Dietary Cholesterol Drives Fatty Liver-associated Liver Cancer by Modulating Gut Microbiota and Metabolites

Running title: Cholesterol, gut microbiome and NAFLD-HCC

Supplementary Methods

Gut microbiota analysis

To determine dietary cholesterol-induced gut microbiota alterations, DNA was extracted from fecal samples using QIAamp DNA Mini Kit. Amplicon library for bidirectional (2×250 bp) sequencing on Illumina MiSeq platform was constructed using universal primers 515f, 5′-GTGCCAGCMGCCGCGGTAA-3′ and 806r, 5′-GGACTACHVGGGTWTCTAAT-3′ targeted across 16S rRNA genes V4 hypervariable regions. Library clean-up and normalisation was performed using SequelPrep Normalization Plate Kit (Thermo Fisher Scientific, Waltham, MA) according to the manufacturer's instructions. Quality filtering and analysis of the 16S rRNA gene sequence data were performed with the Mothur software suite as previously described.[1] Briefly, Needleman-Wunsch alignment algorithm with default parameters, was used to merge paired-end reads into contigs, followed by alignment against SILVA 16S rRNA sequence database (version 123) using NAST algorithm.[2] Contigs that mapped outside V4 region were discarded. The remaining sequences were merged with more abundant sequences having a maximum difference of two nucleotide bases and then screened for chimeric sequences with de novo Uchime.[3] The resulting sequences were assigned to Greengenes taxa (version 13.8). Sequences classified as eukarya, archaea, mitochondria, chloroplast and unknown kingdoms were discarded. The final sequences were assigned to clustered into

Zhang X, *et al*. *Gut* 2021; 70:761–774. doi: 10.1136/gutjnl-2019-319664

operational taxonomy units using the opticlust algorithm of Mothur software suit.[1] An average of 43 763 ± 17155 reads per sample were obtained after quality control steps. Sequences were rarefied to 11 479 sequences to minimize effects of uneven sampling in downstream analysis. Spearman correlation analysis was performed to illustrate the correlation of cholesterol with gut microbiota and metabolites. Functional capacity of the gut microbiota was predicted using PICRUSt,[4] an algorithm that estimates the functional potential of microbial communities given a marker gene survey and a set of sequenced reference genome. Metagenomic sequencing of human stool was performed as previously described.[5]

Metabolomic analysis

LC-MS/MS analyses were performed using an UHPLC system (1290, Agilent Technologies, Santa Clara, CA) with a UPLC BEH Amide column (1.7 rs of eukarya, archaea, mitochondTripleTOF 5600 (Q-TOF, AB Sciex, Toronto, Canada). The mobile phase consisted of 25 mM NH4OAc and 25 mM NH4OH in water (pH = 9.75). Triple TOF mass spectrometer was used for its ability to acquire MS/MS spectra on an information-dependent basis during an LC/MS experiment. In this mode, the acquisition software (Analyst TF 1.7, AB Sciex) continuously evaluates the full scan survey MS data as it collects and triggers the acquisition of MS/MS spectra depending on preselected criteria. In each cycle, 12 precursor ions with intensity greater than 100 were chosen for fragmentation at collision energy of 30V (15 MS/MS events with product ion accumulation time of 50 msec each).

Zhang X, *et al*. *Gut* 2021; 70:761–774. doi: 10.1136/gutjnl-2019-319664

MS raw data files were converted to the mzXML format using ProteoWizard and processed by R package XCMS (version 3.2). The preprocessing generated a data matrix that consisted of the retention time, mass-to-charge ratio (m/z) values, and peak intensity. R package CAMERA was used for peak annotation after XCMS data processing. In-house MS2 database was applied in metabolites identification.

Serum detection

Serum was obtained by centrifuging whole blood samples. Serum level of alpha-fetoprotein (AFP) was measured using Mouse alpha fetoprotein/AFP Quantitikine enzyme-linked immunosorbent assay (ELISA) kit (R&D Systems Inc., Minneapolis, MN) according to manufacturer's instructions. The levels of serum alanine aminotransferase (ALT), aspartate aminotransferase (AST), cholesterol and glucose in serum were determined using Catalyst One Chemistry Analyzer (IDEXX Laboratories, Westbrook, ME). Serum insulin was determined by Mouse Insulin ELISA Kit (Millipore, Billercia, MA).

For glucose tolerance test, mice were fasted overnight by transferring mice to clean cages with no food or feces. Mice were then injected intraperitoneally with 20% glucose solution (2 g/kg body weight glucose) in water. Blood from the tail vein was obtained before, and at 30, 60, 90 and 120 min after glucose injection for determination of blood glucose using Glucose Meter.

Histological analysis

Liver histology was assessed by H&E staining of paraffin-embedded sections

Zhang X, *et al*. *Gut* 2021; 70:761–774. doi: 10.1136/gutjnl-2019-319664

as previously described.[6] Two investigators who were blinded to the treatment independently evaluated the slides and assigned scores for steatosis and inflammation. Steatosis was scored by low to medium power evaluation of parenchymal involvement according to the following criteria: 0 (<5%), 1 (5%-33%), 2 (33%-66%), or 3 (>66%). Inflammation was scored by overall assessment of all inflammatory foci according to the following criteria: 0 (No foci), 1 (<2 foci per 200X field), 2 (2-4 foci per 200X field), 3 (>4 foci per 200X field). Fibrosis was assessed by Sirius Red staining of paraffin embedded sections. Quantitative morphometric measurements were processed. The presence of steatosis was further confirmed using Oil Red O staining of frozen sections. The liver histology was evaluated by Prof. Anthony Chan, who is a liver pathology expert and blinded to the treatment conditions.

Magnetic resonance imaging of mice liver

Mice were anesthetized with 10 mg/kg xylazine (Rompun, Bayer HealthCare, Leverkusen, Germany) and 90 mg/kg ketamine (Ketalar, Pfizer, Hong Kong SAR, China) and scanned using a Philips Achieva 3T scanner (Philips Healthcare, Best, the Netherlands), and a custom-made dedicated mouse body-size radiofrequency coil as the transmitter and receiver. Anatomical imaging included fast spin echo T1-weighted (TR=500 msec, TE=75 msec) and T2 weighted images (TR=2000 msec, TE=100 msec). The slice thickness was 1mm and the in-plane resolution was 0.2×0.2mm.

Cytokine profiling assay

Zhang X, *et al*. *Gut* 2021; 70:761–774. doi: 10.1136/gutjnl-2019-319664

Mouse serum and hepatic inflammatory cytokine levels were measured by MILLIPLEX MAP Mouse Cytokine/Chemokine-Premixed 22 Plex (Millipore). Plates were read on the Bio-Plex 200 System (Bio-Rad Laboratories Inc., Hercules, CA) and the concentrations of 22 mouse inflammatory cytokine/chemokine were calculated. Hepatic IL-6, IL-1α and IL-1β levels were also measured by ELISA (Cusabio, Wuhan, China).

Hepatic measurement of cholesterol, NAD+/NADH, SOD activity, hydroxyproline, triglyceride and lipid peroxidation

Cholesterol/Cholesteryl Ester Quantification Kit (ab65359) (Abcam, Cambridge, MA) was used for the detection of hepatic cholesterol and cholesteryl ester levels. NAD+/NADH ratio was determined by NAD/NADH assay kit (Abcam). Liver tissues samples in NADH/NAD Extraction Buffer were filtered through a 10kD Spin Column (Abcam) to remove enzymes that consume NADH before performing the assay. Total NAD+ and NADH were measured following the manufacturer's instructions. Liver superoxide dismutase (SOD) activity was measured by SOD assay kit (Jiancheng Bioengineering, Nanjing, China). Hepatic hydroxyproline was assayed by hydroxyproline assay kit (Jiancheng Bioengineering) to quantify liver collagen content. Triglyceride was detected using Wako E-test triglyceride Kit (Wako Pure Chemical Industries, Osaka, Japan). Lipid peroxidation was quantified by measuring malondialdehyde using thiobarbituric acid reactive substances (TBARs) assay (Sigma-Aldrich, St Louis, MO). Serum lipopolysaccharides (LPS) concentration in portal vein was detected by an antibody-based biosensors method[7] with mouse LPS ELISA Kit (Cusabio).

