Physiological variation in metabolic phenotyping and functional genomic studies: use of orthogonal signal correction and PLS-DA

C.L. Gavaghan[a,*], I.D. Wilson[b], J.K. Nicholson[a]

[a] *Biological Chemistry, Biomedical Sciences Division, Imperial College of Science, Technology and Medicine, University of London, Sir Alexander Fleming Building, South Kensington, London SW7 2AZ, UK*
[b] *Astra-Zeneca Pharmaceuticals, Department of Safety and Medicine, Alderley Park, Macclesfield, Cheshire SKN 4TG, UK*

Received 19 July 2002; revised 16 August 2002; accepted 4 September 2002

First published online 30 September 2002

Edited by Thomas L. James

Abstract Metabolic phenotyping, or metabotyping, is increasingly being used as a probe in functional genomics studies. However, such profiling is subject to intrinsic physiological variation found in all animal populations. Using a nuclear magnetic resonance-based metabonomic approach, we show that diurnal variations in metabolism can obscure the interpretation of strain-related metabolic differences in two phenotypically normal mouse strains (C57BL10J and Alpk:ApfCD). To overcome this problem, diurnal-related metabolic variation was removed from these spectral data by application of orthogonal signal correction (OSC), a data filtering method. Interpretation of the removed orthogonal variation indicated that diurnal-related variation had been removed and that the AM samples contained higher levels of creatine, hippurate, trimethylamine, succinate, citrate and 2-oxo-glutarate and lower levels of taurine, trimethylamine-*N*-oxide, spermine and 3-hydroxy-*iso*-valerate relative to the PM samples. We propose OSC will have great potential removing confounding variation obscuring subtle changes in metabolism in functional genomic studies and will be of benefit to optimising interpretation of proteomic and genomic datasets.

Key words: Metabonomics; ^{1}H NMR spectroscopy; Physiological variation; Orthogonal signal correction

1. Introduction

Functional genomic studies offer a holistic approach to studying biological systems by expanding the investigation from studying single genes or proteins to studying all genes or proteins at once in a systematic fashion [1]. In order to understand fully the biological function of particular genes it is necessary to correctly place genetic information in the context of biochemical activity at the cellular and whole organism level and metabolic phenotyping, or metabotyping, is increasingly being used as a probe in functional genomics studies [2].

Nuclear magnetic resonance (NMR)-based metabonomic approaches have been developed to aid the understanding of pathophysiological and toxicological processes in which time-related metabolic events in vivo are interrogated using chemometric methods [2]. Metabonomics thus represents a systems approach to understanding metabolic variation in complex multicellular organisms. Changes in metabolism are observable in NMR spectral profiles of biofluids or tissues and these data are reflective of the metabolic status of an organism [3]. Multivariate statistical methods are employed to enhance interpretation of these complex spectral data. In principle, metabonomic technologies can be extended to interpretation of genetic modifications in relation to physiological processes and we have recently demonstrated this potential by characterising metabolic differences in mouse models [4,5].

Interpretation of models of disease or toxicity in metabonomic studies may be confounded by variation contained within the NMR spectral data that is not related to the hypotheses under test. For example, physiological factors can markedly affect the metabolic composition of biological fluids and tissues and normal physiological variation can obscure the metabolic information that relates to the disease processes under study. Such factors include dietary variation [6], exercise and physical activity [7], normally regulated diurnal cycles [8], estrus or menstrual cycle-related processes [8], genetic drift, strain and species differences [6,9]. Similarly, unrelated biochemical consequences arising from gene insertion or deletion in the development of transgenic models may obscure the understanding of the physiological consequences relating to the disease or toxicity processes under investigation. In such cases, methods are required that can remove and minimise interference from unrelated metabolic variation. Ideally, these methods should enable interpretation of unrelated metabolic variation. For example, it may be desirable to characterise the metabolic effects resulting from unwitting interventions during the development of transgenic models.

In this study, we show the benefits of pre-filtering data by orthogonal signal correction (OSC) for the removal and investigation of non-correlated variation contained within spectral data in NMR-based metabonomic studies. OSC was originally applied to the removal of unwanted variation from near-infrared spectral data [10,11] although it has also found applications on data generated by X-ray powder diffraction [12], fluorescence spectroscopy [13] and ^{13}C cross-polarisation/magic angle spinning NMR spectroscopy [14]. We report how variation relating to diurnal cycles obscures the interpretation of strain-related metabolic differences in the phenotypically normal C57BL10J and Alpk:ApfCD mouse strains and demonstrate how OSC can eliminate this non-correlated variation by selectively removing and identifying the metabolic changes relating to diurnal and strain-related variation in ^{1}H NMR spectral data. OSC and related methods may also be applied

*Corresponding author. Fax: (44)-207-594 3226.
E-mail address: c.gavaghan@ic.ac.uk (C.L. Gavaghan).

