ORIGINAL ARTICLE

Molecular phenotyping of a UK population: defining the human serum metabolome

Warwick B. Dunn · Wanchang Lin · David Broadhurst · Paul Begley ·
Marie Brown · Eva Zelena · Andrew A. Vaughan · Antony Halsall ·
Nadine Harding · Joshua D. Knowles · Sue Francis-McIntyre · Andy Tseng ·
David I. Ellis · Steve O'Hagan · Gill Aarons · Boben Benjamin · Stephen Chew-Graham ·
Carly Moseley · Paula Potter · Catherine L. Winder · Catherine Potts ·
Paula Thornton · Catriona McWhirter · Mohammed Zubair · Martin Pan ·
Alistair Burns · J. Kennedy Cruickshank · Gordon C. Jayson · Nitin Purandare ·
Frederick C. W. Wu · Joe D. Finn · John N. Haselden · Andrew W. Nicholls ·
Ian D. Wilson · Royston Goodacre · Douglas B. Kell

Abstract Phenotyping of 1,200 'healthy' adults from the UK has been performed through the investigation of diverse classes of hydrophilic and lipophilic metabolites present in serum by applying a series of chromatography–mass spectrometry platforms. These data were made robust to instrumental drift by numerical correction; this was prerequisite to allow detection of subtle metabolic differences. The variation in observed metabolite relative concentrations between the 1,200 subjects ranged from less than 5 % to more than 200 %. Variations in metabolites could be related to differences in gender, age, BMI, blood pressure, and smoking. Investigations suggest that a sample size of 600 subjects is both necessary and sufficient for robust analysis of these data. Overall, this is a large scale and non-targeted chromatographic MS-based metabolomics study, using samples from over 1,000 individuals, to provide a comprehensive measurement of their serum metabolomes. This work provides an important baseline or reference dataset for understanding the 'normal' relative concentrations and variation in the human serum metabolome. These may be related to our increasing knowledge of

Nitin Purandare is now deceased.

W. B. Dunn · W. Lin · D. Broadhurst · P. Begley · M. Brown ·
E. Zelena · A. A. Vaughan · A. Halsall · N. Harding ·
S. Francis-McIntyre · A. Tseng · D. I. Ellis · S. O'Hagan ·
C. L. Winder · R. Goodacre (✉) · D. B. Kell (✉)
Faculty of Engineering and Physical Sciences, School
of Chemistry, Manchester Institute of Biotechnology,
The University of Manchester, Manchester M1 7DN, UK
e-mail: roy.goodacre@manchester.ac.uk

D. B. Kell
e-mail: dbk@manchester.ac.uk

W. B. Dunn · C. L. Winder · R. Goodacre · D. B. Kell
Faculty of Engineering & Physical Sciences, Manchester Centre
for Integrative Systems Biology, Manchester Institute of
Biotechnology, The University of Manchester,
Manchester M1 7DN, UK

W. B. Dunn · W. Lin · P. Begley · M. Brown
Faculty of Medical and Human Sciences, Centre for
Endocrinology and Diabetes, Institute of Human Development,
The University of Manchester, Manchester, UK

W. B. Dunn · W. Lin · P. Begley · M. Brown
Centre for Advanced Discovery and Experimental Therapeutics
(CADET), Central Manchester University Hospitals NHS
Foundation Trust, Manchester Academic Health Sciences
Centre, Manchester M13 9WL, UK

W. B. Dunn (✉)
School of Biosciences, University of Birmingham, Edgbaston,
Birmingham B15 2TT, UK
e-mail: w.dunn@bham.ac.uk

D. Broadhurst
Department of Medicine, Katz Group Centre for Pharmacy &
Health, University of Alberta, Edmonton, AB T6G 2R3, Canada

J. D. Knowles
Faculty of Engineering & Physical Sciences, School of
Computer Science, The University of Manchester,
Manchester M13 9PL, UK

the human metabolic network map. Information on the Husermet study is available at http://www.husermet.org/. Importantly, all of the data are made freely available at MetaboLights (http://www.ebi.ac.uk/metabolights/).

1 Introduction

The biochemical composition of human cells, tissues and biofluids is highly complex, and their integrative and dynamic interactions (termed the interactome) defines function and phenotype (Vidal et al. 2011). Of these biochemicals, small molecule metabolites are involved in many important processes, from acting as the building blocks for larger biochemicals and structures, in regulation of biochemical processes, and within metabolism to generate essential cellular components (Dunn et al. 2011). The quantitative collection of metabolites in a biological system is defined as the metabolome (Oliver et al. 1998), with sample-specific metabolomes differing in composition both qualitatively and quantitatively. For fundamental reasons, the metabolome is expected [e.g. Kell (2004, 2006a, b), Kell and Westerhoff (1986)] and is indeed found (Raamsdonk et al. 2001), to amplify changes observed in the transcriptome and proteome. The holistic study of the quantitative complement of metabolites in humans provides a sensitive and dynamic snapshot of the human metabolic phenotype (Dunn et al. 2011) [also referred to as the metabotype (Gavaghan et al. 2000)]. Knowledge of

variations in metabotype may be applied in disease risk prediction and diagnosis, in understanding molecular pathophysiology, in interpreting the influence of our environment and lifestyle and in the development and assessment of drug efficacy, toxicity and adverse drug reactions. Metabolomics thus has an important role to play in personalized and stratified medicine (Nicholson et al. 2012; van der Greef et al. 2006).

Both genetics and the environment contribute significantly to human function and phenotype. Recent studies have sought to relate the influence of the genetic fingerprint on metabolism, including through the application of genome-wide association (GWAS)-metabolomics studies (Suhre and Gieger 2012; Suhre et al. 2011). These and other studies have shown the importance of applying metabolomics, alone or as part of integrated multi-omic studies to investigate human phenotypes. The use of ^{1}H NMR spectroscopy to analyse urine samples, collected in large scale epidemiological studies, has revealed interesting trends between populations and provided new biomarkers, related for example to blood pressure differences between individuals and populations (Holmes et al. 2008; Yap et al. 2010). However, whilst robust and precise, ^{1}H NMR spectroscopy does not access the whole metabolome and the use of other metabolite profiling technologies such as gas chromatography–mass spectrometry (GC–MS) and ultra performance liquid chromatography–mass spectrometry (UPLC–MS) offer excellent opportunities for expanding metabolome coverage due to the prior chromatographic separation of the many thousands of small molecules estimated via analysis of the human metabolic network (Kell and Goodacre 2014; Thiele et al. 2013) to be

G. Aarons
Faculty of Medical & Human Sciences, Institute of Inflammation and Repair, Salford Royal Dermatopharmacology, The University of Manchester, Manchester M6 8HD, UK

B. Benjamin · S. Chew-Graham · M. Zubair · A. Burns
Clinical & Cognitive Neurosciences, Faculty of Medical & Human Sciences, Institute of Brain, Behaviour & Mental Health, The University of Manchester, Manchester M13 9PT, UK

C. Moseley · P. Potter · P. Thornton · G. C. Jayson
Faculty of Medical and Human Sciences, Institute of Cancer Sciences, Christie Hospital, The University of Manchester, Manchester M20 4BX, UK

C. Potts · C. McWhirter · F. C. W. Wu · J. D. Finn
Faculty of Medical & Human Sciences, Institute of Human Development, The University of Manchester, Manchester M13 9WL, UK

M. Pan
Neurology and Gastrointestinal Centre of Excellence for Drug Discovery, GlaxoSmithKline, Third Avenue, Harlow, Essex CM19 5AW, UK

J. K. Cruickshank · N. Purandare
Faculty of Medical & Human Sciences, Institute of Cardiovascular Sciences, The University of Manchester, Manchester M13 9PT, UK

J. K. Cruickshank
St Thomas' Hospital, King's College, University of London & King's Health Partners, London SE1 9NH, UK

J. N. Haselden · A. W. Nicholls
Investigative Preclinical Toxicology, GlaxoSmithKline, David Jack Centre for Research and Development, Park Road, Ware, Hertfordshire SG12 0DP, UK

I. D. Wilson
Faculty of Medicine, Department of Surgery & Cancer, Sir Alexander Fleming Building, Imperial College London, London, SW7 2AZ, UK

I. D. Wilson
DMPK Innovative Medicines, AstraZeneca, Cheshire SK10 4TF, UK

in the human metabolome, followed by sensitive MS-based detection. A small scale study to characterize the human serum metabolome has been performed in <150 subjects (including quantification of a subset of metabolites). This study, which employed multiple analytical platforms highlighted the importance of this strategy to broaden the coverage of the metabolome and provided the first experimentally-derived serum metabolome database (Psychogios et al. 2011). However, it is only recently that technological and methodological advances that compensate for unavoidable instrumental drift (Begley et al. 2009; Dunn et al. 2011; Zelena et al. 2009) and provide high quality data have allowed us to study the large populations and numbers of metabolites needed (Broadhurst and Kell 2006) in epidemiological studies with these non-targeted MS-based techniques. Studies applying targeted assays to study low hundreds of metabolites have also been reported (Cheng et al. 2012; Yu et al. 2012).

Here we present data from The Husermet project (http://www.husermet.org/) which has applied non-targeted chromatography-mass spectrometry platforms to study the hydrophilic and lipophilic metabolic complement of serum samples obtained in a large ($n = 1,200$) investigation of the phenotype of a 'healthy' UK adult population. This required the development of substantive methods able to deal with long-term drift observed in such instrumentation. Serum samples were collected from normal healthy adults (that is to say, with no known disease at the time of sampling) of between 19 and 81 years of age over a 4-year period. We describe the variations and the influence of age, gender, BMI, blood pressure and smoking on the human serum metabolome, and the correlation of clinical chemistry measures with hydrophilic and lipophilic metabolites.

2 Results and discussion

2.1 Metabolic phenotyping of 1,200 subjects from a UK population

More than 3,000 serum samples were collected at three separate UK locations across a 4-year period, applying the same standard operating procedure (SOP) at all sites. 1,200 serum samples were selected for a first-pass of data acquisition. GC–MS and UPLC–MS in positive (UPLC–MS(+)) and negative ion (UPLC–MS(−)) ion modes were applied as complementary analytical platforms to profile a diverse range of hydrophilic and lipophilic metabolites present in the serum of 1,200 adult subjects from the UK in the age range of 19–81 years; at the time of sampling all subjects were defined as 'healthy', with no diagnosis of any disease. Data were acquired across 11 months in 10 different analytical experimental batches; each batch was composed of a single serum sample from 120 subjects and analysed across a five-day period. Each batch included the periodic analysis of a pooled quality control (QC) sample) to allow analytical variation to be measured quantitatively within and between these analytical experiments (Dunn et al. 2012). The same pooled QC sample was applied for all analytical experimental runs.

