Association of a Peripheral Blood Metabolic Profile With Coronary Artery Disease and Risk of Subsequent Cardiovascular Events

Svati H. Shah, MD, MHS; James R. Bain, PhD; Michael J. Muehlbauer, PhD; Robert D. Stevens, PhD; David R. Crosslin, PhD; Carol Haynes, BS; Jennifer Dungan, RN, PhD; L. Kristin Newby, MD, MHS; Elizabeth R. Hauser, PhD; Geoffrey S. Ginsburg, MD, PhD; Christopher B. Newgard, PhD; William E. Kraus, MD

Background—Molecular tools may provide insight into cardiovascular risk. We assessed whether metabolites discriminate coronary artery disease (CAD) and predict risk of cardiovascular events.

Methods and Results—We performed mass–spectrometry–based profiling of 69 metabolites in subjects from the CATHGEN biorepository. To evaluate discriminative capabilities of metabolites for CAD, 2 groups were profiled: 174 CAD cases and 174 sex/race-matched controls ("initial"), and 140 CAD cases and 140 controls ("replication"). To evaluate the capability of metabolites to predict cardiovascular events, cases were combined ("event" group); of these, 74 experienced death/myocardial infarction during follow-up. A third independent group was profiled ("event-replication" group; n=63 cases with cardiovascular events, 66 controls). Analysis included principal-components analysis, linear regression, and Cox proportional hazards. Two principal components analysis–derived factors were associated with CAD: 1 comprising branched-chain amino acid metabolites (factor 4, initial $P=0.002$, replication $P=0.01$), and 1 comprising urea cycle metabolites (factor 9, initial $P=0.0004$, replication $P=0.01$). In multivariable regression, these factors were independently associated with CAD in initial (factor 4, odds ratio [OR], 1.36; 95% CI, 1.06 to 1.74; $P=0.02$; factor 9, OR, 0.67; 95% CI, 0.52 to 0.87; $P=0.003$) and replication (factor 4, OR, 1.43; 95% CI, 1.07 to 1.91; $P=0.02$; factor 9, OR, 0.66; 95% CI, 0.48 to 0.91; $P=0.01$) groups. A factor composed of dicarboxylacylcarnitines predicted death/myocardial infarction (event group hazard ratio 2.17; 95% CI, 1.23 to 3.84; $P=0.007$) and was associated with cardiovascular events in the event-replication group (OR, 1.52; 95% CI, 1.08 to 2.14; $P=0.01$).

Conclusions—Metabolite profiles are associated with CAD and subsequent cardiovascular events. (***Circ Cardiovasc Genet*. 2010;3:207-214.**)

Key Words: metabolism ■ risk factors ■ coronary artery disease

Coronary artery disease (CAD) is the leading cause of death in industrialized countries. Many accepted risk factors for CAD are metabolic. However, there remains an incomplete mechanistic understanding of CAD risk and equally important, a need to refine our ability to identify individuals at highest risk of cardiovascular events. Given the complex nature of CAD, evaluation with more comprehensive tools may improve risk stratification and enhance our understanding of the disease process. Metabolomics, the study of small-molecule metabolites, may be particularly useful for the diagnosis of human disease. Studies have demonstrated heritability of metabolites in mice,[1] and we have shown that metabolite profiles are heritable in human families with early-onset CAD,[2] suggesting that the known heritability of CAD may be mediated at least in part through metabolic components measurable in peripheral blood.

Clinical Perspective on p 214

In this study, we performed quantitative profiling of 69 metabolites, including acylcarnitine species (byproducts of mitochondrial fatty acid, carbohydrate, and amino acid oxidation), amino acids, and conventional metabolites, in participants enrolled in the Duke CATHGEN biorepository. Our primary goals were to assess the capability of metabolite

Received January 24, 2009; accepted February 1, 2010.

From the Departments of Medicine (S.H.S., L.K.N., W.E.K.), Nursing (J.D.), and Pharmacology and Cancer Biology (C.B.N.); Center for Human Genetics (S.H.S., D.R.S., C.H., E.R.H.); Sarah W. Stedman Nutrition and Metabolism Center (J.R.B., M.J.M., R.D.S., C.B.N.); Duke Clinical Research Institute (L.K.N.); and Duke Institute for Genome Sciences and Policy (G.S.G.), Duke University, Durham, NC.

Guest Editor for this article was Donna K. Arnett, PhD.

The online-only Data Supplement is available at http://circgenetics.ahajournals.org/cgi/content/full/CIRCGENETICS.109.852814/DC1.

Correspondence to Svati Shah, Duke University Medical Center 3445, Durham, NC 27710. E-mail svati.shah@duke.edu

Circ Cardiovasc Genet **is available at http://circgenetics.ahajournals.org** **DOI: 10.1161/CIRCGENETICS.109.852814**

profiles to discriminate the presence of CAD and to explore the relation of metabolite profiles with subsequent cardiovascular events.

Methods

Study Sample

The CATHGEN biorepository consists of subjects recruited sequentially through the cardiac catheterization laboratories at Duke University Medical Center (Durham, NC). After informed consent was obtained, blood was drawn from the femoral artery at the time of arterial access for catheterization, immediately processed to separate plasma, and frozen at −80°C. Subjects were fasting for a minimum of 6 hours before collection. Clinical data were provided by the Duke Databank for Cardiovascular Disease (DDCD), a database of patients undergoing catheterization at Duke University since 1969. Data were collected for long-term medications, that is, medications at admission (inpatients) or from a clinic note within 1 month prior (outpatients). Follow-up data, including occurrence of myocardial infarction (MI) and death, were collected at 6 months after catheterization and annually thereafter. Vital status was confirmed through the National Death Index. The indication for catheterization for all subjects was clinical concern for ischemic heart disease. Patients with severe pulmonary hypertension or organ transplant were excluded.

