Supplemental -Table of Content

Methods

Microbiome analyses

DNA extraction and 16S sequencing

DNA was extracted from fecal samples using the PowerLyzer PowerSoil DNA isolation kit (MoBio, Carlsbad CA). DNA quality and quantity were measured with the Nanodrop 1000 (Thermo Fisher Scientific, Wilmington DE). For each sample, the V4 region of bacterial 16S rRNA genes was amplified in triplicate reactions using the universal bacterial primer set 515F/806R, which amplifies bacterial and archaeal 16S rRNA genes[1]. PCR reactions contained 12 µl molecular biology grade water (Corning Cellgro, Atlanta GA), 10 µl 5 Prime Hot Master Mix (5 Primer, Inc., Gaithersburg MD), 1.0 µl each of the forward and reverse primers (each at 5 µM final concentration), and 1 µl genomic DNA. Reactions were held at 94°C for 3 min to denature the DNA, run for 35 cycles of amplification at 94°C for 45s, 50°C for 60s, and then 72°C for 90s, and completed with a final extension step of 10 min at 72°C. Amplicons from each sample were quantified using the PicoGreen dsDNA Assay Kit (Life Technologies, Eugene OR). Equal amounts of DNA from each sample were pooled, followed by PCR purification (Qiagen, Hilden, Germany). DNA concentrations in these sub-pools were quantified with the Qubit high sensitivity dsDNA Assay (Life Technologies), and combined at equal concentrations.

DNA sequencing was performed using the NYULMC Genome Technology Center Illumina MiSeq platform. The sequencing reactions generated 151-base-pair forward and reverse reads and a 12-base-pair barcode read.16S rRNA sequencing reads were de-multiplexed and quality filtered using QIIME version 1.9.1 with default settings as described[2]. Filtered reads were clustered into operational taxonomic units (OTUs) using a closed-reference OTU picking algorithm[3] with 97% similarity threshold against the GreenGenes database V13.8[4].

Alpha and beta diversity

Alpha diversity was estimated using rarefied OTU tables at 5,000 sequences per sample, using Faith's phylogenetic diversity metric[5]. Beta diversity was estimated using unweighted UniFrac[6] on the rarefied OTU tables, and principal coordinate analyses (PCoA) was performed using the resulting UniFrac distance matrices. Significance of alpha diversity differences over time was estimated using repeated measures ANOVA on the diversity values obtained at 5,000 sequences/sample. Differences were considered significant, when $p < 0.05$, unless otherwise indicated.

Random forests were constructed using OTUs in each sample as features of the model and as implemented in QIIME through the supervised_learning.py command, with 500 trees and leave-one-out error estimation[7]. Accuracy of the classifier was estimated by dividing the observed error by the percentage of subjects in the majority class, with ratios below 2 indicating poor classification accuracy[7]. The relative importance of each OTU in the model (i.e. its predictive power) was assessed by estimating the expected mean

decrease in accuracy when the OTU is excluded from the model, with higher decreases in accuracy corresponding to larger predictive power.
Analysis of enrichment/depletion of specific bacterial taxa in subgroups of patients (e.g. pre- versus post-surgery) was performed using LEfSe[8]. Significant taxa with linear discriminant log-scores larger than 2.0 were reported. A random forest classifier was used to test differences between subjects from the NYC and the BCN cohorts based on their microbiome composition[7].

Pathway Analyses
Bacterial functions were inferred from the 16S rRNA data using PICRUSt[9]. The OTU table was normalized by the 16S rRNA gene copy number of each OTU, and functions were predicted from the normalized table and expressed as counts of KEGG orthologs (KOs) for each sample, with each KO being annotated to a KEGG pathway. Differential analysis of pathways was performed using STAMP v2 with default parameters[10], and significance of differences in pathway enrichment/depletion was estimated using ANOVA and the Tukey-Kramer post-hoc test.

Network Analyses
Co-occurrence network analyses were performed by filtering out singletons (OTUs with a single count) and summarizing the remaining OTUs at the genus level. Co-occurrences between genera were then estimated using SparCC[11] with 20 iterations, 500 bootstrap replicates, and filtering non-significant correlations ($p < 0.05$, two-sided). The filtered correlations were then imported into Cytoscape v3.0.2[12] to visualize them as a co-occurrence network. The network layout was selected as edge-weighted Spring embedded metrics in Cytoscape. We then identified groups of highly connected bacteria within this network. This problem is equivalent to the problem of finding maximal cliques in a graph: a clique is defined as any subset of vertices (i.e. bacteria) such that they are all adjacent, and a maximal clique is a clique that cannot be further extended. Finding all maximal cliques in a graph can be efficiently performed using the Bron-Kerbosch algorithm[13]. Because bacteria within a maximal clique are, by definition, adjacent to each other, they must also be significantly correlated to each other and thus their correlation with biomarkers is of similar magnitude.

Assays
Determinations of plasma concentrations of glucose by the glucose oxidase method (Analox Instruments, Lunenburg MA), insulin by radioimmunoassay (RIA) (Millipore, St. Charles MO, serum amyloid A (SAA) protein by Elisa (Frederick MD), high-sensitivity C-reactive protein (hs-CRP) by Roche Integra 400 plus (Roche, Indianapolis IN), bile acids by liquid chromatography-mass spectrometry, TMAO, betaine and choline by liquid chromatography/ mass spectrometry in both positive and negative mode, SCFA by electron ionization for gas chromatography-mass spectrometry without derivatization, lipid panel, liver function tests and white cell counts were done by the Columbia University Diabetes Research Center Translational Bioanalytical Core, the Columbia University Irving biomarker core, and the Michigan Metabolomic Core, all certified Core

Laboratories. The homeostatic model for assessment of insulin resistance (HOMA-IR) was calculated[14].

References

[1] Caporaso JG, Lauber CL, Walters WA, et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. The ISME Journal. 2012;6:1621-4.
[2] Clemente JC, Pehrsson EC, Blaser MJ, et al. The microbiome of uncontacted Amerindians. Science Advances. 2015;1:e1500183.
[3] Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinformatics. 2010;26:2460-1.
[4] McDonald D, Price MN, Goodrich J, et al. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. The Isme Journal. 2011;6:610-18.
[5] Faith DP. Conservation evaluation and phylogenetic diversity. Biological Conservation. 1992;61:1-10.
[6] Lozupone C, Knight R. UniFrac: a new phylogenetic method for comparing microbial communities. Applied and environmental microbiology. 2005;71:8228-35.
[7] Knights D, Parfrey LW, Zaneveld J, Lozupone C, Knight R. Human-associated microbial signatures: examining their predictive value. Cell host & microbe. 2011;10:292-6.
[8] Segata N, Izard J, Waldron L, et al. Metagenomic biomarker discovery and explanation. Genome Biology. 2011;12:R60.
[9] Langille MGI, Zaneveld J, Caporaso JG, et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nature biotechnology. 2013;31:814-21.
[10] Parks DH, Tyson GW, Hugenholtz P, Beiko RG. STAMP: statistical analysis of taxonomic and functional profiles. Bioinformatics. 2014;30:3123-4.
[11] Friedman J, Alm EJ. Inferring Correlation Networks from Genomic Survey Data. PLoS Computational Biology. 2012;8:e1002687.
[12] Shannon P, Markiel A, Ozier O, et al. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. Genome Research. 2003;13:2498-504.
[13] Bron C, Kerbosch J. Finding All Cliques of an Undirected Graph. Commun Acm. 1973;16:575-7.
[14] Matthews DR, Hosker JP, Rudenski AS, et al. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. Diabetologia. 1985;28:412-9.