Zhang X, *et al*. *Gut* 2021; 70:761–774. doi: 10.1136/gutjnl-2019-319664

Fluorescence activated cell sorting analysis

Flow cytometry cell sorting was performed to study the immune cell type in the liver tissues of germ-free mice. Cell suspension was filtered with a 70-mm cell strainer, washed with PBS, and resuspended in staining solution for flow cytometry. Cells were then stained with fluorochrome-conjugated monoclonal antibodies: fluorescein isothiocyanate (FITC) anti-mouse CD45 (BioLegend, San Diego, CA).

Immunohistochemistry

Paraffin-embedded liver tissues were used for analyzing Ki-67, AFP, GP73 and α-SMA expression and colon tissues for evaluating E-cadherin expression. After deparaffinization, the slides were heated in an autoclave with sodium citrate for antigen repairing, followed by 1% hydrogen peroxide to abolish endogenous peroxidase activity, and blocked with 2% goat serum. Slides were then incubated with primary antibody including Ki67 (Cell signalling technology, Danvers, MA, #9449,1:400), AFP (Abcam, ab46799, 1:200), GP73 (Santa Cruz Biotechnology, Santa Cruz, CA, sc-365817, 1:100), α-SMA (Santa Cruz Biotechnology, sc-32251, 1: 100), E-Cadherin (Cell signalling technology, 14472,1:100) at 4℃ overnight. Signals were developed with DAB (Millipore).

Western blot analysis of hepatic proteins

Thirty micrograms protein were separated by sodium dodecyl sulfate polyacrylamide gel electrophoresis and transferred onto nitrocellulose membranes (GE Healthcare, Piscataway, NJ). Membranes were incubated with

Zhang X, *et al*. *Gut* 2021; 70:761–774. doi: 10.1136/gutjnl-2019-319664

primary antibody CDC20 (Santa Cruz Biotechnology) overnight at 4°C and then with secondary antibody at room temperature for 1 hour. Proteins of interest were visualized using ECL Plus Western Blot Detection Reagents (GE Healthcare). β-Actin was used as total protein loading control.

Quantification RT-PCR and RT2 Profiler PCR Array Gene Expression

Total RNA was extracted using the TRIzol Reagent (Thermo Fisher Scientific). RNA quality was determined using a spectrophotometer and was reverse transcribed using a complementary DNA conversion kit (Thermo Fisher Scientific). qRT-PCR was performed using SYBR Green Master Mix (Roche, Basel, Switzerland) in the Light Cycler 480 Real-Time PCR System (Roche). mRNA expression of mouse genes was analysed with specific primers listed in online supplementary table S9.

PCR array was performed using the real-time RT2 Profiler PCR Array Mouse Inflammatory Response and Autoimmunity (catalog no. PAMM-077Z; QIAGEN, Hilden, Germany) or RT2 Profiler PCR Array Mouse Cancer Pathway Finder (Catalog no. PAMM-033Z; QIAGEN) in combination with SYBR Green qPCR Mastermix. Cycle threshold values were exported to a table for analysis, which were normalized based on a full panel of reference genes. The fold changes were calculated using the delta-delta cycle threshold method. Genes with fold changes more than 2 were considered to be of biological significance.

RNA Sequencing

Total RNA isolated from 5 HFHC- and 10 HFLC-fed mice were subjected to transcriptome sequencing. Poly-A containing messenger RNA purification, double-stranded complementary DNA synthesis, end repair, 3' end adenylation, adapter ligation, and enrichment of DNA fragments for RNA-seq library construction were performed using the reagents provided in the Illumina TruSeq RNA Sample Preparation Kit. RNA-seq library sequencing was constructed on V4 region and performed on an Illumina MiSeq PE250 as per the manufacturer's instructions. Data was presented as Reads Per Kilobase Million (RPKM).

Cell treatment

The Immortalized human liver epithelial cell line LO2 was purchased from the American Type Culture Collection (ATCC, Manassas, VA) and was treated with cholesterol (Sigma-Aldrich) and Taurocholic acid (TCA, 0.1 μM, 1 μM, Sigma-Aldrich) for 24 h. Two NASH-HCC cell lines HKCI-2 and HKCI-10 were established previously from NASH-HCC patients[8] and were treated with 3-Indolepropionic acid (IPA, 10 μM, 100 μM, Sigma-Aldrich) with or without cholesterol (200 μg/mL). Oil Red O staining was then performed for lipid accumulation assessment. (3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay was performed using Vybrant MTT Cell Proliferation Assay Kit (Thermo Fisher Scientific).

Human subjects

We analyzed fecal metagenomics data from 98 subjects including 59 hypercholesterolemia patients and 39 healthy controls recruited from Prince of

Zhang X, *et al*. *Gut* 2021; 70:761–774. doi: 10.1136/gutjnl-2019-319664

Wales Hospital, Hong Kong. The exclusion criteria are hypertension; impaired glucose regulation; diabetes mellitus; history of gastrointestinal disease and gastrointestinal surgery; abnormal liver and kidney function. Each individual provided written informed consent. The clinical study protocol was approved by Joint Chinese University of Hong Kong – New Territories East Cluster Clinical Research Ethics Committee (CUHK-NTEC CREC).

Statistical analyses

Differences between two groups were compared by Student's t test or Mann-Whitney U test. Multiple group comparisons were made by Kruskal-Wallis test or one-way ANOVA. All statistical tests were performed using GraphPad Prism Software. Data were considered significant at $p < 0.05$.

References:

1. Schloss PD, Westcott SL, Ryabin T, et al. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol 2009;75:7537-41.
2. Pruesse E, Peplies J, Glockner FO. SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. Bioinformatics 2012;28:1823-9.
3. Edgar RC, Haas BJ, Clemente JC, et al. UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 2011;27:2194-200.
4. Langille MG, Zaneveld J, Caporaso JG, et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nat Biotechnol 2013;31:814-21.

5. Yu J, Feng Q, Wong SH, et al. Metagenomic analysis of faecal microbiome as a tool towards targeted non-invasive biomarkers for colorectal cancer. Gut 2017;66:70-8.
6. Zhang X, Shen J, Man K, et al. CXCL10 plays a key role as an inflammatory mediator and a non-invasive biomarker of non-alcoholic steatohepatitis. *J Hepatol* 2014;61:1365-75.
7. Su W, Ding X, Methods of Endotoxin Detection. J Lab Autom. 2015;20:354-64.
8. Xu W, Zhang X, Wu JL, et al. O-GlcNAc transferase promotes fatty liver-associated liver cancer through inducing palmitic acid and activating endoplasmic reticulum stress. *J Hepatol* 2017;67:310-20.

Zhang X, *et al*. *Gut* 2021; 70:761–774. doi: 10.1136/gutjnl-2019-319664

Supplementary Figures

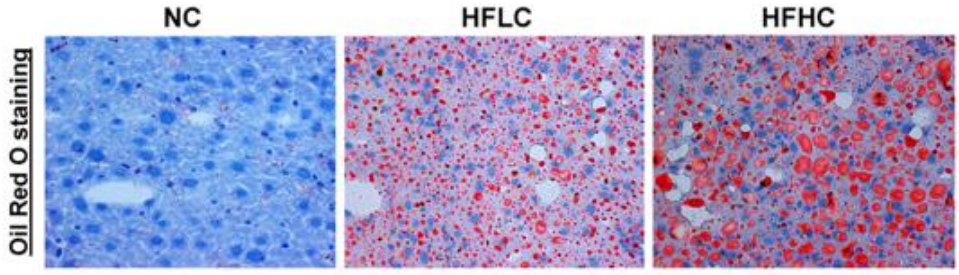

Figure S1 Oil Red O staining of liver sections in mice fed with NC, HFLC or HFHC diet for 14 months. HFHC, high-fat/high-cholesterol diet; HFLC, high-fat/low-cholesterol diet; NC, normal chow.

