Supplementary Data for:

A Novel Bioinformatics Approach to Identify the Consistently Well-performing Normalization Strategy for Current Metabolomic Studies

Qingxia YANG[1,3], Jiajun HONG[2], Yi LI[2], Weiwei XUE[3], Song LI[1,*], Hui YANG[1,*], Feng ZHU[1,2,*]

[1] Multidisciplinary Center for Pituitary Adenomas of Chongqing, Department of Neurosurgery, Xinqiao Hospital, Army Medical University, Chongqing, 400038, China

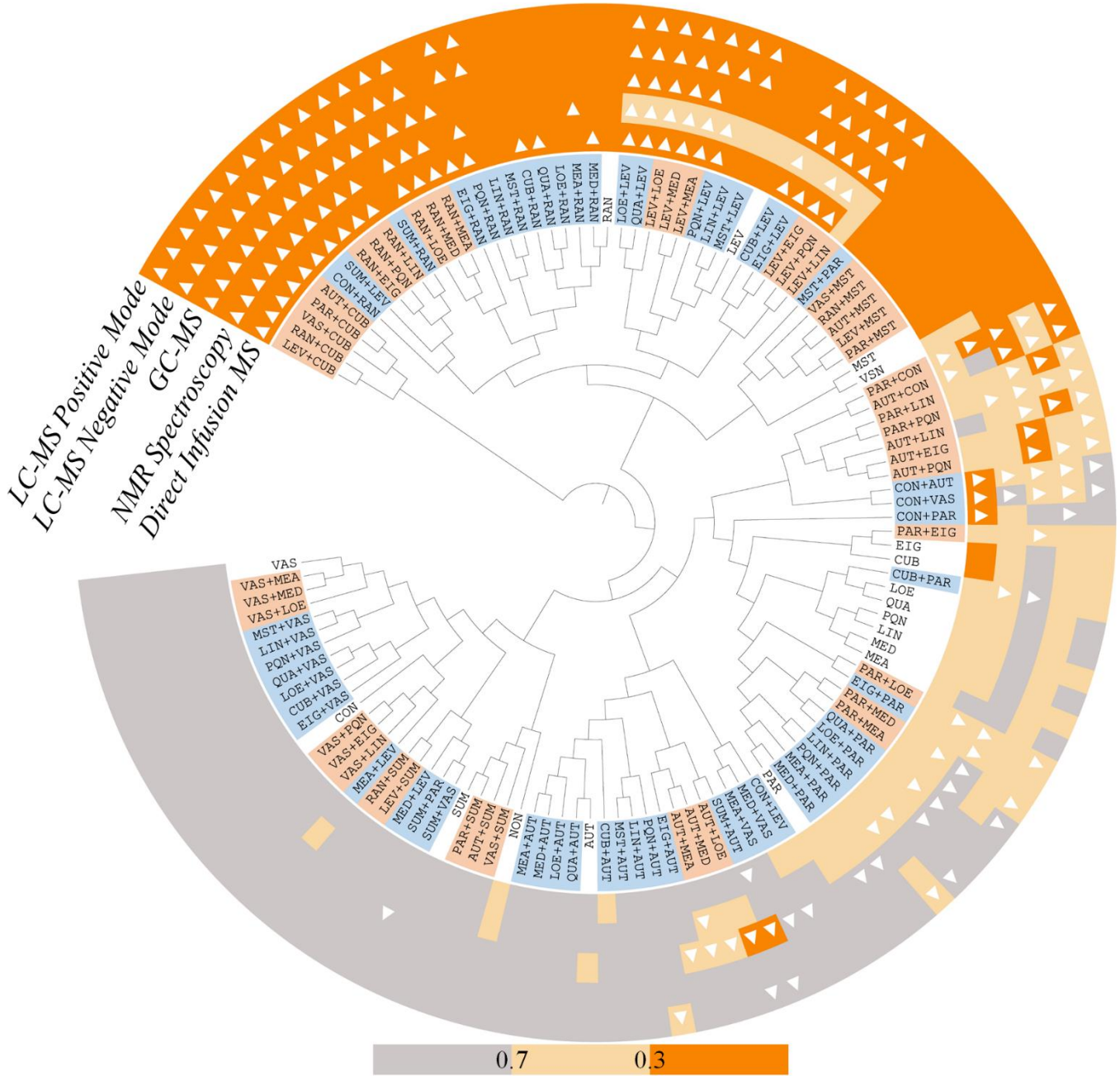
[2] College of Pharmaceutical Sciences, Zhejiang University, Hangzhou 310058, China

[3] School of Pharmaceutical Sciences, Chongqing University, Chongqing 401331, China

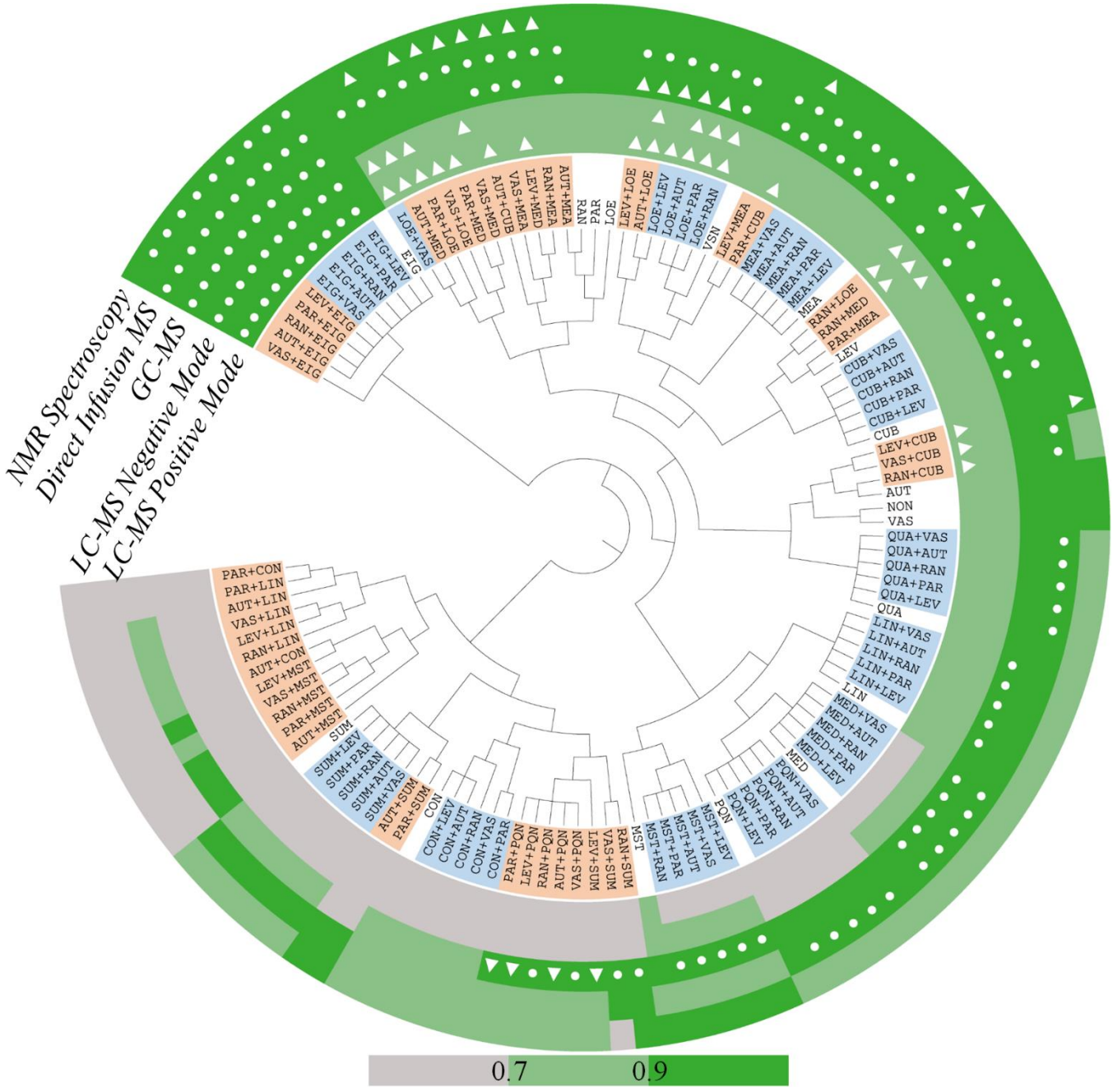
* To whom the correspondence should be addressed: Prof. Feng ZHU, College of Pharmaceutical Sciences, Zhejiang University, Hangzhou 310058, China. E-mail: zhufeng@zju.edu.cn; Prof. Hui YANG, Multidisciplinary Center for Pituitary Adenomas of Chongqing, Department of Neurosurgery, Xinqiao Hospital, Army Medical University, Chongqing, 400038, China. E-mail: huiyangxinqiao@163.com; Prof. Song LI, Multidisciplinary Center for Pituitary Adenomas of Chongqing, Department of Neurosurgery, Xinqiao Hospital, Army Medical University, Chongqing, 400038, China. E-mail: dlisong3@163.com.