Supplemental Table 1: Comparison of Barcelona (BCN) and the New York City (NYC) cohorts at pre-surgery

	BCN	NYC	P
N	12	14	
Age (years)	44.5 (12.2)	39.5 (11.2)	0.258
Gender (% women)	83.3	85.7	
Weight (kg)	118.0 (18.6)	127.4 (14.8)	0.237
BMI (kg/m^2)	45.6 (8.5)	46.5 (3.9)	0.181
Systolic Blood Pressure (mmHg)	128.6 (12.3)	122.4 (12.1)	0.189
Diastolic Blood Pressure (mmHg)	71.4 (17.9)	72.1 (18.6)	0.553
Fasting Glucose (mmol/L)	6.7 (3.6)	6.5 (1.8)	0.237
Fasting Insulin (pmol/L)	190.3 (108.3)	193.1 (76.4)	0.877
HOMA-IR	8.0 (6.0)	8.1 (3.8)	0.328
HbA1c (%)	6.4 (1.2)	6.0 (0.6)	0.486
T-CHOL (mmol/L)	4.8 (1.1)	4.9 (1.1)	0.719
HDL (mmol/L)	1.2 (0.3)	1.3 (0.4)	0.341
LDL (mmol/L)	3.0 (1.0)	2.8 (0.8)	0.703
Triglycerides (mmol/L)	1.7 (1.0)	1.6 (1.2)	0.554
WBC (K/mcL)	8.5 (2.5)	7.8 (2.1)	0.681
SAA (ng/mL)	67.7 (66.9)	50.8 (35.4)	0.797
hs-CRP (mg/L)	7.3 (4.6)	11.4 (9.0)	0.217

Data are presented as mean (SD). SAA: serum amyloid A; hsCRP: ultrasensitive C reactive protein. WBC: white cell count.

<u>Supplemental Table 2:</u> Comparison of the effect of surgical weight loss between the NYC and BCN cohorts.

	BCN	NYC	P
N	12	14	
Δ Weight (kg)	41.7 (12.9)	41.4 (10.4)	0.944
Δ Weight (%)	35.0 (7.5)	32.3 (6.0)	0.324
Δ BMI (kg/m^2)	16.0 (4.6)	15.3 (3.8)	0.627
Δ SBP (mmHg)	1.9 (21.6)	5.3 (10.1)	0.615
Δ DBP (mmHg)	4.3 (18.7)	4.3 (17.8)	0.252
Δ Fasting Glucose (mmol/L)	2.3 (3.7)	1.5 (1.9)	0.483
Δ Fasting Insulin (pmol/L)	101.4 (90.3)	130.6 (66.7)	0.356
Δ HOMA-IR	5.4 (5.2)	6.2 (3.2)	0.629
Δ HbA1c (%)	1.1 (1.2)	0.6 (0.5)	0.272
Δ T-CHOL (mmol/L)	0.1 (1.3)	0.6 (1.0)	0.362
Δ HDL-C (mmol/L)	0.4 (0.3)	0.2 (0.4)	0.311
Δ LDL-C (mmol/L)	0.2 (1.3)	0.4 (0.8)	0.668
Δ Triglycerides (mmol/L)	0.8 (0.9)	0.3 (0.6)	0.186
Δ WBC (K/mcL)	2.0 (1.1)	1.1 (2.2)	0.283
Δ SAA (ng/mL)	48.0 (60.0)	34.0 (30.1)	0.463
Δ hs-CRP (mg/L)	5.9 (3.7)	10.4 (8.9)	0.122

Delta (Δ) change of weight, systolic (SBP) and diastolic (DBP) blood pressure and circulating biomarkers of metabolism and inflammation. Data are presented as mean (SD). WBC: white blood count; SAA: serum amyloid A; hsCRP: ultrasensitive C reactive

<u>Supplemental Table 3</u>: Comparison of the two surgery groups (SG n=7 and RYGB n=5) in the BCN cohort.

	SG		RYGB		
	Pre-surgery	**12 months**	**Pre-surgery**	**12 months**	**P**
N	7	7	5	5	
Age (years)	50.2 (9.4)	N/A	36.6 (11.9)	N/A	0.069
Weight (kg)	113.7 (20.4)	76.3 (16.4)***	123.8 (15.8)	76.3 (5.8) **	0.787
Weight loss (kg)	N/A	37.5 (10.4)	N/A	47.6 (14.9)	0.234
Weight loss (%)	N/A	33.0 (6.9)	N/A	37.7 (8.1)	0.326
Excess weight loss (%)		78.34 (22.88)		75.33 (15.91)	0.707
BMI (kg/m^2)	46.4 (10.3)	31.0 (7.4) ***	44.4 (5.8)	27.6 (4.2)***	0.538
SBP (mmHg)	127.3 (14.5)	131.0 (20.2)	130.4 (9.6)	120.6 (9.2)	0.575
DBP (mmHg)	75.4 (13.9)	79.0 (9.1)	65.8 (23.0)	71.0 (9.2)	0.442
Fasting Glucose (mmol/L)	7.0 (4.5)	4.3 (0.5) *	6.4 (2.2)	4.6 (0.3)	0.406
Fasting Insulin (µU/mL)	25.0 (15.5)	12.6 (7.2)*	30.6 (17.0)	12.9 (5.9)*	0.214
HOMA-IR	6.6 (4.7)	2.5 (1.7)*	9.9 (7.7)	2.7 (1.4)	0.048
HbA1c (%)	6.5 (1.4)	5.4 (0.3) *	6.3 (1.1)	5.3 (0.2)	0.936
T-CHOL (mmol/L)	5.1 (1.0)	5.6 (0.9)	4.4 (1.3)	3.5 (1.1)	0.379
HDL-C (mmol/L)	1.1 (0.3)	1.6 (0.3) **	1.2 (0.2)	1.5 (0.3)	0.242
LDL-C (mmol/L)	3.3 (1.0)	3.5 (0.7)	2.5 (1.0)	1.7 (0.9)	0.610
Triglycerides (mmol/L)	2.0 (1.1)	1.1 (0.6)*	1.4 (0.9)	0.65 (0.2)	0.808
WBC (K/mcL)	9.2 (2.8)	6.3 (1.1) *	7.35 (1.5)	5.6 (1.3)**	0.262
SAA (ng/mL)	83.8 (80.1)	19.6(13.2)*	45.1 (37.6)	19.9 (24.0)	0.338
hs-CRP (mg/L)	9.1 (4.6)	1.9 (1.3)**	4.8 (3.7)	0.6 (0.6)*	0.812

Differences from pre- and post-surgery within in surgical group: * P < 0.05, ** P < 0.01, *** P < 0.001. Data are presented as mean (SD). Analyses performed via Wilcoxon signed-rank test or paired t-test. SBP: systolic blood pressure; DBP: diastolic blood pressure; SAA: serum amyloid A; hsCRP: ultrasensitive C reactive protein; RYGB: Roux-en-Y gastric bypass; SG: sleeve gastrectomy. P values are for differences between surgical groups of delta change from pre-surgery to 12 months. Analyses performed via Wilcoxon rank-sum test or independent sample t-test.

