

Supplementary Data for

**Profiling of polar urine metabolite extracts from Chinese colorectal cancer patients to screen for potential diagnostic
and adverse-effect biomarker**

Yi Deng^{1#}, Houshan Yao^{2#}, Wei Chen¹, Hua Wei¹, Xinxing Li², Feng Zhang¹, Shouhong Gao¹, Huan Man^{1,3}, Jing Chen^{1,3}, Xia Tao¹, Mingming Li^{1}, Wansheng Chen^{1,4*}*

¹Department of Pharmacy, Changzheng Hospital, Secondary Military Medical University, Shanghai, China

²Department of Surgery, Changzheng Hospital, Secondary Military Medical University, Shanghai, China

³College of Chemical and Biological Engineering, Yichun University, Jiangxi Province, China

⁴ Research and Development Center of Chinese Medicine Resources and Biotechnology, Shanghai University of Traditional Chinese Medicine, Shanghai, China

*Corresponding Author:

Mingming Li, Department of Pharmacy, Changzheng Hospital, Secondary Military Medical University, Shanghai, China. Email:

limingming@smmu.edu.cn

Wansheng Chen, Department of Pharmacy, Changzheng Hospital, Secondary Military Medical University, Shanghai, China. Email:

chenwansheng@smmu.edu.cn.

[#]Equal first author

1. Supplementary Tables

Table S1. Clinical information for CRC patients and nonneoplastic controls

Clinical information	CRC patients	Nonneoplastic controls	<i>P</i> -value
In total			
Number	139	50	
Age (media, range)	63, 36 - 87	61, 47 - 89	0.234 ^a
Male/female ratio	91/48	28/22	0.373 ^b
TNM-0	8		
TNM-I	26		
TNM-II	42		
TNM-III	50		
TNM-IV	13		
Training set			
Number	70	25	
Age (media, range)	60, 36 - 87	60, 47 - 78	0.993 ^a
Male/female ratio	46/24	16/9	0.877 ^b
TNM-0	5		
TNM-I	13		
TNM-II	16		
TNM-III	32		
TNM-IV	4		
Testing set			

Number	69	25	
Age (media, range)	65, 40 - 87	61, 49 - 89	0.135 ^a
Male/female ratio	45/24	12/13	0.131 ^b
TNM-0	3		
TNM-I	13		
TNM-II	26		
TNM-III	18		
TNM-IV	9		

^aAge differences between the two groups were analyzed by Student's *t* test. ^bSex differences between the two groups were analyzed by Chi-square test.

Table S2. The grade of AEs in CRC patients with capecitabine-based adjuvant chemotherapy

Grade ^a	HFS	Bone marrow suppression			BMS
		Anemia	Neutropenia	Thrombocytopenia	
NA ^b	1	8	8	8	8
Grade 0	10	23	17	15	11
Grade 1-2	29	9	16	16	19
Grade 3-4	3	3	2	4	5

^aAdverse effects were graded according to the Common Terminology Criteria for Adverse Events (Version 4.0). ^bThese data were missing. Abbreviations: BMS, bone marrow suppression; HFS, hand foot syndrome.

Table S3. Clinical information of participants.

	Group	With AE	Without AE	<i>P</i> -value
HFS	Number	32	10	
	Age (media, range)	58, 37 - 74	57, 36 - 81	0.768 ^a
	Male/female ratio	21/11	6/4	1.000 ^b
	Cycles (media, range)	6.3, 1 - 8	6.5, 1 - 8	0.770 ^a
	TNM-I	1	1	0.656 ^c
	TNM-II	14	4	
	TNM-III	12	4	
	TNM-IV	5	1	
Anemia	Number	12	23	
	Age (media, range)	59, 49 - 74	57, 38 - 81	1.000 ^a
	Male/female ratio	8/4	16/7	0.549 ^b
	Cycles (media, range)	6.2, 2 - 8	6.7, 2 - 8	0.394 ^a
	TNM-I	0	1	0.795 ^c
	TNM-II	6	8	
	TNM-III	2	11	
	TNM-IV	4	3	
Neutropenia	Number	18	17	
	Age (media, range)	59, 37 - 74	57, 38 - 81	0.593 ^a
	Male/female ratio	11/7	13/4	0.471 ^b
	Cycles (media, range)	6.4, 2 - 8	6.7, 2 - 8	0.620 ^a
	TNM-I	1	0	0.902 ^c
	TNM-II	7	7	

	TNM-III	6	7	
	TNM-IV	4	3	
Thrombocytopenia	Number	20	15	
	Age (media, range)	59, 37 - 81	56, 38 - 68	0.329 ^a
	Male/female ratio	13/7	11/4	0.721 ^b
	Cycles (media, range)	7.0, 4 - 8	5.9, 2 - 8	0.092 ^a
	TNM-I	1	0	0.845 ^c
	TNM-II	7	7	
	TNM-III	9	4	
	TNM-IV	3	4	
BMS	Number	24	11	
	Age (media, range)	60, 37 - 81	54, 38 - 69	0.101 ^a
	Male/female ratio	17/7	7/4	0.937 ^b
	Cycles (media, range)	6.7, 2 - 8	6.3, 2 - 8	0.567 ^a
	TNM-I	1	0	0.559 ^c
	TNM-II	9	5	
	TNM-III	9	4	
	TNM-IV	5	2	