0014-5793/02/$22.00
PII: S0014-5793(02)03476-2

to other types of multivariate data sets, such as those generated by transcriptomic and proteomic methods, and hence is likely to be a widely applicable functional genomic tool.

2. Materials and methods

2.1. Animal handling and collection of urine for ^1H NMR spectroscopy

Animal handling of male C57BL0J and Alpk:ApfCD mice, sample collection and storage followed previously reported procedures [4]. The possible loss of amines (due to their volatility) was minimised by immediate collection of urine into Eppendorf containers, which were then sealed and immediately frozen. Urine samples were obtained from 8 week old C57BL10J mice at AM (collected between 9:00 and 10:00, $n=20$) and PM (collected between 13:30 and 14:30, $n=20$) time-points and Alpk:ApfCD mice at AM (collected between 9:00 and 10:00, $n=15$) and PM (collected between 13:30 and 14:30, $n=15$) time-points by minimal manipulation of the lower abdomen.

2.2. Sample preparation and ^1H NMR spectroscopy of whole mouse urine

Samples were randomly selected for sample preparation and ^{1}H NMR acquisition. Due to the small volume mouse urine collected (typically 100 μl) only deuterated solvents were added during sample preparation, thus maximising NMR sensitivity. The collected urine volumes were highly variable (40–340 μl range); deuterium oxide (D_2O) was added to each sample to make the urine sample up to 340 μl total volume. In order to control the sample pH to a narrow range, sodium phosphate buffer (140 μl, 100 mM in D_2O, pD 7.4) was added to the urine samples. The differential dilution of the urine samples with D_2O to make the samples to comparable volumes does not affect the PR analysis because of the normalisation procedure used during data processing and ensured that the phosphate buffer concentration added remained consistent throughout the sample set. Trimethylsilylpropionic acid (50 μl, 0.05 mg/ml in D_2O) was added as an internal reference standard. Samples were prepared and analysed in a random order. ^{1}H NMR spectra were acquired at 300 K on a Bruker DRX-600 spectrometer as previously reported [4].

2.3. Data reduction of the NMR spectra

Data reduction and normalisation procedures were followed as previously reported [4].

2.4. Principal components analysis (PCA) of the ^1H NMR spectral data

The normalised, centred data were analysed by PCA to establish any 'groupings' with respect to strain or diurnal cycles. A PCA model was constructed using all samples. The Scores plot of PC1 versus PC2 was examined for separation or clusters relating to the two strains of mice, C57BL10J and Alpk:ApfCD, and variation related to diurnal differences between the AM and PM urine collections.

2.5. Prediction of classes by partial least squares-discriminant analysis (PLS-DA)

As PCA indicated the presence of strain and diurnal-related separation, data were analysed by PLS-DA with a view to establishing whether the separation between the C57BL10J and Alpk:ApfCD AM and PM clusters was significant by prediction of class.

PLS-DA models were constructed to establish the significance of the effects of strain and diurnal cycles on the urinary metabolite profiles. Thus, to investigate the significance of strain-related variation two PLS-DA model types were calculated, the first (PLS-DA model type 1) related to strain differences in the AM urine collections (C57BL10J, AM = class 1 and Alpk:ApfCD, AM = class 2) and the second (PLS-DA model type 2) included data relating to PM collections (C57BL10J, PM = class 1 and Alpk:ApfCD, PM = class 2). For the purpose of cross-validation and due to the small number of samples, 10 validation PLS-DA models were calculated (five for model type 1 and five for model type 2) excluding 10% of the samples per class in each validation model. To establish whether the differences were significant between strains the PLS-DA validation models were used to predict the strain of the test samples. Class membership was predicted using a value of 0.5 dividing line between the two classes including values predicted using a class membership probability value >0.01.

The same procedure was applied to establish the significance of the effects of diurnal cycles on the urinary metabolite profiles, i.e. two PLS-DA model types were calculated to investigate the differences due to diurnal cycles. The first (PLS-DA model type 3) investigated diurnal variation in the C57BL10J strain (C57BL10J, AM = class 1 and C57BL10J, PM = class 2) and the second (PLS-DA model type 4) investigated diurnal variation in the Alpk:ApfCD strain (Alpk:ApfCD, AM = class 1 and Alpk:ApfCD, PM = class 2). These PLS-DA models were validated as detailed for the investigation between strains.

