Discovery of early-stage biomarkers for diabetic kidney disease using ms-based metabolomics (FinnDiane study)

Kloet, F.M. van der; Tempels, F.W.A.; Ismail, N.; Heijden, R. van der; Kasper, P.T.; Rojas-Cherto, M.; ... ; Hankemeier, T.

ORIGINAL ARTICLE

Discovery of early-stage biomarkers for diabetic kidney disease using ms-based metabolomics (FinnDiane study)

F. M. van der Kloet · F. W. A. Tempels · N. Ismail · R. van der Heijden ·
P. T. Kasper · M. Rojas-Cherto · R. van Doorn · G. Spijksma · M. Koek ·
J. van der Greef · V. P. Mäkinen · C. Forsblom · H. Holthöfer · P. H. Groop ·
T. H. Reijmers · T. Hankemeier

Abstract Diabetic kidney disease (DKD) is a devastating complication that affects an estimated third of patients with type 1 diabetes mellitus (DM). There is no cure once the disease is diagnosed, but early treatment at a sub-clinical stage can prevent or at least halt the progression. DKD is clinically diagnosed as abnormally high urinary albumin excretion rate (AER). We hypothesize that subtle changes in the urine metabolome precede the clinically significant rise in AER. To test this, 52 type 1 diabetic patients were recruited by the FinnDiane study that had normal AER (normoalbuminuric). After an average of 5.5 years of follow-up half of the subjects (26) progressed from normal AER to microalbuminuria or DKD (macroalbuminuria), the other half remained normoalbuminuric. The objective of this study is to discover urinary biomarkers that differentiate the progressive form of albuminuria from non-progressive form of albuminuria in humans. Metabolite profiles of baseline 24 h urine samples were obtained by gas chromatography–mass spectrometry (GC–MS) and liquid chromatography–mass spectrometry (LC–MS) to detect potential early indicators of pathological changes. Multivariate logistic regression modeling of the metabolomics data resulted in a profile of metabolites that separated those patients that progressed from normoalbuminuric AER to microalbuminuric AER from those patients that maintained normoalbuminuric AER with an accuracy of 75% and a precision of 73%. As this data and samples are from an actual patient population and as such, gathered within a less controlled environment it is striking to see that within this profile a number of metabolites (identified as early indicators) have been associated with DKD already in literature, but also that new candidate biomarkers were found. The discriminating metabolites included acyl-carnitines, acyl-glycines and metabolites

F. M. van der Kloet · F. W. A. Tempels · N. Ismail ·
R. van der Heijden · P. T. Kasper · M. Rojas-Cherto ·
R. van Doorn · G. Spijksma · J. van der Greef ·
T. H. Reijmers · T. Hankemeier (✉)
Division Analytical Biosciences, Leiden/Amsterdam
Center for Drug Research, Einsteinweg 55,
2333CC Leiden, The Netherlands
e-mail: hankemeier@lacdr.leidenuniv.nl

P. T. Kasper · M. Rojas-Cherto · R. van Doorn ·
J. van der Greef · T. H. Reijmers · T. Hankemeier
Netherlands Metabolomics Centre, Einsteinweg 55,
2333CC Leiden, The Netherlands

M. Koek · J. van der Greef
TNO Quality of Life, Utrechtseweg 48,
3704 HE Zeist, The Netherlands

V. P. Mäkinen · C. Forsblom · H. Holthöfer · P. H. Groop
Folkhälsan Institute of Genetics, Folkhälsan Research Center,
Biomedicum Helsinki, 1 Haartmaninkatu 8, 00290 Helsinki,
Finland

V. P. Mäkinen · C. Forsblom · P. H. Groop
Division of Nephrology, Department of Medicine,
Helsinki University Central Hospital, Helsinki, Finland

H. Holthöfer
Centre for BioAnalytical Sciences, Dublin City University,
Dublin, Ireland

related to tryptophan metabolism. We found candidate biomarkers that were univariately significant different. This study demonstrates the potential of multivariate data analysis and metabolomics in the field of diabetic complications, and suggests several metabolic pathways relevant for further biological studies.

Keywords Diabetic kidney disease · Nephropathy · Metabolite profile · Metabolomics · Urine · GC–MS · LC–MS · Multivariate data analysis

1 Introduction

As the number of patients with diabetes increases, diabetic kidney disease (DKD) is a growing public health problem. An estimated third of type 1 diabetic patients will develop DKD over the course of several decades after diabetes onset (Finne et al. 2005; Gross et al. 2005). These diabetes patients have a 10-fold risk of premature death due to cardiovascular and other circulatory diseases, and the risk increases even more for those who develop renal failure (Mäkinen et al. 2008; Groop et al. 2009; Soedamah-Muthu et al. 2004). Typical clinical manifestations of DKD include increased urinary albumin excretion rate (AER) and rising blood pressure, histological manifestations include the Kimmelstiel–Wilson nodules in the glomerulus (Kimmelstiell and Wilson 1936). DKD cannot be cured at present, but improved glycemic control and aggressive treatment of high blood pressure can halt the progression of the disease, especially when administered at an early stage of DKD (Lachin et al. 2000; Thomas and Atkins 2006). AER is the primary diagnostic biomarker for DKD in clinical practice. Elevated levels of AER measured at follow-up times with respect to the AER levels measured at baseline indicate if a patient is suffering a progressive form of DKD. For healthy individuals it is not uncommon to find elevated levels of AER which are clinically at the edge of microalbuminuria, but have no damage of the kidney function, therefore AER is not an early predictive marker or a quantitative measure of the kidney function at an early stage of kidney disease (Caramori et al. 2006). It may be possible that subtle alterations in metabolic pathways precede the changes that manifest as macroalbuminuria. These changing levels of the related low-molecular weight metabolite profiles may therefore be useful as early markers of a progressive form of DKD.

