Targeted Clinical Metabolomics Platform for the Stratification of Diabetic Patients

Running head: Clinical metabolomics platform for the diabetes clinic

Linda Ahonen[1,#], Sirkku Jäntti[2], Tommi Suvitaival[1], Simone Theilade[1], Claudia Risz[1], Risto Kostiainen[2], Peter Rossing[1,3], Matej Orešič[4,5], Tuulia Hyötyläinen[6]*

[1]Steno Diabetes Center Copenhagen, 2820 Gentofte, Denmark

[2]Drug Research Program, Division of Pharmaceutical Chemistry and Technology, Faculty of Pharmacy, University of Helsinki, 00014 University of Helsinki, Finland

[3]Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark

[4]Turku Centre for Biotechnology, University of Turku and Åbo Akademi University, Turku 20520, Finland

[5]School of Medical Sciences, Örebro University, 702 81 Örebro, Sweden

[6]Department of Chemistry, Örebro University, 702 81 Örebro, Sweden

[#]Current affiliation: Biosyntia ApS, 2100 Copenhagen, Denmark

*Corresponding author:

Prof. Tuulia Hyötyläinen, Ph.D.

Department of Chemistry, Örebro University, 702 81 Örebro, Sweden

Phone: +46 70 7882165; E-mail: tuulia.hyotylainen@oru.se

Abstract

Background

Several small molecule biomarkers have been reported in the literature for prediction and diagnosis of (pre)diabetes, its co-morbidities and complications. Here, we report the development and validation of a novel, quantitative, analytical method for use in the diabetes clinic. This method enables the determination of a selected panel of 36 metabolite biomarkers from human plasma.

Methods

Based on a review of the literature and our own data, we selected a panel of metabolites indicative of various clinically-relevant pathogenic stages of diabetes. We combined these candidate biomarkers into a single ultra-high-performance liquid chromatography-tandem mass spectrometry (UHPLC-MS/MS) method and optimized it, prioritizing simplicity of sample preparation and time needed for analysis, enabling high-throughput analysis in clinical laboratory settings.

Results

We validated the method in terms of limit of (a) detection (LOD), (b) limit of quantitation (LOQ), (c) linearity (R^2), (d) linear range, and (e) intra- and inter-day repeatability of each metabolite. The method's performance was demonstrated in the analysis of selected samples from a diabetes cohort study. Metabolite levels were associated with clinical measurements and kidney complications in type 1 diabetes (T1D) patients. Specifically, both amino acids and amino acid-related analytes were associated with macro-albuminuria. Additionally, specific bile acids were associated with kidney function, anti-hypertensive medication, statin medication and clinical lipid measurements.

Conclusions

The developed analytical method is suitable for robust determination of selected plasma metabolites in the diabetes clinic.

Background

The incidence of type 2 diabetes (T2D) is rising globally, currently estimated to exceed 450 million patients worldwide. In addition, the prevalence of prediabetes is approximately 2-3 times higher than for diabetes. Prediabetes is a condition with a high risk of progression to T2D, with a yearly conversion rate of 5%-10% *(1, 2)*. It is also known that excessive hepatic fat accumulation is a typical feature of T2D patients and plays an important, pathogenic role in disease development and progression. Particularly, non-alcoholic fatty liver disease (NAFLD) may have an important, deleterious impact on diabetic patients, increasing the risk of cardiovascular complications. Moreover, there is evidence of associations between prediabetes and complications of diabetes such as early nephropathy, small fiber neuropathy, early retinopathy and risk of macrovascular disease *(2)*. Therefore, there is a need for predictive tools for efficient and accurate tracking of the progression from the state of normal glucose tolerance (NGT) to pre-diabetes and finally to T2D, as well as a need for the identification of those individuals with T1D and T2D who are at risk of developing diabetic complications. There is also a need for improved stratification of those individuals who already have the disease based on their risk of developing complications. Finally, there is a pressing need to then tailor intervention strategies to these individuals. Ideally, knowledge about the underlying pathophysiological characteristics associated with either fasting or postprandial glucose dysregulation would be utilized in order to optimize the efficacy of any interventions *(3)*.

The complex etiology of diabetes makes effective screening, diagnosis, prognosis and intervention challenging. Several studies have shown changes in the circulating levels of specific metabolites prior to an individual developing overt T2D. For example, the Framingham Offspring, European Investigation into Cancer and Nutrition (EPIC) Potsdam, Metabolic Syndrome in Men (METSIM), Cardiovascular Risk in Young Finns (CRY) and Southall and Brent Revisited (SABRE) studies have replicated the finding of increased levels of branched-chain amino acids and their derivatives, aromatic amino acids and α-hydroxybutyrate, even years ahead of conversion to overt T2D *(4-9)*. Amino acids, particularly tyrosine, were found to be associated with risk of microvascular disease *(10)*. Also other metabolites, *e.g.* 1,5-anhydroglucitol, norvaline and l-aspartic acid, were found to be associated with macroalbuminuric diabetic kidney disease *(11, 12)*, while glutamine, glutamic

3

acid and symmetric dimethylarginine (ADMA) were suggested as potentially-predictive biomarkers of diabetic complications *(13-15)*. Several metabolites (*e.g.*, α-hydroxybutyrate, β-hydroxypyruvate and 1,5-anhydroglucitol (1,5-AG)), were associated with regulation of glycemic control *(16, 17)*. Many lipids were identified as predictive biomarkers of diabetes. Specifically, triglycerides of low carbon number and double bond count as well as lysophosphatidylcholine, LPC(18:2), were identified as early predictors of T2D *(18, 19)*. Notably, these markers were unaffected by obesity *(18)*. Additionally, bile acids were associated with T2D and insulin resistance *(20, 21)*. Mannose *(22)*, 2-aminoadipic acid *(23, 24)* as well as indoxyl-sulfate and cresyl-sulfate *(26)* were suggested as possible biomarkers and, finally, creatinine*(25)* is already routinely implemented as an estimate of renal function.