Following data pre-processing to construct a robust dataset, 126, 2178 and 2280 *metabolite features* were detected by GC–MS, UPLC–MS(+) and UPLC–MS(−), respectively; due to multiple adducts/fragments etc. during electrospray ionisation (Brown et al. 2009) more than 1,500 metabolites are estimated as being detected. All of these metabolites were detected reproducibly across all analytical experimental batches in a periodically analysed (every 5th injection) single pooled QC sample; this quantifies the variation introduced by sample preparation, data acquisition and data pre-processing. The criterion applied to define reproducible detection was relative standard deviation (RSD) less than 20 % for UPLC–MS and RSD less than 30 % for GC–MS, calculated after signal correction [see Dunn et al. (2011)]. Classes of metabolites detected included amino acids (GC–MS), organic acids (GC–MS), carbohydrates (GC–MS), fatty acids (GC–MS and UPLC–MS), peptides (UPLC–MS), acyl glycerides (UPLC–MS), sphingolipids (UPLC–MS), steroids including vitamin D metabolites (UPLC–MS) and glycerophospholipids (UPLC–MS), representing a diverse set of metabolic pathways and regulatory processes. This allowed many different areas of metabolism and biological function to be investigated simultaneously, so as to identify their importance with regard to the human 'healthy' population phenotype. This approach is in contrast to targeted studies that focus on small segments of metabolism or just a few metabolite classes. Additionally, a variety of exogenous metabolites were also detected including drugs and their metabolic products (e.g. paracetamol, (acetaminophen)). By applying linear discriminant analysis, we concluded that no metabolic differences were observed that could be related to time differences in acquiring the analytical data (see Supplementary Fig. 1), showing for the first time that a metabolome-wide study of large sample sets derived from the human population could be profiled reproducibly via chromatography-mass spectrometry platforms over a period of 11 months. A range of standard clinical chemistry measurements was also performed for all 1,200 subjects (23 assays in total including lipids (LDLC, CHOL, HDLC, TRIG), enzyme concentrations (PLT, ALK, AST, ALT, GGT, ALP, LDH), metabolites (glucose, creatinine, urea), ions (Ca, K, Na, phosphate), blood components (WBC, RBC, HAEM, TBIL) and total protein and albumin) and provided the ability to relate changes in these assays applied in routine clinical use to metabolic pathways

Table 1 Clinical characteristics of the cohort studied defining median and inter-quartile range

Characteristic	
Gender (male:female)	701:490[a]
Age (median, IQR)	48.0 (40.0,60.0)[b]
BMI (median, IQR)	25.63 (23.20,28.71)[b]
Smokers (non:ex:current)	502:163:176[c]
SBP (median, IQR), mmHg	125 (115,137)[d]
DBP (median, IQR), mmHg	76 (70,83)[d]
GLUC (median, IQR), mmol L^{-1}	4.71 (4.20,5.30)[e]
CHOL (median, IQR), mmol L^{-1}	5.10 (4.30,5.80)[f]
TRIG (median, IQR), mmol L^{-1}	1.18 (0.80,1.80)[g]
HDLC (median, IQR), mmol L^{-1}	1.26 (1.00,1.50)[h]
LDLC (median, IQR), mmol L^{-1}	3.2 (2.54,3.77)[i]

[a] Data not available for 8 subjects

[b] Data not available for 4 subjects

[c] Data not available for 355 subjects

[d] Data not available for 179 subjects

[e] Data not available for 175 subjects

[f] Data not available for 262 subjects

[g] Data not available for 326 subjects

[h] Data not available for 347 subjects

[i] Data not available for 360 subjects

and associated mechanisms. All metabolite data and associated demographic/clinical metadata are available at the publically available metabolomics data repository MetaboLights (http://www.ebi.ac.uk/metabolights/; study identifier MTBLS97). The clinical characteristics of the cohort discussed here are provided in Table 1.

2.2 Variability in relative metabolite concentrations

The relative concentrations of metabolites were investigated to derive the cumulative variation associated with background/baseline genetic and environmental influences. The distribution of variation associated with inter-subject variability [as calculated as the relative standard deviation (RSD)] for all 1,200 subjects following signal correction) is shown in Fig. 1. The distribution is skewed to lower RSD values; one interpretation of this is that the serum metabolome is comparatively tightly regulated in "healthy" populations (i.e., subjects with no diagnosed disease at the time of sampling). This could reasonably be expected, with a greater variation observed in the human urine metabolome (Bouatra et al. 2013), a biofluid composed of metabolites that are being excreted from the body. Of course, if the inter-subject variability is equivalent to the technical variability measured by replicate analysis of the same quality control (QC) sample then the metabolite feature contains no biological information. For GC–MS,

UPLC–MS(+) and UPLC–MS(−), respectively, 7 of 126, 71 of 2,181 and 42 of 2,283 *metabolic features* were observed to have an inter-subject RSD/QC RSD <1.5; thus the overwhelming majority of metabolite features reported contain biological information and those metabolite features with a value less than 1.5 were removed from further analyses.

In this distribution (see Fig. 1), those metabolites showing a high variability between subjects describe inter-subject variability which is likely to be caused by environmental and genetic variation. Caffeine showed an RSD greater than 200 % and salicylic acid, probably derived as a result of aspirin use but possibly also via tobacco, had an RSD >800 %, whilst N-methylpyrrolidinine (used in the formulation of drug vehicles) had an RSD of 550 %; such analytes thus show variation related to consumption of pharmaceutical drugs and/or specific food components. Trehalose varied by greater than 200 %, suggesting significant variation in glucose usage and storage properties, albeit trehalose is also used as a food additive. Tetradecanoic, hexadecenoic and eicosanoic acids all had variances greater than 100 %. Oxidised longer chain fatty acids and acyl carnitines also showed higher variations—potentially a sign of oxidative stress or changes in energy production in the body. Glycerophosphoethanolamines and a small peptide (*γ-glutamyl-L-isoleucine* or *γ-glutamyl-L-leucine*) also show high variation. By contrast, the aromatic amino acids (tryptophan, phenylalanine and tyrosine) all showed a low degree of inter-subject variability.

2.3 Metabolite-metabolite correlations

Metabolites do not operate in isolation but through a complex network of interactions, with metabolism being one network, though other networks are observed in biological systems (Camacho et al. 2005), especially through correlation of non-neighbour metabolites indicating their involvement in regulatory pathways [see e.g. Kotze et al. (2013)]. We note also that as reported in Camacho et al. (2005) without clear metabolite linkage, correlations should be treated with caution as correlation does not necessarily equate to causation. To highlight these complex networks we illustrate the 20 metabolites for GC–MS that show the highest pairwise Pearson's correlations. Where a metabolite was detected as more than one 'metabolic feature', only one 'feature' has been included in Fig. 2, the feature with the higher correlation coefficient. The data show the expected correlations between leucine and valine (both involved in branched chain amino acid metabolism) and between different fatty acids and glycerol (related to glycerolipid and glycerophospholipid metabolism). However, and unexpectedly, proline was also correlated with leucine and valine, and phosphate with fatty acids.
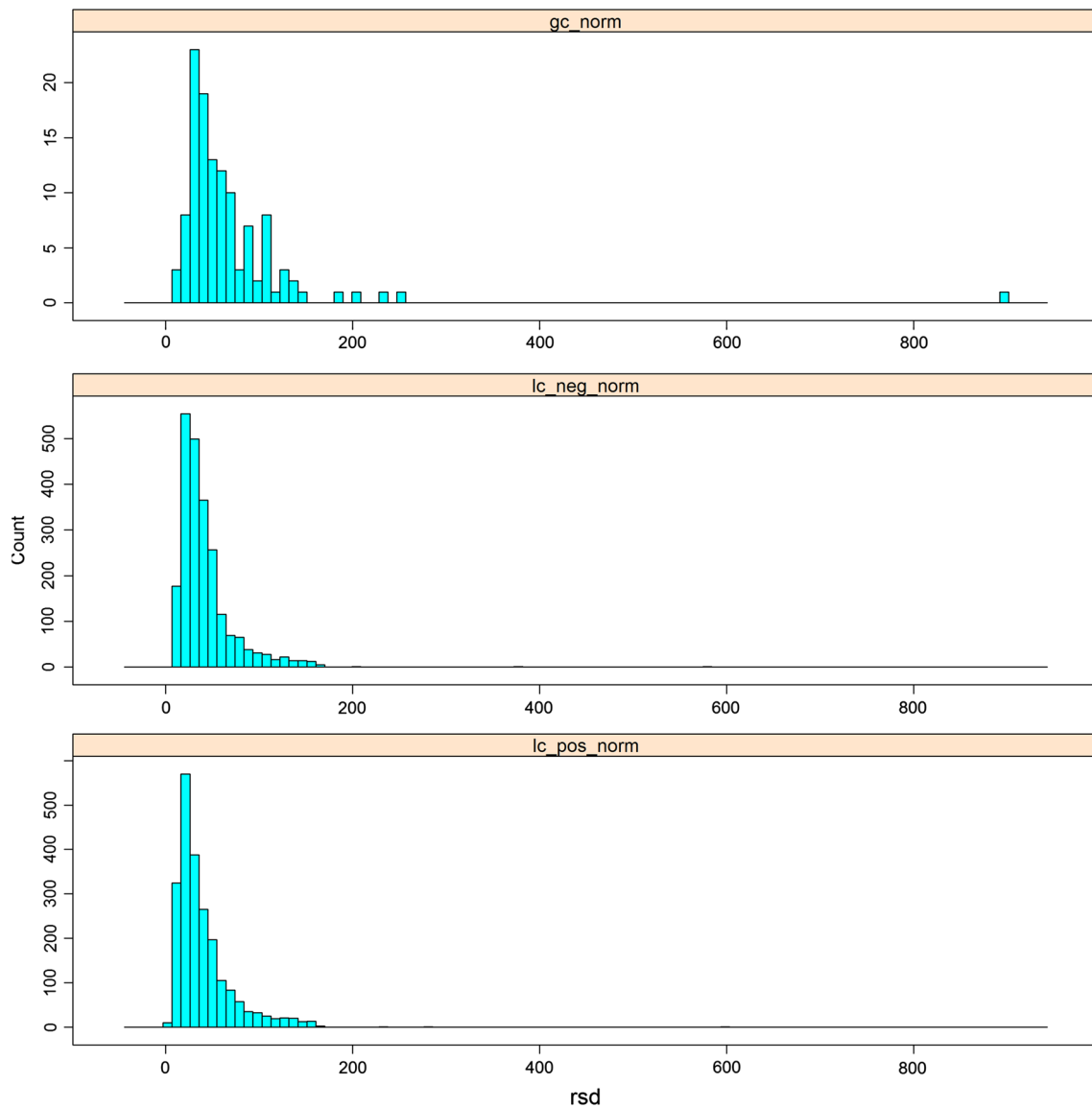
Fig. 1 The distribution of relative standard deviations defining the inter-subject variability in metabolite relative concentrations for each analytical platform applied, following signal correction. The data are shown as distribution plots. *Top plot* GC–MS, *middle plot* UPLC–MS(−), *bottom plot* UPLC–MS(+)

Assessing the UPLC–MS data (Supplementary Fig. 2) we detected expected correlations between fatty acids and oxidized fatty acids, between different sphingolipids, between fatty acids and sphingolipids, between different lyso-glycerophospholipids, between different diacylglycerides and between diacylglycerides and sphingolipids.