2.6. Application of OSC to remove diurnal- and strain-related variations from ^1H NMR spectral data

To remove and characterise the variation in metabolite profiles caused by diurnal cycles, OSC, a spectral filtering method, was applied to the NMR data. In the OSC procedure, the identity of the sample classes is included in the calculation and assigned by a vector, Y. The first step in the OSC procedure calculates the first principal component, or score vector t, that is an optimal linear description of the spectral data and describes maximum separation based on class. The longest vector orthogonal to Y is calculated and this vector, t^*, describes the greatest source of variation that is not related to class yet still provides a good summary of the spectral data [10]. This is done as:

$$t^* = (1 - Y(Y'Y)^{-1}Y')t$$

After this step, the loading vector, p^*, relating to this corrected vector is calculated and the product of the orthogonal score and loading vectors is subtracted from the spectral data. Examination of the OSC scores, t^* and loadings p^* can be useful in determining the source of the removed variation. The residual matrix represents the filtered spectral data and is then used for calculation of PCA or PLS-based models.

To remove diurnal variation from the urinary ^{1}H NMR data by OSC, and focus subsequent analysis on the strain-related metabolic differences, observations relating the Alpk:ApfCD strain were specified as class 1 and the C57BL10J strain as class 2.

To illustrate how OSC can be utilised to selectively remove unwanted non-correlated variation from spectral data, the variation relating to strain-related metabolic differences was removed by application of OSC. The observations relating to the AM urine collections were specified as class 1 and the PM urine collections as class 2. The OSC scores, t^*, and loadings, p^*, from both filtering investigations were examined to investigate the variation removed. The strategy used for the OSC application in this study is shown in Fig. 1.

2.7. PCA of the OSC-filtered ^1H NMR data

PCA was applied to both OSC-filtered NMR data sets. The scores plots of PC1 versus PC2 were examined for separation between the two strains of mice or between the diurnal groups. The corresponding loadings were investigated to establish which NMR spectral regions related to diurnal or strain-related metabolic differences. The OSC scores, t^*, and loadings were also examined to characterise the variation removed from the filtered data.

2.8. Prediction of classes by PLS-DA after application of OSC

To ascertain the validity of the separation between classes in the PCA models calculated after application of OSC, prediction of classes was repeated to (a) investigate the significance of the differences between strains (class 1 = Alpk:ApfCD, AM and PM, class 2 = C57BL10J, AM and PM) and (b) investigate the significance of strain-related differences (class 1 = AM observations and class 2 = PM observations).

Due to the small number of samples and as data relating to both AM and PM observations were included in the PLS-DA models after OSC ($n=70$), five PLS-DA validation models were calculated excluding representative samples from the Alpk:ApfCD, AM; Alpk:ApfCD, PM; C57BL10J, AM and C57BL10J, PM classes. Samples to be excluded from a PLS-DA model, calculated for the purpose of prediction of class identity, were also excluded during the OSC procedure performed prior to PLS-DA. Excluded sample numbers accounted for 10% of the total sample number for class prediction.

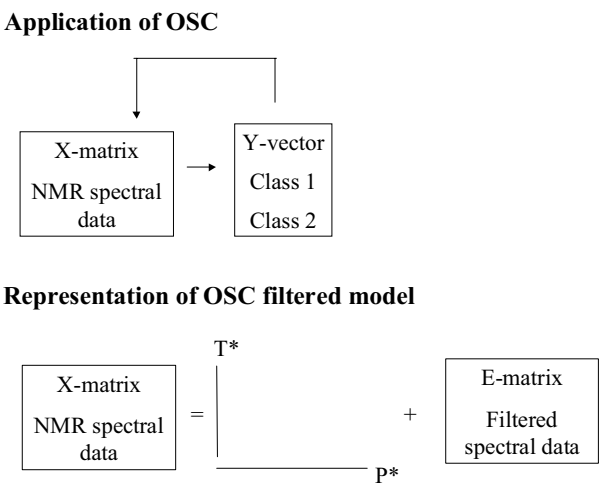

Fig. 1. Strategy for demonstrating the application of OSC to ^{1}H NMR spectral data for the selective removal of confounding variation. 1) OSC was applied to remove confounding orthogonal variation from the ^{1}H NMR urinary spectral data relating to the Alpk:ApfCD and C57BL10J mice. To remove diurnal variation, class 1 = Alpk:ApfCD and class 2 = C57BL10J. To remove strain-related variation, class 1 = AM and class 2 = PM urine collections. 2) The OSC-filtered NMR data matrix was subsequently used for PCA investigations of (a) strain-related differences between the Alpk:ApfCD and C57BL10J mice and (b) the effects of diurnal cycles on the urinary metabolite profiles in both mouse strains. Examination of the OSC scores, t*, and loadings, p*, enabled interpretation of the variation removed by the OSC procedure. These corresponded to t* and p* relating to (a) removed diurnal variation and (b) removed strain-related variation.

