AN ABSTRACT OF THE THESIS OF

Jungmin Choi for the degree of Master of Science in Food Science and Technology presented on May 12, 2020.

Title: High-throughput Sequencing to Characterize the Microbial Diversity and Functional Properties of Cheese

Abstract approved:

_____

Si Hong Park

Cheese is a fermented dairy product that is made from animal milk and is considered to be healthy food due to its available nutrients and potential probiotic characteristics. Since the microbes in the cheese matrix directly contribute to the quality and physicochemical properties of cheese, it is important to understand the microbial properties of cheese.

The purpose of this study understands and compare the microbial compositions of different cheeses (Cheddar, Provolone, and Swiss) and cheese locations (core, rind, and mixed). The second research object was analyzed microbial community shift during Cheddar cheeses making. All the cheese samples produced at the Arbuthnot

Dairy Center at Oregon State University were collected to determine the microbial community structure using 16S rRNA gene amplicon sequencing with the Illumina MiSeq platform (Illumina, San Diego, CA).

To compare the microbial composition of different types of cheese, a total of 225 operational taxonomic units were identified from the 4,675,187 sequencing reads generated. *Streptococcus* was observed to be the most abundant organism in Provolone (72 to 85%) and Swiss (60 to 67%), whereas *Lactococcus* spp. were found to dominate Cheddar cheese (27 to 76%). Species richness varied significantly by cheese. According to alpha diversity analysis, porter-soaked Cheddar cheese exhibited the highest microbial richness, whereas smoked Provolone cheese showed the lowest. Rind regions of each cheese changed color through smoking and soaking for the beverage process. Also, the microbial diversity of the rind region was higher than the core region because smoking and soaking processes directly contacted the rind region of each cheese. The microbial communities of the samples clustered by cheese indicated that, within a given type of cheese, microbial compositions were very similar. Moreover, 34 operational taxonomic units were identified as biomarkers for different types of cheese through the linear discriminant analysis effect size method. Last, both carbohydrate and AA metabolites comprised more than 40% of the total functional annotated genes from 9 varieties of cheese samples. This study provides insight into the microbial composition of different types of cheese, as well as various locations within a cheese, which applies to its safety and sensory quality.

As long as the microbial community shift during Cheddar cheese-making, a total of 773,821 sequencing reads and 271 amplicon sequence variants (ASVs) were

acquired from 108 samples. *Streptococcus* and *Lactococcus* were observed as the most abundant ASVs in the cheese, which were used as the starter lactic acid bacteria (SLAB). *Escherichia coli* was detected in the raw milk, however, it was not detected after inoculating with SLAB. According to an alpha diversity analysis, SLAB inoculation decreased the microbial richness by inhibiting the growth of other bacteria present in the milk. A beta diversity analysis showed that microbial communities before the addition of SLAB, clustered together, as did the samples from cheese making and aging. Non-starter lactic acid bacteria (NSLAB) were detected 15 weeks into aging for the June 6th and June 26th produced cheeses, and 17 weeks into aging for the cheese produced on April 26th. These NSLAB were identified as an unidentified group of *Lactobacillaceae*. This study characterizes the changes in the Cheddar cheese microbiome throughout production from raw milk to a six-month-aged final product.

High-throughput Sequencing to Characterize the Microbial Diversity and Functional Properties of Cheese

by

Jungmin Choi

A THESIS

submitted to

Oregon State University

in partial fulfillment of

the requirements for the

degree of

Master of Science

Presented May 12, 2020

Commencement June 2020

Master of Science thesis of Jungmin Choi presented on May 12th, 2020

APPROVED:

Major Professor, representing Food Science and Technology

Head of the Department of Food Science and Technology

Dean of the Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

Jungmin Choi, Author

ACKNOWLEDGMENTS

First and foremost, I would like to show my sincere gratitude to my supervisor Dr. Si Hong Park, assistant professor of Food Science and Technology, Oregon State University for the continuous support of my study and research. Thanks to his critical feedback for my research, his patience, motivation, and enthusiasm helped me all the time for my study and writing of this thesis.

I also thank my lab members in Dr. Si Hong Park's lab: Sang In Lee, Ian Moppert, Daria Van de Grift, and Bryna Rackerby helped me to work at the lab and stimulating discussions for my research.

It is a pleasure to thank Robin Frojen who provided cheese samples for my research and helped me to access the cheese making facility (Arbuthnot Dairy Center at Oregon State University) and provided practical advice about cheese.

I am deeply grateful to Dr. Lisbeth Goddik, Head of the Department of Food Science and Technology, Oregon State University provided the critical reviews for published manuscripts.

Last but not the least, I am deeply grateful to my parents who supported me to study at Oregon State University.

CONTRIBUTION OF AUTHORS

The authors thank Sang In Lee, Robin Frojen, Lisbeth Goddik, and Si Hong Park who conceived and contributed the research. Sang In Lee helped to conduct experiments and Robin Frojen helped cheese sample collection. Sang In Lee and Bryna Rackerby analyzed the data and drafted the manuscript. Dr. Irene Hanning (University of Tennessee, Knoxville) and Daria Van de Grift, Ian A. Moppert, Si Hong Park, Bryna Rackerby, Robin Frojen, Lisbeth Goddik, Sang-Do Ha and Sushumna Canakapalli (Oregon State University, Corvallis) critically reviewed the manuscript. This research was partially supported by the Korean-American Scientists and Engineers Association (KSEA) Young Investigator Grants (YIG), BUILD (Building University-Industry linkages through Learning and Discovery) Dairy grant, and Oregon State University (Corvallis) start-up package awarded to Si Hong Park and Chung-Ang University Research Program (Anseong, Gyeonggi-Do, South Korea) awarded to Sang-Do Ha.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

Page

LIST OF FIGURES

LIST OF FIGURES (Continued)

Figure Page

LIST OF TABLES

Chapter 1. Microbial communities of a variety of cheeses and comparison between core and rind region of cheeses

Jungmin Choi[1], Sang In Lee[1], Bryna Rackerby[1], Lisbeth Goddik[1], Robin Frojen[1], Sang-Do Ha[2], Jang H. Kim[3], and Si Hong Park[1]*

[1]Department of Food Science and Technology, Oregon State University, Corvallis 97331

[2]Department of Food Science and Technology, Advanced Food Safety Research Group, Brain Korea 21 Plus, Chung-Ang University, South Korea 06974

[3]School of Family and Consumer Sciences, College of Agricultural and Life Sciences, University of Idaho, Moscow 83844

Journal of Dairy Science

https://doi.org/10.3168/jds.2019-17455

Volume 103, Issue 5, May 2020

1.1 Abstract

Understanding the microbial community of cheese is important in the dairy industry, as the microbiota contributes to the safety, quality, and physicochemical and sensory properties of cheese. In this study, the microbial compositions of different cheeses (Cheddar, Provolone, and Swiss cheese) and cheese locations (core, rind, and mixed) collected from the Arbuthnot Dairy Center at Oregon State University were analyzed using 16S rRNA gene amplicon sequencing with the Illumina MiSeq platform (Illumina, San Diego, CA). A total of 225 operational taxonomic units were identified from the 4,675,187 sequencing reads generated. *Streptococcus* was observed to be the most abundant organism in Provolone (72 to 85%) and Swiss (60 to 67%), whereas *Lactococcus* spp. were found to dominate Cheddar cheese (27 to 76%). Species richness varied significantly by cheese. According to alpha diversity analysis, porter-soaked Cheddar cheese exhibited the highest microbial richness, whereas smoked Provolone cheese showed the lowest. Rind regions of each cheese changed color through smoking and soaking for the beverage process. In addition, the microbial diversity of the rind region was higher than the core region because smoking and soaking processes directly contacted the rind region of each cheese. The microbial communities of the samples clustered by cheese, indicated that, within a given type of cheese, microbial compositions were very similar. Moreover, 34 operational taxonomic units were identified as biomarkers for different types of cheese through the linear discriminant analysis effect size method. Last, both carbohydrate and AA metabolites comprised more than 40% of

the total functional annotated genes from 9 varieties of cheese samples. This study provides insight into the microbial composition of different types of cheese, as well as various locations within a cheese, which is applicable to its safety and sensory quality.

Key words: cheese, microbiome, Illumina MiSeq

1.2 Introduction

Cheese is a nutrient-rich food that contains vitamins, minerals, proteins, bioactive peptides, amino acids (AA), fats, and fatty acids (Walther et al. 2008). The microorganisms present in cheese not only influence the flavor profile through the production of volatile compounds (Percival and Percival 2017), but also potentially contribute to human health associated with anti-cancer and cholesterol-lowering properties (Walther et al. 2008; Broadbent et al. 2011; Potočki, 2016).

The microbiomes and metabolomes of cheeses vary and are based on the cheese type as well as environmental and processing conditions such as starter cultures, pasteurization methods, cooking temperatures, and aging conditions (Didienne et al. 2012; Montel et al. 2014; Duru et al. 2018). Bacteria primarily originate from 2 sources: the inoculated starter cultures and the indigenous milk microbiota (Montel et al. 2014; De Filippis et al. 2016). Microorganisms originating from processing environments are also transferred from production surfaces to cheese surfaces where they affect the microbial composition of rinds during aging (Bokulich and Mills, 2013). Bacteria are essential for the formation of cheese and are largely responsible for flavor development and nutritional

benefits (Walther et al. 2008; Didienne et al. 2012; Montel et al. 2014). In addition, microorganisms originated from processing environments dominate both surfaces of cheese and facilities and affect the microbial composition of cheese rinds during aging (Bokulich and Mills, 2013).

The rind and core regions of a given cheese exhibit different microbial compositions. This is due in part to the differences in oxygen availability throughout the cheese. The presence of oxygen at the cheese surface allows for the growth of aerobic organisms that are unable to grow deeper where less oxygen can penetrate. A natural rind is developed through interactions between the surface and the environment during the aging process. Another type of rind is the washed rind, which usually has a sticky texture and strong flavor. When a cheesemaker soaks (or washes) a cheese with a brine or alcohol, the bacteria from the environment or soaking material grow on the surface of the cheese and develop the rind (Donnelly, 2014). Bacteria commonly found on cheese rinds include Gram-negative bacteria, such as *Advenella*, *Psychrobacter*, and *Psychroflexus*, which contain various lipases, proteases, and other enzymes that enhance aging (Schmitz-Esser et al. 2018). Moisture content of cheese also affects the microbial diversity (Pintado and Malcata, 2000). Vacuum packaging affects the microbial and physical properties of the cheese rind, as moisture content and texture do not change significantly while under a vacuum. Conversely, unpackaged cheeses exhibit a substantial loss of water with a concomitant increase in rigidity.

The development of next-generation sequencing technologies has helped researchers obtain genomic information quickly at a low cost and has furthered our understanding of the microbial properties of target food matrices. High-throughput

sequencing has made it possible to explore food microbiomes and to investigate the genomes of individual organisms. Next-generation sequencing can reveal how microbes respond to environmental conditions, allowing cheesemakers to better control microbial growth in their products based on predictions of how conditions will affect the growth of both beneficial and undesirable organisms (Solieri et al. 2013). It is important to understand what conditions can cause changes in the cheese microbiome since some organisms impart beneficial sensory characteristics, whereas others may reduce quality through the production of spoilage compounds. Knowledge of food microbiomes is important with foods that are fermented or aged, as these are primarily microbe-driven processes and the microbes present will drastically affect the outcome (Mayo et al. 2014).

In this study, we examined 3 different cheeses (Cheddar, Provolone, and Swiss) produced plain or treated by soaking in cider, pinot noir (wine), or porter (beer), or by smoking. All the cheese samples were produced at the Arbuthnot Dairy Center at Oregon State University (Corvallis, OR). Microbial communities were analyzed using the 16S rRNA gene amplicon sequencing via the Illumina MiSeq platform (Illumina, San Diego, CA) to determine the differences in microbial communities between the core and rind regions of each cheese. The different types of cheese and their various treatments were compared with one another to better understand which microbial populations affect the properties of cheese.

1.3 Materials and methods

1.3.1 Cheese sampling

A variety of cheeses were obtained from the Arbuthnot Dairy Center at Oregon State University (Corvallis, OR, Figure 1.1). All cheese samples were produced using raw milk pasteurized at 65°C for 30 min, and subsequently cooled to 32°C. Different starter cultures were used depending on cheese type: *Lactococcus lactis* ssp. *lactis* biovar diacetylactis, *Lc. lactis* ssp. *lactis*, *Lc. lactis* ssp. *cremoris*, and *Streptococcus thermophilus* for Cheddar cheese; *S. thermophilus* and *Lactobacillus delbrueckii* ssp. *bulgaricus* for Provolone; and *Lc. Lactis* ssp. *lactis*, *Lc. lactis* ssp. *cremoris*, *S. thermophilus*, *Lb. delbrueckii* ssp. *lactis*, *Lb. helveticus*, and *Lb. delbrueckii* ssp. *lactis* without the addition of propionic acid bacteria for Swiss type cheese. Additionally, the starter cultures were incubated at 32°C for 1 h and rennet was added to coagulate the curd. After curd formation, the curds were cooked (Cheddar and Swiss: 32 to 40°C for 45 min; Provolone: 50°C for 30 min). Using 4.0% (wt/wt) of curd weight, dry salt was added to Cheddar and Provolone cheeses and transferred to a mold to produce hard cheeses. Swiss cheese was formed hard cheese through a mold and transferred to a 25% saturated salt brine solution for 20 h. Afterward, cheeses were sealed in vacuum packaging and stored at 4°C for 3 to 6 mo. To provide flavor and modify the quality, cheese was soaked in beverages including pinot noir [wine, 14% alcohol by volume

Figure 1.1. Nine different varieties of cheese produced at the Arbuthnot Dairy Center at Oregon State University (Corvallis, OR, USA). Sampling included cheese from rind and core regions (A) and from 3 different types of cheese: (B) Cheddar, (C) Provolone, and (D) Swiss cheese.

(ABV)], hard apple cider (cider, 7% ABV), and porter (dark beer, 5% ABV). Fresh cheese blocks were soaked in their respective beverages at room temperature for 48 h. After 48 h, the cheeses were removed from the beverages and allowed to air dry at room temperature for 3 h until aerobic microorganisms can be grown. Traditional smoking methods were also applied to change the properties of Cheddar, Provolone, and Swiss type cheese. Cheeses were smoked at 30°C for 6 h using applewood. Beverage-soaked cheese and smoked cheese samples were sealed with vacuum packaging and stored at 4°C for 3 to 6 mo. Finally, cheeses were divided into 3 primary groups: (1) Cheddar, (2) Provolone, and (3) Swiss. The Cheddar cheeses were subdivided into plain, smoked, and pinot noir-, cider-, and porter-soaked, whereas Provolone and Swiss only included plain and smoked type cheese. Samples were obtained from the core and rind regions of each cheese. Equal parts of rind and core regions were mixed to generate mixed regions of cheeses. Beverage-soaked cheese samples (pinot noir-, cider-, and porter-soaked Cheddar), and smoked cheese (Cheddar, Provolone, and Swiss) were collected, taking the core, mixed, and rind regions separately. We collected only mixed region from plain cheese because no differences were observed between the core and rind regions of plain cheese, whereas smoking and soaking methods have a chance to alter the properties of the surface of the cheese. To avoid any technical errors, samples were obtained in triplicate from all of the cheese types, making a total of 63 samples collected. The DNA was extracted from these samples and a microbiome sequencing library was constructed to compare the differences between the core, rind, and mixed regions.

1.3.2 DNA extraction

For DNA extraction, 1 g of core, rind, and mixed regions of each cheese were separated using a sterile razor blade. Cheese samples were homogenized with 9 mL of 1% tri-citric acid buffer ($C_6H_7Na_3O_8$) while incubating in a 55 to 60°C water bath with vigorous vortexing. Extraction of DNA was performed on the homogenized cheese solutions using the DNeasy PowerFood Microbial Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. The concentration of extracted DNA was measured via a Qubit 4 Fluorometer (Thermo Fisher Scientific, Waltham, MA), and subsequently diluted to achieve a final concentration of 10 ng/μL.

1.3.3 16S rRNA gene-based library preparation

The sequencing library was prepared targeting the V4 region of the 16S rRNA gene as previously reported by Kozich et al. (2013). Briefly, DNA amplicons were generated using a high-fidelity polymerase (AccuPrime, Invitrogen, Carlsbad, CA), and PCR products were confirmed using 1% agarose gel electrophoresis. Amplified DNA samples were normalized using a SequalPrep Normalization kit (Life Technologies, Carlsbad, CA) according to the manufacturer's recommendation. Following normalization, 5 μL of each normalized aliquot from each sample were combined to construct a pooled library and quantified using the KAPA Library Quantification kit (Kapa Biosystems, Woburn, MA). Finally, the library was diluted to the appropriate concentration and sequenced by the Illumina MiSeq.

1.3.4 Microbiome sequencing via Illumina MiSeq

A 20 n*M* pool of the 16S rRNA library and 20 n*M* PhiX control V3 (Illumina) were mixed with 0.2 *N* of fresh NaOH and HT1 buffer (Illumina) to produce the final concentration of 6 p*M*. The resulting library was mixed with the PhiX control v3 (5%, vol/vol, Illumina) and 600 μL loaded on a MiSeq v2 (500 cycle, 2 × 250 bp) reagent cartridge for sequencing. The 16S rRNA amplicon sequences are available at the Sequence Read Archive of the National Center for Biotechnology Information (https://dataview.ncbi.nlm.nih.gov/object/PRJNA591223?reviewer=ktedeigveld5k4n89ljtkr3f0l) under SUB6594737.

1.3.5 Data analyses

Both demultiplexed R1 and R2 raw sequences were acquired directly from the Illumina BaseSpace website, and sequences were analyzed using the Quantitative Insights Into Microbial Ecology 2 (QIIME 2, version 2018.11) open source pipeline (Bolyen et al. 2019). Demultiplexed sequences were joined and denoised for quality control using the DADA2 scripts available in QIIME 2 (v. 2018.11) to cluster operational taxonomic units (OTU) at 100% sequence similarity with default parameters and generate a feature table for further analysis. The processed sequencing data were assigned a taxonomy and aligned to the Greengenes reference database (v. 13.8; http://greengenes.lbl.gov) at 99% sequence similarity (McDonald et al. 2012). For further

statistical analysis and visual exploration, an OTU table with taxa in plain format and metadata file were uploaded to the MicrobiomeAnalyst tool available at http://www.microbiomeanalyst.ca (Dhariwal et al. 2017). Linear discriminant analysis (LDA) of effect size (LEfSe) was applied to determine the most discriminant taxa among different cheese samples based on the relative abundance. The LDA score, which uses statistical significance and biological relevance to find biomarker genes, was set to a default value of 2.0 (Segata et al. 2011).

1.3.6 Predicted functional properties of cheese microbiome

The functional potential of microbiome data from different types of cheese samples was predicted based on the 16S rRNA data using Piphillin (with 97% identity cut-off; Iwai et al. 2016) and phylogenetic investigation of communities by reconstruction of unobserved states 2 (PICRUSt2) based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (Langille et al. 2013). The predicted functional properties (PFP) tools based on 16S rRNA marker genes were used to predict the full genome sequence through comparing the fully identified and sequenced bacteria with unidentified bacteria using the phylogenetic tree (PICRUSt2) or directly compared with each other (Piphillin). Sequences then were compared with the assigned predicted full genome sequencing data within the KEGG or BioCyc reference database to assign functional properties. While PICRUSt2 requires the Greengenes database and preprocessed data from QIIME 2 pipeline, the Piphillin tool does not require any pre-

processed data and support the KEGG and BioCyc reference databases (Iwai et al. 2016; Douglas et al. 2018). Statistical analysis of metagenomic profiles (STAMP) was used to identify microbial relevant functions between the 9 different types of cheese, and it also can be used to generate principal component analysis (PCA) analysis based on the KEGG orthology from Piphillin and PICRUSt2 analyses (Parks et al. 2014).