Running Title: Approach to Identify Well-performing Normalization

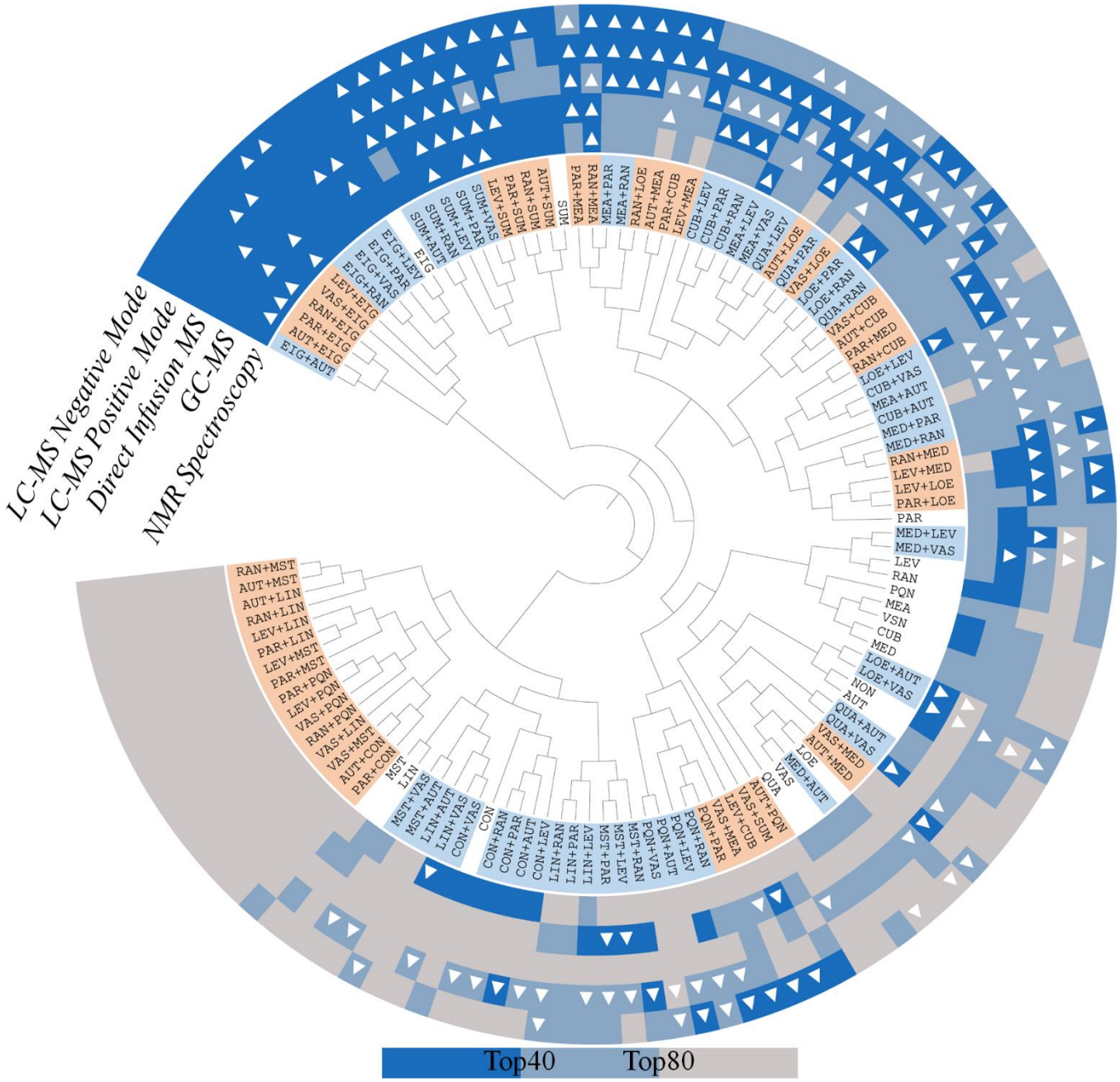
Supplementary Figure S1. The relationship among the performances of all studied strategies identified using the hierarchical clustering of the quantitative metric (PMAD) across five benchmarks representing different analytical platforms. The leaves of this hierarchical tree gave the name of the studied strategies. The background colors of the strategies of a single method, sequential combination of *sample*-based and *metabolite*-based methods, and sequential integration of *metabolite*-based and *sample*-based ones, were white, light blue, and light orange, respectively. The methods with PMAD of superior (≤0.3), good (>0.3 & <0.7) and poor (>0.7) performance were colored by dark orange, light orange, and gray, respectively. If the performance of a combined strategy was better than both single methods within this combination, a triangle was used to highlight that strategy.



Supplementary Figure S2. The relationship among the performances of all studied strategies identified using a hierarchical clustering of the quantitative metric (AUC) across all five benchmarks representing different analytical platforms. The leaves of this hierarchical tree gave the name of the studied strategies. The background colors of the strategies of a single method, sequential combination of *sample*-based and *metabolite*-based methods, and sequential integration of *metabolite*-based and *sample*-based ones, were white, light blue, and light orange, respectively. The methods with AUC of superior (>0.9), good (>0.7 & ≤0.9) and poor (≤0.7) performances were colored by dark green, light green, and gray, respectively. If the AUC values of a combined strategy and any single method in this combination equaled to 1 (perfect classification), a white round dot was applied to highlight that strategy. If the performance of a combined strategy was better than both single methods in the combination, a triangle was used to highlight.



Supplementary Figure S3. The relationship among the performances of all studied strategies identified using the hierarchical clustering of the quantitative metric (CS) across all five benchmarks representing different analytical platforms. The leaves of this hierarchical tree gave the name of the studied strategies. The background colors of the strategies of a single method, sequential combination of *sample*-based and *metabolite*-based methods, and sequential integration of *metabolite*-based and *sample*-based ones, were white, light blue, and light orange, respectively. The methods that ranked to be the top one third, bottom one third, and the remaining one third by the CS values were indicated by dark blue, gray, and light blue color, respectively. If the performance of a combined strategy was better than both single methods within this combination, a triangle was used to highlight that strategy.



Supplementary Table S1. The performances of all normalization strategies across five benchmarks as measured by three criteria. Intragroup variations were assessed by the common logarithm of pooled median absolute deviation (PMAD, the lower the PMADs were, the more thorough the removals of experimentally induced noise were by the studied method); Consistency score (CS) was used to quantitatively measure the overlap among multiple lists of the metabolic markers identified from different partitions of given dataset (the higher the CS values were, the more robust the studied method was in biomarker discovery); The value of area under the ROC curve (AUC) was adopted for achieving the assessments based on support vector machine (the higher the AUC values were, the more capable the studied method was in classifying distinct sample groups). Five benchmarks were named according to **Table 2** as LC-MS Positive Mode, LC-MS Negative Mode, GC-MS, NMR Spectroscopy and Direct Infusion MS by their analytical platform.

