Correlation network analysis for data integration and biomarker selection†

Aram Adourian*[a]**, Ezra Jennings**[a]**, Raji Balasubramanian**[a]**, Wade M. Hines**[a]**, Doris Damian**[a]**, Thomas N. Plasterer**[a]**, Clary B. Clish**[a]**, Paul Stroobant**[a]**, Robert McBurney**[a]**, Elwin R. Verheij**[ab]**, Ivana Bobeldijk**[ab]**, Jan van der Greef**[ab]**, Johan Lindberg**[c]**, Kerstin Kenne**[c]**, Ulf Andersson**[c]**, Heike Hellmold**[c]**, Kerstin Nilsson**[d]**, Hugh Salter**[d]** and Ina Schuppe-Koistinen**[c]

[a]*BG Medicine Inc., 610N Lincoln Street, Waltham, Massachusetts, USA. E-mail: aadourian@bg-medicine.com; Fax: +1-781-895-1119; Tel: +1-781-890-1199*

[b]*Analytical Sciences Department, TNO Quality of Life, P.O. Box 360, Zeist, The Netherlands*

[c]*Molecular Toxicology, Safety Assessment, AstraZeneca R&D, 15185 Södertälje, Sweden*

[d]*Disease Biology, AstraZeneca R&D, 15185 Södertälje, Sweden*

High-throughput biomolecular profiling techniques such as transcriptomics, proteomics and metabolomics are increasingly being used in *in vivo* studies to recognize and characterize effects of xenobiotics on organs and systems. Of particular interest are biomarkers of treatment-related effects which are detectible in easily accessible biological fluids such as blood. A fundamental challenge in such biomarker studies is selecting among the plethora of biomolecular changes induced by a compound and revealed by molecular profiling, to identify biomarkers which are exclusively or predominantly due to specific processes. In this work we present a cross-compartment correlation network approach, involving no *a priori* supervision or design, to integrate proteomic, metabolomic and transcriptomic data for selecting circulating biomarkers. The case study we present is the identification of biomarkers of drug-induced hepatic toxicity effects in a rodent model. Biomolecular profiling of both blood plasma and liver tissue from Wistar Hannover rats administered a toxic compound yielded many hundreds of statistically significant molecular changes. We exploited drug-induced correlations between blood plasma analytes and liver tissue molecules across study animals in order to nominate selected plasma molecules as biomarkers of drug-induced hepatic alterations of lipid metabolism and urea cycle processes.

Introduction

The application of high-throughput molecular profiling techniques such as transcriptomics, proteomics and metabolomics for *in vivo* toxicological studies is currently an active area of exploration.[1-5] Of particular promise is the potential of molecular profiling for identifying novel biological markers (biomarkers) of toxicity-related tissue pathogenesis, particularly in easily accessible biological fluids such as blood or urine. Indeed, accessible biomarkers for more accurate and earlier assessment of the potential of novel pharmaceutical agents to act as toxicants in man are critical to continued improvement of drug development and safety evaluation.

In particular, integrating multiple, disparate high-throughput data sets for biomarker discovery is an area of active research in the pharmaceutical and life sciences communities, and is the subject of

the present study. In particular, we consider here the case of biomarkers for drug-induced liver injury as an exemplary and important field of such research. Biomarkers of adverse hepatic effects that are detectible in accessible biological fluids and which are invariant or highly conserved across species are particularly of interest, due to their availability for repeated measurements and potential for comparison across preclinical models and human subjects. Indeed, candidate biomarkers of liver toxicity derived from molecular profiling of biological fluids such as blood or urine will need to share many characteristics of established hepatic injury markers, such as the transaminase enzymes in blood.[6] However, in contrast to transaminase enzymes, which become manifest primarily after frank tissue damage has occurred, novel molecular biomarkers which can be shown to be specific to early alterations of normal hepatic function promise predictive utility as well as indication of specific biochemical processes responsible for, or leading to, liver injury.

Recently a number of groups have reported on *in vivo* studies of molecular profiling of blood or urine components for the purpose of identifying drug-induced hepatotoxicity biomarkers.[7-9] Such experiments routinely analyze and quantify hundreds of biomolecular species from each sample. An inherent challenge of such studies is discerning which of the plethora of observed biomolecular changes in the biological fluid matrix are exclusively or predominantly of liver origin and reflective of hepatotoxic processes, and which are secondary or unrelated effects due to pharmacology, acute phase response, systemic stress and the like. One key objective of the present study is to address this challenge. In particular, we describe a novel approach to selecting and qualifying molecules for biomarkers among all observed biomolecular alterations in an experiment, and establishing plausibility of such biomarkers in the context of hepatotoxicity.

To these ends, we introduce and demonstrate a new, empirical, systems-based approach to relating observed biomolecular changes in biological fluids to drug-induced toxicity effects in liver tissue for *in vivo* models, using a cross-tissue correlation network-based integration of molecular profiling data. Previous molecular profiling approaches have typically investigated the differential behavior of individual profiled analytes, characterized by univariate or multivariate metrics of 'group mean fold-change' or equivalent, and associated confidence values. These approaches neglect inherent biological variability by averaging analyte concentration levels across organisms within a pre-defined experimental group, with the underlying assumption that similar or identical genotypes should produce similar steady-state responses to treatment under controlled conditions. In the current study we exploit the subtle but inevitable biological variability across study subjects to integrate different types of molecular profiling data and elucidate additional information about molecular responses to a toxicant. Although such data integration requires extremely stringent tolerances on bioanalytical platforms, namely that the measurement variance be consistently less than the individual-to-individual biological variability, the additional information is essential in detecting subtle and early modulations of biochemical pathways and mechanisms.[10-15] Correlation analyses are also appealing for broad molecular profiling analyses in that they are not dependent on absolute quantification of analytes.

In this study we identify candidate biomarkers in blood plasma which we show to be empirically correlated to alterations of normal hepatic function and to drug metabolism within liver tissue. We further provide a correlation network-based integration of molecularly profiled components spanning liver tissue and blood plasma to explore the origins of the candidate biomarkers. Comprehensive molecular profiling analyses were carried out on liver tissue and blood plasma of Wistar Hannover rats administered different dosage regimes of an experimental drug compound which previously had been shown to induce liver steatosis. Molecular profiling of liver tissue samples involved cytosolic and membrane-bound protein profiling using liquid chromatography and mass spectrometry, endogenous metabolite analysis using multiple liquid chromatography-mass spectrometry methods, and gene expression profiling using high density oligonucleotide microarrays. Blood plasma from the same animals was also profiled using mass spectrometric techniques to measure endogenous metabolite levels, as well as conventional clinical chemistries and hematology measurements. Correlation networks derived from these analyses generated a number of candidate plasma biomarkers of drug-induced processes, in particular of perturbations in lipid metabolism and the urea cycle, and have provided new insights into adverse processes related to those candidate biomarkers which merit further investigation.

Results

To characterize the effects of the experimental compound on endogenous liver and plasma analytes after one and seven days of treatment, rats were administered the drug at dosages of 200 mg kg^{-1} per day for 7 days, 20 mg kg^{-1} per day for 7 days or 200 mg kg^{-1} for a single day, with five male rats per dosage group. Histopathology examination revealed centrilobular hepatosteatosis in rats treated with 200 mg kg^{-1} per day of the compound for seven days. Slight centrilobular hypertrophy of liver cells was also observed in these animals. Of the five rats administered 200 mg kg^{-1} for a single day, two exhibited positive multifocal liver fat stains, as did two of the five rats administered 20 mg kg^{-1} per day of the compound for seven days. No other organ toxicities were observed in the treatment groups. In addition to the three treatment groups, five male rats were administered the vehicle solution for seven days as a control group. No liver pathology was observed in animals in this vehicle-administered control group.

The only statistically significant blood chemistry alterations relative to the vehicle-administered group were observed in the group administered 200 mg kg^{-1} per day for 7 days, consisting of a +19% ± 4% increase in aspartate aminotransferase (AST) levels ($p_{FDR} < 0.05$), and a −68% ± 23% decrease in triglyceride levels ($p_{FDR} < 0.05$). Blood glucose, urea, creatine, alanine aminotransferase, bilirubin, cholesterol, albumin, sodium and potassium were not statistically significantly different between the vehicle group and any other treatment group. No treatment-related hematology alterations were noted apart from a slight increase in neutrophil and monocyte counts in the group receiving 200 mg kg^{-1} per day doses for seven days.

