Non-targeted multimodal metabolomics and lipidomics data from ovine rumen fluid fractions

Nikola Palevich[1*], Paul H. Maclean[1], Arvind K. Subbaraj[2] and Mingshu Cao[1].

[1]AgResearch Ltd., Grasslands Research Centre, Palmerston North, 4442, New Zealand.

[2]AgResearch Ltd., Lincoln Research Centre, Christchurch, 7674, New Zealand.

*Corresponding author: E-mail: nik.palevich@agresearch.co.nz

Keywords

Metabolomics, Lipidomics, Ovine, Rumen fluid.

Abstract (250 words maximum)

Background: Metabolomics is a powerful and sensitive approach for investigating low molecular weight metabolite profiles present in rumen biofluids to identify potential roles of metabolites in rumen microbiome and understanding host-level regulatory mechanisms associated with animal production.

Findings: Rumen samples from sheep grazed on a mixed ryegrass and clover pasture diet, were fractionated based on molecular weight and analysed using metabolomics and lipidomics to detect the small molecules present in ovine rumen fluid fractions.

Conclusions: Untargeted metabolomics provides a detailed snapshot of the ovine ruminal fluid metabolome that can be used as reference for future studies on ovine rumen fluid or as a comparator for other ruminant species. All data and metadata are available for download in the MetaboLights database.

Data description

Ruminant livestock are an important component of feeding the growing human population while also being sources of global greenhouse gas emissions. Rumen microbiota breakdown and convert plant polysaccharides from feed into energy sources but also result in methane formation that affect ruminant productivity. As such, characterisation of the low molecular weight metabolites is key to improving our understanding of the rumen metabolome and our pursuit of developing a system-wide picture of rumen metabolism and biology. Metabolomics facilitates our ability to rapidly detect hundreds to thousands of metabolites within a single sample and enables accurate measurement of end-products of complex, genetic, epigenetic and environmental interactions. While rapid developments in genomics have accelerated our knowledge of rumen biology at the molecular level [1-8], there has been less work focusing on the low molecular weight molecules that stem from rumen fermentation of feed, and a complete absence of metabolomics studies on ovine rumen samples [9-11]. To date, only the bovine rumen fluid has been characterised using metabolomics [12-17].

Moreover, past studies on the rumen metabolome have used a 'freeze dry/grind/extract/inject' approach, which ignores the highly biodiverse ecology that is the rumen, with bacteria, fungi, higher order microbes, plant material at different stages of breakdown, all encased in a mammalian environment. As such, there is a need to address these major research and literature gaps as it is the rumen microbial communities that underlie variations in undesirable methane formation and conversion of feed to useful animal products. The presented dataset was initially collected to study the metabolic signatures of the ovine rumen with the purpose of increasing rumen digestibility of various forages and enhancing animal performance. However, significant challenges exist for analysing ruminal biochemistry and identifying metabolites within the chemically complex and heterogenous rumen fluid that have been shown to vary greatly based on dietary factors and host species [1, 10]. Also, an important consideration for metabolome studies is the impact of sample collection prior to extraction.

As such, we performed an analysis of the rumen metabolome using a newly developed *in vitro* system, using different molecular weight cut-off points that represent different fractions of the rumen *in vitro* and could lead to better understanding of how different components of the rumen interact. Studies were conducted with an *in vitro* simulation of the rumen fermentation using a simple *in vitro* artificial rumen system of permeable

continuous culture type which would simulate both the removal of end-products of fermentation and flow of ingesta. Experiments were made to determine the effects of dialysis on the fermentation and microbial population in the continuous culture. In addition we accounted for the macro components of rumen fluid such as large degraded plant material associated with the digestion of common fibrous feeds or associated microbial features and proteins, by obtaining filtrates from dialysis of rumen fluid through membranes with molecular weight cut-offs of 20 kDa, 8-10 kDa and 100 Da. These three fractions are expected to exclude large tannin-rich plants extracts and proteins respectively. We then performed metabolomic profiling of the enriched low molecular weight dialyzed rumen fluid (DRF) fractions to investigate biologically relevant molecules such as phenolics, phospholipids, amino acids, dicarboxylic acids, fatty acids, volatile fatty acids, glycerides, carbohydrates and cholesterol esters. In this study, an untargeted approach using multimodal methodologies including polar and semi-polar-retention chromatographies coupled to mass spectrometry, have been used to detect a wide-range of metabolites encompassing polar, semi-polar compounds and lipid species. A basic overview of the datasets and experimental design is shown in Figure 1.

