Simulated digestion, uptake and colonic fermentation of Gorse (*Ulex europeaus*) suggests safe use as a source of protein concentrates and bioactives.

Ajay Iyer[1,2], Gary J. Duncan[1], Nicholas J. Hayward[1], Charles S. Bestwick[1], Sylvia H. Duncan[1], and Wendy R. Russell[1]

[1]University of Aberdeen, Rowett Institute, Aberdeen AB25 2ZD, Scotland
[2]Conway Institute of Biomolecular and Biomedical Research, University College Dublin (UCD) Belfield, Dublin 4, Ireland.

Abstract

Background: Invasive plants such as Gorse (*Ulex europeaus*) may serve as a potentially sustainable source of edible proteins and bioactives through the production of leaf protein concentrates (GLPC). The digestion, uptake and safety of phytochemicals from GLPC and the protein depleted but bioactive rich fraction (PDF) from Gorse was studied using an in vitro simulated digestion model.

Results: Both the Gorse fractions and their digesta maintained the viability of cultured (Caco-2) intestinal cells above the 80% threshold, suggesting a lack of overt toxicity.
In the digestion model, most of the plant metabolites in the digesta were largely confined to the apical fraction of the Caco-2 monolayer (representing the lumen) and were subsequently expected to be delivered to the colon. Food matrix had a significant but marginal effect on the permeability of metabolites across the Caco-2 monolayer. Lastly, *in vitro* intestinal microbial action appeared to produce beneficial compounds such as enterolactones, and equol.

Conclusion: The Gorse fractions lack general overt toxicity to Caco-2 cells. GLPC and PDF provide metabolic substrates for the colonic microbes to produce secondary compounds that are associated with positive health outcomes.

Keywords: Simulated digestion, gorse, leaf protein concentrates, faecal fermentation, apparent permeability.

1 Introduction

The spread of invasive plants can significantly affect local ecologies and subsequently the economy by depreciating the land capability (Vilà & Ibáñez, 2011), competing with crops of agricultural value (Ziska et al., 2011), and disrupting local niches and ecological webs (Cocîrlea & Oancea, 2022). Among the numerous strategies investigated to counter the spread of invasive species, invasivory has been proposed as an unconventional yet sustainable alternative to incineration, animal grazing, or herbicide use (Iyer et al., 2021). Invasivory is the consumption of invasive species as a means to deplete its population in non-native habitats. The consumption of invasive species does not directly compete with the existing agricultural infrastructure and can be easily set up to complement existing food production capacities. In the context of policy-making and urban foraging communities, invasivory has gained interest (Cerveira et al., 2022; Mishan & Hamada, 2020), although formal scientific investigation on this matter remains limited. Gorse has spread far beyond its native habitat that has resulted in substantial ecological and economic damage to the introduced environment (Broadfield & McHenry, 2019). For example, the regions of Canterbury and Wellington in New Zealand are well known cases of Gorse overwhelming local flora and disrupting the ecology (Came et al., 2020).

Previous research by Iyer et al. (2021) has identified Gorse as a lucrative candidate for protein production. During early growth phases, the leaves of Gorse are rich in protein as well as phytochemicals such as chlorogenic acid, kaempferol and sinapic acid. Leaf protein concentrates (LPC) from Gorse offers an unconventional, yet potentially sustainable and scalable way of obtaining nutrition from invasive plant biomass. Gorse leaf protein concentrate (GLPC) can be prepared at a yield and purity of 5.5% ($w \cdot w^{-1}$) and 96.8% ($w \cdot w^{-1}$) respectively, with a highly favourable amino acid profile (Iyer et al., 2022). While LPC has previously been studied in a human intervention setup (Dewan et al., 2007), it was important to establish the safety of GLPC for application in consumer products.

Establishing food safety in the preliminary laboratory stages of product development involves establishing the absence of potent anti-nutrients, and a precedence of consumption, according to the SAFEST guidelines (Jonas et al., 1996). As of 23 August 2023, no reports of toxicity *via* anti-nutrients in Gorse could be found on literature databases such as The National Library of Medicine (MEDLINE), Cochrane

Central, Scopus and Google Scholar. Conversely though, there is limited evidence of its direct human consumption, with the bulk of the documentation largely focused on animal feed. Gorse continues to remain a subsistence form of animal feed when high quality feeds are unavailable (Rymer, 1979).

The aim of the study was to test if GLPC and plant metabolites were safe for consumption. In this work, GLPC and the protein depleted fractions (PDF), obtained from Gorse leaves, were subjected to an in vitro simulated digestion model to understand their digestion as they progress through the digestive tract. The uptake of metabolites at the intestinal phase in the model was also simulated using polarised Caco-2 cells. Additionally, the cell viability assays using the Caco-2 cell model (at their 36 hour proliferative stage) was used to test for the absence of toxicity.

2 Materials and methods

All chemicals were purchased from Merck (Darmstadt, Germany) unless stated otherwise. Milli-Q®purified water was used for all experiments. To keep the media and water anaerobic, they were boiled for 5 min, allowed to cool for one hour, then flushed with nitrogen for 2 min. Finally, the water was placed in an anaerobic cabinet (Don Whitley Scientific, West Yorkshire, UK) with a gas phase of 10% H_2, 10% CO_2, and 80% N_2, for 48 h before use.

2.1 Gorse sample preparation

Gorse (*Ulex europaeus*) was collected in Aberdeenshire, Scotland (GPS co-ordinate: 57.257, −2.483). Sampling procedures and processing of the Gorse leaves has been previously described by Iyer et al. (2022).

The freeze-dried and milled Gorse leaves (20 g) were first macerated in phosphate buffer solution (PBS; 20 mM; pH 7.0; 200 mL) at 30 °C for 30 min. The sample was centrifuged (3000 × g; 15 min) and the supernatant was recovered and set aside. The retentate was subjected to a second extraction using cellulase (1 mg·g^{-1} dry retentate) in sodium-citrate buffer (50 mM; pH 4.0; 200 mL) at 40 °C for 2 h. The sample was centrifuged (3000 × g; 15 min) and the supernatant was pooled with the previous step. Ethanol (130 mL) was added to the pooled supernatants and allowed to rest at −20 °C overnight. The sample was centrifuged (4000 × g; 30 min). The supernatant (after *in vacuo* ethanol removal) and precipitate were freeze-dried. The mass obtained from dried supernatant was the protein depleted fraction (PDF) and the precipitate was the leaf protein concentrate (GLPC).

2.2 The simulated digestion model

The *in vitro* digestion model used to study the GLPC and PDF samples was based on a previous publication by Minekus et al. (2014) with some modifications using other protocols by (Mulet-Cabero et al., 2020; Brodkorb et al., 2019).

Stock solutions (2.5×, 200 mL) were prepared as follows:

Stock salivary fluid comprised of KCl (18.87 mM), KH_2PO_4 (4.63 mM), $NaHCO_3$ (17.0 mM), $MgCl_2 \cdot 6\,H_2O$ (0.18 mM) and $(NH_4)_2CO_3$ (0.08 mM).

Stock gastric fluid comprised of KCl (8.62 mM), KH_2PO_4 (1.13 mM), $NaHCO_3$ (31.25 mM), NaCl (9.6 mM), $MgCl_2 \cdot 6\,H_2O$ (0.41 mM) and $(NH_4)_2CO_3$ (0.63 mM).

Stock intestinal fluid comprised of KCl (8.5 mM), KH_2PO_4 (1.0 mM), $NaHCO_3$ (106.25 mM), NaCl (48.0 mM), $MgCl_2 \cdot 6\,H_2O$ (0.41 mM) and $(NH_4)_2CO_3$ (0.41 mM).

Calcium chloride ($CaCl_2 \cdot 2\,H_2O$) solution (10 mL, 300.0 mM) was also prepared which was added to the final simulated fluids.

Enzyme solutions α-amylase (0.15 g·mL^{-1}; Merck, Darmstast, Germany), pepsin (0.1 g·mL^{-1}; ≥30 units·mg^{-1}; Merck, Darmstadt, Germany) and pancreatin (0.08 g·mL^{-1}; Merck, Darmstadt, Germany) were prepared using the stock salivary, gastric, and intestinal buffers respectively. Thus the final experimental Simulated Salivary Fluid (SSF) comprised of stock salivary fluid (70 mL), α-amylase solution (10 mL), calcium chloride solution (CaCl, 0.48 mL) with pH adjusted to 7 using sodium hydroxide solution (NaOH, 1 M). Simulated Gastric Fluid (SGF) comprised of stock gastric fluid (64 mL), pepsin solution (16 mL), calcium chloride solution (CaCl, 0.48 mL) and adjusted to pH 4 using hydrochloric acid solution (HCl, 1 M). Simulated Intestinal Fluid (SIF) comprised of stock intestinal fluid (42.5 mL), pancreatin solution (25 mL), calcium chloride solution (CaCl, 0.2 mL) and adjusted to pH 7 using sodium hydroxide solution (NaOH, 1 M).

All simulated fluids were maintained at 37 °C for at least 20 min prior to sample addition. SGF and SIF were prepared and stored in an anaerobic cabinet. To initiate the simulated salivary phase, GLPC/PDF (2 g) was incubated with SSF (20 mL; 37 °C) for 30 s, while stirring on a magnetic stirrer. The relative volume and incubation period was adjusted according to physiological reports (Hughes & Wiles, 1996). The sample and SSF were then transferred to an anaerobic chamber (37 °C) where they were added to SGF (40 mL) and incubated under stirring for 3 h. Samples were finally transferred to the SIF (80 mL)

and incubated for a further 3 h while stirring. These incubation times were based on physiological transit periods in an empty alimentary tract (Schiller et al., 2005). Samples were then centrifuged at 3000 $\times g$ for 15 min and the supernatant was sterilised using 0.2 µm filter. At each stage, 1 mL of sample was taken for the analysis of phenolics *via* LC/MS.

2.3 Cell culture

2.3.1 Media

The complete working medium (CWM) was comprised of DMEM (863.6 mL·L^{-1}; Dulbecco's Modified Eagle Medium; BioWhittaker®; Lonza; Walkersville; USA), foetal Bovine Serum (104.9 mL; HyClone™; GE Healthcare Life Sciences; Utah; USA), penicillin-streptomycin 100× solution (10.5 mL·L^{-1}; Sigma-Aldrich, Pennsylvania; USA) and non-essential amino acids (10.5 mL·L^{-1}; Sigma-Aldrich, Pennsylvania; USA).

Phenol-red free complete working medium (PRF-CWM) was comprised of DMEM (874.1 mL·L^{-1}; DMEM; BioWhittaker®; Lonza; Walkersville; USA), foetal bovine serum (104.9 mL·L^{-1}; HyClone™; GE Healthcare Life Sciences; Utah; USA), penicillin-streptomycin 100× solution (10.5 mL·L^{-1}; Sigma-Aldrich, Pennsylvania; USA), non-essential amino acids (NEAA, 10.5 mL·L^{-1}, Sigma-Aldrich, Pennsylvania, USA) and glutamine (10.5 mL·L^{-1}; Sigma-Aldrich, Pennsylvania; USA).