Supplemental Table 4: Spearman Correlations between individual circulating bile acids and stool bacteria

Months: month since surgery; Taxa: p=phylum, c=class, o=order, f=family, g=genus

Bile acids	Corr coef	P value	Taxa	Fdr	Bonferroni	Months
Conj12AlphaOH	-1	0	p__Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae	0	0	3
Conj12AlphaOH	-1	0	p__Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales	0	0	3
CA	0.991071429	1.44E-05	p__Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__Pantoea	0.005773946	0.005773946	3
Conj12AlphaOH	-0.964285714	0.000454149	p__Proteobacteria	0.045528454	0.182113817	3
Conj12AlphaOH	-0.964285714	0.000454149	p__Proteobacteria;c__Gammaproteobacteria	0.045528454	0.182113817	3
CA	0.964285714	0.000454149	p__Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__Klebsiella	0.045528454	0.182113817	3
CA	-0.964285714	0.000454149	p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__[Ruminococcus]	0.045528454	0.182113817	3
CA	0.964285714	0.000454149	p__Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__Erwinia	0.045528454	0.182113817	3
GCDCA	0.964285714	0.000454149	p__Proteobacteria;c__Gammaproteobacteria;o__Pasteurellales	0.060704606	0.182113817	3
GCDCA	0.964285714	0.000454149	p__Proteobacteria;c__Gammaproteobacteria;o__Pasteurellales;f__Pasteurellaceae	0.060704606	0.182113817	3
GCDCA	0.964285714	0.000454149	p__Proteobacteria;c__Gammaproteobacteria;o__Pasteurellales;f__Pasteurellaceae;g__Aggregatibacter	0.060704606	0.182113817	3
CA	0.955357143	0.000789528	p__Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__Citrobacter	0.063320125	0.316600624	3
Secondary	-0.964285714	0.000454149	p__Proteobacteria	0.091056908	0.182113817	3
Secondary	-0.964285714	0.000454149	p__Proteobacteria;c__Gammaproteobacteria	0.091056908	0.182113817	3
Secondary UnC	-0.928571429	0.002519472	p__Firmicutes;c__Clostridia;o__Clostridiales;f__Christensenellaceae	0.144329776	1	3
Secondary UnC	-0.928571429	0.002519472	p__Firmicutes;c__Clostridia;o__Clostridiales;f__Christensenellaceae;g__	0.144329776	1	3
Secondary UnC	0.964285714	0.000454149	p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Leuconostocaceae;g__Weissella	0.144329776	0.182113817	3
Secondary UnC	0.928571429	0.002519472	p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Alcaligenaceae;g__Sutterella	0.144329776	1	3
Secondary UnC	0.928571429	0.002519472	p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Alcaligenaceae	0.144329776	1	3
Secondary UnC	0.928571429	0.002519472	p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae; g__Prevotella	0.144329776	1	3
Secondary UnC	0.928571429	0.002519472	p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae	0.144329776	1	3
TDCA	0.964285714	0.000454149	p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Porphyromonadaceae	0.182113817	0.182113817	3
Total UC	0.964285714	0.000454149	p__Proteobacteria;c__Deltaproteobacteria;o__Desulfovibrionales;f__Desulfovibrionaceae;g__Bilophila	0.182113817	0.182113817	3
Total C	-0.964285714	0.000454149	p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Roseburia	0.182113817	0.182113817	3
TDCA	-0.964285714	0.000454149	p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__	0.182113817	0.182113817	3
CDCA	0.964285714	0.000454149	p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Faecalibacterium	0.091056908	0.182113817	6
CDCA	0.964285714	0.000454149	p__Actinobacteria	0.091056908	0.182113817	6
GCDCA	0.964285714	0.000454149	p__Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Acidaminococcus	0.182113817	0.182113817	6
Secondary	-0.964285714	0.000454149	p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Ruminococcus	0.182113817	0.182113817	6
Conjugated	-1	0	p__Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__Clostridium	0	0	12
12-Alpha-OH	-1	0	p__Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__Clostridium	0	0	12
Primary UC	1	0	p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae	0	0	12
Non12AlphaOH	-0.973214286	0.000222306	p__Proteobacteria;c__Gammaproteobacteria;o__Aeromonadales;f__Aeromonadaceae;g__	0.044572372	0.089144743	12
TCA	-0.973214286	0.000222306	p__Proteobacteria;c__Gammaproteobacteria;o__Aeromonadales;f__Aeromonadaceae;g__	0.044572372	0.089144743	12

Bile acids	Corr coef	P value	Taxa	Fdr	Bonferroni	Months
Non12AlphaOH	-0.973214286	0.000222306	p__Proteobacteria;c__Gammaproteobacteria;o__Aeromonadales; f__Aeromonadaceae	0.044572372	0.089144743	12
TCA	-0.973214286	0.000222306	p__Proteobacteria;c__Gammaproteobacteria;o__Aeromonadales; f__Aeromonadaceae	0.044572372	0.089144743	12
TCA	-0.964285714	0.000454149	p__Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales; f__Enterobacteriaceae;g__Klebsiella	0.045528454	0.182113817	12
TCA	-0.964285714	0.000454149	p__Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales; f__Enterobacteriaceae;g__Erwinia	0.045528454	0.182113817	12
CA	0.964285714	0.000454149	p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae	0.182113817	0.182113817	12
Total UC	0.964285714	0.000454149	p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae	0.182113817	0.182113817	12
CDCA	0.964285714	0.000454149	p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae	0.182113817	0.182113817	12
SecondConj	-0.964285714	0.000454149	p__Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae; g__Clostridium	0.182113817	0.182113817	12
GDCA	-0.964285714	0.000454149	p__Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae; g__Dialister	0.182113817	0.182113817	12
GCDCA	-0.964285714	0.000454149	p__Actinobacteria;c__Actinobacteria;o__Actinomycetales	0.182113817	0.182113817	12
All Conj	-0.964285714	0.000454149	p__Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae; g__Clostridium	0.182113817	0.182113817	12
All BA	-0.964285714	0.000454149	p__Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae; g__Clostridium	0.182113817	0.182113817	12
GDCA	-0.964285714	0.000454149	p__Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae; g__Clostridium	0.182113817	0.182113817	12

Supplemental Table 5: Bile acids definitions

Unconjugated Bile Acids (Primary is red, Secondary is Blue)

DCA	deoxycholic	Hydroxyl group removed from CA by gut bacteria then becomes DC
CDCA	chenodeoxycholic	
UDCA	ursodeoxycholic	Epimer of CDC
CA	cholic acid	
LCA	lithocholic	Hydroxyl group removed from CDC by gut bacteria then becomes LC

Glycine Conjugated Bile Acids (aka Bile Salts) (Primary is green, Secondary is violet)

GDCA	glyco-deoxycholic	
GCDCA	glyco-chenodeoxycholic	
GUDCA	glyco-ursodeoxycholic	
GCA	glyco-cholic acid	
GLCA	glyco-lithocholic	

