Table of content

Supplementary Methods

1. General population-based cohort

The CKB recruited 512,726 individuals (210,206 men and 302,520 women) aged 30-79 years living in 10 regions across China (5 urban and 5 rural). Details of the CKB design, survey methods, and population characteristics have been described elsewhere. During the baseline survey in 2004-2008, interviewer-administered laptop-based questionnaires on sociodemographic characteristics, smoking and alcohol consumption, diet, physical activity, and medical history, a range of physical measurements including height, weight, waist circumference and blood pressure, and blood samples were collected by trained technicians with standard protocols. After completion of the baseline survey, this project randomly selected 5 to 6% of the individuals who remained in the cohort from each of the 10 study regions to conduct periodic resurveys every 5-6 years. The first resurvey took place between July and October 2008, with 19,788 participants resurveyed. The third resurvey took place between August 2020 and December 2021, with 25,078 participants resurveyed. In the third resurvey, liver scans, objectively-measured physical activity, and sleep data via a wrist-worn accelerometer, saliva and fecal samples, and eye measurements (including visual acuity, intraocular pressure and retinal images) were collected.

For each participant, a 10-ml non-fasting blood sample (with time of last meal recorded) was collected into one EDTA vacutainer. The samples were then kept in a portable, insulated cool box with ice packs (to maintain their temperature at 0-4°C) for up to a few hours before being taken to the local study laboratory for immediate processing. After centrifuging and aliquoting, the four cryovials from each blood sample were stored in a -40°C freezer for 3-4 months, before being couriered on dry ice to the central blood repository in Beijing for long-term storage at -80°C.

2. Assessment of metabolomics

2.1 Chemicals and reagents

All of the standards of targeted metabolites were obtained from Sigma-Aldrich (St. Louis, MO, USA), Steraloids Inc. (Newport, RI, USA) and TRC Chemicals (Toronto, ON, Canada). All the standards were accurately weighed and prepared in water, methanol, sodium hydroxide solution, or hydrochloric acid solution to obtain individual stock solution at a concentration of 5.0 mg/mL. Appropriate amount of each stock solution was mixed to create stock calibration solutions. Formic acid was of (Optima grade and obtained from Sigma-Aldrich (St. Louis, MO, USA). Methanol (Optima LC-MS), acetonitrile (Optima LC-MS), and isopropanol (Optima LC-MS) were purchased from Thermo-Fisher Scientific (FairLawn, NJ, USA). Ultrapure water was produced by a Mill-Q Reference system equipped with a LC-MS Pak filter (Millipore, Billerica, MA, USA)[1].

2.2 Instrumentation

An ultra-performance liquid chromatography coupled to tandem mass spectrometry (UPLC-MS/MS) system (ACQUITY UPLC-Xevo TQ-S, Waters Corp., Milford, MA, USA) was used to quantitate all targeted metabolites in this project. The instrument performance optimization and routine maintenance were performed every week.

2.3 Data analysis

The raw data files generated by UPLC-MS/MS were processed using the MassLynx software (v4.1, Waters, Milford, MA, USA) to perform peak integration, calibration, and quantitation for each metabolite.

Mass spectrometry-based quantitative metabolomics refers to the determination of the concentration of a substance in an unknown sample by comparing the unknown to a set of standard samples of known concentration (i.e., calibration curve). The calibration curve is a plot of how the analytical signal changes with the concentration of the analyte (the substance to be measured). For most analyses a plot of instrument response vs. concentration will show a linear relationship. This yields a model described by the equation y = ax + b, where y is the instrument response e.g., peak height or area, a represents the slope/sensitivity, and b is a constant that describes the background. The analyte concentration (x) of unknown samples may be calculated from this equation.

3. Cross-lagged analysis

The cross-lagged panel model simultaneously estimated the autoregressive and cross-lagged regressive effects of NAFLD and metabolomics at two time points, including (1) autoregression of NAFLD at follow-up on metabolomics at baseline, (2) autoregression of β-value at follow-up on β-value at baseline, (3) cross-lagged regression of β-value at follow-up on NAFLD at baseline (ρ_1), and (4) cross-lagged regression of NAFLD at follow-up on β-value at baseline (ρ_2). The significance and magnitude of ρ_1 and ρ_2 reflected the temporal associations of the two variables. We fitted a structural equation model to estimate all the parameters and statistics above using the R package *lavaan*.

