Electronic Supplementary Material (ESI) for Analyst.
This journal is © The Royal Society of Chemistry 2014

AlRabiah *et al*. Multiple metabolomics of uropathogenic *E. coli* reveal different information content in terms of metabolic potential compared to virulence factors.

Supplementary Information: Experimental

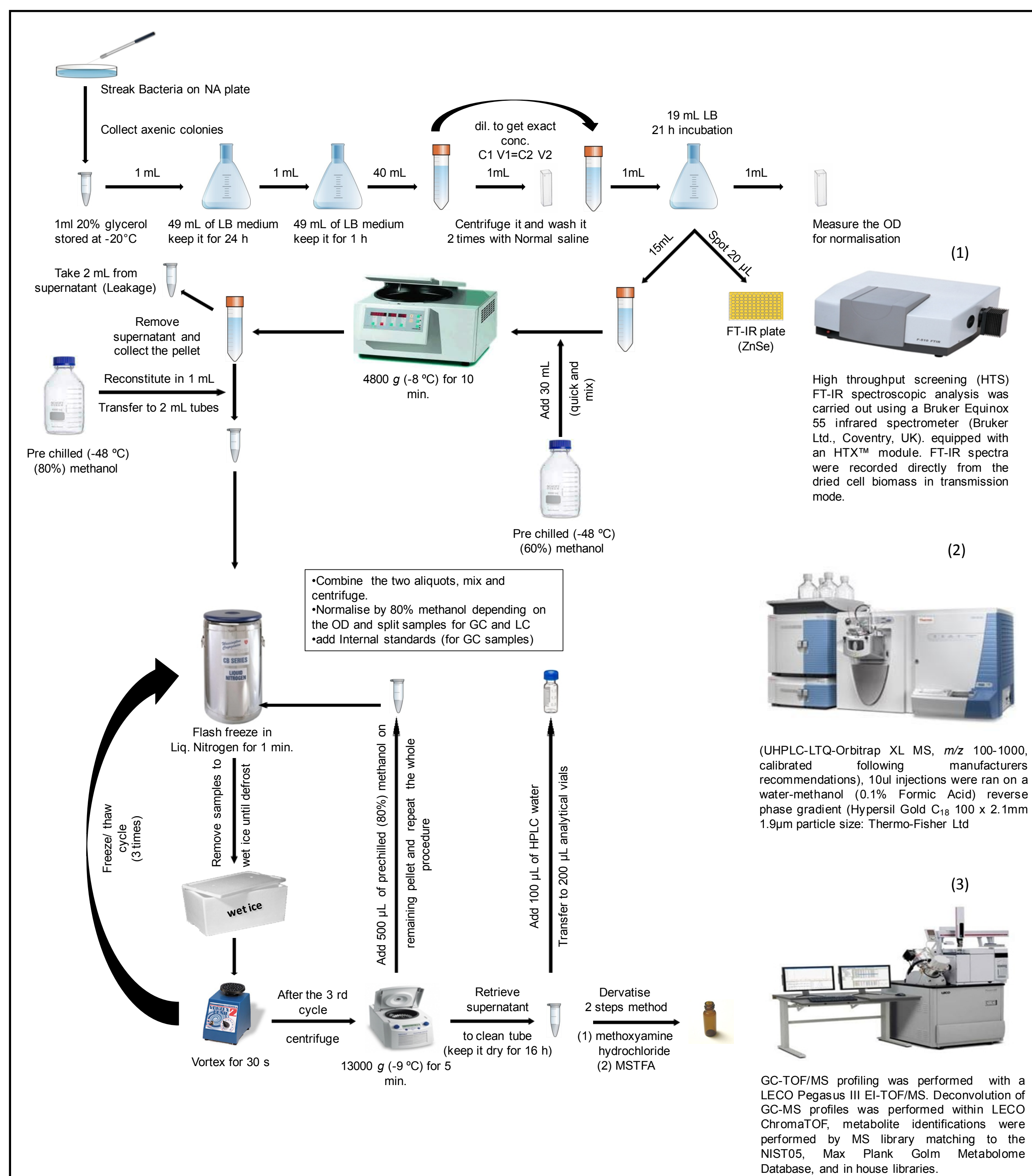

Figure S1 General scheme of sample preparation approach used, including: (1) FT-IR analysis of samples directly from the culture; (2) LC-MS; and (3) GC-MS analysis of samples after quenching and extraction using 60% and 80% cold (-48 ºC) methanol, respectively.