Sample collection

Whole rumen content samples were collected post-mortem and pooled from five sheep grazing *ad libitum* on a ryegrass and clover pasture diet. Approximately 5 L of rumen contents were collected from each animal and filtered through 4 layers of cheesecloth (335 μm mesh) to remove plant material present in the digesta (Figure 1A). The pH of each sample was determined immediately after sampling using a pH meter (6.6). Filtered rumen fluid was immediately processed so as not to alter the fermentative capacity of the ruminal fluid. Sampling was conducted in December 2018 at AgResearch Grasslands Research Centre (Palmerston North, New Zealand) under the approval of the AgResearch Grasslands Animal Ethics Committee.

A method was developed to acquire DRF fractions that enrich for different sized components of rumen fluid for metabolomics and lipidomics analyses (Figure 1B). DRF fractions based on three molecular weight cut-offs (MWCO) were obtained using Spectra-Por® Float-A-Lyzer® G2 dialysis systems with MWCOs of 20 kDa (Z726931, Sigma-Aldrich), 8-10 kDa (Z726605, Sigma-Aldrich) and 100 Da (Z727253, Sigma-Aldrich). Briefly, the pooled and filtered RF contents were mixed and divided evenly into four Schott gas washing bottles fitted with Drechsel type head connections (GL 14, DURAN). The analytical conditions *in vitro* were

identical to the animal's physiological conditions, with a temperature of 39°C in anaerobic conditions obtained by insufflating a stream of O_2-free CO_2 inside the container with constant mixing. To obtain each DRF fraction, five replicates of each individual MWCO apparatus were dialyzed against 10mL of autoclaved phosphate buffered saline (1× PBS, 137 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4, and 2 mM KH_2PO_4, pH 7.4) buffer overnight at 39°C in a water bath. Additional containers were prepared in parallel containing only PBS and distilled water as control treatments. The principles of dialysis allowed small compounds to migrate from high concentration (rumen fluid contents) to low concentration (dialysis systems containing saline). DRF fractions and samples of digesta obtained from different phases of the process were snap-frozen in liquid nitrogen and kept on dry ice until long term storage at -80ºC.

Extractions for non-targeted metabolomic and lipidomic profiling

A combination of multiple metabolomics platforms, or 'multi-modal' strategy, were applied in parallel to the same batch of biological samples to facilitate interpretation and provide extensive coverage of the rumen metabolome. To comprehensively elucidate metabolites associated with DRF fractions we used: hydrophilic interaction liquid chromatography (HILIC) to separate polar compounds [18], ultra-high-performance liquid chromatography (UHPLC) with C18 chromatography to separate semi-polar compounds and a modified C18 phase (CSH-C18) for separation of lipids [19-20]. LC-MS analyses were done in both positive and negative electrospray ionization (ESI) modes.

Upon collection, the DRF fractions were snap-frozen in liquid nitrogen, transferred to glass vials and stored at -80°C until further use. Seven aliquots of 1 mL each (5 for analyses and two for quality control samples (QC)) of each sample were transferred into microcentrifuge tubes. The two QC samples (12) were pooled and solely used for monitoring sample degradation, tracking run-order effects within a batch, and quality control purposes only. Briefly, samples were thawed overnight at 4°C, centrifuged (4°C, 11,000 × *g*) for 10 min and 200 µL of supernatant transferred into a 2 mL micro-centrifuge tube. An extraction solvent comprising 800 µL of chloroform:methanol (1:1; v/v) was added and samples were vortexed (1 min). Sample was diluted with water (400 µL), again vortexed (1 min), and centrifuged (4°C, 11,000 × *g*) for 15 min.

To evaluate the lipidome of the DRF fractions, the lower, organic layer was taken (200 µL), evaporated to dryness under a continuous stream of nitrogen (30°C), and the dried extract was reconstituted in 200 µL of

chloroform:methanol (2:1; *v/v*), with 16:0 d_{31}-18:1 phosphatidylethanolamine (10 µg/mL) as an internal standard. Finally, samples were vortexed (1 min), and 100 µL was transferred to a glass insert in an auto-sampler vial for LC-MS analysis. For polar and semi-polar compounds, hydrophilic interaction liquid chromatography (HILIC) and C18 chromatography were applied, respectively [18-20]. For these analyses, supernatants (200 µL) were mixed with 800 µL of pre-chilled chloroform:methanol (1:1, v/v) containing 1.6 mg/L of internal standards; d5-Ltryptophan, d4-citric acid, d10-leucine, d2-tyrosine, d35-stearic acid, d5-benzoic acid, 13C2-glucose, and d7-alanine. Two aliquots of the upper aqueous layer (200 µL) was taken and evaporated as above, then reconstituted in 200 µL of the extraction solvents (acetonitrile:water containing 0.1% formic acid;1:1 for HILIC and 1:9 for C18, *v/v*).

