Published by Oxford University Press on behalf of the International Epidemiological Association

International Journal of Epidemiology 2008;**37**:i23–i30
doi:10.1093/ije/dym281

A GC-TOF-MS study of the stability of serum and urine metabolomes during the UK Biobank sample collection and preparation protocols

Warwick B Dunn,[1,2]* David Broadhurst,[1] David I Ellis,[1] Marie Brown,[1] Anthony Halsall,[1] Steven O'Hagan,[1] Irena Spasic,[1,2] Andrew Tseng[1] and Douglas B Kell[1,2]

Accepted 10 December 2007

Background The stability of mammalian serum and urine in large metabolomic investigations is essential for accurate, valid and reproducible studies. The stability of mammalian serum and urine, either processed immediately by freezing at −80°C or stored at 4°C for 24 h before being frozen, was compared in a pilot metabolomic study of samples from 40 separate healthy volunteers.

Methods Metabolic profiling with GC-TOF-MS was performed for serum and urine samples collected from 40 volunteers and stored at −80°C or 4°C for 24 h before being frozen at −80°C. Subsequent Wilcoxon rank sum test and Principal Components Analysis (PCA) methods were used to assess whether differences in the metabolomes were detected between samples stored at 4°C for 0 or 24 h.

Results More than 700 unique metabolite peaks were detected, with over 200 metabolite peaks detected in any one sample. PCA and Wilcoxon rank sum tests of serum and urine data showed as a general observation that the variance associated with the replicate analysis per sample (analytical variance) was of the same magnitude as the variance observed between samples stored at 4°C for 0 or 24 h. From a functional point of view the metabolomic composition of the majority of samples did not change in a statistically significant manner when stored under two different conditions.

Conclusions Based on this small pilot study, the UK Biobank sampling, transport and fractionation protocols are considered suitable to provide samples, which can produce scientifically robust and valid data in metabolomic studies.

Keywords Metabolomics, metabolic profiling, GC-MS, univariate analysis, multivariate analysis, biofluid, serum, urine

Introduction

Life is structured on many levels of biological organization. Only in the current post-genomic era, after sequencing of many genomes (http://genomesonline.org/), has the complexity of biological organization at the cellular level become fully recognized. To understand the function and dysfunction of such complex systems requires integrated, systems-level approaches.[1,2] The UK Biobank opens up the opportunity, through the provision of a collection of biofluid and tissue samples representative of the UK population aged 40–69, to study the interaction of the 'omes' (including the genomic and expression levels) and the environment for the integrated study of health and disease in humans. This can be undertaken by studying single 'omic'

[1] Bioanalytical Sciences Group, School of Chemistry, Manchester Interdiscplinary Biocentre, University of Manchester, 131 Princess Street, Manchester, M1 7ND, UK.

[2] Manchester Centre for Integrative Systems Biology, Manchester Interdiscplinary Biocentre, University of Manchester, 131 Princess Street, Manchester, M1 7ND, UK.

* Corresponding author. Manchester Centre for Integrative Systems Biology, Manchester Interdisciplinary Biocentre, University of Manchester, 131, Princess St, MANCHESTER M1 7DN, UK. E-mail: Warwick.Dunn@manchester.ac.uk

functional levels (genome, transcriptome, proteome and metabolome), or alternatively in an integrative systems biology approach quantitatively studying each 'ome' and the inter-ome connectivities.

We have a special interest here in the metabolome, which is defined as the quantitative collection of low molecular weight (MW) compounds present in a biological system, which participate in metabolic reactions required for growth, maintenance and normal function.[3,4] It is considered the optimal level at which to study a cell or organism,[5] since the founders of Metabolic Control Analysis (MCA) long ago proved that while changes in the expression level of individual genes and proteins normally have only marginal effects on fluxes, they can and do have very large changes in the concentrations of metabolic intermediates[6,7] (as measured by the metabolome). Operating at the functional level of the cell, and as the final 'downstream' product of gene expression, the metabolome reflects both genotypic[8] and phenotypic information.[9] The study of the metabolome (termed metabolomics) is the integration of many processes in a complex workflow, in effect a metabolome pipeline,[10] consisting of experimental design, sample collection and preparation, analytical operations, data processing and data storage. Data produced from large-scale metabolomic studies are extensive and require structured storage of both the data and meta-data.[10–12] Powerful multivariate analyses, either supervised or unsupervised, are required to interrogate the data and define structure related to biological similarities or differences of the system.[5,10,13] Appropriate experimental design, sample numbers and statistical analyses are required to ensure generation of accurate and valid hypotheses and biological conclusions.[14] In many metabolomics studies the initial objective is that of an inductive, data-driven, rather than hypothesis-driven, approach to obtain the maximum overview of the biological landscape before hypothesis generation and testing.[15] This is not least because of the fact that it is hard to make hypotheses about molecules one does not even know may exist as they have not been detected or reported (at all or in the relevant system).

