Clinica Chimica Acta 493 (2019) 112–122

Contents lists available at ScienceDirect

Clinica Chimica Acta

journal homepage: www.elsevier.com/locate/cca

Metabolomics profiling and pathway analysis of human plasma and urine reveal further insights into the multifactorial nature of coronary artery disease

Arwa M. Amin[a,b,*,1], Hamza Mostafa[a,*,1], Nor Hayati Arif[c], Muhamad Ali SK Abdul Kader[a,d], Yuen Kah Hay[a]

[a] School of Pharmaceutical Sciences, Universiti Sains Malaysia, Penang, Malaysia
[b] Department of Clinical and Hospital Pharmacy, College of Pharmacy, Taibah University, Al Madinah Al Munawwarah, Saudi Arabia
[c] Psychiatry Department, Hospital Pulau Pinang, Malaysia
[d] Cardiology Department, Hospital Pulau Pinang, Penang, Malaysia

ARTICLE INFO

Keywords:
Coronary artery disease
^{1}H NMR
Metabolomics
Systems biology
Single nucleotide polymorphism
Gut microbiota

ABSTRACT

Background: Coronary artery disease (CAD) claims lives yearly. Nuclear magnetic resonance (^{1}H NMR) metabolomics analysis is efficient in identifying metabolic biomarkers which lend credence to diagnosis. We aimed to identify CAD metabotypes and its implicated pathways using ^{1}H NMR analysis.

Methods: We analysed plasma and urine samples of 50 stable CAD patients and 50 healthy controls using ^{1}H NMR. Orthogonal partial least square discriminant analysis (OPLS-DA) followed by multivariate logistic regression (MVLR) models were developed to indicate the discriminating metabotypes. Metabolic pathway analysis was performed to identify the implicated pathways.

Results: Both plasma and urine OPLS-DA models had specificity, sensitivity and accuracy of 100%, 96% and 98%, respectively. Plasma MVLR model had specificity, sensitivity, accuracy and AUROC of 92%, 86%, 89% and 0.96, respectively. The MVLR model of urine had specificity, sensitivity, accuracy and AUROC of 90%, 80%, 85% and 0.92, respectively. 35 and 12 metabolites were identified in plasma and urine metabotypes, respectively. Metabolic pathway analysis revealed that urea cycle, aminoacyl-tRNA biosynthesis and synthesis and degradation of ketone bodies pathways were significantly disturbed in plasma, while methylhistidine metabolism and galactose metabolism pathways were significantly disturbed in urine. The enrichment over representation analysis against SNPs-associated-metabolite sets library revealed that 85 SNPs were significantly enriched in plasma metabotype.

Conclusions: Cardiometabolic diseases, dysbiotic gut-microbiota and genetic variabilities are largely implicated in the pathogenesis of CAD.

1. Introduction

Coronary artery disease (CAD) is one of the leading causes of death globally [1]. The pathogenesis of CAD includes plaque formation in one of the coronary arteries which if enlarged may rupture and form thrombus leading to narrowing of the artery and consequently ischemia [2,3]. CAD patients tend to have high platelets reactivity, circulating active platelets and formation of monocyte-platelet aggregates [4]. There are multiple risk factors which may accelerate CAD development such as hypertension, dyslipidaemia, presence of pro-inflammatory cytokines and the formation of hyperglycaemia associated advanced glycation end products [3,5]. Early diagnosis of CAD will prevent disease progression and ensure promising management outcomes. Furthermore, identifying subjects who are at high risk or in the stage of developing CAD will help to prevent the disease.

Disease metabolic biomarkers usually lend credence to diagnosis and help to guide therapy [6]. They may also indicate and explain the root cause of the disease. Metabolomics analysis of non-invasive or

* Corresponding authors at: Department of Clinical and Hospital Pharmacy, College of Pharmacy, Taibah University, Al Madinah Al Munawwarah, Saudi Arabia; School of Pharmaceutical Sciences, Universiti Sains Malaysia, 11800, Penang, Malaysia.
E-mail addresses: arwamostafa@taibahu.edu.sa, arwa.amin.m@gmail.com (A.M. Amin), hmam12_pha084@student.usm.my, hamza.mostafa@hotmail.com (H. Mostafa), nyati39@yahoo.com (N.H. Arif), mdali_sheikh@hotmail.com (M.A.S. Abdul Kader), khyuen@usm.my (Y. Kah Hay).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.cca.2019.02.030
Received 9 January 2019; Received in revised form 25 February 2019; Accepted 27 February 2019
Available online 28 February 2019
0009-8981/

minimally invasive samples such as plasma and urine has been proven to be an efficient tool to identify metabolic biomarkers of diseases and drugs response classes [7,8]. In metabolomics analysis, researchers use spectroscopic techniques such as nuclear magnetic resonance (^{1}H NMR) and mass spectroscopy (MS) to identify a metabotype (metabolic fingerprint/metabolic biomarkers) which can phenotype disease or drug response [8]. By indicating the metabotype and further investigating the associated metabolic pathways, better understanding of disease development and progression can be attained. We previously identified metabotypes of clopidogrel high on treatment platelets reactivity (HTPR) in plasma and urine among CAD patients who were planned for elective interventional angiographic procedure (IAP) [9,10]. Metabolic pathway analysis and disease wise classification of the identified metabolites showed that insulin resistance, glucose intolerance, obesity, gut microbiota disturbances, CAD progression and heart failure were associated with clopidogrel HTPR. Furthermore, we found that the *CYP2C19*2* homozygous mutation was associated with clopidogrel HTPR, however, this genotype turned out not to be the perfect predictor of clopidogrel HTPR [11]. While the discovered metabotypes were indicatives of high platelets reactivity as measured by adenosine diphosphate (ADP) stimulation of the P2Y12 receptors, it is not completely understood whether these metabotypes are due to CAD development and progression. Therefore, we sought identifying CAD metabotypes and the associated metabolic pathways among the same patients using discriminant analysis of their samples against the samples of healthy control subjects. Towards that end, we applied metabolic phenotyping and pathway analysis on 50 CAD patients and 50 healthy peers from another metabolomics study which was conducted at the same research lab [9,12]. Our aim was identifying metabotypes and metabolic pathways associated with stable CAD in plasma and urine using ^{1}H NMR metabolic profiling.

2. Methods

This is a sub-analysis study which included the ^{1}H NMR metabolic profiling spectra of 50 stable CAD patients from the clopidogrel HTPR study and 50 healthy controls from the alcohol dependence (AD) study [9,10,12,13]. Subjects recruitment and samples collection for both studies were performed at Hospital Pulau Pinang – Malaysia [9,10,12,13]. Samples preparation and analysis for both studies were conducted at the same lab in Universiti Sains Malaysia using the same metabolomics analysis procedures and conditions. Both studies were approved by the Medical Research and Ethics Committee (MREC) of Ministry of Health Malaysia. All the patients and healthy controls who agreed to participate have signed a consent form prior to sampling. The 50 CAD patients who were selected from clopidogrel HTPR study [10] were CAD patients (aged between 28 and 60 y), diagnosed on previous angiogram, on low dose aspirin, not taking any thienopyridine for at least two weeks before recruitment, having platelets count $>70 \times 10^9$/L, not suffering from current acute coronary syndrome (ACS) and not having a history of stroke in the past three months, un-coalesced peptic ulcer, liver cirrhosis and stage five chronic kidney disease (CKD) based on the National Kidney Foundation, Kidney Disease Quality Outcome Initiative (KDQOI) definition [14]. The 50 controls were Non-alcohol dependent subjects chosen from the AD metabolomics study [13]. The selection criteria were healthy subjects aged between 28 and 60 years old, not suffering from any chronic illnesses, infectious diseases, or taking any medications during or one week prior to sampling. Urine samples were obtained and stored immediately in −80 °C freezer. Blood samples were collected in BD vacutainer® EDTA (Becton Dickinson (BD) vacutainer® Systems). Blood was centrifuged for 10 min at 12000 rpm and 4 °C (MIKRO® 22 R, Hettich Zentrifugen) to separate plasma which was immediately stored in −80 °C freezer.

Plasma and urine samples' preparation and ^{1}H NMR metabolic profiling were performed as previously described protocol [9,10,12,13]. In brief, after thawing, plasma samples were centrifuged at 12,000 rpm and 4 °C for 5 min (MIKRO 22 R, HettichZentrifugen®). Three hundred microliters of the centrifuged plasma were mixed with 300 μL phosphate buffer (pH 7.4). Of the final mixture, 550 μL were transferred to 5 mm NMR tube (BRUKER®, Switzerland). Similarly, urine samples were thawed then centrifuged for 5 min at 4 °C and 12,000 rpm (MIKRO 22 R, HettichZentrifugen®). 400 μL of the centrifuged urine were mixed with 200 μL of phosphate buffer. Five hundred fifty microliters of the resultant mixture was transferred to 5 mm NMR tube (BRUKER®).

Prepared samples were analysed using ^{1}H NMR Bruker 500 MHz (Bruker® BioSpin). TopSpin 3.2 (BRUKER®, BioSpin) software was used for the the acquisition setups. Baseline correction, shimming and locking, were done manually. Water peak suppression was applied to prevent its domination over low concentration metabolites [15]. All the samples were run at 300 K temperature. For plasma samples, the Nuclear Over-hauser enhancement spectroscopy one dimensional (NOESY-Presat-1-D), the Carr Purcell Meiboom Gill one dimensional (CPMG-1D) and the 2-dimensional (2-D) *J*-resolved (*J*-RES) experimental settings were used for ^{1}H NMR analysis [9,12]. For urine samples, the NOESY-Presat-1-D and the *J*-RES experimental settings were used for ^{1}H NMR analysis [10,13]. To confirm the identities of the metabolites in the overlapped spectral regions, we run 2-D Heteronuclear Single Quantum Coherence (HSQC) experiment on selective samples from each group.