1.4 Results

1.4.1 Taxonomic analysis

A total of 4,675,187 sequencing reads were generated from the 63 cheese samples collected in this study, which included 9 different types of cheese (plain, smoked, cider-, porter-, and pinot noir-soaked Cheddar, plain and smoked Provolone, and plain and smoked Swiss). The mean value for the frequency of sequences per sample was 71,081 reads/sample after analysis with QIIME 2. A total of 225 OTU were identified and included 86 OTU at the genus level. The processed sequencing data were compared with the Greengenes reference database (13.8; http://greengenes.lbl.gov; McDonald et al. 2012). In the taxonomic analysis, all data were analyzed at genus level. *Streptococcus* spp., *Lactococcus* spp., unidentified *Lactobacillaceae*, and *Lactobacillus* spp. were the most abundant taxa identified among all the cheese samples (Figure 1.2) and the rest of the OTU were categorized as others (Figure 1.2). The composition of the cheese microbiome was largely dependent on the starter culture used. *Lactococcus lactis* ssp.

lactis biovar diaceylactis, *Lc. lactis* ssp. *lactis*, *Lc. lactis* ssp. *cremoris*, and *S. thermophilus* spp. were used as starter cultures in Cheddar cheese production, while Provolone starter cultures were composed of *S. thermophilus* and *Lb. delbrueckii* ssp. *bulgaricus*, and *cremoris*, *S. thermophilus*, *Lb. delbrueckii* ssp. *lactis*, *Lb. helveticus*, and *Lb. delbrueckii* ssp. *lactis* at once with no additional propionic acid bacteria during manufacturing. *Streptococcus* spp. were present in the highest percentages in most of cheese samples, ranging from 15 to 85% of the relative abundance (Figure 1.2). The only sample with a low relative abundance was the pinot noir-soaked Cheddar, where it ranged from 2 to 13% for both the core and mixed regions of the cheese. The average value for the rind region of the pinot noir-soaked Cheddar cheese was 2%. The second lowest concentration was in the smoked Cheddar samples, where *Streptococcus* spp. comprised 19 to 24% of the organisms present.

The Provolone samples (plain and smoked) exhibited the highest abundance for *Streptococcus* spp., with values ranging from 78 to 83%. All other samples showed between 23 to 67% of relative abundance, with most falling in the upper end of the range. In addition, the relative abundance of *Streptococcus* spp. in the core or mixed regions were higher than in the rind region.

The second most prevalent genus was *Lactococcus* spp., which was present in all samples with values ranging from 1 to 73% (Figure 1.2). Provolone included the fewest *Lactococcus* spp., with less than 1% in all samples. A likely explanation for this is the

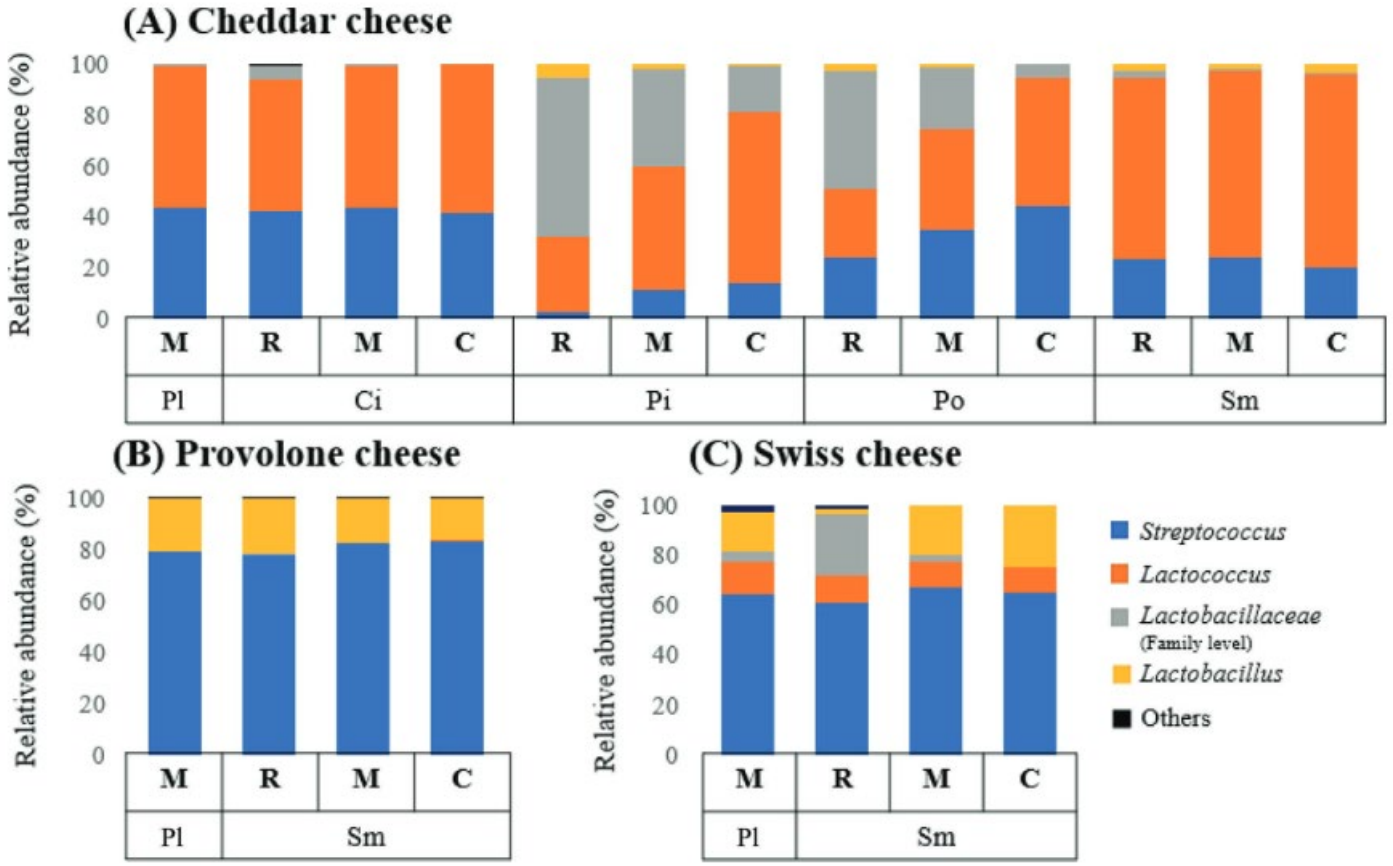


Figure 1.2. Genus level of relative abundance (mean value for triplicate) of the 4 most abundant bacteria (*Streptococcus*, *Lactococcus*, *Lactobacillaceae*, and *Lactobacillus*), which accounted for more than 1% in cheese samples. Pl = plain; Ci = cider-soaked; Pi = pinot noir-soaked; Po = porter-soaked; and Sm = smoked. Samples are separated by the mix (M = core and rind mixed), core (C), and rind (R) regions.

absence of *Lactococcus* spp. in the Provolone starter culture, which contained *S. thermophilus* and *Lb. delbrueckii* ssp. *bulgaricus*. While the highest values reported for *Lactococcus* spp. were from the smoked Cheddar (71 to 76%), all other Cheddar varieties contained only 27 to 67%, and all Swiss cheese samples ranged from 9 to 13%. Additionally, the abundance of *Lactococcus* spp. in the core and mixed regions was higher than the rind regions across all cheese samples. The prevalence of *Lactobacillaceae* family was dominant on the rind region of the porter- and pinot noir-soaked Cheddar cheeses (46 and 63%, respectively) than core regions. This family was also identified on the rind of the smoked Swiss (24%) as well as less abundant in the core region (Figure 1.2).

Lactobacillus spp. were less abundant than *Streptococcus* spp. and *Lactococcus* spp. across all samples, with values ranging from 0 to 22%. *Lactobacillus* spp. were present in low abundance in plain Cheddar cheese samples (less than 0.1%), and more prevalent in pinot noir-soaked (0 to 5%), porter-soaked (0 to 2%), and smoked (2 to 3%) Cheddar cheeses. In addition, the relative abundance of *Lactobacillus* spp. in the rind region of pinot noir-soaked (5%) and porter-soaked (2%) Cheddar cheese was higher than in the core region (0%). Compared with the Cheddar cheese, Swiss (2 to 24%), and Provolone (17 to 22%) cheeses possessed more *Lactobacillus* spp. The relative abundance of *Lactobacillus* spp. in smoked Provolone was higher in the rind region (22%) than the core region (17%) of the same cheese, whereas in smoked Swiss it was higher in the core region (24%) than the rind region (3%).

1.4.2 Alpha diversity

Alpha diversity of the microbial communities was assessed using the Chao1 and Shannon indexes (Figure 1.3). The Chao1 (Figure 1.3A and 1.3B) indicates the microbial richness among the different types of cheese. No significant differences ($P < 0.05$) were observed between plain and smoked Provolone or between plain and smoked Swiss cheese; however, the different types of cheese (Cheddar, Provolone, and Swiss) had significant differences ($P < 0.05$). The richness (Chao1 index) of Swiss and smoked Swiss cheese (5 to 10) was higher than that of Provolone and smoked Provolone (4 to 7). The richness of plain, cider-soaked, and smoked Cheddar cheeses were low (5 to 10), whereas pinot noir- and porter-soaked Cheddar cheeses were high (8 to 10). Finally, the richness of all cheese samples showed high values in the rind region when compared with the core region, except for porter-soaked Cheddar cheese. The Shannon index (Figure 1.3C and 1.3D) accounts for both richness and evenness of OTU. High values for diversity indicate more diverse communities. The Shannon index indicates that the rind is more diverse than the core across all cheeses. The Shannon index of Provolone and smoked Provolone (0.40 to 0.56) were significantly lower ($P < 0.05$) than Swiss and smoked Swiss cheese (0.74 to 1.06). Smoked versus non-smoked samples were not significantly different for any of the cheeses. The Shannon index of Cheddar cheese differed significantly ($P < 0.05$) depending on the treatment. The Shannon index of plain Cheddar (0.69 to 0.71), cider-soaked Cheddar (0.69 to 0.92), and smoked Cheddar (0.67

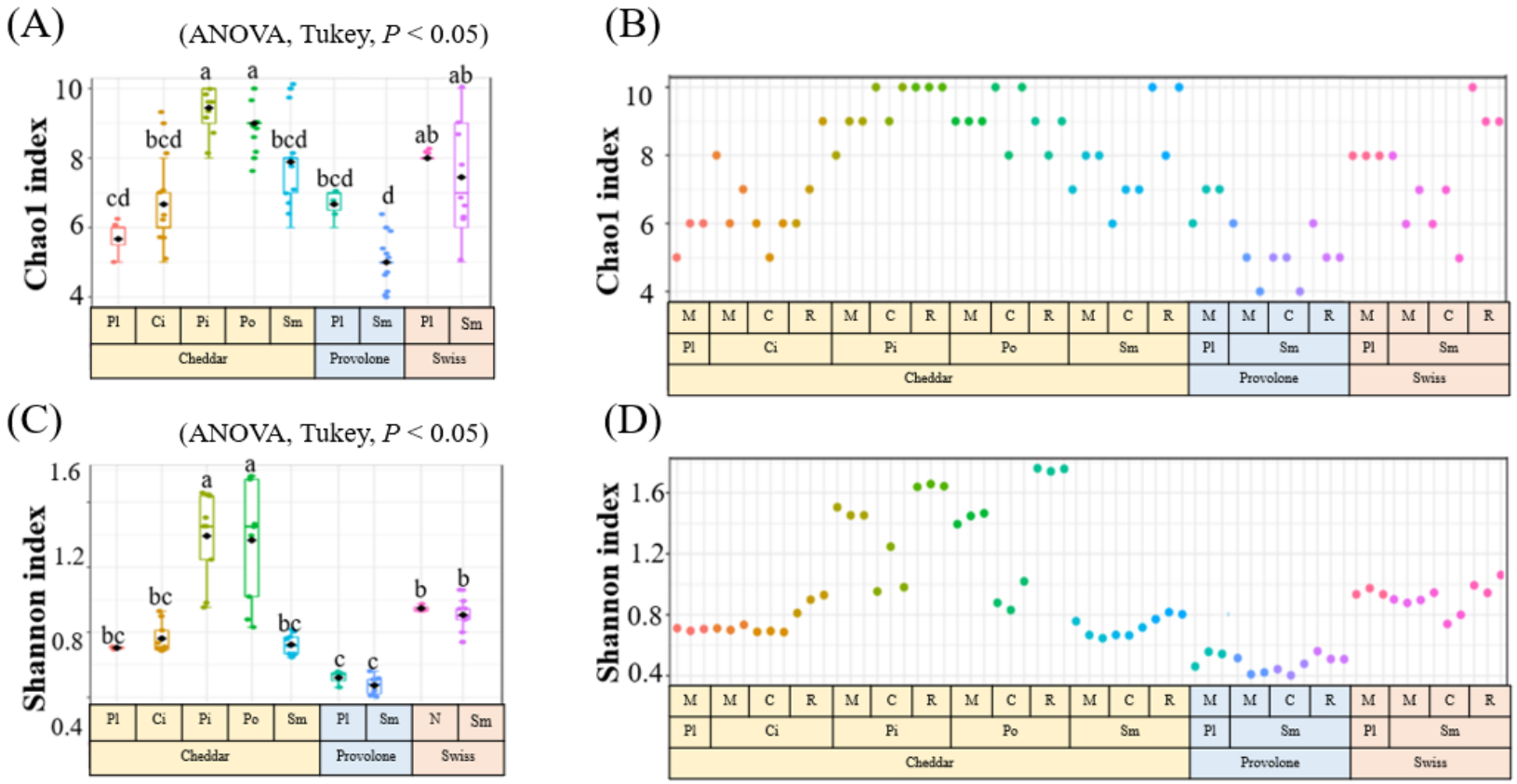


Figure 1.3. Alpha diversity of different types of cheese. The Chao1 index (Figure 1.3A and 1.3B) and Shannon index (Figure 1.3C and 1.3D) of 9 cheese types (Pl = plain cheese; Ci = cider-soaked cheese; Pi = pinot noir-soaked cheese; Po = porter-soaked cheese; and Sm = smoked cheese). Samples are separated by the type of cheese (Figure 1.3A and 1.3C) or type of cheese with mix (M), core (C), and rind (R) regions (Figure 1.3B and 1.3D). Boxes in the plots represent the inter quartile range (IQR) between first and third quartiles, respectively. The horizontal line and dots in the box indicate the median and mean value, respectively. The whiskers mean the lowest and highest value within 1.5 times the IQR from the first and third quartiles, respectively.

to 0.82) cheese were significantly lower than pinot noir- (0.95 to 1.66) and porter-soaked (0.84 to 1.76) Cheddar cheeses.

1.4.3 Beta diversity

When beta diversity parameters were used to assess the microbiome structure of the cheese samples, significant differences were found between the core and rind regions. Bray-Curtis models (Figure 1.4A) derived from the data input into MicrobiomeAnalyst (https://www.microbiomeanalyst.ca) showed obvious grouping between samples. According to the dendrogram (Figure 1.4B), clusters were identified by the types of cheese (Provolone, Swiss, and Cheddar cheese).

1.4.4 Identification of biomarkers

Biomarker bacteria from the different types of cheese were assessed using LEfSe. When LEfSe was applied to the microbiota data of 9 different types of cheese, 34 different taxonomic clades with an LDA score higher than 2.0 were found (Figure 1.5) and a total of 10 biomarkers were identified at the genus level. For instance, 19 OTU in plain Swiss cheese, 1 OTU in smoked Swiss cheese, 6 OTU in smoked Provolone, 1 OTU in smoked Cheddar cheese, 1 OTU in Provolone, 3 OTU in porter-soaked Cheddar, 1 OTU in pinot noir-soaked Cheddar, and 2 OTU in plain Cheddar cheese were identified as the representative bacteria among the cheese samples (Table 1.1).

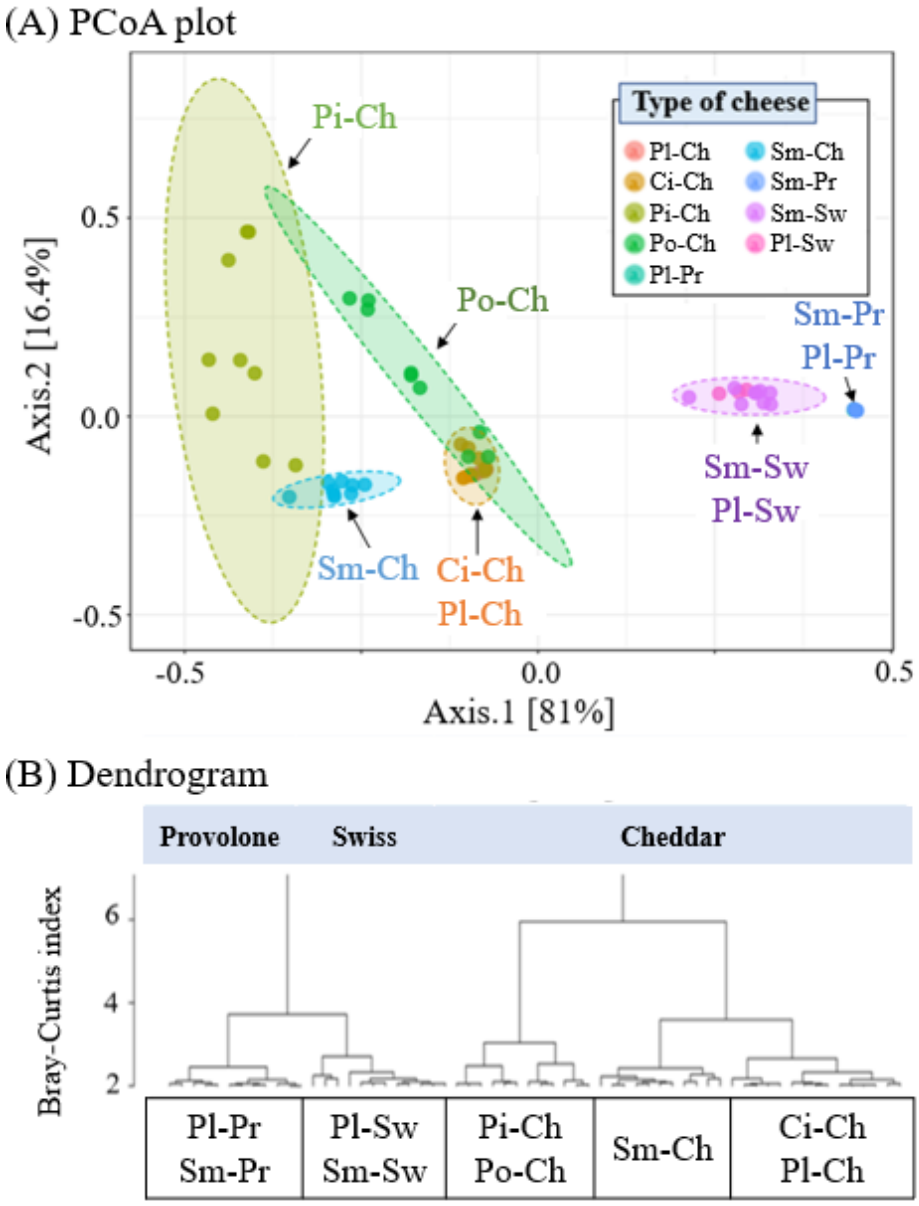


Figure 1.4. (A) Bray-Curtis distance principal coordinate analysis (PCoA) plot shows groupings of similar cheese varieties (Ch = Cheddar; Pl = plain; Ci = cider-soaked; Pi = pinot noir-soaked; Po = porter-soaked; Sm = smoked; Pr = provolone; and Sw = Swiss). All replicates fell within a range of one another, with the Cheddar cheese varieties having the largest amount of variance and (B) dendrogram of represented taxonomic relationships.

