Supporting Information

Walker *et al.*

Materials and methods

Animals: The C57BL/6J and C57BL/6N strains emerged from the ancestral C57BL/6 in the 1940s. C57BL/6J and C57BL/6N mice were bred, housed and handled according to the federal animal welfare guidelines. Mouse husbandry was conducted under a continuously controlled hygiene standard according to the Federation of European Laboratory Animal Science Associations (FELASA). All mice were housed in techniplast cages under standard vivarium conditions (mean ambient temperature of 21 ± 1 °C, 12:12 hours light: dark cycle) and had free access to food and water in the experiments. Genetic background characterization in tail-clip samples from C57BL/6N and C57BL/6J mice bred at our laboratory was performed by Taconic (New York, U.S.A). Illumina medium density SNP panel analysis confirmed 100.0% similarity to the respective genotype reference. Microarray analyses revealed a marked down regulation of nnt-transcripts in C57BL/6J livers compared to C57BL/6N.

UPLC-TOF-MS/MS Experiments

Further MS experiments were followed with pooled cecal samples in order to increase the concentration of metabolites. We evaluated the putative metabolites by applying MS/MS experiments with time of flight (TOF) mass spectrometer (MaXis) (Bruker Daltonics) coupled to the Ultraperformance Liquid Chromatography (UPLC) system (ACQUITY™; Waters, Milford, MA). Before analysis, the mass spectrometer was calibrated with 5 ppm of arginine solution. Before MS/MS analysis separation was performed using reverse chromatography with UPLC system. The fragmentation

experiments were exhibited in automated MS/MS mode of TOF system. Following parameters were applied: capillary voltage of 4000 V and end plate offset to -500 V, dry gas flow rate of 8 L/min, dry gas temperature to 200 °C, nebulizer gas flow rate of 2.0 bar. The mass range was set from 50 to 1200 and scan rate of 5 Hz with a rolling average of 2 by acquiring profile spectra. The parameters for automated MS/MS fragmentation were set to a number of five precursor ions and an intensity cut-off >1000 with a collision energy range between 30 and 40 eVolt.

Pre-processing and statistical analysis of metabolomics data

Before statistically evaluating the experiment, we have adopted several pre-processing tasks. The spectra alignment leads to a sizable number of zero values associated to the variable mass signal. Thus, the variance of this variable needs to be stabilized excluding mass signal values occurring in less than 20% of all samples. The multivariate statistical analysis was performed with either SIMCA-P 12 (Umetrics, Umea, Sweden), Genedata Analyst 7.5 (Genedata Solution in Silico, Basel, Switzerland) or SAS (SAS Institute GmbH, Germany). Heatmap visualisation was performed using Hierarchical Clustering Explorer by normalizing each mass signals by the standard deviation (Jinwook and Shneiderman 2002). The so inferred discriminative variables/mass signals were uploaded into MassTRIX (Suhre and Schmitt-Kopplin 2008). MassTRIX enables to upload high precision mass spectrum data and assign the detected mass signals into metabolites, using the information of the theoretical monoisotopic masses within a selected error in ppm. The information of monoisotopic masses of metabolites is derived from different databases including KEGG (Kyoto Encyclopedia of Genes and Genomes), LIPID MAPS (LIPID Metabolites and Pathways Strategy) and HMDB (Human Metabolome Database)

(Kanehisa and Goto 2000, Sud et al 2007, Wishart et al 2009). Mass signals that were both significant and discriminative have been uploaded into MassTRIX, within an error range of 1 ppm for both modes. *Bacteriodes vulgatus* has been chosen as reference species in the KEGG database. The significance of the discriminative mass signals, derived from S-Plot was confirmed by using of a non-parametric statistical test (Wilcoxon-Mann-Whitney test with $p<0.05$).

Comparison of measured MS/MS data was done by using METLIN database, which also provides amongst metabolite annotation an extensive collection of MS/MS data for metabolites (Smith et al 2005). The annotation of selected parent ions was done within an error range of 10 ppm, searching and comparison of given MS/MS data was done visually. Molecular formula calculation of measured parent and fragment ions was performed using DataAnalysis Version 4.0 SP2 (Bruker Daltonik GmbH, Bremen, Germany).

Elemental composition mass difference network analysis and visualization

The calculation of molecular formulas and construction of connections due to preselected mass differences reference list was done by in-house software NetCalc (Tziotis et al 2011). In detail, for each mass signal, molecular formula was calculated, including carbon, hydrogen, oxygen, nitrogen, sulfur and phosphorus within an error of 0.2 ppm (CHNOSP). Mass signals with valid molecular formulas were taken for network visualization. Possible biochemical transformations (representing mass differences) such as hydroxylation, taurine or sulfate conjugation (OH: 15.994915 (exact mass); $C_2H_5NO_2S$: 107.00410; SO_3: 79.95682) were calculated between mass signals with valid molecular formulas. All mass differences, taken for calculation are summarized in Table S4. The mass differences are subsequently used as

connections between nodes (edges) and afterwards a construction of mass difference network was possible. Network visualization was performed using Gephi 0.8.1 beta version (Bastian et al 2009) by applying the ForceAtlas2 layout, which is based on attraction and repulsion of nodes due to the number of found mass differences, respectively.