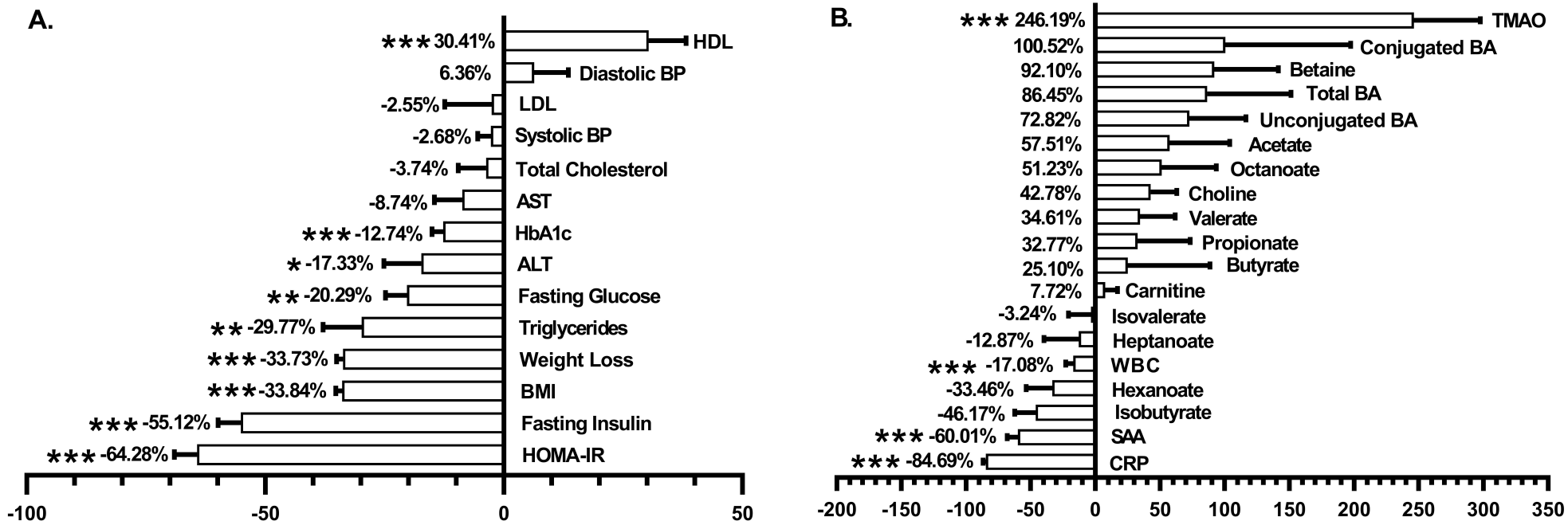
Taurine Conjugated Bile Acids (aka Bile Salts) (Primary is green, Secondary is violet)

TDCA	tauro-deoxycholic	
TCDCA	tauro-chenodeoxycholic	
TUDCA	tauro-ursodeoxycholic	
TCA	tauro-cholic acid	
TLCA	tauro-lithocholic	

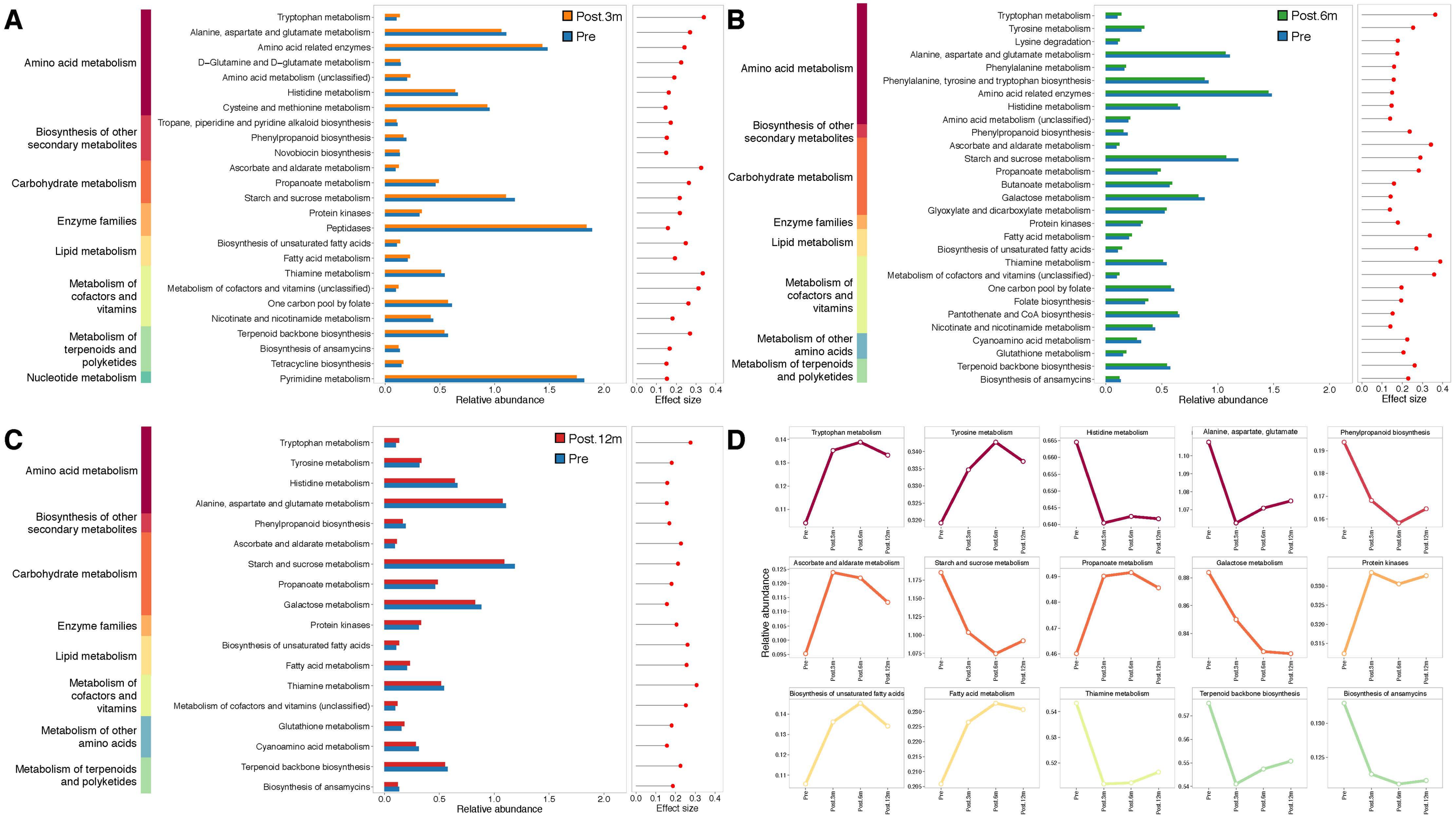
Key

	Primary Bile Acids	Primary bile acids (CDC and CA)
	Primary Bile Salts	Primary bile salts (GCDC, TCDC, GCA, TCA)
	Secondary Bile Acids	Secondary bile acids (DC, UDC, LC)
	Secondary Bile Salts	Secondary bile salts (GDC, TDC, GUDC, TUDC, GLC, TLC)

