Supporting Information:

Autonomous Multi-Modal Metabolomics Data Integration for Comprehensive Pathway Analysis

Tao Huan[†], Amelia Palermo[†], Julijana Ivanisevic[○], Duane Rinehart[†], David Edler[§], Thiery Phommavongsay[†], H. Paul Benton[†], Carlos Guijas[†], Xavier Domingo-Almenara[†], Benedikt Warth[‖], Gary Siuzdak[†, ⊥, *]

[†] Scripps Center for Metabolomics, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, USA

[○] Metabolomics Platform, Faculty of Biology and Medicine, University of Lausanne, Switzerland

[§] Department of Molecular Medicine and Surgery, Karolinska Institute, 171 77 Stockholm, Sweden

[‖] Department of Food Chemistry and Toxicology, Faculty of Chemistry, University of Vienna, Währingerstraße 38, 1090 Vienna, Austria

[⊥] Departments of Chemistry, Molecular and Computational Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, USA

*Author to whom correspondence should be addressed:

Gary Siuzdak

Tel: (858) 784-9415

E-mail: siuzdak@scripps.edu

Internet: https://masspec.scripps.edu/

Table of Contents

Text S-1. Tissue Sample Preparation 3

Text S-2. LC-MS Data Collection 3

1. RP(+) metabolomics analysis 3
2. RP(-) metabolomics analysis 4
3. HILIC(+) metabolomics analysis 4
4. HILIC(-) metabolomics analysis 4

Text S-3. Autonomous MS/MS Data Collection 5

Text S-4. LC-MS Data Processing 6

1. RP (+) metabolomics data processing 6
2. RP(-) metabolomics data processing 6
3. HILIC(+) metabolomics data processing 6
4. HILIC(-) metabolomics data processing 6

Text S-6. Metabolite Identification 7

Table S-1. Detailed clinical features of the colon cancer samples 8

Table S-2. Parameters for multi-modal analysis in the colon cancer study 9

Table S-3. Predicted pathway dysregulations in each analytical mode 10

Table S-5. Predicted pathway dysregulations in the multi-modal XCMS analysis 14

Table S-6. Confirmed metabolites in the colon cancer multi-modal study 18

Table S-7. Top 10 predicted pathway dysregulation ranked by the number of dysregulated genes 19

Text S-1. Tissue Sample Preparation.

For each sample, 10 mg frozen colon tissue was cut and placed in a LC glass vial. Glass beads (5 g) and 600 µL ice cold methanol/water (4:1, v/v) was then added to the same glass vial. Tissue samples were homogenized for 30 seconds for twice at medium speed. The homogenized sample was sonicated in ice bath for 10 min and then kept in -20 ºC freezer overnight for complete protein precipitation. After that, 550 µL of the homogenized solution was transferred to a 1.5 mL eppendorf vial. Another 200 µL ice cold methanol/water (4:1, v/v) was added to rinse off any remaining solutions in glass vials and then combine the wash solution to the same eppendorf vial. The solutions were centrifuged at 13,000 rpm for 15 min and the supernatant was transferred to new glass vials. Supernatant was dried down using Speedvac till complete dryness and then resuspended in H_2O/ACN (50/50, v/v) at a volume of 6.5 µL/mg tissue. The reconstituted solution was then transferred to another 1.5 mL eppendorf vial, vortexed and spin down at 13,000 rpm for 15 min. Supernatant was pipetted into LC insert in glass HPLC vials for LC-MS analysis. A pooled sample was prepared by aliquoting 10 µL metabolite solution from each sample. This pooled sample was used as a quality control (QC) to monitor the instrumental stability throughout the LC-MS analysis of all the colon cancer samples.

Text S-2. LC-MS Data Collection

1. RP(+) metabolomics analysis

For RPL(+) metabolomics analysis, an Agilent ZORBAX 300SB-C18 LC column (300 Å, 5 µm, 150 × 0.5 mm) was used. Mobile phase A was 0.1% formic acid in water and mobile phase B was 0.1% FA in ACN. The LC-gradient was: t = 0.0 min, 95% A; t = 5 min, 95% A; t = 50 min, 5% A; t = 60 min; 5% A; t = 61 min, 95% A; t = 64 min, 95% A. At the end of the LC gradient, a 10-min re-equilibration time at 95% A was applied. The LC flow rate was 20 µL/min. The sample injection volume was 8 µL.