We performed a mediation analysis to assess whether metabolomics were the mediators of the effect of BMI on NAFLD. Only metabolomics that showed a potential effect on NAFLD (i.e. CDCA and hippuric acid) in the cross-lagged analysis were included. The temporal associations between BMI and NAFLD and between BMI and metabolomics were assessed with the CLPM first to deduce the potential mediator. The relationships among exposure X, mediator M and outcome Y at baseline (b) and follow-up (f) were as follows:

$$X_f = \beta_X X_b + \varepsilon_{Xf} \quad (1)$$

$$M_f = \beta_M M_b + a X_b + \varepsilon_{Mf} \quad (2)$$

$$Y_f = \beta_Y Y_b + b M_b + c' X_b + \varepsilon_{Yf} \quad (3)$$

where β is the autoregressive coefficient, ε is the residual item, ab is the direct effect, and c' is the indirect effect. The model fitting and parameter estimation method was similar to that of the cross-lagged analysis.

4. Genotyping and GWAS

4.1 Genotyping

The CKB has conducted three phases of genotyping[2]. A custom-designed biobank array, to provide optimised genome-wide coverage for the Chinese population, was developed by the University of Oxford's Clinical Trial Service Unit and Epidemiological Studies Unit (Oxford, UK) in collaboration with the Beijing Genomics Institute (Shenzhen, China) and Affymetrix (now Thermo Fisher Scientific, Santa Clara, CA, USA). This 700K single nucleotide polymorphism (SNP) array was used to genotype ~32,000 CKB participants in the first phase. A revised and updated version of the original array which covers ~803K SNPs was used to genotype ~69 000 participants in the second and third phases.

Variants with call rate >0.98, plate effect $p>10^{-6}$, batch effect $p>10^{-6}$, Hardy-Weinberg equilibrium (HWE) deviations $p>10^{-6}$ (combined 10 degrees of freedom Chi-squared test from 10 regions) and minor allele frequency (MAF) difference from 1000 Genomes East Asian frequencies <0.2 were identified, resulting in genotypes for 532 415 biallelic variants present on both array versions. The qualified genotypes for each chromosome were phased with SHAPEIT. Then, imputation was performed for each 5-Mb interval with IMPUTE 4 based on haplotypes derived from the 1000 Genomes phase III.

4.2 GWAS of CDCA

We selected variants that did not deviate from HWE ($p>1\times10^{-6}$), per variant missing rates <10%, per sample missing rate <10%, and according to various threshold based on MAF and imputation quality score (INFO) (i.e. INFO>0.3 for MAF>3%, INFO>0.6 for MAF 1-3%, INFO>0.8 for MAF 0.5-1%, and INFO>0.9 for MAF 0.1-0.5%). The genotype-phenotype association test was carried out in 2115 samples from the CKB with eligible CDCA measurement and non-missing genotyping data. We carried out linear mixed model (LMM) association analyses and adjusted for genotyping array, 11 ancestry principal components in the CKB to assess the association between the z-score of log-transformed CDCA and imputed genotype dosages under an additive genetic model by using BOLT-LMM version 2.3. After association analysis, we applied the PLINK clumping function to determine top loci that were independent to each other. Specifically, variants with $p<1\times10^{-5}$, r^2 >0.2 and <500 kb away from the peak were assigned to that peak' s clump. The genes within each clump were identified by the overlap between gene regions and clump region.

5. Mendelian randomization

5.1 Genetic associations of BMI with metabolomics and NAFLD

The genetic associations of BMI with metabolomics and NAFLD were examined by the 2-stage least squares (2SLS) estimator method using individual participant-level data. In the first stage, the association between BMI-GRS and BMI was examined using linear regression, adjusted for age, sex, region, the first 11 principal components, GWAS array type, smoking history, alcohol consumption, and levels of physical activity. In the second stage, the associations of the resulting estimated values with metabolomics and NAFLD were examined using linear and logistic regression, respectively, adjusted for the same covariates. We calculated the genetic estimates per increase in genetically predicted BMI, equivalent to 1-SD baseline BMI.