A number of analytical strategies are employed in the investigation of the metabolome,[16,17] with its study being analytically challenging because of its size (the human metabolome is estimated to contain at least 1500 metabolites not including gut microbiota derived metabolites)[18,19] and large variations in chemical and physical properties (polarity, reactivity, volatility, MW) and concentrations (sub pM-mM). This complexity negatively influences the overall objective of metabolomics, the unbiased quantification and identification of all metabolites in a biological system. Currently no metabolomic strategy or technology can fulfil this goal. An approach employing metabolic profiling (semi or fully quantitative detection of 10s–1000s of metabolites of differing functionalities present in many different biological pathways), in combination with the use of multiple analytical platforms, is preferred if one is to obtain the maximum metabolomic information. The goal of quantifying the complexity and variation of the human serum metabolome is currently being attacked by the HUSERMET project at The University of Manchester (www.husermet.org) using the approach described above.

Of the different analytical technologies employed in metabolomics,[17] chromatography-mass spectrometry systems are considered most favourable for the detection of large numbers of metabolites by providing partial or full separation of metabolites by chromatographic systems prior to sensitive mass spectrometry (MS) detection, that also provides the opportunity to identify metabolites. GC-TOF-MS,[20] GCxGC-TOF-MS,[21] HPLC-MS and the analytically superior UPLC-MS[22] and CE-MS[23] have all been employed in mammalian metabolomic studies, either in a metabolic profiling or targeted analysis approach. Gas chromatography provides the separation of volatile and thermally stable metabolites, generally after chemical derivatization to induce these properties, and has provided the detection of over 900 raw peaks in pooled serum samples[24] and over 4000 peaks in the analytically superior technique of comprehensive GCxGC-MS,[25] though we note that a single metabolite can produce multiple (generally 1–3) products following derivatization. These large numbers of peaks were only achievable after multi-objective optimization of analytical instrumentation, in a closed-loop approach[24,25] requiring minimal human intervention, showing the necessity to operate instrumentation at optimal settings to provide maximal biological information. GC-MS has been employed in metabolic profiling for many years from initial studies in the early 1970s[26] through measurements in tumour biology[27] to applications in the detection of inborn errors of metabolism[28] and the detection of diagnostic or prognostic biomarkers of disease.[20,29]

Mammalian serum and urine are extra-cellular biofluids (the 'exometabolome'), reflecting a picture of metabolic activity at a specific time (serum or plasma) or over a period of time (urine), and can be considered as a kind of metabolic 'snapshot'[7] or 'footprint', similar to the metabolic footprints of yeast,[30] that integrates the metabolic responses of multi-cellular systems. The strategies for sampling and storage of biofluids for metabolomic studies are especially important in comparison with the proteome and transcriptome, as metabolic activity is significantly more rapid (intracellular metabolic reaction half lives are often <1 s). Metabolic activity during sampling and storage requires stopping, or quenching, and further metabolic activity minimized to stop changes in the metabolic profile or to minimize chemical or biologically induced changes

of metabolites either in concentration or structure. To this end, reduced temperatures during sample preparation (4°C) and storage (−80°C) are normally employed.

The study described here, employing metabolomic analytical and data processing strategies, assesses the stability of serum and urine samples of 40 healthy human volunteers using the proposed sample collection, transport and fractionation protocols of the UK Biobank, and in particular assesses whether the process yields samples which are fit for purpose for metabolomic investigations. Specifically, the study was designed using paired samples to determine whether serum and urine samples retain the same metabolomic composition when stored at 4°C for either 0 ($T=0$ h) or 24 ($T=24$ h) hours prior to freezing and transportation. The variance observed by replicate analyses of a single sample (analytical variance) was compared with the variance observed between replicates of paired samples ($T=0$ h and $T=24$ h samples) to assess whether changes in metabolomic composition occurred under the different storage conditions.

Methods

Sample collection and storage

Blood and urine samples were collected from 40 volunteers at room temperature.[31] Blood samples were split into two separate aliquots; one aliquot was allowed to clot for 30 min (4°C) before the serum fraction was separated and frozen immediately at −80°C ($T=0$ h samples). The second aliquot was stored at 4°C for 24 h before the serum fraction was separated and frozen at −80°C ($T=24$ h samples). Urine samples were separated into two aliquots; one aliquot was frozen immediately at −80°C ($T=0$ h samples) and the other stored at 4°C for 24 h before being frozen at −80°C ($T=24$ h samples). Frozen samples were transported in dry ice to Manchester and stored at −80°C.