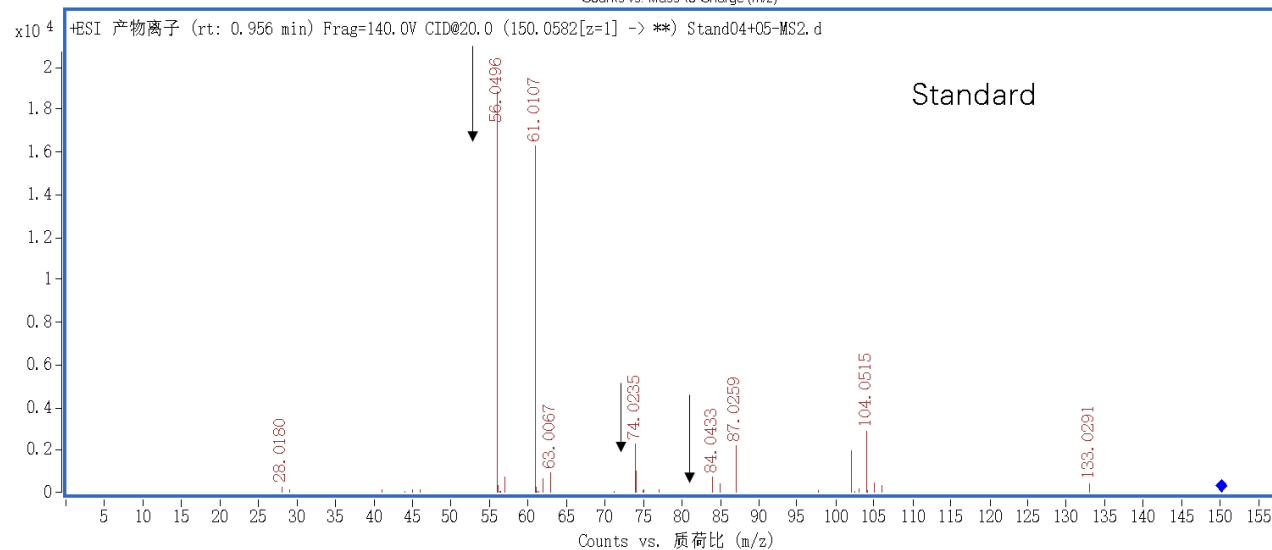
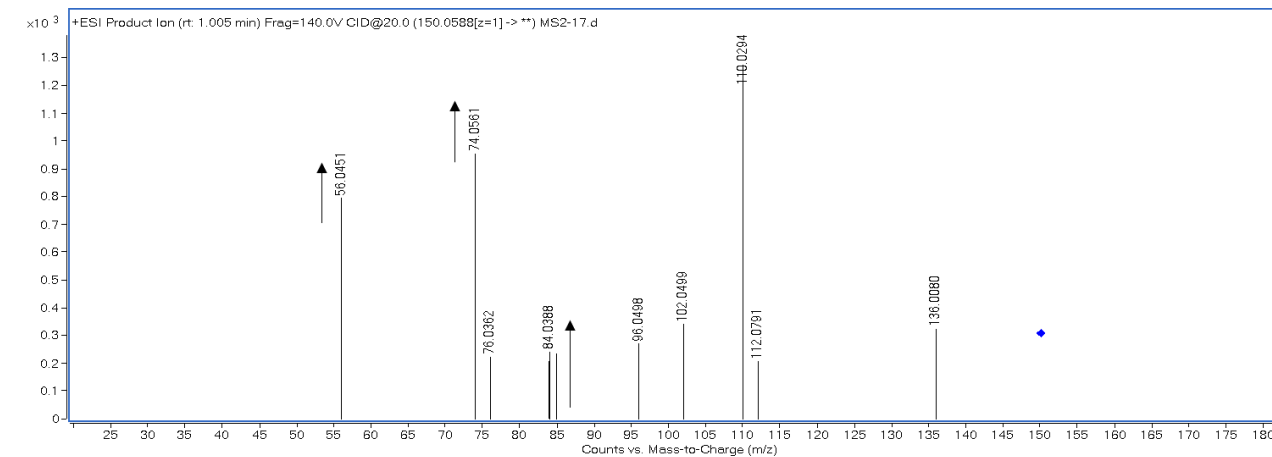
^aAge difference and chemotherapy cycle was analyzed by Student's *t* test. ^bSex difference was analyzed by Chi-square test. ^cComposition difference of the pathological stage was analyzed by Mann-Whitney test. Abbreviations: HFS: hand foot syndrome. BMS: bone marrow suppression.

Table S3. Updated mass spectrums of identified urine metabolites at first revision. For each metabolite identified by standard, the upper figure in each panel shows the spectrum from urine and lower panel shows the spectrum from commercial standards. Compound spectra 1, 2, 10, 19, 30 are obtained by comparison of our experiments with standards. The other metabolites are shown by the spectrum comparison results in HMDB.