5.2 Genetic associations of CDCA with NAFLD and fibrosis

The genetic associations of CDCA with NAFLD and fibrosis were calculated by two-sample MR using summary-level data. Genetic instruments for CDCA were obtained from CDCA GWAS, with the genome-wide significance level at p-value $\leq 1\times10^{-4}$. After quality control, we excluded SNPs in linkage disequilibrium (r^2 ≥ 0.001) based on the 1000 Genomes European reference panel and attained those with the smallest p-values. Finally, there were 85 SNPs associated with CDCA for genetic instruments (**Supplementary Table 5)**. Logistic regressions were further performed to retrieve the SNP-to-outcome estimates, adjusted for the same covariates as in 2SLS and restricted to those whose BMI >28 kg/m^2 without excessive alcohol consumption or positive HBsAg. We used a conventional inverse variance-weighted (IVW) method in which the SNP-to-outcome estimate was regressed on the SNP-to-exposure estimate using logistic regression, with the y-axis intercept forced through the origin. For each outcome, a combined causal estimate was calculated from the causal estimate from each BMI SNP using a random-effects meta-analysis. Odds ratio (OR) and 95% confidence intervals (CIs) per genetically predicted 1-SD increase in CDCA is presented.

References:

1. Xie G, Wang L, Chen T, et al. A metabolite array technology for precision medicine. *Anal Chem* 2021; 93(14): 5709-5717.
2. Walters RG, Millwood IY, Lin K, et al. Genotyping and population characteristics of the China Kadoorie Biobank. *Cell Genom* 2023; 3(8). doi:10.1016/j.xgen.2023.100361

Table S1. List of metabolomic biomarkers quantified by UPLC-MS

Category	Metabolite	Category	Metabolite
Acylcarnitines	Palmitoylcarnitine	Amino acids	Tyrosine
	L-Carnitine		Tryptophan
	Propionylcarnitine		Methionine
	Acetylcarnitine		Lysine
	Tetradecanoylcarnitine		Glycine
	Stearylcarnitine		Arginine
	Decanoylcarnitine		Leucine
	Glutarylcarnitine		Alanine
	Dodecanoylcarnitine		Proline
	Octanoylcarnitine		Histidine
	Valerylcarnitine		Aspartic acid
	Hexanoylcarnitine		Serine
	Malonylcarnitine		Valine
Bile acids	TCA		Isoleucine
	GCDCA		Asparagine
	GDCA		Threonine
	GCA		Phenylalanine
	TCDCA		Glutamine
	CDCA		Glutamic acid
	CA	Amino acids related	Citrulline
	TDCA		Aminoadipic acid
	GLCA-3S		4-Hydroxyproline
	GLCA		1-Methylhistidine
	DCA		Sarcosine
	GUDCA		Ornithine
Biogenic amines	GABA	Carbohydrates and related	Glucose
	Beta-Alanine	Carboxylic acids	Cis-Aconitic acid
Fatty acids-Free/non-covalently bond	Arachidonic acid		Succinic acid
	DHA		Hippuric acid
	EPA	Indoles derivatives	Indole-3-propionic acid
	Myristic acid		Indoleacetic acid
	Oleic acid		
	Dodecanoic acid		

Abbreviations: FDR, false discovery rate; TCA, taurocholic acid; GCDCA, glycochenodeoxycholic acid; GDCA, glycodeoxycholic acid; GCA, glycocholic acid; TCDCA, taurochenodeoxycholic acid; CDCA, chenodeoxycholic acid; CA, cholic acid; TDCA, taurodeoxycholic acid; GLCA, glucuronic acid; DCA, deoxycholic acid; GUDC, glycoursodeoxycholic acid; GABA, gamma-aminobutyric acid; DHA,

docosahexaenoic acid; EPA, eicosapentaenoic acid.