Mass spectrometry-based metabolomics has been extensively applied in disease diagnostics (Johnson 2005; Jankevics et al. 2009). Nevertheless, there are only a handful of similar studies of human DKD (Mäkinen et al. 2006; Mäkinen et al. 2008; Zhang et al. 2009) most of which were based on serum samples and differentiated between patients suffering from DKD and a healthy control group. In this study, urine samples from 52 type 1 diabetic patients from the FinnDiane Study that were clinically defined as having a normal AER (<30 mg/24 h, Mäkinen et al. 2006) were profiled. Half of this group (26 patients) suffered from the progressive form of albuminuria; the other half did not show a progression in albumin excretion. Both gas chromatography–mass spectrometry (GC–MS) and liquid chromatography–mass spectrometry (LC–MS) were used to analyze a wide range of metabolites in these urine samples. The data was from an actual patient population measured within a less controlled environment. As changes in biological samples are often multifactorial (Zhang and Chan 2005; Qiu et al. 2008), both univariate and multivariate data analyses were used. The multivariate metabolite profiles to differentiate between the two groups were found using logistic regression (LR) with variable selection. Based on MS^n fragmentation experiments (LC–MS only), manual interpretation combined with database searches we could identify several of the discriminating compounds that may be relevant for further biological studies.

2 Experimental

2.1 Samples

At baseline, type 1 diabetic patients were recruited by the Finnish Diabetic Nephropathy Study Group (FinnDiane). The initial data collection was cross-sectional (serum and urine samples), but with longitudinal records of albuminuria and clinical history. These study patients suffering from type 1 diabetes mellitus had an age of onset below 35 years and a transition to insulin treatment that occurred within a year after onset. The classification of renal status was made centrally according to urinary albumin excretion rate (AER) in at least two out of three consecutive overnight or 24 h-urine samples. Absence of diabetic kidney disease was defined as AER within the normal range (AER <20 μg/min or <30 mg/24 h). Prospective clinical data were available for a subset of patients, all being male. There were 26 subjects that progressed from normal AER to microalbuminuria (PR) and had urine samples available at the time having normal AER. 26 clinically group-matched (age, diabetes duration, baseline albuminuria status, sex) non-progressive AER (NP) subjects were selected as study reference. Subjects for this group (NP) that had a long follow-up were preferentially included. Table 1 shows some clinical characteristics of these subjects for the non-progressive AER and progressive AER groups. A more detailed table is included in the supplement (Table S1).

Table 1 Clinical characteristics of the subjects at baseline

	Clinical parameters (Normoalbuminuric)	
	Non-progressive	Progressive
Number of samples (Male)	26	26
Age (years)	36 ± 9	35 ± 10
Blood pressure (mm Hg)	80/132 ± 8/10	82/132 ± 11/15
BMI (kg/m^2)	25 ± 2	27 ± 3
Serum creatinine (μmol/l)	88 ± 14	86 ± 13
AER (mg/24 h)	12 ± 6	14 ± 7
HbA$_{1c}$ (%)	8 ± 1	9 ± 1
Diabetes duration (years)	27 ± 6	18 ± 11
Follow-up time (years)	6 ± 1	5 ± 2

2.1.1 GC–MS

All urine samples were processed and analyzed once using a randomized sample sequence over multiple batches. After every 6th study sample a quality control (QC) sample was injected. The QC sample was obtained by taking an aliquot of the same volume of all urine samples from this study. They were prepared once and measured in duplicate. All urine samples were analyzed with GC–MS according to the method described below.

2.1.2 LC–MS

The urine samples of the normal AER subjects were split into two aliquots, which were independently processed by the sample-pretreatment. Each extract was subsequently analyzed by LC–MS once so that for each sample from the normal AER group in total two LC–MS analyses were obtained. A pooled QC sample (the same as for GC–MS) was analyzed after every 6th sample.

2.2 Materials

2.2.1 GC–MS

For the GC–MS analysis, pyridine and N-methyl-N-tri-methylsilyl trifluoroacetamide were obtained from Mallinckrodt Baker BV (Deventer, The Netherlands) and Alltech (Breda, The Netherlands), respectively. Standards were purchased from Sigma-Aldrich (Zwijndrecht, The Netherlands).

2.2.2 LC–MS

For the LC–MS method, LC–MS grade acetonitrile (AcN) and MS-grade water were obtained from Biosolve (Valkenswaard, The Netherlands). All standards were

purchased from Sigma-Aldrich, except for phenylalanine-d5, which was from C/D/N Isotopes Inc. (PointeClaire, Quebec, Canada). Acetic acid, formic acid and sodium hydroxide were obtained from Biosolve (Valkenswaard, The Netherlands), Acros Organics (Geel, Belgium) and Merck (Darmstadt, Germany), respectively.

2.3 Sample preparation

2.3.1 Initial sample preparation

Urine samples (stored at −80°C) from FinnDiane were thawed at room temperature, homogenized using a vortex and centrifuged at 7500 g for 20 min. Specific volumes from the supernatant (typically 100 or 50 μl) were taken and 10 vol% aliquots of a 1 M acetic acid solution (adjusted to pH 6.0 with solid sodium hydroxide) were added and stored in vials. Samples were then stored at −80°C prior to method specific sample preparation.

2.3.2 GC–MS specific sample preparation

Sample preparation for GC–MS analysis was done similar to the method described elsewhere (Koek et al. 2006). In short, samples and standard urine solutions were thawed at room temperature and homogenized using a vortex. All 80 μl samples were mixed with 10 μl solutions containing the internal quality standards leucine-d3, glutamic acid-d3, phenylalanine-d5 and cholic acid-d4 (each present at a concentration of about 250 μg/ml in methanol/water (1:4 v/v)) and subsequently lyophilized at −37°C in autosampler vials. The internal quality standards alanine-d4 and glucose-d7 in pyridine (each about 250 μg/ml) were added to the dry extracts prior to oximation. Oximation (90 min at 40°C) was performed after adding 20 μl of a 56 mg/ml ethoxyamine hydrochloride solution in pyridine and 20 μl of pyridine to the extracts. Next, a 10 μl mixture of trifluoracetylanthraceen, dicyclohexylphthalate (DCHP) and difluorobiphenyl (each at a concentration of 250 μg/ml in pyridine) was mixed with the extracts and the mixtures were silylated for 50 min at 40°C with 200 μl of N-methyl-N-trimethylsilyl trifluoroacetamide. Then the samples were centrifuged (500 g for 20 min) and the supernatant was taken for analysis by GC–MS. The final GC–MS prepared samples were containing standards at a concentration of 10 ng/ml each. More details are given in the supplement.