Results

Identification of diphloretoylputrescine

As described in Materials and methods, we confirmed several metabolites by performing automated MS/MS experiments. We intentionally searched for all mass signals that were elaborated from the S-Plot, illustrated in Figure 2, B. The mass signal of diphloretoylputrescine was 383.197532 with calculated molecular formula of $C_{22}H_{28}N_2O_4$, which was annotated as [#16-Methoxy-2,3-dihydro-3-hydroxytabersonine ($[M-H]^-$) #Rhyncophylline; Rhynchophylline ($[M-H]^-$) #Isorhynchophylline ($[M-H]^-$)] in MassTRIX, which was refused after carefully looking on the acquired MS/MS results, shown in Figure S2, A. The corresponding fragments of the daughter ion with their respective molecular formula are shown in Figure S2 A. The highest fragment (1) is assumed to be derived through a loss of hydroxycinnamylaldehyde of the precursor mass, resulting in (1) and the second fragment of (2). The fragment of (3) was presumed to be derived through C-N cleavage of the precursor mass and corresponding fragment (4). Finally, the possible structure, shown in Figure S2, B of $C_{22}H_{28}N_2O_4$ could be assumed as the most plausible one, consisting of two molecules of hydroxyphenylpropionic acid (also known as phloretic acid or dihydroumaric acid) and a putrescine molecule as the linkage (Figure S2, B). This metabolite is likely derived due to enzymatic condensation of the carboxyl group of hydroxyphenylpropionic acid and aminogroup of putrescine and will be named diphloretoylputrescine.

Figures

Figure S1

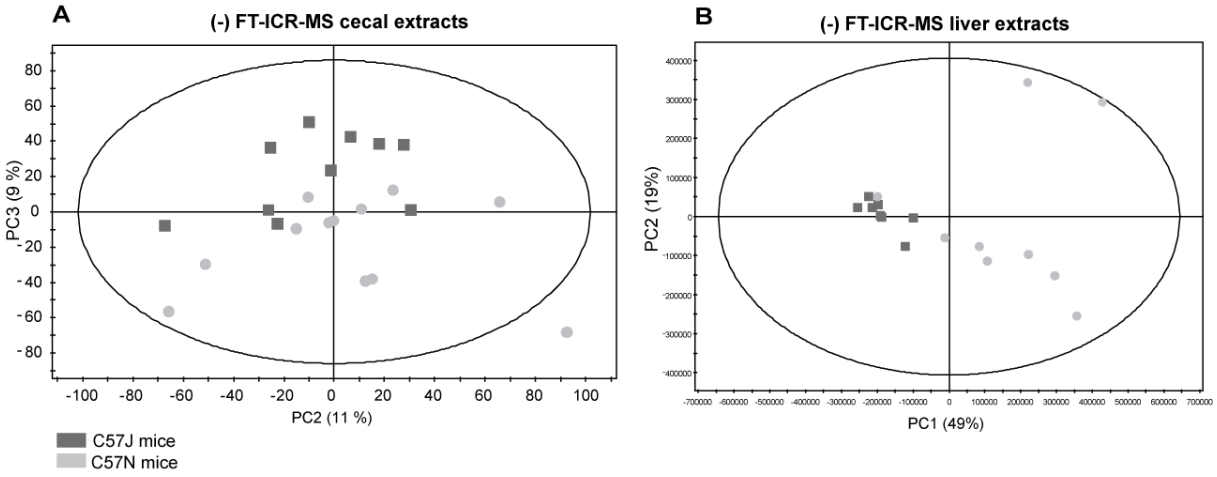

Figure S1 Unsupervised multivariate analysis of cecal and liver extracts, measured in (-) FT-ICR-MS

(A) Principal component analysis (PCA) score scatter plot from an unsupervised multivariate statistical model of cecal and liver samples (B) from C57J and C57N mice, derived from FT-ICR-MS data, points are colored according the group of C57J (dark grey) and C57N (light grey) mice. The plot is showing a possible separation between the two classes by the second principal and third components with $R^2X(cum)=0.41$ and $Q^2(cum)=0.11$ for cecal samples (A). For liver metabolome a possible separation (B) was shown between the groups in first and second component with $R^2X(cum)=0.68$ and $Q^2(cum)=0.53$.