2.3.2 Cell type and handling

Caco-2 cells (at passage number 21) were purchased from ATCC and maintained at −70 °C under glycerol in liquid nitrogen at a concentration of 10^6 cells·mL^{-1}. Basic handling and set up were performed based on established good laboratory practices Natoli et al., 2012. The working culture was adjusted to 10^4 cells·mL^{-1} and incubated in 10 mL CWM. Media was replaced every 48 h.

2.4 Sample and standard preparation

2.4.1 Representative/standard compounds

In a previous publication by Iyer et al. (2021), ten compounds were found to be significantly higher in Gorse compared to other invasive plants found in Scotland, namely, p-coumaric acid, acetovanillone, protocatechuic acid, chlorogenic acid, vanillin, p-hydroxybenzoic acid, sinapic acid, kaempferol, mandelic acid and ferulic acid. Among these however, only six compounds, namely, p-coumaric acid, acetovanillone, vanillin, p-hydroxybenzoic acid, mandelic acid and ferulic acid were detected in the final GLPC and were used as representative sample compounds to study their metabolism as they passed along the digestion model. A standard compound mixture (50 mM, 200 µL, p-coumaric acid, acetovanillone, protocatechuic acid, chlorogenic acid, vanillin, p-hydroxybenzoic acid, sinapic acid, kaempferol, mandelic acid and ferulic acid) was prepared and incubated with the Caco-2 cells for 24 h at 37 °C to assess active cell transport.

2.4.2 Preparation of GLPC and PDF digesta for toxicity testing

Samples obtained at the intestinal phase in the digestion model were collected, freeze-dried and reconstituted in PRF-CWM. The samples were filter-sterilised through a 0.2 µm filter (Millipore, USA). Toxicity was tested on these samples as described above.

2.5 Toxicity

2.5.1 MTT assay

A sterile 96-well plate (Greiner Bio-One, Kremsmünster, Austria) was seeded with Caco-2 cells (6×10^4 cells; 200 µL), which were allowed to adhere to the plate over 36 h. The spent media was aspirated, and cells were washed *in situ* using sterile PBS. Then the digestate samples (200 µL) were loaded and incubated for 24 h. The supernatant medium was drained, and the cells were incubated in PRF-CWM (100 µL) with MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) at a concentration of 0.5 mg·mL^{-1} for 4 h at 37 °C. The supernatant was aspirated and acidified (40 mM,

HCl), and 2-propanol (250 µL) was added to the wells. The 96-well plate was placed on a shaker for 30 min, then centrifuged at 900 × g at 4 °C for 10 min min. The supernatant (150 µL) was recovered and absorbance was measured at 570 nm.

2.6 Metabolite transport across the Caco-2 monolayer

The transport of metabolites in the digestates across the intestinal barrier was simulated using a Caco-2 monolayer on a transwell plate (Hubatsch et al., 2007). Metabolite movement was measured in both directions across the Caco-2 monolayer to establish if transportation was passive or active, and independent or synergistic/ inhibitory. The basolateral fraction represented the blood where metabolites were expected to be measured after being transported across the Caco-2 monolayer. The apical fraction represented the lumen, where metabolites which could not be transported across the Caco-2 monolayer were expected to remain. Under physiological conditions, the apical (luminal) fraction was expected to be delivered to the colon where it interacts with commensal microbes.

2.6.1 Apical to basolateral

Two 24-well plates with filter-inserts (0.4 µm pore size; Greiner Bio-One, Kremsmünster, Austria) were seeded with Caco-2 (10^4 cells·cm^{-2} at passage 27) and incubated for 21 d to facilitate growth and differentiation. The Transepithelial Electrical Resistance (TEER) was measured to be 4879±360 mΩ. Once TEER measurements were found to be constant and consistent across wells, PDF, GLPC, or standard compound (200 µL) was loaded on the upper well. The lower well contained PRF-CWM (1 mL) which was drawn hourly for 4 h and stored immediately at −20 °C. The bottom well was replenished with fresh PRF-CWM with every draw of sample.

2.6.2 Basolateral to apical

In a technique described previously by Stone et al. (2019), wells were initially layered with a poly-L-lysine solution (Sigma-Aldrich, Pennsylvania, USA) for 2 h and allowed to air-dry overnight in sterile conditions.

The inserts were inverted, arranged on the lid, and seeded with Caco-2 (10^4 cells·cm^{-2}, passage 38). The baseplate was gently laid over the inserts as the bottom chamber touched the seed droplet, stabilising them as a hanging droplet. The set-up was allowed to rest at 37 °C for 6 h. A fresh 24-well plate was filled with PRF-CWM (1 mL) and the inserts were gently laid into them and incubated for two weeks to facilitate growth and differentiation. TEER was measured to be 5055±666 mΩ. The subsequent steps were identical to the apical to basolateral part of the experiment.

2.7 Scaled up metabolite transport

The apical to basolateral transport was scaled up using a 100 mm diameter culture dish (Corning, Transwell®, 75 mm insert, polycarbonate, 44 cm^2 surface area) under conditions described above. The TEER was 454±92 mΩ.

2.8 Faecal fermentation

2.8.1 Medium

The M2GSC medium was prepared as described previously by Cummings and Macfarlane (1991). The recipe is as follows (1× concentration):

Bacto™Casitone (10 g·L^{-1}), Bacto™Yeast Extract (2.5 g·L^{-1}), sodium bicarbonate ($NaHCO_3$, 4.0 g·L^{-1}), glucose (2 g·L^{-1}), soluble starch (2 g·L^{-1}), cellobiose (2 g·L^{-1}), clarified rumen fluid (100 mL·L^{-1}), cysteine HCl (1 g·L^{-1}), dipotassium hydrogen phosphate (K_2HPO_4, 0.4 g·L^{-1}), potassium dihydrogen phosphate (K_2HPO_4, 0.4 g·L^{-1}), ammonium sulphate ($(NH_4)_2SO_4$, 0.8 g·L^{-1}), sodium chloride (NaCl, 0.8 g·L^{-1}), magnesium sulphate ($MgSO_4$, 80 mg·L^{-1}), calcium chloride ($CaCl_2$, 80 mg·L^{-1}) and resazurin (1 µg·L^{-1}).

The solution was prepared using water kept under anaerobic conditions for 48 h.

2.8.2 Fermentation sample preparation

The luminal fraction of GLPC and PDF digesta from the preceding Transwell®setup were recovered and freeze-dried. The samples were reconstituted in M2GSC medium and filter sterilised under anaerobic conditions to produce the experimental medium for faecal fermentation.

2.8.3 Faecal inoculation

Faecal samples were obtained from three healthy adult donors, stored at 4 °C, and processed within 5 h of collection. Under sterile conditions, the faecal sample from each donor (2 g) was weighed into DispoMix™tubes (gentleMACS) containing sterile glycerol-PBS (PBS in 30% glycerol, deoxygenated and autoclaved, 10 mL). The samples were blended under high shear using the DispoMix™blender (Medic Tools; Switzerland) for 30 s to obtain faecal slurries.

Experimental medium (5 mL), and faecal slurry (20 µL) were suspended in air-tight, rubber-sealed tubes and incubated at 37 °C for 48 h. Samples were drawn using anaerobic techniques at 0 h, 24 h and 48 h and stored at −70 °C for no longer than 20 d until further processing. For further analysis, samples were allowed to thaw overnight at 4 °C in the dark and centrifuged at 12,500 × g for 5 min at 4 °C. The supernatant was recovered and characterised for metabolites.

2.9 Metabolite characterisation

Phenolic and protein metabolites present in the analytes were characterised using LC-MS as described previously by Russell et al. (2011). Naming schemes used for in house measurement, corresponding trivial names, and the compounds groups are available in **Supplementary Table 1**.

2.10 Short chain fatty acid (SCFA) measurements

Short chain fatty acids were measured using gas chromatography as described by Richardson et al. (1989). Internal standard used was EBA (50 µL; 2-ethyl butyric acid; 1.27% v·v^{-1}). External working

standard; prepared in deoxygenated distilled water, comprised of acetic acid (1.72 mL·L^{-1}, Fisher, UK), propionic acid (1.49 mL·L^{-1}), iso-butyric acid (0.5 mL·L^{-1}), n-butyric acid (0.18 mL·L^{-1}), sodium formate (0.7 g·L^{-1}, BDH Chemicals Ltd., UK), lithium lactate (1.0 g·L^{-1}, BDH Chemicals Ltd., UK), and sodium succinate dibasic hexahydrate (0.3 g·L^{-1}).

External standard or sample (1 mL) was mixed with HCl (hydrochloric acid, 0.5 mL, 9 M) and diethyl ether (2 mL) using a vortex mixer for one minute. The mixture was centrifuged at 3000 $\times$ g for 10 min under ambient condition. The organic layer was recovered and to the aqueous fraction, fresh diethyl ether (2 mL) was added and subjected to extraction once more. The organic layers were pooled. From the pooled organic layer, 800 µL sample and MTBSTFA (100 µL, N-tert-butyldimethylsilyl)-N-methyltrifluoroacetamide) were transferred to amber capped glass vials (Agilent Technologies, USA), heated to 80 °C for 20 min. The vials were allowed to rest under ambient conditions in the dark for 72 h.

Derivatised SCFA samples were measured on an Agilent HP 6890 GC system with autosampler and cooled sample tray (between 8 to 12 °C), using an Agilent 190912-333 HP-1: Methyl-Siloxane column. Samples were heated to 63 °C for 3 min and held at this temperature for a further 3 min. Along a temperature ramp of 10 °C·min^{-1}, column temperature was increased to 190 °C within 17.2 min, and held for 1.5 min. The evaporated sample was further heated to 275 °C and maintained at 135.0 mN·mm^{-2} within the GC. The measurement was performed using Flame Ionization Detector unit (FID).

Owing to constraints in available resources during the Covid-19 pandemic, the triplicate samples generated from the experiment were pooled into a single sample and measured for SCFA.

2.11 Statistical analysis

Statistical analysis was performed using R (Version 3.6.3) (R Core Team, 2021). Data processing and visualisation were performed using the `tidyverse` set of packages (Wickham et al., 2019). Principal component analysis (PCA, univariate-scaled) was performed using packages `factoextra` (Kassambara & Mundt, 2020). Statistical comparison across means was performed using one-way ANOVA and expressed as $F_{(\text{degrees of freedom, residuals})}$=F value, p-value). The raw data, codes and annotations may be found in

the OSF repository (https://osf.io/9qtyz/).