To evaluate the discriminative capability of metabolites for CAD, 2 independent case-control groups were constructed: "initial" (174 CAD cases and 174 CAD-free controls) and "replication" (140 CAD cases and 140 CAD-free controls). For the initial group, sequential cases meeting inclusion criteria were selected: $CAD_{index} \geq 32$ (at least 1 coronary artery with ≥95% stenosis) and age –of onset ≤55 years. CAD_{index} is a numeric summary of angiographic data.[3] Age –of onset was defined as age at first MI, percutaneous coronary intervention, coronary artery bypass grafting (CABG), or at first catheterization meeting the CAD_{index} threshold. Sex- and race-matched controls meeting the following criteria were selected: $CAD_{index} \leq 23$; no coronary artery with >50% stenosis; age –at catheterization ≥61 years; and no history of MI, percutaneous coronary intervention, CABG, or transplant. Given differences in age based on these criteria, results could be confounded by age. Therefore, for the replication group, sequential subjects meeting the same inclusion criteria were selected, but the criterion of age –of onset (cases) or age –at catheterization (controls) was removed and cases/controls were not matched, thereby allowing generalizability of findings to a representative population of patients referred for catheterization. Analyses were also performed by constraining cases to those with a history of MI (n=86 cases in initial group, n=61 in replication group).

To evaluate the capability of metabolites to predict risk of subsequent events, an "event" group was constructed by combining CAD cases from the initial and replication groups (event group, n=314); of these, 74 individuals experienced death or MI during follow-up. To validate findings for the association of metabolites with risk of cardiovascular events, profiling was performed in an independent event case-control group ("event-replication") composed of unique individuals from CATHGEN and meeting the following criteria: ejection fraction >40%, no history of percutaneous coronary intervention or CABG, and no subsequent CABG. Among these, event cases (n=63) experienced death or MI or had percutaneous coronary intervention with acute coronary syndrome within 2 years after catheterization; controls (n=66) were event-free with at least 2 years of follow-up and were matched to cases on age, race, sex, and CAD_{index}.

The Duke Institutional Review Board approved the protocols for CATHGEN and for this study. Informed consent was obtained from each subject.

Metabolite Measurements

Fasting plasma samples were used for quantitative determination of targeted levels for 45 acylcarnitines, 15 amino acids, 5 conventional metabolites (total, LDL, and HDL cholesterol; triglycerides; and glucose), ketones, β-hydroxybutyrate, total free fatty acids, and C-reactive protein (CRP) (online-only Data Supplement 1). Methodology and coefficients of variation have been reported.[2,4] The laboratory (Sarah W. Stedman Nutrition and Metabolism metabolomics/biomarker core) was blinded to case-control status, and the cases/controls were randomly distributed.

Standard methods were used for conventional metabolites with reagents from Roche Diagnostics (Indianapolis, Ind), and for free fatty acids (total) and ketones (total and β-hydroxybutyrate) with reagents from Wako (Richmond, Va). All assays were performed on a Hitachi 911 clinical chemistry analyzer.

For mass spectroscopy–profiled metabolites (acylcarnitines, amino acids), the following protocol was used.[5,6] Proteins were first removed by precipitation with methanol. Aliquots of supernatants were dried and then esterified with hot, acidic methanol (acylcarnitines) or *n*-butanol (amino acids). Analysis was done by tandem mass spectroscopy with a Quattro Micro instrument (Waters Corp, Milford, Mass). Quantification of the "targeted" intermediary metabolites was facilitated by addition of mixtures of known quantities of stable-isotope internal standards (Supplement 2). Details are included in Supplement 3; biological annotation is included in Supplement 4.

Statistical Analysis

Metabolite levels reported as "0" (ie, below the lower limits of quantification) were given a value of limits of quantification/2. Metabolites with >25% of values as "0" were not analyzed (5 acylcarnitines). Metabolites were natural-log transformed to approximate a normal distribution. For analysis of CAD status, generalized linear-regression models were used to assess differences in metabolite levels between CAD cases and controls, unadjusted and adjusted for CAD risk factors not constrained by matching: diabetes, hypertension, dyslipidemia, body mass index (BMI), family history of CAD, and smoking. Analyses of the replication group were further adjusted for race, sex, and age. With log transformation, all significant metabolites showed a normal distribution (Kolmogorov-Smirnov test $P>0.01$), except valine, ketones, and C8, C8:1-OH/C6:1-DC, C10:1, C14:2, C16:1, C16:1-OH/C14:1-DC, and C18-OH/C16-DC acylcarnitines. Visual inspection suggested a grossly normal distribution (Supplement 7). Regardless, we performed sensitivity analyses with nonparametric Wilcoxon tests, which showed results similar to those of the semiparametric linear models, except for valine and C14:2 acylcarnitine, which were not significant in linear regressions ($P=0.10$ and $P=0.06$, respectively) but were significant with these nonparametric tests ($P=0.05$ and $P=0.008$). Analyses were also stratified by diabetes and smoking status.

In exploratory analyses, multivariable models were further adjusted for medications (β-blockers, statins, diabetes medications, aspirin, angiotensin-converting-enzyme inhibitors, nitrates, clopidogrel, and diuretics), use of preprocedural sedation, and continuous intravenous heparin use at time of catheterization. The CATHGEN protocol requires sample collection before supplemental heparin administration during catheterization. Therefore, adjustment for continuous intravenous heparin use at the time of catheterization addresses differences related to heparin. Only 66% of individuals had medication data, hence, medications were coded as not on medication, missing, and on medication.

Given that metabolites reside in overlapping pathways, correlation of metabolites is expected. We used principal-components analysis (PCA) to reduce the large number of correlated variables (Supplement 5) into uncorrelated factors. Factors with higher "eigenvalues" account for larger amounts of variability. Factors with an eigenvalue ≥1.0 were identified, and varimax rotation was performed to produce interpretable factors. Metabolites with a factor load ≥0.4 were reported as composing a factor. Scoring coefficients were constructed from the initial group and used to calculate factor scores for each individual (weighted sum of the standardized metabolites within that factor, weighted on the factor loading for each metabolite) and were also applied to the replication group. Generalized linear-regression models were used to assess the difference in factors