Molecular profiling of plasma samples using mass spectrometry platforms, and of liver samples using mass spectrometry and oligonucleotide microarray platforms, yielded a plethora of biomolecular changes as a result of administration of the compound. Specifically, in the comparison of vehicle-administered rats and rats administered the compound at 200 mg kg^{-1} per day for 7 days, at a univariate significance level of $p_{FDR} < 0.05$ over 200 altered endogenous plasma metabolites were observed, over 300 liver gene transcript changes were detected, and over 400 liver peptides exhibiting abundance change were identified. Similarly, in the 200 mg kg^{-1} for a single day and 20 mg kg^{-1} per day for 7 days dosage groups, over 500 and over 80 analyte abundance changes, respectively, relative to the vehicle-adminstered group at a univariate significance level of $p_{FDR} < 0.05$, were observed across all bioanalytical platforms.

In order to prioritize among the abundance of observed biomolecular changes induced by drug exposure, and to select plasma-based biomarkers related to changes within liver tissue, the partial correlation matrix comprising all of the measured analytes in plasma and tissue was calculated, using all three dosage groups (a total of fifteen animals) and controlling for group mean effects.[16] This yielded 172 correlations spanning analytes in plasma and analytes in liver, and 17 327 correlations among solely liver analytes, at a significance for correlations of $p_{FDR} < 0.15$. The overall correlation network derived from the correlation matrix was not entirely connected; there were many disjoint sub-networks making up the larger network. The focus in this work was on the sub-networks which included the 172 correlations spanning plasma and tissue. The number of sub-networks for consideration was further reduced by requiring that at least one plasma analyte member of each of these sub-networks exhibit a statistically significant ($p_{FDR} < 0.05$) univariate change between the vehicle-administered group and the group receiving 200 mg kg^{-1} per day for seven days. This resulted in nine disjoint networks, of which two are discussed in the present report. The two networks discussed here were selected because they each comprised a larger number of analytes than any of the other seven networks.

Correlation network centered upon hepatic UDP-glucuronosyltransferase and plasma lipids

One plasma-to-liver correlation network which was found to comprise plasma analytes as well as liver analytes was centered upon hepatic UDP-glucuronosyltransferase polypeptide A1 (Ugt1a1) and a number of plasma lipids. This correlation network was pursued because of the established role of Ugt1a1 in hepatic drug metabolism processes,[17] and as such this network most likely reflects liver exposure to the compound. Fig. 1A is a visualization of the correlation matrix centered upon hepatic Ugt1a1, illustrating significant correlations both to other liver analytes as well as to analytes observed in plasma. In Fig. 1A, each observed protein, gene transcript and endogenous metabolite is assigned a node co-ordinate in the two-dimensional plane, and the links between nodes represent

correlation values between pairs of nodes. The network in Fig. 1A has been constrained to comprise only analytes which are separated by one correlation link from Ugt1a1; apart from this constraint this correlation analysis is unsupervised. In order to illustrate a representative data set which gives rise to one of the correlation links in Fig. 1A, Fig. 1B shows a plot of the relative abundances of a plasma triglyceride with fatty acid composition of 52 carbon atoms and four double bonds (52:4 TG) and a liver peptide of Ugt1a1, both of which are represented in the network of Fig. 1A; these two analytes were inversely correlated with a partial correlation value of $r_{partial} = -0.86$ ($p_{FDR} < 0.05$).
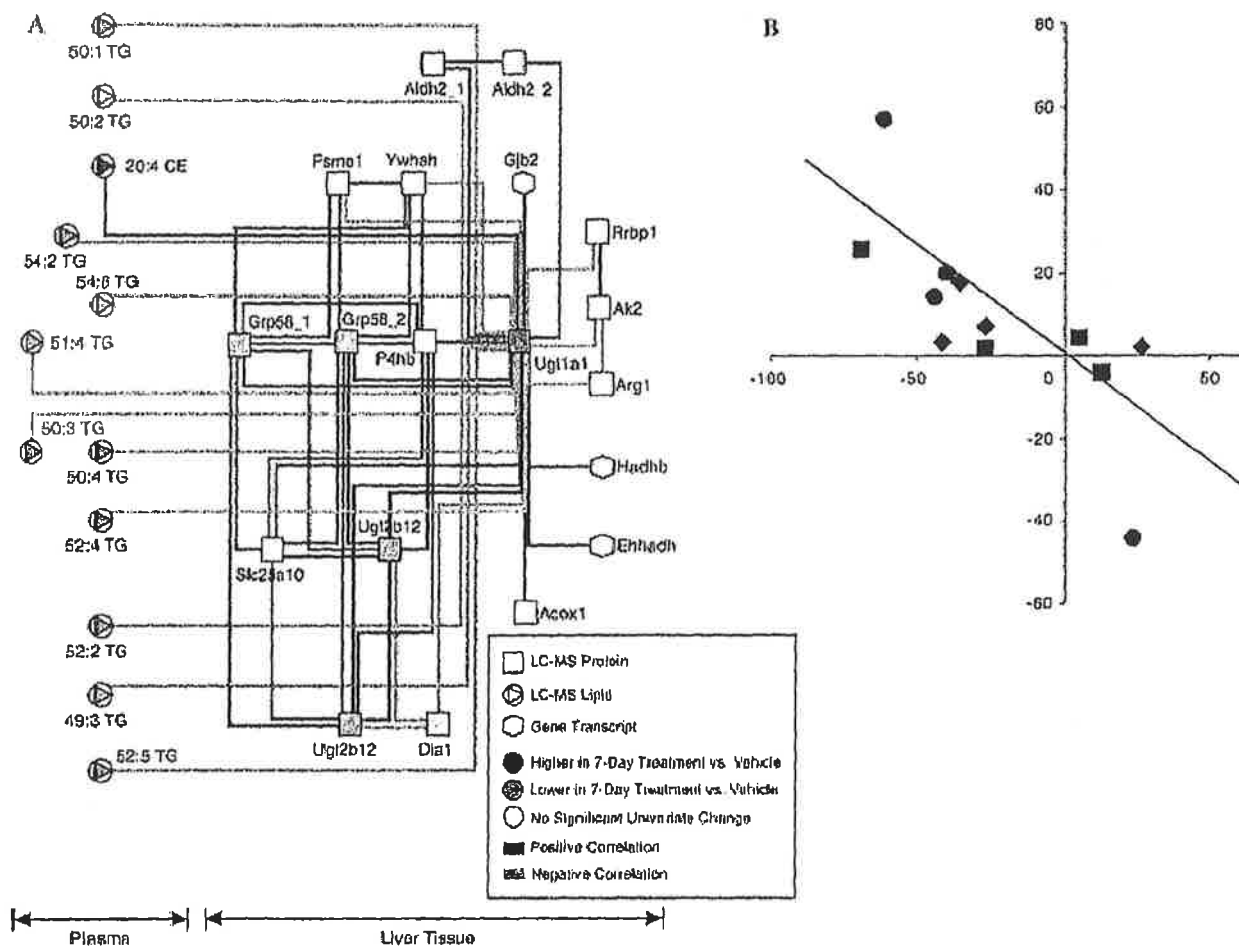

Fig. 1 (a) Ugt1a1-Centered correlation network. In the figure, plasma analytes are to the left, and liver tissue analytes are to the right; topological distance contains no information. Only those analytes within one correlation link from Ugt1a1 are shown, and correlations among plasma analytes are not shown. Network nodes are colored according to significance in the test between the vehicle-administered group and the group receiving 200 mg kg^{-1} per day for 7 days. Analytes in this network are listed in Table 1 and Table 2. Lipid nomenclature indicates total number of carbon atoms and double bonds of all fatty acid chains, separated by a colon; TG = triglyceride, CE = cholesteryl ester. (b) Representative scatter plot of normalized abundances, adjusted for group means (see Methods), of 52:4 TG in the plasma (on the abscissa) and the Ugt1a1 liver peptide AMEIAEALGR (ordinate) of the fifteen drug treated animals in the study. Closed circles (●) represent animals receiving 200 mg kg^{-1} per day for 7 days, diamonds (◆) represent animals receiving 200 mg kg^{-1} for 1 day, and squares (■) represent animals receiving 20 mg kg^{-1} per day for 7 days. The correlation

trendline illustrates the negative partial correlation between these two measurements. Axis units are arbitrary units.