Three univariate-scaled principal component analyses (PCA) were performed on the data set. The first PCA model considered the non-bacterial phase of the model that included the metabolite profile of the initial GLPC/PDF fractions to the apical and basolateral fractions after Caco-2 exposure. The second PCA model considered the apical fraction of the Caco-2 model to the 48 h simulated colonic fermentation. The third PCA model considered the complete set of phases from the GLPC/PDF to the fermentation phase.

3 Results

3.1 Metabolites across the *in vitro* digestion model

The LC/MS platform detected and quantified 181 metabolites in the *in vitro* digestion model. Data for the metabolites measured in each phase of the digestion model, including control measurements, is provided in **Supplementary Tables 2** and **3**. **Supplementary Table 2** presents the metabolites quantified for samples from the initial GLPC/PDF to the Caco-2 monolayer. **Supplementary Table 3** presents metabolites quantified in the faecal fermentation. Corresponding metabolite group-wise mass balance is provided in **Supplementary Table 4** from the initial GLPC/PDF to the Caco-2 monolayer phase. **Supplementary Table 5** shows the group-wise balance of the faecal fermentation phase of the digestion model.

Interesting dynamics were noted for the major metabolite groups. For mandelic acids in the PDF they could only be detected until the gastric phase (**Supplementary Table 4**), and were subsequently absent even in the supernatants in the faecal fermentation phase (**Supplementary Table 5**). Mandelic acids were absent in the GLPC and its digestates. Similarly in the case of phenolic alcohols, they were absent in GLPC and could only be detected in the PDF until the salivary phase (**Supplementary Table 4**). However, they reappeared at 48 h in the colonic fermentation supernatants (**Supplementary Table 5**). This was expected as many microbes are capable of converting flavonoids into corresponding phenolic alcohols.

For analyses, the *in vitro* digestion model was divided into the non-bacterial (initial GLPC/PDF material to Caco-2 monolayer phase) and bacterial phase (Caco-2 apical fraction to 48 h time point of faecal fermentation), as the metabolites were better represented in the PCA model based on their $\cos^2$ values.

3.1.1 Metabolites measured in GLPC/PDF to Caco-2 monolayer model

PCA was performed on the data from **Supplementary Table 2**. The codes, and annotations are provided in **Supplementary Code (code chunk 4 and 5)**. This model captures the variance in

metabolites from GLPC/PDF to the Caco-2 monolayer, and is depicted in Figure 1.

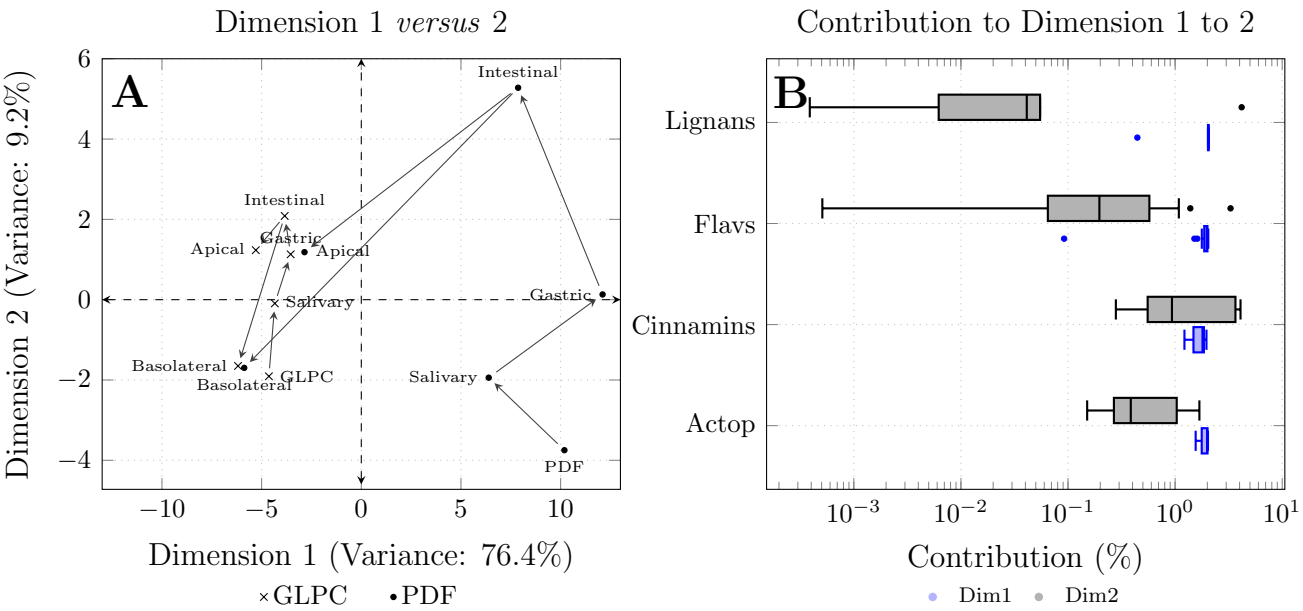


Figure 1: The panel A is the PCA plot of the metabolite profiles from the initial (GLPC/PDF) sample to the Caco-2 transport model. Panel B uses boxplots to show the distribution of the (%) contribution of the compounds belonging to major metabolite groups. Flavs = Flavanoids/coumarins, Cinnamins = Cinnamic acids, Actop = Acetophenones. Boxplots in blue mark represents Dimension 1, while the boxplot marked in black represents Dimension 2.

In Figure 1 panel A, each point represents the metabolite profile measured at each phase of the digestion model. The metabolite profile representing the points for the basolateral fractions of GLPC and PDF very nearly converge, with a distance of 0.3 along the Dimension 1. In contrast the metabolite profile representing the points for the apical fractions of GLPC and PDF are further apart at a distance of 2.4 along Dimension 1. This suggests that irrespective of the disparity in the composition of the metabolites in the initial fractions (separated by a distance of 14.0), the bioavailability of the compounds transported across the Caco-2 monolayer may be limited. This model indicates that much of the metabolites from GLPC/PDF are delivered to the colon for further microbial fermentation.

Metabolite groups such as lignans, flavanoids/coumarins, cinnamic acids, benzoic acids, and acetophenones showed high contribution to the PCA model, particularly along Dimension 1 (Figure 1, panel B). A mass balance of the compounds contributing the highest to the PCA model in Figure 1 is given in Table 1.

Table 1: Mass balance of metabolites across initial sample to Caco-2 transportation phase.

Group	Compound	Sample	Start	Salivary	Gastric	Intestinal	Apical	Basolateral
Lignans	secoisolariciresinol	GLPC	0.18± 0.08	0.15±0.00	0.22± 0.00	0.15± 0.00	0.03±0.00	N/D
		PDF	4.10± 1.68	2.88±0.30	2.88± 0.30	3.45± 0.08	0.92±0.04	0.04±0.01
	lariciresinol	GLPC	0.21± 0.37	0.36±0.02	0.53± 0.21	N/D	N/D	N/D
		PDF	7.12± 2.56	4.74±0.37	7.92± 0.48	5.67± 0.32	0.66±0.57	N/D
	syringaresinol	GLPC	17.13± 8.94	15.56±1.23	19.21± 1.95	16.58± 0.31	1.50±0.06	N/D
		PDF	402.67±179.43	321.13±5.93	430.54±35.71	333.67±37.23	50.53±2.34	18.47±0.54
Indoles	indole-3-carboxylic acid	GLPC	N/D	0.02±0.00	0.05± 0.00	N/D	N/D	N/D
		PDF	0.30± 0.13	0.21±0.02	0.33± 0.00	0.26± 0.00	0.08±0.00	N/D
Flavanoids/ Coumarins	quercetin-3-glucoside	GLPC	0.23± 0.07	0.28±0.04	0.43± 0.01	0.36± 0.02	0.09±0.00	0.05±0.00
		PDF	2.50± 1.22	1.79±0.15	2.90± 0.27	1.72± 0.13	0.62±0.03	N/D
	glycitein	GLPC	1.05± 0.53	1.50±0.01	2.25± 0.07	1.32± 0.06	0.07±0.01	0.03±0.00
		PDF	12.41± 7.98	11.30±0.30	14.16± 3.17	12.18± 0.35	0.94±0.08	0.69±0.02
Benzoic acids	syringin	GLPC	2.41± 1.22	2.04±0.13	2.95± 0.14	2.25± 0.06	0.07±0.06	N/D
		PDF	22.55± 6.97	16.77±0.41	27.20± 3.07	17.76± 1.78	0.96±0.22	1.56±0.07

Values expressed in mean±standard deviation (µg/2g GLPC/PDF).
N/D = not detected.

Data in Table 1 and **Supplementary Table 4**, suggest that the compound groups of indoles, lignans, flavonoids/coumarins, and benzoic acids that were found in high quantities in the PDF fraction, were transported across the Caco-2 monolayer into the basolateral fraction. Individual compounds, namely secoisolariciresinol, lariciresinol, syringaresinol, indole-3-carboxylic acid, quercetin-3-glucoside, glycitein and syringin, showed the most variance and drove the characterisation of the metabolite profile noted in Figure 1.

3.1.2 Metabolites measured from the Caco-2 apical fraction to faecal fermentation stage

A PCA was performed (Figure 2) using data from **Supplementary Table 3** to study the variance in the metabolites from the apical fraction of the Caco-2 monolayer to the faecal fermentation phase.

In Figure 2, the panels visualise the PCA model with combinations of Dimensions 1, 2 and 3. Plots showing Dimensions 1 and 2, and Dimensions 1 and 3 place the apical phases of the GLPC/PDF digestates diametrically opposite their respective 24 h and 48 h fermentate metabolite profiles. When considering Dimensions 1 and 2, and Dimensions 2 and 3, there was a clear distinction between metabolites arising from GLPC and PDF for all three volunteer samples. The contribution of individual compounds in each dimension is provided in **Supplementary Code (code chunk 12 and 13)**. The mass balance of highest contributing compounds to the PCA model across Dimensions 1 to 3 is provided in Table 2.

