Supporting Information

Optimization and application of direct infusion nanoelectrospray HRMS method for large-scale urinary metabolic phenotyping in molecular epidemiology

Elena Chekmeneva [a*], *Gonçalo dos Santos Correia* [a], *Queenie Chan* [b,c], *Anisha Wijeyesekera*[a], *Adrienne Tin*[e], *Jeffery Hunter Young*[e], *Paul Elliott* [b,c,d], *Jeremy K. Nicholson* [a,d], *and Elaine Holmes* [a,d*]

[a] Biomolecular Medicine, Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, South Kensington, London SW7 2AZ, United Kingdom

[b] Department of Epidemiology and Biostatistics, School of Public Health, Faculty of Medicine, Imperial College London, London W2 1PG, United Kingdom

[c] MRC-PHE Centre for Environment and Health, Imperial College London, London W2 1PG, United Kingdom

[d] MRC-NIHR National Phenome Centre, London SW7 2AZ, United Kingdom

[e] Department of Medicine, Johns Hopkins University, School of Medicine, Baltimore, Maryland 21287, United States

*** Corresponding authors:** ***Dr Elena Chekmeneva (email: e.chekmeneva@imperial.ac.uk) and Prof Elaine Holmes (e-mail: elaine.holmes@imperial.ac.uk)***

Table of content:

Materials and methods

1. Materials and preparation of standard solutions

All chemicals and solvents were of the highest purity grade. Labelled and non-labelled standards used for quantification (See Table S1 for details) were purchased from Sigma-Aldrich (Isotec, Gillingham, UK), CK Isotope (Cambridge Isotope Laboratories, Ibstock, UK), ALSA CHIM (Illkirch, France), Biozol (Eching, Germany), TCI-UK (Oxford, UK), Alfa Aesar, TOCRIS (Bristol, UK) , and Acros Organics. Ultrapure HPLC-grade methanol and water were purchased from Sigma-Aldrich (Gillingham, UK).

The standard stock solutions were prepared taking into account all specifications of chemical purity and isotopic enrichment for labelled standards. For the most part, the labelled standard solutions were additionally prepared in deuterated methanol and assessed by ^{1}H NMR for signs of chemical degradation with a time interval of two months corresponding to the time required to complete the analysis of >10000 urine samples from the INTERMAP study. ^{1}H NMR spectra of labelled glutamic acid–d5 and proline betaine-^{13}C2 are shown in the Figure S-2. For the deuterated standards showing unstable behaviour fresh stock solutions were prepared weekly. Generally, the use of ^{13}C and ^{15}N isotope-labelled standards is preferential to deuterated standards in targeted analysis because the latter can interfere with the signal of targeted compound.[1] The stock solutions of the labelled and non-labelled standards were stored at 4 °C in tightly closed containers.

For the preparation of analyte-free matrix in the INTERMAP study, sodium borate solution in ultrapure HPLC-grade water at concentration of 10 g/L (as in the standard procedure for urine sample collection and storage in the INTERMAP project) was diluted 50 times to obtain similar borate concentration as expected in diluted urine samples. For the ARIC study, water with 0.01% of formic acid was used as an analyte-free matrix.

2. Chip based nanoelectrospray MS system

The nanoelectrospray was created and maintained by applying 1.4 kV high voltage and 0.8 psi nitrogen flow controlled by ChipSoft software (version 8.3.1). The sample plate temperature was maintained at 4 °C. The data were collected in high resolution continuum mode with the scan time of 1 s over the mass range of 40 – 600 m/z in negative and positive ion modes with automatic polarity switch infusing 5 μL of a sample. The sampling cone voltage was set at 40 V, and the source offset at 80 V.

Total data acquisition time was of 40 s for the INTERMAP study and 30 s for the ARIC study for each ionisation mode (first negative, and then positive ion mode) but the overall turnaround time for each sample was of ca. 2 minutes to let the instrument automatically

switch the polarity and the voltage settle before acquiring the data in the next polarity. The data for the negative and positive ionisation mode were acquired in two separate files in the MassLynx™ software. The total time for the analysis of a 96 well-plate was less than 4 hours.

Sodium formate solution was used to calibrate the mass spectrometer on a daily basis. The lock-mass function was turned off but data in both modes were recalibrated post-acquisition by in-house software using reference signals of endogenous metabolites present in all urine samples.

MS/MS was performed for targeted peaks in DIMS in Resolution mode in non-spiked pooled urine sample and pooled urine sample spiked with the standards. The optimal collision energy (CE) was selected for each peak between 10-60 eV. For the endogenous metabolites the spectra were compared to the spectra acquired for the standards in neat solvent and to the metabolite fragmentation patterns available in online databases (HMDB[2] and Metlin[3]).

3. Synthetic Test Mixture (TM)

The TM consisted of glutamine, glutamic acid, creatinine, cytidine, citric acid, leucine, phenylalanine, tryptophan, hippuric acid, benzoic acid, and octanoic acid prepared at 25-100 μM concentration in ultrapure water.

For the analysis, the TM was diluted 10 times with ultrapure water. An aliquot of 50 μL was placed in a well and 100 μL of MeOH containing 0.015% of formic acid (final concentration of formic acid in a well was of 0.01%) were added to maintain 1:2 water-methanol proportion required for a stable nESI signal. The intra-day replicate analysis (n=10) of the TM was done in three separate days (total n=30), and the CV% of signal intensities was assessed intra- and inter-day without and with normalisation (to the sum of the intensities of targeted peaks). The TM was also analysed in two separate days (n=8) to estimate the effect of instrument maintenance (mass spectrometer venting after the nitrogen gas supply shutdown event with subsequent detector setup and chip-alignment procedure in the ChipSoft software) on the signal stability.

Results and discussion

1. Parallel assay – standard addition vs. external calibration curves

In this work, the MS quantification was performed by the back-calculation of the ratio of each selected metabolite to its internal standard and the slope of the calibration curve obtained for that metabolite in the pooled urine sample. In order to show the reliability of our method, we compared the method of standard additions using the pooled urine sample as a matrix with the method of external calibration in the analyte-free matrix. For the INTERMAP study, the solution of boric acid of the same concentration as expected in the urinary specimens was prepared and used as an analyte-free matrix. The accuracy of metabolite quantification by both approaches was tested using the validation QC samples prepared at three different levels of concentrations defined for each metabolite from the corresponding linear range (Table 1). Since the QC samples prepared in analyte-free matrix contained only the known spiked concentration of metabolites, without any background level of metabolites encountered in endogenous matrix, they can be used as a standard to prove the reliability of the method of standard additions. Their quantification was performed by using the slope of the standard addition calibration curve in pooled urine sample and external calibration curve in an analyte-free matrix. The comparison of the slopes of calibration curves obtained by both approaches in the INTERMAP study along with the values of relative error (RE%) of measured concentration for low and medium level QC samples prepared in analyte-free matrix and quantified using standard addition and external calibration curves yielded results corresponding to the adapted acceptance criteria (Table S-2). A similar assessment of matrix effect was undertaken for the ARIC pooled SR samples and analyte-free matrix consisting of water with 0.01% formic acid (Table S-3).

2. Analysis of Test Mixture

The TM was analysed over three days to assess intra- (n=10) and inter-day (n=30) variability of signal intensities. Figure S-9(A) demonstrates the inter-day CV% values for the intensities of compounds measured in negative ion mode and the effect of signal normalisation to the sum of intensities of measured compounds. When non-normalised the intensity varied by as much as 25-40% which could be expected in the DIMS experiment taking into account the low concentration of compounds in the TM (μM diluted 10 times). However, after normalisation to the sum of intensities the CV% value reduced to 15-20 with exception of citric acid (CV% = 26). The TM was used to assess the effect of nitrogen gas shutdown event

and subsequent venting of mass spectrometer followed by detector set-up and chip-alignment procedure for the nESI source. Figure S-9(B) shows the normalised intensity values for the n=8 measurements done before and after the instrument maintenance. The CV% values for all the analytes measured in positive and negative ion mode did not exceed 20 which is an acceptable metabolite precision providing the normalisation was done for targeted analytes.

Supplementary Tables

Table S-1. List of metabolites for quantification, their internal standards and the corresponding m/z values

Metabolite	Adduct	m/z	Internal standrad
Hydroxycinnamic acid	M-H	163.0401	m-Coumaric acid-1,2,3-13C3
Acetylcarnitine	M+H	204.1230	Acetyl-d3-L-carnitine
Arginine	M+H	175.1195	L-Arginine-13C6 hydrochloride
Ascorbic acid	M-H	175.0248	L-Ascorbic acid-1-13C
Benzoic Acid	M-H	121.0295	Benzoic acid-2,3,4,5,6-d5
Caffeic acid	M-H	179.0350	Caffeic acid-13C9
Carnitine	M+H	162.1125	DL-Carnitine-(trimethyl-d9)
Cholic acid	M-H	407.2803	Cholic acid-2,2,4,4-d4
Citric acid	M-H	191.0197	Citric acid-2,2,4,4-d4
Citrulline	M+Na	198.0855	L-Citrulline-4,4,5,5-d4
Cotinine	M+H	177.1022	(±)-Cotinine-(methyl-d3)
Creatine	M+H	132.0773	Creatine-d3 (methyl-d3)
Creatinine	M+H	114.0667	Creatinine-d3 (methyl-d3)
Daidzein	M-H	253.0506	Daidzein-d4 (4-hydroxyphenyl-2,3,5,6-d4)
Deoxycholic acid	M-H	391.2854	Deoxycholic acid-2,2,4,4-d4
Genistein	M-H	269.0455	Genistein-d4 (4-hydroxyphenyl-2,3,5,6-d4)
Glutamic acid	M-H	146.0459	DL-Glutamic-2,3,3,4,4-d5 acid
Glycocholic acid	M-H	464.3018	Glycocholic-2,2,4,4-d4 acid
Glycodeoxycholic acid	M-H	448.3068	Glycodeoxycholic-2,2,4,4-d4 acid
Hippuric acid	M-H	178.0511	N-Benzoyl-d5-glycine
Homovanillic acid Homovanillic acid fragment	M-H	181.0501 122.0360	(4-Hydroxy-3-methoxyphenyl-d3)acetic-α,α-d2 acid
Hydroxybenzoic acid	M-H	137.0239	4-Hydroxybenzoic-2,3,5,6-d4 acid
Indoxyl sulfate	M-H	212.0023	Indoxyl-2,4,5,6,7-d5 sulfate
Isovalerylglycine	M-H	158.0823	N-Isovalerylglycine-2,2-d2
2-oxoglutaric acid	M-H	145.0142	Succinic acid-2,2,3,3-d4
Ketoleucine	M-H	129.0557	Ketoleucine-1-13C sodium salt
Kynurenine	M+Na	231.0746	L-Kynurenine sulfate:H2O (ring-d4, 3,3-d2)
Leucine	M-H	130.0874	L-Leucine-1-13C
Malic acid	M-H	133.0137	(S)-(-)-Malic-2,3,3-d3 acid
Methylsuccinic acid	M-H	131.0344	(±)-2-Methyl-d3-succinic-2,3,3-d3 acid
N-acetylneuraminic acid	M-H	308.0982	N-Acetyl-D-[2,3-13C2]neuraminic acid
Nicotinamide	M+H	123.0558	Nicotinamide-2,4,5,6-d4
Nicotine	M+H	163.1230	DL-Nicotine (methyl-d3, 98%)
Nicotinic acid	M+H	124.0399	Nicotinic-d4 acid
Nα-Acetyl-L-ornithine	M+Na	197.0902	Nα-Acetyl-L-ornithine-5,5-d2
Ornithine	M+H	133.0977	L-Ornithine-2,3,3,4,4,5,5-d7 HCl
Phenylacetic acid	M+H	137.0603	Phenylacetic acid-1,2-13C2
Phenylacetylglutamine	M-H	263.1037	Nα-(Phenyl-d5-acetyl)-L-glutamine
Phenylalanine	M-H	164.0712	DL-Phenylalanine-1-13C
Proline betaine	M+H; M+Na	144.1025; 166.0838	Stachydrine-(dimethyl-13C2)