2.4 The effect of sample size

It is becoming increasingly evident that many biological studies are underpowered with regard to their ability to come to a robust and statistically significant and justifiable biological conclusion (Broadhurst and Kell 2006; Button et al. 2013; Dunn et al. 2011; Dunn et al. 2012; Ioannidis 2005; Ioannidis and Panagiotou 2011). It is obvious that sample size in metabolomic studies is an important aspect of experimental design, especially in terms of applying metabolites as predictive biomarkers. Although these issues have been addressed in theory [see Xia et al. (2013) for a detailed discussion], to our knowledge, no previous large-scale studies have assessed the influence of sample size. Thus, we studied the effect of sample size in terms of the prediction power of classification and the consistency of feature selection. The experimental design was to divide the whole sample population into several subsets for classification and feature selection. The results of these subsets

Fig. 2 Heatmap with dendrogram of correlation network for metabolites detected by GC–MS. The twenty unique metabolites with one or more of the highest correlations are depicted. The *lower bar* represents the *colour code* of coefficients from pairwise Pearson's correlations (Color figure online)
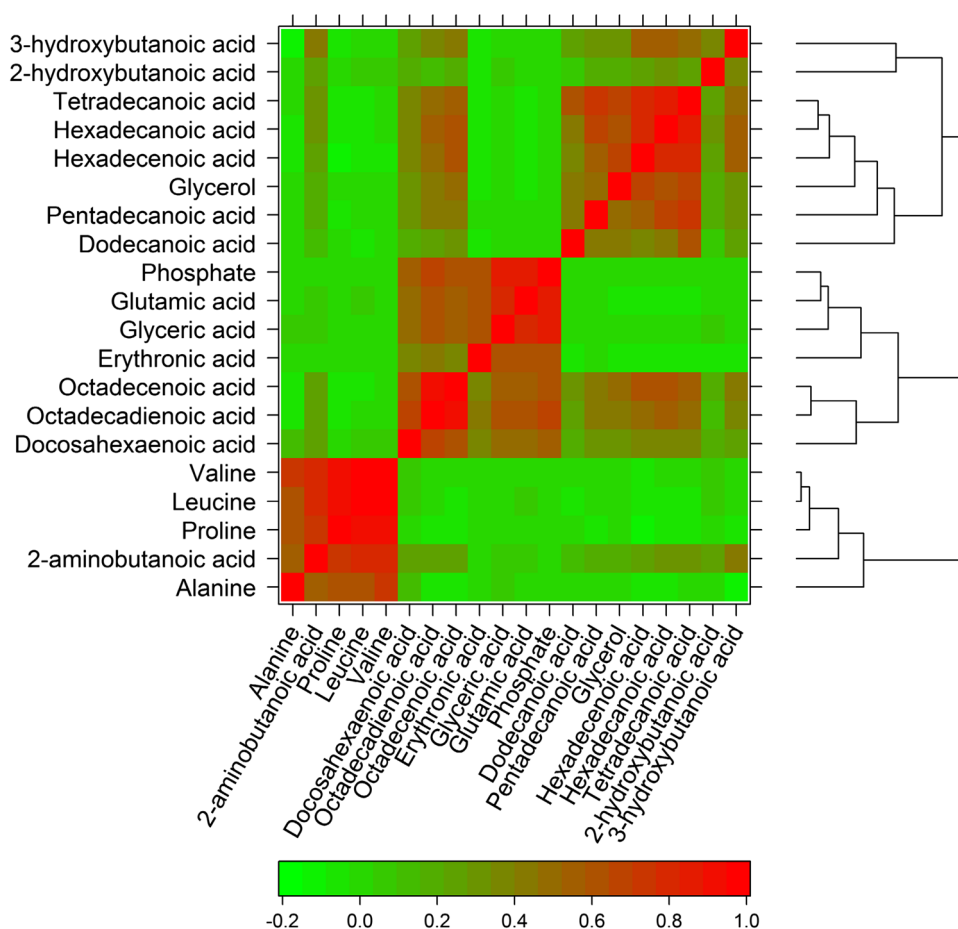
are used to select the smallest subset which has an acceptable performance, comparing this with the whole sample population in both classification and feature selection. Three groups, viz. age, gender and BMI for the three analytical platforms, have been used to evaluate the effects of sample size. Sample size is defined as the sum of samples in both classes in a binary classification and in this study the number of samples in each class was not equivalent (see Supplementary Table 1). In an ideal study the number of samples in each class would be balanced. Figure 3 shows the prediction accuracy using Random Forests (RF) with a 95 % confidence interval in the three groups (age, gender and BMI) for UPLC–MS positive ion mode. At low sample sizes the prediction accuracy was variable, but as the sample size was increased the median accuracy also increased with concomitant decrease in variation. These data showed that a sample size of 600 was appropriate to achieve similar results to those of the whole sample population with *the current dataset* where we are looking for general (i.e. not disease-specific) changes and where the variation is expected to be lower than that for the comparison of two populations such as ones that are 'healthy' and 'diseased'. A previous study based on NMR

data has shown that sample sizes of low thousands of subjects offer sufficient statistical precision to detect biomarkers quantifying predisposition to disease, a different assessment to the one we have performed above (Nicholson et al. 2011). We emphasise that this highlights the requirement to include hundreds of samples in these types of studies but does not suggest that a sample size of 600 is appropriate for all studies [for detailed discussions on this subject see Xia et al. (2013)]. However, the trends observed for all analytical platforms suggested a higher sample size would still slightly increase the prediction accuracy. The same trends were also seen with UPLC–MS(−) as well as for GC–MS. Classification results with RF and Support Vector Machine (SVM) classifiers for all three platforms and the effects of sample size on feature selection are shown in Supplementary Figs. 3 and 4.

2.5 Metabolic characteristics of this UK population

Metabolic characteristics of this subset of the UK population are discussed below. Results of data analysis performed applying consensus feature selection as described in the methods section and associated with the discussions
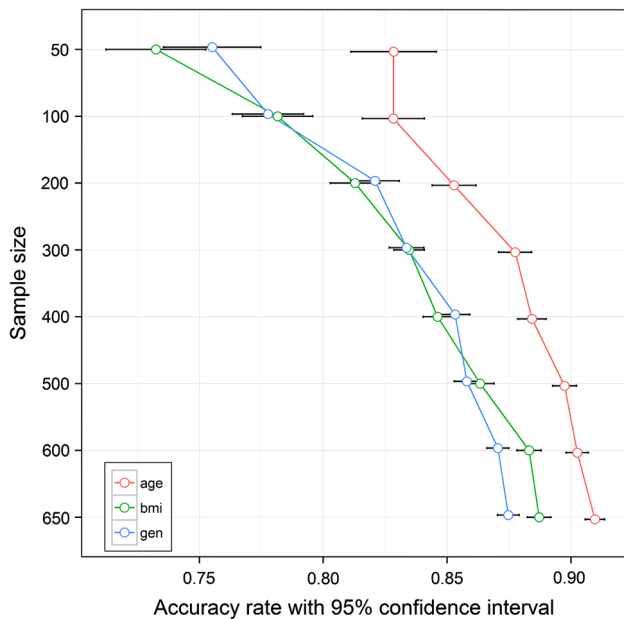
Fig. 3 Classification analysis to assess sample size effects. The accuracy rate of discrimination with 95 % confidence intervals for data acquired applying UPLC–MS(+) for the three parameters of age (age <50 vs. age >65), BMI (BMI <25 vs. BMI >30) and gender (male vs. female). A Random Forest (RF) classifier was employed and 100 bootstrap sample sets were used for the assessment of classification accuracy

related to gender, age, BMI, blood pressure and smoking are available. The results of two-way analysis of variance (ANOVA) and their post hoc analysis by Tukey's HSD ("honestly significant difference") test are available and are summarized, where appropriate, below. All data analysis results are available in supplementary material files 2–4. Where no results for two-way ANOVA are included, metabolites have been defined as biologically important by applying consensus feature selection protocol, but two-way ANOVA has shown no statistical significance with a 'critical' $p < 0.05$ [cf. Broadhurst and Kell (2006)]. Similarly interactions between main effects are only discussed if significant.

2.6 Gender

Two-way ANOVA was performed using Gender (male, female) and Age (four grouped categories: <40, 40–49, 50–64 and >64 years) as the main effects. Many differences in the serum metabolome were observed when comparing the metabolic profiles of males and females. A number of these had been observed previously highlighting the robustness of our study; these included 4-hydroxyphenyllactic acid [$F(1,1123) = 245.1$, $p = 3.9 \times 10^{-50}$], creatinine, citrate, urate [$F(1,1092) = 512.3$, $p = 2.6 \times 10^{-93}$], glycerol [$F(1,1081) = 93.7$, $p = 2.6 \times 10^{-21}$], hexadecenoic acid [$F(1,1097) = 62.8$, $p = 5.5 \times 10^{-15}$] and

tyrosine (Kochhar et al. 2006; Lawton et al. 2008; Slupsky et al. 2007). For glycerol, there was also a significant difference between age categories [$F(3,1081) = 3.1$, $p = 1.1 \times 10^{-12}$]. Tukey post hoc test showed that, independent of gender, comparisons of age categories <40 vs. 40–49 ($p = 0.0005$), <40 vs. 50–64 ($p = 9.1 \times 10^{-12}$), <40 vs. 65–81 ($p = 1.8 \times 10^{-8}$), 40–49 vs. 50–64 ($p = 0.004$) and 40–49 vs. 65–81 ($p = 0.03$) were significant using a critical p value of 0.05. There was also a significant interaction between gender and age categories for urate [$F(3,1092) = 4.8$, $p = 0.002$], glycerol [$F(3,1081) = 2.8$, $p = 0.039$] and hexadecenoic acid [$F(3,1097) = 4.7$, $p = 0.003$]. In our study, 4-hydroxyphenyllactic acid was found to be higher and tyrosine lower in males. Both of these metabolites are structurally related and these differences may reflect differences in gut microfloral co-metabolism, or the effects of alcohol consumption (Liebich and Pickert 1985). However, we observed a multitude of other robust changes related to gender also. Eight diacylglycerides were observed to be higher in relative concentration in the serum of women compared to men including DG(44:6) [$F(1,808) = 276.5$, $p = 1.3 \times 10^{-53}$] and DG(46:2) [$F(1,848) = 206.1$, $p = 5.3 \times 10^{-42}$]). For DG(46:2) there was also a significant difference between age categories [$F(3,848) = 5.8$, $p = 0.0006$] and a significant interaction between gender and age categories [$F(3,848) = 7.5$, $p = 6.0 \times 10^{-5}$]. Tukey post hoc test showed that, independent of gender, comparisons of age categories <40 vs. 65–81 ($p = 0.002$) and 50–64 vs. 65–81 ($p = 0.0009$) were significant using a critical p-value of 0.05. Four fatty acids (for example, hexadecenoic acid as shown above) and thirteen glycerophospholipids (for example, PC(36:2) [$F(1,1103) = 224.8$, $p = 2.2 \times 10^{-46}$]) showed the same trend as diacylglycerides. PC(36:2) also showed a significant difference between age categories [$F(3,1103) = 3.4$, $p = 0.02$] and a significant interaction between gender and age categories [$F(3,1103) = 4.5$, $p = 0.004$]. Tukey post hoc test showed that, independent of gender, comparisons of age categories <40 vs. 40–49 ($p = 0.02$) was significant using a critical p-value of 0.05. Serum creatinine relative concentrations were observed to be higher in females than males and, when integrated with higher phosphate levels, might suggest greater breakdown of creatine phosphate in muscles in females. Caffeine relative concentrations were higher in women [$F(1,847) = 38.3$, $p = 9.6 \times 10^{-10}$] perhaps reflecting coffee/tea/chocolate consumption, as was 2-aminomalonic acid [$F(1,1048) = 87.6$, $p = 4.8 \times 10^{-20}$] which has been associated with atherosclerotic plaques (Rupérez et al. 2012) and renal failure (Mao et al. 2008). For caffeine [$F(3,847) = 9.3$, $p = 5.0 \times 10^{-6}$] and 2-aminomalonic acid [$F(3,1048) = 3.6$, $p = 0.01$] there was also a significant difference between age categories and a significant interaction between