Table 1. Baseline Clinical Characteristics

	Initial Group			Replication Group		
	Cases (n=174)	Controls (n=174)	*P**	Cases (n=140)	Controls (n=140)	*P**
Age, mean (SD), y	48.7 (10.0)	67.8 (5.9)	<0.0001	61.1 (13.0)	60.3 (13.0)	0.69
Age-of-onset, mean (SD), y	45.8 (6.9)	N/A	N/A	57.0 (10.8)	N/A	N/A
Sex, % male	77.0	75.3	0.71	75.0	51.4	<0.0001
Race, % white	66.9	67.8	0.97	77.5	76.9	0.90
Hypertension, %	64.9	68.4	0.50	72.9	64.3	0.12
Diabetes, %	32.2	23.0	0.06	28.6	19.3	0.07
Family history of CAD, %	57.5	22.4	<0.0001	49.3	32.9	0.005
Currently smoking, %	66.1	47.1	0.0004	64.3	39.3	<0.0001
Body mass index, mean (SD)	31.1 (6.8)	29.3 (6.0)	0.02	29.1 (5.8)	31.4 (9.1)	0.03
CAD_{index}, mean (SD)	56.5 (21.9)	6.2 (9.4)	<0.0001	58.6 (21.8)	4.6 (8.6)	<0.0001
No. of coronary arteries w/ ≥75% stenosis, %			<0.0001			<0.0001
0	0	100		0	100	
1	27.0	0		22.1	0	
2	29.3	0		35.0	0	
3	43.7	0		42.9	0	
Ejection fraction, mean (SD)	51.8 (13.0)	59.4 (13.0)	<0.0001	53.5 (15.1)	63.4 (9.5)	<0.0001
History of MI, %	49.4	0	<0.0001	42.9	0	<0.0001
History of dyslipidemia, %	73.6	44.8	<0.0001	60.7	49.3	0.05
Total cholesterol, mean (SD)	178.5 (55.1)	176.1 (39.2)	0.54	177.9 (44.4)	169.9 (38.7)	0.14
LDL cholesterol, mean (SD)	105.3 (39.3)	104.2 (32.1)	0.64	105.9 (36.8)	101.1 (32.5)	0.37
HDL cholesterol, mean (SD)	35.6 (10.8)	48.0 (16.0)	<0.0001	39.3 (12.2)	39.0 (12.6)	0.70
Triglycerides, mean (SD)	157.7 (170.4)	93.2 (60.7)	<0.0001	128.1 (82.4)	119.8 (91.1)	0.17

**P* value for difference between cases and controls.

between cases and controls. All factors were normally distributed (Kolmogorov-Smirnov test $P>0.01$), except for factors 7 to 9; visual inspection showed a grossly normal distribution (Supplement 7). Nonparametric Wilcoxon tests for these factors showed the same results as did linear models.

To further assess the independent capability of metabolite profiles to discriminate CAD cases from controls, logistic-regression models were constructed; in these models, CAD risk factors (BMI, dyslipidemia, hypertension, diabetes, family history, and smoking) were forced into the model and then metabolite factors were added. Receiver operating characteristic (ROC) curves were constructed and measures of model fit calculated. Nonparametric analysis for comparison of the areas under these curves was performed with previously reported methods.[7]

For analysis of subsequent cardiovascular events, cases from initial and replication groups were pooled (event group). The relation between metabolite factors and time –to occurrence of death/MI was assessed with Cox proportional hazards (unadjusted and adjusted for BMI, dyslipidemia, hypertension, diabetes, family history, smoking, age, race, sex, creatinine, ejection fraction, and CAD_{index}). The assumption of proportional hazards was met. For replication in the event-replication group, scoring coefficients from PCA-derived factors constructed in the initial CAD group were used to calculate factor scores in the event-replication group; logistic regression was used to assess the association between factors and case-control status (unadjusted and adjusted for BMI, dyslipidemia, hypertension, diabetes, family history, smoking, creatinine, and ejection fraction).

Because all analyses were exploratory in nature and given the colinearity of the metabolites, 2-sided probability values unadjusted for multiple comparisons are presented; however, results interpreted in the context of Bonferroni correction are reported. Nominal significance was defined as $P\leq0.05$. Bonferroni-corrected probability values were $P<0.0007$ (individual metabolites) and $P<0.004$ (factors). Statistical analyses were performed by D.R.C. and S.H.S. with SAS version 9.1 (Cary, NC). The authors had full access to the data and take responsibility for its integrity. All authors have read and agreed to the manuscript as written.

Results

Patient Populations

Population characteristics for the initial (174 early-onset CAD cases, 174 matched controls) and replication groups (140 CAD cases, 140 controls) are displayed in Table 1 and for the event-replication group in Supplement 8.

Association of Individual Metabolites With CAD

Levels of several amino acids were different between cases and controls in the initial group (Table 2), including the branched-chain amino acids leucine/isoleucine ($P<0.0001$) and valine ($P=0.007$), glutamate/glutamine ($P<0.0001$), proline ($P=0.04$), and methionine ($P=0.05$). Levels of several acylcarnitines were also different between cases and controls in the initial group, including the C16 acylcarnitines (C16:1, $P=0.006$; C16:1-OH/C14:1-DC, $P=0.004$; C16:2, $P=0.05$; and C18-OH/C16-DC, $P=0.003$), C4:Ci4 ($P=0.009$), C8 ($P=0.009$), C8:1-OH/C6:1-DC ($P=0.003$), and C10:1 ($P=0.002$) acylcarnitines (Table 2). For most metabolites, these differences persisted after adjustment for CAD risk factors.

Several of these metabolites were also significant in the replication group in adjusted analyses (with similar direction