Propionylcarnitine	M+H	218.1392	Propionyl-L-carnitine-(N-methyl-d3)
Saccharin	M-H	181.9917	Saccharin-d4 (ring-d4)
Succinic acid	M-H	117.0193	Succinic acid-2,2,3,3-d4
Taurocholic acid	M-H	514.2844	Taurocholic acid-2,2,4,4-d4 acid
Tryptophan	M-H	203.0821	L-Tryptophan-d5
Tyramine	M+H	138.0892	2-(4-Hydroxyphenyl)ethyl-1,1,2,2-d4-amine
Tyrosine	M-H	180.0661	L-Tyrosine-3,3-d2

Table S-2. Slope values of calibration curves obtained by the addition of authentic standards into pooled SR sample from the INTERMAP study and in the analyte-free matrix; the values of relative error (RE%) in measured concentration for the QC samples prepared in analyte-free matrix and quantified using both calibration curves.

Metabolite INTERMAP study	Slope in analyte-free matrix	Slope of the standard addition curve	RE(%) of QC_{low} using analyte-free matrix curve	RE(%) of QC_{med} using analyte-free matrix curve	RE(%) of QC_{low} using standard addition curve	RE(%) of QC_{med} using standard addition curve
Hydroxycinnamic acid	0.33	0.36	2	4	9	5
Acetylcarnitine	0.96	0.97	10	2	17	10
Ascorbic acid	0.41	0.41	6	14	>30	9
Benzoic Acid	0.07	0.04	12	7	>30	>30
Caffeic acid	1.33	1.41	4	1	5	6
Carnitine	0.65	0.66	1	2	22	3
Cholic acid	0.76	0.85	11	1	10	12
Citric acid	0.04	0.04	7	5	24	12
Cotinine	0.70	0.75	1	5	10	3
Creatine	0.38	0.41	>30	5	27	17
Creatinine	0.86	0.63	20	9	14	18
Daidzein	1.87	1.89	3	1	23	7
Deoxycholic acid	0.30	0.36	7	2	3	13
Genistein	1.49	1.75	9	2	23	20
Glutamic acid	0.36	0.33	3	1	12	11
Glycocholic acid	0.65	0.74	9	1	18	15
Glycodeoxycholic acid	0.67	0.73	8	1	16	12
Hippuric acid	1.10	1.14	6	4	23	14

Homovanillic acid	0.03	0.03	>30	1	>30	17
Indoxyl sulfate	2.15	2.14	15	8	22	17
Isovalerylglycine	0.79	0.78	2	1	2	1
2-oxoglutaric acid	0.11	0.19	>30	7	7	16
Ketoleucine	0.10	0.11	0	9	30	8
Kynurenine	0.41	0.37	8	6	30	21
N-acetylneuraminic acid	0.38	0.35	8	9	18	19
Nicotine	0.31	0.34	12	1	18	11
Nicotinic acid	0.06	0.07	22	5	18	5
Phenylacetic acid	0.16	0.17	22	3	24	18
Phenylacetylglutamine	2.71	3.10	23	5	>30	27
Proline betaine	0.99	0.75	6	10	1	13
Propionylcarnitine	1.05	1.06	11	3	22	12
Saccharin	1.38	1.65	8	1	28	21
Succinic acid	0.28	0.24	10	1	8	19
Taurocholic acid	0.52	0.62	11	4	27	25
Tyramine	0.04	0.05	12	1	1	15
Vanillylmandelic acid	0.72	0.84	11	2	>30	11

Table S-3. Slope values of calibration curves obtained by the addition of authentic standards in pooled urine sample from the ARIC study and in the analyte-free matrix. The table also compares concentration of metabolites in total pooled SR sample calculated form the standard addition curve compared to the expected concentration levels (from the literature) and LOQ in µg/mL estimated from the calibration curve obtained in analyte-free matrix.

Metabolite ARIC study	Slope of the standard addition curve	Slope of the curve in analyte-free matrix	Calculated concentration in a pooled SR (µmol/mmol Creatinine*)	Expected normal concentration, µmol/mmol Creatinine	LOQ determined in analyte-free matrix (H_2O), nmol/mL
*Creatinine, mM	0.04	0.05	5.52	10 ±2 mM	37.89
Benzoic Acid	0.03	0.03	114.94	0.1-150	1.51
Citric acid	0.06	0.02	232.06	46.9- 600	3.15
Glutamic acid	0.09	0.08	23.84	3.3-18.4	0.89
Hippuric acid	0.06	0.06	371.34	19-933	1.14
Hydroxybenzoic acid	0.14	0.14	52.70	0.7-29	0.27
Indoxyl sulfate	0.40	0.35	9.00	6.0–64.8	0.28
2-oxoglutaric acid	0.04	0.02	49.06	< 150	0.38
Ketoleucine	0.08	0.06	35.56	0.02-0.5	0.24
Leucine	0.12	0.11	227.19	1.2-19.1	1.57
Malic acid	0.77	0.61	6.04	0.7-5.3	n.a.
Methylsuccinic acid	0.07	0.07	14.17	0.4-10.8	0.31
N-acetylneuraminic acid	0.18	0.17	12.19	2.5-8.6	0.19
Nicotinic acid	0.19	0.17	17.48	0.1-0.8	0.18
Phenylalanine	0.08	0.08	13.43	1.6-18.2	0.51
Succinic acid	0.09	0.08	35.53	0.3-33.3	1.84
Tryptophan	0.15	0.14	8.44	1.3-29.4	0.52
Tyrosine	0.08	0.08	27.63	3-38.7	0.83
Acetylcarnitine	2.34	2.92	4.74	0.4-7.5	1.08
Acetyl-L-ornithine	0.11	0.12	19.92	0.5-2.8	1.68
Arginine	0.87	0.89	1.95	0-23	0.15
Carnitine	0.67	0.67	21.37	0.6–15.2	0.91
Citrulline	0.15	0.16	8.71	0-74.3	1.03
Creatine	0.05	0.05	37.65	10-650	10.54
Nicotinamide	0.16	0.17	21.41	n.a.	1.39
Ornithine	0.11	0.10	49.01	0-22	7.94
Phenylacetic acid	0.06	0.06	27.09	0.3-1.9	2.03

Table S-4. Concentration of metabolites in total pooled SR sample from the INTERMAP study calculated form the standard addition curve compared to the expected concentration levels from the literature[2,4,5] and LOQ in µg/mL estimated from the calibration curve obtained in analyte-free matrix.

Metabolite INTERMAP study	Calculated concentration in a pooled SR, µmol/mmol Creatinine*	Expected normal concentration, µmol/mmol Creatinine	LOQ determined in analyte-free matrix, nmol/mL
***Creatinine mM**	**6.46 mM**	**10 ±2 mM**	20.34
Hydroxycinnamic acid	0.30	0.01-0.17	1.22
Ascorbic acid	55.65	3.8-85.5	1.14
Benzoic Acid	113.27	0.1-150	26.22
Caffeic acid	1.29	0.01-0.38	0.56
Cholic acid	0.94	0.43-0.73	0.24
Citric acid	209.32	46.9- 600	10.94
Daidzein	0.75	0.002–0.064; 2.43 after soya consumption	0.39
Deoxycholic acid	0.49	0.05-0.06	0.25
Genistein	1.29	0.03-0.11; 0.39 after soya consumption	0.37
Glutamic acid	9.57	3.3-18.4	4.08
Glycocholic acid	0.19	0.03±0.057	0.21
Glycodeoxycholic acid	0.61	n.a.	0.22
Hippuric acid	107.34	19-933	10.61
Homovanillic acid	9.98	0.1-10	6.59
Indoxyl sulfate	8.43	6.0–64.8	5.63
Isovalerylglycine	2.79	1.0–10.0	0.63
2-oxoglutaric acid	7.37	0.18-17	2.05
Ketoleucine	38.87	0.02-0.45	0.77
Leucine	49.32	1.2-19.1	n.a
N-acetylneuraminic acid	12.28	2.5-8.6	0.65
Phenylacetylglutamine	29.53	0.5-78	3.03
Saccharin	2.16	n.a.	0.55
Succinic acid	84.77	0.3-33.3	5.08
Taurocholic acid	0.75	n.a.	0.19
Vanillylmandelic acid	4.56	0.6-3	2.52
Acetylcarnitine	1.82	0.4-7.5	0.98
Carnitine	11.13	0.6–15.2	1.24
Cotinine	0.43	*ca*. 0.37-3.0 (active smoker)	0.57
Creatine	127.50	10-650	17.55
Hexoses	587.42	37–501	n.a.
Kynurenine	4.81	1.1–2.5	1.92
Nicotine	0.44	*ca.* 7-9 (active smoker)	0.62

Nicotinic acid	6.10	0.1-0.82	0.29
Phenylacetic acid	0.39	0.3-1.9	2.21
Proline betaine	48.75	0.8-23.7	4.19
Propionylcarnitine	1.48	0.01-0.20	0.92
Tyramine	14.29	0.23-0.78	3.65

Table S-5. Concentration of metabolites in mM calculated by standard addition method for eight pooled SR samples (24 hrs collection) obtained for four populations (USA, UK, PRC, and Japan) and two visits (first – F and repeat – R) from the INTERMAP study and pooled spot urine SR sample from the ARIC study.