Figure S2

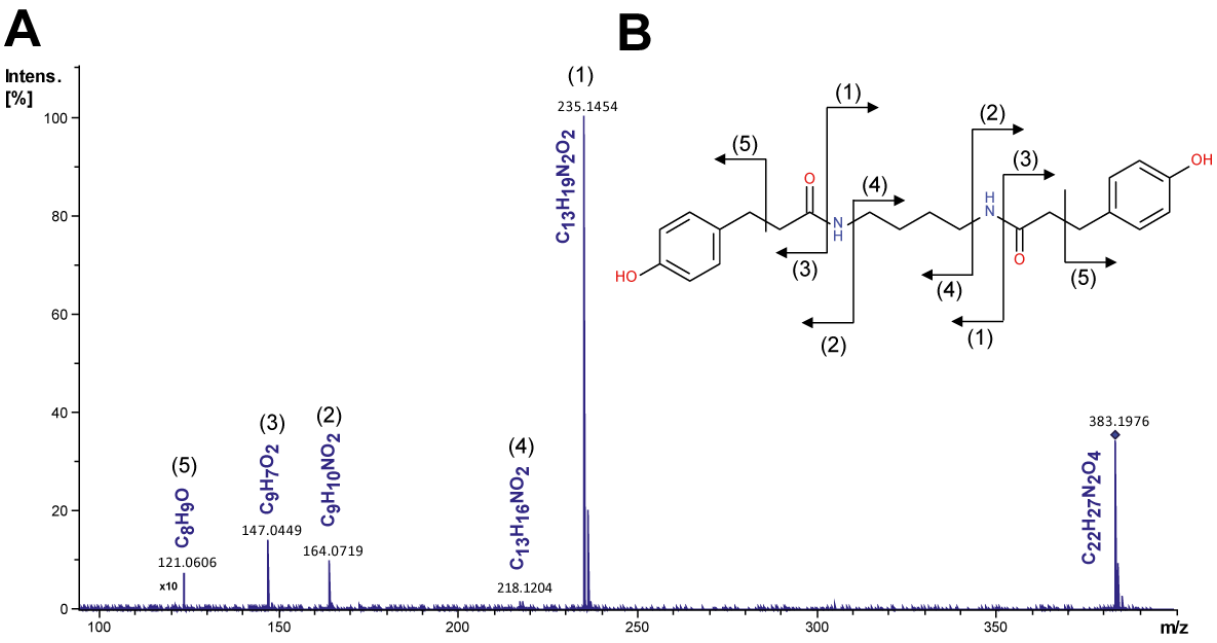

Figure S2 Identification of diphlyretoylputrescine

(A) MS/MS of the novel metabolite called diphloretoylputrescine; (B) Most plausible structure of the diphloretoylputrescine

Figure S3

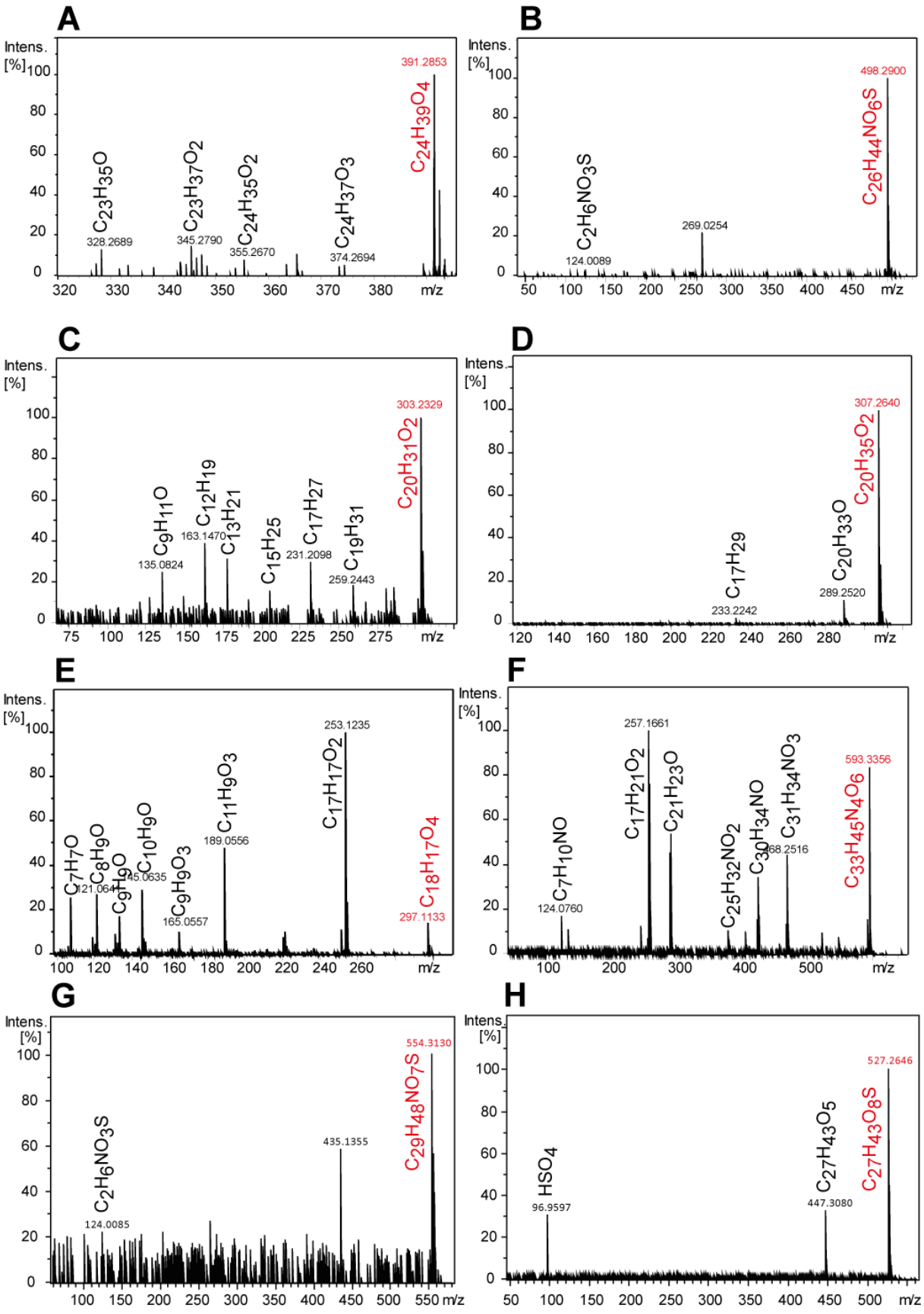
A
Intens. [%]
391.2853
C24H39O4
C23H35O
328.2689
C23H37O2
345.2790
C24H35O2
355.2670
C24H37O3
374.2694
m/z
B
Intens. [%]
498.2900
C26H44NO6S
C2H6NO3S
124.0089
269.0254
m/z
C
Intens. [%]
303.2329
C20H31O2
C9H11O
135.0824
C12H19
163.1470
C13H21
C15H25
231.2098
C17H27
C19H31
259.2443
m/z
D
Intens. [%]
307.2640
C20H35O2
C17H29
233.2242
C20H33O
289.2520
m/z
E
Intens. [%]
253.1235
C17H17O2
C11H9O3
189.0556
C10H9O
C8H9O
121.0647
C7H7O
C9H9O2
145.0635
C9H9O3
165.0557
C18H17O4
297.1133
m/z
F
Intens. [%]
257.1661
C17H21O2
C21H23O
C7H10NO
124.0760
C25H32NO2
C30H34NO
C31H34NO3
368.2516
593.3356
C33H45N4O6
m/z
G
Intens. [%]
554.3130
C29H48NO7S
C2H6NO3S
124.0085
435.1355
m/z
H
Intens. [%]
527.2646
C27H43O8S
C27H43O5
447.3080
HSO4
96.9597
m/z

