Analyse comparative de la prévalence et de la diversité des communautés bactériennes des ensilages et du lait cru bovin

Thèse

Alexandre Jules Kennang Ouamba

Doctorat en sciences et technologie des aliments
Philosophiæ doctor (Ph. D.)

Québec, Canada

Analyse comparative de la prévalence et de la diversité des communautés bactériennes des ensilages et du lait cru bovin

Thèse

Alexandre Jules Kennang Ouamba

Sous la direction de :

Denis Roy

Gisèle LaPointe, codirectrice de recherche

Résumé

Le microbiote du lait cru est un déterminant majeur de sa qualité et de celle des produits dérivés. Dans les fermes laitières, l'hygiène et la santé du pis, les fèces, la litière et les fourrages conservés sont les principales sources de microorganismes qui contaminent l'environnement de l'étable et qui peuvent se retrouver dans le lait. Les ensilages de légumineuses et de graminées produits par fermentation lactique, outre les bactéries lactiques, sont riches en espèces microbiennes dont la prévalence, la diversité et l'abondance dépendent de facteurs incontrôlables tels que les paramètres environnementaux, et de facteurs contrôlables tels que les pratiques de gestion adoptées par les fermiers. Le choix du type de fourrages parmi lesquels le foin, les ensilages d'herbe/légume ou de maïs, l'utilisation ou non d'inoculants, le type d'inoculant commercial utilisé et les types de structures d'entreposage sont autant d'éléments qui influencent la composition microbienne des ensilages et dont la gestion a un impact peu documenté sur la qualité microbiologique du lait cru. Les travaux de cette thèse portent sur l'écologie microbienne des fourrages préservés et l'évaluation de leur contribution à la contamination du lait cru de vache. Pour ce faire, des méthodes de préservation d'échantillons de lait cru à base d'azidiol ou de bronopol ont été évaluées pour leur capacité à maintenir intactes la viabilité et l'abondance des communautés bactériennes présentes. Des échantillons de foin, d'ensilages inoculés et non inoculés, et de lait cru ont été prélevés deux fois, à l'automne 2015 et au printemps 2016 dans 24 fermes laitières au Québec. Les analyses métataxonomiques et des charges microbiennes déterminées par PCR quantitative après le traitement ou non des cellules microbiennes au propidium monoazide ont montré que l'azidiol combiné au diméthyle sulfoxide et une température de -20 °C permet de stabiliser le microbiote du lait cru pendant au moins 30 jours, et que l'azidiol seul maintient l'intégrité des communautés bactériennes pendant 10 jours à 4 °C. Ces études ont également démontré que le séquençage à haut-débit des régions V3-V4 et V6-V8 du gène codant pour l'ARN ribosomique 16S génère des données dont l'exploitation peut conduire à des conclusions plus ou moins divergentes selon les hypothèses de départ. L'importance d'un choix judicieux de la région hypervariable à séquencer est soulignée. Nos résultats ont révélé des différences entre le microbiote du foin et celui des ensilages, ainsi qu'entre les types d'ensilage inoculés et non inoculés. De plus, les rations alimentaires à base de foin, d'ensilage d'herbe/légume non inoculé, d'un mélange d'ensilage d'herbe/légume non inoculé et d'ensilage de maïs inoculé, ou d'un mélange d'ensilage d'herbe/légume et de maïs inoculés et non inoculés partagent jusqu'à 31 % de leur microbiote identifié au niveau de variants de séquences avec le lait cru produit dans les fermes correspondantes. Les taxons plausiblement transférés des fourrages au lait appartiennent surtout aux *Proteobacteria*, *Firmicutes* et *Actinobacteria*. Les résultats obtenus suggèrent que la contamination bactérienne du lait cru par les fourrages se fait de manière aléatoire. Il ressort de nos études que ces taxons supposément transférés des ensilages sont en grande partie responsables des différences observées entre les communautés bactériennes du lait des cinq types de rations. Bien que les structures phylogénétiques des

échantillons de lait produits par les vaches alimentées avec les rations d'ensilages non inoculés ou inoculés se soient montrées significativement différentes, il est difficile de conclure sur l'impact réel des inoculants sur la qualité microbiologique du lait cru. L'analyse des réseaux de co-occurrence et de co-exclusion au sein du microbiote a révélé d'une part, dans les ensilages les interactions entre les bactéries lactiques et non-lactiques qui pourraient considérablement influencer le processus de fermentation et ultimement la qualité du produit final, et d'autre part, dans le lait des niches microbiennes associées aux sites de contamination du lait dans l'environnement à la ferme. Par l'implémentation des méthodes d'analyses multivariées et multi-table intégrant le microbiote et les paramètres physico-chimiques des matrices alimentaires échantillonnées, cette thèse propose une approche d'exploitation des données de métataxonomique permettant d'approfondir nos connaissances de la microbiologie du lait cru et des produits laitiers.

Abstract

The microbiota of raw bovine milk is tightly associated with its quality and that of derivatives. On dairy farms, udder health, faeces, beddings, and fermented forage are among the main sources of milk microbial contaminants. Grass/legume and corn silage obtained by lactic fermentation harbour complex microbial community of which the diversity, the prevalence, and abundance are driven by uncontrollable factors such as environmental conditions, or controllable factors inherent to farm management practices. Forage types including hay, grass/legume or corn silage, the use of microbial additives and their commercial types, and the types of storage structures may influence community assembly of mature silage. However, the impact of forage management practices on the raw milk microbiota is not fully understood. This thesis aimed at investigating the microbial ecology of preserved forage and assessing their contribution to the raw milk contamination. To do so, the ability of milk sample preservation methods based on azidiol or bronopol to maintain viable and stable microbiota over time was assessed. Forage and raw milk samples were collected twice from 24 dairy farms in Quebec, in the fall 2015 and the spring 2016. Analyses of metataxonomic and qPCR-based bacterial load data derived from cells treated or not with propidium monoazide to account for viability showed that azidiol combined with dimethyl sulfoxide prevented the microbiota instability of raw milk stored at -20 °C for at least 30 days. Azidiol used alone was additionally found to keep the microbiota in milk samples intact for up to 10 days when stored at 4 °C. It was demonstrated that for hypothesis-driven microbiota studies, divergent conclusions can be drawn from hight-throughput sequencing of the V3-V4 and V6-V8 hypervariable regions of the 16S rRNA gene pool. The importance of the hypervariable region to target is therefore highlighted. Our study revealed differences between hay microbiota and that of silage, whether inoculated or not. Moreover, forage ration combinations shared up to 31 % of their bacterial phylotypes with raw milk samples produced in the corresponding farms. Taxa that were probably transferred from forage ration combinations to raw milk encompassed the phyla *Proteobacteria*, *Firmicutes*, and *Actinobacteria*. Our results suggested that raw milk contamination on the farm occurs erratically, and that transferred taxa were mainly involved in differential abundance outcomes. Although the microbiota of milk samples associated with the five forage ration combinations exhibited differences in phylogenetic composition, concluding on the effects of microbial additives used for ensiling is a challenge. In this thesis, the analysis of bacteria interaction networks showed that co-occurrence or co-exclusion of lactic and non-lactic bacteria might considerably affect the microbiological quality of mature silage at feed-out. On the other hand, the same analysis performed with milk samples revealed microbial niches associated with milk contamination points on the farm environment. Through the implementation of multivariate multi-table analysis methods that integrated data from the microbiota and from the physicochemical characteristics of the sampled matrices, this thesis suggests methodological approaches that exploit metataxonomic data to deepen our knowledge of the raw milk microbiota and dairy products.

Table des matières

Liste des tableaux

Liste des figures

Liste des abréviations, sigles, acronymes

%	Percent
µL	Microliter
AA	Amino acids
AAB	Acetic acid bacteria
ADF	Acid detergent fiber
ADN/DNA	Acide désoxyribonucléique / Desoxyribonucleic acid
Alt	Altitude
AMS	Automatic milking systems
AN	Avogadro number (6.02×10^{23} mol^{-1})
aNDF	Amylase derived neutral detergent fiber
aNDFom	Ash-corrected NDF
ARNr/ rRNA	Acide ribonucléique ribosomique / Ribosomal ribonucleic acid
ASV	Amplicon sequence variant
avgK	Average degree
AZ4	Azidiol-treated samples stored at 4 °C
AZDm	Azidiol- and DMSO-treated samples
B_Buch500	Biotal buchneri 500
B_supersile	Biotal supersile
Bag_Sil	Bag silo
BgS	Bag silo
BH	Benjamini–Hochberg procedure
BkS	Bunker silo
BR4	Bronopol-treated samples stored at 4 °C
Buch_500	Buchneri 500
Bunk-silo	Bunker silo
C	Corn silage
CCpnA	Canonical correspondence analysis
CCS	Circular consensus sequences
CF	Crude fat
cfu	Colony-forming unit
CHO_DM / ESC_DM	Ethanol-soluble carbohydrates as percent dry matter

CHO_NFC / ESC_NFC	Carbohydrate as % non-fiber carbohydrate
CI	Corn silage inoculated
CLR	Centered log-ration
Conv_Sil	Conventional silo (concrete-stave silo)
CP	Crude protein
CP_ADF	Acid detergent fiber-CP
CP_DM	CP as % dry matter
CP_NH3	Ammonia as % CP
CP_Sol	Soluble CP
CvS	Conventional silo
CvS+BkS	Forage stored in conventional and bunker silo on the same farm
CvS+OLS	Forage stored in conventional and oxygen-limiting silos on the same farm
CvS+RdB,	Forage stored in conventional silo and round bale on the same farm
DADA2	Divisive Amplicon Denoising Algorithm version 2
DIABLO	Data integration analysis for biomarker discovery using latent components
DMSO	Dimethyl sulfoxide
DP	Degradable protein
e.g.	For example
EDTA	Ethylenediaminetetraacetic acid
ESC	Ethanol-soluble carbohydrates
FA	Fatty acids
FDR	False discovery rate
g	Gram
GL	Grass/legume silage
GLC	Combination of grass/legume and corn silage uninoculated
GLCI	Combination of grass/legume uninoculated and corn silage inoculated
GLI	Grass/legume silage inoculated
GLICI	Combination of grass/legume and corn silage inoculated
h	Hour
H	Hay
HCl	Hydrocloric acid
i.e.	That is
ICC	Intraclass correlation coefficient
ITS1	Internal transcribed spacer 1

ITS2	Internal transcribed spacer 2
KEGG	Kyoto encyclopedia of genes and genomes
kg	Kilogram
LA	Lactic acid
LA_VFA	Lactic acid expressed as percent volatile fatty acids
LAB	Lactic acid bacteria
Lact	Lactose
LCBD	Local contribution to beta diversity
LME	Linear mixed effects
Log	Logarithm base 10
Long	Longitude
Ls	Loose
MALDI-TOF	Matrix-assisted laser desorption/ionization coupled with time-of-flight mass spectrometer
MENAP	Molecular ecological network analysis pipeline
mg	Milligram
Mg	Magnesium
min	Minute
mL	Milliliter
mM	Millimolar
MPN	Most probable number
MRS	De Man, Rogosa and Sharpe agar (culture medium)
Mstr	Moisture
NDF	Neutral detergent fiber
NDF_D_30	NDF digestibility at 30 hours
NDF_D_30DM	NDF_D_30 as percent dry matter
NFC	Non-fiber carbohydrates
NH_3	Ammonia
NoPre	No preservative
NT	Amplicon length
OL_Sil	Oxygen limiting silo
OLS	Oxygen-limiting silo
OLSq	Forage stored in oxygen-limiting silo and square bale on the same farm
OTU	Operational taxonomic unit
PCA	Principal component analysis

PCoA	Principal coordinate analysis
PCR	Polymerase chain reaction
pH	Potential of hydrogen
PhILR	Phylogenetic isomeric log-ratio
P_i	Among-module connectivity
PMA	Propidium monoazide
PMA-qPCR	Quantitative polymerase chain reaction from propidium monoazide-treated cells
pMEN	Phylogenetic molecular ecological networks
Prot	Proteins
qPCR	Quantitative polymerase chain reaction
RdB	Round bale
RDP	Ribosomal database project
Rnd_Bal	Round bale
SCC	Somatic cell count
SLRS	Synthetic long-read sequences
sp.	Single unidentified species of a given genus
SP_NH3	Ammonia expressed as percent soluble protein
sPLS	Sparse partial least squares regression
spp.	Multiple species of the same genus
SqB	Square bale
Sqr_Bal	Square bale
Stack_Sil	Stack silo
StS	Stack silo
TDN	Total digestible nutrient
v/v	Volume/volume ratio
V1-V2/V1-V3/ V3-V4/V4/V4-V5/V6-V8/V7-V9	Hypervariable regions of the 16S rRNA gene
VBNC	Viable but not culturable cells
VFA	Volatile fatty acids
w/v	Weight/volume ratio
WSC	Water-soluble carbohydrates
x g	Unit for gravitational acceleration
Z_i	Within-module connectivity
zOTU	Zero-radius operational taxonomic unit

À la mémoire de mon père, Daniel Ouamba.

Remerciements

Je remercie grandement mon directeur de thèse, Professeur Denis Roy, de m'avoir offert l'opportunité de faire un doctorat. Je me souviens de ce jour où, alors que j'achevais mon stage de recherche dans votre laboratoire, vous êtes venu vers moi, avec bienveillance comme d'habitude, mais, cette fois-là, pour me proposer un poste de doctorant. Ce fut avec un « grand oui » que j'ai accepté le défi, sans aucun doute que ce serait une aventure riche d'enseignements et d'apprentissage dans plusieurs aspects de ma vie. En me confiant ce projet de thèse, vous réalisiez sans le savoir, un rêve qui commençait à s'évanouir, celui de me former en bioinformatique appliquée à la microbiologie. Par vos qualités de grand pédagogue, vous m'avez offert un encadrement exceptionnel et un soutien indéfectible. Votre passion pour la recherche a été une source d'inspiration et un incitatif à l'excellence. À travers nos échanges édifiants et sans restriction d'horaire, vous avez aiguisé mon esprit critique. Bien plus, vous m'avez appris à rédiger des demandes de subvention dont trois auxquelles j'ai participé ont connu une suite heureuse. Je m'estime extrêmement chanceux de faire partie de votre équipe, et privilégié de bénéficier de votre savoir immense. Vous êtes véritablement un modèle pour moi, et vous le resterez. Trouvez ici l'expression de ma profonde reconnaissance.

Un grand merci à Prof. Gisèle LaPointe, ma codirectrice de thèse, pour ses encouragements et ses critiques constructives. Malgré la distance, vous avez fait montre d'une grande disponibilité. Nos rencontres bien que fortuites constituaient à chaque fois une occasion de discuter de mon projet. Vos leçons de rédaction scientifique en langue anglaise me seront toujours utiles, et je vous en remercie.

Merci à Prof. Yvan Chouinard pour tout le plaisir que vous avez eu à m'initier aux techniques d'échantillonnage des fourrages préservés et du lait cru dans les fermes laitières. Je remercie également Monsieur Robert Berthiaume dont l'appui au recrutement des fermes a été essentiel.

Je remercie grandement Prof. Jean-Christophe Vuillemard pour ses encouragements et la prélecture de cette thèse.

Merci à Thibault Varin, « mon Thibault », qui m'a montré les bases de l'analyse bio-informatique, et avec qui les discussions ont toujours été fructueuses. Ta générosité a accéléré mon apprentissage. Merci beaucoup.

Merci à Marie Verheyde, Halimatou Diallo et Myriam Laberge pour leur soutien technique. Merci également à Patricia Savard, Annalisse Bertsch, Nassim Mecheri, Camille Cavestri, Maria-Carolina Rodriguez, Elena Pulido et Joseph Lupien-Meilleur pour l'ambiance chaleureuse qui a régné au laboratoire tout au long de

ce projet. Plusieurs d'entre vous ont sacrifié de leur temps pour garder mes filles afin que je puisse faire avancer mon doctorat. Je vous en suis très reconnaissant.

Je suis également reconnaissant envers Mme Guylaine Maheu (Directrice) du CPE Poids Plume et son équipe de « wonder women » pour leurs encouragements. Vos conseils m'ont été très utiles pour la conciliation famille-études.

Un grand merci à toi Miss Mery (Mérilie Gagnon), d'avoir partagé avec moi ce beau projet. Je n'aurais pas voulu une autre coéquipière que toi pour le mener à bien. Ton sens de l'organisation et ta motivation m'ont beaucoup simplifié la vie. Ce fut un réel plaisir de sillonner le Québec en ta compagnie pour la collecte d'échantillons. Je tiens à remercier tes parents et ta sœur de leur gentillesse. C'était toujours un plaisir de les rencontrer lors des voyages d'échantillonnage, et de leur parler, à l'occasion, de mes travaux de recherche.

Je remercie grandement ma famille pour son soutien constant. Merci à ma mère pour ses prières, mes parents adoptifs, Jules-Roger Kuiate et Pélagie Komguem au Cameroun, Guy Fokou et Véronique Tindo en Belgique, pour leurs encouragements et leur présence, ma belle-mère pour son appui, ma conjointe Rosine pour ses sacrifices, mes filles Ivana et Sarah-Danièle, mon fils Jules-Simon et mon filleul Ayel pour tout le bonheur qu'ils m'apportent. Merci à vous tous, mes amis, pour votre présence inconditionnelle. Chacun(e) de vous est un trésor pour moi.

Avant-propos

Cette thèse est constituée de quatre chapitres dont un a été publié et les trois autres en voie de soumission à des revues avec comité de lecture. Pour les trois manuscrits à publier, je serai le premier auteur. Tous les chapitres sont rédigés en anglais. Au début de chaque chapitre, un résumé en français est proposé. L'introduction et la conclusion de la thèse sont rédigées en français. Une section regroupant toutes les références bibliographiques des documents et sites web consultés et cités est présentée après la conclusion de la thèse. Les travaux constituant cette thèse correspondent à environ 50 % de tout le travail qui a été réalisé dans le cadre de ce doctorat. En effet, l'analyse du mycobiote des ensilages et du lait cru par metataxonomique ciblant les régions ITS1 et ITS2 du gène de l'ARNr des levures et moisissures, l'analyse du microbiote de l'ensilage et du lait cru par métagénomique shotgun, et l'évaluation par métataxonomique ciblant la région V3-V4 du gène de l'ARNr 16S de l'impact des litières sur le microbiote du lait cru font partie des travaux dont les résultats ne figurent pas dans cette thèse.

Après la mise en contexte de la problématique et du but de cette thèse, le chapitre 1, présenté sous forme d'article, fait office de la revue de littérature. Intitulée « Farm management practices: potential microbial sources that determine the microbiota of raw bovine milk », cette section résume l'état de l'art des nouvelles connaissances sur les pratiques de gestion de la ferme laitière, notamment la gestion des fourrages et de la litière, et leur impact sur la qualité microbiologique du lait cru. Elle s'appuie particulièrement sur les avancées technologiques récentes en faveur des sciences dites « omics » et de l'identification des bactéries pour proposer de nouvelles approches intégratives d'analyses qui permettront d'approfondir nos connaissances sur l'écologie microbienne des systèmes laitiers. Ce chapitre fait suite à une invitation de l'éditeur du « World Journal of Microbiology and Biotechnology ». L'article sera publié en tant que « invited review » dans la revue ci-dessus mentionnée avec comme coauteurs Mérilie Gagnon, Ph.D., Prof. Gisèle LaPointe, Prof. Yvan Chouinard et Prof. Denis Roy. Ce chapitre se termine par l'énoncé des hypothèses et des objectifs de la thèse.

Le chapitre 2 rapporte l'essentiel de la mise au point des méthodes de conservation d'échantillons de lait cru, d'extraction d'ADN et de séquençage du gène codant pour l'ARNr 16S effectuée préalablement à l'échantillonnage des fermes laitières. Il a fait l'objet d'une publication dans la revue « Microorganisms » (manuscrit soumis le 7 février 2020, accepté le 3 mars 2020, et publié le 5 mars 2020), sous le titre « Optimization of preservation methods allows deeper insights into changes of raw milk microbiota ». La version intégrée dans cette thèse est identique à celle publiée. Pour ce manuscrit dont je suis l'auteur principal, le Dr Thibault Varin a été d'une grande assistance à la prise en main des outils de bioinformatique. Les Profs. Gisèle LaPointe, Simon Dufour, et Denis Roy sont les coauteurs de cet article.

Au chapitre 3, une analyse exhaustive des communautés microbiennes et des interactions entre elles et avec les caractéristiques physicochimiques de cinq types de fourrages est proposée. Ce chapitre intitulé « Metataxonomic insights into the microbial ecology of grass or legume and corn silage produced with and without inoculants in commercial dairy farms » est présenté sous forme d'article et sera soumis pour publication dans la revue « Applied and Environmental Microbiology ».

L'étude comparative du microbiote d'ensilages et de lait cru fait l'objet du chapitre 4 dont le titre est : « Association between forage combination, use of inoculant and raw milk microbiota on dairy farms in Quebec ». Ce chapitre est rédigé sous forme d'article et sera soumis pour publication dans la revue « Applied and Environmental Microbiology ».

Pour les chapitres 3 et 4, la collecte d'échantillons a été réalisée par Dre Mérilie Gagnon et moi-même. Les analyses bioinformatiques et statistiques, la rédaction du manuscrit et son édition ont été effectuées par moi-même. Les coauteurs sont Mérilie Gagnon, Ph.D., Thibault Varin, Ph.D., Prof. Yvan Chouinard, Prof. Gisèle LaPointe et Prof. Denis Roy.

Introduction

Le lait est un composant alimentaire essentiel pour environ 6 milliards de personnes à travers le monde [1]. Au Canada, la production laitière estimée à 76,7 millions d'hectolitres en 2010 est passée à 93,2 millions d'hectolitres en 2020, la majeure partie étant produite au Québec [2,3]. Pour l'industrie laitière, l'aptitude du lait à la transformation est primordiale. Compte tenu de sa richesse en nutriments et de son contenu élevé en eau, le lait est un milieu de culture par excellence pour divers microorganismes qui peuvent être utiles ou indésirables à la fabrication fromagère, ou encore constituer un risque pour la santé du consommateur. Par exemple, les pathogènes tels que *Listeria monocytogenes* [4,5], ou encore les agents d'altération tels que les bactéries appartenant au genre *Pseudomonas* [6,7] peuvent contaminer le lait pendant la traite ou lors du transport vers le lieu de transformation. De plus, les lactobacilles [8] et les bactéries sporulantes telles que les clostridies et les bacilles [9] sont capables de résister au traitement thermique appliqué au cours de la transformation, et ainsi causer des dommages pendant l'affinage des fromages [10], ou réduire la durée de conservation des produits laitiers [11].

À la ferme, les facteurs qui déterminent la qualité microbiologique du lait cru comprennent le microbiote du pis [12–14], la localisation géographique de la ferme [14,15], les saisons [6,16], la période de lactation des vaches [17,18], le sol [12], les fèces [12,13], la qualité de l'air [14], la litière [12,19] et l'alimentation par les pâturages [12,20], les concentrés [21] et les fourrages conservés [12,22]. Le foin et les ensilages constituent les principales formes de conservation des fourrages destinés à l'alimentation des vaches laitières [23–25]. Contrairement au foin qui est obtenu par séchage, les ensilages sont obtenus à la suite d'une fermentation en anaérobie des fourrages frais ou préfanés. Le processus d'ensilage exploite les capacités fermentatives des bactéries lactiques (LAB) épiphytes ou ajoutées comme inoculants pour induire une diminution rapide du pH, limiter la croissance des microorganismes d'altération et ainsi permettre la conservation des fourrages à base de graminées ou de légumineuses. Les LAB utilisées comme inoculants incluent des espèces homofermentaires et des espèces hétérofermentaires facultatives telles que *Lactiplantibacillus plantarum*, ou les hétérofermentaires obligatoires telles que *Lentilactobacillus buchneri*, les formulations commerciales pouvant contenir une combinaison des deux catégories [26]. Les progrès récents en technologie de séquençage à haut-débit ont permis de mieux comprendre la dynamique microbienne lors du processus d'ensilage et le microbiote final des ensilages. Cependant, très peu d'études ont analysé les communautés bactériennes viables des ensilages matures préparés et conservés dans des silos à la ferme.

Les ensilages constituent des sources importantes de bacilles sporulés qui peuvent être transférées au lait. Vissers et al. [27,28] ont associé la concentration élevée des sporulés dans les ensilages à une augmentation de la concentration des spores dans le lait cru. En effet, ces auteurs suggèrent que les spores

bactériennes provenant du sol contaminent les fourrages au champ et lors de la récolte, survivent au processus d'ensilage au cours duquel elles peuvent proliférer, puis parviennent au lait cru via les fèces ou le contact de ces derniers avec le pis. Si le mode de contamination du lait par les spores bactériennes provenant des ensilages est bien connu, il y a plus d'hypothèses que d'évidences sur le transfert des bactéries non sporulantes, notamment les lactobacilles et les entérobactéries, des ensilages vers le lait cru. Les informations concernant l'impact des pratiques de gestion des ensilages telles que l'usage ou non des inoculants sur le microbiote du lait et son aptitude à la transformation sont fragmentaires. L'étude plus approfondie de l'écologie microbienne des ensilages et du lait par l'analyse métataxonomique permettrait de mieux comprendre la structure phylogénétique des communautés bactériennes des ensilages et du lait cru, et ainsi d'identifier les bactéries du lait qui proviendraient des ensilages. C'est dans cette optique que se situe l'objectif principal de cette thèse qui est de déterminer l'impact de l'alimentation des vaches avec les ensilages inoculés ou non inoculés sur la structure des communautés microbiennes du lait destiné à la fabrication fromagère à travers l'étude détaillée de la prévalence et de la diversité des espèces microbiennes présentes dans les deux écosystèmes. Ces informations utiles permettront aux producteurs et transformateurs laitiers d'améliorer des pratiques de contrôle à la ferme en vue de réduire la prévalence des microorganismes qui affectent la qualité du lait et celle des produits laitiers.

Chapitre 1 - Revue de la littérature

Pratiques de gestion à la ferme : influence des sources de contamination sur le microbiote du lait cru de vache

Farm management practices: potential microbial sources that determine the microbiota of raw bovine milk

Alexandre J. K. Ouamba [1,2], Mérilie Gagnon [1,2], Gisèle LaPointe [2,3], P. Yvan Chouinard [2,4], Denis Roy [1,2]

[1]Département des sciences des aliments, Laboratoire de génomique microbienne, Université Laval, 2440 boulevard Hochelaga, Québec G1V 0A6 Canada

[2]Regroupement de recherche pour un lait de qualité optimale (Op+Lait), 3200 rue Sicotte, Saint-Hyacinthe J2S 2M2 Canada

[3]Department of Food Science, University of Guelph, 50 Stone Road E, Guelph N1G 2W1 Canada

[4]Département des sciences animales, Université Laval, 2425 rue de l'Agriculture, Québec G1V 0A6 Canada

1.1 Résumé

Outre les routines de traite, les études récentes ont montré que les systèmes de traite et les conditions d'entreposage influencent grandement le microbiote du lait cru. L'application de la culturomique, une technique d'analyse microbienne à haut-débit basée sur la culture bactérienne a révolutionné notre compréhension des bactéries associées à l'homme. Son extension en microbiologie laitière, ainsi que sa combinaison avec la métataxomique et la métabolomique ouvriront de nouvelles avenues pour contrôler la santé des bovins, pour améliorer les pratiques de gestion et ultimement la qualité du lait. Cette revue résume les résultats récents concernant l'impact des pratiques de gestion sur la diversité des microorganismes qui déterminent la qualité du lait cru de vache. Les derniers développements dans l'intégration des méthodes dites « omics » pour étudier le microbiote de lait cru depuis l'introduction au début des années 90 des méthodes à base de PCR en microbiologie laitière sont soulignés.

1.2 Abstract

Environmental and herd associated factors such as geographical location, climatic conditions, forage types, bedding types, soil, animal genetics, herd size, housing, lactation stage, and udder health are exploited by farmers to dictate specific management strategies that ensure dairy farm profitability and enhance the sustainability of milk production. Along with milking routines, milking systems, and storage conditions, these farming practices greatly influence the microbiota of raw milk as evidenced by several recent studies. During the

past few years, the increased interest in high-throughput sequencing technologies to investigate dairy microbial ecology has improved our understanding of raw milk community dynamics throughout storage and processing. However, knowledge of niche specific community assembly in the farm environment, and on the factors that determine bacteria transfer to the raw milk is still lacking. In addition, species or strain level identification of bacteria using high-throughput sequencing of 16S rRNA gene targeted regions is challenging. Recently, the development and application of culturomics, a high-throughput culture-dependent microbial analysis technique, has revolutionized the understanding of human associated bacteria and extending this technique to dairy microbiology, along with its combination with metataxonomics and metabolomics will open new avenues for controlling cattle health (mastitis), and improving farming practices and ultimately milk quality. This review summarizes the recent findings regarding the impact of farm management practices on the diversity of bacterial species that determine the microbiological quality of raw cow milk. Latest developments in the complementarity of omics methods for investigating the raw milk microbiota since the introduction in the early 1990s of PCR-based methods in dairy microbiology are emphasized.

1.3 Introduction

Bovine milk, a nutritious food commonly consumed worldwide, serves as the raw material for manufacturing a wide range of food products. However, the high nutrient and water content of milk makes it a prime habitat for beneficial, pathogenic, and spoilage bacteria [29–31] that can significantly impact its quality and safety, as well as deteriorate the shelf life of dairy products. The main microorganisms involved in the premature spoilage of dairy products have been extensively reviewed [11]. Psychrotrophic bacteria such as *Pseudomonas* spp. and spore-forming thermoduric bacteria such as *Bacillus* spp. and *Paenibacillus* spp., have been pointed out as the main microbial causes of food loss and waste in the dairy chain [11]. Moreover, it is well known that the abundance of psychrotrophic bacteria increases in raw milk during cold storage [7,18], and that the occurrence of spore-forming bacteria in milk can cause cheese defects [10]. The microbiological quality of raw milk is therefore a key determinant of the quality of processed dairy products such as pasteurized milk, cheese, yogurt, or powdered milk [11,18,32,33].

From the dairy farm to silos at processing facilities, the factors that drive the microbiological quality of raw milk include animal health, farm environment and management practices, and storage and transportation conditions. During the past two decades, several studies provided new insights into our understanding of the occurrence and impact of the raw milk microbiota on dairy products. It was clearly demonstrated that the microbial diversity of raw milk after milking varies in time and between farms [6,18]. Since the pioneering studies on the link between cow teat skin bacteria and the microbiota of raw milk [34–36], contrasting microbial community compositions of the cow teat skin have been reported in the literature [37]. Beyond the biases induced by the variety of microbial analysis techniques used to investigate the microbiota, variation in farming practices

are an important cause of the observed differences in the raw milk community assembly [38]. In this regard, studies by Vissers and colleagues on the contamination of bulk tank milk by *Bacillus cereus* spores [28] and butyric acid bacteria [27] were the first to advise the development of farming practices to improve the quality of raw milk. This review aims to highlight the latest findings on the interplay of farming practices and the safety and quality of raw bovine milk and explore the opportunities of knowledge advancement offered by integrated analytical approaches that combine culture-dependent and culture-independent techniques to enhance our understanding of the dairy microbiology.

1.4 Recent advancements in microbiota analysis methods

The total bacterial count expressed as colony-forming unit or copy numbers per millilitre is generally recognised as an indicator of the quality of raw milk and has been found to correlate with problems in management practices before, during, and after milking [39]. However, it was demonstrated that although viable counts could be similar for milk samples from the same farm or from different farms, bacterial community composition and phylogeny might vary considerably [15,19,40]. Since variation in species composition implies potential variation in the metabolic pathways, proper identification of microorganisms that contaminate raw milk on the farm, either by culture-dependent or -independent approaches, or by both, is crucial to understand the relevance of farming practices to the patterns of microbial occurrence.

1.4.1 Culturomics: a high-throughput like culture-dependent method

Culture-based microbial analysis consists of growing bacteria on culture media under controlled temperature conditions, and isolating single colonies formed for identification and characterisation purposes. Culture-dependent methods have long been applied to investigate the microbiota of raw cow milk (Table 1-1). These techniques that do not generally require highly qualified knowledge allow easy comparison of results from different laboratories. However, bacteria in a viable but not culturable (VBNC) state, the choice of culture media, and growth conditions are among the main factors likely to induce biases in this microbiological analysis approach [41]. For example, Gagnon et al. [22] have demonstrated that the RMS culture medium was not strictly specific to lactic acid bacteria (LAB), as 22 % of isolates identified by partial 16S rRNA gene sequencing were not assigned to the order *Lactobacillales*. Likewise, Mallet et al. [42] found that the total mesophilic bacterial count using the standard plate count agar corresponded to that of Gram-negative bacteria in raw milk. Although chemical compounds can be added to improve the specificity of the culture medium (e.g. vancomycin for LAB count), bacteria may exhibit differential sensitivity at the species or strain level. Moreover, anaerobic bacteria from raw milk are likely to enter a VBNC state in an oxygen-stressed environment. Sample storage at low temperature can affect the viability of raw milk bacteria [43]. Isolation of bacteria from plates by random selection

of single colonies as generally performed does not lead to a representative population of the microbiota, except if a tool such as the Harrison disk is used [19,22].

As mentioned above, each bacterial isolate is associated with a gene pool that determines the relevance of the metabolic activities of the identified species in the dairy chain. Culture-dependent techniques allow targeting specific bacterial groups that can be overlooked by culture-independent approaches despite their potential biological importance for milk and dairy products [21,40]. Microbial culturomics denotes the combination of extensive bacterial culture on a variety of culture media in different conditions with matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry for high-throughput isolate identification [44]. The purpose of applying this approach is to culture and identify unknown bacteria from their natural habitat [45]. Although it was introduced just a decade ago, this technique has been improved with newly developed culture media [45,46] such that recently, Diakite et al. [47] were able to cover 67 % of the reads obtained by metagenomic analysis and to reduce by 22 % the associated unidentified taxa. Most of the studies in which microbial culturomics was implemented as described in the original reports are associated with the human microbiota. The significant contributions of this methodological approach to the improved knowledge of the human microbiome have been extensively reviewed, and it is reported that the repertoire of cultured species isolated from human body sites has been increased by 28 % [48,49]. In contrast, only few studies applying microbial culturomics have been carried out on samples associated with the dairy farm environment [50,51], and nowadays, to our knowledge, none dealing with the udder [52] nor with the raw milk microbiota are available. Combining culturomics with high-throughput sequencing of the 16S rRNA gene pool, Zehavi et al. [50] found that the culturable bacterial species represented only 23 % of the total operational taxonomic units (OTUs) detected from the rumen. Considering the multiple sources of milk contamination on farm, the optimization of microbial culturomics for the dairy environment is encouraged to uncover the uncultured portion of the dairy microbial contaminants. On the other hand, MALDI-TOF is a reliable tool for isolate identification that can be routinely applied in the laboratory. Indeed, extensive culturing of microorganisms as performed in a culturomics approach generates thousands of isolates. Unlike genotypic methods such as restriction fragment length polymorphism, random amplified polymorphic DNA, or multilocus sequence typing that are applied on single isolates for identification and characterisation, rapid identification of these isolates in a high-throughput like manner can be achieved using MALDI-TOF. Dobranić et al. [53] successfully identified enterococci isolates from raw milk using this method. Implementing a culturomics-like approach on raw milk, Frétin et al. [21] assigned isolates to the phylum *Deinococcus-Thermus* that was not detected by the metataxonomic analysis. In addition, out of the 34 genera identified by culture-dependent method, these authors found that 12 were not detected by high-throughput sequencing. Using a similar approach with a Fourier-transform infrared method instead of MALDI-TOF mass spectroscopy for isolate identification from bulk tank milk, Breitenweiser et al. [40] isolated bacterial colonies corresponding to 24 genera and 6 families (*Jonesiaceae, Mycobacteriacea,*

Propionibacteriacea, Tsukamurellaceae, Paenibacillaceae, Morganellaceae) that were not detected by a concurrent metataxonomic analysis.

Table 1-1. Abundance of bacterial groups in bulk tank raw milk according to culture-dependent or independent method used in each study reviewed

Number of herds x sampling period[1]	Geographical site	Method	Targeted bacterial group	Log cfu/mL	Reference
974 raw tanker milk	California, USA	qPCR	Aerobic mesophilic bacteria	3.15	[54]
71 raw silo milk	California, USA	qPCR	Aerobic mesophilic bacteria	4.2	[55]
24 herds x 3	Eastern Canada	Viable counts	Lactic acid bacteria	2.56	[22]
84 herds	Eastern Canada	Viable counts	Anaerobic spore-forming bacteria	0.35	[19]
			Aerobic mesophilic spore-forming bacteria	0.38	
			Aerobic thermophilic spore-forming bacteria	0.35	
70 herds x 3	Eastern Canada	Bactoscan	Total bacteria	4.26	[56]
33 herds	New York state, USA	Viable counts	Aerobic mesophillic spores	1.7	[57]
			Aerobic mesophilic bacteria	3.8	
9 herds (472 samples)	New York state, USA	Viable counts	Aerobic mesophilic bacteria	3.47	[58]
5 herds	New York state, USA	Viable counts	Aerobic mesophillic spores	0.3	[59]
			Aerobic thermophilic spores	0.3	
108 herds x 2	Basse-Normandie, France	Viable counts	Aerobic mesophilic bacteria	3.69	[42]
			Lactoccocci	3.85	
			Lactobacilli	4.21	
			Leuconostoc	3.11	
			Pseudomonas	2.72	
			Gram-negative bacteria	2.88	
			Yeasts	1.92	
			Moulds	0.41	
729 herds	Denmark, Germany, and Netherlands	Viable counts	Aerobic mesophilic bacteria	3.9	[60]
13 herds	Galicia, Spain	Flow cytometry	Total bacteria	4.15-4.70	[61][2]
105 herds x 52	Netherlands	Bactoscan	Total bacteria	3.9	[62]
24 herds	Netherlands	MPN	Butyric acid bacteria spores	-0.3	[27,28]
		Viable counts	*Bacillus cereus* spores	-1.74	

45 farms x3	Norway	Viable counts	Aerobic mesophilic bacteria	4.27	[15]
1 herd x 6	Nothern Italy	MPN	Anaerobic spore-forming bacteria	1.48	[63]
		Petrifilm	Aerobic mesophilic bacteria	3.7	
		Petrifilm	Coliforms	1.24	
		Viable counts	Lactic acid bacteria	3.52	
		Viable counts	Propionic bacteria	1.02	
23 herds	Nothern Italy	MPN	Anaerobic spore-forming bacteria	2.76	[64]
		Viable counts	Lactic acid bacteria	3.8	
		Viable counts	Propionic bacteria	2	
		Petrifilm	Aerobic mesophilic bacteria	4.09	
		Petrifilm	Coliforms	2.15	
2 herds x 2	Southern Germany	Viable counts	Aerobic mesophilic bacteria	4.5	[40]
3 raw milk vending machines x 10	Southern Italy	Viable counts	Aerobic mesophilic bacteria	5	[65]
			Lactic acid bacteria	4	
			Enterococci	2	
			Enterobacteriaceae	1.8	
			Total coliforms	2	
			Fecal coliforms	0.5	
			Pseudomonas	5	
			Coagulase-negative cocci	2	
			Yeasts	5	
1 herd	Croatia	Viable counts	Enterococci	<1-5	[53][2]
			Enterobacteria	<1-4.69	
			Escherichia coli	<1-3.3	
			Staphylococci	<1-4	
			Aerobic mesophilic bacteria	3-5.39	
			Psychrophilic bacteria	<2-5.17	
			Lactic acid bacteria	<2-4.77	
20 herds (150 samples)	Czech Republic	Viable counts	Aerobic mesophilic bacteria	4.18	[66]
			Psychrophilic lipolytic bacteria	2.8	
1 herd	Slovakia	Viable counts	Aerobic mesophilic bacteria	6.71	[67]
			Psychrophilic bacteria	6.62	
			Thermophilic bacteria	3.11	
3 herds x 5	Vologda district, Russia	Viable counts	Aerobic mesophilic bacteria	4.03	[68]
21 herds	Tunisia	Viable counts	Aerobic mesophilic bacteria	6	[69]
		Viable counts	Aerobic mesophilic spores	<1	

3 herds x 12	South-east Victoria, Australia	Viable counts	Aerobic mesophilic bacteria Psychrophilic bacteria Thermoduric psychrophilic bacteria	4-5 2.75-3.8 1.5-2	[70][2]

[1]: When the samples were not bulk tank raw milk, the nature of the sample is given.
[2]: Viable count range

1.4.2 Culture-independent methods

The application of culture-independent techniques in dairy microbiology brought new perspectives on the diversity of raw milk microbiota [71]. High-throughput sequencing has attracted much interest in dairy microbiology to the point that during the last decade, nearly 200 published studies have used this technique to analyse the microbiota of raw milk [37]. Metataxonomics, metagenomics, metatranscriptomics, metaproteomics, and metabolomics are the main high-throughput methods used for microbial community analysis and characterization, and their application in food microbiology has been extensively and critically reviewed [37,72,73]. Here we will mainly focus on new developments in the metataxonomic approaches and combinations with other omics techniques that show high potential for investigating microbial niches associated with raw milk production.

Marchesi and Ravel [74] defined metataxonomics as the high-throughput characterisation of microbial communities based on the amplification and sequencing of taxonomic marker genes, from which all the sequences generated are used to build a metataxonomic tree that illustrates amplicon phylogeny. The most used marker genes for metataxonomic studies include the 16S rRNA gene [75] for bacteria and the internal transcribed spacers (ITS1 and ITS2) for fungi (Hoggard et al 2018). This technique has been widely used in dairy microbiology research (Table 1-2). Regardless of the marker gene of interest, the experimental design of metataxonomic studies including sampling, sample handling and storage, separating dead from living cells, DNA extraction, DNA amplification, sequencing, raw sequence analysis, and subsequent downstream analysis are all subject to biases [37,41,43,76]. Good practices associated with each stage of the metataxonomic analysis process mentioned above have been comprehensively reviewed [76–78]. However, the development of a consensus approach for the 16S rRNA gene high-throughput sequencing is challenging [77]. Using a culture-dependent technique, O'Connell et al. [79] noted an increase in total and psychrotrophic bacteria counts in bulk tank raw milk stored at 4 and 6 °C for 5 days, whereas this was not observed in a concomitant study using a culture-independent approach [17]. Changes in the diversity of milk microbiota after cold storage were not noticed plausibly due to the extraction of total DNA regardless of cell viability. Regarding storage conditions, milk samples that cannot be processed (freezing cell pellets at -80 °C) within the 24h following sampling should preferably be treated with a preservative solution such as azidiol [43]. Depending on the goal of the study, it may be relevant to distinguish dead from live bacterial cells by treatment with a viability dye such as propidium

monoazide (PMA). PMA is a DNA intercalating dye that enters microbial cells with lowered cell membrane integrity, binding to DNA following the exposure to bright visible light at up to 460 nM, and preventing nucleic acid replication during PCR [76,80]. Although PMA has been widely used in combination with quantitative PCR [81–83], its application in high-throughput sequencing studies is limited [37,55].

DNA extraction is a critical stage of the culture-independent microbial community analysis process, particularly for complex food matrices such as bovine milk. In the raw milk matrix, fat globules and casein micelles are dispersed in an aqueous solution of minerals, lactose, and whey proteins. Raw milk genomic DNA is classically isolated from the casein pellet after the whey and fat fractions have been discarded [17,18,40,43]. Although milk fat is known to interfere with the nucleic acid extraction process [84,85], a few studies reported an improvement of qPCR sensitivity when the cream fraction is included in DNA isolation [86–88]. In addition, bacteria such as *Staphylococcus aureus*, *Limosilactobacillus reuteri*, and *Escherichia coli* were previously found to bind to milk fat globules [88–90]. Recently, Lima et al. [87] performed a metataxonomics analysis of the microbiota of raw whole milk from healthy and clinical mastitis cows, along with three milk fractions including fat, casein pellet and casein pellet combined with fat, and reported no significant difference in the mean relative abundance of mastitis pathogens across milk sample-template. These authors also found that 27 % of the families identified in healthy milk samples composed the core microbiota (group of taxa shared by all or most of samples) of milk fractions and accounted for 80 % relative abundance of the total bacterial community. Another recent study on the human milk microbiota recommended the use of whole milk for microbiome investigations, although the authors reported a 39 % reduction in the efficiency of DNA extraction in the presence of the fat fraction [91]. These findings suggest that the bacterial community profile obtained from casein pellets as commonly done in dairy microbiology may not be representative of the real microbiota in milk. However, none of these studies have contrasted their protocols with that involving the use of Ethylenediaminetetraacetic acid (EDTA) to clarify raw milk before pelleting and fat removal as previously described [7,43,92]. Indeed, the addition of EDTA to human breast milk was found to induce a disruption of fat globule membrane with subsequent release of membrane-bound protein and free fatty acids, and a pH reduction [93].

Besides the bacterial cell recovery strategy from raw milk, the success of DNA isolation for community profiling also relies on the method of cell lysis. It is recognised that cell lysis approaches integrating enzymatic, chaotropic, and mechanical lysis will improve DNA extraction from bacteria [17,18,40,43]. Compared with a culture-dependent approach, Breitenwieser et al. [40] could not detect by metataxonomics several taxa from the phylum *Actinobacteria*. Owing to a thick layer of peptidoglycans, cell lysis of *Actinobacteria* may be difficult. As for spore-forming bacteria that may occur in raw milk, DNA extraction should be optimised [94]. Efforts to develop standards of DNA isolation from raw bovine milk should be encouraged.

The most recent developments of high-throughput sequencing of the 16S rRNA gene have offered three strategies that include sequencing one to three adjacent hypervariable regions (short-length amplicon sequencing) using the Illumina platform [95], or the full-length gene either by generating circular consensus sequences (CCS) using the PacBio system [96] or by the construction of synthetic long-read sequences (SLRS) using the LoopSeq technology [97]. An improved version of the Divisive Amplicon Denoising Algorithm implemented in a software package (DADA2) for amplicon error correction and high resolution sample inference [98] was recently adapted to PacBio CCS [96] and LoopSeq SLRS [97] data to provide exact amplicon sequence variants (ASVs) with single-nucleotide resolution from the full-length 16S rRNA gene with a near-zero error rate. A recent comparison of the PacBio CCS and LoopSeq SLRS methods through a DADA2 processing pipeline showed that while both technologies exhibited concordant results, higher accuracy and lower per-base error rate were obtained with the LoopSeq system [97]. Although these technologies may allow accurate species and strain level identification of bacteria from complex community assemblies, they are currently less cost-effective than the Illumina amplicon sequencing strategy for investigating the microbiota diversity in a large-scale study, and this would likely limit their wide application for some time. Illumina partial 16S rRNA gene sequencing is the most common approach of microbial community profiling. It is well known that the choice of the target region (out of the nine available) on the 16S rRNA gene as well as the clustering method (OTUs or ASVs) of the resulting sequences prior to taxonomy inference will have considerable impact on the microbiota diversity and subsequent data interpretation [99,100], particularly for large datasets of complex microbial communities [101–103]. While OTU picking is based either on amplicon similarity level (arbitrarily fixed at >97 % sequence identity) with reference genomes, or on *de novo* clustering (highly similar sequences in a given dataset are clustered together), ASV inference relies on error modelling and reads similarity and abundance to distinguish sequence biological variants with single-nucleotide differences [98,103,104]. Recently, Abellan-Schneyder et al. [105] comprehensively analysed the effect of primers targeting the hypervariable regions V1-V2, V1-V3, V3-V4, V4, V4-V5, V6-V8, and V7-V9, taxonomic classification methods such as OTUs, zero-radius OTUs (zOTUs) [106], and ASV [98], databases for taxonomy inference such as GreenGenes [107], the Ribosomal Database Project (RDP) [108], the Silva database (Quast et al 2013), the genomic-based 16S rRNA Database, 'The All-Species Living Tree' [109], and bioinformatic processing parameters (amplicon truncation length) on the taxonomic composition of three mock communities and 33 human samples [110]. The authors recommended (for human faeces) the V3-V4 region of the 16S rRNA, Silva or RDP databases, ASVs and zOTU clustering methods, and the creation of a specific and complex mock community reflecting the microbial ecosystem of interest for reliability and comparability of results [110]. Although several studies have employed short-read amplicon sequencing to raw milk and dairy product microbial analyses [37,72], studies that have assessed hypervariable regions for dairy microbial profiling are scarce, and no specific variable regions have been recommended. It was reported that the V3-V4 region is the most used [37]. A recent study comparing the V3-V4 and V6-V8 regions

for analysing raw bovine milk microbiota composition showed the superiority of the former for discrimination between samples, while the latter exhibited higher coverage [43]. The ASV clustering method has already been used in dairy microbiology [15,43,111]. Efforts to improve the technical aspects of short-length amplicon sequencing and data analysis process of microbial habitats on dairy farms are highly encouraged [43,99].

The microbial environment on dairy farms is composed of several habitats where interconnections map out the routes of raw milk contamination. From the forage field to the bulk tank, these habitats include soil, crops, silage and other feeds, cow faeces, bedding, cow udder and teat, water, milking clusters, pipeline, bulk tank, and raw milk [52,112,113]. Although several studies have investigated the microbiota inherent to single habitats, few have compared them to determine relative contributions to raw milk contamination [12,14,113]. Using SourceTracker [114] on metataxonomic data from various microbial habitats on dairy farms, Doyle et al. [12] found that the most important sources of raw milk contamination were the teat surface, faeces, grass for outdoor herds, bedding for indoor herds, silage, and soil, respectively. Interestingly, a recent study implemented the same analysis scheme to compare the impact of rumen fluid, airborne dust, water, bedding, and faeces on raw milk microbiota from two dairy farms and found that airborne dust was the main contributor to the milk community on both farms, followed by faeces, bedding, and water in farm 1, and bedding, faeces, and rumen fluid in farm 2 [113]. Moreover, Du et al. [14] using a similar approach found that teat, teat liner, teat dip cup, and dairy hall air differentially impact the milk microbiota from two dairy farms, demonstrating that the contribution of on-farm microbial sources to raw milk contamination may considerably vary depending on the farm configuration. Noteworthy, from the three studies mentioned above, high levels of contribution to milk contamination from unknown sources were noticed, emphasizing the multiplicity of microbial niches on the farm environment, and the relevance of implementing good management practices to limit raw milk contamination on farm. However, the reliability of such studies depends on the efficiency of sequence clustering methods and the accuracy of taxonomy inference, the latter highly relying on the database chosen as mentioned earlier. Currently, there is a growing interest in developing ecosystem-specific databases for improved resolution and accuracy of taxonomic assignment [115–117]. DAIRYdb, a manually curated reference database containing more than 10,000 full-length 16S rRNA gene sequences from dairy products, was recently developed and found to improve the accuracy of taxonomic classification of sequences associated with the dairy environment [118]. The construction of DAIRYdb included sequences from cheese, milk, teat, starter, and whey. However, the performance of this promising database might be reduced when assessing the microbiota of other on-farm microbial habitats such as cow faeces, silage, or bedding, and this could somehow limit its application in comprehensive studies of farm-scale microbial ecology.

Species or strain level identification of dairy microorganisms is essential for understanding their relevance for milk quality, safety, and processing. Although valuable efforts are being made to achieve this goal

with short-length amplicon sequencing, advanced technologies for high-throughput sequencing of the full-length 16S rRNA or other informative genes constitute a promising avenue for metataxonomic studies [97]. However, a decrease in the associated costs is needed for better accessibility. Beyond amplicon sequence clustering and accurate taxonomic classification, subsequent function pathways or metabolic profile prediction can add meaningful interpretation of the potential microbial activities. This can be achieved with bioinformatic tools such as PICRUSt2 [119], Tax4Fun2 [120], Piphillin [121], and more recently IPCO [122], that are all based on the KEGG database [123]. No independent benchmark tests involving all the cited tools are currently available, so readers are referred to the original papers for more information. We recommend testing one or more of these tools with each dataset to make a choice. Unlike metataxonomics that is limited to marker genes, metagenomics characterizes the metagenome from which the taxonomic information at the species and strain levels and all the functional potential of the microbiota are directly available [74]. Excellent reviews on the lessons, methodological advances, challenges, applications, and prospects of microbial metagenomics have been recently published [124–127]. The density and complexity of the information generated by metagenomic studies have enabled the evolution of specific research areas, including metatranscriptomics for characterizing community genes that are expressed at a given time, metaproteomics for generating timepoint protein content, and metabolomics for analysing the metabolite profiles of microbial systems [74]. Applications of these omics methods in dairy microbiology have been comprehensively reviewed [72,73]. Despite its limitations, metataxonomics can be combined with other omics disciplines to deepen our understanding of ecological niches. A study by De Filippis et al. [128] integrated metataxonomics and metatranscriptomics to investigate bacterial community dynamics and associated gene expression from raw bovine milk to a ripened traditional Italian pasta-filata cheese. The authors first identified highly expressed genes during cheese production and ripening, and secondly used three temperature and relative humidity conditions to modulate cheese ripening. They revealed the essential role of non-starter lactic acid bacteria in cheese ripening and demonstrated that variations of ripening conditions can considerably reduce the ripening time without affecting cheese quality. Recently, Bellassi et al. [129] analysed the effects of feeding systems including hay and a mixed forage ration composed of hay and fresh grass on the milk microbial community and chemical composition using metataxonomics and metabolomics. They reported a higher discriminative power of the metabolomic approach compared to amplicon sequencing on milk samples from both feeding systems, indicating that milk samples may exhibit similar taxonomic profiles but different metabolic activities, combining the microbial and bovine metabolic differences. Interestingly, they were able to correlate specific taxa to key metabolic markers of the feeding types tested. Several other multi-omics studies have been employed to advance our knowledge on dairy ecosystems [18,73,130].

In specific hypothesis driven microbiome studies, multi-omics approaches for holistic analysis generate what is now commonly referred to as “big data” from which new knowledge can only be extracted by using appropriate analytical methods that integrate large amounts of multi-table data. Recent advances in multi-table

data integration encompassing important aspects such as methods, tools, application, interpretation, and future approaches have been extensively reviewed elsewhere [131–134]. Tools such as MetaboAnalyst for examining metabolome and other omics data [135], MixOmics that is focused on data exploration, dimension reduction and visualisation, biomarker discovery, and model prediction [136,137], and MaAslin2 that is based on general linear models [138] are promising for dairy research. For instance, a recent study by Afshari et al. [137] used DIABLO (Data Integration Analysis for Biomarker discovery using Latent cOmponents), a multi-omics integrative method, to discriminate between cheddar samples of different sensory quality [139]. The authors reported that higher relative abundance of *Streptococcus salivarius/thermophilus*, enriched amounts of proline, histidine, isoleucine and aspartic acid, and lower quantity of octadecanol and stearic acid distinguished high-quality mature cheddar from low-quality counterparts that exhibited higher relative abundance of *Lactococcus lactis* as determined by DIABLO. The authors concluded that integrated multi-omics analyses could complement the current sensory evaluation of cheese quality [139]. Systematic integration of multi-omics analyses to investigate microbial niches in the dairy farm environment will shed light on relationships among microorganisms and on their interplays with host (cow teat) or abiotic (metabolites, pH, temperature, anaerobia, organic matter, etc.) factors that characterise the living milieu (teat, udder, skin, rumen, faeces, silage, litter, milk, etc.). The outcomes of such studies may include identifying patterns of milk microbial contamination from various microbial niches, patterns of microbial relationships (co-occurrence, co-exclusion) and inter-omics interactions in specific microbial sources, biomarker discovery, prediction of the milk processability and the quality of dairy products, as well as the generation of new research hypotheses. This information would be useful for recommending new farming practices that enhance the quality and safety of raw milk and dairy products.

Table 1-2 Relative abundance of dominant taxa found in raw milk using high-throughput sequencing technology

Number of herds x sampling period[1]	Location	Sequencing platform (hypervariable region)	Dominant taxa	Abundance (%)	Reference
1 herd x 2	British Columbia, Canada	Illumina HiSeq 16S RNA sequencing (V3)	*Bacteroides* *Roseburia* *Bifidobacterium* *Parabacteroides* *Lachnospiraceae*	36 7 4 3 2.5	[13]
45 herds x 3	Norway	Illumina HiSeq 16S RNA sequencing (V3-V4)	*Pseudomonas* *Lactococcus* *Bacillus* *Streptococcus* *Clostridium XI* * *Acinetobacter* *Staphylococcus* *Facklamia*	26.58 12.03 11.3 6.42 5.54 3.44 2.77 2.68	[15]

			Corynebacterium	2.38	
			Paenibacillus	2.23	
974 raw tanker milk samples	California, USA	Illumina HiSeq 16S RNA sequencing (V4)	*Streptococcus*	6.51	[54]
			Clostridiales	6.33	
			Staphylococcus	5.45	
			Ruminococcaceae	4.35	
			Corynebacterium	3.7	
			Turicibacter	2.45	
			Peptostreptococcaceae	2.22	
			Lachnospiraceae	2.03	
2 herds	Southern Germany	Illumina MiSeq 16S RNA sequencing (V3-V4)	*Corynebacterium*	10	[40]
			Caryophanon	7.65	
			Staphylococcus	5.75	
			Streptococcus	5	
			Chryseobacterium	3.5	
			Aerococcus	3	
			Jeotgalicoccus	2.5	
			Pseudomonas	2.5	
			Brachybacterium	2	
			Clostridium	2	
			Facklamia	2	
			Janibacter	2	
1 herd	Ireland	Illumina MiSeq 16S RNA sequencing (V3-V4)	*Ruminococcaceae*	24	[12]
			Rikenellaceae	5.6	
			Lachnospiraceae	5.3	
			Bacteroidaceae	4.8	
			Prevotellaceae	4.4	
			Preptostreptococcaceae	4.4	
1 herd	France	Illumina MiSeq 16S RNA sequencing (V3-V4)	*Sphingomonas*	20.59	[21]
			Lactococcus lactis	11.6	
			Bifidobacterium crudilactis / psychraerophilum	10.7	
			Rhodococcus spp.	8.56	
			Bacillus	6.61	
			Lacticaseibacillus casei	6.61	
			Corynebacterium	5.56	
10 herds x 12	Shanghai, China	Illumina MiSeq 16S RNA sequencing (V3-V4)	*Pseudomonas*	19.6	[6]
			Bacillus	13.8	
			Lactococcus	11.7	
			Acinetobacter	10.2	
67 herds	Ireland	Illumina MiSeq 16S RNA sequencing (V3-V4)	*Pseudomonas*	6.6	[18]
			Acinetobacter	5.2	
			Lactococcus	4.7	
			Corynebacterium	4.2	
			Streptococcus	2.5	
1 dairy facility x 3	Oregon, USA	Illumina MiSeq 16S RNA sequencing (V4)	*Ruminococcaceae*	9.54	[111]
			Peptostreptococcaceae	9.33	
			Lactococcus	4.39	
			Methanobrevibacter	4.17	

			Facklamia	2.51	
			Clostridiales	2.33	
			Corynebacterium	2.18	
			Streptococcus	2.14	
			Escherichia	2.06	
			Solibacillus	2.05	
			Turicibacter	2.03	
			Arthrobacter	2.02	
			Staphylococcus	2.01	
2 dairy facilities	Hungary	Illumina NextSeq (metagenomics)	*Bacilli*	51	[140]
			Gammaproteobacteria	25	
			Actinobacteria	16	
			Flavobacteria	3.5	
			Alphaproteobacteria	2.5	
			Chlorobia	1	
			Clostridiales	0.5	
1 dairy facility (26 raw milk samples)	Mexico	PGM-Ion Torrent (V2, V3, V4, V6, V7, V8 and V9)	*Streptococcus*	23.12	[16]
			Lactococcus	17.53	
			Enterobacteriaceae	17.27	
			Aeromonadaceae	14.82	
			Lactobacillus	0.72	
1 herd and 6 dairy facilities	Ireland	Roche 454 16S RNA sequencing (V4)	*Lactococcus*	55	[141]
			Pseudomonas	26	
			Leuconostoc	6.25	
			Prevotella	5	
			Ruminococcus	2	
			Flavobacterium	2	

[1]: When not bulk tank raw milk, the nature of the sample is given.
*: *Peptostreptococcaceae*

1.5 Bacteria of interest from raw milk

Investigations of the raw milk microbiota have shown high complexity (Table 1-1 and 1-2) but somehow low variability, though different compositions of the core microbial community have been described. Upon comparing five studies that analysed the microbiota of bulk tank milk by high-throughput sequencing, Parente et al. [37] reported that up to 2,000 taxa were identified at the genus level, of which *Enterococcus*, *Facklamia*, *Lactobacillus*, *Lactococcus*, *Paeniclostridium*, *Pseudomonas*, *Psychrobacter*, *Rikenellaceae*, *Rombustia*, *Ruminococcaceae*, *Staphylococcus*, *Strenotrophomonas*, *Streptococcus*, and *Turicibacter* were among the 25 most abundant and prevalent. From almost 1000 tanker truck raw milk samples, Kable et al. [54] reported a core microbiota composed of 29 taxa of which the most abundant included *Streptococcus*, *Staphylococcus,* and unidentified *Clostridiales*. In a study involving 472 bulk tank milk samples collected from 19 herds, *Ruminococcaceae, Acinetobacter, Clostridiales, Bacteroidales, Pseudomonas, Staphylococcus, Lachnospiraceae, Corynebacterium, Planococcaceae, Bacillus, Thermoanaerobacterium*, and 5-7N15 composed the core microbiota [58]. It was recently stated that *Microbacterium, Pediococcus, Fusobacterium,*

Propionibacterium, Acinetobacter, Bifidobacterium, Pseudomonas, Staphylococcus, Streptococcus, Lachnospiraceae, Corynebacterium, Bacteroides, Enterococcus, Ruminococcaceae, Aerococcus, Jeotgalicoccus, Psychrobacter and *Enterobacter* were frequently associated with bovine milk [41]. These taxa commonly found in raw milk encompass cow associated (gut and skin), disease causing, spoilage, beneficial, and psychrotrophic bacteria. In this section only some of the bacteria taxa or phenotypic groups that have been recently mentioned as relevant for milk quality are emphasized. Readers are referred to the review by Quigley et al. [5] for broader information on the microbiota of raw milk.

1.5.1 Psychrotrophs

Psychrotrophic bacteria are among the most studied species of the raw milk microbiota, probably due to their spoilage potential and their enrichment under cold storage conditions as shown for *Pseudomonas* and *Acinetobacter* [15,54]. An increase of 5 to 8 log cfu/mL of *Pseudomonas* upon storing raw milk at 4 °C for 72 h has been reported [65]. Likewise, Vithanage et al. [70] attributed a significant shift of the microbial structure of raw milk stored at 4 °C for 4 days to the proliferation of psychrotrophs. The outgrowth of *Pseudomonas fluorescens* group and *Bacillus* spp. along with the emergence of *Aeromonas*, *Listeria*, and *Stenotrophomonas* occurred at the expense of LAB such as *Streptococcus*, *Enterococcus*, and *Staphylococcu*s, as well as *Enterobacteriaceae* such as *Hafnia*, *Rahnella*, *Klebsiella*, *Enterobacter*, and *Serratia*. Psychrotrophic bacteria such as *Pseudomonas* can produce pigments and heat-stable enzymes. For instance, 50 to 75 % of *Pseudomonas* and *Bacillus* isolates showed persistent proteolytic and lipolytic activity following heat treatment of raw milk at 142 °C for 4 sec [70]. A recent study by Reichler et al. [142] showed that species of the genus *Pseudomonas* still pose major threats to fluid milk processing.

1.5.2 Thermoduric spore-forming bacteria

Thermophilic, mesophilic and psychrotrophic bacteria that can survive pasteurisation and grow in pasteurised milk are called thermoduric, regardless of their capacity to produce spores [143]. Reducing their occurrence in raw milk is therefore crucial. Spore contamination is a major and recurrent problem for the production of powdered milk [144]. The control of butyric acid producing bacteria such as *Clostridium tyrobutyricum* is of great importance to minimize late blowing defects during cheese manufacture [10]. The genus *Bacillus* is dominant among thermoduric spore-forming bacteria, and *Bacillus licheniformis* is one of the most ubiquitous [19,69,144,145]. *Paenibacillus* is also recognised as thermoduric [11,15,146]. Less common bacteria such as *Methylonatrum, Cloacibacillus,* and *Planobacterium* were reported to show significant positive correlation with heat-resistant spore counts in raw milk [6]. Thermoduric bacterial spores may grow and reach 20,000 cfu/mL in approximatively 14 to 21 days following milk pasteurisation [11]. Moreover, thermoduric bacteria of the genera *Bacillus, Paenibacillus, Psychrobacillus*, and *Viridibacillus* have shown lipolytic and proteolytic activity [11,70], especially in extended shelf life products.

1.5.3 Lactic acid bacteria

Within the order *Lactobacillales*, *Lactococcus*, *Pediococcus*, *Enterococcus*, *Streptococcus*, and lactobacilli are the most commonly reported genera in raw milk microbiota [147]. Recently, Zheng et al. [148] proposed an important reclassification of the *Lactobacillus* genus into 25 genera of which 23 are novel, along with the amendment of the family *Lactobacillaceae* to contain all previous members of the families *Lactobacillaceae* and *Leuconostocaceae*. LAB are the most studied bacterial groups of the milk microbiota due to their technological relevance and health-promoting properties [32]. Indeed, among multiple applications, LAB can be used as inoculants for ensiling [26] or starter for cheese manufacture [149]. However, heterofermentative LAB species can be associated with dairy product defects [19,150]. Although LAB are generally considered as mesophilic, they can proliferate in raw milk during cold storage. It was previously found that 87 % of LAB isolates from raw milk were able to grow at 6 °C [70]. Moreover, Kable et al. [54] have demonstrated the enrichment of *Lactobacillales* in milk silos at dairy plants.

The genus *Enterococcus* may be somewhat controversial. While certain strains of the same species can be used as microbial additives in silage or as adjunct starter for cheese-making, others were identified as cow or human pathogens [151]. Moreover, members of this group have exhibited resistance to numerous antibiotics including vancomycin [152], as well as to heat-treatments and salt, which can constitute potential threats to the safety of humans consuming raw milk [19]. The most abundant enterococci in milk include *Enterococcus faecalis*, *Enterococcus faecium*, and *Enterococcus durans* [19,53]. It has been reported that their amount could reach 10^5 cfu/mL in raw milk [53].

1.5.4 Kocuria

The genus *Kocuria* (family *Micrococcaceae*) often occurs in raw milk and several cheese varieties such as Parmesan and raw milk cheese [43,140,153]. While few studies have focused on *Kocuria* as a bacterial group of interest, Gagnon et al. [19] reported that several isolates of the genus collected from bulk tank raw milk were heat-resistant. Similar to *Enterococcus*, *Kocuria* may be opportunistic pathogens [154] and therefore may pose a food safety risk for the consumption of milk and derived products. Strains of *Kocuria varians* have shown resistance to more than five antibiotics [155]. The use of the strain *K. varians* S157 as adjunct for experimental cheese-manufacture induced significant increase in the content of dimethyl disulfide and dimethyl trisulfide compared to other manufacturing conditions, contributing to the typicity of the cheese flavour [156]. However, it has been reported that the presence of such sulfuric compounds in excess may induce cheese defects [157]. A *Kocuria rhizophila* strain isolated from raw milk showed a strong proteolytic activity owing to a heat-stable protease capable of withstanding a 77 °C pasteurisation for 40 min [67].

Besides the pathogens mentioned above, raw milk may harbour several other pathogens such as *Bacillus cereus* that can cause emetic disease, or *Clostridium perfringens* that produces toxins [143]. For a comprehensive overview of raw milk pathogens and subsequent implications on the safety of dairy products, the readers are referred to the reviews by Oliveira et al. [158] and Boor et al. [159].

1.6 On-farm management practices and milk microbiota

Over the past few years, numerous studies have improved our knowledge on a variety of dairy farm management practices by concentrating on the factors (Table 1-3) that determine the microbiological quality of raw milk. In this section, recent advancements on feeding, bedding, milking methods, seasonality, and geographical location will be emphasized.

Table 1-3 Farming associated parameters with potential effects on the milk microbiota

Parameter	Reference
Air quality	[14]
Bedding	[4,12,13,19,56,57,160]
Faeces	[12,13]
Feeding (Concentrates in diet)	[21]
Feeding (Grazing)	[12,27,143]
Feeding (Silage)	[22,27,28,64,146,161]
Geographical site	[14,15,69]
Health status	[53,58]
Lactation stage	[17,18]
Milking system	[15,60–62,162]
Season	[6,16,28,54,55,61,69,70]
Soil	[12]
Teat microbiota	[12–14,21]
Teat preparation	[42,57,59,63,64,163,164]

1.6.1 Feeding

Besides hay and concentrates, silage is a key element for dairy cow feeding [165]. Researchers have extensively studied the microbial ecology of silage [26,166–169], but few studies have investigated the link between silage and the microbiota of raw milk. Evidence from the literature suggest that spores from soil contaminate forage plants and remain alive in silage [158]. Spores that survive the cow gastrointestinal tract are found in faeces, from where they may reach the teat skin and finally enter the milk during milking [146]. In a study by Doyle et al. [12], faeces and teat skin samples from cows housed indoors (fed with silage) or outdoors (fed by grazing) clustered separately from each other following principal coordinate analysis, thus indicating

different community structures. The authors also identified faecal material as a major source of contamination of raw milk. Ensiled forages, particularly corn silage, may harbour high amounts of bacterial spores [63]. Several studies have demonstrated that on dairy farms where cows were fed high-quality silage having low content in bacterial spores, low amounts of spores were found in raw milk [27,28,64]. However, for non-spore forming bacteria, little is known on the identity of those that can be transferred into raw milk during milking. Producing high-quality silage is thus critical to minimize the occurrence of spoilage bacteria in raw milk [170]. Chemical (acids, salts) and microbial (mostly LAB) additives have long been used for this purpose. Currently, silage microbiology research is concentrated on the generation of new additives or their combinations that ensure the production of high-quality silage with improved aerobic stability and enhanced animal performance and productivity [26,171,172]. Finding new microbial inoculants with high efficiency is achieved through a series of selection criteria encompassing strain isolation and characterization, safety evaluation, metabolic efficiency, antagonistic properties, fermentation performance, evaluation of aerobic stability, and assessment of animal performance [172]. Recently, the intensive assessment of the properties of a new combination inoculant containing strains of *Lentilactobacillus buchneri* NCIMB 40788 and *Lentilactobacillus hilgardii* CNCM-I-4785 showed promising results for ensiling whole crop corn [171,173,174]. However, although LAB are also used as inoculants or adjuncts for cheese-making, the screening process of silage microbial additives does not involve the assessment of their ability to interfere with the milk fermentation process or alter the quality of dairy products. *L. hilgardii* has been found among the most abundant non-starter LAB of white brined cheese [175]. Moreover, strains of *L. buchneri* and *L. plantarum* identified in different dairy products have been reported as histamine producers [176]. On the other hand, a recent study by Gagnon et al. [22] focused on variations in the LAB profile from raw milk according to forage types including either hay or grass/legume and corn silage prepared with and without inoculants for herd feeding. The authors found that although LAB composition significantly differed between forage types, similar LAB profiles were observed in the associated milk samples. Interestingly, using random amplified polymorphic DNA typing of LAB isolates, the authors found that only ~6 % of LAB occurring in raw milk might originate from forage types or commercial inoculants. However, heat-resistant LAB from silage may pose threats to milk processing. To prevent such issues, LAB strains not commonly found in raw milk and that are not heat-resistant could be selected as inoculants for ensiling, although silage was identified as a minor contributor to the milk microbiota on dairy farm [12,14].

Grazing is widely practiced around the world [177]. Besides providing cows with fresh forage, access to pastures changes several parameters in farm management, including among others bedding, air quality, and temperature. As reported by Doyle et al. [12], microbial communities in ryegrass and soil were more similar to that of bulk tank milk when cows were grazed, while the microbiota of silage showed a higher similarity level with that of the bulk tank milk when cows were fed indoors. Higher relative abundance of cow associated bacteria such as *Eremococcus*, *Alloiococcus*, *Trichococcus*, *Prevotella* and *Psychrobacter* were observed in the milk

produced by cows housed indoors. On the other hand, for grazing cows, higher relative abundance of environmental bacteria such as *Corynebacteriales*, *Pseudomonas*, *Acinetobacter* and *Lactococcus* were noted in raw milk. In the study by Vissers et al. [27], a reduction in the amount of butyric spores in milk was attributed to the removal of silage from the feeding ration. For high moisture soils with elevated content in spores, access to grazing can increase the number of spores in milk, due to direct contacts of cow teat with soil [143]. Recently, Frétin et al. [21] reported that extensive grazing (grassland only) and semi-extensive grazing (grassland and concentrates) had similar effects on the milk bacterial community assembly across two sampling periods. However, they observed significant differences in the teat microbial composition. Feeding cows with concentrates altered the microbiota of the rumen and gastrointestinal tract, therefore modifying that of faeces. As mentioned above, the faecal bacteria may contaminate the teat, inducing a community shift, but not sufficiently to affect the milk microbiota. Likewise, in another study in which cows had open access to a clover pasture that was separate from the barn which itself was separate from the milking room, Falardeau et al. [13] found that the microbiota of teat milk samples was compositionally similar to those from the pasture. Raw milk after 24 h exposure was similar to air, bulk and truck tanks, cheese plant, and Gruyère cheese. Finally, bulk tank milk samples showed similarity to cow associated microbiota.

1.6.2 Bedding

Although cows may spend much of their time lying down, enabling direct contact of the udder or the teat skin with bedding materials, bedding was identified as a minor source of raw milk contamination [12,13]. Cow bedding likely harbours high amounts of anaerobic spores (estimated using the most probable number method, ~4 log MPN/g), mainly from *C. tyrobutyricum*, *Clostridium butyricum*, *Clostridium beijerinckii*, and *Clostridium sporogenes* [63]. Recycled manure solid (RMS) is an ecological alternative to conventional bedding materials that has been attracting increasing interest from farmers. However, few studies have focused on microbial communities populating RMS, and conflicting results have been reported regarding the impact of this bedding type on the raw milk microbiota. In a cross-sectional study involving 125 farms on which lactating cows bedded either on sawdust (35 %), sand (33 %), or RMS (32 %), Bradley et al. [4] found that while higher bacterial counts were observed in RMS, there was no subsequent higher bacterial load in the associated milk compared to that of sawdust and sand. Conversely, in a more in-depth study on a lower number of farms, Robles et al. [56] investigated the influence of different bedding types including sand used on 12 farms, straw and other dry forage on 33 farms, wood products comprising shavings and sawdust on 17 farms, and RMS on 8 farms. The authors reported that the use of RMS as bedding resulted in significantly higher bacterial counts in the bulk tank milk. Gagnon et al. [19] reported, as have others before [4,57], that milk concentration in bacterial spores was not affected by the bedding material, including RMS. This contrasts with findings from other studies who provided evidence of a direct impact of bedding type including RMS on the raw milk content in spores [38,178]. However,

RMS associated milk samples harboured higher amounts of heat-resistant *Streptococcus* and *E. faecalis*, whereas straw associated milk samples showed higher abundance of heat-resistant *Kocuria* [19]. In a recent study, the use of a bedding composed of 100 kg ground limestone, 25 kg straw, and 15 kg RMS for three months showed a reduction in the incidence of subclinical mastitis measured through the microbiological examination of quarter milk samples, probably due to reduced prevalence of faecal coliforms and streptococci in the formulated bedding compared to conventional straw [160]. The scarce literature currently available on the link between bedding types including RMS and the microbiota of raw milk reports several procedures for making RMS, differences in sampling protocols and study experimental design, as well as differences in other factors related to farming practices [4,19,38,56,57]. Drawing objective conclusions based on the outcomes of these studies regarding the influence of RMS on raw milk and dairy product quality and safety is difficult. Additional research is therefore needed to provide new knowledge and sufficient proof-of-concept demonstrations of possible risks associated with the use of RMS on dairy farms. However, the links between the bedding material and cow cleanliness were evidenced [56,179], and larger stalls have been associated with lower bacterial loads in raw milk [56].

