**Supplementary Table 1**. Bowel preparation adequacy and colonoscopy indications and findings.

|  | **Complete cases(N=59)** | **Lost to follow up(N=20)** | **All cases(N=79)** |
| --- | --- | --- | --- |
| **Bowel preparation adequacy, yes n(%) a** |  |  |  |
|  Fair | 15 (25%) | 5 (25%) | 20 (25%) |
|  Good | 37 (63%) | 11 (55%) | 48 (61%) |
|  Excellent | 7 (12%) | 4 (20%) | 11 (14%) |
| **Indication for colonoscopy, yes n(%)** |  |  |  |
| Positive faecal occult blood test | 24 (41%) | 6 (30%) | 30 (38%) |
| Surveillance colonoscopy c | 19 (32%) | 1 (5%) | 20 (25%) |
| Other symptoms e | 18 (31%) | 8 (40%) | 26 (33%) |
| Change in bowel habits b | 5 (9%) | 7 (35%) | 12 (15%) |
| Iron-deficiency anaemia | 5 (9%) | 1 (5%) | 6 (8%) |
| Suspected lesions d | 2 (3%) | 0 (0%) | 2 (3%) |
| **Colonoscopy outcome, yes n(%) f** |  |  |  |
| No abnormalities detected | 12 (20%) | 7 (35%) | 19 (24%) |
| Diverticular disease | 18 (31%) | 4 (20%) | 22 (28%) |
| Polyps | 37 (63%) | 9 (45%) | 46 (58%) |
| Other g | 16 (27%) | 8 (40%) | 24 (30%) |
| *Notes:* *a Based on an overall Boston Bowel Preparation Scale score.**b Change in bowel habit(s) includes frequent stools, diarrhoea, faecal incontinence.**c Surveillance includes for family history, previous polyps, previous gastric surgery.* *d Clinically or radiologically detected lesions included thickened area on scans and possible rectal mass detected; detected lesions were not associated with a cancer diagnosis.**e Other symptoms included abdominal pain, peri-rectal bleeding, haemorrhoids, anorectal pain, weight loss, vomiting, bloating, pain on defecation.* *f Some participants received more than one diagnosis after colonoscopy, therefore percentages exceed 100%.**g Other outcomes included haemorrhoids, granular mucosa, and muscle hypertrophy.*  |

**Supplementary Table S2.** Unadjusted and adjusted models of the change in beta-diversity of faecal samples collected one week before and one month after bowel preparation and colonoscopy using the Aitchison distance metric.

|  |  |
| --- | --- |
|  | **Aitchison distance** |
| *Predictors* | *Df* | *Sum of squares* | *R2* | *F* | *p-value* |
| **Unadjusted model** |
| Time point | 1 | 1975 | 0.004 | 0.441 | **0.001** |
| Residual | 116 | 519339 | 0.996 |  |  |
| Total | 117 | 521314 | 1.000 |  |  |
| **Adjusted model** |
| Age (years) | 1 | 7610 | 0.015 | 1.749 | 0.246 |
| Sex (male) | 1 | 7228 | 0.014 | 1.661 | 0.604 |
| BMI | 1 | 7549 | 0.015 | 1.735 | 0.374 |
| Diet quality | 1 | 8891 | 0.017 | 2.043 | 0.028 |
| IBS (yes) | 1 | 5302 | 0.010 | 1.219 | 0.287 |
| Time point | 1 | 1976 | 0.004 | 0.454 | 0.001 |
| Residual | 111 | 482935 | 0.926 |  |  |
| Total | 117 | 521314 | 1.000 |  |  |
| *Note: Age at time of recruitment; BMI calculated as weight (kilograms)/height(metres)2 at time of recruitment; Diet quality measured using a Simple Dietary Questionnaire based on previous studies(37).**Abbreviations: BMI, body mass index; CI, confidence interval; Df, degrees of freedom; IBS, irritable bowel syndrome.* |