Figure S3 Selected MS/MS of eight_metabolites, represented each class of described cecal metabolites:

(A) Deoxycholic acid, (B) Taurodeoxycholic acid, (C) Arachidonic acid (C20:4), (D) Eicosadienoic acid (C20:2); (E) Enterolactone, (F) L-Urobilin, (G) Taurodihydroxycholestanoic acid, (H) Dihydroxyoxocholestanoic acid sulfate: MS/MS experiments are performed in negative mode of the TOF-MS instrument, molecular formulas of parent and fragment ions were calculating using SmartFormula function of DataAnalysis; molecular formulas are displayed as deprotonated species.

Tables

Table S1

Mean relative abundances of sequences at the family level showing significant differences between cecum samples of C57J and C57N mice

Family	C57N mice [%]	C57J mice [%]	p-value	adj. p-value
Ruminococcaceae	35.4	15.4	0.0002	0.0043
Helicobacteraceae	0.7	10	0.0004	0.0044
Erysipelotrichaceae	1.4	14.4	0.0008	0.006
Bacteroidaceae	1.4	17.5	0.0025	0.0127
Deferribacteraceae	10.5	2.6	0.0032	0.0139

Table S2 Overview of percentages and p-values of highest abundant OTUs between the cecum samples of the two mouse strains. Taxonomic assignment was based on RDP trainset 7, similarities with the closest cultivated relatives was calculated in ARB database after alignment with SINA aligner

					Taxonomic assignment						
OTU Nr.	C57N mice [%]	C57J mice [%]	p-value[a]	p-value [adj.][b]	Phylum	Class	Order	Family	Genus	Closest cultivated relative similarity [%]	ACC number
Otu6	0	11.2	0.0001	0.00805	*Firmicutes*	*Erysipelo-trichia*	*Erysipelo-trichales*	*Erysipelo-trichaceae*	*Erysipelo-trichaceae_inc.sed.*	*Allobaculum stercoricanis [91.8]*	AJ417075
Otu2	1	12.5	0.0018	0.01656	*Bacteroidetes*	*Bacteroidia*	*Bacteroidales*	*Bacteroidaceae*	*Bacteroides*	*B. fluxus [100]; B.uniformis [99.5]; B. rodentium [99.5]*	AB490802 AB510711 AB531489
Otu12	0.3	4.4	0.0008	0.01309	*Bacteroidetes*	*Bacteroidia*	*Bacteroidales*	*Bacteroidaceae*	*Bacteroides*	*B. acidifaciens [100]; B.xylanisolvens [100]*	EU136694 AB510713
Otu5	0.7	9.7	0.0006	0.01263	*Proteo-bacteria*	*Epsilonproteo-bacteria*	*Campylobacterales*	*Helico-bacteraceae*	*Helicobacter*	*H.hepaticus [100]; H. bilis [97.9]*	AJ007931 AY578097
Otu1	15.8	6.4	0.0011	0.01309	*Firmicutes*	*Clostridia*	*Clostridiales*	*Ruminococcaceae*	*Pseudo-flavonifractor*		
Otu9	3.1	1	0.005	0.02559	*Firmicutes*	*Clostridia*	*Clostridiales*	*Ruminococcaceae*	*Anaerotruncus*	*Anaerotruncus colihominis [95.9]*	ABGD02000031
Otu3	9.9	2.5	0.0042	0.02262	*Deferri-bacteres*	*Deferribacteres*	*Deferribacterales*	*Deferri-bacteraceae*	*Mucispirillum*	*Mucispirillum schaedleri [100]*	AY387670

a: p- values were calculated by using a non-parametric univariate statistical test Wilcoxon-Mann-Whitney test; b: multiple testing corrected by Benjamini-Hochberg algorithm

