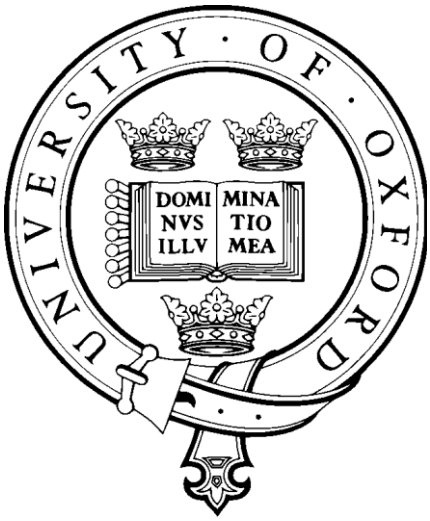
Sequencing of the 16S rRNA gene of the gut microbiota in mice treated with a high fibre diet as a potential endogenous mechanism of radiosensitisation

Chee Kin Then

Balliol College, University of Oxford

Thesis submitted for the degree of Master of Science by Research

Trinity Term 2019



CRUK/MRC Oxford Institute for Radiation Oncology

Department of Oncology, University of Oxford

Abstract

New non-toxic radiosensitisers are needed in the treatment of muscle-invasive bladder cancer because elderly patients are very vulnerable to chemotherapy-related toxicity of currently available radiosensitisers. Histone deacetylase (HDAC) inhibition is a promising mechanism of radiosensitisation. Tumour cells can accumulate butyrate which is produced by the gut microbiota via fibre fermentation to sufficient levels to cause HDAC inhibition, due to the Warburg effect. This has been seen when mice were treated with high-fibre diets. I hypothesise that the gut microbiota in mice fed a high fibre diet provide a sufficiently high systemic butyrate level for this to act as an endogenous mechanism of radiosensitisation. The specific aims are to examine the impact of the diet on the microbiome, and subsequently correlate diet-induced microbiome changes with tumour growth and response to radiation treatment.

RT112 xenografts were used in this study. 16S rRNA gene sequencing was performed on a MiSeq platform and metagenomic analysis was conducted using a QIIME2 platform.

The faecal (n=59) and caecal (n=59) microbiomes from mice were similar and they shared the 3 major taxa (> 80% abundance): *Bacteroidales*, *Clostridiales* and *Verrucomicrobiales*. Significantly higher abundance of *Bacteroides acidifaciens* ($p < 0.001$) existed in the gut microbiome of mice fed soluble high fibre (HF) for 2 weeks. Principal coordinate analysis showed a notable clustering effect within groups, indicating that diet indeed modified the faecal microbiome. Survival analysis by log-rank test showed soluble HF conferred survival benefits with delayed tumour growth after irradiation ($n = 32$, $p = 0.045$). Comparison of the gut microbiome of responders ($n = 4$) and non-responders ($n = 4$) to irradiation in the soluble HF group revealed higher abundance of *B. acidifaciens* in responding mice ($p < 0.05$). Predictive metagenomic profiling showed gut microbiota in responders to be enriched for carbohydrate metabolism. To investigate the correlation between specific bacterial taxa and tumour response to irradiation, all mice fed different diets ($n = 32$) were pooled together. Univariate linear regression revealed a statistically significant positive correlation between the survival time of mice and abundance of the *B. acidifaciens* ($R^2 = 0.528$, $P < 0.001$).

In conclusion, high fibre diets sensitised RT112 xenografts to irradiation by modifying the gut microbiome and this phenotype was associated with higher relative abundance of *B. acidifaciens*. Therefore, *B. acidifaciens* could be a potential radiosensitiser in bladder cancer.

Table of contents

Table of figures

Table of tables

Abbreviations

ABC	ATP-binding cassette
B. acidifaciens	*Bacteroides acidifaciens*
BBN	N-butyl-N-(4-hydroxybutyl)-nitrosamine
CFU	colony formation unit
E. Coli	Escherichia coli
FDA	false-discovery rate
Gy	Gray
HAT	histone acetyltransferase
HDAC	histone deacetylase
HF	high fibre
Insoluble HF	insoluble high fibre
IR	irradiation
KEGG	Kyoto Encyclopedia of Genes and Genomes
LDA	linear discriminant analysis
LEFSe	linear discriminant analysis effect size
LF	low fibre
LF+B	low fibre plus butyrate
MIBC	muscle-invasive bladder cancer
Mixed HF	mixed high fibre
OTU	operational taxonomic unit
PCoA	principal coordinates analysis
PICRUSt	Phylogenetic Investigation of Communities by Reconstruction of Unobserved States
SCFA	short chain fatty acid
SDS	sodium dodecyl sulfate
Soluble HF	soluble high fibre

Introduction

Bladder cancer and the challenges in treating elderly patients

Bladder cancer is the 5th commonest cancer in UK and the incidence in 2015 was 10,171 cases (Meng et al., 2017). The median age of bladder cancer patients in the UK is approximately 73 years old (Erlich & Zlotta, 2016). This disease is more prevalent in men than in women (3.38:1 ratio) and disease incidence increases with age (Figure 1) (Bray et al., 2018). Microscopic or macroscopic (visible) blood in the urine (haematuria) is the most common symptom and haematuria is associated with advanced pathological stage (Ramirez et al., 2016).

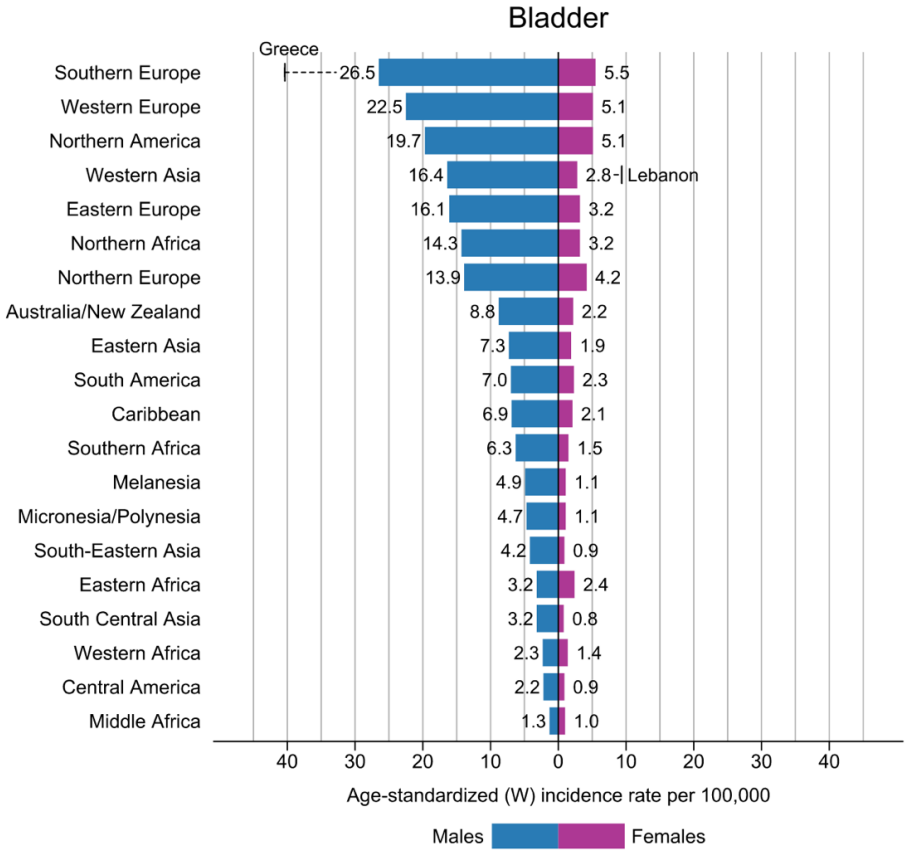


Figure 1. Region-specific incidence by sex for bladder cancer in 2018. Rates are shown in descending order of the world age-standardised rate among men. From Bray et al, 2018.

Bladder cancer can be classified as non-muscle-invasive (75%) and muscle-invasive bladder cancer (MIBC; 25%) (Figure 2) (Sanli et al., 2017). Ta and T1 refer to papillary tumours that are confined to mucosa or have invaded the lamina propria (submucosa). Tis (carcinoma in situ) is a flat tumour confined to mucosa. T2a or T2b tumours have invaded the superficial or deep muscle layer. T3 tumours have invaded beyond the muscularis propria into perivesical tissue. T4a tumours have invaded the prostate, uterus, vagina and/or bowel, whereas T4b tumours have invaded the pelvic or abdominal walls.

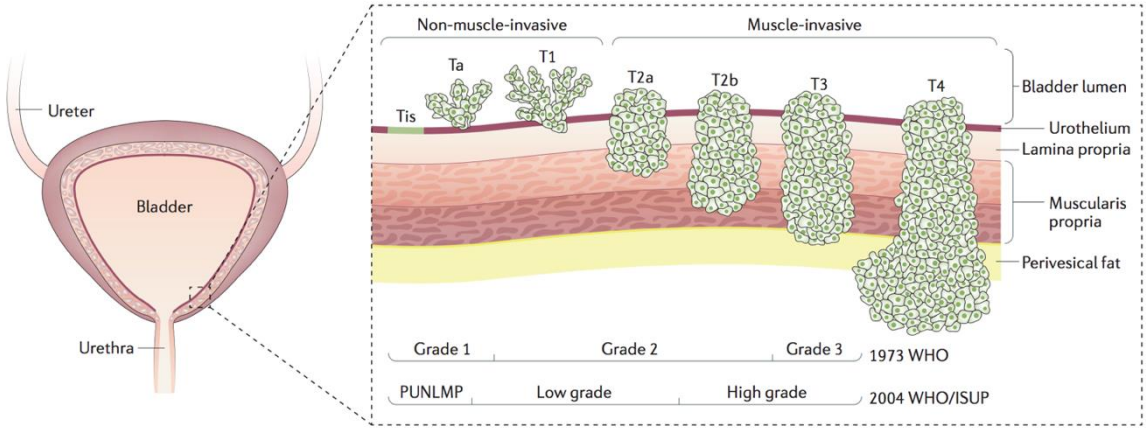


Figure 2. Staging of bladder cancer. It is generally classified as non-muscle-invasive (Tis, Ta, Ta) and muscle invasive-bladder-cancer (T2a, T2b, T3, T4). ISUP, International Society of Urological Pathology; PUNLMP, papillary urothelial malignancy of low malignant potential. From Sanli et al, 2017.

Bladder cancer generally originates from the bladder urothelium and it is a consequence of chronic exposure of the urothelium to carcinogens (Choi et al., 2014). In MIBC, luminal and basal-like subtypes are the two major intrinsic molecular subsets defined by the University of North Carolina (UNC) (Figure 3) (Damrauer et al., 2014), although more recently further subclassification has been described (Robertson et al., 2018).

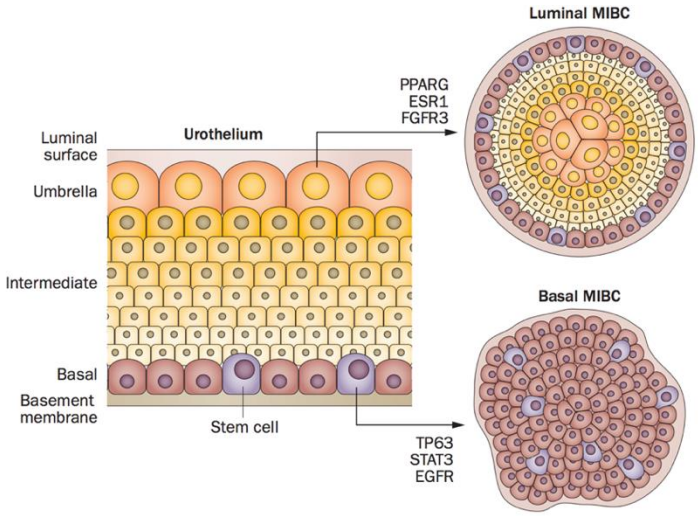


Figure 3. Molecular subtypes of bladder cancer. The biomarkers are PPARG, ESR1 and FGFR3 for luminal MIBC, and TP63, STAT3 and EGFR for basal MIBC. From Choi et al, 2014.

Muscle-invasive bladder cancer has a very poor survival rate, with around 45% five-year survival rate regardless of the type of treatment (Grossman et al., 2003; Stein et al., 2001). Surgical removal of the bladder (cystectomy) is one of the current radical approaches for treatment of MIBC. However, urinary diversion after cystectomy can profoundly reduce the quality of life in elderly patients (Mohamed et al., 2012). Most of MIBC patients responded to cisplatin-based combination chemotherapy, but estimates of disease-specific survival suggest that chemotherapy has only a 5–10% benefit over conventional surgery (Sternberg et al., 2013). Radical radiotherapy has been considered as an alternative to cystectomy, with similar cause-specific survival rates (Kotwal et al., 2008). Chemoradiation has been another promising treatment for MIBC for some years. Radiotherapy with concurrent carbogen and nicotinamide significantly improved overall survival rate, risk of death, and local relapse over radiotherapy alone (Hoskin et al., 2010). In a phase II setting, concurrent gemcitabine-based chemoradiotherapy provided a high response rate in MIBC with acceptable toxicity (Choudhury et al., 2011). Compared with radiotherapy alone, the combination of fluorouracil and mitomycin C with radiotherapy significantly improved locoregional control of bladder cancer (James et al., 2012). However, elderly patients are very vulnerable to treatment-related toxicity, secondary to medical comorbidities and geriatric syndromes (Guancial et al., 2015).

Double-strand break (DSB) response

Radiotherapy, like most anti-tumoural treatments, provides its therapeutic effect by inducing DNA double-strand breaks (DSBs) and thereafter cell death (Baskar et al., 2014). DSBs are predominantly repaired by either homologous recombination or non-homologous end joining (Figure 4) (Lans et al., 2012). Homologous recombination utilises a sister chromatid as a template in late S or G2 phase of the cell cycle to repair DNA damage and it is an error-free repair process. By contrast, non-homologous end joining is a relatively error-prone repair pathway. Non-homologous end joining can rejoin broken DNA ends in all phases of cell cycle. However, DNA trimming, which causes loss of a few nucleotides, is needed before ligation can occur.

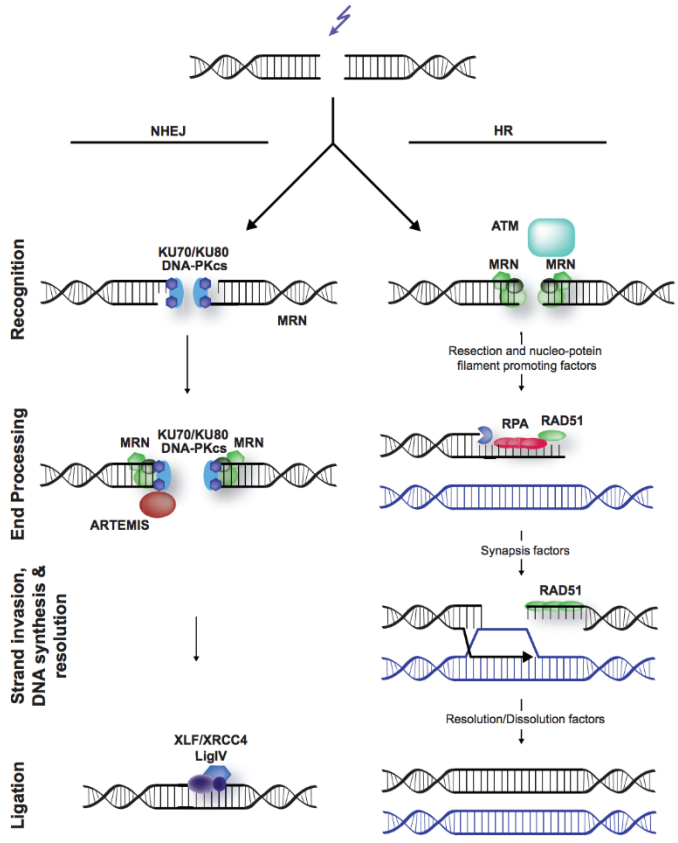


Figure 4. Double-strand break repair. In NHEJ, the broken DNA ends are bound by the KU70/KU80 heterodimer, which recruits DNA-PKcs. The MRN complex further processes the DNA ends prior to ligation executed by XRCC4, XLF/NHEJ1 and the LigIV complex. In contrast to NHEJ, HR is initiated by DNA end-resection, involving the MRN complex. ATM is further recruited to phosphorylate histone H2AX. Single-stranded DNA generated by DNA end-resection is bound by RPA, which is subsequently replaced by RAD51. RAD51 promotes the invasion of the single-stranded DNA into a homologous double-stranded DNA template, leading to repair process. From Lans et al, 2012.

Histone deacetylases and histone deacetylase inhibitors

Histone deacetylases (HDACs) are enzymes which are classified as class I (HDAC1, 2, 3, and 8), IIa (HDAC4, 5, 7, and 9), IIb (HDAC6 and 10), III (sirtuins), and IV (HDAC11) (Figure 5) (Bolden et al., 2006). HDACs have been found to be important cancer therapeutic targets. HDAC expression is aberrant in multiple cancer types, including bladder cancer (Li & Seto, 2016; Ozawa et al., 2010). For example, up-regulation of HDAC2, 8 and down-regulation of HDAC4, 5, and 7 were shown in urothelial cancers (Niegisch et al., 2013). Therefore, histone deacetylase inhibitors have been used as anti-tumoural agents that induce tumour cell death, differentiation or cell cycle arrest (Bolden et al., 2006).

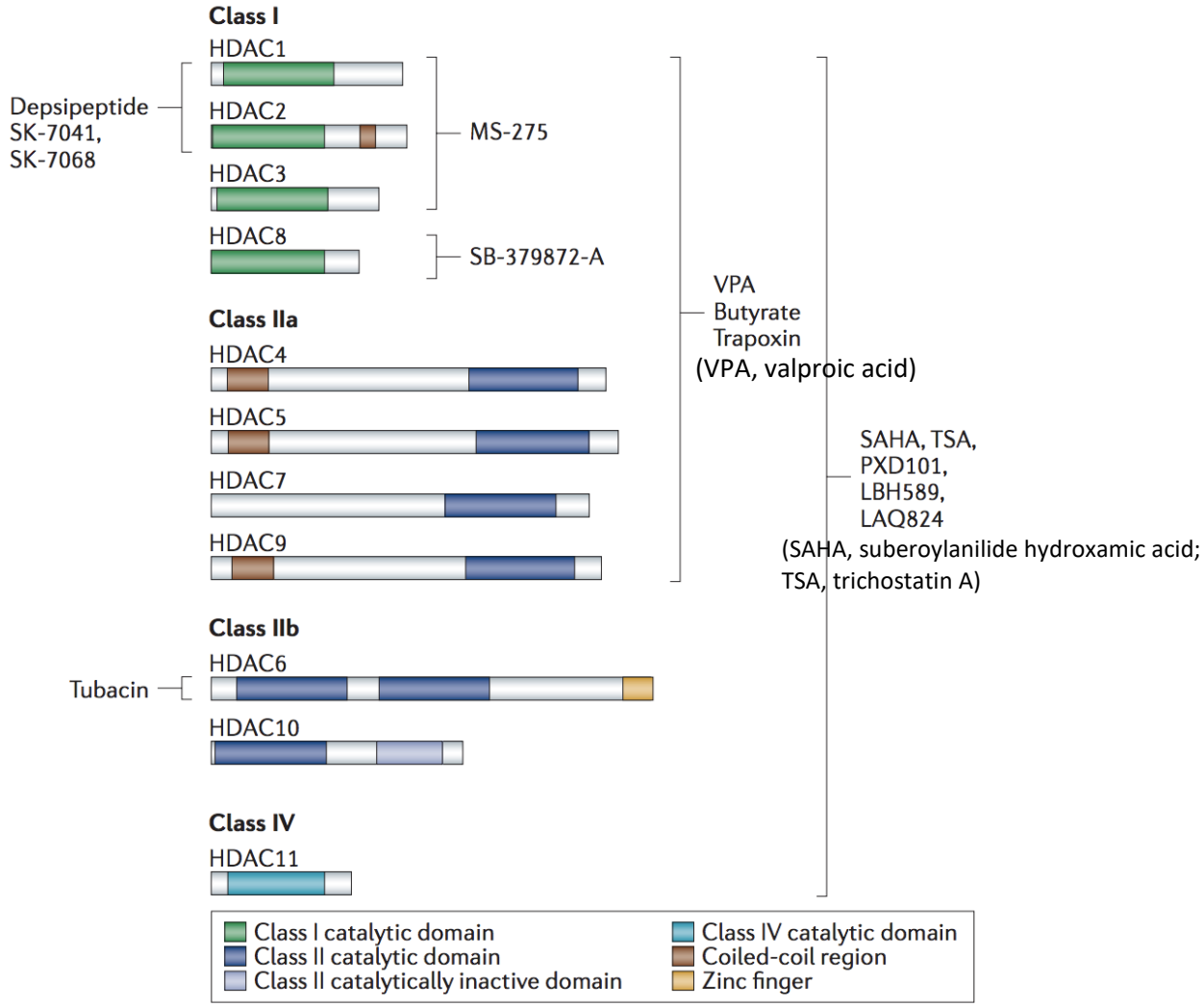


Figure 5. The histone deacetylase family. Class I (HDAC1, 2, 3 and 8), class IIa (HDAC4, 5, 7 and 9), class IIb (HDAC6 and 10) and class IV (HDAC11). Different classes of HDAC inhibitors, eg. SAHA, VPA, are also shown. Adapted from Bolden et al, 2006.

Histone deacetylase (HDAC) inhibitors have been proposed as radiosensitisers, and are thought to act by inhibiting the DNA damage response (Groselj et al., 2013). Groselj et al. demonstrated that panobinostat, a histone deacetylase inhibitor, is a promising radiosensitiser, which delayed the growth of RT112 bladder cancer xenografts but did not exacerbate local acute (3.75 days) and late (12 weeks) toxicity to the intestinal tract (Figure 6) (Groselj et al., 2018). Nicholson et al. proposed a mechanism for HDAC inhibition sensitising bladder cancer to radiation, via cIAP2-mediated downregulation of MRE11 (Nicholson et al., 2017). Taken together, HDAC inhibitors are promising radiosensitisers which enhance anti-tumour response.

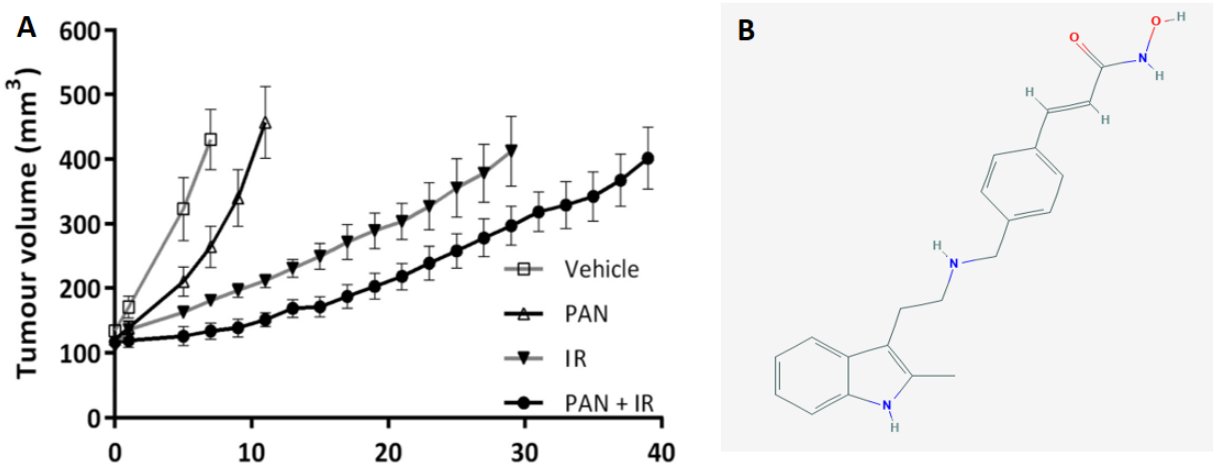


Figure 6. Histone deacetylase inhibition is a promising mechanism of radiosensitisation that enhances the response of bladder tumours to irradiation. (A) RT112 xenografts were treated with vehicle (n =5), panobinostat (n = 7), IR (5x4 Gy; n = 6), or panobinostat + IR (n = 6). From Groselj et al, 2018. (B) Chemical structure of pan-DAC inhibitor panobinostat (LBH589) (NCBI).

Dietary fibre

Dietary fibre is classified as soluble and insoluble and both forms have a physiological impact on human health (Dhingra et al., 2012). Soluble fibre includes inulin, pectin, gum, mucilage, β-glucan, fructo-oligosaccharide, alginate, phyllium, and some hemicelluloses; insoluble fibre includes cellulose, lignin, resistant starch and other hemicelluloses (Figure 7) (Lattimer & Haub, 2010). Soluble fibre can reduce serum cholesterol levels (Gunness & Gidley, 2010), while insoluble fibre increases gut transit time (de Vries et al., 2015). Both of them improve insulin resistance (C. Chen et al., 2016; Weickert & Pfeiffer, 2018), and modulate gut hormones and inflammatory markers (Weickert & Pfeiffer, 2008).

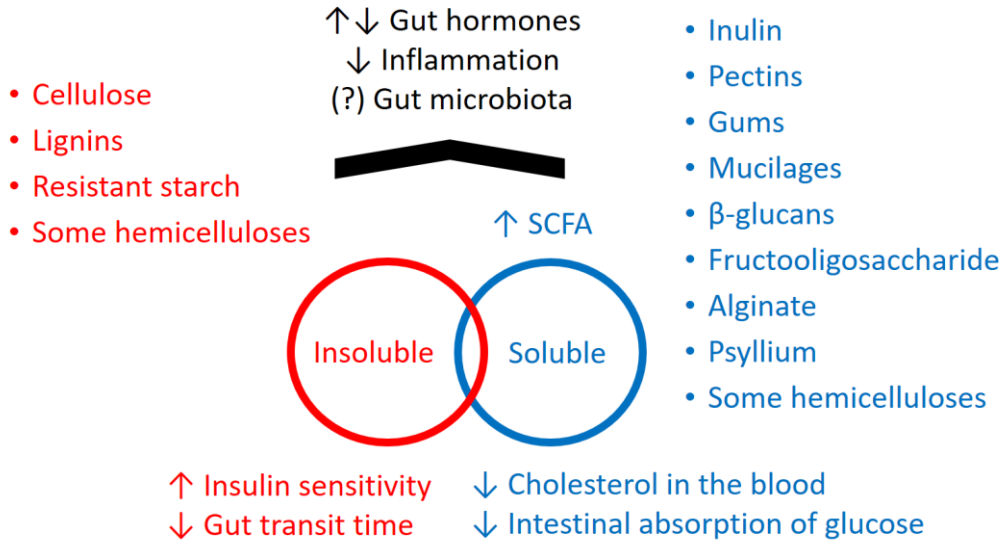


Figure 7. Metabolic effects of dietary fibre. SCFA production tends to be more pronounced with soluble dietary fibre. Adapted from Weickert & Pfeiffer, 2008.

Short chain fatty acids

Acetate (CH3 -COO−), propionate (CH3 -CH2 -COO−) and butyrate (CH3 -CH2 -CH2 -COO−) are the three main short chain fatty acids (SCFAs), which are fermented from carbohydrates by the gut microbiota (Figure 8). These SCFAs infiltrate into the circulatory system and play important roles in human physiology. Acetate increases colonic blood flow and oxygen. Propionate prevents proliferation and induces apoptosis in colorectal cancer cells (Riviere et al., 2016).

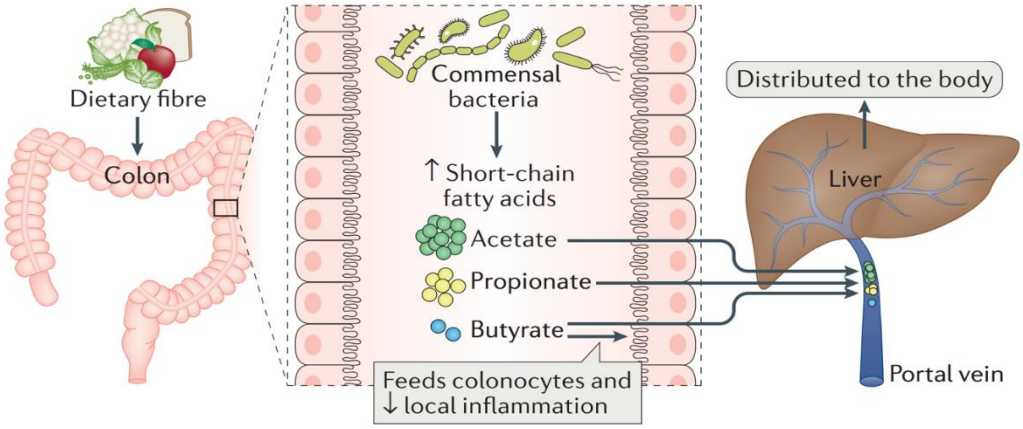


Figure 8. Short-chain fatty acids. Dietary fibre is fermented by commensal bacteria. Short-chain fatty acids (SCFAs), namely, acetate, propionate and butyrate, are generated. Butyrate accounts for 80% of the SCFAs produced by the microbiota. Adapted from Marques et al, 2018.

Butyrate has been proposed to reduce oxidative stress in the colon, improve gut barrier function and protect against colon cancer and colitis (Bourassa et al., 2016). A high fibre diet has been shown to significantly increase SCFA-producing bacteria and SCFA levels (Bishehsari et al., 2018). Moreover, high fibre intake has been demonstrated to significantly increase butyrate levels in plasma and tumours in mice (Figure 9) (Wei et al., 2016).

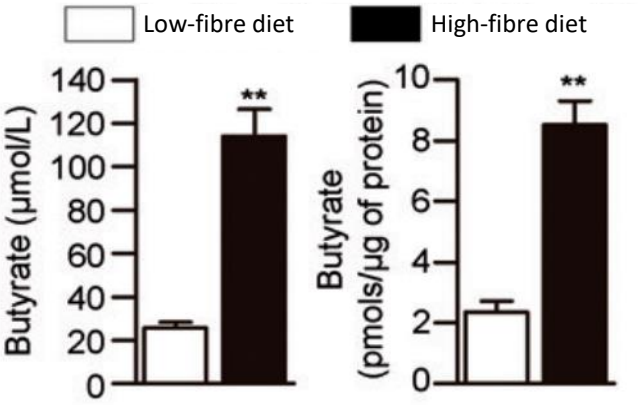


Figure 9. A high fibre diet increased butyrate concentrations in plasma and tumours in mice.
Butyrate levels in the high-fibre diet group were four-fold higher than that in low-fibre diet group. Adapted from Wei et al, 2016.

The Warburg effect, butyrate and histone deacetylase inhibition

The Warburg effect, the metabolic reprogramming of cancer cells to aerobic glycolysis, provides the primary energy source in urothelial cancer cells, promoting tumourigenesis (Massari et al., 2016). High glycolytic flux increases glycine synthesis and purine synthesis, which sustains uncontrolled proliferation of tumour cells (Massari et al., 2016). In normal colonocytes, butyrate can act as the primary energy source, which provides about 70% of ATP production, via mitochondrial β-oxidation (Roediger, 1980) . However, in cancer cells, decreased use of the Kreb's cycle, along with upregulation of the pentose phosphate pathway and fatty-acid synthesis, results in more butyrate accumulating in the nucleus, to a level sufficient for histone deacetylase inhibition (Figure 10) (Donohoe et al., 2012). Butyrate is fermented from dietary fibre. Its levels are increased by consumption of a high fibre diet and it is involved in cellular activities which include mitochondrial activity, G-protein coupled receptors, microbiome homeostasis and HDAC inhibition (Bourassa et al., 2016). Davie proposed that butyrate inhibits HDAC1 and HDAC2 which are recruited by SP1 and SP3, and leads to histone hyperacetylation (Davie, 2003). Butyrate also stimulates the activity of histone acetyltransferase (HAT) by increasing the essential cofactor - acetyl-CoA (Donohoe et al., 2012).