Levels of liver Ugt1a1 protein, as measured by relative quantification of three distinct peptides, were found to be significantly elevated relative to the vehicle-administered group, with maximal level in the 7-day 200 mg kg^{-1} per day treatment group (Fig. 2). The relative level of the gene transcript for Ugt1a1 also exhibited similar increased abundance in all treated groups. Another member of the UDP-glucuronosyltransferase family, Ugt2b12, as measured by the proteomic liquid chromatography-mass spectrometry (LC-MS) platform, is also present in the network and is positively correlated to Ugt1a1 and up-regulated upon treatment; Ugt2b12 appears as two nodes in the network because the detected proteolytic peptides map into two distinct polypeptide isoforms.

Fig. 2 Change in abundance of the three measured peptides from hepatic Ugt1a1 for the single-day (open bars) and seven-day (closed bars) 200 mg kg^{-1} per day treatment groups relative to the vehicle-administered group of animals. Peptide amino acid sequences are indicated. Also shown for comparison is the average change of the five probesets for Ugt1a1 from the oligonucleotide microarray. All changes shown are significantly different relative to the vehicle-administered group ($p_{FDR} < 0.05$), and significant differences ($p_{FDR} < 0.05$) between 1 and 7 days of treatment are denoted by asterisks (*). Error bars represent standard error of the mean.

As shown in Fig. 1A and Table 1, there were twelve analytes measured in plasma which were found to be correlated to hepatic Ugt1a1 protein with an absolute correlation value of $|r_{partial}| > 0.60$ and a false discovery rate adjusted p-value for correlation of less than 0.15. Eleven of these plasma analytes are plasma triglycerides composed of mono- or poly-unsaturated fatty acids, all of which were found to be inversely correlated to hepatic Ugt1a1, while an arachidonic acid (20:4) containing cholesteryl ester was observed to be positively correlated to Ugt1a1 levels (Table 1).

Table 1 Plasma lipids correlated with hepatic Ugt1a1

52:4 TG	52:5 TG	50:1 TG	54:2 TG	54:6 TG	51:4 TG	50:4 TG	50:2 TG	52:2 TG	49:3 TG	50:3 TG	20:4 CE

Correlation to   −0.75 −0.74 −0.72 −0.70 −0.71 −0.68 −0.65 −0.64 −0.64 −0.61 −0.61 +0.70
Ugt1a1[a]

Table 2 Liver analytes correlated with hepatic Ugt1a1

Symbol	Common name	Correlation to Ugt1a1[b]	7-Day 200 mg kg⁻¹ per day treatment group change relative to vehicle group
Ugt2b12_2	UDP glucuronosyltransferase 2B12 precursor	+0.79	+90% ± 28%
Ugt2b12_1	UDP glycosyltransferase 2, polypeptide B4	+0.78	+89% ± 27%
Grp58_1[a]	Glucose regulated protein, 58 kDa	+0.70	+45% ± 11%
Grp58_2[a]	Glucose regulated protein, 58 kDa	+0.72	+43% ± 10%
Gjb2	Gap junction beta-2 (Connexin-26)	+0.73	+42% ± 22%
Dia1	Cytochrome-b5 reductase, microsomal form	−0.73	−24% ± 8%
Hadhb	Hydroxyacyl-coenzyme A dehydrogenase	+0.71	+25% ± 13%*
Acox1	Acyl-coenzyme A oxidase 1, peroxisomal	+0.72	n/s
Ak2	Adenylate kinase isoenzyme 2, mitochondrial	−0.70	n/s
Aldh2_1[a]	Aldehyde dehydrogenase 2, mitochondrial precursor	+0.72	n/s
Aldh2_2[a]	Aldehyde dehydrogenase 2, mitochondrial	+0.72	n/s
Arg1	Arginase 1	−0.70	n/s
Ehhadh	Enoyl-coenzyme A, hydratase	+0.72	n/s
P4hb	Protein disulfide-isomerase A3	+0.70	n/s
Psme1	Protease (prosome, macropain) 28 subunit, alpha	−0.72	n/s
Rrbp1	Ribosome-binding protein 1	−0.73	n/s
Slc25a10	Mitochondrial dicarboxylate carrier	+0.74	n/s
Ywhah	Tyrosine 3-monooxygenase	−0.71	n/s

All eleven of the Ugt1a1-correlated plasma triglycerides were uniformly and statistically significantly lower by approximately 50% in abundance in the plasma of the 7-day treated animals compared to the vehicle-administered animals (Fig. 3A). This level of decrease was similar to that of the triglyceride blood chemistry measure. In contrast, the Ugt1a1-correlated 20:4 cholesteryl ester

was significantly elevated in plasma after 7 days of treatment, $+148\% \pm 51\%$, $p_{FDR} < 0.05$, relative to the vehicle-administered group.
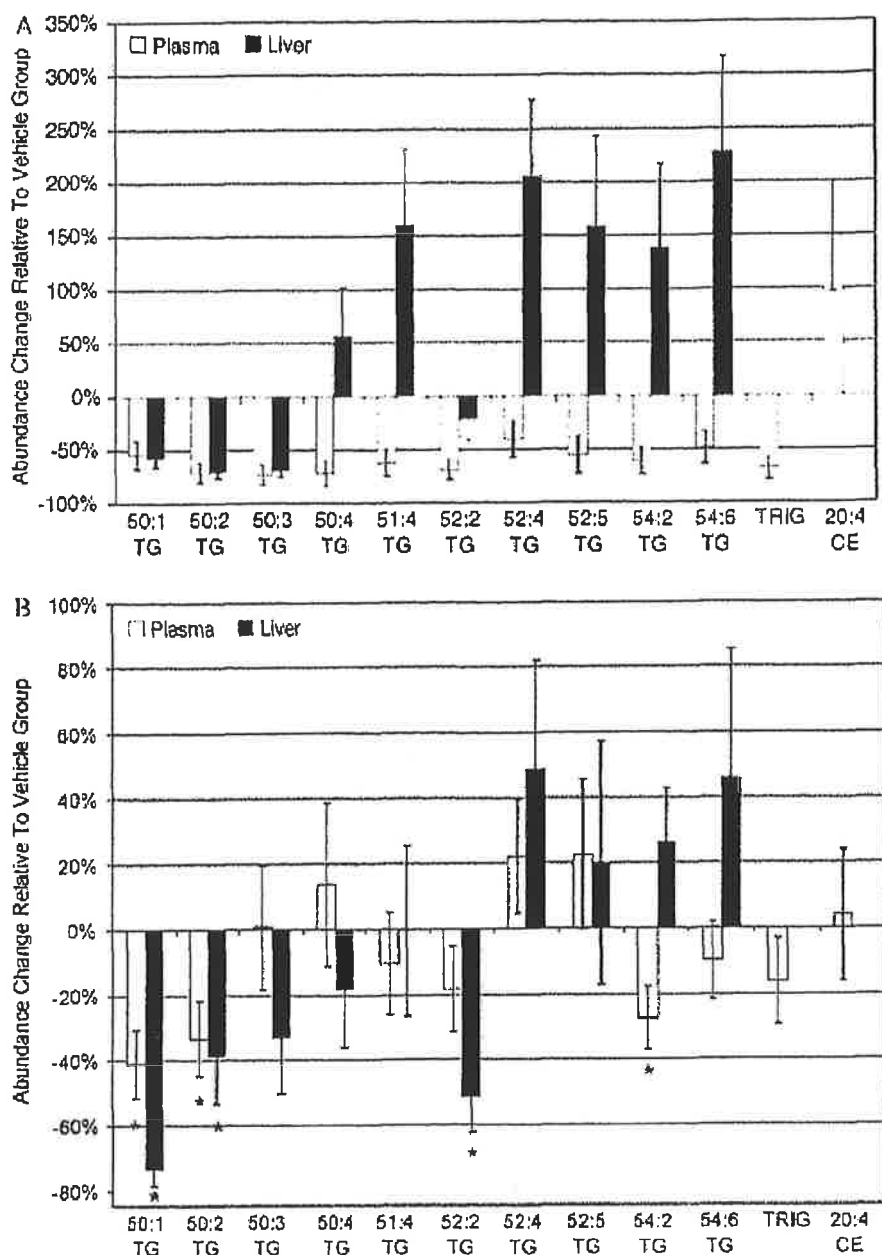

Fig. 3 (a) Change in abundance of Ugt1a1-correlated lipid levels in the 7-day 200 mg kg^{-1} per day treatment group relative to the vehicle-administered group, in plasma (open bars) and liver (closed bars). Also shown for comparison is the blood chemistry measure for triglyceride, TRIG. All changes shown are significantly different relative to the vehicle-administered group ($p_{FDR} < 0.05$). Error bars represent standard error of the mean. (b) Change in abundance of Ugt1a1-correlated lipid levels in the 1-day 200 mg kg^{-1} per day treatment group relative to the vehicle-administered group, in plasma (open bars) and liver (closed bars). Also shown for comparison is the blood chemistry measure for

triglyceride, TRIG. Significant changes relative
to the vehicle-administered group ($p_{FDR} < 0.05$)
are denoted by asterisks (*). Error bars represent
standard error of the mean.

