A whole blood-based transcriptional risk score for nonobese type 2 diabetes predicts dynamic changes in glucose metabolic traits

Running title: Blood-based transcriptional risk score, diabetes

Yanan Hou,1,2* Huajie Dai,1,2* Na Chen,1,2* Zhiyun Zhao,1,2* Qi Wang,1,2 Tianzhichao Hou,1,2 Jie Zheng,1,2 Tiange Wang,1,2 Mian Li,1,2 Hong Lin,1,2 Shuangyuan Wang,1,2 Ruizhi Zheng,1,2 Jieli Lu,1,2 Yu Xu,1,2 Yuhong Chen,1,2 Ruixin Liu,1,2 Guang Ning,1,2 Weiqing Wang,1,2 Yufang Bi,1,2† Jiqiu Wang,1,2† Min Xu1,2†

Author affiliations:

1. Department of Endocrine and Metabolic Diseases, Shanghai Institute of Endocrine and Metabolic Diseases, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China;

2. Shanghai National Clinical Research Center for metabolic Diseases, Key Laboratory for Endocrine and Metabolic Diseases of the National Health Commission of the PR China, Shanghai Key Laboratory for Endocrine Tumor, State Key Laboratory of Medical Genomics, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China.

* These authors contributed equally to the work

† These authors co-supervised the work

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
†Address of Correspondence:

Min Xu, MD&PhD, Professor of medicine, Shanghai National Clinical Research Center for Endocrine and Metabolic Diseases, Shanghai Institute of Endocrine and Metabolic Diseases, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, 200025, China. E-mail: della.xumin@shsmu.edu.cn;

Jiqiu Wang, MD&PhD, Professor of medicine, Department of Endocrine and Metabolic Diseases, Shanghai Institute of Endocrine and Metabolic Diseases, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China, 200025. Email: wangjq@shsmu.edu.cn;

Yufang Bi, MD&PhD, Professor of medicine, Department of Endocrine and Metabolic Diseases, Shanghai Institute of Endocrine and Metabolic Diseases, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China. Email: byf10784@rjh.com.cn.
Abstract

Background: The performance of peripheral blood transcriptional markers in evaluating the risk of type 2 diabetes (T2D) with normal weight is unknown. We developed a whole blood-based transcriptional risk score (wb-TRS) for nonobese T2D and assessed its contributions to disease risk and dynamic changes in glucose metabolism.

Methods and findings: We developed the wb-TRS in 1105 participants aged ≥40 years and in normal weight for up to 10 years from a well-defined community-based cohort with blood transcriptome data and validated it in an external dataset (253 overweight/obese participants from a dietary intervention trial with 3 repeated transcriptome data). Potential biology significance and causal inference were also explored. The wb-TRS included 144 transcripts. Compared to the lowest tertile, wb-TRS in tertile 3 associated with 8.68-folds (95% confidence interval [CI], 3.51-21.5), and each 1-unit increment associated with 2.57-folds (95% CI, 1.86-3.56) higher risk of nonobese T2D, after adjustments for traditional risk factors. Furthermore, baseline wb-TRS was significantly associated with dynamic changes in average, daytime, nighttime and 24h glucose and HbA1c, and area under the curve of glucose measured in the continuous glucose monitoring during 6-month of intervention. The wb-TRS improved the predicting performance for nonobese T2D in a model with fasting glucose, triglycerides and demographic and anthropometric parameters. Mitch analysis implicated oxidative phosphorylation, cholesterol metabolism and mTORC1 signaling involved in nonobese T2D pathogenesis. Transcriptome-wide Mendelian
randomization supported causal effects of gene transcripts such as RAB1A and GCC1-PAX4 on nonobese T2D risk.

**Conclusions:** A whole blood based nonobese T2D associated TRS was validated to predict dynamic changes in glucose metabolism. These findings also suggested several genes and biological pathways that might involve in the pathogenesis of nonobese T2D.

**Keywords:** nonobese type 2 diabetes, glucose metabolic traits, blood transcriptome, transcriptional risk score, risk prediction, transcriptome-wide Mendelian randomization
Abbreviations:

- AUC – area under the curve
- BMI – body mass index
- CPM – count per million
- CI – confidence interval
- DBP – diastolic blood pressure
- DPMH – Dietary Pattern and Metabolic Health
- eQTL – expression quantitative trait locus
- GEE – generalized estimating equation
- HbA1c – hemoglobin A1c
- HOMA-IR – homeostasis model assessment of insulin resistance
- HOMA-β – homeostasis model assessment of beta cell function
- HDL-C – high-density lipoprotein cholesterol
- MET – metabolic equivalent of task
- Mitch analysis – a multi-contrast gene set enrichment analysis
- LDL-C – low-density lipoprotein cholesterol
- LASSO – least absolute shrinkage and selection operator
- OGTT – oral glucose tolerance test
- OR – odds ratio
- RCS – restricted cubic spline
- RNA-seq – RNA sequencing
- SBP – systolic blood pressure
T2D – Type 2 diabetes

TRS – transcriptional risk score

TWMR – Transcriptome-wide Mendelian randomization

TSGE – Tissue-specific gene expression
**Author summary**

**Why was this study done?**

- Blood transcriptome has been emerging as a potential novel approach to understand contribution of genes expression to biological functions and to identify key genes that confer to complex disease susceptibility, such as cardiovascular, psychiatry and chronic obstructive pulmonary disease.
- Rare studies have investigated the performance of the peripheral blood transcriptional markers in evaluating the risk and prediction of type 2 diabetes (T2D).