The ^{1}H NMR CPMG-1-D spectra and the NOESY-Presat-1-D spectra were bucketed (binned) to bin's width of 0.04 ppm by the Analysis of Mixtures (AMIX) software (ver 3.9.13, BRUKER®, BioSpin) to get bucket tables of bins' intensities (Microsoft Excel) [16]. The resultant bucket tables included bins' columns labelled by the central chemical shift (CCS) chemical shift (ppm) of each bin and rows of subjects' samples (cases). The water regions (plasma: 4.65–4.88 ppm and urine: 4.72–4.80 ppm) were excluded from each table. In addition, EDTA peaks were excluded from plasma spectra (free EDTA: 2 singlets at 3.18 ppm and 3.59 ppm; Ca-EDTA: singlet at 2.54 and AB pattern at 3.10 ppm; Mg-EDTA: singlet at 2.68 ppm and AB pattern at 3.20 ppm). Negative values in the resultant tables were converted to zero. Bins were excluded from the bucket tables if they had zero values in >50% of the cases.

Subjects' demographics data were uploaded to statistical software for social sciences (SPSS), version 22, IBM. Mean and standard deviation (SD) were used to describe continuous variables, while frequencies were used for categorical variables. Statistical analysis of the processed ^{1}H NMR spectra was done using unsupervised and supervised multivariate analysis to indicate the discriminating metabolites between CAD patients and controls. This was achieved by developing principle component analysis (PCA) and preliminary orthogonal partial least square discriminant analysis models (OPLS-DA) followed by factor analysis (FA) and multivariate logistic regression (MVLR) to indicate the discriminating metabotypes. The area under receiver operating characteristic (AUROC) was used to evaluate the OPLS-DA and MVLR models [17]. The bucket tables were uploaded to the Soft Independent Modelling of Class Analogy software (SIMCA® 13.0.3, Umetrics®). The pareto scaling was used to scale the data. The unsupervised PCA and the hotelling's plot were used to indicate the intrinsic outliers. After that, the supervised OPLS-DA was run to indicate the best discriminating models. Sensitivity, specificity, accuracy and AUROC which were generated through the SIMCA misclassification tool were used to evaluate each model. The most influencing bins in the models were identified by the variable influence on projection (VIP) plot. Bins with the highest VIP value have the major influence in the model. A cut-off of VIP value >1 was used to select the most influential bins [10]. To further reduce the bins to the most important ones, bins with VIP value >1 were run in univariate logistic regression analysis (ULR) using SPSS, (ver 22) software. Bins with $p < .1$ in the ULR were selected for FA run, followed by MVLR. The $p < .05$ was used for the significance level in the MVLR. Like OPLS-DA models, the sensitivity, specificity, accuracy and AUROC were used to evaluate the final MVLR models. The metabolites

identification was done using representative plasma and urine ^{1}H NMR spectra from each group (CAD and controls). Chenomx 8.2 (Chenomx® Inc.), Human Metabolome Database (HMDB) version 3.6 and Bruker Biofluid Reference Compound Database (B-Biorefcode, Bruker® BioSpin) libraries were used to identify the metabolites.

To investigate the presence of biologically meaningful patterns among CAD metabotypes, the metabolic pathway analysis was performed using the functional enrichment and the pathway analysis tool of the free web-based software MetaboAnalyst 4.0 [18]. The tool through its updated metabolite set libraries allows for indicating metabolic pathways, single nucleotide polymorphisms (SNPs), diseases and locations associated-metabolite sets in the discovered metabotypes. We performed enrichment over representation analysis (ORA) for the list of metabolites which were identified from bins with $p < .1$ in the ULR analysis. Furthermore, to get further insights on the perturbed metabolic pathways, we also performed a pathway topology analysis using the identified metabolites from the MVLR models. The p value and the false discovery rate (FDR) of each perturbed pathway were used to account for significance and false discovery in the pathway analysis, respectively. In this regard, the thresholds of $p < .05$ and FDR < 0.1 were used to consider the pathway is impactful in the discovered metabotypes. Fig. S1 in the online supplementary data demonstrates summary of the steps which were followed to indicate the discriminating multivariate models and the metabolic pathways analysis (see online supplementary material).

3. Results

The mean (SD) age of the CAD group was 56 (6.3), while the mean age of the controls was 39.8 (9.4) y (SD: 8.1683). Males were 44 (88%) and % 32 (64%) of the CAD and controls, respectively. Of all CAD patients, 22 (44%) and 8 (16%) were suffering from type 2 diabetes mellites (T2DM) and stage 3 CKD, respectively. All CAD patients were on aspirin and statins while 41 (82%), 34 (68%) and 6 (12%) were on one of the Beta Blockers (βBs), angiotensin converting enzyme inhibitors (ACEI) and angiotensin II reuptake blockers (ARBs), respectively. The detailed demographics and clinical data of the study's subjects are presented in Table 1. The medications of CAD patients are detailed in Table S.1 (see online supplementary material).

In plasma, the OPLS-DA model discriminated between CAD patients and healthy controls with high accuracy (98%) and AUROC (0.994) (Table 2 and Fig. 1). All the 65 ^{1}H NMR spectral bins with VIP > 1 in this model had strong association with CAD in the ULR ($p < .02$), (see table S.2 in the online supplementary material). From the 65 significant bins in the ULR, 35 metabolites were identified (Table 3). To reduce the number of the bins prior to running MVLR, we applied FA. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy for the FA was >0.6 and the Bartlett's Test of Sphericity had $p < .05$. Upon applying the FA for the 65 bins, eight factors which had eigenvalue >1 explained 93.27% of the total variance. Of these factors, 4 factors (2, 3, 6 & 7) were significantly associated with CAD in the MVLR (see tables S.3 & S.4 in the online supplementary material). The plasma MVLR model had high accuracy (89%) and AUROC (0.96), albeit slightly lower than the OPLS-DA (Table 2). Of the significantly associated factors in the MVLR, factor 3 contained 14 bins corresponding to D-Glucose peaks in the ^{1}H NMR spectra.

In urine, the OPLS-DA model discriminated between CAD patients and healthy controls with high accuracy (98%) and AUROC (0.9996) (Table 2 and Fig. 1). From the 37 bins which had VIP > 1, 19 bins had $p < .1$ in the ULR (see table S.5 in the online supplementary material). From the 19 bins, 12 metabolites were identified (Table 3). We run FA to reduce the number of the bins prior to running the MVLR. The KMO measure of sampling adequacy was >0.6 and the Bartlett's Test of Sphericity had $p < .05$. Four factors had eigenvalues >1 and explained 94.8% of the total variance. Of them, three factors (1, 2 & 3) were significantly associated with CAD in the MVLR (see Tables S.6 & S.7 in the online supplementary material). The accuracy and the AUROC of the MVLR model were 89% and 0.92, respectively. Table 2 details the specificities, sensitivities, accuracies and AUROCs of the OPLS-DA and the MVLR metabotype models. Table 3 presents the list of metabolites which were identified from the ULR significant bins in plasma and urine. Fig. 1 depicts the score scatter plots of the OPLS-DA models of plasma and urine CAD metabotypes.

Table 1
Subjects' demographics and clinical data.

Demographics	CAD Patients No. (%)	Controls No. (%)
Age mean (SD)	56 (6.3)	39.8 (9.4)
Sex:		
Male	44 (88%)	32 (64%)
Female	6 (12%)	18 (36%)
Ethnicity:		
Malay	18 (36%)	24 (48%)
Chinese	15 (30%)	11 (22%)
Indian	17 (34%)	15 (30%)
Smoking:		
Smokers	19 (38%)	15 (30%)
Past-smokers	16 (32%)	4 (8%)
Non-smokers	14 (28%)	31 (62%)
BMI mean (SD)	27.35 (4.17)	22.48 (1.84)
Obesity		
Healthy (18.5–24.9)	16 (32%)	49 (98%)
Over weight (25–29.9)	23 (46%)	1 (2%)
Obese (≥30)	11 (22%)	0
Type 2 DM:		—
Type 2 DM	22 (44%)	
Non-DM	28 (56%)	
Serum Cr/CrCl[a] (total No. 49):		—
Serum Cr (SD)	92.8 (20.7) μmol/L	
CrCl[a] (ABW)	76.4 (19) mL/min	—
CKD (total No. 49):		
CKD stage-3	8 (16%)	—
Non-CKD (Normal renal function)	41 (82%)	

Total number of CAD patients and controls is 50 subjects per each group. Abbreviations: SD: Standard deviation, DM: diabetes mellitus, CKD: chronic kidney disease, Cr: creatinine, Cr Cl: Creatinine clearance, ABW: Adjusted body weight (for obese patients) and No.: number of valid cases.

[a] Creatinine clearance (CrCl) for each patient was calculated using Cockcroft-gault (C-G) equation. The adjusted body weight was used to calculate CrCl for overweight and obese patients.

In plasma, the enrichment ORA of ULR metabolites against the metabolic pathway-associated-metabolite sets library showed that urea cycle and glucose-alanine cycle pathways had the highest fold enrichment, the lowest p value ($p < .001$) and FDR < 0.1 (Table 4 and Fig. 2.A). There were other significant pathways including the glycine-serine metabolism, valine-leucine-isoleucine degradation, alanine metabolism and Warburg effect, however they had FDR > 0.1. Table 4 details the p value and the FDR of the significant pathways. Fig. 2.A depicts the Bar chart of the enrichment ORA for CAD-plasma metabotype against metabolic pathway-associated metabolite sets library in plasma. The enrichment ORA of plasma ULR metabolites against SNPs-associated-metabolite sets library revealed that 85 SNPs had p value <.05, fold enrichment was >30 and FDR < 0.1 (see Fig. S.2 and table E.1 in the online supplementary material). Of the significant SNPs, 11 had very strong significance ($p < .0001$) and very low FDR (FDR < 0.02). Table 5 details the p values, implicated metabolites and FDRs of the 11 SNPs with the lowest p and FDR values. The enrichment ORA of plasma ULR metabolites against diseases-associated-metabolite sets library showed that 51 diseases were enriched in the metabolites' list (p value <.05) (see online supplementary table E.2 and Fig. S.3). Of the significant diseases, 41 diseases had FDR < 0.1. Early markers of myocardial injury and heart failure (HF) had >6-fold enrichment, strong significance ($p < .001$) and extremely low FDR (FDR < 0.001). Hyperornithinemia with gyrate atrophy (HOGA), pyruvate dehydrogenase deficiency (E3), chronic progressive external

Table 2
Specificity, sensitivity, accuracy and AUROC of OPLS-DA and MVLR CAD metabotypes models.

