Technical report: In-gel sample preparation prior to proteomic analysis of bovine faeces increases protein identifications by removal of high molecular weight glycoproteins.

Yixin Huang[1], Richard Burchmore[2], Nicholas N. Jonsson[1], Paul C. D. Johnson[1], P. David Eckersall[1,*]

[1]*Institute of Biodiversity, Animal Health & Comparative Medicine, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow G61 1QH, UK*

[2] *Institute of Infection, Immunity & Inflammation and Glasgow Polyomics, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow G61 1BD, UK*

** Corresponding author: P. David Eckersall; e-mail: David.Eckersall@glasgow.ac.uk*

Abstract: Bovine faecal composition is complex and a knowledge gap exists in the understanding of the bovine faecal proteome. In the present study, in-gel sample preparation (IGSP) of faecal samples prior to proteomics showed an increase in the number of proteins identified in faecal samples compared to those processed by filter-aided sample preparation (FASP). The optimised sample preparation method removed high molecular weight glycoproteins as part of the clean-up process of the faecal samples, and in combination with in-gel digestion before liquid chromatography with tandem mass spectrometry (LC-MS/MS) The use of IGSP led to enhanced protein identification with increases in the number of peptides identified and in the percent coverage of proteins in the bovine faecal samples.

Significance*:* Characterization of faecal proteins has the potential to increase our understanding of host responses to changes such as diet, disease and drug-treatment. In-gel sample preparation prior to proteomics can be used to remove high molecular weight glycoproteins and reduce protein/peptide loss in FASP. This method of sample preparation will have application not only in the investigation of bovine faecal extracts but also in studies where large molecules such as glycoproteins or oligosaccharides could have detrimental influences on sample preparation involving ultrafiltration.

Highlights

- The interference of high molecular weight glycoproteins on preparation of faecal samples has been eliminated.
- In gel sample preparation increased the number of proteins identified and the number of peptides found per protein.
- The bovine faecal proteome is a complex mix of protein from the animal, its feed and ruminal and intestinal bacteria.

Keywords: bovine faeces, mass spectrometry, in-gel sample preparation, filter-aided sample preparation

Filter-aided sample preparation (FASP) is a common proteomic sample preparation method for the generation of tryptic peptides prior to nano liquid chromatography and mass spectrometry (LC-MS/MS) [1]. However, in preliminary experiments we obtained an unexpectedly low number of protein identifications when using FASP to prepare bovine faecal samples for LC-MS/MS. We hypothesised that the low yield of known proteins was a consequence of the complexity of the sample matrix and the potentially extended period over which the matrix components were able to react.

Faeces is a complex, heterogeneous, mixture of compounds that includes proteins from diet, host and microbiome, with a huge range of small, potentially chemically active molecules. Faecal samples comprise components from upper (proximal) and lower (distal) gastrointestinal tract, which have been allowed to interact with each other at body temperature for a period of time that could range from a few minutes in the case of rectal mucus, to a day or more, for the ingested feed components.

Modifications of the preparation procedures for faecal samples were therefore investigated to increase the protein identifications made by LC-MS/MS analysis. Possible causes of the low protein identification using FASP were also investigated. In assessing the protein composition of bovine faecal samples by SDS-PAGE, gels were stained with Coomassie blue for protein or with the periodic acid-Schiff (PAS) stain for glycoprotein. All bovine faecal samples were found to contain PAS-staining material with molecular weights (MW) greater than 200 kDa. The presence of these high MW glycoproteins could have had detrimental effects on the preparation of the faecal samples by FASP for proteomic analysis by LC-MS/MS. An in-gel sample preparation (IGSP) method to remove the high MW glycoproteins before trypsin digestion in the gel was developed and compared to FASP in terms of the number of protein identifications made following either of the preparation methods. The influence of gel pieces on digestion efficiency and peptide recovery was minimized by using the modified method from Goldman *et al*. [2].