Metabolite/Concentration in pooled urine sample, mM	USA-F	USA-R	UK-F	UK-R	PRC-F	PRC-R	J-F	J-R	ARIC
Creatinine	6.27	6.52	8.46	6.79	4.81	7.82	7.56	6.59	8.71
Hydroxycinnamic acid	0.02	0.02	0.03	0.02	0.04	0.03	0.03	0.03	
Ascorbic acid	0.51	0.43	0.49	0.34	0.13	0.24	0.22	0.25	
Benzoic Acid	0.79	1.01	0.81	0.79	0.80	1.35	1.05	0.70	1.52
Citric acid	1.90	1.76	2.65	2.27	1.13	1.96	1.90	1.59	1.14
Daidzein	*n.d.*	*n.d.*	*n.d.*	*n.d.*	*n.d.*	*n.d.*	0.001	0.001	
Glutamic acid	0.14	0.15	0.16	0.14	0.13	0.20	0.19	0.19	0.19
Hippuric acid	0.73	1.08	1.54	1.34	0.47	1.10	0.66	0.51	4.41
Indoxyl sulfate	0.11	0.13	0.11	0.05	0.06	0.06	0.21	0.03	0.05
Isovalerylglycine	0.02	0.02	0.03	0.02	0.03	0.04	0.03	0.04	
2-oxoglutaric acid	0.30	0.25	0.42	0.40	0.32	0.42	0.39	0.31	0.57
Ketoleucine	0.18	0.19	0.22	0.21	0.19	0.26	0.26	0.19	0.06
Leucine	0.45	0.42	0.43	0.40	0.47	0.60	0.58	0.51	0.17
N-acetyl neuraminic acid	0.09	0.09	0.11	0.09	0.07	0.11	0.09	0.07	0.10
Phenylacetylglutamine	0.36	0.35	0.53	0.42	0.20	0.43	0.39	0.27	
Saccharin	0.04	0.04	0.05	0.05	0.01	0.02	0.02	0.02	
Succinic acid	0.41	0.38	0.53	0.36	0.42	0.42	0.37	0.34	0.28
Vanillylmandelic acid	0.03	0.03	0.06	0.03	0.03	0.02	0.02	0.03	
Acetylcarnitine	0.02	0.03	0.04	0.02	0.02	0.02	0.03	0.03	0.04
Carnitine	0.13	0.10	0.13	0.11	0.09	0.10	0.13	0.10	0.15
Cotinine	0.003	0.004	0.004	0.003	0.006	0.007	0.006	0.005	
Creatine	0.96	1.02	1.05	1.07	0.85	0.93	1.67	1.32	0.38
Nicotine	0.007	0.005	0.008	0.003	0.007	0.008	0.008	0.007	
Nicotinic acid	0.05	0.06	0.11	0.06	0.07	0.09	0.10	0.07	0.21
Proline betaine	0.29	0.33	0.26	0.19	0.21	0.23	0.23	0.19	
Tyramine	0.21	0.24	0.22	0.19	0.16	0.22	0.23	0.22	

n.d – < LLOQ

Table S-6. Intra- and inter-day accuracy and precision values for several metabolites measured in positive and negative ion modes in validation QC samples from the USA-F and Japan-F populations from the INTERMAP study.

Metabolite INTERMAP study USA-F and Japan-F populations		Accuracy						Precision, CV%					
		QC high intra-day	QC high inter-day	QC med intra-day	QC med inter-day	QC low intra-day	QC low inter-day	QC high intra-day	QC high inter-day	QC med intra-day	QC med inter-day	QC low intra-day	QC low inter-day
Hydroxycinnamic acid	USA	90	94	119	118	103	105	4	10	6	6	6	9
	Japan	99	98	94	94	98	96	1	3	5	5	10	9
Acetylcarnitine	USA	94	98	118	118	94	105	4	11	4	7	5	17
	Japan	115	113	99	97	109	106	8	9	13	12	15	15
Ascorbic acid	USA	80	76	91	101	81	71	9	22	10	11	>30	>30
	Japan	81	85	118	115	124	105	5	17	9	26	17	25
Benzoic Acid	USA	102	99	117	114	125	121	13	13	13	19	13	19
	Japan	109	108	110	104	120	107	2	6	8	11	14	27
Caffeic acid	USA	107	105	111	112	128	125	2	24	1	28	1	29
	Japan	114	110	97	95	115	109	2	23	6	22	8	25
Carnitine	USA	115	110	104	107	125	117	9	10	10	9	15	17
	Japan	104	103	84	85	83	81	7	6	10	11	15	15
Cholic acid	USA	104	106	89	95	97	96	4	2	1	13	2	3
	Japan	104	99	94	94	92	90	4	5	3	3	12	11
Citric acid	USA	101	107	130	124	122	120	1	14	6	27	28	26
	Japan	119	115	70	70	70	70	4	6	29	27	27	23
Cotinine	USA	101	102	108	109	105	105	1	3	1	10	2	9
	Japan	108	107	96	95	99	96	1	5	6	6	12	13
Creatine	USA	111	110	108	105	88	92	3	4	4	16	9	29
	Japan	104	104	86	87	94	85	7	6	12	11	>30	>30
Creatinine	USA	98	99	104	104	80	89	3	5	3	17	12	25

	Japan	106	106	106	104	130	130	5	4	7	7	7	>30
Daidzein	USA	103	103	112	114	122	124	3	8	2	7	2	15
	Japan	103	101	91	90	106	102	2	4	7	7	10	11
Deoxycholic acid	USA	108	107	89	90	95	101	7	3	2	10	1	2
	Japan	110	108	99	99	101	97	5	6	10	9	23	22
Genistein	USA	105	103	111	113	128	130	2	9	2	8	1	12
	Japan	110	108	94	93	107	107	2	5	4	5	6	8
Glutamic acid	USA	104	103	106	112	112	114	3	8	3	8	5	11
	Japan	98	97	89	91	90	89	3	5	4	5	6	6
Glycocholic acid	USA	95	97	79	88	87	89	7	2	4	20	3	4
	Japan	110	108	97	97	107	107	1	4	4	5	5	9
Glycodeoxycholic acid	USA	99	98	90	97	117	111	3	5	1	14	2	8
	Japan	111	108	102	104	112	106	1	5	14	16	18	19
Hippuric acid	USA	105	105	109	111	93	111	3	9	3	11	3	25
	Japan	119	116	89	90	98	97	8	8	9	9	17	19
Homovanillic acid	USA	106	107	97	99	87	93	11	4	17	15	4	6
	Japan	80	81	87	55	89	87	6	11	11	15	15	16
Indoxyl sulfate	USA	105	104	117	115	109	105	8	11	5	7	2	11
	Japan	106	102	103	101	112	104	9	11	11	11	11	18
Isovalerylglycine	USA	106	104	104	111	122	122	7	10	6	12	2	17
	Japan	99	101	81	86	80	86	6	7	6	10	11	16
2-oxoglutaric acid	USA	91	99	75	82	>130	115	5	11	19	22	>30	>30
	Japan	88	91	<70	<70	84	81	9	16	>30	>30	>30	>30
Ketoleucine	USA	91	94	83	92	72	74	4	6	7	18	28	5
	Japan	90	88	97	94	101	95	4	10	11	12	21	>30
Kynurenine	USA	102	102	116	117	78	72	4	4	4	6	4	10
	Japan	98	98	103	101	97	93	7	7	9	8	18	18
Leucine	USA	103	105	83	87	99	92	3	4	5	11	7	9
	Japan	106	102	96	92	<70	<70	4	8	12	15	>30	>30
N-acetylneuraminic acid	USA	108	104	107	109	87	84	4	10	4	18	11	15

	Japan	113	112	112	108	>130	123	4	10	4	7	6	15
Nicotine	USA	102	103	114	113	101	110	2	4	2	8	2	18
	Japan	110	106	96	95	93	89	1	7	5	6	10	13
Nicotinic acid	USA	107	106	108	110	106	97	7	6	9	16	13	19
	Japan	90	91	88	88	97	89	8	7	2	9	27	25
Phenylacetic acid	USA	110	110	103	103	130	127	4	6	3	4	2	17
	Japan	105	105	87	87	101	99	4	5	6	7	9	8
Phenylacetylglutamine	USA	107	105	109	108	105	108	3	10	4	13	3	15
	Japan	126	122	97	97	130	126	3	7	9	8	18	23
Proline betaine	USA	104	104	104	106	114	106	5	5	5	16	3	29
	Japan	104	101	96	98	117	111	4	8	7	9	5	13
Propionylcarnitine	USA	99	102	117	115	103	108	2	17	4	20	4	19
	Japan	95	94	95	96	101	106	6	10	7	10	12	16
Saccharin	USA	119	111	122	115	123	128	3	14	2	27	9	24
	Japan	126	119	113	107	>130	>130	8	14	11	16	6	26
Succinic acid	USA	96	96	108	109	101	96	5	12	5	13	8	15
	Japan	103	103	96	97	102	98	3	6	8	7	10	12
Taurocholic acid	USA	102	103	75	87	<70	<70	2	3	4	25	9	9
	Japan	104	102	96	96	84	105	3	4	5	6	27	25
Tyramine	USA	102	105	106	108	109	110	4	15	5	9	11	21
	Japan	123	117	107	103	122	115	4	5	8	9	22	28
Vanillylmandelic acid	USA	108	109	109	111	110	119	6	12	7	8	4	20
	Japan	102	101	87	89	97	96	6	8	4	4	5	7

Table S-7. In-study validation as inter-day relative error (RE%) of measured concentration and precision values for metabolites measured in positive and negative ion modes in validation QC samples from the ARIC study.

Metabolite ARIC study	Accuracy, RE% QC in pooled urine sample			Precision, CV% QC in pooled urine sample		
	QC high inter-day	QC med inter-day	QC low inter-day	QC high inter-day	QC med inter-day	QC low inter-day
Acetyl-L-ornithine	10	15	18	11	16	20
Arginine	8	13	17	8	13	20
Benzoic Acid	9	8	30	11	22	>30
Carnitine	29	14	27	5	10	23
Citrulline	8	14	25	10	15	24
Creatine	6	12	13	5	15	15
Creatinine	11	16	>30	9	21	24
Glutamic acid	8	9	10	7	14	13
Hydroxybenzoic acid	4	7	8	5	7	10
Indoxyl sulfate	16	10	9	5	12	12
2-oxoglutaric acid	9	12	19	12	17	18
Leucine	7	14	11	8	12	9
Malic acid	10	11	17	13	14	25
Methylsuccinic acid	7	19	20	9	17	24
N-acetyl neuraminic acid	8	11	14	7	11	16
Nicotinamide	10	20	24	10	24	>30
Nicotinic acid	10	18	21	11	18	16
Ornithine	9	24	32	10	21	31
Phenylacetic acid	8	16	25	8	19	>30
Phenylalanine	6	11	24	7	13	27
Succinic acid	10	12	18	12	18	23
Tryptophan	11	15	19	11	18	25
Tyrosine	9	9	10	7	9	10

Table S-8. Inter-batch precision values (CV%) obtained for validation QC samples and dilution SR series from the USA-F (14 batches) and Japan-F (10 batches) populations from the INTERMAP study and quantified using the calibration curves from each batch.