Chromatography and mass spectrometry spectral acquisition

The chromatographic gradient and other conditions were selected to detect metabolites over a wide polarity range for the non-targeted LC-MS and lipid analyses as previously described [18-20]. For semi-polar compounds, C18 conditions were set as described [18], with extract (2 µL) injected into a 100 mm × 2.1 mm Thermo Hypersil Gold C18 column with 1.9 µm particle size and eluted over a 16 min gradient with a flow rate of 400 µL/min. The mobile phase was a mixture of water with 0.1% formic acid (solvent A), and acetonitrile with 0.1% formic acid (solvent B). Chromatographic gradient and other LC-MS conditions have been previously described [20-21]. For polar compounds, HILIC conditions were set as described [18], with extract (2 µL) injected onto a 100 mm × 2.1 mm ZIC-pHILIC column with 5 µm particle size and eluted over 17 min with solvent gradient from 97% solvent A (1 min), 97%-70% solvent A (1-12 min), 70-10% solvent A (12-14.5 min) to 10% solvent A (14.5-17 min). Mobile phase solvent A was a mixture of acetonitrile with 0.1% formic acid, solvent B was a mixture of water with 16 mM ammonium formate and flowrate was 250 µL/min. Chromatographic gradient and other LC-MS conditions have been previously described [18, 22].

C18 and HILIC column effluents were connected to a high-resolution mass spectrometer (Exactive Orbitrap™, ThermoFisher Scientific, Waltham, MA, USA) mass spectrometer with electrospray ionization, and lipid analysis was conducted on a Q-Exactive quadrupole-high resolution mass spectrometer (ThermoFisher Scientific, Waltham, MA, USA). Both full and data dependent MS^2 ($ddMS^2$) scans were

collected in profile data acquisition mode. For full scan mode, a mass resolution setting of 35,000 was set to record a mass range of *m/z* 200-2000 with a maximum trap fill time of 250 ms. In $ddMS^2$, MS^2 measurements are activated when a set peak intensity threshold is achieved. For $ddMS^2$ scan mode, the same mass resolution setting was maintained with a maximum trap fill time of 120 ms. The isolation window of selected MS^1 scans was ± 1.5 *m/z* with a normalized collision energy of 30. Samples were run in both positive and negative ionization modes separately. Positive ion mode parameters were as follows: spray voltage, 4.0 kV; capillary temperature, 275 °C; capillary voltage, 90 V; tube lens 120 V. Negative ion mode parameters were as follows: spray voltage, -2.5 kV; capillary temperature, 275 °C; capillary voltage, -90 V; tube lens, -100 V. The nitrogen source gas desolvation settings were the same for both modes (arbitrary units): sheath gas, 40; auxiliary gas, 10; sweep gas, 5.

Pooled sample from all conditioned DRF fraction extracts and internal standards were used as controls and samples were run in randomised order to avoid bias due to any inherent variation due to run order. Blank subtraction was applied after internal standard correction. To verify and maintain data quality, the QC sample was injected once every 10 samples. Retention time, signal intensity, and mass error of the internal standard were constantly monitored during the runs. Fragmentation data on approximately 4 samples in total per ionization mode (positive and negative) were used for identification of metabolite ions/classes.

Data quality, processing and analysis

The MS raw data files (Thermo .raw files) were converted to mzXML files using MSConvert function of ProteoWizard™ [23]. These files were uploaded to XCMS Online [24, 25] with suitable parameters for data processing including peak detection, retention time alignment, profile alignment, isotope annotation, grouping and gap filling. The type of adducts generated are dependent on the solvents and eluting conditions used. For this study, [M+H] and $[M+NH_4]$ adducts were selected for negative and positive ionisation modes, respectively. Finally, a retention time tolerance of 0.1 min, and mass error tolerance of ± 10 ppm was allowed. The filtered matrix was normalized by a QC based on LOESS signal correction (QC-RLSC) [26] and all subsequent *m/z* features and retention times with relative standard deviation (RSD) [27] > 0.3 were eliminated. The resultant data matrix was used for downstream statistical analyses and metabolite identification.

Statistical analysis and metabolite annotation/identification

Following the metabolomics data analysis pipeline [28] the raw data from the lipid stream (Figure 1C) were processed and subjected to peak detection, quality control, statistical analysis, and the annotation of the top-ranking peaks. Metabolite features were initially confirmed by matching source-induced fragmentation data against standard MS/MS spectra in the METLIN MS^2 spectral library using XCMS Online [24, 25]. The fold change (FC $\geq$ 2), *p*-value ($p < 0.1$), and intensity thresholds (10,000) were defined for peak ranking. The identity of significantly differential peaks were further determined from the exact mass composition using HMDB [29], LIPID MAPS [30], METLIN [31], LMDB [9], BMDB [32], PubChem and mzCloud databases. All the raw data and metadata reported in this study have been submitted to the MetaboLights database (www.ebi.ac.uk/metabolights) with the study identifier: MTBLS1717.