As part of an ongoing investigation, faecal samples were obtained from two groups of healthy beef cattle that were fed a diet composed of mostly barley cereal grains in which the barley had been treated with either ammonia or a preservative. Fresh faecal samples from each group were collected from the floor following observation of defaecation. Samples were refrigerated immediately after collection, transported directly to the laboratory, and stored at -80˚C until use. The method of faecal protein extraction was modified based on previous studies [3, 4]. Briefly, on thawing at room temperature, 3 g of faeces were mixed with 12 mL of sample buffer (50 mM Tris-HCl, 150 mM NaCl, 0.1% SDS, pH 7.4). One tablet of protease inhibitor (Roche Diagnostics, US) was added to every 25 mL of the sample buffer, followed by disruption by a homogenizer stomacher until there were no hard pellets remaining. The samples were centrifuged at 400 × g at 4 ˚C for 30 min, the supernatant was collected and sonicated on ice using an ultrasonic liquid processor (VC-130, Sonics & Materials, US) at 80% of amplitude for seven times of 5 s run interspersed with 10 s cool down. Samples were then centrifuged at 14,000 × g at 4 ˚C for 30 min. The supernatant was concentrated using an Amicon Ultra-15 centrifugal filter unit (10 kDa cut-off) (Merck & Co., US). The total concentration of protein of each sample was measured by the BCA method (Thermo Scientific, UK) with bovine serum albumin as standard. For this assessment of the benefit of IGSP two faecal extracts (one from each group) were

prepared by FASP and by IGSP, with results of protein identifications combined for each sample preparation method.

Periodic acid-Schiff staining is widely used in histochemistry and histological studies to show the presence of carbohydrates and carbohydrate-containing compounds. The presence of glycoproteins in a selection of the bovine faecal samples was determined by PAS stain method, modified from Segrest and Jackson [5]. Briefly, faecal samples, along with a bovine serum sample (from our previous study [6]) as control material, which has up to 50% of proteins being glycosylated [7], were loaded twice, on left- and right-hand sides of a 4-12% Bis-Tris gel (Invitrogen, UK) and run for 35 min at 200 V. Half of the gel was fixed in the fixative solution (40%, v/v ethanol with 5% v/v acetic acid) overnight and stained in the periodic acid solution (0.7% v/v periodic acid with 5% v/v acetic acid) for 2 h. Following washing in the sodium metabisulfite solution (0.2% w/v sodium metabisulfite with 5% v/v acetic acid) for 3 h (the solution was changed every 30 min), the gel was then stained in Schiff's fuchsin-sulfite reagent (Sigma, UK) overnight. The other half of the gel was stained in 0.1% w/v G250 Coomassie blue (Sigma, UK) for one hour and de-stained in 7.5% acetic acid with 20% methanol overnight (Figure 1a). Comparing the PAS and Coomassie blue-stained gels showed that the bovine faecal samples contained a high abundance of glycoproteins, the majority of which had MW higher than 190 kDa or lower than 10 kDa. This raised the possibility that an in-gel clean-up method could be used by excision of proteins within 10-190 kDa to exclude the highly abundant high MW glycoproteins from further proteomic analysis.

The use of single percentage polyacrylamide gels (10%, Invitrogen, UK) was shown to better restrict entry of the high MW glycoproteins into the gel and help sample preparation. Running the gel for only a short time (4 min) before staining with Coomassie blue meant that all stained proteins could be included in a narrow gel section of 3-5 mm (Figure 1c) and enabled excision of all proteins from 10-190 kDa in this one gel piece [2]. This process therefore concentrated the proteins into a single band, eliminating the high MW glycoproteins from further processing and at the same time minimizing the gel volume for in-gel digestion. Two bovine faecal samples were processed by IGSP (removal of the high MW glycoproteins in combination with in-gel digestion). Three replicates of each sample were run in separate tracks of the 10% polyacrylamide gels. The Coomassie blue-stained portions of each sample track of the gel were excised and processed by in-gel digestion, modified from Shevchenko et. al [8]. Briefly, excised gels were cut into small pieces with the gel pieces of the three replicates of each sample being pooled, followed by washing in 500 µL of 100 mM ammonium bicarbonate (ABC) and 500 µL of 50% of acetonitrile in 100 mM ABC for 30 min on a shaker, respectively. Samples were reduced in 10 µL of 45 mM dithiothreitol in 150 µL of 100 mM ABC at 60 ˚C for 30 min and were then

