Lower gut microbiome diversity and higher abundance of proinflammatory genus *Collinsella* are associated with biopsy-proven non-alcoholic steatohepatitis

Stuart Astbury[1,2,*], *Edmond Atallah*[1,2,*], *Amrita Vijay*[1,3], *Guruprasad P Aithal*[1,2], *Jane I Grove*[1,2,#], *Ana M Valdes*[2,4]

[1]*Nottingham Digestive Diseases Centre, School of Medicine, University of Nottingham, UK*

[2]*National Institute for Health Research (NIHR) Nottingham Biomedical Research Centre, Nottingham University Hospitals NHS Trust and the University of Nottingham, UK*

[3]*Department of Twin Research and Genetic Epidemiology, King's College London, London, UK*

[4]*Division of Rheumatology, Orthopaedics and Dermatology, School of Medicine, University of Nottingham, UK*

[*]Joint first authors

[#]Correspondence: jane.grove@nottingham.ac.uk, Nottingham Digestive Diseases Centre, E Floor, West Block, Queen's Medical Centre, Nottingham, NG7 2UH, UK.

Lower gut microbiome diversity and higher abundance of proinflammatory genus *Collinsella* are associated with biopsy-proven non-alcoholic steatohepatitis

There is increasing evidence for the role of gut microbial composition in the pathogenesis of non-alcoholic fatty liver disease (NAFLD). Non-alcoholic steatohepatitis (NASH) is the most serious form of NAFLD where inflammation causes liver damage that can progress to cirrhosis. We have characterized the gut microbiome composition in UK patients with biopsy-proven NASH (n=65) and compared it to that in healthy controls (n=76). We report a 7% lower Shannon alpha diversity in NASH patients without cirrhosis (n=40) compared to controls ($p=2.7x10^{-4}$) and a 14% drop in NASH patients with cirrhosis (n=25, $p=5.0x10^{-4}$). Beta diversity (Unweighted UniFrac distance) was also significantly reduced in both NASH ($p=5.6x10^{-25}$) and NASH-cirrhosis ($p=8.1x10^{-7}$) groups. The genus most strongly associated with NASH in this study was *Collinsella* (0.29% abundance in controls, 3.45% in NASH without cirrhosis (False Discovery Rate (FDR) p=0.008), and 4.38% in NASH with cirrhosis (FDR p=0.02)). This genus, which has been linked previously to obesity and atherosclerosis, was also positively correlated with fasting levels of triglycerides (p=0.01) and total cholesterol ($p=1.2x10^{-4}$) and negatively correlated with high-density lipoprotein cholesterol ($p=2.8x10^{-6}$) suggesting that some of the pathways present in this microbial genus may influence lipid metabolism in the host. In patients we also found decreased abundance of some of the *Ruminococcaceae* which are known to produce high levels of short chain fatty acids which can lower inflammation. This may thus contribute to pathology associated with NASH.

Keywords: fatty liver; non-alcoholic steatohepatitis; cirrhosis; microbiome; Collinsella

Introduction

Non-alcoholic fatty liver disease (NAFLD) is the most common liver disease worldwide[1–3] and includes a spectrum of pathologies ranging from simple steatosis through non-alcoholic steatohepatitis (NASH), characterised by the presence of inflammation, to cirrhosis[4] which is responsible for 2% of all deaths worldwide[5]. Around one-third of those with early NASH will progress to advanced fibrosis and cirrhosis within 5-10 years[6], but a lack of early symptoms means that patients accumulate significant, sometimes irreversible liver damage by the time they are diagnosed. Therefore, understanding of contributing factors with potential as interventional treatment targets, is urgently sought.

There is a growing body of evidence linking the gut-liver axis to the development of NAFLD. NAFLD has been associated with an increase in intestinal permeability[7], and the subsequent infiltration of Gram-negative bacteria could induce liver inflammation via TLR4 signalling, enabling the release of TNF-α and the activation of hepatic stellate cells[8]. As the liver is linked to the intestine through the portal circulation, gut-derived products including both nutrients from the diet and microbial components will arrive first at the liver[9]. Pathogenic interactions between the liver and microbiota range from straightforward effects, such as increased intestinal permeability leading to passage of Gram-negative bacteria into the portal circulation and subsequent inflammation[10], to more complex interactions with metabolites produced by the microbiota, under the influence of modifiable factors such as diet. Since the establishment of a gut microbiome characteristic of obesity[11], similar links between the intestinal microbiota and the development of NASH have been investigated.

Small-scale studies of the microbiome in NASH have been somewhat contradictory: a study in 16 NASH patients identified a number of associated genera (*Parabacteroides* and *Allisonella*), and a longitudinal study has correlated a reduction of intrahepatic triglycerides with a reduction in *Firmicutes* and increase in *Bacteroidetes*[12]. This has been confirmed in a separate study using real-time PCR for 16S rRNA which also associated a drop in *Bacteroidetes* (in which *Prevotella* was used as a marker for the entire phylum) with NASH[13]. However, an increase in *Bacteroidetes* abundance was subsequently described in NASH patients[14].

Lower alpha diversity has consistently been found to be correlated with obesity and insulin resistance[15], higher visceral fat[16], and a number of inflammatory conditions[17,18]. Since inflammatory processes drive the pathology and progression of NAFLD and are linked to features of the metabolic syndrome including insulin resistance and visceral fat, we hypothesized that patients with NASH will display distinct changes in their microbiota and lower alpha diversity after adjustment for possible confounders such as BMI.

We describe significant differences in the microbiome in a substantial, well-phenotyped cohort of 65 biopsy-proven NASH patients compared to a control cohort of 76 participants, and identify a metabolic potential mechanism by evaluating diet and serum lipids.

Results

Clinical characteristics

The descriptive characteristics of NASH patients and controls are presented in Table 1. The healthy control group was slightly older on average compared to the NASH group, and was predominantly females. The control group had a significantly lower median BMI compared to the NASH group, therefore all analyses were adjusted for age, sex and BMI. All participants in the control group were of White ethnicity; 87% (33/40 NASH and 24/25 NASH-cirrhosis) of the patients were of White ethnicity. No significant differences in demographics were observed between NASH and NASH-cirrhosis groups. Within the NASH group, 1 patient had no fibrosis, 8 patients had fibrosis grade F1, 8 had grade F2, and 23 had grade F3.

Diversity metrics

A significant decrease in alpha diversity, indicated by Shannon index and operational taxonomic units (OTUs), was observed in both NASH and NASH-cirrhosis groups compared to controls (all $p<0.05$;Table 1; Figure 1). A significant decrease in beta diversity (Unweighted UniFrac), indicative of inter-individual diversity, was also observed between control and NASH groups, although it was slightly higher in the NASH-cirrhosis group compared to the NASH group (Figure 1).