MS conditions for RP(+) analysis were set as follows: capillary voltage, 4500; nebulizer gas flow, 1.6 Bar; dry gas, 6.0 L/min at 220 ºC; funnel 1 RF 150 Vpp; funnel 2 RF, 200 Vpp; isCID energy, 0 eV; hexapole RF: 50 Vpp; Quadrupole ion energy, 4 eV; low mass 50 m/z; collision cell energy, 7.0 eV; pre pulse storage 5.0 µs; collision RF, ramp from 350 to 800 Vpp; transfer time ramp from 50 to 100 µs; detection mass range 25 to 1000 m/z; spectra collection rate 2.0 Hz.

2. RP(-) metabolomics analysis

For RP(-) metabolomics analysis, the same Agilent ZORBAX 300SB-C18 LC column (300 Å, 5 µm, 150 × 0.5 mm) was used. Mobile phase A was 20 mM NH_4AC in H_2O (pH 9.7) with 5% ACN and mobile phase B was ACN with 5% H_2O. The LC gradient was: t = 0 min, 95% A; t = 5 min, 95% A; t = 10 min, 25% A; t = 29 min, 5% A; t = 30 min, 95% A; t = 32 min, 95% A. At the end of the LC gradient, a 10-min re-equilibration time at 95% A was applied. The flow rate was 20 µL/min. The sample injection volume was 2 µL.

MS conditions for lipidomic analysis were set as follows: capillary voltage, 4000; nebulizer gas flow, 1.4 Bar; dry gas, 7.0 L/min at 180 ºC; funnel 1 RF 150 Vpp; funnel 2 RF, 200 Vpp; isCID energy, 0 eV; hexapole RF: 50 Vpp; Quadrupole ion energy, 4 eV; low mass 50 m/z; collision cell energy, 8.0 eV; pre pulse storage 5.0 µs; collision RF, ramp from 350 to 850 Vpp; transfer time, ramp from 50 to 100 µs; detection mass range 25 to 1000 m/z; spectra collection rate 2.0 Hz.

3. HILIC(+) metabolomics analysis

For HILIC(+) metabolomics analysis, , a Phenomenex Luna NH2 LC column (100 Å, 3 µm, 150 × 1 mm) was used. Mobile phase A was 0.1% formic acid in water and mobile phase B was 0.1% FA in ACN. The LC-gradient was: t = 0.0 min, 95% A; t = 5 min, 95% A; t = 50 min, 5% A; t = 60 min; 5% A; t = 61 min, 95% A; t = 64 min, 95% A. At the end of the LC gradient, a 10-min re-equilibration time at 95% A was applied. The LC flow rate was 20 µL/min. The sample injection volume was 8 µL.

MS conditions for HILIC(+) analysis were set as follows: capillary voltage, 4500; nebulizer gas flow, 1.4 Bar; dry gas, 7.0 L/min at 180 ºC; funnel 1 RF 150 Vpp; funnel 2 RF, 200 Vpp; isCID energy, 0 eV; hexapole RF: 50 Vpp; Quadrupole ion energy, 4 eV; low mass 50 m/z; collision cell energy, 7.0 eV; pre pulse storage 5.0 µs; collision RF, ramp from 350 to 850 Vpp; transfer time, ramp from 50 to 100 µs; detection mass range 25 to 1000 m/z; spectra collection rate 2.0 Hz.

4. HILIC(-) metabolomics analysis

For HILIC(-) metabolomics analysis, a Phenomenex Luna NH2 LC column (100 Å, 3 µm, 150 × 1 mm) was used. Mobile phase A was 20 mM NH_4AC in H_2O (pH 9.7) with 5% ACN and mobile phase B was ACN with 5% H_2O. The LC gradient was: t = 0.0 min, 5% A; t = 5 min, 5% A; t = 50 min, 95% A; t = 63 min, 95% A. At the end of the LC gradient, a 15-min reequilibration time at

5% A was applied to the HILIC column. The flow rate was 50 µL/min. The sample injection volume was 8 µL.