Zhang X, *et al*. *Gut* 2021; 70:761–774. doi: 10.1136/gutjnl-2019-319664

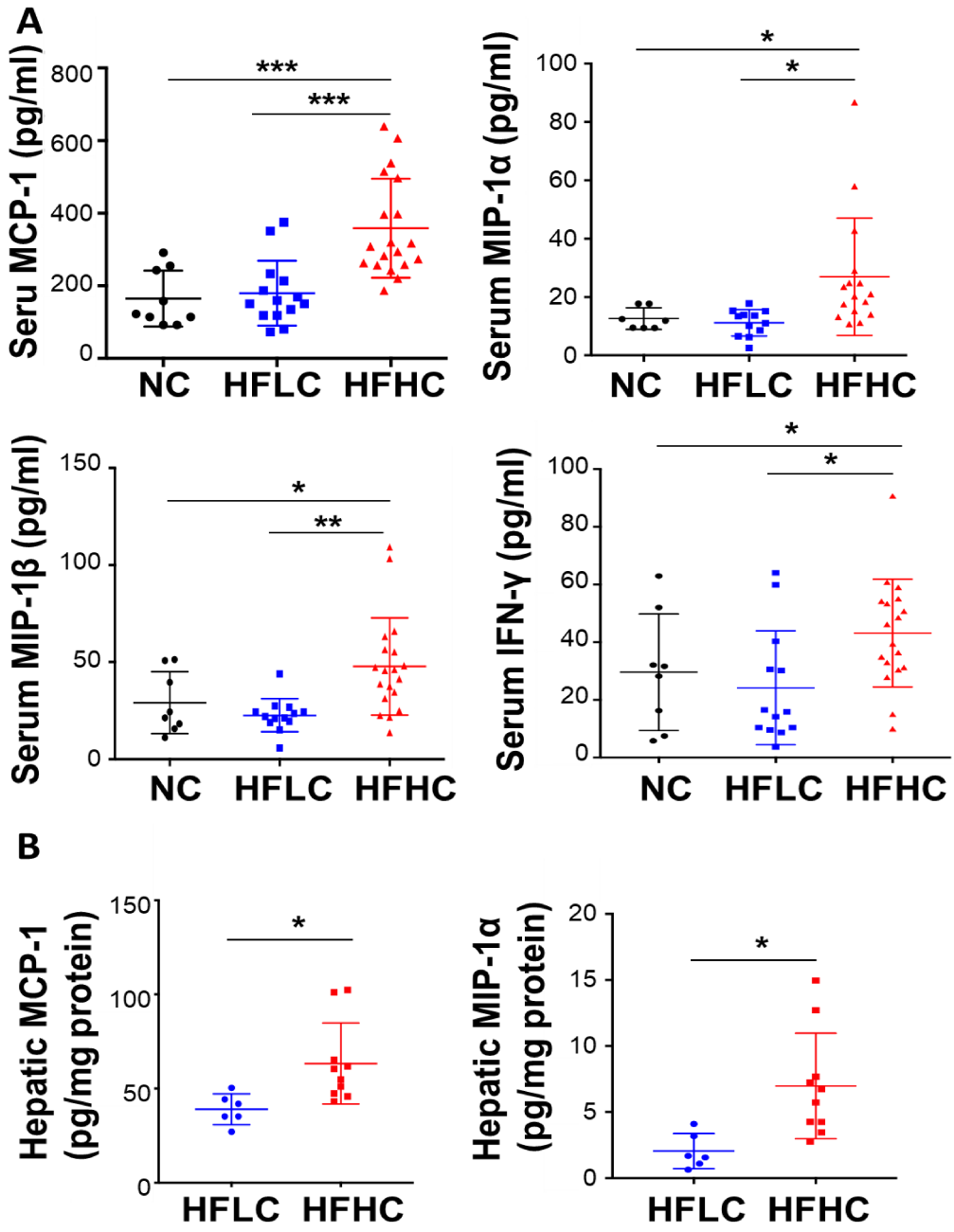

Figure S2 (A) Serum MCP-1, MIP-1α, MIP-1β and IFN-γ protein levels determined by cytokine profiling assay in mice fed NC, HFLC and HFHC diet for 14 months; **(B)** Hepatic MCP-1 and MIP-1α levels by cytokine profiling assay in mice fed HFLC and HFHC diet for 14 months. *p < 0.05, **p < 0.01, ***p < 0.001. HFHC, high-fat/high-cholesterol diet; HFLC, high-fat/low-cholesterol diet; NC, normal chow.

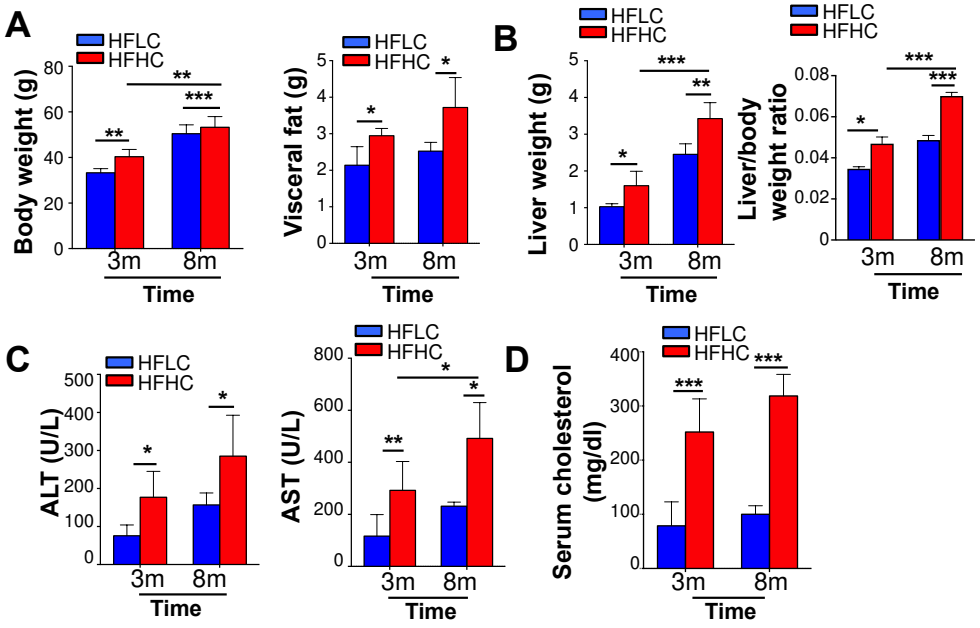

Figure S3 (A) Body weight, visceral fat weight, (**B**) liver weight, liver-to-body weight, (**C**) serum levels of ALT, AST and (**D**) cholesterol in mice fed with HFLC and HFHC for 3 and 8 months *p < 0.05, **p < 0.01, ***p < 0.001. ALT, alanine aminotransferase; AST, aspartate aminotransferase; HFHC, high-fat/high-cholesterol diet; HFLC, high-fat/low-cholesterol diet.