Table S3 Summary with detailed information of putative metabolites that are significantly changed between C57J and C57N mice concerning cecal and liver metabolites displayed in Figure 3 - 5

Cecal metabolites

Compound class	Mass (avg.)	Compound Name	Monoisotopic mass	C57J mice (n=10)	C57N mice (n=12)	p-value	Molecular Formula
C24 Bile acids							
	355.264269	Cholandienoic acid	356.271516	1.14E+07	7.05E+06	0.029559	$C_{24}H_{36}O_2$
	371.259162	Oxocholenoic acid	372.266431	2.45E+07	1.56E+07	0.003005	$C_{24}H_{36}O_3$
	373.274756	Oxocholanoic acid	374.28208	9.79E+07	6.65E+07	n.s.	$C_{24}H_{38}O_3$
	375.29046	Lithocholic acid	376.297729	1.49E+09	9.91E+08	n.s.	$C_{24}H_{40}O_3$
	391.285157	Deoxycholic acid	392.292644	1.57E+10	9.36E+09	0.024968	$C_{24}H_{40}O_4$
	401.233323	Trixoxocholanoic acid	402.240611	1.65E+07	1.89E+07	0.024968	$C_{24}H_{34}O_5$
	407.280086	Cholic acid	408.287559	8.58E+09	7.69E+09	n.s.	$C_{24}H_{40}O_5$
	421.259526	Trihydroxyoxocholanoic acid	422.266825	2.18E+08	2.62E+08	0.004578	$C_{24}H_{38}O_6$
	423.275108	Tetrahydroxycholanoic acid	424.282474	5.43E+08	7.15E+08	0.017608	$C_{24}H_{40}O_6$
C_{24} Taurine conjugated Bile acids							
	478.263234	Taurooxocholenoic acid*	479.27053	5.51E+06	1.04E+06	0.006300	$C_{26}H_{41}O_5N_1S_1$
	480.278841	Taurooxocholanoic acid*	481.286179	7.49E+06	1.20E+06	0.024220	$C_{26}H_{43}O_5N_1S_1$
	482.294473	Taurolithocholic acid	483.301828	1.20E+07	1.78E+06	0.013170	$C_{26}H_{45}O_4N_1S_1$
	494.25797	Taurodioxocholanoic acid*	495.265445	1.65E+07	4.78E+06	0.004580	$C_{26}H_{41}O_6N_1S_1$
	496.273817	Taurohydroxyoxocholanoic acid*	497.281094	1.27E+08	3.75E+07	0.014700	$C_{26}H_{43}O_6N_1S_1$
	498.289477	Taurodeoxycholic acid	499.296743	5.34E+08	2.25E+08	0.005620	$C_{26}H_{45}O_5N_1S_1$
	510.252936	Taurohydroxydioxocholanoic acid*	511.26036	1.95E+07	7.31E+06	0.001230	$C_{26}H_{41}O_7N_1S_1$
	514.284132	Taurocholic acid	515.291658	3.88E+09	1.10E+09	0.004580	$C_{26}H_{45}O_7N_1S_1$
	530.279358	Taurotetrahydroxycholanoic acid*	531.286573	5.93E+07	8.37E+06	0.006320	$C_{26}H_{45}O_8N_1S_1$
	594.240876	Taurocholic acid 3-sulfate	595.248475	6.92E+07	2.92E+05	0.006710	$C_{26}H_{45}O_{10}N_1S_2$
Other conjugated C24 Bile acids							
	448.306922	Glycodeoxycholic Acid	449.314107	1.66E+06	1.00E+06	0.01869	$C_{26}H_{43}O_4N_1$
	464.301993	Glycocholic acid	465.309022	8.09E+06	2.12E+06	0.03425	$C_{26}H_{43}NO_6$
	512.268599	Sulfolithocholylglycine	513.276009	2.02E+08	1.08E+08	0.00686	$C_{26}H_{43}O_7N_1S_1$
	544.258192	Sulfocholylglycine	545.265839	8.95E+06	1.40E+06	0.00093	$C_{26}H_{43}O_9N_1S_1$
C_{27} Taurine conjugated Bile acids							
	520.309819	Taurodihydrocholestenoic acid*	521.317477	2.09E+06	2.91E+05	0.02600	$C_{29}H_{47}O_5N_1S_1$
	522.325673	Taurocholestenoic acid*	523.333126	5.65E+06	1.16E+06	0.01058	$C_{29}H_{49}O_5N_1S_1$
	536.30475	Taurodioxocholestanoic acid	537.312392	8.13E+06	7.68E+05	0.00397	$C_{29}H_{47}O_6N_1S_1$
	538.320689	Taurodihydroxycholestenoic acid	539.328041	2.31E+07	3.57E+06	0.00457	$C_{29}H_{49}O_6N_1S_1$
	540.336074	Taurodihydroxycholestanoic acid	541.343691	1.58E+07	3.69E+06	0.00457	$C_{29}H_{51}O_6N_1S_1$
	552.29969	Taurodihydroxyoxocholestenoic acid	553.307307	1.02E+07	1.54E+06	0.00736	$C_{29}H_{47}O_7N_1S_1$
	554.31551	Taurodihydroxycholestenoic acid	555.322956	4.06E+07	4.26E+06	0.00372	$C_{29}H_{49}O_7N_1S_1$