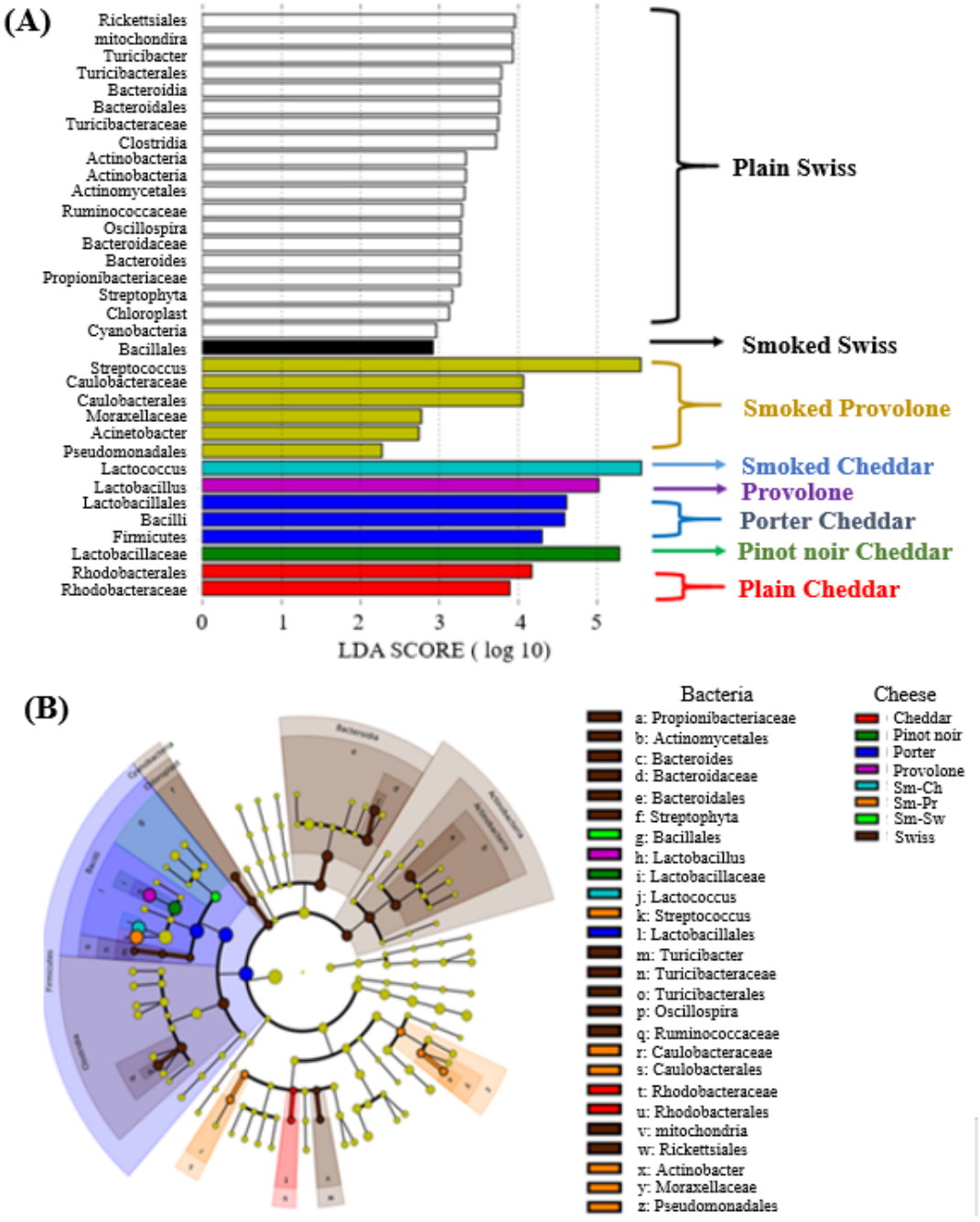


Figure 1.5. Taxonomic differences of cheese microbiota between 9 different types of cheese. (A) The taxa exhibit linear discriminant analysis (LDA) significant differences ($P < 0.01$) among cheese samples (Ch = Cheddar; Pl = plain; Ci = cider-soaked; Pi = pinot noir-soaked; Po = porter-soaked; Sm = smoked; Pr = provolone; and Sw = Swiss), up to genus level, and (B) taxonomic cladogram obtained from linear discriminant analysis effect size analysis of 16S rRNA gene sequencing.

Table 1.1. Relative abundance and potential pathogen of biomarkers identified through LEfSe.

Cheese	Treatment	Strain	Taxonomic level	Relative abundance (%)	Potential pathogen
Cheddar	Plain	*Rhodobacterales*	order	0.003 ± 0.002[e]	-
		Rhodobacteraceae	family	0.003 ± 0.002[e]	-
	Pinot-noir	*Lactobacillaceae*	family	29.717 ± 16.825[d]	-
	Porter	*Firmicutes*	phylum	99.973 ± 0.019[a]	+
		Bacilli	class	99.973 ± 0.029[a]	+
		Lactobacillales	order	99.973 ± 0.029[a]	-
	Smoked	*Lactococcus*	genus	76.124 ± 3.209[c]	-
Provolone	Plain	*Lactobacillus*	genus	20.552 ± 2.797[d]	-
	Smoked	*Streptococcus*	genus	83.158 ± 3.431[b]	-
		Caulobacterales	order	0.007 ± 0.007[e]	+
		Caulobacteraceae	family	0.007 ± 0.007[e]	+
		Pseudomonadales	order	0.006 ± 0.014[e]	+
		Moraxellaceae	family	0.006 ± 0.014[e]	+
		Acinetobacter	genus	0.006 ± 0.014[e]	+
Swiss	Plain	*Bacteroidia*	class	2.078 ± 3.500[e]	+
		Bacteroidales	order	2.078 ± 3.500[e]	+
		Clostridia	class	0.833 ± 1.367[e]	+
		Actinobacteria	phylum	0.382 ± 0.072[e]	-
		Actinobacteria	class	0.382 ± 0.072[e]	-
		Actinomycetales	order	0.382 ± 0.072[e]	-
		Propionibacteriaceae	family	0.380 ± 0.076[e]	+
		Bacteroidaceae	family	0.330 ± 0.538[e]	+
		Bacteroides	genus	0.328 ± 0.540[e]	+
		Ruminococcaceae	family	0.251 ± 0.424[e]	-
		Oscillospira	genus	0.176 ± 0.293[e]	-
		Cyanobacteria	phylum	0.045 ± 0.025[e]	-
		Chloroplast	class	0.045 ± 0.025[e]	-
		Streptophyta	order	0.045 ± 0.025[e]	-
		Rickettsiales	order	0.006 ± 0.005[e]	+
		mitochondria	family	0.006 ± 0.005[e]	-
		Turicibacter	genus	0.006 ± 0.006[e]	-
		Turicibacterales	order	0.006 ± 0.006[e]	-
		Turicibacteraceae	family	0.006 ± 0.006[e]	-
	Smoked	*Bacillales*	order	0.938 ± 0.461[e]	+

[a-f] Means within identical superscripts in each column exhibit no differences at the 95% significance level ($P < 0.05$)

1.4.5 Predictive Metagenomics Profiling (PMP)

The microbiome functioning potential of cheeses was predicted based on predictive metagenomics profiling to compare the different bacterial functions among the 9 cheese samples. Several predictive pathways were significantly enriched in the microbiome data, giving 11 different functions from the Piphillin and PICRUSt2 pipelines. The functions were carbohydrate metabolism, AA metabolism, nucleotide metabolism, metabolism of cofactors and vitamins, energy metabolism, lipid metabolism, metabolism of other AA, glycan biosynthesis and metabolism, metabolism of terpenoids and polyketides, xenobiotics biodegradation metabolism, and biosynthesis of other secondary metabolites (Figure 1.6, Tables 1.2 and 1.3). Additionally, functional property differences between 9 varieties of cheese samples were analyzed and compared through a PCA plot (Figure 1.7). As can be seen in the PCA plot, functional properties were influenced by cheese type. For example, Provolone clustered with smoked Provolone whereas Swiss clustered with smoked Swiss, demonstrating similar functional properties.

Additionally, porter-soaked and pinot noir-soaked cheeses clustered together, whereas cider-soaked and smoked Cheddar cheese each clustered separately. The primary carbohydrate metabolized by these organisms is lactose, which is fermented to produce lactic acid (Figure 1.8). Among carbohydrate metabolism, galactose metabolism (10.06%) and glycolysis/gluconeogenesis (21.82%) were related to lactic-acid-producing mechanisms (Table 1.4).

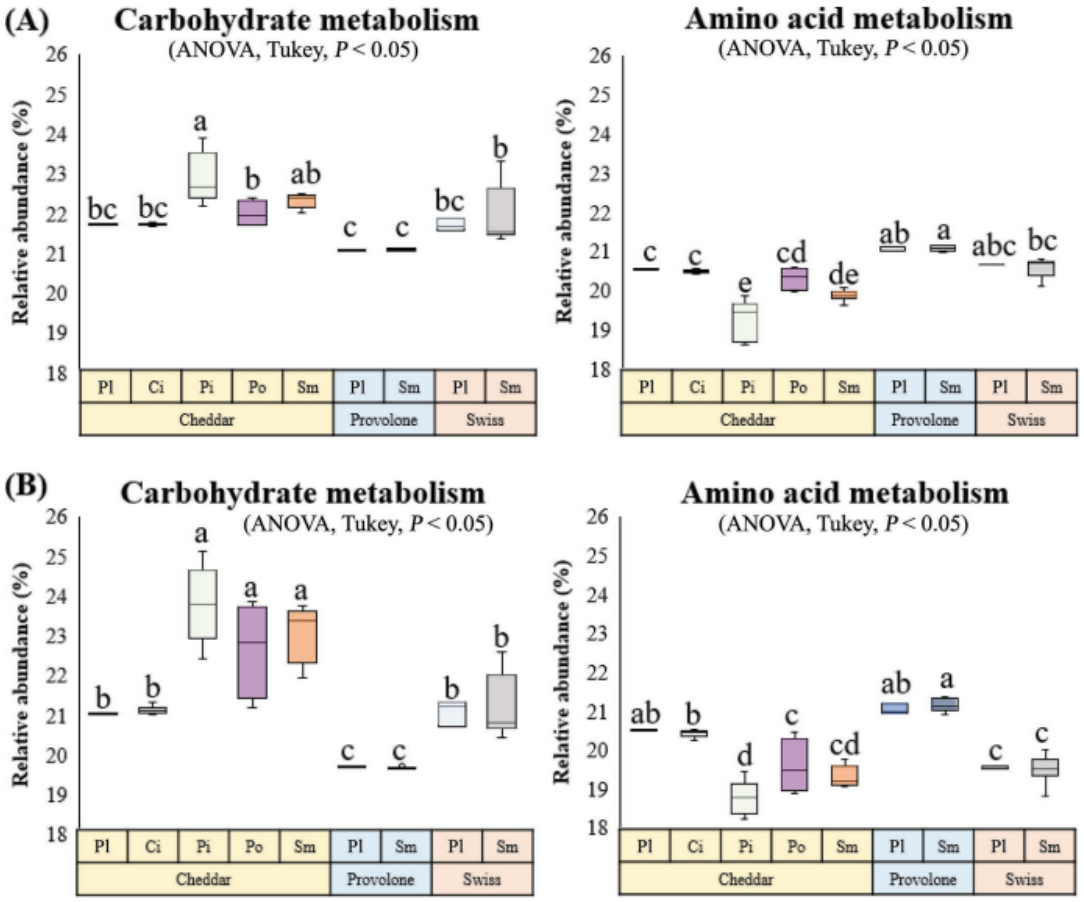


Figure 1.6. Box plot for predicted relative abundance of carbohydrate and AA metabolism annotated genes showing differences in cheese microbiota among 9 different types of cheese (Pl = plain; Ci = cider-soaked; Pi = pinot noir-soaked; Po = porter-soaked; and Sm = smoked) from (A) Piphillin (Iwai et al. 2016) and (B) phylogenetic investigation of communities by reconstruction of unobserved states 2 (PICRUSt2). Boxes in the plots represent the interquartile range (IQR) between first and third quartiles, respectively. The horizontal line in the box indicates the median. The whiskers indicate the lowest and highest value within 1.5 times the IQR from the first and third quartiles, respectively. a–e: Letters within each figure exhibit difference at the 95% significance level ($P < 0.05$).

Table 1.2. Relative abundance (%) of functional properties of bacterial genes in 9 different types of cheese (Pl: plain, Ci: cider-soaked, Pi: pinot noir-soaked, Po: porter-soaked, and Sm: smoked) acquired from Piphillin.

Function	Cheddar					Provolone		Swiss	
	Pl	Ci	Po	Pi	Sm	Pl	Sm	Pl	Sm
Carbohydrate metabolism	21.71 ± 0.01^{bc}	21.72 ± 0.02^{bc}	22.00 ± 0.29^{b}	22.88 ± 0.62^{a}	22.32 ± 0.17^{ab}	21.07 ± 0.02^{c}	21.08 ± 0.03^{c}	21.70 ± 0.16^{bc}	21.94 ± 0.74^{b}
Amino acid metabolism	20.53 ± 0.02^{c}	20.48 ± 0.05^{c}	20.29 ± 0.37^{cd}	19.25 ± 0.48^{e}	19.86 ± 0.13^{de}	21.02 ± 0.06^{ab}	21.06 ± 0.06^{a}	20.65 ± 0.01^{abc}	20.56 ± 0.24^{bc}
Nucleotide metabolism	13.85 ± 0.01^{cde}	13.81 ± 0.02^{cde}	13.67 ± 0.23^{de}	13.00 ± 0.36^{f}	13.49 ± 0.07^{ef}	14.58 ± 0.04^{a}	14.55 ± 0.05^{a}	14.23 ± 0.12^{ef}	14.16 ± 0.51^{f}
Metabolism of cofactors and vitamins	11.33 ± 0.00^{a}	13.81 ± 0.05^{a}	11.2 ± 0.10^{abc}	11.26 ± 0.17^{ab}	11.33 ± 0.06^{a}	11.13 ± 0.01^{cde}	11.13 ± 0.04^{cde}	11.10 ± 0.05^{de}	11.03 ± 0.04^{e}
Energy metabolism	8.78 ± 0.01^{cde}	8.83 ± 0.06^{cd}	8.91 ± 0.23^{bc}	9.59 ± 0.31^{a}	9.16 ± 0.08^{b}	8.35 ± 0.02^{f}	8.34 ± 0.02^{f}	8.57 ± 0.04^{def}	8.54 ± 0.07^{ef}
Lipid metabolism	5.73 ± 0.01^{cdef}	5.77 ± 0.05^{cde}	5.92 ± 0.19^{bcd}	6.40 ± 0.28^{a}	5.98 ± 0.05^{b}	5.47 ± 0.02^{f}	5.46 ± 0.02^{f}	5.58 ± 0.03^{ef}	5.65 ± 0.15^{def}
Metabolism of other amino acids	4.88 ± 0.01^{bc}	4.88 ± 0.02^{bc}	4.94 ± 0.09^{b}	4.85 ± 0.10^{cd}	4.77 ± 0.02^{d}	5.10 ± 0.01^{a}	5.10 ± 0.01^{a}	5.06 ± 0.02^{a}	5.09 ± 0.01^{a}
Glycan biosynthesis and metabolism	4.36 ± 0.01^{cd}	4.35 ± 0.03^{cd}	4.24 ± 0.15^{e}	3.84 ± 0.19^{f}	4.13 ± 0.05^{de}	4.66 ± 0.01^{b}	4.66 ± 0.01^{b}	4.55 ± 0.07^{bc}	4.47 ± 0.15^{abc}
Metabolism of terpenoids and polyketides	3.78 ± 0.01^{c}	3.80 ± 0.02^{c}	3.80 ± 0.14^{c}	4.15 ± 0.08^{b}	3.99 ± 0.04^{a}	3.52 ± 0.01^{de}	3.51 ± 0.01^{e}	3.57 ± 0.02^{de}	3.61 ± 0.06^{d}
Xenobiotics biodegradation and metabolism	3.10 ± 0.00^{ab}	3.09 ± 0.03^{ab}	3.03 ± 0.07^{abc}	2.97 ± 0.17^{bc}	3.14 ± 0.02^{a}	2.99 ± 000^{abc}	2.99 ± 0.00^{abc}	2.95 ± 0.02^{bc}	2.93 ± 0.09^{c}
Biosynthesis of other secondary metabolites	1.95 ± 0.00^{c}	1.95 ± 0.01^{c}	1.95 ± 0.07^{c}	1.83 ± 0.02^{e}	1.85 ± 0.02^{d}	2.10 ± 0.00^{a}	2.11 ± 0.00^{a}	2.04 ± 0.02^{b}	2.04 ± 0.01^{b}

[a-f] Means within identical superscripts in each row exhibit no differences at the 95% significance level ($P < 0.05$)

Table 1.3. Relative abundance (%) of functional properties of bacterial genes in 9 different types of cheese (Pl: plain, Ci: cider-soaked, Pi: pinot noir-soaked, Po: porter-soaked, and Sm: smoked) acquired from PICRUSt2.

