**Application**

Because of the public interest in the type and source of feedstuffs used in livestock production, we assessed the potential for faecal protein signatures to indicate whether specific plant sources were included in the diet of cattle. Many proteins derived from dietary plant materials were found in the faeces of cattle, but limitations on the depth of currently available plant specific information prevents a reliable identification of plant species.

**Introduction**

Our group has previously identified plant proteins such as serpin Z4 in bovine faeces (Huang et al., 2023). Concentrate feeds for cattle include cereals such as barley as a source of starch and energy, with supplementary protein, often as soyabean meal or as by-products of fermentation or oil extraction. However, there is particular public concern about the potential environmental effects of soyabean production (Wilkinson & Lee, 2018). We hypothesised that it might be possible to interrogate faecal proteomic signatures to forensically identify cattle that had been fed on specific plant species. The aim of the study was therefore to apply bioinformatic analysis to a pre-existing library of faecal proteins from barley-fed beef cattle to assess the reliability of assignment of plant species to proteins identified in faeces and to identify potential sources of bias.

**Materials and Methods**

The study was conducted using data generated previously by quantitative tandem mass tagged (TMT) mass spectroscopy of cattle faeces (Huang et al., 2023). In brief, 5 faecal samples were collected from 93 continental crossbred steers fed with an ammonia-treated barley-based diet (ATB) and 5 of 124 steers fed a propionate-treated similar diet (PTB). Both diets were barley cereal, soyabean meal, barley or wheat straw, and mineral premix. The presence of contaminating species was possible, but unlikely. The samples were analysed with TMT 10-plex by Huang et al. (2023). Here, MaxQuant (MQ) and Proteome Discoverer (PD) were both used to search against protein sequence databases including Swiss-Prot and TrEMBL downloaded from UniProt for each of 14 common feed plant species individually followed by quantification using Perseus after MQ. Searches were then conducted using a single, concatenated FASTA file with all 14 species. The abundance of each protein in faeces from cattle fed ATB and PTB was compared using Mann-Whitney U-test with Benjamini-Hochberg correction for FDR. Two potential sources of bias were assessed – the similarity of homologs and orthologs of proteins between species, and the richness of protein databases by species. The mean pairwise-distances of the sequences among 7 species of elongation factor Tu (EF-Tu, reported as present in all 7 species when searched individually), of ABC transporter domain-containing protein including ABC transporter G family member 31-like protein and pleiotropic drug resistance protein 1 and 2 (5 species), of germin-like protein (4 species), of hexosyltransferase (3 species), of pyruvate, phosphate dikinase (3 species), of JmjC domain-containing protein (1 species), and of trichome birefringence-like N-terminal domain-containing protein (1 species) was estimated with the Poisson model in MEGA 11. Spearman’s rank correlation coefficient was determined for 1) the mean pairwise distances for each protein with the number of species in which each protein was found, 2) the total number of protein sequences in each species database with the number of identified proteins returned.

**Results**

More proteins were identified using PD than MQ. Individual database searches identified 8 and 35 barley proteins with MQ and PD respectively, compared to 1 and 11 in the single concatenated searches. No differences in protein abundance by barley treatment (ATB v PTB) were detected (P > 0.05). Each analysis identified serpin family proteins including serpin domain-containing protein, serpin Z4 and serpin Z7. The serpins were not included in the pairwise distance analysis because of the large number of potential orthologs and homologs. EF-Tu was identified from 5 and 7 multiple species in individual searches using both MQ and PD respectively. EF-Tu sequences were similar among 7 plant species, with an overall mean pairwise-distance of 0.17. In contrast, for trichome birefringence-like N-terminal domain-containing protein (found only in barley), the overall mean pairwise-distance was 0.6. The pairwise differences between barley and wheat for these two proteins were 0.004 and 0.09. The correlation between pairwise sequence distance and the number of species in which proteins were identified was -0.69 (not significant, P = 0.09). The number of sequences per species database correlated significantly with the number of identified proteins returned from individual species database searches and from the single concatenated database using PD (*r* = 0.90; P < 0.001, *r* = 0.88; P < 0.001 respectively).

**Conclusions**

Faecal proteomics identified highly abundant plant proteins that were likely to be in the diet, but it also identified proteins from species that were very unlikely to be ingested by cattle in this study. Each analysis found barley serpin Z4, which is the most abundant barley protein in faeces of cattle fed on the barley-based diet. Although the concatenated database search possibly improves peptide match while excluding highly conserved sequences, the number of sequences in the databases seems likely to bias identifications towards species with more sequences; proteins that are highly similar among species or orthologs are also likely to have incorrect species assignments. Current untargeted proteomic methods applied to faeces are unlikely to be useful for forensic assessment of plant species fed to cattle.

**References**

Huang, Y., Jonsson, N. N., McLaughlin, M., Burchmore, R., Johnson, P. C. D., Jones, R. O., Eckersall, P. D., 2023. Quantitative TMT-based proteomics revealing host, dietary and microbial proteins in bovine faeces including barley serpin Z4, a prominent component in the head of beer. J Proteomics, 285, 104941.

Wilkinson, J. M., and Lee, M. R. F., 2018. Review: Use of human-edible animal feeds by ruminant livestock. Animal, 12(8), 1735-1743.