2.3.3 LC–MS specific sample preparation

The FinnDiane urine samples were thawed at room temperature and subsequently mixed using a vortex. Next, the urine was centrifuged (7500 g for 10 min) at room temperature. To obtain the LC–MS samples, 35 μl of the

supernatant was transferred into an autosampler vial and mixed with 10 μl of the internal standards mix (valine-d8, phenylalanine-d5, tryptophan-d5, thymine-d4 and reserpine in water at a concentration of 21, 21, 21, 35 and 7 μg/ml, respectively) and 25 μl water.

For the validation of the LC–MS method (see supplement), two urine samples of one normoalbuminuric diabetic and one healthy volunteer were mixed in a 1:1 ratio. To 500 μl urine, 100 μl of a solution of 1 M acetic acid (adjusted to pH 6.0 with sodium hydroxide), 50 μl of the internal standards mix (valine-d8, phenylalanine-d5, tryptophan-d5, thymine-d4 and reserpine in water at a concentration of 60, 60, 60, 100, 20 μg/ml, respectively) and a varying volume of the calibration mix (phenylalanine, tryptophan and salicylamide in water at a concentration of 100 μg/ml) were added. Subsequently, water was added so that finally 1000 μl of validation sample was obtained for each calibration concentration. The samples were centrifuged (7500 g for 10 min) and the supernatant was analyzed by the LC–MS method. More details are given in the supplement.

2.4 Identification of metabolites

High resolution mass spectra were acquired using the 1200 Agilent gradient LC system coupled to a linear ion trap–Fourier transform (LTQ-FT) hybrid mass spectrometer and a LTQ-orbitrap mass spectrometer (both from Thermo Fisher Waltham, MA) Both systems were equipped with an ionmax ESI source. Spectra were recorded only in positive ESI centroid ion mode, with a source temperature of 275°C, source voltage of 4 kV and a sheath gas of 40 AU.

2.4.1 FT

LTQ-FT was setup for a MS^3 scanning method. Resolution was set to 12500 for all events to decrease scan time. Scan event one was a full scan with a scan range from 120 to 1000 m/z. Scan event two was set to fragment one of the targeted masses with a CID energy of 35% and isolation width of 1.5 m/z. Scan event three was set to Data Dependent scan where it was set to fragment the most intense ion from scan event two with CID of 35% and isolation width of 1.5 m/z. All spectra were recorded in FT-mode with a typical mass accuracy of <1.5 ppm for both full scan and MS/MS spectra.

2.4.2 Orbitrap

The LTQ-orbitrap was setup for (MS^3) HCD fragmentation measurements. Resolution was set to 7500 for all events to decrease scan time. Scan event one was setup for a full scan with a scan range from 120 to 600 m/z. Scan event two was set to fragment one of the targeted masses in HCD fragmentation mode with a HCD energy of 30% and isolation width of 1.5 m/z. Scan event three was set to fragment one of the targeted masses in HCD fragmentation mode with a HCD energy of 50% and isolation width of 1.5 m/z. All spectra were recorded in FT-mode with lock masses 279.15909 and 391.28429 m/z in full scan. Mass accuracy in full scan spectra was <2 ppm and in HCD MS–MS spectra <4 ppm.

2.4.3 Interpretation of the spectra

The spectra were collected and molecular formulae were then calculated by an in-house developed software tool using MS^3 fragmentation data. Spectra together with molecular formulae were further interpreted manually. The compounds that were identified were confirmed by comparison to spectra found in databases such as HMDB and by authentic standards where possible.

2.5 Data processing and analysis

2.5.1 GC–MS

After sample analysis with GC–MS a target table of all relevant peaks (with known and unknown identity) was constructed. For this, a standard target table containing over 300 entries of endogenous plasma metabolites and urine specific peaks was used ultimately leading to a target list of 144 compounds. These compounds were integrated using a reconstructed ion chromatogram of a characteristic ion of each compound. The internal standards were quantified in the same manner. The GC–MS data have been corrected for internal standard response using DCHP followed by QC correction as described earlier (van der Kloet et al. 2009).

2.5.2 LC–MS

After analysis of all samples using LC–MS a small subset of samples that contained 2–3 samples of each albuminuric class was used for screening compounds using software provided by Bruker-Daltonics (DataAnalysis). The m/z resolution was set to 0.01 Da, the minimum S/N ratio was set to 3 and to prevent integration problems the retention time window was set to 20 s. This resulted in a target list of about 600 features, where each feature was characterized by a retention time and a mass, and which could represent a metabolite. One metabolite can have multiple features. This target list was used as input for the software package Quant-Analysis by Bruker-Daltonics to create extracted ion chromatograms (EICs) for all peaks for all LC–MS samples. In this way a peak table was constructed that was further narrowed down by removing features that had lots

of missing values or no response at all in the QC samples (Bijlsma et al. 2006). As the dead time of the LC–MS method was 1 min, features detected with retention times lower than 2 min were discarded. Features with large differences in retention time (RSD higher than 5%) were removed as well. This resulted in a final peak table of 106 features. For the LC–MS data the response has been corrected using the most optimal internal standard as described earlier (van der Kloet et al. 2009).