OTUs significantly associated with NASH

Of the 124 OTUs used in the analysis, 27 were significantly associated with NASH after adjustment for covariates (age, gender and BMI) and false discovery rate (FDR), and 21 of these remained significant in the NASH-cirrhosis group. Several

OTUs were significantly associated with a higher likelihood of NASH: most significantly *Collinsella*, which was also significantly associated with the cirrhosis group (Figures 2 and 3). These OTUs remained associated when patients with white ethnicity only were analysed (data not shown).The genera most significantly associated with the control group relative to NASH were *Clostridium sensu stricto 1*, an unidentified genus belonging to the order *Mollicutes RF9*, *Alistipes* and several *Ruminococcaceae* (Figures 2 and 3).

Figure 4 displays the relative abundance of each OTU significantly associated with presence of NASH. *Collinsella* showed a 12-fold increase in relative abundance in the NASH group, increasing from 0.29% to 3.44%, and to 4.38% in NASH-cirrhosis patients. Conversely, the genera *Allistipes, Peptoclostridium, Paraprevotella* and two *Ruminococcaceae* showed statistically significant reductions in the both NASH groups relative to controls.

A number of genera had significantly reduced abundance in NASH-only participants. However, of these, *Actinomyces, Fusicatenibacter*, an unknnown *Lachnospiracae,* an unknown *Mollicutes RF9, Paraprevotella* and *Ruminococcaceae UCG-010* were not significantly altered in the cirrhosis group compared to controls. Aside from *Mollicutes RF9,* which had a similar relative abundance in the NASH-cirrhosis group to in controls, the remaining 5 genera all displayed similar changes in abundance to NASH patients without cirrhosis, but this did not meet the FDR-adjusted significance threshold.

Microbial associations with dietary components

We then assessed the link between nutrient intake and the abundance of taxa to investigate the potential impact of diet on microbes which could contribute to the development of NASH/NASH-cirrhosis. Figure 5 shows a heatmap summarizing Spearman correlations between nutrient intake calculated from a Food Frequency Questionnaire (FFQ) and the bacterial genera found to be associated with presence of NASH. The genera that showed significantly positive associations with fibre (NSP) were *Eggerthella* (r=0.646; p=0.047) and *Ruminococcaceae UCG-014* (r=0.853; p =0.003), whereas *Collinsella* (r =–0.785; p=0.015) showed a negative correlation. *Collinsella* showed a positive correlation (r=0.647; p=0.051) with total fat although it did not reach statistical significance, whereas *Alistipes* and *Eggerthella* were significantly correlated with fat intake (r=-0.517; p =0.040 and r=-0.717; p=0.033, respectively). The correlation between *Collinsella* and sugar was r=0.512 (p=0.062) and the correlation with total carbohydrates was r=0.326 (p=0.822).

Serum lipids

Based on our finding that the abundance of *Collinsella* is 12-fold higher in NASH patients and correlation with fat intake, we explored the association between *Collinsella* and circulating lipid sub-types to investigate any changes in fat metabolism that may arise which impact on disease pathways (Table 2). Fasting serum levels of triglycerides (TG), and saturated fatty acids were significantly higher in the NASH and NASH-cirrhosis group than in controls (Supplementary Table 1) whereas levels of low density lipoprotein cholesterol (LDL-C) and high density lipoprotein cholesterol (HDL-C) were lower in patients. Correlation between these lipids and presence or absence of NASH/NASH-cirrhosis was then assessed (Supplementary Table 2). HDL-C was the most strongly associated lipid with NAFLD showing a strong negative correlation

(Pearson's r= -0.75 $p=1.9x10^{-17}$), TG and LDL-C and fatty acids were also correlated with presence of disease.

We then tested whether any of these markers were also correlated with the abundance of the genus *Collinsella* (Table 2) and found that HDL-C was inversely associated with its abundance (r=-0.39, $p=2.8x10^{-6}$). A significant positive correlation was seen between the relative abundance of *Collinsella* and both serum TG (r=0.21, p=0.049) and total cholesterol (r=0.328, $p=1.2x10^{-4}$).

Discussion

Links between the microbiome and the development of NAFLD are now well established[19,20], but the details of changes in specific genera is lacking. We aimed to compare a substantial cohort of biopsy-proven NASH and NASH-cirrhosis patients, with a well characterised control cohort sourced from the TwinsUK study.

The genus most strongly associated with NASH in this study, *Collinsella*, has been previously linked to obesity and atherosclerosis in human studies[21–23], showing a 4-fold increase in abundance in obesity and 3.5-fold in type 2 diabetes[24]. It has also been found to be highly abundant in untreated patients with hypercholesterolaemia compared to controls and patients treated with statins[25]. NMR spectroscopy in a mouse model has shown a significant correlation between members of the *Coriobacteriaceae* family, which includes *Collinsella*, and hepatic triglycerides[26]. *Collinsella* has been linked to pro-inflammatory dysbiosis in type 2 diabetes[24] and with circulating insulin[21], suggestive of a mechanism for promotion of NAFLD pathology.

Insulin resistance and dyslipidaemia are characteristic of NAFLD[27,28]. Consistent with this, we show significant increases in total triglycerides, VLDL-C and saturated fatty acids, and concomitant reduction in HDL-C in both NASH and NASH-cirrhosis groups. We also report that the abundance of *Collinsella* is negatively correlated with serum HDL-C levels and positively correlated with TG, which suggests that some of the pathways present in this microbial genus may influence lipid

metabolism in the host. This is consistent with previous work in healthy adults correlating *Collinsella* abundance with total cholesterol[29], and in agreement with animal studies[26,30]. We observed a significant inverse correlation between the relative abundance of *Collinsella* and dietary fibre (Figure 5) as previously reported in obese women[21]. Dietary fibre and the gut derived metabolites that are associated with its breakdown (i.e. short chain fatty acids) have been implicated in the aetiology of metabolic disease including NAFLD[31]. Increasing dietary fibre intake could be considered a strategy for modulating the composition of the gut microbiome[32], reducing dysbiosis with the potential to reduce pathological features.

Collinsella spp. have been shown to metabolise bile acids to oxo-bile acid intermediates. Production of these secondary bile acids may increase intestinal permeability and contribute to the development of NAFLD[33,34]. Administration of the antibiotic rifaximin to decrease 7α-dehydroxylating bacteria (including *Collinsella*) has been linked to a drop in secondary fecal bile acid concentrations[35]. This provides a convincing mechanism by which *Collinsella* could mediate the progression of NAFLD, but further work is required to establish this link and determine the specific species or group of species that could be the cause.

We identified an OTU belonging to the putative order *Mollicutes RF9* significantly associated with the control cohort. Members of the *Mollicutes* class have previously been linked to obesity[36], but their functional roles are unclear[37]. Recent work[38] has correlated abundance of *Mollicutes RF39* with circulating indolepropionic acid, a potent antioxidant produced by the microbiota and associated with a lower risk of type II diabetes[39,40] thus providing a likely mechanism protective for NAFLD.

