Simon Deusch

METAOMIC STUDIES OF THE DIETARY IMPACT ON THE STRUCTURAL AND FUNCTIONAL DIVERSITY OF THE RUMEN MICROBIOME

2018

METAOMIC STUDIES OF THE DIETARY IMPACT ON THE STRUCTURAL AND FUNCTIONAL DIVERSITY OF THE RUMEN MICROBIOME

FACULTY OF AGRICULTURAL SCIENCES

Institute of Animal Science - University of Hohenheim

Feed-Gut Microbiota Interaction

Jun.-Prof. Dr. Jana Seifert

METAOMIC STUDIES OF THE DIETARY IMPACT ON THE STRUCTURAL AND FUNCTIONAL DIVERSITY OF THE RUMEN MICROBIOME

INAUGURAL DISSERTATION

submitted in fulfillment of the requirements for the degree

"Doktor der Agrarwissenschaften" (Dr. sc. agr. / PhD in Agricultural Sciences)

to the

Faculty of Agricultural Sciences

presented by

Simon Deusch (Diploma Agricultural Biologist)

born on 28.08.1985 in Sankt Georgen im Schwarzwald, Germany

2017

This work was funded by a doctoral fellowship of the Faculty of Agricultural Science, University of Hohenheim which is gratefully acknowledged.

The presented studies were realized at the Institutes of Animal Science and Chemistry and at the Agricultural Experiment Station Meiereihof of the University of Hohenheim in Stuttgart, Germany. The experimental work was performed in the time between October 2013 and May 2016 under the supervision of Jun.-Prof. Dr. Jana Seifert (University of Hohenheim, Head of the Junior-Stiftungsprofessur "Feed-Gut Microbiota Interaction").

The Faculty of Agricultural Sciences at the University of Hohenheim accepted this thesis as a doctoral thesis (dissertation) in fulfillment of the regulations to acquire the doctoral degree "Doktor der Agrarwissenschaften" on September 29, 2017.

Dean of the Faculty of Agricultural Sciences:	Prof. Dr. Ralf T. Vögele

---

Date of the oral examination:	October 13, 2017
Chairperson of the oral examination:	Prof. Dr. Jörn Bennewitz

---

Dissertation supervisor:	Jun.-Prof. Dr. Jana Seifert
External reviewer:	Prof. Dr. Qendrim Zebeli
Third reviewer:	Prof. Dr. Markus Rodehutscord

MY FAMILY

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

#	Number
%	Percentage
~	Approximately
°C	Degree Celsius
µl	Microliter
µm	Micrometer
µM	Micromolar
16S rRNA gene	16S ribosomal ribonucleic acid gene
ABC transporter	ATP-binding cassette transporters
ANOSIM	Analysis of similarities
ATP	Adenosine triphosphate
BCFA	Branched-chain fatty acid
bp	Base pairs
CAZymes	Carbohydrate-active enzymes
CBM	Carbohydrate-binding module
cDNA	Complementary deoxyribonucleic acid
CE	Carbohydrate esterase
COG	Clusters of Orthologous Groups
CP	Crude protein
CS	Corn silage-based diet
Da	Dalton
Delta Cn	Delta correlation
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphates
e.g.	For example
ESI	Electrospray ionization
FDR	False discovery rate
FID	Free induction decay
g	Gram
Gb	Gigabase
GC-MS	Gas chromatography-mass spectrometry
GH	Glycoside hydrolases

GIT	Gastrointestinal tract
GS	Grass silage-based diet
GT	Glycosyltransferases
h	Hour
H	Grass hay-based diet
HCD	Higher-energy collisional dissociation
HPLC	High-performance liquid chromatography
Hz	Hertz
KEGG	Kyoto Encyclopedia of Genes and Genomes
kg	Kilogram
KO	KEGG Orthology
L	Liter
LC	Liquid chromatography
LC-MS/MS	Liquid chromatography tandem-mass spectrometry
LFQ	Label-free quantification value
LP	Liquid phase
M	Molar
m/z	Mass to charge ratio
mA	Milliampere
Maldi-ToF	Matrix-assisted laser desorption/ionization
MCP	Microbial crude protein
MCR	Methyl-coenzyme M reductase
MHz	Megahertz
min	Minute
ml	Milliliter
mm	Millimeter
mM	Millimolar
mRNA	Messenger ribonucleic acid
ms	Millisecond
mV	Millivolt
NADH	Nicotinamide adenine dinucleotide
NH_3	Ammonia
nl/min	Nanoliter per minute

NMDS	Non-metric multidimensional scaling
NMR	Nuclear magnetic resonance
nt	Nucleotide
Ø	Average
ORF	Open reading frame
OTU	Operational taxonomic unit
P	Probability value
p. m.	Post meridiem
PCR	Polymerase chain reaction
pH	Potential of hydrogen
PMSF	Phenylmethylsulfonyl fluoride
ppm	Parts per million
PSM	Peptide-spectrum match
r	Resolution
RDP	Ribosomal Database Project
RF	Rumen fluid
RNA	Ribonucleic acid
rpm	Revolutions per minute
s	Second
SCFA	Short-chain fatt acid
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEM	Standard error of the mean
SLH	S-layer homology
SP	Solid phase
spp.	Species
TCA	Trichloroacetic acid
TSP	Trimethylsilylpropanoic acid
U	Enzyme unit
σ	Standard deviation

CHAPTER I

INTRODUCTION

1. INTRODUCTION

1.1 Ruminants

Ruminants, earth's prevalent herbivores (Hungate, 1975) harbor the ability to convert the solar energy stored in structural plant polysaccharides such as cellulose, indigestible for humans, into valuable animal protein in form of dairy and meat products (Haug et al., 2007; Quann et al., 2015). Thereby, ruminant livestock constitutes a significant source of food and also livelihood pivotal particularly in developing countries (McDermott et al., 1999). Globally, ruminant livestock accounts for up to 3.6 billion farm animals comprising about 1.4 billion of *Bos taurus* and *B. indicus* species (Hackmann and Spain, 2010). Moreover, photosynthesis-derived cellulose in plant cell walls is the most abundant polymer and the most renewable source of carbon on earth (Enquist et al., 2003; Pauly and Keegstra, 2008). A steadily growing world population is expected to lead to an increasing demand for animal-based food products (McMichael et al., 2007) which entails an undesired side effect since livestock production causes about 14.5% of the total anthropogenic greenhouse gas emissions (Gerber et al., 2013) and thus, plays an important role in climate change (Ripple et al., 2014). Meat cattle and dairy cows are the largest contributors accounting for about 65% of the total livestock sector emissions of which above 44% are caused solely by enteric methane formation (Gerber et al., 2013). Processes underlying efficiency in ruminant production respectively plant fiber degradation and methane formation are driven by a complex bionetwork consisting of archaea, bacteria, fungi and protists residing in the strictly anaerobic rumen (Hungate, 1966; Mackie, 2000).

Aside from reticulum, omasum and abomasum, the rumen is the largest part (80%) of a highly specialized, four-chambered stomach of the ruminant gastrointestinal tract. The functions of reticulum and rumen are closely interrelated and thus, are often termed rumino-reticulum. Comparable to a continuous bioreactor the animal feed is digested mainly in the rumino-reticulum via chemical and mechanical degradation. The reticulum plays an important role in separating feed components wherein undigested matter is redirected to the rumen and already digested, finer feed particles (< 2 mm) are further passed on to the omasum for dehydration. Contractions of the reticulum enable regurgitation of coarse feed particles for repetitive mechanical degradation by chewing, known as rumination. Likewise, the rumen is continuously in motion mixing the forage and helping the flow of feed particles. With a volume of up to 100 liters, the rumen of cattle contains tremendous amounts (60 to 120 kg) of low quality lignocellulose matter ingested by the animals that like humans, lack the hydrolytic

enzymes to degrade the respective fibrous material (Mackie et al., 2001; Moran, 2005; Puniya et al., 2015). Besides, the rumen epithelial wall contains no glandular cells and hence, no digestive juices are secreted.

Therefore ruminants host a symbiotic microbiome whose members interact synergistically to extract energy from the animal feed in producing highly active lignocellulolytic enzymes resulting in the production of short-chain fatty acids and microbial biomass (Figure 1) which serve as the main sources of energy and amino acids for the animal (Hungate, 1966). Short-chain fatty acids and microbial cells constitute about 70 to 85% of nutrients for the animals (Bergman, 1990; van Houtert, 1993). Moreover, B and K vitamins are produced and toxic compounds such as myco- and phytotoxins are neutralized supporting the animal's health (Mackie et al., 2001). In exchange, ruminants provide a stable and suitable environment for the microbes which is characterized by pH values ranging between 5.5 and 6.9 and high bicarbonate-phosphate-based buffering capacities from vast amounts of saliva released (70 to 150 L per day). Further features are a low redox potential (-350 to -400 mV) favoring anaerobic metabolism, an osmolality of 250 to 350 mOsmol/kg^{-1} and constant temperatures of 38 to 41°C (Puniya et al., 2015). Continuous feed intake with a high throughput of digesta (120 to 200 L daily) and mastication of feed by the animals grants sufficient nutrient supply for the microorganisms. Additionally, indigestible matter, microbial cells as well as soluble fermentation end products are removed constantly (Mackie et al., 2001) which is of great importance to prevent noxious acidification of the rumen due to accumulation of short-chain fatty acids.

The primary short-chain fatty acids produced by microbial fermentation of plant polysaccharides are ketogenic acetate and butyrate as well as glucogenic propionate which are absorbed passively through the rumen epithelium and transported to the liver for metabolic processing yielding energy in form of adenosine triphosphate (ATP) used in the animal's skeletal muscles, adipose tissue and the mammary gland (Figure 1). Butyrate itself yields most ATP but in the rumen wall the largest part of butyrate is converted to β-hydroxybutyrate, a ketone representing an important energy source for epithelial and other cells (Sejrsen et al., 2006). Regarding milk production, acetate and butyrate-derived ketones are involved in the generation of milk fat whereas most propionate is converted to glucose and rather affects lactose content and overall milk yield (Miettinen and Huhtanen, 1996).

Next to degrading plant fibers, rumen microorganisms are even capable of converting low-quality crude protein of animal diets and especially non-protein nitrogen, mainly urea, into microbial protein (Figure 1). This is of high value for the host since it contains

indispensable amino acids like methionine and lysine that cannot be synthesized by the animal itself. Microbial protein synthesis requires nitrogen in form of ammonia obtained from the breakdown of urea and deamination of dietary amino acids. Ammonia concentration is regulated by absorption through the rumen wall (Wallace et al., 1997) since higher concentrations are detrimental to the animal and the microorganisms. In the liver, ammonia is recycled to urea and part of it is brought back to the rumen via saliva and epithelial secretion once more providing a non-protein nitrogen source necessary for microbial protein synthesis (Lapierre and Lobley, 2001). The abomasum, the last part of the ruminant's stomach, is connected to the small intestine and resembles a true stomach similar to that of monogastric animals with a low pH value of about 2. Acidity and protease activity in the abomasum carry out the breakdown of microbial cells flushed in with the digested feed matter enabling initial peptic digestion and subsequently tryptic digestion in the following small intestine with absorption of mainly microbial but as well dietary amino acids (Mackie et al., 2001).

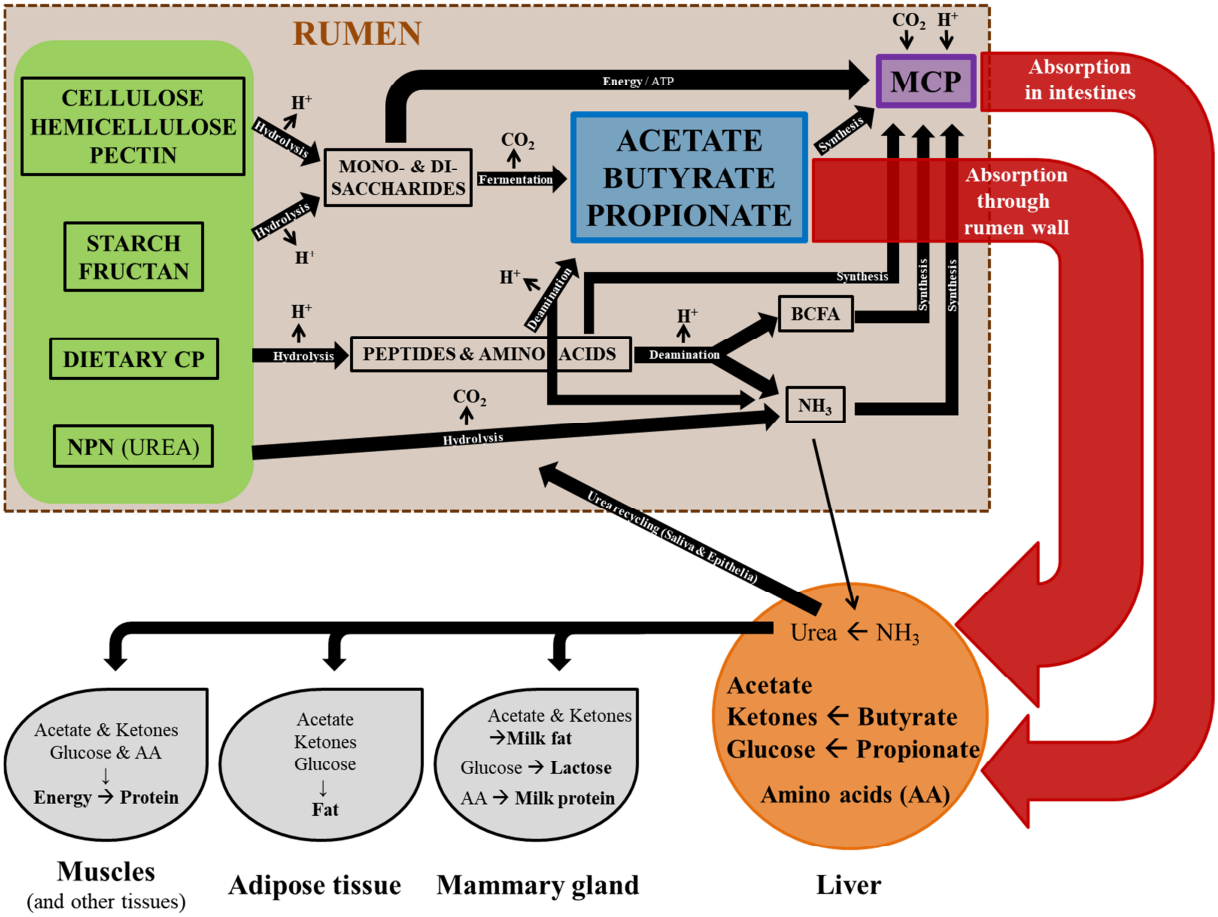

Figure 1| Overview of protein and carbohydrate metabolism in the rumen ecosystem. BCFA, branched-chain fatty acids; MCP/CP, microbial/crude protein; NH_3, ammonia.

Other byproducts of microbial fermentation such as carbon dioxide, hydrogen and formate but as well some acetate, methanol and methylamines are further utilized by methanogenic archaea. Thereby methane is formed which along with excess of carbon dioxide cannot be absorbed by the host animals and is eructed to the atmosphere implementing an energy loss from feed (Johnson and Ward, 1996; Morgavi et al., 2010; Mizrahi, 2011).

First investigations of ruminants date back to the beginning of the 20th century. In the 1950s, Robert E. Hungate initiated the exploration of the complex rumen ecosystem focusing on its microbes revolutionizing anaerobic cultivation techniques and promoting awareness and understanding of rumen microbiology (Hungate, 1947; 1950; Hungate, 1966). Ever since, nutritionists, microbiologists and physiologists tried to improve production efficiency, animal health and methane mitigation strategies by altering the rumen microbiome. The immense amount of various fiber-degrading enzymes produced in the rumen attracted other scientific disciplines aiming at the discovery of new tools relevant for biofuel production. Even today, investigations of the structural and functional diversity of the complex rumen microbiome and its interactions with the host and within the specific ecological niches of the rumen ecosystem are of great interest for animal production, biotechnology and climatology.

1.2 Rumen Microbiome

To obtain energy, the microbial communities of the rumen carry out the hydrolysis of structural plant polysaccharides like cellulose, hemicellulose, pectin and storage polysaccharides such as starch and fructan into monomeric and dimeric sugars that are further fermented yielding short-chain fatty acids, carbon dioxide and hydrogen (Figure 1). In addition, dietary crude protein is hydrolyzed to peptides and amino acids, which are deaminated once again resulting in the production of short-chain fatty acids, carbon dioxide and hydrogen but as well branched-chain fatty acids and ammonia. Hydrogen, some carbon dioxide and ammonia along with peptides, amino acids and small parts of the branched- and short-chain fatty acids as well as other compounds and energy from the fermentation of carbohydrates are utilized for microbial growth respectively protein synthesis (Figure 1). In summary, major end products of the microbiome's metabolism are microbial cells, acetate, butyrate and propionate absorbed by the host as well as gaseous carbon dioxide and methane, which are released to the environment.

However, degradation of the entirety of organic matter ingested by the animals cannot be achieved by a single species and thus, requires collaboration and interaction of a succession of many microorganisms to employ the whole spectrum of functional capabilities

available (Bladen et al., 1961). Synergistically, polymeric carbohydrates are converted into soluble sugars that are further metabolized by other species of the bionetwork, which in return built branched-chain fatty acids, vitamins, ammonia, carbon dioxide as well as other cofactors required by fiber-degrading species (Bryant, 1973). Further microbial interactions for example exist between lactate or succinate producers and species that utilize the respective compounds explaining why the widespread fermentation products lactate and succinate usually do not accumulate in the rumen of healthy animals.

One milliliter of rumen fluid may contain about 10^{12} microbes outnumbering the total people on earth by a factor of above 100 (Prescott, Harley & Klein, 2005). The rumen microbiome exhibits a high biodiversity comprising microorganisms from all domains of life such as eukaryotic, anaerobic fungi and unicellular protists as well as prokaryotic archaea and bacteria. Additionally, viruses, more precisely archaeaphages and bacteriophages belong to the rumen ecosystem. Bacteria are the most numerous and diverse microorganisms in the rumen appearing with up to 10^{11} cells in each gram of rumen matter (Mackie, 2000) with above 200 species (McSweeney et al., 2005) and represent about 50 to 75% of the overall microbial communities (Minato et al., 1966). Archaea are less abundant ranging between 10^7 to 10^9 cells per gram contributing to only a small part of about 1 to 3% of the microbial mass in the rumen (Janssen and Kirs, 2008; Wright and Klieve, 2011). The much larger cells of ciliate protozoa, distributed over 25 genera, appear in comparably smaller numbers ranging between 10^4 and 10^6 cells per gram (Puniya et al., 2015) but can constitute up to 50% of microbial biomass (Newbold et al., 2015). Spores of anaerobic fungi range between 10^3 and 10^5 per gram including six genera (Mackie et al., 2001) whereas phages occur with particle numbers of about 10^8 to 10^9 per gram of rumen fluid (Klieve and Bauchop, 1988).

Besides being most abundant and diverse, bacteria are also the most metabolically active species in the rumen being mainly involved in the degradation of plant fibers (Brulc et al., 2009) as reflected by the fact that up to 75% of the total rumen bacteria live attached to plant particles (Koike et al., 2003). The solid-associated population is further separated into loosely- and tightly-adhered species. Loosely-attached bacteria can be washed off the rumen matter easily (McAllister et al., 1994) whereas tightly-adherent species remain attached but can be removed by chemical or physical treatments such as methylcellulose or chilling (Trabalza-Marinucci et al., 2006). Both solid-associated subpopulations have been estimated to contribute to a total of about 80 to 90% of the endoglucanase activity, 70% of amylase activity (Minato et al., 1966) and 75% of the protease activity (Brock et al., 1982) in the rumen. The liquid-associated bacterial population constitutes up to 30% but plays a minor role

in fiber degradation and is composed of free-floating, planktonic species that feed on soluble components in the rumen fluid and bacteria that detached from the plant particles or live adhered to protozoa and fungal sporangia (Miron et al., 2001). Aside from the extramural bacteria that constitute 99% of rumen bacteria only a minority of about 1% is associated with the rumen epithelium (Czerkawski, 1986; Dehority, 2003). They are believed to remove oxygen and convert urea into carbon dioxide and ammonia used by others for microbial protein synthesis (McAllister et al., 1994).

Apart from allocated space, bacteria can be classified according to their broad functional potential, as there are amongst others cellulolytic, hemicellulolytic, pectinolytic, amylolytic, saccharolytic, proteolytic, lipolytic and ureolytic specialists. Starch and sugar degraders usually constitute the largest part of the ruminal bacterial population and are of great relevance since diets for high-producing ruminants usually contain large amounts of readily fermentable starch and sugars. Despite their undoubtable significance, bacteria specialized for fiber degradation are typically less present, even when straw is fed exclusively (Puniya et al., 2015). Some bacterial species depend on specific substrates such as the major cellulose degrader *Fibrobacter succinogenes* that is restricted to cellulose, hemicellulose and glucose as growth substrates (Hungate, 1950). *Ruminobacter amylophilus* requires starch and maltose (Anderson, 1995) whereas others like *Prevotella ruminicola* are capable of utilizing a broad range of substrates including peptides, proteins, monosaccharides and plant polysaccharides (Miyazaki et al., 1997; Matsui et al., 2000; Purushe et al., 2010). Substrates and fermentation end products of some well-described rumen bacteria are given in Table 1.

Overall, the bacterial community composition of high-producing ruminants is dominated by the phylum of Bacteroidetes, in particular the family of Prevotellaceae (Jami and Mizrahi, 2012; Jami et al., 2014) that includes species with versatile metabolic capabilities like *Prevotella ruminicola*, *P. brevis*, *P. bryantii* and *P. albensis*. The more diverse, Gram-positive Firmicutes constitute the second most abundant phylum and comprise many fiber-degrading bacteria within the families of Clostridiaceae, Eubacteriaceae, Lachnospiraceae and Ruminococcaceae but cover as well bacteria of the class of Negativicutes including the families of Acidaminococcaceae, Selenomonadaceae and Veillonellaceae. The latter ones are represented by species with rather limited abilities in hydrolyzing structural carbohydrates such as the succinate utilizers *Succiniclasticum ruminis* (van Gylswyk, 1995) and *Phascolarctobacterium succinatutens* (Watanabe et al., 2012) as well as the saccharolytic *Selenomonas ruminantium* (Dehority, 2003), the lipolytic *Anaerovibrio lipolyticus* (Prins et al., 1975) and the ureolytic, lactate-fermenting

Megasphaera elsdenii (Counotte et al., 1981).

Table 1| The fermentation characteristics of predominant ruminal prokaryotes. Abbreviations are as follows: A, acetate; AA, amino acids; B, butyrate; Br, branched-chain volatile fatty acids; CE, cellulose; CH_4, methane; CO_2, carbon dioxide; DX, dextrins; E, ethanol; F, formate; GL, glycerol; H_2, hydrogen; HC, hemicellulose; L, lactate; M, methanol; MA, methylamines; OA, organic acids; P, propionate; PC, pectin; S, succinate; ST, starch; SU, sugars; XY, xylans. Adapted from (Russell and Rychlik, 2001).

Phylum	**Species**	**Main substrates**	**End products**
Bacteroidetes	*Prevotella albensis*	ST, PC, XY, SU	S, A, F, P
Bacteroidetes	*Prevotella brevis*	ST, PC, XY, SU	S, A, F, P
Bacteroidetes	*Prevotella bryantii*	ST, PC, XY, SU	S, A, F, P
Bacteroidetes	*Prevotella ruminicola*	ST, PC, XY, SU	S, A, F, P
Fibrobacteres	*Fibrobacter succinogenes*	CE	S, F, A
Firmicutes	*Anaerovibrio lipolytica*	GL, SU	A, S, P
Firmicutes	*Butyrivibrio fibrisolvens*	ST, CE, HC, PC, SU	B, F, A, H_2
Firmicutes	*Clostridium aminophilum*	AA	A, B
Firmicutes	*Clostridium sticklandii*	AA	A, Br, B, P
Firmicutes	*Eubacterium ruminantium*	HC, DX, SU	A, F, B, L
Firmicutes	*Lachnospira multiparus*	PC, SU	L, A, F, H_2
Firmicutes	*Megasphaera elsdenii*	L, SU	P, A, B, Br, H_2
Firmicutes	*Phascolarctobacterium succinatutens*	S	P
Firmicutes	*Peptostreptococcus anaerobius*	AA	Br, A
Firmicutes	*Ruminococcus albus*	CE, HC	A, F, E, H_2
Firmicutes	*Ruminococcus flavefaciens*	CE, HC	S, F, A, H_2
Firmicutes	*Selenomonas ruminantium*	ST, DX, SU, L, S	L, A, P, B, F, H_2
Firmicutes	*Streptococcus bovis*	ST, SU	L, A, F, E
Firmicutes	*Succiniclasticum ruminis*	S	P
Proteobacteria	*Ruminobacter amylophilus*	ST	S, F, A, E
Proteobacteria	*Succinomonas amylolytica*	ST	S, A, P
Proteobacteria	*Succinivibrio dextrinosolvens*	PC, DX, SU	S, A, F, L
Proteobacteria	*Wolinella succinogenes*	OA, H_2, F	S
Euryarchaeota	*Methanobrevibacter ruminantium*	H_2, CO_2, F	CH_4
Euryarchaeota	*Methanomassiliicoccus* spp.	H_2, M, MA	CH_4

Fiber-digesters of the family of Clostridiaceae are *Clostridium cellobioparum*, *C. longisporum* and *C. lochheadii* (Vos et al., 2011) while the family of Eubacteriaceae is represented by the cellulolytic *Eubacterium cellulosolvens* (Prins et al., 1972) but as well by hemicellulolytic bacteria such as *E. xylanophilum* and *E. uniformis* (Van Gylswyk and Van der Toorn, 1985). The Lachnospiraceae family comprises the cellulolytic *Butyrivibrio fibrisolvens* (Bryant and Small, 1956) and pectinolytic *Lachnospira multiparus* (Duskova and Marounek, 2001). The major cellulose degraders in the rumen are *Ruminococcus flavefaciens* (Latham and Wolin, 1977) and *R. albus* (Suen et al., 2011) of the Ruminococcaceae family but as well the Gram-negative *Fibrobacter succinogenes* (Shinkai et al., 2009) of the Fibrobacteres phylum. The phylum of Proteobacteria includes rather starch and sugar fermenting organisms such as *Succinivibrio dextrinosolvens* (O'Herrin and Kenealy, 1993), *S. amylolytica* (Bryant et al., 1958) and *Ruminobacter amylophilus* (Stackebrandt and Hippe, 1986), all members of the Succinivibrionaceae family.

The less numerous ruminal archaea are mostly methanogenic species of the orders of Methanobacteriales, Methanomicrobiales, Methanosarcinales and Methanomassiliicoccales all belonging to the phylum of Euryarchaeota (Puniya et al., 2015; St-Pierre et al., 2015). The by far most abundant archaeal genus is the hydrogenotrophic *Methanobrevibacter* (Janssen and Kirs, 2008; Leahy et al., 2013; St-Pierre and Wright, 2013; Seedorf et al., 2015; Danielsson et al., 2017) of the Methanobacteriales order represented by the species *Methanobrevibacter ruminantium* and *M. gottschalkii*. Other genera of the order of Methanobacteriales present in the rumen are *Methanobacterium* and *Methanosphaera*. The order of Methanomicrobiales comprises two families less frequent in the rumen, the likewise hydrogenotrophic Methanomicrobiaceae (Oren, 2014a) and Methanospirillaceae (Oren, 2014b). Hydrogenotrophic methanogens are generally prevailing in the rumen and use fermentation products of other species in particular hydrogen, carbon dioxide and formate for ATP synthesis resulting in production of methane. Formate, a fermentation intermediate (Hungate et al., 1970), thereby is also oxidized to hydrogen and carbon dioxide. Hydrogenotrophic methanogens live in a synergistic relationship with bacteria that produce hydrogen during plant cell wall degradation and some are endo- and ectosymbiotically associated with protozoa (Lange et al., 2005). Up to 37% of the ruminal methane emissions are formed by the symbiotic relationship of protozoa and methanogens in which protozoa transfer hydrogen directly to attached methanogens which use the hydrogen as electron donor for reducing carbon dioxide to methane during their energy production (Wright and Klieve, 2011). In this, the interspecies hydrogen transfer minimizes the partial pressure of hydrogen which is

important for maintaining the redox potential and pH of the rumen ensuring the function of other fermenting species (Lange et al., 2005) and influencing the array of fermentation products as well as overall fermentation rate (van Lingen et al., 2016). Other highly abundant methanogens of the methylotrophic order of Methanomassiliicoccales of the class of Thermoplasmata obligately utilize substrates such as methylamines and methanol (Seedorf et al., 2015; Sollinger et al., 2016). This applies as well to the order of Methanosarcinales that farther includes acetoclastic species like *Methanosarcina barkeri* and *M. acetivorans* (Maeder et al., 2006). In the rumen acetoclastic methanogenesis plays a minor role since the growth rate of acetate-utilizing archaea is lower than the passage rate of rumen digesta (Janssen and Kirs, 2008).

Eukaryotic protozoa were the first rumen microorganisms to be observed (Gruby and Delafond, 1843) probably due to their conspicuously large (20 to 200µm) and highly motile cells. The vast majority of rumen protozoa are ciliates but include as well a few smaller flagellates, which utilize only soluble nutrients and thus, are less important for overall ruminal fermentation (Puniya et al., 2015). Ciliated protozoa are capable of engulfing comparably large molecules such as starch granules, small plant particles and bacterial cells which constitute their major source of protein (Williams and Coleman, 1992). Structural plant polysaccharides are actively digested within the protozoal cells and protozoa are able to store and utilize large amounts of starch. This regulates the rate of carbohydrate fermentation in slowing down bacterial short-chain fatty acid production, which might prevent acidification of the rumen, especially if vast amounts of starch are given in the diet (Puniya et al., 2015). Protozoal fermentation products are similar to those of bacteria including the formation of hydrogen, which is directly utilized by associated methanogens. Protein utilization of protozoa is generally inefficient resulting in the excretion of high amounts of amino acids, which are again available for bacterial protein synthesis. Thus, protozoa play an important role in nitrogen recycling in the rumen (Jouany, 1996; Koenig et al., 2000; Lapierre and Lobley, 2001). The uptake of bacterial cells impacts the bacterial nitrogen turnover (Bach et al., 2005) and incorporated protein is not available for intestinal digestion by the ruminants since the agile protozoal cells actively try to prevent being washed out of the rumen (Puniya et al., 2015). Removal of ciliate protozoa from the rumen resulted in a 30% increase of microbial protein supply and 11% reduction of methane emission (Newbold et al., 2015). Despite their large biomass and functional capabilities, the existence of rumen protozoa is not required for survival of the host but is thought to be beneficial for a more stable ruminal fermentation and the overall wellbeing of the animals (Williams and Coleman, 1992).

The least abundant and diverse rumen fungi appear exclusively in herbivores and are the only obligately anaerobic fungi possessing hydrogenosomes instead of mitochondria (Puniya et al., 2015). Supposed to be the most efficient fiber degraders within the rumen (Akin et al., 1988; Akin and Borneman, 1990), anaerobic fungi exhibit a broad spectrum of hydrolytic enzymes for instance cellulases, hemicellulases, xylanases, pectinases, proteases, amylases, amyloglycosidases, esterases and mannases necessary for plant biomass degradation (Puniya et al., 2015). Ruminal fungi are significantly entangled in the breakdown of lignin-containing plant material (Akin and Borneman, 1990; Thareja et al., 2006; Sirohi et al., 2012) and in this support bacterial access to cellulose (Borneman et al., 1991) and therefore the overall fermentation of cellulolytic material.

Ruminal viruses, lytic and lysogenic archaeaphages and bacteriophages are considered pathogens specific for the prokaryotes present in the rumen and capable of lysing bacterial and archaeal cells. The specificity of phages might be used for removing undesired bacteria or methanogens from the rumen ecosystem (Bach et al., 2002; McAllister and Newbold, 2008). Additionally, phages mediate horizontal gene transfers and may be involved in maintaining the genetic heterogeneity and equilibrium in the rumen ecosystem enabling rapid adaption to changes in environmental conditions like shifts in diet composition (Swain et al., 1996).

1.3 Dietary Impact

Amongst many aspects such as age and species of the host, the diet fed to the animals is the main factor shaping the structure and the functions of the rumen microbiome (Tajima et al., 2001; Ley et al., 2008; Ellison et al., 2014; Henderson et al., 2015). The most common forages used for high-performance dairy cows and fattening cattle across Europe and North America are corn silage and grass silage (Wilkinson et al., 1996) whereas grass hay plays an important role in raw milk-based food production. Silage-derived, non-pasteurized milk may cause problems involving spores of certain *Clostridium* species (Cremonesi et al., 2012; La Torre et al., 2016). Further, high-producing dairy cows and meat cattle require large amounts of energy in form of readily available starch and soluble sugars resulting in an increased supply of concentrate used in the respective diets. In consequence, ruminal fermentation rates increase and more short-chain fatty acids are produced in less time, which can cause a drop in rumen pH and subsequently initiate a detrimental rumen acidosis affecting health and performance of the animals. Additionally, lower fiber contents in the animal's diet reduce the rate of rumination, which in turn influences salivation and thus, further decreases the buffering capacities of the rumen (Hook et al., 2011b; Dijkstra et al., 2012).

Regarding the rumen microbiome, dietary changes may be equated with changes in environmental conditions causing a disturbance in the rumen ecosystem (Monteils et al., 2012) which requires adaptation of the microbial communities shifting towards a more suitable structure for digesting the respective compounds. Fiber-rich diets such as grass-silage and grass hay contain higher amounts of structural carbohydrates and more crude protein when compared to corn silage (Phipps et al., 2010) and thus, promote an increase in abundance and metabolic activity of cellulolytic, hemicellulolytic, pectinolytic and proteolytic bacterial specialists. On the other hand, grain-based diets like corn silage but as well concentrate contain more storage carbohydrates (Juniper et al., 2008) such as starch and sugars and rather stimulate the activity and abundance of amylolytic and saccharolytic bacteria. Microbial fermentation of structural polysaccharides yields mainly acetate whereas fermentation of starch and soluble sugars results in greater proportions of propionate and butyrate as well as higher overall fermentation rates (Czerkawski, 1986). Further, propionate formation requires hydrogen and in this competes with the most significant hydrogenotrophic methanogenesis in reducing hydrogen availability for the respective methanogenic archaea (van Gastelen et al., 2015). Studies confirmed reduced methane emissions of ruminants fed with corn silage-based diets (Beauchemin and McGinn, 2005; van Gastelen et al., 2015). Degradation of crude protein in grass silage and grass hay is faster and more feasible when compared to corn silage whereas readily available energy in corn silage provides a more balanced reservoir of energy and protein sources which results in a higher efficiency of microbial protein synthesis even though corn silage contains less crude protein (Givens and Rulquin, 2004).

Diet composition, in particular the source of carbohydrates and fiber content used in animal diets is the major factor influencing the abundance and metabolic activity of specific bacterial groups in the rumen. Those in turn determine the array of fermentation end products including enteric methane formation hence affecting as well milk composition (Nielsen et al., 2006; Livingstone et al., 2015) and animal health (Russell and Rychlik, 2001; Gressley et al., 2011) and therefore the overall efficiency and performance of the animals.

Studies examining the dietary impact on fermentation characteristics and animal performance are numerous (Abrahamse et al., 2008; Whelan et al., 2011; Brask et al., 2013; Dewhurst, 2013; Hassanat et al., 2013; Zhu et al., 2013). The same applies to investigations that focus on animal feed-induced shifts in rumen microbial community composition in various contexts (Tajima et al., 2001; Fernando et al., 2010; Kong et al., 2010; Pitta et al., 2010; de Menezes et al., 2011; Ann Huws et al., 2012; Belanche et al., 2012; Carberry et al.,

2012; Petri et al., 2013; Thoetkiattikul et al., 2013; Zhang et al., 2013; Lengowski et al., 2016). Nevertheless, most studies available analyzing nutrition- or ration-related effects on the rumen microbial communities are restricted to nucleic acid-based approaches with limited functional insights. Currently, there are methods available targeting particularly the functional profiles of the active rumen microbial communities under certain environmental conditions (Deusch et al., 2015).

1.4 Methodology

Traditionally, the first studies of rumen microbes were based on classical anaerobic cultivation of pure cultures to characterize isolated strains of bacteria (Hungate, 1950; Bryant, 1959; Scott and Dehority, 1965; Hungate, 1966), archaea (Vogels and Stumm, 1980; McInerney et al., 1981), protozoa (Oxford, 1951; Sugden and Oxford, 1952) and fungi (Orpin, 1975; 1976). Cultivation of microorganisms is an essential technique to obtain detailed information about physiological properties of single microorganisms (Palleroni, 1997). However, besides being costly and laborious in execution, anaerobic cultivation methods suffer from severe drawbacks concerning the analysis of total microbial communities since the compositions of culture media and other environmental parameters *in vitro* do not resemble *in situ* conditions. These limitations distort the outcome of such investigations (Nocker et al., 2007). Beyond, only a fragment of the total microbial diversity is cultivable while the majority remains inaccessible via cultivation-based methods (Amann et al., 1995; Creevey et al., 2014). A minority of about 7 to 11% of the rumen bacteria is considered cultivable (Edwards et al., 2004; Kim et al., 2014). The rumen microbiome comprises myriads of different microorganism hence, only restricted knowledge about actual composition of the microbial communities and the totality of processes and interactions carried out in the complex rumen ecosystem can be obtained from information about individual, cultivable species.

Deviations between *in situ* and cultivable diversity have led to the widespread application of more precise, culture-independent molecular methods to analyze the structure of the most abundant and active microbial communities in the rumen represented by archaea and bacteria. Fingerprinting techniques, for instance terminal restriction fragment length polymorphism (Khafipour et al., 2009; Castro-Carrera et al., 2014), denaturing gradient gel electrophoresis (Sadet et al., 2007; Petri et al., 2012) as well as quantitative real-time polymerase chain reaction (Singh et al., 2014; Lengowski et al., 2016) and others like fluorescence *in situ* hybridization (Belenguer et al., 2010; Valle et al., 2015) and microarrays

(Kim et al., 2014; Patra and Yu, 2015) are all based on nucleic acids. More precisely, the 16S rRNA gene, a ribosomal gene that is present in all prokaryotes exhibiting conserved and variable regions, is used to determine evolutionary relationships and phylogenetic similarities between species (Youssef et al., 2009). In particular next-generation sequencing approaches of polymerase chain reaction (PCR)-amplified 16S rRNA gene have provided unique and invaluably detailed knowledge about the composition of the rumen microbiome elucidating the relative abundance of species and the overall diversity and richness (Edwards et al., 2004; Mackie and Cann, 2005; Lee et al., 2012; Shabat et al., 2016).

Nevertheless, only limited functional information can be obtained from such data since, next to the bias introduced by DNA extraction (Bergmann et al., 2010) and PCR amplification (Hong et al., 2009; Huber et al., 2009), most 16S rRNA gene sequences stem from cloning of environmental DNA and do not directly relate to already cultured prokaryotes. Functional capabilities concerning ecological niches are often strain-specific and even closely related species exhibit a vastly different functional potential (Violle et al., 2012; Zhu et al., 2015). Additionally, many 16S rRNA gene sequences identified are not present in databases like the Ribosomal Database Project (Cole et al., 2009) representing novel, completely undescribed species. Further, DNA-based methods may include dead, lysed and dormant cells not contributing to actual functional processes and interactions.

Following the dogma of molecular biology, structural information obtained from the 16S rRNA gene can be extended by messenger RNA/complementary DNA- (Wang et al., 2011; Kang et al., 2013) and protein-based approaches which rather target the metabolically active consortia of a microbiome and provide a more reliable analysis of functions and expressed metabolic pathways within the rumen ecosystem. Moreover, bioinformatic and technical advances paved the way for *Omics*-technologies (Seifert et al., 2013) such as metagenomics, metatranscriptomics and metaproteomics as well as metabolomics (García-Cañas et al., 2014). High-throughput next-generation shotgun sequencing of total DNA does not require prior knowledge about the microbial communities and is not restricted to specific genes like the 16S rRNA gene allowing to bypass cloning and PCR-related bias resulting in a more comprehensive functional prediction that includes all genes potentially present (Brooks et al., 2015; Jovel et al., 2016). Moreover, metatranscriptomic analyses can access the expression of genes present in the rumen under certain conditions (Bashiardes et al., 2016; Li et al., 2016; Li and Guan, 2017). Metagenomic studies revealed further insights into the phylogenetic diversity and potential functions of rumen microbial communities (Brulc et al., 2009; Hess et al., 2011; Ferrer et al., 2012; Pope et al., 2012; Wang et al., 2013; Denman et

al., 2015) and constructed comprehensive metagenomic data publicly available (Meyer et al., 2008) that include also sequences of uncultured and novel microorganisms (Raes and Bork, 2008; Rinke et al., 2013). Increasing numbers of metagenomic datasets provide a growing availability of reference sequences essential for appropriate metaproteomic studies, which enable the link between genetic potential and functional expression of microbial communities in specific environments (Seifert et al., 2013).

Tandem mass spectrometry-based shotgun metaproteomics access the entirety of proteins expressed by the microbial communities of a given ecosystem and provide increased information about taxonomic diversity, actual functional profiles and interactions of the most active fractions of the investigated prokaryotic communities (Wilmes and Bond, 2006; Hettich et al., 2013; Seifert et al., 2013; Wilmes et al., 2015; Muth et al., 2016). However, suitable reference databases (Tanca et al., 2016) and specific sample preparation procedures (Seifert et al., 2013) to reduce co-extraction of eukaryotic proteins and increase the coverage of prokaryotic cells, especially for the highly heterogeneous rumen sample matter (Deusch and Seifert, 2015) are required. Furthermore, humic compounds present in the rumen might interfere with the metaproteomic workflow (Chourey et al., 2010; Heyer et al., 2013) that includes proteolytic digestion of extracted proteins and subsequent measurements of peptides by a one-dimensional liquid chromatography directly coupled to the mass spectrometric analysis (Seifert et al., 2013).

In contrast to proteins, metabolites cannot be assigned to specific strains or species and thus, conclusions about metabolic activity of specific members of complex communities based on metabolites are impossible (Seifert et al., 2013). Nevertheless, microbially produced metabolites such as short-chain fatty acids directly affect the host (Koh et al., 2016) and constitute the link between dietary impact on the structure and metabolic activity of the microbiome and its influence on the host. Modern high-throughput nuclear magnetic resonance-based metabolomic approaches enable the quantification of high numbers of metabolites and other substances present in rumen fluid (Ametaj et al., 2010; Lee et al., 2012; Saleem et al., 2013) and represent valuable data regarding the activity and abundance of specific metabolic pathways despite the lack of phylogenetic information.

Nowadays, a combination of different state of the art *Omics*-technologies represents the most effective and powerful methodology to investigate the microbiomes of highly heterogeneous and complex ecosystems likes the rumen (Figure 2) to obtain the most precise and comprehensive insights (Lamendella et al., 2012; Morgan and Huttenhower, 2014; Deusch et al., 2015).

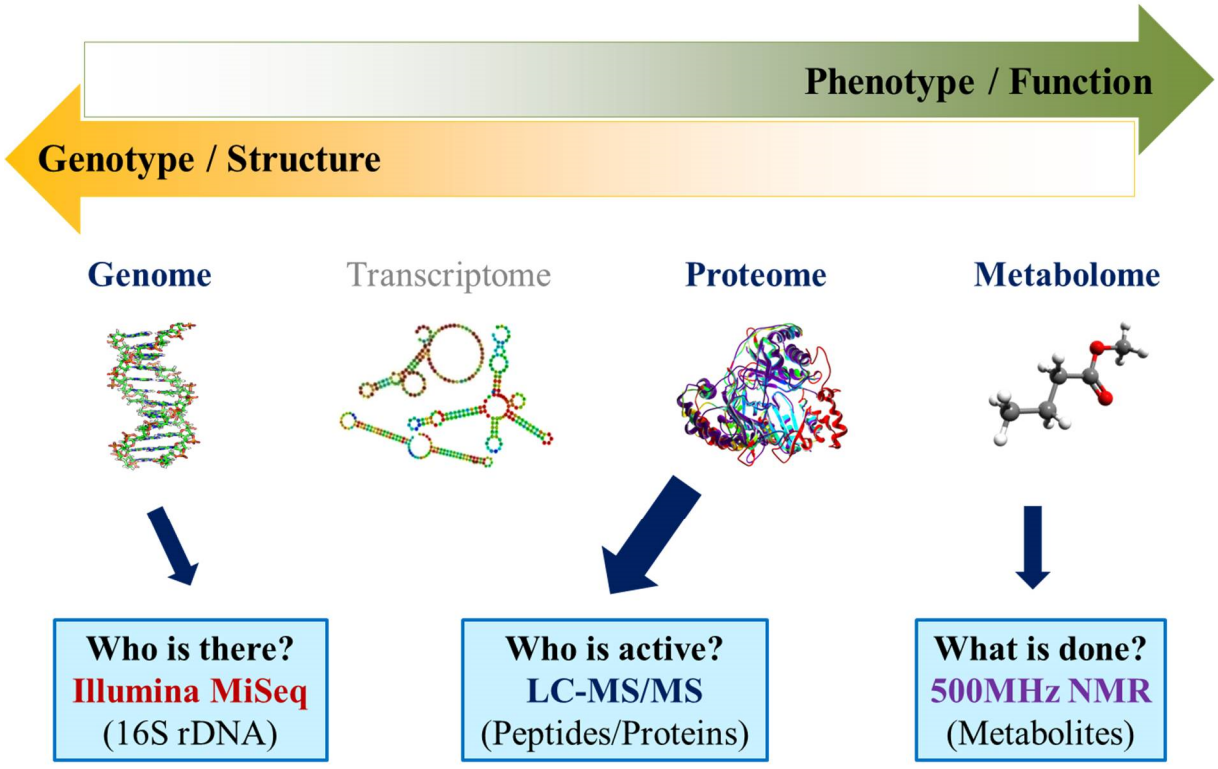

Figure 2| Combination of different *Omics*-technologies to study the rumen microbiome.

1.5 Objectives

During the last decade, the application of *Omics*-technologies and their combination were commonly used to investigate the microbial communities of the human gastrointestinal tract but available studies in livestock science employing state of the art methods are rare. The first step of this project was to provide an overview about *Omics*-based research of the gut microbiome of farm animals including as well pre-*Omics*-approaches, further comparing and discussing their major findings. At that time, most studies considered relied on stand-alone, mainly DNA-based methods and functional metagenomics to investigate the rumen microbial communities. This clearly emphasized the importance of protein-based analyses such as metaproteomics and the integration of various methods to study the rumen microbiome to gain deeper, more complete insights into the actual functions carried out by specific groups of the prokaryotic communities. The first manuscript with the title "**News in livestock research – use of *Omics*-technologies to study the microbiota in the gastrointestinal tract of farm animals**" was published in the Computational and Structural Biotechnology Journal in 2015.

The next step of this work was to achieve a feasible mass spectrometry-based metaproteome characterization of the prokaryotic communities in the rumen requiring the development of specific sample preparation protocols and workflows to overcome the typical heterogeneity of the rumen sample matter. Further, different ecological niches within the rumen had to be taken in account to obtain representative samples. The tightly fiber-adherent bacterial consortia required detachment from plant-particles prior to the separation from the residual matter. The second manuscript "**Catching the tip of the iceberg – Evaluation of sample preparation protocols for metaproteomic studies of the rumen microbiota**" was published in the journal of Proteomics in 2015.

The final step and the major aim of this project was the in depth analysis of the most active members of the complex bionetwork inside the rumen consisting of archaea and bacteria and their adaptive response to the most common forages, corn silage, grass silage and grass hay in order to extend knowledge about dietary effects to improve the efficiency in animal production. So far, the best of my knowledge, there are no studies investigating the metaproteome expressed by the entirety of archaeal and bacterial communities in the different phases of the rumen ecosystem under varying dietary influence. Structural and functional data obtained from the metaproteomic approach were further supplemented by 16S rRNA gene-based analyses of the archaeal and bacterial community structures and the metabolomes of rumen fluids were quantified by nuclear magnetic resonance to obtain insights as accurate and complementary as possible (Figure 2). The third manuscript carrying the title "**A Structural and Functional Elucidation of the Rumen Microbiome Influenced by Various Diets and Microenvironments**" was published in Frontiers in Systems Microbiology in 2017.

CHAPTER II

FIRST MANUSCRIPT

NEWS IN LIVESTOCK RESEARCH – USE OF *OMICS*-TECHNOLOGIES TO STUDY THE MICROBIOTA IN THE GASTROINTESTINAL TRACT OF FARM ANIMALS

Simon Deusch[1], Bruno Tilocca[1], Amélia Camarinha-Silva[1], Jana Seifert[1]

[1]Institute of Animal Science, University of Hohenheim, Stuttgart, Germany

Published in:
Computational and Structural Biotechnology Journal (2015)
Volume 13, 2015, Pages 55-63
doi: 10.1016/j.csbj.2014.12.005
Accepted: December 19, 2014

Open access.
The original publication is available at https://doi.org/10.1016/j.csbj.2014.12.005.

2. NEWS IN LIVESTOCK RESEARCH – USE OF *OMICS*-TECHNOLOGIES TO STUDY THE MICROBIOTA IN THE GASTROINTESTINAL TRACT OF FARM ANIMALS

2.1 Abstract

Technical progress in the field of next-generation sequencing, mass spectrometry and bioinformatics facilitates the study of highly complex biological samples such as taxonomic and functional characterization of microbial communities that virtually colonize all present ecological niches. Compared to the structural information obtained by metagenomic analyses, metaproteomic approaches provide, in addition, functional data about the investigated microbiota. In general, integration of the main *Omics*-technologies (genomics, transcriptomics, proteomics and metabolomics) in live science promises highly detailed information about the specific research object and helps to understand molecular changes in response to internal and external environmental factors.

The microbial communities settled in the mammalian gastrointestinal tract are essential for the host metabolism and have a major impact on its physiology and health. The microbiotas of livestock like chicken, pig and ruminants are becoming a focus of interest for veterinaries, animal nutritionists and microbiologists. While pig is more often used as an animal model for human-related studies, the rumen microbiota harbors a diversity of enzymes converting complex carbohydrates into monomers, which bears high potential for biotechnological applications.

This review will provide a general overview about the recent *Omics*-based research of the microbiota in livestock including its major findings. Differences concerning the results of pre-*Omics*-approaches in livestock as well as the perspectives of this relatively new *Omics*-platform will be highlighted.

© 2014 Deusch et al. Published by Elsevier B.V. on behalf of the Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

2.2 Introduction

The methodology to study the microbial communities (microbiota) inhabiting the gastrointestinal tract (GIT) of livestock was changing from classic cultivation techniques and pure culture characterization to state of the art *Omics*-approaches (Figure 3). Despite cultivation being a sound technique to characterize the physiological properties of

microorganisms [1], there are severe drawbacks in using this as a tool for characterizing bacterial communities. Typically, the culture media do not resemble *in situ* conditions and in some cases the carbon richness is higher than the substrates found *in situ*, allowing the growth of only a small fraction of the community while suppressing other members [2]. In the past, cultivation studies have contributed to our understanding of the gut microbiota, but the limits of these methods directed us to an inaccurate and incomplete knowledge of a niche where most microbiota still remain unknown. The inconsistency between *in situ* and cultivable diversity has resulted in the widespread use of culture-independent molecular approaches [3,4]. Microbial community profiling methods (16S ribosomal RNA gene based approaches) have become important tools to characterize microbial communities and the interactions between the microorganisms present in the GIT. In addition, the complexity of the microbial processes harbors new enzymatic functions, which are of interest for biotechnological applications. Overall, the analysis of the microbiota is important to improve animal nutrition strategies and animal health. This knowledge can be used to modulate the microbiota to reduce antibiotic treatments and, in the case of ruminants, to inhibit the formation of emission gases. Thus, the progress of *Omics*-technologies and the availability of bioinformatic tools to evaluate big datasets demand their use in these fields of research.

Two pyrosequencing techniques, 454 (Roche) and sequencing by synthesis (Illumina), are mainly used for (meta-)genomic and (meta-)transcriptomic projects. Both systems have unique features, such as short paired-end reads (max. 2 × 300 bp) with Illumina *vs*. long read length (600-800 bp) with 454. The latter one is more feasible in terms of shotgun sequencing studies (see below) [5,6], while Illumina provides barcoding strategies and bigger data sets that are more favorable to analyze hundreds of samples in targeted sequencing projects [6]. Two other techniques, that were not frequently applied for metagenomic studies of animal microbiota, are the Ion Torrent (Life Technologies) and the PacBio (Pacific biosciences). All techniques are continuously improving and a state of the art overview is given by C. Knief [6] or can be found at the respective company webpages.

The gene of choice to analyze the phylogenetic composition of a microbial community is the 16S rRNA gene, a ribosomal gene in prokaryotes characterized by conserved and variable sequence regions, which is used to calculate evolutionary relationships and similarities between the species [7]. There are a couple of techniques in molecular ecology, such as fingerprinting methods, microarrays and fluorescence *in situ* hybridization which use the 16S rRNA gene as a target molecule. In this review, we focus mainly on next-generation sequencing methods to describe the microbial community structure.

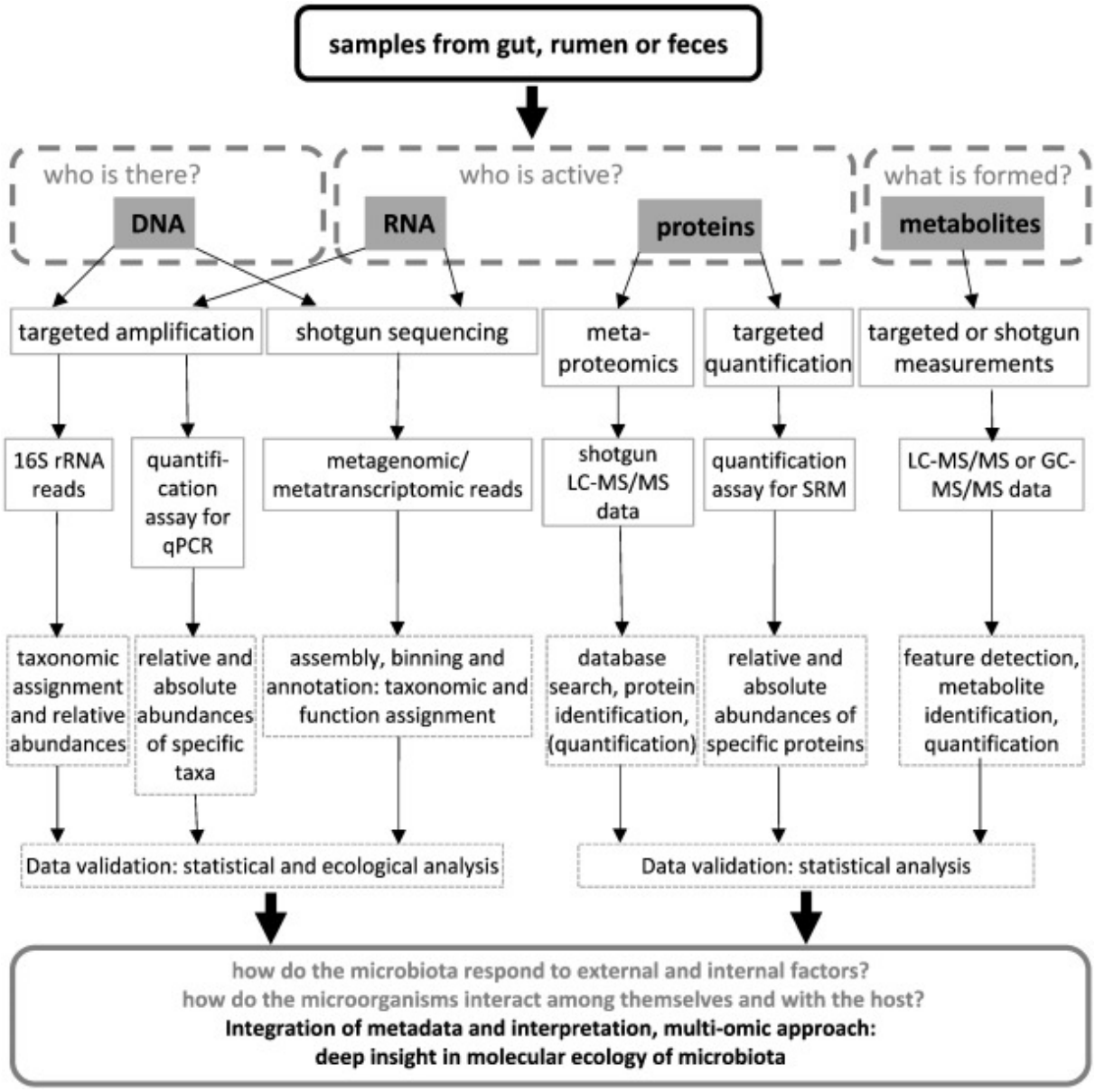

Figure 3| Workflow of possible methods to study the structure and function of the microbiota in farm animals.

Nowadays the total diversity of a microbiological sample is analyzed preferably by pyrosequencing of the 16S rRNA gene, obtained by amplification of extracted DNA. The active fraction of the community is analyzed using mRNA/cDNA. Subsequent to pyrosequencing, quality filtering and denoising processes have to be applied. The reads should be checked for chimeras and clustered to operational taxonomic units (OTU) in order to assign the respective taxonomies to the sequences. There is a diverse range of bioinformatics tools available in free software platforms such as Mothur, QIIME, RDP pipeline, LIBSHUFF, UniFrac and MEGAN that support data analysis and convert data to formats that can later be used in statistical packages like R, Metastats or Primer-E. A detailed

overview of the methods can be found in several review papers [5,8]. These pipelines should be used with special care as it is not only important to make sense of all the raw data, but also to ensure that the final picture is a direct reflection of the original raw data collected and thus of the original community structure of the sample. The output data reveal ecological indices, relative abundance values of the identified taxa and enable a pre-selection for a targeted quantitative PCR (real-time PCR) approaches if necessary.

In addition to the phylogenetic structure of the community, the analysis of encoded and expressed metabolic pathways is the second objective. Metagenomic or metatranscriptomic data are obtained by shotgun pyrosequencing of the total DNA and cDNA, respectively. Reads have to be quality filtered, assembled to contigs, binned and assigned to taxonomies and possible gene functions. As the assembly requires sequence reads with appropriate length, so far 454 pyrosequencing was the method of choice as it produces reads up to 800 bp (see above). Due to the progress in data generation and bioinformatic processing Illumina pyrosequencing is recently used as well. Several tools are available for the annotation of open reading frames on the contigs, MG-RAST [9], MEGAN5 [10], IMG/M [11], Metarep [12] and MicroScope [13]. CAMERA portal [14] was shut down in July 2014. These tools can also be used for metabolic pathway reconstruction. This is usually done based on the KEGG database [15] or the subsystem classification of SEED [16].

In addition to metatranscriptomic studies, the community activity can be assessed based on expressed proteins and formed metabolites. Metaproteomic studies investigate the protein inventory of a specific sample at a certain point of time [17]. This allows the identification of the active microbial fraction and their expressed metabolic pathways. The first key step is to find an optimized sample preparation protocol to avoid co-extraction of eukaryotic proteins and to get a purified protein sample. The following workflow depends on the available technical equipment [18]. In a gel-based approach, proteins are separated and proteolytically digested into peptides followed by a one-dimensional liquid chromatography directly coupled to the mass spectrometric analysis (LC-MS/MS). In a gel-free approach, peptides are prepared by in-solution digestion directly in the protein mix. Peptides are separated by two-dimensional LC and measured by MS/MS analysis. The protein identification is the second big challenge as it is highly depending on the available sequence database which can either be used from public resources or sample-specific sequences. An overview of available bioinformatic tools and workflows are given in [18,19]. The coverage of metaproteomic studies of complex microbial samples, such as feces or rumen contents, is still low. Since there is a high species diversity and cell density in these types of samples, only

abundant proteins are identified while rare species, that may have important metabolic functions, are missed. Targeted proteomic approaches, like selective reaction monitoring (SRM) can be used to specifically detect and quantify proteins of interest [20]. Metabolomic approaches are becoming more interesting for microbial ecology studies as the technical progress allows a comprehensive analysis of hundreds to thousands of metabolites. NMR- and MS-based methods are available and their application to detect defined groups of metabolites is reviewed by Xie et al. [21].

The following sections will provide an insight into the ongoing research of the microbiota of the gastrointestinal tract of livestock animals with special emphasis to the use of *Omics*-technologies and their importance for the understanding of these niches.

2.3 The Microbiota of Chicken

The chicken intestinal environment comprises a vast and diverse assemblage of microorganisms living not as single species populations, but rather in complex communities comprising multiple species that include animal and human pathogens. Intricate networks of interactions between the microorganisms and their environment shape the respective communities and are important for animal welfare and food safety reasons. The chicken GIT consists of more than 900 species of bacteria. This diverse microbiota helps not only the breakdown and digestion of food but also plays an important role concerning the growth and health of the host [22].

In the past, the chicken GIT microbial community was studied by culture-based methods. These studies discovered that 10-60% of cecal bacteria can be cultured [23,24] and about 45% could be assigned to the genus level [22]. The profiles of the different gut sections are nowadays studied using cultivation-independent methods like clone libraries [25-27], denaturing gradient gel electrophoresis (DGGE) [28], temperature gradient electrophoresis (TGGE) [27], terminal restriction fragment length polymorphism (T-RFLP) [4,29-31], quantitative PCR (qPCR) [32], microarrays [33], next-generation sequencing [34-39] and metaproteomics [40].

Regarding the research in microbial ecology of the chicken GIT, several studies focused on the influence of diet [25,31,32,41], antimicrobial feed additives [29,34] host genotype [38,42], gender [38], spatial microbial diversity [25,28,31,43], age [28], development and temporal microbial variations [22,26,28]. It is important to take into consideration that all these factors may change the bacterial community of each section. Sklan et al. showed that the different sections of the chicken GIT are highly inter-connected [44].

However, because of the high diversity within each section, it has been suggested to analyze them as separate ecosystems [28]. It was demonstrated that the microbial communities colonizing the GIT of chicken benefit the host [29,31,36]. Nevertheless, two recent studies revealed that this colonization can also harm the host [35,45].

After hatching, the colonization of the chicken GIT begins. This is a moment of great importance regarding the establishment of the microbial communities. Although the colonization of the chickens by maternally derived bacteria is low, some studies postulated that the microbial community structure of the small intestine settles within two weeks. Older studies showed that cecal bacteria need longer time to develop [40,46]. The gut is colonized by commensal, transient and pathogenic microorganisms. Commensal microorganisms are beneficial to the host as they provide amino acids, short-chain fatty acids and vitamins [40]. Stanley et al. observed inter-individual GIT variation between microbial groups and also differences between groups of birds from replicate trials. It was suggested that the hygiene levels of the new hatcheries might cause highly variable gut microbial community [37].

The chicken gut is divided in three upper segments: crop, proventriculus and gizzard. The crop is a food storage muscular pouch related to the breakdown of starch and the fermentation of lactate. Digestion starts in the proventriculus while the gizzard grinds food. Because of its lower pH and fermentation activity, the gizzard functions as microbial barrier. Similar microbial communities were found in the crop and gizzard. Lactobacilli, facultative and microaerophilic bacteria are the most dominant bacteria present in this two segments. Other abundant species belonged to Clostridiaceae, *Enterococcus* and in the case of the crop also *Bifidobacterium* and Enterobacteriaceae (Figure 4) [25,46,47]. The small intestine is relatively long and has a constant diameter. It consists of three parts: the duodenum, jejunum and ileum where the nutrient absorption and food digestion occurs. Due to the low pH, pancreatic and bile secretions, the bacterial density in the duodenum is comparably low. Besides *Lactobacillus* as the main colonizer of the jejunum (reaching coverage of up to 99%), *Streptococcus* was identified as well. Amit-Romach et al. has shown that the relative proportion of *Lactobacillus* spp. in duodenum and jejunum increases within age [48]. The chicken's ileum harbors *Lactobacillus* in higher abundance (> 68%) and in lower abundances *Streptococcus*, Enterobacteriaceae and Clostridiaceae [28,43]. Lu et al. demonstrated that during all different stages of microbial community development in the ileum Lactobacilli were dominant [26]. This gut section is also known to be colonized by novel butyrate producing bacteria that may play an important role regarding the availability of nutrients, absorption rate and chicken performance [47].

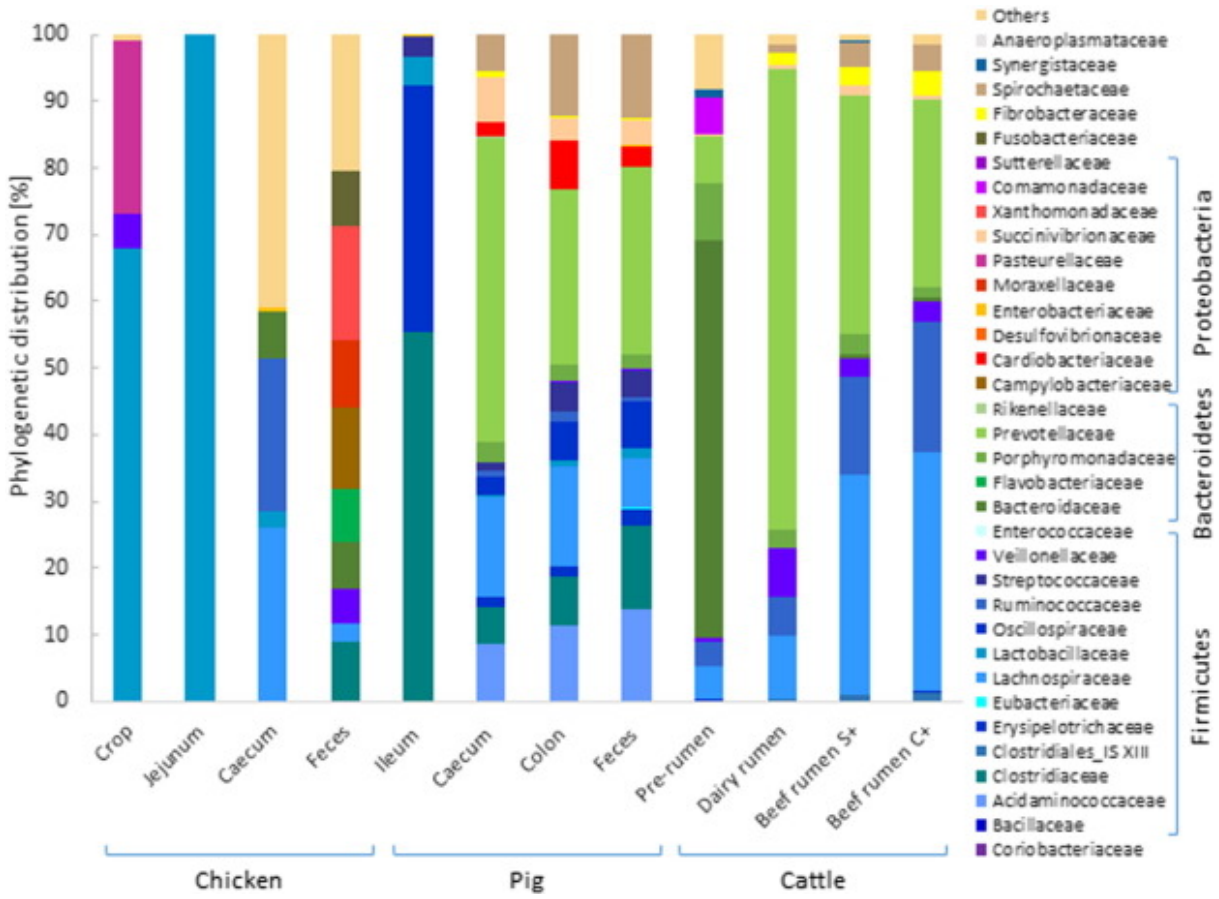

Figure 4| Phylogenetic distribution of bacterial families in different GIT sections of chickens, pigs and cows. Chicken's crop, jejunum and cecum data arise from the analysis of V1-V3 16S rRNA region as performed by Videnska et al. [51], Stanley et al. [36] and Sergeant et al. [39], respectively. All pig's data arise from the study performed by Looft et al. [57] on V1-V3 16S rRNA region. Cow's data derive from the work performed by Wu et al. [89] on V3-V5 16S rRNA region.

Chickens have two ceca which are important for recycling urea, the absorption of water, and digestion of cellulose, starch and polysaccharides. These two fermentation chambers have the highest bacterial density and are colonized by obligate anaerobes like *Clostridium*, Bacteroidetes, and *Bifidobacterium* (Figure 4) [42]. Recently, 16S rDNA amplicon pyrosequencing studies estimated a bacterial population of about 700 species [39]. This wealth of microorganisms makes the ceca an important study site and a reservoir rich in unknown and uncultured microorganisms and pathogens [30,39,46,47]. Qu et al. proved that mobile DNA elements are the cause of functional microbiome evolution and that horizontal gene transfers and the metavirulomes of cecal microbiomes were related to the host environment [49]. A metagenomic analysis of the chicken cecum using the Illumina MiSeq 2000 system revealed a relatively high proportion of sequences encoding glycosyl hydrolases

that were identified by sequence comparison with the carbohydrate active enzymes (CAZy) database (Figure 5) [39]. More than 200 genes of non-starch polysaccharide-degrading enzymes were identified indicating a great potential for xylane degradation compared to a lower cellulolytic potential in the ceca. This is also congruent to the comparative study of Waite and Taylor describing an abundance of β-xylosidase and β-glucosidase in grain-fed chickens [50]. Both studies also described the presence of genes involved in propionate and lactate production [39,50].

Chicken feces samples are colonized by *Lactobacillus*, *Clostridium*, *Faecalibacterium*, *Ruminococcus*, *Bacillus*, *Eubacterium*, and *Fusobacterium* (Figure 4). Here the microbiota is not stable and it has been proposed that these fluctuations are related to the emptying of the previous gut sections [25]. A recent study in meta-analysis of the avian gut microbiota showed that genes related to cytokine receptors and cell adhesion grouping into "signaling molecules and interaction" were less present in fecal samples indicating a lower potential of host/bacteria interactions [50]. The only metaproteomic study using a chicken fecal sample identified about 3487 proteins in total [40]. Bacterial proteins mainly belonged to *Lactobacillus* and *Clostridium*. Gene ontology analyses showed that the majority encodes for stress-related proteins like chaperons and proteases as well as enzymes involved in glycolysis [40].