2.5.3 GC–MS and LC–MS sample normalization

For further statistical analysis the feature representing glucose (indicative for diabetes) was removed from the data. For healthy subjects the response of urine samples is often normalized by its creatinine level. In the case of varying degrees of kidney failure (i.e. microalbuminuric or macroalbuminuric), the creatinine levels cannot be used for normalization because they show irregular behavior due to diabetes and/or medication. Furthermore, a recent report shows that the creatinine concentration may vary over a large scale and that variances in urinary metabolite concentrations are not due to urine dilution effects; rather, they reflect actual metabolite variances (Jankevics et al. 2009; Saude et al. 2007). To circumvent this, as a means of normalization per sample, row scaling (i.e. subject normalization) was applied by taking the sum of the peak areas of all the components measured and dividing the response of each metabolite in a sample by this sum (Kemperman et al. 2007). For LC–MS measurements duplicates were averaged per metabolite after visual inspection. All data were auto-scaled prior to further multivariate statistical analysis.

All multivariate data analyses were performed using Matlab 2008a (Mathworks 2008). The PCA charts were created using the PLS-toolbox 5.5.2 from Eigenvector Research (2008).

3 Results and discussion

3.1 Data processing and quality

The LC–MS method was set up and successfully validated for urine (for details, see supplement), while the GC–MS method validation results were published earlier (Koek et al. 2006). To monitor the stability of the analytical system, quality control (QC) samples were measured during both GC–MS and LC–MS analyses. In these QC samples the responses of each compound should be constant over time. The stability of the response per compound is expressed in relative standard deviation (RSD) values, in which case for each metabolite the standard deviation of

the response in all QC samples is divided by the average of the response in all QC samples. Large RSD values of response indicate poor repeatability, which can be assigned to instability of the analytical system or due to other variations such as instability of the analyte, etc. For GC–MS a total of 144 compounds were measured of which 106 had an RSD of less than 10% in the QC samples. The remainder was predominantly in the RSD range of 10–20% (for 9 compounds the RSD value was larger than 30%). For LC–MS, 106 features were found of which 65 had an RSD value of 10–20% and of which the remainder was predominantly in the 20–30% RSD range (for 8 compounds the RSD value was larger). 130 compounds (GC–MS) and 89 features (LC–MS) were selected for further data analyses based on RSD values less than or equal to 25%.

3.2 GC–MS results

Univariate tests for significant difference between non-progressive and progressive AER subjects [t-tests and Wilcoxon tests (Miller and Miller 2005)] were executed for each compound at a 95% significance level. The probability of a Type I error (i.e. the error of rejecting a null hypothesis when it is actually true) was further reduced using the Bonferroni approach by setting the significance level to $\alpha = 0.05/130 \approx 0.00039$. None of the tested compounds showed a significant difference.

Principal component analysis (PCA), an unsupervised multivariate data analysis method, was used to investigate whether there was an apparent metabolomic separation between the non-progressive normo AER subjects vs. the progressive AER subjects. Using the GC–MS data of all metabolites, PC1 vs. PC2 did not show any clear separation
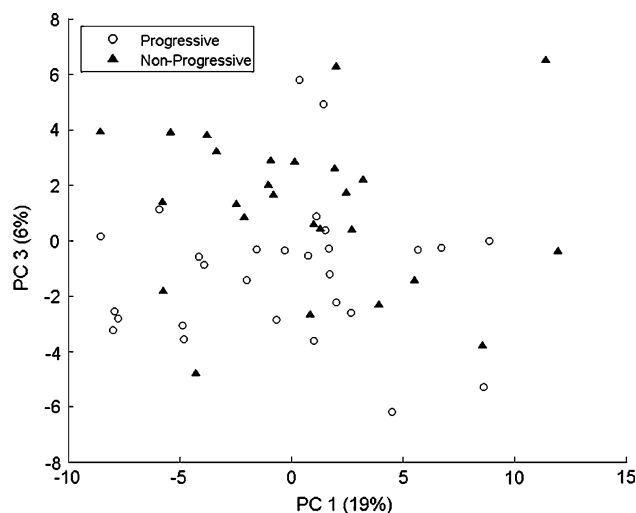

Fig. 1 PCA score plot of the GC–MS data of urine from normal AER subjects using 130 compounds

between the two groups, whereas PC1 vs. PC3 (Fig. 1) showed some clustering of the two classes.

Because no clear separation was visible in the PCA score plot, supervised multivariate data analysis was used. As the classification problem had a dichotomous outcome (progressive or non-progressive), we used multivariate logistic regression (LR) in which the data was modeled in such a manner that the predicted outcome was always bounded between zero and one (corresponding with the 2 groups) (Westerhuis et al. 2008; Hastie et al. 2008). The choice of logistic regression was made because an ordinary linear regression method assumes that in the population a normal distribution of error values around the dependent variable is associated with each independent variable, and that the dispersion of the error values for each of these independent values is the same. However, the distribution of errors for any independent value cannot be normal when the distribution has only two values (Pampel 2000). To prevent over-fitting we used cross-validation followed by permutation tests. The choice of logistic regression in combination with variable selection allowed us to deal with the heterogeneity of the data and obtain stable cross validated models (for details see supplement).

The comparison between normal AER PR and normal AER NP subjects rendered a cross-validated logistic regression model with an accuracy of 65% and a precision of 64%. Ultimately 65 out of the 130 available metabolites were left in the final model. Accuracy can be viewed as the "overall effectiveness of a classifier" and precision as "class agreement of the prediction with a specific class (e.g. progressive, non-progressive)" (Sokolova and Lapalme 2009). Details regarding the cross-validation, the exact definition of accuracy and precision can be found in the supplement (Table S3).

To obtain a list of candidate biomarkers that form the predictive metabolic profile using this multivariate model, the significant contribution of each of these biomarkers to the model and the model itself was evaluated as described below. Permutation tests by means of randomizing the class membership vector were performed to evaluate the significance of the logistic regression model differentiating the non-progressive from the progressive normoalbuminuric subjects (see Supplement). The unpermuted model showed a tendency towards being significant (16 out of 100 permuted models gave equal or better classification results).