Metabotype model	OPLS-DA model				MVLR model			
	Specificity	Sensitivity	Accuracy	AUROC	Specificity	Sensitivity	Accuracy	AUROC
Plasma	100%	96%	98%	0.994	92%	86%	89%	0.96
Urine	100%	96%	98%	0.9996	90%	80%	85%	0.92

Abbreviations: OPLS-DA: Orthogonal partial least square discriminant analysis, MVLR: Multivariate logistic regression, AUROC: Area under receiver operating characteristic.

ophthalmoplegia and Kearns-Sayre syndrome, pyruvate dehydrogenase E3-binding protein deficiency and pyruvate carboxylase deficiency had fold enrichment ≥12, strong significance ($p < .0001$) and very low FDR (FDR < 0.01). Diabetes Mellitus (DM) - maturity onset diabetes of the young (MODY) had p value of 0.00096, fold enrichment >4.5 and FDR 0.016. The enrichment ORA of plasma ULR metabolites against locations-associated-metabolite sets library showed that prostate, pancreas and mitochondria were significantly enriched in the metabolites' list (p value <.05), however, only prostate had FDR < 0.1 (see also Fig. S.4 in the online supplementary material for the bar chart depiction of the analysis). Pathway topology analysis using the identified metabolites from the MVLR models showed that perturbations in aminoacyl-tRNA biosynthesis, synthesis and degradation of ketone bodies, propanoate metabolism, arginine and proline metabolism, alanine-aspartate-glutamate metabolism, valine-leucine-isoleucine biosynthesis, D-glutamine and D-glutamate metabolism, valine-leucine-isoleucine degradation and butanoate metabolism had $p < .05$ and FDR < 0.1. There were other significantly perturbed pathways; glycine-serine-threonine metabolism and glycolysis or gluconeogenesis, however they had FDR > 0.1. Pathway topology analysis is detailed in Table 6. The pathways are ranked relatively based on the level of significance and impact. Fig. 3.A depicts graphical overview for the metabolic pathway topology analysis of CAD plasma metabotype.

In urine, the enrichment ORA of ULR metabolites against the metabolic pathway-associated-metabolite sets library showed that the methylhistidine metabolism pathway had significantly high fold enrichment (>40-fold enrichment, ($p < .001$) and FDR < 0.1 (Table 4 and Fig. 2.B)). Another significant pathway was histidine metabolism, however it had FDR > 0.1. Table 4 details the p values and the FDR values of the two significant pathways. Fig. 2.B depicts the Bar chart of the ORA for CAD-urine metabotype against metabolic pathway-associated metabolite sets library in urine. The enrichment ORA of urine ULR metabolites against SNPs-associated-metabolite sets library revealed that 341 SNPs had $p < .05$, however none of them had FDR < 0.1 (see Fig. S.5 and table E.3 in the online supplementary material). The enrichment ORA of urine ULR metabolites against diseases-associated-metabolite sets library showed that 11 diseases were significantly enriched in the metabolites' list (p value <.05) (see online supplementary table E.4 and Fig. S.6). Of them, obesity and degradation of skeletal muscle had >40-fold-enrichment and FDR < 0.1. Among the significant diseases, there were T2DM, Alzheimer's disease, DM-MODY, type I and glucoglycinuria, however they had FDR > 0.1. The enrichment ORA of urine ULR metabolites against locations-associated-metabolite sets library showed that prostate and muscle were enriched in the metabolites' list (p value <.05), however, both had FDR > 0.1 (see also Fig. S.7 in the online supplementary material for the bar chart depiction of the analysis). Pathway topology analysis using the identified metabolites from the MVLR models showed that perturbations in galactose metabolism, histidine metabolism and amino-sugar and nucleotide-sugar metabolism had $p < .05$. None of them had and FDR < 0.1. The analysis is detailed in Table 6. Fig. 3.B depicts graphical overview for the metabolic pathway topology analysis of CAD urine metabotype.

4. Discussion

^{1}H NMR metabolomics approach was applied to identify novel CAD metabotypes in plasma and urine and the implicated pathways in the discovered metabotypes. The discriminating OPLS-DA and the MVLR models (metabotypes) had high accuracies and AUROCs. These metabotypes, presuming further validation in future studies, can be clinically useful in CAD diagnosis and management. The identified biomarkers and the pathway analysis revealed further insights on the pathophysiology of CAD and the multifactorial predisposing factors. Many of the biomarkers of the discovered CAD metabotypes are known to be associated with cardiovascular, atherosclerotic and cardiometabolic diseases. For instance, the discovered metabotypes contained dyslipidaemia biomarkers; mainly triglycerides, very low-density

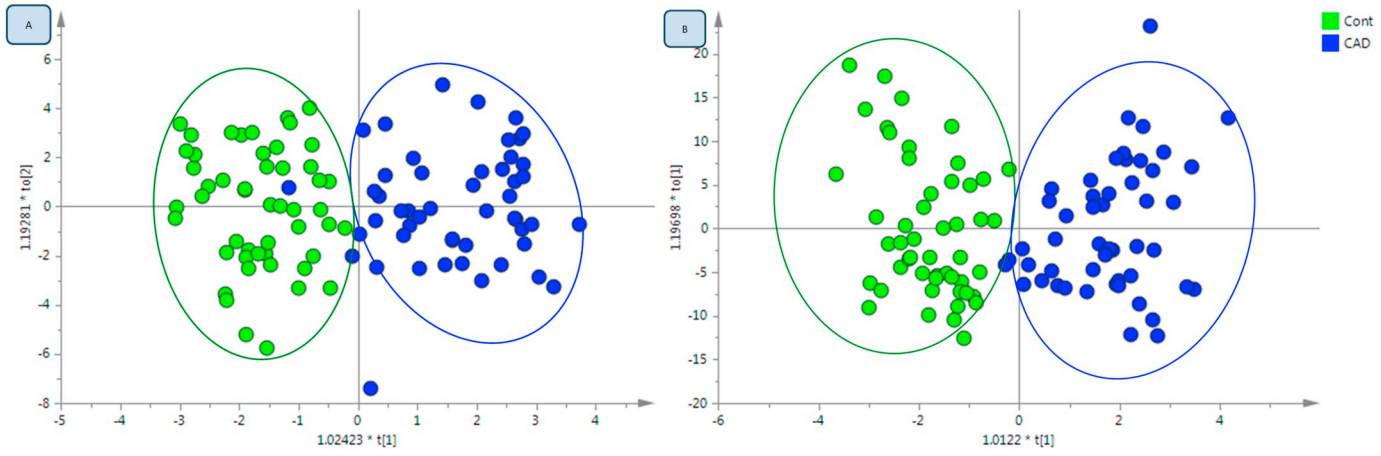

Fig. 1. Score scatter plot of the OPLS-DA models for plasma and urine CAD metabotypes.
A: Score plot of OPLS-DA plasma model. B: Score plot of OPLS-DA urine model. Abbreviations: Cont: Control, CAD: coronary artery disease. For further details on these models accuracy and AUROC, refer to table 2.

Table 3
Plasma and urine bins' regions with p value $<.1$ in the univariate logistic regression analysis and their corresponding identified metabolites.

Bins' ppm[a]	Identified metabolites
Plasma	
0.94	L-Isoleucine, L-Leucine
0.98, 1.02	L-Valine, L-Isoleucine
1.06	L-Valine
1.10	3-methyl-2-oxovalerate
1.14	2,3-butanediol
1.18	3-hydroxybutyric acid, *N*-acetyl-threonine
1.26	Lipid (VLDL/LDL Cholesterol), Oleic acid
1.30	L-Lactic acid, L-Threonine, Lipid (VLDL/LDL Cholesterol), Oleic acid
1.42	L-Lysine
1.46	L-Alanine, L-Lysine, L-Isoleucine
1.50	L-Alanine, L-Lysine
1.54	L-Lysine, Lipid
1.58	Lipid
1.62	2-oxovalerate, Oleic acid
1.66, 1.70	L-Lysine, L-Leucine, Oleic acid
1.74, 1.82, 1.86	L-Lysine, L-Ornithine
1.98, 3.38	L-Proline
2.18	Triacylglycerol
2.22	Triacylglycerol, L-Valine, Acetone
2.26	L-Valine, Acetoacetate
2.30	L-Glutamic acid, L-Proline, 3-hydroxybutyric acid
2.34	L-Glutamic acid, L-Proline, 3-hydroxybutyric acid, pyruvate, Oleic acid
2.38	L-Glutamic acid, L-Proline, 3-hydroxybutyric acid, L-Glutamine, 3-oxoadipic acid, Oleic acid
2.78	Polyunsaturated fatty acid, 3-oxoadipic acid
2.82	3-oxoadipic acid
3.26, 3.50, 3.54. 3.82, 3.86, 3.90, 4.62, 4.66, 5.22	D-Glucose
3.42	D-Glucose, L-Proline
3.46	D-Glucose, Acetoacetate, 3-oxoadipic acid
3.70	D-Glucose, DMG, L-Leucine, PAG
3.74	D-Glucose, L-Lysine, Mannose
3.78	L-Glutamine, L-Alanine, L-ornithine, D-Glucose, Mannose, PAG
4.02	Creatinine, Choline, PA
4.06	L-Lactic acid, Choline
4.22	L-Threonine, *N*-acetyl-threonine
4.26	Sn-Glycerophosphocholine, L-Threonine, *N*-acetyl-threonine
4.46	Salidroside
5.30	Unsaturated Fatty acid
5.70, 5.74, 5.78, 5.82	Urea
7.30	PA
7.34, 7.38, 7.42	PA, PAG
8.42	Formic acid
4.50, 4.54, 7.22, 7.26	Un-identified bins
Urine	
3.26	D-Glucose
4.62	D-Glucose
4.9	Mannose
5.06	Riboflavin
5.14	Mannose
5.18	Mannose
5.34	Chlorogenic acid
7.02	Salicylamide, 1-methylhistidine, 4-hydroxyhippuric
7.06	L-Histidine, Salicylamide, 3-methylhistidine
7.50	4-hydroxyhippuric, Hippuric acid, 3-indoxylsulfate
7.78	4-hydroxyhippuric, Salicylamide, L-Histidine
8.94	1-Methylnicotinamide
4.54, 4.58, 4.70, 4.86, 4.94, 4.98, 5.02	Unidentified bins