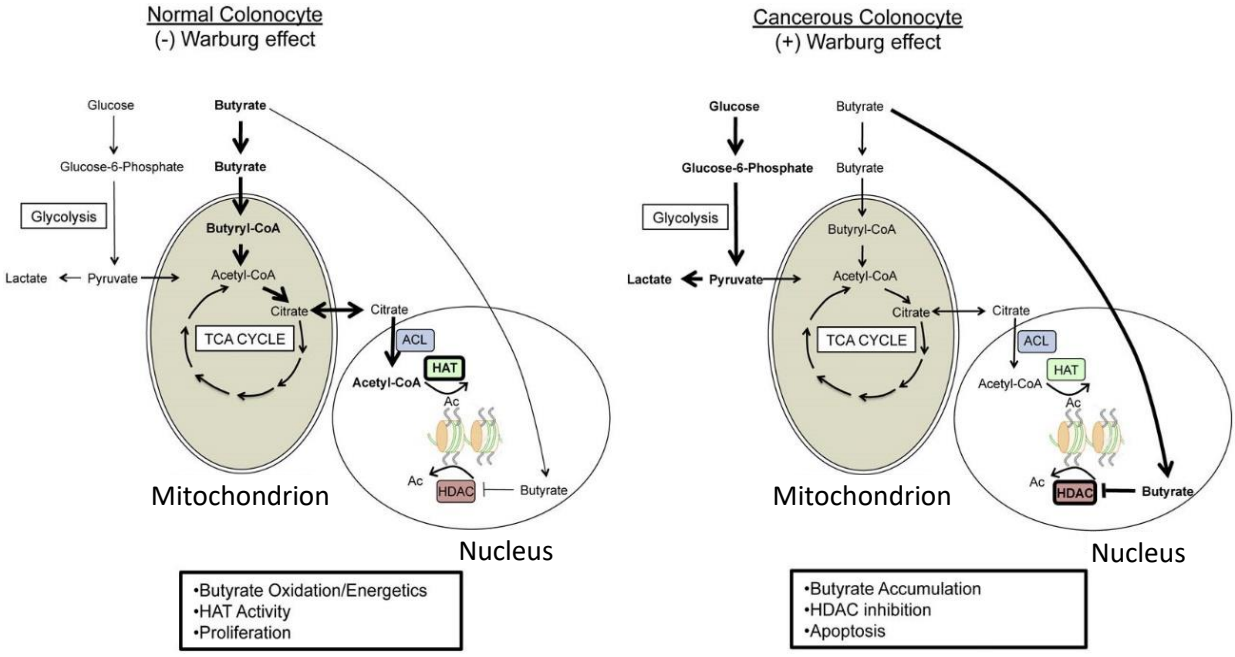


Figure 10. The Warburg effect causes butyrate accumulation and HDAC inhibition. Normal colonocytes consume butyrate as their primary energy source. Due to the Warburg effect, cancerous colonocytes consume glucose as their primary energy source so that butyrate accumulates to levels sufficient to cause HDAC inhibition. Adapted from Donohoe et al, 2012.

Gut microbiota and gut microbiome

The gut microbiota comprises more than 100 trillion microbes and contains 10 times more cells than the total number of somatic and germ cells in the body (Savage, 1977). The gut microbiome contains 100 times the number of genes in the human genome (Backhed et al., 2005) (see glossary for definitions of terms). The density of bacteria in the colon approaches 10^{11}–10^{12} cells/ml (Whitman et al., 1998). As we develop and change from infancy to old age, so does our microbiota. More than 90% of all bacterial phylogenetic types (phylotypes) belong to just two of the 70 known divisions (phyla) in the domain: the *Bacteroidetes* and the *Firmicutes*. The gut microbial community needs to resist perturbation by opportunistic or fast growing species and maintain its diversity and stability (Lozupone et al., 2012). The gut microbiota has a huge beneficial impact on human physiological functions, such as maintaining mucous membrane integrity, extraction of nutrients for metabolite production, regulation of blood glucose and lipid profiles, protection against enteropathogens and preservation of normal immune function (Okumura & Takeda, 2018). The complex interactions between microorganisms and humans are not yet fully understood.

Taxonomy and phylogenetics

Taxonomy is the identification system used to classify biological organisms (see glossary for further explanation). In this classification, taxonomic rank is the relative level of a taxon (a group of organisms) based on a taxonomic hierarchy. Examples of taxonomic ranks are kingdom, phylum, class, order, family, genus and species (Figure 11).

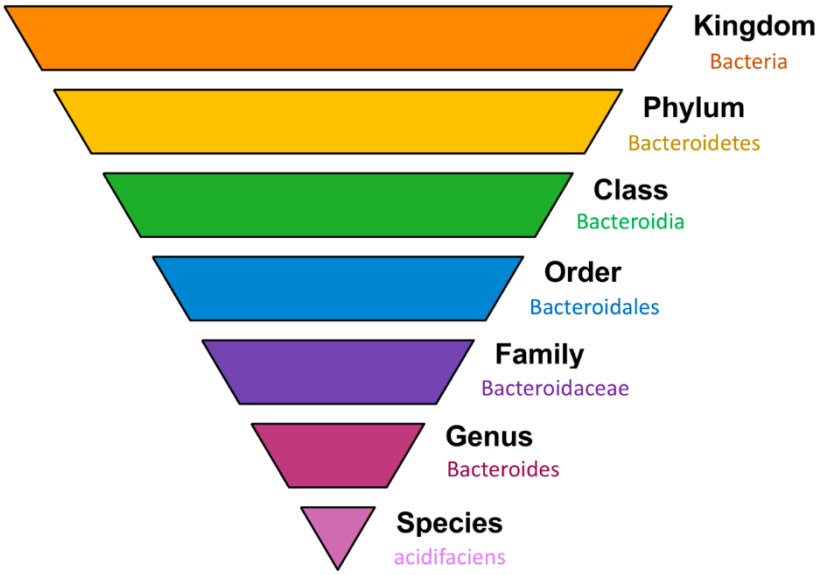


Figure 11. Taxonomic classification. The major ranks are applied to *Bacteroides acidifaciens*.

Phylogenetics is the study of the evolutionary history and relationships among species. These relationships are discovered through DNA sequences under a model of evolution. The result of these analyses is a phylogeny (or phylogenetic tree) – a diagrammatic depiction of the evolutionary relationships of different species.

Butyrate-producing colonic bacterial species

Butyrate producers form a functional cohort in the gut microbiota, and the *Lachnospiraceae* and *Ruminococcaceae*, comprising around 20% of the total bacteria, are the major butyrate-producing families. By mapping the bacterial genome to the genes which are involved in the butyrate synthesis pathways (Figure 12), the potential butyrate producers have been identified as belonging to the bacterial phyla *Firmicutes, Actinobacteria, Bacteroidetes, Fusobacteria, Proteobacteria, Spirochaetes* and *Thermotogae*.

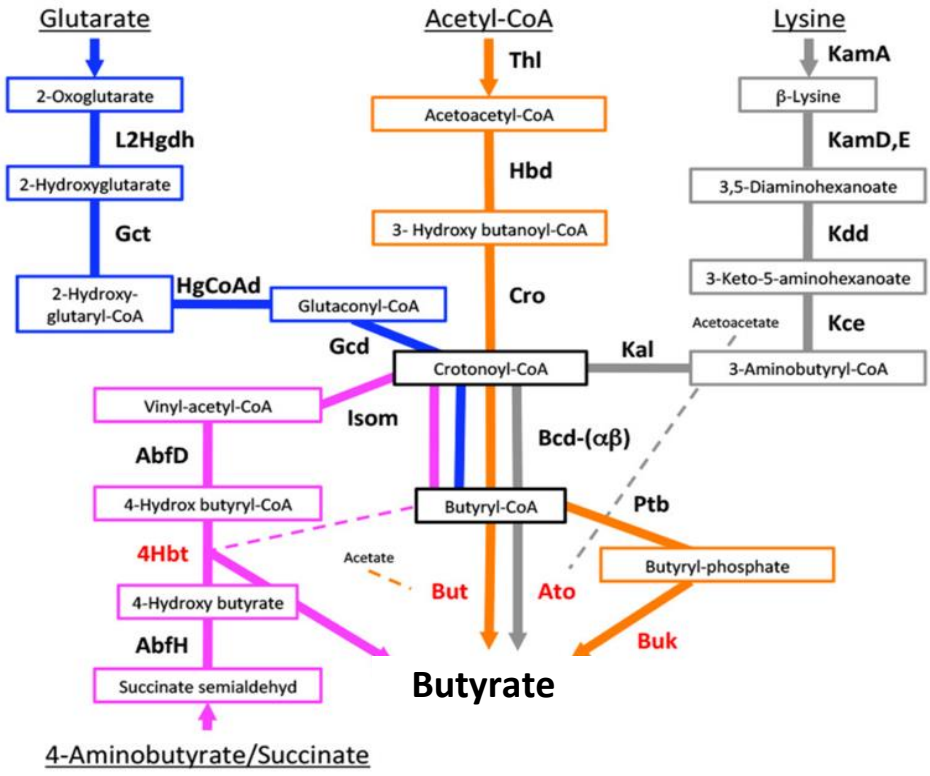


Figure 12. Pathway of butyrate synthesis. Four different butyrate synthesis pathways could be classified by their substrates – glutarate, acetyl-coA, lysine and 4-aminobutyrate/succinate. L2Hgdh, 2-hydroxyglutarate dehydrogenase; Gct, glutaconate CoA transferase (α, β subunits); HgCoAd, 2-hydroxy-glutaryl-CoA dehydrogenase (α, β, γ subunits); Gcd, glutaconyl-CoA decarboxylase (α, β subunits); Thl, thiolase; hbd, β-hydroxybutyryl-CoA dehydrogenase; Cro, crotonase; Bcd, butyryl-CoA dehydrogenase (including electron transfer protein α, β subunits); KamA, lysine-2,3-aminomutase; KamD,E, β-lysine-5,6-aminomutase (α, β subunits); Kdd, 3,5-diaminohexanoate dehydrogenase; Kce, 3-keto-5-aminohexanoate cleavage enzyme; Kal, 3-aminobutyryl-CoA ammonia lyase; AbfH, 4-hydroxybutyrate dehydrogenase; AbfD, 4-hydroxybutyryl-CoA dehydratase; Isom, vinylacetyl-CoA 3,2-isomerase (same protein as AbfD): 4Hbt, butyryl-CoA:4-hydroxybutyrate CoA transferase; But, butyryl-CoA:acetate CoA transferase; Ato, butyryl-CoA:acetoacetate CoA transferase (α, β subunits); Ptb, phosphate butyryltransferase; Buk, butyrate kinase. Cosubstrates for individual butyryl-CoA transferases are shown. Adapted from Vital et al, 2014.

Metagenomics and 16S rRNA sequencing

16S ribosomal RNA (rRNA) is a component of the 30S small subunit of the prokaryotic ribosome and is coded by the 16S rRNA gene (Figure 13). Bacterial 16S rRNA gene sequencing is the general approach used to survey microbial communities by reconstructing phylogenies (Yarza et al., 2014), using the species-specific sequences within given hypervariable regions.

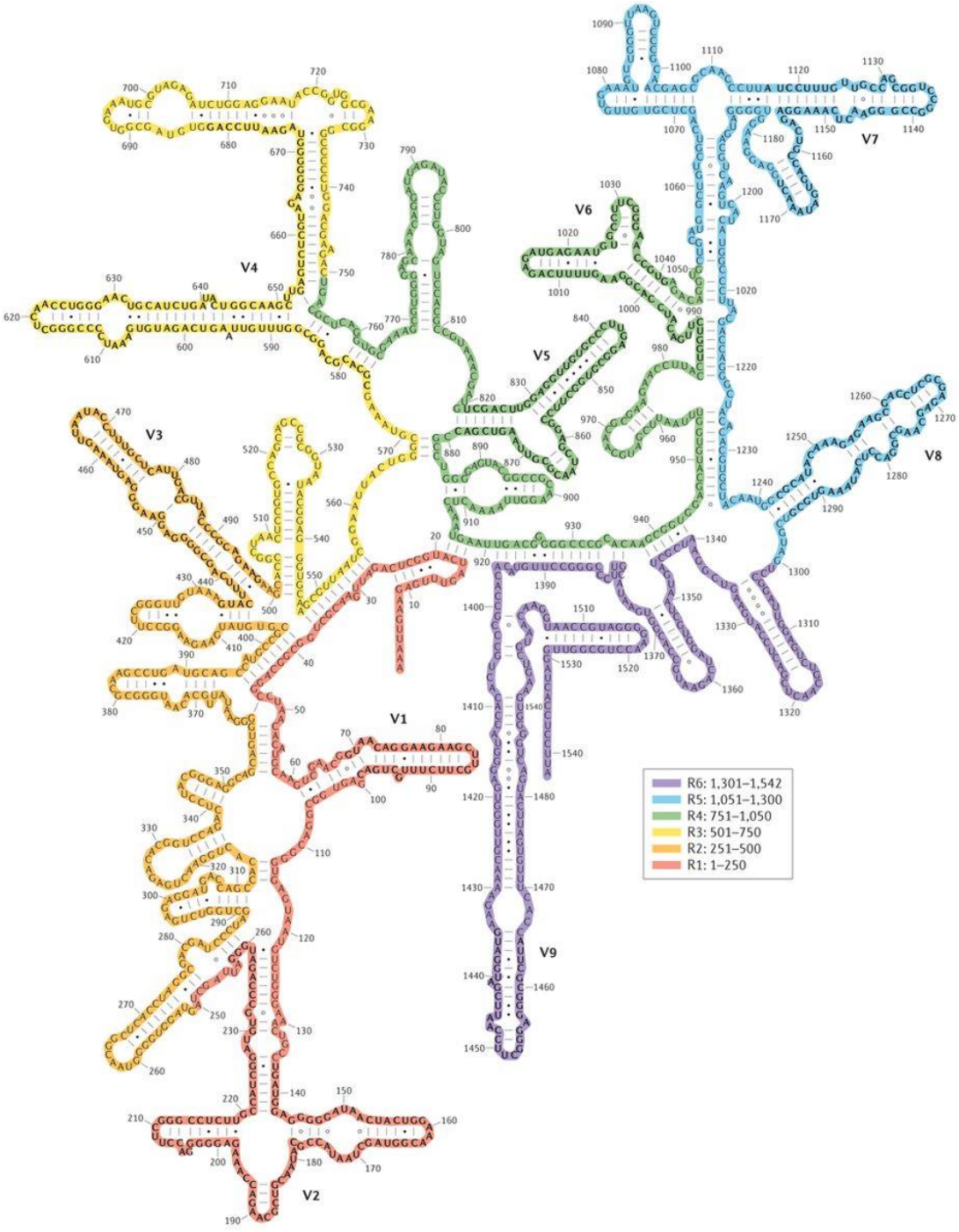


Figure 13. 16S rRNA. Secondary structure of 16S rRNA of E. coli which contains nine hypervariable regions in the six R fragments. From Yarza et al, 2014.

Linear discriminant analysis effect size

Linear discriminant analysis effect size (LEfSe) is an algorithm which is used to discover the differentially abundant taxa between two or more biological conditions (Segata et al., 2011a). It validates statistical significance, biological consistency and effect relevance using the non-parametric factorial Kruskal Wallis (KW) sum-rank test. Cladograms are used to show the taxa characterising differences between different study groups (Figure 14) and linear discriminant analysis (LDA) is used to estimate the effect size of each differentially abundant feature.

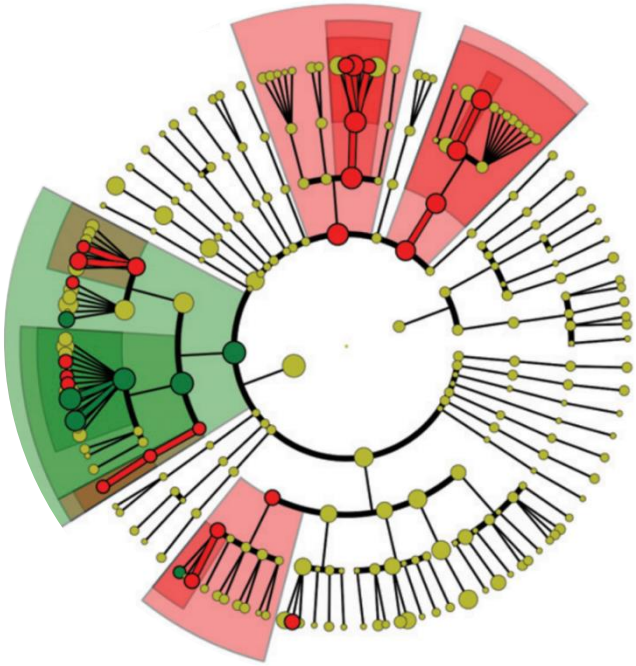


Figure 14. Cladogram. The cladograms depict the taxa (highlighted by small circles) characterising different abundance between different biological conditions. Differences are represented in the color of the most abundant class (red and green indicating two different groups, yellow indicating non-significant).

Alpha and beta diversity

Evaluation of diversity is a common approach to assess the changes in community. Diversity implies a measure of both species richness and evenness. Species richness is a measure of the total number of species in a community; evenness expresses how evenly the species in a community are distributed among each other. Alpha diversity refers to diversity within a particular community (Figure 15). Beta diversity is species diversity between different communities.

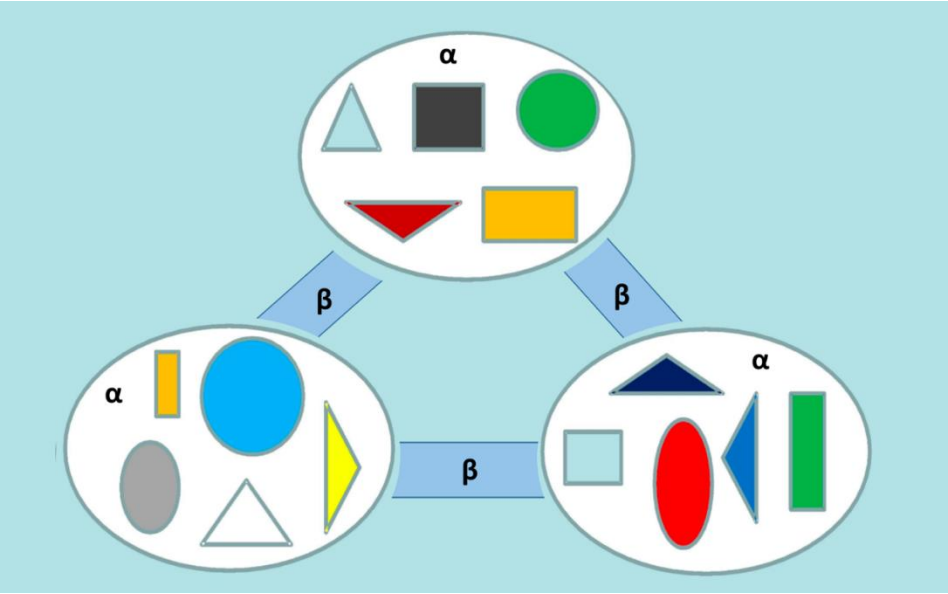


Figure 15. Alpha and beta diversity. Alpha diversity refers to a measure of similarity within communities. Beta diversity refers to a measure of similarity between communities.

Predictive functional profiling

PICRUSt (Phylogenetic Investigation of Communities by Reconstruction of Unobserved States) allows prediction of metagenome functional compositions based on the marker gene tree of 16S rRNA (Greengenes) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) Ortholog database (Figure 16) (M. G. Langille et al., 2013). The hierarchy of KEGG orthology contains four levels (Kanehisa et al., 2004). The top level (level 1) comprises five categories (carbohydrate and lipid metabolism, energy metabolism, environmental information processing, genetic information processing, nucleotide and amino acid metabolism) and the second level (level 2) comprises 24 subcategories. The third level corresponds to individual pathway maps and the fourth level corresponds to KO (KEGG Orthology) entries.

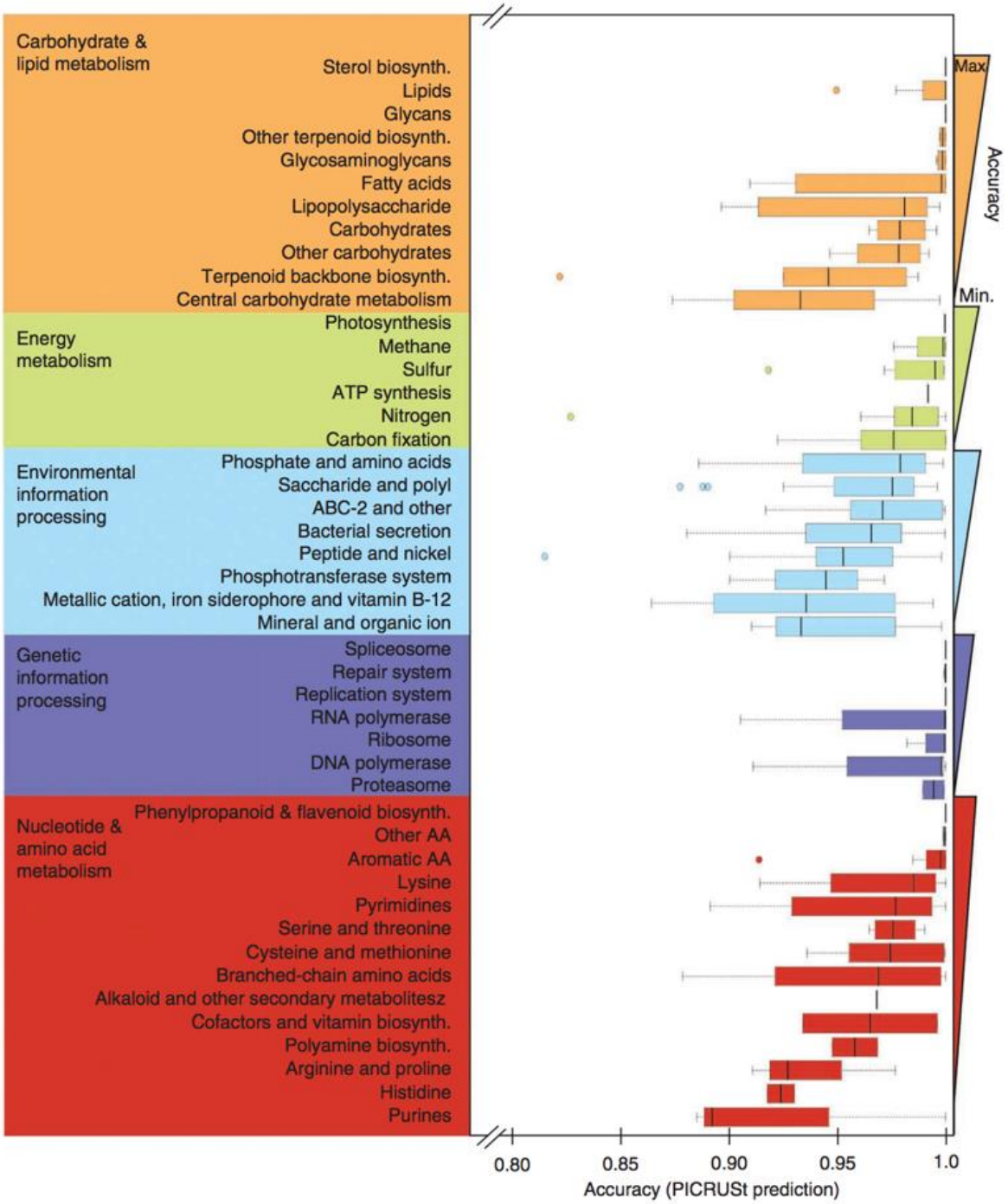


Figure 16. Metagenome prediction by PICRUSt. This computational approach is used to prediction functional composition in terms of carbohydrate and lipid metabolism, energy metabolism, environmental information processing, genetic information processing, and nucleotide and amino acid metabolism. All of them have an accuracy of more than 0.8. From Langille et al, 2013.

Aims of this study

My hypothesis is that the gut microbiota in mice fed a high fibre diet provides a sufficiently high systemic butyrate level for this to act as an endogenous mechanism of radiosensitisation. The aims of this study were:

1) to examine the impact of the diet on the microbiome before and after irradiation, and
2) subsequently to correlate diet-induced microbiome changes with tumour growth and response to radiation treatment.

Materials

Cell lines

- RT112 bladder cancer cells (Margaret Knowles, University of Leeds, UK)

Chemicals

- 98% Sodium butyrate (Sigma Aldrich)
- Phenol red-free matrigel (BD Biosciences)
- RPMI medium (Sigma Aldrich)
- Agarose, electrophoresis grade (MP Biomedicals, USA)
- V3F and V3R primers (Sigma Aldrich)

Devices

- Gulmay-320 cabinet irradiator (Xstrahl Inc, UK)
- Eppendorf concentrator 5301 (Eppendorf North America Inc, USA)
- NanoDrop One spectrophotometer (Thermo Fisher Scientific)
- Veriti PCR Thermal Cyclers (Thermo Fisher Scientific)

Diets

- see Table A1 for details (Research Diets Inc, USA)

Kits

- DNeasy PowerSoil DNA Isolation Kit (QIAGEN Ltd, UK)
- Phire Tissue Direct PCR Master Kit (Thermo Fisher Scientific)

Mouse strains

- CD1-nude mice (Charles River Laboratories, USA)

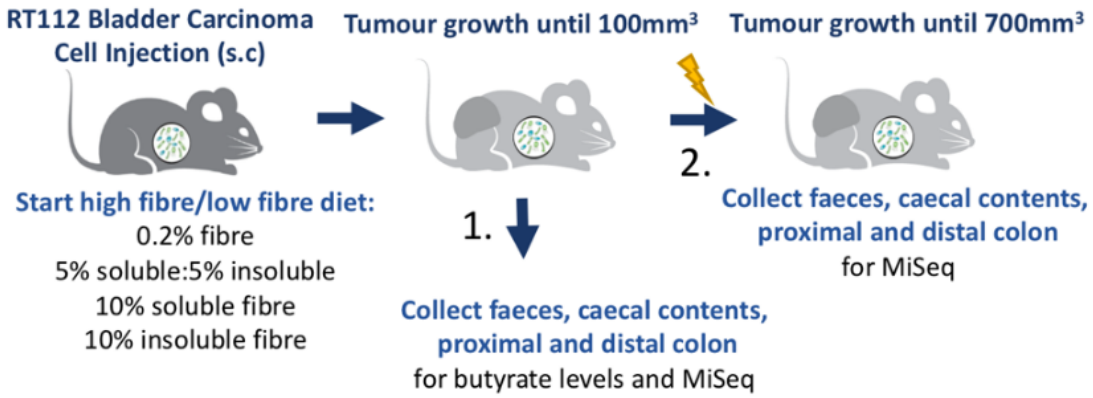
Software

- QIIME2 (Bolyen E, 2018)
- LEfSe (Segata et al., 2011b)
- PICRUSt (M. G. I. Langille et al., 2013)
- Graphpad Prism version 7.00 (GraphPad Software, USA, www.graphpad.com)

Methods

Overview of study design

Two cohorts of mice were studied. Four types of samples were collected, namely faeces, caecal contents, and tissue from proximal and distal colon. The samples from the first cohort were collected when the tumours had grown to 100 mm^3, and consisted of 15 mice in five groups (n = 3 per group) fed diets comprising: low fibre (< 0.2% cellulose), low fibre with butyrate, high mixed fibre (5% cellulose, 5% inulin), high insoluble fibre (10% cellulose), high soluble fibre (10% inulin) (Figure 17, Table 1, Table A1). In low fibre with butyrate group, drinking water was added with butyrate (Sigma Aldrich) at a final concentration of 300 mM. The samples from the second cohort were collected when the tumours had reached 700 mm^3 either without (n = 3) or following (n = 8) radiation, and consisted of four groups fed diets comprising: low fibre, high soluble fibre, high insoluble fibre and high mixed fibre. 16S rRNA gene sequencing was performed on a MiSeq platform to investigate the bacterial components of the mouse gut microbiome. Metagenomic analysis was conducted using a QIIME2 platform



to identify the features (differentially abundant taxa or diversity index) that correlated to the phenotype.

Figure 17. Schema of sample collection and analyses. Two cohorts of gut microbiome were analysed, with the faecal and caecal content samples collected when tumours reached 100 mm^3 and 700 mm^3 respectively.

Table 1. Overview of the study and sample sizes

	Sample collected when tumour growth until	Group	Radiation	Sample size
1.	100mm^3	Low fibre (0.2% cellulose)		3
		Low fibre with butyrate (0.2% cellulose)		3
		High mixed fibre (5% inulin, 5% cellulose)		3
		High insoluble fibre (10% cellulose)		3
		High soluble fibre (10% inulin)		3
2.	700mm^3	Low fibre (0.2% cellulose)		3
		High mixed fibre (5% inulin, 5% cellulose)		3
		High insoluble fibre (10% cellulose)		3
		High soluble fibre (10% inulin)		3
		Low fibre (0.2% cellulose)	+	8
		High mixed fibre (5% inulin, 5% cellulose)	+	8
		High insoluble fibre (10% cellulose)	+	8
		High soluble fibre (10% inulin)	+	8

Mice and mouse diets

CD1-nude female mice were housed in a temperature-controlled environment with a 12-h reversed-phase light/dark cycle (lights on 21:00 h) and provided with food and water ad libitum. At 7 to 8 weeks of age, mice were injected subcutaneously with RT112 bladder cancer cells and started receiving either low fibre diets (2 g cellulose/3850 kcal) or high insoluble fibre diets (100 g cellulose/3850 kcal) or high soluble fibre diets (100 g inulin/3850 kcal) or high mixed fibre diets (50 g cellulose + 50 g inulin / 3850 kcal) for a maximum time of 9 weeks or until they were culled when the tumours reached 700 mm^3. Faeces, caecal contents, and proximal and distal colons from the first cohort were taken when the tumour reached 100 mm^3 (each group n = 3) prior to the irradiation to investigate the microbiome at baseline. Faeces, caecal contents, and proximal and distal colons from the second cohort were taken when the tumour reached 700 mm^3 after ionising radiation (each group n = 8) or without irradiation (each group n = 3) or at the end of study (9 weeks after xenograft) to study the association between the gut microbiome composition and tumour response.

Xenograft model and irradiation method

All *in vivo* studies were performed by Dr. Salome Paillas. Mice were injected subcutaneously under anaesthesia with 5 x 10^6 human bladder cancer cells (RT112) with phenol red-free Matrigel (BD Biosciences) and RPMI medium at a total volume of 100 ul (1:1) in the right flank. Tumour growth was measured three times a week using calipers using (a x b x c x π/6). To assess the tumour effects of different dietary fibre to irradiation *in vivo*, mice received ionising radiation to the tumour (6 Gy, single fraction, 300 kV, using a Gulmay-320 cabinet irradiator).

Microbiome sample collection and DNA extraction

All samples were transported on ice and kept at -20°C for less than 2 hours before DNA extraction. Bacterial genomic DNA was extracted using a DNeasy PowerSoil DNA Isolation Kit (QIAGEN Ltd, Manchester, UK), as per the Human Microbiome Project (Human Microbiome Project, 2012a). Briefly, by adding sodium dodecyl sulfate (SDS), microbial cells were lysed by mechanical disruption with a ceramic bead set on 3,000 rpm for 10 minutes, followed by binding of DNA tightly to a silica membrane in a Spin Filter device at high salt concentrations. Eventually, DNA was collected into sterile elution buffer and quantified using a NanoDrop

spectrophotometer. All DNA samples were kept at -80°C. Those sent for sequencing to Omega Bioservices (Georgia, USA) were dried in the eppendorf concentrator at a rotational speed of 1,400 rpm and centrifugal force of 240 x g for 1 hour at 30°C.

Identification and quantification of bacterial DNA

I conducted quantification of the microbiota of the contents of the intestinal tracts and the intestinal wall of the proximal and distal colon (tissue) by PCR of 16s rRNA. This was performed on genomic DNA extracted as described above. The PCR was performed using primers - V3F (CCAGACTCCTACGGGAGGCAG) and V3R (CGTATTACCGCGGCTGCTG) (H. Y. Wang et al., 2014). All primers were purchased from Sigma. For each sample, Phire Tissue Direct PCR Master Mix (Thermo Fisher Scientific) was used to amplify the 16S rRNA gene hypervariable V3 region (product size = 200 bp). PCR amplifications were performed using the following conditions: 98°C for 5 minutes followed by 35 cycles at 98°C for 5 seconds each, 66.3°C for 5 seconds, and 72°C for 30 seconds and a final extension step at 72°C for 1 minute. The amplification products were visualised on a 1% agarose gel after electrophoretic migration of 5 µl of amplified material. I created a standard curve from serial dilutions of Escherichia Coli from 1 x 10^2, 1 x 10^4, 1 x10^6 colony-forming units (CFU) which was quantified by CFU assay. All samples were run in duplicate.

Bacterial 16S rRNA sequencing

16S rRNA gene sequencing methods were adapted from the methods developed for the NIH-Human Microbiome Project (Human Microbiome Project, 2012a, 2012b). The amplification and sequencing of 16S rRNA gene V3V4 region were done commercially by Omega Bioservices (Georgia, USA) on a MiSeq platform (Illumina, Inc, San Diego, CA) using the 2x300 bp paired-end protocol, yielding paired-end reads with near-complete overlap. The primers containing adapters for Miseq sequencing were used for amplification and single-end barcodes, allowing pooling and direct sequencing of PCR products (Caporaso et al., 2012). PBS negative controls were included to eliminate the confounding effects of environmental contamination.