Furthermore, using the regression vectors obtained from the permutation tests the significance of the contribution to the logistic regression model for each of the 65 compounds was determined. In total 34 compounds were found to be significant with a P-value lower than 0.05 (at most 5 out of 100 permuted models had a larger regression coefficient than the unpermuted model, see Supplement). Table 2 lists those compounds (21 in total) that were identified from this list of 34 compounds, ranked by their significance, together

with their up-regulation (i.e. the relative concentration increased for the PR samples compared to the NP samples), their multivariate P-value and t-test P-value.

3.3 LC–MS results

With a Bonferroni corrected α of 0.00056 (0.05/89), 3 out of the 89 features showed significantly different group means with either a t-test or a Wilcoxon test (Fig. 2). As most of the statistically relevant features so far had not been identified in our lab for the LC–MS method, they were subjected to identification using high resolution MS and multi-stage MS/MS (see Experimental). Table 3 lists these three compounds, their respective P-values and their up-regulation (i.e. the relative concentration increased for the PR samples compared to the NP samples).

The PCA score plot of LC–MS results of the normal AER subjects (Fig. 3) showed some clustering along the diagonal of the first and second principle component of the progressive and non-progressive subjects.

Analogue to the GC–MS data analysis, LR with variable selection was used. The resulting model contained 42 features. The accuracy of the cross-validated logistic regression model for the binary classification of NP vs. PR was 75% with a precision of 73%.

To evaluate the significance of metabolites contributing to the LC–MS based LR model differentiating between non-progressive and progressive normoalbuminuric subjects, and the model itself, permutations test were performed. The model was found to be significant; only 4 out of 100 permuted models gave equal or better classification results (see Supplement). Using the regression vectors from the permutation tests, the significance of each of the 42 features was determined. 14 features were significant (at most 5 out of 100 permuted models had a larger regression coefficient than the unpermuted mode, see supplement). High resolution MS and multi-stage MS/MS (see Experimental) were used to identify these features. Table 4 lists 8 of these 14 features that were identified ranked by significance, together with their up-regulation (i.e. the relative concentration increased for the PR samples compared to the NP samples) and the univariate t-test P-value. Literature study revealed that several compounds that were identified with a multivariate significance higher than a P-value of 0.05 could be linked to DKD. As these compounds also contributed to the model, these compounds were added to Table 4. Note that the compounds that showed a univariate significance were also included in the LR model (Table 4).

3.4 Strengths and weaknesses

Both GC–MS and LC–MS data suffer from heterogeneity in the (samples of the) study population. The origin of this

Table 2 Metabolites discriminating progressive and non-progressive normal AER subjects using the logistic regression of GC–MS data

Compound	Up-regulation[a]	Significance[b]	t-test P-value
4-Oxoproline[c]	−1	0	0.03
Pseudouridine[d]	1	0	0.41
3,4,5-Trihydroxypentanoic acid[e]	−1	0	0.09
Deoxyfructose[c]	1	0	0.80
3-Hydroxy-3-(3-hydroxyphenyl) propanoic acid[e]	1	0	0.58
L-Valine[d]	1	0.01	0.25
2,3-Dihydroxy-3-methylbutanoate[c]	−1	0.01	0.27
5-Hydroxymethyl-2-furancarboxylic acid[e]	−1	0.02	0.09
Galactonic acid[e]	−1	0.02	0.01
2-Hydroxyvaleric acid[c]	1	0.02	0.58
N-formylproline or N-ethylproline[c]	−1	0.02	0.00
2-Hydroxyglutaric acid[d]	1	0.02	0.51
N-(3-hydroxybenzoyl)glycine[e]	1	0.02	0.47
L (+) Arabinose[d]	−1	0.03	0.13
Benzoic acid[d]	−1	0.03	0.04
3-Hydroxyphenylacetic acid[d]	−1	0.03	0.01
Glucuronide compound[c]	1	0.03	0.32
D-Glutamicacid[d]	−1	0.05	0.04
Gluconicacid[d]	1	0.05	0.04
Glycolicacid[d]	−1	0.05	0.05
L-Cystine[d]	1	0.05	0.37

[a] Increased concentration for progressive subjects, i.e. positive for progression: 1, decreased concentration for progressive subjects: −1

[b] The number of times that the metabolite in a permuted model had a larger regression coefficient than the unpermuted model divided by the total number of permutations executed

[c] Compounds were characterized, and only the class of metabolite could be suggested

[d] Compounds were identified, and the identity confirmed by an authentic standard

[e] Compounds were annotated based on elemental composition and by comparison to reference libraries

heterogeneity is the result of multiple factors, the main ones are: (i) the difficulty to obtain an exact kidney phenotype, as discussed in the introduction; AER is the primary diagnostic biomarker in clinical practice, but its usefulness as an early marker is limited due to high natural variance (Caramori et al. 2006); and (ii) the uncontrolled environment in which the urine samples were taken (e.g. urine metabolite concentrations depend on diet, variation due to slight differences in sampling 24 h urine, etc.). This heterogeneity and the small number of prospective samples severely complicated proper statistical analyses. In cross-validating the binary classification models, the leave two out strategy was used (1 sample was left out for each class). In order to select those features/compounds in the model that were specific for the whole dataset and not just for a few samples, a variable selection method (see above) was included. It turned out that when a subset of variables was used more accurate predictive models were obtained that were less susceptible to different training/test-set schemes (data not shown). Although the number of prospective samples is small the nature of the data (i.e. urine samples at

an early stage of diabetic kidney disease) and the fact that some of the found biomarkers were already related to DKD the obtained results certainly give rise to future research.