Sample type INTERMAP study	**Dilution QC 1:100 (n=11)**		**Dilution QC 1:50 (n=11)**		**Dilution QC 1:30 (n=11)**		**Dilution QC 1:20 (n=11)**		**Validation QC low (n=22)**		**Validation QC medium (n=22)**		**Validation QC high (n=22)**	
	USA (n=14)	Japan (n=10)	USA (n=14)	Japan (n=10)	USA (n=14)	Japan (n=10)	USA (n=14)	Japan (n=10)	USA (n=14)	Japan (n=10)	USA (n=14)	Japan (n=10)	USA (n=14)	Japan (n=10)
Hydroxycinnamic acid	27%	22%	22%	17%	18%	12%	16%	14%	13%	11%	11%	10%	11%	12%
Ascorbic acid	18%	29%	21%	27%	21%	18%	>30%	23%	24%	18%	11%	11%	7%	13%
Benzoic Acid	18%	18%	20%	14%	16%	13%	10%	16%	16%	15%	13%	14%	15%	17%
Caffeic acid	19%	21%	13%	20%	17%	14%	12%	18%	7%	12%	9%	9%	6%	14%
Citric acid	21%	19%	27%	22%	16%	13%	22%	15%	26%	20%	22%	19%	11%	14%
Daidzein	18%	17%	15%	24%	21%	25%	21%	11%	11%	13%	12%	11%	10%	15%
Genistein	19%	n.d.	21%	n.d.	21%	n.d.	18%	n.d.	13%	n.d.	13%	n.d.	9%	n.d.
Glutamic acid	22%	>30%	19%	15%	18%	12%	14%	6%	10%	8%	8%	7%	7%	11%
Hippuric acid	21%	20%	21%	28%	21%	15%	20%	10%	17%	21%	15%	18%	13%	17%
Homovanillic acid	18%	27%	25%	17%	18%	12%	13%	14%	24%	18%	18%	14%	13%	13%
Indoxyl sulfate	26%	>30%	20%	18%	23%	30%	21%	21%	14%	23%	13%	16%	12%	19%
Isovalerylglycine	14%	10%	14%	16%	12%	8%	8%	10%	9%	10%	8%	9%	8%	13%
2-oxoglutaric acid	28%	>30%	27%	>30%	25%	30%	29%	27%	>30%	30%	>30%	22%	22%	22%
Ketoleucine	23%	>30%	25%	26%	14%	22%	11%	23%	24%	27%	28%	23%	12%	16%
Leucine	26%	24%	25%	18%	22%	12%	15%	13%	19%	16%	14%	11%	10%	11%
N-acetylneuraminic acid	19%	11%	21%	14%	20%	10%	15%	11%	15%	12%	11%	9%	9%	13%
Phenylacetylglutamine	21%	20%	18%	>30%	21%	18%	21%	11%	22%	20%	16%	13%	13%	21%
Saccharin	19%	15%	18%	19%	23%	10%	21%	12%	17%	15%	14%	11%	10%	15%
Succinic acid	20%	25%	25%	25%	14%	16%	8%	16%	14%	17%	11%	14%	10%	17%
Vanillylmandelic acid	16%	11%	17%	11%	14%	10%	13%	11%	10%	9%	9%	11%	9%	15%
Acetylcarnitine	17%	20%	17%	25%	17%	17%	12%	15%	18%	20%	24%	16%	10%	18%
Carnitine	8%	18%	15%	23%	13%	17%	11%	14%	13%	20%	23%	13%	9%	15%

Cotinine	23%	16%	19%	21%	21%	15%	17%	10%	7%	14%	19%	10%	5%	14%
Creatine	22%	26%	23%	29%	23%	20%	19%	15%	21%	20%	25%	11%	11%	13%
Creatinine	19%	21%	19%	29%	19%	21%	19%	18%	18%	24%	27%	17%	14%	21%
Kynurenine	n.d.	21%	n.d.	17%	n.d.	19%	n.d.	18%	n.d.	11%	n.d.	9%	n.d.	12%
Nicotine	16%	21%	21%	22%	24%	22%	25%	18%	12%	17%	21%	13%	5%	15%
Nicotinic acid	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	17%	16%	24%	10%	11%	17%
Phenylacetic acid	26%	79%	23%	>30%	27%	20%	42%	20%	14%	13%	22%	12%	6%	15%
Proline betaine	19%	26%	20%	>30%	19%	23%	15%	19%	14%	23%	25%	16%	11%	23%
Propionylcarnitine	19%	22%	20%	21%	19%	16%	15%	15%	14%	22%	21%	18%	11%	17%
Tyramine	n.d.	n.d.	n.d.	n.d.	>30%	>30%	24%	24%	24%	18%	26%	12%	13%	17%

Table S-9. Inter-batch precision values (CV%) obtained for pooled SR and dilution SR series from the ARIC study and quantified using the calibration curves from each batch (n).

ARIC study metabolite	Pooled SR n=30	Dilution SR 1/10 n=12	Dilution SR 1/50 n=12
Benzoic Acid	13%	8%	20%
Citric acid	20%	11%	10%
Glutamic acid	10%	10%	17%
Hippuric acid	11%	10%	20%
Hydroxybenzoic acid	7%	6%	13%
Indoxyl sulfate	8%	7%	13%
2-oxoglutaric acid	12%	9%	19%
Ketoleucine	13%	14%	25%
Leucine	14%	11%	22%
Malic acid	11%	10%	16%
Methylsuccinic acid	13%	10%	19%
N-acetylneuraminic acid	7%	10%	12%
Nicotinic acid	14%	11%	12%
Phenylalanine	10%	9%	15%
Succinic acid	12%	10%	>30%
Tryptophan	8%	8%	11%
Tyrosine	17%	12%	18%
Acetylcarnitine	28%	24%	26%
Arginine	7%	10%	26%
Carnitine	12%	7%	15%
Citrulline	17%	20%	24%
Creatine	9%	15%	12%
Creatinine	9%	12%	14%
Nicotinamide	21%	30%	30%
Ornithine	15%	10%	18%
Phenylacetic acid	10%	12%	14%

Supplementary Figures

Figure S-1. Flowchart representing DI-nESI-HRMS method optimization and validation process

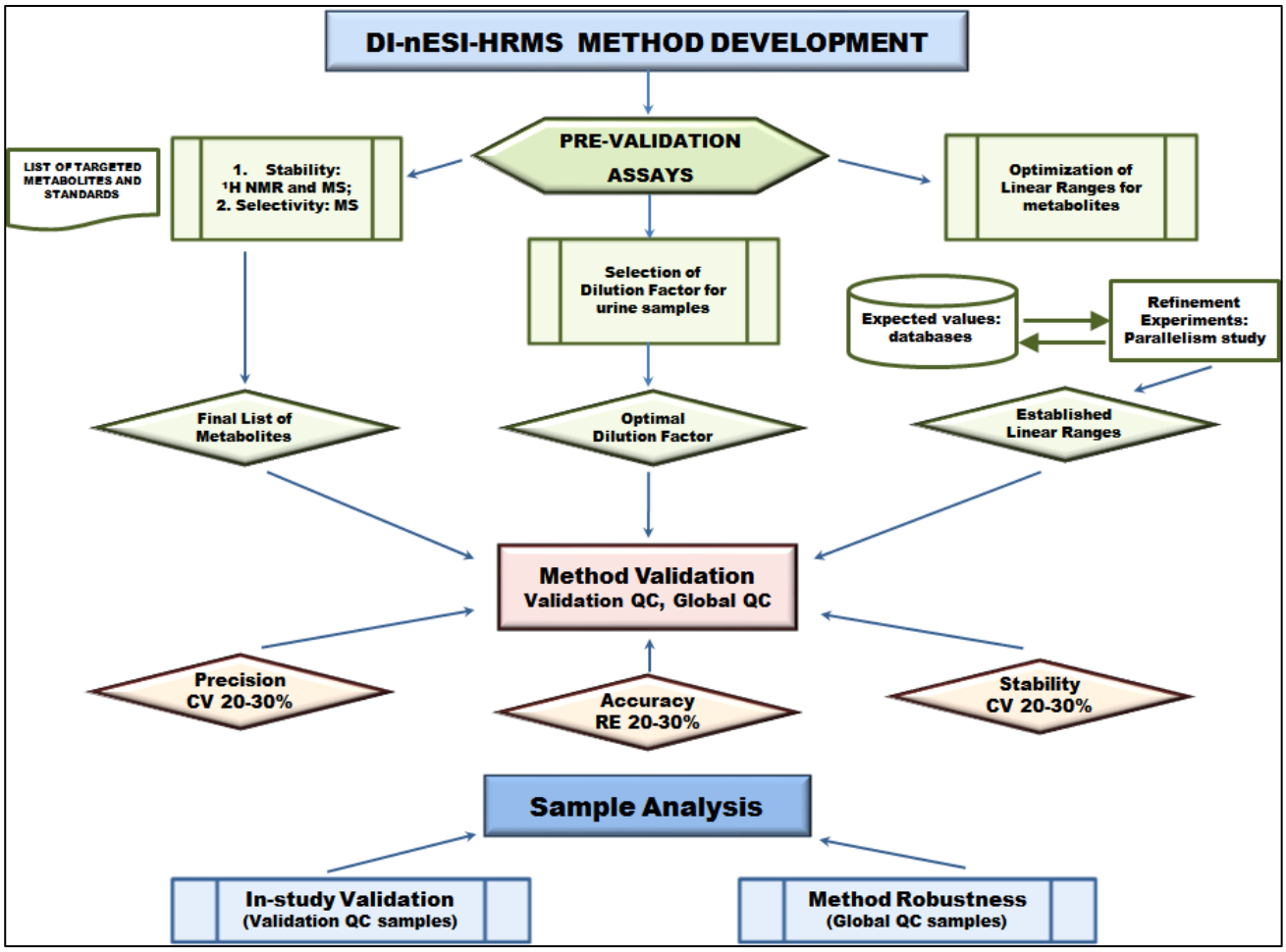


Figure S-2. ^{1}H NMR spectra of labelled glutamic acid–d5 standard (A) and labelled proline betaine-^{13}C2 standard (B). The lower panel in each spectrum shows the first measurement, and the upper panel – the second measurement two months later.