Most of the studies described above have been performed with non-targeted metabolomics methods, using workflows which are difficult to apply in routine clinical laboratory settings. Herein, our goal was to develop a fast and robust method for quantitative analysis of a selected panel of metabolite biomarkers, which are informative as to the prediction and diagnosis of (pre)diabetes and its co-morbidities / complications, as well as in follow-up of interventions. We developed a method which includes 36 metabolites, representing several metabolite classes, including amino acids, bile acids, carnitines, phenolic compounds and small organic acids. The method is based on simple sample preparation and fast, quantitative ultra-high-performance liquid chromatography coupled to tandem mass spectrometry (UHPLC-MS/MS) analysis. Both sample preparation and the subsequent analyses were optimized and validated. Additionally, the method was demonstrated in a subset of samples from a cohort of diabetic patients, who were observed at the Steno Diabetes Center Copenhagen between 2009 and 2011 *(27)*.

Materials and methods

Chemicals and standard solutions

Information on the chemicals used for this work is given in **Supplemental Methods**. Additionally, the structures of the compounds (**Supplemental Figures S1A-C**) and the preparation of standards and stock solutions is also described in **Supplemental Methods** and in **Tables 1** and **2**.

Samples

Plasma samples from a previously-described cohort *(27, 28)* were used for validation of the method. In short, during 2009-2011, a total of 1285 patients were invited to enter a study examining diabetic complications at the Steno Diabetes Center Copenhagen (SDCC). The study conformed to the Declaration of Helsinki and was approved by the Danish National Committee on Biomedical Research Ethics (2009-056; NCT01171248). Additionally, all patients gave written, informed consent. Of the invited 1285 patients, 676 accepted to participate and for our purposes, to demonstrate method functionality, a subset of 50 patient samples was analyzed. In addition to these plasma samples, pooled plasma samples from the SDCC were used for method development and validation as well as for quality control. All plasma samples were stored at -80 °C until analysis.

Sample preparation

Sample preparation included protein precipitation and derivatization. 10 µL of 1 M 5-sulphosalisylic acid dehydrate (SSA) solution was added to 30 µL of plasma sample, samples were vortex mixed and centrifuged at 9000 RCF (5 minutes at 4 °C) after which 20 µL of the upper phase was collected, After addition of 20 µL of the ISTD MIX 20 µL of a 6-aminoquinoline-N-hydroxy-succinimidyl carbamate-reagent (AQC-reagent) (5 mg mL^{-1}, at 55 °C) was added, and the samples were vortex mixed and stored at -80 °C until analysis.

The samples in the validation study were randomized before sample preparation and again before analysis. Calibration curves were created at the beginning and at the end of the sample analyses. Additionally, blank samples and pooled plasma samples were included in the analytical sequence for quality control purposes. Samples were injected three times, resulting in three technical replicate measurements for each of the 50 samples.

Ultra high-performance liquid chromatography (UHPLC) - Mass spectrometry

The UHPLC-QQQ-MS system is described in detail in **Supplemental Methods**. Briefly, the UHPLC analysis was performed on a Kinetex® F5 column (100 × 2.1 mm, particle size 1.7 µm) with a flow rate of 0.4 mL min^{-1} and an injection volume of 2 µL. H$_2$O + 0.1% HCOOH (A) and

ACN:IPA (2:1, v/v) + 0.1% HCOOH (B) were used as the mobile phases for gradient elution. Both positive and negative ion mode MS- and MS/MS-spectra (scan range *m/z* 40-600) were acquired (**Supplemental Tables S1** and **S2**).

Results

Based on our earlier diabetes-related studies, as well as on the results published in the literature, we selected 36 specific metabolites for this study (**Table 1**) *(2, 4-7, 9, 10, 13, 16-20, 22, 23, 31-33)*. Our aim was to develop a robust and fast analytical assay in terms of both sample preparation and analysis, for quantitative determination of these selected metabolites. However, analyzing both highly polar and nonpolar metabolites in a single method is highly problematic. As some of the candidate biomarkers (*e.g.*, very polar sugar derivatives and neutral lipids such as triacylglycerols) would have required a second sample preparation step and/or analytical method, these were excluded from the final method. The method was validated in terms of (a) limit of detection (LOD), (b) limit of quantitation (LOQ), (c) linearity (R^2) and linear range and (d) intra- and inter-day repeatability of each analyte.

Sample preparation

The main goal in the selection of conditions for sample preparation was the development of a workflow that is simple, robust and feasible to automate, while taking into consideration the LC-MS method as well. Here, we combined a simple protein precipitation with acid followed by derivatization of amino acids and structurally-related compounds. For the protein precipitation, acidic conditions were chosen, as protein precipitation with methanol or acetonitrile would have required evaporation of the solvent prior to derivatization and analysis. The amount of derivatization reagent, the amount and type of the solvent and buffer as well as the time for the derivatization reaction were optimized. Since the derivatization reagent has an impact on the MS detection, the conditions were optimized to decrease ion suppression as well as to improve the overall robustness of the method. Dry ACN was used for dissolving the AQC-reagent, as even trace amounts of water in the solvent can react with the reagent. The final sample preparation conditions included protein precipitation with SSA, followed by neutralization and pH adjustment using a mixture of carbonate buffer and NaOH) prior to the derivatization with AQC in anhydrous

ACN. The MS spectra showed that only amino acids and related compounds with amino acid functionality (namely the amino acids, AADA, ADMA, SDMA, kynurenine and taurine) were derivatized and not any of the other targeted compounds.

LC-MS

MS- and MS/MS-spectra were acquired for each of the analytes in order to select optimal precursor and product ions for selected reaction monitoring (SRM) analyses (**Supplemental Figures S2A-E**). Depending on the ionization properties of the different analytes, protonated ($[M+H]^+$) or deprotonated ($[M-H]^-$) molecules were chosen as precursor ions (**Supplemental Tables S1** and **S2**). MS/MS-spectra were acquired and the most selective and intense product ions were selected for SRM analyses. When possible, one ion transition was chosen for quantification and another ion transition was chosen as the qualifying ion transition to ensure correct measurements of the analytes. Finally, the analysis parameters (fragmentor voltage, collision energy, cell accelerator voltage) were optimized for each ion transition (**Supplemental Tables S1** and **S2**). All the derivatized amino acids and related compounds (**Figure 1**) produced the product ion $[M-H-170]^-$. These were then selected for SRM analyses together with one other diagnostic product ion (where possible). Among the bile acids, CDCA and UDCA were not fragmented and, therefore, the only chosen product ions for these two analytes were their deprotonated molecules. For isomeric compounds (GCDCA, GDCA and GUDCA; TCDCA, TDCA and TUDCA) the MS and MS/MS-spectra are similar to the same three main product ions and their separation depends on chromatographic separation. In addition, TCA shows the same three main product ions as TCDCA, TDCA and TUDCA, but has different precursor ion.