Table 2. Association of Individual Metabolites With CAD

	Initial Group				Replication Group			
Metabolite	CAD Cases	Controls	Unadjusted *P*	Adjusted *P**	Cases	Controls	Unadjusted *P*	Adjusted *P*†
Amino acids								
PRO	190.2 (56.4)	177.5 (41.3)	0.04	0.13	197.0 (75.9)	173.9 (44.3)	0.001	0.03
LEU/ILE	**175.1 (39.5)**	**158.7 (36.3)**	**<0.0001**	**0.004**	**183.6 (52.7)**	**162.3 (35.5)**	**<0.0001**	**0.002**
VAL	259.1 (58.1)	242.2 (54.9)	0.007	0.26	256.8 (63.7)	266.2 (51.5)	0.10	0.05
MET	25.8 (5.6)	24.6 (4.8)	0.05	0.14	26.6 (7.7)	24.0 (5.2)	0.003	0.03
GLX	**151.1 (41.7)**	**125.5 (39.0)**	**<0.0001**	**<0.0001**	**129.7 (30.5)**	**120.3 (31.4)**	**0.005**	**0.02**
CIT	36.4 (12.2)	39.9 (11.2)	0.002	0.003	39.7 (12.0)	37.8 (10.8)	0.21	0.86
Acylcarnitines								
C2	9.10 (4.25)	9.90 (3.98)	0.02	0.01	10.73 (4.86)	8.76 (3.83)	<0.0001	<0.0001
C4:Ci4	0.22 (0.12)	0.18 (0.10)	0.009	0.03	0.22 (0.11)	0.20 (0.11)	0.06	0.15
C5	0.104 (0.095)	0.087 (0.047)	0.01	0.08	0.092 (0.045)	0.101 (0.045)	0.05	0.004
C8	0.107 (0.057)	0.123 (0.076)	0.009	0.04	0.129 (0.121)	0.124 (0.106)	0.54	0.17
C8:1-OH/C6:1-DC	0.030 (0.027)	0.032 (0.019)	0.003	0.005	0.026 (0.013)	0.028 (0.012)	0.47	0.56
C10:1	0.174 (0.097)	0.193 (0.083)	0.002	0.002	0.200 (0.096)	0.198 (0.096)	0.87	0.47
C14:2	0.039 (0.027)	0.044 (0.028)	0.02	0.02	0.039 (0.028)	0.047 (0.025)	0.06	0.42
C14:1-OH	**0.012 (0.006)**	**0.014 (0.007)**	**0.04**	**0.03**	**0.012 (0.006)**	**0.015 (0.007)**	**0.002**	**0.006**
C16:2	0.0088 (0.0065)	0.0092 (0.0053)	0.05	0.03	0.0098 (0.0059)	0.0117 (0.007)	0.03	0.11
C16:1	0.0258 (0.0171)	0.0262 (0.0124)	0.006	0.07	0.0286 (0.0154)	0.0321 (0.0138)	0.03	0.24
C16:1-OH/C14:1-DC	**0.0087 (0.0036)**	**0.0091 (0.0031)**	**0.004**	**0.01**	**0.0088 (0.0040)**	**0.0096 (0.0040)**	**0.008**	**0.01**
C18-OH/C16-DC	0.008 (0.009)	0.007 (0.004)	0.003	0.02	0.007 (0.003)	0.008 (0.004)	0.005	0.03
Ketones	289.8 (345.0)	324.3 (286.0)	0.04	0.14	319.2 (289.9)	313.1 (279.0)	0.87	0.44
β-Hydroxybutyrate	199.7 (271.6)	237.1 (235.3)	0.01	0.05	211.5 (205.7)	202.9 (199.8)	0.63	0.29

Means and SDs for metabolites significantly different between cases and controls in the initial group are presented. Results for these analytes in the replication group are also presented. All values are in millimolar for amino acids and micromolar for acylcarnitines. Analytes in bold show association across both datasets in adjusted analyses (with consistent direction of effect).

*Adjusted for diabetes, hypertension, smoking, dyslipidemia, family history of CAD, and BMI.

†Adjusted for age, race, sex, diabetes, hypertension, smoking, dyslipidemia, family history of CAD, and BMI.

of effect), including the amino acids leucine/isoleucine and glutamate/glutamine, and the long-chain acylcarnitines C14:1-OH and C16:1-OH/C14:1-DC (Table 2). In unadjusted analyses, these metabolites, amino acids methionine and proline, and C16:2 and C16:1 acylcarnitine were significant in both groups.

Further adjustment for lipids (total, LDL, and HDL cholesterol, and triglycerides) resulted in similar results, although with attenuation of association for leucine/isoleucine in the initial group (Supplements 9 and10). Analyses stratified by diabetes (Supplements 11 and 12) suggested some heterogeneity of association by diabetes. For example, leucine/isoleucine and C16:1-OH/C14:1-DC showed stronger associations in nondiabetics. Analyses stratified by smoking status suggested no difference in smokers and nonsmokers (Supplements 14 and 15).

Unbiased PCA

PCA identified 12 factors, comprising collinear metabolites (Table 3), grouping in biologically plausible factors. Three factors were significantly different between cases and controls in the initial group in adjusted analyses: factor 1 (medium-chain acylcarnitines), factor 4 (branched-chain amino acids and related metabolites), and factor 9 (arginine, histidine, citrulline, and Ci4-DC:C4DC). Of these factors, 2 factors (4 and 9) remained 'significant in the replication group. Factor 1 was only weakly significant in the replication group (unadjusted $P=0.15$, adjusted $P=0.03$).

Further adjustment for lipids showed a continued association with CAD, although factor 4 was not significant in the initial group (initial group: factor1, $P=0.0002$; factor 4, $P=0.59$; factor 9, $P=0.02$; replication group: factor1, $P=0.01$; factor 4, $P=0.02$; and factor 9, $P=0.004$). Although we adjusted for diabetes, given studies showing relations between metabolites with insulin resistance,[4,8,9] we further adjusted the base multivariable model for fasting glucose. These analyses revealed a continued significant association with CAD (initial group: factor 1, $P=0.02$; factor 4, $P=0.02$; factor 9, $P=0.003$; replication group: factor 1, $P=0.03$; factor 4, $P=0.05$; factor 9, $P=0.02$).

Stratified analyses suggested a stronger association between factors 4 and 9 with CAD in nondiabetics compared with diabetics (Supplement 13), with minimal/no association in diabetics and no consistent differences in association with CAD by smoking status (Supplement 16).