Sample preparation of *Escherichia coli* inoculates for metabolic fingerprinting and metabolic profiling

The whole protocol adopted for metabolomics is provided pictorially in Figure S1 and specific details are discussed below:

Following the protocol of [1], 49 mL of Lysogeny Broth (LB) was used. This was prepared by adding 5 g yeast extract (Amersham Life Sciences, Cleveland, USA) to 10 g tryptone (Formedia, Hunstanton, UK) and 10 g sodium chloride (Fisher Scientific Ltd., Loughborough, UK) and dissolving this in 1 L of reverse osmosis water followed by autoclaving at 121 ºC, for 15 min at 15 psi. This was inoculated with 1 mL of bacterial stock (20% [v/v] glycerol) of each isolate and incubated for 24 h in a shaking incubator at 37 ºC and 200 rpm. The overnight culture of each isolate (1 mL) was incubated for an additional hour at 37 ºC at 200 rpm after dilution with fresh media (49 mL). These axenic overnight cultures were washed three times with physiological saline (0.9% [w/v] NaCl) then diluted with the same washing solvent to adjust the bacterial turbidity to 0.5 McFarland standard (OD 0.1 ±0.02) optical density (OD) at 600 nm using a Biomate 5 (Thermo, Hemel Hempstead, UK) in order to standardise the size of the inocula to be used in subsequent experiments (Figure S1).

The next stage was to inoculate 19 mL of LB media with 1 mL of the experimental inocula. Each isolate was prepared in four biological replicates and the 20 mL was incubated for 21 h at 37 ºC and 200 rpm. For each biological replicate the 21 h culture was split for FT-IR, GC-MS and LC-MS to ensure that results were obtained from the same biological cultures. Starting with FT-IR, 20 µL from each replicate was spotted directly on the FT-IR plate (in 3 different wells and each well was analysed 3 times) following the method of [1]. For GC-MS and LC-MS 15 mL from each replicate was collected and quenched according to the procedures developed by [2]. The collected cultures were quenched in 30 mL of 60% cold methanol (-48 ºC) and rapidly mixed, after which the quenched culture was centrifuged at 4800 *g* and -9 ºC for 10 min. Following this, the supernatant was quickly removed and the bacterial pellets that remained were centrifuged for another 2 min and the remaining supernatant was removed. It was possible at this point to sample the quench supernatant to determine whether there had been any leakage of metabolites. The bacterial pellets were stored at -80 ºC overnight then metabolite extraction was applied following [2] method (the

only change was 80% cold methanol (80:20 methanol-water (vol/vol)) used instead of 100% methanol). The bacterial pellets were suspended in 750 µL of 80% methanol (80:20 methanol-water (vol/vol)) at -48 ºC, put into 2 mL tubes then liquid nitrogen was used to flash freeze. After this, they were put on wet ice and when they were partially defrosted the samples were thoroughly vortexed for about 30 s.

The cycle of freeze-thawing and vortexing was repeated twice more to ensure the maximum possible intracellular metabolites were extracted from within the cells. The suspensions were centrifuged at 13000 *g* and -9 ºC for 5 min. The supernatants were collected and placed in clean 2 mL tubes then held on dry ice. The pellet had 750 µL of 80% methanol (80:20 methanol-water (vol/vol)) (-48 ºC) added to it and the entire process was repeated. The second extraction aliquot was mixed with the first one, which was held on dry ice and was subsequently thoroughly vortexed (Figure S1). 800 µL of each extract of the GC-MS samples that had been normalised to equivalent OD and made up with 80% methanol (80:20 methanol-water (vol/vol)) was spiked with 100 µL of internal standard (0.3 mg/L succinic-d_4 acid, malonic-d_2 acid and glycine-d_5 in HPLC grade water). 300 µL of each extract (normalised to OD by addition of 80% methanol (80:20 methanol-water (vol/vol))) was collected for LC-MS samples. QC samples were created by combining *ca*. 100 µL from each sample (normalised to OD) and mixing thoroughly. The QC mix was divided into 11 QC samples for LC-MS, each containing 300 µL of the QC mix, and 2 QC samples for GC-MS, each containing 800 µL of the QC mix. All samples were dried for 16 h using a speed vacuum concentrator (Eppendorf 5301, Eppendorf, Cambridge, UK) operated at 30 ºC. For GC-MS samples, a chemical derivatisation with two stages was used as many metabolite classes within central metabolism are non-volatile. Prior to analysis of reaction products with GC-MS, carbonyl moieties were substituted by means of methoxyamination and followed by a per-silylation reaction. Samples taken from -80 ºC storage were put into a speed vacuum concentrator for 30 min to eliminate any remaining condensation. The next stage was to dissolve the extracts in 50 μL of 20 mg/mL *O*-methoxylamine hydrochloride in pyridine after which they were vortexed, and incubated at 60 ºC for 30 min in a dry-block heater. Then 50 μL of MSTFA was added, and the extracts were further mixed and incubated at 60 ºC for 30 min. When that was complete, 20 μL of retention index marker solution (0.3 mg/mL docosane, nonadecane, decane, dodecane, and pentadecane in pyridine) was added followed by 15 min of centrifugation at 15,800 *g*. The resultant supernatant (100 µL) from this process was transferred to 2 mL amber glass GC-MS vials fitted with 200 µL inserts before being