3. Results

3.1. ^{1}H NMR spectroscopy of urine samples from C57BL10J and Alpk:ApfCD mice at AM and PM time-points

Visual inspection of the ^{1}H NMR urine spectra revealed changes in the patterns associated with a variety of differences between the two strains and diurnal variation (Fig. 2). Inter-animal variation in urine composition and the complexity of the spectra, coupled with differences in the volume of urine excreted, makes visual comparison of spectra difficult, hence chemometric methods of spectral analysis were employed.

3.2. PCA of ^{1}H NMR urinary data

The scores plot of PC1 versus PC2 (Fig. 3) showed separation between the C57BL10J and Alpk:ApfCD strains. Distinct separation of the Alpk:ApfCD, AM cluster from both the AM and PM C57BL10J cluster occurred in PC1. Although PC2 described variation relating to the spectral differences between the Alpk:ApfCD, PM cluster with both AM and PM C57BL10J classes this separation was less distinct

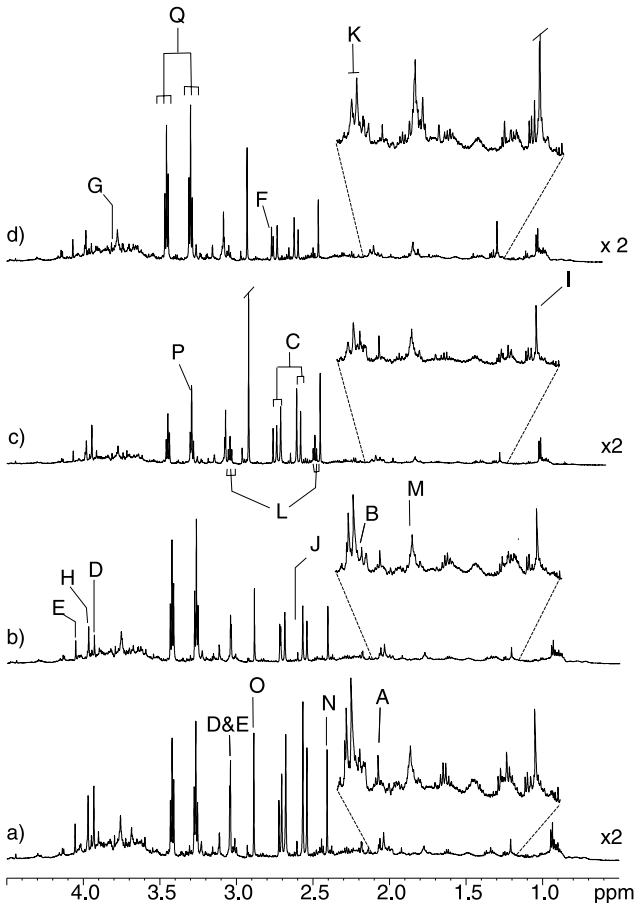

Fig. 2. 600 MHz ^{1}H NMR spectra of representative Alpk:ApfCD and C57BL10J mouse urine collected at AM and PM time-points. Key: (a) C57BL10J AM; (b) C57BL10J PM; (c) Alpk:ApfCD AM; (d) Alpk:ApfCD PM; A = acetate; B = acetamide; C = citrate; D = creatine; E = creatinine; F = dimethylamine; G = guanodinoacetic acid; H = hippurate; I = 3-hydroxy-*iso*-valerate; J = methylamine; K = *N*-acetyl glycoproteine; L = 2-oxo-glutarate; M = spermine; N = succinate; O = trimethylamine; P = trimethylamine-*N*-oxide; Q = taurine.

and the C57BL10J AM cluster was only partially separated from the Alpk:ApfCD PM cluster. Separation of the C57BL10J AM and PM groups appeared to be partially separated in PC2 and formed a tighter cluster than the two Alpk:ApfCD clusters.

3.3. Investigation of the diurnal- and strain-related variations by PLS-DA

PLS-DA was applied to these data to investigate whether the C57BL10J and Alpk:ApfCD strains showed significantly different urinary metabolite profiles. PLS-DA validation mod-

Table 1
Summary of prediction results investigating the diurnal effects in two strains of mice by PLS-DA

PLS-DA model	Percentage of correct predicted classifications	Percentage of predictions correct with significance < 0.01
C57BL10J, AM versus Alpk:ApfCD, AM	100	90
C57BL10J, PM versus Alpk:ApfCD, PM	100	90
C57BL10J, AM versus C57BL10J, PM	100	100
Alpk:ApfCD, AM versus Alpk:ApfCD, PM	100	80

The prediction results are an average of ten test samples calculated from predictions of five validation models per PLS-DA model stated. Two test samples were used per PLS-DA model for prediction of class membership.