Antibiotic growth promoters improve chicken growth performance and health status. The inclusion of penicillin in diets increases the body weight of chickens and also the Firmicutes to Bacteroidetes ratio in ceca. These effects might be caused by a reduction of the weight of the small intestine and the thickness of the gut wall, increasing the absorption of nutrients. The addition of the antibiotics tetracycline and streptomycin also induces a rapid shift in microbial community, increasing the prevalence of Lactobacillales and Enterobacteriales in fecal samples. The restoration of the microbial community after usage of these antibiotics was observed after removing the therapy [51].

In the era of next-generation sequencing, high-throughput technologies have brought an immense contribution in characterizing the poultry microbiota, bridging genomics, immunology, physiology, host and environmental factors to give a precious insight into animal production, food safety and public health.

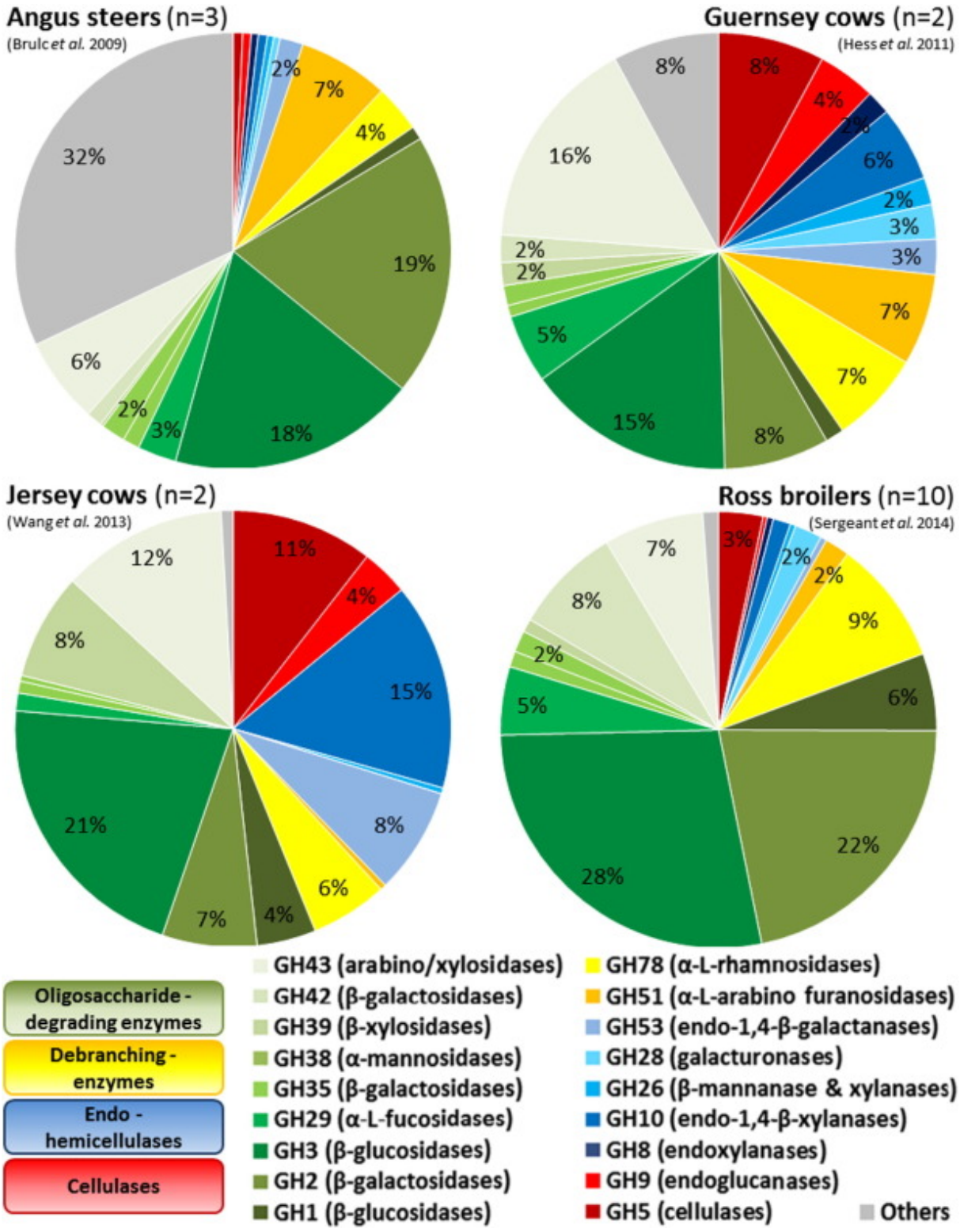

Figure 5| Abundance of glycoside hydrolase (GH)-families in metagenomes of bovine rumen and chicken cecum. The percentage of each GH-group relative to the total number of GH-families identified in each metagenomic dataset is shown grouped according to major activity [90]. Brulc et al. [72] [Angus steers] — Pyrosequencing data (shotgun sequencing using GS20 from 454 Life Science) of 4 metagenomic samples; the mean of three fiber-adherent and one pooled liquid sample is shown. The average size of the metagenomes was 0.026 Gb. The

samples were obtained from three 5-year-old Angus Simmental Cross steers maintained on grass-legume hay. Hess et al. [70] [Guernsey cows] — Massively parallel shotgun sequencing using Illumina GAIIx and HiSeq 2000 was applied on metagenomic samples of the fiber-adherent rumen microbiota of two Guernsey cows kept on a mixed diet containing 60% fiber. The total metagenome size was 268 Gb. Wang et al. [91] [Jersey cows] — All samples were pooled at equal amount and pyrosequenced with the Roche GS FLX Titanium system. Average size of metagenomes was 0.49 Gb. Rumen digesta samples were collected from two Jersey cows fed mainly Timothy grass hay *ad libitum*. Sergeant et al. [39] [Ross broilers] — Cecal samples were collected from 10 Ross broilers consuming a wheat based diet with 5% maize which contained ionophores but no antibiotics. Sequencing was carried out on the Illumina Miseq 2000 system.

2.4 The Microbiota of Pig

Pigs harbor a complex gut microbiota, which establishes strong and complex interactions with the host. Since the importance of these interactions and their implication in nutritional, immunological and physiological functions became more relevant, several research groups started to focus on the characterization of the porcine gut microbiota by using different methods. In the past, members of the porcine gut microbiota were investigated by cultivation attempts that are limited to a small fraction as it is difficult to achieve optimal growth conditions in vitro [52]. However, cultural methods are still used and flanked with cultivation-independent techniques. Furthermore, isolation attempts of novel species are still necessary to describe novel metabolic functions by physiological tests. Disadvantages of the culture-based methods triggered a wider use of cultivation-independent methods for the investigation of gut microbiota in the last two decades [4]. qPCR [53-55], T-RFLP [53,55] and microarrays [56] were used to study the porcine microbiota. A comparison between culture-based and fluorescence *in situ* hybridization combined with flow cytometry detection (FCM-FISH) methods were performed by Collado and Sanz [52] and revealed a better sensitivity with the FCM-FISH technique. Currently, several studies applied *Omics*-technologies such as metagenomics [57-62] and metabolomics [63-66]. To our knowledge, no metaproteomic and metatranscriptomic study on pig's gut microbiota was published so far.

Most investigated sections within the pig's GIT are ileum (small intestine), cecum and colon (large intestine) (Figure 4). Phylogenetic characterization, based on amplification of the V1-V3 region of 16S rRNA gene and pyrosequencing of the amplicons, showed both longitudinal and radial differences along the GIT [57]. The ileum lumen samples, for

example, revealed a lower diversity in terms of richness and abundance when compared with other gut sections. This comprises almost exclusively Firmicutes and Proteobacteria, whereas the phylum level profiles of the cecum and mid-colon are highly congruent and include mainly Firmicutes, Proteobacteria, Bacteroidetes and Spirochetes. Other phyla such as Fibrobacteres, Actinobacteria, Tenericutes, Synergistetes and Planctomycetes are present but their sequences constitute less than 1% of total rRNA gene sequences [57]. Interestingly, mucosa-associated bacterial communities along GIT are different from those present in the lumen. However, statistically significant differences were found solely in the ileum between the mucosal and luminal communities and most lumen-associated bacteria were also found at mucosal level. Total DNA sequencing using 454 pyrosequencing and a subsequent SEED subsystem annotation of metagenomic sequences from GIT sections showed that unlike samples from the large intestine, the ileum microbiota was completely devoid of enzymes for pectin and hemicelluloses degradation [57]. By contrast, all sites encode starch-degrading enzymes. Members of Bacteroidetes represented about half of the microbiome in large intestine sections and harbored enzymes for polysaccharide degradation. The ileum was enriched in Firmicutes-associated genes of numerous bacterial ABC transporters for monosaccharides and amino acid uptake and bacterial carbohydrate transport phosphotransferase systems showing a preference for the metabolization of easily accessible low molecular weight molecules by Firmicutes species. Therefore, a clear separation of the carbohydrate degradation steps based on the phylogenetic level in the pig GIT can be made, starting with the conversion of polysaccharides to oligosaccharides by pathways encoded in Bacteroidetes and followed by the uptake and fermentation of monosaccharides by metabolic processes encoded in Firmicutes.

Concerning fecal-associated microbiota, shotgun metagenomics analysis followed by sequence annotation using both MG-RAST and JGI IMG/M-ER pipelines [59] showed that metagenomic swine fecal datasets were dominated by the phyla Firmicutes and Bacteroidetes. Numerically-abundant bacterial orders revealed that Clostridiales, unclassified Firmicutes, Bacteroidales, Spirochaetales, unclassified Gammaproteobacteria, and Lactobacillales were the top six most abundant bacterial orders. Archaeal sequences constituted less than 1% of total 16S rRNA gene sequences, and were dominated by the Methanomicrobia and Thermococci [59]. Annotation pipelines used by Lamendella and co-workers have shown that carbohydrate metabolism was the most abundant SEED subsystem, representing 13% of swine fecal metagenomes [59]. Other abundant functional genes were associated with the subsystem cell wall and capsule, stress, and virulence. Additionally, 75% to 90% of

metagenomic reads could not be assigned to subsystems, suggesting the need for improving binning and coding region prediction algorithms to annotate these unknown sequences [59].

Structure and activity of GIT microbiota can differ significantly between animals depending on the breed, diet, health status, age and environment [56-58]; suggesting the investigation of pig's gut microbiota as a powerful and versatile tool to predict effects of new feeding/breeding strategies and also perform studies on animal welfare. A study investigating diet-induced obesity in pigs identified an increase in proportion of the phyla Firmicutes compared to Bacteroidetes by T-RFLP and qPCR approaches [55]. This study also points towards high fat/high caloric diets as a main factor changing the gut microbial community composition. In addition, non-targeted metabolite profiling approaches used by Hanhineva et al. discovered that metabolic effects of high fat diets causing obesity were observed in all examined biofluids (plasma, urine, and bile) [66]. 16S rRNA sequencing investigations were performed to observe possible effects of genetically modified maize on the intestinal microbiota either in short [67] or long-term [60] pig feeding studies. Similar levels of overall biodiversity for both treatments (isogenic vs. Bt-maize) were determined; moreover, no statistical differences occurred in microbiota composition except for the genus *Holdemania* that was more abundant in isogenic group. However, the authors argued that this difference may be related to the changing of the maize source during the animal's early life, when the gut microbiota has not completely developed [60].

Several other studies investigated how different diet composition can affect porcine gut microbiota in order to draw either a balanced diet able to ensure a higher animal growth rate [53,61,63,64], or cost-effective [60] and environmental friendly diets [54,61]. Another point of interest is the potential of the intestinal microbiota to improve the animal's health status by stimulating the growth of beneficial commensal on the expense to opportunistic pathogens [53,54].

Since the importance of gut microbiota in animal production was clarified, the study of in-feed antibiotic (AB) effects on porcine gut microbiota is now of great importance. Nowadays various groups focus on understanding how the use of antibiotics promotes animal growth and how it affects the gut microbiota in short- [58] and long-term treatments [56]. It is also of interest if different effects occur depending on genetic background, age, and/or environment where the animal is bred [58]. Particular attention is attributed to the investigation of gut microbiota development of AB-treated sow's offspring in order to understand how imprinting mechanisms can be impaired in AB-treated pregnant sows [56]. However, more investigation in this field is required, not only due to its importance to human

health. Further studies to analyze the active fraction of the microbiota in the porcine gut by using metatranscriptomics and metaproteomics have to be done in the future.

2.5 The Microbiota in the Rumen

Over 3.5 billion domesticated ruminants worldwide including cattle, sheep and goats (http://faostat.fao.org/) constitute a highly significant source of food products to humans. These animals host a complex gut microbiome (comprising about 10^{10} bacteria, 10^{7} archaea, 10^{8} protozoa and 10^{3} fungal spores per ml rumen fluid [68]) which in exchange provides various enzymes essential for the breakdown of plant fibers into volatile fatty acids and microbial crude protein. The microbial community composition and the active metabolic pathways involved in ruminal microbial metabolism were studied intensively during the last years and are of great interest to animal nutrition [69], biotechnology [70] and climatology [71].

In cell numbers bacteria are most abundant representing over 95% of microorganisms within the rumen ecosystem [72] and were first described using classical microbiology methods [73]. Over 200 bacterial species from the rumen were cultivated and most of them have been described physiologically [74]. Nevertheless, nucleic acid based approaches revealed that culture-dependent methods can only detect around 11% of the present bacterial phylogeny, thus yielding imprecise and incomplete datasets [75]. For example, the cultivable genus *Ruminococcus* was believed to play a major role in ruminal cellulose degradation but actually appeared only below quantities of 2% [76].

Combinations of high throughput *Omics*-technologies in rumen microbial ecology provide a deeper insight into the symbiotic host-microbe relationship and the impact of nutritional strategies on the animal performance [77]. Comparisons between studies are challenging due to numerous analysis steps, varying methods and sampling strategies. Additionally the structure of the rumen microbiota differs significantly across individual animals [78] and depends on the substrates provided by specific diets [75].

Investigations of the rumen biology usually focus on bacterial or archaeal communities neglecting eukaryotic microorganisms. In order to characterize the entire rumen community, barcoded amplicons from all three domains of life were mixed and analyzed *via* Multiplex 454 Titanium pyrosequencing [79]. Twelve DNA samples from 11 ruminants out of three different species kept on various diets were processed indicated positive associations of *Methanobrevibacter ruminatium* and the Fibrobacteraceae family. The phylogenetic distribution was determined considering 257,485 bacterial, 125,052 archaeal, 45,231

protozoal and 186,485 fungal sequencing reads using the QIIME software package [79].

A comparable high-throughput approach analyzed the gut bacteria, archaea and fungi of 12 beef cows *via* 454 pyrosequencing concluding that in comparison with the bacterial community, archaea and fungi were more consistent during dietary alteration in liquid and solid fractions [80]. DNA sequences were processed using Mothur and CD-HIT suite. Observed species richness based on the V1-V3 region of the 16S rRNA gene accounted for 1,903 to 2,432 bacterial OTUs and between 8 and 13 archaeal OTUs per sample. Fungal OTUs based on 18S rRNA gene ranged from 21 to 40 [80].

Similar species richness was determined, with more than 1,000 OTUs, by a pyrotag sequencing approach of DNA extracts from plant fiber material placed in the rumen for 72 h. The same material was used for a deep sequencing approach of the total DNA detecting a huge number of CAZymes (Figure 5) and allowing the assembly of 15 genomes of uncultured bacteria [70].

The diversity of the bacterial community structure was analyzed in liquid and solid fractions of the rumen via metagenomic approaches [72,81] and confirmed the previous findings of a DGGE-ARISA study [82]. Bacteria more abundant in solid fractions, as *Ruminococcus* spp., *Fibrobacter succinogenes* and *Selenomonas ruminatium*, are more likely to be involved in the degradation of polysaccharides. The average number of identified sequences per animal within diet and fraction ranged from 1,822 in the Bermuda grass liquid fraction to 3,675 in the wheat solid fraction [81].

A PCR-DGGE fingerprint study indicated that the bacterial community structure of three Holstein cows did not change among five different gut sampling locations and three daily time points. Anyhow, a greater community shift was observed between individuals fed the same diet concluding that the deviation between animals is greater than the differences between fractions or time points [83].

An Illumina GAIIx-based study applied massively parallel sequencing to establish quantitative rumen microbiome profiles [84]. Eleven rumen fluid samples of three dairy cows resulted in more than 6million reads of 146 bp length in each library. Commonly applied freeware was used to process the obtained sequence data. It was confirmed that the variation in rumen microbial metagenomes of different animals was greater within samples of the same rumen [84].

Furthermore, differences in rumen microbial ecology of 16 Holstein Friesian dairy cows kept on an equal diet were determined by bacterial tag-encoded amplicon pyrosequencing from the V2 and V3 regions of the 16S rRNA gene. In total 162,000

sequencing reads were filtered using the QIIME pipeline yielding 4,986 OTUs overall. The samples had an average of 1,800 OTUs but shared only 154 OTUs out of 32 genera. This comparably small core microbiome suggests a high functional similarity between individuals despite the actually observed phylogenetic differences [78].

The rumen microbiotas of three steers consuming a common diet were investigated by a full-length 16S rDNA clone library approach and 454 pyrosequencing of the total DNA [72]. Most sequences (64%) aligned to 59 OTUs are present in all libraries, whereas 273 OTUs containing 10% of sequences belonged to a single library. Besides, a wide range of unique glycoside hydrolase catalytic modules with 3,800 sequences belonging to 35 glycoside hydrolase families were found to be present in the bovine microbiomes [72].

The rumen microbiome represents an important source of novel enzymes promising for biotechnological applications (Fig. 3). A deep sequencing approach using paired-end Illumina sequencing of DNA extracts obtained from plant fiber-adherent bacteria of a cow rumen yielded in 268 Gb of metagenomic DNA [70]. 27,755 putative CAZy genes were identified after sequence analyses showing a sequence similarity of less than 95% for 99% of the sequences. To discover new enzyme activities 90 ORFs were selected for protein expression studies and 57 of the expressed proteins showed clear cellulolytic activities. This study demonstrated for the first time the benefit of deep metagenomic sequencing and activity screenings in the discovery of novel enzymes from the cow rumen [70].

Ferrer et al. used metagenomic libraries and functional screening assays for the detection of novel glycosyl hydrolases (GH) [85]. They discovered a multifunctional enzyme of GH family 43 belonging to Clostridiales and showing unusually broad substrate specificity. The 3D structure of the enzyme was modeled to determine the substrate binding sites and catalytic domains. These activity-based screening studies showed clear benefits to discover new metabolic functions besides the sole sequence analyses of DNA or RNA extracts.

Along with the microbial community composition two studies analyzed the rumen microbial metabolic profile *via* NMR [86,87]. Thereby Lee et al. [86] suggested that the bovine host breeds are overlaying specific diets as major factor in determining the bacterial community structure and their metabolite profiles. Zhao et al. [87] was able to associate several metabolites with specific diets containing different types of roughages.

One study providing valuable information for milk production investigated the bacterial communities of 15 dairy cows *via* pyrosequencing and compared to production parameters and milk composition [88]. 141,344 reads averaging 338 bp in length were obtained detecting 17 bacterial phyla in total of which only seven were present in all cows.

The results indicated that the ratio of Firmicutes to Bacteroidetes was clearly associated with milk fat content, but most other taxa were rather related to the residual feed intake phenotype. Elucidating the role of rumen microbiota in shaping host physiological parameters may promote better agricultural yield through modulation of bacterial community structure [88].

2.6 Concluding Remarks

The most extensive surface in the animal body is the GIT that harbors an immense variety and amount of microorganisms. Internal and external factors can unbalance this dynamic and complex niche and thereby, also disturb or improve the animal's health status.

Until recently, comparative studies of the microbiota were done between a few points of time and samples, sometimes even pooled samples were used. The results are often contradictory depending on the used animal (breed, age, gender etc.), the experimental setup (feeding and sampling), and used DNA extraction and sequencing method (target region of the 16S rDNA gene sequence). Therefore, it is hard to compare those studies and correlate them with each other. Nowadays, *Omics*-methods offer the advantage of being able to reliably measure and compare hundreds of samples simultaneously with low costs per sample. The millions of sequence reads available through pyrosequencing methods exceed the depths necessary to describe microbial community compositions of a few samples by far. Therefore, inter- and intra-population similarities, temporal dynamics and effects of external factors on the GIT community of livestock should be addressed with the comparison of a broad array of samples. Requirements to deeply cover the phylogenetic diversity are optimized nucleic acid extraction methods and amplification strategies, especially the choice of the amplification region within the 16S rRNA.

Metagenomic sequencing and genome assemblies of uncultured prokaryotes already allows the detection of potential functions of the microbiota, but the analysis of the active fraction of the microbiota in the GIT of animals is still in their infancy. Metatranscriptomic and metaproteomic analysis should gain more importance within the next years to grant deeper insights into the expressed pathways and community interaction mechanisms. Labeling and imaging techniques will support the description of the *in vivo* activity of the communities and of single members. Combination of the collected data will support modeling approaches to detect microbial response mechanisms towards different feeding strategies, pathogens, antibiotics or environmental changes. When compared to the human gut, the analysis of livestock GIT was clearly neglected in the past years, but mainly due to the functional diversity, it should become of interest for future analyses.

2.7 Acknowledgments

We greatly acknowledge the financial support by the Carl-Zeiss- Stiftung and the Ellrichshausen'sche Stiftung.

2.8 References

[1] Palleroni NJ. Prokaryotic diversity and the importance of culturing. *Antonie Van Leeuwenhoek* 1997; 72:3-19.

[2] Nocker A, Burr M, Camper AK. Genotypic microbial community profiling: a critical technical review. *Microb Ecol* 2007; 54:276-89.

[3] Hugenholtz P, Goebel BM, Pace NR. Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. *J Bacteriol* 1998; 180:4765-74.

[4] Zoetendal EG, Cheng B, Koike S, Mackie RI. Molecular microbial ecology of the gastrointestinal tract: from phylogeny to function. *Curr Issues Intest Microbiol* 2004; 5:31-47.

[5] Thomas T, Gilbert J, Meyer F. Metagenomics — a guide from sampling to data analysis. *Microb Inf Exp* 2012; 2:3.

[6] Knief C. Analysis of plant microbe interactions in the era of next generation sequencing technologies. *Front Plant Sci* 2014; 5:216.

[7] Youssef N, Sheik CS, Krumholz LR, Najar FZ, Roe BA, Elshahed MS. Comparison of species richness estimates obtained using nearly complete fragments and simulated pyrosequencing-generated fragments in 16S rRNA gene-based environmental surveys. *Appl Environ Microbiol* 2009; 75:5227-36.

[8] Kim M, Lee KH, Yoon SW, Kim BS, Chun J, Yi H. Analytical tools and databases for metagenomics in the next-generation sequencing era. *Genomics Inf* 2013; 11:102-13.

[9] Meyer F, Paarmann D, D'Souza M, Olson R, Glass EM, Kubal M. The metagenomics RAST server — a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinforma* 2008; 9:386.

[10] Huson DH, Weber N. Microbial community analysis using MEGAN. *Methods Enzymol* 2013; 531:465-85.

[11] Markowitz VM, Chen IM, Chu K, Szeto E, Palaniappan K, Grechkin Y. IMG/M: the integrated metagenome data management and comparative analysis system. *Nucleic Acids Res* 2012; 40:D123-9.

[12] Goll J, Rusch DB, Tanenbaum DM, Thiagarajan M, Li K, Methe BA. METAREP: JCVI metagenomics reports — an open source tool for high-performance comparative metagenomics. *Bioinformatics* 2010; 26:2631-2.

[13] Vallenet D, Belda E, Calteau A, Cruveiller S, Engelen S, Lajus A. MicroScope — an integrated microbial resource for the curation and comparative analysis of genomic and metabolic data. *Nucleic Acids Res* 2013; 41:D636-47.

[14] Seshadri R, Kravitz SA, Smarr L, Gilna P, Frazier M. CAMERA: a community resource for metagenomics. *PLoS Biol* 2007; 5:e75.

[15] Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M. KEGG for linking genomes to life and the environment. *Nucleic Acids Res* 2008; 36:D480-4.

[16] Overbeek R, Begley T, Butler RM, Choudhuri JV, Chuang HY, Cohoon M. The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Res* 2005; 33:5691-702.

[17] Wilmes P, Bond PL. Towards exposure of elusivemetabolicmixed-culture processes: the application of metaproteomic analyses to activated sludge. *Water Sci Technol* 2006; 54:217-26.

[18] Seifert J, Herbst FA, Halkjaer Nielsen P, Planes FJ, Jehmlich N, Ferrer M. Bioinformatic progress and applications in metaproteogenomics for bridging the gap between genomic sequences and metabolic functions in microbial communities. *Proteomics* 2013; 13:2786-804.

[19] Abraham PE, Giannone RJ, Xiong W, Hettich RL. Metaproteomics: extracting and mining proteome information to characterize metabolic activities in microbial communities. *Curr Protoc Bioinforma* 2014; 46:13 26 11-4.

[20] Maiolica A, Junger MA, Ezkurdia I, Aebersold R. Targeted proteome investigation via selected reaction monitoring mass spectrometry. *J Proteomics* 2012; 75:3495-513.

[21] Xie G, Zhang S, Zheng X, Jia W. Metabolomics approaches for characterizing metabolic interactions between host and its commensal microbes. *Electrophoresis* 2013; 34:2787-98.

[22] Apajalahti J, Kettunen A, Graham H. Characteristics of the gastrointestinal microbial communities, with special reference to the chicken. *Worlds Poult Sci J* 2004; 60: 223-32.

[23] Mead GC. Microbes of the avian cecum: types present and substrates utilized. *J Exp Zool Suppl* 1989; 3:48-54.

[24] Salanitro JP, Blake IG, Muirhead PA. Studies on the cecal microflora of commercial broiler chickens. *Appl Microbiol* 1974; 28:439-47.

[25] Sekelja M, Rud I, Knutsen SH, Denstadli V, Westereng B, Naes T. Abrupt temporal fluctuations in the chicken fecal microbiota are explained by its gastrointestinal origin. *Appl Environ Microbiol* 2012; 78:2941-8.

[26] Lu J, Idris U, Harmon B, Hofacre C, Maurer JJ, Lee MD. Diversity and succession of the intestinal bacterial community of the maturing broiler chicken. *Appl Environ Microbiol* 2003; 69:6816-24.

[27] Zhu XY, Zhong T, Pandya Y, Joerger RD. 16S rRNA-based analysis of microbiota from the cecum of broiler chickens. *Appl Environ Microbiol* 2002; 68:124-37.

[28] van der Wielen PW, Keuzenkamp DA, Lipman LJ, van Knapen F, Biesterveld S. Spatial and temporal variation of the intestinal bacterial community in commercially raised broiler chickens during growth. *Microb Ecol* 2002; 44:286-93.

[29] Torok VA, Allison GE, Percy NJ, Ophel-Keller K, Hughes RJ. Influence of antimicrobial feed additives on broiler commensal posthatch gut microbiota development and performance. *Appl Environ Microbiol* 2011; 77:3380-90.

[30] Torok VA, Hughes RJ, Mikkelsen LL, Perez-Maldonado R, Balding K, MacAlpine R. Identification and characterization of potential performance-related gut microbiotas in broiler chickens across various feeding trials. *Appl EnvironMicrobiol* 2011; 77:5868-78.

[31] Torok VA, Ophel-Keller K, Loo M, Hughes RJ. Application of methods for identifying broiler chicken gut bacterial species linked with increased energy metabolism. *Appl Environ Microbiol* 2008; 74:783-91.

[32] Singh P, Karimi A, Devendra K, Waldroup PW, Cho KK, Min Kwon Y. Influence of penicillin on microbial diversity of the cecal microbiota in broiler chickens. *Poult Sci* 2013; 92:272-6.

[33] van der Hoeven-Hangoor E, van der Vossen JM, Schuren FH, Verstegen MW, de Oliveira JE, Montijn RC. Ileal microbiota composition of broilers fed various commercial diet compositions. *Poult Sci 2013*; 92:2713-23.

[34] Danzeisen JL, Kim HB, Isaacson RE, Tu ZJ, Johnson TJ. Modulations of the chicken cecal microbiome and metagenome in response to anticoccidial and growth promoter treatment. *PLoS One* 2011; 6:e27949.

[35] Singh KM, Shah T, Deshpande S, Jakhesara SJ, Koringa PG, Rank DN. High through put 16S rRNA gene-based pyrosequencing analysis of the fecal microbiota of high FCR and low FCR broiler growers. *Mol Biol Rep* 2012; 39:10595-602.

[36] Stanley D, Denman SE, Hughes RJ, Geier MS, Crowley TM, Chen H. Intestinal microbiota associated with differential feed conversion efficiency in chickens. *Appl Microbiol Biotechnol* 2012; 96:1361-9.

[37] Stanley D, Geier MS, Hughes RJ, Denman SE, Moore RJ. Highly variable microbiota development in the chicken gastrointestinal tract. *PLoS One* 2013; 8:e84290.

[38] Zhao L,Wang G, Siegel P, He C, Wang H, ZhaoW. Quantitative genetic background of the host influences gut microbiomes in chickens. *Sci Rep* 2013; 3:1163.

[39] Sergeant MJ, Constantinidou C, Cogan TA, Bedford MR, Penn CW, Pallen MJ. Extensive microbial and functional diversity within the chicken cecal microbiome. *PLoS One* 2014; 9:e91941.

[40] Tang Y, Underwood A, Gielbert A, Woodward MJ, Petrovska L. Metaproteomics analysis reveals the adaptation process for the chicken gut microbiota. *Appl Environ Microbiol* 2014; 80:478-85.

[41] Janczyk P, Halle B, Souffrant WB. Microbial community composition of the crop and ceca contents of laying hens fed diets supplemented with Chlorella vulgaris. *Poult Sci* 2009; 88:2324-32.

[42] Meng H, Zhang Y, Zhao L, Zhao W, He C, Zhai Z. Body weight selection affects quantitative genetic correlated responses in gut microbiota. *PLoS One* 2014; 9:e89862.

[43] Gong J, Si W, Forster RJ, Huang R, Yu H, Yin Y. 16S rRNA gene-based analysis of mucosa-associated bacterial community and phylogeny in the chicken gastrointestinal tracts: from crops to ceca. *FEMS Microbiol Ecol* 2007; 59:147-57.

[44] Sklan D, Shachaf B, Baron J, Hurwitz S. Retrograde movement of digesta in the duodenum of the chick: extent, frequency, and nutritional implications. *J Nutr* 1978; 108:1485-90.

[45] Yang CM, Cao GT, Ferket PR, Liu TT, Zhou L, Zhang L. Effects of probiotic, Clostridium butyricum, on growth performance, immune function, and cecal microflora in broiler chickens. *Poult Sci* 2012; 91:2121-9.

[46] Rehman HU, Vahjen W, Awad WA, Zentek J. Indigenous bacteria and bacterial metabolic products in the gastrointestinal tract of broiler chickens. *Arch Anim Nutr* 2007; 61:319-35.

[47] Stanley D, Hughes RJ, Moore RJ. Microbiota of the chicken gastrointestinal tract: influence on health, productivity and disease. *Appl Microbiol Biotechnol* 2014; 98: 4301-10.

[48] Amit-Romach E, Sklan D, Uni Z. Microflora ecology of the chicken intestine using 16S ribosomal DNA primers. *Poult Sci* 2004; 83:1093-8.

[49] Qu A, Brulc JM, Wilson MK, Law BF, Theoret JR, Joens LA. Comparative Metagenomics reveals host specific metavirulomes and horizontal gene transfer elements in the chicken cecum microbiome. *PLoS One* 2008; 3:e2945.

[50] Waite DW, Taylor MW. Characterizing the avian gut microbiota: membership, driving influences, and potential function. *Front Microbiol* 2014; 5:223.

[51] Videnska P, Faldynova M, Juricova H, Babak V, Sisak F, Havlickova H. Chicken faecal microbiota and disturbances induced by single or repeated therapy with tetracycline and streptomycin. *BMC Vet Res* 2013; 9:30.

[52] Collado MC, Sanz Y. Characterization of the gastrointestinal mucosa-associated Microbiota of pigs and chickens using culture-based and molecular methodologies. *J Food Prot* 2007; 70:2799-804.

[53] Metzler-Zebeli BU, Hooda S, Pieper R, Zijlstra RT, van Kessel AG, Mosenthin R. Nonstarch polysaccharides modulate bacterial microbiota, pathways for butyrate production, and abundance of pathogenic Escherichia coli in the pig gastrointestinal tract. *Appl Environ Microbiol* 2010; 76:3692-701.

[54] Metzler-Zebeli BU, Mann E, Schmitz-Esser S, Wagner M, Ritzmann M, Zebeli Q. Changing dietary calcium-phosphorus level and cereal source selectively alters abundance of bacteria and metabolites in the upper gastrointestinal tracts of weaned pigs. *Appl Environ Microbiol* 2013; 79:7264-72.

[55] Pedersen R, Andersen AD, Molbak L, Stagsted J, Boye M. Changes in the gut microbiota of cloned and non-cloned control pigs during development of obesity: gut microbiota during development of obesity in cloned pigs. *BMC Microbiol* 2013; 13:30.

[56] Arnal ME, Zhang J, Messori S, Bosi P, Smidt H, Lalles JP. Early changes in microbial colonization selectively modulate intestinal enzymes, but not inducible heat shock proteins in young adult swine. *PLoS One* 2014; e87967.

[57] Looft T, Allen HK, Cantarel BL, Levine UY, Bayles DO, Alt DP. Bacteria, phages and pigs: the effects of in-feed antibiotics on the microbiome at different gut locations. *ISME J* 2014; 8:1566-76.

[58] Kim HB, Borewicz K, White BA, Singer RS, Sreevatsan S, Tu ZJ. Microbial shifts in the swine distal gut in response to the treatment with antimicrobial growth promoter, tylosin. *Proc Natl Acad Sci U S A* 2012; 109:15485-90.

[59] Lamendella R, Domingo JW, Ghosh S, Martinson J, Oerther DB. Comparative fecal metagenomics unveils unique functional capacity of the swine gut. *BMC Microbiol* 2011; 11:103.

[60] Buzoianu SG, Walsh MC, Rea MC, O'Sullivan O, Crispie F, Cotter PD. The effect of feeding Bt MON810 maize to pigs for 110 days on intestinal microbiota. *PLoS One* 2012; 7:e33668.

[61] Mann E, Schmitz-Esser S, Zebeli Q, Wagner M, Ritzmann M, Metzler-Zebeli BU. Mucosa-associated bacterial microbiome of the gastrointestinal tract of weaned pigs and dynamics linked to dietary calcium-phosphorus. *PLoS One* 2014; 9:e86950.

[62] Su Y, Bian G, Zhu Z, Smidt H, ZhuW. Early methanogenic colonisation in the faeces of Meishan and Yorkshire piglets as determined by pyrosequencing analysis. *Archaea* 2014; 547908.

[63] Pieper R, Boudry C, Bindelle J, Vahjen W, Zentek J. Interaction between dietary protein content and the source of carbohydrates along the gastrointestinal tract of weaned piglets. *Arch Anim Nutr* 2014; 68:263-80.

[64] Pieper R, Neumann K, Kroger S, Richter JF, Wang J, Martin L. Influence of fermentable carbohydrates or protein on large intestinal and urinary metabolomic profiles in piglets. *J Anim Sci* 2012; 90(Suppl. 4):34-6.

[65] Nielsen KL, Hartvigsen ML, Hedemann MS, Laerke HN, Hermansen K, Bach Knudsen KE. Similar metabolic responses in pigs and humans to breads with different contents and compositions of dietary fibers: a metabolomics study. *Am J Clin Nutr* 2014; 99:941-9.

[66] Hanhineva K, Barri T, KolehmainenM, Pekkinen J, Pihlajamaki J, Vesterbacka A. Comparative non targeted profiling of metabolic changes in tissues and biofluids in highfat diet-fed Ossabaw pig. *J Proteome Res* 2013; 12:3980-92.

[67] Buzoianu SG, Walsh MC, Rea MC, O'Sullivan O, Cotter PD, Ross RP. High-throughput sequence-based analysis of the intestinal microbiota of weanling pigs fed genetically modified MON810 maize expressing Bacillus thuringiensis Cry1Ab (Bt maize) for 31 days. *Appl Environ Microbiol* 2012; 78:4217-24.

[68] Mackie RI. Molecular ecology and diversity in gut microbial ecosystems. Ruminant physiology: digestion, metabolism, growth and reproduction 2000; 2000. p. 61-77.

[69] Carberry CA, Kenny DA, Han S, McCabe MS, Waters SM. Effect of phenotypic residual feed intake and dietary forage content on the rumen microbial community of beef cattle. *Appl Environ Microbiol* 2012; 78:4949-58.

[70] Hess M, Sczyrba A, Egan R, Kim TW, Chokhawala H, Schroth G. Metagenomic discovery of biomass-degrading genes and genomes from cow rumen. *Science* 2011; 331:463-7.

[71] Thorpe A. Enteric fermentation and ruminant eructation: the role (and control?) of methane in the climate change debate. *Clim Change* 2009; 93:407-31.

[72] Brulc JM, Antonopoulos DA, Miller ME, Wilson MK, Yannarell AC, Dinsdale EA. Genecentric metagenomics of the fiber-adherent bovine rumen microbiome reveals forage specific glycoside hydrolases. *Proc Natl Acad Sci U S A* 2009; 106:1948-53.

[73] Hungate RE, Bryant MP, Mah RA. The rumen bacteria and protozoa. *Annu Rev Microbiol* 1964; 18:131-66.

[74] Russell JB, Hespell RB. Microbial rumen fermentation. *J Dairy Sci* 1981; 64:1153-69.

[75] Fernando SC, Purvis II HT, Najar FZ, Sukharnikov LO, Krehbiel CR, Nagaraja TG. Rumen microbial population dynamics during adaptation to a high-grain diet. *Appl Environ Microbiol* 2010; 6:7482-90.

[76] Krause DO, Dalrymple BP, SmithWJ, Mackie RI, McSweeney CS. 16S rDNA sequencing of *Ruminococcus albus* and *Ruminococcus flavefaciens*: design of a signature probe and its application in adult sheep. *Microbiology* 1999; 145(Pt 7):1797-807.

[77] Morgavi DP, Kelly WJ, Janssen PH, Attwood GT. Rumen microbial (meta) genomics and its application to ruminant production. *Animal* 2013; 7:184-201.

[78] Jami E, Mizrahi I. Composition and similarity of bovine rumen microbiota across individual animals. *PLoS One* 2012; 7:e33306.

[79] Kittelmann S, Seedorf H, Walters WA, Clemente JC, Knight R, Gordon JI. Simultaneous amplicon sequencing to explore co-occurrence patterns of bacterial, archaeal and eukaryotic microorganisms in rumen microbial communities. *PLoS One* 2013; 8:e47879.

[80] Fouts DE, Szpakowski S, Purushe J, Torralba M, Waterman RC, MacNeil MD. Next generation sequencing to define prokaryotic and fungal diversity in the bovine rumen. *PLoS One* 2012; 7:e48289.

[81] Pitta DW, Pinchak E, Dowd SE, Osterstock J, Gontcharova V, Youn E. Rumen bacterial diversity dynamics associated with changing from Bermuda grass hay to grazed winter wheat diets. *Microb Ecol* 2010; 59:511-22.

[82] Larue R, Yu Z, Parisi VA, Egan AR, MorrisonM. Novel microbial diversity adherent to plant biomass in the herbivore gastrointestinal tract, as revealed by ribosomal intergenic spacer analysis and rrs gene sequencing. *Environ Microbiol* 2005; 7:530-43.

[83] LiM, Penner GB, Hernandez-Sanabria E, Oba M, Guan LL. Effects of sampling location and time, and host animal on assessment of bacterial diversity and fermentation parameters in the bovine rumen. *J Appl Microbiol* 2009; 107:1924-34.

[84] Ross EM, Moate PJ, Bath CR, Davidson SE, Sawbridge TI, Guthridge KM. High throughput whole rumen metagenome profiling using untargeted massively parallel sequencing. *BMC Genet* 2012; 13:53.

[85] Ferrer M, Ghazi A, Beloqui A, Vieites JM, Lopez-Cortes N, Marin-Navarro J. Functional metagenomics unveils a multifunctional glycosyl hydrolase from the family 43 catalysing the breakdown of plant polymers in the calf rumen. *PLoS One* 2012; 7: e38134.

[86] Lee HJ, Jung JY, Oh YK, Lee SS, Madsen EL, Jeon CO. Comparative survey of rumen Microbial communities and metabolites across one caprine and three bovine groups, using bar-coded pyrosequencing and (1)H nuclear magnetic resonance spectroscopy. *Appl Environ Microbiol* 2012; 78:5983-93.

[87] Zhao S, Zhao J, Bu D, Sun P, Wang J, Dong Z. Metabolomics analysis reveals large effect of roughage types on rumen microbial metabolic profile in dairy cows. Lett *Appl Microbiol* 2014; 59:79-85.

[88] Jami E, White BA, Mizrahi I. Potential role of the bovine rumen microbiome in modulating milk composition and feed efficiency. *PLoS One* 2014; 9:e85423.

[89] Wu S, Baldwin RL, Li W, Li C, Connor EE, Li RW. The bacterial community Composition of the bovine rumen detected using pyrosequencing of 16S rRNA genes. *Metagenomics* 2012; 1 [Article ID: 235571].

[90] Allgaier M, Reddy A, Park JI, Ivanova N, D'Haeseleer P, Lowry S. Targeted discovery of glycoside hydrolases from a switchgrass-adapted compost community. *PLoS One* 2010; 5:e8812.

[91] Wang L, Hatem A, Catalyurek UV, Morrison M, Yu Z. Metagenomic insights into the carbohydrate-active enzymes carried by the microorganisms adhering to solid digesta in the rumen of cows. *PLoS One* 2013; 8:e78507.

CHAPTER III

SECOND MANUSCRIPT

CATCHING THE TIP OF THE ICEBERG – EVALUATION OF SAMPLE PREPARATION PROTOCOLS FOR METAPROTEOMIC STUDIES OF THE RUMEN MICROBIOTA

Simon Deusch[1], Jana Seifert[1]

[1]Institute of Animal Science, University of Hohenheim, Stuttgart, Germany

Published in:
Proteomics (2015)
Volume 15, 2015, Pages 3590-3595
doi: 10.1002/pmic.201400556
Accepted: March 10, 2015

With kind permission of the Wiley-VCH Verlag GmbH & Co. KGaA.
The original publication is available at http://dx.doi.org/10.1002/pmic.201400556.

3. CATCHING THE TIP OF THE ICEBERG – EVALUATION OF SAMPLE PREPARATION PROTOCOLS FOR METAPROTEOMIC STUDIES OF THE RUMEN MICROBIOTA

3.1 Abstract

Various metabolic processes are performed in the rumen caused by a complex microbiota comprising bacteria, archaea, protozoa, and fungi. Thus, the description of the active microbial fraction and their functions are of great interest for animal nutrition, biotechnology, and climatology. Metaproteomic studies of the rumen microbiota are challenged by the need of optimized sample preparation protocols in order to retrieve an enhanced amount of prokaryotic instead of plant- and bovine-derived cells before protein extraction and subsequent LC-MS/MS analysis. The present study evaluates three different protocols applied to the rumen microbiota either attached to plant fibers or present as planktonic cells. The findings of our work suggest the integration of cheesecloth-gauze filtration in sample preparation to achieve a better protein identification ratio. Our data have been deposited to ProteomeXchange with identifier PXD001526 (http://proteomecentral.proteomexchange.org/dataset/PXD001526).

3.2 Introduction

The rumen represents a huge fermentation chamber within the gastrointestinal tract of ruminants hosting various microbial species. The complexity of enzymatic processes to digest food and to produce metabolites absorbed by the host is of interest for the discovery of new enzymatic functions and inhibiting the release of undesirable by-products (e.g. methane emission). So far, the rumen microbial communities and their enzymatic inventories were studied by next-generation sequencing methods and functional metagenomics [1-3] but, to the best of our knowledge, no study of the rumen metaproteome was published before. This is probably due to the manifold compositions of rumen samples, which afford the separation of prokaryotic cells from the residual matter prior to protein extraction, aiming at the highest possible coverage of the present microorganisms and in this way also increase the reliability and the numbers of prokaryotic protein identifications. Fortunately, technical progress in MS and an increasing availability of reference sequences enhance the adequate utilization of metaproteomic approaches, even concerning highly heterogeneous samples [4].

To obtain representative samples, it is likewise important to consider that rumen bacteria attach to different substrates present in the rumen [5] and hence, can be roughly

grouped into (i) free-living bacteria in ruminal fluid, (ii) bacteria loosely attached to feed particles and (iii) firmly adhered species [6,7]. Thereby the particle-associated fractions constitute up to 80% of the total microbial matter [6,8,9] whereas the number of bacteria attached to protozoa, fungi, or the rumen epithelium are negligible [6].

In the past, several studies described optimized methods for isolation of rumen microorganisms including various chemical and physical treatments, e.g. methylcellulose, chilling, or blending, for both, the solid and the fluid fraction of the rumen [10-13]. Here we focused on improving the identification ratio of prokaryotic to eukaryotic proteins. Thus, two sample preparation protocols for the fiber adherent fractions and one protocol to analyze the liquid-associated bacteria and archaea were established and evaluated.

3.3 Materials and Methods

Solid fraction rumen samples and rumen juice were collected from a fistulated Holstein-Frisian cow (approved by Regierungspräsidium Stuttgart; V263/09 TE) 1 h prior to morning feeding at the University of Hohenheim. Samples were frozen immediately at -80°C. The animal was fed twice a day with 4.5 kg grass hay diet containing 12% concentrate (20% soybean, 17% corn, 25% barley, 28% wheat, 4% molasses, 6% minerals).

A workflow of the following protocols is shown in Figure 6. The first protocol (RS-H) investigating the particle-associated ruminal species included Histodenz™ density-gradient centrifugation as this was successfully applied for cell separations from rat intestinal samples before [14]. A total amount of two times 3 g of the frozen, solid rumen digesta were suspended each in 30 mL ice-cold detachment-buffer (0.2 M NaCl; 50 mM Tris-HCl; pH 8) containing 0.1% methylcellulose 400 cP (MC buffer) by vortexing. The suspensions were subjected to a sonication bath for 1 min and shaken horizontally overnight at 4°C. Subsequently, the samples were centrifuged twice for 5 min at 100 × g and 4°C and the supernatants were filtered sequentially through sterile 100 and 40 µm polyethylene (PE) filters. The filtrates were pelleted by centrifugation for 20 min at 12,000 × g and 4°C and stored at -20°C. The frozen pellets were resuspended in 3.5 mL ice-cold detachment buffer by vigorous vortexing, and underlayed with an equal volume of detachment-buffer containing 50% (v/w) Histodenz™. After 45 min centrifugation at 10,000 × g and 4°C the interphases were sampled, pooled, and again centrifuged for 20 min at 12,000 x g and 4°C. The supernatant was discarded and the obtained cell pellet was frozen at -20°C.

The second protocol (RS-C) for the solid rumen fraction utilized common cheesecloth gauze as reported by [12,13]. In total 6 g of frozen, solid rumen digesta were suspended as

described above. The suspension was subjected to a sonication bath for 1 min and shaken horizontally for 3 h at 4°C. Subsequently, the sample was squeezed through two layers of sterile gauze and washed twice with 25 mL detachment buffer while squeezing repeatedly. The filtrate was centrifuged for 15 min at 200 × g and 4°C, the supernatant was filtered through a sterile 40 µm PE filter and centrifuged for 20 min at 12,000 × g and 4°C to finally store the cells at -20°C.

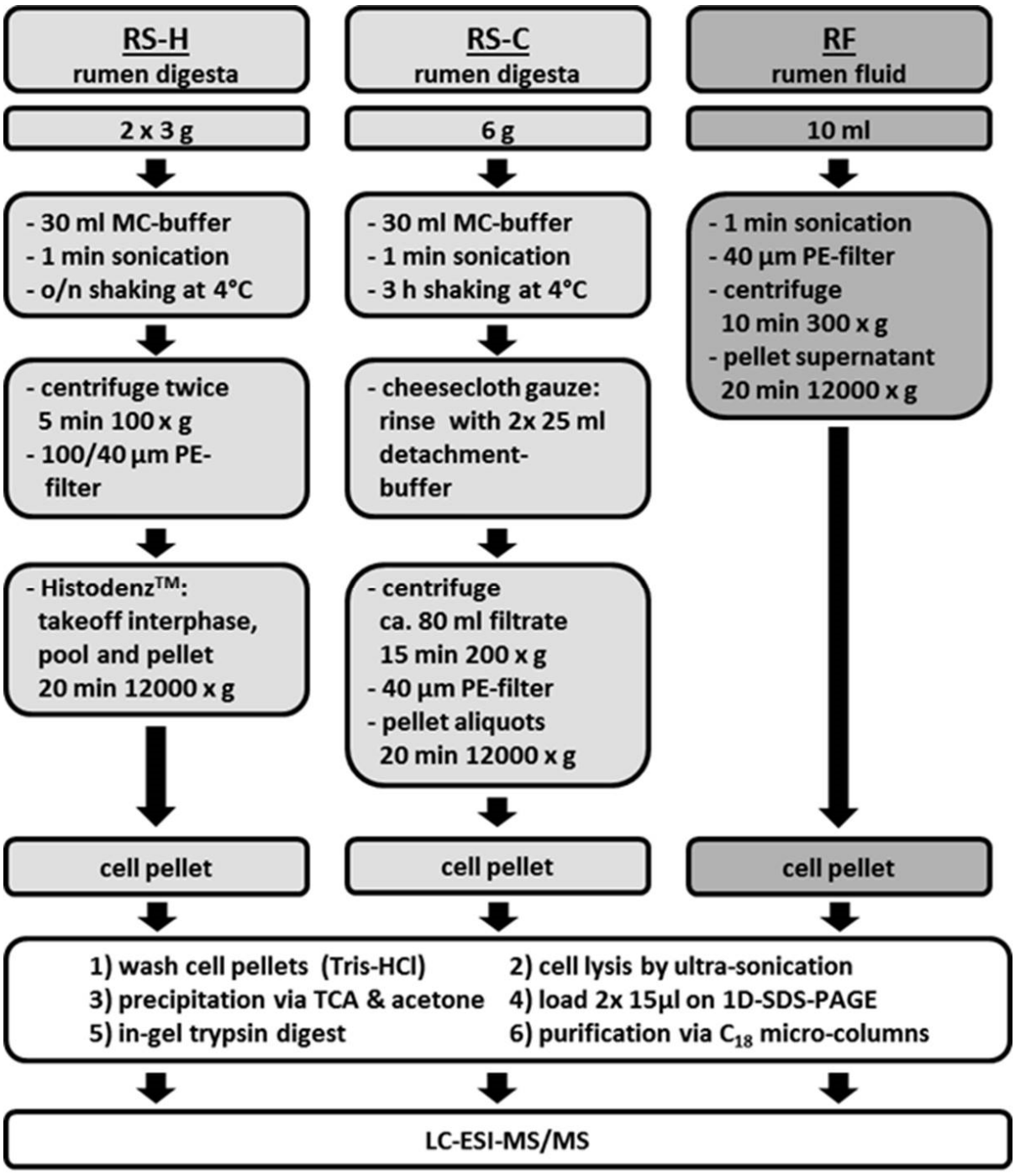

Figure 6| Scheme of applied workflows to process samples from rumen solid (RS) and rumen fluid (RF).

The protocol (RF) for the planktonic cells subjected 10 mL of rumen fluid to 1 min sonication bath and 1 min vortexing prior to filtering through a sterile 40 µm PE filter. The filter surface was carefully scratched with a sterile spatula to prevent clogging. The filtrate was centrifuged for 10 min at 300 × g and 4°C and 2 mL of the supernatant were pelleted for 20 min at 12,000 × g and 4°C. Finally, the supernatant was discarded and the pellet was frozen at -20°C.

Cell pellets obtained from each protocol (RS-H, RS-C, RF) were washed three times in 500 µL 50 mM Tris-HCl (pH 7.5; 0.1 mg/mL chloramphenicol; 1 mM PMSF), resuspended in 300 µL of 20 mM Tris-HCl (pH 7.5; 2% SDS), and shaken in a Thermo-Mixer (Eppendorf) for 10 min at 60°C and 1,200 rpm, and 1 mL 20 mM Tris-HCl (pH 7.5; 0.1 mg/mL $MgCl_2$; 1 mM PMSF; 1 µL/mL Benzonase) was added and cells were lysed by ultrasonication on ice, four times 2 min (amplitude 60%; cycle 0.5), followed by shaking for 10 min at 37°C and 1,200 rpm. Samples were centrifuged at 10,000 × g and 4°C for 10 min, proteins in the supernatant were precipitated using 20% TCA at 4°C for 30 min and centrifuged for 15 min at 12,000 × g and 4°C. Protein pellets were washed twice in ice-cold acetone and dried by vacuum centrifugation. Prior to 1D-SDS-PAGE (12%, 20 min at 20 mA and 30 min at 40 mA) pellets were resuspended in 35 µL Laemmli-buffer by 5 min sonication bath and vortexing. Samples were incubated for 5 min at 95°C to reduce disulfide bonds. Two times 15 µL were applied on the gel. Each lane representing one sample was cut into five equal-sized pieces and subjected to in-gel trypsin digestion overnight [15]. Obtained peptides were purified and desalted by C_{18} microcolumns (Millipore), dried and prior to LC-MS/MS measurement reconstituted in 20 µL 0.1% TFA.

Out of each sample, 5 µL were injected, ionized peptides were analyzed and fragmented by a Q-Exactive Plus mass spectrometer (Thermo Fisher Scientific). Details of the data acquisition are described in the Supporting Information (Chapter 3.7).

Raw MS and MS/MS data were processed by Thermo Proteome Discoverer software (v. 1.4.1.14), Mascot (v. 2.4), and independently searched against the NCBInr databases (v. July 12, 2014) for bacteria (~31 × 10^6 sequences), archaea (~8.5 × 10^5 sequences), and eukaryota (~12 × 10^6 sequences). Oxidation of methionine was set as variable modification and carbamidomethylation of cysteine as fixed modification. Precursor ion tolerance was defined at 10 ppm and fragment ion tolerance to 0.02 Da. Furthermore, all peaks besides the top 12 peaks per 100 Da in each MS/MS were removed to denoise spectra before identification. Using Thermo Proteome Discoverer, the default filter was set to one peptide per protein and a MASCOT significance threshold of 0.05 including a peptide, protein, and

PSM FDR below 1%. Protein grouping was enabled with a minimum peptide confidence of medium and a delta Cn better than 0.15. Strict maximum parsimony principle was applied. The MS proteomics data have been deposited to the ProteomeXchange Consortium (http://www.proteomexchange.org) via the PRIDE partner repository [16] with the dataset identifier PXD001526.

Evaluation of performance was done in determining the number of possible peptide and protein identifications, the ratio between prokaryotic and eukaryotic proteins, and in defining the bias toward specific phylogenetic groups as well as the pI, protein mass, subcellular localization, and cluster of orthologous groups (COG) analysis.

3.4 Results

The RS-C protocol using cheesecloth-gauze yielded the highest number of total, as well as archaeal and bacterial protein identifications with up to approximately 2,300 prokaryotic proteins, respectively approximately 4,800 peptides identified in one of the duplicates. Comparing the protocols for the fiberadherent fraction, the RS-C protocol reached not only higher numbers of total identifications, but also showed a 25% increase for the identification ratio of prokaryotic to eukaryotic proteins. On the other hand, the RF protocol for the rumen fluid showed the lowest number of total protein identifications, but as expected, resulted in the best ratio of prokaryotic to eukaryotic protein identifications since the sample matter contains less plant-derived material (Table 2).

None of the protocols biased toward a specific protein weight or pI range, nor were there any disparities concerning the numbers of peptide-spectra matches (Table 2).

Phylogenetic analyses were performed using Unipept [17]. The overall distribution of bacterial groups is in consistency with literature reports and recent 16S rDNA sequencing data of other rumen studies [1-3,18]. Slight distinctions in taxonomic distribution appeared due to varying numbers of protein identifications or different sample compositions, ergo solid, and liquid rumen fractions. As assumed, the more proteins identified the more precise is the subsequent phylogenetic analysis. This might explain the divergence in bacterial species distribution in between the two RS protocols (Figure 7A), as well as most of the negligible changes in archaeal community structure across all protocols and replicates (Figure 7B). Mainly *Prevotella* sp. belonging to the phyla of Bacteroidetes are known to be more abundant in rumen fluid, whereas *Fibrobacter succinogenes* and several gram-positive families belonging to the phyla of Firmicutes are more numerous in solid fractions [18] which is in consistency with our findings (Figure 7A and C). In 2004, Shin et al. reported, based on 16S

rDNA gene sequences, that members of the family Methanomicrobiaceae are the predominant archaeal species in rumen fluid [19]. This supports our data pointing toward an increased abundance of Methanomicrobiaceae in the fluid rumen fraction (Figure 7B). Eukaryotic phylogeny was determined at phyla level as most peptides identified belong to the group of viridiplantae and only negligible information concerning protozoal and bovine proteins is expected. Generally, using publicly available databases, the taxonomic association of proteomic data is always restricted to known organisms. Proteins of unknown or poorly characterized species will not be detected by metaproteomic approaches.

Table 2| Number (#) of protein and peptide identifications, average (Ø) pI, molecular weight (MW) and peptide-spectra matches (PSM) including standard deviation (σ) of each protocol for identifications based on bacteria, archaea or eukaryota sequence databases. The average identification ratios of prokaryotic to eukaryotic proteins are given in the third column.

			Protein-IDs (#)	**Peptide-IDs** (#)	**Ø Ratio Protein-IDs** (BacArc : Euk)	**Ø pI** (pH)	σ pI	**Ø MW** (kDa)	σ MW	**Ø PSM** (#)	σ PSM
RS-H	Bacteria	A	1379	3141	2.02	6.98	2.09	37.85	27.16	5.33	7.29
		B	1336	2935		6.95	2.04	39.09	28.47	5.53	8.27
	Archaea	A	96	1798		6.20	1.50	43.94	21.32	5.08	6.50
		B	101	1774		6.32	1.54	43.22	21.61	4.94	6.31
	Eukaryota	A	742	2336		7.30	1.85	39.81	40.05	15.64	27.73
		B	701	2067		7.15	1.81	38.92	41.03	12.12	17.94
RS-C	Bacteria	A	2138	4792	2.50	7.16	2.15	34.07	27.06	5.57	8.54
		B	1925	4390		7.14	2.16	33.83	29.70	5.16	7.73
	Archaea	A	159	2531		6.35	1.73	41.85	24.38	4.92	7.05
		B	150	2524		6.51	1.71	41.75	23.05	4.73	6.11
	Eukaryota	A	929	2781		7.33	1.81	38.23	42.28	12.43	19.13
		B	818	2618		7.41	1.83	38.87	42.90	10.68	17.82
RF	Bacteria	A	1399	3162	4.01	7.06	2.03	38.87	28.79	6.78	12.07
		B	1266	2921		7.10	2.03	36.79	26.28	6.99	12.46
	Archaea	A	134	1434		6.27	1.47	45.52	24.05	5.30	8.34
		B	151	1453		6.35	1.51	40.81	21.39	5.44	8.40
	Eukaryota	A	323	1007		6.92	1.73	48.80	68.05	11.92	29.65
		B	412	1197		7.04	1.78	44.63	48.07	12.92	26.90

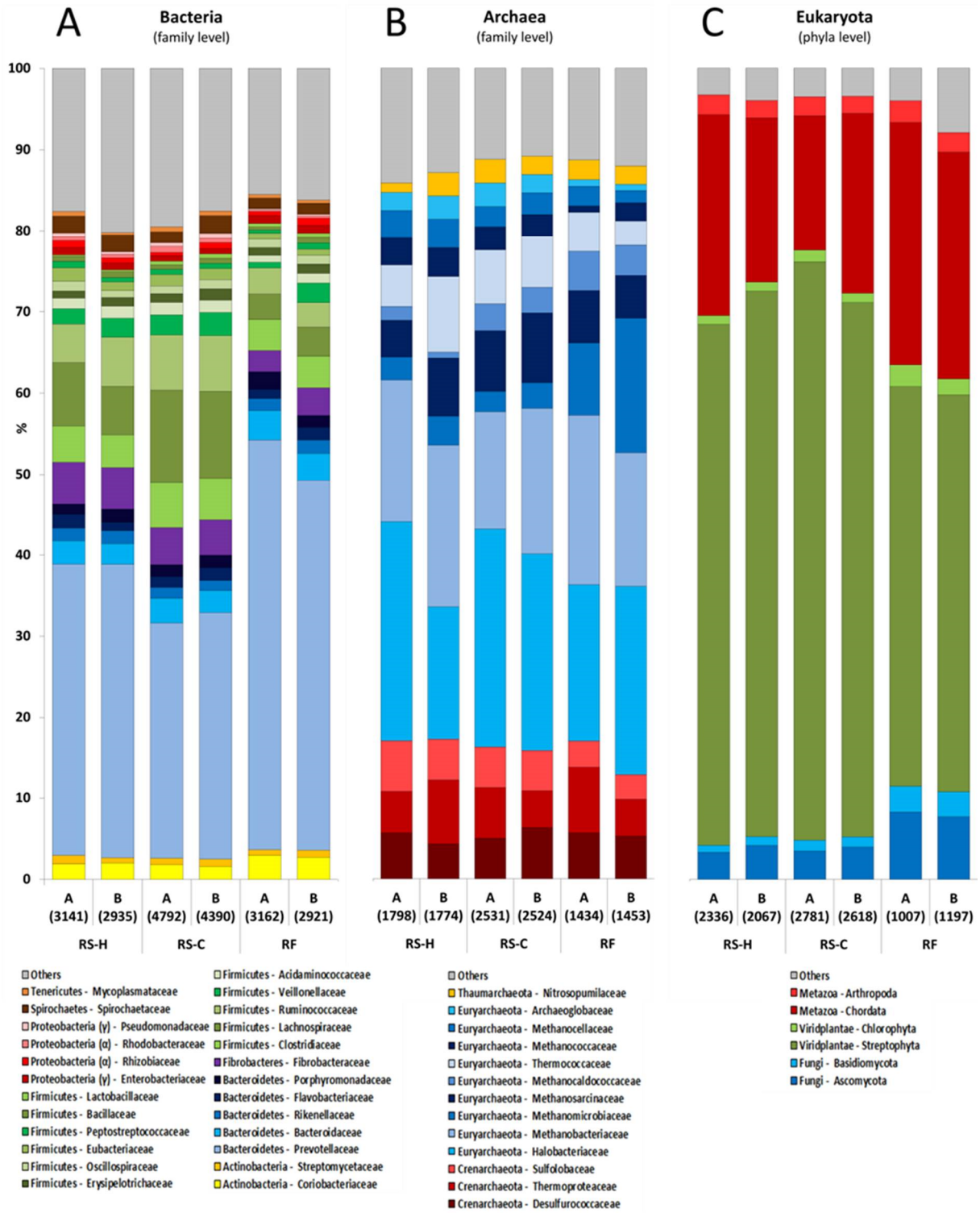

Figure 7| The phylogenetic distribution (Unipept) of each protocol (RS-H, RS-C, RF) and repetition is shown at family level for bacteria (**A**) and archaea (**B**); eukaryota (**C**) were analyzed at phyla level. The numbers of total peptide identifications are given in brackets.

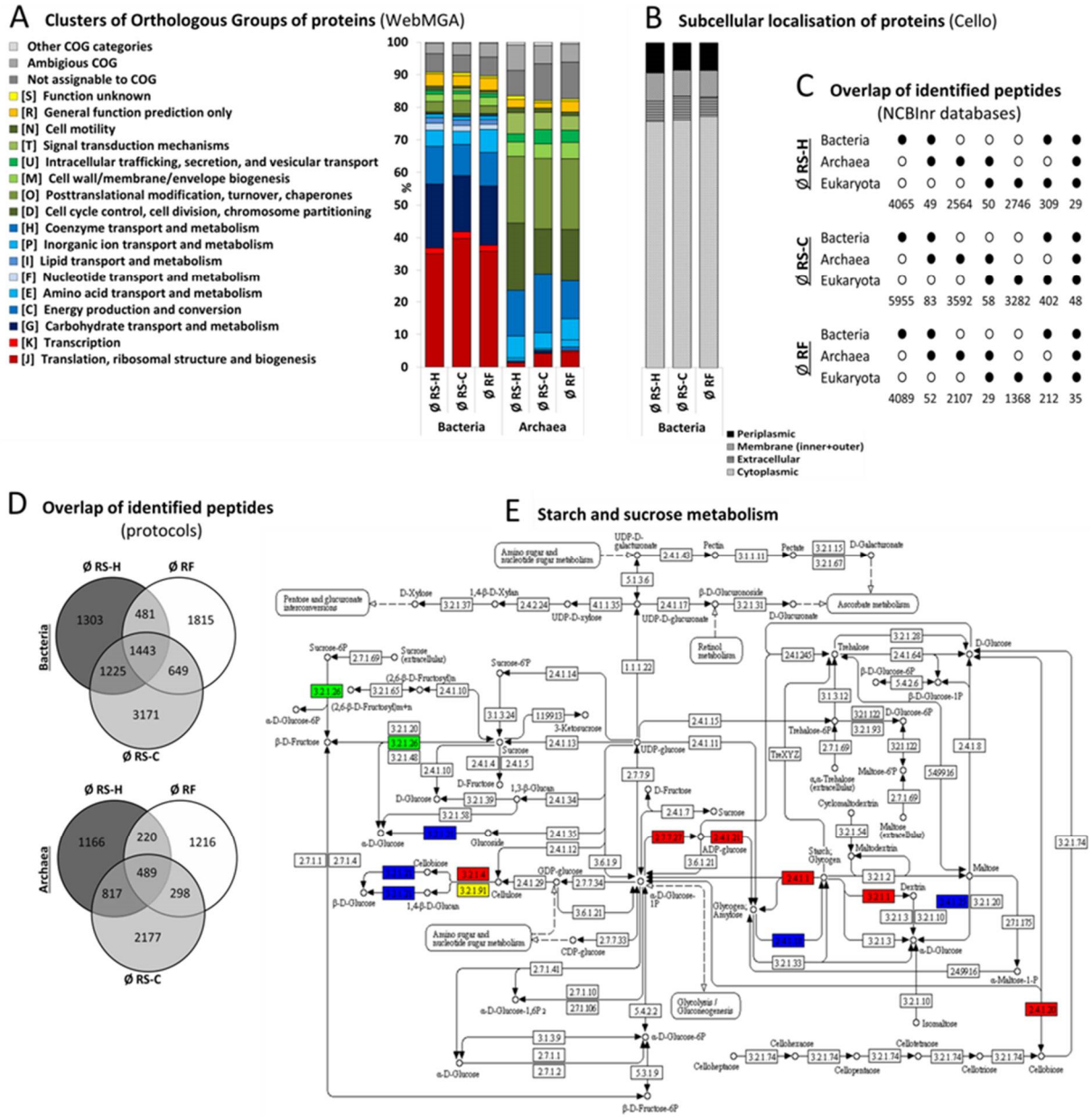

Figure 8| Clusters of orthologous groups of archaeal and bacterial proteins (**A**) identified in average (Ø) with each protocol are given in percentage (WebMGA). The subcellular localizations of identified bacterial proteins (Cello) are shown in the middle (**B**). Overlap of peptide identifications across all NCBInr databases is indicated by dot plots on the right side (**C**) while overlap of prokaryotic peptide identifications across protocols is shown by Venn diagrams (**D**). Pathway-map of starch and sucrose metabolism (**E**). CAZymes active in carbohydrate metabolisms were identified in all protocols as highlighted in red. Enzymes identified in both rumen solid protocols are shown in yellow. Proteins found exclusively in the RS-C and RF are marked in blue and the green-labeled enzymes are derived from rumen fluid.

The prokaryotic protein functions were grouped according to COG using WebMGA server [20] with an e-value cut-off of 10^{-3} keeping the best hit. There are no differences in COG grouping between protocols, but across kingdoms (Figure 8A). The most frequent bacterial protein functions are involved in carbohydrate transport and metabolism [G], energy production and conversion [C], and translation processes [J]. In contrast, archaeal proteins were mainly active in cell cycle control, cell division [D], and PTM [O] (Figure 8A). Sample preparation had no impact on the subcellular localization of proteins (Figure 8B) and all NCBInr database searches identified mostly unique peptides (Figure 8C). The overlap in prokaryotic peptide identifications across protocols is shown by Venn diagrams (Figure 8D). Bacterial carbohydrate-active enzymes (CAZymes) [21] were found in all samples (Figure 8E) and grouped mainly into different glycosyl hydrolase families of the identified *Prevotella* and *Ruminococcus* species.

3.5 Conclusion

Although the identification ratio of proteins was improved, the total number of identifications still only depicts a sparse part of the actually active microbiota, elucidating the obstacle in obtaining representative metaproteome data. Nonetheless, future attempts to improve proteome coverage of such complex samples should include extended LC-MS/MS analysis time and the use of a sample-specific metagenomic databases.

3.6 Acknowledgments

We would like to thank the Carl-Zeiss-Stiftung for funding J.S. and S.D. and the PRIDE team for their support. Special thanks to J. Pfannstiel and W. X. Schulze for their great support.

3.7 Supporting Information

Of each sample 5 µL were injected by the auto-sampler and peptides were eluted into a separation column (Easy-Spray Pepmap, C_{18}, 75 µm x 50 cm (RS-H) / 25 cm (RS-C) / 15 cm (RF), 2µm, Thermo Fisher Scientific). Chromatography was performed with 0.1% formic acid in solvent A (100% water) and B (100% acetonitrile). For the solid fractions (RS-H and RS-C) the solvent B gradient was set from 1 to 40% during the first 115 min and subsequently increased to 90% within 5 min plus an additional 5 min at 90% using a nano-high pressure liquid chromatography system (Ultimate3000 UHPLC, Thermo Fisher Scientific). For the RF protocol, the solvent B gradient was set from 1% to 40% during the first 115 min and

subsequently increased to 45% within 15 min followed by 2 min to 90% plus an additional 5 min at 90%. Ionized peptides were analyzed and fragmented by a Q-Exactive Plus mass spectrometer (Thermo Fisher Scientific). For analysis continuous scanning of eluted peptide ions was carried out between 300-1600 m/z automatically switching to MS/MS higher energy collisional dissociation (HCD) mode and twelve MS/MS events per survey scan. For MS/MS HCD measurements the dynamic precursor exclusion was set to 20 s and peptide match was enabled.

3.8 References

[1] Brulc, J. M., Antonopoulos, D. A., Miller, M. E., Wilson, M. K., *et al.*, Gene-centric metagenomics of the fiber-adherent bovine rumen microbiome reveals forage specific glycoside hydrolases. *Proc. Natl. Acad. Sci. US A.* 2009, *106*, 1948-1953.