analysed. With the employed GC-MS analytical method, it is possible to achieve a throughput of 40 samples a day; therefore, for greater chemical stability of samples, randomised batches of 40 samples per day were derivatised throughout the period of analysis, and QC samples were also derivatised across multiple batches, to enable derviatisation (technical) and instrument (analytical) error to be measured.

Fourier transform infrared (FT-IR) spectroscopy

Aliquots of 20 µL of the bacterial preparations were spotted directly onto clean 96-well ZnSe plates (Bruker Ltd, Coventry, UK) and were dried in an oven for 45 min at 40 ºC (as detailed in [1]). A Bruker Equinox 55 infrared spectrometer (Bruker Ltd., Coventry, UK) equipped with a HTX™ module was used for high throughput screening (HTS) FT-IR spectroscopic analysis using the method of [3]. The spectra were collected in the range of 4000-600 cm^{-1}, and 64 co-adds were taken at 4 cm^{-1} resolution.

FT-IR data were converted after analysis to tab delimited files before being analysed in MATLAB 2010a (The Mathworks Inc., Natwick, USA). FT-IR data were baseline corrected using the extended multiplicative scatter correction (EMSC) algorithm [4] and CO_2 signals removed as described in [1] before performing multivariate analysis (*vide infra*).

Gas chromatography-mass spectrometry (GC-MS)

A LECO Pegasus III TOF/MS was used to conduct GC-TOF/MS. It was operated in GC-MS mode (Leco Corp., St. Joseph, MO), with a Gerstel MPS-2 autosampler (Gerstel, Baltimore, MD) and an Agilent 6890N GCxGC (operated in GC mode) with a split/splitless injector and Agilent LPD split-mode inlet liner (Agilent Technologies, Stockport, UK). A 30 m × 0.25 mm × 0.25 µm VF17-MS bonded phase capillary column (Varian, Oxford, UK) was used at a constant helium carrier gas flow of 1mL per min. The temperature program was as follows - 4 min hold at 70 ºC, 20 ºC/min to 300 ºC, 4 min hold. A split ratio of 4:1 was used for sample injections of 1 μL. The operational temperature of the injector was 280 ºC, and after 30 s, a 25 mL/min gas saver flow was used, and the transfer line was held at 240 ºC. The mass spectrometer had a source temperature of 220 ºC and was operated at 70 eV ionisation energy, and acquired *m*/*z* 45-600 at 20 Hz. The full details of the GC-MS protocol was published [5, 6] and these follow the accepted Metabolomics Standards Initiative guidelines.[7]

Liquid chromatography-mass spectrometry (LC-MS)

UHPLC-MS analysis was carried out on an Accela UHPLC autosampler system coupled to an electrospray LTQ-Orbitrap XL hybrid mass spectrometry system (ThermoFisher, Bremen, Germany). Analysis was carried out in both positive and negative ESI modes whilst each run was completely randomised to negate for any bias. A gradient type UHPLC method was used during each run as is previously described by [6, 8]. 10 µL of the extract was injected onto a Hypersil GOLD UHPLC C_{18} column (length 100mm, diameter 2.1 mm, particle size 1.9 µm, Thermo-Fisher Ltd. Hemel Hempsted, UK) held at a constant temperature of 50 ºC whilst a solvent flow rate of 400 $\mu L/min^{-1}$ was used to drive the chromatographic separation.

