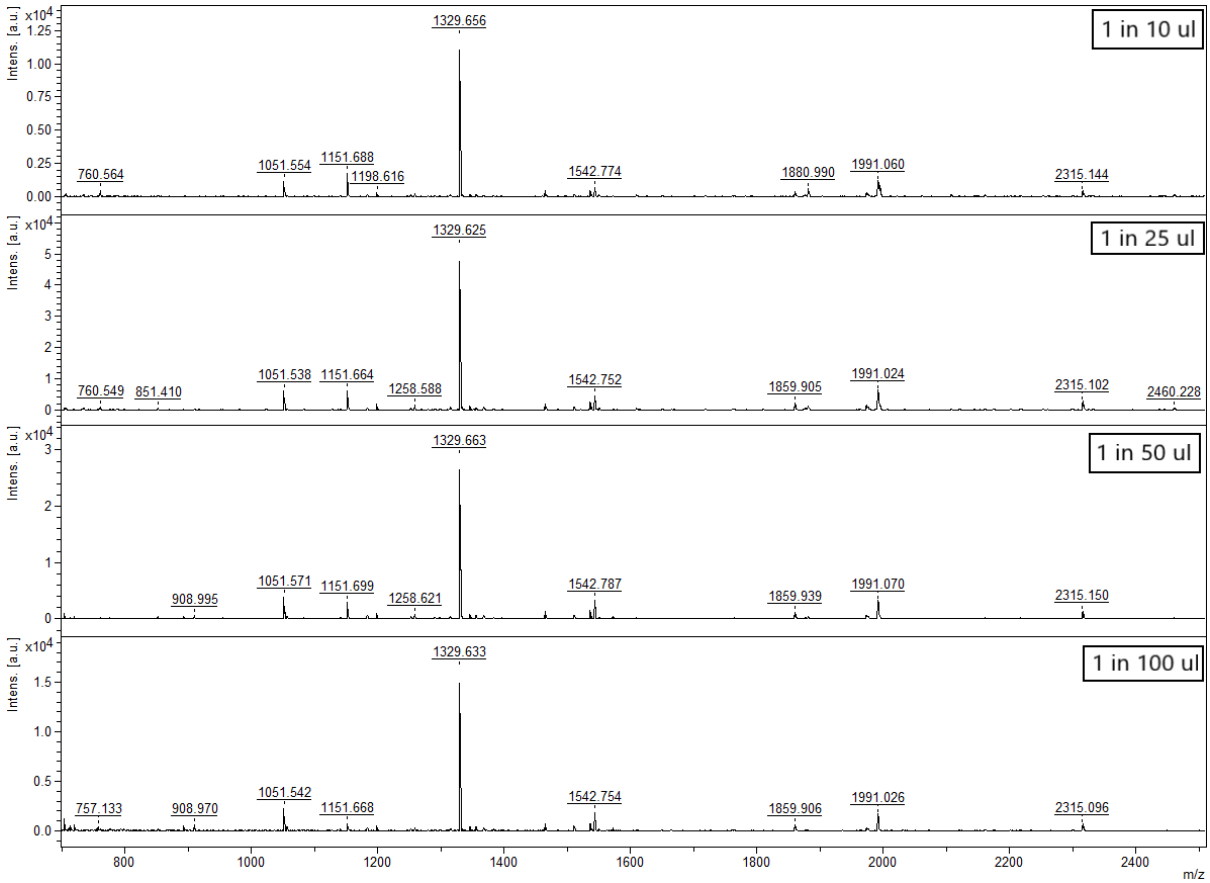
Appendix 3: *Samples mixed with both peptide standard (a) and BSA (b) using the matrix HCCA, detected more peaks consistently and had a higher signal intensity for the detected peaks. HCCA mixed with peptide standard detected 55 peaks, averaged across all replicates, while DHB detected 48 peaks. HCCA mixed with BSA detected an average of 79 peaks, and DHB mixed with BSA detected 64 peaks.*



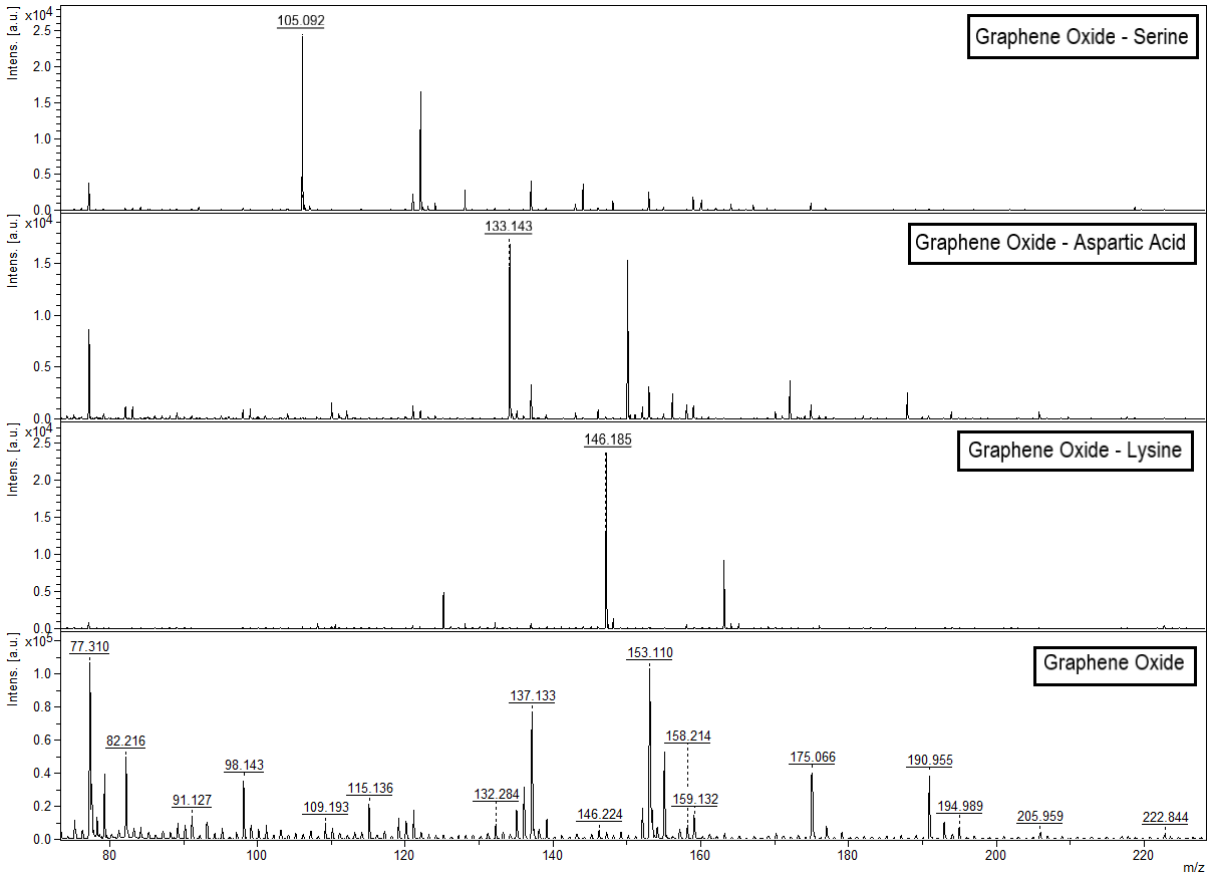
Appendix 4: *Principal component analysis (PCA) was performed on the processed data to determine the reproducibility of replicates and the similarity between samples prepared using the same matrix. Some replicates were removed due to poor quality spectra identified during processing, The BSA samples group together for both matrices, but the samples spiked with the peptide standard are grouped apart. The samples prepared using just the DHB matrix cluster closer to the BSA and peptide groups, suggesting that they are more similar in peptide composition than those prepared using HCCA. This indicates that there may peaks in the spiked samples that originate from the DHB matrix. This provided insights into the level of interference that may be carried over from the matrix used. This may prove problematic when screening large numbers of samples if there are similarities in the fingerprint of samples prepared using DHB, which may interfere with sample discrimination.*



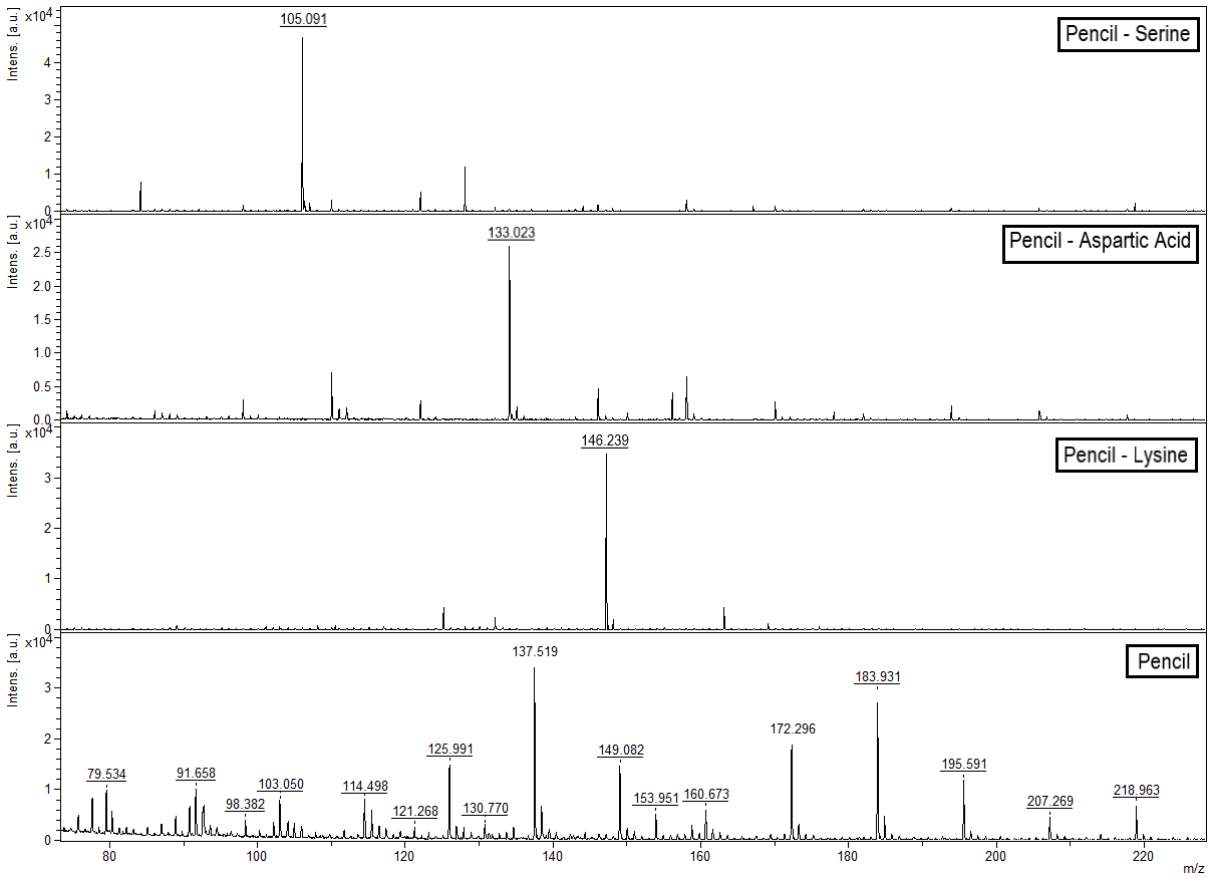
Appendix 5: *Ultrafiltration devices were trialled with a 3kDa (Pall) and 10kDa (Millipore, Inc, Billerica, MA) cut off; however, these devices removed a large number of ions in the ultrafiltrate that were present in the unfiltered samples, as well as introducing some suppressing and interfering ions.*



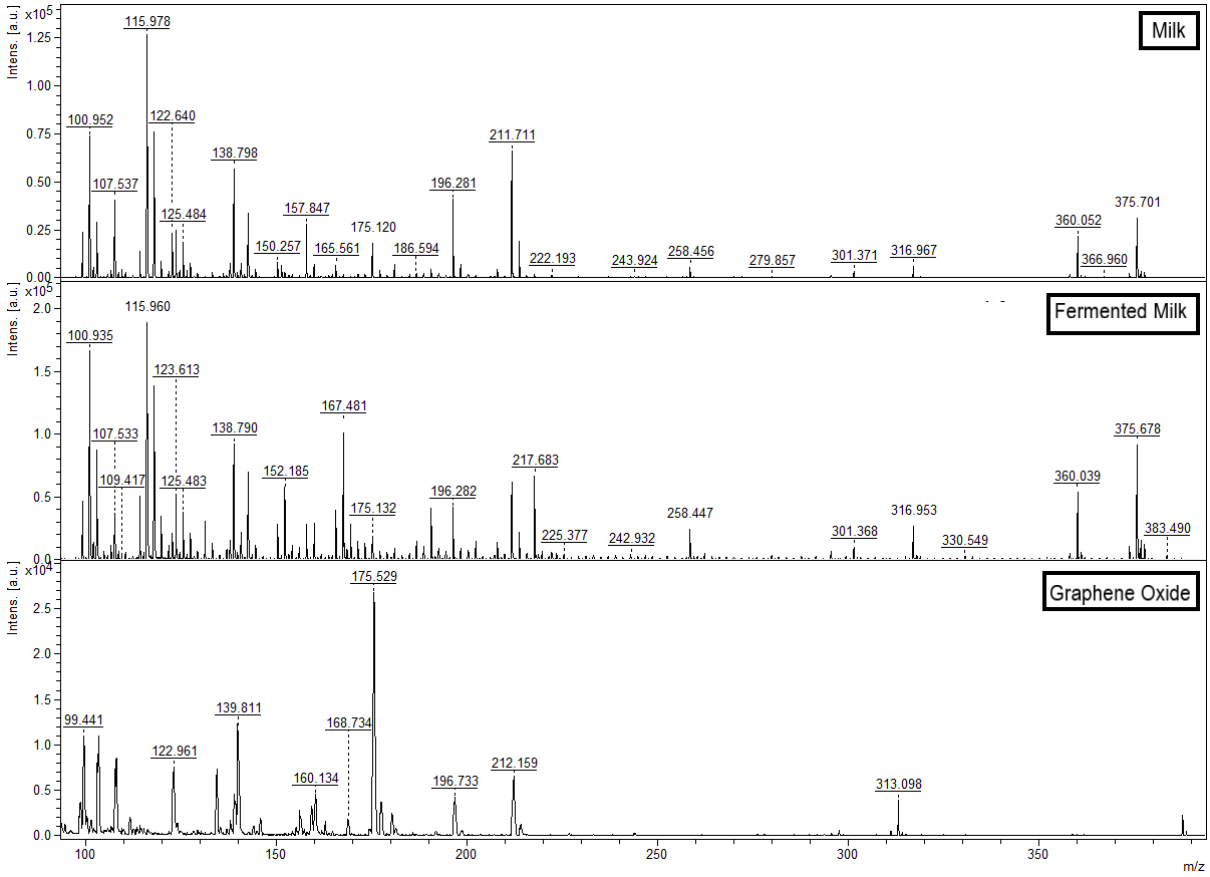
Appendix 6: *Dilutions trialled to establish suitable sample dilution. Four dilutions were trialled – 1 in 10 ul, 1 in 25 ul, 1 in 50 ul, 1 in 100 ul, sample: matrix. 1 in 25 ul was selected.*