[2] Ferrer, M., Ghazi, A., Beloqui, A., Vieites, J. M., *et al.*, Functional metagenomics unveils a multifunctional glycosyl hydrolase from the family 43 catalysing the breakdown of plant polymers in the calf rumen. *PLoS One* 2012, *7*, e38134.

[3] Hess, M., Sczyrba, A., Egan, R., Kim, T. W., *et al.*, Metagenomic discovery of biomass-degrading genes and genomes from cow rumen. *Science* 2011, *331*, 463-467.

[4] Seifert, J., Herbst, F.-A., Nielsen, P. H., Planes, F. J., *et al.*, Bioinformatic progress and applications in metaproteogenomics for bridging the gap between genomic sequences and metabolic functions in microbial communities. *Proteomics* 2013, *13*, 2786-2804.

[5] McAllister, T. A., Bae, H. D., Jones, G. A., Cheng, K. J., Microbial attachment and feed digestion in the rumen. *J. Anim. Sci.* 1994, *72*, 3004-3018.

[6] Miron, J., Ben-Ghedalia, D., Morrison, M., Invited review: adhesion mechanisms of rumen cellulolytic bacteria. *J. Dairy Sci.* 2001, *84*, 1294-1309.

[7] Cheng, K. J., McAllister, T. A., in: Hobson, P. N., Stewart, C. S. (Eds.), *The Rumen Microbial Ecosystem*, Springer Netherlands 1997, pp. 492-522.

[8] Forsberg, C. W., Lam, K., Use of adenosine 5'-triphosphate as an indicator of the microbiota biomass in rumen contents. *Appl. Environ. Microbiol.* 1977, *33*, 528-537.

[9] Craig, W. M., Broderick, G. A., Ricker, D. B., Quantitation of microorganisms associated with the particulate phase of ruminal ingesta. *J. Nutr.* 1987, *117*, 56-62.

[10] Leedle, J. A. Z., Barsuhn, K., Butine, T. J., Comparison of treatments for the in-vitro detachment of cellulolytic bacteria from ruminal digesta samples. *Curr. Microbiol.* 1987, *15*, 129-132.

[11] Ranilla, M. J., Carro, M. D., Diet and procedures used to detach particle-associated microbes from ruminal digesta influence chemical composition of microbes and estimation of microbial growth in Rusitec fermenters. *J. Anim. Sci.* 2003, *81*, 537-544.

[12] Trabalza-Marinucci, M., Poncet, C., Delval, E., Fonty, G., Evaluation of techniques to detach particle-associated microorganisms from rumen contents. *Anim. Feed Sci. Tech.* 2006, *125*, 1-16.

[13] Whitehouse, N. L., Olson, V. M., Schwab, C. G., Chesbro, W. R.*, et al.*, Improved techniques for dissociating particle-associated mixed ruminal microorganisms from ruminal digesta solids. *J. Anim. Sci.* 1994, *72*, 1335-1343.

[14] Haange, S. B., Oberbach, A., Schlichting, N., Hugenholtz, F.*, et al.*, Metaproteome analysis and molecular genetics of rat intestinal microbiota reveals section and localization resolved species distribution and enzymatic functionalities. *J.Proteome Res.* 2012, *11*, 5406-5417.

[15] Jehmlich, N., Schmidt, F., Hartwich, M., von Bergen, M.*, et al.*, Incorporation of carbon and nitrogen atoms into proteins measured by protein-based stable isotope probing (Protein-SIP). *Rapid Comm. Mass Sp.* 2008, *22*, 2889-2897.

[16] Vizcaino, J. A., Cote, R. G., Csordas, A., Dianes, J. A.*, et al.*, The PRoteomics IDEntifications (PRIDE) database and associated tools: status in 2013. *Nucleic Acids Res.* 2013, *41*, D1063-1069.

[17] Mesuere, B., Debyser, G., Aerts, M., Devreese, B. et al., The Unipept metaproteomics analysis pipeline. *Proteomics* 2014. doi: 10.1002/pmic.201400361

[18] Pitta, D. W., Pinchak, E., Dowd, S. E., Osterstock, J.*, et al.*, Rumen bacterial diversity dynamics associated with changing from bermudagrass hay to grazed winter wheat diets. *Microb. Ecol.* 2010, *59*, 511-522.

[19] Shin, E. C., Choi, B. R., Lim, W. J., Hong, S. Y.*, et al.*, Phylogenetic analysis of archaea in three fractions of cow rumen based on the 16S rDNA sequence. *Anaerobe* 2004, *10*, 313-319.

[20] Wu, S., Zhu, Z., Fu, L., Niu, B., Li, W., WebMGA: a customizable web server for fast metagenomic sequence analysis. *BMC Genomics* 2011, *12*, 444.

[21] Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., Henrissat, B., The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 2014, *42*, D490-495.

CHAPTER IV

THIRD MANUSCRIPT

A STRUCTURAL AND FUNCTIONAL ELUCIDATION OF THE RUMEN MICROBIOME INFLUENCED BY VARIOUS DIETS AND MICROENVIRONMENTS

Simon Deusch[1], Amélia Camarinha-Silva[1], Jürgen Conrad[2], Uwe Beifuss[2], Markus Rodehutscord[1], Jana Seifert[1]

[1]Institute of Animal Science, University of Hohenheim, Stuttgart, Germany
[2]Institute of Chemistry, University of Hohenheim, Stuttgart, Germany

Published in:
Frontiers in Systems Microbiology (2017)
Volume 8, 2017, Article 1605
doi: 10.3389/fmicb.2017.01605
Accepted: August 07, 2017

Open access.
The original publication is available at https://doi.org/10.3389/fmicb.2017.01605.

4. A STRUCTURAL AND FUNCTIONAL ELUCIDATION OF THE RUMEN MICROBIOME INFLUENCED BY VARIOUS DIETS AND MICROENVIRONMENTS

4.1 Abstract

The structure and function of the microbiome inhabiting the rumen are, amongst other factors, mainly shaped by the animal's feed intake. Describing the influence of different diets on the inherent community arrangement and associated metabolic activities of the most active ruminal fractions (bacteria and archaea) is of great interest for animal nutrition, biotechnology, and climatology. Samples were obtained from three fistulated Jersey cows rotationally fed with corn silage, grass silage or grass hay, each supplemented with a concentrate mixture. Samples were fractionated into ruminal fluid, particle-associated rumen liquid, and solid matter. DNA, proteins and metabolites were analyzed subsequently. DNA extracts were used for Illumina sequencing of the 16S rRNA gene and the metabolomes of rumen fluids were determined by 500 MHz-NMR spectroscopy. Tryptic peptides derived from protein extracts were measured by LC-ESI-MS/MS and spectra were processed by a two-step database search for quantitative metaproteome characterization. Data are available via ProteomeXchange with the identifier PXD006070. Protein- and DNA-based datasets revealed significant differences between sample fractions and diets and affirmed similar trends concerning shifts in phylogenetic composition. Ribosomal genes and proteins belonging to the phylum of Proteobacteria, particularly Succinivibrionaceae, exhibited a higher abundance in corn silage-based samples while fiber-degraders of the Lachnospiraceae family emerged in great quantities throughout the solid phase fractions. The analysis of 8,163 quantified bacterial proteins revealed the presence of 166 carbohydrate active enzymes in varying abundance. Cellulosome affiliated proteins were less expressed in the grass silage, glycoside hydrolases appeared in slightest numbers in the corn silage. Most expressed glycoside hydrolases belonged to families 57 and 2. Enzymes analogous to ABC transporters for amino acids and monosaccharides were more abundant in the corn silage whereas oligosaccharide transporters showed a higher abundance in the fiber-rich diets. Proteins involved in carbon metabolism were detected in high numbers and identification of metabolites like short-chain fatty acids, methylamines and phenylpropionate by NMR enabled linkage between producers and products. This study forms a solid basis to retrieve deeper insight into the complex network of microbial adaptation in the rumen.

4.2 Introduction

Ruminant livestock with about 3.6 billion farm animals globally (Hackmann and Spain, 2010) represents an important source of human food since these animals have the ability to convert plant-derived non-starch polysaccharides, indigestible for humans, to usable food products in form of milk and meat. This includes also an undesired side effect as ruminants release a substantial portion of methane, a potential greenhouse gas, to the atmosphere (McMichael et al., 2007). The underlying metabolic processes are driven by a complex microbial network consisting of archaea, bacteria, fungi, and protists residing in the strictly anaerobic rumen (Hungate, 1966; Mackie, 2000). Composition and activity of the rumen microbiome are, among other factors, primarily shaped by the diet (Ley et al., 2008; Henderson et al., 2015) and play an important role regarding the animals' health (Russell and Rychlik, 2001; Gressley et al., 2011) as well as feed efficiency and emission of environmentally harmful substances (Mizrahi, 2011; Shabat et al., 2016). Furthermore, the huge amount of fiber-degrading enzymes expressed in the rumen serves as a unique resource for the discovery of new lignocellulolytic enzymes useful for biofuel production (Brulc et al., 2009; Hess et al., 2011; Ferrer et al., 2012).

Bacteria, the most abundant (Krause and Russell, 1996; Mackie, 2000), diverse (McSweeney et al., 2005) and metabolically active species in the rumen are mainly responsible for the degradation and fermentation of plant fibers and proteins ingested by the animals (Hungate, 1966; Brulc et al., 2009). Bacterial species attached to feed particles constitute up to 75% of the total microbial population (McAllister et al., 1994; Koike et al., 2003). Others are free floating in the rumen fluid or live associated to fungi, protists, and the rumen epithelium (McAllister et al., 1994; Miron et al., 2001). Besides, bacteria can be classified according to their functional potential as there are, amongst others, fibrolytic, amylolytic, proteolytic, and saccharolytic species. Generally, starch and sugar degraders constitute the largest part of the ruminal bacterial population and are of great importance since diets for high-producing ruminants usually contain large amounts of readily fermentable starch and sugars. Despite their undoubtable significance, bacteria specialized for fiber degradation are typically less present (Puniya et al., 2015). Degradation of the entire organic matter taken in by the host animals cannot be achieved by a single organism but requires the functional capacities and cooperation of a succession of many microorganisms (Bladen et al., 1961). Hence, to obtain energy, bacterial communities interact synergistically in building diverse fibrolytic enzymes that finally yield in the production of short chain fatty acids and microbial protein which serve as the main energy and amino acid sources for the host

(Hungate, 1966; Mackie, 2002).

Various methods have been employed to study the rumen microbiome ranging from classical cultivation (Bryant, 1959; Hungate et al., 1964) to molecular approaches including next generation sequencing (Edwards et al., 2004; Mackie and Cann, 2005; Creevey et al., 2014) and functional metagenomics (Brulc et al., 2009; Hess et al., 2011; Ferrer et al., 2012) as well as metabolomics (Ametaj et al., 2010; Saleem et al., 2013). Furthermore, investigations of diet induced shifts in microbial community composition of the rumen in different contexts are numerous (Tajima et al., 2001; Fernando et al., 2010; Kong et al., 2010; Pitta et al., 2010; de Menezes et al., 2011; Ann Huws et al., 2012; Belanche et al., 2012; Carberry et al., 2012; Petri et al., 2013; Thoetkiattikul et al., 2013; Zhang et al., 2013; Lengowski et al., 2016) but generally rumen studies are restricted to nucleic acids-based approaches with limited functional insights.

Bioinformatic and technical progress in mass spectrometry as well as a growing availability of reference sequences facilitate metaproteomic studies yielding increased information about taxonomic diversity, actual functional profiles, and interactions of the most active fractions of the investigated microbiota (Hettich et al., 2013; Seifert et al., 2013; Muth et al., 2016; Tanca et al., 2016). However, investigations of the prokaryotic rumen metaproteome are challenged by the complexity of rumen samples, which requires specific sample preparation procedures to separate archaeal and bacterial cells from the residual matter prior to protein extraction. Moreover, humic compounds are present and interfere with the metaproteomic workflow (Chourey et al., 2010; Heyer et al., 2013). Nevertheless, LC-ESI-MS/MS-based rumen studies have already been implemented successfully (Deusch and Seifert, 2015). The combination of different up to date *Omics*-technologies represents the most powerful tool to analyze the microbiome of complex ecosystems like the rumen (Lamendella et al., 2012; Deusch et al., 2015), but studies of the ruminal prokaryotic communities that include multiple approaches and state of the art methods are rare.

So far, to the best of our knowledge, there are no publications investigating the impact of the most common forages used as feed for dairy cows and fattening cattle on the metaproteome expressed by the entirety of archaeal and bacterial communities in the different phases of the rumen ecosystem. Complementary, structural and functional information obtained from the mass spectrometry-based analysis of the rumen metaproteome targeting the most metabolically active prokaryotes was further supplemented with Illumina MiSeq sequencing of the 16S rRNA gene that includes all cells present. Additionally, metabolome patterns were investigated by 500 MHz NMR spectroscopy. The aim of this investigation was

to provide deeper insights into the complicated microbial network of the rumen ecosystem and its response to different animal diets to improve efficiency in animal production.

4.3 Methods

4.3.1 Ethics Statement

The animals of this study were kept according to the German Animal Welfare legislation at the Agricultural Experiment Station Meiereihof of the University of Hohenheim in Stuttgart, Germany. The experimental procedures and treatments were authorized by the Regierungspräsidium Stuttgart in Germany as previously reported (Lengowski et al., 2016).

4.3.2 Animals and Diets

To access the dietary and host-related impact on the rumen microbiome a Latin square design using three rumen cannulated lactating Jersey cows was applied. Animals were fed rotationally with three different diets for *ad libitum* consumption and free access to drinking water. Feed was given once daily at 7.30 a.m. Based on dry matter, diets consisted of 52% concentrate mixture and 48% of either corn silage, grass silage, or grass hay. The concentrate was composed of 19% wheat, 19% barley, 7% soybean meal, 6% molasses, and 1% vitamin mineral premix. The corn silage-based diet was supplemented with urea to obtain a balanced nitrogen content in comparison to the grass silage- and grass hay-based diets. The chemical characteristics of the experimental diets are shown in Table 3.

Table 3 | The measured chemical characteristics of the three forage sources and the thereof calculated properties of the final total mixed rations fed to the cows. CS, corn silage; GS, grass silage; H, grass hay.

	Basic forages			**Total mixed rations**		
Components (% of dry matter)	**CS**	**GS**	**H**	**CS diet**	**GS diet**	**H Diet**
Dry matter (%)	38.3	61.8	97.0	63.4	60.6	71.4
Crude ash	4.1	11.9	8.8	5.3	8.9	7.4
Ether extract	3.2	3.7	2.5	2.8	3.2	2.6
Crude protein	7.8	12.8	13.1	14.2	15.3	15.4
Neutral detergent fiber (organic)	41.3	52.9	58.9	27.2	33.2	36.0
Acid detergent fiber (organic)	21.0	31.8	33.6	13.3	18.8	19.6
Acid detergent lignin	1.7	2.6	3.0	1.0	1.4	1.6
Non-fiber carbohydrates	43.7	18.7	16.7	50.5	39.5	38.6

4.3.3 Sampling

Samples were taken 5 h after feeding at 12.30 p.m. with a preceding adaptation time of 20 days for each diet. A quantity of 200 g of rumen matter was taken from five different positions (cranial, caudal, dorsal, ventral, medial) each and squeezed vigorously by hand using disposable polyethylene gloves to obtain the particle-associated liquid phase (LP) sample fraction, the remains constituted the solid phase (SP) sample fraction. Equal parts of the obtained LP and SP sample fractions were pooled across the five rumen positions. As a third sample fraction ventral rumen fluid (RF) was collected using a vacuum pump. Two times 27 samples (of three cows, diets and sample fractions) with 40 ml of the RF and LP sample fractions and 20 g of the SP sample fractions were frozen immediately at -80°C until further processing. The bacterial populations in the rumen can be subdivided into planktonic species, free-living in the RF and the fiber-adherent communities which can be further separated into groups of loosely and tightly attached species supposed to be present in the LP and the SP sample fractions, respectively (McAllister et al., 1994).

4.3.4 Illumina Amplicon Sequencing

Using the FastDNA™ SPIN Kit for Soil (MP Biomedical, Solon, OH, USA) 250 mg of defrosted and vortexed samples were used for DNA extraction according to the manufacturer's instruction with slight modifications as described in Burbach et al. (2016). The extraction protocol included a bead-beating step for improved mechanical disruption of Gram-positive bacteria as suggested by Henderson et al. (2013). Quality and purity of DNA extracts were analyzed using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). Illumina library preparation by PCR amplification of the V1-2 region of the 16S rRNA gene was performed as reported recently (Camarinha-Silva et al., 2014). The archaeal community was amplified using the previously described primers Arch349 and Arch806 (Lee et al., 2012). The forward primer contained a 6-nt barcode, a 2-nt linker and both primers comprised sequences complementary to the Illumina specific adaptors (Camarinha-Silva et al., 2014). The PCR mixture of a total volume of 20 µl contained PrimeSTAR HS DNA polymerase (2.5 U, Clontech Laboratories, Mountain View, CA, USA), 2.5 mM dNTP mixture, 0.2 µM primers and 1 µl of template DNA. An initial denaturation at 95°C for 3 min was followed by 20 cycles of denaturation at 98°C for 10 s, annealing at 59°C for 10 s, extension at 72°C for 45 s and a final extension for 2 min at 72°C. One microliter of the PCR product was used for a second PCR (15 cycles) under the same conditions with the reverse primer containing a sequence that integrated the Illumina multiplexing sequence and

Illumina index primers (Camarinha-Silva et al., 2014). Integrity of amplicons was analyzed by gel electrophoresis, purified, and normalized using SequalPrep Normalization Kit (Invitrogen Inc., Carlsbad, CA, USA). Samples were pooled and sequenced using 250 bp paired-end sequencing chemistry on an Illumina MiSeq platform. Sequences were processed using the MOTHUR software pipeline (Kozich et al., 2013). Sequences were excluded if they had any primer or barcode mismatch or *N* character, aligned, checked for chimeras using UCHIME (Edgar et al., 2011) and clustered into operational taxonomic units (OTUs) at ≥ 97% similarity. Low abundance OTUs (< 0.05% of total reads) were removed and a total of 1,484 bacterial and 626 archaeal phylotypes were taxonomically assigned using the naïve Bayesian RDP classifier (Wang et al., 2007) and the RDP database (Cole et al., 2014). Sequences were submitted to European Nucleotide Archive under the study accession number PRJEB19491. The mean number of sequence reads for bacteria and archaea was 41,410 ± 1,689 and 14,986 ± 1,713 respectively.

4.3.5 Sample Preparation for Mass Spectrometry

Samples were thawed and vortexed prior to sample preparation as described by Deusch and Seifert (2015). To detach firmly fiber-associated bacteria, 4 g of each SP sample were shaken horizontally for 2 h at 4°C in 35 ml precooled 50 mM Tris-HCl (pH 8; 0.2 M NaCl; 0.1% methylcellulose 400cP). To dilute the liquid fractions, 5 ml of the respective buffer were added to 8 g of the RF and LP samples. All samples were sonicated briefly for 1 min and pressed through two-layered sterile cheesecloth. Residues were rinsed again with 30 ml of the above mentioned buffer and pressed vigorously. Obtained filtrates were centrifuged at 200 x g for 10 min at 4°C and supernatants were further filtered through sterile 40 µm PE filters. Cells were pelleted at 10,000 x g for 15 min at 4°C and washed three times in 1 ml 50 mM Tris-HCl (pH 7.5; 0.1 mg/ml chloramphenicol; 1 mM PMSF). Subsequently, aliquoted cell pellets were stored at -20°C. Protein extraction was performed as described previously (Deusch and Seifert, 2015). Cell pellets were resuspended by vortexing in 200 µl 50 mM Tris-HCl (pH 7.5; 0.1 mg/ml chloramphenicol; 1 mM PMSF) and 300 µl of 20 mM Tris-HCl (pH 7.5; 2% SDS) were added. After shaking in a Thermo-Mixer (Eppendorf) for 10 min at 60°C and 1,200 rpm 1 ml of 20 mM Tris-HCl (pH 7.5; 0.1 mg/ml $MgCl_2$; 1 mM PMSF; 1 µl/ml Benzonase, Novagen) was added. Cells were lysed by ultra-sonication on ice, four times 2 min (amplitude 60%; cycle 0.5) followed by shaking in a Thermo-Mixer for 10 min at 37°C and 1,200 rpm. Samples were centrifuged at 10,000 × g for 10 min at 4°C and proteins in the supernatant were precipitated for 30 min at 4°C using 20% precooled trichloroacetic acid.

Subsequently, precipitates were centrifuged at 12,000 × g for 15 min at 4°C, protein pellets were washed twice in precooled acetone and dried by vacuum centrifugation. Protein pellets were resuspended in 35 µl Laemmli buffer by 5 min sonication bath and vortexing followed by incubation for 5 min at 95°C to reduce disulfide bonds. Twenty microliters were purified with a short run on a one-dimensional sodium dodecyl sulfate polyacrylamide gel electrophoresis (1D-SDS-PAGE; 4% stacking gel, 20 mA; 12% running gel, 40 mA). Each gel lane of 0.5 cm length representing one sample was cut out and subjected to in-gel trypsin (Promega, Madison, USA) digestion overnight (Jehmlich et al., 2008). Obtained peptides were purified and desalted using Stage tips equipped with five layers of Empore™ SPE Disks (C_{18}; diameter 47 mm; thickness 0.5 mm) as described in detail by Rappsilber et al. (2007).

4.3.6 LC-ESI-MS/MS Measurements

LC-ESI-MS/MS analyses were performed in technical duplicates on an EasyLC 1000 nano-UHPLC (Thermo Scientific) coupled to a Q Exactive HF mass spectrometer (Thermo Scientific). Prior to LC-ESI-MS/MS measurements peptides were reconstituted in 20 µl 0.1% formic acid and 4 µl were injected by the autosampler. Separations of the peptide mixtures were done on a 20 cm fused silica emitter of 75 µm inner diameter (Proxeon Biosystems), in-house packed with reversed-phase ReproSil-Pur 120 C_{18}-AQ 1.9 µm resin (Dr. Maisch GmbH). Peptide mixtures were injected onto the separation column in HPLC solvent A (0.1% formic acid) at a flow rate of 500 nl/min and eluted with a solvent B (80% acetonitrile in 0.1% formic acid) gradient of 1-33% within the first 73 min followed by an increase to 50% within 3 min plus an additional 3 min at 90%. The Q Exactive HF was operated in the positive ion mode. Full scan was acquired in the mass range from 300 to 1,650 m/z in the Orbitrap mass analyzer at a resolution of r = 120,000 followed by higher energy collisional dissociation (HCD) fragmentation of the twelve most intense precursor ions. High resolution MS/MS spectra were acquired with a resolution of r = 30,000. The target values were 3 x 10^6 charges for the MS scans and 1 x 10^5 charges for the MS/MS scans with a maximum fill time of 25 and 45 ms, respectively. The dynamic precursor exclusion was set to 30 s and peptide match was enabled.

4.3.7 Bioinformatic Data Analysis

To improve the false discovery rate of peptide identifications and enhance the confidence of protein identifications, a two-step search approach was applied to create an artificial metagenome (Jagtap et al., 2013; Hansen et al., 2014). Therewith, the size of the

search databases was reduced and simultaneously the sample-specificity was increased. First, all 54 raw data files were processed separately by Thermo Proteome Discoverer software (v. 1.4.1.14), Mascot engine (v. 2.4) in searching independently against the UniProtKB/TrEMBL databases (v. April 28, 2016) for bacteria (Taxonomy ID 2; 40,026,301 sequences) and archaea (Taxonomy ID: 2157; 1,200,545 sequences). Oxidation of methionine was set as variable modification and carbamidomethylation of cysteine as fixed modification. Precursor ion tolerance was defined at 10 ppm and fragment ion tolerance at 0.02 Da with two missed trypsin cleavages. Furthermore, all peaks besides the top 12 peaks per 100 Da in each MS/MS were removed to denoise spectra before identification and the Percolator node was activated with a false discovery rate of 1%. Using Thermo Proteome Discoverer, protein grouping was enabled with a minimum PSM confidence of medium and a delta Cn better than 0.15, strict maximum parsimony principle was applied. As a second step, the protein identifications inferred from the previous process were used to create sample-specific databases for label-free quantification (LFQ) of proteins via MaxQuant (v. 1.5.3.8) as previously demonstrated (Cox et al., 2014). The final in-house databases contained 22,331 bacterial and 818 archaeal protein sequences. The LFQ modality of MaxQuant was enabled with a minimum ratio count of two. Matching between runs with a match time window of 0.7 min and re-quantification was applied. Technical duplicates were combined to one experiment. Oxidation of methionine was set as variable modification with a maximum of five modifications per peptide and carbamidomethylation of cysteine was set as fixed modification. Besides, the default settings of MaxQuant were kept which included two missed trypsin cleavages, fully tryptic peptides, a peptide and protein false discovery rate below 1%, at least one peptide per protein (Gupta and Pevzner, 2009), a precursor mass tolerance of 4.5 ppm after mass recalibration and a fragment ion mass tolerance of 20 ppm. The phylogenetic composition was inferred from the proteins quantified by MaxQuant as annotated in the UniProtKB/TrEMBL database. KEGG Orthology (KO) identifiers and Cluster of Orthologous Groups of proteins (COG) were assigned using WebMGA (Wu et al., 2011) with an e-value cutoff of 10^{-3} considering exclusively the best hits. In order to link the quantified proteins to pathway maps of carbohydrate metabolism, KO identifiers were translated and grouped manually to the respective KEGG REACTION numbers as defined by the KO database. Carbohydrate-active enzymes (CAZymes) were annotated by searching the quantified bacterial proteins against the database for automated CAZyme annotation with hidden Markov models (dbCAN HMMs v. 5.0, based on the CAZyDB v. July 15, 2016) using hmmscan of the HMMER3 software package (Yin et al., 2012) and considering entirely the best e-value hits. The mass spectrometry data have been

deposited to the ProteomeXchange Consortium via the PRIDE partner repository (Vizcaino et al., 2016) with the dataset identifier PXD006070.

4.3.8 Nuclear Magnetic Resonance Spectroscopy

The RF samples were defrosted, vortexed, and filtered through sterile 100 µm PE filters. Filtrates were centrifuged at 13,000 x g for 30 min at 4°C. Obtained supernatants were sterilized by passing through a 0.22 µm syringe filter and 3 ml of each RF sample were dried completely by vacuum centrifugation at room temperature overnight. By vigorous vortexing and 5 min brief sonication, dehydrated RF samples were reconstituted in 1.5 ml 50 mM sodium phosphate buffer (pH 6.5) in 99.9% deuterium oxide (Sigma-Aldrich, Germany) containing 5 mM 3-(trimethylsilyl)propionic-2,2,3,3-d4 acid sodium salt (TSP; Sigma-Aldrich, Germany) as an internal chemical shift reference and quantification standard. Subsequently, dissolved samples were centrifuged at 13,000 x g for 30 min at 4°C and 1 ml supernatant was transferred to a 5 mm glass NMR tube for measuring at 500 MHz using a Varian INOVA NMR spectrometer (Agilent Technologies). All ^{1}H-NMR spectra were acquired at 25°C using the first transient of the noesy presaturation pulse sequence (Saude et al., 2006). Each spectrum was collected with 32 transients using a 4 s acquisition time, 1 s recycle delay and a mixing time of 0.1 s at a spectral width of 6,490 Hz. Spectral assignments were performed by 2D homonuclear and heteronuclear NMR: DQFCOSY, gHSQCAD, gHSQCTOCSY, as well as gHMBCAD were run using CHEMPACK 7.2 pulse sequences implemented in VnmrJ 4.2 (Agilent Technologies Inc., Santa Clara, CA, USA). Additionally, ^{1}H-NMR spectra were imported into the Chenomx NMR Suite 8.2 software (database available at pH 6.5, Chenomx Inc., Edmonton, AB, Canada) for quantification (Weljie et al., 2006; Wishart, 2008) as described in Ametaj et al. (2010). Spectra were referenced to TSP (δ 0.0 ppm) for chemical shift and quantification. Prior to spectral analysis, all free induction decays (FIDs) were automatically zero-filled to 64 k data points, corrected for phase and baseline distortions and a line broadening of 0.5 Hz was applied. Concentrations of identified metabolites were divided by a factor of two since dried RF samples were reconstituted in half of the initial volume.

4.3.9 Statistical Analyses

The LFQ abundance values of proteins and the OTU counts were analyzed using the Primer 6 (v. 6.1.16) and Permanova+ (v. 1.0.6) statistical software package (PRIMER-E, Plymouth, UK). Non-metric multidimensional scaling (NMDS) was performed using the

Bray-Curtis similarity matrix (Bray and Curtis, 1957). One-way analysis of similarities (ANOSIM) was used to determine statistical differences in protein and OTU abundance between diets, sample fractions and host animals (Clarke and Warwick, 2001). One-way analysis of variance (ANOVA) with post-hoc Tukey HSD (Honestly Significant Difference) was used for pairwise comparisons of the abundance means of taxonomic groups in diets and sample fractions and to test differences of metabolite concentrations in diets (IBM SPSS Statistics, Version 20.0. Armonk, NY: IBM Corp).

4.4 Results

4.4.1 Metaproteomics and Amplicon Sequencing of the Rumen Microbiome

Over all samples 8,163 bacterial and 358 archaeal proteins were quantified by mass spectrometric measurements of the peptides and a two-step search identification and quantification process. Illumina MiSeq sequencing of the V1-2 region of 16S rRNA gene resulted in 1,484 bacterial and 626 archaeal OTUs assigned (Chapter 4.9, Table S1). Regarding the abundances of bacterial proteins and OTUs, NMDS plots based on the Bray Curtis similarity revealed diet induced shifts and variations between sample fractions with similar trends for both, the metaproteomic dataset (Figure 9A) and the DNA-based approach (Figure 9B).

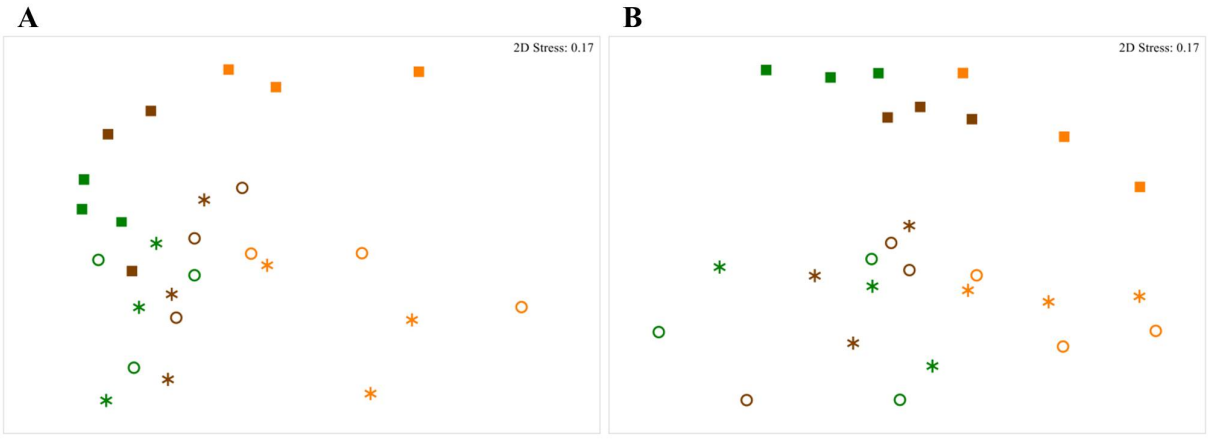

Figure 9 | Non-metric multidimensional scaling (NMDS) plots of the metaproteome (**A**) and the bacterial community structure (**B**). Yellow, corn silage-based diet; green, grass silage-based diet; brown, grass hay-based diet. Squares, solid phase; stars, liquid phase; circles, rumen fluid.

There were no dietary, sample fraction or host related effects on the abundance of archaeal proteins and OTUs. ANOSIM verified significant differences in bacterial protein abundances

regarding the diets with an R-value of 0.600 ($P = 0.0001$) for the corn silage- and grass silage-based samples and 0.442 ($P = 0.0001$) for the corn silage- and hay-derived samples. Respectively, the numbers of assigned bacterial OTUs of the corn silage- and grass silage-based samples differed with an R-value of 0.500 ($P = 0.0003$) and the corn silage- and hay-derived samples showed an R-value of 0.558 ($P = 0.0001$). Furthermore, ANOSIM confirmed significant variations in protein abundances between sample fractions with an R-value of 0.362 ($P = 0.0001$) for the LP and SP fractions. The RF and SP fractions differed with an R-value of 0.321 ($P = 0.0008$). Concerning the sequencing data, the R-value for the LP and SP fractions was 0.596 ($P = 0.0001$) and 0.561 ($P = 0.0002$) for the RF and SP fractions. Complete statistical information of the bacterial protein and OTU abundances considering diets, sample fractions and the individual cows is shown in Table 4.

4.4.2 Dietary Impact on Community Structure and Variations in Fractions

Taxonomic information was obtained from quantified proteins as annotated by the UniProtKB/TrEMBL database and from RDP retrieved OTU assignments of the Illumina amplicon sequencing. The taxonomic distribution of archaeal and bacterial proteins and OTUs in each sample at phyla, class, order, and family level as well as the corresponding numbers of proteins and OTUs are listed in Chapter 4.9, Table S2. Overall, the bacterial community composition was dominated by the phylum of Bacteroidetes followed by the phylum of Firmicutes. Less abundant phyla were Actinobacteria, Elusimicrobia, Proteobacteria, Spirochaetes, Synergistetes, Tenericutes, and Verrucomicrobia (Chapter 4.9, Table S2).

However, the bacterial community structure averaged over the three animals per treatment revealed concordant tendencies concerning the dietary influence and variations in different microenvironments for both applied methods. Figure 10 shows the average abundance in diets and sample fractions of bacterial phyla, orders and families commonly identified by metaproteomics on the left and 16S rRNA gene sequencing on the right side. Proteins and OTUs belonging to the phylum of Proteobacteria including the family of Succinivibrionaceae were significantly more abundant throughout the samples of the corn silage-based diet when compared to the grass silage- and hay-based samples with $P < 0.01$ (LFQs) and $P < 0.05$ (OTUs). Within the corn silage-derived samples the respective proteins and OTUs were more abundant in the SP fraction (Figure 10). The abundance of OTUs assigned to Succinivibrionaceae was higher when compared to the LFQ-values of the corresponding proteins. Likewise, the family of Acidaminococcaceae belonging to the phylum of Firmicutes showed a higher protein ($P < 0.01$) and OTU abundance in the corn

silage-based diet when compared to the grass silage- and hay-based samples (Figure 10). The abundance of the respective OTUs was least in the SP fraction of all diets whereas the LFQ-values of the corresponding proteins were highest in the SP fractions (Figure 10). The Firmicutes family of Selenomonadaceae was identified exclusively in the protein-based dataset (Figure 10). Selenomonadaceae proteins exhibited a lower abundance in the grass silage-derived sample fractions when compared to the corn silage- and hay-based samples ($P < 0.01$). Members of the order of Clostridiales including the families of Lachnospiraceae and Ruminococcaceae constituted major parts of the Firmicutes phylum (Figure 10) and revealed the highest LFQ-values and numbers of OTUs in the SP fractions of all diets when compared to the respective RF and LP fractions with $P < 0.01$ (LFQs) and $P < 0.05$ (OTUs).

Table 4 | Analysis of similarity of the metaproteomic- and the 16S rRNA gene-based datasets. The upper part of the table shows the global *R* statistics and the respective probability values (*P*) for the main factors (diets, sample fractions and individual cows). The lower part shows the *R* statistics and corresponding probability values (*P*) for the pairwise group comparisons within the main factors. Significantly different main factors and groups within are marked (*). CS, corn silage-based diet; GS, grass silage-based diet; H, grass hay-based diet. RF, rumen fluid; LP, liquid phase; SP, solid phase. Cows are indicated by no. 23, 28 and 59.

		LFQ-values of 8,163 bacterial proteins		**Abundance of 1,484 bacterial OTUs**	
		R	***P***	***R***	***P***
	Diets *	0.418	0.0001	0.366	0.0001
	Fractions *	0.244	0.0004	0.379	0.0001
	Cows	0.143	0.0090	0.176	0.0070
Diets	**GS:H**	0.248	0.0100	0.078	0.1630
	CS:GS *	0.600	0.0001	0.500	0.0003
	CS:H *	0.442	0.0001	0.558	0.0001
Fractions	**RF:LP**	0.053	0.1870	-0.045	0.6980
	RF:SP *	0.321	0.0008	0.561	0.0002
	LP:SP *	0.362	0.0001	0.596	0.0001
Cows	**28:59**	0.088	0.1090	0.133	0.0790
	23:59	0.169	0.0280	0.207	0.0280
	23:28	0.180	0.0150	0.202	0.0270

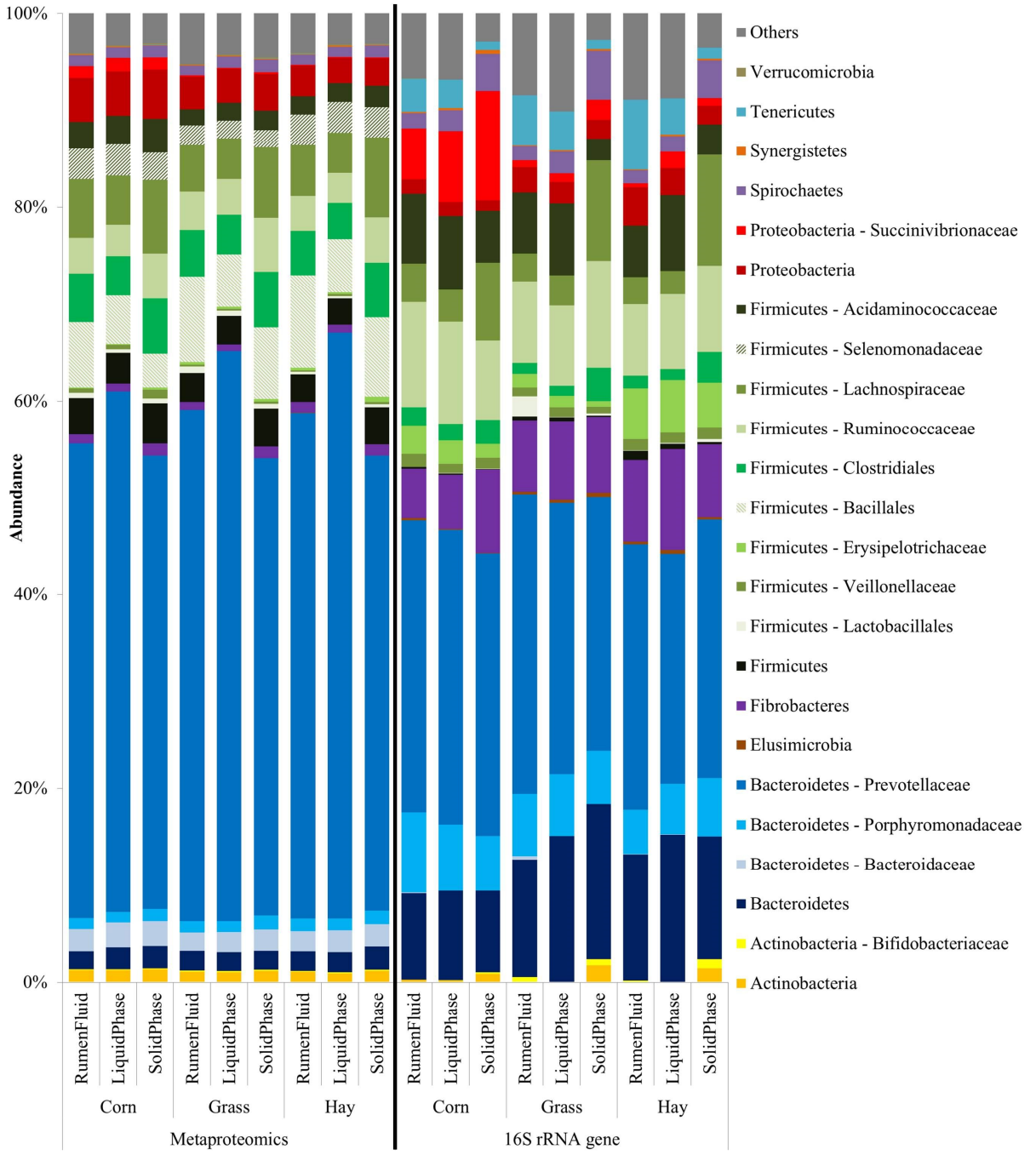

Figure 10 | Bacterial phylogenetic distribution in diets and sample fractions (n = 3) obtained from the metaproteome (8,163 proteins) on the left and the 16S rRNA gene sequencing (1,484 OTUs) on the right. The order of Bacillales and the family of Selenomonadaceae were found exclusively in the metaproteomic dataset (shaded coloring).

No OTUs were assigned to the Firmicutes order of Bacillales although highly abundant proteins were identified by mass spectrometry. Within diets, Bacillales proteins were more abundant in the RF fractions. The Erysipelotrichaceae family was more abundant in the hay-based diet when compared to the corn silage- and grass silage-based samples (Figure 10)

emphasized primarily by the abundance of assigned OTUs ($P < 0.01$). Proteins and OTUs belonging to the order of Lactobacillales and to the family of Veillonellaceae constituted only a small part of the Firmicutes phylum as determined by both approaches (Chapter 4.9, Table S2). Proteins of the Veillonellaceae family were more abundant in the corn silage-based diet when compared to the grass silage- and hay-based samples ($P < 0.05$). The abundance of OTUs belonging to the phylum of Fibrobacteres was higher when compared to the LFQ-values of the corresponding proteins (Figure 10). According to the 16S rRNA gene sequencing, the phylum of Fibrobacteres was less abundant in the RF and LP fractions of the corn silage-based diet when compared to the respective fractions of the grass silage- and hay-based diets. Within diets, the abundance of Fibrobacteres proteins was higher in the SP fractions of all diets when compared to the RF and LP samples ($P < 0.05$). Related to the overall bacterial community structure, OTU and LFQ-values assigned to the phylum of Elusimicrobia exhibited a low abundance across all samples (Chapter 4.9, Table S2). Regardless the dietary treatments, the Prevotellaceae family dominated the phylum of Bacteroidetes and the overall bacterial community composition as determined by the protein- and DNA-based approaches (Figure 10). Within diets, Prevotellaceae proteins showed the highest LFQ-values in the LP fractions ($P < 0.05$). The family of Porphyromonadaceae exhibited a higher abundance of OTUs when compared to the LFQ-values of the respective proteins. In the same way, sequences assigned to the phylum of Bacteroidetes were more abundant than the LFQ-values of the corresponding proteins (Figure 10). In contrast, the family of Bacteroidaceae showed a higher abundance of proteins when compared to the abundance of the corresponding OTUs (Figure 10). OTU and LFQ-values of Bifidobacteriaceae were higher in the grass silage- and hay-based samples when compared to the corn silage-based fractions (Chapter 4.9, Table S2). The abundance of OTUs and LFQs of the Actinobacteria phylum including the family of Bifidobacteriaceae was higher in the SP fractions of all diets when compared to the respective RF and LP fractions (Figure 10). Within diets, the phylum of Spirochaetes exhibited the highest abundance in the SP fractions as found by both methods. Similarly, the phylum of Synergistetes revealed the highest abundance of proteins and particularly OTUs in the SP fractions when compared to respective the RF and LP fractions (Figure 10). Proteins assigned to the phylum of Tenericutes were low abundant when compared to the corresponding OTU abundance. OTUs of the phylum of Tenericutes were more abundant in the RF and LP fractions when compared to the respective SP sample fractions whereas the LFQ-values of the corresponding proteins were higher in the SP fractions of all diets (Figure 10). In contrast, proteins of the Verrucomicrobia phylum were

more abundant in comparison to the respective OTUs (Chapter 4.9, Table S2).

Figure 11 shows the average abundance in diets and sample fractions of archaeal phyla, classes and families identified by metaproteomics on the left and 16S rRNA gene sequencing on the right. The metaproteomics-based approach identified 358 archaeal proteins that distributed over all archaeal phyla including as well the family of Thermococcaceae whereas the 626 OTUs were exclusively assigned to three families of methanogens: Methanobacteriaceae, Methanosarcinaceae, and Methanomassiliicoccaceae (Figure 11). Proteins belonging to the phyla of Thaumarchaeota and Crenarchaeota constituted a minor part of the total archaeal protein abundance (Figure 11). The LFQ-values of unclassified archaeal proteins were above 16% in each sample (Chapter 4.9, Table S2). Within the phylum of Euryarchaeota, proteins belonging to the families of Methanobacteriaceae, Methanosarcinaceae, Thermococcaceae, and the class of Methanomicrobia were most abundant while OTUs of the family of Methanobacteriaceae prevailed with above 72% in each sample (Chapter 4.9, Table S2).

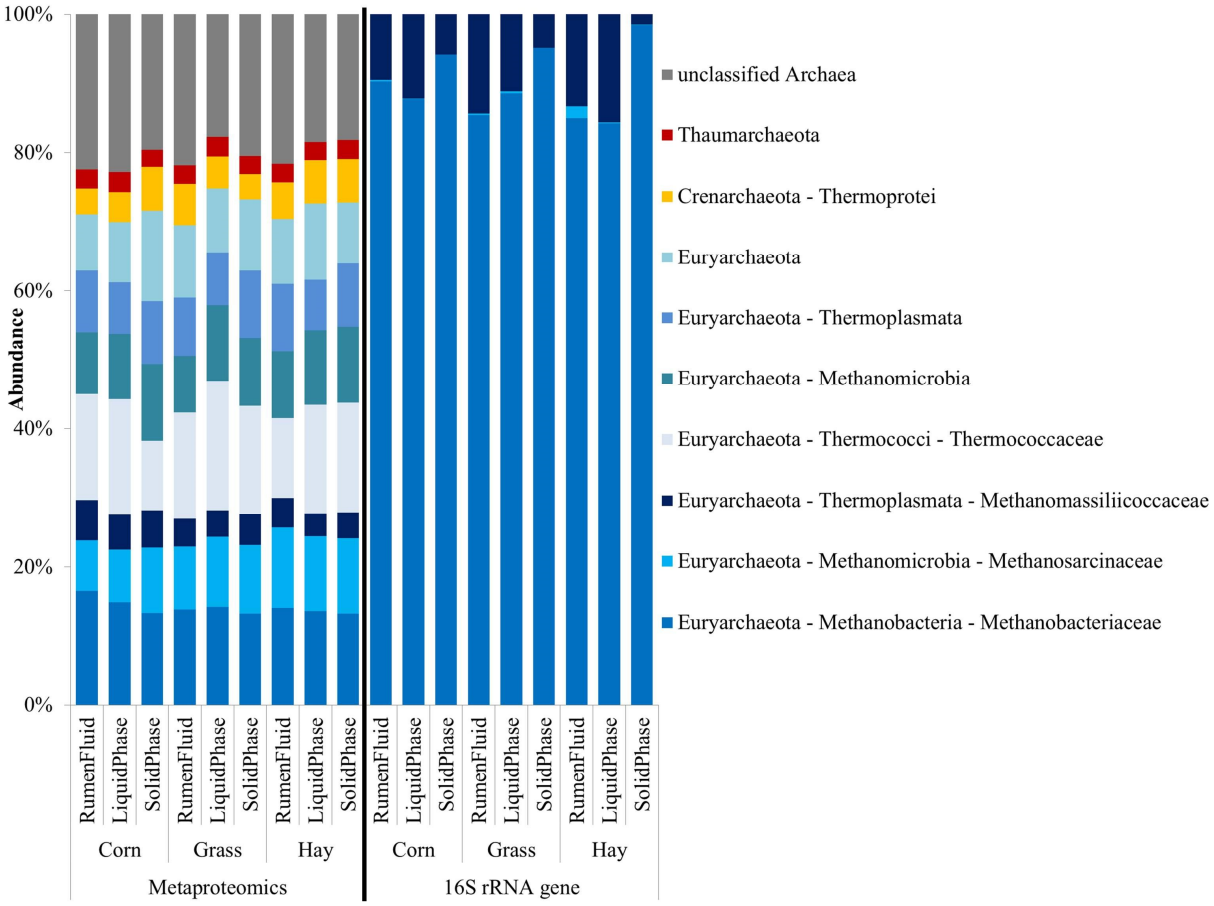

Figure 11 | Archaeal phylogenetic distribution in diets and sample fractions (n = 3) obtained from the metaproteome (358 proteins) on the left and the 16S rRNA gene sequencing (626 OTUs) on the right.

4.4.3 Carbohydrate-Active Enzymes

CAZyme annotation with hidden Markov models identified a total of 166 bacterial proteins in five CAZy-categories (Chapter 4.9, Table S3). A majority of 91 proteins were assigned to 16 glycoside hydrolase (GH) families and 38 proteins belonged to seven glycosyltransferase (GT) families. Furthermore, 16 proteins were assigned to three families of carbohydrate esterases (CE) and two proteins fell into two families of carbohydrate-binding modules (CBM). In addition, 19 proteins with sequence similarity to S-layer homology domains (SLH) were present in high abundance throughout diets and sample fractions. The percentage abundance of the respective bacterial CAZymes in diets and sample fractions in relation to the total abundance of CAZymes is shown in Figure 12A. The phylogenetic origin at phyla level of the five CAZy-categories in diets and sample fractions is depicted in Figure 12B. The ratio of the most abundant CAZy-categories of GH, mainly produced by Bacteroidetes species and SLH, almost exclusively derived from Firmicutes species varied between the diets (Figures 12A,B). Proteins of the SLH category were less abundant in the grass silage-based sample fractions when compared to the corn silage- and hay-based samples. Contrarily, proteins of the GH category were more abundant in the grass silage-based diet in comparison to the corn silage- and hay-based samples (Figure 12A). Proteobacteria-derived proteins belonging to the GT category were more abundant in the corn silage-based samples when compared to grass silage- and hay-based samples (Figure 12B). Proteins assigned to the CE category were more abundant in the LP and SP fractions of the corn silage-based diet when compared to the respective fractions of the grass silage and hay-based diets (Figure 12A). Moreover, the LP and SP fractions of the corn silage-based diet exhibited a higher abundance of proteobacterial proteins in the CE category when compared to the remaining samples (Figure 12B). CBM related proteins were more abundant in the sample fractions of the grass silage-based diet when compared to the corn silage- and hay-based diets (Figure 12A). Across all samples, the family GH57, based on seven proteins like alpha-amylases, was most abundant (Figure 12A). The most expressed GT family was GT35 including 18 proteins of different glycogen and starch phosphorylases. Within the category of CE the most expressed family was CE1 with eight proteins. The CBM32 family was most abundant in the SP fraction of the grass silage-based diet (Figure 12A). Details about the proteins assigned to the different CAZy-categories including the phylogenetic origin and the respective functions are listed in Chapter 4.9, Table S3.

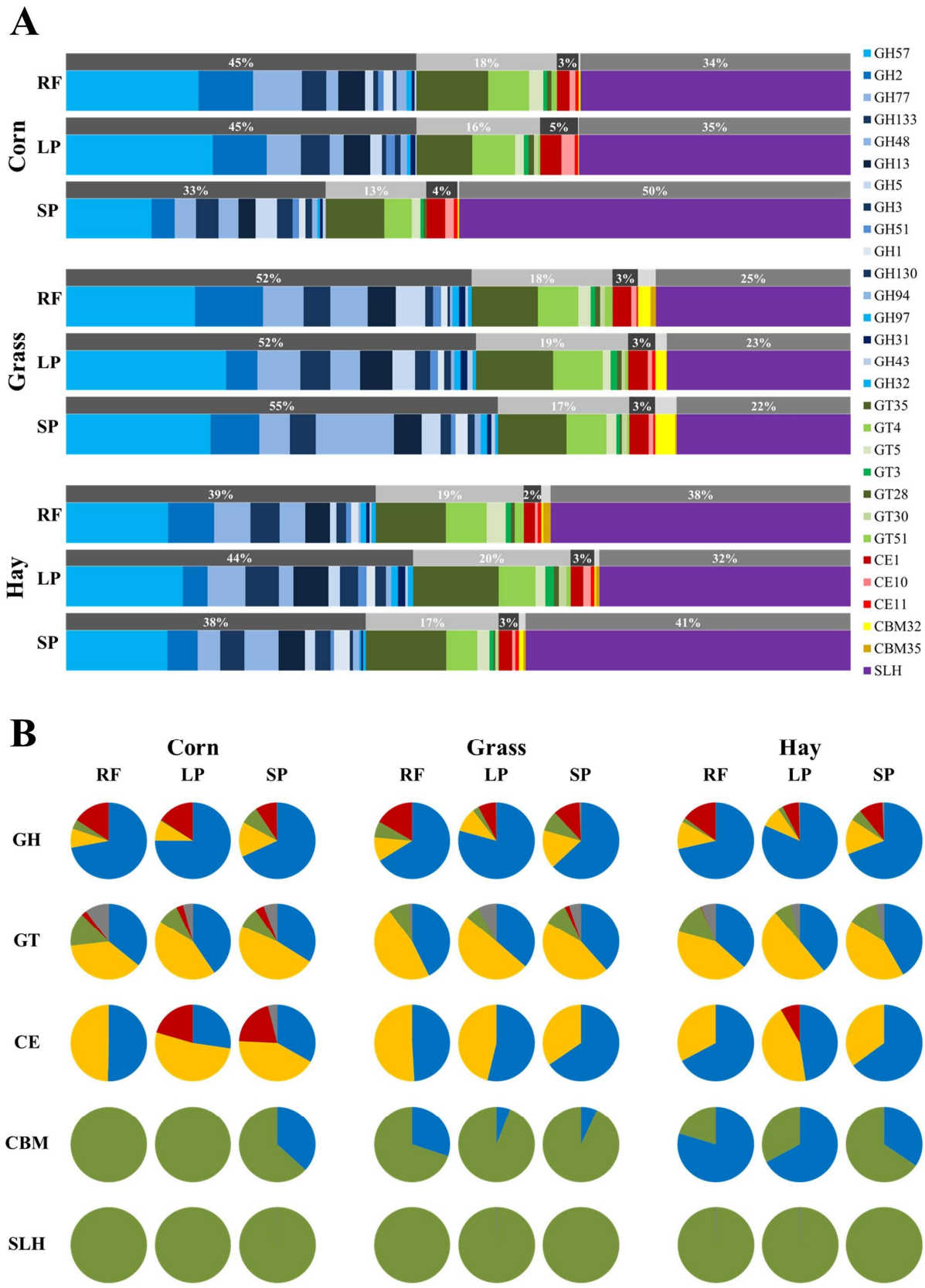

Figure 12 | Abundance of 166 CAZymes in diets and sample fractions (n = 3). Glycoside hydrolases, GH (91 proteins); glycosyltransferases, GT (38 proteins); carbohydrate esterases, CE (16 proteins); carbohydrate-binding modules, CBM (2 proteins) and S-layer homology

domains, SLH (19 proteins) (**A**). Pie charts depict the phylogenetic origin of CAZymes of the respective categories in diets and sample fractions at phyla level: blue, Bacteroidetes; green, Firmicutes; red, Proteobacteria; yellow, environmental samples; gray, others (**B**). RF, rumen fluid; LP, liquid phase; SP, solid phase.

4.4.4 ABC Transporters

WebMGA assigned 7,745 and 336 KO identifiers to 8,163 bacterial and 358 archaeal proteins respectively (Chapter 4.9, Table S4), revealing a total of 170 bacterial proteins with sequence similarity to ABC transporters (Chapter 4.9, Table S5) as defined by the KO database. Figure 13 depicts the average abundance of the respective membrane transporters in diets and sample fractions in relation to the maximum LFQ-values within each group of transporters and the origin at bacterial phyla level. Analyses affirmed 106 proteins to the group of oligosaccharide, polyol, and lipid transporters that include several subunits of multiple sugar transport systems and of cellobiose, arabinose/lactose, maltose/maltodextrin, sorbitol/mannitol, and galactose oligomer transporters. The respective proteins were more abundant in the hay- and grass silage-based diets when compared to the corn silage-based diet (Figure 13). Within diets, the highest abundance was observed in the SP fractions. Oligosaccharide, polyol, and lipid transporters originated mainly from Firmicutes species. In contrast, the group of monosaccharide transporters, based on 20 proteins, was more abundant in the corn silage-based samples when compared to the grass silage- and hay-based diets (Figure 13). This group includes subunits of ribose, rhamnose, methyl-galactoside, and *sn*-glycerol 3-phosphate transport systems. Proteobacterial proteins showed an increased contribution regarding monosaccharide transporters when compared to the respective phylogenetic origin of the group of oligosaccharide, polyol, and lipid transporters (Figure 13). Phosphate and amino acid transporters based on 27 proteins showed the highest abundance in corn silage-derived samples and were mainly produced by Proteobacteria species (Figure 13). Mineral and organic ion transporters including 17 proteins were almost exclusively produced by Firmicutes species and showed the highest LFQ-values in the corn silage-derived SP fraction (Figure 13).

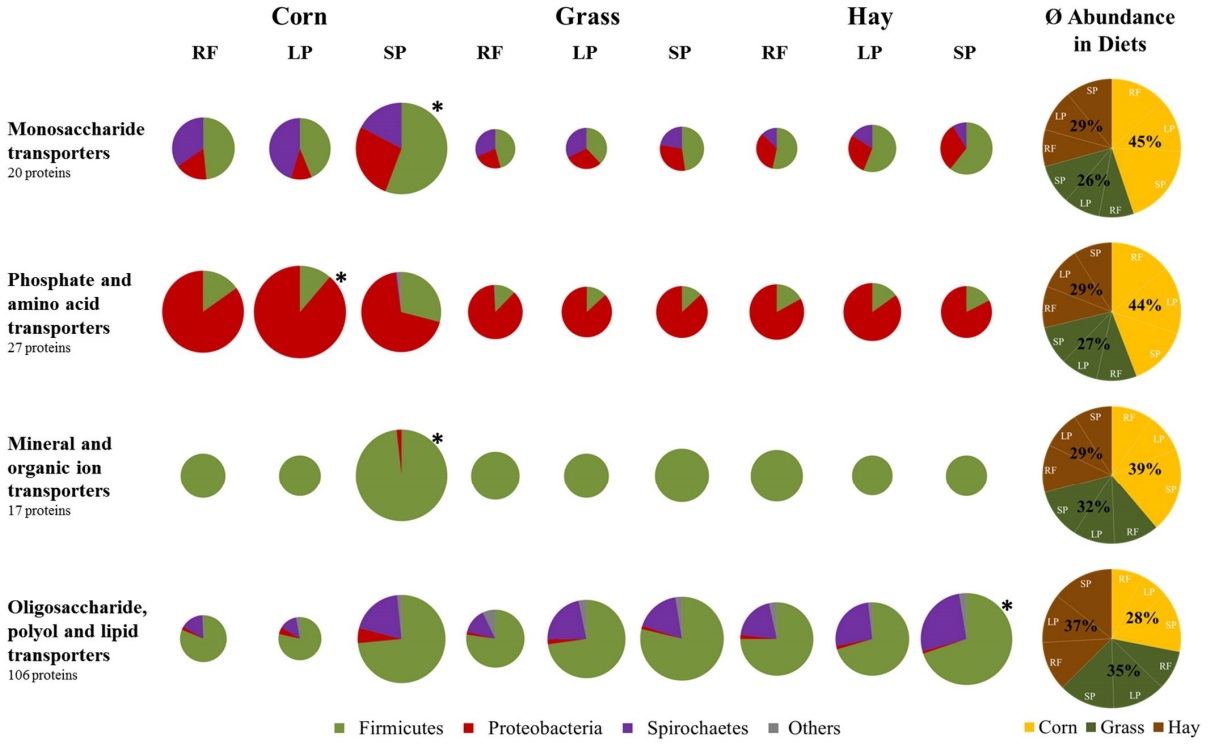

Figure 13 | Varying abundance of bacterial ABC transporters in diets and sample fractions (n = 3). Pie sizes indicate the abundance in relation to the sample with the maximum LFQ-values (*) in each group. Phylogenetic origin of proteins in the respective samples is shown at phyla level. RF, rumen fluid; LP, liquid phase; SP, solid phase.

4.4.5 Archaeal and Bacterial Enzymes Involved in Carbon Metabolism

Varying abundance, in diets and sample fractions, of enzymes involved in the carbon metabolism are shown in Figure 14A. To visualize the abundance of the respective enzymes in the pathway map, a total of 1,121 bacterial proteins assigned to 80 KO identifiers carrying out 70 KEGG REACTIONS were arranged into 60 functional groups (Chapter 4.9, Table S6). A few KOs were assigned to more than one KEGG REACTION and thus, appear more than once in the grouping. Furthermore, 28 archaeal proteins of eight KOs involved in the methane metabolism were grouped to five KEGG REACTIONS (Chapter 4.9, Table S6). Additionally, eight compounds identified by NMR are shown in Figure 14A. Twenty-four enzymes of the groups 8 and 12 involved in the phosphate acetyltransferase-acetate kinase pathway were present in all diets and sample fractions, whereas only one acetyl-CoA synthetase from an uncultured bacterium in group 9 was found in the SP fraction of the hay-based diet (Figure 14A). The abundance of seven proteins of group 6 carrying out the oxidation of pyruvate to acetyl-CoA was higher in the RF and SP fractions of the grass silage-based diet and higher in the LP fraction of the hay-based diet when comparing to the other fractions and diets.

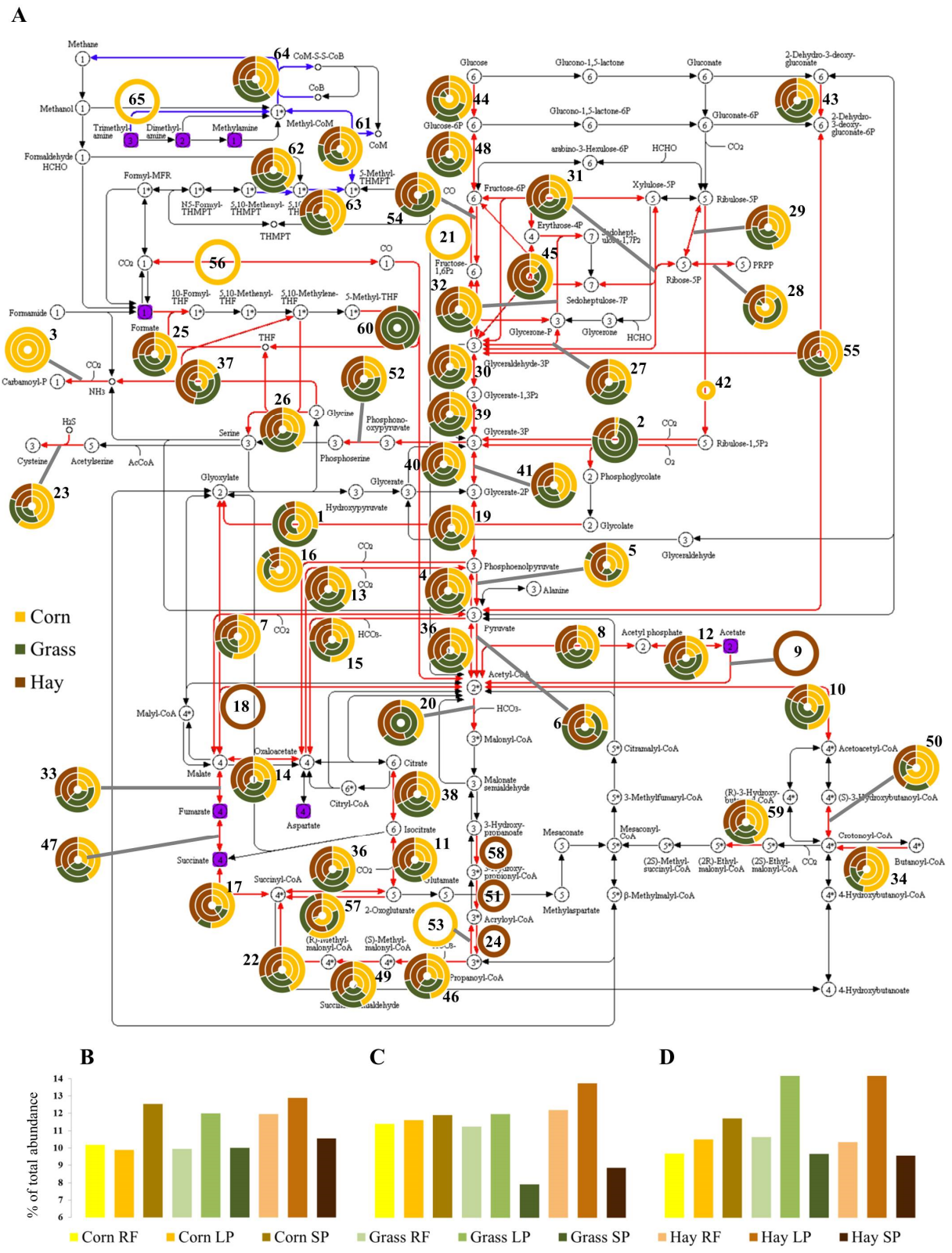

Figure 14 | (**A**) Abundance of archaeal (blue) and bacterial (red) enzymes (n = 3) involved in carbon metabolism, grouped and numbered according to the respective reactions (Table S6). Outer circle, solid phase; middle circle, liquid phase; inner circle, rumen fluid. Purple, compounds identified by NMR. The percentage of total LFQ-values of groups involved in (**B**)

the Embden-Meyerhof-Parnas pathway (5, 19, 27, 30, 32, 39, 41, 44, 48, 54), (**C**) citrate cycle (11, 14, 17, 33, 38, 47, 57) and (**D**) pentose phosphate pathway (29, 31, 45, 48). RF, rumen fluid; LP, liquid phase; SP, solid phase.

Conversion of acetyl-CoA to malonyl-CoA employing two proteins of group 20 produced by *Bacillus* and *Prevotella* species, exhibited abundance exclusively in the grass silage-derived RF fraction and regarding the corn silage-based diet, were present only in the SP fraction (Figure 14A). A Tenericutes-derived carbamate kinase in group 3 producing carbamoyl phosphate, a metabolite in nitrogen disposal through the urea cycle, was found exclusively in the fractions of the corn silage-based diet (Figure 14A). Summing up the LFQ-values of groups involved in the Embden-Meyerhof pathway (5, 19, 27, 30, 32, 39, 41, 44, 48, 54) consisting of a total of 417 proteins, the abundance within diets was the highest in the LP fractions of the grass silage- and hay-based diets and the corn silage-derived SP fraction (Figure 14B). The sum of abundance of groups belonging to the citrate cycle (11, 14, 17, 33, 38, 47, 57) based on a total of 76 proteins exhibited lower LFQ-values in the SP fractions of the grass silage- and hay-based diets when compared to the respective RF and LP fractions and the corn silage-based samples (Figure 14C). Looking at the groups 29, 31, 45, 48 of the pentose phosphate pathway based on 34 proteins, the sum of LFQ-values was higher in the LP fractions of the grass silage- and hay-based diets when compared to the remaining samples (Figure 14D).

4.4.6 Enzymes of Short-Chain Fatty Acid Production

Bacterial enzymes involved in the production of acetate, butyrate, propionate and formic acid (Chapter 4.9, Table S7) were retrieved using COG assignments of WebMGA (Chapter 4.9, Table S4). The LFQ-values of proteins belonging to the respective COGs were summarized as described in Polansky et al. (2015). Figure 15 depicts the phylogenetic origin of the corresponding enzymes averaged over the diets. In total 45 enzymes involved in butyrate production were dominated by Firmicutes species constituting above 90% of the total abundance in all diets whereas enzymes from Actinobacteria appeared exclusively in the corn silage-based diet (Figure 15A). Five proteins of Bacteroidetes species showed a maximum of 3.9% of abundance in the grass silage-based samples and exhibited LFQ-values below 0.3% in the corn silage-based diet. Thirteen proteins of the Lachnospiraceae family exhibited the highest LFQ-values in the hay-based diet with 59% of the total abundance, decreasing in the grass (56%) and corn (31%) silage-based diets. Likewise, the abundance of two enzymes

derived from Ruminococcaceae was higher in the grass silage- and hay-based diet (1.7%) when compared to the average in the corn silage-based diet (1.3%). In contrast, five enzymes of Eubacteriaceae showed the highest LFQ-values in the corn silage-based diet and decreased in the grass silage- and hay-derived samples. A 3-hydroxybutyryl-CoA dehydrogenase of *Megasphaera elsdenii* belonging to the family of Veillonellaceae was present in high abundance in the corn silage-based diet (24%) while the LFQ-values in the grass silage- and hay-based diets accounted for 0.4% and 0.1% respectively (Figure 15A). Propionate production based on 52 proteins was dominated by 25 enzymes from Bacteroidetes species,

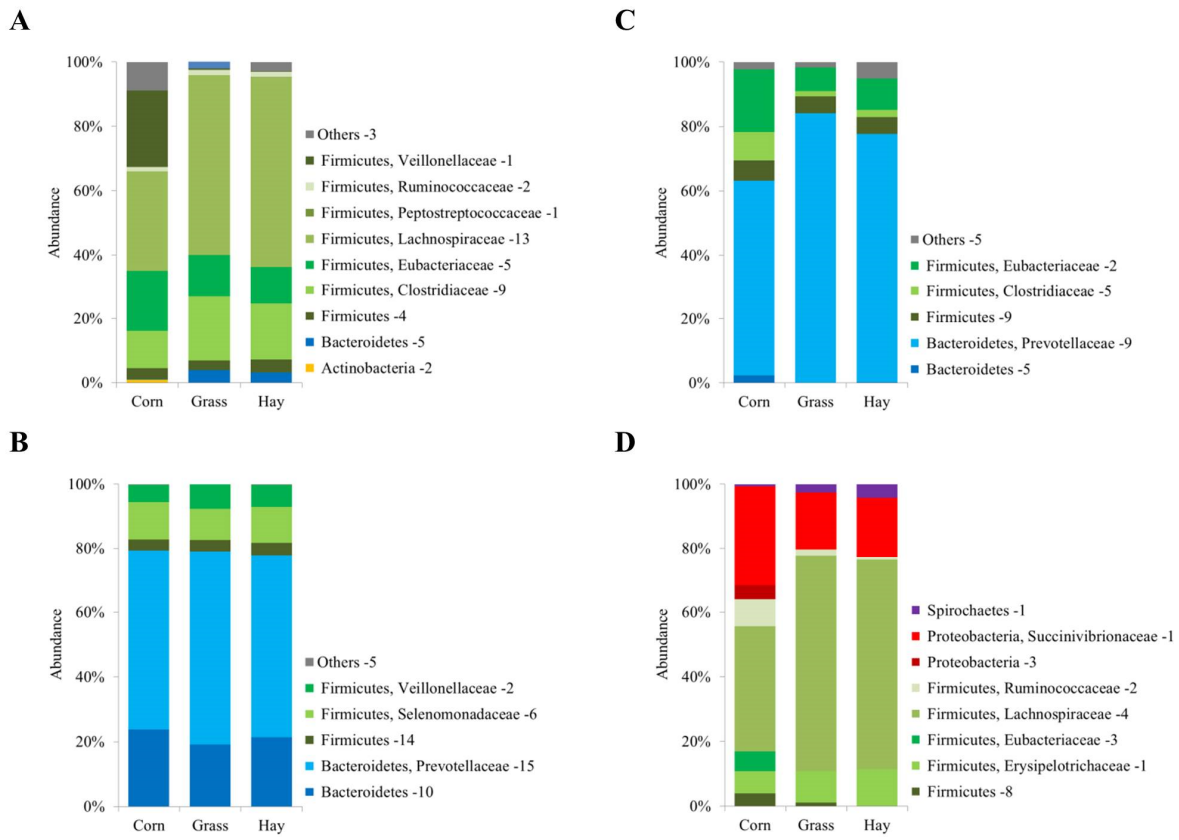

Figure 15 | Average abundance and phylogenetic origin of bacterial enzymes involved in (**A**) butyrate, (**B**) propionate, (**C**) acetate, and (**D**) formic acid production in diets (n = 9). Numbers of assigned proteins are shown in the respective phylogenetic legends. For butyrate biosynthesis: COG4770 (acetyl/propionyl-CoA carboxylase), COG3426 (butyrate kinase), COG1250/COG1024 (3-hydroxyacyl-CoA dehydrogenase), COG0183 (acetyl-CoA acetyltransferase). For propionate biosynthesis: COG4799 (acetyl-CoA carboxylase), COG2185/COG1884 (methylmalonyl-CoA mutase). For acetate production: COG1012 (NAD-dependent aldehyde dehydrogenase), COG0282 (acetate kinase), COG0280 (phosphotransacetylase), and for formic acid COG1882 (formate acetyltransferase).

mainly Prevotellaceae, composing above 75% of the total abundance in all samples (Figure 15B). Correspondingly, nine enzymes of the Prevotellaceae family prevailed in acetate production constituting above 60% of abundance in the corn silage-based diet increasing to 77% in the hay-based diet and 84% in the grass silage-based diet (Figure 15C). Two enzymes of *Eubacterium* species showed an abundance of 20% while decreasing in the hay- and the grass silage-based diets with 10 and 7% respectively. Concerning the 23 enzymes involved in formic acid production, four proteins from Lachnospiraceae species constituted 39% of abundance in the corn silage-based diet whereas reaching 65% in the hay- and 67% in the grass silage-derived samples (Figure 15D). A formate C-acetyltransferase of *Succinatimonas hippie* belonging to the family of Succinivibrionaceae reached 31% of the total abundance in the corn silage-based diet and decreased to 19% in the hay- and 18% in the grass silage-based diet. Three formate acetyltransferases of *Eubacterium* species showed the maximum abundance of 6% in the corn silage-derived samples and were not present in the grass silage- and hay-based diets (Figure 15).