alkylated in 20 µL of 100 mM iodoacetamide in the dark for 30 min. Samples were washed in 500 µL of 50% acetonitrile in 100 mM ABC for 30 min on a shaker, shrunk by acetonitrile and were then completely dried down in a vacuum centrifuge. Samples were incubated in 120 µL of 0.05 µg/µL trypsin in 25 mM ABC overnight. The supernatant was collected and the rest of the gel pieces were submerged in 40 µL of 5% formic acid for 20 min on a shaker, and incubated with 80 µL of 5% acetonitrile for another 20 min. The supernatant was pooled with the previously collected supernatant, and were dried down in a vacuum centrifuge. For comparison, the same two bovine faecal samples were processed by FASP method. One hundred µg of proteins from each sample were mixed with 5 µL of SDT-lysis buffer (4% SDS, 100 mM dithiothreitol (DTT), 100 mM Tris-HCl, pH 7.6) and were moved to the filter unit (10,000 MWCO, Expedeon, UK). Two hundred µL of urea buffer (8 M urea in 100 mM Tris-HCl, pH 8.5) were added to the samples and were centrifuged at 13,000 rpm for 15 min (repeated). Following an addition of 100 µL of 50 mM iodoacetamide (in urea buffer), samples were incubated in darkness for 20 mins. Samples were centrifuged at 13,000 rpm for 10 min and the flow-through was discarded. Samples were washed with 100 µL of urea buffer and 100 µL of 50 mM ABC three times respectively, for 15 min each at 13,000 rpm. Each sample was digested by 1 µg of trypsin (in 50 mM ABC) overnight at 37 ˚C. The filter units with digested samples were transferred into new eppendorfs and samples were collected by centrifuge (13,000 rpm for 10 min). Fifty µL of 10% acetonitrile were added to each sample and centrifuged at 13,000 rpm for 10 min, the flow-through was collected and pooled with the previously collected digested samples. Samples were mixed with 1 µL of 1 % trifluoroacetic acid, and were dried down in a vacuum centrifuge. Five µg of peptide samples prepared from both methods were analysed by a nanoflow ultrahigh-performance liquid chromatography-electrospray ionisation-tandem mass spectrometry (nUHPLC-ESI-MS/MS). The peptides were solubilized in 20 µL 5% acetonitrile with 0.5% formic acid using the auto-sampler of a nanoflow uHPLC system (RSLCnano, Thermo Scientific, UK) and were detected online by ESI-MS with an Orbitrap Elite MS (Thermo Scientific, UK). Ionisation of LC eluent was performed by interfacing the LC coupling device to an NanoMate Triversa (Advion Bioscience) with an electrospray voltage of 1.7 kV. An injection volume of 5 µL of the reconstituted protein digest was desalted and concentrated for 10 min on a C18 trap column (Acclaim PepMap C18 100Å 0.3 × 5 mm, 5µM particle size, Thermo Scientific, UK) using a flow rate of 25 µL/min with 1% acetonitrile with 0.1% formic acid. Peptide separation was performed on an Acclaim PepMap C18 100Å phase column (50 cm × 75 µm, particle size 3 µm, Thermo Scientific, UK) using a solvent gradient at a fixed solvent flow rate of 0.3 µL/min for analytical column. The solvent composition was (A) 0.1% formic acid and (B) 0.08% formic acid in 80% acetonitrile. The solvent gradient was 4% B for 12 min, 4 to 60% B for 90 min, 60 to 99% B for 14 min and held at 99% B for 5 min. A further 9 min at initial conditions for column re-equilibration was used before the next

injection. The Orbitrap Elite acquires a full-scan MS in the range 300 to 2,000 m/z for a high-resolution precursor scan at 60,000 RP (at 400 m/z), while simultaneously acquiring up to the top 15 precursors which are isolated at 0.7 m/z width and subjected to CID fragmentation (35% NCE) in the linear ion trap using rapid scan mode. Singly charged ions are excluded from selection, while selected precursors are added to a dynamic exclusion list for 30 s.

Raw data generated by LC-MS/MS were imported into Proteome Discoverer (version 2.4, Thermo Scientific, UK). Faecal samples contain proteins from the animal, from the plant-based diet (mainly barley plus other species) and microorganisms ingested with the diet or resident in the gastrointestinal tract. Therefore, the data were assessed using Sequest HT engine to interrogate sequences in the Swissprot databases of bovine, barley and bacterial proteins. For the latter the database search focused on known ruminal and faecal microorganisms (*Ruminococcaceae, Lachnospiraceae, Clostridiaceae, Prevotellaceae, Bacterioidaceae, Spirochaetaceae*) [9]. Precursor mass tolerance was set as 10 ppm and the fragment mass tolerance as 0.6 Da. Carbamidomethylation of cysteine was specified as fixed modification, and oxidation of methionine, deamidation of asparagine/glutamine and acetylation of lysine and N-term were set as dynamic modifications. Data organization and graphing (package *ggVennDiagram*) were performed in R software (version 4.0.3) [10].