Upon the observation of the significant inverse correlation of the eleven plasma lipids to Ugt1a1,
we proceeded to examine the corresponding abundances of these lipid species in the liver tissue (Fig.
3A and 3B). With the exception of 49:3 TG and 20:4 CE, all of the plasma lipids correlated to
Ugt1a1 were also able to be measured in liver tissue in the course of this study. Six of the Ugt1a1-
correlated triglycerides were observed to exhibit significant accumulation in the liver tissue of the 7-
day treatment group animals compared to the vehicle-administered animals (Fig. 3A). In contrast,
three triglycerides, 50:1 TG, 50:2 TG, and 50:3 TG, were found to be significantly depleted in liver
tissue upon 7 days of treatment relative to the vehicle-administered group (Fig. 3A).

After a single day of administration, the effects of the compound on the Ugt1a1-correlated plasma
triglycerides were found to be quite different than the effects observed after seven days of treatment
(Fig. 3B). In particular, two of the Ugt1a1-correlated triglycerides, 50:1 TG and 52:2 TG, were
observed to be more depleted in liver after one day of treatment than after seven days of treatment.
Levels of the Ugt1a1-correlated triglycerides in the plasma of the one-day treatment group were
either not statistically significantly changed from the vehicle group, or were significantly depleted in
abundance compared to the vehicle-administered animals. The largest one-day triglyceride plasma
level reduction was that of 50:1 TG which was observed to decrease by −41% ± 11% ($p_{FDR} < 0.05$), a
significantly larger reduction after one day of treatment than that of triglyceride as measured by
blood chemistry (Fig. 3B). The level in plasma of the Ugt1a1-correlated 20:4 cholesteryl ester was
not statistically significantly changed after one day of treatment compared to the vehicle-
administered group.

Although not part of the correlation network in Fig. 1A, the abundances in liver tissue of eight
observed phosphatidylcholine (PC) species and one octadecenoic acid containing 18:1 lyso-
phosphatidylcholine (18:1 LPC) were also found to be significantly altered. Six of these PCs and the
18:1 LPC were found to be decreased in liver tissue after seven days of treatment by between
approximately 40% and 80% (Fig. 4). In contrast, two PC species containing four double-bonds,
37:4 PC and 38:4 PC, were found to be significantly accumulated in the liver tissue upon seven days
of treatment (Fig. 4). Of these nine lipids, only 37:4 PC was statistically significantly altered after a
single day of treatment of 200 mg kg^{-1}, being elevated relative to the vehicle group by +43% ± 25%
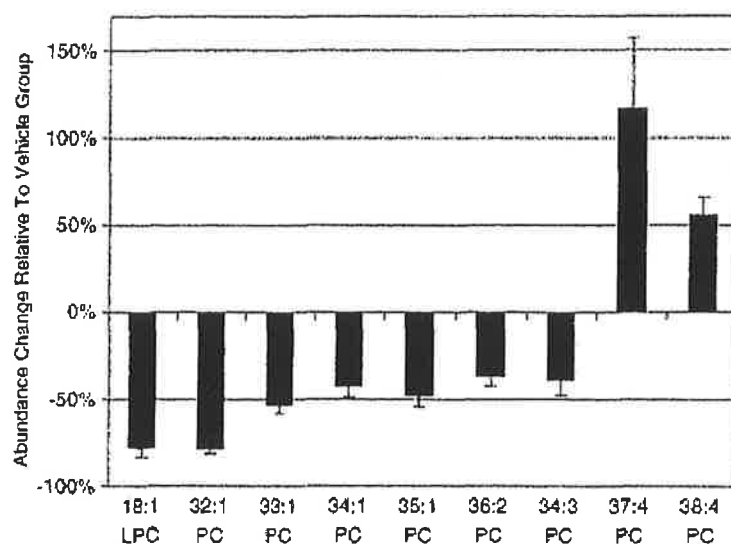($p_{FDR} < 0.05$).

Fig. 4 Change in liver tissue abundance of
observed PC and lyso-PC (LPC) levels in the 7-
day 200 mg kg^{-1} per day treatment group relative
to the vehicle-administered group. All shown

changes are significant ($p_{FDR} < 0.05$). Error bars represent standard error of the mean.

Correlation network involving plasma ornithine and hepatic urea cycle proteins

A second plasma-to-liver correlation network in the present study was centered upon measured plasma levels of ornithine. Fig. 5 is the network visualization of the correlation matrix centered upon plasma ornithine measurements showing correlation to liver analytes. There are four distinct measurements of ornithine in Fig. 5 because plasma levels of ornithine were detected using a gas chromatographic-mass spectrometric molecular profiling technique which involves molecular fragmentation prior to detection. Fig. 6 shows the experimental observations from this study overlaid on a schematic diagram of the urea cycle.
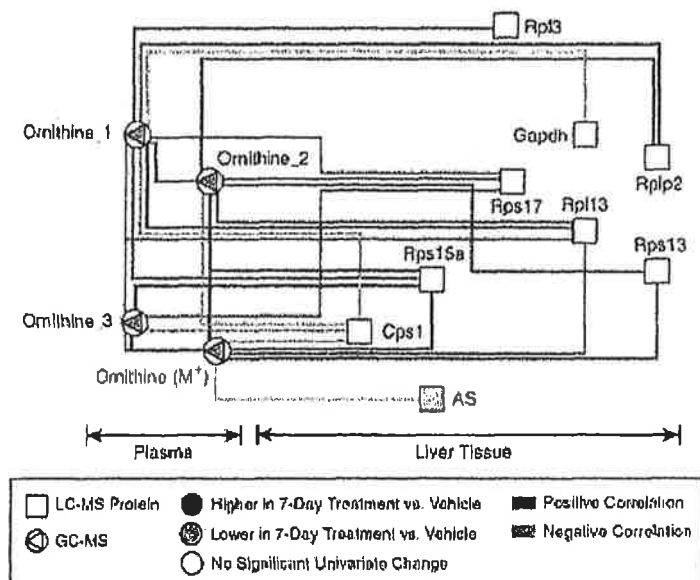
Fig. 5 Plasma ornithine-centered correlation network. 'Ornithine (M⁺)' denotes the intact ornithine molecular ion; the other three ornithine nodes are fragment ions of ornithine as detected by the mass spectrometer. Only those analytes within one correlation link from a measured ornithine ion are shown. Network nodes are colored according to significance in the test between the vehicle-administered group and the group receiving 200 mg kg⁻¹ per day for 7 days. Analytes in this network are listed in Table 3. Topological distance between network nodes contains no information.
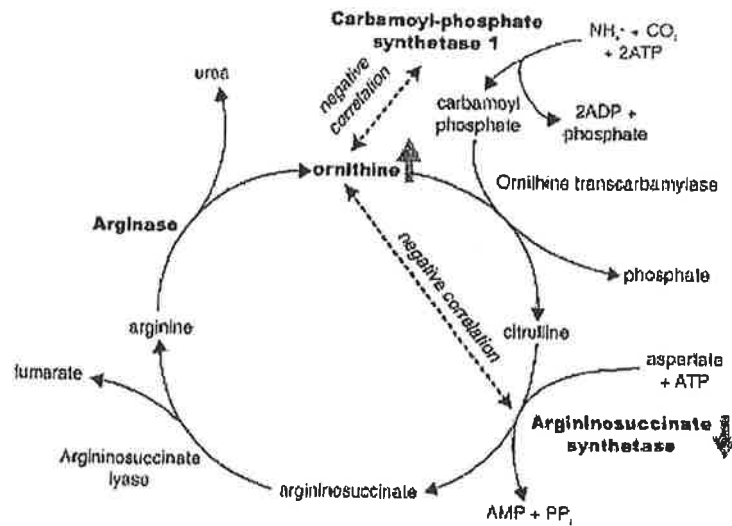
Fig. 6 Representation of the urea cycle, with components observed in the present study in bold. Dashed arrows indicate observed correlations in the ornithine-centered correlation network; block arrows indicate direction of univariate abundance change in the 7-day 200 mg kg^{-1} per day treatment group relative to the vehicle-administered group.