Table S2. Characteristics of study participants in CKB

	Without NAFLD (n=607)	NAFLD (n=651)	Overall (n=1258)
Age (SD), y	53.3 (9.2)	52.1 (8.6)	52.1 (8.9)
Women (%)	61.7	75.5	68.9
Urban region (%)	13.3	19.2	16.4
Ever regular smoking (%)			
Men	81.2	74.0	78.3
Women	0.8	1.3	1.0
Weekly alcohol consumption (%)			
Men	40.0	24.2	33.2
Women	0.9	0.8	0.9
Blood and anthropometry			
Systolic blood pressure (SD), mmHg	123.6 (17.7)	126.3 (18.0)	123.0 (17.8)
Diastolic blood pressure (SD), mmHg	75.8 (9.0)	76.6 (8.8)	76.2 (8.9)
BMI (SD), kg/m^2	21.5 (2.6)	23.7 (2.7)	22.6 (2.8)
Waist circumference (SD), cm	73.9 (8.2)	79.5 (7.7)	76.8 (8.4)
Biochemistry			
ALT (SD), U/L	15.4 (8.4)	17.8 (10.8)	16.7 (9.7)
AST (SD), U/L	22.8 (7.9)	22.9 (7.6)	22.9 (7.8)
GGT (SD), U/L	17.6 (14.0)	20.8 (15.6)	19.3 (14.8)
Total cholesterol (SD), mmol/L	4.2 (0.8)	4.4 (0.8)	4.3 (0.8)
HDL-C (SD), mmol/L	1.3 (0.3)	1.2 (0.3)	1.3 (0.3)
LDL-C (SD), mmol/L	2.4 (0.7)	2.5 (0.7)	2.5 (0.7)
Triglycerides (SD), mmol/L	1.2 (0.6)	1.5 (0.8)	1.4 (0.7)
Albumin (SD), g/L	46.0 (2.4)	46.2 (2.3)	46.1 (2.3)
Creatinine (SD), umol/L	63.5 (12.1)	64.3 (11.3)	63.9 (11.7)
Fasting glucose (SD), mmol/L	4.5 (0.8)	4.5 (0.7)	4.5 (0.8)

Adjusted for age, sex, and region (urban/rural) when appropriate. *P*-values between groups were <0.05 except for smoking and drinking status in women, systolic blood pressure, AST, and fasting glucose.

Abbreviations: SD, standard deviation; BMI, body mass index; ALT, alanine transaminase; AST, aspartate aminotransferase; GGT, gamma-glutamyl transpeptidase; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol.

Table S3. Mendelian randomization of the associations between BMI and metabolomic biomarkers in CKB

Metabolite	IVW		MR-Egger		Weighted median	
	SD difference	*p*-value	SD difference	*p*-value	SD difference	*p*-value
CDCA	0.30 (0.03, 0.57)	0.031	-0.22 (-0.75, 0.31)	0.415	0.09 (-0.22, 0.39)	0.578
Palmitoylcarnitine	0.42 (0.14, 0.70)	0.003	0.50 (-0.04, 1.03)	0.069	0.24 (-0.05, 0.53)	0.098
Hippuric acid	-0.32 (-0.60, -0.05)	0.023	-0.17 (-0.71, 0.37)	0.537	-0.20 (-0.48, 0.08)	0.163
DHA	0.15 (-0.13, 0.42)	0.300	-0.09 (-0.61, 0.43)	0.740	0.25 (-0.04, 0.53)	0.087
Proline	-0.06 (-0.32, 0.19)	0.627	0.19 (-0.03, 0.71)	0.479	0.16 (-0.13, 0.45)	0.274

SD difference of log-transformed metabolite concentration per 1-SD increase BMI. 2 stage least square (2SLS) was used, both stages adjusted for age, sex, 10 regions, 11 national principle components, GWAS array type, smoking status, alcohol status, and physical activity.