	556.330806	Taurotrihydroxycholestanoic acid	557.338606	2.49E+07	3.79E+06	0.02253	$C_{29}H_{51}NO_7S$
	570.310363	Taurotetrahydroxycholestenoic acid	571.317871	1.19E+07	1.43E+06	0.00774	$C_{29}H_{49}O_8N_1S_1$
	572.326197	Taurotetrahydroxycholestanoic acid	573.333521	6.63E+06	1.33E+06	0.04102	$C_{29}H_{51}O_8N_1S_1$
Sulfates of C_{27} Bile acids							
	511.273369	Dihydroxycholestenoic acid sulfate	512.280759	1.06E+08	7.57E+07	0.0176	$C_{27}H_{44}O_7S_1$
	527.268213	Dihydroxyoxocholestanoic acid sulfate	528.275674	9.67E+08	6.10E+08	0.0296	$C_{27}H_{44}O_8S_1$
Fatty acids							
	297.279888	Pristanic acid	298.28717	2.27E+08	2.75E+08	0.02497	$C_{19}H_{38}O_2$
	299.201693	Retinoic Acid	300.208919	5.40E+07	2.74E+07	0.00686	$C_{20}H_{28}O_2$
	301.217307	Eicosapentaenoic acid (C20:5)	302.224568	5.46E+07	2.91E+07	0.01761	$C_{20}H_{30}O_2$
	303.232832	Arachidonic Acid (C20:4)	304.240217	2.96E+09	9.30E+08	0.01761	$C_{20}H_{32}O_2$
	307.264205	Eicosadienoic acid (C20:2)	308.271516	2.01E+09	3.69E+09	0.0101	$C_{20}H_{36}O_2$
	319.227887	Hydroxyeicosatetraenoic acid	320.235132	2.68E+07	1.42E+07	0.0084	$C_{20}H_{32}O_3$
	327.233001	Docosahexanoic acid (C22:6)	328.240217	5.76E+08	1.59E+08	0.0122	$C_{22}H_{32}O_2$
	333.279892	Docosatrienoic acid (C22:3)	334.287165	9.55E+07	3.03E+08	0.0019	$C_{22}H_{38}O_2$
	335.222773	Leuktotriene B4	336.230047	4.92E+07	2.45E+07	0.0046	$C_{20}H_{32}O_4$
	335.295584	Docosadienoic acid (C22:2)	336.302814	7.78E+07	1.14E+08	0.02101	$C_{22}H_{40}O_2$
	351.21769	Hydroxy Leuktotriene B4	352.224974	5.72E+08	2.99E+08	0.01470	$C_{20}H_{32}O_5$
Endocannabinoids							
	324.290713	N-Oleoylethanolamine	325.298063	5.25E+06	1.31E+07	0.029560	$C_{20}H_{39}O_2N_1$
	326.306606	Stearoylethanolamine	327.313713	1.24E+07	3.29E+07	0.006860	$C_{20}H_{41}NO_2$
	352.322092	Anandamide (20:1)	353.329362	8.98E+06	1.42E+07	0.017610	$C_{22}H_{43}O_2N_1$
	380.353453	Erucicoylethanolamine	381.36066	4.69E+05	1.79E+06	0.021210	$C_{24}H_{47}O_2N_1$
	410.236976	N-arachidonoyl taurine	411.2443	7.64E+06	5.91E+05	0.032172	$C_{22}H_{37}NO_4S$
	435.25176	Lysophosphatidic Acid (18:1)	436.258977	9.42E+06	2.10E+07	0.004530	$C_{21}H_{41}O_7P_1$
Urobilinoids							
	583.25568	Bilirubin	584.263472	2.37E+07	2.92E+07	n.s.	$C_{33}H_{36}N_4O_6$
	587.28717	D-Urobilin	588.29477	2.69E+06	8.17E+06	0.006980	$C_{33}H_{40}N_4O_6$
	589.30305	D-Urobilinogen	590.310419	2.04E+07	1.06E+08	0.014670	$C_{33}H_{42}N_4O_6$
	591.31879	I-Urobilinogen	592.326068	2.12E+08	6.32E+08	0.012220	$C_{33}H_{44}N_4O_6$
	593.33436	L-Urobilin	594.341718	5.20E+08	9.73E+08	0.024970	$C_{33}H_{46}N_4O_6$
	595.34993	L-Urobilinogen	596.357367	5.05E+07	9.64E+07	0.021010	$C_{33}H_{48}N_4O_6$
Phenyl containing metabolites							
	297.113129	Enterolactone	298.1205028	1064943510	544119521.7	0.034857	$C_{18}H_{18}O_4$
	301.144571	Enterodiol	302.151809	29490958.7	14861003.58	0.008351	$C_{18}H_{22}O_4$
	383.197532	Diphloretoylputrescine	384.2048968	88953507.4	155890.25	0.000117	$C_{22}H_{28}N_2O_4$
Alpha oxidation metabolites							
	295.300709	Phytol	296.307916	20013279	38160069.5	0.014699	$C_{20}H_{40}O$
	297.279888	Pristanic acid	298.28717	227472118.4	275316008	0.024968	$C_{19}H_{38}O_2$
	311.295613	Phytanic acid	312.302831	1142037498	1048226437	0.468257	$C_{20}H_{40}O_2$
	327.290557	Hydroxylphytanic acid	328.297745	152262240	137092838.3	0.843191	$C_{20}H_{40}O_3$

