Supplementary material 1: Supplementary Methods

Ethical approval

Ethical approval for the study was granted by the clinical research ethics committee of the Cork teaching hospitals under project code APC073. This study was carried out in accordance with the ethical standards of the research committee and with the 1964 Helsinki declaration and its later amendments. All subjects gave written informed consent prior to the beginning of the study (see supplementary materials 2). The authors declare no conflict of interest.

Experimental design

Irish athletes preparing for and/or participating in major international competition (including the 2016 Summer Olympics) were recruited. Faecal and urine samples were collected from female (N=14) and male (N=23) athletes from 16 different sports (field events (N=1), boxing (N=1), cycling (N=1), fencing (N=2), hockey (N=10), judo (N=2), marathon (N=3), Paralympic cycling (N=2), Paralympic swimming (N=2), pentathlon (N=1), race walking (N=2), rowing (N=2), middle distance running (N=4), swimming (N=2), taekwondo (N=1) and triathlon (N=1); see supplementary materials 3 for participant characteristics summary). All samples were stored initially on a short term basis at -20°C before storage at -80°C pre-processing. The mean total storage time of faecal samples (between initial storage and DNA extraction) was 15(±9) days. The mean age of the athletes was 27 (±5) years. Gut health questionnaires (see supplementary materials 4 for questionnaire and supplementary materials 5 for responses) and food frequency questionnaires (see supplementary materials 6 for questionnaire and supplementary materials 7 for responses) were completed by all participants. Food frequency questionnaires were converted to nutrient intake values using FETA (Food Frequency Questionnaire European Prospective Investigation into Cancer and Nutrition Tool for Analysis) [1]. A subset of samples (N=14) were collected during the 2016 Summer Olympics in Rio de Janeiro, Brazil. Two individuals (participant OM008 and OM033) provided samples at more than one time point (2 and 3 time points, respectively).

DNA extraction from faecal samples

Faecal samples were homogenised and DNA extraction was completed using the repeated bead beating plus column (RBBC) method, as previously described [2]. Briefly, 0.25g of faecal sample was added into a sterile 2ml screw-cap tube containing 0.25g of zirconia beads (0.25g of 0.1mm, 0.125g of 1.5mm and a single 2.5mm bead). Following the addition of 1ml of lysis buffer (500mM NaCl, 50mM Tris-HCl, 50mM EDTA and 4% sodium dodecyl sulphate (SDS)) samples were homogenized for 3 minutes on a Mini Beadbeater (BioSpec, Bartlesville, USA). Samples were incubated at 70°C for 15 minutes to lyse the cells. Following centrifugation (16,000xg; 5 minutes; 4°C) supernatant was removed and the bead beating steps were repeated, this time using 200µl of lysis buffer. Supernatant from both bead beating steps was pooled and 10M ammonium acetate was added. Samples were centrifuged (16,000xg; 10 minutes; 4°C) and resulting supernatant was treated with one volume of isopropanol. DNA was pelleted and washed with 70% ethanol before being dissolved in Tris-EDTA. DNA was RNase and proteinase K treated. DNA was finally washed with 100% ethanol and buffers AW1 and AW2 (QIAamp DNA stool mini kit, Qiagen, England) before being eluted in 200µl AE buffer (Qiagen). DNA was stored at -20°C before library preparation.

Library preparation

Samples were prepared for Illumina NextSeq shotgun sequencing as per modified Nextera XT DNA library preparation protocol (Illumina, California, USA). Briefly, 0.2ng/µl of sample DNA was tagmented for 7.5 minutes at 55°C. Tagmented DNA was amplified and index adapters (i5 and i7) added. Agencourt AMPure beads (Beckman Coulter, Clare, Ireland) were used to clean up prepared library DNA. DNA samples were diluted to equimolar concentrations before pooling. Prepared library was sequenced on the NextSeq platform using standard Illumina protocols at the Teagasc sequencing facility. The datasets generated and analysed during the current study are available in the European Nucleotide Archive under the study accession number PRJEB32794.

Bioinformatic analysis

Resulting FastQ reads were quality checked by first removing contaminating human derived reads using NCBI Best Match Tagger (BMTagger). Resulting reads were trimmed and poor quality and duplicate reads removed using a combination of Picard and SAM tools [3, 4]. Taxonomic classifications of trimmed reads were determined using Kraken 2 and Bracken [5, 6]. Functional profiling was completed using the Human Microbiome Project Unified Metabolic Analysis Network 2 (HUMAnN2; http://huttenhower.sph.harvard.edu/humann2)[7, 8].

Analysis of resulting taxonomic and functional tables was completed in R (R Foundation for Statistical Computing, Vienna, 2012) (version 3.5.1). Sports were classified into sports classification groups (SCGs), as previously outlined (see figure 1.A)[9]. SCGs used in this study were established based on the dynamic and static components of exercise involved within an individual sport. A highly dynamic exercise involves changes in muscle length and joint movement, which is associated with increased oxygen uptake resulting in increased cardiac output. A highly static exercise involves the development of intramuscular strength with minimal changes in muscular length or joint movement. Static exercise is associated with increased voluntary contraction which results in an increased blood pressure load [9]. Statistical analysis was completed using SCGs with groups B1 and B2 being grouped as B (moderate dynamic component) as only 1 individual could be classified to SCG B2. Data was also analysed for potential confounding factors including sex and location of sample collection, where information was available. Alpha and beta diversities were computed in R using vegan (version 2.5-2) [10]. A hclust tree was created in the R package ape (version 5.1) and visualised using iToL [11, 12]. LEfSe was used to determine species and pathways which characterise specific groups [13]. LEfSe allows for the identification of pathways or taxa which characterise groups. The significance threshold was set at 0.05 and an LDA effect size threshold of 3 was used for discriminative features. Spearman correlations were calculated in R using the RcmdrMisc package (version 1.0-10) and corrected for multiple comparisons using Holms method. A heatmap was created to visualise functional pathways from HUMAnN2 using the pheatmap package in R.

Metabolomics analysis

Urine samples underwent GC-MS, UPLC-MS and ^{1}H-NMR analysis, while faecal samples underwent ^{1}H-NMR and UPLC-MS analysis in Imperial College London, as previously described [14-16].

<u>^{1}H-NMR spectroscopy metabolic profiling analysis</u>

Briefly, urine and faecal water samples were prepared with a pH 7.4 phosphate buffer for ^{1}H-NMR spectroscopy as described previously by Garcia-Perez et al [17]. Samples were analysed at 300K on a 600 MHz spectrometer (Bruker BioSpin, Karlsruhe, Germany), following a previously described method [14]. A standard one-dimensional pulse sequence with saturation of the water resonance was utilised: RD – gz,1 – 90° – t – 90° – tm – gz,2 – 90° – ACQ, whereby the relaxation delay (RD) was set at 4 s, 90° represents the applied 90° radio frequency pulse, interpulse delay (t_1) was set to an interval of 4 µs, mixing time (tm) was 10 ms, magnetic field gradients (g_{z1} and g_{z2}) were applied for 1 ms and the acquisition period (AQA) was 2.7 s. The receiver gain was set to 90.5 across the board. Suppression of the water signal was achieved through continuous wave irradiation at the water resonance frequency using 25Hz radio-frequency pulse strength during the RD and tm. Each urine spectrum was acquired using 4 dummy scans, 32 scans, 64K time domain points and with a spectral window set to 20 ppm for urine. Prior to Fourier transformation, the free induction decays were multiplied by an exponential function corresponding to a line broadening of 0.3 Hz. ^{1}H NMR spectra were automatically phased and digitized over the range δ -0.5 to 9.5, and each spectrum was baseline corrected. Spectra were subsequently referenced to the internal chemical shift reference (trimethylsilyl-[2, 2, 3, 3,-2H4]-propionate, TSP) at δ 0.0. Spectral regions corresponding to the internal standard (δ -0.5 to 0.5) and water (δ 4.5 to 5.5) were excluded. Confirmation of metabolite identities in the NMR data was obtained using 1D ^{1}H NMR and 2D ^{1}H-^{1}H NMR and ^{1}H-^{13}C NMR experiments. In addition, statistical tools such as SubseT Optimization by Reference Matching (STORM) and Statistical Total Correlation SpectroscopY (STOCSY) were also applied [18, 19].