Because only sparse fragmentation spectra (MS^2) were collected for positive mode ions and no MS^2 for negative mode ions in this study we used the correlation structure of peaks for metabolite identification. Unless otherwise specified, we carried out MSI (Metabolomics Standards Initiative) level 2 metabolite annotation as evidenced by chromatographic behaviour, library search based on accurate mass (< 10 ppm error) and the match of isotopic peak intensity between experimental and theoretical spectrum. For peaks detected with C18 and HILIC chromatography, after the filtering of ^{13}C isotopic peaks and early eluting peaks (<1 min) were left for annotation in each ionization stream. Metabolite identification was based on the match of *m/z* ([M+H], [M+NH4] or [M-H], within 5 ppm accuracy) with those in HMDB and in-house retention time databases. In addition, we exploited the correlation structure of top peaks to assist the annotation. Correlated peaks due to coelution may indicate the similar physio-chemical properties of eluting metabolites, or in-source fragment ions [28]. The presence of in-source fragment peaks helps identify the molecular ion for correct annotation of the metabolite. On the other hand, peak correlation among biological samples may suggest the origin of candidate metabolites, providing additional information for annotation.

For identifying significant peaks (FDR corrected *p* value ≤ 0.05, FC ≥ 1) as potential ovine rumen fluid metabolome products, a multi-group comparison between the DRF fractions (20 kDa, 8-10 kDa, 100 Da) groups was performed (Figure 2). An initial prerequisite for our downstream comparative multimodal analyses was that a peak needed to be present in ≥ 3 of the 5 biological replicates across each of the metabolomics

analysis. Univariate hypothesis testing of the peaks was carried out using Permutation ANOVAs as implemented in the lmPerm R package version 2.1 [33] with 1 million permutations, after which, significance was determined for the peaks using the Benjamini-Hochberg FDR-corrected [34] *p*-values (FDR *p* value < 0.05). A total of 675 peaks for HILIC, 144 C18 peaks and 454 peaks for LIPID were identified (both positive and negative modes), with putative annotations were quantified with varying degree of confidence. Overall, a combined total of 1,454 peaks were identified with no putative annotations across all metabolomics analyses. This suggests that further correlation between metabolomics and transcriptomics data to aid in identifying these unknown features is required and is an area of active research.

Potential use

To date, the majority of published ruminant metabolomics work has focused on the processed bovine ruminal fluid metabolome [14-17]. This suggests that greater consideration needs to be given to what components are extracted for understanding the rumen, as otherwise we may be mainly extracting the degraded pasture metabolome that is destined for return to the pasture, rather than what is available for uptake by the ruminant. Our data provides the first view of the ovine rumen metabolome during *in vitro* continuous culture using multimodal metabolomics. Separating out these components has important implications for both animal production, health assessment, disease diagnosis, bioproduct characterization and biomarker discovery for desirable economic traits (e.g. feed efficiency and milk production). In particular, this dataset will serve as a valuable resource in greenhouse gas research for future mining of targets for interventions that target the rumen, including the search for potential vaccine candidates for methane mitigation strategies.

Future studies can employ various comprehensive bioinformatics tools to mine the presented raw data for any metabolites that may be present in the ovine rumen, rather than focus solely on the top list of peaks. Also, future efforts can employ more updated and curated databases such as BMDB [32] where Bovine rumen metabolites and spectral data are archived. In addition, data can also be compiled for the other common livestock species namely goats, horses and pigs, to compose an open access, comprehensive livestock metabolome database focused on extremely low molecular weight metabolites. When coupled with the advent of sequencing technologies and availability of genomic and transcriptomic datasets, is a comprehensive

approach for the identification of these peaks and characterization of the ovine rumen fluid molecular mechanisms.

Different molecular weight cut-offs to separate out different components of the rumen led to different metabolite and lipid profiles. While this finding is not surprising, it does suggest that fractionating rumen fluid prior to extraction and analysis could give insights into metabolic activity occurring in relation to the solid residue present compared to the more active liquid matrix. Biologically this may be due to the effects of dialysis on fermentation patterns and subsequent shifts in microbial population during continuous culture. For example, accumulation of particular fermentation end-products and variations in the composition of volatile fatty acids (VFAs) can rapidly inhibit microbial growth within the rumen. Further work is needed to establish whether higher molecular weight fractions can provide insight into the degree of breakdown of plant material.