Sample preparation

Serum and urine samples were thawed and stored on ice (4°C) and prepared in a similar approach as previously described.[29] The full sample preparation protocol is available in Supplementary data A.

GC-TOF-MS analysis

Lyophilized samples were chemically derivatized and analysed by GC-TOF-MS (Agilent 6890 gas chromatograph and Leco Pegasus III time-of-flight mass spectrometer) as previously described.[24] The complete sample derivatization protocol is available in Supplementary data B. All samples were randomised and three technical replicates were analysed within 24 h of chemical derivatization completion.

Raw data were processed using LECO ChromaTof v2.12 and its associated chromatographic deconvolution algorithm, with the baseline set at 1.0, data point averaging of 5 and average peak width of 2.5. A reference database was prepared, incorporating the mass spectrum and retention index of all metabolite peaks detected in a random selection of 10 serum and 10 urine samples so to allow detection of all metabolites present, whether expected or not expected from the study of metabolic pathways. Each metabolite peak in the reference database was searched for in each urine and serum sample and if matched (retention index deviation $<\pm 10$; mass spectral match >750) the peak area was reported and the response ratio relative to the internal standard (peak area-metabolite/peak area-succinic d_4 acid internal standard) calculated. These data (matrix of N samples $\times$ P metabolite peaks) representing normalized peak lists were exported in ASCII format for further analysis.

Multivariate and univariate data analysis

Many multivariate data reduction and analysis techniques are available for studying metabolomic datasets.[5,10,13,16] Principal Components Analysis (PCA) is an unsupervised technique that assumes no *a priori* knowledge of class structure (in this study samples stored at 4°C for 0 or 24 h) and acts to reduce the dimensionality of multivariate data whilst attempting to preserve as much of the cross-dimensional variance as possible. PCA reduces the dimensionality of the multivariate data and projects the resulting data into a lower dimensional space, usually two. These dimensions are termed the principal components. Those samples that are observed in close proximity (small between-sample distances) in PCA space can be described as metabolically similar when compared with samples that occupy widely different positions. Each subject was represented by six spectra [2 time points ($T=0$ h, $T=24$ h) $\times$ 3 technical replicates]. All data sets were column normalized to median absolute deviation.

Univariate statistical analysis was performed in order to assess the characteristics of each independent peak within the metabolomic fingerprint. As there is no guarantee of normality in the data a non-parametric, rank-based analysis of variance was used (Wilcoxon rank sum test[32]). For each metabolite, the null hypothesis that the sample concentrations at each storage temperature came from populations with the same mean was tested. Generally if a P-value of <0.01 is calculated then the two sample populations for that metabolite are deemed to be significantly different; however when many parallel tests are performed care has to be taken regarding type I errors (i.e. falsely rejecting the null hypothesis). To help reduce the possibility of type I errors the modified critical P-value, α, was modified using a Bonferroni correction which sets the P-value for the

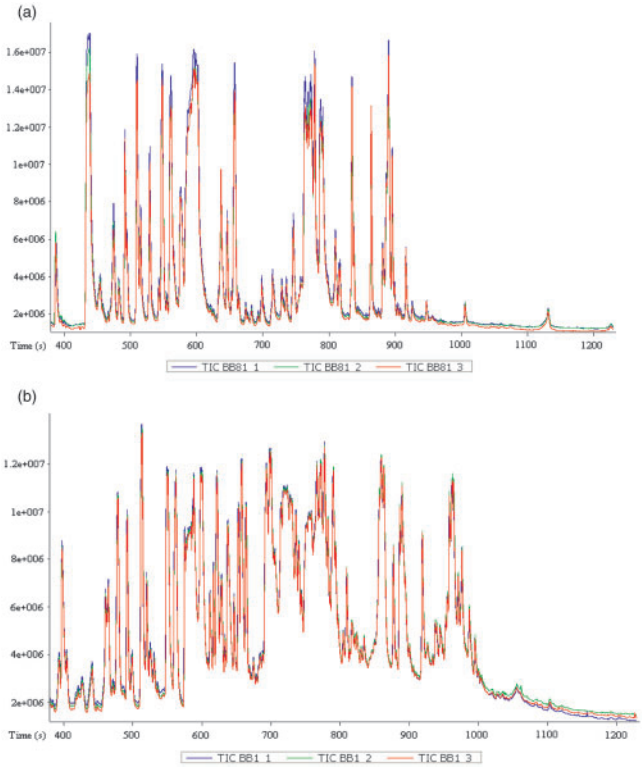

Figure 1 Typical GC-TOF-MS total ion current chromatograms for three technical replicates of a single volunteer for (**a**) serum and (**b**) urine