gender and age categories for caffeine [$F(3,847) = 6.3$, $p = 0.0003$] and 2-aminomalonic acid [$F(3,1048) = 24.3$, $p = 3.5 \times 10^{-15}$]. Tukey post hoc test for caffeine showed that, independent of gender, comparisons of age categories <40 vs. 40–49 ($p = 8.2 \times 10^{-5}$), <40 vs. 50–64 ($p = 0.0002$) and <40 vs. 65–81 ($p = 1.4 \times 10^{-5}$) were significant using a critical p-value of 0.05. Tukey post hoc test for 2-aminomalonic acid showed that, independent of gender, comparisons of age categories <40 vs. 50–64 ($p = 0.03$) and 40–49 vs. 50–64 ($p = 0.03$) were significant using a critical p-value of 0.05. Three glycerol-like metabolites (glyceric acid [$F(1,1107) = 9.1, p = 0.003$], glycerol [$F(1,1081) = 93.7$, $p = 2.6 \times 10^{-21}$] and glycerol-3-phosphate [$F(1,1127) = 11.8, p = 0.0006$]) were present in greater amounts in the serum of women compared to men, suggesting differences in glycerol metabolism and potentially related to differences in the rate of glycerolipid and glycerophospholipid synthesis. For glycerol [$F(3,1081) = 20.1, p = 1.1 \times 10^{-12}$] and glyceric acid [$F(3,1107) = 6.8$, $p = 0.0001$] there was also a significant difference between age categories. There was also a significant interaction between gender and age categories for glycerol [$F(3,1081) = 2.8$, $p = 0.04$] and glycerol-3-phosphate [$F(3,1127) = 8.7, p = 1.1 \times 10^{-5}$]. Tukey post hoc tests showed that, independent of gender, comparisons of age categories for glycerol [<40 vs. 40–49 ($p = 0.0005$), <40 vs. 50–64 ($p = 9.1 \times 10^{-12}$), <40 vs. 65–81 ($p = 1.8 \times 10^{-8}$), 40–49 vs. 50–64 ($p = 0.004$) and 40–49 vs. 65–81 ($p = 0.03$)], glycerol-3-phosphate [<40 vs. 50–64 ($p = 0.04$)] and glyceric acid [<40 vs. 40–49 ($p = 0.006$), <40 vs. 50–64 ($p = 0.0002$), <40 vs. 65–81 ($p = 0.005$)] were significant using a critical p-value of 0.05. Methionine sulfoxide, also present in greater amounts in the serum of women [$F(1,901) = 20.3, p = 7.7 \times 10^{-6}$], is an oxidation product of methionine and is considered to be a marker of oxidative stress (Bachi et al. 2013) (Fig. 4). Other gender-specific changes in the metabolome as a function of age, BMI and BP were also observed and are discussed below.

2.7 Age

We assessed age-related changes through the comparison of all subjects below the age of 50 years with all subjects older than 64 years. Two-way ANOVA was performed using Gender and Age (two categories: <50 years, and >64 years) as the main effects. Different classes of metabolites showed changes related to age, with some changes not being gender-related and others being specific to one gender. For example, citric acid showed a general increase with age for both males and females [$F(1,779) = 79.8$, $p = 3.1 \times 10^{-18}$] and therefore is probably not thus a biomarker for pancreatic cancer (Bathe et al. 2011); visually the rate of increase was greater in
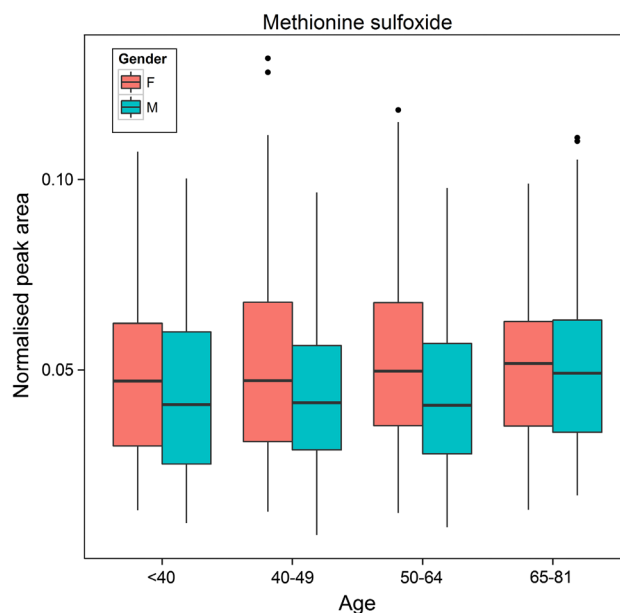

Fig. 4 A boxplot showing the distribution of methionine sulfoxide for males and females across different age categories. For *each box*, the *central line* is the median, the edges of the *box* are the upper and lower quartiles, the whiskers extend the *box* by a further $\pm 1.5 \times$ interquartile range (IQR), and outliers (>1.5 × IQR) are plotted as individual points. Data were analysed using 2-way ANOVA showing a significant difference between males and females, [$F(1,901) = 20.3, p = 7.7 \times 10^{-6}$]. There was no significant difference between age categories and no significant interaction between gender and age categories

females than in males (Fig. 5). Citrate has previously been shown to be related to age, along with other metabolites also observed in our study. These include serine [$F(1,755) = 6.5, p = 0.011$], phosphate, aspartate, erythritol/threitol [$F(1,743) = 171.0$, $p = 2.6 \times 10^{-35}$], caffeine [$F(1,565) = 8.8, p = 0.0032$], hexadecenoic acid, glycerol-3-phosphate, histidine, tryptophan [$F(1,778) = 39.1, p = 0.0007$], tyrosine [$F(1,788) = 39.1, p = 6.8 \times 10^{-10}$] and threonine [$F(1,778) = 3.9, p = 0.05$] (Lawton et al. 2008; Menni et al. 2013). There was a significant difference between gender categories for serine [$F(1,755) = 7.4, p = 0.007$], erythritol/threitol [$F(1,743) = 10.5, p = 0.001$], caffeine [$F(1,565) = 24.3, p = 1.1 \times 10^{-6}$] and tryptophan [$F(1,778) = 55.4, p = 2.6 \times 10^{-13}$]. There was also a significant interaction between gender and age categories for caffeine [$F(1,565) = 17.6, p = 3.2 \times 10^{-5}$].

Age-related changes in amino acids were also observed. These changes included tryptophan [$F(1,778) = 11.7, p = 0.0007$]; also showed a significant difference between gender categories [$F(1,778) = 55.4, p = 2.6 \times 10^{-13}$] which decreases with age and tyrosine [$F(1,788) = 39.1, p = 6.8 \times 10^{-10}$] which increases with age (as shown in Fig. 6), threonine and serine which both decreased with age and methionine and cysteine [$F(1,785) = 16.0$,
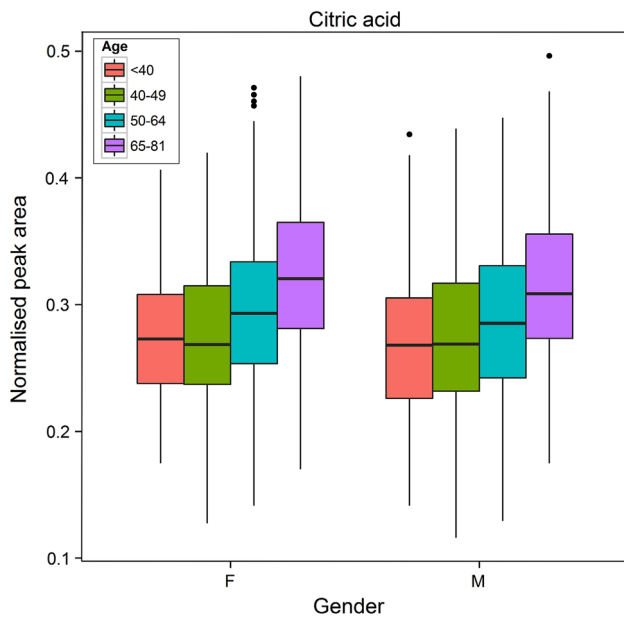
Fig. 5 A boxplot showing the distribution of citric acid for males and females across different age categories. For *each box*, the *central line* is the median, the edges of the *box* are the upper and lower quartiles, the whiskers extend the box by a further $\pm 1.5 \times$ interquartile range (IQR), and outliers are plotted as individual points ($>1.5 \times$ IQR). Data were analysed using 2-way ANOVA. There was a significant difference between males and females ($F(1,779) = 79.8$, $p = 3.1 \times 10^{-18}$). There was no significant difference between age categories and no significant interaction between gender and age categories

$p = 7.1 \times 10^{-5}$] which also both decreased with age. Cysteine also showed a significant difference between gender categories [$F(1,785) = 12.9$, $p = 0.0003$] and showed a significant interaction between gender and age categories [$F(1,785) = 4.8$, $p = 0.03$]. Vitamin D metabolites also show decreases with age in both males and females, and have been related to the onset of the metabolic syndrome [e.g. Lee et al. (2009), Lu et al. (2009)] and this observation might argue for the benefits of vitamin supplementation in older people. For example, 24-Hydroxygeminivitamin D_3 showed a difference between age categories [$F(1,703) = 52.2$, $p = 1.3 \times 10^{-12}$], gender categories [$F(1,703) = 36.8$, $p = 2.2 \times 10^{-9}$] and a significant interaction between age and gender categories [$F(1,703) = 5.7$, $p = 0.02$]. Different fatty acids showed either increases or decreases with age (e.g. octadecadienoic acid increased with age [$F(1,763) = 8.6$, $p = 0.003$]), but no correlation between age and carbon number, nor degree of saturation, was observed for fatty acids. Erythritol and/or threitol showed an increase (as shown above) with age as did inositol [$F(1,779) = 151.8$, $p = 5.5 \times 10^{-32}$], which also showed a significant interaction between age and gender categories [$F(1,779) = 11.3$, $p = 0.0008$].

These two changes are consistent with the age-dependent increases in classes of carbohydrates that underpin diabetic complications (Brownlee 2001).

2.8 BMI

While gender and age are independent variables, the body mass index (BMI) is not (although is taken as such for the purposes of this study where one class is BMI <25 and the other class is BMI >30). Nonetheless, with obesity becoming a growing problem in developed and developing countries, even in children (Friend et al. 2013), the measurement of BMI and its relationship to the serum metabolome has become of increasing importance. As is well known, increased BMI is correlated to increases in body fat, greater risk of insulin resistance and metabolic disorders including diabetes and cardiovascular diseases [e.g. Pradhan (2007)]. It should be remembered that BMI is linked to excess weight and the associated risk of insulin resistance and metabolic disorders. BMI is not directly correlated to adiposity as a higher BMI can be related to excess bone, muscle or fat and does not take into account the distribution of the latter and its influence on metabolic diseases. However, BMI provides a readily available surrogate measure of overall body fatness in large-scale studies and was therefore chosen as an appropriate surrogate marker in this study. Two-way ANOVA was performed using BMI (<25 vs. >30) and gender as the main effects.