Method	LC-MS Positive Mode			LC-MS Negative Mode			GC-MS			Direct Infusion MS			NMR Spectroscopy		
	PMAD	CS	AUC	PMAD	CS	AUC	PMAD	CS	AUC	PMAD	CS	AUC	PMAD	CS	AUC
AUT	-0.09	76.10	0.72	-0.07	74.70	0.75	-0.18	79.30	0.99	-0.08	89.30	1.00	-0.12	69.20	0.91
AUT+CON	-0.59	73.70	0.53	-0.39	66.20	0.53	-0.88	60.30	0.88	-0.37	47.80	0.63	-0.01	58.90	0.67
AUT+CUB	-15.46	90.00	0.77	-15.38	83.20	0.73	-15.53	80.10	1.00	-15.92	91.70	1.00	-16.04	68.40	0.97
AUT+EIG	-0.39	174.20	1.00	-0.33	276.70	1.00	-0.62	121.20	1.00	-0.30	113.20	1.00	-0.25	81.80	1.00
AUT+LIN	-0.39	54.20	0.42	-0.53	67.20	0.50	-0.52	78.00	0.79	-0.13	49.80	0.57	-0.16	55.20	0.38
AUT+LOE	-0.15	83.40	0.75	-0.13	83.30	0.78	-0.55	88.70	1.00	-0.11	96.20	1.00	-0.20	67.10	0.91
AUT+MEA	-0.13	88.70	0.71	-0.12	89.60	0.74	-0.37	84.00	1.00	-0.08	93.70	1.00	-0.16	66.90	0.94
AUT+MED	-0.14	74.40	0.76	-0.13	75.90	0.77	-0.39	77.50	0.94	-0.09	93.30	1.00	-0.17	66.20	0.95
AUT+MST	-2.71	56.50	0.55	-2.54	66.50	0.46	-1.69	75.10	0.95	-0.79	51.50	0.57	-1.25	54.10	0.43
AUT+PQN	-0.36	62.80	0.56	-0.44	70.20	0.61	-0.59	70.10	1.00	-0.22	94.20	0.85	-0.25	57.20	0.73
AUT+SUM	2.75	77.40	0.59	2.60	117.20	0.64	1.03	102.40	0.63	0.89	86.86	1.00	1.29	74.20	0.90
CON	0.14	71.90	0.61	0.11	76.80	0.69	2.77	70.30	0.88	-0.11	52.90	0.86	0.04	86.20	0.78
CON+AUT	-0.14	84.10	0.61	-0.38	76.60	0.69	-0.36	70.70	0.88	-0.99	52.90	0.86	-0.17	71.80	0.78
CON+LEV	0.01	83.30	0.61	0.07	73.40	0.69	-0.08	63.90	0.88	-0.21	52.60	0.86	-0.05	73.20	0.78

CON+PAR	-0.02	83.10	0.61	-0.14	76.60	0.69	1.20	71.40	0.88	-0.56	53.10	0.86	-0.29	73.20	0.80
CON+RAN	-0.93	83.10	0.61	-1.42	76.70	0.69	-0.86	72.10	0.88	-1.95	52.90	0.86	-0.85	73.50	0.79
CON+VAS	-0.14	79.30	0.61	-0.48	76.00	0.69	-0.43	74.50	0.88	-1.45	53.60	0.86	-0.03	71.80	0.80
CUB	-0.26	73.60	0.71	-0.29	74.80	0.78	-0.06	84.30	1.00	-0.73	88.10	1.00	-0.26	67.70	0.90
CUB+AUT	-0.03	80.60	0.71	-0.03	72.90	0.78	-0.12	85.50	1.00	-0.16	91.50	1.00	-0.11	65.00	0.90
CUB+LEV	-1.03	87.90	0.71	-1.06	80.60	0.78	-1.22	96.50	1.00	-0.86	94.50	1.00	-0.43	68.90	0.90
CUB+PAR	-0.16	89.30	0.71	-0.17	81.20	0.78	-0.10	91.80	1.00	-0.48	94.00	1.00	-0.31	67.50	0.90
CUB+RAN	-0.76	87.80	0.71	-0.78	78.60	0.78	-0.68	89.20	1.00	-0.94	94.00	1.00	-0.77	67.40	0.90
CUB+VAS	1.05	82.00	0.71	1.10	77.00	0.78	1.07	86.40	1.00	0.63	92.20	1.00	0.28	68.50	0.90
EIG	-0.36	199.10	1.00	-0.38	308.10	1.00	-0.13	126.40	1.00	-0.55	118.30	1.00	-0.34	75.10	1.00
EIG+AUT	-0.03	177.30	1.00	-0.04	281.10	1.00	-0.36	119.10	1.00	-0.08	115.50	1.00	-0.12	72.10	1.00
EIG+LEV	-1.11	194.60	1.00	-1.14	305.80	1.00	-1.31	126.50	1.00	-0.60	119.60	1.00	-0.36	74.40	1.00
EIG+PAR	-0.20	196.90	1.00	-0.22	299.00	1.00	-0.26	120.20	1.00	-0.33	118.80	1.00	-0.33	72.90	1.00
EIG+RAN	-0.82	187.90	1.00	-0.83	312.90	1.00	-0.92	121.20	1.00	-0.82	117.60	1.00	-0.79	72.80	1.00
EIG+VAS	1.14	195.30	1.00	1.16	313.60	1.00	0.73	117.10	1.00	0.49	116.40	1.00	0.20	74.60	1.00
LEV	-0.87	71.80	0.71	-0.91	81.90	0.81	-0.94	91.50	0.99	-0.54	93.10	1.00	-0.37	70.70	0.91
LEV+CUB	-16.29	73.40	0.76	-16.29	88.20	0.72	-16.54	91.80	0.97	-16.37	86.80	1.00	-16.37	66.30	0.91
LEV+EIG	-1.13	204.80	1.00	-1.14	307.90	1.00	-1.30	127.60	1.00	-0.75	106.10	1.00	-0.45	75.80	1.00
LEV+LIN	-1.06	67.60	0.62	-1.30	71.40	0.48	-1.43	75.70	0.74	-0.52	51.10	0.64	-0.29	57.20	0.67
LEV+LOE	-0.92	79.90	0.78	-0.96	85.20	0.77	-1.43	90.50	1.00	-0.58	97.30	1.00	-0.45	69.30	0.92
LEV+MEA	-0.90	90.60	0.74	-0.94	79.90	0.77	-1.05	87.20	1.00	-0.54	106.30	1.00	-0.39	62.80	0.95
LEV+MED	-0.91	79.40	0.73	-0.95	84.40	0.73	-1.02	90.30	1.00	-0.55	99.80	1.00	-0.41	66.20	0.96