4.4.7 Metabolites in Rumen Fluid

NMR spectroscopy allowed the quantification of 20 different compounds in all RF samples. Two-dimensional NMR spectroscopy further validated the presence of 12 compounds (*) including the major short-chain fatty acids acetate, butyrate, propionate, and valerate (Table 5). There were no statistically significant alterations of the metabolites regarding the different diets as determined by one-way ANOVA ($P > 0.05$). Overall, acetate was detected in highest concentrations ranging from 54.86 mM in the corn silage-based samples to 59.70 mM in the grass silage-derived RF samples. Propionate concentration was lower in the RF samples of the hay-based diet (13.86 mM) when compared to 15.22 and 15.94 mM in the corn silage- and grass silage-based samples, respectively (Table 5). The ratio of acetate to propionate concentration was highest in the hay-based samples and lowest in the corn silage-based samples. Butyrate was more abundant in the grass silage- and hay-derived RF samples (16.84 and 15.81 mM) when compared to 14.98 mM in the corn silage-based samples (Table 5). Valerate concentration was higher in corn silage- and hay-derived RF fractions (5.27 mM and 4.15 mM) when compared to 2.99 mM in the grass silage-based diet. The corn silage-based samples exhibited the highest concentration of succinate (0.38 mM) in comparison to the grass silage- and hay-based samples with 0.1 and 0.25 mM respectively (Table 5). Higher amounts of lactate were found in in the hay diet-based samples (5.25 mM) when compared to the corn silage- (1.18 mM) and grass silage-based samples (0.84 mM).

Methylamine, dimethylamine, and trimethylamine were less abundant in the corn silage-based samples when compared to the grass silage and hay-based RF fractions (Table 5). Several compounds identified in low concentrations like aspartate, formate, fumarate, methylamines, and succinate are involved in the carbon metabolism as indicated in Figure 14.

Table 5 | Average (Ø) concentration (mM) and standard error of mean (SEM) of compounds identified in the RF fractions (n = 3) are shown. CS, corn silage-based diet; GS, grass silage-based diet; H, grass hay-based diet; RF, rumen fluid. *Compounds validated by 2D-NMR spectroscopy. There was no statistically significant difference in metabolite abundance between the diets ($P > 0.05$).

Metabolite	Average concentration (mM) and SEM (n = 3)					
	Corn RF		Grass RF		Hay RF	
	Ø	SEM	Ø	SEM	Ø	SEM
2-Phenylpropionate *	0.14	0.02	0.12	0.05	0.15	0.05
3-Phenylpropionate *	0.38	0.05	0.40	0.08	0.39	0.10
Acetate *	54.86	3.79	59.70	8.49	55.95	3.41
Adipate	1.78	0.67	1.66	0.34	1.45	0.58
Aspartate	0.23	0.07	0.39	0.15	0.31	0.12
Butyrate *	14.98	2.20	16.84	5.58	15.81	1.54
Dimethylamine *	0.10	0.07	0.29	0.08	0.28	0.05
Formate	0.14	0.03	0.10	0.02	0.11	0.02
Fumarate	0.02	0.00	0.01	0.00	0.02	0.00
Isobutyrate *	1.43	0.57	1.29	0.42	1.51	0.49
Isovalerate *	1.31	0.26	1.09	0.32	1.02	0.04
Lactate	1.18	1.11	0.84	0.45	5.25	4.35
Methylamine *	0.49	0.15	0.94	0.12	0.70	0.16
Phenylacetate *	0.13	0.05	0.17	0.11	0.19	0.09
Pimelate	4.54	0.53	3.13	0.70	3.73	0.19
Propionate *	15.22	3.17	15.94	3.34	13.86	1.30
Succinate	0.38	0.38	0.10	0.05	0.25	0.11
Trimethylamine *	0.17	0.13	0.59	0.35	0.62	0.49
Urea	2.74	1.21	2.68	1.54	2.98	0.86
Valerate *	5.27	0.95	2.99	0.42	4.15	0.31
Acetate : Propionate	3.60		3.75		4.04	

4.5 Discussion

In this study, the dietary effects of the most common forages in cattle production on the structure and function of the archaeal and bacterial communities inhabiting different microenvironments of the rumen ecosystem were analyzed by a combination of shotgun-metaproteomics, Illumina amplicon sequencing and nuclear magnetic resonance to provide deeper insights into the complex microbial adaptation to varying substrates. In general, the bacterial community composition of the rumen samples was dominated by Bacteroidetes and Firmicutes species as reported by other nucleic acid-based studies (Jami and Mizrahi, 2012; Jami et al., 2014). However, the bacterial community arrangement changed significantly in response to varying diets and differed between the rumen sample fractions as confirmed by ANOSIM for both, the 16S rRNA gene and the metaproteomic analysis. There was no significant effect of the host animals on the inherent bacterial community structure (Table 4).

4.5.1 Bacterial Community Composition and Activity is Influenced by Dietary Treatments

Dietary impact is the main factor shaping bacterial communities in the rumen (Ley et al., 2008; Henderson et al., 2015). The higher amount of non-fiber carbohydrates, mainly starch and sugars, in the corn silage-based diet (Table 3) might be responsible for the increased LFQ and OTU abundances of members of the Proteobacteria phylum and the family of Succinivibrionaceae (Hespell, 1992; Bryant, 2015). Typical members of the Succinivibrionaceae family are *Succinimonas amylolytica* and *Ruminobacter amylophilus*. The latter one is restricted to starch and maltose as fermentation substrates and produces succinic, formic and acetic acids (Hamlin and Hungate, 1956). *S. amylolytica* was shown to increase in abundance when substrates contained starch with succinic, acetic and propionic acid being the main fermentation products (Bryant et al., 1958). Thus, the higher amount of non-fiber carbohydrates in the corn silage-based diet (Table 3) and the associated rise in abundance of Proteobacteria and Succinivibrionaceae species (Figure 10) might have caused an increased production of succinate. The NMR-based metabolomic analysis determined the highest concentration of succinate in the corn silage-based RF fractions (Table 5). Moreover, the rise of Proteobacteria in the fractions of the corn silage-based diet (Figure 10) is in concordance with the increased abundance of monosaccharide transporters in the respective samples and the higher proteobacterial contribution to monosaccharide transporters when compared to the group of oligosaccharide transporters (Figure 13). Additionally, the abundance of proteobacterial glycosyltransferases and particularly carbohydrate esterases

increased in the corn silage-derived LP and SP fractions (Figure 12).

Beyond, the higher abundance of the Proteobacteria phylum and the Succinivibrionaceae family in the corn silage-based samples and the therewith associated increase in succinate formation might explain the rise of OTUs and proteins of the Acidaminococcaceae family in the samples of the corn silage-based diet (Figure 16). All OTUs of the Acidaminococcaceae family were assigned to the genus of the succinate-fermenting *Succiniclasticum* (van Gylswyk, 1995) and 48 of 70 identified proteins were produced by the asaccharolytic and succinate-utilizing *Phascolarctobacterium* species (Watanabe et al., 2012). *Succiniclasticum* and *Phascolarctobacterium* species of the Acidaminococcaceae family ferment succinate to produce propionate (Figure 16), the most important carbon source for the ruminant's gluconeogenesis (Yost et al., 1977). In total six proteins of *Phascolarctobacterium* species involved in propionate production were identified (Chapter 4.9, Table S7). There are reports about reduced methane emissions under corn silage-based dietary regimen (Beauchemin and McGinn, 2005; van Gastelen et al., 2015). A metagenomics study linked a decrease in methane emissions and abundance of methanogenic archaea to an increased abundance of the Succinivibrionaceae family (Wallace et al., 2015) whose members utilize hydrogen to produce succinate which is then rapidly converted to propionate and in this compete with the most common, hydrogenotrophic methanogenesis (Liu and Whitman, 2008; McCabe et al., 2015). Concerning the present study, the ratio of acetate to propionate, an indicator for methanogenic activity, was lowest in the corn silage-based diet (Table 5). Furthermore, the increased abundance levels of the Succinivibrionaceae family (Figure 10), the higher amount of succinate (Table 5) and the increased abundance of six propionate-producing *Phascolarctobacterium* proteins in the fractions of the corn silage-based diet may indicate a consistency with the above-mentioned investigations.

No OTUs were assigned to the Firmicutes order of Bacillales or the family of Selenomonadaceae with its typical members of the rumen ecosystem, *Selenomonas ruminantium* and *Anaerovibrio lipolytica*. In contrast, 219 proteins were assigned to the family of Selenomonadaceae with the majority belonging to saccharolytic *Selenomonas* and lipolytic *Anaerovibrio* species, respectively. This may emphasize the benefits and the necessity of applying multiple, complementary methods to investigate the microbiomes of complex ecosystems like the rumen. A previous study analyzed the identical sample material by quantitative real-time PCR and reported no dietary impact on the abundance of *S. ruminantium* (Lengowski et al., 2016). Contrarily, the LFQ-values of proteins of the Selenomonadaceae family assessed in the present study revealed a lower abundance

throughout the grass silage-based diet when compared to the corn silage- and the hay-based samples (Figure 10) including proteins involved in propionate formation (Chapter 4.9, Table S7) and with sequence similarity to SLH (Chapter 4.9, Table S3). The Gram-negative staining but phylogenetically Gram-positive rumen anaerobe *S. ruminantium* exhibits peptidoglycan-associated proteins with SLH that play an important role in the maintenance of the cell surface structure (Kojima et al., 2010). Furthermore, *S. ruminantium* species are characterized by their ability to use a broad range of substrates including the fermentation products of other bacteria (Bryant, 1956; Cotta, 1990; Rasmussen, 1993). Cross-feeding between *S. ruminantium* and *Butyrivibrio fibrisolvens* was reported before (Cotta and Zeltwanger, 1995). Another study showed that the co-cultivation of *S. ruminantium* and *B. fibrisolvens* promoted the growth of *S. ruminantium* (Cotta, 1992). Correspondingly, the LFQ-values of 124 Butyrivibrio proteins assessed by metaproteomics exhibited the lowest abundance in the samples of the grass silage-based diet. Thus, the lower abundance of the Selenomonadaceae family in the grass silage-based diet might be linked to a similarly decreased abundance of the *Butyrivibrio* proteins in the respective samples.

The family of Erysipelotrichaceae of the Clostridium subphylum cluster XVII as well belonging to the phylum of Firmicutes exhibited the highest abundance of LFQs and particularly OTUs in the sample fractions of the hay-based diet (Figure 10). Most of the respective OTUs were assigned to the genera of *Sharpea* and *Kandleria* while most proteins were produced by species belonging to the genera of *Coprobacillus*, *Catenibacterium*, and *Eggerthia*. Due to phenotypic, chemotaxonomic and phylogenetic data, it was suggested that *Lactobacillus catenaformis* and *L. vitulinus* should be reclassified into the genera of *Eggerthia* and *Kandleria*, respectively (Salvetti et al., 2011). Similar to *Lactobacillus* species, most members of Erysipelotrichaceae family probably ferment a wide range of sugars to produce mainly lactic acid (Figure 16) as reported for *Sharpea azabuensis* (Morita et al., 2008). This assumption may be supported by the higher amounts of lactate identified in the RF samples of the hay-based diet (Table 5). Moreover, studies of the sheep rumen linked a higher abundance of the Erysipelotrichaceae family in low-methane emitting animals to an increased lactic acid production in which less hydrogen and thus less methane is formed (Kittelmann et al., 2014; Kamke et al., 2016). The higher abundance of the Erysipelotrichaceae family, mainly *Sharpea* species was further associated with an increased abundance of *Megasphaera* species that convert formed lactate to butyrate (Kamke et al., 2016). In contrast, the higher abundance of the Erysipelotrichaceae family in samples of the hay-based diet determined in the present study was not accompanied by an increased abundance of *Megasphaera* species or the

Veillonellaceae family that exhibited the highest abundance in the corn silage-based samples (Chapter 4.9, Table S2). Fernando et al. (2010) reported a similar increase of *Megasphaera* species upon high grain diets. It was hypothesized that the enrichment of the Erysipelotrichaceae family is related to the smaller rumen size and higher ruminal turnover rates of low-methane emitting sheep which favors microorganisms with fast, hetero-fermentative growth on sugars (Kittelmann et al., 2014; Kamke et al., 2016). However, in the present study the rise of the Erysipelotrichaceae family in the larger rumen of the Jersey cows was observed exclusively under hay-based dietary regimen and the ratio of acetate to propionate, an indicator for methanogenic activity was highest in the hay-based diet (Table 5). The findings of the present study suggest further microbiological investigations of the yet sparsely described family of Erysipelotrichaceae to obtain a more valid basis regarding their functional capabilities and their role within the rumen ecosystem.

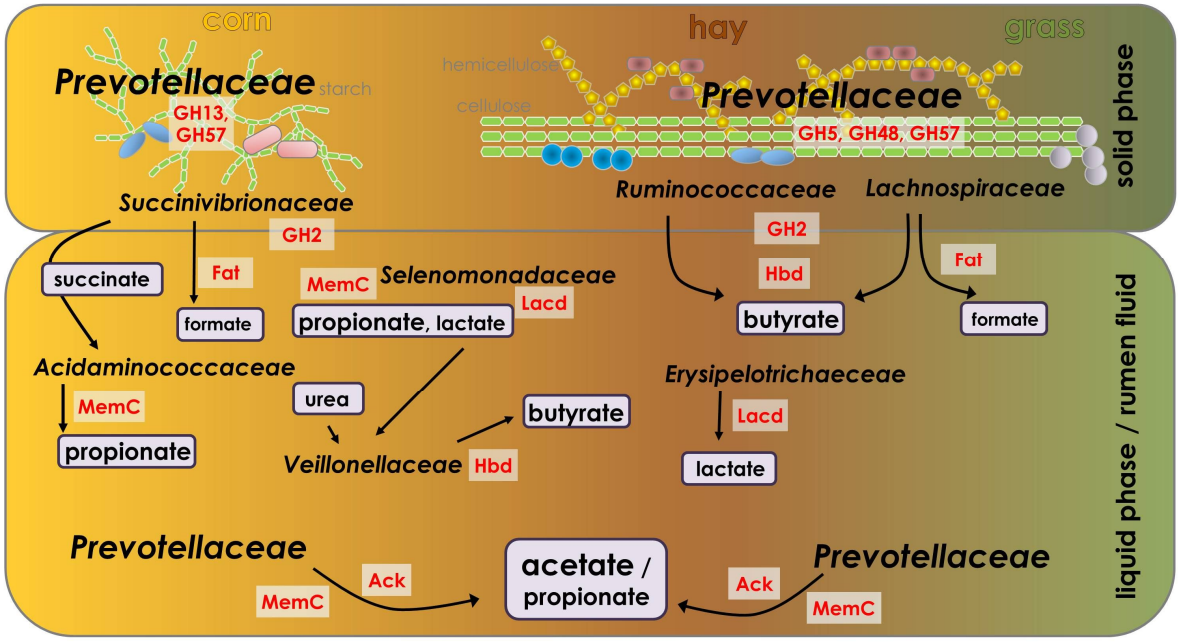

Figure 16 | Consolidation of metaproteomic, 16S rDNA sequencing and metabolomic analyses showing the most active and abundant rumen bacteria in the respective diets and sample fractions and their main fermentation products. Character size of bacterial family names and metabolites are in accordance to their abundance (LFQ for proteins, concentration for metabolites). Ack, acetate kinase; Fat, formate acetyltransferase; GH, glycosyl hydrolase; Hbd, hydroxybutyryl-CoA dehydrogenase; Lacd, lactate dehydrogenase; MemC methylmalonyl-CoA mutase.

Bifidobacterial genomes comprise features necessary for the metabolism of plant-derived complex carbohydrates like glycoside hydrolases and sugar ABC transporters (Pokusaeva et al., 2011). This might explain the higher abundance of proteins and OTUs of the family of Bifidobacteriaceae in the more fiber-rich grass silage- and hay-based samples when compared to the corn silage-based diet (Chapter 4.9, Table S2). Moreover, a *Bifidobacterium*-derived glycosyltransferase of the family GT4 and oligosaccharide, polyol, and lipid transporters from Bifidobacteriaceae were more abundant in the grass silage-based diet and the hay-based samples when compared to the relatively low abundance in the corn-silage based diet. Furthermore, the polysaccharide-degrading capabilities of Actinobacteria including the Bifidobacteriaceae family might be reflected by the higher abundance in the SP fractions of all diets (Figure 10). Contrarily, the study of de Menezes et al. (2011) reported a higher abundance of 16S rRNA gene sequences of Actinobacteria in the LP fractions.

4.5.2 Bacterial Community Composition and Activity is Influenced by the Microenvironment

Besides the effect of animal feed composition, large variations in community structure between the different microenvironments of the rumen ecosystem were observed emphasizing the importance of sample fractionation in rumen studies to cover the effects of treatments throughout the whole ecosystem and its specific functional niches. The difference in bacterial community composition between sample fractions was shown to be greater than the difference between the same fractions of individual animals (Kong et al., 2010). Variations between the different microenvironments have already been discovered before the advent of 16S rRNA gene-based community analysis since the chemical composition of firmly-attached bacteria was shown to be different from the liquid-associated population whereas loosely-attached bacteria were rather similar to the liquid-associated population (Legay-Carmier and Bauchart, 1989) which confirms the findings of the present study. Within diets, the more fiber-rich SP fractions revealed a significant increase in abundance of polysaccharide degrading species of the Firmicutes order of Clostridiales including the families of Lachnospiraceae and Ruminococcaceae as found by OTU and protein abundance levels (Figure 10). The order of Clostridiales includes several cellulolytic *Clostridium* species (Vos et al., 2011) and as well fiber-degrading members of the Eubacteriaceae family (Prins et al., 1972; Van Gylswyk and Van der Toorn, 1985). Most of the OTUs within the order of Clostridiales, excepting Lachnospiraceae and Ruminococcaceae, were assigned to unclassified Clostridiales but no OTUs were assigned to the family of Eubacteriaceae while

most respective proteins were produced by *Eubacterium* and *Clostridium* species. The family of Lachnospiraceae comprises the prominent cellulolytic *B. fibrisolvens* (Bryant and Small, 1956) and the pectinolytic *Lachnospira multiparus* (Duskova and Marounek, 2001). Members of the Lachnospiraceae family exhibit strong hydrolyzing activities with multiple sets of carbohydrate-active enzymes (Stackebrandt, 2014) which may explain the increased abundance in the SP fractions of all diets (Figure 10). The group of oligosaccharide, polyol and lipid transporters contained 27 Lachnospiraceae proteins that exhibited the highest abundance in the SP fractions of all diets (43%) when compared to the LP (29%) and the RF (28%) fractions. A similar increase in abundance of Lachnospiraceae species in the fiber-adherent fractions was reported by Larue et al. (2005). Butyrate formation seemed to be performed by the members of the Lachnospiraceae in the two fiber-rich diets (Figure 16) as proteins involved in butyrate formation revealed the highest abundance (Figure 15) and butyrate concentration was higher in the RF fractions of these samples (Table 5). Comparably, protein and OTU abundances of the Ruminococcaceae family with its prominent cellulolytic representatives *Ruminococcus flavefaciens* and *Ruminococcus albus* were not significantly influenced by the dietary treatments but the respective proteins were more abundant in the SP fraction of all diets. A comprehensive study of Henderson et al. (2015) reported a relatively even distribution of *Ruminococcus* species across different diets and host animals comparable to the finding of the present study. The presence of phenylpropionate identified by NMR (Table 5) was reported to be essential for adherence to and degradation of cellulose by *R. albus* (Stack and Hungate, 1984). The phylum of Fibrobacteres is represented by the major cellulose degrader *Fibrobacter succinogenes* that is restricted to cellulose, hemicellulose or glucose as growth substrates (Hungate, 1950; Puniya et al., 2015). However, there were no significant dietary effects on the abundance of Fibrobacteres proteins and OTUs but within diets, the LFQ-values reached their maximum in the SP fractions (Figure 10) similar to the findings of a previous DNA sequencing study (de Menezes et al., 2011). Large differences between protein- and the OTU-based abundances might be explained by the fact that metaproteomic investigations depend on the amount and quality of reference sequences available for database searches (Seifert et al., 2013; Tanca et al., 2013). Currently only 15 genomes and thus, comparably small numbers of annotated protein sequences are available for the phylum of Fibrobacteres which limits mass spectrometry-based identifications. In contrast, 1,863 different 16S rRNA gene sequences of the phylum of Fibrobacteres are deposited in the RDP database.

The phylum of Spirochaetes was more abundant in the SP fractions as determined by both methods (Figure 10) confirming the findings of de Menezes et al. (2011). The majority of proteins and OTUs belonged to *Treponema* species. *Treponema bryantii* is known to interact with cellulolytic bacteria (Stanton and Canale-Parola, 1980) and *T. succinifaciens* is able to ferment carbohydrates (Cwyk and Canale-Parola, 1979) which may explain the increased abundance throughout the SP fractions. Comparably, the low abundant phylum of Synergistetes, characterized in 2009, revealed the highest LFQ-values of proteins and particularly OTUs in the SP fractions (Chapter 4.9, Table S2). There is not much information about the members of this phylum that are present in many anaerobic ecosystems including the gastrointestinal tract of animals but usually appear in low abundance within the respective environments (Jumas-Bilak and Marchandin, 2014). *Synergistes jonesii* was first isolated from the rumen of goat and did not ferment carbohydrates but is thought to be involved in the degradation of plant-derived toxins such as mimosine and thus, might be beneficial for the host animal (Allison et al., 1992). The overall low abundant phyla of Elusimicrobia and Verrucomicrobia were not affected by the different diets or sample fractions. However, the abundance of Elusimicrobia proteins was lower when compared to the respective OTUs whereas the abundance of Verrucomicrobia proteins was higher when compared to the corresponding OTUs (Chapter 4.9, Table S2). A study employing total RNA sequencing as well as targeted RNA- and DNA amplicon sequencing identified the phyla of Elusimicrobia and Verrucomicrobia exclusively in the RNA-based datasets and proposed a higher activity of the respective phyla in the rumen (Li et al., 2016). The present data may confirm a high metabolic activity of the Verrucomicrobia phylum due to the higher abundance of proteins when compared to the respective OTUs.

The Prevotellaceae family comprises common rumen bacteria such as *Prevotella ruminicola*, *P. brevis*, *P. bryantii*, and *P. albensis* and was shown to increase in abundance upon inclusion of concentrate in diets (Henderson et al., 2015). In the present study, the Prevotellaceae family constituted the most dominant bacterial family within the rumen ecosystem as reported before (Kim et al., 2011; Jami and Mizrahi, 2012). The abundance of proteins and OTUs assigned to the Prevotellaceae family was not affected by the different diets but within diets, the LFQ-values were highest in all LP fractions (Figure 10). A higher abundance of the Prevotellaceae family in the liquid rumen fractions has been reported before (Whitford et al., 1998; Kocherginskaya et al., 2001; Pitta et al., 2010). Members of the Prevotellaceae family are characterized by their versatile metabolic capabilities and their ability to utilize a broad range of substrates including peptides, proteins, monosaccharides,

and plant polysaccharides (Miyazaki et al., 1997; Matsui et al., 2000; Purushe et al., 2010) and thus, may not be primarily affected by changes in diet composition. The majority of identified CAZymes were produced by Prevotellaceae species including 46 glycoside hydrolases, 14 glycosyltransferases and six carbohydrate esterases emphasizing their functional prevalence (Chapter 4.9, Table S3). Furthermore, most enzymes involved in acetate and propionate formation were derived from Prevotellaceae species (Figure 16).

The high abundance of polysaccharide-degrading bacteria in the SP fractions described above might have caused an increased availability of monosaccharides and a higher abundance of proteins involved in glycolysis by the Embden-Meyerhof-Parnas pathway. Considering the different texture and the higher amounts of starch in the corn silage-based diet when compared to the more fibrous grass silage- and hay-based diets most monosaccharides possibly were present in the SP fraction of the corn silage-based diet whereas most monosaccharides from the degradation of structural plant polysaccharides were rather present in the LP fractions of the grass silage- and hay-based diets (Figure 14B). Similar to reports of Pitta et al. (2010), more fibrous hay diets included the development of a digesta mat with clearly separated phases whereas wheat-based rumen content was more homogenized without a distinct fibrous mat. Proteins involved in the citric acid cycle showed a remarkably low abundance in the SP fractions of the grass silage- and hay-based diets (Figure 14C) pointing toward a low abundance of substrates like succinate which is in accordance with the abundance values of proteins of the Embden-Meyerhof-Parnas pathway. This might explain as well the high abundance of the pentose phosphate pathway in the LP fractions of the grass silage- and hay-based diets that probably contained most sugars within the fiber-rich diets (Figure 14D).

Most studies of the rumen metabolome identified higher numbers of different metabolites by a combination of NMR- and more sensitive MS-based approaches (Ametaj et al., 2010; Saleem et al., 2012, 2013). However, despite the statistical insignificancies the metabolite concentrations assessed by NMR in the present study further supported the findings of the 16S rRNA gene sequencing and the metaproteomic analysis. Insignificant metabolite patterns might also be related to the 52% identical composition of the experimental diets used in the present study.

4.5.3 Archaeal Community Differs in Sequence and Protein Composition

There were no diet, sample fraction or host related shifts in community structure of archaea probably due to the less versatile metabolic capabilities when compared to bacteria

(Henderson et al., 2015) and the relatively low numbers of identified proteins and OTUs. The findings of the present study support the results of a previous study that analyzed the abundance of total methanogens and the Rumen Cluster C in the same sample material using quantitative real-time PCR (Lengowski et al., 2016). However, in contrast to the bacterial datasets, the archaeal community composition inferred from the quantified proteins differed clearly from the structure obtained by Illumina amplicon sequencing (Figure 11). OTUs were exclusively assigned to three families of methanogens with members of the family Methanobacteriaceae and Methanomassiliicoccaceae being dominant throughout all diets and sample fractions similar to the results of other 16S rRNA gene-based studies (Janssen and Kirs, 2008; Seedorf et al., 2015). On the other hand, the phylogenetic composition inferred from the quantified archaeal proteins depicted a higher diversity including the presence of the Crenarchaeota and Thaumarchaeota phyla that were identified in low abundance in the rumen ecosystem before (Shin et al., 2004; Wang et al., 2016; Jin et al., 2017). The different phylogenetic distributions of archaeal proteins and OTUs might be attributed to the prevalence of *Methanobrevibacter ruminantium* and *M. gottschalkii* of the Methanobacteriaceae family which dominate the ruminal archaeal community (Janssen and Kirs, 2008; St-Pierre and Wright, 2013; Seedorf et al., 2015; Danielsson et al., 2017). High numbers of the respective 16S rRNA genes might have prevented sufficient sequencing reads of low abundant 16S rRNA genes in the present study. Moreover, some archaea are endo- and ectosymbiotically linked to protozoa (Lange et al., 2005). In particular species of Methanobacteriaceae family were found to be associated with protozoa (Janssen and Kirs, 2008). The specific sample preparation protocols for the shotgun metaproteomic analysis focusing on the enrichment of prokaryotic cells might have caused bias against protozoa-associated archaea, which could further explain the differences between the DNA- and protein-derived datasets. The metaproteomic approach identified the alpha, beta and gamma subunits of the methyl-coenzyme M reductase based on 15 proteins (Chapter 4.9, Table S7) that constituted 23% of the total abundance of 55 proteins that belonged to the methane metabolism as defined by the KEGG database. Eleven proteins were produced by species of the Methanobacteriaceae family and four proteins were assigned to the Thermoplasmata phylum. The methyl-coenzyme M reductase catalyzes the final step in the formation of methane (Ragsdale, 2014) and the mcrA gene, encoding the alpha subunit of the methyl-coenzyme M reductase, is preferably used as a functional marker since present in all methanogens (Denman et al., 2007). The abundance of the subunits of the methyl-coenzyme M reductase (Table S1) was lowest in the corn silage-based diet corresponding to the ratio of

acetate to propionate that was lowest in the corn silage-based diet as well (Table 5).

Next to the hydrogenotrophic methanogenesis, methylotrophic methanogens like the family of Methanomassiliicoccaceae belonging to the class of Thermoplasmata utilize compounds like methylamine, dimethylamine, and trimethylamine as their major energy and carbon sources. Trimethylamines are formed by bacteria from plant-derived glycine betaine and cholin (Neill et al., 1978; Mitchell et al., 1979). A trimethylamine corrinoid protein MttC of the methanogenic *archaeon ISO4-H5* was identified in the metaproteomic analysis and the NMR-based analysis found higher concentrations of methylamine, dimethylamine and trimethylamine in the grass silage- and hay-based diets (Table 5) which is supported by other reports that indicate higher glycine betaine concentrations in grass as when compared to corn (Lerma et al., 1991; Chendrimada et al., 2002). The ratio of acetate to propionate was higher in the two fiber-rich diets that further contained more than twice as much methylamines when compared to the corn silage-based diet suggesting a potentially higher activity of methylotrophic methanogenesis in the fiber-rich diets. A previous study suggested to focus on methylamines rather than on hydrogen for mitigating methane emission from the rumen (Poulsen et al., 2013).

4.6 Conclusion

The obtained datasets revealed significant alterations of the structure and function of the microbial communities in response to the dietary treatments as determined unanimously by the protein- and the DNA-based analyses. Certain contrasts between the methods employed clearly emphasized the benefits of using combinations of complementary methods to study the microbiome of complex ecosystems like the rumen. Moreover, tremendous variations in community composition and functional patterns regarding the different microenvironments within the rumen were observed by both methods prompting for the necessity of sample fractionation in rumen studies to cover the effects of applied treatments throughout the whole ecosystem. The role of low abundant phyla such as Elusimicrobia, Synergistetes, Tenericutes, and Verrucomicrobia within the rumen ecosystems remains comparably vague since not much information about the respective species is available. This suggests microbiological investigations of the respective species in pure and mixed cultures to obtain more accurate data regarding their functional capabilities. This study may provide deeper insights into the complicated network of bacterial interactions and adaptions to various substrates. Bioinformatic and technical progress may enhance the metaproteomic coverage of future studies.

4.7 Author Contributions

The authors' contributions are as follows: MR and JS designed the research. SD, AC, JC, UB, MR, and JS conducted the research. SD, AC, JC, and JS analyzed the data. SD and JS wrote the manuscript. All authors read and approved the final version of the manuscript.

4.8 Acknowledgments

The authors gratefully acknowledge the financial support of the Carl Zeiss Stiftung. We would like to thank the ProteomeCenter in Tübingen for their excellent support.

4.9 Supplementary Material

The Supplementary Material for this article can be found online at:
http://journal.frontiersin.org/article/10.3389/fmicb.2017.01605/full#supplementary-material

Table S1 | List of identified proteins and OTUs including the LFQ-values determined by MaxQuant.

Table S2 | Percentage abundance of archaeal and bacterial proteins and OTUs at phyla, class, order, and family level.

Table S3 | dbCAN assignments. List of identified CAZymes.

Table S4 | WebMGA assignments. List of KEGG Orthology (KO) identifiers and Cluster of Orthologous Groups of proteins (COG) assigned to archaeal and bacterial proteins.

Table S5 | List of ABC transporters.

Table S6 | KEGG REACTIONS in carbon metabolism.

Table S7 | List of enzymes involved in short-chain fatty acid production.

4.10 References

Allison, M. J., Mayberry, W. R., McSweeney, C. S., and Stahl, D. A. (1992). *Synergistes jonesii*, gen. nov., sp. nov.: a rumen bacterium that degrades toxic pyridinediols. *Syst. Appl. Microbiol.* 15, 522–529. doi: 10.1016/S0723-2020(11)80111-6

Ametaj, B. N., Zebeli, Q., Saleem, F., Psychogios, N., Lewis, M. J., Dunn, S. M., et al. (2010). Metabolomics reveals unhealthy alterations in rumen metabolism with increased proportion of cereal grain in the diet of dairy cows. *Metabolomics* 6, 583–594. doi: 10.1007/s11306-010-0227-6

Ann Huws, S., Chiariotti, A., Sarubbi, F., Carfi, F., and Pace, V. (2012). Effects of feeding Mediterranean buffalo sorghum silage versus maize silage on the rumen microbiota and milk fatty acid content. *J. Gen. Appl. Microbiol.* 58, 107–112. doi: 10.2323/jgam.58.107

Beauchemin, K. A., and McGinn, S. M. (2005). Methane emissions from feedlot cattle fed barley or corn diets. *J. Anim. Sci.* 83, 653–661. doi: 10.2527/2005.833653x

Belanche, A., de la Fuente, G., Pinloche, E., Newbold, C. J., and Balcells, J. (2012). Effect of diet and absence of protozoa on the rumen microbial community and on the representativeness of bacterial fractions used in the determination of microbial protein synthesis. *J. Anim. Sci.* 90, 3924–3936. doi: 10.2527/jas.2011-4802

Bladen, H. A., Bryant, M. P., and Doetsch, R. N. (1961). A study of bacterial species from the rumen which produce ammonia from protein hydrolyzate. *Appl. Microbiol.* 9, 175–180.

Bray, J. R., and Curtis, J. T. (1957). An ordination of the upland forest communities of southern Wisconsin. *Ecol. Monogr.* 27, 325–349. doi: 10.2307/1942268

Brulc, J. M., Antonopoulos, D. A., Miller, M. E., Wilson, M. K., Yannarell, A. C., Dinsdale, E. A., et al. (2009). Gene-centric metagenomics of the fiber-adherent bovine rumen microbiome reveals forage specific glycoside hydrolases. *Proc. Natl. Acad. Sci. U.S.A.* 106, 1948–1953. doi: 10.1073/pnas.0806191105

Bryant, M. P. (1956). The characteristics of strains of *Selenomonas* isolated from bovine rumen contents. *J. Bacteriol.* 72, 162–167.

Bryant, M. P. (1959). Bacterial species of the rumen. *Bacteriol. Rev.* 23, 125–153.

Bryant, M. P. (2015). "Succinivibrio," in *Bergey's Manual of Systematics of Archaea and Bacteria*, ed W. B. Whitman (New York, NY: John Wiley & Sons, Ltd). 581–586.

Bryant, M. P., and Small, N. (1956). The anaerobic monotrichous butyric acid-producing curved rod-shaped bacteria of the rumen. *J. Bacteriol.* 72, 16–21.

Bryant, M. P., Small, N., Bouma, C., and Chu, H. (1958). *Bacteroides ruminicola*. n. sp. and *Succinimonas amylolytica*; the new genus and species; species of succinic acid-producing anaerobic bacteria of the bovine rumen. *J. Bacteriol.* 76, 15–23.

Burbach, K., Seifert, J., Pieper, D. H., and Camarinha-Silva, A. (2016). Evaluation of DNA extraction kits and phylogenetic diversity of the porcine gastrointestinal tract based on Illumina sequencing of two hypervariable regions. *Microbiologyopen* 5, 70–82. doi: 10.1002/mbo3.312

Camarinha-Silva, A., Jauregui, R., Chaves-Moreno, D., Oxley, A. P., Schaumburg, F., Becker, K., et al. (2014). Comparing the anterior nare bacterial community of two discrete human populations using Illumina amplicon sequencing. *Environ. Microbiol.* 16, 2939–2952. doi: 10.1111/1462-2920.12362

Carberry, C. A., Kenny, D. A., Han, S., McCabe, M. S., and Waters, S. M. (2012). Effect of phenotypic residual feed intake and dietary forage content on the rumen microbial community of beef cattle. *Appl. Environ. Microbiol.* 78, 4949–4958. doi: 10.1128/AEM.07759-11

Chendrimada, T. P., Neto, M. G., Pesti, G. M., Davis, A. J., and Bakalli, R. I. (2002). Determination of the betaine content of feed ingredients using high-performance liquid chromatography. *J. Sci. Food Agric.* 82, 1556–1563. doi: 10.1002/jsfa.1214

Chourey, K., Jansson, J., VerBerkmoes, N., Shah, M., Chavarria, K. L., Tom, L. M., et al. (2010). Direct cellular lysis/protein extraction protocol for soil metaproteomics. *J. Proteome Res.* 9, 6615–6622. doi: 10.1021/pr100787q

Clarke, K. R., and Warwick, R. M. (2001). *Change in Marine Communities: An Approach to Statistical Analysis and Interpretation, 2nd Edn*. Plymouth: PRIMER-E.

Cole, J. R., Wang, Q., Fish, J. A., Chai, B., McGarrell, D. M., Sun, Y., et al. (2014). Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res.* 42, D633–D642. doi: 10.1093/nar/gkt1244

Cotta, M. A. (1990). Utilization of nucleic acids by *Selenomonas ruminantium* and other ruminal bacteria. *Appl. Environ. Microbiol.* 56, 3867–3870.

Cotta, M. A. (1992). Interaction of ruminal bacteria in the production and utilization of maltooligosaccharides from starch. *Appl. Environ. Microbiol.* 58, 48–54.

Cotta, M. A., and Zeltwanger, R. L. (1995). Degradation and utilization of xylan by the ruminal bacteria *Butyrivibrio fibrisolvens* and *Selenomonas ruminantium*. *Appl. Environ. Microbiol.* 61, 4396–4402.

Cox, J., Hein, M. Y., Luber, C. A., Paron, I., Nagaraj, N., and Mann, M. (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell. Proteom.* 13, 2513–2526. doi: 10.1074/mcp.M113.031591

Creevey, C. J., Kelly, W. J., Henderson, G., and Leahy, S. C. (2014). Determining the culturability of the rumen bacterial microbiome. *Microb. Biotechnol.* 7, 467–479. doi: 10.1111/1751-7915.12141

Cwyk, W. M., and Canale-Parola, E. (1979). *Treponema succinifaciens* sp. nov., an anaerobic spirochete from the swine intestine. *Arch. Microbiol.* 122, 231–239. doi: 10.1007/BF00411285

Danielsson, R., Dicksved, J., Sun, L., Gonda, H., Müller, B., Schnürer, A., et al. (2017). Methane production in dairy cows correlates with rumen methanogenic and bacterial community structure. *Front. Microbiol.* 8:226. doi: 10.3389/fmicb.2017.00226

de Menezes, A. B., Lewis, E., O'Donovan, M., O'Neill, B. F., Clipson, N., and Doyle, E. M. (2011). Microbiome analysis of dairy cows fed pasture or total mixed ration diets. *FEMS Microbiol. Ecol.* 78, 256–265. doi: 10.1111/j.1574-6941.2011.01151.x

Denman, S. E., Tomkins, N. W., and McSweeney, C. S. (2007). Quantitation and diversity analysis of ruminal methanogenic populations in response to the antimethanogenic compound bromochloromethane. *FEMS Microbiol. Ecol.* 62, 313–322. doi: 10.1111/j.1574-6941.2007.00394.x

Deusch, S., and Seifert, J. (2015). Catching the tip of the iceberg - evaluation of sample preparation protocols for metaproteomic studies of the rumen microbiota. *Proteomics* 15, 3590–3595. doi: 10.1002/pmic.201400556

Deusch, S., Tilocca, B., Camarinha-Silva, A., and Seifert, J. (2015). News in livestock research - use of Omics-technologies to study the microbiota in the gastrointestinal tract of farm animals. *Comput. Struct. Biotechnol. J.* 13, 55–63. doi: 10.1016/j.csbj.2014.12.005

Duskova, D., and Marounek, M. (2001). Fermentation of pectin and glucose, and activity of pectin-degrading enzymes in the rumen bacterium *Lachnospira multiparus*. *Lett. Appl. Microbiol.* 33, 159–163. doi: 10.1046/j.1472-765x.2001.00970.x

Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C., and Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27, 2194–2200. doi: 10.1093/bioinformatics/btr381

Edwards, J. R., McEwan, N. J., Travis, A., and John Wallace, R. (2004). 16S rDNA library-based analysis of ruminal bacterial diversity. *Anton. Van Leeuw.* 86, 263–281. doi: 10.1023/B:ANTO.0000047942.69033.24

Fernando, S. C., Purvis, H. T. II., Najar, F. Z., Sukharnikov, L. O., Krehbiel, C. R., Nagaraja, T. G., et al. (2010). Rumen microbial population dynamics during adaptation to a high-grain diet. *Appl. Environ. Microbiol.* 76, 7482–7490. doi: 10.1128/AEM.00388-10

Ferrer, M., Ghazi, A., Beloqui, A., Vieites, J. M., Lopez-Cortes, N., Marin-Navarro, J., et al. (2012). Functional metagenomics unveils a multifunctional glycosyl hydrolase from the family 43 catalysing the breakdown of plant polymers in the calf rumen. *PLoS ONE* 7:e38134. doi: 10.1371/journal.pone.0038134

Gressley, T. F., Hall, M. B., and Armentano, L. E. (2011). Ruminant Nutrition Symposium: productivity, digestion, and health responses to hindgut acidosis in ruminants. *J. Anim. Sci.* 89, 1120–1130. doi: 10.2527/jas.2010-3460

Gupta, N., and Pevzner, P. A. (2009). False discovery rates of protein identifications: a strike against the two-peptide rule. *J. Proteome Res.* 8, 4173–4181. doi: 10.1021/pr9004794

Hackmann, T. J., and Spain, J. N. (2010). Invited review: ruminant ecology and evolution: perspectives useful to ruminant livestock research and production. *J. Dairy Sci.* 93, 1320–1334. doi: 10.3168/jds.2009-2071

Hamlin, L. J., and Hungate, R. E. (1956). Culture and physiology of a starch-digesting bacterium (*Bacteroides amylophilus* n. sp.) from the bovine rumen. *J. Bac.* 72, 548–554.

Hansen, S. H., Stensballe, A., Nielsen, P. H., and Herbst, F. A. (2014). Metaproteomics: evaluation of protein extraction from activated sludge. *Proteomics* 14, 2535–2539. doi: 10.1002/pmic.201400167

Henderson, G., Cox, F., Ganesh, S., Jonker, A., Young, W., and Janssen, P. H. (2015). Rumen microbial community composition varies with diet and host, but a core microbiome is found across a wide geographical range. *Sci. Rep.* 5:14567. doi: 10.1038/srep14567

Henderson, G., Cox, F., Kittelmann, S., Miri, V. H., Zethof, M., Noel, S. J., et al. (2013). Effect of DNA extraction methods and sampling techniques on the apparent structure of cow and sheep rumen microbial communities. *PLoS ONE* 8:e74787. doi: 10.1371/journal.pone.0074787

Hespell, R. B. (1992). "The Genera Succinivibrio and Succinimonas," in *The Prokaryotes: A Handbook on the Biology of Bacteria: Ecophysiology, Isolation, Identification, Applications*, eds. A. Balows, H.G. Trüper, M. Dworkin, W. Harder, and K.-H. Schleifer. (New York, NY: Springer), 3979–3982.

Hess, M., Sczyrba, A., Egan, R., Kim, T. W., Chokhawala, H., Schroth, G., et al. (2011). Metagenomic discovery of biomass-degrading genes and genomes from cow rumen. *Science* 331, 463–467. doi: 10.1126/science.1200387

Hettich, R. L., Pan, C., Chourey, K., and Giannone, R. J. (2013). Metaproteomics: harnessing the power of high performance mass spectrometry to identify the suite of proteins that control metabolic activities in microbial communities. *Anal.Chem.* 85, 4203–4214. doi: 10.1021/ac303053e

Heyer, R., Kohrs, F., Benndorf, D., Rapp, E., Kausmann, R., Heiermann, M., et al. (2013). Metaproteome analysis of the microbial communities in agricultural biogas plants. *N. Biotechnol.* 30, 614–622. doi: 10.1016/j.nbt.2013.01.002

Hungate, R. E. (1950). The anaerobic mesophilic cellulolytic bacteria. *Bacteriol. Rev.* 14, 1–49.

Hungate, R. E. (1966). *The Rumen and Its Microbes*. New York, NY; London: Academic Press.

Hungate, R. E., Bryant, M. P., and Mah, R. A. (1964). The rumen bacteria and protozoa. *Annu. Rev. Microbiol.* 18, 131–166. doi: 10.1146/annurev.mi.18.100164.001023

Jagtap, P., Goslinga, J., Kooren, J. A., McGowan, T., Wroblewski, M. S., Seymour, S. L., et al. (2013). A two-step database search method improves sensitivity in peptide sequence matches for metaproteomics and proteogenomics studies. *Proteomics* 13, 1352–1357. doi: 10.1002/pmic.201200352

Jami, E., and Mizrahi, I. (2012). Composition and similarity of bovine rumen microbiota across individual animals. *PLoS ONE* 7:e33306. doi: 10.1371/journal.pone.0033306

Jami, E., White, B. A., and Mizrahi, I. (2014). Potential role of the bovine rumen microbiome in modulating milk composition and feed efficiency. *PLoS ONE* 9:e85423. doi: 10.1371/journal.pone.0085423

Janssen, P. H., and Kirs, M. (2008). Structure of the archaeal community of the rumen. *Appl. Environ. Microbiol.* 74, 3619–3625. doi: 10.1128/AEM.02812-07

Jehmlich, N., Schmidt, F., Hartwich, M., von Bergen, M., Richnow, H. H., and Vogt, C. (2008). Incorporation of carbon and nitrogen atoms into proteins measured by protein-based stable isotope probing (Protein-SIP). *Rapid Commun. Mass Spectrom.* 22, 2889–2897. doi: 10.1002/rcm.3684

Jin, D., Zhao, S., Zheng, N., Bu, D., Beckers, Y., Denman, S. E., et al. (2017). Differences in ureolytic bacterial composition between the rumen digesta and rumen wall based on ureC gene classification. *Front. Microbiol.* 8:385. doi: 10.3389/fmicb.2017.00385

Jumas-Bilak, E., and Marchandin, H. (2014). "The phylum *Synergistetes*," in *The Prokaryotes: Other Major Lineages of Bacteria and the Archaea*, eds E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt, and F. Thompson (Berlin; Heidelberg: Springer), 931–954.

Kamke, J., Kittelmann, S., Soni, P., Li, Y., Tavendale, M., Ganesh, S., et al. (2016). Rumen metagenome and metatranscriptome analyses of low methane yield sheep reveals a *Sharpea*-enriched microbiome characterised by lactic acid formation and utilisation. *Microbiome* 4:56. doi: 10.1186/s40168-016-0201-2

Kim, M., Morrison, M., and Yu, Z. (2011). Status of the phylogenetic diversity census of ruminal microbiomes. *FEMS Microbiol. Ecol.* 76, 49–63. doi: 10.1111/j.1574-6941.2010.01029.x

Kittelmann, S., Pinares-Patiño, C. S., Seedorf, H., Kirk, M. R., Ganesh, S., McEwan, J. C., et al. (2014). Two different bacterial community types are linked with the low-methane emission trait in sheep. *PLoS ONE* 9:e103171. doi: 10.1371/journal.pone.0103171

Kocherginskaya, S. A., Aminov, R. I., and White, B. A. (2001). Analysis of the rumen bacterial diversity under two different diet conditions using denaturing gradient gel electrophoresis, random sequencing, and statistical ecology approaches. *Anaerobe* 7, 119–134. doi: 10.1006/anae.2001.0378

Koike, S., Yoshitani, S., Kobayashi, Y., and Tanaka, K. (2003). Phylogenetic analysis of fiber-associated rumen bacterial community and PCR detection of uncultured bacteria. *FEMS Microbiol. Lett.* 229, 23–30. doi: 10.1016/S0378-1097(03)00760-2

Kojima, S., Ko, K. C., Takatsuka, Y., Abe, N., Kaneko, J., Itoh, Y., et al. (2010). Cadaverine covalently linked to peptidoglycan is required for interaction between the peptidoglycan and the periplasm-exposed S-layer-homologous domain of major outer membrane protein Mep45 in *Selenomonas ruminantium*. *J. Bacteriol.* 192, 5953–5961. doi: 10.1128/JB.00417-10

Kong, Y., Teather, R., and Forster, R. (2010). Composition, spatial distribution, and diversity of the bacterial communities in the rumen of cows fed different forages. *FEMS Microbiol. Ecol.* 74, 612–622. doi: 10.1111/j.1574-6941.2010.00977.x

Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., and Schloss, P. D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl. Environ. Microbiol.* 79, 5112–5120. doi: 10.1128/AEM.01043-13

Krause, D. O., and Russell, J. B. (1996). How many ruminal bacteria are there? *J. Dairy Sci.* 79, 1467–1475. doi: 10.3168/jds.S0022-0302(96)76506-2

Lamendella, R., VerBerkmoes, N., and Jansson, J. K. (2012). – Omics' of the mammalian gut—new insights into function. *Curr. Opin. Biotechnol.* 23, 491–500. doi: 10.1016/j.copbio.2012.01.016

Lange, M., Westermann, P., and Ahring, B. K. (2005). Archaea in protozoa and metazoa. *Appl. Microbiol. Biotechnol.* 66, 465–474. doi: 10.1007/s00253-004-1790-4

Larue, R., Yu, Z., Parisi, V. A., Egan, A. R., and Morrison, M. (2005). Novel microbial diversity adherent to plant biomass in the herbivore gastrointestinal tract, as revealed by ribosomal intergenic spacer analysis and rrs gene sequencing. *Environ. Microbiol.* 7, 530–543. doi: 10.1111/j.1462-2920.2005.00721.x

Lee, H. J., Jung, J. Y., Oh, Y. K., Lee, S. S., Madsen, E. L., and Jeon, C. O. (2012). Comparative survey of rumen microbial communities and metabolites across one caprine and three bovine groups, using bar-coded pyrosequencing and (1)H nuclear magnetic resonance spectroscopy. *Appl. Environ. Microbiol.* 78, 5983–5993. doi: 10.1128/AEM.00104-12

Legay-Carmier, F., and Bauchart, D. (1989). Distribution of bacteria in the rumen contents of dairy cows given a diet supplemented with soya-bean oil. *Br. J. Nutr.* 61, 725–740. doi: 10.1079/BJN19890158

Lengowski, M. B., Witzig, M., Möhring, J., Seyfang, G. M., and Rodehutscord, M. (2016). Effects of corn silage and grass silage in ruminant rations on diurnal changes of microbial populations in the rumen of dairy cows. *Anaerobe* 42, 6–16. doi: 10.1016/j.anaerobe.2016.07.004

Lerma, C., Rich, P. J., Ju, G. C., Yang, W.-J., Hanson, A. D., and Rhodes, D. (1991). Betaine deficiency in maize: complementation tests and metabolic basis. *Plant Physiol.* 95, 1113–1119. doi: 10.1104/pp.95.4.1113

Ley, R. E., Hamady, M., Lozupone, C., Turnbaugh, P. J., Ramey, R. R., Bircher, J. S., et al. (2008). Evolution of mammals and their gut microbes. *Science* 320, 1647–1651. doi: 10.1126/science.1155725

Li, F., Henderson, G., Sun, X., Cox, F., Janssen, P. H., and Guan, L. L. (2016). Taxonomic assessment of rumen microbiota using total RNA and targeted amplicon sequencing approaches. *Front. Microbiol.* 7:987. doi: 10.3389/fmicb.2016.00987

Liu, Y., and Whitman, W. B. (2008). Metabolic, phylogenetic, and ecological diversity of the methanogenic archaea. *Ann. N. Y. Acad. Sci.* 1125, 171–189. doi: 10.1196/annals.1419.019

Mackie, R. I. (2000). "Molecular ecology and diversity in gut microbial ecosystems," in *Ruminant Physiology: Digestion, Metabolism, Growth and Reproduction*, eds C. R. Bell, M. Brylinsky, and P. Johnson-Green (Halifax: Atlantic Canada Society for Microbial Ecology), 61–77.

Mackie, R. I. (2002). Mutualistic fermentative digestion in the gastrointestinal tract: diversity and evolution. *Integr. Comp. Biol.* 42, 319–326. doi: 10.1093/icb/42.2.319

Mackie, R. I., and Cann, I. K. O. (2005). *A Review of Gastrointestinal Microbiology with Special Emphasis on Molecular Microbial Ecology Approaches.* Dordrecht: Springer.

Matsui, H., Ogata, K., Tajima, K., Nakamura, M., Nagamine, T., Aminov, R. I., et al. (2000). Phenotypic characterization of polysaccharidases produced by four *Prevotella* type strains. *Curr. Microbiol.* 41, 45–49. doi: 10.1007/s002840010089

McAllister, T. A., Bae, H. D., Jones, G. A., and Cheng, K. J. (1994). Microbial attachment and feed digestion in the rumen. *J. Anim. Sci.* 72, 3004–3018. doi: 10.2527/1994.72113004x

McCabe, M. S., Cormican, P., Keogh, K., O'Connor, A., O'Hara, E., Palladino, R. A., et al. (2015). Illumina MiSeq phylogenetic amplicon sequencing shows a large reduction of an uncharacterised *Succinivibrionaceae* and an increase of the *Methanobrevibacter gottschalkii* clade in feed restricted cattle. *PLoS ONE* 10:e0133234. doi: 10.1371/journal.pone.0133234

McMichael, A. J., Powles, J. W., Butler, C. D., and Uauy, R. (2007). Food, livestock production, energy, climate change, and health. *Lancet* 370, 1253–1263. doi: 10.1016/S0140-6736(07)61256-2

McSweeney, C. S., Blackall, L. L., Collins, E., Conlan, L. L., Webb, R. I., Denman, S. E., et al. (2005). Enrichment, isolation and characterisation of ruminal bacteria that degrade non-protein amino acids from the tropical legume *Acacia angustissima*. *Anim. Feed Sci. Technol.* 121, 191–204. doi: 10.1016/j.anifeedsci.2005.02.018

Miron, J., Ben-Ghedalia, D., and Morrison, M. (2001). Invited review: adhesion mechanisms of rumen cellulolytic bacteria. *J. Dairy Sci.* 84, 1294–1309. doi: 10.3168/jds.S0022-0302(01)70159-2

Mitchell, A. D., Chappell, A., and Knox, K. L. (1979). Metabolism of betaine in the ruminant. *J. Anim. Sci.* 49, 764–774. doi: 10.2527/jas1979.493764x

Miyazaki, K., Martin, J. C., Marinsek-Logar, R., and Flint, H. J. (1997). Degradation and utilization of xylans by the rumen anaerobe *Prevotella bryantii* (formerly *P. ruminicola* subsp. *brevis*) B14. *Anaerobe* 3, 373–381. doi: 10.1006/anae.1997.0125

Mizrahi, I. (2011). "The role of the rumen microbiota in determining the feed efficiency of dairy cows," in *Beneficial Microorganisms in Multicellular Life Forms*, eds E. Rosenberg and U. Gophna (Berlin; Heidelberg: Springer), 203–210.

Morita, H., Shiratori, C., Murakami, M., Takami, H., Toh, H., Kato, Y., et al. (2008). *Sharpea azabuensis* gen. nov., sp. nov., a Gram-positive, strictly anaerobic bacterium isolated from the faeces of thoroughbred horses. *Int. J. Syst. Evol. Microb.* 58, 2682–2686. doi: 10.1099/ijs.0.65543-0

Muth, T., Renard, B. Y., and Martens, L. (2016). Metaproteomic data analysis at a glance: advances in computational microbial community proteomics. *Expert. Rev. Proteomics* 13, 757–769. doi: 10.1080/14789450.2016.1209418

Neill, A. R., Grime, D. W., and Dawson, R. M. (1978). Conversion of choline methyl groups through trimethylamine into methane in the rumen. *Biochem. J.* 170, 529–535. doi: 10.1042/bj1700529

Petri, R. M., Schwaiger, T., Penner, G. B., Beauchemin, K. A., Forster, R. J., McKinnon, J. J., et al. (2013). Characterization of the core rumen microbiome in cattle during transition from forage to concentrate as well as during and after an acidotic challenge. *PLoS ONE* 8:e83424. doi: 10.1371/journal.pone.0083424

Pitta, D. W., Pinchak, E., Dowd, S. E., Osterstock, J., Gontcharova, V., Youn, E., et al. (2010). Rumen bacterial diversity dynamics associated with changing from bermudagrass hay to grazed winter wheat diets. *Microb. Ecol.* 59, 511–522. doi: 10.1007/s00248-009-9609-6

Pokusaeva, K., Fitzgerald, G. F., and van Sinderen, D. (2011). Carbohydrate metabolism in Bifidobacteria. *Genes Nutr.* 6, 285–306. doi: 10.1007/s12263-010-0206-6

Polansky, O., Sekelova, Z., Faldynova, M., Sebkova, A., Sisak, F., and Rychlik, I. (2015). Important metabolic pathways and biological processes expressed by chicken cecal microbiota. *Appl. Environ. Microbiol.* 82, 1569–1576. doi: 10.1128/AEM.03473-15

Poulsen, M., Schwab, C., Jensen, B. B., Engberg, R. M., Spang, A., Canibe, N., et al. (2013). Methylotrophic methanogenic *Thermoplasmata* implicated in reduced methane emissions from bovine rumen. *Nat. Commun.* 4, 1428. doi: 10.1038/ncomms2432

Prins, R. A., Van Vugt, F., Hungate, R. E., and Van Vorstenbosch, C. J. (1972). A comparison of strains of *Eubacterium cellulosolvens* from the rumen. *Anton. Leeuw. J. Microb.* 38, 153–161. doi: 10.1007/BF02328087

Puniya, A. K., Singh, R., and Kamra, D. N. (2015). *Rumen Microbiology: From Evolution to Revolution.* New Delhi: Springer India.

Purushe, J., Fouts, D. E., Morrison, M., White, B. A., Mackie, R. I., Coutinho, P. M., et al. (2010). Comparative genome analysis of *Prevotella ruminicola* and *Prevotella bryantii*: insights into their environmental niche. *Microb. Ecol.* 60, 721–729. doi: 10.1007/s00248-010-9692-8

Ragsdale, S. W. (2014). "Biochemistry of methyl-coenzyme M reductase: the nickel metalloenzyme that catalyzes the final step in synthesis and the first step in anaerobic oxidation of the greenhouse gas methane," in *The Metal-Driven Biogeochemistry of Gaseous Compounds in the Environment*, eds P. M. H. Kroneck and M. E. S. Torres (Dordrecht: Springer), 125–145.

Rappsilber, J., Mann, M., and Ishihama, Y. (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* 2, 1896–1906. doi: 10.1038/nprot.2007.261

Rasmussen, M. A. (1993). Isolation and characterization of *Selenomonas ruminantium* strains capable of 2-deoxyribose utilization. *Appl. Environ. Microbiol.* 59, 2077–2081.

Russell, J. B., and Rychlik, J. L. (2001). Factors that alter rumen microbial ecology. *Science* 292, 1119–1122. doi: 10.1126/science.1058830

Saleem, F., Ametaj, B. N., Bouatra, S., Mandal, R., Zebeli, Q., Dunn, S. M., et al. (2012). A metabolomics approach to uncover the effects of grain diets on rumen health in dairy cows. *J. Dairy Sci.* 95, 6606–6623. doi: 10.3168/jds.2012-5403

Saleem, F., Bouatra, S., Guo, A., Psychogios, N., Mandal, R., Dunn, S., et al. (2013). The bovine ruminal fluid metabolome. *Metabolomics* 9, 360–378. doi: 10.1007/s11306-012-0458-9

Salvetti, E., Felis, G. E., Dellaglio, F., Castioni, A., Torriani, S., and Lawson, P. A. (2011). Reclassification of *Lactobacillus catenaformis* (Eggerth 1935, Moore and Holdeman 1970) and *Lactobacillus vitulinus* (Sharpe et al. 1973) as *Eggerthia catenaformis* gen. nov., comb. nov. and *Kandleria vitulina* gen. nov., comb. nov., respectively. *Int. J. Syst. Evol. Microbiol.* 61, 2520–2524. doi: 10.1099/ijs.0.029231-0

Saude, E. J., Slupsky, C. M., and Sykes, B. D. (2006). Optimization of NMR analysis of biological fluids for quantitative accuracy. *Metabolomics* 2, 113–123. doi: 10.1007/s11306-006-0023-5

Seedorf, H., Kittelmann, S., and Janssen, P. H. (2015). Few highly abundant operational taxonomic units dominate within rumen methanogenic archaeal species in New Zealand sheep and cattle. *Appl. Environ. Microbiol.* 81, 986–995. doi: 10.1128/AEM.03018-14

Seifert, J., Herbst, F.-A., Nielsen, P. H., Planes, F. J., Jehmlich, N., Ferrer, M., et al. (2013). Bioinformatic progress and applications in metaproteogenomics for bridging the gap between genomic sequences and metabolic functions in microbial communities. *Proteomics* 13, 2786–2804. doi: 10.1002/pmic.201200566

Shabat, S. K. B., Sasson, G., Doron-Faigenboim, A., Durman, T., Yaacoby, S., Berg Miller, M. E., et al. (2016). Specific microbiome-dependent mechanisms underlie the energy harvest efficiency of ruminants. *ISME J.* 10, 2958–2972. doi: 10.1038/ismej.2016.62

Shin, E. C., Choi, B. R., Lim, W. J., Hong, S. Y., An, C. L., Cho, K. M., et al. (2004). Phylogenetic analysis of archaea in three fractions of cow rumen based on the 16S rDNA sequence. *Anaerobe* 10, 313–319. doi: 10.1016/j.anaerobe.2004.08.002

Stack, R. J., and Hungate, R. E. (1984). Effect of 3-phenylpropanoic acid on capsule and cellulases of *Ruminococcus albus* 8. *Appl. Environ. Microbiol.* 48, 218–223.

Stackebrandt, E. (2014). "The family Lachnospiraceae," in *The Prokaryotes: Firmicutes and Tenericutes*, eds E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt, and F. Thompson (Berlin; Heidelberg: Springer), 197–201.

Stanton, T. B., and Canale-Parola, E. (1980). *Treponema bryantii* sp. nov., a rumen spirochete that interacts with cellulolytic bacteria. *Arch. Microbiol.* 127, 145–156. doi: 10.1007/BF00428018

St-Pierre, B., and Wright, A. D. (2013). Diversity of gut methanogens in herbivorous animals. *Animal* 7, 49–56. doi: 10.1017/S1751731112000912

Tajima, K., Aminov, R. I., Nagamine, T., Matsui, H., Nakamura, M., and Benno, Y. (2001). Diet-dependent shifts in the bacterial population of the rumen revealed with Real-Time PCR. *Appl. Environ. Microbiol.* 67, 2766–2774. doi: 10.1128/AEM.67.6.2766-2774.2001

Tanca, A., Palomba, A., Deligios, M., Cubeddu, T., Fraumene, C., Biosa, G., et al. (2013). Evaluating the impact of different sequence databases on metaproteome analysis: insights from a lab-assembled microbial mixture. *PLoS ONE* 8:e82981. doi: 10.1371/journal.pone.0082981

Tanca, A., Palomba, A., Fraumene, C., Pagnozzi, D., Manghina, V., Deligios, M., et al. (2016). The impact of sequence database choice on metaproteomic results in gut microbiota studies. *Microbiome* 4:51. doi: 10.1186/s40168-016-0196-8

Thoetkiattikul, H., Mhuantong, W., Laothanachareon, T., Tangphatsornruang, S., Pattarajinda, V., Eurwilaichitr, L., et al. (2013). Comparative analysis of microbial profiles in cow rumen fed with different dietary fiber by tagged 16S rRNA gene pyrosequencing. *Curr. Microbiol.* 67, 130–137. doi: 10.1007/s00284-013-0336-3

van Gastelen, S., Antunes-Fernandes, E. C., Hettinga, K. A., Klop, G., Alferink, S. J., Hendriks, W. H., et al. (2015). Enteric methane production, rumen volatile fatty acid concentrations, and milk fatty acid composition in lactating Holstein-Friesian cows fed grass silage- or corn silage-based diets. *J. Dairy Sci.* 98, 1915–1927. doi: 10.3168/jds.2014-8552

van Gylswyk, N. O. (1995). *Succiniclasticum ruminis* gen. nov., sp. nov., a ruminal bacterium converting succinate to propionate as the sole energy-yielding mechanism. *Int. J. Syst. Bacteriol.* 45, 297–300. doi: 10.1099/00207713-45-2-297

Van Gylswyk, N. O., and Van der Toorn, J. J. T. K. (1985). *Eubacterium uniforme* sp. nov. and *Eubacterium xylanophilum* sp. nov., fiber-digesting bacteria from the rumina of sheep fed corn stover. *Int. J. Syst. Evol. Microb.* 35, 323–326. doi: 10.1099/00207713-35-3-323

Vizcaino, J. A., Csordas, A., del-Toro, N., Dianes, J. A., Griss, J., Lavidas, I., et al. (2016). 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res.* 44, D447–D456. doi: 10.1093/nar/gkw880

Vos, P., Garrity, G., Jones, D., Krieg, N. R., Ludwig, W., Rainey, F. A., et al. (2011). *Bergey's Manual of Systematic Bacteriology: Volume 3: The Firmicutes*. New York, NY: Springer.

Wallace, R. J., Rooke, J. A., McKain, N., Duthie, C. A., Hyslop, J. J., and Ross, D. W. (2015). The rumen microbial metagenome associated with high methane production in cattle. *BMC Genomics* 16:839. doi: 10.1186/s12864-015-2032-0

Wang, L., Xu, Q., Kong, F., Yang, Y., Wu, D., Mishra, S., et al. (2016). Exploring the goat rumen microbiome from seven days to two years. *PLoS ONE* 11:e0154354. doi: 10.1371/journal.pone.0154354

Wang, Q., Garrity, G. M., Tiedje, J. M., and Cole, J. R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* 73, 5261–5267. doi: 10.1128/AEM.00062-07

Watanabe, Y., Nagai, F., and Morotomi, M. (2012). Characterization of *Phascolarctobacterium succinatutens* sp. nov., an asaccharolytic, succinate-utilizing bacterium isolated from human feces. *Appl. Environ. Microbiol.* 78, 511–518. doi: 10.1128/AEM.06035-11

Weljie, A. M., Newton, J., Mercier, P., Carlson, E., and Slupsky, C. M. (2006). Targeted profiling: quantitative analysis of 1H NMR metabolomics data. *Anal. Chem.* 78, 4430–4442. doi: 10.1021/ac060209g

Whitford, M. F., Forster, R. J., Beard, C. E., Gong, J., and Teather, R. M. (1998). Phylogenetic analysis of rumen bacteria by comparative sequence analysis of cloned 16S rRNA genes. *Anaerobe* 4, 153–163. doi: 10.1006/anae.1998.0155

Wishart, D. S. (2008). Quantitative metabolomics using NMR. *Trends Anal. Chem.* 27, 228–237. doi: 10.1016/j.trac.2007.12.001

Wu, S., Zhu, Z., Fu, L., Niu, B., and Li, W. (2011). WebMGA: a customizable web server for fast metagenomic sequence analysis. *BMC Genomics* 12:444. doi: 10.1186/1471-2164-12-444

Yin, Y., Mao, X., Yang, J., Chen, X., Mao, F., and Xu, Y. (2012). dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* 40, W445–W451. doi: 10.1093/nar/gks479

Yost, W. M., Young, J. W., Schmidt, S. P., and McGilliard, A. D. (1977). Gluconeogenesis in ruminants: propionic acid production from a high-grain diet fed to cattle. *J. Nutr.* 107, 2036–2043

Zhang, R., Zhu, W., Zhu, W., Liu, J., and Mao, S. (2013). Effect of dietary forage sources on rumen microbiota, rumen fermentation and biogenic amines in dairy cows. *J. Sci. Food Agric.* 94, 1886–1895. doi: 10.1002/jsfa.6508

CHAPTER V

DISCUSSION

5. DISCUSSION

The influence of the varying diets on the structure and functional profiles of the archaeal and bacterial communities taking into account the different ecological niches in the rumen as well as associated shifts in metabolome composition of the rumen fluid are the subject matter of the following discussion. Moreover, the rumen core microbiome, present in all animals, diets and sample fractions as well as the potential impact of individual animals on their inherent microbiome, the overall experimental design and the different methods including sample preparation procedures applied to investigate the prokaryotic communities and the metabolomes will be debated in the ensuing sections.