In the optimization of the LC-MS method, different columns (Ascentis Express RP-Amide, Poroshell 120 SB-AQ, Acclaim RSLC PolarAdvantage, Acclaim Trinity P2 and Kinetex® F5 column) and different LC modes were tested. Based on the resolution of the chromatographic separation, the Kinetex® F5 column was chosen for further optimization. The conditions were optimized to include sufficient retention for the most polar compounds. Therefore, the gradient elution was initiated at 99% of the aqueous eluent. The UHPLC method showed good chromato-graphic performance (**Figure 2**), fulfilling general acceptance criteria for an analytical method (**Section 3.3**). For a few of the analytes, the resolution was, however, insufficient to achieve

7

baseline separation and due to very similar MS/MS-spectra these analytes (α(R)-OHB & α(S)-OHB, Leu & Ile, TDCA & TCDCA, GCDCA & GDCA, ADMA & SDMA) were quantified together.

Method validation

The quantitative performance of the developed UHPLC-ESI-MS/MS method was evaluated with respect to (a) limit of detection (LOD), (b) limit of quantitation (LOQ), (c) linearity (R^2) and linear range and (d) intra- and inter-day repeatability (**Table 3**). The LODs (at $S/N \geq 3$) were measured from standard samples and are remarkably different for different analytes, with the lowest LOD being < 2.5 ng mL^{-1} (being the lowest measured concentration) for Ala, AzelA, GCDCA, GDCA, Leu, Ile, N-MNA, Phe and TCA. The highest LOD, on the other hand, was 25000 ng mL^{-1} for α(R)-OHB & α(S)-OHB. These results indicate an acceptable sensitivity within the concentration ranges normally detected in blood samples for most of the analytes, *(34)* except for α(R)-OHB & α(S)-OHB, for which the LODs are clearly higher than the lowest amounts previously detected in blood samples (3120 ng mL^{-1}). Calibration curves and the intra- and inter-day repeatability were determined by using normalized peak areas. For the analytes which were quantified together (*i.e.,* GCDCA and GDCA, ADMA and SDMA, and TDCA and TCDCA), only one ISTD was used. The ISTDs used for GCDCA and GDCA, ADMA and SDMA, and TDCA and TCDCA were GDCA-d6, ADMA-d7 and TCDCA-d9, respectively. Additionally, for a few analytes (*i.e.,* GBB, Crea, α(R)-OHB, α(S)-OHB and β-OHB), the ISTD signal was not repeatable and therefore, the validation parameters of these analytes were measured without normalization to an ISTD. The calibration curves were determined within a concentration range of 2.5-75000 ng mL^{-1}. The linear ranges showed a broad variation between the different analytes (**Table 3**). This variability is, however, acceptable since the biological concentration of each analyte is within the measured linear range. The coefficients of determination (R^2) were higher than 0.97 for all analytes and above 0.99 for most analytes. This shows that the method has both good linearity and quantitation ability for each analyte, with accuracies well within the general requirement of 80-120%.

For the repeatability studies, three standard samples (c= 100 ng mL^{-1}, 1000 ng mL^{-1} and 10000 ng mL^{-})^1 were analyzed in four consecutive runs and in three runs on five consecutive days for intra-day and inter-day repeatability measurements, respectively. Relative standard deviations (%RSD)

8

were calculated for both the intra- and inter-day studies (**Table 3**). The %RSDs for the intra-day repeatability studies were generally below 1.5% and 20.8% for the retention times and normalized peak area ratios, respectively. There are a few exceptions to these results for the analytes with no internal standards (Crea, GBB and β-OHB). The %RSDs for the intra-day and inter-day repeatability for these three analytes was between 17.8 and 20.9% and between 3.5 and 24.5%, respectively.

Feasibility of the method for the analysis of samples from a diabetes cohort

The feasibility of the developed UHPLC-ESI-MS/MS method for the analysis of biological samples was demonstrated by analyzing plasma samples from individuals with diabetes who had a wide range of albuminuria. Albuminuria is a pathological condition where the protein albumin is present in the urine in abnormal amounts. It is a sign of diabetic kidney disease, which often occurs in especially subjects with type 1 diabetes *(28)*. In healthy subjects (normo-albuminuric), only trace amounts of albumin (< 30 mg/24 hours) are present in the urine while subjects with elevated amounts of albumin in the urine, on the other hand, can be classified as either micro-albuminuric (c = 30 – 299 mg/24 hours) or macro-albuminuric (c ≥ 300 mg/24 hours).

In total, 50 samples were selected from a previously-described study cohort of a total of 676 participants *(28)*. The subset was created with computational sampling, aiming at finding a small random subset of the cohort, where the distributions of potentially confounding clinical variables are as similar as possible between the two study groups. This allowed us to study associations between metabolites and albuminuria even in this small sample set whilst avoiding the confounding effects of other factors. The clinical variables assessed were age, antihypertensive medication, BMI, duration of diabetes, glycated hemoglobin (HbA1c), insulin day dose, sex, smoking, systolic blood pressure, total cholesterol and total triglycerides.

Selection of the best random subsample was done in four steps: (1) In total, 1 million n=25+25 sub-samples were drawn with random sampling, (2) the correlation between each clinical variable and the albuminuria group variable was computed for each subsample, (3) the highest absolute value of correlation in each subsample was identified, and (4) the random subsample with the

lowest value of maximum correlation was selected for being the least-confounded random subset for analysis.

Computational selection resulted in a balanced subset of samples from 25 normo-albuminuric and 25 macro-albuminuric participants. The highest Pearson correlation to the albuminuria group variable among the clinical variables was 0.21 for total triglycerides. All other clinical variables had a lower absolute correlation to the group variable, suggesting that the selected small subset was not confounded by imbalance in the clinical characteristics.