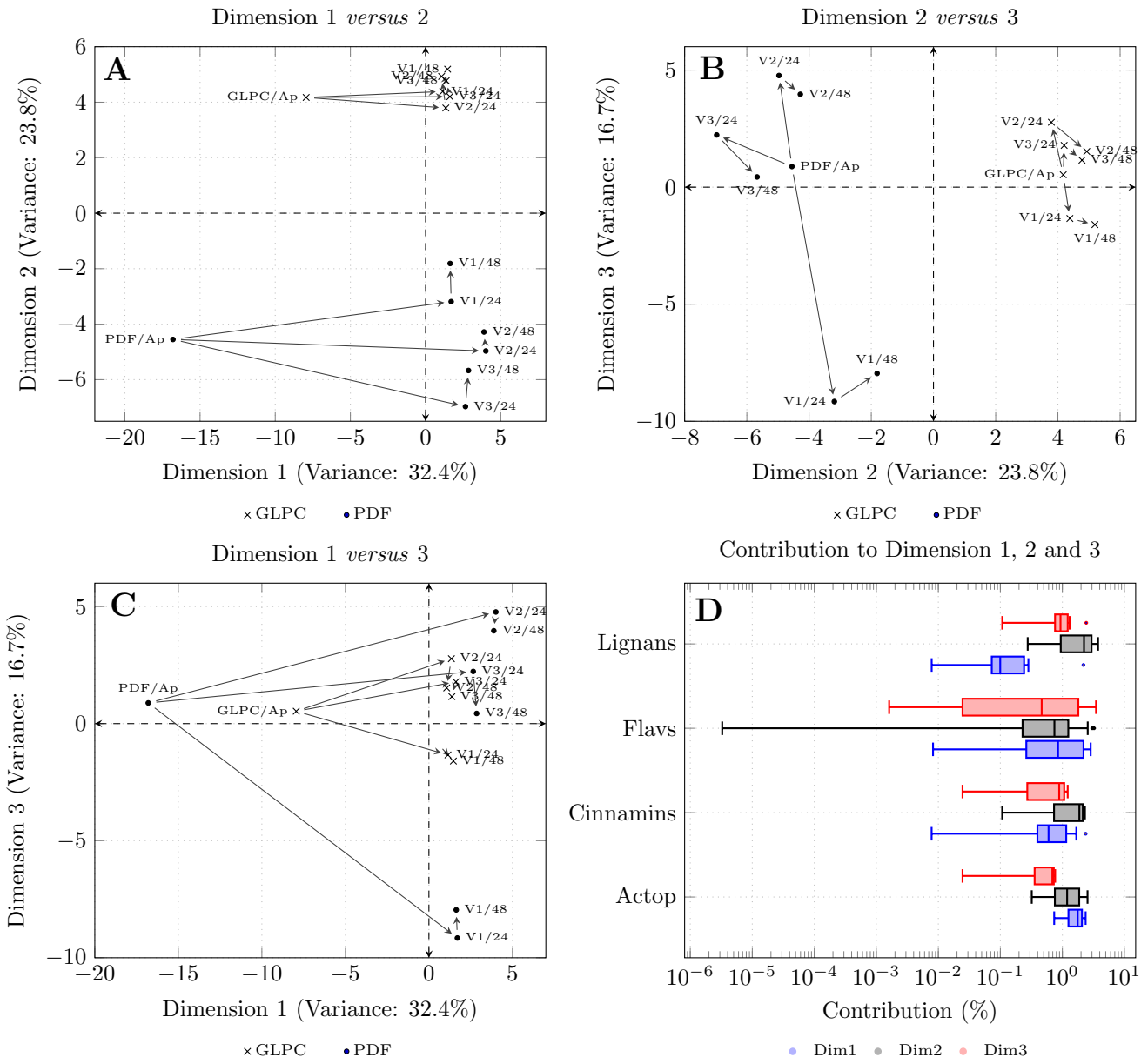


Figure 2: The panels A, B and C show the PCA models from the perspective of Dimensions 1 and 2, Dimensions 2 and 3, and Dimensions 1 and 3 respectively. Panel D shows the distribution of the (%) contribution of the compounds belonging to major metabolite groups. Flavs = Flavanoids/ coumarins, Cinnamins = Cinnamic acids, Actop = Acetophenones, PDF/AP=Protein Depleted Fraction from apical phase of the Caco-2 transport setup, GLPC/AP=Gorse leaf protein concentrate from apical phase of the Caco-2 transport setup. Boxplots in blue mark represents Dimension 1, the boxplot marked in black represents Dimension 2, and the boxplot marked in red represents Dimension 3.

Table 2: Mass of notable compounds in the faecal fermentation (µg/2g) GLPC/PDF

Group	**Compound**	**Volunteer**	**PDF**			**GLPC**		
			Apical	24 h	48 h	**Apical**	24 h	48 h
Benzoic Acids	2, 5-dihydroxybenzoic acid	V1	N/D	1.21±0.10	1.17±0.06	N/D	N/D	0.05±0.00
		V2		0.36±0.02	0.43±0.06		0.03±0.00	0.04±0.00
		V3		0.46±0.03	0.53±0.03		0.04±0.00	0.05±0.01
	vanillic acid	V1	5.12±0.43	23.45±1.00	21.39±0.66	0.42±0.03	0.63±0.02	1.11±0.14
		V2		5.13±0.05	6.42±0.58		0.23±0.02	0.38±0.02
		V3		9.85±0.14	9.74±0.13		0.43±0.01	0.42±0.04
Flavanoids/ Coumarins	daidzein	V1	1.36±0.03	2.27±0.09	2.03±0.15	0.13±0.01	0.20±0.01	0.18±0.01
		V2		1.06±0.04	1.04±0.08		0.16±0.01	0.10±0.01
		V3		1.83±0.02	1.91±0.13		0.16±0.00	0.16±0.02
	equol	V1	N/D	0.00±0.00	0.00±0.00	N/D	N/D	N/D
		V2		0.29±0.01	0.31±0.03			
		V3		0.00±0.00	0.00±0.00			
	genistein	V1	0.43±0.02	2.62±0.07	2.04±0.18	0.26±0.01	0.28±0.02	0.26±0.02
		V2		0.87±0.05	0.81±0.05		0.15±0.00	0.07±0.01
		V3		1.41±0.01	1.61±0.13		0.15±0.00	0.17±0.01
Lignans	enterodiol	V1	N/D	0.00±0.00	0.00±0.00	N/D	N/D	N/D
		V2		0.03±0.00	0.03±0.00			
		V3		0.00±0.00	0.00±0.00			
	enterolactone	V1	N/D	0.10±0.00	0.08±0.01	N/D	0.10±0.01	0.10±0.02
		V2		0.19±0.02	0.19±0.01		0.11±0.01	0.09±0.00
		V3		0.12±0.01	0.11±0.00		0.12±0.01	0.11±0.01
Phenyl-propionic acids	4- hydroxy- 3-methoxyphenyl-propionic acid	V1	N/D	4.39±0.09	3.75±0.07	N/D	0.51±0.01	0.46±0.05
		V2		1.04±0.01	1.11±0.10		0.34±0.02	0.21±0.00
		V3		0.42±0.02	2.93±0.04		0.29±0.00	0.35±0.04

N/D = Not Detected.

3.1.3 PCA including all phases of the digestion model

The metabolites across the entire *in vitro* digestion model were analysed within a single PCA model to identify compounds driving the categorisation of the metabolite profile in each phase of the digestion. This PCA model is shown in **Supplementary Figure 2** and the contribution of each variable (metabolite) is provided in **Supplementary Code (code chunk 7)**. Major compounds driving categorisation were: histamine, putresine, 4-ethylphenol, 3-methyl-indole, protocatachaldehyde, enterodiol and equol. However, the $\cos^2$ values of the points representing the metabolite profile in the colonic fermentation phase was <0.5 indicating poor representation in the model.

3.2 Permeability across the Caco-2 cell culture model

3.2.1 Transport using standard compounds

The permeability of standard compounds across the Caco-2 model was tested individually, as well as an equimolar (50 mM each compound) mixture and the P_{app} values are provided in Table 3. This was performed to check if the permeability of these abundant compounds was additive (i.e. independent of the presence of other compounds) or non-additive (i.e. synergistic, inhibitory, etc.).

In Table 3, chlorogenic acid, kaemperfol and protocatechuic acid showed non-finite permeability ratios either as a single compound or a mixture. This suggests that efflux proteins governs the active and unidirectional transport of these compounds (Farrell et al., 2012). In the case of acetovanillone, the ratio remained constant for individual and mixed conditions suggesting the action of the efflux proteins on this compound as well. Moreover, it was the only compound with similar $P_{app\ (A \rightarrow B)}$ values in single and mixed compound conditions suggesting that its permeability is additive. In the case of p-hydroxybenzoic acid, the $P_{app\ (B \rightarrow A)}$ values were statistically similar suggesting that permeability along that direction was also additive. Interestingly though, the $P_{app\ (A \rightarrow B)}$ of vanillin and kaempferol was inhibited when compounds were mixed.

Table 3: Apparent permeability ($\times 10^{-6}\,cm^{-1}\cdot s^{-1}$) of standard compounds either as single compounds or a 50 mM equimolar mixture.

Compound	**Single**			**Mixed**		
	$\mathbf{P_{app\ (A\rightarrow B)}}$	$\mathbf{P_{app(B\rightarrow A)}}$	**Ratio**	$\mathbf{P_{app\ (A\rightarrow B)}}$	$\mathbf{P_{app(B\rightarrow A)}}$	**Ratio**
mandelic acid	1.28±0.03	0.19±0.02	23.10	7.46	NA	-
chlorogenic acid	0.44±0.01	NA	-	4.33	NA	-
sinapic acid	1.11±0.01	1.75±0.26	0.60	13.13	NA	-
kaempferol	4.36±0.26	NA	-	NA	NA	-
p-hydroxybenzoic acid	1.34±0.08	2.42±3.62	0.55	17.96	0.38±0.06	47.26
ferulic acid	8.43±0.03	17.61±2.55	0.48	16.42	2.34±0.09	7.02
acetovanillone	15.18±0.59	35.58±0.48	0.43	10.47	24.74±0.41	0.42
protocatechuic acid	1.96±0.03	NA	-	14.31	NA	-
vanillin	19.59±0.31	38.29±0.16	0.51	NA	NA	-
p-coumaric acid	19.11±0.25	15.01±1.01	1.27	11.91	1.07±0.05	11.13

Values expressed in mean±standard deviation $\times 10^{-6}\,cm^{-1}\cdot s^{-1}$.
Ratio=$\frac{P_{app\ (A\rightarrow B)}}{P_{app\ (B\rightarrow A)}}$.
N/D=not detected.
Values marked with indicate statistically similar P_{app} values in single compounds as well as mixtures using Student's t-test.

3.2.2 Transport of GLPC/PDF digesta

The P_{app} values of compounds in the PDF and GLPC experimental digesta are given in **Supplementary Table 6**. Among the standard compounds, only ferulic acid (only in PDF $P_{app\ (A\rightarrow B)}$ = 11.89±0.38×$10^{-6}\,cm^{-1}\cdot s^{-1}$) and p-hydroxybenzoic acid (PDF $P_{app\ (A\rightarrow B)}$ = 2.34±0.55×$10^{-6}\,cm^{-1}\cdot s^{-1}$; GLPC $P_{app\ (A\rightarrow B)}$ = 2.62±1.55×$10^{-10}\,cm^{-1}\cdot s^{-1}$) could be detected in the digesta, and showed transport in at least one direction. In both cases, the P_{app} values were significantly different from the values in Table 3, albeit closer to the single compound setup. This suggests that transport of compounds in the digesta across the Caco-2 layer were synergistic/competitive . Furthermore, a one-way ANOVA on the permeability ratio between the GLPC and PDF indicated significant difference for the same compound, ($F_{(1,7)}$ = 11.39, p = 0.012; excluding single direction transport). This suggests a matrix effect on the permeability of these compounds.

