



Published in final edited form as:

Diabetologia. 2022 April ; 65(4): 657–674. doi:10.1007/s00125-021-05641-x.

Circulating metabolite profile in young adulthood identifies long-term diabetes susceptibility: the Coronary Artery Risk Development in Young Adults (CARDIA) study

Venkatesh L. Murthy¹, Matthew Naylor², Mercedes Carnethon³, Jared P. Reis⁴, Donald Lloyd-Jones³, Norrina B. Allen³, Robert Kitchen⁵, Paolo Piaggi⁶, Lyn M. Steffen⁷, Ramachandran S. Vasan^{8,9}, Jane E. Freedman¹⁰, Clary B. Clish¹¹, Ravi V. Shah¹⁰

¹Department of Medicine and Radiology, University of Michigan, Ann Arbor, MI, USA

²Section of Cardiovascular Medicine, Department of Medicine, Boston University School of Medicine, Boston, MA, USA

³Northwestern University, Chicago, IL, USA

⁴National Heart, Lung, and Blood Institute, Bethesda, MD, USA

⁵Simches Cardiovascular Research Center, Massachusetts General Hospital, Boston, MA, USA

⁶Department of Information Engineering, University of Pisa, Pisa, Italy

⁷University of Minnesota School of Public Health, Minneapolis, MN, USA

⁸Section of Preventive Medicine and Epidemiology, Boston University School of Medicine, Boston, MA, USA

⁹Framingham Heart Study, Framingham, MA, USA

¹⁰Vanderbilt Translational and Clinical Cardiovascular Research Center, Vanderbilt University School of Medicine, Nashville, TN, USA

¹¹Broad Institute of Harvard and MIT, Cambridge, MA, USA

Abstract

Aims/hypothesis—The aim of this work was to define metabolic correlates and pathways of diabetes pathogenesis in young adults during a subclinical latent phase of diabetes development.

Methods—We studied 2083 young adults of Black and White ethnicity in the prospective observational cohort Coronary Artery Risk Development in Young Adults (CARDIA) study (mean \pm SD age 32.1 \pm 3.6 years; 43.9% women; 42.7% Black; mean \pm SD BMI 25.6 \pm 4.9 kg/m²) and 1797 Framingham Heart Study (FHS) participants (mean \pm SD age 54.7 \pm 9.7 years;

[✉]Venkatesh L. Murthy, vlmurthy@med.umich.edu; Ravi V. Shah, ravi.shah@vumc.org. Venkatesh L. Murthy and Ravi V. Shah contributed equally to this work.

Contribution statement VLM, CBC, RSV and RVS contributed to the conception, design and acquisition of data. All authors made substantial contributions to analysis and interpretation of data, drafting of the article critically for intellectual content, and gave approval of the version to be published. VLM is the guarantor of this work.

Supplementary Information The online version contains peer-reviewed but unedited supplementary material available at <https://doi.org/10.1007/s00125-021-05641-x>.

52.1% women; mean \pm SD BMI 27.4 ± 4.8 kg/m²), examining the association of comprehensive metabolite profiles with endophenotypes of diabetes susceptibility (adipose and muscle tissue phenotypes and systemic inflammation). Statistical learning techniques and Cox regression were used to identify metabolite signatures of incident diabetes over a median of nearly two decades of follow-up across both cohorts.

Results—We identified known and novel metabolites associated with endophenotypes that delineate the complex pathophysiological architecture of diabetes, spanning mechanisms of muscle insulin resistance, inflammatory lipid signalling and beta cell metabolism (e.g. bioactive lipids, amino acids and microbe- and diet-derived metabolites). Integrating endophenotypes of diabetes susceptibility with the metabolome generated two multi-parametric metabolite scores, one of which (a proinflammatory adiposity score) was associated with incident diabetes across the life course in participants from both the CARDIA study (young adults; HR in a fully adjusted model 2.10 [95% CI 1.72, 2.55], $p < 0.0001$) and FHS (middle-aged and older adults; HR 1.33 [95% CI 1.14, 1.56], $p = 0.0004$). A metabolite score based on the outcome of diabetes was strongly related to diabetes in CARDIA study participants (fully adjusted HR 3.41 [95% CI 2.85, 4.07], $p < 0.0001$) but not in the older FHS population (HR 1.15 [95% CI 0.99, 1.331], $p = 0.07$).

Conclusions/interpretation—Selected metabolic abnormalities in young adulthood identify individuals with heightened diabetes risk independent of race, sex and traditional diabetes risk factors. These signatures replicate across the life course.

Keywords

Diabetes; Metabolomics; Prevention

Introduction

The transition from normoglycaemia to clinically diagnosed diabetes has a long subclinical ‘latent’ period, during which sufficient physiological reserve exists to counteract worsening peripheral insulin resistance and progressive pancreatic beta cell loss and failure. In this subclinical phase, circulating biomarkers of beta cell failure and molecular biomarkers of disease risk (e.g. genomics, metabolites and proteins) have been proposed to clarify early pathogenesis of diabetes. Limited data using genomics and metabolite profiling in young adults have suggested that a complex underlying molecular architecture of diabetes propensity and obesity may exist for over a decade prior to clinically overt diabetes [1, 2]. Furthermore, broad evidence supports the notion that structural phenotypes of ‘metabolic health’ (alterations in adipose tissue and muscle distribution/fat content) may inform diabetes risk before frank dysglycaemia occurs [3]. Nevertheless, most studies examining metabolite profiles have focused on mid-life, where much of the metabolic risk may already be ‘encoded’ in traditional clinical risk factors.

The primary hypothesis of this study was that metabolic phenotypes reflecting the earliest stages of diabetes development in young adulthood would identify biomarkers and pathways of diabetes risk. We establish molecular signatures of future diabetes risk based on association between circulating polar and lipid metabolites in young adulthood and multiple endophenotypes implicated in diabetes (e.g. adiposity, inflammation, muscle phenotypes)

prospectively over a follow-up period of over 20 years, highlighting known and novel pathways implicated by associated metabolites. Finally, to study the early metabolic origins of diabetes, we investigate the association of these metabolite signatures with incident diabetes in two separate cohorts across the lifespan: a large cohort of young White and Black adults (Coronary Artery Risk Development in Young Adults study [CARDIA]); and a validation cohort of middle-aged adults (the Framingham Heart Study [FHS]).

Methods

The general flow of methods for the study is shown in Fig. 1. The datasets analysed during the current study are available from the CARDIA study (www.cardia.dopm.uab.edu) or FHS (www.framinghamheartstudy.org) directly.

CARDIA study

CARDIA is a prospective cohort study of 5115 White and Black participants (age 18–30 years, recruited in 1985–1986 from four field centres: Birmingham, AL; Chicago, IL; Minneapolis, MN; Oakland, CA). Here, we quantified metabolites in 2358 individuals in a fasting state (>8 h) at the year 7 visit, selected across a spectrum of metabolic risk with cases additionally selected to enrich for individuals with incident cardiovascular disease during follow-up [4]. We excluded individuals who had diabetes (fasting blood sugar ≥ 7 mmol/l or self-reported diabetes at the study visit or any CARDIA study examination prior to year 7) and those who did not have contemporary (year 7) BMI, glucose, parental history of diabetes, or diabetes at or before year 7, yielding 2083 participants. The final study subsample had characteristics similar to those of the full cohort attending the year 7 visit (electronic supplementary material [ESM] Table 1). Assessment of standard cardiovascular risk factors as well as computed tomography (CT) characterisation of adipose, liver, and muscle tissue have been described previously [5–7],

Metabolite profiling was performed via LC-MS (Broad Institute, Cambridge, MA), in four platforms: (1) lipids (C8-positive); (2) a variety of amino acid metabolites, acylcarnitines, dipeptides and other cationic polar metabolites (hydrophilic interaction liquid chromatography [HILIC]-positive); (3) fatty acids, eicosanoids and bile acids (C18-negative); and (4) sugars, organic acids, nucleotides and anionic polar metabolites (HILIC-negative). Methods for metabolite quantification and analysis have been published previously [8, 9]. Metabolite profiling in the FHS has been described previously [1].

FHS

We studied 1797 participants from the FHS Second Generation (Offspring) cohort who underwent fasting metabolite profiling at the fifth examination cycle and were free of prevalent diabetes (defined as medication use for diabetes, fasting glucose ≥ 7 mmol/l or random glucose ≥ 11.1 mmol/l on or before examination 5; ESM Table 2) [1]. Metabolite profiling in the FHS was done previously in a case-cohort design, with a case-control match for selected conditions including incident diabetes to which a set of randomly selected control individuals was added [1]. For this study, we combined individuals across the

subsamples. All individuals provided written informed consent, and CARDIA and FHS studies were approved by the Institutional Review Board at each participating institution.

Endophenotypes and clinical outcomes

Endophenotypes in CARDIA included waist circumference (at year 7, 10, 15, 20, 25 and 30 visits), C-reactive protein (year 7, 15, 20 and 25), IL-6 (year 20) and CT-based measures at year 25, including liver attenuation, pericardial fat volume, abdominal fat volume, visceral fat volume, subcutaneous fat volume, intermuscular fat volume, psoas fat and muscle volume, paraspinous fat and muscle volume, lateral oblique fat and muscle volume, rectus fat and muscle volume, and visceral/subcutaneous fat volume ratio. Of note, we included fat and muscle measures given their postulated connection to diabetes.

Our primary clinical outcome in CARDIA study participants was incident diabetes, defined as the composite of self-reported diabetes, fasting glucose ≥ 7 mmol/l or 2 h glucose of ≥ 11.1 mmol/l on an OGTT (if performed). Age of onset of diabetes was taken as self-reported age of onset of diabetes or, if not reported, age at first examination with self-reported diabetes or with plasma glucose measurements consistent with diabetes. Time to diabetes or censor was measured as the difference in age from the year 7 study visit to censoring (loss to follow-up) or to the onset of diabetes. Of note, while we were unable to differentiate type 2 diabetes from type 1 diabetes, most cases of diabetes identified at the period in adulthood that we have investigated is likely to be type 2 [10]. For FHS participants, our primary outcome for validation was diabetes (defined as medication use for diabetes, fasting glucose ≥ 7 mmol/l, random glucose ≥ 11.1 mmol/l or HbA_{1c} $\geq 6.5\%$, where measured). Survival time was taken as the difference in age at the examination 5 study visit (for the Offspring cohort) to age at censoring or age of onset of diabetes.

Statistical analysis

Exposure preparation—CT fat and muscle volumes were scaled by height raised to the 2.7 power. We selected the power of 2.7 for allometric height indexing, as this minimised the relationship between height and individual CT measures (relative to a linear or 1.7 power), thereby making them less body-size dependent. After scaling for body size where appropriate, distributions of continuous phenotypes and covariates were graphically examined. Continuous endpoints and covariates were subjected to hyperbolic arcsine transformation to improve normality. All continuous phenotypes were then mean-centred and standardised for further analysis.