1.6.3 Milking

The milking environment provides ideal conditions for bacterial growth, as it is composed of several microbial niches that can directly or indirectly share their contents. The maintenance of proper hygiene in this area is therefore essential for controlling environmental pathogens and reducing milk contamination rates. As mentioned earlier, the amount of dirt on the udder or teat has been positively correlated with the level of raw milk contamination [57,64,163]. The degree of cleanliness of the milker's hands also impacts the microbiological quality of raw milk, and recommendations have been made for proper hand hygiene or wearing gloves [164]. Implementing a combination of interventions including training milking staff to efficient teat-end cleaning and a laundering approach using specific recipe of detergent and chlorine bleach for cleaning towels resulted in a 37 and 40 % decrease in the mesophilic and thermophilic contents, respectively, in bulk tank raw milk compared to that before the interventions [59]. However, a previous study by Doyle et al. [12] demonstrated that teat preparation before milking (including water wash, forestripping, disinfection and drying with paper towel) did not show a major impact on milk microbiota compared to untreated teats. According to the authors, this contrasting result might be explained by high standards of hygiene measures implemented in the experimental farm in which their study was carried out. Zucali et al. [64] showed that the implementation of milking practices including use of gloves, dry udder clean, forestripping, pre-dipping, and post-dipping on farms resulted in reduced milk contamination by anaerobic-spore forming bacteria. Pre-dipping has been associated with low abundance of aerobic mesophilic bacteria, *Lactococcus*, yeasts, and moulds, whereas post-dipping has been correlated with low levels of presumed cheese ripening bacteria and *Leuconostoc* [42]. Similarly, comparing methods of teat

preparation, Bava et al. [63] have demonstrated that the application of pre-dipping, wiping, forestripping, and post-dipping as a milking routine considerably reduced the amount of aerobic spore-forming bacteria, total aerobic bacteria, and propionibacteria in milk compared to the application of forestripping only, or the combination of forestripping with post-dipping. However, these treatments did not affect the occurrence of LAB and coliforms in raw milk. Based on the authors' observations, it seemed that implementing the complete milking routine of teat hygiene prevented the occurrence of *C. tyrobutyricum* in milk. Although several protocols of teat preparation are being implemented on dairy farms, particular attention should be paid to this critical phase of milk production to limit the bacterial load of milk.

The types of milking system and housing are important factors that shape the milking environment. Automatic milking systems (AMS) have gained in frequency of installation and it has been estimated that about 38,000 milking robots are currently operational around the world [180]. While high milking frequency achieved with AMS may possibly reduce the time bacteria stay in the teat, it may concomitantly favour bacterial access to the teat canal that stays open after milking [180]. Moreover, stagnant fluids such as residual milk in the milking robot may result in a higher amount of spoilage bacteria in raw milk [180]. In a recent study, Skeie et al. [15] collected the bulk tank raw milk from 45 dairy farms implementing three types of housing and milking systems including free roaming with parlor, free roaming with automatic milking, and stall with pipeline milking. They showed that milk samples from automated milking had higher total mesophilic bacteria count than the two other milking systems and attributed the observed difference to poor teat cleaning and the absence of teat drying with AMS. However, the wider range of viable counts reported from farms implementing a stall housing with pipeline milking system might reflect a higher variation in milking routines in conventional systems. Skeie et al. [15] also reported significant effects of the types of housing and milking systems on milk microbiota at the genus level, along with significant variations in the occurrence of *Pseudomonas*, *Lactococcus*, *Acinetobacter*, *Facklamia*, and *Psychrobacter* sequence variants. Comparing bulk tank raw milk samples from farms using AMS or conventional milking systems, Johansson et al. [162] found higher proteolytic activity that was not associated with enzymes (plasmin and plasminogen) naturally present in milk on AMS farms. Recently, a systematic review by Cogato et al. [181] examined the tendencies and gaps associated with AMS in industrial and scientific research. Although the authors might have not considered mining words associated with microbial ecology of raw milk in the text examined, they revealed that interests were more focused on animal welfare and productivity than on milk quality. They therefore highlighted the need for more research on cleaning operations and animal health to improve milk quality. This supports concerns about the microbiological quality of raw milk produced using AMS as discussed above.

Regardless of the milking system, biofilm formation on milking equipment is of major concern for animal and human health, as well as for milk quality. However, few studies have been devoted to biofilms formed on

the inner surfaces of the milking tools, and consequently their real impact on milk and dairy product quality is not well known. Using a combination of culture-dependent and -independent techniques, Weber et al. [182] recently investigated the living biomass associated with biofilms on the milking machines within two farms. They found that *Actinobacteria* could be among the primary colonizers of stainless-steel surfaces inside the milking equipment, and that the genera *Bacillus*, *Kocuria*, *Microbacterium*, *Staphylococcus*, *Acinetobacter*, *Chryseobacterium*, and *Pseudomonas* were frequently detected. Most of these bacteria were also found on ultrafiltration membranes of a laboratory-scale model system used for dairy fluid filtration [183].

1.6.4 Geographical location and seasonality

Variations in dairy microbiology research outcomes from the literature according to countries, regions and other types of geographical localisations seems apparent (Table 1-2). Geographical location has been commonly used as a categorical variable for farms to investigate and explain the differences in microbial occurrence and community diversity. For instance, Skeie et al. [15] have demonstrated a significant effect of geographical location on the bacterial composition of raw milk samples collected from 45 dairy farms that implemented different housing and milking system types and that were distributed in two geographical areas in Norway. The authors further reported the genera *Pseudomonas, Bacillus, Staphylococcus, Paenibacillus, Psychrobacter, Chrysobacterium, Aerococcus* and *Rhizobium* as differentially abundant across the two regions. They also found that the number of sequence variants of the genera *Pseudomonas, Lactococcus, Bacillus*, and *Paenibacillus* were significantly increased or lowered among the two areas. However, because both regions shared similar altitude and climatic conditions, the authors explained the observed regional variations using farm associated factors that were not evaluated in their study, such as farming practices and sources of contamination. Similarly, Kmiha et al. [69] analysed the occurrence of spore-forming bacteria in raw milk collected from six regions located in the north and northwest Tunisia, and found a regional effect on the incidence of heat-resistant spores capable of surviving a 100 °C treatment for 40 min. However, although the impact of farm location on milk quality was part of their research questions, the authors did not provide any information either on the characteristics of the different geographical locations, or on the on-farm management practices. In contrast, using a polyphasic approach, Fricker et al. [184] analysed 48 raw milk samples from farm and dairy tanks located in Germany, Austria, and Norway, and found that although both milk types exhibited distinct microbial communities, a significant variation according to the geographical region was not observed. Again, farm characteristics and management practices were not described. As recently confirmed by Nikoloudaki et al. [185], it is well known that raw milk microbial assembly results from numerous farming conditions, management practices, and sources of contamination. Geographical location describes a specific position on the earth and is generally associated with climatic and geographic parameters that can impact or define farming conditions. Technically, each farm is associated with unique geographic coordinates. As discussed above, the term

geographical location has referred to a number of types of places such as country or regions. From a microbiology viewpoint, particularly in dairy microbiology, using this term or its equivalents as a categorical variable rather than farming associated factors or conditions to describe microbial diversity of the farm environment may be misleading and less meaningful for data interpretation or knowledge discovery. A better usage of geographical location can be exemplified by the study by Du et al. [14], in which nine sources of milk contamination from the milking environment were analysed in two Chinese dairy farms using a polyphasic approach (culture-dependent and 16S rRNA gene sequencing) to determine the contribution of each microbial niche sampled to the raw milk microbiota. Interestingly, the authors contrasted the results obtained from the farm located in Shijing village (Henan province) with those from the other farms located in Fuyu village (Heilongjiang province), the first geographical location being about 2,000 km away from the second on a direct line. Using the SourceTracker tool on high-throughput data, they found that on the farm from Shijing, the teat liner was by far the main contributor (at ~60 %) to milk microbiota, followed by dairy hall air (~15 %) and teat (~10 %); while on the Fuyu farm, teat dip cup (~35 %), teat (~20 %), and cowshed air (~13 %) were the main sources of contamination.

Seasonality in the diversity of raw milk microbiota has been widely demonstrated [6,16,54,186]. In a longitudinal study involving 10 dairy farms, Li et al. [6] showed that milk microbial communities sampled throughout 12 months clustered into two groups, one comprising fall and spring, and the other summer and winter samples. Kable et al. [54] have estimated that seasonal variation could account for 5 % of the variation in raw tanker milk (sampled consecutively during fall, spring, and summer across two years) microbial composition and structure. Higher viable counts of raw milk mesophilic and thermoduric bacteria were obtained in summer compared to winter and spring [69,70], while for psychrotrophs, higher viable counts were obtained in winter [70]. Bacteria that correlated with warmer temperatures included *Bacillus thuringiensis*, *B. licheniformis*, *Bacillus pumilus*, *Bacillus subtilis*, and *Paenibacillus*, whereas *Pseudomonas*, *Acinetobacter*, *Psychrobacter*, *B. cereus*, *Bacillus weitenstephensis*, *Bacillus circulans*, *Actinobacteria*, and *Propionibacterium* were found associated with cooler temperatures [6,28,69,70]. Besides temperature, humidity was revealed as a key factor that determines the seasonal variation in microbial diversity. Accordingly, while the abundance of *Chitinophaga* and *Niastella* were associated with low humidity, higher temperature and humidity levels generally favoured the proliferation of most bacteria [6]. Ruvalcaba-Gomez et al. [16] analysed the microbial composition of raw milk samples across dry (from November to May) and rainy (from June to October) seasons and found higher abundance of *Streptococcaceae* and *Lactococcus* in the dry season, whereas *Aeromonadaceae* and *Acinetobacter* were more abundant in the rainy season. Although weather conditions including temperature and humidity seem to consistently modulate bacterial groups such as mesophilic or psychrotrophic bacteria, a core microbiota may exist for raw milk. Li et al. [6] reported that the genera *Acinetobacter* and *Pseudomonas* occurred in 112 raw bulk tank milk sampled over 12 months. In the study by Kable et al. [54], a wider core microbiota of 29 taxa

among which the more abundant included unidentified *Clostridiales*, *Peptostreptococcaceae*, and *Ruminococcaceae*, *Staphylococcus*, *Streptococcus*, *Turicibacter*, unidentified *Lachnospiraceae*, and *Corynebacterium* was identified in 899 raw milk tanker trucks sampled over fall, spring, and summer. However, there are still gaps in the knowledge of the core microbiota at the species or strain level.

1.7 Conclusion

Research outcomes reviewed above clearly show that the complexity of raw milk microbiota reflects the farm environment from which it originates. Current knowledge on the routes of raw milk contamination provides evidence of interplay of farming conditions and management practices and the quality of raw milk. Therefore, improving the quality and safety of raw milk and dairy products requires a better understanding of the multitude of ecological niches that make up the farm microbial environment. The latter can be achieved by employing integrated high-throughput approaches that target both the microbiota and the metabolome of the matrix. As for the RMS reviewed above, more in-depth investigation of microbial dynamics during processing and of the final microbiota is needed to fully assess the associated risks for herd health and milk safety from a human perspective.

Our current understanding of the pattern of raw milk contamination on farm is challenged by the limitations of species or strain level identification of bacteria using the affordable amplicon sequencing technology (as have done most studies in dairy microbiology). Pending the cost reduction of metagenomic (or metataxonomic based on full-length sequencing of the 16S rRNA gene pool) studies that allow finer resolution, the culturomics approach can be used to uncover and characterize the hidden microbiota (uncultured, undetected, or ambiguously classified microorganisms) associated with milk and sources of contamination. Silage exemplifies important sources of microorganisms with high fermentative potential. Promising inoculants for ensiling should be screened for their possible interference with milk processing. More research is needed to fill the gaps of bacteria interactions in silage as well as in other microbial sources on the farm. Future research should exploit network-based analysis techniques to decipher interconnections among these niches in relation to raw milk microbiota. This would likely generate relevant knowledge to recommend new intervention strategies or farming decisions to ensure the production of high-quality milk and dairy products.

1.8 Hypothesis and objectives

1.8.1 General hypothesis

The comparative analysis of the prevalence and diversity of microbial communities in preserved forages and raw milk provides information on the impact of forage management practices on cow's milk microbiota.

1.8.2 Principal objective

The main objective is to evaluate the impact of forage management practices on the raw cow's milk microbiota and to identify taxa from forage that can enter the raw milk.

1.8.3 Specific objectives

- To develop a method of preserving the microbiota of raw cow milk samples.
- To assess the diversity of viable microbial communities occurring in hay, grass/legume and corn silage from commercial dairy farms, and to evaluate changes in the silage microbiota driven by the use of inoculants.
- To investigate the impact of feeding dairy cows with dry or ensiled forage, whether inoculated or uninoculated, on raw milk microbiota.

Chapitre 2 - Optimisation des méthodes de préservation du microbiote du lait cru

Optimization of preservation methods allows deeper insights into changes of raw milk microbiota

Alexandre J. K. Ouamba [1,2], Gisèle LaPointe [2,3], Simon Dufour [2,4] and Denis Roy [1,2]

[1] Département des Sciences des aliments, Laboratoire de génomique microbienne, Université Laval, Quebec, QC G1V 0A6, Canada

[2] Regroupement FRQ-NT Op+lait, Saint-Hyacinthe, Quebec, QC J2S 2M2, Canada

[3] Department of Food Science, University of Guelph, 50 Stone Road E, Guelph, ON N1G 2W1, Canada

[4] Département de Pathologie et Microbiologie, Faculté de Médecine Vétérinaire, Université de Montréal, Saint-Hyacinthe, Quebec, QC J2S 2M2, Canada

2.1 Résumé

L'instabilité temporelle du microbiote du lait cru affecte la fiabilité des études de sa composition. La capacité de l'azidiol et du bronopol à maintenir intact le microbiote du lait cru conservé à 4 °C, ou du mélange azidiol et diméthyle sulfoxyde à -20 °C a été évaluée. Les aliquotes de 5, 10 et 30 jours post conservation ont été traités avec le propidium monazide, puis analysés par séquençage à haut-débit des régions V3-V4 et V6-V8 du gène codant pour l'ARNr 16S. Ces deux régions ne conduisent pas nécessairement au même résultat. Après 5 jours de stockage, le microbiote d'échantillons non préservés a présenté une diminution considérable de la diversité et de la structure microbiennes. Le mélange azidiol et DMSO a permis de mieux stabiliser le microbiote du lait cru, de même que l'azidiol utilisé seul pendant 10 jours. L'ajout d'un antifongique permettrait de prolonger la durée de conservation des échantillons.

2.2 Abstract

The temporal instability of raw milk microbiota drastically affects the reliability of microbiome studies. However, little is known about the microbial integrity in preserved samples. Raw cow milk samples were preserved with azidiol or bronopol and stored at 4 °C, or with dimethyl sulfoxide (DMSO) or a mixture of azidiol and DMSO and stored at -20 °C for up to 30 days. Aliquots of five, ten- and thirty-day post-storage were treated with propidium monoazide (PMA), then analysed by sequencing the 16S rRNA gene V3-V4 and V6-V8 regions. The V6-V8 gave a higher richness and lower diversity than the V3-V4 region. After five-day storage at 4 °C, the microbiota of unpreserved samples was characterized by a drastic decrease in diversity, and a significant shift in community structure. The treatment with azidiol and DMSO conferred the best community stabilization in

preserved raw milk. Interestingly, the azidiol treatment performed as well for up to ten days, thus appearing as a suitable alternative. However, neither azidiol nor bronopol could minimize fungal proliferation as revealed by PMA-qPCR assays. This study demonstrates the preservative ability of a mixture of azidiol and DMSO and provides deeper insights into the microbial changes occurring during the cold storage of preserved raw milk.

2.3 Introduction

The interest in culture-independent approaches to study the microbiological quality of dairy milk has continued to rise, especially with the advent of high-throughput -omics technologies in the field of dairy science and technology, including metatranscriptomics, metaproteomics, and metagenomics [73]. Closely related to the logistics of implementing farming practices across the dairy production chain, microorganisms found in bulk tank milk have been shown to accurately typify dairy farms and the various inherent sources of contamination [12,187]. Raw milk samples thus generally harbour diverse microbial communities of which most can easily develop due to high nutrient and water contents of the matrix. Whether they are spoilage microorganisms [31], disease causing [30], health promoting or of technological interest [29,188], microbial dynamics is still occurring in raw milk once refrigerated immediately after milking [7,189] and during transportation to dairy processing facilities [54]. This temporal instability of milk microbiota can considerably affect sample microbial integrity, especially in prolonged sampling or when the sample delivery time to the laboratory is delayed, as it frequently happens in large-scale studies. Keeping intact the microbial communities in collected samples till processing highly contributes to guarantee the quality, validity, and reliability of derived information. In this context, freezing biological samples at -20 °C or -80 °C immediately after collection has been generally accepted as the gold standard approach for preserving microbiome profile over time [190–193]. However, the implementation and efficient management of such low temperature conditions when sampling in remote locations are hard to achieve, particularly if relatively large volumes of liquids are collected. Alternatively, several preservative chemicals have been developed to prevent milk spoilage or preserve milk constituents, among which bronopol and azidiol are commonly used for milk chemical analysis [194].

The biocidal or bacteriostatic activity of bronopol (2-bromo-2-nitropropane-1,3-diol) has been shown to depend on the initial bacterial load in raw milk, unlike azidiol (a mixture of chloramphenicol and sodium azide) which exerts bacteriostatic activity [195]. Although these preservatives have a broad spectrum of activity against Gram-positive and Gram-negative bacteria, their impact on the abundance and diversity of raw milk microbiota is not well known. Previous investigation outcomes on their ability to preserve microbial populations based on culture-independent techniques, once only limited to a few species for DNA isolation and PCR assays [196], are now expanding to high-throughput sequencing of whole communities [189,197–199]. However, few of these studies dealt with dairy milk, and none of them, to our knowledge, analysed the temporal variations of viable microbial communities in refrigerated milk preserved with azidiol or bronopol, using high-throughput sequencing

technologies. Furthermore, if samples were immediately frozen after collection, there would be no assurance of temperature stability over the course of sampling and sample conveyance due to prevailing environmental conditions and transportation time to the laboratory. Additionally, freezing milk samples may affect the viability of some coliforms, psychrotrophic microorganisms or bacteria belonging to the genus Mycoplasma [200,201]. In a previous study on the cryopreservation of anaerobic ammonium-oxidizing bacteria, Heylen et al. [202] recommended the use of dimethyl sulfoxide (DMSO) over glycerol as cryopreservative agent to improve bacterial recovery. For viable microbiota profiling, the effectiveness of propidium monoazide (PMA) has been largely demonstrated, particularly in combination with quantitative PCR. The concept of viability-PCR was first introduced by Nogva et al. in 2003 [203] as they used ethidium monoazide in conjunction with PCR to specifically amplify DNA from viable bacterial cells (i.e., with non-compromised membrane integrity), and later by Nocker et al. in 2006 [80] for the same purpose, but with PMA instead, showing improved specificity [see reviews by Fittipaldi et al. [76] and Elizaquivel et al. [204] for more information on practical aspects and challenges associated with viability dye]. If azidiol- or bronopol-treated samples are suitable for culture-independent analytical methods, no information has been reported on the suitability of PMA for use concurrently. Hence, we hypothesized that a mixture of DMSO as cryopreservative and azidiol as preservative will permit a better stability of the milk microbial composition, diversity, and viability when samples are stored at -20 °C.

In order to get more insight into the ability of preservation methods to stabilize the microbiota in raw dairy milk, we analysed the temporal stability/instability of viable microbial communities in preserved milk samples collected from six dairy farms (Fig. 2-1). We first compared community compositions in fresh and unpreserved milk samples. We then evaluated four preservative substances including azidiol (AZ4), bronopol (BR4), and a mixture of azidiol and dimethyl sulfoxide (AZDm). Community profiles obtained after five, ten and thirty days of storage at 4 °C or -20 °C were compared to those at day zero and to fresh unpreserved samples. Prior to DNA extraction, samples were treated with PMA to account for cell viability. Viability-qPCR assays were performed on different bacterial groups as well as total fungi to complement community profiling derived from high-throughput sequencing of the 16S rRNA gene. Doing this, we also sought to evaluate whether 16S rRNA sequencing of the V3-V4 and V6-V8 regions provides consistent results.

2.4 Materials and Methods

2.4.1 Milk sampling

Cow milk samples were collected from six dairy farms in Quebec, Canada. At each farm, about 1500 mL of raw milk were collected in sterilized two-litre glass bottles directly from the bulk tank, immediately after milking, and milk agitation. Prior to milk collection, the bulk tank valve was abundantly washed with tap water

and approximately 100 mL of milk were discarded prior to sampling. Samples were then placed in a 4 °C electric cooler and transferred to the laboratory where they were processed the same day.

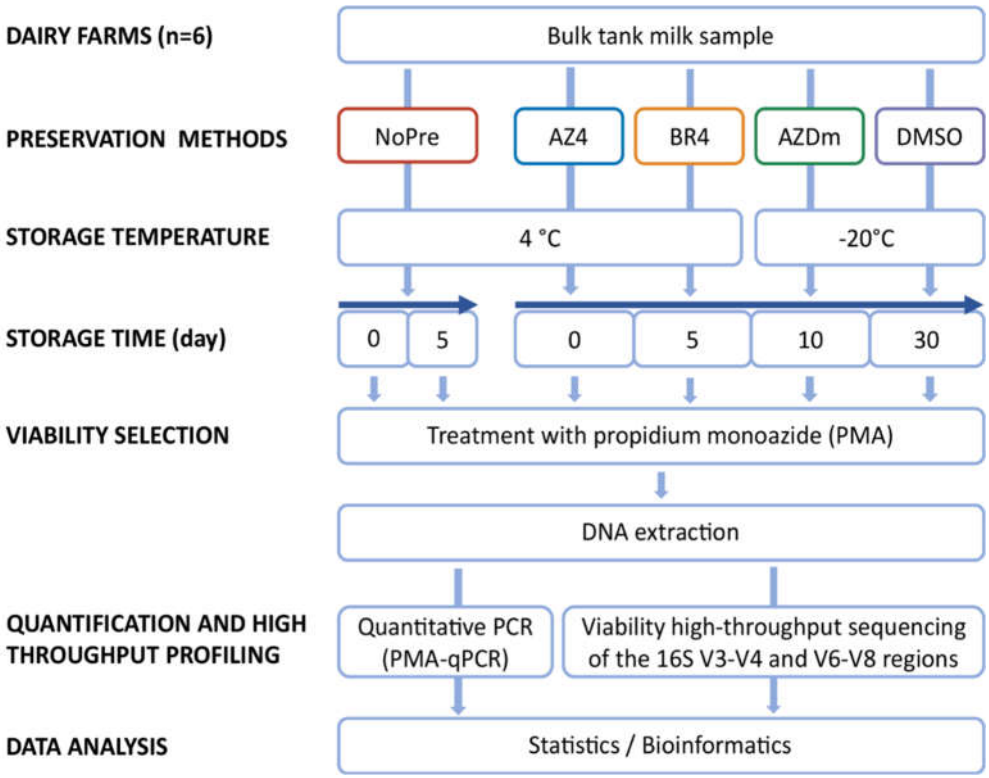


Figure 2-1: Experimental design of the study. Raw milk samples collected from six dairy farms were aliquoted and treated with four preservation methods, and then stored in a refrigerator or in a freezer according to the preservative used. Preservation include NoPre: No treatment, AZ4: Azidiol, BR4: Bronopol, AZDm: Combination of azidiol and dimethyl sulfoxide, and DMSO: dimethyl sulfoxide. Preservative-free samples were stored for five days, and preserved samples for up to 30 days. In addition to freshly treated aliquots, sub-samples were taken from the storage after five, 10 and 30 days, then treated with propidium monoazide and subjected to microbial analyses through quantitative PCR and 16S rRNA sequencing of the V3-V4 and V6-V8 regions. Data were analysed to evaluate the microbiota stabilizing effects of the preservation methods as depicted by the two hypervariable regions.

2.4.2 Preservation treatments and experimental design

Each milk sample was mixed by inversion and immediately divided into 52 aliquots of 30 mL to apply the different preservative methods. Treatments included preservative-free milk (NoPre), azidiol with storage at 4 °C (AZ4), bronopol with storage at 4 °C (BR4), dimethyl sulfoxide with storage at -20 °C (DMSO), and a mixture of azidiol and DMSO with storage at -20 °C (AZDm). Azidiol was prepared as described previously [196] by dissolving 4.5 g of trisodium citrate 5,5-hydrate (VWR International, Radnor, Pennsylvanie, USA), 1.8 g of sodium azide (VWR International), 0.075 g of chloramphenicol (MilliporeSigma, Burlington, Massachusetts, USA), 0.035 g of bromophenol blue (MilliporeSigma) and 10 mL of ethanol (Greenfield Global, Toronto, Canada)

in 100 mL of sterile distilled water. To obtain bronopol, a 5 % (weight/vol) solution of 2-bromo-2-nitro-1,3-propanediol (VWR International) was prepared with sterile distilled water. Treatment AZ4 was achieved by adding 3.33 µL of azidiol per mL of milk aliquot (for a final concentration of 0.06 mg/100 mL). For BR4 treatment, 4 µL of bronopol solution per mL of milk aliquot were added (for a final concentration of 20 mg/100 mL). Milk treated with DMSO contained dimethyl sulfoxide (VWR International) at 5 % [v/v] (22). Milk samples treated with AZDm were obtained by adding azidiol and DMSO to reach final concentrations of 0.06 mg/100 mL and 5 % respectively. Treated and preservative-free milk aliquots were vigorously shaken to homogenize the preparation. All the experiments were carried out aseptically on ice.

Treatments NoPre were processed at time zero and day 5 only, while all other treated aliquots were processed at time zero and kept at corresponding temperatures for 5, 10 and 30 days. All time zero treated aliquots were processed within an hour. Three replicate aliquots of each timepoint per treatment were prepared and randomly distributed in a 4 °C refrigerated incubator or a -20 °C freezer accordingly. Upon five days post-treatment and following timepoints, the three replicates per timepoint and per treatment were removed from the incubator or freezer, and, when available, the remaining aliquots were randomly displaced to account for temperature variations. Frozen aliquots were thawed for two hours in a refrigerator. Aliquots of the same treatment for a given timepoint were then pooled for microbial cell recovery and genomic DNA extraction.

2.4.3 DNA extraction and sequencing of the 16S rRNA gene

Treated and untreated pooled aliquots were fully homogenized, and 10 mL of each mixture were clarified with 2 mL of 500 mM ethylenediamine tetra-acetic acid [7] at pH 8.0 (Millipore Sigma) and centrifuged at 12 000 x *g* at 4 °C for 15 minutes. Cell pellets were then washed twice with 1 mL sucrose buffer (sucrose 12 % [w/v], 25 mM Tris-HCl pH 8.0). Cell pellets were treated with PMA using a lamp instrument (halogen 500 W, Ingenia Biosystems, Barcelona, Spain) as previously described [82]. PMA treated cells were kept at -80 °C until use. The protocol used for DNA extraction was adapted from that described by Quigley et al. [205] using a combination of chemical and mechanical lyses provided in the DNeasy PowerFood Microbial Kit (Qiagen, Hilden, Germany), and enzymatic lysis with lysozyme (MilliporeSigma), mutanolysin from *Streptomyces* (MilliporeSigma) and proteinase K (MilliporeSigma) with some modifications. Briefly, chemical and enzymatic lyses were achieved once by suspending cell pellets in 450 µL of warmed (heated at 37 °C instead of 55 °C as recommended for 10 min to prevent enzyme denaturation) solution MBL (provided in the extraction kit) containing 9 mg lysozyme and 50 U mutanolysin. The suspension was then incubated at 37 °C for 1 hour, followed by the addition of 25 µL of proteinase K (20 mg/mL) and another 1-hour incubation at 55 °C. Afterwards, mechanical lysis was performed using microbeads and a vortex adapter. This step and the remaining were all performed as described in the Qiagen instruction manual. DNA was quality checked with a NanoDrop

spectrophotometer ND 1000 (Thermo Fisher Scientific, Burlington, ON, Canada) and quantified with a Qubit fluorometer 2.0 (Invitrogen., Burlington, ON, Canada), then stored at -20 °C until use.

DNA samples were submitted to the *Plateforme d'Analyses Génomiques* of University Laval (Quebec, Canada) for library preparation and 16S rRNA gene sequencing on the Illumina MiSeq platform, targeting two non-overlapping hypervariable regions namely V3-V4 and V6-V8. Primer pairs 341F (5'-CCTACGGGNGGCWGCAG-3') / 805R (5'-GACTACHVGGGTATCTAATCC-3') and B969F (5'-ACGCGHNRAACCTTACC-3') / BA1406R (5'-ACGGGCRGTGWGTRCAA -3') were used to amplify the V3-V4 and V6-V8 regions respectively.

2.4.4 Quantitative PCR

Gene copy number representing biomass variation across time of total lactic acid bacteria, total acetic acid bacteria, *Enterobacteriaceae*, *Pseudomonas* spp., total bacteria and total fungi in samples from all preservation methods were assessed by quantitative PCR (qPCR) on a ViiA7 system (Thermo Fisher Scientific). Species or group specific primer sets used for amplification are listed in Table 2-1. All qPCR amplifications were performed in triplicate with mixtures (10 µL) comprised of 3.6 µL UltraPure DNAse and RNAse free distilled water (Invitrogen), 5 µL PowerUp SYBR Green master mix (Thermo Fisher Scientific), 0.2 µL of each forward and reverse primer at 10 nM, and 1 µL DNA template. For each primer set, running parameters were set up as described in the corresponding reference provided in Table 2-1. Standard curves were constructed by using ten-fold serial dilutions of known genomic DNA concentrations extracted from single cultures of microorganisms (Table 2-1). For a standard curve, the gene copy number was determined with the following formula:

$$\frac{[\mathbf{DNA}]^{\circ} \times \mathbf{AN}}{\mathbf{660} \times \mathbf{NT}}$$

where [DNA]° is the DNA concentration (g/L), "AN" the Avogadro number (6.02 x 10^{23} mol^{-1}), "660" the average molar weight of a base pair (g/mol/base pair), and "NT" the amplicon length (base pair). Sample microbial biomass expressed in gene copy number per millilitre milk was calculated based on each standard curve's equation.

Table 2-1: PCR primers used for species or group quantification.

Microorganisms	Primer name	Primer sequence (5'-3')	Gene	Amplicon size	Reference
Lactobacillus buchneri	LBF2 LBR1	GAAACAGGTGCTAATACCGTATAACAACCA CGCCTTGGTAGGCCGTTACCTTACCAACA	16S rRNA	129	[206]
Lactobacillus plantarum	LacPlan1F LacPlan1R	AGGCGCGGCTGATGTCA CGCGATTGTCTTGGTTTTGTT	recA	68	[207]
Lactic acid bacteria	WLAB1 WLAB2	TCCGGATTTATTGGGCGTAAAGCGA TCGAATTAAACCACATGCTCCA	16S rRNA	407	[208]
Acetic acid bacteria	AQ1F AQ2R	TCAAGTCCTCATGGCCCTTATG CGCCATTGTAGCACGTGTGTA	16S rRNA	55	
Enterobacteriaceae	rplP 1F rplP 185R	ATGTTACAACCAAAGCGTACA TTACCYTGACGCTTAACTGC	rplP	185	[209]
Pseudomonas spp	Pse435F Pse686R	ACTTTAAGTTGGGAGGAAGGG ACACAGGAAATTCCACCACCC	16S rRNA	251	[210]
Total bacteria	Uni334F Uni514R	ACTCCTACGGGAGGCAGCAGT ATTACCGCGGCTGCTGGC	16S rRNA	180	[211]
Total fungi	ITS1f 5.8s	TCCGTAGGTGAACCTGCGG CGCTGCGTTCTTCATCG	rRNA	300	[212]

2.4.5 16S rRNA gene bioinformatics and data analysis

Adapter and primer removal from demultiplexed sequences were performed using the Cutadapt (version 2.3) tool [213]. Reads corresponding to the V3-V4 and V6-V8 hypervariable regions were separately analysed following the DADA2 (version 1.12) pipeline [98] in R environment. For both datasets, taxonomy was assigned to the resulting amplicon sequence variants (ASVs) using the Silva version 132 DADA2-formatted reference databases. Sequence alignment was performed using the DECIPHER (version 2.12.0) package [214] and the phangorn (version 2.5.3) package [215] was used to construct the phylogenetic tree as previously described [216]. Processed sequences were then imported into the phyloseq (version 1.28.0) package [217] for downstream analyses of bacterial communities for which ASVs were identified at least at the phylum level and occurred in at least two of the six fresh milk samples (i.e., NoPre treatment at day zero) with greater than 0.001 % relative abundance.

We compared the V3-V4 and V6-V8 bacterial profiles by calculating corresponding Chao1, Shannon and Inverse Simpson indices on fresh milk samples using the phyloseq package. To visualize community composition and diversity variation between both hypervariable regions, heat trees (i.e. cladogram allowing the visualization of differential abundance and diversity of taxonomic data) were constructed using the Metacoder (version 0.3.2) R package [218]. A custom function was implemented in R in order to combine the two datasets.

Significant differences between taxa median proportions were calculated using a Wilcoxon rank sum test on a sub-dataset composed of the core microbiota identified for both regions. Only significant features obtained after FDR correction of p-values were considered. Heat trees were also constructed for taxa specific to each region. A Venn diagram comparing the number of taxa identified by both regions at the genus level was constructed using an online tool (http://bioinformatics.psb.ugent.be/webtools/Venn/).

To characterize microbial alteration over time in milk without preservatives, we compared community composition and structure of five-day stored milk with fresh samples. Besides the above-mentioned alpha-diversity measures, beta-diversity was assessed by plotting principal coordinate analysis (PCoA) with weighted and unweighted UniFrac distance metrics [219]. For alpha- and beta-diversities, a Wilcoxon rank sum test with FDR correction was performed using the ggpubr (version 0.2) R package [220]. Biomarkers distinguishing fresh from five-day stored raw milk samples were identified by applying the linear discriminant analysis effect size (LEfSe) algorithm [221] available in the Galaxy platform [222] with default parameters. Differential abundance of taxa between both conditions and data visualization were performed using a combination of R packages phyloseq, ggplot2 version 3.1.1 [223] and DESeq2 version 1.24.0 [224].

The effect of preservation chemicals on the microbiota upon contact with milk samples was evaluated by analysing variation in alpha- and beta-diversities before and immediately after preservatives were added to milk aliquots. Alpha-diversity measures including Chao1 and Shannon indices were computed. Differential abundance of taxa between fresh and preserved samples was calculated, and community composition, taxa prevalence and distribution among treatments were visualized using a heat map constructed with the R package ComplexHeatmap 2.1.0 [225].

The ability of preservation methods to maintain stable milk microbiota over time was assessed by linear mixed-effects (LME) modelling using the R package nlme 3.1-140 [226]. Computing variation of diversity metrics across time was adapted from a previously described approach [227]. Briefly, models were fitted based on log- or arcsine-transformed data for variance stabilization or normality prerequisite. The random effect for the intercept and the slope was defined by dairy farm. Computed alpha-diversity measures included Chao1, Shannon and Inverse Simpson indices. Beta-diversity trends and instability (i.e., distance variation within communities for consecutive timepoints) were evaluated using Jensen-Shannon divergence as well as weighted and unweighted UniFrac distances. We used the function emmeans() with FDR correction of p-values implemented in the R package emmeans 1.3.5 [228] for multiple comparison tests among storage timepoints. For beta-diversity trends, significance of the linear model was evaluated with a permutation test. PCoA was also performed on weighted UniFrac distance to estimate and visualize differences in microbial community structure

over time for each preservation method. Significance of differences between communities at the four timepoints was evaluated using a Kruskal Wallis test implemented in the R package ggpubr.

Intraclass correlation coefficients (ICC) with a one-way model was also used to assess microbial temporal stability at different taxonomic levels in preserved and non-preserved milk aliquots. We used the function icc() of the R package irr 0.84.1 [229] as previously described [230] to compute single score ICCs as an index of interrater reliability of preserved and non-preserved milk microbial community abundances over time. For each timepoint and preservation method, the calculated index quantifies abundance variability for a given taxonomic level among samples from different farms, taken as biological replicates. Intraclass correlation calculations were performed using default parameters. Their values range between 0 and 1. For ICC values less than 0.5, or between 0.5 and 0.75, or between 0.75 and 0.9, or above 0.9, they describe poor, moderate, good or excellent reliability, respectively [231].

Taxa differential abundances between consecutive and non-consecutive timepoints were computed using the DESeq2 algorithm and feature dynamics (i.e., for each preservation method, variations in abundance and prevalence across time of taxa that undergone at least once a log2-fold change) were visualized with a heatmap.

2.5 Results

2.5.1 High-throughput sequencing of the 16S rRNA hypervariable regions V3-V4 and V6-V8 reveals different taxonomic profiles

Sequence processing with the DADA2 pipeline resulted in 1,882,615 reads with an average of 17,432 per sample for the V3-V4 region, whereas for the V6-V8 region, 3,555,435 reads with an average of 32,921 per sample were obtained. From these reads, 712 and 977 ASVs were inferred for V3-V4 and V6-V8 regions, respectively.

We compared the taxonomic profiles resulting from 16S rRNA sequencing of V3-V4 and V6-V8 regions in unpreserved milk samples at day zero. Consistently with the number of ASVs, the estimated richness (Chao1 index) was lower for the data from the V3-V4 region compared to the profiles obtained from the V6-V8 region (Fig. 2-2A). However, microbial communities appeared more diverse and evenly distributed (higher Shannon and Inverse Simpson indices) with the V3-V4 than the V6-V8 region.

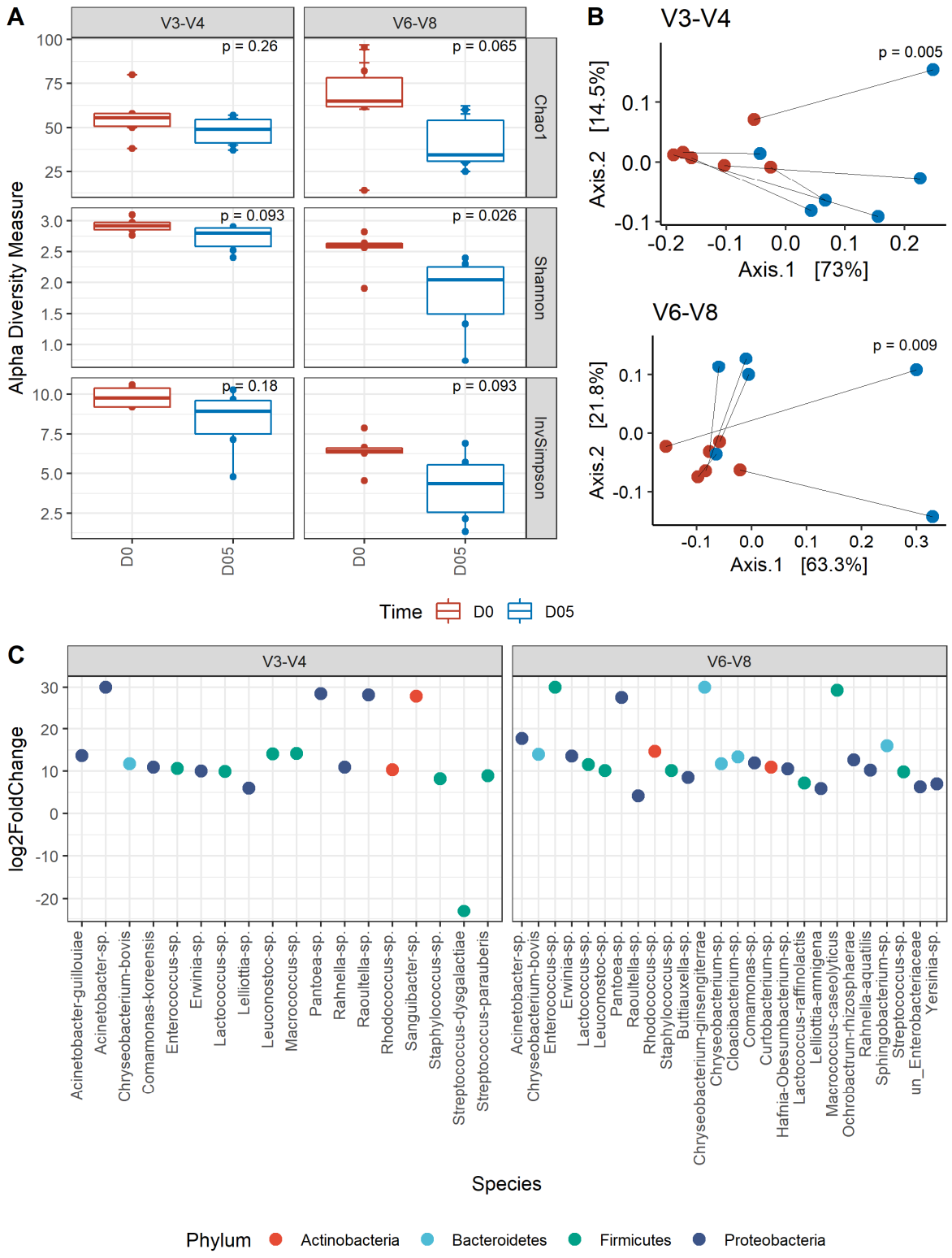


Figure 2-2: Changes in the microbial diversity of unpreserved raw milk samples. (A) Alpha-diversity measures of fresh (D0) and five-day stored (D05) raw milk based on V3-V4 (left) or V6-V8 (right) datasets. P-values indicate the significance of differences between timepoints obtained from the Wilcoxon rank sum test. (B) Principal

coordinate analysis on weighted UniFrac distances showing shifts in each unpreserved raw milk sample between fresh (red) and five-day stored (blue) based on V3-V4 (top) and V6-V8 (bottom) datasets. P-values indicate the significance of differences between timepoints obtained from the Kruskal Wallis test. (C) Visualization of species level log2-fold changes in relative taxa abundance between day 0 and day 5 based on V3-V4 (left) and V6-V8 (right) datasets. Species are coloured by corresponding Phylum.

Overall, 53 genera were identified for the V3-V4 region, compared to 82 for V6-V8, both regions sharing 40 genera (Fig. 2-3A). Among the 13 taxa identified only in the V3-V4 region, *Actinobacteria* (*Micrococcales*, *Corynebacteriales*, *Frankiales*) and *Proteobacteria* (*Betaproteobacteriales*) were the most abundant, most of ASVs being distributed among *Actinobacteria* and *Proteobacteria* (Fig. 2-3B). As shown in Fig. 2-3C, the core microbiota down to the species level (i.e., species identified in both regions) was dominated by the phyla *Proteobacteria* and *Firmicutes*, while *Actinobacteria* and *Bacteroidetes* contained fewer ASVs. At the genus level, the ten most abundant taxa for the V3-V4 region included *Serratia*, *Pseudomonas*, *Lelliottia*, *Lactococcus*, an unidentified *Enterobacteriaceae*, *Cedecea*, *Aeromonas*, *Yersinia*, *Janthinobacterium* and *Kocuria*. Except for the genus *Lactococcus* that was not identified in one sample, these genera together with *Methylobacterium* were also the most prevalent. However, the list of the top ten genera for the V6-V8 region was not the same, as it consisted of *Serratia*, *Pseudomonas*, *Rahnella*, *Sediminibacterium*, *Lelliottia*, *Cedecea*, *Aeromonas*, *Lactococcus*, *Raoultella* and *Janthinobacterium*. These genera together with *Yersinia* and *Delftia* were among the most prevalent. Within the core microbiota, differentially abundant taxa were identified. They include *Methylobacterium* sp., *Stenotrophomonas maltophilia*, *Janthinobacterium* sp., *Pseudomonas* sp., *Cedecea davisae*, *Lelliottia* sp., *Yersinia* sp., and an unidentified *Enterobacteriaceae* that were significantly more abundant ($p<0.05$) in V3-V4 sequence profiles, and *Janthinobacterium lividum*, *Aeromonas* sp., *Lelliottia amnigena*, *Rhanella* sp., *Rhanella aquatilis* and *Chryseobacterium* sp. that were significantly more abundant ($p<0.05$) in the V6-V8 dataset. Of the 42 uniquely identified taxa from the V6-V8 profiles, *Proteobacteria* (*Enterobacteriales*, *Betaproteobacteriales*) and *Bacteroidetes* (*Chitinophagales*) were the most abundant (Fig. 2-3D). Most of ASVs were assigned to *Proteobacteria*, *Actinobacteria* and *Bacteroidetes* phyla.

2.5.2 Short-term storage of unpreserved milk resulted in significant alteration of the microbial community structure

Regardless of the sequenced hypervariable region, the estimated richness (Chao1) and diversity metrics (Shannon and InvSimpson) significantly decreased ($p<0.05$) after five days of storage at 4 °C when no preservatives were used (Fig. 2-2A). This indicates a drastic loss of diversity and thus the emergence of a few dominant species. Consequently, a significant shift ($p\leq0.008$ for V3-V4, $p\leq0.02$ for V6-V8) in community structure occurred during storage (Fig. 2-2B). In order to identify taxa that underwent significant increase or decrease in abundance and thus could explain the observed changes in community structure, we performed differential abundance analyses between fresh and stored milk samples without preservatives. As shown in Fig.

2-2C which illustrates fold changes for the V3-V4 region, *Streptococcus dysgalactiae* was reduced more than 20-fold in abundance at five days post-storage, while the 17 other bacterial groups changed by more than five-log2. *Chryseobacterium bovis*, *Pantoea* sp., *Raoultella* sp. and *Sanguibacter* sp. had the highest (> 25) $\log_2$-fold change. Most of these bacteria, particularly species of *Lactococcus*, *Acinetobacter*, *Rahnella*, *Pantoea*, *Enterococcus,* and *Lelliottia* became predominant after five days of storage (Fig. A1). For the V6-V8 region on the other hand (Fig. 2-2C), *Raoultella* sp. had a less than five-fold increase in abundance, whereas the majority of the 25 other taxa exhibited between five- and twenty-fold increases in abundance. *Chryseobacterium ginsengiterrae*, *Enterococcus* sp., *Macrococcus caseolyticus,* and *Pantoea* sp. exhibited the highest (> 25) fold changes. As for V3-V4 sequencing, most of these taxa, specifically species of *Lactococcus*, *Sphingobacterium*, *Acinetobacter*, *Yersinia*, *Erwinia* and *Rahnella aquatilis* were predominant in stored milk (Fig. A2). These observations are consistent with LEfSe on V3-V4 sequences that reveals *Lactococcus*, *Yersinia*, *Enterococcus*, *Macrococcus,* and *Leuconostoc* as biomarkers of five-day stored milk without preservative at 4 °C (Fig. A3-A). For the V6-V8 region, only *Acinetobacter*, *Macrococcus,* and *Rhodococcus* were identified as discriminatory (Fig. A3-B).

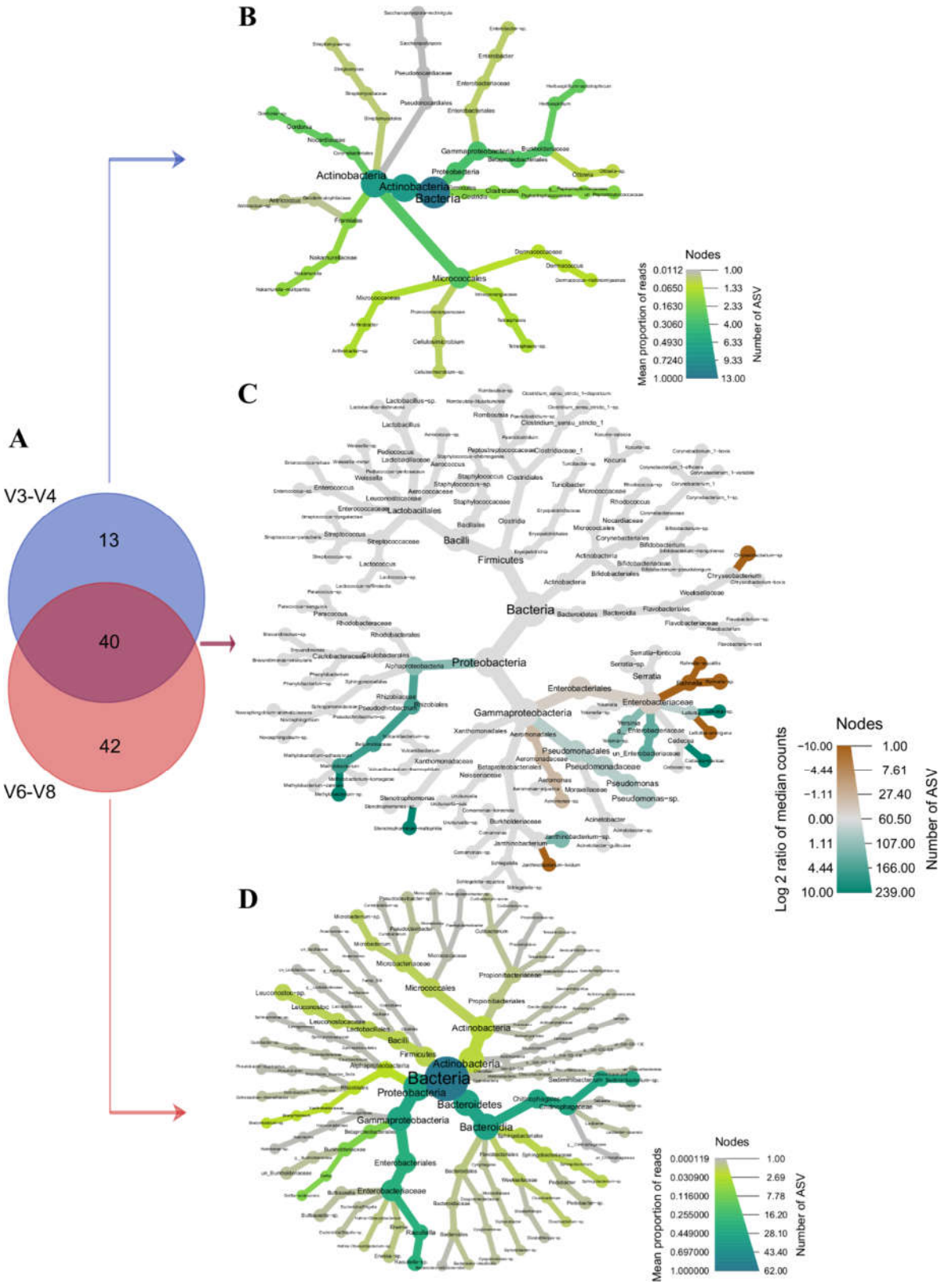


Figure 2-3: Core and differential microbial profiles obtained by partial sequencing of the 16S rRNA gene. (A) Venn diagram of taxa detected by sequencing the hypervariable regions V3-V4 and V6-V8. Values indicate the

number of shared (purple) and uniquely identified (blue or salmon) taxa at the genus level. (B) Heat tree illustrating the microbial profile from taxa uniquely detected by V3-V4 sequencing. The colour indicates the mean proportion of reads and the node size the number of ASV detected. (C) Differential abundance in the core microbiota of the V3-V4 and V6-V8 datasets. Taxa coloured green are significantly abundant ($p<0.05$) in the V3-V4 dataset, while those coloured brown are significantly abundant ($p<0.05$) in the V6-V8 dataset. The node size indicates the number of ASV. (D) Taxonomic distribution of taxa detected only in the V6-V8 dataset. The colour and node size have respectively the same meaning as for (B).

2.5.3 Impact of preservative chemicals on the microbiota once mixed with milk

In order to assess the stabilizing properties of preservation methods, we first analysed their effect on microbial communities once mixed with fresh milk. Regardless of the sequenced region, neither the alpha-diversity metrics, including Chao1 estimates and Shannon index, nor the beta-diversity measures, including unweighted- (Fig. A4 A) and weighted-UniFrac distances, were significantly different before and immediately after addition of preservatives (Fig 4A-B). However, log2-fold increases in taxa abundance were observed in the V3-V4 dataset for treatments AZ4 (*Rothia*, *Leuconostoc* and *Clostridium_sensu_stricto_1*), BR4 (*Clostridium_sensu_stricto_1*), and AZDm (*Clostridium_sensu_stricto_1*), while for the V6-V8 sequenced region, only the genus *Cloacibacterium* exhibited log2-fold reduction in abundance for treatment AZ4 (Fig. A4 B-C). To further investigate the extent to which microbial communities were sensitive to preservative chemicals, we also performed qPCR assays to quantify and compare community gene copy numbers between treated and untreated samples immediately after treatment application. Preservative chemicals tended to increase viable bacterial load (total bacteria group), with significant differences found for treatments AZ4 ($p<0.01$) and AZDm ($p<0.05$) compared to NoPre (Fig. 2-4C). No significant differences were found for any treatment for acetic acid bacteria (AAB), *Enterobacteriaceae* and *Pseudomonas* sub-groups. Unlike other bacterial sub-groups, lactic acid bacteria (LABs) appeared more sensitive to the chemicals, with a significant decrease ($p<0.05$), close to a log of their viable load under the effect of treatment BR4. Fungal sensitivity to preservative chemicals was also tested, for which biomass quantities in treated samples were comparable to those without treatment (NoPre), although within-sample variability was higher for BR4 and AZ4 treatments (Fig. 2-4C).

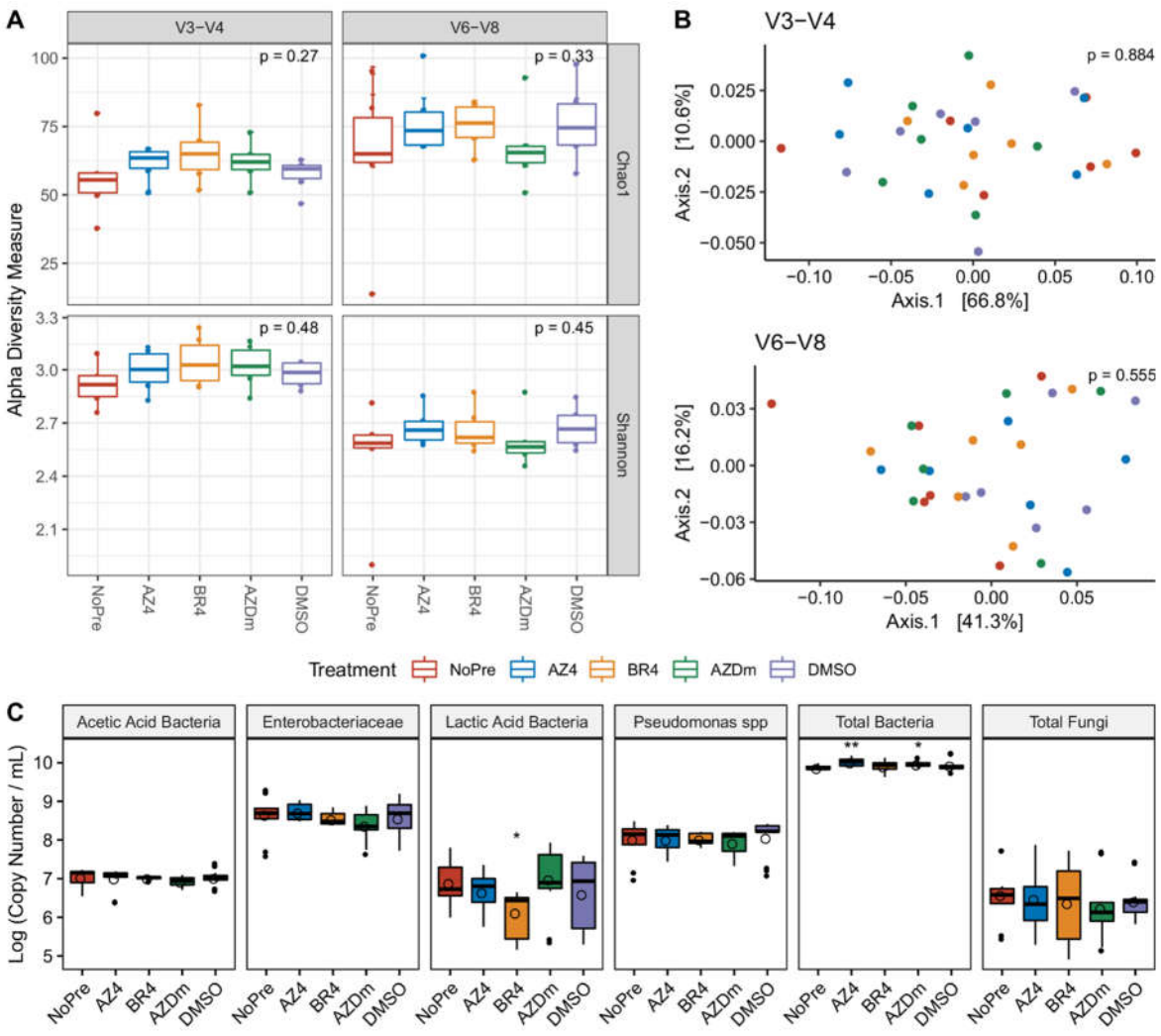


Figure 2-4: Changes occurring in the microbial composition and structure of raw milk once subjected to preservatives. Treatments include NoPre (untreated raw milk), AZ4 (raw milk treated with azidiol), BR4 (Bronopol-treated raw milk), AZDm (Raw milk treated with a mixture of azidiol and dimethyl sulfoxide), and DMSO (raw milk treated with dimethyl sulfoxide). (A) Visualization of alpha-diversity measures following treatments as described in the V3-V4 (left) and V6-V8 (right) datasets. (B) Principal coordinate analysis on weighted UniFrac distances based on the V3-V4 (top) and the V6-V8 (bottom) datasets. (C) Boxplot showing the bacterial and fungal quantification by PMA-qPCR. Asterisks above boxes indicate a significant difference compared to NoPre, obtained by performing a Wilcoxon rank test. P-values are flagged with asterisks as follows: *, p<0.05; **, p<0.01.

2.5.4 Impact of preservation methods on microbial community stability over time

Following the assessment of the preservatives' effects on the microbiota upon addition to milk, microbial community stability was evaluated at four timepoints over 30 days by comparing community diversity and structure between fresh and stored samples. Because higher discriminative power was obtained with the V3-V4 based dataset, only V3-V4 figures are presented in this section; all V6-V8 homologues are available as

supplemental materials. As shown in Fig. 2-5A, no significant change in alpha-diversity measure was observed during the storage of treated samples, except the Chao1 estimate ($p=0.04$) for DMSO treatment. However, we found no significant difference after performing a post-hoc test accounting for the multiple comparisons made (Fig. A5-A). Consistent results were also obtained with the V6-V8 based dataset, with no significant difference found for any of the alpha-diversity metrics, nor preservation methods (Fig. A6-A).

The ability of preservation methods to stabilize milk microbial communities was also evaluated by measuring Jensen-Shannon divergence, weighted- and unweighted-UniFrac distance discrepancies between serial timepoints for each single sample. Accordingly, data were recorded at three timepoints (Fig. 2-5B). Only Jensen-Shannon divergence in treatment AZ4 showed a significant difference in distance measure between consecutive timepoints ($p=0.01$). Most aliquots in other conditions remained relatively stable with little changes over time, although regression across timepoints may show more or less pronounced trends. Concerning Jensen-Shannon divergence measures on treatment AZ4, a post-hoc test revealed that overall milk community structure at five days post-storage was comparable to those at 10 and 30 days, while the 10- and 30-day samples were significantly different from one another (Fig. A5-B). Similar results were obtained with the V6-V8 sequence dataset regarding Jensen-Shannon divergence (Fig. A6-B), except that post-hoc analysis following significance found in treatment AZ4 did not reveal significantly different community structures across timepoints (Fig. A7-A).

Using the same distance measures as for evaluating single sample microbial stability over time, we investigated the stabilizing effects of preservation methods on community structures by comparing between-sample diversity across timepoints. Significant shifts in community structure were observed in treatments BR4 ($p=0.002$) and AZDm ($p=0.02$) when Weighted-UniFrac distance was used (Fig. 2-5C). For treatment BR4, pairwise phylogenetic distances tended to increase over storage time, while for treatment AZDm, the opposite was observed, indicating for the latter a better control on microbial dynamics. This is consistent with the post-hoc test that showed no significant difference across timepoints for AZDm treated samples, while for BR4 treated samples, community structure at day 30 was significantly different from all preceding timepoints (Fig. A5-C). Considering unweighted-UniFrac measure, there was no significant difference in any of the four treatments, implying relatively stable phylogenetic diversity over time. Jensen-Shannon divergence showed that significant differences in community structure arose during storage in AZ4 ($p=0.009$) and BR4 ($p<0.001$) treated milk (Fig. 2-5C). As illustrated in Fig. A5-C, these results were mostly explained by a significant difference between community structure at day 30 compare to day zero for AZ4, and between days zero, five, and ten for BR4. Unlike alpha-diversity and within sample stability over time, beta-diversity analyses on the V6-V8 related dataset showed different results (Fig. A6-C).

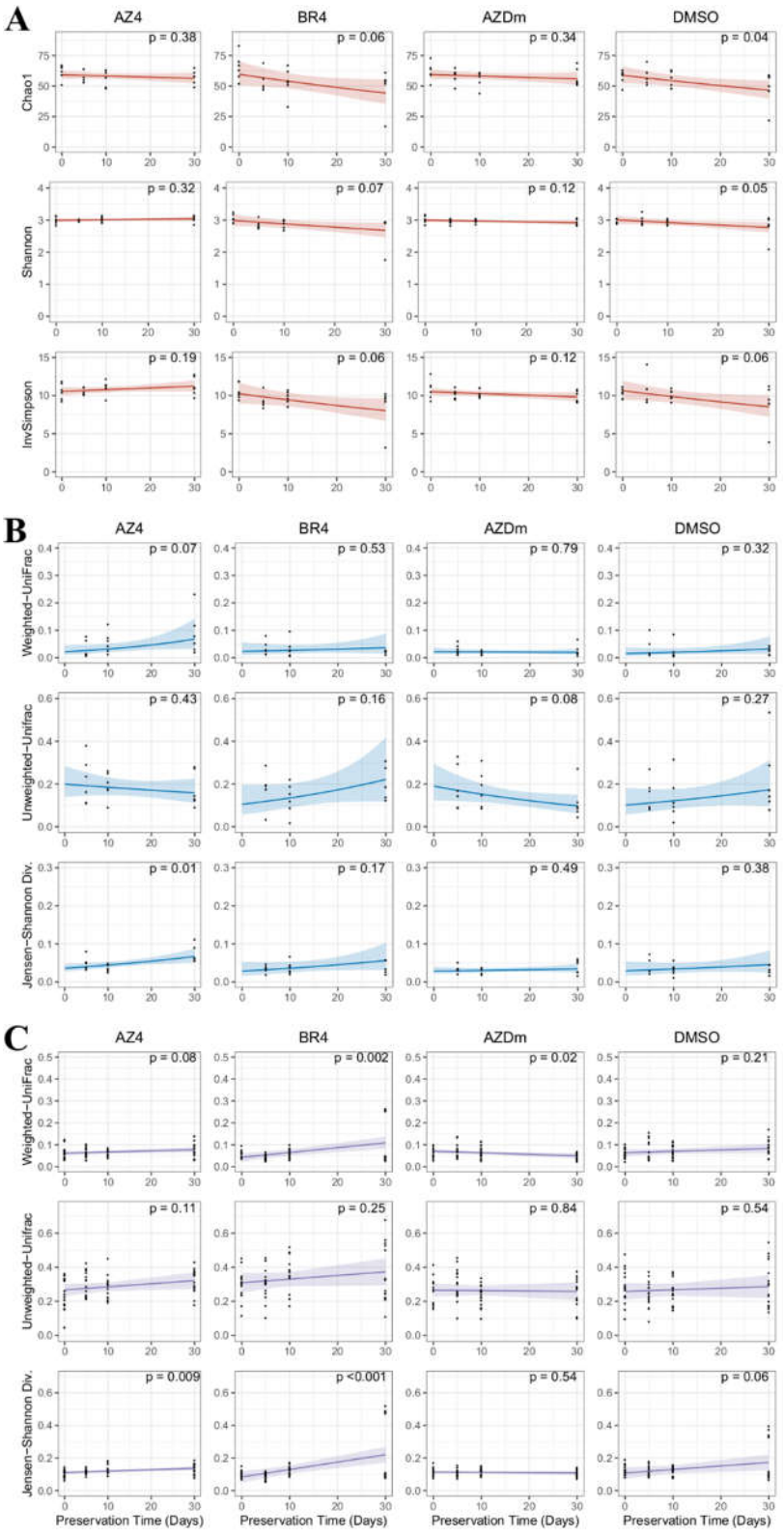


Figure 2-5: Temporal stability of microbial communities in preserved raw milk across 30 days of storage. Treatments include NoPre (untreated raw milk), AZ4 (raw milk treated with azidiol), BR4 (Bronopol-treated raw

milk), AZDm (Raw milk treated with a mixture of azidiol and dimethyl sulfoxide), and DMSO (raw milk treated with dimethyl sulfoxide). (A) Variations in alpha-diversity measures. For Chao1 estimates, Shannon and InvSimpson indices, salmon lines represent linear mixed-effects fit against the storage time and the 95 % confidence interval is shaded. (B) Diversity trends between samples of the same farm at consecutive timepoints. Blue lines represent linear mixed-effects fit against the storage time and the 95 % confidence interval is shaded. (C) Diversity trends between sample aliquots from different farms at consecutive timepoints. Purple lines represent linear mixed-effects fit against the storage time and the 95 % confidence interval is shaded.

Using weighted-UniFrac distance, significant shifts in microbial communities were found for AZDm ($p \leq 0.004$) and DMSO ($p=0.02$) treated samples over time, while with unweighted-UniFrac, significant changes ($p=0.03$) in community composition were found in AZDm treated samples only. For Jensen-Shannon divergence, significant changes ($p=0.01$) were found for AZDm treated samples. Despite these results, based on regression lines across timepoints, there was no obvious trend in community dynamics for any treatment and distance method. However, post-hoc analyses performed on weighted-UniFrac distance distinguished community structures at different timepoints in AZDm and DMSO treatments, but not in AZDm considering unweighted-UniFrac and Jensen-Shannon measures (Fig. A7-B).

Temporal changes in microbial composition at selected taxonomic levels were evaluated by quantifying consistency in taxa relative abundance through the calculation of ICC values at each timepoint for preserved and unpreserved samples. The rationale for treatment NoPre was to demonstrate variation in bacterial abundance that occur in the absence of any preservation measure other than storage at 4 °C. Intra class correlation results showed that for treatment NoPre, abundance consistencies at phylum down to genus levels varied from near excellent (≈ 0.9) to poor (<0.5) after only five days of storage, while at the ASV level, consistency decreased from moderate (≈ 0.6) to poor (<0.3), indicating drastic alteration of milk microbiota integrity during storage without preservatives (Fig. 2-6A). No such extent of microbiome community denaturation was noticed in treated samples, except for treatment BR4 where consistency at the ASV level dropped from good (≈ 0.71) at day five to poor (≈ 0.42) at day 30, while remaining good despite temporal instability (ICC values ranging from ≈ 0.9 at day five to ≈ 0.8 at day 30) at higher taxonomic levels. Treatment AZDm exhibited the most stable consistency in taxa abundance overall, followed by treatments AZ4 and DMSO. Contrasting results were obtained for the V6-V8 sequences dataset, where ICC values across time for all treated samples remained above 0.8 (good consistency) from phylum to species level, and above 0.6 (moderate consistency) at the ASV level, with no ICC drop of more than 0.2 (Fig. A8-A).

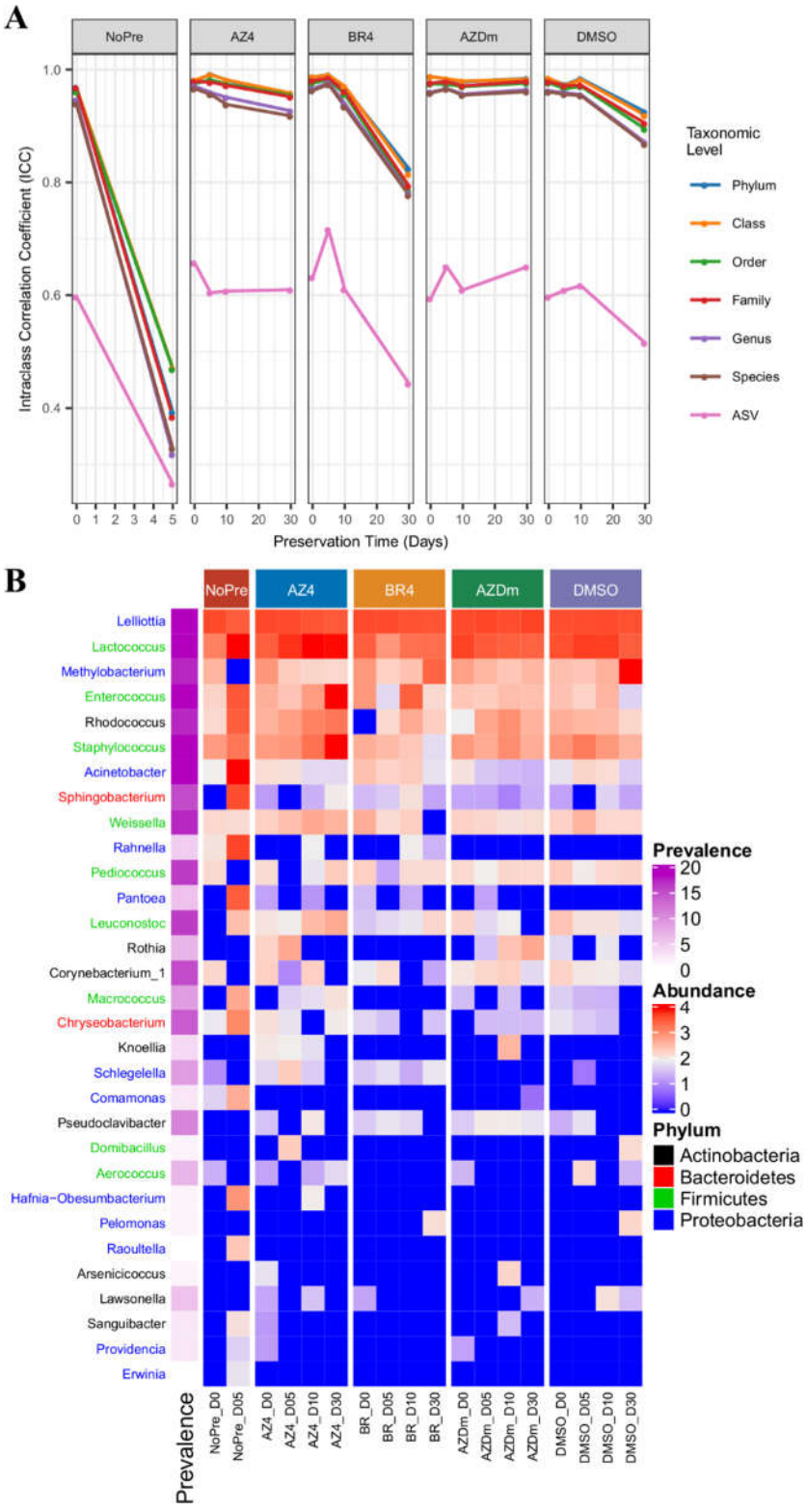


Figure 2-6: Community instability and taxa dynamics in unpreserved and preserved milk samples over storage time. Treatments include NoPre (untreated raw milk), AZ4 (raw milk treated with azidiol), BR4 (Bronopol-treated raw milk), AZDm (Raw milk treated with a mixture of azidiol and dimethyl sulfoxide), and DMSO (raw milk treated

with dimethyl sulfoxide). (A) Intraclass correlation coefficients computed for all the taxonomic levels between samples of different farms and plotted against the storage time. (B) Prevalence and abundance of taxa that underwent ≥ log2-fold changes. Each taxon at the genus level is coloured by its corresponding phylum.

To identify taxa specifically at the genus level whose changes in abundance across all consecutive timepoints may explain observed community shifts, we performed DESeq2-based differential abundance analysis. Among the fifteen most abundant taxa, *Proteobacteria* including *Lelliottia*, *Methylobacterium*, *Acinetobacter*, *Rahnella* and *Pantoea*, *Firmicutes* including *Lactococcus*, *Enterococcus*, *Staphylococcus*, *Weissella*, *Leuconostoc* and *Macrococcus*, and to a lesser extent *Actinobacteria* including *Rhodococcus*, *Rothia,* and *Corynebacterium*_1 were the main drivers of community structure instability in preserved and unpreserved milk samples (Fig. 2-6B). Genera exhibiting the highest log2-fold changes include *Lactococcus*, *Enterococcus,* and *Staphylococcus* for treatment AZ4, and *Methylobacterium* for treatment DMSO. Considering the V6-V8 based dataset on the other hand, the top fifteen most abundant taxa explaining structural changes in microbial community over time were largely dominated by *Proteobacteria* (Fig. A8-B). However, genera showing the highest log2-fold change include *Lactococcus*, *Staphylococcus,* and *Enterococcus* for treatment AZ4.

As a complementary approach to high throughput sequencing based analysis of microbial temporal stability, we performed PMA-qPCR to quantify variation in abundance of selected taxonomic groups over time. Viable microbial loads for taxonomic groups including AAB, *Enterobacteriaceae*, LABs, *Pseudomonas*, total bacteria, and total fungi, were expressed in gene copy number per mL of milk sample. AAB average loads were relatively stable over time for all treatments, although significant differences ($p<0.05$) were found at days 30 for treatment BR4 and 5 for treatment AZDm compared to day zero, respectively (Fig. 2-7). Unlike AAB communities, *Enterobacteriaceae* were slightly less sensitive to preservatives, particularly at days 5 for treatment BR4 and 10 for treatment AZDm where significant increases in bacterial loads ($p<0.001$, $p<0.05$, respectively) were observed. LAB were less sensitive to treatment AZ4, for which significant increases in community loads were observed at day 30 compare to day zero. For the *Pseudomonas* group, significant increases in community loads were noted at day 5 for treatments AZ4 and BR4 and AZDm ($p<0.05$, $p<0.001$, $p<0.01$, respectively). Overall, viability-PCR for total bacteria showed that communities were relatively stable for all treatments, except treatment BR4 for which a significant increase in bacterial load was observed at day 5, even though variation amplitudes were only a few tenths of a log. On the other hand, fungal communities were stable under preservation methods involving freezing (AZDm and DMSO) but grew dynamically under treatments AZ4 and BR4, which both involved refrigeration at 4 °C. Compared to day zero, fungal loads significantly increased at days 10 and 30 ($p<0.05$, $p<0.001$, respectively) for treatment AZ4, and days 5, 10, and 30 ($p<0.05$, $p<0.0001$, $p<0.0001$, respectively) for treatment BR4.