Other compounds of interest identified in the digesta profile were daidzein, genistein, apigenin and vitexin which demonstrated low $P_{app\ (A\rightarrow B)}$, but high $P_{app\ (B\rightarrow A)}$ values and thus are expected to be confined to the lumen. Similarly, although daidzein and genistein may be associated to beneficial health outcomes (Poschner et al., 2017), they are more likely to remain in the lumen and be delivered to the

colon where they are metabolised by the gut microbiota.

3.3 Toxicity of plant metabolites to Caco-2 cells

The data provided in the **Supplementary Table 7** shows no significant loss in cell survival for the experimental extracts and their respective digesta. Cell viability did not significantly drop lower than 80% for any treatment, except for camptothecin which was the toxicity-positive control. The tolerance limit of 80% was set by precedence (Farrell et al., 2012) where a significant drop in viability, in tandem with low TEER, suggested to the loss of the monolayer integrity.

The mean cell viability for the pure compounds as well as the digesta was 94.7±13.0% which was not significantly lower than 100% viability using one sample t-test (t = -2.24, df= 29, p= 0.033). For the initial GLPC/PDF samples, the mean viability was 87.1±14.4% and their corresponding digesta showed 84.9±2.8%. A one sample t-test for the Caco-2 viability when exposed to GLPC/PDF digesta was significantly lower than 100% (t = -17.18, df= 9, p= 3.45×10^{-8}), but higher than 80% (t= 5.51, df= 9, p= 3.76×10^{-4}) as well. Toxicity may have been contributed by the enzymes-only control used in the experiment which showed a Caco-2 viability of 28.2±0.5%. This suggests that PDF and GLPC did not adversely affect cell viability in the Caco-2 model. Analyses are shown in **Supplementary Codes (code chunk 16)**.

3.4 Short chain fatty acids (SCFA) profile

The short chain fatty acid content in the faecal fermentation stage was measured at 0 h, 24 h and 48 h (**Supplementary Table 8**). To understand the overall trends in the SCFA profile, a PCA (Figure 3) was performed. The codes and annotations are provided in **Supplementary Codes (code chunk 14)**.

Dimension 1 of the PCA model (Figure 3 panel B, blue boxplots) is mainly influenced by the bacterial metabolism of carbohydrates and proteins, while the separation along Dimension 2 was driven by dicarboxylic acids (Figure 3, panel B, black boxplots).

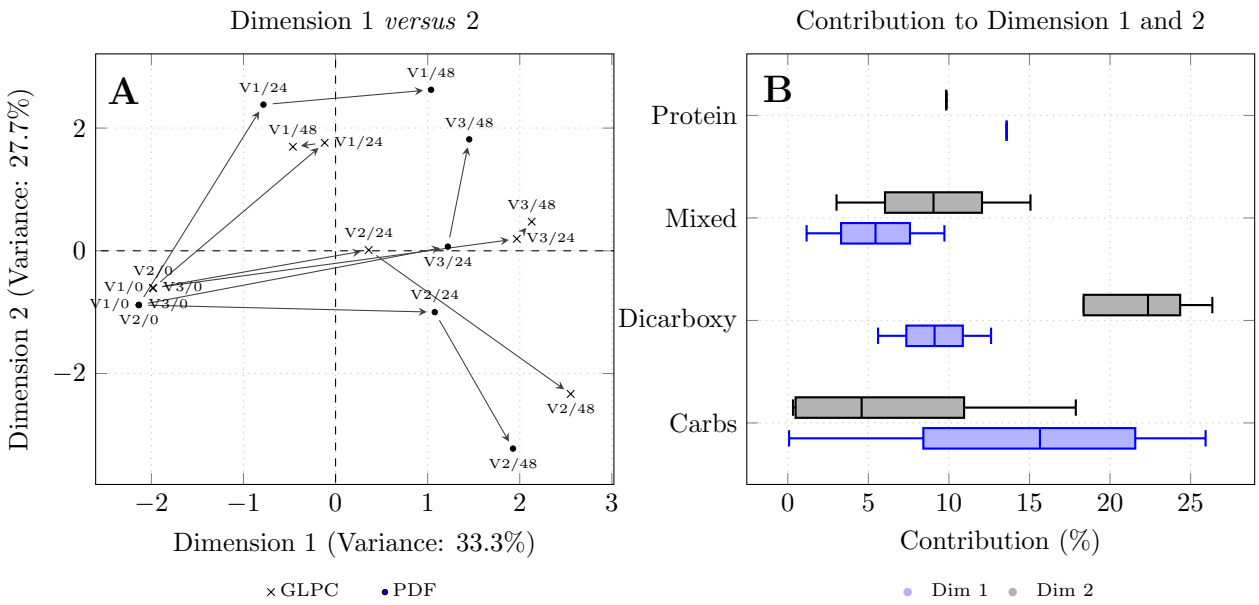


Figure 3: PCA profile of the short chain fatty acids (SCFA) measured in the colonic phase of the in vitro digestion model. The left panel shows the PCA profile of the SCFA at different fermentation times for each volunteer. In panel B, SCFAs are grouped according to the substrates from which they are produced (or dicarboxylic acids that are metabolic intermediates), and their contribution to the PCA profile in the left panel. Carbs=carbohydrates, Dicarboxy=dicarboxylic acids, Mixed= either protein or carbohydrates may be the substrate for the SCFA. Boxplots in blue mark represents Dimension 1, while the boxplot marked in black represents Dimension 2.

In Figure 3 panel B, the median contribution of SCFAs arising from bacterial metabolism of carbohydrates and protein digestion was high for dimension 1 (marked in blue boxplot), while for dimension 2, dicarboxylic acids showed the highest contribution (marked in black boxplot). In the left panel, the relative placement of the points representing the SCFA profile indicates the changes as colonic fermentation occurs over 48 hours. While the points representing the 48 h profile for all volunteers moved away from their starting point, they differentiate among each other along dimension 2. This suggests that dicarboxylic concentrates drives categorisation between the individual volunteers.

The microbiota of volunteer 2 produced significantly higher butyrate (14.62 mg) compared to volunteer 1 (1.93 mg) and 3 (3.89 mg) at the 48 h time point in the presence of the PDF fraction (**Supplementary Table 8**). Conversely, the microbiota from volunteer 2 showed lower levels of dicarboxylic acids (succinate+lactate = 15.99 mg) compared to volunteer 1 (49.58 mg) and 3 (33.02 mg) after 48 h of colonic fermentation. These trends were similar in the presence of the GLPC fractions as well.

4 Discussion

Simulated digestion models provide a controlled and physiologically relevant environment for studying the interactions and evaluating the food safety of novel foods or additives, eliminating the need for animal or human experiments. Leaf protein concentrates from Gorse (GLPC) and its protein depleted fraction (PDF) were subjected to simulated *in vitro* digestion to study the metabolites released and produced as they passed through the alimentary canal.

While previous studies have focused on the digestibility of leaf protein concentrates and used models largely to demonstrate the bio-availability of the protein (Kaur et al., 2021), the work here explored the wider suite of metabolites generated across the digestion process in an attempt to establish benignity. Although recent studies by Prado Massarioli et al. (2023) have used similar *in vitro* digestion models to study the resultant metabolites, the use of anaerobic conditions in this work allowed for closer replication of physiological conditions, and aided *in situ* preservation of metabolites, otherwise lost to oxidation.

4.1 Enzymatic digestion enhances release of metabolites from the food matrix

Based on the mass balance provided in **Supplementary Table 4**, the recovery of the metabolite groups transported to the basolateral phase by the Caco-2 monolayer (indicating their uptake) shows a relatively high value for benzoic acid (12.9%) with flavonoids/coumarins in a distant second with (7.1%) for GLPC. In the case of PDF, cinnamic acids showed the highest recovery in the intestinal phase at 207.5% relative to the initial content. This may be a result of their release from the matrix of residual proteins. For example, phenyllactic acid was high in the basolateral phase of the PDF digesta (424.1±8.4 µg), showing a 6× increase in recovery compared to the initial PDF sample (68.9±0.2 µg) as given in **Supplementary Table 3**. Phenyllactic acids are plant auxins that are phenylalanine catabolites. The Caco-2 cell line used to model the intestinal enterocyte, does not express the necessary enzymes such as phenylalanine hydroxylase (PAH) to produce phenylpyruvic acids (Lichter-Konecki et al., 1999), that is subsequently catabolised to phenyllactic acids. It is likely that freeze-drying and re-dissolution of the PDF intestinal

phase digesta in Dulbecco medium for the Caco-2 setup, may have released phenyllactic acids from the matrix.

4.2 Resurgence of plant metabolites in colonic fermentates

The quantities of flavonoids/ coumarins, benzoic acids, and cinnamic acids dropped significantly post exposure to the Caco-2 monolayer (**Supplementary Table 2** and **5**). The combined quantities of compounds in the apical and basolateral fractions were lower than that measured in the preceding intestinal phase. The quantities of these compound groups rose again during the 48 h faecal fermentation (accounting for quantities measured in the controls). This observation may be explained by conjugation of compounds (glucuronidatin, sulphation, methylation) by the Caco-2 monolayer used to simulate the intestinal lining, and the subsequent deconjugation by microbial action during fermentation.

The glucuronidation of xenobiotic compounds by the Caco-2 cell line is well documented and the mechanisms triggering the relevant biochemical systems have been detailed previously by Abid et al. (1995). Likewise, the modification of dietary phenolics by the colonic microbes has been described by Russell et al. (2013), and remains a topic of further investigation.

4.3 Permeability of the Caco-2 monolayer was sensitive to composite samples

The rate of transport of compounds across a given surface area (transport velocity), is known as apparent permeability (P_{app}). The ratio of $P_{app(A\rightarrow B)}$ to $P_{app(B\rightarrow A)}$ indicates the nature of metabolite interaction with the cell surface receptors. A ratio of <1 indicates interaction of metabolites with the efflux transporters (Turner et al., 2000). In other words, the Caco-2 monolayer actively attempts to stop the movement of these metabolites from the apical phase to the basolateral. A ratio of >1 suggests either higher uptake or efflux activity ensuring metabolites do not permeate from the basolateral into the apical phase (Farrell et al., 2012; Čvorović et al., 2018).

The general trend noted in Table 3, was that for a given compound, the $P_{app\ (A\rightarrow B)}$ was higher in mixed

compound conditions compared to single compounds. Similarly, $P_{app(B \rightarrow A)}$ dropped below detection levels in mixed conditions compared to single standards. Thus while $P_{app(A \rightarrow B)}$ value for chlorogenic acid was comparable to that previously observed by Hithamani et al. (2017), its bioavailability from diet is expected to be significantly different.