Several genera belonging to the order *Clostridiales* were also associated with the control group in this study. Of these, we found that *Ruminococcaceae* showed significant positive correlations with fibre intake. Several *Ruminococcaceae* have previously been inversely associated with cirrhosis in small studies including patients with a variety of aetiologies[41,42], and with NAFLD[43,44]. *Ruminococcaceae* are mainly responsible for fermenting fibre and other plant components of the diet such as inulin and cellulose, producing short-chain fatty acids (SCFA) which can both be utilised by the host for energy[45], and display anti-inflammatory properties in the gut[46].

We report a lower alpha diversity in NASH patients, further reduced in those with cirrhosis, similar to previous work comparing NAFLD patients without advanced fibrosis to those with cirrhosis [47]. We also demonstrate parallel changes in beta diversity, with a decline in our 40 NASH patients (of which 23 had advanced fibrosis) compared to controls but with no further reduction in cirrhotics. This indicates that the microbiome in the general population is both more diverse and more stable than that of both NASH and NASH-cirrhosis patients, with inter-individual variation increasing in cirrhosis. In agreement with earlier publications, we found a notable decrease in *Bacteroides* abundance in both the NASH group and NASH-cirrhosis groups. Unlike a previous report in paediatric NAFLD[48], we did not see a significant decrease in *Oscillspira*; increases in *Blautia* and *Ruminococcus* were observed but did not achieve statistical significance. Increases in *Blautia* and *Ruminococcus* were similarly observed in a cirrhosis group in the aforementioned Caussy *et al.* study[47], but not in their cohort without advanced fibrosis. This suggests that some of these changes may reflect

different stages of severity of liver disease and highlights the importance of in-depth hepatic phenotyping.

We acknowledge several limitations in our study. The control group is predominantly female whereas both NASH and cirrhosis groups are more evenly gender balanced. The effect of gender on the microbiome is well established[49,50] and may therefore have some effect on our results, but this has been adjusted for in all statistical comparisons. There are also acknowledged issues with using 16S reads as a proxy indicator for abundance[51] compared to absolute quantification with genus or species-specific PCR primers or a metagenomic sequencing approach. The study presented here is also purely observational; a germ-free animal model would be required to definitively verify the links we propose between *Collinsella* abundance, NASH and lipid metabolism. Furthermore, the associations between nutrients and the gut microbiome would require further evaluations to explore their interdependence in the aetiology of NASH and NASH-cirrhosis. The diet-gut microbiome associations presented are based on statistical associations from cross-sectional observations, and do not prove that causal relationships exist. Further explorations into causal impacts would require dietary interventions. In order to untangle the effect of nutrients on host-microbiome crosstalk, future studies should also consider looking at fecal metabolites such as SCFAs. Fecal bile acid data was also not available for this study; this would be required to study the metabolic effects of *Collinsella* in more detail, as well as the gut-liver axis in general[52]. Thus, our study could be improved by looking at bacterial metabolic responses to diet which may be more informative in assessing the effect of certain nutrients on health, compared to the gut microbiota itself.

Overall, we have determined the relationship between gut microbiome composition and diversity and disease in the largest biopsy-proven NASH cohort to date. We confirm the findings from previous studies[14,47] showing a reduced alpha diversity in NASH patients and further decrease in cirrhosis patients. We also report, for the first time, that the genus *Collinsella* is significantly increased in NASH patients with and without cirrhosis. This association with pathology is suggested to be linked with effects on host lipid metabolism and diet.

Materials and methods

Study participants

The 65 adult patients included in this study were prospectively recruited from hepatology clinics at Nottingham University Hospitals NHS Trust. Patients gave informed consent to participate in research studies in accordance with the 1975 Declaration of Helsinki. The studies were approved by NHS Regional Ethics Committees (REC reference 12/WM/0288 and GM010201). NASH was diagnosed on the basis of the following criteria: appropriate exclusion of other causes of liver disease including alcohol, drugs, autoimmune or viral hepatitis, or cholestatic or metabolic/genetic liver disease; a weekly ethanol consumption of <140g in women and <210g in men and a liver biopsy showing steatohepatitis with or without cirrhosis. The Fatty Liver Inhibition of Progression (FLIP) Consortium algorithm was used as a diagnostic indicator where histological diagnosis of NASH requires presence of steatosis, ballooning and lobular inflammation, and those without all three features are identified as having fatty liver while those lacking steatosis (<5% hepatocytes with fat accumulation) are not considered as having NASH[53]. Fibrosis was scored following the

CRN grading system[54] by a single pathologist. Patients were stratified into two groups: NASH (40 patients without cirrhosis) and NASH with cirrhosis (25 patients). All patients in the NASH-cirrhosis group had compensated cirrhosis at the time of sample collection.

76 healthy controls were included from the UK Adult Twin Registry (TwinsUK), a large cohort of volunteer adult twins from the United Kingdom hosted by the Department of Twin Research and Genetic Epidemiology at St. Thomas' Hospital, King's College London [55].This cohort contained 13 twin pairs (9 monozygotic pairs and 4 dizygotic pairs); all others had a twin who was not included.

Sequencing

Fecal samples were provided at study visits and immediately frozen at -80°C. DNA was extracted by homogenising in lysis buffer (500mM NaCl, 50mM Tris-HCl, 50mM EDTA, 4% SDS) with bead beating (MagNA Lyser, Roche) followed by extraction using the QIAmp DNA Stool Mini Kit (Qiagen; cat 51504). DNA was diluted to 20ng/µl for 16S rRNA amplification and sequencing at the Genetic Laboratory, Erasmus Medical Center, Rotterdam, the Netherlands. Water negative controls were included from extraction, through PCR to sequencing and select samples were sequenced in duplicate for quality control.

The V4 region of the 16S ribosomal RNA gene was amplified using universal primers 355F (CCAGACTCCTACGGGAGGCAGC) and 806R (GGACTACHVGGGTWTCTAAT). Amplified DNA was sequenced on the MiSeq platform (Illumina). Read filtering and clustering was carried out using the MYcrobiota pipeline[56]. Briefly, chimeric sequences were filtered using the VSEARCH algorithm within Mothur, and reads were clustered into operational taxonomic units (OTUs) using

closed-reference clustering against the SILVA database v132 based on a 97% similarity. Diversity metrics (Shannon index, observed OTUs and Unweighted UniFrac) were calculated by rarefying the OTU table down to 7000 sequences per sample 50 times and taking the average. These analyses were carried out in QIIME 2 (v2018.11).

Lipid measurements

Serum lipids were quantified using a high-throughput ^{1}H NMR platform (Nightingale Health, Helsinki, Finland) using 500Mhz and 600MhH protein nuclear magnetic resonance spectroscopy as described previously[57,58].

Diet-microbiome analysis

Self-reported dietary intake was assessed by food frequency questionnaires following the EPIC-Norfolk guidelines[59]. Individuals answered questions on intake of 131 food groups and nutrient intakes (weight/day) were derived from the UK Nutrient Database using the FETA software (version 2)[60]. A total of thirty-three nutrients were estimated from this calculation. Associations between relative abundances of bacterial taxa shown to be significantly associated with NASH or the control group, and nutrient intakes (adjusted for total energy (kcal)) intake were examined by Spearman correlation.