Function	Cheddar					Provolone		Swiss	
	Pl	Ci	Po	Pi	Sm	Pl	Sm	Pl	Sm
Carbohydrate metabolism	21.03 ± 0.02^{b}	21.11 ± 0.10^{b}	22.62 ± 0.93^{a}	23.76 ± 0.93^{a}	23.04 ± 0.68^{a}	19.68 ± 0.02^{c}	19.66 ± 0.02^{c}	21.08 ± 0.32^{b}	21.21 ± 0.80^{b}
Amino acid metabolism	20.49 ± 0.03^{ab}	20.41 ± 0.09^{b}	19.57 ± 0.62^{c}	18.78 ± 0.42^{d}	19.30 ± 0.68^{cd}	21.04 ± 0.14^{ab}	21.14 ± 0.16^{a}	19.55 ± 0.04^{c}	19.50 ± 0.34^{c}
Nucleotide metabolism	13.36 ± 0.02^{de}	13.29 ± 0.10^{c}	14.12 ± 0.37^{cd}	14.26 ± 0.27^{bc}	12.63 ± 0.21^{e}	15.36 ± 0.08^{a}	15.28 ± 0.09^{a}	15.01 ± 0.42^{ab}	15.17 ± 0.89^{a}
Metabolism of cofactors and vitamins	11.55 ± 0.01^{a}	11.57 ±0.03^{a}	10.02 ± 0.60^{bc}	10.02 ± 0.63^{c}	11.20 ± 0.23^{ab}	10.50 ± 0.08^{bc}	10.56 ± 0.09^{bc}	10.45 ± 0.17^{bc}	10.31 ± 0.45^{c}
Energy metabolism	9.27 ± 0.01^{d}	9.35 ± 0.12^{cd}	9.64 ± 0.13^{cd}	9.64 ± 0.09^{abc}	9.76 ± 0.12^{ab}	9.60 ± 0.14^{bcd}	9.50 ± 0.17^{bcd}	9.93 ± 0.11^{a}	9.84 ± 0.34^{ab}
Lipid metabolism	5.69 ± 0.01^{d}	5.72 ± 0.03^{d}	6.25 ± 0.19^{bc}	6.25 ± 0.08^{a}	6.03 ± 0.05^{b}	5.32 ± 0.05^{e}	5.29 ± 0.06^{e}	5.74 ± 0.05^{d}	5.83 ± 0.11^{cd}
Metabolism of other amino acids	4.48 ± 0.00^{a}	4.45 ± 0.05^{a}	3.25 ± 0.49^{bc}	3.25 ± 0.43^{c}	4.34 ± 0.06^{a}	4.08 ± 0.05^{ab}	4.11 ± 0.05^{ab}	3.99 ± 0.18^{ab}	3.82 ± 0.11^{b}
Glycan biosynthesis and metabolism	4.81 ± 0.01^{b}	4.72 ± 0.04^{b}	5.18 ± 0.16^{a}	5.18 ± 0.13^{a}	4.80 ± 0.02^{b}	5.06 ± 0.03^{a}	5.04 ± 0.03^{a}	5.05 ± 0.07^{a}	5.12 ± 0.03^{a}
Metabolism of terpenoids and polyketides	3.87 ± 0.001^{a}	3.87 ± 0.01^{a}	3.36 ± 0.22^{bc}	3.36 ± 0.19^{c}	3.82 ± 0.04^{a}	3.76 ± 0.00^{ab}	3.76 ± 0.00^{ab}	3.79 ± 0.06^{a}	3.79 ± 0.21^{ab}
Xenobiotics biodegradation and metabolism	3.41 ± 0.00^{a}	3.37 ± 0.07^{a}	3.32 ± 0.05^{a}	3.32 ± 0.06^{ab}	3.20 ± 0.07^{bc}	3.38 ± 0.03^{a}	3.40 ± 0.03^{a}	3.13 ± 0.06^{c}	3.18 ± 0.11^{c}
Biosynthesis of other secondary metabolites	2.03 ± 0.01^{b}	2.19 ± 0.02^{bc}	2.19 ± 0.08^{a}	2.19 ± 0.12^{a}	1.89 ± 0.04^{c}	2.23 ± 0.03^{a}	2.25 ± 0.04^{a}	2.27 ± 0.03^{a}	2.25 ± 0.09^{a}

$^{a\text{-}e}$ Means within identical superscripts in each row exhibit no differences at the 95% significance level ($P < 0.05$)

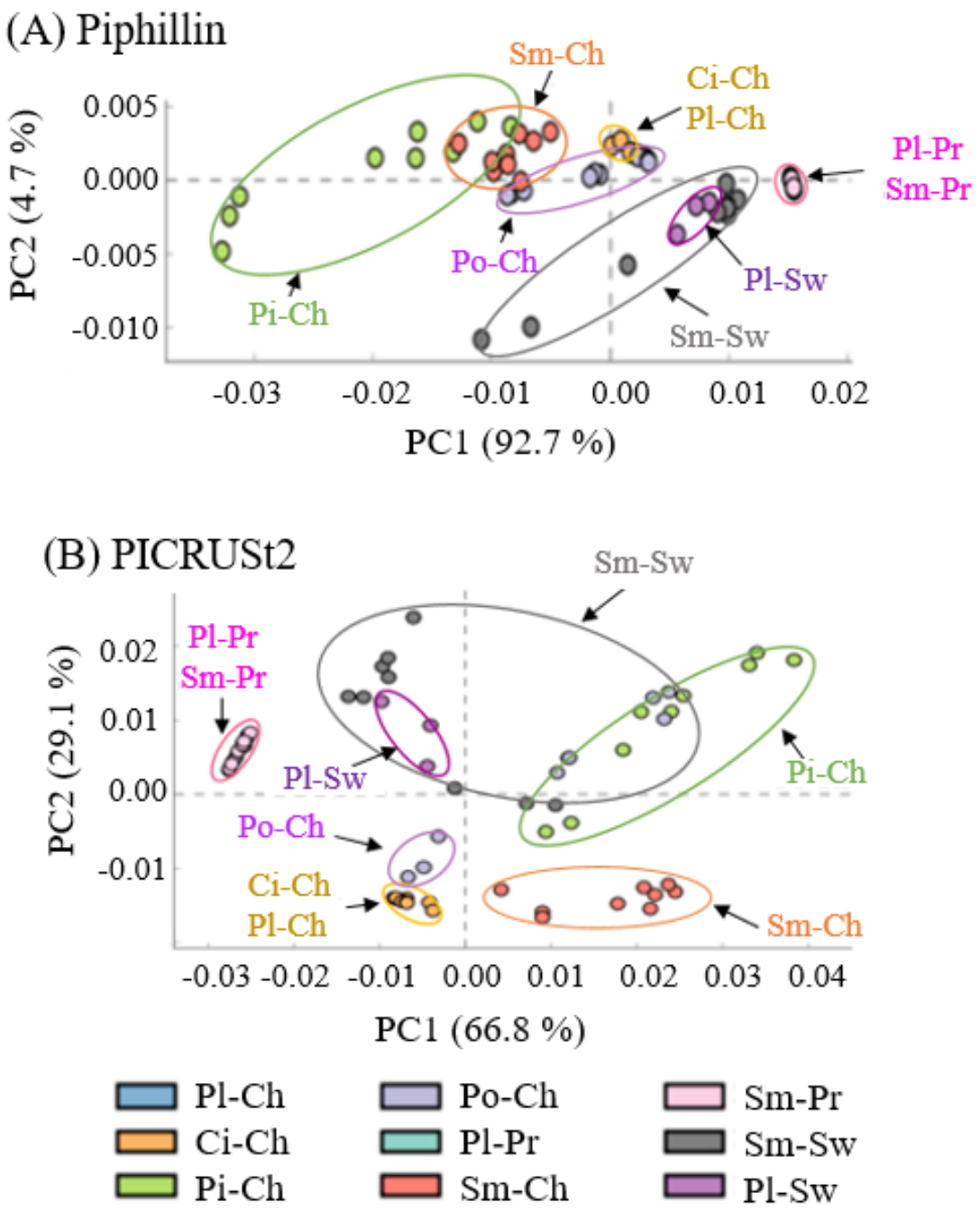


Figure 1.7. PCA plot for functional annotated genes differences of cheese microbiota between 9 different types of cheese (Ch: Cheddar, Pl: plain cheese, Ci: cider-soaked cheese, Pi: pinot noir-soaked cheese, Po: porter-soaked cheese, Sm: smoked cheese, Pr: Provolone, and Sw: Swiss) from (A) Piphillin and (B) PICRUSt2.

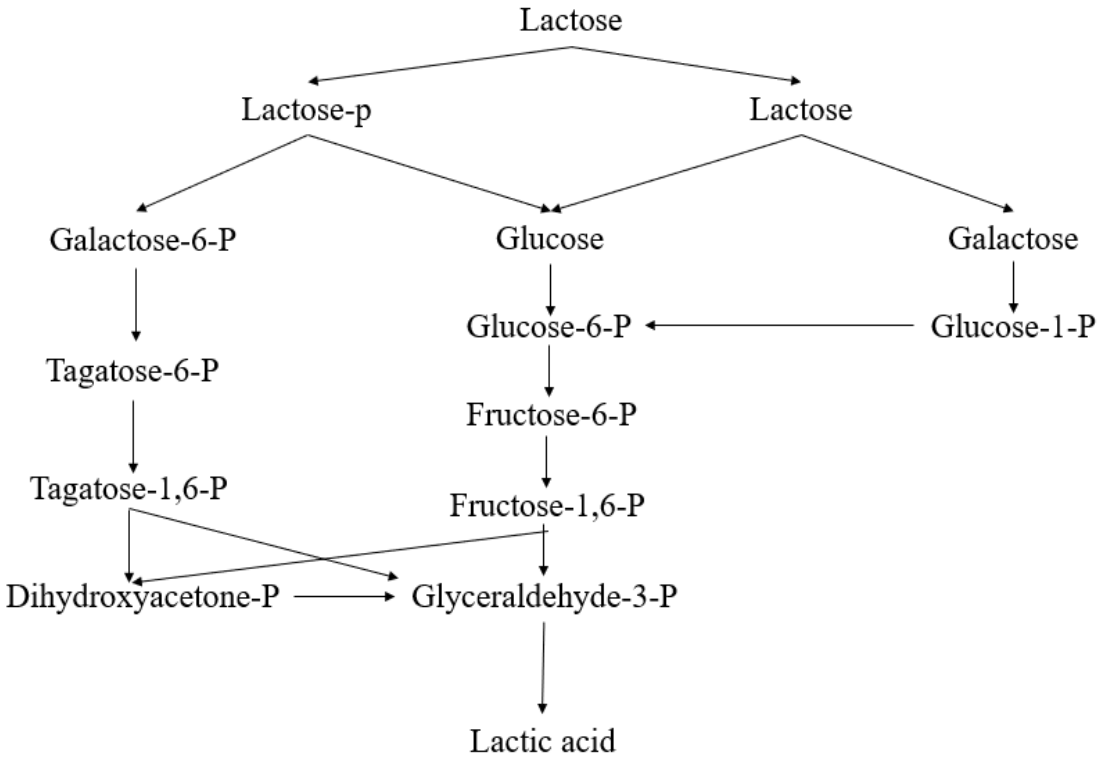


Figure 1.8. Lactic acid production pathway from cheese samples through LAB.

Table 1.4. Relative abundance (%) of lactic acid fermentation related bacterial genes in 9 different types of cheese (Pl: plain, Ci: cider-soaked, Pi: pinot noir-soaked, Po: porter-soaked, and Sm: smoked).

Function	Cheddar					Provolone		Swiss	
	Pl	Ci	Po	Pi	Sm	Pl	Sm	Pl	Sm
Galactose metabolism	8.82 ± 0.00 [ab]	12.20 ± 3.38 [ab]	7.06 ± 2.17 [ab]	5.68 ± 1.41 [a]	9.70 ± 6.96 [ab]	10.23 ± 2.07 [ab]	15.62 ± 6.04 [b]	9.82 ± 2.54 [ab]	10.51 ± 5.32 [ab]
Glycolysis/Gluconeogenesis	19.71 ± 0.00 [ab]	27.11 ± 7.64 [ab]	15.04 ± 5.15 [a]	11.08 ± 3.72 [a]	20.55 ± 14.50 [ab]	23.38 ± 4.74 [ab]	35.70 ± 13.79 [b]	21.09 ± 5.14 [ab]	21.89 ± 9.80 [ab]
Total	28.53 ± 0.0	39.31 ± 11.02	22.10 ± 7.32	16.76 ± 5.10	30.26 ± 21.46	33.61 ± 6.81	51.32 ± 19.83	30.91 ± 7.68	32.40 ± 15.08

[a-e] Means within identical superscripts in each row exhibit no differences at the 95% significance level ($P < 0.05$)

1.5 Discussion

The cheese microbiome plays a key role in determining the organoleptic and physicochemical properties of cheese, affecting both its quality and safety (Yeluri Jonnala et al. 2018). The development of next-generation sequencing technologies has allowed the characterization of microbial communities in cheeses collected around the world to become an active area of research (Dugat-Bony et al. 2016). Wolfe et al. (2014) sequenced 137 different cheese rind communities from 10 countries to identify the dominant bacterial community members. De Filippis et al. (2014) delineated the microbial community properties of 3 popular Italian cheeses, mozzarella, Grana Padano, and Parmigiano Reggiano. In the present study, we analyzed the microbial populations of 9 different cheeses within 3 varieties, which were acquired from the Arbuthnot Dairy Center at Oregon State University (Corvallis). The high-throughput sequencing approach was used to generate a list of 225 I, which were processed using the Greengenes database.

Of these 225 I, only 9 were found to represent 1% or more of the overall microbial community. The dominant I present in cheeses were identified as *Streptococcus* spp., *Lactococcus* spp., an unspecified group of the family *Lactobacillaceae*, and *Lactobacillus* spp. The 2 most abundant I identified were *Streptococcus* spp. and *Lactococcus* spp., both of which are commonly used as cheese starter cultures. Cheeses made with combined single strains of *Lc. Lactis* and *Lc. Cremoris* as starters develop desirable flavors of fermented dairy products. In this study, *Lactococcus* was used as a

starter lactic acid bacteria for all the cheese samples (Schleifer et al. 1985). The species *S. thermophilus* is widely used for the preparation of several dairy products, such as fermented milks, yogurts, and cheeses (Mora et al. 2002; Dugat-Bony et al. 2016). Lactic acid bacteria (LAB) are a heterogeneous group of microorganisms that convert carbohydrates into lactic acid. They contain both pathogenic and beneficial organisms, including *S. thermophilus* and *Lc. Lactis*, which are used in milk fermentation (Bolotin et al. 2001). There are 2 subspecies of *Lc. Lactis*: *Lc. Lactis* ssp. *Lactis* and *Lc. Lactis* ssp. *Cremoris*, which were originally classified as *S. lactis* and *S. cremoris* (Schleifer et al. 1985).

Of the I present at over 1% of the total microbial composition, some were specific to certain cheeses in a genus level. For example, *Lactobacillus* spp. were found in all samples of Swiss cheese. According to Takano (2002), *Lb. helveticus*, which is used as starter bacteria for producing Swiss cheese, is responsible for the proteolytic generation of antihypertensive peptides during the fermentation of milk. *Lactobacillus* spp. were found throughout all Swiss and Provolone cheese samples, ranging from a low of 1 to 3% on the rind region of smoked Swiss cheese to a high of 24% in the core region of the smoked Swiss cheese. To make the Cheddar cheeses for this study, *Lactococcus* spp. and *Streptococcus* spp. were used as starter cultures. *Lactobacillaceae* were present in high abundance (46 and 65%) on the rind regions of the porter-and pinot noir-soaked Cheddar cheeses; however, they were found at low abundance in plain (0.3%) and cider-soaked (5%) Cheddar cheeses. *Lactobacillaceae* are common nonstarter lactic acid bacteria (NSLAB). The origin of NSLAB is debated, but every cheese sample has NSLAB. Although raw milk is a significant environmental contamination or contamination during

the cheese making process or storage (Gobbetti et al. 2015). *Lactobacillaceae* grow well in carbohydrate-containing substrates, such as dairy products, grain products, beer, and wine (Felis and Pot, 2014). Because the cheeses produced for this study were made using pasteurized milk, the native microbiota of raw milk could not be a major source of NSLAB. Though some organisms may survive during pasteurization, the production environment is likely the primary source of *Lactobacillaceae* in the untreated Cheddar cheese. It may be present only in low levels in the cheeses as there are low levels in the production environment or competition from starter cultures. Higher levels of *Lactobacillaceae* in the porter and pinot noir-soaked Cheddar cheeses were expected, as it has been noted that these organisms can grow in beer and wine. The soaking treatments likely inoculate the cheese surface with these organisms. Following this logic, one would expect to see higher levels of *Lactobacillaceae* in the rind of the cider-soaked samples as well; however, this was not the case. This may be due to the addition of a large quantity of salt to the cider to combat pH problems. Differences were found in the microbial communities present in the rind versus the core regions, with the rinds showing a higher level of diversity than the core regions. The rind microbiome further varied based on the type of rind, degree of aging, and environmental conditions (Yeluri Jonnala et al. 2018). While the cheese samples from Wolfe et al. (2014) made a distinctive rind microbiota through exposing cheese surface in the air during aging, beverage-soaked cheese samples in this study did not exhibit noticeable rind formation. Beverage-soaked cheese samples have exposed to the air for 3 h and ripened with a vacuum-sealed packaging for 3 to 6 mo at 4°C. This might inhibit the development of proliferative rind microbiota compared with the previous Wolfe et al. (2014) study. However, we found that beverage-soaked

treatment on cheese surface can lead more profound microbial variations between the core and rind regions, which matched with the result from the previous study (Wolfe et al. 2014). For example, significant differences in microbial richness between the core and rind were seen in smoked cheeses. According to De Filippis et al. (2016) and Wolfe et al. (2014), the lower water activity and higher oxygen concentration can alter the microbial communities between the core and rind regions. Soaking into the different beverages (cider, porter, and pinot noir) can change the microbial composition and diversity of cheese. The relative abundance of unidentified *Lactobacillaceae* was increased after soaking in the beverage, and all results indicated the relative abundance of *Lactobacillaceae* from rind regions was higher than the mixed and core regions. In the previous study, Cousin et al. (2017) analyzed the microbial property of hard apple cider that main bacterial community was identified to the *Lactobacillaceae* and *Acetobacteraceae* family. It is related to our result that the increase of relative abundance of *Lactobacillaceae* in the rind region of cider-soaked Cheddar cheese. Porter beer is a red-brown acidic beer that commonly uses yeast and LAB as starter cultures. *Pediococcus damnosus* and *Lactobacillus* belong to *Lactobacillaceae* family were identified to the most common bacterial group found in the maturation phase of porter beer (Bokulich and Bamforth, 2017). The high relative abundance of *Lactobacillaceae* in the rind region of porter-soaked Cheddar cheese corresponded to the previous study.

Malolactic fermentation converts malic acids to lactic acids, led by *Ornococcus oeni* and other LAB. These are important bacterial strains to produce red wine flavor, and LAB occupied a high portion of microbial populations in red wine (Bokulich et al. 2016). The increase of *Lactobacillaceae* in pinot noir-soaked Cheddar cheese in this study

originated from pinot noir wine. Based on the previous studies, the changes of microbial populations in the rind and core regions of cheese might be directly influenced by soaking materials. According to the previous studies, the microbial populations of cheese were significantly determined by cheese making and ripening processes (Porcellato and Skeie, 2016; De Filippis et al. 2016). Porcellato and Skeie (2016) found that scalding temperature (37 and 39°C) during ripening up to 3 mo is the most important factor for the establishment of Dutch-type cheese microbiome. A high number of *Lactobacillus* spp. was detected during ripening at high scalding temperature (39°C). In this study, because all cheeses were aged under same conditions up to 3 to 6 mo, the aging condition was not considered the factor that might affect the microbiome of different types of cheese. Based on the Shannon index, no significant differences ($P > 0.05$) were observed between smoked and non-smoked Cheddar, Provolone, and Swiss cheeses. This indicates that smoking did not alter the microbial diversity of cheese because they were not exposed to other environmental microbes. However, the Shannon index of pinot noir- (1.39), and porter-soaked (1.37) Cheddar cheeses were significantly increased ($P < 0.05$) compared with the plain Cheddar cheese (0.70), indicating that soaking in pinot noir or porter increased the microbial diversity. Additionally, the Shannon index of the rind regions was higher than the core regions across all samples. The richness increased in the rind region may be partly due to the availability of oxygen, which allows the growth of aerobic bacteria that cannot survive in the core (Donnelly, 2014). Interaction with environmental sources is another factor that may increase microbial diversity in the rind region. The rind is contact with the surface of cheese throughout production and aging and can readily be contaminated with microbes from the environment. On the other hand, the core region

can be preserved and is composed of organisms present at the time of shaping. Analysis of the beta diversity indicated that the various groups of cheese were strongly related to one another in microbial composition. The same group of cheese samples was clustered on Bray-Curtis principal coordinate analysis plots (Figure 1.4). Both pinot noir and porter-soaked Cheddar cheeses showed a noticeable difference between the core and rind regions compared with other groups of cheese. The porter- and pinot noir-soaking treatments changed the beta diversity in the Cheddar cheeses. Moreover, the dendrogram indicated that each group of cheese exhibited high similarity in microbial composition. Provolone and Swiss cheeses were especially similar, whereas the Cheddar cheese samples did not cluster with the other 2 varieties of cheese. According to previous studies (Wolfe et al. 2014; Dugat-Bony et al. 2016), the microbial diversity of cheese was significantly influenced by processing, type of cheese, and moisture content.