Comparisons between the use of FASP and IGSP for the identified bovine and bacterial proteins (master proteins with at least two unique peptides) are shown in Venn diagrams (Figure 2), and comparisons for all the master proteins are shown in Venn diagrams in Supplementary Figure 1. Data for analysis combined results of the two samples that were either processed by FASP or IGSP method. For each of the databases interrogated, more proteins were identified using the IGSP method than the FASP method. For the identified master proteins with at least two unique peptides, the number of bovine proteins increased around four-fold from 14 with FASP to 57 by IGSP, while increasing similarly from 18 to 63 proteins for the bacterial proteins identified. For the barley proteins in the faeces, only serpin-Z4, serpin-Z7 and alpha-amylase/trypsin inhibitor CMb were found in samples that were digested by FASP. In comparison, another five barley proteins (alpha-amylase/trypsin inhibitor CMd, alpha-amylase inhibitor BDAI-1, alpha-amylase inhibitor BMAI-1, phytepsin and signal recognition particle 54 kDa protein 3) were found in the samples that were prepared by IGSP. As examples of these findings, Table 1 shows a selection of 38 identified proteins separated by origin (barley, bovine or bacteria) and by presence in samples prepared by FASP or IGSP. The proteins being selected on the basis of the number of peptides identified in the IGSP groups. It was noticeable that with IGSP, not only were more proteins identified but that for proteins identified when prepared by both methods, the number of peptides and % coverage for each protein was greater in samples

prepared by IGSP. For example for the barley protein serpin-Z4, the number of peptides increased from 4 to 7 and coverage increased from 11% to 25%, for bovine protein alpha-2 macroglobulin, the number of peptides increased from 5 to 28 with the % coverage increasing from 4% to 26% and for bacterial protein phosphoenolpyruvate carboxykinase (*Agathobacter rectalis*), the number of peptides increased from 3 to 9 with the % coverage increasing from 7% to 24%. Of the 19 proteins of Table 1 that were identified in proteins when prepared by both FASP and IGSP, the number of peptides increased by an average of 2.9 times and the coverage by an average of 3.5 times when IGSP was used. The full list of proteins with at least two peptides identified in each of the two methods is given in Supplementary Table S1.

There were four bacterial proteins shown in Table 1 that were only found in samples prepared by FASP. These proteins could have been partially digested to peptides in the intestine but held by non-covalent, protein-protein interaction in large molecular complexes which would be retained by the filter in FASP. In contrast, in the IGSP the complexes would be broken down by the presence of SDS and reducing agent and the released peptides, migrating with the dye front in SDS-PAGE would not be included in the gel piece excised prior to trypsinisation. The loss of identification of such proteins is a limitation of the method but greatly outweighed by the many more proteins identified with IGSP.

In-gel sample preparation, by running samples for a short-distance in the polyacrylamide gel and excising the entire protein staining bands prior to in-gel digestion [11], combined with bottom-up proteomics analysis, had been reported to identify with a large number of high-confidence peptides and proteins in human cell lines [12], and to increase the depth of analysis of plasma sample, which have a large dynamic range of protein abundances. In the present study, the IGSP method not only avoided the influence of MS-incompatible detergents, buffers or salts [2], but also removed high-abundance and high MW glycoproteins that may affect the identification of lower-abundance proteins. The recognition of the role of high MW glycoproteins in poor protein identification using the FASP method and the ability of IGSP to overcome this problem can contribute to proteomic investigation of faecal samples and other samples where this might occur. Although this study compared IGSP to FASP in the preparation of two samples, the use of IGSP has consistently given a greater yield of protein identification when used for preparation of multiple faecal extract samples for quantitative proteomic study.

The high MW glycoproteins are probably intestinal mucins, but remain to be identified, and could have detrimental effects on the use of FASP in at least two ways. The presence of such large molecules can block the pores in the filtration devices used such that the efficiency of filtration is greatly reduced leading to loss of trypsin-digested peptides for MS analysis. Furthermore, mucin has an inhibitory

effect on protein digestion by trypsin as one of its important roles in the intestine is to resist endogenous proteases such as trypsin in order to protect the intestine while food is digested [13]. Mucins remain intact in the presence of digestive enzymes so will have a similar effect on the trypsin used in FASP. Removal of the mucin by use of IGSP overcomes both obstacles to the proteomic study of bovine faecal samples. The identification of plant-based diet (barley) proteins in the faeces provided new areas for the study of animal digestion and absorption. Interestingly, the serine protease inhibitors identified in bovine faecal samples, serpin-Z4 and serpin-Z7 also survive through malting and brewing in the beer making process, and are proteins found in beer froth [14, 15]. The effects of plant-derived protease inhibitors on protein digestibility have aroused interest in human gastrointestinal health studies [16]. In conclusion, the use of IGSP in proteomics improved protein identification in bovine faeces compared to proteomics based on FASP method and could benefit future studies in quantitative protein investigations of bovine faeces.

Acknowledgement

We are grateful for support from Suzanne McGill and Stefan Weidt from Glasgow Polyomics, University of Glasgow, and Nicola Brady from Institute of Biodiversity, Animal Health & Comparative Medicine, University of Glasgow.