No.	Metabolites	Mass spectrum of identified metabolites	Fit (%)	RFit (%)
1	Pyroglutamate	<p>Mass spectrum of identified metabolites for Pyroglutamate. The top panel shows the experimental spectrum from urine (+ESI Product Ion (rt: 1.005 min) Frag=140.0V CID@20.0 (130.0493[z=1] -> **) MS2-11.d) with peaks at m/z 56.0488, 58.0534, 68.0488, 70.0517, 72.0417, 84.0415, 86.0574, 85.0466, and 93.0403. The bottom panel shows the standard spectrum (+ESI Product Ion (rt: 1.076 min) Frag=140.0V CID@20.0 (130.0493[z=1] -> **) WorklistData-LJG-MS2.d) with peaks at m/z 28.0175, 41.0389, 56.0488, 57.0529, 84.0449, 85.0463, and 130.0493. Both panels have a y-axis labeled 'Counts vs. Mass-to-Charge (m/z)' and an x-axis labeled 'Counts vs. Mass-to-Charge (m/z)'.</p>	0.99	0.99

2

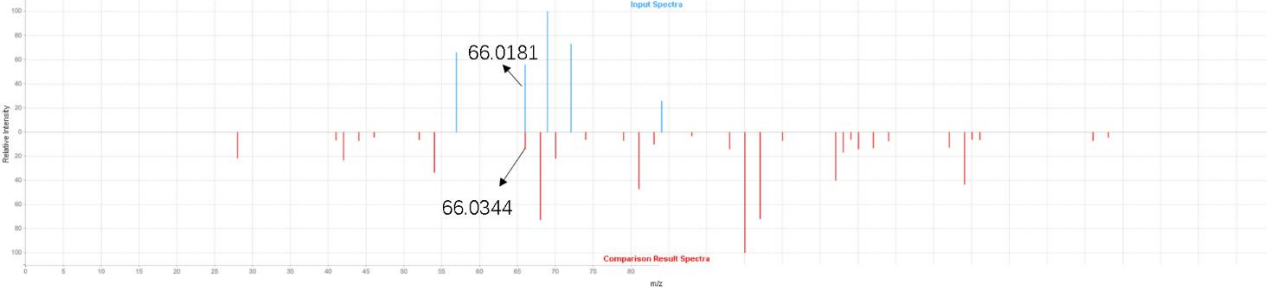
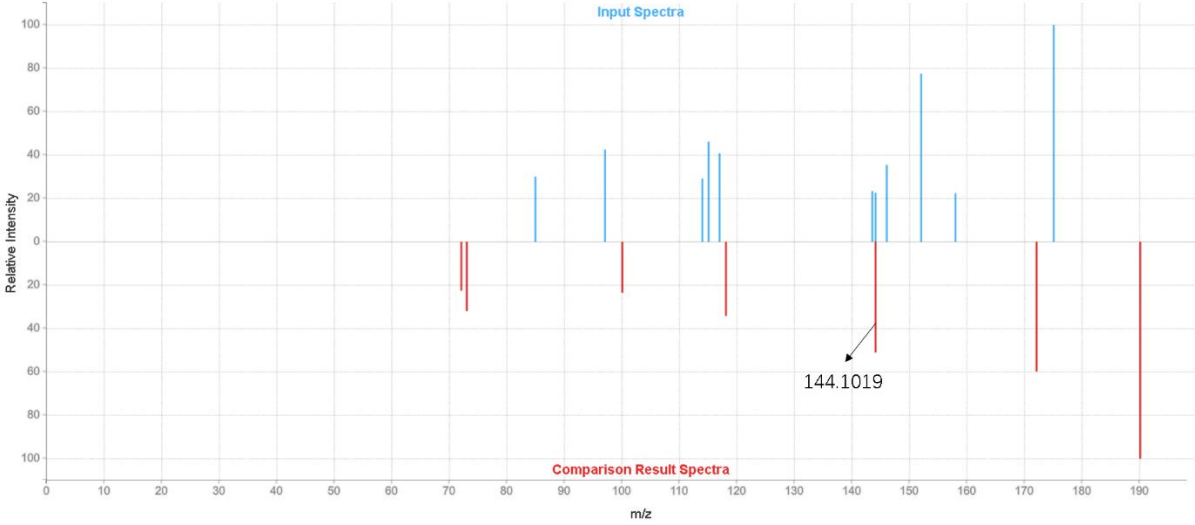
L-Methionine



0.99

0.99

3	5-Acetamidovalerate	<p>Mass spectrum plot for 5-Acetamidovalerate. The y-axis represents Relative Intensity (0 to 100) and the x-axis represents m/z (0 to 170). The plot compares an input spectrum (blue) with a comparison result spectrum (red). Key peaks are labeled with their m/z values: 55.0535, 55.0542, 56.0508, 56.0495, 70.0655, 70.0651, 100.0760, 100.0757, 101.0674, and 101.0597.</p>	0.97	0.71
4	S-(2-carboxypropyl)-Cysteamine	<p>Mass spectrum plot for S-(2-carboxypropyl)-Cysteamine. The y-axis represents Relative Intensity (0 to 100) and the x-axis represents m/z (0 to 170). The plot compares an input spectrum (blue) with a comparison result spectrum (red). Key peaks are labeled with their m/z values: 87.0428, 87.0441, 147.0467, and 56.0474.</p>	0.95	0.48