Associations between metabolite concentrations and relevant clinical variables were tested with metabolite-specific mixed-effects models using the R-package limma *(35)*. Metabolite concentrations entered the model as the dependent variable, participant identity as the random effect and the following clinical variables as fixed effects: albuminuria group, age, BMI, estimated globular filtration rate (eGFR; kidney function), glycated hemoglobin (HbA1c; glycemic control), sex, systolic blood pressure, total cholesterol, total triglycerides. Significance tests of coefficients were corrected for multiple testing over the metabolites with the Benjamini-Hochberg method *(36)*.

Associations indicated by significant model coefficients (multiple-testing-corrected $p < 0.05$) were visualized as a bipartite network (**Figure 3**) between clinical variables and metabolites with the R-package ggplot2 *(37)*. Strength (log-10-transformed coefficients) and the signs of each association were shown in the width and the color of the line, respectively. Metabolomic associations to albuminuria group and eGFR, which are the key variables of the present study, were highlighted with opaque lines.

The target panel included metabolites which have previously been associated particularly with kidney functions The analysis resulted in estimates of concentrations of the measured metabolites in 50 participants with T1D. Macro-albuminuria, which is an indicator of kidney disease, was associated with elevated GCDCA & GDCA, Tyr and decreased Kynu (**Figure 3**). Estimated globular filtration rate (eGFR; kidney function) , was associated with ADMA & SDMA, Cit, Gln, taurine and Tyr. Glycated hemoglobin (HbA1c; glucose control) was associated with decreased GCDCA & GCDA, Glu and HCit. Smoking was associated with elevated Glu, α-OHB and

10

decreased Gln as well as to a disruption in the balance of the bile acids GCDCA and GDCA. Although no metabolomic associations were found with age or BMI in this small sub-study, several metabolites were associated with sex, statin medication, systolic blood pressure, total cholesterol and total triglycerides. It should, however, be noted that as our target panel is based on reported markers of (pre)diabetes and diabetic complications, and does not cover the entire metabolome, a comprehensive pathway analysis could be biased and not fully reliable. The quantitative results are presented in **Supplemental Table S3**.

Overall, our results agree with the previously-published results. Elevated levels of phenylalanine and of arginine, citrulline and ornithine have been reported in T2D patients with macroalbuminuria and microalbuminuric patients, respectively *(38)*. Additionally, increased levels of plasma homocysteine have been found to be related to macroalbuminuria in diabetic patients *(39)*. ADMA, on the other hand, has been suggested as a candidate biomarker for diabetic kidney complications, whilst elevated levels of ADMA have been shown to predict a more accelerated course of renal function loss and promoted the development of renal damage *(14, 15, 40)*. Our results are also in line with literature data showing that bile acid metabolism is altered in T2D patients *(41)*. Overall, our results demonstrate that the developed analytical method is feasible for performing targeted metabolomic analysis of plasma samples from diabetic patients, and that it can be used for the stratification of diabetic patients.

Conclusions

The method developed here proved to be fast (with a sample analysis time of less than 10 minutes) and robust, and thus suitable for routine analyses in the diabetes clinic. Validation of the method showed that the selected panel of markers can be effectively used for classification of subjects with diabetic complications, such as macro-albuminuria. Further evaluation of the clinical relevance of the method is now merited in order to evaluate the full potential of this diagnostic panel in the stratification of prediabetes, metabolic and diabetic complications.

Acknowledgements

Author contributions

1) Linda Ahonen: method development and acquisition of data, analysis and interpretation of data, writing the manuscript

2) Sirkku Jäntti: design of the analytical work, method development and acquisition of data, critical revision of the manuscript

3) Claudia Risz: acquisition of data, critical revision of the manuscript

4) Tommi Suvitaival: design of the clinical sub-cohort, analysis and interpretation of data, critical revision of the manuscript

5) Simone Theilade: design and establishment of the clinical cohort, critical revision of the manuscript

6) Peter Rossing: design of the clinical cohort, critical revision of the manuscript

7) Risto Kostiainen: design of the analytical work, critical revision of the manuscript

8) Matej Oresic: design of the work, critical revision of the manuscript

9) Tuulia Hyötyläinen: design of the work, development of the methodology, interpretation of data, writing the manuscript

References

1. Tuso P. Prediabetes and lifestyle modification: Time to prevent a preventable disease. The Permanente Journal 2014;18:88-93.

2. Bansal N. Prediabetes diagnosis and treatment: A review. World Journal of Diabetes 2015;6:296-303.

3. Kristine F, Adam H, Thomas PJS. Heterogeneity of pre-diabetes and type 2 diabetes: Implications for prediction, prevention and treatment responsiveness. Current Diabetes Reviews 2016;12:30-41.

4. Tillin T, Hughes AD, Wang Q, Würtz P, Ala-Korpela M, Sattar N, et al. Diabetes risk and amino acid profiles: Cross-sectional and prospective analyses of ethnicity, amino acids and diabetes in a south asian and european cohort from the sabre (southall and brent revisited) study. Diabetologia 2015;58:968-79.

5. Würtz P, Soininen P, Kangas AJ, Rönnemaa T, Lehtimäki T, Kähönen M, et al. Branched-chain and aromatic amino acids are predictors of insulin resistance in young adults. Diabetes Care 2012.

6. Wang TJ, Larson MG, Vasan RS, Cheng S, Rhee EP, McCabe E, et al. Metabolite profiles and the risk of developing diabetes. Nature Medicine 2011;17:448.

7. Stančáková A, Civelek M, Saleem NK, Soininen P, Kangas AJ, Cederberg H, et al. Hyperglycemia and a common variant of gckr are associated with the levels of eight amino acids in 9,369 finnish men. Diabetes 2012.

8. Suvitaival T, Mantere O, Kieseppä T, Mattila I, Pöhö P, Hyötyläinen T, et al. Serum metabolite profile associates with the development of metabolic co-morbidities in first-episode psychosis. Translational Psychiatry 2016;6:e951.

9. Floegel A, Stefan N, Yu Z, Mühlenbruch K, Drogan D, Joost H-G, et al. Identification of serum metabolites associated with risk of type 2 diabetes using a targeted metabolomic approach. Diabetes 2013;62:639-48.