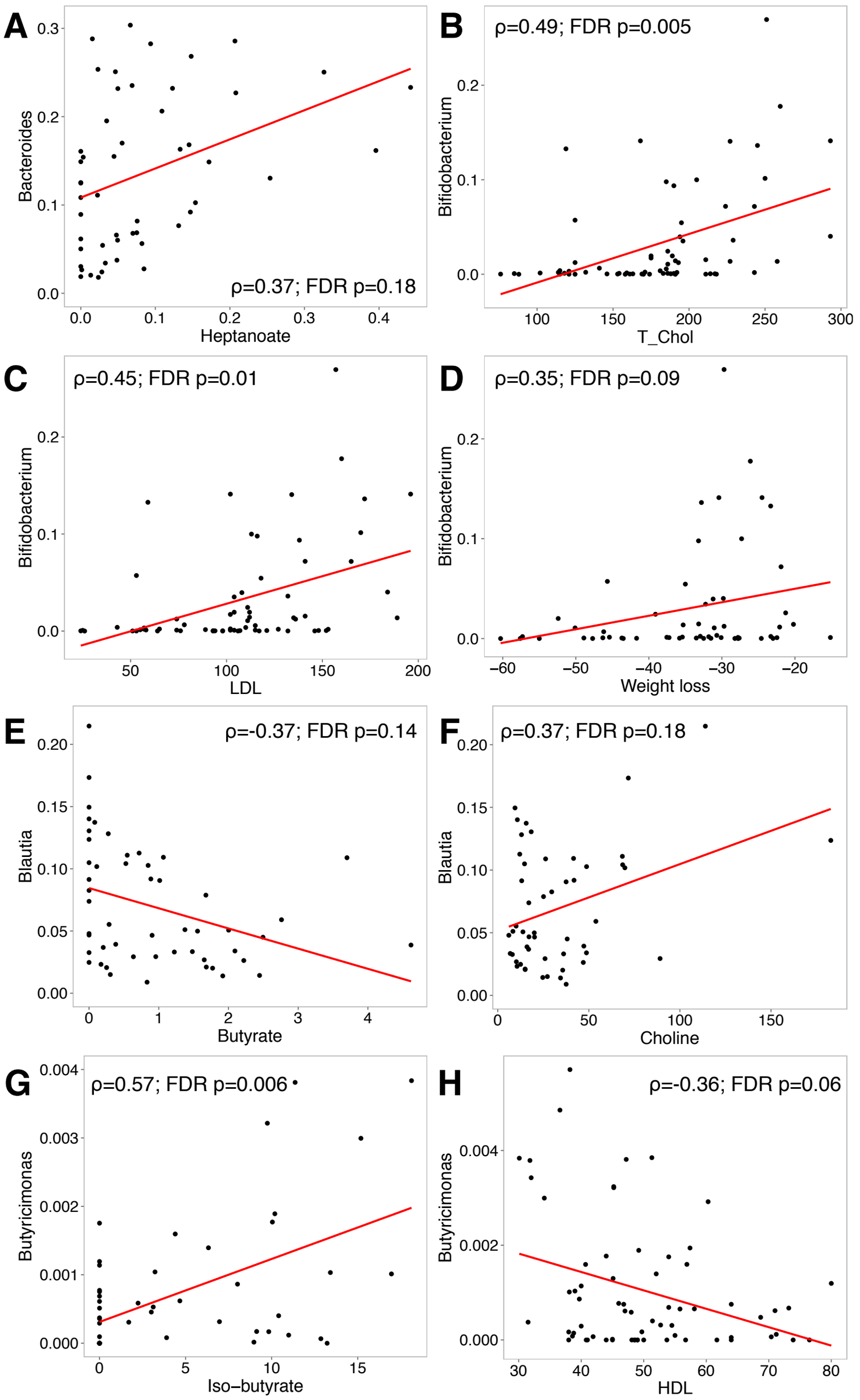
Supplemental Figure 1: Magnitude of change (express as percent increase and percent decrease) of clinical characteristics and metabolic biomarkers (A) and circulating biomarkers of inflammation and microbial metabolism (B), for the entire cohort (n=26), one year after surgery. *p<0.05, **p<0.01, ***p<0.001.


A.
*** 30.41%
HDL
6.36%
Diastolic BP
-2.55%
LDL
-2.68%
Systolic BP
-3.74%
Total Cholesterol
-8.74%
AST
*** -12.74%
HbA1c
* -17.33%
ALT
** -20.29%
Fasting Glucose
** -29.77%
Triglycerides
*** -33.73%
Weight Loss
*** -33.84%
BMI
*** -55.12%
Fasting Insulin
*** -64.28%
HOMA-IR
-100
-50
0
50
B.
*** 246.19%
TMAO
100.52%
Conjugated BA
92.10%
Betaine
86.45%
Total BA
72.82%
Unconjugated BA
57.51%
Acetate
51.23%
Octanoate
42.78%
Choline
34.61%
Valerate
32.77%
Propionate
25.10%
Butyrate
7.72%
Carnitine
-3.24%
Isovalerate
-12.87%
Heptanoate
*** -17.08%
WBC
-33.46%
Hexanoate
-46.17%
Isobutyrate
*** -60.01%
SAA
*** -84.69%
CRP
-200
-150
-100
-50
0
50
100
150
200
250
300
350

Supplemental Figure 2: Pathway Analysis
Relative abundance of all microbial metabolism pathways significantly (false discovery rate (FDR) $p < 0.05$) enriched between pre-surgery and 3m (A), pre-surgery and 6m (B), and pre-surgery and 12m (C). Effect size of differences between pre- and post-surgery, estimated as η^2. Change overtime of the relative abundance of pathways differentially enriched pre- and post-surgery (D).