Appendix 7: *Two matrices were shortlisted: graphene oxide and 6B graphite pencil. Using the graphene oxide as matrix, the three tested amino acids were detected with minimal background interference from the matrix. The labelled peak indicates the analyte. A blank sample of the graphene oxide spectrum is included to indicate the level of background noise introduced by the matrix when used with the analyte.*

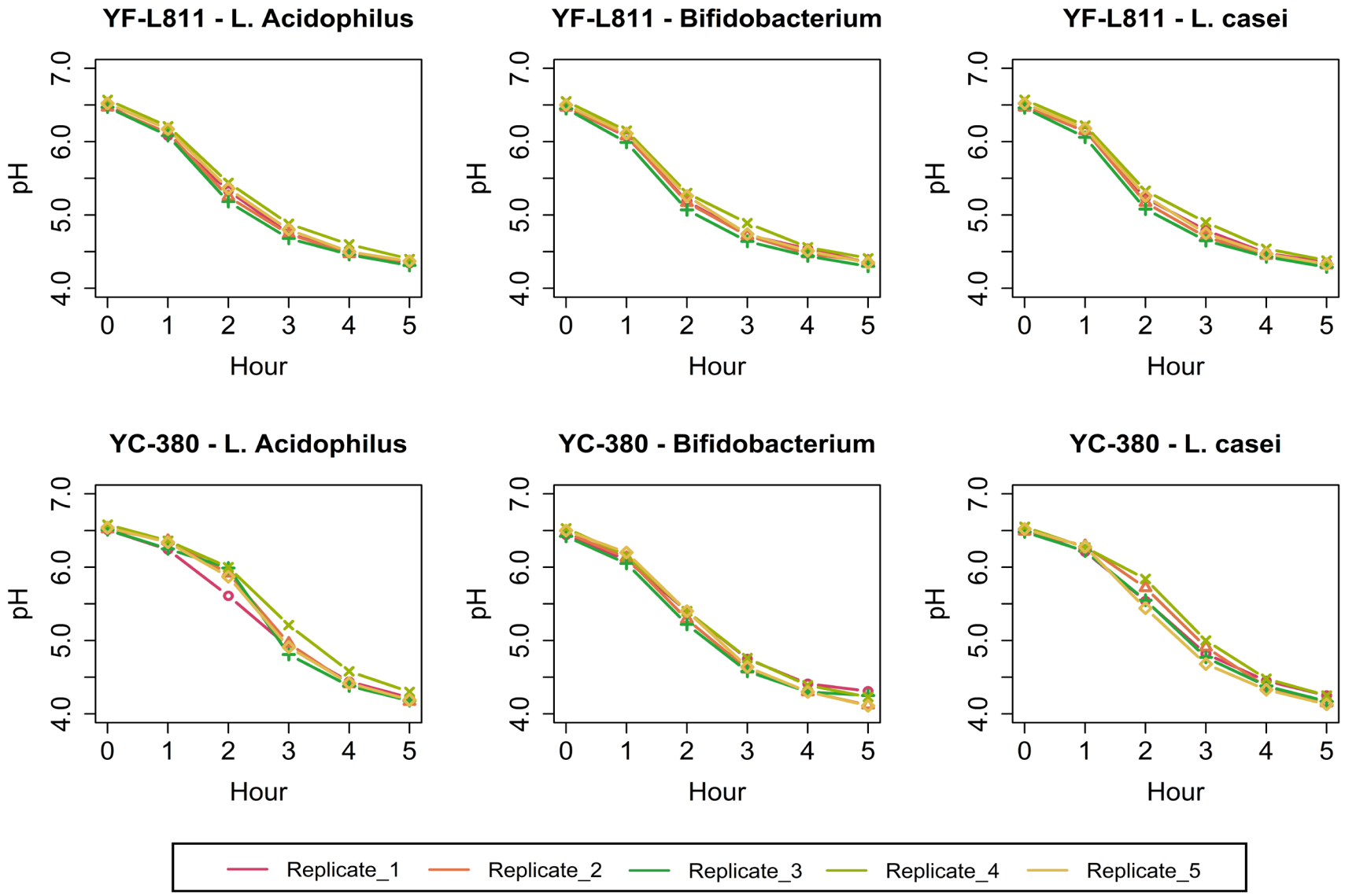


Appendix 8: *Each amino acid was also detected with the 6B graphite pencil as matrix. There was minimal background interference from the pencil matrix, indicating its suitability for the analysis of low molecular weight compounds using MALDI-TOF.*

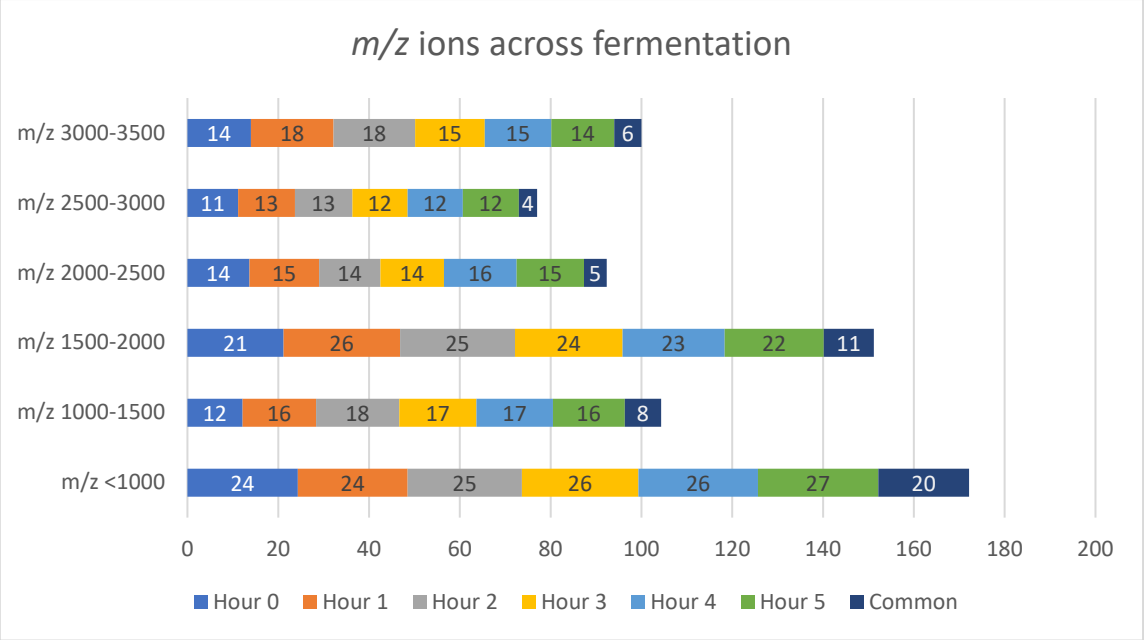


Appendix 9: *The sample preparation was adequate for fingerprinting the smaller compounds in milk and fermented milk samples using graphene oxide as matrix. There was minimal interference from the graphene oxide matrix in the milk and fermented milk spectra. There were many ions in common between the milk and fermented milk, although the fermented milk appears more abundant in ions, particularly in the m/z 100-200 region. From this work, we could establish an appropriate matrix for the analysis of both peptides and small compounds in milk and fermented milk, that lends itself to rapid and automated acquisition.*

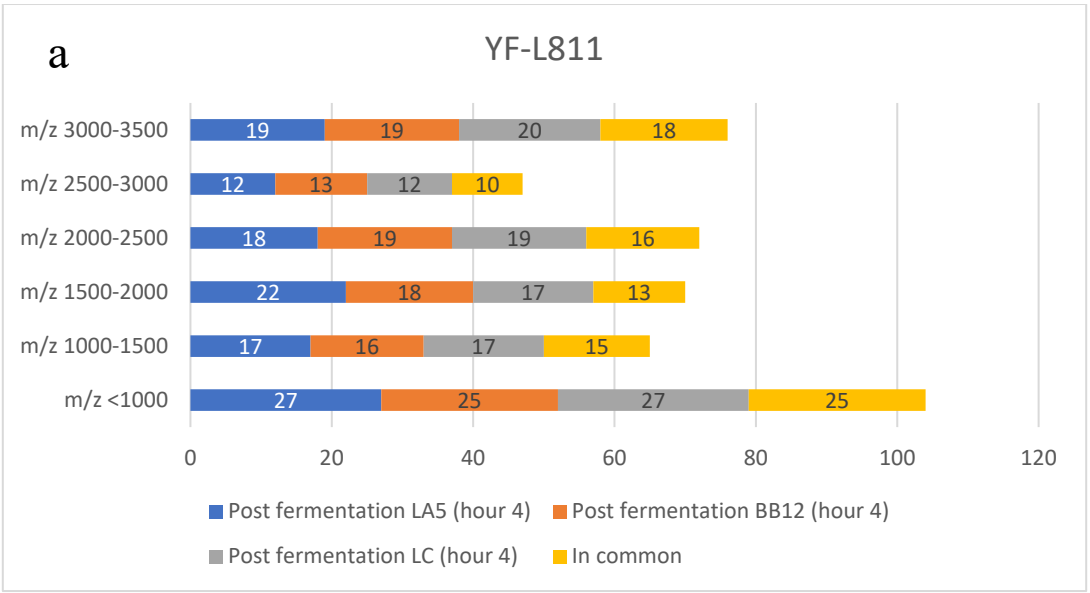
Chapter 4:



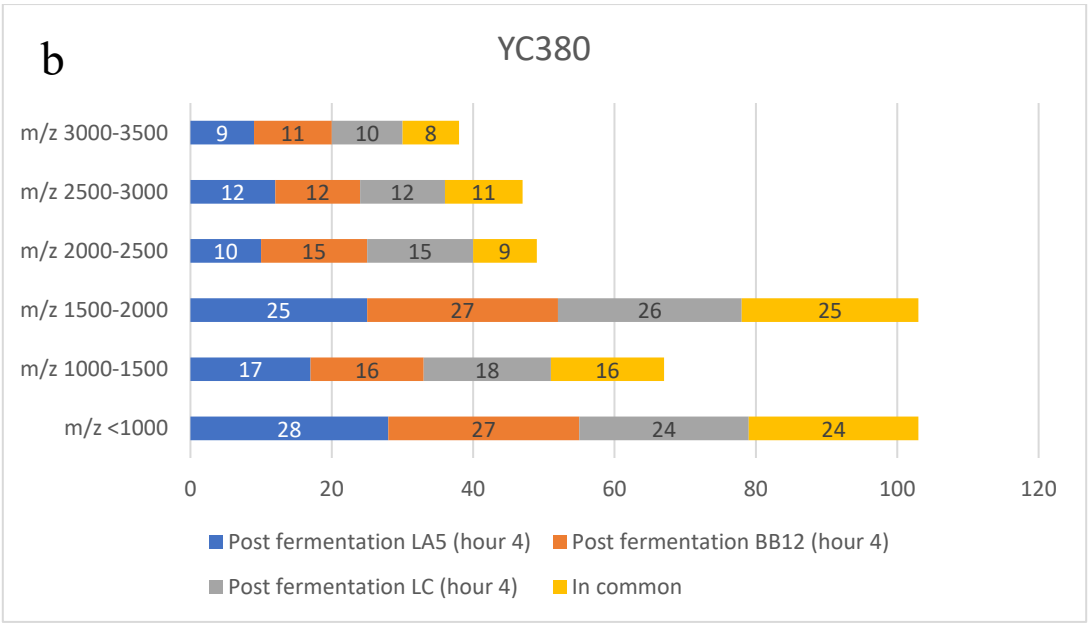
Appendix 10: *pH readings of each fermented milk type. These are an average of three measurements. An ANOVA of the pH readings by replicate for each fermented milk type was not significant. (P = 0.999, 0.899, 0.909, 0.851, 0.997, and 0.962, respectively, for each fermented milk type).*