Plasma levels of ornithine were consistent across the four detected ornithine ions, and exhibited statistically significant elevation only in the 200 mg kg^{-1} per day 7-day treatment group relative to the vehicle-administered group of +39% ± 11% (p_{FDR} < 0.05), averaged over all four ornithine measurements (Table 3). The other members of this correlation network were all proteins measured using the mass spectrometric proteomics platform applied to liver tissue samples. Nine liver proteins were thus found to be correlated to plasma ornithine, of which three, AS (arginosuccinate synthetase), Gapdh (glyceraldehyde-3-phosphate dehydrogenase) and Rps17 (ribosomal protein S17), exhibited a significant change (p_{FDR} < 0.05) between the seven-day treatment group and the vehicle-administered group. Arginosuccinate synthetase protein levels were determined by six observed peptides; all other protein abundances in this network were determined based on the average abundances of at least three peptide measurements each.

Table 3 Analytes in plasma ornithine correlation network

Symbol	Common name	7-Day 200 mg kg^{-1} per day treatment group change relative to vehicle group[a]
Ornithine (M$^+$)	Plasma L-ornithine molecular ion	+44% ± 13%
Ornithine_1	Plasma L-ornithine molecular fragment ion	+37% ± 10%
Ornithine_2	Plasma L-ornithine molecular fragment ion	+38% ± 10%
Ornithine_3	Plasma L-ornithine molecular fragment ion	+38% ± 11%
AS	Arginosuccinate synthetase	−52% ± 7%
Gapdh	Glyceraldehyde-3-phosphate dehydrogenase	−36% ± 10%
Rps17	Ribosomal protein S17	+28% ± 11%
Cps1	Carbamoyl-phosphate synthetase 1,	n/s

	mitochondrial	
Rps13	Ribosomal protein S13	n/s
Rps15a	Ribosomal protein S15a	n/s
Rpl3	Ribosomal protein L3	n/s
Rpl13	Ribosomal protein L13	n/s
Rplp2	Ribosomal protein, large P2	n/s

[a] Significant univariate changes in abundance between 7-day treatment group and vehicle-administered group are shown ($p_{FDR} <$ 0.05); n/s = not significant ($p_{FDR} > 0.05$). The only member of this network that exhibited significant univariate change between the 1-day treatment group and the vehicle-administered group was arginosuccinate synthetase ($-40\% \pm 10\%$, $p_{FDR} < 0.05$).

Three significant negative correlations were measured in this network, between plasma ornithine and the hepatic proteins AS, Cps1 (carbamoyl-phosphate synthetase 1, mitochondrial) and Gapdh. In addition, it was found that all four measures of plasma ornithine were mutually positively correlated, as expected, with an average mutual correlation coefficient of $r_{partial} = +0.98$, $p_{FDR} < 0.05$, over all six possible pairwise correlations.

Discussion

In this work we have explored the utility of correlation networks spanning plasma and tissue to select among a plethora of drug-induced changes and prioritize plasma-based biomarkers of drug-induced alteration of hepatic function. These correlation networks were generated solely based on empirical measurements, with no *a priori* supervision or design. One such correlation network was seen to be centered upon hepatic UDP-glucuronosyltransferase A1 (Fig. 1A) and is of particular interest because of the established role of this enzyme in catalyzing a common conjugation reaction in drug metabolism.[12] Additional unpublished data together with the data presented here have demonstrated that the drug compound is a potent enzyme inducer, and thus Ugt1a1-induction and the associated correlation network which are observed reflect hepatic drug exposure. The identification of specific plasma lipids whose abundances vary co-ordinately with that of Ugt1a1, namely the eleven triglycerides and the arachidonic acid containing cholesteryl ester of the correlation network of Fig. 1A, provided sensitive surrogates in plasma for drug-induced hepatic Ugt1a1 production in the present study. The negative, or inverse, nature of the correlation between the eleven plasma triglycerides and hepatic Ugt1a1, together with the observation that these eleven triglycerides were uniformly lower in the plasma and generally accumulated in liver tissue of the 7-day treatment group (Fig. 3A), is consistent with the histopathological observations of fatty liver tissue.

In the 7-day treatment group, the eleven Ugt1a1-correlated triglycerides exhibited reductions in plasma of similar magnitude as the blood chemistry measurement of triglyceride (Fig. 3A). However, at the earlier one-day treatment point (Fig. 3B) three of the Ugt1a1-correlated triglycerides, 50:1 TG, 50:2 TG and 54:2 TG, exhibited a statistically significant reduction in circulating levels of between 30% and 40%, while the blood chemistry triglyceride measurement was not significantly altered at one day of treatment; indeed none of the conventional blood chemistry measures were significantly changed after one day after treatment. From a biomarker perspective, the 50:1 TG, 50:2 TG and 54:2 TG Ugt1a1-correlated triglyceride species may indeed be more sensitive early blood-based biomarkers of perturbed lipid homeostasis.

Interestingly, lower plasma levels of the eleven Ugt1a1-correlated triglycerides in the 7-day treatment group do not completely coincide with a simple triglyceride sequestration phenomenon, and exhibit a fatty-acid-dependent accumulation or depletion in the liver. Levels of three of the plasma triglycerides in the Ugt1a1 network, specifically those comprising fatty acids with relatively fewer unsaturated double bonds and fewer carbon atoms—50:1 TG, 50:2 TG, and 50:3 TG—were in fact observed to be decreased rather than increased in liver tissue after seven days of treatment, and by approximately the same amount as the depletion in plasma, indicating that the export process for these triglycerides was not as substantially affected.

The observed pattern of hepatic triglyceride accumulation in the Ugt1a1 network prompted us to investigate changes in phosphatidylcholine (PC) levels in liver. Synthesis of PCs is an important

component of very low-density lipoprotein (VLDL) formation, and as such PCs are essential for proper triglyceride export. The changes in hepatic PC levels upon treatment (Fig. 4) may indicate a significant, and fatty-acid composition-dependent, alteration of phosphatidylcholine synthesis.

Similar results have been described in recent studies. Watkins *et al.*[18] report that inhibition of phosphatidylethanolamine-*N*-methyltransferase (PEMT), a component in one of the primary pathways for hepatic PC biosynthesis, effects a significant and specific accumulation of relatively long-chain poly-unsaturated fatty acids in hepatic TGs. The measured positive co-ordination between plasma 20:4 cholesteryl ester (CE) accumulation and Ugt1a1 induction by treatment is consistent with recent reports of a relationship between PEMT activity and PC biosynthesis necessary to ensure proper flux of cholesterol esters between liver and plasma;[18,19] indeed, the accumulation of 20:4 CE and its correlation with Ugt1a1 drug metabolism may indicate a drug-induced inability to maintain this flux at normal levels. Taken together, one hypothesis that may be pursued in subsequent studies is that the drug compound or its metabolites impair triglyceride export by altering the synthesis of phosphatidylcholine, impairing VLDL formation and triglyceride export from the liver and contributing to the observed steatosis.