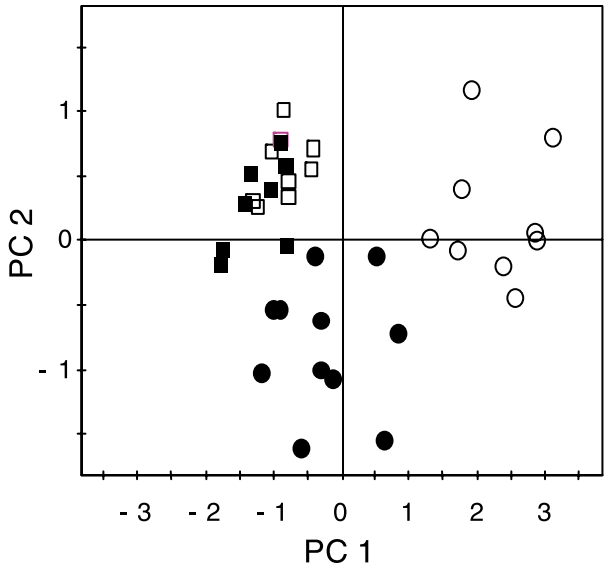

Fig. 3. PC scores plot derived from the ^{1}H NMR spectra of urine samples obtained from Alpk:ApfCD and C57BL10J mice collected at AM and PM time-points. Key: ○ = Alpk:ApfCD, AM; ● = Alpk:ApfCD, PM; □C57BL10J, AM; ■ = C57BL10J, PM.

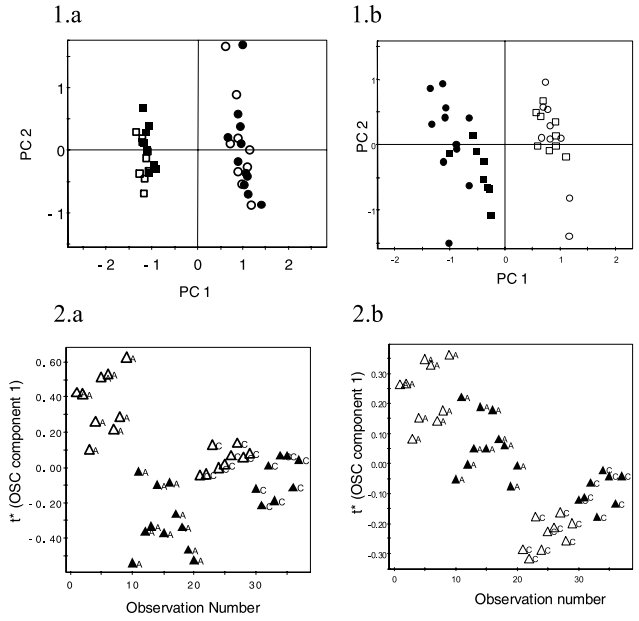

Fig. 4. Scores plots generated from application of OSC to remove (a) diurnal-related variation and (b) strain-related variation. (1) PC scores plot from the ^{1}H NMR spectra of urine samples obtained from Alpk:ApfCD and C57BL10J mice after application of OSC. Key: ○ = Alpk:ApfCD, AM; ● = Alpk:ApfCD, PM; □ = C57BL10J, AM; ■ = C57BL10J, PM. (2) Plot of OSC scores relating to the variation removed by OSC. Key: △A = Alpk:ApfCD, AM; ▲A = Alpk:ApfCD, PM; △C = C57BL10J, AM; ▲C = C57BL10J, PM.

els were calculated to (a) investigate the significance of the differences between the strains at each time-point and (b) investigate the significance of the diurnal effects in each strain. The prediction results (Table 1) show that for one-component models, the differences between the Alpk:ApfCD and C57BL10J strains, at both AM and PM time-points, can be predicted accurately in 100% of cases and predicted 90% accurately with significance ≤ 0.01. The differences between the diurnal cycles in each strain of mouse are predicted accurately in 100% of cases and predicted $\geq 80\%$ accurately with significance ≤ 0.01.