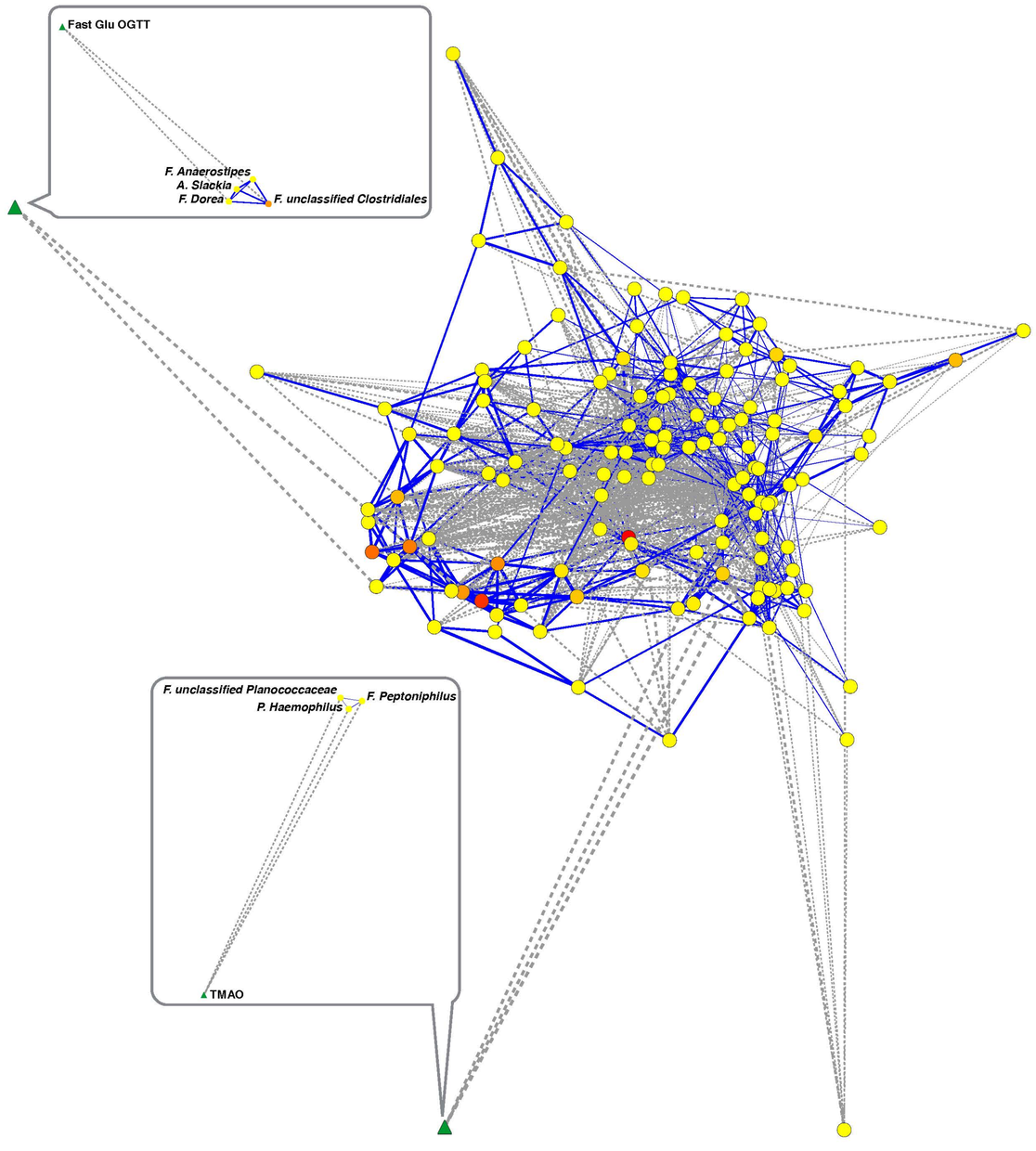

A
Amino acid metabolism
Biosynthesis of other secondary metabolites
Carbohydrate metabolism
Enzyme families
Lipid metabolism
Metabolism of cofactors and vitamins
Metabolism of terpenoids and polyketides
Nucleotide metabolism
Tryptophan metabolism
Alanine, aspartate and glutamate metabolism
Amino acid related enzymes
D−Glutamine and D−glutamate metabolism
Amino acid metabolism (unclassified)
Histidine metabolism
Cysteine and methionine metabolism
Tropane, piperidine and pyridine alkaloid biosynthesis
Phenylpropanoid biosynthesis
Novobiocin biosynthesis
Ascorbate and aldarate metabolism
Propanoate metabolism
Starch and sucrose metabolism
Protein kinases
Peptidases
Biosynthesis of unsaturated fatty acids
Fatty acid metabolism
Thiamine metabolism
Metabolism of cofactors and vitamins (unclassified)
One carbon pool by folate
Nicotinate and nicotinamide metabolism
Terpenoid backbone biosynthesis
Biosynthesis of ansamycins
Tetracycline biosynthesis
Pyrimidine metabolism
Post.3m
Pre
0.0 0.5 1.0 1.5 2.0
Relative abundance
0.0 0.1 0.2 0.3 0.4
Effect size
B
Amino acid metabolism
Biosynthesis of other secondary metabolites
Carbohydrate metabolism
Enzyme families
Lipid metabolism
Metabolism of cofactors and vitamins
Metabolism of other amino acids
Metabolism of terpenoids and polyketides
Tryptophan metabolism
Tyrosine metabolism
Lysine degradation
Alanine, aspartate and glutamate metabolism
Phenylalanine metabolism
Phenylalanine, tyrosine and tryptophan biosynthesis
Amino acid related enzymes
Histidine metabolism
Amino acid metabolism (unclassified)
Phenylpropanoid biosynthesis
Ascorbate and aldarate metabolism
Starch and sucrose metabolism
Propanoate metabolism
Butanoate metabolism
Galactose metabolism
Glyoxylate and dicarboxylate metabolism
Protein kinases
Fatty acid metabolism
Biosynthesis of unsaturated fatty acids
Thiamine metabolism
Metabolism of cofactors and vitamins (unclassified)
One carbon pool by folate
Folate biosynthesis
Pantothenate and CoA biosynthesis
Nicotinate and nicotinamide metabolism
Cyanoamino acid metabolism
Glutathione metabolism
Terpenoid backbone biosynthesis
Biosynthesis of ansamycins
Post.6m
Pre
0.0 0.5 1.0 1.5 2.0
Relative abundance
0.0 0.1 0.2 0.3 0.4
Effect size
C
Amino acid metabolism
Biosynthesis of other secondary metabolites
Carbohydrate metabolism
Enzyme families
Lipid metabolism
Metabolism of cofactors and vitamins
Metabolism of other amino acids
Metabolism of terpenoids and polyketides
Tryptophan metabolism
Tyrosine metabolism
Histidine metabolism
Alanine, aspartate and glutamate metabolism
Phenylpropanoid biosynthesis
Ascorbate and aldarate metabolism
Starch and sucrose metabolism
Propanoate metabolism
Galactose metabolism
Protein kinases
Biosynthesis of unsaturated fatty acids
Fatty acid metabolism
Thiamine metabolism
Metabolism of cofactors and vitamins (unclassified)
Glutathione metabolism
Cyanoamino acid metabolism
Terpenoid backbone biosynthesis
Biosynthesis of ansamycins
Post.12m
Pre
0.0 0.5 1.0 1.5 2.0
Relative abundance
0.0 0.1 0.2 0.3 0.4
Effect size
D
Relative abundance
Tryptophan metabolism
Tyrosine metabolism
Histidine metabolism
Alanine, aspartate, glutamate
Phenylpropanoid biosynthesis
Ascorbate and aldarate metabolism
Starch and sucrose metabolism
Propanoate metabolism
Galactose metabolism
Protein kinases
Biosynthesis of unsaturated fatty acids
Fatty acid metabolism
Thiamine metabolism
Terpenoid backbone biosynthesis
Biosynthesis of ansamycins
Pre
Post.3m
Post.6m
Post.12m

Supplemental Figure 3: Correlations Analyses
Partial correlations between bacterial taxa and biomarkers significant after false discovery rate (FDR) correction.
Bacteroides with heptanoate (A); Bifidobacterium with total cholesterol (B), LDL cholesterol (C) and weight loss (D); Bautia with butyrate (E) and choline (F); Butyricimonas with isobutyrate (G), and HDL (H).

