Endocrinology. First published ahead of print March 27, 2008 as doi:10.1210/en.2007-1747

Bioinformatics analysis of targeted metabolomics – uncovering old and new tales of diabetic mice under medication

Elisabeth Altmaier[1], Steven L. Ramsay[2,3], Armin Graber[2,4], Hans-Werner Mewes[1,5], Klaus M. Weinberger[2], Karsten Suhre[1,6,*]

[1] Institute of Bioinformatics and Systems Biology, Helmholtz Zentrum München, German Research Center for Environmental Health (GmbH), Ingolstädter Landstraße 1, 85764 Neuherberg, Germany

[2] Biocrates life sciences AG, Innrain 66, A-6020, Innsbruck, Austria

[3] now at CSL Limited, 45 Poplar Road, Parkville, VIC 3052, Australia

[4] now at Institute for Bioinformatics, UMIT, A-6060 Hall in Tirol, Austria

[5] Department of Genome-oriented Bioinformatics, Wissenschaftszentrum Weihenstephan, Technische Universität München, D-85350 Freising, Germany

[6] Faculty of Biology, University of Munich (LMU), Großhaderner Straße 2, D-82152 Planegg-Martinsried, Germany

[*] Corresponding author: Prof. Dr. Karsten Suhre, Tel. +49-89-3187-2627, Fax: +49-89-3187-3585, e-mail: karsten.suhre@helmholtz-muenchen.de

Grants:

The first author (E.A.) is funded by Fugato-Qualipid grant FKZ 0313391.

Keywords: Type II Diabetes, Metabolism, Phospholipids, Acylcarnitines, Reducing Sugars, Data Analysis, ANOVA, Rosiglitazone, Diabetic Mice, Targeted Metabolomics, Drug Testing, Biomarker Discovery, Electrospray Ionization (ESI), Tandem Mass Spectrometry

Acknowledgements: The authors wish to thank Michael Mader, Gabi Kastenmüller, and Werner Römisch-Margl for helpful discussions, and Robert Meyer for critical reading of the manuscript.

Running head: Targeted metabolomics of diabetic mice

Disclosure statement:

K.W. is an employee of *Biocrates life sciences AG*. This private company offers products and services in the field of targeted quantitative metabolomics research. S.L.R. and A.G. were previously employed by *Biocrates life sciences AG*. E.A., H.W.M., and K.S. have nothing to disclose.

Copyright (C) 2008 by The Endocrine Society

Abstract

Metabolomics is a powerful tool for identifying both known and new disease-related perturbations in metabolic pathways. In preclinical drug testing, it has a high potential for early identification of drug off-target effects. Recent advances in high-precision high-throughput mass spectrometry have brought the Metabolomics field to a point where quantitative, targeted, metabolomic measurements with ready-to-use kits allow for the automated in-house screening for hundreds of different metabolites in large sets of biological samples. Today, the field of metabolomics is, arguably, at a point where transcriptomics was about five years ago. This being so, the field has a strong need for adapted bioinformatics tools and methods. In this paper, we describe a systematic analysis of a targeted quantitative characterization of more than 800 metabolites in blood plasma samples from healthy and diabetic mice under rosiglitazone treatment. We show that known and new metabolic phenotypes of diabetes and medication can be recovered in a statistically objective manner. We find that concentrations of methylglutarylcarnitine are oppositely impacted by rosiglitazone treatment of both healthy and diabetic mice. Analyzing ratios between metabolite concentrations dramatically reduces the noise in the dataset, allowing for the discovery of new potential biomarkers of diabetes, such as the N-hydroxyacyloylsphingosyl-phosphocholines SM(OH)28:0 and SM(OH)26:0. Using a hierarchical clustering technique on partial eta-squared values, we identify functionally related groups of metabolites, indicating a diabetes-related shift from lysophosphatidylcholine to phosphatidylcholine levels. The bioinformatics data analysis approach introduced here can be readily generalized to other drug-testing scenarios and to other medical disorders.

INTRODUCTION

Metabolomics is defined as the comprehensive quantitative measurement of low molecular weight compounds systematically covering the key metabolites, which ideally represent the whole range of pathways of intermediary metabolism. In a systems biology approach, it provides a functional readout of changes determined by genetic blueprint, regulation, protein abundance and modification, and environmental influence. Other functional genomics technologies, such as transcriptomics and proteomics are highly valuable, but merely indicate the potential cause for phenotypic response. They do not necessarily predict drug effects, toxicological response or disease states at the phenotype level unless functional validation is added. Metabolomics can bridge this information gap by depicting such functional information since metabolite differences in biological fluids and tissues provide the closest link to the various phenotypic responses. It is thus clear that a promising approach to identifying possible functional relationships between medication and medical phenotype lies in the extensive characterization of the largest possible number of metabolites from relevant or potentially impacted metabolic pathways.

In recent years, electrospray ionization (ESI) tandem mass spectrometry (MS/MS), often applied in concert with an initial liquid or gas phase chromatography purification step, has been used in a number of metabolomics studies. Rolinski et al. (1) showed that ESI-MS/MS is a highly sensitive, linear, and sufficiently precise method for the quantitative determination of amino acids and acylcarnitines in mouse blood. The method also allows large-scale screening applications when speed and cost effectiveness are mandatory. More generally, these authors suggest that ESI-MS/MS may be used to improve screening for inherited metabolic diseases, such as amino acid disorders, disorders related to carnitine metabolism, peroxisomal disorders, disorders in cholesterol, steroid, and lipid metabolism, lysosomal storage disorders, congenital disorders of glycosylation, and disorders of purine and pyrimidine metabolism. For instance, Butler at al. (2) analyzed serum samples from patients with clinical vasculitis compared to healthy individuals. In a pilot study, Paige et al. (3) evaluated the potential of metabolomics in a study of older depressed patients, and Kaddurah-Daouk et al. (4) used a specialized metabolomics platform for the quantification of over 300 polar and nonpolar lipid metabolites to evaluate global lipid changes in schizophrenia both before and after treatment

with three commonly used atypical antipsychotics. Rozen et al. (5) show that in motor neuron diseases, and, more specifically, in amyotrophic lateral sclerosis (ALS), perturbations of the metabolome that are characteristic for the disease and/or a given drug treatment can be identified. Using a commercial metabolomics platform, Lawton et al. (6) identified compounds that show statistically significant changes in ALS compared to the healthy controls, which can be considered as ALS biomarker candidates. To better understand the metabolic side effects of protease inhibitors on glucose and lipid homeostasis, Flint et al. (7) determined the endogenous metabolome of hepatocytes and adipocytes treated with atazanavir or lopinavir, and compared them to the plasma metabolome of HIV patients treated with these drugs.

Recently, quantitative targeted metabolomics using multiplexed tandem mass spectroscopy has been developed to the point where it can now be applied on a high-throughput basis. This allows the measurement of hundreds of metabolites in a fully automated manner for many samples at a time (8). Today, almost any pharmaceutical-research company or any medical laboratory can access this technology either on a fee-for-service basis or by using in-house instrumentation and prefabricated measurement kits that contain all required reactants, standards, and protocols, much like the microarray kits used in transcriptomics. This technological advance brings metabolomics to a point of sophistication requiring more advanced bioinformatic and biostatistical methods to cope with the ever-increasing amount of incoming metabolic information. Indeed, one can argue that metabolomics today is at the point where transcriptomics was five years ago. In this paper, we present a new bioinformatics approach to this challenge, addressing the analysis of a metabolomics dataset from diabetic mice under drug treatment.

Metabolic disorders such as type II diabetes are among the prime candidate diseases for which a largely improved understanding can be expected to be gained from a truly holistic metabolomics approach. One concrete question concerns the effect of insulin-sensitizing drugs on the metabolomic pathways, which are possibly affected downstream of the insulin signaling pathway itself. Rosiglitazone, marketed by GlaxoSmithKline under the name Avandia, is one such anti-diabetic drug from the thiazolidinedione class. Its mechanism of action is by activation of the intracellular receptor peroxisome proliferator-activated receptor gamma (PPARγ). PPARγ agonists such as Rosiglitazone induce the expression of the adipocyte-derived hormone adiponectin and increase plasma adiponectin levels. Adiponectin has been shown to improve insulin sensitivity and low adiponectin levels have been associated with obesity and an increased diabetes risk in both, humans and animals. Apart from the effect of Rosiglitazone on insulin resistance (9), it induces the maturation of small adipocytes (10). This leads to an increased formation of adipocytes, which in turn increases the storage potential of the adipose tissue (11). Thus, activating PPARγ by rosiglitazone leads to a redistribution of fatty acids from non-adipose tissue to adipose tissue with the consequence of a reduced availability of triglycerides in plasma (10).