Additional adjustment for 10 classes of medications had minimal influence on the relation between factors and CAD in the initial group (factor 1, $P=0.009$; factor 4, $P=0.03$;

Table 3. Principal Components Analysis

					Initial Group				Replication Group			
					CAD		MI		CAD		MI	
Factor	Name	Individual Components*	Eigen-Value	Var**	Unadjusted	Adjusted†	Unadjusted	Adjusted†	Unadjusted	Adjusted†	Unadjusted	Adjusted†
1	Medium-chain acylcarnitines	C8, C10:1, C12, C10, C12:1, C10-OH:C8DC, C6-DC, C8:1-DC, C14:1, C14:2, C8:1-OH/C6:1-DC, C2 acylcarnitines	12.45	0.21	0.001	0.01	0.01	0.06	0.15	0.03	0.59	0.18
2	Long-chain acylcarnitines	C18:1, C18:2, C18, C16, C16:1, C20:4, C14:1, C14:2, C16:2, C14:1-OH	5.78	0.10	0.28	0.34	0.21	0.14	0.03	0.01	0.08	0.05
3	Long-chain dicarboxyl/hydroxyl acylcarnitines	C18-OH/C16-DC, C20-OH/C18-DC, C20:1-OH/C18:1-DC, C16-OH/C14-DC, C18:1-OH/C16:1-DC, C14-OH/C12-DC, C12-OH:C10-DC, C14:1-OH, C20	4.75	0.08	0.10	0.36	0.04	0.13	<0.0001	0.004	0.03	0.21
4	BCAA related	Phe, Tyr, leu/Ile, Met, Val, C5, Ala	2.87	0.05	0.002	0.02	0.0002	0.01	0.01	0.03	0.006	0.005
5	Ketone related	Ket, Hbut, Ala (-), C2, C4:OH, C14:1	2.24	0.04	0.18	0.33	0.02	0.12	0.54	0.41	0.07	0.06
6	Various	C8:1, C10:3	1.92	0.03	0.56	0.75	0.92	0.28	0.79	0.92	0.76	0.89
7	Amino acids	Ser, Gly, FFA (-)	1.71	0.03	0.19	0.13	0.59	0.42	0.04	0.18	0.28	0.60
8	Dicarboxyls	C5-DC, C8:1-OH/C6:1-DC, Cit, C8:1-DC, C6-DC	1.41	0.02	0.73	0.34	0.59	0.25	0.05	0.57	0.002	0.04
9	Urea cycle related	Arg, His, Cit, Ci4-DC:C4DC (-)	1.33	0.02	0.0004	0.004	0.0006	0.01	0.01	0.01	0.003	0.006
10	Short-chain acylcarnitines	C3, C4:Ci4, C5	1.22	0.02	0.02	0.19	0.03	0.23	0.72	0.92	0.27	0.48
11	Various	C5:1, C18:2-OH (-), C22 (-)	1.15	0.02	0.62	0.13	0.95	0.13	0.03	0.01	0.13	0.12
12	Various	Asx, C22	1.08	0.02	0.12	0.83	0.15	0.80	<0.0001	<0.0001	0.01	0.05

Results of PCA are presented. *P* values for the difference in the mean value of the factors between cases and controls for the initial and replication groups are presented.

*Analytes with factor load ≥0.4 for that factor are listed, in order of magnitude of load for that factor; analytes with a negative factor load for that factor are annotated with a (-).

**Proportion of variance explained by that factor.

†Adjusted for diabetes, hypertension, smoking, dyslipidemia, family history, and BMI; replication group additionally adjusted for age, race, and sex.

factor 9, $P=0.003$) but were no longer significant in the replication group (factor 1, $P=0.02$; factor 4, $P=0.19$; factor 9, $P=0.14$). We also performed similar analyses, restricted to those individuals with available medication data, in the combined datasets to optimize power (n=416). These results showed a continued association between factors 4 and 9 with CAD (factor 4: unadjusted model, $P=0.0009$; model adjusted for CAD risk factors, $P=0.03$; model adjusted for CAD risk factors and medications, $P=0.05$; factor 9: unadjusted model, $P=0.0003$; model adjusted for CAD risk factors, $P=0.002$; model adjusted for CAD risk factors and medications, $P=0.007$).

Results presented are unadjusted for multiple comparisons. We used PCA to account for the colinearity of metabolites. Of the individual metabolites, only glutamate/glutamine would survive Bonferroni correction. Factors 4 and 9 would survive Bonferroni correction at the level of factors ($P<0.004$).

Association of Metabolite Profiles With Prevalent MI

To examine the association of these metabolites with a more severe phenotype, we evaluated the relation of the PCA-derived factors in cases with a prior history of MI compared with CAD-free controls (initial group n=86 MI cases, replication group n=61 MI cases). The 2 factors (4 and 9) associated with CAD were also associated with MI in both groups (Table 3).

Assessment of Model Fit and ROC Curves for CAD

To further quantify the independent association of metabolite factors with CAD, the following logistic-regression models were constructed: (1) clinical model; (2) clinical model plus factors 4 and 9; and (3) clinical model plus all metabolite factors. Factors 4 and 9 were independently associated with CAD in both the initial group (factor 4: odds ratio [OR], 1.42; 95% CI, 1.09 to 1.84, $P=0.01$; factor 9: OR, 0.69, 95% CI, 0.53 to 0.90, $P=0.006$) and the replication group (factor 4: OR, 1.42; 95% CI, 1.06 to 1.89, $P=0.02$; factor 9: OR, 0.67; 95% CI, 0.48 to 0.92, $P=0.01$). Measure of model fit and ROC curves (Figure 1) in the initial group showed modestly greater discriminative capability for models containing factors 4 and 9 (c-statistic 0.778), with some improvement with the addition of all factors (c-statistic 0.804) compared with the model containing clinical variables only (c-statistic 0.756; $P=0.06$ for comparison of clinical model with clinical model plus factors 4 and 9; $P=0.003$ for comparison of clinical model with clinical model plus all factors). In the replication group, there was a slightly higher c-statistic with the addition of factors 4 and 9 to the clinical model (c-statistic 0.773) than for the clinical model alone (c-statistic 0.743) but more dramatic improvement with addition of all factors (c-statistic 0.874; $P=0.04$ for comparison of clinical model to clinical model plus factors 4 and 9; and $P<0.0001$ for comparison of clinical model to clinical model plus all factors).