3.5 Biological context of the new candidate biomarkers

Reviewing the metabolites in Table 2 it is interesting to note that many of the GC–MS compounds are carboxylic compounds and/or acidic metabolites that are prevalently detected in urine (Lawson et al. 1976) (5-hydroxymethyl-2-furancarboxylic acid, benzoic acid and hippuric acid). Others are endogeneous amino acids (valine, serine). From the identified compounds, we have not found any documented direct relation to DKD. Deoxyfructose could be related to a phenomenon called "diabetic stress" in which deoxyglucosone is converted to the less reactive deoxyfructose (Knecht et al. 1992). Galactonic acid has been associated with diabetic retinopathy (Kador et al. 2002). Rainey et al. already suggested an interaction of 5-hydroxymethyl-2-furancarboxylic acid with galactonic acid (Mrocheck and Rainey 1972).

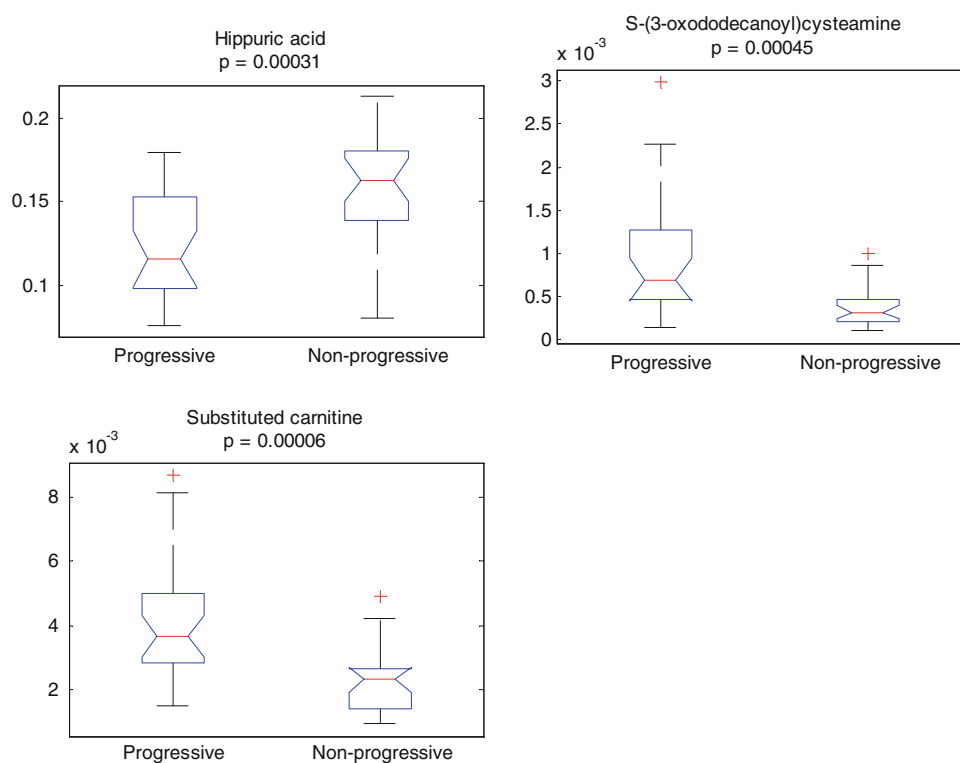**Fig. 2** Boxplots of the 3 compounds that showed significant group means

Table 3 The 3 metabolites that show a statically relevant difference between the group means of the progressive group and non-progressive normoalbuminuric group

Compound	*P*-value		Up-regulation[a]
	t-test	Wilcoxon	
Substituted carnitine[b]	0.00000592	0.00000368	1
Hippuric acid[c]	0.00003066	0.00004267	−1
S-(3-oxododecanoyl) cysteamine[c]	0.00004540	0.00003714	1

[a] 1 means increased concentration for progressive subjects, i.e. positive for progression, −1 means decreased concentration for progressive subjects

[b] Compounds were characterized, and only the class of metabolite can suggested

[c] Compounds were annotated based on elemental composition and MS/MS (e.g. by comparison to reference libraries)

The metabolites in LC–MS (Table 4) are either: (1) acylcarnitines, (2) acyl-glycines, (i.e. salicyluric acid, hippuric acid, (2-phenylacetoxy- propionyl) glycine and 3-methylcrotonylglycine) and (3) compounds related to the tryptophan metabolism (i.e. tryptophan, indoleacetic acid and kynurenic acid).

It is perhaps not that straightforward to link the metabolites to a specific pathway as a recent study already demonstrated that many of the acylglycines and tryptophan metabolites in mammalian blood have shown a relation to gut "microbiome" (Wikoff et al. 2009) which could explain a presence in urine by regular excretion.

In general, acyl-carnitines are formed in the fatty acid metabolism pathway to transport the long-chain acyl

groups of fatty acids into mitochondria, where these groups are broken down through ß-oxidation to acetate to obtain energy via the citric acid cycle. Under normal homeostasis conditions, carnitine is eliminated by excretion in urine, in both free and esterified forms, mainly as acetylcarnitine (Chalmers et al. 1984). A higher acylcarnitine to carnitine ratio in urine in relation to plasma is suggested to be the result of a less efficient reabsorption of acylcarnitines or of a renal acylation of carnitine followed by leakage of the locally formed acylcarnitine product into urine (Vernez 2005; Wagner et al. 1986; Rebouche and Seim 1998) .

Glycine conjugation is an effective detoxification system for preventing accumulation of acyl-CoA esters in several inherited metabolic disorders. Acylglycines in urine
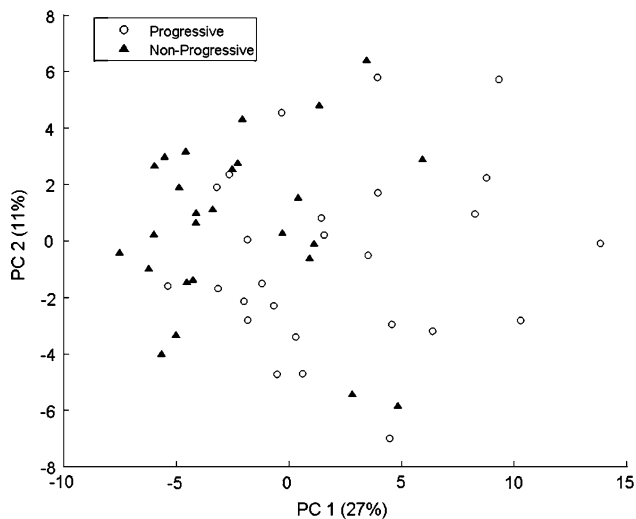
Fig. 3 PCA score plots of the LC–MS data of the urine samples from Normal AER subjects using 89 features

have been reported as the direct expression of accumulation of the correspondent acyl-CoA esters in the mitochondrion (Bonafe et al. 2000).