Conflict of interests

The authors declare no conflict of interests.

References

[1] J.R. Wisniewski, A. Zougman, N. Nagaraj, M. Mann, Universal sample preparation method for proteome analysis, Nat Methods 6(5) (2009) 359-363.

[2] A.R. Goldman, L.A. Beer, H.Y. Tang, P. Hembach, D. Zayas-Bazan, D.W. Speicher, Proteome analysis using Gel-LC-MS/MS, Curr Protoc Protein Sci 96(1) (2019) e93.

[3] G. Debyser, B. Mesuere, L. Clement, J. Van de Weygaert, P. Van Hecke, G. Duytschaever, M. Aerts, P. Dawyndt, K. De Boeck, P. Vandamme, B. Devreese, Faecal proteomics: A tool to investigate dysbiosis and inflammation in patients with cystic fibrosis, J Cyst Fibros 15(2) (2016) 242-250.

[4] X. Zhang, L. Li, J. Mayne, Z. Ning, A. Stintzi, D. Figeys, Assessing the impact of protein extraction methods for human gut metaproteomics, J Proteomics 180 (2018) 120-127.

[5] J.P. Segrest, R.L. Jackson, Molecular weight determination of glycoproteins by polyacrylamide gel electrophoresis in sodium dodecyl sulfate, Methods Enzymol. 28 (1972) 54-63.

[6] R. Turk, N. Rosic, J. Kules, A. Horvatic, A. Gelemanovic, A. Galen, B.B. Ljubic, M. Benic, V. Stevanovic, V. Mrljak, C.C. Chadwick, P.D. Eckersall, Milk and serum proteomes in subclinical and clinical mastitis in Simmental cows, J Proteomics 244 (2021) 104277.

[7] S. Fanayan, M. Hincapie, W.S. Hancock, Using lectins to harvest the plasma/serum glycoproteome, Electrophoresis 33(12) (2012) 1746-54.

[8] A. Shevchenko, H. Tomas, J. Havlis, J.V. Olsen, M. Mann, In-gel digestion for mass spectrometric characterization of proteins and proteomes, Nat Protoc 1(6) (2006) 2856-2860.

[9] D.B. Holman, K.E. Gzyl, A meta-analysis of the bovine gastrointestinal tract microbiota, FEMS Microbiol Ecol 95(6) (2019) fiz072.

[10] R.C. Team, R: A language and environment for statistical computing, Vienna, Austria, 2020.

[11] L.A. Beer, B. Ky, K.T. Barnhart, D.W. Speicher, In-depth, reproducible analysis of human plasma using IgY 14 and SuperMix Immunodepletion, Methods Mol Biol 1619 (2017) 81-101.

[12] A.M. Gholami, H. Hahne, Z. Wu, F.J. Auer, C. Meng, M. Wilhelm, B. Kuster, Global proteome analysis of the NCI-60 cell line panel, Cell Rep 4(3) (2013) 609-620.

[13] M. E. V. Johansson, H. Sjövall, G.C. Hansson, The gastrointestinal mucus system in health and disease, Nat Rev Gastroenterol Hepatol 10(6) (2013) 352-361.

[14] T. Iimure, N. Nankaku, N. Hirota, Z. Tiansu, T. Hoki, M. Kihara, K. Hayashi, K. Ito, K. Sato, Construction of a novel beer proteome map and its use in beer quality control, Food Chemistry 118(3) (2010) 566-574.

[15] D.E. Evans, M.C. Sheehan, D.C. Stewart, The impact of malt derived proteins on beer foam quality. Part II- The influence of malt foam-positive proteins and non-starch polysaccharides on beer foam quality, J. Inst. Brew. 105(2) (1999) 171-177.

[16] A. Karlund, I. Paukkonen, C. Gomez-Gallego, M. Kolehmainen, Intestinal exposure to food-derived protease inhibitors: Digestion physiology- and gut health-related effects, Healthcare (Basel) 9(8) (2021) 1002.

Table 1: The improvements of in-gel sample preparation (IGSP) in faecal protein identifications (peptide coverage percentage and numbers) from bovine, barley and bacteria databases compared to filter-aided sample preparation (FASP).