Single metabolite to endophenotype regression—The first step was to examine relationships between metabolites and endophenotypes on a single-metabolite basis (using linear or logistic regression). Six sets of nested models were constructed with varying combinations of adjustments (across age, sex, race, year 7 BMI, year 7 glucose, parental history of diabetes, year 7 waist circumference, year 7 C-reactive protein, and/or the Framingham Diabetes Score [coefficients taken from <https://framinghamheartstudy.org/fhs-risk-functions/diabetes/>, accessed 14 January 2021]). In each case, a Benjamini–Hochberg false discovery rate (FDR) correction was applied within each combination of endophenotype and set of adjustments to control for multiplicity.

Defining metabolite scores for diabetes—Metabolites across all four platforms were combined (duplicates included) to generate scores. To generate a multivariable score for diabetes susceptibility, we employed two approaches: (1) directly fitting a Cox survival model for incident diabetes; and (2) combining regression models that were fit on endophenotypes related to diabetes susceptibility to build a score for diabetes. In both approaches, we used least absolute shrinkage and selection operator (LASSO) regression to select metabolites most explanatory for each given outcome (either diabetes itself or each endophenotype), using repeated cross-validation to optimise hyperparameters. For the second approach, we then applied principal components analysis (PCA) to metabolite coefficients from LASSO models of endophenotypes to organise the strength of the metabolome–endophenotype relationships into a metabolic ‘signature.’ This two-staged approach (penalised regression followed by PCA) has been employed previously to generate summary scores based on high-dimensional metabolomic data across various, related clinical phenotypes [8]. The principal components that result are loaded on adiposity, muscle or inflammatory outcomes, with individual metabolites given a numerical weight for each principal component. Scores were then calculated as a linear sum of log-transformed and scaled metabolite level multiplied by its principal component-defined weight, across all metabolites separately for derived scores. Scores were then mean-centred/standardised for analysis. The relationship between these scores and incident diabetes was evaluated using Cox regression with the same adjustment models described above in single-metabolite regressions. Unadjusted survival plots across tertiles of the scores for each outcome were generated using the Kaplan–Meier method. Effect modification was evaluated for age, sex and race on each of these scores. Model R^2 was computed using the method of Xu and O’Quigley, and C statistic was computed by the method of Harrell. Net reclassification improvement (NRI) was computed at the 75th percentile of follow-up time using methods for censored data.

FHS validation—Because not all of the identified metabolites included in scores in CARDIA were measured in the FHS validation cohort, we identified matching metabolites and then fit ‘reduced’ metabolite scores in CARDIA to apply to FHS using linear regression. In this approach, the ‘full’ scores in CARDIA were the response (dependent) variables and metabolites in the full score that were also available for validation in FHS were predictor (independent) variables. We measured correlation of this ‘reduced’ score against the ‘full’ score in CARDIA (derivation cohort) to test whether a score based on the smaller set of metabolites carried to validation would still be highly correlated with the full score. These reduced metabolite scores were subsequently computed in the FHS, standardised, and validated against incident diabetes using Cox regression, adjusted for age, sex, fasting glucose and parental history of diabetes and either BMI or waist circumference. We explored effect modification by age, sex and race (race in the CARDIA cohort only) using multiplicative interaction terms. We evaluated NRI and model discrimination and fit as in CARDIA.

R (version 4.1.0; R Foundation for Statistical Computing, Vienna, Austria) was used for statistical analyses. All models were built on complete cases. A two-sided p value less than

0.05 was considered significant, with type 1 multiplicity corrections using the FDR method, as specified above.

Results

Clinical features of the study sample

The characteristics of the CARDIA subsample at the time of metabolite collection (year 7) are shown in ESM Table 1. The CARDIA subsample for this study was composed of young adults with an approximately equal distribution by sex and race (mean \pm SD age 32.1 ± 3.6 years; 43.9% women; 42.7% Black) and a mean \pm SD BMI of 25.6 ± 4.9 kg/m². The study sample exhibited a low prevalence of hypertension and normal fasting blood glucose (mean \pm SD 5.0 ± 0.4 mmol/l) and lipid levels. The characteristics of the FHS subsample are shown in ESM Table 2. In general, included FHS participants were older (mean \pm SD age 54.7 ± 9.7 years), heavier (mean \pm SD BMI 27.4 ± 4.8 kg/m²) and had a worse cardiometabolic risk profile.

Metabolites associated with adiposity and proinflammatory endophenotypes identify pathways of diabetes susceptibility in young adulthood

A large number of significant metabolite associations was shared across proinflammatory endophenotypes (waist circumference, visceral and intermuscular adiposity, hepatic steatosis, inflammation; ESM Fig. 1). Branched-chain amino acids (muscle insulin resistance [11]), dimethylguanidinovaleric acid (hepatic steatosis [12–14]), and glutamate and tyrosine (pancreatic beta cell dysfunction and decreased insulin secretion [15]) were associated with greater adiposity and/or inflammation, while glycine (associated with decreased oxidative stress and diabetes risk [16–18]), glutamine [19] and 1-methylnicotinamide (anti-inflammatory [20]) were lower in individuals with greater visceral adiposity (Table 1). We also observed many lipids related to endophenotypes in previous reports in smaller or older populations, including proinflammatory bioactive sphingomyelins and ceramides [21, 22], long-chain triacylglycerols and diacylglycerols of low saturation, and lipid plasmalogens linked to inflammation or insulin resistance (e.g. C36:2 phosphatidylcholine [PC] plasmalogen [23]). In addition, we identified several metabolites associated with adiposity or inflammation in young adults that have not been widely reported previously, including microbe- or diet-derived metabolites (cinnamoylglycine [24], hippurate [25], trigonelline [26, 27]) and several biologically active phospholipids (lysophosphatidylcholines [LPCs] and lysophosphatidylethanolamines [LPEs] [28–31]).

The metabolome in young adulthood identifies distinct phenotypic axes relevant to diabetes pathogenesis

We next sought to identify integrated relationships across multiple metabolites and phenotypes to define metabolic predisposition to diabetes using a two-stage LASSO–PCA approach (Fig. 1). Results of the LASSO regression for each endophenotype (subset to the top 100 largest magnitude penalised regression coefficients across endophenotypes) are shown in Fig. 2 and visually suggest that clusters of metabolites may be concordantly related to distinct proinflammatory phenotypes. We describe two major ‘phenotypic

principal components' (Fig. 3a), explaining 59.2% of the total variance in the metabolome–endophenotype relationship. Principal component 1 appeared to reflect proinflammatory adiposity, with high loadings on waist circumference, C-reactive protein and IL-6, and visceral-like fat. Principal component 2 was highly loaded on muscle volumes and appeared to reflect an android tissue distribution. The weightings for the 100 most heavily weighted metabolites in principal components are shown in Fig. 3b; correlations between metabolite scores and individual phenotypes are shown in ESM Fig. 2. We observed a slightly lower proinflammatory adiposity score in White people relative to Black people, and a widely different distribution of android scores in men vs women (consistent with greater muscle mass in men; ESM Fig. 3). We observed a mild to moderate degree of correlation between the metabolite scores and BMI or waist circumference (9–56% variance explained by BMI or waist circumference for both scores; ESM Fig. 4). Metabolites associated with proinflammatory adiposity and android distribution (summarised in Table 1) identified both familiar and novel pathways relevant to diabetes across a wide array of diabetes pathophysiology, including gut microbial products, muscle tissue insulin resistance and fatty liver disease.

Coefficients for metabolites for each score are provided in ESM Table 3, and full regression results are shown in ESM Tables 4–30.

Composite metabolite-based scores in young adulthood are associated with long-term diabetes development over >20 years

Over a median follow-up of 23 years (IQR 18, 23 years) in CARDIA study participants, we observed 239 incident diabetes cases. As expected, in models fit on incident diabetes within CARDIA (Cox LASSO), we observed strong associations with incident diabetes with all adjustments (Table 2). We examined the association of each metabolite-based score with diabetes and found that phenotype-based scores were prognostic for incident diabetes within CARDIA participants (Table 2, Fig. 4; HR of 2.10 per 1 SD higher proinflammatory adiposity score and HR of 2.24 per SD higher android score). The scores improved risk discrimination by *C*-index, increasing from 0.762 to 0.781 ($p=0.02$) for the proinflammatory adiposity score and to 0.778 ($p=0.03$) for the android score. Favourable continuous NRI range was strong for the proinflammatory adiposity score (NRI 0.404 [95% CI 0.253, 0.568]) and moderate for the android score (NRI 0.284 [95% CI 0.105, 0.417]), compared with a model with clinical risk factors including BMI (similar findings with models with waist circumference. Table 2). There was also a modest interaction between the two metabolomic scores developed on subclinical phenotypes in the CARDIA participants, with additive effects of adverse changes in both ($p=0.01$, ESM Fig. 5). No statistically significant interactions were seen in the scores with age, race or sex.

We next proceeded to external validation of the metabolite scores in FHS participants. The Cox LASSO models fit on diabetes in CARDIA study participants were associated with diabetes in minimally adjusted (age- and sex-adjusted) but not fully adjusted models in FHS participants (Table 3). The distribution of each metabolite-based score across sex and age in FHS is shown in ESM Fig. 6, demonstrating similar patterns by sex and no large relationship with age. Correlation of the final ('reduced') scores for the android and proinflammatory

scores were excellent in CARDIA study participants ($R=0.97$ for the final reduced vs full score in both cases). We replicated the relationship between the proinflammatory adiposity score and incident diabetes in the FHS participants (283 incident cases over a median of 18 years [IQR 12, 20]; ESM Fig. 6). The proinflammatory adiposity score was associated with increased risk of incident diabetes (HR per 1 SD higher score 1.33, $p=0.0004$) after adjustment for risk factors (Table 3), with improvement of discrimination based on C -index increase from 0.836 to 0.841 ($p=0.05$). The relationship of the android score with diabetes in FHS participants was significant in models adjusted for age and sex but not independently significant of all risk factors including BMI (in adjusted model: HR 1.19, $p=0.09$; Table 3). Interactions with age and sex were not significant.