MS conditions for HILIC(-) were set as follows: capillary voltage, 4500; nebulizer gas flow, 1.6 Bar; dry gas, 6.0 L/min at 220 ºC; funnel 1 RF 150 Vpp; funnel 2 RF, 200 Vpp; isCID energy, 0 eV; hexapole RF: 50 Vpp; Quadrupole ion energy, 4 eV; low mass 50 m/z; collision cell energy, 7.0 eV; pre pulse storage 5.0 µs; collision RF, ramp from 350 to 800 Vpp; transfer time, ramp from 50 to 100 µs; detection mass range 25 to 1000 m/z; spectra rate 2.0 Hz.

Text S-3. Autonomous MS/MS Data Collection

Autonomous MS/MS data collection was performed using the pooled colon cancer sample in each of the analytical modes. Data dependent MS/MS spectra acquisition mode was selected, in which each MS1 survey scan was followed by 10 MS/MS scan of the most abundant metabolic peaks in the MS1 survey scan. An averaged collision energy, 20 – 50 eV was applied for generating MS/MS spectra. Smart exclusion was applied so that after a precursor ion appears in three continues MS1 survey scan with sufficient abundance to trigger MS/MS analysis, it would be excluded from MS/MS analysis for 0.2 min. MS1 scan was collected at 8 Hz. If the MS1 precursor intensity larger than 1e5, MS/MS spectra were collected at 12 Hz. On the other hand, if the MS1 precursor intensity was less than 1e5, MS/MS spectra were collected at 4 Hz.

Text S-4. LC-MS Data Processing

1. RP (+) metabolomics data processing

RP(+) metabolomic data processing parameter: centWave for feature detection (Δ m/z = 10 ppm, minimum peak width = 5 s, and maximum peak width = 60 s); obiwarp settings for retention time correction (profStep = 1); parameters for chromatogram alignment, including mzwid = 0.015, minfrac = 0.5, and bw = 5. The relative quantification of metabolite features was based on extracted ion chromatogram (EIC) areas. parametric t-test was used to find the metabolic difference in between normal and tumor colon tissue samples.

2. RP(-) metabolomics data processing

RP(-) metabolomic data processing parameter: centWave for feature detection (Δ m/z = 10 ppm, minimum peak width = 5 s, and maximum peak width = 60 s); obiwarp settings for retention time correction (profStep = 1); parameters for chromatogram alignment, including mzwid = 0.015, minfrac = 0.5, and bw = 5. The relative quantification of metabolite features was based on extracted ion chromatogram (EIC) areas. Paired parametric t-test was used to find the metabolic difference in between normal and tumor colon tissue samples.

3. HILIC(+) metabolomics data processing

HILIC(+) metabolomic data processing parameter: centWave for feature detection (Δ m/z = 10 ppm, minimum peak width = 10 s, and maximum peak width = 120 s); obiwarp settings for retention time correction (profStep = 1); parameters for chromatogram alignment, including mzwid = 0.015, minfrac = 0.5, and bw = 30. The relative quantification of metabolite features was based on extracted ion chromatogram (EIC) areas. Paired parametric t-test was used to find the metabolic difference in between normal and tumor colon tissue samples.

4. HILIC(-) metabolomics data processing

HILIC(-) metabolomic data processing parameter: centWave for feature detection (Δ m/z = 10 ppm, minimum peak width = 10 s, and maximum peak width = 120 s); obiwarp settings for retention time correction (profStep = 1); parameters for chromatogram alignment, including mzwid = 0.015, minfrac = 0.5, and bw = 30. The relative quantification of metabolite features was based on extracted ion chromatogram (EIC) areas. Paired parametric t-test was used to find the metabolic difference in between normal and tumor colon tissue samples.