Xcalibur software (Thermo-Fisher Ltd. Hemel Hempsted, UK) was used as the operating system for the Thermo LTQ-Orbitrap XL MS system following the method described in [8].

Data processing was initiated by the conversion of the standard UHPLC raw files into the universal NetCDF format *via* the software conversion tool within Xcalibur. Subsequently, in house peak deconvolution software containing the XCMS algorithm (http://masspec.scripps.edu/xcms/xcms.php) was used for peak picking as described previously.[8, 9] The output from this system resulted in a Microsoft Excel based data matrix of mass spectral features with related accurate *m/z* and retention time pairs. Data from the internally pooled QC samples was then used to align for instrument drift and quality control (*via* application of an in-house Matlab script [9]). The data matrix was also signal corrected to remove peaks that exceeded the 20% RSD threshold within QC samples across the analytical run. Normalisation of each peak within the samples was achieved using the mean peak area whilst putative identification of metabolite features were performed applying the PUTMEDID-LCMS set of workflows as previously described.[10] Ambiguity arising from the same *m/z* ratio can lie within lipid identification due to differing points of unsaturation and multiple isomeric identifications. Multiple adducts of the same lipid can also occur due to the presence of different charged (composite) species (i.e. protonated and sodiated ions).

Supplementary Information: Results

Table S1 Genetic backgrounds mediating quinolone resistance in the ST131 uropathogenic *Escherichia coli* (UPEC) isolates used in this study.

DNA Ext. NO	Quinolones resistance mechanisms
2	mutation on both gyrA and parC
25	mutation on both gyrA and parC
48	*aac(6')lb-cr*
52	mutation on gyrA
75	mutation on both gyrA and parC
124	*aac(6')lb-cr*
160	*aac(6')lb-cr*
164	mutation on both gyrA and parC
183	mutation on both gyrA and parC
184	mutation on both gyrA and parC
230	*aac(6')lb-cr*

Table S2 Virulence factors for the *Escherichia coli* isolates used in this study.

Isolate no. → **VF ↓**	2	25	48	52	75	124	160	164	183	184	230
papAH	0	0	0	0	0	0	0	0	0	0	0
papC	0	0	0	0	0	0	0	0	0	0	0
papEF	0	0	0	0	0	0	0	0	0	0	0
papG II,III	0	0	0	0	0	0	0	0	0	0	0
papG I	0	0	0	0	0	0	0	0	0	0	0
allele-I	0	0	0	0	0	0	0	0	0	0	0
allele-II	0	0	0	0	0	0	0	0	0	0	0
allele-III	0	0	0	0	0	0	0	0	0	0	0
sfa/focDE	0	0	0	0	0	0	0	0	0	0	0
sfaS	0	0	0	0	0	0	0	0	0	0	0
focG	0	0	0	0	0	0	0	0	0	0	0
afa/draBC	0	0	0	1	1	0	1	1	0	0	0
bmaE	0	0	0	0	0	0	0	0	0	0	0
gafD	0	0	0	0	0	0	0	0	0	0	0
nfaE	0	0	0	0	0	0	0	0	0	0	0
fimH	1	1	1	1	1	1	1	1	1	1	1
hlyA	0	0	0	0	0	0	0	0	0	0	0
cnfI	0	0	0	0	0	0	0	0	0	0	0
cdtB	0	0	0	0	0	0	0	0	0	0	0
fyuA	1	1	1	1	1	1	1	1	1	1	1
iutA	1	1	1	1	1	1	1	1	1	1	1
kpsMT II	1	1	1	0	0	0	0	0	1	1	1
kpsMT iii	0	0	0	0	0	0	0	0	0	0	0
kpsMT k1	0	0	0	0	0	0	0	0	0	0	0
kpsMT K5	1	1	1	0	0	0	0	0	1	1	1
Rfc	0	0	0	0	0	0	0	0	0	0	0
ibeA	0	0	0	0	0	0	0	0	0	0	0
cvaC	0	0	0	0	0	0	0	0	0	0	0
traT	1	1	1	1	1	1	0	0	0	1	0
PAI	1	1	0	1	1	1	1	1	1	1	1
VF SCORE	7	7	6	6	6	5	5	5	6	7	6

Data obtained from [11, 12]

Virulence factors (VF):

Adhesion genes:	*fimH, papAH, papC, papEF, papG, alleles I, alleles II, alleles III, sfaS, focG, sfa/focDE, afa/draBC, bmaE, nfaE, gaf D.*
Toxin genes:	*cnf1, cdtB, hlyA.*
Siderophore genes:	*fyuA, iutA*
Capsule synthesis genes:	*kpsMT II, kpsMT III, kpsMT K1, kpsMT K5, rfc*
Miscellaneous genes:	*cvaC, traT, ibeA, PAI*

Table S3 Results of metabolic/biochemical profiling of the *Escherichia coli* isolates used in this study.