10. Welsh P, Rankin N, Li Q, Mark PB, Würtz P, Ala-Korpela M, et al. Circulating amino acids and the risk of macrovascular, microvascular and mortality outcomes in individuals with type 2 diabetes: Results from the advance trial. Diabetologia 2018;61:1581-91.

11. Tavares G, Venturini G, Padilha K, Zatz R, Pereira AC, Thadhani RI, et al. 1,5-anhydroglucitol predicts ckd progression in macroalbuminuric diabetic kidney disease: Results from non-targeted metabolomics. Metabolomics 2018;14:39.

12. Yu B, Zheng Y, Nettleton JA, Alexander D, Coresh J, Boerwinkle E. Serum metabolomic profiling and incident ckd among african americans. Clin J Am Soc Nephrol 2014;9:1410-7.

13. Rhee SY, Jung ES, Park HM, Jeong SJ, Kim K, Chon S, et al. Plasma glutamine and glutamic acid are potential biomarkers for predicting diabetic retinopathy. Metabolomics 2018;14:89.

14. Lajer M, Tarnow L, Jorsal A, Teerlink T, Parving H-H, Rossing P. Plasma concentration of asymmetric dimethylarginine (adma) predicts cardiovascular morbidity and mortality in type 1 diabetic patients with diabetic nephropathy. 2008;31:747-52.

15. Zobel EH, von Scholten BJ, Reinhard H, Persson F, Teerlink T, Hansen TW, et al. Symmetric and asymmetric dimethylarginine as risk markers of cardiovascular disease, all-cause mortality and deterioration in kidney function in persons with type 2 diabetes and microalbuminuria. 2017;16:88.

16. Yousri NA, Mook-Kanamori DO, Selim MME-D, Takiddin AH, Al-Homsi H, Al-Mahmoud KAS, et al. A systems view of type 2 diabetes-associated metabolic perturbations in saliva, blood and urine at different timescales of glycaemic control. Diabetologia 2015;58:1855-67.

17. Gall WE, Beebe K, Lawton KA, Adam K-P, Mitchell MW, Nakhle PJ, et al. A-hydroxybutyrate is an early biomarker of insulin resistance and glucose intolerance in a nondiabetic population. PLoS ONE 2010;5:e10883.

18. Suvitaival T, Bondia-Pons I, Yetukuri L, Poho P, Nolan JJ, Hyotylainen T, et al. Lipidome as a predictive tool in progression to type 2 diabetes in finnish men. Metabolism 2018;78:1-12.

19. Wang-Sattler R, Yu Z, Herder C, Messias AC, Floegel A, He Y, et al. Novel biomarkers for pre-diabetes identified by metabolomics. Molecular Systems Biology 2012;8.

20. Carter TC, Rein D, Padberg I, Peter E, Rennefahrt U, David DE, et al. Validation of a metabolite panel for early diagnosis of type 2 diabetes. Metabolism 2016;65:1399-408.

21. Haeusler RA, Astiarraga B, Camastra S, Accili D, Ferrannini E. Human insulin resistance is associated with increased plasma levels of 12α-hydroxylated bile acids. Diabetes 2013;62:4184-91.

22. Sone H, Shimano H, Ebinuma H, Takahashi A, Yano Y, Iida KT, et al. Physiological changes in circulating mannose levels in normal, glucose-intolerant, and diabetic subjects. Metabolism 2003;52:1019-27.

23. Wang TJ, Ngo D, Psychogios N, Dejam A, Larson MG, Vasan RS, et al. 2-aminoadipic acid is a biomarker for diabetes risk. The Journal of Clinical Investigation 2013;123:4309-17.

24. Klein MS, Shearer J. Metabolomics and type 2 diabetes: Translating basic research into clinical application. Journal of Diabetes Research 2016;2016:10.

25. Kim SH, Lee KA, Jin HY, Baek HS, Park TS. The relationship between anemia and the initiation of dialysis in patients with type 2 diabetic nephropathy. Diabetes Metab J 2015;39:240-6.

26. Barrios C, Beaumont M, Pallister T, Villar J, Goodrich JK, Clark A, et al. Gut-microbiota-metabolite axis in early renal function decline. PLOS ONE 2015;10:e0134311.

27. Theilade S, Lajer M, Persson F, Joergensen C, Rossing P. Arterial stiffness is associated with cardiovascular, renal, retinal, and autonomic disease in type 1 diabetes. 2013;36:715-21.

28. Theilade S, Hansen TW, Goetze JP, Rossing P. Increased plasma concentrations of midregional proatrial natriuretic peptide is associated with risk of cardiorenal dysfunction in type 1 diabetes. American Journal of Hypertension 2015;28:772-9.

29. MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, et al. Skyline: An open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics (Oxford, England) 2010;26:966-8.

30. Team RC. R: A language and environment for statistical computing, r foundation for statistical computing. Vienna, Austria 2017.

31. Du M-R, Yan L, Li N-S, Wang Y-J, Zhou T, Jiang J-L. Asymmetric dimethylarginine contributes to retinal neovascularization of diabetic retinopathy through ephrinb2 pathway. Vascular Pharmacology 2018;108:46-56.

32. Haeusler RA, Astiarraga B, Camastra S, Accili D, Ferrannini E. Human insulin resistance is associated with increased plasma levels of 12a-hydroxylated bile acids. Diabetes 2013;62:4184-91.

33. Larsen MB, Henriksen JE, Grauslund J, Peto T. Prevalence and risk factors for diabetic retinopathy in 17 152 patients from the island of funen, denmark. Acta Ophthalmologica 2017;95:778-86.

34. Wishart DS, Feunang YD, Marcu A, Guo AC, Liang K, Vázquez-Fresno R, et al. Hmdb 4.0: The human metabolome database for 2018. Nucleic acids research 2018;46:D608-D17.

35. Ritchie ME, Smyth GK, Phipson B, Wu D, Hu Y, Shi W, Law CW. Limma powers differential expression analyses for rna-sequencing and microarray studies. Nucleic Acids Research 2015;43:e47-e.

36. Hochberg YBaY. Controlling the false discovery rate: A practical and powerful approach to multipletesting. Journal of the Royal Statistical Society Series B (Methodological) 1995;57:12.