GC-MS urinary targeted analysis

Urine samples analysed by GC coupled to MS using a previously described method [15]. The GC- MS system consisted of an Agilent 7890 B (Agilent Technologies, Palo Alto, CA, USA), equipped with an automatic liquid sampler model 7693 coupled to an Agilent 7000 C single quadruple mass selective detector. Acquisition was done by Mass Hunter workstation software for Q (Agilent Technologies, Palo Alto, CA). The GC was fitted with aDB-5 MS capillary column (30 m 0.25 mm i.d., 0.25 mm film thickness; (5%-phenyl)-methylpolysiloxane bonded and cross-linked; Agilent J&W Scientific, Folsom, CA) and helium was used as carrier gas at 1 mL/min. 1 µL of each derivatised samples was injected in splitless mode injector temperature of 270°C. A glass liner ultra-inert, split less, single taper and glass wool was used to avoid a possible contamination with non-volatile elements into the column. Chromatograms obtained were uploaded directly into the MassHunter Workstation Software Version B.07.1 (Agilent Technologies Inc.) for data processing as previously described [15].

UPLC-MS urinary and faecal analysis

Urine and faecal samples were analysed by Liquid Chromatography couple to Mass Spectrometry using a previously described method by Visser et al [16]. An Acquity UPLC system coupled to a Xevo tandem MS (Waters Co, Milford MA, USA) was used. 5µL of derivatized urine and 2.5 µL of derivatized faecal sample were injected into a reversed-phase column (Acquity BEH C18 100mm x 2.1mm, 1.7µm). The desolvation gas flow was set at 900 ml/min, and the cone gas flow at 40 ml/min. The desolvation temperature was 650 ◦C. Data was collected in multiple reaction monitoring, with an electrospray ionization source in positive mode. The capillary voltage was 2700 V, the cone voltage 18V and the collision voltage (CV). Spectral acquisitions were performed using the Waters MassLynx v4.1 software.

Data analysis

Resulting data sets were imported into R for multivariate and univariate data analysis (R Foundation for Statistical Computing, Vienna, 2012) (version 3.5.1). PCA and OPLS-DA analyses were performed using the ropls package to observe clustering based on sex, location of sample collection and SCGs. Multidimensional scaling (MDS) plots were computed in R using vegan (version 2.5-2). Statistical analysis was completed using the Kruskal-Wallis test, with a Mann Whitney U test performed for significant results to determine the groups between which this difference applied. All reported p values were adjusted using Benjamini-Hochberg. The significance threshold was set at 0.05.

References

1. Mulligan, A.A., et al., *A new tool for converting food frequency questionnaire data into nutrient and food group values: FETA research methods and availability.* BMJ Open, 2014. **4**(3): p. e004503.
2. Yu, Z. and M. Morrison, *Improved extraction of PCR-quality community DNA from digesta and fecal samples.* Biotechniques, 2004. **36**(5): p. 808-12.
3. BroadInstitute. *Picard*. 17.01.2019]; Available from: http://broadinstitute.github.io/picard/.
4. Li, H., et al., *The Sequence Alignment/Map format and SAMtools.* Bioinformatics, 2009. **25**(16): p. 2078-9.
5. Wood, D.E. and S.L. Salzberg, *Kraken: ultrafast metagenomic sequence classification using exact alignments.* Genome Biology, 2014. **15**(3): p. R46.
6. Lu, J., et al., *Bracken: estimating species abundance in metagenomics data.* PeerJ Computer Science, 2017.
7. Abubucker, S., et al., *Metabolic reconstruction for metagenomic data and its application to the human microbiome.* PLoS Computational Biology, 2012. **8**(6): p. e1002358.
8. Franzosa, E.A., et al., *Species-level functional profiling of metagenomes and metatranscriptomes.* Nature Methods, 2018. **15**(11): p. 962-968.
9. Mitchell, J.H., et al., *Task Force 8: classification of sports.* Journal of the American College of Cardiology, 2005. **45**(8): p. 1364-7.
10. Oksanen, J., et al., *Community Ecology Package*. 2018.
11. Letunic, I. and P. Bork, *Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees.* Nucleic Acids Res, 2016. **44**: p. 242-245.
12. Paradis, E., J. Claude, and K. Strimmer, *APE: Analyses of Phylogenetics and Evolution in R language.* Bioinformatics, 2004. **20**(2): p. 289-90.
13. Segata, N., et al., *Metagenomic biomarker discovery and explanation.* Genome Biology, 2011. **12**(6): p. R60.
14. Dona, A.C., et al., *Precision high-throughput proton NMR spectroscopy of human urine, serum, and plasma for large-scale metabolic phenotyping.* Anal Chem, 2014. **86**(19): p. 9887-94.
15. Zhao, L., et al., *High Throughput and Quantitative Measurement of Microbial Metabolome by Gas Chromatography/Mass Spectrometry Using Automated Alkyl Chloroformate Derivatization.* Anal Chem, 2017. **89**(10): p. 5565-5577.
16. Visser, W.F., et al., *A sensitive and simple ultra-high-performance-liquid chromatography-tandem mass spectrometry based method for the quantification of D-amino acids in body fluids.* J Chromatogr A, 2011. **1218**(40): p. 7130-6.
17. Barton, W., et al., *The microbiome of professional athletes differs from that of more sedentary subjects in composition and particularly at the functional metabolic level.* Gut, 2017.
18. Cloarec, O., et al., *Statistical total correlation spectroscopy: an exploratory approach for latent biomarker identification from metabolic 1H NMR data sets.* Analytical Chemistry, 2005. **77**(5): p. 1282-9.
19. Posma, J.M., et al., *Subset optimization by reference matching (STORM): an optimized statistical approach for recovery of metabolic biomarker structural information from 1H NMR spectra of biofluids.* Analytical Chemistry, 2012. **84**(24): p. 10694-701.

Supplementary material 2: Consent form

SUBJECT INFORMATION SHEET

Study Title: **OlympicMet**

Protocol No.: APC073

Principal Investigator: Prof Fergus Shanahan

Co-investigators Dr. Orla O'Sullivan

Dr. Paul Cotter

Dr. Sharon Madigan

Dr. Rod McLoughlin

Dr. Sarah Jane Cullen

Dr. Adrian McGoldrick

Ciara O'Donovan

Site of Investigation: APC, Microbiome Institute, University College Cork (UCC), Teagasc Food Research Centre, Fermoy, Co. Cork, Racing Academy and Centre of Education (RACE), The Curragh, Co Kildare and Irish Sports Institute, National Sports Campus, Dublin 15

Why is this study being run?

The objective of this study is to establish the microbiome of elite athletes in the preparation for and/or during major competition and to determine any influences of exercise and diet on host metabolism, body composition or immune function.

If you agree to participate in the study you will be asked questions about your diet and general health and your eligibility to participate in the study will be determined.

The visits will take place at the Irish Sports Institute, Racing Academy and Centre of Education (RACE), The Curragh, Co Kildare or other affiliated training facility. This study will involve at least 1 study over 12 months with possible follow up visits in the following 2 year At least 20 healthy elite athletes (males and females) who have qualified for major competition will be enrolled.

Study Procedure

You will be advised of the purpose of the study by an investigator from the study team and the procedures which will be undertaken. You will be given a copy of the Subject Information sheet, which will explain what is required from you. If interested, you will then be requested to read and sign the Informed Consent form, and receive a signed copy. Your suitability to participate in the study will be checked and you will be asked some general personal questions. The investigator dealing with you will discuss all samples and that we wish to take. In addition you may be asked to attend follow up visits for up to two years.

Visit Outline

You will be given, in advance, a container and instructions for collecting a stool sample and urine sample in a quick and easy, hygienic device, in your own home.