Isolate no. → / MT ↓	2	25	48	52	75	124	160	164	183	184	230
APPA	0	0	0	0	0	0	0	0	0	0	0
ADO	0	0	0	0	0	0	0	0	0	0	0
PyrA	0	0	0	0	0	0	0	0	0	0	0
IARL	0	0	0	0	0	0	0	0	0	0	0
dCEL	0	0	0	0	0	0	0	0	0	0	0
BGAL	1	1	1	1	1	1	1	1	1	1	1
H2S	0	0	0	0	0	0	0	0	0	0	0
BNAG	0	0	0	0	0	0	0	0	0	0	0
AGLTp	0	0	0	0	0	0	0	0	0	0	0
dGLU	1	1	1	1	1	1	1	1	1	1	1
GGT	0	0	0	0	0	0	0	0	0	0	0
OFF	1	1	1	1	1	1	1	1	1	1	1
BGLU	0	0	0	0	0	0	0	0	0	0	0
dMAL	1	1	1	1	1	1	1	1	1	1	1
dMAN	1	1	1	1	1	1	1	1	1	1	1
dMNE	1	1	1	1	1	1	1	1	1	1	1
BXYL	0	0	0	0	0	0	0	0	0	0	0
BAlap	0	0	0	0	0	0	0	0	0	0	0
ProA	1	1	1	1	0	1	1	1	1	0	0
LIP	0	0	0	0	0	0	0	0	0	0	0
PLE	0	0	0	0	0	0	0	0	0	0	0
TyrA	1	1	1	1	1	1	1	1	1	1	1
URE	0	0	0	0	0	0	0	0	0	0	0
dSOR	1	1	1	1	1	1	1	1	1	1	1
SAC	1	1	1	1	1	1	1	1	1	1	1
dTAG	0	0	0	0	0	0	0	0	0	0	1
dTRE	1	1	1	1	1	1	1	1	1	1	1
CIT	0	0	0	0	0	0	0	0	0	0	0
MNT	0	0	0	0	0	0	0	0	0	0	0
5KG	0	0	0	1	0	0	0	0	0	0	0
ILATk	1	1	1	1	1	1	1	1	1	1	1
AGLU	0	0	0	0	0	0	0	0	0	0	0
SUCT	1	1	0	1	1	1	1	1	1	1	0
NAGA	0	0	0	0	0	0	0	0	0	0	0
AGAL	1	1	1	1	1	1	1	1	1	1	1
PHOS	1	0	0	0	0	0	1	1	0	1	1
GlyA	0	0	0	0	0	0	1	1	0	0	0
ODC	1	1	1	1	1	1	1	1	1	1	1
LDC	0	1	1	0	1	1	1	1	1	1	1
IHISa	0	0	0	0	0	0	0	0	0	0	0
CMT	1	1	1	1	1	1	1	1	1	1	1
BGUR	1	0	0	1	1	1	0	0	0	1	0
O129R	1	0	1	1	1	1	1	1	1	1	1

Isolate no. → MT ↓	2	25	48	52	75	124	160	164	183	184	230
GGAA	0	0	0	0	0	0	0	0	0	0	0
IMLTa	0	0	0	0	0	0	1	1	1	0	0
ELLM	1	1	1	1	1	1	1	1	1	1	1
ILATa	0	0	0	0	0	0	1	1	1	0	0

Data obtained from [11]

Metabolic Tests (MT):