Our objective here is to systematically identify relevant diabetes biomarkers and to pinpoint metabolic pathways that are affected by rosiglitazone treatment in healthy and in diabetic mice. We first discuss individual metabolite concentrations and show that known metabolic responses to diabetes and/or drug treatment are indeed recovered from the experimental data. We then analyze all possible ratios between all metabolite pairs, since, in some cases, such ratios are known to be better indicators for disease as are absolute concentrations. For example, the ratio between tyrosine and phenylalanine is a widely used biomarker for phenylketonuria (12). Finally, we introduce a more general data analysis approach based on metabolite ratios and a derived statistical parameter, the partial eta-square value, which allows for automatic identification of groups of metabolites that respond to disease or drug treatment in a correlated manner.

MATERIAL & METHODS

Experimental setup

The dataset used in this study is issued from a pre-clinical trial of a candidate anti-diabetes drug, where untreated animals and animals

treated with rosiglitazone are used as a reference. Here we only analyze the reference data obtained from a targeted quantitative metabolomics characterization of plasma samples. A set of four different animal groups (wild type/mutant diabetes; untreated/treated with rosiglitazone) with 10 animals per group (9 in the mutant-treated group) was used. The four groups will be referred to as follows: M-U (mutant untreated), W-U (wild type untreated), M-T (mutant treated with rosiglitazone) and W-T (wild type treated with rosiglitazone). As the diabetic animal model the classical male homozygous db-/db- mouse from a C57BL/6 background was used (13). The control group consisted of heterozygous male db-/db+ mice. The animals were offered water and food *ad libitum* to develop the described diabetes phenotype. Treatment was started at the age of 10 weeks. At this point, the db-/db- mice were clearly diabetic, as was confirmed by a doubled body weight and doubled plasma glucose concentrations (Fig. 1) of the db-/db- mice when compared to the wild type. Koranyi *et al.* (14) reported that db-/db- mice fed *ad libitum* are severely insulin resistant at the age of five weeks, evidenced by hyperglycemia and a doubled body weight when compared to the wild type. They further reported a number of additional clinical parameters that testify of the diabetes state of these animals, i.e. highly increased plasma and pancreatic insulin levels. These parameters were therefore not measured here (see also (15) and references therein for more physiological details on the db-/db- mouse model). The insulin-sensitizer rosiglitazone was given at conventional animal testing doses (see for example (16)). Information on the exact dosage was not available due to confidentiality issues related to the preclinical testing of the target drug. Medication was administered daily in the morning, for a period of 10 days. All animals were sacrificed in the morning to avoid influence of circadian rhythm in the animals' metabolism. All animal experimentation was conducted in accordance with accepted standards of humane animal care.

Metabolite profiling

Targeted metabolite profiling by electrospray ionization (ESI) tandem mass spectrometry (MS/MS) was performed at *Biocrates life sciences GmbH*, Austria. The technique is described in detail by patent US 2007/0004044 (accessible online at http://www.freepatentsonline.com/20070004044.html). Briefly, a targeted profiling scheme is used to quantitatively screen for known small molecule metabolites using multiple reaction monitoring, neutral loss and precursor ion scans. The quantification of the metabolites of the biological sample is achieved by reference to appropriate internal standards. The method is proven to be in conformance with 21CFR (Code of Federal Regulations) Part 11 and has been used in the past in different academic and industrial applications (2, 17).

Metabolite spectrum

Concentrations of all analyzed metabolites are reported in μM. In total, 802 different metabolites were screened and detected: 18 amino acids, 50 reducing mono-, di- and oligosaccharides (abbreviated as H*n* for *n*-hexose, dH for desoxyhexose, UA for uronic acid, HNAc for N-acetylglucosamine), 16 acylcarnitines (C*x:y*, where *x* denotes the number of carbons in the side chain and *y* the number of double bonds), five hydroxylacyl-carnitines (C(OH)*x:y*), five dicarboxylacyl-carnitines (C*x:y*-DC), free carnitine (C0) and 707 lipids. These lipids are subdivided into 82 different ceramides (Cer) and glucosyl-ceramides (GlcCer), 110 different sphingo-myelins (SM*x:y*) and sphingomyelin-derivatives, such as N- hydroxyldicarbo-acyloylsphingosyl-phosphocholine (SM(OH,COOH)*x:y*) and N- hydroxyl-acyloylsphingosyl-phosphocholine (SM (OH)*x:y*), 95 glycero-phosphatidic acids (PA), 85 glycero-phosphatidylcholines (PC), 103 glycero-phosphatidylethanolamines (PE), eleven phosphatidylglycerols (PG), 177 glycero-phosphatidylinositols (PI), glycero-phosphatidylinositol-bisphosphate (PIP2) and -triphosphate (PIP3), and 44 glycero-phosphatidylserines (PS). Glycero-phospho-lipids are further differentiated with respect to the presence of ester (*a*) and ether (*e*) bonds in the glycerol moiety, where two letters (*aa*, *ea*, or *ee*) denote that the first as well as the second position of the glycerol unit are bound to a fatty acid residue, while a single letter (*a* or *e*) indicates a bond with only one fatty acid residue. E.g. PC_ea_33:1 denotes a plasmalogen phosphatidylcholine with 33 carbons in the two fatty acid side chains and a single double bond in one of them. In some cases, the mapping of metabolite names to

individual masses can be ambiguous. For example, stereo-chemical differences are not always discernable, neither are isobaric fragments. A nomenclature of all screened-for metabolites is provided as supplementary material (Tab. S1).

Bioinformatic data analysis

All possible ratios between pairs of metabolite concentrations are computed and analyzed. Individual metabolite concentrations are naturally included in this framework by adding a virtual metabolite "1" to the list of variables. The virtual metabolite "1" concentration is set to unity for all observations. This dataset of (803^2) metabolite ratios is then analyzed using a two-way ANOVA with factors "state" (mutant diabetes/wild type; M/W) and "medication" (untreated/treated with rosiglitazone; U/T). Rather than looking at raw p-values, which would require additional correction for multiple testing, we use the partial η^2 of the ANOVA as the primary descriptor to identify interesting compounds, since η^2 is a good estimator of how much of the observed variance in the data can actually be explained by the respective factor.

To select metabolite-ratios that yield more information than single metabolites alone, the gain in η^2 ("η^2-gain") is computed for every metabolite-ratio as follows:

$$\eta^2 - gain\left(\frac{M_1}{M_2} \right) = \min \left(\frac{\eta^2\left(\frac{M_1}{M_2} \right)}{\eta^2\left(\frac{M_1}{1} \right)}, \frac{\eta^2\left(\frac{M_1}{M_2} \right)}{\eta^2\left(\frac{1}{M_2} \right)} \right)$$

where M_1 and M_2 denote any pair of metabolites. η^2-gain values that are highly above unity indicate pairs of metabolites that are candidates for being linked through closely related metabolic pathways.

In order to identify not only isolated metabolites that are affected by "state" or "medication", but also groups of metabolites that display a similar behavior and groups that are correlated with other groups, a matrix of η^2 values is generated. The rows of this matrix are labeled by the metabolites at the numerator position of the ratios and the columns are labeled by the metabolites at the denominator position. The matrix entries are the η^2 values for the metabolite pair and the selected factor. In principle, there is one matrix for every factor and every possible interaction in the regression model used in the ANOVA. Here, we only analyze the two matrices for the factors "state" and "medication".

A two-dimensional hierarchical average linkage clustering using a Euclidean distance is applied to each η^2-matrix. Alternative and more sophisticated clustering methods can also be applied in the future. The central idea introduced here is to use metabolite ratios in the ANOVA and then to cluster the resulting matrices of η^2 values. Each cluster, defined by all connected pairs of metabolites, which all have an η^2 value higher than 0.3, then defines two groups of metabolites that display similar profiles within the groups. These two groups are likely to interact through some metabolic pathway or regulatory mechanism.

RESULTS

First, we analyze single metabolite profiles and their response to disease and medication by comparing the observations with already known results. A summary of these results is given in Tables 1-3. In a second step, we analyze ratios of metabolite concentrations. We show how this approach allows us to find metabolite pairs that are likely to be linked through some (short) metabolic pathway. Here we focus on already known facts, but also indicate some new findings (Tab. 4&5). Finally, we apply a cluster analysis technique to the matrix of η^2 values, which is defined by all metabolite pairs. The resulting clusters identify groups of metabolites that exhibit similar response profiles to disease or drug treatment (Tab. 6&7). The objective and automated identification of such groups of metabolites is the methodological goal of this paper. We show that this method not only allows us to recover already known facts (from the previous steps) but also to identify new groups of potential biomarkers of disease and/or drug on-target and off-target effects.