Gene name	Protein	FASP		IGSP	
		Coverage (%)	Peptides	Coverage (%)	Peptides
Barley: in FASP and IGSP					
IAT2	Alpha-amylase/trypsin inhibitor CMb	18	2	28	3
PAZ1	Serpin-Z4	11	4	25	7
PAZ7	Serpin-Z7	8	3	11	4
Barley: only in IGSP					
IAT3	Alpha-amylase/trypsin inhibitor CMd	-	-	40	4
IAD1	Alpha-amylase inhibitor BDAI-1	-	-	30	3
Bovine: in FASP and IGSP					
ALB	Albumin	12	8	57	29
MPTX	Mucosal pentraxin	17	4	49	8
ANXA4	Annexin A4	10	3	34	9
ENPP3	Ectonucleotide pyrophosphatase/phosphodiesterase family member 3	5	4	30	16
A2M	Alpha-2-macroglobulin	4	5	26	28
SERPINA1	Alpha-1-antiproteinase	5	2	25	8
ANPEP	Aminopeptidase N	2	2	10	8
C3	Complement C3	3	5	9	12
Bovine: only in IGSP					
LYZ3	Lysozyme C-3	-	-	61	6
S100A9	Protein S100-A9	-	-	41	4
PIGR	Polymeric immunoglobulin receptor	-	-	28	13

ALPI	Intestinal-type alkaline phosphatase	-	-	28	9
SERPINA3-1	Serpin A3-1	-	-	23	8
DPEP1	Dipeptidase 1	-	-	20	6
LTF	Lactotransferrin	-	-	18	9
DPP4	Dipeptidyl peptidase 4	-	-	15	9
Bacteria: in FASP and IGSP					
pckA	Phosphoenolpyruvate carboxykinase (ATP) (*Lachnospira eligens*)	8	3	31	10
tuf	Elongation factor Tu (*Agathobacter rectalis*)	6	2	26	7
pckA	Phosphoenolpyruvate carboxykinase (ATP) (*Agathobacter rectalis*)	7	3	24	9
ilvC	Ketol-acid reductoisomerase (NADP(+)) (*Clostridium botulinum*)	13	4	22	6
tuf1	Elongation factor Tu (*Clostridium perfringens*)	6	2	19	5
fusA	Elongation factor G (*Lachnoclostridium phytofermentans*)	4	2	14	9
argC	NAD(P)-specific glutamate dehydrogenase (*Prevotella ruminicola*)	5	2	12	4
pnp	Polyribonucleotide nucleotidyltransferase (*Bacteroides thetaiotaomicron*)	3	2	7	4
Bacteria: only in FASP					

rpsM	30S ribosomal protein S13 (Bacteroides thetaiotaomicron)	14	2	-	-
gpmI	2,3-bisphosphoglycerate-independent phosphoglycerate mutase (*Bacteroides vulgatus*)	4	3	-	-
pfp	Pyrophosphate--fructose 6-phosphate 1-phosphotransferase (*Spirochaeta thermophila*)	4	2	-	-
tcdA	Toxin A (*Clostridium novyi*)	1	2	-	-
Bacteria: only in IGSP					
gdh	NAD-specific glutamate dehydrogenase (*Clostridium symbiosum*)	-	-	32	11
pckA	Phosphoenolpyruvate carboxykinase (ATP) (*Bacteroides fragilis*)	-	-	10	5
ppdK	Pyruvate, phosphate dikinase (*Clostridium symbiosum*)	-	-	9	6
fucI	L-fucose isomerase (*Bacteroides thetaiotaomicron*)	-	-	9	5