5	Methylhistidine	 <p>Mass spectrum plot for Methylhistidine. The y-axis is Relative Intensity (0 to 100) and the x-axis is m/z (0 to 80). The plot shows two spectra: Input Spectra (blue) and Comparison Result Spectra (red). Key peaks are labeled at m/z 66.0181 and 66.0344.</p>	0.82	0.07
6	N-lactoyl-Valine	 <p>Mass spectrum plot for N-lactoyl-Valine. The y-axis is Relative Intensity (0 to 100) and the x-axis is m/z (0 to 190). The plot shows two spectra: Input Spectra (blue) and Comparison Result Spectra (red). A key peak is labeled at m/z 144.1019.</p>	0.9	0.53

7	N-Acetylamino-octanoic acid	<p>Mass spectrum plot for N-Acetylamino-octanoic acid. The y-axis is 'Relative intensity' from 0 to 100. The x-axis is 'm/z' from 0 to 190. A blue line represents the 'Input Spectra' and a red line represents the 'Comparison Result Spectra'. Labeled peaks include: 43.0172, 43.0542, 84.9599, 85.1012, 114.1283, and 114.1277.</p>	0.9	0.67
8	N-lactoyl-Leucine	<p>Mass spectrum plot for N-lactoyl-Leucine. The y-axis is 'Relative intensity' from 0 to 100. The x-axis is 'm/z' from 0 to 210. A blue line represents the 'Input Spectra' and a red line represents the 'Comparison Result Spectra'. Labeled peaks include: 57.0294, 57.0699, 68.9930, 68.0699, 84.0773, 84.0808, 86.0289, 86.0964, 114.0459, 114.0918, 186.1125, and 204.1166.</p>	0.97	0.69

9	4-Hydroxy-3-methoxy-cinnamoylglycine	<p>Input Spectra</p> <p>177.0540</p> <p>177.0552</p> <p>Comparison Result Spectra</p> <p>Relative Intensity</p> <p>m/z</p>	0.95	0.57
---	--------------------------------------	--	------	------

10	Alpha-N- Phenylacetyl-L- glutamine	<div data-bbox="548 263 1834 694"> <p>+ESI Product Ion (rt: 4.972 min) Frag=160.0V CID@40.0 (265.1180[z=1] -> **) MS2-01.d</p> <p>Counts vs. Mass-to-Charge (m/z)</p> </div> <div data-bbox="548 710 1834 1141"> <p>+ESI Product Ion (rt: 4.898 min) Frag=160.0V CID@40.0 (265.1187[z=1] -> **) Stand-02-MS2.d</p> <p>Standard</p> <p>Counts vs. Mass-to-Charge (m/z)</p> </div>	0.99	0.99
----	--	--	------	------

11	N-Acetyltyrosine	<p>Mass spectrum plot for N-Acetyltyrosine. The y-axis is 'Relative Intensity' (0-100) and the x-axis is 'm/z' (0-180). The plot shows two spectra: 'Input Spectra' (blue) and 'Comparison Result Spectra' (red). Key peaks are labeled: 69.0689, 130.0704, 86.0964, and 86.0970.</p>	0.85	0.52
12	Indoleacetic acid	<p>Mass spectrum plot for Indoleacetic acid. The y-axis is 'Relative Intensity' (0-100) and the x-axis is 'm/z' (0-180). The plot shows two spectra: 'Input Spectra' (blue) and 'Comparison Result Spectra' (red). Key peaks are labeled: 130.0652 and 130.0657.</p>	0.94	0.46

13	Octenoylglycine	<p>Mass spectrum plot for Octenoylglycine. The y-axis represents Relative Intensity (0 to 100) and the x-axis represents m/z (0 to 210). The plot shows two spectra: Input Spectra (blue) and Comparison Result Spectra (red). Key peaks are labeled with their m/z values: 55.0540, 55.0542, 84.9595, and 85.1012.</p>	0.94	0.56
14	N-Propionylmethionine	<p>Mass spectrum plot for N-Propionylmethionine. The y-axis represents Relative Intensity (0 to 100) and the x-axis represents m/z (0 to 210). The plot shows two spectra: Input Spectra (blue) and Comparison Result Spectra (red). Key peaks are labeled with their m/z values: 56.9636 and 57.0335.</p>	0.86	0.15

15	N-Acetyltryptophan		0.9	0.75
16	Pyro-L-glutaminy-L-glutamine		0.97	0.8

17	Hydroxyphenylacetylglycine	<p>Mass spectrum plot for Hydroxyphenylacetylglycine. The x-axis represents m/z from 0 to 230, and the y-axis represents Relative Intensity from 0 to 100. The plot compares an 'Input Spectra' (blue bars) with a 'Comparison Result Spectra' (red bars). Key peaks are labeled with arrows: 135.0428 (blue) and 135.0441 (red).</p>	0.81	0.7
18	Heptenecoylglycine	<p>Mass spectrum plot for Heptenecoylglycine. The x-axis represents m/z from 0 to 190, and the y-axis represents Relative Intensity from 0 to 100. The plot compares an 'Input Spectra' (blue bars) with a 'Comparison Result Spectra' (red bars). Key peaks are labeled with arrows: 55.0539 (blue), 55.0542 (red), 83.0853 (blue), and 83.0855 (red).</p>	0.89	0.39