Perturbations in amino acid and acylcarnitine concentrations conform with the diabetes phenotype

In our dataset, the diabetes phenotype reveals itself through high sugar concentrations and reduced concentrations of the glucogenic

amino acids glycine, serine and alanine in diabetic mice when compared to the wild type mice (Fig. 1). This observation may be explained by an impaired uptake of glucose by insulin-resistant cells, which induces hepatic gluconeogenesis, a process that then consumes glucogenic amino acids to initiate the production of the glucose precursors pyruvate and 3-phosphoglycerate (18). In contrast, plasma levels of the branched chain amino acids (BCAA) leucine/isoleucine and valine are increased in diabetic mice. This is consistent with earlier reports of abnormally high BCAA plasma concentrations observed in experimental diabetic (streptozotocin-induced) rats (19) and in insulin-dependent type I diabetic human patients (20). Arginine levels in the diabetic mice are found to be decreased, a fact which is already known from experimental diabetic rats (21) as well as from diabetic human patients (22). On the other hand, ornithine levels are found to be increased in diabetic mice. This observation suggests that the activity of the arginase (EC 3.5.3.1), which catalyses the reaction from arginine to ornithine is increased in diabetes (23). An alternative interpretation, an increased degradation of arginine to citrulline by the NO synthase (EC 1.14.13.39) can be excluded, since the concentrations of citrulline do not show any significant difference among the four mouse groups. The fact that the activity of ornithine decarboxylase (EC 4.1.1.17) is reduced in numerous tissues in diabetic rats (24) supports our interpretation of the increased ornithine levels (Fig. 2).

In addition to the established facts about the amino acids discussed above, the methylmalonylcarnitine (C3-DC-M) shows a significant difference with a high η^2 value for the factor "state" (Tab. 1; Fig. 3). An explanation for the increase of this metabolite could be a ketosis induced in the diabetic mice, since an increased glucose production at limited glucose utilization generally results in hyperglycemia. Simultaneously, the oxidation of an increased amount of free fatty acids provided by lipolysis facilitates gluconeogenesis and produces ketone bodies like acetoacetate (25). C3-DC-M can be converted to acetoacetate (26) and may thus be taken as an indicator of ketosis. In the past, diabetic ketosis was thought to be limited mostly to patients with type I diabetes mellitus, but recent studies support the observation of ketosis also in type II diabetes (27). Besides C3-DC-M, plasma concentrations of three other short-chained acylcarnitines, namely hydroxylpropionylcarnitine (C3(OH)), pimeloylcarnitine (C7-DC) and butenoyl-carnitine (C4:1) are also significantly increased in diabetic mice (Tab. 1; Fig. 3). To our knowledge, no straight-forward mechanistic explanation for these observations can be given. Replication in independent experiments with diabetic mice and then in human diabetic patients are in order and may constitute a promising avenue to further extend our functional understanding of the underlying disease-relevant pathways in which these metabolites are involved.

Metabolites that are impacted by drug treatment are identified

Rosiglitazone influences the lipid metabolism by activating the transcription factor PPARγ, which induces adipocyte maturation (10). This newly-formed adipocytes increase the lipid storage potential of the adipose tissue (11). Our data shows that rosiglitazone treatment significantly reduces plasma concentrations of numerous acylcarnitines, in particular those of myristoyl- (C14), palmitoyl- (C16) and stearoylcarnitines (C18) with saturated side chains, and of hexadecenylcarnitine (C16:1), octadecenylcarnitine (C18:1), and octadecadienylcarnitine (C18:2) with unsaturated side chains; and this both in diabetic and in healthy mice (Tab. 2). This observation agrees with the expected effect of rosiglitazone on reducing the concentrations of non-esterified fatty acids in individuals with type II diabetes (28), since free long-chain fatty acids are metabolized to acylcarnitines when transported into mitochondria for β-oxidation. An alternative (or additional) explanation is corroborated by the increase of free carnitine (C0) concentrations in liver tissue of the treated mouse group (unpublished data). This increase could be an indicator for an inhibitory effect of rosiglitazone on the carnitine palmitoyltransferase I (EC 2.3.1.21), which catalyzes the reaction of acyl-CoA with free carnitine to acylcarnitine. Further observations, such as reduced concentrations of glycero-phosphatidylcholines (PC) and increased plasma levels of glycero-phosphatidylinositol (PI), phosphatidyl-inositolbi- (PIP2), and triphosphates (PIP3)

(Tab. 2), can be interpreted with respect to the above-mentioned overall effect of rosiglitazone on the organism's lipid metabolism. PI, PIP2, and PIP3 additionally play an important role as second messengers, for instance in the sensitizing of cells for glucose uptake by rosiglitazon treatment (29-31).

An additional outcome of the variance analysis is the possibility to find metabolites that are oppositely affected by treatment with rosiglitazone in healthy and in diabetic mice (η^2 values for interaction terms; Tab. 3). One interesting example is methylglutarylcarnitine (C5-M-DC), where treatment with rosiglitazone increases plasma concentrations of C5-M-DC in diabetic mice while it decreases these concentrations in healthy mice (Fig. 4). Such interactions are of prime interest in drug testing. They may be indicators of potential side effects that otherwise will only become apparent during phase II clinical testing. Or contrarily, these interactions may lead to premature abandonment of drug testing in phase I if moderate side effects are observed in healthy individuals whereas such effects might not be present in diseased patients. A complete set of all significant η^2 values for all metabolite concentrations is available as supplementary material (S2).

Metabolite concentration ratios allow identification of new potential biomarkers

We now turn to the analysis of relationships between pairs of metabolites. The strength of detecting significant pairs of metabolites by using ratios between pairs of metabolite concentrations is exemplified by the ratio of the sphingomyelins SM(OH)28:0 and SM(OH)26:0 (Fig. 5). The concentrations of these compounds exhibit large variations in the metabolite concentrations within the groups, thereby masking any potential difference between the four mouse groups. In contrast, the ratios between the concentrations of those compounds show considerably lower variation within the animal groups, allowing for the identification of significant differences with respect to the factor "state". This situation can be described in an idealized manner by a steady-state approximation: in the case of metabolite pairs showing low intra-group variability in their ratio, but large inter-group variation, the equilibrium between the two metabolites would then be governed by the corresponding factor ("state" or "medication"). In this example, the plasma concentrations of the single metabolites SM(OH)28:0 and SM(OH)26:0 vary strongly in the group of untreated mutants. In contrast, the variation of the ratio SM(OH)28:0 / SM(OH)26:0 is very low in all four groups, especially in the group of untreated mutants, making the impact of the factor "state" on this ratio clearly discernable. Indeed, in this case lower values of this ratio in the diabetes groups are indicative of less long chain fatty acids and thus of a higher activity of the beta-oxidation, as can be expected in diabetes. This ratio is thus a potential biomarker candidate for a shift in the beta-oxidation activity in diabetic mice. The effect of information gain and reduction of variability through the use of concentration ratios is also impressively demonstrated by a second example, the ratio SM(OH)28:1 and SM(OH)26:0 (Fig. 6), where a similar biochemical explanation may hold.

Other pairs of metabolites for which the concentration ratios are found to be significantly different between the two goups comprise a number of glycero-phosphatidylcholines (Tab. 4). As we will discuss below, differences in these metabolites' concentrations are possibly due to a differential regulation of certain acyl-transferase enzymes. Here we only note that these metabolites were not identified in the classical analysis of single metabolite concentrations, which again illustrates that metabolite ratios may be valuable potential biomarkers, e.g. for enzyme activity or regulation.

With regard to the factor "medication", metabolite ratios with a strongly increased η^2 value were observed for a number of glycerophospholipids, in particular the glycero-phosphatidylinositols with one and two ester-bonded side chains (PI_a and PI_aa), PIP2 and PIP3, and the glycero-phosphatidylcholines PC_a, PC_aa, and PC_ee (Tab. 5). This is likely a consequence of a series of interrelated modifications of the overall lipid metabolism induced by the effect of rosiglitazone on adipocite maturation. A more detailed analysis of the underlying processes of these modifications based on the present dataset would now be possible, but is beyond the scope of the present paper. The full

set of all η^2 values for all metabolite pairs is therefore provided as supplementary material (S2) and may serve as a starting point for further investigations.

Functionally-related groups of metabolites can be identified by clustering of the η^2 matrix of the concentration ratios

Up to this point, we considered only individual concentration ratios. This confines the analysis to single metabolite pairs. However, since the analyzed factors "state" or "medication" often affect larger groups of metabolites in a similar manner, we aspire to find such groups that exhibit similar behavior (within the groups as well as between them) by using a clustering approach (see methods). The clustered η^2 matrix for the factor "state" allows identifying at least 13 groups of metabolites with similar profiles (Tab. 6). Some of these clusters consist of groups of metabolites that have already been discussed above. For instance, group 4 and group 6 relate the glycogenic amino acids serine and alanine to a set of sugar variables. Group 7 involves the branched chain amino acids, but also ornithine, which, as discussed above, is linked to arginine in group 5. In the clustering matrix for the factor "medication", four corresponding groups of metabolites could be identified (Tab. 7). Two groups (group 14 & group 15) mainly consist of glycero-phosphatidylinositols (PI, PIP2, PIP3), while two other groups (group 16 & group 17) mostly contain long-chain acylcarnitines and glycero-phosphocholines (PC_aa, PC_ea, PC_ee). In mice treated with rosiglitazone, the concentrations of the long-chain acylcarnitines and the glycero-phosphocholines are decreased, whereas the glycero-phosphatidylinositol levels are increased. The rediscovery of these already known metabolic interactions through a quasi-objective and automated clustering method confirms the success of this approach and shows the available potential. In particular, it is now possible to use the reported groups and the relationships between them as a starting point for further in-depth analysis, knowing that the impact of the factors "state" (diabetes) and "medication" (treatment with rosiglitazone) is statistically significant by construction in all cases.