Legend to Figures

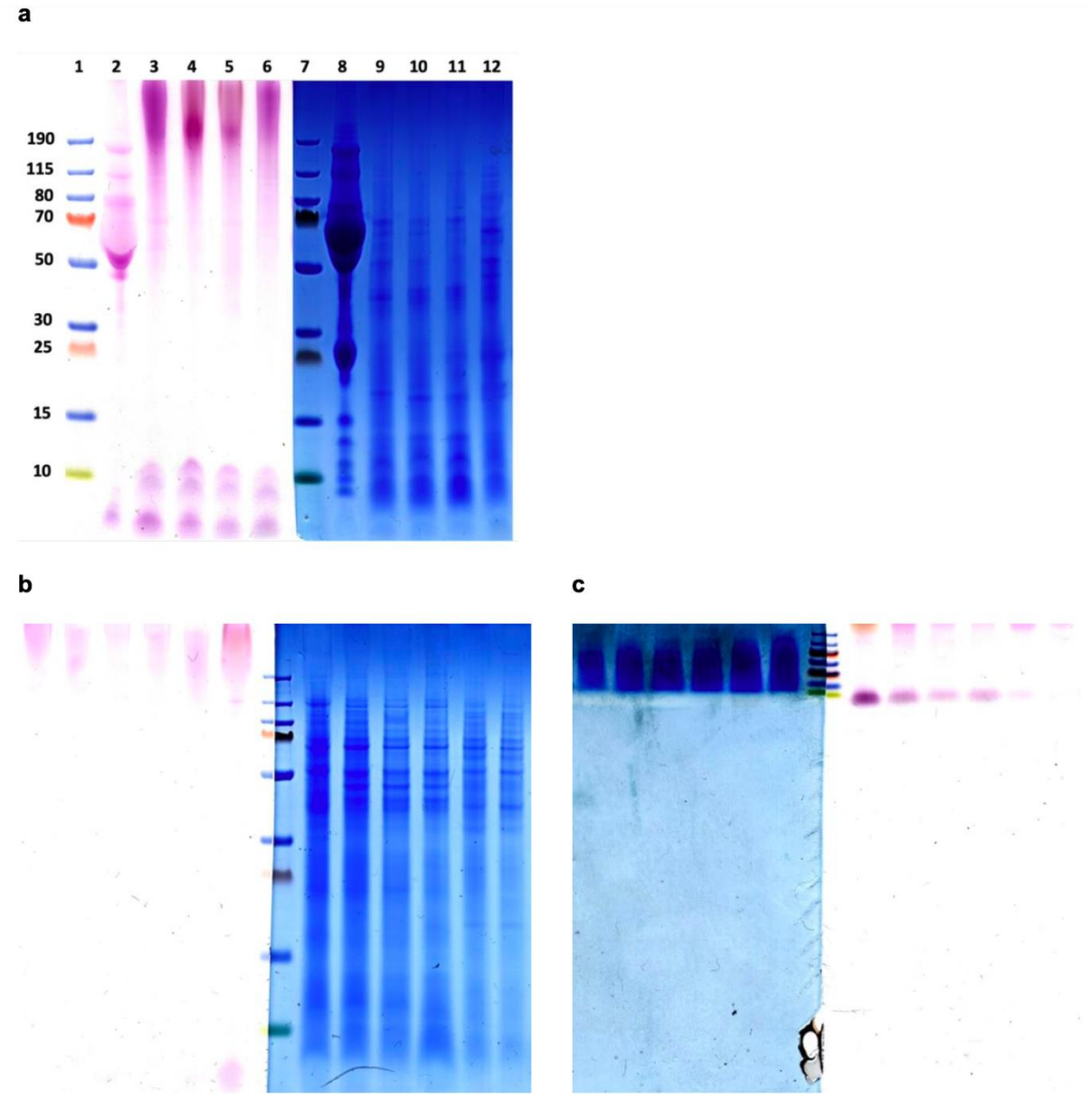
Figure 1: (a) Samples in the gradient gel were stained by periodic acid-Schiff (left) and by Coomassie blue (right). A bovine serum sample was run as references (lanes 2 and 8). Lanes 3 to 6 were loaded with the same bovine faecal samples as lanes 9 to 12 respectively. Faecal samples were run on 10% Bis-Tris gels at 200 V for (b) 35 min and (c) 4 min, and were stained by Coomassie blue and Periodic acid-Schiff respectively.

Figure 2: Comparisons between filter-aided sample preparation (FASP) and in-gel sample preparation (IGSP) in the bovine faecal sample identifications: (a) bovine proteins and (b) bacterial proteins (*Ruminococcaceae, Lachnospiraceae, Clostridiaceae, Prevotellaceae, Bacterioidaceae, Spirochaetaceae*). Proteins assessed here were the master proteins that had at least two unique peptides identified.

Supplementary Figure 1: Comparisons between filter-aided sample preparation (FASP) and in-gel sample preparation (IGSP) in the bovine faecal sample identifications: (a) bovine proteins; (b) barley proteins and (c) bacterial proteins (*Ruminococcaceae, Lachnospiraceae, Clostridiaceae, Prevotellaceae, Bacterioidaceae, Spirochaetaceae*). Proteins compared here were all master proteins.

291 **Figure 1**

292



293

294

295 **Figure 1:** (a) Samples in the gradient gel were stained by periodic acid-Schiff (left) and by Coomassie blue (right). A bovine
296 serum sample was run as references (lanes 2 and 8). Lanes 3 to 6 were loaded with the same bovine faecal samples as lanes
297 9 to 12 respectively. Faecal samples were run on 10% Bis-Tris gels at 200 V for (b) 35 min and (c) 4 min, and were stained by
298 Coomassie blue and Periodic acid-Schiff respectively.