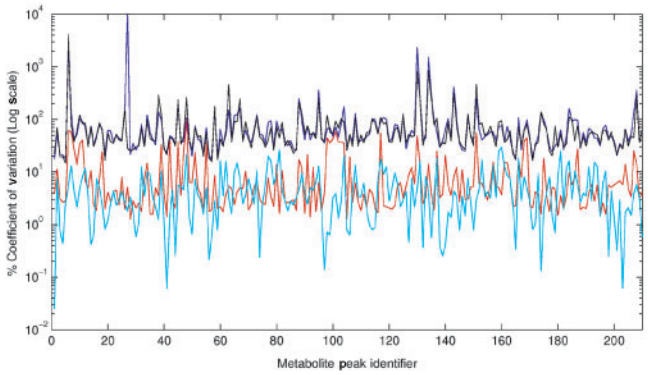

Figure 2 Variance observed in the dataset for serum. The abscissa shows the metabolite peak identifier and ordinate shows the coefficient of variance (non-parametric) depicting (i) analytical variance (red); (ii) inter-subject variability for $T = 0$ h samples (black); (iii) inter-subject variability for $T = 24$ h samples (blue); (iv) variability between the two class medians (cyan). Serum shows a smaller analytical variance when compared with inter-subject variance (for a given class) across the 210 peaks that were detected in the majority of samples (i.e. peaks with <50% missing values). The coefficient of variation between the two class medians is of the same order of magnitude as the analytical coefficient of variation, indicating that there is little measurable difference between the two sample populations ($T = 0$ h and $T = 24$ h)

entire set of n comparisons equal to α; thus, requiring the modified P-value value for each comparison to be α/n.[14] With ~200 statistically usable peaks detected for both the urine and serum experiments, the Bonferroni modified critical P-value was set conservatively to 0.01/200, or, 5×10^{-5}.

Results

GC-TOF-MS is a powerful analytical tool employed in metabolomic studies, providing in this study the detection of 706 different metabolite peaks with over 210 peaks detected in each sample. Typical total ion current (TIC) chromatograms for urine and serum are shown in Figure 1. Definitive identification of metabolite peaks is highly recommended and uses two orthogonal properties of the metabolite, retention time (related to volatility/polarity) and mass spectrum (related to metabolite structure). Matching of mass spectra and retention indices of peaks in a sample to those available in mass spectral/retention index libraries prepared by analysis of pure (authentic) chemical standards with the same analytical conditions as for the analysis of samples provides a higher accuracy of identification. Preliminary identification can be undertaken by matching of mass spectra only using commercially or publicly available mass spectral libraries or laboratory-prepared mass spectral libraries, but requires subsequent definitive identification. Supplementary data C details all metabolites detected including both endogenous metabolites (amino acids, organic acids, sugars, sugar alcohols and lipids) and various exogenous pharmaceutical-related metabolites (paracetamol, diethazine and dothiepin). This shows the power of the techniques to detect both endogenous and exogenous compounds. Metabolites were detected in the μM to mM range in this study.

The main objective of the study was to observe whether the metabolomic composition of serum and urine change when stored at 4°C for 24 h before being frozen at −80°C, in comparison with being frozen immediately at −80°C after sample preparation. A study was designed using paired samples: samples frozen immediately after preparation ($T = 0$ h) and samples stored at 4°C for 24 h before being frozen ($T = 24$ h).

Figure 2 details the variability observed for different classes investigated for serum. Triplicate analyses of the same derivatized sample (technical replicates, shown in red) provided typical coefficients of variations (COV) of <20% for the majority of metabolite peaks, which were detected in >50% of all samples. The reproducibility of multiple sample preparation procedures performed on the same sample is equivalent to the reproducibility of replicate analyses (data not shown). Similar levels of variance were observed when comparing the population medians for $T = 0$ h and $T = 24$ h samples (cyan), showing that there is