5.1 Experimental Design

Regarding the significance and meaningfulness of the outcomes of this work, one of the most relevant aspects to be considered is the experimental design including the used animals, respectively the breed, age and lactating status, different diet compositions and rumen sampling strategies as well as the overall conditions of the experimental trial. Employing a Latin square design based on three Jersey cows fed rotationally with three different rations provided access to diet- and host-related effects on the rumen microbiome (Figure 17).

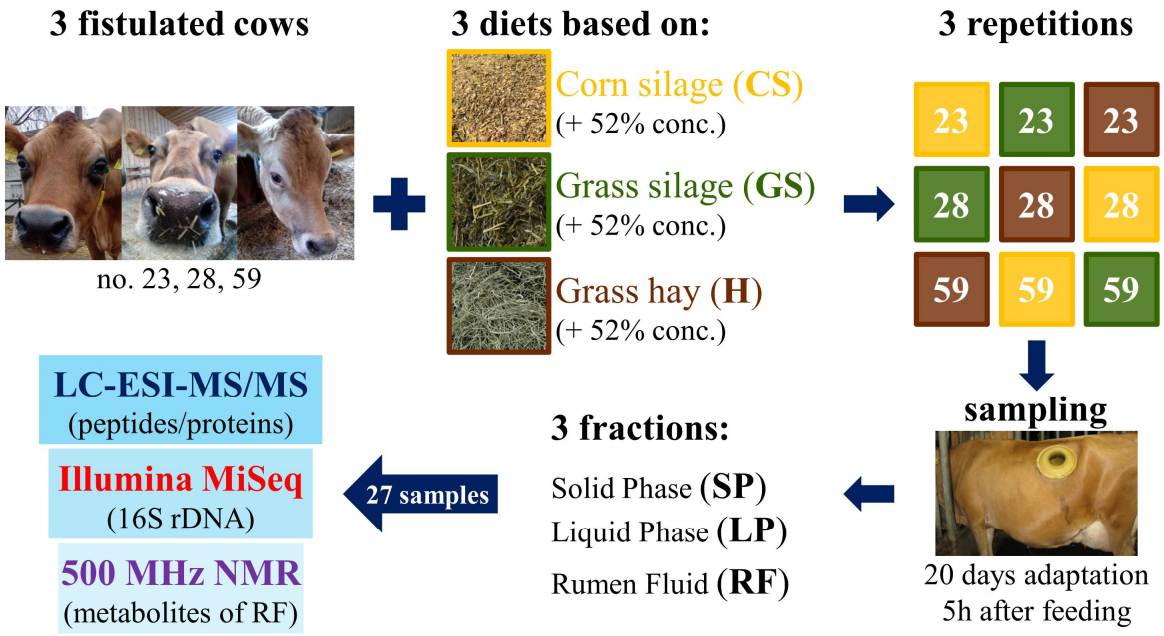

Figure 17| Experimental design based on three Jersey cows, diets, rotational repetitions and sample fractionations.

5.1.1 Animals

The Jersey dairy cattle are considered to be one of the oldest cattle breeds and the second most distributed breed in the world (Felius, 1995; Rimas and Fraser, 2008). The good reputation of the Jersey variety is based on their outstanding economy of production. Milk production of Jerseys is more efficient and exhibits a higher quality based on greater fat and protein contents but also lower overall milk yields when compared to the most popular breed of Holstein-Friesian cattle. Nevertheless, Jerseys produce more milk in relation to body weight with a lower environmental impact than any other dairy breed (Capper and Cady, 2012). Due to the smaller size, more milking cows can be kept per area reducing the maintenance requirements. Remarkable grazing abilities, a higher heat tolerance and fertility along with fewer cases of labour dystocia and less susceptibility to mastitis make Jersey cattle interesting for crossbreeding with other dairy and even meat cattle breeds (Briggs and Briggs, 1980; American Jersey Cattle Association, National All-Jersey Inc., 2006).

The efficiency in milk production, their widespread use across the world and the valuable breeding potential make Jersey cows a suitable variety to investigate the influence of commonly used animal feeds on the rumen microbiota. The three cannulated Jersey cows used in this study (no. 23, 28 and 59) were selected in order to reduce the risk of possible deviations between the animals. Details about the individual animals are listed in Table 6.

Table 6| The characteristics of cows. Age in months, days in lactation and lactation period at the begin of the experimental trial as well as the average daily milk yield, temperature, uptake of total mixed ration and body weight during the experiment.

		24/02/2014		24/02/2014 - 25/04/2014			
Jersey dairy cows	**Age** (mo)	**Lactation** (d)	**Lactation period**	**Milk** (L/d)	**Temp.** (°C)	**TMR** (kg/d)	**Weight** (kg)
23	58	146	3	15.7 *± 3.3*	37.8 *± 0.4*	29.0 *± 7.4*	605 *± 40*
28	69	222	2	12.7 *± 1.2*	38.1 *± 0.2*	20.9 *± 4.9*	449 *± 42*
59	64	234	3	21.4 *± 3.6*	38.3 *± 0.3*	31.3 *± 12.3*	563 *± 16*

Next to the potential impact of individual animals on their microbiome (Guan et al., 2008; Roehe et al., 2016), recent studies investigating Holstein-Frisian and Jersey cows under identical dietary regimen found differences between the breeds regarding the archaeal and bacterial community compositions (King et al., 2011; Cersosimo et al., 2016; Paz et al., 2016). Thus, it is important to consider that, besides the factor of specific dietary treatments the results of this study cannot be transferred unrestrictedly to other cattle breeds or ruminant species.

5.1.2 Diets

Corn silage, grass silage and grass hay are the most commonly used forages for dairy cows and fattening cattle in Europe and North America (Wilkinson et al., 1996). Variations in the chemical composition of the respective forages affect the rumen microbial communities in different directions. In practice however, forages are usually supplemented with concentrate to meet the requirements of high-performance dairy cows. Thus, the main experiments of the current thesis (Chapter 4) are based on diets, respectively total mixed rations containing about 48% of the above-mentioned forages each supplemented with about 52% of identical concentrate (Table 7) in order to obtain conditions as close as possible to the actual situation in animal production. However, interactions or reactions between the concentrate mixture and the corresponding forage sources are possible and cannot be excluded, but were not further considered in this study. The corn silage-based diet (CS) was supplemented with urea to obtain more balanced nitrogen content (Table 7) in comparison to the grass silage (GS)- and grass hay (H)-based diets. Urea constitutes an important non-protein nitrogen source for microbial protein synthesis.

The chemical compositions of total mixed rations, respectively experimental diets were calculated from measured analyses of single components (Table 3) according to the guidelines of the Association of German Agricultural Analytic and Research Institutes (VDLUFA, 2006; Kirchgeßner et al., 2014). The experimental diets differed mainly in neutral detergent fiber, acid detergent fiber and acid detergent lignin, which increased gradually from CS over GS to the H diet (Table 3). The crude protein content of the CS diet was lower than in the more fiber-rich GS and H diets whereas the CS diet contained most non-fiber carbohydrates when compared to the GS and H diets. The GS diet contained most inorganic substances and most crude fat.

Table 7| Composition ratios of the three experimental diets used in this study. CS, corn silage-based diet; GS, grass silage-based diet; H, grass hay-based diet.

	Components (% of dry matter)	**CS diet**	**GS diet**	**H diet**
Forage source	**Corn silage**	47.64	-	-
	Grass silage	-	48.05	-
	Grass hay	-	-	48.05
Concentrate composition	**Barley**	18.36	18.52	
	Wheat	18.36	18.52	
	Soybean meal	7.44	7.51	
	Molasses	5.96	6.01	
	Urea	0.60	-	
	Limestone	0.69	0.45	
	Sodium chloride	0.25	0.25	
	Vitamin mineral premix	0.70	0.70	

5.1.3 Trial Conditions and Sampling Procedure

The three Jersey cows were kept in a herd of six cows in a free-stall barn with *ad libitum* consumption and free access to drinking water. The troughs were filled once daily after morning milking and transponders restricted access only to one trough per cow. Samples were taken 5 h after feeding at 12.30 p.m. after an adaptation time of 20 days for each diet respectively (Figure 17).

Distinct microbial groups such as bacteria or protozoa may need different adaptation times and reports about the time required for the rumen microbiome to re-stabilize after dietary changes are controversial but are generally thought to occur within several days (Hackmann, 2015). Different rumen sampling techniques such as access via cannula or esophageal tubing were shown to be neglectable, not influencing the final microbial community compositions (Lodge-Ivey et al., 2009; Paz et al., 2016). Moreover, different sampling locations within the rumen did not significantly influence the bacterial diversity (Li et al., 2009).

However, there are reports of variations between planktonic, loosely- and tightly-particle-associated bacterial populations (Larue et al., 2005; de Menezes et al., 2011; Kim et al., 2011a; Stiverson et al., 2011). This has already been discovered before the advent of 16S rRNA gene-based community analysis since the chemical composition of firmly-attached

bacteria was shown to be different from the liquid-associated population whereas loosely-attached bacteria were rather similar to the liquid-associated population (Legay-Carmier and Bauchart, 1989). The variety between bacterial communities of distinct sample fractions of the rumen was shown to be even larger as the variety amongst distinct individuals (Kong et al., 2010).

To obtain most representative samples for the present study, rumen matter was collected from five different positions (cranial, caudal, dorsal, ventral, and medial) through a rumen cannula. The samples were squeezed vigorously by hand using disposable polyethylene gloves to obtain the liquid phase (LP) sample fractions, the remains constituted the solid phase (SP) sample fractions. Equal parts of the LP and SP sample fractions were pooled across the five positions. As a third sample fraction ventral rumen fluid (RF) was collected using a vacuum pump. Considering the different functional niches of bacterial populations in the rumen requires subdivision into planktonic species, free-living in the RF and the fiber-adherent communities which should be further separated into the groups of loosely- and tightly-attached species supposed to be present in the LP and the SP fractions, respectively (McAllister et al., 1994). Throughout the whole sampling and pooling procedures, sterile 1000 ml glass beakers, spatulas and hoses were used. Squeezing and pooling was performed as quick as possible to reduce exposure to oxygen and the risk of contamination. In total 27 samples (of three cows, diets and sample fractions) were frozen immediately at -80°C until further processing for Illumina MiSeq amplicon sequencing of the 16S rRNA gene, tandem mass spectrometry-based shotgun metaproteomics and nuclear magnetic resonance measurements of the metabolites in the nine RF fractions (Figure 17).

5.2 Consideration of Applied Methods

The methods applied to analyze the different biomolecules like DNA, protein and metabolites and the therewith associated, specific sample preparation procedures as well as the subsequent bioinformatic analyses and taxonomic assignments are of fundamental importance concerning the accuracy and reliability of the present data and will be discussed in the following paragraphs.

5.2.1 Illumina MiSeq Amplicon Sequencing

Using fragments of the 16S rRNA gene for characterizing prokaryotic communities involves many processing steps each of which each can potentially influence the final community structure and diversity observed. Issues to be considered are related to DNA

extraction method (Henderson et al., 2013) and DNA concentration (Chandler et al., 1997), primer design (Tremblay et al., 2015) and amplified regions of the 16S rRNA gene (Kumar et al., 2011), polymerase-chain reaction (PCR) settings (Ishii and Fukui, 2001) and amplicon purification steps. Moreover the used sequencing platforms, the size of prepared libraries (Gihring et al., 2012), various bioinformatic data analyses and taxonomic assignment procedures (Jovel et al., 2016) including 16S rRNA gene reference databases as well as multiple copies of the 16S rRNA gene (Kembel et al., 2012) have to be considered.

5.2.1.1 DNA Extraction, Quantification and Purity

First, sufficient quantity and quality of extracted DNA are fundamental to all downstream steps. The rumen sample matter is highly heterogeneous and microorganisms are present in different proportions exhibiting a huge diversity involving distinct morphologies, cell wall compositions and chemical characteristics (Legay-Carmier and Bauchart, 1989). These properties affect representative DNA extractions, respectively cell disruption and lysis steps and hence, may favor or discriminate against certain microbial groups or microorganisms which distorts the final image of the microbial community structure (Yu and Morrison, 2004; Bergmann et al., 2010; Henderson et al., 2013; Villegas-Rivera et al., 2013). Studies employing different DNA extraction methods are not comparable (Henderson et al., 2013). Using identical DNA extraction protocols may not guarantee a proper representation of the microbial community structure but enables comparisons between samples or studies. Additionally, DNA extracts are accessible for contaminations with foreign microbial DNA due to non-specific marker genes. Sources of contamination may include the PCR reagents, ultrapure water and the extraction reagents.

In this study the FastDNA™ SPIN Kit for Soil (MP Biomedical, Solon, OH, USA) and 250 mg of each sample fraction were used for DNA extraction according to the manufacturer's instruction with slight modifications as previously described (Burbach et al., 2016). The protocol included a bead-beating step for improved mechanical disruption of Gram-positive bacteria (Feinstein et al., 2009; Smith et al., 2011; Wesolowska-Andersen et al., 2014) which typically exhibit more stable cell walls with thick layers of peptidoglycan. The results of the Illumina MiSeq sequencing of the V1 to V2 regions of the 16S rRNA gene revealed representative proportions of expected ruminal genera belonging to the respective Gram-positive and Gram-negative bacterial consortia indicating a proper DNA extraction procedure (Chapter 5.3).

5.2.1.2 PCR, Regions of the 16S rRNA Gene and Primers

Substances such as humic acids (Bergmann et al., 2010), polysaccharides, urea and bile salts might be present in the sample and inhibit subsequent PCR amplification (Schrader et al., 2012). Other parameters influencing the efficiency of mixed template PCR amplifications are the guanine-cytosine content of the target regions (Pinto and Raskin, 2012), DNA concentration (Polz and Cavanaugh, 1998) and thermocycling conditions (Ishii and Fukui, 2001; Kalle et al., 2013).

Today, the most used sequencing platforms for studies of the gut microbiome are Roche 454 GS-FLX, Illumina MiSeq, Ion Torrent PGM and PacBio SMRT (de la Cuesta-Zuluaga and Escobar, 2016). Most of the next generation sequencing platforms cannot cover the whole length of the 16S rRNA gene (ca. 1,500 base pairs) and thus, only parts of the gene, the hypervariable V1 to V9 regions are used preferably for modern sequencing technologies (Youssef et al., 2009). The respective 16S rRNA gene can be inferred from the surrogate sequence of the hypervariable regions (Juste et al., 2008). Nevertheless, the reliability of the respective correlations differs between the hypervariable regions yielding contrary counts of operational taxonomic units from different hypervariable regions (Schloss, 2010; Yarza et al., 2014). The selection of the hypervariable region of the 16S rRNA gene for amplification influences the determination of microbial community structures (Kumar et al., 2011). Here we chose the V1 to V2 region for amplification of the archaeal and bacterial 16S rRNA gene as used and suggested by other studies (Liu et al., 2007; Camarinha-Silva et al., 2014).

For amplification of the selected hypervariable regions specific broad-range or universal primers, flanking the sequence of interest, must be used. The flanked regions are very constant over a broad scope of bacterial groups enabling proper annealing and amplification of the target sequence (Wang and Qian, 2015). However, mutations may occur in the conserved regions as well and primers with a low coverage may result in biased amplification of the template DNA again causing discrimination or preference of certain microbial groups (Cai et al., 2013; Klindworth et al., 2013).

In this study, to target the V1 to V2 region of the archaeal 16S rRNA gene the primers Arch349F and Arch806R were used as described previously (Lee et al., 2012) and the V1 to V2 region of the 16S rRNA gene of bacteria was amplified using the primers described by Camarinha-Silva et al. (2014). The forward primers contained a six nucleotide barcode (Meyer and Kircher, 2010), a two nucleotide linker (Hamady et al., 2008) and both primers contained sequences complementary to the Illumina specific adaptors.

5.2.1.3 Data Processing and Taxonomic Assignments

Bioinformatic pipelines for the processing of raw sequences (Schloss et al., 2009; Kozich et al., 2013) include the removal of low quality reads, sequencing errors, chimeric sequences (Edgar et al., 2011) and the clustering into operational taxonomic units by various algorithms (DeSantis et al., 2006b; Wang et al., 2007). Subsequently taxonomic assignments using comprehensive 16S rRNA gene reference databases such as Greengenes (DeSantis et al., 2006a), the Ribosomal Database Project (Cole et al., 2014) as well as SILVA (Quast et al., 2013) are performed considering also multiple copies of 16S rRNA gene (Kembel et al., 2012). The choice of database can have stronger effects on the outcome than different assignment methods (Liu et al., 2008). To characterize and compare different microbial communities commonly ecological and statistical evaluations in from of α and β diversity are applied. Sequence processing and assignments of the current study are described in Chapter 4.3.4. Different prokaryotic species exhibit varying numbers of 16S rRNA gene copies that also depend on the respective growth stage (Ogata et al., 1997; Fogel et al., 1999; Peterka and Avgustin, 2001; Tajima et al., 2001; Sirohi et al., 2012). Moreover, the different copies of the 16S rRNA gene of a species may exhibit varying sequences causing an identification of several ribotypes for a single microorganism, which may affect the precision of the microbiome structure depiction (Dahllöf, 2002; Crosby and Criddle, 2003). A more precise picture of the abundance and diversity of the investigated communities can be obtained by considering the different numbers of 16S rRNA gene copies (Kembel et al., 2012). Nevertheless, the numbers of 16S rRNA gene copies are known for only a few ruminal species (Acinas et al., 2004) and were generally not considered in most studies of the rumen microbiome which also applies to the present study.

5.2.2 Shotgun Metaproteomics

Shotgun or bottom-up metaproteomics based on tandem mass spectrometry stands for the characterization of the entire protein complement of an environmental microbiota at a given point in time entailing phylogenetic and functional information about the active consortia of the respective microbiota (Wilmes and Bond, 2004; 2006; Hettich et al., 2013; Seifert et al., 2013; Tanca et al., 2014; Wilmes et al., 2015).

Focusing at the actual functions out of numerous possibilities that are encoded by the genomes of a microbiome is essential for an improved understanding of the rumen microbiome including the interactions of its different members as well as host-microbe interactions. For this purpose, proteins are best suited since they represent the function of a

cell. Shifts in microbiome structure are not necessarily linked to alterations of functional profiles or pathways due to the functional redundancy of microorganisms.

The intention of shotgun metaproteomics is the identification of peptides retrieved from tryptically digested proteins. For separation, the peptides are injected onto a nano-high performance liquid chromatography (nano-HPLC) that is coupled to a tandem mass spectrometer (MS/MS). Subsequently, peptides are identified by database searches using the obtained MS/MS spectra and therefrom the proteins present in the sample are derived (Chapter 5.2.2.3).

Developments in metaproteomics include the step from two-dimensional gel-electrophoresis and matrix assisted laser desorption/ionization time-of flight (MALDI-ToF) mass spectrometry to one-dimensional gel-electrophoresis combined with liquid chromatography and tandem mass spectrometry (GeLC-MS/MS). In contrast to the limited analysis of single protein-spots from two-dimensional gel-electrophoresis (Snelling and Wallace, 2017) the GeLC-MS/MS approach can analyze complex and comprehensive mixtures of peptides within a single measurement run (Dzieciatkowska et al., 2014). Technical progress in liquid chromatography (LC) involving reverse phase columns based on C_{18} molecules enabled the separation of thousands of peptides due to their hydrophobicity. Moreover, improved accuracy of mass determination by mass spectrometers, higher resolutions and dynamic ranges, promoted the successfully feasible application of shotgun metaproteomic studies respectively GeLC-MS/MS approaches (VerBerkmoes et al., 2009; Yates et al., 2009) as employed within this project (Chapter 3 and Chapter 4). GeLC-MS/MS approaches are not restricted to microbiome analyses only but were also used to investigate the ruminal epithelial tissue (Yang et al., 2013).

Important steps of the present shotgun metaproteomic workflow are depicted in Figure 18. First steps were the rumen sample collection as described in Chapter 5.1.3 and the particularly crucial sample preparation procedure as described in the following paragraph (Chapter 5.2.2.1). Further steps included the unbiased proteome extraction and the capture, purification and separation of extracted proteins by one -dimensional sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) followed by tryptic in-gel digestion and peptide extraction including peptide purification (Chapter 5.2.2.2). Final steps were the LC-MS/MS measurements and the subsequent bioinformatic protein identification for defining a sample specific in-house search database required for reliable protein quantification. The closing step comprised the bioinformatic data analyses of the functions of quantified proteins to find biological relevancies (Chapter 5.2.2.3).

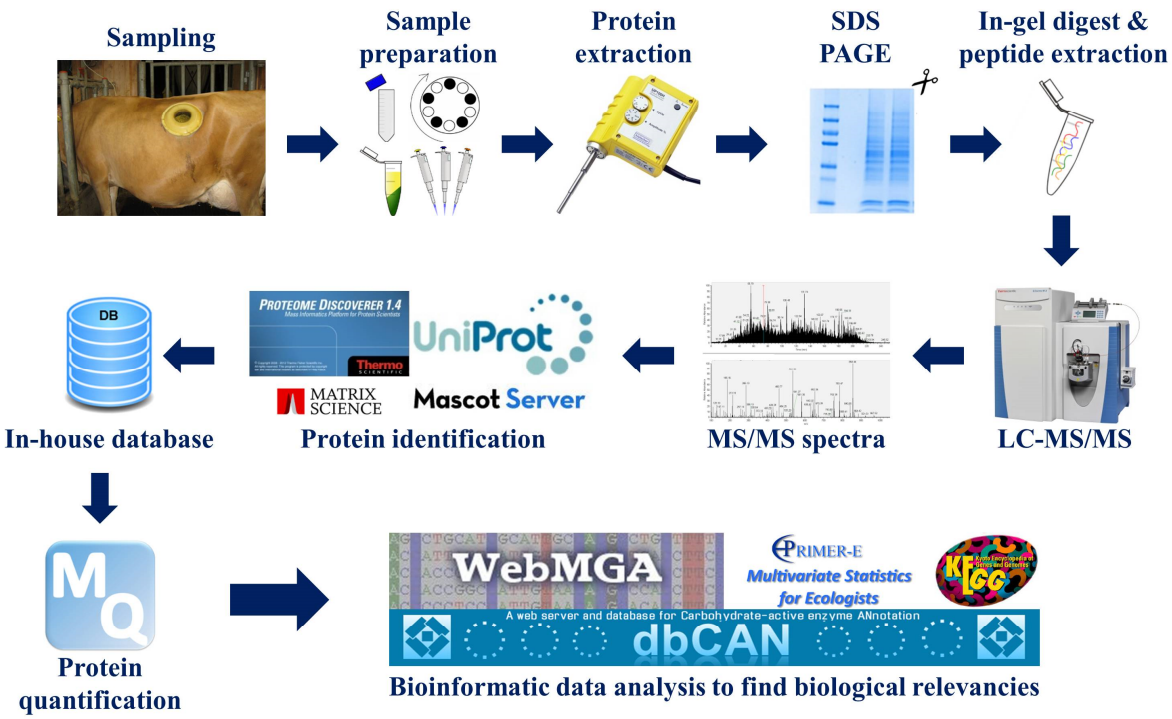

Figure 18| Scheme of the main crucial steps of the present metaproteomic workflow.

5.2.2.1 Sample Preparation

The rumen matter is highly heterogeneous comprising eukaryotic cells from rumen fungi and protozoa as well as prokaryotic cells of archaea and bacteria. Moreover, epithelial cells of the animals and enormous amounts of plant cells from ingested feed are present. Metaproteomic studies require the extraction of proteins preferably of the desired target organisms to increase the coverage of the respective metaproteome as well as the reliability of protein identifications. This implicates the avoidance of undesired proteins that might be present in the environmental sample. Thus, optimized sample preparation protocols are necessary in order to retrieve, in this case, enhanced amounts of prokaryotic instead of plant- and bovine-derived cells before protein extraction and subsequent LC-MS/MS analysis. There is no universal sample preparation method and in contrast to nucleic acids, proteins cannot be enriched or amplified by PCR, hence different environmental samples require the establishment of individual sample preparation protocols (de Castro et al., 2013).

Prokaryotic cells exhibit about 10^3 smaller sizes and masses when compared to eukaryotic cells. Hence, there is the possibility to separate prokaryotic cells from the residual matter by filtration and centrifugation. However, up to 75% of the rumen bacteria live attached to the plant material (Koike et al., 2003) and have to be detached prior to the enrichment procedure. Furthermore, to obtain representative samples, it is important to

consider the different ecological niches within the rumen as described in Chapter 5.1.3. Several studies evaluated methods for isolating rumen microorganisms that include various chemical and physical treatments, e.g. methylcellulose, chilling or blending (Leedle et al., 1987; Whitehouse et al., 1994; Ranilla and Carro, 2003; Trabalza-Marinucci et al., 2006).

Within this project, we focused on improving the identification numbers and the ratio of prokaryotic to eukaryotic proteins. Different sample preparation protocols for the fiber-adherent and the liquid-associated bacteria and archaea were established and evaluated by the number of possible protein identifications, the ratio between prokaryotic and eukaryotic proteins and in analyzing bias towards specific phylogenetic groups, subcellular localizations and Cluster of Orthologous Groups. Using methylcellulose and common cheesecloth for the detachment and separation of bacterial cells, the identification ratio of prokaryotic to eukaryotic protein was improved (Deusch and Seifert, 2015), in particular for the fiber-adherent bacterial fractions that typically contain more plant-derived proteins (Chapter 3). Concerning the main experiments (Chapter 4) an average identification ratio of archaeal and bacterial proteins to eukaryotic proteins of 2.8 to 1 was accomplished.

Nevertheless, the total numbers of identified proteins depicted only a meager part of the entire prokaryotic metaproteome actually present emphasizing a major obstacle in conducting representative metaproteomic studies requiring further improvements such as enhanced mass spectrometric measurements and more precise, sample-specific search databases for the bioinformatic identification and quantification steps. The respective findings were published in the journal of Proteomics in 2015 (Chapter 3). Simultaneously, this was the first published analysis of the ruminal, prokaryotic metaproteome using shotgun metaproteomics (Snelling and Wallace, 2017; Wallace et al., 2017).

5.2.2.2 Protein Extraction, SDS-PAGE, Digestion and Purification

Prokaryotic cells enriched from rumen samples (Chapter 5.2.2.1) were subjected to protein extraction including the subsequent denaturation and breakage of disulfide bonds, which is necessary for a suitable enzymatic digestion by trypsin.

Proper protein extraction is of significant importance in shotgun metaproteomic studies since all proteins will be assembled from digested peptides (Chapter 5.2.2.3). The first step of the protein extraction is the lysis of the cells. Due to the huge diversity of cell types and different cell membranes such as from Gram-positive and Gram-negative bacteria present in the samples of the current study (Chapter 5.2.1.1), a combination of physical and chemical lysis using SDS and sonication as well as thermal treatments were applied. Combined

extraction methods may enhance a representative protein coverage (Hansen et al., 2014). The taxonomic distribution of identified proteins (Chapter 5.3) indicated a proper protein extraction procedure based on representative proportions of expected ruminal genera belonging to Gram-positive and Gram-negative bacteria.

SDS denatures proteins by disrupting non-covalent bonds and causes a similar negative charge of all proteins enabling separation of proteins in SDS-PAGE according to size exclusively. Moreover, SDS also solubilizes membrane-associated proteins (Seddon et al., 2004). Further, an antibiotic (chloramphenicol), the protease-inhibitor phenylmethylsulfonyl fluoride (PMSF) and a nuclease (benzonase) were added to the samples to prevent protein degradation and eliminate nucleic acids. The protein extracts were precipitated using trichloroacetic acid (TCA) and washed thoroughly with chilled acetone to remove the chemicals used during protein extraction (Chourey et al., 2010; Tanca et al., 2014).

The next step included the capture, purification and separation of the protein extracts by SDS-PAGE, which was challenged by the presence of interfering humic compounds in the protein extract as stated before (Heyer et al., 2013). Moreover, humic substances were reported to possibly degrade proteins (Solaiman et al., 2007) and may impede the subsequent LC-MS/MS measurements (Chourey et al., 2010). Potentially interfering substances still present after precipitation might be entrapped in the gel matrix, further purifying the protein sample. Excised gel pieces were subjected to in gel-digestion using trypsin as demonstrated before (Jehmlich et al., 2008). Peptides from in-gel digested proteins were eluted, purified and desalted using Stage tips as described by Rappsilber et al. (2007).

Alternatively, the filter-aided sample preparation (FASP) method (Wisniewski et al., 2009) combining the advantages of in-gel and in-solution digestion might be used for sample cleanup and protein digestion.

Trypsin is used preferably in metaproteomic studies due its specificity and the generation of positively charged C- and N-termini improving the quality of MS/MS spectra since usually, the positive ion mode is used in shotgun metaproteomics (Mann et al., 2001). Trypsin cleaves proteins at the carboxyl side of arginine and lysine generating short peptides of a length of about seven to 20 amino acids further promoting MS/MS analyses (Saveliev et al., 2012; Wu et al., 2012). Besides trypsin, other proteases might be used (Giansanti et al., 2016).

Apart from technical advances, the capacities of nano-HPLC and mass spectrometers are still limited suggesting the fractionation of complex protein or peptide mixtures prior to the LC-MS/MS analysis, which can be achieved by extended gel-electrophoresis and

partitioning of the gel for separate analysis. Alternatively, a complex peptide mixture of a short-run gel electrophoresis can be analyzed using an extended LC gradient respectively measurement time or by fractionation of peptide mixtures using strong cation exchange (SCX) chromatography before LC-MS/MS analysis.

5.2.2.3 LC-MS/MS Measurements and Bioinformatic Analyses

To improve the throughput in LC-MS/MS analyses of shotgun metaproteomics, peptides instead of total proteins are analyzed. However, top-down proteomics of intact proteins include improved detection of small proteins, proteoforms and post-translational modifications and may complement the bottom-up approaches (Catherman et al., 2014).

Mass analysis or more precise the determination of the mass to charge ratio (m/z) of peptides employs electromagnetic fields in a vacuum, therefore peptides have to be electrically charged (Walther and Mann, 2010). Thus, as depicted in Figure 19, in bottom-up tandem mass spectrometry peptides are separated by nano-HPLC followed by ionization and transfer to the gas phase with an ion source such as the electrospray ionization (Fenn et al., 1989). Mass analyzers isolate the charged peptides according to their m/z ratio and subsequently, the mass detector measures the respective intensities. From the first stage of mass analysis (MS1) the precise masses of the peptides are given. During the second stage of the mass analysis (MS2), each of the positively charged precursor ions is further fragmented along its backbone into daughter ions, for example by higher energy collisional dissociation (HCD) that generates mostly b- (N-terminus) and y-ions (C-terminus). The obtained MS/MS spectra then provide additional information about the amino acid sequence of the respective peptide.

State of the art mass spectrometers comprise a precise mass analyzer for measuring the mass of peptide precursor ions and a second mass analyzer for rapid detection of the masses of daughter or fragment ions. Variants of the ion source, mass analyzer and detector are available and are used in different combinations. There are as well different approaches for the fragmentation of ions (Walther and Mann, 2010; Graham et al., 2011; Guthals and Bandeira, 2012). A highly precise mass analyzer is the orbitrap also used within this project where ions circulate around an electrode spindle and the m/z ratios are determined via Fourier transformation of the frequency signals (Walther and Mann, 2010).

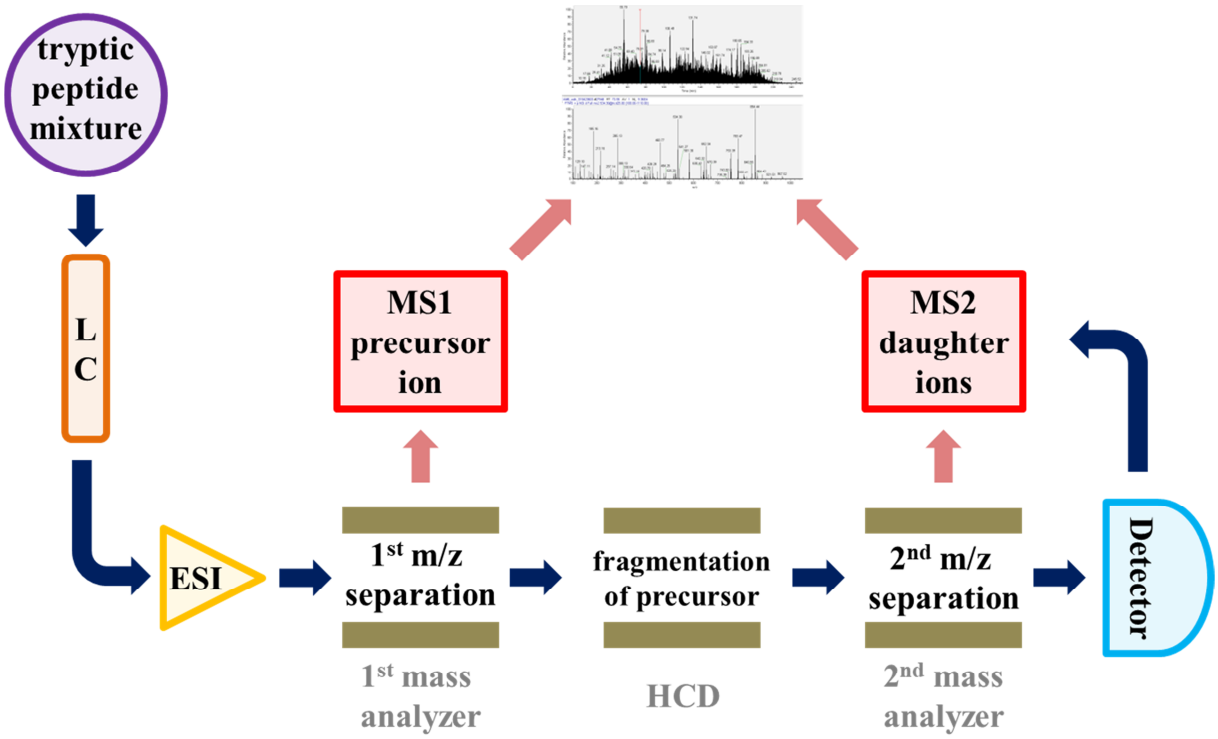

Figure 19| Simplified scheme of spectra acquisition in liquid chromatography tandem mass spectrometry with electrospray ionization (LC-ESI-MS/MS). HCD, higher energy collisional dissociation.

Using specific algorithms and software such as X!Tandem, SEQUEST, OMSSA or Mascot (Eng et al., 1994; Perkins et al., 1999; Craig and Beavis, 2004; Geer et al., 2004) peptide-spectra matches (PSMs) are performed in which each of the obtained MS/MS spectra is compared to protein or nucleic acid databases such as the UniprotKB/TrEMBL database or sample-specific metagenomes to find the best matching peptide sequence. This requires the *in silico* digestion of all entries in the respective databases knowing the specific cleavage sites from trypsin and therefrom the masses of the peptides are calculated. If the calculated peptide mass corresponds to a measured peptide mass, the masses of the respective daughter ions are calculated as well and compared to the experimentally observed values of the fragment ions. In this way, peptide sequences are identified and the entirety of identified peptides is then used to infer the proteins that were present in the analyzed sample (Walther and Mann, 2010; Cottrell, 2011; Mann, 2016). Other possibilities for peptide identification are de novo sequencing and spectral libraries, which both are not commonly used in metaproteomic studies (Muth et al., 2016). Regarding the current project, Mascot in combination with ProteomeDiscoverer was utilized for peptide and protein identification against the NCBInr database (Chapter 3.3) or the UniprotKB/TrEMBL databases (Chapter 4.3.7).

Many proteins contain unique peptides that occur exclusively in the respective proteins, but there are as well peptides that occur in more than one protein causing ambiguity and limit the confidence regarding the origin of the respective peptide. The small peptides generated by trypsin cleavage (Chapter 5.2.2.2) may even increase the number of redundant peptides that potentially belong to more than one protein. This issue is gaining even more importance when considering that microorganisms often exhibit similar protein sequences and that large-sized search databases are used in metaproteomic studies. Moreover, only peptides or proteins present in the databases can be identified and thus, the specificity and size of the database is crucial concerning the reliability of peptide and protein identifications (Vaudel et al., 2011; Tanca et al., 2016). Furthermore, PSM is a probability measure that includes false-positive matches, which can be reduced in employing decoy databases and generally more precise databases (Yadav et al., 2012; Muth et al., 2015). PSMs require statistical validation by suitable false discovery rate (FDR) thresholds.

Addressing these issues, the sizes of the search databases for the main experiment (Chapter 4) were reduced and simultaneously the sample-specificity was increased by using a two-step search approach (Jagtap et al., 2013; Hansen et al., 2014). In this process, the false discovery rates of PSM respectively peptide identifications and associated protein identifications were improved enhancing the confidence of identifications. The first, weakly filtered survey search against all archaeal or bacterial protein entries in the UniprotKB/TrEMBL database yielded an artificial metagenome or in-house database reduced in size but with much higher sample specificity that was subsequently used for more precise protein quantification applying strict filter criteria (Figure 18).

Regarding the first metaproteomic analysis (Chapter 3), only a single search run against the large NCBInr database was performed which consequently lead to an increased probability of false positive protein identifications. However, the focus of the first metaproteomic study was on the development of sample preparation protocols and improved identification ratios (Chapter 5.2.2.1) not requiring precise search databases.

For the comparison of different samples, quantification of the identified proteins is required. Quantification can be achieved in various ways such as label-free quantification by software like MaxQuant (Cox and Mann, 2008; Cox et al., 2014) that was also used for the main experiment of this project (Chapter 4). The label-free quantification intensity values (LFQ) of each protein in each sample obtained from MaxQuant are based on precursor signal intensity and were used for the subsequent taxonomic and functional analyses (Tyanova et al., 2016). Labeling approaches employ metabolic labeling such as SILAC or chemical labeling

like ICAT, iTRAQ and TMT and are more accurate in terms of quantification as reviewed before (Bantscheff et al., 2007; Schulze and Usadel, 2010).

The functions of genes or proteins are often inferred via sequence homology analysis from already annotated genes and proteins introducing bias regarding the functional analyses. Rapidly increasing numbers of sequenced genomes and therewith-associated hypothetical proteins are not experimentally verified and rely mainly on sequence similarities.

However, to link proteins to their respective function there are several databases with different classifications systems available. Cluster of Orthologous Genes (COG), Kyoto Encyclopedia of Genes and Genomes (KEGG), the Carbohydrate-Active EnZymes (CAZymes) database (CAZy db) or eggNOG that is based on COG but includes more microbial sequences (Kanehisa and Goto, 2000; Tatusov et al., 2000; Jensen et al., 2008; Cantarel et al., 2009). Nevertheless, these databases lack regular updates and hence exhibit a deficient coverage especially concerning the increasing numbers of sequenced genomes. Moreover, the MetaProteomeAnalyzer, a metaproteomic pipeline covering the all steps from PSM to functional analyses including several search algorithms was released recently (Muth et al., 2015). This program was tested intensively to analyze the raw data and search results of the current study but all attempts failed due to numerous Java programming errors that occurred within the first editions of the MetaProteomeAnalyzer. For that reason, the COG, KEGG and CAZy databases were used for the functional assignments and classification of proteins using the software WebMGA (Wu et al., 2011) and dbCAN (Yin et al., 2012). Detailed information about the mass spectrometric measurements, protein identification and quantification procedures and the functional analyses are described in Chapter 4.3.6 and Chapter 4.3.7 which also includes more precise information about the sample preparation, protein extraction and trypsin digestion of the main experiment.

5.2.3 Nuclear Magnetic Resonance

Organisms inhabit different ecological niches being exposed to various selective pressures and accordingly differ as well in their metabolic requirements. Metabolites represent the biochemical propellant at cell level and hence are an elementary part of all physiological processes (Aldridge and Rhee, 2014). Metabolomic analyses, the identification and quantification of small molecules (< 1,500 Da) linking genotype and phenotype (Fiehn, 2002) employ either NMR- or MS-based approaches that represent the leading technologies for characterizing the metabolomes of complex mixtures in biological fluids. The MS-based approach is usually coupled with gas chromatography (GC-MS) or liquid chromatography

(LC-MS). The advantages of NMR-based analyses lie within quantification and reproducibility covering as well a wide dynamic range analyzing samples in a non-destructive way whereas MS-based measurements are more sensible and can detect higher numbers of different compounds (Aldridge and Rhee, 2014). Moreover, two-dimensional NMR allows the determination of yet unknown structures and does not require prior knowledge about the metabolites. Ionization or derivatization of the investigated compounds is not required (Markley et al., 2017). Comprehensive studies available employed a combination of NMR and GC-MS/MS to investigate the metabolomes of rumen fluid including as well dietary effects (Ametaj et al., 2010; Saleem et al., 2012; Saleem et al., 2013).

For NMR analyses of complex mixtures of metabolites, the most significant nuclei are ^{1}H and ^{13}C with ^{1}H being more sensitive appearing with a natural presence of about 100%. The most common variant of NMR-based metabolomics is the one-dimensional ^{1}H NMR. Many ^{1}H signals overlap preventing proper identification of specific compounds and cover only a small range of 10 ppm whereas ^{13}C signals appear in a range of 200 ppm thus, are better resolved but ^{13}C exhibits a low natural abundance of 1.1% hence, is less sensitive. Improved, more definite identification of metabolites in complex mixtures can be achieved by two-dimensional, homonuclear and heteronuclear NMR (Markley et al., 2017). In this, compound identification can be maximized while minimizing false-positives in matching the ^{1}H and ^{13}C signals (Cui et al., 2008).

5.2.3.1 Sample Preparation for NMR Analysis

The RF samples were filtered and sterilized for subsequent dehydration by vacuum centrifugation. The presence of solid particles disturbs the homogeneity of the magnetic field due to the different magnetic susceptibility of particles when compared to the solutions. The samples were reconstituted in a deuterium-based sodium phosphate buffer to improve the spectral resolution and to minimize large signals from regular hydrogen that otherwise would swamp the sample signals. The NMR signal of deuterium is further used as field frequency lock for stabilization of the magnetic field strength to keep the resonance frequency constant. The detailed sample preparation procedure is described in Chapter 4.3.8.

Dehydration of the samples might have caused a loss of volatile compounds such as ethanol or methanol but the common ruminal metabolites were not affected as indicated by the numbers and concentrations of identified metabolites (Table 5). Moreover, preliminary test experiments (data not included) revealed no differences in regard to the presence of expected metabolites between of vacuum-dried and freeze-dried samples. Furthermore,

preliminary comparisons between the RF sample fractions and the LP sample fractions did not show any significant differences concerning the obtained spectra and the presence of metabolites (data not included).

A previous study investigating the metabolome of rumen fluid used lyophilized sample material analyzed by one-dimensional ^{1}H NMR exclusively and identified more amino acids but lacks verification from two-dimensional NMR experiments (Lee et al., 2012). Another study based on one-dimensional ^{1}H NMR spectroscopy added comparably small volumes of deuterium buffers to the original rumen fluid samples identifying up to 51 different compounds but generally lacks proper evidence regarding the presence of the respective metabolites (Zhao et al., 2014).

5.2.3.2 Metabolite Identification and Quantification

Well-resolved ^{1}H spectra with low signal-to-noise ratios allow identification and quantification based on peak positions and intensities. This may involve the manual analysis of a spectroscopist or the utilization of software like the Chenomx NMR Suite (Chenomx, Edmonton, Canada) that includes a database containing standard spectra of many metabolites at various magnetic field strengths and pH values (Mercier et al., 2011). For manual identification of compounds, there are public reference databases containing standard spectra of common compounds (Ulrich et al., 2008; Wishart et al., 2009).

However, for an absolute quantification a standard with specific concentration has to be added. Concerning the present study, the RF samples were spiked with 5 mM 3-(trimethylsilyl)propionic-2,2,3,3-d4 acid sodium salt (TSP) as an internal chemical shift reference and quantification standard.

Metabolites such as methylamine, dimethylamine and trimethylamine exhibit different singlet signals at specific chemical shifts. Nevertheless, the respective ^{1}H spectra do not contain much information thus, additional confirmation by two-dimensional ^{1}H and ^{13}C experiments is required to verify the expected chemical shifts of the methyl carbons (Dona et al., 2016). Moreover, low abundant compounds or compounds with overlapping signals need to be analyzed by two-dimensional NMR spectroscopic methods.

The metabolomic experiment (Chapter 4) verified 12 out of 20 identified compounds by two-dimensional homonuclear and heteronuclear NMR with DQFCOSY, gHSQCAD, gHSQCTOCSY, as well as gHMBCAD pulse sequences. The quantification of verified identifications was carried out using the Chenomx NMR Suite 8.2 software, database available at pH 6.5 as reported by other studies (Weljie et al., 2006; Wishart, 2008; Ametaj et

al., 2010; Saleem et al., 2013). A detailed description of the identification and quantification procedures can be found in Chapter 4.3.8.

However, most studies of the rumen metabolome identified higher numbers of different metabolites by a combination of NMR- and MS-based approaches. Stand-alone NMR analyses are challenged by the two-dimensional ^{1}H and ^{13}C experiments for proper verification of the presence of distinct metabolites and generally lack sensitivity for the detection of low abundant compounds (Aldridge and Rhee, 2014). Concerning the present study, the obtained spectra exhibited high signal-to-noise ratios limiting the reliable identification and verification of further low abundant metabolites. Nevertheless, the most important short-chain fatty acids of the rumen were quantified confidently and provided valuable additional information on the activity of certain microbial groups.

Moreover, a parallel study determined the amount of acetate, butyrate, propionate and valerate as well as the corresponding isoforms in the identical sample material by gas chromatography-mass spectrometry (GC-MS) which revealed strongly consistent abundance patterns (Gero Seyfang, oral communication).

5.3 The Rumen Core Microbiome

During the first weeks of a calf's life, the rumen is not functional and suckled milk does not pass through the rumen since the esophageal groove is closed by reflex. At that stage, the rumen still is comparably small and important components like the rumen wall villi, responsible for the absorption of nutrients have not formed yet (Van Soest, 1994). The physiological and structural development towards a completely functional rumen with age is associated with the establishment of microorganisms (Malmuthuge et al., 2015) and fermentation products, respectively. Microbes start colonizing the rumen gastrointestinal tract one to five days past parturition (Fonty et al., 1983; Minato et al., 1992; Skillman et al., 2004; Jami et al., 2013) and once fully established and developed, form relatively stable microbial communities. Thus, polysaccharide-degrading and cellulolytic capacities are present in the rumen from the very beginning since the respective microbial species appeared in the rumen during the first days of the calves' life and grew abundant after two to three weeks (Fonty et al., 1983; Minato et al., 1992). Yet, each individual microbiome exhibits unique and variable elements (Ley et al., 2006). The abundance of specific members of the rumen microbiome fluctuates regularly during the day (Dehority and Orpin, 1997; Lengowski et al., 2016) and parts of the community structure can vary between individual animals (Welkie et al., 2010; Jami and Mizrahi, 2012), different host species (Paz et al., 2016) and geographic areas

(Henderson et al., 2015). In particular, dietary influences shape the rumen microbiome but beyond all impact factors, a relatively constant core microbiome is maintained (Henderson et al., 2015). The core microbiome comprises species that are considered to possess a crucial functional potential or are at least main contributors to ruminal fermentation since present under all different conditions (Weimer, 2015). Further, the ruminal communities are resilient, reestablishing the original bacterial community structure after external perturbations (Li et al., 2012; Weimer, 2015). A meta-study of all publicly available rumen 16S rRNA gene sequences deposited in the RDP database found in total 19 bacterial phyla (Kim et al., 2011b). However, it was proposed, that the core communities of the rumen belong to eight bacterial phyla (Sitao et al., 2012) and generally comprise the two most dominant phyla of Bacteroidetes and Firmicutes that make up the vast majority of 16S rRNA gene sequences found within the gastrointestinal tract of ruminants, but as well of monogastric animals (Flint et al., 2008).

The core microbiome of the main study (Chapter 4) was assessed by metaproteomics and 16S rRNA gene sequencing and included species that were present in all animals, diets and ecological niches. Overall 27 samples, 8,163 bacterial and 358 archaeal proteins were quantified and 1,484 bacterial and 626 archaeal operational taxonomic units (OTUs) could be assigned. In contrast, the core metaproteome consisted of only 822 bacterial and 41 archaeal proteins, which appeared in high abundance constituting about 78% and 68% of the total label-free quantification (LFQ) values of all proteins. The core communities based on 16S rRNA gene comprised 208 bacterial and only three archaeal OTUs making up 49% and 7% of the total OTU abundance. Despite the huge differences in numbers and abundances of proteins and OTUs, the taxonomic distribution of the core bacterial communities at phyla level was relatively similar to the structure of the overall microbiome regarding both methods respectively (Figure 20).

However, this changed when comparing the bacterial core community composition with the overall structure at family level, which revealed some more drastic changes in the protein- and DNA-based analyses (Figure 21). Proteins of the family of Prevotellaceae revealed an increased contribution of about 7% in the core metaproteome and proteins assigned to the phylum of Firmicutes increased fourfold while the LFQ abundances of most other phyla and families including Acidaminococcaceae, Lachnospiraceae, Ruminococcaceae and Selenomonadaceae remained relatively constant. In contrast, proteins of the Firmicutes orders of Bacillales and Clostridiales decreased drastically in abundance in the core metaproteome (Figures 21A,C). The generally low abundant Erysipelotrichaceae and

Veillonellaceae proteins as well as the phyla of Elusimicrobia and Tenericutes decreased in abundance or were even absent in the core metaproteome. These findings underpin the essential functional potential of the Prevotellaceae family and certain families of the Firmicutes phylum in particular of yet unclassified Firmicutes species dominating the rumen microbiome (Flint et al., 2008).

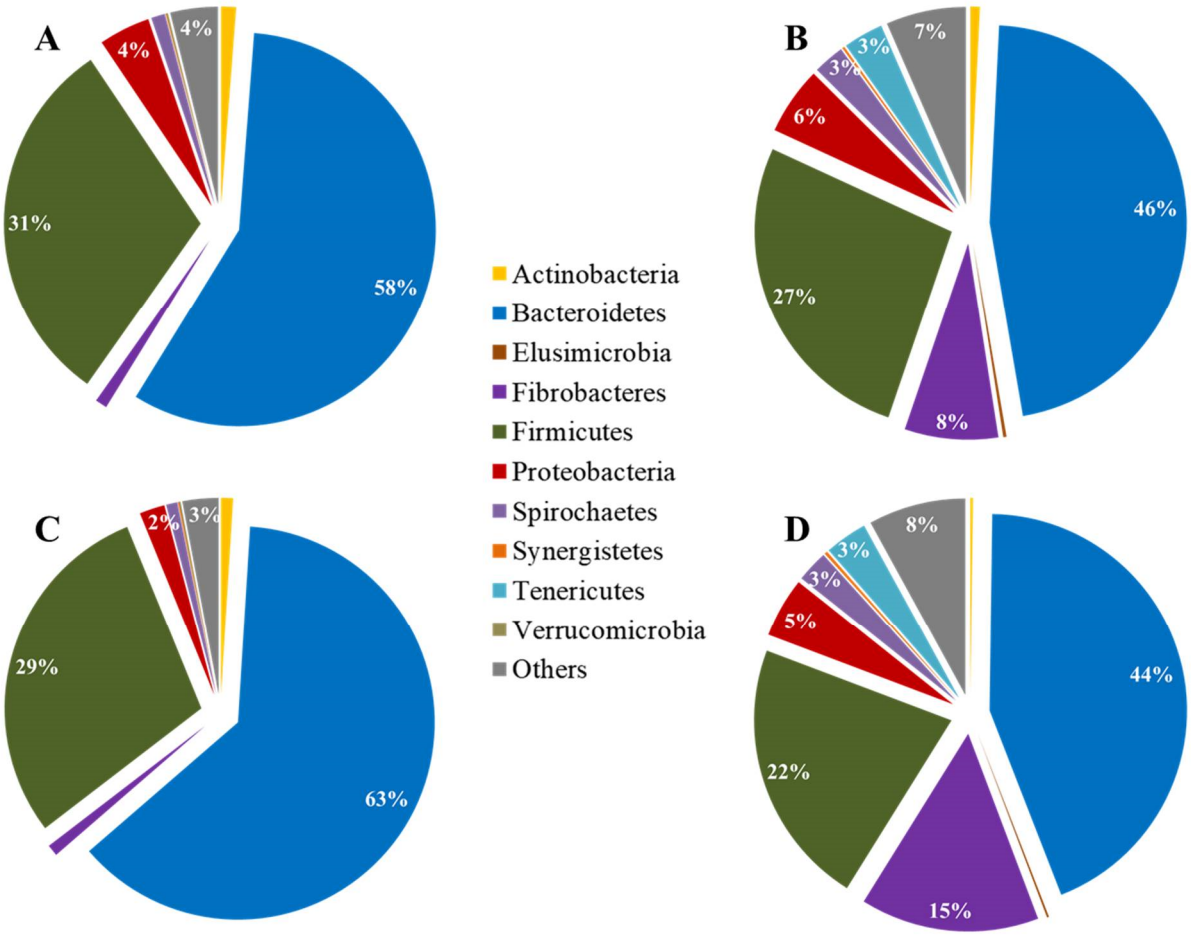

Figure 20| Pie charts of the overall (n = 27) bacterial community structure at phyla level obtained from 8,163 proteins (**A**) and 1,484 OTUs (**B**) while the core microbiome consisted of 822 bacterial proteins (**C**) and 208 OTUs (**D**). The core microbiome represented 78% of the total protein abundance and 49% of the total OTU abundance.

Concerning the 16S rRNA gene-based dataset (Figures 21B,D), an almost doubled increase in OTU abundance of the Fibrobacteres phylum was observed in the core community structure emphasizing the crucial fiber-degrading capabilities of these species. Similar to the observation in the core metaproteome, the abundance of Prevotellaceae OTUs increased about 5% in the 16S rRNA gene-based core community. The abundance of OTUs assigned to the

Bacteroidetes phylum decreased about 50%. Comparable to the shifts in the protein-based dataset, the family of Erysipelotrichaceae, missing in the core metaproteome, exhibited a remarkable decrease of above 97% of abundance in the OTU-based core community (Figures 21B,D). Contrarily to the core metaproteome, there was no relevant change of OTU abundances of the Veillonellaceae family, Elusimicrobia and Tenericutes phyla in the 16S rRNA gene-based dataset. Further, an increased OTU abundance of the families of Acidaminococcaceae and Succinivibrionaceae in the DNA-based core community structure was observed while OTUs of important fiber-degrading families like Lachnospiraceae and Ruminococcaceae decreased slightly.

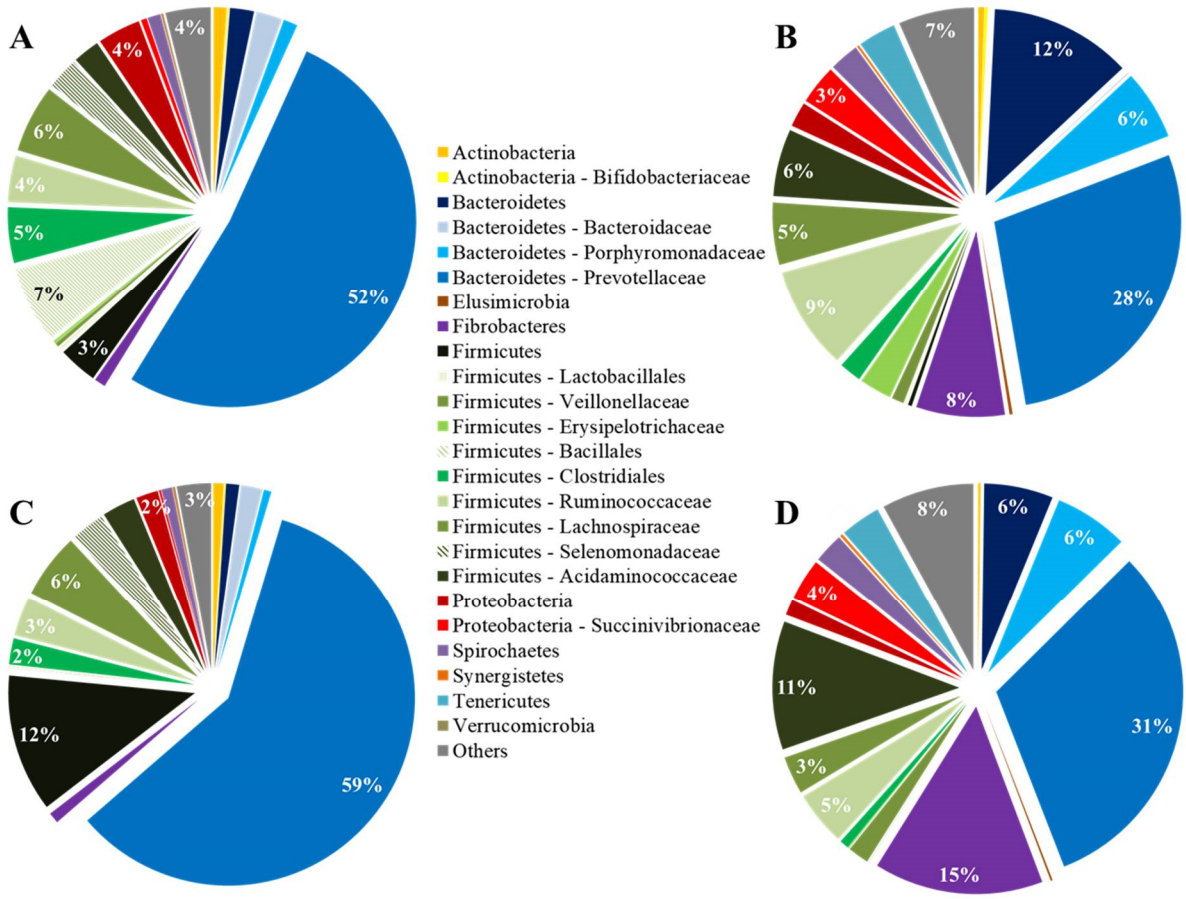

Figure 21| Pie charts of the overall (n = 27) bacterial community structure at phyla and family level based on 8,163 proteins (**A**) and 1,484 OTUs (**B**) while the core microbiome comprised 822 bacterial proteins (**C**) and 208 OTUs (**D**).

DNA- and protein-derived core communities both were dominated by the phyla of Bacteroidetes and Firmicutes and further included the less abundant phyla of Actinobacteria, Fibrobacteres, Proteobacteria, Spirochaetes and Synergistetes (Figure 20). Yet, the bacterial core metaproteome of the present study did not contain any Elusimicrobia and Tenericutes-derived proteins while OTUs of the respective phyla appeared in the DNA-based core communities. On the other hand, none of the OTUs of the core communities were assigned to the phylum of Verrucomicrobia whereas the core metaproteome contained Verrucomicrobia-derived proteins.

However, disregarding the shared or overall bacterial taxa composition, differences between the protein- and the DNA-based abundances of specific bacterial families and phyla accentuate the importance of combining different methods to investigate the rumen microbial communities. The metaproteomic approach rather targets the metabolically active consortia while the 16S rRNA gene sequencing includes all cells present disregarding their functional contribution (Hettich et al., 2013; Wilmes et al., 2015). Huge variations between LFQ and OTU abundances of the Fibrobacteres phylum can be explained by the fact that metaproteomic investigations depend on the amount and quality of reference sequences available for database searches (Tanca et al., 2013; Muth et al., 2016). Currently only about 15 genomes and thus, comparably small numbers of annotated protein sequences are available for the phylum of Fibrobacteres which limits mass spectrometry-based identifications. In contrast, 1,863 different 16S rRNA gene sequences of the phylum of Fibrobacteres are deposited in the RDP database. On the other hand, the metaproteomic approach identified proteins belonging to the Firmicutes order of Bacillales and the family of Selenomonadaceae, common in the rumen, which were not identified by 16S rRNA gene sequencing.

Generally, the average overall bacterial phyla distributions obtained from the DNA- and protein-based approaches of this study are similar to findings of other studies employing 16S rRNA gene sequencing and comparable diet compositions respectively higher amounts of concentrate as shown in Figure 22.

Variations in the abundance of certain phyla and the ratio of Bacteroidetes to Firmicutes species between the studies considered in Figure 22 might be attributed to different DNA extraction methods (Henderson et al., 2013), chosen primer sets (Cai et al., 2013) or amplified 16S rRNA gene sequence regions (Schloss, 2010). Other potentially influencing factors are the different databases used for taxonomic assignments (Liu et al., 2008) or deviations in diet composition. Further, several studies reported that the phylum of Fibrobacteres, that includes the major cellulolytic bacterium *Fibrobacter succinogenes*,

fluctuates remarkably across cows and diets (Jami and Mizrahi, 2012) and was not even identified in the fiber-adherent fraction of a previous metagenomic study (Brulc et al., 2009).

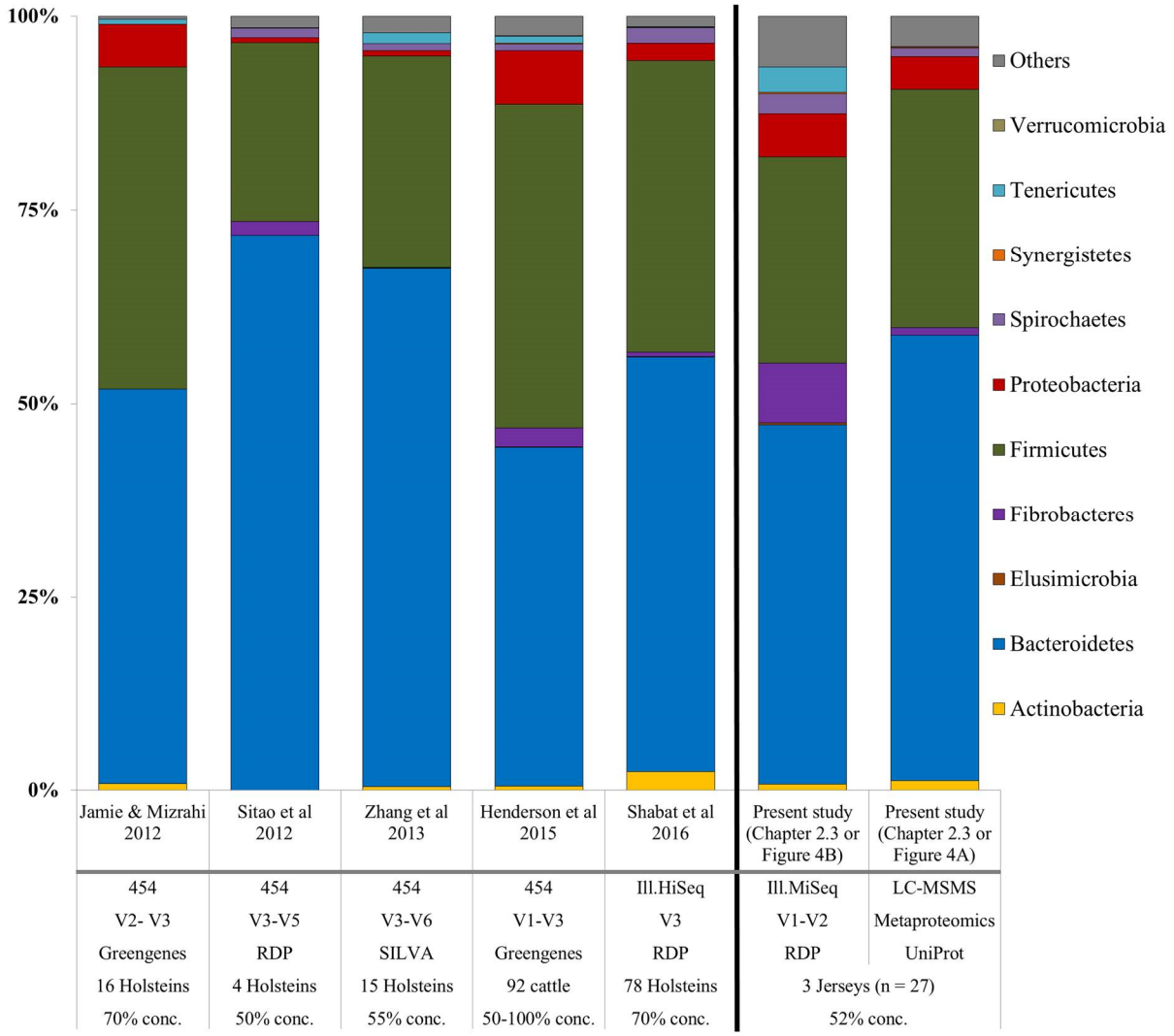

Figure 22| Comparison of averagely identified bacterial phyla in 16S rRNA gene-based studies of the microbiome of *B. taurus* species fed with high concentrate diets (Jami and Mizrahi, 2012; Sitao et al., 2012; Zhang et al., 2013; Henderson et al., 2015; Shabat et al., 2016). The sequencing platforms, regions of the 16S rRNA gene, databases for taxonomic assignments, numbers of cattle and the proportions of concentrate in diets are shown.

The overall and core archaeal community structures obtained from the mass spectrometry-based approach are shown in Figure 23. There were no considerable shifts in taxonomic distribution of the archaeal core metaproteome at phyla, order or family level. The archaeal core metaproteome constituted 68% of the total archaeal protein abundance. The 16S rRNA gene-based archaeal community structures were not included due to the low number

and abundance of only three archaeal OTUs in the core community that exclusively belonged to the order of Thermoplasmata, respectively to the family of Methanomassiliicoccaceae. Further, all 626 OTUs of the overall archaeal community structure were exclusively assigned to three families of methanogens, Methanobacteriaceae, Methanosarcinaceae and Methanomassiliicoccaceae as described in Chapter 4.4.2. Nevertheless, these findings are in concordance with reports of other studies employing 16S rRNA gene sequencing (Janssen and Kirs, 2008; Seedorf et al., 2015).

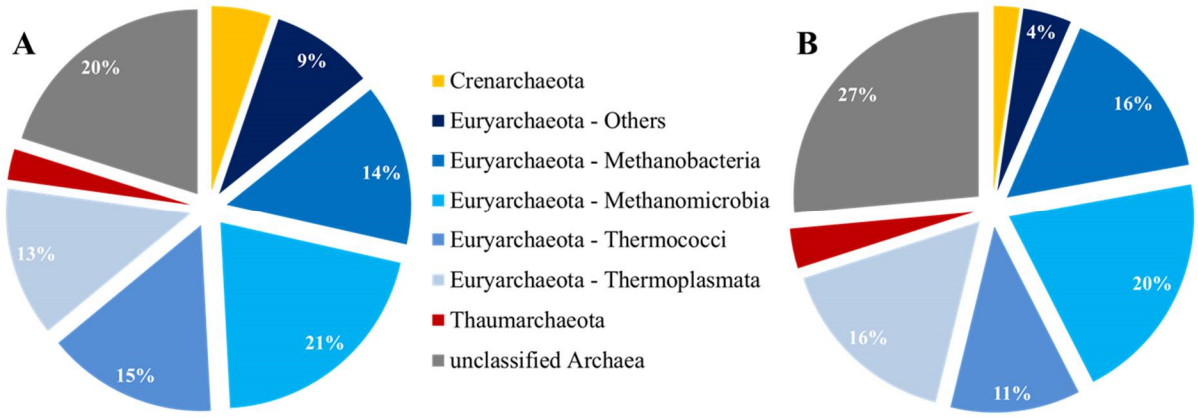

Figure 23| Pie charts of the overall (n = 27) archaeal community structure at phyla and order level obtained from 358 proteins (**A**) while the archaeal core metaproteome consisted of 41 archaeal proteins (**B**) representing 68% of the total protein abundance.

5.4 Potential Influence of Individual Host Animals

Besides the shared rumen core microbiome, diurnal fluctuations and variations between diverse rumen fractions as well as age, breed, dietary and other environmental impact factors, each individual animal is considered to maintain its own characteristic microbiome composition (Ley et al., 2006; Ley et al., 2008; Weimer et al., 2010) which responds differently to shifts in diet composition (Zhou et al., 2012). Variations between the animals concern the proportions of specific groups of the core communities and further include the presence and abundance of bacterial taxa that do not occur in other animals. Nevertheless, it was shown that the groups of bacteria differing between animals are often phylogenetically related exhibiting similar genetic features respectively a redundancy in functional potentials driven by the particular ecological niches (Jami and Mizrahi, 2012; Weimer, 2015). Host specificity is not limited to bacteria only but was also observed for archaea and protozoa (Zhou et al., 2012).

However, the animals' rumen constitutes the framing environment for the microbial ecosystem and hence variations in host genomes are likely to influence the structure and activity of the inherent microbiome (Benson et al., 2010; Weimer, 2015). A study based on 146 Holstein-Friesian dairy cows fed with identical rations found large variations between individual animals regarding their feed conversion efficiency that correlated with several characteristics of the rumen microbiome composition (Shabat et al., 2016). Additionally, a study under identical dietary regimen identified varying ratios of Firmicutes to Bacteroidetes species among 15 Holstein-Friesian dairy cows that highly correlated with the daily milk fat yield (Jami et al., 2014). Identical dietary treatments did not compulsorily cause identical microbiome structures (Hernandez-Sanabria et al., 2010; Welkie et al., 2010). Nonetheless, these investigations lack information about the potential impact of the host genome on the rumen microbiome. In general, evidences proofing a regulation of the rumen microbiome by the host genome are quite uncertain and rare (Roehe et al., 2016).

Previous investigations stated an influence of host genomics on feed conversion efficiency and methane emissions regardless the changes in function and structure of the rumen microbiome (Robinson et al., 2014; Pickering et al., 2015) whereas a recent study declared that the abundance of archaea and associated methane emission is determined by the individual host genome (Roehe et al., 2016).

Regarding the present study, no data concerning host genome and related effects on the rumen microbiome were collected, but differences in physiological parameters between the three cows (23, 28 and 59) such as the average daily milk yield and consumption of total mixed rations throughout the period of the experiment were observed (Table 6).

The proportions of average daily milk yield to consumption of total mixed rations or body weight during the period of the experiment in relation to the mean of ratios of the three cows indicated a below average feed conversion performance of cow no. 23 (Figure 24). Additionally, cow no. 23 was at an earlier stage (146 days) of the lactation period when compared to the remaining animals (222 and 234 days). Similar patterns were found in the average abundance ratio of Bacteroidetes to Firmicutes species based on the nine samples per cow when compared to the average of abundance ratios of the three cows as determined by both, the metaproteomic and the 16S rRNA gene-based approach (Figure 24).