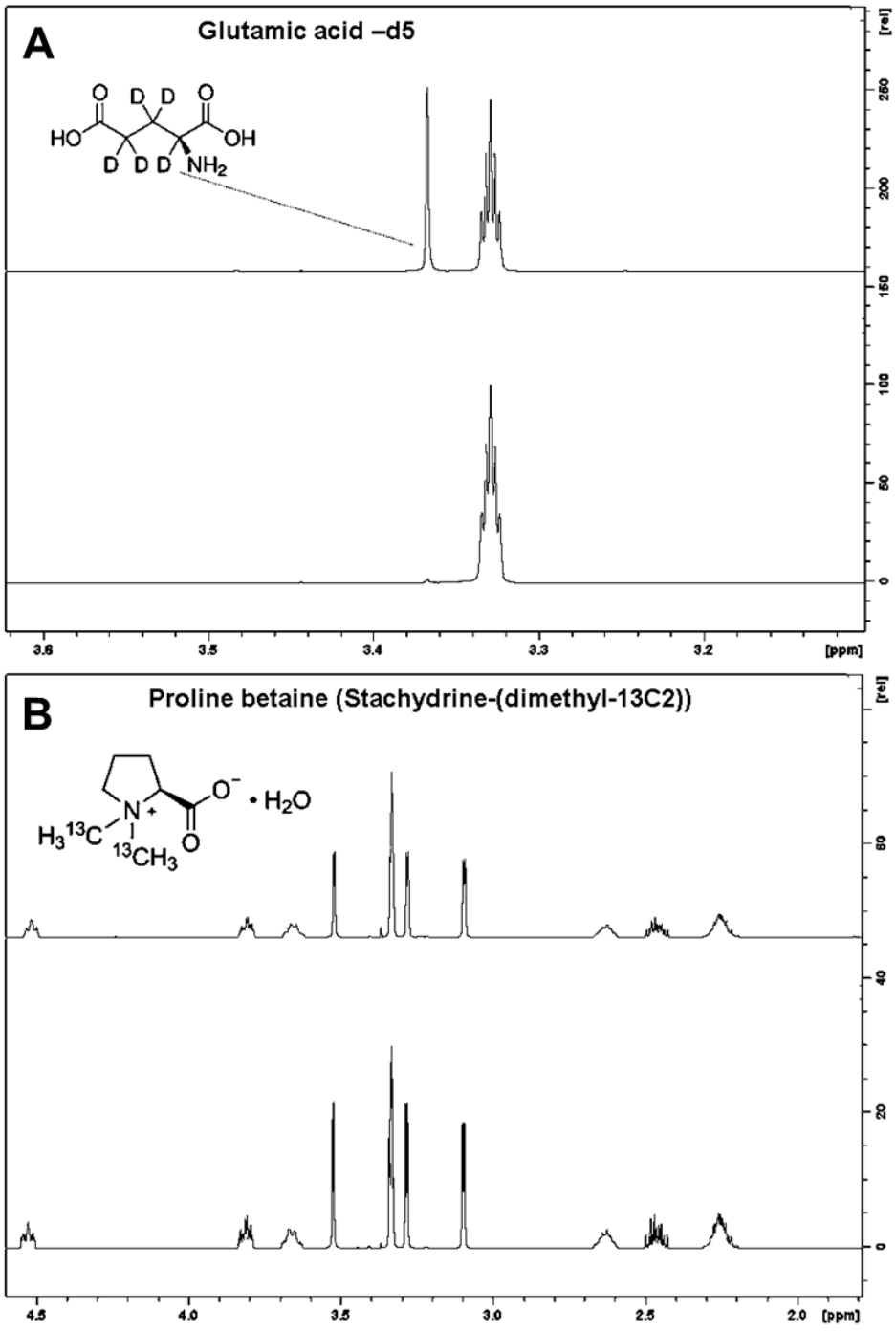
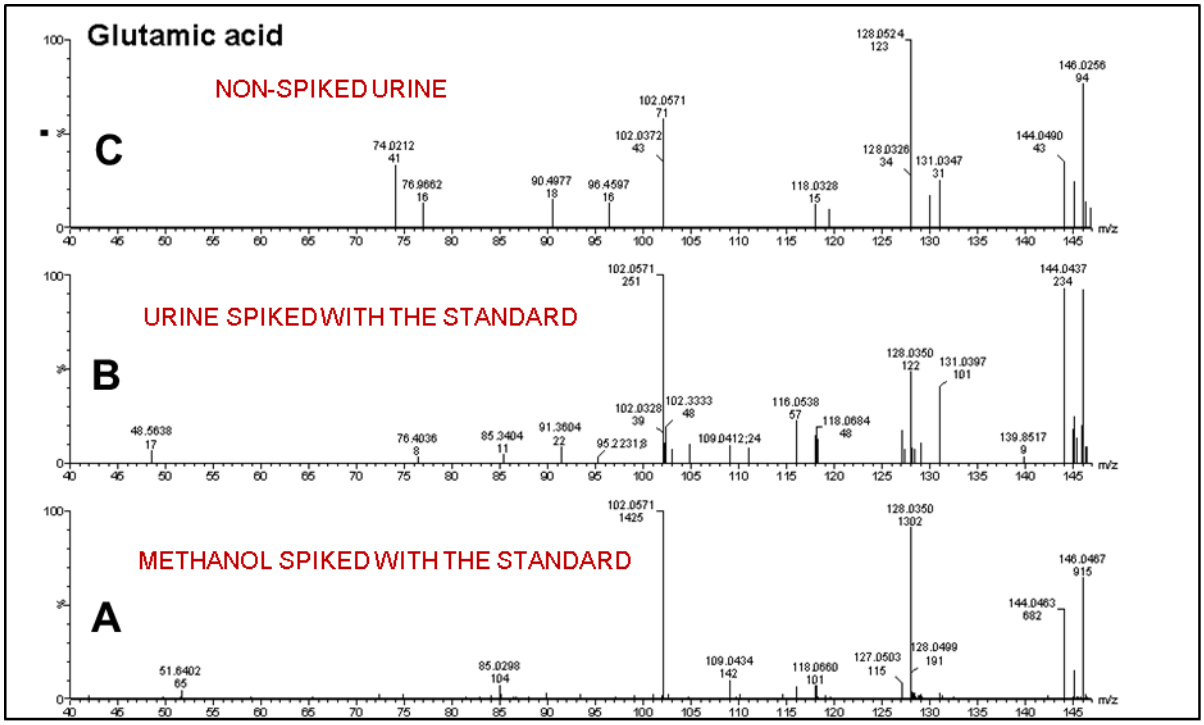


Figure S-3. MS/MS spectra obtained for the standards in neat methanol (A), spiked (B) and non-spiked pooled urine sample (C) from the INTERMAP study for glutamic acid and phenylacetylglutamine. The spectra were obtained by direct infusion in Resolution mode on the Waters Synapt G2-S (Q-ToF).



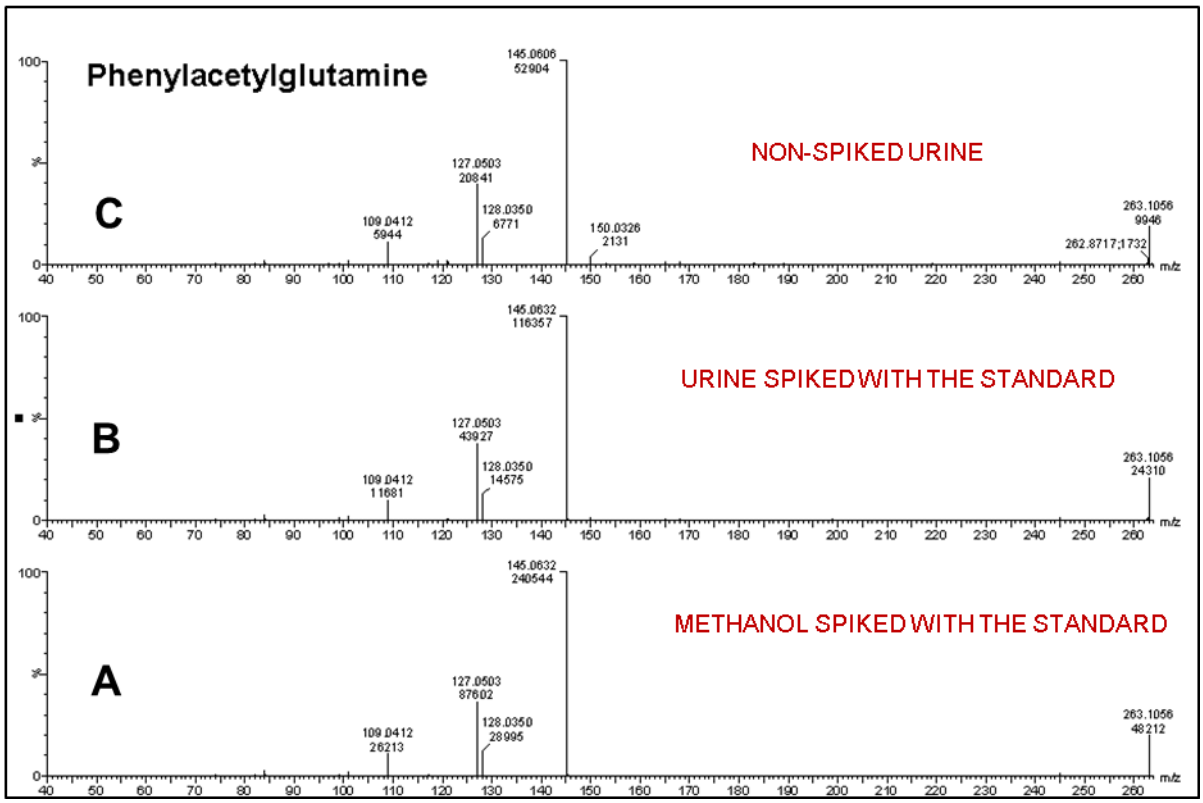


Figure S-4. The ratio of intensities of metabolites and internal standards at dilution 1/20 to dilution 1/50 in the pooled urine samples from the Japan (blue) and USA (red) INTERMAP populations. The ratios of intensities for stable isotope labelled standards are marked with an (*).