19	Creatinine	<div><p>+ESI Product Ion (rt: 0.622 min) Frag=130.0V CID@10.0 (114.0663[z=1] -> **) MS2-01.d</p><p>Counts vs. Mass-to-Charge (m/z)</p><table><tr><th>m/z</th><th>Relative Intensity (approx. x 10⁻⁴)</th></tr><tr><td>5</td><td>1.5</td></tr><tr><td>56.0498</td><td>0.1</td></tr><tr><td>58.0657</td><td>0.1</td></tr><tr><td>60.0566</td><td>0.4</td></tr><tr><td>61.0414</td><td>0.1</td></tr><tr><td>71.0612</td><td>0.2</td></tr><tr><td>74.0610</td><td>0.1</td></tr><tr><td>76.0402</td><td>0.1</td></tr><tr><td>84.0458</td><td>0.1</td></tr><tr><td>86.0716</td><td>0.1</td></tr><tr><td>88.0375</td><td>0.1</td></tr><tr><td>98.9860</td><td>0.1</td></tr><tr><td>113.0826</td><td>0.1</td></tr><tr><td>114.0666</td><td>0.3</td></tr><tr><td>116.1063</td><td>0.1</td></tr><tr><td>118.0870</td><td>1.2</td></tr><tr><td>119.0900</td><td>0.1</td></tr><tr><td>121.0514</td><td>0.1</td></tr><tr><td>130.1098</td><td>0.3</td></tr></table></div> <div><p>+ESI Product Ion (rt: 0.600 min) Frag=130.0V CID@10.0 (114.0656[z=1] -> **) WorklistData--JSG-MS2.d</p><p>Counts vs. Mass-to-Charge (m/z)</p><table><tr><th>m/z</th><th>Relative Intensity (approx. x 10⁻³)</th></tr><tr><td>44.0487</td><td>1.7</td></tr><tr><td>61.0075</td><td>0.3</td></tr><tr><td>84.9599</td><td>0.4</td></tr><tr><td>114.0665</td><td>1.4</td></tr><tr><td>121.0611</td><td>1.7</td></tr></table></div>	m/z	Relative Intensity (approx. x 10 ⁻⁴)	5	1.5	56.0498	0.1	58.0657	0.1	60.0566	0.4	61.0414	0.1	71.0612	0.2	74.0610	0.1	76.0402	0.1	84.0458	0.1	86.0716	0.1	88.0375	0.1	98.9860	0.1	113.0826	0.1	114.0666	0.3	116.1063	0.1	118.0870	1.2	119.0900	0.1	121.0514	0.1	130.1098	0.3	m/z	Relative Intensity (approx. x 10 ⁻³)	44.0487	1.7	61.0075	0.3	84.9599	0.4	114.0665	1.4	121.0611	1.7	0.99	0.99
m/z	Relative Intensity (approx. x 10 ⁻⁴)																																																							
5	1.5																																																							
56.0498	0.1																																																							
58.0657	0.1																																																							
60.0566	0.4																																																							
61.0414	0.1																																																							
71.0612	0.2																																																							
74.0610	0.1																																																							
76.0402	0.1																																																							
84.0458	0.1																																																							
86.0716	0.1																																																							
88.0375	0.1																																																							
98.9860	0.1																																																							
113.0826	0.1																																																							
114.0666	0.3																																																							
116.1063	0.1																																																							
118.0870	1.2																																																							
119.0900	0.1																																																							
121.0514	0.1																																																							
130.1098	0.3																																																							
m/z	Relative Intensity (approx. x 10 ⁻³)																																																							
44.0487	1.7																																																							
61.0075	0.3																																																							
84.9599	0.4																																																							
114.0665	1.4																																																							
121.0611	1.7																																																							