At the study visit visit, you will bring your stool and urine sample, produced within the previous 12 hours. Alternatively you will collect these with special kits provided by the study tem and post them to the research centre in Fermoy, Co. Cork (kits and stamps envelopes will be provided by the study team). Details of your current and/or past medications will also be collected. You will be asked to give a blood sample (up to 20ml which is 4 teaspoons), which will be analysed for biomarkers of health and disease. Your diet will be recorded using the Food Frequency Questionnaire that will be administered during this visit. Your weight and height will be measured and your perceived gut health will be established with a gut health questionnaire.

Should further visits be required the investigator will discuss and plan these with you at your convenience.

DNA, urine and faecal samples will be stored for further testing should it be required. Samples will be destroyed on request.

What happens if I start the study and change my mind later?

You do not have to take part in the study, participation is entirely voluntary. Refusal to participate, or discontinuing participation at any time, will involve no penalty, loss of benefits or denial of treatment or services by the Cork Teaching Hospital or the participating investigators.

Will I experience any unpleasant side effects?

During the collection of your blood sample or insertion of the cannula, you may experience a slight scratch, which may be uncomfortable for a moment but quickly passes.

Funding of trial

There are no cost implications for the Health Services Executive (HSE) or to you. The management of patients and investigative tests will comply with current standards of care. Cost of research tests will be incurred by the Alimentary Pharmabiotic Centre, University College Cork.

Confidentiality

Any personal information which you volunteer to this study will be treated in accordance with the Data Protection Act 1988 and Data Protection (Amendment) Act 2003. Accordingly, the study team shall only process such information in line with the purpose for which you provided it. Your information shall not be disclosed to third parties unless you are specifically informed otherwise or we require a third party to process the information on our behalf (in which case the University requires these third parties to comply with its instructions and we require that they do not use your personal information for their own business purposes, unless you have explicitly consented to the use of your personal information in this way). You have the right to review your data and request a copy and or you have the right to request deletion of your data

All the information gathered from this study will be stored on a computer, paper files and will be treated confidentially. You will be identified only by a subject number.

In the event of any publication regarding this study, your identity will not be disclosed.

What happens if there is anything I do not understand?

If there is anything you are not sure about, the investigators will be happy to explain in more detail to yourself or your relatives, guardians (or legal representative if required). The study will be fully explained to you before you decide if you want to take part.

Dr. Orla O'Sullivan, Research Fellow,

Tel: 025 42556

Dr. Paul Cotter, Principal Research Officer

Tel: 025 42694

Prof Fergus Shanahan,

Consultant Gastroenterologist,

Tel: 021 4901226

Protocol No.: APC073

CONSENT BY SUBJECT FOR PARTICIPATION IN A HUMAN INTERVENTION STUDY

Protocol Number: APC073

Patient Name: ________________________________

Title of protocol: Study Title: OlympicMet

Principal Investigator: Prof Fergus Shanahan (021-4901226)

Participation in this study is voluntary and you may withdraw at any time for any reason

The research project and procedure associated with it have been fully explained to me. All experimental procedures have been identified and no guarantees have been given about the possible results. I have had the opportunity to ask questions concerning any and all aspects of the project and any procedures involved. I am aware that participation is voluntary and I may withdraw consent at any time. I am aware that my decision not to participate or to withdraw will not restrict my access to health care services normally available to me. Confidentiality of records concerning my involvement in this project will be maintained in an appropriate manner. I understand that the investigators have such insurance as is required by law in the event of injury resulting from this research.

I, the undersigned, hereby consent to participate as a subject in the above described project conducted at the Cork University Hospital and University College Cork. I have received a copy of this consent form for my records. I understand that if I have any questions concerning this research, I can contact the Doctor listed below. If I have further queries concerning my rights in connection with the research, I can contact the Clinical Research Ethics Committee of the Cork Teaching Hospitals, Lancaster Hall, 6 Little Hanover Street, Cork

Analyses of all samples and information collected will be conducted in the CUMH and/or APC. On some occasions the analyses may be done in collaboration with third parties, including commercial partners, which may require samples to be shipped to these organisations. Please note that data protection laws in these countries may not be as stringent.

Protocol No.: APC073

In addition samples and data will be stored and may be used in other research studies. In all cases, samples and data will be coded with anonymized identifier numbers.

(please tick box if you agree) ☐

Access to samples and/or data will require approval from the Director and/or Executive Management Group pf the APC, Microbiome Institute, Cork

I agree that results and microbial isolates may be used in the development of health promoting products

(please tick box if you agree) ☐

I agree that my contact details may be made available for recruitment to further studies and also allow a follow up check on my status following my participation by the this project or its CREC approved collaborators

(please tick box if you agree) ☐

After reading the entire consent form, if you have no further questions about giving consent, please sign the where indicated.

Subject's Signature: ______________________________ Date: ___________________

dd mon yy

NAME (BLOCK LETTERS): _________________________ Time: ___________________

Investigator's Signature: __________________________ Date: ___________________

dd mon yy

NAME (BLOCK LETTERS) __________________________ Time: ___________________

Supplementary material 3: Participant characteristics and data collected

Total participants	37
Total samples	40
Age	27 (±5) years
Sex	
Male	23
Female	14
Location collected	
Brazil	14
Ireland	26
Food frequency questionnaires completed	35 individuals (38 FFQs)
Gut health questionnaires completed	36 individuals (39 GHQs)

Disparity in numbers of individuals involved versus forms collected from arises from the collection of 2 and 3 samples from two participants. Number given for sex is given by participants involved. Numbers given for location is given as samples collected per location.

Supplementary material 4: Gut health questionnaire

Gut Health Questionnaire

Code: ______________________________

Date: ______________________________

DIRECTIONS:

This questionnaire asks you to assess how you have been feeling **during the last week.** All information is held in strict confidence. Take all the time you need to complete this questionnaire.

Section A:

For each question, place the number that best describes your symptoms beside the question:

0 = No or rarely - you have never experienced the symptom or the symptom is familiar to you but you perceive it as insignificant (monthly or less).

1 = Occasionally - Symptom comes and goes and is linked in your mind to stress, diet, fatigue or some identifiable trigger.

2 = Often - Symptom occurs 2-3 times per week and/or with a frequency that bothers you enough that you would like to do something about it.

3 = Frequently - Symptom occurs 4 or more times per week and/or you are aware of the symptom every day, or it occurs with regularity on a monthly or cyclical basis.

1. Do you feel any abdominal pain or discomfort during or after eating? __________
2. Do you feel any abdominal bloating, excessive burping or belching after meals? __________
3. Do you feel any increase flatulence or wind? __________
4. Do you feel any abdominal grumbling or gurgling? __________
5. Do you ever feel the urgency to open your bowels? __________
6. Do you ever feel like you don't fully empty your bowels (An inability to pass your entire stool)? __________
7. Do you ever feel nausea, vomiting, or heartburn? __________

8. Do you ever feel a loss of appetite? __________

9. Do you ever feel tired or lethargic? __________

Section B:

Currently, how often do you pass a bowel action (please tick a box)

- Once a week ☐
- Once every 4 - 6 days ☐
- Once every 2 – 3 days ☐
- Once every 1 – 2 days ☐
- Once a day ☐
- 2 – 3 times per day ☐
- 4 – 6 times per day ☐
- Other ____