Text S-6. Metabolite Identification

Metabolite identification was performed based on MS/MS spectra matching in METLIN to confirm the identities of the statistically significant metabolic features from all LC-MS analyses. In specific, experimental MS/MS spectra generated in autonomous MS/MS analysis were matched against MS/MS spectra of metabolite standards archive in METLIN mass spectra database. Mass tolerance for precursor ion was 10 ppm and mass tolerance for fragment ions was 30 ppm. METLIN mass spectra database is one of the most comprehensive metabolite spectra library, including over 16,000 metabolites individually analyzed and another 200,000 metabolites with *in silico* MS/MS data.

Table S-1. Detailed clinical features of the colon cancer samples

Code no	date of birth year,month,day	male / female	date of surgery	localization	T stage
1	1942 07 08	male	2015 08 24	cecum (right)	T3
2	1940 08 11	male	2015 07 15	ascendens (right)	T4b
3	1925 09 30	male	2014 11 20	cecum (right)	T3c
4	1949 04 07	female	2014 11 19	cecum (right)	T3b
5	1946 12 04	female	2014 10 22	ascendens (right)	T3a
6	1933 02 16	male	2014 10 22	ceecum (right)	T3a
7	1929 05 03	female	2015 06 03	ascendens(right)	T4b
8	1945 02 23	female	2015 07 07	cecum (right)	T4b
9	1942 12 07	female	2014 10 10	cecum (right)	T3b
10	1922 09 21	female	2014 10 02	cecum (right)	T4b

Table S-2. Parameters for multi-modal analysis in the colon cancer study

	Analysis	Polarity	Mass tolerance (ppm)	Intensity cutoff	P value cutoff
1	Metabolomic RP(+)	Positive	5	10000	0.001
2	Metabolomic RP(-)	Negative	5	10000	0.001
3	Metabolomic HILIC(+)	Positive	5	10000	0.005
4	Metabolomic HILIC(-)	Negative	5	10000	0.01

Table S-3. Predicted pathway dysregulations in each analytical mode

	RP(+) metabolomics		
	Pathway	**No. of dysregulated metabolic features**	**p-value**
1	bile acid biosynthesis, neutral pathway	6	3.3E-03
2	sucrose degradation	3	3.5E-03
3	4-hydroxy-2-nonenal detoxification	2	4.9E-03
4	lactose degradation III	2	4.9E-03
5	D-galactose degradation V (Leloir pathway)	2	1.3E-02
6	trehalose degradation	2	1.3E-02
7	adenine and adenosine salvage III	2	1.9E-02
8	leukotriene biosynthesis	3	1.9E-02
9	pregnenolone biosynthesis	2	1.9E-02
10	adenosine nucleotides degradation	2	2.8E-02
11	nicotine degradation IV	2	3.8E-02

	RP(-) metabolomics		
	Pathway	**No. of dysregulated metabolic features**	**p-value**
1	adenosine nucleotides degradation	4	8.4E-05
2	urate biosynthesis/inosine 5'-phosphate degradation	3	1.3E-04
3	UMP biosynthesis	3	1.3E-04
4	adenine and adenosine salvage III	3	1.5E-04
5	sucrose degradation	4	1.9E-04
6	lactate fermentation (reoxidation of cytosolic NADH)	2	2.1E-04
7	lactose degradation III	2	2.1E-04
8	purine ribonucleosides degradation to ribose-1-phosphate	3	2.3E-04
9	2-amino-3-carboxymuconate semialdehyde degradation to glutaryl-CoA	2	3.3E-04
10	methylglyoxal degradation I	2	5.0E-04

11	trehalose degradation	2	5.0E-04
12	dolichyl-diphosphooligosaccharide biosynthesis	2	7.4E-04
13	methylglyoxal degradation VI	2	7.4E-04
14	UTP and CTP <i>de novo</i> biosynthesis	2	7.4E-04
15	guanosine nucleotides degradation	2	1.1E-03
16	D-galactose degradation V (Leloir pathway)	2	1.1E-03
17	UTP and CTP dephosphorylation I	2	2.0E-03