All 16S rRNA gene-based metagenomic analysis was conducted by myself using a QIIME2 platform (Hall & Beiko, 2018). I clustered quality filtered sequences with >97% identity into bins known as Operational Taxonomic Units (OTUs), using open-reference OTU picking. I

obtained the relative abundance of each OTU from all samples. In the taxonomic analysis, I classified the microbiome at the phylum, class, order, family, genus and species levels by referring to the Greengenes database (Balvociute & Huson, 2017).

The analysis pipeline is illustrated below (Figure 18):

i. All sequences were trimmed to a length of 240, since the quality dropped at this length based on the sequence quality plots.
ii. De-noised sequencing errors by using the "Deblur" plugin in QIIME2 (Nearing et al., 2018).
iii. Taxonomic assignment was performed with Greengenes (DeSantis et al., 2006) by the "feature-classifier" command.
iv. To visualise the differences in microbial composition between gut contents and tissue, a taxonomic profile was generated by conducting differential abundance analysis using balances (balances are useful tools to investigate the relationship between proportions) in gneiss (a toolkit to study phylogenetic data).
v. To identify the features characterising the differences between different groups, the LEfSe method of analysis was performed to compare abundances of all bacterial clades (Segata et al., 2011a). By validation using the Kruskal-Wallis test at the α setting of 0.05, effect size was obtained by LDA based on the significantly different vectors resulting from the comparison of abundances between groups.
vi. To validate the significance of enrichment of bacterial taxa among different groups, discrete false-discovery rates (DS-FDR) were calculated (Jiang et al., 2017)
vii. Generated phylogenetic tree by using the "phylogeny" plugin in QIIME2
viii. To investigate the alpha and beta diversity, the diversity commands of "alpha-group-significance" and "beta-group-significance" were used to obtain Peilous's evenness, Faith's phylogenetic diversity, Shannon's index, and Bray-Curtis dissimilarity. A principal coordinates (PCoA) plot was obtained by using the Emperor Tool based on the results of Bray-Curtis dissimilarities.
ix. The OTU table was rarefied using the "alpha-rarefraction" command in QIIME2. The alpha rarefraction plot showed the richness of the samples with increasing sequence count.

x. To predict the metagenome functional profiles, PICRUSt, a bioinformatics software package, was used to collapse predicted functions (KEGG Orthology; KO) based on 16s rRNA surveys into higher categories (KEGG pathway) after picking OTUs and normalisation (M. G. Langille et al., 2013).

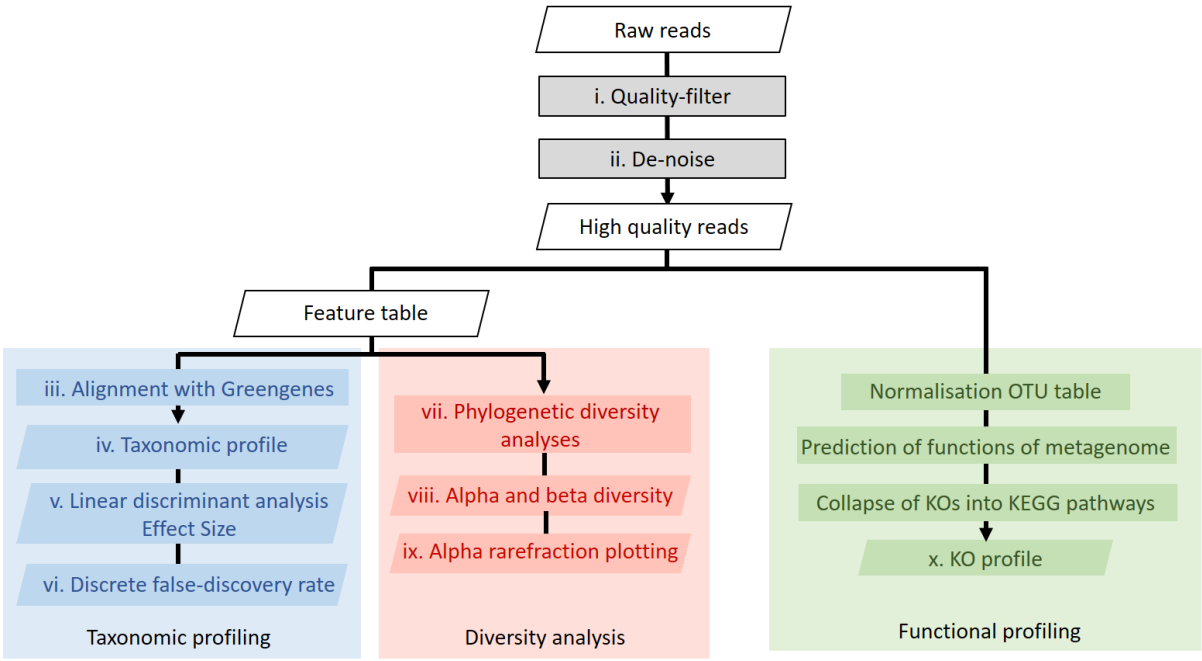


Figure 18. Flow chart of metagenomic analysis pipeline. Raw reads were first proceeded through a quality-filter and further de-noised to generate high quality reads and they were analysed further to investigate taxonomic profiles, diversity and functional profiles. Roman numbers are used to match the analysis pipeline text and Figure 18.

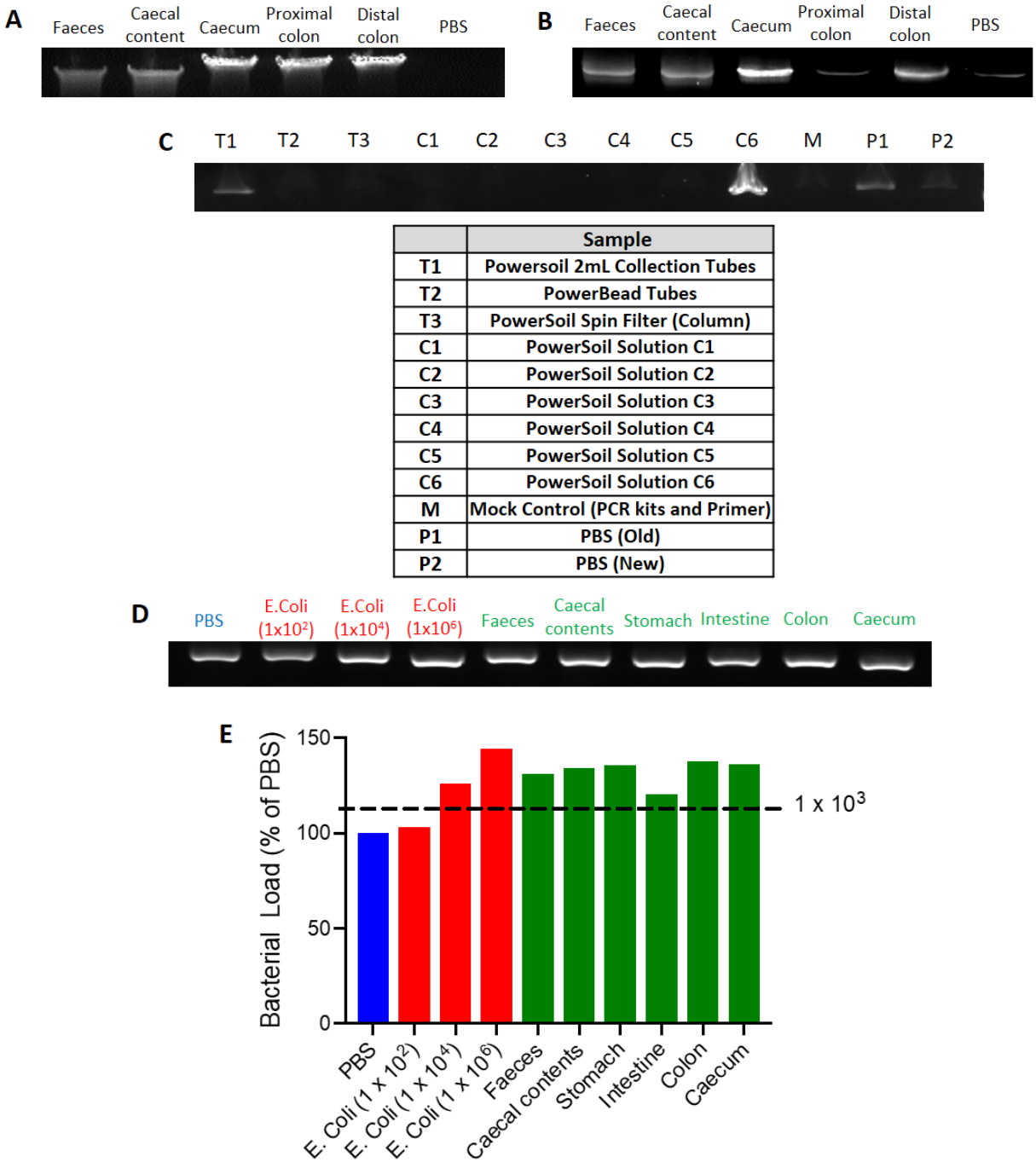
Statistical analysis

Alpha diversity was compared using the Mann-Whitney test. All mice were classified into high, intermediate or low diversity groups based on tertiles of distribution. Time to quadruple in volume was defined as the interval (in days) from the date of irradiation (growth to 100 mm^3) to the date of tumour growth to 400mm^3. Tumour growth curves were analysed for each group, and their slopes were compared using ANOVA. The LEfSe method of analysis was applied to determine the difference in faecal taxa, using the Kruskal-Wallis test. Significantly different taxa presented from the previous comparison were used as input for LDA, which produced an LDA score. Volcano plots showed the significance of the taxa which are different among different groups, with $\log_{10}$(FDR-adjusted p-values) on the y-axis and median-adjusted effect sizes on the x-axis. In addition, mice were also classified as having high or low abundance of *B. acidifaciens* based on the median relative abundance of this taxon in the gut microbiome sample. All analyses were conducted in QIIME2 and GraphPad Prism version 8.0 (La Jolla, CA).

Results

The issue of contamination

Genomic DNA was successfully extracted using a QIAGEN Powersoil kit and its presence was demonstrated by gel electrophoresis of processed samples from faeces, caecal contents, and the proximal and distal colon (n = 1 mouse; Figure 19A). To identify bacterial DNA, amplification of 16s rRNA and gel electrophoresis were performed. Bacterial DNA was found in all samples, including the PBS control (Figure 19B). To elucidate the issue of contamination from the DNA extraction kit or other laboratory reagents, all containers and reagents were tested and bacterial DNA was found in the collection tube, elution buffer and PBS (Figure 19C). By quantification of bacterial loads, I determined that our mouse samples contained more than 10^4 bacterial colony forming units (CFUs) which appeared to override contaminating species in the sample microbial communities (Figure 19D & 19E).

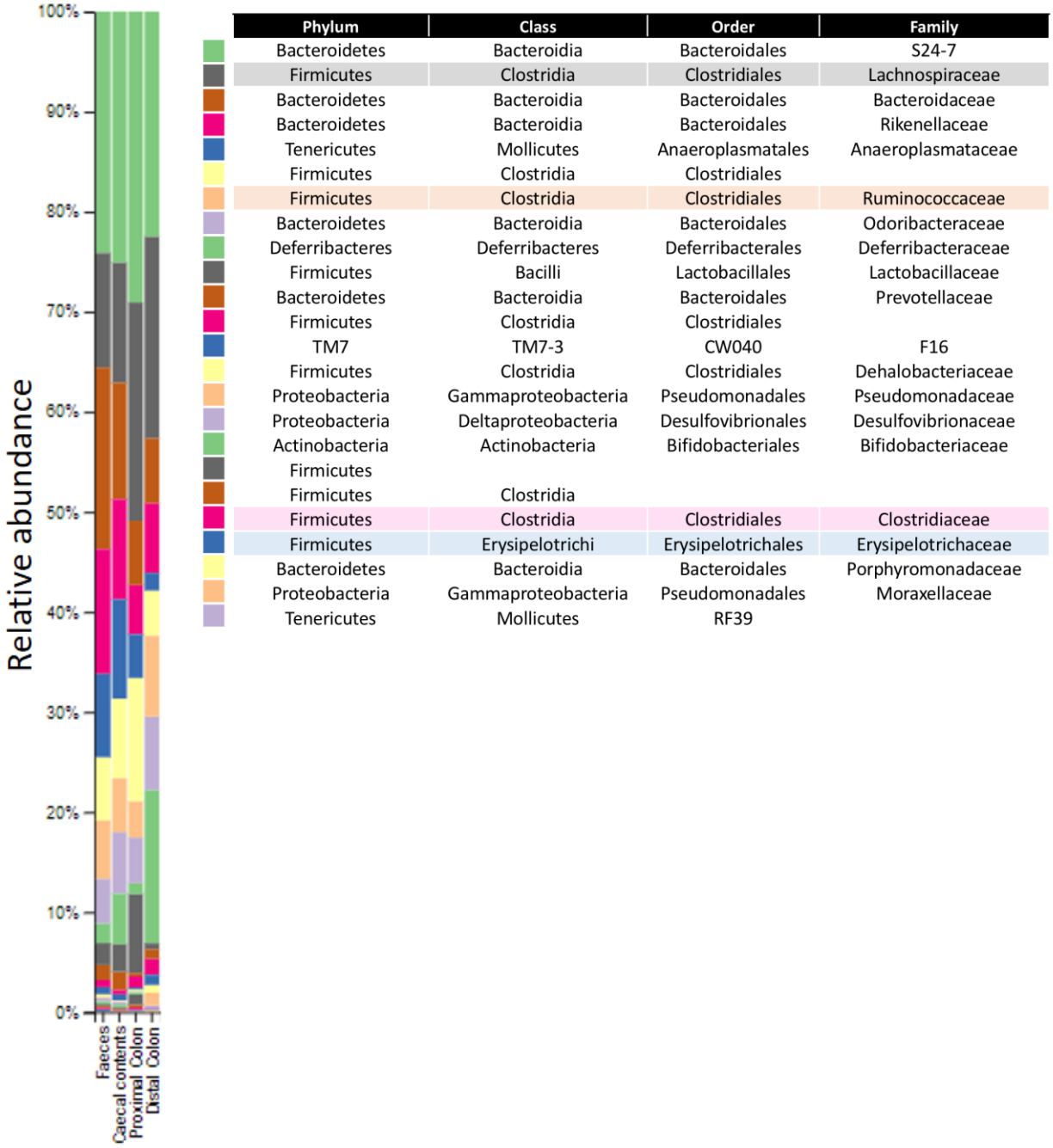


	Sample
T1	Powersoil 2mL Collection Tubes
T2	PowerBead Tubes
T3	PowerSoil Spin Filter (Column)
C1	PowerSoil Solution C1
C2	PowerSoil Solution C2
C3	PowerSoil Solution C3
C4	PowerSoil Solution C4
C5	PowerSoil Solution C5
C6	PowerSoil Solution C6
M	Mock Control (PCR kits and Primer)
P1	PBS (Old)
P2	PBS (New)

Figure 19. The issue of contamination and quantification of DNA load from samples. (A) Gel electrophoresis of whole genomic DNA from different samples (n = 1 mouse). (B) Gel electrophoresis of bacterial DNA after amplification of 16s rRNA (n = 1 mouse). (C) Determining the presence of bacterial DNA contamination for all components of QIAGEN Powersoil DNA extraction kits. (D, E) Quantification of bacterial load from different tissue and luminal contents from mice, with E. Coli (1×10^2, 1×10^4, 1×10^6 CFUs) as controls (n =1 mouse).

Validation of sequencing and analysis pipeline using luminal content and tissue samples

Bacteroidetes and *Firmicutes* were the most dominant phyla of bacteria, and the *Lachnospiraceae*, *Ruminococcaceae*, *Clostridiaceae* and *Erysipelotrichaceae* families were the butyrate-producing colon bacterial families present (Figure 20). No significant difference in alpha diversity was observed between luminal content and tissue samples (n = 2 per group) (Figure 21A-C), but a slight clustering effect (bacterial components within groups were close to each other) in terms of beta diversity was found between these two groups, indicating an increased similarity in the bacterial components within these groups, compared to that between groups (Figure 21D). *Clostridiaceae* and *Candidatus arthromitus* were enriched in the luminal contents, and *Bacteroidetes*, *Bacteroidia*, and *Bacteroidales* were enriched in tissues, indicating the gut microbiota from luminal contents and tissues were different (Figure 22). All bacteria from the contents and tissues were predicted to be mainly involved in carbohydrate metabolism, membrane transport and amino acid metabolism (Figure 23).



Phylum	Class	Order	Family
Bacteroidetes	Bacteroidia	Bacteroidales	S24-7
Firmicutes	Clostridia	Clostridiales	Lachnospiraceae
Bacteroidetes	Bacteroidia	Bacteroidales	Bacteroidaceae
Bacteroidetes	Bacteroidia	Bacteroidales	Rikenellaceae
Tenericutes	Mollicutes	Anaeroplasmatales	Anaeroplasmataceae
Firmicutes	Clostridia	Clostridiales	
Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
Bacteroidetes	Bacteroidia	Bacteroidales	Odoribacteraceae
Deferribacteres	Deferribacteres	Deferribacterales	Deferribacteraceae
Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae
Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae
Firmicutes	Clostridia	Clostridiales	
TM7	TM7-3	CW040	F16
Firmicutes	Clostridia	Clostridiales	Dehalobacteriaceae
Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae
Proteobacteria	Deltaproteobacteria	Desulfovibrionales	Desulfovibrionaceae
Actinobacteria	Actinobacteria	Bifidobacteriales	Bifidobacteriaceae
Firmicutes			
Firmicutes	Clostridia		
Firmicutes	Clostridia	Clostridiales	Clostridiaceae
Firmicutes	Erysipelotrichi	Erysipelotrichales	Erysipelotrichaceae
Bacteroidetes	Bacteroidia	Bacteroidales	Porphyromonadaceae
Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae
Tenericutes	Mollicutes	RF39	

Figure 20. Stack bar plot of phylogenetic composition of luminal content and tissue samples. Common bacterial taxa at the family in luminal contents (faeces and caecal contents) and tissue samples (proximal and distal colon) by 16s rRNA sequencing. Butyrate-producing colonic bacteria are highlighted in the table (n = 1 mouse).

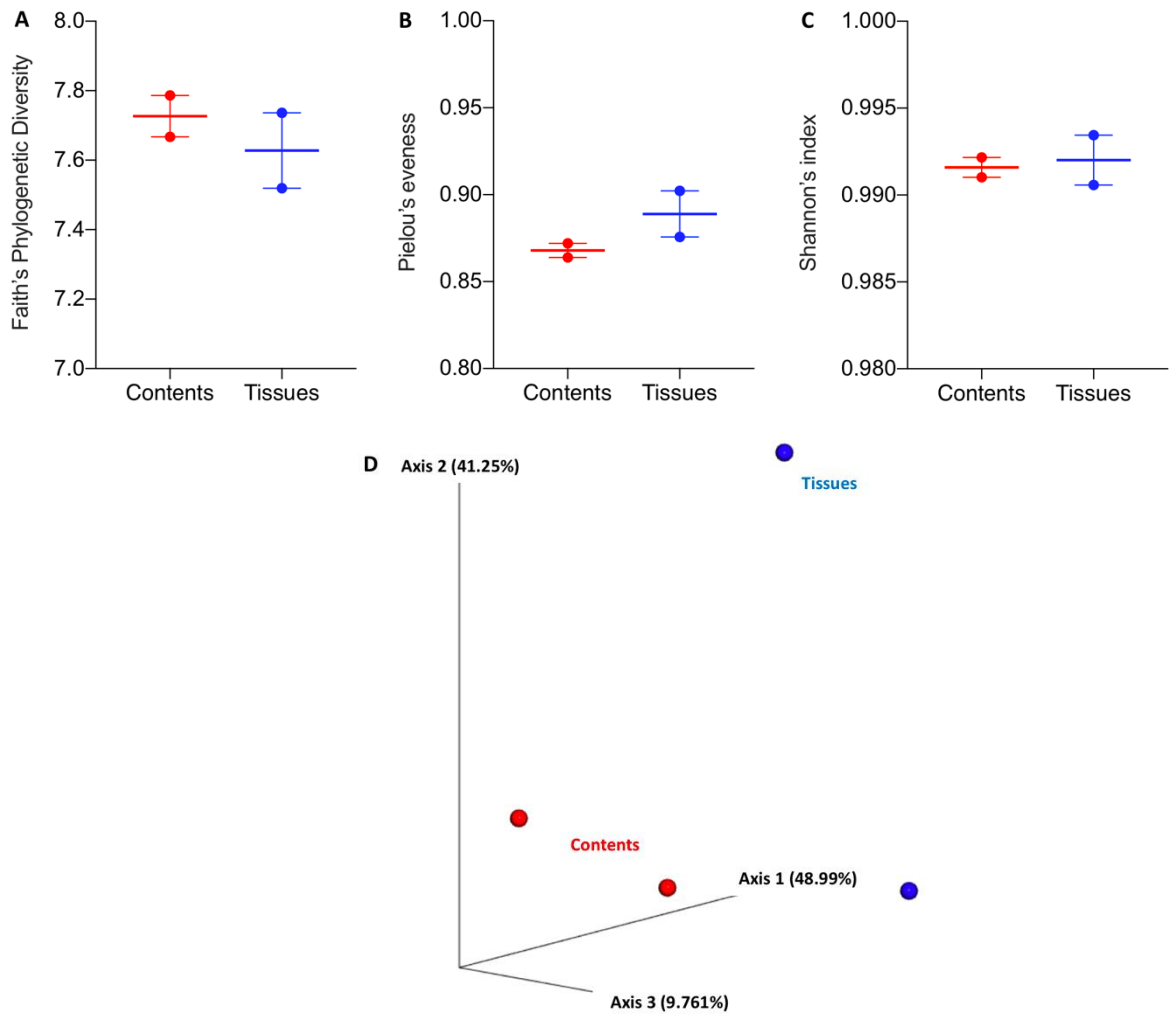


Figure 21. Alpha and beta diversity of luminal content and tissue samples. (A-C) Faith's phylogenetic diversity, Pielou's evenness and Shannon's index of gut microbiome in luminal contents and tissue samples by Mann-Whitney U rank sum (MW) test. Error bars represent the interquartile range of diversity scores. (D) Principal coordinate analysis of luminal contents and tissue sample using Bray-Curtis dissimilarity. Dots of luminal contents were labeled in red and dots of tissue samples were labeled in blue.

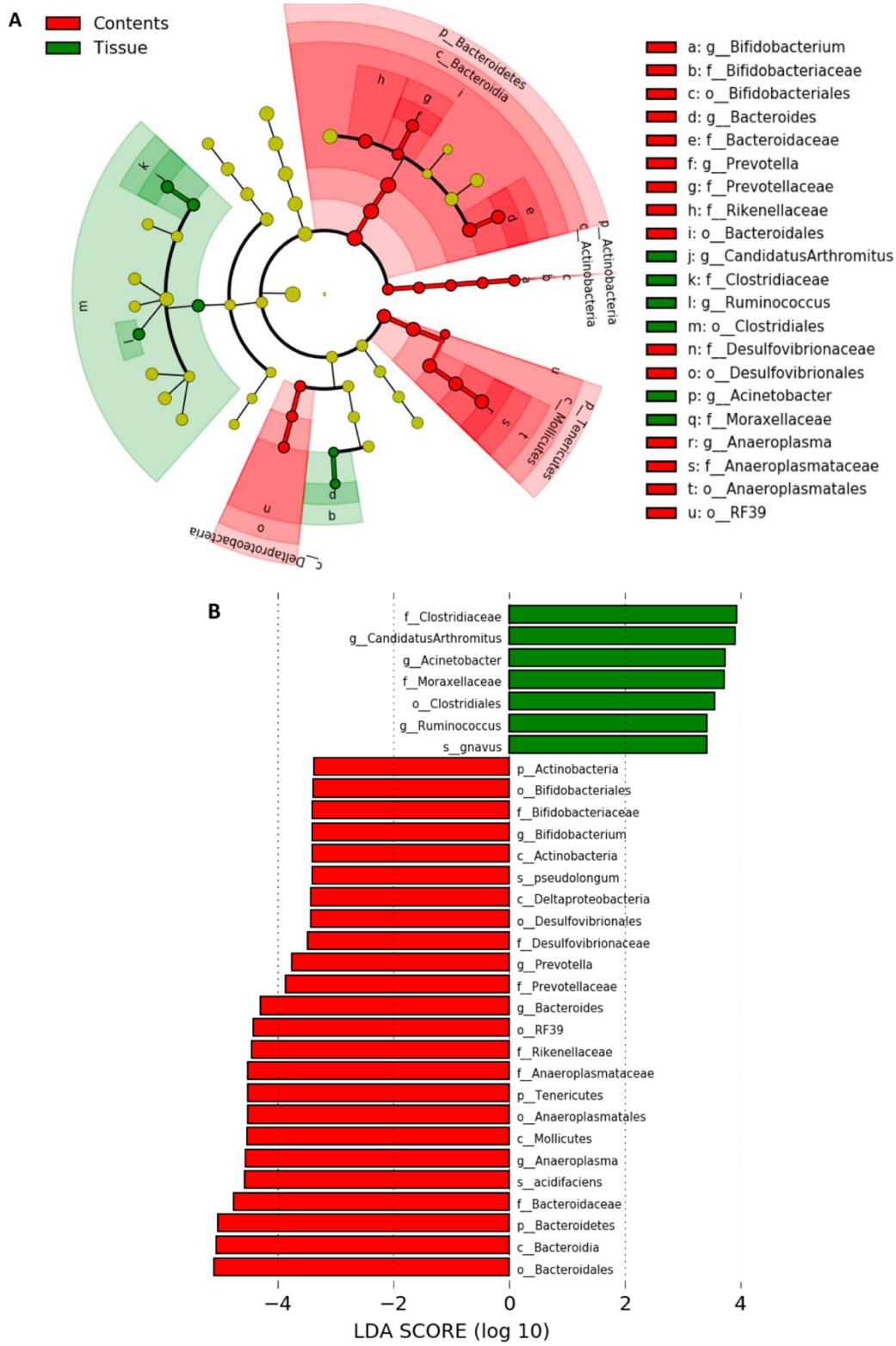


Figure 22. Differences in bacterial components of luminal content and tissue samples. (A) Taxonomic cladogram from LEfSe showing differences of taxa in luminal contents and tissue samples at order (o), family (f) and genus (g) levels. Dot size is proportional to the abundance of the taxon. (B) LDA scores computed for differentially abundant taxa in the microbiomes of luminal contents (red) and tissue (green) at phylum (p), class (c), order (o), family (f), genus (g) and species (s) levels. Length indicates effect size associated with a taxon and p = 0.5 for Kruskal-Wallis test. Letter in front of taxa denotes taxonomic level.

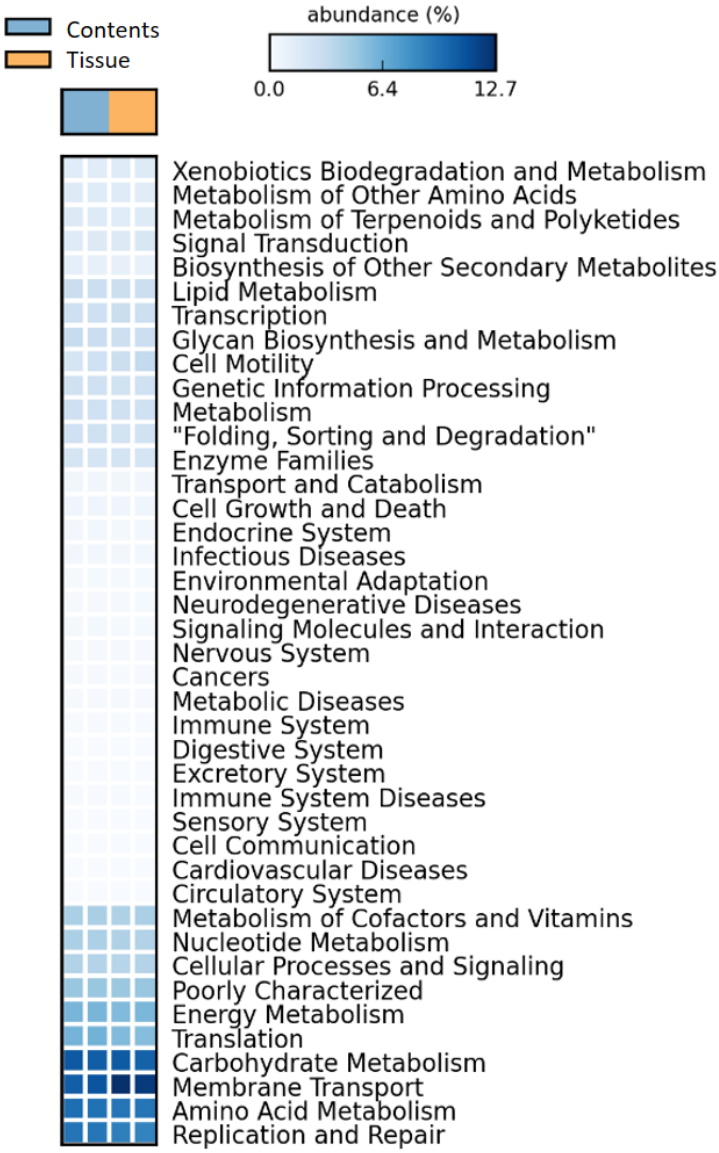


Figure 23. Predictive functional metagenome profile of luminal content and tissue samples. Heatmap of metabolic pathways obtained from PICRUSt analysis depicting the relative expression of the indicated genes.

Environmental microbiome

All DNA samples were classified by the date of collection, to observe the influence of this factor on the gut microbiome. Principal coordinate analysis revealed no cluster effect within samples collected on the same day, and this result proved that there was minimal contamination of the gut microbiome (Figure 24). The PBS control was processed identically to the luminal content and tissue samples from the start of the DNA extraction. The community microbiome in the negative control was very different from the gut microbiome of the mice (Figure 25). The sequence counts and nucleic acid amounts in the PBS controls were extremely low, compared to than that in the faecal microbiota (Figure 26). Therefore, the environmental microbiota had minimal impact on the analysis of the gut microbiota of interest in this study. The quality scores of sequencing in this study were high and the read depth was sufficient to enable a full analysis of the gut microbiota (Figure 27-30).

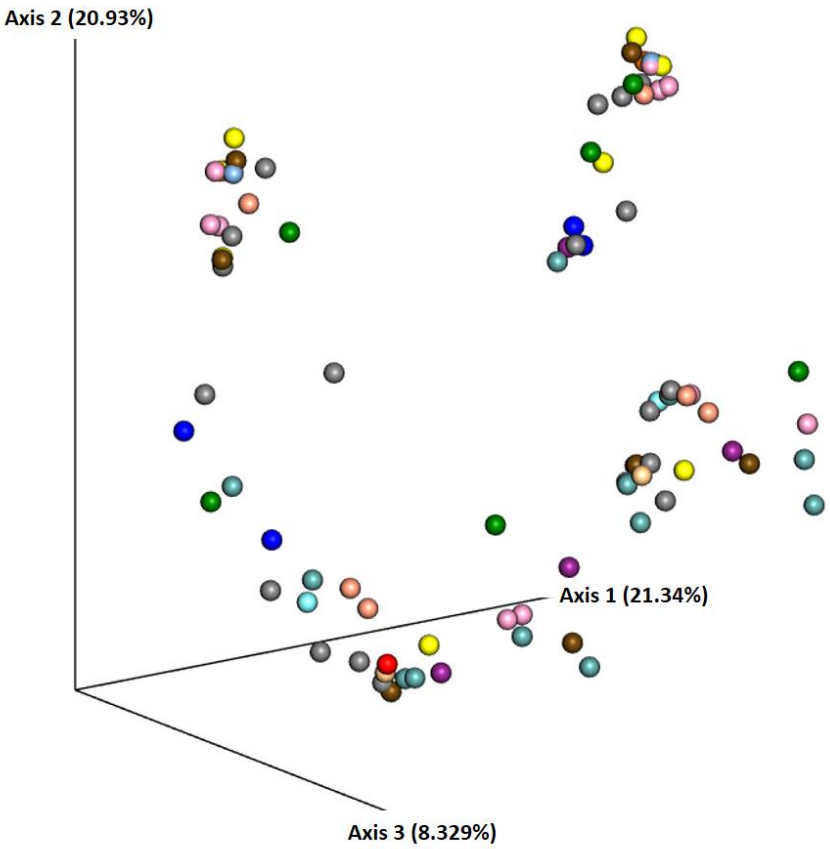


Figure 24. Beta diversity of the gut microbiome of samples collected from different dates. Principal coordinate analysis of faecal and caecal microbiomes using Bray-Curtis dissimilarity (Each colour denotes a specific date of sample collection).