This study is the first report of the ovine rumen metabolome, and while due to the collection method used is not directly comparable to previously published bovine rumen metabolomes, represents an important reference dataset. Our dataset should enable microbiologists and livestock researchers to conduct more targeted metabolomic studies to identify low molecular weight features where further metabolome coverage is required. Further research will be done on this prototype model, including additional data mining and annotation of the metabolite features and investigation into the microbial community profiles within the different dialysis contents using next-generation sequencing technology, as these microbiome signatures may have critical implications into various livestock metabolomics applications.

Availability of supporting data and materials

Supporting data and corresponding metadata are available in the MetaboLights database [MTBLS1717].

Abbreviations

m/z: Mass-to-charge ratio; MS: Mass spectrometry; DRF: Dialyzed rumen fluid; MWCO: Molecular weight cut-offs.

Competing interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Authors' contributions

NP conceived the project, designed and conducted all experiments, analyzed and interpreted the data, and wrote the manuscript. PHM provided statistical and bioinformatics support for the project. AKS and MC provided guidance with metabolomics and lipidomics analysis. All authors critically revised and approved the final manuscript.

Acknowledgments

Special thanks to Linley Schofield in the Rumen Microbiology lab for providing the dialysis equipment. We thank Hailey Gillespie and Trevor Holloway for their timely assistance with the collection of rumen fluid for this study. We thank Alastair Ross and Karl Fraser for their revisions on the manuscript. This work was supported by the AgResearch Ltd Strategic Science Investment Fund (SSIF, Grant No. C10X1702) to NP. This work was also supported in part by the Agricultural and Marketing Research and Development Trust (AGMARDT) Postdoctoral Fellowship Programme (Grant No. P17001) to NP.

References

1. Henderson G, Cox F, Ganesh S, Jonker A, Young W, Collaborators GRC, et al. Rumen microbial community composition varies with diet and host, but a core microbiome is found across a wide geographical range. Scientific reports. 2015;5 14567:1-13.

2. Seshadri R, Leahy SC, Attwood GT, Teh KH, Lambie SC, Cookson AL, et al. Cultivation and sequencing of rumen microbiome members from the Hungate1000 Collection. Nature Biotechnology. 2018;36 4:359-67.

3. Nayfach S, Roux S, Seshadri R, Udwary D, Varghese N, Schulz F, et al. A genomic catalog of Earth's microbiomes. Nature Biotechnology. 2021;39 4:499-509.

4. Borges DGL, Echeverria JT, de Oliveira TL, Heckler RP, de Freitas MG, Damasceno-Junior GA, et al. Discovery of potential ovicidal natural products using metabolomics. PLOS ONE. 2019;14 1:e0211237.

5. Palevich N, Kelly WJ, Leahy SC, Altermann E, Rakonjac J and Attwood GT. The complete genome sequence of the rumen bacterium *Butyrivibrio hungatei* MB2003. Standards in Genomic Sciences. 2017;12 1:72.

6. Palevich N, Kelly WJ, Ganesh S, Rakonjac J and Attwood GT. *Butyrivibrio hungatei* MB2003 competes effectively for soluble sugars released by *Butyrivibrio proteoclasticus* B316^{T} during growth on xylan or pectin. Applied and Environmental Microbiology. 2019;85 3:e02056-18.

7. Palevich N, Kelly WJ, Leahy SC, Denman S, Altermann E, Rakonjac J, et al. Comparative genomics of rumen *Butyrivibrio* spp. uncovers a continuum of polysaccharide-degrading capabilities. Applied and environmental microbiology. 2019;86 1.

8. Palevich N, Maclean PH, Kelly WJ, Leahy SC, Rakonjac J and Attwood GT. Complete genome sequence of the polysaccharide-degrading rumen bacterium *Pseudobutyrivibrio xylanivorans* MA3014 reveals an incomplete glycolytic pathway. Genome Biology and Evolution. 2020;12 9:1566-72.

9. Goldansaz SA, Guo AC, Sajed T, Steele MA, Plastow GS and Wishart DS. Livestock metabolomics and the livestock metabolome: A systematic review. PLOS ONE. 2017;12 5:e0177675.

10. Huws SA, Creevey CJ, Oyama LB, Mizrahi I, Denman SE, Popova M, et al. Addressing global ruminant agricultural challenges through understanding the rumen microbiome: past, present, and future. Frontiers in microbiology. 2018;9:2161.

11. Mao SY, Huo WJ and Zhu WY. Microbiome-metabolome analysis reveals unhealthy alterations in the composition and metabolism of ruminal microbiota with increasing dietary grain in a goat model. Environmental Microbiology. 2016;18 2:525-41.