Tryptophan metabolism changes with DKD have been reported before (Bonafe et al. 2000), where tryptophan plasma concentration in animals with experimental renal failure decreased while a simultaneous increase of metabolites related to the kynurenine pathway (e.g. kynurenic acid) in plasma were observed. It was demonstrated in animals that the content of kynurenic acid in kidneys is the highest among all tissues (Lou et al. 1994). Furthermore, it has been known that kynurenic acid is the main metabolite excreted from organism (rats) by means of tubular secretion. In renal failure this mechanism is considerably impaired, which in consequence leads to excessive accumulation of this substance in the organism (Pawlak et al. 2002).

Of interest are the elevated levels of 2-(2-phenylacetoxy)propionylglycine in the progressive patients. In the absence of medium chain acyl-CoA dehydroxygenase (MCAD) phenylpropionacid is converted to phenylpropionglycine instead of benzoic acid. Phenylpropionglycine is detected only in the urine of MCAD-deficient patients and has been used as a biomarker for the diagnosis of this condition (Wikoff 2009). However, in summary, it should be mentioned that further research is required to investigate the biochemical context of all the candidate biomarkers in more detail.

4 Conclusions

It was demonstrated that based on LC–MS measurements of urine samples a statistically significant multivariate model could be constructed to distinguish between

Table 4 Identified compounds from the metabolites discriminating most between progressive and non-progressive normoalbuminuric subjects using logistic regression model of LC–MS data

Compound	Up-regulation[a]	Significance[b]	t-test P-value
Tryptophan[c]	1	0	0.00
Salicyluric acid[c]	1	0.01	0.09
Substituted carnitine[d]	1	0.01	0.00
S-(3-oxododecanoyl) cysteamine[e]	1	0.01	0.00
Hippuric acid[e]	−1	0.02	0.00
Substituted carnitine[d]	1	0.02	0.00
N-methyl guanosine[e]	1	0.04	0.00
Substituted carnitine[d]	1	0.05	0.00
Kynurenic acid[c]	1	0.08	0.18
2-(2-phenylacetoxy)propionylglycine[e]	1	0.09	0.45
Substituted carnitine[d]	−1	0.09	0.36
Indoleacetic acid[c]	−1	0.1	0.15
3-methylcrotonylglycine[e]	1	0.1	0.03
Heptanoylcarnitine[d]	1	0.1	0.00

[a] 1 means increased concentration for progressive subjects, i.e. positive for progression, −1 means decreased concentration for progressive subjects

[b] The number of times that a metabolite in a permuted model had a larger regression coefficient than the unpermuted model divided by the total number of permutations executed

[c] Compounds were identified, and the identity confirmed by an by an authentic standard based on MS/MS and retention time

[d] Compounds were characterized, and only the class of metabolite can suggested

[e] Compounds were annotated based on elemental composition and MS/MS (e.g. by comparison to reference libraries)

progressive and non-progressive subjects within the normal AER group with an accuracy of 75%. Many of the compounds contributing to the model could be grouped in one of three classes, i.e. acyl-carnitines, acyl-glycines and compounds related to the tryptophan metabolism. All of the compounds that were measured that show a univariate significant difference between the two groups were included in the metabolic profile defined by the multivariate model. The metabolic profile also included metabolites that did not show a univariate significant difference and emphasizes the additional benefit of multivariate statistics over univariate statistics alone in preventing overlooking candidate biomarkers.

Future research will focus on the discovery of additional biomarkers using complementary metabolomics platforms and the validation of the explorative biomarker profiles with a validation set. In addition, more effort will be directed to the biological interpretation: it will be investigated which pathways were involved in the biochemical changes associated with the onset, development and progression of DKD, and whether these changes are the same during onset and progression, or if different changes of biochemistry occur at the different stages of DKD, e.g. due to the disease pathology or due to the use of medication after onset of DKD. In summary, the results obtained demonstrate the potential of metabolomics in the study of diabetic complications, as subtle changes in the urine metabolome precede the clinically significant rise in AER.

References

Bijlsma, S., et al. (2006). Large-scale human metabolomics studies: A strategy for data (pre) processing and validation. *Analytical Chemistry, 78*, 567–574.

Bonafe, L., et al. (2000). Evaluation of urinary acylglycines by electrospray tandem mass spectrometry in mitochondrial energy metabolism defects and organic acidurias. *Molecular Genetics and Metabolism, 311*, 302–311.

Caramori, M. Luiza, Fioretto, P., & Mauer, M. (2006). Enhancing the predictive value of urinary albumin for diabetic nephropathy. *Journal of the American Society of Nephrology, 17*(2), 339–352.

Chalmers, R. A., et al. (1984). Urinary excretion of L-carnitine and acylcarnitines by patients with disorders of organic acid metabolism: Evidence for secondary insufficiency of L-carnitine. *Pediatric Research, 18*(12), 1325–1328.

Eigenvector Research. (2008). *PLS toolbox.* Available at www.eigenvector.com

Finne, P., et al. (2005). Incidence of end-stage renal disease in patients with type 1 diabetes. *Journal of the American Medical Association, 294*(14), 1782–1787.

Groop, P.-H., et al. (2009). The presence and severity of chronic kidney disease predicts all-cause mortality in type 1 diabetes. *Baseline, 58*(7), 1651–1658.