[a] The chemical shift of the bin (ppm) corresponds to the central chemical shift of the bin's range (i.e. 3.78 ppm: 3.76 ppm – 3.80 ppm). The metabolites were identified and confirmed by overlaying of 1-D CPMG (Carr Purcell Meiboom Gill one dimensional) in plasma, 2-D J-resolved (J-RES) and 2-D Heteronuclear Single Quantum Coherence (HSQC) spectra with metabolites platforms such as; B-BIOREFCODE, Chenomx and HMDB. For further details on the ULR- p values, please refer to the supplementary data. Abbreviations: ppm: chemical shift, DMG: *N*,*N*-dimethylglycine, PA: L-Phenylalanine, PAG: Phenylacetylglycine. **bins' ppm** in **bold** indicate presence of the bin in multivariate logistic regression (MVLR) model.

lipoprotein (VLDL) and low density lipoprotein (LDL), a well-known risk factors of CAD [19]. Branched chain amino acids (BCAAs); L-leucine, L-isoleucine and L-valine, and their related catabolites such as phenylalanine are well-documented proteinogenic-amino-acid trait of CAD risk factors, CAD development, CAD severity, myocardial infarction and death [20–22]. Noteworthily, a previous study by Yang and colleagues indicated that BCAAs were strongly associated with the increase in carotid intima-media thickness [21]. Elevations of phenylacetylglycine, acetone, ornithine and creatinine are commonly discussed biomarkers in heart failure (HF) [23,24]. Moreover, acetone and 3-hydroxybutyrate are associated with elevated myocardial energy expenditure in HF [25]. Elevations in choline and 3-indoxylsulfate are associated with unstable angina [26]. D-glucose peaks were majorly presented in one of the associated factors in our CAD plasma model (factor 3 in Tables S.3 and S.4). This is in line with previous reports on the association of blood glucose level with CAD development and progression in diabetic and non-diabetic patients [27–30]. In their study on CAD patients, Yubero-Serrano and colleagues found that glucose concentration at 120 min post oral glucose tolerance test was an independent predictor of increased intima-media thickness of common carotid arteries (IMMT-CC) [27]. Noteworthily, T2DM – CAD patients with HbA1c $>$ 6.5 had significant increase in IMMT-CC compared to CAD patients with T2DM - HbA1c $<$ 6.5, pre-diabetes and normoglycemia. Taken together, the collective spectrum of CAD metabotypes' biomarkers represents a distinctive phenotype of the complexity of CAD development and progression.

Using MetaboAnalyst 4.0, a web-based software of pathway analysis to identify associated metabolite sets, SNPs, diseases, locations and perturbed metabolic pathways gave profound interpretation and understanding of the associations between the discovered biomarkers and CAD. One of the remarkable findings in our study is the pathway analysis of CAD plasma metabotype against SNPs-associated metabolite set library. This analysis revealed that there are 85 SNPs that are significantly associated with CAD plasma metabotype, particularly 11 SNPs with high significance ($p < .0001$) (Tables; 5, E.1 & Fig. S.2). To the best of our knowledge, this is the first study to report an association between these SNPs and CAD. However, some of these SNPs had been found to be associated with other conditions. For instance, *rs2588400* which had fold-enrichment $>$60, very high significance and very low FDR in our study was found to be associated with vitiligo in Koreans; an East Asian population [31]. Since our subjects are East Asians as well, it is interesting to discover that this genetic variant (*rs2588400*) is associated with CAD. In fact, vitiligo has been reported to be associated with cardiovascular disease (CVD) and DM [32,33]. Therefore, further experimental investigation of the clinical utility of this SNP as a biomarker of CAD can be instigating for future research. Another SNP is the *rs2194980* which had fold-enrichment $>$30, very high significance and very low FDR. This SNP was found to be associated with the perturbation in metabolite levels, particularly the amino acid tyrosine [34,35]. None of the SNPs which had significant association with CAD urine metabotype had FDR $<$ 0.1 so it is less likely that they have impactful role in the CAD metabotype-SNPs pathway link.

The findings of the ORA of CAD metabotypes against the metabolic pathway-associated-metabolite sets library and the topology pathway analysis had many common pathways that were indicated previously in the metabotypes of clopidogrel HTPR [9,10]. Pathways such as urea cycle, glucose-alanine cycle, aminoacyl-tRNA biosynthesis, glycine-serine-threonine metabolism, synthesis and degradation of ketone bodies, valine-leucine-isoleucine biosynthesis and degradation, butanoate metabolism, D-glutamine and D-glutamate metabolism, propanoate metabolism, arginine and proline metabolism, glycolysis or gluconeogenesis and alanine-aspartate-glutamate metabolism were

Table 4
Enrichment analysis of CAD metabotypes – Pathway associated metabolite sets.

Pathway Associated Metabolite sets[a]	Implicated Metabolites[b]	Total[c]	Hits[d]	Raw *p*[e]	Holm *p*[f]	FDR[g]
Plasma CAD metabotype						
Urea Cycle	L-Glutamic acid; L-Alanine; Ornithine; Pyruvic acid; Urea, L-Glutamine	29	6	0.000196	0.0192	0.0192
Glucose-Alanine cycle	D-Glucose; L-Glutamic acid; L-Alanine; Pyruvic acid	13	4	0.000522	0.0507	0.0256
Glycine and serine metabolism	Dimethylglycine; L-Glutamic acid; L-Alanine; L-Threonine; Pyruvic acid; Ornithine; *N*-acetylthreonine[h]	59	7[h]	0.00925	0.888	0.246
Valine, leucine and isoleucine degradation	Acetoacetic acid; L-Glutamic acid; 3-Methyl-2-oxovaleric acid; L-Valine; L-Leucine; L-Isoleucine	60	6	0.01	0.954	0.246
Alanine Metabolism	L-Alanine; L-Glutamic acid; Pyruvic acid	17	3	0.0152	1	0.299
Warburg Effect	D-Glucose; L-Lactic acid; L-Glutamic acid; L-Glutamine; Pyruvic acid	58	5	0.0342	1	0.559
Urine CAD Metabotype						
Methylhistidine Metabolism	L-Histidine; 3-Methylhistidine	4	2	0.000746	0.0731	0.0731
Histidine Metabolism	L-Histidine; 1-Methylhistidine; 3-Methylhistidine	43	3	0.0117	1	0.571

This Metabolites enrichment analysis was generated by MetaboAnalyst 4.0 web-based program using the identified metabolites from bins with p value < .1 in the univariate logistic regression (ULR) analysis (i.e. Identified metabolites are presented in Table 3). See Fig. 2 for bar chart visualization of this analysis.

[a] Only pathway associated metabolite sets with $p < .05$ are shown in the table.
[b] Implicated metabolites from CAD metabotypes in the associated metabolites set.
[c] Total number of metabolites in the metabolites set.
[d] Hits: number of metabolites from CAD metabotype involved in the metabolites set.
[e] Raw p: original p value calculated from the enrichment analysis.
[f] Holm p: adjusted raw p value by Holm-Bonferroni method.
[g] FDR: false discovery rate.
[h] *N*-acetylthreonine is not available in MetaboAnalyst 4.0 database, however it is a product of threonine metabolism and known metabolite of the glycine and serine metabolism pathway. We considered *N*-acetylthreonine in the hits count, but it was not auto-included in the calculation of p value and FDR by MetaboAnalyst 4.0 since it was not available in its database. Had it has been included in that analysis, the Glycine and serine metabolism pathway may have had lower p value.

significantly associated with clopidogrel HTPR plasma metabotypes [9]. Similarly, galactose metabolism and histidine metabolism pathways were significantly associated with clopidogrel HTPR urine metabotypes [10]. Considering that clopidogrel HTPR was assessed using platelets function assessment kit of platelets P2Y12 receptors' activity (VerifyNow® P2Y12 kit) [10], one can infer that the aforementioned pathways which are common with CAD metabotypes, are metabolic-phenotype of increased platelets reactivity. Indeed, CAD patients tend to have high platelets reactivity with usual circulation of active platelets [4], and perhaps patients who endure extreme increase in platelets reactivity suffer low response to antiplatelets as well. Another notable finding which may substantiate our understanding of CAD associated platelets reactivity metabolic-phenotype is that many of those common metabolic pathways had been reported to be perturbed in other well-known risk factors of CAD. Metabolic pathways of urea cycle, aminoacyl-tRNA biosynthesis, glycine-serine-threonine metabolism, arginine-proline metabolism, valine-leucine-isoleucine biosynthesis and degradation, synthesis and degradation of ketone bodies, alanine-aspartate-glutamate metabolism, histidine metabolism and glycolysis or gluconeogenesis were found to be perturbed in T2DM and insulin