Table S4. Observational associations between metabolomic biomarkers and NAFLD in CKB

Metabolite	OR (95% CI)	*p*-value
CDCA	1.17 (1.03, 1.32)	0.016
Palmitoylcarnitine	1.12 (0.99, 1.26)	0.065
Hippuric acid	0.82 (0.73, 0.92)	0.001
DHA	0.96 (0.85, 1.07)	0.425
Proline	0.99 (0.87, 1.13)	0.868

OR of NAFLD per 1-SD increase log-transformed metabolite concentration. Logistic regression was adjusted for age, sex, 10 regions, smoking status, alcohol status, and physical activity.

Table S5. Genetic instruments of CDCA

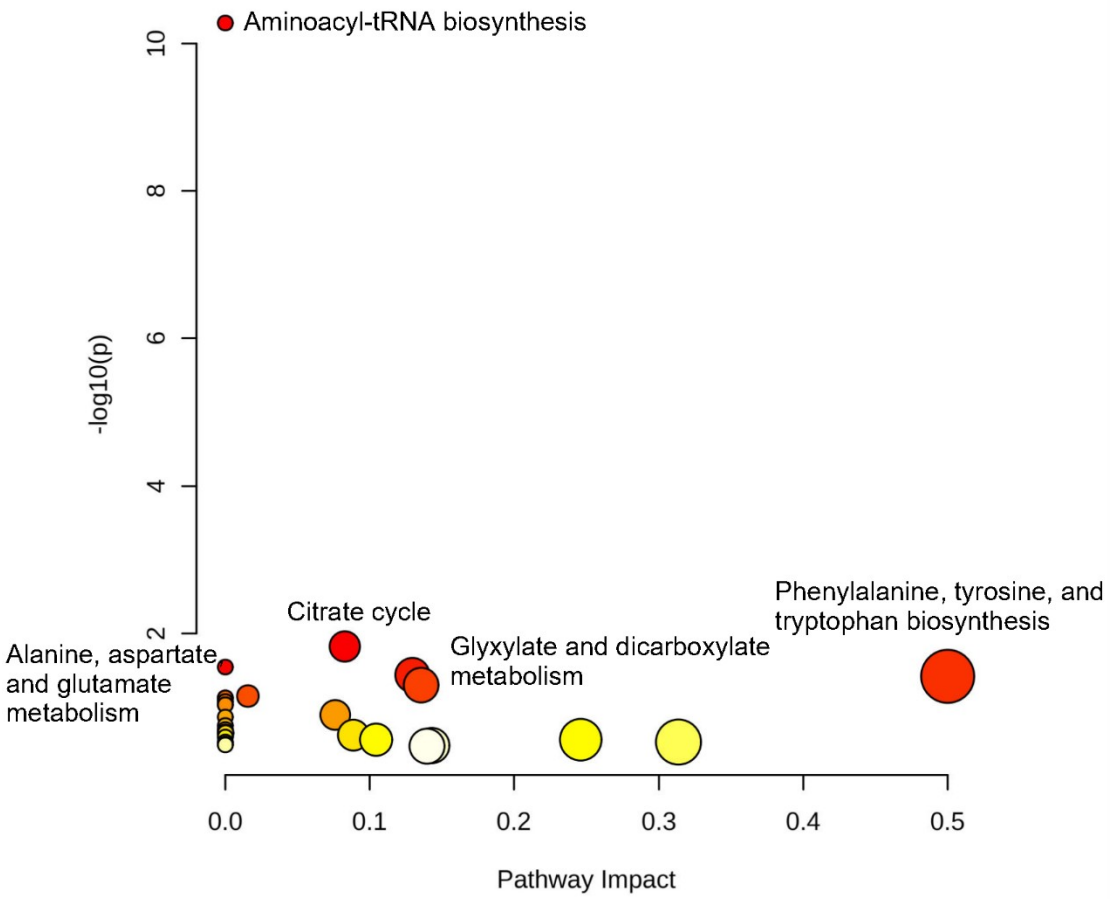
SNP	Effect allele	Other allele	Beta	SE	*P*-value	SNP	Effect allele	Other allele	Beta	SE	*P*-value
rs532148656	T	G	0.302	0.073	3.1E-05	rs39517	C	T	0.314	0.077	4.4E-05
rs531739947	G	A	0.542	0.131	3.4E-05	rs3001645	T	C	0.171	0.042	4.5E-05
rs537743188	T	A	1.822	0.463	8.4E-05	rs2116735	A	G	2.077	0.510	4.5E-05
rs2191323	C	T	0.189	0.035	6.7E-08	rs9306023	G	A	0.136	0.034	4.8E-05
rs10819283	A	G	0.196	0.043	4.1E-06	rs9910688	T	C	0.329	0.081	5.0E-05
rs4147736	A	G	0.160	0.035	5.0E-06	rs12067409	C	T	0.262	0.065	5.3E-05
rs12040283	C	T	0.200	0.044	5.5E-06	rs9589217	T	C	0.153	0.038	5.3E-05
rs113638355	G	A	0.190	0.042	5.7E-06	rs78750043	G	T	0.192	0.048	5.4E-05
rs72960488	T	A	0.308	0.068	5.9E-06	rs4773340	G	A	0.277	0.068	5.4E-05
rs1765132	G	A	0.181	0.041	7.9E-06	rs1401016	A	T	0.728	0.181	5.7E-05
rs112641336	G	A	0.191	0.043	8.1E-06	rs17075181	G	A	0.123	0.031	5.9E-05
rs12567117	C	T	0.224	0.050	8.4E-06	rs12918185	A	G	0.267	0.067	5.9E-05
rs34943535	T	G	0.146	0.033	8.8E-06	rs7641154	G	A	0.143	0.036	6.0E-05
rs6889982	G	A	0.241	0.054	9.6E-06	rs3787590	C	T	0.212	0.053	6.0E-05
rs6864571	C	G	0.162	0.037	1.0E-05	rs77932068	C	T	0.129	0.032	6.2E-05
rs17628836	A	G	0.399	0.091	1.0E-05	rs12590682	A	G	0.142	0.036	6.3E-05
rs62098118	T	C	0.144	0.033	1.1E-05	rs10880014	T	C	0.407	0.102	6.4E-05
rs28627942	T	C	0.159	0.036	1.3E-05	rs4784757	A	G	0.158	0.039	6.5E-05
rs12692516	T	C	0.139	0.032	1.4E-05	rs111367688	G	A	1.171	0.294	6.7E-05
rs1439866	C	T	0.136	0.031	1.4E-05	rs7011671	C	T	0.12	0.03	6.7E-05