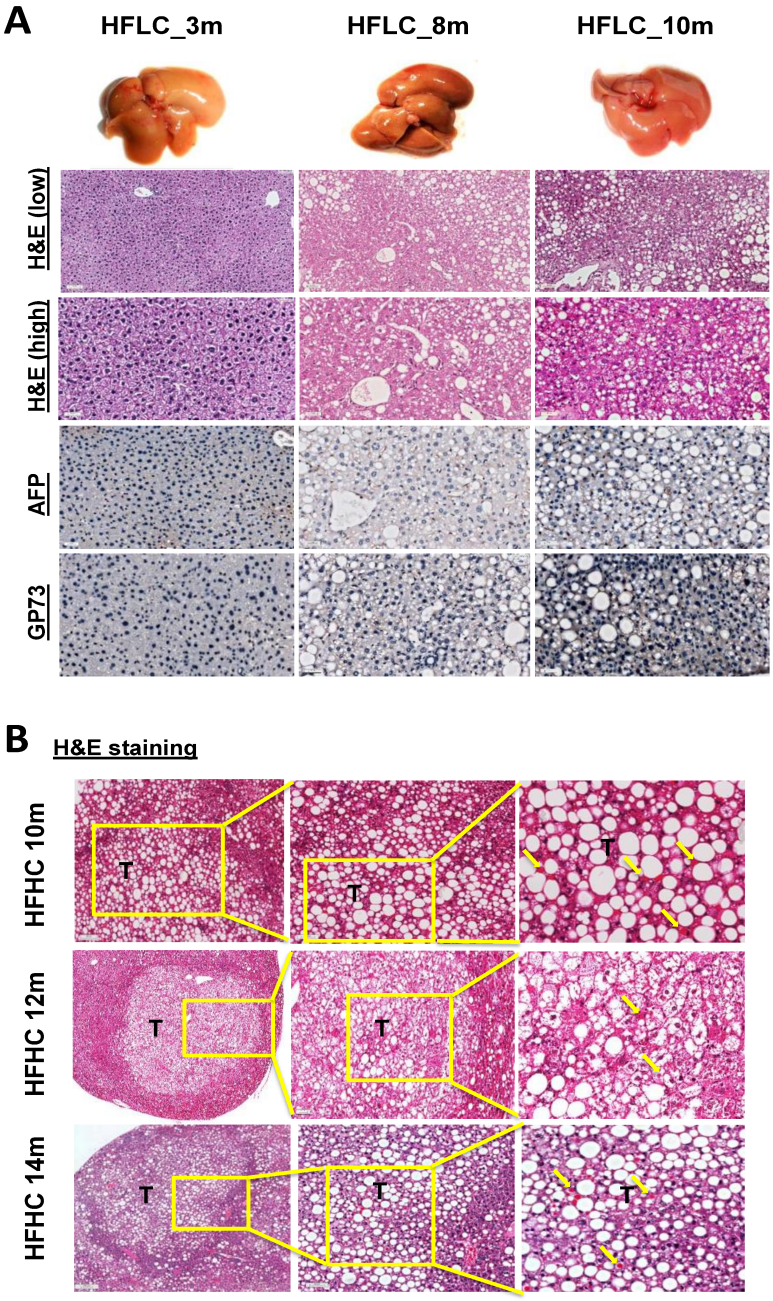

Figure S4 (A) Representative gross morphology, H&E staining, and IHC staining of tumour markers AFP and GP73 in the liver of mice fed with HFLC for 3,8 and 10 months**; (B)** H&E staining of well-demarcated tumours from 10, 12 and 14 months HFHC-fed mice. Yellow arrows indicate Mallory hyaline. AFP, alpha-fetoprotein; GP73, Golgi protein 73; HFHC, high-fat/high-cholesterol diet; HFLC, high-fat/low-cholesterol diet; IHC, immunohistochemistry.

.

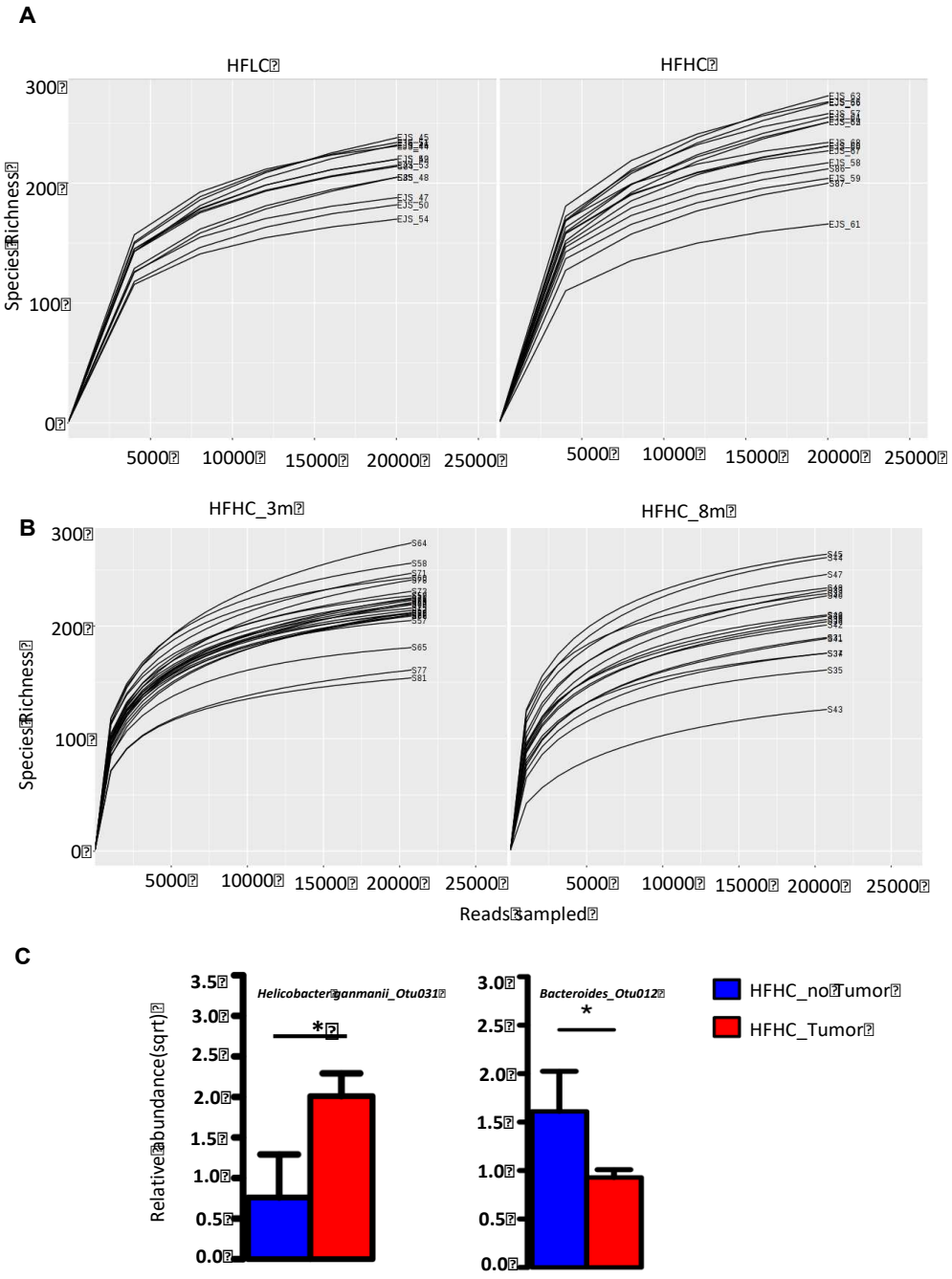

Figure S5 Rarefaction analysis shows that observed number of OTUs reached saturation in **(A)** 14 months HFLC and HFHC-fed mice and **(B)** 3 months and 8 months HFHC-fed mice**; (C)** *Helicobacter ganmanii_Otu031* was observed to be enriched in HFHC-fed mice with tumour compared with HFHC-fed mice without tumour while *Bacteroides_Otu012* was observed to be reduced in HFHC-fed mice with tumour compared with HFHC-fed mice without tumour. *p < 0.05. HFHC, high-fat/high-cholesterol diet; HFLC, high-fat/low-cholesterol diet; OTUs, operational taxonomic units.