**What did the researchers do and find?**

- We developed a whole blood-based transcriptional risk score (wb-TRS) for T2D risk in a longitudinal community-based cohort.
- The wb-TRS was positively associated with risk of T2D and glucose metabolic traits. Furthermore, the baseline wb-TRS as well associated with dynamic changes of glucose metabolism measured in the continuous glucose monitoring in response to 6-month dietary intervention.
- The wb-TRS performed better and was complementary to fasting glucose, triglycerides and the demographic and anthropometric factors in predicting T2D.
- Transcriptome-wide Mendelian randomization and Mitch analysis and supported several gene transcripts (RAB1A and GCC1-PAX4) and biological pathways (oxidative phosphorylation, mTORC1 signaling, and cholesterol metabolism) that might involve in the pathogenesis of T2D.
What do these findings mean?

- Blood transcriptome highly related to and reflect risk of T2D and predict dynamic changes in glucose metabolism. It may lend insight into discerning potential causal genes and biological mechanisms in the pathogenesis of nonobese T2D.
Introduction

Type 2 diabetes (T2D) has been a worldwide epidemic in recent decades leading high rates of early mortality and co-morbidity and possessing a considerable health and economic burden for the patients [1]. Early prediction and detection of individuals with glucose dysregulation and type 2 diabetes is extremely important for the prevention and management of these disorders. Previous reliable risk prediction models usually comprise the anthropometric measurements and traditional laboratory data. Recent studies have provided evidence of the prediction effect by including genetic factors and metabolic molecules in the models [2, 3]. However, as an essential biological and pathogenesis process, the performance and effect of gene transcription profile in evaluating T2D risk has not been well examined yet [4].

Tissue-specific gene expression (TSGE) is phenotypic manifestations integrating both genetic and environmental variations, which confers further insights to the intricate pathophysiology of diseases [5]. However, most tissues such as pancreas, liver, etc. cannot be readily obtained without invasive procedures in humans. As an alternative, blood transcriptome could better predict TSGE for ~60% of the genes on average, with up to 81.7% in skeletal muscle, 71.5% in visceral adipose tissue, and 52.4% in pancreas [5]. Blood transcriptome has been emerging as a potential novel approach for complex diseases to understand contribution of genes expression to biological functions and to identify key genes that confer to complex disease susceptibility [6-9], such as cardiovascular, psychiatry and chronic obstructive pulmonary disease. However, rare studies have investigated the performance of the
peripheral blood transcriptional markers in evaluating the risk and prediction of T2D.

Though obesity is one of its major risk factors, the T2D risk is still high in normal weight, especially in east Asians. One in three of diagnosed T2D had a body mass index (BMI) of less than 25 kg/m², indicating the importance of metabolic health in the pathogenesis regardless of adiposity [10, 11]. Studies focused on refining T2D subtypes highlighting nonobese T2D, which characterized by a lower BMI, early onset of age, severe insulin deficiency, and serious complications [12-15]. The east Asians developed T2D risk at a relatively lower level of BMI [16]. The existing evidence indicated that specific genes, pathways, and biological mechanisms might underly the pathophysiology of nonobese T2D.

Thus, we developed a whole blood-based transcriptional risk score (wb-TRS) of nonobese T2D in 1105 participants aged 40 and above in an ongoing longitudinal Chinese cohort, in which the participants have maintained normal weight for up to 10 years. To validate its performance in a more general age and BMI range, we examined its predictive value on dynamic change in glucose metabolism in an independent external dataset (253 young overweight/obese participants from a dietary intervention trial with 3 repeated whole blood-based transcriptome data). Furthermore, we explored the potential genes and biological pathways that may conferred disease’s susceptibility and pathogenesis and performed transcriptome-wide Mendelian randomization (TWMR) to investigate additional evidence supporting a causal links between gene transcripts and T2D risk in a separate East Asian population and European population. The wb-TRS would provide an additional layer of biological
interpretability for T2D.

Methods

Study population

The current analysis was based on one of the follow-up circles of our previous community-based cohort of in Shanghai, China [17-19]. Briefly, all the participants have taken part in a comprehensive examination that included a questionnaire and biochemical measurements at baseline in 2008-2009, the first (5-years) in 2013 and second (10-years) follow-up visits in 2018, respectively. The participants with known T2D before baseline examination were excluded.

According to the information at baseline and the 5-year follow-up examination, at the 10-year follow-up visit, we selected the participants whose BMI <25kg/m² in past 10 years to take part in the examination sequentially and collected the whole blood samples for RNA-sequencing (RNA-seq). Finally, 1183 participants attended the follow-up visit finished the blood sampling for RNA-seq and 1105 with high-quality sequencing data were included in the present analysis. The flowchart of study participants was shown in Fig S1.

The study protocol was approved by the Institutional Review Board of Rui-Jin Hospital affiliated to Shanghai Jiao Tong University School of Medicine. All participants gave the written consents for the study.