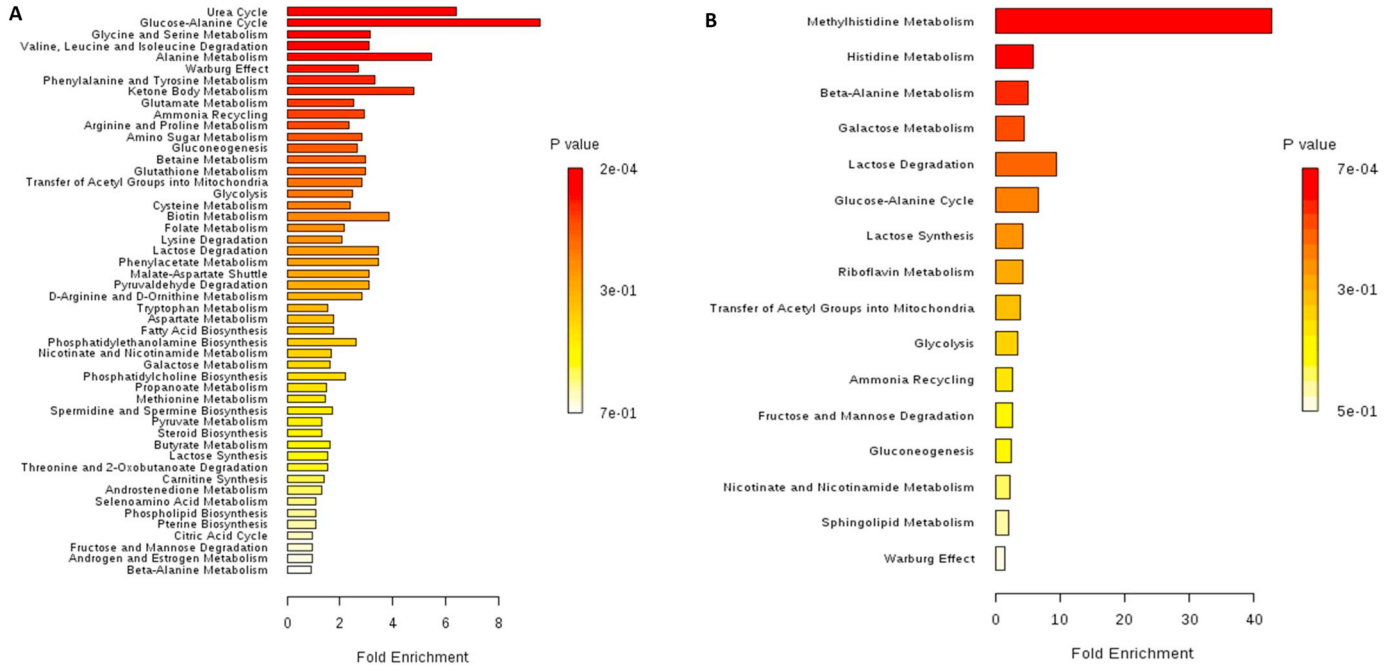

Fig. 2. Graphical overview of the enrichment over representation analysis (ORA) for metabolic pathways associated with CAD metabotypes.
A: Bar chart of the ORA for the CAD plasma metabotype against metabolic pathway-associated metabolite sets library in plasma. B: Bar chart of the ORA for CAD urine metabotype against metabolic pathway-associated metabolite sets library in urine. Pathway associated metabolites sets are sorted based on fold enrichment and p value. For further details on the p value of each metabolite set, refer to table 4. Graphs were generated by MetaboAnalyst 4.0 web-based software.

Table 5
Enrichment ORA of plasma metabotype against SNPs-associated-metabolite sets library.

SNP associated metabolite sets[a]	Implicated Metabolites[b]	Total[c]	Hits[d]	Raw *p*[e]	Holm *p*[f]	FDR[g]
rs169712	L-Proline; L-Leucine; L-Valine; L-Glutamic acid	5	4	2.09E-07	0.00092	0.00092
rs2194980	L-Valine; L-Phenylalanine; L-Glutamic acid; L-Leucine	8	4	2.83E-06	0.0125	0.00227
rs10874325	L-Leucine; Ornithine; L-Valine	3	3	3.09E-06	0.0136	0.00227
rs11939527	L-Proline; L-Leucine; L-Valine	3	3	3.09E-06	0.0136	0.00227
rs2588400	L-Proline; Ornithine; L-Glutamine	3	3	3.09E-06	0.0136	0.00227
rs8041815	L-Proline; Ornithine; L-Leucine	3	3	3.09E-06	0.0136	0.00227
rs1392880	L-Valine; L-Leucine; L-Glutamic acid	4	3	1.22E-05	0.0538	0.00674
rs2567397 | rs7664292	L-Valine; L-Glutamic acid; L-Leucine	4	3	1.22E-05	0.0538	0.00674
rs2300701	L-Leucine; L-Phenylalanine; L-Valine	5	3	3.02E-05	0.133	0.0133
rs313408	L-Leucine; L-Valine; L-Alanine	5	3	3.02E-05	0.133	0.0133

This enrichment ORA of CAD plasma metabotype against SNPs-associated-metabolite set library was generated by MetaboAnalyst 4.0 web-based software using the list of identified metabolites from bins with $p < .1$ in the univariate logistic regression (UVLR) analysis (i.e. Identified bins are presented in Table 3).

[a] Only SNPs with $p < .0001$ are shown in the table.

[b] Implicated metabolites from CAD metabotype.

[c] Total number of metabolites in the SNP-associated-metabolite set.

[d] Hits: number of metabolites from CAD metabotype involved in the SNP-associated-metabolite set.

[e] Raw *p*: original *p* value calculated from the ORA.

[f] Holm *p*: adjusted raw *p* value by Holm-Bonferroni method.

[g] FDR: false discovery rate. All of SNPs associated metabolite sets in urine had FDR > 0.1 so they were not presented in this table, see online supplementary table E.3 for further details.

resistance [36–38]. Since advanced glycation-end-products which are formed due to persistent hyperglycaemia lead to platelets activation and impaired fibrinolysis [5,39], 2 distinctive features of CAD [4,40], it is very conceivable that CAD and T2DM share platelets reactivity metabolic-phenotype. Moreover, persistent insulin resistance leads to poor outcome in CAD patients [41]. Other metabolic conditions such as obesity and metabolic syndrome have common altered metabolic pathways with CAD as well. Perturbations in aminoacyl-tRNA biosynthesis, glycolysis or gluconeogenesis, glycine-serine-threonine metabolism, synthesis and degradation of ketone bodies and propanoate metabolism pathways were identified previously in the metabotypes of metabolic syndrome and obesity [42–45]. Patients who are suffering from metabolic syndrome and obesity usually suffer from insulin resistance and glucose intolerance [42,46]. Obesity and metabolic syndrome are associated with platelets activity and endothelial dysfunction [47,48]. While near to half of our CAD patients were suffering from T2DM, majority of them were overweight and obese (Table 1). In other words, many of our patients endured at least one cardiometabolic disease that interferes with platelets reactivity. Taken together, these findings substantiate the role of glucose intolerance and insulin resistance in CAD development and the importance of early management of these risk factors in CAD prevention and treatment.

The ORA of CAD metabotypes against disease-associated-metabolite sets library revealed many significant disease associations. Considering the identified CAD biomarkers and the metabolic pathways, many of disease-associations such as HF, early markers of myocardial injury, DM-MODY, DM-non-insulin and DM-insulin dependent, T2DM and obesity were not far from expected. Besides what we discussed early on cardiovascular diseases' biomarkers, many of the metabolites in CAD metabotypes were commonly described in T2DM, insulin resistance and obesity. Metabolites such as; D-glucose, mannose, 3-hydroxybutyrate, acetoacetate, acetone, BCAAs, L-phenylalanine, L-proline, L-ornithine, L-alanine, L-lysine, 2-oxovalerate, pyruvate, L-threonine, L-histidine, L-glutamic acid, dimethylglycine (DMG), choline, L-lactic acid, phenylacetylglycine and creatinine are increasingly discussed biomarkers of T2DM [36–38,49–52]. Most, if not all, of these biomarkers are remarkably present in the metabolic traits of insulin resistance, metabolic syndrome and obesity [42,43,46,51–53]. The second most significant disease association with urine CAD metabotype was the degradation of skeletal muscle. Since all of our CAD patients were on statins, this association might be due to statins' muscle toxicity side-effect. While it may not be symptomatic or appear in creatine kinase plasma levels, statin induced myopathy may manifest in the metabolic biomarkers due to alteration in muscle metabolism and inflammation [54,55]. Furthermore, DM, other metabolic diseases and genetic predispositions increase the risk of statin induced myopathy [54]. Given that, majority of our patients were overweight or obese and near to half were T2DM, it could be inferred that large proportion of the patients were affected by statins induced muscle toxicity.

The ORA of CAD metabotypes against location-associated-metabolite sets library showed that prostate was the most significant location in plasma and urine. The correlations between cardiovascular diseases and both prostatic cancer and benign prostatic hyperplasia (BPH) had been documented, however, often mitigated by some controversy in the literature [56,57]. Noteworthy, most of our CAD patients were male. The second significant location in plasma metabotype was the pancreas followed by the mitochondria. This can be justified by the association between DM and CAD which was clearly present in the discovered biomarkers and the metabolic pathways. Furthermore, it is well documented that mitochondrial dysfunction and oxidative stress are associated with CAD and cardiometabolic diseases [58–60]. The muscle was the second significant location in urine metabotype. This may add to our speculation for the phenotype of statins induced myopathy among our patients [54,55].

The discovered plasma and urine metabotypes in the current study suggest potential role of dysbiotic gut microbiota in CAD. The metabotypes contained dysbiotic gut microbiota derived metabolites such as choline, L-phenylalanine, phenylacetylglycine, creatinine, D-glucose, L-leucine, L-isoleucine, L-alanine, L-valine, L-threonine, L-lactic acid, acetoacetate, 3-hydroxybutyric acid, pyruvate, L-proline, L-histidine, 3-indoxylsulfate, hippuric acid and 4-hydroxyhippuric acid [61,62]. Although trimethylamine-N-oxide (TMAO), a well-known pro-atherogenic and cardiovascular events-associated gut-microbiota derived metabolite [63], was not present in the discovered metabotypes, creatinine and L-valine which are present were found to be associated with TMAO level in patients suffering from T2DM, metabolic syndrome and pre-diabetes [62]. This potential role of dysbiotic gut microbiota in cardiometabolic diseases highlights new perspectives of treatment and prevention where dietary intervention can be used to modulate dysbiotic gut microbiota [64,65].