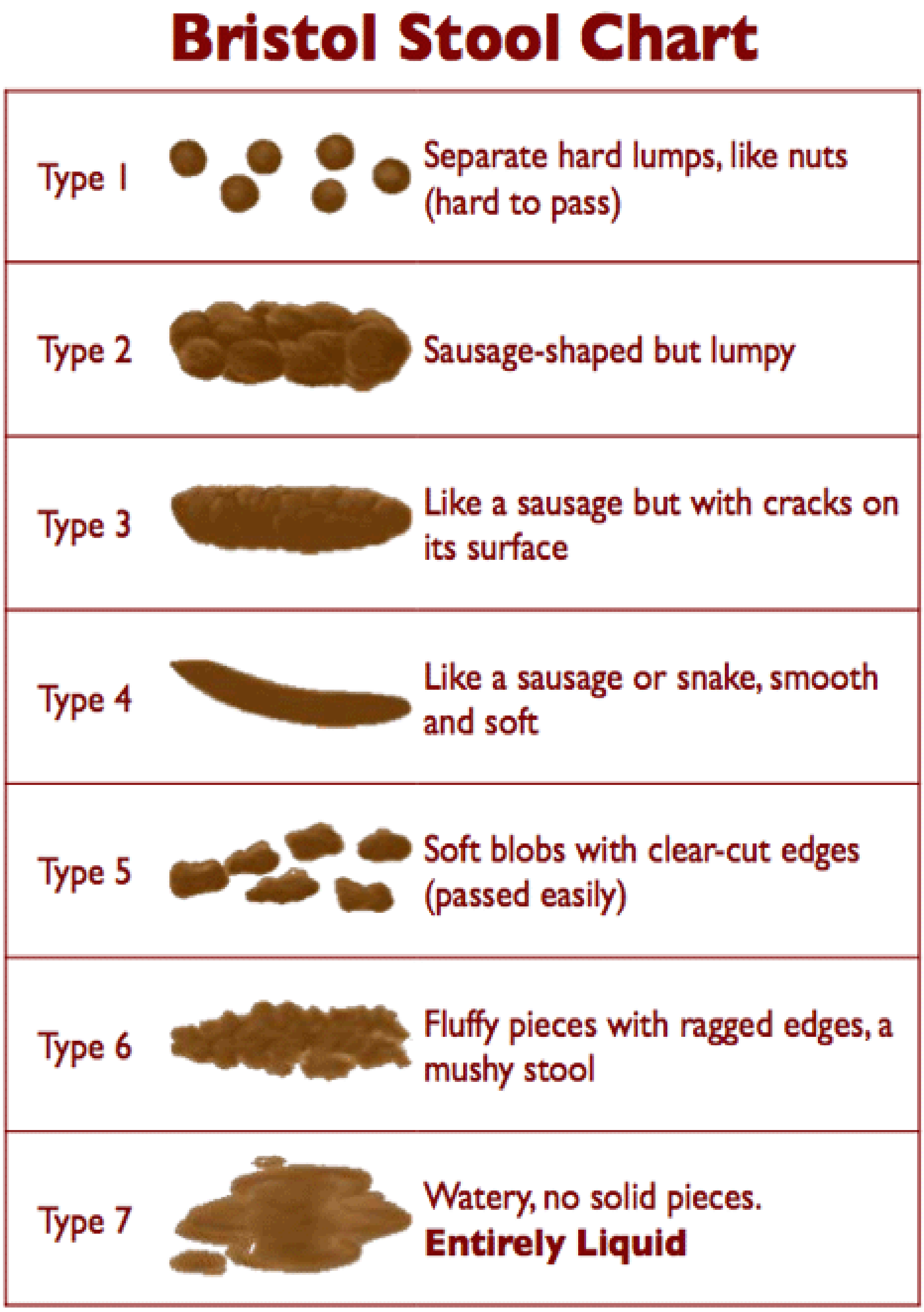
Section C:

Do you currently have satisfactory relief of your gut symptoms? (Please tick a box)

- Yes ☐
- No ☐
- If no why ____

Section D:

Please tick the box that best describes your current stool type?



Section E: Medications

1. When did you last take antibiotics?

   Over 12 months ago ☐

   6-12 months ago ☐

   3-6 months ago ☐

   Less than 3 months ago ☐

   Unsure ☐

   Other ____________________________________

2. When did you last take antifungal medication? (Ingestion only)

   Over 12 months ago ☐

   6-12 months ago ☐

   3-6 months ago ☐

   Less than 3 months ago ☐

   Unsure ☐

   Other ____________________________________

3. Have you taken any other medication in the last 6 months?

   Yes ☐

   Please specify _________________________________

   No ☐

   Rather not say ☐

4. Do you/have you ever taken supplements containing probiotics?

No ☐

Yes

Daily ☐

Weekly ☐

In the last 3 months ☐

Over 3 months ago ☐

Please provide brand name ______________________________

Supplementary material 5: Responses to gut health questionnaire

Sample	SCG	A.1	A.2	A.3	A.4	A.5	A.6	A.7	A.8	A.9	B	C	D	E.1	E.2	E.3	E.4
OL002	C2	1	0	0	0	0	0	2	0	1	6	1	2	2	6	1	2
OL021	C1	1	0	0	0	1	0	0	1	1	6	1	4	4	4	1	1
OL022	C3	1	2	3	1	3	3	0	0	0	0	0	5	3	5	2	4
OL023	A3	1	0	0	1	1	0	0	0	1	6	1	3	2	5	1	4
OL025	C2	1	2	2	3	1	1	0	0	0	5	2	3	1	6	1	2
OM001	C3	0	1	0	1	0	1	0	1	1	6	1	4	1	1	1	4
OM002	C3	0	0	0	0	0	0	0	0	0	6	1	3	6	6	2	2
OM004	C1	0	0	1	1	1	0	0	0	2	5	1	6	1	1	2	1
OM005	B2	0	0	0	0	0	1	0	0	1	6	1	2	1	6	1	1
OM006	C2	0	0	1	0	2	0	0	0	1	6	1	3	4	4	2	1
OM007	C2	2	2	0	1	2	1	0	0	1	6	1	6	1	1	1	2
OM008	C2	1	0	0	3	1	0	1	1	1	5	1	4	1	1	1	1
OM009	C2	0	0	1	1	0	1	1	0	0	4	1	4	1	1	2	5
OM014	C3	0	3	3	0	2	2	0	0	2	7	1	5	4	3	2	5
OM018	B1	0	0	0	0	0	0	0	2	0	5	1	4	0	0	0	2
OM019	A3	2	0	1	1	2	1	1	1	2	6	2	5	1	1	1	1
OM020	C1	0	0	0	0	0	0	0	0	1	6	1	6	1	6	2	1
OM022	C1	0	0	0	2	2	2	0	0	2	5	1	3	1	1	2	1
OM032	C2	0	1	1	1	1	0	0	1	1	1	1	1	2	5	2	1
OM033	C2	0	0	1	1	1	0	0	0	1	6	1	6	1	1	1	2
OM034	C2	0	1	1	1	0	0	0	1	1	3	1	5	4	5	1	1
OM035	C3	1	1	1	1	1	0	0	0	1	5	1	4	3	6	1	4
OM037	C3	2	2	3	3	3	0	3	0	2	6	1	3	1	5	2	4
OM040	C3	0	0	0	0	0	0	0	0	1	5	1	2	1	1	2	2
OM046	A3	0	0	1	1	0	0	0	1	1	6	1	4	2	6	2	4
OM058	C1	0	1	0	0	1	1	0	0	0	5	1	6	2	5	2	2
OM059	C1	1	2	2	1	1	1	0	0	1	6	1	4	2	1	1	2
OM061	C1	0	0	1	1	1	1	0	0	0	6	1	2	1	5	2	3
OM062	C1	1	1	1	0	1	0	0	0	1	6	1	4	1	1	1	1
OM063	C1	0	0	1	1	2	1	0	1	2	6	1	6	2	2	2	5
OM064	C1	0	0	1	0	0	1	0	1	0	6	1	4	3	1	2	1
OM065	C1	0	0	0	0	1	0	1	2	1	5	1	4	1	1	1	2
OM068	C1	0	3	2	1	2	1	1	1	2	6	1	4	4	5	1	2
OM069	C1	1	2	2	2	3	0	1	1	3	6	1	4	5	1	2	2
OM070	C1	1	2	4	2	1	2	0	1	1	6	1	5	5	1	2	2
OM072	B1	0	1	0	1	0	0	0	0	1	5	1	3	4	5	2	1