In this study a range of amino acids showed either an increase (cysteine [$F(1,690) = 18.8$, $p = 1.6 \times 10^{-5}$], cystine [$F(1,686) = 16.9$, $p = 4.4 \times 10^{-5}$], glutamine, tyrosine [$F(1,695) = 62.6$, $p = 9.9 \times 10^{-15}$], phenylalanine [$F(1,687) = 28.4$, $p = 1.4 \times 10^{-7}$] and valine [$F(1,685) = 32.0$, $p = 2.2 \times 10^{-8}$]) or decrease (asparagine [$F(1,687) = 12.8$, $p = 0.0004$], histidine, serine [$F(1,670) = 4.1$, $p = 0.04$] and phosphoserine [$F(1,498) = 29.6$, $p = 8.3 \times 10^{-8}$]) in relative amounts as BMI increased in one or both genders. Cysteine [$F(1,690) = 11.6$, $p = 0.0007$], valine [$F(1,685) = 53.9$, $p = 6.0 \times 10^{-13}$], serine [$F(1,670) = 6.9$, $p = 0.009$] and phosphoserine [$F(1,498) = 6.1$, $p = 0.01$] also showed a significant difference between gender categories and there was a significant interaction between gender and BMI categories for tyrosine [$F(1,695) = 4.3$, $p = 0.04$] and phosphoserine [$F(1,498) = 5.7$, $p = 0.02$]. Valine, tyrosine and phenylalanine have been strongly linked as early makers of insulin resistance and markers of risk for the development of diabetes (Newgard et al. 2009; Wang et al. 2011). Phosphoserine can be associated with cysteine production, serine metabolism or as a byproduct of protein degradation. Short-chain organic acids (including acetate [$F(1,645) = 38.4$, $p = 1.1 \times 10^{-9}$], 2-aminobutanoic acid

Fig. 6 A boxplot showing the distribution of tyrosine and tryptophan for males and females across different age categories. For *each box*, the *central line* is the median, the edges of the *box* are the upper and lower quartiles, the whiskers extend the *box* by a further $\pm 1.5 \times$ interquartile range (IQR), and outliers are plotted as individual points ($>1.5 \times$ IQR). Data were analysed using 2-way ANOVA. There was a significant difference across age categories (<50 years vs. >64 years) for tryptophan [$F(1,778) = 11.7$, $p = 0.0007$] and tyrosine [$F(1,788) = 39.1$, $p = 6.8 \times 10^{-10}$]. There was a significant difference across gender categories for tryptophan [$F(1,788) = 55.4$, $p = 2.6 \times 10^{-13}$]. There was no significant interaction between gender and age categories for tryptophan or tyrosine
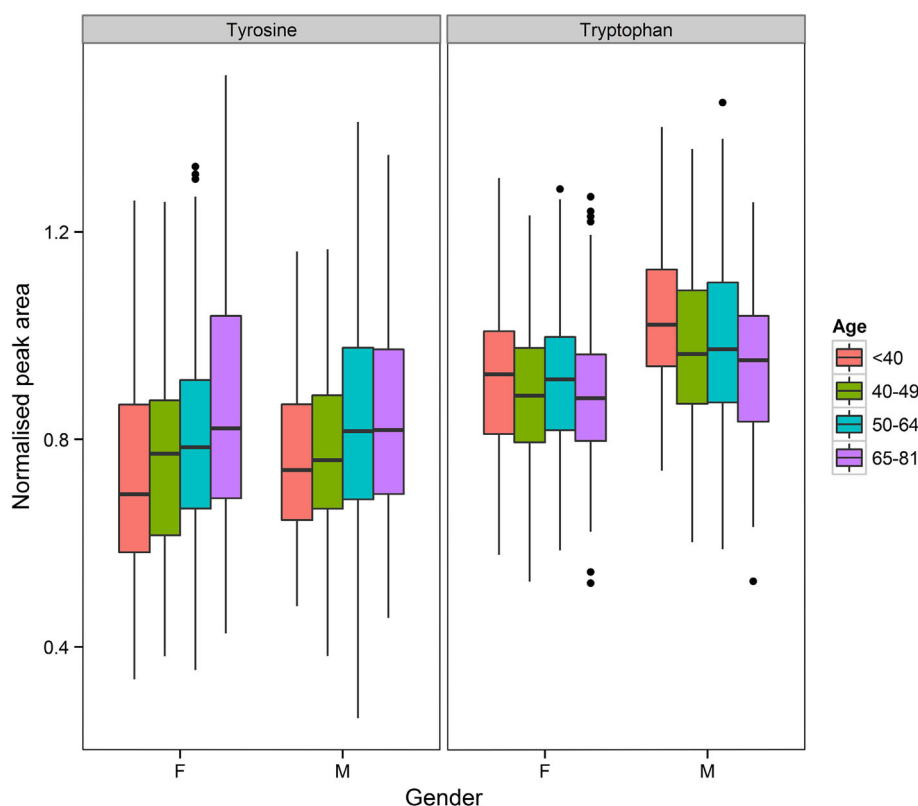
[$F(1,637) = 8.9$, $p = 0.003$] and 2-aminomalonic acid [$F(1,642) = 57.2, p = 1.4 \times 10^{-13}$]) showed a decrease in relative concentration with increasing BMI. 2-Aminomalonic acid also showed a significant difference between gender categories [$F(1,642) = 34.3$, $p = 7.4 \times 10^{-9}$]. Four diacylglycerides show a decrease as BMI increased, for example, DG(44:6) showed a statistically significant difference applying 2-way ANOVA [$F(1,489) = 57.0$, $p = 2.1 \times 10^{-13}$] and also showed a significant difference between gender categories [$F(1,489) = 143.5, p = 3.6 \times 10^{-29}$]. Five sphingolipids show a decrease as BMI increased, for example, SM(d18:1/24:1) showed a statistically significant difference applying 2-way ANOVA [$F(1,518) = 36.0$, $p = 3.8 \times 10^{-9}$] and also showed a significant difference between gender categories [$F(1,518) = 88.5$, $p = 1.6 \times 10^{-19}$]. Four lyso-glycerophospholipids show a decrease as BMI increased, for example, lysoPC(18:2) showed a statistically significant difference applying 2-way ANOVA [$F(1,693) = 88.4$, $p = 7.6 \times 10^{-20}$] and also showed a significant difference between gender categories [$F(1,693) = 27.1$, $p = 2.6 \times 10^{-7}$]. Three fatty acids show a decrease as BMI increased, for example, dodecanoic acid showed a statistically significant difference applying 2-way ANOVA [$F(1,658) = 20.4$, $p = 7.4 \times 10^{-6}$] and also showed a significant difference between gender categories [$F(1,658) = 34.9$, $p = 5.7 \times 10^{-9}$]. Citrate and fructoselysine-3-phosphate

showed female-specific decreases as a function of BMI. The latter is observed in increased concentrations in tissue and biofluids of diabetic subjects as an Advanced-Glycation Endproduct (AGE) (Delpierre and Van Schaftingen 2003). Glycerol [$F(1,655) = 43.9$, $p = 7.1 \times 10^{-11}$] and glycerol-3-phosphate showed male-specific increases in amounts (2-way ANOVA results for comparison of gender for glycerol was $F(1,655) = 91.2$, $p = 2.6 \times 10^{-20}$). Glutamine and glutamate showed an increase and a decrease respectively and threonine showed a decrease as BMI increased. Correlation analysis showed that diglycerides, glycerophosphocholines, sphingomylenins, tyrosine, tyrosyl-arginine and urate also correlated with BMI (Supplementary Fig. 5).

2.9 Blood pressure

Elevated blood pressure (BP) is associated with an increased risk of cardiovascular diseases [e.g. He and Whelton (1999)]. In the UK, up to 38 % of the population is considered hypertensive at one stage or another of their lives, with a greater prevalence of high blood pressure in men. Here we found that a range of metabolic classes in serum were altered in relation to increasing blood pressure when comparing normal blood pressure (systolic = 90–120 mmHg) versus hypertension (systolic >140 mmHg). Two-way ANOVA was performed using

Blood Pressure (Normal; Hypertension) and Gender as the main effects.

Methionine sulfoxide was negatively correlated with BP, in both males and females, and methionine showed an increase in relative concentration with blood pressure. One interpretation is that reactive oxygen species that do not oxidize methionine may damage other tissues, leading to a range of disorders (Kell 2009).

Multiple amino acids showed changes including a decrease in cysteine [$F_{(1,589)} = 11.5$, $p = 0.0007$] and lysine in both males and females whilst other changes were gender specific (e.g., decreased alanine [$F_{(1,567)} = 13.3$, $p = 0.0003$] and increased tryptophan in males only and increased histidine and decreased threonine in females only). Cysteine also showed a significant difference between gender categories [$F_{(1,589)} = 10.2$, $p = 0.001$] and there was a significant interaction between gender and BP categories for cysteine [$F_{(1,589)} = 8.6$, $p = 0.003$]. Lactate relative concentrations were increased in both genders [$F_{(1,587)} = 9.3$, $p = 0.002$] whilst acetate [$F_{(1,543)} = 10.1$, $p = 0.002$] decreased in females only. Citrulline increased ($F_{(1,592)} = 5.7, p = 0.02$) and showed a significant difference between gender categories [$F_{(1,592)} = 3.9$, $p = 0.05$] across both sexes as BP increased, while erythritol/threitol [$F_{(1,554)} = 10.1$, $p = 0.002$] showed a interaction between gender and BP which was statistically significant [$F_{(1,554)} = 5.8$, $p = 0.02$], and erythronic acid/threonic acid decreased in both genders. Glyceric acid and glycerol-3-phosphate both increased and sucrose decreased in both males and females. Other changes included decreases in indole-acetate [$F_{(1,561)} = 9.7$, $p = 0.002$] in males only. Correlation analysis showed links to elevated BP to urate, triacylglycerides, dipeptides, glycerophosphocholines and 4-hydroxyphenyl-lactic acid (Supplementary Fig. 5).

2.10 Smoking

Smoking is an important risk factor in cancer and cardiovascular diseases; the metabolic disturbances associated with smoking can have important roles in the onset and progression of these diseases. Two-way ANOVA was performed using smoking (non-smoker, ex-smoker, smoker) and gender as the main effects. Correlation analysis showed links between smoking status and salicylic acid, assumedly derived from aspirin and the lifestyle influences on the metabolic phenotype. Smoking was also correlated with the two aromatic amino acids tyrosine ([$F_{(2,796)} = 3.7$, $p = 0.02$], Tukey post hoc tests showed that, independent of gender, comparisons of smoking categories were statistically significant for smokers vs. non-smokers ($p = 0.02$)) and tryptophan (elevated in smokers). Tryptophan has been associated with smoking initiation and nicotine dependence