Further analysis of the cheese microbiome with LEfSe (Figure 1.5) found significant differences in bacterial abundance among different types of cheese. Biomarker or biological markers imply a measurable indicator of a certain biological state or condition. The LEfSe is a tool used to find biomarkers between 2 or more groups using relative abundance. To identify biomarkers in cheese samples, LEfSe was used to compare the relative abundance of bacteria from each cheese samples and find bacterial strains that were specific and in high abundance in each cheese. The biomarker is widely used in the clinical field as an indicator to diagnose a target sample (Segata et al. 2011). Biomarkers present at more than 1% in the cheese samples included *Streptococcus* spp., *Lactococcus* spp., *Lactobacillaceae*, and *Lactobacillus* spp., which were identified as biomarkers in the smoked Provolone, smoked Cheddar, pinot noir-soaked Cheddar, and

Provolone cheeses, respectively. Although *Streptococcus* spp. were used as starter cultures for all the 3 types of cheese (Cheddar, Provolone, and Swiss), the relative abundance of *Streptococcus* spp. was the highest in Provolone. In Provolone cheese, the smoking treatment decreased the relative abundance of *Lactobacillus* spp., whereas *Streptococcus* spp. increased. Similarly, *Lactococcus* spp. and *Streptococcus* spp., the predominant organisms in Cheddar cheeses, were affected by smoking. Smoking decreased the relative abundance of *Streptococcus* spp. and increased the relative abundance of *Lactococcus* spp. In the previous study (Majcher et al. 2011), the number of lactobacilli, lactococci, streptococci, and enterococci were decreased during the smoking process. Warm smoking (25 to 35°C) leads to an increase in phenolic compounds formed during the smoking process that are known to have bactericidal properties. Thus, this treatment may have affected the microbial diversity. However, the decreasing ratio of each bacteria was different depending on the strain, which led to changes in the relative abundance of bacterial composition. Procedures such as soaking or smoking played an important role in shifting the microbial composition as well as changing biomarkers.

Amplicon-based microbiome sequencing using the 16S rRNA gene is a powerful tool to assess and compare microbial community structure and diversity in a certain ecosystem. Although the 16S rRNA gene amplicon sequencing is widely used to characterize the microbial taxonomic composition and phylogenetic diversity (Ashauer et al. 2015), it is difficult to provide direct evidence of functional capabilities of microbiota (Iwai et al. 2016). The rapid growth in the number of sequenced genomes makes it possible to infer which functions are associated with a marker gene based on its sequence similarity when comparing it to a reference genome. In the PFP of the cheese

microbiome, the highest number of sequencing reads was assigned to carbohydrate metabolism and AA metabolism from both the PICRUSt2 and Piphillin analyses. It is not surprising that carbohydrate metabolism was the most abundant annotated functional property of the microbes present in cheese samples, as fermentation of carbohydrates is the key functional activity of cheese starter cultures. The primary carbohydrate metabolized by these organisms is lactose, which is fermented into lactic acid and other metabolites (Porcellato and Skeie, 2016; Mataragas et al. 2018; Bautista-Gallego et al. 2019). Lactic acid fermentation in dairy products is a metabolic process that uses lactose to produce lactate in a lactic acid solution. It is an anaerobic fermentation reaction that occurs in a fermented food such as cheese and yogurt. Lactose [β-d-galactopyranosyl-(1→4)-dglucose] is the primary sugar composed of glucose and galactose naturally found in milk and dairy products. Lactic acid bacteria metabolize lactose, glucose, and galactose to produce lactic acids through galactose and glycolysis/gluconeogenesis metabolisms (Figure 1.8). Among carbohydrate metabolites, functional genes related to lactic acid production accounted for more than 30% of the metabolomes. Another important role of the cheese microbiota is proteolysis and AA metabolism, which are important for texture and flavor development during cheese ripening (Ardo, 2006). In Figure 1.7, the differences in functional properties between different types of cheese were displayed in a PCA plot. Although the PCA plots from Piphillin and PICRUSt2 did not match exactly, both samples were clustered by the type of cheese. In a previous study, Wolfe et al. (2014) analyzed taxonomic diversity and functional properties of 137 different cheese rinds using a shotgun sequencing method. According to their metagenomic results, both taxonomic diversity and functional properties were related to

each other and were strongly affected by the type of cheese. Functional potentials clustered by rind type, moisture content, and cheese making procedures (Wolfe et al. 2014; Dugat-Bony et al. 2016). Our findings correspond with previous reports that microbial diversity and functional property were significantly affected by the type of cheese.

1.6 Conclusions

The present study improves the understanding of composition, diversity, and functional properties of microbiota from different types of cheese through 16S rRNA gene-based microbiome sequencing using the Illumina MiSeq platform. We compared the microbial community differences among 9 different types of cheese acquired from the Arbuthnot Dairy Center at Oregon State University (Corvallis, OR). In this study, we found that different types of cheese exhibited significant changes in microbial community structure, biomarkers, microbial diversity, and PFP, depending on the moisture content, rind formation, and color of rinds. This study provides better insight into the microbial properties of different cheese types, as well as locations within the cheese, to help manage the quality of cheese. In addition, future study is needed to investigate the potential influence of chemical compositional variations between the rind and the core regions of cheese on the microbiome.

1.7 References

Aßhauer KP, Wemheuer B, Daniel R, Meinicke P (2015) Tax4Fun: predicting functional profiles from metagenomic 16S rRNA data. Bioinformatics. 31: 2882-2884.

Ardö Y (2006) Flavour formation by amino acid catabolism. Biotechnology advances 24: 238-242

Bautista-Gallego J, Ferrocino I, Botta C, Ercolini D, Cocolin L, Rantsiou K (2019) Probiotic potential of a *Lactobacillus rhamnosus* cheese isolate and its effect on the fecal microbiota of healthy volunteers. Food research international 119: 305-314

Bamforth CW, Nicholas AB (2017) Brewing Microbiology: Current Research, Omics and Microbial Ecology. Caister Academic Press, Wymondham, UK

Bolotin A, Wincker P, Mauger S, Jaillon O, Malarme K, Weissenbach J, Ehrlich SD, Sorokin A (2001) The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ssp. *Lactis* IL1403. Genome research 11: 731-753

Bokulich NA, Mills DA (2013) Facility-specific "house" microbiome drives microbial landscapes of artisan cheesemaking plants. Appl Environ Microbiol 79: 5214-5223

Bokulich NA, Collins TS, Masarweh C, Allen G, Heymann H, Ebeler SE, Mills DA (2016) Associations among wine grape microbiome, metabolome, and fermentation behavior suggest microbial contribution to regional wine characteristics. Mbio 7(3): e00631-16

Broadbent JR, Cai H, Larsen R, Hughes J, Welker D, de Carvalho V, Tompkins T, Ardö Y, Vogensen F, de Lorentiis A (2011) Genetic diversity in proteolytic enzymes and amino acid metabolism among *Lactobacillus helveticus* strains. Journal of dairy science 94: 4313-4328

Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP (2016) DADA2: high-resolution sample inference from Illumina amplicon data. Nature methods 13: 581-583

Cousin FJ, Le Guellec R, Schlusselhuber M, Dalmasso M, Laplace JM, Cretenet M (2017) Microorganisms in fermented apple beverages: current knowledge and future directions. Microorganisms 5: 39

De Filippis F, Genovese A, Ferranti P, Gilbert JA, Ercolini D (2016) Metatranscriptomics reveals temperature-driven functional changes in microbiome impacting cheese maturation rate. Scientific reports 6: 21871

Dhariwal A, Chong J, Habib S, King IL, Agellon LB, Xia J (2017) MicrobiomeAnalyst: a web-based tool for comprehensive statistical, visual and meta-analysis of microbiome data. Nucleic acids research 45: W180-W188

Didienne R, Defargues C, Callon C, Meylheuc T, Hulin S, Montel MC (2012) Characteristics of microbial biofilm on wooden vats ('gerles') in PDO Salers cheese. International journal of food microbiology 156: 91-101

Donnelly CW (2014) Cheese and Microbes. ASM Press, Washington, DC.

Douglas GM, Beiko RG, Langille MG (2018) Predicting the functional potential of the microbiome from marker genes using PICRUSt. Methods in molecular biology 1846: 169-177

Dugat-Bony E, Garnier L, Denonfoux J, Ferreira S, Sarthou AS, Bonnarme P, Irlinger F (2016) Highlighting the microbial diversity of 12 French cheese varieties. International journal of food microbiology 238: 265-273

Duru IC, Laine P, Andreevskaya M, Paulin L, Kananen S, Tynkkynen S, Auvinen P, Smolander OP (2018) Metagenomic and metatranscriptomic analysis of the microbial community in Swiss-type Maasdam cheese during ripening. International journal of food microbiology 281: 10-22.

Felis GE, Pot B (2014) Lactic Acid Bacteria: Biodiversity and Taxonomy. Wiley. Hoboken, NJ. 245-247.

Gobbetti M, de Angelis M, di Cagno R, Mancini L, Fox PF (2015) Pros and cons for using non-starter lactic acid bacteria (NSLAB) as secondary/adjunct starters for cheese ripening. Trends in food science and technology 45: 167-178

Iwai S, Weinmaier T, Schmidt BL, Albertson DG, Poloso NJ, Dabbagh K, de Santis TZ (2016) Piphillin: improved prediction of metagenomic content by direct inference from human microbiomes. PLoS ONE 11: e0166104

Jonnala BRY, McSweeney PL, Sheehan JJ, Cotter PD (2018) Sequencing of the Cheese Microbiome and its relevance to Industry. Frontiers in microbiology 9

Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD (2013) Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. Appl Environ Microbiol 01043-01013

Langille MG, Zaneveld J, Caporaso JG, McDonald D, Knights D, Reyes JA, Clemente JC, Burkepile DE, Thurber RLV, Knight R (2013) Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nature Biotechnology 31: 814

Majcher MA, Goderska K, Pikul J, Jeleń HH (2011) Changes in volatile, sensory and microbial profiles during preparation of smoked ewe cheese. Journal of the science of food and agriculture 91(8): 1416-1423

Mataragas M, Alessandria V, Ferrocino I, Rantsiou K, Cocolin L (2018) A bioinformatics pipeline integrating predictive metagenomics profiling for the analysis of 16S rDNA/rRNA sequencing data originated from foods. Food microbiology 76: 279-286

Mayo B, Rachid CTCC, Alegría Á, Leite, A, Peixoto SR, Delgado S (2014) Impact of next generation sequencing techniques in food microbiology. Current Genomics 15: 293-309

McDonald D, Price MN, Goodrich J, Nawrocki EP, de Santis TZ, Probst A, Andersen GL, Knight R, Hugenholtz P (2012) An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. The ISME Journal 6: 610

Montel MC, Buchin S, Mallet A, Delbes-Paus C, Vuitton DA, Desmasures N, Berthier F (2014) Traditional cheeses: rich and diverse microbiota with associated benefits. International journal of food microbiology 177: 136-154

Mora D, Fortina M, Parini C, Ricci G, Gatti M, Giraffa G, Manachini P (2002) Genetic diversity and technological properties of *Streptococcus thermophilus* strains isolated from dairy products. Journal of applied microbiology 93: 278-287

Parks DH, Tyson GW, Hugenholtz P, Beiko RG (2014) STAMP: statistical analysis of taxonomic and functional profiles. Bioinformatics 30: 3123-3124

Percival B, Percival F (2017) Reinventing the Wheel: Milk, Microbes, and the Fight for Real Cheese. Vol. 65. Univ of California Press. Oakland, CA.

Pintado ME, Xavier M (2000) Characterization of whey cheese packaged under vacuum. Journal of food protection 63: 216-221

Porcellato D, Skeie SB (2016) Bacterial dynamics and functional analysis of microbial metagenomes during ripening of Dutch-type cheese. International dairy journal 61: 182-188

Potočki S (2016) Potential health benefits of sphingolipids in milk and dairy products. Mljekarstvo 66: 251-261

Schleifer K, Kraus J, Dvorak C, Kilpper-Bälz R, Collins M, Fischer W (1985) Transfer of *Streptococcus lactis* and related streptococci to the genus *Lactococcus* gen. nov. Systematic and Applied Microbiology 6: 183-195

Schmitz-Esser S, Dzieciol M, Nischler E, Schornsteiner E, Bereuter O, Mann E, Wagner M (2018) Abundance and potential contribution of Gram-negative cheese rind bacteria from Austrian artisanal hard cheeses. International journal of food microbiology 266: 95-103

Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, Huttenhower C (2011) Metagenomic biomarker discovery and explanation. Genome Biology 12: R60

Takano T (2002) Anti-hypertensive activity of fermented dairy products containing biogenic peptides. Antonie van. Leeuwenhoek 82: 333-340

Walther B, Schmid A, Sieber R, Wehrmüller K (2008) Cheese in nutrition and health. Dairy Science and Technology 88: 389-405

Wolfe BE, Button JE, Santarelli M, Dutton RJ (2014) Cheese rind communities provide tractable systems for in situ and in vitro studies of microbial diversity. Cell 158: 422-433

Shah N, Jelen P (1990) Survival of lactic acid bacteria and their lactases under acidic conditions. Journal of food science 55: 506-509

Solieri L, Dakal TC, Giudici P (2013) Next-generation sequencing and its potential impact on food microbial genomics. Annals of Microbiology 63: 21-37

Chapter 2. Assessment of overall microbial community shift during Cheddar cheese production from raw milk to aging

Jungmin Choi[1], Sang In Lee[1], Bryna Rackerby[1], Robin Frojen[1], Lisbeth Goddik[1], Sang-Do Ha[2] and Si Hong Park[1]*

[1]Department of Food Science and Technology, Oregon State University, Corvallis, OR, United States.

[2]School of Food Science and Technology, Advanced Food Safety Research Group, Brain Korea 21 Plus, Chung-Ang University, Anseong, Gyeonggi-Do, South Korea.

Applied Microbiology and Biotechnology

https://doi.org/10.1007/s00253-020-10651-7

2.1 Abstract

Cheese is a fermented dairy product that is made from animal milk and is considered to be a healthy food due to its available nutrients and potential probiotic characteristics. Since the microbes in the cheese matrix directly contribute to the quality and physicochemical properties of cheese, it is important to understand the microbial properties of cheese. In this study, Cheddar cheeses produced on three different dates at the Arbuthnot Dairy Center at Oregon State University were collected to determine the microbial community structure. A total of 773,821 sequencing reads and 271 amplicon sequence variants (ASVs) were acquired from 108 samples. *Streptococcus* and *Lactococcus* were observed as the most abundant ASVs in the cheese, which were used as the starter lactic acid bacteria (SLAB). *Escherichia coli* was detected in the raw milk; however, it was not detected after inoculating with SLAB. According to an alpha diversity analysis, SLAB inoculation decreased the microbial richness by inhibiting the growth of other bacteria present in the milk. A beta diversity analysis showed that microbial communities before the addition of SLAB clustered together, as did the samples from cheese making and aging. Non-starter lactic acid bacteria (NSLAB) were detected 15 weeks into aging for the June 6th and June 26th produced cheeses, and 17 weeks into aging for the cheese produced on April 26th. These NSLAB were identified as an unidentified group of *Lactobacillaceae*. This study characterizes the changes in the Cheddar cheese microbiome over the course of production from raw milk to a 6-month-aged final product.

Keywords: Cheese, microbiota, high-throughput sequencing, starter bacteria, non-starter bacteria, aging

2.2 Introduction

According to the Food and Agriculture Organization (FAO), the global cheese market value could increase to 124 billion dollars by 2022 (Gosalvitr et al. 2019). In 2018, the USA and European Union (EU), the major cheese producers and consumers, made 5,878 and 10,160 tons of cheese, respectively (USDA 2018). Cheddar cheese is one of the most widely produced and consumed cheeses in the USA and the United Kingdom (UK), with its production in the UK increasing by 30% in the last five years (Gosalvitr et al. 2019).

Cheese is a biologically and biochemically dynamic matrix in which the microbiota structure and activity are influenced by manufacturing practices (De Filippis et al. 2016). The cheese microbiota not only provides flavor through the production of volatile compounds (Percival and Percival 2017), but also has anti-cancer and cholesterol-lowering properties (Broadbent et al. 2011; Potočki 2016; Walther et al. 2008). The microbial dynamics of cheese are influenced by the interactions among factors such as starter lactic acid bacteria (SLAB), cheese making techniques, storage conditions, and the presence of non-starter lactic acid bacteria (NSLAB) (Beresford et al. 2001; Porcellato and Skeie 2016; Van Hoorde et al. 2010).

Characterization of the cheese microbiota is important in the industry, as certain microorganisms impart beneficial sensory characteristics while others may reduce quality

(Mayo et al. 2014). Development of a next generation sequencing (NGS) technology allows researchers produce a large amount of genomic information quickly at a low cost, which can improve the understanding of the microbial properties within the target matrix. The advent of high throughput sequencing has made it possible to explore the microbiomes, the combined genetic material of the microorganisms present in a particular environment, of various foods and investigate the genomes of individual organisms. The NGS technique can reveal how microbes respond to environmental conditions, allowing cheesemakers to improve their control of microbial growth in products based on predictions of how conditions will impact both beneficial and undesirable microbes (Solieri et al. 2013).

Recently, several studies have been done to characterize the microbial populations and properties of dairy products such as milk and cheese using a NGS technique. Wolfe et al. (2014) evaluated the microbial compositions in 137 different cheeses collected from 10 countries and identified 24 genera of bacteria and fungi are dominant in communities. Another study completed by Dugat-Bony et al. (2015), microbial community changes in cheese over aging were evaluated using metagenomic and metatranscriptomic analyses.

In this study, we examined three different batches of Cheddar cheese produced on different days (April 26th, June 6th, and June 26th in 2018) at the Arbuthnot Dairy Center at Oregon State University (Corvallis, OR, USA). Samples were collected from each batch at 36 stages, from raw milk to 26 weeks after aging. These samples were used to analyze the microbial community shift during cheese production, providing a better understanding of the microbial properties of Cheddar cheese throughout processing.

2.3 Materials and methods

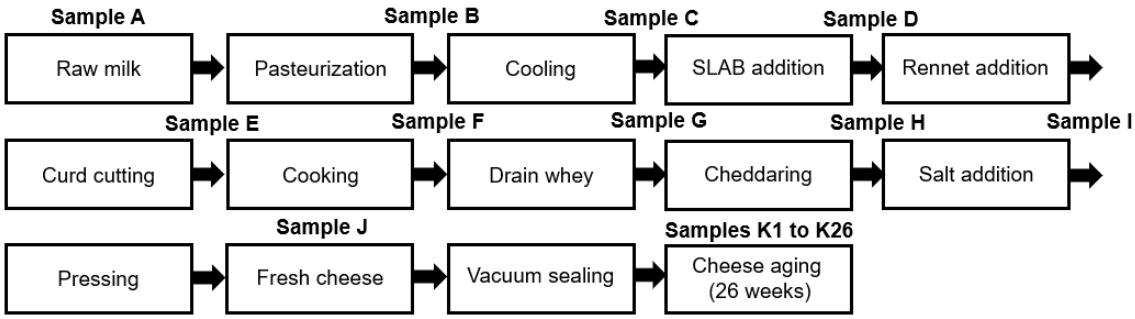
2.3.1 Cheddar cheese making and sampling

Figure 2.1 depicts a flow chart for the production of the three separate batches of Cheddar cheese made at Oregon State University. Raw milk was sampled (Sample A) prior to the milk being transferred to the cheese vat, where the temperature was set at 63 ℃ for 30 min to pasteurize the milk. Pasteurized milk was sampled (Sample B) and allowed to cool down for one hour. After cooling, samples were collected (Sample C) and the SLAB mesophilic commercial starter series CHOOZIT™ RA (mixture of *Lactococcus lactis* subsp. *lactis*, *Lactococcus lactis* subsp. *cremoris*, and *Streptococcus thermophilus*) and MD starter culture (*Lactococcus lactis* subsp. *lactis* biovar diacetylactis; Danisco, Copenhagen, Denmark) were inoculated at 1.6×10^6 and 2.0×10^6 cells/ml into the pasteurized milk, respectively. One hour after inoculating the milk with the SLAB, samples were collected (Sample D) and 0.01% rennet (v/w) was added to coagulate milk for 45 min. The coagulum was then cut into 2.54 × 2.54 cm cube-shaped particles and samples were collected (Sample E). The curds were then stirred while heating from 32 to 37 ℃ for 45 min to cook the curd. Curd samples were also collected (Sample F). After cooking, the whey was separated from the curd until the pH reached 6.3, at which point an additional curd sample was collected (Sample G). The removed whey was discarded. The curd was then cut into proportional loaves with a curd knife and allowed to rest for 5 min. Cheddaring was then carried out until the pH of the curds

reached to 5.3 to 5.4 and a portion of the curds was collected (Sample H). After the addition of salt (3.5% NaCl in total curd weight, w/w), a sample was collected (Sample I) and the curds were placed into molds and pressed for 16 h at room temperature before another sample (Sample J) was taken. The Cheddar cheese was vacuum packed and stored at 4 °C for 6-months of aging. The Cheddar cheese samples were collected each week for 26 weeks while aging (samples K1 to K26). In order to understand the microbial property changes of Cheddar cheese, samples were collected from 3 different stages, from raw milk to pasteurized and cooled milk (Step 1, Samples A to C), from inoculation with SLAB to fresh cheese (Step 2, Samples D to J), and aging cheese up to 26 weeks of aging (Step 3, Samples K1 to K26; Figure 2.1).