This observation aligns well with the expected increase in the expression of efflux transporters in response to increased phenolic load (Huang et al., 2008), and that the chosen standard compounds were non-conjugated phytochemicals. The review by Williamson et al. (2018) provides a good description of the fate of dietary phenolics in the human body. Non-conjugated phenolic entities are transported across the intestinal lining, where first conjugations occur, into the blood from where they are transported to the liver and are conjugated further. These conjugated moieties are transported back into the alimentary canal where they are further metabolised by the gut microbiota.

To replicate this effect *in vitro*, the basolateral fraction of the Caco-2 monolayer can be exposed to the HepG2 cell line followed by a re-exposure to the Caco-2 monolayer to study transport of the conjugated moieties from the basolateral fraction to the apical fraction. Unfortunately, this arm of the experiment could not be executed owing to limitations faced during the pandemic. Nonetheless within the performed experiments, partial dynamics of the major compounds identified in Gorse (Iyer et al., 2021) could be examined.

4.4 Most metabolites in the digesta were delivered to the colon

The PCA of the metabolite profile across the entire digestion model is shown in **Supplementary Figure 2**. The $\cos^2$ values were below even the lenient threshold of 0.5 for the profiles representing the colonic fermentation phase of the model (high quality models generally filter above 0.85). The $\cos^2$ value is a measure of how well an individual is represented in the PCA model and consequently, how well its variance is captured in the model. By bifurcating the model into non-bacterial and bacterial phases, the metabolite profile of the digesta in each phase is better represented in the model and more internally comparable. The microbes in the colonic fermentation phase have a far more diverse metabolic capability compared to the enzymatic treatment and the mammalian Caco-2 cell line and thus may be prudent to

analyse these results independently.

In Figure 1, as GLPC and PDF pass through the digestion model, their overall metabolite profile begins to converge at the basolateral phase of the Caco-2 model. There is low variance between the transported metabolites for PDF and GLPC digesta, suggesting that irrespective of how variable the starting phenolic profile was, the metabolites transported across the Caco-2 monolayer is highly regulated and restrictive. On the other hand the profiles of the apical phase (which represent the fraction expected to stay in the lumen) showed greater variance between each other. Most ingested plant metabolites are thus expected to be delivered to the colon.

4.5 Beneficial microbial metabolites were noted in faecal fermentates

In the faecal fermentation phase of the digestion model, for all volunteers, bacterial metabolism of the phytochemicals such as pinoresinol and secoisolariciresinol to enterolignans such as enterodiol and enterolactone was observed. These metabolites are known to contribute towards beneficial effects, although this model may not truly represent the scale of production and uptake in the human body. Another case is the conversion of daidzein to equol that was noted only in the case for volunteer 2. Despite accounting for the vanillin from the diet, an excess was noted in the supernatants. This may be an indication of the conversion of protocatachaldehyde to vanillin through microbial action.

4.6 SCFA profile may compliment the phytochemical metabolite profile

The values noted for SCFA were based on pooled fermentate samples. Thus, while statistical comparison was not possible, their PCA profile could provide some indication of the overall trends in the bacterial metabolite trends. Formate, acetate, propionate and butyrate were SCFAs arising from carbohydrate digestion. Valerate was a product of protein metabolism. Iso-butyrate and iso-valerate have mixed substrates and may arise either from carbohydrate and protein metabolism. Lastly, dicarboxylic acids

are metabolic intermediates. As shown in panel A of Figure 3, Dimension 1 was largely influenced by carbohydrate and protein metabolites while Dimension 2 was largely influenced by the dicarboxylic acids. Complimenting the phenolic metabolites noted previously, the SCFA profile of the fermentates for volunteer 2 was characterised by low intermediates such as dicarboxylic acids and higher carbohydrate metabolites; particularly butyrate. Correspondingly, the microbiota of volunteer 2 was capable of converting daidzein to equol unlike volunteer 1 and volunteer 3.

4.7 Digesta showed low toxicity to Caco-2 cells

Cellular toxicity was measured as a function of the viability of proliferative Caco-2 cells. In the digesta samples noted in **Supplementary Table 7**, there was no case of overt toxicity noted when exposed to Caco-2. Any toxicity measured was largely associated to the enzymes used in the digestion model rather than the extract digesta.

5 Conclusions

The absence of prior reports on toxicity from Gorse, the absence of overt toxicity in proliferative Caco-2 cells, and the absence of any overtly toxic metabolites noted across the simulated digestion model suggests that the gorse leaf protein concentrates (GLPC) and metabolite-rich protein depleted fraction (PDF) may be safe for human consumption.

A majority of the Gorse bioactives are confined to the apical fraction of the digestion model and are expected to be delivered to the colon. While a highly composite bioactive profile appears to enhance permeability for most compounds, it is further modulated by the sample matrix (*viz.* gorse proteins). Gut microbial composition could potentially synergistically enhance health benefits through the production of enterolactones and equol.

Supplementary Materials

OSF respository link: All supplementary materials may be found at the OSF respository at https://osf.io/9qtyz/

Supplementary Codes: The codes and output for all statistical analyses performed in R is provided in the file `Supplementary Code.pdf`.

Supplementary Materials: The supplementary materials (`Supplementary.pdf` file) contains the the Supplementary Figures 1 and 2, and caption for all the Supplementary Tables.

Supplementary Table 1: The Supplementary Table 1 (`Supplementary_ Table1.csv`) shows the chemical names used for in house identification and data processing, and their corresponding trivial names and compound groups.

Supplementary Table 2: The Supplementary Table 2 (`Supplementary_ Table2.csv`) shows the metabolites measured in the fractions (pg·µL^{-1}) in the *in vitro* digstion model from the initial GLPC/PDF to the Caco2 monolayer.

Supplementary Table 3: The Supplementary Table 3 (`Supplementary_ Table3.csv`) shows the metabolites measured in the fractions (pg·µL^{-1}) in the *in vitro* digstion model from the initial GLPC/PDF to the Caco2 monolayer.

Supplementary Table 4: Groupwise summation of compounds (µg/2g) of initial GLPC/PDF to the Caco2 monolayer phase.

Supplementary Table 5: Groupwise sum of compounds (µg/2g) of Apical fraction of Caco2 phase to 48 hour colonic fermentation phase.

Supplementary Table 6: Apparent permeability of compounds in the Gorse digestates.

Supplementary Table 7: The Supplementary Table 7 (`Supplementary_ Table7.csv`) shows Caco2

viability (%) across experimental fractions at 24 h incubation.

Supplementary Table 8: SCFA content of samples across faecal fermentation (mM/5mL).

Supplementary Figure 1: Process flow of the digestion. Arrows marked in dot-and-dash represent anaerobic setup and those marked in plain represent aerobic setup.

Supplementary Figure 2: Univariate scaled principal component analysis (PCA) of the metabolite profile measured at each phase of the digestion process of the initial GLPC/PDF.

Acknowledgements

The authors thank Jenny Martin, Freda Farquharson, and Gillian Donachie for their support with the faecal fermentation experiment. Louise Cantlay helped measure metabolites using the LC/MS. Dr. Teresa Grohmann, Mr. Timo Kramer and Dr. Elaine Hillesheim helped with the correction and editing of the manuscript.

Funding

This work was funded by the Scottish Government through Rural and Environment Science and Analytical Services Division (RESAS), as part of its strategic funding programme.

References

Abid, A., Bouchon, I., Siest, G., & Sabolovic, N. (1995). Glucuronidation in the caco-2 human intestinal cell line: Induction of UDP-glucuronosyltransferase 16. *Biochemical Pharmacology*, *50*(4), 557–561. https://doi.org/10.1016/0006-2952(95)00162-S

Broadfield, N., & McHenry, M. T. (2019). A World of Gorse: Persistence of Ulex europaeus in Managed Landscapes. *Plants*, *8*(11), 523. https://doi.org/10.3390/plants8110523

Brodkorb, A., Egger, L., Alminger, M., Alvito, P., Assunção, R., Ballance, S., Bohn, T., Bourlieu-Lacanal, C., Boutrou, R., Carrière, F., Clemente, A., Corredig, M., Dupont, D., Dufour, C., Edwards, C., Golding, M., Karakaya, S., Kirkhus, B., Le Feunteun, S., ... Recio, I. (2019). INFOGEST static in vitro simulation of gastrointestinal food digestion. *Nature Protocols*, *14*(4), 991–1014. https://doi.org/10.1038/s41596-018-0119-1

Came, H., Warbrick, I., McCreanor, T., & Baker, M. (2020). From gorse to ngahere: An emerging allegory for decolonising the New Zealand health system. *The New Zealand Medical Journal*, *133*(1524), 102–110.

Cerveira, I., Baptista, V., Teodósio, M. A., & Morais, P. (2022). What's for dinner? Assessing the value of an edible invasive species and outreach actions to promote its consumption. *Biological Invasions*, *24*(3), 815–829. https://doi.org/10.1007/s10530-021-02685-3

Cocîrlea, M. D., & Oancea, S. (2022). Impact of invasive plant species on ecosystems, biodiversity loss and their potential industrial applications. *Current Trends in Natural Sciences*, *11*(21), 88–103. https://doi.org/10.47068/ctns.2022.v11i21.009

Cummings, J., & Macfarlane, G. (1991). The control and consequences of bacterial fermentation in the human colon. *Journal of Applied Bacteriology*, *70*(6), 443–459. https://doi.org/10.1111/j.1365-2672.1991.tb02739.x

Čvorović, J., Ziberna, L., Fornasaro, S., Tramer, F., & Passamonti, S. (2018). Bioavailability of Flavonoids: The Role of Cell Membrane Transporters. In *Polyphenols: Mechanisms of Action in Human Health and Disease* (pp. 295–320). Elsevier. https://doi.org/10.1016/B978-0-12-813006-3.00022-2

Dewan, P., Kaur, I., Chattopadhya, D., A Faridi, M. M., & Agarwal, K. N. (2007). A pilot study on the effects of curd (dahi) & leaf protein concentrate in children with protein energy malnutrition (PEM). *The Indian Journal of Medical Research*, *126*(3), 199–203.