This study showed that ^{1}H NMR metabolomics profiling of plasma and urine was robust in phenotyping stable CAD with high accuracy. It provided deep insights on the implicated pathways and the probable predisposing genetic and non-genetic factors. Our findings may lead to

Table 6
Metabolic pathways associated with CAD Plasma and Urine metabotypes.

Metabolic Pathway[a]	Implicated Metabolites[b]	Total[c]	Hits[d]	Raw p[e]	−Log (p)[f]	Holm p[g]	FDR[h]	Impact[i]
Plasma Metabotype								
Aminoacyl-tRNA biosynthesis	L-Threonine; L-Lysine; L-Alanine; L-Valine; L-Glutamine; L-Glutamic acid; L-leucine; L-Proline	75	8	6.07E−07	14.315	4.85E−05	4.85E−05	0.11268
Synthesis and degradation of ketone bodies	Acetoacetic acid; 3-Hydroxybutyric acid; Acetone	6	3	2.19E−05	10.728	0.001732	0.000877	0.7
Propanoate metabolism	L-Lactic acid; Acetoacetic acid; L-Valine; Acetone	35	4	0.000447	7.7136	0.034843	0.011912	0.04317
Arginine and proline metabolism	L-Glutamine; L-Glutamic acid; Ornithine; L-Proline; Urea	77	5	0.001142	6.7749	0.087938	0.022841	0.29805
Alanine, aspartate and glutamate metabolism	L-Alanine; L-Glutamine; L-Glutamic acid	24	3	0.001949	6.2403	0.14815	0.031188	0.44065
Valine, leucine and isoleucine biosynthesis	L-Threonine; L-Leucine; L-Valine	27	3	0.002757	5.8936	0.20678	0.03676	0.0265
D-Glutamine and D-glutamate metabolism	L-Glutamine; L-Glutamic acid	11	2	0.005814	5.1475	0.43023	0.066445	0.13904
Valine, leucine and isoleucine degradation	Acetoacetic acid; L-Leucine; L-Valine	40	3	0.008484	4.7695	0.61936	0.075417	0.02232
Butanoate metabolism	3-Hydroxybutyric acid; Acetoacetic acid; L-Glutamic acid	40	3	0.008484	4.7695	0.61936	0.075417	0.0451
Glycine, serine and threonine metabolism	Dimethylglycine; L-Threonine; Choline; *N*-acetylthreonine[j]	48	4	0.014027	4.2668	0.99593	0.11222	0.13081
Glycolysis or Gluconeogenesis	L-Lactic acid; D-Glucose	31	2	0.04307	3.1449	1	0.31323	0
Urine Metabotype								
Galactose metabolism	D-Glucose; D-Mannose	41	2	0.00743	4.9022	0.59439	0.34115	0.00276
Histidine metabolism	Histidine; 1-Methylhistidine	44	2	0.008529	4.7643	0.67377	0.34115	0.14548
Amino-sugar and nucleotide-sugar metabolism	D-Mannose; D-Glucose	88	2	0.03206	3.4402	1	0.85492	0

This pathway analysis was generated by MetaboAnalyst 4.0 web-based software using identified metabolites from multivariate logistic regression (MVLR) results of bins associated with CAD (Bins are presented in bold in Table 3).

[a] Only metabolic pathways with $p < .05$ are shown in the table.
[b] Implicated metabolites from CAD metabotype.
[c] Total number of metabolites in the pathway.
[d] Hits: number of metabolites from CAD metabotype involved in the pathway.
[e] Raw p: original p value calculated from the pathway analysis.
[f] -log (p): negative log of (p) value.
[g] Holm p: adjusted raw p value by Holm-Bonferroni method.
[h] FDR: false discovery rate.
[i] Impact: impact of the pathway as calculated from pathway topology analysis.
[j] *N*-acetylthreonine was not available in MetaboAnalyst 4.0 database, however, it is a product of threonine metabolism and known metabolite of the Glycine; serine and threonine metabolism pathway. Had it been included in the pathway analysis the Glycine; serine and Threonine metabolism pathway may had more significance.

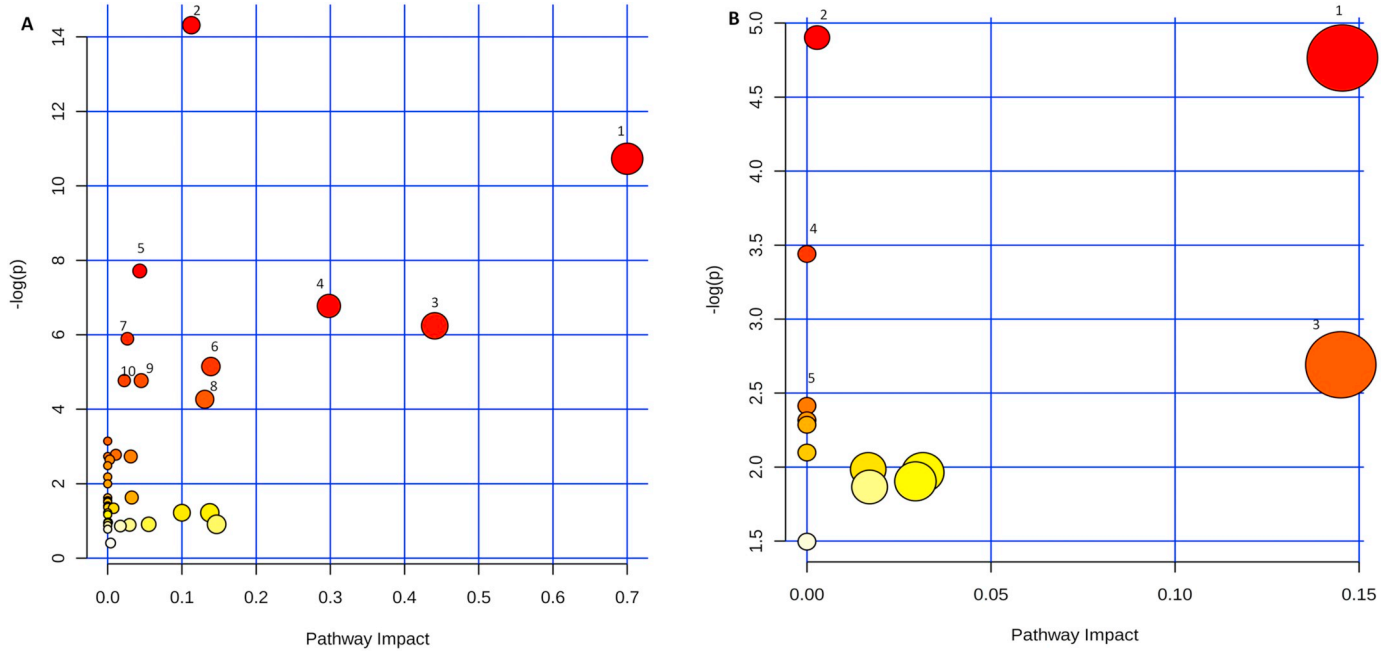

Fig. 3. Graphical overview of metabolic pathway topology analysis of CAD metabotypes*
A: Graphical overview of metabolic pathway topology analysis for metabolites associated with CAD in plasma multivariate logistic regression (MVLR) model; (1) Synthesis and degradation of ketone bodies, (2) Aminoacyl t-RNA biosynthesis, (3) Alanine, Aspartate and glutamate metabolism, (4) Arginine and proline metabolism, (5) Propanoate metabolism, (6) D-Glutamine and D-Glutamate metabolism, (7) Valine, leucine and isoleucine biosynthesis, (8) Glycine, serine and threonine metabolism, (9) Butanoate metabolism, (10) Valine, leucine and isoleucine degradation. B: Graphical overview of metabolic pathway topology analysis for metabolites associated with CAD in urine MVLR model; (1) Histidine metabolism, (2) Galactose metabolism, (3) Riboflavin metabolism, (4) Amino sugar and nucleotide sugar metabolism, (5) beta-alanine metabolism. *Graphs were generated by MetaboAnalyst 4.0 web-based software. For further details on the p value of each metabolic pathway, refer to table 6.

new perspective where some protective measures can be applied to prevent CAD development in risk groups. However, further investigations of the identified biomarkers and pathways in an independent and large group are warranted to validate the findings of the current study. In spite of the meaningful findings of this study, it has its limitations. The MVLR metabotypes' models had lower accuracies and AUROCs than the OPLS-DA models. This could be due to the loss of some spectral regions after running the ULR and FA. However, sometimes the lost spectral regions are noises which lose the significance of their association upon running ULR. Although we matched the healthy controls with CAD in terms of sex, ethnicities and age, this could not be perfectly achieved. Despite avoiding the selection of very old CAD patients and very young healthy controls, mean age of controls was lower than CAD. This was because choosing older age healthy subjects who are not on any medications and not having history of any chronic disease was challenging. Have had we included elder controls, they might have had confounding diseases and medications. Furthermore, using highly-matched age, sex and ethnicity groups may cause overemphasizing of methods artificially [66]. Although the enrichment ORA provided profound information on CAD-associated-SNPs in our study, future investigation of these associations using experimental analysis is required prior to establishing them as CAD-genetic markers. While ^{1}H NMR has shown to be fast and reliable tool for metabolomics analysis, using other spectroscopic instruments to analyse the samples such as mass spectroscopy (MS) may discover more biomarkers and pathways.

Acknowledgments

This work was funded by Universiti Sains Malaysia, Penang, Malaysia. The authors declare that they have no conflict of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cca.2019.02.030.