Supplementary material 6: Food frequency questionnaire

Please fill in the following based on the previous months eating habits

		M = month		**W = week**	**D = Day**						
ID:											**Date:**
Meat, fish, poultry - Medium serving (size deck of cards)	**Never**	**<1 / M**	**1/M**	**1/W**	**2/W**	**5/W**	**1/D**	**2/D**	**4/D**	**6+/D**	**Comment**
Beef (roast / steak)											
Beef: stew											
Beef burger (1 burger)											
Pork (roast / chops / escalopes)											
Lamb (roast / chops / stew)											
Chicken or other poultry e.g. turkey: roast											
Breaded chicken, chicken nuggets, chicken burger											
Bacon / Ham											
Processed meat (Corned beef, Luncheon meats, sausages)											
Savoury pies (e.g. meat, pork, steak & kidney, sausage roll)											
Liver, heart, kidney, paté											

Fish fried (batter/breadcrumbs), baked, grilled, fingers/cakes											
White fish, fresh or frozen (e.g. cod, haddock, plaice, sole)											
Oily fish, fresh or canned (e.g. mackerel, kippers, tuna, salmon, sardines, herring)											
Shellfish (e.g. crab, prawns, mussels)											
Bread and savoury biscuits - One slice or one biscuit	**Never**	**<1 / M**	**1/M**	**1/W**	**2/W**	**5/W**	**1/D**	**2/D**	**4/D**	**6+/D**	**Comment**
White bread and rolls (including ciabatta & pannini)											
Brown bread and rolls											
Wholemeal bread and rolls											
Wheat-free; Rye brad; spelt bread (specify)											
Cream crackers, cheese biscuits											
Crisp bread, e.g. Ryvita											
Pancakes, muffins, oatcakes											
Scone (white)											
Scone (brown)											
Cereals - One medium sized bowl	**Never**	**<1 / M**	**1/M**	**1/W**	**2/W**	**5/W**	**1/D**	**2/D**	**4/D**	**6+/D**	**Comment**
Non-ready to eat - Porridge, Readybrek											

High fibre (All Bran/flakes, Weetabix, Shredded wheat)										
Cornflakes, Rice Krispies										
Muesli (e.g. Country Store, Aplen)										
Sugar-coated cereals (e.g. Frosties, Crunchy nut cornflakes)										
Potatoes, rice and pasta - Medium serving = a cupful	**Never**	**<1 / M**	**1/M**	**1/W**	**2/W**	**5/W**	**1/D**	**2/D**	**4/D**	**6+/D**	**Comment**
Boiled, instant or jacket potatoes											
Mashed potatoes											
Chips / Roast potatoes											
Potato Salad											
White Rice											
Brown Rice											
White/yellow or green pastas (e.g. spagetti, noodles)											
Wholemeal pasta											
Lasagne (meat based)											
Lasagne (vegetarian)											
Pizza											
Dairy products and fat	**Never**	**<1 / M**	**1/M**	**1/W**	**2/W**	**5/W**	**1/D**	**2/D**	**4/D**	**6+/D**	**Comment**
Cream (tablespoon)											

Full-fat yoghurt or Greek style yoghurt (125g carton)											
Dairy desserts (125g carton)											
Cheddar cheese (medium serving)											
Low-fat cheddar cheese (medium serving)											
Cottage cheese											
Eggs as boiled, fried, scrambled, poached (one)											
Quiche (medium serving)											
Light salad cream or light mayonnaise (tablespoon)											
Salad cream, mayonnaise (tablespoon)											
French dressing (tablespoon)											
Other salad dressing (tablespoon)											
The following on bread or vegetables (teaspoon):											
Butter											
Lite Butter e.g. Dawn Lite, Connacht Gold											
Sunflower margarine e.g. Flora											
Low-fat margarine (e.g. lowlow)											
Cholesterol Lowering e.g. Flora Pro-Active, DairyGold											

Heart											
Cream & Veg. Oil spread e.g. Kerrymaid, Dairy Gold											
Olive oil spread e.g. Golden Olive											
Fruit - 1 fruit or a medium serving (a cupful)	**Never**	**<1 / M**	**1/M**	**1/W**	**2/W**	**5/W**	**1/D**	**2/D**	**4/D**	**6+/D**	**Comment**
Apples											
Pears											
Oranges, satsumas, mandarins											
Grapefruit											
Bananas											
Grapes											
Melon											
Peaches, plums											
Apricots											
Strawberries, raspberries, kiwi fruit											
Blueberries											
Tinned fruit (specify)											
Dried fruit e.g. raisins											
Frozen fruit (specify)											
Vegetables (Fresh, frozen, tinned) - 2 tablespoons	**Never**	**<1 / M**	**1/M**	**1/W**	**2/W**	**5/W**	**1/D**	**2/D**	**4/D**	**6+/D**	**Comment**

Carrots											
Spinach											
Broccoli, spring greens, kale											
Brussel sprouts											
Cabbage											
Peas											
Green beans, broad beans, runner beans											
Courgettes											
Cauliflower											
Parsnips, turnips											
Leeks											
Onions											
Garlic											
Mushrooms											
Sweet peppers											
Green salad, lettuce											
Cucumber, celery											
Tomatoes											
Sweetcorn											
Beetroot											
Coleslaw											

Baked beans											
Dried lentils, beans, peas											
Tofu, soya meat, TVP, vegeburger											
Sweets and snacks - medium serving	**Never**	**<1 / M**	**1/M**	**1/W**	**2/W**	**5/W**	**1/D**	**2/D**	**4/D**	**6+/D**	**Comment**
Chocolate coated sweet biscuits e.g. digestive (one)											
Plain biscuit e.g. marietta, digestives, rich tea (one)											
Cakes e.g. fruit, sponge											
Buns, pastries e.g. croissants, doughnuts											
Fruit pies, tarts, crumbles											
Sponge puddings											
Milk puddings e.g. rice, custard, trifle											
Ice cream, choc ices, frozen desserts											
Chocolates, singles or squares											
Sweets, toffees, mints											
Sugar added to tea coffee, cereals (teaspoons)											
Sugar substitute e.g. canderel (teaspoon)											
Crisps or other packet snacks											
Peanuts or other nuts											

Soups, sauces and spreads	Never	<1 / M	1/M	1/W	2/W	5/W	1/D	2/D	4/D	6+/D	Comment
Vegetable soups: homemade/fresh (1 bowl)											
Vegetable soups: tinned/packet (1 bowl)											
Meat or cream soups: homemade / fresh (1 bowl)											
Meat or cream soups: tinned / packet (1 bowl)											
Sauces e.g. white, cheese, gravy (tablespoon)											
Tomato based sauces e.g. pasta sauces											
Curry-type sauces											
Pickles, chutney (tablespoon)											
Marmite, Bovril (tablespoon)											
Jam, marmalade, honey, syrup (teaspoon)											
Peanut butter (teaspoon)											
Drinks	**Never**	**<1 / M**	**1/M**	**1/W**	**2/W**	**5/W**	**1/D**	**2/D**	**4/D**	**6+/D**	**Comment**
Tea (cup)											
Herbal tea (cup)											
Whole milk (cup) - cow, goat, soya, rice milk - specify											
Semi-skimmed milk (cup)											
Coffee instant (cup)											
Coffee ground (cup)											

Coffee, decaffeinated (cup)											
Coffee whitener e.g. Coffee-mate (teaspoon)											
Cocoa, Hot Chocolate (cup)											
Horlicks, Ovaltine (cup)											
Wine (glass)											
Beer, Larger or Cider (half pint)											
Low alcohol / alcohol free beer / larger (half pint)											
Port, Sherry, Vermouth, liqueurs (glass)											
Spirits e.g. Gin, Whiskey (Single measure)											
Low calorie or diet soft fizzy drink (glass)											
Fizzy Soft drinks e.g. (Coca Cola - Glass)											
Pure fruit drinks e.g. orange juice (small glass)											
Fruit squash (small glass)											
Functional Foods / Supplements	**Never**	**<1 / M**	**1/M**	**1/W**	**2/W**	**5/W**	**1/D**	**2/D**	**4/D**	**6+/D**	**Comment**
Probiotic Yoghurts / cheese / milk											
Vitamin supplements (details)											
Mineral supplements (details)											
Ready meals	**Never**	**<1 / M**	**1/M**	**1/W**	**2/W**	**5/W**	**1/D**	**2/D**	**4/D**	**6+/D**	**Comment**