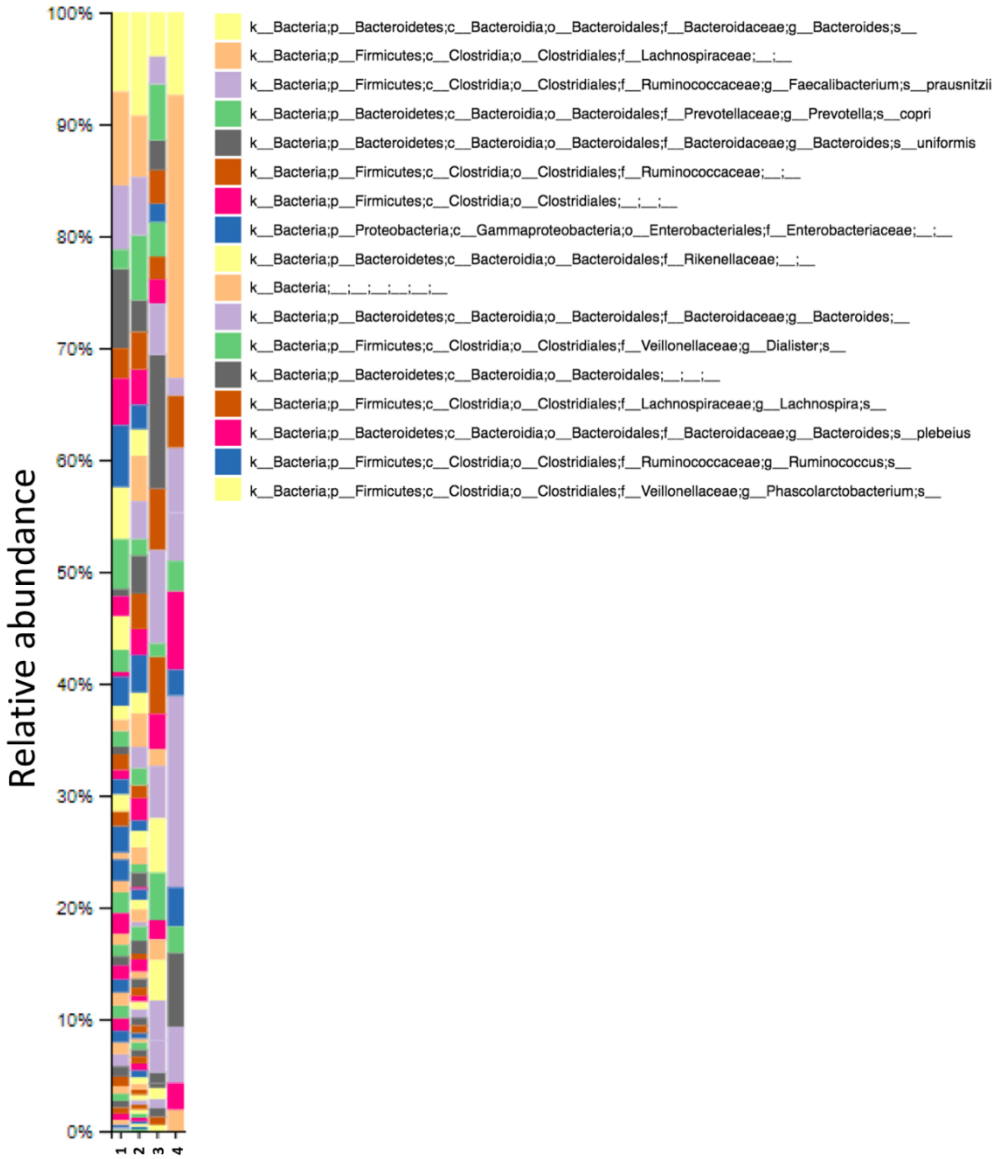


Figure 25. Phylogenetic composition of the environmental microbiome. Common bacterial taxa at the species level in 4 samples of PBS as negative controls of DNA extraction by 16s rRNA sequencing.

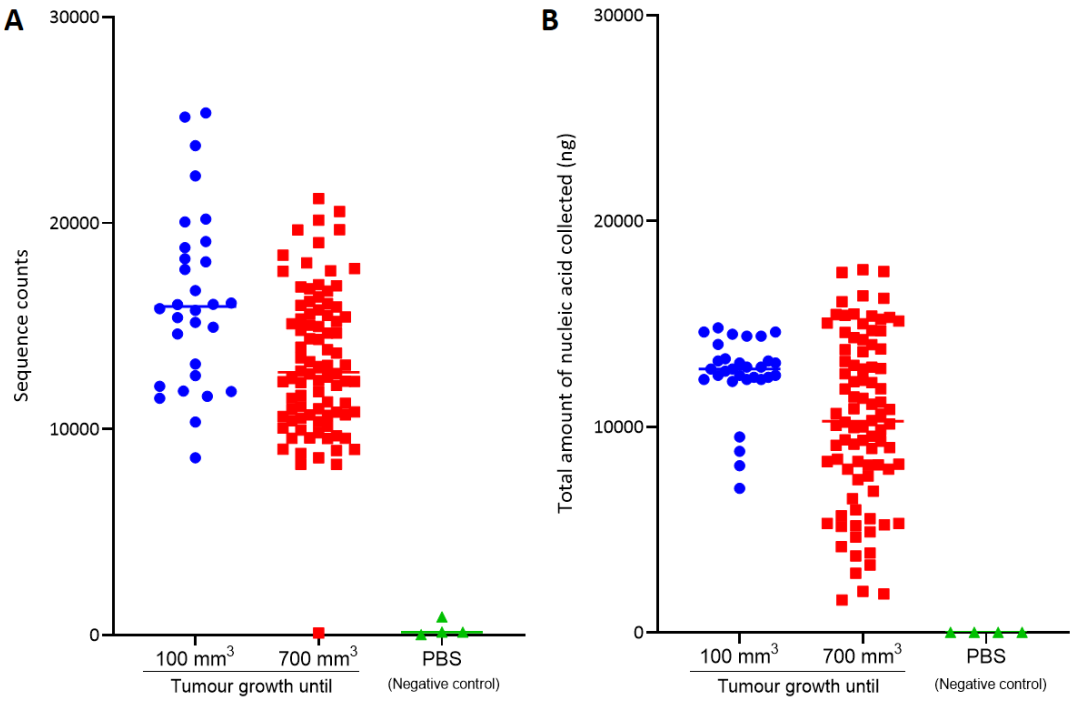


Figure 26. Sequence counts and total amount of nucleic acids in all samples. Comparison of sequence counts on the Miseq platform and total amount of nucleic acids quantified by PicoGreen assay in all samples which were collected when the tumours reached 100 mm^3 and 700 mm^3.

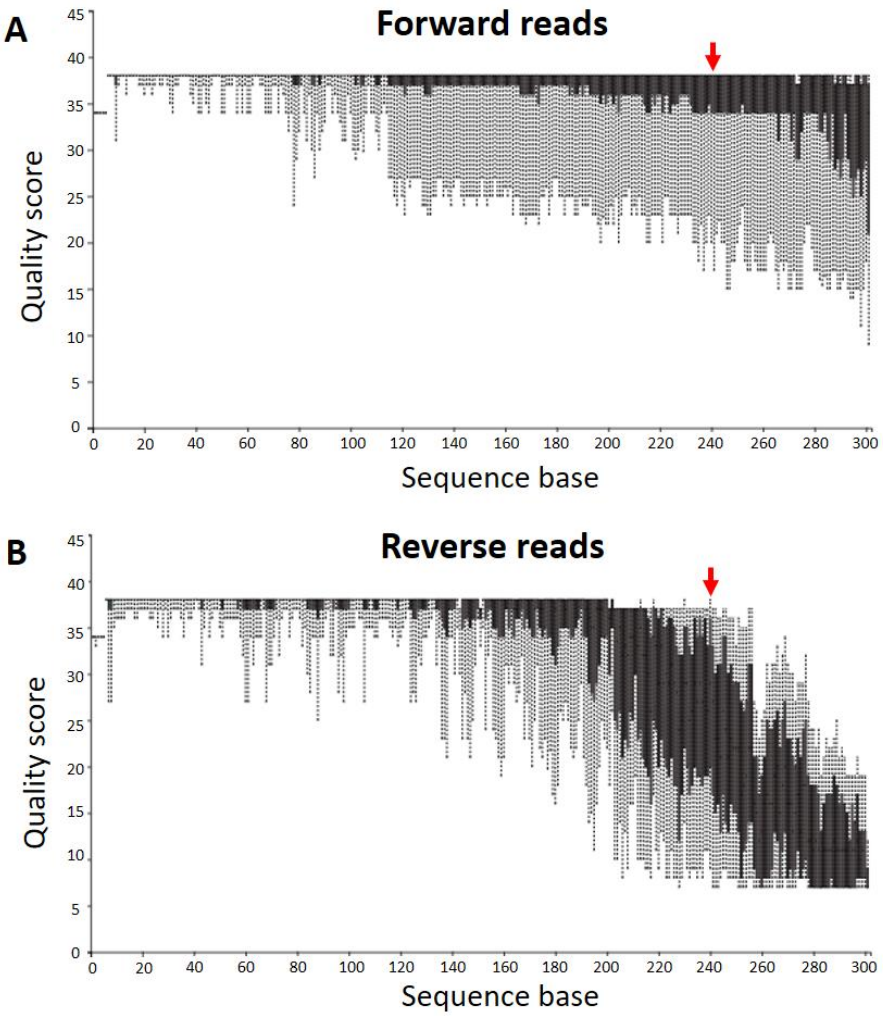


Figure 27. Sequencing quality scores of samples when the tumours reached 100 mm^3. Truncation was performed at a length of 240 because the quality score dropped off at that point.

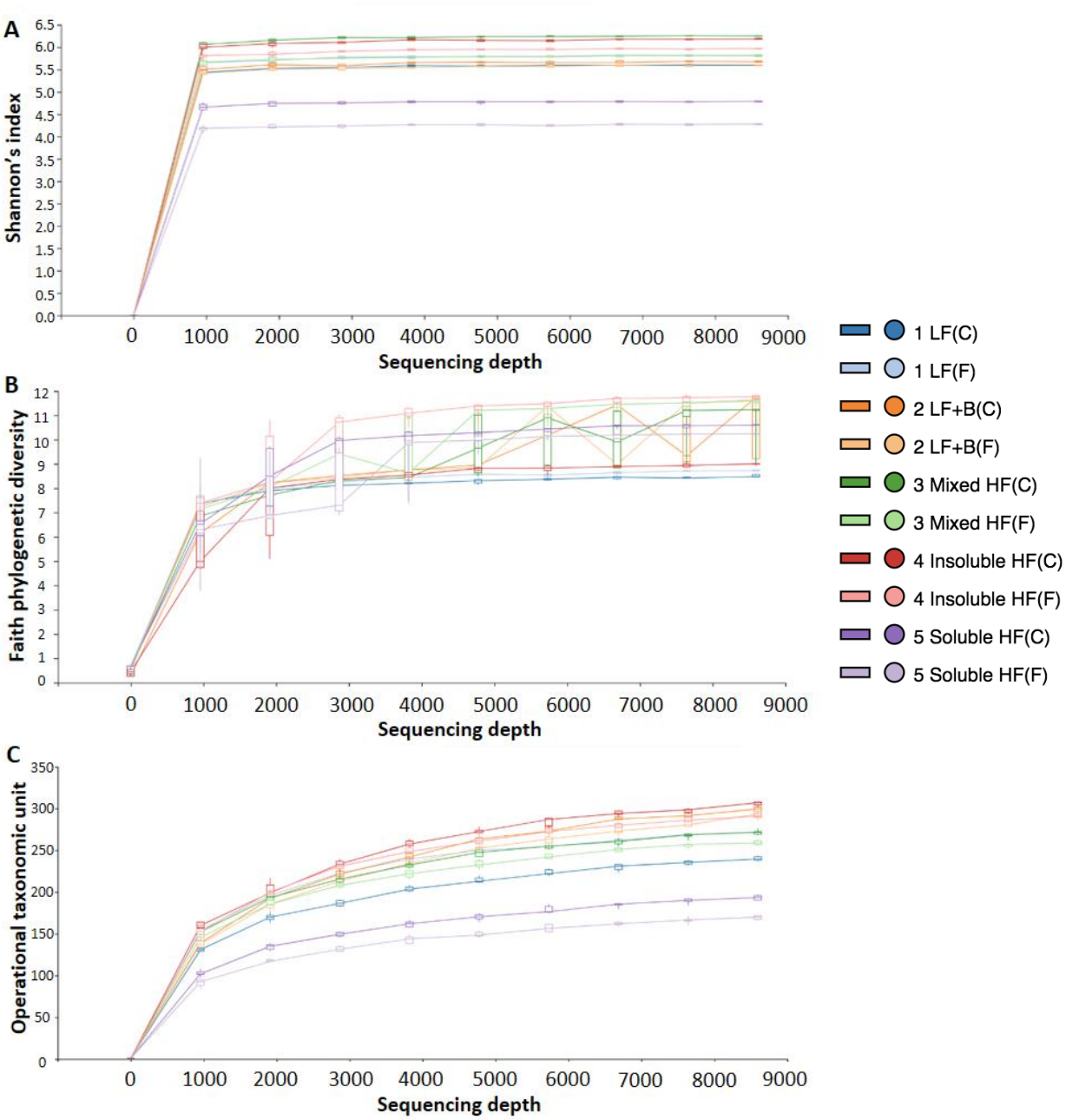


Figure 28. Alpha rarefraction plotting of samples collected when the tumours reached 100 mm^3. (A) Shannon's index, (B) Faith's phylogenetic diversity, and (C) operational taxonomic units (OTUs = species) were calculated at consecutive sequencing depth. In the figure legends, 'C' denotes caecal microbiome and 'F' denotes faecal microbiome.

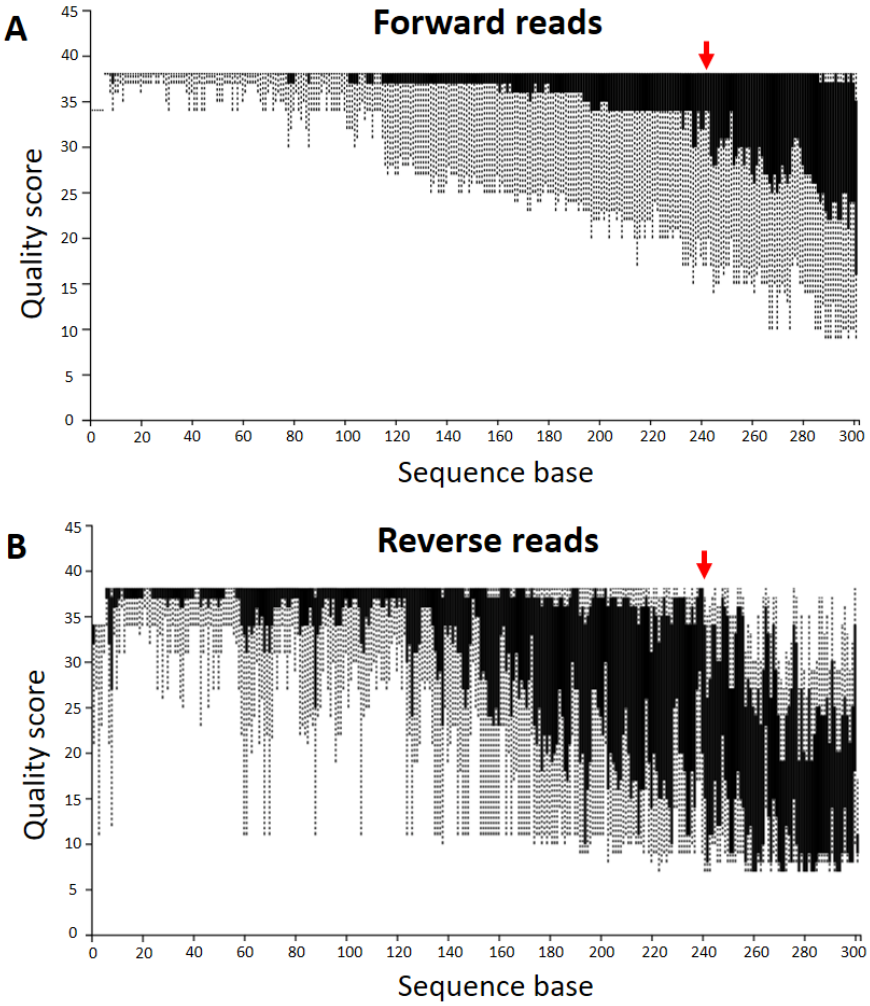


Figure 29. Sequencing quality scores of samples collected when the tumours reached 700 mm^3. Truncation was performed at a length of 240 because the quality score dropped off at that point.

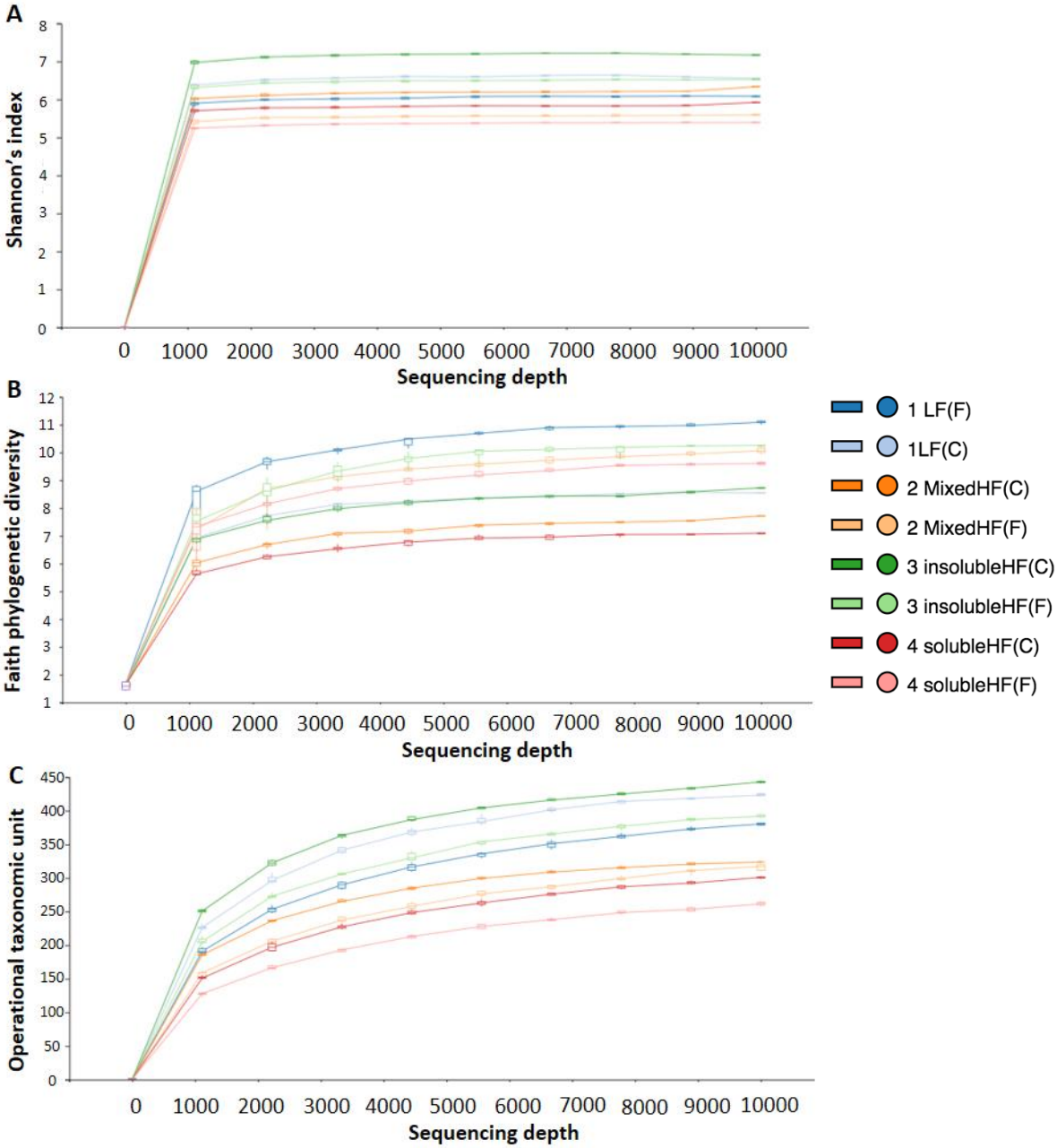


Figure 30. Alpha rarefraction plotting of sequencing of samples collected when the tumours reached 700 mm^3. (A) Shannon's index, (B) Faith's phylogenetic diversity, and (C) operational taxonomic units (OTUs = species) were calculated at consecutive sequencing depth.

Similar bacterial components in the faecal and caecal microbiomes

By investigation of the faecal and caecal microbiomes in six to ten-week-old C57BL/6 mice, Pang et al. revealed that both profiles were not correlated with each other (Pang et al., 2012). I compared faecal and caecal content profiles, using the samples collected when the tumours reached 100 mm^3 and 700 mm^3. From the gut microbiota of faeces and caecal contents collected when tumours grew until they reached 100 mm^3, both profiles shared the same four major populations. These four taxa namely, *Bacteroidales*, *Clostridiales*, *Verrucomicrobiales* and *Erysipelotrichales,* had more than 80% relative abundance of bacteria at the order level (Figure 31). From the samples collected when the tumours had grown to 700 mm^3, three bacteria orders, namely, *Bacteroidales*, *Clostridiales*, and *Verrucomicrobiales,* occupied more than 90% of the relative abundance of the microbiome (Figure 32). Faecal and caecal microbiota were shown to have similar bacterial components. However, the microbiota were different when comparing mice with 100 mm^3 tumours and mice with 700 mm^3 tumours.

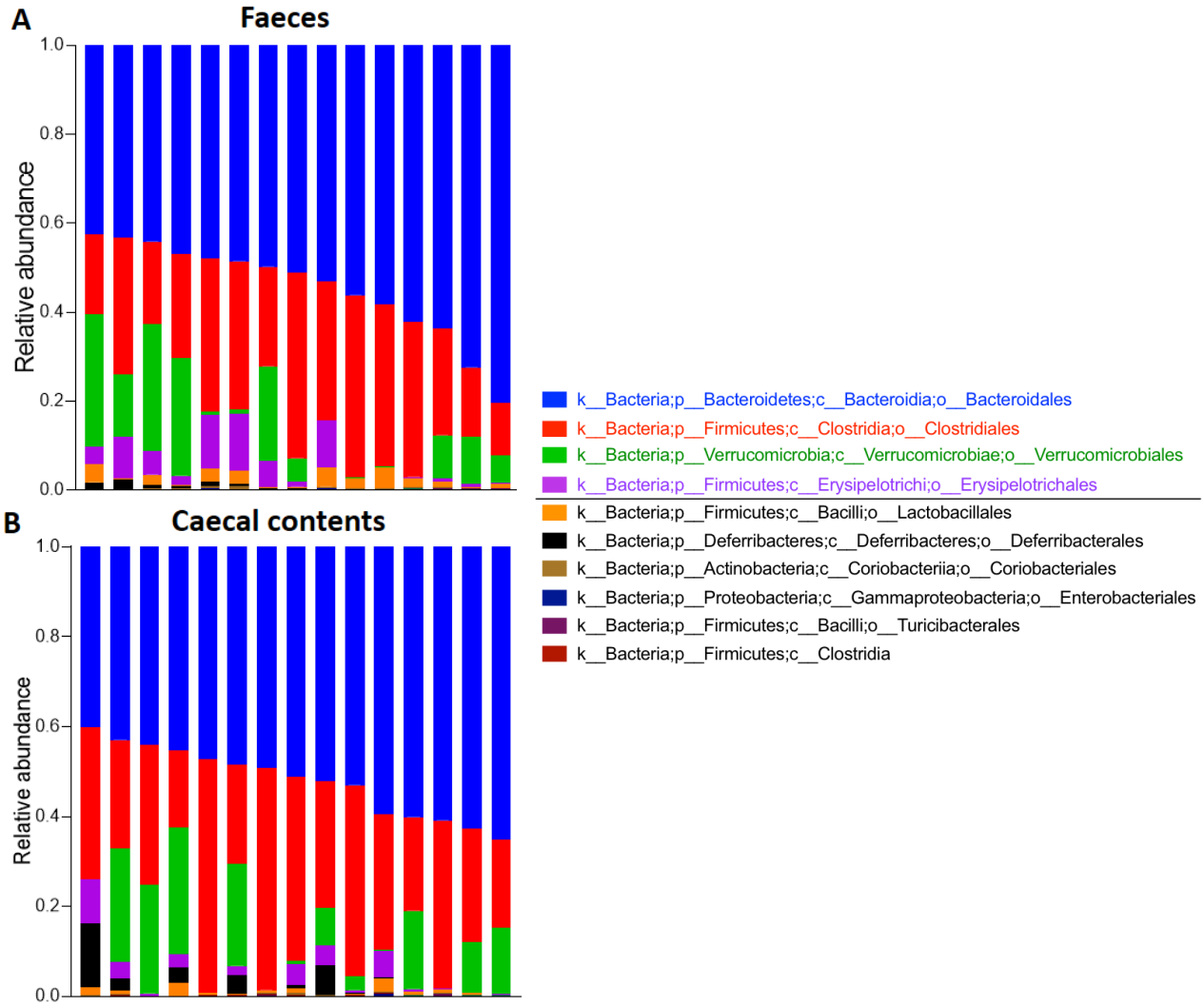


Figure 31. Phylogenetic composition of faecal and caecal microbiomes at the order level when tumours reached 100 mm^3. Samples were collected from (A) faeces (n = 15) and (B) caecal contents (n = 15), and sorted by *Bacteroidales*.

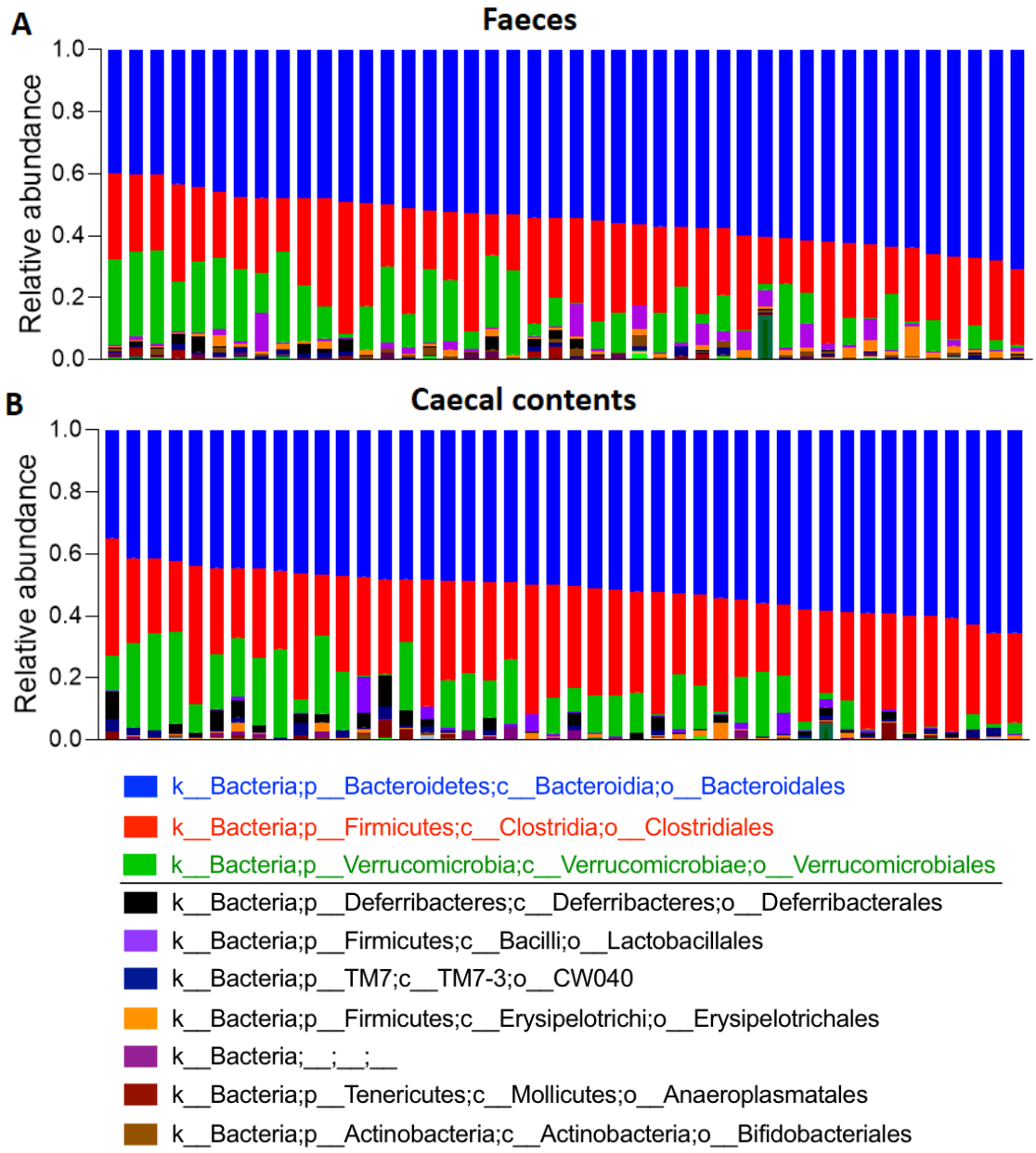


Figure 32. Phylogenetic composition of faecal and caecal microbiomes at the order level when tumours reached 700 mm^3. Samples were collected from (A) faeces (n = 44) and (B) caecal contents (n = 44), and sorted by *Bacteroidales.*

The diet did not affect the time for tumours to reach 100 mm^3

Mice were injected subcutaneously with RT112 bladder carcinoma cells, and at the same time, they started receiving diets with different ratios of dietary fibre, including low fibre, low fibre with butyrate, high mixed fibre, high soluble fibre and high insoluble fibre. The mice were culled when their tumours reached 100 mm^3. On average, the tumours took 12.8 ± 1.4 days to reach that size. No significant difference of time to culling was shown among the different diet groups (Figure 33).

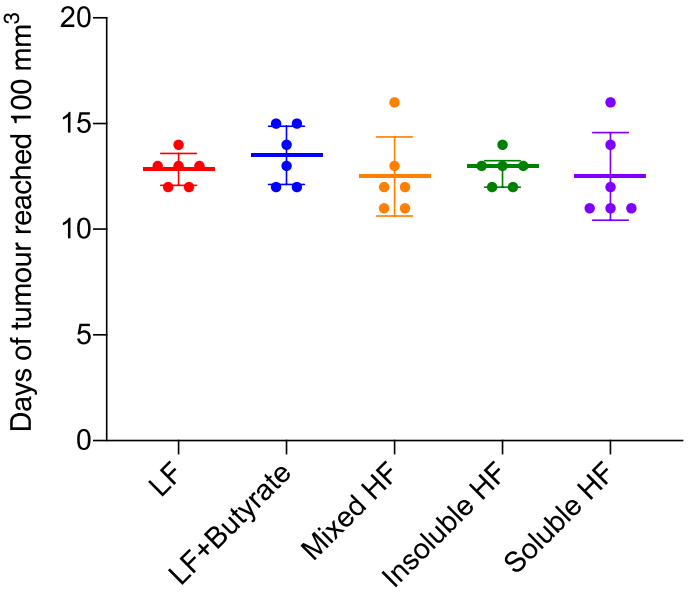


Figure 33. Time for the tumours to grow to 100 mm^3. All of the mice were culled between 11 to 16 days after tumour inoculation, average 12.8 (STD ± 1.4) days.

The landscape and diversity of the gut microbiome of samples collected when the tumours reached 100 mm^3

Faeces and caecal contents were collected when the tumours reached 100 mm^3. Abundance analysis revealed the five bacterial taxa with the highest abundance to be *Bacteroides acidifaciens*, *Parabacteroides*, *Akkermansia muciniphila*, *Lachnospiraceae* and *S24-7* (Figure 34). The *Porphyromonadaceae, Lachnospiraceae, Erysipelotrichaceae* and *Ruminococceae* were the families that related to butyrate production (Table A2). In terms of alpha diversity, the soluble high fibre group had a lower Pielou's evenness score and Shannon's index, but a similar Faith's phylogenetic diversity with the others (Figure 35A-C). This could be due to the higher abundance of *B. acidifaciens,* which lowered the diversity within groups. In terms of beta diversity, principal coordinate analysis showed a notable cluster effect among different groups which indicates that samples within groups were more similar to each other than that from the other groups (Figure 36). This indicated that the faecal microbiome was indeed modified in this study, which might be a diet-effect or a cage-effect (see later).