37. Wickham H. Ggplot2: Elegant graphics for data analysis. Springer-Verlag 2016.

38. Chuang W-H, Arundhathi A, Lu C, Chen C-C, Wu W-C, Susanto H, et al. Altered plasma acylcarnitine and amino acid profiles in type 2 diabetic kidney disease. 2016;12:108.

39. Mao S, Xiang W, Huang S, Zhang A. Association between homocysteine status and the risk of nephropathy in type 2 diabetes mellitus. Clinica Chimica Acta 2014;431:206-10.

40. Fliser D, Kronenberg F, Kielstein JT, Morath C, Bode-Böger SM, Haller H, Ritz E. Asymmetric dimethylarginine and progression of chronic kidney disease: The mild to moderate kidney disease study. 2005;16:2456-61.

41. Prawitt J, Caron S, Staels BJCDR. Bile acid metabolism and the pathogenesis of type 2 diabetes. 2011;11:160.

Table 1. Standard compounds acquired for quality control and for quantitation.

Compound	Abbreviation	Group	Vendor	Solvent, stock solution
L-Glutamine	Gln	Amino acids + related metabolites	Sigma-Aldrich	H_2O
Glycine	Gly			0.1 M HCl
L-Alanine	Ala			
L-Leucine	Leu			
L-Isoleucine	Ile			
L-Phenylalanine	Phe			
L-Tryptophan	Trp			
L-Tyrosine	Tyr			
L-Glutamic Acid	Glu			
L-Citrulline	Cit			
L-Homocitrulline	HCit		SCB	
Asymmetric dimethylarginine	ADMA			
Symmetric dimethylarginine	SDMA			
DL-2-Aminoadipic Acid	AADA		Sigma-Aldrich	
L-Kynurenine	Kynu			
Taurine	Taurine			
Deoxycholic Acid	DCA	Bile acids	Sigma-Aldrich	MeOH
Glychochenodeoxycholic Acid	GCDCA			
Glycodeoxycholic Acid	GDCA			
Glycocholic Acid	GCA			
Taurodeoxycholic Acid	TDCA			
Taurochenodeoxycholic Acid	TCDCA			
Deoxychenocholic Acid	CDCA			
Cholic Acid	CA			
Taurocholic Acid	TCA			
Glycoursodeoxycholic Acid	GUDCA		CIL	
Ursodeoxycholic Acid	UDCA			
Tauroursodeoxycholic Acid	TUDCA			
Creatinine	Crea	Other metabolites	Sigma-Aldrich	10% MeOH
Indoxyl Sulfate	IndS			
N-methyl-nicotinamide	N-MNA		SCB	
Gamma-butyrobetaine	GBB			
Azelaic Acid	AzelA	Small organic acids	Sigma-Aldrich	MeOH
L-3-hydroxybutyric Acid	β-OHB			10% MeOH
R-2-hydroxybutyric Acid	α(R)-OHB			
S-2-hydroxybutyric Acid	α(S)-OHB			

Vendor information:

Sigma-Aldrich (Steinheim, Germany)

SCB: Santa Cruz Biotechnology, Inc. (Dallas, TX, USA)

CIL: Cambridge Isotope Laboratories Inc. (Tewksbury, MA, USA)

Table 2. Internal standards, with concentrations in ISTD MIX, acquired for quality control and for quantitation.

Internal standard	Abbreviation	Group	Vendor	Solvent, stock solution	Concentration in ISTD MIX (ng mL^{-1})
d5-Glutamine	d5-Gln	Amino acids + related metabolites	CIL	H$_2$O	30 000
d10-L-Leucine	d10-Leu		CDN	0.1 M HCl	5 000
^{2}H$_4$-L-Homocitrulline	2H4-HCit		Alsachim		
Glycine-1-^{13}C,2,2-d$_2$	13C, d2-Gly		Sigma-Aldrich		
d4-DL-Alanine	d4-Ala				
d5-L-Glutamic Acid	d5-Glu				
d10-Isoleucine	d10-Ile		CIL		
d5-L-Phenylalanine	d5-Phe				500
d8-Tryptophan	d8-Trp				5 000
d7-Tyrosine	d7-Tyr				
d4-Citrulline	d4-Cit				500
d3-L-2-Aminoadipic Acid	d3-AADA				10 000
d7-Asymmetric dimethylarginine	d7-ADMA				5 000
^{13}C$_6$-Kynurenine	13C6-Kynu		Alsachim		30 000
d4-Taurine	d4-Taurine				500
d4-Deoxycholic Acid	d4-DCA	Bile acids	CDN	MeOH	500
d4-Glycocholic Acid	d4-GCA				250
d4-Deoxychenocholic Acid	d4-CDCA				500
d4-Glycoursodeoxycholic Acid	d4-GUDCA				5 000
d4-Cholic Acid	d4-CA				500
d4-Ursodeoxycholic Acid	d4-UDCA				250
d4-Glychochenodeoxycholic Acid	d4-GCDCA		CIL		5 000
d6-Glycodeoxycholic Acid	d6-GDCA				30 000
d9-Taurochenodeoxycholic Acid	d9-TCDCA				500
d4-Taurocholic Acid	d4-TCA				
d4-Tauroursodeoxycholic Acid	d4-TUDCA				250
d5-Creatinine	d5-Crea	Polar metabolites	CDN	10% MeOH	10 000
d4-N-methyl-nicotinamide	d4-N-MNA				250
d9-Gamma-butyrobetaine	d9-GBB				500
d4-Indoxyl Sulfate	d4-IndS		Sigma-Aldrich		5 000
d14-Azelaic Acid	d14-AzelA	Small organic acids	CDN	MeOH	5 000
d4-3-Hydroxybutyric Acid	d4-β-OHB			10% MeOH	100 000

d3-2-Hydroxybutyric Acid	d3-α-OHB				

Vendor information:

Sigma-Aldrich (Steinheim, Germany)

CDN: C/D/N Isotopes, Inc. (Quebec, Canada)

CIL: Cambridge Isotope Laboratories Inc. (Tewksbury, MA, USA)

Alsachim (Illkirch Graffenstaden, France)

SCB: Santa Cruz Biotechnology, Inc. (Dallas, TX, USA)

Table 3. Limit of detection (LOD), limit of quantitation (LOQ), linearity (R^2), linear range, intra- and inter-day repeatability.