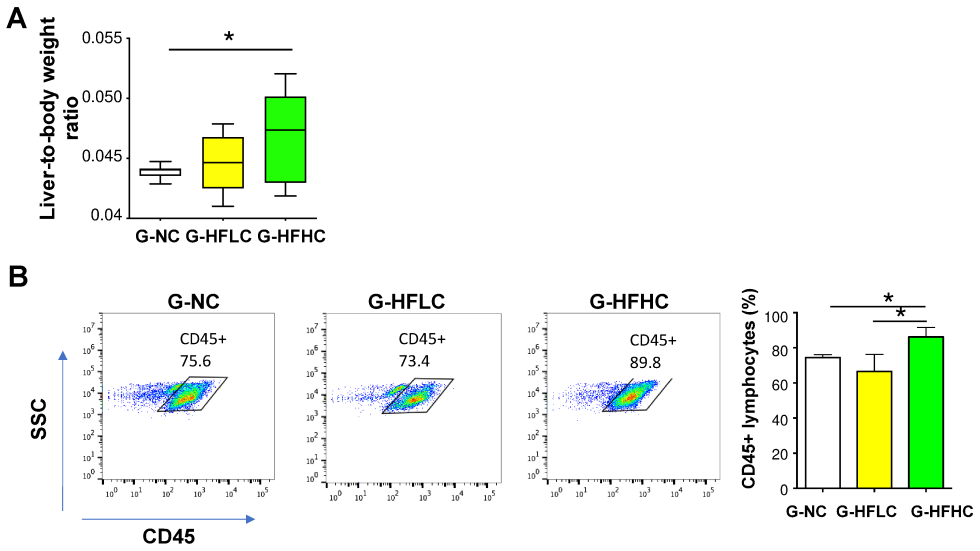

Figure S6 (A) Liver-to-body weight ratio in G-NC, G-HFLC and G-HFHC mice; **(B)** Flow cytometry analysis of $CD45^+$ lymphocytes in G-NC, G-HFLC and G-HFHC mice at 10 months after fecal microbiota transplantation. *p < 0.05. HFHC, high-fat/high-cholesterol diet; HFLC, high-fat/low-cholesterol diet; NC, normal chow.

Zhang X, *et al*. *Gut* 2021; 70:761–774. doi: 10.1136/gutjnl-2019-319664

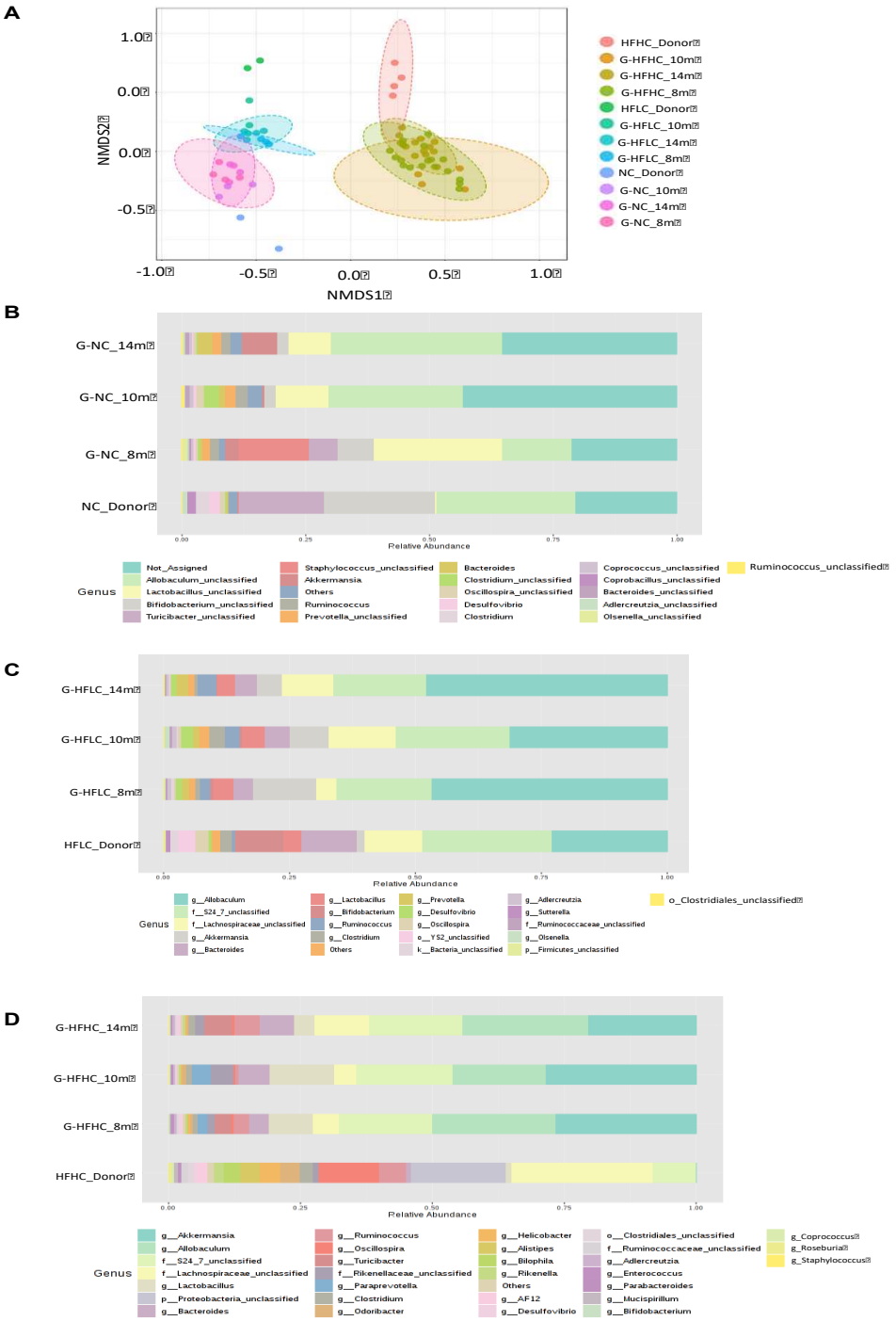

Figure S7 (**A**) β-diversity blot and (**B-D)** the composition of transferred ecosystem in conventional donor mice and germ-free mice transplanted with stool from NC, HFLC and HFHC-fed mice at 8, 10 and 14 months. HFHC, high-fat/high-cholesterol diet; HFLC, high-fat/low-cholesterol diet; NC, normal chow.

Zhang X, *et al*. *Gut* 2021; 70:761–774. doi: 10.1136/gutjnl-2019-319664

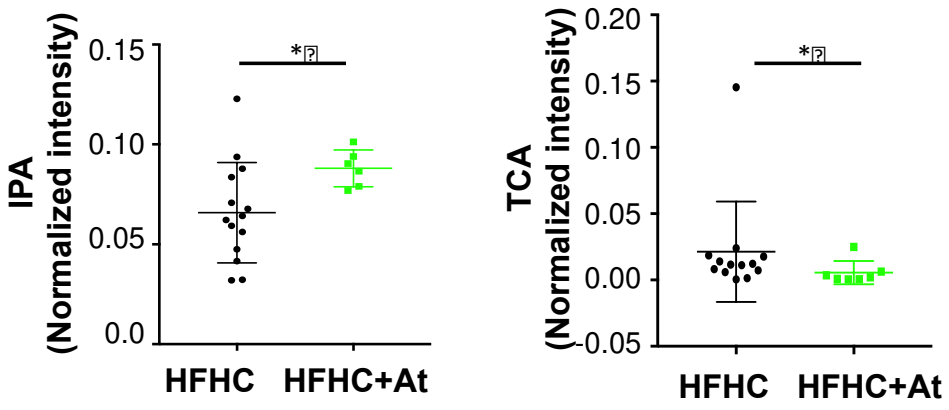

Figure S8 Serum IPA is up-regulated while TCA is down-regulated after Atorvastatin treatment in HFHC-fed mice. *p < 0.05. At, Atorvastatin; HFHC, high-fat/high-cholesterol diet; IPA, 3-indolepropionic acid; TCA, taurocholic acid.