12. Saleem F, Ametaj B, Bouatra S, Mandal R, Zebeli Q, Dunn S, et al. A metabolomics approach to uncover the effects of grain diets on rumen health in dairy cows. Journal of Dairy Science. 2012;95 11:6606-23.

13. Li H, Yu Q, Li T, Shao L, Su M, Zhou H, et al. Rumen microbiome and metabolome of Tibetan Sheep (Ovis aries) reflect animal age and nutritional requirement. Frontiers in Veterinary Science. 2020;7.

14. Saleem F, Bouatra S, Guo AC, Psychogios N, Mandal R, Dunn SM, et al. The bovine ruminal fluid metabolome. Metabolomics. 2013;9 2:360-78.

15. Artegoitia VM, Foote AP, Lewis RM and Freetly HC. Rumen fluid metabolomics analysis associated with feed efficiency on crossbred steers. Scientific reports. 2017;7 1:1-14.

16. De Almeida RTR, Do Prado RM, Porto C, Dos Santos GT, Huws SA and Pilau EJ. Exploring the rumen fluid metabolome using liquid chromatography-high-resolution mass spectrometry and Molecular Networking. Scientific reports. 2018;8 1:1-8.

17. Sun H-Z, Wang D-M, Wang B, Wang J-K, Liu H-Y, Guan LL, et al. Metabolomics of four biofluids from dairy cows: potential biomarkers for milk production and quality. Journal of proteome research. 2015;14 2:1287-98.

18. Pahalagedara ASNW, Flint S, Palmer J, Subbaraj A, Brightwell G, Gupta TB. Antimicrobial Activity of Soil Clostridium Enriched Conditioned Media Against *Bacillus mycoides*, *Bacillus cereus*, and *Pseudomonas aeruginosa*. Frontiers in Microbiology. 2020;11.

19. Su M, Subbaraj AK, Fraser K, Qi X, Jia H, Chen W, et al. Lipidomics of Brain Tissues in Rats Fed Human Milk from Chinese Mothers or Commercial Infant Formula. Metabolites. 2019;9 11:253.

20. Subbaraj AK, Huege J, Fraser K, Cao M, Rasmussen S, Faville M, et al. A large-scale metabolomics study to harness chemical diversity and explore biochemical mechanisms in ryegrass. Communications biology. 2019;2 1:1-17.

21. Fraser K, Harrison SJ, Lane GA, Otter DE, Hemar Y, Quek S-Y, et al. Non-targeted analysis of tea by hydrophilic interaction liquid chromatography and high resolution mass spectrometry. Food chemistry. 2012;134 3:1616-23.

22. Fraser K, Lane GA, Otter DE, Hemar Y, Quek S-Y, Harrison SJ, et al. Analysis of metabolic markers of tea origin by UHPLC and high resolution mass spectrometry. Food Research International. 2013;53 2:827-35.

23. Holman JD, Tabb DL and Mallick P. Employing ProteoWizard to convert raw mass spectrometry data. Current protocols in bioinformatics. 2014;46 1:13.24. 1-13.24. 9.

24. Forsberg EM, Huan T, Rinehart D, Benton HP, Warth B, Hilmers B, et al. Data processing, multi-omic pathway mapping, and metabolite activity analysis using XCMS Online. Nature protocols. 2018;13 4:633.

25. Smith CA, Want EJ, O'Maille G, Abagyan R and Siuzdak G. XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. Analytical chemistry. 2006;78 3:779-87.

26. Dunn WB, Wilson ID, Nicholls AW and Broadhurst D. The importance of experimental design and QC samples in large-scale and MS-driven untargeted metabolomic studies of humans. Bioanalysis. 2012;4 18:2249-64.

27. Parsons HM, Ekman DR, Collette TW and Viant MR. Spectral relative standard deviation: a practical benchmark in metabolomics. Analyst. 2009;134 3:478-85.

28. Cao M, Fraser K, Jones C, Stewart A, Lyons T, Faville M, et al. Untargeted metabotyping *Lolium perenne* reveals population-level variation in plant flavonoids and alkaloids. Frontiers in plant science. 2017;8:133.

29. Wishart DS, Feunang YD, Marcu A, Guo AC, Liang K, Vázquez-Fresno R, et al. HMDB 4.0: the human metabolome database for 2018. Nucleic acids research. 2018;46 D1:D608-D17.

30. Fahy E, Sud M, Cotter D and Subramaniam S. LIPID MAPS online tools for lipid research. Nucleic acids research. 2007;35 suppl_2:W606-W12.

31. Guijas C, Montenegro-Burke JR, Domingo-Almenara X, Palermo A, Warth B, Hermann G, et al. METLIN: a technology platform for identifying knowns and unknowns. Analytical chemistry. 2018;90 5:3156-64.