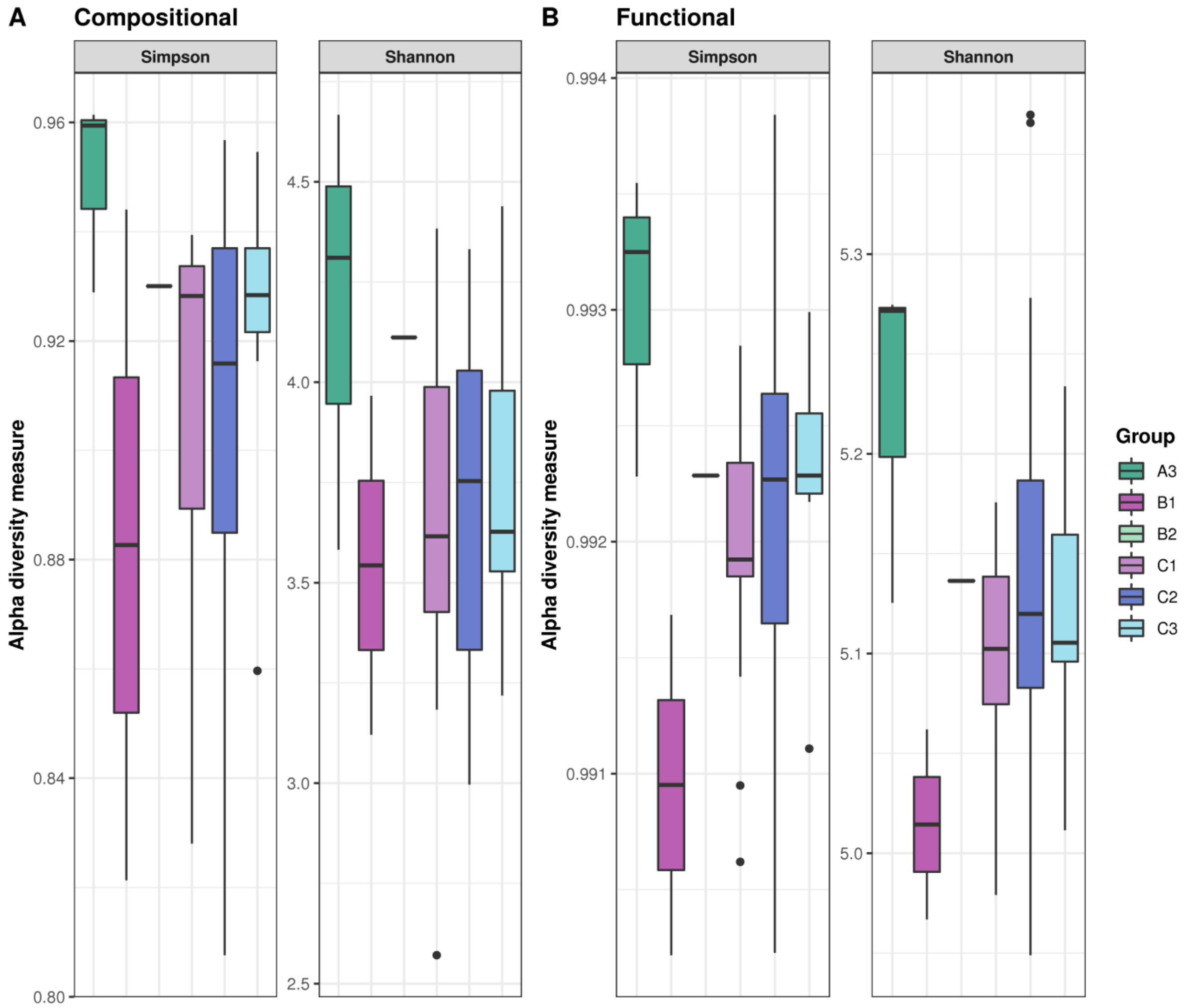
Ready meal (specify)											
Takeaway (specify)											

Supplementary material 8: Dietary intakes compared using Kruskal-Wallis between sports classification groups

Energy, nutrients and food groups	P value
Energy (kcal)	0.233
Energy (kj)	0.233
Protein(g)	0.249
Alcohol(g)	0.216
Total carbohydrate (g)	0.255
Starch (g)	0.233
Englyst fibre(g)	0.164
Total fat(g)	0.216
Monounsaturated fat (g)	0.216
Polyunsaturated fat (g)	0.216
Saturated fat(g)	0.216
Calcium (mg)	0.216
Iron(mg)	0.164
Potassium(mg)	0.216
Carotene(µg)	0.216
Folate (mcg)	0.164
Vitamin C(mg)	0.255
Vitamin D (mcg)	0.139
Vitamin E(mg)	0.216
Alcoholic beverages (g)	0.216
Cereals and cereal products (g)	0.216
Eggs and egg dishes (g)	0.588
Fats and oils (g)	0.216
Fish and fish products (g)	0.164
Fruit (g)	0.216
Meat and meat products (g)	0.216
Milk and milk products(g)	0.216
Non-alcoholic beverages (g)	0.216
Nuts and seeds(g)	0.188
Potatoes(g)	0.728
Soups and sauces(g)	0.216
Sugars, preserves and snacks(g)	0.588
Vegetables (g)	0.164

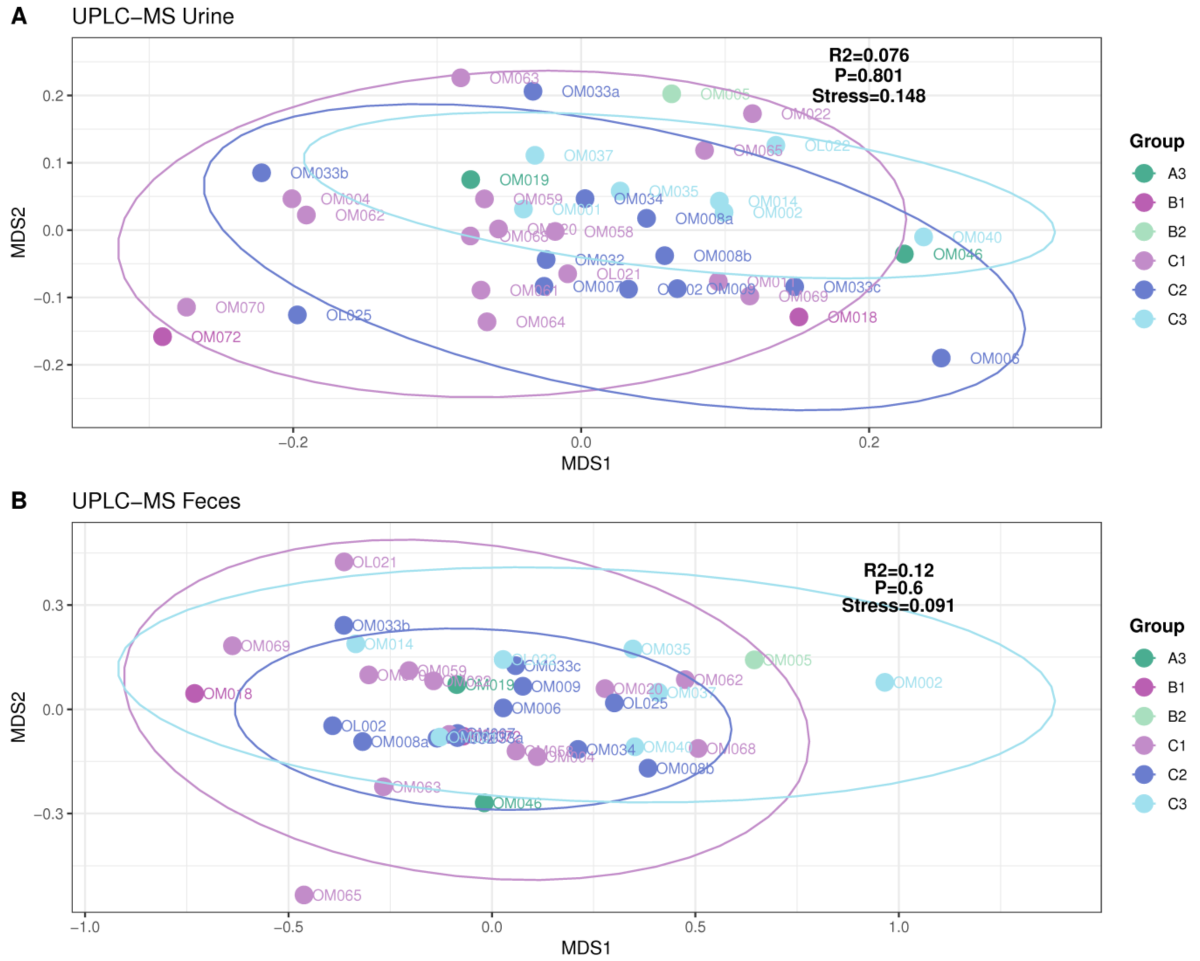
Dietary intakes compared by sports classification group (A3 (N=3), B (grouped samples with a moderate dynamic component for statistical analysis; N=3), C1 (N=15), C2 (N=11), C3 (N=7)) (See figure 1A for more information regarding groupings) using the Kruskal-Wallis test and adjusted using Benjamini-Hochberg.