**Supplementary Table S3. Adjusted analyses of the change in differential abundance of taxa at the genus level.** Taxa were adjusted for age, sex, body mass index, diet quality, and irritable bowel syndrome at baseline, and for multiple comparisons using Benjamini-Hochberg correction (q<0.05).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Family* | *Genus* | *Coefficient (SE)* | *p-value* | *q-value* |
| **Lower** |
| Lachnospiraceae | Unidentified | −0.36 (0.09) | <0.001 | <0.001 |
| **Higher** |
| Propionibacteriaceae | Cutibacterium | 0.30 (0.04) | <0.001 | <0.001 |
| Prevotellaceae | Prevotella 9 | 0.24 (0.07) | 0.001 | 0.001 |
| Veillonellaceae | Megamonas | 0.23 (0.03) | <0.001 | <0.001 |
| Flavobacteriaceae | Uncultured | 0.23 (0.03) | <0.001 | <0.001 |
| Uncultured | Unidentified | 0.22 (0.03) | <0.001 | <0.001 |
| Unidentified | Unidentified | 0.21 (0.05) | <0.001 | <0.001 |
| Ruminococcaceae | Ruminococcaceae UCG-009 | 0.21 (0.05) | <0.001 | <0.001 |
| Burkholderiaceae | Oxalobacter | 0.21 (0.06) | <0.001 | 0.001 |
| Lachnospiraceae | Lactonifactor | 0.20 (0.06) | 0.001 | 0.001 |
| Ruminococcaceae | CAG 352 | 0.20 (0.05) | <0.001 | 0.001 |
| Christensenellaceae | Unidentified | 0.20 (0.04) | <0.001 | <0.001 |
| Lachnospiraceae | GCA 900066575 | 0.20 (0.05) | <0.001 | 0.001 |
| Eubacteriaceae | Eubacterium | 0.19 (0.05) | <0.001 | 0.001 |
| Eggerthellaceae | Uncultured | 0.19 (0.04) | <0.001 | <0.001 |
| Eggerthellaceae | Gordonibacter | 0.19 (0.04) | <0.001 | <0.001 |
| Erysipelotrichaceae | Solobacterium | 0.17 (0.04) | <0.001 | <0.001 |
| Rikenellaceae | Rikenellaceae RC9 gut group | 0.16 (0.05) | 0.001 | 0.001 |
| Uncultured | Gut metagenome | 0.16 (0.04) | 0.001 | 0.001 |
| Saccharimonadaceae | Uncultured | 0.16 (0.05) | 0.001 | 0.001 |
| Lachnospiraceae | Lachnospiraceae UCG-010 | 0.15 (0.04) | 0.001 | 0.001 |
| Veillonellaceae | Megasphaera | 0.14 (0.04) | 0.001 | 0.001 |

**Supplementary Table S4.** **Adjusted analyses of the change in differential abundance of MetaCyc pathways.** Analyses were further adjusted for age, sex, body mass index, diet quality, and irritable bowel syndrome at baseline, and for multiple comparisons using Benjamini-Hochberg correction (q<0.05).

|  |  |  |  |
| --- | --- | --- | --- |
| *Pathway* | *Coefficient (SE)* | *p-value* | *q-value* |
| **Higher** |  |
| PWY.5741 | 0.43 (0.08) | <0.001 | <0.001 |
| PWY.3801 | 0.41 (0.06) | <0.001 | <0.001 |
| LPSSYN.PWY | 0.39 (0.10) | <0.001 | <0.001 |
| PWY.6728 | 0.38 (0.09) | <0.001 | <0.001 |
| PWY.1422 | 0.37 (0.05) | <0.001 | <0.001 |
| PWY.6397 | 0.37 (0.05) | <0.001 | <0.001 |
| PWY.6731 | 0.30 (0.06) | <0.001 | <0.001 |
| PWY.7031 | 0.29 (0.07) | <0.001 | <0.001 |
| PWY.4361 | 0.29 (0.06) | <0.001 | <0.001 |
| PWY.7527 | 0.28 (0.07) | <0.001 | <0.001 |