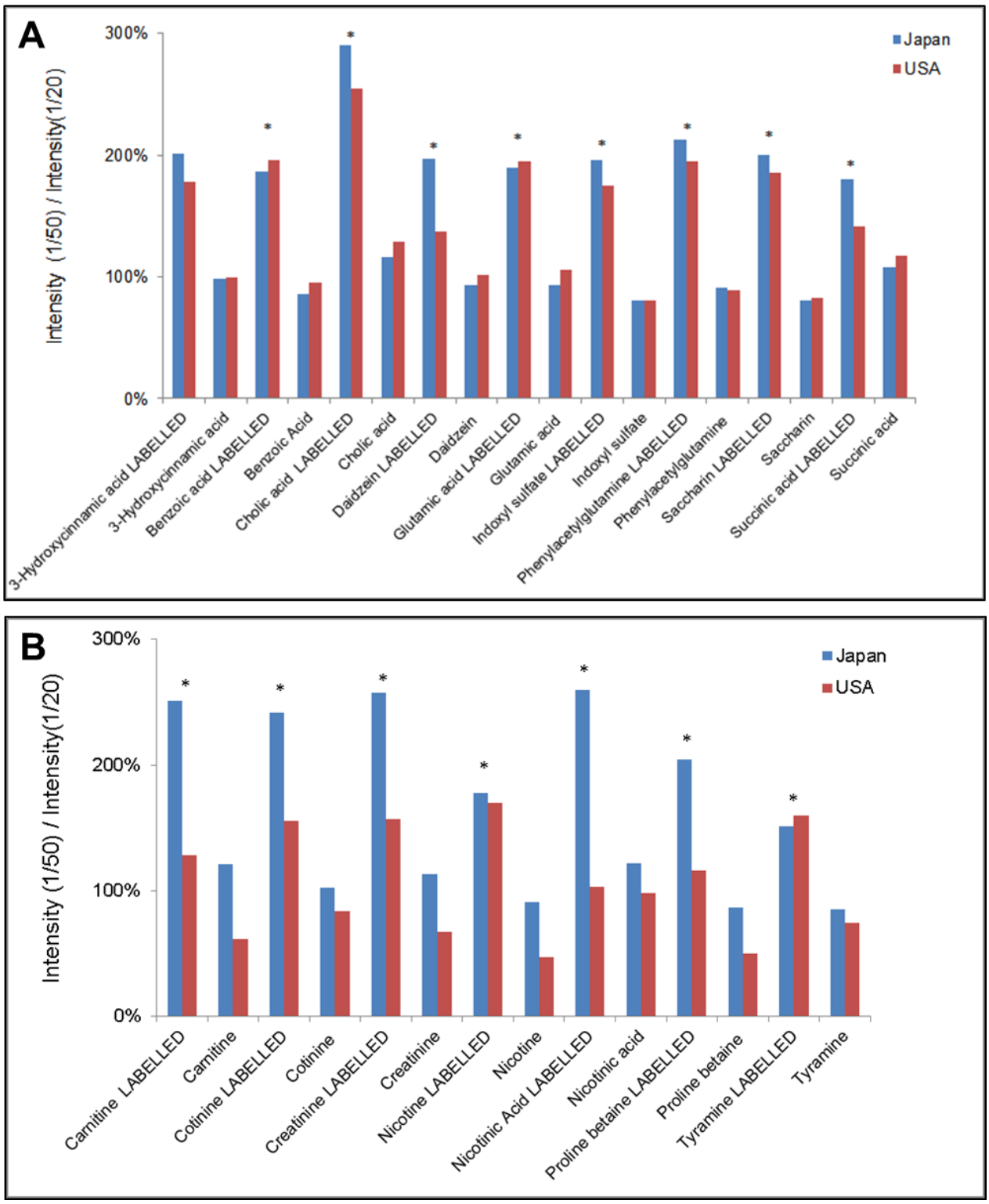


Figure S-5. Calibration curves obtained in pooled urine sample by the method of standard additions (open squares) and in analyte-free matrix (open circles) for cholic acid, hippuric acid, and creatinine. SIL – stable isotope labelled standard.

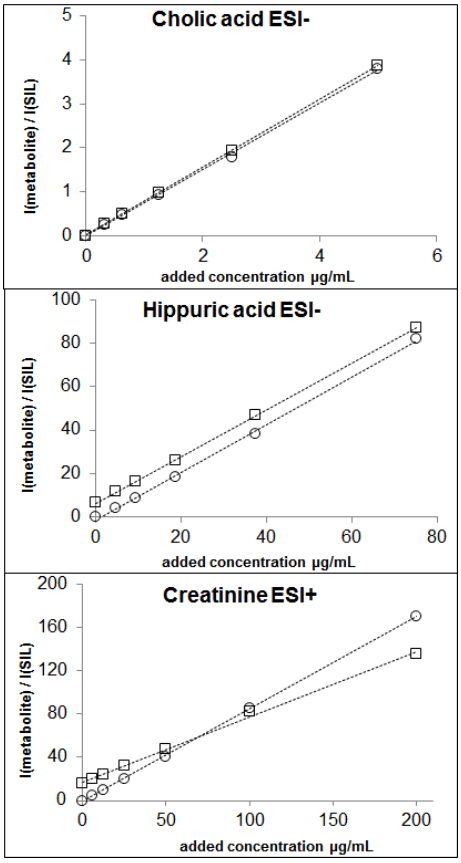


Figure S-6. Precision (as CV %) of slope values and concentration of the QC samples for some metabolites measured in Japan-F pooled urine samples from three validation series prepared and measured to assess the assay stability (freshly prepared, three freeze-thaw cycles, long-term storage at -80°C).

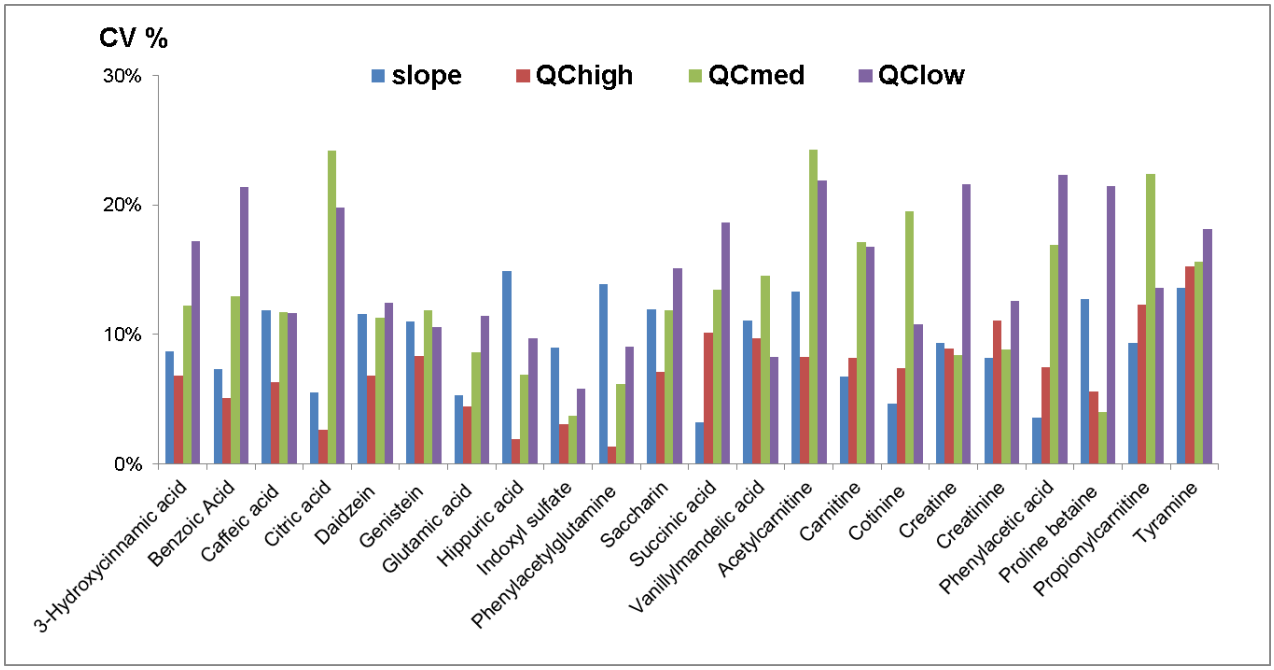


Figure S-7. Concentration values of nicotine (A), acetylcarnitine (B), and creatinine (C) measured in positive ion mode, and of caffeic acid (D), glutamic acid (E), and hippuric acid (F) measured in negative ion mode in the pooled SR samples in 11 batches from the USA-F population of the INTERMAP study. The straight lines represent the mean concentration ± 2 standard deviation.