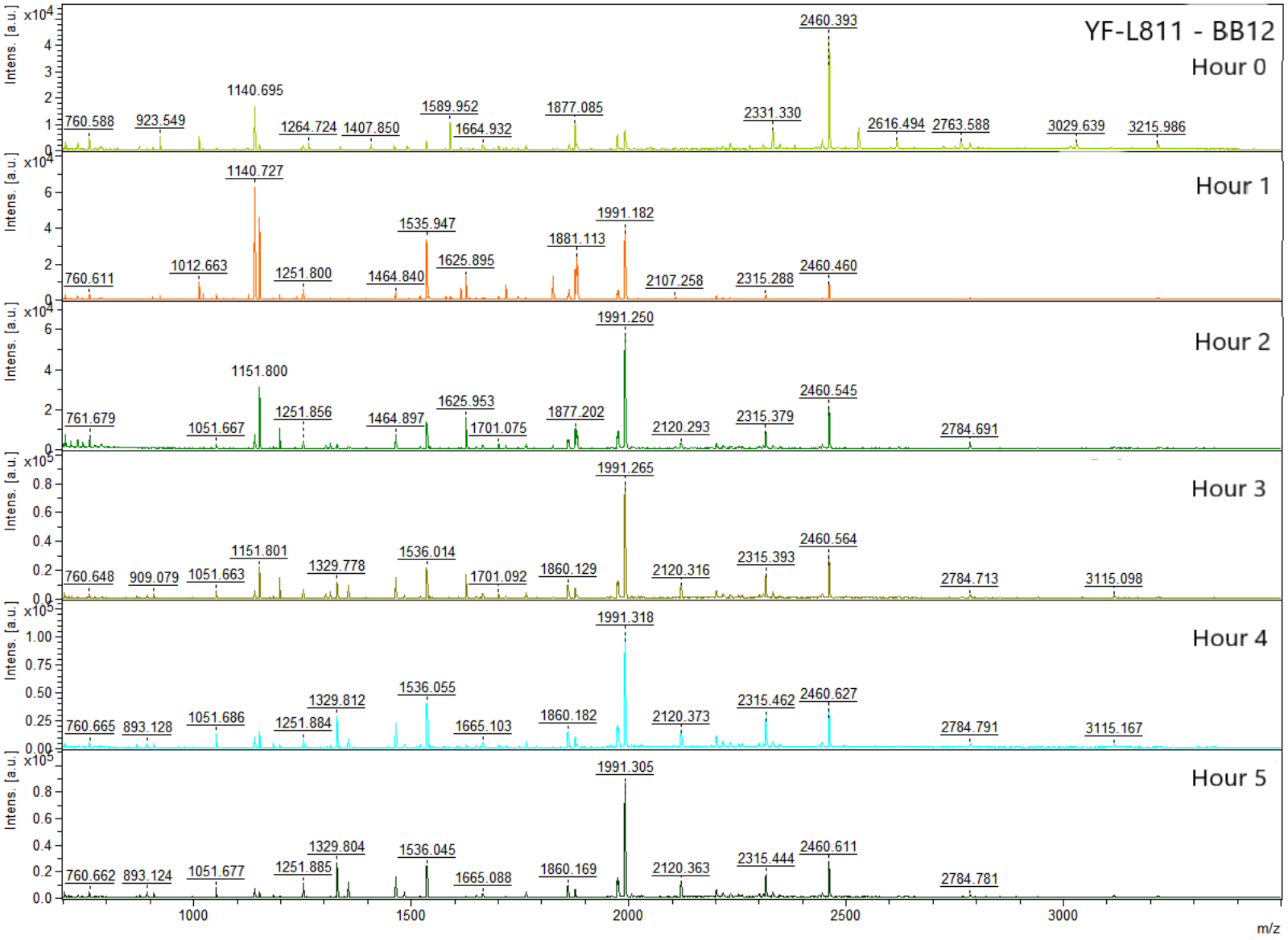
Appendix 11 *Barchart of m/z ions detected by MALDI-TOF throughout fermentation across all samples. m/z ions are grouped by size. Chart is colour-coded by fermentation time, with m/z ions in common between all samples also highlighted.*



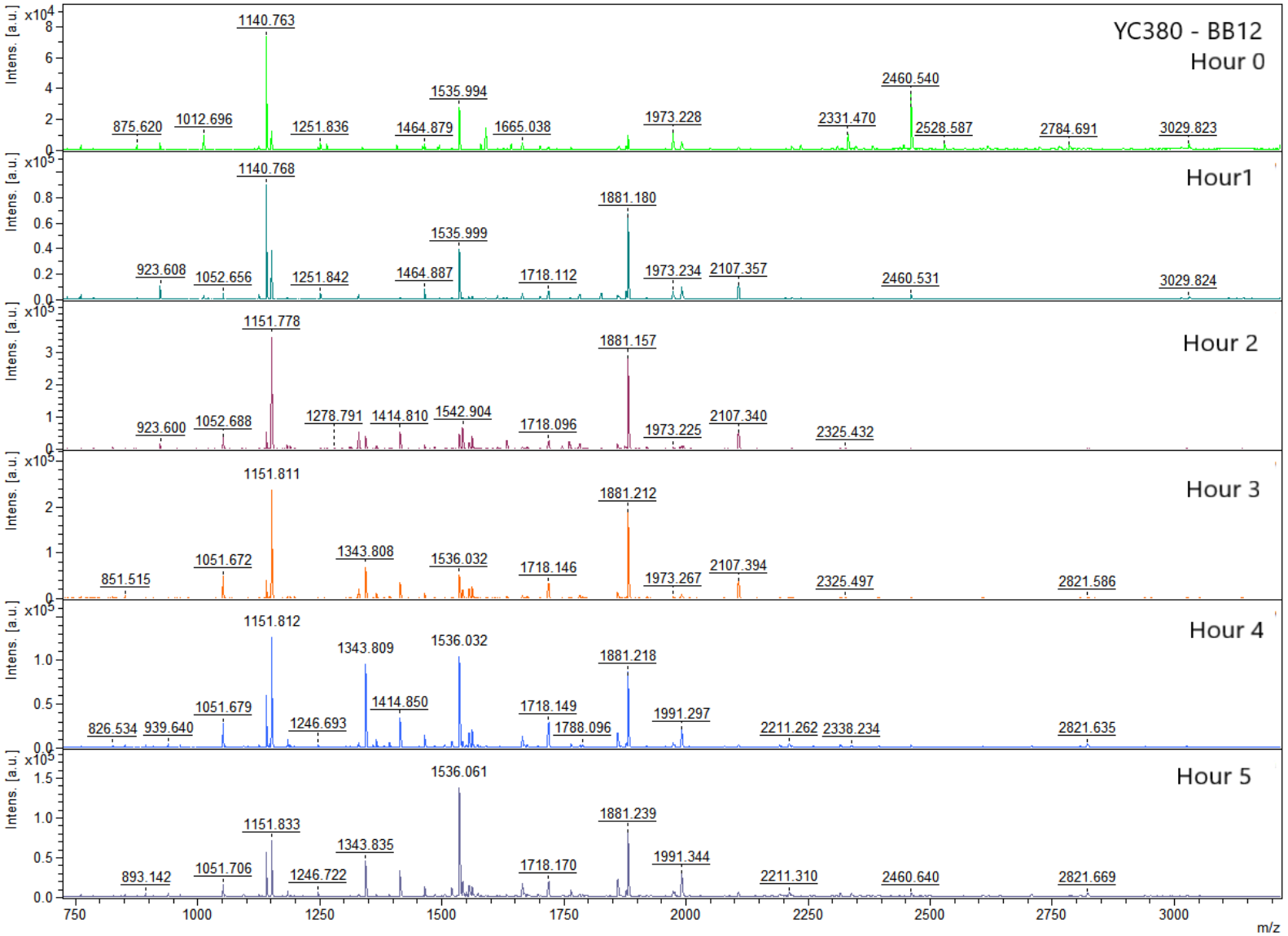
Appendix 12a *m/z ions detected for YF-L811, for fermented samples. The bar chart is colour-coded according to probiotic type used and the number of ions in common across all samples.*



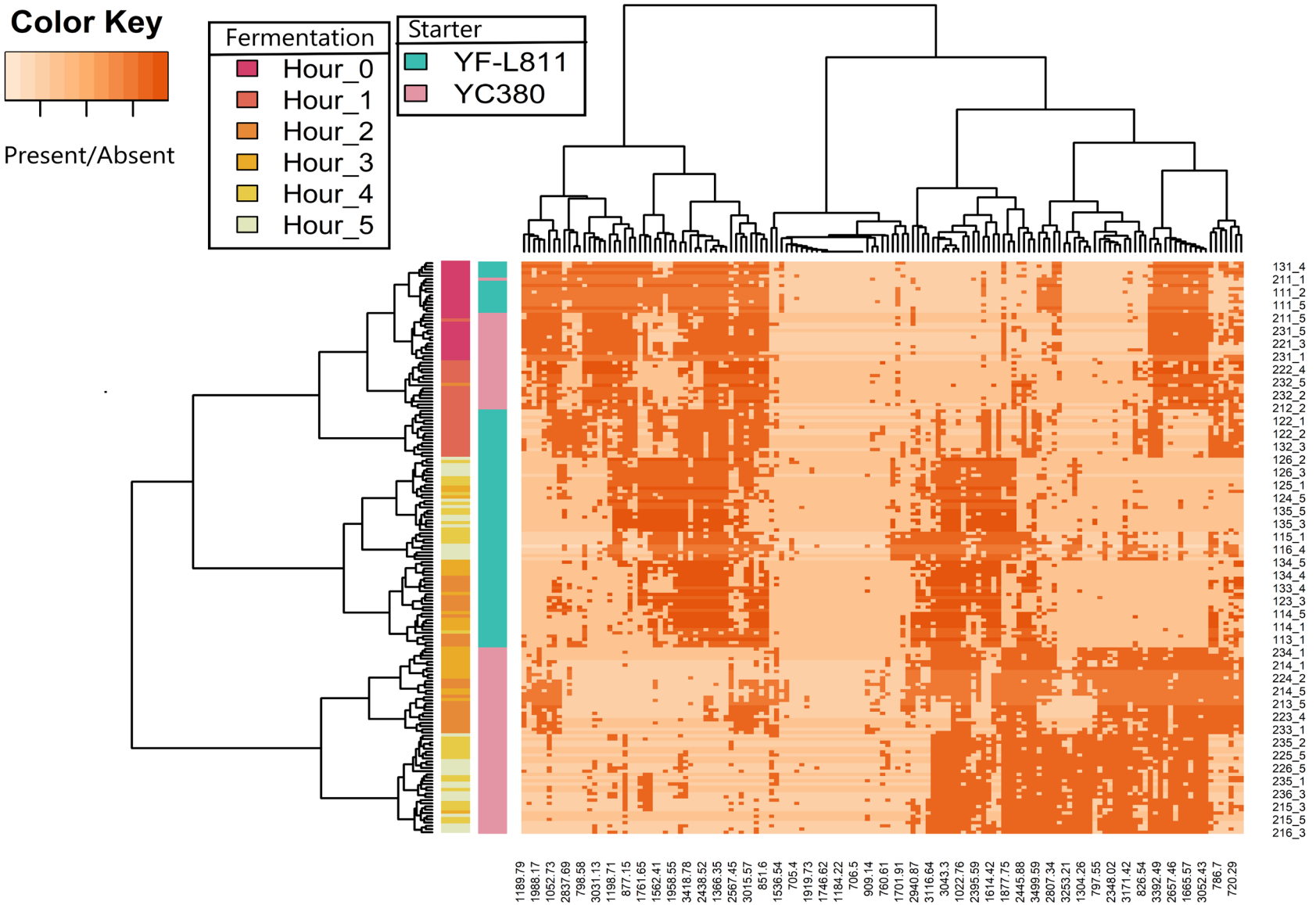
Appendix 12b *m/z ions detected for YC380, for fermented samples. The bar chart is colour-coded according to probiotic type used and the number of ions in common across all samples.*