3.4. *OSC-PCA application to ^{1}H NMR urinary data to remove diurnal- and strain-related variation*

The ^{1}H NMR urinary data were filtered using OSC to remove (a) diurnal- and (b) strain-related variation. Following OSC, PCA models were calculated to examine variation relating to (a) strain differences and (b) diurnal cycles (Fig. 1). Resultant scores plots of PC1 versus PC2 (Fig. 4(1)) showed distinct separation in PC1 between (a) C57BL10J and Alpk:ApfCD strains and (b) AM and PM observations, respectively. All PCA components were examined and confirmed that (a) diurnal- and (b) strain-related separations were effectively removed during the OSC procedure for both investigations.

The OSC scores, t*, and loadings, p*, were examined in both analyses to investigate the variation removed during the OSC procedure. The OSC scores, t*, obtained for the OSC-PCA investigation of strain-related differences (Fig. 4(2a)) showed distinct separation between the Alpk:ApfCD, AM and Alpk:ApfCD, PM observations and apparent partial separation between the C57BL10J, AM and C57BL10J, PM observations, reflective of the diurnal-related separation observed in the scores plots from PCA prior to OSC (Fig. 3). Examination of the OSC loadings (Fig. 5) revealed the differences in the spectral regions describing diurnal-related variation.

The OSC scores, t*, obtained for the OSC-PCA investigation of diurnal-related differences (Fig. 4(2b)) showed partial separation between the Alpk:ApfCD and C57BL10J strains, indicating that strain-related variation had been effectively filtered from the spectral data. Examination of the OSC loadings and interpretation of the respective metabolite changes associated with the strain-related separation observed in the OSC scores confirmed that strain-related variation had been removed.

As pre-filtering the data by OSC is a PLS-based supervised method, the interpretation of the PCA models investigating strain- and diurnal-related variations were validated by prediction of class using PLS-DA. The prediction results (Table 2) for one-component models predicted $\geq 95\%$ correctly for the strain or diurnal cycle with significance ≤ 0.01.

Table 2
Summary of PLS-DA prediction results investigating biochemical changes associated with diurnal cycles and strain differences

PLS-DA class prediction	Percentage of correct predicted classifications	Percentage of predictions correct with significance < 0.01
Prediction of strain	100	95
Prediction of diurnal cycle	100	100

For each class the percentage reported is an average obtained from five training PLS-DA models. These were calculated after an OSC procedure removing either variation relating to diurnal cycles or strain differences.

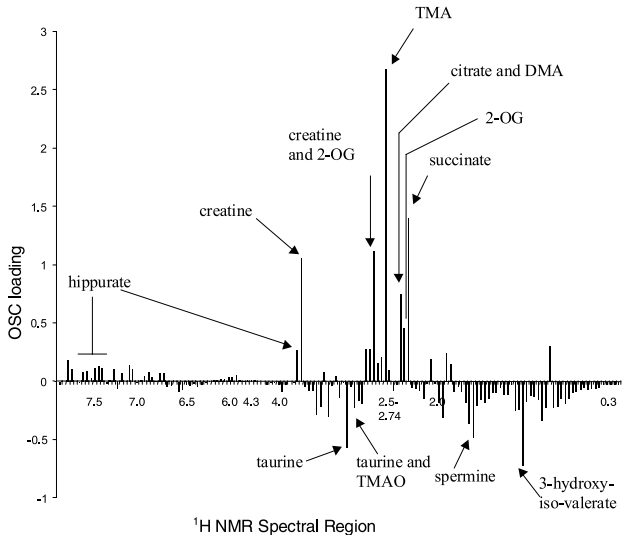

Fig. 5. Plot of OSC loadings showing NMR spectral descriptors relating to diurnal variation removed during OSC procedure. Each bar represents a spectral region covering 0.04 ppm. This plot shows how the ^{1}H NMR metabolite profiles of the AM and PM urine samples differ, where a positive value indicates that relatively greater urinary concentration of a metabolite is present in the AM sample.

4. Discussion

4.1. Diurnal variation in the biochemical composition of urine

We have shown that there is marked diurnal variation in metabolite excretion profiles both in Alpk:ApfCD and C57BL10J mice. PCA revealed that such variation appeared greater in the Alpk:ApfCD strain compared to the C57BL10J strain as indicated by the tighter clustering of the individuals and partial overlap of the AM and PM observations in the C57BL10J strain. Prediction of class by PLS-DA established that the separation observed in the scores plots from PCA (Fig. 2) was due to significant diurnal-related variation in the urinary ^{1}H NMR metabolite profiles in both mouse strains. Further interpretation of the PCA model to establish the metabolic differences between the two mouse strains was confounded by the influence of diurnal variation, illustrating how normal physiological variation factors may hinder metabonomic analyses. It is axiomatic that gene expression and protein synthesis must also be affected by such normal physiological variation and hence we can infer that interpretation of genomic and proteomic data may benefit from OSC treatment. By removing confounding variation from these data sets interpretation would be focused to the disease model or toxicity effect under investigation.