Supplementary material 9: Alpha diversity



Diversity measures of gut microbiome composition (A) and functional potential (B) as determined using vegan in R coloured by group. Species level assignments from Bracken following taxonomic classification with Kraken 2 were analysed with vegan to determine Shannon and Simpson diversities for compositional data. Unstratified pathway abundance (reads per million) assignments as determined using HUMAnN2 were analysed with vegan to determine Shannon and Simpson diversities for functional data. No significant differences were established in alpha diversity by sex, location of sample collection, sport or sports classification group.

Supplementary Material 10: Multidimensional scaling (MDS) result from UPLC-MS analysis

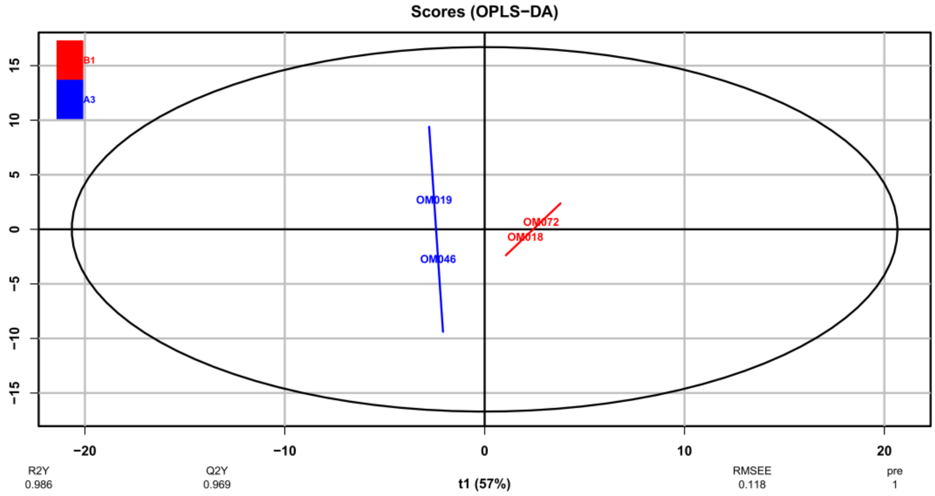


Multidimensional scaling (MDS) result from UPLC-MS analysis of (A) urine and (B) faecal samples, as determined using vegan in R, show the differences between metabolites contained in samples based on sports classification group (A3 (N=3), B1 (N=2), B2 (N=1), C1 (N=15), C2 (N=12), C3 (N=7)) (See figure 1A for more information regarding groupings). Clustering of samples was not observed based on sports classification group for UPLC-MS analysis of urine and faecal samples.

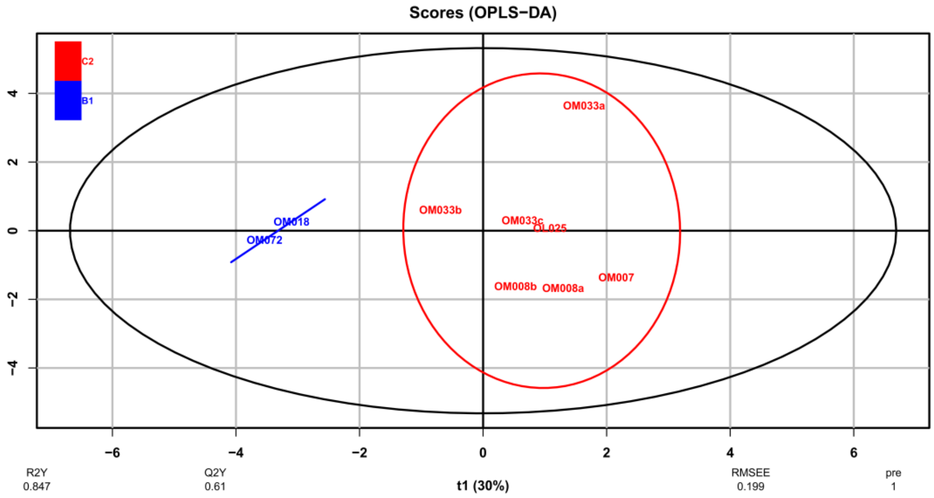
Supplementary Material 11: OPLS-DA analysis of GC-MS urine data. OPLS-DA analysis was performed using the ropls package in R to analyse the differences between sports classification groups (A3 (N=3), B1 (N=2), B2 (N=1), C1 (N=15), C2 (N=12), C3 (N=7)) (See figure 1.A for more information regarding groupings) using paired analysis.