Zhang X, *et al*. *Gut* 2021; 70:761–774. doi: 10.1136/gutjnl-2019-319664

Supplementary Tables

Table S1 Differentially abundant bacteria at phylum level

Taxa.Phylum	FDR	HFLC_mean	HFHC_mean	HFLC_median	HFHC_median
Proteobacteria	0.0009	1.86	7.73	1.89	6.09
Actinobacteria	0.0061	4.66	0.6	3.64	0.51
Deferribacteres	0.0092	0.0046	1.36	0	0.75
Tenericutes	0.0093	0.14	0.012	0.09	0
Verrucomicrobia	0.0099	1.34	0.019	1.21	0.02
Cyanobacteria	0.011	1.41	0.005	0.98	0
Firmicutes	0.011	48.78	61.22	47.16	61.31
Bacteroidetes	0.015	41.26	28.02	43.8	31.52

Zhang X, *et al. Gut* 2021; 70:761–774. doi: 10.1136/gutjnl-2019-319664

Table S2 Differentially abundant bacterial classes

Taxa.class	FDR	HFLC_mean	HFHC_mean	HFLC_median	HFHC_median
Clostridia	3.00E-06	15.68	46.48	15.41	47.03
Deltaproteobacteria	0.0004	0.25	2.84	0.16	2.58
Betaproteobacteria	0.0029	0.86	0.13	1.1	0.03
Actinobacteria	0.0029	4.43	0.046	3.52	0.03
Alphaproteobacteria	0.0066	0.38	0.047	0.37	0.03
Coriobacteriia	0.0066	0.23	0.55	0.12	0.5
Epsilonproteobacteria	0.0066	0.32	4.44	0.05	2.81
Deferribacteres	0.0066	0.0046	1.36	0	0.75
Mollicutes	0.0078	0.14	0.012	0.09	0
Erysipelotrichi	0.0078	23.23	7.93	22.78	0.6
Bacteroidia	0.0078	41.14	26.57	43.66	29.72
Verrucomicrobiae	0.0079	1.34	0.019	1.21	0.02
X4C0d2	0.011	1.41	0.005	0.98	0

Zhang X, *et al. Gut* 2021; 70:761–774. doi: 10.1136/gutjnl-2019-319664

Table S3 Differentially abundant bacterial orders

Taxa.order	FDR	HFLC_mean	HFHC_mean	HFLC_median	HFHC_median
Clostridiales	5.00E-06	15.68	46.47	15.41	47.03
Gemellales	0.0002	0.04	0.52	0.01	0.43
Desulfovibrionales	0.0004	0.25	2.84	0.16	2.58
Burkholderiales	0.0036	0.85	0.12	1.09	0.03
Bifidobacteriales	0.0037	4.41	0.0062	3.52	0
RF32	0.009	0.38	0.042	0.37	0.02
RF39	0.009	0.13	0.00062	0.09	0
Coriobacteriales	0.009	0.23	0.55	0.12	0.5
Campylobacterales	0.009	0.32	4.44	0.05	2.81
Deferribacterales	0.009	0.0046	1.36	0	0.75
Erysipelotrichales	0.012	23.23	7.93	22.78	0.6
Bacteroidales	0.012	41.14	26.57	43.66	29.72
Verrucomicrobiales	0.012	1.34	0.019	1.21	0.02
YS2	0.017	1.41	0.005	0.98	0
Turicibacterales	0.065	0.41	0.054	0.24	0.05

Zhang X, *et al. Gut* 2021; 70:761–774. doi: 10.1136/gutjnl-2019-319664

Table S4 Differentially abundant bacterial families

Taxa.family	FDR	HFLC_mean	HFHC_mean	HFLC_median	HFHC_median
Ruminococcaceae	6.00E-05	2.05	10.14	1.76	9.5
Streptococcaceae	6.00E-05	0.34	2.5	0.16	2.46
Gemellaceae	0.0003	0.04	0.52	0.01	0.43
Desulfovibrionaceae	0.0005	0.22	2.7	0.13	2.43
Odoribacteraceae	0.0006	0.053	1.35	0.03	1.42
Lachnospiraceae	0.0007	12.95	28.42	13.23	27.9
Mogibacteriaceae	0.0007	0.02	0.15	0.02	0.12
Rikenellaceae	0.0017	0.82	6.06	0.48	5.19
Paraprevotellaceae	0.0018	1.14	0.088	1.2	0.06
Alcaligenaceae	0.0022	0.84	0.11	1.06	0.01
Peptostreptococcaceae	0.0022	0.00077	0.1	0	0.11
Christensenellaceae	0.0023	0.0031	0.046	0	0.035
Bifidobacteriaceae	0.0026	4.41	0.0062	3.52	0
Bacteroidaceae	0.0059	10.3	2.1	11.07	1.25
Lactobacillaceae	0.0064	5.1	1.05	3	1
Coriobacteriaceae	0.0089	0.23	0.55	0.12	0.5
Helicobacteraceae	0.0089	0.32	4.44	0.05	2.81
Deferribacteraceae	0.0089	0.0046	1.36	0	0.75
Erysipelotrichaceae	0.013	23.23	7.93	22.78	0.6
Verrucomicrobiaceae	0.014	1.34	0.019	1.21	0.02
Dehalobacteriaceae	0.017	0.046	0.1	0.03	0.08
S247	0.02	28.13	16.23	29.26	14.29
Leuconostocaceae	0.02	0	0.0062	0	0

Zhang X, *et al. Gut* 2021; 70:761–774. doi: 10.1136/gutjnl-2019-319664

Table S5 Differentially abundant bacterial genera

Taxa.genus	FDR	HFLC_mean	HFHC_mean	HFLC_median	HFHC_median
Ruminococcus	3.00E-05	0.67	6.94	0.61	7.04
Oscillospira	5.00E-05	1.35	8.66	1.14	7.7
Lactococcus	0.0002	0.28	2.21	0.14	2.2
Gemella	0.0003	0.04	0.52	0.01	0.43
Anaerovorax	0.0007	0.00077	0.023	0	0.025
Bilophila	0.0007	0.038	1.85	0.04	1.5
Odoribacter	0.0008	0.052	1.34	0.03	1.42
Prevotella	0.0012	1.15	0.005	1.21	0
Adlercreutzia	0.003	0.12	0.5	0.04	0.49
Sutterella	0.0032	0.84	0.11	1.06	0
AF12	0.0032	0.00077	0.86	0	0.78
Bifidobacterium	0.0041	4.41	0.0062	3.52	0
Roseburia	0.0041	0.00077	0.44	0	0.39
Alistipes	0.0041	0.29	2.82	0.05	2.4
Desulfovibrio	0.0067	0.18	0.84	0.1	0.61
Bacteroides	0.0081	10.3	2.1	11.07	1.25
Lactobacillus	0.0082	5.05	1.04	2.99	0.99
Defluviitalea	0.012	0.012	0.00062	0.01	0
Mucispirillum	0.012	0.0046	1.36	0	0.75
Dorea	0.012	0.16	0.56	0.02	0.47
Helicobacter	0.012	0.32	4.37	0.05	2.69
Paraprevotella	0.014	0.0054	0.085	0	0.055
Rikenella	0.018	0.012	0.86	0.01	0.42

Zhang X, *et al. Gut* 2021; 70:761–774. doi: 10.1136/gutjnl-2019-319664

Akkermansia	0.019	1.34	0.019	1.21	0.02
Anaerotruncus	0.022	0	0.035	0	0.015
Dehalobacterium	0.022	0.046	0.1	0.03	0.08

Zhang X, *et al. Gut* 2021; 70:761–774. doi: 10.1136/gutjnl-2019-319664

Table S6 Clinical features for the recruited subjects

Variable	Healthy control (n = 39)	Hypercholesterolemia patients (n = 59)	*P* value
Mean age, years ± SD	66.74 ± 7.32	64.59 ± 6.88	0.1492
Gender			
Male	19	22	
Female	20	37	
Total Cholesterol (mmol/L)	4.42 ± 0.53	6.0 ± 0.56	<0.0001
LDL-Cholesterol (mmol/L)	2.37 ± 0.50	3.83 ± 0.61	<0.0001
HDL-Cholesterol (mmol/L)	1.59 ± 0.47	1.65 ± 0.51	0.58
Triglyceride (mmol/L)	1.08 ± 0.70	1.24 ± 0.64	0.2427

Zhang X, *et al. Gut* 2021; 70:761–774. doi: 10.1136/gutjnl-2019-319664

Table S7 Metabolite profiling of the serum from HFLC and HFHC fed mice by UHPLC-QTOF-MS (HFHC vs HFLC)