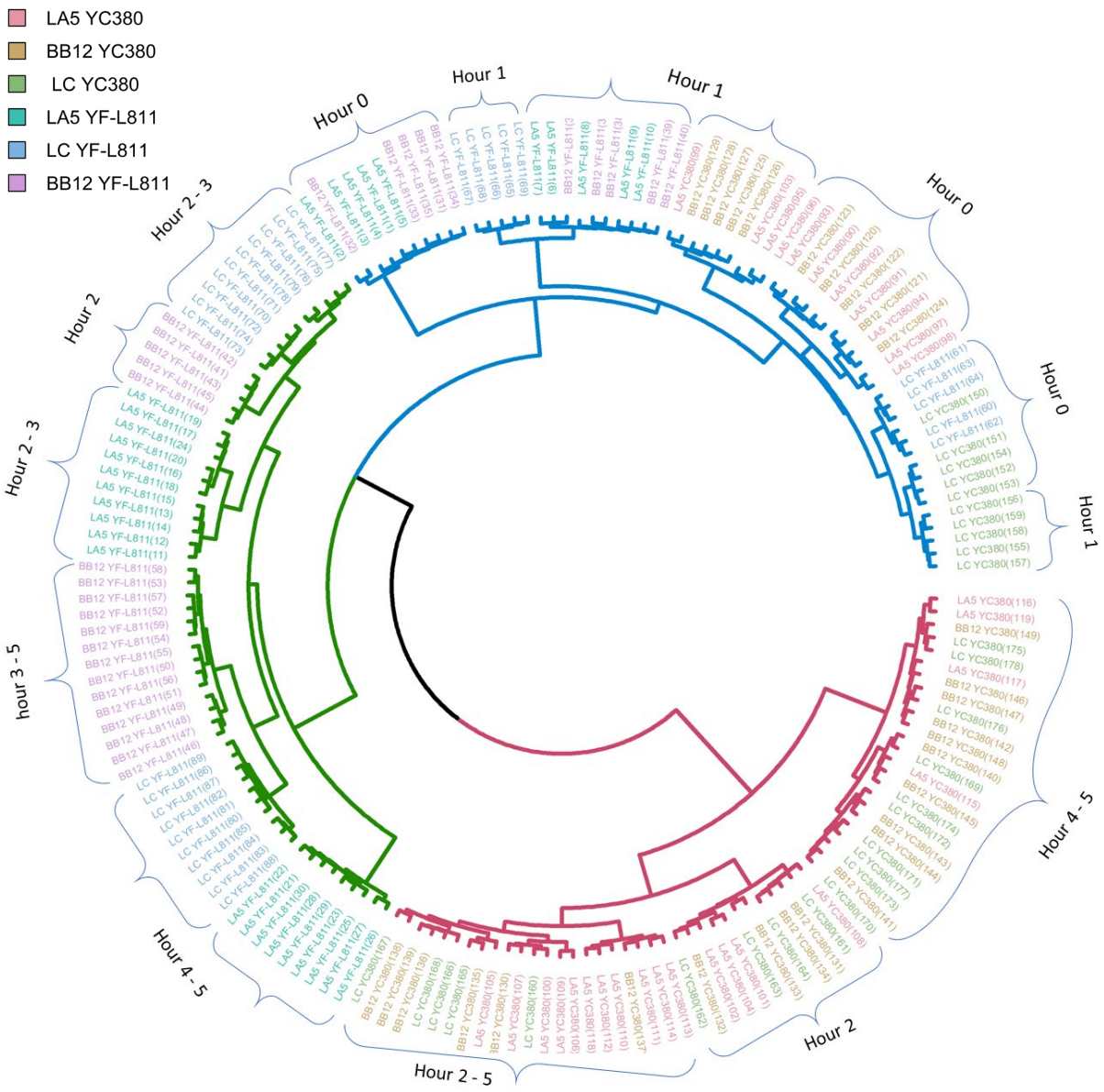
Appendix 13 *MALDI-TOF spectra - Starter YF-L811, Probiotic BB12*



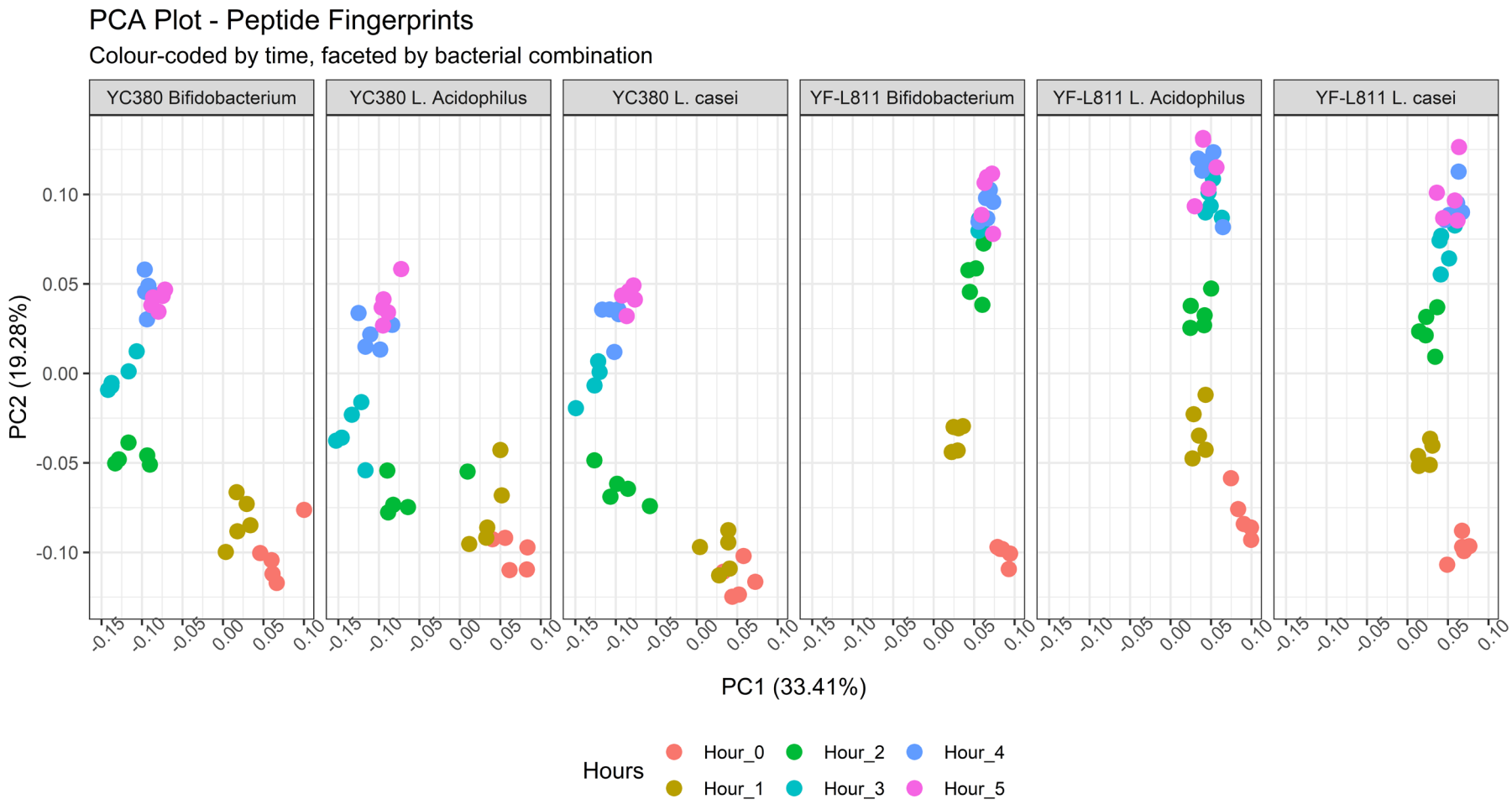
Appendix 14 *MALDI-TOF spectra - Starter YC380, Probiotic BB12*



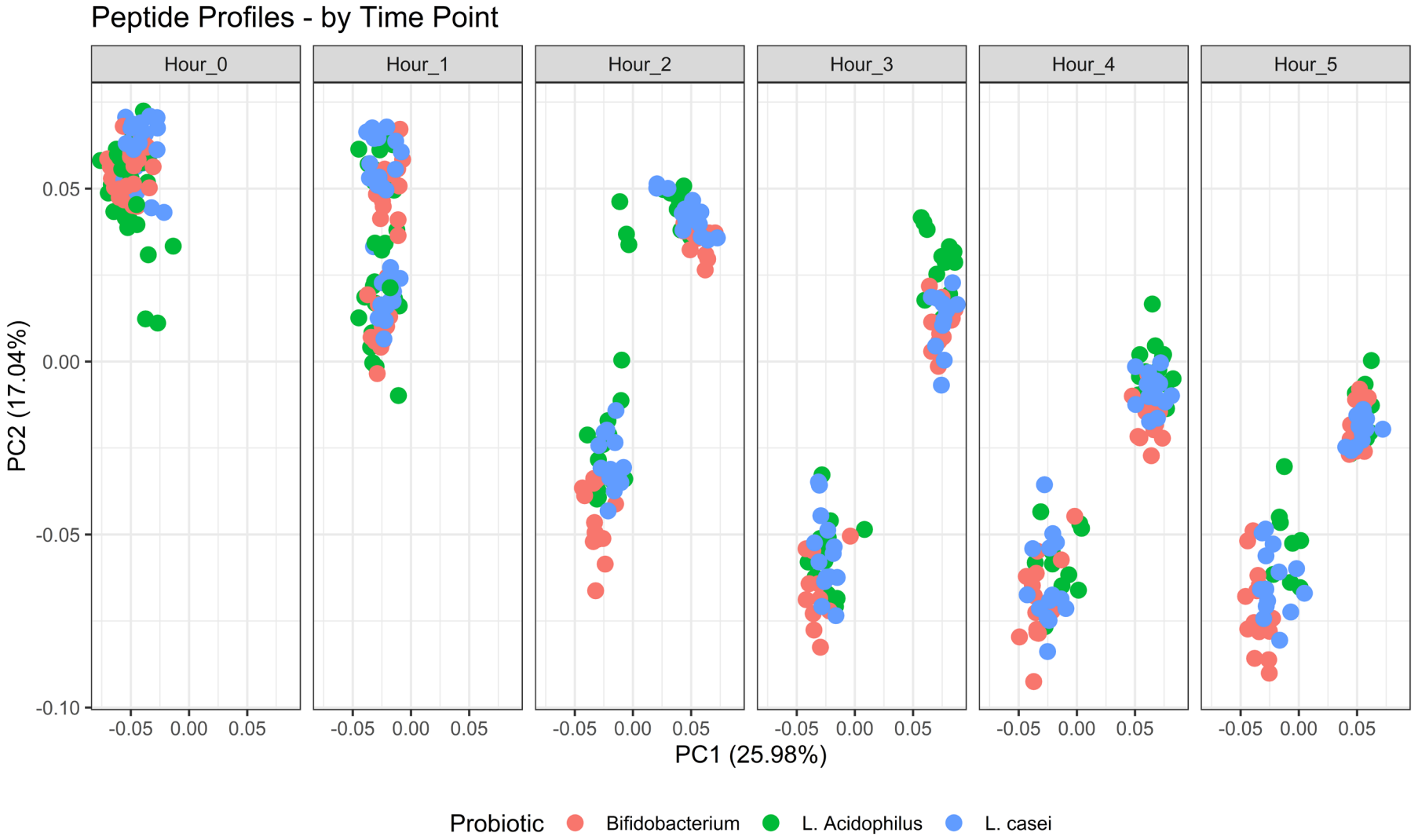
Appendix 15 *Hierarchically clustered heatmap of the peaks present across all samples. The rows are samples and columns are m/z values. The rows are annotated by starter culture and fermentation time. The heatmap is colour-coded by peaks present (light-orange) and absent (dark-orange).*



Appendix 16 *Hierarchical clustering of peptide fingerprints (averaged across the technical repeats) presented in a circlised layout. Dendrogram was generated using a Bray-Curtis distance and Ward.D2 linkage. Colours correspond to the bacterial combination. The dendrogram is annotated with labels corresponding to smaller sub-clusters. Potential outliers are highlighted.*



Appendix 17a *PCA plots of peptide fingerprints, faceting by bacterial combination and colour-coded by fermentation time point.*



Appendix 17b *PCA plots of peptide fingerprints, faceting by time and colour-coded by probiotic culture. This includes three MALDI technical repeats.*