Data collection
A standard questionnaire was adopted to collect information on sociodemographic characteristics, disease and medical history, and lifestyles. Current smoking or drinking was defined as participants consumed cigarettes or alcohol regularly in the past 6 months. We used Global Physical Activity Questionnaire (GPAQ) to obtain frequency, duration, and intensity of physical activity. Physical activity level was evaluated as Metabolic Equivalent of Task (MET) minutes per week. Body height, weight and waist circumference were measured using a standard protocol. BMI was calculated as body weight (kg) divided by squared height (m²). Three seated systolic- and diastolic blood pressure (SBP and DBP) were measured (Omron HEM-752 Fuzzy; Omron, Dalian, China) at the non-dominant arm consecutively with 1 min interval after at least 5-minutes rest. The three readings were averaged for analysis.

**Biochemical measurements**

Venous blood samples were obtained after an overnight fasting (>10 hours). All participants underwent a 75 g oral glucose tolerance test (OGTT). Fasting and 2h-OGTT glucose was measured using glucose oxidase method on an automated analyzer (ADVIA-1650 Chemistry System; Bayer); fasting and 2h-OGTT insulin using electrochemiluminescence assay on an automated analyzer (Roche Diagnostics, Basel, Switzerland); and hemoglobin A1c (HbA1c) using high pressure liquid chromatography on an automated analyzer (VARIANT II Hemoglobin Testing System, Bio-Rad Laboratories). Fasting triglycerides, total cholesterol, low-density and high-density lipoprotein cholesterol (LDL-C and HDL-C) were measured on the
automated analyzer (ADVIA-1650 Chemistry System; Bayer).

**Definition of T2D**

T2D was diagnosed as fasting glucose $\geq 7.0$ mmol/l, 2h-OGTT glucose $\geq 11.1$ mmol/l, HbA1c $\geq 6.5\%$, and/or taking antidiabetic therapy [20]. The homeostasis model assessment of insulin resistance (HOMA-IR) was calculated as fasting glucose $(\text{mmol/l}) \times \text{fasting insulin (uIU/ml)} / 22.5$. The HOMA index for beta cell function $(\text{HOMA-}\beta)$ was calculated as $20 \times \text{fasting insulin (uIU/ml)} / (\text{fasting glucose [mmol/l]} - 3.5)$.

**RNA-seq**

The whole blood samples were collected by using the Paxgene RNA tubes according to the manufacture’s protocol, shipped and stored in -80°C before RNA extracting. RNA was extracted by using the PAXgene Blood miRNA Kit Secondary BD Hemogard Closures (Catalog no 763134, Qiagen, Germany). Nanodrop One spectrometer (Thermo Fisher Scientific Inc., USA) and Qubit®3.0 Fluorometer (Life Technologies, CA, USA) were used to determine RNA purity and concentration. Isolated total RNA was analyzed on an Agilent Bioanalyzer 2100 (Agilent technologies, Santa Clara, CA, US) for RNA integrity number quality check. We excluded 78 samples with low RNA quality. 1ug RNA was used as input material for the RNA-seq preparations. RNA-seq strand-specific libraries were constructed using the VAHTS Total RNA-
seq (H/M/R) Library Prep Kit (Vazyme, China). Briefly, RNA was purified by magnetic beads after removal of rRNA. Then RNA was fragmented into small pieces using divalent cations for 8 minutes at 94°C. The cleaved RNA fragments were copied into first strand cDNA using reverse transcriptase and random primers. Second strand cDNA synthesis was subsequently performed using DNA Polymerase I and RNase H. These cDNA fragments then go through the end repair process, addition of a single ‘A’ base, and ligation of the adapters. The products were purified and enriched by PCR to create the final cDNA library. Purified libraries were quantified and validated by Qubit® 3.0 Fluorometer and Agilent 2100 bioanalyzer to confirm the insert size and calculate the mole concentration. Cluster was generated by cBot after the library diluted to 10 pM. Cluster were then sequenced in 150-bp paired-end mode on the Illumina NovaSeq 6000 platform (Illumina, USA) generating arounds 12 gigabases (Gb) reads per samples.

The base call and sequence quality across all generated FASTQ data was evaluated using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc). We removed adapters, low quality reads, ribosome RNA reads, and reads with a sequence length less than 25 to obtain clean reads. Clean reads were mapped to human GRCh38.102 with Hisat2 (Hierarchical Indexing for Spliced Alignment of Transcripts, version 2.0.5) using default parameters. Overall, we obtained RNA-seq data with base mass (Q30) ≥90%, average mapped ratio was 94%. Count data were filtered to include transcripts with > 1 count per million (CPM) in 99% of T2D cases or non-T2D controls.
wb-TRS construction

The 1105 participants with blood-based transcriptome data were randomly assigned to training (75%) and testing (25%) dataset (R package caret) (Fig 1). To identify genes that could better distinguish the diseases status, we included the transcripts with gene expression level >1CPM in 99% of T2D cases or non-T2D controls and with a nominal difference P value <0.01 between these two groups in the wb-TRS features screening progress (n=757 gene transcripts). The wb-TRS construction was carried out in the training dataset and validated in testing and an external dataset. A multi-contrast gene set enrichment (Mitch) analysis [21] and Reactome pathway analysis [22] were performed to identify key genes and pathways.

In detail, we performed a penalized logistic regression framework using R package glmnet to develop the wb-TRS for nonobese T2D. We adopted the least absolute shrinkage and selection operator (LASSO) regression model to model training, which provides shrinks coefficients toward zero and automated feature selection. We adopted 10-fold cross-validation procedure to determine the lambda (0.006) that maximized the area under curve (AUC) for predicting T2D (Fig S2). The regression model with maximum AUC was used to calculate the wb-TRS.