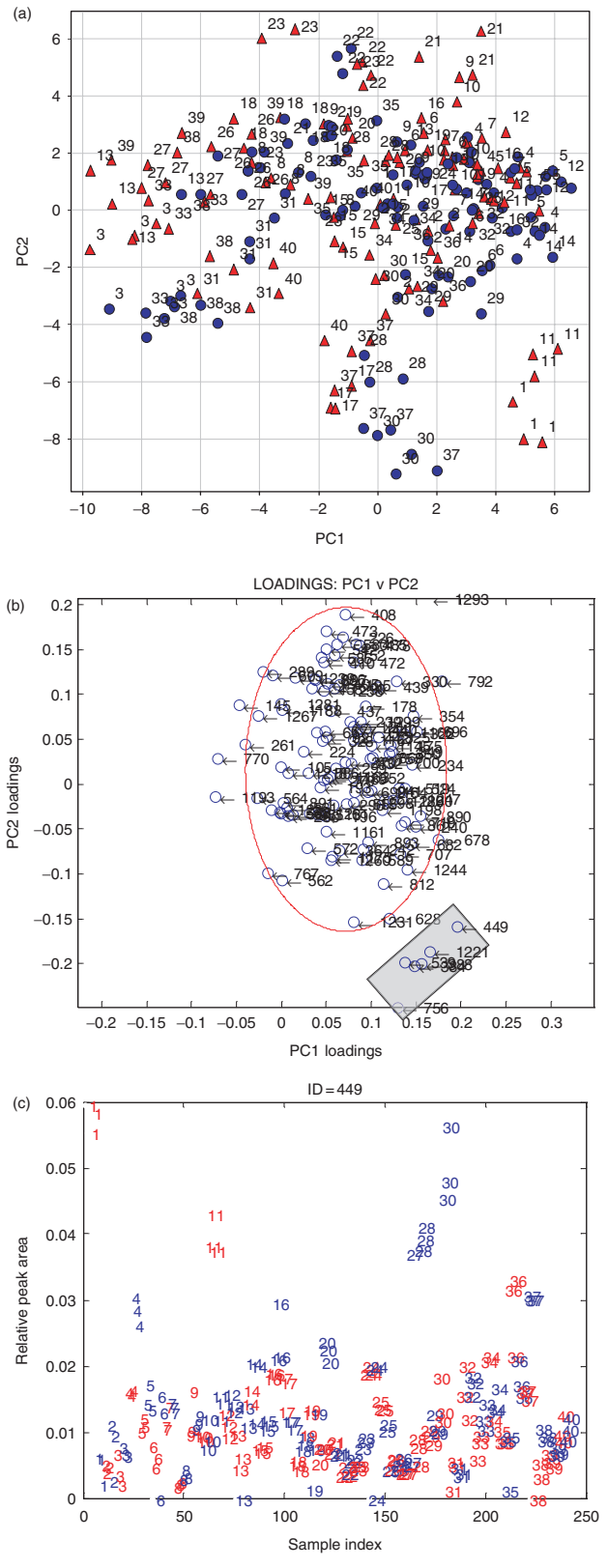

Figure 3 (**a**) Principal Components Analysis scores plot for serum for all 40 volunteers studied. Red depicts $T = 0$ h and blue depicts $T = 24$ h samples. PCs 1 and 2 represent, respectively 8 and 5% of the variance (no further clustering at higher dimensions). Labels represent subject identifiers. Each sample was analysed in triplicate and all three technical replicate analyses are shown. (**b**) Loadings plots for Principal Components 1 and 2. Shaded area indicates the metabolites responsible for the erroneous samples in the PCA scores plot. Specific numbers refer to spectral peaks. (**c**) Actual relative peak areas detected for metabolite peak 449. The order of appearance across the x-axis does not reflect order of analysis but are ordered in pairs ($T = 0$ followed by $T = 24$) with respect to subject id

no additional detectable variability when comparing $T = 0$ h and $T = 24$ h samples. Therefore these samples are metabolically similar. However, inter-volunteer variability is significantly greater (as shown by the black and blue lines) and hence the metabolomes of different volunteers are significantly more variable when compared with the variation caused by the storage of samples at 4°C for 0 and 24 h. In other words the variability due to general differences in metabolism for this human population is much greater than the variability due to storage temperature within an individual's metabolism and also much greater than the variation of the hypothetical average individual due to storage temperature (calculated by median metabolite level for each class).

In order to look at the multivariate nature of the metabolomic data (rather than splitting it into individual metabolites) PCA was used. PCA plots of serum data are shown in Figure 3, it can be observed that the distances between the three technical replicate data points within a class depicting the analytical variability (either red $T = 0$ h or blue $T = 24$ h) are on average of similar magnitude to the distances between the sample class which depict variability resulting from the two different sample storage conditions (shown by the differences between the two coloured classes for the same subject). There were no obvious trends across the whole sample population. Although only PC1 and PC2 are shown, which contain most variance, higher PCs (to PC5) also show that the effect of storage at different temperatures is minimal compared with inter-subject variability, for serum. Out of the 40 volunteers a small number are in fact separated by relatively large distance (samples 1, 11, 28 and 30). If the loadings plot is examined (Figure 3b) it can be seen that this may be explained by a small number of metabolites, relative to the large number of metabolites detected, that are outside the 2 SD zone (see shaded area). Upon closer inspection of these individual metabolites (for example peak 449 - Figure 3c) it can be seen that indeed these samples appear as outliers to a general distribution. If these distributions are tested (using either ANOVA or Wilcoxon rank sum test) the resulting P-values are above the critical level and therefore not statistically significant. As a further statistical test, univariate Wilcoxon rank sum test was performed on all of the reported metabolites using the null hypothesis reported earlier. Out of 554 tested