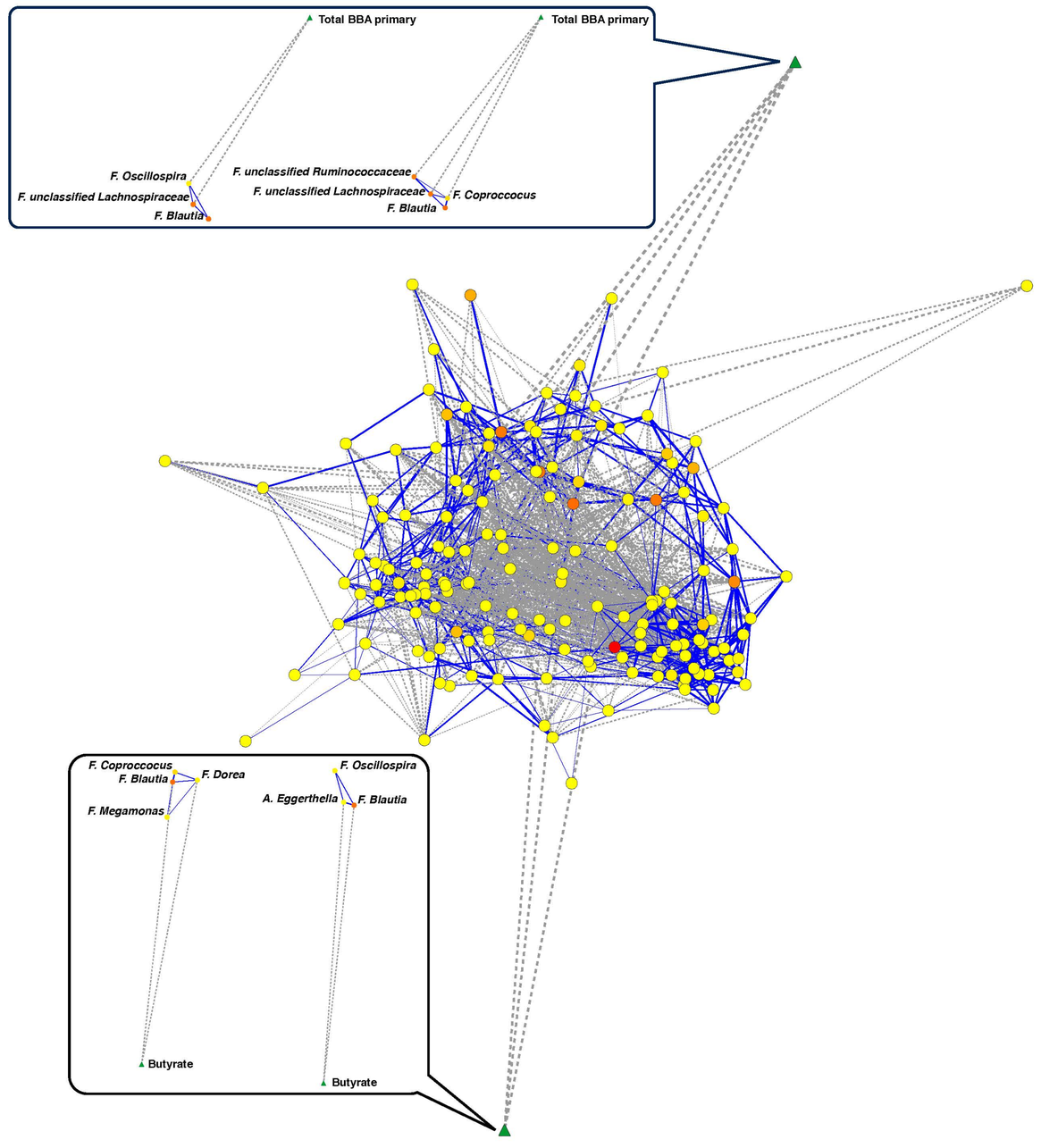

A
Bacteroides
Heptanoate
ρ=0.37; FDR p=0.18
B
ρ=0.49; FDR p=0.005
Bifidobacterium
T_Chol
C
ρ=0.45; FDR p=0.01
Bifidobacterium
LDL
D
ρ=0.35; FDR p=0.09
Bifidobacterium
Weight loss
E
ρ=-0.37; FDR p=0.14
Blautia
Butyrate
F
ρ=0.37; FDR p=0.18
Blautia
Choline
G
ρ=0.57; FDR p=0.006
Butyricimonas
Iso−butyrate
H
ρ=-0.36; FDR p=0.06
Butyricimonas
HDL

Supplemental Figure 4: Network Analyses
Network analysis of co-occurring bacteria at pre-surgery (A), 3m (B), 6m (C), 12m (D) post-surgery.
Circles represent bacterial taxa at the genus level, with color indicating relative abundance from red (high) to yellow (low). Edges between nodes indicate significant correlations (blue: positive; dashed gray: negative), with length of the edge representing strength of the correlation (shorter edges indicating higher correlation). Triangles represent biomarkers that are significantly correlated with bacterial clusters, magnified in the sub-panels.