Metabolites	FC	log2(FC)	p.ajusted	-log(p)
Ergothioneine	0.19227	-2.3788	6.66E-06	5.1768
4-Pyridoxic acid.1	3.5102	1.8116	1.47E-05	4.8317
Monoethylglycylxylidide MEGX	0.31307	-1.6754	1.48E-05	4.83
Primaquine	1.6572	0.72876	1.48E-05	4.83
4-Pyridoxic acid	3.9046	1.9652	2.02E-05	4.695
3-Ureidopropionate.1	1.7887	0.83887	2.02E-05	4.695
Dapsone	3.6675	1.8748	3.00E-05	4.5225
N-Acetyl-L-Histidine.1	1.6355	0.70977	8.82E-05	4.0546
3-Indolepropionic acid	0.11671	-3.0989	0.00012693	3.8964
2E-Eicosenoic acid	1.6225	0.69818	0.00012693	3.8964
Flumequine	0.45436	-1.1381	0.0001475	3.8312
N-Acetyl-L-Histidine	1.5273	0.61096	0.0001475	3.8312
Cyclohexylamine	0.53851	-0.89296	0.00028511	3.545
Maltitol	2.2687	1.1819	0.0004191	3.3777
3-Aminopropanesulphonic Acid	2.0546	1.0389	0.0004191	3.3777
Cholesterol	1.6826	0.75065	0.00056583	3.2473
3-Carboxypropyltrimethylammonium cation	0.27648	-1.8548	0.00073858	3.1316
13S-HODE	0.4203	-1.2505	0.00094262	3.0257
Saccharin	2.8602	1.5161	0.00095244	3.0212
3-Ureidopropionate	1.974	0.98114	0.00095244	3.0212
D-glucosamine 6-phosphate	0.39333	-1.3462	0.00095354	3.0207
Galactonic acid	1.7022	0.76739	0.0013587	2.8669

Zhang X, *et al. Gut* 2021; 70:761–774. doi: 10.1136/gutjnl-2019-319664

Nervonic acid	2.1918	1.1321	0.0017002	2.7695
1-Methyladenosine	1.9871	0.99069	0.0018042	2.7437
7-Oxocholesterol	2.6777	1.421	0.0018581	2.7309
1-Oleoyl-sn-glycero-3-phosphocholine	2.0085	1.0061	0.0019883	2.7015
D-Mannitol	1.6381	0.71201	0.0022459	2.6486
11Z, 14Z-Eicosadienoic Acid	1.5433	0.62606	0.0031841	2.497
D-Quinovose	0.76915	-0.37866	0.0032892	2.4829
Pentadecanoic Acid.1	0.73664	-0.44097	0.0036911	2.4328
L-Cysteine	0.51998	-0.94346	0.004116	2.3855
L-Phenylalanine.1	1.3473	0.43012	0.004116	2.3855
Deoxythymidine 5'-phosphate dTMP	0.48225	-1.0522	0.0045282	2.3441
N-Acetylmannosamine	2.1466	1.1021	0.0046692	2.3308
Phthalic acid Mono-2-ethylhexyl Ester	1.6153	0.69177	0.0060631	2.2173
Indolelactic acid	1.7874	0.83789	0.0062456	2.2044
Taurochenodeoxycholate	4.3531	2.122	0.0068544	2.164
Stearidonic Acid.1	0.67248	-0.57243	0.0075188	2.1239
L-Proline.1	2.0247	1.0177	0.0088351	2.0538
Urocanic acid	0.63213	-0.66172	0.0088351	2.0538
5-Fluoro-5'-Deoxyuridine	1.9255	0.94521	0.0089537	2.048
N-Acetylglutamine.1	1.6306	0.70537	0.011058	1.9563
Met-Tyr	2.273	1.1846	0.011078	1.9555
gamma-L-Glutamyl-L-phenylalanine	1.6335	0.70797	0.011078	1.9555
O-Acetyl-L-serine	1.5688	0.64963	0.011078	1.9555
2-Hydroxyadenine	1.4254	0.51139	0.011739	1.9304
Hydroxyphenyllactic acid	1.6956	0.76181	0.012567	1.9008

Zhang X, *et al. Gut* 2021; 70:761–774. doi: 10.1136/gutjnl-2019-319664

Salicylic acid	1.5987	0.6769	0.012567	1.9008
N-Acetyl-L-alanine	1.3748	0.45925	0.012893	1.8897
N-Acetyl-L-phenylalanine	1.3133	0.39314	0.013885	1.8575
L-Methionine.1	2.3358	1.2239	0.014152	1.8492
Psychosine	3.0068	1.5882	0.015578	1.8075
trans-cinnamate	1.2696	0.34434	0.015578	1.8075
Creatine	1.3636	0.44746	0.01609	1.7934
N-Acetyl-DL-methionine	2.253	1.1719	0.01974	1.7047
L-Valine.1	1.4544	0.54041	0.01974	1.7047
L-Methionine	2.0648	1.046	0.021164	1.6744
Ile-Met	1.8052	0.85215	0.021164	1.6744
Dihomo-gamma-Linolenic Acid	1.5527	0.63482	0.021164	1.6744
Arg-Cys	1.4483	0.53441	0.021164	1.6744
Taurolithocholic acid	1.4875	0.57288	0.02137	1.6702
N6-Acetyl-L-lysine	2.7307	1.4493	0.023293	1.6328
Hippuric acid	0.47554	-1.0724	0.024488	1.611
gamma-L-Glutamyl-L-valine	1.5589	0.64056	0.025501	1.5934
1-Stearoyl-2-hydroxy-sn-glycero-3-phosphocholine	1.6247	0.70018	0.026053	1.5841
Pantothenate.1	1.4917	0.57699	0.026053	1.5841
Glycerol 1-myristate	0.70413	-0.50609	0.026053	1.5841
AFMK	1.3016	0.38032	0.026053	1.5841
Erucic acid.1	1.7374	0.79689	0.02718	1.5658
L-Leucine.1	1.3412	0.42357	0.031176	1.5062
Sepiapterin	1.2967	0.37486	0.031532	1.5012
Nname, cis-9, 10-Epoxystearic acid	0.7355	-0.4432	0.031813	1.4974

Famciclovir	1.7937	0.84296	0.0328	1.4841
Serotonin	0.59612	-0.74633	0.033096	1.4802
Biliverdin	0.53985	-0.88936	0.034116	1.467
Tyr-Thr	0.57619	-0.79537	0.03478	1.4587
1-methylguanosine	1.252	0.32422	0.035218	1.4532
Glycocholic acid	3.6327	1.8611	0.036445	1.4384
Alpha-N-Phenylacetyl-L-glutamine	1.6041	0.68176	0.036445	1.4384
Glycerol 3-phosphate	1.5	0.58494	0.037004	1.4317
Taurodeoxycholic acid	2.427	1.2792	0.038368	1.416
2-Hydroxy-3-methylbutyric acid	1.535	0.61825	0.039199	1.4067
6-Benzylaminopurine	1.3866	0.47155	0.039199	1.4067
Acetylcarnitine.1	2.0095	1.0068	0.041887	1.3779
1-Aminocyclopropanecarboxylic acid	1.7788	0.83093	0.042159	1.3751
5-Hydroxytryptophol 5HTOL	0.597	-0.74419	0.044459	1.352
5-Methylcytidine	1.4986	0.58358	0.044459	1.352
Val-Tyr	1.7379	0.79736	0.044485	1.3518
Glycerophosphocholine	1.3064	0.3856	0.045849	1.3387
Succinate	1.483	0.56854	0.049279	1.3073
3-Hydroxydodecanoic acid	0.73764	-0.439	0.049279	1.3073
L-Norleucine.1	1.3349	0.41675	0.049279	1.3073
Isradipine	1.6995	0.76511	0.049775	1.303
Cytidine.1	0.76429	-0.38782	0.049775	1.303
6-Hydroxydopamine	0.56261	-0.82979	0.05059	1.2959
His-Tyr	1.3123	0.39215	0.050827	1.2939
Glycyl-L-leucine	1.6199	0.69592	0.056172	1.2505