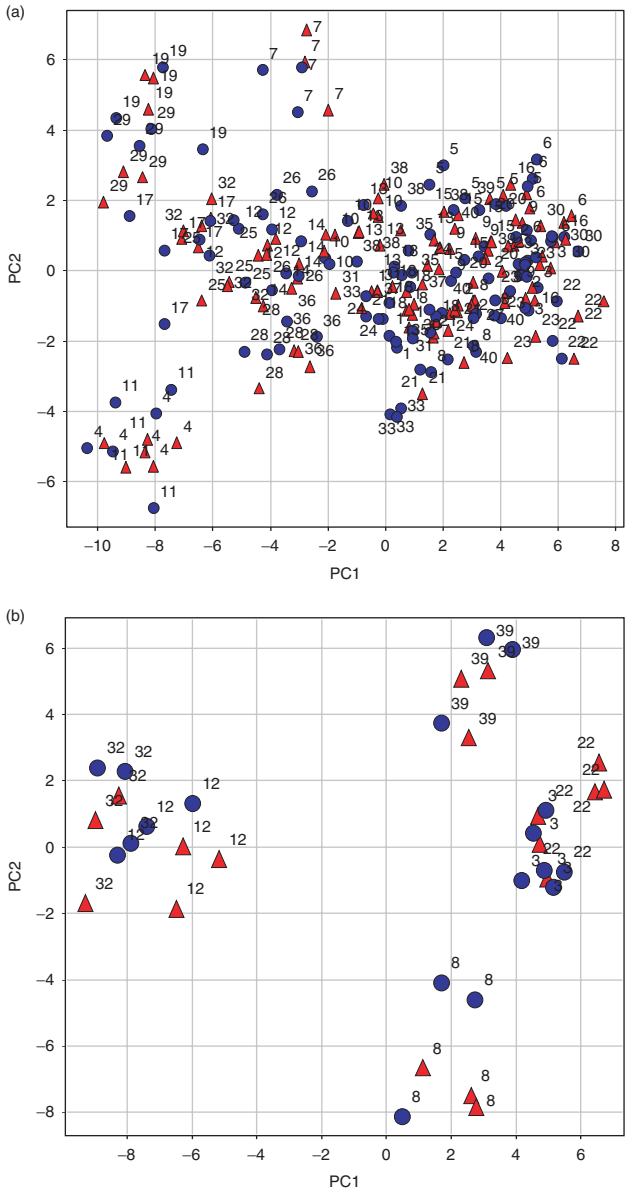

Figure 4 Principal Components Analysis plots for urine for (**a**) all 40 volunteers studied and (**b**) six randomly chosen volunteers. Red depicts $T=0$ h and blue depicts $T=24$ h. PCs 1 and 2 represent, respectively (a) 12 and 3% and (b) 19 and 8% of the variance. Labels represent subject identifiers. Each sample was analysed in triplicate and all three replicate analyses are shown

peaks, 5 produced a P-value <0.01. However, using correction for multiple tests no peak had a modified critical P-value $<5\times10^{-5}$ showing that these are probably random correlations. These data are shown in Supplementary data D.

PCA was also performed with the urine data as shown in Figure 4. Urine samples showed a similar trend to that observed for serum, where analytical variability and biological variability are of similar magnitude and are both much smaller than the variability observed between the averaged urine metabolomes of different volunteers. Again a few samples were observed as outliers, though no metabolite peak showed a modified critical P-value $<5\times10^{-5}$ using the univariate Wilcoxon rank sum test and a similar conclusion as for serum can be upheld. Although only PC1 and PC2 are shown, which contain most variance, higher PCs (to PC5) also show that the effect of storage at different temperatures is minimal compared with inter-subject variability, for urine.

Discussion

A metabolomic study, employing metabolic profiling with GC-TOF-MS, was performed to assess the stability of serum and urine when stored at 4°C for two different time periods (0 and 24 h) before being frozen at −80°C. More than 700 unique metabolite peaks (>200 peaks per sample) were detected with high reproducibility and studied, using multivariate and univariate methods, to assess possible changes in the metabolomes of these samples. No differences in the variance were observed between technical replicates, and the variance between samples stored at 4°C for 0 or 24 h were detected for serum or urine. Although PCA analysis showed a small number of samples as being metabolically different for serum and urine, univariate analysis showed these differences to be related to a small number of metabolites and not to be statistically significant. Therefore the general observation for serum and urine was minimal change in the metabolomic composition for these biofluids was detected. Importantly, when studying the intra- vs inter-subject variability for serum and urine it can easily be observed that the variance in the metabolome of a single subject stored at −80°C or 4°C for 24 h is small when compared with the variance in the metabolomes of 40 healthy volunteers and it is known that as well as different genotypes, many phenotypic factors influence the composition of the human metabolome including diet, health and lifestyle[33] and diurnal and oestrus cycles.[34] Therefore in well-designed metabolomic studies, where samples are derived from many volunteers, changes in the metabolome caused by storage at 4°C for 24 h are seen to be minimal when compared with the variability observed between subjects. In metabolomic studies involving the investigation of biomarkers of disease, drug toxicity or environmental stresses the variation between classes of healthy and diseased individual's has been shown to be even greater in many studies.[20,29,35]