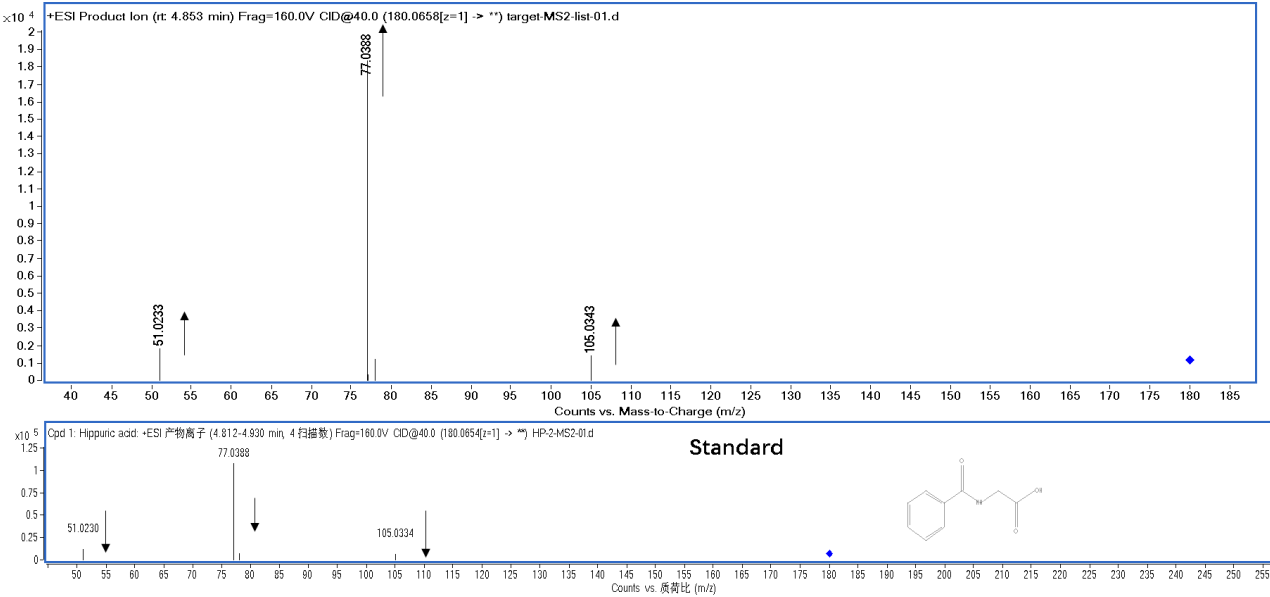
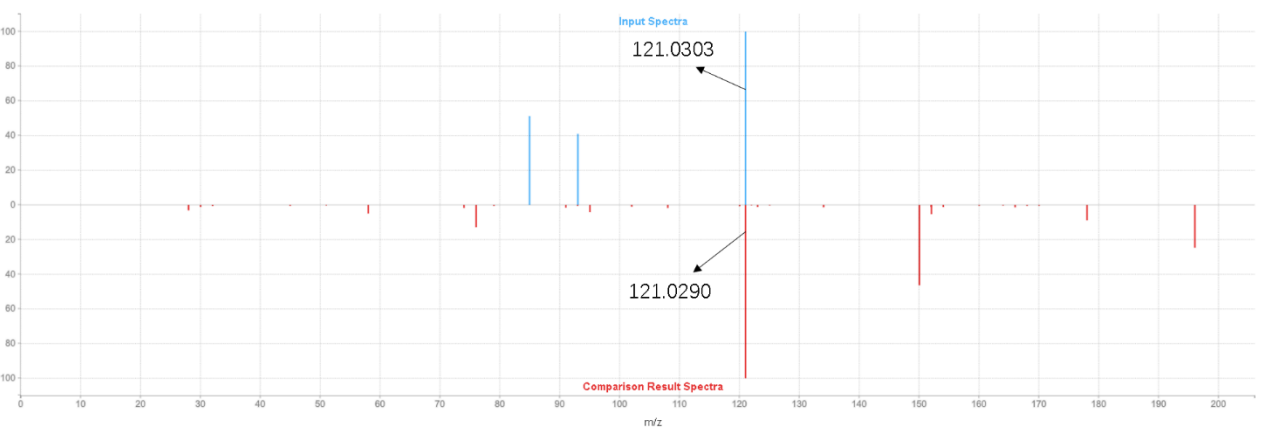
20	Indolylacryloylglycine	<p>Mass spectrum plot for Indolylacryloylglycine. The y-axis represents Relative Intensity (0 to 100) and the x-axis represents m/z (0 to 180). The plot compares an input spectrum (blue) with a comparison result spectrum (red). Key peaks are labeled with their m/z values: 115.0528, 115.0542, 130.0657, 130.0651, 142.0667, and 142.0651.</p>	0.92	0.68
21	8-Hydroxy-5,6-octadienoic	<p>Mass spectrum plot for 8-Hydroxy-5,6-octadienoic. The y-axis represents Relative Intensity (0 to 100) and the x-axis represents m/z (0 to 170). The plot compares an input spectrum (blue) with a comparison result spectrum (red). Key peaks are labeled with their m/z values: 87.0263 and 87.0446.</p>	0.84	0.49

22	N-Heptanoylglycine	<p>Mass spectrum comparison for N-Heptanoylglycine. The plot shows Relative Intensity (0 to 100) versus m/z (0 to 190). The 'Input Spectra' (blue) has major peaks at m/z 84.9651 and 170.1327. The 'Comparison Result Spectra' (red) has major peaks at m/z 85.1012 and 170.1116. The spectra are highly similar.</p>	0.89	0.51
23	cis-4-Decenedioic acid	<p>Mass spectrum comparison for cis-4-Decenedioic acid. The plot shows Relative Intensity (0 to 100) versus m/z (0 to 210). The 'Input Spectra' (blue) has major peaks at m/z 155.0702 and 155.1072. The 'Comparison Result Spectra' (red) has major peaks at m/z 155.1072 and 155.0702. The spectra are highly similar.</p>	0.86	0.39

24	Alanylasparagine	<p>Mass spectrum plot for Alanylasparagine. The y-axis represents Relative Intensity (0 to 100) and the x-axis represents m/z (0 to 150). The plot shows two spectra: Input Spectra (blue) and Comparison Result Spectra (red). Key peaks are labeled with their m/z values: 84.9587 and 84.0398.</p>	0.84	0.64
25	25 1-Methyl-2-nonyl-4(1H)-quinolinone	<p>Mass spectrum plot for 25 1-Methyl-2-nonyl-4(1H)-quinolinone. The y-axis represents Relative Intensity (0 to 100) and the x-axis represents m/z (0 to 290). The plot shows two spectra: Input Spectra (blue) and Comparison Result Spectra (red). Key peaks are labeled with their m/z values: 85.0281, 97.1014, 85.1017, and 97.1017.</p>	0.98	0.76

26	3,4-Methylenesebacic acid		0.93	0.68
27	2-trans,4-cis-Decadienoylcarnitine		0.99	0.87

28	4-Hydroxy-(3',4'-dihydroxyphenyl)-valeric acid		0.95	0.47
29	N ¹ N ¹² Diacetylspermine		0.96	0.41