Statistical analysis

OTUs with a relative abundance of <0.1% in every sample were removed, and relative OTU abundances were inverse normal transformed before further analysis. Associations between NASH and NASH-cirrhosis groups and OTU abundance at genus level were predicted using a general linear model adjusted for age, gender and BMI. OTUs were assumed to be significantly associated with NASH or cirrhosis with a p

value <0.05 after adjusting for FDR. All statistical analyses were carried out in R v3.5.2.

References

1. Browning JD, Cohen JC, Hobbs HH. Patatin-like phospholipase domain-containing 3 and the pathogenesis and progression of pediatric nonalcoholic fatty liver disease. Hepatology 2010; 52:1189–92.

2. Chalasani N, Younossi Z, Lavine JE, Diehl AM, Brunt EM, Cusi K, Charlton M, Sanyal AJ. The diagnosis and management of non-alcoholic fatty liver disease: Practice Guideline by the American Association for the Study of Liver Diseases, American College of Gastroenterology, and the American Gastroenterological Association. Hepatology 2012; 55:2005–23.

3. Kelly T, Yang W, Chen C-S, Reynolds K, He J. Global burden of obesity in 2005 and projections to 2030. Int J Obes 2008; 32:1431–7.

4. Higuera-de la Tijera F, Servín-Caamaño AI. Pathophysiological mechanisms involved in non-alcoholic steatohepatitis and novel potential therapeutic targets. World J Hepatol 2015; 7:1297–301.

5. Murray CJL, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, Ezzati M, Shibuya K, Salomon JA, Abdalla S, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet Lond Engl 2012; 380:2197–223.

6. Calzadilla Bertot L, Adams LA. The Natural Course of Non-Alcoholic Fatty Liver Disease. Int J Mol Sci [Internet] 2016 [cited 2019 Jul 18]; 17. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4881593/

7. Miele L, Valenza V, La Torre G, Montalto M, Cammarota G, Ricci R, Mascianà R, Forgione A, Gabrieli ML, Perotti G, et al. Increased intestinal permeability and tight junction alterations in nonalcoholic fatty liver disease. Hepatology 2009; 49:1877–87.

8. Csak T, Ganz M, Pespisa J, Kodys K, Dolganiuc A, Szabo G. Fatty acid and endotoxin activate inflammasomes in mouse hepatocytes that release danger signals to stimulate immune cells. Hepatology 2011; 54:133–44.

9. Brandl K, Kumar V, Eckmann L. Gut-liver axis at the frontier of host-microbial interactions. Am J Physiol-Gastrointest Liver Physiol 2017; 312:G413–9.

10. Moreira APB, Texeira TFS, Ferreira AB, Peluzio M do CG, Alfenas R de CG. Influence of a high-fat diet on gut microbiota, intestinal permeability and metabolic endotoxaemia. Br J Nutr 2012; 108:801–9.

11. Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JI. An obesity-associated gut microbiome with increased capacity for energy harvest. Nature 2006; 444:1027–131.

12. Wong VW-S, Tse C-H, Lam TT-Y, Wong GL-H, Chim AM-L, Chu WC-W, Yeung DK-W, Law PT-W, Kwan H-S, Yu J, et al. Molecular Characterization of the Fecal Microbiota in Patients with Nonalcoholic Steatohepatitis – A Longitudinal Study. PLoS ONE 2013; 8:e62885.

13. Mouzaki M, Comelli EM, Arendt BM, Bonengel J, Fung SK, Fischer SE, McGilvray ID, Allard JP. Intestinal microbiota in patients with nonalcoholic fatty liver disease. Hepatology 2013; 58:120–7.

14. Zhu L, Baker SS, Gill C, Liu W, Alkhouri R, Baker RD, Gill SR. Characterization of gut microbiomes in nonalcoholic steatohepatitis (NASH) patients: a connection between endogenous alcohol and NASH. Hepatology 2013; 57:601–609.

15. Le Chatelier E, Nielsen T, Qin J, Prifti E, Hildebrand F, Falony G, Almeida M, Arumugam M, Batto J-M, Kennedy S, et al. Richness of human gut microbiome correlates with metabolic markers. Nature 2013; 500:541–6.

16. Beaumont M, Goodrich JK, Jackson MA, Yet I, Davenport ER, Vieira-Silva S, Debelius J, Pallister T, Mangino M, Raes J, et al. Heritable components of the human fecal microbiome are associated with visceral fat. Genome Biol 2016; 17:189.

17. Morgan XC, Tickle TL, Sokol H, Gevers D, Devaney KL, Ward DV, Reyes JA, Shah SA, LeLeiko N, Snapper SB, et al. Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. Genome Biol 2012; 13:R79.

18. Scher JU, Sczesnak A, Longman RS, Segata N, Ubeda C, Bielski C, Rostron T, Cerundolo V, Pamer EG, Abramson SB, et al. Expansion of intestinal Prevotella copri correlates with enhanced susceptibility to arthritis. eLife [Internet] 2013 [cited 2013 Nov 13]; 2. Available from: http://elife.elifesciences.org/content/2/e01202

19. Schnabl B, Brenner DA. Interactions Between the Intestinal Microbiome and Liver Diseases. Gastroenterology 2014; 146:1513–24.

20. Safari Z, Gérard P. The links between the gut microbiome and non-alcoholic fatty liver disease (NAFLD). Cell Mol Life Sci 2019; 76:1541–58.

21. Gomez-Arango LF, Barrett HL, Wilkinson SA, Callaway LK, McIntyre HD, Morrison M, Nitert MD. Low dietary fiber intake increases Collinsella abundance in the gut microbiota of overweight and obese pregnant women. Gut Microbes 2018; 9:189–201.

22. Karlsson FH, Fåk F, Nookaew I, Tremaroli V, Fagerberg B, Petranovic D, Bäckhed F, Nielsen J. Symptomatic atherosclerosis is associated with an altered gut metagenome. Nat Commun 2012; 3:1245.

23. Kourosh A, Luna RA, Balderas M, Nance C, Anagnostou A, Devaraj S, Davis CM. Fecal microbiome signatures are different in food-allergic children compared to siblings and healthy children. Pediatr Allergy Immunol Off Publ Eur Soc Pediatr Allergy Immunol 2018; 29:545–54.

24. Candela M, Biagi E, Soverini M, Consolandi C, Quercia S, Severgnini M, Peano C, Turroni S, Rampelli S, Pozzilli P, et al. Modulation of gut microbiota dysbioses in type 2 diabetic patients by macrobiotic Ma-Pi 2 diet. Br J Nutr 2016; 116:80–93.

25. Khan TJ, Ahmed YM, Zamzami MA, Siddiqui AM, Khan I, Baothman OAS, Mehanna MG, Kuerban A, Kaleemuddin M, Yasir M. Atorvastatin Treatment Modulates the Gut Microbiota of the Hypercholesterolemic Patients. Omics J Integr Biol 2018; 22:154–63.