Statistical analysis

Multivariable logistic regression analyses were used to assess the risk of nonobese T2D in relation to wb-TRS. Odds ratio (OR) and the corresponding 95% confidence
interval (CI) were calculated. The linear regression model was to assess the association of wb-TRS with the cardiometabolic traits. The restricted cubic spline (RCS) analysis was used to test any nonlinear associations between wb-TRS and glucose metabolism. The receiver operator characteristic curves (ROC) were used to evaluate and compare the predictive values of 6 models for nonobese T2D (R package pROC) including: 1. primary model (age, sex, BMI and family history of diabetes); 2. wb-TRS model; 3. fasting glucose and triglycerides model; 4. primary model adding fasting glucose and triglycerides (full); 5. primary model adding wb-TRS; 6. primary model adding fasting glucose and triglycerides (full), and wb-TRS.

Independent validation

To validate the predictive value of wb-TRS for the dynamic changes in glucose metabolism, we performed analysis using data of 3 repeated blood-based transcriptome data and glucose metabolism in the Dietary Pattern and Metabolic Health (DPMH) study, a parallel-arm randomized controlled feeding trial conducted in 2019 (URL: https://www.clinicaltrials.gov; Unique identifier: NCT03856762) [23]. A total of 253 overweight/obese participants (aged 25 to 60 years, BMI ≥ 24.0 kg/m², fasting glucose ≥ 5.6 mmol/l) with transcriptome data and glucose metabolic traits at baseline, the first (3-months) and second (6-months) follow-up visits were included in the validation analysis. The whole blood RNA was collected and underwent RNA-seq following the same procedure with the above cohort study. We adopted generalized estimating equations (GEE) models to assess the
The association of three-repeated measurements of wb-TRS with three-repeated measurements of the glucose metabolism during the 6-months intervention. We further fixed the generalized linear models to assess the effects of baseline wb-TRS on dynamic changes of glucose metabolism at 3- and 6-months of intervention.

**The biological pathway and mechanism analysis**

To gain biological insight behind the wb-TRS, we performed several bioinformatics analyses. We divided the participants in the initial cohort into 3 subgroups based on wb-TRS tertiles. We performed a multi-contrast gene set enrichment analysis by using R package mitch (v.1.4.0). The Mitch analysis uses a rank-MANOVA statistical approach to identify sets of genes that exhibit joint enrichment or divergent responses across multiple contrast. The input data was the log₂(fold change) of all detected transcripts (n=54,458, including protein coding gene, lncRNA, and snRNA) for comparison between the bottom and mid tertile, and the top and mid tertile, respectively. The input genesets provided were 50 genesets: h.all.v7.5.1.symbols.gmt in Molecular Signatures Database (MSigDB). Enrichment score (ES) was obtained and ranged from -1 to 1. Meanwhile, we used the 144 gene transcripts included in the wb-TRS as input profiles to perform Reactome pathway analysis on reactome website ([https://reactome.org/PathwayBrowser/#TOOL=AT](https://reactome.org/PathwayBrowser/#TOOL=AT)).

**Transcriptome-wide Mendelian randomization**

We performed a two-sample transcriptome-wide Mendelian randomization (TWMR)
to identify causal associations of gene transcripts included in wb-TRS with T2D risk using the TwoSampleMR. The summary statistics of genome-wide association study of T2D was from the Asian Genetic Epidemiology Network (AGEN) [16] and the DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) [24], respectively. We selected the significant expression quantitative trait loci (eQTLs with \( P < 5 \times 10^{-8} \)) obtained from eQTLGen Consortium [25] and with linkage disequilibrium (LD) based on \( r^2 < 0.01 \) as valid instruments variants (IVs) used in TWMR. When no shared eQTLs were available between exposures and outcomes, proxies with LD \( r^2 > 0.8 \) were substituted. TWMR causality tests were performed using the inverse variance weighted method by meta-analyzing the effect of each IV from Wald ratio. Heterogeneity and pleiotropy of the results were assessed using Cochran’s Q statistic or MR-Egger intercept, respectively. All analyses were performed using SAS (version 9.3; SAS Institute, Cary, NC, USA) and R software (version 3.5.1; R Core Team). A two-sided P value < 0.05 was considered statistically significant.

**Results**

**Clinical characteristics of study participants**

In the 1105 participants, 305 were defined as nonobese T2D (27.6%), 737 (66.7%) were women, with an average age of 64.5 (standard deviation [SD], 7.0) years old. The balanced clinical characteristics between the training and testing dataset indicated a low risk of selection bias introduced by random sampling (Table S1).
Construction and distribution of wb-TRS

Finally, 144 important features (gene transcripts) were included in the wb-TRS, which weights were displayed in Table S2. The wb-TRS was estimated as gene expression level of transcript1 × weight1 + gene expression level of transcript2 × weight2 + ... + gene expression level of transcript144 × weight144.

The wb-TRS was positively associated with age (β 0.98, SE 0.15, P < 0.0001), and was slightly higher in men (-1.7, SD 1.5) than that in women (-2.1, SD 1.3). The mean wb-TRS was -0.5 (SD 0.9) in T2D and -2.5 (SD 1.1) in non-diabetes in training dataset, and -0.9 (SD 1.4) in T2D and -2.4 (SD 1.3) in non-diabetes in testing dataset.