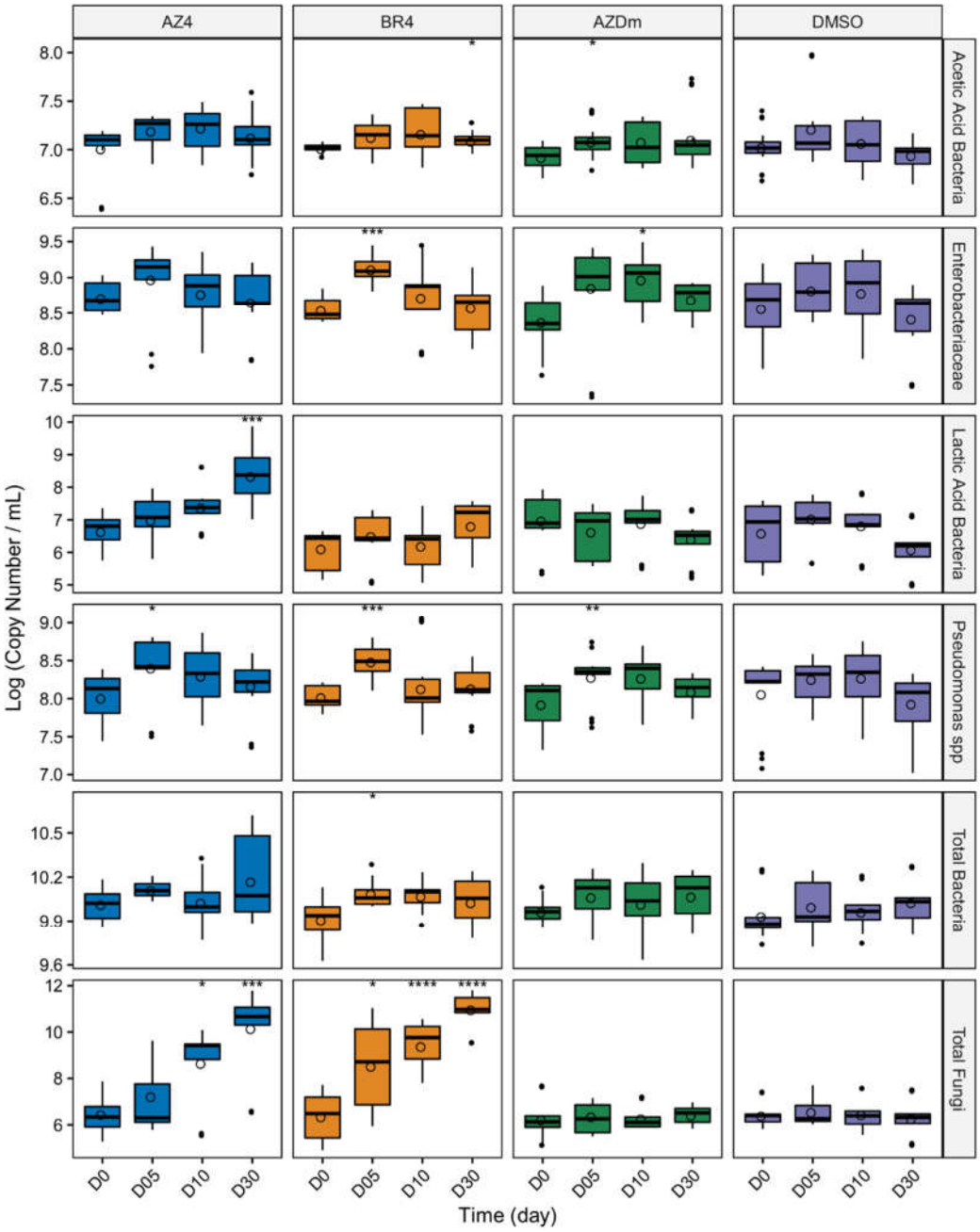


Figure 2-7: Quantification of viable microbial groups/sub-groups over time. Log copy numbers/mL are plotted against the preservation time. Copy numbers calculated at timepoints D05, D10 and D30 are compared to that at D0 using a Wilcoxon rank test. Asterisks above boxes indicate significant difference compared to D0, and flag p-values as follows: *, $p<0.05$; **, $p<0.01$; ***, $p<0.001$; ****, $p<0.0001$.

2.6 Discussion

In the current study, the "viability high-throughput sequencing" approach was used to assess the efficiency of raw milk preservation methods. By analogy to viability-PCR, "viability high-throughput sequencing" refers to high-throughput sequencing of DNA from live microbial cells as depicted by PMA. Regarding technical limitations associated with the use of PMA [76], a previous study conducted in our laboratory [82] has optimized the application of PMA-qPCR on milk and cheese samples, and since then it has become routine practice on various matrices [81,83,232,233]. We showed that V3-V4 and V6-V8 hypervariable regions of the 16S rRNA gene do not draw identical pictures of the community profiles of fresh raw milk samples. Indeed, we found different, but overlapping lists of the top ten predominant taxa in fresh raw milk for both regions. Moreover, compared to the V3-V4 region, the microbial profile using the V6-V8 region exhibited a higher richness, but lower alpha-diversity, indicating large dominance by few taxa. These results are consistent with previous studies which compared single or combined 16S rRNA hypervariable regions [100,234].

On the other hand, in unpreserved raw milk conserved at 4 °C for five days, we observed significant drops in richness and alpha-diversity measures for both hypervariable regions. This was accompanied by a highly significant shift in community structure again revealed by both variable regions. These results clearly show that storing raw milk samples at 4 °C is insufficient to stabilize the microbial community for up to five days, consistent with previous studies [7,189]. More importantly, our results confirm those from previous studies which reported the predominance of a diversified portion of the microbiota in refrigerated milk, including psychrotrophic and mesophilic bacteria [7,189]. Accordingly, we identified the genera *Lactococcus*, *Yersinia*, *Enterococcus*, *Macrococcus*, and *Leuconostoc* in V3-V4 sequences, and in V6-V8 sequences the genera *Acinetobacter*, *Macrococcus*, and *Rhodococcus* as biomarkers of five-day stored raw milk, thus presumably spoiled raw milk. Once more, the identification of microbial taxa that characterize potentially spoiled raw milk during or after storage and transportation under refrigeration to dairy plants may provide different results depending on the sequenced region. Comparison of published results from partial 16S rRNA high-throughput sequencing is problematic due to the use of a wide variety of analytical techniques and bioinformatic pipelines. A consensus on the most appropriate hypervariable regions that best describe microbial communities, particularly for dairy associated bacteria, is yet to be established [99,235]. Identifying the best hypervariable region between V3-V4 and V6-V8 was, however, not the scope of the current study. Nevertheless, due to its greater coverage, the V6-V8 region seems more adapted for exploring the microbial composition of milk. For hypothesis driven microbiome studies aiming at characterizing or comparing different biological or environmental conditions associated with milk, the V3-V4 region would be preferred, as it appeared statistically more discriminative. However, the phylogenetic resolution provided by amplicon sequencing is a limiting factor in the microbial profiling and the evaluation of its dynamics in various ecosystems. With the recently developed approach that combines circular

consensus sequencing from the Pacbio platform and the DADA2 algorithm to better classify long-read sequences up to sub-species level [96], targeting the 16S rRNA gene along its full length will lead to better resolution in species identification.

We assessed the immediate effects of preservative substances on milk microbiota before investigating their stabilizing ability. Our results showed that none of the preservatives adversely affected milk microbiota composition and structure once added to samples. This corroborates previous studies that showed that bacterial viability and recovery were not hampered within 24 to 48 h after azidiol was mixed with milk samples [189,201]. Our results indicate no negative interactions between the preservatives tested and the ability of PMA to target viable cells. Nevertheless, some biases were noted for a few taxa where abundance was significantly increased when AZ4, BR4, or AZDm treatments were applied, but not sufficiently to induce significant differences at the community level. The same trend was also noted for viable total bacterial loads determined by PMA-qPCR on AZ4 and AZDm treated milk aliquots. Even though the immediate bactericidal effect of treatment BR4 was significantly higher on LAB compared to other sub-groups quantified, the microbiota composition and evenness (alpha-diversity) in BR4-treated milk remained stable over the course of storage, as for that of other treated milk samples, regardless of the sequenced region. In contrast, the community structure (beta-diversity) varied significantly across time in BR4-treated milk, showing the highest variability in consistency measured through ICC analyses for all taxonomic levels. Our results showed that BR4 treatment is inappropriate to preserve milk samples intended for microbiome analyses. They confirmed those from Sierra et al. who used automated flow cytometry to evaluate the effect of bronopol on total bacterial counts in preserved goat milk [236]. The authors did not recommend bronopol as a preservation method preceding total bacterial counts because of its biocidal effect.

We also demonstrated, as have previous studies on other matrices [190–193], that freezing, particularly with a cryoprotectant such as DMSO [201], is the gold standard for preserving dairy milk samples. Indeed, we found that milk microbial community composition and structure were stable when frozen, supplemented or not with azidiol. However, our analyses on α- and β-diversity measures showed that although not always significant, there was a higher trend of differentiation between microbial communities across timepoints for DMSO-treated samples, compared to AZDm treatment. The tendency was more accentuated between the 10 and 30 days of storage. This indicates, as expected, that sample treatment with azidiol before freezing provides a better preservation of the microbial community in milk samples. To our knowledge, this is the first report on the preservative properties of azidiol combined with DMSO before freezing. We also found a better stability of the milk microbiota in AZ4-treated samples compared to DMSO-treated ones, specifically for up to 10 days post-storage. Our findings corroborate those of Martins et al. [195] who showed that the total bacterial counts in azidiol treated milk remained stable under refrigeration between 1 ° and 4 °C for up to 1 week. For large-scale

studies in remote areas, sample refrigeration would be more easily and reliably achieved than freezing. Therefore, in this instance, treatment AZ4 would be preferred to AZDm, particularly to minimize the bacterial proliferation for short-term preservation before sample analysis. Besides bacteria, variation of fungal loads across time in treated and untreated samples were assessed by PMA-qPCR. We found that only treatments AZDm and DMSO were able to stabilize the total fungal community, while treatments AZ4 and BR4 were not able to stabilize them for at least up to 5 days. These latter treatments lacked freezing, which is well known for inhibiting fungal growth. In addition, the lack of any fungistatic or fungicidal substance certainly explains yeast and mould proliferation during milk storage. Several studies, as reviewed by Frey-Klett et al. [237], have demonstrated that fungi and bacteria occurring in the same environment can interact to favour their growth through diverse physical and chemical mechanisms. Because fungal communities constitute an important part of raw milk microbiota [5], fungistatic substances (not fungicidal) should be used to limit their proliferation during sample storage. Indeed, preventing fungal growth in stored milk sample would logically improve azidiol efficiency, and ultimately allow long-term storage of AZ4-treated raw milk for microbial analyses.

Besides high throughput marker gene or shotgun sequencing, other omics approaches, as reviewed by Tilocca et al. [73], have been used to characterize the microbiota of dairy milk and products. Whether they involve gene expression analyses of microbial communities (metatranscriptomics), their metabolites identification and quantification (meta-metabolomics) or the proteome analysis (metaproteomics), they all depend on the microbiome stability for consistent and reliable study results. Sodium azide, one of the main constituents of azidiol, has been successfully used to minimize bacterial growth during proteomic studies on distinct matrices such as milk [238] or urine [239]. Also, chloramphenicol, another constituent of azidiol, has been extensively used in studies characterizing proteome dynamics of microbe's antibiotic resistance. To our knowledge, no incompatibility has been reported between sodium azide and chloramphenicol. Therefore, like the recently developed metaproteomic approach to decipher enzyme modulation by lysozyme treatment of Grana Padano cheese samples [130], future related large scales studies might obviously benefit from azidiol preservative abilities to fix microbial communities in collected samples. However, depending on the analysed matrix and the research objectives, optimizations would be needed to ensure efficiency and practicality of the preservation method.

2.7 Conclusions

Using a "viability high-throughput sequencing" approach combined with viability-PCR, we demonstrated that a combination of azidiol and DMSO to preserve the microbiota of raw milk at freezing temperature was more efficient than simple freezing without a preservative agent. Interestingly, we showed that the microbiota in azidiol-treated milk was equally stable for up to 10 days at refrigeration temperature of 4 °C. Our findings suggest that in the latter condition, combining azidiol with a fungistatic substance would improvably extend sample

preservation time. Being among the most targeted hypervariable regions in high-throughput amplicon sequencing of the 16S rRNA gene, we proved that V3-V4 and V6-V8 regions do not always provide the same picture from the same microbial populations, highlighting caution in choosing the right region to sequence according to research objectives if short-read sequencing is to be performed. Based on the above findings, we encourage the use of azidiol for optimal conservation of raw milk microbiota intended for culture-dependent and -independent analyses in large-scale epidemiological or longitudinal studies.

Chapitre 3 – Écologie microbienne de fourrages d'herbes et de maïs ensilés avec ou sans inoculants

Metataxonomic insights into the microbial ecology of grass or legume and corn silage produced with and without inoculants

Alexandre J. K. Ouamba [1,2], Mérilie Gagnon [1,2], Thibault Varin [1], P. Yvan Chouinard [2,3], Gisèle LaPointe [2,4], Denis Roy [1,2]

[1]Département des sciences des aliments, Laboratoire de génomique microbienne, Université Laval, 2440 boulevard Hochelaga, Québec G1V 0A6 Canada

[2]Regroupement de recherche pour un lait de qualité optimale (Op+Lait), 3200 rue Sicotte, Saint-Hyacinthe J2S 2M2 Canada

[3]Département des sciences animales, Université Laval, 2425 rue de l'Agriculture, Québec G1V 0A6 Canada

[4]Department of Food Science, University of Guelph, 50 Stone Road E, Guelph N1G 2W1 Canada

3.1 Résumé

Le microbiote des fourrages conservés détermine leur qualité, et peut affecter la composition microbienne et l'aptitude du lait à la transformation. La diversité et la viabilité des communautés bactériennes de cinq types de fourrages incluant le foin et les ensilages inoculés ou non ont été analysées par métataxonomique dans 24 fermes laitières échantillonnées à deux reprises. Il en ressort que les bactéries lactiques dominent le microbiote d'ensilages, alors que dans le foin, les genres *Pantoea* et *Sphingomonas* sont dominants. Une abondance élevée des genres *Pediococcus*, *Weissella* et *Bacillus* caractérise les ensilages d'herbe/légume comparé à ceux du maïs où sont observées des proportions élevées du genre *Acetobacter*. Les inoculants commerciaux n'affectent pas systématiquement le microbiote d'ensilages, et seraient par ailleurs ubiquitaires. Cependant, l'analyse des réseaux d'interactions bactériennes a révélé des différences de co-occurrence et de rôle topologique bactériens entre les ensilages inoculés et non inoculés.

3.2 Abstract

The microbiota of preserved forage is a key determinant of its quality, and it can critically affect raw milk microbial composition and processibility. Here, we comprehensively assessed the diversity of viable bacterial communities of hay and grass or legume and corn silage to deepen our knowledge of how conservation processes and inoculant addition drive microbial occurrence patterns on dairy farms. Samples of eight different forage types were collected at the feed-out phase from 24 dairy farms over two sampling periods and analysed by high-throughput sequencing and quantitative PCR after being treated with propidium monoazide to account

for viable cells. We found consistent significant differences between hay and silage community structures across sampling periods. Silage was generally dominated by lactic acid bacteria (LAB), while *Pantoea*, *Sphingomonas*, *Curtobacterium*, and *Methylobacterium* were the main co-dominant genera in hay. Grass or legume and corn silage exhibited phylogenetically dissimilar microbial profiles, the former being characterized by high relative abundance of *Pediococcus*, *Weissella*, and *Bacillus*, and the latter by high proportions of *Acetobacter*. The use of commercial microbial additives including either *Lentilactobacillus buchneri* alone or in combination with *Lactiplantibacillus plantarum*, *Lacticaseibacillus casei*, *Pediococcus pentosaceus*, and *Enterococcus faecium* in different formulations did not systematically improve silage microbial profile, especially when corn-based forages were produced. However, inoculating grass or legume silage tended to reduce the occurrence and abundance of *Weissella*, but inconsistently prevented that of *Bacillus*. Moreover, the core *Lactobacillales* phylotypes were the dominant LAB in uninoculated and inoculated grass or legume and corn silage with up to 96 % and 95 % relative abundance, respectively, indicating either the ubiquity of inoculants or the high competitiveness of epiphytes. Forage physicochemical parameters as well as sampling periods, storage forms, and inoculants associated with specific taxa. Variations in taxa co-occurrence patterns and topological roles between uninoculated and inoculated silage types demonstrated the usefulness of the network analysis for deciphering silage microbial ecology in large-scale facilities. The integration of management practices and forage physicochemical parameters with microbial dynamics and interactions during forage processing stages is needed to fully decipher taxa roles in silage fermentation.

3.3 Introduction

Forage conservation is critical to ensure a proper yearlong availability of feed for dairy cattle. This is particularly important in cold-weather areas characterised by rough winter conditions and limited growing seasons such as those prevailing in North America [240]. If the origin and patterns of phyllosphere community assembly are unclear [241], the fate of microorganisms dwelling on fresh plants throughout the preservation processes contributes to the quality of conserved forages and determines associated potential risks to animal and human health [242]. Due to the development and the application of molecular techniques to gain insight into the microbiome associated with preserved forages [167], management systems have gained substantial upgrades intended to improve safety and production yields, as well as beneficial effects to animals [26,170,243–245]. Management practices and environmental factors [240] unavoidably alter the microbial content of preserved forages. These feeds therefore constitute important vehicles for various microorganisms including bacteria and fungi from the growing fields to dairy barns, and ultimately to milk and dairy products [9,112,170].

Hay and silage are the main forms of forage conservation in dairy production systems [23]. Although haymaking consists of drying forage crops in order to suppress enzymatic and microbial activities, hay still harbours a viable microbiota which composition and structure are not well known [246,247]. On the other hand,

ensiling is based on the fermentative properties of epiphytic microorganisms, particularly lactic acid bacteria (LAB), that metabolize water soluble carbohydrates (WSC) into organic acids under anaerobic conditions wherein the rapid decline in pH is a key determinant of silage quality. The genera *Lactiplantibacillus, Lacticaseibacillus, Lentilactobacillus* (formerly *Lactobacillus*), other lactobacilli, *Pediococcus*, *Weissella*, *Leuconostoc*, *Enterococcus*, *Streptococcus*, and *Lactococcus* are generally associated with silage [170,245].

However, not all LAB strains can induce fast pH decrease during the early stages of fermentation, and in naturally fermenting forage crops, they may be outcompeted by undesirable acid-tolerant bacteria including *Enterobacteriaceae*, acetic acid bacteria (AAB), and spore-forming bacteria associated with poor quality silage [243,245]. Microbial additives encompassing homofermentative or facultative heterofermentative LAB (*Lactiplantibacillus plantarum*, *Lacticaseibacillus casei*, *Pediococcus* spp.), obligate heterofermentative LAB (*Lentilactobacillus buchneri*, *Lentilactobacillus hilgardii*), combination inoculants, and non-LAB inoculants (*Bacillus subtilis*) have been proposed to enhance forage crops fermentation and improve the aerobic stability of silage, as well as its safety and nutritional value [26,245]. From a microbiological viewpoint, the efficiency of these commercial inoculants whether as single or as multi-species formula have been generally assessed in controlled laboratory conditions that do not mimic the various management practices and changing environmental factors observed in large-scale ensiling [23,248,249]. Consequently, few studies have evaluated the microbial communities populating silage prepared in various farm-scale silo types [248,250], and none of this kind have focussed on associations between physicochemical characteristics and the viable microbiota at feed-out, while contrasting uninoculated versus inoculated silage.

Gagnon et al. [22] recently used a culture-based approach to analyse LAB communities occurring in hay and grass/legume and corn silage produced with and without inoculation at different dairy farms. They revealed that while *L. casei/paracasei* and *L. plantarum* were common among all forage types, *Enterococcus mundtii*, *L. pentosus*, *Companilactobacillus tucceti*, *Lactococcus lactis*, and *Leuconostoc mesenteroides/pseudomesenteroides* were only identified in hay, whereas *L. buchneri* group was specific to silage regardless of their type and inoculation status. A better knowledge of the viable microbial communities of preserved forages including hay and grass/legume or corn silage at feed-out from farm-scale facilities will provide new insights into their microbial loads and help evaluate the effectiveness of forage management practices implemented on dairy farms. This information is useful for fine-tuning the search for new additives and may help forage makers adjust processing routines to improve the hygienic quality, nutritive value, and aerobic stability of conserved forages. The current study complements that from Gagnon et al. [22] by implementing a viability high-throughput sequencing approach combined with viability-PCR [43,251] on the same samples to provide a comprehensive and comparative analysis of the microbial ecology of hay and grass/legume or corn silage produced with or without inoculants. Therefore, this study aimed to assess the diversity of microbial communities

occurring in various types of preserved forages and gain insight into how inoculants shape the microbiota of mature silage in commercial settings.

3.4 Materials and methods

3.4.1 Dairy farm selection and sample collection

Farm recruitment and sampling were carried out as previously described [22]. Briefly, the 24 tie-stall herds from the province of Quebec (Canada) were selected based on the forage harvested as hay or silage. Accordingly, farms were grouped into five feeding typologies comprising herds fed with either hay as the unique type of forage (H), or silage as the main forage source, the latter including grass/legume silage (GL), grass/legume silage with corn silage (C) as supplement, grass/legume silage with corn silage inoculated at harvest (CI) as supplement, and grass/legume silage inoculated at harvest (GLI) with CI as supplement.

Sampling was carried out in the fall 2015 and the spring 2016. On each farm, forage samples were collected during the same visit as previously described [22]. Within the farms enrolled in the study, eight storage forms or silo types included loose and baled hay as well as wrapped square/round bales, concrete-stave silo, oxygen-limiting silo, pressed bag silo, bunker silo, and stack silo for silage (Table 3-1). For inoculated silage, the commercial inoculants used to control the ensiling process included 11C33, 11CFT and 11G22 (Pioneer, Johnston, IA), as well as Biotal Buchneri 500 and Biotal Supersile (Lallemand Animal Nutrition, Milwaukee, WI). From the 500 g of each forage sample collected, a subsample was sent to Lactanet laboratories (Sainte-Anne-de-Bellevue, Qc, Canada) for infrared quantification of organic compounds, moisture, pH, and minerals, as well as the estimation of fermentation acids when applicable. All the above-mentioned sample and farm information composed the sample metadata that were further integrated in the dataset and used where applicable as categorical or quantitative variables for data analysis.

Table 3-1: Distribution of forage storage forms and silo types

	Storage form / Silo type							
Feed group[a]	**Loose**	**Small bale**	**Wrapped square/round bale**	**Concrete-stave silo**	**Oxygen-limiting silo**	**Pressed silo**	**Bunker silo**	**Stack silo**
H	2	10						
GL			11	10	2	4		2
GLI				14	2			
C				4		2		2
CI				14			2	

[a]H, hay; GL, non-inoculated grass, or legume silage; GLI, inoculated grass/legume silage; C, non-inoculated corn silage; CI, inoculated corn silage.

3.4.2 DNA extraction and high-throughput sequencing of the 16S rRNA gene pool.

For forage samples, 30 g of each were homogenized in 270 mL of peptone buffer solution as previously described [22]. Aliquots of 2 mL immediately taken from the suspension were centrifuged at 12,000 x g for 15 min at 4 °C. The pellets were washed twice with 1 mL sucrose buffer (sucrose 12 % [w/v], 25 mM Tris-HCl pH 8.0). Half of the cell suspension were treated with propidium monoazide (PMA) as previously described [43]. All non-PMA and PMA treated cells were stored at –80 °C. Genomic DNA was then extracted using the DNeasy PowerFood Microbial Kit (Qiagen, Hilden, Germany) combined with enzymatic lysis with mutanolysin from *Streptomyces* (MilliporeSigma), lysozyme (MilliporeSigma), and proteinase K (MilliporeSigma) following the steps previously described [43].

Sequencing was performed in a single run on Illumina MiSeq at the *Plateforme d'Analyses Génomiques*, Laval University (Quebec, Canada). Primer pairs 347F (5'-GGAGGCAGCAGTRRGGAAT)/803R (5'-CTACCRGGGTATCTAATCC) were used to amplify the V3-V4 region of the 16S rRNA gene.

3.4.3 Bioinformatic and statistical analyses

The Cutadapt (version 2.3) software [213] was used to remove adapters and primers from demultiplexed sequences. Sequencing reads were modelled and denoised with the DADA2 (version 1.14) pipeline [98] developed for R. After constructing the merged sequence table and removing chimeras, the Silva version 132 DADA2-formatted reference databases (down to genus and species levels) were used for taxonomy assignment. Sequence alignment and phylogenetic tree construction were performed using the DECIPHER (version 2.14.0) package [214] and phangorn (version 2.5.5) package [215], respectively, as previously described [216]. The resulting amplicon sequence variants (ASVs) and phylogenetic tree were further processed with the phyloseq (version 1.30.0) package [217] for alpha- and beta-diversity analyses of the forage microbiota.

As a preprocessing step in the data analysis, community differences between non-PMA and PMA treated samples were compared, as well as between sampling periods, by computing alpha- and beta-diversity measures using the Phyloseq package. Centered Log-Ratio (CLR) and Phylogenetic Isomeric Log-Ratio Transform (PhILR) were applied to the ASV table prior to beta-diversity analyses for which Aitchison and Euclidean distances were used, respectively, to account for the compositional nature of microbial data [252,253].

Microbial communities in forage types were characterized by assessing and comparing alpha-diversity measures including Chao1, Shannon and Inverse Simpson indices, and beta-diversity measures including the sample local contribution to beta-diversity (LCBD), principal coordinate analysis (PCoA), and principal component analysis (PCA). LCBD is another way of assessing the beta diversity that provides comparative indicators of the uniqueness of a community profile in a single sample among groups [254]. LCBD indices were

computed on Hellinger-transformed data with the microbiomeSeq R package [255]. PCA was performed on CLR and PCoA on PhILR normalized data as described above to assess the dissimilarities of community structures among forage types. Significant differences were determined at 0.05 threshold from false discovery rate (FDR) corrected p-values after a Kruskal–Wallis test computed with the R package ggpubr [220]. Microbial differential abundance testing was performed for paired combinations between uninoculated and inoculated forage types using the R package ALDEx2 4.0 [256]. The functional metagenomic content of forage samples was predicted using the Piphillin software [121] through the Piphillin online server (http://piphillin.secondgenome.com/). The differential abundance of microbial functional features was computed with ALDEx2 as described above. The abundance and effect size of differentially abundant taxa or functional features between uninoculated and inoculated forage types were visualized using a heatmap constructed with the Bioconductor package ComplexHeatmap 2.4.2 [225].

Categorical and quantitative metadata related to farms and forage samples were selectively integrated in multivariate multi-table analyses based on sparse partial least squares regression (sPLS) and canonical correspondence analysis (CCpnA) to determine their association with microbial features. The analysis scheme was partially inspired from a previously described study [257]. Relevant metadata variables were selected for multivariate multi-table analyses by permutational multivariate analysis of variance (adonis) on CLR normalized data. The function *adonis2()* from vegan (version 2.5-6) R package [258] was used on Aitchison distance matrix computed in the phyloseq package. Significant variables ($p<0.05$) identified using adonis test were retained for sPLS based analyses performed with MixOmics (version 6.1.1) R package [136]. To complement the sPLS approach, general linear models implemented in MaAsLin2 R package (https://github.com/biobakery/MaAsLin2) were also used to determine multivariate associations between forage metadata and ASVs based on CLR normalized data. For MaAslin2 analyses, ASVs significantly associated with adonis selected variables were limited to those which BH corrected p-values were lower than 0.25. CCpnA analysis of each forage type was performed as previously described [257] on corresponding dataset reduced to all ASVs resulting from both sPLS and MaAsLin2, as well as all variables selected based on adonis test. Besides the CCnpA triplot result that illustrates how microbiome patterns are related to farm and forage metadata, prevalence, abundance, and distribution of all selected ASVs among each forage type were visualised in a heatmap constructed with the ComplexHeatmap package.

Microbial communities in uninoculated and inoculated forages (grass/legume and corn silage) were further characterized by investigating co-occurrence patterns using network analyses. The network inference was performed by constructing the phylogenetic molecular ecological networks (pMENs) using the online molecular ecological network analysis pipeline (MENAP, http://ieg4.rccc.ou.edu/mena) as previously described [259,260]. For pMENs construction, only ASV with at least 40 % prevalence within a silage group were

considered. The similarity matrix that measures the degree of concordance between the abundance profiles of ASV across samples was obtained based on Pearson correlation coefficients calculated from CLR-transformed data. Depending on the structure of the dataset corresponding to each silage group, the random matrix theory-based approach [261,262] was used to automatically determine the appropriate threshold value for network structure. The fast greedy modularity optimization procedure [263] was used to detect network modules defined as a group of highly interconnected nodes that share few or no connections with other nodes outside the group. In the built network, each node corresponds to a single ASV. The Maslov-Sneppen approach [264] was used to generate 100 randomly rewired networks for each pMEN obtained. Global network and individual node properties were calculated based on similarity matrices. Node topological roles for a given network were determined based on within-module connectivity (Z_i) and among-module connectivity (P_i) indices and visualized with a scatterplot constructed using the ggplot2 (version 3.3.0) R package [223]. Accordingly, nodes were assigned four different topological roles including module hubs, network hubs, peripherals, and connectors as previously defined [265]. Peripheral ASVs also considered as specialists, have both low Z_i and low P_i indices ($Z_i \leq 2.5$, $P_i \leq 0.62$) and characterized by few links almost always with other ASVs within their module. ASVs assigned the connector role have low Z_i and high P_i indices ($Z_i \leq 2.5$, $P_i > 0.62$) and are highly connected to several modules. Module hubs have high Z_i and low P_i indices ($Z_i > 2.5$, $P_i \leq 0.62$) and are highly connected to several ASVs within their module. Connectors and module hubs are both considered as generalist species. Network hubs also defined as super-generalists have high Z_i and high P_i indices ($Z_i > 2.5$, $P_i > 0.62$), thus playing both connector and module hub roles. Habitat generalists refers to species, in this study ASVs, that are largely distributed across samples within a group, thus having high prevalence, while habitat specialists are restricted to few samples in the group they belong to, thus having low prevalence, but occur at high relative abundance [266]. Module hubs, network hubs, and connectors that have high values of either Zi or Pi, or both high connectivity indices, are generally considered as keystones species.

Besides functional analyses, phenotypic traits of forage microbiome were predicted and compared using the BugBase tool [267]. BugBase's algorithm relies on software such as PICRUSt [268], IMG [269], KEGG [123], and PATRIC [270] to predict phenotypes and corresponding microbial contributors at the phylum level. Phenotypic traits including biofilm formation, Gram staining, oxygen tolerance, pathogenic potential, mobile element content, and oxidative stress tolerance were then predicted from CLR transformed data. Prior to BugBase analyses, chimera free sequences derived from the DADA2 pipeline were mapped to the Greengenes 97 % reference database for format compatibility requirements.

3.4.3 Quantitative PCR

Quantitative polymerase chain reaction (qPCR) was performed to determine copy numbers of specifically targeted genes of *L. buchneri*, *L. plantarum*, lactic acid bacteria (LAB), acetic acid bacteria (AAB),

Enterobacteriaceae, *Pseudomonas*, total bacteria, and total fungi in all PMA free and PMA treated forage, using specific primer pairs as previously described [43]. Amplification reactions were carried out in duplicate using a ViiA7 system (Thermo Fisher Scientific, Burlington, ON, Canada). Each reaction mixture of 10 µL total volume was composed of 3.6 µL UltraPure DNAse and RNAse free distilled water (Thermo Fisher Scientific), 5µL PowerUp SYBR Green master mix (Thermo Fisher Scientific), 0.2µL of each primer at 10 nM, and 1 µL DNA template. Bacterial loads were reported as gene copy number per milligram of forage.

3.5 Results

3.5.1 Bacterial diversity in forage types

A total of 81 forage samples were collected from 24 dairy farms over two sampling periods. During both periods, H samples exhibited significantly higher Chao1, Shannon, and inverse Simpson indices ($p<0.01$) compared with other forage types which did not significantly differ from each other (Fig. 3-1A-B). However, inconsistent diversity trends were observed between uninoculated and inoculated silage across sampling periods.

Forage types collected in the fall season (Fig. 3-1C-D) exhibited highly dissimilar community structures ($p<1e-04$). While there was a clear separation between H and silage, taxonomic compositions in GL and C were not significantly different from those of their inoculated counterparts. Principal coordinate analysis (PCoA) based on PhILR-transformed data revealed significant differences between uninoculated silage and inoculated counterparts (Fig. 3-1D), indicating that taxa occurring in the compared habitats were not phylogenetically related. In contrast, silage microbial communities from the spring season were compositionally similar (Fig. 3-1E) and phylogenetically related (Fig. 3-1F), while being almost all significantly different from H.

Taxonomic profiles were assessed at the genus level, concomitantly with sample LCBD indices. The H samples collected across both periods (Fig. 3-2A-B) showed high LCBD indices. The genera *Pantoea*, *Sphingomonas*, *Curtobacterium*, *Methylobacterium*, and *Pseudomonas* were variably dominant across both sampling periods. Among GL samples collected in the fall, four showed highly distinctive microbial profiles, with communities variably co-dominated by *Bacillus* and *Saccharopolyspora*, or *Weissella* and *Pediococcus*, or by *Pediococcus*, *Methylobacterium,* and *Enterococcus*. Most of the other GL samples were dominated by *Lactobacillus*, *Weissella*, *Pediococcus,* or *Lactococcus*. In the spring, GL samples with high LCBD indices showed dominance of *Serratia*, *Pseudomonas* and an unclassified *Enterobacteriaceae*, or *Pediococcus* and *Weissella* (Fig 2B). Other samples were generally dominated by *Lactobacillus*, *Weissella*, and *Enterococcus*. For GLI samples collected in the fall, the most distinctive ones showed dominance of either *Bacillus*, *Pediococcus*, or co-dominance of *Lactobacillus*, *Corynebacterium,* and *Staphylococcus*, compared with few other samples showing *Lactobacillus* as the sole dominant genus (Fig. 3-2A).

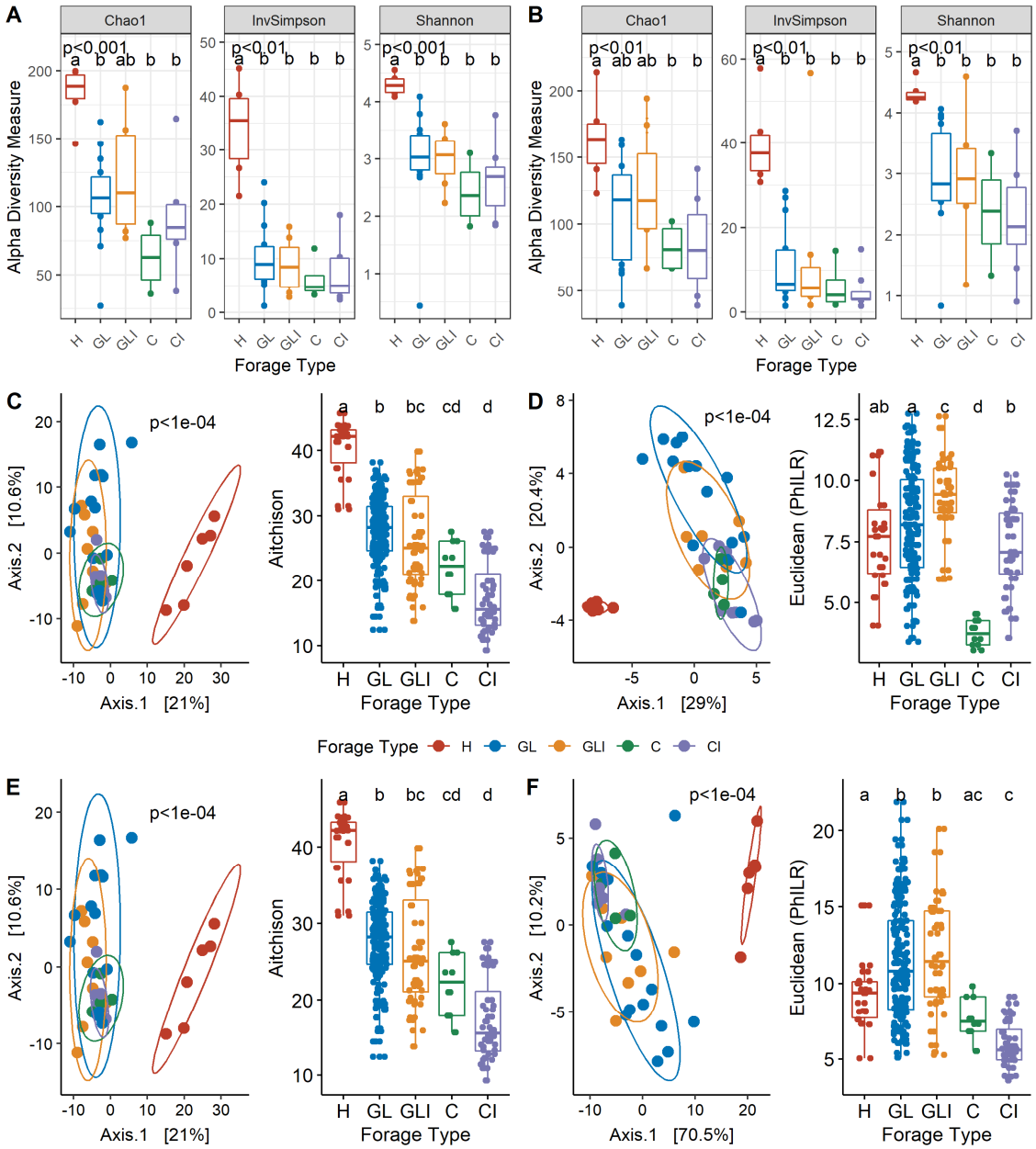


Figure 3-1: Microbial diversity according to forage type. Alpha-diversity measures of forage types in the fall (A) and the spring (B). Principal component analysis (left) on CLR (C) or PhILR (D) transformed data with corresponding post-hoc tests (right) for the fall. Principal component analysis (left) on CLR (E) or PhILR (F) transformed data with corresponding post-hoc tests (right) for the spring. p-values indicate the significance of differences between groups from the Kruskal Wallis test.

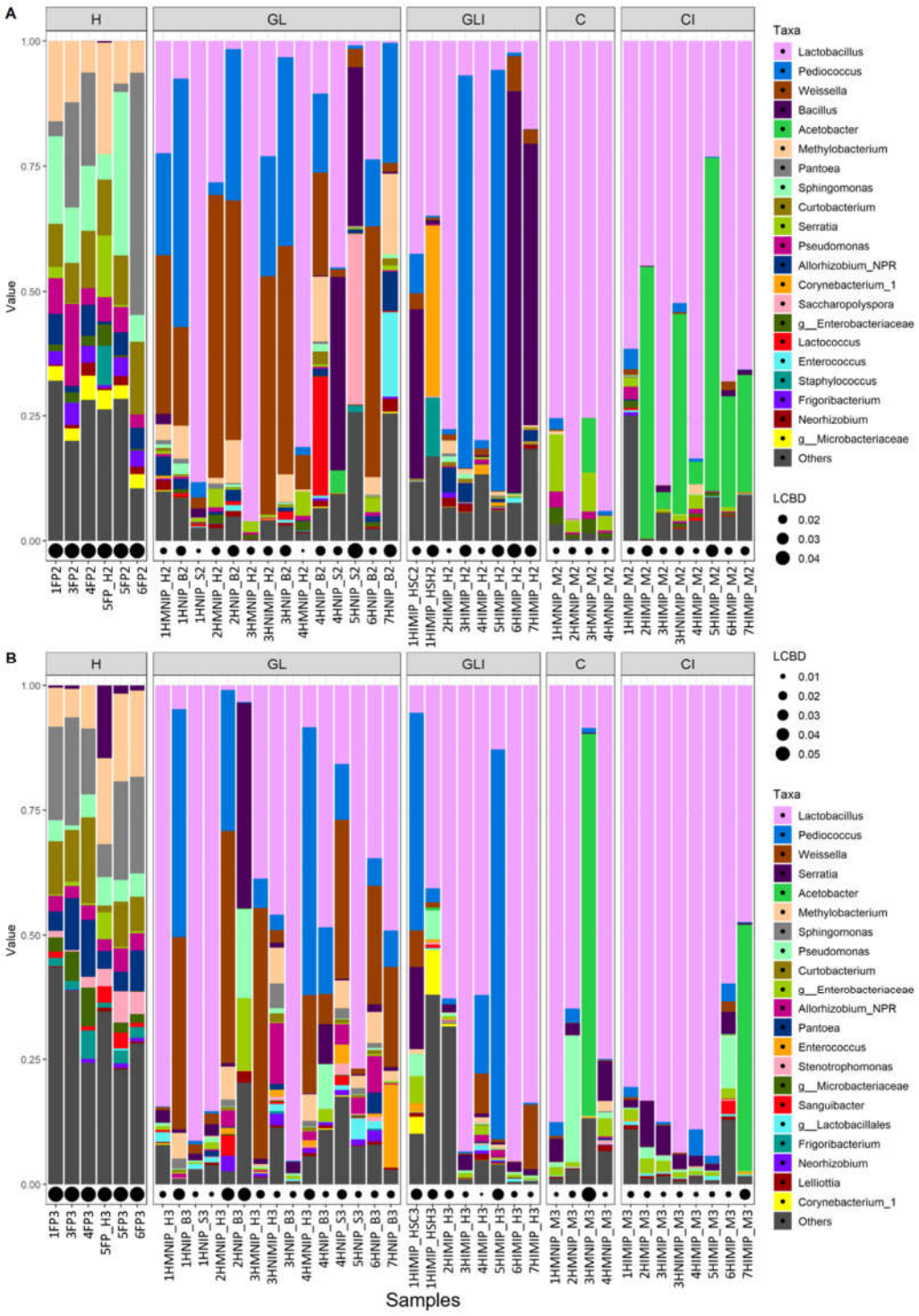


Figure 3-2: Forage microbial profiles across sampling periods. Relative abundance of the top 21 most abundant genera within forage samples collected in the fall (A) and the spring (B). LCBD indices represent the sample local contribution to the beta-diversity between groups. The higher the index, the more unique is the sample microbial profile and the higher its contribution to the beta-diversity between groups.

In the spring samples, *Lactobacillus* was the most abundant genus, followed by *Pediococcus*. Samples with high LCBD indices were dominated by either *Pediococcus* alone or in codominance with *Serratia* and *Weissella* (Fig. 3-2B). C samples collected in the fall were the most homogenous with *Lactobacillus* as the main dominant genus, although *Serratia* often occurred with considerable relative abundance (Fig. 3-2A). Among C samples collected in the spring, one exhibited high abundance of *Acetobacter*, other samples showing large proportions of *Lactobacillus* or *Pseudomonas* as dominant or subdominant genera, respectively (Fig. 3-2B). Finally, CI samples collected during fall were broadly dominated by *Lactobacillus*, although some with higher LCBD indices exhibited large proportions of *Acetobacter* as either dominant or codominant (Fig. 3-2A). For CI samples collected in the spring, *Lactobacillus* was almost the sole dominant genus, except in one sample showing *Acetobacter* as codominant (Fig. 3-2B).

A differential abundance analysis performed to identify taxa significantly enriched between uninoculated and inoculated silage revealed that in the fall, phylotypes (ASV) of *Pediococcus pentosaceus* and *Weissella* were significantly more abundant in GL compared with GLI samples, except for two GLI samples inoculated with Biotal Buchneri 500 (Fig. 3-3A). In the spring, *Lactobacillus* and some *Proteobacteria* were significantly more abundant in GLI samples (Fig. 3-3B). For corn silage, *Loigolactobacillus coryniformis* and *Lactobacillus* phylotypes exhibiting high prevalence (83-90 %) were significantly more abundant in C compared with CI samples in the fall (Fig. 3-3C). Surprisingly in the spring, while some highly prevalent (83-100 %) phylotypes of *Proteobacteria* were significantly more abundant in CI samples, phylotypes of *Lactobacillus* and *Proteobacteria* were significantly more abundant in C samples (Fig. 3-3D).

LAB including *L. buchneri* and *L. plantarum* were significantly enriched ($p<0.001$) in ensiled forages compared with H samples (Fig. 3-4A-B). However, inconsistent enrichment of LAB was noted among silage across sampling periods. Ensiling significantly reduced AAB loads compared with H ($p<0.05$), but differences between inoculated silage and uninoculated counterparts were inconsistent across both periods. Ensiling significantly reduced *Pseudomonas* levels compared with H ($p<0.05$) and loads of this group tended to increase with inoculation, as were *Enterobacteriaceae* levels. While total bacteria load tended to increase with inoculation between silage across both sampling periods, contrasting patterns of significant variations ($p<0.05$) of fungi loads were noted (Fig. 3-4A-B). We found similar patterns of bacterial load variations between forage types within PMA-treated (Fig. 3-4A-B) and PMA-free (data not shown) samples, although the latter group broadly exhibited higher load levels.

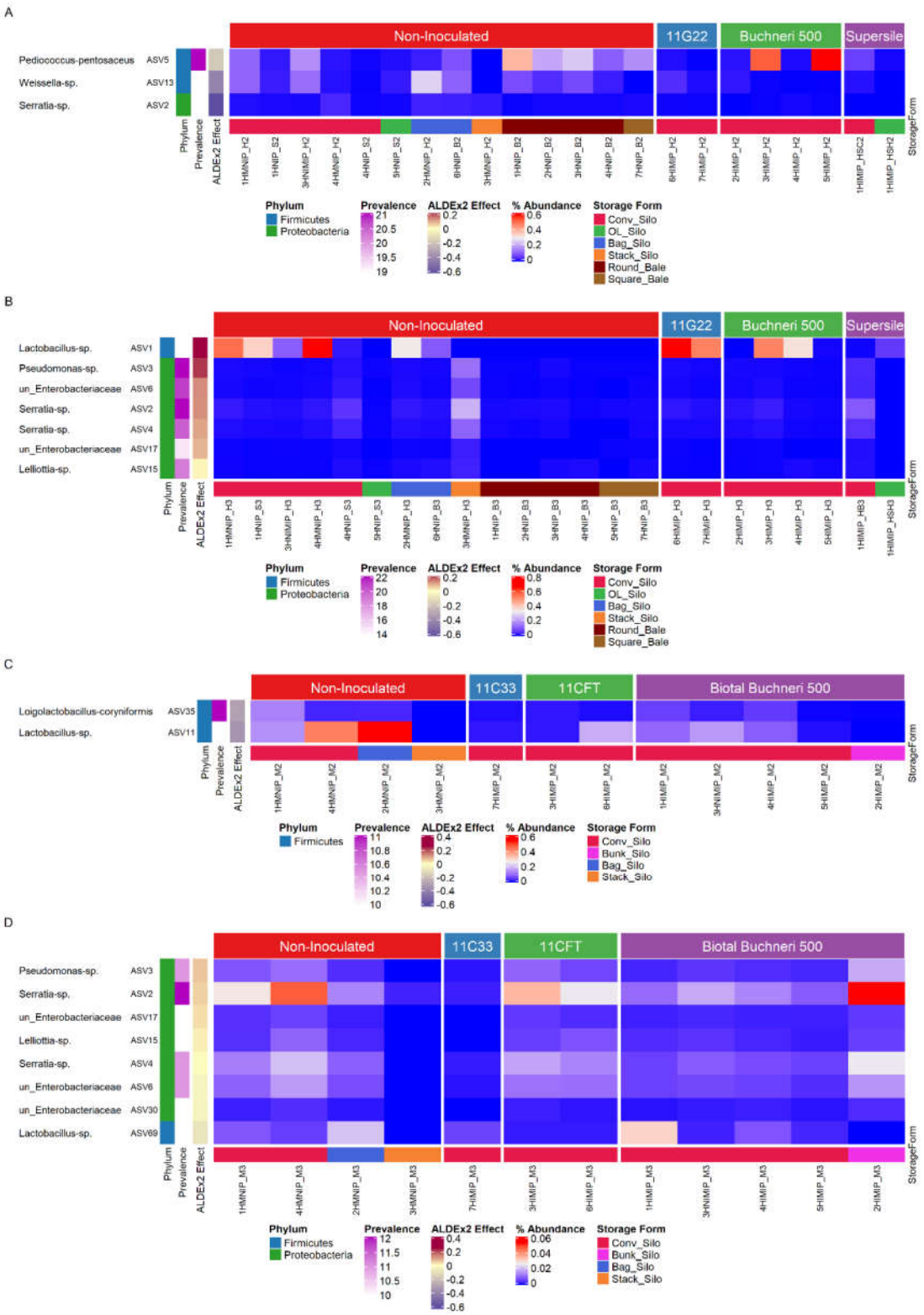


Figure 3-3: Distribution of differentially abundant ASV between uninoculated and inoculated silage. Relative abundance of ASV significantly enriched between uninoculated and inoculated grass/legume silage in the fall (A) and the spring (B). Relative abundance of ASV significantly enriched between uninoculated and inoculated corn silage in the fall (D) and the spring (E).

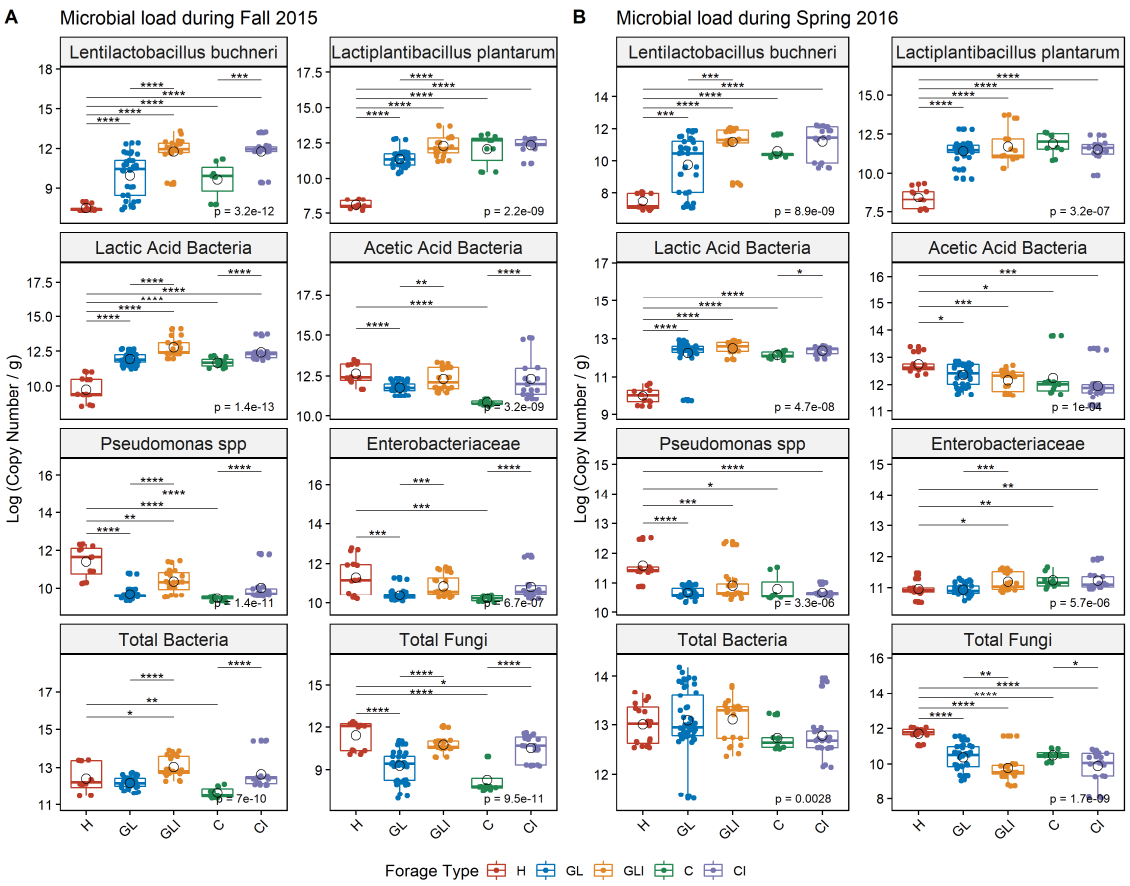


Figure 3-4: Quantification of viable microbial groups across sampling periods. Copy numbers are compared between H and each silage type, and between uninoculated and inoculated silage in the fall (A) and the spring (B) using the Kruskal Wallis test. Asterisks above boxes indicate significant differences and flag p-values from a Wilcoxon rank tests as follows: *, p < 0.05; **, p < 0.01; ***, p < 0.001; ****, p < 0.0001.

3.5.2 Abundance profile of the core phylotypes from the order *Lactobacillales* in uninoculated and inoculated silage

In addition to differential abundance testing, specific and shared LAB phylotypes were identified within ensiled forages to understand how they differentially occur between uninoculated and inoculated silage. The analysis was limited to the order *Lactobacillales* as it is the lowest taxonomic classification level gathering the genera *Lactobacillus*, *Pediococcus*, and *Enterococcus*. Broadly, of the 2980 unique ASVs composing the whole PMA-treated forage dataset, about 30.67 % (914) were assigned to the order *Lactobacillales*. From the 913 *Lactobacillales* occurring in silage, 836 were found in grass/legume silage, of which 356 were specific to GL, 223 to GLI, and 257 shared between both, while from the 241 unique ASVs that occurred in corn silage, 33 were specific to C, 123 to CI, and 85 shared between both (data not shown).

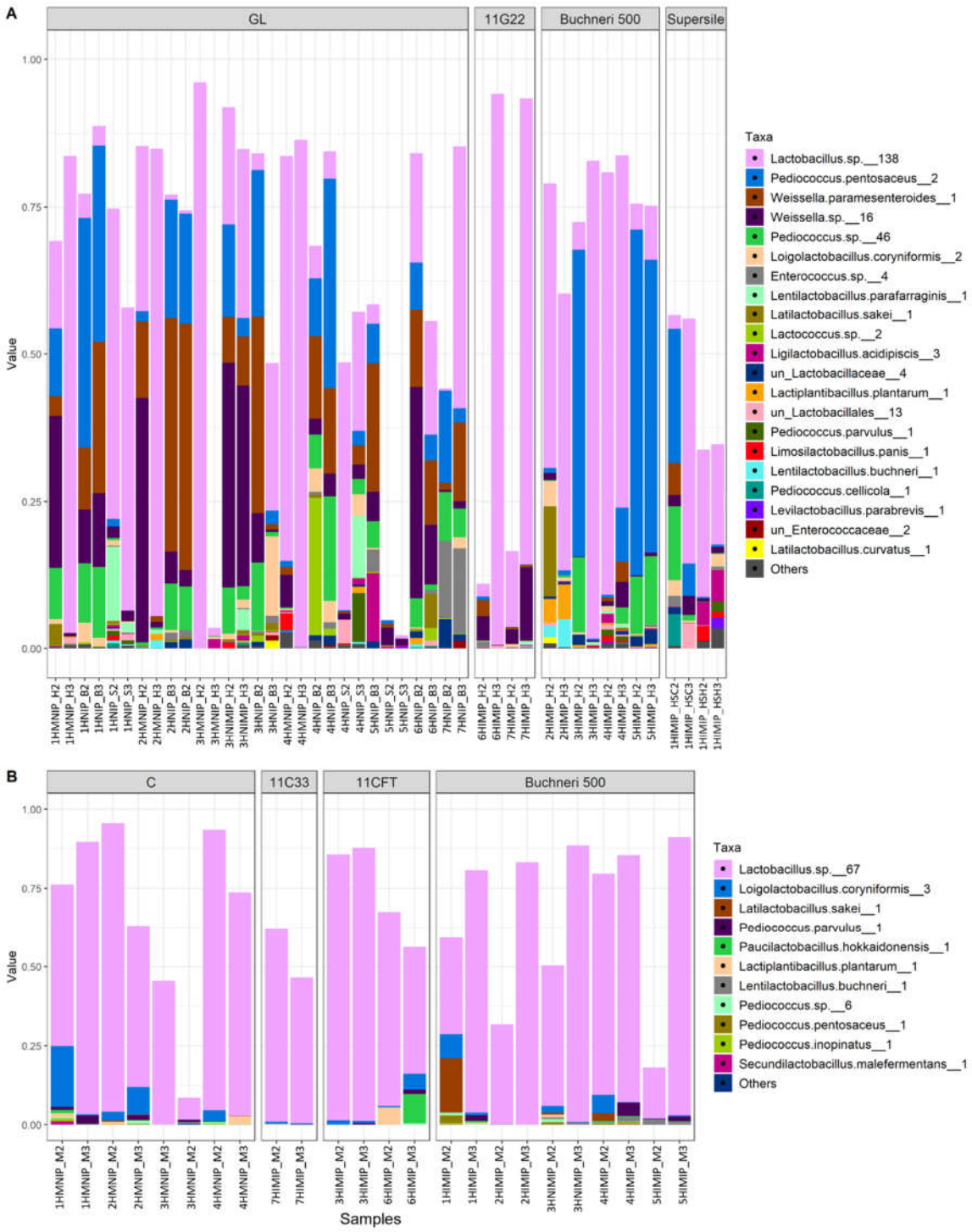


Figure 3-5: Core LAB community profiling. Relative abundance of the dominating core ASV between uninoculated and inoculated grass/legume (A) or corn (B) silage. Numeric values following taxonomic names represent corresponding ASV numbers.

The dominating core LAB among grass/legume silage (i.e., LAB occurring both in GL and GLI) totalled up to 96 % of the LAB communities and included *Lactobacillus* spp. represented by 138 phylotypes, *P.*

pentosaceus represented by 2 phylotypes, *Weissella parasenteroides* counting two phylotypes, and other *Weissella* and *Pediococcus* spp. represented by 16 and 46 phylotypes, respectively (Fig. 3-5A). For corn silage, the core LAB phylotypes in both silage types dominated the LAB communities with up to 95 % relative abundance (Fig. 3-5B), except for samples where LAB were not the dominant taxa (Fig. 3-2A-B). Within corn silage, the dominant LAB populations were composed of *Lactobacillus* spp. represented by 67 phylotypes, *L. coryniformis* represented by three phylotypes, and to a certain extent *Latilactobacillus sakei* and *Paucilactobacillus hokkaidonensis* each counting 1 phylotype (Fig. 3-5B).

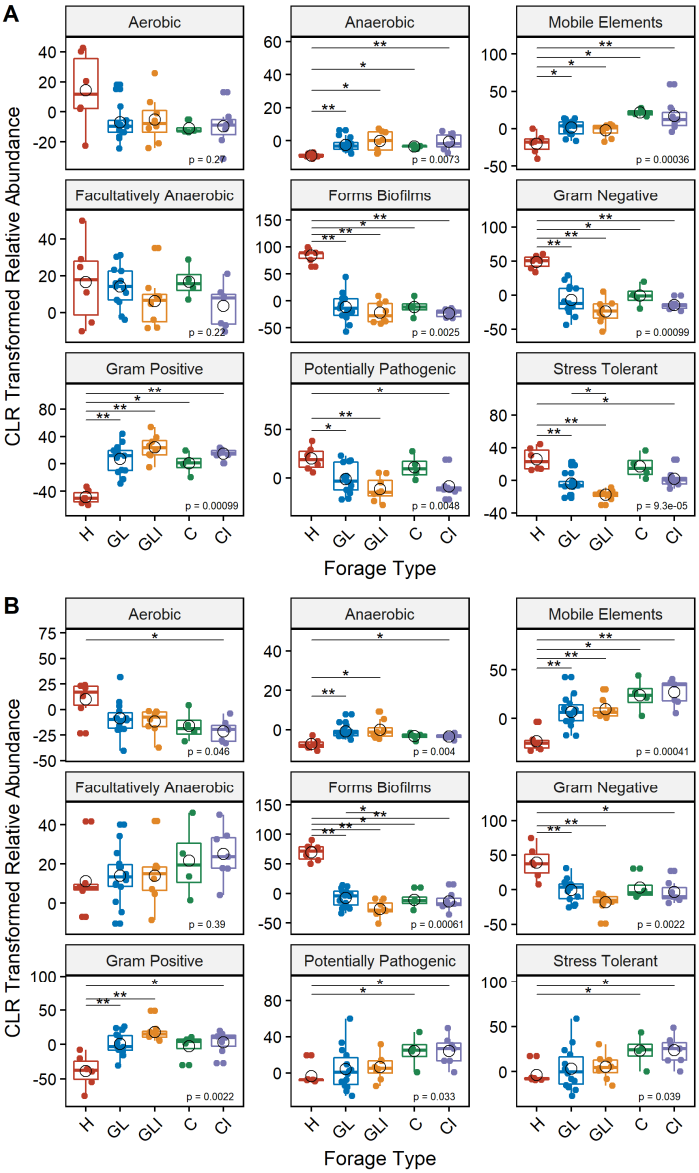


Figure 3-6: Microbial phenotypic composition predicted by BugBase. Proportions of microbial phenotypic traits compared between H and each silage type, and between uninoculated and inoculated silage in the fall (A) and

the spring (B) using the Kruskal Wallis test. Asterisks above boxes indicate significant differences and flag p-values from a Wilcoxon rank tests as follows: *, $p < 0.05$; **, $p < 0.01$.

3.5.3 Prediction of phenotypic traits and function pathways

The analysis of microbial phenotypic traits using BugBase showed that anaerobic, containing mobile elements, biofilm forming capacity, Gram-negative, Gram positive, potentially pathogenic, and stress tolerant phenotypes were differentially distributed among forage types (Fig. 3-6A-B). Across both sampling periods, biofilm forming capacity, and Gram-negative phenotypes were significantly enriched in H compared with silage (GL, GLI, C, and CI). This could be attributable to higher relative abundances of *Actinobacteria* and *Proteobacteria* in H samples for biofilm forming capacity and Gram-negative phenotypes, respectively (Fig. B1 A-B). A significant enrichment of Gram-positive phenotypes was observed in silage compared with H across both sampling periods, a difference attributable to the higher abundance of *Firmicutes* in silage (Fig. B1 A-B). Analysing the predicted function pathways, we found that biofilm forming associated features were significantly enriched in uninoculated grass/legume silage (Fig. B2-B3), while other features associated with the metabolism of macromolecules (See “Annexe B” for details) were significantly enriched in the inoculated counterparts (Fig. B4 A-B).

3.5.4 Forage physicochemical characteristics associated with microbial communities.

Forage analyses using near infrared showed that the physicochemical traits of H samples were broadly consistent across the sampling periods (Table B1). Comparing uninoculated silage with inoculated counterparts, ammonia content was significantly ($p<0.05$) higher in GLI and CI compared with GL and C silage in the spring, respectively, as were acetic acid in GLI in the spring and butyric acid in CI in the fall compared with GL and C silage, respectively (Table B2-3). The dry matter content of grass/legume silage ranged between 29 and 65 % for GL samples, and between 32 and 57 % for GLI. pH values ranged between 3.7 and 5.3 for GL samples, and between 3.8 and 4.7 for GLI. Butyric acid was detected in less than 38 % of GL samples with amounts ranging between 0.42 and 1.1 %, and in more than 62 % of GLI samples with amounts ranging from 0.42 to 0.99 % dry matter. In contrast, butyric corn silage was detected in only 25 % of C samples with amounts of butyric acid ranging between 1.09 and 1.96 % dry matter, and none among CI samples. However, corn silage exhibited lower dry matter content, ranging between 20 and 41 %, as well as lower pH values ranging from 3.5 to 3.9.

To gain insight into how forage end products and characteristics associate with microbial communities, categorical and quantitative metadata related to farms and forage samples were selectively integrated in multivariate multi-table analyses based on sPLS and CCpnA. General linear models implemented in MaAsLin2 were also used to determine multivariate associations between forage metadata and ASVs based on CLR normalized data. This analysis scheme was implemented for all forage types separately. H microbial

communities were grouped into two clusters (Fig. 3-7A, B5 A). Moisture content and positively correlated phylotypes of *Serratia*, *Enterobacteriaceae*, *Nocardioides* and *Yersinia*, as well as those negatively correlated including *Spirosoma*, *Intrasporangiaceae*, and *Methylobacterium* were the main contributors to the formation of cluster 1 (Fig. B5 B-C, B6). Magnesium content and positively (*Frigoribacterium*, *Curtobacterium*, *Pantoea*, *Rhizobiaceae*, *Microbacteriaceae*, *Allorhizobium* group and *Pantoea agglomerans*) or negatively (*Pseudomonas*) correlated phylotypes contributed the most to the formation of cluster 2 (Fig. B5 B-C, B6). Among all the taxa found to associate with the selected H variables, phylotypes of *Methylobacterium* (cluster 1) or *Curtobacterium*, *Pantoea*, *Serratia* (cluster 2) where the most prevalent and abundant (See "Annexe B" for details).

Selected taxa of the GL microbiota formed three clusters (Fig. 3-7B, B7 A). Cluster 1 was mainly driven by the variables lactic acid (LA), moisture, fatty acid (FA), crude fat (CF), and volatile fatty acid (VFA), and the positively associated taxa including phylotypes of *Lactobacillus*, *Weissella*, and *Carnobacterium* (Fig. 3-7B, B7 B-C, B8). For cluster 2, the pH and positively correlated taxa including phylotypes of *Methylobacterium*, *Sphingomonas*, *Curtobacterium*, *Allorhizobium* group, *Methylobacterium adhaesivum*, *Lactobacillus* and *Pediococcus* were the main contributors. The two variables ammonia expressed as percentage of crude proteins (CP_NH_3) and soluble crude proteins (CP_Sol), and positively associated taxa including phylotypes of *Weissella*, *Aeriscardovia*, *Lactobacillus*, *Corynebacterium*, *Pediococcus*, *Lactobacillus kefiranofaciens*, *Serratia*, *Brevibacterium* and *Brachybacterium* were the main contributors to cluster 3 (Fig. 3-7B, B7 B-C, B8). Phylotypes of *Lactobacillus*, *Pediococcus* and *Weissella paramesenteroides* were the most abundant and prevalent of all the GL selected taxa (See "Annexe B" for details).

For GLI, ASVs differentially associated with silage parameters also formed three clusters as for GL (Fig. 3-7C, B9 A). The phylotypes of *Lactobacillus*, *L. parafarraginis*, *Weissella*, *Ligilactobacillus acidipiscis*, and *Pediococcus* that positively associated with the variables CP_NH_3, ammonia (NH_3), ammonia expressed as percentage of soluble protein (SP_NH_3), and acetic acid (AA) all constituted the main contributors to cluster 1 (Fig. 3-7C, B9 B-C, B10). The variables ethanol soluble carbohydrates as percentage of dry matter (ESC_DM) and ESC as percentage of non fiber carbohydrate (ESC_NFC) positively correlated with phylotypes of *Methylobacterium*, *Pediococcus*, *Lactobacillus*, *Luteimonas aestuarii*, *Sphingomonas phyllosphaerae*, *Neorhizobium*, and *Rhodococcus* all contributed the most to Cluster 2. Cluster 3 was mainly driven by the variables LA, LA expressed as percentage of VFA (LA_VFA), and CP_Sol, together with the positively associated phylotypes of *Lactobacillus*, *Clostridiaceae*, *Enterobacteriaceae*, *Lactobacillales*, *Frigoribacterium faeni*, *Microbacteriaceae*, and *L. plantarum* (Fig. 3-7C, B9 B-C, B10). The most prevalent of the selected ASVs included phylotypes of *Lelliotia*, *Serratia*, *L. acidipiscis*, and *Weissella* (See "Annexe B" for details).

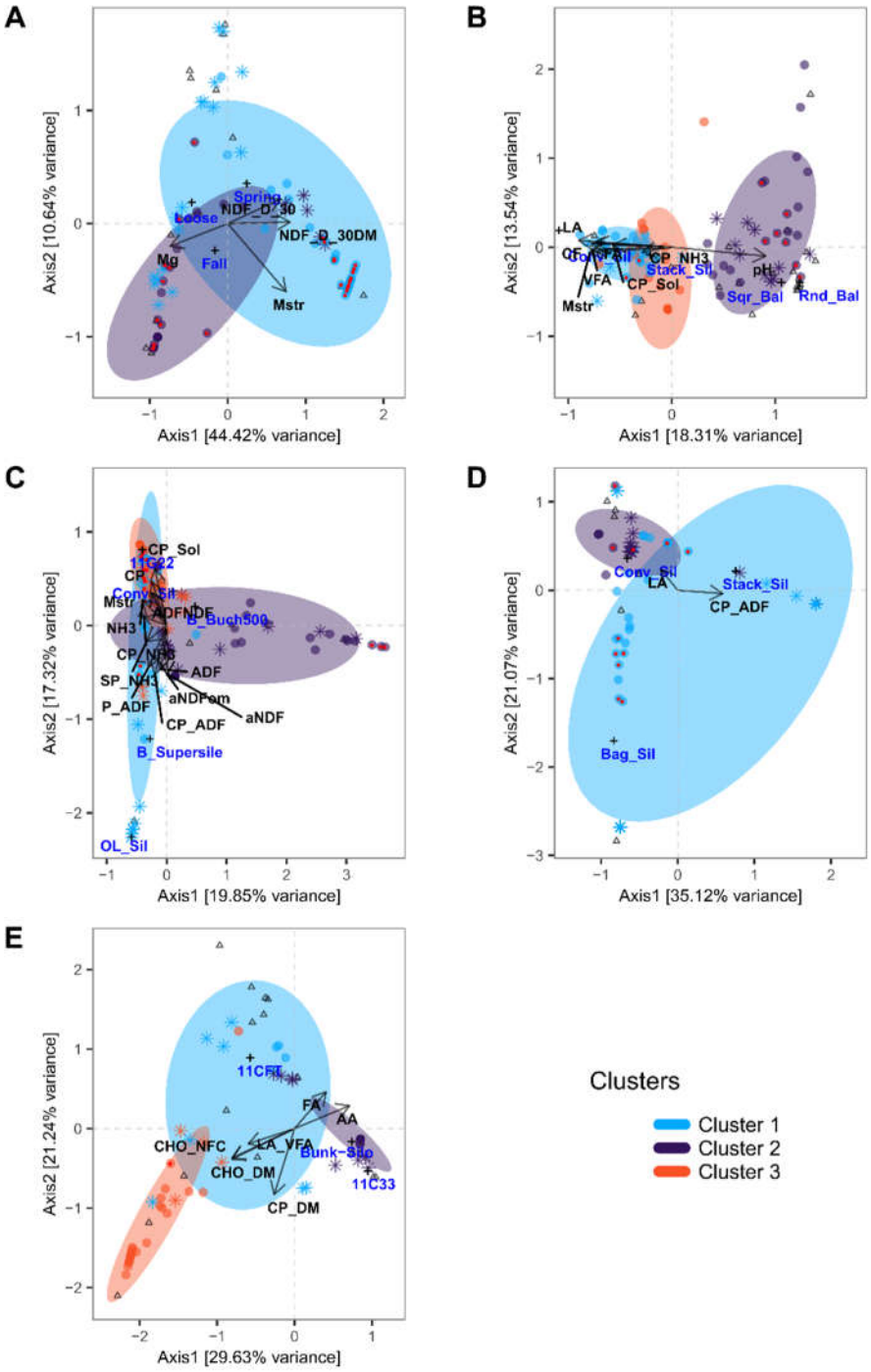


Figure 3-7: Canonical correspondence analysis. Triplots illustrating canonical relationships between bacterial ASV (round and star shaped points) and physicochemical parameters (arrows), inoculants (+), forage storage form (+), or sampling periods (+) for H (A), GL (B), GLI (C), C (D), and CI (E). Sampling periods include fall 2015 and spring 2016. Storage forms include loose, square bales (Sqr_Bal), conventional silo (Conv_Sil), bag silo (Bag_Sil), stack silo (Stack_Sil), round bale (Rnd_Bal), oxygen limiting silo (OL_Sil), and bunker silo (Bunk-silo).

Inoculants include Biotal Buchneri 500 (B_Buch500), Biotal Supersile (B_supersile), 11G22, 11C33, 11CFT. Forage physicochemical parameters include amino acids (AA), acid detergent fiber (ADF), ADF as % neutral detergent fiber (ADFNDF), amylase derived NDF (aNDF), ash-corrected NDF (aNDFom), crude fat (CF), ethanol soluble carbohydrate as % dry matter (CHO_DM), carbohydrate as % non-fibrous carbohydrate (CHO_NFC), crude protein (CP), acid detergent fiber-CP (CP_ADF), CP as % dry matter (CP_DM), ammonia as % CP (CP_NH3), soluble CP (CP_Sol), fatty acids (FA), lactic acid (LA), LA as % volatile fatty acids (LA_VFA), magnesium (Mg), moisture (Mstr), NDF digestibility at 30 hours (NDF_D_30), NDF_D_30 as % dry matter (NDF_D_30DM), ammonia (NH3), degradable protein (DP), pH (pH), ammonia as % soluble protein (SP_NH3), volatile fatty acids (VFA). Round points depict taxa selected by sPLS and star shaped points those selected by MaAsLin approaches. Round points with red centre are taxa selected by both methods. Triangles represent samples.

In the case of C, we obtained two clusters (Fig. 3-7D, B11 A) driven by both LA and acid detergent fiber-crude protein (CP_ADF). Phylotypes of *Lactobacillus*, *Acetobacter*, and *Serratia* positively correlated with LA, and those negatively correlated including *Pediococcus* and *Pc. hokkaidonensis* mostly contributed to the formation of cluster 1. Taxa mostly contributing to the formation of cluster 2 included phylotypes of *Lactobacillus* among which *Lt. sakei*, all positively associated with CP_ADF (Fig. 3-7D, B11 B-C, B12). Phylotypes of *Lactobacillus* and *Serratia* were among the more prevalent and abundant (See “Annexe B” for details).

Finally for CI, the selected ASVs formed three clusters (Fig. 3-7E, B13 A). The parameters AA and FA mainly contributed to clusters 1 and 2. While a few phylotypes of *Lactobacillus* positively correlated with AA and FA mainly drove cluster 1, those of *Comamonas jiangduensis*, *Ameyamaea*, and *Acinetobacter gerneri* were the main contributors to cluster 2 (Fig. 3-7E, B13 B-C, B14). For cluster 3, the parameters carbohydrates (CHO_DM and CHO_NFC), CP_DM, and LA_VFA and positively associated phylotypes of *Lelliottia*, *Enterobacter*, *Raoultella terrigena*, *Enterobacteriaceae*, and *Vagococcus fluvialis* were the main drivers (Fig. 3-7E, B13 B-C, B14). Among selected taxa, phylotypes of *Lactobacillus*, *Acetobacter*, and *Serratia* were the most abundant and prevalent.

3.5.5 Molecular ecological network analyses

To investigate how silage bacteria co-occur in the presence or absence of inoculants, a network was constructed for each silage type. The analysis of topological properties revealed that the GLI network had a higher average degree (avgK) and a lower average geodesic path (distance between nodes) than that of GL, thus appearing more complex and denser (Table 3-2). The 80 nodes composing the GL network totalled 8 modules (group of ASVs sharing more links among themselves than with others outside the group) and 405 links, of which only 18.8 % were positive (Fig. B15 A). It appeared that some *Lactobacillales* including phylotypes of *Pediococcus, Enterococcus, Weissella*, *L. sakei*, *L. coryniformis*, and unidentified *Lactobacillaceae* co-occurred with those of *Proteobacteria* including *Methylobacterium*, *Pantoea*, *Sphingomonas*, *Stenetrophomonas*, and *Allorhizobium*, or *Actinobacteria* comprising *Rhodococcus* and *Curtobacterium*

(modules 1, 3 and 6). Module 5 exclusively composed of *Lactobacillaceae* exhibited co-occurrence between phylotypes of *L. buchneri* and those of lactobacilli (Fig. B15 A). On the other hand, the GLI network was composed of 47 nodes grouped into two modules, and 848 links from which only 7.2 % were positive (Table 3-2). In contrast to that of GL, the GLI network involved higher amounts of positive relationships among *Firmicutes* (Fig. B15 B). Curiously, phylotypes corresponding to *L. buchneri* and *L. plantarum* were found in distinct modules, sharing no relationships; though both co-occurred with the same phylotype of *Lactobacillus*. In addition, fewer phylotypes of *Proteobacteria* sharing positive interactions with *Firmicutes* were observed. Another particularity of the GLI network is the co-occurrence between some phylotypes of *Weissella* and those of *Bacillus* (see "Annexe B" for details). Analysing taxa topological roles, we identified 47 keystone phylotypes distributed as network hubs and connectors in the GL pMEN, of which 64 % were *Firmicutes*, 34 % *Proteobacteria* and the remaining *Actinobacteria* (Table B4). Among network hubs were phylotypes of *Weissella* and *Sphingomonas*, while connectors mostly included *Firmicutes* such as *L. plantarum*, *L. buchneri*, *Bacillus*, *Weissella*, *Lactococcus* and other lactobacilli phylotypes (see "Annexe B" for details). Conversely, no keystone species were observed in the GLI network. However, compared with the GL network, there was a shift of topological roles in GLI, so that nodes formed a cluster around the value 0.5 along the Pi axis (Fig. 3-8). Consequently, all keystone phylotypes of the GL network also found in GLI changed their topological roles to peripherals (Fig. 3-8).