Discussion

The principal aim of this study was to identify comprehensive profiles comprised of known and novel metabolites associated with endophenotypes of diabetes susceptibility (e.g. systemic inflammation, body composition) in young Black adults and young White adults early in the course of disease. In addition to extending metabolites previously implicated in diabetes to a younger, biracial cohort (CARDIA study), we identified many novel metabolites in pathways implicated in diabetes not previously widely reported for further investigation (e.g. bioactive lipid species, microbe-derived/diet-derived metabolites). Multi-metabolite scores derived from phenotypes were independently associated with long-term diabetes development from young adulthood to mid-life and provided incremental risk stratification beyond clinical risk factors in the CARDIA study, with validation in an independent sample of individuals across a broad age range (FHS). Of note, the proinflammatory score was associated with diabetes across both cohorts, while the android score was only significant in the CARDIA study participants. In addition, metabolite scores fit directly on diabetes were associated with diabetes in the CARDIA study participants after full adjustment but not in the FHS participants. Collectively, the observation that metabolite-based signatures in young adulthood provide prognostic value for diabetes incremental to race, sex and traditional diabetes risk factors over two decades highlights a potential role for metabolic risk stratification at an early stage when intervention may be most likely to reap maximal benefits.

The current study offers several important areas of novelty to the burgeoning literature on translatable, precision approaches to diabetes prevention. First, quantification of the circulating metabolome in CARDIA study participants (a young adult sample of Black and White individuals at high lifetime metabolic risk) extends the current literature on metabolomic risk prediction to an early period in diabetes development. Indeed, much of the work in large cohorts examining metabolite profiles has focused on mid-life, where much of the metabolic risk may already be 'encoded' in traditional clinical risk factors. Accordingly, the predictive performance of prior multi-metabolite scores in diabetes prediction in large, heterogeneous community-based samples across a wide age spectrum has been modest, with limited added prognostic value for metabolites [1,2]. Of note, recent seminal studies have suggested that metabolite-based diabetes prediction may offer greater value in individuals who are younger [32] or who are without prevalent dysglycaemia [33], possibly due to decreased sensitivity of clinical factors alone, differences among cohorts

(race, age, geographic location), analytical follow-up or methodology. Indeed, seminal trans-cohort work by Ahola-Olli and colleagues [32] involving 229 metabolites (polar/lipid metabolites, fatty acids) measured by NMR demonstrated a similar effect size for diabetes, though restricted to White European study participants. The results in the CARDIA study participants are fundamentally useful: (1) they extend the utility of metabolite-based markers of diabetes susceptibility across race, sex and geography; and (2) they demonstrate further comprehensive discovery of many new metabolites and metabolic pathways (due to a broader metabolome coverage) not previously implicated in diabetes in human investigation. The consistency of our findings with previously published work (see review [34]) lends to the validity of the approach and the importance of the shared metabolites in the pathogenesis of diabetes. Further studies based on these results can therefore prioritise metabolites for downstream mechanistic discovery and clinical application, thereby limiting the cost and scope of metabolite profiling performed in the clinical context.

In addition to a broader metabolome coverage in a young adult cohort not previously reported, we used integrative statistical techniques to develop multi-metabolite scores rooted in precise measures of body composition, inflammation and anthropometry (endophenotypes) to focus discovery on preclinical, quantitative pathophysiology heterogeneous across individuals and central to diabetes. The use of continuous phenotypes as outcomes to guide metabolite discovery (as opposed to dysglycaemia alone) not only improves statistical power but also may uncover novel metabolite patterns related to diabetes susceptibility. Accordingly, many identified metabolites in the identified pathophenotypic principal components in the CARDIA study participants have not previously been widely associated with diabetes (Table 1). Certainly, this pattern is consistent with observations from human genetics suggesting that diabetes susceptibility loci may not necessarily reflect glycaemic traits alone [35]. Indeed, a heterogeneity of metabolic alterations that precede diabetes (e.g. microbial diversity, intestinal function, innate immunity) are uniquely captured by the approach here, highlighting opportunities for broad downstream mechanistic discovery.

Furthermore, apart from metabolites with well-known association with diabetes (e.g. branched-chain amino acids; Table 1), we uncovered metabolites implicated in traditional mechanisms of diabetes pathogenesis across multiple tissues, including muscle insulin resistance (e.g. asymmetric dimethylarginine [36, 37]), oxidative stress and inflammation (bilirubin/biliverdin [38], glycine [16, 17, 39]), and intestinal inflammation and permeability (asparagine [40]). Moreover, we identified several additional metabolite classes with a suspected mechanistic role in diabetes not previously widely reported, including microbe- or food-derived products and bioactive lipids. Sphingomyelins and ceramide species (mostly related to inflammation/adiposity in the CARDIA study) have been implicated in insulin resistance [41], though with variable associations in humans [42]. Several LPCs and LPEs (biologically active lipid byproducts of plasma membrane metabolism) were associated with inflammation and adiposity in CARDIA study participants. Despite known roles of some LPCs in innate immune activation [43], LPCs associated with lower inflammation or lower incident diabetes risk in CARDIA participants are modifiable with weight loss in studies of obesity (e.g. LPC 18:1, LPC 20:4 [44]), underscoring the complexity and context-dependence of lipid signalling. Select plasmalogens (a suspected

buffer to proinflammatory lipid oxidation [45] associated with insulin resistance in animal models [23]) were also associated with adiposity phenotypes in the CARDIA study participants. Moreover, eicosanoid metabolic byproducts of arachidonic acid linked to diabetes complications [46, 47] were associated with diabetes in CARDIA study participants (e.g. 5-hydroxyeicosatetraenoic acid [5-HETE]). In addition, 12,13-diHOME (12, 13-dihydroxy-9-octadecenoic acid; a lipokine released by brown adipose tissue during exercise [9, 48], implicated in brown fat activation [49] and skeletal muscle metabolic efficiency [48]) also exhibited favourable weighting in the proinflammatory metabolite score. Finally, we identified microbial and food-derived metabolites, some of which may be functional or modifiable by dietary intake, that are linked to obesity and insulin sensitivity (cinnamoylglycine, hippurate, trigonelline, histidine). Additional work to understand mechanistic implications of these physiologically plausible associations is warranted.

The novelty of this study comes in the uniqueness of the sample (one of the largest young White and Black American adult populations with decades of follow-up for events), broad metabolome coverage (permitting discovery of novel metabolite–phenotype associations central to diabetes risk) and statistical approaches used to unify phenotypes, metabolites and clinical risk. Nevertheless, these results have important limitations. While our replication was not across race (FHS), we observed no evidence of effect modification by race in CARDIA study participants, suggesting that the metabolic abnormalities present in young adulthood may be similarly associated with diabetes risk in Black adults and White adults. We acknowledge that the use of CT-defined fat distributions at year 25 to derive metabolite-based scores for diabetes that are then applied across the entirety of follow-up does introduce some bias, in that adiposity itself may influence and be influenced by metabolic dysfunction and diabetes. In addition, given that metabolites are quantified on a relative scale, clinical translation of metabolic risk scores require establishing absolute quantification, normative bounds across age and sex (among other confounders of interest) and testing across a broad array of samples. Nevertheless, identifying robust, minimal signatures of risk that are biologically plausible in the pathogenesis of diabetes is an important first step to prioritising those metabolites relevant for absolute quantification and further downstream validation. We recognise that the sample studied here did not have a high proportion of individuals with severe obesity and did not study longitudinal modification of activity and diet as a method to alter the metabolome and mitigate diabetes risk. Of note, recent reports suggest that acute exercise may alter the circulating levels of some metabolites implicated in diabetes risk [9]. Finally, using novel techniques, we observed validation of the proinflammatory (but not the android) score across both cohorts. These findings may represent the differences between the two cohorts (e.g. age, sex, distribution of adiposity), and further validation by different methods and in larger, diverse cohorts would be necessary.

In conclusion, in a large, biracial cohort of young adults, we used broad metabolite platform and integrative analytic approaches across multiple, diverse phenotypes of metabolic risk to uncover novel associations between the circulating metabolome and selected endophenotypes reflecting metabolic risk, specifying pathways of insulin resistance, oxidative stress, hepatic steatosis, and inflammatory lipid signalling. Multi-metabolite scores

derived from these approaches (specifically proinflammatory adiposity) were associated with incident diabetes, suggesting the importance of inflammatory metabolic alterations early in adulthood to diabetes risk. Given the potential impact of nutrition- and activity-based interventions on several identified metabolic pathways, further work targeting precision clinical applications and molecular discovery based on metabolic profiles early in adulthood diagnostically and therapeutically are warranted.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We acknowledge the participants, coordinators and investigators of the CARDIA study and the FHS, without whom these efforts would not have been possible.

Funding

VLM and RVS are supported in part by grants from the National Institute on Aging (R01 AG059729) and National Heart, Lung and Blood Institute (R01 HL136685). MN is supported by grant K23 HL138260 from the National Heart, Lung, and Blood Institute. The CARDIA study is conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with the University of Alabama at Birmingham (HHSN268201800005I & HHSN268201800007I), Northwestern University (HHSN268201800003I), University of Minnesota (HHSN268201800006I) and Kaiser Foundation Research Institute (HHSN268201800004I). The FHS acknowledges the support of contracts NO1-HC-2519 5 , HHSN268201500001I and 75N92019D00031 from the National Heart, Lung and Blood Institute and grant R01DK081572 for this research. RSV is supported in part by the Evans Medical Foundation and the Jay and Louis Coffman Endowment from the Department of Medicine, Boston University School of Medicine. This manuscript has been reviewed by CARDIA for scientific content. The views expressed in this manuscript are those of the authors and do not necessarily represent the views of the NHLBI (the study funder); the National Institutes of Health; or the US Department of Health and Human Services. The study sponsors/funders were not involved in the design of the study; the collection, analysis, and interpretation of data; writing the report; and did not impose any restrictions regarding the publication of the report.

Authors' relationships and activities

VLM owns stock in General Electric and Cardinal Health and stock options in Ionetix and has received research grants and speaking honoraria from Siemens Medical Imaging and expert testimony fees on behalf of Jubilant Draximage. VLM has served on medical advisory boards for Ionetix and Curium. RVS is supported in part by grants from the National Institutes of Health and the American Heart Association. In the past 36 months, RVS has served as a consultant for Myokardia (concluded 28 February 2021) and Best Doctors (completed June 2021), and has been on a scientific advisory board for Amgen (concluded February 2018). RVS is a co-inventor on a patent for ex-RNAs signatures of cardiac remodelling. All other authors declare that there are no relationships or activities that might bias, or be perceived to bias, their work.

Data availability

The datasets analysed during the current study are available from the CARDIA study (www.cardia.dopm.uab.edu) or FHS (www.framinghamheartstudy.org) directly.