A cluster that provides presumably new information on diabetes is constituted by the groups 1 and 2. Group 1 consists of the phosphatidylcholines PC_aa, PC_ea and PC_ee, whereas group 2 comprises mostly the lysophosphatidylcholines PC_e and PC_a. Some of the metabolites in these two groups have already been identified above as being relevant to diabetes, based on concentration ratios (see above; Tab. 4). In diabetic mice, the plasma concentrations of metabolites from group 1 are increased, whereas the plasma concentrations of metabolites from group 2 are decreased with respect to the wild type. Most metabolite ratios in this cluster have an η^2-gain greater than two, indicating that the metabolites from both groups are connected by some direct metabolic pathways. Indeed, the enzymes alkylglycerophosphocholine O-acyltransferase (EC 2.3.1.63) (32, 33) and 1-acylglycerophosphocholine O-acyltransferase (EC 2.3.1.23) (34, 35) are known to catalyze the conversion of PC_e and PC_a to PC_ea and PC_aa (Fig. 7). Thus, our observations suggest that, in diabetic mice, the regulation of these reactions is modified in a way that the equilibrium is shifted to the PC_aa/PC_ea/PC_ee side. The precise functional background of this observation requires further investigation, which may lead to new insight into the metabolic pathways that are perturbed in diabetic patients. This example highlights how an objective bioinformatics analysis of high-throughput targeted metabolomics experiments can lead to the formulation of new and testable hypotheses.

DISCUSSION

In this paper, we have presented a bioinformatics analysis of what can be considered as a standard experimental setting of a preclinical drug testing experiment with two independent factors, "state" and "medication". Targeted quantitative metabolomics, covering a wide range of more than 800 relevant metabolites, measured in a reproducible manner with generally low in-group variability, provides an excellent test bed for automated and objective analysis methods.

Our first objective here was to recover already known metabolic effects of diabetes and of the treatment with the anti-diabetes drug

rosiglitazone. This objective has been reached in several instances. The analysis of single metabolites recovered known pathways affected by diabetes, namely gluconeogenesis, the impaired uptake of branched-chain amino acids (BCAA) by muscle tissue, the diminished activity of ornithine decarboxylase, the increased activity of the arginase and the indication of an increase in ketone bodies. The analysis of single metabolites with respect to the factor "medication" shows effects that can be explained by the impact of rosiglitazone on the non-esterified fatty acid metabolism, followed by a decrease in long-chain acylcarnitine concentrations.

Our second objective was to identify new compounds and pathways that may be related to diabetes. By submitting all possible ratios between metabolite pairs to a two-factor variance analysis (ANOVA), and then focusing specifically on metabolite pairs with a high η^2-gain, we identify relations between compounds that were not apparent in the initial analysis of single metabolites. For instance, in diabetic mice the beta-oxidation appears to be significantly increased, as can be deduced from the concentration shift of sphingomyelins SM(OH)28:0 and SM(OH)28:1 to SM(OH)26:0. This makes the ratio of this metabolic pair a potential biomarker for diabetes. With respect to the factor "state", newly identified metabolites that are affected by rosiglitazone treatment, but which were not apparent in the analysis of single metabolites, are mainly groups of glycero-phosphocholines and a few sphingomyelins. These results suggest an impact of medication on some PIP2, PIP3, PI and PC compounds. Few of the latter molecules were already apparent in the analysis of single metabolites.

An inherent property of a multifactor ANOVA is the possibility to identify metabolites that are affected by interactions between the factors. Here, methylglutarylcarnitine (C5-M-DC) has been identified as a prime example, where treatment with rosiglitazone has an increasing effect on C5-M-DC levels in diabetic mice, but a decreasing effect in wild type mice. In general, such differently affected metabolites can be indicators for undesired side effects of medication. These metabolites thus become valuable early indicators of potential issues with a new drug that, otherwise, may only be identified late in clinical drug testing, since the effect of the drug on healthy candidates in phase I would be opposite to what would be seen in phase II testing of affected patients. Our approach presents a straightforward way to spot such potentially problematic metabolites.

In a more exploratory vein, we sought to identify groups of potentially interrelated metabolites that could then be used to enhance our understanding of principal metabolic mechanisms underlying the disease or the medication under question. To achieve this goal, we applied a standard hierarchical clustering method to the matrices that are constituted by the different η^2-values from the ANOVA of the metabolite ratios (see methods). As a first verification of this approach, we identified groups of metabolites that were already known from the previous analysis of single metabolites and metabolite pairs, e.g. the amino acids and sugars, which are indicative of an increased gluconeogenesis in diabetes, and are combined in a single cluster. In addition, new groups of metabolites could be detected, including different groups of glycero-phosphocholines related to the diabetes "state". The discovery of these new groups generalizes the results from the analysis of ratios and, therefore, makes it possible to describe the behavior of entire classes of metabolites. This approach can lead to a better understanding of the underlying metabolic processes in case of disease or of drug treatment. One example of such a process that should now be further investigated depicts the impact of diabetes on the regulation of the metabolic pathways that involve the enzymes 1-alkylglycerophosphocholine O-acyltransferase (EC 2.3.1.63) and 1-acylglycerophosphocholine O-acyltransferase (EC 2.3.1.23) in the conversion of PC_e and PC_a to PC_ea and PC_aa, respectively (Fig. 7).

Future efforts will concentrate on testing and developing these analysis methods for use in diabetic human studies. The design of such clinical studies will be challenged by the variability resulting from genetic differences as well as variations in nutritional and sampling conditions when compared to pre-clinical studies on mouse models.

REFERENCES

1. **Rolinski B, Arnecke R, Dame T, Kreischer J, Olgemoller B, Wolf E, Balling R, Hrabe de Angelis M, Roscher AA** 2000 The biochemical metabolite screen in the Munich ENU Mouse Mutagenesis Project: determination of amino acids and acylcarnitines by tandem mass spectrometry. Mamm Genome 11:547-551
2. **Butler A, Schneider R, Floyd E, Roffi C, Johnson K, Weinberger K, Graber A, Delnomdedieu M** 2006 Metabonomic Profiling to identify Biomarkers of vascular Injury in Canines and Humans. Cambridge Healthtech Institute's 7th Annual, Metabolic Markers for Drug Development and Clinical Studies, Orlando, USA, 4-5 December 2006
3. **Paige LA, Mitchell MW, Krishnan KR, Kaddurah-Daouk R, Steffens DC** 2007 A preliminary metabolomic analysis of older adults with and without depression. Int J Geriatr Psychiatry 22:418-423
4. **Kaddurah-Daouk R, McEvoy J, Baillie RA, Lee D, Yao JK, Doraiswamy PM, Krishnan KRR** 2007 Metabolomic mapping of atypical antipsychotic effects in schizophrenia. Mol Psychiatry 12:934–945
5. **Rozen S, Cudkowicz ME, Bogdanov M, Matson WR, Kristal BS, Beecher C, Harrison S, Vouros P, Flarakos J, Vigneau-Callahan K, Matson TD, Newhall KM, Beal MF, Brown RH, Kaddurah-Daouk R** 2005 Metabolomic analysis and signatures in motor neuron disease. Metabolomics 1:101-108
6. **Lawton KA, Paige LA, Mitchell MW, Kaddurah-Daouk R, Brown RH, Milburn M, Cudkowicz ME** 2006 Identification of Metabolic Biomarkers for Amyotrophic Lateral Sclerosis. The Second Scientific Meeting of the Metabolomics Society, Harvard Medical School - The Conference Center , Boston, USA, 24-29 June 2006
7. **Flint OP, Elosua C, Parker RA, Hamilton CM, Hamilton MA** 2006 HIV Protease Inhibitors and the Hepatic Metabolome. 13th Conference on Retroviruses and Opportunistic Infections, Denver, Colorado, 2006
8. **Ramsay SL, Stoeggl WM, Weinberger K, Graber A, Guggenbichler W** 2007 Apparatus and method for analyzing a metabolite profile. In: United States Patent US 20070004044
9. **Hernandez R, Teruel T, Lorenzo M** 2003 Rosiglitazone produces insulin sensitisation by increasing expression of the insulin receptor and its tyrosine kinase activity in brown adipocytes. Diabetologia 46:1618-1628
10. **Okuno A, Tamemoto H, Tobe K, Ueki K, Mori Y, Iwamoto K, Umesono K, Akanuma Y, Fujiwara T, Horikoshi H, Yazaki Y, Kadowaki T** 1998 Troglitazone increases the number of small adipocytes without the change of white adipose tissue mass in obese Zucker rats. J Clin Invest 101:1354-1361
11. **de Souza CJ, Eckhardt M, Gagen K, Dong M, Chen W, Laurent D, Burkey BF** 2001 Effects of pioglitazone on adipose tissue remodeling within the setting of obesity and insulin resistance. Diabetes 50:1863-1871
12. **Chace DH, Sherwin JE, Hillman SL, Lorey F, Cunningham GC** 1998 Use of phenylalanine-to-tyrosine ratio determined by tandem mass spectrometry to improve newborn screening for phenylketonuria of early discharge specimens collected in the first 24 hours. Clin Chem 44:2405-2409
13. **Hummel KP, Dickie MM, Coleman DL** 1966 Diabetes, a new mutation in the mouse. Science 153:1127-1128
14. **Koranyi L, James D, Mueckler M, Permutt MA** 1990 Glucose Transporter Levels in Spontaneously Obese (db/db) Insulin-resistant Mice. J Clin Invest 85:962-967
15. **Kumar S, McCue P, Dunn SR** 2003 Diabetic kidney disease in the db/db mouse. Am J Physiol Renal Physiol 284:F1138–F1144
16. **Chaput E, Saladin R, Silvestre M, Edgar AD** 2000 Fenofibrate and Rosiglitazone Lower Serum Triglycerides with Opposing Effects on Body Weight. Biochemical and Biophysical Research Communications 271:445–450
17. **Bernardo K, Stoeggl W, Marini G, Zaccaria P, Weinberger K, Graber A, Ramsay SL** 2006 A Fully Automated Sample Preparation and Electrospray Ionisation Tandem Mass Spectrometry Based Approach to Identify and Quantify Lipids - An Application to Monitor