Farrell, T. L., Poquet, L., Dew, T. P., Barber, S., & Williamson, G. (2012). Predicting Phenolic Acid Absorption in Caco-2 Cells: A Theoretical Permeability Model and Mechanistic Study. *Drug Metabolism and Disposition*, *40*(2), 397–406. https://doi.org/10.1124/dmd.111.041665

Hithamani, G., Kizhakayil, D., & Srinivasan, K. (2017). Uptake of phenolic compounds from plant foods in human intestinal Caco-2 cells. *Journal of Biosciences*, *42*(4), 603–611. https://doi.org/10.1007/s12038-017-9705-6

Huang, R. S., Murry, D. J., & Foster, D. R. (2008). Role of xenobiotic efflux transporters in resistance to vincristine. *Biomedicine & Pharmacotherapy*, *62*(2), 59–64. https://doi.org/10.1016/j.biopha.2007.04.011

Hubatsch, I., Ragnarsson, E. G. E., & Artursson, P. (2007). Determination of drug permeability and prediction of drug absorption in Caco-2 monolayers. *Nature Protocols*, *2*(9), 2111–2119. https://doi.org/10.1038/nprot.2007.303

Hughes, T., & Wiles, C. (1996). Clinical measurement of swallowing in health and in neurogenic dysphagia. *QJM*, *89*(2), 109–116. https://doi.org/10.1093/qjmed/89.2.109

Iyer, A., Bestwick, C. S., Duncan, S. H., & Russell, W. R. (2021). Invasive Plants Are a Valuable Alternate Protein Source and Can Contribute to Meeting Climate Change Targets. *Frontiers in Sustainable Food Systems*, *5*, 575056. https://doi.org/10.3389/fsufs.2021.575056

Iyer, A., Guerrier, L., Leveque, S., Bestwick, C. S., Duncan, S. H., & Russell, W. R. (2022). High throughput method development and optimised production of leaf protein concentrates with potential to support the agri-industry. *Journal of Food Measurement and Characterization*, *16*(1), 49–65. https://doi.org/10.1007/s11694-021-01136-w

Jonas, D., Antignac, E., Antoine, J.-M., Classen, H.-G., Huggett, A., Knudsen, I., Mahler, J., Ockhuizen, T., Smith, M., Teuber, M., Walker, R., & De Vogel, P. (1996). The safety assessment of novel foods. *Food and Chemical Toxicology*, *34*(10), 931–940. https://doi.org/10.1016/S0278-6915(96)00061-0

Kassambara, A., & Mundt, F. (2020). *Factoextra: Extract and Visualize the Results of Multivariate Data Analyses* (Version 1.0.7). https://CRAN.R-project.org/package=factoextra
Programmers: _:n3203.

Kaur, L., Lamsar, H., López, I. F., Filippi, M., Ong Shu Min, D., Ah-Sing, K., & Singh, J. (2021). Physico-Chemical Characteristics and In Vitro Gastro-Small Intestinal Digestion of New Zealand Ryegrass Proteins. *Foods*, *10*(2), 331. https://doi.org/10.3390/foods10020331

Lichter-Konecki, U., Hipke, C. M., & Konecki, D. S. (1999). Human Phenylalanine Hydroxylase Gene Expression in Kidney and Other Nonhepatic Tissues. *Molecular Genetics and Metabolism*, *67*(4), 308–316. https://doi.org/10.1006/mgme.1999.2880

Minekus, M., Alminger, M., Alvito, P., Ballance, S., Bohn, T., Bourlieu, C., Carrière, F., Boutrou, R., Corredig, M., Dupont, D., Dufour, C., Egger, L., Golding, M., Karakaya, S., Kirkhus, B., Le

Feunteun, S., Lesmes, U., Macierzanka, A., Mackie, A., ... Brodkorb, A. (2014). A standardised static *in vitro* digestion method suitable for food – an international consensus. *Food Funct.*, *5*(6), 1113–1124. https://doi.org/10.1039/C3FO60702J

Mishan, L., & Hamada, K. (2020). When Invasive Species Become the Meal [newspaper]. *The New York Times: T Magazine*. Retrieved August 8, 2023, from https://www.nytimes.com/2020/10/02/t-magazine/eating-invasive-species.html

Mulet-Cabero, A.-I., Egger, L., Portmann, R., Ménard, O., Marze, S., Minekus, M., Le Feunteun, S., Sarkar, A., Grundy, M. M.-L., Carrière, F., Golding, M., Dupont, D., Recio, I., Brodkorb, A., & Mackie, A. (2020). A standardised semi-dynamic *in vitro* digestion method suitable for food – an international consensus. *Food & Function*, *11*(2), 1702–1720. https://doi.org/10.1039/C9FO01293A

Natoli, M., Leoni, B. D., D'Agnano, I., Zucco, F., & Felsani, A. (2012). Good Caco-2 cell culture practices. *Toxicology in Vitro*, *26*(8), 1243–1246. https://doi.org/10.1016/j.tiv.2012.03.009

Poschner, S., Maier-Salamon, A., Zehl, M., Wackerlig, J., Dobusch, D., Pachmann, B., Sterlini, K. L., & Jäger, W. (2017). The Impacts of Genistein and Daidzein on Estrogen Conjugations in Human Breast Cancer Cells: A Targeted Metabolomics Approach. *Frontiers in Pharmacology*, *8*, 699. https://doi.org/10.3389/fphar.2017.00699

Prado Massarioli, A., Giovanini De Oliveira Sartori, A., Francetto Juliano, F., Eduardo Pedroso Gomes Do Amaral, J., Cavalcanti Dos Santos, R., Maria De Lima, L., & Matias De Alencar, S. (2023). Simulated gastrointestinal digestion/Caco-2 cell model to predict bioaccessibility and intestinal permeability of p-coumaric acid and p-coumaroyl derivatives in peanut. *Food Chemistry*, *400*, 134033. https://doi.org/10.1016/j.foodchem.2022.134033

R Core Team. (2021). *R: A Language and Environment for Statistical Computing* (Version 4.1.2). Vienna, Austria. https://www.R-project.org/

Programmers: _:n2750.

Richardson, A., Calder, A., Stewart, C., & Smith, A. (1989). Simultaneous determination of volatile and non-volatile acidic fermentation products of anaerobes by capillary gas chromatography. *Letters in Applied Microbiology*, *9*(1), 5–8. https://doi.org/10.1111/j.1472-765X.1989.tb00278.x

Russell, W. R., Duncan, S. H., Scobbie, L., Duncan, G., Cantlay, L., Calder, A. G., Anderson, S. E., & Flint, H. J. (2013). Major phenylpropanoid-derived metabolites in the human gut can arise from microbial fermentation of protein. *Molecular Nutrition & Food Research*, *57*(3), 523–535. https://doi.org/10.1002/mnfr.201200594

Russell, W. R., Gratz, S. W., Duncan, S. H., Holtrop, G., Ince, J., Scobbie, L., Duncan, G., Johnstone, A. M., Lobley, G. E., Wallace, R. J., Duthie, G. G., & Flint, H. J. (2011). High-protein, reduced-carbohydrate weight-loss diets promote metabolite profiles likely to be detrimental to colonic health. *The American Journal of Clinical Nutrition*, *93*(5), 1062–1072. https://doi.org/10.3945/ajcn.110.002188

Rymer, L. (1979). Ethnobotany and Native Distribution of Gorse (*Ulex europaeus* L.) in Britain. *Environmental Conservation*, *6*(3), 211–213. https://doi.org/10.1017/S0376892900003064

Schiller, C., Frohlich, C.-P., Giessmann, T., Siegmund, W., Monnikes, H., Hosten, N., & Weitschies, W. (2005). Intestinal fluid volumes and transit of dosage forms as assessed by magnetic resonance imaging. *Alimentary Pharmacology and Therapeutics*, *22*(10), 971–979. https://doi.org/10.1111/j.1365-2036.2005.02683.x

Stone, N. L., England, T. J., & O'Sullivan, S. E. (2019). A Novel Transwell Blood Brain Barrier Model Using Primary Human Cells. *Frontiers in Cellular Neuroscience*, *13*, 230. https://doi.org/10.3389/fncel.2019.00230

Turner, J. R., Black, E. D., Ward, J., Tse, C.-M., Uchwat, F. A., Alli, H. A., Donowitz, M., Madara, J. L., & Angle, J. M. (2000). Transepithelial resistance can be regulated by the intestinal brush-border Na^+/H^+ exchanger NHE3. *American Journal of Physiology-Cell Physiology*, *279*(6), C1918–C1924. https://doi.org/10.1152/ajpcell.2000.279.6.C1918

Vilà, M., & Ibáñez, I. (2011). Plant invasions in the landscape. *Landscape Ecology*, *26*(4), 461–472. https://doi.org/10.1007/s10980-011-9585-3

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., … Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Williamson, G., Kay, C. D., & Crozier, A. (2018). The Bioavailability, Transport, and Bioactivity of Dietary Flavonoids: A Review from a Historical Perspective. *Comprehensive Reviews in Food Science and Food Safety*, *17*(5), 1054–1112. https://doi.org/10.1111/1541-4337.12351

Ziska, L. H., Blumenthal, D. M., Runion, G. B., Hunt, E. R., & Diaz-Soltero, H. (2011). Invasive species and climate change: An agronomic perspective. *Climatic Change*, *105*(1-2), 13–42. https://doi.org/10.1007/s10584-010-9879-5

List of Tables

1 Mass balance of metabolites across initial sample to Caco-2 transportation phase. . . . 15

2 Mass of notable compounds in the faecal fermentation (µg/2g) GLPC/PDF 18

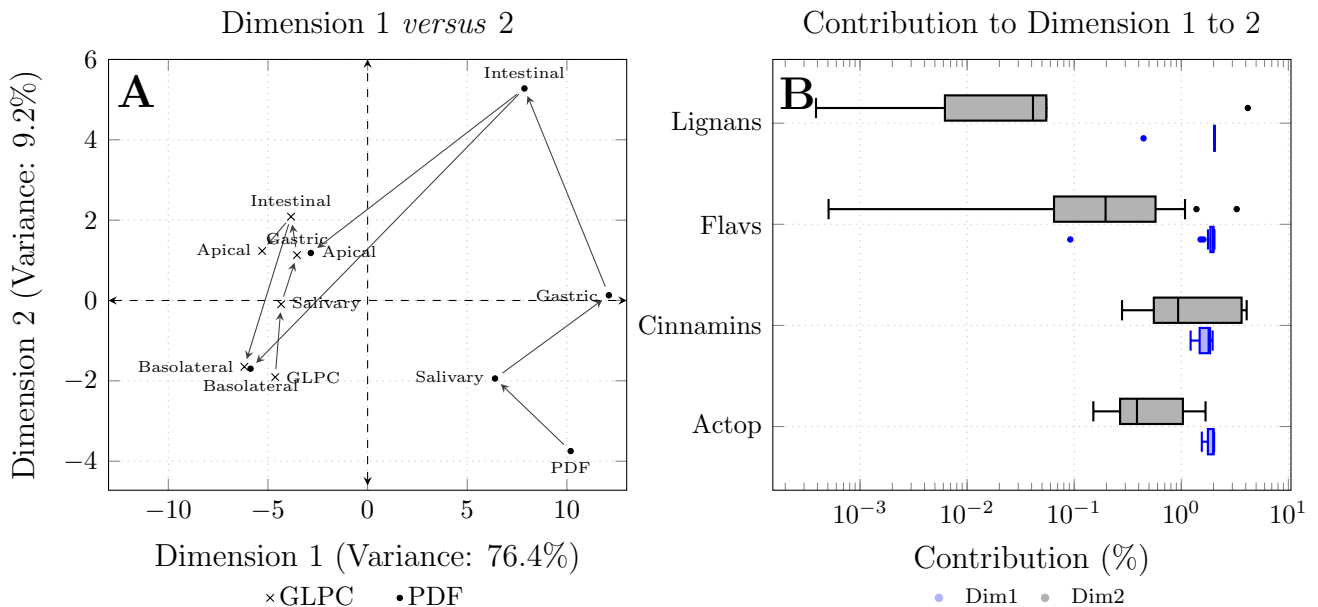
3 Apparent permeability ($\times\ 10^{-6}\,\mathrm{cm}^{-1}\cdot\mathrm{s}^{-1}$) of standard compounds either as single compounds or a 50 mM equimolar mixture. 20