Table 3-2: Topological properties of the empirical pMENs in grass/legume and corn silage microbial communities and their associated random pMENs

	Empirical networks						**Random networks**		
Silage type	Similarity threshold	Network size	Average degree (avgK)	Average path	Average clustering coefficient	Modularity (No. of modules)	Average path	Average clustering coefficient	Modularity
GL	0.81	80	10.1	2.4	0.5	0.2 (8)	2.1 ± 0.03	0.5 ± 0.02	0.2 ± 0.006
GLI	0.31	47	36.1	1.2	0.8	0.02 (2)	1.2 ± 0	0.8 ± 0.002	0.01 ± 0.01
C	0.31	25	17.4	1.3	0.8	0.0 (1)	1.3 ± 0	0.8 ± 0.004	0.01 ± 0.02
CI	0.31	55	37.3	1.3	0.8	0.03 (2)	1.3 ± 0	0.8 ± 0.003	0.03 ± 0.01

For corn silage, microbial communities formed a larger and more complex network in CI than in C (Table 3-2). The 25 nodes composing the C network formed a single module, involving 217 links of which 22 % were positive (Fig. B15 C). All *Proteobacteria* phylotypes (*Serratia*, *Pseudomonas*, and unclassified *Enterobacteriaceae*) involved in this network positively interacted with each other and co-occurred with *Firmicutes* including *L. coryniformis*, and two *Lactobacillus* phylotypes. Neither *L. buchneri* nor *L. plantarum* were

involved in this network. The CI network totalled 55 nodes composing two modules, and 1026 links of which 9.8 % were positive (Fig. B15 D). Unlike in the GLI network, *L. buchneri* and *L. plantarum* found in different modules (modules 1 and 2, respectively) co-occurred and were both involved in positive relationships with *P. parvulus* and some lactobacilli. However, most phylotypes of *Proteobacteria* including *Serratia*, *Pseudomonas*, *Acetobacter*, and *Yersinia* variably co-occurred among themselves and with some lactobacilli (See "Annexe B" for details). No keystone phylotypes were found in the C and CI networks.

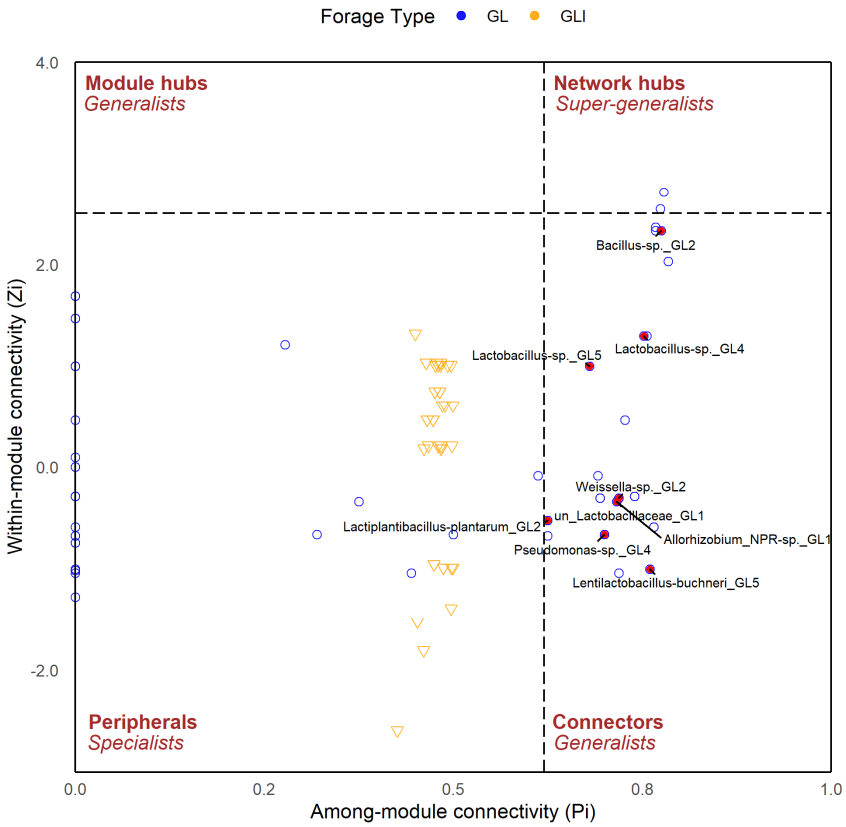


Figure 3-8: Distribution of network topological roles of grass/legume silage. Labelled ASV (red points) depict those that exhibited different topological roles in GL compared to GLI.

3.6 Discussion

In this study, we implemented a viability high-throughput sequencing approach combined with viability-PCR [43,251] to provide a comprehensive and comparative analysis of the viable microbial ecology of hay and grass/legume or corn silage produced with or without inoculants at commercial farm-scale facilities.

Recently, Daniels et al. [247] analysed the viable microbiota of commercial Meadow and Italian ryegrass hay and revealed *Proteobacteria*, *Cyanobacteria*, *Bacteroidetes*, and *Actinobacteria* as the predominant phyla. In our study, *Cyanobacteria* were detected as the rarest taxa and were discarded from the dataset upon

abundance filtering, while the phyla *Proteobacteria*, *Actinobacteria*, and *Bacteroidetes* were respectively the more prevalent and abundant in hay. However, the differences found among hay community profiles could be attributable to fluctuations in abundance of predominant genera, including *Sphingomonas*, *Methylobacterium*, *Pantoea*, *Curtobacterium*, and *Pseudomonas*. Behrendt et al. [271] analysed the microbial community of grass phyllosphere using a culture-dependent method and identified *Pseudomonas*, *Stenotrophomonas*, *Pantoea*, *Clavibacter*, and *Curtobacterium* as the predominant genera. Although not among the most abundant, *Stenotrophomonas* and *Clavibacter* were also detected in our study, thus indicating that the microbiota of hay might consistently reflect the epiphytic communities of plants at harvest. Moreover, except *Stenotrophomonas*, we found that specific phylotypes of the above-mentioned genera correlated with hay moisture content, suggesting bacterial growth in less dried hay. The observed differences between H samples could therefore be also explained by various environmental and farmers' management factors including, but not limited to, plant species, management practices, geographical location, climatic conditions, moisture concentration at harvest, and storage form [240], that drive the incidence and abundance of epiphytic microorganisms on plants before harvest or during processing. Although LAB were not among the dominant taxa, their prevalence and abundance were found to vary between H samples and across sampling periods as revealed by a concomitant study on the same forage samples, where Gagnon et al. [22] identified LAB communities through culture-based techniques. *W. paramesenteroides/thailandensis*, *P. pentosaceus*, *L. casei/paracasei*, *Lc. mesenteroides/pseudomesenteroides*, *Lpb. pentosus,* and *E. casseliflavus/gallinarum/faecium* were found as the predominant cultivable LAB in hay. In this study, the detection of *Weissella*, *Lactococcus*, *Enterococcus*, and unidentified *Lactobacillales* as the sole representatives of LAB in hay might be due to the occurrence of other LAB under the detection threshold. For survey studies, this emphasizes the relevance of combining both culture-dependent and high-throughput sequencing approaches to deepen our understanding of microbial community compositions and functions. However, none of the LAB were associated with hay moisture content. Despite the observed differences among hay samples, their microbial composition and structure clearly discriminated them from those of silage across both sampling periods, revealing haymaking and ensiling as strikingly distinct processes that differentially alter the epiphytic microbiota of fresh forage plants.

Grass/legume forage has higher buffering capacity and lower WSC content than corn forage, offering different habitats for microorganisms [170]. These factors impact the rate of pH decrease during the first stages of fermentation [170,272], thus differentially modulating the growth of microorganisms depending on their initial abundance on the pre-ensiled forages. This could explain the differences in phylogenetic composition and community structure between grass/legumes and corn silage observed in our study. We also found that most GL silage was generally dominated by *Lactobacillus* alone or in co-dominance with either *Weissella*, *Pediococcus* or both, while few samples inconsistently exhibited co-dominance of *Bacillus*, *Saccharopolyspora*, *Lactococcus*, *Serratia*, *Methylobacterium*, or *Enterococcus* across both sampling periods. Previous studies have

shown that LAB, preferably lactobacilli are the main microorganisms expected to dominate the microbiota of good-quality silage [245]. High prevalence and relative abundance of *Weissella*, *Pediococcus*, *Enterococcus*, and *Lactococcus* were also found by Gagnon et al. [22] using a culture-dependent approach. Besides the high buffering capacity, this observation presumably illustrates grass/legume silage for which the fresh forage phyllosphere contained insufficient amounts of *Lactobacillus* to lead the first stages of the fermentation process [207,273]. Such conditions might also favour the growth of undesirable bacteria such as *Bacillus* and *Serratia* [245], as revealed in our study. Although several studies have reported the occurrence of *Methylobacterium* in silage [274,275], none have described its role in the fermentation process. The genus *Saccharopolyspora* represented by three phylotypes of *Saccharopolyspora rectivirgula* (formerly *Micropolyspora faeni*) and that has been associated with moist hay, compost, or straw [276,277] is a thermophilic *Actinobacteria* identified as a major cause of extrinsic allergic alveolitis (farmer's lung disease) in dairy barns [278,279]. The occurrence of this pathogen as co-dominant bacteria in GL silage and in lower abundance in GLI samples suggests its fermentative capability. Although this pathogen has already been identified in corn silage [280], no studies had reported its occurrence in grass/legume silage. In the case of C silage, the microbiota was dominated by either *Lactobacillus*, or *Acetobacter,* or both, while *Pseudomonas* and *Serratia* sporadically occurred with considerably high relative abundances. Our results corroborate those of Guan et al. [281] who found *Lactobacillus* and *Acetobacter* as predominant bacteria in naturally fermented corn silage. These authors also highlighted the inconsistent incidence of *Acetobacter* between laboratory- and large-scale bunker silos. In the current study, two phylotypes of this genus were detected in C silage processed in stack silos. Interestingly, 15 phylotypes of *Acetobacter*, among which some with higher abundance levels, were identified in CI silage that were inoculated with either Biotal Buchneri 500, 11CFT, or 11C33, and were stored in stack, bunker, or conventional silos (concrete-stave silos). These results indicate that species of *Acetobacter* might outcompete LAB even with the addition of inoculants, or that oxygen infiltration into the forage might have favoured their proliferation. Moreover, viable-PCR analyses confirmed inconsistent enrichment of microbial load including that of LAB, AAB, *Pseudomonas*, *Enterobacteriaceae*, total bacteria, and total fungi across sampling periods in inoculated silage. We found similar patterns of bacterial load variation between forage types for PMA-free samples (data not shown) that broadly exhibited higher loads compared with PMA-treated samples. These observations suggest that prevailing weather conditions (temperature fluctuation) might impair the efficiency of the inoculant during ensiling and consequently favour undesirable AAB and acid tolerant *Proteobacteria* in mature silage. Except for LAB, the effects of these microorganisms during ensiling is not clear [282–284], and further researches are needed to better understand their function and interplays with LAB during silage fermentation. Although other studies had reported a high prevalence of *Pseudomonas* and *Serratia* throughout the ensiling process of corn silage [282,285], like other *Proteobacteria*, their effects on the silage fermentation process are not well understood.

We found that phylotypes dominating the *Lactobacillales* community in uninoculated and inoculated silage were members of the core LAB, clearly indicating the ubiquity of inoculant species in the phyllosphere. This suggests that inoculant related species, if occurring in sufficient amounts relative to the epiphytic communities, microbial additives would not be necessary to obtain quality silage. The recently reported successful ensiling using a transplanted epiphytic microbiota as the sole source of microorganisms supports this idea [286,287]. Moreover, uninoculated GL and C silage with good fermentation profiles were observed in this study. Conditions for successful natural fermentation are not always met and ensiling without additives may be therefore most often associated with increased risk of economic losses [243,244]. Consequently, microbial additives employed to dominate the communities, minimize the occurrence of undesirable microorganisms, and drive the fermentation process [288] are generally recommended [243]. However, if this is particularly true for grass or legume silage, microbial additives might not necessarily improve the quality of corn silage, as recently revealed by a meta-analysis [168,289]. The current study showing dominant or codominant *Acetobacter* in CI silage particularly in the fall corroborates this finding, though some C silage exhibited undesirable microbial profiles, mostly in the spring.

Our study showed that in GL and GLI silage, specific phylotypes of *Lactobacillus*, *Weissella*, *Lelliottia*, *Serratia*, and an unidentified *Enterobacteriaceae* were positively associated with LA and moisture contents, and negatively associated with pH. Although some *Proteobacteria* can produce lactic acid [245], acetic acid is their primary product during ensiling [290]. Therefore, the correlation of corresponding phylotypes with LA could result from their growth and subsequent acid production in the early stages of fermentation. The relatively low abundance of these phylotypes in mature silage could be due to the inhibitory effect of accumulating lactic acid and lower pH level during ensiling [291]. We also found that phylotypes of *P. pentosaceus*, *W. paramesenteroides*, *Lactobacillus*, *Pediococcus*, *Methylobacterium*, and *Lactococcus* positively correlated with pH, and were negatively associated with LA, indicating their sensitivity to lower pH levels or higher amounts of LA. Similar results were obtained by Ogunade et al. [249,274]. On the other hand, while phylotypes of *Pediococcus* were codominant in some GL samples, this genus largely dominated three GLI samples inoculated with Biotal Buchneri 500 or Biotal Supersile that include *P. pentosaceus* and *P. acidilactici* in their formulation, respectively. If the genus *Pediococcus* is known to dominate the microbiota in the early stages of fermentation although inconsistently [173,207,274], the factors that favour its dominance in mature silage are not well understood [292]. In addition to *Methylobacterium*, other *Proteobacteria* including *Pantoea*, *Sphingomonas*, and *Stenotrophomonas*, and *Actinobacteria* including *Rhodococcus* and *Curtobacterium* co-occurred with most of LAB phylotypes within the GL, as have phylotypes of *Pseudomonas* and *Acetobacter* within the C communities. Higher abundance of these *Proteobacteria* were observed in samples wherein lactobacilli were not predominant. This suggests that co-occurring LAB probably do not induce rapid decrease in pH during ensiling, as some lactobacilli strains more adapted for ensilage would have done. The decreased rates of co-occurrence between

Firmicutes and *Proteobacteria* observed in GLI and CI communities supports this hypothesis. The observed phenomenon could be explained by dominating lactobacilli from inoculation; as more lactic acid is subsequently produced during ensiling, most *Proteobacteria* are inhibited and consequently, increased patterns of co-exclusion appear to the detriment of co-occurrence.

However, the enhancement of positive interactions between LAB due to inoculants certainly prompt, in addition to related species, other acid-tolerant bacteria such as *Bacillus* in GLI, *Acetobacter* in CI, or *Serratia* in both. This might explain the observed co-occurrence between phylotypes of *Weissella* and *Bacillus* in GLI, or *Lactobacillus* and *Acetobacter* in CI communities. Although *Acetobacter* has been frequently identified in corn silage [281,283,293], there is still no consensus on the role played by this genus in silage. While *Acetobacter* was reported as aerobic, du Toit et al. [294] demonstrated the effective survival of *Acetobacter pasteurianus* under anaerobic conditions. In their attempts to validate the finding that AAB including *A. pasteurianus* might increase silage aerobic stability through acetic acid production, a theory refuted by a previous contradictory finding on the capabilities of these bacteria to initiate aerobic spoilage [295], Queiroz et al. [296] found no effect on the aerobic stability following silage inoculation with the bacteria. In our study, the genus *Acetobacter* occurred in all the 16 CI samples, in which it dominated or co-dominated the microbial community of more than 37 % without considerable effects on fermentation characteristics. Hence, the conditions under which AAB, specifically *Acetobacter* species, could drive the silage fermentation process, or improve aerobic stability, or even initiate spoilage are not clear. Likewise, bacilli are known to produce butyric, acetic, or lactic acids, as well as antibacterial substances. However, beneficial effects of *Bacillus* have been reported [245].

We found that the addition of inoculants in grass/legume and corn silage drastically changed bacterial interconnection patterns compared with uninoculated counterparts, resulting in increased network density and complexity levels, as well as in a modified modularity and taxa topological roles. As suggested by Ma et al. [297], in addition to beta diversity, microbial co-occurrence networks could be used to characterize community assemblage depending on the environment. The observed network modules that have been interpreted as microbial niches [298,299] generally contained desirable and undesirable bacteria interconnected with positive or negative links. While positive interactions among bacterial phylotypes might indicate cooperation, nutritional cross-feeding, co-colonization, or co-survival in similar environments, negative associations might result from bacteriocin or other substance production, competition, changing environment, or overpopulation of a niche [300]. These modules could reflect heterogeneities of fermentation processes undergone during ensiling or reveal the main players of that fermentation as pictured at feed-out of mature silage community composition. If keystone species identified in the GL network can be essential to its stability [298], the interpretation of their ecological relevance is not evident [299,301]. In the context of silage, thanks to the plethora of studies conducted on the fermentation process and subsequent effects on animal performance [25,26,240], there is no doubt of

the type of microbial community expected in a successful mature silage, although this is not always obtained despite inoculation. For instance, our study revealed phylotypes of *L. buchneri* and *L. plantarum* as keystone species of the GL network that do not show any type of cooperation with undesirable bacteria. This finding confirms the ecological importance of the two taxa from which specific strains are currently used as inoculants. Other keystone species from the same network among which are phylotypes of *L. coryniformis*, *Weissella* sp, *Lactobacillus* spp. were found to cooperate with at least one undesirable bacterial phylotype. Isolates of such keystone species although belonging to the LAB community could not be theoretically selected as candidate inoculants. Obviously, keystone species such as *Serratia* spp. *Pantoea* spp. or *Pseudomonas* spp. are undesirable in silage and therefore illustrate key taxa to inhibit or suppress during fermentation. On the other hand, phylotypes of the genus *Methylobacterium* of which some were identified as keystone species shared positive and negative interactions with other bacteria regardless of the phylum they belong to. Species of the genus *Methylobacterium* are methanotrophic bacteria commonly associated with pre-ensiled forage plants [249,302]. In this study, the positive correlation of this genus with pH is in accordance with its neutrophilic characteristics. Since *Methylobacterium* species are aerobic [302], their occurrence in GL samples as dominant or co-dominant taxa might be explained by the presence of oxygen during ensiling. Rigorous experiments should therefore be carried out to link keystone species derived from co-occurrence network topological roles to a particular role in the ecosystem dynamics and stability of preserved forages. In this study, the observed disparities of taxa occurrence, abundance, or dominance, as well as differential associations with forage metadata within and across the identified forms of forage storage depict the relevance of epiphytic microbial communities, microbial additives, and management conditions on the end products and microbiome structures of mature silage, and ultimately on their aerobic stability after feed-out. We also demonstrated how bacterial occurrence is highly variable, particularly LAB communities in farm-scale mature silage. Integrating silage associated parameters mentioned above in a time-varying network analysis approach [301,303] to decipher the temporal variations of microbial interactions would help fill the gaps in the current knowledge of microbial interplays and complex succession throughout ensiling.

3.7 Conclusion

In summary, hay microbiota characterized with high abundance of *Sphingomonas*, *Methylobacterium*, *Curtobacterium,* and *Pantoea* is significantly different from that of ensiled forages. The use of inoculants at commercial farm-scale facilities may unpredictably affect the microbiota composition of mature silage. While LAB were underrepresented in hay, they were inconsistently enriched in inoculated compared with uninoculated silage and were ubiquitous, thus probably mostly epiphytic instead of from commercially-made additives. The microbiota of grass/legume silage was variably dominated or co-dominated by *Lactobacillus* or *Pediococcus*, specifically exhibiting higher abundance of *Weissella* or *Bacillus* in uninoculated and inoculated silage,

respectively. On the other hand, *Lactobacillus* and *Acetobacter* inconsistently dominated the microbial communities of corn silage regardless of inoculation. Besides microbiome composition and structure, the analysis of co-occurrence and co-exclusion patterns among community assemblages clearly distinguished uninoculated from inoculated silage. Our study provides a better knowledge of how inoculants used for ensiling modulate bacterial communities populating preserved mature forages at feed-out in commercial dairy farms. Further investigations integrating management practices and silage physicochemical parameters with microbial dynamics and interactions throughout silage fermentation and post-feedout periods are needed to fully understand biological processes involved for high quality silage.

3.8 Acknowledgements

We are grateful to the Natural Sciences and Engineering Research Council of Canada (Ottawa, ON, Canada), Novalait (Quebec, QC, Canada), Agriculture and Agri-Food Canada (Ottawa, ON, Canada), and the Fonds de recherche du Québec – Nature et technologies (Quebec, QC, Canada) for their financial contribution. We also thank the Regroupement de recherche pour un lait de qualité optimale (Op+lait) for granting a scholarship to Alexandre Kennang during his study program. We express our thanks to Robert Berthiaume from Lactanet (Sainte-Anne-de-Bellevue, Canada), for his kind assistance in farm selection, and to the dairy farmers who agreed to collaborate. We sincerely thank Marie Verheyde, Myriam Laberge, and Halimatou Diallo for their kind assistance.

Chapitre 4 – Relation entre le type de fourrages, l'utilisation d'inoculants et le microbiote du lait cru

Association between forage combination, use of silage inoculant, and raw milk microbiota

Alexandre J. K. Ouamba [1,2], Mérilie Gagnon [1,2], Thibault Varin [1], P. Yvan Chouinard [2,3], Gisèle LaPointe [2,4], Denis Roy [1,2]

[1]Département des sciences des aliments, Laboratoire de génomique microbienne, Université Laval, 2440 boulevard Hochelaga, Québec G1V 0A6 Canada

[2]Regroupement de recherche pour un lait de qualité optimale (Op+Lait), 3200 rue Sicotte, Saint-Hyacinthe J2S 2M2 Canada

[3]Département des sciences animales, Université Laval, 2425 rue de l'Agriculture, Québec G1V 0A6 Canada

[4]Department of Food Science, University of Guelph, 50 Stone Road E, Guelph N1G 2W1 Canada

4.1 Résumé

La gestion des fourrages dans les fermes influence la qualité du lait cru. Cependant, les taux de transfert microbiens du fourrage au lait sont méconnus. Cette étude examine la relation entre la composition microbienne des fourrages et celle du lait cru afin d'améliorer nos connaissances sur les bactéries viables transférées au lait. Les rations alimentaires à base de fourrages étaient constituées de foin (H), d'ensilage d'herbe/légume non inoculé (GL), d'un mélange de GL et d'ensilage de maïs non inoculé (GLC) ou inoculé (GLCI), ou d'un mélange d'ensilage d'herbe/légume et de maïs inoculés (GLICI). Deux séries d'échantillonnage ont été effectuées dans 24 fermes laitières. Des différences significatives, principalement dues aux protéobactéries, ont été observées entre les échantillons de lait des fermes GLC et GLICI. Bien que la contamination du lait par les bactéries des ensilages semble aléatoire, ces deux matrices peuvent partager jusqu'à 31 % des phylotypes identifiés.

4.2 Abstract

Forage management on dairy farms can impact on the quality of raw milk. However, little is known on the microbial transfer rates from forage to milk. This study examined the community composition and structure of both cow forage and milk obtained from the same farm to improve our knowledge of the viable bacteria from forage that contaminate raw milk. Forage rations were composed of either hay (H) or grass/legume silage (GL) as the only forage type, or a combination of grass/legume and corn silage uninoculated (GLC) or inoculated (GLCI or GLICI). Samples of the forage types composing the ration and the associated raw milk from 24 dairy farms were collected twice (in the fall and the spring) and analysed using 16S rRNA gene-targeted amplicon

sequencing coupled with treatment with propidium monoazide to account for living cells. Three community types separating H, GL, and GLICI forages were identified. While the H community type was co-dominated by *Enterobacteriaceae*, *Microbacteriaceae*, *Beijerinckiaceae*, and *Sphingomonadaceae*, GLC and GLICI associated communities showed high relative abundances of *Leuconostocaceae* and *Acetobacteraceae*, respectively. Raw milk samples were not grouped in the same way, but GLC milk was significantly different from that of GLICI across both sampling periods. The differences observed between these raw milk microbiota groups were driven by *Enterobacteriaceae* and *Proteobacteria*, instead of lactic acid bacteria as it would have been projected specifically for the GLICI group for which the associated forage contained silage prepared with commercial inoculants. Indeed, a clearly defined pattern of raw milk contamination by bacteria from silage was not observed. Of the 113 phylotypes that were shared between H, GLCI, GL, GLICI, and GLC forage rations and corresponding milk, bacterial transfer rates at the level of amplicon variants were estimated at 18, 19, 21, 30, and 31 %, respectively. These results show the relevance of forage in the total mixed ration as one of the sources of bacteria that contaminate milk on farm.

4.3 Introduction

The microbiological quality of raw milk is essential for its safety and processability. On dairy farms, the complex community of raw milk [5] gradually builds up as the milk is collected from the mammary gland of the cow [304,305] through the teat canal and the milking equipment [35,306] to a cooled bulk tank. Factors inherent to the cow such as the health status of the udder [52,307] or the lactation stage [17,18], or from environmental origin including air, pasture, faeces, bedding, teat surface, water, and feed [5,19,143,308] as well as management practices [309], have been found to influence the occurrence of microorganisms in raw milk. The impact of seasons and weather conditions on changes in the milk microbiota throughout the year has been demonstrated [6,12,54]. Milk refrigeration upon milking as it is recommended favours the proliferation of psychrotrophic bacteria [7,43,70,310], but also mesophiles including *Lactococcus*, *Enterococcus*, *Streptococcus*, or *Lactobacillus* that can withstand temperatures as low as 4°C [7,43,70,189]. Manifestly, microbial dynamics persists throughout the transport chain to such an extent that the raw milk collected from farms by conveying trucks shows a distinct community structure from that in the storage silos at processing plants [13,18,54]. Despite the observed differences, most of the taxa that dominated the microbiota of milk from silos at processing facilities, among which were *Pseudomonas*, *Lactococcus*, and *Acinetobacter*, had been found predominant in bulk or tanker milk as well. Moreover, McHugh et al. [18] reported that none of the genera uniquely occurring in milk from silos at more than 5 % relative abundance had been previously identified in the supplying tanker or bulk tank. These findings identify the on-farm environment as a key place to carry out effective measures to ensure raw milk quality.

Management practices implemented to maintain dairy farm profitability and improve the quality of products encompass a variety of measures among which housing, use of antibiotics, milking routine, bedding, cow hygiene, and herd nutrition are the most important of those associated with changes in the raw milk microbiota [17,38,52]. Dry and ensiled grass or legume, which can be supplemented with corn silage, constitute the most common feed components for dairy cows [311]. In a previous study (Chapter 2), we used high-throughput sequencing of the 16S rRNA gene pool to compare the composition and structure of the viable bacterial communities populating farm-scale produced hay and grass/legume or corn silage, among which the last two were ensiled with or without microbial additives. We found, as have others [245,247,249] that these distinct forage types also harboured phylogenetically different community assemblies which included, besides technologically relevant bacteria [29], pathogenic and spoilage microorganisms that can contaminate raw milk and cause serious defects during milk processing [112,242]. Doyle et al. [12] previously investigated various sources of raw milk contamination at the farm, and found that grass silage was a minor contributor, after teat surface and faeces, to the microbiota of bulk tank milk produced by cows housed indoors. However, our knowledge on the prevalence and diversity of raw milk microbial species that originate from forage types including hay and grass/legume or corn silage is limited. Moreover, despite the increasing interest for silage inoculants with improved fermentative capabilities and high potential for silage aerobic stability and animal productivity [171,174,284,312], little is known about the impact of these commercial inoculants on the raw milk microbiota and processability. Recently, Gagnon [22] et al. applied a culture-dependent approach targeting lactic acid bacteria (LAB) to identify common phylotypes between the five forage types that were given to the cows and the milk they produced, respectively. Corroborating the results from Doyle et al. [12] specifically for the LAB fraction of the microbiota, these authors found that isolates probably originating from forage represented about 6 % of the observed LAB community in milk. Importantly, they did not find any significant impact of the increased proportion of lactobacilli in inoculated silage on the LAB community of corresponding milk samples. This emphasizes the lack of knowledge of the patterns of raw milk contamination on farm that are driven by silage management practices.

The main objective of this study was to investigate the impact of feeding dairy cows with dry or ensiled forage, whether inoculated or uninoculated, on raw milk microbiota. Using high-throughput sequencing and qPCR both based on viable cells [43,251], we examined the same samples as those used by Gagnon et al. [22], that were collected in fall 2015 and spring 2016. Forage samples and derived microbial sequences analysed in the current study were also identical to those of the previous chapter (Chapter 2). However, to expand on the previous chapter, forage samples are now grouped according to forage ration types (see materials and methods below) by merging sequence data where applicable to build a new dataset that matched milk samples at the farm level. This study therefore provides a comparative viewpoint of the living fraction of the microbial community populating five forage types fed to cows and associated raw milk from commercial dairy farms by: (1) analysing

the bacterial profiles and community assembly of forage types containing either hay or grass/legume silage as the sole forage types, or a mixture of grass/legume and corn silage each supplemented or not with inoculants when ensiling, (2) assessing the compositional and structural changes of the microbiota from raw milk associated with the five forage types, and (3) identifying the bacterial phylotypes shared among forage rations and corresponding milk samples, and finally determining their transfer rates for each forage combination.

4.4 Materials and methods

4.4.1 Forage and milk sampling

The forage types and raw milk samples analysed in this study are identical, as mentioned above, to those examined by Gagnon et al. [22] and which were collected from 24 dairy farms in the fall 2015 and the spring 2016. The same forage samples were also analysed in a previous study [Chapter 2]. Sampling and sample treatments were therefore processed in the same way as previously described for forage [Chapter 2], but differently for milk [22], since the distinct processing steps required for DNA extraction were performed as previously described [43]. Briefly, the 24 dairy farms implemented five cow feeding practices defined by using either hay (H) or grass/legume silage uninoculated (GL) as the sole forage types, or by adding to GL or grass/legume silage inoculated (GLI) corn silage uninoculated (C) or inoculated (CI) at proportions ranging from 38 to 74 % relative to the total mixture obtained. The forage rations therefore included H, GL, GLC, GLCI, and GLICI feeding combinations counting 5, 7, 4, 1, and 7 herds, respectively, for each of the two sampling periods. Farmers used eight forage storage forms, including concrete-stave silo, oxygen limiting silo, pressed silo, bunker silo, wrapped square/round bale, stack silo, loose, and small bale as previously described [Chapter 2]. Commercial inoculants used for ensiling included Biotal Buchneri 500 and Biotal Supersile from Lallemand Animal Nutrition (Milwaukee, WI), and 11C33, 11CFT, and 11G22 from Pioneer (Johnston, IA).

Milk samples collected from bulk tanks (100 mL) were conveyed refrigerated to the laboratory as described previously [22,43]. A subsample was sent to Lactanet laboratories (Sainte-Anne-de-Bellevue, QC, Canada) for the quantification of fat, protein, somatic cells (SCC), urea, and lactose. Moreover, for each milk sample, four cell pellets, individually prepared from 10 mL aliquots were treated with propidium monoazide (PMA) as previously described [43] to account for viable cells. Four other pellets were processed without PMA treatment. PMA-treated and PMA-free pellets were stored at -80°C until DNA extraction.

4.4.2 DNA extraction, sequencing, and PCR quantification

Genomic DNA extraction from forage and milk samples was performed using the DNeasy PowerFood Microbial Kit (Qiagen, Hilden, Germany) following enzymatic lysis with mutanolysin from *Streptomyces* (MilliporeSigma), lysozyme (MilliporeSigma), and proteinase K (MilliporeSigma) as previously described

[Chapter 2, 43]. Genomic DNA of PMA-free and PMA-treated samples for forage and milk were sent for sequencing at the *Plateforme d'Analyses Génomiques* of Laval University (Quebec, Canada). The V3-V4 region of the 16S rRNA gene was amplified using the 347F (5'-GGAGGCAGCAGTRRGGAAT) and 803R (5'-CTACCRGGGTATCTAATCC) primers.

Using specific primer sets as described previously [Chapter 2], milk loads in *Lactiplantibacillus plantarum*, *Lentilactobacillus buchneri*, lactic acid bacteria (LAB), acetic acid bacteria (AAB), *Pseudomonas*, *Enterobacteriaceae*, total bacteria, and total fungi in the PMA-free and PMA-treated samples were determined by quantitative PCR (qPCR). qPCR results were expressed as log copy numbers per millilitre of milk as previously described [43].

4.4.3 Bioinformatic and data analyses

Raw sequences were quality checked and processed using the software FastQC (Version 0.11.9), Cutadapt (version 2.3), and the DADA2 pipeline [98,213] as previously described [43]. The same forage sequences analysed previously [Chapter 2] were used in this study. However, sequence data associated with forage were simply merged (1/1 ratio of read counts) according to the feeding combination described above to form a new dataset. Although forage sequences were examined separately from those of milk, their corresponding sequence tables were merged using the function *mergeSequenceTables()* before removing chimeras and assigning taxonomy as described in the DADA2 tutorial for big data. Forage and milk processed sequences were therefore assigned ASV (amplicon sequence variant) names at once, and downstream analyses regarding alpha and beta diversity were performed using the phyloseq (version 1.30.0) package [217]. Based on a preprocessing analysis establishing the significant effect of sampling period on the microbial structure, the alpha- and beta-diversity metrics of milk microbial communities were analysed separately for each sampling periods on the PMA-treated dataset. For the alpha-diversity analysis, Chao1, Shannon, and Inverse Simpson indices were computed. Local contribution to beta-diversity (LCBD) and principal coordinate or component analyses (PCoA, PCA) were used to capture the beta-diversity as previously described [Chapter 2, 255]. Sequence data were transformed by Centered Log-Ratio (CLR) or by Phylogenetic Isomeric Log-Ratio (PhILR) to assess the compositional and the phylogenetic structures of the milk microbiota [252,253]. Statistical analysis of group comparisons was performed using the Kruskal-Wallis and post-hoc tests with false discovery rate correction of the p-values. The R package ALDEx2 4.0 was used to compute taxa differential abundance [256] at the ASV level. Function pathway prediction was performed on PMA-free data using the package Piphillin [121], to account for the whole functional potential of the microbiota, irrespective of viability. The R package ComplexHeatmap 2.4.2 [225] was used for data visualization.

In addition to the differential abundance testing described above, the PhILR transformed milk dataset was analysed using a sparse logistic regression model implemented in the glmnet (version 3.0.2) R package [313] to identify "balances" that discriminate milk samples between forage combinations [253]. In the PhILR transform approach, the phylogenetic information is used to transform the microbiome data into an unconstrained space with an orthogonal basis [253]. The resulting log-ratios of the geometric mean relative abundances of adjacent clades are called "balances" [253].

Sparse partial least squares regression (sPLS) coupled with general linear models, and canonical correspondence analysis (CCpnA) were computed to determine associations between milk microbial communities, milk parameters, and farm characteristics as previously described [Chapter 2], using MixOmics (version 6.11) (37), MaAsLin2 (https://github.com/biobakery/MaAsLin2), and phyloseq. A heatmap was used to visualise the abundance level and distribution of the selected ASVs among milk samples and feeding typologies.

The online molecular ecological network analysis pipeline (MENAP, http://ieg4.rccc.ou.edu/mena) was used to construct ASV co-occurrence networks for each typology on PMA-treated data as described in Chapter 2. For each network, the node (representing an ASV) parameters including the within-module connectivity (Z_i) and among-module connectivity (P_i) indices were used to determine topological roles, that included module hubs, network hubs, peripherals, and connectors [265].

To gain deeper insight into the similarities of microbial community composition among forage types regardless of the predefined feeding combinations, we applied a partitioning clustering approach using the partitioning around medoids (PAM) algorithm on Euclidean distances calculated from PhILR transformed data. We used the PAM algorithm implemented in the factoextra 1.0.7 [314] R package by applying the *eclust()* function that provides at once the information on gap statistics to get the optimum number of clusters, silhouette analysis, and cluster graph, which are accessed through the functions *fviz_gap_stat(), fviz_silhouette()*, and *fviz_cluster()* respectively. In addition to silhouette analysis, ordination in the PhILR space was performed as previously described to validate the cluster analysis results. Forage rations (whether from composite or single forage type) were then classified into different forage ration community types defined by the number of clusters obtained.

Co-occurring ASVs among forage types composing each forage combination and the associated milk samples were investigated by calculating intersects with the software VENN DIAGRAMS available online at http://bioinformatics.psb.ugent.be/webtools/Venn/. The abundance and distribution of shared ASVs among forage and corresponding milk in each combination were visualized by constructing a heatmap and a chord diagram using Complexheatmap [225] and Circos v0.63-9 [315], respectively.

4.5 Results

4.5.1 Microbial community types of forage combinations

Significant differences in the phylogenetic structure of microbial communities were found between H and GL, GLC, or GLICI forage combinations, as well as between those of GL and GLC or GLICI (Fig. 4-1A, supplemental Fig. C1 A). Bacterial pools provided by GLC and GLICI forage rations were phylogenetically similar. Cluster analysis of forage rations at the farm level showed three community types. This result was validated by gap statistics, silhouette analysis (supplemental Fig. C1 B-C), and a PCoA evaluating cluster separation (Fig. 4-1B). Cluster 1 gathered all samples of the H type, broadly exhibiting co-dominance of *Enterobacteriaceae*, *Microbacteriaceae*, *Sphingomonadaceae*, *Beijerinckiaceae*, and *Pseudomonadaceae* (Fig. 4-1C). Most forage rations containing a mixture of grass/legume and corn silage when both were inoculated composed cluster 2. Samples in this cluster were either largely dominated by *Lactobacillaceae* or co-dominated by *Lactobacillaceae* and *Acetobacteraceae* or *Bacillaceae*. Cluster 3 gathered almost all GL samples, characterized by relatively high proportions of *Beijerinckiaceae*, *Rhizobiaceae*, and *Enterococcaceae*. Additionally, forage rations in this cluster were either largely dominated by *Lactobacillaceae* or co-dominated by *Lactobacillaceae* and *Leuconostocaceae*. Forage rations co-dominated by *Acetobacteraceae* and *Enterobacteriaceae* (sample 3HMNIP3) or by *Bacillaceae* and *Pseudonocardiaceae* (sample 5HNIP2) exhibited the highest LCBD indices compared with others within clusters 2 and 3, respectively.

4.5.2 Diversity of raw milk microbial communities

Similar alpha-diversity metrics (Chao1, Shannon, and inverse Simpson indices) were observed between raw milk samples across forage ration combinations (supplemental Fig. C2). Beta diversity analysis performed at the genus level showed that in the fall, GL, GLC, and GLICI milk samples harboured similar microbial community structures, each significantly different from that of H milk (Fig. 4-2A). In the spring, while H milk samples showed similar community structures with GL and GLC, a significant difference was observed between H and GLICI (Fig. 4-2B). Interestingly, GL, GLC and GLICI milk samples exhibited significantly different community structures ($p<0.05$) from each other. However, in the fall, milk microbial communities of H compared with GL, or GLC versus GLICI were phylogenetically similar, while those of H and GL were significantly different from GLC and GLICI (Fig. 4-2C). In the spring, H, GL, and GLC milk samples were phylogenetically similar, but each was significantly different from GLICI (Fig. 4-2D). Regardless of the sampling periods, *Serratia*, *Pseudomonas*, unclassified *Enterobacteriaceae*, and *Cellulosimicrobium* generally dominated milk samples. Samples exhibiting higher LCBD indices were enriched in either one or combinations of the genera *Lactobacillus*, *Lactococcus*, *Escherichia*/*Shigella* group, *Kocuria*, *Novosphingobium*, *Vulcaniibacterium*, *Schlegelella*, *Rothia* or *Psychrobacter* (Fig. 4-2E-F).

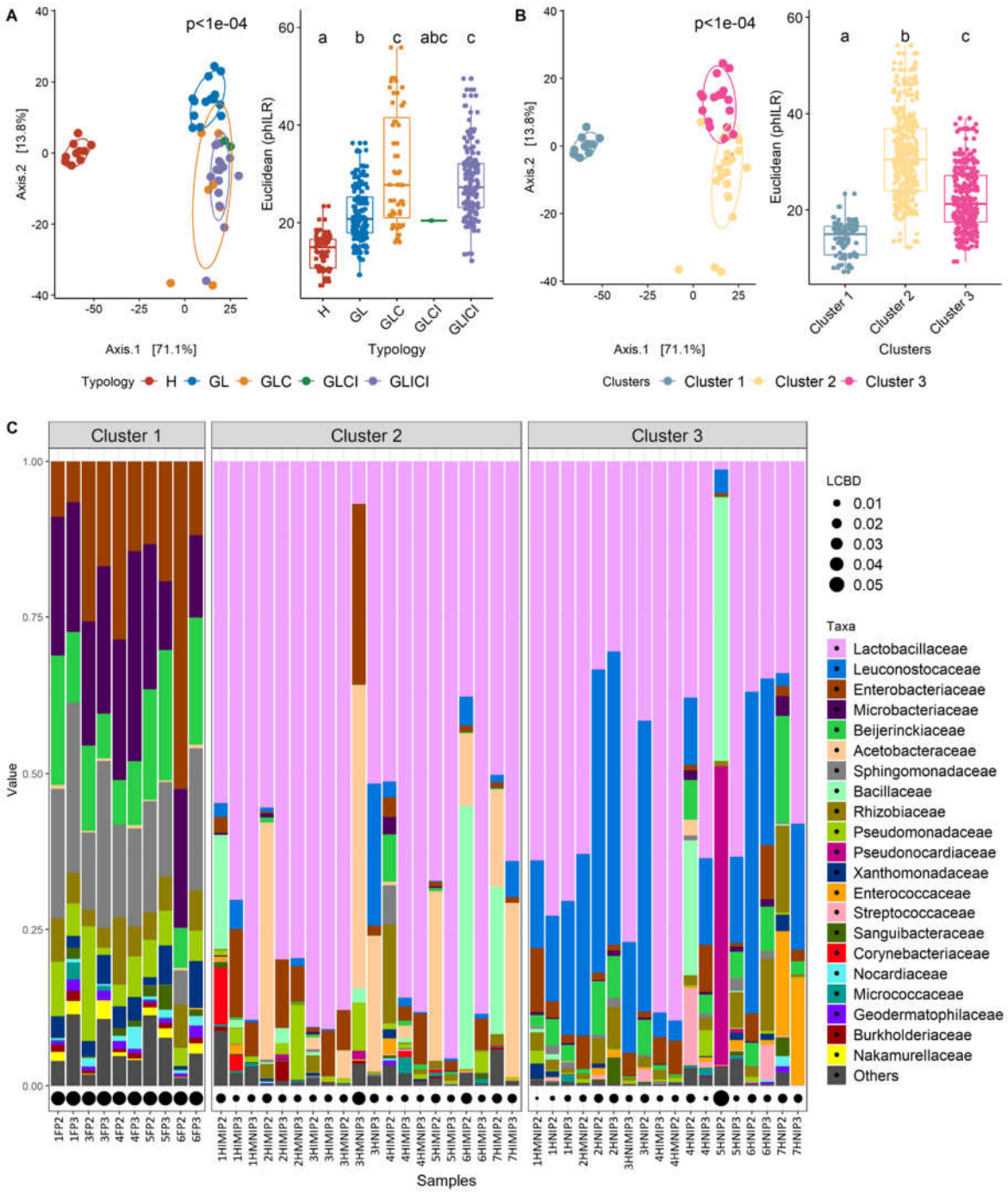


Figure 4-1: Forage ration community types. Principal component analysis based on PhILR transformed data (left) with corresponding post-hoc tests (right) for forage ration types (A) and derived clusters (B). Forage feeding typologies or clusters associated with different letters are significantly different according to a multiple comparison analysis using the Wilcoxon rank sum test with FDR correction. (C) Relative abundance of the 21 more abundant families occurring in the defined clusters. Local contribution to beta diversity values denote the indices of sample local contribution to the observed beta-diversity between groups. Values are proportional to sample contribution to beta-diversity. Cluster analysis results are provided by supplemental Fig. C1 in Annex C.

Almost all the ASVs found to be differentially abundant between milk samples in the fall and in the spring (96 % of 47 ASVs) were *Proteobacteria*, and the remaining *Actinobacteria* (supplemental Fig. C3 A-B). Among these taxa, *Pseudomonas*, unclassified *Enterobacteriaceae, Serratia*, and *Cellulosimicrobium* were the most abundant. Compared with GL or GLC, most of the differentially abundant taxa exhibited higher relative abundance in GLICI milk samples across both sampling periods.

We identified 19 balances that discriminated between milk samples associated with the forage types (supplemental Fig. C4). Among balances, mostly *Proteobacteria* were involved at all taxonomic levels. However, concerning *Firmicutes*, we found that the abundance of *Clostridium disporicum*, *Clostridium*, *Paeniclostridium*, *Coprococcus*, *Romboutsia sedimentorum*, *Romboutsia*, *Veillonella dispar*, and an unclassified *Peptostreptococcaceae* relative to the *Actinobacteria* including *Bifidobacterium*, *Cellulosimicrobium*, *Kocuria*, and other members of the *Actinobacteria* class distinguished GLICI or GLC from GL or H milk samples. Moreover, the level of *Leuconostoc* relative to that of *Weissella* also separated GLICI from GLC milk samples.

PMA-qPCR analyses performed to estimate live microbial loads in milk revealed that neither *L. buchneri*, *L. plantarum*, LAB, AAB, *Pseudomonas* spp., *Enterobacteriaceae*, nor total bacteria varied significantly between milk samples in the fall (Fig. 4-3A). However, a significantly lower abundance ($p<0.05$) of total fungi was observed in GLICI compared with H milk samples (Fig. 4-3A). In the spring, a significant enrichment of LAB was observed in GLICI compared with GL milk samples ($p<0.05$), as were *Pseudomonas* in GLICI compared with GLC and GL milk samples, respectively ($p<0.05$), and *Enterobacteriaceae* in GLICI compared with GL milk samples ($p<0.05$). Total bacterial loads were significantly higher in GLICI or GLC compared with GL ($p<0.05$) and H ($p<0.0001$), respectively (Fig. 4-3B). Although not significant, *L. buchneri* and *L. plantarum* levels were consistently higher in GLC and GLICI milk samples, respectively, across both sampling periods.

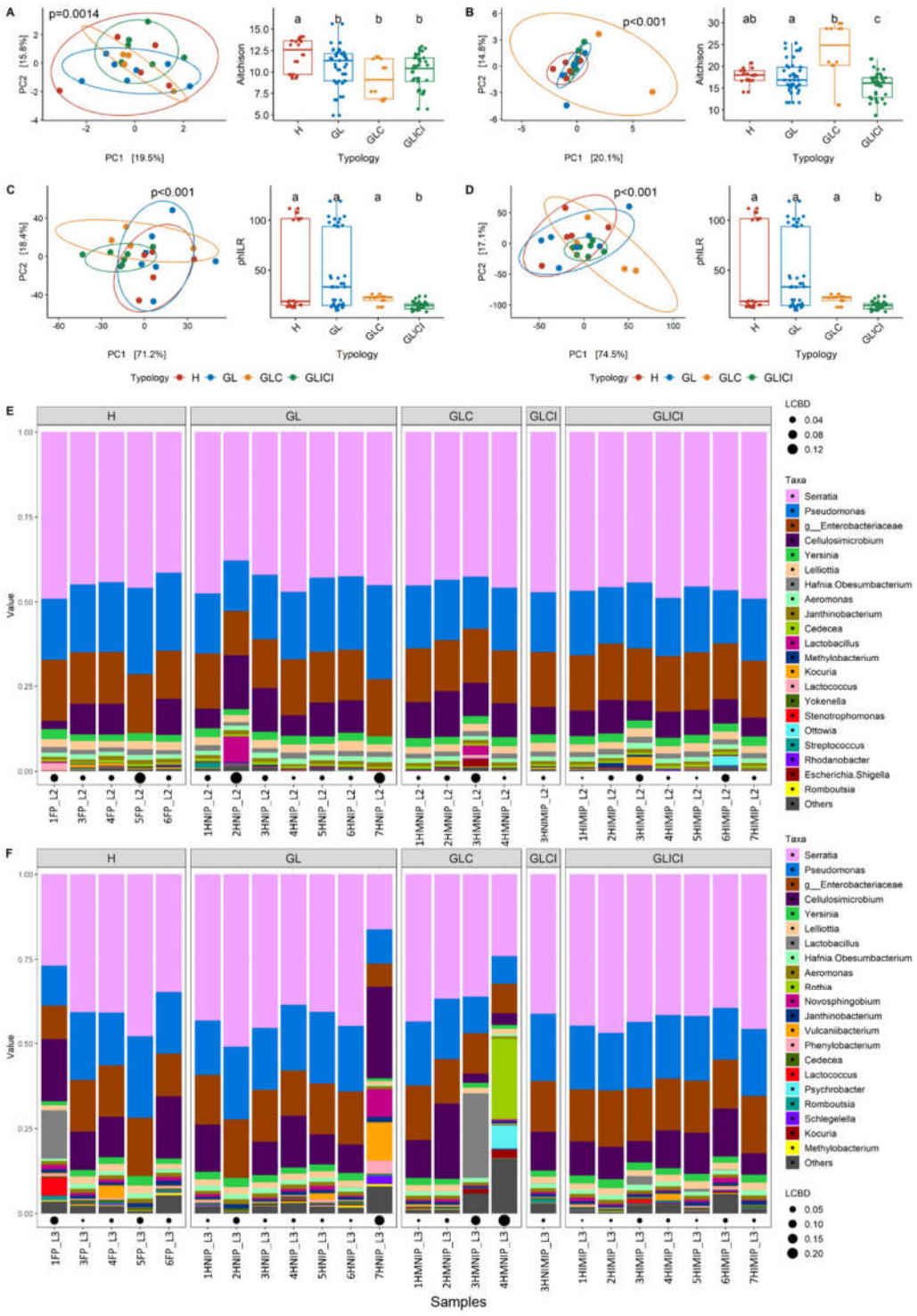


Figure 4-2: Milk community diversity and composition across forage types. Principal component analysis (left) with corresponding post-hoc test (right) based on CLR transformed data of milk samples in the fall (A) and the spring (B) or based on PhILR transformed data in the fall (C) and the spring (D). Milk groups associated with different letters are significantly different according to a multiple comparison analysis using the Wilcoxon rank sum test with FDR correction. Relative abundance of the 21 more abundant genera in the microbiota of milk samples collected in the fall (E) and the spring (F). Local contribution to beta diversity values denote the indices

of sample local contribution to the observed beta-diversity between groups. Values are proportional to sample contribution to the beta diversity.

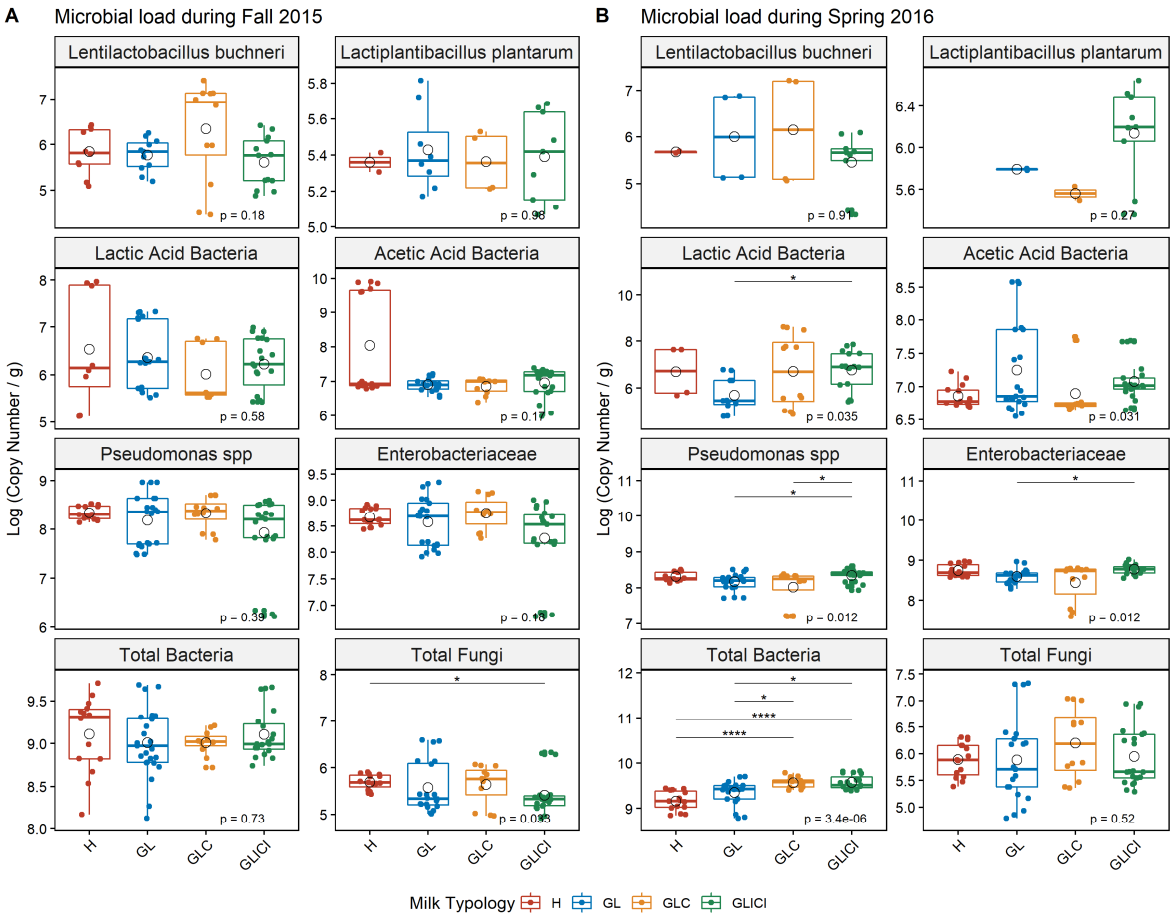


Figure 4-3: Comparative analysis of milk viable microbial loads across forage types. Microbial loads expressed in log copy numbers are compared between H milk sample and those from cows fed with the three other forage types, or between GLICI milk samples and those from GLC or GL in the fall (A) and the spring (B). p values are obtained following a Kruskal Wallis test. Asterisks above boxes indicate significant differences and flag p-values from a Wilcoxon rank tests as follows: *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ****, $p < 0.0001$.

4.5.3 Distribution of *Lactobacillales* in raw milk and prediction of function pathways

The milk dataset was comprised of 569 ASVs of which those assigned to the order *Lactobacillales* (~11 %) accounted for only 1 % of total relative abundance. Broadly, 28 ASVs occurred in H milk samples, 25 in GL, 20 in GLC, and 18 in GLICI milk samples (Fig. 4-4A). The core LAB among milk samples included *Lactobacillus* represented by two phylotypes, and *Lactococcus* represented by a single phylotype. Within each feeding combination, none of the core LAB phylotypes, those shared between group pairs, nor those found specific to a group (Fig. 4-4B-C, supplemental Fig. C5A-B) were consistently detected among milk samples. Moreover,

several samples harboured one or two phylotypes of either *Lactococcus*, *Lactobacillus*, *Leuconostoc*, *Streptococcus*, *Weissella*, or *Enterococcus* as the sole representatives of the LAB community.

Regarding predicted metabolic pathways, we found that compared with GL and GLC, GLICI milk samples were significantly enriched in specific features associated with the biosynthesis of secondary metabolites, metabolism of cofactors and vitamins, carbohydrate metabolism, energy metabolism, biofilm formation, amino acid metabolism, thermogenesis, and nucleotide metabolism across sampling periods (supplemental Fig. C6-7).

4.5.4 Milk characteristics and association with microbial communities

Milk content in fat, protein, urea, and lactose, as well as SCC did not vary significantly between forage types, nor across sampling periods (supplemental Table C1). Regardless of the feeding type, the multivariate multi-table analyses revealed three clusters of milk microbial communities (supplemental Fig. C8 A). SCC and fat content were the main variables contributing to the formation of cluster 1, while protein and lactose content highly contributed to the formation of cluster 3. Among the 78 taxa selected by the sPLS approach, i) *Firmicutes* including *Staphylococcus*, *Enterococcus*, and *Streptococcus dysgalactiae*, ii) *Actinobacteria* represented by *Corynebacterium bovis*, *Corynebacterium glutamicum*, and *Frigoribacterium*, and iii) *Proteobacteria* comprising *Pantoea*, *Acinetobacter*, *Acinetobacter bouvetii*, and *Acinetobacter guillouiae* mostly contributed to the separation of clusters on component 1 (supplemental Fig. C8 B). These taxa that were all associated with cluster 1 positively correlated with SCC and fat content. On component 2, i) *Actinobacteria* including *Bifidobacterium pseudolongum*, *Rhodococcus*, *Corynebacterium*, and *Bifidobacterium mongoliense*, ii) *Proteobacteria* comprising *Ottowia*, *Pseudomonas*, and *Comamonas koreensis*, and iii) *Bacteroidetes* including *Chryseobacterium* were the main contributors to the cluster separation (supplemental Fig. C8 B). Cluster 2 was mainly composed of taxa that showed negative correlations with the milk parameters mentioned above (supplemental Fig. C8 C), while taxa composing cluster 3 were roughly positively associated with protein, lactose, and longitude. The MaAsLin2 approach revealed 38 microbial features significantly associated with the use of inoculants, milk fat and protein content, as well as the SCC. Mostly i) *Firmicutes* including phylotypes of unclassified *Peptostreptococcaceae*, *Lactobacillus*, *Staphylococcus*, *Enterococcus*, *W5053*, unclassified *Carnobacteriaceae*, *Lactococcus*, and *S. dysgalactiae*, and ii) *Actinobacteria* comprising *Brachybacterium*, *B. mongoliense*, *Micrococcus*, and *Corynebacterium* positively correlated with inoculants. In addition to phylotypes of cluster 1, those of *Corynebacterium humireducens* and *B. pseudolongum*, found in cluster 3, positively correlated with SCC.

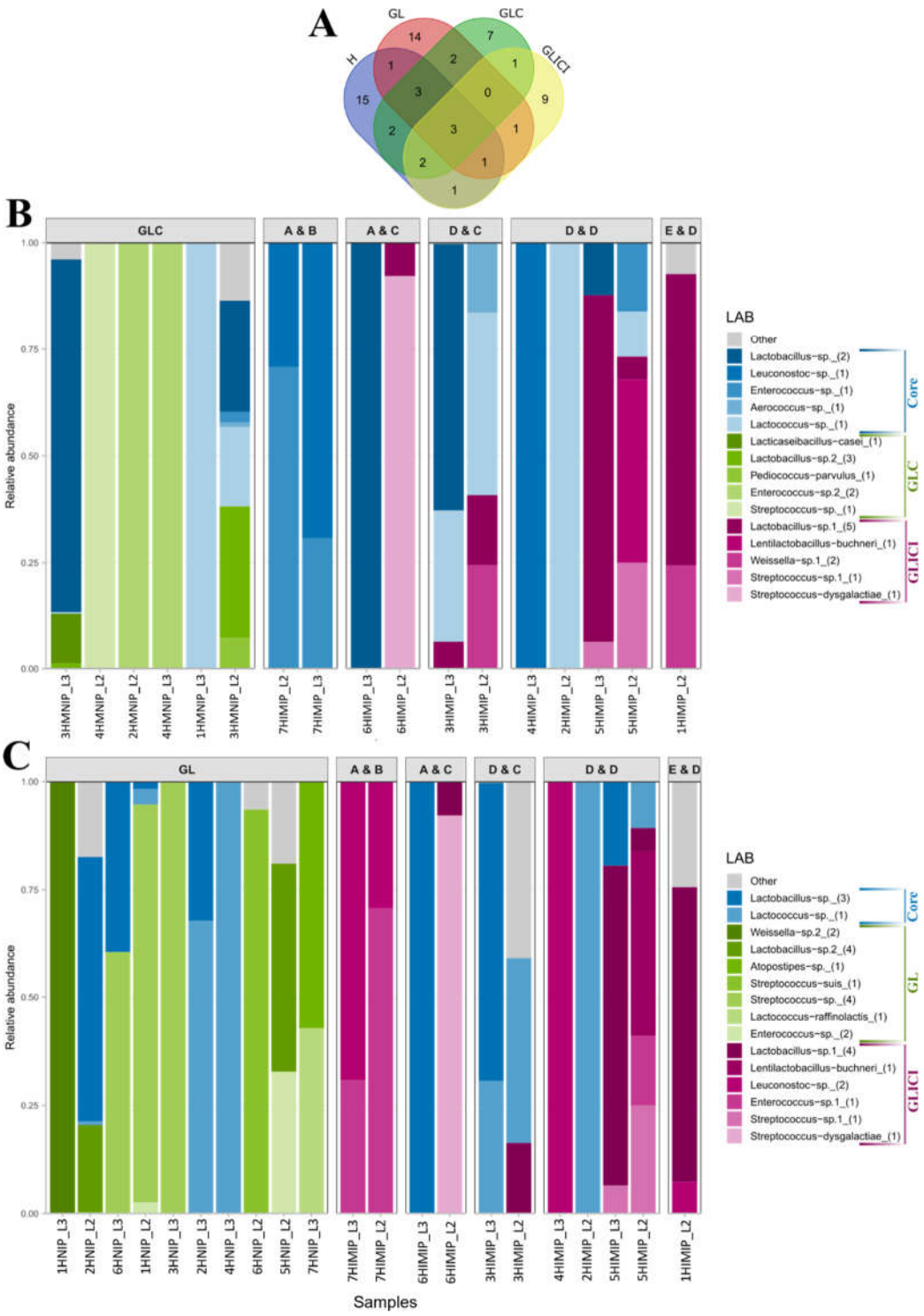


Figure 4-4: Distribution and proportion of LAB communities among milk groups. (A) Venn diagram of LAB phylotypes among milk groups. Proportion of the core and unique LAB phylotypes among GLICI and GLC (B) or GL (C) milk samples. Milk samples originating from inoculated forage types are separated according to the inoculants used for the grass or legume (first letter) and corn (second letter) ensilage. Accordingly, “A” = 11G22, “B” = 11C33, “C” = 11CFT, “D” = Biotal Buchneri 500, and “E” = Biotal Supersile. Taxa are coloured according

to their unicity to a group (GLC versus GLICI, GLICI versus GL) or whether they were shared (core) by group pairs.

The CCpnA result (Fig. 4-5) also showed three clusters and revealed that in addition to milk SCC and fat content, ASVs composing cluster 1 were correlated with forage storage forms. Cluster 3 that mainly correlated with protein and lactose contents were also found to be more associated with the spring season than with the fall. Interestingly, ASVs composing this cluster were also found to correlate more with the inoculant Biotal Buchneri 500 and to a lesser extent with inoculants 11G22 and 11C33 used concomitantly in grass/legume and corn silage. Taxa gathered into cluster 2 correlated with inoculants 11G22 and 11CFT, as well as Biotal Buchneri 500 and 11CFT, both inoculant pairs used concurrently on two GLICI farms. ASVs significantly correlated with milk components and categorical variables were not the most abundant, although encompassing the major phyla detected (supplemental Fig. C9). However, the greatest prevalent ASVs were distributed in cluster 3, of which the most abundant included phylotypes of unclassified *Enterobacteriaceae*, *B. pseudolongum*, *Aerococcus*, and *C. humireducens*. On the other hand, ASVs composing cluster 1 were generally the less prevalent, but seemed to exhibit higher prevalence within the H forage. The most abundant ASVs of this cluster included *S. dysgalactiae*, *Staphylococcus*, and *C. glutamicum*.

4.5.5 Milk microbial network analysis and topological roles

The GL milk microbial community network exhibited a higher size and average geodesic path but contained a lower average degree than that of GLC, indicating a lower complexity level (supplemental Table C2). The 135 nodes composing the GL milk network formed four modules and mostly co-excluded rather than co-occurred, as 80 % of the 1089 links connecting them were negative (supplemental Fig. C10 A). The GL nodes belonged to *Proteobacteria*, *Actinobacteria*, *Firmicutes*, and *Fusobacteria*, each totalling 61, 37, 36 and 1 phylotypes, respectively. *Proteobacteria*, including phylotypes of *Aeromonas*, *Cedecea*, *Hafnia*, *Janthinobacterium*, *Lelliothia*, *Methylobacterium*, *Pseudomonas*, *Serratia*, *Yersinia*, and other unclassified *Enterobacteriaceae* essentially populated modules 2 and 4 within the network, sharing only positive relationships among themselves. The *Lactobacillus* phylotype exhibited essentially negative relationships with other taxa including *Proteobacteria*, *Actinobacteria*, and *Firmicutes*, the latter including *Aerococcus*, *Lachnospiraceae*_NK3A20 group, *Pediococcus*, *Ruminococcaceae*_NK4A214 group, *Ruminococcaceae*_UCG-005, *Staphylococcus*, *Weissella,* and an unclassified *Carnobacteriaceae*. The *P. pentosaceus* phylotype shared positive relationships with *Firmicutes* identified as *Staphylococcus* and *Carnobacteriaceae*, and negative relationships with other *Firmicutes* including *Staphylococcus*, *Jeotgalicoccus*, and *Aerosphaera*. Besides phylotypes assigned to the genera *Lactobacillus* and *Pediococcus*, no other phylotypes related to inoculants used by these farmers were involved in this network. The analysis of topological roles identified two phylotypes of *Pseudomonas* as keystone species of the module hubs category (Fig. 4-6).

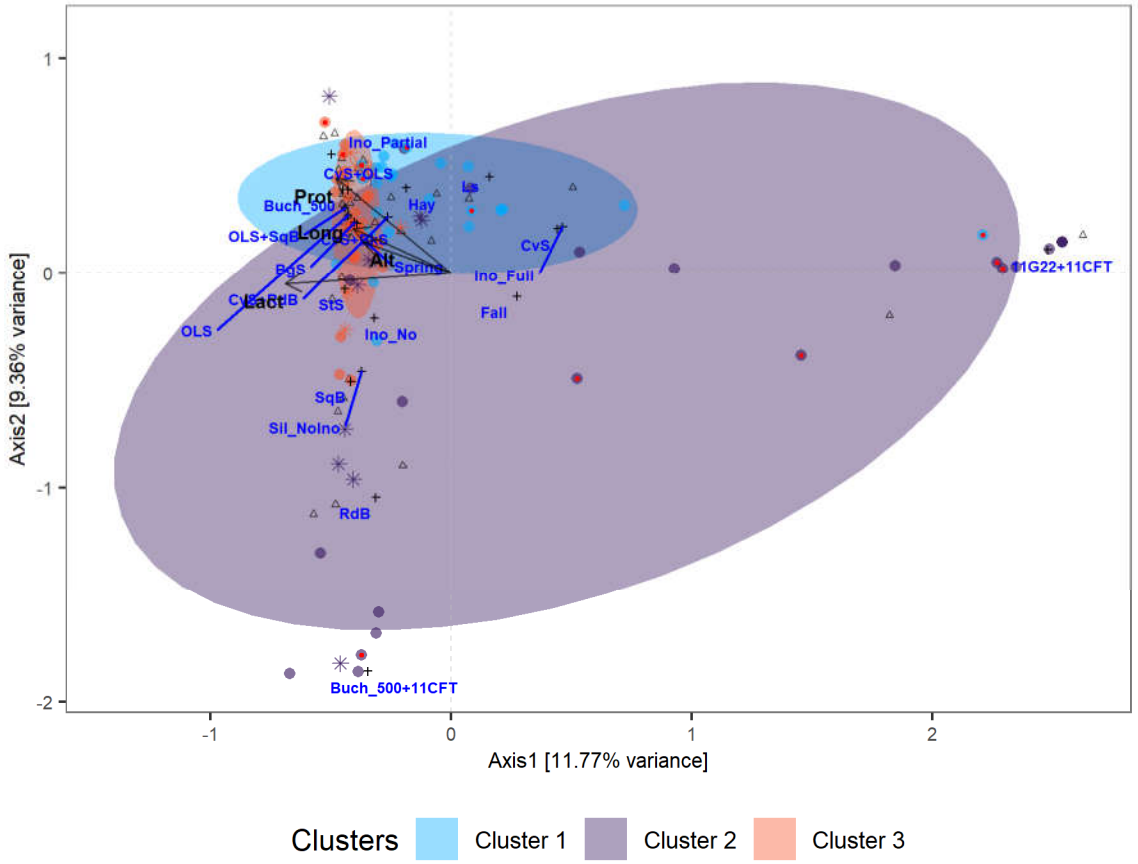


Figure 4-5: Canonical correspondence analysis based on the selected taxa. Triplot of canonical relationships between phylotypes (round and star shaped points) and milk parameters (arrows), inoculants (+), forage storage form (+), or sampling periods (+) for milk samples associated with all the feeding combinations. Fall and Spring denote the fall 2015 and the spring 2016 seasons, respectively. Storage forms include loose (Ls), square bales (SqB), conventional silo (CvS), bag silo (BgS), stack silo (StS), round bale (RdB), oxygen limiting silo (OLS), and bunker silo (BkS). Buch_500, 11CFT, and 11G22 denote the inoculants Biotal Buchneri 500, 11CFT, and 11G22. Ino_Partial (partially inoculated) and Ino_Full (fully inoculated) correspond to mixed forage rations for which one (corn silage) or both forage types (with grass or legume silage) were inoculated. Milk parameters included proteins (Prot), lactose (Lact), sample altitude (Alt), and sample longitude (Long). Round points depict taxa selected by the sPLS and star shaped points those selected by the MaAsLin2 approaches. Round points with red centre are taxa selected by both methods. Triangles correspond to samples.

The GLC milk network formed a single module (supplemental Fig. C10 B) composed of 57 nodes (supplemental Table C2) and 1067 links, the latter of which 74 % were positive, indicating more co-occurrence than co-exclusion between phylotypes. Nodes were assigned to the phyla *Proteobacteria*, *Firmicutes*, and *Actinobacteria*, which were represented by 40, 10 and 7 phylotypes, respectively. *Firmicutes* included phylotypes of *Lactococcus*, *Staphylococcus*, *Streptococcus*, *Turicibacter*, *Romboutsia*, *Ruminococcaceae*_UCG-005 group, *Facklamia tabacinasalis* and *Aerococcus*. Inoculant related phylotypes were not involved in the GLC network.

The analysis of topological roles did not reveal keystone species, as all nodes were classified as peripherals (Fig. 4-6). However, there was a considerable shift in the extent of topological roles between GLC nodes and those of the GL network.

For the GLICI network, 135 nodes (supplemental Table C2) were interconnected with 1041 links illustrating more co-exclusion (85 %) than co-occurrence. This network was co-dominated by the phyla *Proteobacteria*, *Firmicutes*, and *Actinobacteria*, which were represented by 51, 42, and 40 phylotypes, respectively. The phyla *Bacteroidetes* and *Patescibacteria*, each represented by a single phylotype, were the less prevalent. This network thus exhibited a higher diversity level compared with those of GL and GLC. The GLICI network was composed of two disconnected modules, of which the smallest composed of 26 essentially co-occurrent nodes was dominated by *Proteobacteria* (supplemental Fig. C10 C).

In the larger module, two unconnected phylotypes of *Lactobacillus* shared principally negative relationships with *Firmicutes* including phylotypes classified as *Acidaminococcaceae*, *Carnobacteriaceae*, *Clostridiaceae*, *Enterococcaceae*, *Erysipelotrichaceae*, *Lachnospiraceae*, *Leuconostaceae*, *Peptostreptococcaceae*, *Ruminococcaceae*, and *Staphylococcaceae*, as well as phylotypes of the phyla *Actinobacteria*, *Proteobacteria* and *Patescibacteria*.

P. pentosaceus represented by a single phylotype shared negative relationships with other *Firmicutes* including phylotypes of *Staphylococcus*, *Streptococcus*, *Enterococcus*, *Jeotgalicoccus*, *Carnobacteriaceae*, and *Lactobacillus*, as well as other phylotypes classified as *Actinobacteria* and *Proteobacteria*. The genus *Enterococcus*, represented by two phylotypes, also shared essentially negative relationships with other *Firmicutes* including phylotypes of *Staphylococcus*, *Streptococcus*, *Weissella*, *Jeotgalicoccus*, *Carnobacteriaceae*, *Enterococcus*, *Ruminococcaceae*_UCG-005 group, *Leuconostoc*, *P. pentosaceus*, *Lactobacillus*, *Atopostipes*, *Clostridium_sensus_stricto*, *Paeniclostridium*, and *Cellulosillyticum*, as well as phylotypes of the phyla *Actinobacteria*, *Proteobacteria*, and *Patescibacteria*. Co-occurrence relationships involving *Firmicutes* were observed between phylotypes of *Staphylococcus* and *Actinobacteria* (including *Corynebacterium casei*, *Corynebacterium*, *Micrococcus*, *Brachybacterium*, and *Brevibacterium senegalense*) or *Proteobacteria* (including *Budvicia*, between *Romboutsia* and *Clostridium_sensu_stricto*), and between *Cellulosilyticum* and *Candidatus Saccharimonas*. The network topological analysis identified three module hubs (keystones species) including the phylotypes of *Pseudomonas*, *Lactobacillus*, and *Rhodococcus*. All the other nodes in the network were peripherals.

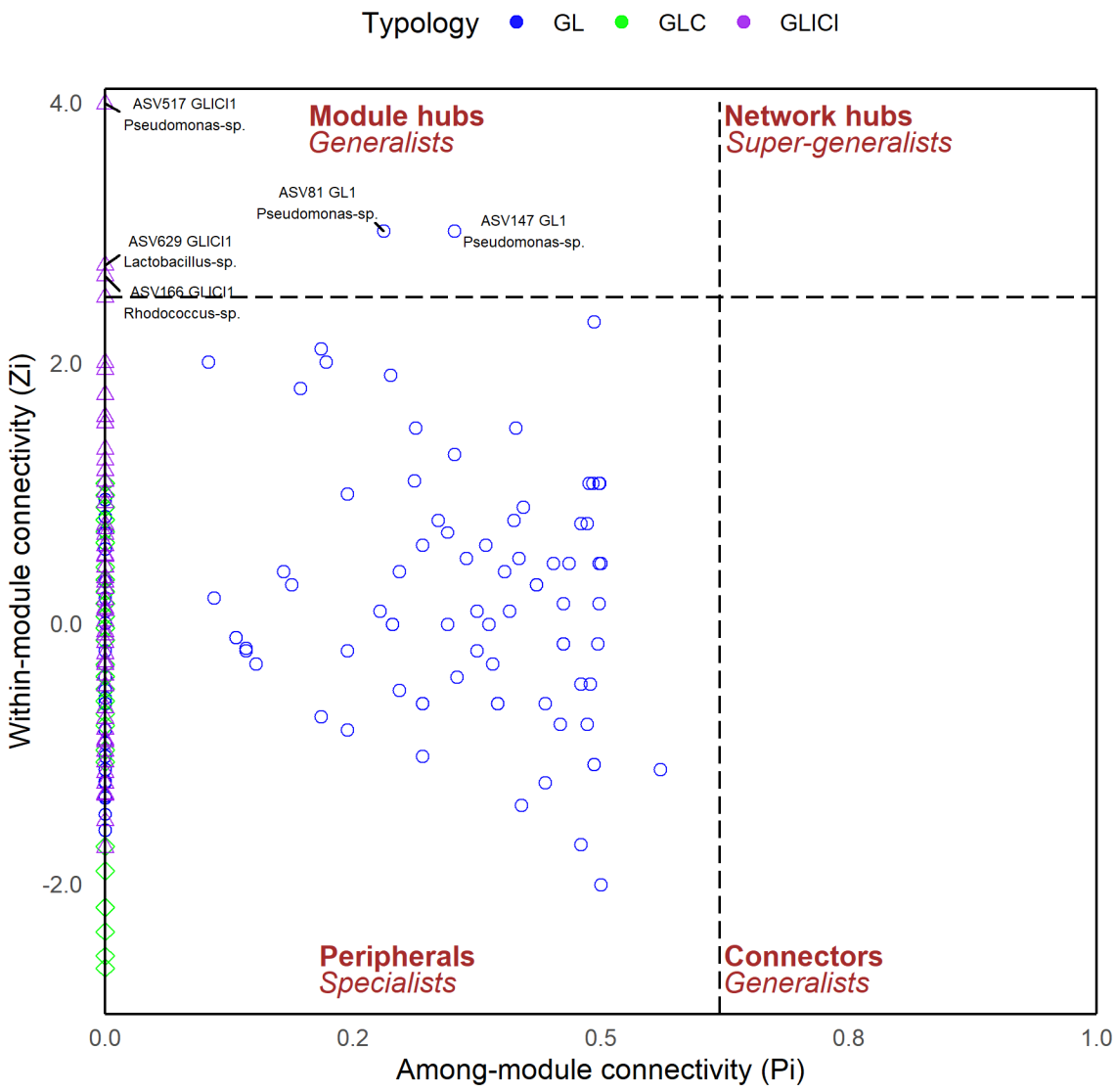


Figure 4-6: Network topological roles of the milk microbiota. Labelled ASVs represent module hubs (keystone species).