2.3.2 Cheese processing plant environmental samples

Environmental sites were identified and swabbed throughout the Arbuthnot Dairy Center at Oregon State University. All sites were sampled with sterile cotton-tipped swabs (Puritan Medical, Guilford, ME, USA). Cotton tips were soaked in sterile phosphate-buffered saline (PBS) and streaked across a 10 × 10 cm (or more) area of the target surfaces (24 locations and utensils, Figure 2.2) in perpendicular directions, ensuring full contact of all parts of swab tip on the surface. Swab tips were stored in 15 ml tubes until further experiments. Swab tips were inoculated into 5 ml of De Man, Rogosa and Sharpe (MRS) broth (2% dextrose, 1% peptic digest of animal tissue, 1% beef extract, 0.5% yeast extract, 0.5% sodium acetate, 0.2% disodium phosphate, 0.2%



Step	Step 1	Step 2	Step 3
Sample	A, B, and C	D, E, F, G, H, I, and J	K1 to K26
Description	Before SLAB addition	Cheese making process	Aging (6 months)

Figure 2.1. Flowchart for manufacturing and sampling stages for Cheddar cheese in the Arbuthnot Dairy Center at Oregon State University.

ammonium citrate, 0.1% Tween® 80, 0.01% magnesium sulfate, and 0.005% manganese sulfate) and incubated anaerobically at 37 ℃ for 72 h.

2.3.3 PCR to confirm lactic acid bacteria (LAB) in the cheese making facility

To identify which LAB were present in the cheese facility, a PCR assay with enrichment broth was applied using LAB specific primers. Five primer pairs of representative LAB for identifying *Lactococcus* spp., *Streptococcus* spp., *Lactobacillus* spp., *Pediococcus* spp. and *Leuconostoc* spp. at the genus level were used. After incubating in MRS broth for 72 h, all cultures with visible growth were selected and had target sequences amplified through the colony PCR assay. PCR was performed in 20 µl total volume, containing 1 µl of grown culture, 10 µl of 2 x GoTaq® Green Master Mix (Promega, Madison, WI, USA), and 500 nM of forward and reverse primers (Table 2.1). PCR conditions varied based on which targeting primer pairs were used (Table 2.2).

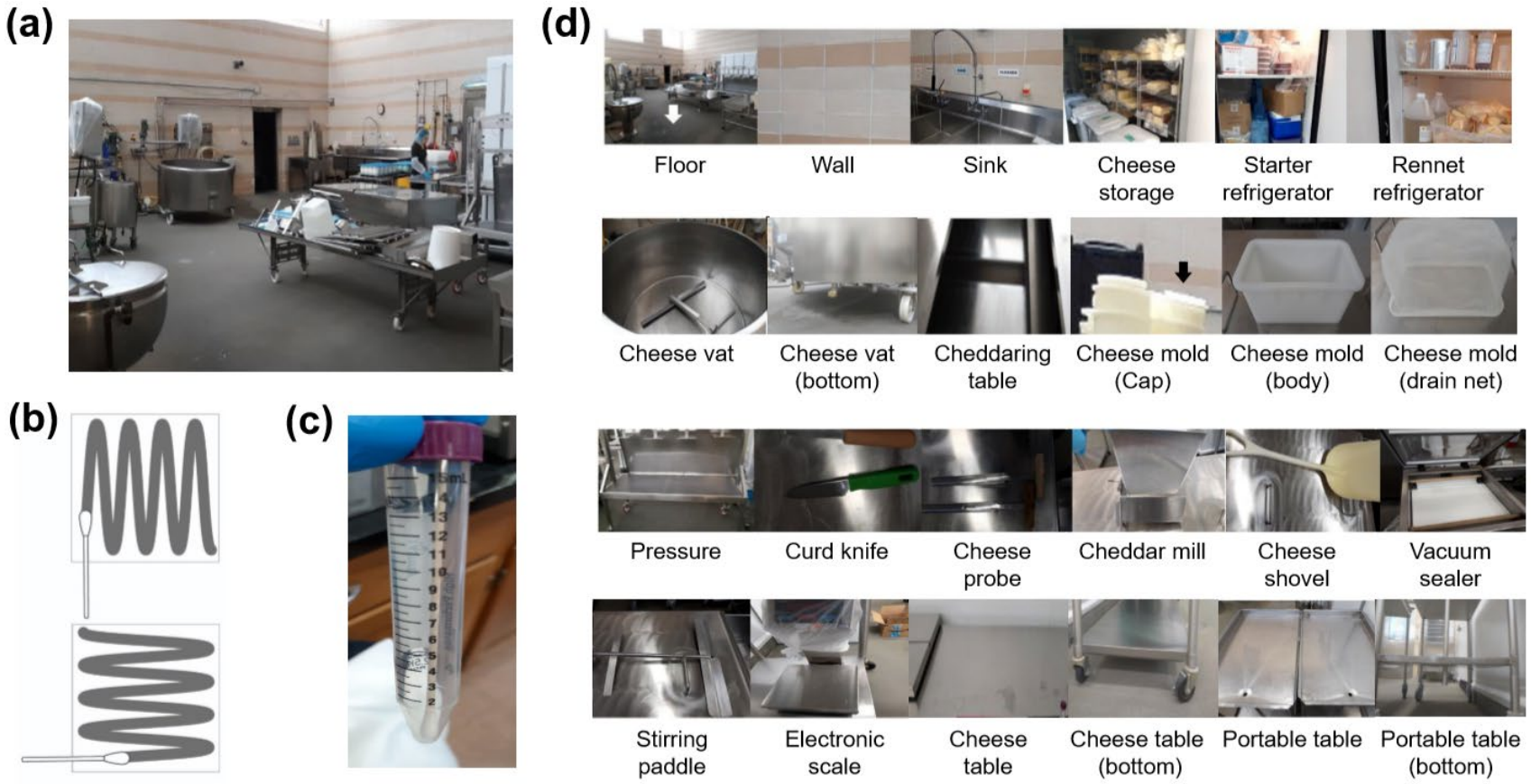


Figure 2.2. Environmental sites swabbed throughout the Arbuthnot Dairy Center at Oregon State University. (a) Overview of cheese making facility, (b) swab plan, (c) storage cotton swab in a tube, and (d) swabbed locations in the cheese facility.

Table 2.1. Nucleotide sequences of primer sets used in this study to identify cultures from environmental swabs.

Strain	Primer	Primer sequence (5'-3')	Size	Reference
Lactobacillus	LbLMA1	CTC AAA ACT AAA CAA AGT TTC	250 bp	(Dubernet et al. 2002)
	R16-1	CTT GTA CAC ACC GCC CGT CA		
Lactococcus	L1	CAG GGC ACG TTG AAA AGT GCT T	570 bp	(Deasy et al. 2000)
	L2	GTG TGA CAT CAC TAA CTT CGC		
Streptococcus	Str-791F	AAC TCT GTT GTT AGA G	804 bp	(Kang et al. 2006)
	Str-1595R	ATC TCT AGG AAT AGC AC		
Pediococcus	Pedio-3F	CTG AAT GAG ATT TTA ACA CG	1200 bp	(Singh et al. 2008)
	Pedio-3R	GGT TTT AAG AGA TTA GCT		
Leuconostoc	LeuF	CGA AAG GTG CTT GCA CCT TTC AAG	976 bp	(Jang et al. 2003)
	LeuR	TTT GTC TCC GAA GAG AAC A		

Table 2.2. PCR conditions which targeting different LAB used in this study.

		***Lactobacillus* spp.**		***Lactococcus* spp.**		***Streptococcus* spp.**		***Pediococcus* spp.**		***Leuconostoc* spp.**	
		Temperature	**Time**	**Temperature**	**Time**	**Temperature**	**Time**	**Temperature**	**Time**	**Temperature**	**Time**
Pre-denaturation		95 ℃	5 min	95 ℃	5 min	95 ℃	3 min	94 ℃	5 min	95 ℃	2 min
30 cycles	Denaturation	95 ℃	30 s	94 ℃	1 min	94 ℃	1 min	94 ℃	30 s	94 ℃	1 min
	Annealing	55 ℃	30 s	55 ℃	1 min	50 ℃	1 min	55 ℃	30 s	55 ℃	1 min
	Extension	72 ℃	30 s	72 ℃	1 min	72 ℃	1 min	72 ℃	30 s	72 ℃	1 min
Final-extension		72 ℃	7 min	72 ℃	5 min	72 ℃	10 min	72 ℃	7 min	72 ℃	10 min

2.3.4 DNA extraction

For DNA extraction, 30 ml of raw, pasteurized, cooled, and post-inoculated milk were collected, as well as 1 g of cooked curd, curd with whey removed, cheddared curd, salted curd, fresh cheese, and aged cheese (from week 1 to week 26, for each week, separately), respectively. Thirty ml of four liquid samples were centrifuged at 13,000 ×*g* for 5 min. The lipid layer was scraped with a disposable loop, and the supernatant was discarded. The milk pellet was dissolved into 1.8 ml of water and centrifuged again at 13,000 ×*g* for 2 min. The lipid layer was scraped and the supernatant discarded. 1 g of each solid cheese samples were homogenized in 9 ml of 2% trisodium citric acid buffer ($C_6H_7Na_3O_8$) with vigorous vortexing with incubating in a 55 to 60 ℃ water bath until all the cheese samples were totally melted. A 1.8 ml portion of cheese solution was centrifuged at 13,000 ×*g* for 2 min, and the supernatant was discarded. The DNeasy PowerFood Microbial Kit (Qiagen, Hilden, Germany) was used to isolate DNA from both liquid and solid samples according to the manufacturer's instructions. The concentration of extracted DNA was measured via a Qubit 4 Fluorometer (Thermo Fisher Scientific, Waltham, MA, USA), and subsequently diluted to achieve a final concentration of 10 ng/ul.

2.3.5 16S rRNA gene-based library preparation

A sequencing library was prepared based on the V4 region of 16S rRNA (Kozich et al. 2013). The V4 region of the 16S rRNA gene of each sample was amplified using a high-fidelity polymerase (AccuPrime, Invitrogen, Carlsbad, CA, USA). Amplified DNA fragments were confirmed via gel electrophoresis with 1% agarose gel. Amplified PCR products were normalized through a SequalPrep™ Normalization Kit (Life Technologies, Carlsbad, CA, USA) following the manufacture's instruction. After normalization, 5 µl of each normalized aliquot were combined to make a library pool and quantified using a KAPA Library Quantification Kit (Kapa Biosystems, Woburn, MA, USA).

2.3.6 Microbiome sequencing via an Illumina MiSeq platform

A 20 nM sample of the pooled library was mixed with 20 nM of PhiX control v3 (10%, v/v, Illumina, San Diego, CA, USA) in 0.2 N of fresh NaOH and HT1 buffer (Illumina) until a final concentration of the pooled library was 7.8 pM. The 600 µl of the mixture was transferred into an Illumina MiSeq® v2 (2 x 250 bp, 500 cycles) reagent cartridge. The sequences of the 16S rRNA library are available at the Sequence Read Archive (SRA) of the National Center for Biotechnology Information (NCBI) under SUB6885679.

2.3.7 Data analyses

Both of the demultiplexed R1 and R2 raw sequences files as a fastq format were directly downloaded from the Illumina BaseSpace website (https://basespace.illumina.com/dashboard). The Quantitative Insights into Microbial Ecology 2 (QIIME 2) version 2019.01 (Bolyen et al. 2019) as an open-source pipeline was used to analyze the sequence data. Demultiplexed sequences were joined and denoised based on quality scores through DADA2 scripts available in QIIME 2. DADA2 scripts were merged to 100% sequence homology to construct an amplicon sequence variant (ASV) table. The taxonomy of sequence was acquired from the Greengenes reference database (version 13.8, http://greengenes.lbl.gov) using an ASV table at 99% sequence similarity. To visualize the processed data, the ASV table was uploaded with taxa in plain format and metadata files were imported to the MicrobiomeAnalyst tool available at http://www.microbiomeanalyst.ca (Dhariwal et al. 2017). Alpha and beta diversities are exported from MicrobiomeAnalyst. Analysis of variance (ANOVA) and analysis of group similarity (ANOSIM) tests were applied to assess the significant difference for alpha and beta diversities, respectively. To identify the most discriminant taxa among cheese samples from different steps based on the relative abundance, Linear Discriminant Analysis (LDA) of Effect Size (LEfSe) was used with an LDA score of 2.0 (Segata et al. 2011). The functional properties of the cheese microbiome were predicted using the 16S rRNA data via a phylogenetic investigation of communities by reconstruction of unobserved states 2 (PICRUSt2) and a Piphillin (with 97% identity cut-

off) (Iwai et al. 2016) based analysis with the KEGG database (Langille et al. 2013). Statistical Analysis of Metagenomic Profiles (STAMP) was used to compare microbial relevant functional changes during the cheese making process by generating a principal component analysis (PCA) plot based on the KEGG orthology (KO) from PICRUSt2 and Piphillin (Parks et al. 2014). Tukey-Kramer (ANOVA) statistical test was used for the PCA plot in STAMP.

2.4 Results

Cheddar cheeses were produced in three different batches made on April 26th, June 6th, and June 26th in 2018. In order to understand any potential source for NSLAB from the cheese making facilities, swabs from equipment and facility surfaces were collected and inoculated into MRS media. A total of 24 locations were selected and swabbed in triplicate (a total of 72 samples; Figure. 2.2).

2.4.1 Taxonomy analysis

There were a total of 773,821 sequencing reads generated from 108 samples (3 batches of cheese, 36 samples per batch) collected during this study from raw milk to aging. The mean value of the frequency of sequences per sample was found to be 7,165 reads/sample after data were analyzed using QIIME 2 pipeline. A total of 271 ASVs were identified by QIIME 2. The processed sequencing data was aligned to the Greengenes

reference database (McDonald et al. 2012). The highest bacterial diversity came from raw and pasteurized milk before the addition of SLAB (Step 1), while most of the bacteria in cheese samples consisted primarily of *Streptococcus* spp. and *Lactococcus* spp. after SLAB was added (Steps 2 and 3).

In our study, the relative abundance of organisms was dependent on the SLAB, which consisted of *L. lactis* subsp. *lactis* biovar diacetylactis*, L. lactis* subsp. *lactis, L. lactis* subsp. *cremoris*, and *S. thermophilus*. These organisms were found to be the dominant bacteria in Cheddar cheese after inoculation through the aging. During the cheese making step (Step 2), the relative abundance of *Streptococcus* tended to decrease while relative abundance of *Lactococcus* tended to increase among all the cheese batches. The relative abundance of SLAB (*Lactococcus* and *Streptococcus*) fluctuated during the aging step. Fifteen weeks of aging after cheese production, unidentified *Lactobacillaceae* (NSLAB) were detected in the cheese samples. *Lactobacillaceae* were not present in the raw milk and were not intentionally introduced during the cheese making process (Figure 2.3). *Escherichia coli* was detected in the raw and pasteurized milk samples. The relative abundance of *E. coli* was the highest (1 to 23%) before inoculation with SLAB (Step 1), dropped to less than 0.1% during cheese manufacturing (Step 2) and became undetectable (0%) in aged cheese samples (Step 3).

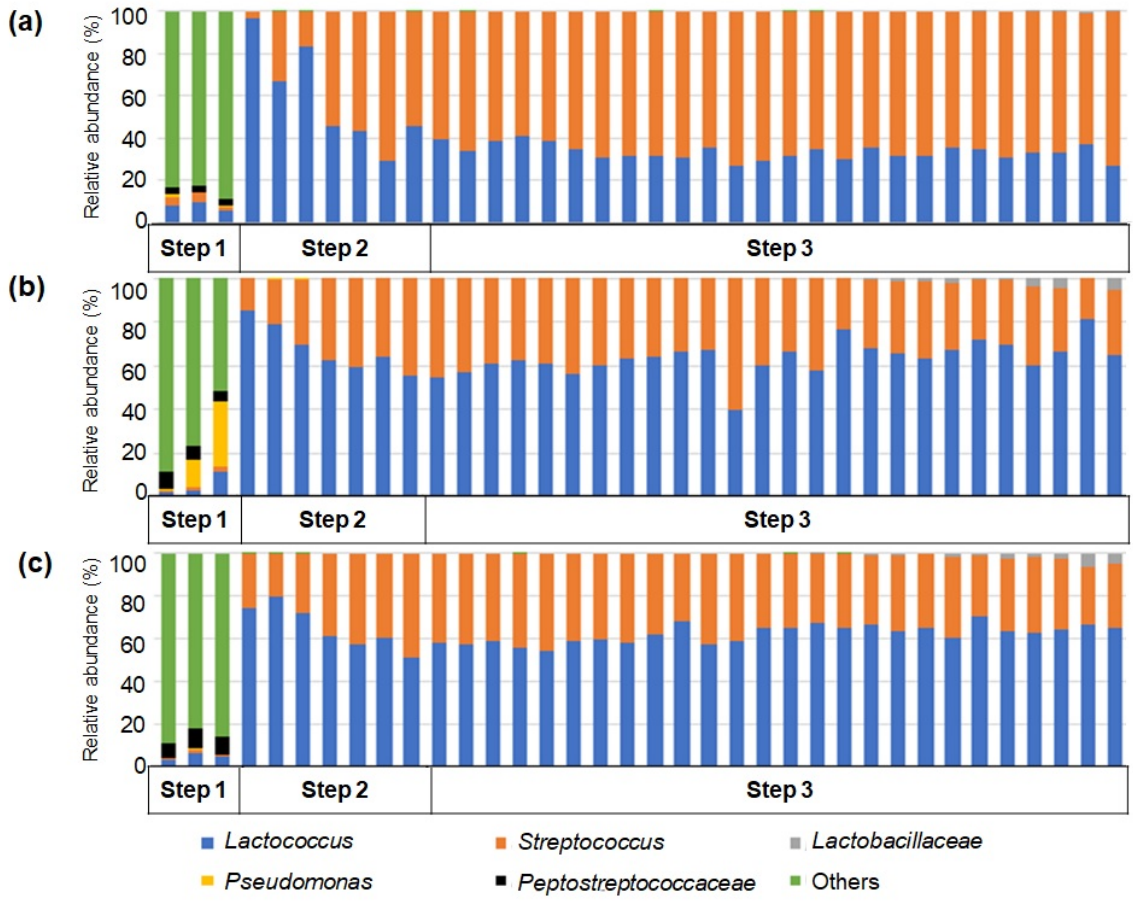


Figure 2.3. Taxonomic bar plot of relative frequency of microbial counts identified in three different batches of cheese showing the top 5 representative taxa at the genus level; (a) April 26th, (b) June 6th, and (c) June 26th in 2018 produced Cheddar cheese. Samples were separated by the before adding SLAB (Step 1), cheese sample during making procedure (Step 2), and aging (Step 3).