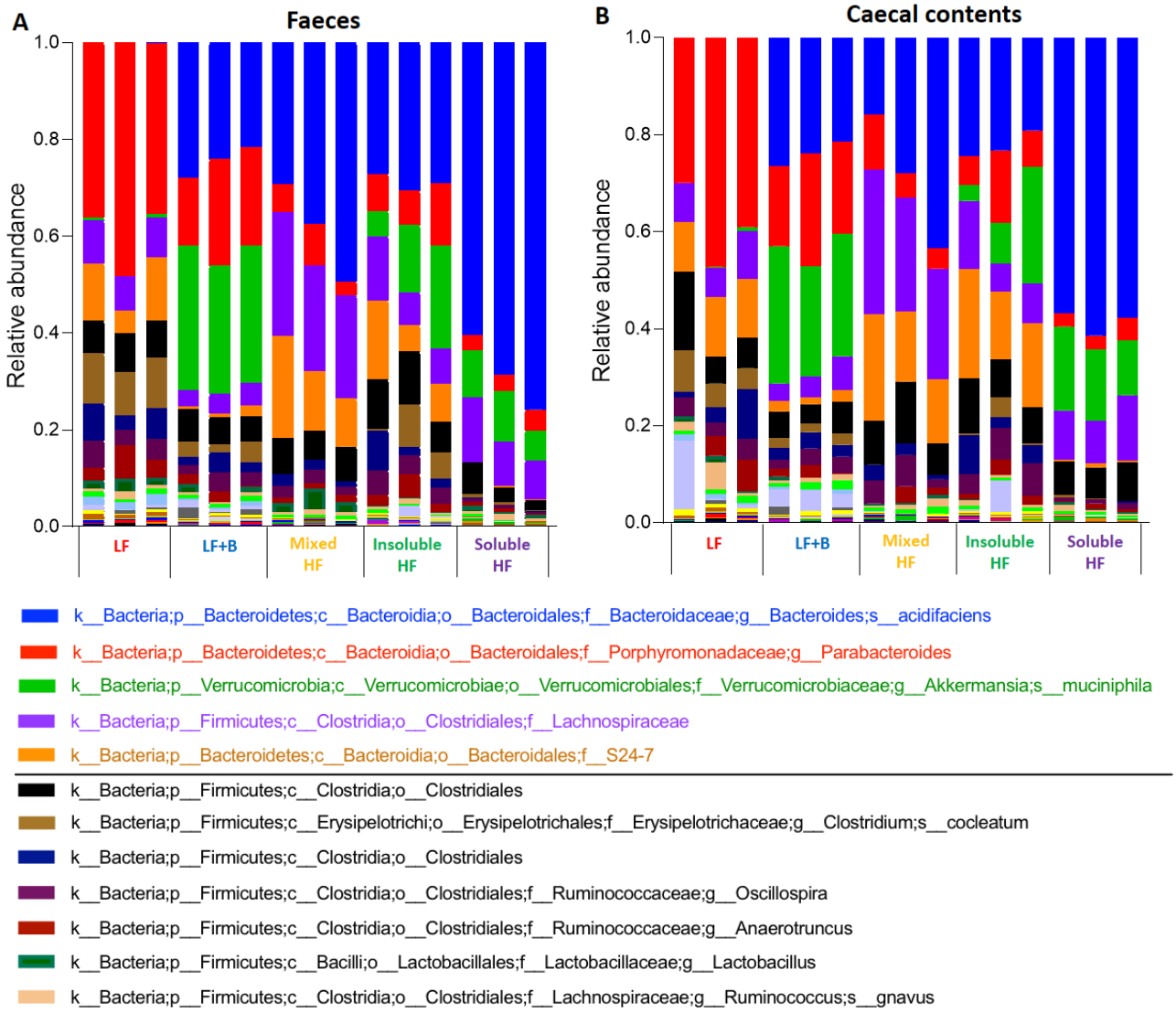


Figure 34. Stacked bar plot of phylogenetic composition of common bacterial taxa at the species level when tumours reached 100 mm^3. Samples were collected from mice fed with low fibre, low fibre with butyrate, high mixed fibre, high insoluble fibre and high soluble fibre diets (n = 3 for each group).

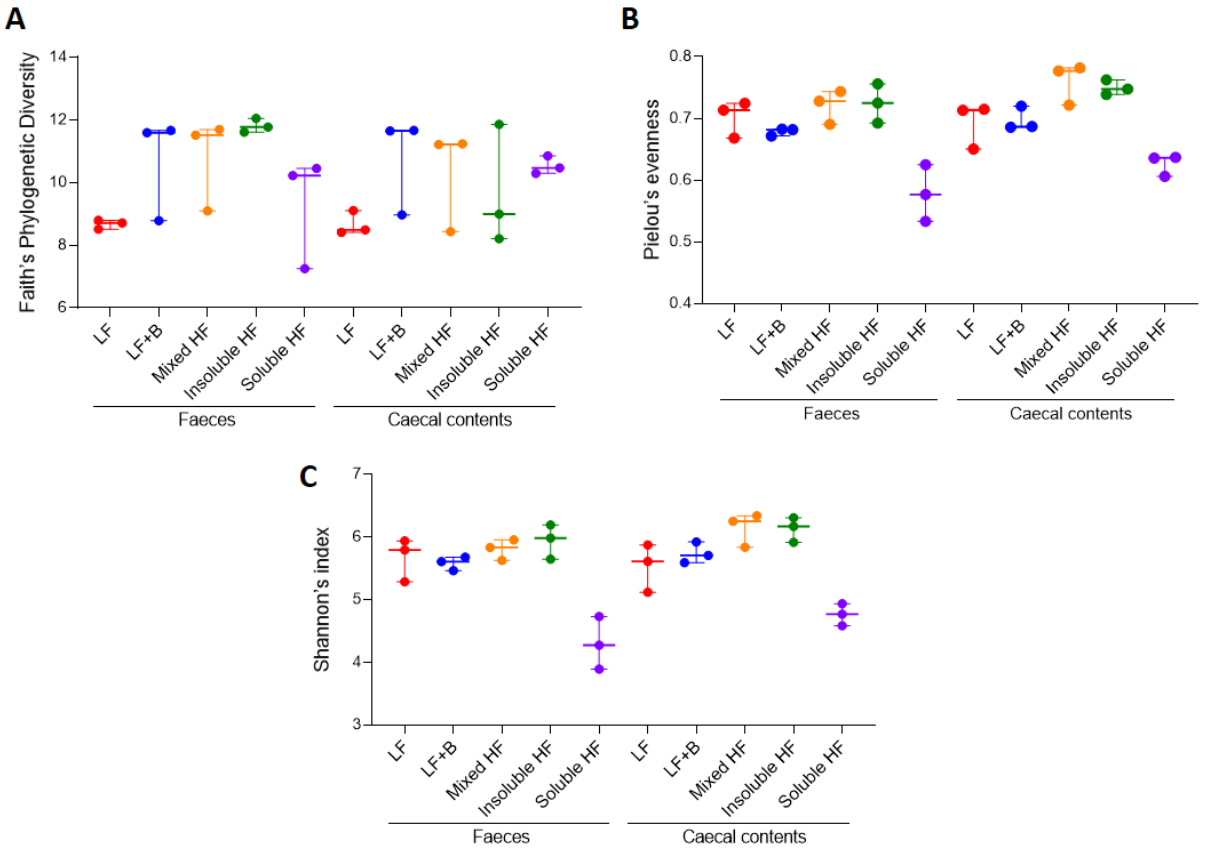


Figure 35. Alpha diversity of faecal and caecal microbiomes when tumours reached 100 mm³. (A-C) Faith's phylogenetic diversity, Pielou's evenness and Shannon's index of faecal and caecal microbiomes by Kruskal-Wallis test. Error bars represent the interquartile range of diversity scores.

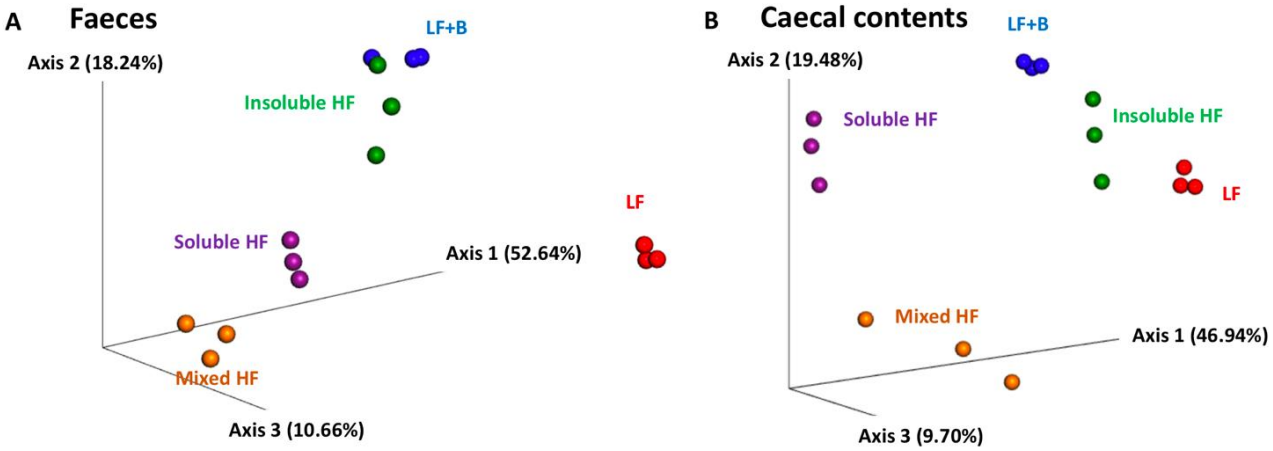


Figure 36. Beta diversity of the faecal and caecal microbiomes when the tumours reached 100 mm^3. Principal coordinate analysis of faecal and caecal microbiomes using Bray-Curtis dissimilarity. A notable clustering effect by diets was seen in the gut microbiome.

Enrichment of bacterial taxa of the gut microbiome in samples collected when the tumours reached 100 mm^3

A taxonomic cladogram of LEFse of the faecal microbiome at the order, family and genus levels showed that the low fibre diet group was enriched with *Adlercreutzia, Coriobacteriaceae, Coriobacteriales, Parabacteroides, Porphyromonadaceae, Anaerotruncus, Oscillospira, Ruminococcaceae*; the low fibre with butyrate diet group was enriched with *Akkermansia, Verrucomicrobiaceae, Verrucomicrobiales*; the high mixed fibre diet group was enriched with S24-7*, Lactobacillus, Lachnospiraceae*; the high soluble fibre diet group was enriched with *Bacteroidaceae* and *Bacteroidales* (Figure 37). To further investigate the abundance of specific taxa in different diet groups (Figure 38), the results showed that the high soluble fibre diet significantly increased *B. acidifaciens* abundance ($p < 0.001$); the low fibre diet increased *Parabacteroides* abundance ($p < 0.001$); the low fibre with butyrate increased *Akkermansia muciniphila* abundance ($p < 0.001$); the high mixed fibre diet increased *Lachnospiraceae* abundance ($p = 0.005$).

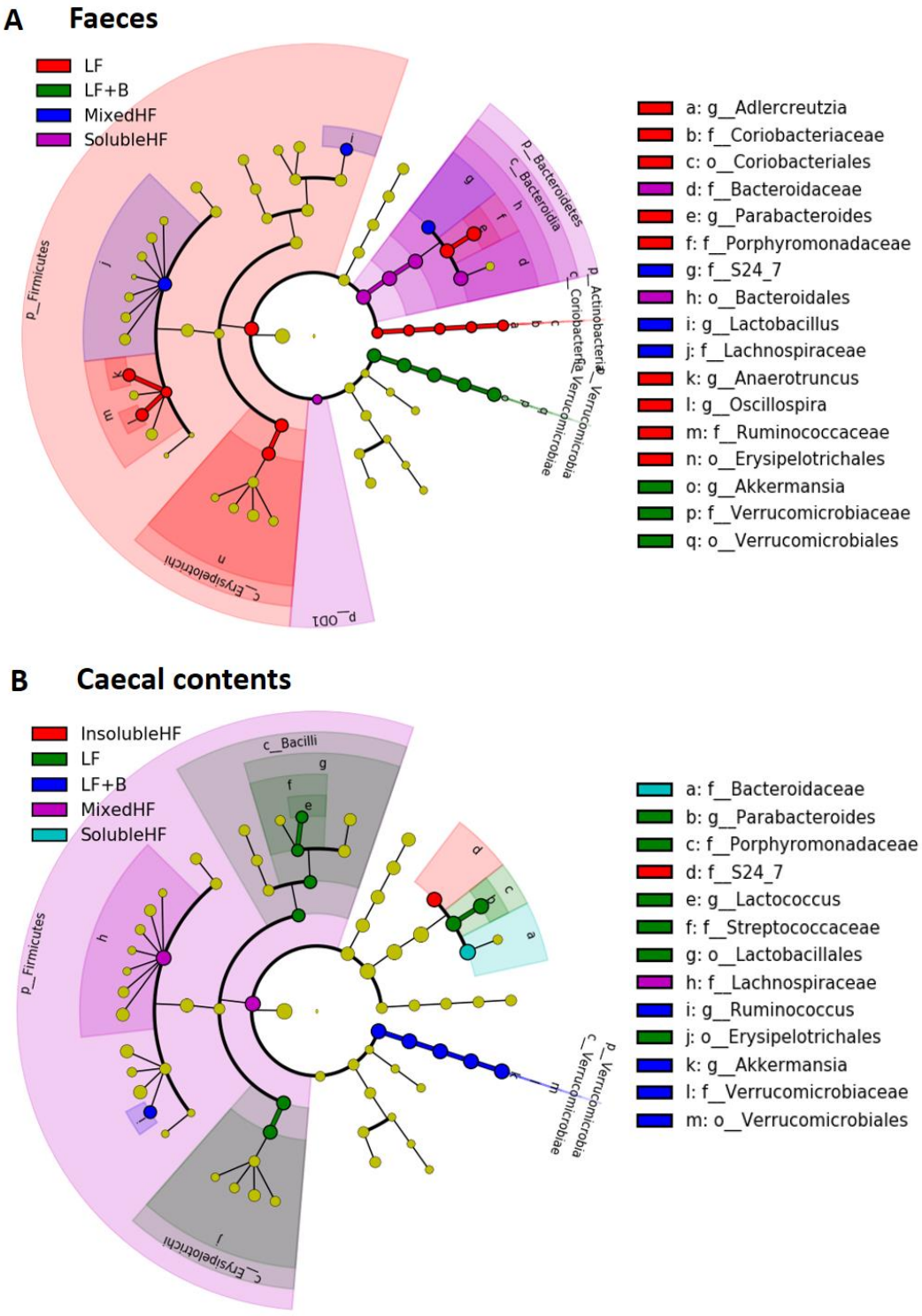


Figure 37. Taxonomic cladogram from LEfSe showing differences in faecal and caecal taxa when the tumours reached 100 mm^3. Dot size is proportional to the abundance of the taxon.

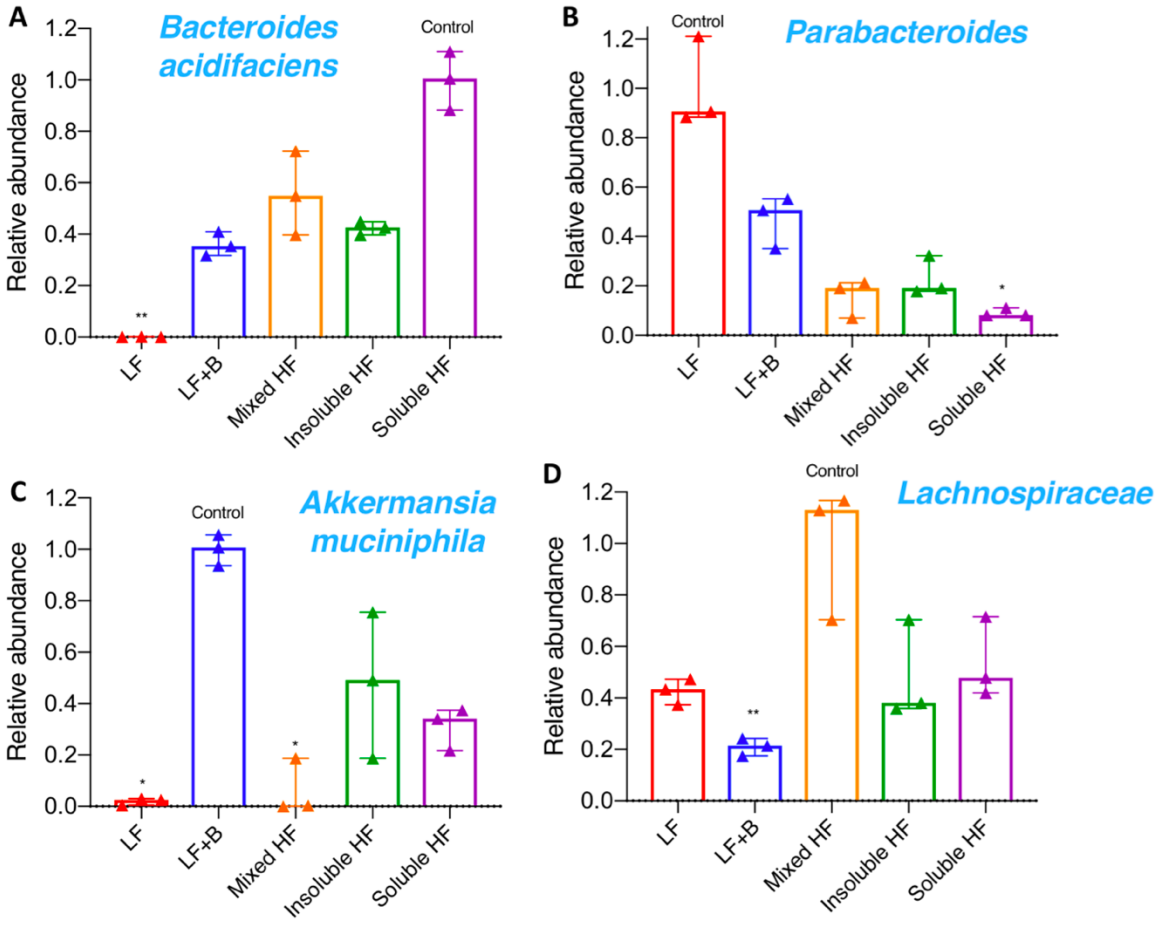


Figure 38. Differentially abundant taxa in the faecal microbiome when the tumours reached 100 mm^3. All comparisons among different diet groups was performed by Kruskal-Wallis test and Dunn's multiple comparison tests. All tests compared each median with the 'control' denoted. The diet with the highest abundance of a taxa was denoted as the control. *P < 0.05; **P < 0.01.

Metagenomic functional prediction of faecal microbiome when the tumours reached 100 mm³

By referring to KEGG orthology database level 2, the gut microbiome of the low fibre group was enriched with membrane transport pathways and the gut microbiome of the high soluble fibre group was enriched with carbohydrate metabolism pathways (Figure 39).

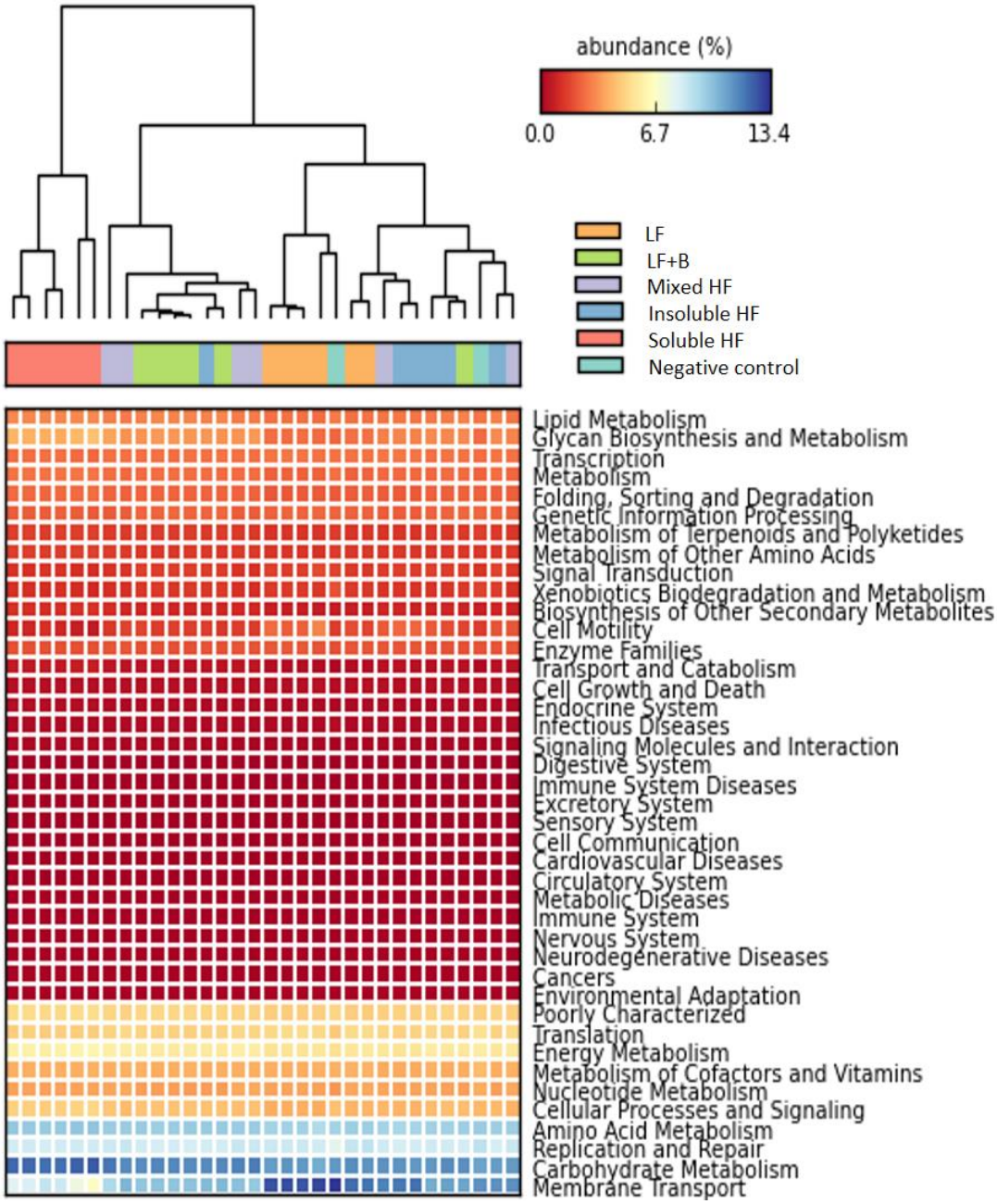


Figure 39. Metagenomic functional prediction by PICRUSt of faecal microbiome when the tumours reached 100 mm³. Prediction from 16S rRNA surveys by referring to KEGG orthology database level 2.

Growth delay in irradiated bladder cancer cell xenografts in the high soluble fibre group

To investigate the effect of different diets on the tumour response to irradiation, tumour growth was monitored to 700 mm^3. Slopes of the tumour growth curves were obtained using linear regression to indicate the tumour progression rates. Tumour growth curves revealed that the high soluble fibre diet group had the lowest growth rate. The slopes were 4.4 ± 1.3 for LF, 16.1 ± 1.7 for mixed HF, 28.7 ± 1.3 for insoluble HF and 0.4 ± 1.5 for soluble HF (Figure 40A). Kaplan Meier survival curves revealed that the soluble HF group had the longest median survival time (7.5 days for LF, 7 days for mixed HF, 10 days for insoluble HF, 11.5 days for soluble HF) and the p-value of the log-rank test was 0.005 (Figure 40B).

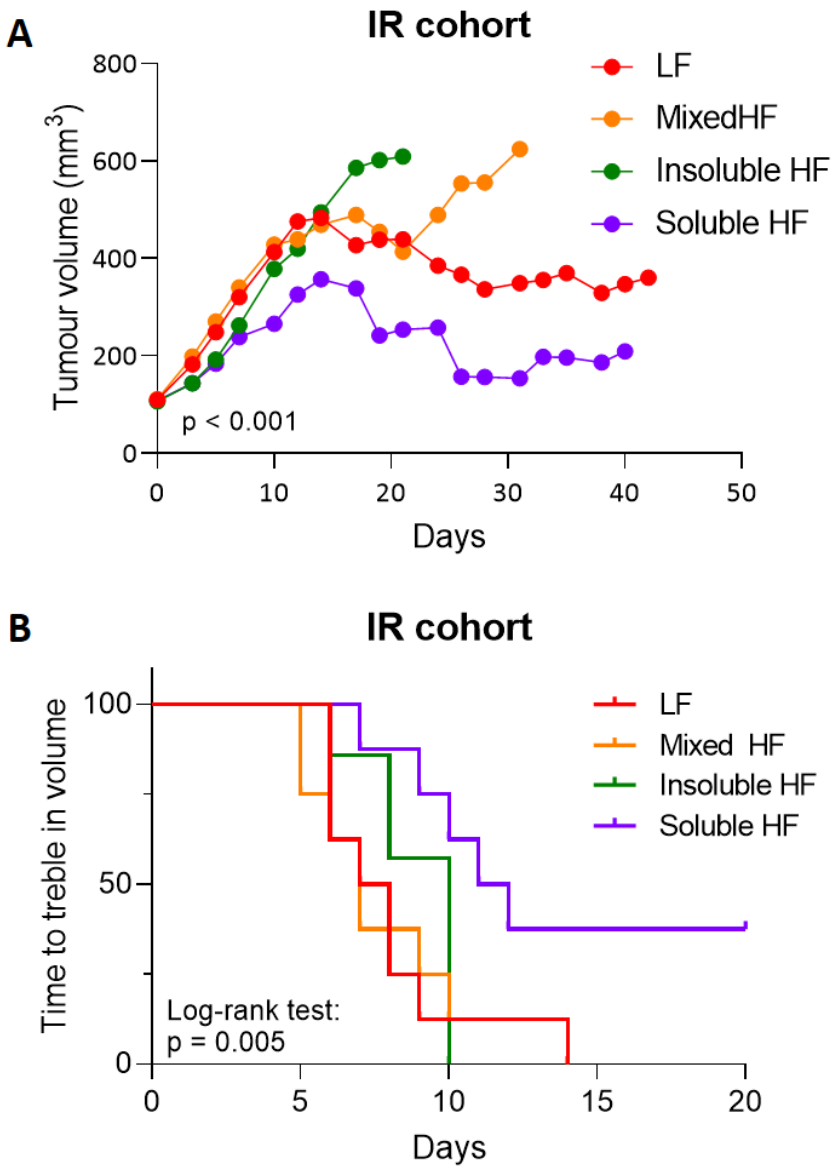


Figure 40. Soluble HF causes increased growth delay in irradiated bladder cancer cell xenografts. (A) Treatment of irradiated (6 Gy) RT112 xenografts with low fibre, high mixed fibre, high insoluble fibre and high soluble fibre diets (n = 8 for each group). Slopes of tumour curves were calculated by linear regression to represent tumour growth rates and compared by ANOVA test. (B) Kaplan–Meier survival curve of mice with RT112 xenografts showing plots of time to treble tumour volume.

The landscape and diversity of the gut microbiome of samples collected when the tumours reached 700 mm^3

When the tumours reached 700 mm^3, abundance analysis of the gut microbiome of both IR and non-IR cohorts revealed that the top 6 bacterial taxa with the highest abundance were *S24-7, Akkermansia muciniphila, Bacteroides*, *Lachnospiraceae, Clostridiales* and *B. acidifaciens* (Figure 41). The *Lachnospiraceae, Porphyromonadaceae, Rikenellaceae* and *Ruminococceae* were the families related to butyrate production (Table A3). In terms of alpha diversity, no significant difference in Faith's phylogenetic diversity, Pielou's evenness score and Shannon's index was shown among the different diet groups (Figure 42A-C). In terms of beta diversity, principal coordinate analysis using Bray-Curtis dissimilarity showed a notable cluster effect among different groups, which indicates that samples within groups were more similar to each other than those from the other groups (Figure 43). This indicated that the diets still affected the bacterial compositions of the gut microbiome when mice got older until the tumours reached 700 mm^3.

Enrichment of bacterial taxa of the gut microbiome of samples collected when the tumours reached 700 mm^3

Taxonomic cladograms of LEFse of the faecal microbiome showed that the low fibre diet increased *Adlercreutzia, Coriobacteriales, Bacteroides, Lactococcus, Streptococcaceae, Lactobacillales, Oscillospira, Ruminococcus, Ruminococcaceae, Clostridiales, Clostridium*; the high mixed fibre increasing *Odoribacter, Odoribacteraceae, F16, CW040*; the high insoluble fibre diet increased *Rikenellaceae, Mucispirillum, Deferribacteraceae, Deferribacterales, Ruminococcus, Mogibacteriaceae, Coprobacillus, Anaeroplasma, Anaeroplasmataceae, Anaeroplasmatales, Akkermansia, Verrucomicrobiaceae, Verrucomicrobiales*; and the high soluble fibre diet increased *S24-7* (Figure 44).

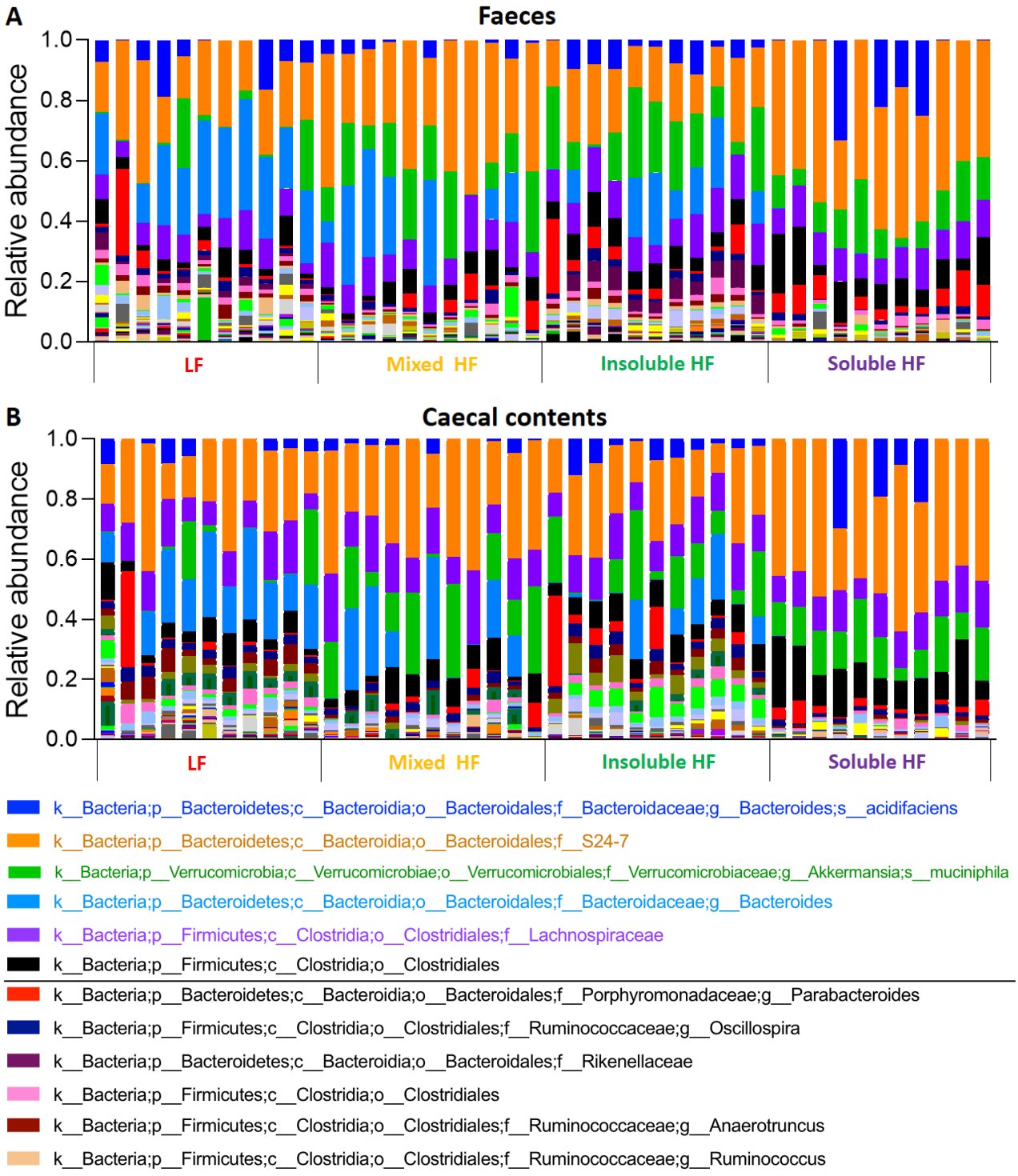


Figure 41. Stacked bar plot of phylogenetic composition of common bacterial taxa at the species level when tumours reached 700 mm^3. Samples were collected from mice fed with low fibre, high mixed fibre, high insoluble fibre and high soluble fibre diets (n = 8 for each group).