previously (Wang and Li 2010) and our data show decreases in the metabolically related indole-acetate and indole-propionate ([$F_{(2,757)} = 1.4$, $p = 1.3 \times 10^{-5}$]; indole propionate also showed a significant difference between gender categories [$F_{(1,757)} = 4.7$, $p = 0.03$] and there was a significant interaction between gender and BP categories for indole-propionate [$F_{(2,757)} = 6.2$, $p = 0.002$]; Tukey post hoc tests showed that, independent of gender, comparisons of smoking categories were statistically significant for smokers vs. non-smokers ($p = 7.7 \times 10^{-6}$) and smokers vs ex-smokers ($p = 0.04$)). Statistical analysis also showed decreases in other amino acids including aspartate, histidine and lysine in smokers. Glycerol ([$F_{(2,759)} = 3.3$, $p = 0.04$]; glycerol also showed a significant difference between gender categories [$F_{(1,759)} = 40.8$, $p = 2.9 \times 10^{-10}$]; Tukey post hoc tests showed that, independent of gender, comparisons of smoking categories were statistically significant for smokers vs. non-smokers ($p = 0.03$)) and glycerol-3-phosphate were decreased in smokers as were a number of fatty acids (for example, octadecenoic acid [$F_{(2,759)} = 3.3$, $p = 0.04$]; Tukey post hoc tests showed that, independent of gender, comparisons of smoking categories were statistically significant for non-smokers vs. ex-smokers ($p = 0.02$)). Lactate [$F_{(2,778)} = 3.5$, $p = 0.03$; Tukey post hoc tests showed that, independent of gender, comparisons of smoking categories were statistically significant for non-smokers vs. smokers ($p = 0.03$)] and citrate [$F_{(2,800)} = 3.9$, $p = 0.02$; Tukey post hoc tests showed that, independent of gender, comparisons of smoking categories were statistically significant for non-smokers vs. smokers ($p = 0.01$)] are also decreased in smokers as is inositol [$F_{(2,784)} = 15.7, p = 2.0 \times 10^{-7}$; Tukey post hoc tests showed that, independent of gender, comparisons of smoking categories were statistically significant for non-smokers vs. smokers ($p = 9.2 \times 10^{-8}$) and for ex-smokers vs. smokers ($p = 0.0006$)]. Biotin was decreased in smokers [$F_{(2,814)} = 20.0$, $p = 3.2 \times 10^{-9}$; Tukey post hoc tests showed that, independent of gender, comparisons of smoking categories were statistically significant for non-smokers vs. smokers ($p = 1.3 \times 10^{-9}$) and for ex-smokers vs. smokers ($p = 0.001$)] and this has been shown previously in women (Sealey et al. 2004). Finally caffeine is present at lower relative concentrations in smokers [$F_{(2,655)} = 8.1$, $p = 0.0003$; also showed a significant difference between gender categories [$F_{(1,655)} = 32.5$, $p = 1.8 \times 10^{-8}$]; Tukey post hoc tests showed that, independent of gender, comparisons of smoking categories were statistically significant for smokers vs. non-smokers ($p = 0.001$) and smokers vs. ex-smokers ($p = 0.0006$]) which is unexpected as there is a logical lifestyle link between coffee drinkers and smokers; however this may show a change in rates of caffeine metabolism in smokers.
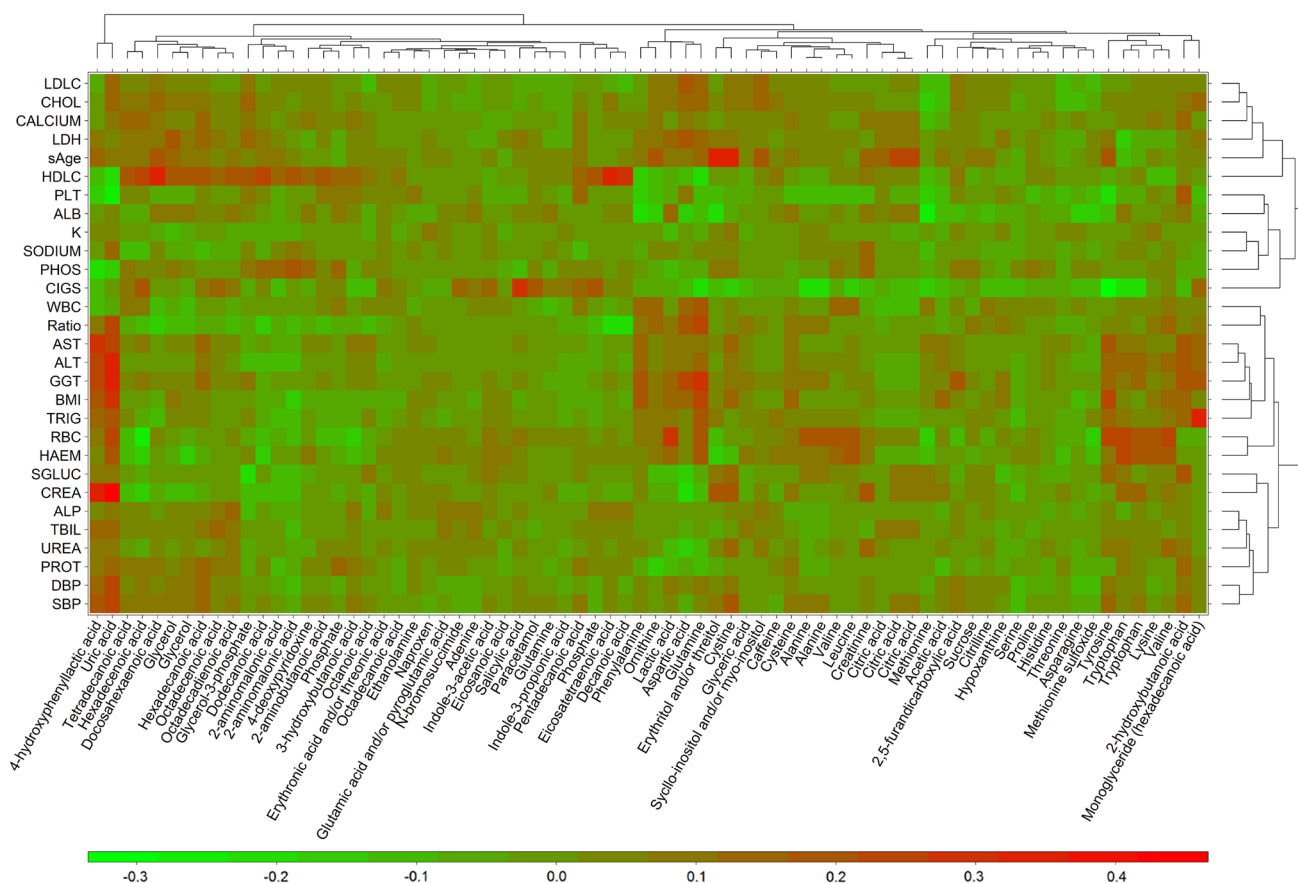
Fig. 7 Heatmap with dendrogram of Pearson's correlation analysis between metabolites detected by GC–MS and clinical chemistry data. The arrangement of the clusters are produced by hierarchical clustering on both metabolites and clinical chemistry data. The *lower bar* represents the *colour code* of coefficients from pairwise Pearson's correlations between GC–MS data and the clinical chemistry data (Color figure online)

2.11 Correlations between clinical chemistry and metabolic profiling data

In addition to metabolite profiling, each sample was also subjected to a panel of conventional clinical chemistry assays. This was to enable positive and negative correlations (if any) to these standard clinical diagnostics and the broader metabolic phenotypes to be determined. This ability to anchor newer methods of volunteer/patient phenotyping, in this case metabotyping, with currently used "best practice" represents an important step towards obtaining wider acceptance of the utility of the metabolite profiling approach. The results of this for the correlation of clinical chemistry with GC–MS analysis is illustrated in Fig. 7 (UPLC–MS correlation in Supplementary Fig. 5).

An obvious area where such correlations would be expected is across lipid (and particularly cholesterol) metabolism. As might be expected correlations emerged from the metabotypes determined here between total cholesterol concentrations in serum and the amounts of monoglycerides and diglycerides present. There were also

positive correlations between circulating high density lipoprotein cholesterol (HDLC) and relative concentrations of fatty acids, diglycerides, phosphotidylcholines, sphingomylenins and triglycerides, although we were unable to find any correlations for the low density lipoprotein (LDLC). Triglycerides as determined by standard clinical chemistry assays were associated with raised di- and monoglycerides, phosphatidylcholines, sphingomylenins and urate. As discussed above, there was also a correlation of diastolic blood pressure with urate, triglycerides and phosphatidylcholines.

Another set of interesting correlations relating to organ function was seen when some of the clinical markers for liver function were examined. For example, amongst a range of other correlations, both AST and ALT were associated with relative concentrations of urate and 4-hydroxyphenyllactic acid. ALT, in addition, also covaried with acylglycerides and the PC/PE ratio. As observed above, systolic blood pressure (SBP) was also associated with 4-hydroxyphenyllactic acid. Another liver enzyme, GGT varied with diglycerides, glycerophosphocholines,

urate, tyrosyl-arginine, aspartate and glutamate whilst no correlations were seen for LDH. In the case of renal function creatinine and urea concentrations were both associated with circulating dipeptides and hexanoylglycine, with creatinine also covarying with phosphatidylcholines, sphingomyelenins, urate, erythritol/threitol and triglycerides. Correlations for many other clinical chemistry markers, for e.g., serum glucose, with circulating metabolites were also found.

2.12 Concluding statement and future roles for The Husermet Project

The importance of the Husermet project is that it has developed the tools and resources to collect and provide metabolic profiles based on chromatography-mass spectrometry for a large human population (Dunn et al. 2011). This is a vital prerequisite to well-powered studies that can complement the large-scale but necessarily qualitative studies of genome sequence variation now appearing. Here we describe how these have been applied to profile a sample of the 'normal' UK population and for these 1,200 healthy individuals define biologically important metabolic changes associated with age and gender as well to link metabolic changes with disease risk factors including BMI, blood pressure and smoking. It was noteworthy that a significant number of metabolites known to be associated with insulin resistance and the metabolic syndrome did indeed increase with age, indicating the great dangers of a diabesity epidemic in the UK. Additionally, we have correlated metabolic variations with clinical chemistry measurements to indicate metabolic disturbances associated with differences in these variables. 'Omics' measurements are normally hypothesis-generating rather than hypothesis testing (Kell and Oliver 2004), although it is always gratifying to be able to reproduce known and published data, as many examples illustrated above have done.

Most importantly, the Husermet protocol has developed a dataset made publicly available through MetaboLights (Haug et al. 2013) so this large resource can be applied as required by the scientific community. An obvious next step is the integration of our data with those for the recently published human metabolic network reconstruction (Swainston et al. 2013) and the other small molecules with which it interacts (Kell 2013; Kell et al. 2013).

3 Materials and methods

3.1 Ethics statement

Written informed consent was obtained from each study participant and the study conformed to the principles set out in the WMA Declaration of Helsinki and the NIH Belmont report. The study was approved by the Stockport Local Research Ethics Committee.

3.2 Sample collection

Following assessment of suitable plasticwares such that any plasticizers, phthalates etc. were minimal or absent, serum was collected from 1,200 subjects following appropriate ethical approval of the study; informed consent was acquired from all subjects. A range of clinical parameters were acquired (including age, gender, BMI and smoking status). No data related to medication or food intake were collected. Approximately 10 mL of blood was drawn into serum collection tubes (Greiner, Stonehouse, UK) and was allowed to clot on ice at 4 °C for a minimum of 1 h. The serum fraction was separated by centrifugation (2,500×g, 4 °C, 15 min) and 500 μL volumes were aliquoted into separate cryovials (Greiner, Stonehouse, UK). Serum was processed and frozen at −80 °C within 6 h of blood collection. All samples were transported to The University of Manchester on dry ice and stored at −80 °C. Samples were analysed within 2 years of sample collection.

3.3 Sample preparation

All samples were prepared according to a SOP as described previously (Dunn et al. 2011) and will not be described in detail here. In summary, serum was allowed to thaw on ice followed by addition of 1,200 μL of methanol and 200 μL of internal standard solution (0.167 mg mL^{-1} malonic acid d_2, succinic acid d_4, glycine d_5, citric acid d_4, D-fructose $^{13}C_6$, L-tryptophan d_5, L-lysine d_4, L-alanine d_7, stearic acid d_{35}, benzoic acid d_5 and octanoic acid d_{15}) to 400 μL of serum. The sample was vortex mixed and following centrifugation, four 370 μL aliquots were transferred to separate tubes and dried in a centrifugal vacuum evaporator for 18 h. Quality control (QC) samples were prepared applying a pooled serum sample (Sigma-Aldrich; S7023) as described above.