OSC was applied to the NMR spectral data to remove the influence of diurnal metabolic perturbations and examination of the OSC scores (Fig. 4(2a)) showed separation between the Alpk:ApfCD diurnal groups and apparent partial separation of the C57BL10J diurnal groups, confirming that the variation removed during OSC related to diurnal cycles. Furthermore, the separation between the Alpk:ApfCD and C57BL10J diurnal groups reflected that observed in the PCA scores plot (Fig. 3) prior to OSC. Thus, the Alpk:ApfCD diurnal groups were distinctly separated in the OSC scores plot whereas the C57BL10J diurnal groups showed partial overlap. Examination of the OSC loadings revealed the spectral regions responsible for the differences in metabolite profiles caused by diurnal variation in the Alpk:ApfCD and C57BL10J mouse strains and the relative metabolite changes contained within these spectral regions for AM and PM urinary profiles were identified (Fig. 5). The AM samples contained relatively higher levels of tricarboxylic acid cycle intermediates (citrate and 2-oxo-glutarate), indicating that the metabolite profile differences reflected the increased metabolic activity of mice during the night.

Following the OSC procedure removing diurnal-related variation, PCA resulted in distinct strain-related separation of the C57BL10J and Alpk:ApfCD groups in PC1, and the scores plot of PC1 versus PC2 (Fig. 4(1a)) indicated that diurnal-related variation had been removed. The PC1 loadings and corresponding urinary metabolic differences between the strains were examined and the results were compared to previous work investigating strain differences between male C57B10J and Alpk:ApfCD strains [4]. Generally, the strain-related metabolic differences concurred with previously reported results. The Alpk:ApfCD mice contained relatively higher levels of creatine and lower levels of guanidinoacetic acid, supporting the previously reported suggestion that guanidinoacetate methyltransferase-dependent pathways are more active in the Alpk:ApfCD mice. However, although the levels of trimethylamine-*N*-oxide (TMAO) were still relatively higher in the Alpk:ApfCD mice compared to the C57BL10J mice, in the present study the range of TMAO concentrations present in the Alpk:ApfCD varied greatly and in some cases, the urinary TMAO levels were similar to those in the C57BL10J group. In the previous study, TMAO concentrations were significantly higher in the Alpk:ApfCD mice [4]. The ratio of trimethylamine (TMA) concentrations was still an important discriminating factor between the strains in the present study but this ratio was inverse to earlier findings, thus the Alpk:ApfCD mice contained higher levels of TMA than the C57BL10J mice. The TMA:TMAO ratios in the C57BL10J mice were comparable in both previous and present studies [4]. TMA concentrations were consistently higher relative to TMAO concentrations in both studies.

Urinary TMA levels are known to originate from the action of intestinal microflora on precursors such as choline [15,16]. Studies have shown that animal handling can affect colon temperature and this is referred to as stress-induced hypothermia (SIH) [17]. There is the possiblity that the animals from this study may have experienced SIH. However, from preliminary studies (data not shown), where urine was collected from mice in the afternoon, and therefore not exposed to any stress from handling prior to urine collection, we observed differences in the urinary composition compared to AM urine collections obtained from different animals. As the animals used in this study were from colonies which are handled regularly and have been maintained since 1967, we previously proposed that it would be unlikely that the observed changes in TMA:TMAO ratios were due to a systematic difference in gut microfloral activity [4]. However, a recently reported study by Gareipy et al. investigated differences in adult stress regulation between highly aggressive and non-aggressive mice and they showed that the systemic effects of handling are modulated by genetic background [18]. These findings indicate that the C57BL10J and Alpk:ApfCD strains may experience different levels of stress upon handling and hence exhibit different SIH responses. Thus, we cannot exclude the possibility that handling has a minor effect on metabolism, in particular

metabolism of the methylamines by the gut microflora, and on the urinary metabolite profiles obtained. The scores obtained from PC models constructed prior to and post application of OSC and the OSC scores t^* generated were examined for any separation indicative of a SIH response. No trend or separation according to the order of urine collections was observed in the AM and PM urine collections in both strains. Thus, we conclude that stress effects resulting from the animal handling procedure the animals in the current study were exposed to were minimal and the separation between the AM and PM urine collections is attributed to genuine diurnal metabolic regulation. However, the work by Gareipy et al. does indicate that the systemic effects of animal handling should be considered in all studies where metabonomic, proteomic and genomic data are used.