The Ugt1a1-centered network of Fig. 1A also contains components within liver tissue which are involved in mitochondrial and peroxisomal fatty acid beta-oxidation, indicating that these processes were co-ordinately affected by drug metabolism as well. In the correlation network, the gene transcript for hydroxyacyl-coenzyme A dehydrogenase (Hadhb) was observed to be increased in abundance as a result of drug treatment and was found to be positively correlated with Ugt1a1 induction. Hadhb is a multienzyme complex subunit involved in mitochondrial fatty acid oxidation. There were also components of peroxisomal beta-oxidation present in the Ugt1a1-centered network. The three enzymes of the classic peroxisomal beta-oxidation cycle are acyl-coenzyme A oxidase, enoyl-coenzyme A hydratase, and 3-ketoacyl-CoA thiolase. In the Ugt1a1 network we observed two of these enzymes positively correlated to Ugt1a1: acyl-coenzyme A oxidase 1 (Acox1), the enzyme which catalyzes the first step in peroxisomal beta-oxidation of fatty acids, and enoyl-coenzyme A hydratase (Ehhadh), although enoyl-coenzyme A hydratase is also found in the mitochondria. While these were significantly positively correlated to Ugt1a1 and therefore drug metabolism processes, they did not exhibit a statistically significant univariate change upon drug treatment; this may represent a case in which there was a lack of sufficient statistical power to distinguish a univariately significant group-based change, but sufficient statistical power to identify significant pairwise correlation. Peroxisomes are a primary site for the beta-oxidation of very long-chain fatty acids,[20] and the fatty-acid composition dependence of accumulated liver TGs observed in Fig. 3A may further reflect modulated peroxisomal activity. Taken together, the Ugt1a1 network findings may implicate mitochondrial and peroxisomal beta-oxidation processes being co-ordinately affected by treatment.

Both the Ugt1a1-centered network and the plasma ornithine centered network of Fig. 5 include components of the urea cycle. Because urea cycle enzymes exist almost exclusively in the liver, they and their products are of special interest as specific markers of liver injury.[21] The observed treatment-related inverse correlations in the ornithine-centered network between plasma ornithine and argininosuccinate synthetase (AS), and between plasma ornithine and carbamoyl-phosphate synthase I (Cps1), are consistent with a drug-induced deficiency in the first steps of the urea cycle. The observed increase in plasma ornithine levels of approximately +40% in the 7-day 200 mg kg^{-1} per day treatment group relative to the vehicle group, together with the observed decrease in AS protein levels in the same comparison of approximately −50%, is also consistent with the observed inverse relationship between ornithine and AS. While ornithine is not a direct substrate of AS, it is a substrate of ornithine transcarbamylase, the enzyme which precedes AS in the urea cycle and produces citrulline, the substrate for AS.

It has been reported that juvenile visceral steatosis mice have lower levels of urea cycle enzymes in the liver.[22] These mice have a defect in a carnitine transporter, which results in both steatosis and lower levels of all urea cycle enzymes.[23,24] It has also been shown that long-chain fatty acids can suppress glucocorticoid-mediated induction of genes encoding urea cycle enzymes.[25] Taken together, these studies indicate that there may be a link between fatty acid metabolism and the urea cycle that is observable through monitoring of plasma biomarkers such as ornithine.

Plasma ornithine is also seen to be positively correlated with a number of ribosomal proteins. These findings are intriguing given the role of ornithine as a starting material for polyamine

biosynthesis and the reported potential relationship between ribosomal protein expression and the polyamine synthesis pathway.[26,27]

Experimental

Animal experiments

Twenty (20) male Wistar Hannover rats, approximately 2 months of age at study initiation (220 g to 300 g total mass each) were used. Animals were housed 2 or 3 per cage during an acclimatization period of 10 days prior to dosing. During the course of the study, 5 animals per group were housed in individual metabolism cages. A first group of five animals was dosed by oral gavage in 10 ml kg^{-1} vehicle solution only (0.5% w/v hydroxypropyl methylcellulose in 0.1% w/v aqueous polysorbate 80); a second group of five animals was dosed with 200 mg kg^{-1} per day of an experimental drug compound in 10 ml kg^{-1} vehicle solution for 7 days beginning on day 1 and necropsied on day 8 when tissues were taken for pathology; a third group of five animals was dosed with 20 mg kg^{-1} per day with the experimental drug compound in 10 ml kg^{-1} vehicle solution for 7 days beginning on day 1 and necropsied on day 8; a fourth group of five animals received a single dose of the compound (200 mg kg^{-1}) on day 1 and were necropsied 24 hours after dosing. All experiments were approved by and performed in compliance with institutional guidelines of the Molecular Toxicology and Safety Assessment Division of AstraZeneca R&D.

Sample collection

Blood plasma samples were taken from the orbital plexus prior to necropsy. Selected organs for histopathology, including liver, kidney, heart, adrenal glands, lungs, pancreas, spleen, stomach, thymus and testes, were retained at necropsy. Samples were fixed in 10% (w/v) phosphate-buffered formalin and embedded in paraffin wax. Sections were cut and stained with haematoxylin and eosin for light microscopy. Samples from the liver were also freeze-sectioned and stained for fat with Oil Red-O. In addition, sections from the liver (right lateral lobe) were cut into 100 mg pieces and fresh frozen in liquid nitrogen at necropsy for proteomics, transcriptomics and metabolomics.

LC-MS analysis of plasma and liver lipids

Plasma samples were prepared for lipid analysis by adding 600 μl of isopropanol to 150 μl of whole plasma. Samples were vortexed to mix, and centrifuged to precipitate and remove proteins, and the supernatant was removed for analysis. To prepare samples for LC-MS analysis, 400 μl of water were added to 100 μl of the supernatent, and 200 μl of this mixture were transferred to a Waters 717 autosampler and separated on a Waters 600-MS HPLC system with a C4 column at 1 ml min^{-1} with a gradient from 5% methanol–water to 100% methanol containing 10 mM ammonium acetate and 0.1% formic acid. The output of the HPLC was directed to a ThermoFinnigan TSQ 700/7000 (ThermoFinnigan, San Jose CA) equipped with electrospray to aquire LC-MS and MS/MS spectra. The scan cycle consisted of a single full scan (1 s per scan) mass spectrum acquired over m/z 200–1700 in the positive ion mode.

Twenty-five milligrams of thawed liver tissue was homogenized by ultrasound-based acoustic disruption using a Covaris E100 instrument (Covaris Inc., Woburn, MA). Homogenate samples were treated with isopropanol, vortexed and centrifuged to precipitate protein and to extract lipids, and the supernatant was collected for analysis as described above. The LC-MS analysis proceeded as described above for plasma.

LC-MS data files were converted to CDF format using Xcalibur software (ThermoFinnigan, San Jose, CA). The converted files were evaluated with in-house peak detection, peak integration and peak alignment software (BG Medicine, Waltham MA and TNO Pharma, Zeist, The Netherlands). This resulted in approximately five-hundred molecular ion peaks per plasma sample data set and a similar number per liver sample data set. Noise filtering was subsequently performed for each mass trace data set by applying an in-house signal processing algorithm (BG Medicine, Waltham, MA and TNO Pharma, Zeist, The Netherlands) that is based on information entropy evaluation.[11,28] Subsequently, a minimum ion count intensity threshold was applied, which resulted in 168 aligned

peaks per sample for plasma analyses, and 112 aligned peaks per sample for liver tissue analyses. The resulting LC-MS data set peak intensities were normalized using an ANOVA multiplicative model with parameters estimated using an interative maximum likelihood method.[29,30]

Identification of molecular ion peaks was performed by tandem mass spectrometry and by matching with in-house mass spectra databases. Subsequently, non-monoisotopic peaks were removed from the data set, and in the case of multiple ion adduct peaks, the highest intensity peak was selected. Thus the final data set for plasma lipids consisted of 80 unique identified lipids, and the final data set for liver tissue comprised 43 unique identified lipids. All subsequent statistical analyses were performed on these identified lipid analytes. Exogenous calibration standards for every analyte were not included in the analyses, and as such quantification is relative and not absolute.