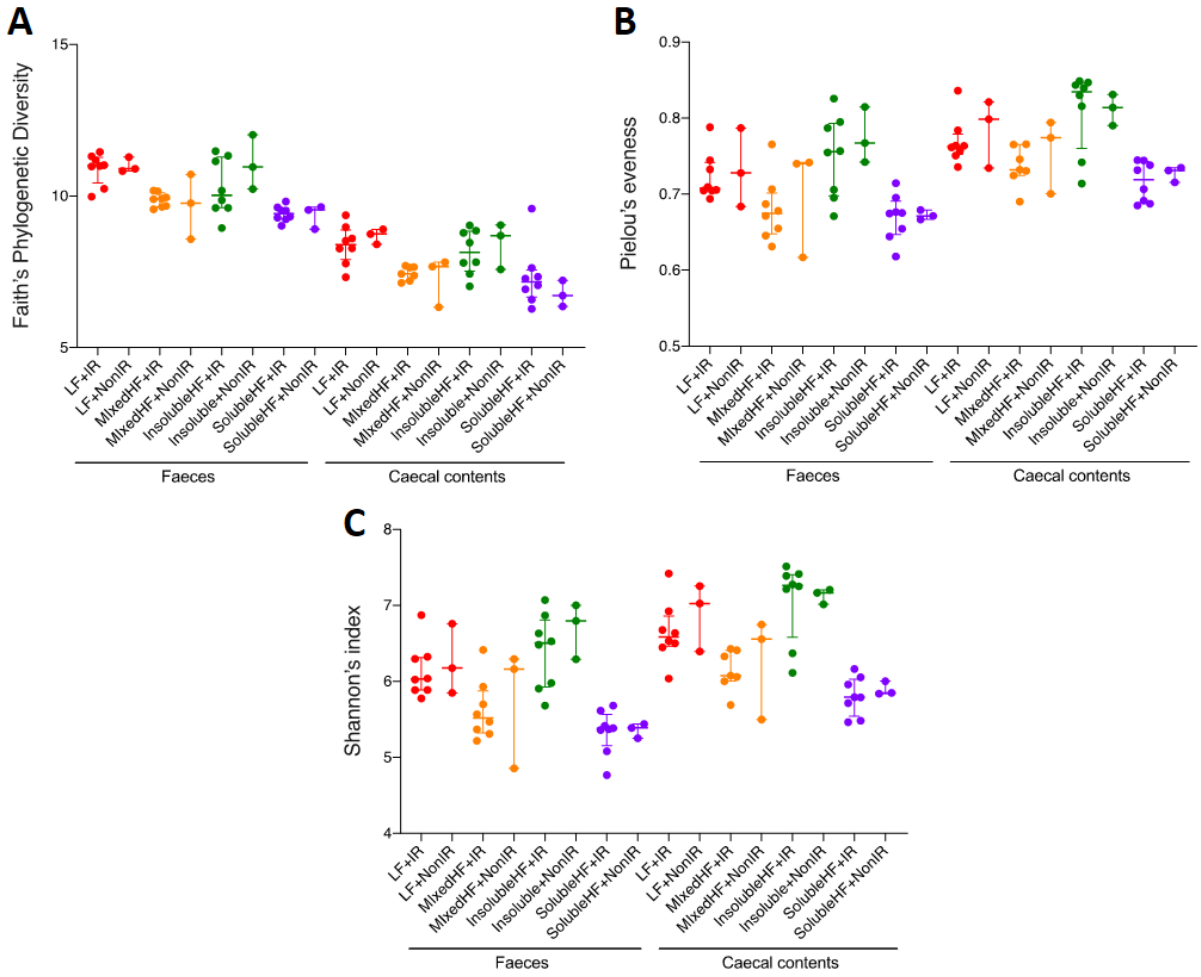


Figure 42. Alpha diversity of the faecal and caecal microbiomes when the tumours reached 700 mm^3. (A-C) Faith's phylogenetic diversity, Pielou's evenness and Shannon's index of faecal and caecal microbiomes by Kruskal-Wallis test. Error bars represent the interquartile range of diversity scores.

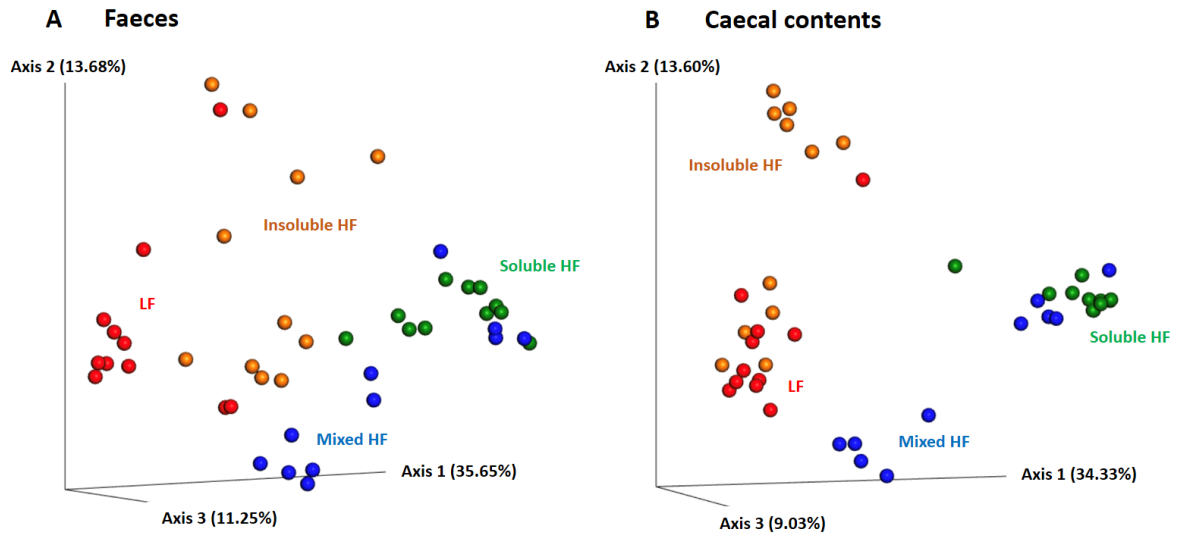


Figure 43. Beta diversity of the faecal and caecal microbiomes when the tumours reached 700 mm^3. Principal coordinate analysis of faecal and caecal microbiomes using Bray-Curtis dissimilarity.

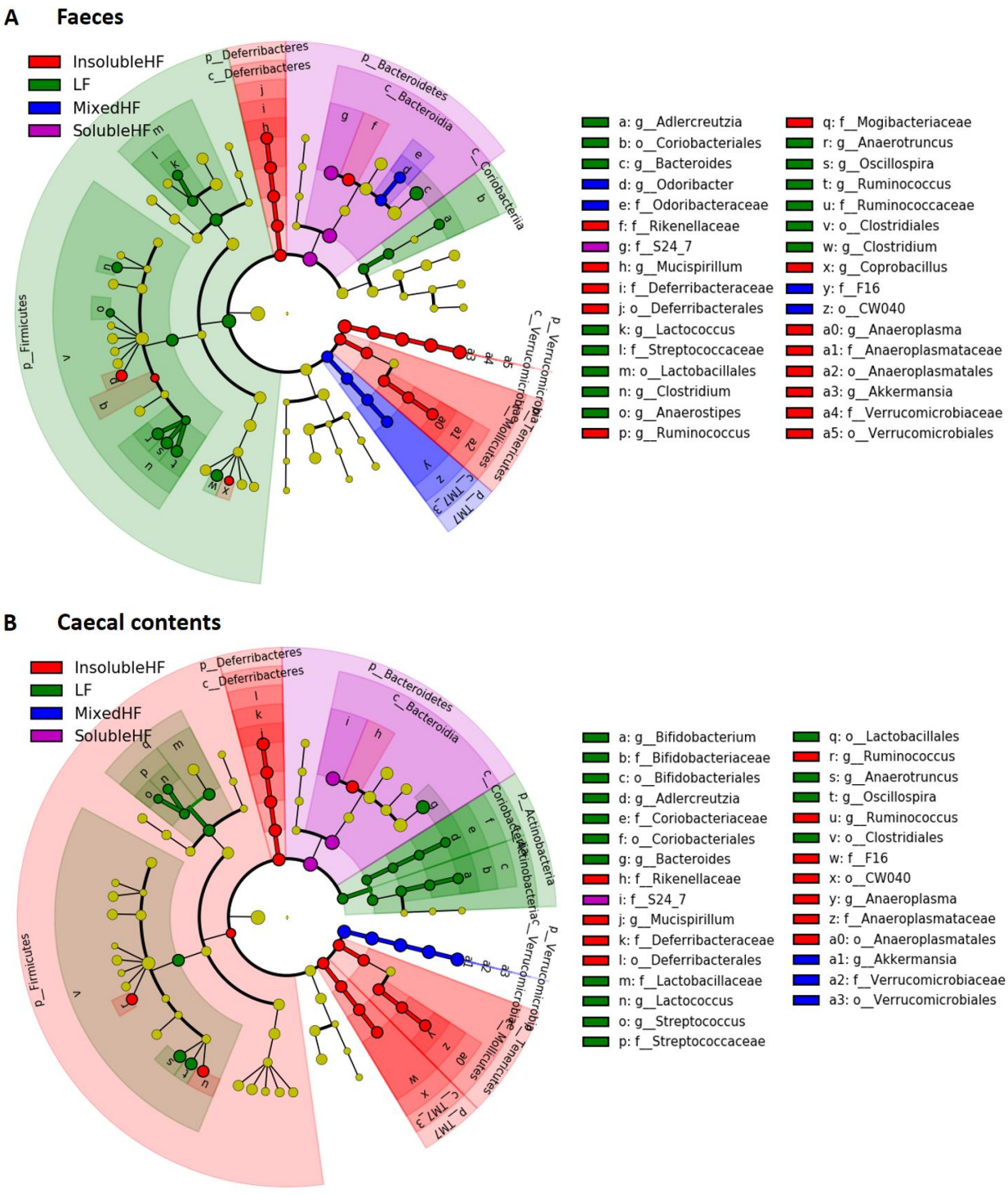


Figure 44. Taxonomic cladogram from LEfSe showing differences in faecal and caecal taxa in the IR cohort when the tumours reached 700 mm^3. Dot size is proportional to the abundance of the taxon.

Stratification of the high soluble fibre group based on tumour response to irradiation

The slopes of the tumour growth curves were used to estimate the tumour growth rate. Four mice were classified as responders with smaller slopes; their slopes were 7.3 ± 1.2, -0.9 ± 0.6, -4.8 ± 0.5, -5.2 ± 0.9, and the other four mice were classified as non-responders with larger slopes; their slopes were 34.6 ± 3.0, 31.1 ± 2.6, 23.3 ± 1.0, 33.8 ± 2.9 (Figure 45).

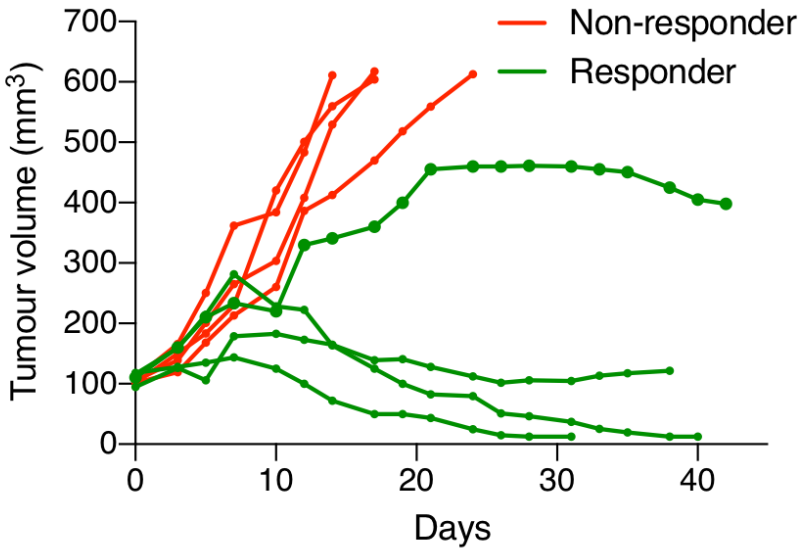


Figure 45. Tumour growth in the soluble HF group with irradiation. Mice were stratified into responders and non-responders based on the tumour response to irradiation.

Diversity and enrichment of the faecal microbiome by response in the soluble HF group

In terms of alpha diversity, no significant difference of Faith's phylogenetic diversity, Pielous's evenness or Shannon's index was shown between responders and non-responders (Figure 46). In terms of beta diversity, principal coordinate analysis of Bray-Curtis dissimilarity showed the gut microbiome of responders and non-responders were more similar within groups than between groups (Figure 47). Linear discriminant analysis showed responding mice were enriched with *Bacteroidaceae, Flavobacterium, Flavobacteriales, Lactococcus, Streptococcus, Streptococcaceae, Allobaculum*, *Erysipelotrichales*, and non-responding mice were enriched with *Bifidobacterium, Bidifobacteriaceae, Bifidobacteriales, Parabacteroides, Porphyromonadaceae, Lactobacillus, Lactobacillaceae* and *Lactobacillales* (Figure 48A). In terms of effect size, *B. acidifaciens* and *Bacteroidaceae* had the largest enrichment in responders, and *Parabacteroides* and *Porphyromonadaceae* had the largest enrichment in non-responders (Figure 48B).

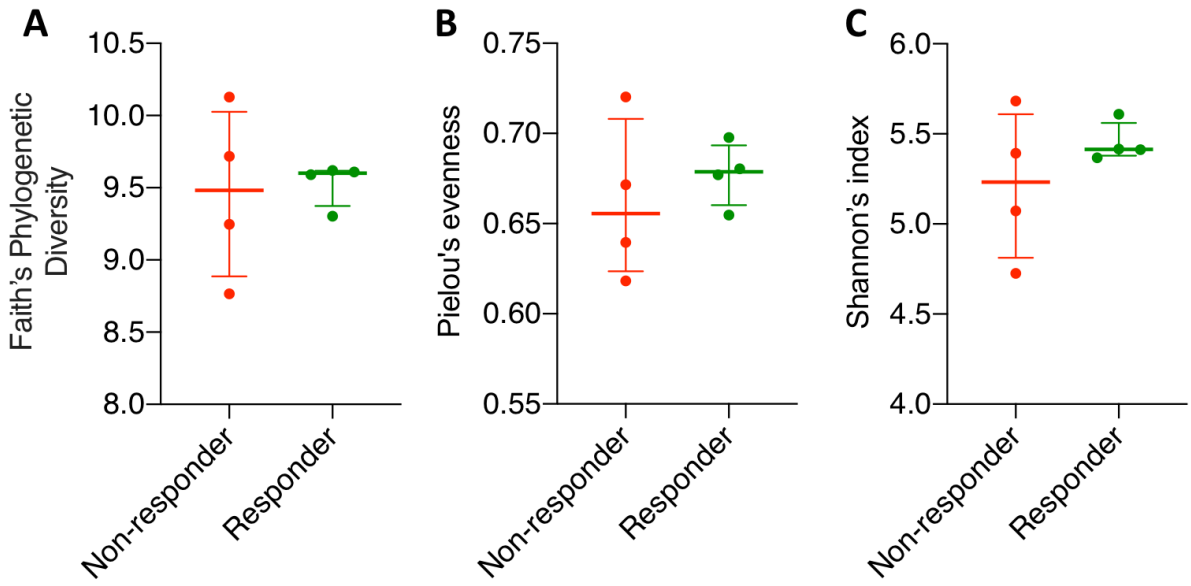


Figure 46. Alpha diversity of the faecal microbiome in responders and non-responders in the soluble HF group. (A-C) Faith's phylogenetic diversity, Pielou's evenness and Shannon's index of faecal microbiome in responders and non-responders by Kruskal-Wallis test. Error bars represent the interquartile range of diversity scores.

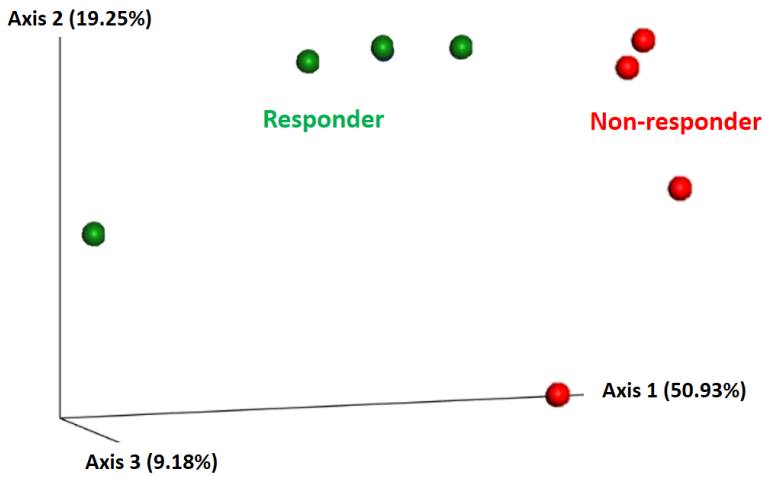


Figure 47. Beta diversity of the faecal microbiome in responders and non-responders in the soluble HF group. Principal coordinate analysis of faecal samples (n = 8) in soluble HF group by response using Bray-Curtis dissimilarity.

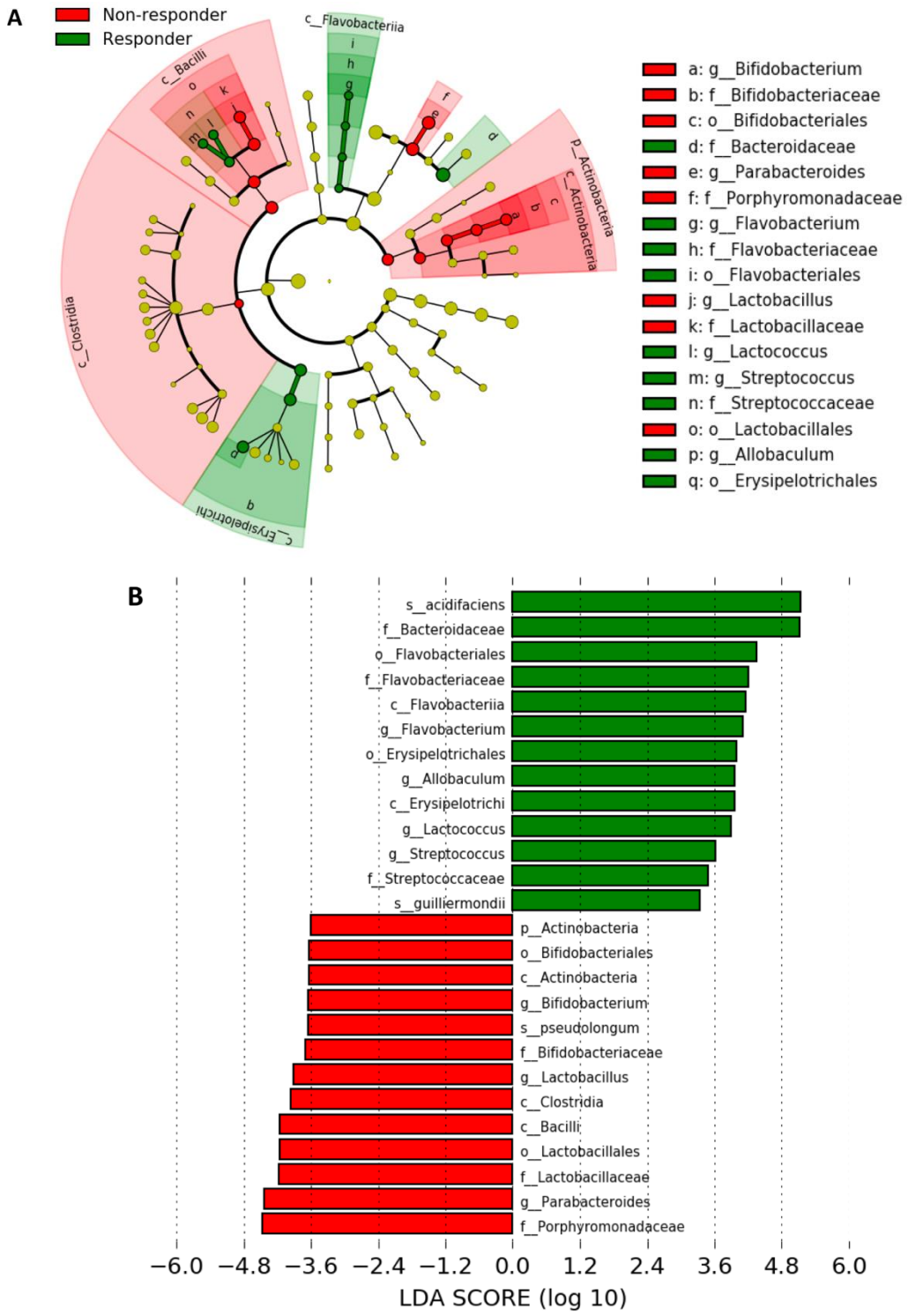


Figure 48. Differences in bacterial components in responders and non-responders in the soluble HF group. (A) Taxonomic cladogram from LEfSe showing differences of taxa in responders and non-responders of soluble HF group. Dot size is proportional to the abundance of the taxon. (B) Linear discriminant analysis (LDA) scores computed for differentially abundant taxa in the microbiomes of responders (green) and non-respondners (red). Length indicates the effect size associated with a taxon and p = 0.05 for Kruskal-Wallis test.

Discrete false-discovery rate and metagenomics functional prediction of the faecal microbiome by response in the soluble HF group

To further explore these findings, the discrete false-discovery rates within all taxonomic levels were calculated (Figure 49). The results showed that the *B. acidifaciens* species and the *Allobaculum* genus had p-values of less than 0.05 in responders, and *Lactobacillus* and *Parabacteroides* had this p-value in non-responders. Functional prediction at KEGG level 2 revealed that the faecal microbiome in responders was enriched by carbohydrate metabolism pathways and in non-responders by membrane transport pathways (Figure 50). To further explore this at KEGG level 3, enrichment of the amino sugar and nucleotide sugar metabolism pathway was shown in responders and enrichment of the ATP-binding cassette (ABC) transporter pathway was shown in non-responders (Figure 51).

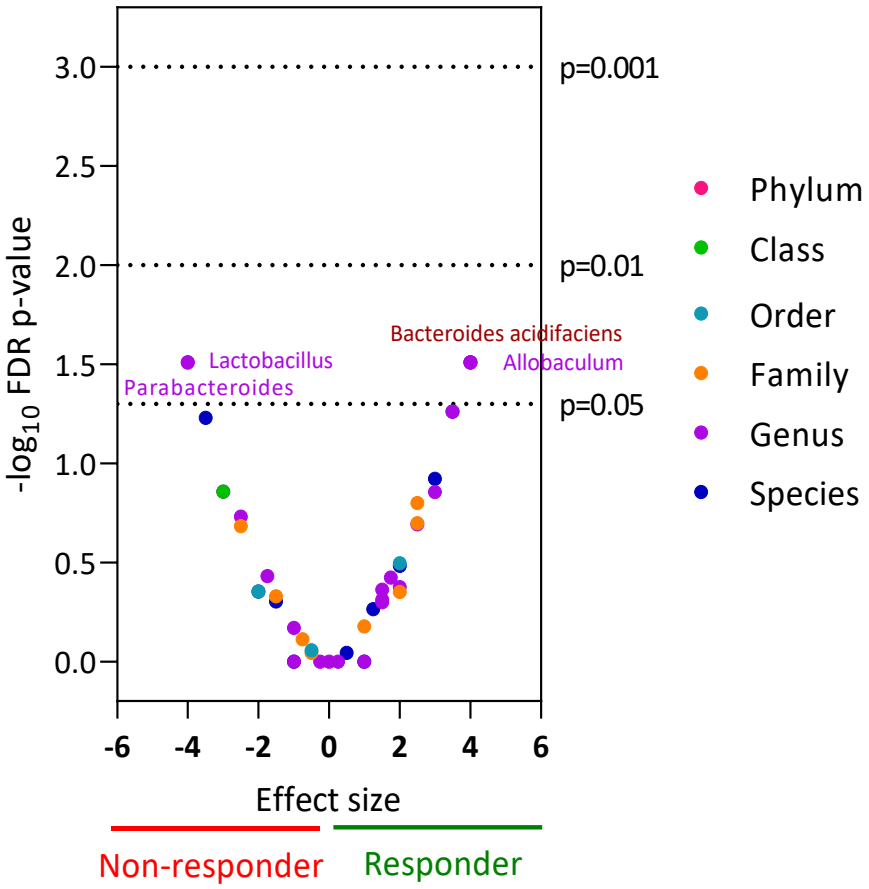


Figure 49. Discrete false-discovery rate of different abundant taxa in responders and non-responders in the soluble HF group. Differential abundance testing within all taxonomic levels in responders versus non-responders by Mann-Whitney U test.

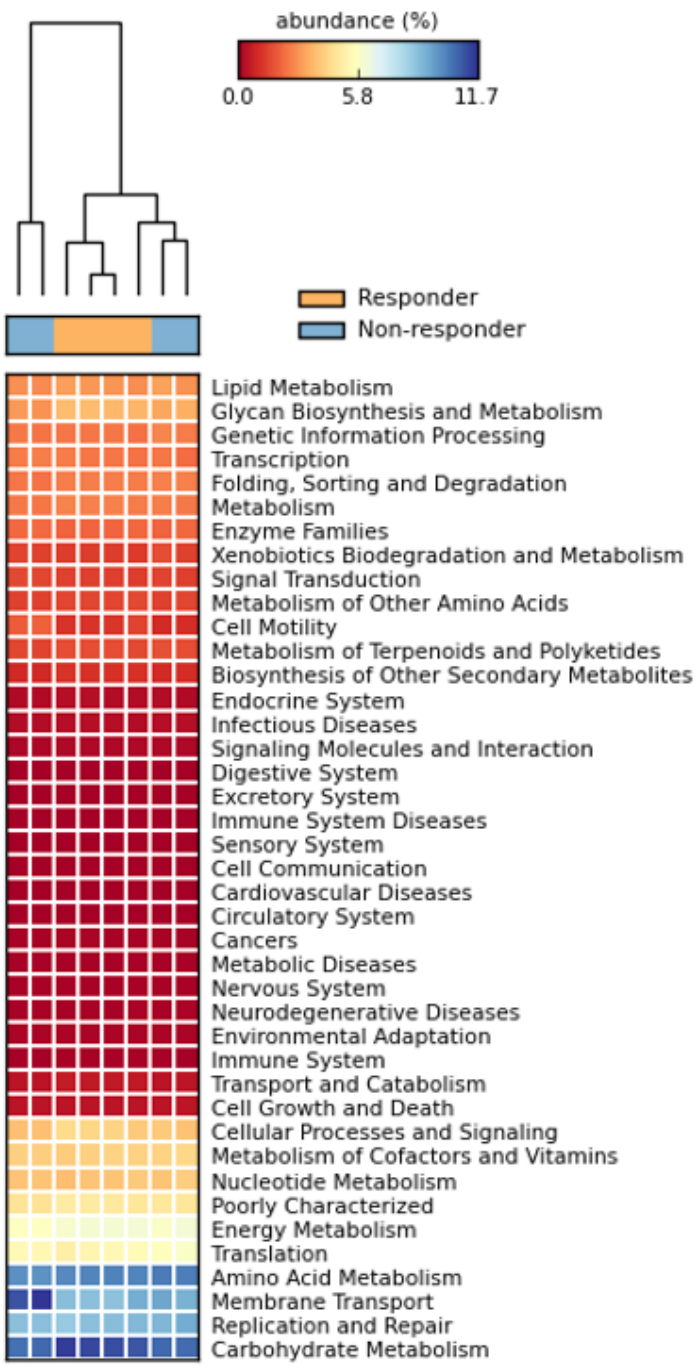


Figure 50. Metagenomic functional prediction by PICRUSt of the faecal microbiome in responders and non-responders in the soluble HF group (level 2). Prediction from 16S rRNA surveys in the metagenomes of the faecal microbiome from responders (n = 4) and non-responders (n = 4) in the soluble HF group by referring to KEGG database level 2. Columns represent mice (responders, orange; non-responders, blue), and rows represent enrichment of predicted KEGG pathways (red, low enrichment; yellow, medium enrichment; blue, high enrichment).

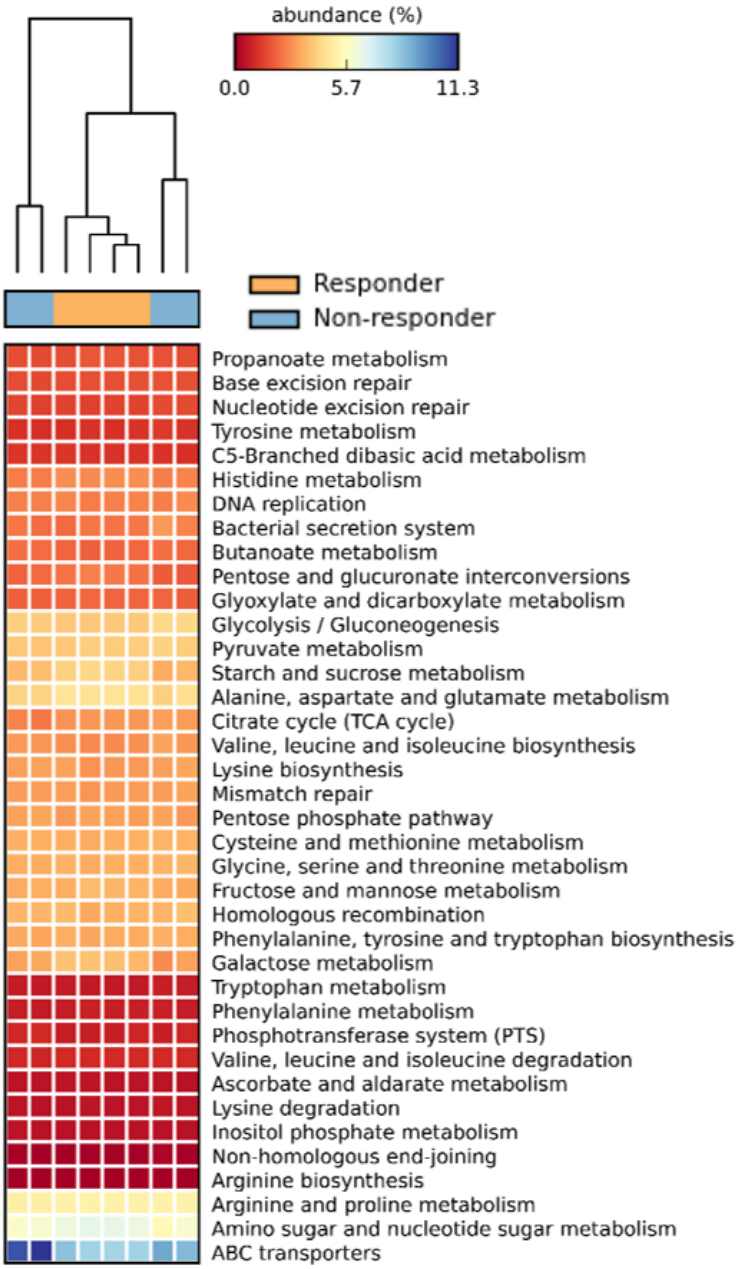


Figure 51. Metagenomic functional prediction by PICRUSt of the faecal microbiome in responders and non-responders in the soluble HF group (level 3). Prediction from 16S rRNA surveys in the metagenomes of faecal microbiome from responders (n = 4) and non-responders (n = 4) in the soluble HF group by referring to KEGG database level 3. Columns represent mice (responders, orange; non-responders, blue), and rows represent enrichment of predicted KEGG pathways (red, low enrichment; yellow, medium enrichment; blue, high enrichment).