Lipid Changes in Cardiovascular Disease Patients. 17th International Mass Spectrometry Conference, Prague, Czech Republic, 25 Aug - 1 Sep 2006
18. **Magnusson I, Rothman DL, Katz LD, Shulman RG, Shulman GI** 1992 Increased rate of gluconeogenesis in type II diabetes mellitus. A 13C nuclear magnetic resonance study. J Clin Invest 90:1323-1327
19. **Glanville NT, Anderson GH** 1985 The effect of insulin deficiency, dietary protein intake, and plasma amino acid concentrations on brain amino acid levels in rats. Can J Physiol Pharmacol 63:487-494
20. **Borghi L, Lugari R, Montanari A, Dall'Argine P, Elia GF, Nicolotti V, Simoni I, Parmeggiani A, Novarini A, Gnudi A** 1985 Plasma and skeletal muscle free amino acids in type I, insulin-treated diabetic subjects. Diabetes 34:812-815
21. **Mans AM, DeJoseph MR, Davis DW, Hawkins RA** 1987 Regional amino acid transport into brain during diabetes: effect of plasma amino acids. Am J Physiol 253:E575-583
22. **Grill V, Bjorkman O, Gutniak M, Lindqvist M** 1992 Brain uptake and release of amino acids in nondiabetic and insulin-dependent diabetic subjects: important role of glutamine release for nitrogen balance. Metabolism 41:28-32
23. **Kashyap SR, Lara A, Zhang R, Park YM, Defronzo RA** 2007 Insulin Reduces Plasma Arginase Activity in Type 2 Diabetes Patients. Diabetes Care published online ahead of print October 10, 2007
24. **Conover CA, Rozovski SJ, Belur ER, Aoki TT, Ruderman NB** 1980 Ornithine decarboxylase activity in insulin-deficient states. Biochem J 192:725-732
25. **Wolfsdorf J, Glaser N, Sperling MA** 2006 Diabetic ketoacidosis in infants, children, and adolescents: A consensus statement from the American Diabetes Association. Diabetes Care 29:1150-1159
26. **Zhang W, Reynolds KA** 2001 MeaA, a putative coenzyme B12-dependent mutase, provides methylmalonyl coenzyme A for monensin biosynthesis in Streptomyces cinnamonensis. J Bacteriol 183:2071-2080
27. **Newton CA, Raskin P** 2004 Diabetic ketoacidosis in type 1 and type 2 diabetes mellitus: clinical and biochemical differences. Arch Intern Med 164:1925-1931
28. **Raskin P, Rappaport EB, Cole ST, Yan Y, Patwardhan R, Freed MI** 2000 Rosiglitazone short-term monotherapy lowers fasting and post-prandial glucose in patients with type II diabetes. Diabetologia 43:278-284
29. **Bevan P** 2001 Insulin signalling. J Cell Sci 114:1429-1430
30. **Iwata M, Haruta T, Usui I, Takata Y, Takano A, Uno T, Kawahara J, Ueno E, Sasaoka T, Ishibashi O, Kobayashi M** 2001 Pioglitazone ameliorates tumor necrosis factor-alpha-induced insulin resistance by a mechanism independent of adipogenic activity of peroxisome proliferator--activated receptor-gamma. Diabetes 50:1083-1092
31. **Rieusset J, Auwerx J, Vidal H** 1999 Regulation of gene expression by activation of the peroxisome proliferator-activated receptor gamma with rosiglitazone (BRL 49653) in human adipocytes. Biochem Biophys Res Commun 265:265-271
32. **Waku K, Nakazawa Y** 1970 Acyltransferase activity to 1-O-alkyl-glycero-3-phosphorylcholine in sarcoplasmic reticulum. J Biochem (Tokyo) 68:459-466
33. **Waku K, Nakazawa Y** 1972 Acyltransferase activity to 1-acyl-, 1-O-alkenyl-, and 1-O-alkyl-glycero-3-phosphorylcholine in Ehrlich ascites tumor cells. J Biochem (Tokyo) 72:495-497
34. **Bell RM, Coleman RA** 1980 Enzymes of glycerolipid synthesis in eukaryotes. Annu Rev Biochem 49:459-487
35. **Miki Y, Hosaka K, Yamashita S, Handa H, Numa S** 1977 Acyl-acceptor specificities of 1-acylglycerolphosphate acyltransferase and 1-acylglycerophosphorylcholine acyltransferase resolved from rat liver microsomes. Eur J Biochem 81:433-441

FIGURES LEGENDS

Figure 1: Gluconeogenesis: box plots of plasma concentrations [μM] of Hexose (H1) and of the glycogenic amino acids glycine, serine, and alanine in mutant (M) and wild type (W) mice, untreated (U) and treated with rosiglitazone (T). Box plots of body weight (bottom right) of the db-/db- mice show that these mice are highly obese when compared to the wild type.

Figure 2: The ureacycle: plasma concentrations [μM] of arginine and citrulline and of the non proteinogenic amino acid ornithine.

Figure 3: Plasma concentrations [μM] of methylmalonylcarnitine (C3-DC-M), hydroxylpropionylcarnitine (C3(OH)), pimeloylcarnitine (C7-DC) and butenoylcarnitine (C4:1).

Figure 4: Plasma concentrations [μM] of methylglutarylcarnitine (C5-M-DC).

Figure 5: Plasma concentrations [μM] of N-hydroxyacyloylsphingosyl-phosphocholine (SM(OH)28:0) and (SM(OH)26:0) – top row – and the ratio of both concentrations – bottom row.

Figure 6: Plasma concentrations [μM] of N-hydroxyacyloylsphingosyl-phosphocholine (SM(OH)28:1) and (SM(OH)26:0) – top row – and the ratio of both concentrations – bottom row.

Figure 7: Conversion of PC_e to PC_ea and of PC_a to PC_aa by the enzymes 1-acylglycerophosphocholine O-acyltransferase (EC 2.3.1.23) and 1-alkylglycerophosphocholine O-acyltransferase (EC 2.3.1.63).

Figure 8: Example of clustering by η^2 (M/W). The full matrix for both factors is available in a zoomable format as supplementary material (Fig. S3/S4).

TABLES 1-7

Metabolite	Abbreviation	η^2 (M/W)	η^2 (U/T)	η^2 (M/W*U/T)
hexose	H1	0,644	0,012	0,050
glycine	Gly	0,625	0,036	0,033
serine	Ser	0,214	0,045	0,023
alanine	Ala	0,187	0,220	0,077
leucine/ isoleucine	Leu/Ile	0,626	0,164	0,001
valine	Val	0,607	0,106	0,005
phenylalanine	Phe	0,735	0,084	0,001
ornithine	Orn	0,518	0,001	0,062
arginine	Arg	0,341	0,390	0,006
methylmalonylcarnitine	C3-DC-M	0,693	0,032	0,102
hydroxylpropionylcarnitine	C3(OH)	0,540	0,011	<0,001 .
pimeloylcarnitine	C7-DC	0,654	0,004	0,050
butenoylcarnitine	C4:1	0,558	0,003	0,037

Table 1: Partial η^2 values for the two factors "state" (mutant or wild type; M/W) and "medication" (untreated or treated with rosiglitazone; U/T) and for the interaction of both factors (M/W * U/T). Shown are selected metabolites that display a strong difference in plasma concentrations between wild type and diabetic mice. The complete table of all η^2 values for all measured metabolites is provided as supplementary material (Tab. S2).