Test	Abb.	Test	Abb.	Test	Abb.
Ala-phe-pro arylamidase	APPA	Glutamyl arylamidse Pna	AGLTp	Citrate (sodium)	CIT
L-pyrrolydonyl-arylamidase	PyrA	β-glucosidase	BGLU	α-glucosidase	AGLU
L-arabitol	IARL	β-xylosidase	BXYL	β-N-acetyl-galactosaminidase	NAGA
D-cellobiose	dCEL	β-alanine arylamidase pN	BAlap	L-histidine assimilation	IHISa
H2S production	H2S	Lipase	LIP	Glu-gly-arg- arylamidase	GGAA
β-N-acetyl-glucosaminidase	BNAG	Palatinose	PLE		
Adonitol	ADO	Urease	URE	Phosphatase	PHOS
Beta-Galactosidase	BGAL	D-sorbitol	dSOR	GlycineArylamidase	GlyA
D-Glucose	dGLU	Saccharose/Sucrose	SAC	OrnithineDecarboxylse	ODC
Gamma-Glutamyl-Transferase	GGT	D-Tagatose	dTAG	Lysine Decarboxylase	LDC
Fermentation/Glucose	OFF	D-Trehalose	dTRE	Courmarate	CMT
D-Maltose	dMAL	Malonate	MNT	Beta-Glucoronidase	BGUR
D-Mannitol	dMAN	5-Keto-D-Gluconate	5KG	O/129 Resistance	O129R
D-Mannose	dMNE	L-Lactate alkalinisation	ILATk	L-Malate assimilation	IMLTa
L-ProlineArylamidase	ProA	Succinate alkalinisation	SUCT	Ellman	ELLM
Tyrosine Arylamidase	TyrA	Alpha-Galactosidase	AGAL	L-Lactate assimilation	ILATa

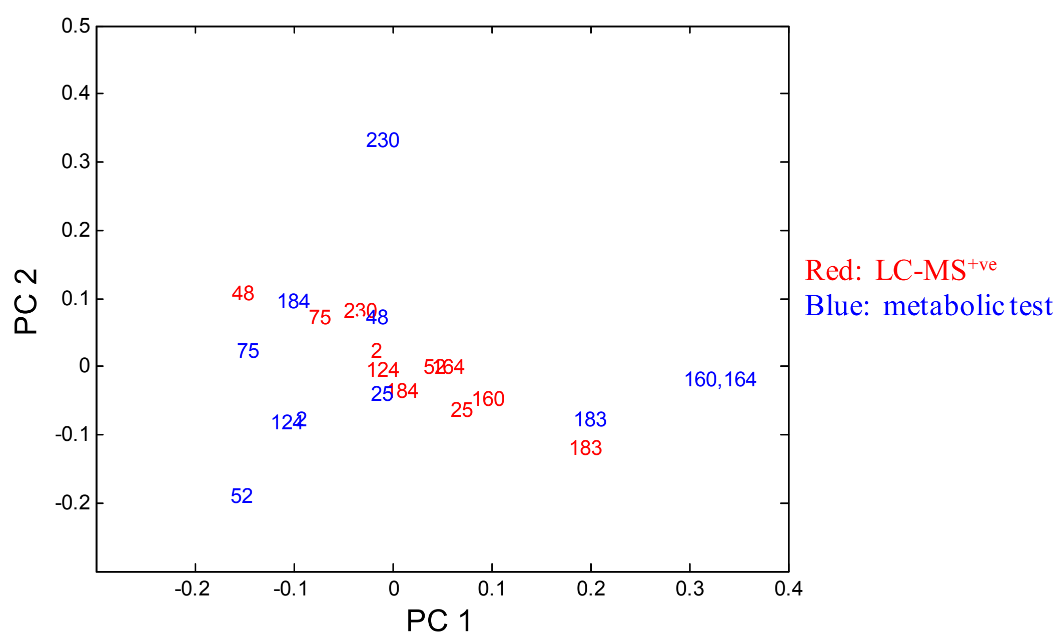

Figure S2 Superimposed scatter plots of PCoA scores of the first two components of the metabolic test and Procrustean transformed LC-MS positive mode data.