Fig S3 displays the violin plot of the wb-TRS and gene transcripts most negatively (FUOM) and positively (RPF1) associated with T2D in training dataset.

Association of wb-TRS with nonobese T2D

Table 1 shows association of wb-TRS with nonobese T2D risk in testing dataset. Each 1-unit increment in wb-TRS (one score) was associated with a 2.49-folds higher risk of nonobese T2D (95%CI 1.84-3.36), after adjustments of age, sex, BMI and family history of diabetes. Additional adjustment for potential confounders including lifestyles, education levels, waist circumference, SBP and DBP, and lipids profiles did not change the results appreciably (OR 2.57, 95% CI 1.86-3.56). Specifically, each 1-unit increment in wb-TRS was associated with 3.40-folds higher risk of newly diagnosed nonobese T2D (95%CI 2.00-5.81), and 2.33-folds higher risk of diabetes.
with duration ≥ 5 years (95% CI 1.56-3.48). Compared with the bottom tertile, the middle or top tertile of wb-TRS was associated with 2.46- (95% CI 0.95-6.35) and 8.68-folds (95% CI 3.51-21.5) higher risk of nonobese T2D.

We then examined the associations of wb-TRS with nonobese T2D risk stratified by several traditional risk factors of T2D (Fig S4). A stronger association was identified in age <65 years and women (a marginal interaction between TRS and sex was observed with P for interaction =0.051).

**Association of wb-TRS with glucose metabolism**

In both training and testing dataset, the wb-TRS was positively associated with fasting glucose, 2h-OGTT glucose, and HbA1c; but inversely associated with HOMA-β, in the fully adjusted model (Table 2). The RCS analysis showed similar results, except for that it demonstrated a J shaped relationships between wb-TRS and HbA1c and logHOMA_β (all P for non-linear ≤ 0.001) (Fig 2).

The GEE analysis by using the 3 repeated measurements of glucose metabolism traits and wb-TRS at baseline, 3 month- and 6 month-visits during the intervention showed that each 1-unit increment in wb-TRS was associated with 0.11mmol/l (SE 0.05) and 0.15mmol/l (SE 0.02) increment of 1h- and 2h-OGTT glucose, respectively. Interestingly, the baseline wb-TRS significantly affected the dynamic changes of cardiometabolic traits at 6- months of dietary intervention (Table 3). Baseline wb-TRS significantly associated with changes in average daytime, nighttime and 24h glucose and HbA1c levels and the AUC of glucose measured in the continuous
glucose monitoring at 6 month of diet intervention (all P<0.03).

**wb-TRS performed well in predicting T2D**

In both training and testing dataset, wb-TRS performed well in predicting nonobese T2D, even when adding wb-TRS into a model with fasting glucose, triglycerides, and the primary model (Fig 3). In total dataset, a model with wb-TRS along with fasting glucose, triglycerides, and variables in primary model significantly improved the AUC (0.93), as compared to the primary model (AUC =0.65), the fasting glucose and triglycerides model (AUC=0.82), and the fasting glucose, triglycerides adding to the primary model (full model, AUC =0.84) for predicting T2D (Table S3). The wb-TRS improved predicting performance in a model with fasting glucose, triglycerides, and the demographic and anthropometry parameters (AUC-TRS+full: 0.93 [0.91-0.95] vs. AUC-full: 0.84 [95% CI 0.81-0.87]; P < 0.0001).

**The biological pathway and mechanism analysis**

The Mitch analysis compared the bottom versus middle and top versus middle tertiles of wb-TRS (Fig 4). The top versus middle tertile group comparison demonstrated downregulation of MYC target V2, and the bottom versus middle tertile group comparison demonstrated upregulation of oxidative phosphorylation, mTORC1 signaling, cholesterol metabolism, peroxisome, fatty acid metabolism, etc. (Fig 4A). The significant metabolic pathways were displayed in Fig 4B. The small GTPase (RAB1A gene) enriched in mTORC1 signaling, fatty acid oxidation gene (CPT1A)
enriched in oxidative phosphorylation pathway, and cholesterol efflux gene (ABCA1)

enriched in cholesterol metabolism pathway.

The Reactome analyses suggested that the 144 gene transcripts included in wb-TRS
were involved in NR1H2 (LXRβ) and NR1H3 (LXRα) regulate gene expression to
limit cholesterol uptake, regulation of lipid metabolism by PPAR alpha pathway, etc.
The cholesterol uptake suppression gene (MYLIP) enriched in LXRβ and LXRα
regulate gene expression to limit cholesterol uptake pathway (Table S4).

Causal relationships between gene transcripts included in wb_TRS and T2D

Gene transcripts (n=19) with no validated IVs were removed from TWMR analysis.

Among the 125 remaining gene transcripts, 6 gene transcripts (GCC1-PAX4,
AKR1A1, AAMP, RAB1A, ITGA1, and ZNF639) were causally linked to T2D risk in
AGEN and 1 gene transcript (RAB1A) was successfully validated in DIAGRAM
(Table S5, P < 0.0004 =0.05/125, where 125 corresponds to the number of gene
transcripts tested in TWMR). Specifically, RAB1A enriched in mTOCR1 signaling is
causally associated with 38% (95% CI: 1.18-1.62, P=8.0×10⁻⁵) and 47% (95% CI:
1.21-1.77, P=5.4×10⁻⁵) increase of T2D risk in AGEN and DIAGRAM, respectively.
GCC1-RAX4 converted 17% decrease of T2D risk (95% CI: 0.80-0.88, P=5.2×10⁻¹⁵)
in AGEN.