26. Claus SP, Ellero SL, Berger B, Krause L, Bruttin A, Molina J, Paris A, Want EJ, Waziers I de, Cloarec O, et al. Colonization-Induced Host-Gut Microbial Metabolic Interaction. mBio 2011; 2:e00271-10.

27. Sanyal AJ, Campbell–Sargent C, Mirshahi F, Rizzo WB, Contos MJ, Sterling RK, Luketic VA, Shiffman ML, Clore JN. Nonalcoholic steatohepatitis: Association of insulin resistance and mitochondrial abnormalities. Gastroenterology 2001; 120:1183–92.

28. Katsiki N, Mikhailidis DP, Mantzoros CS. Non-alcoholic fatty liver disease and dyslipidemia: An update. Metabolism 2016; 65:1109–23.

29. Lahti L, Salonen A, Kekkonen RA, Salojärvi J, Jalanka-Tuovinen J, Palva A, Orešič M, de Vos WM. Associations between the human intestinal microbiota, Lactobacillus rhamnosus GG and serum lipids indicated by integrated analysis of high-throughput profiling data. PeerJ 2013; 1:e32.

30. Martínez I, Wallace G, Zhang C, Legge R, Benson AK, Carr TP, Moriyama EN, Walter J. Diet-Induced Metabolic Improvements in a Hamster Model of Hypercholesterolemia Are Strongly Linked to Alterations of the Gut Microbiota. Appl Environ Microbiol 2009; 75:4175–84.

31. Canfora EE, Meex RCR, Venema K, Blaak EE. Gut microbial metabolites in obesity, NAFLD and T2DM. Nat Rev Endocrinol 2019; 15:261–73.

32. Makki K, Deehan EC, Walter J, Bäckhed F. The Impact of Dietary Fiber on Gut Microbiota in Host Health and Disease. Cell Host Microbe 2018; 23:705–15.

33. Doden H, Sallam LA, Devendran S, Ly L, Doden G, Daniel SL, Alves JMP, Ridlon JM. Metabolism of Oxo-Bile Acids and Characterization of Recombinant 12α-Hydroxysteroid Dehydrogenases from Bile Acid 7α-Dehydroxylating Human Gut Bacteria. Appl Environ Microbiol 2018; 84.

34. Stenman LK, Holma R, Eggert A, Korpela R. A novel mechanism for gut barrier dysfunction by dietary fat: epithelial disruption by hydrophobic bile acids. Am J Physiol-Gastrointest Liver Physiol 2012; 304:G227–34.

35. Kakiyama G, Pandak WM, Gillevet PM, Hylemon PB, Heuman DM, Daita K, Takei H, Muto A, Nittono H, Ridlon JM, et al. Modulation of the fecal bile acid profile by gut microbiota in cirrhosis. J Hepatol 2013; 58:949–55.

36. Turnbaugh PJ, Bäckhed F, Fulton L, Gordon JI. Diet-Induced Obesity Is Linked to Marked but Reversible Alterations in the Mouse Distal Gut Microbiome. Cell Host Microbe 2008; 3:213–23.

37. Turnbaugh PJ. Microbes and Diet-Induced Obesity: Fast, Cheap, and Out of Control. Cell Host Microbe 2017; 21:278–81.

38. Menni C, Hernandez MM, Vital M, Mohney RP, Spector TD, Valdes AM. Circulating levels of the anti-oxidant indoleproprionic acid are associated with higher gut microbiome diversity. Gut Microbes 2019; :1–8.

39. de Mello VD, Paananen J, Lindström J, Lankinen MA, Shi L, Kuusisto J, Pihlajamäki J, Auriola S, Lehtonen M, Rolandsson O, et al. Indolepropionic acid and novel lipid metabolites are associated with a lower risk of type 2 diabetes in the Finnish Diabetes Prevention Study. Sci Rep 2017; 7:46337.

40. Valdes AM, Walter J, Segal E, Spector TD. Role of the gut microbiota in nutrition and health. The BMJ [Internet] 2018 [cited 2019 Apr 2]; 361. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6000740/

41. Bajaj JS, Ridlon JM, Hylemon PB, Thacker LR, Heuman DM, Smith S, Sikaroodi M, Gillevet PM. Linkage of gut microbiome with cognition in hepatic encephalopathy. Am J Physiol-Gastrointest Liver Physiol 2011; 302:G168–75.

42. Bajaj JS, Hylemon PB, Ridlon JM, Heuman DM, Daita K, White MB, Monteith P, Noble NA, Sikaroodi M, Gillevet PM. Colonic mucosal microbiome differs from stool microbiome in cirrhosis and hepatic encephalopathy and is linked to cognition and inflammation. Am J Physiol-Gastrointest Liver Physiol 2012; 303:G675–85.

43. Jiang W, Wu N, Wang X, Chi Y, Zhang Y, Qiu X, Hu Y, Li J, Liu Y. Dysbiosis gut microbiota associated with inflammation and impaired mucosal immune function in intestine of humans with non-alcoholic fatty liver disease. Sci Rep 2015; 5:8096.

44. Raman M, Ahmed I, Gillevet PM, Probert CS, Ratcliffe NM, Smith S, Greenwood R, Sikaroodi M, Lam V, Crotty P, et al. Fecal Microbiome and Volatile Organic Compound Metabolome in Obese Humans With Nonalcoholic Fatty Liver Disease. Clin Gastroenterol Hepatol 2013; 11:868-875.e3.

45. Scheppach W, Weiler F. The butyrate story: old wine in new bottles? Curr Opin Clin Nutr Metab Care 2004; 7:563.

46. Kles KA, Chang EB. Short-Chain Fatty Acids Impact on Intestinal Adaptation, Inflammation, Carcinoma, and Failure. Gastroenterology 2006; 130:S100–5.

47. Caussy C, Tripathi A, Humphrey G, Bassirian S, Singh S, Faulkner C, Bettencourt R, Rizo E, Richards L, Xu ZZ, et al. A gut microbiome signature for cirrhosis due to nonalcoholic fatty liver disease. Nat Commun 2019; 10:1406.

48. Chierico FD, Nobili V, Vernocchi P, Russo A, Stefanis CD, Gnani D, Furlanello C, Zandonà A, Paci P, Capuani G, et al. Gut microbiota profiling of pediatric nonalcoholic fatty liver disease and obese patients unveiled by an integrated meta-omics-based approach. Hepatology 2017; 65:451–64.

49. Haro C, Rangel-Zúñiga OA, Alcalá-Díaz JF, Gómez-Delgado F, Pérez-Martínez P, Delgado-Lista J, Quintana-Navarro GM, Landa BB, Navas-Cortés JA, Tena-Sempere M, et al. Intestinal Microbiota Is Influenced by Gender and Body Mass Index. PLOS ONE 2016; 11:e0154090.

50. Falony G, Joossens M, Vieira-Silva S, Wang J, Darzi Y, Faust K, Kurilshikov A, Bonder MJ, Valles-Colomer M, Vandeputte D, et al. Population-level analysis of gut microbiome variation. Science 2016; 352:560–4.