Gross, J. L., et al. (2005). Diabetic nephropathy: Diagnosis, prevention, and treatment. *Diabetes Care, 28*(1), 164–176.

Hastie, T., Tibshirani, R., & Friedman, J. (2008). *The elements of statistical learning* (2nd ed.). New York: Springer.

Jankevics, A., et al. (2009). Chemometrics and intelligent laboratory systems metabolomic studies of experimental diabetic urine samples by 1 H NMR spectroscopy and LC/MS method. *Chemometrics and Intelligent Laboratory Systems, 97*(1), 11–17.

Johnson, D. W. (2005). Contemporary clinical usage of LC/MS: Analysis of biologically important carboxylic acids. *Clinical Biochemistry, 38*, 351–361.

Kador, P. F. et al. (2002). Effect of galactose diet removal on the progression of retinal vessel changes in galactose-fed dogs. *Investigative Ophthalmology, 43*, 1916–1921.

Kemperman, R. F. J., et al. (2007). Comparative urine analysis by liquid chromatography-mass spectrometry and multivariate statistics: Method development, evaluation, and application to proteinuria research articles. *Journal of Proteome Research, 6*, 194–206.

Kimmelstiell, P., & Wilson, C. (1936). Benign and malignant hypertension and nephrosclerosis. *American Journal of Pathology, 12*, 4548.

Knecht, K. J., et al. (1992). Detection of 3-deoxyfructose and 3-deoxyglucosone human urine and plasma: Evidence for intermediate stages of the maillard reaction in viva. *Review Literature and Arts of the Americas, 294*(1), 130–137.

Koek, M. M., et al. (2006). Microbial metabolomics with gas chromatography/mass spectrometry. *Analytical Chemistry, 78*(4), 1272–1281.

Lachin, J. M., Genuth, S., Cleary, P., Davis, M. D., et al. (2000). Retinopathy and nephropathy in patients with type 1 diabetes four years after a trial of intensive therapy. *New England Journal of Medicine, 342*, 381–389.

Lawson, A. M., Chalmers, R. A., & Watts, R. W. E. (1976). Urinary organic acids in man. I. Normal patterns. *Clinical Chemistry, 1287*, 1283–1287.

Lou, G. L., Pinsky, C., & Sitar, D. S. (1994). Kynurenic acid distribution into brain and peripheral tissues of mice. *Canadian Journal of Physiology and Pharmacology, 72*(2), 161–167.

Mäkinen, V. P., et al. (2006). Diagnosing diabetic nephropathy by 1 H NMR metabonomics of serum. *Magnetic Resonance Materials in Physics Biology and Medicine, 19*, 281–296.

Mäkinen, V. P, Forsblom, C., et al. (2008). Metabolic phenotypes, vascular complications, and premature deaths in a population of 4,197 patients with type 1 diabetes. *Diabetes, 57*, 2480–2487.

Mäkinen, V. P., Soininen, P., et al. (2008). H NMR metabonomics approach to the disease continuum of diabetic complications and premature death. *Molecular Systems Biology, 4*(167), 1–12.

Mathworks. (2008). *Matlab.* Available at www.mathworks.com.

Miller, J. N., & Miller, J. C. (2005). *Statistics and chemometrics for analytical chemistry* (5th ed.). England: Pearson Education Limited.

Mrocheck, J. E., & Rainey, W. T. J. (1972). Identification and biochemical significance of substituted furans in human urine. *Clinical Chemistry, 18*(8), 821–828.

Pampel, F. C. (2000). *Logistic regression a primer, thousand oaks: Sage university papers on quantitative applications in the social sciences*. Thousand Oaks: Sage Publications.

Pawlak, D., et al. (2002). Tryptophan metabolism via the kynurenine pathway in experimental. *Nephron, 90*, 328–335.

Qiu, Y., et al. (2008). Multivariate classification analysis of metabolomic data for candidate biomarker discovery in type 2 diabetes mellitus. *Metabolomics, 4*(4), 337–346.

Rebouche, C. J., & Seim, H., (1998). Carnitine metabolism and its regulation in microorganisms and mammals. *The Annual Review of Nutrition, 18*, 39–61.

Saude, E. J., et al. (2007). Variation of metabolites in normal human urine. *Metabolomics, 3*(4), 439–451.

Soedamah-Muthu, S. S., et al. (2004). Risk factors for coronary heart disease in type 1 diabetic patients in Europe: The EURODIAB prospective complications study. *Diabetes Care, 27*(2), 530–537.

Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing and Management, 45*(4), 427–437.

Thomas, M. C., & Atkins, R. C. (2006). Blood pressure lowering for the prevention and treatment of diabetic kidney disease. *Drugs, 66*(17), 2213–2234.

van der Kloet, F. M., et al. (2009). Analytical error reduction using single point calibration for accurate and precise metabolomic phenotyping research articles. *Journal of Proteome Research, 8*, 5132–5141.

Vernez, L. (2005). *Analysis of carnitine and acylcarnitines in biological fluids and application to a clinical study*. PhD Thesis, University of Basel, Basel.

Wagner, S., Deufel, T., & Guder, W. (1986). Carnitine metabolism in isolated rat kidney cortex tubules. *Biological Chemistry Hoppe-Seyler, 367*, 75–79.

Westerhuis, J. A., et al. (2008). Discriminant Q2 (DQ2) for improved discrimination in PLSDA models. *Metabolomics, 4*(4), 293–296.

Wikoff, W. R., et al. (2009). Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. *Proceedings of the National Academy of Sciences of the United States of America, 106*(10), 3698–3703.

Zhang, Z., & Chan, D. W. (2005). Cancer proteomics: In pursuit of "true" biomarker discovery. *Biomarkers, 14*, 2283–2286.

Zhang, J., et al. (2009). Metabonomics research of diabetic nephropathy and type 2 diabetes mellitus based on UPLC-oaTOF-MS system. *Analytica Chimica Acta, 650*(14), 16–22.