The present study indicates that the urinary products from the methylamine pathway are strongly influenced by diurnal cycles and that the influence of diurnal cycles is more pronounced in the Alpk:ApfCD strain. The urine collections in these studies do not contain time-averaging effects that are observed when urine is collected over defined time periods from animals placed in metabolic cages, as typified in metabonomic studies in rat models. In this study urine was collected at a fixed time interval for each animal by manipulation of the lower abdomen, thus enhancing the sensitivity of the metabonomic analysis to the influence of diurnal cycles. This increased sensitivity is also an important consideration for the application of metabonomic studies in humans, as biofluid measurements (e.g. plasma, sera) are at fixed time intervals and are representative of the metabolite 'fingerprint' at the time of sampling.

We have shown in this study the confounding influence of diurnal cycles in the urinary metabolite profiles in the Alpk:ApfCD and C57BL10J mouse strains and how the Alpk:ApfCD urinary metabolite profiles were more influenced by diurnal cycles. Application of OSC to NMR spectral data demonstrated how this approach can selectively remove confounding variation in metabonomic studies, focusing interpretation of metabolite excretion profiles to strain-related differences between the phenotypically normal Alpk:ApfCD and C57BL10J mouse strains. The potential for metabotyping mice using NMR of biofluids in the evaluation of animal models of disease and drug efficacy has been proposed in earlier work [4]. We have shown here that OSC improves the effectiveness of the use of metabonomics in functional genomic studies and for minimising confounding variation with significant exogenous components and influences.

Acknowledgements: This work was supported by Astra-Zeneca.

References

[1] Heiter, P. and Boguski, M. (1997) Science 278, 601–602.
[2] Nicholson, J.K., Connelly, J., Lindon and Holmes, E. (2002) Nat. Rev. 1, 153–161 and references cited therein.
[3] Nicholson, J.K., Buckingham, M.J. and Sadler, P.J. (1983) J. Biochem. 211, 606–615.
[4] Gavaghan, C.L., Holmes, E., Lenz, E., Wilson, I.D. and Nicholson, J.K. (2000) FEBS Lett. 484, 169–174.
[5] Griffin, J.L., Williams, H.J., Sang, E., Clarke, K., Rae, C. and Nicholson, J.K. (2001) Anal. Biochem. 293, 16–21.
[6] Gavaghan, C.L., Nicholson, J.K., Connor, S.C., Wilson, I.D., Wright, B. and Holmes, E. (2001) Anal. Biochem. 291, 245–252.
[7] Holmes, E., Foxall, P.J.D., Nicholson, J.K., Nield, G.H., Brown, S.M., Beddell, C.R., Sweatman, B.C., Rahr, E., Lindon, J.C., Spraul, M. and Neidig, P. (1994) Anal. Biochem. 220, 284–296.
[8] Bollard, M.E., Holmes, E., Lindon, J.C., Mitchell, S.C., Branstetter, D., Zhang, W. and Nicholson, J.K. (2001) Anal. Biochem. 295, 194–202.
[9] Holmes, E., Nicholls, A.W., Lindon, J.C., Connor, S.C., Connelly, J.C., Haselden, J.N., Damment, S.J.P., Spraul, M., Neidig, P. and Nicholson, J.K. (2000) Chem. Res. Toxicol. 13 (6), 771–778.
[10] Wold, S., Antti, H. and Lindgren, F. (1998) Chemometr. Intell. Lab. 44, 175–185.
[11] Fearn, T. (1999) Chemometr. Intell. Lab. 50, 47–52.
[12] Artusson, T., Hagman, A., Bjork, S., Trygg, J., Wold, S. and Jacobsson, S.P. (2000) Appl. Spectrosc. 54 (8), 1222–1230.
[13] Eriksson, L., Trygg, J., Johansson, E., Bro, R. and Wold, S. (2000) Anal. Chim. Acta 420, 181–195.
[14] Hauksson, J.B., Edlund, U. and Trygg, J. (2001) Magn. Reson. Chem. 39, 267–275.
[15] Asatoor, A.M. and Simenhoff, M.I. (1965) Biochim. Biophys. Acta 111, 384–392.
[16] Baker, J.R. and Chaykin, S. (1962) J. Biol. Chem. 237, 1309–1313.
[17] Zethof, T.J., Van der Heyden, J.A., Tolboom, J.T. and Oliver, B. (1994) Physiol. Behav. 55, 109–115.
[18] Gariepy, J.L., Rodriguiz, R.M. and Jones, B.C. (2002) Pharmacol. Biochem. Behav. 73, 7–17.