51. Tkacz A, Hortala M, Poole PS. Absolute quantitation of microbiota abundance in environmental samples. Microbiome 2018; 6:110.

52. Liu H-X, Keane R, Sheng L, Wan Y-JY. Implications of microbiota and bile acid in liver injury and regeneration. J Hepatol 2015; 63:1502–10.

53. Bedossa P, Poitou C, Veyrie N, Bouillot J-L, Basdevant A, Paradis V, Tordjman J, Clement K. Histopathological algorithm and scoring system for evaluation of liver lesions in morbidly obese patients. Hepatol Baltim Md 2012; 56:1751–9.

54. Brunt EM, Kleiner DE, Wilson LA, Belt P, Neuschwander-Tetri BA, NASH Clinical Research Network (CRN). Nonalcoholic fatty liver disease (NAFLD) activity score and the histopathologic diagnosis in NAFLD: distinct clinicopathologic meanings. Hepatol Baltim Md 2011; 53:810–20.

55. Moayyeri A, Hammond CJ, Valdes AM, Spector TD. Cohort Profile: TwinsUK and healthy ageing twin study. Int J Epidemiol 2013; 42:76–85.

56. Boers SA, Hiltemann SD, Stubbs AP, Jansen R, Hays JP. Development and evaluation of a culture-free microbiota profiling platform (MYcrobiota) for clinical diagnostics. Eur J Clin Microbiol Infect Dis Off Publ Eur Soc Clin Microbiol 2018; 37:1081–9.

57. Valdes AM, Ravipati S, Pousinis P, Menni C, Mangino M, Abhishek A, Chapman V, Barrett DA, Doherty M. Omega-6 oxylipins generated by soluble epoxide hydrolase are associated with knee osteoarthritis. J Lipid Res 2018; 59:1763–70.

58. Soininen P, Kangas AJ, Würtz P, Tukiainen T, Tynkkynen T, Laatikainen R, Järvelin M-R, Kähönen M, Lehtimäki T, Viikari J, et al. High-throughput serum NMR metabonomics for cost-effective holistic studies on systemic metabolism. Analyst 2009; 134:1781–5.

59. EPIC-Norfolk nutritional methods: food frequency questionnaire [Internet]. [cited 2019 Jul 10]; Available from: http://www.srl.cam.ac.uk/epic/nutmethod/FFQ.shtml

60. FETA software for analysis of EPIC-Norfolk food frequency questionnaire data [Internet]. MRC Epidemiol. Unit [cited 2019 Aug 14]; Available from: http://www.mrc-epid.cam.ac.uk/ffq-epic-tool-for-analysis-feta/

Figure legends

Figure 1: Diversity metrics in control, NASH and NASH-cirrhosis groups. **A**: Alpha diversity measured using the Shannon index, **B**: Beta diversity measured using Unweighted UniFrac distances between all samples in each group. Diversity indicies were calculated using the OTU table rarefied to 7000 sequences per sample 50 times and taking the average.

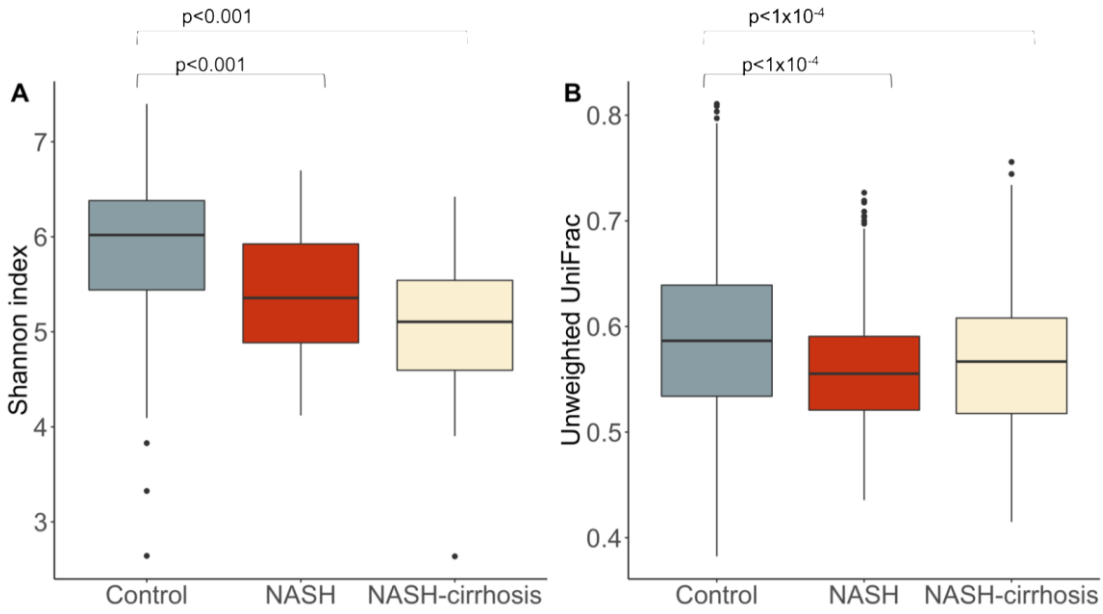


Figure 2: Heat map showing all OTUs (clustered at genus level) significantly associated with NASH or the control group and association with presence of NASH, NASH-cirrhosis, BMI, hypertension, total fasted triglycerides (TG), high density lipoprotein (HDL), low density lipoprotein (LDL) and fasted glucose. Values are beta coefficients from linear models adjusted for BMI, age and gender (associations between genera and BMI are adjusted for age and gender only). P values shown in parentheses are adjusted for FDR.