	HILIC(+) metabolomics		
	Pathway	**No. of dysregulated metabolic features**	**p-value**
1	purine deoxyribonucleosides degradation	5	7.7E-05
2	purine ribonucleosides degradation to ribose-1-phosphate	6	8.5E-05
3	histidine degradation	3	1.5E-04
4	adenosine nucleotides degradation	4	2.2E-04
5	dopamine degradation	3	2.7E-04
6	adenine and adenosine salvage I	3	4.8E-04
7	guanosine ribonucleotides <i>de novo</i> biosynthesis	3	4.8E-04
8	sucrose degradation	3	8.7E-04
9	guanosine nucleotides degradation	3	8.7E-04
10	purine deoxyribonucleosides salvage	3	8.7E-04
11	<i>S</i>-methyl-5'-thioadenosine degradation	2	9.2E-04
12	diphthamide biosynthesis	2	9.2E-04
13	lactose degradation III	2	9.2E-04
14	tRNA splicing	2	9.2E-04
15	adenine and adenosine salvage II	2	9.2E-04
16	homocarnosine biosynthesis	2	9.2E-04
17	adenine and adenosine salvage III	3	1.5E-03
18	bile acid biosynthesis, neutral pathway	6	1.6E-03
19	icosapentaenoate biosynthesis II (metazoa)	2	2.3E-03
20	γ-linolenate biosynthesis	2	2.3E-03

21	L-serine degradation	2	2.3E-03
22	uracil degradation	2	2.3E-03
23	coenzyme A biosynthesis	2	2.3E-03
24	7-(3-amino-3-carboxypropyl)-wyosine biosynthesis	2	2.3E-03
25	lipoate biosynthesis and incorporation	2	2.3E-03
26	UDP-<i>N</i>-acetyl-D-galactosamine biosynthesis II	3	2.7E-03
27	tRNA charging	3	2.7E-03
28	CMP-<i>N</i>-acetylneuraminate biosynthesis I (eukaryotes)	3	2.7E-03
29	pyrimidine deoxyribonucleosides degradation	2	5.0E-03
30	pyrimidine ribonucleosides degradation	2	5.0E-03
31	(S)-reticuline biosynthesis	2	5.0E-03
32	leucine degradation	2	5.0E-03
33	stearate biosynthesis	2	5.0E-03
34	trehalose degradation	2	5.0E-03
35	4-hydroxybenzoate biosynthesis	2	9.5E-03
36	guanine and guanosine salvage	2	9.5E-03
37	adenosine ribonucleotides <i>de novo</i> biosynthesis	2	9.5E-03
38	tyrosine degradation	2	9.5E-03
39	lysine degradation II (pipecolate pathway)	3	1.6E-02
40	D-galactose degradation V (Leloir pathway)	2	1.6E-02
41	urate biosynthesis/inosine 5'-phosphate degradation	2	1.6E-02
42	nicotine degradation III	4	1.7E-02
43	TCA cycle	3	2.5E-02
44	noradrenaline and adrenaline degradation	2	3.7E-02
45	gluconeogenesis	2	3.7E-02

	HILIC(-) metabolomics		
	Pathway	**No. of dysregulated metabolic features**	**p-value**
1	adenosine nucleotides degradation	6	8.6E-04