LEV+MST	-2.64	70.40	0.54	-2.50	68.70	0.48	-1.98	73.90	0.96	-0.81	50.40	0.67	-1.27	56.80	0.60
LEV+PQN	-1.12	63.20	0.56	-1.26	73.60	0.61	-1.48	69.20	1.00	-0.68	52.00	0.85	-0.46	57.90	0.73
LEV+SUM	1.94	120.60	0.54	1.71	155.20	0.55	0.23	127.10	1.00	0.44	95.00	0.87	1.13	80.95	0.77
LIN	-0.14	69.20	0.71	-0.19	71.70	0.71	0.27	84.10	1.00	-0.45	50.00	0.93	-0.34	67.30	0.87
LIN+AUT	-0.14	74.30	0.71	-0.10	78.50	0.71	-0.12	74.00	1.00	-0.07	50.50	0.93	-0.12	66.00	0.87
LIN+LEV	-0.91	78.90	0.71	-0.95	81.60	0.71	-0.93	89.70	1.00	-0.53	50.90	0.93	-0.37	67.60	0.87
LIN+PAR	-0.15	77.80	0.71	-0.15	82.40	0.71	0.07	85.70	1.00	-0.28	50.00	0.93	-0.35	66.70	0.87
LIN+RAN	-1.00	79.30	0.71	-0.93	83.40	0.71	-0.68	83.40	1.00	-0.79	50.60	0.93	-0.77	66.80	0.87
LIN+VAS	0.72	71.50	0.71	0.88	75.00	0.71	0.82	75.60	1.00	0.45	51.90	0.93	0.20	72.00	0.87
LOE	-0.15	71.60	0.71	-0.19	76.00	0.76	-0.04	73.30	0.98	-0.51	93.70	1.00	-0.34	67.50	0.96
LOE+AUT	-0.05	74.80	0.75	-0.05	78.10	0.78	-0.10	79.90	1.00	-0.07	92.60	1.00	-0.12	72.50	0.94
LOE+LEV	-0.90	82.10	0.76	-0.94	77.70	0.78	-1.21	87.40	1.00	-0.59	95.20	1.00	-0.41	72.70	0.91
LOE+PAR	-0.11	83.10	0.75	-0.14	88.30	0.78	-0.07	82.20	1.00	-0.30	96.20	1.00	-0.34	73.20	0.94
LOE+RAN	-0.81	79.20	0.75	-0.82	88.40	0.78	-0.66	82.60	1.00	-0.76	96.50	1.00	-0.78	73.50	0.94
LOE+VAS	0.91	73.30	0.75	0.99	81.60	0.77	1.07	77.40	0.94	0.49	94.70	1.00	0.25	73.70	0.94
MEA	-0.15	73.40	0.77	-0.19	70.90	0.77	0.11	83.60	1.00	-0.41	89.60	1.00	-0.14	70.90	0.95
MEA+AUT	-0.08	82.40	0.77	-0.07	79.30	0.77	-0.11	84.90	1.00	-0.06	93.20	1.00	-0.14	70.40	0.95
MEA+LEV	0.66	85.70	0.77	0.48	82.20	0.77	0.45	87.70	1.00	0.53	97.60	1.00	-0.24	72.30	0.95
MEA+PAR	-0.12	88.20	0.77	-0.15	84.60	0.77	-0.01	88.20	1.00	-0.26	97.30	1.00	-0.20	70.20	0.95
MEA+RAN	-0.87	87.20	0.77	-0.87	85.70	0.77	-0.66	85.20	1.00	-0.79	95.80	1.00	-0.82	70.40	0.95
MEA+VAS	0.05	85.70	0.77	0.16	81.50	0.77	0.01	87.30	1.00	-0.10	93.70	1.00	0.25	69.90	0.95
MED	-0.16	70.30	0.68	-0.20	73.70	0.73	0.12	84.50	1.00	-0.42	90.90	1.00	-0.23	68.90	0.88

MED+AUT	-0.09	69.30	0.68	-0.07	75.90	0.73	-0.11	77.50	1.00	-0.08	90.80	1.00	-0.09	67.50	0.88
MED+LEV	0.53	73.40	0.68	1.70	78.20	0.73	0.55	89.40	1.00	0.85	95.50	1.00	0.15	69.90	0.88
MED+PAR	-0.13	77.10	0.68	-0.15	82.10	0.73	0.00	86.70	1.00	-0.27	94.80	1.00	-0.28	68.30	0.88
MED+RAN	-0.88	77.40	0.68	-0.87	83.20	0.73	-0.68	84.80	1.00	-0.80	94.20	1.00	-0.76	67.70	0.88
MED+VAS	0.04	73.70	0.68	0.16	77.70	0.73	-0.05	92.20	1.00	-0.28	92.10	1.00	0.12	69.20	0.88
MST	-4.13	76.10	0.74	-3.94	69.10	0.77	-2.61	84.50	1.00	-2.20	53.10	0.86	-2.24	65.30	0.93
MST+AUT	-0.14	79.70	0.68	-0.10	73.20	0.77	-0.12	76.60	1.00	-0.07	54.20	0.86	-0.12	64.40	0.93
MST+LEV	-0.91	83.90	0.69	-0.95	75.80	0.77	-0.93	94.40	1.00	-0.53	53.90	0.86	-0.37	64.80	0.93
MST+PAR	-2.14	83.20	0.68	-2.03	77.40	0.77	-1.37	93.60	1.00	-1.15	53.20	0.86	-1.29	65.20	0.93
MST+RAN	-1.00	86.20	0.68	-0.93	80.20	0.77	-0.68	88.80	1.00	-0.79	54.00	0.86	-0.77	64.70	0.93
MST+VAS	0.72	79.20	0.68	0.88	72.50	0.77	0.82	78.60	1.00	0.45	54.30	0.86	0.20	69.60	0.93
NON	2.97	75.30	0.76	2.96	74.30	0.71	7.32	74.30	1.00	-0.42	91.30	1.00	-0.23	72.80	0.93
PAR	-0.11	82.40	0.70	-0.13	83.80	0.74	0.01	90.80	0.99	-0.27	92.50	1.00	-0.31	69.30	0.96
PAR+CON	-0.59	68.80	0.55	-0.48	81.60	0.59	-0.62	63.50	0.69	-0.45	50.00	0.62	-0.59	56.40	0.51
PAR+CUB	-15.48	91.90	0.75	-15.43	92.40	0.78	-15.41	88.00	1.00	-16.09	93.80	1.00	-16.26	69.20	0.97
PAR+EIG	-0.35	182.40	1.00	-0.33	271.10	1.00	-0.21	125.10	1.00	-0.49	116.80	1.00	-0.33	77.40	1.00
PAR+LIN	-0.36	69.60	0.50	-0.55	68.10	0.53	-0.41	77.20	0.69	-0.31	50.40	0.61	-0.27	58.40	0.41
PAR+LOE	-0.17	81.20	0.76	-0.18	85.20	0.78	-0.38	87.90	0.96	-0.31	97.80	1.00	-0.41	68.20	0.91
PAR+MEA	-0.15	86.20	0.71	-0.18	84.40	0.80	-0.14	91.10	1.00	-0.27	95.50	1.00	-0.30	70.00	0.93
PAR+MED	-0.16	79.00	0.74	-0.18	79.20	0.73	-0.15	86.40	0.95	-0.29	96.90	1.00	-0.33	68.60	0.97
PAR+MST	-2.68	70.80	0.56	-2.52	75.60	0.57	-1.84	75.20	0.95	-0.82	52.30	0.53	-1.24	57.50	0.59
PAR+PQN	-0.38	63.30	0.56	-0.48	72.20	0.61	-0.46	70.60	1.00	-0.42	50.50	0.85	-0.40	57.90	0.73