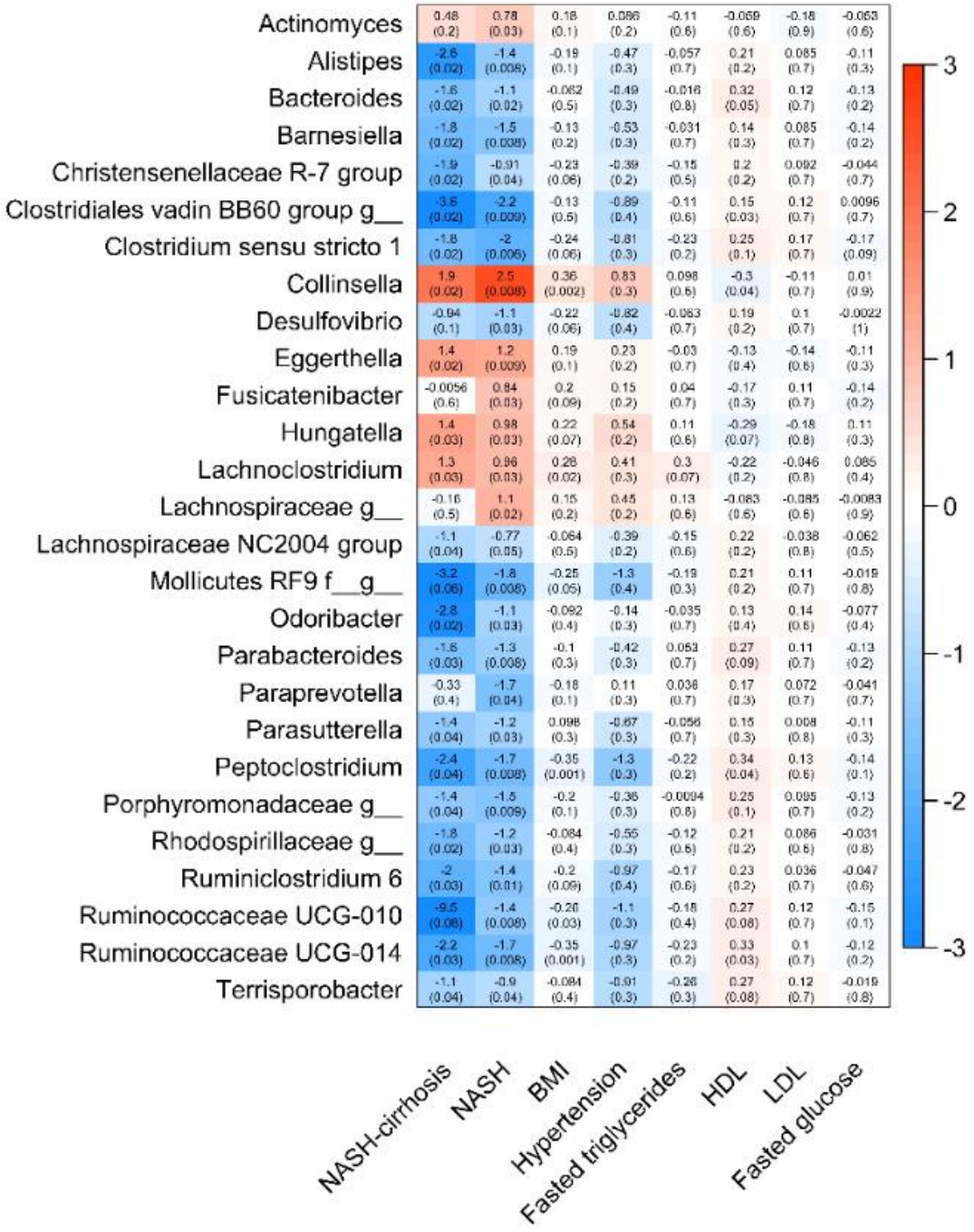


Figure 3: Forest plot of effect sizes with 95% confidence intervals showing association of all significant (FDR adjusted p <0.05) OTUs with NASH and NASH-cirrhosis compared to control cohort. Smaller dots on the NASH-cirrhosis arm indicate a loss of statistical significance.

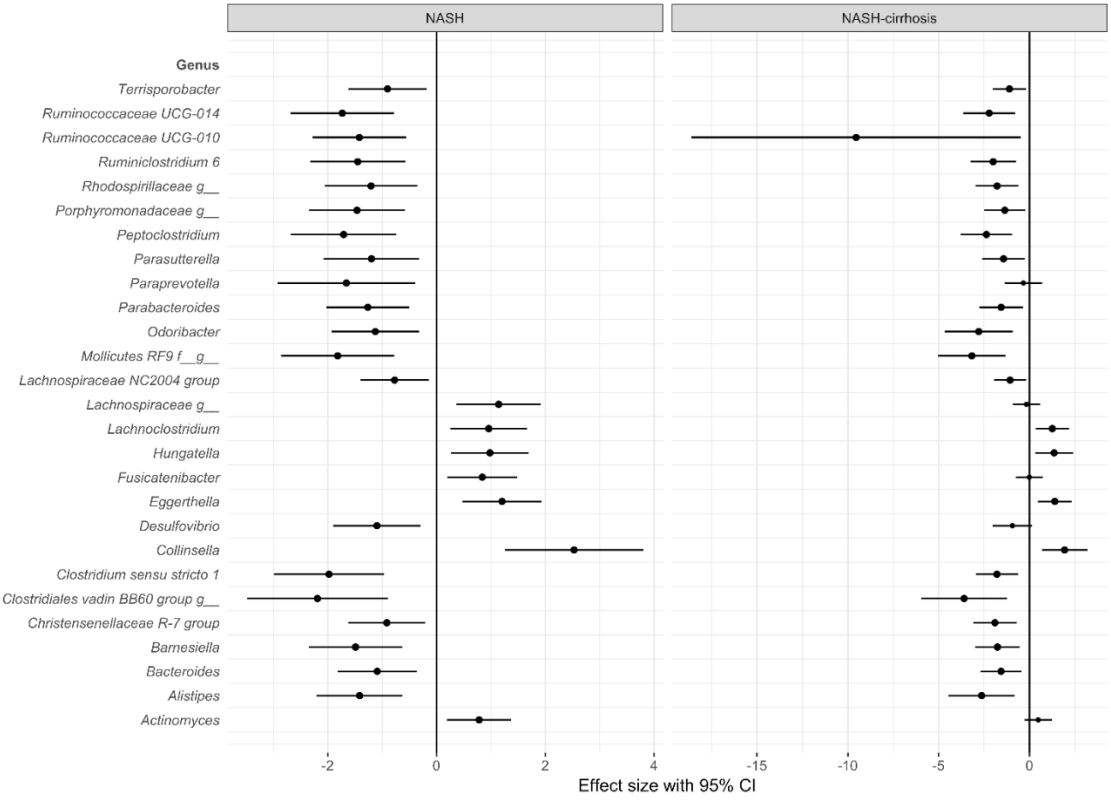


Figure 4: Relative abundance of all OTUs significantly associated (FDR p<0.05) with the NASH group. *p<0.05, **p<0.01.