2.4.2 Alpha diversity

Alpha diversity of the microbial communities was analyzed using the Chao1 and Shannon indices (Fig. 2.4a and b). Both models were derived from the data input into MicrobiomeAnalyst (https://www.microbiomeanalyst.ca). The Chao1 index indicates the richness of microbial communities from the raw milk to the 6-month-aged cheese. The richness of milk samples (Step 1) was significantly higher (30 to 110) than the post-SLAB (Steps 2 and 3) samples (2 to 6). The Shannon index accounts for both richness and evenness of ASVs; high values for diversity represent more diverse communities. The Shannon alpha diversity index prior to the addition of SLAB (Step 1) was significantly higher (2.8 to 4.3) compared to post-inoculation samples (Steps 2 and 3), which showed a significantly lower alpha diversity index (less than 0.8).

2.4.3 Beta diversity

Beta diversity outputs were used to assess the microbiome structure of the cheese samples and determine whether there were significant differences between processing steps. The Jaccard index model, which was derived from MicrobiomeAnalyst, showed obvious groupings between samples (Figure 2.4c and d). According to the Jaccard index model, the groupings of the samples tightly clustered by cheese making step (Steps 1 to

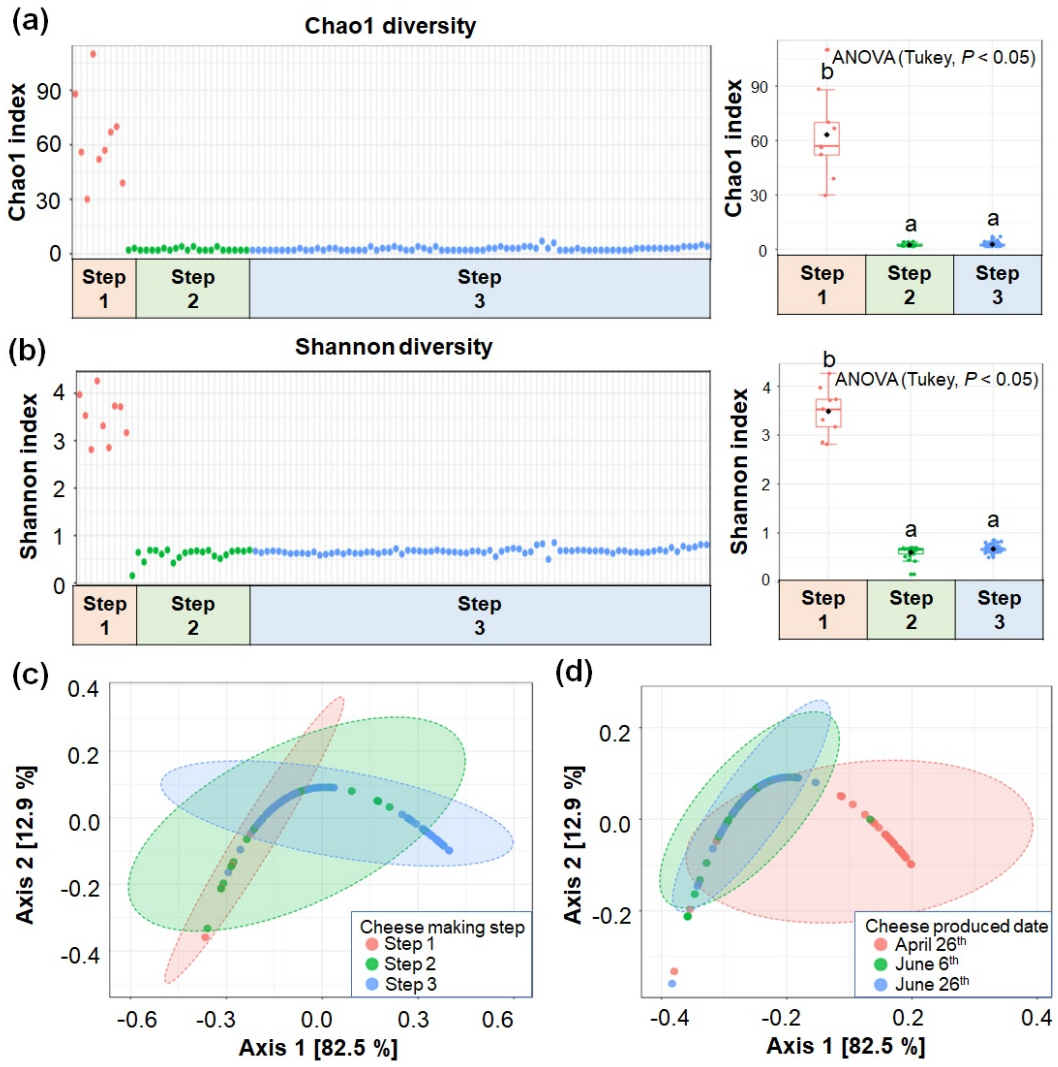


Figure 2.4. Alpha diversity of Cheddar cheese. The (a) Chao1 index and (b) Shannon index of Cheddar cheese. Samples were categorized by before adding SLAB (Step 1), cheese sample during making procedure (Step 2), and aging (Step 3). Beta diversity, principal coordinates analysis (PCoA) of Cheddar cheese were ordinated based on the Jaccard distance matrix. (c) Samples were categorized by Step 1, 2, and 3. (d) Samples were separated by April 26th, June 6th, and June 26th in 2018 produced Cheddar cheese.

3) and changed sequentially according to the cheese-manufacturing step. Additionally, the aged cheese samples from April 26th produced a unique cluster that was separate from cheeses produced on June 6th and June 26th, which formed a second cluster. This indicates that the microbial relationship between cheese groups is dependent both on cheese making steps and produced batches.

2.4.4 Biomarkers of cheese

Biomarkers from the raw milk to 6-months-aged cheese were assessed using LEfSe (http://huttenhower.sph.harvard.edu/lefse/) which is an algorithm for discovering biomarkers from target samples. Biomarkers are bacterial communities that are significantly and relatively highly abundant in two or more samples that help to explain conditions of the sample source (Segata et al. 2011). LEfSe was applied to the microbiome data from the raw milk through the 6-months-aged samples of Cheddar cheese. LEfSe identified a total of 6 differentially abundant taxonomic clades with an LDA score higher than 2.0 (Figure 2.5). *Bacilli*, *Streptococcaceae*, and *Lactobacillales* were identified as biomarkers after the addition of SLAB (Step 2) while *Flavobacteriales*, *Flavobacteriia*, and *Firmicutes* were identified as biomarkers prior to inoculation (Step 1).

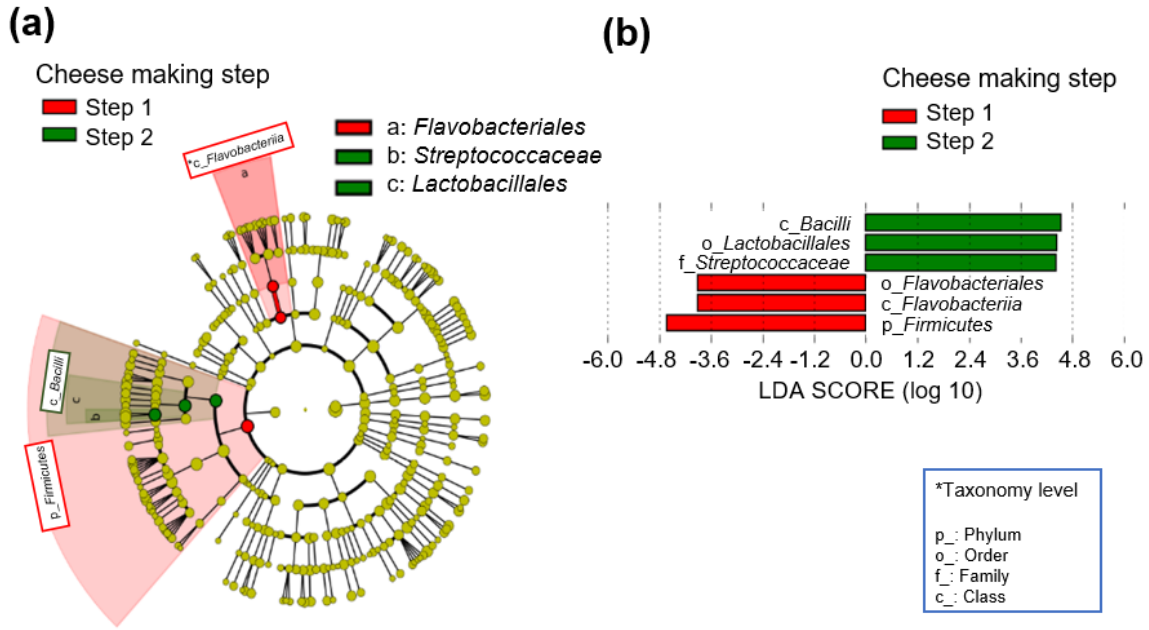


Figure 2.5. Biomarkers of cheese microbiota during cheese manufacturing and aging presented as a taxonomic cladogram obtained from LEfSe analysis of 16S sequencing through (a) cladogram plot and (b) LEfSe result.

2.4.5 Environmental contamination analysis

To identify any potential NSLAB contamination sources, swabs were taken from facility surfaces and equipment used in the cheese making process. A total of 24 locations with three sampling points per location were selected. Of these samples, cultures were obtained from 41 out of 72 inoculated into MRS broth. These samples were analyzed for five representative LAB (*Lactococcus* spp., *Lactobacillus* spp., *Streptococcus* spp., *Pediococcus* spp., and *Leuconostoc* spp.) at the genus level using a PCR assay. Of the 41 samples, two were identified as *Streptococcus* spp., which came from cheese mold (body) and shovel samples (Table 2.3).

2.4.6 Bacterial function in cheese

To understand bacterial functional properties and their changes during the processing of Cheddar cheese from raw milk through 6 months of aging, the cheese's functional potentials and compositions were predicted using 16S rRNA marker genes from Piphillin and PICRUSt2. Several predicted pathways were significantly enriched when the microbiome data was collapsed into 11 different functions (Table 2.4). These 11 different functions included: nucleotide metabolism, carbohydrate metabolism, amino acid metabolism, metabolism of other amino acids, metabolism of cofactors and vitamins, glycan biosynthesis and metabolism, energy metabolism, lipid metabolism, biosynthesis of other secondary metabolites, metabolism of terpenoids and polyketides and

Table 2.3. Five representative LAB (*Lactobacillus* spp., *Lactococcus* spp., *Streptococcus* spp., *Pediococcus* spp., and *Leuconostoc* spp.) detected from swab samples of a cheese making facility using a PCR assay.

Location	**Culture positive on MRS media**	**PCR result**				
		***Lactobacillus* spp.**	***Lactococcus* spp.**	***Streptococcus* spp.**	***Pediococcus* spp.**	***Leuconostoc* spp.**
Table	2/3	-*	-	-	-	-
Vacuum sealer	1/3	-	-	-	-	-
Sink	3/3	-	-	-	-	-
Cheese storage	1/3	-	-	-	-	-
Floor	3/3	-	-	-	-	-
Wall	3/3	-	-	-	-	-
Rennet refrigerator	0/3	-	-	-	-	-
Starter refrigerator	3/3	-	-	-	-	-
Scale	2/3	-	-	-	-	-
Portable table	3/3	-	-	-	-	-
Under table	2/3	-	-	-	-	-
Under vat	0/3	-	-	-	-	-
Under portable table	0/3	-	-	-	-	-
Cheese knife	2/3	-	-	-	-	-
Cheese probe	3/3	-	-	-	-	-
Stirring paddle	3/3	-	-	-	-	-
Cheddar mill	1/3	-	-	-	-	-
Mold body	2/3	-	-	+**	-	-
Mold cap	0/3	-	-	-	-	-
Mold inner net	1/3	-	-	-	-	-
Pressure	0/3	-	-	-	-	-
Cheese vat	2/3	-	-	-	-	-
Cheddaring table	2/3	-	-	-	-	-
Shovel	2/3	-	-	+	-	-

* Negative PCR results according to the 1% agarose gel electrophoresis.

** Positive PCR results according to the 1% agarose gel electrophoresis.

Table 2.4. Relative abundance (%) of functional properties of bacterial genes in three different steps of Cheddar cheese making and aging (step 1: before adding SLAB, step 2: cheese sample during manufacture, and step 3: aging cheese)

Category	PICRUSt2			Piphillin		
	Step 1	Step 2	Step 3	Step 1	Step 2	Step 3
Carbohydrate metabolism	12.34 ± 0.43^{a}	14.23 ± 0.14^{b}	14.15 ± 0.14^{b}	20.22 ± 0.86^{A}	21.85 ± 0.25^{B}	21.73 ± 0.22^{B}
Amino acid metabolism	14.25 ± 0.39^{b}	13.27 ± 0.13^{a}	13.17 ± 0.11^{a}	21.84 ± 0.52^{B}	20.30 ± 0.42^{A}	20.53 ± 0.33^{A}
Nucleotide metabolism	18.48 ± 1.06^{a}	19.68 ± 0.17^{b}	19.81 ± 0.14^{b}	13.86 ± 0.45^{A}	13.73 ± 0.24^{A}	13.86 ± 0.18^{A}
Metabolism of cofactors and vitamins	13.84 ± 0.25^{b}	11.95 ± 0.09^{a}	11.99 ± 0.10^{a}	11.89 ± 0.32^{B}	11.37 ± 0.09^{A}	11.31 ± 0.06^{A}
Energy metabolism	7.54 ± 0.12^{b}	6.73 ± 0.10^{a}	6.66 ± 0.10^{a}	9.56 ± 0.30^{B}	8.92 ± 0.27^{A}	8.75 ± 0.20^{A}
Lipid metabolism	6.60 ± 0.23^{b}	5.87 ± 0.06^{b}	5.91 ± 0.05^{a}	5.88 ± 0.31^{B}	5.82 ± 0.18^{AB}	5.73 ± 0.14^{A}
Metabolism of other amino acids	12.13 ± 0.73^{a}	12.53 ± 0.10^{b}	12.48 ± 0.10^{b}	4.44 ± 0.19^{A}	4.82 ± 0.10^{B}	4.88 ± 0.07^{B}
Glycan biosynthesis and metabolism	$5.39 \pm .049^{a}$	7.01 ± 0.15^{b}	7.10 ± 0.15^{b}	3.23 ± 0.22^{A}	4.28 ± 0.16^{B}	4.37 ± 0.13^{B}
Metabolism of terpenoids and polyketides	3.09 ± 0.02^{a}	3.02 ± 0.01^{b}	3.02 ± 0.01^{b}	3.74 ± 0.11^{A}	3.87 ± 0.17^{B}	3.77 ± 0.12^{AB}
Xenobiotics biodegradation and metabolism	3.19 ± 0.33^{b}	2.71 ± 0.08^{a}	2.67 ± 0.08^{a}	2.97 ± 0.22^{A}	3.14 ± 0.05^{B}	3.10 ± 0.04^{B}
Biosynthesis of other secondary metabolites	3.16 ± 0.10^{b}	2.99 ± 0.07^{a}	3.03 ± 0.07^{a}	2.35 ± 0.19^{B}	1.90 ± 0.09^{A}	1.96 ± 0.06^{A}

[a-b] Identical superscripts in each row exhibit no differences at the 95% significance level in PICRUSt2 results ($P < 0.05$).
[A-B] Identical superscripts in each row exhibit no differences at the 95% significance level in Piphillin results ($P < 0.05$).

xenobiotics biodegradation and metabolism. Functional property differences between the three different cheese manufacturing steps were analyzed and compared through a PCA plot (Figure 2.6). According to the PCA plot, the functional properties of Step 1 differed significantly from the functional properties of Steps 2 and 3.

2.5 Discussion

108 samples were taken from 3 different batches of Cheddar cheeses produced at the Arbuthnot Dairy Center at Oregon State University, beginning with raw milk and following through the end of 26 weeks of aging. High throughput sequencing was used to generate a list of 271 ASVs from the Greengenes database. The majority of the ASVs were from raw or pasteurized milk before the addition of SLAB (Step 1) and were not detected after the addition SLAB (Steps 2 and 3). The most prevalent bacteria in the Cheddar cheese were *Streptococcus* spp. and *Lactococcus* spp., both of which came from the SLAB (composed of *L. lactis* subsp. *lactis* biovar diacetylactis*, L. lactis* subsp. *lactis, L. lactis* subsp. *cremoris*, and *S. thermophilus*). The relative abundance of *Lactococcus* was higher than *Streptococcus* after being added to the pasteurized milk, however the relative abundance of *Lactococcus* decreased and *Streptococcus* increased during cheese making procedure. According to Jonnala et al. (2018), the combination of pressure, temperature and pH determines the bacterial composition during cheese making. In this study, the increased temperature during cooking (Sample F) and the decreased pH by

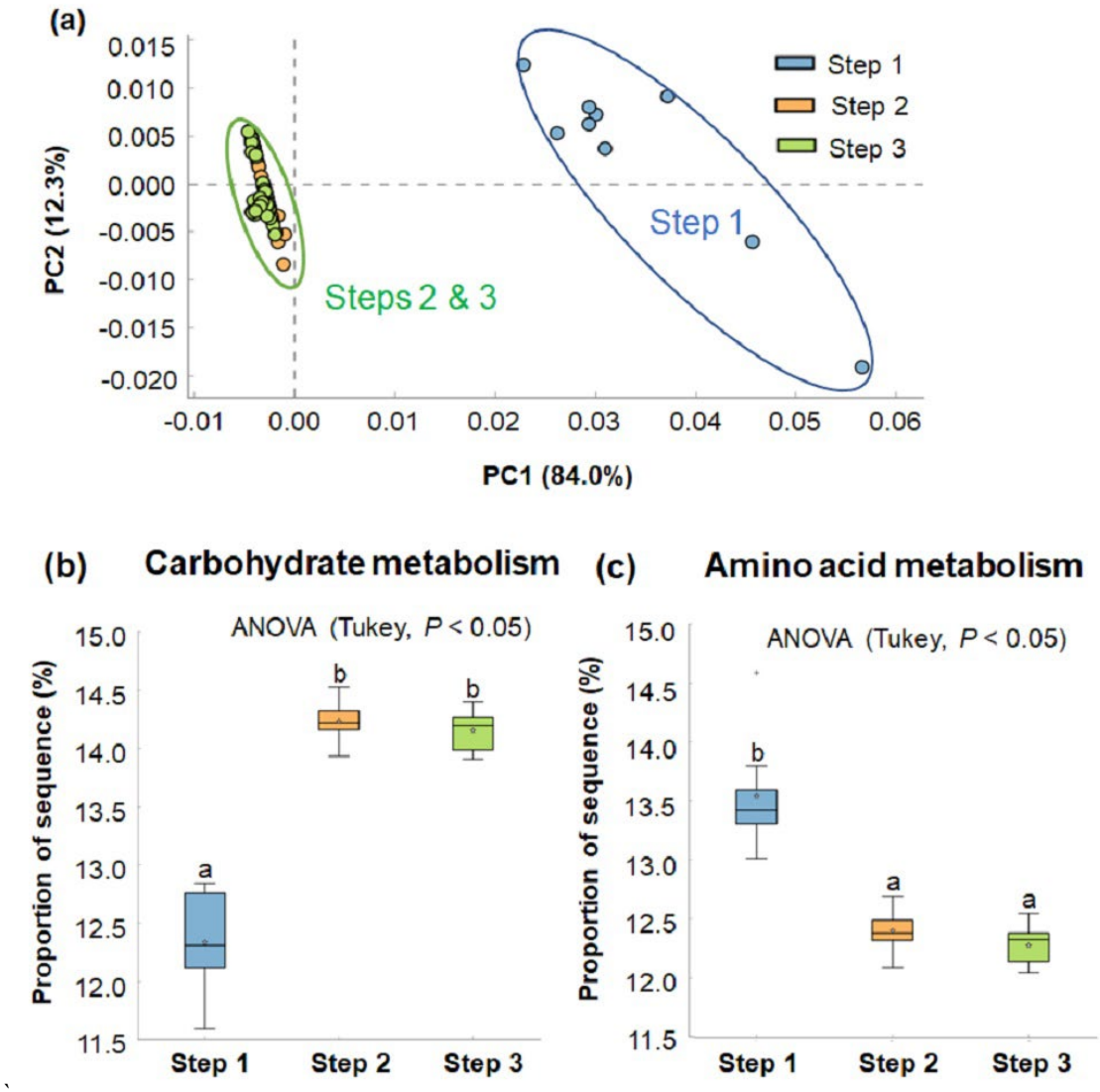


Figure 2.6. Changes of functional properties of cheese microbiota during cheese manufacturing and aging obtained from PICRUSt2 analysis of 16S rRNA marker gene sequencing. Cheese samples were categorized by Step 1 (before adding SLAB), Step 2 (cheese sample during making procedure), and Step 3 (aging). (a) PCA plot for functional annotated genes differences of cheese microbiota, boxplot for the proportion of (b) carbohydrate metabolism related sequences and (c) amino acid metabolism related sequences.

draining whey and cheddaring (Samples G and H) influenced the microbial changes during cheese making. *Streptococcus* appears to adjust to the stresses of cheese making better than *Lactococcus*. During the aging step, the relative abundances of *Lactococcus* and *Streptococcus* changed stochastically. After 15 weeks of aging, a number of unidentified *Lactobacillaceae* (NSLAB) were detected; these were not present in the raw milk or in the SLAB.