**Supplementary Table S5. Associations between the change in depressive symptom scores and change in the relative abundance of taxa at the genus level one month after bowel preparation and colonoscopy compared to baseline.** Linear regression analyses were adjusted for age, sex, body mass index, diet quality, irritable bowel syndrome at baseline, and for multiple comparisons using Benjamini-Hochberg correction (q<0.05).

|  |  |  |  |
| --- | --- | --- | --- |
| *Genus* | *Coefficient (SE)* | *p-value* | *q-value* |
| **Negatively associated** |
| Ruminococcaceae UCG-009 | -0.94 (0.23) | <0.001 | 0.028 |
| Harryflintia | -0.57 (0.19) | 0.004 | 0.215 |
| Unidentified | -0.36 (0.17) | 0.036 | 0.841 |
| Uncultured | -0.23 (0.08) | 0.005 | 0.215 |
| Klebsiella | -0.14 (0.05) | 0.006 | 0.230 |
| **Positively associated** |
| Granulicatella | 0.22 (0.10) | 0.036 | 0.841 |
| Uncultured | 0.15 (0.04) | 0.001 | 0.103 |
| Haemophilus | 0.13 (0.06) | 0.044 | 0.841 |

**Supplementary Material**

**RStudio libraries:**

tidyverse

Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G,

Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K,

Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K,

Yutani H (2019). “Welcome to the tidyverse.” \_Journal of Open Source Software\_,

\*4\*(43), 1686. doi:10.21105/joss.01686 <https://doi.org/10.21105/joss.01686>.

knitr

Xie Y (2023). \_knitr: A General-Purpose Package for Dynamic Report Generation in

R\_. R package version 1.43, <https://yihui.org/knitr/>.

Xie Y (2015). \_Dynamic Documents with R and knitr\_, 2nd edition. Chapman and

Hall/CRC, Boca Raton, Florida. ISBN 978-1498716963, <https://yihui.org/knitr/>.

Xie Y (2014). “knitr: A Comprehensive Tool for Reproducible Research in R.” In

Stodden V, Leisch F, Peng RD (eds.), \_Implementing Reproducible Computational

Research\_. Chapman and Hall/CRC. ISBN 978-1466561595.

microbiomeMarker

Cao Y, Dong Q, Wang D, Zhang P, Liu Y, Niu C (2022). “microbiomeMarker: an

R/Bioconductor package for microbiome marker identification and visualization.”

\_Bioinformatics\_. doi:10.1093/bioinformatics/btac438

<https://doi.org/10.1093/bioinformatics/btac438>.

phyloseq

McMurdie PJ, Holmes S (2013). “phyloseq: An R package for reproducible

interactive analysis and graphics of microbiome census data.” \_PLoS ONE\_,

\*8\*(4), e61217. <http://dx.plos.org/10.1371/journal.pone.0061217>.

vegan

Oksanen J, Simpson G, Blanchet F, Kindt R, Legendre P, Minchin P, O'Hara R,

Solymos P, Stevens M, Szoecs E, Wagner H, Barbour M, Bedward M, Bolker B,

Borcard D, Carvalho G, Chirico M, De Caceres M, Durand S, Evangelista H,

FitzJohn R, Friendly M, Furneaux B, Hannigan G, Hill M, Lahti L, McGlinn D,

Ouellette M, Ribeiro Cunha E, Smith T, Stier A, Ter Braak C, Weedon J (2022).

\_vegan: Community Ecology Package\_. R package version 2.6-4,

<https://CRAN.R-project.org/package=vegan>.

Maaslin2

Mallick H, Rahnavard A, McIver LJ, Ma S, Zhang Y, Nguyen LH, Tickle TL, Weingart

G, Ren B, Schwager EH, Chatterjee S, Thompson KN, Wilkinson JE, Subramanian A,

Lu Y, Waldron L, Paulson JN, Franzosa EA, Bravo HC, Huttenhower C (2021).