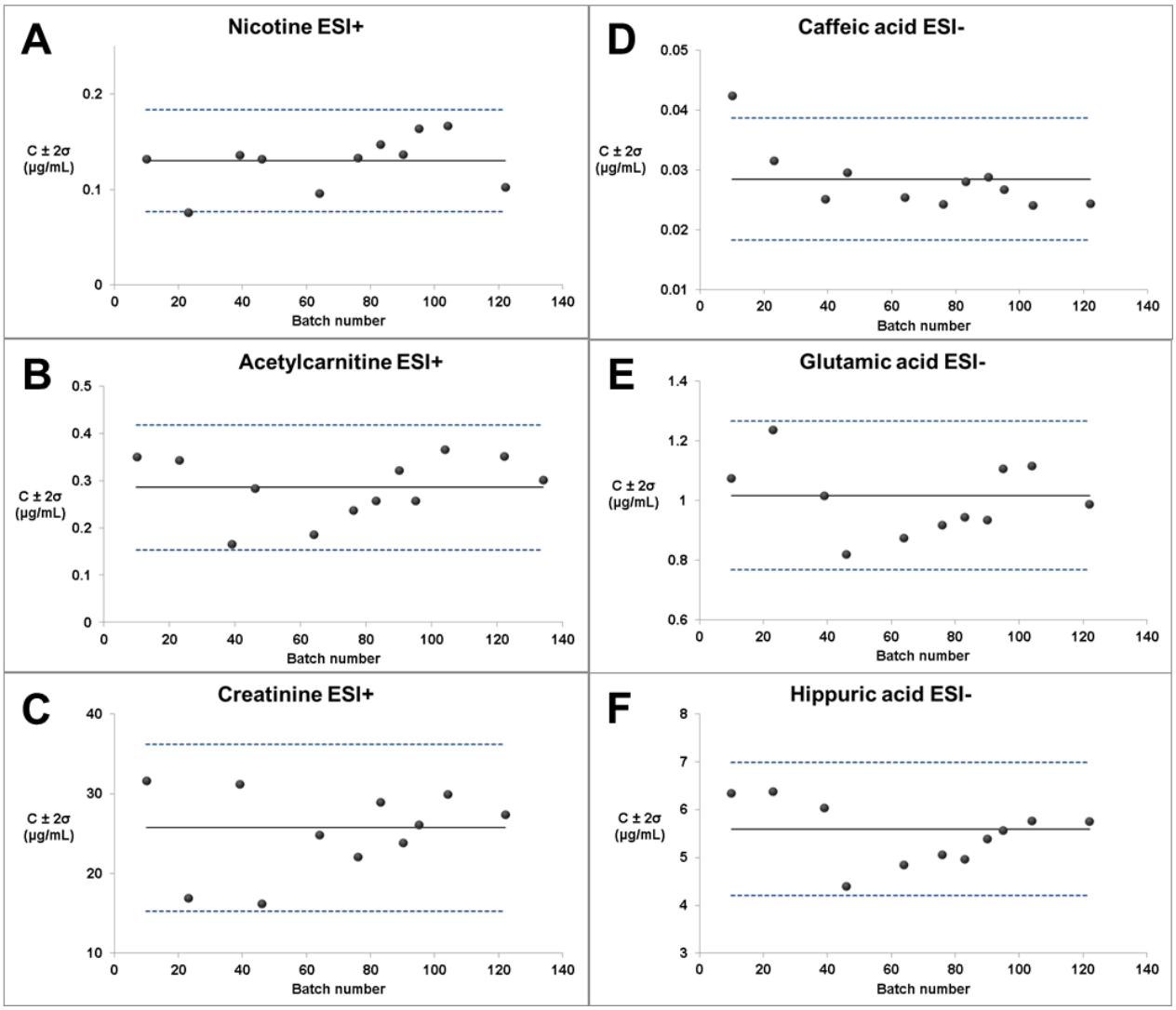
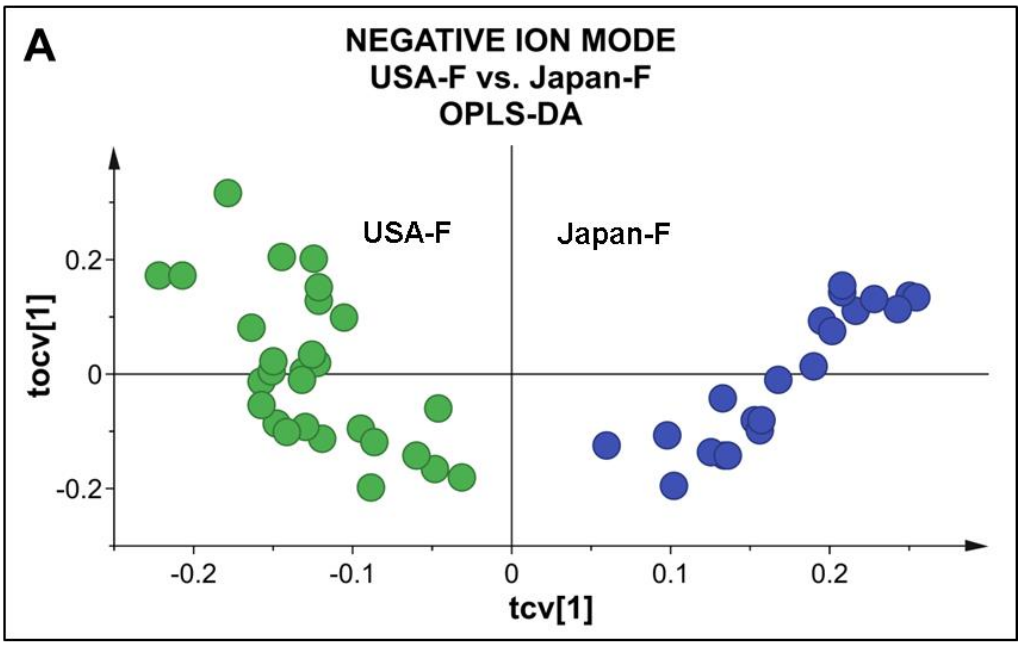
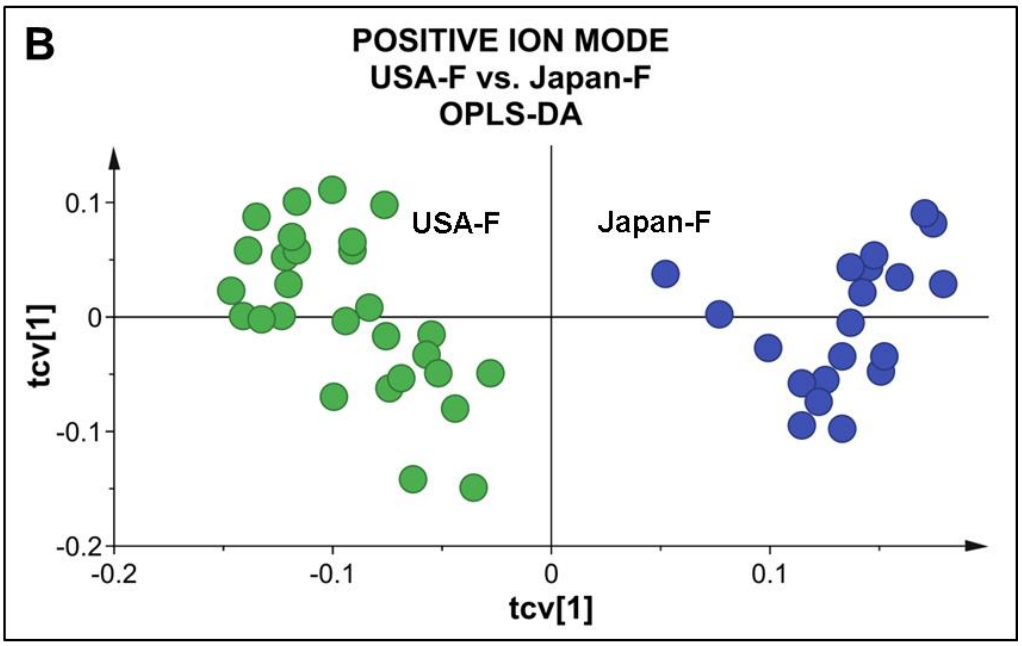


Figure S-8. OPLS-DA cross-validated score plots obtained for the full-scan global profile analysis of the QC samples from the USA-F and Japan-F populations from the INTERMAP study in negative (A) and positive (B) ion modes. The model characteristics are presented in the Figure.



$R^2X=0.611$ $R^2Y=0.933$ $Q^2Y=0.908$



$R^2X=0.365$ $R^2Y=0.966$ $Q^2Y=0.921$

Figure S-9. Test Mixture (TM) analysis: inter-day CV% (n=30) values for the intensities of compounds measured in negative ion mode and the effect of signal normalisation to the sum of intensities (A). CV% of normalised intensity values for the n=8 measurements done before and after the instrument maintenance (B).

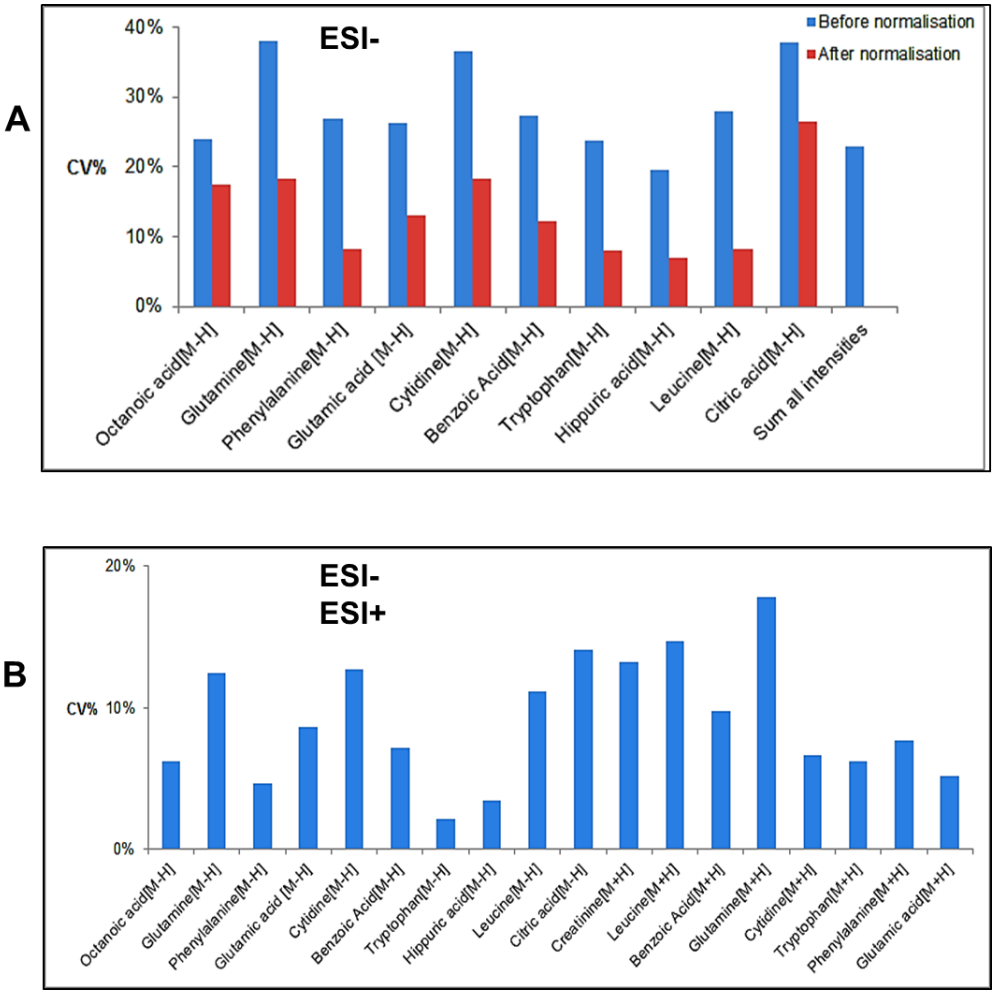


Figure S-10. Box plots showing difference in population (Japan and USA) and gender metabolite concentration levels obtained from DI-nESI-HRMS by the standard addition method.