PAR+SUM	2.72	89.90	0.58	2.50	138.20	0.44	1.20	108.20	0.63	0.71	95.49	1.00	1.17	72.05	0.91
PQN	-0.16	70.80	0.69	-0.20	76.90	0.68	0.33	84.70	0.98	-0.44	81.90	1.00	-0.30	67.50	0.87
PQN+AUT	-0.17	76.60	0.69	-0.11	84.70	0.68	-0.08	74.60	0.98	-0.09	82.10	1.00	-0.14	64.40	0.87
PQN+LEV	-0.92	77.40	0.69	-0.95	84.50	0.68	-0.88	89.90	0.98	-0.55	82.40	1.00	-0.41	65.70	0.87
PQN+PAR	-0.17	77.90	0.69	-0.17	85.60	0.68	0.11	84.90	0.98	-0.29	82.50	1.00	-0.35	66.40	0.87
PQN+RAN	-1.04	77.20	0.69	-0.94	89.40	0.68	-0.65	83.10	0.98	-0.83	82.00	1.00	-0.80	66.40	0.87
PQN+VAS	0.68	74.20	0.69	0.88	79.80	0.68	0.84	77.10	0.98	0.42	82.10	1.00	0.20	67.20	0.87
QUA	-0.15	71.60	0.76	-0.19	68.40	0.76	-0.01	79.60	0.99	-0.46	92.50	1.00	-0.30	67.30	0.88
QUA+AUT	-0.05	63.70	0.76	-0.05	73.40	0.76	-0.15	79.30	0.99	-0.10	92.00	1.00	-0.14	68.00	0.88
QUA+LEV	-0.90	84.54	0.76	-0.94	80.50	0.76	-1.18	88.60	0.99	-0.56	96.40	1.00	-0.41	68.60	0.88
QUA+PAR	-0.11	85.36	0.76	-0.13	80.60	0.76	-0.08	83.00	0.99	-0.30	96.40	1.00	-0.35	67.70	0.88
QUA+RAN	-0.81	83.07	0.76	-0.82	83.30	0.76	-0.70	81.40	0.99	-0.83	95.80	1.00	-0.78	70.10	0.88
QUA+VAS	0.91	64.90	0.76	0.99	75.20	0.76	0.92	75.30	0.99	0.40	93.20	1.00	0.20	72.20	0.88
RAN	-0.89	67.80	0.70	-0.87	83.20	0.74	-0.72	88.90	1.00	-0.80	92.20	1.00	-0.80	70.70	0.95
RAN+CUB	-16.27	77.90	0.72	-16.19	76.60	0.73	-16.00	87.10	0.97	-16.64	95.10	1.00	-16.74	68.50	0.89
RAN+EIG	-1.16	170.10	1.00	-1.12	264.50	1.00	-1.14	127.20	1.00	-1.03	119.90	1.00	-0.92	80.90	1.00
RAN+LIN	-1.17	52.00	0.57	-1.31	69.70	0.45	-1.07	77.50	0.70	-0.83	50.50	0.60	-0.82	59.90	0.57
RAN+LOE	-0.94	87.70	0.72	-0.93	85.50	0.80	-1.07	85.90	1.00	-0.83	98.90	1.00	-0.86	68.80	0.96
RAN+MEA	-0.92	89.20	0.73	-0.92	84.90	0.73	-0.87	89.80	0.98	-0.81	95.10	1.00	-0.83	71.40	0.97
RAN+MED	-0.93	77.50	0.74	-0.92	84.70	0.79	-0.89	85.90	0.96	-0.82	95.60	1.00	-0.84	68.00	0.97
RAN+MST	-2.69	57.40	0.60	-2.51	64.20	0.49	-1.72	75.50	0.96	-0.76	50.00	0.58	-1.24	57.50	0.54
RAN+PQN	-1.16	64.80	0.56	-1.23	67.40	0.61	-1.13	70.30	1.00	-0.94	50.10	0.85	-0.92	58.40	0.73

RAN+SUM	1.96	100.70	0.53	1.76	142.60	0.59	0.50	102.70	1.00	0.16	84.05	1.00	0.61	75.45	0.92
SUM	2.69	97.70	0.59	2.47	116.60	0.53	1.41	104.80	0.84	0.61	92.40	0.99	1.28	78.10	0.86
SUM+AUT	-0.15	111.50	0.59	-0.05	154.40	0.53	-0.63	86.80	0.84	-0.09	108.60	0.99	-0.10	75.30	0.86
SUM+LEV	-1.29	109.50	0.59	-1.28	160.40	0.53	-1.43	133.10	0.84	-1.15	96.10	0.99	-0.67	89.00	0.86
SUM+PAR	1.27	109.50	0.59	1.21	158.80	0.53	0.39	116.70	0.84	0.26	95.20	0.99	0.59	75.60	0.86
SUM+RAN	-1.14	115.20	0.59	-1.02	178.40	0.53	-1.09	97.80	0.84	-0.79	96.80	0.99	-0.76	74.05	0.86
SUM+VAS	0.99	118.70	0.59	1.17	155.50	0.53	0.17	118.70	0.84	0.97	95.50	0.99	0.46	84.70	0.86
VAS	0.78	72.40	0.74	0.90	77.40	0.70	0.76	77.00	1.00	0.43	91.30	1.00	0.19	70.20	0.92
VAS+CUB	-14.58	82.40	0.76	-14.37	87.40	0.72	-14.37	81.80	0.98	-15.39	90.90	1.00	-15.73	68.40	0.87
VAS+EIG	0.46	220.70	1.00	0.61	299.90	1.00	0.31	97.60	1.00	0.17	100.50	1.00	0.07	74.70	0.99
VAS+LIN	0.35	57.30	0.62	0.32	76.30	0.50	0.70	61.90	0.74	0.28	53.80	0.66	0.03	66.20	0.58
VAS+LOE	0.72	85.20	0.79	0.84	81.90	0.75	0.51	82.50	0.91	0.41	100.40	1.00	0.12	65.00	0.97
VAS+MEA	0.75	72.80	0.72	0.87	86.00	0.72	0.70	83.70	1.00	0.43	82.40	1.00	0.16	64.10	0.97
VAS+MED	0.74	68.60	0.73	0.86	78.30	0.74	0.69	76.90	0.93	0.42	95.70	1.00	0.15	67.10	0.95
VAS+MST	-2.81	65.50	0.58	-2.60	69.10	0.54	-1.19	60.10	0.86	-0.81	54.10	0.61	-1.26	61.40	0.56
VAS+PQN	0.48	63.20	0.56	0.51	69.90	0.61	0.48	68.50	1.00	0.30	52.00	0.85	0.05	60.50	0.73
VAS+SUM	3.71	76.10	0.57	3.59	93.50	0.59	2.04	87.30	1.00	1.41	69.20	0.98	1.59	66.00	0.50
VSN	-4.01	73.90	0.76	-3.68	72.30	0.78	-4.13	84.30	1.00	-0.25	92.20	1.00	-0.17	72.40	0.93

Supplementary Method S1. Normalization Methods Analyzed in This Study

A. ***Sample-based Normalization Methods*** (11 methods in total)

1. ***Contrast*** **(Contrast Normalization, CON)**

CON is a sample-based normalization[1], which adopts *MA*-plots assuming the presence of non-linear biases[2]. The inputs could be logged and transformed into the contrast space based on an orthonormal transformation matrix[2]. But the log function applied in this method cannot process zeros and negative numbers, which requires the conversion of non-positive numbers to an extremely small value[2]. CON has been employed to reveal the role of polychlorinated biphenyls in non-alcoholic fatty liver disease[3]. CON is defined as follows (*i* and *j* represent the metabolite and sample, respectively).

Data matrix uses the alternative matrix:

$$\mathrm{Z} = \log(\mathrm{X}) \cdot \mathrm{T}$$

$$z_{ij} = \sum_{k=1}^{J} log\,(x_{ik}) m_{kj} = \left[\vec{x}_{sum,}\vec{x}_{cont_1}, \ldots \vec{x}_{cont_{J-1}}\right]$$

where $\mathrm{T} = (\mathrm{t_{ij}})$ is the orthonormal transformation matrix; Then evaluating contrasts by computing multi-loess fits $\left[\hat{\vec{x}}_{cont_1}, \ldots \hat{\vec{x}}_{cont_{J-1}}\right]$, using the Euclidean distance $\epsilon = \sqrt{\sum_{j=1}^{J-1}\left(\hat{\vec{x}}_{cont_j} - \vec{x}_{cont_j}\right)^2}$. Loess $\left[\tilde{\vec{x}}_{sum}, \tilde{\vec{x}}_{cont_1}, \ldots \tilde{\vec{x}}_{cont_{J-1}}\right]$ map smoothly from contrasts to zero:

$$\left[\vec{x}_{sum}, \hat{\vec{x}}_{cont_1}, \ldots \hat{\vec{x}}_{cont_{J-1}}\right] \cdot T \mapsto [\vec{x}_{sum}, 0, \ldots 0] \cdot \mathrm{T}$$

CON expands the MA-plots to several dimensions and converts the data into set of rows representing orthonormal contrasts. But the use of log function impedes the handling of negative values and zeros.