The primary function of SLAB is to produce enough lactic acid to reduce the pH of the milk to the desired level. In this study, the desired pH was around 5.3. SLAB also contribute to cheese aging, as their enzymes are involved in proteolysis, lipolysis, and the conversion of amino acids into flavor compounds; these processes alter the chemical, microbiological and sensory properties of the end products (Caplice and Fitzgerald 1999; Cogan et al. 2007; Leroy and De Vuyst 2004). For example, the species *S. thermophilus,* which was used as a SLAB in this Cheddar cheese, is widely used in fermented dairy products, including yogurt and other cheeses. In addition to producing lactic acid, *S. thermophilus* also produces exopolysaccharides which improve the texture of fermented products (Almirón-Roig et al. 2000; Mora et al. 2002). *Lactococcus lactis* subsp. *lactis* and *Lactococcus lactis* subsp. *cremoris* are extensively used in SLAB cultures for fermented dairy products including cheese, sour cream and butter (Beresford et al. 2001).

NSLAB are important in cheese making because they have the potential to impact flavor development during aging (Cotter and Beresford 2017). In previous studies, NSLAB isolated from cheese belonged to a very heterogeneous group; frequently they were members of *Lactobacillus* spp., which included *Lactobacillus casei*, *Lactobacillus paracasei*, *Lactobacillus plantarum*, *Lactobacillus pentosus*, *Lactobacillus curvatus*,

Lactobacillus rhamnosus, *Lactobacillus fermentum*, *Lactobacillus parabuchneri* and *Lactobacillus brevis* (Quigley et al. 2012; Settanni and Moschetti 2010). This is consistent with our finding of unidentified *Lactobacilliceae* during fermentation. The origin of our NSLAB was not determined, however NSLAB can be introduced from a variety of sources. Raw milk is the most well-known source of NSLAB (Montel et al. 2014). NSLAB also have been isolated from various locations within cheese facilities, such as floors, drains, surfaces of equipment used in cheese manufacture, and packaging equipment (Bokulich and Mills 2013). Provolone and Swiss cheese made in the same facility as our Cheddar cheese use *Lactobacillus delbrueckii* subsp. *bulgaricus* and *L. delbrueckii* subsp. *lactis*, and *Lactobacillus helveticus* as SLAB. It is possible that cross-contamination from these other starter cultures could be a potential source of NSLAB in our Cheddar cheese samples.

Salmonella, *Brucella*, *Mycobacterium*, *Coxiella*, *Listeria*, *E. coli*, *Campylobacter*, *Corynebacterium*, *Yersinia*, and *Bacillus* have been identified as bacterial pathogens in raw milk (Claeys et al. 2013). In our sequencing data, only *E. coli* was observed in the raw and pasteurized milk before the addition of SLAB. After inoculation with SLAB (*Lactococcus* spp. and *Streptococcus* spp.), the relative abundance of *E. coli* decreased to less than 1% during manufacturing (Step 2) and was not detected (0%, Step 3) during the aging process. Pathogens can be introduced into cheese through their presence in raw milk and can survive into subsequent steps of the cheese making process (Johnson et al. 1990). Despite a variety of sources for pathogenic contamination, factors such as pH, salt content, and water activity can prevent the growth of unwanted bacteria in aged hard cheeses (D'Amico and Donnelly 2017). In addition, LAB can produce bacteriocins or

bacteriocin-like substances (BLS) that inhibit the growth of foodborne pathogens (Yang et al. 2012). According to Trmčić et al. (2016), none of the 9 hard cheeses produced from pasteurized milk were positive for coliforms, a common indicator for the presence of pathogens, while 4 of 19 soft cheeses and 3 of 11 fresh cheeses were positive for coliforms. The lack of pathogens in their hard cheeses was consistent with our findings that no pathogens were present after 26 weeks of aging.

According to the alpha diversity given by the Chao1 and Shannon indices, the samples taken prior to the addition of SLAB (Step 1) scored significantly higher than samples taken after inoculation (Steps 2 and 3). SLAB, which dominate the cheese microbiome after inoculation, produces bacteriocins. This combined with the low pH, high salt content, and low water activity of hard cheeses inhibits the growth of other bacteria in the cheese (Yang et al. 2012), leading to a decrease in the microbial richness of cheese after the addition of SLAB. The beta diversity of samples taken before inoculation with SLAB also showed a significantly different microbial community structure compared to samples taken after adding SLAB. According to the Randazzo et al. (2002), the initial microbiota diversity in the raw milk was remarkably different from that of the final aged cheese since most of the bacteria in the raw milk were no longer detected as SLAB dominated during the cheese making process. These findings were consistent with our beta diversity results. The similarity of cheeses is also influenced by whether or not the cheeses came from the same batch. According to the Alessandria et al. (2016), microbial properties of raw milk impact the development of the microbiota during aging. Furthermore, the cheese making facility also influences the microbial composition of the final product.

The NGS data using LEfSe found multiple biomarkers that differed significantly between different stages of cheese manufacture and aging. A total of 6 different bacteria were identified as biomarkers for the cheese making process. Before adding the starter culture (Step 1), *Flavobacteriales* (order), *Flavobacteriia* (class), and *Firmicutes* (phylum) were identified as biomarkers. However, after inoculation, *Bacilli* (class), *Streptococcaceae* (family), and *Lactobacillales* (order) were identified as biomarkers. According to LEfSe, the SLAB became biomarkers after inoculation.

In the predicted functional properties of the cheese microbiome, almost half of the DNA reads were assigned to the nucleotide metabolism (19.67%), carbohydrate metabolism (14.02%), and amino acid metabolism (13.28%). Carbohydrate metabolism is a key process in cheese fermentation since it converts the lactose in the milk and curd into lactic acid (Porcellato and Skeie 2016). According to Figure 2.6b, the relative abundance of genes related to carbohydrate metabolism were significantly lower in samples taken before the addition of starter culture (Step 1, 18.48%) compared to afterwards (Steps 2 and 3, 19.68 and 19.81%, respectively). This is not surprising since the *Streptococcus* spp. and *Lactococcus* spp. from the starter culture are representative of the lactic acid bacteria that are primarily used to ferment lactose into lactic acid. Another important role of the cheese microbiota is related to amino acid metabolism (13.28%). Catabolism of amino acids is important for the production of flavor compounds in cheese, which can improve quality. While the addition of SLAB increased carbohydrate related functions, genes related to amino acid metabolism decreased since the primary goal of SLAB inoculation is to produce lactic acid.

The present study improves our understanding of how the composition, diversity, and functional properties of the cheese microbiota change throughout the cheese making process, beginning with raw milk and proceeding to a 6-months-aged final product. In this study, we found that SLAB dominated the post-inoculation cheese microbiota, and that the addition of SLAB caused changes in microbial community structure, biomarkers, microbial diversity, and predicted functional properties. Additionally, unidentified *Lactobacillaceae* known as NSLAB were detected 15 weeks after aging in the June 6th and June 26th produced cheeses and 17 weeks after aging in the April 26th cheese samples.

2.6 References

Alessandria V, Ferrocino I, De Filippis F, Fontana M, Rantsiou K, Ercolini D, Cocolin L (2016) Microbiota of an Italian Grana-like cheese during manufacture and ripening, unraveled by 16S rRNA-based approaches. Appl Environ Microbiol 82(13):3988-3995.

Almirón-Roig E, Mulholland F, Gasson MJ, Griffin AM (2000) The complete cps gene cluster from *Streptococcus thermophilus* NCFB 2393 involved in the biosynthesis of a new exopolysaccharide. Microbiology 146(11):2793-2802

Beresford TP, Fitzsimons NA, Brennan NL, Cogan TM (2001) Recent advances in cheese microbiology. Int Dairy J 11:259-274

Bokulich NA, Mills DA (2013) Facility-specific "house" microbiome drives microbial landscapes of artisan cheesemaking plants. Appl Environ Microbiol 79(17):5214-5223

Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, Bai Y, Bisanz JE, Bittinger K, Brejnrod A, Brislawn CJ, Brown CT, Callahan BJ, Caraballo-Rodríguez AM, Chase J, Cope EK, Da Silva R, Diener C, Dorrestein PC, Douglas GM, Durall DM, Duvallet C, Edwardson CF, Ernst M, Estaki M, Fouquier J, Gauglitz JM, Gibbons SM, Gibson DL, Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S, Holste H, Huttenhower C, Huttley GA, Janssen S, Jarmusch AK, Jiang L, Kaehler BD, Kang KB, Keefe CR, Keim P, Kelley ST, Knights D, Koester I, Kosciolek T, Kreps J, Langille MGI, Lee J, Ley R, Liu YX, Loftfield E, Lozupone C, Maher M, Marotz C, Martin BD, McDonald D, McIver LJ, Melnik AV, Metcalf JL, Morgan SC, Morton JT, Naimey AT, Navas-Molina JA, Nothias LF, Orchanian SB, Pearson T, Peoples SL, Petras D, Preuss ML, Pruesse E, Rasmussen LB, Rivers A, Robeson MS, Rosenthal P, Segata N, Shaffer M, Shiffer A, Sinha R, Song SJ, Spear JR, Swafford AD, Thompson LR, Torres PJ, Trinh P, Tripathi A, Turnbaugh PJ, Ul-Hasan S, van der Hooft JJJ, Vargas F, Vázquez-Baeza Y, Vogtmann E, von Hippel M, Walters W, Wan Y, Wang M, Warren J, Weber KC, Williamson CHD, Willis AD, Xu ZZ, Zaneveld JR, Zhang Y, Zhu Q, Knight R, Caporaso JG (2019) Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol 37: 852–857

Broadbent JR, Cai H, Larsen R, Hughes J, Welker D, De Carvalho V, Tompkins T, Ardö Y, Vogensen F, De Lorentiis A (2011) Genetic diversity in proteolytic enzymes and amino acid metabolism among *Lactobacillus helveticus* strains. J Dairy Sci 94(9):4313-4328

Caplice E, Fitzgerald GF (1999) Food fermentations: role of microorganisms in food production and preservation. Int J Food Microbiol 50(1-2):131-149

Claeys WL, Cardoen S, Daube G, De Block J, Dewettinck K, Dierick K, De Zutter L, Huyghebaert A, Imberechts H, Thiange P (2013) Raw or heated cow milk consumption: Review of risks and benefits. Food Control 31(1):251-262

Cogan T, Beresford T, Steele J, Broadbent J, Shah N, Ustunol Z (2007) Invited review: advances in starter cultures and cultured foods. J Dairy Sci 90(9):4005-4021

Cotter PD, Beresford TP (2017) Microbiome changes during ripening. In: McSweeny PLH, Fox PF, Cotter PD, Everett DW (eds) Cheese. Elsevier, Amsterdam, pp 389-409

D'Amico DJ, Donnelly CW (2017) Growth and survival of microbial pathogens in cheese. In: McSweeny PLH, Fox PF, Cotter PD, Everett DW (eds) Cheese. Elsevier, Amsterdam pp 573-594

Deasy BM, Rea MC, Fitzgerald GF, Cogan TM, Beresford TP (2000) A rapid PCR based method to distinguish between *Lactococcus* and *Enterococcus.* Syst Appl Microbiol 23.4: 510-522.

Dugat-Bony E, Straub C, Teissandier A, Onesime D, Loux V, Monnet C, Irlinger F, Landaud S, Leclercq-Perlat MN, Bento P, Fraud S, Gibrat J-F, Aubert J, Fer F, Guédon E, Pons N, Kennedy S, Beckerich J, Swennen D, Bonnarme P (2015) Overview of a surface-ripened cheese community functioning by meta-omics analyses. PLoS ONE 10(4):e0124360.

De Filippis F, Genovese A, Ferranti P, Gilbert JA, Ercolini D (2016) Metatranscriptomics reveals temperature-driven functional changes in microbiome impacting cheese maturation rate. Sci Rep 6:21871

Dhariwal A, Chong J, Habib S, King IL, Agellon LB, Xia J (2017) MicrobiomeAnalyst: a web-based tool for comprehensive statistical, visual and meta-analysis of microbiome data. Nucleic Acids Res 45(W1):W180-W188

Dubernet, Ségolène, Nathalie Desmasures, Micheline Guéguen (2002) A PCR-based method for identification of *lactobacilli* at the genus level. FEMS Microbiol Lett 214.2: 271-275.

Gosalvitr P, Cuellar-Franca R, Smith R, Azapagic A (2019) Energy demand and carbon footprint of cheddar cheese with energy recovery from cheese whey. Energy Procedia 161:10-16

Iwai S, Weinmaier T, Schmidt BL, Albertson DG, Poloso NJ, Dabbagh K, DeSantis TZ (2016) Piphillin: improved prediction of metagenomic content by direct inference from human microbiomes. PLoS ONE 11(11):e0166104

Jang J, Kim B, Lee J, Han H (2003) A rapid method for identification of typical *Leuconostoc* species by 16S rDNA PCR-RFLP analysis. J Microbiol Method 55(1): 295-302.

Johnson EA, Nelson JH, Johnson M (1990) Microbiological safety of cheese made from heat-treated milk, Part II. Microbiology. J Food Prot 53(6):519-540

Jonnala BRY, Mcsweeney PLH, Sheehan JJ, Abram F (2018) Sequencing of the cheese microbiome and its relevance to industry. Front Microbiol 9:1020

Kang BC, Song ES, Jang HJ, Hyung PJ, Park HK, Kim CM (2006) Development of genus- and species-specific probe design system for pathogen detection based on 23S rDNA. J. Microbiol. Biotechnol 16(5), 740-747.

Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD (2013) Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. Appl Environ Microbiol 01043-13

Langille MG, Zaneveld J, Caporaso JG, McDonald D, Knights D, Reyes JA, Clemente JC, Burkepile DE, Thurber RLV, Knight R (2013) Predictive functional profiling

of microbial communities using 16S rRNA marker gene sequences. Nat Biotechnol 31(9):814

Leroy F, De Vuyst L (2004) Lactic acid bacteria as functional starter cultures for the food fermentation industry. Trends Food Sci Technol 15(2):67-78

Mayo B, TCC Rachid C, Alegría Á, MO Leite A, S Peixoto R, Delgado S (2014) Impact of next generation sequencing techniques in food microbiology. Curr Genomics 15(4):293-309

Montel M-C, Buchin S, Mallet A, Delbes-Paus C, Vuitton DA, Desmasures N, Berthier F (2014) Traditional cheeses: rich and diverse microbiota with associated benefits. Int J Food Microbiol 177:136-154

Mora D, Fortina M, Parini C, Ricci G, Gatti M, Giraffa G, Manachini P (2002) Genetic diversity and technological properties of *Streptococcus thermophilus* strains isolated from dairy products. J Appl Microbiol 93(2):278-287

Parks DH, Tyson GW, Hugenholtz P, Beiko RG (2014) STAMP: statistical analysis of taxonomic and functional profiles. Bioinformatics 30(21):3123-3124

Percival B, Percival F (2017) Reinventing the wheel: Milk, microbes, and the fight for real cheese. Univ of California Press, California

Porcellato D, Skeie SB (2016) Bacterial dynamics and functional analysis of microbial metagenomes during ripening of Dutch-type cheese. Int Dairy J 61:182-188

Potočki S (2016) Potential health benefits of sphingolipids in milk and dairy products. Mljekarstvo 66(4):251-261

Quigley L, O'Sullivan O, Beresford TP, Ross RP, Fitzgerald GF, Cotter PD (2012) High-throughput sequencing for detection of subpopulations of bacteria not previously associated with artisanal cheeses. Appl Environ Microbiol 78(16):5717-5723

Randazzo CL, Torriani S, Akkermans AD, de Vos WM, Vaughan EE (2002) Diversity, dynamics, and activity of bacterial communities during production of an artisanal Sicilian cheese as evaluated by 16S rRNA analysis. Appl Environ Microbiol 68(4):1882-1892

Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, Huttenhower C (2011) Metagenomic biomarker discovery and explanation. Genome Biol 12(6):R60

Settanni L, Moschetti G (2010) Non-starter lactic acid bacteria used to improve cheese quality and provide health benefits. Food Microbiol 27(6):691-697

Singh AK, Ramesh A (2008) Succession of dominant and antagonistic lactic acid bacteria in fermented cucumber: insights from a PCR-based approach. Food microbiol 25.2: 278-287

Solieri L, Dakal TC, Giudici P (2013) Next-generation sequencing and its potential impact on food microbial genomics. Ann Microbiol 63(1):21-37

Trmčić A, Chauhan K, Kent D, Ralyea R, Martin N, Boor K, Wiedmann M (2016) Coliform detection in cheese is associated with specific cheese characteristics, but no association was found with pathogen detection. J Dairy Sci 99(8):6105-6120

United States Department of Agriculture (2018) Dairy: World markets and trades. https://apps.fas.usda.gov/psdonline/circulars/dairy.pdf. Accessed 7 February 2020

Van Hoorde K, Heyndrickx M, Vandamme P, Huys G (2010) Influence of pasteurization, brining conditions and production environment on the microbiota of artisan Gouda-type cheeses. Food Microbiol 27(3):425-433

Walther B, Schmid A, Sieber R, Wehrmüller K (2008) Cheese in nutrition and health. Dairy Sci Technol 88(4-5):389-405

Wolfe BE, Button JE, Santarelli M, Dutton RJ (2014) Cheese rind communities provide tractable systems for in situ and in vitro studies of microbial diversity. Cell 158(2):422-33

Yang E, Fan L, Jiang Y, Doucette C, Fillmore S (2012) Antimicrobial activity of bacteriocin-producing lactic acid bacteria isolated from cheeses and yogurts. AMB Express 2(1):48