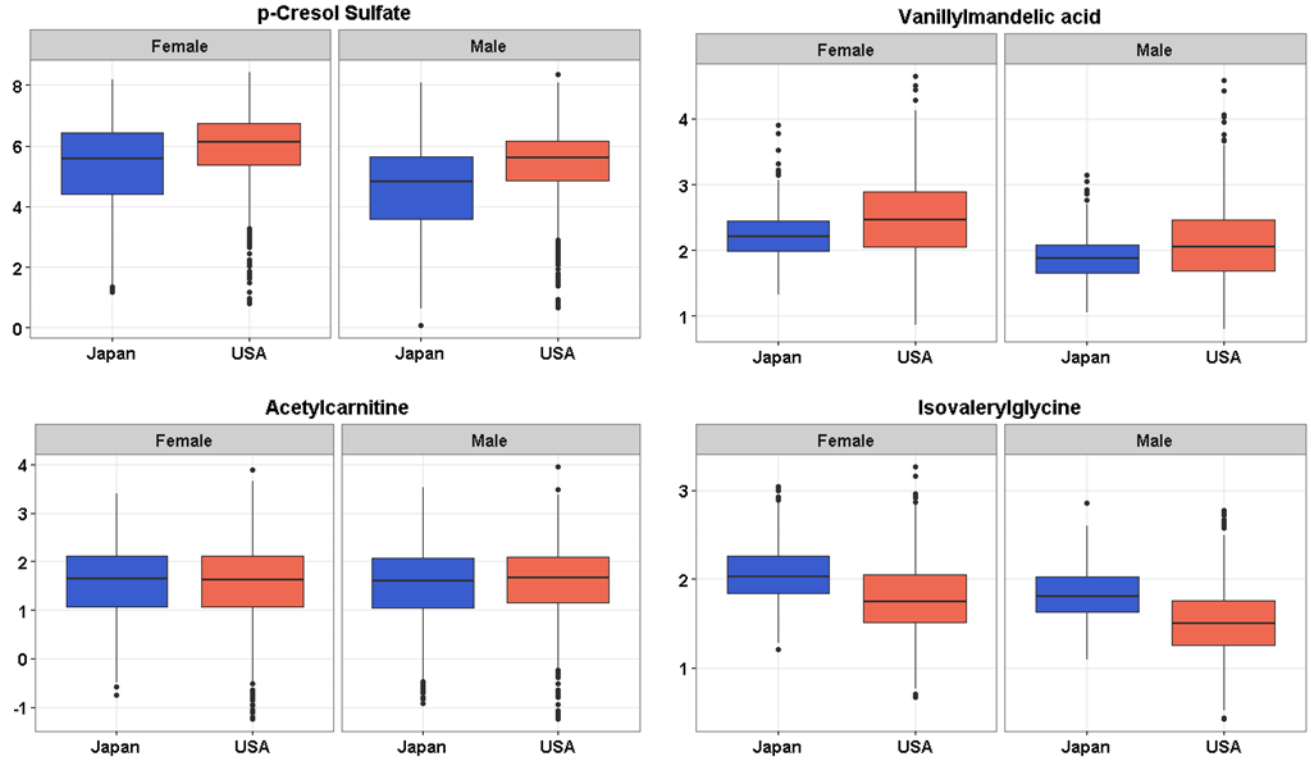


Figure S-11. OPLS-DA loadings S plots comparing features (m/z) from the global profiles of urine samples from the USA and Japanese populations obtained by DI-nESI-HRMS method in positive (A) and negative (B) ionisation modes.

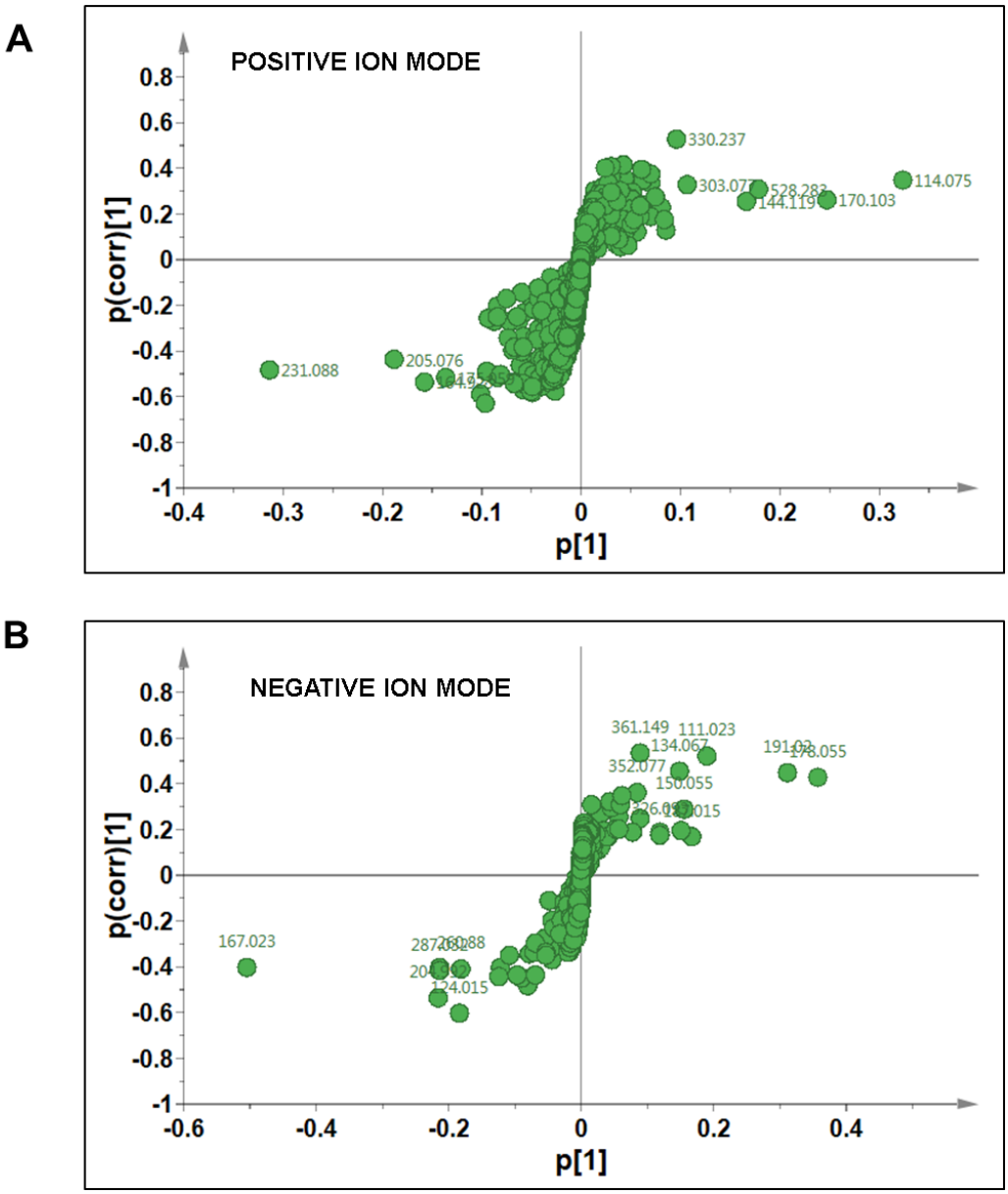
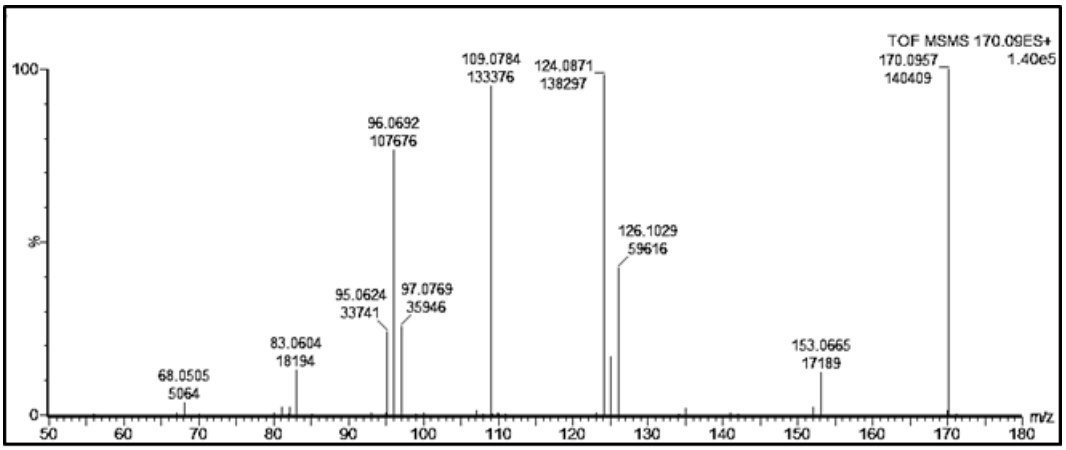
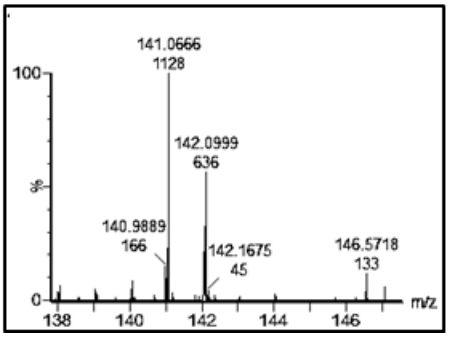


Figure S-12. MS/MS spectrum (CE=15eV) of the features m/z 170.10 detected in positive ion mode. The fragmentation gave rise to the ions at m/z 153.07, m/z 141.07 (zoomed in), m/z 126.10, m/z 124.09, m/z 109.07, m/z 97.08, m/z 96.07, m/z 95.06, and m/z 83.06.





Supplementary References

1. Stokvis, E.; Rosing, H.; Beijnen, J. H., Stable isotopically labeled internal standards in quantitative bioanalysis using liquid chromatography/mass spectrometry: necessity or not? *Rapid Communications in Mass Spectrometry* **2005,** *19* (3), 401-407.
2. Bouatra, S.; Aziat, F.; Mandal, R.; Guo, A. C.; Wilson, M. R.; Knox, C.; Bjorndahl, T. C.; Krishnamurthy, R.; Saleem, F.; Liu, P.; Dame, Z. T.; Poelzer, J.; Huynh, J.; Yallou, F. S.; Psychogios, N.; Dong, E.; Bogumil, R.; Roehring, C.; Wishart, D. S., The Human Urine Metabolome. *Plos One* **2013,** *8* (9), e73076, 1-28.
3. Smith, C. A.; O'Maille, G.; Want, E. J.; Qin, C.; Trauger, S. A.; Brandon, T. R.; Custodio, D. E.; Abagyan, R.; Siuzdak, G., METLIN - A metabolite mass spectral database. *Ther Drug Monit* **2005,** *27* (6), 747-751.
4. Berman, T.; Goldsmith, R.; Goen, T.; Spungen, J.; Novack, L.; Levine, H.; Amitai, Y.; Shohat, T.; Grotto, I., Urinary concentrations of environmental contaminants and phytoestrogens in adults in Israel. *Environ Int* **2013,** *59*, 478-484.
5. Sarkar, M.; Muhammad-Kah, R.; Liang, Q. W.; Kapur, S.; Feng, S. X.; Roethig, H., Evaluation of spot urine as an alternative to 24 h urine collection for determination of biomarkers of exposure to cigarette smoke in adult smokers. *Environ Toxicol Phar* **2013,** *36* (1), 108-114.