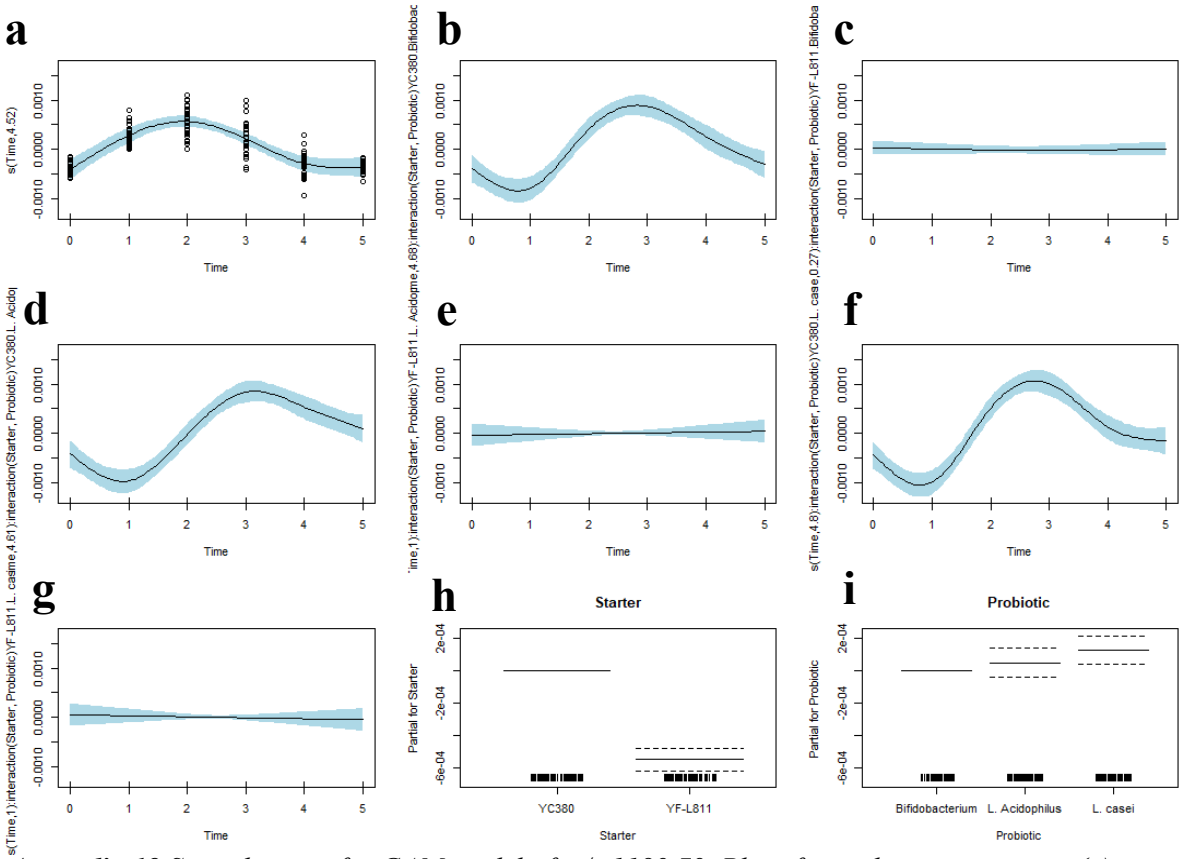
MALDI-TOF *m/z* ion	Multiple linear regression model: Adj_R2	Significance of multiple linear model: p-Value	Generalised additive model: Adj_R2	Significance of multiple linear model vs Generalised additive model
704.16	0.56	1.03E-16	0.7	1.54E-10
705.4	0.30	2.08E-01	0.38	8.27E-05
706.5	0.57	2.87E-04	0.73	2.83E-16
715.42	0.52	3.64E-25	0.64	1.09E-09
716.63	0.43	2.28E-13	0.56	1.85E-08
720.29	0.55	4.02E-04	0.71	1.18E-13
732.49	0.67	8.66E-02	0.75	6.09E-07
733.55	0.39	3.25E-04	0.41	2.50E-01
734.59	0.60	6.64E-12	0.7	3.16E-09
744.56	0.65	1.14E-02	0.76	9.20E-11
745.53	0.53	2.38E-02	0.6	3.16E-04
747.5	0.58	2.27E-06	0.85	2.16E-58
758.49	0.68	3.28E-07	0.78	1.60E-12
760.61	0.60	3.18E-01	0.82	5.70E-34
772.54	0.33	6.37E-02	0.41	1.68E-04
775.72	0.61	1.08E-13	0.75	3.03E-14
786.7	0.12	1.02E-01	0.73	8.01E-69
787.62	0.28	1.33E-01	0.66	2.52E-32
788.58	0.76	1.14E-07	0.71	NA
797.55	0.38	2.24E-16	0.73	5.41E-39
798.58	0.56	6.51E-01	0.66	8.45E-08
826.54	0.52	2.60E-01	0.76	5.58E-30
843.5	0.58	1.21E-30	0.7	2.31E-11
851.6	0.53	5.70E-17	0.72	3.00E-20
877.15	0.32	1.42E-10	0.63	1.01E-25
893.21	0.66	1.25E-27	0.71	3.32E-04
909.14	0.53	1.27E-16	0.61	1.43E-05
923.65	0.57	1.22E-26	0.85	9.59E-61
947.09	0.67	2.15E-19	0.87	1.50E-52
1012.77	0.55	9.25E-30	0.94	2.96E-245
1022.76	0.41	8.45E-09	0.77	1.54E-50
1051.71	0.55	7.90E-30	0.76	2.60E-27
1052.73	0.28	2.28E-05	0.75	2.92E-59
1066.13	0.15	3.87E-02	0.91	5.39E-296
1125.7	0.40	2.02E-18	0.91	1.53E-186
1140.79	0.48	1.43E-22	0.91	8.78E-171
1151.89	0.30	5.72E-01	0.95	0.00E+00
1184.22	0.59	3.56E-28	0.71	1.43E-12
1189.79	0.16	3.39E-01	0.88	8.73E-212
1198.71	0.44	4.80E-04	0.85	1.46E-95
1252.37	0.80	8.22E-08	0.86	5.65E-14
1304.26	0.32	8.60E-13	0.78	1.65E-65
1314.36	0.44	4.08E-04	0.9	4.29E-161
1330.23	0.68	1.38E-23	0.75	2.69E-07
1344.28	0.77	2.88E-24	0.93	1.64E-85

MALDI-TOF *m/z* ion	Multiple linear regression model: Adj_R2	Significance of multiple linear model: p-Value	Generalised additive model: Adj_R2	Significance of multiple linear model vs Generalised additive model
1366.35	0.44	3.88E-06	0.84	1.08E-84
1393.55	0.51	1.61E-06	0.59	3.37E-05
1415.33	0.79	4.79E-21	0.93	9.47E-73
1465.38	0.81	1.42E-30	0.85	2.73E-08
1485.64	0.49	5.32E-17	0.51	1.10E-01
1495.42	0.64	5.92E-38	0.92	7.73E-117
1521.51	0.71	2.29E-22	0.92	4.71E-80
1536.54	0.72	1.29E-17	0.93	2.38E-97
1543.43	0.37	1.88E-06	0.72	2.76E-39
1556.43	0.80	2.78E-18	0.93	1.04E-62
1562.41	0.58	2.47E-05	0.91	3.78E-127
1590.58	0.59	9.64E-35	0.91	4.61E-115
1614.42	0.22	2.86E-10	0.89	1.33E-218
1626.4	0.41	3.50E-06	0.89	1.03E-151
1633.88	0.22	8.42E-03	0.82	1.80E-111
1650.53	0.64	8.15E-15	0.92	5.52E-128
1665.57	0.77	2.14E-14	0.94	2.78E-109
1701.91	0.73	3.70E-34	0.84	1.29E-17
1718.67	0.77	5.46E-07	0.83	6.72E-08
1746.62	0.07	2.02E-02	0.78	8.79E-109
1761.65	0.11	9.84E-02	0.72	3.00E-71
1764.61	0.67	1.02E-12	0.82	3.22E-25
1782.74	0.49	1.76E-03	0.9	2.33E-137
1788.87	0.32	2.03E-03	0.82	1.01E-90
1794.91	0.43	4.98E-01	0.9	1.03E-164
1826.59	0.27	4.80E-08	0.94	0.00E+00
1860.65	0.86	3.19E-73	0.93	1.05E-22
1863.63	0.55	1.05E-10	0.91	1.78E-132
1877.75	0.56	7.39E-11	0.78	1.35E-32
1881.77	0.51	1.13E-02	0.92	1.55E-183
1919.73	0.41	1.97E-01	0.88	4.42E-138
1958.55	0.50	2.78E-04	0.65	1.67E-11
1973.78	0.76	6.64E-12	0.85	1.52E-14
1976.61	0.82	1.55E-15	0.88	1.11E-12
1988.17	-0.01	5.56E-01	0.59	1.59E-46
1991.8	0.80	9.37E-19	0.91	6.15E-35
1994.81	0.45	9.19E-01	0.7	7.99E-26
2107.84	0.44	1.78E-05	0.79	5.23E-52
2120.84	0.87	7.25E-26	0.86	8.88E-01
2192.11	0.69	1.18E-38	0.82	4.29E-21
2201.87	0.84	1.15E-03	0.8	NA
2216.78	0.71	8.34E-08	0.83	5.37E-18
2217.75	0.73	1.41E-08	0.85	1.56E-21
2235.06	0.66	1.07E-20	0.84	9.91E-35
2253.19	0.75	9.46E-22	0.93	1.03E-91

MALDI-TOF *m/z* ion	Multiple linear regression model: Adj_R2	Significance of multiple linear model: p-Value	Generalised additive model: Adj_R2	Significance of multiple linear model vs Generalised additive model
2261.13	0.69	5.28E-40	0.87	5.99E-41
2300.61	0.88	7.84E-24	0.88	9.71E-01
2309.61	0.50	4.94E-01	0.79	7.08E-41
2316.09	0.89	5.17E-38	0.92	4.81E-11
2332.36	0.44	2.41E-08	0.93	2.48E-227
2348.02	0.53	3.91E-17	0.87	3.13E-89
2383.03	0.54	1.39E-29	0.91	1.76E-136
2395.59	0.82	4.42E-50	0.94	1.71E-69
2438.52	0.68	3.88E-01	0.83	2.04E-28
2445.88	0.68	1.42E-16	0.86	5.55E-40
2461.91	0.67	1.07E-21	0.88	1.44E-55
2498.27	0.43	5.07E-18	0.76	7.68E-42
2567.45	0.70	1.44E-32	0.83	1.73E-26
2607.68	0.75	1.29E-24	0.91	1.87E-58
2618.45	0.39	6.66E-15	0.93	1.60E-288
2621.55	0.79	2.95E-02	0.85	1.36E-11
2657.46	0.45	2.83E-21	0.87	3.65E-103
2724.62	0.55	2.57E-28	0.93	9.48E-178
2764.83	0.44	1.18E-17	0.89	1.86E-133
2767.91	0.50	9.55E-22	0.92	9.21E-162
2785.82	0.70	1.07E-16	0.79	2.40E-11
2796.66	0.81	1.92E-03	0.86	1.98E-08
2807.34	0.86	1.99E-41	0.94	2.72E-40
2822.89	0.81	3.42E-43	0.89	8.72E-20
2837.69	0.44	1.71E-22	0.78	1.90E-47
2854.27	0.89	1.70E-16	0.82	NA
2896.4	0.59	5.91E-28	0.92	2.53E-135
2911.28	0.60	3.03E-23	0.88	1.02E-77
2940.87	0.65	7.50E-38	0.82	1.85E-27
2951.98	0.49	2.47E-11	0.76	2.61E-34
3015.57	0.64	2.89E-38	0.94	2.41E-175
3024.12	0.65	4.70E-40	0.92	2.49E-114
3026.91	0.59	1.89E-25	0.8	4.25E-30
3031.13	0.63	3.15E-38	0.95	8.49E-239
3043.3	0.67	3.75E-03	0.83	2.23E-30
3052.43	0.44	2.08E-09	0.84	1.21E-77
3111.11	0.74	7.53E-48	0.9	2.65E-52
3116.64	0.67	2.81E-10	0.85	2.02E-35
3126.49	0.81	1.64E-02	0.83	2.89E-03
3140.25	0.25	4.38E-03	0.57	2.04E-20
3155.09	0.70	6.55E-16	0.83	1.58E-21
3171.42	0.69	7.75E-01	0.81	8.31E-20
3217.83	0.56	4.59E-21	0.89	3.22E-103
3223.18	0.74	1.62E-47	0.8	2.57E-08
3238.59	0.77	7.12E-54	0.84	3.51E-12

MALDI-TOF *m/z* ion	Multiple linear regression model: Adj_R2	Significance of multiple linear model: p-Value	Generalised additive model: Adj_R2	Significance of multiple linear model vs Generalised additive model
3253.21	0.45	1.08E-18	0.78	9.55E-47
3307.15	0.16	1.31E-06	0.49	3.27E-18
3346.38	0.85	7.70E-29	0.94	1.42E-51
3392.49	0.70	4.50E-34	0.78	3.39E-10
3397.42	0.45	2.12E-15	0.75	3.33E-35
3418.78	0.72	9.89E-38	0.81	1.09E-12
3452.88	0.52	2.55E-06	0.73	1.40E-21
3499.59	0.18	2.83E-01	0.23	4.95E-03

Appendix 18*: A full list of m/z ions detected by MALDI-TOF MS, after processing. Reported are the adjusted R^2 obtained from the full multiple linear regression model, the significance (p-value) of the multiple linear regression model, the adjusted R^2 obtained from the generalised additive model, and the p-value reported from an ANOVA of the multiple linear regression model vs. generalised additive model.*



Appendix 19 *Smooth terms for GAM model of m/z 1189.79. Plot of signal intensity ~ time* ***(a)****, smooth term for Time * interaction (YC380 *Bifidobacterium)* ***(b)****, smooth term for Time* interaction (YF-L8110 *Bifidobacterium)* ***(c)****, smooth term for Time * interaction (YC380 *. L. acidophilus)* ***(d)****, smooth term for Time * interaction (YF-L811 *. L. Acidophilus)* ***(e)****, smooth term for Time * interaction (YC380 *.L. casei)* ***(f)****, smooth term for Time * interaction (YF-L811 *.L. casei)* ***(g)****, effect of starter culture* ***(h)****, effect of probiotic culture* ***(i).***

Inspection of the smooth terms indicates that the signal intensity behaves differently over time for each bacterial combination (Appendix 19). Starter YC380 has a higher intensity over time, relative to YF-L811, and no differences could be seen between the probiotics, although LC was marginally higher in intensity compared to BB12 and LA5. A good smooth should have some complexity to it, i.e., should not be linear, and the smooth term should be significant, i.e., non-significant terms in this case indicates there is no certainty as to the shape or direction of the effect. In this case, the smooth terms for the interaction of YF-L811 and time were linear and non-significant, suggesting a lack of complexity (i.e., linearity), and lack of certainty as to the direction of the effect. YC380 smooth terms were non-linear and significant, indicating complexity.

GAM interactions for 1189 It would appear that time has a non-linear effect on the peptide intensity but behaves differently for each starter culture and probiotic combination. In YF-L811, it is a straight, linear line. We can also see that there are different effects on the two starter cultures. Starter YC380 has a higher intensity relative to YF-L811. There are no major differences for the probiotic, although *L. casei* appears slightly higher in intensity.

The starter and probiotics were input as categorical variables. This fits a model then with a fixed effect for each level of the category. It appears that YF-L811 has a negative effective; YC380 has a positive effect, i.e. has a higher intensity over time. There are no major effects from the probiotics, these are all positive effects.

A significant smooth term is one where you cannot draw a horizontal line through the 95% confidence interval. The interaction of time with the YF-L811 samples is a significant term as a line can be drawn through the confidence interval. Looking at the summary of the model, the interaction smooth terms for YC380 are significant. The YF-L811 smooths have an edf of 1, indicating a straight line/linearity.

In the case of YF-L811, the smooth terms are linear and non-significant. Indicating that there is no complexity to the smooth, and there is not certainty as to the shape or direction of its effect. YC380 is non-linear and significant, indicating it has some complexity.