2	purine ribonucleosides degradation to ribose-1-phosphate	6	9.4E-04
3	adenine and adenosine salvage III	5	1.0E-03
4	adenine and adenosine salvage I	4	1.2E-03
5	L-serine degradation	3	1.7E-03
6	pyrimidine deoxyribonucleosides degradation	4	1.7E-03
7	D-galactose degradation V (Leloir pathway)	4	1.7E-03
8	purine deoxyribonucleosides degradation	4	1.7E-03
9	UDP-<i>N</i>-acetyl-D-galactosamine biosynthesis II	4	2.5E-03
10	uracil degradation	3	3.0E-03
11	terminal <i>O</i>-glycans residues modification	3	3.0E-03
12	sucrose degradation	4	3.8E-03
13	valine degradation	3	5.4E-03
14	mucin core 1 and core 2 <i>O</i>-glycosylation	4	5.4E-03
15	CMP-<i>N</i>-acetylneuraminate biosynthesis I (eukaryotes)	3	5.4E-03
16	purine deoxyribonucleosides salvage	3	5.4E-03
17	PRPP biosynthesis	2	9.5E-03
18	lactose degradation III	2	9.5E-03
19	coenzyme A biosynthesis	2	9.5E-03
20	adenine and adenosine salvage II	2	9.5E-03
21	glycoaminoglycan-protein linkage region biosynthesis	2	9.5E-03
22	chondroitin biosynthesis	2	9.5E-03
23	D-glucuronate degradation	3	9.7E-03
24	methylglyoxal degradation VI	3	9.7E-03
25	guanosine nucleotides degradation	3	1.7E-02
26	urate biosynthesis/inosine 5'-phosphate degradation	3	1.7E-02
27	thymine degradation	3	2.3E-02
28	lactate fermentation (reoxidation of cytosolic NADH)	2	2.3E-02
29	dolichyl-diphosphooligosaccharide biosynthesis	2	2.3E-02
30	ketogenesis	2	2.3E-02
31	methylglyoxal degradation I	2	2.3E-02
32	citrulline-nitric oxide cycle	2	2.3E-02
33	urea cycle	2	2.3E-02

34	heparan sulfate biosynthesis (late stages)	2	2.3E-02
35	pyrimidine ribonucleosides degradation	2	4.6E-02
36	UDP-<i>N</i>-acetyl-D-glucosamine biosynthesis II	2	4.6E-02
37	ketolysis	2	4.6E-02
38	7-(3-amino-3-carboxypropyl)-wyosine biosynthesis	2	4.6E-02
39	ethanol degradation II	2	4.6E-02
40	ethanol degradation IV	2	4.6E-02
41	oxidative ethanol degradation III	2	4.6E-02
42	lipoate biosynthesis and incorporation	2	4.6E-02

	RP(+) lipidomics		
	Pathway	**No. of dysregulated metabolic features**	**p-value**
1	bile acid biosynthesis, neutral pathway	7	4e-3
2	pregnenolone biosynthesis	2	8e-2
3	thyroid hormone metabolism II (via conjugation and/or degradation)	2	2e-2

Table S-5. Predicted pathway dysregulations in the multi-modal XCMS analysis

	Pathway	**No. of dysregulated metabolic features**	**p-value**
1	tRNA charging	19	1.4E-03
2	phenylalanine degradation IV (mammalian, via side chain)	10	1.4E-03
3	arginine biosynthesis IV	8	1.4E-03
4	glycine betaine degradation	8	1.4E-03
5	citrulline biosynthesis	8	1.4E-03
6	lysine degradation II	7	1.4E-03
7	purine ribonucleosides degradation to ribose-1-phosphate	7	1.4E-03
8	proline biosynthesis I	5	1.4E-03

9	cysteine biosynthesis/homocysteine degradation	5	1.4E-03
10	asparagine biosynthesis I	5	1.4E-03
11	adenosine nucleotides degradation II	6	1.4E-03
12	proline biosynthesis II (from arginine)	6	1.4E-03
13	uridine-5'-phosphate biosynthesis	6	1.4E-03
14	uracil degradation II (reductive)	4	1.4E-03
15	thymine degradation	4	1.4E-03
16	salvage pathways of pyrimidine deoxyribonucleotides	5	1.4E-03
17	adenine and adenosine salvage III	5	1.4E-03
18	histidine degradation III	5	1.4E-03
19	arginine degradation VI (arginase 2 pathway)	5	1.4E-03
20	UDP-<i>N</i>-acetyl-D-glucosamine biosynthesis II	5	1.4E-03
21	adenine and adenosine salvage I	4	1.4E-03
22	xanthine and xanthosine salvage	4	1.4E-03
23	valine degradation I	6	1.4E-03
24	(S)-reticuline biosynthesis II	6	1.4E-03
25	5-aminoimidazole ribonucleotide biosynthesis I	5	1.4E-03
26	tyrosine degradation I	5	1.4E-03
27	urea cycle	6	1.4E-03
28	citrulline-nitric oxide cycle	4	1.4E-03
29	L-glutamine biosynthesis II (tRNA-dependent)	3	1.4E-03
30	salvage pathways of pyrimidine ribonucleotides	6	1.4E-03
31	histidine degradation VI	7	1.4E-03
32	4-hydroxybenzoate biosynthesis	5	1.4E-03
33	pyrimidine ribonucleotides interconversion	4	1.4E-03
34	leucine degradation I	4	1.4E-03
35	alanine biosynthesis II	3	1.4E-03
36	glutathione biosynthesis	3	1.4E-03
37	proline degradation	3	1.4E-03