Metabolite	Abbreviation	η^2 (M/W)	η^2 (U/T)	η^2 (M/W*U/T)
octadecenylcarnitine	C18:1	0,041	0,548	0,002
stearoylcarnitine	C18:0	0,047	0,511	0,047
octadecadienylcarnitine	C18:2	0,003	0,475	<0,001 .
palmitoylcarnitine	C16:0	0,017	0,461	0,013
myristoylcarnitine	C14:0	0,010	0,447	0,008
tetradecenylcarnitine	C14:1	0,298	0,315	0,039
hexadecenylcarnitine	C16:1	0,172	0,292	0,018
diacylphosphatidylinositol triphosphate	PIP3_aa_C42:3	0,179	0,439	0,129
diacylphosphatidylinositol bisphosphate	PIP2_aa_C36:0	0,068	0,338	0,078
dialkylphosphatidyl-choline	PC_ee_C42:6	0,093	0,381	0,089
diacylphosphatidylcholine	PC_aa_C40:5	0,164	0,352	0,122
diacylphosphatidylcholine	PC_aa_C40:6	0,070	0,350	0,060
dialkylphosphatidyl-choline	PC_ee_C42:5	0,176	0,334	0,117

Table 2: As table 1; shown here are selected metabolites that display a strong difference in plasma concentrations between untreated mice and mice treated with rosiglitazone.

Metabolite	Abbreviation	η^2 (M/W)	η^2 (U/T)	η^2 (M/W*U/T)
diacylphosphatidylinositol bisphosphate	PIP2_aa_C24:1	0,063	0,005	0,274
alkylacylphosphatidylcholine	PC_ea_C20:4	<0.001 .	0,073	0,259
methylglutarylcarnitine	C5_M_DC	0,004	0,068	0,232

Table 3: As table 1; shown here are metabolites that display a strong interaction between both factors, but only a small contribution of one factor alone.

metabolite in numerator	metabolite in denominator	η^2 M/W	η^2 U/T	η^2 M/W * U/T	η^2 gain M/W	η^2 gain U/T	η^2 gain M/W * U/T
SM_(OH)_28:0	SM_(OH)_26:0	0,791	0,176	0,007	3,8	75,6	0,1
SM_(OH)_28:1	SM_(OH)_26:0	0,729	0,072	0,067	4,5	13,3	0,5
SM_C26:4	SM_(OH)_28:0	0,736	0,481	0,296	6,2	7,4	3,2
PC_ea_C38:8	PC_e_C32:1	0,736	0,220	0,072	2,6	0,9	0,9
PC_aa_C36:1	PC_e_C32:1	0,733	0,189	0,052	2,7	0,9	0,8
PC_ee_C38:1	PC_e_C32:1	0,727	0,153	0,031	2,9	0,8	0,6
PC_ee_C38:2	PC_e_C32:1	0,722	0,276	0,121	2,8	1,1	1,4
PC_aa_C36:2	PC_e_C32:1	0,722	0,263	0,114	2,7	1,1	1,3
PC_ee_C40:3	PC_e_C32:1	0,706	0,238	0,102	2,4	1,1	1,0
PC_a_C18:2	PC_a_C22:6	0,706	0,544	0,352	8,5	8,8	41,1
PC_a_C22:6	PC_a_C18:2	0,703	0,410	0,026	49,5	11,6	2,7
PA_a_C26:6	SM_C22:3	0,698	0,611	0,789	10,4	18,9	10,5
UA_HNS_UA_HNAc_UA	PS_aa_C38:1	0,736	0,178	0,192	4,1	6,7	4,7

Table 4: Selected pairs of metabolites that show a strong gain in partial η^2 values for *M/W,* when considering ratios between concentrations. The variability of the concentration of each of these compounds cannot be explained by the respective factor(s) alone, whereas their ratios clearly can. The complete table of all η^2 values for all measured metabolite concentration ratios is available as supplementary material (Tab. S2).

metabolite in numerator	metabolite in denominator	η^2 M/W	η^2 U/T	η^2 M/W * U/T	η^2 gain M/W	η^2 gain U/T	η^2 gain M/W * U/T
PG_aa_C44:7	PI_aa_C40:3	0,191	0,822	0,020	0,5	2,1	0,1
PIP2_aa_C36:0	UA_HNAc	0,008	0,774	0,008	0,1	2,2	0,1
PI_a_C18:0	PG_aa_C44:7	0,184	0,715	0,227	1,4	5,5	3,5
UA_HNAc	PIP3_aa_C20:1	0,269	0,706	0,452	1,8	3,0	11,9
PIP2_aa_C36:0	PE_a_C30:3	0,037	0,701	0,212	0,5	2,0	2,7
PC_a_C18:1	PC_aa_C40:6	0,135	0,683	0,356	0,8	2,3	8,3
UA_HNAc	PS_aa_C36:1	0,447	0,660	0,253	1,8	4,0	39,9
PC_a_C16:0	PC_aa_C40:6	0,008	0,653	0,206	0,3	2,2	4,8
PC_a_C16:0	PC_ee_C40:4	0,057	0,641	0,15	2,2	2,2	2,2
PC_a_C16:0	PC_aa_C38:4	0,058	0,638	0,149	2,2	2,1	1,9
PC_a_C18:1	PC_ee_C40:4	0,115	0,615	0,240	0,6	2,1	3,5
PC_a_C18:1	PC_aa_C38:4	0,114	0,612	0,237	0,6	2,0	3,0
LacCer_C18:0	SM_(OH)_26:3	0,347	0,612	0,583	1,5	2,1	4,5
PI_aa_C40:3	PE_aa_C40:5	0,029	0,612	0,030	0,1	2,6	0,3
PA_a_C26:6	SM_C22:3	0,698	0,611	0,789	10,4	18,8	10,5
UA_HNAc	PIP2_aa_C28:0	0,517	0,604	0,459	4,3	3,7	60,8

Table 5: As Tab. 5, but for U/T.

group 1	group 2	group 5	group 8	group 9
PC_aa_C34:0	PC_a_C20:4	Alanine	C0	C4_OH
PC_aa_C34:1	PC_a_C22:6	Arginine	C2:0	C8:0
PC_aa_C34:2	PC_e_C20:0	Serine	C3	C8_DC
PC_aa_C36:1	PC_e_C22:1		C5_M_DC	C10:1
PC_aa_C36:2	PC_e_C26:0		C6	C14:0
PC_aa_C36:3	PC_e_C30:0	**group 6**	Citrulline	C16:0
PC_aa_C38:3	PC_e_C30:1	C3_DC_M	Lysine	C18:0
PC_aa_C38:5	PC_e_C32:0	C7_DC	Tryptophan	C18:2
PC_aa_C40:5	PC_e_C32:1	H1	Tyrosine	Aspartic_Acid
PC_ea_C36:7		H3_HNAc2_NANA	PC_ea_C30:7	Glutamic_Acid
PC_ea_C36:8	**group3**	HNAc_H2_dH	PC_ee_C30:0	Histidine
PC_ea_C38:8	SM_C28_2	PC_aa_C36:5	PIP3_aa_C20:7	Methionine
PC_ee_C36:0	SM_OH_COOH_C24:0	PC_ee_C38:5		Proline
PC_ee_C36:1	SM_OH_COOH_C24:1			Threonine
PC_ee_C36:2	SM_OH_COOH_C24:4			
PC_ee_C36:3		**group 7**		
PC_ee_C38:1	**group 4**	C3_OH		
PC_ee_C38:2	Alanine	C4:1		
PC_ee_C38:3	Arginine	Leucine/Isoleucine		
PC_ee_C40:3	Serine	Ornithine		
PC_ee_C40:5	UA_HNS_UA_HNAc_UA	Phenylalanine		
PC_ee_C42:5	Pentose	Valine		

(6a)

group 10	group 11	group 12	group 13
PC_aa_C32:0	GlcCer_OH_COOH_C8:1	GlcCer_OH_COOH_C8:1	PC_aa_C34:1
PC_aa_C38:2	GlcCer_OH_COOH_C20:2	GlcCer_OH_COOH_C20:2	PC_aa_C36:1
PC_aa_C38:6	GlcCer_C24:0	PC_a_C16:0	PC_aa_C36:2
PC_ea_C32:8	HNAc_H4_dH2	PC_a_C18:0	PC_aa_C36:3
PC_ea_C34:7	PC_a_C16:0	SM_OH_28:0	PC_aa_C38:3
PC_ee_C32:1	PC_a_C18:0	SM_OH_28:1	PC_aa_C38:5
PC_ee_C34:0	PI_aa_C18:1		PC_ea_C36:8
PC_ee_C40:2	PI_aa_C20:6		PC_ea_C38:8
PC_ee_C40:6	PIP2_aa_C36:1		PC_ee_C36:1
			PC_ee_C38:1
			PC_ee_C38:2
			PC_ee_C38:3
			PC_ee_C40:3
			PC_ee_C40:5