rs4953470	T	C	0.513	0.119	1.6E-05	rs113234609	A	G	0.376	0.092	7.1E-05
rs140551001	A	G	0.327	0.076	1.6E-05	rs61881664	C	T	0.255	0.064	7.3E-05
rs62541880	T	C	0.186	0.043	1.6E-05	rs115718639	A	T	0.588	0.148	7.4E-05
rs697746	T	A	0.138	0.032	1.7E-05	rs60185446	T	C	1.176	0.297	7.5E-05
rs2478804	C	T	0.295	0.069	1.7E-05	rs188614021	G	A	0.155	0.039	7.5E-05
rs2286956	A	G	0.194	0.045	1.8E-05	rs4871142	T	G	0.134	0.034	7.7E-05
rs11908635	C	T	0.153	0.036	1.9E-05	rs10817084	A	G	0.146	0.037	7.7E-05
rs115227145	C	T	0.253	0.059	2.0E-05	rs10469208	A	C	0.549	0.139	7.7E-05
rs176379	G	A	0.224	0.052	2.0E-05	rs11541286	C	T	0.857	0.217	7.8E-05
rs6699319	T	C	0.172	0.040	2.1E-05	rs10845895	G	A	0.205	0.052	8.2E-05
rs13252823	C	T	0.202	0.047	2.2E-05	rs4558132	T	C	0.168	0.043	8.3E-05
rs7848692	C	G	0.272	0.064	2.3E-05	rs3911492	C	T	0.143	0.036	8.4E-05
rs12580089	T	C	0.158	0.037	2.5E-05	rs202232579	C	T	0.518	0.132	8.4E-05
rs10838596	G	A	0.150	0.036	2.7E-05	rs62089137	G	A	0.127	0.032	8.4E-05
rs73079525	A	C	0.216	0.052	3.0E-05	rs62429292	T	C	0.271	0.069	8.5E-05
rs17047930	G	A	0.292	0.070	3.1E-05	rs6988626	T	A	0.196	0.050	8.5E-05
rs654704	C	A	0.140	0.034	3.3E-05	rs678450	A	G	0.156	0.040	8.5E-05
rs7702063	G	A	0.190	0.046	3.4E-05	rs9812529	C	A	0.129	0.033	8.6E-05
rs34042015	T	A	0.636	0.154	3.5E-05	rs57851852	T	C	0.187	0.048	9.0E-05
rs200267100	T	C	0.137	0.033	3.9E-05	rs16982983	C	G	0.130	0.033	9.3E-05
rs10220164	A	G	0.268	0.065	4.1E-05	rs17279312	C	G	0.359	0.092	9.5E-05