Abbreviations

CARDIA	Coronary Artery Risk Development in Young Adults
CT	Computed tomography
FDR	False discovery rate

FHS	Framingham Heart Study
HILIC	Hydrophilic interaction liquid chromatography
LASSO	Least absolute shrinkage and selection operator
LPC	Lysophosphatidylcholine
LPE	Lysophosphatidylethanolamine
NRI	Net reclassification improvement
PC	Phosphatidylcholine
PCA	Principal components analysis

References

1. Wang TJ, Larson MG, Vasan RS et al. (2011) Metabolite profiles and the risk of developing diabetes. *Nat Med* 17(4):448–453. 10.1038/nm.2307 [PubMed: 21423183]
2. Walford GA, Porneala BC, Dauriz M et al. (2014) Metabolite traits and genetic risk provide complementary information for the prediction of future type 2 diabetes. *Diabetes Care* 37(9):2508–2514. 10.2337/dcl4-0560 [PubMed: 24947790]
3. Shah RV, Murthy VL, Abbasi SA et al. (2014) Visceral adiposity and the risk of metabolic syndrome across body mass index: the MESA study. *JACC Cardiovasc Imaging* 7(12): 1221–1235. 10.1016/j.jcmg.2014.07.017 [PubMed: 25440591]
4. Murthy VL, Abbasi SA, Siddique J et al. (2016) Transitions in metabolic risk and long-term cardiovascular health: coronary artery risk development in young adults (CARDIA) study. *J Am Heart Assoc* 5(10):e003934. 10.1161/JAHA.116.003934 [PubMed: 27737876]
5. VanWagner LB, Ning H, Lewis CE et al. (2014) Associations between nonalcoholic fatty liver disease and subclinical atherosclerosis in middle-aged adults: the coronary artery risk development in young adults study. *Atherosclerosis* 235(2):599–605. 10.1016/j.atherosclerosis.2014.05.962 [PubMed: 24956534]
6. Terry JG, Shay CM, Schreiner PJ et al. (2017) Intermuscular adipose tissue and subclinical coronary artery calcification in midlife: the CARDIA study (coronary artery risk development in young adults). *Arterioscler Thromb Vasc Biol* 37(12):2370–2378. 10.1161/ATVBAHA.117.309633 [PubMed: 29025708]
7. Hill JO, Sidney S, Lewis CE, Tolan K, Scherzinger AL, Stamm ER (1999) Racial differences in amounts of visceral adipose tissue in young adults: the CARDIA (coronary artery risk development in young adults) study. *Am J Clin Nutr* 69(3):381–387. 10.1093/ajcn/69.3.381 [PubMed: 10075320]
8. Murthy VL, Reis JP, Pico AR et al. (2020) Comprehensive metabolic phenotyping refines cardiovascular risk in young adults. *Circulation*. 142(22):2110–2127. 10.1161/CIRCULATIONAHA.120.047689 [PubMed: 33073606]
9. Naylor M, Shah RV, Miller PE et al. (2020) Metabolic architecture of acute exercise response in middle-aged adults in the community. *Circulation*. 142 (20): 1905–1924. 10.1161/CIRCULATIONAHA.120.050281 [PubMed: 32927962]
10. Koopman RJ, Mainous AG 3rd, Diaz VA, Geesey ME (2005) Changes in age at diagnosis of type 2 diabetes mellitus in the United States, 1988 to 2000. *Ann Fam Med* 3(1):60–63. 10.1370/aftn.214 [PubMed: 15671192]
11. Newgard CB, An J, Bain JR et al. (2009) A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. *Cell Metab* 9(4):311–326. 10.1016/j.cmet.2009.02.002 [PubMed: 19356713]
12. O’Sullivan JF, Morningstar JE, Yang Q et al. (2017) Dimethylguanidino valeric acid is a marker of liver fat and predicts diabetes. *J Clin Invest* 127(12):4394–4402. 10.1172/JCI95995 [PubMed: 29083323]

13. Ottosson F, Ericson U, Almgren P et al. (2019) Dimethylguanidino Valerate: a lifestyle-related metabolite associated with future coronary artery disease and cardiovascular mortality. *J Am Heart Assoc* 8(19):e012846. 10.1161/JAFLA.119.012846 [PubMed: 31533499]
14. Robbins JM, Herzig M, Morningstar J et al. (2019) Association of Dimethylguanidino Valeric Acid with Partial Resistance to metabolic health benefits of regular exercise. *JAMA Cardiol* 4(7):636–643. 10.1001/janr.2019.1573 [PubMed: 31166569]
15. Vangipurapu J, Stancakova A, Smith U, Kuusisto J, Laakso M (2019) Nine amino acids are associated with decreased insulin secretion and elevated glucose levels in a 7.4-year follow-up study of 5,181 Finnish men. *Diabetes* 68(6): 1353–1358. 10.2337/db18-1076 [PubMed: 30885989]
16. McCarty MF, DiNicolantonio JJ (2013) The cardiometabolic benefits of glycine: is glycine an ‘antidote’ to dietary fructose? *Open Heart* 1(1):e000103. 10.1136/openhrt-2014-000103
17. Floegel A, Stefan N, Yu Z et al. (2013) Identification of serum metabolites associated with risk of type 2 diabetes using a targeted metabolomic approach. *Diabetes* 62(2):639–648. 10.2337/db12-0495 [PubMed: 23043162]
18. El-Hafidi M, Franco M, Ramirez AR et al. (2018) Glycine increases insulin sensitivity and glutathione biosynthesis and protects against oxidative stress in a model of sucrose-induced insulin resistance. *Oxidative Med Cell Longev* 2018:2101562. 10.1155/2018/2101562
19. Petrus P, Lecoutre S, Dollet L et al. (2019) Glutamine links obesity to inflammation in human White adipose tissue. *Cell Metab* 31(2): 375–390 [PubMed: 31866443]
20. Chlopicki S, Swies J, Mogielnicki A et al. (2007) 1-Methylnicotinamide (MNA), a primary metabolite of nicotinamide, exerts anti-thrombotic activity mediated by a cyclooxygenase-2/prostacyclin pathway. *Br J Pharmacol* 152(2):230–239. 10.1038/sj.bjp.0707383 [PubMed: 17641676]
21. Lemaitre RN, Yu C, Hoofnagle A et al. (2018) Circulating sphingolipids, insulin, HOMA-IR, and HOMA-B: the strong heart family study. *Diabetes* 67(8): 1663–1672. 10.2337/db17-1449 [PubMed: 29588286]
22. Sokolowska E, Blachnio-Zabielska A (2019) The role of ceramides in insulin resistance. *Front Endocrinol (Lausanne)* 10:577. 10.3389/fendo.2019.00577 [PubMed: 31496996]
23. Fujisaka S, Avila-Pacheco J, Soto M et al. (2018) Diet, genetics, and the gut microbiome drive dynamic changes in plasma metabolites. *Cell Rep* 22(11):3072–3086. 10.1016/j.celrep.2018.02.060 [PubMed: 29539432]
24. Cirulli ET, Guo L, Leon Swisher C et al. (2019) Profound perturbation of the metabolome in obesity is associated with health risk. *Cell Metab* 29(2):488–500 e482. 10.1016/j.cmet.2018.09.022 [PubMed: 30318341]
25. Pallister T, Jackson MA, Martin TC et al. (2017) Hippurate as a metabolomic marker of gut microbiome diversity: modulation by diet and relationship to metabolic syndrome. *Sci Rep* 7(1):13670. 10.1038/s41598-017-13722-4 [PubMed: 29057986]
26. Zhou JY, Du XH, Zhang Z, Qian GS (2017) Trigonelline inhibits inflammation and protects beta cells to prevent fetal growth restriction during pregnancy in a mouse model of diabetes. *Pharmacology* 100(5–6):209–217. 10.1159/000479088 [PubMed: 28723688]
27. Li Y, Li Q, Wang C, Lou Z, Li Q (2019) Trigonelline reduced diabetic nephropathy and insulin resistance in type 2 diabetic rats through peroxisome proliferator-activated receptor-gamma. *Exp Ther Med* 18(2): 1331–1337. 10.3892/etm.2019.7698 [PubMed: 31363374]
28. Del Bas JM, Caimari A, Rodriguez-Naranjo MI et al. (2016) Impairment of lysophospholipid metabolism in obesity: altered plasma profile and desensitization to the modulatory properties of n-3 polyunsaturated fatty acids in a randomized controlled trial. *Am J Clin Nutr* 104(2):266–279. 10.3945/ajcn.116.130872 [PubMed: 27305954]
29. Wallace M, Morris C, O’Grada CM et al. (2014) Relationship between the lipidome, inflammatory markers and insulin resistance. *Mol Bio Syst* 10(6): 1586–1595. 10.1039/c3mb70529c
30. Semba RD, Gonzalez-Freire M, Moaddel R et al. (2018) Altered plasma amino acids and lipids associated with abnormal glucose metabolism and insulin resistance in older adults. *J Clin Endocrinol Metab* 103(9):3331–3339. 10.1210/jc.2018-00480 [PubMed: 29947780]