Liver metabolites

Compound class	Mass (avg.)	Compound Name	Monoisotopic mass	C57J mice (n=9)	C57N mice (n=9)	p-value	Molecular Formula
C_{24} Bile acids							
	355.264264	Cholandienoic acid	356.271516	17263469.22	26538010.44	0.0071	$C_{24}H_{36}O_2$
	421.259583	Trihydroxyoxocholanoic acid	422.266825	2922257.778	5075806.444	0.0379	$C_{24}H_{38}O_6$
	423.27525	Tetrahydroxycholanoic acid	424.282474	5331722.444	9031965.778	0.0092	$C_{24}H_{40}O_6$
C_{24} Taurine conjugated Bile acids							
	478.263312	Taurooxocholenoic acid*	479.27053	5748338.889	11983061.44	0.0013	$C_{26}H_{41}O_5N_1S_1$
	480.278958	Taurooxocholanoic acid*	481.286179	9290278.889	16690640.56	0.0031	$C_{26}H_{43}O_5N_1S_1$
	482.294598	Taurolithocholic acid*	483.301828	13649887.56	25799485	0.0017	$C_{26}H_{45}O_4N_1S_1$
	494.258236	Taurodioxocholanoic acid*	495.265445	17022386.44	67782992.22	0.0017	$C_{26}H_{41}O_6N_1S_1$
	496.273915	Taurohydroxyoxocholanoic acid*	497.281094	148457734.7	653905144.9	0.0041	$C_{26}H_{43}O_6N_1S_1$
	498.289549	Taurodeoxycholic acid	499.296743	1467297301	5015382734	0.0054	$C_{26}H_{45}O_5N_1S_1$
	510.253166	Taurohydroxydioxocholanoic acid*	511.26036	21612837.78	86546642	0.0023	$C_{26}H_{41}O_7N_1S_1$
	514.284366	Taurocholic acid	515.291658	6824737109	32612640256	0.0031	$C_{26}H_{45}O_7N_1S_1$
	530.279368	Taurotetrahydroxycholanoic acid*	531.286573	20724202.11	68405396.22	0.0071	$C_{26}H_{45}O_8N_1S_1$
	594.241164	Taurocholic acid 3-sulfate	595.248475	860191.1111	5388764	0.0023	$C_{26}H_{45}O_{10}N_1S_2$
	471.242178	Sulfodeoxycholic acid	472.249461	0	2691166.556	0.0045	$C_{24}H_{40}O_7S_1$
	512.268788	Sulfolithocholylglycine	513.276009	582727987.6	1710337072	0.0071	$C_{26}H_{43}O_7N_1S_1$
	528.263706	Sulfodeoxycholylglycine	529.270924	6083828.778	19780599.44	0.0041	$C_{26}H_{43}O_8N_1S_1$
C_{27} Bile acids							
	433.332423	Dihydroxycholestanoic acid	434.33961	3418680	1468676.667	0.0052	$C_{27}H_{46}O_4$
	465.322061	Tetrahydroxycholestanoic acid	466.329439	10333437.78	16274843.56	0.0092	$C_{27}H_{46}O_6$
C_{27} Taurine conjugated Bile acids							
	536.304948	Taurodioxocholestanoic acid*	537.312392	1468867.778	4333670.333	0.0034	$C_{29}H_{49}O_6N_1S_1$
	538.320777	Taurodihydroxycholestenoic acid	539.328041	5055143.556	11447127.78	0.0031	$C_{29}H_{51}O_6N_1S_1$
	540.336458	Taurodihydroxycholestanoic acid	541.343691	3020145.778	7982480.556	0.0054	$C_{29}H_{53}O_6N_1S_1$
	552.300103	Taurodihydroxyoxocholestenoic acid*	553.307307	527308.8889	4214648.667	0.0012	$C_{29}H_{47}O_7N_1S_1$
	554.315757	Taurodihydroxycholestenoic acid*	555.322956	7282979	21595340.67	0.0023	$C_{29}H_{49}O_7N_1S$
	570.310652	Taurotetrahydroxycholestenoic acid*	571.317871	240135.3333	3128848.333	0.0031	$C_{29}H_{49}O_8N_1S_1$
Sulfates of C_{27} Bile acids							
	511.273586	Dihydroxycholestenoic acid sulfate*	512.280759	4084308.556	10314849.56	0.0009	$C_{27}H_{44}O_7S_1$
	527.268481	Dihydroxyoxocholestanoic acid sulfate*	528.275674	5326271.556	15981330.78	0.0054	$C_{27}H_{44}O_8S_1$
Fatty acids							
	299.201656	Retinoic Acid	300.208919	9482634.889	5528281	0.0193	$C_{20}H_{28}O_2$
	301.217283	Eicosapentaenoic acid (C20:5)	302.224568	26310711.11	96625823.56	0.0305	$C_{20}H_{30}O_2$
	303.232947	Arachidonic Acid (C20:4)	304.240217	429306588.4	1113333703	0.0031	$C_{20}H_{32}O_2$
	307.264248	Eicosadienoic acid (C20:2)	308.271516	71268765.33	153701189.3	0.0017	$C_{20}H_{36}O_2$
	327.232941	Docosahexanoic acid (C22:6)	328.240217	597811264	849598545.8	0.0469	$C_{22}H_{32}O_2$
	333.279915	Docosatrienoic acid (C22:3)	334.287165	11871785.11	26822734.67	0.0041	$C_{22}H_{38}O_2$
	335.295565	Docosadienoic acid (C22:2)	336.302814	4426979.222	9696600	0.0009	$C_{22}H_{40}O_2$
	435.251741	Lysophosphatidic Acid (18:1)	436.258977	17937294	13480449.89	0.0152	$C_{21}H_{41}O_7P_1$
	591.318904	I-Urobilinogen	592.326068	913467.4444	23332797.11	0.0017	$C_{33}H_{44}N_4O_6$
	593.334604	L-Urobilin	594.341718	4685591.111	49994084.11	0.0009	$C_{33}H_{46}N_4O_6$
	595.350206	L-Urobilinogen	596.357367	242354.3333	7881763.111	0.0008	$C_{33}H_{48}N_4O_6$
Alpha oxidation metabolites							
	297.279887	Pristanic acid	298.28717	1.49E+07	1.57E+07	n.s.	$C_{19}H_{38}O_2$
	311.295549	Phytanic acid	312.302831	52409692.44	83564167.56	0.0003	$C_{20}H_{40}O_2$
	327.290417	Hydroxyphytanic acid	328.297745	1.65E+06	1.77E+06	n.s.	$C_{20}H_{40}O_3$