Summary of GAM model for *m/z ion* 1189:

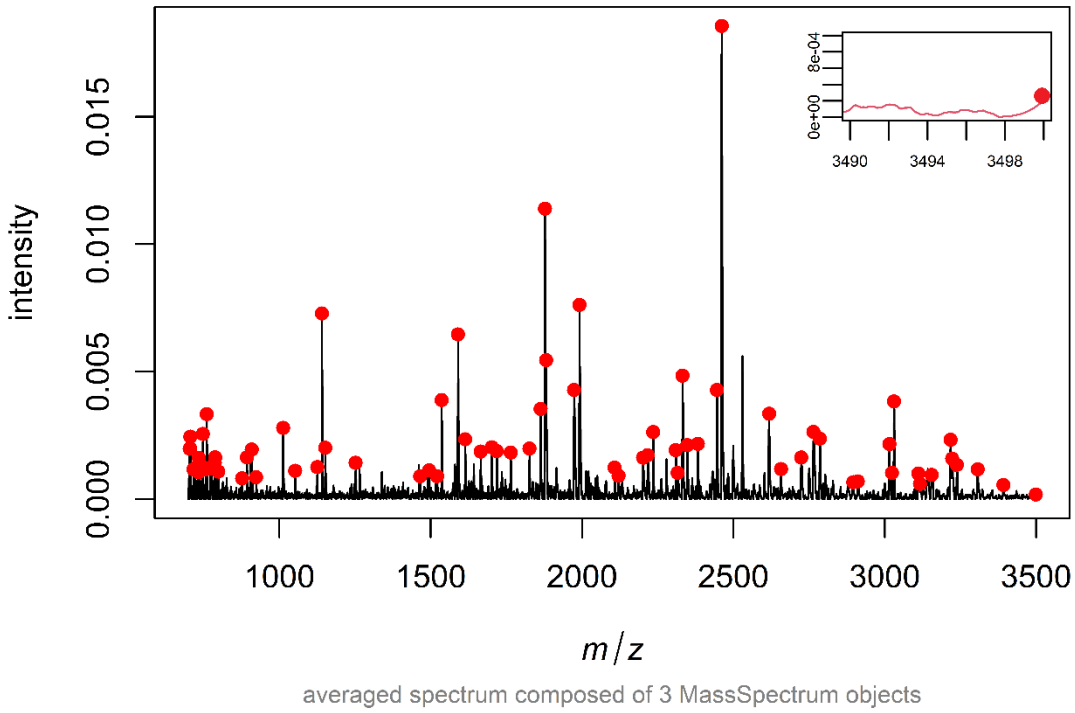
Parametric coefficients	***Estimate***	***Std. Error***	***t-value***	***Pr(>|t|)***
(Intercept)	1.12E-03	3.57E-05	31.373	<2e-16
Starter [YF-L811]	-5.48E-04	3.58E-05	-15.303	<2e-16
Probiotic [L.Acidophilus]	4.97E-05	4.38E-05	1.135	0.2583
Probiotic [L. casei]	1.29E-04	4.40E-05	2.931	0.0039

Approximate significance of smooth terms:	***edf***	***Ref.df***	***F***	***p-value***
s(Time)	4.5198	4.8619	39.657	<2e-16
s(Time) * interaction(Starter,Probiotic)YC380.Bifidobacterium	4.6794	4.9377	23.171	<2e-16
s(Time) * interaction(Starter, Probiotic)YF-L811.Bifidobacterium	0.2669	0.4821	0.05	0.876
s(Time) * interaction(Starter, Probiotic)YC380.L. Acidophilus	4.6108	4.9134	21.587	<2e-16
s(Time) * interaction(Starter, Probiotic)YF-L811.L. Acidophilus	1	1	0.128	0.721
s(Time) * interaction(Starter, Probiotic)YC380.L. casei	4.7959	4.9705	28.675	<2e-16
s(Time) * interaction(Starter, Probiotic)YF-L811.L. casei	1	1	0.199	0.656

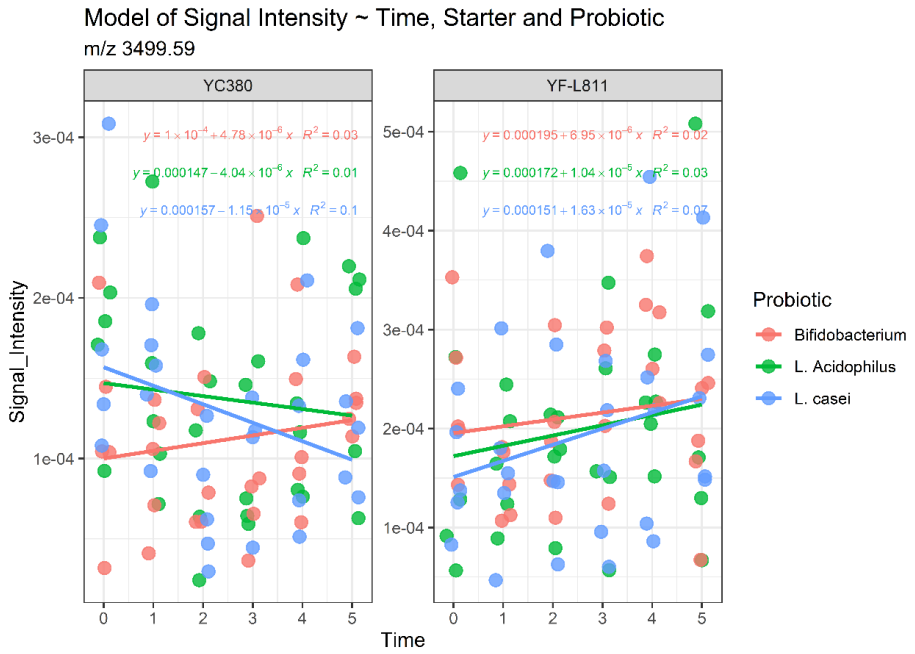
Appendix 20 *Table 17 Summary of GAM model on peptide 1189*

Smooth Term	*k'*	*edf*	*k-index*	*p-value*
s(Time)	5	4.52	1.21	0.99
s(Time) * interaction(Starter,Probiotic)YC380.Bifidobacterium	5	4.679	1.21	1
s(Time) * interaction(Starter, Probiotic)YF-L811.Bifidobacterium	5	0.267	1.21	1
s(Time) * interaction(Starter, Probiotic)YC380.L. Acidophilus	5	4.611	1.21	0.99
s(Time) * interaction(Starter, Probiotic)YF-L811.L. Acidophilus	5	1	1.21	0.99
s(Time) * interaction(Starter, Probiotic)YC380.L. casei	5	4.796	1.21	0.99
s(Time) * interaction(Starter, Probiotic)YF-L811.L. casei	5	1	1.21	1

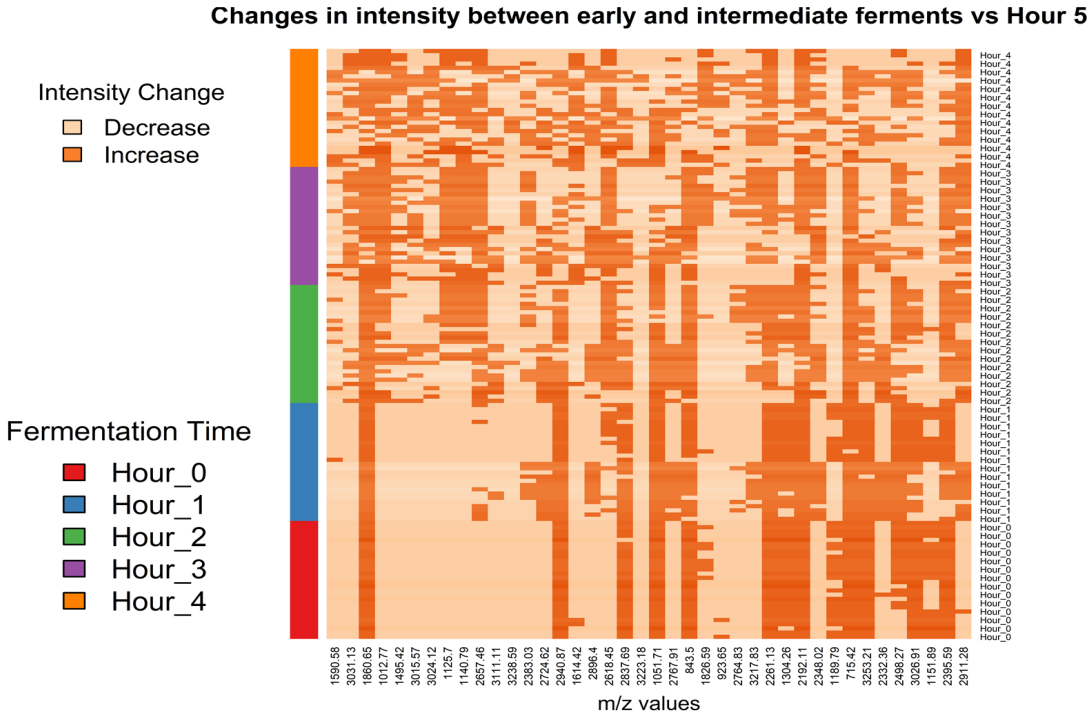
Appendix 21 *Gam.Check results for GAM model on peptide 1189. Low p-values indicate that k is too low, especially when edf is close to k'*



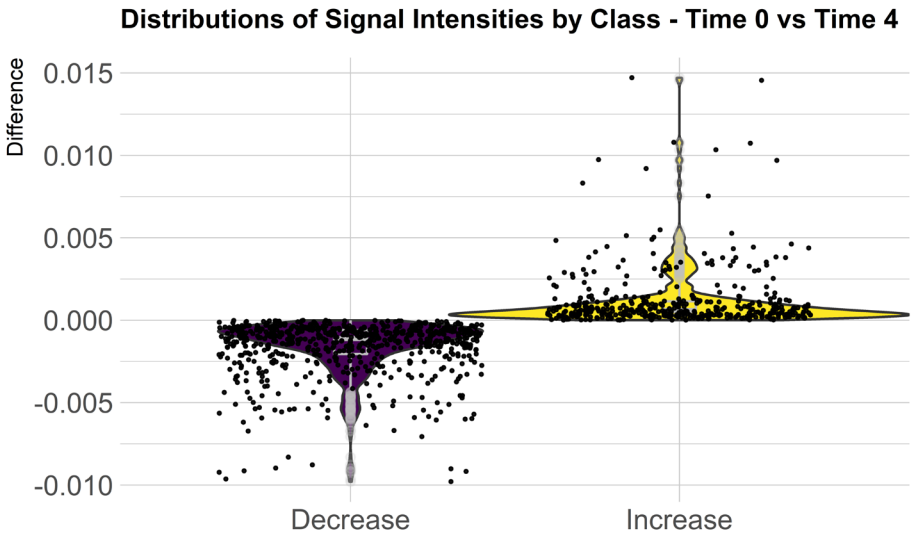
Appendix 22 *MALDI-TOF processed spectra. Peaks are highlighted in red. Some peaks are notably low intensity (inset).*



Appendix 23 L*inear model of m/z 3499.59 ~ Fermentation time, starter and probiotic..*



Appendix 24 *Heatmap of class labels for each m/z value across fermentation time points, comparing change in signal intensity between hour 0 vs hour 5, hour 1 vs hour 5, hour 2 vs hour 5, hour 3 vs hour 5 and hour 4 vs hour 5. The fermentation times are coloured-coded along the rows, m/z values are labelled along the columns and are organised by lowest-highest molecular weight. The heatmap is coloured by the class label: "decreasing" or "increasing".*



Appendix 25 *Violin plot of distribution of signal intensities for each class, for all instances of the top 40 m/z values between time 0 and 4.*

MASSEY UNIVERSITY
COLLEGE OF SCIENCES
TE WĀHANGA PŪTAIAO

School of Food & Advanced Technology
Microbiology Laboratory
COLLEGE OF SCIENCES
Manager: Ann-Marie Jackson
T: +64 6 951 7115
Email: A.M.Jackson@massey.ac.nz

TO:	Fionnuala Murphy	AT:	AgResearch
SUBJECT:	Final Report	REPORT DATE:	20/10/2020
MICRO JOB #:	J20-222	TEST DATE:	19/10/2020 → 20/10/2020
Sample Type:	fermented milks for sensory testing	DATE Received at LAB:	19/10/2020
Lab Testing by:	Kylie Evans	Results Read by:	Kylie Evans

No	Sample Name	Solid / Liquid	Single / Composite	Aerobic PC cfu/mL or g	Anaerobic PC cfu/mL or g	*Bacillus cereus* cfu/mL or g	*Bifido-bacteria* cfu/mL or g	Coliforms		Enterobact-eriaceae cfu/mL or g
								Total MPN/10mL	*E. coli* MPN/10mL or g	
1	Sample 1	L	S					< 3		
2	Sample 2	L	S					< 3		
3	Sample 3	L	S					< 3		
4	Sample 4	L	S					< 3		
5	Sample 5	L	S					< 3		
6	Sample 6	L	S					< 3		

References / Methodology:

General Microbiological Testing: Compendium of Methods for the Microbiological Examination of Foods. 4th Edition. American Public Health Assoc. Edited by Downes and Ito, 2001.