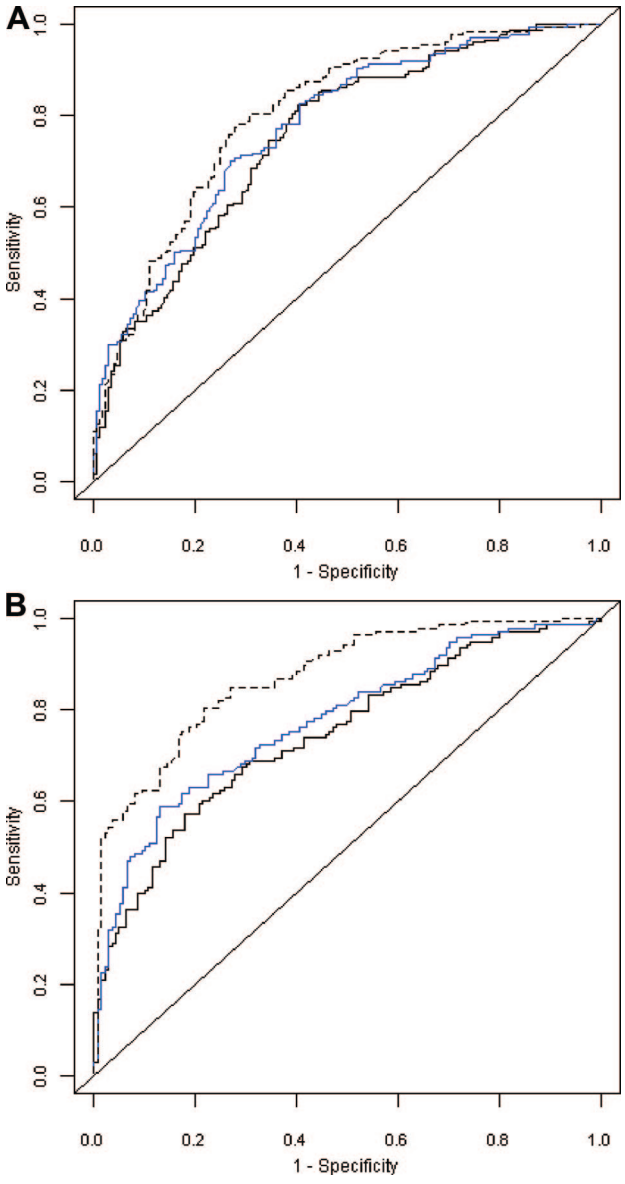

Figure 1. ROC curves for metabolite factors and CAD. ROC curves and measures of model fit (c-statistic) are presented for: clinical model inclusive of CAD risk factors (diabetes, hypertension, dyslipidemia, smoking, BMI, family history; for the replication group, age, race, and sex are also included, black solid line); a model inclusive of traditional risk factors plus metabolite factors 4 and 9 (blue line); and a model inclusive of traditional risk factors plus all metabolite factors (black, dashed line), for both the initial (A) and replication (B) groups.

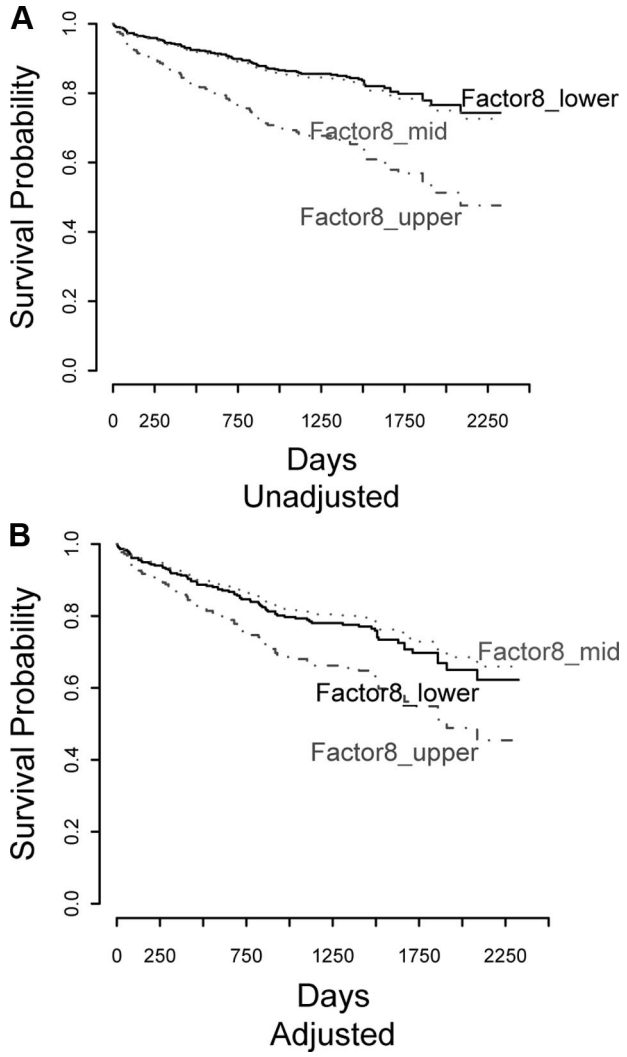

Figure 2. Cox proportional-hazards model for predictive capability of metabolite factor 8 for cardiovascular events. Unadjusted (A) and adjusted (B) survival curves (adjusted for BMI, CAD severity, hypertension, dyslipidemia, diabetes, smoking, family history, ejection fraction, serum creatinine, subsequent CABG, age, race, and sex) are presented for factor 8.

Given that it is a standard of care to measure lipids in patients in whom a diagnosis of CAD is being considered and that CRP is a recognized biomarker of cardiovascular disease, we reconstructed these models including lipids and CRP. These analyses revealed a higher clinical model fit in both initial and replication groups (c-statistic 0.842 and 0.778, respectively). The addition of factors 4 and 9 to the clinical model inclusive of lipids and CRP resulted in no improvement in the discriminative ability of the model in the initial group (c-statistic 0.848, $P=0.31$ for comparison with clinical model), with some improvement with the addition of all factors (c-statistic 0.865, $P=0.01$ for comparison with clinical model). However, the magnitude of improvement in the clinical model with the addition of metabolite factors remained similar and large in the replication group (c-statistics: clinical model inclusive of lipids and CRP, 0.778; clinical model plus factors 4 and 9, 0.799, $P=0.08$; clinical model plus all metabolite factors, 0.900, $P=0.0001$ for comparison with clinical model).