31. Han MS, Lim YM, Quan W et al. (2011) Lysophosphatidylcholine as an effector of fatty acid-induced insulin resistance. *J Lipid Res* 52(6): 1234–1246. 10.1194/jlr.M014787 [PubMed: 21447485]
32. Ahola-Olli AV, Mustelin L, Kalimeri M et al. (2019) Circulating metabolites and the risk of type 2 diabetes: a prospective study of 11,896 young adults from four Finnish cohorts. *Diabetologia* 62(12):2298–2309. 10.1007/s00125-019-05001-w [PubMed: 31584131]
33. Merino J, Leong A, Liu CT et al. (2018) Metabolomics insights into early type 2 diabetes pathogenesis and detection in individuals with normal fasting glucose. *Diabetologia* 61(6): 1315–1324. 10.1007/s00125-018-4599-x [PubMed: 29626220]
34. Harneed A, Mojsak P, Buczynska A, Suleria HAR, Kretowski A, Ciborowski M (2020) Altered metabolome of lipids and amino acids species: a source of early signature biomarkers of T2DM. *J Clin Med* 9(7):2257. 10.3390/jcm9072257
35. Dimas AS, Lagou V, Barker A et al. (2014) Impact of type 2 diabetes susceptibility variants on quantitative glycemic traits reveals mechanistic heterogeneity. *Diabetes* 63(6):2158–2171. 10.2337/db13-0949 [PubMed: 24296717]
36. Lee W, Lee HJ, Jang HB et al. (2018) Asymmetric dimethylarginine (ADMA) is identified as a potential biomarker of insulin resistance in skeletal muscle. *Sci Rep* 8(1):2133. 10.1038/s41598-018-20549-0 [PubMed: 29391561]
37. Sydow K, Mondon CE, Schrader J, Konishi H, Cooke JP (2008) Dimethylarginine dimethylaminohydrolase overexpression enhances insulin sensitivity. *Arterioscler Thromb Vasc Biol* 28(4): 692–697. 10.1161/ATVBAHA.108.162073 [PubMed: 18239148]
38. Ikeda N, Inoguchi T, Sonoda N et al. (2011) Biliverdin protects against the deterioration of glucose tolerance in db/db mice. *Diabetologia* 54(8):2183–2191. 10.1007/s00125-011-2197-2 [PubMed: 21614569]
39. Wheeler MD, Thurman RG (1999) Production of superoxide and TNF-alpha from alveolar macrophages is blunted by glycine. *Am J Phys* 277(5):L952–L959. 10.1152/ajplung.1999.277.5.L952
40. Chen S, Liu Y, Wang X et al. (2016) Asparagine improves intestinal integrity, inhibits TLR4 and NOD signaling, and differently regulates p38 and ERK1/2 signaling in weanling piglets after LPS challenge. *Innate Immun* 22(8):577–587. 10.1177/1753425916664124
41. Teruel T, Hernandez R, Lorenzo M (2001) Ceramide mediates insulin resistance by tumor necrosis factor-alpha in brown adipocytes by maintaining Akt in an inactive dephosphorylated state. *Diabetes* 50(11):2563–2571. 10.2337/diabetes.50.11.2563 [PubMed: 11679435]
42. Torretta E, Barbacini P, Al-Daghri NM, Gelfi C (2019) Sphingolipids in obesity and correlated co-morbidities: the contribution of gender, age and environment. *Int J Mol Sci* 20(23):5901. 10.3390/ijms20235901
43. Chiurchiu V, Leuti A, Maccarrone M (2018) Bioactive lipids and chronic inflammation: managing the fire within. *Front Immunol* 9: 38. 10.3389/finrmu.2018.00038 [PubMed: 29434586]
44. Heimerl S, Fischer M, Baessler A et al. (2014) Alterations of plasma lysophosphatidylcholine species in obesity and weight loss. *PLoS One* 9(10):e111348. 10.1371/journal.pone.0111348 [PubMed: 25340546]
45. Braverman NE, Moser AB (2012) Functions of plasmalogen lipids in health and disease. *Biochim Biophys Acta* 1822 9:1442–1452. 10.1016/j.bbdis.2012.05.008 [PubMed: 22627108]
46. Paynter NP, Balasubramanian R, Giulianini F et al. (2018) Metabolic predictors of incident coronary heart disease in women. *Circulation* 137(8):841–853. 10.1161/CIRCULATIONAHA.117.029468 [PubMed: 29459470]
47. Al-Shabrawey M, Mussell R, Kahook K et al. (2011) Increased expression and activity of 12-lipoxygenase in oxygen-induced ischemic retinopathy and proliferative diabetic retinopathy: implications in retinal neovascularization. *Diabetes* 60(2):614–624. 10.2337/db10-0008 [PubMed: 21228311]
48. Stanford KI, Lynes MD, Takahashi H et al. (2018) 12,13-diHOME: An exercise-induced Lipokine that increases skeletal muscle fatty acid uptake. *Cell Metab* 27(5): 1111–1120 e1113. 10.1016/j.cmet.2018.03.020 [PubMed: 29719226]

49. Lynes MD, Leiria LO, Lundh M et al. (2017) The cold-induced lipokine 12,13-diHOME promotes fatty acid transport into brown adipose tissue. *Nat Med* 23(5):631–637. 10.1038/nm.4297 [PubMed: 28346411]
50. Liu X, Zheng Y, Guasch-Ferre M et al. (2019) High plasma glutamate and low glutamine-to-glutamate ratio are associated with type 2 diabetes: case-cohort study within the PREDIMED trial. *Nutr Metab Cardiovasc Dis* 29(10): 1040–1049. 10.1016/j.numecd.2019.06.005 [PubMed: 31377179]
51. Zhou M, Shao J, Wu CY et al. (2019) Targeting BCAA catabolism to treat obesity-associated insulin resistance. *Diabetes* 68(9): 1730–1746. 10.2337/db18-0927 [PubMed: 31167878]
52. Ito T, Schaffer SW, Azuma J (2012) The potential usefulness of taurine on diabetes mellitus and its complications. *Amino Acids* 42(5): 1529–1539. 10.1007/s00726-011-0883-5 [PubMed: 21437784]
53. Maleki V, Mahdavi R, Hajizadeh-Sharafabad F, Alizadeh M (2020) The effects of taurine supplementation on oxidative stress indices and inflammation biomarkers in patients with type 2 diabetes: a randomized, double-blind, placebo-controlled trial. *Diabetol Metab Syndr* 12:9. 10.1186/s13098-020-0518-7 [PubMed: 32015761]
54. Hellmuth C, Kirchberg FF, Lass N et al. (2016) Tyrosine is associated with insulin resistance in longitudinal Metabolomic profiling of obese children. *J Diabetes Res* 2016:2108909. 10.1155/2016/2108909 [PubMed: 26881241]
55. Feng RN, Niu YC, Sun XW et al. (2013) Histidine supplementation improves insulin resistance through suppressed inflammation in obese women with the metabolic syndrome: a randomised controlled trial. *Diabetologia* 56(5):985–994. 10.1007/s00125-013-2839-7 [PubMed: 23361591]
56. Niu YC, Feng RN, Hou Y et al. (2012) Histidine and arginine are associated with inflammation and oxidative stress in obese women. *Br J Nutr* 108(1):57–61. 10.1017/S0007114511005289 [PubMed: 21996294]
57. Ottosson F, Smith E, Melander O, Fernandez C (2018) Altered asparagine and glutamate homeostasis precede coronary artery disease and type 2 diabetes. *J Clin Endocrinol Metab* 103(8): 3060–3069. 10.1210/jc.2018-00546 [PubMed: 29788285]
58. Darabi Z, Darand M, Yari Z et al. (2019) Inflammatory markers response to citrulline supplementation in patients with non-alcoholic fatty liver disease: a randomized, double blind, placebo-controlled, clinical trial. *BMC Res Notes* 12(1):89. 10.1186/s13104-019-4130-6 [PubMed: 30767788]
59. Pallister T, Jackson MA, Martin TC et al. (2017) Untangling the relationship between diet and visceral fat mass through blood meta-bolomics and gut microbiome profiling. *Int J Obes* 41(7): 1106–1113. 10.1038/ijo.2017.70
60. Varadaiah YGC, Sivanesan S, Nayak SB, Thirumalarao KR (2019) Purine metabolites can indicate diabetes progression. *Arch Physiol Biochem* 1–5. 10.1080/13813455.2019.1663219
61. Strom K, Morales-Alamo D, Ottosson F et al. (2018) N(1)-methylnicotinamide is a signalling molecule produced in skeletal muscle coordinating energy metabolism. *Sci Rep* 8(1):3016. 10.1038/s41598-018-21099-1 [PubMed: 29445118]
62. Hong S, Moreno-Navarrete JM, Wei X et al. (2015) Nicotinamide N-methyltransferase regulates hepatic nutrient metabolism through Sirt1 protein stabilization. *Nat Med* 21 (8):887–894. 10.1038/nm.3882 [PubMed: 26168293]
63. Zhang J, Chen Y, Liu C, Li L, Li P (2020) N(1)-Methylnicotinamide improves hepatic insulin sensitivity via activation of SIRT1 and inhibition of FOXO1 acetylation. *J Diabetes Res* 2020:1080152. 10.1155/2020/1080152 [PubMed: 32280711]
64. Chaker L, Ligthart S, Korevaar TI et al. (2016) Thyroid function and risk of type 2 diabetes: a population-based prospective cohort study. *BMC Med* 14(1):150. 10.1186/s12916-016-0693-4 [PubMed: 27686165]
65. Panveloski-Costa AC, Silva Teixeira S, Ribeiro IM et al. (2016) Thyroid hormone reduces inflammatory cytokines improving glycaemia control in alloxan-induced diabetic wistar rats. *Acta Physiol (Oxford)* 217(2): 130–140. 10.1111/apha.12647
66. Panveloski-Costa AC, Serrano-Nascimento C, Bargi-Souza P, Poyares LL, Viana GS, Nunes MT (2018) Beneficial effects of thyroid hormone on adipose inflammation and insulin sensitivity of obese Wistar rats. *Physiol Rep* 6(3):e13550. 10.14814/phy2.13550

67. Fernandez-Garcia JC, Delpino-Rius A, Samarra I et al. (2019) Type 2 diabetes is associated with a different pattern of serum poly-amines: a case(-)control study from the PREDIMED-plus trial. *J Clin Med* 8(1):71. 10.3390/jcm8010071
68. Gao M, Zhao W, Li C et al. (2018) Spermidine ameliorates non-alcoholic fatty liver disease through regulating lipid metabolism via AMPK. *Biochem Biophys Res Commun* 505(1):93–98. 10.1016/j.bbrc.2018.09.078 [PubMed: 30241944]
69. Zhuang R, Ge X, Han L et al. (2019) Gut microbe-generated metabolite trimethylamine N-oxide and the risk of diabetes: a systematic review and dose-response meta-analysis. *Obes Rev* 20(6):883–894. 10.1111/obr.12843 [PubMed: 30868721]
70. Vasan SK, Noordam R, Gowri MS, Neville MJ, Karpe F, Christodoulides C (2019) The proposed systemic thermogenic metabolites succinate and 12,13-diHOME are inversely associated with adiposity and related metabolic traits: evidence from a large human cross-sectional study. *Diabetologia* 62(11):2079–2087. 10.1007/s00125-019-4947-5 [PubMed: 31309263]
71. Liu P, Zhu W, Chen C et al. (2020) The mechanisms of lysophosphatidylcholine in the development of diseases. *Life Sci* 247: 117443. 10.1016/j.lfs.2020.117443 [PubMed: 32084434]
72. Zhu D, Ran Y (2012) Role of 15-lipoxygenase/15-hydroxyeicosatetraenoic acid in hypoxia-induced pulmonary hypertension. *J Physiol Sci* 62(3): 163–172. 10.1007/s12576-012-0196-9 [PubMed: 22331435]
73. Puri P, Wiest MM, Cheung O et al. (2009) The plasma lipidomic signature of non-alcoholic steatohepatitis. *Hepatology* 50(6): 1827–1838. 10.1002/hep.23229 [PubMed: 19937697]
74. Boden G, Chen X, Ruiz J, White JV, Rossetti L (1994) Mechanisms of fatty acid-induced inhibition of glucose uptake. *J Clin Invest* 93(6):2438–2446. 10.1172/JCII17252 [PubMed: 8200979]
75. Shimabukuro M, Zhou YT, Levi M, Unger RH (1998) Fatty acid-induced beta cell apoptosis: a link between obesity and diabetes. *Proc Natl Acad Sci U S A* 95(5):2498–2502. 10.1073/pnas.95.5.2498 [PubMed: 9482914]
76. Egan BM, Greene EL, Goodfriend TL (2001) Nonesterified fatty acids in blood pressure control and cardiovascular complications. *Curr Hypertens Rep* 3(2): 107–116. 10.1007/s11906-001-0021-y [PubMed: 11276390]
77. Forouhi NG, Koulman A, Sharp SJ et al. (2014) Differences in the prospective association between individual plasma phospholipid saturated fatty acids and incident type 2 diabetes: the EPIC-InterAct case-cohort study. *Lancet Diabetes Endocrinol* 2(10): 810–818. 10.1016/S2213-8587(14)70146-9 [PubMed: 25107467]
78. Heim M, Johnson J, Boess F et al. (2002) Phytanic acid, a natural peroxisome proliferator-activated receptor (PPAR) agonist, regulates glucose metabolism in rat primary hepatocytes. *FASEB J* 16(7):718–720. 10.1096/fj.01-0816fje [PubMed: 11923221]

Research in context

What is already known about this subject?