LC-MS analysis of liver proteins

One-hundred milligrams of tissue dissected from thawed liver sample were cut into four 25 mg pieces. Each piece was homogenized by ultrasound-based acoustic disruption using a Covaris E100 instrument (Covaris Inc., Woburn, MA). Unbroken cells, nuclei and extracellular debris were removed using low-g centrifugation, and the resulting homogenate was subjected to membrane/cytosol fractionation using high-speed ultracentrifugation. Cytosolic proteins were subjected to C4 reversed-phase column chromatography to isolate three fractions for analysis. Each of the three protein cytosolic fractions was digested using trypsin, and from each fraction acidic peptides were isolated by anion exchange chromatography and desalted by reversed-phase column chromatography prior to LC-MS analysis. The membrane-bound protein fraction was isolated by R1-C18 reversed-phase chromatography and trypsin digested. Digestion reagents and undigested and partially digested materials were separated from the tryptic peptide fraction by HPLC chromatography. The resulting membrane tryptic peptide fraction was dried *in vacuo*. Cytosolic peptides from the three acidic peptide sets and the membrane tryptic peptide fraction were measured by reversed-phase LC-ESI-MS using a Waters HPLC system and Micromass Q-Tof mass spectrometer (Waters Corp., Milford MA). Raw LC-MS data files were analyzed using in-house peak detection, peak integration and peak alignment software (BG Medicine, Waltham MA). Noise was filtered from the LC-MS data sets using the entropy-based signal processing algorithm described above. Subsequently applying a minimum ion count intensity threshold, and requiring peptide ion peaks to be present in at least 17 of the 20 samples, yielded 2689 aligned LC-MS peptide ion peaks across the four fractions per sample. The resulting data set peak intensities were normalized using the ANOVA multiplicative model as described above.

Peptide MS/MS spectra were acquired during LC-MS/MS analyses that were interspersed among LC-MS profiling scans at a frequency of approximately one MS/MS acquisition per three MS acquisitions. Spectra were converted to Sequest .dta files using MassLynx PeptideAuto software (Waters Corp., Milford MA). In-house proteomics data analysis algorithms and software (BG Medicine, Waltham, MA) were used to incorporate Sequest, Mascot, X!Tandem and manual MS/MS spectrum annotation results for final peptide identification. Peptides were matched to proteins and isoforms using in-house algorithms (BG Medicine, Waltham, MA) incorporating public sequence databases. Searches were initially filtered to include only mammalian sequences, and subsequently to the *Rattus norvegicus* genome. Identifications were made of 1723 peptides, and all subsequent statistical analyses were performed on 1723 distinct individual peptide measurements, and not on reconstructed protein measurements. Protein or peptide calibration standards were not included in the mass spectrometric analyses, and as such quantification is relative and not absolute.

Plasma GC-MS analysis

Plasma samples were extracted with methanol, and extracts were centrifuged to remove precipitated proteins. Extract supernatants were derivatized by oximation and subsequent silylation. Nine internal standards were added, namely cholic acid-D4, alanine-D4, leucine-D3, DFBP, Phe-D5-2TMS, glutamic acid-D3, trifluoroacetylanthracene, glucose-D7, and DCHP. The derivatized samples were loaded in an ATAS Focus autosampler and molecules were separated on an Agilent 6890 system (Agilent, Santa Clara, CA) with a programmed temperature vaporizer injector. Molecules were analyzed with electron impact ionization on an Agilent 5973 MSD system (Agilent, Santa Clara,

CA) in full scan monitoring mode. Ion peaks measured across samples were aligned using an in-house algorithm which uses the exogenous internal standards to yield alignment coefficients for each data set. After further applying a minimum ion count intensity threshold, and requiring ion peaks to be present in at least 17 of the 20 samples, 924 ion peaks remained in the data set. Normalization scaling factors for each sample data set were calculated using internal standards and maximum likelihood optimization as described above and in Hartemink *et al.*[29] Ion peaks were associated with parent analytes using an in-house database of electron impact ionization spectra which were matched to the obtained fragmentation spectra; these 427 identified ion peaks were used as input to subsequent statistical analyses. Exogenous calibration standards for every analyte were not included, and as such quantification is relative and not absolute.

Liver mRNA analysis

RNA was isolated from frozen liver tissue using an internal protocol (AstraZeneca, Södertälje, Sweden). Affymetrix GeneChip oligonucleotide microarray analysis was carried out on labeled and fragmented cRNA prepared from total RNA with the Affymetrix U34A array, version December 2003, and analyzed using the Affymetrix GeneChip Scanner system (Affymetrix, Santa Clara, CA). The CEL files were quantile normalized, and positional-dependent nearest-neighbor output was computed for subsequent analysis as reported elsewhere.[31] The resulting quantification from the Affymetrix GeneChip platform is relative and not absolute.

Statistical and correlation network analyses

Molecular profiling data were natural-logarithm transformed and analyzed for each bioanalytical platform by univariate analysis of variance (ANOVA) and ANOVA contrasts. Because the number of analytes measured with any of the platforms used far exceeds the number of samples, the p-value resulting from ANOVA was adjusted to the false discovery rate-adjusted p-value (p_{FDR}).[32] Univariate results are expressed as mean ± standard error of mean. Partial pairwise Pearson correlations controlling for treatment group means, $r_{partial}$, were calculated for all pairs of analytes measured.[16] Partial correlations set the mean value of the analyte level to zero for each of the three treatment groups; this is done in order to avoid trivial correlations that are driven by group mean differences. Statistical significances of correlation values were determined using the Student's t-transformation and were also adjusted to the false discovery rate-adjusted p-value.[32] Data from the vehicle group were not used in correlation calculations in this study; only the fifteen animals treated with the drug were considered (namely, five from each of the three dosage groups of 200 mg kg^{-1} per day for 7 days, 20 mg kg^{-1} per day for 7 days and 200 mg kg^{-1} for a single day).

The criteria for inclusion in the correlation networks discussed in the text were as follows. In order to merit inclusion in the correlation network of Fig. 1A, a plasma analyte must have exhibited a partial correlation averaged across the three observed Ugt1a1 peptides of Fig. 2 of $|r_{partial}| \geq 0.60$ group, with a false discovery rate-adjusted p-value less than 0.15; for analyte correlations within liver tissue, an analyte must have exhibited a partial correlation averaged across all observed peptides or probesets, for proteins and gene transcripts respectively, of $|r_{partial}| \geq 0.70$ with a false discovery rate-adjusted p-value less than 0.15. In Fig. 1A, correlations between plasma analytes are not shown, for clarity. For the correlation network of Fig. 5, the partial correlation, $r_{partial}$, of each liver peptide to each plasma ornithine fragment was calculated. In order for a liver protein to merit inclusion into the final correlation network, the average over all $r_{partial}$ values across all peptides of that protein and all observed plasma ornithine fragments must have been $|r_{partial}| \geq 0.60$ with a false discovery rate-adjusted p-value less than 0.15. Correlation network visualizations were generated using in-house software developed by BG Medicine (BG Medicine, Waltham, MA).

Conclusion

In summary, the use of correlation network analyses spanning different tissues to prioritize and select plausible candidate biomarkers in biological fluids, as demonstrated in the present study, has been shown to be a practical and compelling approach to the integration and application of high

throughput molecular profiling data. While empirical correlations do not confer information about causality or directionality, they are shown here to bestow an important level of plausibility that the selected plasma biomarkers reflect processes relevant to hepatic drug exposure.

Indeed, expecting that observed correlations are consequences of the underlying biochemical reactions and enzymatic regulations of metabolism, a number of intriguing relationships were revealed which would merit subsequent study, such as the role of aldehyde dehydrogenases in the Ugt1a1 network, or the drug-induced inverse correlation of glyceraldehyde-3-phosphate dehydrogenase (Gapdh) to plasma ornithine. However, these explorations were ancillary to the discovery and selection of plasma biomarkers in the current study.

The challenge of sorting through often hundreds of observed biomolecular changes upon drug exposure in pursuit of an appropriate biomarker in drug discovery and development is a significant one for scientists, industry and regulators alike. It has become evident that large group-specific univariate changes often do not capture the relevant complexity of biological processes, and that further empirical approaches such as the one presented here will be necessary for successful biomarker selection.

References

1 W. H. Heijne, A. S. Kienhuis, B. van Ommen, R. H. Stierum and J. P. Groten, Systems toxicology: applications of toxicogenomics, transcriptomics, proteomics and metabolomics in toxicology, *Expert Rev. Proteomics*, 2005, **2**(5), 767–780.

2 L. Hood and R. M. Perlmutter, The impact of systems approaches on biological problems in drug discovery, *Nat. Biotechnol.*, 2004, **22**(10), 1215–1217 [Links].