30	Hippurate	 <p>The figure displays two mass spectra for Hippurate. The top spectrum, titled '+ESI Product Ion (rt: 4.853 min) Frag=160.0V CID@40.0 (180.0658[z=1] -> **) target-MS2-list-01.d', shows relative intensity (x10⁻⁴) versus mass-to-charge ratio (m/z) from 40 to 185. Key peaks are labeled at m/z 51.0233, 77.0388 (base peak), and 105.0343. The bottom spectrum, titled 'Cpd 1: Hippuric acid: +ESI 产物离子 (4.812-4.930 min, 4 扫描数) Frag=160.0V CID@40.0 (180.0654[z=1] -> **) HP-2-MS2-01.d', shows relative intensity (x10⁻⁵) versus m/z from 50 to 255. Key peaks are labeled at m/z 51.0230, 77.0388, and 105.0334. A chemical structure of Hippuric acid is shown to the right of the bottom spectrum.</p>	0.99	0.99
31	Hydroxyhippurate	 <p>The figure compares two mass spectra for Hydroxyhippurate. The top spectrum, labeled 'Input Spectra' in blue, shows relative intensity versus m/z from 0 to 200. Key peaks are labeled at m/z 121.0303 and 121.0290. The bottom spectrum, labeled 'Comparison Result Spectra' in red, shows relative intensity versus m/z from 0 to 200. Key peaks are labeled at m/z 121.0303 and 121.0290.</p>	0.97	0.65

32	Prolyl-Valine	<p>Mass spectrum plot for Prolyl-Valine. The y-axis represents Relative Intensity (0 to 100) and the x-axis represents m/z (0 to 150). The plot shows two spectra: Input Spectra (blue) and Comparison Result Spectra (red). Key peaks are labeled with their m/z values:</p> <ul style="list-style-type: none"> 55.0185 55.0542 68.0507 68.0495 70.1370 72.0831 72.0808 70.0651 98.0618 98.0600 	0.97	0.9
33	Aspartylphenylalanine	<p>Mass spectrum plot for Aspartylphenylalanine. The y-axis represents Relative Intensity (0 to 100) and the x-axis represents m/z (0 to 160). The plot shows two spectra: Input Spectra (blue) and Comparison Result Spectra (red). Key peaks are labeled with their m/z values:</p> <ul style="list-style-type: none"> 120.0803 130.0556 130.0750 120.0402 166.0846 165.9910 	0.95	0.83

34	Phenylacetylglutamine		0.79	0.47
35	Humulinic acid A		0.99	0.72

36	Ubiquinone-1		0.97	0.41
37	4-Pyridoxic acid		0.98	0.96

38	alpha-D-Glucose	<p>Mass spectrum comparison for alpha-D-Glucose. The plot shows Relative Intensity (0 to 100) versus m/z (0 to 150). The 'Input Spectra' (blue) has major peaks at m/z 84.9609 and 116.9772. The 'Comparison Result Spectra' (red) has major peaks at m/z 85.0000 and 117.0000. Arrows point from the input peak labels to the corresponding comparison peaks.</p>	0.82	0.85
39	3-Hydroxydodecane dioic acid	<p>Mass spectrum comparison for 3-Hydroxydodecane dioic acid. The plot shows Relative Intensity (0 to 100) versus m/z (0 to 270). The 'Input Spectra' (blue) has major peaks at m/z 211.1263 and 233. The 'Comparison Result Spectra' (red) has major peaks at m/z 211.1329 and 233. Arrows point from the input peak labels to the corresponding comparison peaks.</p>	0.74	0.66

40	Indoxyl	<p>Mass spectrum comparison for Indoxyl. The plot shows Relative Intensity (0 to 100) on the y-axis and m/z (0 to 130) on the x-axis. The 'Input Spectra' (blue) has major peaks at m/z 51.0226 and 77.0378. The 'Comparison Result Spectra' (red) has major peaks at m/z 51.0000 and 77.000. Arrows point from the labels to the corresponding peaks in both spectra.</p>	0.85	0.71
41	Glutamylproline	<p>Mass spectrum comparison for Glutamylproline. The plot shows Relative Intensity (0 to 100) on the y-axis and m/z (0 to 250) on the x-axis. The 'Input Spectra' (blue) has major peaks at m/z 85.0222 and 244. The 'Comparison Result Spectra' (red) has major peaks at m/z 85.084 and 244. Arrows point from the labels to the corresponding peaks in both spectra.</p>	0.96	0.74