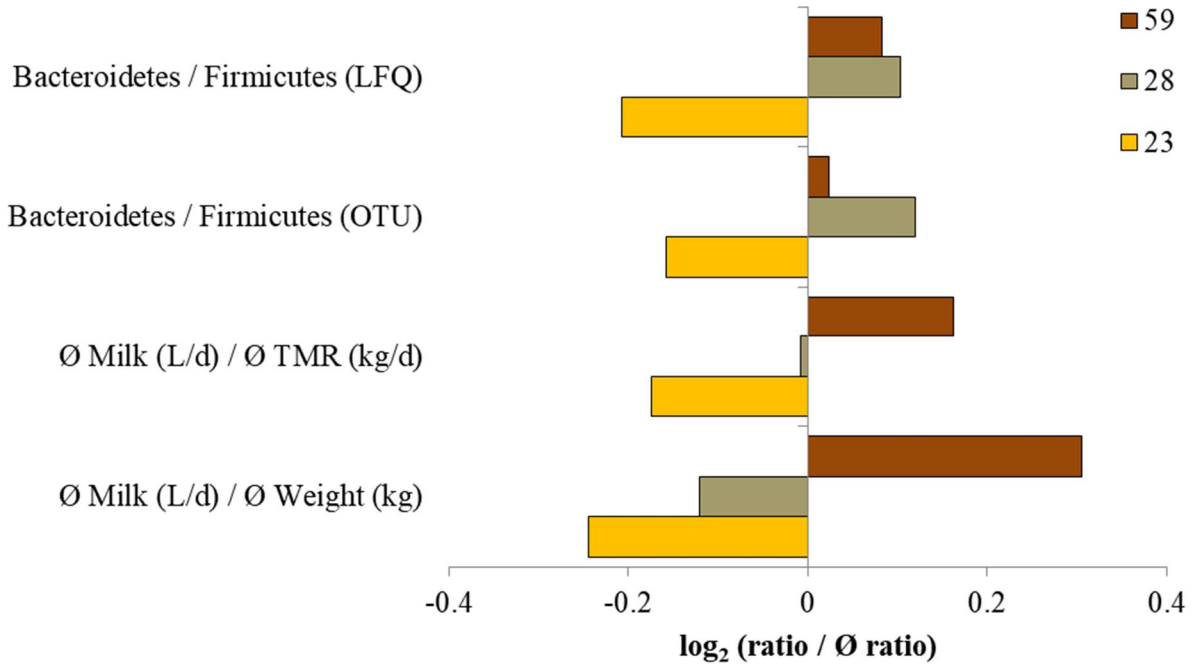

Figure 24| The average abundance ratio of Firmicutes to Bacteroidetes of all samples per cow (n = 9) in relation to the mean of the three respective ratios (0) for the mass spectrometry-based LFQ and 16S rRNA gene-based OTU abundance values. Further, the ratio of daily milk yield (L/d) to uptake of total mixed ration (kg/d) and body weight (kg) during the experimental trial (n = 60) in relation to the mean of the respective ratios (0) are shown for each cow.

The below average feed conversion efficiency as well as the lower average abundance ratio of Bacteroidetes to Firmicutes indicated a deviation of cow no. 23 when compared to the cows no. 28 and 59. Moreover, in the course of the metaproteomic analyses, variations in the abundance of 71 quantified proteins assigned to the genus *Bos* revealed a statistically significant separation of cow no. 23 when compared to no. 28 and 59 regardless the dietary treatments and different rumen fractions (Figure 25). The observed variations in feed conversion efficiency and abundance ratio of Bacteroidetes to Firmicutes within the cows as well as the differential abundance of proteins assigned to the genus *Bos* allowed the suspicion of further, host-related deviations concerning the overall microbiome structure.

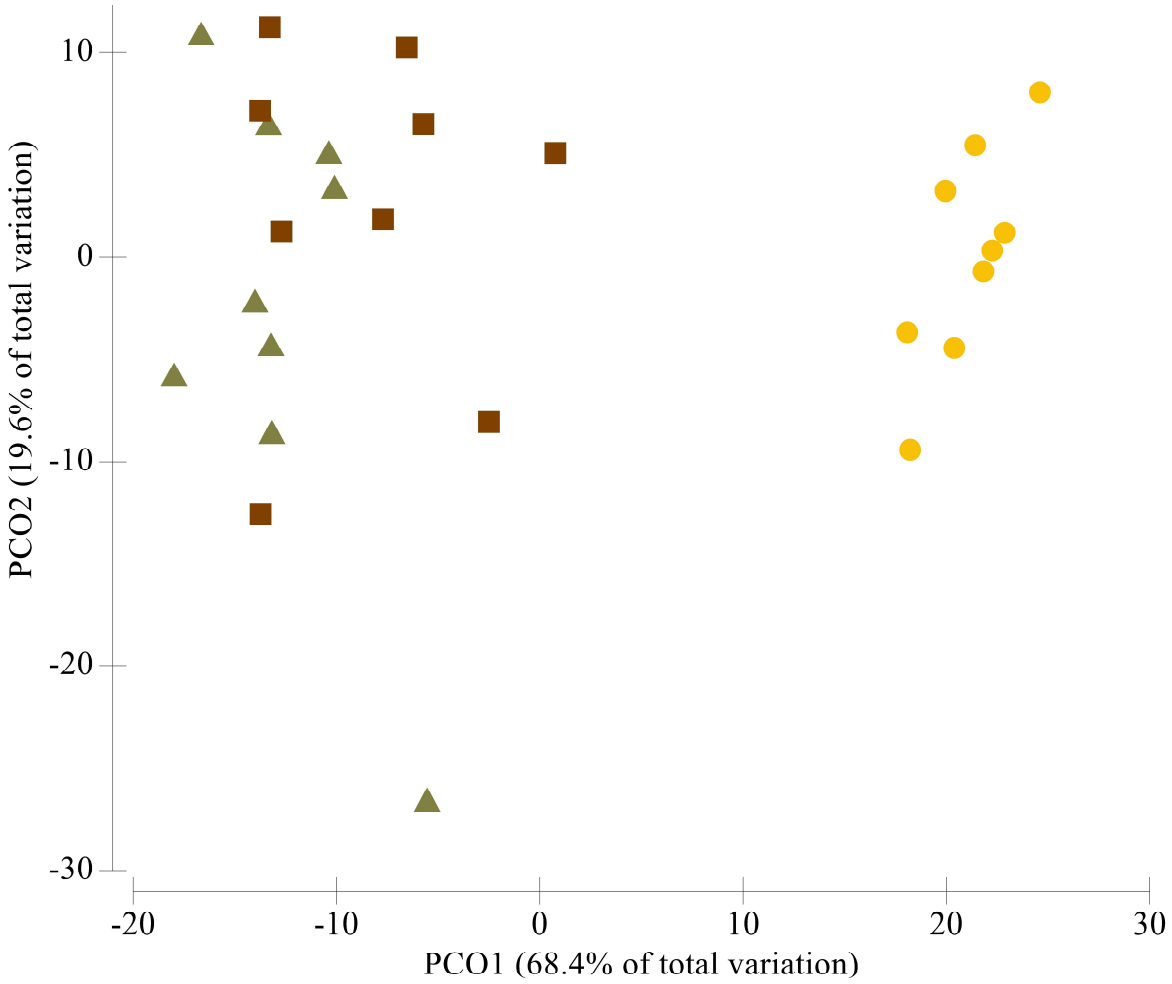

Figure 25| Principal coordinate analyses showing the abundance of 71 quantified *Bos* proteins identified in all 27 samples within the course of the metaproteomic analyses. Yellow dots, cow no. 23; grey triangles, cow no. 28 and brown squares, cow no. 59.

In order to depict the individuality and associated variations of the microbiome compositions, the overall bacterial community structure at phyla, order and family level of the entire 27 samples analyzed within this project are shown, of the metaproteomic approach in Figure 26 and of the 16S rRNA gene sequencing in Figure 27. Concerning the metaproteomic dataset (Figure 26), variations between the individual cows were rather restricted to the abundance ratio of Bacteroidetes to Firmicutes. In contrast, the DNA-based dataset (Figure 27) revealed remarkable outliers such as the two samples GS_RF_23 and H_RF_23 of cow no. 23 showing an increased abundance of OTUs belonging to the order of Lactobacillales and the phylum of Tenericutes respectively. Further variations between cows in the OTU-based microbiome structures existed in the abundance of the Succinivibrionaceae family within the CS diet-based samples in which the abundance of OTUs of the Succinivibrionaceae family was comparably lower in all samples of cow no. 23.

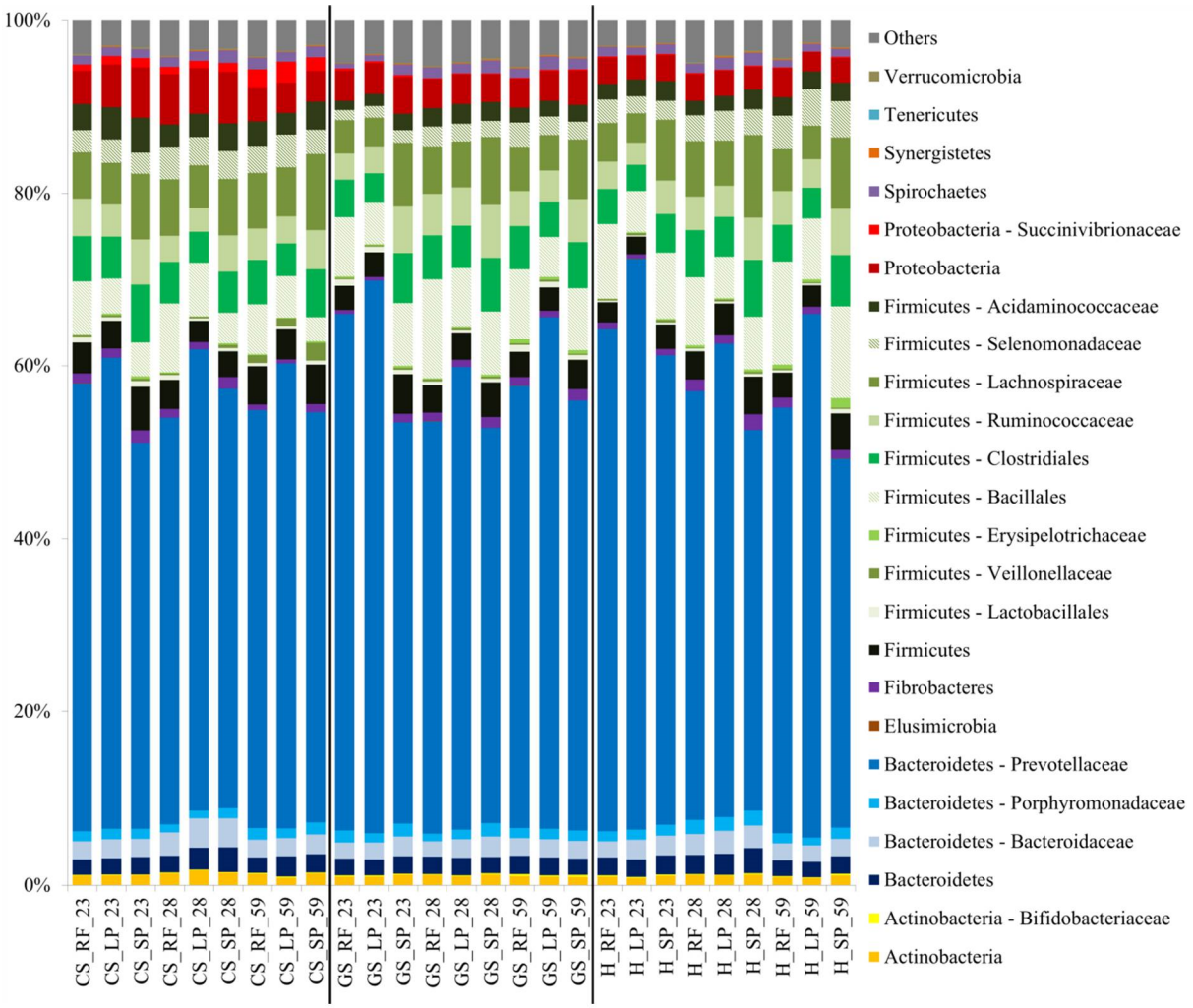

Figure 26| Bacterial phylogenetic distribution of the entire 27 samples based on 8,163 proteins. CS, corn silage-based diet; GS, grass silage-based diet; H, grass hay-based diet. RF, rumen fluid; LP, liquid phase; SP, solid phase. Cows are indicated by no. 23, 28 and 59.

The order of Lactobacillales contains bacteria that produce lactic acid such as *Streptococcus bovis* and *Lactobacillus* spp. (Asanuma and Hino, 2002) that are known to be involved in the development of subacute ruminal acidosis (Nagaraja and Titgemeyer, 2007), the most significant nutritional disease affecting dairy cows (Morgante et al., 2007; Plaizier et al., 2012). Several studies investigated the influence of subacute ruminal acidosis on the rumen microbiome (Hook et al., 2011a; Mao et al., 2013; Wetzels et al., 2016; Wetzels et al., 2017). Species of the phylum of Tenericutes, respectively the class of Mollicutes typically lack a cell wall and are commonly known as parasites of eukaryotic hosts whereas being rather commensals in the rumen ecosystem (Razin et al., 1998; Brown et al., 2015).

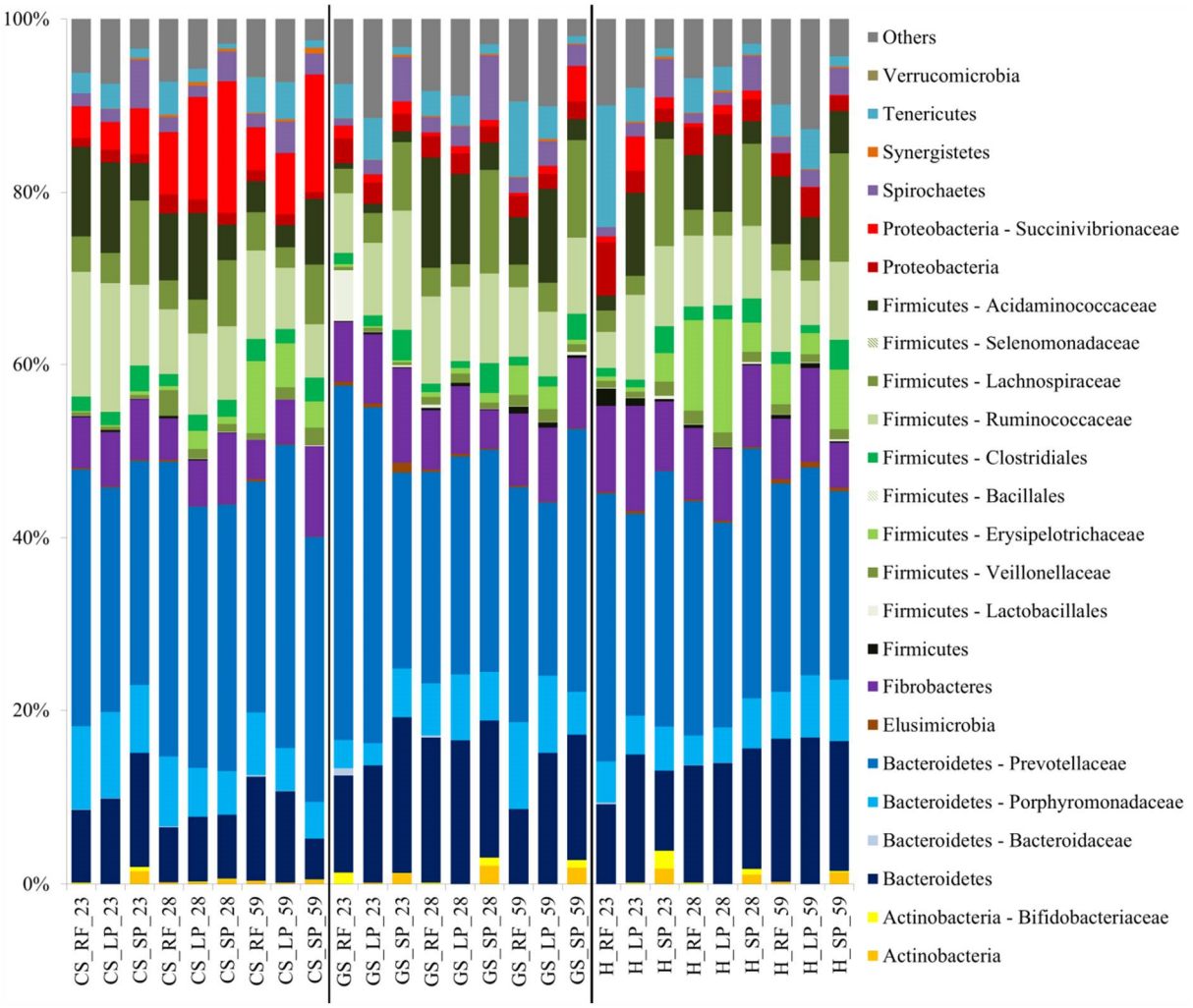

Figure 27| Bacterial phylogenetic distribution of the entire 27 samples based on 1,484 OTUs. CS, corn silage-based diet; GS, grass silage-based diet; H, grass hay-based diet. RF, rumen fluid; LP, liquid phase; SP, solid phase. Cows are indicated by no. 23, 28 and 59.

Despite the variations in feed conversion efficiency, ratio of Bacteroidetes to Firmicutes (Figure 24), abundance patterns of host proteins (Figure 25) and aberrations in bacterial community composition of certain samples (Figure 27) there were no significant statistical differences between the bacterial community structures in regard to the three cows. Although, the community compositions of the cows no. 28 and 59 were more similar to each other as when compared to no. 23 respectively (Table 4). In fact, statistical analyses of similarity (ANOSIM) revealed significant variations in the DNA- and protein-based bacterial community compositions within the diets and the rumen sample fractions but not between the individual cows (Table 4). Thus, the primary focus of the present study lies on the effects induced by the different diets and ecological niches respectively rumen sample fractions (Chapter 4) which will be discussed in detail in the following paragraph including as well the

functional information obtained from the metaproteomic approach and the results of the metabolome analysis. There were no diet-, sample fraction- or host-related statistical significances concerning the archaeal community composition accessed in all 27 samples probably due to the low amount of identified proteins and assigned OTUs, hence no statistical details are shown.

5.5 Dietary Impact and Variations in Ecological Niches

Concerning the main experiments of the present study (Chapter 4), the dietary- and rumen sample fraction-related impact was significantly greater than the effects of the individual host animals on the rumen bacterial community structure as found for both, the DNA- and the protein-based dataset (Table 4, Figure 9). In terms of dietary influence and variations between distinct ecological niches, similar results were found by several studies employing various methods to explore the rumen microbiome (Kong et al., 2010; Pitta et al., 2010; Kim et al., 2011a; Patel et al., 2014b; Henderson et al., 2015).

The dietary impact on the bacterial metaproteome was found to be greater than the influence of the distinct rumen sample fractions whereas the DNA-based bacterial community composition revealed a slightly larger, yet comparably negligible effect of the ecological niches when compared to the variations between diets (Table 4, Figure 9). Once more, these findings emphasize the importance and advantages of combining different but complementary methods to investigate the rumen microbiome.

Within diets, the LFQ abundances of 8,163 bacterial proteins quantified by the mass spectrometry-based two-step search process and the 1,484 bacterial OTU abundances obtained from Illumina MiSeq sequencing of the V1 to V2 region of 16S rRNA gene differed significantly between the CS and the GS diets as well as between the CS and the H diets. The bacterial community structures of the more fiber-rich GS and H diets were not significantly different from each other (Table 4, Figure 9).

In terms of the rumen sample fractions, the bacterial community structures and metaproteoms of the SP were significantly different from those of the RF and the LP that exhibited no significant variation as found by both methods (Table 4, Figure 9). Differences in bacterial community composition of the three rumen sample fractions are in concordance with a study analyzing the chemical composition of the respective communities showing that the firmly-attached bacteria (SP) differed from the liquid-associated population (RF) whereas loosely-attached bacteria (LP) were rather similar to the liquid-associated population (Legay-Carmier and Bauchart, 1989). Other studies reporting structural shifts in the bacterial

communities in regard to the different ecological niches within the rumen ecosystem are numerous (Larue et al., 2005; Kong et al., 2010; de Menezes et al., 2011; Kim et al., 2011a; Stiverson et al., 2011).

The LFQ-values of 358 quantified archaeal proteins and the abundance of 626 archaeal OTUs revealed no significant variation between the different diets and rumen sample fractions considering all 27 samples.

5.5.1 Shifts in Bacterial and Archaeal Community Composition

Averaged over the three animals per treatment the bacterial community compositions revealed unanimous tendencies concerning the dietary influence and rumen sample fractions for both applied methods. Figure 10 shows the average abundance of bacterial phyla, orders and families commonly identified by metaproteomics on the left and 16S rRNA gene sequencing on the right side.

The abundances of proteins and OTUs of the phylum of Proteobacteria and particularly the family of Succinivibrionaceae with its considerable representatives *Ruminobacter amylophilus* and *Succinimonas amylolytica* increased in all fractions of the CS diet (Figure 10) that contained more none-structural carbohydrates, mainly starch, when compared to the more fiber-rich GS and H diets. As shown in Table 1, starch and therefrom derived maltose are the fermentation substrates required by *R. amylophilus* that produces succinic, formic and acetic acids (Hamlin and Hungate, 1956). Moreover, *S. amylolytica* was shown to increase in abundance on starch (Bryant et al., 1958) with succinic, acetic and propionic acid being the main fermentation products. Potential availability of readily accessible sugars from the fermentation of starch might be responsible for the increased abundance of amylolytic and saccharolytic species of the phylum of Proteobacteria and the Succinivibrionaceae family (Hespell, 1992; Bryant, 2015). In consequence, more succinate was produced in the fractions of the CS diet as indicated by the NMR-based metabolomic analysis (Table 5). Beyond, this might explain the comparable rise of OTUs and proteins of the Acidaminococcaceae family in samples of the CS diet (Figure 10). All respective OTUs were assigned to the genus of the succinate-fermenting *Succiniclasticum* (van Gylswyk, 1995) and 48 of 70 identified proteins were produced by the asaccharolytic and succinate-utilizing *Phascolarctobacterium* species (Watanabe et al., 2012).

Recent studies reported about reduced methane emissions under CS diet-based dietary regimen (Beauchemin and McGinn, 2005; van Gastelen et al., 2015). A metagenomics study linked a decrease in methane emissions and abundance of methanogenic archaea to an

increased abundance of the Succinivibrionaceae family (Wallace et al., 2015) whose members utilize hydrogen to produce succinate which is rapidly converted to propionate and in this compete with the most common, hydrogenotrophic methanogenesis (Liu and Whitman, 2008; McCabe et al., 2015). Concerning the present study, the ratio of acetate to propionate, an indicator for methanogenic activity, was lowest in the CS diet (Table 5). Furthermore, the increased abundance levels of the Succinivibrionaceae family (Figure 10) and the higher amount of succinate (Table 5) in the fractions of the CS diet may underpin the findings of the above-mentioned investigations. Moreover, *Succiniclasticum* and *Phascolarctobacterium* species of the Acidaminococcaceae family use succinate to produce propionate, the most important carbon source for the ruminant's gluconeogenesis (Yost et al., 1977).

Succinivibrionaceae proteins were also found in the samples of the first metaproteomic study (Chapter 3) obtained from a non-lactating Holstein-Friesian cow fed with grass hay and 12% concentrate (Figure 28). Regardless the different diet compositions, host breeds and search databases the abundance of Succinivibrionaceae was similar to the low abundance observed in the GS and H diets of the main experiment (Figure 28). In contrast, the abundance of Proteobacteria proteins in the samples of the Holstein-Frisian cow was even higher than the abundance in the CS diet samples of the main experiment (Figure 28). These remarkable differences might be explained by the distinct databases used for protein identification (Chapter 3) and quantification (Chapter 4). While the experiment described in Chapter 4 employed a customized, sample-specific database derived from on UniprotKB, the first metaproteomic analysis (Chapter 3) was performed using all bacterial entries of the large NCBInr database that contains about 180 million proteobacterial protein sequences thus, increasing the probability of false-positive protein identifications (Chapter 5.2.2.3).

No 16S rRNA gene sequences (OTUs) were assigned to the Firmicutes order of Bacillales and the family of Selenomonadaceae with its typical members of the rumen ecosystem, *Selenomonas ruminantium* and *Anaerovibrio lipolytica*. In contrast, 219 proteins were assigned to the family of Selenomonadaceae with the bulk of 131 and 40 proteins belonging to saccharolytic *Selenomonas* and lipolytic *Anaerovibrio* species, respectively (Figure 10) again pronouncing the benefits and necessity of applying multiple, complementary methods to investigate the microbiomes of complex ecosystems like the rumen. However, the first metaproteomic analysis (Chapter 3) did not identify any Selenomonadaceae proteins either (Figure 28), probably due to the inconveniently large search databases (Chapter 5.2.2.3).

A recent study investigated the same sample material used for the main experiment (Chapter 4) by quantitative real-time PCR and stated no dietary impact on the abundance of *S. ruminantium* (Lengowski et al., 2016). Contrarily, the LFQ-values of proteins of the Selenomonadaceae family, assessed by the metaproteomic approach, revealed a lower abundance throughout the GS diet fractions when compared to the CS and the H diet samples (Figure 10). *S. ruminantium* species are characterized by their ability to use a broad range of substrates such as different monosaccharides, disaccharides, lactate, glycerol, dextrins and nucleic acids including as well the fermentation products of other bacteria (Bryant, 1956; Linehan et al., 1978; Cotta, 1990; Rasmussen, 1993). Cross-feeding between *S. ruminantium* and *Butyrivibrio fibrisolvens* was reported before (Cotta and Zeltwanger, 1995) and another study showed that the co-cultivation of *S. ruminantium* and *B. fibrisolvens* promoted growth of *S. ruminantium* (Cotta, 1992). Correspondingly, the LFQ-values of 124 *B. fibrisolvens* proteins assessed by metaproteomics exhibited the lowest average abundance in the samples of the GS diet. Thus, the lower average abundance of the Selenomonadaceae family in the GS diet samples might be linked to a similarly decreased abundance of the 124 *B. fibrisolvens* proteins.

The family of Erysipelotrichaceae of the Clostridium subphylum cluster XVII as well belonging to the phylum of Firmicutes exhibited the highest abundance of LFQs and particularly OTUs in the sample fractions of the H diet (Figure 10). Most of the respective OTUs were assigned to the genera of *Sharpea* and *Kandleria* while most proteins were produced by species belonging to the genera of *Coprobacillus*, *Catenibacterium* and *Eggerthia*. Due to phenotypic, chemotaxonomic and phylogenetic data, it was suggested that *Lactobacillus catenaformis* and *L. vitulinus* should be reclassified into the genera of *Eggerthia* and *Kandleria*, respectively (Salvetti et al., 2011). Corresponding to *Lactobacillus* species, most members of Erysipelotrichaceae family probably ferment a wide range of sugars to produce mainly lactic acid as it was reported for *Sharpea azabuensis* (Morita et al., 2008). This assumption is supported by the higher amounts of lactate identified via NMR in two out of three RF samples of the H diet (Table 5).

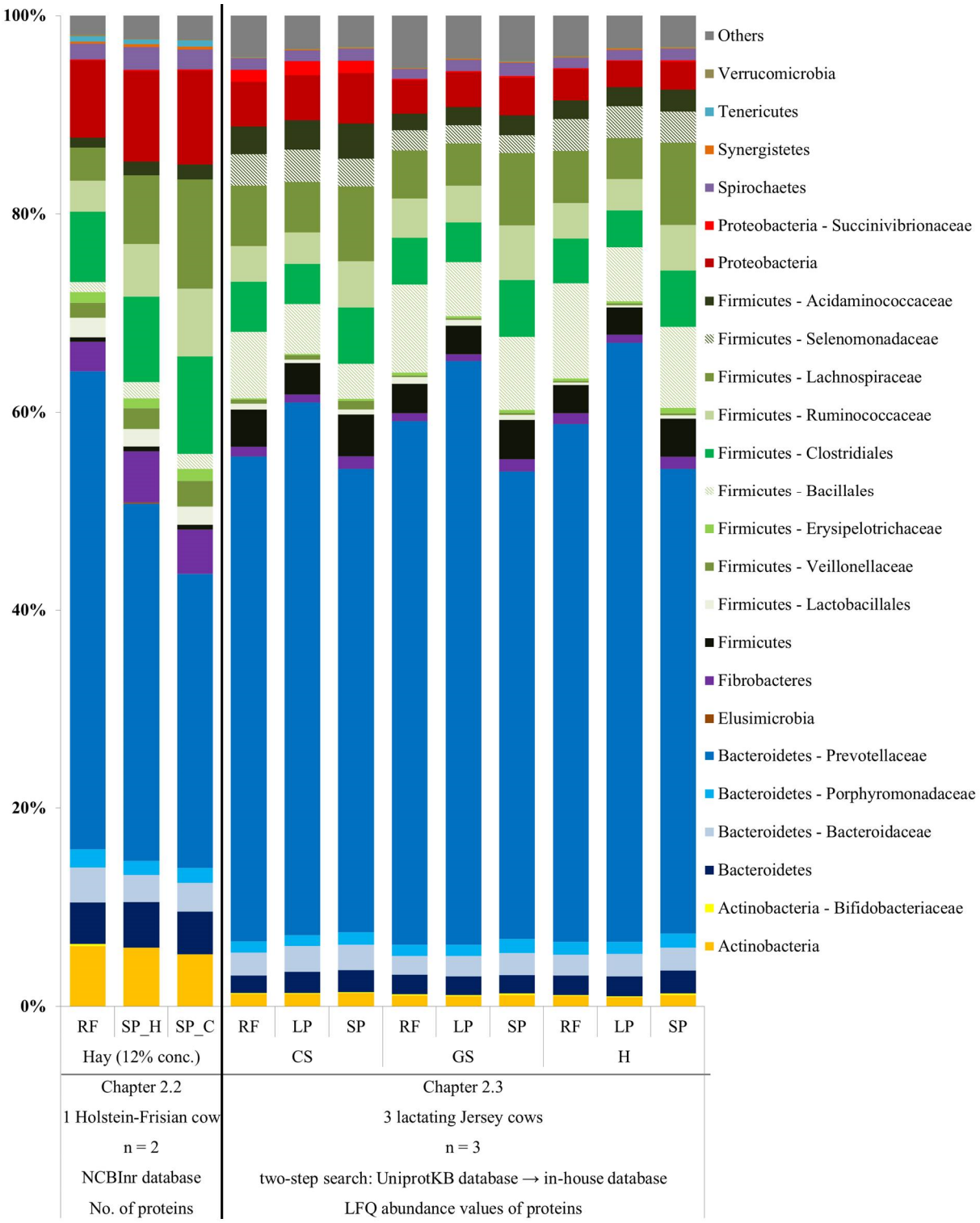

Figure 28| Comparison of the phylogenetic origin of bacterial proteins averagely identified in the first metaproteomic study (Chapter 3) and the main experiment (Chapter 4) at phyla, order and family level. The diets, breed of host animals, numbers of samples, search databases and the type of abundance data are indicated. RF, rumen fluid; SP_H, solid phase extracted with Histodenz protocol; SP_C, solid phase extracted with cheesecloth protocol. LP, liquid phase; SP, solid phase.

The abundance of the family of Bifidobacteriaceae was higher in the samples of the more fiber-rich GS and H diets when compared to the fractions of the CS diet as found by both methods (Figure 10). Genomes of *Bifidobacteria* comprise features necessary for the metabolism of plant-derived complex carbohydrates like glycoside hydrolases and sugar ABC transporters (Pokusaeva et al., 2011) as underpinned by the comparably high abundance of OTUs closely related to *Bifidobacteria* in the SP fractions of all dietary treatments. Similarly, the protein and OTU abundances assigned to the phylum of Actinobacteria were highest in all SP fractions (Figure 10). The average abundance of the phylum of Actinobacteria was much higher in the first metaproteomic analysis when compared to the main experiment (Figure 28) again indicating the necessity of precise search databases. The NCBInr database contains about 42 million more protein sequences of Actinobacteria when compared to the UniprotKB/TrEMBL database thus, increasing the probability of false-positive identifications.

The throughout low abundant Veillonellaceae family exhibited inconsistent LFQ and OTU abundances across the different sample fractions and diets. In contrast, a 16S rRNA gene-based study reported a significant dietary impact on the Veillonellaceae family (Zhang et al., 2013). Of 53 proteins assigned to the Veillonellaceae family 36 belonged to ureolytic and lactate-fermenting *Megasphaera* species. Lactate is among the most preferred substrates of *Megasphaera elsdenii* that is able to ferment about 97% of the rumen lactate (Counotte et al., 1981) relieving the risk of rumen acidosis (Kung and Hession, 1995) thus, being a highly beneficial member of the rumen ecosystem (Aikman et al., 2011). Thirteen out of 19 OTUs of the Veillonellaceae family could not be assigned to a specific genus, indicating greater proportions of yet unclassified species within the family of Veillonellaceae as observed by other studies (Zhang et al., 2013; Henderson et al., 2015).

Next to the influence of the dietary treatments on the bacterial community structure, expressive variations in abundance of certain families regarding the rumen sample fractions respectively ecological niches were observed highlighting the importance of sample fractionation in rumen studies to access the effects of treatments throughout the whole ecosystem and its specific functional niches.

Apart the dietary effects, the more fiber-rich SP fractions clearly revealed an increase of abundance of cellulolytic, hemicellulolytic and pectinolytic species of the Firmicutes order of Clostridiales including the families of Lachnospiraceae and Ruminococcaceae as supported by OTU and protein abundance levels (Figure 10). The order of Clostridiales includes several cellulolytic *Clostridium* species (Vos et al., 2011) and as well fiber-degrading members of the Eubacteriaceae family (Prins et al., 1972; Van Gylswyk and Van der Toorn, 1985). Most of

the OTUs within the order of Clostridiales, excepting Lachnospiraceae and Ruminococcaceae, were assigned to unclassified Clostridiales but no OTUs were assigned to the family of Eubacteriaceae while most respective proteins were produced by *Clostridium* and *Eubacterium* species. This indicates again the advantage of complementing DNA- and protein-based approaches.

The Lachnospiraceae family comprises the prominent cellulolytic *Butyrivibrio fibrisolvens* (Bryant and Small, 1956) and the pectinolytic *Lachnospira multiparus* (Duskova and Marounek, 2001) and exhibits strong hydrolyzing activities with multiple sets of carbohydrate-active enzymes (Stackebrandt, 2014) explaining the increased abundance in the SP fractions of all diets (Figure 10). Another study confirmed the increased abundance of Lachnospiraceae species in the fiber-adherent fractions (Larue et al., 2005).

The same trend was observed for Ruminococcaceae family that includes major cellulose degraders (Latham and Wolin, 1977; Suen et al., 2011). Previous studies reported a higher abundance in the fiber-rich fractions as well. The vast majority of the respective proteins and OTUs were assigned to *Ruminococcus* species. The NMR-based metabolome analysis identified phenylpropionate (Table 5), a compound essential for the adherence to and degradation of cellulose by *R. albus* (Stack and Hungate, 1984).

The increased abundance of the Clostridiales order including the families of Lachnospiraceae and Ruminococcaceae in the SP fraction was as well observed in the first metaproteomic investigation (Figure 28) and was confirmed by other studies (Larue et al., 2005; de Menezes et al., 2011; Kim et al., 2011a).

Another considerable cellulose-degrading specialist within the rumen ecosystem is the Gram-negative *Fibrobacter succinogenes* of the Fibrobacteres phylum (Hungate, 1950; Shinkai et al., 2009). However, there was no obvious dietary influence on the abundance of Fibrobacteres-derived proteins but the respective LFQ-values reached the maximum in the SP fractions of all diets (Figure 10) similar to the findings of a previous study (de Menezes et al., 2011). The average abundance per diet of OTUs assigned to Fibrobacteres increased with the fiber content from the CS to the GS diet and was highest in the samples of the H diet. There are reports about a decreased abundance of *Fibrobacter* upon inclusion of concentrate in diets (Henderson et al., 2015). Regardless the different databases used for protein identification that both contained comparably small numbers of Fibrobacteres protein sequences, a higher abundance of the Fibrobacteres phylum was observed in the samples of the first metaproteomic analysis containing less concentrate (Figure 28). Large differences in abundance between the metaproteomic and the 16S rRNA gene-based approaches were

discussed in Chapter 5.3 including reports about noteworthy inconsistencies in abundance of the Fibrobacteres phylum among cows and diets (Jami and Mizrahi, 2012).

The phylum of Spirochaetes appeared in higher abundance in the SP fractions of all diets (Figure 10) as determined by both methods confirming the results of another study (de Menezes et al., 2011). Most proteins and OTUs of the respective phylum belonged to *Treponema* species. *Treponema bryantii* was reported to interact with cellulolytic bacteria (Stanton and Canale-Parola, 1980) and *T. succinifaciens* exhibits carbohydrate-fermenting capabilities (Cwyk and Canale-Parola, 1979) which may verify the increased abundance throughout the SP fractions. The same trend was observed in the samples of the first metaproteomic analysis revealing a higher abundance in the SP samples (Figure 28).

Comparably, the phylum of Synergistetes, characterized in 2009, revealed the highest abundance of proteins and particularly OTUs in the SP fractions (Figure 10). There is not much information about the members of the phylum of Synergistetes which are present in many anaerobic ecosystems including the gastrointestinal tract of animals but usually appear in low abundance within the respective environments (Jumas-Bilak and Marchandin, 2014). *Synergistes jonesii* was first isolated from the rumen of goat, did not ferment carbohydrates and is thought to be involved in the degradation of plant-derived toxins beneficial for the host animal (Allison et al., 1992).

OTUs assigned to the phylum of Tenericutes were highly abundant in the RF and LP fractions of all diets whereas the LFQ-values of the respective proteins were highest in the SP fractions of all diets (Figure 10). Generally, the abundance of Tenericutes OTUs was much higher when compared to the abundances of Tenericutes-derived proteins. As mentioned in Chapter 5.4, species of the phylum of Tenericutes respectively the class of Mollicutes, typically do not possess cell walls (Razin et al., 1998; Brown et al., 2015) which might increase the susceptibility for distortions potentially introduced during protein or DNA extractions.

The Prevotellaceae family comprises common rumen bacteria such as *Prevotella ruminicola*, *P. brevis*, *P. bryantii* and *P. albensis* and constitutes the most dominant bacterial family within the rumen. The abundance of proteins and OTUs assigned to the Prevotellaceae family was not affected by the different diets but within diets, the LFQ-values were highest in all LP fractions (Figure 10). Members of the Prevotellaceae family are characterized by their versatile metabolic capabilities and their ability to utilize a broad range of substrates including peptides, proteins, monosaccharides and plant polysaccharides (Miyazaki et al., 1997; Matsui et al., 2000; Purushe et al., 2010) and thus, may not be greatly affected by changes in diet

composition.

The LFQ and OTU abundances of the Bacteroidetes families of Bacteroidaceae and Porphyromonadaceae were not affected by diets or sample fractions but revealed differences between the DNA- and protein-based datasets. The LFQ-values of proteins of the Bacteroidaceae family were much higher when compared to the corresponding OTUs while Porphyromonadaceae-derived proteins revealed a lower abundance when related to the respective OTUs.

Similarly, the overall low abundant phyla of Elusimicrobia and Verrucomicrobia were not affected by the different diets or sample fractions. However, the abundance of Elusimicrobia proteins was much lower when compared to the respective OTUs whereas the abundance of Verrucomicrobia proteins was much higher when compared to the corresponding OTUs (Figure 10). A recent study employing total RNA sequencing as well as targeted RNA- and DNA amplicon sequencing identified the phyla of Elusimicrobia and Verrucomicrobia exclusively in the RNA-based datasets and proposed a higher activity of the respective phyla in the rumen (Li et al., 2016).

Figure 11 shows the average abundance of archaeal phyla, classes and families commonly identified by metaproteomics and 16S rRNA gene sequencing in diets and sample fractions. The archaeal community composition revealed no shifts in regard to the dietary treatments or ecological niches probably due to the less versatile metabolic capabilities when compared to bacteria (Henderson et al., 2015) and the comparably low numbers of identified proteins and OTUs. However, these results are consistent with the findings of a recent study that analyzed the abundance of total methanogens and Rumen Cluster C in the same sample material using quantitative real-time PCR (Lengowski et al., 2016).

Contrarily to the bacterial datasets, the archaeal community structures inferred from the quantified proteins differed clearly from the 16S rRNA gene-based analysis (Figure 11). All 626 OTUs were entirely assigned to three families of methanogens with the members of the family Methanobacteriaceae and Methanomassiliicoccaceae being dominant throughout all diets and sample fractions similar to previous reports (Janssen and Kirs, 2008; Seedorf et al., 2015). On the other hand, the metaproteomic analysis indicated greater archaeal diversity with most proteins being assigned to yet unclassified archaeal species (Figure 11). Moreover, the phyla of Crenarchaeota and Thaumarchaeota were found to be present, which is supported by other 16S rRNA gene studies that identified the respective phyla in low abundance in the rumen ecosystem (Shin et al., 2004; Wang et al., 2016; Jin et al., 2017). Averaged across all 27 samples, the most abundant groups within the phylum of Euryarchaeota were the families

of Thermococcaceae, Methanobacteriaceae and the order of Methanomicrobia (Figure 11).

The different phylogenetic distributions of archaeal proteins and OTUs might be attributed to the prevalence of *Methanobrevibacter ruminantium* and *M. gottschalkii* of the Methanobacteriaceae family which dominate the ruminal archaeal community (Janssen and Kirs, 2008; St-Pierre and Wright, 2013; Seedorf et al., 2015; Danielsson et al., 2017). High numbers of the respective 16S rRNA genes might have prevented sufficient sequencing reads of low abundant 16S rRNA genes in the current study. Moreover, some archaea are endo- and ectosymbiotically linked to protozoa (Lange et al., 2005), in particular species of Methanobacteriaceae family were found to be associated with protozoa (Janssen and Kirs, 2008). The specific sample preparation protocols for the shotgun metaproteomic analysis focusing on the enrichment of prokaryotic cells might have cause bias against protozoa-associated archaea, which could further explain the differences between the DNA- and protein-derived datasets.

However, the combination of protein- and DNA-based approaches may be advantageous for investigations of the rumen archaeal communities promising a higher coverage of the taxonomic diversity including as well functional information.

Comparing the phylogenetic distribution of archaeal proteins obtain from the first metaproteomic analysis (Chapter 3) and the main experiment (Chapter 4) revealed enormous variations as shown in Figure 29. Species of the phylum of Crenarchaeota and the class of Halobacteria appeared in much higher abundance in the first dataset whereas proteins of common rumen archaea such as the class of Thermoplasmata and Thermococci exhibited a drastically decreased abundance. Halobacteria usually inhabit hypersaline lakes and are not considered members of the rumen ecosystem. The NCBInr database contains about 1.1 million protein sequences for Halobacteria whereas only about 80,000 sequences are available for Thermoplasmata. This indicates again the importance of employing sample-specific search databases in metaproteomic studies to minimize the probability of false-positive identifications.

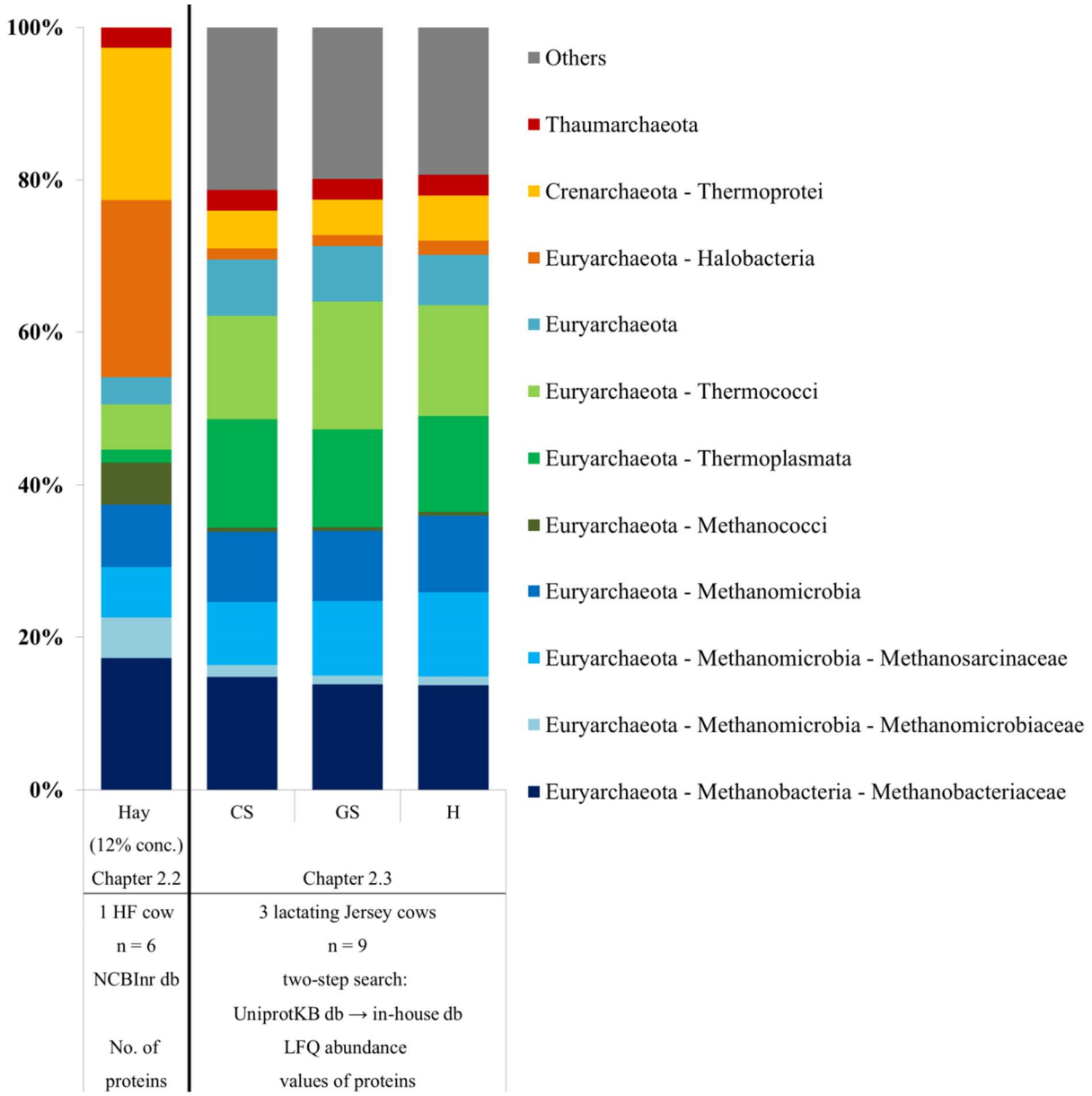

Figure 29| Comparison of the phylogenetic origin of archaeal proteins identified in the first metaproteomic study (Chapter 3) and the main experiment (Chapter 4) averaged over diets at phyla, class and family level. The diets (CS, corn silage-based diet; GS, grass silage-based diet; H, grass hay-based diet), breed of host animals (HF, Holstein-Frisian), numbers of samples, used search databases and type of abundance data are indicated.

5.5.2 Changes in Specific Functional Profiles

Apart from the dietary- and sample fraction-related effects on the protein- and DNA-based phylogenetic structures of the bacterial communities within the rumen, also shifts in specific functional profiles obtained from the metaproteomic analysis were observed (Chapter 4). The advantages of metaproteomic studies lie in the additional functional information crossing the gap between gene sequences obtained from metagenomic studies and the actually present phenotypes further involving the ability to access metabolic functions in the active microbial communities (Seifert et al., 2013).

5.5.2.1 Carbohydrate-Active Enzymes

In the course of the functional analyses carbohydrate-active enzymes (CAZymes), involved in the metabolism of carbohydrates, were annotated by searching the entirety of 8,163 quantified bacterial proteins against the database for automated CAZyme annotation with hidden Markov models (dbCAN HMMs v. 5.0, based on the CAZy db v. July 15, 2016). Using hmmscan of the HMMER3 software package (Yin et al., 2012) exclusively the best e-value hits were considered as described in Chapter 4.3.7.

Overall, 166 bacterial CAZymes were identified of which 91 distributed over 16 glycoside hydrolase (GH) families and 38 proteins belonged to seven different glycosyltransferase (GT) families. Moreover, 16 proteins were assigned to three families of carbohydrate esterases (CE), two proteins fell into two families of carbohydrate-binding modules (CBM) and 19 proteins with sequence similarity to S-layer homologies (SLH), involved in adhesion mechanisms of cellulosomes (Bayer et al., 1998; An and Friedman, 2000) were present (Table 8). Details about the abundance of CAZymes in the respective families of the five CAZy-categories as well as the corresponding taxonomic origin and contribution at phyla level in the diets and sample fractions are depicted in Figure 12. Further detailed descriptions can be found in Chapter 4.4.3.

The ratio of the most abundant CAZy-categories of GH, mainly produced by Bacteroidetes species and SLH, almost exclusively from Firmicutes species differed remarkably between the diets (Figure 12). The lower abundance of the Selenomonadaceae family throughout the sample fractions of GS diet (Figure 10) is underpinned by the low abundance of the SLH category in the respective sample fractions (Figure 12) that includes nine proteins of the Selenomonadaceae family (Table 8). The Gram-negative staining but phylogenetically Gram-positive rumen anaerobe *Selenomonas ruminantium* exhibits peptidoglycan-associated proteins with S-layer homology domains that play an important role

in the maintenance of the cell surface structure (Kojima et al., 2010).

Table 8| The numbers of identified CAZymes in the respective CAZy-categories and the corresponding bacterial origin at phyla and family level. GH, glycoside hydrolases; GT, glycosyltransferases; CE, carbohydrate esterases; CBM, carbohydrate-binding modules; SLH, S-layer homologies.

CAZyme categories (no. of families)	**GH** (16)	**GT** (7)	**CE** (3)	**CBM** (2)	**SLH**
Actinobacteria		3			
Bacteroidetes	5	2	1		
Bacteroidetes - Bacteroidaceae	4	1	1		
Bacteroidetes - Porphyromonadaceae	3	1			
Bacteroidetes - Prevotellaceae	46	13	6	1	
Fibrobacteres	1		1		
Firmicutes		1			4
Firmicutes - Acidaminococcaceae					2
Firmicutes - Lachnospiraceae	2	4			
Firmicutes - Ruminococcaceae	5				
Firmicutes - Selenomonadaceae	1	1			9
Firmicutes - Veillonellaceae					2
Firmicutes - Bacillales		2		1	
Firmicutes - Clostridiales	4				1
Proteobacteria	4	4	2		
Spirochaetes	1				
Synergistetes					1
Unclassified / environmental samples	15	6	5		

In total four GHs originated from the Clostridiales order, five GHs were produced by the Ruminococcaceae family and four GTs as well as two GHs were derived from the Lachnospiraceae family (Table 8). However, the most CAZymes belonged to the Prevotellaceae family including 46 GHs, 14 GTs and six CEs emphasizing their versatile metabolic activity. Six GTs and five CEs assigned to yet undescribed environmental samples (Table 8) appeared in high abundance dominating the CAZy-categories of GT and CE disregarding diets and sample fractions (Figure 12).

The category of CE exhibited the highest abundance throughout the sample fractions of the CS diet (Figure 12). The increased proteobacterial contribution to the CAZy-categories of GT and particularly CE in the LP and SP sample fractions of the CS diet (Figure 12) may support the observed rise of species of the Proteobacteria phylum in the sample fractions of the CS diet as shown in Figure 10. Two glycosyltransferases of the most abundant GT families, GT35 and GT4, were produced by *Bifidobacterium* species and were averagely more abundant in the GS diet supporting the overall higher abundance of the Bifidobacteriaceae family in the more fiber-rich GS diet (Chapter 5.5.1).

Biofuels produced from plant material constitute an important, renewable energy source (Blanch et al., 2008; Christopher et al., 2014). Nevertheless, the low activity of hydrolytic enzymes currently available and the therefore inefficient degradation of recalcitrant lignocellulose matter limit the production of biofuels (Hess et al., 2011). The discovery of new, more efficient enzymes from naturally occurring, biomass-degrading microbial communities such as the rumen microbiome represents a favorable strategy (Rubin, 2008). Hence, several functional metagenomic studies focused on the investigation and discovery of CAZymes within the rumen (Brulc et al., 2009; Hess et al., 2011; Ferrer et al., 2012; Pope et al., 2012; Wang et al., 2013; Patel et al., 2014a; Jose et al., 2017). Moreover, a recent metatranscriptomic study of the rumen microbial consortia assessed the presence of CAZymes based on messenger RNAs (Li and Guan, 2017).

Figure 30 shows a comparison of the abundance of 30 commonly identified GH families from the rumen microbiome of bovines. All respective studies employed the CAZy database (Lombard et al., 2014) for annotation but using different assignment algorithms. The numbers of overall identified GH families varies considerably between these studies with the lowest number of 24 identified GH families from the present study, even though hits with slightly higher e-values and low abundance were included for the comparison. Overall, the distribution of GH families obtained from the metagenomic studies appears similar to the results of the metaproteome-based dataset with a lower abundance of cellulases and hemicellulases and higher amounts of oligosaccharide-degrading enzymes. However, endo-hemicellulases appeared in a remarkably low abundance in the metaproteomic dataset of the current study (Figure 30). The metatranscriptomic study identified a comparably high number of endoglucanases within family GH9 and less oligosaccharide-degrading enzymes. GHs of the family GH2 were found in higher abundance in all metagenomics studies and the metaproteome-based dataset whereas appearing in low abundance in the metatranscriptomic dataset. Debranching enzymes were more abundant in the metagenomics analyses as when

compared to the metatranscriptomic and the current metaproteomic study (Figure 30).

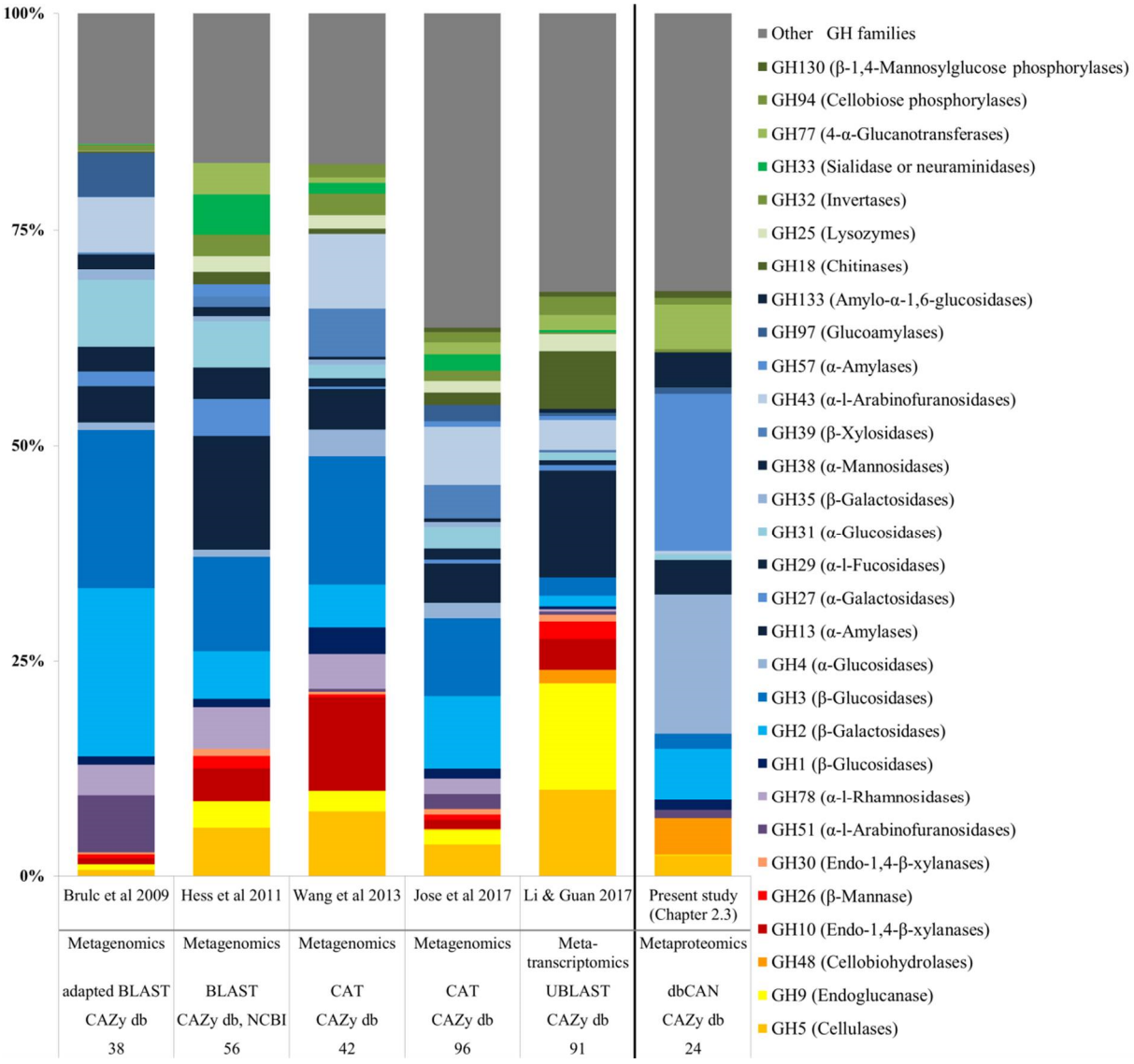

Figure 30| Comparison of commonly identified glycoside hydrolase families of recent studies investigating the rumen microbiome of *B. taurus* species (Brulc et al., 2009; Hess et al., 2011; Wang et al., 2013; Jose et al., 2017; Li and Guan, 2017). The respective methodologies, assignment algorithms and databases as well as the number of totally identified glycoside hydrolase families are given. Color code: yellowish, cellulases; reddish, endo-hemicellulases; purplish, debranching enzymes; bluish, oligosaccharide-degrading enzymes; greenish, uncategorized glycoside hydrolase families. CAT, CAZymes Analysis Toolkit (Park et al., 2010); dbCAN, dbCAN HMMs v. 5.0 (Yin et al., 2012).

The differences in abundance of certain GH families might be attributed to the different dietary regimen and ruminant breeds used in the respective studies. Furthermore, the different biomolecules analyzed and assignment procedures applied allow variations to be expected. As well, the identification of CAZymes is based on sequence homologies entailing the probability of incorrect assignments. Nevertheless, regarding the present study, most GHs were produced by species of the Bacteroidetes phylum also constituting a large part of the GTs which is in accordance with the study of Jose et al. (2017). Moreover, CAZymes identified in the study of Patel et al. (2014a) originated mainly from the phylum of Bacteroidetes.

5.5.2.2 ABC Transporters

ATP-binding cassette transporters (ABC transporters) usually comprise several subunits including integral membrane proteins and ATPases to catalyze the transport of compounds across the cell membranes using energy from the hydrolysis of ATP (Davidson et al., 2008; ter Beek et al., 2014). Further, ABC transporters are essential for the viability of cells since counteracting to detrimental changes within the cell (Poolman et al., 2004). Thus, ABC transporters are as well involved in resistance mechanisms of pathogenic microorganisms to antibiotics (Jones and George, 2004). Moreover, ABC importers play a crucial role in the uptake of external nutrients such as sugars, amino acids and trace metals (Cui and Davidson, 2011). Membrane proteins contain less lysine and arginine residues and thus, tend to be resistant to trypsin digestion as performed during the metaproteomic workflow (Saveliev et al., 2012). The amphipathic nature of membrane proteins make common shotgun metaproteomic-based identifications difficult (Santoni et al., 2000). Nevertheless, subunits of the ABC transporter complexes appeared to be accessible via mass spectrometry-based analyses.

A total of 170 bacterial proteins exhibited sequence similarities to ABC transporters as defined by the KEGG database. The average abundance of the respective membrane transporters in relation to the maximum LFQ-values within each group of transporters and the corresponding origin of proteins at bacterial phyla level in diets and sample fractions are depicted in Figure 13. Details about the bioinformatic assignment procedures are described in Chapter 4.3.7.

Overall, 106 proteins were assigned to the group of oligosaccharide, polyol, and lipid transporters that include several subunits of multiple sugar transport systems and of cellobiose, arabinose/lactose, maltose/maltodextrin, sorbitol/mannitol and galactose oligomer

transporters. The respective proteins were averagely more abundant in samples of the fiber-rich GS and H diets when compared to the CS diet. Within diets, the highest abundance was observed in the SP sample fractions. Major contributors were Firmicutes species followed the phyla of Spirochaetes and Proteobacteria (Figure 13). The increased abundance of the Lachnospiraceae family in the SP sample fractions of all diets shown in Figure 10 is supported by the higher abundance of 27 Lachnospiraceae proteins belonging to the group of oligosaccharide, polyol and lipid transporters in the SP fractions of all diets. Nine Spirochaetes proteins within the group of oligosaccharide, polyol and lipid transporters revealed the lowest abundance in all RF sample fractions confirming the abundances patterns described in Chapter 5.5.1. As mentioned in Chapter 5.5.1, genomes of *Bifidobacteria* comprise features for sugar ABC transporters (Pokusaeva et al., 2011) as verified by four Bifidobacteriaceae proteins belonging to the group of oligosaccharide, polyol and lipid transporters.

Conversely, the group of monosaccharide transporters comprising 20 proteins exhibited a higher abundance in all sample fractions of the CS diet (Figure 13), which might be attributed to the higher amount of starch present and the associated increased availability of monosaccharides. This group comprises subunits of ribose, rhamnose, methyl-galactoside and *sn*-glycerol 3-phosphate transport systems. Proteobacteria showed an increased contribution regarding monosaccharide transporters when compared to the respective phylogenetic origin of the group of oligosaccharide, polyol and lipid transporters. This may underpin the increased abundance of Proteobacteria and the included Succinivibrionaceae family in the sample fractions of the CS diet (Figure 10).

Similarly, the group of phosphate and amino acid transporters based on 27 proteins showed the highest abundance in samples of the CS diet (Figure 13). Phosphate and amino acid transporters were mainly derived from Proteobacteria and to a lesser extent from Firmicutes species.

Mineral and organic ion transporters from 17 proteins, almost exclusively produced by Firmicutes species, showed the by far highest LFQ-values in the SP fraction of the CS diet (Figure 13). Most proteins of this group were assigned to Iron (III) transporter systems. Uptake of iron is essential for almost all microorganism since it is required in many metabolic processes (Cornelis and Andrews, 2010).

Generally, the distinct abundance patterns of proteins within the groups of monosaccharide and oligosaccharide, polyol and lipid transporters may reflect the dietary influence on the bacterial communities with a higher availability of monosaccharides inferred from starch fermentation in the CS diet samples whereas more complex oligosaccharides from fiber-degradation are present in the GS and H diets.

Proteins with sequence homology to ABC transporter subunits were also identified within the first metaproteomic analysis (Chapter 3). As depicted in Figure 31, the numbers of proteins belonging to the group of oligosaccharide, polyol and lipid transporters increased in the SP fractions as observed for the experiment described in Chapter 4 (Figure 13). Moreover, most ABC transporter proteins were derived from Firmicutes species similar to the main experiment (Figure 13) which is further confirmed by a 16S rRNA gene-based study of the rumen microbiome of sheep (Wang et al., 2017) that used the PICRUSt tool for functional prediction (Langille et al., 2013).

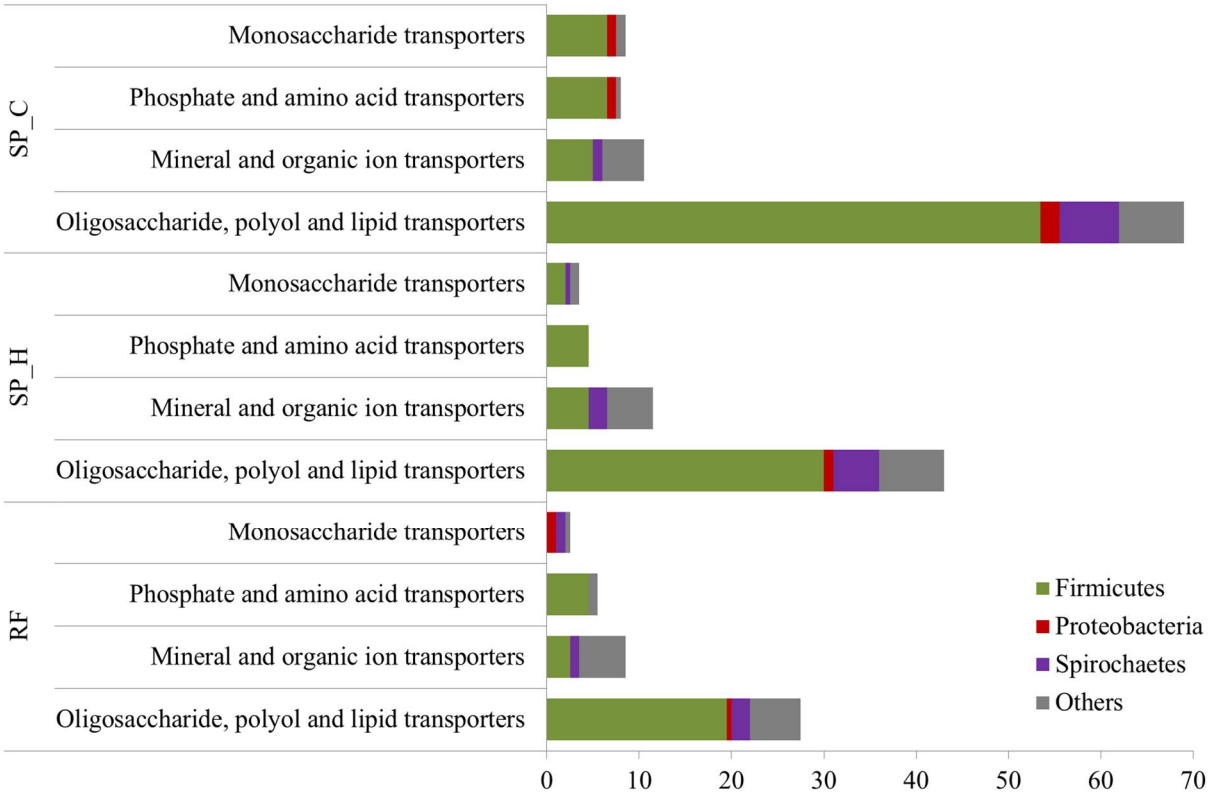

Figure 31| Numbers of bacterial ABC transporters in different groups averagely (n = 2) identified in samples of the first metaproteomic analysis (Chapter 3). Phylogenetic origin of proteins in the respective samples is shown at phyla level. RF, rumen fluid; SP_H, solid phase extracted with Histodenz protocol; SP_C, solid phase extracted with cheesecloth protocol.

5.5.2.3 Enzymes Involved in Carbon Metabolism

A fundamental aspect of all life is the carbon metabolism that comprises the biochemical pathways for carbon usage respectively glycolysis including the Embden-Meyerhof-Parnas (EMP) pathway, the pentose phosphate pathway and the citrate cycle. Moreover, the pathways of methane metabolism are included. The EMP pathway is the most widespread variant of glycolysis converting glucose to pyruvate yielding energy in form of ATP and NADH (nicotinamide adenine dinucleotide). A parallel pathway is the rather anabolic pentose phosphate pathway forming NADPH (nicotinamide adenine dinucleotide phosphate), pentoses and ribose 5-phosphate which servers as precursor for nucleotide synthesis. The citrate cycle uses acetate in form of acetyl-CoA derived from pyruvate and H_2O producing carbon dioxide and NADH respectively ATP and involves as well succinate.

Sums of 1,121 bacterial and 28 archaeal proteins were assigned to 80 and eight KO identifiers respectively, carrying out 70 and five KEGG REACTIONS that belong to the pathways of the carbon metabolism as defined by the KEGG database. To visualize the abundance of the respective enzymes in the carbon metabolism pathway map, the respective KEGG REACTIONS were arranged into 60 bacterial and five archaeal functional groups. Varying abundance, in diets and sample fractions, of enzymes in the respective functional groups involved in the carbon metabolism are shown in Figure 14A. Additionally, eight compounds identified by NMR (Table 5) are included in the pathway map of the carbon metabolism (Figure 14A). Details about the proteins, the KO identifiers and KEGG RACTIONS in the functional grouping are listed in supplementary Table S6 of Chapter 4. A few KOs were assigned to more than one KEGG REACTION and thus, appear more than once in the grouping. A detailed description of differential abundance of specific functional groups of the carbon metabolism can be found in Chapter 4.4.5.

The sums of LFQ-values of 417 proteins belonging to the functional groups involved in the EMP pathway revealed the highest abundance in the LP fractions of the GS and the H diets whereas being most abundant in the SP fraction of the CS diet (Figure 14B). The LFQ-values of groups belonging to the citrate cycle comprising a total of 76 proteins exhibited a remarkably lower abundance in the SP fractions of the GS and H diets when compared to the respective RF and LP fractions and the CS sample fraction (Figure 14C). Looking at the groups of the pentose phosphate pathway based on 34 proteins, the sum of LFQ-values showed higher abundance in the LP fractions of the GS and H diets with similar patterns to the EMP pathway (Figure 14D).

The increased abundance of polysaccharide-degrading bacteria in the SP fractions described in Chapter 5.5.1 might have caused an increased availability of monosaccharides and a higher abundance of proteins involved in glycolysis by the EMP pathway. Considering the more mushy texture and the higher amounts of starch in the CS diet most monosaccharides possibly were present in the SP fraction of the CS diet whereas most monosaccharides from degradation of structural plant polysaccharides were rather present in the LP fractions of the more fibrous-structured GS and H diets (Figure 14B). This might explain as well the high abundance of the pentose phosphate pathway in the LP fractions of the GS and H diets that probably contained most sugars within the fiber-rich diets (Figure 14D). Proteins involved in the citric acid cycle showed a remarkably low abundance in the SP fractions of the GS and H diets pointing towards a low abundance of substrates like acetate and succinate, which is in accordance with the abundance values of proteins of EMP and pentose phosphate pathway (Figures 14B,C,D).