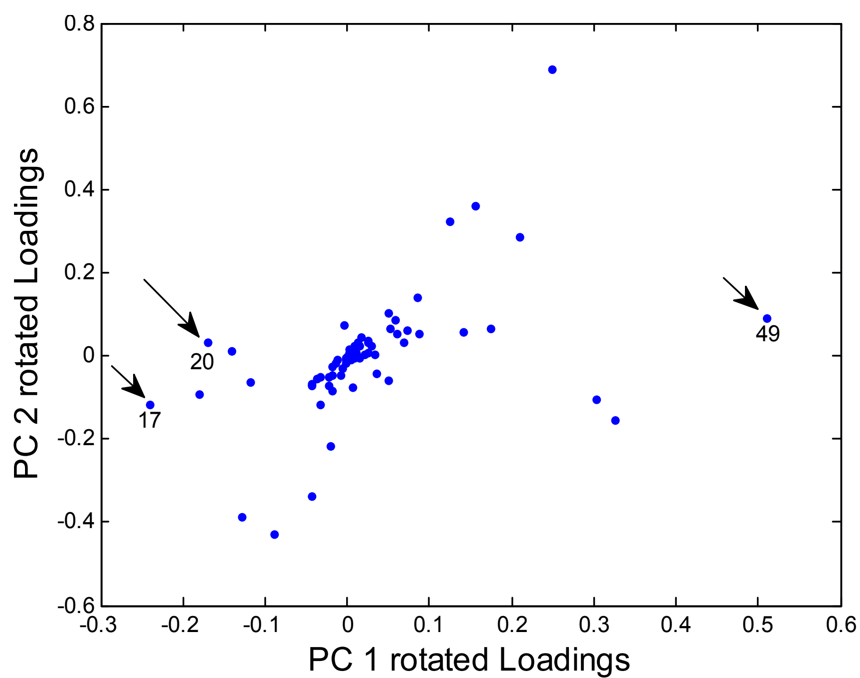

Figure S3: Rotated loading plot of GC-MS data showing the highly correlated metabolites, which were associated with metabolic profiling data.

References

1. H. AlRabiah, E. Correa, M. Upton and R. Goodacre, *Analyst*, 2013, **138**, 1363-1369.
2. C. L. Winder, W. B. Dunn, S. Schuler, D. Broadhurst, R. Jarvis, G. M. Stephens and R. Goodacre, *Anal. Chem.*, 2008, **80**, 2939-2948.
3. C. L. Winder, S. V. Gordon, J. Dale, R. G. Hewinson and R. Goodacre, *Microbiology*, 2006, **152**, 2757-2765.
4. H. Martens and T. Nęs, *Multivariate Calibration*, Wiley, 1992.
5. P. Begley, S. Francis-McIntyre, W. B. Dunn, D. I. Broadhurst, A. Halsall, A. Tseng, J. Knowles, R. Goodacre and D. B. Kell, *Anal. Chem.*, 2009, **81**, 7038-7046.
6. W. B. Dunn, D. Broadhurst, P. Begley, E. Zelena, S. Francis-McIntyre, N. Anderson, M. Brown, J. D. Knowles, A. Halsall and J. N. Haselden, *Nat. protoc.*, 2011, **6**, 1060-1083.
7. L. W. Sumner, A. Amberg, D. Barrett, M. H. Beale, R. Beger, C. A. Daykin, T. W. M. Fan, O. Fiehn, R. Goodacre, J. L. Griffin, T. Hankemeier, N. Hardy, J. Harnly, R. Higashi, J. Kopka, A. N. Lane, J. C. Lindon, P. Marriott, A. W. Nicholls, M. D. Reily, J. J. Thaden and M. R. Viant, *Metabolomics*, 2007, **3**, 211-221.
8. D. C. Wedge, J. W. Allwood, W. Dunn, A. A. Vaughan, K. Simpson, M. Brown, L. Priest, F. H. Blackhall, A. D. Whetton, C. Dive and R. Goodacre, *Anal. Chem.*, 2011, **83**, 6689-6697.
9. W. B. Dunn, D. Broadhurst, M. Brown, P. N. Baker, C. W. G. Redman, L. C. Kenny and D. B. Kell, *J. Chrom B*, 2008, **871**, 288-298.
10. M. Brown, D. C. Wedge, R. Goodacre, D. B. Kell, P. N. Baker, L. C. Kenny, M. A. Mamas, L. Neyses and W. B. Dunn, *Bioinformatics*, 2011, **27**, 1108-1112.
11. T. M. Gibreel, A. R. Dodgson, J. Cheesbrough, F. J. Bolton, A. J. Fox and M. Upton, *Journal of Clinical Microbiology*, 2012, **50**, 3202-3207.
12. T. M. Gibreel, A. R. Dodgson, J. Cheesbrough, A. J. Fox, F. J. Bolton and M. Upton, *Journal of Antimicrobial Chemotherapy*, 2012, **67**, 346-356.