2. Supplementary Figures

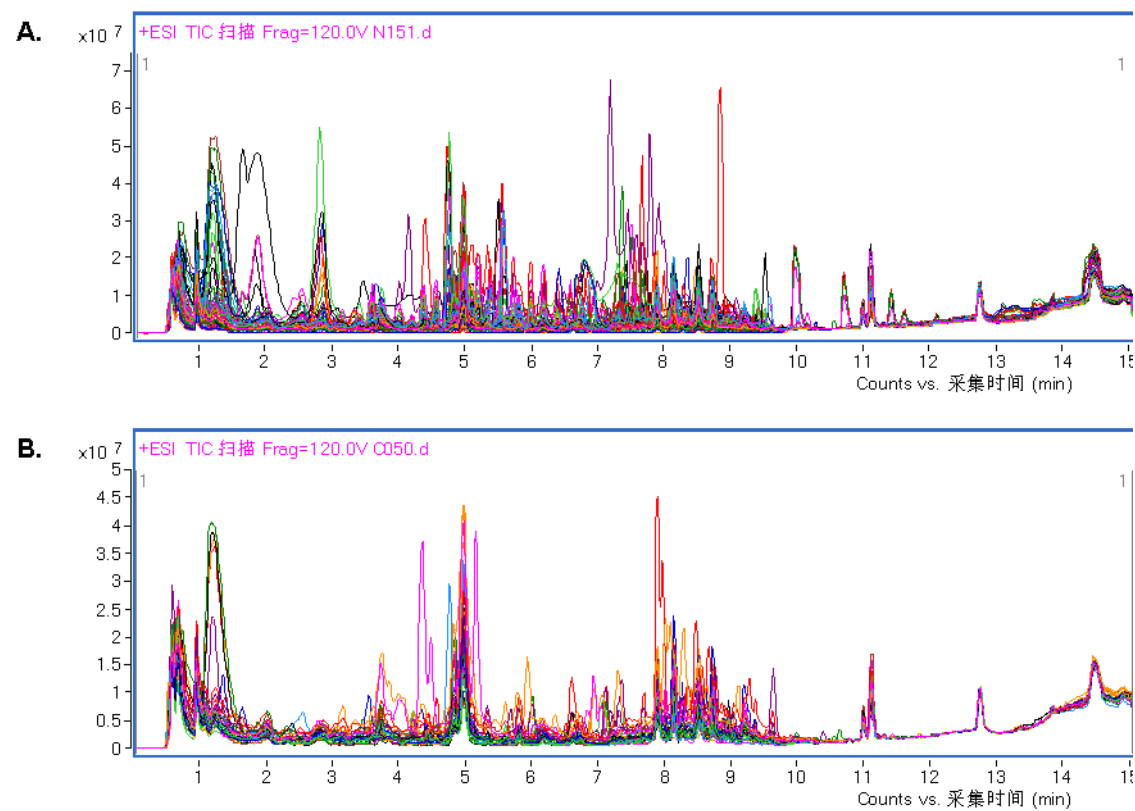


Fig. S1. Total ion chromatograms (TICs) of the urine metabolic profiles obtained from A1. CRC patients, B. nonneoplastic controls.

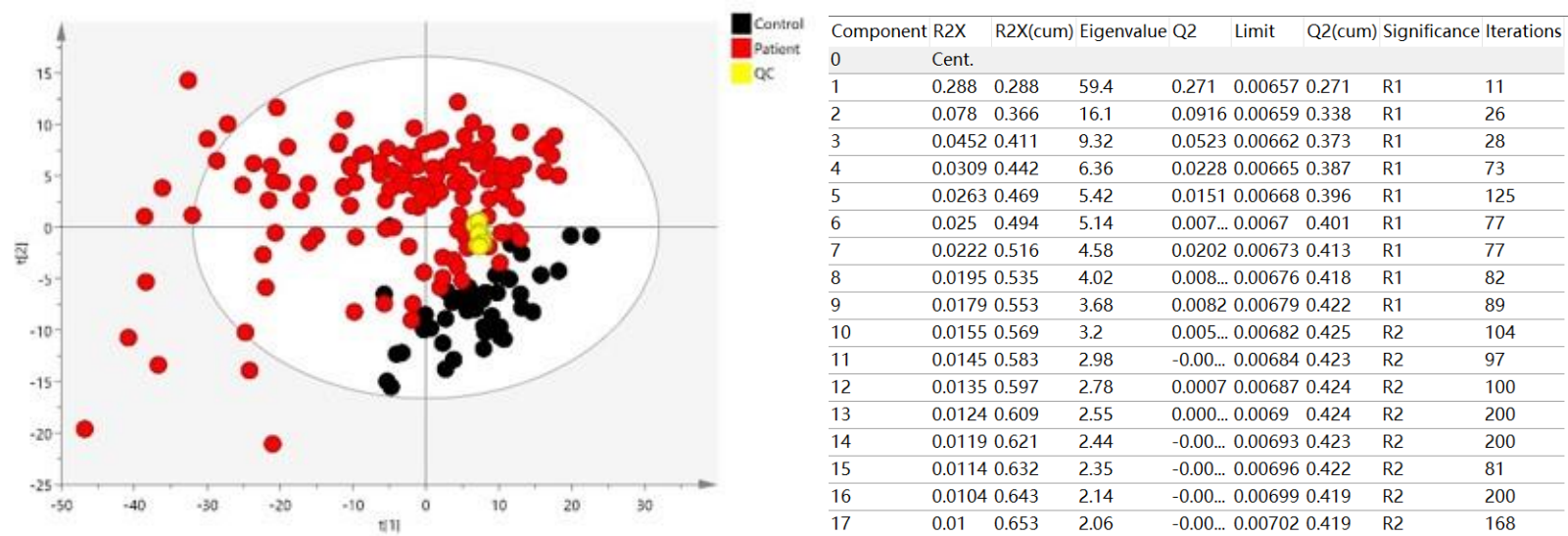


Fig. S2. Principal component analysis (PCA) of QC and experimental samples.

PCA plot (left) and the table of fitness scores (right) are presented. The PCA with all features indicates a tight clustering of QC samples. Black: QC samples, red: CRC patients, yellow: control patients.