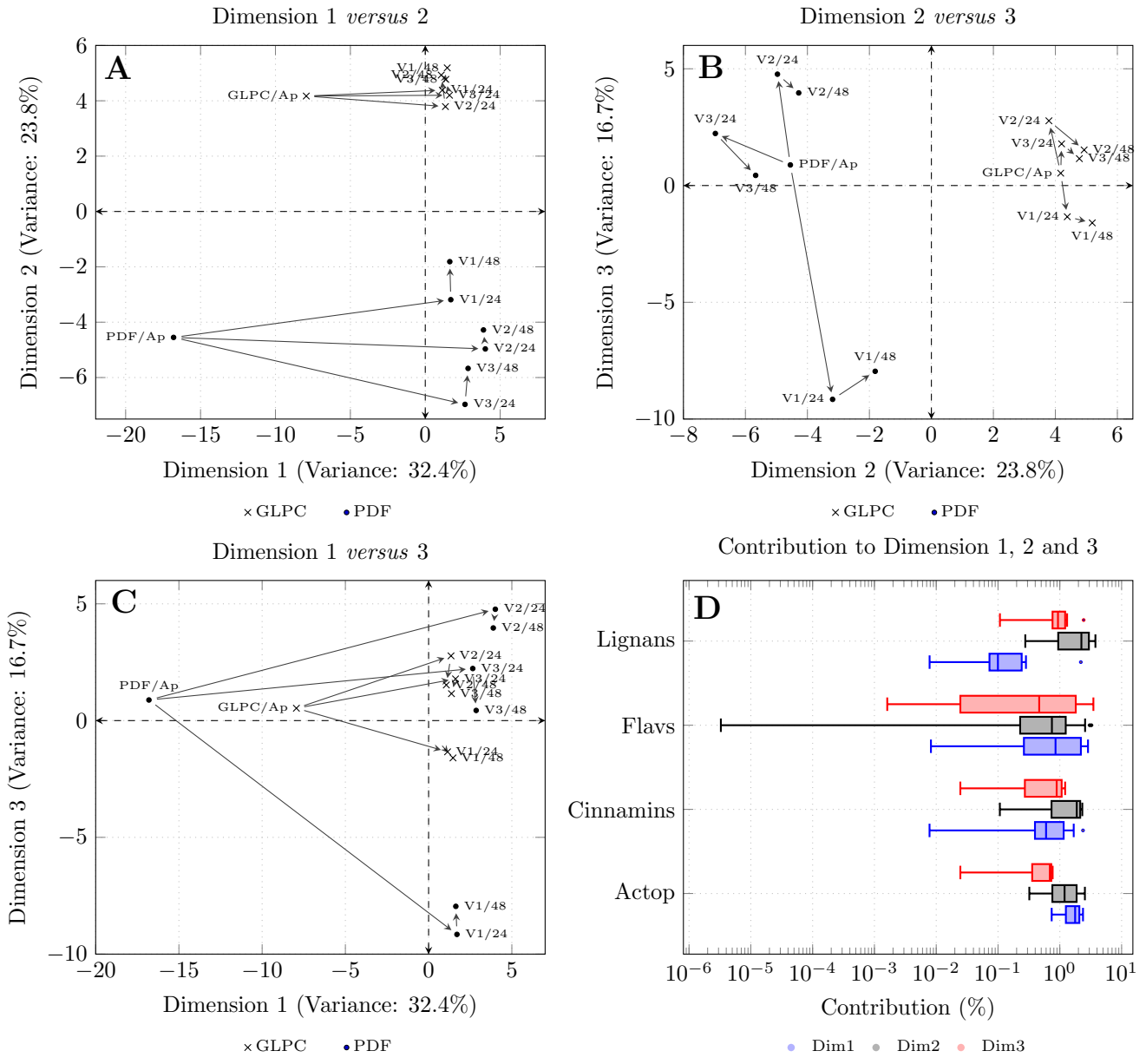
List of Figures

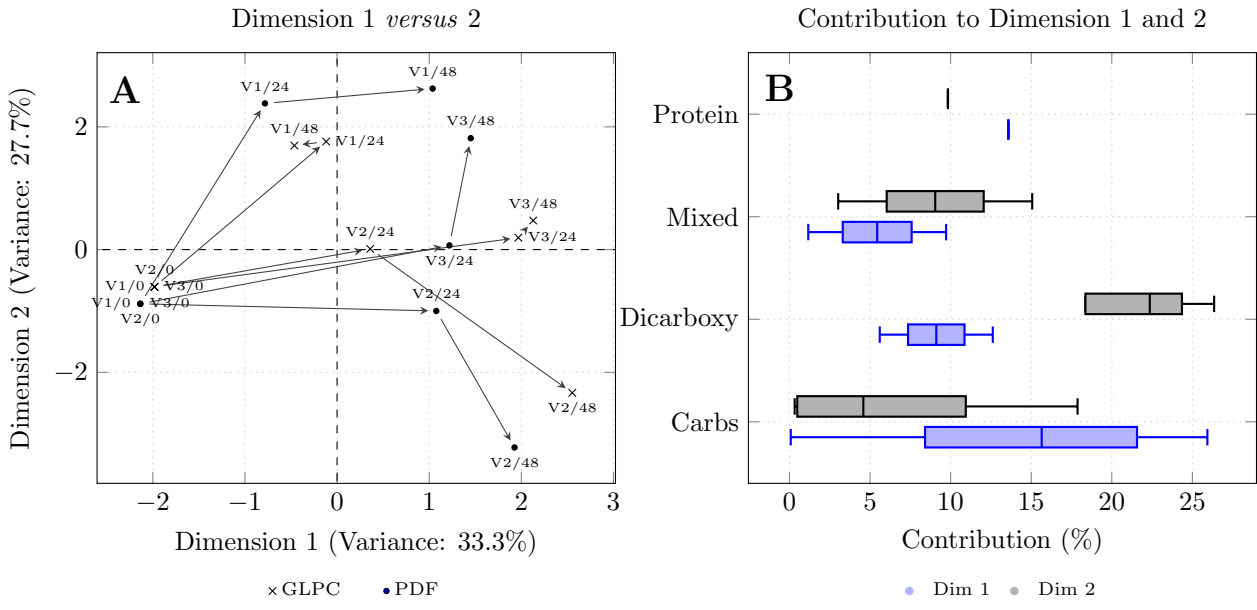
1 The panel A is the PCA plot of the metabolite profiles from the initial (GLPC/PDF) sample to the Caco-2 transport model. Panel B uses boxplots to show the distribution of the (%) contribution of the compounds belonging to major metabolite groups. Flavs = Flavanoids/coumarins, Cinnamins = Cinnamic acids, Actop = Acetophenones. Boxplots in blue mark represents Dimension 1, while the boxplot marked in black represents Dimension 2. 14

2 The panels A, B and C show the PCA models from the perspective of Dimensions 1 and 2, Dimensions 2 and 3, and Dimensions 1 and 3 respectively. Panel D shows the distribution of the (%) contribution of the compounds belonging to major metabolite groups. Flavs = Flavanoids/ coumarins, Cinnamins = Cinnamic acids, Actop = Acetophenones, PDF/AP=Protein Depleted Fraction from apical phase of the Caco-2 transport setup, GLPC/AP=Gorse leaf protein concentrate from apical phase of the Caco-2 transport setup. Boxplots in blue mark represents Dimension 1, the boxplot marked in black represents Dimension 2, and the boxplot marked in red represents Dimension 3. 17

3 PCA profile of the short chain fatty acids (SCFA) measured in the colonic phase of the in vitro digestion model. The left panel shows the PCA profile of the SCFA at different fermentation times for each volunteer. In panel B, SCFAs are grouped according to the substrates from which they are produced (or dicarboxylic acids that are metabolic intermediates), and their contribution to the PCA profile in the left panel. Carbs=carbohydrates, Dicarboxy=dicarboxylic acids, Mixed= either protein or carbohydrates may be the substrate for the SCFA. Boxplots in blue mark represents Dimension 1, while the boxplot marked in black represents Dimension 2. 22


Dimension 1 versus 2
A
Intestinal
Intestinal
Gastric
Apical
Apical
Salivary
Gastric
Basolateral
GLPC
Basolateral
Salivary
PDF
6
4
2
0
−2
−4
−10
−5
0
5
10
Dimension 2 (Variance: 9.2%)
Dimension 1 (Variance: 76.4%)
×GLPC
•PDF
Contribution to Dimension 1 to 2
B
Lignans
Flavs
Cinnamins
Actop
10^{-3}
10^{-2}
10^{-1}
10^{0}
10^{1}
Contribution (%)
Dim1
Dim2


Dimension 1 versus 2
A
Dimension 2 (Variance: 23.8%)
Dimension 1 (Variance: 32.4%)
GLPC/Ap
PDF/Ap
V1/48
V1/24
V2/48
V2/24
V3/48
V3/24
GLPC
PDF
Dimension 2 versus 3
B
Dimension 3 (Variance: 16.7%)
Dimension 2 (Variance: 23.8%)
Dimension 1 versus 3
C
Dimension 3 (Variance: 16.7%)
Dimension 1 (Variance: 32.4%)
Contribution to Dimension 1, 2 and 3
D
Lignans
Flavs
Cinnamins
Actop
Contribution (%)
Dim1
Dim2
Dim3


Dimension 1 *versus* 2
A
V1/24
V1/48
V1/48
V1/24
V3/48
V3/48
V3/24
V2/24
V3/24
V2/0
V1/0
V3/0
V1/0
V3/0
V2/0
V2/24
V2/48
V2/48
2
0
−2
−2
−1
0
1
2
3
Dimension 2 (Variance: 27.7%)
Dimension 1 (Variance: 33.3%)
× GLPC
• PDF
Contribution to Dimension 1 and 2
B
Protein
Mixed
Dicarboxy
Carbs
0
5
10
15
20
25
Contribution (%)
Dim 1
Dim 2