Zhang X, *et al. Gut* 2021; 70:761–774. doi: 10.1136/gutjnl-2019-319664

n-Propyl cinnamate	0.62714	-0.67315	0.056172	1.2505
Glu-Ser	1.5025	0.58735	0.056172	1.2505
Harmane	1.4804	0.56603	0.056172	1.2505
Theaflavin	1.4249	0.51091	0.056172	1.2505
Ile-Pro	0.70857	-0.49701	0.056172	1.2505
1-O-cis-9-Octadecenyl-2-O-acetyl-sn-glycero-3-phosphocholine	1.3356	0.41753	0.056172	1.2505
Lys-Cys	1.5027	0.58754	0.062556	1.2037
L-Proline	1.4393	0.52534	0.062582	1.2036
.gamma.-L-Glu-.epsilon.-L-Lys	1.9648	0.97436	0.063111	1.1999
5'-Deoxyadenosine.1	1.5635	0.64479	0.064258	1.1921
Citramalic acid	1.5074	0.59207	0.064258	1.1921
N-Trishydroxymethylmethyl-2-aminoethanesulfonic acid TES	1.4196	0.50547	0.06498	1.1872
Ile-Ala	1.3013	0.38	0.064995	1.1871
PC160/160	1.3354	0.41728	0.066042	1.1802
Dulcitol	1.4063	0.4919	0.066783	1.1753
Lomefloxacin	0.54614	-0.87266	0.06833	1.1654
Altretamine	0.38781	-1.3666	0.071927	1.1431
Undecanedioic acid	0.64167	-0.64009	0.075071	1.1245
4-Aminophenol	0.6663	-0.58575	0.080694	1.0932
trans-2-Octenoic acid, ethyl ester	1.2884	0.36559	0.080694	1.0932
L-Aspartate	1.5139	0.59828	0.091297	1.0395
9S-HODE	0.54005	-0.88884	0.091911	1.0366
Sphingosine	1.6775	0.74628	0.097342	1.0117
L-Glutamate.1	1.3357	0.41757	0.10729	0.96946
N-Acetyl-L-glutamate	1.3605	0.44411	0.11125	0.95369

Taurocholic acid	6.492	2.6987	0.11304	0.94679
DL-Methionine sulfoxide	1.7107	0.77461	0.11979	0.92157
1-Palmitoyl-2-hydroxy-sn-glycero-3-phosphoethanolamine	1.3724	0.45666	0.11979	0.92157
DL-2-Aminoadipic acid.1	1.3193	0.39979	0.12468	0.9042
Cytidine	0.78687	-0.3458	0.12665	0.89741
Sucrose	3.4095	1.7696	0.12821	0.89209
D-Lyxose	1.3148	0.39487	0.1315	0.88107
Glyceric acid	1.4071	0.49268	0.13202	0.87936
Galactinol	3.2428	1.6972	0.13323	0.8754
Phe-Tyr	1.4099	0.49554	0.135	0.86967
N-Acetyl-L-glutamate.1	1.3813	0.46603	0.13526	0.86882
5-L-Glutamyl-L-alanine	1.742	0.80076	0.1405	0.85234
Thymidine	1.7058	0.77043	0.14357	0.84293
Glutathione disulfide	1.3612	0.44493	0.15075	0.82173
Nicotinate D-ribonucleotide	1.3341	0.41586	0.15075	0.82173
2'-Deoxy-D-ribose	0.49436	-1.0164	0.16176	0.79114
2-Methylbutyroylcarnitine	1.6145	0.69112	0.16595	0.78002
N-Acetyl-L-tyrosine	1.4054	0.49096	0.16595	0.78002
Deoxycytidine	0.76141	-0.39325	0.16595	0.78002
DL-Methionine sulfoxide.1	1.5768	0.65696	0.16667	0.77814
Docosatrienoic Acid	1.4999	0.58488	0.16786	0.77505
Acetylcarnitine	1.2715	0.3465	0.17359	0.76047
Pyrrolidine	1.2566	0.32948	0.17359	0.76047
Sotalol	1.2871	0.36416	0.17602	0.75443
EDTA	1.3469	0.42967	0.17613	0.75416

Acetylglycine	0.77854	-0.36115	0.17802	0.74953
Phenyllactic acid	1.487	0.57244	0.18275	0.73815
R-2-Hydroxycaprylic acid	0.70155	-0.51138	0.18503	0.73275
Phosphorylcholine.1	1.3039	0.38282	0.18503	0.73275
L-Tyrosine.1	1.3349	0.41671	0.19925	0.70061
Fosfomycin	1.3332	0.41487	0.19925	0.70061
Uric acid	0.75718	-0.4013	0.19925	0.70061
DL-O-tyrosine.1	1.2777	0.3536	0.20584	0.68648
Acetoacetic acid	0.73106	-0.45194	0.20613	0.68585
Cysteine-S-sulfate	0.74281	-0.42893	0.2105	0.67675
Glutathione disulfide.1	1.3342	0.41603	0.2105	0.67675
Nicotinamide N-oxide	0.77466	-0.36836	0.23075	0.63685
3-Methoxy-4-Hydroxyphenylglycol Sulfate.1	0.41773	-1.2593	0.23783	0.62373
Ile-Tyr	1.4623	0.54821	0.23783	0.62373
Pro-Tyr	1.3453	0.42788	0.23783	0.62373
Chenodeoxycholate	1.4636	0.54949	0.23976	0.62022
alpha-ketoglutarate	0.4643	-1.1069	0.24141	0.61724
Tauroursodeoxycholic acid	6.0636	2.6002	0.24299	0.61442
Triethanolamine	0.77677	-0.36444	0.24299	0.61442
5-Hydroxyindoleacetate	1.3176	0.39793	0.24978	0.60245

Zhang X, *et al. Gut* 2021; 70:761–774. doi: 10.1136/gutjnl-2019-319664

Table S8 cDNA expression array analysis of liver tissues from germ-free mice transplanted with stool from HFLC and HFHC-fed mice with tumour

		G-HFHC *vs* G-HFLC	
Gene Symbol	**Fold Regulation**	**Function**	**Expression**
Fos	5.06	Promote HCC through cholesterol accumulation	Macrophages and Neutrophils
Ccl12	2.00	Pro-inflammatory cytokine	Macrophages
Cxcr1	1.96	Promote steatohepatitis	Neutrophils
Ccl1	1.84	Pro-inflammatory cytokine	Macrophages
Myd88	1.80	Cause IL-6 indution	Macrophages
Il1β	1.78	Promote steatohepatitis	Macrophages and Neutrophils
Cxcl10	1.69	A key cytokine in NASH	Macrophages
C3ar1	1.60	G protein-coupled receptor protein involved in the complement system	Macrophages

Zhang X, *et al. Gut* 2021; 70:761–774. doi: 10.1136/gutjnl-2019-319664

Table S9 Mouse primer sequences for real-time qRT-PCR

Gene Symbol	Forward Primer	Reverse Primer
Cyp7a1	CACCATTCCTGCAACCTTCTGG	ATGGCATTCCCTCCAGAGCTGA
Cyp8b1	CATGAAGGCTGTGCGTGAGGAA	CATCACGCTGTCCAACACTGGA
Cyp27a1	TCAGGAGACCATCGGCACCTTT	CCAGTCACTTCCTTGTGCAAGG
Cyp7b1	CGGAAATCTTCGATGCTCCAAAG	GCTTGTTCCGAGTCCAAAAGGC
α-SMA	TGCTGACAGAGGCACCACTGAA	CAGTTGTACGTCCAGAGGCATAG
GAPDH	CATCACTGCCACCCAGAAGACTG	ATGCCAGTGAGCTTCCCGTTCAG

Zhang X, *et al. Gut* 2021; 70:761–774. doi: 10.1136/gutjnl-2019-319664