Moreover, the alpha, beta and gamma subunits of the methyl-coenzyme M reductase (MCR) were identified based on 15 highly abundant proteins that constituted 23% of the total abundance of 55 proteins that belonged to the methane metabolism as defined by the KEGG database. The MCR catalyzes the final step in the formation of methane (Jaun and Thauer, 2007; Ragsdale, 2014) and the mcrA gene, encoding the alpha subunit of MCR, is preferably used as a functional marker since present in all methanogens (Friedrich, 2005; Denman et al., 2007). A study analyzing the same sample material investigated the abundance of total methanogens in targeting the mcrA gene by quantitative real-time PCR (Lengowski et al., 2016) confirming the results of the current study regarding dietary and fraction-related effects. Eleven MCR-derived proteins were produced by species of the Methanobacteriaceae family and four proteins were assigned to the phylum of Thermoplasmata.

Carbon is the terminal electron acceptor for the most prevalent, hydrogenotrophic methanogenesis in the rumen requiring carbon dioxide, hydrogen or formic acid as identified by NMR (Table 5). Besides, methylotrophic methanogens like the family of Methanomassiliicoccaceae belonging to the class of Thermoplasmata utilize compounds like methylamine, dimethylamine and trimethylamine, as well identified by NMR in the RF fractions of all diets (Table 5), as their major energy and carbon sources. Trimethylamines are formed by bacteria from plant-derived glycine betaine and cholin (Neill et al., 1978; Mitchell et al., 1979). Methylamine, dimethylamine and trimethylamine were found in higher concentrations in the GS and H diets (Table 5) which is supported by other reports that indicate higher glycine betaine concentrations in grass as when compared to corn (Lerma et

al., 1991; Chendrimada et al., 2002). However, a trimethylamine corrinoid protein MttC of methanogenic archaeon ISO4-H5 was identified. Moreover, the ratio of acetate to propionate, an indicator for methanogenic activity, was higher in the more fiber-rich GS and H diets (Table 5) that further contained more than twice as much methylamines. Thus, a higher activity of methylotrophic methanogenesis in the fiber-rich diets can be assumed. A previous study suggested focusing on methylamines rather than on hydrogen for mitigating methane emission from the rumen (Poulsen et al., 2013).

5.5.2.4 Enzymes of Short-Chain Fatty Acid Production

The majority of metabolites produced by bacteria in the rumen are the short-chain fatty acids acetate, butyrate and propionate, which represent the primary source of energy to the ruminant host animal (Bergman, 1990) and further the fermentation intermediate formate (Hungate et al., 1970) that is also used as substrate in hydrogenotrophic methanogenesis. Microbial short-chain fatty acids influence many physiological processes directly affecting the host (Koh et al., 2016) and serve as a link between dietary impact on the structure and metabolic activity of the microbiome and its influence on the host. In contrast to metabolites, enzymes involved in their production can be assigned to a specific taxonomic origin.

Bacterial enzymes involved in the production of acetate, butyrate, propionate and formic acid were retrieved using the COG assignments from WebMGA (Wu et al., 2011). The LFQ-values of proteins belonging to COGs modified from a previous study (Polansky et al., 2015) were summarized over the diets including the phylogenetic origin at phyla and family level and are depicted in Figure 15. A detailed description can be found in Chapter 4.4.6.

In total 35 and 45 enzymes involved in acetate (Figure 15C) and butyrate production (Figure 15A) respectively were identified. Moreover, 52 proteins belonged to propionate production (Figure 15B) and 23 enzymes were associated with the generation of formic acid (Figure 15D).

Six proteins of *Phascolarctobacterium* species involved in propionate production appeared in highest abundance in the CS diet, which may support the assumed rise of the Acidaminococcaceae family as consequence of increased amounts of succinate (Table 5) produced from Proteobacteria and the family of Succinivibrionaceae in the fractions of the CS diet as described in Chapter 5.5.1. Moreover, four enzymes associated with propionate formation belonged to *Selenomonas* species and exhibited the lowest abundance in the GS diet (Figure 15B) similar to the decreased abundance of Selenomonadaceae proteins in the SLH category in the fraction of the GS diet described in Chapter 5.5.2.1. These findings might

support the overall lower abundance of the Selenomonadaceae family observed in Figure 10. A 3-hydroxybutyryl-CoA dehydrogenase of *Megasphaera elsdenii*, involved in butyrate formation, was found in a remarkably high abundance of 24% in the CS diet when compared to the abundance of 0.4% and 0.1% in the GS and H diets respectively (Figure 15A). Butyrate is a common fermentation product of *Megasphaera elsdenii* (Table 1). Thirteen Lachnospiraceae proteins dominated in butyrate formation and revealed the highest abundance of 56% and 58% in the fiber-rich GS and H diets (Figure 15A) underpinned by the higher amounts of butyrate in the respective RF fractions as determined by NMR (Table 5). Considering the abundance of proteins, enzymes of the Prevotellaceae family prevailed in acetate and propionate production constituting above 60% and 55% of the total abundance in all diets (Figures 15B,C) again reflecting the functional dominance of *Prevotella* species within the rumen ecosystem as supported by the numbers and abundances of CAZymes (Chapter 5.5.2.1) and the overall abundance obtained from the 16S rRNA gene and protein-based approaches (Figure 10).

5.5.3 Variation of Metabolites in Rumen Fluid

Using the Chenomx NMR Suite 8.2 software, 20 different compounds were identified in all RF samples. Two-dimensional NMR spectroscopy further validated the presence of 12 compounds (*) including the major short-chain fatty acids acetate, butyrate, propionate and valerate (Table 5). There were no statistically significant alterations of the rumen metabolome patterns regarding the different diets. A detailed description of the amounts of identified metabolites in the RF fractions can be found in Chapter 4.4.7.

Despite the statistical insignificancies, some patterns in metabolome composition might support findings of the 16Sr DNA and the metaproteome analyses. The observed rise of the phylum of Proteobacteria and the family of Succinivibrionaceae in the CS diet fractions might be interlinked with the increased abundance of the Acidaminococcaceae family due to the enhanced production of succinate (Chapter 5.5.1). Succinate concentration averagely was highest in the RF fraction of the CS diet (Table 5). Moreover, six proteins of *Phascolarctobacterium* species that use succinate to produce propionate were involved in propionate production and appeared in highest abundance in the CS diet (Chapter 5.5.2.4). Most lactate was present in the RF samples of the H diet (Table 5), which might support the increased abundance of the family of Erysipelotrichaceae that assumedly produce lactate (Chapter 5.5.1). Thirteen Lachnospiraceae proteins dominated in butyrate formation revealing the highest abundance in GS and H diets (Chapter 5.5.2.4) underpinned by the higher amounts

of butyrate in the respective RF fractions as determined by NMR. As mentioned in Chapter 5.5.1, *Ruminococcus albus* requires phenylpropionate for the adherence to and degradation of cellulose (Stack and Hungate, 1984). Higher amounts of methylamines in the fiber-rich GS and H diets may confirm the findings of previous studies (Chapter 5.5.2.3).

5.6 Conclusions and Perspectives

Progress in efficiency of animal production requires optimized feeding strategies, which presuppose an improved understanding of the dietary impact on the complex bionetwork residing in the rumen. Targeting the actual functions out of numerous possibilities that are encoded by the genomes of the rumen microbiome is essential to obtain enhanced knowledge about the manifold interactions of different members of the microbial communities including as well host-microbe interactions. Therefore, proteins are best suited since representing the actual function of investigated cells further entailing phylogenetic information.

Thus, the major aim of the present project was the initial establishment of suitable mass spectrometry-based shotgun metaproteomics to investigate the rumen microbiome providing deeper functional insights. Despite the challenging obstacles such as the heterogeneity of the rumen sample matter, the presence of inhibiting substances and demanding bioinformatic data processing steps, a workflow to characterize the prokaryotic rumen metaproteome was successfully implemented forming a solid basis for further investigations.

Thereupon, the dietary impact of the most common forages in ruminant production on the structure and the function of the rumen microbiome was analyzed by a combination of complementary, state of the art methods. The archaeal and bacterial communities inhabiting different ecological niches within the rumen were investigated using GeLC-MS/MS-based metaproteomics flanked by Illumina MiSeq amplicon sequencing of the 16S rRNA gene. Moreover, the metabolic profiles of the rumen fluid were assessed by nuclear magnetic resonance to obtain most precise and comprehensive insights.

The obtained datasets revealed significant alterations of the structure and function of the microbial communities in response to the dietary treatments as determined unanimously by the protein- and the DNA-based analyses. Certain contrasts between the different methods employed clearly emphasized the benefits of using combinations of complementary methods to study the microbiome of complex ecosystems like the rumen. Moreover, tremendous variations in community composition and functional patterns regarding the different

ecological niches within the rumen were observed by both methods prompting for the necessity of sample fractionation in rumen studies to cover the effects of applied treatments throughout the whole ecosystem. On top of the structural information obtained from DNA- and protein-based analyses, the metaproteomic approach provided additional functional insights elucidating the varying abundance of carbohydrate-active enzymes, ATP-binding cassette transporters, metabolic pathways of the carbon metabolism and enzymes involved in short-chain fatty acid production further supplemented by the metabolomics analysis.

Regarding the influence of different animal feeds, the abundance of proteobacterial species and the family of Succinivibrionaceae was significantly altered upon the increased availability of starch in the CS diet further promoting the formation of succinate. This may be linked to cross-feeding of succinate-utilizing and propionate-producing species of the Acidaminococcaceae family that exhibited a similar rise in abundance in response to CS diet.

Moreover, the yet sparsely described family of Erysipelotrichaceae was promoted by the H diet assumptively being involved in lactic acid production, which suggests further investigations of the respective species.

Prominent members of the bacterial rumen community such as the family of Selenomonadaceae have not been identified by the 16S rRNA gene-based analysis. Selenomonadaceae proteins were less abundant in GS diet, which correlated with the lower abundance of *Butyrivibrio* proteins in the GS diet accentuating cross-feeding between the respective species.

The Bifidobacteriaceae family exhibited a higher abundance in the more fiber-rich GS and H diets emphasizing the polysaccharide-degrading capabilities of this species. Species of the order of Clostridiales including the Lachnospiraceae and Ruminococcaceae families exhibited an increased abundance throughout the SP sample factions regardless the dietary treatments confirming their crucial fiber-degrading potential.

The most dominant ruminal bacteria of the Prevotellaceae family did not respond to dietary changes reflecting their versatile metabolic capabilities. Prevotellaceae species were more present in the RF and LP fractions not being primarily involved in fiber-degradation.

The role of low abundant phyla such as Elusimicrobia, Synergistetes, Tenericutes and Verrucomicrobia within the rumen ecosystems remains comparably vague since not much information about the respective species is available. This suggests microbiological investigations of the respective species in pure and mixed cultures to obtain a more valid basis regarding their functional capabilities.

The archaeal community did not respond to dietary changes and exhibited similar distributions across the ecological niches, which may be attributed to their less versatile metabolic capabilities when compared to bacteria. Moreover, differences in the DNA- and protein-based phylogenetic distributions suggest further technical and methodological improvements.

The current project focused exclusively on the most significant members of the rumen microbiome, namely archaea and bacteria. However, further metaproteomic investigations might include as well anaerobic fungi and particularly ruminal protozoa to gain more information about the interactions of methanogens and ciliate protozoa and their role in within the rumen ecosystem including the energy loss from methane emissions.

Concerning the outcome of this project, it might be important to consider that each of the experimental diets comprised 52% of identical concentrate mixture thus, differing less in chemical composition. Future studies of the dietary impact on the rumen microbiome could obtain more significant results in applying more different dietary treatments, which might provide deeper insights into the adaptation responses of the microbial communities.

Furthermore, a parallel study highlighted the inevitability of repeated sampling time points to cover treatment effects on a diurnal basis, which could not be addressed in the current project due to the complex, labor and particularly time intensive analysis procedures.

In the course of this project, no significant host-related effects on the rumen microbiome were detected. Nevertheless, the used animals exhibited differences regarding certain physiological parameters. To obtain more comprehensive and confident data, future studies may include a higher number of animals to increase the statistical reliability.

In terms of the rumen microbiome, the microorganisms attached to the epimural layer were not investigated in the current project. However, shotgun metaproteomic analyses might provide detailed insights into the functions of the epimural microbiome located at the rumen wall that besides scavenging of oxygen and the hydrolysis of urea is involved in many crucial processes. Despite the comparably low overall abundance, the epimural communities are closest to the host and probably are most involved in host-microbe interactions. Biofilm formation at the mucosa of the rumen by the epimural communities may serve as a protection of the epithelium against pathogens and nutritional diseases such as the subacute ruminal acidosis. Shotgun metaproteomics harbors the potential to elucidate fundamental host-microbe interactions at the interface between host and feed. Furthermore, the probably competitive relation of the epimural microbiome with the residual communities might be investigated using metaproteomics.

The results of the present project may provide deeper insights into the complicated network of bacterial interactions and adaptions to various substrates. So far, to the best of my knowledge, this was the first study investigating the impact of the most common forages used as feed for dairy cows and fattening cattle on the metaproteome expressed by the entirety of archaeal and bacterial communities in the different phases of the rumen ecosystem. The advantages of shotgun metaproteomic-based investigations and the combination of complementary methods were clearly demonstrated. Bioinformatic and technical progress may enhance the metaproteomic coverage of future studies.

CHAPTER VI

BIBLIOGRAPHY

6. BIBLIOGRAPHY

Abrahamse, P.A., Vlaeminck, B., Tamminga, S., and Dijkstra, J. (2008). The effect of silage and concentrate type on intake behavior, rumen function, and milk production in dairy cows in early and late lactation. *J Dairy Sci* 91(12)**,** 4778-4792. doi: 10.3168/jds.2008-1350.

Acinas, S.G., Marcelino, L.A., Klepac-Ceraj, V., and Polz, M.F. (2004). Divergence and redundancy of 16S rRNA sequences in genomes with multiple rrn operons. *J Bacteriol* 186(9)**,** 2629-2635. doi: 10.1128/JB.186.9.2629-2635.2004.

Aikman, P.C., Henning, P.H., Humphries, D.J., and Horn, C.H. (2011). Rumen pH and fermentation characteristics in dairy cows supplemented with Megasphaera elsdenii NCIMB 41125 in early lactation. *J Dairy Sci* 94(6)**,** 2840-2849. doi: 10.3168/jds.2010-3783.

Akin, D.E., and Borneman, W.S. (1990). Role of rumen fungi in fiber degradation. *J Dairy Sci* 73(10)**,** 3023-3032. doi: 10.3168/jds.S0022-0302(90)78989-8.

Akin, D.E., Borneman, W.S., and Windham, W.R. (1988). Rumen fungi: morphological types from Georgia cattle and the attack on forage cell walls. *Biosystems* 21(3-4)**,** 385-391. doi: 10.1016/0303-2647(88)90037-8.

Aldridge, B.B., and Rhee, K.Y. (2014). Microbial metabolomics: innovation, application, insight. *Curr Opin Microbiol* 19**,** 90-96. doi: 10.1016/j.mib.2014.06.009.

Allison, M.J., Mayberry, W.R., McSweeney, C.S., and Stahl, D.A. (1992). Synergistes jonesii, gen. nov., sp.nov.: A Rumen Bacterium That Degrades Toxic Pyridinediols. *Syst Appl Microbiol* 15(4)**,** 522-529. doi: 10.1016/S0723-2020(11)80111-6.

Amann, R.I., Ludwig, W., and Schleifer, K.H. (1995). Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol Rev* 59(1)**,** 143-169.

American Jersey Cattle Association, National All-Jersey Inc. (2006). Jersey Directory. American Jersey Cattle Association, National All-Jersey Inc. www.usjersey.com

Ametaj, B.N., Zebeli, Q., Saleem, F., Psychogios, N., Lewis, M.J., Dunn, S.M., et al. (2010). Metabolomics reveals unhealthy alterations in rumen metabolism with increased proportion of cereal grain in the diet of dairy cows. *Metabolomics* 6(4)**,** 583-594. doi: 10.1007/s11306-010-0227-6.

An, Y.H., and Friedman, R.J. (2000). Handbook of Bacterial Adhesion: Principles, Methods, and Applications. Humana Press. doi: 10.1007/978-1-59259-224-1.

Anderson, K.L. (1995). Biochemical analysis of starch degradation by Ruminobacter amylophilus 70. *Appl Environ Microbiol* 61(4)**,** 1488-1491.

Ann Huws, S., Chiariotti, A., Sarubbi, F., Carfi, F., and Pace, V. (2012). Effects of feeding Mediterranean buffalo sorghum silage versus maize silage on the rumen microbiota and milk fatty acid content. *J Gen Appl Microbiol* 58(2)**,** 107-112. doi: 10.2323/jgam.58.107.

Asanuma, N., and Hino, T. (2002). Regulation of fermentation in a ruminal bacterium, Streptococcus bovis, with special reference to rumen acidosis. *Anim Sci J* 73(5)**,** 313-325. doi: 10.1046/j.1344-3941.2002.00044.x.

Bach, A., Calsamiglia, S., and Stern, M.D. (2005). Nitrogen metabolism in the rumen. *J Dairy Sci* 88 Suppl 1**,** E9-21. doi: 10.3168/jds.S0022-0302(05)73133-7.

Bach, S.J., McAllister, T.A., Veira, D.M., Gannon, V.P.J., and Holley, R.A. (2002). Transmission and control of Escherichia coli O157:H7 — A review. *Can J Anim Sci* 82(4)**,** 475-490. doi: 10.4141/A02-021.

Bantscheff, M., Schirle, M., Sweetman, G., Rick, J., and Kuster, B. (2007). Quantitative mass spectrometry in proteomics: a critical review. *Anal Bioanal Chem* 389(4)**,** 1017-1031. doi: 10.1007/s00216-007-1486-6.

Bashiardes, S., Zilberman-Schapira, G., and Elinav, E. (2016). Use of Metatranscriptomics in Microbiome Research. *Bioinform Biol Insights* 10**,** 19-25. doi: 10.4137/BBI.S34610.

Bayer, E.A., Chanzy, H., Lamed, R., and Shoham, Y. (1998). Cellulose, cellulases and cellulosomes. *Curr Opin Struct Biol* 8(5)**,** 548-557. doi: 10.1016/S0959-440X(98)80143-7.

Beauchemin, K.A., and McGinn, S.M. (2005). Methane emissions from feedlot cattle fed barley or corn diets. *J Anim Sci* 83(3)**,** 653-661. doi:10.2527/2005.833653x.

Belanche, A., de la Fuente, G., Pinloche, E., Newbold, C.J., and Balcells, J. (2012). Effect of diet and absence of protozoa on the rumen microbial community and on the representativeness of bacterial fractions used in the determination of microbial protein synthesis. *J Anim Sci* 90(11)**,** 3924-3936. doi: 10.2527/jas.2011-4802.

Belenguer, A., Toral, P.G., Frutos, P., and Hervas, G. (2010). Changes in the rumen bacterial community in response to sunflower oil and fish oil supplements in the diet of dairy sheep. *J Dairy Sci* 93(7)**,** 3275-3286. doi: 10.3168/jds.2010-3101.

Benson, A.K., Kelly, S.A., Legge, R., Ma, F., Low, S.J., Kim, J., et al. (2010). Individuality in gut microbiota composition is a complex polygenic trait shaped by multiple environmental and host genetic factors. *Proc Natl Acad Sci USA* 107(44)**,** 18933-18938. doi: 10.1073/pnas.1007028107.

Bergman, E.N. (1990). Energy contributions of volatile fatty acids from the gastrointestinal tract in various species. *Physiol Rev* 70(2)**,** 567-590.

Bergmann, I., Mundt, K., Sontag, M., Baumstark, I., Nettmann, E., and Klocke, M. (2010). Influence of DNA isolation on Q-PCR-based quantification of methanogenic Archaea in biogas fermenters. *Syst Appl Microbiol* 33(2)**,** 78-84. doi: 10.1016/j.syapm.2009.11.004.

Bladen, H.A., Bryant, M.P., and Doetsch, R.N. (1961). A Study of Bacterial Species from the Rumen Which Produce Ammonia from Protein Hydrolyzate. *Appl Microbiol* 9(2), 175-180.

Blanch, H.W., Adams, P.D., Andrews-Cramer, K.M., Frommer, W.B., Simmons, B.A., and Keasling, J.D. (2008). Addressing the Need for Alternative Transportation Fuels: The Joint BioEnergy Institute. *ACS Chem Biol* 3(1), 17-20. doi: 10.1021/cb700267s.

Borneman, W.S., Ljungdahl, L.G., Hartley, R.D., and Akin, D.E. (1991). Isolation and characterization of p-coumaroyl esterase from the anaerobic fungus Neocallimastix strain MC-2. *Appl Environ Microbiol* 57(8), 2337-2344.

Brask, M., Lund, P., Hellwing, A.L.F., Poulsen, M., and Weisbjerg, M.R. (2013). Enteric methane production, digestibility and rumen fermentation in dairy cows fed different forages with and without rapeseed fat supplementation. *Anim Feed Sci Tech* 184(1-4), 67-79. doi: ht10.1016/j.anifeedsci.2013.06.006.

Briggs, H.M., and Briggs, D.M. (1980). Modern breeds of livestock. Forth Edition. MacMillan Company.

Brock, F.M., Forsberg, C.W., and Buchanan-Smith, J.G. (1982). Proteolytic activity of rumen microorganisms and effects of proteinase inhibitors. *Appl Environ Microbiol* 44(3), 561-569.

Brooks, J.P., Edwards, D.J., Harwich, M.D., Jr., Rivera, M.C., Fettweis, J.M., Serrano, M.G., et al. (2015). The truth about metagenomics: quantifying and counteracting bias in 16S rRNA studies. *BMC Microbiol* 15, 66. doi: 10.1186/s12866-015-0351-6.

Brown, D.R., Bradbury, J.M., and Johansson, K.-E. (2015). Anaeroplasma. Bergey's Manual of Systematics of Archaea and Bacteria. doi: 10.1002/9781118960608.gbm01257.

Brulc, J.M., Antonopoulos, D.A., Miller, M.E., Wilson, M.K., Yannarell, A.C., Dinsdale, E.A., et al. (2009). Gene-centric metagenomics of the fiber-adherent bovine rumen microbiome reveals forage specific glycoside hydrolases. *Proc Natl Acad Sci U SA* 106(6), 1948-1953. doi: 10.1073/pnas.0806191105.

Bryant, M.P. (1956). THE CHARACTERISTICS OF STRAINS OF SELENOMONAS ISOLATED FROM BOVINE RUMEN CONTENTS. *J Bacteriol* 72(2), 162-167.

Bryant, M.P. (1959). Bacterial species of the rumen. *Bacteriol Rev* 23(3), 125-153.

Bryant, M.P. (1973). Nutritional requirements of the predominant rumen cellulolytic bacteria. *Fed Proc* 32(7), 1809-1813.

Bryant, M.P. (2015). Succinivibrio. Bergey's Manual of Systematics of Archaea and Bacteria. doi: 10.1002/9781118960608.gbm01087.

Bryant, M.P., and Small, N. (1956). The Anaerobic Monotrichous butyric acid-producing curved rod-shaped bacteria of the rumen. *J Bacteriol* 72(1), 16-21.

Bryant, M.P., Small, N., Bouma, C., and Chu, H. (1958). Bacteroides ruminicola n. sp. ans succinimonas amylolytica the new genus and species: Species of Succinic Acid-Producing Anaerobic Bacteria of the Bovine Rumen. *J Bacteriol* 76(1)**,** 15-23.

Burbach, K., Seifert, J., Pieper, D.H., and Camarinha-Silva, A. (2016). Evaluation of DNA extraction kits and phylogenetic diversity of the porcine gastrointestinal tract based on Illumina sequencing of two hypervariable regions. *MicrobiologyOpen* 5(1)**,** 70-82. doi: 10.1002/mbo3.312.

Cai, L., Ye, L., Tong, A.H., Lok, S., and Zhang, T. (2013). Biased diversity metrics revealed by bacterial 16S pyrotags derived from different primer sets. *PLoS One* 8(1)**,** e53649. doi: 10.1371/journal.pone.0053649.

Camarinha-Silva, A., Jauregui, R., Chaves-Moreno, D., Oxley, A.P., Schaumburg, F., Becker, K., et al. (2014). Comparing the anterior nare bacterial community of two discrete human populations using Illumina amplicon sequencing. *Environ Microbiol* 16(9)**,** 2939-2952. doi: 10.1111/1462-2920.12362.

Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009). The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res* 37(Database issue)**,** D233-D238. doi: 10.1093/nar/gkn663.

Capper, J.L., and Cady, R.A. (2012). A comparison of the environmental impact of Jersey compared with Holstein milk for cheese production. *J Dairy Sci* 95(1)**,** 165-176. doi: 10.3168/jds.2011-4360.

Carberry, C.A., Kenny, D.A., Han, S., McCabe, M.S., and Waters, S.M. (2012). Effect of phenotypic residual feed intake and dietary forage content on the rumen microbial community of beef cattle. *Appl Environ Microbiol* 78(14)**,** 4949-4958. doi: 10.1128/aem.07759-11.

Castro-Carrera, T., Toral, P.G., Frutos, P., McEwan, N.R., Hervas, G., Abecia, L., et al. (2014). Rumen bacterial community evaluated by 454 pyrosequencing and terminal restriction fragment length polymorphism analyses in dairy sheep fed marine algae. *J Dairy Sci* 97(3)**,** 1661-1669. doi: 10.3168/jds.2013-7243.

Catherman, A.D., Skinner, O.S., and Kelleher, N.L. (2014). Top Down Proteomics: Facts and Perspectives. *Biochem Biophys Res Commun* 445(4)**,** 683-693. doi: 10.1016/j.bbrc.2014.02.041.

Cersosimo, L.M., Bainbridge, M.L., Kraft, J., and Wright, A.-D.G. (2016). Influence of periparturient and postpartum diets on rumen methanogen communities in three breeds of primiparous dairy cows. *BMC Microbiol* 16(1)**,** 78. doi: 10.1186/s12866-016-0694-7.

Chandler, D.P., Fredrickson, J.K., and Brockman, F.J. (1997). Effect of PCR template concentration on the composition and distribution of total community 16S rDNA clone libraries. *Mol Ecol* 6(5)**,** 475-482.

Chendrimada, T.P., Neto, M.G., Pesti, G.M., Davis, A.J., and Bakalli, R.I. (2002). Determination of the betaine content of feed ingredients using high-performance liquid chromatography. *J Sci Food Agr* 82(13)**,** 1556-1563. doi: 10.1002/jsfa.1214.

Chourey, K., Jansson, J., VerBerkmoes, N., Shah, M., Chavarria, K.L., Tom, L.M., et al. (2010). Direct cellular lysis/protein extraction protocol for soil metaproteomics. *J Proteome Res* 9(12)**,** 6615-6622. doi: 10.1021/pr100787q.

Christopher, L.P., Yao, B., and Ji, Y. (2014). Lignin Biodegradation with Laccase-Mediator Systems. *Front Energy Res* 2(12). doi: 10.3389/fenrg.2014.00012.

Cole, J.R., Wang, Q., Cardenas, E., Fish, J., Chai, B., Farris, R.J., et al. (2009). The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Res* 37(Database issue)**,** D141-145. doi: 10.1093/nar/gkn879.

Cole, J.R., Wang, Q., Fish, J.A., Chai, B., McGarrell, D.M., Sun, Y., et al. (2014). Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res* 42(Database issue)**,** D633-642. doi: 10.1093/nar/gkt1244.

Cornelis, P., and Andrews, S.C. (2010). Iron Uptake and Homeostasis in Microorganisms. Caister Academic Press.

Cotta, M.A. (1990). Utilization of nucleic acids by Selenomonas ruminantium and other ruminal bacteria. *Appl Environ Microbiol* 56(12)**,** 3867-3870.

Cotta, M.A. (1992). Interaction of ruminal bacteria in the production and utilization of maltooligosaccharides from starch. *Appl Environ Microbiol* 58(1)**,** 48-54.

Cotta, M.A., and Zeltwanger, R.L. (1995). Degradation and utilization of xylan by the ruminal bacteria Butyrivibrio fibrisolvens and Selenomonas ruminantium. *Appl Environ Microbiol* 61(12)**,** 4396-4402.

Cottrell, J.S. (2011). Protein identification using MS/MS data. *J Proteomics* 74(10)**,** 1842-1851. doi: 10.1016/j.jprot.2011.05.014.

Counotte, G.H.M., Prins, R.A., Janssen, R.H.A.M., and deBie, M.J.A. (1981). Role of Megasphaera elsdenii in the Fermentation of dl-[2-(13)C]lactate in the Rumen of Dairy Cattle. *Appl Environ Microbiol* 42(4)**,** 649-655.

Cox, J., Hein, M.Y., Luber, C.A., Paron, I., Nagaraj, N., and Mann, M. (2014). Accurate Proteome-wide Label-free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ. *Mol Cell Proteomics : MCP* 13(9)**,** 2513-2526. doi: 10.1074/mcp.M113.031591.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* 26. doi: 10.1038/nbt.1511.

Craig, R., and Beavis, R.C. (2004). TANDEM: matching proteins with tandem mass spectra. *Bioinformatics* 20(9)**,** 1466-1467. doi: 10.1093/bioinformatics/bth092.

Creevey, C.J., Kelly, W.J., Henderson, G., and Leahy, S.C. (2014). Determining the culturability of the rumen bacterial microbiome. *Microb Biotechnol* 7(5)**,** 467-479. doi: 10.1111/1751-7915.12141.

Cremonesi, P., Vanoni, L., Silvetti, T., Morandi, S., and Brasca, M. (2012). Identification of Clostridium beijerinckii, Cl. butyricum, Cl. sporogenes, Cl. tyrobutyricum isolated from silage, raw milk and hard cheese by a multiplex PCR assay. *J Dairy Res* 79(3)**,** 318-323. doi: 10.1017/s002202991200026x.

Crosby, L.D., and Criddle, C.S. (2003). Understanding bias in microbial community analysis techniques due to rrn operon copy number heterogeneity. *Biotechniques* 34(4)**,** 790-794, 796, 798 passim.

Cui, J., and Davidson, A.L. (2011). ABC solute importers in bacteria. *Essays Biochem* 50(1)**,** 85-99. doi: 10.1042/bse0500085.

Cui, Q., Lewis, I.A., Hegeman, A.D., Anderson, M.E., Li, J., Schulte, C.F., et al. (2008). Metabolite identification via the Madison Metabolomics Consortium Database. *Nat Biotechnol* 26(2)**,** 162-164. doi: 10.1038/nbt0208-162.

Cwyk, W.M., and Canale-Parola, E. (1979). Treponema succinifaciens sp. nov., an anaerobic spirochete from the swine intestine. *Arch Microbiol* 122(3)**,** 231-239.

Czerkawski, J.W. (1986). An Introduction to Rumen Studies. Pergamon Press, Oxford/New York.

Dahllöf, I. (2002). Molecular community analysis of microbial diversity. *Curr Opin Biotechnol* 13(3)**,** 213-217.

Danielsson, R., Dicksved, J., Sun, L., Gonda, H., Müller, B., Schnürer, A., et al. (2017). Methane Production in Dairy Cows Correlates with Rumen Methanogenic and Bacterial Community Structure. *Front Microbiol* 8**,** 226. doi: 10.3389/fmicb.2017.00226.

Davidson, A.L., Dassa, E., Orelle, C., and Chen, J. (2008). Structure, Function, and Evolution of Bacterial ATP-Binding Cassette Systems. *Microbiol Mol Biol Rev : MMBR* 72(2)**,** 317-364. doi: 10.1128/MMBR.00031-07.

de Castro, A.P., Sartori da Silva, M.R., Quirino, B.F., and Kruger, R.H. (2013). Combining "omics" strategies to analyze the biotechnological potential of complex microbial environments. *Curr Protein Pept Sci* 14(6)**,** 447-458.

de la Cuesta-Zuluaga, J., and Escobar, J.S. (2016). Considerations For Optimizing Microbiome Analysis Using a Marker Gene. *Front Nutr* 3**,** 26. doi: 10.3389/fnut.2016.00026.

de Menezes, A.B., Lewis, E., O'Donovan, M., O'Neill, B.F., Clipson, N., and Doyle, E.M. (2011). Microbiome analysis of dairy cows fed pasture or total mixed ration diets. *FEMS Microbiol Ecol* 78(2)**,** 256-265. doi: 10.1111/j.1574-6941.2011.01151.x.

Dehority, B.A. (2003). Rumen Microbiology. Nottingham University Press.

Dehority, B.A., and Orpin, C.G. (1997). "Development of, and natural fluctuations in, rumen microbial populations," in The Rumen Microbial Ecosystem, eds. P.N. Hobson & C.S. Stewart. (Dordrecht: Springer Netherlands), 196-245.

Denman, S.E., Martinez Fernandez, G., Shinkai, T., Mitsumori, M., and McSweeney, C.S. (2015). Metagenomic analysis of the rumen microbial community following inhibition of methane formation by a halogenated methane analog. *Front Microbiol* 6**,** 1087. doi: 10.3389/fmicb.2015.01087.

Denman, S.E., Tomkins, N.W., and McSweeney, C.S. (2007). Quantitation and diversity analysis of ruminal methanogenic populations in response to the antimethanogenic compound bromochloromethane. *FEMS Microbiol Ecol* 62(3)**,** 313-322. doi: 10.1111/j.1574-6941.2007.00394.x.

DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K., et al. (2006a). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* 72(7)**,** 5069-5072. doi: 10.1128/aem.03006-05.

DeSantis, T.Z., Jr., Hugenholtz, P., Keller, K., Brodie, E.L., Larsen, N., Piceno, Y.M., et al. (2006b). NAST: a multiple sequence alignment server for comparative analysis of 16S rRNA genes. *Nucleic Acids Res* 34(Web Server issue)**,** W394-399. doi: 10.1093/nar/gkl244.

Deusch, S., and Seifert, J. (2015). Catching the tip of the iceberg - evaluation of sample preparation protocols for metaproteomic studies of the rumen microbiota. *Proteomics* 15(20)**,** 3590-3595. doi: 10.1002/pmic.201400556.

Deusch, S., Tilocca, B., Camarinha-Silva, A., and Seifert, J. (2015). News in livestock research - use of *Omics*-technologies to study the microbiota in the gastrointestinal tract of farm animals. *Comput Struct Biotechnol J* 13**,** 55-63. doi: 10.1016/j.csbj.2014.12.005.

Dewhurst, R. (2013). Milk production from silage: comparison of grass, legume and maize silages and their mixtures. *Agr Fodd Sci Finland* (1)**,** 57-69%V 22.

Dijkstra, J., Ellis, J.L., Kebreab, E., Strathe, A.B., López, S., France, J., et al. (2012). Ruminal pH regulation and nutritional consequences of low pH. *Anim Feed Sci Tech* 172(1-2)**,** 22-33. doi: 10.1016/j.anifeedsci.2011.12.005.

Dona, A.C., Kyriakides, M., Scott, F., Shephard, E.A., Varshavi, D., Veselkov, K., et al. (2016). A guide to the identification of metabolites in NMR-based metabonomics/metabolomics experiments. *Comput Struct Biotechnol J* 14**,** 135-153. doi: 10.1016/j.csbj.2016.02.005.

Duskova, D., and Marounek, M. (2001). Fermentation of pectin and glucose, and activity of pectin-degrading enzymes in the rumen bacterium Lachnospira multiparus. *Lett Appl Microbiol* 33(2)**,** 159-163.

Dzieciatkowska, M., Hill, R., and Hansen, K.C. (2014). GeLC-MS/MS analysis of complex protein mixtures. *Methods Mol Biol* 1156**,** 53-66. doi: 10.1007/978-1-4939-0685-7_4.

Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., and Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27(16)**,** 2194-2200. doi: 10.1093/bioinformatics/btr381.

Edwards, J., R McEwan, N., J Travis, A., and John Wallace, R. (2004). 16S rDNA library-based analysis of ruminal bacterial diversity. *Anton Leeuw J Microb* 86(3)**,** 263-281. doi: 10.1023/b:anto.0000047942.69033.24.

Ellison, M.J., Conant, G.C., Cockrum, R.R., Austin, K.J., Truong, H., Becchi, M., et al. (2014). Diet alters both the structure and taxonomy of the ovine gut microbial ecosystem. *DNA Res* 21(2)**,** 115-125. doi: 10.1093/dnares/dst044.

Eng, J.K., McCormack, A.L., and Yates, J.R. (1994). An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J Am Soc Mass Spectrom* 5(11)**,** 976-989. doi: 10.1016/1044-0305(94)80016-2.

Enquist, B.J., Economo, E.P., Huxman, T.E., Allen, A.P., Ignace, D.D., and Gillooly, J.F. (2003). Scaling metabolism from organisms to ecosystems. *Nature* 423(6940)**,** 639-642. doi: 10.1038/nature01671.

Feinstein, L.M., Sul, W.J., and Blackwood, C.B. (2009). Assessment of bias associated with incomplete extraction of microbial DNA from soil. *Appl Environ Microbiol* 75(16)**,** 5428-5433. doi: 10.1128/aem.00120-09.

Felius, M. (1995). Cattle breeds: an encyclopedia. Misset.

Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F., and Whitehouse, C.M. (1989). Electrospray ionization for mass spectrometry of large biomolecules. *Science* 246(4926)**,** 64-71.

Fernando, S.C., Purvis, H.T., 2nd, Najar, F.Z., Sukharnikov, L.O., Krehbiel, C.R., Nagaraja, T.G., et al. (2010). Rumen microbial population dynamics during adaptation to a high-grain diet. *Appl Environ Microbiol* 76(22)**,** 7482-7490. doi: 10.1128/aem.00388-10.

Ferrer, M., Ghazi, A., Beloqui, A., Vieites, J.M., Lopez-Cortes, N., Marin-Navarro, J., et al. (2012). Functional metagenomics unveils a multifunctional glycosyl hydrolase from the family 43 catalysing the breakdown of plant polymers in the calf rumen. *PLoS One* 7(6)**,** e38134. doi: 10.1371/journal.pone.0038134.

Fiehn, O. (2002). Metabolomics--the link between genotypes and phenotypes. *Plant Mol Biol* 48(1-2)**,** 155-171. doi: 10.1023/a:1013713905833.

Flint, H.J., Bayer, E.A., Rincon, M.T., Lamed, R., and White, B.A. (2008). Polysaccharide utilization by gut bacteria: potential for new insights from genomic analysis. *Nat Rev Microbiol* 6(2)**,** 121-131. doi: 10.1038/nrmicro1817.

Fogel, G.B., Collins, C.R., Li, J., and Brunk, C.F. (1999). Prokaryotic Genome Size and SSU rDNA Copy Number: Estimation of Microbial Relative Abundance from a Mixed Population. *Microb Ecol* 38(2)**,** 93-113. doi: 10.1007/s002489900162.

Fonty, G., Gouet, P., Jouany, J.P., and Senaud, J. (1983). Ecological Factors Determining Establishment of Cellulolytic Bacteria and Protozoa in the Rumens of Meroxenic Lambs. *Microbiology* 129(1)**,** 213-223. doi: doi:10.1099/00221287-129-1-213.

Friedrich, M.W. (2005). "Methyl-Coenzyme M Reductase Genes: Unique Functional Markers for Methanogenic and Anaerobic Methane-Oxidizing Archaea," in *Methods Enzymol*. ed. Jared R. L., editor. (Waltham, MA: Academic Press) 428-442.

García-Cañas, V., Cifuentes, A., and Simó, C. (2014). Applications of Advanced Omics Technologies: From Genes to Metabolites. Elsevier Science.

Geer, L.Y., Markey, S.P., Kowalak, J.A., Wagner, L., Xu, M., Maynard, D.M., et al. (2004). Open mass spectrometry search algorithm. *J Proteome Res* 3(5)**,** 958-964. doi: 10.1021/pr0499491.

Gerber, P.J., Steinfeld, H., Henderson, B., Mottet, A., Opio, C., Dijkman, J., Falcucci, A. & Tempio, G. 2013. Tackling climate change through livestock - A global assessment of emissions and mitigation opportunities. Food and Agriculture Organization of the United Nations (FAO), Rome.

Giansanti, P., Tsiatsiani, L., Low, T.Y., and Heck, A.J.R. (2016). Six alternative proteases for mass spectrometry-based proteomics beyond trypsin. *Nat. Protocols* 11(5)**,** 993-1006. doi: 10.1038/nprot.2016.057.

Gihring, T.M., Green, S.J., and Schadt, C.W. (2012). Massively parallel rRNA gene sequencing exacerbates the potential for biased community diversity comparisons due to variable library sizes. *Environ Microbiol* 14(2)**,** 285-290. doi: 10.1111/j.1462-2920.2011.02550.x.

Givens, D.I., and Rulquin, H. (2004). Utilisation by ruminants of nitrogen compounds in silage-based diets. *Anim Feed Sci Tech* 114(1)**,** 1-18. doi: 10.1016/j.anifeedsci.2003.09.005.

Graham, C., McMullan, G., and Graham, R.L. (2011). Proteomics in the microbial sciences. *Bioeng Bugs* 2(1)**,** 17-30. doi: 10.4161/bbug.2.1.14413.

Gressley, T.F., Hall, M.B., and Armentano, L.E. (2011). Ruminant Nutrition Symposium: Productivity, digestion, and health responses to hindgut acidosis in ruminants. *J Anim Sci* 89(4)**,** 1120-1130. doi: 10.2527/jas.2010-3460.

Gruby D., Delafond H. M. O. (1843). Recherches sur des animalcules se développantcaz en grand nombre dans l'estomac et dans les intestins, pendant la digestion des animaux herbivores et carnivores. *Compt Rend Acad Sci* 17, 1304-1308.

Guan, L.L., Nkrumah, J.D., Basarab, J.A., and Moore, S.S. (2008). Linkage of microbial ecology to phenotype: correlation of rumen microbial ecology to cattle's feed efficiency. *FEMS Microbiol Lett* 288(1)**,** 85-91. doi: 10.1111/j.1574-6968.2008.01343.x.

Guthals, A., and Bandeira, N. (2012). Peptide identification by tandem mass spectrometry with alternate fragmentation modes. *Mol Cell Proteomics* 11(9), 550-557. doi: 10.1074/mcp.R112.018556.

Hackmann, T.J. (2015). "Time required for adaptation of rumen fermentation and the rumen microbiome", in: *American Dairy Science Association & American Society of Animal Science, Joint Annual Meeting.* (Orlando, Florida, USA: J. Anim. Sci. Vol. 93, Suppl. s3; J. Dairy Sci. Vol. 98, Suppl. 2).

Hackmann, T.J., and Spain, J.N. (2010). Invited review: ruminant ecology and evolution: perspectives useful to ruminant livestock research and production. *J Dairy Sci* 93(4), 1320-1334. doi: 10.3168/jds.2009-2071.

Hamady, M., Walker, J.J., Harris, J.K., Gold, N.J., and Knight, R. (2008). Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. *Nat Methods* 5(3), 235-237. doi: 10.1038/nmeth.1184.

Hamlin, L.J., and Hungate, R.E. (1956). CULTURE AND PHYSIOLOGY OF A STARCH-DIGESTING BACTERIUM (BACTEROIDES AMYLOPHILUS N. SP.) FROM THE BOVINE RUMEN. *J Bacteriol* 72(4), 548-554.

Hansen, S.H., Stensballe, A., Nielsen, P.H., and Herbst, F.A. (2014). Metaproteomics: Evaluation of protein extraction from activated sludge. *Proteomics* 14(21-22), 2535-2539. doi: 10.1002/pmic.201400167.

Hassanat, F., Gervais, R., Julien, C., Masse, D.I., Lettat, A., Chouinard, P.Y., et al. (2013). 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* 96(7), 4553-4567. doi: 10.3168/jds.2012-6480.

Haug, A., Høstmark, A.T., and Harstad, O.M. (2007). Bovine milk in human nutrition - a review. *Lipids Health Dis* 6, 25-25. doi: 10.1186/1476-511X-6-25.

Henderson, G., Cox, F., Ganesh, S., Jonker, A., Young, W., and Janssen, P.H. (2015). Rumen microbial community composition varies with diet and host, but a core microbiome is found across a wide geographical range. *Sci Rep* 5, 14567. doi: 10.1038/srep14567.

Henderson, G., Cox, F., Kittelmann, S., Miri, V.H., Zethof, M., Noel, S.J., et al. (2013). Effect of DNA extraction methods and sampling techniques on the apparent structure of cow and sheep rumen microbial communities. *PLoS One* 8(9), e74787. doi: 10.1371/journal.pone.0074787.

Hernandez-Sanabria, E., Guan, L.L., Goonewardene, L.A., Li, M., Mujibi, D.F., Stothard, P., et al. (2010). Correlation of Particular Bacterial PCR-Denaturing Gradient Gel Electrophoresis Patterns with Bovine Ruminal Fermentation Parameters and Feed Efficiency Traits. *Appl Environ Microbiol* 76(19), 6338-6350. doi: 10.1128/AEM.01052-10.

Hespell, R.B. (1992). "The Genera Succinivibrio and Succinimonas," in The Prokaryotes: A Handbook on the Biology of Bacteria: Ecophysiology, Isolation, Identification, Applications, eds. A. Balows, H.G. Trüper, M. Dworkin, W. Harder & K.-H. Schleifer. (New York, NY: Springer New York), 3979-3982.

Hess, M., Sczyrba, A., Egan, R., Kim, T.W., Chokhawala, H., Schroth, G., et al. (2011). Metagenomic discovery of biomass-degrading genes and genomes from cow rumen. *Science* 331(6016), 463-467. doi: 10.1126/science.1200387.

Hettich, R.L., Pan, C., Chourey, K., and Giannone, R.J. (2013). Metaproteomics: harnessing the power of high performance mass spectrometry to identify the suite of proteins that control metabolic activities in microbial communities. *Anal Chem* 85(9), 4203-4214. doi: 10.1021/ac303053e.

Heyer, R., Kohrs, F., Benndorf, D., Rapp, E., Kausmann, R., Heiermann, M., et al. (2013). Metaproteome analysis of the microbial communities in agricultural biogas plants. *N Biotechnol* 30(6), 614-622. doi: 10.1016/j.nbt.2013.01.002.

Hong, S., Bunge, J., Leslin, C., Jeon, S., and Epstein, S.S. (2009). Polymerase chain reaction primers miss half of rRNA microbial diversity. *ISME J* 3(12), 1365-1373. doi: 10.1038/ismej.2009.89.

Hook, S.E., Steele, M.A., Northwood, K.S., Dijkstra, J., France, J., Wright, A.D., et al. (2011a). Impact of subacute ruminal acidosis (SARA) adaptation and recovery on the density and diversity of bacteria in the rumen of dairy cows. *FEMS Microbiol Ecol* 78(2), 275-284. doi: 10.1111/j.1574-6941.2011.01154.x.

Hook, S.E., Steele, M.A., Northwood, K.S., Wright, A.D., and McBride, B.W. (2011b). Impact of high-concentrate feeding and low ruminal pH on methanogens and protozoa in the rumen of dairy cows. *Microb Ecol* 62(1), 94-105. doi: 10.1007/s00248-011-9881-0.

Huber, J.A., Morrison, H.G., Huse, S.M., Neal, P.R., Sogin, M.L., and Mark Welch, D.B. (2009). Effect of PCR amplicon size on assessments of clone library microbial diversity and community structure. *Environ Microbiol* 11(5), 1292-1302. doi: 10.1111/j.1462-2920.2008.01857.x.

Hungate, R.E. (1947). Studies on Cellulose Fermentation: III. The Culture and Isolation for Cellulose-decomposing Bacteria from the Rumen of Cattle. *J Bacteriol* 53(5), 631-645.

Hungate, R.E. (1950). The Anaerobic mesophilic cellulolytic bacteria. *Bacteriol Rev* 14(1), 1-49.

Hungate, R.E. (1966). The rumen and its microbes. New York, London: Academic Press.

Hungate, R.E. (1975). The Rumen Microbial Ecosystem. *Annual Review of Ecology and Systematics* 6, 39-66.

Hungate, R.E., Smith, W., Bauchop, T., Yu, I., and Rabinowitz, J.C. (1970). Formate as an Intermediate in the Bovine Rumen Fermentation. *J Bacteriol* 102(2), 389-397.

Ishii, K., and Fukui, M. (2001). Optimization of Annealing Temperature To Reduce Bias Caused by a Primer Mismatch in Multitemplate PCR. *Appl Environ Microbiol* 67(8), 3753-3755. doi: 10.1128/AEM.67.8.3753-3755.2001.

Jagtap, P., Goslinga, J., Kooren, J.A., McGowan, T., Wroblewski, M.S., Seymour, S.L., et al. (2013). A two-step database search method improves sensitivity in peptide sequence matches for metaproteomics and proteogenomics studies. *Proteomics* 13(8), 1352-1357. doi: 10.1002/pmic.201200352.

Jami, E., Israel, A., Kotser, A., and Mizrahi, I. (2013). Exploring the bovine rumen bacterial community from birth to adulthood. *ISME J* 7(6), 1069-1079. doi: 10.1038/ismej.2013.2.

Jami, E., and Mizrahi, I. (2012). Composition and Similarity of Bovine Rumen Microbiota across Individual Animals. *PLoS One* 7(3), e33306. doi: 10.1371/journal.pone.0033306.

Jami, E., White, B.A., and Mizrahi, I. (2014). Potential Role of the Bovine Rumen Microbiome in Modulating Milk Composition and Feed Efficiency. *PLoS One* 9(1), e85423-Article No.: e85423. doi: 10.1371/journal.pone.0085423.

Janssen, P.H., and Kirs, M. (2008). Structure of the archaeal community of the rumen. *Appl Environ Microbiol* 74(12), 3619-3625. doi: 10.1128/aem.02812-07.

Jaun, B., and Thauer, R.K. (2007). "Methyl-Coenzyme M Reductase and its Nickel Corphin Coenzyme F430 in Methanogenic Archaea," in Nickel and Its Surprising Impact in Nature. John Wiley & Sons, Ltd), 323-356.

Jehmlich, N., Schmidt, F., Hartwich, M., von Bergen, M., Richnow, H.H., and Vogt, C. (2008). Incorporation of carbon and nitrogen atoms into proteins measured by protein-based stable isotope probing (Protein-SIP). *Rapid Commun Mass Spectrom* 22(18), 2889-2897. doi: 10.1002/rcm.3684.

Jensen, L.J., Julien, P., Kuhn, M., von Mering, C., Muller, J., Doerks, T., et al. (2008). eggNOG: automated construction and annotation of orthologous groups of genes. *Nucleic Acids Res* 36(Database issue), D250-254. doi: 10.1093/nar/gkm796.

Jin, D., Zhao, S., Zheng, N., Bu, D., Beckers, Y., Denman, S.E., et al. (2017). Differences in Ureolytic Bacterial Composition between the Rumen Digesta and Rumen Wall Based on ureC Gene Classification. *Front Microbiol* 8(385). doi: 10.3389/fmicb.2017.00385.

Johnson, D.E., and Ward, G.M. (1996). Estimates of animal methane emissions. *Environ Monit Assess* 42(1-2), 133-141. doi: 10.1007/bf00394046.

Jones, P.M., and George, A.M. (2004). The ABC transporter structure and mechanism: perspectives on recent research. *Cell Mol Life Sci CMLS* 61(6), 682-699. doi: 10.1007/s00018-003-3336-9.

Jose, V.L., Appoothy, T., More, R.P., and Arun, A.S. (2017). Metagenomic insights into the rumen microbial fibrolytic enzymes in Indian crossbred cattle fed finger millet straw. *AMB Express* 7, 13. doi: 10.1186/s13568-016-0310-0.

Jouany, J.P. (1996). Effect of rumen protozoa on nitrogen utilization by ruminants. *J Nutr* 126(4 Suppl)**,** 1335s-1346s.

Jovel, J., Patterson, J., Wang, W., Hotte, N., O'Keefe, S., Mitchel, T., et al. (2016). Characterization of the Gut Microbiome Using 16S or Shotgun Metagenomics. *Front Microbiol* 7**,** 459. doi: 10.3389/fmicb.2016.00459.

Jumas-Bilak, E., and Marchandin, H. (2014). "The Phylum Synergistetes," in The Prokaryotes: Other Major Lineages of Bacteria and The Archaea, eds. E. Rosenberg, E.F. DeLong, S. Lory, E. Stackebrandt & F. Thompson. (Berlin, Heidelberg: Springer Berlin Heidelberg), 931-954.

Juniper, D.T., Browne, E.M., Bryant, M.J., and Beever, D.E. (2008). Digestion, rumen fermentation and circulating concentrations of insulin, growth hormone and IGF-1 in steers fed diets based on different proportions of maize silage and grass silage. *Animal* 2(6)**,** 849-858. doi: 10.1017/s1751731108001857.

Juste, A., Thomma, B.P., and Lievens, B. (2008). Recent advances in molecular techniques to study microbial communities in food-associated matrices and processes. *Food Microbiol* 25(6)**,** 745-761. doi: 10.1016/j.fm.2008.04.009.

Kalle, E., Gulevich, A., and Rensing, C. (2013). External and semi-internal controls for PCR amplification of homologous sequences in mixed templates. *J Microbiol Methods* 95(2)**,** 285-294. doi: 10.1016/j.mimet.2013.09.014.

Kanehisa, M., and Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res* 28(1)**,** 27-30.

Kang, S.H., Evans, P., Morrison, M., and McSweeney, C. (2013). Identification of metabolically active proteobacterial and archaeal communities in the rumen by DNA- and RNA-derived 16S rRNA gene. *J Appl Microbiol* 115(3)**,** 644-653. doi: 10.1111/jam.12270.

Kembel, S.W., Wu, M., Eisen, J.A., and Green, J.L. (2012). Incorporating 16S gene copy number information improves estimates of microbial diversity and abundance. *PLoS Comput Biol* 8(10)**,** e1002743. doi: 10.1371/journal.pcbi.1002743.

Khafipour, E., Li, S., Plaizier, J.C., and Krause, D.O. (2009). Rumen Microbiome Composition Determined Using Two Nutritional Models of Subacute Ruminal Acidosis. *Appl Environ Microbiol* 75(22)**,** 7115-7124. doi: 10.1128/AEM.00739-09.

Kim, M., Morrison, M., and Yu, Z. (2011a). Phylogenetic diversity of bacterial communities in bovine rumen as affected by diets and microenvironments. *Folia Microbiol* 56(5)**,** 453. doi: 10.1007/s12223-011-0066-5.

Kim, M., Morrison, M., and Yu, Z. (2011b). Status of the phylogenetic diversity census of ruminal microbiomes. *FEMS Microbiol Ecol* 76(1)**,** 49-63. doi: 10.1111/j.1574-6941.2010.01029.x.

Kim, M., Wang, L., Morrison, M., and Yu, Z. (2014). Development of a phylogenetic microarray for comprehensive analysis of ruminal bacterial communities. *J Appl Microbiol* 117(4)**,** 949-960. doi: 10.1111/jam.12598.

King, E.E., Smith, R.P., St-Pierre, B., and Wright, A.-D.G. (2011). Differences in the Rumen Methanogen Populations of Lactating Jersey and Holstein Dairy Cows under the Same Diet Regimen. *Appl Environ Microbiol* 77(16)**,** 5682-5687. doi: 10.1128/AEM.05130-11.

Kirchgeßner, M., Stangl, G., Schwarz, F.J., Roth, F.X., Südekum, K.H., and Eder, K. (2014). *Tierernährung.* DLG-Verlag GmbH.

Klieve, A.V., and Bauchop, T. (1988). Morphological diversity of ruminal bacteriophages from sheep and cattle. *Appl Environ Microbiol* 54(6)**,** 1637-1641.

Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., et al. (2013). Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res* 41(1)**,** e1. doi: 10.1093/nar/gks808.

Koenig, K.M., Newbold, C.J., McIntosh, F.M., and Rode, L.M. (2000). Effects of protozoa on bacterial nitrogen recycling in the rumen. *J Anim Sci* 78(9)**,** 2431-2445.

Koh, A., De Vadder, F., Kovatcheva-Datchary, P., and Bäckhed, F. (2016). From Dietary Fiber to Host Physiology: Short-Chain Fatty Acids as Key Bacterial Metabolites. *Cell* 165(6)**,** 1332-1345. doi: 10.1016/j.cell.2016.05.041.

Koike, S., Yoshitani, S., Kobayashi, Y., and Tanaka, K. (2003). Phylogenetic analysis of fiber-associated rumen bacterial community and PCR detection of uncultured bacteria. *FEMS Microbiol Lett* 229(1)**,** 23-30.

Kojima, S., Ko, K.C., Takatsuka, Y., Abe, N., Kaneko, J., Itoh, Y., et al. (2010). Cadaverine covalently linked to peptidoglycan is required for interaction between the peptidoglycan and the periplasm-exposed S-layer-homologous domain of major outer membrane protein Mep45 in Selenomonas ruminantium. *J Bacteriol* 192(22)**,** 5953-5961. doi: 10.1128/jb.00417-10.

Kong, Y., Teather, R., and Forster, R. (2010). Composition, spatial distribution, and diversity of the bacterial communities in the rumen of cows fed different forages. *FEMS Microbiol Ecol* 74(3)**,** 612-622. doi: 10.1111/j.1574-6941.2010.00977.x.

Kozich, J.J., Westcott, S.L., Baxter, N.T., Highlander, S.K., and Schloss, P.D. (2013). Development of a Dual-Index Sequencing Strategy and Curation Pipeline for Analyzing Amplicon Sequence Data on the MiSeq Illumina Sequencing Platform. *Appl Environ Microbiol* 79(17)**,** 5112-5120. doi: 10.1128/AEM.01043-13.

Kumar, P.S., Brooker, M.R., Dowd, S.E., and Camerlengo, T. (2011). Target region selection is a critical determinant of community fingerprints generated by 16S pyrosequencing. *PLoS One* 6(6)**,** e20956. doi: 10.1371/journal.pone.0020956.

Kung, L., Jr., and Hession, A.O. (1995). Preventing in vitro lactate accumulation in ruminal fermentations by inoculation with Megasphaera elsdenii. *J Anim Sci* 73(1)**,** 250-256.

La Torre, A., Bassi, D., Zotta, T., Orru, L., Lamontanara, A., and Cocconcelli, P.S. (2016). Draft Genome Sequence of Clostridium sporogenes Strain UC9000 Isolated from Raw Milk. *Genome Announc* 4(2). doi: 10.1128/genomeA.00244-16.

Lamendella, R., VerBerkmoes, N., and Jansson, J.K. (2012). 'Omics' of the mammalian gut—new insights into function. *Curr Opin Biotechnol* 23. doi: 10.1016/j.copbio.2012.01.016.

Lange, M., Westermann, P., and Ahring, B.K. (2005). Archaea in protozoa and metazoa. *Appl Microbiol Biotechnol* 66(5)**,** 465-474. doi: 10.1007/s00253-004-1790-4.

Langille, M.G.I., Zaneveld, J., Caporaso, J.G., McDonald, D., Knights, D., Reyes, J.A., et al. (2013). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat Biotech* 31(9)**,** 814-821. doi: 10.1038/nbt.2676.

Lapierre, H., and Lobley, G.E. (2001). Nitrogen Recycling in the Ruminant: A Review. *J Dairy Sci* 84**,** E223-E236. doi: 10.3168/jds.S0022-0302(01)70222-6.

Larue, R., Yu, Z., Parisi, V.A., Egan, A.R., and Morrison, M. (2005). Novel microbial diversity adherent to plant biomass in the herbivore gastrointestinal tract, as revealed by ribosomal intergenic spacer analysis and rrs gene sequencing. *Environ Microbiol* 7(4)**,** 530-543. doi: 10.1111/j.1462-2920.2005.00721.x.

Latham, M.J., and Wolin, M.J. (1977). Fermentation of cellulose by Ruminococcus flavefaciens in the presence and absence of Methanobacterium ruminantium. *Appl Environ Microbiol* 34(3)**,** 297-301.

Leahy, S.C., Kelly, W.J., Ronimus, R.S., Wedlock, N., Altermann, E., and Attwood, G.T. (2013). Genome sequencing of rumen bacteria and archaea and its application to methane mitigation strategies. *Animal* 7 Suppl 2**,** 235-243. doi: 10.1017/s1751731113000700.

Lee, H.J., Jung, J.Y., Oh, Y.K., Lee, S.-S., Madsen, E.L., and Jeon, C.O. (2012). Comparative Survey of Rumen Microbial Communities and Metabolites across One Caprine and Three Bovine Groups, Using Bar-Coded Pyrosequencing and (1)H Nuclear Magnetic Resonance Spectroscopy. *Appl Environ Microbiol* 78(17)**,** 5983-5993. doi: 10.1128/AEM.00104-12.

Leedle, J.Z., Barsuhn, K., and Butine, T. (1987). Comparison of treatments for the in vitro detachment of cellulolytic bacteria from ruminal digesta samples. *Curr Microbiol* 15(3)**,** 129-131. doi: 10.1007/BF01577259.

Legay-Carmier, F., and Bauchart, D. (1989). Distribution of bacteria in the rumen contents of dairy cows given a diet supplemented with soya-bean oil. *Br J Nutr* 61(3)**,** 725-740.

Lengowski, M.B., Witzig, M., Möhring, J., Seyfang, G.M., and Rodehutscord, M. (2016). Effects of corn silage and grass silage in ruminant rations on diurnal changes of microbial populations in the rumen of dairy cows. *Anaerobe* 42**,** 6-16. doi: 10.1016/j.anaerobe.2016.07.004.

Lerma, C., Rich, P.J., Ju, G.C., Yang, W.-J., Hanson, A.D., and Rhodes, D. (1991). Betaine Deficiency in Maize: Complementation Tests and Metabolic Basis. *Plant Physiol* 95(4)**,** 1113-1119. doi: 10.1104/pp.95.4.1113.

Ley, R.E., Hamady, M., Lozupone, C., Turnbaugh, P.J., Ramey, R.R., Bircher, J.S., et al. (2008). Evolution of mammals and their gut microbes. *Science* 320(5883)**,** 1647-1651. doi: 10.1126/science.1155725.

Ley, R.E., Turnbaugh, P.J., Klein, S., and Gordon, J.I. (2006). Microbial ecology: human gut microbes associated with obesity. *Nature* 444. doi: 10.1038/4441022a.

Li, F., and Guan, L.L. (2017). Metatranscriptomic Profiling Reveals Linkages between the Active Rumen Microbiome and Feed Efficiency in Beef Cattle. *Appl Environ Microbiol* 83(9). doi: 10.1128/aem.00061-17.

Li, F., Henderson, G., Sun, X., Cox, F., Janssen, P.H., and Guan, L.L. (2016). Taxonomic Assessment of Rumen Microbiota Using Total RNA and Targeted Amplicon Sequencing Approaches. *Front Microbiol* 7(987). doi: 10.3389/fmicb.2016.00987.

Li, M., Penner, G.B., Hernandez-Sanabria, E., Oba, M., and Guan, L.L. (2009). Effects of sampling location and time, and host animal on assessment of bacterial diversity and fermentation parameters in the bovine rumen. *J Appl Microbiol* 107(6)**,** 1924-1934. doi: 10.1111/j.1365-2672.2009.04376.x.

Li, R.W., Wu, S., Baldwin, R.L.t., Li, W., and Li, C. (2012). Perturbation dynamics of the rumen microbiota in response to exogenous butyrate. *PLoS One* 7(1)**,** e29392. doi: 10.1371/journal.pone.0029392.

Linehan, B., Scheifinger, C.C., and Wolin, M.J. (1978). Nutritional Requirements of Selenomonas ruminantium for Growth on Lactate, Glycerol, or Glucose. *Appl Environ Microbiol* 35(2)**,** 317-322.

Liu, Y., and Whitman, W.B. (2008). Metabolic, phylogenetic, and ecological diversity of the methanogenic archaea. *Ann N Y Acad Sci* 1125**,** 171-189. doi: 10.1196/annals.1419.019.

Liu, Z., DeSantis, T.Z., Andersen, G.L., and Knight, R. (2008). Accurate taxonomy assignments from 16S rRNA sequences produced by highly parallel pyrosequencers. *Nucleic Acids Res* 36(18)**,** e120-e120. doi: 10.1093/nar/gkn491.

Liu, Z., Lozupone, C., Hamady, M., Bushman, F.D., and Knight, R. (2007). Short pyrosequencing reads suffice for accurate microbial community analysis. *Nucleic Acids Res* 35(18)**,** e120. doi: 10.1093/nar/gkm541.

Livingstone, K.M., Humphries, D.J., Kirton, P., Kliem, K.E., Givens, D.I., and Reynolds, C.K. (2015). Effects of forage type and extruded linseed supplementation on methane production and milk fatty acid composition of lactating dairy cows. *J Dairy Sci* 98(6)**,** 4000-4011. doi: 10.3168/jds.2014-8987.

Lodge-Ivey, S.L., Browne-Silva, J., and Horvath, M.B. (2009). Technical note: Bacterial diversity and fermentation end products in rumen fluid samples collected via oral lavage or rumen cannula. *J Anim Sci* 87(7)**,** 2333-2337. doi: 10.2527/jas.2008-1472.

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., and Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* 42(D1)**,** D490-D495. doi: 10.1093/nar/gkt1178.

Mackie, R.I. (2000). Molecular ecology and diversity in gut microbial ecosystems. Ruminant Physiology: Digestion, Metabolism, Growth and Reproduction 2000:61-77.

Mackie, R.I., and Cann, I.K.O. (2005). A review of gastrointestinal microbiology with special emphasis on molecular microbial ecology approaches. Springer, Po Box 17, 3300 Aa Dordrecht, Netherlands.

Mackie, R.I., McSweeney, C.S., and Aminov, R.I. (2001). "Rumen," in *eLS*. John Wiley & Sons, Ltd.

Maeder, D.L., Anderson, I., Brettin, T.S., Bruce, D.C., Gilna, P., Han, C.S., et al. (2006). The Methanosarcina barkeri genome: comparative analysis with Methanosarcina acetivorans and Methanosarcina mazei reveals extensive rearrangement within methanosarcinal genomes. *J Bacteriol* 188(22)**,** 7922-7931. doi: 10.1128/jb.00810-06.

Malmuthuge, N., Griebel, P.J., and Guan, L.L. (2015). The Gut Microbiome and Its Potential Role in the Development and Function of Newborn Calf Gastrointestinal Tract. *Front Vet Sci* 2(36). doi: 10.3389/fvets.2015.00036.

Mann, M. (2016). Origins of mass spectrometry-based proteomics. *Nat Rev Mol Cell Biol* 17(11)**,** 678-678. doi: 10.1038/nrm.2016.135.

Mann, M., Hendrickson, R.C., and Pandey, A. (2001). Analysis of proteins and proteomes by mass spectrometry. *Annu Rev Biochem* 70**,** 437-473. doi: 10.1146/annurev.biochem.70.1.437.

Mao, S.Y., Zhang, R.Y., Wang, D.S., and Zhu, W.Y. (2013). Impact of subacute ruminal acidosis (SARA) adaptation on rumen microbiota in dairy cattle using pyrosequencing. *Anaerobe* 24**,** 12-19. doi: 10.1016/j.anaerobe.2013.08.003.

Markley, J.L., Brüschweiler, R., Edison, A.S., Eghbalnia, H.R., Powers, R., Raftery, D., et al. (2017). The future of NMR-based metabolomics. *Curr Opin Biotech* 43**,** 34-40. doi: 10.1016/j.copbio.2016.08.001.

Matsui, H., Ogata, K., Tajima, K., Nakamura, M., Nagamine, T., Aminov, R.I., et al. (2000). Phenotypic Characterization of Polysaccharidases Produced by Four Prevotella Type Strains. *Curr Microbiol* 41(1)**,** 45-49. doi: 10.1007/s002840010089.

McAllister, T.A., Bae, H.D., Jones, G.A., and Cheng, K.J. (1994). Microbial attachment and feed digestion in the rumen. *J Anim Sci* 72(11)**,** 3004-3018.

McAllister, T.A., and Newbold, C.J. (2008). Redirecting rumen fermentation to reduce methanogenesis. *Aust J Exp Agr* 48(2)**,** 7-13. doi: 10.1071/EA07218.

McCabe, M.S., Cormican, P., Keogh, K., O'Connor, A., O'Hara, E., Palladino, R.A., et al. (2015). Illumina MiSeq Phylogenetic Amplicon Sequencing Shows a Large Reduction of an Uncharacterised Succinivibrionaceae and an Increase of the Methanobrevibacter gottschalkii Clade in Feed Restricted Cattle. *PLoS One* 10(7)**,** e0133234. doi: 10.1371/journal.pone.0133234.

McDermott, J.J., Randolph, T.F., and Staal, S.J. (1999). The economics of optimal health and productivity in smallholder livestock systems in developing countries. *Rev Sci Tech* 18(2)**,** 399-424.

McInerney, M.J., Mackie, R.I., and Bryant, M.P. (1981). Syntrophic association of a butyrate-degrading bacterium and methanosarcina enriched from bovine rumen fluid. *Appl Environ Microbiol* 41(3)**,** 826-828.

McMichael, A.J., Powles, J.W., Butler, C.D., and Uauy, R. (2007). Food, livestock production, energy, climate change, and health. *The Lancet* 370(9594)**,** 1253-1263. doi: 0.1016/S0140-6736(07)61256-2.

McSweeney, C.S., Denman, S.E. and Mackie, R.I. (2005). Rumen bacteria. In Methods in gut microbial ecology for ruminants (pp. 23-37). Springer Netherlands.

Mercier, P., Lewis, M.J., Chang, D., Baker, D., and Wishart, D.S. (2011). Towards automatic metabolomic profiling of high-resolution one-dimensional proton NMR spectra. *J Biomol NMR* 49(3-4)**,** 307-323. doi: 10.1007/s10858-011-9480-x.

Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E.M., Kubal, M., et al. (2008). The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* 9**,** 386. doi: 10.1186/1471-2105-9-386.

Meyer, M., and Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb Protoc* 2010(6)**,** pdb.prot5448. doi: 10.1101/pdb.prot5448.

Miettinen, H., and Huhtanen, P. (1996). Effects of the Ratio of Ruminal Propionate to Butyrate on Milk Yield and Blood Metabolites in Dairy Cows. *J Dairy Sci* 79(5)**,** 851-861. doi: 10.3168/jds.S0022-0302(96)76434-2.

Minato, H., Endo, A., Ootomo, Y., and Uemura, T. (1966). ECOLOGICAL TREATISE ON THE RUMEN FERMENTATION: The amylolytic and cellulolytic activities of the fractionated bacterial portions attached to rumen solids. *J Gen Appl Microbiol* 12(1)**,** 53-69. doi: 10.2323/jgam.12.53.

Minato, H., OTSUKA, M., SHIRASAKA, S., ITABASHI, H., and MITSUMORI, M. (1992). Colonization of microorganisms in the rumen of young calves. *J Gen Appl Microbiol* 38(5)**,** 447-456.

Miron, J., Ben-Ghedalia, D., and Morrison, M. (2001). Invited review: adhesion mechanisms of rumen cellulolytic bacteria. *J Dairy Sci* 84(6)**,** 1294-1309. doi: 10.3168/jds.S0022-0302(01)70159-2.