(6b)

	group 1	group 2	group 3	group 4	group 5	group 6	group 7	group 8	group 9	group 10	group 11	group 12	group 13
group 1													
group 2	x												
group 3	x												
group 4													
group 5													
group 6		x	x	x	x								
group 7					x								
group 8						x	x						
group 9						x	x						
group 10						x							
group 11						x							
group 12	x												
group 13		x	x							x		x	

(6c)

Table 6: (a), (b) Groups of metabolites that are found to cluster with high η^2 values for the factor "state" (M/W); (c) groups that cluster together.

group 14	group 15	group 16	group 17
PI_aa_C20:7	C4:1	Arginine	PC_aa_C18:7
PI_aa_C32:1	C6_OH	C14:0	PC_aa_C36:1
PI_aa_C40:4	Ornithine	C14:1	PC_aa_C36:2
PIP2_aa_C26:6	PC_e_C22:1	C16:0	PC_aa_C38:3
PIP2_aa_C26:8	PC_e_C26:0	C16:1	PC_aa_C38:4
PIP2_aa_C36:1	PC_e_C30:0	C18:0	PC_aa_C38:6
PIP3_aa_C10:2	PC_e_C30:1	C18:1	PC_ea_C18:0
PIP3_aa_C20:4	PI_a_C20:0	C18:2	PC_ea_C38:8
PIP3_aa_C36:4	PI_aa_C38:2	PC_aa_C40:5	PC_ee_C20:7
PS_a_C26:0	PI_aa_C40:3	PC_aa_C40:6	PC_ee_C38:1
PIP3_aa_C18:8	PI_e_C22:0	PC_ee_C42:5	PC_ee_C38:2
PIP3_aa_C20:2	PIP2_aa_C16:4	PC_ee_C42:6	PC_ee_C40:3
PC_a_C18:0	PIP2_aa_C26:0		PC_ee_C40:4
PC_e_C32:0	PIP2_aa_C28:0		
SM_C20:2	PIP3_aa_C20:0		
SM_C26:3	PIP3_aa_C20:6		
SM_C30:4	PIP3_aa_C20:7		
SM_OH_28:0	PIP3_aa_C42:3		
SM_OH_COOH_16:3			
SM_OH_COOH_26:0			

(7a)

	group 14	group 15	group 16	group 17
group 14				
group 15				
group 16	X	X		
group 17	X			

(7b)

Table 7: (a) Groups of metabolites that clusters for the factor "medication" (U/T); (b) groups that cluster together.

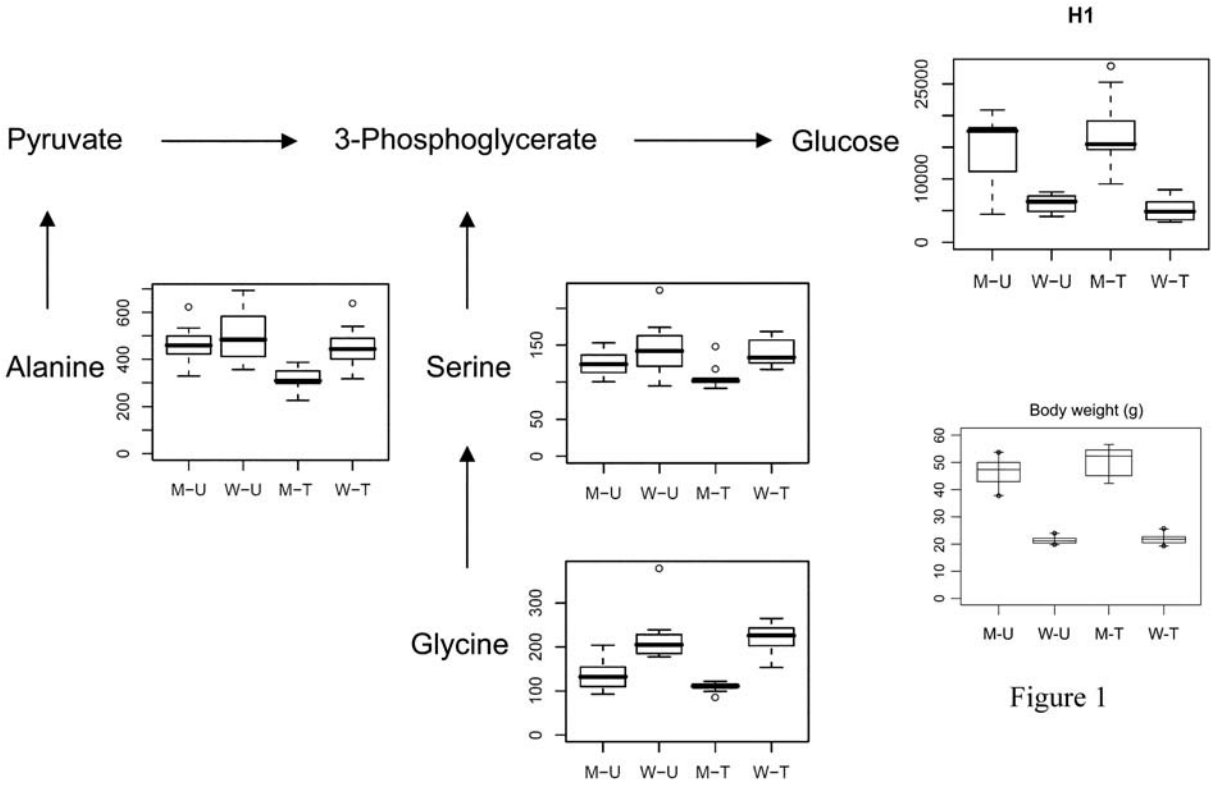

Pyruvate
3-Phosphoglycerate
Glucose
Alanine
Serine
Glycine
H1
Body weight (g)
M−U
W−U
M−T
W−T

Figure 1

Figure 2

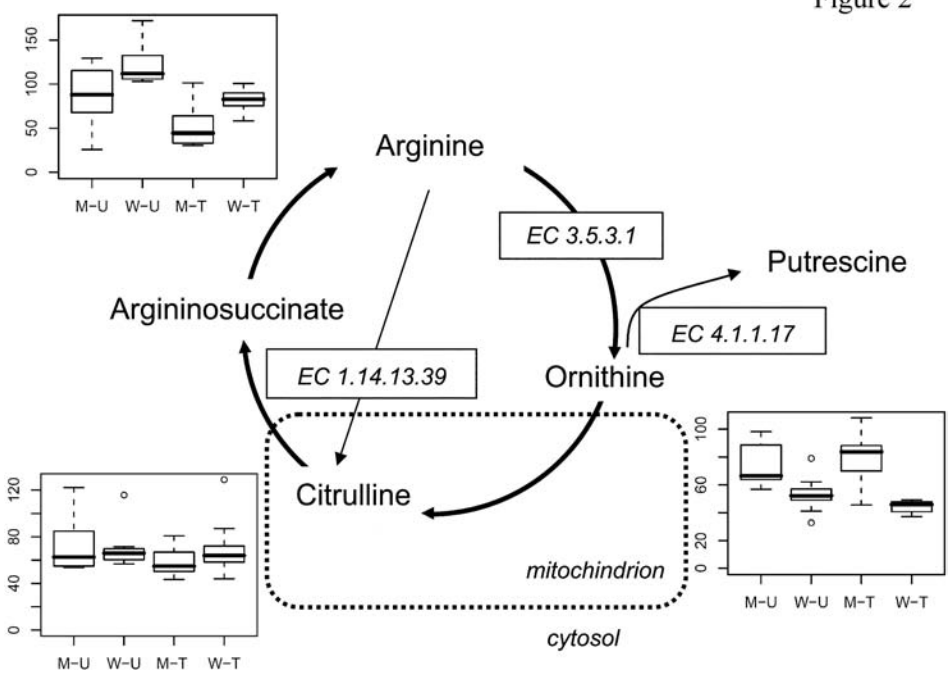

Arginine
Argininosuccinate
Citrulline
Ornithine
Putrescine
EC 3.5.3.1
EC 4.1.1.17
EC 1.14.13.39
mitochindrion
cytosol
M−U
W−U
M−T
W−T