Correlation between *B. acidifaciens* abundance and mouse survival time in IR and non-IR cohorts in different diet groups

To explore how specific bacterial taxa affect mouse survival time, the correlation between *B. acidifaciens* abundance and time to culling was investigated, because *B. acidifaciens* was the 'top hit' for responders in the soluble HF group. *B. acidifaciens* abundance was evenly distributed among all cages, and some mice (n = 4) in the soluble HF group had the highest abundance (Figure 52). Some mice in the non-IR cohort lived as long as those in the IR cohort, ie. >40 days, which may be a reflection of 6 Gy being a relatively low dose of irradiation. In the IR cohort, the time of culling positively correlated with *B. acidifaciens* abundance (R^2 = 0.528, $p < 0.001$). However, a similar correlation was not seen in the non-IR cohort (R^2 = 0.085, p = 0.357). Using the time for tumours to treble in volume as the outcome measure, mice with high *B. acidifaciens* abundance had a significantly prolonged median survival time (Log-rank test: $p < 0.001$) (Figure 53), and a similar finding was seen in the IR cohort (p = 0.003), but not in the non-IR cohort (p = 0.236). *B. acidifaciens* abundance was evenly distributed across different cages (Figure 54) and this suggested that the existence of this taxa in the gut microbiome was not a cage-specific effect.

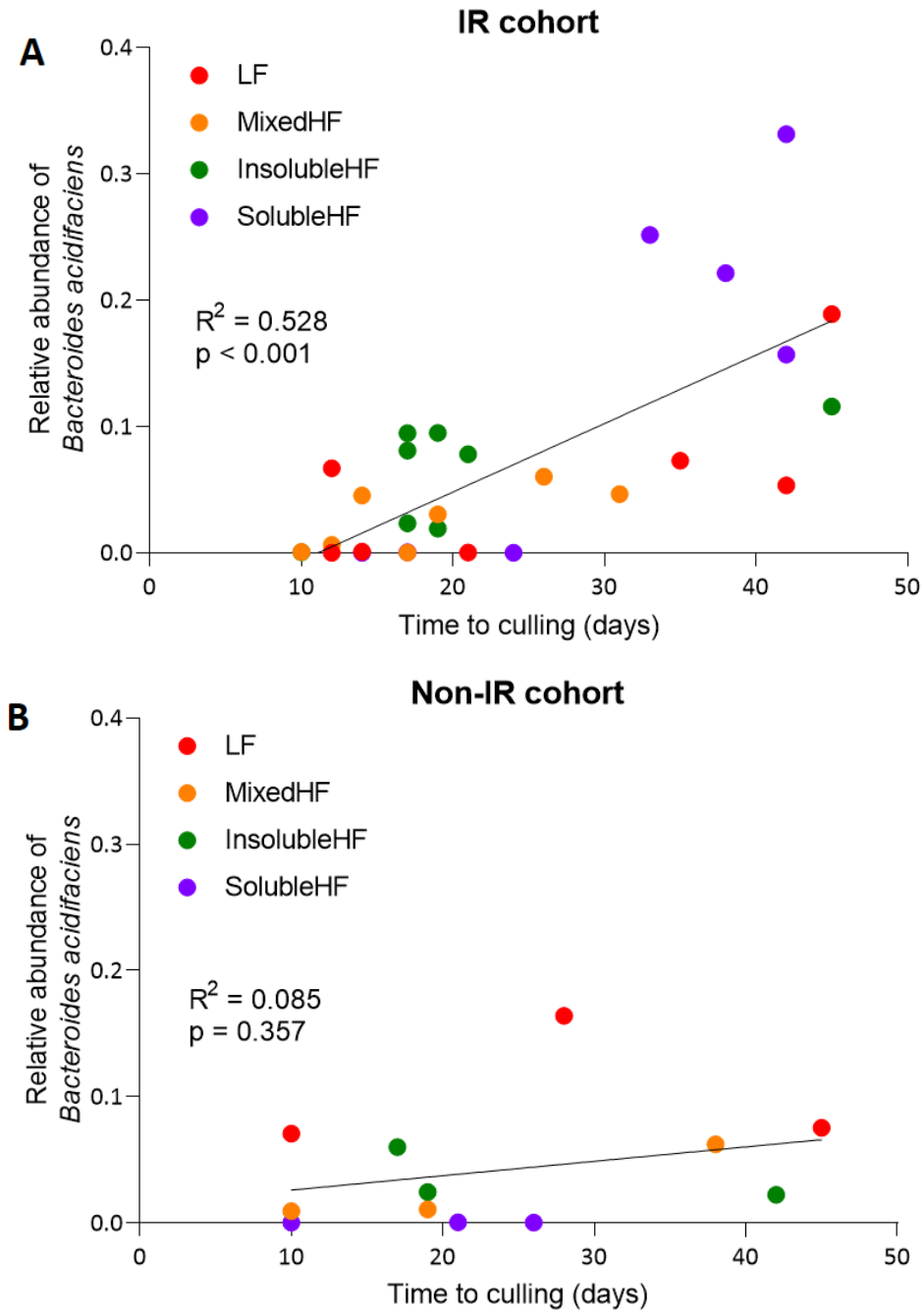


Figure 52. Correlation of time to culling with *B. acidifaciens* abundance in IR and Non-IR cohorts from different diet groups. (A & B) Correlation between time of culling versus *B. acidifaciens* abundance (IR cohort, $R^2 = 0.53$, $P < 0.001$; Non-IR cohort, $R^2 = 0.09$, $P = 0.36$) in the faecal microbiome.

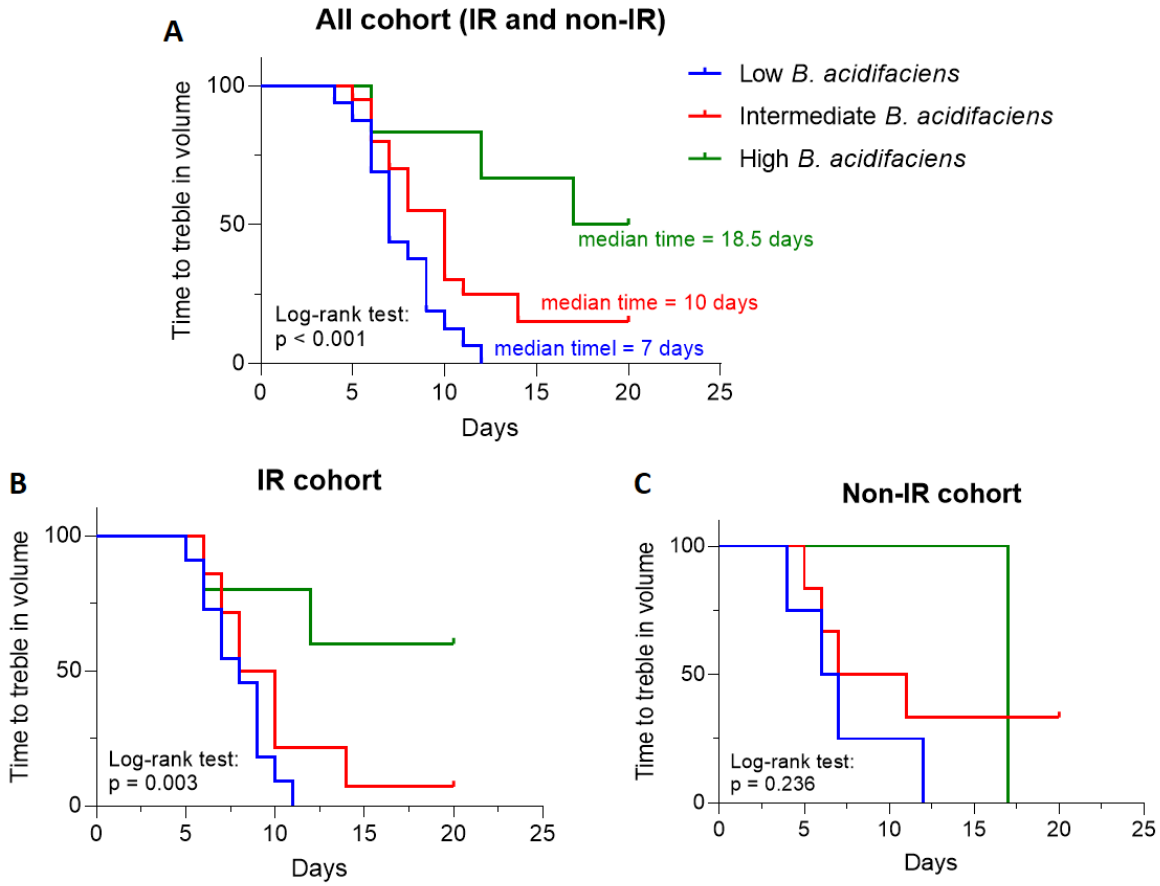


Figure 53. Kaplan-Meier (KM) plots of time for tumours to treble in volume based on *B. acidifaciens* abundance in IR and Non-IR cohorts from different diet groups combined. Comparison KM plots by log-rank test in mice with high abundance (green = undefined), intermediate abundance (red), or low abundance (blue) of *B. acidifaciens* in (A) all, (B) IR and (C) non-IR cohorts.

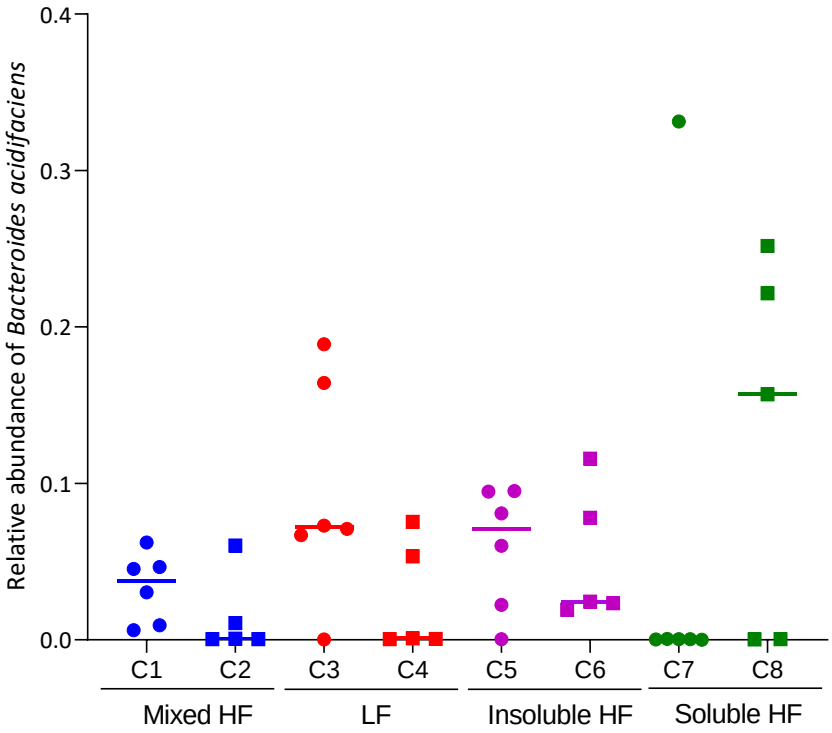


Figure 54. Relative abundance of B. acidifaciens of the faecal microbiome in mice from different cages. Comparison of *B. acidifaciens* abundance among mice from cage 1 (C1) to cage 8 (C8) by Kruskal-Wallis test.

Correlation between *Parabacteroides* genus abundance and mouse survival time in IR and non-IR cohorts in different diet groups

To further identify specific unfavourable bacterial taxa which might offset the effect of irradiation, the *Parabacteroides* genus*,* was selected, as the 'top hit' for non-responders in the soluble HF group. It was investigated for the association between its abundance and time to culling (Figure 55). In the IR cohort, the time to culling positively correlated with the *Parabacteroides* genus abundance (R^2 = 0.164, p = 0.022). However, a similar correlation was not seen in the non-IR cohort (R^2 = 0.084, p = 0.360). Using the time for tumours to treble in volume as the outcome measure, mice in the low *Parabacteroides* genus abundance group had no significant difference in median survival time compared to the high abundance group (Log-rank test: p = 0.374) (Figure 56). *Parabacteroides* genus abundance was evenly distributed among different cages (Figure 57) and this suggested that the existence of this taxon in the gut microbiome was not a cage-specific effect.

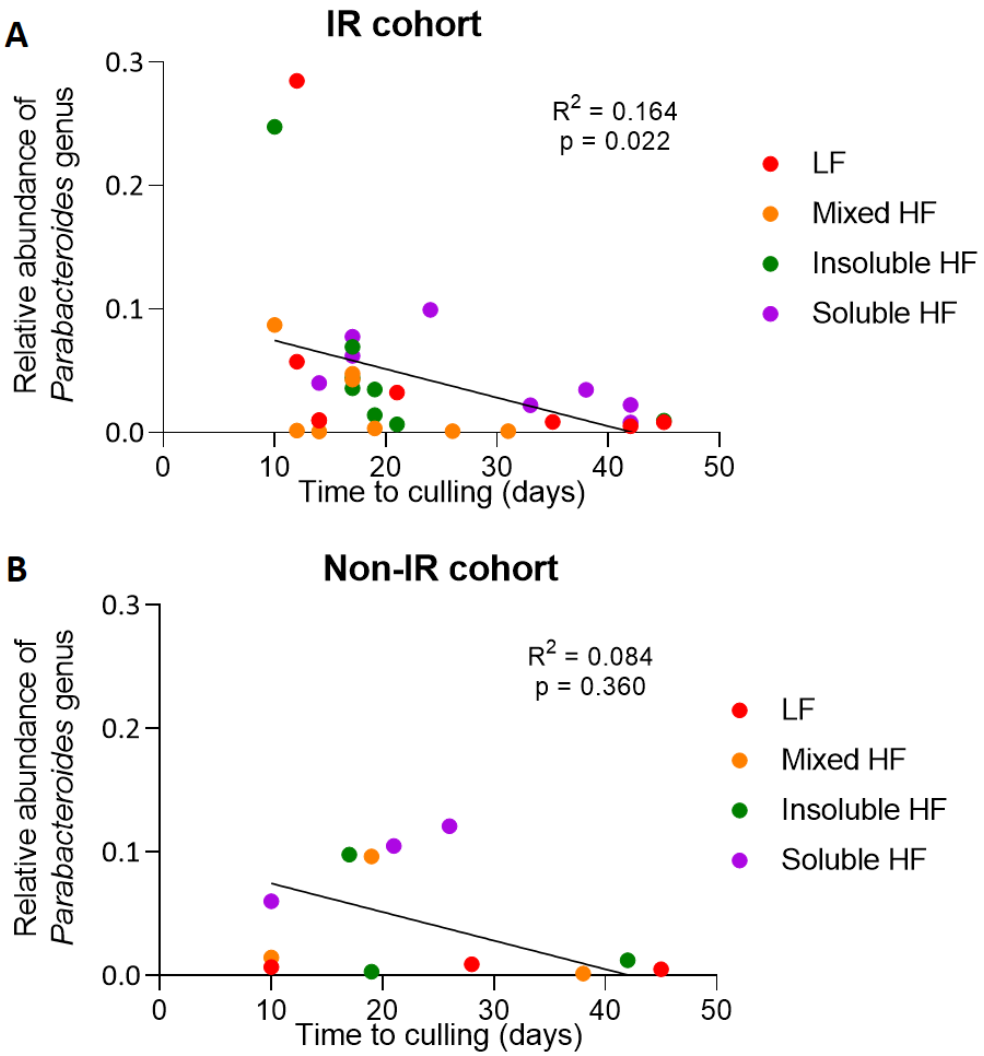


Figure 55. Correlation of time to culling with *Parabacteroides* genus abundance in IR and non-IR cohorts from different diet groups. (A & B) Correlation between time of culling versus *Parabacteroides* genus abundance (IR cohort, $R^2 = 0.164$, P = 0.022; Non-IR cohort, $R^2 = 0.084$, P = 0.360) in the faecal microbiome.

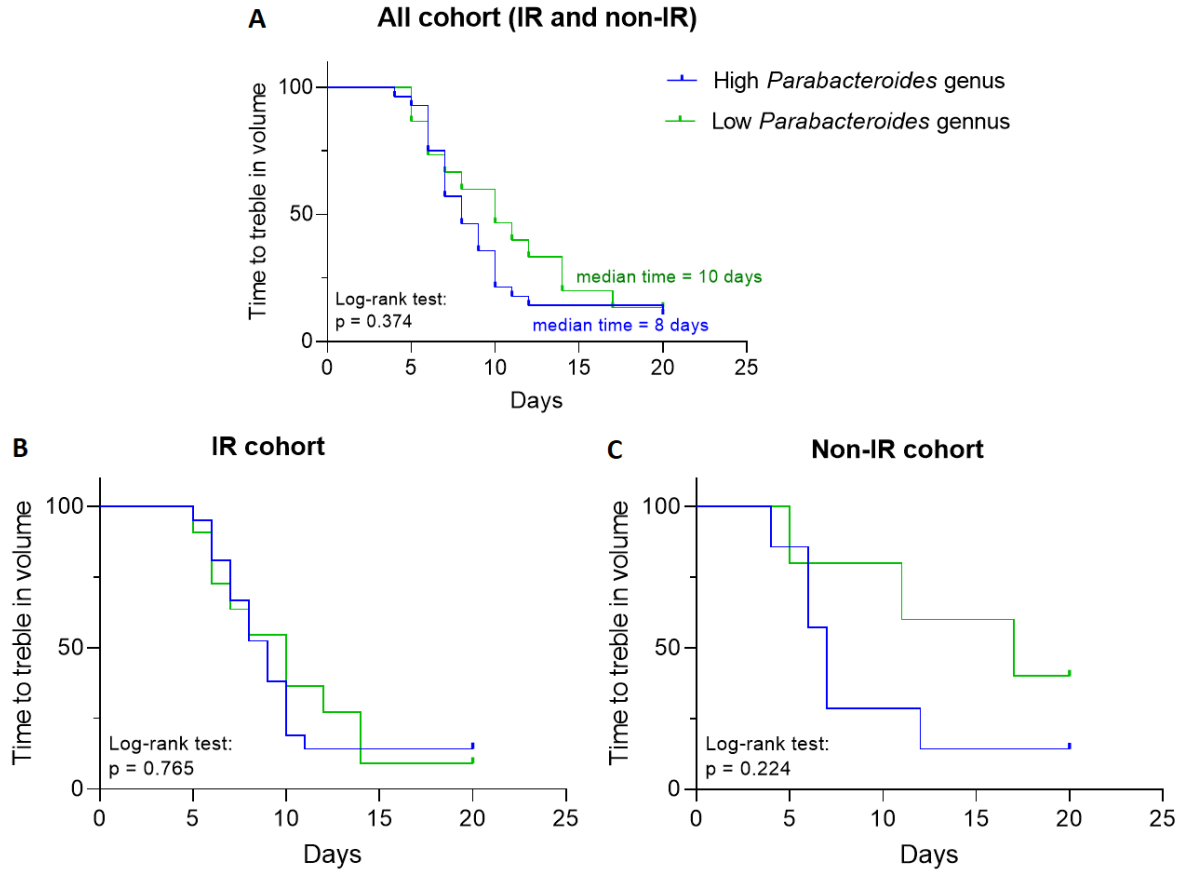


Figure 56. Kaplan-Meier (KM) plots of time for tumours to treble in volume based on *Parabacteroides* genus abundance in IR and Non-IR cohorts from different diet groups combined. Comparison KM plots by log-rank test in mice with high abundance (blue), or low abundance (green) of *Parabacteroides* genus in (A) all, (B) IR and (C) non-IR cohorts.

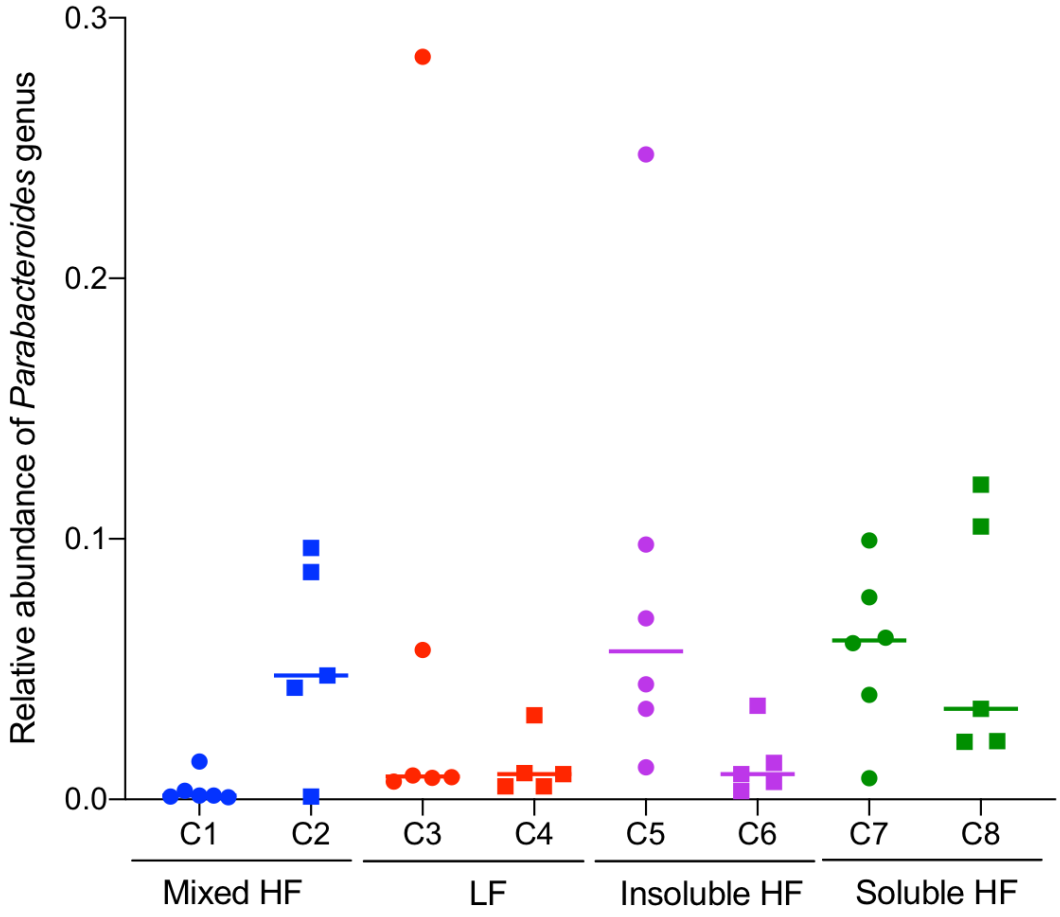


Figure 57. Relative abundance of *Parabacteroides* genus of the faecal microbiome in mice from different cages. Comparison of *Parabacteroides* genus abundance among mice from cage 1 (C1) to cage 8 (C8) by Kruskal-Wallis test.

Discussion

In 2016, Wei et al. demonstrated that up to four-fold-higher levels of butyrate accumulated in the lymphoma flank xenograft tumours of mice fed with high fibre diets. Butyrate has been proposed to increase histone acetylase inhibition, which is a mechanism of radiosensitisation. Therefore, the gut microbiota in mice fed a high fibre diet might act as an endogenous radiosensitiser, which provide a sufficiently high systemic butyrate level.

In this study, faecal and caecal microbiomes were investigated in mice fed with low fibre diets, low fibre diets with butyrate, high mixed fibre diets, high insoluble fibre diets and high soluble fibre diets, and profiles of both microbiomes were correlated with each other. Gut microbiome was shaped in mice fed with different modified diets for two weeks and a pattern was seen of homogenous gut microbiota in samples within groups. A more distinct bacterial taxon was seen in each group: enrichment of *B. acidifaciens* in soluble HF, *Parabacteroides* in LF, *Akkermansia muciniphila* in LF with butyrate, and *Lachnospiraceae* in mixed HF. To date, there is limited literature studying the effects of dietary fibre intake or the gut microbiota on tumour growth. Hardman et al. suggested that breast cancer xenografts grew more slowly in mice fed with fish oil concentrate than the tumours of control mice (Hardman et al., 2002). In contrast, Cougnoux et al. proposed that Colibactin-producing E. coli enhanced tumour growth of colon cancer xenografts (Cougnoux et al., 2014). No significant difference for the time to reach 100 mm^3 was seen in mice from different groups in this study. This demonstrated that diet and the diet-modified microbiome had no effect on tumour growth, to that point.

The tumours were irradiated when they reached a volume of 100 mm^3 and monitored until 700 mm^3. As the mice aged as the irradiated tumours reached 700 mm^3, all mice were enriched in a pattern from the S24-7 family. This indicated that their gut microbiomes had altered compared to those when the tumour volume was 100 mm^3. Although the gut microbiomes became more heterogeneous, a notable cluster effect still existed in samples within groups based on Bray-Curtis dissimilarity by principal coordinate analysis. Mice responding to irradiation in the soluble HF group were enriched with *B. acidifaciens* and non-responding mice were enriched with *Parabacteroides* genus. A predictive metagenomics study of the gut microbiome in responders was enriched with the carbohydrate metabolism pathways. This suggested that fermentation of carbohydrate, possibly soluble fibre, was

enriched in responders and its final products, short-chain fatty acids or butyrate, might contribute to the radiosensitisation. When the mice from different diet groups were pooled, *B. acidifaciens* abundance was positively correlated with survival time and mice with high *B. acidifaciens* had the longest median survival times using Kaplan-Meier survival analysis.

Early studies showed that dietary fibre, such as wheat bran, can be protective against colonic adenomas. Besides, systemic reviews have also provided some evidence that the dietary fibre intake is inversely associated with the risk of breast cancer (Aune et al., 2012), pancreatic cancer (C. H. Wang et al., 2015), prostate cancer (Deschasaux et al., 2014) and ovarian cancer (Huang et al., 2018). Most of the early studies are limited to epidemiology studies. Other healthy lifestyle factors, such as exercise, not smoking, abstention from alcohol, etc., rather than high fibre intake, could be major confounding biases in these studies. One of the major mechanisms of dietary fibre intake in protection against cancer is the modulation of the gut microbiota (Zeng et al., 2014), via short chain fatty acid production and immunomodulation. Dysbiosis has been linked to colorectal cancer, and a recent study proposed stage-distinct alterations in the tumourigenesis of colorectal cancer (Yachida et al., 2019); *Atopobium parvulum* and *Actinomyces odontolyticus* were significantly enriched in multiple polypoid adenomas and intramucosal carcinomas, and *Fusobacterium nucleatum* increased continuously from intramucosal carcinoma to more advanced stages. *Helicobacter pylori* is the most famous and early-defined link between bacteria and cancer (Uemura et al., 2001). Nowadays, *H. pylori* eradication has become one of the most promising treatment in the prevention of gastric cancer (Suzuki & Matsuzaki, 2018).

In this study, short-term high fibre intake alone and its modifying influence on the gut microbiota had no effect on tumour progression rate (to 100 mm^3). Gopalakrishnan et al. demonstrated that the gut microbiota can enhance melanoma responses to anti-PD-L1, possibly via enhanced systemic and anti-tumour immunity (Gopalakrishnan, Spencer, et al., 2018). Sivan et al. suggested that commensal *Bifidobacterium* are associated with facilitation of anti-PD-L1 efficacy, via enhanced cytotoxic T cell accumulation in the tumour microenvironment (Sivan et al., 2015). The immune-stimulatory effect of CTLA-4 blockage was found to be dependent on *B. thetaiotaomicron* or *B. fragilis* (Vetizou et al., 2015). All of these recent studies proposed that a modified gut microbiota can augment the efficacy of anti-tumoural treatment. However, most of them are limited to chemotherapy and

immunotherapy (Alexander et al., 2017). To date, this is the first study to suggest that a high soluble fibre diet, with the related modified gut microbiome, can act as a radiosensitiser.

B. acidifaciens abundance was enhanced by a high soluble fibre diet in this study and *B. acidifaciens* was identified as a potential radiosensitiser because its abundance was positively correlated with tumour response to irradiation and survival time in the IR cohort, analysis of different diets. This bacterium was first isolated in 2000 and got its name because the isolates reduced the pH level of peptone-yeast broth with Fildes' digest (Miyamoto & Itoh, 2000). Another previous study also revealed that *B. acidifaciens* was enriched in normal humans, compared to patients with inflammatory bowel disease (Ott et al., 2004). Yang et al. revealed that *B. acidifaciens* increased insulin sensitivity and further prevented obesity (Yang et al., 2017). However, the effect of this bacteria on tumour growth is still controversial. A study showed that increasing *B. acidifaciens* was seen in hepatocellular carcinoma which was induced by a streptozotocin-high fat diet (Xie et al., 2016). However, *B. acidifaciens* reduces isoflavones, a chemical that is associated with increased risk of breast cancer (Renouf & Hendrich, 2011). In contrast, *B. acidifaciens* was shown to contribute to the anti-tumour effect of medicinal Gynostemma saponins (L. Chen et al., 2015). In terms of the immune context, IL-6 and IL-10 were enhanced by *B. acidifaciens* via enhancement of MHC class II molecules expression and also co-stimulatory molecules (i.e., CD80 and CD86) on antigen presenting cells (Tsuda et al., 2007). *B. acidifaciens* has also been shown to promote IgA antibody production in the large intestine (Yanagibashi et al., 2009; Yanagibashi et al., 2013). Therefore, there are several possible mechanisms underlying the radiosensitising effects of high soluble fibre diets or B. acidifaciens: (1) enhancement of butyrate or other short-chain fatty acids in tumours, (2) suppress overgrowth of unfavourable bacteria, possibly inhibiting *Parabacteroides* genus, or (3) enhancement of anti-tumoural immunity.

In this study, two findings, namely, (1) the positive correlation between the time of culling and the *B. acidifaciens* abundance, and (2) the dose dependent effect of *B. acidifaciens* abundance on time for tumours to treble in volume, supported the hypothesis that the gut microbiota was a potential radiosensitiser. Environmental contamination is an inevitable issue in microbiome studies (Eisenhofer et al., 2019). To minimise the influence of contamination in this study, bacterial loads of samples were quantified and appropriate negative controls were included. Bacterial loads from luminal content and tissue samples

contained more than 10^4 CFUs which overrode the environmental bacteria communities. Furthermore, the bacterial compositions of the PBS negative controls were very different from those of the study groups, so the environmental microbiome was considered not to be a bias in this study. Inhibited immunity in the CD-1 nude mice was a potentially major limitation in this study because innate immunity plays a role in shaping the gut microbiota and adaptive immunity has a strong impact on tumour progression or treatment response.

Although a strong correlation between *B. acidifaciens* abundance and tumour response to irradiation was seen, further studies are still needed to validate a causal relationship between *B. acidifaciens* and radiosensitisation of bladder tumours. In collaboration with Professor Kevin Foster, I will assess the effect of *B. acidifaciens* supernatant on radiosensitivity of bladder cancer cell lines *in vitro* using clonogenic assays. Studying bacterial metabolites in supernatant is a more precise way of studying the effect of microbiota on treatment response (Figure 58). Besides, the soluble HF diet was assumed to act as a radiosensitiser via increasing butyrate levels and hence DNA damage in the tumour. I will validate the butyrate levels in tumours and plasma using HPLC, and DNA damage levels in tumours using a γH2AX assay. Mounting evidence suggests that the gut microbiota modulates tumour response to treatment (Gopalakrishnan, Helmink, et al., 2018; Gopalakrishnan, Spencer, et al., 2018), I will conduct *in vivo* studies to investigate the effects of *B. acidifaciens* supernatant orally (metabolites) on the immune context of nitrosamine (BBN)-induced bladder tumour development in K5.Stat3-transgenic mice.