- The development of diabetes is heralded by a long subclinical phase in which subclinical physiological alterations in metabolism contribute to insulin resistance and dysglycaemia
- These features may be especially important in younger individuals at high lifetime metabolic risk
- Circulating metabolites may provide additional information to understand the pathophysiology and risk stratification in young adults at risk for diabetes

What is the key question?

- What is the association of the circulating metabolome with composite phenotypes of diabetes susceptibility (and diabetes itself) in young adults across race?

What are the new findings?

- We studied >2000 participants in the Coronary Artery Risk Development in Young Adults (CARDIA) study using plasma metabolite profiling to build metabolomic signatures of clinical diabetes risk, demonstrating their association to long-term diabetes risk additive to clinical risk factors
- Metabolites implicated in clinical risk comprised both known (e.g, branched-chain amino acids) and novel (e.g, microbe-derived) metabolites
- A metabolic signature representing proinflammatory adiposity was associated with risk of diabetes in young adults in the CARDIA study, with replication in middle-aged and older adults in the Framingham Heart Study

How might this impact on clinical practice in the foreseeable future?

- Development of targeted metabolite profiles with absolute quantification of metabolites may provide insight into risk, independent of known risk factors, early in adulthood as a way to interrupt risk of diabetes

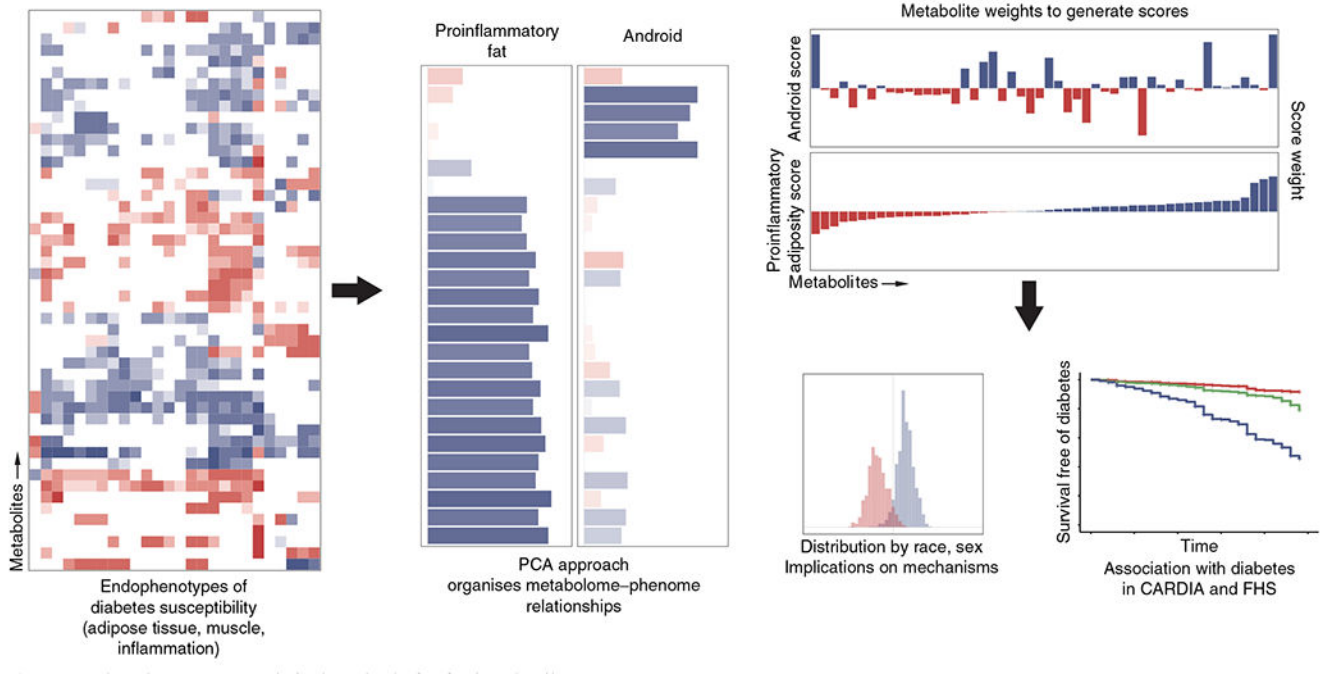


Fig. 1. Study scheme. See Statistical methods for further details

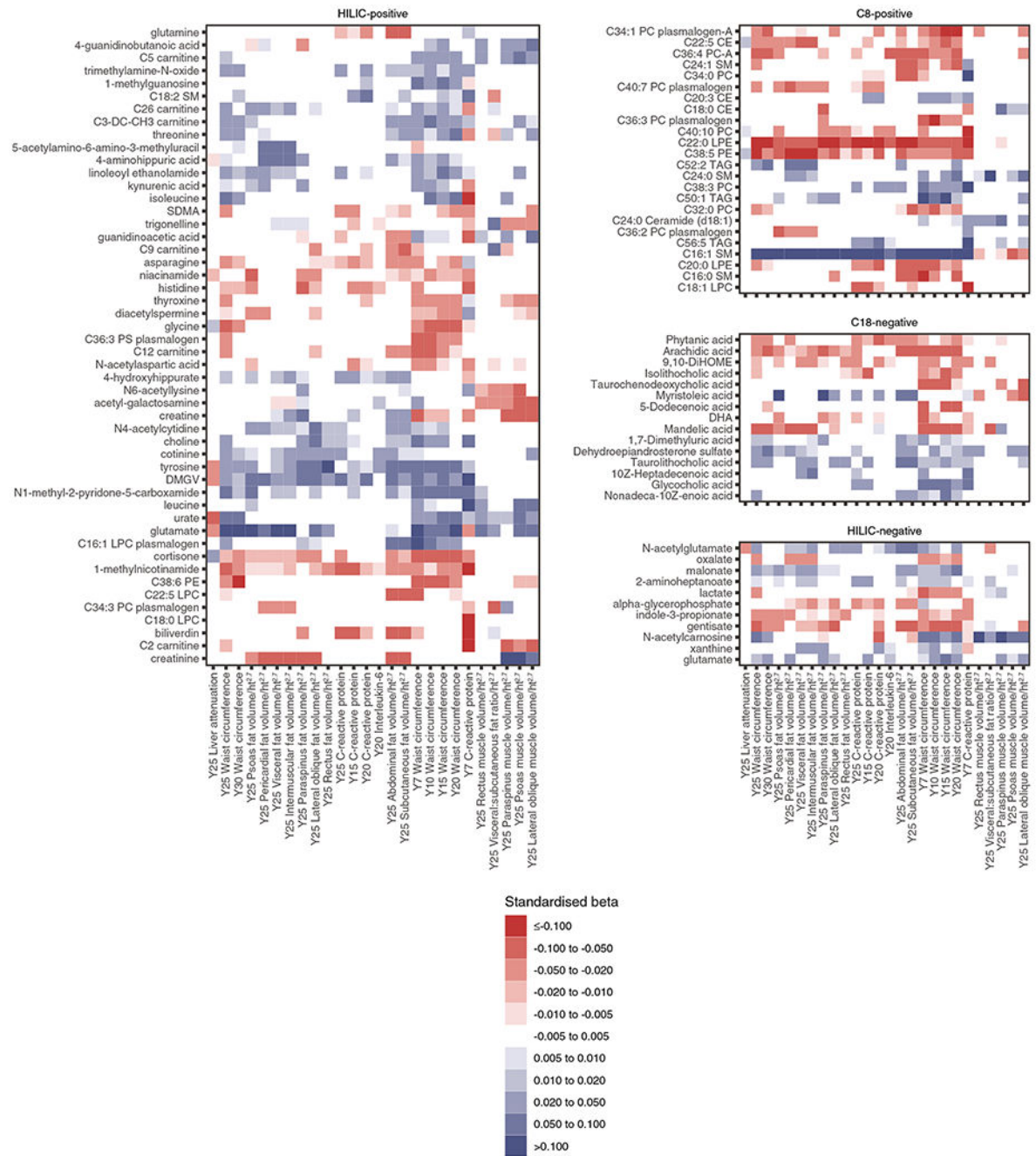


Fig. 2. Heatmap of penalised regression coefficients (LASSO), representing the 100 largest magnitude penalised regression coefficients for metabolites across endophenotypes. Each column specifies a given endophenotype. Each row is a metabolite, ordered using complete-link-age clustering. Regression models are unadjusted and the metabolites shown are those that survive LASSO regression. These penalised regression coefficients therefore represent the joint multivariable relationship between the endophenotypes and the metabolome and are entered into PCA. This represents the first step in our analytical

strategy (from Fig. 1). CE, cholesterol ester; DHA, docosahexaenoic acid; 9,10-diHOME, 9,10-dihydroxy-12-octadecenoic acid; DMGV, dimethylguanidinovalerate; ht., height; PE, phosphatidylethanolamine; PS, phosphatidylserine; SDMA, symmetric dimethylarginine; SM, sphingomyelin; TAG, triacyl-glycerol; Y, CARDIA examination year