3 T. G. Kleno, B. Kiehr, D. Baunsgaard and U. G. Sidelmann, Combination of 'omics' data to investigate the mechanism(s) of hydrazine-induced hepatotoxicity in rats and to identify potential biomarkers, *Biomarkers*, 2004, **9**(2), 116–138 [Links].

4 D. G. Robertson, Metabonomics in toxicology: a review, *Toxicol. Sci.*, 2005, **85**(2), 809–822 [Links].

5 J. van der Greef and R. N. McBurney, Innovation: rescuing drug discovery *in vivo* systems pathology and systems pharmacology, *Nat. Rev. Drug Discovery*, 2005, **4**(12), 961–967 [Links].

6 D. E. Amacher, A toxicologist's guide to biomarkers of hepatic response, *Hum. Exp. Toxicol.*, 2002, **21**(5), 253–262 [Links].

7 D. E. Amacher, R. Adler, A. Herath and R. R. Townsend, Use of proteomic methods to identify serum biomarkers associated with rat liver toxicity or hypertrophy, *Clin. Chem.*, 2005, **51**(10), 1796–1803 [Links].

8 J. Gao, L. A. Garulacan, S. M. Storm, G. J. Opiteck, Y. Dubaquie, S. A. Hefta, D. M. Dambach and A. R. Dongre, Biomarker discovery in biological fluids, *Methods*, 2005, **35**(3), 291–302 [Links].

9 R. J. Mortishire-Smith, G. L. Skiles, J. W. Lawrence, W. Spence, A. W. Nicholls, B. A. Johnson and J. K. Nicholson, Use of metabonomics to identify impaired fatty acid metabolism as the mechanism of a drug-induced toxicity, *Chem. Res. Toxicol.*, 2004, **17**(2), 165–173 [Links].

10 C. B. Clish, E. Davidov, M. Oresic, T. N. Plasterer, G. Lavine, T. Londo, M. Meys, P. Snell, W. Stochaj, A. Adourian, X. Zhang, N. Morel, E. Neumann, E. Verheij, J. T. Vogels, L. M. Havekes, N. Afeyan, F. Regnier, J. van der Greef and S. Naylor, Integrative biological analysis of the APOE*3-leiden transgenic mouse, *Omics*, 2004, **8**(1), 3–13.

11 E. Davidov, C. B. Clish, M. Oresic, M. Meys, W. Stochaj, P. Snell, G. Lavine, T. R. Londo, A. Adourian, X. Zhang, M. Johnston, N. Morel, E. W. Marple, T. N. Plasterer, E. Neumann, E. Verheij, J. T. Vogels, L. M. Havekes, J. van der Greef and S. Naylor, Methods for the differential integrative omic analysis of plasma from a transgenic disease animal model, *Omics*, 2004, **8**(4), 267–288.

12 K. Morgenthal, W. Weckwerth and R. Steuer, Metabolomic networks in plants: transitions from pattern recognition to biological interpretation, *Biosystems*, 2006, **83**(2–3), 108–117 [Links].

13 M. Oresic, C. B. Clish, E. J. Davidov, E. Verheij, J. Vogels, L. M. Havekes, E. Neumann, A.

Adourian, S. Naylor, J. van der Greef and T. Plasterer, Phenotype characterisation using integrated gene transcript, protein and metabolite profiling, *Appl. Bioinf.*, 2004, **3**(4), 205–217.

14 R. Steuer, J. Kurths, O. Fiehn and W. Weckwerth, Observing and interpreting correlations in metabolomic networks, *Bioinformatics*, 2003, **19**(8), 1019–1026 [Links].

15 A. J. Butte and I. S. Kohane, Mutual information relevance networks: functional genomic clustering using pairwise entropy measurements, *Pac. Symp. Biocomput. 2000*, 2000, 418–429.

16 H. Blalock, *Causal inferences in nonexperimental research*, UNC Press, Chapel Hill, NC, 1961.

17 G. J. Dutton, *Glucuronidation of Drugs and Other Compounds*, CRC Press, Boca Raton, Florida, 1980.

18 S. M. Watkins, X. Zhu and S. H. Zeisel, Phosphatidylethanolamine-*N*-methyltransferase activity and dietary choline regulate liver-plasma lipid flux and essential fatty acid metabolism in mice, *J. Nutr.*, 2003, **133**(11), 3386–3391 [Links].

19 A. A. Noga and D. E. Vance, A gender-specific role for phosphatidylethanolamine *N*-methyltransferase-derived phosphatidylcholine in the regulation of plasma high density and very low density lipoproteins in mice, *J. Biol. Chem.*, 2003, **278**(24), 21851–21859 [Links].

20 T. Hu, P. Foxworthy, A. Siesky, J. V. Ficorilli, H. Gao, S. Li, M. Christe, T. Ryan, G. Cao, P. Eacho, M. D. Michael and L. F. Michael, Hepatic peroxisomal fatty acid beta-oxidation is regulated by liver X receptor-alpha, *Endocrinology*, 2005, **146**(12), 5380–5387 [Links].

21 M. Ikemoto, S. Tsunekawa, Y. Toda and M. Totani, Liver-type arginase is a highly sensitive marker for hepatocellular damage in rats, *Clin. Chem.*, 2001, **47**(5), 946–948 [Links].

22 M. Tomomura, Y. Imamura, M. Horiuchi, T. Koizumi, H. Nikaido, J. Hayakawa and T. Saheki, Abnormal expression of urea cycle enzyme genes in juvenile visceral steatosis (jvs) mice, *Biochim. Biophys. Acta*, 1992, **1138**(2), 167–171 [Links].

23 Y. Imamura, T. Saheki, H. Arakawa, T. Noda, T. Koizumi, H. Nikaido and J. Hayakawa, Urea cycle disorder in C3H-H-2 degree mice with juvenile steatosis of viscera, *FEBS Lett.*, 1990, **260** (1), 119–121 [Links].

24 K. Yokogawa, M. Yonekawa, I. Tamai, R. Ohashi, Y. Tatsumi, Y. Higashi, M. Nomura, N. Hashimoto, H. Nikaido, J. Hayakawa, J. Nezu, A. Oku, M. Shimane, K. Miyamoto and A. Tsuji, Loss of wild-type carrier-mediated L-carnitine transport activity in hepatocytes of juvenile visceral steatosis mice, *Hepatology*, 1999, **30**(4), 997–1001 [Links].

25 M. Tomomura, A. Tomomura, M. A. Dewan and T. Saheki, Long-chain fatty acids suppress the induction of urea cycle enzyme genes by glucocorticoid action, *FEBS Lett.*, 1996, **399**(3), 310–312 [Links].

26 C. A. Panagiotidis, S. C. Huang and E. S. Canellakis, Relationship of the expression of the S20 and L34 ribosomal proteins to polyamine biosynthesis in *Escherichia coli*, *Int. J. Biochem. Cell Biol.*, 1995, **27**(2), 157–168 [Links].

27 S. C. Huang, C. A. Panagiotidis and E. S. Canellakis, Transcriptional effects of polyamines on ribosomal proteins and on polyamine-synthesizing enzymes in *Escherichia coli*, *Proc. Natl. Acad. Sci. U. S. A.*, 1990, **87**(9), 3464–3468 [Links].

28 R. R. Coifman and M. V. Wickerhauser, Wavelets, adapted waveforms and de-noising, *Electroencephalogr. Clin. Neurophysiol.*, 1996, **45**, 57–78 [Links].

29 A. Hartemink, D. Gifford, T. Jaakkola and R. Young, Maximum likelihood estimation of optimal scaling factors for expression array normalization, *Proc. SPIE–Int. Soc. Opt. Eng.*, 2001, **4266**, 132–140 [Links].

30 M. K. Kerr, M. Martin and G. A. Churchill, Analysis of variance for gene expression microarray data, *J. Comput. Biol.*, 2000, **7**, 819–837 [Links].

31 L. Zhang, M. F. Miles and K. D. Aldape, A model of molecular interactions on short oligonucleotide microarrays, *Nat. Biotechnol.*, 2003, **21**(7), 818–821 [Links].

32 Y. Benjamini and Y. Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing, *J. R. Stat. Soc. B*, 1995, **57**, 289–300.

Footnote