Compound	Linearity (R^2) Range (ng mL^{-1})	LOD (ng/ml^{-1})	%RSD_Rt, intra-day	%RSD_Area, Intra-day ($N = 4$)			%RSD_Rt, inter-day	%RSD_Area, Inter-day ($N = 15$)		
				100 ng mL^{-1}	1000 ng mL^{-1}	10000 ng mL^{-1}		100 ng mL^{-1}	1000 ng mL^{-1}	10000 ng mL^{-1}
α(R)-OHB & α(S)-OHB	0.995 25 000 – 75 000	25 000	-	-	-	-	-	-	-	-
AADA	0.984 5 000 – 75 000	500	0.2 ($N = 4$)	-	-	9.1	0.1 ($N = 15$)	-	-	8.7
ADMA & SDMA	0.992 2 500 – 50 000	500	0.2 ($N = 8$)	-	5.6	0.8	0.2 ($N = 30$)	-	8.5	4.2
Ala	0.996 500 – 50 000	< 2.5	0.2 ($N = 8$)	-	4.5	3.0	0.1 ($N = 30$)	-	9.8	13.6
AzelA	0.995 500 – 10 000	< 2.5	0.5 ($N = 8$)	-	11.4	3.9		-	15.8	8.4
β-OHB	0.970 2 500 – 75 000	75	0.6 ($N = 4$)	-	-	20.9	1.2 ($N = 15$)	-	-	24.5
CA	0.996 10 – 10 000	7.5	0.2 ($N = 12$)	2.4	3.1	5.2	0.7 ($N = 45$)	20.8	18.1	20.2
CDCA	0.999 25 – 2 500	7.5	0.2 ($N = 8$)	4.0	4.9	-	0.2 ($N = 45$)	4.3	5.1	14.2
Cit	0.984 500 – 10 000	250	0.2 ($N = 8$)	-	7.7	6.7	0.2 ($N = 30$)	-	9.1	8.3
Crea	0.973 250 – 7 500	25	0.8 ($N = 4$)	-	17.8	-	0.0 ($N = 15$)	-	3.5	-
DCA	0.996 5 – 2 500	2.5	0.2 ($N = 8$)	5.8	6.1	-	0.3 ($N = 30$)	4.3	8.3	-
GBB	0.974 250 – 10 000	50	0.5 ($N = 8$)	-	18.7	15.9	1.5 ($N = 30$)	-	27.3	28.5
GCA	0.997 50 – 25 000	25	0.1 ($N = 12$)	4.6	4.2	4.2	0.4 ($N = 45$)	6.8	5.1	7.3
GCDCA & GDCA	0.997 25 – 2 500	< 2.5	0.2 ($N = 8$)	1.9	4.2	-	0.5 ($N = 30$)	16.4	16.1	-
Gln	0.987 750 – 50 000	5	0.1 ($N = 8$)	-	5.0	7.7	0.5 ($N = 30$)	-	10.5	11.5
Glu	0.990 750 – 75 000	500	0.2 ($N = 8$)	-	13.9	10.7	0.3 ($N = 30$)	-	10.9	5.2

Compound	Linearity (R^2) Range (ng mL^{-1})	LOD (ng/ml^{-1})	%RSD_Rt, intra-day	%RSD_Area, Intra-day ($N = 4$)			%RSD_Rt, inter-day	%RSD_Area, Inter-day ($N = 15$)		
				100 ng mL^{-1}	1000 ng mL^{-1}	10000 ng mL^{-1}		100 ng mL^{-1}	1000 ng mL^{-1}	10000 ng mL^{-1}
Gly	0.993 7 500 – 75 000	1 000	0.03 ($N = 4$)	-	-	16.2	0.6 ($N = 15$)	-	-	19.6
GUDCA	0.994 75 – 10 000	25	0.1 ($N = 12$)	5.0	9.0	10.6	0.3 ($N = 45$)	13.1	10.9	6.2
HCit	0.995 500 – 25 000	250	0.2 ($N = 8$)	-	8.3	2.6	0.5 ($N = 30$)	-	11.1	16.4
IndS	0.986 5 000 – 75 000	750	0.3 ($N = 4$)	-	-	11.3	0.3 ($N = 15$)	-	-	15.4
Kynu	0.993 500 – 75 000	250	0.2 ($N = 8$)	-	11.2	7.4	0.4 ($N = 30$)	-	7.7	4.4
Leu & Ile	0.997 25 – 75 000	< 2.5	0.4 ($N = 12$)	4.6	4.3	1.5	0.5 ($N = 45$)	13.0	14.0	5.7
N-MNA	0.998 25 – 10 000	< 2.5	0.5 ($N = 12$)	1.6	6.4	3.7	1.0 ($N = 45$)	20.1	18.5	6.5
Phe	0.995 250 – 25 000	< 2.5	0.4 ($N = 0.4$)	-	5.9	6.6	0.4 ($N = 30$)	-	10.1	4.6
Taurine	0.994 250 – 25 000	10	0.2 ($N = 8$)	-	8.3	5.7	0.5 ($N = 30$)	-	8.4	8.7
TCA	0.983 2 500 – 25 000	< 2.5	0.1 ($N = 4$)	-	-	4.5	0.3 ($N = 15$)	-	-	15.5
TDCA & TCDCA	0.984 1 000 – 25 000	10	0.7 ($N = 8$)	-	0.4	5.7	0.7 ($N = 30$)	-	2.6	4.2
Trp	0.996 25 – 25 000	25	0.4 ($N = 12$)	9.0	2.9	4.7	0.5 ($N = 45$)	18.8	5.4	5.3
TUDCA	0.990 250 – 10 000	10	0.1 ($N = 8$)	-	5.5	4.5	0.7 ($N = 30$)	-	1.8	3.0
Tyr	0.992 50 – 75 000	25	0.2 ($N = 12$)	10.3	9.1	4.4	0.3 ($N = 45$)	5.7	8.4	3.4
UDCA	0.991 50 – 50 000	25	0.2 ($N = 12$)	1.5	3.5	3.3	0.2 ($N = 45$)	3.5	10.3	5.6

Figure captions

Figure 1. Derivatization reaction of amino acids and amino acid related compounds.