Figure 3

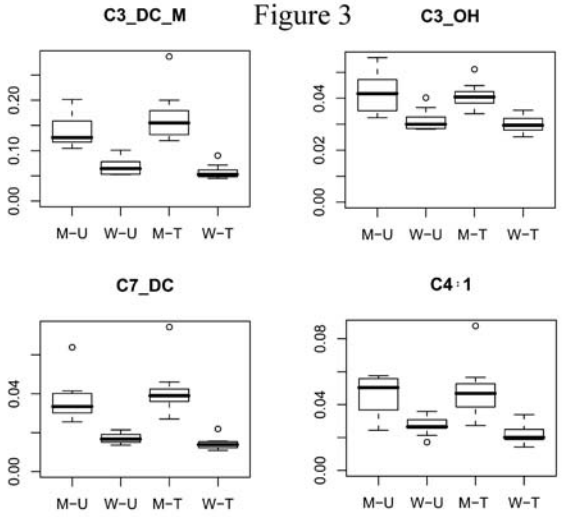
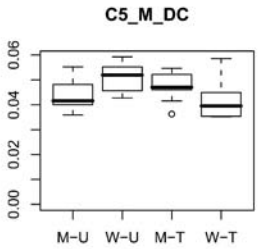

C3_DC_M
0.00
0.10
0.20
M-U
W-U
M-T
W-T
C3_OH
0.00
0.02
0.04
M-U
W-U
M-T
W-T
C7_DC
0.00
0.04
M-U
W-U
M-T
W-T
C4:1
0.00
0.04
0.08
M-U
W-U
M-T
W-T

Figure 4



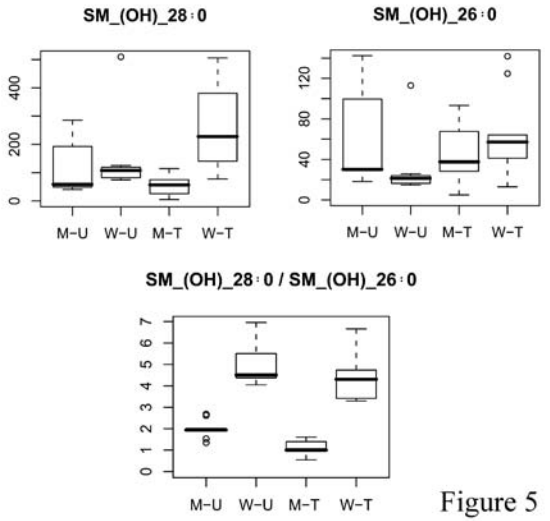

SM_(OH)_28:0
0
200
400
M-U
W-U
M-T
W-T
SM_(OH)_26:0
0
40
80
120
M-U
W-U
M-T
W-T
SM_(OH)_28:0 / SM_(OH)_26:0
0 1 2 3 4 5 6 7
M-U
W-U
M-T
W-T

Figure 5

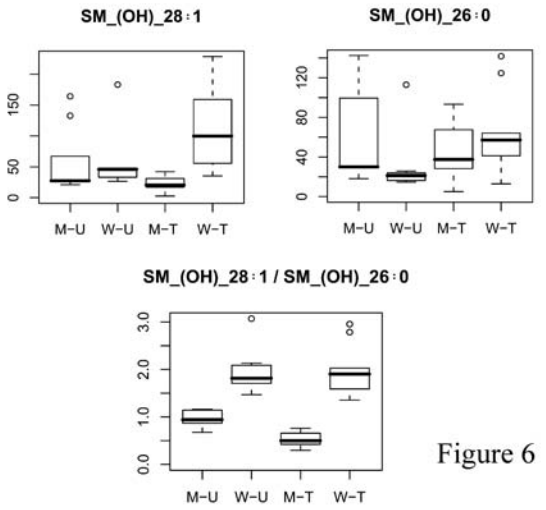

SM_(OH)_28:1
0
50
150
M-U
W-U
M-T
W-T
SM_(OH)_26:0
0
40
80
120
M-U
W-U
M-T
W-T
SM_(OH)_28:1 / SM_(OH)_26:0
0.0
1.0
2.0
3.0
M-U
W-U
M-T
W-T

Figure 6

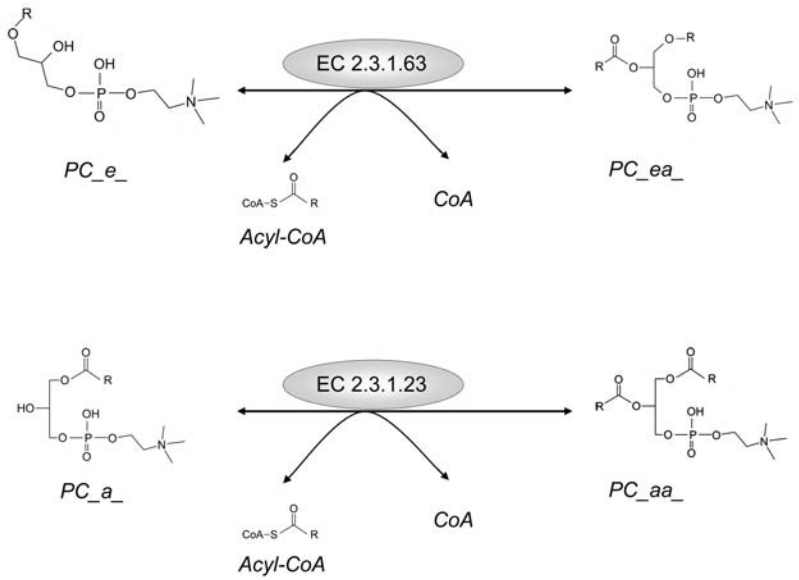


Figure 7

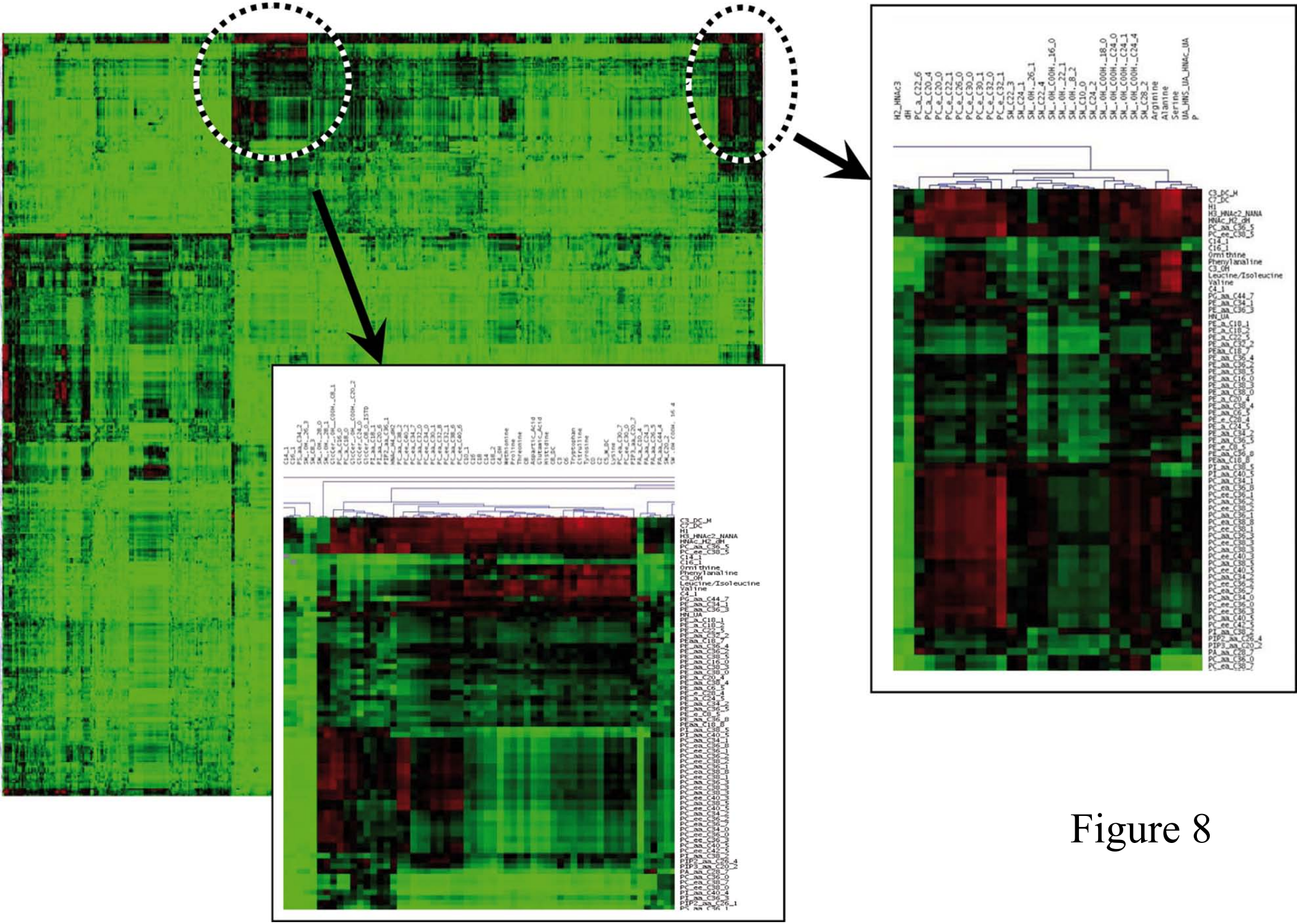


Figure 8