38	alanine degradation III	3	1.4E-03
39	UDP-<i>N</i>-acetyl-D-galactosamine biosynthesis II	4	1.4E-03
40	CMP-<i>N</i>-acetylneuraminate biosynthesis I (eukaryotes)	5	1.4E-03
41	arginine degradation I (arginase pathway)	4	1.4E-03
42	sucrose degradation V (mammalian)	4	1.4E-03
43	molybdenum cofactor biosynthesis	4	1.5E-03
44	lipoate biosynthesis and incorporation II	3	1.5E-03
45	methionine salvage II (mammalia)	3	1.5E-03
46	glutamate degradation II	3	1.5E-03
47	threonine degradation II	3	1.5E-03
48	methionine degradation I (to homocysteine)	4	1.5E-03
49	D-glucuronate degradation I	3	1.5E-03
50	glutathione-mediated detoxification I	3	1.5E-03
51	NAD biosynthesis from 2-amino-3-carboxymuconate semialdehyde	4	1.5E-03
52	4-hydroxyproline degradation I	4	1.5E-03
53	L-dopa degradation	4	1.5E-03
54	histamine degradation	4	1.5E-03
55	selenocysteine biosynthesis II (archaea and eukaryotes)	2	1.6E-03
56	glutamine biosynthesis I	2	1.6E-03
57	alanine biosynthesis III	2	1.6E-03
58	adenine and adenosine salvage VI	2	1.6E-03
59	L-serine degradation	2	1.6E-03
60	glutamine degradation I	2	1.6E-03
61	L-cysteine degradation II	2	1.6E-03
62	asparagine degradation I	2	1.6E-03
63	glutamate dependent acid resistance	2	1.6E-03
64	folate polyglutamylation	3	1.6E-03
65	glucose and glucose-1-phosphate degradation	3	1.6E-03
66	isoleucine degradation I	3	1.6E-03
67	glutamate degradation III (via 4-aminobutyrate)	3	1.6E-03
68	L-cysteine degradation III	3	1.6E-03
69	L-cysteine degradation I	3	1.6E-03
70	β-alanine degradation I	3	1.6E-03
71	aspartate degradation II	3	1.6E-03

72	methylglyoxal degradation VI	3	1.6E-03
73	4-aminobutyrate degradation I	3	1.6E-03
74	serine biosynthesis	3	1.8E-03
75	galactose degradation I (Leloir pathway)	3	1.8E-03
76	putrescine degradation III	4	1.8E-03
77	urate biosynthesis/inosine 5'-phosphate degradation	4	1.8E-03
78	4-hydroxy-2-nonenal detoxification	3	1.8E-03
79	guanosine nucleotides degradation III	4	1.8E-03
80	thio-molybdenum cofactor biosynthesis	2	1.9E-03
81	tyrosine biosynthesis IV	3	1.9E-03
82	PRPP biosynthesis I	2	1.9E-03
83	lactose degradation III	2	1.9E-03
84	pyruvate fermentation to lactate	2	1.9E-03
85	spermine and spermidine degradation I	4	2.1E-03
86	arsenate detoxification I (glutaredoxin)	3	2.1E-03
87	noradrenaline and adrenaline degradation	6	2.2E-03
88	glycoaminoglycan-protein linkage region biosynthesis	3	2.5E-03
89	glycine biosynthesis I	2	2.5E-03
90	glycine biosynthesis III	3	2.5E-03
91	aspartate biosynthesis	2	2.5E-03
92	4-hydroxyphenylpyruvate biosynthesis	2	2.5E-03
93	trehalose degradation II (trehalase)	2	2.5E-03
94	phenylalanine degradation I (aerobic)	3	2.5E-03
95	methylglyoxal degradation I	2	2.5E-03
96	ketogenesis	2	2.5E-03
97	inosine-5'-phosphate biosynthesis II	4	2.6E-03
98	folate transformations I	4	2.6E-03
99	nicotine degradation IV	8	2.9E-03
100	catecholamine biosynthesis	6	3.4E-03
101	heparan sulfate biosynthesis (late stages)	3	3.6E-03
102	taurine biosynthesis	2	3.6E-03
103	guanine and guanosine salvage I	3	3.6E-03
104	phosphatidylethanolamine biosynthesis II	2	3.6E-03
105	ceramide biosynthesis	2	3.6E-03