Metabolite Factors and Risk of Subsequent Cardiovascular Events

During a median of 2.72 years of follow-up, 74 of 314 CAD cases had an incident cardiovascular event. In unadjusted comparisons, factor 8 (short-chain dicarboxylacylcarnitines) was highly associated with occurrence of death or MI (Figure 2; highest vs lowest tertile hazard ratio [HR], 2.50; 95% CI, 1.47 to 4.17; $P=0.0008$; highest vs middle tertile HR, 2.33; 95% CI, 1.39 to 3.85; $P=0.002$). The strength of this association was somewhat attenuated after adjustment for

CAD risk factors, CAD_{index}, age, race, sex, ejection fraction, creatinine, and treatment with CABG after catheterization (highest vs lowest tertile: HR, 1.67; 95% CI, 0.88 to 3.13; $P=0.11$; highest vs middle tertile: HR, 1.89; 95% CI, 1.09 to 3.33; $P=0.03$). Factor 1 was also associated with the occurrence of death/MI (highest vs lowest tertile HR, 1.85; 95% CI, 1.06 to 3.23; $P=0.03$; highest vs middle tertile HR, 1.79; 95% CI, 1.02 to 3.03; $P=0.04$) but was no longer significant after adjustment ($P=0.14$ and 0.05, respectively).

To validate these findings, we performed metabolomic profiling in an independent dataset (event-replication group). Factor 8 was associated with cardiovascular events (unadjusted OR, 1.52; 95% CI, 1.08 to 2.14; $P=0.01$; adjusted OR, 1.82; 95% CI, 1.08 to 3.50; $P=0.03$), with higher scores in cases who experienced subsequent cardiovascular events versus event-free controls. Individual metabolites within the factor were also significantly different ($P<0.05$) between cases and controls, with a similar direction of effect as observed in the original event dataset.

Discussion

We demonstrate in this study that peripheral blood metabolite profiles are independently associated with the presence of CAD and add to the discriminative capability for CAD compared with models containing clinical variables only. Furthermore, we report a specific metabolite cluster that independently predicts subsequent cardiovascular events in individuals with CAD. These findings hold promise for better defining the mechanisms of disease, improving clinical diagnostic models, and better risk stratification of patients once diagnosed with CAD.

Studies of nuclear magnetic resonance–based metabolomic profiling have uncovered differences in myocardial ischemia[10] and have been suggested to predict the presence and severity of CAD.[11] However, another group has refuted the CAD findings.[12] In both studies, the majority of analytes within the profile remained unidentified, thus limiting utility for understanding pathophysiology or as potential diagnostic/prognostic tools. In this study, we performed quantitative and targeted analysis of a discrete set of metabolites by reporting on multiple pathways of lipid, protein, and glucose metabolism and identified groups of metabolites that are associated with CAD and predictive of clinical events in patients with this malady.

For example, factor 4 scores (and individual metabolites composing the factor) were associated with CAD. This factor seems to report on branched-chain amino acid catabolism, previously implicated in obesity and insulin signaling.[4,13] We have previously observed clustering of these metabolites[2,4]; although C5 and valine showed an inconsistent direction of effect in the 2 groups, factor 4 scores and the other metabolites comprising the factor were consistent in both groups (higher in cases), suggesting that this group of metabolites is providing biological insight into the branched-chain amino acid pathway in CAD. Factor 9 was also associated with CAD in both groups; given the strong load of arginine, this factor may be reporting on amino acid catabolism and subsequent metabolism of ammonia derived from these amino acids in the urea cycle. Two of the metabolites composing this factor showed an inconsistent direction of effect across the 2 groups; however, we note that neither metabolite was significant as an individual discriminator. Furthermore, we show that a signature composed of dicarboxylacylcarnitines was predictive of cardiovascular events in individuals with CAD. Dicarboxylacylcarnitines are thought to be derived by carboxylation of long-chain acyl coenzyme A species by P450 enzymes of the endoplasmic reticulum, followed by metabolism of these modified fatty acids to smaller coenzyme A species in peroxisomes and their equilibration with cognate acylcarnitine species.[14] Thus, these metabolites could reflect alterations in mitochondrial fatty acid oxidation, carboxylation at the endoplasmic reticulum, or peroxisomal metabolism. Further studies are required to understand the mechanistic significance of these metabolites in cardiovascular disease.

The potential implications for our findings are multiple. They suggest that simple metabolite profiles could be further developed to aid in the diagnosis of CAD and to identify patients to target for more aggressive therapies. Furthermore, we have previously shown that these profiles are heritable in families with CAD.[2] Combined with this study, these results suggest that measurement of metabolite profiles may be useful for genetic studies of CAD.

We note that addition of metabolite profiles to a clinical model inclusive of lipids did not improve the ROC curve in the initial group, although it did in the replication group. This may be because metabolite profiles do not add independent information in patients with early-onset CAD in whom a higher burden of risk factors exists (hence, the higher c-statistic for the clinical model in the initial group). However, our results suggest that metabolite signatures are important independent discriminators of CAD in a more generalizable population (ie, replication group). Alternatively, this may be a phenomenon of the small sample sizes we had available for these studies and the constrained methodology that was used. Larger studies in broader populations will provide further insights into these relations.

Several limitations of this study deserve mention. First, we did not adjust for multiple comparisons but used PCA to account for the colinearity of metabolites. Only glutamate/glutamine would survive Bonferroni adjustment for multiple comparisons at an individual metabolite level. However, given the colinearity of metabolites, this correction is too conservative. Both factors 4 and 9 would survive Bonferroni correction at the level of factors. More important, we observed similar associations of both factors 4 and 9 with CAD in an independent dataset. Second, our results could be confounded by other factors; however, we have taken great care to explicitly evaluate and adjust for such confounders, including comorbidities and medications; all patients were fasting; and blood samples were collected under similar conditions. There was some attenuation of association in the initial group after adjustment for lipids, suggesting that metabolomic profiles could be reporting on lipid-related pathways. However, a large proportion of patients were being treated with dyslipidemia medications; therefore, lipids are reflective of underlying lipid risk and variability associated with medications. In fact, studies have shown a paradoxical

effect, with dyslipidemia being associated with better outcomes.[15] Hence, lipid-adjusted results should be interpreted with caution. Third, interindividual variability of these metabolites in a general population and biological variability by sex, race, and other demographics are unknown, although investigations are ongoing. Finally, our groups were not population based but were selected on the basis of referral for cardiac catheterization for concern of cardiac ischemia. However, the angiographic phenotype with true lack of disease in controls is a strength. Despite these potential limitations, our study is the largest to date to study peripheral blood metabolite profiles in cardiovascular disease.

In summary, we have observed that simple metabolite profiles are independently associated with CAD and the occurrence of subsequent cardiovascular events. These profiles point toward potential diverse and novel mechanisms of CAD pathophysiology and the opportunity for improved risk stratification.