Supplementary Material 11: OPLS-DA analysis of GC-MS urine data

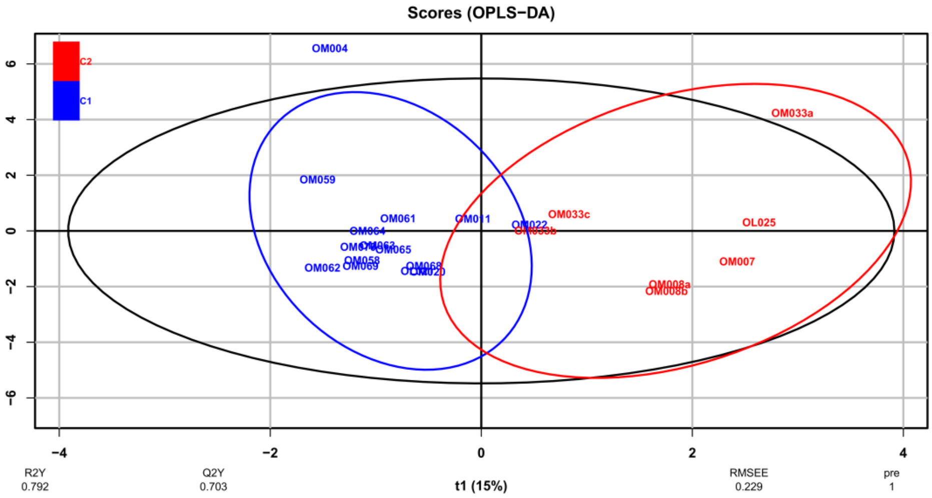
A



B



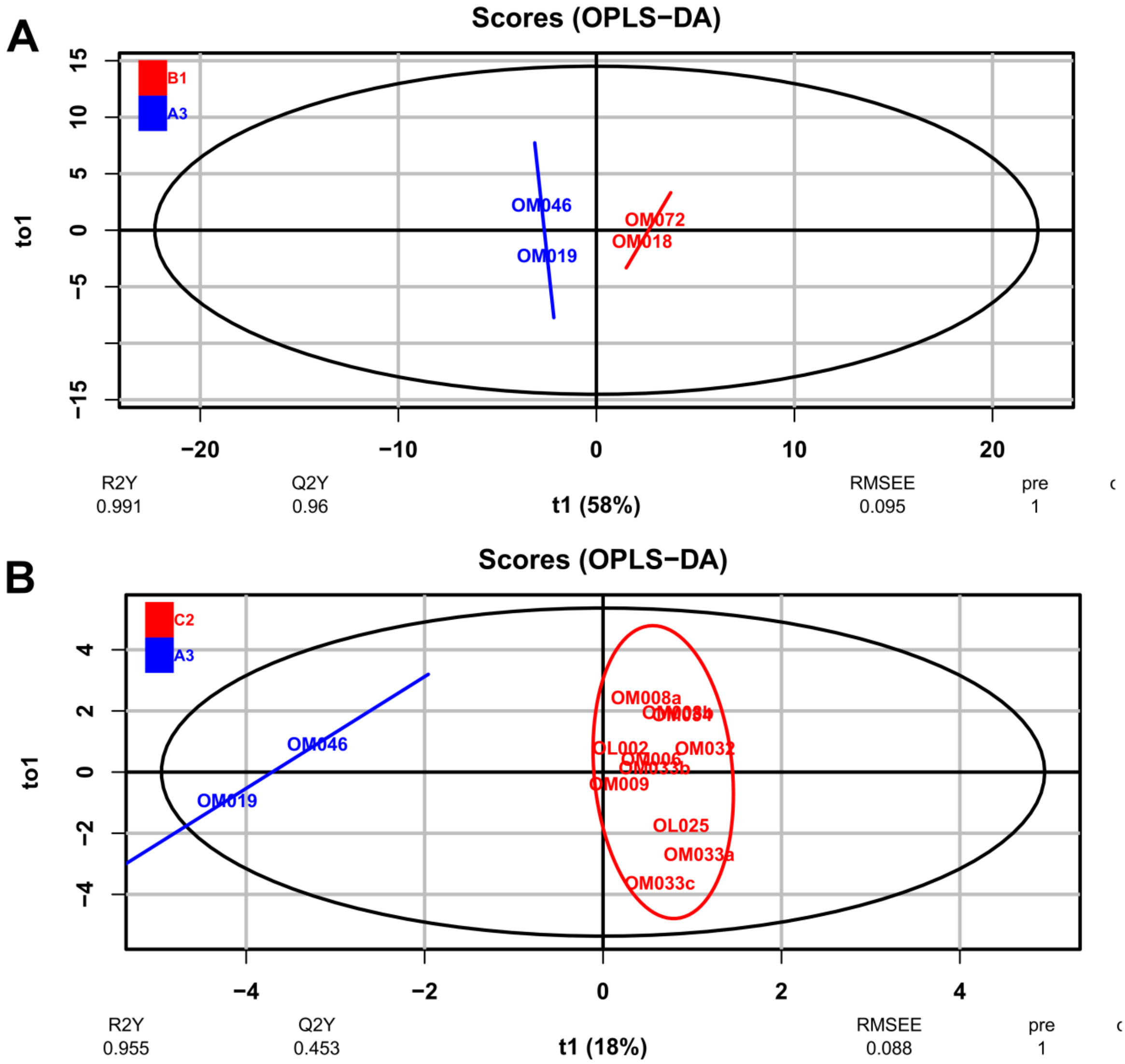
C



OPLS-DA analysis of GC-MS urine data. OPLS-DA analysis was performed using the ropls package in R to analyse the differences between sports classification groups (A3 (N=3), B1 (N=2), B2 (N=1), C1

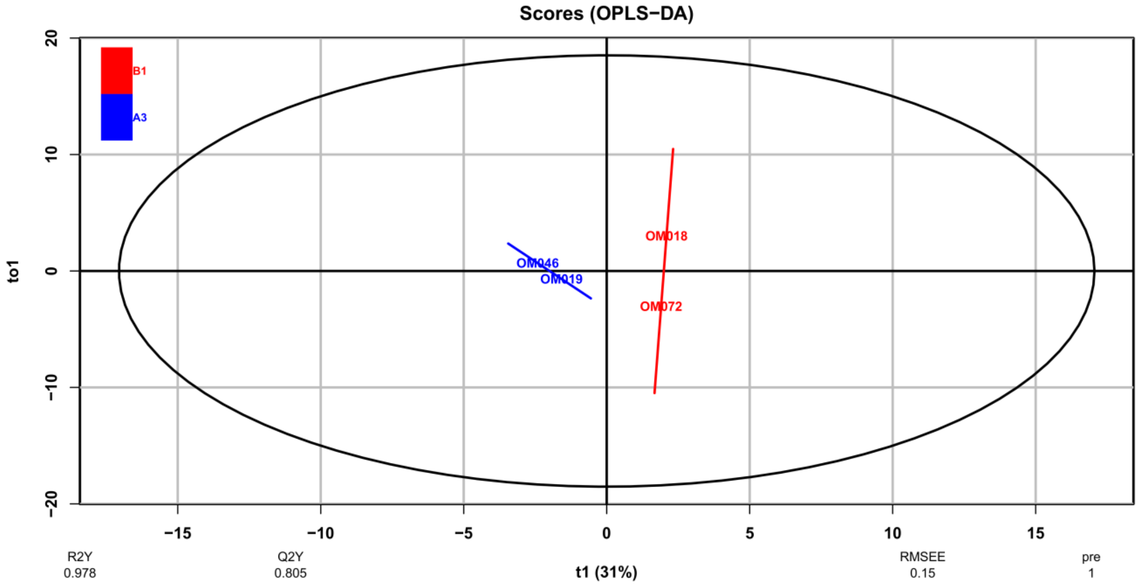
(N=15), C2 (N=12), C3 (N=7)) (See figure 1A for more information regarding groupings) using paired analysis.

Supplementary Material 12: OPLS-DA analysis of NMR urine data



OPLS-DA analysis of NMR urine data. OPLS-DA analysis was performed using the ropls package in R to analyse the differences between sports classification groups (A3 (N=3), B1 (N=2), B2 (N=1), C1 (N=15), C2 (N=12), C3 (N=7)) (See figure 1A for more information regarding groupings) using paired analysis.

Supplementary Material 13: OPLS-DA analysis of NMR faeces data



OPLS-DA analysis of NMR faeces data. OPLS-DA analysis was performed using the ropls package in R to analyse the differences between sports classification groups (A3 (N=3), B1 (N=2), B2 (N=1), C1 (N=15), C2 (N=12), C3 (N=7)) (See figure 1A for more information regarding groupings) using paired analysis.

Supplementary material 14: Significant differences in metabolites overall

Metabolite	Material	Analysis	P value
Creatinine	Feces	NMR	0.016979
Pyroglutamic acid	Urine	GC-MS	0.005045
α-hydroxybutyric acid	Urine	GC-MS	0.009892
3-aminoisobutanoic acid	Urine	GC-MS	0.00366
γ-aminobutyric acid	Urine	GC-MS	0.006636
Ictaconic acid	Urine	GC-MS	0.007733
4-hydroxybenzouc acid	Urine	GC-MS	0.004581
L-Phenylalanine	Urine	GC-MS	0.006021
L-Tryptophan	Urine	GC-MS	0.003693
3 indolepropionic acid	Urine	GC-MS	0.018406
3-hydroxyphenylacetic acid	Urine	GC-MS	0.004851
Phenyllactic acid	Urine	GC-MS	0.004086
Succinic acid	Urine	GC-MS	0.003413
Cis-aconitate	Urine	GC-MS	0.009654
Succinate	Urine	NMR	0.0072
Hippurate	Urine	NMR	0.007106
Lactate	Urine	NMR	0.00578
Acetate	Urine	NMR	0.0225
Creatinine	Urine	NMR	0.0457
Citrate	Urine	NMR	0.0219
L-Ile	Urine	UPLC-MS	0.043851

Significantly different metabolites between sports classification groups (A3 (N=3), B1 (N=2), B2 (N=1), C1 (N=15), C2 (N=12), C3 (N=7)) (See figure 1A for more information regarding groupings).Statistical analysis was completed between sports classification groupings using the Kruskal-Wallis test, with a Mann Whitney U test performed for significant results to determine the groups between which this difference applied (as shown in figure 3D). All reported p values were adjusted using Benjamini-Hochberg.