2. ***Cubic Splines*** **(CUB)**

CUB is one of the sample-based normalizations assuming the existence of non-linear relation between baseline and individual spectra[2, 4]. The aim of CUB is to make the distribution of metabolite intensities similar across all samples, which regards the geometric or arithmetic mean of the concentrations of each metabolite across all samples as the baseline sample[5]. Then, a set of evenly distributed quantiles from both the baseline and the target samples is used to fit a smooth cubic spline[5]. Finally, the spline function generator uses the generated set of interpolated splines to fit the parameters of a natural cubic spline[5]. CUB has been used to reduce variability in DNA microarray experiments and metabolomics profiling[4, 6], and it was defined as follows:

$$x_{target_i} = \frac{1}{J}\sum_{j=1}^{J} x_{ij}$$

Evenly spaced set of N quantiles of the target sample and sample j: $(x_{target_n}, x_{jn})_{n=1\ldots N}$. And the use of cubic spline generator for each iteration k = 1…K: $c_{jk} = f(x_{target_n}, x_{jn})$ leads to the interpolated spline $s_j = \frac{1}{K}\sum_{k=1}^{K} c_{jk}$. In the n-th interval, the spline of sample j is defined by the parameters $\vec{a}_{jn}$

$$s_{jn}(x) = a_{jn_1} + a_{jn_2}(x - x_{target_n}) + a_{jn_3}(x - x_{target_n})^2 + a_{jn_4}(x - x_{target_n})^3$$

The normalized intensities are: $\tilde{x}_{ij} = s_j(x_{ij})$, and a set of evenly distributed quantiles is taken from both the target spectrum and the sample spectrum and used to fit a smooth cubic spline.

3. *Cyclic Loess* (Cyclic Locally Weighted Regression, LOE)

LOE is sample-based normalization[7] based on MA-plots constituting logged Bland-Altman plots[2]. It can estimate a regression surface using multivariate smoothing procedure[8]. However, LOE is one of the most time-consuming methods among all normalizations and the amount of time consumed grows exponentially as the number of sample increases[9]. LOE has been applied in metabolomic profiling to remove the systematic effect[10]. The Cyclic Loess was defined as:

$$M_{ij_1j_2} = log_2(x_{ij_1}) - log_2(x_{ij_2}) = log_2(\frac{x_{ij_1}}{x_{ij_2}})$$

$$A_{ij_1j_2} = \frac{1}{2}\left(log_2(x_{ij_1}) + log_2(x_{ij_2})\right) = \frac{1}{2}log_2(x_{ij_1}x_{ij_2})$$

4. *EigenMS* (EIG)

EIG is a sample-based normalization method and is able to remove bias of unknown complexity from metabolomics data[11]. This method could increase the sensitivity in differential analysis and preserve the original differences at the same time[12]. EIG achieves its efficacy by the following 3 steps[11, 13]: (**1**) preserving true differences in the metabolomics data by estimating treatment effects with an ANOVA model; (**2**) singular value decomposition of the residual matrix is used to determine bias trends in the data; (**3**) the number of bias trends is estimated via a permutation test and the effects of the bias trends are eliminated. It has been applied in quantitative label-free proteomics profiling[12] and metabolomics studies[11]. The equation of EIG is defined as follows. Estimate B by least squares: $\hat{B} = RV_0$, and V_0 is an orthonormal matrix.

$$\tilde{x}_{ij} = x_{ij} - \hat{B}V_0$$

5. *Linear Baseline* (Linear Baseline Scaling, LIN)

As one sample-based normalization method[14], LIN assumes that there is a constant linear relationship between each metabolite of a given sample and the baseline, so it can map each sample to the baseline by the scaling factor[2]. The median of each metabolite across all samples is determined as the baseline. The scaling factor is computed as the ratio of the mean intensity of the baseline to the mean intensity of each sample[2], and the intensities of all the samples are multiplied by their particular scaling factors[2]. However, all these calculation based on the assumption of a linear correlation among samples might be oversimplified[2]. It has been applied to identify differential metabolomics profiles[15] and normalize nuclear magnetic resonance (NMR)-based metabolomics data[16] and MS-based metabolomics data[10]. The equation is defined as following:

$$\tilde{x}_{ij} = \frac{\bar{x}_{baseline}}{\bar{x}_j} \times x_{ij}$$

6. *Mean Normalization* (MEA)

MEA is sample-based method to eliminate the background effect by normalizing the data using mean value of all signals[17]. The mean intensity of all variables is determined by intensity of each metabolite in a given sample[18]. The means of the intensities for each sample are forced to be equal to one another using MEA on the purpose to make the samples comparable[10]. As a result, the mean of all abundances in one sample equals one[18]. This method has been applied to normalize the metabolomics data[10].

$$\tilde{x}_{ij} = x_{ij} - mean\left(x_{ij:i=1,\dots,I}\right)$$

7. *Median Normalization* (MED)

MED, a sample-based normalization method, has a basic assumption that the samples of input dataset are separated by a constant[19]. The median of the metabolite abundances in each sample equals one[19]. For MED, it is more practical than the Total Sum Normalization especially in situations where several saturated abundances may be associated with some of the factors of interest[19]. MED has been used in proteomics analysis[20] and metabolomics analysis[10]. The equation is defined as:

$$\tilde{x}_{ij} = x_{ij} - median\left(x_{ij:i=1,\dots,I}\right)$$

8. *MSTUS* (MS Total Useful Signal, MST)

MST is sample-based normalization method making the assumption that the number of increased and decreased metabolic signals is relatively equivalent[21, 22]. Using MST, each metabolite concentration is divided by the sum of the concentrations for all measured metabolites in a given sample[5]. However, the validity of this hypothesis is questionable since an increase in the concentration of one metabolite

may not necessarily be accompanied by the decrease in that of another metabolite[22, 23]. MST is typical used to normalize NMR-based metabolomic data[24] and MS-based metabolomics data[11].

$$\tilde{x}_{ij} = \frac{x_{ij}}{\sum_{i=1}^{I} x_{ij}}$$

9. *PQN* (Probabilistic Quotient Normalization, PQN)

PQN is one sample-based normalization method based on the assumption that biologically interesting changes in concentration influence only parts of the NMR spectrum, while dilution effects will affect all metabolite signals[25]. PQN transforms the metabolomics spectra according to an overall estimation on the most probable dilution[25]. There are three main steps in the procedure of PQN[2]: (**1**) performing an integral normalization of each spectrum, then select a reference spectrum such as median spectrum; (**2**) calculating the quotient between a given test spectrum and the reference spectrum, then estimating the median of all quotients for each variable; (**3**) dividing all variables of test spectrum with a median quotient. PQN is reported to be a robust method to account for dilution of complex biological mixtures for NMR and MS metabolomics analysis[25, 26]. The reference spectrum used in PQN was defined as:

$$x_{ref,i} = \frac{1}{J}\sum_{j=1}^{J} x_{ij}$$

10. *Quantile* (Quantile Normalization, QUA)

As one sample-based normalization method, the goal of QUA is to achieve the same distributions of metabolic feature intensities among all samples[14], and the quantile-quantile plot is applied to visualize the distribution similarity[2]. QUA is motivated by the idea that the distribution of two data vectors is the same if the quantile-quantile plot is a straight diagonal line, while a common and non-data driven distribution is generated[14]. It has been adopted for high density oligonucleotide array data based on variances[14], improving NMR-based metabolomics analysis[2] and reducing non-biological systematic variation for MS-based metabolomics data[27]. QUA was defined as:

$$\tilde{x}_{ij} = \frac{1}{J}\sum_{l=1}^{J} x_{k_l,l}$$

11. *Total Sum* (SUM)

SUM is a sample-based method relying on the self-averaging property, which normalizes the dataset by calculating and using the sum of the squares of the data[19]. After normalization, the sum of squares of all variables in a sample equals one[19, 28]. A sample-specific constant assigns an appropriate weight

to each sample, and it aims to minimize possible differences in concentration between samples[28]. So far, SUM has been widely used to correct the dataset for metabolomics studies[29].