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

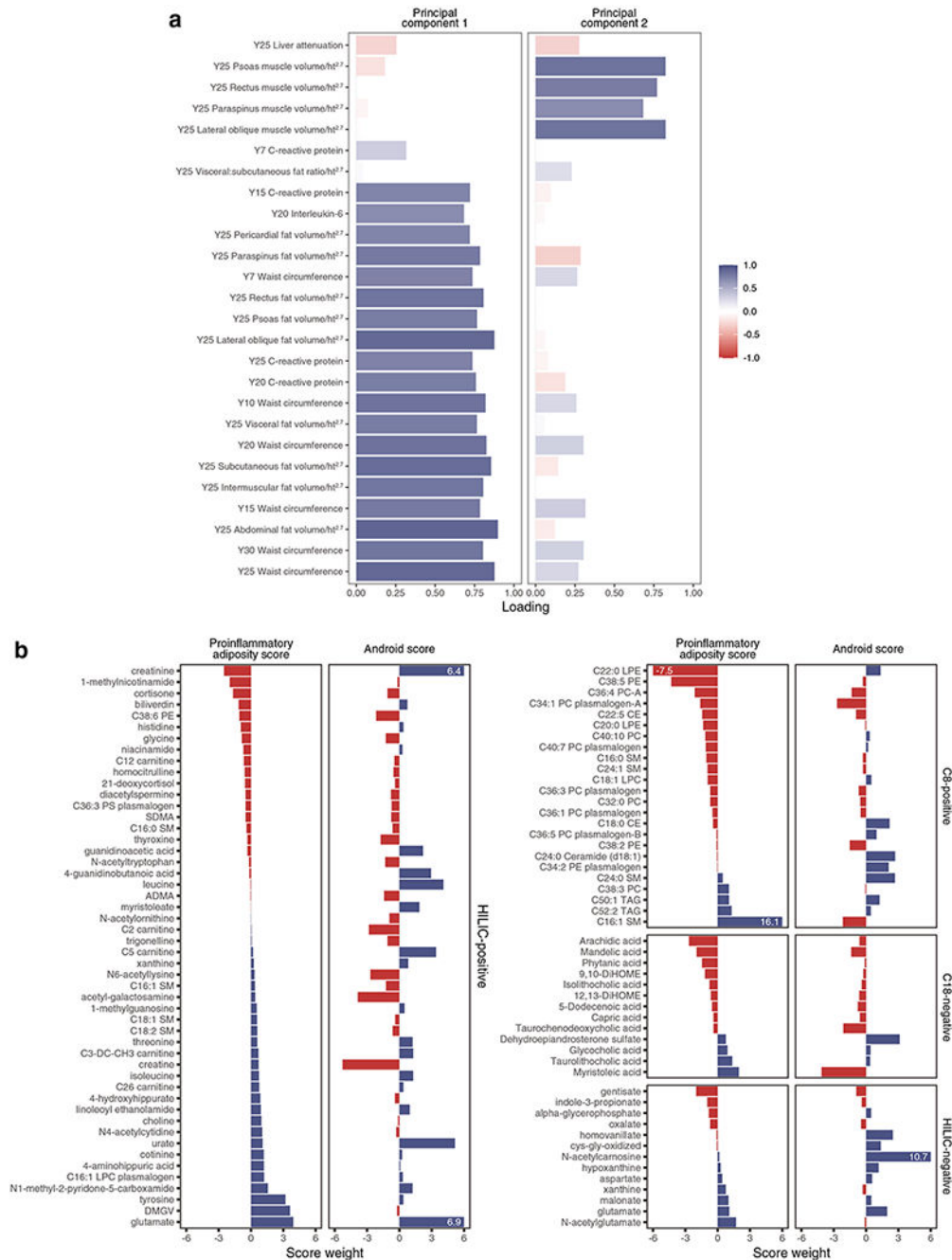


Fig. 3. PCA of penalised regression coefficients. This approach effectively reduces the dimensionality of the metabolome–endophenotype relation, described by principal components having numerical loadings to describe their dependence on metabolism and phenotype, (a) Loadings of the first two principal components in a varimax rotation PCA of these regression coefficients (59.2% total variance explained in the metabolome–endophenotype relationship) on each phenotype: principal component 1 was labelled ‘proinflammatory adiposity’, given its high loadings on visceral and visceral-like depots;

principal component 2 was labelled ‘android’, given its high loading on muscle depots, (b) Weights of selected metabolites. All weights shown in ESM Table 3. For weights that would extend beyond the horizontal axis limits (−6 to +6), we indicate the number, and the bar is at the limit. These ‘weights’ are then multiplied by each respective metabolite level and added across all constituent metabolites to generate a metabolite-based score of proinflammatory adiposity or android-like body composition principal component. ADMA, asymmetric dimethylarginine; CE, cholesterol ester; 9,10-diHOME, 9,10-dihydroxy- 12-octadecenoic acid; 12,13-diHOME, 12,13-dihydroxy-9-octadecenoic acid; DMGV, dimethylguanidinovalerate; ht., height; PE, phosphatidylethanolamine; SM, sphingomyelin; SDMA, symmetric dimethylarginine; TAG, triacylglycerol

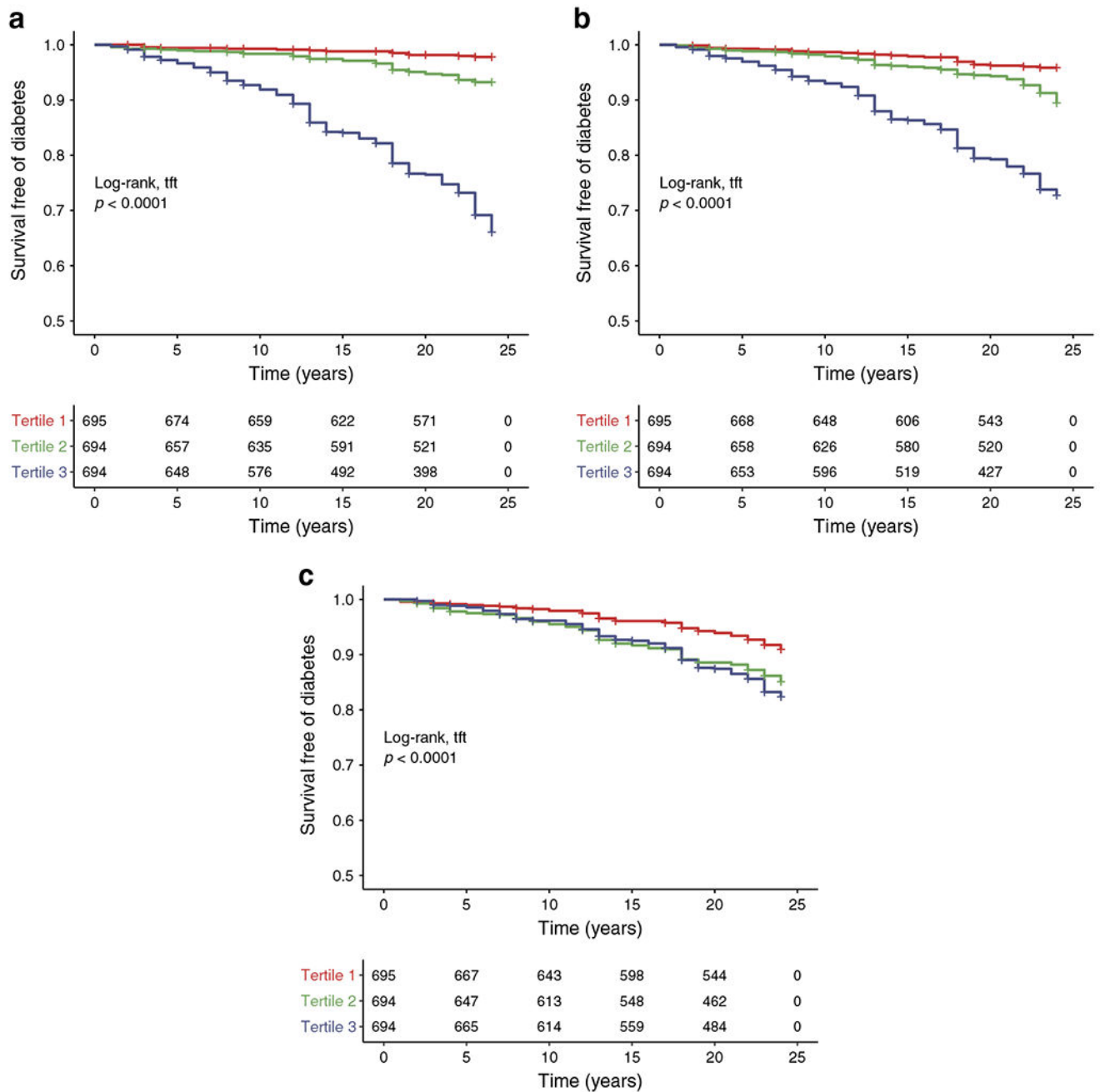


Fig. 4. Diabetes-free survival in CARDIA study participants by metabolite-based scores for endophenotypes and diabetes. Survival free of incident diabetes over time in is shown by tertile of the diabetes (a), proinflammatory adiposity (b) and android metabolite (c) scores. The numbers in the table below the plots represent the numbers at risk. Tft, test for trend

Table 1
Clinical/pathogenic correlates of metabolites associated with diabetes and related phenotypes

Metabolite	Proinflammatory adiposity score	Android score	Clinical/pathogenic correlates
Amino acids and derived products			
Glutamate	↑	↑	Associated with diabetes risk [50]; glutamine associated with fat inflammation [19]
Glutamine	↓	↑	
Glycine	↓	↓	Complex effects on glucose control via alterations in liver fat, decreased abdominal adiposity, lowered oxidative stress, GLP-1 secretion [16]; associated with decreased diabetes risk [17]
BCAAs (leucine, isoleucine, valine) ^a	↑	↑	Branched-chain amino acids: greater diabetes risk [1]; in context of high-fat diet, BCAA feeding potentiates insulin resistance (via mTOR phosphorylation and alteration in muscle metabolism) [11]; catabolism of BCAAs may also be impacted in obesity and contribute to diabetes [51]
Taurine	↓	↓	Improvement in insulin resistance in rodents; variable effects on body mass [52]; taurine supplementation associated with reduced inflammation and oxidative stress in humans with diabetes [53]
Tyrosine	↑	↑	Associated with insulin resistance in children [54] and in adults (also decreased insulin secretion [15])
Histidine	↓	↑	Administration to obese women relieved insulin resistance [55]; low levels have been associated with inflammation and oxidative stress [56]
Asparagine	↓	↓	Lower levels associated with incident diabetes [57]; protective roles in maintaining intestinal epithelial barrier via downregulation of intestinal TLR4 [40]
ADMA	↑	↑	Increased in obesity and impacts insulin resistance in muscle, decreasing glucose uptake [36]
DMGV	↑	↓	Associated with increased liver fat and diabetes, modifies the benefit of exercise on insulin sensitivity and can be lowered with exercise [12, 14]
Citrulline	↓	↓	Member of the NO pathway; supplementation reduces inflammatory markers after 12 weeks in non-alcoholic fatty liver disease [58]
Hippurate	↑	↓	Microbial product; higher levels may reduce risk of the metabolic syndrome [25]; associated with less visceral fat [59]
Cinnamoylglycine	↓	↓	Gut microbial product; associated with decreased BMI [24]
Other metabolites			
Urate ^a and related purine catabolites (e.g. xanthine, hypoxanthine, inosine)	↑	↑	Induction of diabetes increases circulating purine catabolites [60]
1-Methylnicotinamide	↓	↓	Metabolic effects via enhancing muscle lipolysis (myokine) [61]; supplementation in model systems reduces cholesterol and increases hepatic insulin sensitivity via enhancing sirtuin-1 stability [62, 63]
Thyroxine	↓	↓	Associated with lower incident diabetes [64]; precursors (T3) reduce systemic inflammation [65], dysglycaemia and adipose tissue inflammation in rodents [66]
Diacetylspermine	↓	↓	Polyamine; slightly higher levels of diacetylspermine derivatives may be present in individuals with diabetes and precursor spermine related to risk of diabetes [67]; supplementation of spermidine limits diet-induced obesity-related fatty liver disease via AMPK phosphorylation in model system [68]