4.5.6 Shared bacteria between forage types and the corresponding raw milk

Comparing the bacterial communities between preserved forage and associated raw milk within the same feeding combination, we identified common phylotypes between both ecosystems. The proportions of shared and unique phylotypes between forage and milk varied across feeding typologies (Fig. 4-7A). Since a phylotype represents a unique sequence variant in the whole dataset and given that the raw milk microbiota originates from the dairy farm and its vicinity, we assumed that the concurrent occurrence of a phylotype in both forage and milk was plausibly the consequence of a transfer from forage to milk, with no assumption on the mode of transfer. Lower bacterial transfer rates from forage to milk were observed in H, GLCI, and GL feeding combinations, where milk samples shared 18, 19, and 21 % of their microbial community with corresponding

forage rations, respectively. Higher bacterial transfer rates from forage rations to milk were observed for GLC and GLICI forage combinations, as milk samples shared 31 and 30 % of their bacterial phylotypes with the respective associated forage.

A total of 113 phylotypes shared between forage and milk were identified, encompassing *Proteobacteria* (74), *Firmicutes* (26), and *Actinobacteria* (13) (Fig. 4-7B, Fig. C11). *Proteobacteria*, including phylotypes of *Serratia* and an unclassified *Enterobacteriaceae*, and *Actinobacteria* including *Cellulosimicrobium* largely dominated the milk microbial community considered to originate from forage. However, *Firmicutes* included, but not limited to, phylotypes of *Lactobacillus*, *P. pentosaceus*, *L. buchneri*, *Weissella*, *Lactococcus*, *Enterococcus*, *Leuconostoc*, *Clostridium*, and *Staphylococcus*. Interestingly, although these phylotypes were found among the core microbiota between forage and milk samples, they were not necessarily the most abundant in forage. Likewise, *P. pentosaceus* found at high relative abundance in the GLICI forage ration was not detected in the associated milk, whereas it was detected in milk samples from three other forage combinations showing lower relative abundance in forage (e.g., GL, GLC, and GLCI rations). Additionally, phylotypes of the phylum *Firmicutes* were not identified in all forage rations and milk samples across the feeding combinations (supplemental Fig. C11). Surprisingly, 96 % of the phylotypes significantly enriched in milk samples from the H, GL, GLC, and GLICI feeding combinations when compared to each other (supplemental Fig. C3 A-B) were members of those presumably transferred from forage (supplemental Fig. C11). Moreover, of the 39 phylotypes significantly enriched in the GLICI milk compared with H, GL, and GLC counterparts regardless of the sampling periods (supplemental Fig. C3 A-B), 92 % were part of the GLICI forage ration. Likewise, for the GLC type, 93 % of the 40 phylotypes enriched in milk were shared with silage. Lower proportions of 62 % out of 39 phylotypes and 63 % out of 41 phylotypes were observed in the GL and H typologies, respectively. Another noteworthy observation is the occurrence of phylotypes of *Rhodococcus*, *A. guillouiae*, *Staphylococcus*, *Acinetobacter, Pseudoclavibacter*, an unidentified *Enterobacteriaceae*, and *Enterococcus* that correlated with milk SCC among those presumably transferred from forage. Other phylotypes including *Brevundimonas*, *Frigoribacterium*, *Pantoea*, and *Jeotgalicoccus* that positively correlated with milk SCC were also identified in feeding types but were not found in corresponding milk samples (data not shown).

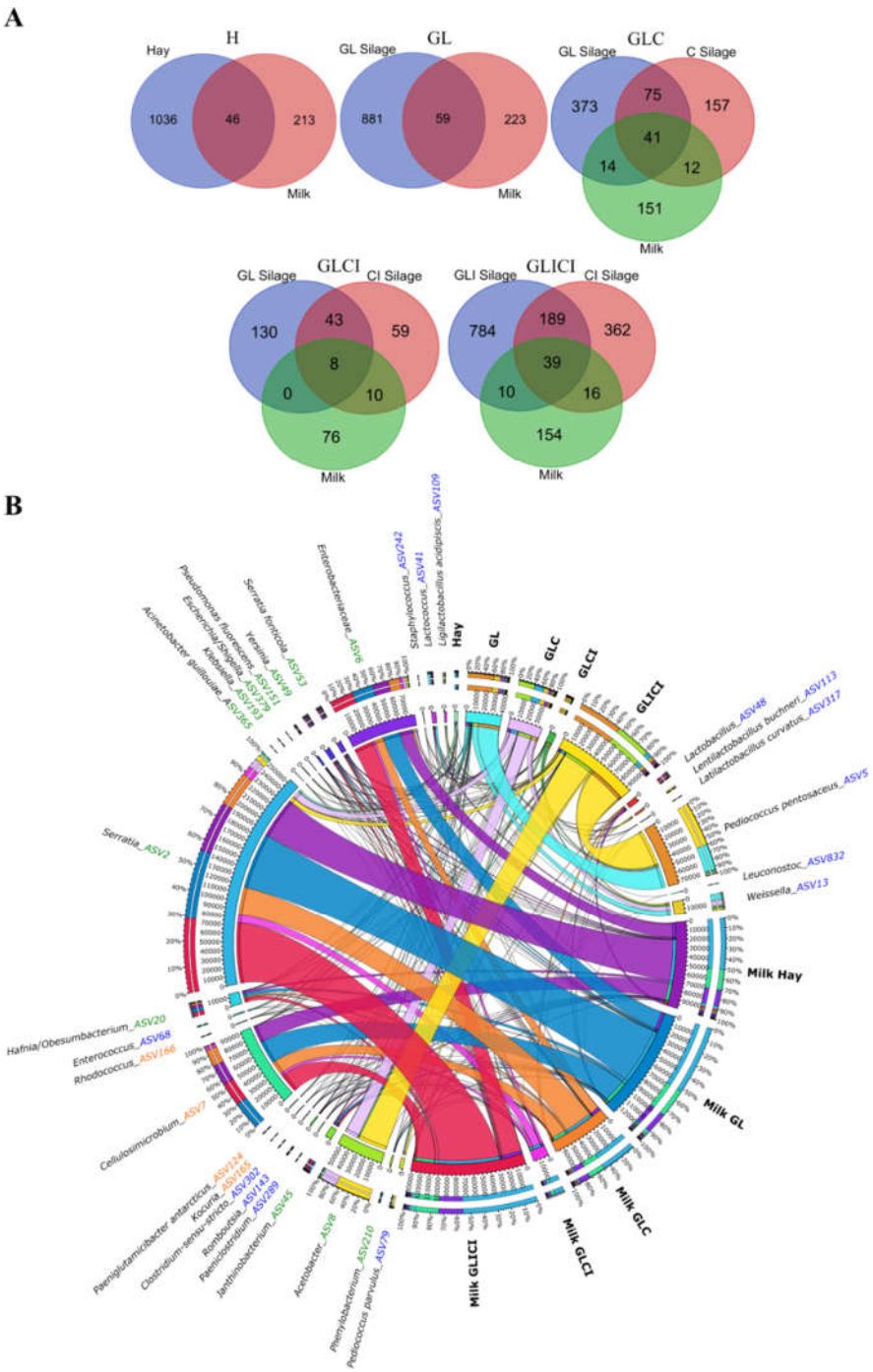


Figure 4-7: The core phylotypes between forage types and corresponding milk samples. (A) Venn diagrams showing the number of phylotypes shared between forage types and corresponding milk samples. (B) Chord diagram illustrating the distribution of the 30 more abundant phylotypes potentially transferred from forage types to milk.

4.6 Discussion

We recently analysed the microbiota of hay and ensiled forage types including grass/legume and corn silage both uninoculated or inoculated at harvest and found significant differences in the compositional and phylogenetic structures [Chapter 2]. In the dairy farms from which forage samples were collected, farmers implemented when applicable a mixed forage ration by adding corn silage to grass/legume silage, whether inoculated or not. In this study, we found that microbial communities composing these forage rations could be grouped into three community types, broadly distinguishing H, GL, and GLICI from each other. Most of the GLC rations exhibited high similarity to GLICI, while GLCI rations were evenly distributed between silage community types. This suggests that in mixed forage rations involving grass/legume and corn silage, inoculating one or both types of feed would not lead to much difference in the resulting bacterial community composition compared to that of GLC. The observed phenomenon could be explained by the high prevalence and high relative abundance of *Acetobacteraceae* in corn silage, or *Leuconostocaceae* in uninoculated grass/legume silage as previously reported [22,281,293]. As discussed earlier [Chapter 2], the genus *Weissella* was the main *Leuconostocaceae* in GL silage and was found to co-occur with undesirable taxa such as *Enterobacteria* in silage. This observation therefore emphasizes the advantages of microbial additives when ensiling grass/legume forage plants to obtain good-quality silage [243,288]. Further investigations involving more farm types will improve our understanding of the microbiological quality and aerobic stability of mixed forage rations and refine our findings on their community types.

Milk samples were not clustered otherwise than by the formation of two groups that perfectly matched the sampling periods (fall and spring), which are sub-categories of a predefined variable in our experimental design. Only the *a priori* defined feeding combinations were used for the statistical analysis of diversity measures for milk microbiota, regardless of the forage ration community types described above. The observed inconsistency in beta diversity patterns among feeding types associated milk samples collected in the fall and the spring corroborates previous studies in which variations of the milk microbiota across seasons was demonstrated [6,12,316]. Indeed, Li et al. [6] showed that during cooler seasons, the growth of psychrotrophic bacteria such as *Pseudomonas* was favoured, whereas *Firmicutes* and *Proteobacteria* correlated with higher temperatures. However, in this study, significantly higher loads of total bacteria, LAB, *Pseudomonas*, and *Enterobacteriaceae* were observed in GLICI milk samples compared with GL, as were *Pseudomonas* in GLICI compared with GLC, and total bacteria in GLC compared with GL in the spring compared to the fall.

Using high-throughput sequencing techniques, several studies have demonstrated the high complexity and diversity of bovine raw milk microbiota, from which the genera *Pseudomonas* and *Acinetobacter* are generally considered as the more common [5,6,17,72]. In the current study, the genus *Acinetobacter* was not found among the top 20 more abundant genera of the PMA-treated nor the PMA-free datasets, from which five

and 23 phylotypes, respectively, were identified. Instead, milk samples were co-dominated by the phylotypes of *Pseudomonas*, unidentified *Enterobacteriaceae*, *Serratia*, and to a lesser extent *Cellulosimicrobium*. Moreover, these genera which have been associated with refrigerated raw milk [7,43,70,310] encompassed the more abundant ASVs differentially enriched across milk groups in both sampling periods, particularly in GLICI compared with GLC. *Serratia liquefaciens* and *Pseudomonas* spp. have recently been reported as the predominant psychrotrophic bacteria that produce heat-resistant proteolytic and lipolytic enzymes with high spoilage potential [310,317]. In a concomitant study on forage microbiota (Chapter 2), it was found that phylotypes of *Pseudomonas* and *Serratia* were among the *Proteobacteria* that positively correlated with *Firmicutes*, and that members of these genera were generally more abundant in inoculated grass/legume silage compared with uninoculated counterparts. This suggests that in addition to warmer temperatures in the spring compared to fall, the farms using the GLICI forage ration along with the implemented farming practices probably contributed to higher loads of *Pseudomonas* and *Enterobacteriaceae* in corresponding milk samples.

However, we observed erratic patterns of raw milk contamination by *Lactobacillales*, exhibiting only three phylotypes as the core LAB from which two were assigned to the genus *Lactobacillus* and the last to *Lactococcus*. Similar results showing the thinness of the core microbiota from 112 cow milk samples were reported by Li et al. [6]. These authors found that *Acinetobacter* and *Pseudomonas* were the sole genera shared by all milk samples they analysed. Regarding LAB, the same authors reported a high prevalence of the genus *Lactococcus* (~99 %). Kable et al. reported a more diverse core microbiota encompassing 29 taxa at the genus level (from 899 raw milk samples), among which unidentified *Aerococcaceae*, *Enterococcus*, and *Streptococcus* represented the LAB community [54]. Recently, Parente et al. used the FoodMicrobionet database to analyse the results from five studies that examined a total of 199 bulk tank milk samples from various regions around the world, and found that the genera *Pseudomonas*, *Streptococcus*, *Lactococcus*, and *Acinetobacter* showed the highest prevalence rates of ~98, ~97, ~93, ~93 %, respectively [37]. Using a culture-dependent approach, Gagnon et al. analysed 1,239 LAB isolates from 48 bulk tank milk samples and found that *Lacticaseibacillus casei/paracasei*, *P. pentosaceus*, *Weissella paramesenteroides/thailandensis*, and *Lactococcus lactis* were the most prevalent with 60, 42, 40, and 30 % prevalence rates, respectively [22]. Moreover, these authors revealed that despite the substantial enrichment of 35 % in lactobacilli in the microbiota of inoculated grass/legume silage compared with uninoculated counterparts, the associated milk samples exhibited similar LAB profiles. These findings support the hypothesis that there is not a clearly defined pattern of raw milk contamination on dairy farms, particularly for LAB.

In this study, some phylotypes of *Streptococcus*, *Staphylococcus*, *Corynebacterium*, *Enterobacteriaceae*, *Acinetobacter*, *Jeotgalicoccus*, *Rothia*, *Rhodococcus*, *Aerosphaera*, *Budvicia*, *Brevundimonas*, *Solibacillus*, *Pseudoclavibacter*, *Enterococcus*, *Frigoribacterium*, *Pantoea*, and *Bifidobacterium*

positively correlated with SCC. The genera *Streptococcus*, *Staphylococcus*, *Corynebacterium*, *Enterococcus*, *Acinetobacter*, as well as members of *Enterobacteriaceae* have been commonly identified as mastitis causing agents [52,306,318,319]. Regarding these taxa, our findings corroborate those of Rodrigues et al. [58] who reported positive correlations with milk SCC. The genera *Jeotgalicoccus*, *Bifidobacterium*, and *Solibacillus* were found among the most abundant taxa of the bovine teat microbiota [305,306], while *Brevundimonas* was identified as a dominant taxa in clinical mastitis samples [320]. In the current study, except for *Enterobacteriaceae*, none of the phylotypes encompassing these taxa were found differentially abundant among milk samples associated with the five forage combinations.

We found that phylotypes of *Pseudomonas*, *Rhodococcus*, *B. mongoliense*, *Kocuria*, *Chryseobacterium bovis*, *Brachybacterium*, *Antricoccus*, *Comamonas koreensis*, *Chryseobacterium*, *Lactobacillus*, and *Ottowia* negatively correlated with proteins. Apart from *Ottowia* and *Antricoccus*, the proteolytic activity of these taxa has been previously demonstrated [189,321–323]. Although these taxa correlated with some of the inoculants used for ensiling, none of the corresponding phylotypes mentioned above were differentially enriched between milk from the five feeding types. The genus *Ottowia* has been identified in a variety of environments such as activated sludge [324], archaeological dental calculus [325], or minced pork [326], whereas *Antricoccus* was reported in natural caves [327] and toothbrushes [328]. Although we did not find any report of the occurrence of these two genera in raw milk, the significance of their potential proteolytic activities should be investigated.

Several studies have analysed changes in the microbial communities of raw milk throughout the production chain from milking on dairy farms [7,43] to storage in silos at the processing facilities [18,54]. In this study, since milk samples were collected directly after milking, they correspond to time zero on the time scale associated with the production and transportation chain from milking to processing plants. Therefore, the constructed networks relate much more to bacterial entry into milk from the sources of contamination (during milking and in the bulk tank) than to a substantial growth of microorganisms in the milk matrix. In this context, because the network modules can be regarded as microbial niches [299], the phylotypes they contain, whether co-occurring or co-excluding, would logically share the ability to grow and populate the respective habitats they originate from. Hence, taxa co-occurrence would mainly reflect a milk co-contamination or a concurrent occurrence of taxa at the sources of contamination. For instance, most phylotypes found to correlate with SCC and that would plausibly originate from the internal and external areas around the teat or the udder shared the same modules across all networks, notably modules 1 and 3 in the GL or module 1 in the GLICI networks. Interestingly, these modules also contained typical taxa associated with the skin such as *Aerococcus* or *Acinetobacter* [308,329], or those known as mastitis pathogens located in the teat canal such as *Streptococcus*, *Escherichia*, *Staphylococcus*, and *Corynebacterium* [52,306,318,319]. The analysis of topological roles derived from the networks revealed phylotypes of *Pseudomonas*, *Rhodococcus*, and *Lactobacillus* as module hubs,

indicating their importance for their network stability [298]. Because these taxa may occur in a wide variety of contamination sources, it is difficult to interpret their relevance in their original communities. However, the observed variations of network topologies improve our understanding of differences between raw milk microbial communities associated with forage ration combinations [297].

Although silage has been considered as one of the most important sources of spore-forming and thermoduric bacteria that contaminate raw milk on the dairy farm [9,112,143,330], Doyle et al. [12] analysed several microbial habitats on the farm environment and found silage as a minor contributor to the milk microbiota compared with the teat surface and herd faeces. Using a culture-dependent method and random amplified polymorphic DNA, Gagnon et al. [22] recently confirmed their finding regarding LAB communities found in raw milk and that probably originated from silage. Here, using a metataxonomic approach, we complement these previous findings by providing more insight into the rates of microbial transfer from five forage ration combinations to corresponding raw milk on dairy farms.

Our results show that silage-based forage rations, particularly GLC and GLICI, share more phylotypes with raw milk produced on corresponding farms compared to that observed in the milk from cows fed a H ration. Among the 113 presumably transferred phylotypes, *Proteobacteria* were by far the most represented compared to *Firmicutes* and *Actinobacteria*, each at 65, 23, and 12 %, respectively. Rather than observing a significant enrichment of *Lactobacillaceae* in milk samples from the GLICI forage type as they dominated the microbiota of the corresponding forage ration, phylotypes assigned to *Enterobacteriaceae* (mainly *Serratia*, unidentified *Enterobacteriaceae*, *Yersinia*, and *Hafnia-Obesumbacterium*), *Pseudomonadaceae* (*Pseudomonas*), *Promicromonosporaceae* (*Cellulosimicrobium*), and *Aeromonadaceae* (*Aeromonas*) were listed among the differentially abundant taxa and were the most represented. Interestingly, 92 % of phylotypes enriched in the GLICI milk, among which all those cited above, were identified in the microbiota of the associated forage ration, as were 93 % of those enriched in the GLC milk. However, these proportions were considerably reduced in the ration involving a single forage type (H or GL). These findings clearly demonstrate that although bacteria from forage may represent a low proportion of the associated raw milk microbiota, they may be the main taxa distinguishing between milk from different feeding combinations. Our results show that the mixture of grass/legume and corn silage significantly impacts the raw milk microbiota compared with a single forage-based ration. Considering the case of GLICI versus GLC milk samples, it appears that differences in their microbial communities were mostly driven by higher relative abundances of *Proteobacteria* in the GLICI forage type. However, this is not specific to GLICI milk as similar observations can be made when the same comparison is performed between other feeding combinations. Therefore, based on the current study, it is difficult to provide reliable explanations on a direct influence of silage inoculants on raw milk microbiota upon milking.

However, the significance of other sources of milk contamination may explain the observed low proportions of the shared phylotypes relative to those uniquely occurring in the forage types or milk samples from the same feeding combination. These sources of microorganisms include the bedding material, faeces, cow skin, water, humans, milking machines and pipelines, bulk tank, air, pasture, and other feed components [5,19,143,308]. These observations suggest a lack of correlation between taxa abundance in forage and their abundance in milk. Supporting this hypothesis, Driehuis et al. [112] reported in their review dealing with the impact of silage on the quality of dairy foods that raw milk contamination by aerial spores from silage or by direct contact of raw milk with silage are negligeable when milking hygiene is properly applied. Bacteria from silage therefore take indirect milk contamination routes, possibly involving sporadic transfer of silage onto the bedding or directly to the cow skin (of which the teat surface is cleaned before milking), improper human handling, or via faeces that can contaminate the bedding and the teat surface. Indeed, spore forming bacteria from silage were found to withstand harsh conditions along the cow gastrointestinal tract and subsequently end up in the faeces [290,331]. Although still under debate, it should be considered that silage bacteria may translocate via the entero-mammary pathway previously described [41,72,332]. On a dairy farm, the interconnections among the microbial sources, which by themselves can be selective habitats, might explain why a clear pattern of milk contamination by silage bacteria was not in evidence in this study.

4.7 Conclusion

The microbiota of forages can be grouped into three community types broadly distinguishing between H, GL, and GLICI samples, GLC showing high similarity with GLICI. However, a subsequent classification of microbial communities in milk associated with the forage ration combinations was not observed. Since milk samples exhibited significant variation in microbial community across sampling periods, we can presume that seasonality would have greater influence on the milk microbiota than forage rations. Nevertheless, the effect of forage ration combinations on the milk microbiota appeared more substantial in the spring, as significantly higher loads of LAB, *Pseudomonas*, *Enterobacteriaceae*, and total bacteria were observed in GLICI compared with milk samples associated with other feeding combinations. This study was carried out using freshly produced bulk tank milk, for which we demonstrated erratic patterns of contamination at the farm. Bacteria from forage rations encompassing H, GL, GLC, GLCI, and GLICI may account for up to 31 % of microbial community in the corresponding milk. Trends of direct contamination of milk by forage bacteria were not evidenced for any of the 113 phylotypes presumably transferred from forage to milk. Although significant differences were observed between GLICI and GLC milk samples, they were driven more by *Enterobacteriaceae* and other *Proteobacteria*, rather than by LAB communities. Drawing reliable conclusions on the influence of silage inoculants on the raw milk microbial community is therefore challenging. Additional milk samples taken at the end of the transport chain from dairy farms to processing plants may reveal further effects of transport on raw milk microbiota. Further

investigations involving more farm types and the integration of metagenomics and metabolomics would be needed to better understand the impact of cow feeding with inoculated silage on milk quality and processability.

Conclusion

Les travaux réalisés dans le cadre de cette thèse avaient pour but d'améliorer le niveau de connaissance sur le microbiote des fourrages préservés (foin, ensilages d'herbe/légume et de maïs inoculés ou non inoculés) sur le microbiote du lait cru. Ces travaux étaient axés sur l'hypothèse que l'analyse de la diversité et de la prévalence des communautés microbiennes présentes dans les ensilages et le lait cru permet de connaître l'impact de ces fourrages fermentés sur la qualité microbiologique du lait. Afin d'obtenir le maximum d'information concernant le lien entre les pratiques de gestion des fourrages dans les fermes laitières et la composition bactérienne du lait cru, la métataxonomique a été placée au cœur des investigations. Par cette approche d'analyse des communautés microbiennes dite culture-indépendante qui s'appuie sur le séquençage à haut-débit d'une portion du gène de l'ARNr 16S, nous avons démontré que les espèces bactériennes présentes dans les fourrages sec ou fermentés peuvent être transférées au lait cru. Les résultats obtenus suggèrent que dans l'environnement des fermes laitières, ces transferts se feraient de manière sporadique. L'intégration des paramètres physicochimiques des fourrages et du lait cru à l'analyse des données issues du séquençage par des méthodes statistiques multivariées et multi-table ont permis de déterminer les corrélations entre ces paramètres et les taxons (définis au niveau de variant de séquence), et ainsi de mieux comprendre l'écologie microbienne des fourrages et du lait cru.

L'originalité de cette thèse se décline en trois thèmes principaux, à savoir la préservation de l'intégrité du microbiote du lait cru conservé aux fins d'analyses microbiologiques, l'effet des pratiques de gestion des fourrages préservés sur leur microbiote viable, et l'impact de ces pratiques de gestion sur la composition microbienne du lait cru. Compte tenu de l'inévitable croissance microbienne observée dans le lait cru au cours du temps malgré la réfrigération, ou encore l'altération des parois cellulaires des bactéries lors de la congélation, le maintien de la diversité et de l'abondance des espèces microbiennes qu'il contient est primordial et constitue un gage de la qualité et de fiabilité des résultats d'analyse du microbiote. Tout en mettant en évidence l'altération du microbiote du lait lors de sa conservation sous réfrigération (4 °C) ou sous congélation (-20 °C), nous avons démontré que l'azidiol utilisé seul ou en combinaison avec le DMSO permet de maintenir l'intégrité et la viabilité du microbiote à court ou moyen terme. Nos résultats suggèrent également qu'en plus de l'azidiol et du DMSO, l'ajout d'un composé antifongique pour limiter la croissance des levures et des champignons microscopiques permettrait une plus longue durée de conservation d'échantillons de lait. De futurs travaux réalisés dans cette perspective amélioreront les conditions de préservation d'échantillons de lait cru, particulièrement lors d'études d'envergure impliquant la microbiologie du lait ou des produits laitiers. Par ailleurs, nos travaux montrent clairement que le choix de la région du gène de l'ARNr 16S à séquencer peut considérablement affecter l'interprétation des résultats et rendre inappropriée la comparaison de ces derniers avec ceux disponibles dans la littérature. La convergence des efforts de recherche vers un consensus méthodologique des pratiques

d'analyse du microbiote du lait et des produits laitiers de la collecte d'échantillons à l'analyse des séquences est vivement souhaitable.

Il ressort de l'analyse du microbiote des fourrages que le foin présente une structure microbienne très différente de celle des ensilages. Bien que les inoculants soient recommandés pour améliorer la qualité des ensilages, nos résultats ont montré que leur usage n'induit pas systématiquement une modification conséquente du microbiote. D'autres facteurs tels que le type de silo ou la période d'échantillonnage (automne ou printemps) influenceraient concurremment la composition et la structure microbiennes des ensilages d'herbe/légume et de maïs, et ainsi masqueraient les effets spécifiques des inoculants. Toutefois, les analyses de réseaux ayant révélé des différences topologiques entre les ensilages inoculés et non inoculés, il se pourrait que les inoculants aient une influence considérable sur la co-occurrence des espèces bactériennes. Ceci démontre l'intérêt de l'analyse des réseaux de co-occurrence ou de co-exclusion pour une meilleure compréhension de l'écologie microbienne des ensilages. Les résultats obtenus montrent également qu'au niveau des variants de séquences, seules les phylotypes présents à la fois dans les ensilages inoculés et non inoculés sont les plus dominants au sein des bactéries lactiques. Cependant, l'inoculation des ensilages d'herbe ou de légumineuse réduit grandement la prévalence et l'abondance des espèces du genre *Weissella*, alors que dans le cas des ensilages de maïs, les lactobacilles sont dominants avec ou sans l'ajout d'inoculants. Malgré l'application d'inoculants lors de la préparation des ensilages de maïs, la prolifération des bactéries du genre *Acetobacter* semble récurrente, particulièrement à l'automne. L'ambigüité autour des propriétés fermentatives de cette bactérie ou de sa capacité à initier la détérioration aérobie des ensilages de maïs nécessite d'avantages de travaux de recherches visant à préciser les conditions de développement de cette bactérie et déterminer ses interactions (compétition, coopération) avec les bactéries lactiques. De tels travaux impliqueraient évidemment la recherche de nouveaux inoculants plus compétitifs et capables de prolonger par leur action inhibitrice, la stabilité aérobie des ensilages.

Les pratiques de gestion des fourrages à la ferme impliquent aussi la préparation des rations pour les vaches. Au cours de nos travaux, cinq types de rations à base de fourrage ont été recensés, à savoir le foin (H), les ensilages d'herbe/légume non inoculés (GL), les mélanges d'ensilages d'herbe/légumineuse non inoculés et de maïs non inoculés (GLC), les mélanges d'ensilage d'herbe/légumineuse non inoculés et de maïs inoculés (GLCI), et les mélanges d'ensilage d'herbe/légumineuse et de maïs inoculés (GLICI). L'analyse microbiologique de ces rations a révélé trois types de communauté microbienne distinguant principalement H, GL, et GLICI les unes des autres. Cependant, seules les populations bactériennes présentes dans les échantillons de lait associés aux fermes GLC et GLICI ont présenté des différences significatives entre elles. Toutefois, il est intéressant de noter que la charge bactérienne des lactobacilles pourtant utilisés comme inoculants n'est pas plus importante dans les échantillons de lait des fermes GLICI comparé à ceux des fermes GLC. De plus, le microbiote du lait varie plus en fonction de la période d'échantillonnage (automne, printemps) que selon le type

de fourrage dans la ration. Nos travaux montrent qu'au niveau de la ferme laitière, les transferts de bactéries des ensilages au lait cru ne semblent pas suivre un modèle prédictible, car s'effectuant possiblement de manière aléatoire. Néanmoins, il apparaît que les rations pour vaches à base principalement d'ensilages, contribuent jusqu'à 31 % à la formation des communautés bactériennes du lait cru. De plus, environ 92 % des phylotypes enrichis uniquement dans l'une ou l'autre des fermes GLC et GLICI sont également présents dans les ensilages correspondants. Ainsi, plutôt que par le nombre d'espèces bactériennes ou leur abondance respective, l'importance des rations à base d'ensilages d'herbe/légumes et de maïs comme source de contamination du lait se démontre plus par l'introduction dans le lait d'espèces bactériennes qui déterminent les dissimilarités du microbiote. D'après nos résultats, les genres bactériens potentiellement transférés des ensilages au lait sont surtout des *Proteobacteria* telles que *Serratia*, *Pseudomonas* et *Lelliottia*, les *Firmicutes* telles que *Lactobacillus*, *Lactococcus*, *Clostridium* et *Romboutsia*, et les *Actinobacteria* telles que *Cellulosimicrobium*. Il est important de noter que parmi les lactobacilles, un phylotype de *Lentilactobacillus buchneri* a été identifié dans les échantillons d'ensilages et de lait des fermes GLICI. Considérant le caractère sporadique du transfert de bactéries d'ensilages vers le lait cru, toute bactérie, y compris l'inoculant, présente dans l'ensilage peut contaminer le lait dépendamment de ses aptitudes à survivre ou à se développer dans les sites de contamination. L'analyse des réseaux de co-occurrence des phylotypes a révélé dans l'environnement à la ferme des niches microbiennes qui rappellent en effet les sources de contamination du lait telles que les surfaces intérieures et extérieures du pis. Cependant, les résultats obtenus ne permettent pas d'établir avec précision l'impact des inoculants sur la composition microbienne du lait. Toutefois, nos travaux démontrent clairement que les ensilages constituent une source de contamination du lait par les espèces potentiellement pathogènes pour les animaux, d'altération et d'intérêt technologique pour le lait et les produits dérivés. Outre les bactéries, les levures et moisissures font partie intégrante du microbiote des ensilages et du lait cru. Si leur importance dans la détérioration des ensilages est bien connue, peu d'informations sont disponibles quant à leur diversité et leur importance pour l'aptitude du lait à la transformation ou encore pour la qualité des produits laitiers. Des travaux semblables à ceux présentés dans cette thèse, orientés sur l'impact des fourrages préservés sur les communautés fongiques du lait cru permettraient d'améliorer nos connaissances des microorganismes qui déterminent la qualité microbiologique du lait.

Bibliographie

1. Muehlhoff, E.; Bennett, A.; McMahon, D. *Milk and dairy products in human nutrition.* Food and Agricultural Organization of the United Nations (FAO), 2013.

2. Jacques, L.-S.; Mondon, M.; LeBreton, M.; Karras, T. *Les retombées économiques de l'industrie laitière au Canada*. ÉcoRessources Consultants, 2011.

3. Statistiques laitières et informations sur les marchés - centre canadien d'information laitière. Available online: https://www.dairyinfo.gc.ca/fra/statistiques-laitieres-et-informations-sur-les-marches/?id=1502465642636 (accessed on Apr 14, 2021).

4. Bradley, A.J.; Leach, K.A.; Green, M.J.; Gibbons, J.; Ohnstad, I.C.; Black, D.H.; Payne, B.; Prout, V.E.; Breen, J.E. The impact of dairy cows' bedding material and its microbial content on the quality and safety of milk - a cross sectional study of UK farms. *Int. J. Food Microbiol.* 2018, *269*, 36–45.

5. Quigley, L.; O'Sullivan, O.; Stanton, C.; Beresford, T.P.; Ross, R.P.; Fitzgerald, G.F.; Cotter, P.D. The complex microbiota of raw milk. *FEMS Microbiol. Rev.* 2013, *37*, 664–98.

6. Li, N.; Wang, Y.; You, C.; Ren, J.; Chen, W.; Zheng, H.; Liu, Z. Variation in raw milk microbiota throughout 12 months and the impact of weather conditions. *Sci. Rep.* 2018, *8*, 2371.

7. Rasolofo, E.A.; St-Gelais, D.; LaPointe, G.; Roy, D. Molecular analysis of bacterial population structure and dynamics during cold storage of untreated and treated milk. *Int. J. Food Microbiol.* 2010, *138*, 108–18.

8. Desfossés-foucault, É.; Lapointe, G.; Roy, D. Dynamics and rRNA transcriptional activity of lactococci and lactobacilli during Cheddar cheese ripening. *Int. J. Food Microbiol.* 2013, *166*, 117–124.

9. Julien, M.-C.; Dion, P.; Lafrenière, C.; Antoun, H.; Drouin, P. Sources of clostridia in raw milk on farms. *Appl. Environ. Microbiol.* 2008, *74*, 6348–57.

10. Podrzaj, L.; Burtscher, J.; Küller, F.; Domig, K.J. Strain-dependent cheese spoilage potential of *Clostridium tyrobutyricum*. *Microorganisms* 2020, *8*, 1836.

11. Martin, N.H.; Torres-Frenzel, P.; Wiedmann, M. Invited review: controlling dairy product spoilage to reduce food loss and waste. *J. Dairy Sci.* 2021, *104*, 1251–1261.

12. Doyle, C.J.; Gleeson, D.; O'Toole, P.W.; Cotter, P.D. Impacts of seasonal housing and teat preparation on raw milk microbiota: a high-throughput sequencing study. *Appl. Environ. Microbiol.* 2017, *83*, e02694-16.

13. Falardeau, J.; Keeney, K.; Trmčić, A.; Kitts, D.; Wang, S. Farm-to-fork profiling of bacterial communities associated with an artisan cheese production facility. *Food Microbiol.* 2019, *83*, 48–58.

14. Du, B.; Meng, L.; Liu, H.; Zheng, N.; Zhang, Y.; Guo, X.; Zhao, S.; Li, F.; Wang, J. Impacts of milking and housing environment on milk microbiota. *Animals* 2020, *10*, 1–11, 2339.

15. Skeie, S.B.; Håland, M.; Thorsen, I.M.; Narvhus, J.; Porcellato, D. Bulk tank raw milk microbiota differs within and between farms: a moving goalpost challenging quality control. *J. Dairy Sci.* 2019, *102*, 1959–1971.

16. Ruvalcaba-Gómez, J.M.; Delgado-Macuil, R.J.; Zelaya-Molina, L.X.; Maya-Lucas, O.; Ruesga-Gutiérrez, E.; Anaya-Esparza, L.M.; Villagrán-de la Mora, Z.; López-de la Mora, D.A.; Arteaga-Garibay, R.I. Bacterial succession through the artisanal process and seasonal effects defining bacterial communities of raw-milk Adobera cheese revealed by high-throughput DNA sequencing. *Microorganisms* 2020, *9*, 24.

17. Doyle, C.J.; Gleeson, D.; O'Toole, P.W.; Cotter, P.D. High-throughput metataxonomic characterization of the raw milk microbiota identifies changes reflecting lactation stage and storage conditions. *Int. J. Food Microbiol.* 2017, *255*, 1–6.

18. McHugh, A.J.; Feehily, C.; Fenelon, M.A.; Gleeson, D.; Hill, C.; Cotter, P.D. Tracking the dairy microbiota from farm bulk tank to skimmed milk powder. *mSystems* 2020, *5*, e00226-20.

19. Gagnon, M.; Hamelin, L.; Fréchette, A.; Dufour, S.; Roy, D. Effect of recycled manure solids as bedding on bulk tank milk and implications for cheese microbiological quality. *J. Dairy Sci.* 2020, *103*, 128–140.

20. Panelli, S.; Brambati, E.; Bonacina, C.; Feligini, M. Diversity of fungal flora in raw milk from the Italian Alps in relation to pasture altitude. *Springerplus* 2013, *2*, 405.

21. Frétin, M.; Martin, B.; Rifa, E.; Isabelle, V.M.; Pomiès, D.; Ferlay, A.; Montel, M.C.; Delbès, C. Bacterial community assembly from cow teat skin to ripened cheeses is influenced by grazing systems. *Sci. Rep.* 2018, *8*, 1–11.

22. Gagnon, M.; Ouamba, A.J.K.; LaPointe, G.; Chouinard, P.Y.; Roy, D. Prevalence and abundance of

lactic acid bacteria in raw milk associated with forage types in dairy cow feeding. *J. Dairy Sci.* 2020, *103*, 5931–5946.

23. Muck, R.E.; Shinners, K.J. Conserved forage (silage and hay): progress and priorities. In Proceedings of the 19th international grassland congress; São Pedro, Brazil, 2001; pp. 753–762.

24. Muck, R.E. Recent advances in silage microbiology. *Agric. Food Sci.* 2013, *22*, 3–15.

25. Wilkinson, J.M.; Rinne, M. Highlights of progress in silage conservation and future perspectives. *Grass Forage Sci.* 2018, *73*, 40–52.

26. Muck, R.E.; Nadeau, E.M.G.; McAllister, T.A.; Contreras-Govea, F.E.; Santos, M.C.; Kung, L. Silage review: recent advances and future uses of silage additives. *J. Dairy Sci.* 2018, *101*, 3980–4000.

27. Vissers, M.M.M.; Driehuis, F.; Te Giffel, M.C.; De Jong, P.; Lankveld, J.M.G. Minimizing the level of butyric acid bacteria spores in farm tank milk. *J. Dairy Sci.* 2007, *90*, 3278–3285.

28. Vissers, M.M.M.; Te Giffel, M.C.; Driehuis, F.; De Jong, P.; Lankveld, J.M.G. Minimizing the level of *Bacillus cereus* spores in farm tank milk. *J. Dairy Sci.* 2007, *90*, 3286–3293.

29. Colombo, M.; Castilho, N.P.A.; Todorov, S.D.; Nero, L.A. Beneficial properties of lactic acid bacteria naturally present in dairy production. *BMC Microbiol.* 2018, *18*, 219.

30. Bernier Gosselin, V.; Dufour, S.; Adkins, P.R.F.; Middleton, J.R. Persistence of coagulase negative staphylococcal intramammary infections in dairy goats. *J. Dairy Res.* 2019, *86*, 211–216.

31. Yuan, L.; Sadiq, F.A.; Burmolle, M.; Wang, N.; He, G. Insights into psychrotrophic bacteria in raw milk: a review. *J. Food Prot.* 2019, *82*, 1148–1159.

32. Gobbetti, M.; Di Cagno, R.; Calasso, M.; Neviani, E.; Fox, P.F.; De Angelis, M. Drivers that establish and assembly the lactic acid bacteria biota in cheeses. *Trends Food Sci. Technol.* 2018, *78*, 244–254.

33. Murphy, S.C.; Martin, N.H.; Barbano, D.M.; Wiedmann, M. Influence of raw milk quality on processed dairy products: how do raw milk quality test results relate to product quality and yield? *J. Dairy Sci.* 2016, *99*, 10128–10149.

34. Verdier-Metz, I.; Michel, V.; Delbès, C.; Montel, M.-C. Do milking practices influence the bacterial diversity of raw milk? *Food Microbiol.* 2009, *26*, 305–310.

35. Verdier-Metz, I.; Gagne, G.; Bornes, S.; Monsallier, F.; Veisseire, P.; Delbès-Paus, C.; Montel, M.-C. Cow teat skin, a potential source of diverse microbial populations for cheese production. *Appl. Environ. Microbiol.* 2012, *78*, 326–233.

36. Vacheyrou, M.; Normand, A.-C.; Guyot, P.; Cassagne, C.; Piarroux, R.; Bouton, Y. Cultivable microbial communities in raw cow milk and potential transfers from stables of sixteen French farms. *Int. J. Food Microbiol.* 2011, *146*, 253–262.

37. Parente, E.; Ricciardi, A.; Zotta, T. The microbiota of dairy milk: a review. *Int. Dairy J.* 2020, *107*, 104714.

38. Murphy, S.; Kent, D.; Martin, N.; Evanowski, R.; Patel, K.; Godden, S.; Wiedmann, M. Bedding and bedding management practices are associated with mesophilic and thermophilic spore levels in bulk tank raw milk. *J. Dairy Sci.* 2019, *102*, 6885–6900.

39. Jayarao, B.M.; Pillai, S.R.; Sawant, A.A.; Wolfgang, D.R.; Hegde, N. V. Guidelines for monitoring bulk tank milk somatic cell and bacterial counts. *J. Dairy Sci.* 2004, *87*, 3561–3573.

40. Breitenwieser, F.; Doll, E. V.; Clavel, T.; Scherer, S.; Wenning, M. Complementary use of cultivation and high-throughput amplicon sequencing reveals high biodiversity within raw milk microbiota. *Front. Microbiol.* 2020, *11*, 1557.

41. Oikonomou, G.; Addis, M.F.; Chassard, C.; Nader-Macias, M.E.F.; Grant, I.; Delbès, C.; Bogni, C.I.; Le Loir, Y.; Even, S. Milk microbiota: what are we exactly talking about? *Front. Microbiol.* 2020, *11*, 60.

42. Mallet, A.; Guéguen, M.; Kauffmann, F.; Chesneau, C.; Sesboué, A.; Desmasures, N. Quantitative and qualitative microbial analysis of raw milk reveals substantial diversity influenced by herd management practices. *Int. Dairy J.* 2012, *27*, 13–21.

43. Kennang Ouamba, A.J.; LaPointe, G.; Dufour, S.; Roy, D. Optimization of preservation methods allows deeper insights into changes of raw milk microbiota. *Microorganisms* 2020, *8*, 368.

44. Lagier, J.-C.; Armougom, F.; Million, M.; Hugon, P.; Pagnier, I.; Robert, C.; Bittar, F.; Fournous, G.; Gimenez, G.; Maraninchi, M.; et al. Microbial culturomics: paradigm shift in the human gut microbiome study. *Clin. Microbiol. Infect.* 2012, *18*, 1185–1193.

45. Lagier, J.C.; Dubourg, G.; Million, M.; Cadoret, F.; Bilen, M.; Fenollar, F.; Levasseur, A.; Rolain, J.M.; Fournier, P.E.; Raoult, D. Culturing the human microbiota and culturomics. *Nat. Rev. Microbiol.* 2018, *16*, 540–550.

46. Afouda, P.; Hocquart, M.; Pham, T.P.T.; Kuete, E.; Ngom, I.I.; Dione, N.; Valles, C.; Bellali, S.; Lagier, J.C.; Dubourg, G.; et al. Alcohol pretreatment of stools effect on culturomics. *Sci. Rep.* 2020, *10*, 1–10.

47. Diakite, A.; Dubourg, G.; Dione, N.; Afouda, P.; Bellali, S.; Ngom, I.I.; Valles, C.; Million, M.; Levasseur, A.; Cadoret, F.; et al. Extensive culturomics of 8 healthy samples enhances metagenomics efficiency. *PLoS One* 2019, *14*, e0223543.

48. Bilen, M.; Dufour, J.C.; Lagier, J.C.; Cadoret, F.; Daoud, Z.; Dubourg, G.; Raoult, D. The contribution of culturomics to the repertoire of isolated human bacterial and archaeal species. *Microbiome* 2018, *6*, 94.

49. Fonkou, M.D.M.; Dufour, J.C.; Dubourg, G.; Raoult, D. Repertoire of bacterial species cultured from the human oral cavity and respiratory tract. *Future Microbiol.* 2018, *13*, 1611–1624.

50. Zehavi, T.; Probst, M.; Mizrahi, I. Insights into culturomics of the rumen microbiome. *Front. Microbiol.* 2018, *9*, 1999.

51. Pereira, A.C.; Cunha, M. V. An effective culturomics approach to study the gut microbiota of mammals. *Res. Microbiol.* 2020, *171*, 290–300.

52. Derakhshani, H.; Fehr, K.B.; Sepehri, S.; Francoz, D.; Buck, J. De; Barkema, H.W.; Plaizier, J.C.; Khafipour, E. Invited review : microbiota of the bovine udder : contributing factors and potential implications for udder health and mastitis susceptibility. *J. Dairy Sci.* 2018, *101*, 10605–10625.

53. Dobranić, V.; Kazazić, S.; Filipović, I.; Mikulec, N.; Zdolec, N. Composition of raw cow's milk microbiota and identification of enterococci by MALDI-TOF MS - short communication. *Vet. Arh.* 2016, *86*, 581–590.

54. Kable, M.E.; Srisengfa, Y.; Laird, M.; Zaragoza, J.; McLeod, J.; Heidenreich, J.; Marco, M.L. The core and seasonal microbiota of raw bovine milk in tanker trucks and the impact of transfer to a milk processing facility. *MBio* 2016, *7*, e00836-16.

55. Kable, M.E.; Srisengfa, Y.; Xue, Z.; Coates, L.C.; Marco, M.L. Viable and total bacterial populations undergo equipment- and time-dependent shifts during milk processing. *Appl. Environ. Microbiol.* 2019, *85*, e00270-19.

56. Robles, I.; Kelton, D.F.; Barkema, H.W.; Keefe, G.P.; Roy, J.P.; von Keyserlingk, M.A.G.; DeVries, T.J. Bacterial concentrations in bedding and their association with dairy cow hygiene and milk quality. *Animal* 2020, *14*, 1052–1066.

57. Miller, R.A.; Kent, D.J.; Boor, K.J.; Martin, N.H.; Wiedmann, M. Different management practices are associated with mesophilic and thermophilic spore levels in bulk tank raw milk. *J. Dairy Sci.* 2015, *98*, 4338–4351.

58. Rodrigues, M.X.; Lima, S.F.; Bicalho, R.C. The microbiome of bulk tank milk : characterization and associations with somatic cell count and bacterial count. *J. Dairy Sci.* 2017, *100*, 2536–2552.

59. Evanowski, R.L.; Kent, D.J.; Wiedmann, M.; Martin, N.H. Milking time hygiene interventions on dairy farms reduce spore counts in raw milk. *J. Dairy Sci.* 2020, *103*, 4088–4099.

60. de Koning, K.; Slaghuis, B.; van der Vorst, Y. Robotic milking and milk quality: effects on bacterial counts, somatic cell counts, freezing point and free fatty acids. *Ital. J. Anim. Sci.* 2003, *2*, 291–299.

61. Castro, Á.; Pereira, J.M.; Amiama, C.; Barrasa, M. Long-term variability of bulk milk somatic cell and bacterial counts associated with dairy farms moving from conventional to automatic milking systems. *Ital. J. Anim. Sci.* 2017, *17*, 218–225.

62. Klungel, G.H.; Slaghuis, B.A.; Hogeveen, H. The effect of the introduction of automatic milking systems on milk quality. *J. Dairy Sci.* 2000, *83*, 1998–2003.

63. Bava, L.; Colombini, S.; Zucali, M.; Decimo, M.; Morandi, S.; Silvetti, T.; Brasca, M.; Tamburini, A.; Crovetto, G.M.; Sandrucci, A. Efficient milking hygiene reduces bacterial spore contamination in milk. *J. Dairy Res.* 2017, *84*, 322–328.

64. Zucali, M.; Bava, L.; Colombini, S.; Brasca, M.; Decimo, M.; Morandi, S.; Tamburini, A.; Crovetto, G.M. Management practices and forage quality affecting the contamination of milk with anaerobic spore-forming bacteria. *J. Sci. Food Agric.* 2015, *95*, 1294–1302.

65. Tremonte, P.; Tipaldi, L.; Succi, M.; Pannella, G.; Falasca, L.; Capilongo, V.; Coppola, R.; Sorrentino, E. Raw milk from vending machines: effects of boiling, microwave treatment, and refrigeration on microbiological quality. *J. Dairy Sci.* 2014, *97*, 3314–3320.

66. Mikulová, M. Content of free fatty acids, lipolytic bacteria and somatic cells in relation to milking technology. *J. Agrobiol.* 2011, *28*, 49–54.

67. Pukančíková, L.; Lipničanová, S.; Kačániová, M.; Chmelová, D.; Ondrejovič, M. Natural microflora of raw cow milk and their enzymatic spoilage potential. *Nov. Biotechnol. Chim.* 2016, *15*, 142–155.

68. Tyapugin, E.A.; Tyapugin, S.E.; Simonov, G.A.; Uglin, V.K.; Nikiforov, V.E.; Serebrova, I.S. Comparative evaluation of technological factors affecting milk production and quality with various milking technologies. *Russ. Agric. Sci.* 2015, *41*, 266–270.

69. Kmiha, S.; Aouadhi, C.; Klibi, A.; Jouini, A.; Béjaoui, A.; Mejri, S.; Maaroufi, A. Seasonal and regional occurrence of heat-resistant spore-forming bacteria in the course of ultra-high temperature milk production in Tunisia. *J. Dairy Sci.* 2017, *100*, 6090–6099.

70. Vithanage, N.R.; Dissanayake, M.; Bolge, G.; Palombo, E.A.; Yeager, T.R.; Datta, N. Biodiversity of culturable psychrotrophic microbiota in raw milk attributable to refrigeration conditions, seasonality and their spoilage potential. *Int. Dairy J.* 2016, *57*, 80–90.

71. Quigley, L.; O'Sullivan, O.; Beresford, T.P.; Ross, R.P.; Fitzgerald, G.F.; Cotter, P.D. Molecular approaches to analysing the microbial composition of raw milk and raw milk cheese. *Int. J. Food Microbiol.* 2011, *150*, 81–94.

72. Addis, M.F.; Tanca, A.; Uzzau, S.; Oikonomou, G.; Bicalho, R.C.; Moroni, P. The bovine milk microbiota: insights and perspectives from -omics studies. *Mol. Biosyst.* 2016, *12*, 2359–2372.

73. Tilocca, B.; Costanzo, N.; Morittu, V.M.; Spina, A.A.; Soggiu, A.; Britti, D.; Roncada, P.; Piras, C. Milk microbiota: characterization methods and role in cheese production. *J. Proteomics* 2020, *210*, 103534.

74. Marchesi, J.R.; Ravel, J. The vocabulary of microbiome research: a proposal. *Microbiome* 2015, *3*, 31.

75. Yang, B.; Wang, Y.; Qian, P.Y. Sensitivity and correlation of hypervariable regions in 16S rRNA genes in phylogenetic analysis. *BMC Bioinformatics* 2016, *17*, 1–8.

76. Fittipaldi, M.; Nocker, A.; Codony, F. Progress in understanding preferential detection of live cells using viability dyes in combination with DNA amplification. *J. Microbiol. Methods* 2012, *91*, 276–89.

77. Pollock, J.; Glendinning, L.; Wisedchanwet, T.; Watson, M. The madness of microbiome: attempting to find consensus "best practice" for 16S microbiome studies. *Appl. Environ. Microbiol.* 2018, *84*, e02627-17.

78. Hornung, B.V.H.; Zwittink, R.D.; Kuijper, E.J. Issues and current standards of controls in microbiome research. *FEMS Microbiol. Ecol.* 2019, *95*, 1–7.

79. O'Connell, A.; Ruegg, P.L.; Jordan, K.; O'Brien, B.; Gleeson, D. The effect of storage temperature and

duration on the microbial quality of bulk tank milk. *J. Dairy Sci.* 2016, *99*, 3367–3374.

80. Nocker, A.; Cheung, C.-Y.; Camper, A.K. Comparison of propidium monoazide with ethidium monoazide for differentiation of live vs. dead bacteria by selective removal of DNA from dead cells. *J. Microbiol. Methods* 2006, *67*, 310–320.

81. Gagnon, M.; Savard, P.; Rivière, A.; LaPointe, G.; Roy, D. Bioaccessible antioxidants in milk fermented by *Bifidobacterium longum* subsp. longum strains. *Biomed Res. Int.* 2015, *2015*, 1–12.

82. Desfossés-Foucault, E.; Dussault-Lepage, V.; Le Boucher, C.; Savard, P.; Lapointe, G.; Roy, D. Assessment of probiotic viability during Cheddar cheese manufacture and ripening using propidium monoazide-PCR quantification. *Front. Microbiol.* 2012, *3*, 350.

83. Rivière, A.; Gagnon, M.; Weckx, S.; Roy, D.; De Vuyst, L. Mutual cross-feeding interactions between *Bifidobacterium longum* subsp. longum NCC2705 and *Eubacterium rectale* ATCC 33656 explain the bifidogenic and butyrogenic effects of arabinoxylan oligosaccharides. *Appl. Environ. Microbiol.* 2015, *81*, 7767–7781.

84. Sayed, I.M.; Hammam, A.R.A.; Elfaruk, M.S.; Alsaleem, K.A.; Gaber, M.A.; Ezzat, A.A.; Salama, E.H.; Elkhawaga, A.A.; El-Mokhtar, M.A. Enhancement of the molecular and serological assessment of hepatitis E virus in milk samples. *Microorganisms* 2020, *8*, 1231.

85. Cheema, A.S.; Stinson, L.F.; Lai, C.T.; Geddes, D.T.; Payne, M.S. DNA extraction method influences human milk bacterial profiles. *J. Appl. Microbiol.* 2021, *130*, 142–156.

86. Herthnek, D.; Nielsen, S.S.; Lindberg, A.; Bölske, G. A robust method for bacterial lysis and DNA purification to be used with real-time PCR for detection of *Mycobacterium avium* subsp. paratuberculosis in milk. *J. Microbiol. Methods* 2008, *75*, 335–340.

87. Lima, S.F.; De Souza Bicalho, M.L.; Bicalho, R.C. Evaluation of milk sample fractions for characterization of milk microbiota from healthy and clinical mastitis cows. *PLoS One* 2018, *13*, e0193671.

88. Sun, L.; Dicksved, J.; Priyashantha, H.; Lundh; Johansson, M. Distribution of bacteria between different milk fractions, investigated using culture-dependent methods and molecular-based and fluorescent microscopy approaches. *J. Appl. Microbiol.* 2019, *127*, 1028–1037.

89. Ali-Vehmas, T.; Westphalen, P.; Myllys, V.; Sandholm, M. Binding of *Staphylococcus aureus* to milk fat globules increases resistance to penicillin-G. *J. Dairy Res.* 1997, *64*, 253–260.

90. Brisson, G.; Payken, H.F.; Sharpe, J.P.; Jiménez-Flores, R. Characterization of *Lactobacillus reuteri* interaction with milk fat globule membrane components in dairy products. *J. Agric. Food Chem.* 2010, *58*, 5612–5619.

91. Stinson, L.F.; Ma, J.; Rea, A.; Dymock, M.; Geddes, D.T. Centrifugation does not remove bacteria from the fat fraction of human milk. *Sci. Rep.* 2021, *11*, 572.

92. Lima, S.F.; Teixeira, A.G.; Lima, F.S.; Ganda, E.K.; Higgins, C.H.; Oikonomou, G.; Bicalho, R.C. The bovine colostrum microbiome and its association with clinical mastitis. *J. Dairy Sci.* 2017, *100*, 3031–3042.

93. Ogundele, M.O. Cytotoxicity of EDTA used in biological samples: Effect on some human breast-milk studies. *J. Appl. Toxicol.* 1999, *19*, 395–400.

94. Esteban, M.; Marcos, P.; Horna, C.; Galan-Malo, P.; Mata, L.; Pérez, M.D.; Calvo, M.; Sánchez, L. Evaluation of methods for DNA extraction from *Clostridium tyrobutyricum* spores and its detection by qPCR. *J. Microbiol. Methods* 2020, *169*, 105818.

95. Kozich, J.J.; Westcott, S.L.; Baxter, N.T.; Highlander, S.K.; Schloss, P.D. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the miseq illumina sequencing platform. *Appl. Environ. Microbiol.* 2013, *79*, 5112–5120.

96. Callahan, B.J.; Wong, J.; Heiner, C.; Oh, S.; Theriot, C.M.; Gulati, A.S.; McGill, S.K.; Dougherty, M.K. High-throughput amplicon sequencing of the full-length 16S rRNA gene with single-nucleotide resolution. *Nucleic Acids Res.* 2019, *47*, e103–e103.

97. Callahan, B.J.; Grinevich, D.; Thakur, S.; Balamotis, M.A.; Yehezkel, T.B. Ultra-accurate microbial amplicon sequencing directly from complex samples with synthetic long reads. *bioRxiv* 2021, 9, 130.

98. Callahan, B.J.; McMurdie, P.J.; Rosen, M.J.; Han, A.W.; Johnson, A.J.A.; Holmes, S.P. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* 2016, *13*, 581–583.

99. Teng, F.; Darveekaran Nair, S.S.; Zhu, P.; Li, S.; Huang, S.; Li, X.; Xu, J.; Yang, F. Impact of DNA extraction method and targeted 16S-rRNA hypervariable region on oral microbiota profiling. *Sci. Rep.* 2018, *8*, 16321.

100. Chen, Z.; Hui, P.C.; Hui, M.; Yeoh, Y.K.; Wong, P.Y.; Chan, M.C.W.; Wong, M.C.S.; Ng, S.C.; Chan, F.K.L.; Chan, P.K.S. Impact of preservation method and 16S rRNA hypervariable region on gut

microbiota profiling. *mSystems* 2019, *4*, e00271-18.

101. Joos, L.; Beirinckx, S.; Haegeman, A.; Debode, J.; Vandecasteele, B.; Baeyen, S.; Goormachtig, S.; Clement, L.; De Tender, C. Daring to be differential: metabarcoding analysis of soil and plant-related microbial communities using amplicon sequence variants and operational taxonomical units. *BMC Genomics* 2020, *21*, 733.

102. Barnes, C.J.; Rasmussen, L.; Asplund, M.; Knudsen, S.W.; Clausen, M.L.; Agner, T.; Hansen, A.J. Comparing DADA2 and OTU clustering approaches in studying the bacterial communities of atopic dermatitis. *J. Med. Microbiol.* 2020, *69*, 1293–1302.

103. Callahan, B.J.; McMurdie, P.J.; Holmes, S.P. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J.* 2017, *11*, 2639–2643.

104. Fricker, A.M.; Podlesny, D.; Fricke, W.F. What is new and relevant for sequencing-based microbiome research? A mini-review. *J. Adv. Res.* 2019, *19*, 105–112.

105. Bolyen, E.; Rideout, J.R.; Dillon, M.R.; Bokulich, N.A.; Abnet, C.C.; Al-Ghalith, G.A.; Alexander, H.; Alm, E.J.; Arumugam, M.; Asnicar, F.; et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* 2019, *37*, 852–857.

106. Edgar, R.C. Updating the 97% identity threshold for 16S ribosomal RNA OTUs. *Bioinformatics* 2018, *34*, 2371–2375.

107. DeSantis, T.Z.; Hugenholtz, P.; Larsen, N.; Rojas, M.; Brodie, E.L.; Keller, K.; Huber, T.; Dalevi, D.; Hu, P.; Andersen, G.L. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* 2006, *72*, 5069–5072.

108. Cole, J.R.; Wang, Q.; Fish, J.A.; Chai, B.; McGarrell, D.M.; Sun, Y.; Brown, C.T.; Porras-Alfaro, A.; Kuske, C.R.; Tiedje, J.M. Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res.* 2014, *42*, D633–D642.

109. Yarza, P.; Richter, M.; Peplies, J.; Euzeby, J.; Amann, R.; Schleifer, K.; Ludwig, W.; Glöckner, F..; Rossello-Mora, R. The All-Species Living Tree Project: a 16S rRNA-based phylogenetic tree of all sequenced type strains. *Syst. Appl. Microbiol.* 2008, *31*, 241–250.

110. Abellan-Schneyder, I.; Matchado, M.S.; Reitmeier, S.; Sommer, A.; Sewald, Z.; Baumbach, J.; List, M.; Neuhaus, K. Primer, pipelines, parameters: issues in 16S rRNA gene sequencing. *mSphere* 2021, *6*,

e01202-20.

111. Choi, J.; In Lee, S.; Rackerby, B.; Frojen, R.; Goddik, L.; Ha, S.-D.; Park, S.H. Assessment of overall microbial community shift during Cheddar cheese production from raw milk to aging. *Appl. Microbiol. Biotechnol.* 2020, *104*, 6249–6260.

112. Driehuis, F. Silage and the safety and quality of dairy foods: a review. *Agric. Food Sci.* 2013, *22*, 16–34.

113. Wu, H.; Nguyen, Q.D.; Tran, T.T.M.; Tang, M.T.; Tsuruta, T.; Nishino, N. Rumen fluid, feces, milk, water, feed, airborne dust, and bedding microbiota in dairy farms managed by automatic milking systems. *Anim. Sci. J.* 2019, *90*, 445–452.

114. Knights, D.; Kuczynski, J.; Charlson, E.S.; Zaneveld, J.; Mozer, M.C.; Collman, R.G.; Bushman, F.D.; Knight, R.; Kelley, S.T. Bayesian community-wide culture-independent microbial source tracking. *Nat. Methods* 2011, *8*, 761–765.

115. F Escapa, I.; Huang, Y.; Chen, T.; Lin, M.; Kokaras, A.; Dewhirst, F.E.; Lemon, K.P. Construction of habitat-specific training sets to achieve species-level assignment in 16S rRNA gene datasets. *Microbiome* 2020, *8*, 65.

116. Rohwer, R.R.; Hamilton, J.J.; Newton, R.J.; McMahon, K.D. TaxAss: leveraging a custom freshwater database achieves fine-scale taxonomic resolution. *mSphere* 2018, *3*, 327–345.

117. Dueholm, M.S.; Andersen, K.S.; McIlroy, S.J.; Kristensen, J.M.; Yashiro, E.; Karst, S.M.; Albertsen, M.; Nielsen, P.H. Generation of comprehensive ecosystem-specific reference databases with species-level resolution by high-throughput full-length 16S rRNA gene sequencing and automated taxonomy assignment (Autotax). *MBio* 2020, *11*, 1–14.

118. Meola, M.; Rifa, E.; Shani, N.; Delbès, C.; Berthoud, H.; Chassard, C. DAIRYdb: a manually curated reference database for improved taxonomy annotation of 16S rRNA gene sequences from dairy products. *BMC Genomics* 2019, *20*, 560.

119. Douglas, G.M.; Maffei, V.J.; Zaneveld, J.R.; Yurgel, S.N.; Brown, J.R.; Taylor, C.M.; Huttenhower, C.; Langille, M.G.I. PICRUSt2 for prediction of metagenome functions. *Nat. Biotechnol.* 2020, *38*, 685–688.

120. Wemheuer, F.; Taylor, J.A.; Daniel, R.; Johnston, E.; Meinicke, P.; Thomas, T.; Wemheuer, B. Tax4Fun2: prediction of habitat-specific functional profiles and functional redundancy based on 16S rRNA gene sequences. *Environ. Microbiome* 2020, *15*, 11.

121. Narayan, N.R.; Weinmaier, T.; Laserna-Mendieta, E.J.; Claesson, M.J.; Shanahan, F.; Dabbagh, K.; Iwai, S.; Desantis, T.Z. Piphillin predicts metagenomic composition and dynamics from DADA2-corrected 16S rDNA sequences. *BMC Genomics* 2020, *21*, 56.

122. Das, M.; Ghosh, T.S.; Jeffery, I.B. IPCO: Inference of pathways from co-variance analysis. *BMC Bioinformatics* 2020, *21*, 62.

123. Kanehisa, M.; Goto, S.; Sato, Y.; Furumichi, M.; Tanabe, M. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.* 2012, *40*, D109–D114.

124. Thomas, T.; Gilbert, J.; Meyer, F. Metagenomics - a guide from sampling to data analysis. *Microb. Inform. Exp.* 2012, *2*, 3.

125. Laudadio, I.; Fulci, V.; Stronati, L.; Carissimi, C. Next-generation metagenomics: methodological challenges and opportunities. *Omi. A J. Integr. Biol.* 2019, *23*, 327–333.

126. Kumar Awasthi, M.; Ravindran, B.; Sarsaiya, S.; Chen, H.; Wainaina, S.; Singh, E.; Liu, T.; Kumar, S.; Pandey, A.; Singh, L.; et al. Metagenomics for taxonomy profiling: tools and approaches. *Bioengineered* 2020, *11*, 356–374.

127. Kirubakaran, R.; ArulJothi, K.N.; Revathi, S.; Shameem, N.; Parray, J.A. Emerging priorities for microbial metagenome research. *Bioresour. Technol. Reports* 2020, *11*, 100485.

128. De Filippis, Francesca, Alessandro Genovese , Pasquale Ferranti, Jack A. Gilbert, D.E. Metatranscriptomics reveals temperature-driven functional changes in microbiome impacting cheese maturation rate. *Sci. Rep.* 2016, *6*, 1–12.

129. Bellassi, P.; Rocchetti, G.; Nocetti, M.; Lucini, L.; Masoero, F.; Morelli, L. A combined metabolomic and metagenomic approach to discriminate raw milk for the production of hard cheese. *Foods* 2021, *10*, 109.

130. Soggiu, A.; Piras, C.; Mortera, S.L.; Alloggio, I.; Urbani, A.; Bonizzi, L.; Roncada, P. Unravelling the effect of clostridia spores and lysozyme on microbiota dynamics in Grana Padano cheese: a metaproteomics approach. *J. Proteomics* 2016, *147*, 21–27.

131. Misra, B.B.; Langefeld, C.; Olivier, M.; Cox, L.A. Integrated omics: Tools, advances and future approaches. *J. Mol. Endocrinol.* 2019, *62*, R21–R45.

132. Subramanian, I.; Verma, S.; Kumar, S.; Jere, A.; Anamika, K. Multi-omics data integration, interpretation,

and its application. *Bioinform. Biol. Insights* 2020, *14*, 117793221989905.

133. Eicher, T.; Kinnebrew, G.; Patt, A.; Spencer, K.; Ying, K.; Ma, Q.; Machiraju, R.; Mathé, E.A. Metabolomics and multi-omics integration: A survey of computational methods and resources. *Metabolites* 2020, *10*, 202.

134. Graw, S.; Chappell, K.; Washam, C.L.; Gies, A.; Bird, J.; Robeson, M.S.; Byrum, S.D. Multi-omics data integration considerations and study design for biological systems and disease. *Mol. Omi.* 2021, 17, 170-185.

135. Chong, J.; Soufan, O.; Li, C.; Caraus, I.; Li, S.; Bourque, G.; Wishart, D.S.; Xia, J. MetaboAnalyst 4.0: Towards more transparent and integrative metabolomics analysis. *Nucleic Acids Res.* 2018, *46*, W486–W494.

136. Rohart, F.; Gautier, B.; Singh, A.; Lê Cao, K.A. mixOmics: an R package for 'omics feature selection and multiple data integration. *PLoS Comput. Biol.* 2017, *13*, e1005752.

137. Singh, A.; Shannon, C.P.; Gautier, B.; Rohart, F.; Vacher, M.; Tebbutt, S.J.; Cao, K.A.L. DIABLO: an integrative approach for identifying key molecular drivers from multi-omics assays. *Bioinformatics* 2019, *35*, 3055–3062.

138. Mallick, H.; Rahnavard, A.; McIver, L.J.; Ma, S.; Zhang, Y.; Tickle, T.L.; Weingart, G.; Ren, B.; Schwager, E.H.; Thompson, K.N.; et al. Multivariable association discovery in population-scale meta-omics. *bioRxiv* 2021.

139. Afshari, R.; Pillidge, C.J.; Dias, D.A.; Osborn, A.M.; Gill, H. Biomarkers associated with cheese quality uncovered by integrative multi-omic analysis. *Food Control* 2021, *123*, 107752.

140. Tóth, A.G.; Csabai, I.; Krikó, E.; Tőzsér, D.; Maróti, G.; Patai, Á. V.; Makrai, L.; Szita, G.; Solymosi, N. Antimicrobial resistance genes in raw milk for human consumption. *Sci. Rep.* 2020, *10*, 7464.

141. Quigley, L.; McCarthy, R.; O'Sullivan, O.; Beresford, T.P.; Fitzgerald, G.F.; Ross, R.P.; Stanton, C.; Cotter, P.D. The microbial content of raw and pasteurized cow milk as determined by molecular approaches. *J. Dairy Sci.* 2013, *96*, 4928–4937.

142. Reichler, S.J.; Trmčić, A.; Martin, N.H.; Boor, K.J.; Wiedmann, M. *Pseudomonas fluorescens* group bacterial strains are responsible for repeat and sporadic postpasteurization contamination and reduced fluid milk shelf life. *J. Dairy Sci.* 2018, *101*, 7780–7800.

143. Gleeson, D.; O'Connell, A.; Jordan, K. Review of potential sources and control of thermoduric bacteria in bulk-tank milk. *Irish J. Agric. Food Res.* 2013, *52*, 217–227.

144. Miller, R.A.; Kent, D.J.; Watterson, M.J.; Boor, K.J.; Martin, N.H.; Wiedmann, M. Spore populations among bulk tank raw milk and dairy powders are significantly different. *J. Dairy Sci.* 2015, *98*, 8492–8504.

145. Coorevits, A.; De Jonghe, V.; Vandroemme, J.; Reekmans, R.; Heyrman, J.; Messens, W.; De Vos, P.; Heyndrickx, M. Comparative analysis of the diversity of aerobic spore-forming bacteria in raw milk from organic and conventional dairy farms. *Syst. Appl. Microbiol.* 2008, *31*, 126–140.

146. Driehuis, F.; Hoolwerf, J.; Rademaker, J.L.W. Concurrence of spores of *Clostridium tyrobutyricum*, *Clostridium beijerinckii* and *Paenibacillus polymyxa* in silage, dairy cow faeces and raw milk. *Int. Dairy J.* 2016, *63*, 70–77.

147. Bintsis, T. Lactic acid bacteria as starter cultures: an update in their metabolism and genetics. *AIMS Microbiol.* 2018, *4*, 665–684.

148. Zheng, J.; Wittouck, S.; Salvetti, E.; Franz, C.M.A.P.; Harris, H.M.B.; Mattarelli, P.; O'Toole, P.W.; Pot, B.; Vandamme, P.; Walter, J.; et al. A taxonomic note on the genus *Lactobacillus*: description of 23 novel genera, emended description of the genus *Lactobacillus Beijerinck* 1901, and union of *Lactobacillaceae* and *Leuconostocaceae*. *Int. J. Syst. Evol. Microbiol.* 2020, *70*, 2782–2858.

149. Wilkinson, M.G.; LaPointe, G. Invited review: starter lactic acid bacteria survival in cheese: new perspectives on cheese microbiology. *J. Dairy Sci.* 2020, *103*, 10963–10985.

150. Banks, J.M.; Williams, A.G. The role of the nonstarter lactic acid bacteria in Cheddar cheese ripening. Int. J. Dairy Technol.; 2004; 57, 145–152.

151. Cameron, M.; Saab, M.; Heider, L.; McClure, J.T.; Rodriguez-Lecompte, J.C.; Sanchez, J. Antimicrobial susceptibility patterns of environmental streptococci recovered from bovine milk samples in the Maritime Provinces of Canada. *Front. Vet. Sci.* 2016, *3*, 79.

152. Monticelli, J.; Knezevich, A.; Luzzati, R.; Di Bella, S. Clinical management of non-faecium non-faecalis vancomycin-resistant enterococci infection. Focus on *Enterococcus gallinarum* and *Enterococcus casseliflavus/flavescens*. *J. Infect. Chemother.* 2018, *24*, 237–246.

153. Fox, P.F.; Guinee, T.P.; Cogan, T.M.; McSweeney, P.L.H.; Fox, P.F.; Guinee, T.P.; Cogan, T.M.;

McSweeney, P.L.H. Microbiology of cheese ripening. In *Fundamentals of cheese science*; Fox, P.F., Guinee, T.P., Cogan, T.M., McSweeney, P.L.H., Eds.; Springer US: Boston, MA, MA, 2017; pp. 333–390 ISBN 978-1-4899-7681-9.

154. Kandi, V.; Palange, P.; Vaish, R.; Bhatti, A.B.; Kale, V.; Kandi, M.R.; Bhoomagiri, M.R. Emerging bacterial infection: identification and clinical significance of *Kocuria* species. *Cureus* 2016, *8*, e731.

155. Rodríguez-Alonso, P.; Fernández-Otero, C.; Centeno, J.A.; Garabal, J.I. Antibiotic resistance in lactic acid bacteria and *Micrococcaceae* / *Staphylococcaceae* isolates from artisanal raw milk cheeses, and potential implications on cheese making. *J. Food Sci.* 2009, *74*, M284–M293.

156. Centeno, J.A.; Garabal, J.I.; Docampo, F.; Lorenzo, J.M.; Carballo, J. Recovering traditional raw-milk Tetilla cheese flavour and sensory attributes by using *Kocuria varians* and *Yarrowia lipolytica* adjunct cultures. *Int. J. Food Microbiol.* 2017, *251*, 33–40.

157. Kilcawley, K.N. Cheese flavour. In *Fundamentals of cheese science*; Fox, P.F., Guinee, T.P., Cogan, T.M., McSweeney, P.L.H., Eds.; Springer US: Boston, MA, MA, 2017; pp. 443–474.

158. Oliveira, R.B.A.; Margalho, L.P.; Nascimento, J.S.; Costa, L.E.O.; Portela, J.B.; Cruz, A.G.; Sant'Ana, A.S. Processed cheese contamination by spore-forming bacteria: a review of sources, routes, fate during processing and control. *Trends Food Sci. Technol.* 2016, *57*, 11–19.

159. Boor, K.J.; Wiedmann, M.; Murphy, S.; Alcaine, S. A 100-Year review: microbiology and safety of milk handling. *J. Dairy Sci.* 2017, *100*, 9933–9951.

160. Zigo, F.; Sasáková, N.; Gregová, G.; Výrostková, J.; Ondrasovičová, S. Effects of using an alternative bedding composition on the levels of indicator microorganisms and mammary health in dairy farm conditions. *Agriculture* 2020, *10*, 1–15.

161. Albonico, F.; Barelli, C.; Albanese, D.; Manica, M.; Partel, E.; Rosso, F.; Ripellino, S.; Pindo, M.; Donati, C.; Zecconi, A.; et al. Raw milk and fecal microbiota of commercial alpine dairy cows varies with herd, fat content and diet. *PLoS One* 2020, *15*, e0237262.

162. Johansson, M.; Lundh, Å.; De Vries, R.; Sjaunja, K.S. Composition and enzymatic activity in bulk milk from dairy farms with conventional or robotic milking systems. *J. Dairy Res.* 2017, *84*, 154–158.

163. Palii, A.P.; Mihalchenko, S.A.; Chechui, H.F.; Reshetnichenko, A.P.; Rozum, Y.E.; Bredykhin, V. V; Bogomolov, O. V; Denicenko, S.A.; Mitiashkina, T.Y.; Sychov, A.I. Milking and udder health assessment

in industrial farming. 2020, *10*, 375–381.

164. Palii, A.P.; Nanka, O. V; Kovalchuk, Y.O.; Kovalchuk, A.O.; Kalabska, V.S.; Kholod, I. V; Pobirchenko, O.M.; Umrihina, O.S.; Poliakov, A.M.; Ishchenko, K. V; et al. Microbial contamination of cow's milk and operator hygiene. *Ukr. J. Ecol.* 2020, *10*, 392–397.

165. Schingoethe, D.J. A 100-year review: total mixed ration feeding of dairy cows. *J. Dairy Sci.* 2017, *100*, 10143–10150.

166. Fabiszewska, A.U.; Zielińska, K.J.; Wróbel, B. Trends in designing microbial silage quality by biotechnological methods using lactic acid bacteria inoculants: a minireview. *World J. Microbiol. Biotechnol.* 2019, *35*, 76.

167. McAllister, T.A.; Dunière, L.; Drouin, P.; Xu, S.; Wang, Y.; Munns, K.; Zaheer, R. Silage review: using molecular approaches to define the microbial ecology of silage. *J. Dairy Sci.* 2018, *101*, 4060–4074.