Table S5. Continued

SNP	Effect allele	Other allele	Beta	SE	*P*-value	SNP	Effect allele	Other allele	Beta	SE	*P*-value
rs10971393	A	G	0.197	0.051	9.8E-05	rs55774621	T	C	0.122	0.031	9.6E-05
rs151141078	T	C	0.184	0.047	9.8E-05						

Figure S1. Pathway analysis of differential metabolomic biomarkers pre- and post-surgery



Pathway Name	Match status	*p*-value
Aminoacyl-tRNA biosynthesis	9/48	5.26E-11
Citrate cycle (TCA cycle)	2/20	0.015
Alanine, aspartate, and glutamate metabolism	2/28	0.029
Glyoxylate and dicarboxylate metabolism	2/32	0.037
Phenylalanine, tyrosine, and trypotophan biosynthesis	1/4	0.038

Figure S2. The significant differences in 19 metabolomic biomarkers before and after surgery

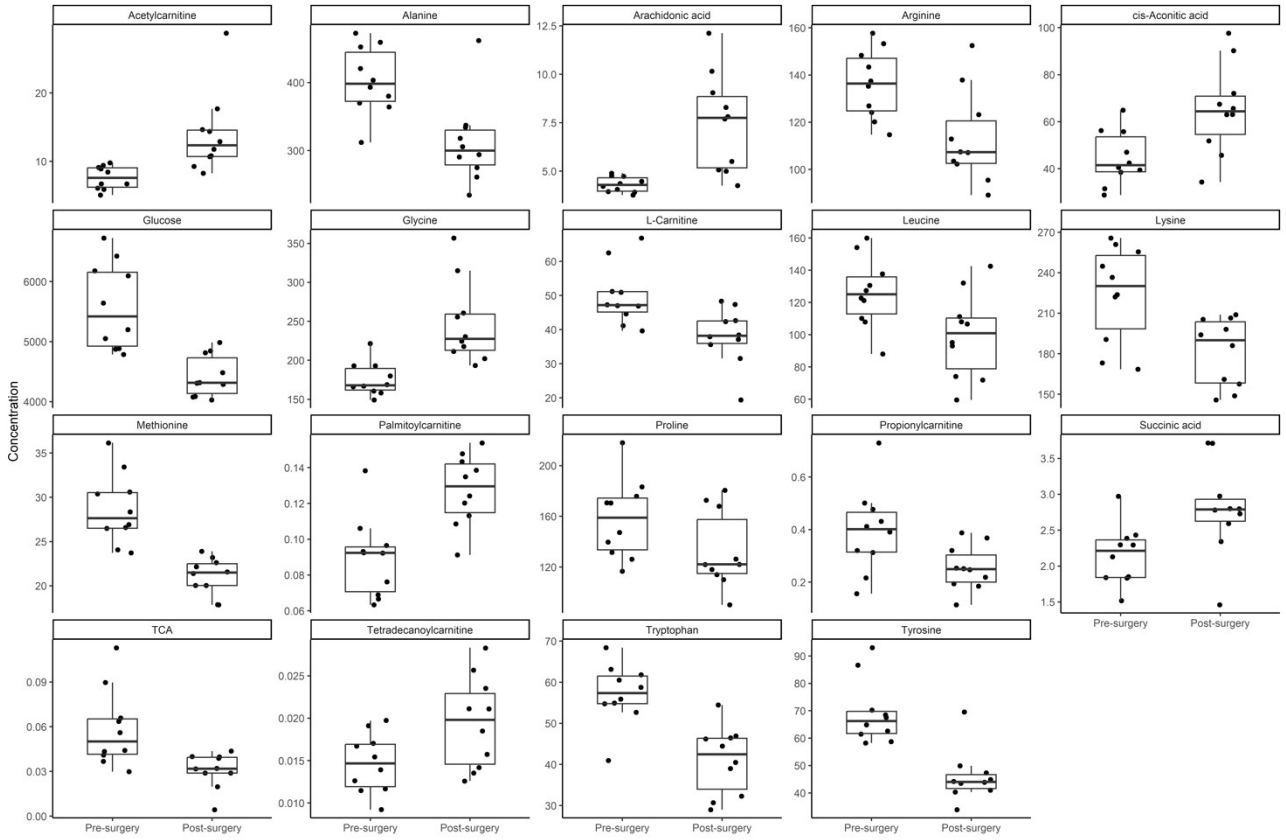
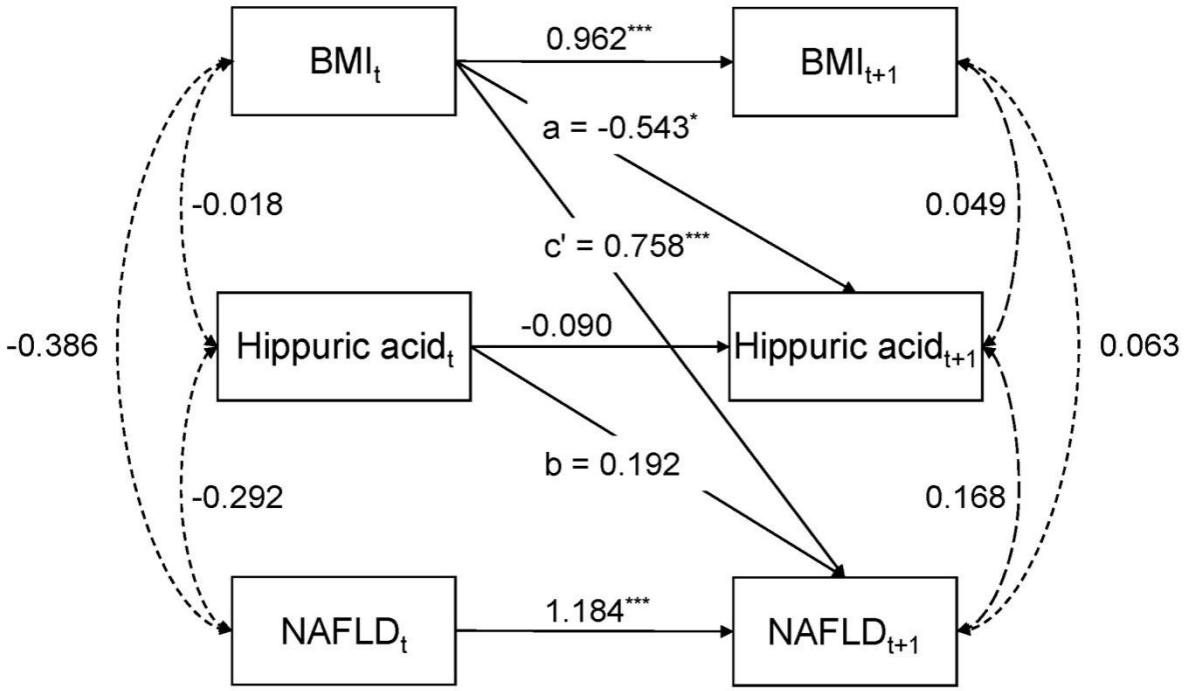


Figure S3. Cross-lagged panel model of the mediation effect of hippuric acid between BMI and NAFLD remission



Proportion mediated for hippuric acid = ab/(ab+c') * 100% = -15.9%
(*P* = 0.271)