32. Foroutan A, Fitzsimmons C, Mandal M, et al. The Bovine Metabolome. Metabolites. 2020;5 10:E233.

33. Wheeler B, Torchiano M and Torchiano M. Package 'lmPerm'. R package version, 2.1. 0. 2016.

34. Benjamini Y and Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal statistical society: series B (Methodological). 1995;57 1:289-300.

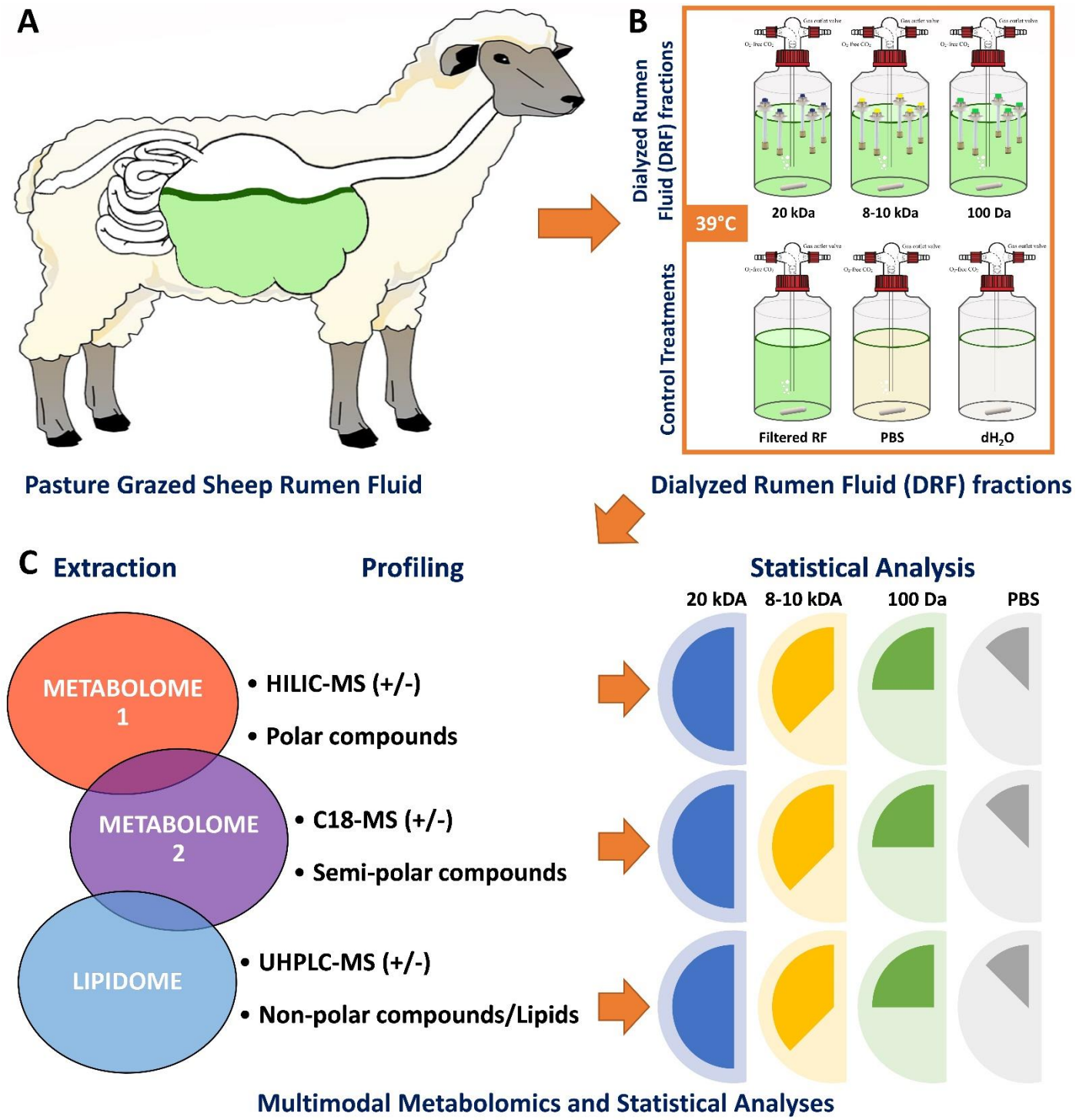

Figure 1. Overview of the experimental design and multimodal metabolomics workflow. (A) New Zealand pasture-fed sheep used for this study. **(B)** Dialyzed rumen fluid (DRF) fractions were obtained under anaerobic rumen conditions (39°C and CO_2) using dialysis systems at three molecular weight cut-offs. **(C)** Schematic of a multimodal metabolomics workflow and statistical analyses used to process data integrated from multiple analytical approaches.