3.4 Data acquisition

Data were acquired on three analytical platforms (UPLC–MS positive and negative ion modes and GC–MS) according to a SOP as described previously (Dunn et al. 2011) and will not be described in detail here. Samples for UPLC–MS analysis were reconstituted in 100 or 200 μL of water for negative and positive ion modes, respectively and analysed applying reversed-phase UPLC–MS (Waters Acquity UPLC coupled to a Waters LCT mass spectrometer) with a 22 (positive ion mode) or 24 (negative ion

mode) minute analysis time. 10 QC samples were analysed at the start of each analytical batch to condition the analytical system and a QC sample was analysed every 5th injection. Samples for GC–MS analysis were prepared applying a two-stage chemical derivatisation procedure (oximation followed by trimethylsilylation) and followed by analysis applying an electron ionisation GC-ToF–MS system (Agilent 6890 N GC coupled to a LECO Pegasus III mass spectrometer). For GC-ToF–MS, 5 QC samples were analysed at the start of each analytical batch to condition the analytical system and a QC sample was analysed every 5th injection. Samples from 1,200 subjects were analysed in 10 different analytical experimental batches, with 120 subject samples analysed in each batch and each batch consisting of analysis across a five day period. Two experimental runs consisting of 60 subjects in each run was operated for UPLC–MS and four experimental runs consisting of 30 subjects in each run was operated for GC–MS. Each batch of 120 subjects was prepared such that it contained a near-random selection of subjects according to the traits in which we are interested (viz. age, gender, BMI, blood pressure, smoking); this was to ensure that any failed batch would not compromise the overall study.

3.5 Data pre-processing

Data were pre-processed and integrated according to a SOP as described previously (Dunn et al. 2011). UPLC–MS data were converted from the raw instrument datafile to Net-CDF files and subsequently XCMS was applied for peak deconvolution and alignment separately for each analytical batch. Due to the untargeted nature of the UPLC–MS analysis, the number and identity of common peaks detected in each batch differed considerably. Thus, each of the 20 batched XCMS chromatographic peak-area data matrices consisted of N_b metabolite features (where $b = 1…20$; with N_b associated m/z and retention times) × 85 samples (60 subjects plus 25 integrated QC samples). GC–MS data were deconvolved and matched to a reference database of 259 metabolites applying ChromaTof (Leco) separately for each analytical batch. This produced 20 chromatographic peak-area data matrices of 259 metabolite features (with associated EI-MS spectrum and retention index) × 80 samples (60 subjects plus 20 integrated QC samples). If a given metabolite was not detected in a given batch then the associated matrix element was replaced with a missing value (NaN; not-a-number).

3.6 Quality assurance, signal correction, batch integration and metabolite identification

For both the GC-Tof–MS and UPLC–MS instrumentation, analytical reproducibility had to be assessed robustly to ensure that data were of comparable high quality within and between analytical batches. The use of periodic analysis of a standard, biologically identical, QC sample within and across all batches, and subsequent statistical assessment of individual peak area variation within and between batches is now highly recommended as a standard quality assurance strategy in metabolite profiling (Dunn et al. 2012). Following preliminary studies (for example, Begley et al. 2009) it has been determined that a tolerance of 20 % RSD for UPLC–MS and 30 % RSD for GC-Tof–MS are acceptable guidelines. Peaks that did not meet acceptable quality thresholds were removed prior to further data analysis. For this study each of the 20 batches was assessed individually, and then data for peaks of high quality were matched across batches. Additionally, it has been shown that for both GC-Tof–MS and UPLC–MS instrumentation there is time dependent non-linear peak area attenuation for many detected metabolite features within a given batch (Begley et al. 2009; Zelena et al. 2009). This problem is compounded with the use of multiple batches, where step changes in instrument sensitivity may be expected. As a pre-processing countermeasure against these phenomena each metabolite feature of a given experimental batch, after XCMS deconvolution, was normalised to the QC sample using robust Locally Weighted Scatterplot Smoothing (LOESS) signal correction (QC-RLSC). Here LOESS was performed on the QC data with respect to the order of injection. A cubic spline correction curve for the whole analytical run was then interpolated, to which the total data set for that peak was normalized. Using this procedure any attenuation of peak response over an analytical run (i.e. confounding factor due to injection order) was minimized, whilst robustly avoiding fitting the correction curve to random measurement error. Normalizing to the QC correction curve also allowed simple data concatenation of high-quality metabolite features across multiple batches. Once combined into a single multi-batch data matrix, each metabolite feature was un-normalized using the overall estimation of expected QC peak area (in this case the median peak area across all batches). Comprehensive details of the quality assurance, signal correction, and batch integration have been described previously (Dunn et al. 2011). For this study a total of 259, 7813 and 7914 unique metabolic features were present in the raw data for GC–MS, UPLC–MS+ and UPLC–MS− respectively. After signal correction, quality assurance, and batch integration there were 126, 2181 and 2283 metabolite features available for further statistical analysis. Each of these features was present in a minimum of 80 % of the samples analyzed. Identification and annotation of metabolites was performed as described previously (Dunn et al. 2011). For UPLC–MS data, the accurate measurement of m/z followed by grouping of different metabolite features

based on retention time similarity, response correlation and expected m/z differences and the matching of the defined molecular formula for each group of features to those present in a revised MMD database was performed (Brown et al. 2011). For UPLC–MS all metabolite identifications are reported as level 2 (metabolite reported) or level 4 (no metabolite reported) according to the recommendations of the Chemical Analysis Group of the Metabolomics Standards Initiative (MSI) (Sumner et al. 2007). For GC–MS, the electron impact (EI) mass spectrum and retention index were compared to either an in-house EI mass spectral library constructed with authentic chemical standards or other available EI mass spectral libraries (NIST05, Golm Metabolome Database (Kopka et al. 2005)). For GC–MS all metabolites are either identified (MSI level 1; if matched to a metabolite in the in-house library which was constructed applying the same analytical conditions), annotated (MSI level 2; mass spectrum matched to NIST05 or Golm Metabolome Database) or unidentified (MSI level 4).

3.7 Data availability

All metabolite data and associated demographic/clinical metadata are available at the publically available metabolomics data repository MetaboLights (http://www.ebi.ac.uk/metabolights/; study identifier MTBLS97).

3.8 Data analysis

All data analysis follows MSI reporting guidelines (Goodacre et al. 2007). The data from each platform was integrated into single data matrix of 1,187 subjects by 4,261 metabolite features. There were a maximum of 20 % missing values for each metabolite feature and missing values were imputed using the mean value for a given metabolite feature for all subjects. Before statistical analysis each metabolite feature was autoscaled (normalized to unit variance). Initially, for each metabolite feature in turn, the distributions of the classification groups in a given clinical hypothesis (Age; Gender; BMI; Blood Pressure; Smoking) were compared using either the non-parametric Mann–Whitney U test, or Kruskal–Wallis test, depending on the number of groups in the comparison. Additionally, 2-way ANOVA was performed to investigate interactions between clinical variables with respect to metabolite relative concentrations. For all reported 2-way ANOVA results data normality (approximate) was checked, and assured, using Q–Q plots (data not shown).

In order to reduce the high dimensional data set down to a manageable, size a consensus feature selection protocol was implemented for each clinical hypothesis. In this protocol three modeling techniques were utilized: (1) Non-parametric univariate hypothesis testing (as described above), (2) Random Forests (RF) (Breiman 2001) and (3) Partial Least Squares Discriminant Analysis (PLS-DA) (Wold et al. 2001). For a given classification problem, and associated data set, each of these modeling techniques provided a ranked list of metabolite features in order of importance. In order to avoid model over-fitting, and possible false discovery, bootstrap resampling was performed for each modeling technique (Efron and Tibshirani 1993). For both classification and feature selection, 100 bootstrap resamplings (with replacement) were made. The resulting ranked lists of features were averaged using the Borda count consensus voting system (Dwork 2001), resulting in a single aggregated ranked list of metabolite importance. The optimal subset of metabolites for each clinical hypothesis was then found from this rank list using forward selection remodeling. Starting with the most important feature, and adding the next important feature one at a time, a series of classification models were built and associated classification accuracy tested (Cho et al. 2004). The optimal number of metabolite features was at the inflection point in the curve of classification accuracy versus the number of features. On average, across all the clinical hypotheses tested, the inflection point was found at 30 metabolite features with accuracy slightly above 75 %, shown in Supplementary Fig. 6. Therefore we used 30 metabolite features found in GC–MS, positive UPLC–MS and negative UPLC–MS for further annotation analysis in this study. To assess the effectiveness of the feature selection, we applied two classifiers, Random Forest (RF) and Support Vector Machines (SVM) (Cristianini and Shawe-Taylor 2000) to discriminate the categorical groups: age (age <50 and age >65), BMI (BMI <25 and BMI >30) and gender (male and female). A bootstrap re-sampling method was employed to evaluate the performances of the two classifiers. The results shown in Supplementary Fig. 7 reveal that the discrimination with feature selection is much better than those without feature selection, especially for both positive and negative LC-MS data sets.

All annotated metabolites were analysed further by Pearson correlation analysis. We applied two correlation analyses, one between identified metabolite pairs and another between identified metabolites and clinical chemistry data. To visualise the correlation results, we used a heatmap of correlation coefficients. We also applied a hierarchical clustering technique to re-order the correlation coefficients in the heatmap, to highlight the relationship between the variables used.

For large-scale studies of the human population sample size is very important and we therefore studied sample size effects in both classification and feature selection. Selecting sample size ranges varying from 50 to 650 (in steps of 50), we again classified three groups on the basis of age,

BMI and gender for three analytical platforms by two classifiers, viz. RF and SVM with 100 bootstrap sample sets. Using the same sample size ranges as for classification, the feature selections were performed using three methods: Wilcoxon test, RF and PLS, combined with a bootstrap re-sampling technique. To examine sample size effects for feature selection, we used correlation analysis to validate the consistency of feature selection on the sample subsets with sample size changing from 50 to 650. The correlation analysis was performed on the aggregated full ranking lists obtained from the three feature selections.

References

Bachi, A., Dalle-Donne, I., & Scaloni, A. (2013). Redox proteomics: Chemical principles, methodological approaches and biological/biomedical promises. *Chemical Reviews, 113*, 596–698. doi:10.1021/cr300073p.

Bathe, O. F., et al. (2011). Feasibility of identifying pancreatic cancer based on serum metabolomics. *Cancer Epidemiology, Biomarkers and Prevention, 20*, 140–147. doi:10.1158/1055-9965.EPI-10-0712.

Begley, P., et al. (2009). Development and performance of a gas chromatography–time-of-flight mass spectrometry analysis for large-scale non-targeted metabolomic studies of human serum. *Analytical Chemistry, 81*, 7038–7046.

Bouatra, S., et al. (2013). The human urine metabolome. *PLoS One, 8*, e73076. doi:10.1371/journal.pone.0073076.

Breiman, L. (2001). Random forests. *Machine Learning, 45*, 5–32.

Broadhurst, D., & Kell, D. B. (2006). Statistical strategies for avoiding false discoveries in metabolomics and related experiments. *Metabolomics, 2*, 171–196. doi:10.1007/s11306-006-0037-z.

Brown, M., et al. (2009). Mass spectrometry tools and metabolite-specific databases for molecular identification in metabolomics. *Analyst, 134*, 1322–1332.

Brown, M., et al. (2011). Automated workflows for accurate mass-based putative metabolite identification in LC/MS-derived metabolomic datasets. *Bioinformatics, 27*, 1108–1112. doi:10.1093/bioinformatics/btr079.

Brownlee, M. (2001). Biochemistry and molecular cell biology of diabetic complications. *Nature, 414*, 813–820. doi:10.1038/414813a.

Button, K. S., et al. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience, 14*, 365–376. doi:10.1038/nrn3475.