Metabolite	Proinflammatory adiposity score	Android score	Clinical/pathogenic correlates
Trigonelline	↑	↓	Diet-related metabolite; reduces inflammation and oxidative stress and protects from diabetes-related complications [26, 27]
TMAO	↑	↑	Gut microbial product of red meat consumption; associated with higher diabetes risk [69]
12,13-diHOME	↓	↓	Lipoic acid catabolite; increased during exercise [9]; involved in brown fat activation [49] and inversely related to adiposity [70]
Lipids			
LPCs	Single metabolite ^b		Bioactive lipid derivatives of plasma membranes, with complex roles of LPCs in insulin resistance [28–31] and inflammation [71], levels of select LPCs are higher in normal-weight individuals relative to obese individuals, and change with weight loss [44]
LPEs	Single metabolite ^b		
Sphingomyelins/ceramide metabolites	Single metabolite ^b		Implication in obesity and insulin resistance [21]; ceramides impact insulin signalling [41]
Plasmalogens	Single metabolite ^b		Association of select plasmalogens with insulin resistance in animal models [23]; potential function as a sink for oxidative stress [45]
Fatty acids and eicosanoids			
Arachidonic acid catabolites (5-, 15-HETE)	Single metabolite ^b		5-HETE is an eicosanoid product of 5-lipoxygenase central to inflammation, linked to diabetes complications (including cardiovascular [46] and retinal [47]); 15-HETE elevated at sites of inflammation [72] and is associated with fatty liver disease [73]
Long-chain fatty acids (oleic acid, palmitic acid, palmitoleic acid, arachidic acid)	Single metabolite ^b		NEFA, which can promote insulin resistance [74] and beta cell apoptosis [75]; may be important in diabetes-related metabolic diseases (e.g. hypertension [76]); several linked to decreased diabetes risk (e.g. arachidic acid) in large populations [77]
Phytanic acid	Single metabolite ^b		Chlorophyll metabolite; potential role in activation of hepatic PPARs and increased glucose uptake in the liver [78]

Metabolite score weights are presented in ESM Table 3

^aThe weighting is displayed for a given metabolite class (e.g. for branched-chain amino acids, isoleucine is representative)

^b'Single metabolite' signifies that these classes of metabolites were significant in single metabolite regressions

↑, metabolites with positive weights; ↓, metabolites with negative weights; ADMA, asymmetric dimethylarginine; AMPK, AMP-activated protein kinase; BCAAs, branched-chain amino acid; DMGM, dimethylguanidinovaleic acid; GLP-1, glucagon-like peptide-1; PPAR, peroxisome proliferator-activated receptor; TLR4, toll-like receptor 4; TMAO, trimethylamine-N-oxide

Table 2

Survival analysis for incident diabetes in CARDIA study participants

Model description	HR		C-index		p value	NRI	R ² (base)	R ² (with score)
	HR [95% CI]	p value	C-index, base (95% CI)	C-index (95% CI)				
Diabetes score								
Unadjusted	3.74 (3.26, 4.29)	<0.0001	-	0.814 (0.787, 0.842)	<0.0001	1.052 (0.896, 1.168)	0.79	0.79
Age-, sex- and race-adjusted	3.73 (3.22, 4.32)	<0.0001	0.627 (0.591, 0.663)	0.816 (0.789, 0.844)	<0.0001	0.940 (0.810, 1.056)	0.2	0.79
Fully adjusted with BMI	3.41 (2.85, 4.07)	<0.0001	0.762 (0.731, 0.794)	0.827 (0.800, 0.853)	<0.0001	0.795 (0.654, 0.925)	0.59	0.82
Fully adjusted with waist circumference	3.43 (2.84, 4.14)	<0.0001	0.769 (0.738, 0.800)	0.826 (0.799, 0.852)	<0.0001	0.768 (0.619, 0.892)	0.62	0.81
Fully adjusted with waist circumference & CRP	3.56 (2.93, 4.32)	<0.0001	0.773 (0.741, 0.804)	0.826 (0.799, 0.852)	<0.0001	0.686 (0.557, 0.835)	0.63	0.82
Age-, sex-, race- and FHS DM risk score ^a -adjusted	3.57 (3.07, 4.15)	<0.0001	0.686 (0.653, 0.719)	0.818 (0.790, 0.845)	<0.0001	0.923 (0.779, 1.045)	0.34	0.8
Proinflammatory adiposity score								
Unadjusted	2.38 (2.12, 2.67)	<0.0001	-	0.741 (0.707, 0.775)	<0.0001	0.780 (0.635, 0.891)	0.57	0.57
Age-, sex- and race-adjusted	2.24 (1.99, 2.53)	<0.0001	0.627 (0.591, 0.663)	0.751 (0.718, 0.784)	<0.0001	0.672 (0.553, 0.800)	0.2	0.6
Fully adjusted with BMI	2.10 (1.72, 2.55)	<0.0001	0.762 (0.731, 0.794)	0.781 (0.750, 0.812)	0.02	0.404 (0.253, 0.568)	0.59	0.68
Fully adjusted with waist circumference	1.86 (1.52, 2.28)	<0.0001	0.769 (0.738, 0.800)	0.781 (0.750, 0.812)	0.06	0.387 (0.209, 0.528)	0.62	0.67
Fully adjusted with waist circumference & CRP	1.88 (1.50, 2.36)	<0.0001	0.773 (0.741, 0.804)	0.781 (0.749, 0.812)	0.18	0.294 (0.123, 0.433)	0.63	0.67
Age-, sex-, race- and FHS DM risk score ^a -adjusted	2.13 (1.88, 2.41)	<0.0001	0.686 (0.653, 0.719)	0.760 (0.728, 0.793)	<0.0001	0.597 (0.466, 0.739)	0.34	0.62
Android score								
Unadjusted	1.39 (1.21, 1.59)	<0.0001	-	0.589 (0.552, 0.626)	<0.0001	0.083 (-0.040, 0.226)	0.095	0.095
Age-, sex- and race-adjusted	3.56 (2.73, 4.66)	<0.0001	0.627 (0.591, 0.663)	0.719 (0.685, 0.753)	<0.0001	0.489 (0.339, 0.625)	0.2	0.45
Fully adjusted with BMI	2.24 (1.67, 3.00)	<0.0001	0.762 (0.731, 0.794)	0.778 (0.749, 0.808)	0.03	0.284 (0.105, 0.417)	0.59	0.64
Fully adjusted with waist circumference	2.10 (1.56, 2.81)	<0.0001	0.769 (0.738, 0.800)	0.781 (0.751, 0.811)	0.06	0.270 (0.106, 0.432)	0.62	0.65
Fully adjusted with waist circumference & CRP	2.14 (1.59, 2.87)	<0.0001	0.773 (0.741, 0.804)	0.785 (0.756, 0.815)	0.03	0.260 (0.110, 0.422)	0.63	0.67
Age-, sex-, race-, and FHS DM risk score ^a -adjusted	3.12 (2.37, 4.10)	<0.0001	0.686 (0.653, 0.719)	0.738 (0.705, 0.770)	0.0004	0.410 (0.257, 0.545)	0.34	0.5

Each HR represents the hazard for a 1-SD change in the corresponding metabolite score (proinflammatory adiposity or android score) in a model with the noted adjustments (in the 'model description'); the 'model description' notes the other clinical covariates that form the 'base' model for each comparison (e.g. C-index or NRI comparison)

^aThe FHS DM score is described in the Methods text under 'Single metabolite to endophenotype regression'

CRP, C-reactive protein; DM, type 2 diabetes

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3

Survival analysis for incident diabetes in FHS participants

Model description	HR		C-index		NRI		R ² (with score)	
	HR (95% CI)	p value	C-index, base (95% CI)	C-index (95% CI)	p value	R ² (base)	R ² (with score)	
Diabetes score								
Unadjusted	1.75 (1.56, 1.96)	<0.0001	-	0.666 (0.635, 0.697)	<0.0001	0.495 (0.360, 0.647)	0.27	
Age- and sex-adjusted	1.70 (1.52, 1.92)	<0.0001	0.607 (0.575, 0.639)	0.689 (0.661, 0.718)	<0.0001	0.451 (0.297, 0.588)	0.33	
Fully adjusted with BMI	1.15 (0.99, 1.33)	0.07	0.836 (0.811, 0.861)	0.840 (0.816, 0.864)	0.01	0.042 (-0.083, 0.262)	0.77	
Fully adjusted with waist circumference	1.10 (0.95, 1.27)	0.21	0.841 (0.817, 0.866)	0.843 (0.819, 0.867)	0.09	0.017 (-0.133, 0.235)	0.79	
Proinflammatory adiposity score								
Unadjusted	1.88 (1.67, 2.12)	<0.0001	-	0.686 (0.655, 0.717)	<0.0001	0.503 (0.351, 0.658)	0.32	
Age- and sex-adjusted	1.95 (1.71, 2.21)	<0.0001	0.607 (0.575, 0.639)	0.711 (0.682, 0.740)	<0.0001	0.526 (0.378, 0.675)	0.4	
Fully adjusted with BMI	1.33 (1.14, 1.56)	0.0004	0.836 (0.811, 0.861)	0.841 (0.817, 0.865)	0.05	0.115 (-0.020, 0.373)	0.78	
Fully adjusted with waist circumference	1.23 (1.05, 1.45)	0.01	0.841 (0.817, 0.866)	0.845 (0.821, 0.869)	0.07	0.156 (-0.057, 0.343)	0.79	
Android score								
Unadjusted	1.27 (1.13, 1.43)	<0.0001	-	0.566 (0.533, 0.599)	<0.0001	0.250 (0.116, 0.402)	0.055	
Age- and sex-adjusted	1.23 (1.04, 1.46)	0.02	0.607 (0.575, 0.639)	0.616 (0.583, 0.648)	0.20	0.059 (-0.086, 0.185)	0.14	
Fully adjusted with BMI	1.19 (0.97, 1.45)	0.09	0.836 (0.811, 0.861)	0.837 (0.812, 0.861)	0.53	0.060 (-0.114, 0.273)	0.77	
Fully adjusted with waist circumference	1.19 (0.98, 1.45)	0.08	0.841 (0.817, 0.866)	0.841 (0.817, 0.866)	0.82	0.067 (-0.098, 0.267)	0.79	

Each HR represents the hazard for a 1-SD change in the corresponding metabolite score (proinflammatory adiposity or android score) in a model with the noted adjustments (in the 'model description'); the 'model description' notes the other clinical covariates that form the 'base' model for each comparison (e.g. C-index or NRI comparison)