168. Oliveira, A.S.; Weinberg, Z.G.; Ogunade, I.M.; Cervantes, A.A.P.; Arriola, K.G.; Jiang, Y.; Kim, D.; Li, X.; Gonçalves, M.C.M.; Vyas, D.; et al. Meta-analysis of effects of inoculation with homofermentative and facultative heterofermentative lactic acid bacteria on silage fermentation, aerobic stability, and the performance of dairy cows. *J. Dairy Sci.* 2017, *100*, 4587–4603.

169. Duniere, L.; Xu, S.; Long, J.; Elekwachi, C.; Wang, Y.; Turkington, K.; Forster, R.; McAllister, T.A. Bacterial and fungal core microbiomes associated with small grain silages during ensiling and aerobic spoilage. *BMC Microbiol.* 2017, *17*, 50.

170. Dunière, L.; Sindou, J.; Chaucheyras-Durand, F.; Chevallier, I.; Thévenot-Sergentet, D. Silage processing and strategies to prevent persistence of undesirable microorganisms. *Anim. Feed Sci. Technol.* 2013, *182*, 1–15.

171. Nair, J.; Huaxin, N.; Andrada, E.; Yang, H.E.; Chevaux, E.; Drouin, P.; McAllister, T.A.; Wang, Y. Effects of inoculation of corn silage with *Lactobacillus hilgardii* and *Lactobacillus buchneri* on silage quality, aerobic stability, nutrient digestibility, and growth performance of growing beef cattle. *J. Anim. Sci.* 2020, *98*, 1–11.

172. Carvalho, B.F.; Sales, G.F.C.; Schwan, R.F.; Ávila, C.L.S. Criteria for lactic acid bacteria screening to enhance silage quality. *J. Appl. Microbiol.* 2021, *130*, 341–355.

173. Drouin, P.; Tremblay, J.; Chaucheyras-Durand, F. Dynamic succession of microbiota during ensiling of

whole plant corn following inoculation with *Lactobacillus buchneri* and *Lactobacillus hilgardii* alone or in combination. *Microorganisms* 2019, *7*, 595.

174. Drouin, P.; Tremblay, J.; Renaud, J.; Apper, E. Microbiota succession during aerobic stability of maize silage inoculated with *Lentilactobacillus buchneri* NCIMB 40788 and *Lentilactobacillus hilgardii* CNCM-I-4785. *Microbiologyopen* 2020, e1153.

175. Petrova, P.; Ivanov, I.; Tsigoriyna, L.; Valcheva, N.; Vasileva, E.; Parvanova-Mancheva, T.; Arsov, A.; Petrov, K. Traditional Bulgarian dairy products: ethnic foods with health benefits. *Microorganisms* 2021, *9*, 480.

176. Moniente, M.; García-Gonzalo, D.; Ontañón, I.; Pagán, R.; Botello-Morte, L. Histamine accumulation in dairy products: microbial causes, techniques for the detection of histamine-producing microbiota, and potential solutions. *Compr. Rev. Food Sci. Food Saf.* 2021, *20*, 1481–1523.

177. Wilkinson, J.M.; Lee, M.R.F.; Rivero, M.J.; Chamberlain, A.T. Some challenges and opportunities for grazing dairy cows on temperate pastures. *Grass Forage Sci.* 2020, *75*, 1–17.

178. Martin, N.H.; Kent, D.J.; Evanowski, R.L.; Zuber Hrobuchak, T.J.; Wiedmann, M. Bacterial spore levels in bulk tank raw milk are influenced by environmental and cow hygiene factors. *J. Dairy Sci.* 2019, *102*, 9689–9701.

179. Godbout, S.; Adam, S.; Belzile, L.; Brassard, P.; Chauveau, L.; Cote, C.; Duchaine, C.; Dufour, S.; Duquette, K.; Fecteau, G.; et al. *Litière recyclée en production laitière : état des connaissances et balises*; 2019; https://www.agrireseau.net/bovinslaitiers/documents/101036/litiere-recyclee-en-production-laitiere-etat-des-connaissances-et-balises.

180. Hogenboom, J.A.; Pellegrino, L.; Sandrucci, A.; Rosi, V.; D'Incecco, P. Invited review: hygienic quality, composition, and technological performance of raw milk obtained by robotic milking of cows. *J. Dairy Sci.* 2019, *102*, 7640–7654.

181. Cogato, A.; Brščić, M.; Guo, H.; Marinello, F.; Pezzuolo, A. Challenges and tendencies of automatic milking systems (AMS): a 20-years systematic review of literature and patents. *Animals* 2021, *11*, 1–21.

182. Weber, M.; Liedtke, J.; Plattes, S.; Lipski, A. Bacterial community composition of biofilms in milking machines of two dairy farms assessed by a combination of culture-dependent and –independent methods. *PLoS One* 2019, *14*, e0222238.

183. Chamberland, J.; Lessard, M.H.; Doyen, A.; Labrie, S.; Pouliot, Y. Biofouling of ultrafiltration membrane by dairy fluids: Characterization of pioneer colonizer bacteria using a DNA metabarcoding approach. *J. Dairy Sci.* 2017, *100*, 981–990.

184. Fricker, M.; Skånseng, B.; Rudi, K.; Stessl, B.; Ehling-Schulz, M. Shift from farm to dairy tank milk microbiota revealed by a polyphasic approach is independent from geographical origin. *Int. J. Food Microbiol.* 2011, *145*, S24-30.

185. Nikoloudaki, O.; Lemos Junior, W.J.F.; Borruso, L.; Campanaro, S.; De Angelis, M.; Vogel, R.F.; Di Cagno, R.; Gobbetti, M. How multiple farming conditions correlate with the composition of the raw cow's milk lactic microbiome. *Environ. Microbiol.* 2021, *23*, 1702–1716.

186. Nguyen, T.T.; Wu, H.; Nishino, N. An investigation of seasonal variations in the microbiota of milk, feces, bedding, and airborne dust. *Asian-Australas J Anim Sci* 2020, *33*, 1858–1865.

187. Feligini, M.; Panelli, S.; Sacchi, R.; Ghitti, M.; Capelli, E. Tracing the origin of raw milk from farm by using automated ribosomal intergenic spacer analysis (ARISA) fingerprinting of microbiota. *Food Control* 2015, *50*, 51–56.

188. Moineau-Jean, A.; Champagne, C.P.; Roy, D.; Raymond, Y.; LaPointe, G. Effect of Greek-style yoghurt manufacturing processes on starter and probiotic bacteria populations during storage. *Int. Dairy J.* 2019, *93*, 35–44.

189. Hahne, J.; Isele, D.; Berning, J.; Lipski, A. The contribution of fast growing, psychrotrophic microorganisms on biodiversity of refrigerated raw cow's milk with high bacterial counts and their food spoilage potential. *Food Microbiol.* 2019, *79*, 11–19.

190. McEniry, J.; O'Kiely, P.; Clipson, N.J.W.; Forristal, P.D.; Doyle, E.M. Assessing the impact of various ensilage factors on the fermentation of grass silage using conventional culture and bacterial community analysis techniques. *J. Appl. Microbiol.* 2010, *108*, 1584–1593.

191. Blekhman, R.; Tang, K.; Archie, E.A.; Barreiro, L.B.; Johnson, Z.P.; Wilson, M.E.; Kohn, J.; Yuan, M.L.; Gesquiere, L.; Grieneisen, L.E.; et al. Common methods for fecal sample storage in field studies yield consistent signatures of individual identity in microbiome sequencing data. *Sci. Rep.* 2016, *6*, 31519.

192. Song, S.J.; Amir, A.; Metcalf, J.L.; Amato, K.R.; Xu, Z.Z.; Humphrey, G.; Knight, R. Preservation methods differ in fecal microbiome stability, affecting suitability for field studies. *mSystems* 2016, *1*, e00021-16.

193. Vandeputte, D.; Tito, R.Y.; Vanleeuwen, R.; Falony, G.; Raes, J. Practical considerations for large-scale gut microbiome studies. *FEMS Microbiol. Rev.* 2017, *41*, S154–S167.

194. Upadhyay, N.; Goyal, A.; Kumar, A.; Ghai, D.L.; Singh, R. Preservation of milk and milk products for analytical purposes. *Food Rev. Int.* 2014, *30*, 203–224.

195. Martins, M.E.P.; Nicolau, E.S.; Mesquita, A.J.; Neves, R.B.S.; Oliveira, J.P. Bronopol and azidiol chemicals: time and temperature influence in the total bacterial count of raw milk. *Ciência Anim. Bras.* 2009, *10*, 627–633.

196. Amores, J.; de la Fe, C.; Gómez-Martín, Á.; Corrales, J.C.; Contreras, A.; Sánchez, A. Preserved goat milk as a valid sample for the PCR detection of *Mycoplasma agalactiae*. *Small Rumin. Res.* 2011, *99*, 61–64.

197. Lackey, K.A.; Williams, J.E.; Price, W.J.; Carrothers, J.M.; Brooker, S.L.; Shafii, B.; McGuire, M.A.; McGuire, M.K. Comparison of commercially-available preservatives for maintaining the integrity of bacterial DNA in human milk. *J. Microbiol. Methods* 2017, *141*, 73–81.

198. Sharma, M.; An, D.; Liu, T.; Pinnock, T.; Cheng, F.; Voordouw, G. Biocide-mediated corrosion of coiled tubing. *PLoS One* 2017, *12*, e0181934.

199. Pan, Y.; Wu, Y.; Li, X.; Zeng, J.; Lin, X. Continuing impacts of selective inhibition on bacterial and fungal communities in an agricultural soil. *Microb. Ecol.* 2019, 1–9.

200. Zouharova-Hubaakova, M.; Ryšánek, D. Effects of freezing milk samples on the recovery of alimentary pathogens and indicator microorganisms. *Acta Vet. Brno* 2007, *76*, 301–307.

201. Amores, J.; Sanchez, A.; Martin, A.G.; Corrales, J.C.; Contreras, A.; de la Fe, C. Viability of Mycoplasma agalactiae and *Mycoplasma mycoides* subsp. *capri* in goat milk samples stored under different conditions. *Vet. Microbiol.* 2010, *145*, 347–350.

202. Heylen, K.; Ettwig, K.; Hu, Z.; Jetten, M.; Kartal, B. Rapid and simple cryopreservation of anaerobic ammonium-oxidizing bacteria. *Appl. Environ. Microbiol.* 2012, *78*, 3010–3013.

203. Nogva, H.K.; Drømtorp, S.M.; Nissen, H.; Rudi, K. Ethidium monoazide for DNA-based differentiation of viable and dead bacteria by 5'-nuclease PCR. *Biotechniques* 2003, *34*, 804–813.

204. Elizaquível, P.; Aznar, R.; Sánchez, G. Recent developments in the use of viability dyes and quantitative

PCR in the food microbiology field. *J. Appl. Microbiol.* 2014, *116*, 1–13.

205. Quigley, L.; O'Sullivan, O.; Beresford, T.P.; Paul Ross, R.; Fitzgerald, G.F.; Cotter, P.D. A comparison of methods used to extract bacterial DNA from raw milk and raw milk cheese. *J. Appl. Microbiol.* 2012, *113*, 96–105.

206. Schmidt, R.J.; Hu, W.; Mills, J.A.; Kung, L. The development of lactic acid bacteria and *Lactobacillus buchneri* and their effects on the fermentation of alfalfa silage. *J. Dairy Sci.* 2009, *92*, 5005–5010.

207. Stevenson, D.M.; Muck, R.E.; Shinners, K.J.; Weimer, P.J. Use of real time PCR to determine population profiles of individual species of lactic acid bacteria in alfalfa silage and stored corn stover. *Appl. Microbiol. Biotechnol.* 2006, *71*, 329–338.

208. Andorrà, I.; Landi, S.; Mas, A.; Guillamón, J.M.; Esteve-Zarzoso, B. Effect of oenological practices on microbial populations using culture-independent techniques. *Food Microbiol.* 2008, *25*, 849–856.

209. Takahashi, H.; Saito, R.; Miya, S.; Tanaka, Y.; Miyamura, N.; Kuda, T.; Kimura, B. Development of quantitative real-time PCR for detection and enumeration of *Enterobacteriaceae*. *Int. J. Food Microbiol.* 2017, *246*, 92–97.

210. Roosa, S.; Wauven, C. Vander; Billon, G.; Matthijs, S.; Wattiez, R.; Gillan, D.C. The *Pseudomonas* community in metal-contaminated sediments as revealed by quantitative PCR: A link with metal bioavailability. *Res. Microbiol.* 2014, *165*, 647–656.

211. Amann, R.I.; Binder, B.J.; Olson, R.J.; Chisholm, S.W.; Devereux, R.; Stahl, D.A. Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. *Appl. Environ. Microbiol.* 1990, *56*, 1919–1925.

212. Fierer, N.; Jackson, J.A.; Vilgalys, R.; Jackson, R.B. Assessment of soil microbial community structure by use of taxon-specific quantitative PCR assays. *Appl. Environ. Microbiol.* 2005, *71*, 4117–4120.

213. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 2011, *17*, 10.

214. Wright, E.S. DECIPHER: harnessing local sequence context to improve protein multiple sequence alignment. *BMC Bioinformatics* 2015, *16*, 322.

215. Schliep, K.P. phangorn: phylogenetic analysis in R. *Bioinformatics* 2011, *27*, 592–593.

216. Callahan, B.J.; Sankaran, K.; Fukuyama, J.A.; McMurdie, P.J.; Holmes, S.P. Bioconductor workflow for microbiome data analysis: from raw reads to community analyses. *F1000Research* 2016, *5*, 1492.

217. McMurdie, P.J.; Holmes, S. phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 2013, *8*, e61217.

218. Foster, Z.S.L.; Sharpton, T.J.; Grünwald, N.J. Metacoder: An R package for visualization and manipulation of community taxonomic diversity data. *PLOS Comput. Biol.* 2017, *13*, e1005404.

219. Lozupone, C.; Knight, R. UniFrac: a new phylogenetic method for comparing microbial communities. *Appl. Environ. Microbiol.* 2005, *71*, 8228–8235.

220. Kassambara, A. ggpubr: "ggplot2" based publication ready plots. R package version 0.2 Available online: https://cran.r-project.org/web/packages/ggpubr/index.html (accessed on Jun 12, 2019).

221. Segata, N.; Izard, J.; Waldron, L.; Gevers, D.; Miropolsky, L.; Garrett, W.S.; Huttenhower, C. Metagenomic biomarker discovery and explanation. *Genome Biol.* 2011, *12*, R60.

222. Afgan, E.; Baker, D.; Batut, B.; van den Beek, M.; Bouvier, D.; Čech, M.; Chilton, J.; Clements, D.; Coraor, N.; Grüning, B.A.; et al. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res.* 2018, *46*, W537–W544.

223. Wickham, H. *ggplot2*; 1st ed.; Springer-Verlag New York, 2009; ISBN 978-0-387-98141-3.

224. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014, *15*, 550.

225. Gu, Z.; Eils, R.; Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 2016, *32*, 2847–2849.

226. Pinheiro, J.; Bates, D.; DebRoy, S.; Sarkar, D.; R Core Team nlme: linear and nonlinear mixed effects models. R package version 3.1-140 Available online: https://cran.r-project.org/package=nlme (accessed on Jun 12, 2019).

227. DiGiulio, D.B.; Callahan, B.J.; McMurdie, P.J.; Costello, E.K.; Lyell, D.J.; Robaczewska, A.; Sun, C.L.; Goltsman, D.S.A.; Wong, R.J.; Shaw, G.; et al. Temporal and spatial variation of the human microbiota during pregnancy. *Proc. Natl. Acad. Sci.* 2015, *112*, 11060–11065.

228. Russell, L.; Henrik, S.; Paul, B.; Maxime, H. emmeans: Estimated marginal means, aka least-squares

means. *R package version*, 2018, 1, 3.

229. Gamer, M.; Lemon, J.; Singh, I. irr: Various coefficients of interrater reliability and agreement. 2012. https://cran.r-project.org/web/packages/irr/index.html.

230. Bender, J.M.; Li, F.; Adisetiyo, H.; Lee, D.; Zabih, S.; Hung, L.; Wilkinson, T.A.; Pannaraj, P.S.; She, R.C.; Bard, J.D.; et al. Quantification of variation and the impact of biomass in targeted 16S rRNA gene sequencing studies. *Microbiome* 2018, *6*, 155.

231. Koo, T.K.; Li, M.Y. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J. Chiropr. Med.* 2016, *15*, 155–63.

232. Lupien-Meilleur, J.; Roy, D.; Lagacé, L. Viability of probiotic bacteria in a maple sap beverage during refrigerated storage. *LWT - Food Sci. Technol.* 2016, *74*, 160–167.

233. Roy, D.; Savard, P.; Guertin, N.; Martoni, C.J.; Jones, M.L.; Champagne, C.P. Viability of *Lactobacillus reuteri* ncimb 30242 during storage in fruit juice and soy beverage. *J. Microbiol. Biotechnol. Food Sci.* 2016, *05*, 320–325.

234. Walker, A.W.; Martin, J.C.; Scott, P.; Parkhill, J.; Flint, H.J.; Scott, K.P. 16S rRNA gene-based profiling of the human infant gut microbiota is strongly influenced by sample processing and PCR primer choice. *Microbiome* 2015, *3*, 26.

235. Tremblay, J.; Singh, K.; Fern, A.; Kirton, E.S.; He, S.; Woyke, T.; Lee, J.; Chen, F.; Dangl, J.L.; Tringe, S.G. Primer and platform effects on 16S rRNA tag sequencing. *Front. Microbiol.* 2015, *6*, 771.

236. Sierra, D.; Sánchez, A.; Contreras, A.; Luengo, C.; Corrales, J.C.; de la Fe, C.; Guirao, I.; Morales, C.T.; Gonzalo, C. Short communication: Effect of storage and preservation on total bacterial counts determined by automated flow cytometry in bulk tank goat milk. *J. Dairy Sci.* 2009, *92*, 4841–4845.

237. Frey-Klett, P.; Burlinson, P.; Deveau, A.; Barret, M.; Tarkka, M.; Sarniguet, A. Bacterial-fungal interactions: hyphens between agricultural, clinical, environmental, and food microbiologists. *Microbiol. Mol. Biol. Rev.* 2011, *75*, 583–609.

238. Zhang, L.; Boeren, S.; van Hooijdonk, A.C.M.; Vervoort, J.M.; Hettinga, K.A. A proteomic perspective on the changes in milk proteins due to high somatic cell count. *J. Dairy Sci.* 2015, *98*, 5339–5351.

239. Thongboonkerd, V.; Saetun, P. Bacterial overgrowth affects urinary proteome analysis:

Recommendation for centrifugation, temperature, duration, and the use of preservatives during sample collection. *J. Proteome Res.* 2007, *6*, 4173–4181.

240. Bernardes, T.F.; Daniel, J.L.P.; Adesogan, A.T.; McAllister, T.A.; Drouin, P.; Nussio, L.G.; Huhtanen, P.; Tremblay, G.F.; Bélanger, G.; Cai, Y. Silage review: unique challenges of silages made in hot and cold regions. *J. Dairy Sci.* 2018, *101*, 4001–4019.

241. Rastogi, G.; Coaker, G.L.; Leveau, J.H.J. New insights into the structure and function of phyllosphere microbiota through high-throughput molecular approaches. *FEMS Microbiol. Lett.* 2013, *348*, 1–10.

242. Driehuis, F.; Wilkinson, J.M.; Jiang, Y.; Ogunade, I.; Adesogan, A.T. Silage review: animal and human health risks from silage. *J. Dairy Sci.* 2018, *101*, 4093–4110.

243. Borreani, G.; Tabacco, E.; Schmidt, R.J.; Holmes, B.J.; Muck, R.E. Silage review: factors affecting dry matter and quality losses in silages. *J. Dairy Sci.* 2018, *101*, 3952–3979.

244. Ogunade, I.M.; Martinez-Tuppia, C.; Queiroz, O.C.M.; Jiang, Y.; Drouin, P.; Wu, F.; Vyas, D.; Adesogan, A.T. Silage review: mycotoxins in silage: Occurrence, effects, prevention, and mitigation. *J. Dairy Sci.* 2018, *101*, 4034–4059.

245. Ávila, C.L.S.; Carvalho, B.F. Silage fermentation - updates focusing on the performance of micro-organisms. *J. Appl. Microbiol.* 2020, *128*, 966–984.

246. Moore-Colyer, M.; Longland, A.; Harris, P.; Zeef, L.; Crosthwaite, S. Mapping the bacterial ecology on the phyllosphere of dry and post soaked grass hay for horses. *PLoS One* 2020, *15*, e0227151.

247. Daniels, S.; Hepworth, J.; Moore-Colyer, M. The haybiome: characterising the viable bacterial community profile of four different hays for horses following different pre-feeding regimens. *PLoS One* 2020, *15*, e0242373.

248. Kraut-Cohen, J.; Tripathi, V.; Chen, Y.; Gatica, J.; Volchinski, V.; Sela, S.; Weinberg, Z.; Cytryn, E. Temporal and spatial assessment of microbial communities in commercial silages from bunker silos. *Appl. Microbiol. Biotechnol.* 2016, *100*, 6827–6835.

249. Ni, K.; Minh, T.T.; Tu, T.T.M.; Tsuruta, T.; Pang, H.; Nishino, N. Comparative microbiota assessment of wilted Italian ryegrass, whole crop corn, and wilted alfalfa silage using denaturing gradient gel electrophoresis and next-generation sequencing. *Appl. Microbiol. Biotechnol.* 2017, *101*, 1385–1394.

250. Dos Santos, A.D.O.; Junior, G.S.D.; Pereira, M.N.; Schwan, R.F.; Da Silva Ávila, C.L. A survey of whole-plant corn silages from Minas Gerais dairy farms. *Sci. Agric.* 2020, *77*, 2020.

251. Nocker, A.; Richter-heitmann, T.; Montijn, R.; Schuren, F.; Kort, R. Discrimination between live and dead cells in bacterial communities from environmental water samples analyzed by 454 pyrosequencing. *Int. Microbiol.* 2010, *13*, 59–65.

252. Gloor, G.B.; Macklaim, J.M.; Pawlowsky-Glahn, V.; Egozcue, J.J. Microbiome datasets are compositional: and this is not optional. *Front. Microbiol.* 2017, *8*, 2224.

253. Silverman, J.D.; Washburne, A.D.; Mukherjee, S.; David, L.A. A phylogenetic transform enhances analysis of compositional microbiota data. *Elife* 2017, *6*, e21887.

254. Legendre, P.; De Cáceres, M. Beta diversity as the variance of community data: dissimilarity coefficients and partitioning. *Ecol. Lett.* 2013, *16*, 951–963.

255. Alfred, S.; William T., S.; Umer, Z.I. microbiomeSeq: An R package for microbial community analysis in an environmental context Available online: http://userweb.eng.gla.ac.uk/umer.ijaz/projects/microbiomeSeq_Tutorial.html#references (accessed on Jul 9, 2020).

256. Fernandes, A.D.; Reid, J.N.S.; Macklaim, J.M.; McMurrough, T.A.; Edgell, D.R.; Gloor, G.B. Unifying the analysis of high-throughput sequencing datasets: characterizing RNA-seq, 16S rRNA gene sequencing and selective growth experiments by compositional data analysis. *Microbiome* 2014, *2*, 15.

257. Ingham, A.C.; Kielsen, K.; Cilieborg, M.S.; Lund, O.; Holmes, S.; Aarestrup, F.M.; Müller, K.G.; Pamp, S.J. Specific gut microbiome members are associated with distinct immune markers in pediatric allogeneic hematopoietic stem cell transplantation. *Microbiome* 2019, *7*, 131.

258. Jari, O.; F. Guillaume, B.; Michael, F.; Roeland, K.; Pierre, L.; Dan, M.; Peter R., M.; R. B., O.; Gavin L., S.; Peter, S.; et al. vegan: community ecology package. R package version 2.5-6, 2019, 1–298.

259. Zhou, J.; Deng, Y.; Luo, F.; He, Z.; Yanga, Y. Phylogenetic molecular ecological network of soil microbial communities in response to elevated CO_2. *MBio* 2011, *2*, e00122-11.

260. Deng, Y.; Jiang, Y.H.; Yang, Y.; He, Z.; Luo, F.; Zhou, J. Molecular ecological network analyses. *BMC Bioinformatics* 2012, *13*, 113.

261. Luo, F.; Yang, Y.; Zhong, J.; Gao, H.; Khan, L.; Thompson, D.K.; Zhou, J. Constructing gene co-expression networks and predicting functions of unknown genes by random matrix theory. *BMC Bioinformatics* 2007, *8*, 299.

262. Luo, F.; Zhong, J.; Yang, Y.; Scheuermann, R.H.; Zhou, J. Application of random matrix theory to biological networks. *Phys. Lett. Sect. A Gen. At. Solid State Phys.* 2006, *357*, 420–423.

263. Clauset, A.; Newman, M.E.J.; Moore, C. Finding community structure in very large networks. *Phys. Rev. E - Stat. Physics, Plasmas, Fluids, Relat. Interdiscip. Top.* 2004, *70*, 6.

264. Maslov, S.; Sneppen, K. Specificity and stability in topology of protein networks. *Science (80-.).* 2002, *296*, 910–913.

265. Olesen, J.M.; Bascompte, J.; Dupont, Y.L.; Jordano, P. The modularity of pollination networks. *Proc. Natl. Acad. Sci. U. S. A.* 2007, *104*, 19891–19896.

266. Barberán, A.; Bates, S.T.; Casamayor, E.O.; Fierer, N. Using network analysis to explore co-occurrence patterns in soil microbial communities. *ISME J.* 2012, *6*, 343–351.

267. Ward, T.; Larson, J.; Meulemans, J.; Hillmann, B.; Lynch, J.; Sidiropoulos, D.; Spear, J.R.; Caporaso, G.; Blekhman, R.; Knight, R.; et al. BugBase predicts organism-level microbiome phenotypes. *bioRxiv* 2017, 133462.

268. Langille, M.G.I.; Zaneveld, J.; Caporaso, J.G.; McDonald, D.; Knights, D.; Reyes, J.A.; Clemente, J.C.; Burkepile, D.E.; Vega Thurber, R.L.; Knight, R.; et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat. Biotechnol.* 2013, *31*, 814–821.

269. Chen, I.-M.A.; Chu, K.; Palaniappan, K.; Pillay, M.; Ratner, A.; Huang, J.; Huntemann, M.; Varghese, N.; White, J.R.; Seshadri, R.; et al. IMG/M v.5.0: an integrated data management and comparative analysis system for microbial genomes and microbiomes. *Nucleic Acids Res.* 2019, *47,* D666-D677.

270. Snyder, E.E.; Kampanya, N.; Lu, J.; Nordberg, E.K.; Karur, H.R.; Shukla, M.; Soneja, J.; Tian, Y.; Xue, T.; Yoo, H.; et al. PATRIC: the VBI PathoSystems Resource Integration Center. *Nucleic Acids Res.* 2007, *35*, D401–D406.

271. Behrendt, U.; Müller, T.; Seyfarth, W. The influence of extensification in grassland management on the populations of micro-organisms in the phyllosphere of grasses. *Microbiol. Res.* 1997, *152*, 75–85.

272. Weinberg, Z.G.; Muck, R.E. New trends and opportunities in the development and use of inoculants for silage. *FEMS Microbiol. Rev.* 1996, *19*, 53–68.

273. Yang, L.; Yuan, X.; Li, J.; Dong, Z.; Shao, T. Dynamics of microbial community and fermentation quality during ensiling of sterile and nonsterile alfalfa with or without *Lactobacillus plantarum* inoculant. *Bioresour. Technol.* 2019, *275*, 280–287.

274. Ogunade, I.M.; Jiang, Y.; Pech Cervantes, A.A.; Kim, D.H.; Oliveira, A.S.; Vyas, D.; Weinberg, Z.G.; Jeong, K.C.; Adesogan, A.T. Bacterial diversity and composition of alfalfa silage as analyzed by Illumina MiSeq sequencing: effects of *Escherichia coli* O157:H7 and silage additives. *J. Dairy Sci.* 2018, *101*, 2048–2059.

275. Yang, F.; Wang, Y.; Zhao, S.; Wang, Y. *Lactobacillus plantarum* inoculants delay spoilage of high moisture alfalfa silages by regulating bacterial community composition. *Front. Microbiol.* 2020, *11*, 1989.

276. Ranalli, G.; Grazia, L.; Roggeri, A. The influence of hay-packing techniques on the presence of *Saccharopolyspora rectivirgula*. *J. Appl. Microbiol.* 1999, *87*, 359–365.

277. Duchaine, C.; Mériaux, A.; Brochu, G.; Bernard, K.; Cormier, Y. *Saccharopolyspora rectivirgula* from Quebec dairy barns: application of simplified criteria for the identification of an agent responsible for farmer's lung disease. *J. Med. Microbiol.* 1999, *48*, 173–180.

278. Lecours, P.B.; Veillette, M.; Marsolais, D.; Duchaine, C. Characterization of bioaerosols from dairy barns: reconstructing the puzzle of occupational respiratory diseases by using molecular approaches. *Appl. Environ. Microbiol.* 2012, *78*, 3242–3248.

279. Schultz, C. Comparing silo filler disease with farmer's lung disease. *Gen. Intern. Med. Clin. Innov.* 2016, *1*, 76–78.

280. Unaogu, I.C.; Gugnani, H.C.; Lacey, J. Occurrence of thermophilic actinomycetes in natural substrates in Nigeria. *Antonie Van Leeuwenhoek* 1994, *65*, 1–5.

281. Guan, H.; Yan, Y.; Li, X.; Li, X.; Shuai, Y.; Feng, G.; Ran, Q.; Cai, Y.; Li, Y.; Zhang, X. Microbial communities and natural fermentation of corn silages prepared with farm bunker-silo in Southwest China. *Bioresour. Technol.* 2018, *265*, 282–290.

282. Hu, Z.; Chang, J.; Yu, J.; Li, S.; Niu, H. Diversity of bacterial community during ensiling and subsequent exposure to air in whole-plant maize silage. *Asian-Australasian J. Anim. Sci.* 2018, *31*, 1464–1473.

283. Dolci, P.; Tabacco, E.; Cocolin, L.; Borreani, G. Microbial dynamics during aerobic exposure of corn silage stored under oxygen barrier or polyethylene films. *Appl. Environ. Microbiol.* 2011, *77*, 7499–507.

284. Guan, H.; Shuai, Y.; Yan, Y.; Ran, Q.; Wang, X.; Li, D.; Cai, Y.; Zhang, X. Microbial community and fermentation dynamics of corn silage prepared with heat-resistant lactic acid bacteria in a hot environment. *Microorganisms* 2020, *8*, 719.

285. Keshri, J.; Chen, Y.; Pinto, R.; Kroupitski, Y.; Weinberg, Z.G.; Sela, S. Microbiome dynamics during ensiling of corn with and without *Lactobacillus plantarum* inoculant. *Appl. Microbiol. Biotechnol.* 2018, *102*, 4025–4037.

286. Ali, N.; Wang, S.; Zhao, J.; Dong, Z.; Li, J.; Nazar, M.; Shao, T. Microbial diversity and fermentation profile of red clover silage inoculated with reconstituted indigenous and exogenous epiphytic microbiota. *Bioresour. Technol.* 2020, *314*, 123606.

287. Yuan, X.J.; Li, J.F.; Dong, Z.H.; Shao, T. The reconstitution mechanism of napier grass microiota during the ensiling of alfalfa and their contributions to fermentation quality of silage. *Bioresour. Technol.* 2020, *297*, 122391.

288. Wang, X.; Haruta, S.; Wang, P.; Ishii, M.; Igarashi, Y.; Cui, Z. Diversity of a stable enrichment culture which is useful for silage inoculant and its succession in alfalfa silage. *FEMS Microbiol. Ecol.* 2006, *57*, 106–115.

289. Bernardi, A.; Härter, C.J.; Silva, A.W.L.; Reis, R.A.; Rabelo, C.H.S. A meta-analysis examining lactic acid bacteria inoculants for maize silage: effects on fermentation, aerobic stability, nutritive value and livestock production. *Grass Forage Sci.* 2019, *74*, 596–612.

290. Pahlow, G.; Muck, R.E.; Driehuis, F.; Oude Elferink, S.J.W.H. Microbiology of Ensiling. In *Silage Science and Technology*; Buxton, D.R., Muck, R.E., Harrison, J.H., Eds.; American Society of Agronomy, Inc.: Madison, Wisconsin, USA, 2003; Vol. 42, pp. 1–62.

291. da Silva, T.C.; da Silva, L.D.; Santos, E.M.; Oliveira, J.S. Importance of the fermentation to produce high-quality silage. In *Fermentation Processes*; InTech, 2017.

292. Nascimento Agarussi, M.C.; Gomes Pereira, O.; Paula, R.A. de; Silva, V.P. da; Santos Roseira, J.P.; Fonseca e Silva, F. Novel lactic acid bacteria strains as inoculants on alfalfa silage fermentation. *Sci. Rep.* 2019, *9*, 1–9.

293. Li, Y.; Nishino, N. Monitoring the bacterial community of maize silage stored in a bunker silo inoculated with *Enterococcus faecium*, *Lactobacillus plantarum* and *Lactobacillus buchneri*. *J. Appl. Microbiol.* 2011, *110*, 1561–1570.

294. Du Toit, W.J.; Pretorius, I.S.; Lonvaud-Funel, A. The effect of sulphur dioxide and oxygen on the viability and culturability of a strain of *Acetobacter pasteurianus* and a strain of *Brettanomyces bruxellensis* isolated from wine. *J. Appl. Microbiol.* 2005, *98*, 862–871.

295. Spoelstra, S.F.; Courtin, M.G.; Van Beers, J.A.C. Acetic acid bacteria can initiate aerobic deterioration of whole crop maize silage. *J. Agric. Sci.* 1988, *111*, 127–132.

296. Queiroz, O.C.M.; Arriola, K.G.; Daniel, J.L.P.; Adesogan, A.T. Effects of 8 chemical and bacterial additives on the quality of corn silage. *J. Dairy Sci.* 2013, *96*, 5836–5843.

297. Ma, B.; Wang, Y.; Ye, S.; Liu, S.; Stirling, E.; Gilbert, J.A.; Faust, K.; Knight, R.; Jansson, J.K.; Cardona, C.; et al. Earth microbial co-occurrence network reveals interconnection pattern across microbiomes. *Microbiome* 2020, *8*, 82.

298. Faust, K.; Raes, J. Microbial interactions: from networks to models. *Nat. Rev. Microbiol.* 2012, *10*, 538–50.

299. Röttjers, L.; Faust, K. From hairballs to hypotheses–biological insights from microbial networks. *FEMS Microbiol. Rev.* 2018, *42*, 761–780.

300. Faust, K.; Sathirapongsasuti, J.F.; Izard, J.; Segata, N.; Gevers, D.; Raes, J.; Huttenhower, C. Microbial co-occurrence relationships in the human microbiome. *PLoS Comput. Biol.* 2012, *8*, e1002606.

301. Layeghifard, M.; Hwang, D.M.; Guttman, D.S. Disentangling interactions in the microbiome: a network perspective. *Trends Microbiol.* 2017, *25*, 217–228.

302. Knief, C.; Delmotte, N.; Chaffron, S.; Stark, M.; Innerebner, G.; Wassmann, R.; Von Mering, C.; Vorholt, J.A. Metaproteogenomic analysis of microbial communities in the phyllosphere and rhizosphere of rice. *ISME J.* 2012, *6*, 1378–1390.

303. Faust, K.; Lahti, L.; Gonze, D.; de Vos, W.M.; Raes, J. Metagenomics meets time series analysis: unraveling microbial community dynamics. *Curr. Opin. Microbiol.* 2015, *25*, 56–66.

304. Porcellato, D.; Meisal, R.; Bombelli, A.; Narvhus, J.A. A core microbiota dominates a rich microbial

diversity in the bovine udder and may indicate presence of dysbiosis. *Sci. Rep.* 2020, *10*, 1–14.

305. Rault, L.; Lévêque, P.A.; Barbey, S.; Launay, F.; Larroque, H.; Le Loir, Y.; Germon, P.; Guinard-Flament, J.; Even, S. Bovine teat cistern microbiota composition and richness are associated with the immune and microbial responses during transition to once-daily milking. *Front. Microbiol.* 2020, *11*, 3202.

306. Falentin, H.; Rault, L.; Nicolas, A.; Bouchard, D.S.; Lassalas, J.; Lamberton, P.; Aubry, J.M.; Marnet, P.G.; Le Loir, Y.; Even, S. Bovine teat microbiome analysis revealed reduced alpha diversity and significant changes in taxonomic profiles in quarters with a history of mastitis. *Front. Microbiol.* 2016, *7*, 480.

307. Oikonomou, G.; Bicalho, M.L.; Meira, E.; Rossi, R.E.; Foditsch, C.; Machado, V.S.; Gustavo, A.; Teixeira, V.; Santisteban, C.; Schukken, Y.H.; et al. Microbiota of cow's milk; distinguishing healthy, sub-clinically and clinically diseased quarters. *PLoS One* 2014, *9*, e85904.

308. Derakhshani, H.; Plaizier, J.C.; De Buck, J.; Barkema, H.W.; Khafipour, E. Composition of the teat canal and intramammary microbiota of dairy cows subjected to antimicrobial dry cow therapy and internal teat sealant. *J. Dairy Sci.* 2018, *101*, 10191–10205.

309. Monsallier, F.; Verdier-Metz, I.; Agabriel, C.; Martin, B.; Montel, M.C. Variability of microbial teat skin flora in relation to farming practices and individual dairy cow characteristics. *Dairy Sci. Technol.* 2012, *92*, 265–278.

310. Salgado, C.A.; Baglinière, F.; Vanetti, M.C.D. Spoilage potential of a heat-stable lipase produced by *Serratia liquefaciens* isolated from cold raw milk. *Lwt* 2020, *126*, 109289.

311. Hassanat, F.; Gervais, R.; Julien, C.; Massé, D.I.; Lettat, A.; Chouinard, P.Y.; Petit, H. V.; Benchaar, C. Replacing alfalfa silage with corn silage in dairy cow diets: effects on enteric methane production, ruminal fermentation, digestion, N balance, and milk production. *J. Dairy Sci.* 2013, *96*, 4553–4567.

312. Wilkinson, J.M.; Muck, R.E. Ensiling in 2050: some challenges and opportunities. *Grass Forage Sci.* 2019, *74*, 178–187.

313. Friedman, J.; Hastie, T.; Tibshirani, R. Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* 2010, *33*, 1–22.

314. Kassambara, A.; Mundt, F. factoextra: extract and visualize the results of multivariate data analyses. *R package version, 2017, 1,* 337-354.

315. Krzywinski, M.; Schein, J.; Birol, I.; Connors, J.; Gascoyne, R.; Horsman, D.; Jones, S.J.; Marra, M.A. Circos: an information aesthetic for comparative genomics. *Genome Res.* 2009, *19*, 1639–1645.

316. Zucali, M.; Bava, L.; Tamburini, A.; Brasca, M.; Vanoni, L.; Sandrucci, A. Effects of season, milking routine and cow cleanliness on bacterial and somatic cell counts of bulk tank milk. *J. Dairy Res.* 2011, *78*, 436–41.

317. Machado, S.G.; da Silva, F.L.; Bazzolli, D.M.S.; Heyndrickx, M.; Costa, P.M. d. A.; Vanetti, M.C.D. *Pseudomonas* spp. and *Serratia liquefaciens* as predominant spoilers in cold raw milk. *J. Food Sci.* 2015, *80*, M1842–M1849.

318. Gonçalves, J.L.; Tomazi, T.; Barreiro, J.R.; Beuron, D.C.; Arcari, M.A.; Lee, S.H.I.; Martins, C.M. de M.R.; Araújo Junior, J.P.; dos Santos, M.V. Effects of bovine subclinical mastitis caused by *Corynebacterium* spp. on somatic cell count, milk yield and composition by comparing contralateral quarters. *Vet. J.* 2016, *209*, 87–92.

319. Dufour, S.; Dohoo, I.R.; Barkema, H.W.; DesCôteaux, L.; DeVries, T.J.; Reyher, K.K.; Roy, J.P.; Scholl, D.T. Epidemiology of coagulase-negative staphylococci intramammary infection in dairy cattle and the effect of bacteriological culture misclassification. *J. Dairy Sci.* 2012, *95*, 3110–3124.

320. Kuehn, J.S.; Gorden, P.J.; Munro, D.; Rong, R.; Dong, Q.; Plummer, P.J.; Wang, C.; Phillips, G.J. Bacterial community profiling of milk samples as a means to understand culture-negative bovine clinical mastitis. *PLoS One* 2013, *8*, e61959.

321. Hantsis-Zacharov, E.; Halpern, M. Culturable psychrotrophic bacterial communities in raw milk and their proteolytic and lipolytic traits. *Appl. Environ. Microbiol.* 2007, *73*, 7162–7168.

322. Meng, L.; Liu, H.; Dong, L.; Zheng, N.; Xing, M.; Zhang, Y.; Zhao, S.; Wang, J. Identification and proteolytic activity quantification of *Pseudomonas* spp. isolated from different raw milks at storage temperatures. *J. Dairy Sci.* 2018, *101*, 2897–2905.

323. Baur, C.; Krewinkel, M.; Kranz, B.; von Neubeck, M.; Wenning, M.; Scherer, S.; Stoeckel, M.; Hinrichs, J.; Stressler, T.; Fischer, L. Quantification of the proteolytic and lipolytic activity of microorganisms isolated from raw milk. *Int. Dairy J.* 2015, *49*, 23–29.

324. Spring, S.; Jäckel, U.; Wagner, M.; Kämpfer, P. *Ottowia thiooxydans* gen. nov., sp. nov., a novel facultatively anaerobic, N2O-producing bacterium isolated from activated sludge, and transfer of

Aquaspirillum gracile to *Hylemonella gracilis* gen. nov., comb. nov. *Int. J. Syst. Evol. Microbiol.* 2004, *54*, 99–106.

325. Jersie-Christensen, R.R.; Lanigan, L.T.; Lyon, D.; Mackie, M.; Belstrøm, D.; Kelstrup, C.D.; Fotakis, A.K.; Willerslev, E.; Lynnerup, N.; Jensen, L.J.; et al. Quantitative metaproteomics of medieval dental calculus reveals individual oral health status. *Nat. Commun.* 2018, *9*, 1–12.

326. Cauchie, E.; Delhalle, L.; Baré, G.; Tahiri, A.; Taminiau, B.; Korsak, N.; Burteau, S.; Fall, P.A.; Farnir, F.; Daube, G. Modeling the growth and interaction between *Brochothrix thermosphacta*, *Pseudomonas* spp., and *Leuconostoc gelidum* in minced pork samples. *Front. Microbiol.* 2020, *11*, 639.

327. Lee, S.D. Antricoccus suffuscus gen. Nov, sp. Nov, isolated from a natural cave. *Int. J. Syst. Evol. Microbiol.* 2015, *65*, 4410–4416.

328. Lee, S.Y.; Lee, S.Y. Assessment of bacterial contamination of toothbrushes using Illumina MiSeq. *Oral Biol. Res.* 2019, *43*, 180–188.

329. Braem, G.; De Vliegher, S.; Verbist, B.; Heyndrickx, M.; Leroy, F.; De Vuyst, L. Culture-independent exploration of the teat apex microbiota of dairy cows reveals a wide bacterial species diversity. *Vet. Microbiol.* 2012, *157*, 383–390.

330. Gopal, N.; Hill, C.; Ross, P.R.; Beresford, T.P.; Fenelon, M.A.; Cotter, P.D. The prevalence and control of *Bacillus* and related spore-forming bacteria in the dairy industry. *Front. Microbiol.* 2015, *6*, 1418.

331. Driehuis, F.; Hoolwerf, J.; Rademaker, J.L.W. Concurrence of spores of *Clostridium tyrobutyricum*, *Clostridium beijerinckii* and *Paenibacillus polymyxa* in silage, dairy cow faeces and raw milk. *Int. Dairy J.* 2016, *63*, 70–77.

332. Rodríguez, J.M. The origin of human milk bacteria : is there a bacterial entero-mammary pathway during late pregnancy and lactation? *Adv. Nutr.* 2014, *5*, 779–784.

Annexe A – Matériel supplémentaire du chapitre 2

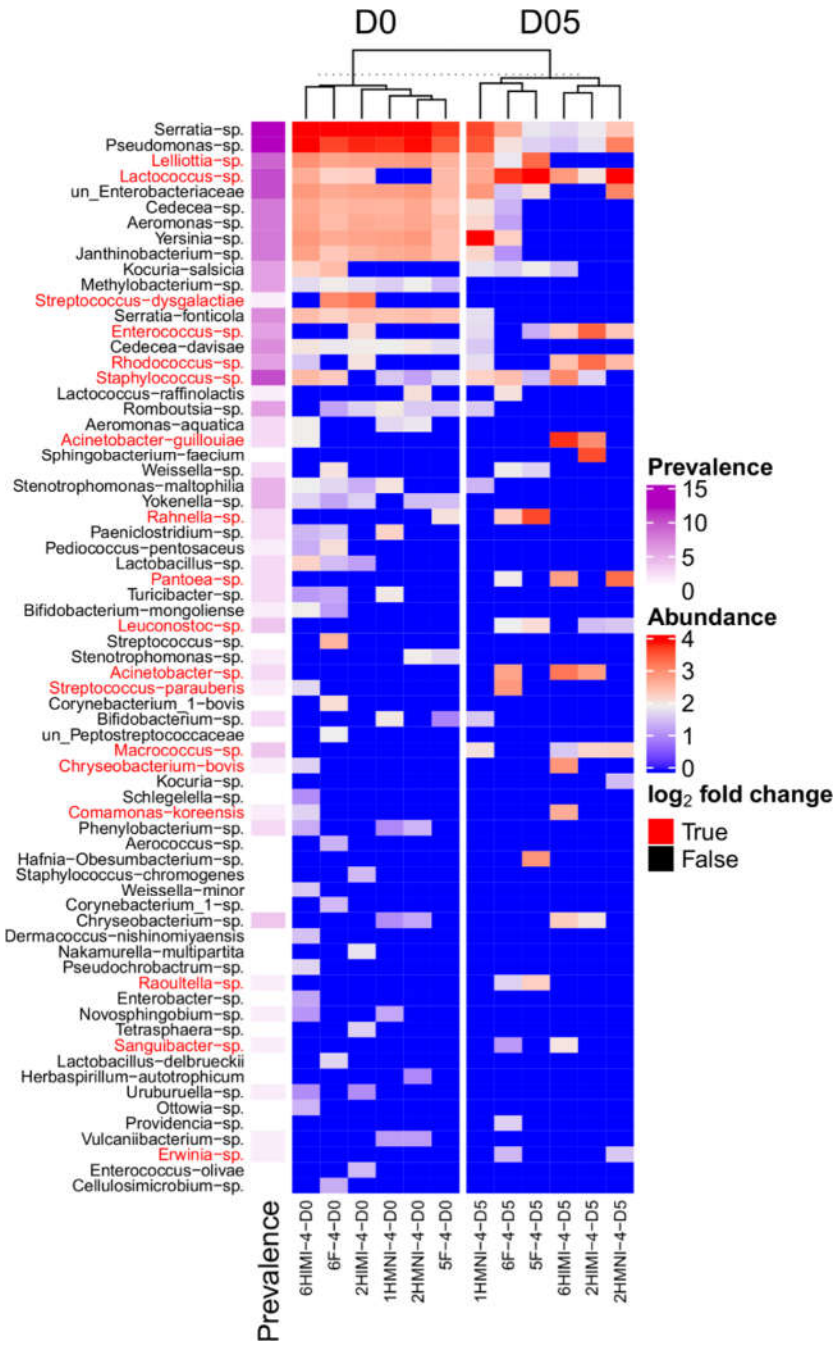


Figure A1: Heat map showing the distribution, prevalence, and abundance of taxa at the species level in fresh and five-day stored unpreserved raw milk based on the V3-V4 dataset. Taxa coloured red are those that underwent log2-fold change during the storage.

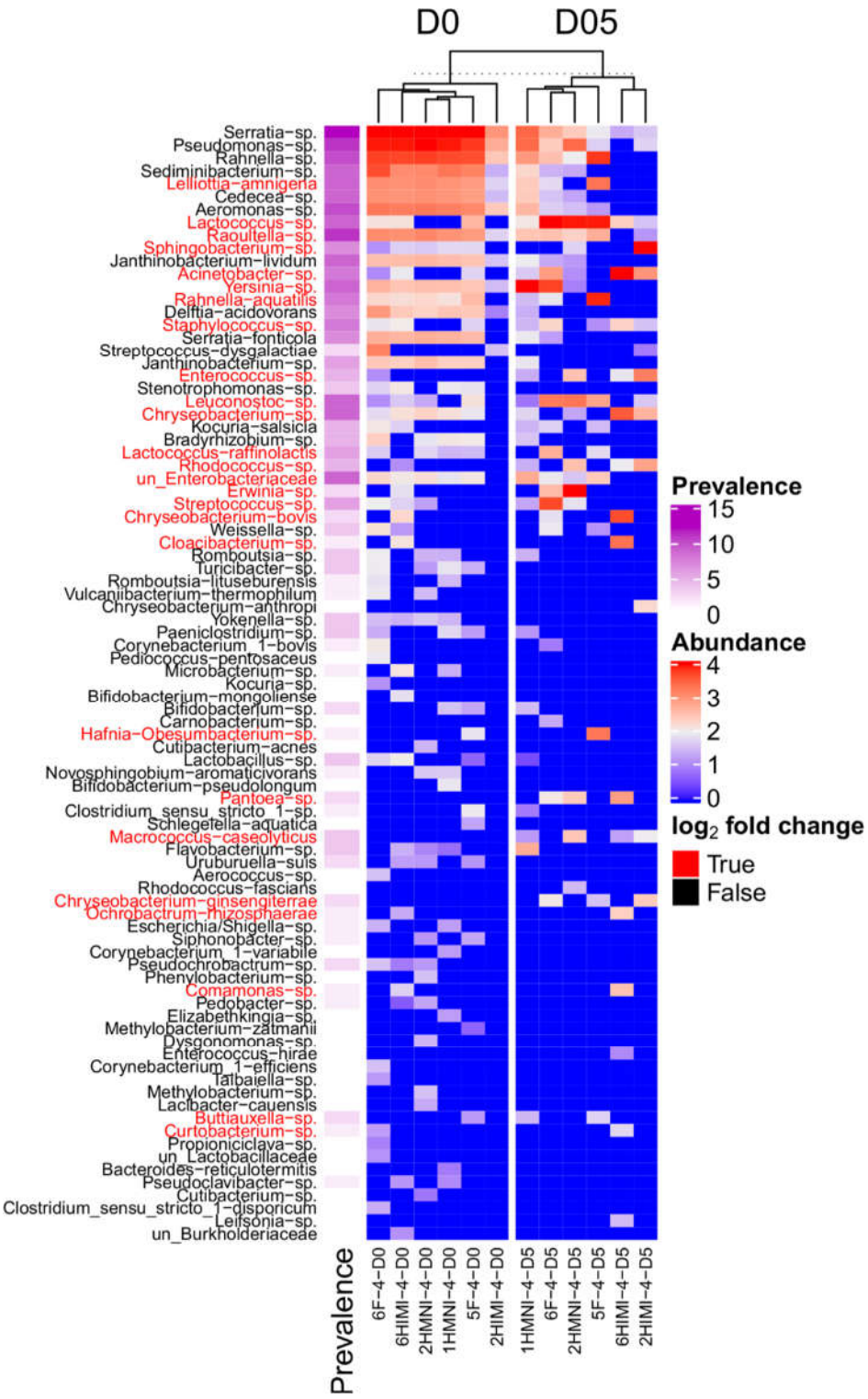


Figure A2: Heat map showing the distribution, prevalence, and abundance of taxa at the species level in fresh and five-day stored unpreserved raw milk based on the V6-V8 dataset. Taxa coloured red are those that underwent log2-fold change during the storage.

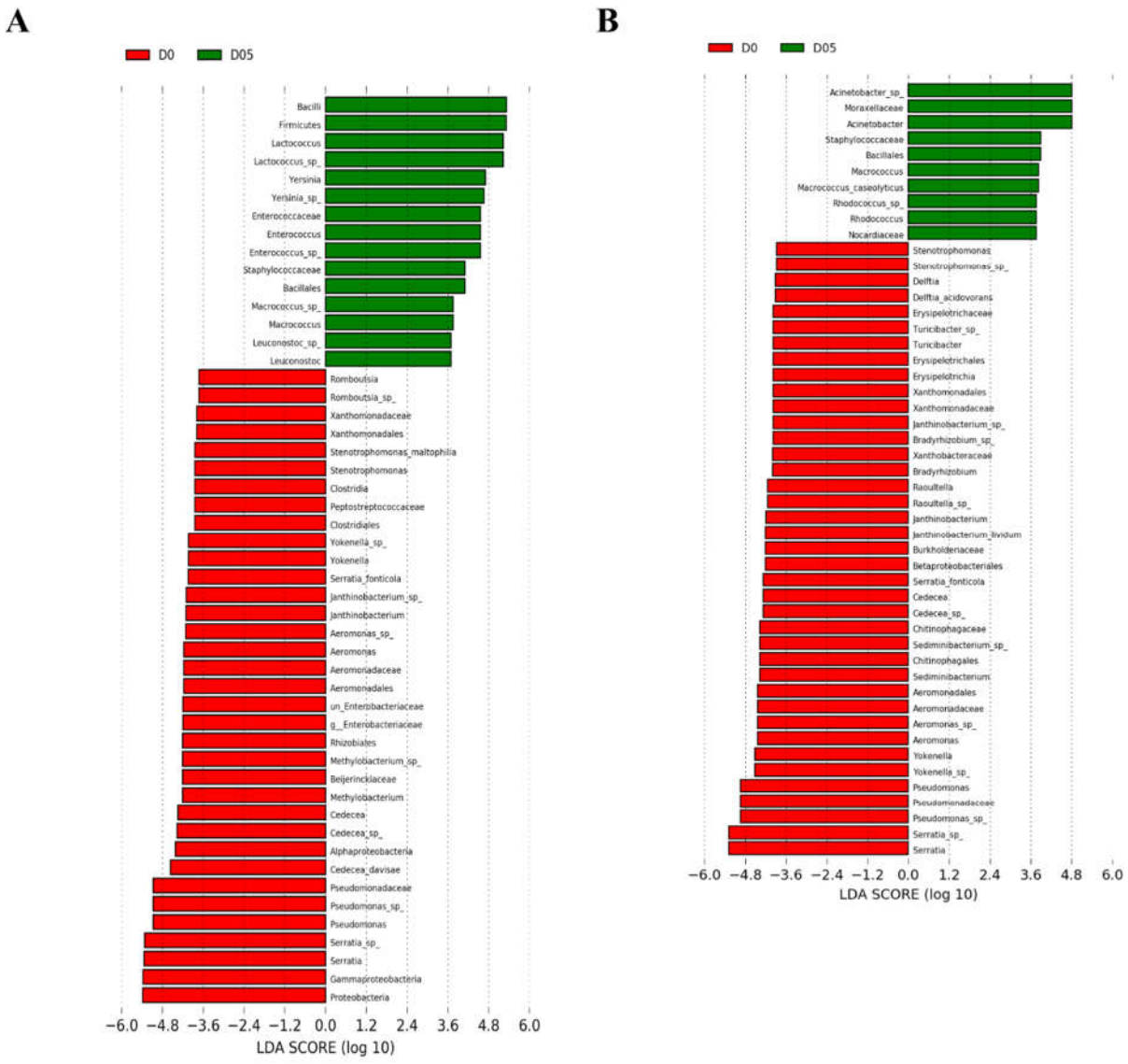


Figure A3: Taxonomic markers characterizing fresh (coloured red) versus five-day stored (coloured green) raw unpreserved milk as depicted by high-throughput sequencing of the 16S rRNA gene targeting the V3-V4 (A) and V6-V8 (B) hypervariable regions.

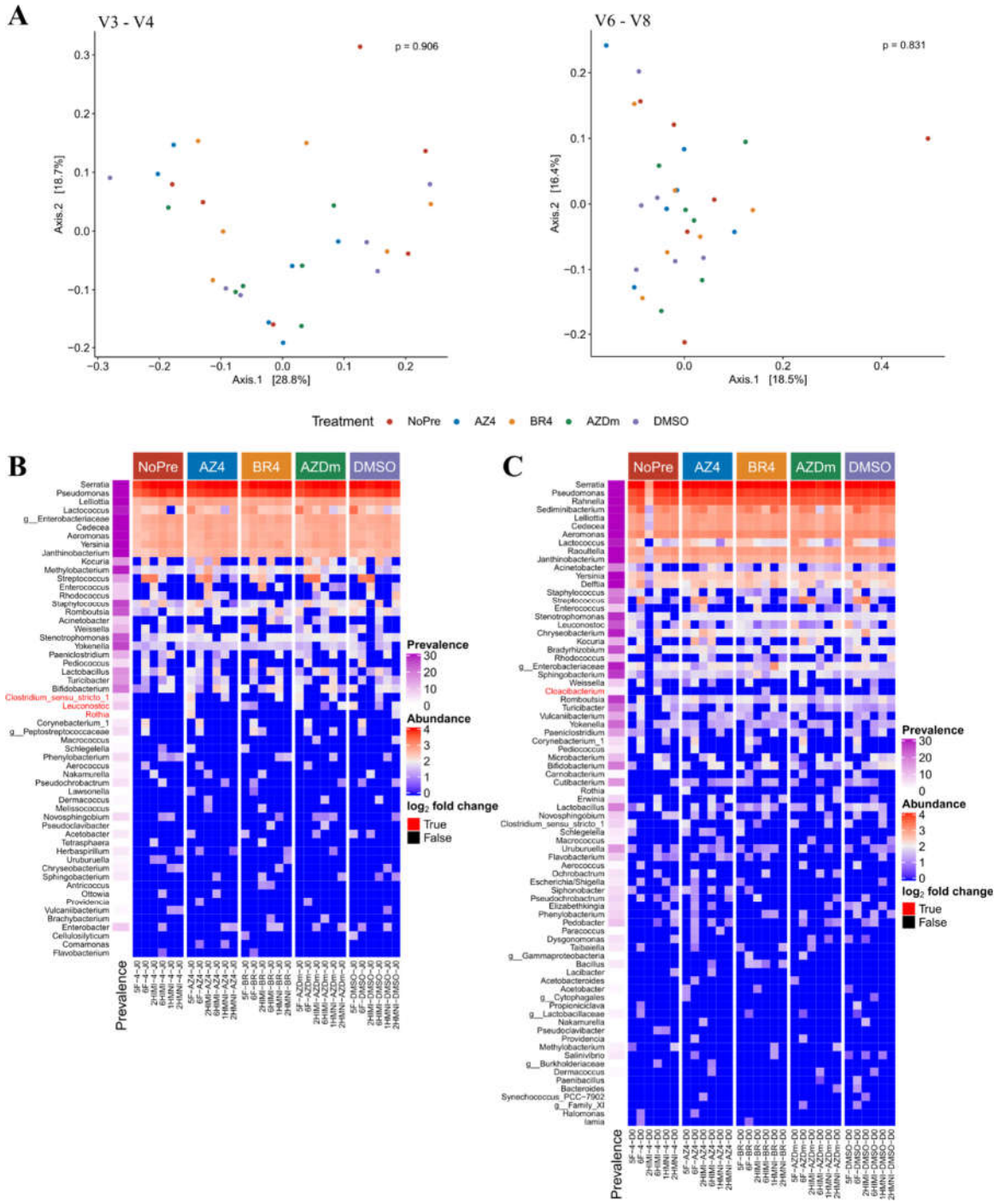


Figure A4: Changes occurring in the microbial composition and structure of raw milk once subjected to preservatives. Treatments include NoPre (untreated raw milk), AZ4 (raw milk treated with azidiol), BR4 (Bronopol-treated raw milk), AZDm (Raw milk treated with a mixture of azidiol and dimethyl sulfoxide), and DMSO (raw milk treated with dimethyl sulfoxide). (A) Principal coordinate analysis on unweighted UniFrac distances based on the V3-V4 (left) and the V6-V8 (right) datasets. (B-C) Prevalence and abundance of taxa (coloured in red) that underwent ≥ log2-fold changes based on the V3-V4 (B) and V6-V8 (C) datasets respectively.

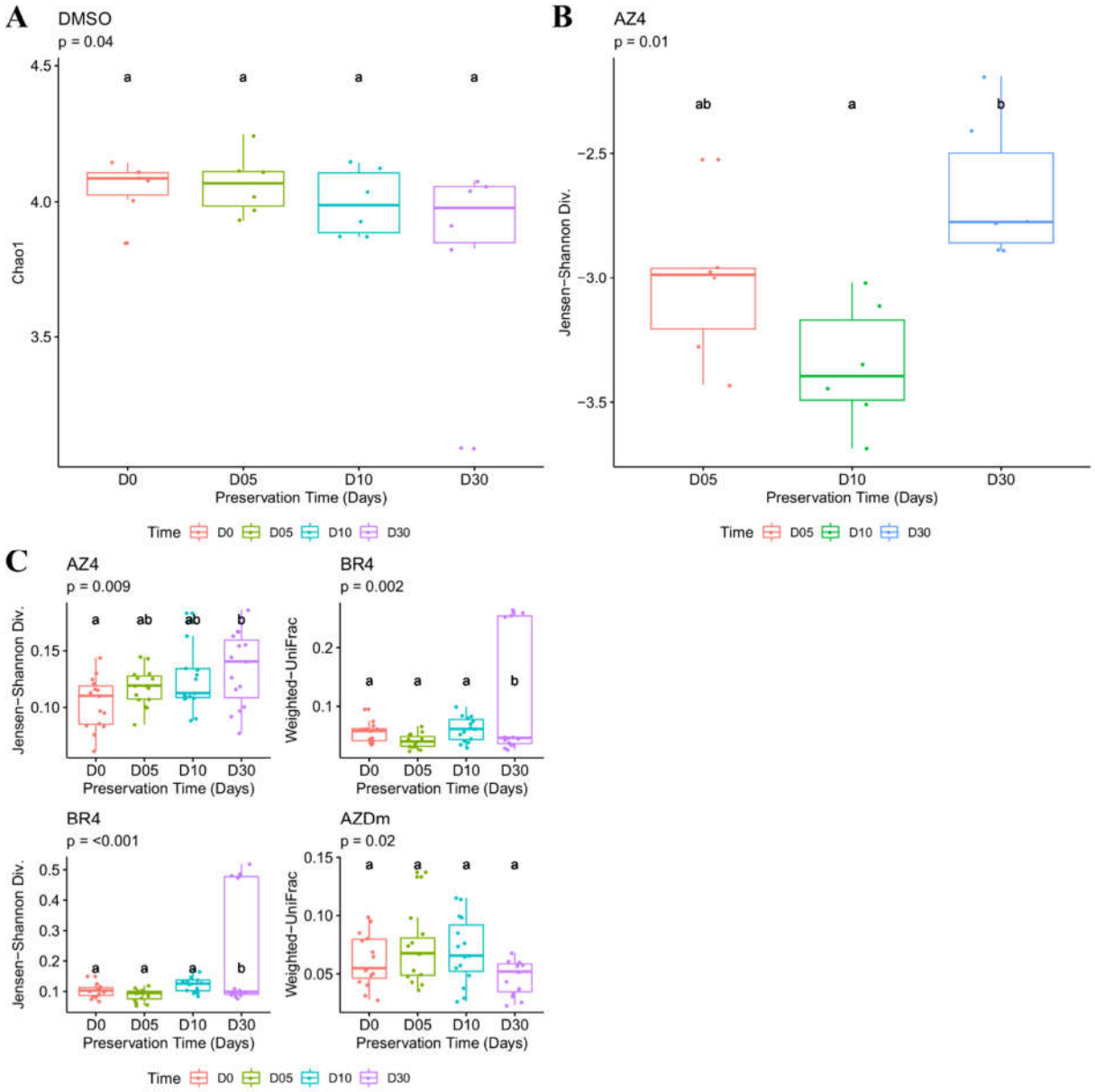


Figure A5: Box plots illustrating multiple comparison tests between storage timepoints of preserved raw milk as described by amplicon sequencing of the V3-V4 variable region. Letters above a box indicate significance group. Timepoints with different letters were significantly different. P-values provided are derived from the corresponding lme analysis shown in Fig. 2-5 of the main text. (A) Multiple comparisons based on the Chao1 estimates analysed by lme modelling and visualized in Fig. 2-5A. (B) Multiple comparisons based on the Jensen-Shannon divergence measures computed within aliquots of the same sample for treatment AZ4 as illustrated in Fig. 2-5B. (C) Multiple comparisons based β-diversity measures computed between aliquots of all samples for treatments AZ4, BR4 and AZDm as illustrated in Fig. 2-5C. NoPre: No treatment, AZ4: Azidiol, BR4: Bronopol, DMSO: Dimethyl sulfoxide, AZDm: Azidiol & DMSO.

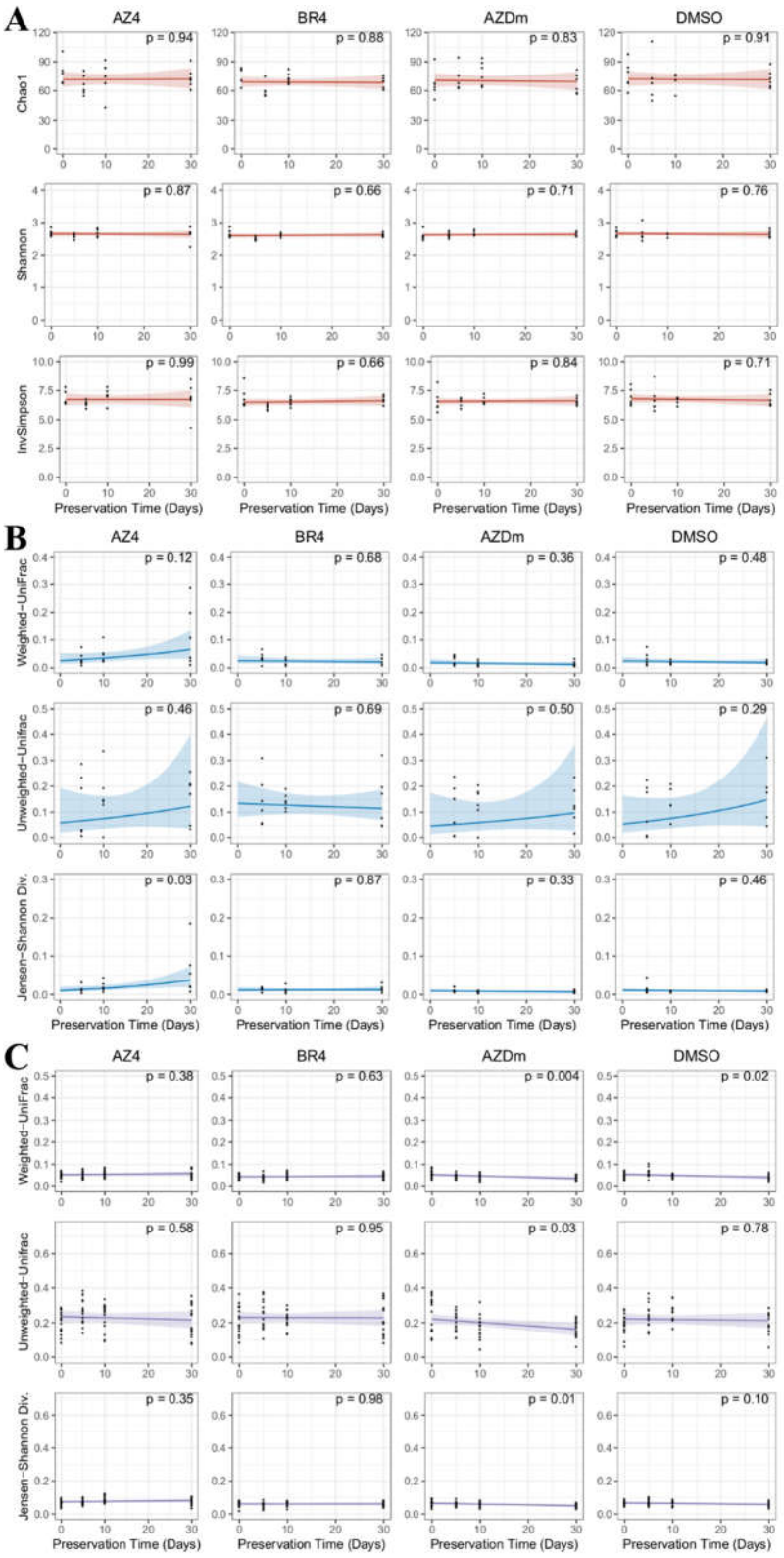


Figure A6: Temporal stability of microbial communities in preserved raw milk across 30 days of storage as described by 16S rRNA gene sequencing of the V6-V8 region. (A) Variation in alpha diversity measures during

the storage of treated raw milk. For Chao1 estimates, Shannon, and InvSimpson indices, salmon lines represent linear mixed-effects fit against the storage time and the 95 % confidence interval is shaded. (B) Diversity trends between samples from the same farm at consecutive timepoints. Blue lines represent linear mixed-effects fit against the storage time and the shade the 95 % confidence interval. (C) Diversity trends between samples from different farms at consecutive timepoints. Purple lines represent linear mixed-effects fit against the storage time and the 95 % confidence interval is shaded. NoPre: No treatment, AZ4: Azidiol, BR4: Bronopol, DMSO: Dimethyl sulfoxide, AZDm: Azidiol & DMSO.

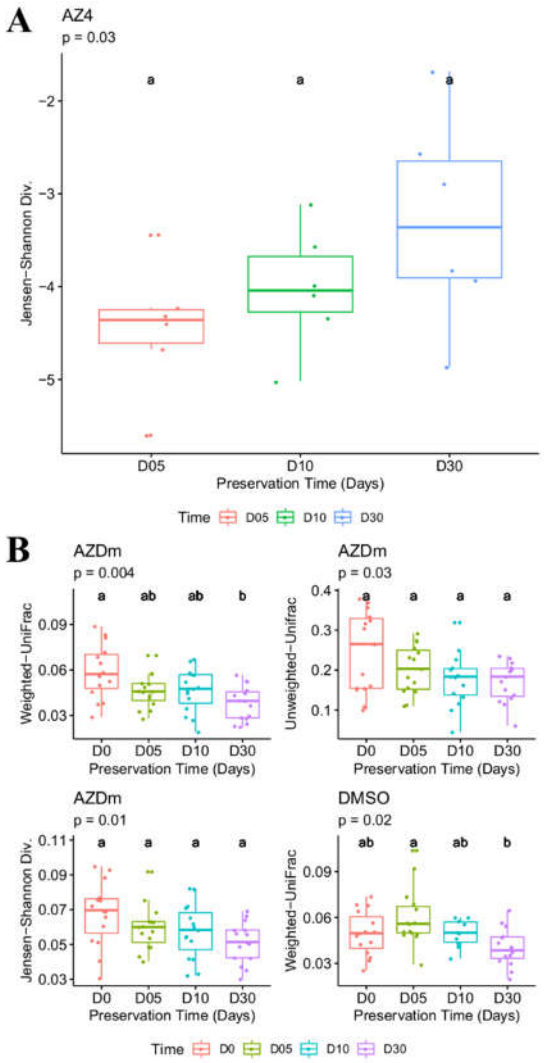


Figure A7: Box plots illustrating multiple comparison tests between storage timepoints of preserved raw milk as described by amplicon sequencing of the V6-V8 variable region. Letters above a box indicate significance group. Timepoints with different letters were significantly different. P-values provided were derived from the corresponding lme analysis shown in Fig. A5. (A) Multiple comparisons based on the Jensen-Shannon divergence measures computed within aliquots of the same sample for treatment AZ4 as illustrated in Fig. A5-A. (B) Multiple comparisons based β-diversity measures computed between aliquots of all samples for treatments AZ4, BR4 and AZDm as illustrated in Fig. A5-C.

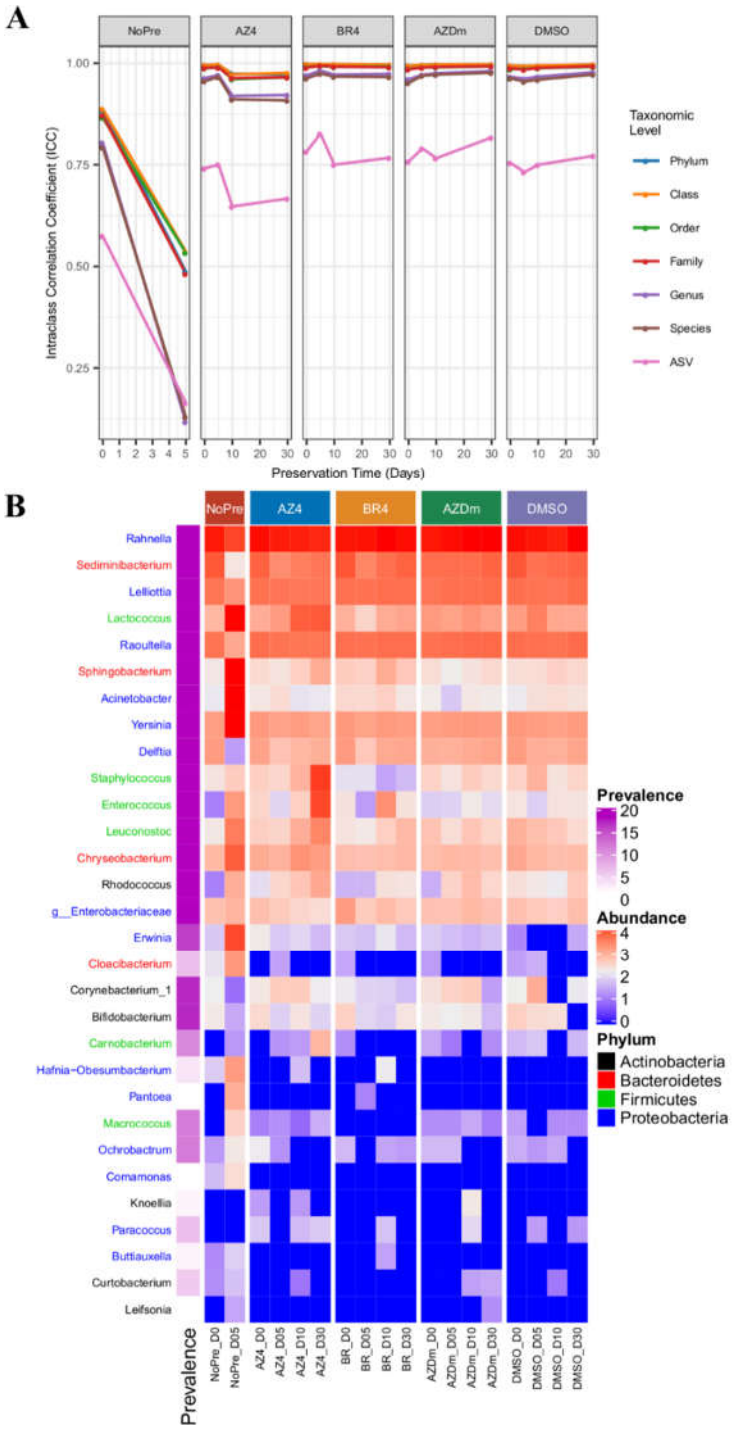


Figure A8: Community instability and taxa dynamics in unpreserved and preserved milk samples over the storage time as depicted by high-throughput amplicon sequencing of the 16S rRNA targeting the V6-V8 region. (A) Intraclass correlation coefficients computed for all the taxonomic levels between samples from different farms and plotted against storage time for preservation-free and preserved samples. (B) Prevalence and abundance of taxa that underwent ≥ log2-fold changes in unpreserved and preserved raw milk during storage. Each taxon at the genus level is coloured by its corresponding phylum.

Annexe B – Matériel supplémentaire du chapitre 3

Table B1: Chemical composition of hay in the Fall 2015 and Spring 2016.