Camacho, D., de la Fuente, A., & Mendes, P. (2005). The origins of correlations in metabolomics data. *Metabolomics, 1*, 53–63.

Cheng, S. S., et al. (2012). Metabolite profiling identifies pathways associated with metabolic risk in humans. *Circulation, 125*, 2222–2231. doi:10.1161/Circulationaha.111.067827.

Cho, J. H., Lee, D., Park, J. H., & Lee, I. B. (2004). Gene selection and classification from microarray data using kernel machine. *FEBS Letters, 571*, 93–98. doi:10.1016/j.febslet.2004.05.087.

Cristianini, N., & Shawe-Taylor, J. (2000). *An introduction to support vector machines and other kernel-based learning methods.* Cambridge: Cambridge University Press.

Delpierre, G., & Van Schaftingen, E. (2003). Fructosamine 3-kinase, an enzyme involved in protein deglycation. *Biochemical Society Transactions, 31*, 1354–1357.

Dunn, W. B., Broadhurst, D. I., Atherton, H. J., Goodacre, R., & Griffin, J. L. (2011a). Systems level studies of mammalian metabolomes: The roles of mass spectrometry and nuclear magnetic resonance spectroscopy. *Chemical Society Reviews, 40*, 387–426.

Dunn, W. B., Wilson, I. D., Nicholls, A. W., & Broadhurst, D. (2012). The importance of experimental design and QC samples in large-scale and MS-driven untargeted metabolomic studies of humans. *Bioanalysis, 4*, 2249–2264. doi:10.4155/bio.12.204.

Dunn, W. B., et al. (2011b). Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. *Nature Protocols, 6*, 1060–1083.

Dwork, C., Kumar, R., Naor, M., Sivakumar, D. (2001). Rank aggregation methods for the Web. In *Proceedings of the 10th International World Wide Web Conference* (pp. 613–622).

Efron, B., & Tibshirani, R. J. (1993). *Introduction to the bootstrap.* London: Chapman and Hall.

Friend, A., Craig, L., & Turner, S. (2013). The prevalence of metabolic syndrome in children: A systematic review of the literature. *Metabolic Syndrome and Related Disorders, 11*, 71–80. doi:10.1089/met.2012.0122.

Gavaghan, C. L., Holmes, E., Lenz, E., Wilson, I. D., & Nicholson, J. K. (2000). An NMR-based metabonomic approach to investigate the biochemical consequences of genetic strain differences: Application to the C57BL10J and Alpk: ApfCD mouse. *FEBS Letters, 484*, 169–174.

Goodacre, R., et al. (2007). Proposed minimum reporting standards for data analysis in metabolomics. *Metabolomics, 3*, 231–241.

Haug, K., et al. (2013). MetaboLights-an open-access general-purpose repository for metabolomics studies and associated meta-data. *Nucleic Acids Research, 41*, D781–D786. doi:10.1093/Nar/Gks1004.

He, J., & Whelton, P. K. (1999). Elevated systolic blood pressure and risk of cardiovascular and renal disease: Overview of evidence from observational epidemiologic studies and randomized controlled trials. *American Heart Journal, 138*, 211–219.

Holmes, E., et al. (2008). Human metabolic phenotype diversity and its association with diet and blood pressure. *Nature, 453*, 396–400. doi:10.1038/nature06882.

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine, 2*, e124.

Ioannidis, J. P. A., & Panagiotou, O. A. (2011). Comparison of effect sizes associated with biomarkers reported in highly cited individual articles and in subsequent meta-analyses. *JAMA, 305*, 2200–2210.

Kell, D. B. (2004). Metabolomics and systems biology: Making sense of the soup. *Current Opinion in Microbiology, 7*, 296–307.

Kell, D. B. (2006a). Metabolomics, modelling and machine learning in systems biology: Towards an understanding of the languages of cells. The 2005 Theodor Bücher lecture. *FEBS Journal, 273*, 873–894.

Kell, D. B. (2006b). Systems biology, metabolic modelling and metabolomics in drug discovery and development. *Drug Discovery Today, 11*, 1085–1092.

Kell, D. B. (2009). Iron behaving badly: Inappropriate iron chelation as a major contributor to the aetiology of vascular and other progressive inflammatory and degenerative diseases. *BMC Medical Genomics, 2*, 2.

Kell, D. B. (2013). Finding novel pharmaceuticals in the systems biology era using multiple effective drug targets, phenotypic screening, and knowledge of transporters: Where drug discovery went wrong and how to fix it. *FEBS Journal, 280*, 5957–5980.

Kell, D. B., Dobson, P. D., Bilsland, E., & Oliver, S. G. (2013). The promiscuous binding of pharmaceutical drugs and their transporter-mediated uptake into cells: What we (need to) know and how we can do so. *Drug Discovery Today, 18*, 218–239.

Kell, D. B., & Goodacre, R. (2014). Metabolomics and systems pharmacology: Why and how to model the human metabolic network for drug discovery. *Drug Discovery Today, 19*, 171–182.

Kell, D. B., & Oliver, S. G. (2004). Here is the evidence, now what is the hypothesis? The complementary roles of inductive and hypothesis-driven science in the post-genomic era. *BioEssays, 26*, 99–105.

Kell, D. B., & Westerhoff, H. V. (1986). Metabolic control theory: Its role in microbiology and biotechnology. *FEMS Microbiology Reviews, 39*, 305–320.

Kochhar, S., Jacobs, D. M., Ramadan, Z., Berrux, F., Fuerholz, A., & Fay, L. B. (2006). Probing gender-specific metabolism differences in humans by nuclear magnetic resonance-based metabonomics. *Analytical Biochemistry, 352*, 274–281. doi:10.1016/j.ab.2006.02.033.

Kopka, J., et al. (2005). GMD@CSB.DB: The Golm metabolome database. *Bioinformatics, 21*, 1635–1638.

Kotze, H. L., et al. (2013). A novel untargeted metabolomics correlation-based network analysis incorporating human metabolic reconstructions. *BMC Systems Biology, 7*, 107. doi:10.1186/1752-0509-7-107.

Lawton, K. A., et al. (2008). Analysis of the adult human plasma metabolome. *Pharmacogenomics, 9*, 383–397. doi:10.2217/14622416.9.4.383.

Lee, D. M., et al. (2009). Vitamin D, parathyroid hormone and the metabolic syndrome in middle-aged and older European men. *European Journal of Endocrinology, 161*, 947–954. doi:10.1530/EJE-09-0496.

Liebich, H. M., & Pickert, A. (1985). Gas chromatographic profiling of phenolic acids in urine of patients with cirrhosis of the liver. *Journal of Chromatography, 338*, 25–32.

Lu, L., et al. (2009). Plasma 25-hydroxyvitamin D concentration and metabolic syndrome among middle-aged and elderly Chinese individuals. *Diabetes Care, 32*, 1278–1283. doi:10.2337/Dc09-0209.

Mao, Y. Y., et al. (2008). A pilot study of GC/MS-based serum metabolic profiling of acute rejection in renal transplantation. *Transplant Immunology, 19*, 74–80. doi:10.1016/j.trim.2008.01.006.

Menni, C., et al. (2013). Metabolomic markers reveal novel pathways of ageing and early development in human populations. *International Journal of Epidemiology, 42*, 1111–1119. doi:10.1093/ije/dyt094.

Newgard, C. B., et al. (2009). A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans

and contributes to insulin resistance. *Cell Metabolism, 9*, 311–326. doi:10.1016/j.cmet.2009.02.002.

Nicholson, J. K., Everett, J. R., & Lindon, J. C. (2012). Longitudinal pharmacometabonomics for predicting patient responses to therapy: Drug metabolism, toxicity and efficacy. *Expert Opinion on Drug Metabolism & Toxicology, 8*, 135–139. doi:10.1517/17425255.2012.646987.

Nicholson, G., et al. (2011). Human metabolic profiles are stably controlled by genetic and environmental variation. *Molecular Systems Biology, 7*, 525.

Pradhan, A. (2007). Obesity, metabolic syndrome, and type 2 diabetes: Inflammatory basis of glucose metabolic disorders. *Nutrition Reviews, 65*, S152–S156.

Psychogios, N., et al. (2011). The human serum metabolome. *PLoS One, 6*, e16957. doi:10.1371/journal.pone.0016957.

Raamsdonk, L. M., et al. (2001). A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations. *Nature Biotechnology, 19*, 45–50.

Rupérez, F. J., et al. (2012). Metabolomic study of plasma of patients with abdominal aortic aneurysm. *Analytical and Bioanalytical Chemistry, 403*, 1651–1660. doi:10.1007/s00216-012-5982-y.

Sealey, W. M., Teague, A. M., Stratton, S. L., & Mock, D. M. (2004). Smoking accelerates biotin catabolism in women. *American Journal of Clinical Nutrition, 80*, 932–935.

Slupsky, C. M., et al. (2007). Investigations of the effects of gender, diurnal variation, and age in human urinary metabolomic profiles. *Analytical Chemistry, 79*, 6995–7004. doi:10.1021/ac0708588.

Suhre, K., & Gieger, C. (2012). Genetic variation in metabolic phenotypes: Study designs and applications. *Nature Reviews Genetics, 13*, 759–769. doi:10.1038/Nrg3314.

Suhre, K., et al. (2011). Human metabolic individuality in biomedical and pharmaceutical research. *Nature, 477*, 54–60.

Sumner, L. W., et al. (2007). Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics, 3*, 211–221. doi:10.1007/s11306-007-0082-2.

Swainston, N., Mendes, P., & Kell, D. B. (2013). An analysis of a 'community-driven' reconstruction of the human metabolic network. *Metabolomics, 9*, 757–764.

Thiele, I., et al. (2013). A community-driven global reconstruction of human metabolism. *Nature Biotechnology, 31*, 419–425.

van der Greef, J., Hankemeier, T., & McBurney, R. N. (2006). Metabolomics-based systems biology and personalized medicine: Moving towards n = 1 clinical trials? *Pharmacogenomics, 7*, 1087–1094.

Vidal, M., Cusick, M. E., & Barabási, A. L. (2011). Interactome networks and human disease. *Cell, 144*, 986–998.

Wang, J., & Li, M. D. (2010). Common and unique biological pathways associated with smoking initiation/progression, nicotine dependence, and smoking cessation. *Neuropsychopharmacology, 35*, 702–719. doi:10.1038/npp.2009.178.

Wang, T. J., et al. (2011). Metabolite profiles and the risk of developing diabetes. *Nature Medicine, 17*, 448–453.

Wold, S., Sjöström, M., & Eriksson, L. (2001). PLS-regression: A basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems, 58*, 109–130. doi:10.1016/S0169-7439(01)00155-1.

Xia, J. G., Broadhurst, D. I., Wilson, M., & Wishart, D. S. (2013). Translational biomarker discovery in clinical metabolomics: An introductory tutorial. *Metabolomics, 9*, 280–299. doi:10.1007/s11306-012-0482-9.

Yap, I. K. S., et al. (2010). Metabolome-wide association study identifies multiple biomarkers that discriminate north and south Chinese populations at differing risks of cardiovascular disease: INTERMAP study. *Journal of Proteome Research, 9*, 6647–6654. doi:10.1021/pr100798r.

Yu, Z. H., et al. (2012). Human serum metabolic profiles are age dependent. *Aging Cell, 11*, 960–967. doi:10.1111/j.1474-9726.2012.00865.x.

Zelena, E., et al. (2009). Development of a robust and repeatable UPLC–MS method for the long-term metabolomic study of human serum. *Analytical Chemistry, 81*, 1357–1364.