“Multivariable Association in Population-scale Meta-omics Studies.” \_PLOS

Computational Biology\_, \*17\*, e1009442. doi:10.1371/journal.pcbi.1009442

<https://doi.org/10.1371/journal.pcbi.1009442>,

<http://huttenhower.sph.harvard.edu/maaslin2>.

Mallick H, Rahnavard A, McIver LJ (2020). \_MaAsLin 2: Multivariable Association

in Population-scale Meta-omics Studies.\_. R/Bioconductor package,

<http://huttenhower.sph.harvard.edu/maaslin2>.

DirichletMultinomial

Morgan M (2023). \_DirichletMultinomial: Dirichlet-Multinomial Mixture Model

Machine Learning for Microbiome Data\_. doi:10.18129/B9.bioc.DirichletMultinomial

<https://doi.org/10.18129/B9.bioc.DirichletMultinomial>, R package version

1.42.0, <https://bioconductor.org/packages/DirichletMultinomial>.

PICRUSt2

Yang C, Zhang L (2023). \_ggpicrust2: Make 'PICRUSt2' Output Analysis and

Visualization Easier\_. R package version 1.7.1,

<https://CRAN.R-project.org/package=ggpicrust2>.

geepack

Halekoh U, Højsgaard S, Yan J (2006). “The R Package geepack for Generalized

Estimating Equations.” \_Journal of Statistical Software\_, \*15/2\*, 1-11.

Yan J, Fine JP (2004). “Estimating Equations for Association Structures.”

\_Statistics in Medicine\_, \*23\*, 859-880.

Yan J (2002). “geepack: Yet Another Package for Generalized Estimating

Equations.” \_R-News\_, \*2/3\*, 12-14.

mice

van Buuren S, Groothuis-Oudshoorn K (2011). “mice: Multivariate Imputation by

Chained Equations in R.” \_Journal of Statistical Software\_, \*45\*(3), 1-67.

doi:10.18637/jss.v045.i03 <https://doi.org/10.18637/jss.v045.i03>.

table1

Rich B (2023). \_table1: Tables of Descriptive Statistics in HTML\_. R package

version 1.4.3, <https://CRAN.R-project.org/package=table1>.

pairwiseAdonis

Martinez Arbizu P (2017). \_pairwiseAdonis: Pairwise Multilevel Comparison using

Adonis\_. R package version 0.4.1.

sjPlot

Lüdecke D (2023). \_sjPlot: Data Visualization for Statistics in Social Science\_.

R package version 2.8.14, <https://CRAN.R-project.org/package=sjPlot>.

ggprism

Dawson C (2022). \_ggprism: A 'ggplot2' Extension Inspired by 'GraphPad Prism'\_.

R package version 1.0.4, <https://CRAN.R-project.org/package=ggprism>.

patchwork

Pedersen T (2022). \_patchwork: The Composer of Plots\_. R package version 1.1.2,

<https://CRAN.R-project.org/package=patchwork>.

padr

Thoen E (2022). \_padr: Quickly Get Datetime Data Ready for Analysis\_. R package

version 0.6.2, <https://CRAN.R-project.org/package=padr>.

xtable

Dahl D, Scott D, Roosen C, Magnusson A, Swinton J (2019). \_xtable: Export Tables

to LaTeX or HTML\_. R package version 1.8-4,

<https://CRAN.R-project.org/package=xtable>.

parallel

R Core Team (2023). \_R: A Language and Environment for Statistical Computing\_. R

Foundation for Statistical Computing, Vienna, Austria.

<https://www.R-project.org/>.

reshape2

Wickham H (2007). “Reshaping Data with the reshape Package.” \_Journal of

Statistical Software\_, \*21\*(12), 1-20. <http://www.jstatsoft.org/v21/i12/>.

magrittr

Bache S, Wickham H (2022). \_magrittr: A Forward-Pipe Operator for R\_. R package

version 2.0.3, <https://CRAN.R-project.org/package=magrittr>.

ggpubr

Kassambara A (2023). \_ggpubr: 'ggplot2' Based Publication Ready Plots\_. R

package version 0.6.0, <https://CRAN.R-project.org/package=ggpubr>.