$$\tilde{x}_{ij} = x_{ij} - \sum_{i=1}^{I} {x_{ij}}^2$$

B. ***Metabolite-based Normalization Methods*** (5 methods in total)

1. ***Auto Scaling*** **(Unit Variance Scaling, AUT)**

AUT is one of the metabolite-based normalization methods to adjust metabolic variance and use the standard deviation as a sole scaling factor[2, 30]. It scales all metabolites to the unit variance and all the metabolites are regarded to be equally important and are comparably scaled[31]. The standard deviation of all metabolites can become 1 after the normalization[2]. The disadvantage of AUT resides in that the analytical errors may be amplified due to the dilution effects during normalization[13]. It has been used to facilitate bladder cancer diagnosis based on the gas sensor array[32] and to identify urinary nucleoside markers from urogenital cancer patients[33]. AUT is defined as following, where $\bar{x}_i$ and s_i represent the mean and standard deviation of certain metabolite.

$$\tilde{x}_{ij} = \frac{x_{ij} - \bar{x}_i}{s_i}$$

2. ***Level Scaling*** **(LEV)**

LEV is metabolite-based normalization method, transforms metabolic signal variation relative to the average metabolic signal. So the resulting values are the changing values in percentages compared to the mean concentration[34]. LEV is suitable for the circumstances when huge relative variations are of great interest and identification of biomarkers focusing on relative response[34]. The weakness of this method is inflation of the measurement errors[34]. LEV has been applied to identify urinary nucleoside markers from urogenital cancer patients in metabolomic analysis[33]. Equation is defined as following, where $\bar{x}_i$ represents the average metabolic signal for certain metabolite in all samples.

$$\tilde{x}_{ij} = \frac{x_{ij} - \bar{x}_i}{\bar{x}_i}$$

3. ***Pareto Scaling*** **(PAR)**

PAR is a metabolite-based normalization method regarding the square root of the standard deviation of certain metabolite as the scaling factor[35]. It is capable of reducing the weight of a large fold change in metabolite signals, but the dominant weight of extremely large fold change may still be unchanged[2].

The disadvantage of PAR is the sensitivity to large fold changes[34]. PAR has been used to reduce the mask effect from the abundant metabolite for metabolomics dataset[36]. The equation of PAR is defined as following, where $\bar{x}_i$ and s_i represent the mean and standard deviation of certain metabolite.

$$\tilde{x}_{ij} = \frac{x_{ij} - \bar{x}_i}{\sqrt{s_i}}$$

4. *Range Scaling* (RAN)

RAN is a metabolite-based normalization method. Using RAN, the measured intensity is divided by the range of those intensities over all samples[37]. The biological range (difference between the minimal and maximal concentration of a certain metabolite) is regarded as the scaling factor[34]. The advantage of RAN is that the relative concentration for each metabolite is generated after removing instrumental response factor[37]. The property of RAN is that the variation level for metabolites are treated equally[37]. But the disadvantage is the sensitivity to outlier because there are only 2 values being used to estimate the biological range[34]. RAN has been widely used in metabolomics studies[37]. RAN is defined as:

$$\tilde{x}_{ij} = \frac{x_{ij} - \bar{x}_i}{x_{i_{max}} - x_{i_{min}}}$$

5. *Vast Scaling* (Variable Stability Scaling, VAS)

VAS is a metabolite-based normalization method, which weights each variable according to a metric of its stability[38]. VAS mainly uses the standard deviation to normalize and focuses on stable variables which do not show significantly strong variation, while coefficient of variation is regarded as scaling factors[34]. VAS is not appropriate when normalizing large induced variation without group structure[34]. It has been applied to enhance the multivariate models which are built for classification and biomarker identification in metabolomics analysis[31, 38]. The equation is defined as following, where $\bar{x}_i$ and s_i represent the mean and standard deviation of certain metabolite.

$$\tilde{x}_{ij} = \frac{\left(x_{ij} - \bar{x}_i\right)}{s_i} \times \frac{\bar{x}_i}{s_i}$$

C. *Sample & Metabolite-based Normalization Methods* (1 method in total)

1. *VSN* (Variance Stabilization Normalization, VSN)

VSN is non-linear method and a sample and metabolite-based normalization method, aiming to keep the variance constant over the entire data range[2, 39]. It approaches the logarithm for large values and uses the inverse hyperbolic sine to remove heteroscedasticity[2]. For small intensities, it performs linear transformation behavior to keep variances unchanged[2]. VSN is originally developed for normalizing

single and two-channel microarray data[40], while currently it also has been used to determine metabolic profiles of liver tissue during early cancer development[41]. The equation is defined, where a_j and b_j are determined using a robust maximum likelihood estimator such that the variance is constant.

$$\tilde{x}_{ij} = \operatorname{arsinh}\left(a_j + b_j x_{ij}\right),\ \operatorname{arsinh}(t) = \log\left(t + \sqrt{t^2 + 1}\right)$$

Supplementary Method S2. Multiple Criteria for Assessing Normalization Performance

Criterion (*Ca*): method's ability to reduce the intragroup variations among samples in each group[42]

The normalization performance of the studied methods is evaluated using intragroup variation among samples. Low intragroup variation means high similarity of samples and reproducibility of analysis[12, 42]. The *pooled median absolute deviation* (PMAD) is adopted as the measure of intragroup variability in this study. PMAD represents variability among samples and helps in the selection of normalization methods by the low intragroup variability. The lower value of PMAD indicates better normalization performance by removing the experimentally induced noise.

Criterion (*Cb*): method's consistency in discovering metabolic markers from different datasets[43]

Consistency score is used to quantitatively measure the consistency of the metabolic markers among different datasets[43]. Consistency score are performed according to the three steps: several sub-datasets are generated by the random sampling of the whole dataset. Then, all metabolite are ranked according to *q*-values (the fold change would be considered when *q*-values of different metabolites are the same). So, there are several lists of differential metabolites for sub-datasets. Finally, the consistency score is calculated using these differential metabolites in each sub-dataset according to the equation as follows:

$$S = \sum_{i=2}^{C} \sum_{S \in I_i} 2^{i-2} \cdot n_S$$

where C is the number of the sub-datasets, I_i indicates a set of significant metabolites containing the intersections of any i sub-datasets, and n_S refers to the number of metabolites in a intersection S. Generally, a normalization method is more robust if more metabolic markers are shared by more sub-datasets with a higher consistency score.

Criterion (*Cc*): method's classification capacity based on the identified markers[18, 31, 44]

Area under the curve (AUC) value of receiver operating characteristic (ROC) curve based on support vector machine (SVM) are provided in this study[45]. The classification capacity of metabolic markers using normalization method is higher if AUC value is higher. Classification capacity is performed by the following steps: Firstly, differential metabolic markers are identified by the *partial least squares discriminant analysis*. Then, SVM classifier is constructed based on the identified differential markers. After *k*-folds cross validation, a method with larger area under ROC curve and higher AUC value is recognized as better performed one.