299

Figure 2

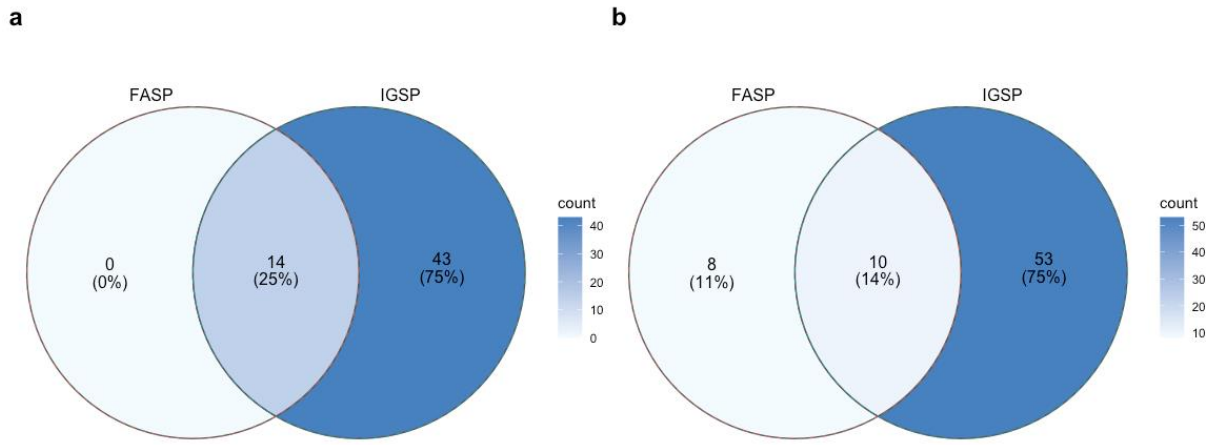
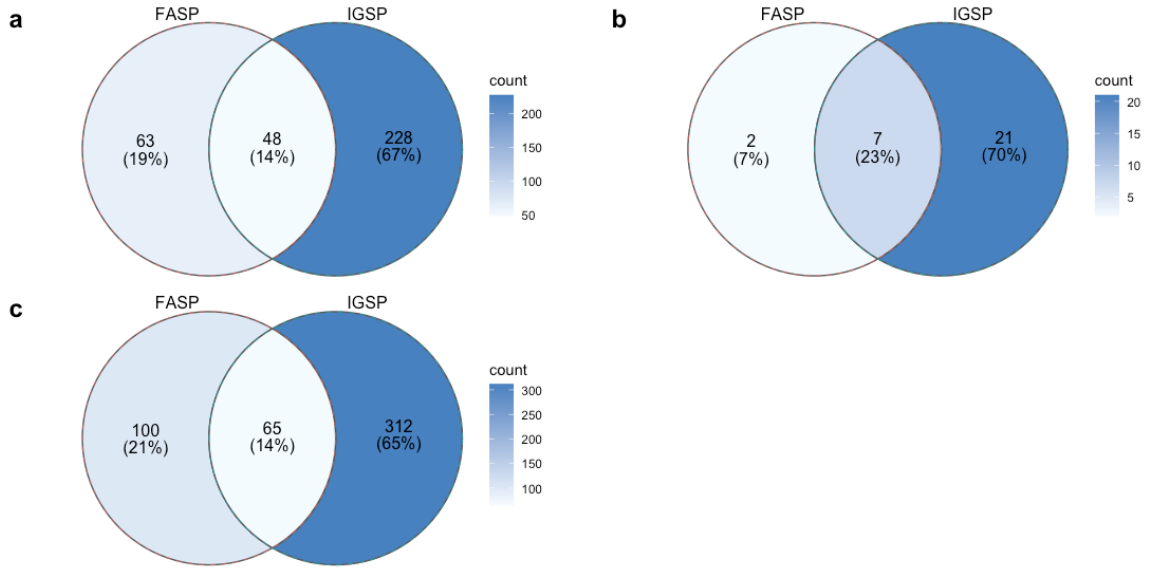


Figure 2: Comparisons between filter-aided sample preparation (FASP) and in-gel sample preparation (IGSP) in the bovine faecal sample identifications: (a) bovine proteins and (b) bacterial proteins (*Ruminococcaceae, Lachnospiraceae, Clostridiaceae, Prevotellaceae, Bacterioidaceae, Spirochaetaceae*). Proteins assessed here were the master proteins that had at least two unique peptides identified.

Supplementary information



Supplementary Figure 1: Comparisons between filter-aided sample preparation (FASP) and in-gel sample preparation (IGSP) in the bovine faecal sample identifications: (a) bovine proteins; (b) barley proteins and (c) bacterial proteins (*Ruminococcaceae, Lachnospiraceae, Clostridiaceae, Prevotellaceae, Bacterioidaceae, Spirochaetaceae*). Proteins compared here were all identified master proteins.