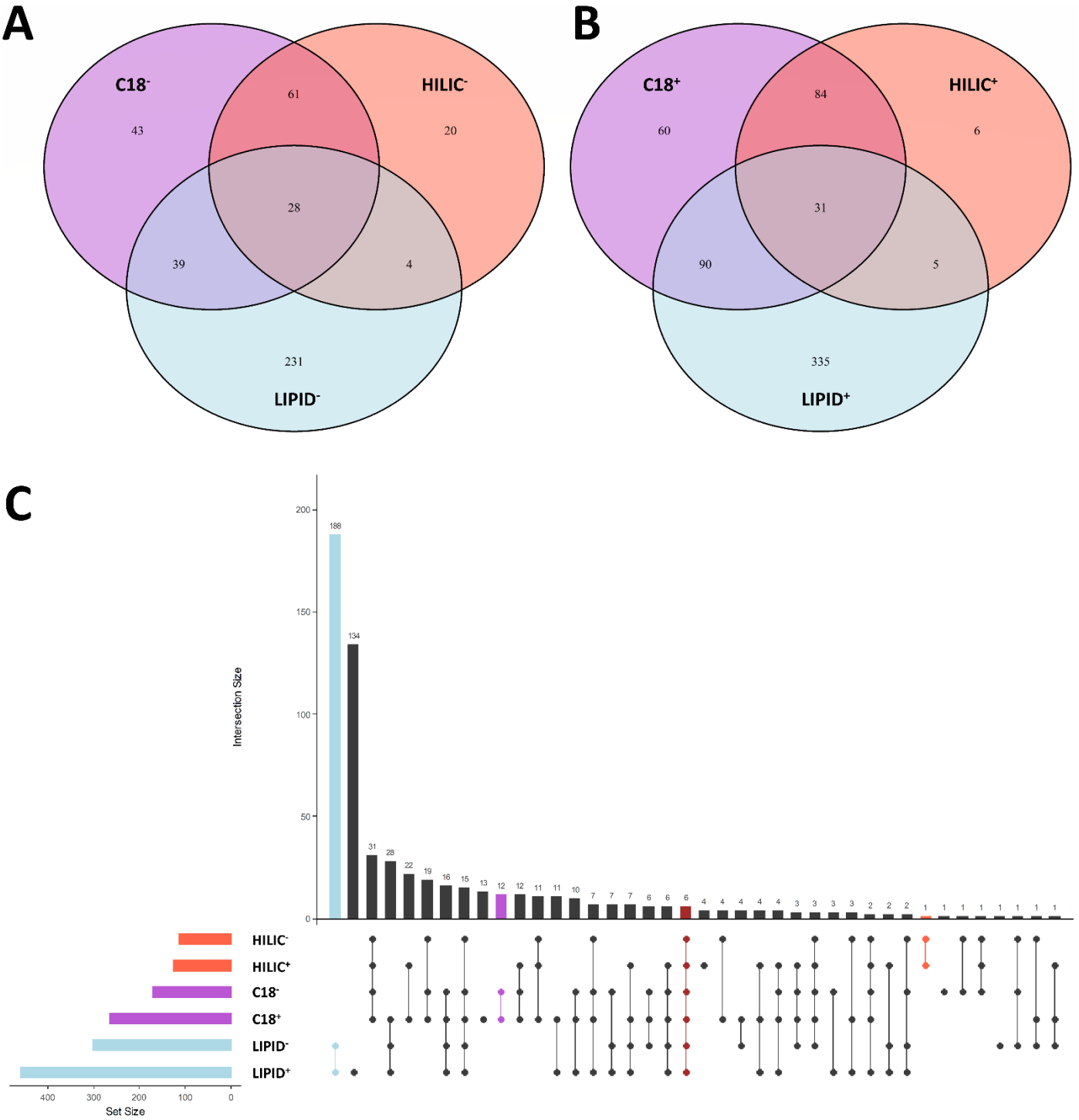

Figure 2. Unique and mutual peaks identified across all metabolomics streams. Venn diagrams compiling the Polar (HILIC), semi-polar (C18) and lipids (LIPID) molecular features determined from all DRF fractions for the negative **(A)** and positive **(B)** streams. UpsetR barchart analysis **(C)** showing the presence and numbers of quantified molecular features determined for different metabolomics streams (i.e. HILIC (orange), C18 (purple) and LIPID (blue). Connected dots display shared molecular features between or among metabolomics streams, and the total number of features in a particular metabolomics stream is shown in the set size. Coloured connected dots indicate the molecular features identified in both the negative (-) and positive (+) extractions of a metabolomics stream.