Discussion

In this study, we developed a wb-TRS for T2D risk in a longitudinal community-
based cohort in which the participants were middle aged and elderly and maintained nonobese for up to 10 years. The wb-TRS was positively associated with risk of T2D and could reflect dynamic change in glucose metabolism such as fasting and 2h-post loading glucose, HbA1c, and beta cell dysfunction in the observational cohort. Furthermore, the baseline wb-TRS well associated with the changes of glucose metabolism measured in the continuous glucose monitoring in response to the 6-month dietary intervention in a panel of younger and overweight participants. The wb-TRS performed better and was a complementarity to fasting glucose, triglycerides and the demographic and anthropometric factors in predicting T2D. The wb-TRS supported several gene transcripts (RAB1A and GCC1-RAX4) and biological pathways (oxidative phosphorylation, cholesterol metabolism, and mTORC1 signaling) that might involve in the etiology of nonobese T2D.

To the best of our knowledge, the present study is the first to examine the effects of wb-TRS on nonobese T2D risk and dynamic change of related cardiometabolic traits by using peripheral blood transcriptional markers. Recently, several studies proposed the concept of TRS or polygenetic TRS, which used cumulative gene expression information to provide an additional layer of biological interpretability and performed better than both traditional risk factors and polygenic risk score for predicting Crohn's disease, chronic obstructive pulmonary disease, and lung function [8, 26-28]. Our study is in line with previous studies showing transcriptional variance is a reliable intermediary mediating associations of genotypes with complex phenotypes [5, 6, 29] and provides evidence of implement of blood transcriptional profile could yield a
promising better predictive effect on T2D and the dynamic changes in related cardiometabolic traits.

In our study, the wb-TRS could well predict dynamic changes in glucose metabolism in response to 6 months of 25% calorie restriction dietary intervention in overweight/obese subjects [23]. In our trial (DPMH), a 25% calorie restriction produced an appreciable weight reduction and improved glucose homeostasis [23]. Calorie restriction improves cardiometabolic risk including atherogenic dyslipidemia (elevated triglycerides and fasting glucose, reduced HDL-C, diabetes, hypertension, and large waist circumference) [30, 31]. Previous findings from the Comprehensive Assessment of Long-term Effects of Reducing Intake of Energy (CALERIETM) study, a 24-month, multicenter, randomized controlled calorie restriction trial, showed that in young to middle aged, normal to slightly overweight adults, two years of >12% caloric restriction improved traditional cardiometabolic risk factors within ranges considered normal [31]. Recent study suggested that in healthy adults without obesity, a 12% calorie restriction led to improved reducing inflammation and branched chain amino acids and shifting lipoproteins from atherogenic to cholesterol transporting [32]. We suspected that the genes and pathways included in the wb-TRS played key roles in cardiometabolic beneficial effects in response to the calorie restriction dietary intervention, which might help to explore the underlying mechanisms for the beneficial effects.

Our results indicated several potential pathways and gene transcripts that might involve in the etiology process of nonobese T2D, which suggested that transcriptional
information in blood could be capable of reflecting diseases risk and could bridge the
eknowledge gap on how gene expression affects nonobese T2D. We found mTORC1
signaling directly regulates β-cell functional maturation and identity maintenance
[33]. Of note, by using TWMR analysis, we also demonstrated that the RAB1A gene
encoding a small GTPase, which is enriched in mTORC1 signaling, was positively
and causally associated with T2D risk in both east Asians and Europeans. Another
T2D susceptibility gene (GCC1-PAX4) was also causally associated with T2D risk in
east Asians [34]. PAX4 implicated in a Japanese case of maturity onset diabetes of the
young that is only 30kb further away from GCC1 is an outstanding candidate for T2D
given its involvement in pancreatic islet development [34]. This molecular
mechanism of nonobese T2D was consistent with that the nonobese Asian Indians
with young-onset diabetes have lower genetically determined beta cell function [35].
We identified oxidative phosphorylation and cholesterol metabolism pathways
involved in nonobese T2D. Of note, fatty acid oxidation gene (CPT1A) enriched in
oxidative phosphorylation pathway might mediate mitochondrial overload increasing
insulin resistance in muscle, lipid accumulation in liver, and lipotoxicity in pancreatic
beta cells thus involved in homeostasis of glucose metabolism [36, 37]. Cholesterol
efflux gene (ABCA1) enriched in cholesterol metabolism pathway might via its
downregulation to inhibit cholesterol transportation to HDL and excretion to bile and
faeces, and then increase lipids profile and deteriorates cardio-metabolism [38].
In Reactome analysis, cholesterol uptake suppression gene (MYLIP) enriched in
LXRβ and LXRα regulate gene expression to limit cholesterol uptake pathway.
Inhibition of MYLIP increases cellular cholesterol accumulation and leads to pancreatic β-cell dysfunction and insulin resistance [39]. We suspected that cholesterol metabolism in liver and pancreas and beta cell dysfunction involved in nonobese T2D but warrant further study in the future. Besides, the consistency of transcriptional levels of these genes in whole-blood and specific tissue is another interesting issue.