Acknowledgments

We are grateful for the time devoted by study participants.

Sources of Funding

This work was supported by the American Heart Association (FTF Award, to Dr Shah) and an unrestricted educational grant to Duke University from Medtronic, Inc (to Dr Shah).

Disclosures

None.

References

1. Ferrara CT, Wang P, Neto EC, Stevens RD, Bain JR, Wenner BR, Ilkayeva OR, Keller MP, Blasiole DA, Kendziorski C, Yandell BS, Newgard CB, Attie AD. Genetic networks of liver metabolism revealed by integration of metabolic and transcriptional profiling. *PLoS Genet*. 2009;4:e1000034.
2. Shah SH, Hauser ER, Bain JR, Muehlbauer MJ, Haynes C, Stevens RD, Wenner BR, Dowdy ZE, Granger CB, Ginsburg GS, Newgard CB, Kraus WE. High heritability of metabolomic profiles in families burdened with premature cardiovascular disease. *Mol Syst Biol*. 2009;5:258–266.
3. Smith LR, Harrell FE Jr, Rankin JS, Califf RM, Pryor DB, Muhlbaier LH, Lee KL, Mark DB, Jones RH, Oldham HN. Determinants of early versus late cardiac death in patients undergoing coronary artery bypass graft surgery. *Circulation*. 1991;8:III-245–III-253.
4. Newgard CB, An J, Bain JR, Muehlbauer MJ, Stevens RD, Lien LF, Haqq AM, Shah SH, Arlotto M, Slentz CA, Rochon J, Gallup D, Ilkayeva O, Wenner BR, Yancy WS Jr, Eisenson H, Musante G, Surwit RS, Millington DS, Butler MD, Svetkey LP. A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. *Cell Metab*. 2009;9:311–326.
5. An J, Muoio DM, Shiota M, Fujimoto Y, Cline GW, Shulman GI, Koves TR, Stevens R, Millington D, Newgard CB. Hepatic expression of malonyl-CoA decarboxylase reverses muscle, liver and whole-animal insulin resistance. *Nat Med*. 2004;10:268–274.
6. Chace DH, Hillman SL, Millington DS, Kahler SG, Roe CR, Naylor EW. Rapid diagnosis of maple syrup urine disease in blood spots from newborns by tandem mass spectrometry. *Clin Chem*. 1995;41:62–68.
7. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44:837–845.
8. Koves TR, Ussher JR, Noland RC, Slentz D, Mosedale M, Ilkayeva O, Bain J, Stevens R, Dyck JR, Newgard CB, Lopaschuk GD, Muoio DM. Mitochondrial overload and incomplete fatty acid oxidation contribute to skeletal muscle insulin resistance. *Cell Metab*. 2008;7:45–56.
9. Shaham O, Wei R, Wang TJ, Ricciardi C, Lewis GD, Vasan RS, Carr SA, Thadhani R, Gerszten RE, Mootha VK. Metabolic profiling of the human response to a glucose challenge reveals distinct axes of insulin sensitivity. *Mol Syst Biol*. 2008;4:214–222.
10. Sabatine MS, Liu E, Morrow DA, Heller E, McCarroll R, Wiegand R, Berriz GF, Roth FP, Gerszten RE. Metabolomic identification of novel biomarkers of myocardial ischemia. *Circulation*. 2005;112:3868–3875.
11. Brindle JT, Antti H, Holmes E, Tranter G, Nicholson JK, Bethell HW, Clarke S, Schofield PM, McKilligin E, Mosedale DE, Grainger DJ. Rapid and noninvasive diagnosis of the presence and severity of coronary heart disease using 1H-NMR-based metabonomics. *Nat Med*. 2002;8:1439–1444.
12. Kirschenlohr HL, Griffin JL, Clarke SC, Rhydwen R, Grace AA, Schofield PM, Brindle KM, Metcalfe JC. Proton NMR analysis of plasma is a weak predictor of coronary artery disease. *Nat Med*. 2006;12:705–710.
13. Um SH, D'Alessio D, Thomas G. Nutrient overload, insulin resistance, and ribosomal protein S6 kinase 1, S6K1. *Cell Metab*. 2006;3:393–402.
14. Ferdinandusse S, Denis S, Van Roermund CW, Wanders RJ, Dacremont G. Identification of the peroxisomal beta-oxidation enzymes involved in the degradation of long-chain dicarboxylic acids. *J Lipid Res*. 2004;45:1104–1111.
15. Wang TY, Newby LK, Chen AY, Mulgund J, Roe MT, Sonel AF, Bhatt DL, DeLong ER, Ohman EM, Gibler WB, Peterson ED. Hypercholesterolemia paradox in relation to mortality in acute coronary syndrome. *Clin Cardiol*. 2009;32:E22–E28.

CLINICAL PERSPECTIVE

Coronary artery disease (CAD) is the leading cause of death in industrialized countries, and many accepted risk factors for CAD are metabolic. However, we have an incomplete mechanistic understanding of CAD risk, and equally important, there is a need to refine our ability to identify individuals at highest risk of cardiovascular events. New molecular profiling tools may help improve risk stratification and enhance our understanding of the disease process. Metabolomics is a novel technology used to study the small-molecule metabolites that are byproducts of cellular metabolism and that may be particularly useful for diagnosis of human disease. Therefore, in this study, we performed quantitative metabolomic profiling of 69 metabolites, including acylcarnitine species (byproducts of mitochondrial metabolism), amino acids, and conventional metabolites. Profiling was performed in peripheral blood samples from 2 independent CAD case-control datasets of participants enrolled in the Duke CATHGEN biorepository of patients undergoing cardiac catheterization. We observed that 2 metabolomic biosignatures were independently associated with CAD in both case and control datasets, 1 comprising metabolites from the branched-chain amino acid pathway and 1 from the urea cycle pathway. The metabolomic signatures also seemed to add discriminative capability to models with clinical risk factors alone. A third metabolomic biosignature in the blood was able to independently predict who would experience future cardiovascular events (death or myocardial infarction). Hence, these metabolomic signatures may represent useful markers for the presence of CAD, potentially improve risk stratification, and identify novel mechanisms of CAD pathophysiology.