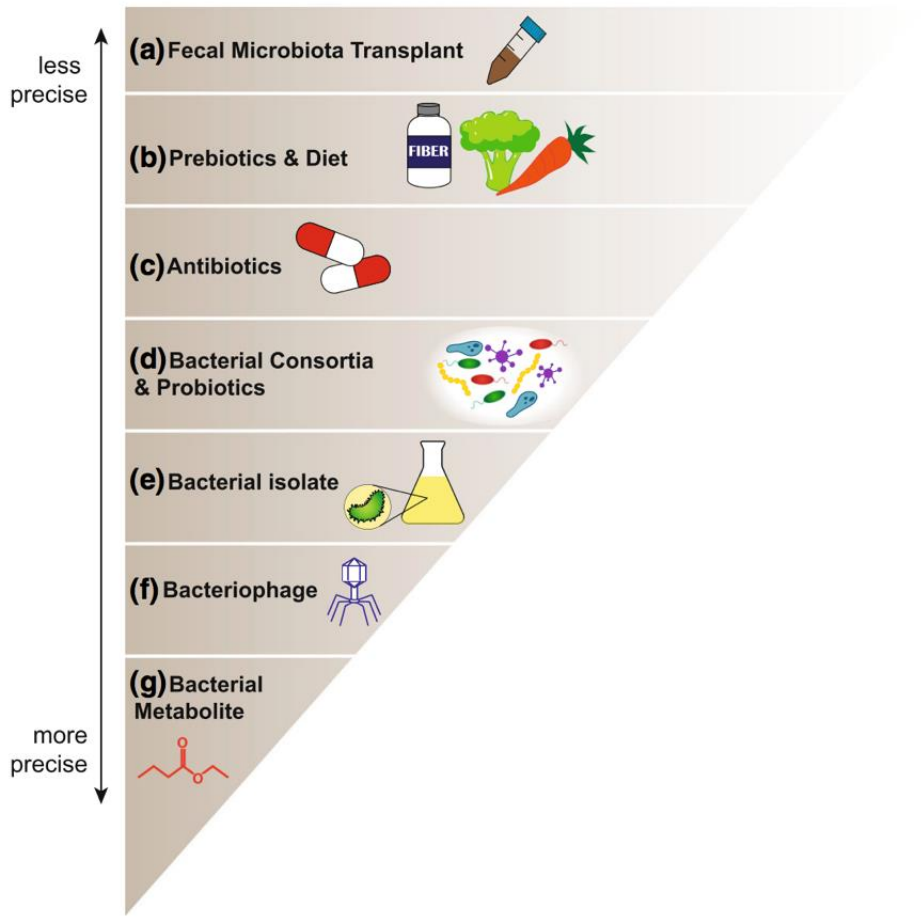


Figure 58. Different approaches in studying the microbiota-oriented interventions. (a) faecal microbiota transplantation (FMT), (b) modulating macronutrient or prebiotic intake, (c) targeting broad classes of bacteria with antibiotics, (d) administration of a select number of known beneficial bacterial species, (e) a single defined bacterial isolate, (f) bacteriophages or viruses that infect and kill selected bacteria, or (g) a single microbial metabolite. From Fessler et al. (2019).

Conclusions

I have demonstrated that a high soluble fibre diet significantly increased *B. acidifaciens* in the gut microbiota of CD-1 nude mice. In mice fed with a high soluble fibre diet, increased responses of RT112 xenografts to irradiation were seen and this phenotype was associated with higher relative abundance of *B. acidifaciens*. Metagenomic studies revealed functional differences in gut bacteria in responders, including enrichment of carbohydrate metabolism.

References

Alexander, Wilson, Teare, Marchesi, Nicholson, et al. (2017). Gut microbiota modulation of chemotherapy efficacy and toxicity. *Nat Rev Gastroenterol Hepatol, 14*(6), 356-365.

Aune, Chan, Greenwood, Vieira, Rosenblatt, et al. (2012). Dietary fiber and breast cancer risk: a systematic review and meta-analysis of prospective studies. *Ann Oncol, 23*(6), 1394-1402.

Backhed, Ley, Sonnenburg, Peterson, & Gordon. (2005). Host-bacterial mutualism in the human intestine. *Science, 307*(5717), 1915-1920.

Balvociute, & Huson. (2017). SILVA, RDP, Greengenes, NCBI and OTT - how do these taxonomies compare? *Bmc Genomics, 18*.

Baskar, Dai, Wenlong, Yeo, & Yeoh. (2014). Biological response of cancer cells to radiation treatment. *Front Mol Biosci, 1*, 24.

Bishehsari, Engen, Preite, Tuncil, Naqib, et al. (2018). Dietary Fiber Treatment Corrects the Composition of Gut Microbiota, Promotes SCFA Production, and Suppresses Colon Carcinogenesis. *Genes, 9*(2).

Bolden, Peart, & Johnstone. (2006). Anticancer activities of histone deacetylase inhibitors. *Nat Rev Drug Discov, 5*(9), 769-784.

Bolyen E. (2018). QIIME 2: Reproducible, interactive, scalable, and extensible microbiome data science. *PeerJ Preprints*, 6:e27295v27292.

Bourassa, Alim, Bultman, & Ratan. (2016). Butyrate, neuroepigenetics and the gut microbiome: Can a high fiber diet improve brain health? *Neuroscience Letters, 625*, 56-63.

Bray, Ferlay, Soerjomataram, Siegel, Torre, et al. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin, 68*(6), 394-424.

Caporaso, Lauber, Walters, Berg-Lyons, Huntley, et al. (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J, 6*(8), 1621-1624.

Chen, Tai, Brar, Leung, & Hsiao. (2015). Tumor grafting induces changes of gut microbiota in athymic nude mice in the presence and absence of medicinal Gynostemma saponins. *PLoS One, 10*(5), e0126807.

Chen, Zeng, Xu, Zheng, Liu, et al. (2016). Therapeutic effects of soluble dietary fiber consumption on type 2 diabetes mellitus. *Exp Ther Med, 12*(2), 1232-1242.

Choi, Czerniak, Ochoa, Su, Siefker-Radtke, et al. (2014). Intrinsic basal and luminal subtypes of muscle-invasive bladder cancer. *Nat Rev Urol, 11*(7), 400-410.

Choudhury, Swindell, Logue, Elliott, Livsey, et al. (2011). Phase II study of conformal hypofractionated radiotherapy with concurrent gemcitabine in muscle-invasive bladder cancer. *J Clin Oncol, 29*(6), 733-738.

Cougnoux, Dalmasso, Martinez, Buc, Delmas, et al. (2014). Bacterial genotoxin colibactin promotes colon tumour growth by inducing a senescence-associated secretory phenotype. *Gut, 63*(12), 1932-1942.

Damrauer, Hoadley, Chism, Fan, Tiganelli, et al. (2014). Intrinsic subtypes of high-grade bladder cancer reflect the hallmarks of breast cancer biology. *Proc Natl Acad Sci U S A, 111*(8), 3110-3115.

Davie. (2003). Inhibition of histone deacetylase activity by butyrate. *Journal of Nutrition, 133*(7 Suppl), 2485S-2493S.

de Vries, Miller, & Verbeke. (2015). Effects of cereal fiber on bowel function: A systematic review of intervention trials. *World J Gastroenterol, 21*(29), 8952-8963.

DeSantis, Hugenholtz, Larsen, Rojas, Brodie, et al. (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Applied and Environmental Microbiology, 72*(7), 5069-5072.

Deschasaux, Pouchieu, His, Hercberg, Latino-Martel, et al. (2014). Dietary total and insoluble fiber intakes are inversely associated with prostate cancer risk. *Journal of Nutrition, 144*(4), 504-510.

Dhingra, Michael, Rajput, & Patil. (2012). Dietary fibre in foods: a review. *J Food Sci Technol, 49*(3), 255-266.

Donohoe, Collins, Wali, Bigler, Sun, et al. (2012). The Warburg effect dictates the mechanism of butyrate-mediated histone acetylation and cell proliferation. *Mol Cell, 48*(4), 612-626.

Eisenhofer, Minich, Marotz, Cooper, Knight, et al. (2019). Contamination in Low Microbial Biomass Microbiome Studies: Issues and Recommendations. *Trends Microbiol, 27*(2), 105-117.

Erlich, & Zlotta. (2016). Treatment of bladder cancer in the elderly. *Investig Clin Urol, 57 Suppl 1*, S26-35.

Fessler, Matson, & Gajewski. (2019). Exploring the emerging role of the microbiome in cancer immunotherapy. *J Immunother Cancer, 7*(1), 108.

Gopalakrishnan, Helmink, Spencer, Reuben, & Wargo. (2018). The Influence of the Gut Microbiome on Cancer, Immunity, and Cancer Immunotherapy. *Cancer Cell, 33*(4), 570-580.

Gopalakrishnan, Spencer, Nezi, Reuben, Andrews, et al. (2018). Gut microbiome modulates response to anti-PD-1 immunotherapy in melanoma patients. *Science, 359*(6371), 97-103.

Groselj, Ruan, Scott, Gorrill, Nicholson, et al. (2018). Radiosensitization In Vivo by Histone Deacetylase Inhibition with No Increase in Early Normal Tissue Radiation Toxicity. *Mol Cancer Ther, 17*(2), 381-392.

Groselj, Sharma, Hamdy, Kerr, & Kiltie. (2013). Histone deacetylase inhibitors as radiosensitisers: effects on DNA damage signalling and repair. *Br J Cancer, 108*(4), 748-754.

Grossman, Natale, Tangen, Speights, Vogelzang, et al. (2003). Neoadjuvant chemotherapy plus cystectomy compared with cystectomy alone for locally advanced bladder cancer. *N Engl J Med, 349*(9), 859-866.

Guancial, Roussel, Bergsma, Bylund, Sahasrabudhe, et al. (2015). Bladder cancer in the elderly patient: challenges and solutions. *Clin Interv Aging, 10*, 939-949.

Gunness, & Gidley. (2010). Mechanisms underlying the cholesterol-lowering properties of soluble dietary fibre polysaccharides. *Food Funct, 1*(2), 149-155.

Hall, & Beiko. (2018). 16S rRNA Gene Analysis with QIIME2. *Methods Mol Biol, 1849*, 113-129.

Hardman, Munoz, & Cameron. (2002). Role of lipid peroxidation and antioxidant enzymes in omega 3 fatty acids induced suppression of breast cancer xenograft growth in mice. *Cancer Cell Int, 2*(1), 10.

Hoskin, Rojas, Bentzen, & Saunders. (2010). Radiotherapy with concurrent carbogen and nicotinamide in bladder carcinoma. *J Clin Oncol, 28*(33), 4912-4918.

Huang, Wang, Shang, Lin, Yang, et al. (2018). Association between dietary fiber intake and risk of ovarian cancer: a meta-analysis of observational studies. *J Int Med Res, 46*(10), 3995-4005.

Human Microbiome Project. (2012a). A framework for human microbiome research. *Nature, 486*(7402), 215-221.

Human Microbiome Project. (2012b). Structure, function and diversity of the healthy human microbiome. *Nature, 486*(7402), 207-214.

James, Hussain, Hall, Jenkins, Tremlett, et al. (2012). Radiotherapy with or without chemotherapy in muscle-invasive bladder cancer. *N Engl J Med, 366*(16), 1477-1488.

Jiang, Amir, Morton, Heller, Arias-Castro, et al. (2017). Discrete False-Discovery Rate Improves Identification of Differentially Abundant Microbes. *mSystems, 2*(6).

Kanehisa, Goto, Kawashima, Okuno, & Hattori. (2004). The KEGG resource for deciphering the genome. *Nucleic Acids Research, 32*, D277-D280.

Kotwal, Choudhury, Johnston, Paul, Whelan, et al. (2008). Similar treatment outcomes for radical cystectomy and radical radiotherapy in invasive bladder cancer treated at a United Kingdom specialist treatment center. *Int J Radiat Oncol Biol Phys, 70*(2), 456-463.

Langille, Zaneveld, Caporaso, McDonald, Knights, et al. (2013). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nature Biotechnology, 31*(9), 814-+.

Langille, Zaneveld, Caporaso, McDonald, Knights, et al. (2013). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat Biotechnol, 31*(9), 814-821.

Lans, Marteijn, & Vermeulen. (2012). ATP-dependent chromatin remodeling in the DNA-damage response. *Epigenetics Chromatin, 5*, 4.

Lattimer, & Haub. (2010). Effects of dietary fiber and its components on metabolic health. *Nutrients, 2*(12), 1266-1289.

Li, & Seto. (2016). HDACs and HDAC Inhibitors in Cancer Development and Therapy. *Cold Spring Harb Perspect Med, 6*(10).

Lozupone, Stombaugh, Gordon, Jansson, & Knight. (2012). Diversity, stability and resilience of the human gut microbiota. *Nature, 489*(7415), 220-230.

Massari, Ciccarese, Santoni, Iacovelli, Mazzucchelli, et al. (2016). Metabolic phenotype of bladder cancer. *Cancer Treat Rev, 45*, 46-57.

Meng, Li, Rao, Wang, & Fu. (2017). A Chinese Herbal Formula, Gengnianchun, Ameliorates beta-Amyloid Peptide Toxicity in a Caenorhabditis elegans Model of Alzheimer's Disease. *Evid Based Complement Alternat Med, 2017*, 7480980.

Miyamoto, & Itoh. (2000). Bacteroides acidifaciens sp. nov., isolated from the caecum of mice. *Int J Syst Evol Microbiol, 50 Pt 1*, 145-148.

Mohamed, Diefenbach, Goltz, Lee, Latini, et al. (2012). Muscle invasive bladder cancer: from diagnosis to survivorship. *Adv Urol, 2012*, 142135.

NCBI.

Nearing, Douglas, Comeau, & Langille. (2018). Denoising the Denoisers: an independent evaluation of microbiome sequence error-correction approaches. *Peerj, 6*.

Nicholson, Jevons, Groselj, Ellermann, Konietzny, et al. (2017). E3 Ligase cIAP2 Mediates Downregulation of MRE11 and Radiosensitization in Response to HDAC Inhibition in Bladder Cancer. *Cancer Res, 77*(11), 3027-3039.

Niegisch, Knievel, Koch, Hader, Fischer, et al. (2013). Changes in histone deacetylase (HDAC) expression patterns and activity of HDAC inhibitors in urothelial cancers. *Urol Oncol, 31*(8), 1770-1779.

Okumura, & Takeda. (2018). Maintenance of intestinal homeostasis by mucosal barriers. *Inflamm Regen, 38*, 5.

Ott, Musfeldt, Wenderoth, Hampe, Brant, et al. (2004). Reduction in diversity of the colonic mucosa associated bacterial microflora in patients with active inflammatory bowel disease. *Gut, 53*(5), 685-693.

Ozawa, Tanji, Kikugawa, Sasaki, Yanagihara, et al. (2010). Inhibition of bladder tumour growth by histone deacetylase inhibitor. *BJU Int, 105*(8), 1181-1186.

Pang, Vogensen, Nielsen, & Hansen. (2012). Faecal and caecal microbiota profiles of mice do not cluster in the same way. *Lab Anim, 46*(3), 231-236.

Ramirez, Gupta, Canter, Harrow, Dobbs, et al. (2016). Microscopic haematuria at time of diagnosis is associated with lower disease stage in patients with newly diagnosed bladder cancer. *BJU Int, 117*(5), 783-786.

Renouf, & Hendrich. (2011). Bacteroides uniformis is a putative bacterial species associated with the degradation of the isoflavone genistein in human feces. *Journal of Nutrition, 141*(6), 1120-1126.

Riviere, Selak, Lantin, Leroy, & De Vuyst. (2016). Bifidobacteria and Butyrate-Producing Colon Bacteria: Importance and Strategies for Their Stimulation in the Human Gut. *Front Microbiol, 7*, 979.

Robertson, Kim, Al-Ahmadie, Bellmunt, Guo, et al. (2018). Comprehensive Molecular Characterization of Muscle-Invasive Bladder Cancer. *Cell, 174*(4), 1033.

Roediger. (1980). Role of anaerobic bacteria in the metabolic welfare of the colonic mucosa in man. *Gut, 21*(9), 793-798.

Sanli, Dobruch, Knowles, Burger, Alemozaffar, et al. (2017). Bladder cancer. *Nat Rev Dis Primers, 3*, 17022.

Savage. (1977). Microbial ecology of the gastrointestinal tract. *Annu Rev Microbiol, 31*, 107-133.

Segata, Izard, Waldron, Gevers, Miropolsky, et al. (2011a). Metagenomic biomarker discovery and explanation. *Genome Biol, 12*(6), R60.

Segata, Izard, Waldron, Gevers, Miropolsky, et al. (2011b). Metagenomic biomarker discovery and explanation. *Genome Biology, 12*(6).

Sivan, Corrales, Hubert, Williams, Aquino-Michaels, et al. (2015). Commensal Bifidobacterium promotes antitumor immunity and facilitates anti-PD-L1 efficacy. *Science, 350*(6264), 1084-1089.

Stein, Lieskovsky, Cote, Groshen, Feng, et al. (2001). Radical cystectomy in the treatment of invasive bladder cancer: long-term results in 1,054 patients. *J Clin Oncol, 19*(3), 666-675.

Sternberg, Bellmunt, Sonpavde, Siefker-Radtke, Stadler, et al. (2013). ICUD-EAU International Consultation on Bladder Cancer 2012: Chemotherapy for urothelial carcinoma-neoadjuvant and adjuvant settings. *Eur Urol, 63*(1), 58-66.

Suzuki, & Matsuzaki. (2018). Gastric cancer: evidence boosts Helicobacter pylori eradication. *Nat Rev Gastroenterol Hepatol, 15*(8), 458-460.

Tsuda, Hosono, Yanagibashi, Hachimura, Hirayama, et al. (2007). Prior stimulation of antigen-presenting cells with Lactobacillus regulates excessive antigen-specific cytokine responses in vitro when compared with Bacteroides. *Cytotechnology, 55*(2-3), 89-101.

Uemura, Okamoto, Yamamoto, Matsumura, Yamaguchi, et al. (2001). Helicobacter pylori infection and the development of gastric cancer. *N Engl J Med, 345*(11), 784-789.

Vetizou, Pitt, Daillere, Lepage, Waldschmitt, et al. (2015). Anticancer immunotherapy by CTLA-4 blockade relies on the gut microbiota. *Science, 350*(6264), 1079-1084.

Wang, Du, Li, Zhang, Zhang, et al. (2014). Comparative analysis of microbiome between accurately identified 16S rDNA and quantified bacteria in simulated samples. *Journal of Medical Microbiology, 63*, 433-440.

Wang, Qiao, Wang, & Zhou. (2015). Dietary fiber intake and pancreatic cancer risk: a meta-analysis of epidemiologic studies. *Sci Rep, 5*, 10834.

Wei, Sun, Yu, Yang, & Ai. (2016). Butyrate production from high-fiber diet protects against lymphoma tumor. *Leuk Lymphoma, 57*(10), 2401-2408.

Weickert, & Pfeiffer. (2008). Metabolic effects of dietary fiber consumption and prevention of diabetes. *J Nutr, 138*(3), 439-442.

Weickert, & Pfeiffer. (2018). Impact of Dietary Fiber Consumption on Insulin Resistance and the Prevention of Type 2 Diabetes. *Journal of Nutrition, 148*(1), 7-12.

Whitman, Coleman, & Wiebe. (1998). Prokaryotes: The unseen majority. *Proceedings of the National Academy of Sciences of the United States of America, 95*(12), 6578-6583.

Xie, Wang, Liu, Wei, Chen, et al. (2016). Distinctly altered gut microbiota in the progression of liver disease. *Oncotarget, 7*(15), 19355-19366.

Yachida, Mizutani, Shiroma, Shiba, Nakajima, et al. (2019). Metagenomic and metabolomic analyses reveal distinct stage-specific phenotypes of the gut microbiota in colorectal cancer. *Nat Med, 25*(6), 968-976.

Yanagibashi, Hosono, Oyama, Tsuda, Hachimura, et al. (2009). Bacteroides induce higher IgA production than Lactobacillus by increasing activation-induced cytidine deaminase expression in B cells in murine Peyer's patches. *Biosci Biotechnol Biochem, 73*(2), 372-377.

Yanagibashi, Hosono, Oyama, Tsuda, Suzuki, et al. (2013). IgA production in the large intestine is modulated by a different mechanism than in the small intestine: Bacteroides acidifaciens promotes IgA production in the large intestine by inducing germinal center formation and increasing the number of IgA+ B cells. *Immunobiology, 218*(4), 645-651.

Yang, Lee, Kim, Lee, Ryu, et al. (2017). Gut commensal Bacteroides acidifaciens prevents obesity and improves insulin sensitivity in mice. *Mucosal Immunol, 10*(1), 104-116.

Yarza, Yilmaz, Pruesse, Glockner, Ludwig, et al. (2014). Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat Rev Microbiol, 12*(9), 635-645.

Zeng, Lazarova, & Bordonaro. (2014). Mechanisms linking dietary fiber, gut microbiota and colon cancer prevention. *World J Gastrointest Oncol, 6*(2), 41-51.

Appendix

Table A1. Rodent diets used in the study with varying levels of cellulose or inulin per 4000 kcal.

	Low fibre		High insoluble fibre		High soluble fibre		High mixed fibre	
	2 gm Cellulose/3850kcal		100 gm Cellulose/3850kcal		100 gm Inulin/3850kcal		50 gm Cellulose + 50 gm Inulin/3850kcal	
	gm%	kcal%	gm%	kcal%	gm%	kcal%	gm%	kcal%
Protein	21	20	19	20	20	20	20	20
Carbohydrate	67	64	61	64	59	60	60	62
Fat	7	16	7	16	7	16	7	16
Total		100		100		96		96
kcal/gm	4.20		3.81		3.95		3.88	
Ingredient	**gm**	**kcal**	**gm**	**kcal**	**gm**	**kcal**	**gm**	**kcal**
Casein	200	800	200	800	200	800	200	800
L-Cystein	3	12	3	12	3	12	3	12
Corn starch	397.486	1590	397.486	1590	359.986	1440	378.736	1515
Maltodextrin 10	132	528	132	528	132	528	132	528
Sucrose	100	400	100	400	100	400	100	400
Cellulose, BM200	2	0	100	0	0	0	50	0
Inulin	0	0	0	0	100	0	50	75
Soybean oil	70	630	70	630	70	630	70	630
t-Butylhydroquinone	0.014	0	0.014	0	0.014	0	0.014	0
Mineral mix S10022G	35	0	35	0	35	0	35	0
Vitamin mix V10037	10	40	10	40	10	40	10	40
Choline bitartrate	2.5	0	2.5	0	2.5	0	2.5	0
Total	**952**	**4000**	**1050**	**4000**	**1012.5**	**4000**	**1031.25**	**4000**
Total Cellulose (gm/kg diet)	2.1		95.2		0		48.5	
Inulin (gm/kg diet)	0		0		98.8		48.5	

These diets were formulated by Jia-Yu Ke at Research Diets, Inc. on 11/6/2018.

Table A2. Bacterial composition in the faecal microbiome when the tumours reached 100 mm^3.

	Phylum	Class	Order	Family	Genus	Species
	Bacteroidetes	Bacteroidia	Bacteroidales	Bacteroidaceae	Bacteroides	acidifaciens
	Bacteroidetes	Bacteroidia	Bacteroidales	Porphyromonadaceae	Parabacteroides	
	Verrucomicrobia	Verrucomicrobiae	Verrucomicrobiales	Verrucomicrobiaceae	Akkermansia	muciniphila
	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae		
	Bacteroidetes	Bacteroidia	Bacteroidales	S24-7		
	Firmicutes	Clostridia	Clostridiales			
	Firmicutes	Erysipelotrichi	Erysipelotrichales	Erysipelotrichaceae	Clostridium	cocleatum
	Firmicutes	Clostridia	Clostridiales			
	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae	Oscillospira	
	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae	Anaerotruncus	
	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	
	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae	Ruminococcus	gnavus

Potentially butyrate-producing families are highlighted.

Table A3. Bacterial composition in the faecal microbiome when the tumours reached 700 mm^3.

	Phylum	Class	Order	Family	Genus	Species
	Bacteroidetes	Bacteroidia	Bacteroidales	Bacteroidaceae	Bacteroides	acidifaciens
	Bacteroidetes	Bacteroidia	Bacteroidales	S24-7		
	Verrucomicrobia	Verrucomicrobiae	Verrucomicrobiales	Verrucomicrobiaceae	Akkermansia	muciniphila
	Bacteroidetes	Bacteroidia	Bacteroidales	Bacteroidaceae	Bacteroides	
	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae		
	Firmicutes	Clostridia	Clostridiales			
	Bacteroidetes	Bacteroidia	Bacteroidales	Porphyromonadaceae	Parabacteroides	
	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae	Oscillospira	
	Bacteroidetes	Bacteroidia	Bacteroidales	Rikenellaceae		
	Firmicutes	Clostridia	Clostridiales			
	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae	Anaerotruncus	
	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae	Ruminococcus	

Potentially butyrate-producing families are highlighted.

Table A4. Information on mouse diets, cages, B. acidifaciens abundance and time of culling.

Diet	**Low fibre**										
Cage	3						4				
ID	C3.3	C3.1	C3.10	C3.30	C3.NM	C3.4	C4.10	C4.1	C4.NM	C4.30	C4.3
IR/Non-IR	Non-IR	IR	IR	IR	IR	Non-IR	Non-IR	IR	IR	IR	IR
B.acidifaciens abundance	0.164145	0.00019252	0.06690705	0.18902547	0.07300095	0.07087815	0.07545676	0.00037169	0.053362	0.00095565	0.00062727
Parabacteroides genus abundance	0.00914106	0.28499005	0.05735844	0.00829265	0.00856735	0.00690319	0.00503422	0.0323372	0.00501694	0.0100624	0.00968776
Time for tumour to treble in volume	17	6	6	6	14	5	20	9	8	7	8
Time to culling (days)	28	12	12	45	35	10	45	21	42	14	14

Diet	**Mixed high fibre**										
Cage	1						2				
ID	C1.NM	C1.10	C1.3	C1.1	C1.30	C1.4	C2.10	C2.NM	C2.3	C2.30	C2.1
IR/Non-IR	IR	IR	IR	Non-IR	Non-IR	IR	IR	IR	IR	Non-IR	IR
B.acidifaciens abundance	0.04540676	0.03044898	0.04655061	0.00935197	0.06229295	0.00622417	0.00076678	0.00042641	0.000356	0.01072546	0.0602382
Parabacteroides genus abundance	0.00078854	0.00326635	0.00116054	0.01453621	0.0015381	0.00150742	0.08723605	0.04763645	0.04282154	0.09647947	0.00114521
Time for tumour to treble in volume	6	7	10	4	20	5	5	9	7	7	14
Time to culling (days)	14	19	31	10	38	12	10	17	17	19	26

Diet	**Insoluble high fibre**										
Cage	5						6				
ID	C5.3	C5.30	C5.1	C5.10	C5.NM	C5.4	C6.NM	C6.1	C6.3	C6.30	C6.10
IR/Non-IR	Non-IR	IR	Non-IR	IR	IR	IR	IR	IR	IR	IR	Non-IR
B.acidifaciens abundance	0.06019205	0.0808974	0.02231957	0.0950628	0.00051919	0.09485331	0.01921414	0.02348066	0.07808474	0.11586863	0.02449694
Parabacteroides genus abundance	0.09782184	0.06949722	0.01230705	0.0347914	0.24760466	0.04419443	0.01402447	0.03597438	0.00672774	0.00966851	0.0031871
Time for tumour to treble in volume	6	8	20	8	6	10	10	10	10		11
Time to culling (days)	17	17	42	19	10	17	19	17	21	45	19

Diet	**Soluble high fibre**										
Cage	7						8				
ID	C7.1	C7.10	C7.3	C7.NM	C7.30	C7.4	C8.NM	C8.1	C8.10	C8.3	C8.30
IR/Non-IR	Non-IR	IR	IR	IR	IR	IR	IR	Non-IR	IR	IR	Non-IR
B.acidifaciens abundance	0.00044366	0.00047901	0.00019435	0.00067822	0.33135489	0	0.22149428	0.00048807	0.25171378	0.15695771	0.00027062
Parabacteroides genus abundance	0.06002028	0.07760014	0.09937808	0.06201959	0.00811843	0.04020364	0.0347429	0.10482646	0.02209507	0.02237381	0.12076314
Time for tumour to treble in volume	6	10	11	9	12	7	20	7	20	20	12
Time to culling (days)	10	17	24	17	42	14	38	21	33	42	26

Glossary

Abundance	A measure of bacterial numbers and mass per unit volume.
Alignment	The comparison between sequences (nucleic acids or amino acids); the result will be some common letters within their sequence structures.
Alpha diversity	A measure of similarity within communities.
Alpha rarefraction	Primarily used to determine if the richness of the samples has been fully observed or sequenced.
Beta diversity	A measure of similarity between communities.
Bray-Curtis dissimilarity	Fraction of overabundant counts.
Chimera	A first source of error originates from chimera formation within the PCR amplification step, thereby creating a chimeric sequence which consists of two or more segments from distinct species.
Cladogram	An evolutionary tree that plots evolutionary relationships among organisms and branch points show different relationships in time.
CFU	A unit used to estimate the number of viable bacteria or fungal cells in a sample. Viable is defined as the ability to multiply via binary fission under the controlled conditions.
Clustering	Classification, or putting samples into (perhaps hierarchical) classes, is often useful when one wishes to assign names to, or to map, ecological communities.
Deblur	Deblur uses sequence error profiles to associate erroneous sequence reads with the true biological sequence from which they are derived, resulting in high quality sequence variant data.
Depth	The number of unique reads that include a given nucleotide in the reconstructed sequence.
Diversity	A measure of the number of species in a community, and a measure of the abundance of each species.
DS-FDR	A method of understanding the rate of type I errors in null hypothesis testing when performing multiple comparisons. It greatly improves the power to detect differential taxa by exploiting the discreteness of the data.
Effect size	A quantitative measure of the magnitude of a phenomenon.
Evenness	A description of the distribution of abundance across the species in a community.
Faith's phylogenetic diversity	An alpha diversity metric that uses a phylogenetic tree to compute sample diversity.
Greengenes	A 16S rRNA gene database which addresses the limitations of public repositories by providing chimera screening,

	standard alignment, and taxonomic classification using multiple published taxonomies.
KEGG	A database resource for understanding high-level functions and utilities of biological systems, such as the cell, the organism and the ecosystem, from molecular-level information, especially large-scale molecular datasets generated by genome sequencing and other high-throughput experimental technologies.
LEfSe	A method for determining the features (organisms, clades, operational taxonomic units, genes, or functions) most likely to explain differences between classes by coupling standard tests for statistical significance with additional tests encoding biological consistency and effect relevance.
Library	A collection of DNA fragments with known adapter sequences at one or both ends that is derived from a single clinical or environmental sample.
Metagenomics	A molecular tool used to analyse DNA acquired from environmental samples, in order to study the community of micro-organisms present, without the necessity of obtaining pure cultures.
Microbiome	A collective term for all of the genetic material within a microbiota and this can also be referred to as the metagenome of the microbiota.
Microbiota	A collective term for the micro-organisms that live in or on the human body.
MiSeq	A DNA-to-data sequencing platform, integrating cluster generation, amplification, sequencing, and data analysis into a single instrument.
OTU	A cluster of similar sequence variants of the 16S rDNA marker gene sequence. Each of these clusters is intended to represent a taxonomic unit of a bacteria species or genus depending on the sequence similarity threshold.
PCoA	A method of representing the distances between samples in a low-dimensional, Euclidean space. In particular, it maximizes the linear correlation between the distances in the distance matrix, and the distances in a space of low dimension (typically, 2 or 3 axes are selected).
Phylogenetics	A study of the evolutionary history and relationships among individuals or groups of organisms (e.g. species, or populations). These relationships are discovered through phylogenetic inference methods that evaluate observed heritable traits, such as DNA sequences or morphology under a model of evolution of these traits. The result of these analyses is a phylogeny (also known as a phylogenetic tree).
Phylogeny	The history of descent of a group of taxa such as species from their common ancestors, including the order of

	branching and sometimes the absolute times of divergence; also applied to the genealogy of genes derived from a common ancestral gene.
PICRUSt	A bioinformatics software package designed to predict metagenome functional content from marker gene (e.g., 16S rRNA) surveys and full genomes.
Pielou's evenness	An alpha diversity metric that measures relative evenness of species richness.
Plugin	A method providing microbiome (i.e. domain-specific) analysis functionality that is accessible to users through a variety of interfaces built around the QIIME 2 framework.
Q-score	A quality control metric for DNA sequencing that is logarithmically related to the base calling error probabilities and serves as a measurement of read accuracy.
Rarefaction	A resampling approach to develop a curve to identify and allow comparisons among samples using the minimum sample size of all the collections.
Reads	Inferred sequences of base pairs corresponding to part of or all of a single DNA fragment.
Relative abundance	The percent composition of an organism of a particular kind relative to the total number of organisms in the area.
Richness	The number of species in a community.
Sequence alignment	Regions of local similarity of DNA, RNA or amino acid sequences.
Shannon's index	An alpha diversity metric that accounts for both abundance and evenness of the taxa present.
Taxon	A group of one or more populations of an organism or organisms seen by taxonomists to form a unit.
Taxonomy	The science of naming, defining (circumscribing) and classifying groups of biological organisms on the basis of shared characteristics. Organisms are grouped together into taxa (singular: taxon) and these groups are given a taxonomic rank; groups of a given rank can be aggregated to form a super-group of higher rank, thus creating a taxonomic hierarchy. The principal ranks in modern use are domain, kingdom, phylum (division is sometimes used in botany in place of phylum), class, order, family, genus, and species.
Volcano plot	A type of scatter-plot that is used to quickly identify changes in large data sets composed of replicate data.