References:

1. Astrand, M. Contrast normalization of oligonucleotide arrays. *J. Comput. Biol.* **10**, 95-102 (2003).
2. Kohl, S.M. et al. State-of-the art data normalization methods improve NMR-based metabolomic analysis. *Metabolomics* **8**, 146-160 (2012).
3. Shi, X. et al. Metabolomic analysis of the effects of polychlorinated biphenyls in nonalcoholic fatty liver disease. *J. Proteome Res.* **11**, 3805-3815 (2012).
4. Workman, C. et al. A new non-linear normalization method for reducing variability in DNA microarray experiments. *Genome Biol.* **3**, research0048 (2002).
5. Saccenti, E. Correlation Patterns in Experimental Data Are Affected by Normalization Procedures: Consequences for Data Analysis and Network Inference. *Journal of proteome research* **16**, 619-634 (2017).
6. Contrepois, K., Jiang, L. & Snyder, M. Optimized Analytical Procedures for the Untargeted Metabolomic Profiling of Human Urine and Plasma by Combining Hydrophilic Interaction (HILIC) and Reverse-Phase Liquid Chromatography (RPLC)-Mass Spectrometry. *Molecular & cellular proteomics : MCP* **14**, 1684-1695 (2015).
7. Dudoit, S., Yang, Y.H., Callow, M.J. & Speed, T.P. Statistical methods for identifying differentially expressed genes in replicated cDNA microarray experiments. *Stat Sinica* **12**, 111-139 (2002).
8. Cleveland, W.S. & Devlin, S.J. Locally weighted regression: an approach to regression analysis by local fitting. *Journal of the American statistical association* **83**, 596-610 (1988).
9. Ballman, K.V., Grill, D.E., Oberg, A.L. & Therneau, T.M. Faster cyclic loess: normalizing RNA arrays via linear models. *Bioinformatics* **20**, 2778-2786 (2004).
10. Ejigu, B.A. et al. Evaluation of normalization methods to pave the way towards large-scale LC-MS-based metabolomics profiling experiments. *Omics : a journal of integrative biology* **17**, 473-485 (2013).
11. Karpievitch, Y.V., Nikolic, S.B., Wilson, R., Sharman, J.E. & Edwards, L.M. Metabolomics data normalization with EigenMS. *PloS one* **9**, e116221 (2014).
12. Valikangas, T., Suomi, T. & Elo, L.L. A systematic evaluation of normalization methods in quantitative label-free proteomics. *Briefings in bioinformatics* (2016).
13. Karpievitch, Y.V. et al. Normalization of peak intensities in bottom-up MS-based proteomics using singular value decomposition. *Bioinformatics* **25**, 2573-2580 (2009).
14. Bolstad, B.M., Irizarry, R.A., Astrand, M. & Speed, T.P. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19**, 185-193 (2003).
15. Yuan, Y. et al. Metabolomic analyses of banana during postharvest senescence by ^{1}H-high resolution-NMR. *Food Chem.* **218**, 406-412 (2017).
16. Backshall, A., Sharma, R., Clarke, S.J. & Keun, H.C. Pharmacometabonomic profiling as a predictor of

toxicity in patients with inoperable colorectal cancer treated with capecitabine. *Clin. Cancer Res.* **17**, 3019-3028 (2011).

17. Andjelkovic, V. & Thompson, R. Changes in gene expression in maize kernel in response to water and salt stress. *Plant cell reports* **25**, 71-79 (2006).

18. De Livera, A.M. et al. Normalizing and integrating metabolomics data. *Analytical chemistry* **84**, 10768-10776 (2012).

19. De Livera, A.M., Olshansky, M. & Speed, T.P. Statistical analysis of metabolomics data. *Metabolomics Tools for Natural Product Discovery: Methods and Protocols*, 291-307 (2013).

20. Ting, L. et al. Normalization and statistical analysis of quantitative proteomics data generated by metabolic labeling. *Molecular & cellular proteomics : MCP* **8**, 2227-2242 (2009).

21. Warrack, B.M. et al. Normalization strategies for metabonomic analysis of urine samples. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **877**, 547-552 (2009).

22. Jacob, C.C., Dervilly-Pinel, G., Biancotto, G. & Le Bizec, B. Evaluation of specific gravity as normalization strategy for cattle urinary metabolome analysis. *Metabolomics* **10**, 627-637 (2014).

23. Chen, Y. et al. Combination of injection volume calibration by creatinine and MS signals' normalization to overcome urine variability in LC-MS-based metabolomics studies. *Anal. Chem.* **85**, 7659-7665 (2013).

24. Craig, A., Cloarec, O., Holmes, E., Nicholson, J.K. & Lindon, J.C. Scaling and normalization effects in NMR spectroscopic metabonomic data sets. *Analytical chemistry* **78**, 2262-2267 (2006).

25. Dieterle, F., Ross, A., Schlotterbeck, G. & Senn, H. Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in ^{1}H NMR metabonomics. *Anal. Chem.* **78**, 4281-4290 (2006).

26. Kirwan, J.A., Weber, R.J., Broadhurst, D.I. & Viant, M.R. Direct infusion mass spectrometry metabolomics dataset: a benchmark for data processing and quality control. *Scientific data* **1**, 140012 (2014).

27. Lee, J. et al. Quantile normalization approach for liquid chromatography-mass spectrometry-based metabolomic data from healthy human volunteers. *Analytical sciences : the international journal of the Japan Society for Analytical Chemistry* **28**, 801-805 (2012).

28. De Livera, A.M. et al. Statistical methods for handling unwanted variation in metabolomics data. *Analytical chemistry* **87**, 3606-3615 (2015).

29. Vogl, F.C. et al. Evaluation of dilution and normalization strategies to correct for urinary output in HPLC-HRTOFMS metabolomics. *Analytical and bioanalytical chemistry* **408**, 8483-8493 (2016).

30. Hu, C.X. & Xu, G.W. Mass-spectrometry-based metabolomics analysis for foodomics. *Trac-Trend Anal. Chem.* **52**, 36-46 (2013).

31. Gromski, P.S., Xu, Y., Hollywood, K.A., Turner, M.L. & Goodacre, R. The influence of scaling metabolomics data on model classification accuracy. *Metabolomics* **11**, 684-695 (2015).

32. Weber, C.M. et al. Evaluation of a gas sensor array and pattern recognition for the identification of bladder cancer from urine headspace. *Analyst* **136**, 359-364 (2011).

33. Struck, W. et al. Liquid chromatography tandem mass spectrometry study of urinary nucleosides as potential cancer markers. *J. Chromatogr. A* **1283**, 122-131 (2013).

34. van den Berg, R.A., Hoefsloot, H.C.J., Westerhuis, J.A., Smilde, A.K. & van der Werf, M.J. Centering, scaling, and transformations: improving the biological information content of metabolomics data. *Bmc Genomics* **7**, 142 (2006).

35. Eriksson, L. et al. Using chemometrics for navigating in the large data sets of genomics, proteomics, and metabonomics (gpm). *Anal. Bioanal. Chem.* **380**, 419-429 (2004).

36. Yang, J., Zhao, X., Lu, X., Lin, X. & Xu, G. A data preprocessing strategy for metabolomics to reduce the mask effect in data analysis. *Front. Mol. Biosci.* **2**, 4 (2015).

37. Smilde, A.K., van der Werf, M.J., Bijlsma, S., van der Werff-van der Vat, B.J. & Jellema, R.H. Fusion of mass spectrometry-based metabolomics data. *Anal. Chem.* **77**, 6729-6736 (2005).

38. Keun, H.C. et al. Improved analysis of multivariate data by variable stability scaling: application to NMR-based metabolic profiling. *Anal Chim Acta* **490**, 265-276 (2003).

39. Huber, W., von Heydebreck, A., Sultmann, H., Poustka, A. & Vingron, M. Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics* **18 Suppl 1**, S96-104 (2002).

40. Kultima, K. et al. Development and evaluation of normalization methods for label-free relative quantification of endogenous peptides. *Mol. Cell Proteomics* **8**, 2285-2295 (2009).

41. Ibarra, R. et al. Metabolomic analysis of liver tissue from the VX2 rabbit model of secondary liver tumors. *HPB Surg.* **2014**, 310372 (2014).

42. Chawade, A., Alexandersson, E. & Levander, F. Normalyzer: a tool for rapid evaluation of normalization methods for omics data sets. *Journal of proteome research* **13**, 3114-3120 (2014).

43. Luan, H. et al. Non-targeted metabolomics and lipidomics LC-MS data from maternal plasma of 180 healthy pregnant women. *GigaScience* **4**, 16 (2015).

44. Risso, D., Ngai, J., Speed, T.P. & Dudoit, S. Normalization of RNA-seq data using factor analysis of control genes or samples. *Nature biotechnology* **32**, 896-902 (2014).

45. Zhou, X., Oshlack, A. & Robinson, M.D. miRNA-Seq normalization comparisons need improvement. *Rna* **19**, 733-734 (2013).