The major strengths of our current study lie in our adopted analytical approaches as well as the well-characterized study cohorts. We adopted high-throughput RNA-seq technology, a penalized regression model with 10-fold cross validation method to develop the wb-TRS, and successfully validated its performance in an independent external cohort, which guaranteed the statistical power and robustness of our findings and conclusions. In terms of the study cohorts, our initial cohort included T2D whose BMI maintained normal for up to 10 years, which are expected to better capture underling pathophysiology mechanism of nonobese T2D.

Our study, however, also has several limitations. First, transcriptome data were measured only at 10 years follow-up visit in initial cohort, which could not establish a causal relationship between the wb-TRS and nonobese T2D risk. It would also be greater interesting to see whether transcriptional change would continue to be associated with incidence risk of nonobese T2D. Second, complex diseases are final phenotypes of transcriptional dysregulation in multiple tissues. Though an appreciable proportion of genes expressed in tissues is expressed in peripheral whole blood [40], it still should be better to interpret the present wb-TRS as one of potential essential
complementary prediction tools for nonobese T2D, rather than presenting an exact
pathophysiological process of it. Nevertheless, the individual’s blood transcriptome
performed well in predicting tissue-specific expression and performed as well as
measured tissue expression in predicting complex diseases [5]. Third, we did not
incorporate metabolic markers in the analysis, which restrained us from comparing
the predictive effect between them. Jointly considering the genetic and metabolic
profiling and transcriptional level of alterations would provide more precise and
accurate evaluation of T2D risk and prediction efficacy. Last, both the initial and
validation cohort consisted of participants from east Asia, cautions should be
exercised to generalize the finding to other ethnicity groups.

In conclusion, we developed a wb-TRS that comprising a panel of genes based on
their blood expression levels, and that may involve important biological function
pathways was positively associated with nonobese T2D risk and predicted dynamic
change in glucose metabolic traits. The study supported that blood transcriptome
highly related to and reflect impaired glucose metabolism. It may lend insight into
discerning potential causal genes and biological mechanisms in the pathogenesis of
nonobese T2D.
Acknowledgments

We thank all the study participants for their participation and contribution.

Data availability

The eQTLs data used in this study are available on the eQLTGen Consortium website [https://eqtlgen.org/]. The summary statistics of genome-wide association study of T2D was available in the Asian Genetic Epidemiology Network (AGEN) [https://blog.nus.edu.sg/agen/] and the DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) [http://www.diagram-consortium.org/downloads.html/]. RNA-seq related additional data reported in this manuscript are available from the corresponding author upon reasonable request.

Funding statement

This work was funded by the National Natural Science Foundation of China (82270859, 91957124, 91857205, 81930021, 81970728, 82070880 and 82088102), the Shanghai Municipal Education Commission–Gaofeng Clinical Medicine Grant Support (20152508 Round 2), the Shanghai Shenkang Hospital Development Center (SHDC12019101, SHDC2020CR1001A, and SHDC2020CR3064B). MX, JW, ML, TW, ZZ, RL, YX, JL, YB, WW, and GN are members of innovative research team of high-level local universities in Shanghai.

Duality of interest
No potential conflicts of interest relevant to this article were reported.

**Contribution statement**

M.X., Y.B., and J.W. contributed to the conception and design of the study. Y.H., H.D., and N.C. performed statistical analysis and interpreted data. Y.H., and M.X. drafted the manuscript. G.N. and W.W. revised the manuscript critically for important intellectual content. Z.Z., Q.W., T.H., J.Z., T.W., M.L., H.L., S.W., R.Z., J.L., Y.X., Y.C., and R.L. contributed to the acquisition of data. All authors were involving in revising the manuscript and had final approval of the version to be published. M.X., Y.B. and J.W. are the guarantors of this work and, as such, had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.


34. Cho YS, Chen CH, Hu C, Long J, Ong RT, Sim X, et al. Meta-analysis of genome-


cholesterol on insulin secretion: a novel mechanism for pancreatic beta-cell dysfunction.


Supporting information

Table S1. Clinical characteristics of study participants.

Table S2. Weights of 144 gene transcripts included in wb-TRS, ordered from most negative to most positive associated with nonobese type 2 diabetes.

Table S3. Receiver operator characteristic curves for prediction of nonobese type 2 diabetes.

Table S4. Top 30 Reactome pathways associated with genes in wb-TRS.

Table S5. Transcriptome wide Mendelian randomization shows the causal effect of gene transcripts included in wb_TRS on type 2 diabetes.

Fig S1. Flowchart of study participants.

Fig S2. Cross-validation results for least absolute shrinkage and selection operator (LASSO) logistics-model for nonobese type 2 diabetes in training dataset.

Fig S3. Violin plot of wb-TRS and gene transcripts most negatively (FUOM) and positively (RPF1) associated with nonobese type 2 diabetes.