In conclusion, small metabolic changes observed for a small number of metabolite peaks in this study are acceptable in metabolomic studies employing large sample sizes, where variability between subjects is greater than variability associated with sample storage and preparation. We recognise that in a study of this

type there can be a high level of false positives when sample sizes are small, and that these conclusions could be overturned for small subsets of metabolites if the sample size were to have been increased considerably.[36] We note too the absence of bias in this design.[37] Thus, on the basis of the present observations, and within the statistical limitations presented, we conclude that the UK Biobank sample collection, transport and fractionation protocols, involving the storage of serum and urine samples at 4°C for 24 hours and well-controlled UK-based transport, are suitable for high-resolution metabolomic studies.

Supplementary data

Supplementary data are available at *IJE* online.

Acknowledgements

We thank the BBSRC, EPSRC and the RSC for financial support. D.B.K. thanks the Royal Society and the Wolfson Foundation for a Research Merit Award.

Conflict of interest: None declared.

References

[1] Kell DB. Metabolomics, modelling and machine learning in systems biology - towards an understanding of the languages of cells. *FEBS J* 2006;**273:**873–94.

[2] Castrillo JO, Oliver SG. Yeast as a touchstone in post-genomic research: strategies for integrative analysis in functional genomics. *J Biochem Mol Biol* 2004;**37:**93–106.

[3] Beecher CWW. The human metabolome. In: Harrigan GG and Goodacre R (ed.). *Metabolic Profiling:Its Role in Biomarker Discovery and Gene Function Analysis*. London: Kluwer Academic Publishers, 2003. pp. 311–19.

[4] Oliver SG, Winson MK, Kell DB, Baganz F. Systematic functional analysis of the yeast genome. *Trends Biotechnol* 1998;**16:**373–78.

[5] Kell DB. Metabolomics and systems biology: making sense of the soup. *Curr Opin Microbiol* 2004;**7:**296–307.

[6] Kell DB, Westerhoff HV. Metabolic control theory: its role in microbiology and biotechnology. *FEMS Microbiol Rev* 1986;**39:**305–20.

[7] Kell DB, Mendes P. Snapshots of systems - metabolic control analysis and biotechnology in the post-genomic era. In: Cornish-Bowden A and Cárdenas ML (ed.). *Technological and Medical Implications of Metabolic Control Analysis*. Dordrecht: Kluwer Academic Publishers, 1999. pp. 3–25.

[8] Raamsdonk LM, Teusink B, Broadhurst D *et al*. A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations. *Nat Biotechnol* 2001;**19:**45–50.

[9] Catchpole GS, Beckmann M, Enot DP *et al*. Hierarchical metabolomics demonstrates substantial compositional similarity between genetically modified and conventional potato crops. *Proc Natl Acad Sci* 2005;**102:**14458–62.

[10] Brown M, Dunn WB, Ellis DI *et al*. A metabolome pipeline: from concept to data to knowledge. *Metabolomics* 2005;**1:**39–51.

[11] Lindon JC, Nicholson JK, Holmes E *et al*. Summary recommendations for standardization and reporting of metabolic analyses. *Nat Biotechnol* 2005;**23:**833–38.

[12] Spasic I, Dunn WB, Velarde G *et al*. MeMo: a hybrid SQL/XML approach to metabolomic data management for functional genomics. *BMC Bioinformatics* 2006;**7:**281.

[13] Kell DB. Metabolomics and machine learning: explanatory analysis of complex metabolome data using genetic programming to produce simple, robust rules. *Mol Biol Rep* 2002;**29:**37–41.

[14] Broadhurst D, Kell DB. Statistical strategies for avoiding false discoveries in metabolomics and related experiments. *Metabolomics* 2006;**2:**171–96.

[15] Kell DB, Oliver SG. Here is the evidence, now what is the hypothesis? The complementary roles of inductive and hypothesis-driven science in the post-genomic era. *Bioessays* 2004;**26:**99–105.

[16] Goodacre R, Vaidyanathan S, Dunn WB, Harrigan GG, Kell DB. Metabolomics by numbers: acquiring and understanding global metabolite data. *Trends Biotechnol* 2004;**22:**245–52.