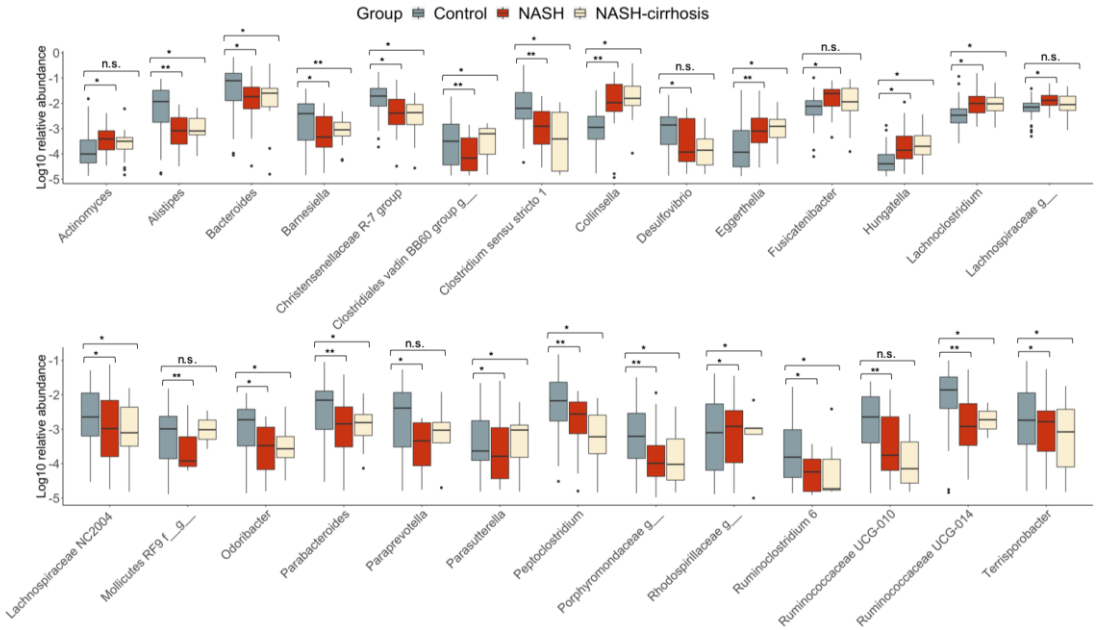
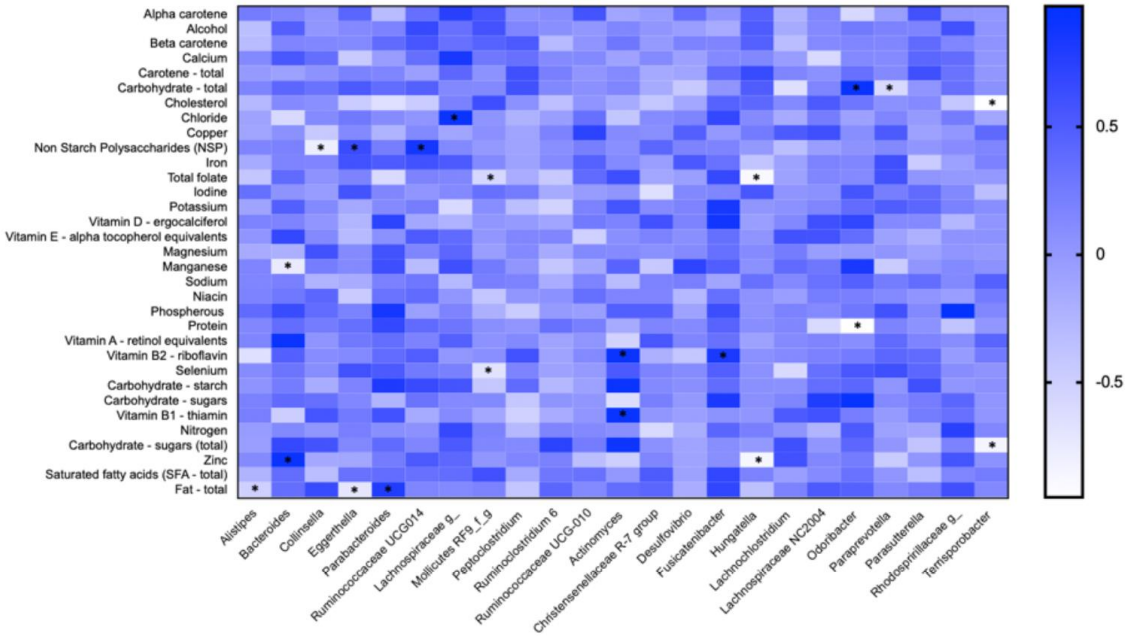


Figure 5: Heatmap depicting the correlation between the abundance of genera significantly associated with NASH and NASH-cirrhosis and dietary components. The intensity of the colour represents the degree of association between bacterial abundances and nutrients as measured by Spearman's correlations. The asterisk symbols indicate the associations that are significant after adjusting for FDR.



Tables

Table 1: Characteristics of the study population**.**

Characteristic	**Healthy (n=76)**	**NASH (n=40)**	**NASH-cirrhosis (n=25)**
Sex (female/male)	67/9	20/20	15/10
Mean age (mean±SEM)	66.3±1.06	60.3±1.6[a]	64.0±2.0
Mean BMI (kg/m^2, mean±SEM)	26.5±0.5	34.9±1.1[a]	34.0±1.1[b]
Type 2 Diabetes	0	22[a]	19[b]
Hypertension	1	19[a]	14[b]
Shannon diversity (mean±SEM)	5.83±0.1	5.32±0.1[a]	5.18±0.2[b]
Observed OTUs (mean±SEM)	405.3±10.9	334.8±9.5[a]	313.7±16.9[b]

NASH: non-alcoholic steatohepatitis; SEM: standard error of the mean; OTUs: operational taxonomic unit.[a]$p<0.05$ Healthy vs NASH, [b]$p<0.05$, Healthy vs NASH-cirrhosis, [c]$p<0.05$ NASH vs NASH-cirrhosis, ANOVA with *a priori* contrasts or Kruskal-Wallace followed by Mann-Whitney U test as appropriate based on distribution.

Table 2: Correlation between relative abundance of *Collinsella* and serum lipids in the total study population.

Metabolite	**Correlation with *Collinsella* relative abundance (Pearson's R and p value)**
Triglycerides (mmol/L)	0.22 (0.01)
Total cholesterol (mmol/L)	0.328 ($1.2x10^{-4}$)
VLDL-C (mmol/L)	0.14 (0.11)
HDL-C (mmol/L)	-0.39 ($2.8x10^{-6}$)
LDL-C (mmol/L)	-0.16 (0.071)
Saturated fatty acids (mmol/L)	0.09 (0.32)
Remnant cholesterol (mmol/L)	0.0021 (0.98)

VLDL-C:very low density cholesterol;HDL-C: high density lipoprotein cholesterol;

LDL-C: low density lipoprotein cholesterol.

Acknowledgements

We are grateful to Melanie Lingaya for technical assistance including DNA extraction. We thank the clinical fellows (Naaventhan Palaniyappan and Robert Scott) and nurse team at the NIHR Nottingham Biomedical Research Centre for assisting with recruitment. Sample collection was funded by the National Institute for Health Research (NIHR) Nottingham Biomedical Research Centre (reference no: BRC-1215-20003). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health. The team at the Genetic Laboratory, Erasmus

Medical Center, Rotterdam, are acknowledged for 16S amplification and sequencing. We thank Philip Kaye (Department of Histopathology, Nottingham University Hospitals NHS Trust) for scoring all patient biopsies.

Funding

This work was funded by the NIHR Nottingham Biomedical Research Centre (BRC-1215-20003, flexifund award), the Nottingham Molecular Pathology Node (MRC/EPSRC grant MR/N005953/1), the MRC AimHy study (MR/M016560/1) and a CDRF project grant. TwinsUK receives funding from the Wellcome Trust and the EU Seventh Framework Programme. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Contributions

JIG, AMV, GPA: Obtained funding, developed concept and designed the study. JIG: study co-ordination. SA, EA, AV, JIG, AMV: Analysed and interpreted the data and drafted the manuscript. All authors reviewed and revised the final manuscript.

Disclosure statement

AMV is a consultant for Zoe Global Ltd.