References

[1] World Health Organization, Cardiovascular diseases (CVDs), http://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds), (2017) , Accessed date: February 2018.
[2] American Heart Association, Coronary Artery Disease - Coronary Heart Disease, http://www.heart.org/HEARTORG/Conditions/More/MyHeartandStrokeNews/Coronary-Artery-Disease—Coronary-Heart-Disease_UCM_436416_Article.jsp, (2014) , Accessed date: 23 August 2014.
[3] P. Libby, P. Theroux, Pathophysiology of coronary artery disease, Circulation 111 (25) (2005) 3481–3488.
[4] M.I. Furman, S.E. Benoit, M.R. Barnard, C.R. Valeri, M.L. Borbone, R.C. Becker, H.B. Hechtman, A.D. Michelson, Increased platelet reactivity and circulating monocyte-platelet aggregates in patients with stable Coronary Artery Disease, J. Am. Coll. Cardiol. 31 (2) (1998) 352–358.
[5] V.P. Singh, A. Bali, N. Singh, A.S. Jaggi, Advanced Glycation end products and diabetic complications, The Korean Journal of Physiology & Pharmacology : Official Journal of the Korean Physiological Society and the Korean Society of Pharmacology 18 (1) (2014) 1–14.
[6] N. Semmar, Metabotype concept: flexibility, usefulness and meaning in different biological populations, in: D.U. Roessner (Ed.), Metabolomics, InTech, 2012.
[7] A.M. Amin, L. Sheau Chin, D. Azri Mohamed Noor, M.A. SK Abdul Kader, Y. Kah Hay, B. Ibrahim, The personalization of clopidogrel antiplatelet therapy: the role of integrative pharmacogenetics and pharmacometabolomics, Cardiol. Res. Pract. 2017 (2017) 17.
[8] J. Lindon, J. Nicholson, E. Holmes, The Handbook of Metabonomics and Metabolomics, (2006).
[9] A.M. Amin, L. Sheau Chin, C.-H. Teh, H. Mostafa, D.A. Mohamed Noor, M.A.S.K. Abdul Kader, Y. Kah Hay, B. Ibrahim, Pharmacometabolomics analysis of plasma to phenotype clopidogrel high on treatment platelets reactivity in coronary artery disease patients, Eur. J. Pharm. Sci. 117 (2018) 351–361.
[10] A.M. Amin, L. Sheau Chin, C.-H. Teh, H. Mostafa, D.A. Mohamed Noor, M.A. Sk Abdul Kader, Y. Kah Hay, B. Ibrahim, 1H NMR based pharmacometabolomics analysis of urine identifies metabolic phenotype of clopidogrel high on treatment platelets reactivity in coronary artery disease patients, J. Pharm. Biomed. Anal. 146 (2017) 135–146.