Figure 2. Chromatograms representing the chromatographic separation of the analytes. The peak numbers correspond to the following analytes: (1) Crea, (2) GBB, (3) α(R)-OHB & α(S)-OHB, (4) β-OHB, (5) N-MNA, (6) Kynu, (7) Leu & Ile, (8) Phe, (9) AzelA, (10) Trp, (11) TUDCA, (12) TCA, (13) GCA, (14) GUDCA, (15) TDCA & TCDCA, (16) CA, (17) CDCA, (18) GCDCA & GDCA, (19) UDCA, (20) DCA, (21) Gly, (22) Gln, (23) ADMA & SDMA, (24) Taurine, (25) Phe, (26) Gln, (27) HCit, (28) Ala, (29) AADA, (30) IndS and (31) Tyr.

Figure 3. Associations between clinical measurements (left) and the quantified analytes (right) in the T1D cohort. The lines indicate statistical associations (red: positive association and blue: inverse/opposite association; line width: strength of the association). Associations directly related to diabetic kidney disease are highlighted with opaque lines.
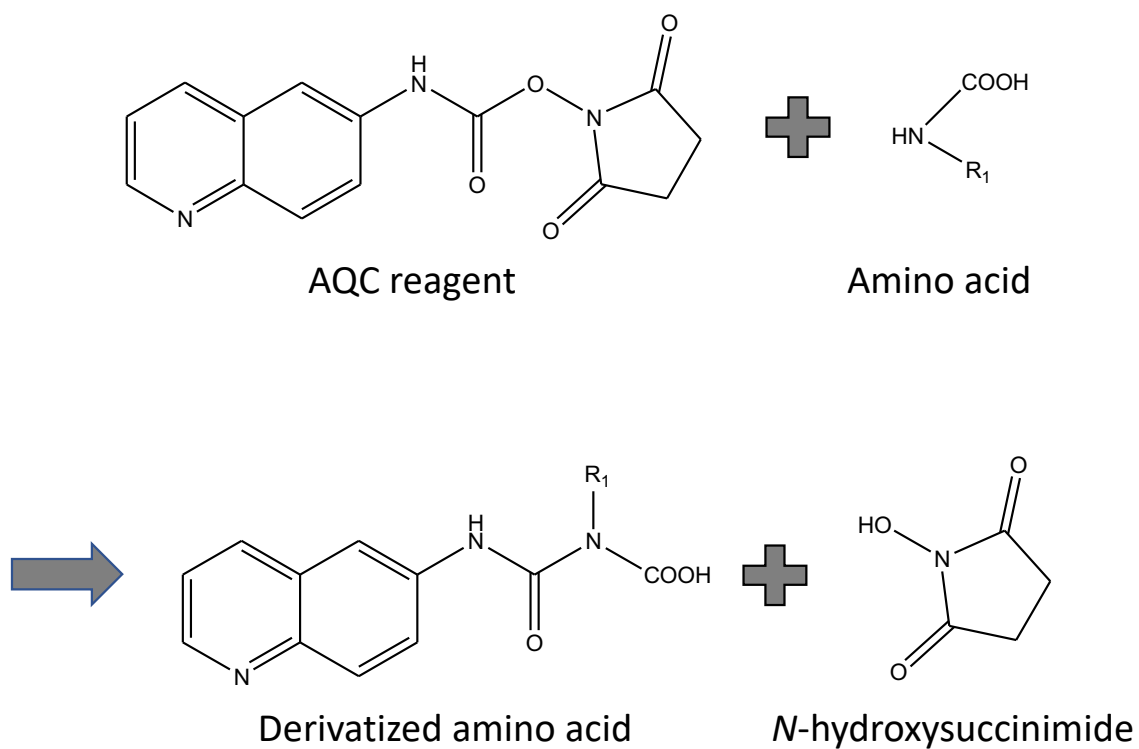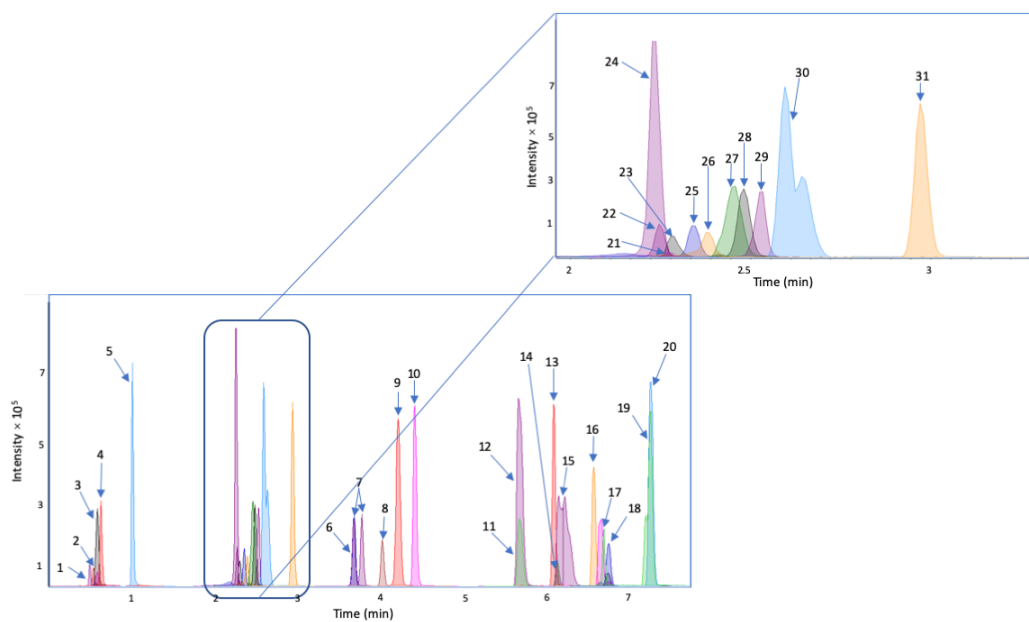
FIGURE 1

AQC reagent   Amino acid

Derivatized amino acid   *N*-hydroxysuccinimide

FIGURE 2

FIGURE 3