Parameters[1]	Fall 2015 (n=6)	Spring 2016 (n=6)	p value
Moisture (%)	13.4 ± 3.2	12.7 ± 3.1	0.75
Crude protein (%)	11.0 ± 1.6	12.7 ± 2.1	0.15
Soluble protein (%)	8.6 ± 13.6	3.6 ± 0.2	0.15
ESC (%)	8.2 ± 1.5	8.3 ± 2.0	0.42
NFC (%)	23.2 ± 2.3	25.6 ± 3.1	0.13
ADF (%)	38.3 ± 2.4	37.2 ± 3.2	0.42
aNDF (%)	63.0 ± 1.6	57.2 ± 5.2	0.04
TDN (%)	60.6 ± 1.8	60.1 ± 1.9	0.57
Calcium (%)	0.5 ± 0.1	0.7 ± 0.1	0.11
Phosphorus (%)	0.2 ± 0.0	0.3 ± 0.0	0.17
Magnesium (%)	0.2 ± 0.0	0.2 ± 0.0	0.52
Potassium (%)	1.6 ± 0.3	2.2 ± 0.3	0.04

Within a row, each mean is given with the corresponding standard deviation. p values were obtained after performing a Wilcoxon Rank Sum Test across sampling periods. [1] ESC, ethanol soluble carbohydrates; NFC, non-fibre carbohydrates; ADF, acid detergent fiber; aNDF, amylase derived neutral detergent fiber; TDN, total digestible nutrient.

Table B2: Chemical and fermentation characteristics of GL and GLI silage in the Fall 2015 and Spring 2016

Parameters[1]	Fall 2015		p value[3]	Spring 2016		p value[3]
	GL (n=14)[2]	GLI (n=8)[2]		GL (n=15)	GLI (n=8)	
Moisture (%)	55.1 ± 10.6	59.1 ± 5.1	0.41	51.5 ± 10.3	58.1 ± 8.0	0.14
Crude Protein (%)	18.1 ± 2.3	18.5 ± 2.3	0.76	17.3 ± 3.3	19.3 ± 2.7	0.23
ADF-CP (%)	10.5 ± 2.3	9.7 ± 2.8	0.94	10.9 ± 4.3	9.5 ± 3.2	0.36
Soluble Protein CP (%)	55.4 ± 8.5	56.3 ± 4.3	0.91	51.7 ± 8.1	59.5 ± 8.0	<0.05
Ammonia (%)	1.4 ± 0.6	1.5 ± 0.5	0.54	1.2 ± 0.6	1.8 ± 0.5	<0.05
Ammonia CP (%)	7.6 ± 3.2	8.0 ± 2.1	0.35	6.6 ± 2.5	9.3 ± 2.4	<0.05
Ammonia SP (%)	13.6 ± 4.1	14.3 ± 4.1	0.68	12.7 ± 3.9	16.2 ± 5.2	0.10
ADF (%)	34.2 ± 2.6	34.2 ± 4.4	0.71	34.6 ± 3.5	34.6 ± 4.3	0.87
aNDF (%)	47.3 ± 5.3	46.8 ± 5.1	0.73	50.2 ± 7.6	47.2 ± 8.1	0.60
Crude Fat (%)	3.5 ± 0.6	3.6 ± 0.2	0.68	3.4 ± 0.4	3.7 ± 0.7	0.30
Fatty Acid (%)	2.1 ± 0.5	2.1 ± 0.3	0.83	2.1 ± 0.3	2.1 ± 0.4	0.72
Volatile Fatty Acid (%)	7.1 ± 3.5	9.3 ± 1.2	0.39	7.0 ± 3.5	10.9 ± 3.2	<0.05
NFC Carbohydrates (%)	24.9 ± 4.3	24.4 ± 4.7	0.83	23.9 ± 3.8	22.7 ± 4.7	0.38
Carbohydrates (%)	15.8 ± 6.4	12.4 ± 5.9	0.21	18.4 ± 8.0	10.1 ± 5.7	<0.05
pH	4.5 ± 0.4	4.3 ± 0.2	0.18	4.5 ± 0.4	4.4 ± 0.2	0.44
Lactic Acid (%)	4.4 ± 2.1	5.3 ± 0.9	0.17	4.4 ± 2.3	6.1 ± 2.2	0.15
VFA Lactic Acid (%)	55.9 ± 15.8	57.4 ± 8.0	0.83	63.1 ± 20.0	55.6 ± 9.3	0.22
Acetic Acid (%)	3.1 ± 2.2	3.5 ± 1.2	0.29	2.3 ± 1.7	4.4 ± 1.7	<0.05
Butyric Acid (%)	0.3 ± 0.4	0.4 ± 0.3	0.72	0.1 ± 0.2	0.3 ± 0.3	0.24

[1] ADF, acid detergent fiber; aNDF, amylase derived neutral detergent fibre; ADF-CP, acid detergent fiber-protein; Soluble Protein CP, soluble crude protein fraction; Ammonia CP, crude protein associated ammonia; Ammonia SP, soluble protein associated ammonia; NFC carbohydrates, ethanol-soluble carbohydrates as percentage of non-fiber carbohydrates; VFA Lactic Acid, lactic acid as percentage of total volatile fatty acid. [2] No significant difference was obtained across sampling periods for GL and GLI, respectively. [3] p values were obtained after performing a Wilcoxon Rank Sum Test.

Table B3: Chemical and fermentation characteristics of C and CI silage in the Fall 2015 and Spring 2016

Parameters[1]	Fall 2015		p	Spring 2016		p
	C (n=4)[2]	CI (n=8)[2]	value[3]	C (n=4)	CI (n=8)	value[3]
Moisture (%)	68 ± 7.8	62.2 ± 3.6	0.20	64.5 ± 6.1	61.1 ± 2.1	0.20
Crude Protein (%)	7. 5 ± 1.6	7.6 ± 0.7**	0.30	7.4 ± 1.0	6.8 ± 0.7	0.10
ADF-CP (%)	12.2 ± 1.0	10.2 ± 3.9	0.50	12.3 ± 0.5	12.3 ± 1.3	0.73
Soluble Protein CP (%)	46.2 ± 2.9	51.6 ± 5.6	0.05	49.3 ± 5.9	59.9 ± 7.3	<0.05
Ammonia (%)	0.7 ± 0.3	1.0 ± 0.3	0.20	0.8 ± 0.1	1.1 ± 0.3	<0.01
Ammonia CP (%)	9.0 ± 1.6	12.6 ± 3.1*	<0.05	10.3 ± 1.5	16.3 ± 2.7	<0.05
Ammonia SP (%)	19.5 ± 3.1	24.1 ± 4.6	<0.05	21.1 ± 2.6	27.3 ± 4.1	<0.05
ADF (%)	26.3 ± 8	22.5 ± 2.1	0.30	24.9 ± 3.8	24.3 ± 6.6	0.61
aNDF (%)	44.9 ± 11.1	39.1 ± 2.3	0.20	43 ± 5.8	35.3 ± 11.7	0.17
Crude Fat (%)	2.5 ± 0.1	2.8 ± 0.2	<0.05	2.5 ± 0.1	2.8 ± 0.3	0.31
Fatty Acid (%)	2.1 ± 0.5	2.6 ± 0.1	0.06	2.4 ± 0.1	2.7 ± 0.3	0.11
Volatile Fatty Acid (%)	6.4 ± 2.5	4.9 ± 1.4	0.40	5.1 ± 0.6	6.1 ± 2.7	0.31
NFC Carbohydrates (%)	43.4 ± 12.7	48.4 ± 2.5	0.61	45.2 ± 6.4	49.4 ± 4.2	0.31
Carbohydrates (%)	3.8 ± 4.4	1.7 ± 0.8	0.27	2.6 ± 0.9	1.5 ± 1.3	0.06
pH	3.8 ± 0.2	3.7 ± 1.1	0.10	3.9 ± 0.0	3.9 ± 0.2	0.61
Lactic Acid (%)	4.9 ± 1.7	2.6 ± 1.5	0.10	3.4 ± 0.5	4.1 ± 1.9	0.61
VFA Lactic Acid (%)	77.5 ± 5.2*	52.8 ± 24.1	<0.05	66.0 ± 7.7	59.2 ± 15.6	0.31
Acetic Acid (%)	1.3 ± 1.2	2.3 ± 1.3	0.30	1.8 ± 0.5	2.6 ± 0.8	0.09
Butyric Acid (%)	ND	ND	-	ND	ND	-
Propanediol (%)	0.9 ± 1.3	1.8 ± 0.9	0.20	1.6 ± 0.4	2 .0± 0.9	0.2

[1] ADF, acid detergent fiber; aNDF, amylase derived neutral detergent fibre; ADF-CP, acid detergent fiber-protein; Soluble Protein CP, soluble crude protein fraction; Ammonia CP, crude protein associated ammonia; Ammonia SP, soluble protein associated ammonia; NFC carbohydrates, ethanol-soluble carbohydrates as percentage of non-fiber carbohydrates; VFA Lactic Acid, lactic acid as percentage of total volatile fatty acid. [2] Within a row, the superscript following a mean and corresponding standard deviation indicates the significance of the difference with the same category during the next sampling period. [3] p values were obtained after performing a Wilcoxon Rank Sum Test.

Table B4: Keystone phylotypes of the GL pMEN

Keystones	Phylum	Family / Species	Phylotypes[1]	Module
Network hubs	*Firmicutes*	*Weissella* sp.	ASV248	3
	Proteobacteria	*Sphingomonas* sp.	ASV27	2
Connectors	*Actinobacteria*	*Rhodococcus* sp.	ASV183	1
	Firmicutes	*Bacillus* sp.	ASV21	2
		Lentilactobacillus buchneri	ASV113	5
		Loigolactobacillus coryniformis	ASV101	3
		Lactiplantibacillus plantarum	ASV66	2
		Lactobacillus spp.	ASV163	1
			ASV178	2
			ASV93, ASV637	3
			ASV97, ASV24, ASV226	4
			ASV47, ASV118	5
		Lactococcus sp.	ASV440	6
		Pediococcus spp.	ASV152, ASV474, SV221, ASV244	1
			ASV204	2
		Lactobacillaceae	ASV125, ASV324	1
		Weissella spp.	ASV275, ASV262, ASV290, ASV67, ASV186, ASV190, ASV198	2
			ASV600	6
	Proteobacteria	*Aeromonas* sp.	ASV29	4
		Allorhizobium group	ASV111	1
		Methylobacterium adhaesivum	ASV179, ASV128	1
		Methylobacterium sp.	ASV94	2
		Pantoea spp.	ASV56, ASV39	3
		Pseudomonas spp.	ASV14, ASV3	4
		Serratia sp.	ASV9	4
		Stenotrophomonas spp.	ASV162, ASV115	2
		Enterobacteriaceae	ASV17, ASV25, ASV30	4

[1] Phylotypes highlighted in red are GL connectors that changed their topological role to peripherals in the GLI network.

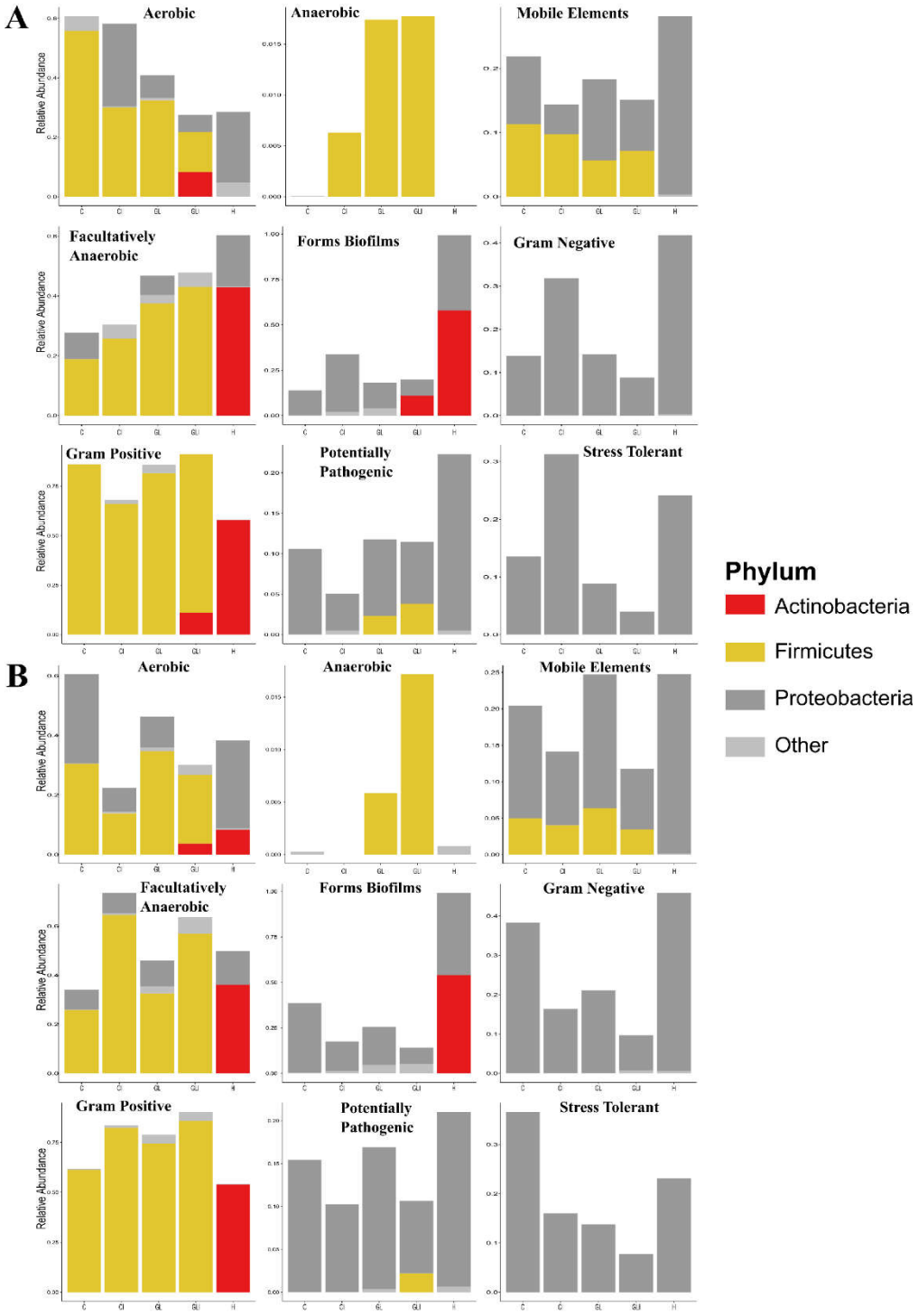


Figure B1: Predicted contribution to the phenotypic traits inferred by BugBase. Relative abundance of the phyla that contributed to microbial phenotypic traits in preserved forage in the fall 2015 (A) and spring 2016 (B).

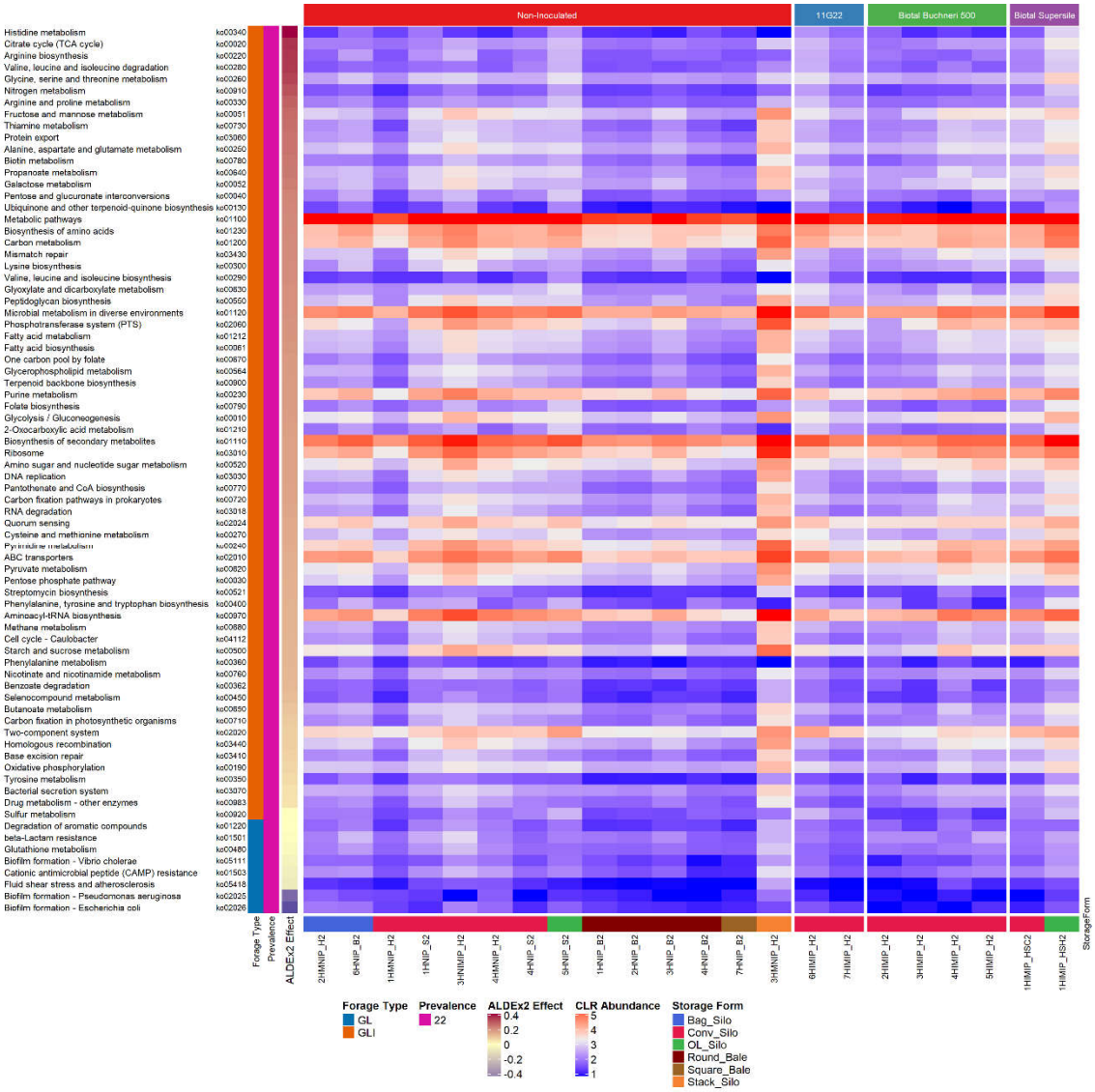


Figure B2: Distribution of differentially abundant function pathways predicted using Piphillin in inoculated and uninoculated grass/legume silage in the fall 2015.

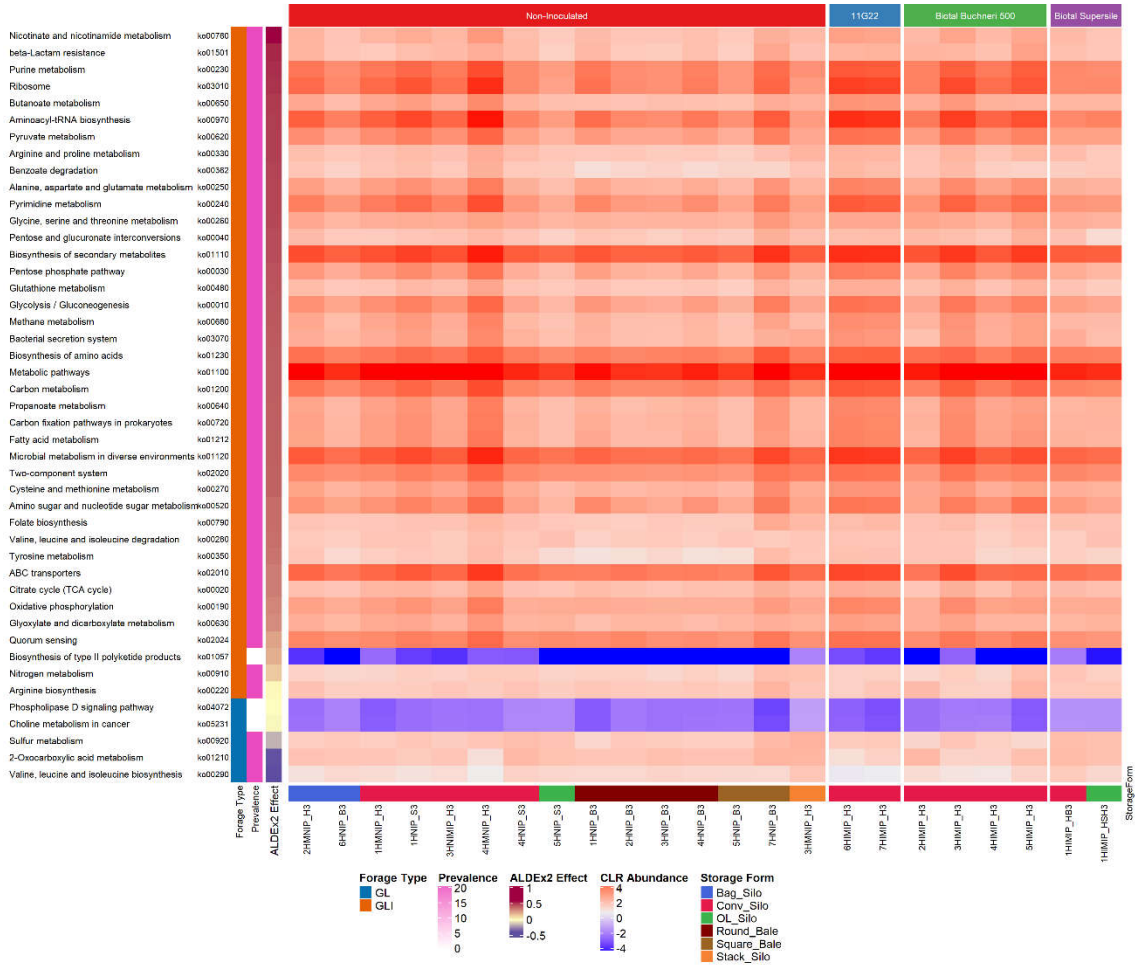


Figure B3: Distribution of differentially abundant function pathways predicted using Piphillin in inoculated and uninoculated grass/legume silage in the spring 2016.

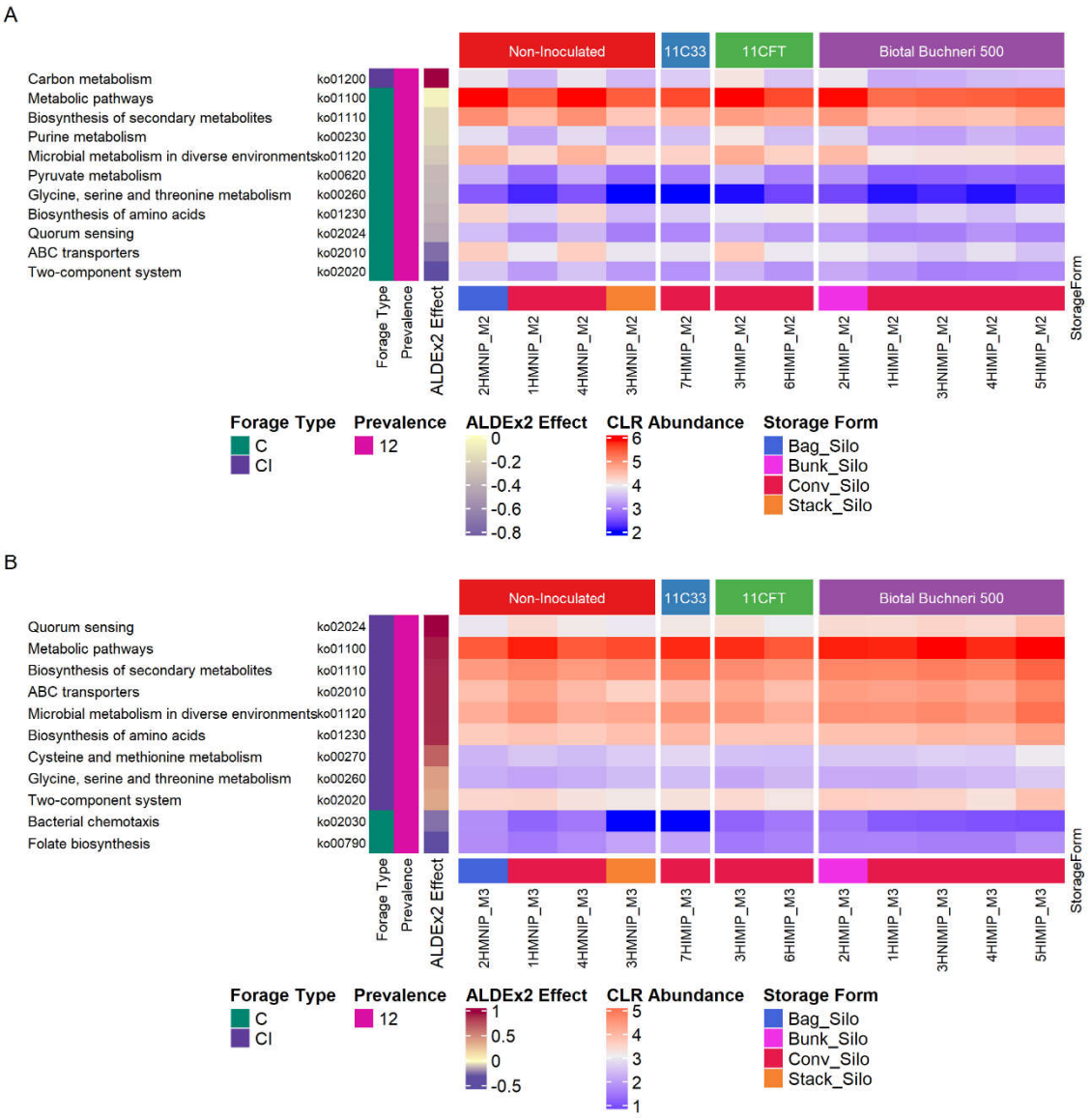


Figure B4: Distribution of differentially abundant function pathways predicted using Piphillin. Inoculated and uninoculated corn silage in the fall 2015 (A) or spring 2016 (B).

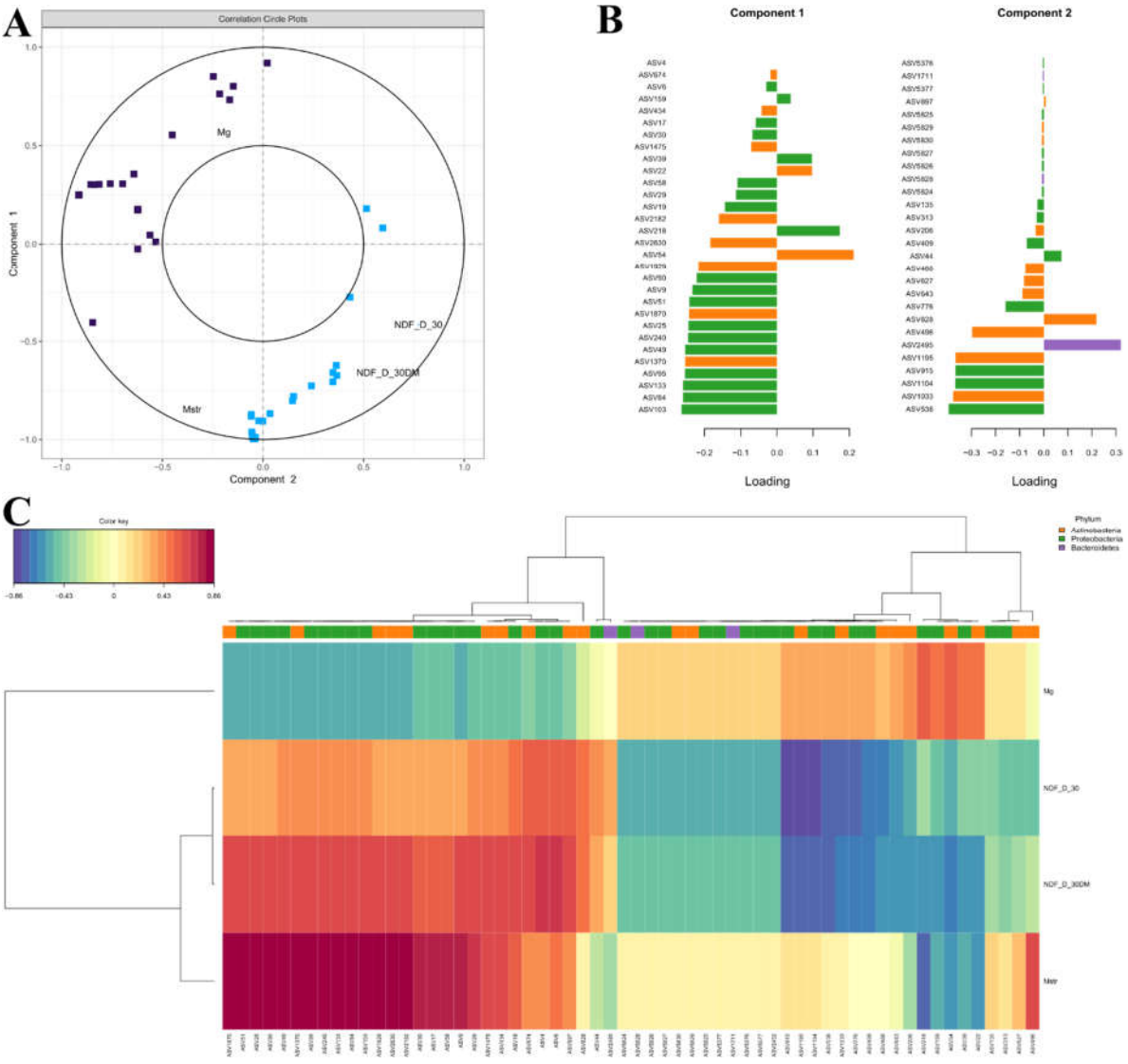


Figure B5: Sparse partial least square regression analysis of hay physicochemical parameters and bacterial ASV. (A) Correlation circle plot based on the first two dimensions. (B) Loading plots of ASV showing their contributions to the first (left) and second (right) components. (C) Clustered image map showing correlations between physicochemical parameters and ASV.

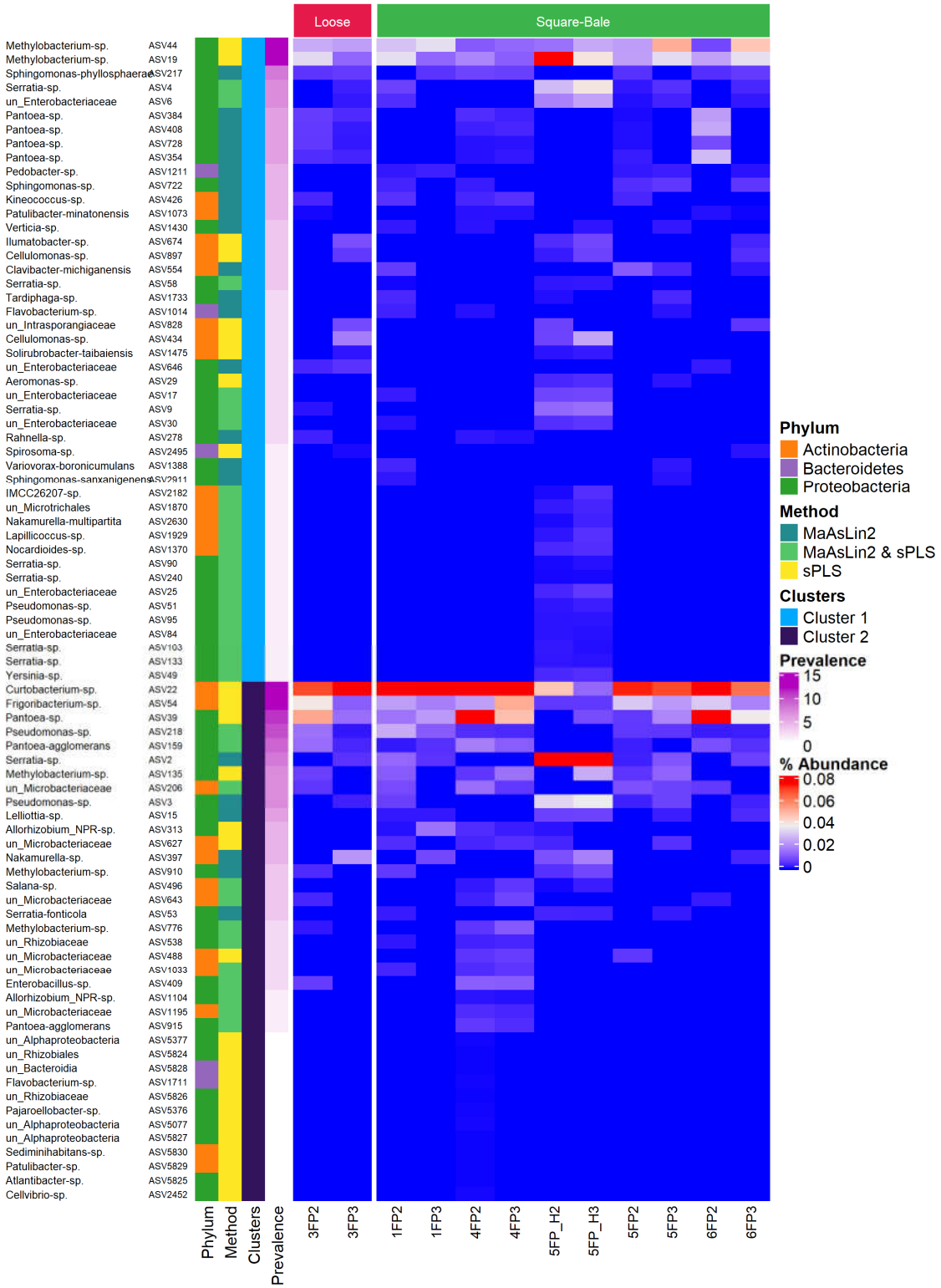


Figure B6: Prevalence, abundance, and distribution across clusters of ASV significantly correlated with hay physicochemical parameters.

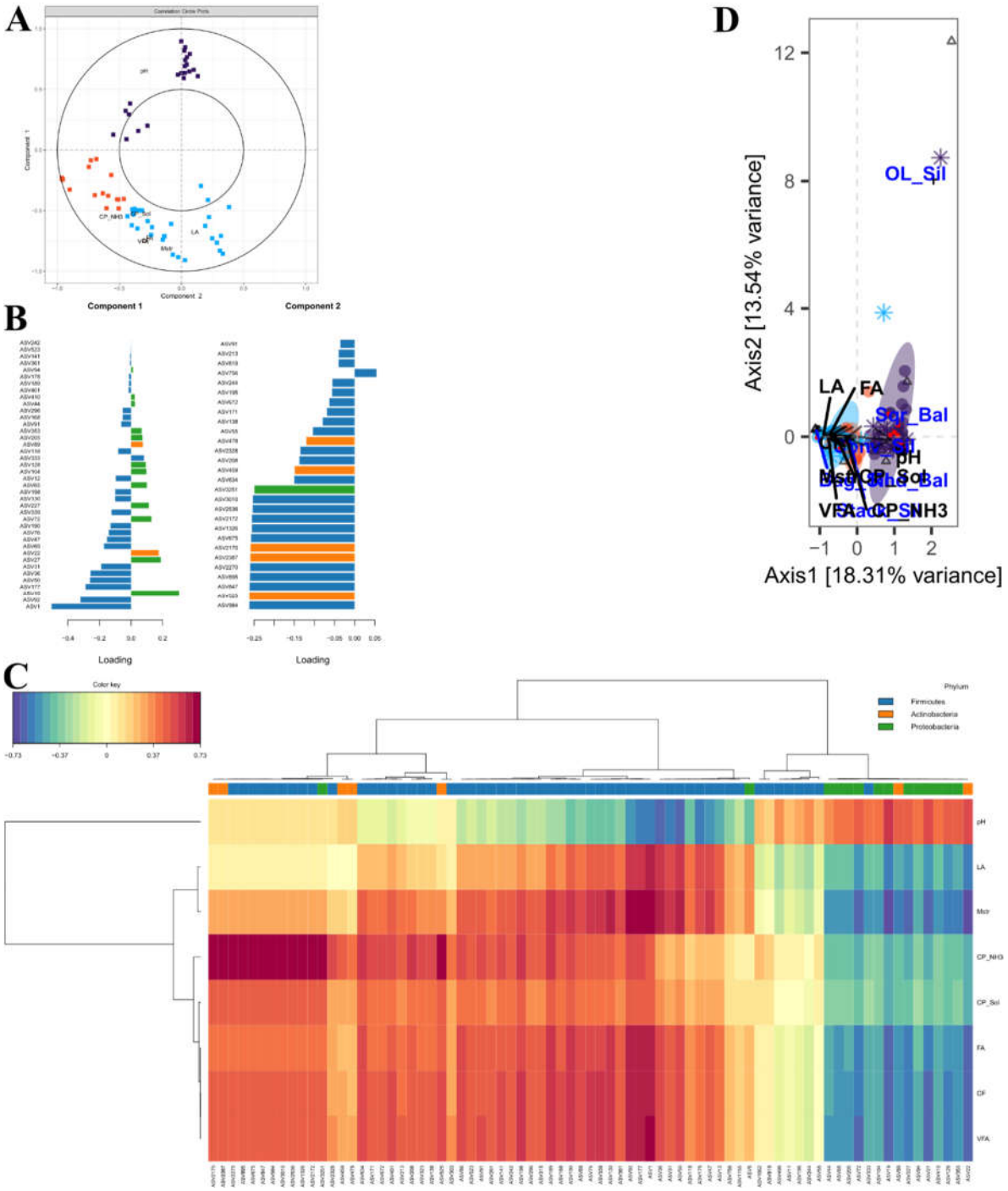


Figure B7: Sparse partial least square regression analysis of physicochemical parameters and bacterial ASV for uninoculated grass/legume silage. (A) Correlation circle plot based on the first two dimensions. (B) Loading plots of ASV showing their contributions to the first (left) and second (right) components. (C) Clustered image map showing correlations between physicochemical parameters and ASV. (D) Illustration of the complete CCpnA plot showing canonical relationships associated with the sub-variable Oxygen limiting silo (OL_Silo). Abbreviations on the triplot are the same as those provided in Fig. 3-7.

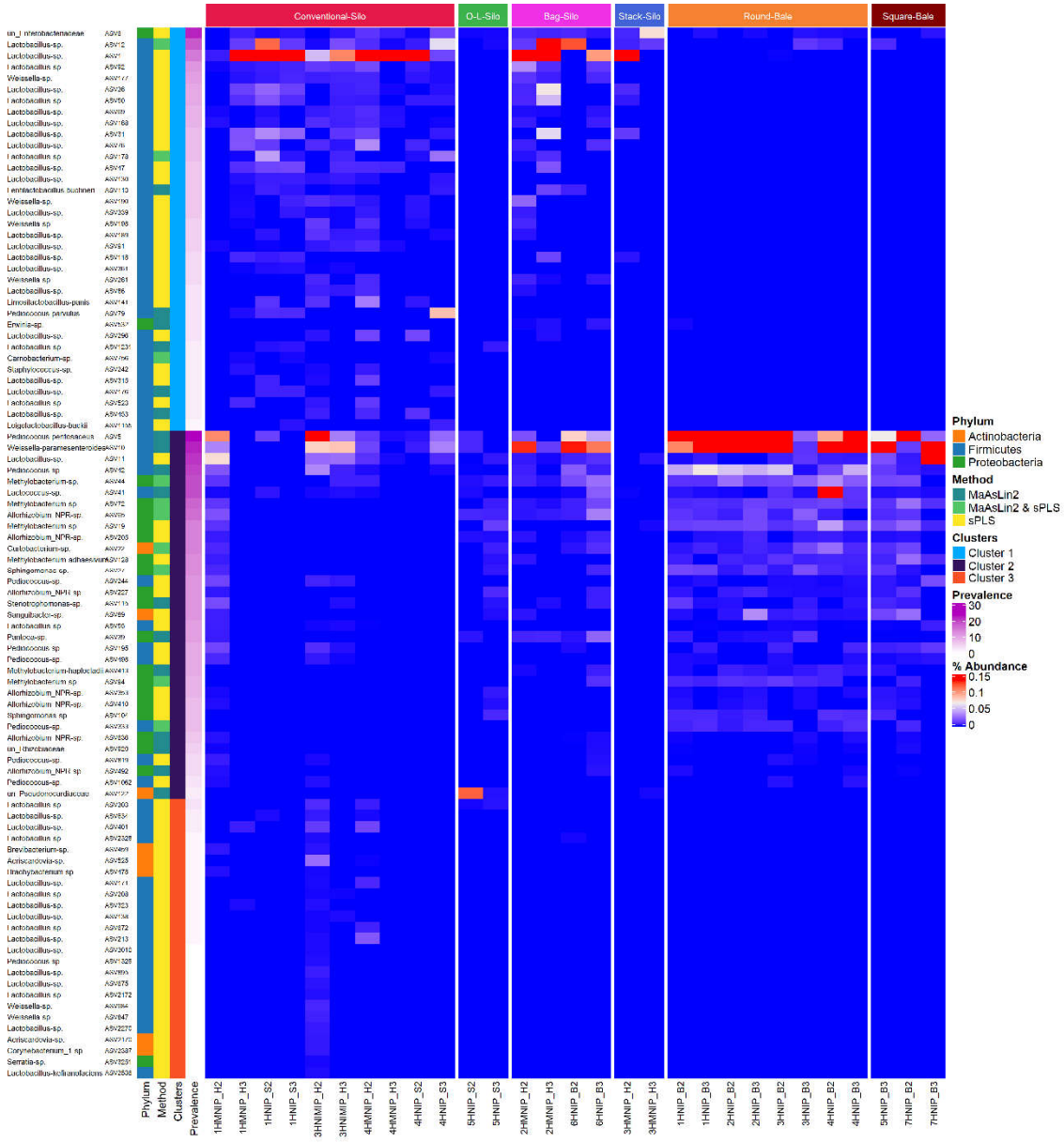


Figure B8: Prevalence, abundance, and distribution across clusters of ASV significantly correlated with physicochemical parameters for uninoculated grass/legume silage.

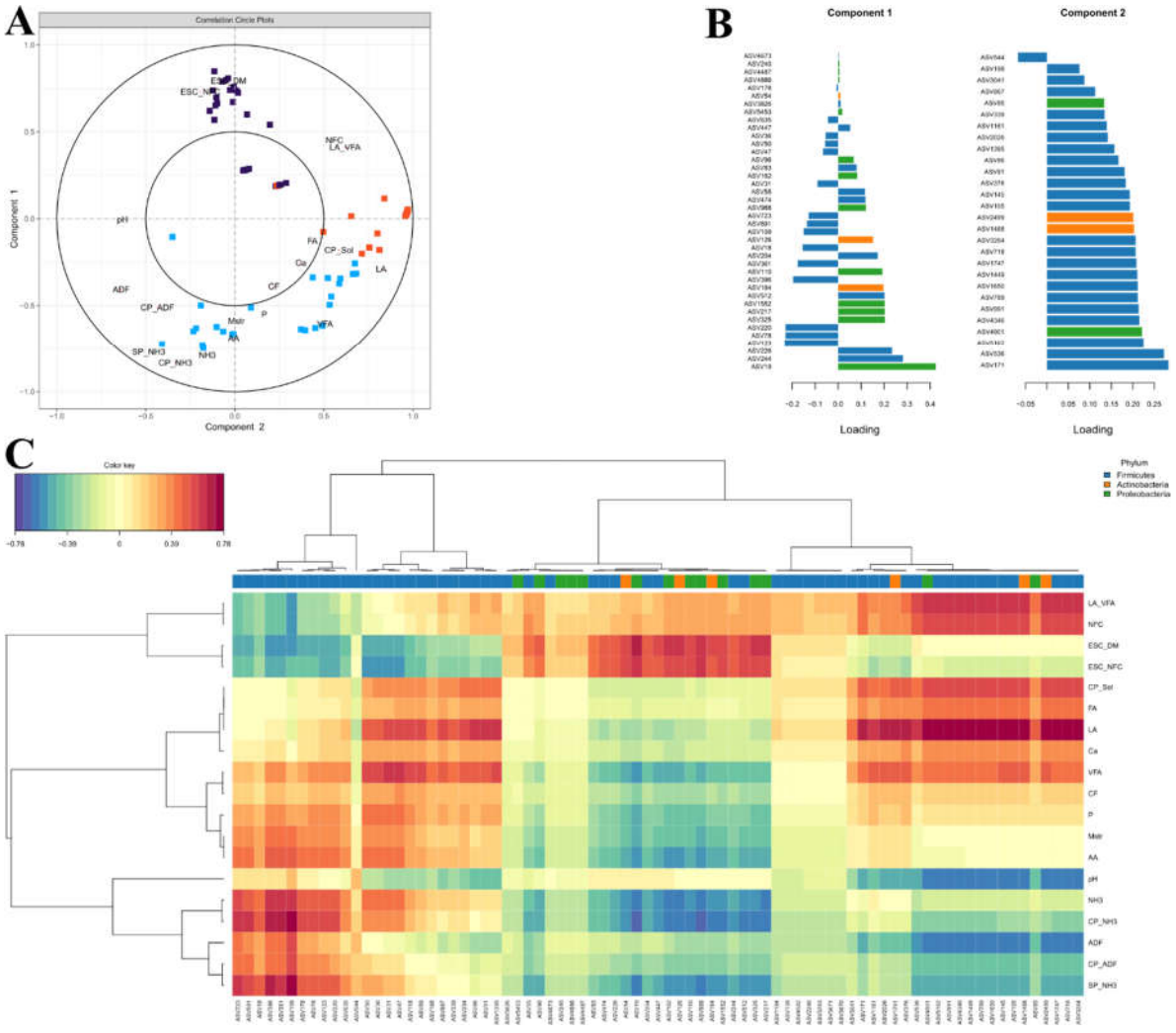


Figure B9: Sparse partial least square regression analysis of physicochemical parameters and bacterial ASV for inoculated grass/legume silage. (A) Correlation circle plot based on the first two dimensions. (B) Loading plots of ASV showing their contributions to the first (left) and second (right) components. (C) Clustered image map showing correlations between physicochemical parameters and ASV.

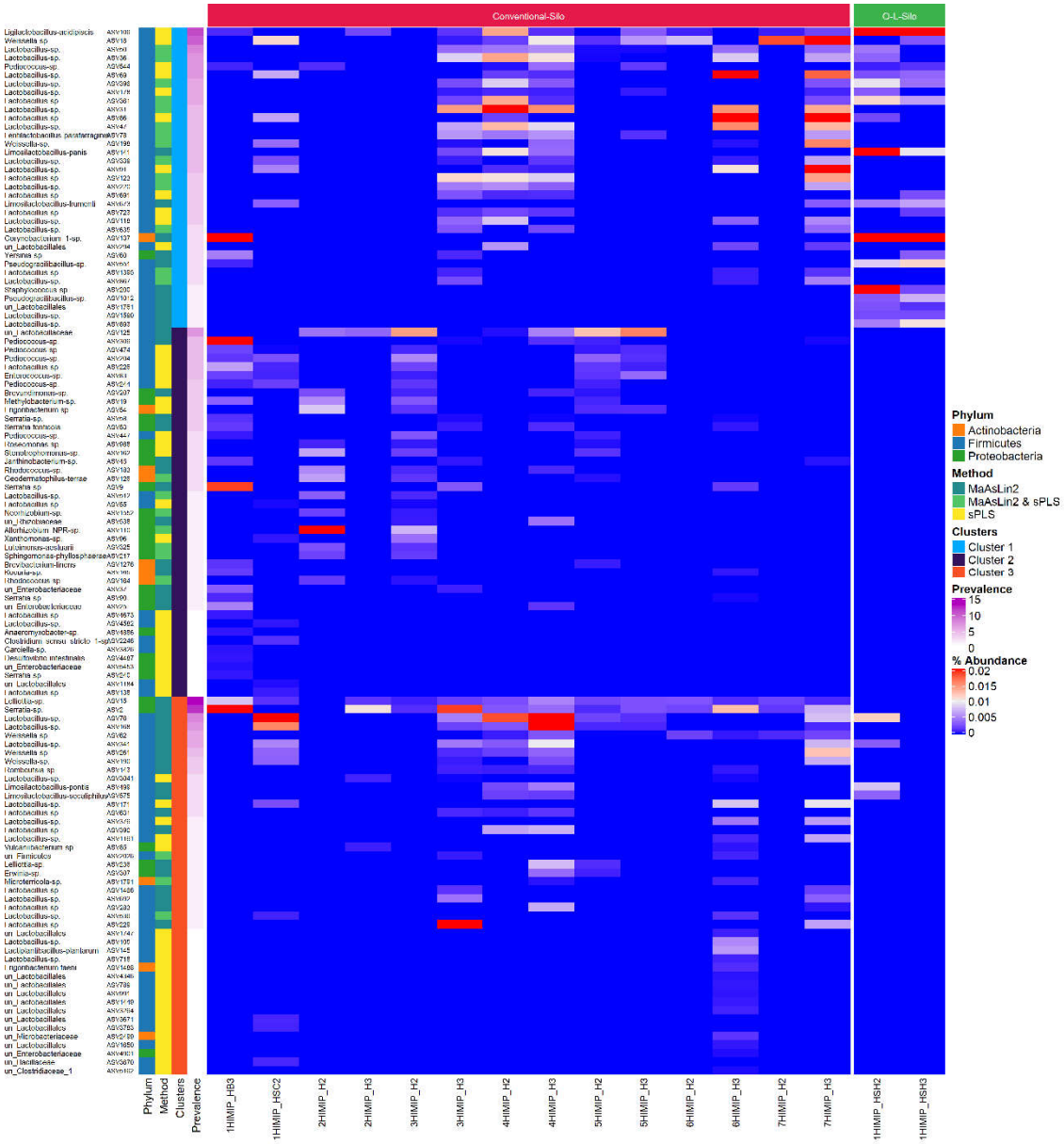


Figure B10: Prevalence, abundance, and distribution across clusters of ASV significantly correlated with physicochemical parameters for inoculated grass/legume silage.

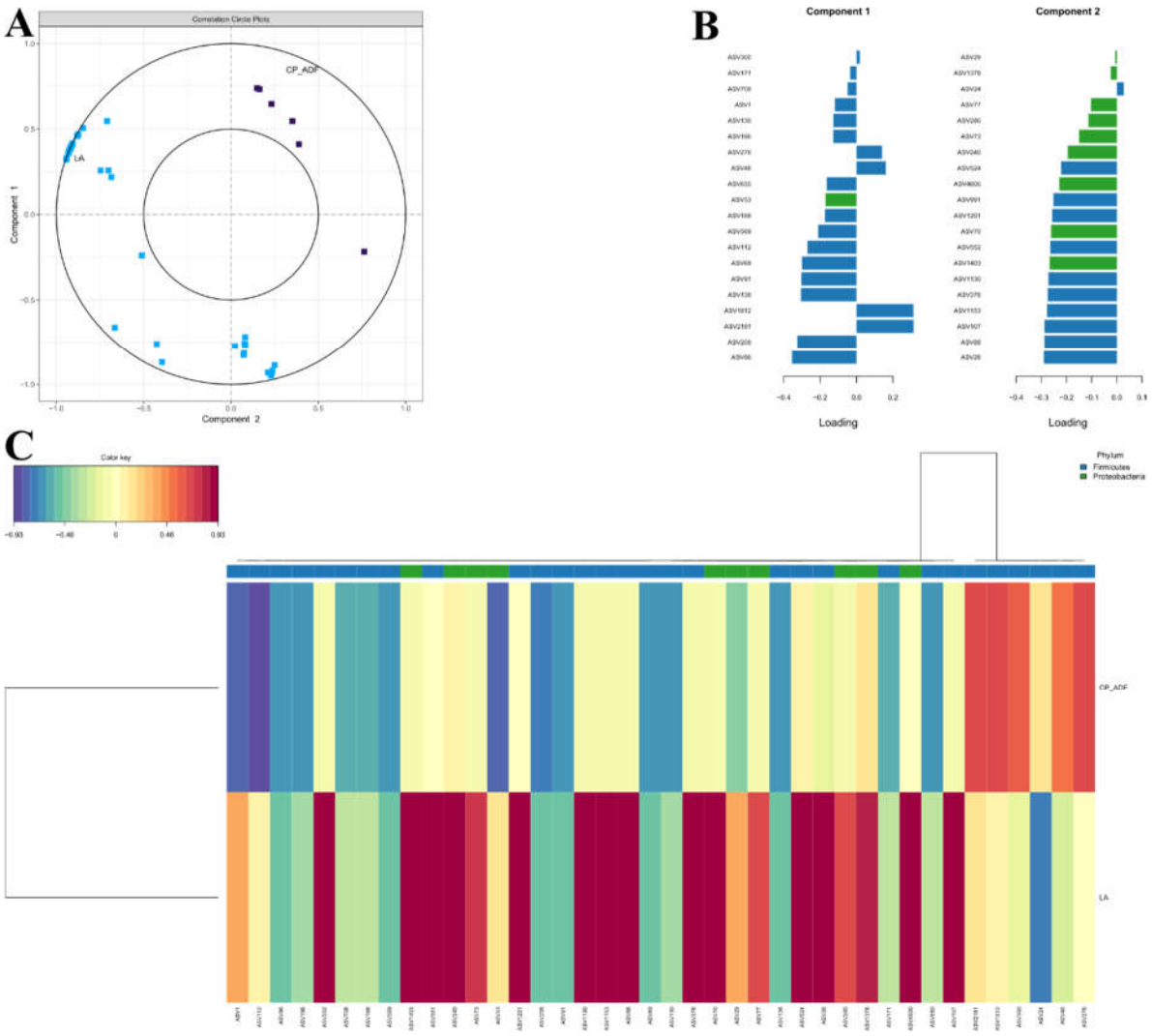


Figure B11: Sparse partial least square regression analysis of physicochemical parameters and bacterial ASV for uninoculated corn silage. (A) Correlation circle plot based on the first two dimensions. (B) Loading plots of ASV showing their contributions to the first (left) and second (right) components. (C) Clustered image map showing correlations between physicochemical parameters and ASV.

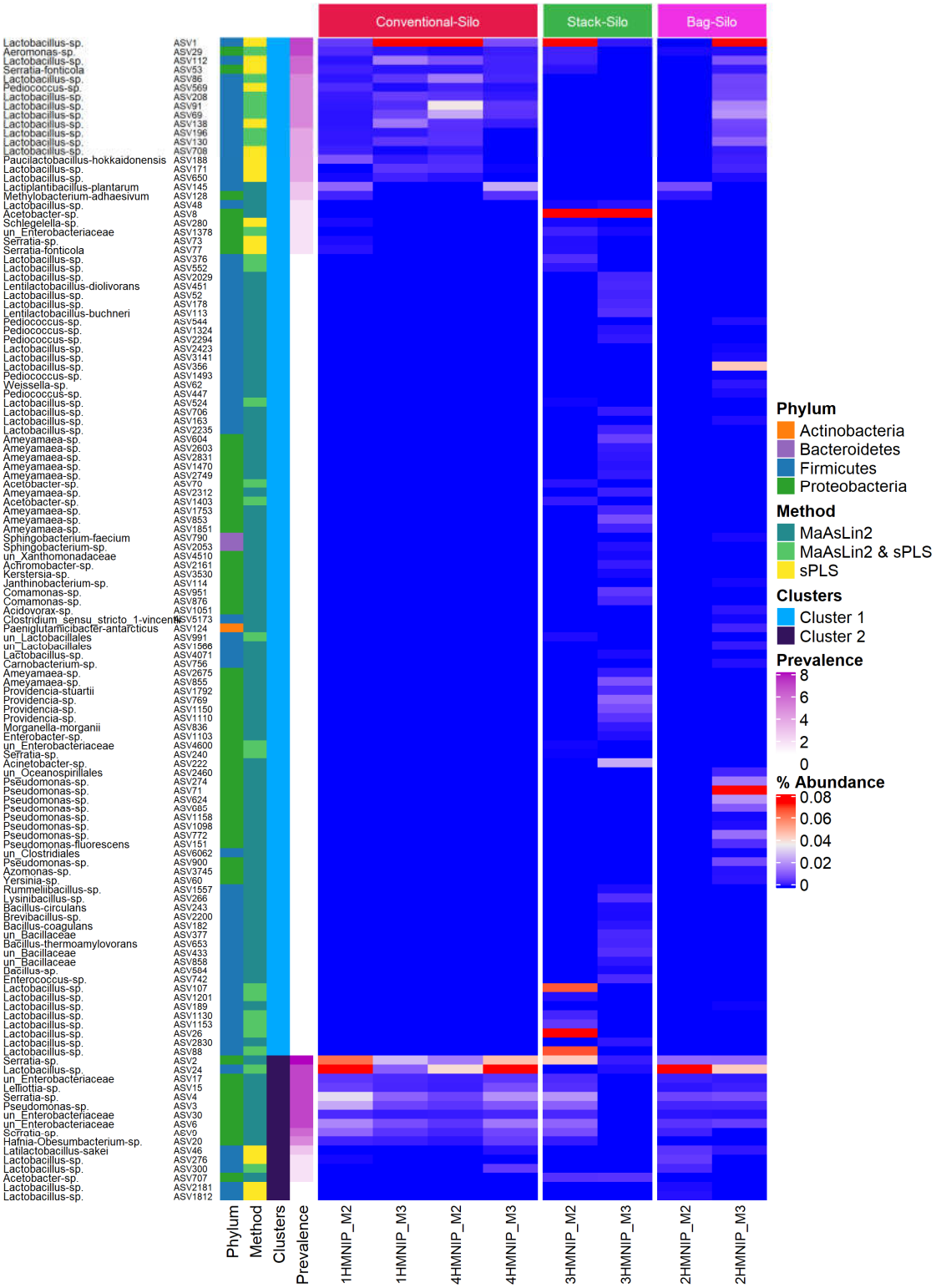


Figure B12: Prevalence, abundance, and distribution across clusters of ASV significantly correlated with physicochemical parameters for uninoculated corn silage.

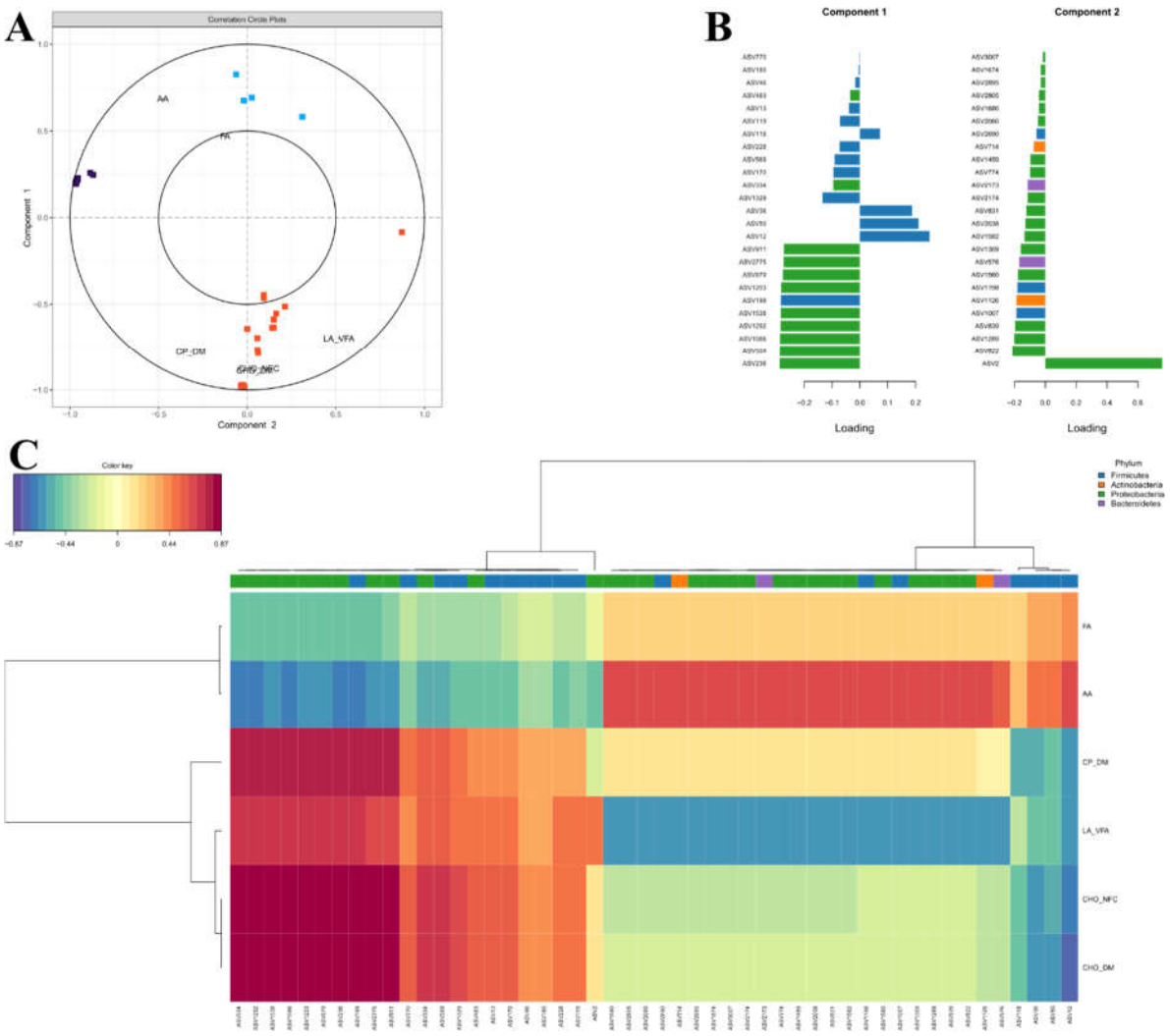


Figure B13: Sparse partial least square regression analysis of physicochemical parameters and bacterial ASV for inoculated corn silage. (A) Correlation circle plot based on the first two dimensions. (B) Loading plots of ASV showing their contributions to the first (left) and second (right) components. (C) Clustered image map showing correlations between physicochemical parameters and ASV.

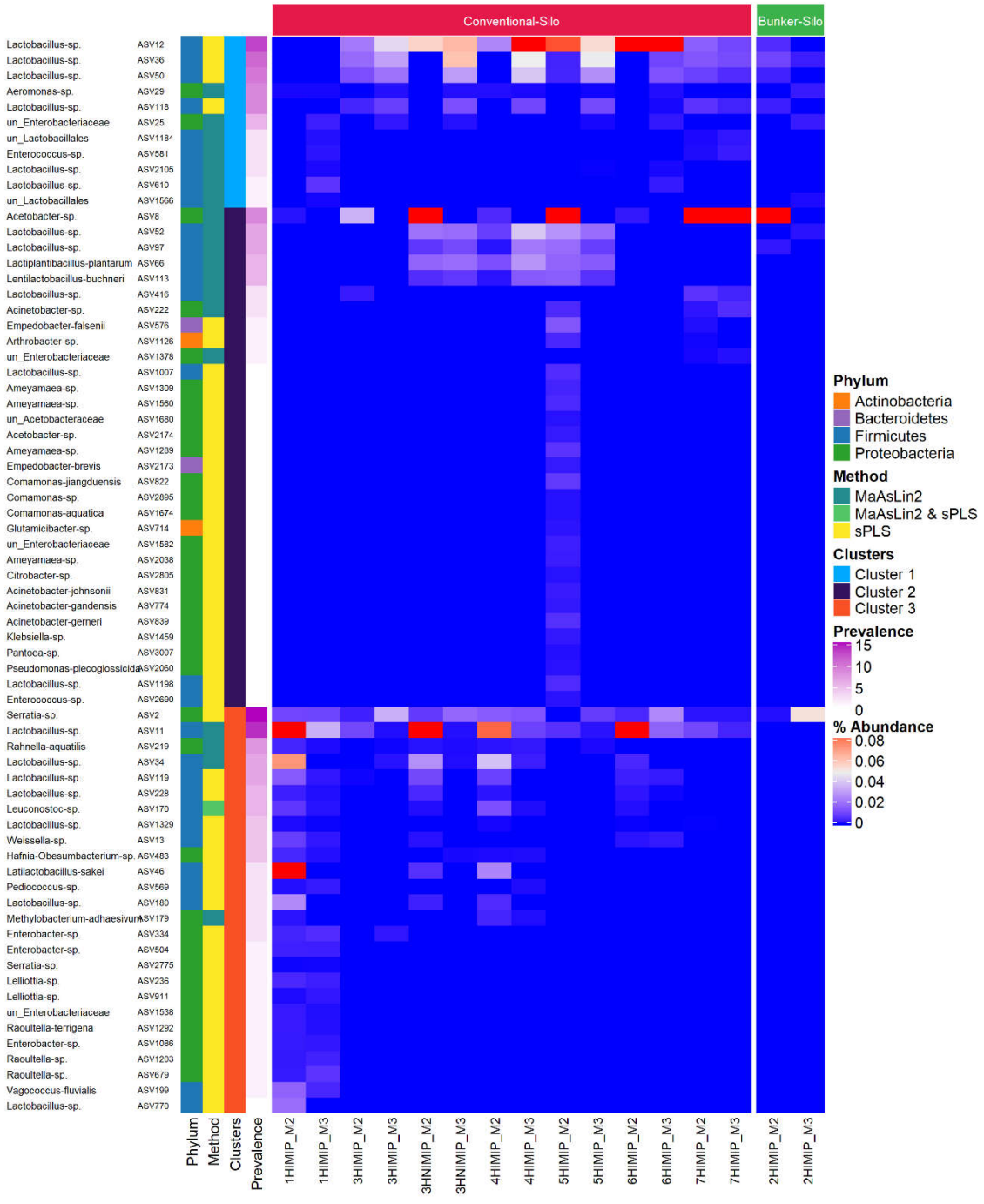
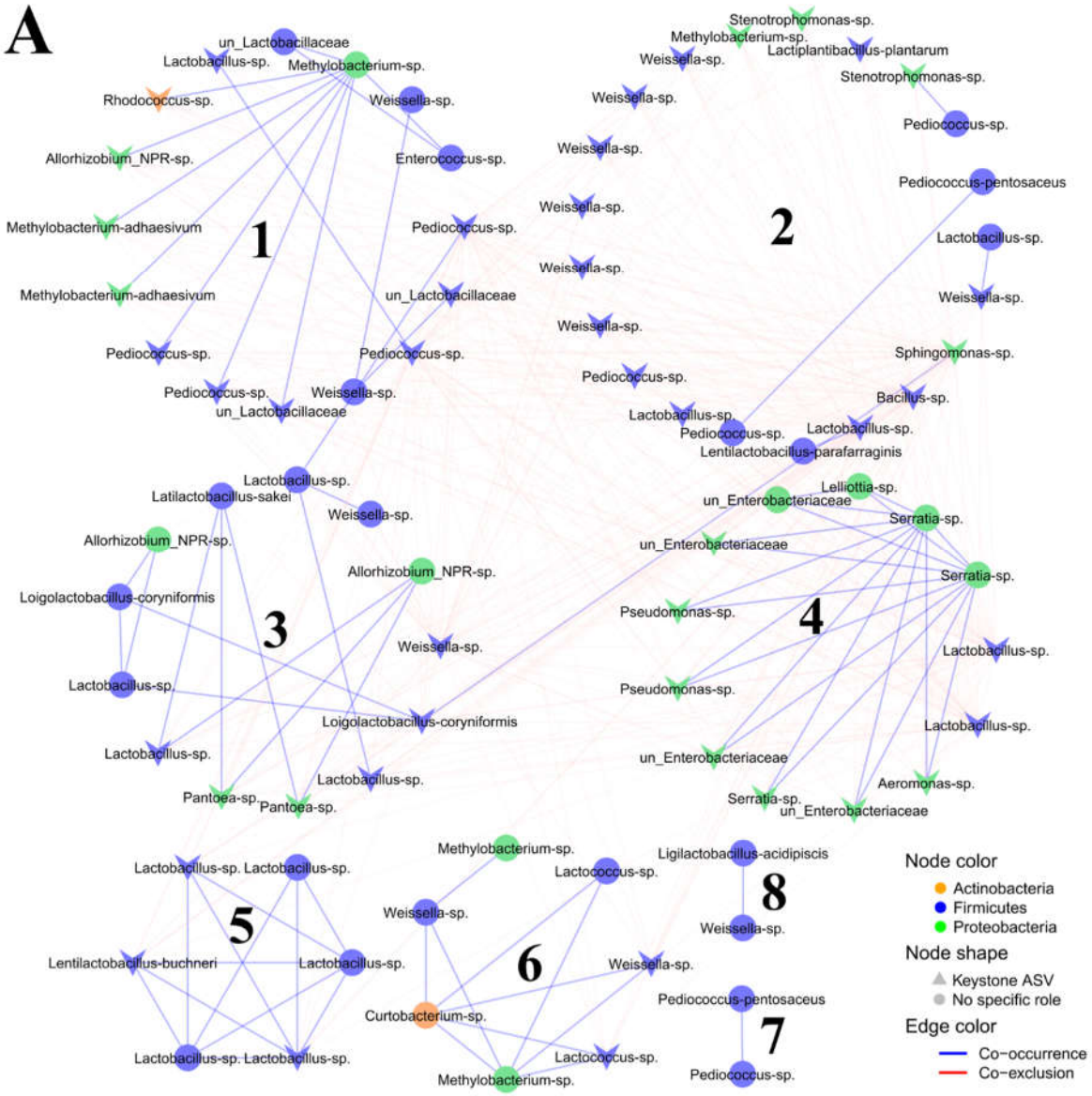
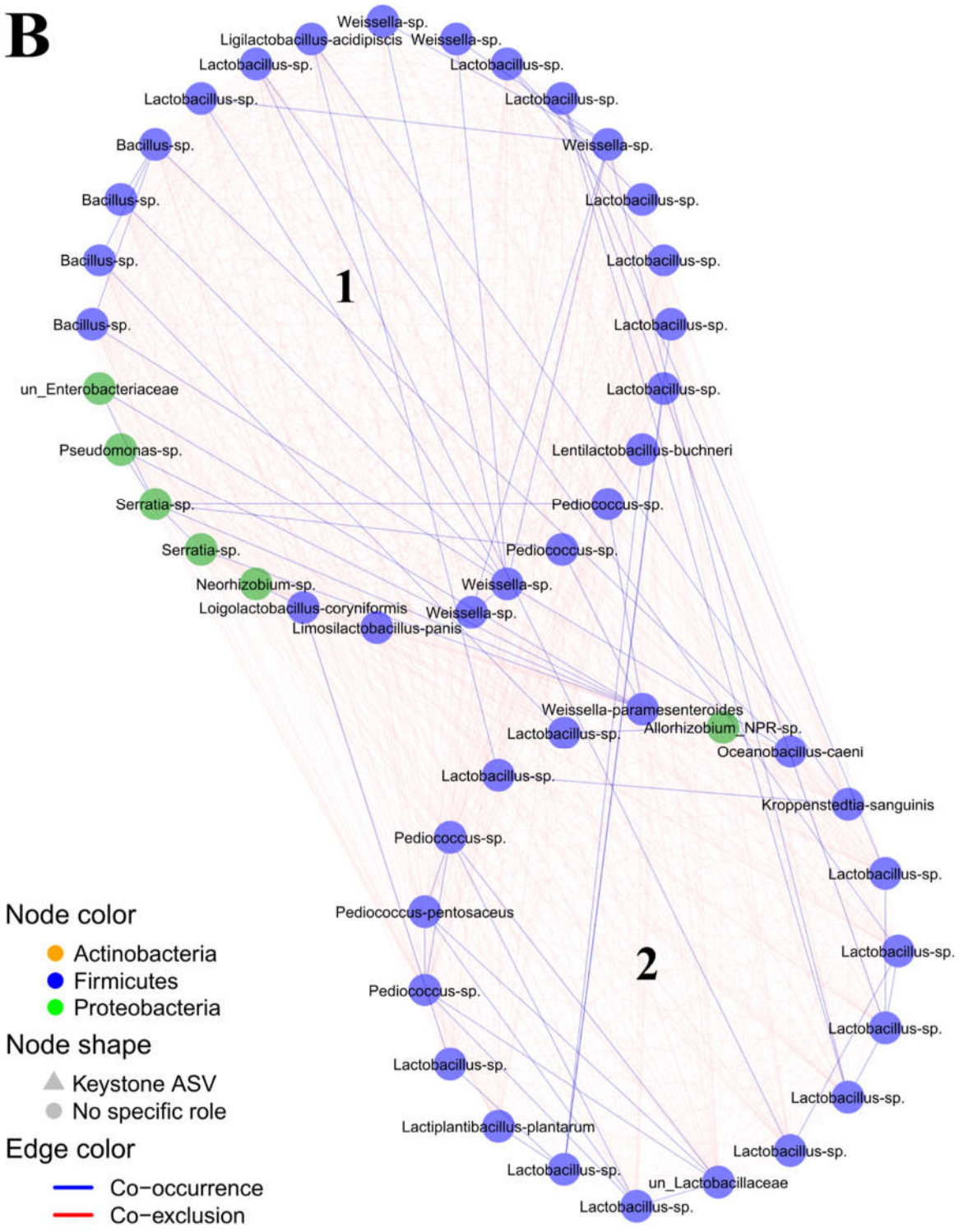


Figure B14: Prevalence, abundance, and distribution across clusters of ASV significantly correlated with physicochemical parameters for inoculated corn silage.