Table S4 Mass difference list used for generation of the mass difference network illustrated in Figure 6, B

Name	Mass difference exact mass (theoretical)	Elemental composition CHNOSP
Hydrogenation	2.01565	H_2
Sulfurdioxide	63.961903	SO_2
Methylation	14.01565	CH_2
ß-Oxidation	28.0313	C_2H_4
Hydroxylation	15.994915	O
Oxidation	13.979265	$+O - H_2$
Glycine conjugation	57.021465	C_2H_3NO
Homocysteine conjugation	117.02483	C_4H_7NOS
Taurine conjugation	107.0041	$C_2H_5NO_2S$
Water	18.010565	H_2O
Sulfur	31.97207	S
Sulfonation	79.95682	SO_3
Hypotaurine conjugation	91.009186	C_2H_5NOS
Cysteine conjugation	103.009186	C_3H_5NOS
Carboxylation	43.98983	CO_2
Phosphorylation	79.96633	HPO_3
^{13}C isotope	1.003355	^{13}C
Glucosamination	163.08446	$C_6H_{13}NO_4$
Glucuronidation	176.03209	$C_6H_8O_6$
Glucose conjugation	162.05283	$C_6H_{10}O_5$
^{34}S isotope	1.995796	^{34}S
Tertiary amidation	14.003074	N
Secondary amidation	15.010899	NH
Primary amidation	16.018724	NH_2

References

Bastian M, Heymann S, Jacomy M (2009). *Gephi: An Open Source Software for Exploring and Manipulating Networks*.

Jinwook S, Shneiderman B (2002). Interactively exploring hierarchical clustering results [gene identification]. *Computer* **35:** 80-86.

Kanehisa M, Goto S (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research* **28:** 27-30.

Smith CA, Maille GO, Want EJ, Qin C, Trauger SA, Brandon TR *et al* (2005). METLIN: A Metabolite Mass Spectral Database. *Therapeutic Drug Monitoring* **27:** 747-751.

Sud M, Fahy E, Cotter D, Brown A, Dennis EA, Glass CK *et al* (2007). LMSD: LIPID MAPS structure database. *Nucleic Acids Research* **35:** D527-D532.

Suhre K, Schmitt-Kopplin P (2008). MassTRIX: mass translator into pathways. *Nucleic acids research* **36:** W481-484.

Tziotis D, Hertkorn N, Schmitt-Kopplin P (2011). Kendrick-analogous network visualisation of ion cyclotron resonance Fourier transform mass spectra: improved options for the assignment of elemental compositions and the classification of organic molecular complexity. *Eur J Mass Spectrom* **17:** 415-421.

Wishart DS, Knox C, Guo AC, Eisner R, Young N, Gautam B *et al* (2009). HMDB: a knowledgebase for the human metabolome. *Nucleic Acids Research* **37:** D603-D610.