[17] Dunn WB, Ellis DI. Metabolomics: current analytical platforms and methodologies. *Trends Analyt Chem* 2005;**24:**285–94.

[18] Duarte NC, Becker SA, Jamshidi N *et al*. Global reconstruction of the human metabolic network based on genomic and bibliomic data. *Proc Nat Acad Sci USA* 2007;**104:**1777–82.

[19] Nicholson JK, Holmes E, Wilson ID. Gut microorganisms, mammalian metabolism and personalized health care. *Nat Rev Microbiol* 2005;**3:**431–38.

[20] Kenny LC, Dunn WB, Ellis DI *et al*. Novel biomarkers for preeclampsia detected using metabolomics and machine learning. *Metabolomics* 2005;**1:**227–34.

[21] Welthagen W, Shellie RA, Spranger J, Ristow M, Zimmermann R, Fiehn O. Comprehensive two-dimensional gas chromatography-time-of-flight mass spectrometry (GC x GC-TOF) for high resolution metabolomics:biomarker discovery on spleen tissue extracts of obses NZO compared to lean C57BL/6 mice. *Metabolomics* 2005;**1:**65–73.

[22] Wilson ID, Plumb R, Granger J, Major H, Williams R, Lenz EA. HPLC-MS-based methods for the study of metabonomics. *J Chromatogr B Analyt Technol Biomed Life Sci* 2005;**817:**67–76.

[23] Soga T, Ohashi Y, Ueno Y, Naraoka H, Tomita M, Nishioka T. Quantitative metabolome analysis using capillary electrophoresis mass spectrometry. *J Proteome Res* 2003;**2:**488–94.

[24] O'Hagan S, Dunn WB, Brown M, Knowles JD, Kell DB. Closed-loop, multiobjective optimization of analytical instrumentation: gas chromatography/time-of-flight mass spectrometry of the metabolomes of human serum and of yeast fermentations. *Anal Chem* 2005;**77:**290–303.

[25] O'Hagan S, Dunn WB, Knowles JD *et al*. Closed-loop, multiobjective optimization of two-dimensional gas chromatography/mass spectrometry for serum metabolomics. *Anal Chem* 2007;**79:**464–76.

[26] Horning EC, Horning MG. Human metabolic profiles obtained by Gc and Gc/Ms. *J Chromatogr Sci* 1971;**9:**129.

[27] Jellum E, Bjornson I, Nesbakken R, Johansson E, Wold S. Classification of human cancer cells by means of capillary gas chromatography and pattern recognition analysis. *J Chromatogr* 1981;**217:**231–37.

[28] Kuhara T. Diagnosis and monitoring of inborn errors of metabolism using urease-pretreatment of urine, isotope dilution, and gas chromatography-mass spectrometry. *J Chromatogr B Analyt Technol Biomed Life Sci* 2002; **781:**497–517.

[29] Underwood BR, Broadhurst D, Dunn WB *et al*. Huntington disease patients and transgenic mice have similar pro-catabolic serum metabolite profiles. *Brain* 2006;**129:**877–86.

[30] Kell DB, Brown M, Davey HM, Dunn WB, Spasic I, Oliver SG. Metabolic footprinting and systems biology: the medium is the message. *Nat Rev Microbiol* 2005;**3:**557–65.

[31] Peakman TC, Elliott P. The UK Biobank sample handling and storage validation studies. *Int J Epidemiol* 2008; **37(suppl 1):**i2–i6.

[32] Graham RC. *Data Analysis for the Chemical Sciences: A Guide to Statistical Techniques*. New York: VCH Publishers Inc, 1993.

[33] Lenz EM, Bright J, Wilson ID *et al*. Metabonomics, dietary influences and cultural differences: a H-1 NMR-based study of urine samples obtained from healthy British and Swedish subjects. *J Pharm Biomed Anal* 2004;**36:**841–49.

[34] Bollard ME, Holmes E, Lindon JC *et al*. Investigations into biochemical changes due to diurnal variation and estrus cycle in female rats using high-resolution H-1 NMR spectroscopy of urine and pattern recognition. *Anal Biochem* 2001;**295:**194–202.

[35] Lenz EM, Bright J, Knight R, Wilson ID, Major H. A metabonomic investigation of the biochemical effects of mercuric chloride in the rat using H-1 NMR and HPLC-TOF/MS: time dependant changes in the urinary profile of endogenous metabolites as a result of nephrotoxicity. *Analyst* 2004;**129:**535–41.

[36] Wacholder S, Chanock S, Garcia-Closas M, El Ghormli L, Rothman N. Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. *J Natl Cancer Inst* 2004; **96:**434–42.

[37] Ioannidis JP. Why most published research findings are false. *PLoS Med* 2005;**2:**e124.