A
1
2
3
4
5
6
7
8
un_Lactobacillaceae
Lactobacillus-sp.
Methylobacterium-sp.
Rhodococcus-sp.
Weissella-sp.
Allorhizobium_NPR-sp.
Enterococcus-sp.
Methylobacterium-adhaesivum
Pediococcus-sp.
Methylobacterium-adhaesivum
un_Lactobacillaceae
Pediococcus-sp.
Pediococcus-sp.
Pediococcus-sp.
Weissella-sp.
un_Lactobacillaceae
Lactobacillus-sp.
Latilactobacillus-sakei
Weissella-sp.
Allorhizobium_NPR-sp.
Allorhizobium_NPR-sp.
Loigolactobacillus-coryniformis
Weissella-sp.
Lactobacillus-sp.
Loigolactobacillus-coryniformis
Lactobacillus-sp.
Lactobacillus-sp.
Pantoea-sp.
Pantoea-sp.
Stenotrophomonas-sp.
Methylobacterium-sp.
Lactiplantibacillus-plantarum
Weissella-sp.
Stenotrophomonas-sp.
Weissella-sp.
Pediococcus-sp.
Weissella-sp.
Pediococcus-pentosaceus
Weissella-sp.
Lactobacillus-sp.
Weissella-sp.
Weissella-sp.
Sphingomonas-sp.
Pediococcus-sp.
Bacillus-sp.
Lactobacillus-sp.
Pediococcus-sp.
Lactobacillus-sp.
Lentilactobacillus-parafarraginis
Lelliottia-sp.
un_Enterobacteriaceae
Serratia-sp.
un_Enterobacteriaceae
Serratia-sp.
Pseudomonas-sp.
Lactobacillus-sp.
Pseudomonas-sp.
Lactobacillus-sp.
un_Enterobacteriaceae
Aeromonas-sp.
Serratia-sp.
un_Enterobacteriaceae
Lactobacillus-sp.
Lactobacillus-sp.
Lentilactobacillus-buchneri
Lactobacillus-sp.
Lactobacillus-sp.
Lactobacillus-sp.
Methylobacterium-sp.
Lactococcus-sp.
Weissella-sp.
Curtobacterium-sp.
Weissella-sp.
Lactococcus-sp.
Methylobacterium-sp.
Ligilactobacillus-acidipiscis
Weissella-sp.
Pediococcus-pentosaceus
Pediococcus-sp.
Node color
Actinobacteria
Firmicutes
Proteobacteria
Node shape
Keystone ASV
No specific role
Edge color
Co−occurrence
Co−exclusion


B
Weissella-sp.
Ligilactobacillus-acidipiscis
Weissella-sp.
Lactobacillus-sp.
Lactobacillus-sp.
Lactobacillus-sp.
Lactobacillus-sp.
Bacillus-sp.
Weissella-sp.
Bacillus-sp.
Lactobacillus-sp.
Bacillus-sp.
Lactobacillus-sp.
1
Bacillus-sp.
Lactobacillus-sp.
un_Enterobacteriaceae
Lactobacillus-sp.
Pseudomonas-sp.
Lentilactobacillus-buchneri
Serratia-sp.
Pediococcus-sp.
Serratia-sp.
Pediococcus-sp.
Neorhizobium-sp.
Weissella-sp.
Loigolactobacillus-coryniformis
Weissella-sp.
Limosilactobacillus-panis
Weissella-paramesenteroides
Lactobacillus-sp.
Allorhizobium_NPR-sp.
Oceanobacillus-caeni
Lactobacillus-sp.
Kroppenstedtia-sanguinis
Pediococcus-sp.
Lactobacillus-sp.
Pediococcus-pentosaceus
Lactobacillus-sp.
2
Pediococcus-sp.
Lactobacillus-sp.
Lactobacillus-sp.
Lactiplantibacillus-plantarum
Lactobacillus-sp.
Lactobacillus-sp.
un_Lactobacillaceae
Lactobacillus-sp.
Node color
Actinobacteria
Firmicutes
Proteobacteria
Node shape
Keystone ASV
No specific role
Edge color
Co−occurrence
Co−exclusion

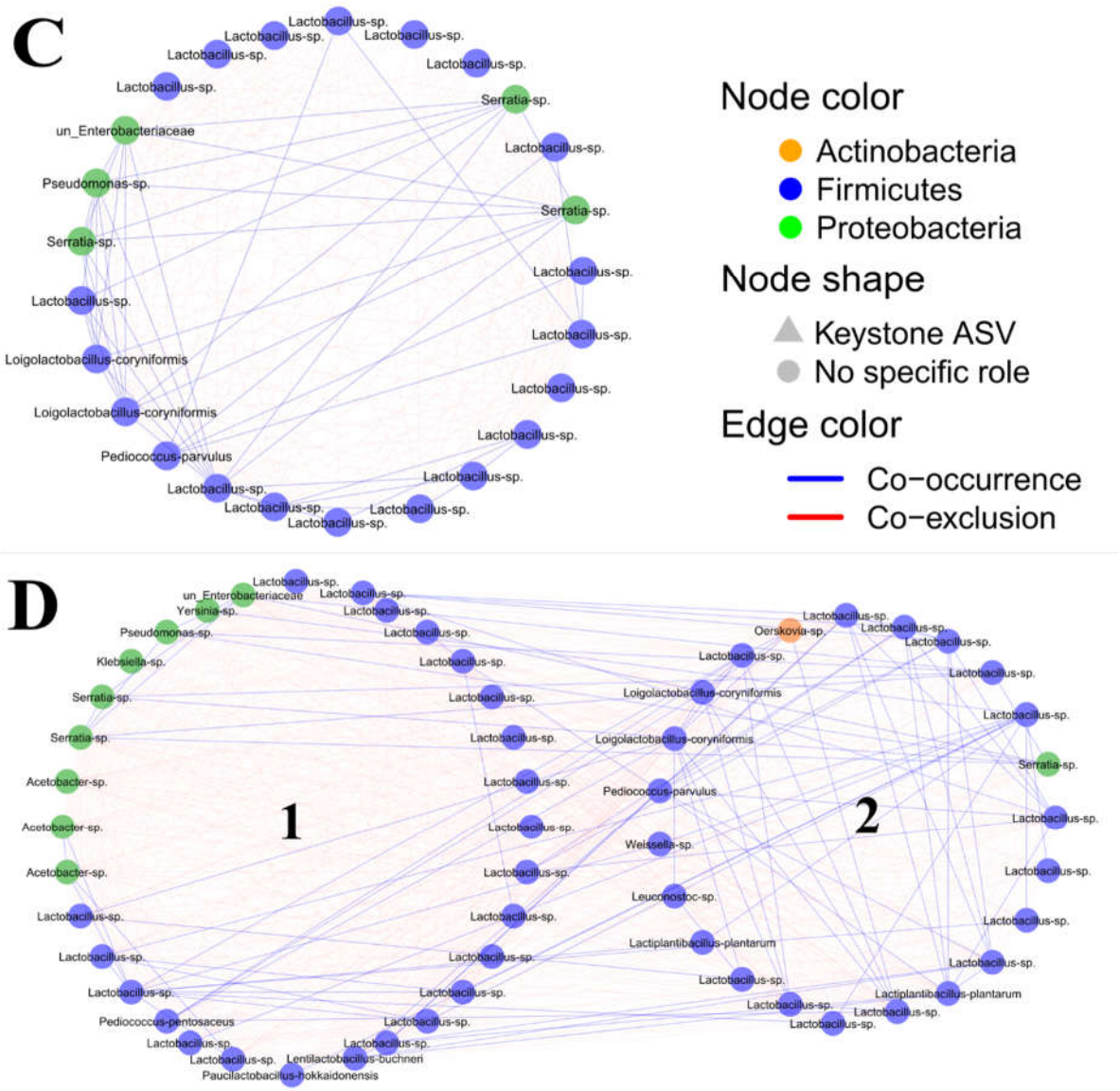


Figure B15: Network analysis of bacterial occurrence. Co-occurring and co-excluding ASV in uninoculated grass/legume (A), inoculated grass/legume (B), uninoculated corn (C), and inoculated corn (D) silage. Co-occurrence relationships (blue lines) are emphasized over co-exclusion (red).

B-1 Supplemental Results

B-1.1 Forage physicochemical characteristics associated with microbial communities

Among the 60 microbial features selected by the sPLS approach within H community assemblies, *Proteobacteria* including an unclassified *Rhizobiaceae*, a species of the group *Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium* (*Allorhizobium-NPR* sp.), and *Pantoea agglomerans*, *Actinobacteria* including unclassified *Microbacteriaceae*, and *Bacteroidetes* represented by *Spirosoma* sp. exhibited the highest loading weights in sPLS component 2 (Fig. B5 B), showing their high contribution in the separation of clusters in component 2 (Fig. B5 B-C). The MaAsLin2 approach selected 55 microbial features significantly associated with magnesium, moisture, and NDF contents, as well as the storage form of hay. MaAsLin2 is a complementary approach to sPLS in that it can integrate categorical variables in the analysis model. The top 50 features with significant associations are shown in the supplementary data (Fig. B5 D). *Actinobacteria* including *Nocardioides* sp. and *Proteobacteria* including *Serratia* sp., *Pseudomonas* sp., and *Yersinia* sp. that were found significantly associated with moisture content (Fig. B5 D) are expected to be part of cluster 1. The canonical relationships between taxa found to be significantly associated with quantitative and categorical variables identified after sPLS and MaAsLin2 analyses were modelled using CCpnA. This analysis revealed that within the variable describing hay storage forms detected by MaAsLin2 analysis, loose hay, not square bales, correlated with community variation. CCpnA also confirmed the impact of sampling periods on taxa occurrence and abundance within H microbiotas.

For the GL samples, *Firmicutes* represented by *Lactobacillus* spp., *Proteobacteria* including *Methylobacterium* sp. and *Sphingomonas* sp., and *Actinobacteria* represented by *Curtobacterium* sp. contributed the much to the separation between clusters in component 1 (Fig. B7 B). On the other hand, *Firmicutes* including *Weissella* spp., *Lactobacillus* spp., and *Pediococcus* sp., *Actinobacteria* comprising *Aeriscardovia* spp. and *Corynebacterium* sp., and finally *Proteobacteria* represented by *Serratia* sp. exhibited the highest contribution to cluster separation in component 2 (Fig. B7 B). Of the 77 ASV distributed among the three clusters, *Firmicutes* almost essentially composed cluster 1, indicating distinctively higher correlations between putative taxa and moisture, FA and VFA (Fig. B7 C). *Proteobacteria*, mostly positively correlated with pH, mainly composed cluster 2. Other *Firmicutes* showing stronger associations with CP_NH3 mainly composed cluster 3. Complementarily, MaAsLin2 analysis identified 27 ASV significantly associated with pH values, LA, moisture, crude fat (CF), and ammonia (NH3). For instance, while some *Firmicutes* including *Lactobacillus* spp., *L. buchneri*, *Pediococcus parvulus*, and *Carnobacterium* sp. were found negatively correlated with pH, hence indicating their high contribution to silage low pH, others such as *Weissella paramesenteroides*, *Lactococcus* sp., *Pediococcus* sp., and *P. pentosaceus* were positively associated with pH, thus under certain circumstances

involved in silage high pH outcomes. Consequently, the latter ASV would drop into cluster 2, and the remaining into cluster 1. The analysis of these canonical relationships also revealed that in addition to pH, ASV composing cluster 2 were correlated with silage storage forms including square bales, round bales, and oxygen-limiting silos (Fig. 3-8B, Fig. B7 D). On the other hand, all ASV composing cluster 1 and most ASV composing cluster 3, principally those identified by the sPLS approach and generally highly correlated with moisture, LA, FA, CP, and VFA, were also correlated with conventional and stack silos (Fig. 3-8B). It appears that cluster 1, specifically composed of ASV that were correlated with LA and moisture, are largely dominated by *Firmicutes*, although *Erwinia* sp., and another unclassified *Enterobacteriaceae* were present. Almost all ASV in this cluster largely occur in conventional silos (concrete-stave silos), bag silos, and oxygen-limiting silos. ASV composing cluster 2 broadly exhibited highest prevalence levels, particularly within round bales, square bales, and bag silo groups, wherein they generally occurred at low relative abundance, except *P. pentosaceus*, *W. paramesenteroides*, and *Pediococcus* sp. that exhibited highest relative abundance (Fig B8). ASV that fell into cluster 3 were generally the least prevalent and abundant, occurring almost essentially in conventional and oxygen-limiting silos.

In the case of GLI samples, sPLS regression selected for 79 ASVs of which *Proteobacteria* represented by *Methylobacterium* sp and *Firmicutes* including *Pediococcus* sp. and *Lactobacillus* sp. exhibited high contributions for the separation of clusters in component 1, while *Firmicutes* including *Lactobacillus* spp. and an unclassified *Clostridiaceae* contributed the most to cluster separation in component 2 (Fig. B9 B). The MaAsLin2 analysis resulted in the selection of 67 ASV that variably correlated with the type of inoculant used during ensiling, the form of storage applied, or ESC, LA, acetic acid (AA), calcium, and phosphorus contents (Fig. B9 C). Some *Firmicutes* including *Lactobacillus* spp. and *Weissella* sp. showed positive correlations with both AA and LA contents, indicating the influence of the ensiling environment on the species development. The ASV composing cluster 1 were mostly associated with silage storage forms including oxygen-limiting and conventional silos, as well as inoculant types including 11G22 and Biotal Supersile, while ASV in cluster 2 mostly associated with the inoculant Biotal Buchneri 500. However, cluster 1 grouped ASV with the highest prevalence levels and broader distribution patterns among conventional and oxygen-limiting silos (Fig. B10). Except for *Corynebacterium* sp. and *Yersinia* sp. that belong to the phyla *Actinobacteria* and *Proteobacteria* respectively, all ASV in this cluster were *Firmicutes* of the genera *Lactobacillus*, *Ligilactobacillus*, *Limosilactobacillus*, *Lentilactobacillus*, *Weissellla*, *Pseudogracilibacillus*, *Pediococcus*, and *Staphylococcus*. Within cluster 2, all ASV, practically absent from oxygen-limiting silos, were mostly *Proteobacteria*, while in cluster 3, ASV mostly included *Firmicutes* with few occurrences in oxygen-limiting silos.

Within the selected microbiota of uninoculated silage, *Firmicutes* represented by *Lactobacillus* spp. contributed the most to the separation of clusters on component 1, while other *Lactobacillus* spp. contributed the most to cluster separations on component 2 (Fig. B11 B). The 40 ASV composing the sPLS derived clusters

belonged to the phyla *Firmicutes* and *Proteobacteria* (Fig. B11 C). The MaAsLin2 modelling revealed 113 ASV significantly associated with CP_ADF, LA, and the silage storage forms. Compared with the sPLS approach, additional phyla including *Actinobacteria* and *Bacteroidetes* were selected. Most of these taxa were positively correlated to silage storage forms, while only six including *Lactobacillus* spp., *Acetobacter* spp., an unclassified *Lactobacillales,* and *Enterobacteriaceae* positively correlated with LA content. On the other hand, among ASV negatively associated with LA were *Firmicutes* and mostly *Proteobacteria*. *Firmicutes* included *Pedicoccus* sp., *L. diolivorans*, *L. buchneri*, and other lactobacilli. *Proteobacteria* were represented by *Ameyamaea* spp., *Achromobacter* sp., *Comamonas* sp., *Providencia stuartii*, *Providencia* spp., and *Morganella morganii*. Taxa grouped in cluster 1 appeared to mostly associate with bag and stack silos, while those in cluster 2 correlated with conventional silos. As illustrated in Fig. B12, taxa that significantly correlated with C metadata were not uniformly distributed among silage storage forms. Broadly, lower prevalence levels were observed within members of cluster 1 compared with those in cluster 2, and taxa with the highest relative abundance were not necessarily the more prevalent. The latter included as members of cluster 1 *Lactobacillus* sp. occurring in conventional, stack, and bag silos, *Acetobacter* sp. occurring in stack silos, *Pseudomonas* sp. occurring in a Bag silo, and *Lactobacillus* spp. occurring in a stack silo. Among ASV composing cluster 2, those with highest relative abundance included *Serratia* sp. occurring mostly in conventional and stack silos, and *Lactobacillus* sp. occurring mostly in conventional and bag silos.

Of the 50 ASV identified by the sPLS approach (Fig. B13 C) within the CI microbial community, *Proteobacteria* including *Lelliottia* sp., *Enterobacter* spp., *Raoultella terrigena*, and an unclassified *Enterobacteriaceae*, and *Firmicutes* represented by *Lactobacillus* spp. and *Vagococcus fluvialis* highly contributed to the separation of clusters on the sPLS component 1, while a *Proteobacteria* identified as *Serratia* sp. exhibited a strong contribution to cluster separation on component 2 (Fig. B13 B). Modelling the relationships between the selected variables and microbial communities using MaAsLin2, we identified 20 ASV significantly associated with CP, AA, and ESC, as well as the silage storage form and the type of inoculant used for ensiling. *Proteobacteria* represented by *Methylobacterium adhaesivum*, and *Firmicutes* including *Lactobacillus* sp. and *Leuconostoc* sp. were negatively correlated with AA, indicating their probable sensitivity to high amounts of AA. Moreover, ASV composing cluster 1 mostly associated with the inoculant 11CFT, while those found in cluster 2 mostly associated with the inoculant 11C33 and with bunker silo. As depicted in Fig. B14, of all the taxa selected, *Proteobacteria* and *Firmicutes* were the most represented. However, *Firmicutes* which mostly composed cluster 1 were the more prevalent, being largely distributed across conventional and bunker silos. The most abundant taxa included *Lactobacillus* sp. occurring in conventional silos, *Acetobacter* sp. occurring in conventional silos and a bunker silo, and *Lactobacillus* sp. and *Latilactobacillus sakei* both occurring in conventional silos.

B-1.2 Molecular ecological network analyses

Within the GL network, we found that the node corresponding to *L. buchneri* positively interacted with *Lactobacillus* spp. and showed negative relationships with other *Lactobacillus* spp. as well as *Pediococcus* spp., *Weissella* sp., *Bacillus* sp., and *Sphingomonas* sp. *Lactiplantibacillus plantarum*, represented by a single phylotype, exhibited only negative relationships with four ASVs including *Pediococcus* sp., *Lactobacillus* sp., *Bacillus* sp., and *Sphingomonas* sp. In addition to *Sphingonomas* sp., other *Proteobacteria* phylotypes including *M. adhaesivum*, *Allorhizobium* group., *Pantoea* spp., and *Stenotrophomonas* positively interacted with *Firmicutes* comprising *Pediococcus* spp., *Enterococcus* sp., *Loigolactobacillus coryniformis*, and *Lt. sakei* (Fig. B15 A).

On the other hand, the GLI network (Fig. B15 B) was composed of 47 nodes of which more than 87 % were *Firmicutes* and the remaining *Proteobacteria*. *L. buchneri* only co-occurred with *Lactobacillus* sp., while sharing negative relationships with more than 87 % of nodes, including *Proteobacteria* (*Pseudomonas* sp., *Serratia* spp., *Allorhizobium* group) and *Firmicutes* (*Bacillus* spp., *Lactobacillus* spp., *Lactobacillus acidipiscis*, *Limosilactobacillus panis*, *L. coryniformis*, *Pediococcus* spp., *P. pentosaceus*, *Weissella* spp., *W. paramesenteroides*, *Oceanobacillus caeni*, and *Kroppenstedtia sanguinis*). Like *L. buchneri*, *L. plantarum* exhibited the same pattern of interactions with other nodes in the network, sharing a unique positive relationship with a phylotype of *Lactobacillus* sp. and negative relationships with the same taxa as *L. buchneri*, except *L. coryniformis*. Of the 47 keystones phylotypes identified in the GL pMEN, the two network hubs included a *Firmicutes* classified as *Weissella* sp. and a *Proteobacteria* classified as *Sphingomonas* sp. (Table B4). Connectors were composed of *Actinobacteria* including *Rhodococcus* sp., *Firmicutes* represented by a phylotype of *Bacillus* sp., *L. buchneri*, *L. plantarum*, *L. coryniformis*, *Lactococcus* sp., nine phylotypes of *Lactobacillus* spp., eight phylotypes of *Weissella* spp., and two phylotypes of unclassified *Lactobacillaceae*, and *Proteobacteria* including a phylotype of *Aeromonas* sp., *Allorhizobium* group, *Methylobacterium* sp., *Serratia* sp., two phylotypes of *M. adhaesivum*, *Pantoea* spp., *Pseudomonas* spp., *Stenotrophomonas* spp., and three phylotypes of unclassified *Enterobacteriaceae* (Table B4).

The C network was composed of 80 % Firmicutes including 17 phylotypes of *Lactobacillus* spp., two phylotypes of *L. coryniformis*, and one phylotype of *Pediococcus parvulus*, and 20 % *Proteobacteria* represented by one phylotype of *Pseudomonas* sp., three phylotypes of *Serratia* spp., and one phylotype of unclassified *Enterobacteriaceae*. All the five phylotypes of *Proteobacteria* positively interacted with each other, and with *Firmicutes* including principally *Lactobacillus* spp. and *L. coryniformis*, indicating that these lactobacilli covaried with *Proteobacteria*. However, the other phylotypes of lactobacilli and that of *P. parvulus* negatively interacted with *Proteobacteria* (Fig. B15 C). On the other hand, 55 nodes composing the CI network were subdivided into two modules (Fig. B15 D), involving mostly *Firmicutes* (80 %) and *Proteobacteria* (18.18 %), the phylum

Actinobacteria being represented by a single phylotype. Among the *Firmicutes* contained in the CI network were one phylotype of *L. buchneri*, two phylotypes of *L. plantarum*, two phylotypes of *L. coryniformis*, one phylotype of *Paucilactobacillus hokkaidonensis*, 34 phylotypes of *Lactobacillus* spp., one phylotype of *P. parvulus*, *P. pentosaceus*, *Weissella* sp., and *Leuconostoc* sp. *Proteobacteria* of this network included three phylotypes of *Acetobacter* spp. and *Serratia* spp., one phylotype of *Klepsiella* sp., *Pseudomonas* sp., *Yersinia* sp., and an unclassified *Enterobacteriaceae*. *Oerskovia* sp. was the sole *Actinobacteria* of the CI network. *L. buchneri* positively interacted with *Firmicutes* only, including *P. parvulus*, *Lactobacillus* sp., and *L. plantarum*, while exhibiting negative relationships with almost all *Proteobacteria* and *Actinobacteria*, as well as most of other *Firmicutes* including *L. plantarum*, Pl. *hokkaidonensis*, *L. coryniformis*, and *P. pentosaceus*. *L. plantarum* shared positive relationships with *P. parvulus* and eight other lactobacilli, and negative relationships with all *Proteobacteria*, *Actinobacteria*, and most of the other *Firmicutes*. The two phylotypes of *L. plantarum* had different node degrees, probably indicating different ecological roles in the silage fermentation process. Concerning *P. pentosaceus*, another species constitutive of some inoculants, positive relationships were shared with only few *Firmicutes* comprising *L. coryniformis*, while negative relationships were shared with most *Proteobacteria* and other *Firmicutes*.

Annexe C – Matériel supplémentaire du chapitre 4

Table C1: Milk characteristics in the fall 2015 and the spring 2016

Parameters	Fall 2015				p value	Spring 2016				p value
	H (n=5)	GL (n=7)	GLC (n=4)	GLICI (n=7)		H (n=5)	GL (n=7)	GLC (n=4)	GLICI (n=7)	
Fat (%)	4.27 ± 0.40	4.02 ± 0.10	4.01 ± 0.20	4.01 ± 0.10	0.32	4.20 ± 0.20	4.01 ± 0.10	3.83 ± 0.20	4.03 ± 0.20	0.05
Protein (%)	3.30 ± 0.01	3.2 ± 0.10	3.30 ± 0.05	3.30 ± 0.10	0.12	3.30 ± 0.10	3.20 ± 0.10	3.20 ± 0.10	3.3 ± 0.10	0.08
Urea-N (mg/dl)	12.60 ± 4.40	12.10 ± 2.50	9.70 ± 1.70	11.20 ± 2.01	0.40	13.80 ± 2.50	11.70 ± 3.50	10.40 ± 3.20	10.10 ± 1.80	0.18
Lactose (%)	4.50 ± 0.10	4.60 ± 0.10	4.60 ± 0.01	4.50 ± 0.10	0.23	4.50 ± 0.13	5.00 ± 0.03	4.60 ± 0.10	4.60 ± 0.10	0.32
SCC (x10^4/ml)	29.60 ± 228.80	14.59 ± 47.50	17.95 ± 80.10	17.97 ± 44.8	0.66	17.44 ± 108.20	16.46 ± 104.50	12.43 ± 56.60	15.31± 29.60	0.70

SCC, somatic cell counts. Mean values are followed with corresponding standard deviation.

Table C2: Topological properties of the empirical pMENs in milk microbial communities and their associated random pMENs

	Empirical networks						Random networks		
Feeding type	Similarity threshold	Network size	Average degree (avgK)	Average path	Average clustering coefficient	Modularity (No. of modules)	Average path	Average clustering coefficient	Modularity
GL	0.81	135	16.13	3.06	0.32	0.32 (4)	2.14 ± 0.01	0.23 ± 0.01	0.16 ± 0.05
GLC	0.37	57	37.44	1.33	0.78	0.00 (1)	1.33 ± 0.00	0.78 ± 0.01	0.02 ± 0.02
GLICI	0.89	135	15.42	1.98	0.28	0.20 (2)	2.13 ± 0.01	0.25 ± 0.01	0.17 ± 0.01

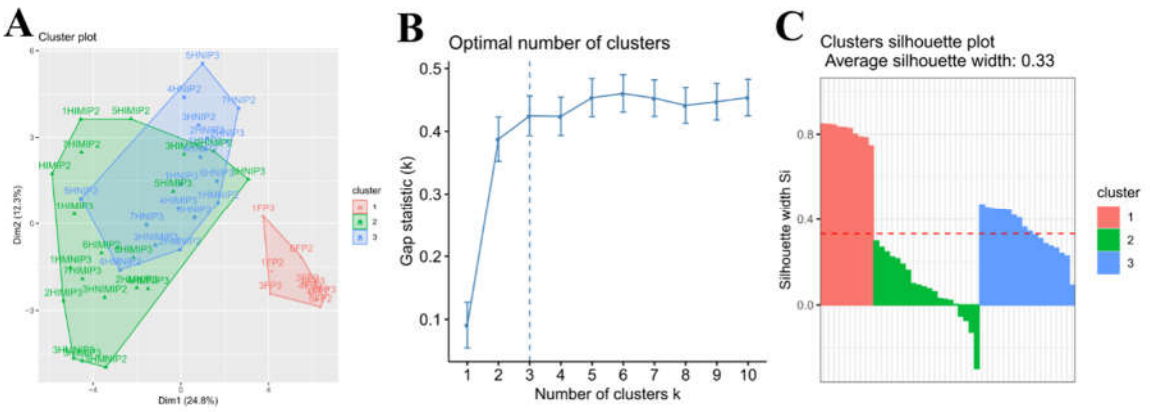


Figure C1: Cluster analysis of forage ration types. Clustering was computed using the portioning around medoids algorithm. (A) Principal component analysis showing the separation of forage ration bacterial communities in three distinct clusters. Gap statistics (B) was performed to estimate the right number of clusters and the silhouette analysis (C) was used for clustering validation.

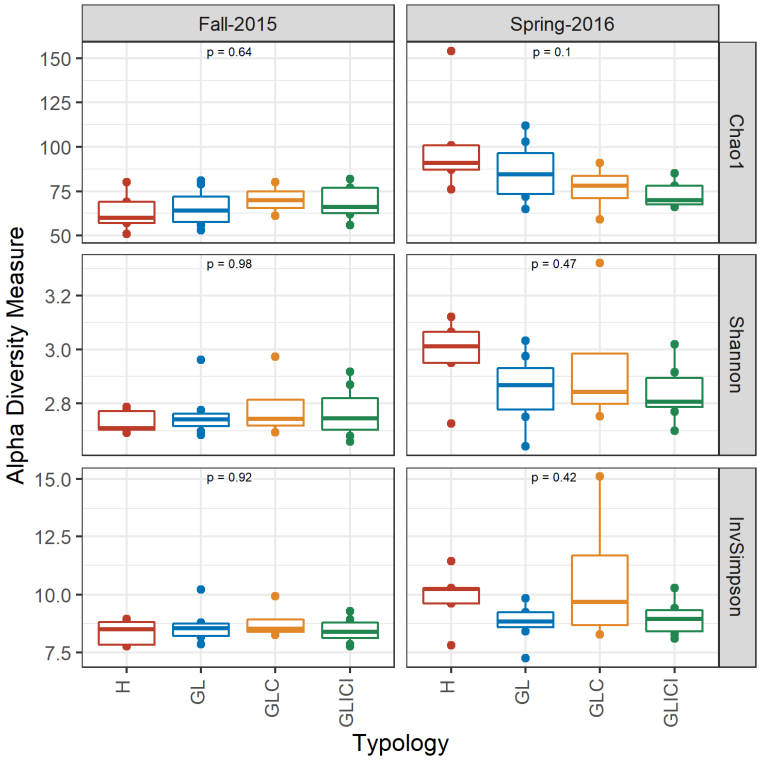


Figure C2: Alpha-diversity of milk bacterial communities. p values indicate the significance of the Kruskal Wallis test performed to compare milk samples from distinct feeding combinations.

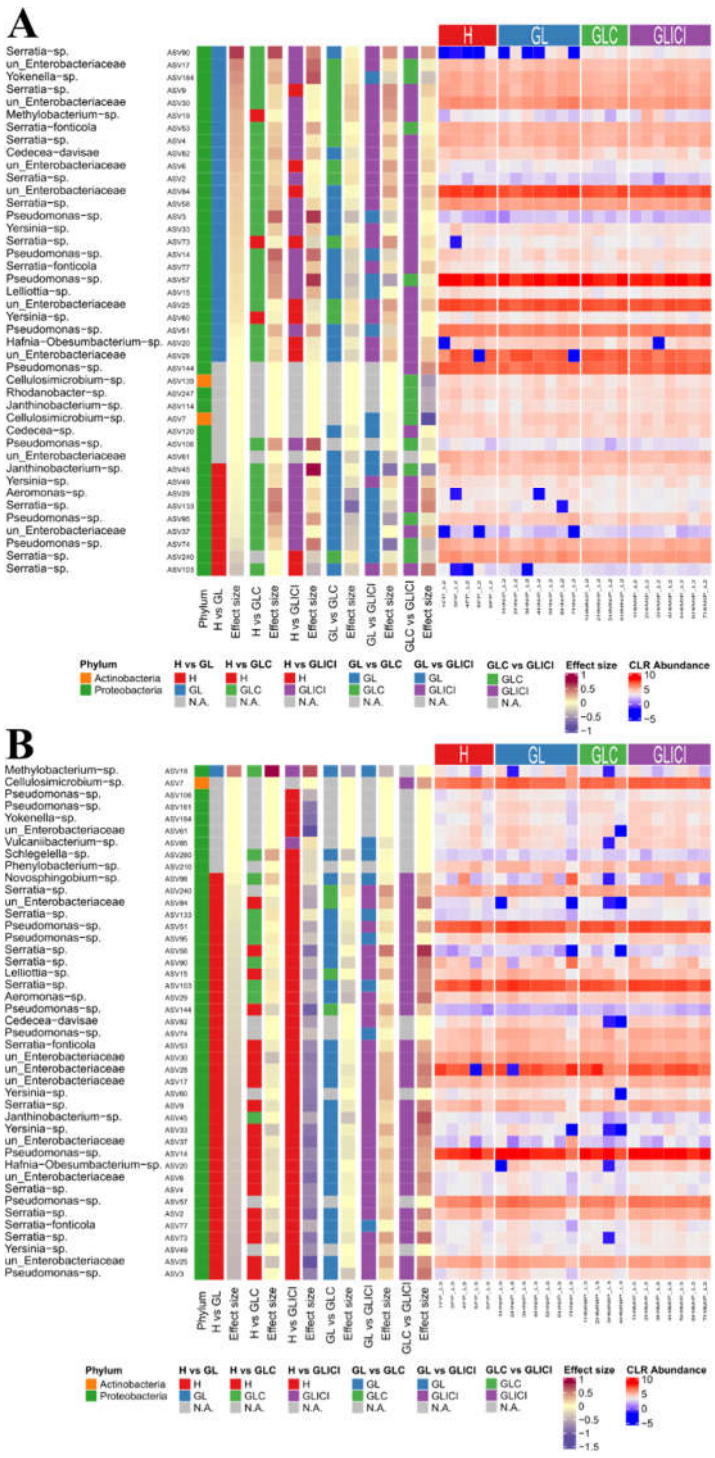


Figure C3: Differentially abundant taxa among milk samples. Differential abundance was performed using the ADEx2 algorithm on microbial community data from milk samples collected in the fall (A) and the spring (B). The heatmap illustrates the distribution and abundance of differentially abundant taxa across feeding types. Corresponding paired group comparisons, effect sizes, and taxonomic classification (phylum level) are displayed as heatmap annotation on the left.

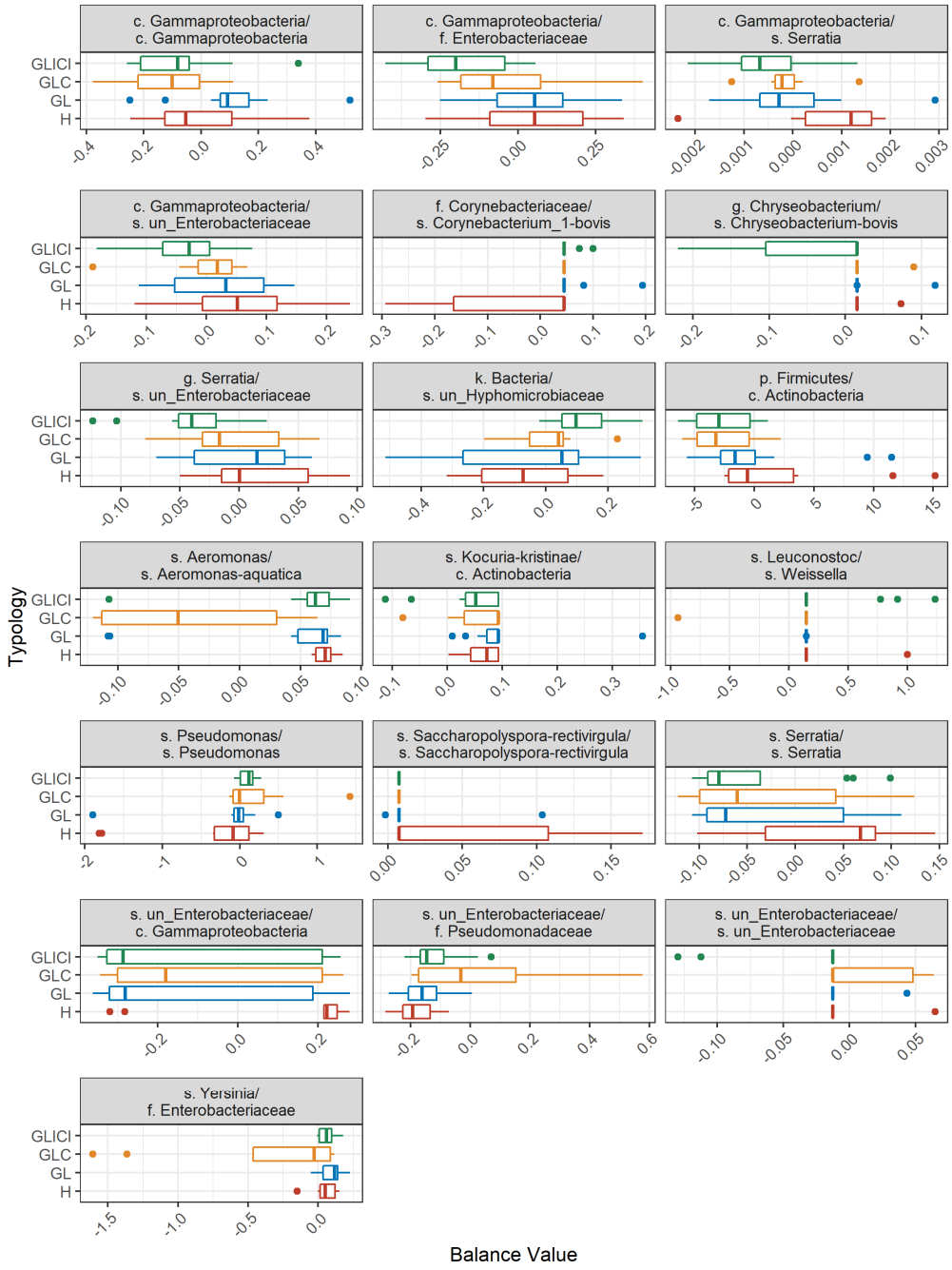


Figure C4: Phylogenetic balances separating milk microbiota by forage type. For example, the balance of the phylum *Firmicutes* relative to the class *Actinobacteria* discriminates H from GLICI and GLC milk samples. *Firmicutes* are represented by *Clostridium disporicum*, *Clostridium* sp., *Paeniclostridium* sp., *Coprococcus* sp., *Romboutsia sedimentorum*, *Romboutsia* sp., *Veillonella dispar*, and an unclassified *Peptostreptococcaceae* while *Actinobacteria* include *Bifidobacterium* spp., *Cellulosimicrobium* spp., *Kocuria* spp., *Corynebacterium* spp., *Saccharopolyspora rectivirgula*, and more.

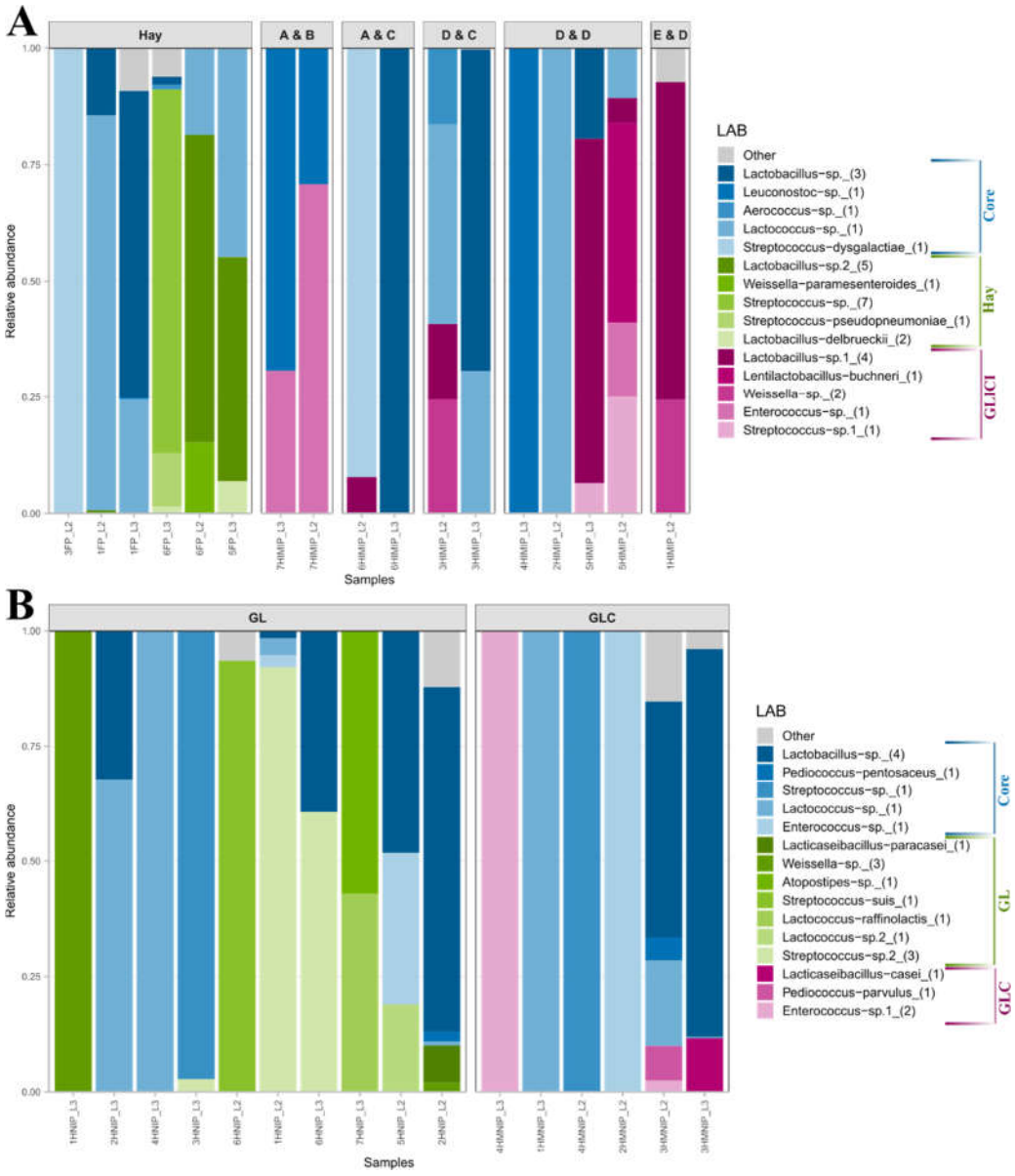


Figure C5: Distribution and proportion of LAB communities among milk groups. Proportion of the core and unique LAB phylotypes among GLICI and H (A) or among GLC and GL (B) milk samples. Milk samples from inoculated forage ration combinations are separated according to the inoculant brands used for the grass or legume (first letter) and corn (second letter) silage. Accordingly, “A” = 11G22, “B” = 11C33, “C” = 11CFT, “D” = Biotal Buchneri 500, and “E” = Biotal Supersile. Taxa are coloured according to whether they are unique to one or the other group (GLICI or H, GLC or GL) or whether they belong to both (core).

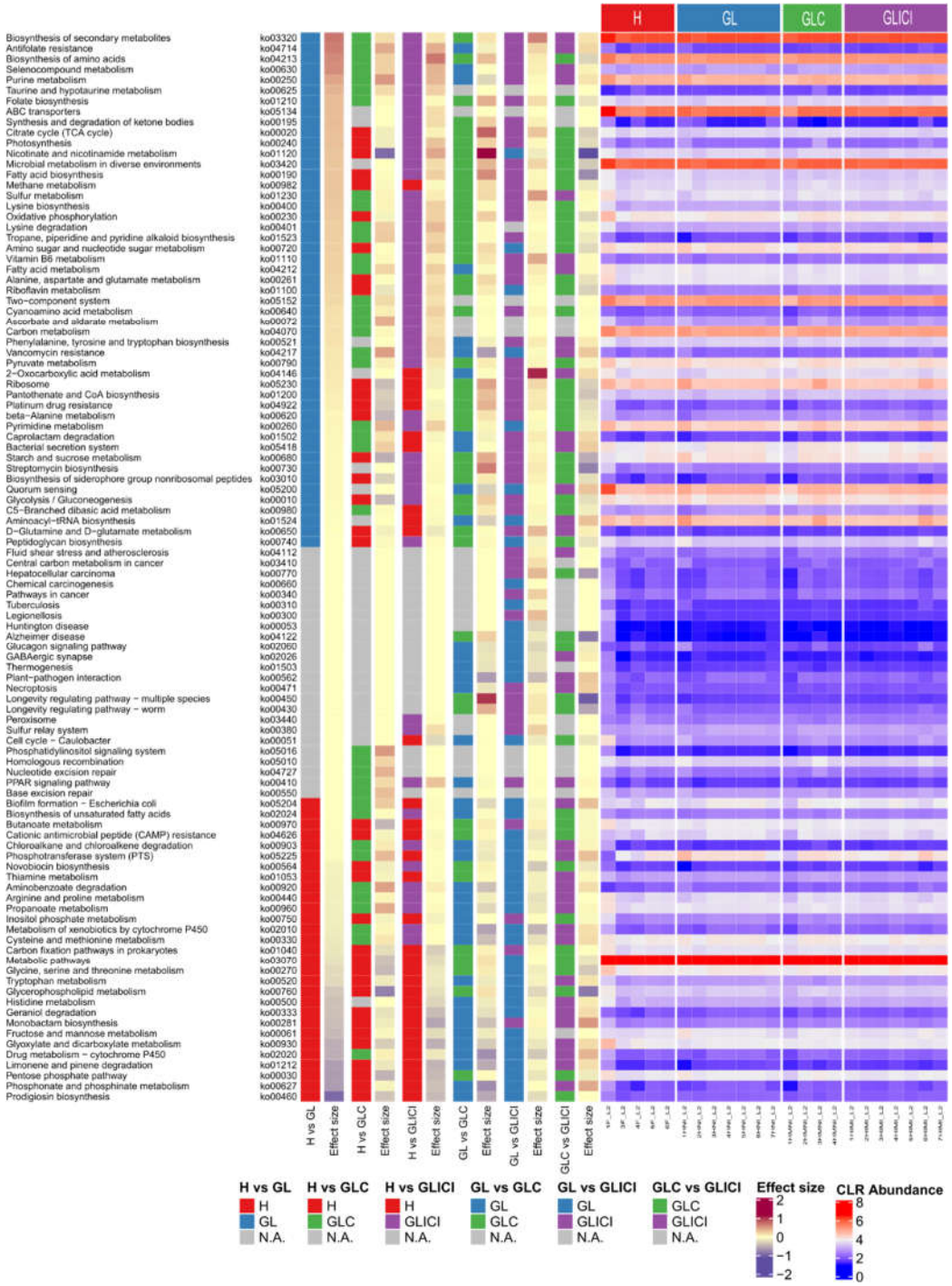


Figure C6: Differentially abundant function pathways among milk samples for fall. Differential abundance was performed using the ADEx2 algorithm on milk sample microbial community data. The heatmap illustrates the distribution and abundance of differentially abundant features across feeding types. Corresponding paired group comparisons, effect sizes, and taxonomic classification (phylum level) are displayed as heatmap annotation on the left.

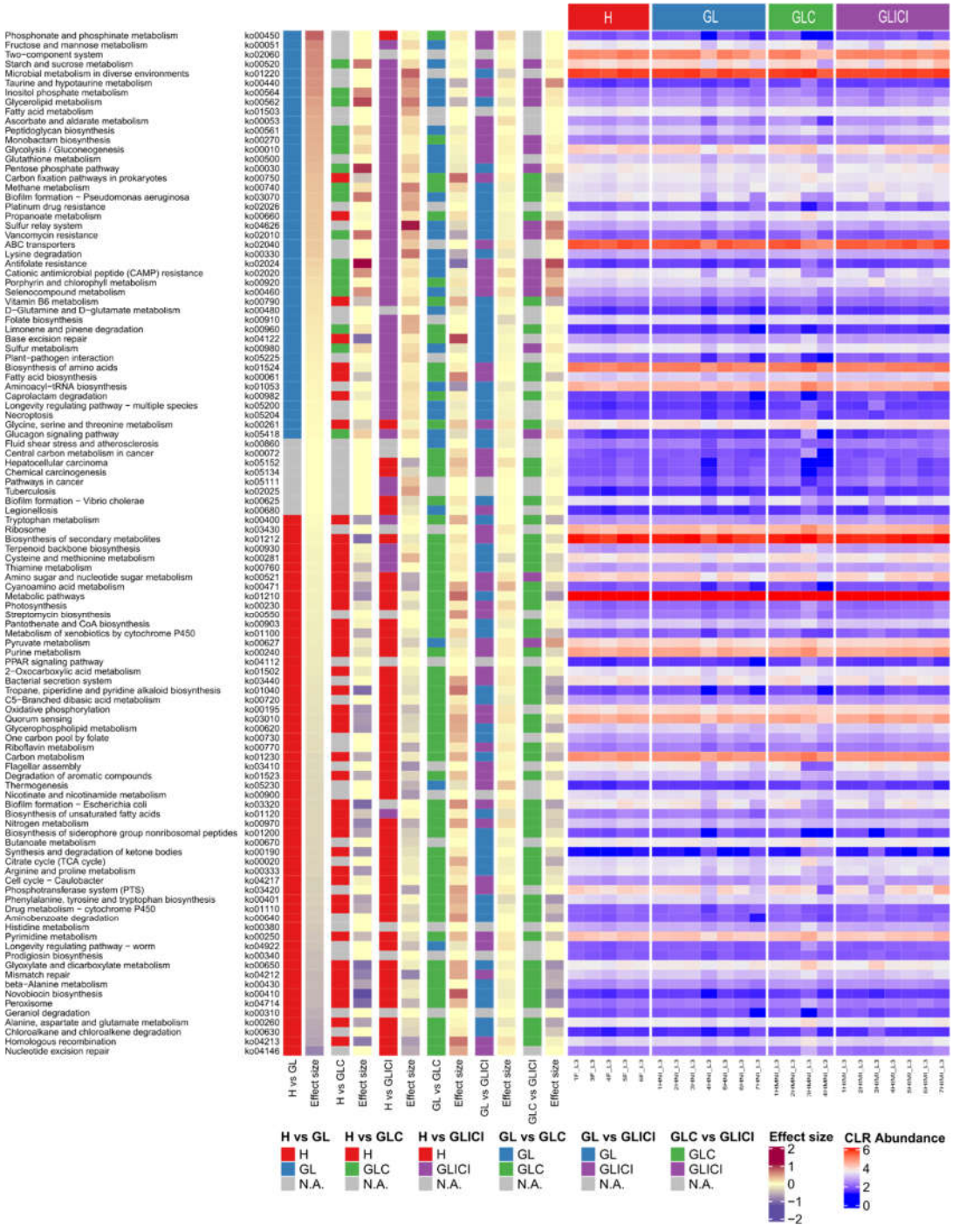


Figure C7: Differentially abundant function pathways among milk samples for spring. Differential abundance was performed on milk sample data using the ADEx2 algorithm. The heatmap illustrates the distribution and abundance of differentially abundant features across feeding types. Corresponding paired group comparisons, effect sizes, and taxonomic classification (phylum level) are displayed as heatmap annotation on the left.

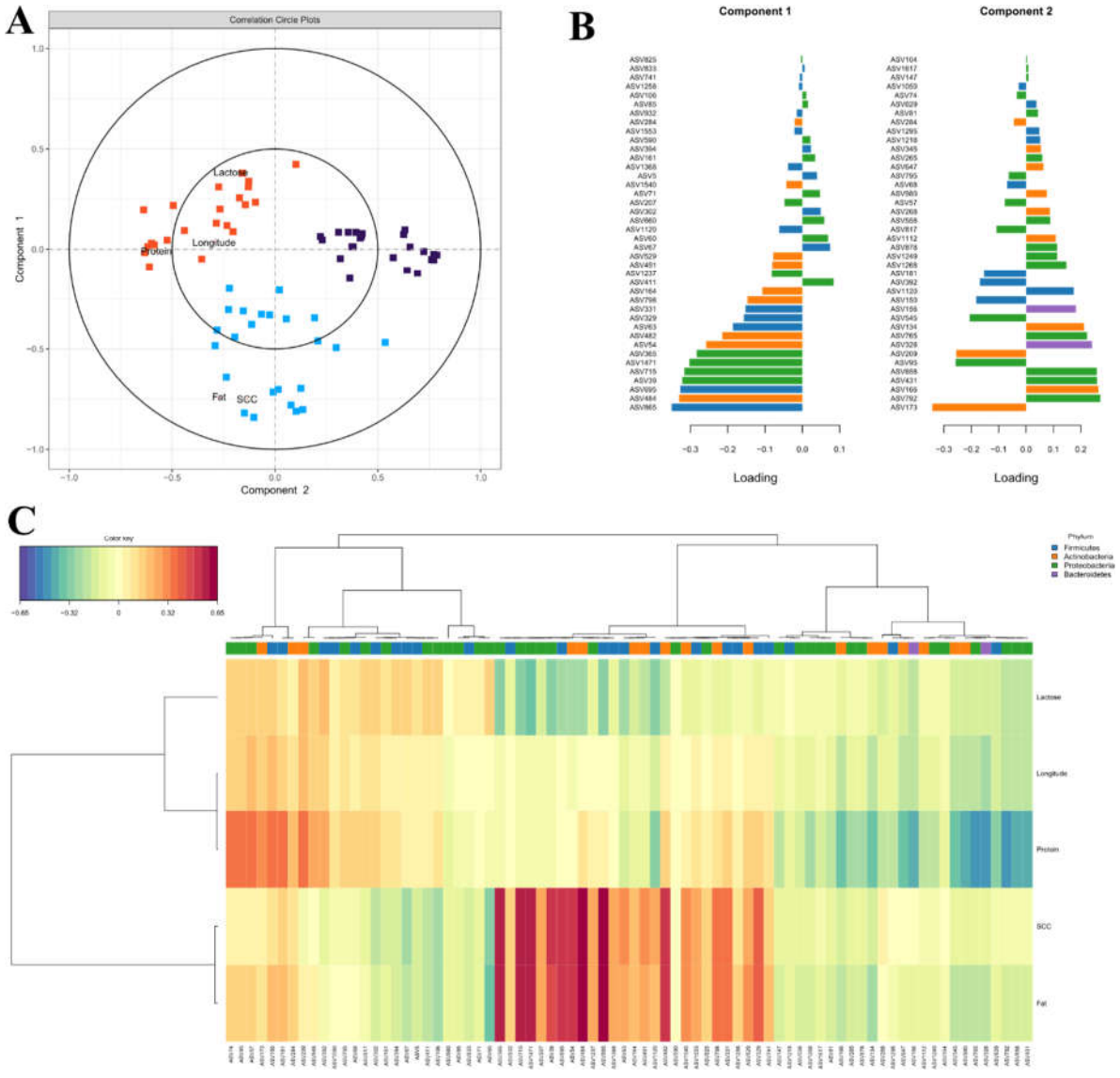


Figure C8: Sparse partial least square regression on milk parameters and ASV. (A) Correlation circle plot based on dimensions 1 and 2. (B) ASV contribution loading plots on the first (left) and second (right) components. (C) Clustered image map of correlations between milk parameters and ASV.

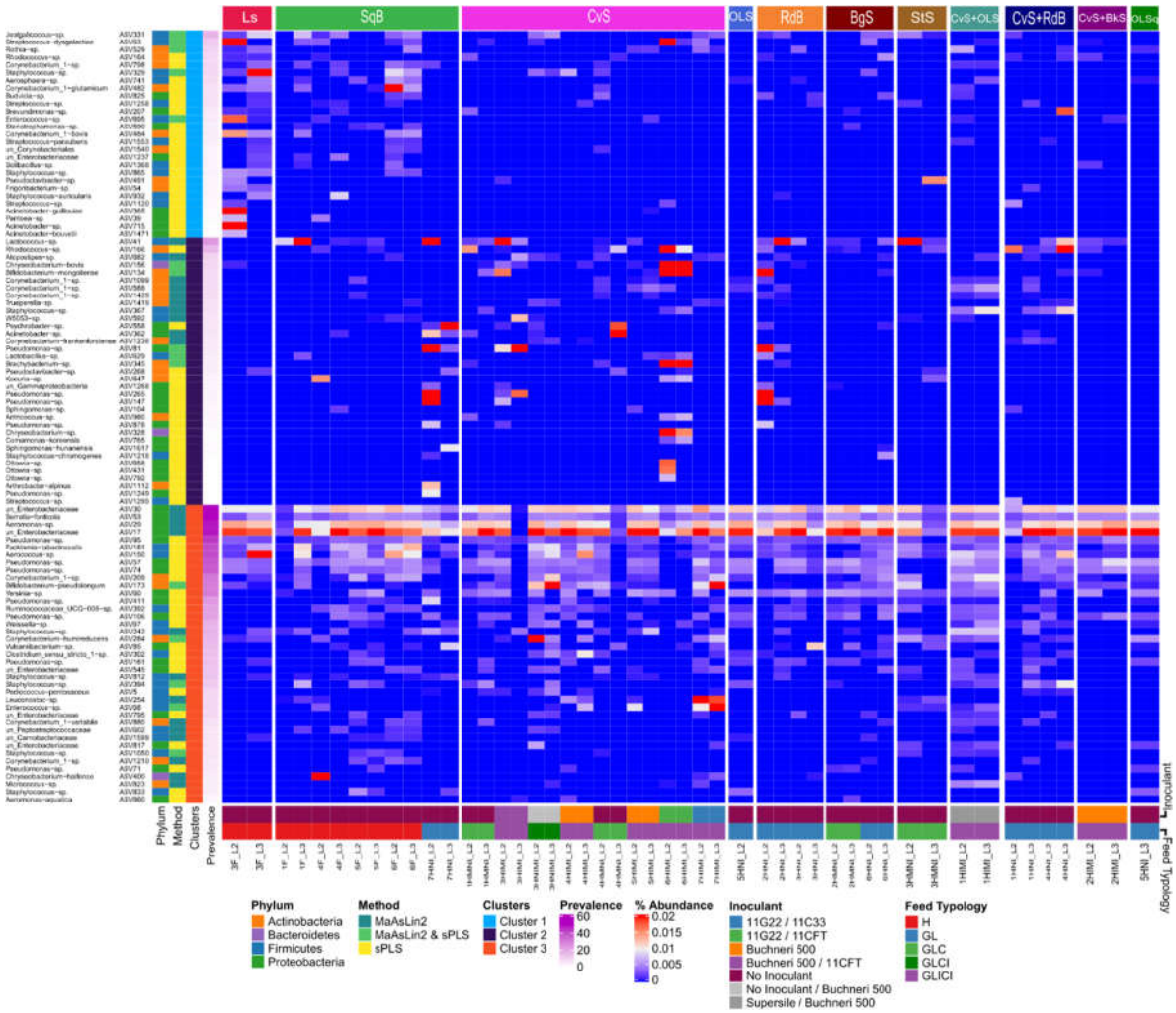
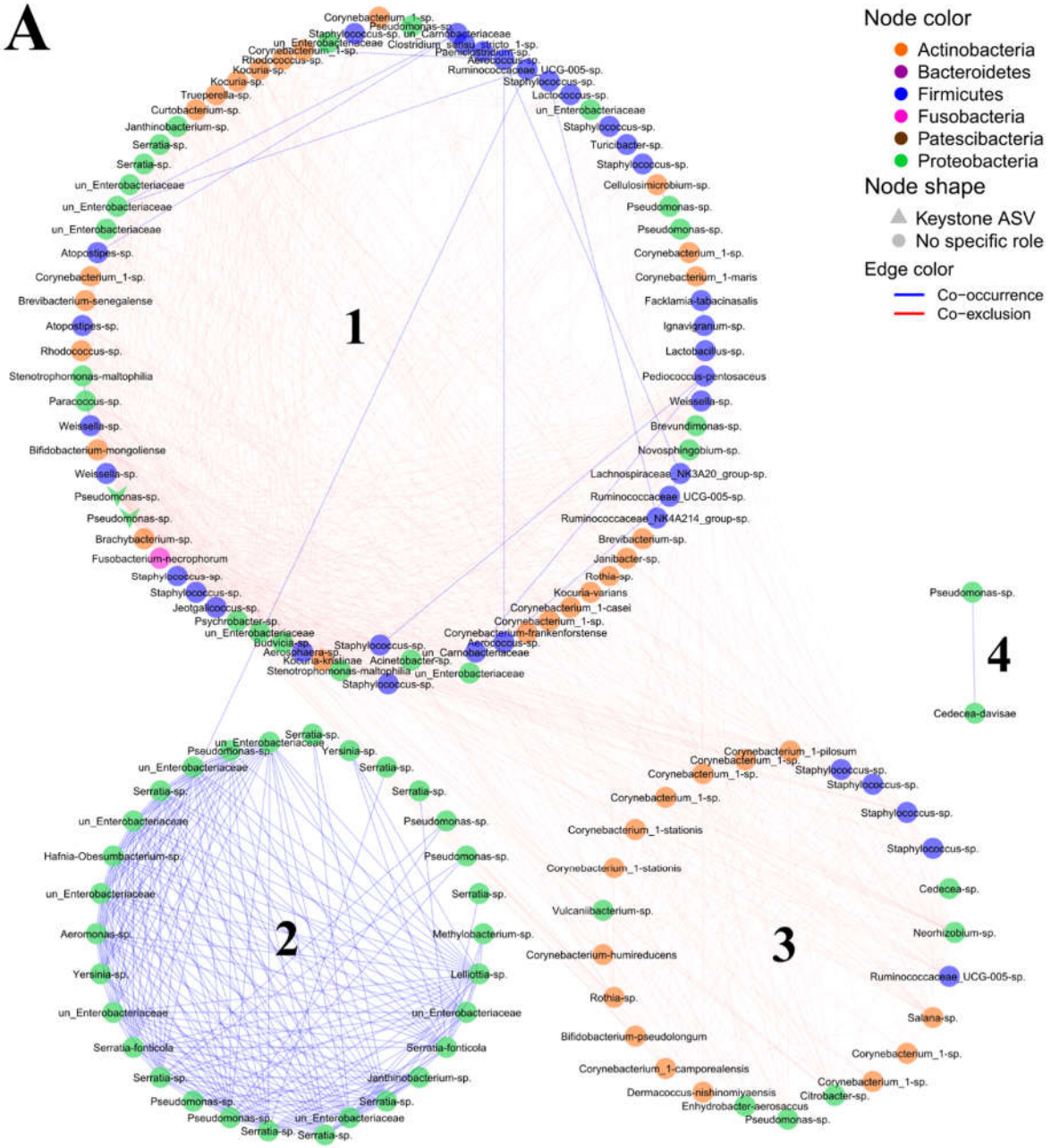
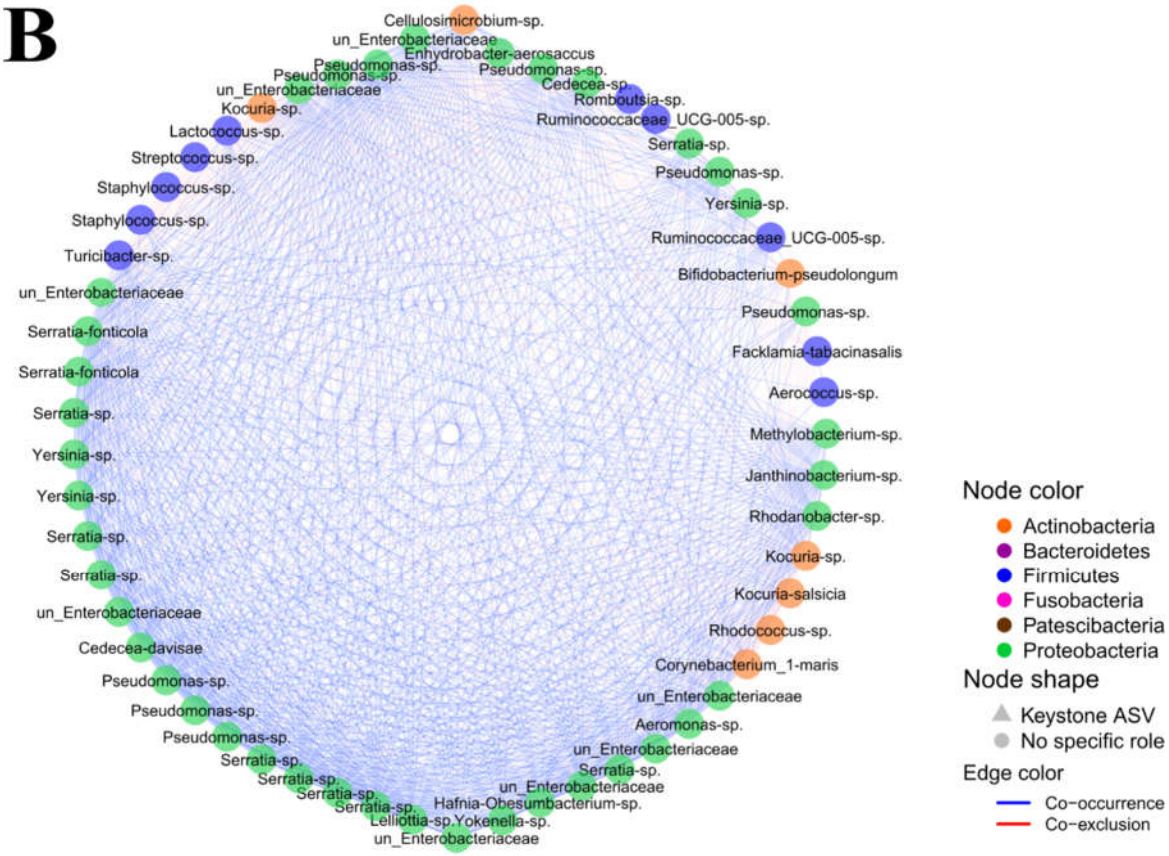


Figure C9: Distribution, prevalence, and abundance of phylotypes significantly correlated with milk parameters. Milk samples are grouped according to the type of forage storage form prevailing in the dairy farm. Accordingly, Ls = loose hay, SqB = square bales, CvS = conventional silo, OLS = oxygen-limiting silo, RdB = round bales, BgS = bag silo, StS = stack silo, CvS+OLS = Forage stored in conventional silo and oxygen-limiting silos on the same farm; similarly for CvS+RdB, CvS+BkS wherein BkS = bunker silo, OLSq where OL and Sq denote OLS and SqB, respectively. The information on inoculants used for ensiling and feed typologies are provided as heatmap annotation on the bottom. Associated clusters determined by sparse partial least square regression are provided as heatmap annotation on the left.


A
Node color
Actinobacteria
Bacteroidetes
Firmicutes
Fusobacteria
Patescibacteria
Proteobacteria
Node shape
Keystone ASV
No specific role
Edge color
Co−occurrence
Co−exclusion
1
2
3
4
Corynebacterium_1-sp.
Pseudomonas-sp.
Staphylococcus-sp.
un_Carnobacteriaceae
un_Enterobacteriaceae
Clostridium_sensu_stricto_1-sp.
Paeniclostridium-sp.
Aerococcus-sp.
Ruminococcaceae_UCG-005-sp.
Staphylococcus-sp.
Lactococcus-sp.
un_Enterobacteriaceae
Staphylococcus-sp.
Turicibacter-sp.
Staphylococcus-sp.
Cellulosimicrobium-sp.
Pseudomonas-sp.
Pseudomonas-sp.
Corynebacterium_1-sp.
Corynebacterium_1-maris
Facklamia-tabacinasalis
Ignavigranum-sp.
Lactobacillus-sp.
Pediococcus-pentosaceus
Weissella-sp.
Brevundimonas-sp.
Novosphingobium-sp.
Lachnospiraceae_NK3A20_group-sp.
Ruminococcaceae_UCG-005-sp.
Ruminococcaceae_NK4A214_group-sp.
Brevibacterium-sp.
Janibacter-sp.
Rothia-sp.
Kocuria-varians
Corynebacterium_1-casei
Corynebacterium_1-sp.
Corynebacterium-frankenforstense
Aerococcus-sp.
un_Carnobacteriaceae
Acinetobacter-sp.
un_Enterobacteriaceae
Staphylococcus-sp.
Budvicia-sp.
Aerosphaera-sp.
Kocuria-kristinae
Stenotrophomonas-maltophilia
Staphylococcus-sp.
Jeotgalicoccus-sp.
Psychrobacter-sp.
un_Enterobacteriaceae
Brachybacterium-sp.
Fusobacterium-necrophorum
Staphylococcus-sp.
Staphylococcus-sp.
Weissella-sp.
Pseudomonas-sp.
Pseudomonas-sp.
Bifidobacterium-mongoliense
Weissella-sp.
Paracoccus-sp.
Stenotrophomonas-maltophilia
Rhodococcus-sp.
Atopostipes-sp.
Brevibacterium-senegalense
Corynebacterium_1-sp.
Atopostipes-sp.
un_Enterobacteriaceae
un_Enterobacteriaceae
un_Enterobacteriaceae
Serratia-sp.
Janthinobacterium-sp.
Curtobacterium-sp.
Trueperella-sp.
Kocuria-sp.
Kocuria-sp.
Rhodococcus-sp.
Corynebacterium_1-sp.
Corynebacterium_1-sp.
un_Enterobacteriaceae
Clostridium_sensu_stricto_1-sp.
Staphylococcus-sp.
un_Carnobacteriaceae
Pseudomonas-sp.
Cedecea-davisae
Serratia-sp.
un_Enterobacteriaceae
Pseudomonas-sp.
Yersinia-sp.
Serratia-sp.
Serratia-sp.
Pseudomonas-sp.
Pseudomonas-sp.
Serratia-sp.
Methylobacterium-sp.
Lelliottia-sp.
un_Enterobacteriaceae
Serratia-fonticola
Janthinobacterium-sp.
Serratia-sp.
Pseudomonas-sp.
Pseudomonas-sp.
un_Enterobacteriaceae
Serratia-sp.
Serratia-sp.
un_Enterobacteriaceae
Hafnia-Obesumbacterium-sp.
un_Enterobacteriaceae
Aeromonas-sp.
Yersinia-sp.
un_Enterobacteriaceae
Serratia-fonticola
Serratia-sp.
Corynebacterium_1-pilosum
Corynebacterium_1-sp.
Corynebacterium_1-sp.
Staphylococcus-sp.
Staphylococcus-sp.
Corynebacterium_1-sp.
Staphylococcus-sp.
Corynebacterium_1-stationis
Staphylococcus-sp.
Corynebacterium_1-stationis
Cedecea-sp.
Vulcanibacterium-sp.
Neorhizobium-sp.
Corynebacterium-humireducens
Ruminococcaceae_UCG-005-sp.
Rothia-sp.
Salana-sp.
Bifidobacterium-pseudolongum
Corynebacterium_1-sp.
Corynebacterium_1-camporealensis
Corynebacterium_1-sp.
Dermacoccus-nishinomiyaensis
Citrobacter-sp.
Enhydrobacter-aerosaccus
Pseudomonas-sp.

B


Cellulosimicrobium-sp.
un_Enterobacteriaceae
Enhydrobacter-aerosaccus
Pseudomonas-sp.
Pseudomonas-sp.
Pseudomonas-sp.
Cedecea-sp.
Romboutsia-sp.
un_Enterobacteriaceae
Ruminococcaceae_UCG-005-sp.
Kocuria-sp.
Serratia-sp.
Lactococcus-sp.
Pseudomonas-sp.
Streptococcus-sp.
Yersinia-sp.
Staphylococcus-sp.
Ruminococcaceae_UCG-005-sp.
Staphylococcus-sp.
Bifidobacterium-pseudolongum
Turicibacter-sp.
Pseudomonas-sp.
un_Enterobacteriaceae
Facklamia-tabacinasalis
Serratia-fonticola
Aerococcus-sp.
Serratia-fonticola
Methylobacterium-sp.
Serratia-sp.
Janthinobacterium-sp.
Yersinia-sp.
Rhodanobacter-sp.
Yersinia-sp.
Kocuria-sp.
Serratia-sp.
Kocuria-salsicia
Serratia-sp.
Rhodococcus-sp.
un_Enterobacteriaceae
Corynebacterium_1-maris
Cedecea-davisae
un_Enterobacteriaceae
Pseudomonas-sp.
Aeromonas-sp.
Pseudomonas-sp.
un_Enterobacteriaceae
Pseudomonas-sp.
Serratia-sp.
Serratia-sp.
un_Enterobacteriaceae
Serratia-sp.
Hafnia-Obesumbacterium-sp.
Serratia-sp.
Lelliottia-sp.Yokenella-sp.
un_Enterobacteriaceae
Node color
Actinobacteria
Bacteroidetes
Firmicutes
Fusobacteria
Patescibacteria
Proteobacteria
Node shape
Keystone ASV
No specific role
Edge color
Co-occurrence
Co-exclusion

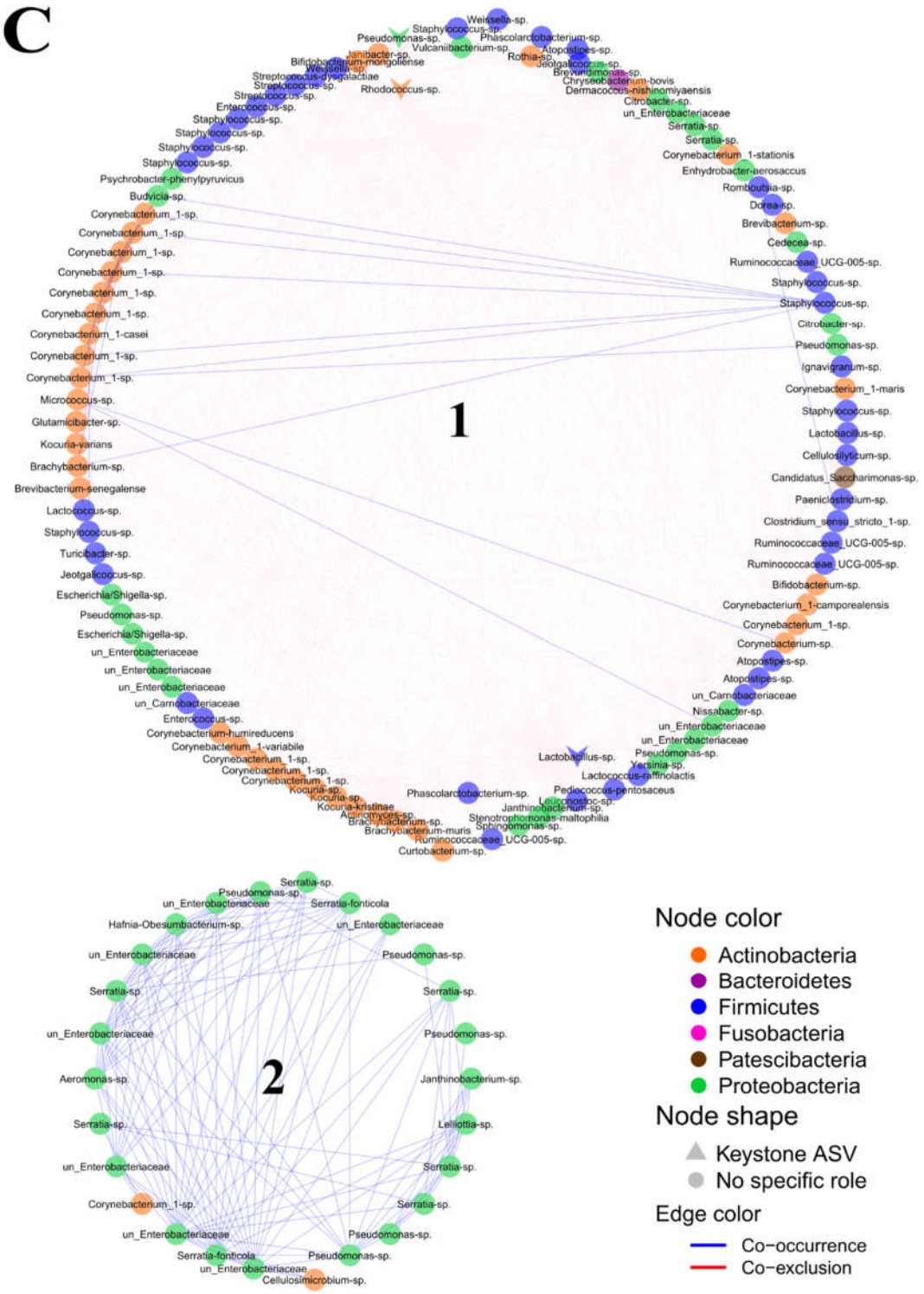


Figure C10: Molecular networks of milk taxa occurrence. Co-occurring and co-excluding phylotypes of milk samples associated with the GL (A), GLC (B), and GLICI feeding typologies(C). Edges comprise blue lines representing co-occurrence relationships, and red lines representing the co-exclusions. Modules are numbered when occurring more than once in a network.

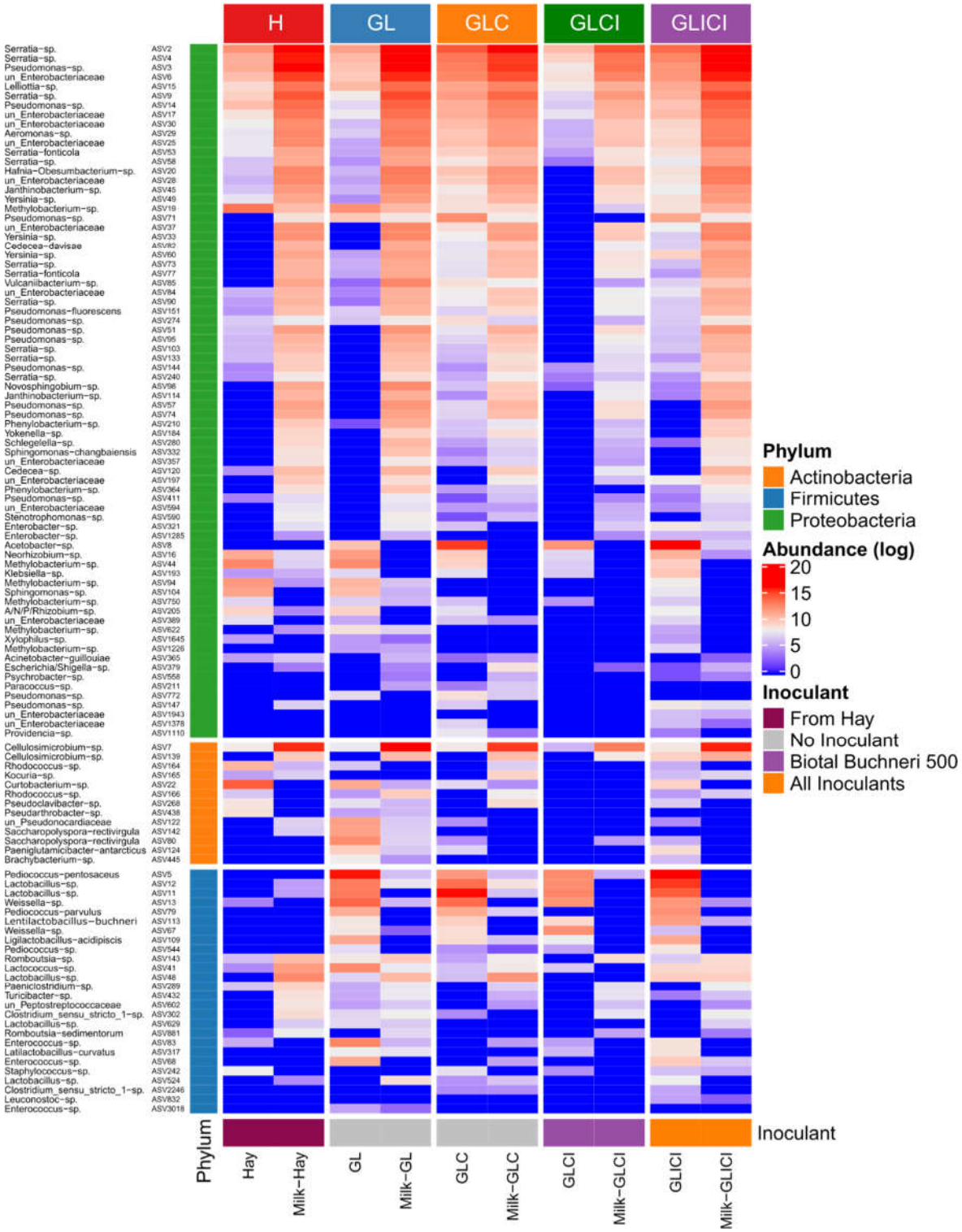


Figure C11: Distribution, prevalence, and abundance of shared phylotypes between forage types and milk samples.