Aerobic Plate Count (APC): Standard Plate Count Agar for 48 hrs @ 30°C
Aerobic Plate Count (APC) for cosmetics: Bacteriological Analytical Manual, 8th Edition, Revision A, 1998. Chapter 23
Aerobic Plate Count on Petrifilm™: 3M™ Quickswab™ onto 3M AC Petrifilm™ for 48 hrs @ 30°C (milk/dairy products) or 35°C (other samples)
Anaerobic Plate Count (AnPC): Standard Plate Count Agar for 48 hrs @ 30°C under anaerobic conditions
Bacillus cereus: Mannitol Egg Yolk Polymixine (MYP) Agar 24hrs @ 30°C, +ve control NCTC 11143 (NZRM 3429)

Micro Job # J20-222 **Client Reference: fermented milks for sensory testing**

Appendix 26 *Microbial results from Consumer Trial for YF-L811 samples*

MASSEY UNIVERSITY
COLLEGE OF SCIENCES
TE WĀHANGA PŪTAIAO

School of Food & Advanced Technology
Microbiology Laboratory
COLLEGE OF SCIENCES
Manager: Ann-Marie Jackson
T: +64 6 951 7115
Email: A.M.Jackson@massey.ac.nz

TO:	Fionnuala Murphy	AT:	AgResearch
SUBJECT:	Final Report	REPORT DATE:	3/11/2020
MICRO JOB #:	J20-231	TEST DATE:	2/11/2020 → 3/11/2020
Sample Type:	fermented milks for sensory testing	DATE Received at LAB:	2/11/2020
Lab Testing by:	Kylie Evans	Results Read by:	Kylie Evans

No	Sample Name	Solid / Liquid	Single / Composite	Aerobic PC cfu/mL or g	Anaerobic PC cfu/mL or g	*Bacillus cereus* cfu/mL or g	*Bifido-bacteria* cfu/mL or g	Coliforms		Enterobact-eriaceae cfu/mL or g
								Total MPN/10mL	*E. coli* MPN/10mL or g	
1	Sample 1	L	S					< 3		
2	Sample 2	L	S					< 3		
3	Sample 3	L	S					< 3		
4	Sample 4	L	S					< 3		
5	Sample 5	L	S					< 3		
6	Sample 6	L	S					< 3		

References / Methodology:

General Microbiological Testing: Compendium of Methods for the Microbiological Examination of Foods. 4th Edition. American Public Health Assoc. Edited by Downes and Ito, 2001.

Aerobic Plate Count (APC): Standard Plate Count Agar for 48 hrs @ 30°C
Aerobic Plate Count (APC) for cosmetics: Bacteriological Analytical Manual, 8th Edition, Revision A, 1998. Chapter 23
Aerobic Plate Count on Petrifilm™: 3M™ Quickswab™ onto 3M AC Petrifilm™ for 48 hrs @ 30°C (milk/dairy products) or 35°C (other samples)
Anaerobic Plate Count (AnPC): Standard Plate Count Agar for 48 hrs @ 30°C under anaerobic conditions
Bacillus cereus: Mannitol Egg Yolk Polymixine (MYP) Agar 24hrs @ 30°C, +ve control NCTC 11143 (NZRM 3429)

Micro Job # J20-231 **Client Reference: fermented milks for sensory testing**

Appendix 27 *Microbial results from Consumer Trial for YC380 samples*

MASSEY UNIVERSITY
COLLEGE OF SCIENCES
TE WĀHANGA PŪTAIAO

School of Food & Advanced Technology
Microbiology Laboratory
COLLEGE OF SCIENCES
Manager: Ann-Marie Jackson
T: +64 6 951 7115
Email: A.M.Jackson@massey.ac.nz

TO:	Fionnuala Murphy	AT:	AgResearch
SUBJECT:	Final Report	REPORT DATE:	21/09/10
MICRO JOB #:	J20-187	TEST DATE:	17/04/20 → 19/4/20
Sample Type:	Fermented milk	DATE Received at LAB:	17/04/20
Lab Testing by:	Kylie Evans	Results Read by:	Ann-Marie Jackson

No	Sample Name	Solid / Liquid	Single / Composite	Aerobic PC cfu/mL or g	Anaerobic PC cfu/mL or g	*Bacillus cereus* cfu/mL or g	*Bifido-bacteria* cfu/mL or g	Coliforms		Enterobact-eriaceae cfu/mL or g
								Total MPN/10mL	*E. coli* MPN/10mL	
1	Sample 1	L	S					240	9.1	
2	Sample 2	L	S					> 2400	< 3	
3	Sample 3	L	S					< 3	< 3	
4	Sample 4	L	S					1100	< 3	
5	Sample 5	L	S					6.2	< 3	
6	Sample 6	L	S					3	< 3	

References / Methodology:

General Microbiological Testing: Compendium of Methods for the Microbiological Examination of Foods. 4th Edition. American Public Health Assoc. Edited by Downes and Ito, 2001.

Aerobic Plate Count (APC): Standard Plate Count Agar for 48 hrs @ 30°C
Aerobic Plate Count (APC) for cosmetics: Bacteriological Analytical Manual, 8th Edition, Revision A, 1998. Chapter 23
Aerobic Plate Count on Petrifilm™: 3M™ Quickswab™ onto 3M AC Petrifilm™ for 48 hrs @ 30°C (milk/dairy products) or 35°C (other samples)
Anaerobic Plate Count (AnPC): Standard Plate Count Agar for 48 hrs @ 30°C under anaerobic conditions
Bacillus cereus: Mannitol Egg Yolk Polymixine (MYP) Agar 24hrs @ 30°C, +ve control NCTC 11143 (NZRM 3429)

Micro Job # J20-187 **Client Reference: Fermented milk**

Appendix 28 *Microbial results from Consumer Trial – preliminary testing showing contamination of samples*

Appendix 29: YF-L811 Bitter Responses for each sample type

Pairing Responses	**Sample 1** 18	**Sample 2** 23
Pairing Responses	**Sample 1** 8	**Sample 3** 33
Pairing Responses	**Sample 1** 8	**Sample 4** 33
Pairing Responses	**Sample 1** 9	**Sample 5** 32
Pairing Responses	**Sample 1** 7	**Sample 6** 34
Pairing Responses	**Sample 2** 9	**Sample 3** 32
Pairing Responses	**Sample 2** 9	**Sample 4** 32
Pairing Responses	**Sample 2** 12	**Sample 5** 29
Pairing Responses	**Sample 2** 10	**Sample 6** 31
Pairing Responses	**Sample 3** 14	**Sample 4** 27
Pairing Responses	**Sample 3** 14	**Sample 5** 27
Pairing Responses	**Sample 3** 15	**Sample 6** 26
Pairing Responses	**Sample 4** 19	**Sample 5** 22
Pairing Responses	**Sample 4** 19	**Sample 6** 22
Pairing Responses	**Sample 5** 16	**Sample 6** 25

Appendix 30: YF-L811 Bitter Responses d` values for each sample type

d' (Sample 1 > Sample 2)	**d' (Sample 2 > Sample 1)**
-0.22	0.22
d' (Sample 1 > Sample 3)	**d' (Sample 3 > Sample 1)**
-1.22	1.22
d' (Sample 1 > Sample 4)	**d' (Sample 4 > Sample 1)**
-1.22	1.22
d' (Sample 1 > Sample 5)	**d' (Sample 5 > Sample 1)**
-1.09	1.09
d' (Sample 1 > Sample 6)	**d' (Sample 6 > Sample 1)**
-1.35	1.35
d' (Sample 2 > Sample 3)	**d' (Sample 3 > Sample 2)**
-1.09	1.09
d' (Sample 2 > Sample 4)	**d' (Sample 4 > Sample 2)**
-1.09	1.09
d' (Sample 2 > Sample 5)	**d' (Sample 5 > Sample 2)**
-0.77	0.77
d' (Sample 2 > Sample 6)	**d' (Sample 6 > Sample 2)**
-0.98	0.98
d' (Sample 3 > Sample 4)	**d' (Sample 4 > Sample 3)**
-0.58	0.58
d' (Sample 3 > Sample 5)	**d' (Sample 5 > Sample 3)**
-0.58	0.58
d' (Sample 3 > Sample 6)	**d' (Sample 6 > Sample 3)**
-0.48	0.48
d' (Sample 4 > Sample 5)	**d' (Sample 5 > Sample 4)**
-0.13	0.13
d' (Sample 4 > Sample 6)	**d' (Sample 6 > Sample 4)**
-0.13	0.13
d' (Sample 5 > Sample 6)	**d' (Sample 6 > Sample 5)**
-0.39	0.39

Appendix 31: YF-L811 Bitter Responses – ranking score for each participant for each sample type

Ranking Score	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6
YS001	0	1	2	3	4	5
YS002	0	1	2	4	3	5
YS003	2	2	2	4	2	3
YS004	2	1	2	1	5	4
YS005	0	1	3	2	4	5
YS006	1	1	2	4	3	4
YS007	0	1	3	3	4	4
YS008	0	1	2	4	4	4
YS009	0	2	3	2	3	5
YS010	2	2	4	4	1	2
YS011	2	3	3	3	2	2
YS012	0	3	3	3	3	3
YS014	1	2	5	2	2	3
YS015	2	0	4	3	4	2
YS016	1	0	2	4	4	4
YS017	0	1	2	4	4	4
YS018	3	1	3	4	1	3
YS019	1	2	2	5	2	3
YS020	3	5	4	2	1	0
YS021	1	0	3	2	4	5
YS022	0	1	4	2	4	4
YS023	2	2	4	2	2	3
YS024	4	5	2	3	1	0
YS025	2	1	2	4	3	3
YS026	0	1	3	3	5	3
YS027	1	4	3	4	2	1
YS028	0	1	2	3	4	5
YS031	1	0	3	3	3	5
YS032	3	0	1	5	4	2
YS033	1	0	2	3	5	4
YS034	0	1	2	4	3	5
YS035	2	1	5	4	2	1
YS036	0	4	2	3	3	3
YS037	1	1	1	5	3	4
YS038	0	1	2	3	5	4
YS039	1	1	4	3	1	5
YS041	1	1	1	3	5	4
YS042	1	2	2	3	3	4
YS043	4	3	4	3	1	0
YS048	1	0	2	3	4	5
YS049	4	3	1	1	3	3
Sum	50	63	108	130	126	138

Appendix 32: YF-L811 Bitter Responses – calculation of Rank Sum to input for Friedman analysis

Selected	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6
YS001	0	2	4	6	8	10
YS002	0	2	4	8	6	10
YS003	4	4	4	8	4	6
YS004	4	2	4	2	10	8
YS005	0	2	6	4	8	10
YS006	2	2	4	8	6	8
YS007	0	2	6	6	8	8
YS008	0	2	4	8	8	8
YS009	0	4	6	4	6	10
YS010	4	4	8	8	2	4
YS011	4	6	6	6	4	4
YS012	0	6	6	6	6	6
YS014	2	4	10	4	4	6
YS015	4	0	8	6	8	4
YS016	2	0	4	8	8	8
YS017	0	2	4	8	8	8
YS018	6	2	6	8	2	6
YS019	2	4	4	10	4	6
YS020	6	10	8	4	2	0
YS021	2	0	6	4	8	10
YS022	0	2	8	4	8	8
YS023	4	4	8	4	4	6
YS024	8	10	4	6	2	0
YS025	4	2	4	8	6	6
YS026	0	2	6	6	10	6
YS027	2	8	6	8	4	2
YS028	0	2	4	6	8	10
YS031	2	0	6	6	6	10
YS032	6	0	2	10	8	4
YS033	2	0	4	6	10	8
YS034	0	2	4	8	6	10
YS035	4	2	10	8	4	2
YS036	0	8	4	6	6	6
YS037	2	2	2	10	6	8
YS038	0	2	4	6	10	8
YS039	2	2	8	6	2	10
YS041	2	2	2	6	10	8
YS042	2	4	4	6	6	8
YS043	8	6	8	6	2	0
YS048	2	0	4	6	8	10
YS049	8	6	2	2	6	6

Appendix 32: YF-L811 Bitter Responses – calculation of Rank Sum to input for Friedman analysis continued

Unselected	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6
YS001	5	4	3	2	1	0
YS002	5	4	3	1	2	0
YS003	3	3	3	1	3	2
YS004	3	4	3	4	0	1
YS005	5	4	2	3	1	0
YS006	4	4	3	1	2	1
YS007	5	4	2	2	1	1
YS008	5	4	3	1	1	1
YS009	5	3	2	3	2	0
YS010	3	3	1	1	4	3
YS011	3	2	2	2	3	3
YS012	5	2	2	2	2	2
YS014	4	3	0	3	3	2
YS015	3	5	1	2	1	3
YS016	4	5	3	1	1	1
YS017	5	4	3	1	1	1
YS018	2	4	2	1	4	2
YS019	4	3	3	0	3	2
YS020	2	0	1	3	4	5
YS021	4	5	2	3	1	0
YS022	5	4	1	3	1	1
YS023	3	3	1	3	3	2
YS024	1	0	3	2	4	5
YS025	3	4	3	1	2	2
YS026	5	4	2	2	0	2
YS027	4	1	2	1	3	4
YS028	5	4	3	2	1	0
YS031	4	5	2	2	2	0
YS032	2	5	4	0	1	3
YS033	4	5	3	2	0	1
YS034	5	4	3	1	2	0
YS035	3	4	0	1	3	4
YS036	5	1	3	2	2	2
YS037	4	4	4	0	2	1
YS038	5	4	3	2	0	1
YS039	4	4	1	2	4	0
YS041	4	4	4	2	0	1
YS042	4	3	3	2	2	1
YS043	1	2	1	2	4	5
YS048	4	5	3	2	1	0
YS049	1	2	4	4	2	2

Appendix 32: YF-L811 Bitter Responses – calculation of Rank Sum to input for Friedman analysis continued