Mitchell, A.D., Chappell, A., and Knox, K.L. (1979). Metabolism of betaine in the ruminant. *J Anim Sci* 49(3)**,** 764-774.

Miyazaki, K., Martin, J.C., Marinsek-Logar, R., and Flint, H.J. (1997). Degradation and Utilization of Xylans by the Rumen AnaerobePrevotella bryantii(formerlyP. ruminicolasubsp.brevis) B14. *Anaerobe* 3(6)**,** 373-381. doi: 10.1006/anae.1997.0125.

Mizrahi, I. (2011). "The Role of the Rumen Microbiota in Determining the Feed Efficiency of Dairy Cows," in Beneficial Microorganisms in Multicellular Life Forms, eds. E. Rosenberg & U. Gophna. (Berlin, Heidelberg: Springer Berlin Heidelberg), 203-210.

Monteils, V., Rey, M., Silberberg, M., Cauquil, L., and Combes, S. (2012). Modification of activities of the ruminal ecosystem and its bacterial and protozoan composition during repeated dietary changes in cows. *J Anim Sci* 90(12)**,** 4431-4440. doi: 10.2527/jas.2011-4321.

Moran, J. (2005). Tropical Dairy Farming: Feeding Management for Small Holder Dairy Farmers in the Humid Tropics. Land Links.

Morgan, X.C., and Huttenhower, C. (2014). Meta'*omic* analytic techniques for studying the intestinal microbiome. *Gastroenterology* 146(6)**,** 1437-1448.e1431. doi: 10.1053/j.gastro.2014.01.049.

Morgante, M., Stelletta, C., Berzaghi, P., Gianesella, M., and Andrighetto, I. (2007). Subacute rumen acidosis in lactating cows: an investigation in intensive Italian dairy herds. *J Anim Physiol Anim Nutr (Berl)* 91(5-6)**,** 226-234. doi: 10.1111/j.1439-0396.2007.00696.x.

Morgavi, D.P., Forano, E., Martin, C., and Newbold, C.J. (2010). Microbial ecosystem and methanogenesis in ruminants. *Animal* 4(7)**,** 1024-1036. doi: 10.1017/s1751731110000546.

Morita, H., Shiratori, C., Murakami, M., Takami, H., Toh, H., Kato, Y., et al. (2008). Sharpea azabuensis gen. nov., sp. nov., a Gram-positive, strictly anaerobic bacterium isolated from the faeces of thoroughbred horses. *Int J Syst Evol Micr* 58(12)**,** 2682-2686. doi: 10.1099/ijs.0.65543-0.

Muth, T., Kolmeder, C.A., Salojarvi, J., Keskitalo, S., Varjosalo, M., Verdam, F.J., et al. (2015). Navigating through metaproteomics data: a logbook of database searching. *Proteomics* 15. doi: 10.1002/pmic.201400560.

Muth, T., Renard, B.Y., and Martens, L. (2016). Metaproteomic data analysis at a glance: advances in computational microbial community proteomics. *Expert Rev Proteomics* 13(8)**,** 757-769. doi: 10.1080/14789450.2016.1209418.

Nagaraja, T.G., and Titgemeyer, E.C. (2007). Ruminal acidosis in beef cattle: the current microbiological and nutritional outlook. *J Dairy Sci* 90 Suppl 1**,** E17-38. doi: 10.3168/jds.2006-478.

Neill, A.R., Grime, D.W., and Dawson, R.M. (1978). Conversion of choline methyl groups through trimethylamine into methane in the rumen. *Biochem J* 170(3)**,** 529-535.

Newbold, C.J., de la Fuente, G., Belanche, A., Ramos-Morales, E., and McEwan, N.R. (2015). The Role of Ciliate Protozoa in the Rumen. *Front Microbiol* 6**,** 1313. doi: 10.3389/fmicb.2015.01313.

Nielsen, T.S., Straarup, E.M., Vestergaard, M., and Sejrsen, K. (2006). Effect of silage type and concentrate level on conjugated linoleic acids, trans-C18:1 isomers and fat content in milk from dairy cows. *Reprod Nutr Dev* 46(6)**,** 699-712. doi: 10.1051/rnd:2006044.

Nocker, A., Burr, M., and Camper, A.K. (2007). Genotypic Microbial Community Profiling: A Critical Technical Review. *Microb Ecol* 54(2)**,** 276-289. doi: 10.1007/s00248-006-9199-5.

O'Herrin, S.M., and Kenealy, W.R. (1993). Glucose and carbon dioxide metabolism by Succinivibrio dextrinosolvens. *Appl Environ Microbiol* 59(3)**,** 748-755.

Ogata, K., Aminov, R.I., Nagamine, T., Sugiura, M., Tajima, K., Mitsumori, M., et al. (1997). Construction of a Fibrobacter succinogenes genomic map and demonstration of diversity at the genomic level. *Curr Microbiol* 35(1)**,** 22-27.

Oren, A. (2014a). "The Family Methanomicrobiaceae," in The Prokaryotes: Other Major Lineages of Bacteria and The Archaea, eds. E. Rosenberg, E.F. DeLong, S. Lory, E. Stackebrandt & F. Thompson. (Berlin, Heidelberg: Springer Berlin Heidelberg), 231-246.

Oren, A. (2014b). "The Family Methanospirillaceae," in The Prokaryotes: Other Major Lineages of Bacteria and The Archaea, eds. E. Rosenberg, E.F. DeLong, S. Lory, E. Stackebrandt & F. Thompson. (Berlin, Heidelberg: Springer Berlin Heidelberg), 283-290.

Orpin, C.G. (1975). Studies on the rumen flagellate Neocallimastix frontalis. *J Gen Microbiol* 91(2)**,** 249-262. doi: 10.1099/00221287-91-2-249.

Orpin, C.G. (1976). Studies on the rumen flagellate Sphaeromonas communis. *J Gen Microbiol* 94(2)**,** 270-280. doi: 10.1099/00221287-94-2-270.

Oxford, A.E. (1951). The conversion of certain soluble sugars to a glucosan by holotrich ciliates in the rumen of sheep. *J Gen Microbiol* 5(1)**,** 83-90. doi: 10.1099/00221287-5-1-83.

Palleroni, N.J. (1997). Prokaryotic diversity and the importance of culturing. *Anton Leeuw J Microb* 72(1)**,** 3-19. doi: 10.1023/a:1000394109961.

Park, B.H., Karpinets, T.V., Syed, M.H., Leuze, M.R., and Uberbacher, E.C. (2010). CAZymes Analysis Toolkit (CAT): web service for searching and analyzing carbohydrate-active enzymes in a newly sequenced organism using CAZy database. *Glycobiology* 20(12)**,** 1574-1584. doi: 10.1093/glycob/cwq106.

Patel, D.D., Patel, A.K., Parmar, N.R., Shah, T.M., Patel, J.B., and Pandya, P.R. (2014a). Microbial and Carbohydrate Active Enzyme profile of buffalo rumen metagenome and their alteration in response to variation in the diet. *Gene* 545. doi: 10.1016/j.gene.2014.05.003.

Patel, V., Patel, A.K., Parmar, N.R., Patel, A.B., Reddy, B., and Joshi, C.G. (2014b). Characterization of the rumen microbiome of Indian Kankrej cattle (Bos indicus) adapted to different forage diet. *Appl Microbiol Biotechnol* 98. doi: 10.1007/s00253-014-6153-1.

Patra, A.K., and Yu, Z. (2015). Essential oils affect populations of some rumen bacteria in vitro as revealed by microarray (RumenBactArray) analysis. *Front Microbiol* 6**,** 297. doi: 10.3389/fmicb.2015.00297.

Pauly, M., and Keegstra, K. (2008). Cell-wall carbohydrates and their modification as a resource for biofuels. *Plant J* 54(4)**,** 559-568. doi: 10.1111/j.1365-313X.2008.03463.x.

Paz, H.A., Anderson, C.L., Muller, M.J., Kononoff, P.J., and Fernando, S.C. (2016). Rumen Bacterial Community Composition in Holstein and Jersey Cows Is Different under Same Dietary Condition and Is Not Affected by Sampling Method. *Front Microbiol* 7**,** 1206. doi: 10.3389/fmicb.2016.01206.

Perkins, D.N., Pappin, D.J., Creasy, D.M., and Cottrell, J.S. (1999). Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 20(18)**,** 3551-3567. doi: 10.1002/(sici)1522-2683(19991201)20:18<3551::aid-elps3551>3.0.co;2-2.

Peterka, M., and Avgustin, G. (2001). Ribosomal RNA genes from Prevotella bryantii: organization and heterogeneity. *Folia Microbiol (Praha)* 46(1)**,** 67-70.

Petri, R.M., Forster, R.J., Yang, W., McKinnon, J.J., and McAllister, T.A. (2012). Characterization of rumen bacterial diversity and fermentation parameters in concentrate fed cattle with and without forage. *J Appl Microbiol* 112(6)**,** 1152-1162. doi: 10.1111/j.1365-2672.2012.05295.x.

Petri, R.M., Schwaiger, T., Penner, G.B., Beauchemin, K.A., Forster, R.J., McKinnon, J.J., et al. (2013). Characterization of the core rumen microbiome in cattle during transition from forage to concentrate as well as during and after an acidotic challenge. *PLoS One* 8(12)**,** e83424. doi: 10.1371/journal.pone.0083424.

Phipps, R.H., Sutton, J.D., and Jones, B.A. (2010). Forage mixtures for dairy cows: the effect on dry-matter intake and milk production of incorporating either fermented or urea-treated whole-crop wheat, brewers' grains, fodder beet or maize silage into diets based on grass silage. *J Anim Sci* 61(3)**,** 491-496. doi: 10.1017/S1357729800014053.

Pickering, N.K., Chagunda, M.G., Banos, G., Mrode, R., McEwan, J.C., and Wall, E. (2015). Genetic parameters for predicted methane production and laser methane detector measurements. *J Anim Sci* 93(1)**,** 11-20. doi: 10.2527/jas.2014-8302.

Pinto, A.J., and Raskin, L. (2012). PCR biases distort bacterial and archaeal community structure in pyrosequencing datasets. *PLoS One* 7(8)**,** e43093. doi: 10.1371/journal.pone.0043093.

Pitta, D.W., Pinchak, E., Dowd, S.E., Osterstock, J., Gontcharova, V., Youn, E., et al. (2010). Rumen bacterial diversity dynamics associated with changing from bermudagrass hay to grazed winter wheat diets. *Microb Ecol* 59(3)**,** 511-522. doi: 10.1007/s00248-009-9609-6.

Plaizier, J.C., Khafipour, E., Li, S., Gozho, G.N., and Krause, D.O. (2012). Subacute ruminal acidosis (SARA), endotoxins and health consequences. *Anim Feed Sci Tech* 172(1-2)**,** 9-21. doi: 10.1016/j.anifeedsci.2011.12.004.

Pokusaeva, K., Fitzgerald, G.F., and van Sinderen, D. (2011). Carbohydrate metabolism in Bifidobacteria. *Genes Nutr* 6(3)**,** 285-306. doi: 10.1007/s12263-010-0206-6.

Polansky, O., Sekelova, Z., Faldynova, M., Sebkova, A., Sisak, F., and Rychlik, I. (2015). Important Metabolic Pathways and Biological Processes Expressed by Chicken Cecal Microbiota. *Appl Environ Microbiol* 82(5)**,** 1569-1576. doi: 10.1128/aem.03473-15.

Polz, M.F., and Cavanaugh, C.M. (1998). Bias in Template-to-Product Ratios in Multitemplate PCR. *Appl Environ Microbiol* 64(10)**,** 3724-3730.

Poolman, B., Spitzer, J.J., and Wood, J.M. (2004). Bacterial osmosensing: roles of membrane structure and electrostatics in lipid-protein and protein-protein interactions. *Biochimica et Biophysica Acta (BBA) - Biomembrane* 1666(1-2)**,** 88-104. doi: 10.1016/j.bbamem.2004.06.013.

Pope, P.B., Mackenzie, A.K., Gregor, I., Smith, W., Sundset, M.A., McHardy, A.C., et al. (2012). Metagenomics of the Svalbard reindeer rumen microbiome reveals abundance of polysaccharide utilization loci. *PLoS One* 7(6)**,** e38571. doi: 10.1371/journal.pone.0038571.

Poulsen, M., Schwab, C., Jensen, B.B., Engberg, R.M., Spang, A., Canibe, N., et al. (2013). Methylotrophic methanogenic Thermoplasmata implicated in reduced methane emissions from bovine rumen. *Nat Commun* 4**,** 1428. doi: 10.1038/ncomms2432.

Prescott, L.M.; Harley, J.P.; Klein, D.A. Microbiology, Sixth edition, McGraw Hill International edition, New York. (2005).

Prins, R.A., Lankhorst, A., van der Meer, P., and Van Nevel, C.J. (1975). Some characteristics of Anaerovibrio lipolytica a rumen lipolytic organism. *Anton Leeuw J Microb* 41(1)**,** 1-11.

Prins, R.A., Van Vugt, F., Hungate, R.E., and Van Vorstenbosch, C.J. (1972). A comparison of strains of Eubacterium cellulosolvens from the rumen. *Anton Leeuw J Microb* 38(2)**,** 153-161.

Puniya, A.K., Singh, R., and Kamra, D.N. (2015). Rumen Microbiology: From Evolution to Revolution. New Delhi: Springer India.

Purushe, J., Fouts, D.E., Morrison, M., White, B.A., Mackie, R.I., Coutinho, P.M., et al. (2010). Comparative Genome Analysis of Prevotella ruminicola and Prevotella bryantii: Insights into Their Environmental Niche. *Microb Ecol* 60(4)**,** 721-729. doi: 10.1007/s00248-010-9692-8.

Quann, E.E., Fulgoni, V.L., and Auestad, N. (2015). Consuming the daily recommended amounts of dairy products would reduce the prevalence of inadequate micronutrient intakes in the United States: diet modeling study based on NHANES 2007-2010. *Nutr J* 14(1)**,** 90. doi: 10.1186/s12937-015-0057-5.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 41(Database issue)**,** D590-596. doi: 10.1093/nar/gks1219.

Raes, J., and Bork, P. (2008). Molecular eco-systems biology: towards an understanding of community function. *Nat Rev Micro* 6(9)**,** 693-699.

Ragsdale, S.W. (2014). "Biochemistry of Methyl-Coenzyme M Reductase: The Nickel Metalloenzyme that Catalyzes the Final Step in Synthesis and the First Step in Anaerobic Oxidation of the Greenhouse Gas Methane," in The Metal-Driven Biogeochemistry of Gaseous Compounds in the Environment, eds. P.M.H. Kroneck & M.E.S. Torres. (Dordrecht: Springer Netherlands), 125-145.

Ranilla, M.J., and Carro, M.D. (2003). Diet and procedures used to detach particle-associated microbes from ruminal digesta influence chemical composition of microbes and estimation of microbial growth in Rusitec fermenters. *J Anim Sci* 81(2)**,** 537-544.

Rappsilber, J., Mann, M., and Ishihama, Y. (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protocols* 2(8)**,** 1896-1906.

Rasmussen, M.A. (1993). Isolation and characterization of Selenomonas ruminantium strains capable of 2-deoxyribose utilization. *Appl Environ Microbiol* 59(7)**,** 2077-2081.

Razin, S., Yogev, D., and Naot, Y. (1998). Molecular biology and pathogenicity of mycoplasmas. *Microbiol Mol Biol Rev* 62(4)**,** 1094-1156.

Rimas, A., and Fraser, E.D.G. (2008). Beef: the untold story of how milk, meat, and muscle shaped the world. William Morrow.

Rinke, C., Schwientek, P., Sczyrba, A., Ivanova, N.N., Anderson, I.J., Cheng, J.F., et al. (2013). Insights into the phylogeny and coding potential of microbial dark matter. *Nature* 499(7459)**,** 431-437. doi: 10.1038/nature12352.

Ripple, W.J., Smith, P., Haberl, H., Montzka, S.A., McAlpine, C., and Boucher, D.H. (2014). Ruminants, climate change and climate policy. *Nature Clim. Change* 4(1)**,** 2-5. doi: 10.1038/nclimate2081.

Robinson, D.L., Goopy, J.P., Donaldson, A.J., Woodgate, R.T., Oddy, V.H., and Hegarty, R.S. (2014). Sire and liveweight affect feed intake and methane emissions of sheep confined in respiration chambers. *Animal* 8(12)**,** 1935-1944. doi: 10.1017/S1751731114001773.

Roehe, R., Dewhurst, R.J., Duthie, C.A., Rooke, J.A., McKain, N., Ross, D.W., et al. (2016). Bovine Host Genetic Variation Influences Rumen Microbial Methane Production with Best Selection Criterion for Low Methane Emitting and Efficiently Feed Converting Hosts Based on Metagenomic Gene Abundance. *PLoS Genet* 12(2)**,** e1005846. doi: 10.1371/journal.pgen.1005846.

Rubin, E.M. (2008). Genomics of cellulosic biofuels. *Nature* 454(7206)**,** 841-845.

Russell, J.B., and Rychlik, J.L. (2001). Factors that alter rumen microbial ecology. *Science* 292(5519)**,** 1119-1122.

Sadet, S., Martin, C., Meunier, B., and Morgavi, D.P. (2007). PCR-DGGE analysis reveals a distinct diversity in the bacterial population attached to the rumen epithelium. *Animal* 1(7)**,** 939-944. doi: 10.1017/s1751731107000304.

Saleem, F., Ametaj, B.N., Bouatra, S., Mandal, R., Zebeli, Q., Dunn, S.M., et al. (2012). A metabolomics approach to uncover the effects of grain diets on rumen health in dairy cows. *J Dairy Sci* 95(11)**,** 6606-6623. doi: 10.3168/jds.2012-5403.

Saleem, F., Bouatra, S., Guo, A., Psychogios, N., Mandal, R., Dunn, S., et al. (2013). The Bovine Ruminal Fluid Metabolome. *Metabolomics* 9(2)**,** 360-378. doi: 10.1007/s11306-012-0458-9.

Salvetti, E., Felis, G.E., Dellaglio, F., Castioni, A., Torriani, S., and Lawson, P.A. (2011). Reclassification of Lactobacillus catenaformis (Eggerth 1935) Moore and Holdeman 1970 and Lactobacillus vitulinus Sharpe et al. 1973 as Eggerthia catenaformis gen. nov., comb. nov. and Kandleria vitulina gen. nov., comb. nov., respectively. *Int J Syst Evol Microbiol* 61(Pt 10)**,** 2520-2524. doi: 10.1099/ijs.0.029231-0.

Santoni, V., Molloy, M., and Rabilloud, T. (2000). Membrane proteins and proteomics: un amour impossible? *Electrophoresis* 21(6)**,** 1054-1070. doi: 10.1002/(sici)1522-2683(20000401)21:6<1054::aid-elps1054>3.0.co;2-8.

Saveliev, S., Engel, L., Strauss, E., Jones, R., and Rosenblatt, M. (2012). The Advantages to Using Arg-C, Elastase, Thermolysin and Pepsin for Protein Analysis. Promega Corporation Web site. www.promega.de/resources/pubhub/the-advantages-to-using-arg-c-elastase-thermolysin-and-pepsin-for-protein-analysis/.

Schloss, P.D. (2010). The effects of alignment quality, distance calculation method, sequence filtering, and region on the analysis of 16S rRNA gene-based studies. *PLoS Comput Biol* 6(7)**,** e1000844. doi: 10.1371/journal.pcbi.1000844.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., et al. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* 75(23)**,** 7537-7541. doi: 10.1128/aem.01541-09.

Schrader, C., Schielke, A., Ellerbroek, L., and Johne, R. (2012). PCR inhibitors - occurrence, properties and removal. *J Appl Microbiol* 113(5)**,** 1014-1026. doi: 10.1111/j.1365-2672.2012.05384.x.

Schulze, W.X., and Usadel, B. (2010). Quantitation in mass-spectrometry-based proteomics. *Annu Rev Plant Biol* 61**,** 491-516. doi: 10.1146/annurev-arplant-042809-112132.

Scott, H.W., and Dehority, B.A. (1965). Vitamin Requirements of Several Cellulolytic Rumen Bacteria. *J Bacteriol* 89(5)**,** 1169-1175.

Seddon, A.M., Curnow, P., and Booth, P.J. (2004). Membrane proteins, lipids and detergents: not just a soap opera. *Biochim Biophys Acta (BBA) - Biomembranes* 1666(1-2)**,** 105-117. doi: 10.1016/j.bbamem.2004.04.011.

Seedorf, H., Kittelmann, S., and Janssen, P.H. (2015). Few Highly Abundant Operational Taxonomic Units Dominate within Rumen Methanogenic Archaeal Species in New Zealand Sheep and Cattle. *Appl Environ Microbiol* 81(3)**,** 986-995. doi: 10.1128/AEM.03018-14.

Seifert, J., Herbst, F.-A., Nielsen, P.H., Planes, F.J., Jehmlich, N., Ferrer, M., et al. (2013). Bioinformatic progress and applications in metaproteogenomics for bridging the gap between genomic sequences and metabolic functions in microbial communities. *Proteomics* 13(18-19, Sp. Iss. SI)**,** 2786-2804. doi: 10.1002/pmic.201200566.

Sejrsen, K., Hvelplund, T., and Nielsen, M.O. (2006). Ruminant Physiology: Digestion, Metabolism and Impact of Nutrition on Gene Expression, Immunology and Stress. Wageningen Academic Publishers.

Shabat, S.K.B., Sasson, G., Doron-Faigenboim, A., Durman, T., Yaacoby, S., Berg Miller, M.E., et al. (2016). Specific microbiome-dependent mechanisms underlie the energy harvest efficiency of ruminants. *ISME J* 10(12)**,** 2958-2972. doi: 10.1038/ismej.2016.62.

Shin, E.C., Choi, B.R., Lim, W.J., Hong, S.Y., An, C.L., Cho, K.M., et al. (2004). Phylogenetic analysis of archaea in three fractions of cow rumen based on the 16S rDNA sequence. *Anaerobe* 10(6)**,** 313-319. doi: 10.1016/j.anaerobe.2004.08.002.

Shinkai, T., Ohji, R., Matsumoto, N., and Kobayashi, Y. (2009). Fibrolytic capabilities of ruminal bacterium Fibrobacter succinogenes in relation to its phylogenetic grouping. *FEMS Microbiol Lett* 294(2)**,** 183-190.

Singh, K.M., Pandya, P.R., Tripathi, A.K., Patel, G.R., Parnerkar, S., Kothari, R.K., et al. (2014). Study of rumen metagenome community using qPCR under different diets. *Meta Gene* 2**,** 191-199. doi: 10.1016/j.mgene.2014.01.001.

Sirohi, S.K., Singh, N., Dagar, S.S., and Puniya, A.K. (2012). Molecular tools for deciphering the microbial community structure and diversity in rumen ecosystem. *Appl Microbiol Biotechnol* 95(5)**,** 1135-1154. doi: 10.1007/s00253-012-4262-2.

Sitao, W., Ransom L., B.t., Weizhong, L., Congjun, L., Erin E., C., and Robert W., L. (2012). The Bacterial Community Composition of the Bovine Rumen Detected Using Pyrosequencing of 16S rRNA Genes. *Metagenomics* Vol. 1 (2012), Article ID 235571, **,** 11 pages.

Skillman, L.C., Evans, P.N., Naylor, G.E., Morvan, B., Jarvis, G.N., and Joblin, K.N. (2004). 16S ribosomal DNA-directed PCR primers for ruminal methanogens and identification of methanogens colonising young lambs. *Anaerobe* 10(5)**,** 277-285. doi: 10.1016/j.anaerobe.2004.05.003.

Smith, B., Li, N., Andersen, A.S., Slotved, H.C., and Krogfelt, K.A. (2011). Optimising bacterial DNA extraction from faecal samples: comparison of three methods. *Open Microbiol J* 5**,** 14-17. doi: 10.2174/1874285801105010014.

Snelling, T.J., and Wallace, R.J. (2017). The rumen microbial metaproteome as revealed by SDS-PAGE. *BMC Microbiol* 17(1)**,** 9. doi: 10.1186/s12866-016-0917-y.

Solaiman, Z., Kashem, M.A., and Matsumoto, I. (2007). "Environmental Proteomics: Extraction and Identification of Protein in Soil," in Advanced Techniques in Soil Microbiology, eds. A. Varma & R. Oelmüller. (Berlin, Heidelberg: Springer Berlin Heidelberg), 155-166.

Sollinger, A., Schwab, C., Weinmaier, T., Loy, A., Tveit, A.T., Schleper, C., et al. (2016). Phylogenetic and genomic analysis of Methanomassiliicoccales in wetlands and animal intestinal tracts reveals clade-specific habitat preferences. *FEMS Microbiol Ecol* 92(1). doi: 10.1093/femsec/fiv149.

St-Pierre, B., Cersosimo, L.M., Ishaq, S.L., and Wright, A.-D.G. (2015). Toward the identification of methanogenic archaeal groups as targets of methane mitigation in livestock animalsr. *FrontMicrobiol* 6**,** 776. doi: 10.3389/fmicb.2015.00776.

St-Pierre, B., and Wright, A.D. (2013). Diversity of gut methanogens in herbivorous animals. *Animal* 7 Suppl 1**,** 49-56. doi: 10.1017/s1751731112000912.

Stack, R.J., and Hungate, R.E. (1984). Effect of 3-Phenylpropanoic Acid on Capsule and Cellulases of Ruminococcus albus 8. *Appl Environ Microbiol* 48(1)**,** 218-223.

Stackebrandt, E. (2014). "The Family Lachnospiraceae," in The Prokaryotes: Firmicutes and Tenericutes, eds. E. Rosenberg, E.F. DeLong, S. Lory, E. Stackebrandt & F. Thompson. (Berlin, Heidelberg: Springer Berlin Heidelberg), 197-201.

Stackebrandt, E., and Hippe, H. (1986). Transfer of Bacteroides amylophilus to a new genus Ruminobacter gen. nov., nom. rev. as Ruminobacter amylophilus comb. nov. *Syst Appl Microbiol* 8(3)**,** 204-207. doi: 10.1016/S0723-2020(86)80078-9.

Stanton, T.B., and Canale-Parola, E. (1980). Treponema bryantii sp. nov., a rumen spirochete that interacts with cellulolytic bacteria. *Arch Microbiol* 127(2)**,** 145-156.

Stiverson, J., Morrison, M., and Yu, Z. (2011). Populations of select cultured and uncultured bacteria in the rumen of sheep and the effect of diets and ruminal fractions. *Int J Microbiol* 2011**,** 750613. doi: 10.1155/2011/750613.

Suen, G., Stevenson, D.M., Bruce, D.C., Chertkov, O., Copeland, A., Cheng, J.F., et al. (2011). Complete genome of the cellulolytic ruminal bacterium Ruminococcus albus 7. *J Bacteriol* 193(19)**,** 5574-5575. doi: 10.1128/jb.05621-11.

Sugden, B., and Oxford, A.E. (1952). Some cultural studies with holotrich ciliate protozoa of the sheep's rumen. *J Gen Microbiol* 7(1-2)**,** 145-153. doi: 10.1099/00221287-7-1-2-145.

Swain, R.A., Nolan, J.V., and Klieve, A.V. (1996). Natural variability and diurnal fluctuations within the bacteriophage population of the rumen. *Appl Environ Microbiol* 62(3)**,** 994-997.

Tajima, K., Aminov, R.I., Nagamine, T., Matsui, H., Nakamura, M., and Benno, Y. (2001). Diet-dependent shifts in the bacterial population of the rumen revealed with real-time PCR. *Appl Environ Microbiol* 67(6)**,** 2766-2774. doi: 10.1128/aem.67.6.2766-2774.2001.

Tanca, A., Palomba, A., Deligios, M., Cubeddu, T., Fraumene, C., Biosa, G., et al. (2013). Evaluating the impact of different sequence databases on metaproteome analysis: insights from a lab-assembled microbial mixture. *PLoS One* 8. doi: 10.1371/journal.pone.0082981.

Tanca, A., Palomba, A., Fraumene, C., Pagnozzi, D., Manghina, V., Deligios, M., et al. (2016). The impact of sequence database choice on metaproteomic results in gut microbiota studies. *Microbiome* 4(1)**,** 51. doi: 10.1186/s40168-016-0196-8.

Tanca, A., Palomba, A., Pisanu, S., Deligios, M., Fraumene, C., Manghina, V., et al. (2014). A straightforward and efficient analytical pipeline for metaproteome characterization. *Microbiome* 2(1)**,** 49. doi: 10.1186/s40168-014-0049-2.

Tatusov, R.L., Galperin, M.Y., Natale, D.A., and Koonin, E.V. (2000). The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res* 28(1)**,** 33-36.

ter Beek, J., Guskov, A., and Slotboom, D.J. (2014). Structural diversity of ABC transporters. *Journal Gen Physiol* 143(4)**,** 419-435. doi: 10.1085/jgp.201411164.

Thareja, A., Puniya, A.K., Goel, G., Nagpal, R., Sehgal, J.P., Singh, P.K., et al. (2006). In vitro degradation of wheat straw by anaerobic fungi from small ruminants. *Arch Anim Nutr* 60(5)**,** 412-417. doi: 10.1080/17450390600884443.

Thoetkiattikul, H., Mhuantong, W., Laothanachareon, T., Tangphatsornruang, S., Pattarajinda, V., Eurwilaichitr, L., et al. (2013). Comparative analysis of microbial profiles in cow rumen fed with different dietary fiber by tagged 16S rRNA gene pyrosequencing. *Curr Microbiol* 67(2)**,** 130-137. doi: 10.1007/s00284-013-0336-3.

Trabalza-Marinucci, M., Poncet, C., Delval, E., and Fonty, G. (2006). Evaluation of techniques to detach particle-associated microorganisms from rumen contents. *Anim Feed Sci Tech* 125(1-2)**,** 1-16. doi: 10.1016/j.anifeedsci.2005.05.009.

Tremblay, J., Singh, K., Fern, A., Kirton, E., He, S., Woyke, T., et al. (2015). Primer and platform effects on 16S rRNA tag sequencing. *Front Microbiol* 6(771). doi: 10.3389/fmicb.2015.00771.

Tyanova, S., Temu, T., and Cox, J. (2016). The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc* 11(12)**,** 2301-2319. doi: 10.1038/nprot.2016.136.

Valle, E.R., Henderson, G., Janssen, P.H., Cox, F., Alexander, T.W., and McAllister, T.A. (2015). Considerations in the use of fluorescence *in situ* hybridization (FISH) and confocal laser scanning microscopy to characterize rumen methanogens and define their spatial distributions. *Can J Microbiol* 61(6)**,** 417-428. doi: 10.1139/cjm-2014-0873.

van Gastelen, S., Antunes-Fernandes, E.C., Hettinga, K.A., Klop, G., Alferink, S.J., Hendriks, W.H., et al. (2015). Enteric methane production, rumen volatile fatty acid concentrations, and milk fatty acid composition in lactating Holstein-Friesian cows fed grass silage- or corn silage-based diets. *J Dairy Sci* 98(3)**,** 1915-1927. doi: 10.3168/jds.2014-8552.

van Gylswyk, N.O. (1995). Succiniclasticum ruminis gen. nov., sp. nov., a ruminal bacterium converting succinate to propionate as the sole energy-yielding mechanism. *Int J Syst Bacteriol* 45(2)**,** 297-300. doi: 10.1099/00207713-45-2-297.

Van Gylswyk, N.O., and Van der Toorn, J.J.T.K. (1985). Eubacterium uniforme sp. nov. and Eubacterium xylanophilum sp. nov., Fiber-Digesting Bacteria from the Rumina of Sheep Fed Corn Stover. *Int J Syst Evol Micr* 35(3)**,** 323-326. doi: doi:10.1099/00207713-35-3-323.

van Houtert, M.F.J. (1993). The production and metabolism of volatile fatty acids by ruminants fed roughages: A review. *Anim Feed Sci Tech* 43(3)**,** 189-225. doi: 10.1016/0377-8401(93)90078-X.

van Lingen, H.J., Plugge, C.M., Fadel, J.G., Kebreab, E., Bannink, A., and Dijkstra, J. (2016). Thermodynamic Driving Force of Hydrogen on Rumen Microbial Metabolism: A Theoretical Investigation. *PLoS ONE* 11(10)**,** e0161362. doi: 10.1371/journal.pone.0161362.

Van Soest, P.J. (1994). Nutritional Ecology of the Ruminant. Comstock Pub.

Vaudel, M., Burkhart, J.M., Sickmann, A., Martens, L., and Zahedi, R.P. (2011). Peptide identification quality control. *Proteomics* 11(10)**,** 2105-2114. doi: 10.1002/pmic.201000704.

VDLUFA (2006). Handbuch der Landwirtschaftlichen Versuchs-und Untersuchungsmethodik (VDLUFA-Methodenbuch), Band III Die chemische Untersuchung von Futtermitteln.

VerBerkmoes, N.C., Denef, V.J., Hettich, R.L., and Banfield, J.F. (2009). Systems Biology: Functional analysis of natural microbial consortia using community proteomics. *Nat Rev Micro* 7(3)**,** 196-205.

Villegas-Rivera, G., Vargas-Cabrera, Y., González-Silva, N., Aguilera-García, F., Gutiérrez-Vázquez, E., Bravo-Patiño, A., et al. (2013). Evaluation of DNA extraction methods of rumen microbial populations. World Journal of Microbiology and Biotechnology 29(2)**,** 301-307. doi: 10.1007/s11274-012-1183-2.

Violle, C., Enquist, B.J., McGill, B.J., Jiang, L., Albert, C.H., Hulshof, C., et al. (2012). The return of the variance: intraspecific variability in community ecology. *Trends Ecol Evol* 27(4)**,** 244-252. doi: 10.1016/j.tree.2011.11.014.

Vogels, G.D., and Stumm, C. (1980). Interactions between methanogenic bacteria and hydrogenic ciliates in the rumen. *Anton Leeuw J Microb* 46(1)**,** 108.

Vos, P., Garrity, G., Jones, D., Krieg, N.R., Ludwig, W., Rainey, F.A., et al. (2011). Bergey's Manual of Systematic Bacteriology: Volume 3: The Firmicutes. Springer New York.

Wallace, R.J., Onodera, R., and Cotta, M.A. (1997). "Metabolism of nitrogen-containing compounds," in The Rumen Microbial Ecosystem, eds. P.N. Hobson & C.S. Stewart. (Dordrecht: Springer Netherlands), 283-328.

Wallace, R.J., Rooke, J.A., McKain, N., Duthie, C.A., Hyslop, J.J., and Ross, D.W. (2015). The rumen microbial metagenome associated with high methane production in cattle. *BMC Genomics* 16. doi: 10.1186/s12864-015-2032-0.

Wallace, R.J., Snelling, T.J., McCartney, C.A., Tapio, I., and Strozzi, F. (2017). Application of meta-omics techniques to understand greenhouse gas emissions originating from ruminal metabolism. *Genet Sel Evol* 49(1)**,** 9. doi: 10.1186/s12711-017-0285-6.

Walther, T.C., and Mann, M. (2010). Mass spectrometry-based proteomics in cell biology. *J Cell Biol* 190(4)**,** 491-500. doi: 10.1083/jcb.201004052.

Wang, L., Hatem, A., Catalyurek, U.V., Morrison, M., and Yu, Z. (2013). Metagenomic insights into the carbohydrate-active enzymes carried by the microorganisms adhering to solid digesta in the rumen of cows. *PLoS One* 8(11)**,** e78507. doi: 10.1371/journal.pone.0078507.

Wang, L., Xu, Q., Kong, F., Yang, Y., Wu, D., Mishra, S., et al. (2016). Exploring the Goat Rumen Microbiome from Seven Days to Two Years. *Plos One* 11(5)**,** e0154354. doi: 10.1371/journal.pone.0154354.

Wang, P., Qi, M., Barboza, P., Leigh, M.B., Ungerfeld, E., Selinger, L.B., et al. (2011). Isolation of high-quality total RNA from rumen anaerobic bacteria and fungi, and subsequent detection of glycoside hydrolases. *Can J Microbiol* 57(7)**,** 590-598. doi: 10.1139/w11-048.

Wang, Q., Garrity, G.M., Tiedje, J.M., and Cole, J.R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* 73(16)**,** 5261-5267. doi: 10.1128/aem.00062-07.

Wang, Y., Cao, P., Wang, L., Zhao, Z., Chen, Y., and Yang, Y. (2017). Bacterial community diversity associated with different levels of dietary nutrition in the rumen of sheep. *Appl Microbiol Biotechnol* 101(9)**,** 3717-3728. doi: 10.1007/s00253-017-8144-5.

Wang, Y., and Qian, P.-Y. (2015). "Conserved Regions in 16S Ribosome RNA Sequences and Primer Design for Studies of Environmental Microbes," in Encyclopedia of Metagenomics: Genes, Genomes and Metagenomes: Basics, Methods, Databases and Tools, ed. K.E. Nelson. (Boston, MA: Springer US), 106-110.

Watanabe, Y., Nagai, F., and Morotomi, M. (2012). Characterization of Phascolarctobacterium succinatutens sp. nov., an Asaccharolytic, Succinate-Utilizing Bacterium Isolated from Human Feces. *Appl Environ Microbiol* 78(2)**,** 511-518. doi: 10.1128/AEM.06035-11.

Weimer, P.J. (2015). Redundancy, resilience, and host specificity of the ruminal microbiota: implications for engineering improved ruminal fermentations. *Front Microbiol* 6**,** 296. doi: 10.3389/fmicb.2015.00296.

Weimer, P.J., Stevenson, D.M., Mantovani, H.C., and Man, S.L. (2010). Host specificity of the ruminal bacterial community in the dairy cow following near-total exchange of ruminal contents. *J Dairy Sci* 93(12)**,** 5902-5912. doi: 10.3168/jds.2010-3500.

Weljie, A.M., Newton, J., Mercier, P., Carlson, E., and Slupsky, C.M. (2006). Targeted profiling: quantitative analysis of 1H NMR metabolomics data. *Anal Chem* 78(13)**,** 4430-4442. doi: 10.1021/ac060209g.

Welkie, D.G., Stevenson, D.M., and Weimer, P.J. (2010). ARISA analysis of ruminal bacterial community dynamics in lactating dairy cows during the feeding cycle. *Anaerobe* 16(2)**,** 94-100. doi: 10.1016/j.anaerobe.2009.07.002.

Wesolowska-Andersen, A., Bahl, M.I., Carvalho, V., Kristiansen, K., Sicheritz-Ponten, T., Gupta, R., et al. (2014). Choice of bacterial DNA extraction method from fecal material influences community structure as evaluated by metagenomic analysis. *Microbiome* 2**,** 19. doi: 10.1186/2049-2618-2-19.

Wetzels, S.U., Mann, E., Metzler-Zebeli, B.U., Pourazad, P., Qumar, M., Klevenhusen, F., et al. (2016). Epimural Indicator Phylotypes of Transiently-Induced Subacute Ruminal Acidosis in Dairy Cattle. *Front Microbiol* 7**,** 274. doi: 10.3389/fmicb.2016.00274.

Wetzels, S.U., Mann, E., Pourazad, P., Qumar, M., Pinior, B., Metzler-Zebeli, B.U., et al. (2017). Epimural bacterial community structure in the rumen of Holstein cows with different responses to a long-term subacute ruminal acidosis diet challenge. *J Dairy Sci* 100(3)**,** 1829-1844. doi: 10.3168/jds.2016-11620.

Whelan, S.J., Mulligan, F.J., Flynn, B., McCarney, C., and Pierce, K.M. (2011). Effect of forage source and a supplementary methionine hydroxy analog on nitrogen balance in lactating dairy cows offered a low crude protein diet. *J Dairy Sci* 94(10)**,** 5080-5089. doi: 10.3168/jds.2011-4174.

Whitehouse, N.L., Olson, V.M., Schwab, C.G., Chesbro, W.R., Cunningham, K.D., and Lykos, T. (1994). Improved techniques for dissociating particle-associated mixed ruminal microorganisms from ruminal digesta solids. *J Anim Sci* 72(5)**,** 1335-1343.

Williams, A.G., and Coleman, G.S. (1992). The Rumen Protozoa. Springer New York.

Wilmes, P., and Bond, P.L. (2004). The application of two-dimensional polyacrylamide gel electrophoresis and downstream analyses to a mixed community of prokaryotic microorganisms. *Environ Microbiol* 6(9)**,** 911-920. doi: 10.1111/j.1462-2920.2004.00687.x.

Wilmes, P., and Bond, P.L. (2006). Metaproteomics: studying functional gene expression in microbial ecosystems. *Trends Microbiol* 14(2)**,** 92-97. doi: 10.1016/j.tim.2005.12.006.

Wilmes, P., Heintz-Buschart, A., and Bond, P.L. (2015). A decade of metaproteomics: where we stand and what the future holds. *Proteomics* 15(20)**,** 3409-3417. doi: 10.1002/pmic.201500183.

Wilkinson, J. M., F. Wadephul, and J. Hill. 1996. Silages in Europe:
A survey of 33 countries. Chalcombe Publications, Lincoln, UK.

Wishart, D.S. (2008). Quantitative metabolomics using NMR. *Trac-Tren Anal Chem* 27(3)**,** 228-237. doi: 10.1016/j.trac.2007.12.001.

Wisniewski, J.R., Zougman, A., Nagaraj, N., and Mann, M. (2009). Universal sample preparation method for proteome analysis. *Nat Meth* 6(5)**,** 359-362. doi: 10.1038/nmeth.1322.

Wright, A.-D.G., and Klieve, A.V. (2011). Does the complexity of the rumen microbial ecology preclude methane mitigation? *Anim Feed Sci Tech* 166-167**,** 248-253. doi: 10.1016/j.anifeedsci.2011.04.015.

Wu, C., Tran, J.C., Zamdborg, L., Durbin, K.R., Li, M., Ahlf, D.R., et al. (2012). A protease for 'middle-down' proteomics. *Nat Meth* 9(8)**,** 822-824. doi: 10.1038/nmeth.2074.

Wu, S., Zhu, Z., Fu, L., Niu, B., and Li, W. (2011). WebMGA: a customizable web server for fast metagenomic sequence analysis. *BMC Genomics* 12**,** 444. doi: 10.1186/1471-2164-12-444.

Yadav, A.K., Kumar, D., and Dash, D. (2012). Learning from decoys to improve the sensitivity and specificity of proteomics database search results. *PLoS One* 7(11), e50651. doi: 10.1371/journal.pone.0050651.

Yang, Y., Wang, J., Yuan, T., Bu, D., Yang, J., and Sun, P. (2013). Proteome profile of bovine ruminal epithelial tissue based on GeLC-MS/MS. *Biotechnol Lett* 35(11), 1831-1838. doi: 10.1007/s10529-013-1291-6.

Yarza, P., Yilmaz, P., Pruesse, E., Glockner, F.O., Ludwig, W., Schleifer, K.H., et al. (2014). Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat Rev Microbiol* 12(9), 635-645. doi: 10.1038/nrmicro3330.

Yates, J.R., Ruse, C.I., and Nakorchevsky, A. (2009). Proteomics by Mass Spectrometry: Approaches, Advances, and Applications. *Annu Rev Biomed Eng* 11(1), 49-79. doi: 10.1146/annurev-bioeng-061008-124934.

Yin, Y., Mao, X., Yang, J., Chen, X., Mao, F., and Xu, Y. (2012). dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res* 40(Web Server issue), W445-451. doi: 10.1093/nar/gks479.

Yost, W.M., Young, J.W., Schmidt, S.P., and McGilliard, A.D. (1977). Gluconeogenesis in ruminants: propionic acid production from a high-grain diet fed to cattle. *J Nutr* 107(11), 2036-2043.

Youssef, N., Sheik, C.S., Krumholz, L.R., Najar, F.Z., Roe, B.A., and Elshahed, M.S. (2009). Comparison of species richness estimates obtained using nearly complete fragments and simulated pyrosequencing-generated fragments in 16S rRNA gene-based environmental surveys. *Appl Environ Microbiol* 75(16), 5227-5236. doi: 10.1128/aem.00592-09.

Yu, Z., and Morrison, M. (2004). Improved extraction of PCR-quality community DNA from digesta and fecal samples. *Biotechniques* 36(5), 808-812.

Zhang, R., Zhu, W., Zhu, W., Liu, J., and Mao, S. (2013). Effect of dietary forage sources on rumen microbiota, rumen fermentation and biogenic amines in dairy cows. *J Sci Food Agric* 94(9), 1886-1895. doi: 10.1002/jsfa.6508.

Zhao, S., Zhao, J., Bu, D., Sun, P., Wang, J., and Dong, Z. (2014). Metabolomics analysis reveals large effect of roughage types on rumen microbial metabolic profile in dairy cows. *Lett Appl Microbiol* 59(1), 79-85. doi: 10.1111/lam.12247.

Zhou, M., Hünerberg, M., Beauchemin, K.A., McAllister, T.A., Okine, E.K., and Guan, L.L. (2012). Individuality of ruminal methanogen/protozoa populations in beef cattle fed diets containing dried distillers' grain with solubles. *Acta Agr Scand A-An* 62(4), 273-288. doi: 10.1080/09064702.2013.788206.

Zhu, A., Sunagawa, S., Mende, D.R., and Bork, P. (2015). Inter-individual differences in the gene content of human gut bacterial species. *Genome Biol* 16, 82. doi: 10.1186/s13059-015-0646-9.

Zhu, W., Fu, Y., Wang, B., Wang, C., Ye, J.A., Wu, Y.M., et al. (2013). Effects of dietary forage sources on rumen microbial protein synthesis and milk performance in early lactating dairy cows. *J Dairy Sci* 96(3)**,** 1727-1734. doi: 10.3168/jds.2012-5756.

CHAPTER VII

SUMMARY

7. SUMMARY

Globally, ruminant livestock constitutes a significant source of food for humans. Under the aspect of a growing world population, an increased demand for animal-based food products is expected. Production efficiency and associated health status of the animals as well as the emission of greenhouse gases are mainly determined by the rumen microbiome. The structure and activity of the microbial communities in the rumen ecosystem in turn are mostly influenced by the animal's feed intake.

The most widely used forage sources for ruminant production in Europe and North America are corn silage, grass silage and grass hay characterized by varying chemical compositions. Different proportions of plant-derived polysaccharides, fibers and crude protein have diverse effects on the rumen microbiome. Progress in animal production requires optimized feeding strategies, which presuppose an improved understanding of the dietary impact on the complex bionetwork residing in the rumen.

A broad range of different methods are applicable to investigate archaea and bacteria which represent the most active and significant members of the rumen microbiome. Yet, most rumen studies available are restricted to nucleic acid-based approaches with limited functional insights. To improve knowledge about the prokaryotic communities and their adaptation responses to different animal feeds, it is essential to focus on the actual functions out of numerous possibilities that are encoded by the genomes of the rumen microbiome. In this way, further insights into the interactions of different members of the rumen ecosystem as well as host-microbe interactions can be obtained. Therefore, proteins are best suited since representing the actual function of investigated cells combined with phylogenetic information.

The major aim of this project was the feasible, first-time establishment of a metaproteomics-based characterization of the ruminal prokaryotic communities to further investigate the dietary impact on the prokaryotic rumen metaproteome.

The first part of this project was providing an overview about research that used state of the art technologies to investigate the microbiome of the gastrointestinal tract of farm animals. Different *Omics*-technologies and their combination are commonly applied in studies of the human gut microbiome but comparable research in livestock science is rare and in the best case employed functional metagenomics. Thus, as well pre-*Omics*-approaches were included in review, further comparing and discussing their major findings. At that time, the

considered studies relied mainly on stand-alone, DNA-based molecular methods, which clearly emphasized the importance of introducing contemporary methods such as shotgun metaproteomics to study the rumen microbiome and to gain deeper, more complete insights into the actual functions carried out by the specific members of the prokaryotic communities.

The second part of the current project focused on a suitable, mass spectrometry-based analysis of the prokaryotic communities in the rumen ecosystem. Metaproteomic studies are challenged by the heterogeneity of the rumen sample matter that contains, besides archaeal and bacterial cells, also eukaryotic cells from rumen fungi and protozoa as well as enormous amounts of plant cells from ingested feed and epithelial cells of the animals. Shotgun metaproteomic studies require the extraction of proteins preferably of the desired target organisms to increase the coverage of the respective metaproteome and the reliability of subsequent protein identifications. This entails the avoidance of undesired proteins present in the rumen samples. In contrast to nucleic acids, proteins cannot be enriched or amplified by PCR and there is no universal sample preparation method. Thus, optimized sample preparation protocols are necessary in order to retrieve enhanced amounts of prokaryotic instead of plant-derived or other eukaryotic cells before protein extraction and subsequent LC-MS/MS analysis. Moreover, individual sample preparation protocols have to be developed considering as well the different ecological niches within the rumen ecosystem. Various sample preparation protocols for the fiber-adherent and the liquid-associated bacteria and archaea were formulated and evaluated in regard to the number of possible peptide and protein identifications, the ratio between prokaryotic and eukaryotic proteins as well as in defining bias towards specific phylogenetic groups. Using methylcellulose for the detachment of the fiber-adherent bacterial consortia and common cheesecloth, filters and centrifugation steps for the separation of prokaryotic cells from the residual matter significantly improved the identification ratio of prokaryotic to eukaryotic proteins. The respective findings were published in the journal of Proteomics in 2015 and simultaneously constituted the first study of the ruminal, prokaryotic metaproteome characterized by shotgun metaproteomics. Nevertheless, the amount of identified proteins depicted only a meager part of the entirety of prokaryotic proteins actually expressed in the rumen accentuating a general obstacle in conducting representative metaproteomic studies. This indicated the need of further improvements such as enhanced mass spectrometric measurements and more precise, sample-specific search databases for the bioinformatic processing steps.

The final step and the major aim of this project was the in depth analysis of the metaproteome of archaea and bacteria and their adaptive response to the most common forages, corn silage, grass silage and grass hay accessing as well host-related influences and variations between different ecological niches within the rumen. Improved mass-spectrometric measurements and the construction of an artificial metagenome in form of a customized, sample-specific in-house database for enhanced bioinformatic identification and quantification of proteins yielded comprehensive datasets comprising 8,163 bacterial and 358 archaeal proteins that were identified across 27 samples from three different rumen fractions of three Jersey cows, fed rotationally with three different diets. To obtain insights as accurate and complementary as possible, the functional and structural data of the metaproteomic analysis was further flanked by 16S rRNA gene-based analyses of the archaeal and bacterial community structures and the metabolomes of the rumen fluid fractions were quantified by nuclear magnetic resonance. The third manuscript carrying the title “A structural and functional elucidation of the rumen microbiome influenced by various diets and allocated space” was submitted to Frontiers in Systems Microbiology in March 2017. So far, to the best of our knowledge, there are no studies investigating the metaproteome expressed by the entirety of archaeal and bacterial communities in the different phases of the rumen ecosystem under varying dietary influence. Dietary treatments revealed significant variations in the metaproteome composition and community structures of ruminal bacteria. Host-related effects were not significant. Protein- and DNA-based approaches confirmed an increased abundance of the phylum of Proteobacteria and the family of Succinivibrionacea throughout the sample fractions of the corn silage-based diet, which potentially enhanced succinate production as indicated by the metabolomic analysis. This might be linked to a comparable rise of the family of Acidaminococcacea that utilize succinate to produce propionate. The abundance of the family of Erysipelotrichaceae was promoted in the samples of the hay-based diet as observed by both methods. Similar to *Lactobacillus* species, most members of Erysipelotrichaceae group probably produce mainly lactic acid as supported by the higher amounts of lactate in the samples of the hay-based diet. The most dominant bacterial family in the rumen ecosystem, the Prevotellaceae family did not respond to dietary changes probably due to their versatile metabolic capabilities. Nevertheless, the Prevotellaceae family appeared most abundant in the loosely fiber-adherent sample fractions regardless the dietary treatments. Several fiber-degraders such as the order of Clostridiales and the included families of Ruminococcaceae and Lachnospiraceae were found in increased abundance in the tightly fiber-adherent sample fractions as revealed by proteins and OTUs further indicating the

importance of sample fractionation in rumen studies. The common rumen family of Selenomonadaceae was identified exclusively in the metaproteomic approach emphasizing the necessity of combining different *Omics*-technologies to investigate the microbiomes of complex ecosystems. The rumen archaea were not influenced by diets or sample fractions, which might be explained by the low numbers of proteins and OTUs identified and by the less versatile metabolic capabilities when compared to bacteria. Functional analyses of the metaproteome revealed an increased abundance of monosaccharide transporters in the corn-silage-based diet whereas oligosaccharide transporters were more abundant in the fiber-rich grass- and hay-based diets. Bioinformatic analysis identified a total of 166 carbohydrate-active enzymes, such as glycoside hydrolases, glycosyltransferases, carbohydrate esterases and carbohydrate-binding modules. A high number of proteins involved in carbon metabolism were identified revealing different abundances of proteins belonging to the Embden-Meyerhof-Parnas pathway, the pentose phosphate pathway and the citrate cycle. Several enzymes involved in short-chain fatty acid production were identified.

In conclusion, within this project the application of shotgun metaproteomics to characterize the prokaryotic rumen metaproteome was successfully implemented and the results clearly emphasized the benefits of using complementary, state of the art methods to study the microbiome of complex ecosystems like the rumen. The obtained datasets revealed a significant impact of the most common forage sources on the bacterial communities in the rumen. Considering the specific functional niches of the rumen microbiome have been shown to be of great importance. This study may provide deeper insights into the complicated network of bacterial interactions and adaptions to various substrates. Bioinformatic and technical progress may enhance the metaproteomic coverage of future studies.

CHAPTER VIII

ZUSAMMENFASSUNG

8. ZUSAMMENFASSUNG

Wiederkäuer stellen weltweit eine wichtige Nahrungsquelle für Menschen dar. Aufgrund einer wachsenden Weltbevölkerung wird eine erhöhte Nachfrage nach tierischen Nahrungsmittelprodukten erwartet. Die Produktionseffizienz und der damit verbundene Gesundheitszustand der Tiere, sowie die Emission von Treibhausgasen, werden vor allem durch das ruminale Mikrobiom bestimmt. Die Struktur und die Aktivität der mikrobiellen Gemeinschaften im Ökosystem des Pansens werden wiederum hauptsächlich durch die Futteraufnahme des Tieres beeinflusst.

Die gebräuchlichsten Futtermittel in der Rinderproduktion in Europa und Nordamerika sind Maissilage, Grassilage und Heu, die sich in ihrer chemischen Zusammensetzung unterscheiden. Die verschiedenen Anteile an Polysacchariden, Rohfasern und Rohprotein haben unterschiedlich Auswirkungen auf das Pansenmikrobiom. Um weitere Fortschritte in der Tierproduktion zu erzielen, werden verbesserte Fütterungsstrategien benötigt, was eine umfassendere Kenntnis über den Einfluss der Futtermittel auf das komplexe Zusammenspiel der Mikroorganismen im Pansen voraussetzt.

Zur Untersuchung der aktivsten und bedeutendsten Mitglieder des ruminalen Mikrobioms, der Archaeen und Bakterien, steht eine Vielfalt verschiedenster Methoden zur Auswahl. Die meisten verfügbaren Studien sind jedoch auf nukleinsäuren-basierte Ansätze beschränkt. Um neue Erkenntnisse über die prokaryotischen Gemeinschaften und deren Anpassung an verschiedene Futtermittel zu erlangen, ist es von grundlegender Bedeutung, sich auf die von einer Vielzahl in den Genomen des ruminalen Mikrobioms kodierten Möglichkeiten, auf die tatsächlich vorhandenen Funktionen zu konzentrieren. Somit können tiefere Einblicke in die Interaktionen verschiedener Gruppen des Pansenökosystems sowie Wechselwirkungen von Mikroorganismen und dem Wirtstier gewonnen werden. Proteine eignen sich hervorragend zu diesem Zweck, da sie die tatsächliche Funktion der zu untersuchenden Zellen darstellen und zusätzlich phylogenetische Information mit sich bringen.

Das Hauptziel des vorliegenden Projektes war eine praktikable, proteinbasierte Charakterisierung der prokaryotischen Gemeinschaften im Pansen, um in weiteren Schritten die Auswirkungen verschiedener Futtermittel auf das ruminale Metaproteom zu untersuchen.

Der erste Schritt war es, eine zusammenfassende Übersicht bezüglich der bereits vorhandenen Studien des Mikrobioms im Gastrointestinaltrakt von Nutztieren nach aktuellem Stand der Wissenschaft zu erstellen. Die Kombination verschiedener Omik-Technologien kommt bei der Untersuchung des menschlichen Darmmikrobioms regelmäßig zum Einsatz. Vergleichbare Studien sind in der Nutztierwissenschaft jedoch kaum vorhanden und beschränken sich auf einige wenige funktionelle Untersuchungen des Metagenomes. Daher wurden auch weniger zeitgemäße Arbeiten für Vergleiche und Diskussion in den Bericht miteinbezogen. Zum damaligen Zeitpunkt verwendeten die meisten Arbeiten einfache, nukleinsäure-basierte Methoden, was die Notwendigkeit der Einführung zeitgemäßer Methoden, wie der Shotgun Metaproteomik, zur Untersuchung des Pansenmikrobioms klar verdeutlichte.

Der zweite Teil des vorliegenden Projektes bestand darin, eine geeignete massenspektrometrische Untersuchung des Pansenmikrobioms zu ermöglichen. Die Herausforderung bei metaproteomischen Untersuchungen besteht dabei in der Komplexität der Pansenmaterie, die eukaryotische Zellen von Pilzen und Protozoen sowie prokaryotische Zellen der Archaeen und Bakterien umfasst. Des Weiteren sind riesige Mengen an pflanzlichen und tierischen Zellen vorhanden. Repräsentative metaproteomische Untersuchungen benötigen die Proteine der zu untersuchenden Organismen, ungewollte Proteine sollen vermieden werden. Im Gegensatz zu Nukleinsäuren können Proteine nicht via PCR angereichert werden. Daher werden spezielle Probenaufbereitungsverfahren benötigt, um die prokaryotischen Zellen noch vor der Proteinextraktion und massenspektrometrischen Messung von den übrigen Zellen abzutrennen. Außerdem müssen die bestimmten ökologischen Nischen im Pansen bei der Probenaufbereitung berücksichtigt werden. Verschiedene Probenaufbereitungsverfahren wurden entwickelt und hinsichtlich des Verhältnisses von prokaryotischen zu eukaryotischen Proteinidentifikationen und der Gesamtmenge an Identifikationen bewertet. Unter Verwendung von Methylcellulose zur Ablösung der angehefteten Bakterien und gewöhnlichen Gazen zur Abtrennung der restlichen Materie konnten signifikant verbesserte Identifikationsraten erzielt werden. Die betreffenden Ergebnisse wurden in dem wissenschaftlichen Magazin Proteomics im Jahr 2015 veröffentlicht und stellten gleichzeitig die erste metaproteomische Untersuchung von Pansenmikroorganismen dar. Die Gesamtanzahl der identifizierten Proteine umfasste jedoch nur einen Bruchteil der tatsächlich im Pansen vorhandenen Proteine, was weitere Verbesserungen bezüglich der massenspektrometrischen Messungen und der Suchdatenbanken für die bioinformatischen Verarbeitungsschritte voraussetzte.

Der letzte und wichtigste Schritt in diesem Projekt war die detaillierte Untersuchung des prokaryotischen Metaproteoms und dessen Anpassungsreaktionen auf die gebräuchlichsten Futtermittel: Maissilage, Grassilage und Heu. Des Weiteren wurden der Einfluss der Wirtstiere und die Unterschiede zwischen den ökologischen Nischen untersucht. Verbesserte massenspektrometrische Messungen und das Erstellen einer probenspezifischen Datenbank für die bioinformatischen Schritte der Proteinidentifikation und -quantifizierung ergaben umfassende Datensätze mit 8163 bakteriellen Proteinen und 358 archäischen Proteinen, die über 27 Proben aus drei Pansenfraktionen, von drei Kühen und drei Futtermitteln quantifiziert wurden. Um möglichst umfassende und akkurate Ergebnisse zu erzielen, wurde die metaproteomische Untersuchung mit einer 16S rRNA-basierten Analyse der Archaeen- und Bakteriengemeinschaften ergänzt. Zusätzlich wurden die Metabolite im Pansensaft der jeweiligen Proben via Kernspinresonanz bestimmt. Das dritte Manuskript mit dem Titel “A structural and functional elucidation of the rumen microbiome influenced by various diets and allocated space” wurde im März 2017 bei dem Magazin Frontiers in Systems Microbiology eingereicht. Nach meinem besten Wissen ist dies die erste Untersuchung diätischer Auswirkungen auf das prokaryotische Metaproteom im Pansen unter Berücksichtigung der verschiedenen ökologischen Nischen. Es zeigten sich signifikante Einflüsse der diätischen Behandlungen auf das Metaproteom und die bakterielle Gemeinschaftsstruktur im Pansen. Der Einfluss individueller Versuchstiere war nicht signifikant. Eindeutige Unterschiede bestanden auch zwischen den ökologischen Nischen. Protein- und DNA-basierte Datensätze bestätigten eine erhöhte Häufigkeit des Phylums der Proteobakterien und der Familie der Succinivibrionaceae in allen Fraktionen der Maissilage, was wahrscheinlich einen Anstieg in der Produktion von Succinat verursachte, was wiederum für ein erhöhtes Vorkommen der Familie der Acidaminococcacea in den Fraktionen der Maissilage verantwortlich war. Die Familie der Erysipelotrichaceae war häufiger in den Heuproben vorzufinden. Vergleichbar mit *Lactobacillus* Spezies produzieren Erysipelotrichaceae vermutlich Milchsäure, was durch die Kernspinresonanzanalyse bestätigt wurde. Die dominanteste Familie der Prevotellaceae wurde nicht von den verschiedenen Futtermitteln beeinflusst, wahrscheinlich auf Grund ihrer vielfältigen metabolischen Kapazitäten. Viele faserabbauende Spezies der Ordnung Clostridiales und den Familien Lachnospiraceae und Ruminococcaceae wiesen eine erhöhte Häufigkeit in den Proben der festen Phasen auf, was über beide Methoden bestätigt wurde und die Bedeutung von Probenunterteilungen in Untersuchungen des Pansenökosystems hervorhebt. Die gewöhnlich im Pansen vertretene Familie der Selenomonadaceae wurde ausschließlich im

metaproteomischen Datensatz identifiziert, was die Notwendigkeit der Anwendung verschiedener Methoden zur Untersuchung der mikrobiellen Gemeinschaften des Pansens klar betont. Die Archaeengemeinschaft wurde nicht von den verschiedenen Diäten beeinflusst und zeigte keine Unterschiede hinsichtlich der verschiedenen Probenfraktionen. Ursachen könnten das geringe metabolische Potential der Archaeen und die geringen Mengen an identifizierten Proteinen und 16S rRNA Gensequenzen sein. Funktionelle Analysen des Metaproteoms ergaben eine erhöhte Häufigkeit von Monosaccharidtransportern in den Probenfraktionen der Maissilage, während Oligosaccharidtransporter häufiger in den faserreichen Grassilage- und Heuproben vorkamen. Bioinformatische Auswertungen des Metaproteoms identifizierten 166 kohlenhydrat-aktive Enzyme und eine große Anzahl von Enzymen des Kohlenstoffmetabolismus sowie weitere Enzyme, die an der Bildung kurzkettiger Fettsäuren beteiligt sind.

In diesem Projekt wurde die Einführung der Shotgun Metaproteomik zur Untersuchung der Pansenmikroorganismen erfolgreich umgesetzt. Zudem verdeutlichen die erzielten Ergebnisse die Vorteile der Kombination verschiedener moderner Methoden zur Untersuchung mikrobieller Gemeinschaften. Eindeutige diätische Einflüsse auf das Pansenmikrobiom und grundlegende Unterschiede zwischen den ökologischen Nischen wurden erfasst. Diese Studie ermöglicht tiefere Einblicke in das komplizierte Netzwerk der Mikroorganismen im Pansen. Weitere technische und bioinformatische Fortschritte könnten die Abdeckung des ruminalen Metaproteoms künftiger Studien noch erweitern.

CHAPTER IX

APPENDIX

ACKNOWLEDGEMENTS

The quality of experience and learning within a demanding project like a PhD is greatly influenced by the people involved in the work. I am fortunate and privileged by having been surrounded with patient, positive and smart scientists.

First, I want to gratefully and sincerely thank Jun.-Prof. Dr. Jana Seifert for the opportunity and her trust to work on this project. Her encouragement and challenges, the continuous guidance and excellent supervision throughout the years and during the completion of my dissertation were crucial to this project and my personal progress.

Moreover, I would like to thank Dr. Amélia Camarinha-Silva for her assistance, scientific advices and motivating optimism as well as Prof. Dr. Markus Rodehutscord for providing the basis for the execution of the experiments.

I am grateful to Dr. Jürgen Conrad for his cooperation and support regarding the nuclear magnetic resonance, Dr. Florian-Alexander Herbst for his helpful advices concerning the bioinformatics analyses and Dr. Olaf Tyc for his interest, suggestions and scientific exchange.

I would like to thank all members and colleagues of the "Feed-Gut Microbiota Interaction" and the "Young Scientist Group Microbial Ecology" for providing an enjoyable, social and friendly environment. It was a pleasure!

Special and deepest thanks go to my mother, father and my siblings for their fundamental support over all these years. This doctorate would have been impossible without their understanding, indulgence and support. I want to thank my close friends for their patience and for not taking my absence badly.

Last but not least, I am most deeply grateful to Verena Wilhelm. She had to endure my moods, cheered me up and cared. Thanks for your comprehensive patience, understanding and all the support.

CURRICULUM VITAE

Personal Data:

Name:	Simon Deusch
Date of birth:	28 Aug 1985
Place of birth:	St. Georgen im Schwarzwald
Nationality:	German
Address:	Wiesenstrasse 22, 78112 St. Georgen, Germany
Email:	deusch@uni-hohenheim.de, deuschsimon@gmail.com
Phone:	+4977242285, +4917629335180

Work Experience:

Postdoc position at the department of Feed-Gut Microbiota Interaction (Oct 2017 – now)
Institute of Animal Science, University of Hohenheim, Stuttgart, Germany

PhD student at the department of Feed-Gut Microbiota Interaction (Jul 2013 – Oct 2017)
Institute of Animal Science, University of Hohenheim, Stuttgart, Germany

Employee and ERASMUS internship within the BACSIN project (Apr 2011 – May 2012)
Netherlands Institute of Ecology, Department of Microbial Ecology, Wageningen, NL

Employee as student assistant at the IT-service and Data-center (Mar 2008 – Mar 2011)
KIM, University of Hohenheim, Stuttgart, Germany

Education:

Doctorate at the Faculty of Agricultural Science (Jul 2013 - Oct 2017)
Department of Feed-Gut Microbiota Interaction, Institute of Animal Science
University of Hohenheim, Stuttgart, Germany
Degree Dissertation:
"Metaomic studies of the dietary impact on the structural and functional diversity of the rumen microbiome"

Diploma Agricultural Biologist (Apr 2006 - Mar 2013)
Focus: Microbiology and Biochemistry
University of Hohenheim, Stuttgart, Germany
Degree Dissertation:
"The role of catabolic genes in phyllosphere colonization of Arthrobacter chlorophenolicus A6"
Netherlands Institute of Ecology, Department of Microbial Ecology, Wageningen, NL

St. Georgen, 19th of June 2017

Simon Deusch

AFFIDAVIT

Bei der von mir eingereichten Dissertation zum Thema „*Metaomic studies of the dietary impact on the structural and functional diversity of the rumen microbiome*“ handelt es sich um meine eigenständig erbrachte Leistung.

Ich habe nur die angegebenen Quellen und Hilfsmittel benutzt und mich keiner unzulässigen Hilfe Dritter bedient. Insbesondere habe ich wörtlich oder sinngemäß aus anderen Werken übernommene Inhalte als solche kenntlich gemacht.

Ich habe nicht die Hilfe einer kommerziellen Promotionsvermittlung oder -beratung in Anspruch genommen.

Die Bedeutung der eidesstattlichen Versicherung und der strafrechtlichen Folgen einer unrichtigen oder unvollständigen eidesstattlichen Versicherung sind mir bekannt.

Die Richtigkeit der vorstehenden Erklärung bestätige ich. Ich versichere an Eides Statt, dass ich nach bestem Wissen die reine Wahrheit erklärt und nichts verschwiegen habe

St. Georgen, den 19. Juni 2017 _____

Simon Deusch

Die Universität Hohenheim verlangt eine Eidesstattliche Versicherung über die Eigenständigkeit der erbrachten wissenschaftlichen Leistungen, um sich glaubhaft zu versichern, dass die Promovendin bzw. der Promovend die wissenschaftlichen Leistungen eigenständig erbracht hat.

Weil der Gesetzgeber der Eidesstattlichen Versicherung eine besondere Bedeutung beimisst und sie erhebliche Folgen haben kann, hat der Gesetzgeber die Abgabe einer falschen eidesstattlichen Versicherung unter Strafe gestellt. Bei vorsätzlicher (also wissentlicher) Abgabe einer falschen Erklärung droht eine Freiheitsstrafe bis zu drei Jahren oder eine Geldstrafe.

Eine fahrlässige Abgabe (also Abgabe, obwohl Sie hätten erkennen müssen, dass die Erklärung nicht den Tatsachen entspricht) kann eine Freiheitsstrafe bis zu einem Jahr oder eine Geldstrafe nach sich ziehen.

Die entsprechenden Strafvorschriften sind in § 156 StGB (falsche Versicherung an Eides Statt) und in § 161 StGB (Fahrlässiger Falscheid, fahrlässige falsche Versicherung an Eides Statt) wiedergegeben.

§ 156 StGB: Falsche Versicherung an Eides Statt

Wer vor einer zur Abnahme einer Versicherung an Eides Statt zuständigen Behörde eine solche Versicherung falsch abgibt oder unter Berufung auf eine solche Versicherung falsch aussagt, wird mit Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft.

§ 161 StGB: Fahrlässiger Falscheid, fahrlässige falsche Versicherung an Eides Statt:

Abs. 1: Wenn eine der in den §§ 154 und 156 bezeichneten Handlungen aus Fahrlässigkeit begangen worden ist, so tritt Freiheitsstrafe bis zu einem Jahr oder Geldstrafe ein.
Abs. 2: Straflosigkeit tritt ein, wenn der Täter die falsche Angabe rechtzeitig berichtigt. Die Vorschriften des § 158 Absätze 2 und 3 gelten entsprechend.

Ich habe die Belehrung zur Eidesstattlichen Versicherung zur Kenntnis genommen.

St. Georgen, den 19. Juni 2017 _____

Simon Deusch