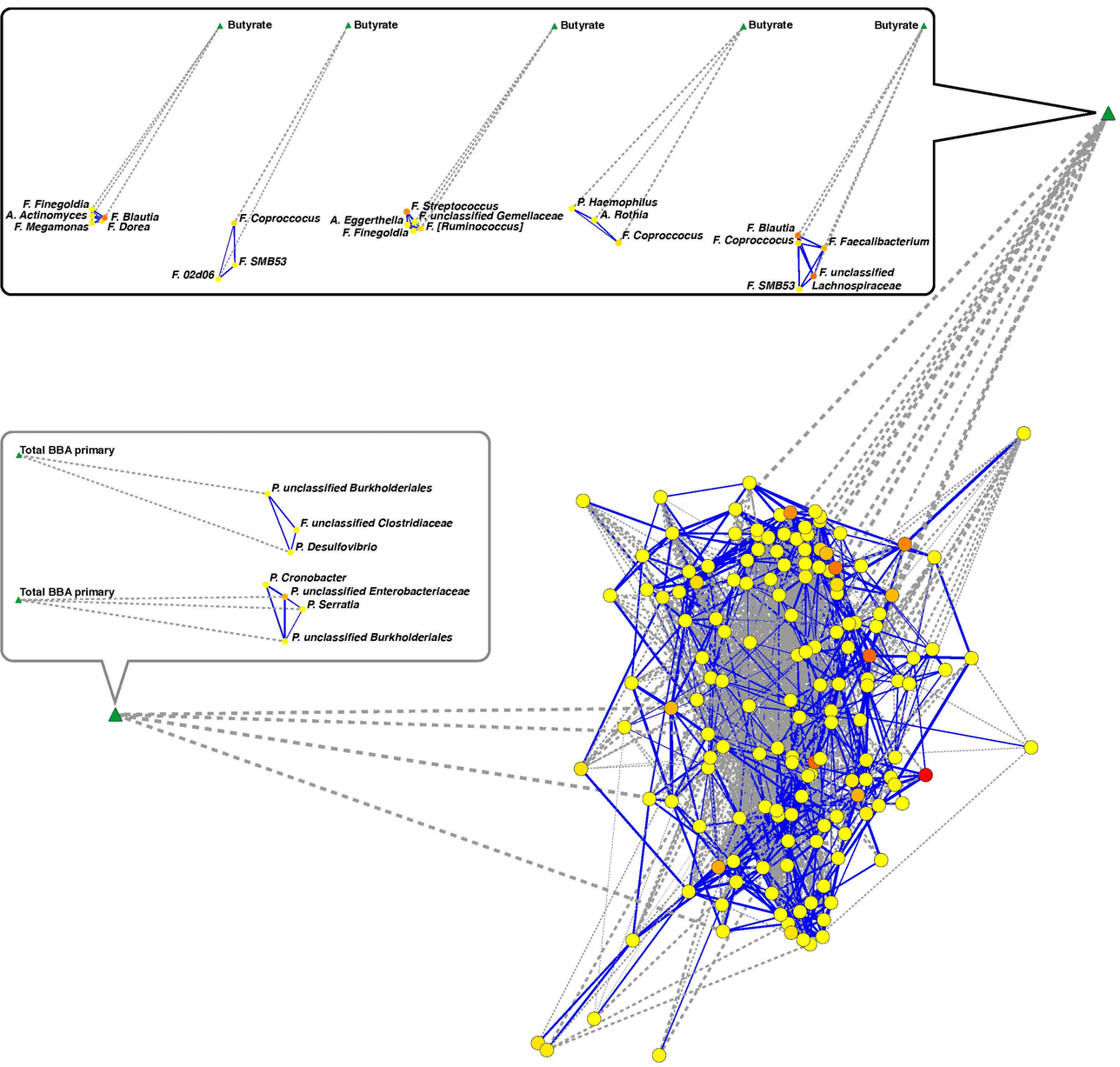
A


Fast Glu OGTT
F. Anaerostipes
A. Slackia
F. Dorea
F. unclassified Clostridiales
F. unclassified Planococcaceae
P. Haemophilus
P. Peptoniphilus
TMAO

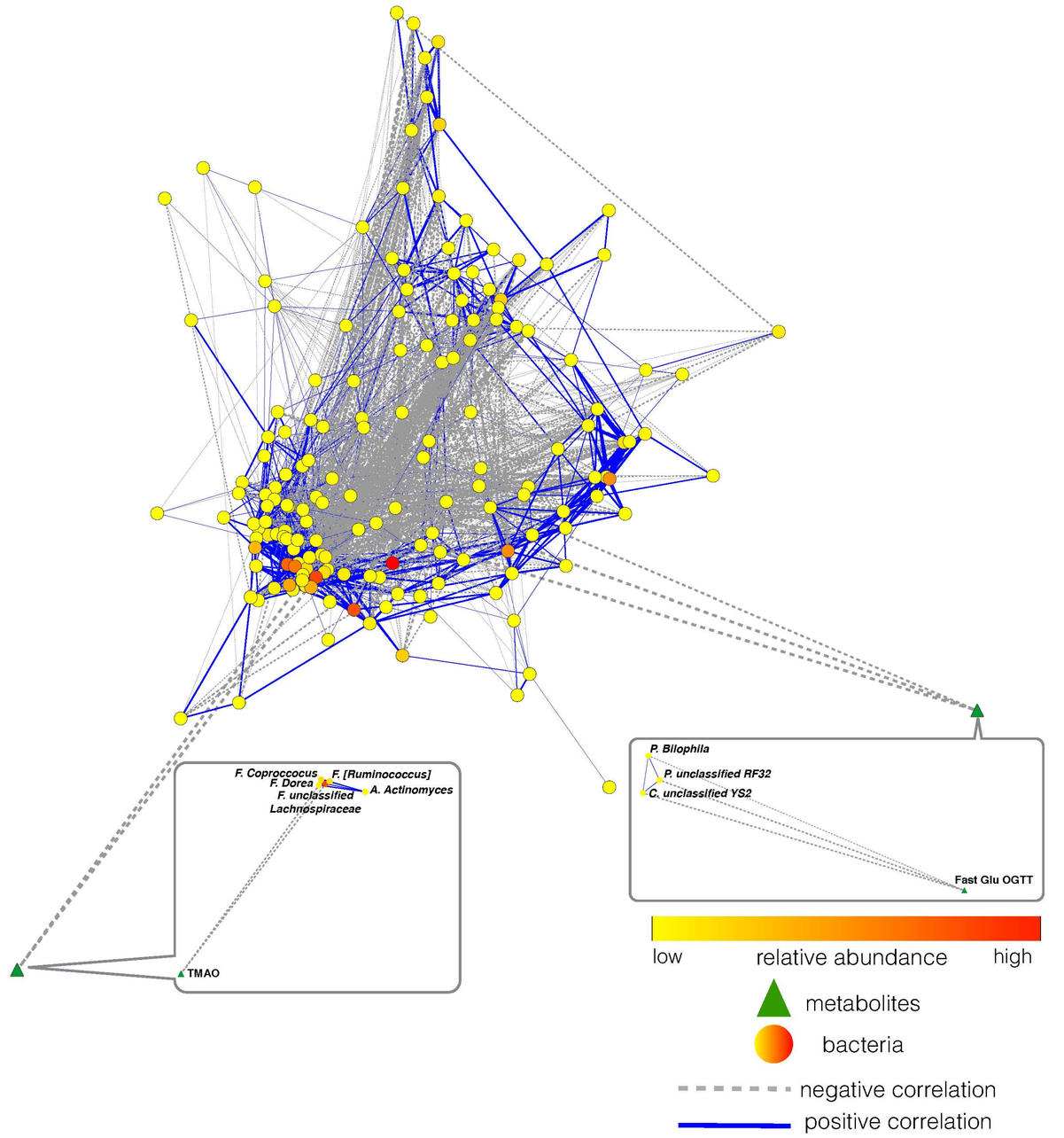
B


Total BBA primary
Total BBA primary
F. Oscillospira
F. unclassified Lachnospiraceae
F. Blautia
F. unclassified Ruminococcaceae
F. unclassified Lachnospiraceae
F. Coproccocus
F. Blautia
F. Coproccocus
F. Blautia
F. Dorea
F. Megamonas
F. Oscillospira
A. Eggerthella
F. Blautia
Butyrate
Butyrate

C


Butyrate
Butyrate
Butyrate
Butyrate
Butyrate
F. Finegoldia
A. Actinomyces
F. Megamonas
F. Blautia
F. Dorea
F. Coproccocus
F. 02d06
F. SMB53
A. Eggerthella
F. Finegoldia
F. Streptococcus
F. unclassified Gemellaceae
F. [Ruminococcus]
P. Haemophilus
A. Rothia
F. Coproccocus
F. Blautia
F. Coproccocus
F. Faecalibacterium
F. SMB53
F. unclassified Lachnospiraceae
Total BBA primary
P. unclassified Burkholderiales
F. unclassified Clostridiaceae
P. Desulfovibrio
Total BBA primary
P. Cronobacter
P. unclassified Enterobacteriaceae
P. Serratia
P. unclassified Burkholderiales

D


F. Coproccocus
F. Dorea
F. unclassified Lachnospiraceae
F. [Ruminococcus]
A. Actinomyces
TMAO
P. Bilophila
P. unclassified RF32
C. unclassified YS2
Fast Glu OGTT
low
relative abundance
high
metabolites
bacteria
negative correlation
positive correlation