106	tetrapyrrole biosynthesis II	2	3.6E-03
107	tetrahydrobiopterin biosynthesis II	2	3.6E-03
108	ethanol degradation II	2	3.6E-03
109	ethanol degradation IV	2	3.6E-03
110	oxidative ethanol degradation III	2	3.6E-03
111	ascorbate recycling (cytosolic)	3	5.5E-03
112	ketolysis	2	5.5E-03

Table S-6. Confirmed metabolites in the colon cancer multi-modal study

Name	Neutral mass	RP(+)		RP(-)		HILIC(+)		HILIC(-)	
		FD	p-value	FD	p-value	FD	p-value	FD	p-value
Pyruvate	88.0160			1.5	6E-4			2	5E-5
Lactate	90.0317			1.8	3E-4			2.1	1.8E-4
Putrescine	88.1000	4.0	3.9E-3						
Spermidine	145.1579	6.3	5.5E-4			2.3	1E-2		
Diacetylspermine	286.2369	22.7	8.5E-3			20	5E-4		
Ornithine	132.0899	1.8	1.3E-2	2.2	1.3E-4	1.6	9E-3		
Arginine	174.1117	1.6	9.1E-3	2	4E-3	1.6	5E-2	1.6	5E-2
Citrulline	175.0957	2.4	1.6E-2			1.5	8E-3		
Fumarate	116.0110							2.2	6E-3
Succinate	118.0266							1.7	2E-2
Glutamate	147.0532	2.7	7E-5	1.8	3E-3	2.5	6E-6	2.4	1E-3
Proline	115.0633	2.1	5.5E-3	2.2	2E-4	3	7E-5	2	5E-2
Hydroxyproline	131.0582							2.5	9E-3
Aspartate	133.0375	2.6	1E-2			2.2	3E-2	2.3	5E-3
Leucine	131.0946	2	2.9E-3	1.5	8E-3	1.7	2E-3	1.9	1E-2
Valine	117.0790	2	1.5E-2	1.5	5E-3			2.1	4E-3
Alanine	89.0480			1.7	8E-3	1.8	7E-3	2.2	4E-3

Table S-7. Top 10 predicted pathway dysregulation ranked by the number of dysregulated genes

Pathway	Overlapping genes	Total genes	Overlapping proteins	Total proteins
tRNA charging	23	39	18	39
glutathione-mediated detoxification	17	24	7	28
triacylglycerol biosynthesis	16	27	5	27
3-phosphoinositide biosynthesis	14	28	2	29
fatty acid β-oxidation	13	17	10	24
adenosine ribonucleotides de novo biosynthesis	13	27	14	27
D-myo-inositol-5-phosphate metabolism	13	20	4	20
bile acid biosynthesis, neutral pathway	10	14	4	13
triacylglycerol degradation	10	15	1	16
retinol biosynthesis	10	18	5	18
TCA cycle	10	17	5	12
glycolysis	10	24	7	24
CDP-diacylglycerol biosynthesis	10	22	4	22
stearate biosynthesis	9	12	8	29
γ-linolenate biosynthesis	9	14	5	14
valine degradation	9	13	2	10
putrescine degradation III	9	10	4	10
serotonin degradation	8	9	6	9
gluconeogenesis	8	24	5	24