[11] A.M. Amin, L. Sheau Chin, D.A. Mohamed Noor, H. Mostafa, M.A.S.K. Abdul Kader, Y. Kah Hay, B. Ibrahim, The effect of CYP2C19 genetic polymorphism and non-genetic factors on clopidogrel platelets inhibition in East Asian coronary artery disease patients, Thromb. Res. 158 (2017) 22–24.
[12] H. Mostafa, A.M. Amin, C.-H. Teh, A. V., N.H. Murugaiyah, B. Ibrahim Arif, Plasma metabolic biomarkers for discriminating individuals with alcohol use disorders from social drinkers and alcohol-naive subjects, J. Subst. Abus. Treat. 77 (2017) 1–5.
[13] H. Mostafa, A.M. Amin, C.H. Teh, V. Murugaiyah, N.H. Arif, B. Ibrahim, Metabolic phenotyping of urine for discriminating alcohol-dependent from social drinkers and alcohol-naive subjects, Drug Alcohol Depend. 169 (2016) 80–84.
[14] A.S. Levey, K.-U. Eckardt, Y. Tsukamoto, A. Levin, J. Coresh, J. Rossert, D.d. Zeeuw, T.H. Hostetter, N. Lameire, G. Eknoyan, Definition and classification of chronic kidney disease: a position statement from Kidney Disease: Improving Global Outcomes (KDIGO), Kidney Int. 67 (6) (2005) 2089–2398.
[15] T.L. Hwang, A.J. Shaka, Water suppression that works. Excitation sculpting using arbitrary wave-forms and pulsed-field gradients, J. Magn. Reson. Ser. A 112 (2) (1995) 275–279.
[16] A. Craig, O. Cloarec, E. Holmes, J.K. Nicholson, J.C. Lindon, Scaling and normalization effects in NMR spectroscopic metabonomic data sets, Anal. Chem. 78 (7) (2006) 2262–2267.
[17] E. Szymańska, E. Saccenti, A.K. Smilde, J.A. Westerhuis, Double-check: validation of diagnostic statistics for PLS-DA models in metabolomics studies, Metabolomics 8 (1) (2012) 3–16.
[18] J. Chong, O. Soufan, C. Li, I. Caraus, S. Li, G. Bourque, D.S. Wishart, J. Xia, MetaboAnalyst 4.0: towards more transparent and integrative metabolomics analysis, Nucleic Acids Res. 46 (W1) (2018) W486–W494.
[19] M. Rizzo, K. Berneis, Low-density lipoprotein size and cardiovascular risk assessment, QJM 99 (1) (2006) 1–14.
[20] S. Bhattacharya, C.B. Granger, D. Craig, C. Haynes, J. Bain, R.D. Stevens, E.R. Hauser, C.B. Newgard, W.E. Kraus, L.K. Newby, S.H. Shah, Validation of the association between a branched chain amino acid metabolite profile and extremes of coronary artery disease in patients referred for cardiac catheterization, Atherosclerosis 232 (1) (2014) 191–196.
[21] R. Yang, J. Dong, H. Zhao, H. Li, H. Guo, S. Wang, C. Zhang, S. Wang, M. Wang, S. Yu, W. Chen, Association of branched-chain amino acids with carotid intima-media thickness and coronary artery disease risk factors, PLoS One 9 (6) (2014) e99598.
[22] S.H. Shah, J.-L. Sun, R.D. Stevens, J.R. Bain, M.J. Muehlbauer, K.S. Pieper, C. Haynes, E.R. Hauser, W.E. Kraus, C.B. Granger, C.B. Newgard, R.M. Califf, L.K. Newby, Baseline metabolomic profiles predict cardiovascular events in patients at risk for coronary artery disease, Am. Heart J. 163 (5) (2012) 844–850.e1.
[23] S.-M. Kang, J.-C. Park, M.-J. Shin, H. Lee, J. Oh, G.-S. Hwang, J.H. Chung, 1 H nuclear magnetic resonance based metabolic urinary profiling of patients with ischemic heart failure, Clin. Biochem. 44 (4) (2011) 293–299.
[24] M.-L. Cheng, C.-H. Wang, M.-S. Shiao, M.-H. Liu, Y.-Y. Huang, C.-Y. Huang, C.-T. Mao, J.-F. Lin, H.-Y. Ho, N.-I. Yang, Metabolic disturbances identified in plasma are associated with outcomes in patients with heart failure: diagnostic and prognostic value of metabolomics, J. Am. Coll. Cardiol. 65 (15) (2015) 1509–1520.
[25] Z. Du, A. Shen, Y. Huang, L. Su, W. Lai, P. Wang, Z. Xie, Z. Xie, Q. Zeng, H. Ren, D. Xu, 1H-NMR-based metabolic analysis of human serum reveals novel markers of myocardial energy expenditure in heart failure patients, PLoS One 9 (2) (2014) e88102.
[26] M. Sun, X. Gao, D. Zhang, C. Ke, Y. Hou, L. Fan, R. Zhang, H. Liu, K. Li, B. Yu, Identification of biomarkers for unstable angina by plasma metabolomic profiling, Mol. BioSyst. 9 (12) (2013) 3059–3067.
[27] E.M. Yubero-Serrano, J. Delgado-Lista, J.F. Alcala-Diaz, A. Garcia-Rios, A.I. Perez-Caballero, R. Blanco-Rojo, F. Gomez-Delgado, C. Marin, F.J. Tinahones, J. Caballero, J.M. Ordovas, B. van Ommen, F. Perez-Jimenez, P. Perez-Martinez, J. Lopez-Miranda, A dysregulation of glucose metabolism control is associated with carotid atherosclerosis in patients with coronary heart disease (CORDIOPREV-DIAB study), Atherosclerosis 253 (2016) 178–185.
[28] M. Yüksel, A. Yıldız, M. Oylumlu, A. Akyüz, M. Aydın, H. Kaya, H. Acet, N. Polat, M.Z. Bilik, S. Alan, The association between platelet/lymphocyte ratio and coronary artery disease severity, Anatolian Journal of Cardiology 15 (8) (2016) 640–647.
[29] Q. Yan, W.Q. Gu, J. Hong, Y.F. Zhang, Y.X. Su, M.H. Gui, Y. Zhang, Z.N. Chi, Y.W. Zhang, X.Y. Li, G. Ning, Coronary angiographic studies of impaired glucose regulation and coronary artery disease in Chinese nondiabetic subjects, Endocrine 36 (3) (2009) 457–463.
[30] C. Nielson, T. Lange, N. Hadjokas, Blood glucose and coronary artery disease in nondiabetic patients, Diabetes Care 29 (5) (2006) 998–1001.
[31] K.A. Cheong, N.-H. Kim, M. Noh, A.-Y. Lee, Three new single nucleotide polymorphisms identified by a genome-wide association study in Korean patients with Vitiligo, J. Korean Med. Sci. 28 (5) (2013) 775–779.
[32] T. Lotti, A.M. D'Erme, Vitiligo as a systemic disease, Clin. Dermatol. 32 (3) (2014) 430–434.
[33] Y. Gupta, A.C. Ammini, Vitiligo, hypothyroidism and cardiomyopathy, Indian J. Endocrinol. Metab. 16 (3) (2012) 463–465.
[34] O. Alsmadi, S.E. John, G. Thareja, P. Hebbar, D. Antony, K. Behbehani, T.A. Thanaraj, Genome at juncture of early human migration: a systematic analysis of two whole genomes and thirteen exomes from Kuwaiti population subgroup of inferred Saudi Arabian tribe ancestry, PLoS One 9 (6) (2014) e99069.
[35] K. Suhre, C. Gieger, Prediction of lipid-metabotype-related physiological susceptibilities, Google Patents, 2012.
[36] S.C. Connor, M.K. Hansen, A. Corner, R.F. Smith, T.E. Ryan, Integration of metabolomics and transcriptomics data to aid biomarker discovery in type 2 diabetes, Mol. BioSyst. 6 (5) (2010) 909–921.
[37] N.A. Yousri, D.O. Mook-Kanamori, M.M.E.-D. Selim, A.H. Takiddin, H. Al-Homsi, K.A. Al-Mahmoud, E.D. Karoly, J. Krumsiek, K.T. Do, U. Neumaier, A systems view of type 2 diabetes-associated metabolic perturbations in saliva, blood and urine at different timescales of glycaemic control, Diabetologia 58 (8) (2015) 1855–1867.
[38] R.M. Salek, M.L. Maguire, E. Bentley, D.V. Rubtsov, T. Hough, M. Cheeseman, D. Nunez, B.C. Sweatman, J.N. Haselden, R.D. Cox, S.C. Connor, J.L. Griffin, A metabolomic comparison of urinary changes in type 2 diabetes in mouse, rat, and human, Physiol. Genomics 29 (2) (2007) 99–108.
[39] F.K. Keating, B.E. Sobel, D.J. Schneider, Effects of increased concentrations of glucose on platelet reactivity in healthy subjects and in patients with and without diabetes mellitus, Am. J. Cardiol. 92 (11) (2003) 1362–1365.
[40] R.B. Francis Jr., D. Kawanishi, T. Baruch, P. Mahrer, S. Rahimtoola, D.I. Feinstein, Impaired fibrinolysis in coronary artery disease, Am. Heart J. 115 (4) (1988) 776–780.
[41] Y. Kitta, T. Nakamura, M. Uematsu, W. Sugamata, J. Deyama, D. Fujioka, Y. Saito, K. Kawabata, J.-e. Obata, K. Kugiyama, Insulin resistance negatively affects long-term outcome in non-diabetic patients with coronary artery disease after therapies to reduce atherosclerotic risk factors, J. Cardiol. 62 (6) (2013) 348–353.
[42] Z. Lin, C.M. Vicente Goncalves, L. Dai, H.M. Lu, J.H. Huang, H. Ji, D.S. Wang, L.Z. Yi, Y.Z. Liang, Exploring metabolic syndrome serum profiling based on gas chromatography mass spectrometry and random forest models, Anal. Chim. Acta 827 (2014) 22–27.
[43] B. Xie, M.J. Waters, H.J. Schirra, Investigating potential mechanisms of obesity by metabolomics, J Biomed Biotechnol 2012 (2012) 10.
[44] M.S.A. Palmnäs, K.A. Kopciuk, R.A. Shaykhutdinov, P.J. Robson, D. Mignault, R. Rabasa-Lhoret, H.J. Vogel, I. Csizmadi, Serum metabolomics of activity energy expenditure and its relation to metabolic syndrome and obesity, Sci. Rep. 8 (1) (2018) 3308.
[45] S. Mayengbam, J.E. Lambert, J.A. Parnell, J.M. Tunnicliffe, J. Han, T. Sturzenegger, H.J. Vogel, J. Shearer, R.A. Reimer, Dietary fiber supplementation normalizes serum metabolites of adults with overweight/obesity in a 12-week randomized control trial, The FASEB Journal 31 (1_supplement) (2017) 433.5-433.5.
[46] L. Men, Z. Pi, Y. Zhou, M. Wei, Y. Liu, F. Song, Z. Liu, Urine metabolomics of high-fat diet induced obesity using UHPLC-Q-TOF-MS, J. Pharm. Biomed. Anal. 132 (2017) 258–266.
[47] G. Anfossi, I. Russo, M. Trovati, Platelet dysfunction in central obesity, Nutr. Metab. Cardiovasc. Dis. 19 (6) (2009) 440–449.
[48] F. Santilli, N. Vazzana, R. Liani, M.T. Guagnano, G. Davi, Platelet activation in obesity and metabolic syndrome, Obes. Rev. 13 (1) (2012) 27–42.
[49] T.S. Kern, R.L. Engerman, Abnormal amino acid concentrations in plasma and urine of experimentally diabetic dogs, Res. Exp. Med. (Berl.) 182 (3) (1983) 185–192.
[50] O. Kövamees, A. Shemyakin, J. Pernow, Amino acid metabolism reflecting arginase activity is increased in patients with type 2 diabetes and associated with endothelial dysfunction, Diab. Vasc. Dis. Res. 13 (5) (2016) 354–360.
[51] S. Polakof, D. Dardevet, B. Lyan, L. Mosoni, E. Gatineau, J.-F. Martin, E. Pujos-Guillot, A. Mazur, B. Comte, Time course of molecular and metabolic events in the development of insulin resistance in fructose-fed rats, J. Proteome Res. 15 (6) (2016) 1862–1874.
[52] Y. Zhou, L. Qiu, Q. Xiao, Y. Wang, X. Meng, R. Xu, S. Wang, R. Na, Obesity and diabetes related plasma amino acid alterations, Clin. Biochem. 46 (15) (2013) 1447–1452.
[53] H. Kim, M. Lee, H. Park, Y. Park, J. Shon, K.-H. Liu, C. Lee, Urine and serum metabolite profiling of rats fed a high-fat diet and the anti-obesity effects of caffeine consumption, Molecules 20 (2) (2015) 3107.
[54] M.T. Janis, R. Laaksonen, M. Oresic, Metabolomic strategies to identify tissue-specific effects of cardiovascular drugs, Expert Opin. Drug Metab. Toxicol. 4 (6) (2008) 665–680.
[55] R. Laaksonen, M. Katajamaa, H. Päivä, M. Sysi-Aho, L. Saarinen, P. Junni, D. Lütjohann, J. Smet, R. Van Coster, T. Seppänen-Laakso, T. Lehtimäki, J. Soini, M. Orešič, A systems biology strategy reveals biological pathways and plasma biomarker candidates for potentially toxic statin-induced changes in muscle, PLoS One 1 (1) (2006) e97.
[56] J.-A. Thomas, L. Gerber, L.L. Bañez, D.M. Moreira, R.S. Rittmaster, G.L. Andriole, S.J. Freedland, Prostate cancer risk in men with baseline history of coronary artery disease: results from the reduce study, Cancer Epidemiology Biomarkers Prevention 21 (4) (2012) 576–581.
[57] O.F. Karatas, O. Bayrak, E. Cimentepe, D. Unal, An insidious risk factor for cardiovascular disease: benign prostatic hyperplasia, Int. J. Cardiol. 144 (3) (2010) 452.
[58] E.J. Lesnefsky, S. Moghaddas, B. Tandler, J. Kerner, C.L. Hoppel, Mitochondrial dysfunction in cardiac disease: ischemia–reperfusion, aging, and heart failure, J. Mol. Cell. Cardiol. 33 (6) (2001) 1065–1089.
[59] W.G. Hunter, J.P. Kelly, R.W. McGarrah III, M.G. Khouri, D. Craig, C. Haynes, O. Ilkayeva, R.D. Stevens, J.R. Bain, M.J. Muehlbauer, Metabolomic profiling identifies novel circulating biomarkers of mitochondrial dysfunction differentially elevated in heart failure with preserved versus reduced ejection fraction: evidence for shared metabolic impairments in clinical heart failure, J. Am. Heart Assoc. 5 (8) (2016) e003190.
[60] P. Bullon, H.N. Newman, M. Battino, Obesity, diabetes mellitus, atherosclerosis and chronic periodontitis: a shared pathology via oxidative stress and mitochondrial dysfunction? Periodontol. 64 (1) (2014) 139–153.
[61] P. Vernocchi, F. Del Chierico, L. Putignani, Gut microbiota profiling: metabolomics based approach to unravel compounds affecting human health, Front. Microbiol. 7 (1144) (2016).
[62] E. Org, Y. Blum, S. Kasela, M. Mehrabian, J. Kuusisto, A.J. Kangas, P. Soininen, Z. Wang, M. Ala-Korpela, S.L. Hazen, M. Laakso, A.J. Lusis, Relationships between

gut microbiota, plasma metabolites, and metabolic syndrome traits in the METSIM cohort, Genome Biol. 18 (1) (2017) 70.
[63] W.H. Tang, Z. Wang, X.S. Li, Y. Fan, D.S. Li, Y. Wu, S.L. Hazen, Increased trimethylamine N-oxide portends high mortality risk independent of glycemic control in patients with type 2 diabetes mellitus, Clin. Chem. 63 (1) (2017) 297–306.
[64] P. Hernández-Alonso, D. Cañueto, S. Giardina, J. Salas-Salvadó, N. Cañellas, X. Correig, M. Bulló, Effect of pistachio consumption on the modulation of urinary gut microbiota-related metabolites in prediabetic subjects, J. Nutr. Biochem. 45 (2017) 48–53.
[65] M.-l. Chen, L. Yi, Y. Zhang, X. Zhou, L. Ran, J. Yang, J.-d. Zhu, Q.-y. Zhang, M.-t. Mi, Resveratrol attenuates trimethylamine-N-oxide (TMAO)-induced atherosclerosis by regulating TMAO synthesis and bile acid metabolism via remodeling of the gut microbiota, MBio 7 (2) (2016) e02210–e02215.
[66] P. Bernini, I. Bertini, C. Luchinat, L. Tenori, A. Tognaccini, The cardiovascular risk of healthy individuals studied by NMR metabonomics of plasma samples, J. Proteome Res. 10 (11) (2011) 4983–4992.