Ranking Score	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6
YS001	5	6	7	8	9	10
YS002	5	6	7	9	8	10
YS003	7	7	7	9	7	8
YS004	7	6	7	6	10	9
YS005	5	6	8	7	9	10
YS006	6	6	7	9	8	9
YS007	5	6	8	8	9	9
YS008	5	6	7	9	9	9
YS009	5	7	8	7	8	10
YS010	7	7	9	9	6	7
YS011	7	8	8	8	7	7
YS012	5	8	8	8	8	8
YS014	6	7	10	7	7	8
YS015	7	5	9	8	9	7
YS016	6	5	7	9	9	9
YS017	5	6	7	9	9	9
YS018	8	6	8	9	6	8
YS019	6	7	7	10	7	8
YS020	8	10	9	7	6	5
YS021	6	5	8	7	9	10
YS022	5	6	9	7	9	9
YS023	7	7	9	7	7	8
YS024	9	10	7	8	6	5
YS025	7	6	7	9	8	8
YS026	5	6	8	8	10	8
YS027	6	9	8	9	7	6
YS028	5	6	7	8	9	10
YS031	6	5	8	8	8	10
YS032	8	5	6	10	9	7
YS033	6	5	7	8	10	9
YS034	5	6	7	9	8	10
YS035	7	6	10	9	7	6
YS036	5	9	7	8	8	8
YS037	6	6	6	10	8	9
YS038	5	6	7	8	10	9
YS039	6	6	9	8	6	10
YS041	6	6	6	8	10	9
YS042	6	7	7	8	8	9
YS043	9	8	9	8	6	5
YS048	6	5	7	8	9	10
YS049	9	8	6	6	8	8
Rank Sum	255	268	313	335	331	343

Appendix 33: YF-L811 Bitter Responses – Friedman analysis

Samples	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6
Rank Sum	255	268	313	335	331	343
Rank Sum Sq	65025	71824	97969	112225	109561	117649

P = Total Panellists		43				
T = Treatments		6				
R = Rank Sum						
4/pt	0.01626016					
Rank Sum Sq	65025	71824	97969	112225	109561	117649
Sum Rank Sum Sq	574253					
9p (t-1) ^ 2	9225					

Test Stat =	112.447154

Difference Between Rank Sums:	Difference
Sample 1 vs 2	13
Sample 1 vs 3	58
Sample 1 vs 4	80
Sample 1 vs 5	76
Sample 1 vs 6	88
Sample 2 vs 3	45
Sample 2 vs 4	67
Sample 2 vs 5	63
Sample 2 vs 6	75
Sample 3 vs 4	22
Sample 3 vs 5	18
Sample 3 vs 6	30
Sample 4 vs 5	4
Sample 4 vs 6	8
Sample 5 vs 6	12

df = t -1 = 5	0.10	0.05	0.01
table chi-sq	9.24	11.1	15.1

conclusion: Test Stat (112.475 is > than the chi-squate critical values and is significant.

Appendix 34: YF-L811 Bitter Responses – HSD analysis

HSD			
critical val from table - q alpha (0.95), t (6) inf	from table is = 4.03		
sq(pt)/4		61.5	7.842193571

HSD Value calculated ->	**31.60404009**

Comparing the Difference between samples and HSD Value

Difference Between Rank Sums:	Difference	Greater than HSD Value?
Sample 1 vs 2	13	FALSE
Sample 1 vs 3	58	TRUE
Sample 1 vs 4	80	TRUE
Sample 1 vs 5	76	TRUE
Sample 1 vs 6	88	TRUE
Sample 2 vs 3	45	TRUE
Sample 2 vs 4	67	TRUE
Sample 2 vs 5	63	TRUE
Sample 2 vs 6	75	TRUE
Sample 3 vs 4	22	FALSE
Sample 3 vs 5	18	FALSE
Sample 3 vs 6	30	FALSE
Sample 4 vs 5	4	FALSE
Sample 4 vs 6	8	FALSE
Sample 5 vs 6	12	FALSE

** this is with 2 participants removed -- 44, 45

Appendix 35: YF-L811 Flavour intensity– Friedman analysis

Samples	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6
Rank Sum	263	253	278	315	356	380
Rank Sum Sq	69169	64009	77284	99225	126736	144400

P = Total Panellists	41
T = Treatments	6
R = Rank Sum	

4/pt	0.016260163					
Rank Sum Sq	69169	64009	77284	99225	126736	144400
Sum Rank Sum Sq	580823					
9p (t-1) ^ 2	9225					

Test Stat =	219.2764228

df = t -1 = 5	0.10	0.05	0.01
table chi-sq	9.24	11.1	15.1

conclusion: Test Stat (219.276) is > than the chi-squate critical values and is significant.

Appendix 36: YF-L811 Flavour intensity– HSD analysis

HSD			
critical val from table - q alpha (0.95), t (6) inf	from table is = 4.03		
sq(pt)/4		61.5	7.842193571

HSD Value calculated ->	**31.60404009**

Comparing the Difference between samples and HSD Value

Difference Between Rank Sums:	Difference	Greater than HSD Value?
Sample 1 vs 2	10	FALSE
Sample 1 vs 3	15	FALSE
Sample 1 vs 4	52	TRUE
Sample 1 vs 5	93	TRUE
Sample 1 vs 6	117	TRUE
Sample 2 vs 3	25	FALSE
Sample 2 vs 4	62	TRUE
Sample 2 vs 5	103	TRUE
Sample 2 vs 6	127	TRUE
Sample 3 vs 4	37	TRUE
Sample 3 vs 5	78	TRUE
Sample 3 vs 6	102	TRUE
Sample 4 vs 5	41	TRUE
Sample 4 vs 6	65	TRUE
Sample 5 vs 6	24	FALSE

Appendix 37 YC380 Bitter – Friedman analysis

Samples	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6
Rank Sum	255	274	338	354	373	386
Rank Sum Sq	65025	75076	114244	125316	139129	148996

P = Total Panellists		44				
T = Treatments		6				
R = Rank Sum						
4/pt	0.015151515					
Rank Sum Sq	65025	75076	114244	125316	139129	148996
Sum Rank Sum Sq	667786					
9p (t-1) ^ 2	9900					

Test Stat =	217.969697

Difference Between Rank Sums:	Difference
Sample 1 vs 2	19
Sample 1 vs 3	83
Sample 1 vs 4	99
Sample 1 vs 5	118
Sample 1 vs 6	131
Sample 2 vs 3	64
Sample 2 vs 4	80
Sample 2 vs 5	99
Sample 2 vs 6	112
Sample 3 vs 4	16
Sample 3 vs 5	35
Sample 3 vs 6	48
Sample 4 vs 5	19
Sample 4 vs 6	32
Sample 5 vs 6	13

df = t -1 = 5	0.10	0.05	0.01
table chi-sq	9.24	11.1	15.1

conclusion: Test Stat (217.9) is > than the chi-square critical values and is significant.

Appendix 38 YC380 Bitter – HSD analysis

HSD		
	from table is = 4.03	
critical val from table - q alpha (0.95), t (6)		
inf	66	8.124038405
sq(pt)/4		
HSD Value calculated ->	**32.73987477**	

Comparing the Difference between samples and HSD Value

Difference Between Rank Sums:	Difference	Greater than HSD Value?
Sample 1 vs 2	19	FALSE
Sample 1 vs 3	83	TRUE
Sample 1 vs 4	99	TRUE
Sample 1 vs 5	118	TRUE
Sample 1 vs 6	131	TRUE
Sample 2 vs 3	64	TRUE
Sample 2 vs 4	80	TRUE
Sample 2 vs 5	99	TRUE
Sample 2 vs 6	112	TRUE
Sample 3 vs 4	16	FALSE
Sample 3 vs 5	35	TRUE
Sample 3 vs 6	48	TRUE
Sample 4 vs 5	19	FALSE
Sample 4 vs 6	32	FALSE
Sample 5 vs 6	13	FALSE

Appendix 39 YC380 Flavour intensity – Friedman analysis

Samples	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6
Rank Sum	253	277	315	343	376	416
Rank Sum Sq	64009	76729	99225	117649	141376	173056
P = Total Panellists	45					
T = Treatments	6					
R = Rank Sum						
4/pt	0.015151515					
				117649	141376	173056
Rank Sum Sq	64009	76729	99225			
Sum Rank Sum Sq	672044					
9p (t-1) ^ 2	9900					
Test Stat =	282.4848485					
df = t -1 = 5	0.10	0.05	0.01			
table chi-sq	9.24	11.1	15.1			

conclusion: Test Stat (282.48) is > than the chi-square critical values and is significant.

Appendix 40 YC380 Flavour intensity – HSD analysis

HSD		
critical val from table - q alpha (0.95), t (6) inf	from table is = 4.03	
sq(pt)/4	66	8.124038405

HSD Value calculated ->	**32.73987477**

Comparing the Difference between samples and HSD Value

Difference Between Rank Sums:	Difference	Greater than HSD Value?
Sample 1 vs 2	24	FALSE
Sample 1 vs 3	62	TRUE
Sample 1 vs 4	90	TRUE
Sample 1 vs 5	123	TRUE
Sample 1 vs 6	163	TRUE
Sample 2 vs 3	38	TRUE
Sample 2 vs 4	66	TRUE
Sample 2 vs 5	99	TRUE
Sample 2 vs 6	139	TRUE
Sample 3 vs 4	28	FALSE
Sample 3 vs 5	61	TRUE
Sample 3 vs 6	101	TRUE
Sample 4 vs 5	33	TRUE
Sample 4 vs 6	73	TRUE
Sample 5 vs 6	40	TRUE

Appendix 41 MALDI-TOF processing, R code used to process – as an exemplar, the code used to process data in Chapter 4 is presented.

```
# Processing MALDI-TOF data - before technical averaging


SampleDir<-("C://Users//murphyf//OneDrive                                    -
AgResearch//MALDI_Data_R_Drive//Time Course Exp Nov2019//All Reps Jan 2020")


#### SampleNames <- factor(sapply(samples,function(x)metaData(x)$file))

library(MALDIquantForeign)

library(MALDIquant)

# importing samples using MALDIquantForeign

samples <- importMzXml(SampleDir)

# importing metaData (MetaData_Samples)

metaSamples =read.csv(file.choose( ), header=T)

# some samples were removed - one whole day's rep was removed. Bio Rep 2,
Maldi Rep 2

metaSamples$MALDI_Bio_Rep<-
paste0(metaSamples$Bio_Rep,metaSamples$MALDI_Rep)

which(metaSamples$MALDI_Bio_Rep == "11")

which(metaSamples$MALDI_Bio_Rep == "22")

metaSamples[,643]

metaSamples[643,]

metaSamples[642,]

metaSamples[750,]

metaSamples[751,]

750-643

643-750

spectra <- samples[-643:-750]

type <- metaSamples[-643:-750,]

# screening spectra for poor quality spectra. removed from the metadata and
spectra

sc.results <- screenSpectra(spectra, meta = type)

(sc.results     <-     screenSpectra(SamplesForProcessing,     meta     =
metaSamples$SampleName, threshold = 3, estimator = "MAD", method = "Hampel"))

# summary of faulty spectra

summary(sc.results)

plot(sc.results, labels = TRUE)
```

```
# Filtered list of mass spectra

spectra <- sc.results$fspectra

type <- sc.results$fmeta

# getting sample info - spectra to be averaged

SampleReps <- type$Sample_to_Ave

# MALDIquant processing

spectra2 <- transformIntensity(spectra, method="sqrt")

spectra2 <- MALDIquant::smoothIntensity(spectra2 ,method = "SavitzkyGolay",halfWindowSize = 20)

spectra2 <- MALDIquant::removeBaseline(spectra2,method = "TopHat")

spectra2 <- MALDIquant::calibrateIntensity(spectra2,method = "TIC")

spectra2<- MALDIquant::alignSpectra(spectra2,halfWindowSize=20,SNR=2,tolerance=0.002, warpingMethod="lowess")

avgSpectra <- MALDIquant::averageMassSpectra(spectra2, labels=SampleReps, method="mean")

names(avgSpectra)

# peak detection - I had tried SNR = 1 too, but it's allowing peaks that are very low intensity (false peaks) so I'm opting for a slightly more strict SNR.

peaks <- MALDIquant::detectPeaks(avgSpectra,method = "SuperSmoother",halfWindowSize = 20,SNR = 2)

peaks <- MALDIquant::binPeaks(peaks, tolerance=0.002)

peaks <- MALDIquant::filterPeaks(peaks, minFrequency=0.25)

featureMatrix <- MALDIquant::intensityMatrix(peaks2, avgSpectra)


rownames(featureMatrix) <- SampleNames


#end result is feature matrix with 490 samples and 143 peaks.

> dim(featureMatrix)

[1] 490 143


# making a dupe of the featurematrix

featMat <- featureMatrix


# renaming these rownames is important because the order of the outputted avgSpectra/Feature matrix was not in order.
```

```
rownames(featMat) <- SampleNames


which(rownames(featMat) == "124_4_3")

which(rownames(featMat) == "124_4_2")


# Samples "124_4_3"  and   "124_4_2" (125 and 126 in featureMatrix), and "234_4_3", "234_4_2" (456 after others (125, 126) are removed)

featMat <- featMat[-125, 126]

featMat <- featMat[-456,457]


# getting names of the filtered featMat. Getting info on hours/starter/bioRep and MALDIrep

SampleNamesNew <- rownames(featMat)

length(SampleNamesNew)

dim(featMat)

HoursNew <- substr(SampleNamesNew, 3,3)

StarterNew <- substr(SampleNamesNew, 1,1)

Bio_RepNew <- substr(SampleNamesNew, 5,5 )

MALDI_RepNew <- substr(SampleNamesNew, 7,7 )


# performing DA to extract the top 40 most discriminant peaks by time point

library('sda')

ddar   <-   sda.ranking(Xtrain=featMat,   L=as.factor(HoursNew),fdr=FALSE, diagonal=TRUE) # DDA

Top40_ddar <- rownames(ddar[1:40,])

Top40Peaks <- featMat[,c(Top40_ddar)]

colnames(Top40Peaks) <- round(as.numeric(colnames(Top40Peaks)), digits=2)
```