Fig S4. Stratified analysis. Data are odds ratio (OR) and 95% confidence interval (CI) calculated using multivariable logistic regression models. The results were adjusted for age, sex, body mass index, family history of diabetes, waist circumference, smoking and drinking status, education level, physical activity, systolic and diastolic blood pressure, total cholesterol, triglycerides, high-density and low-density lipoprotein cholesterol.
Figure Legends

Fig 1. The schematic study design. At the final follow-up visit (10-year’s), the whole blood samples of the participants with BMI <25kg/m² during the past 10 years were collected and underwent RNA sequencing (RNA-seq). The 1105 nonobese participants with RNA-seq data were randomly split into training (75%) and testing (25%) set. Transcripts with gene expression level >1 CPM in 99% of T2D cases or non-T2D controls and with a nominal difference P value <0.01 between these two groups in the wb-TRS features screening progress. In training dataset, we trained the LASSO logistic-model and adopted 10-fold cross-validation to conduct the transcriptional risk score (wb-TRS). In testing and external dataset, we tested the performance of wb-TRS. Finally, Mitch analysis and Reactome pathway analysis were performed to identify key genes and pathways. Abbreviations: T2D, type 2 diabetes; BMI, body mass index; CPM, counts per million; LASSO, least absolute shrinkage and selection operator; TRS, transcriptional risk score; Mitch analysis, multi-contrast gene set enrichment analysis.

Fig 2. Restricted cubic spline analysis for associations between the wb-TRS and cardiometabolic traits in testing dataset. Data were adjusted for age, sex, body mass index, family history of diabetes, waist circumference, smoking and drinking status, education level, physical activity, systolic and diastolic blood pressure, triglycerides, low-density lipoprotein cholesterol and anti-diabetic therapy.

Fig 3. Receiver operator characteristic curves for prediction of type 2 diabetes. Primary model comprised age, sex, BMI and family history of diabetes.
Abbreviations: wb-TRS, whole blood-based transcriptional risk score; BMI, body mass index; FBG, fasting glucose index; TG, triglycerides; and AUC, area under the curve.

**Fig 4. Mitch analysis.** The gene expression level in the top and bottom tertile of the wb-TRS were compared with those in the middle tertile. **A**, a heatmap of enrichment scores for gene expression contrasts between these two comparisons with hierarchical clustering. Cells in red indicate a positive enrichment score, whereas blue indicated a negative enrichment score. **B**, the significant metabolic pathways. Abbreviations: botVmid, bottom versus middle tertile; topVmid, top versus middle tertile.
Fig 1.

Type 2 diabetes (T2D) and non-diabetic participants (aged ≥40 ys at baseline, BMI ≤25kg/m² during the past 10 years)

Whole blood samples collected by using the Paxgene RNA tubes at the last follow up visit

RNA-sequencing (RNA-seq) and biochemical measurements

Primary study cohort
Participants with RNA-seq data (N=1105)

Validation: DPMH study
Participants with 3 repeated RNA-seq data (N=253)

Step 1: Filter >1 CPM in 99% of T2D or non-T2D (11021 genes)
Step 2: Filter difference P value <0.01 (757 genes)

Training dataset (n=829, 75%)

Construction of the wb-TRS
LASSO logistic-model in training set 10-fold cross-validation (144 genes)

Testing dataset (n=276, 25%)

Performance of the wb-TRS
Association study
ROC analysis

Biology mechanism analysis
(Mitch, Reactome)
Fig 2.

- **Fasting glucose, mmol/l**:
  - P for linearity < 0.001
  - P for non-linearity = 0.07

- **2h-OGTT glucose, mmol/l**:
  - P for linearity < 0.001
  - P for non-linearity = 0.26

- **HbA1C, %**:
  - P for linearity < 0.001
  - P for non-linearity = 0.0003

- **Log_{10} Fasting insulin, ul/l/ml**:
  - P for linearity = 0.30
  - P for non-linearity = 0.12

- **Log_{10} HOMA-IR**:
  - P for linearity = 0.28
  - P for non-linearity = 0.34

- **Log_{10} HOMA-B**:
  - P for linearity < 0.001
  - P for non-linearity = 0.03
Fig 3.

A. Training dataset

B. Testing dataset
Fig 4.

A. Color Key and Histogram

B. Pathways of Metth analysis

<table>
<thead>
<tr>
<th>Pathway</th>
<th>FDR</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oxidative Phosphorylation</td>
<td>3.99E-25</td>
<td>0.47</td>
</tr>
<tr>
<td>MTORC1 Signaling</td>
<td>1.66E-21</td>
<td>0.43</td>
</tr>
<tr>
<td>Adipogenesis</td>
<td>2.33E-16</td>
<td>0.37</td>
</tr>
<tr>
<td>Glycogenesis</td>
<td>4.84E-16</td>
<td>0.36</td>
</tr>
<tr>
<td>Myogenesis</td>
<td>6.34E-14</td>
<td>0.28</td>
</tr>
<tr>
<td>Fatty Acid Metabolism</td>
<td>3.01E-11</td>
<td>0.34</td>
</tr>
<tr>
<td>Peroxisome</td>
<td>4.11E-11</td>
<td>0.41</td>
</tr>
<tr>
<td>Cholesterol Homeostasis</td>
<td>2.56E-06</td>
<td>0.33</td>
</tr>
<tr>
<td>Bile Acid Metabolism</td>
<td>6.51E-06</td>
<td>0.27</td>
</tr>
<tr>
<td>Pancreas Beta Cells</td>
<td>6.75E-01</td>
<td>0.08</td>
</tr>
</tbody>
</table>
Table 1. Adjusted ORs (95% CI) of type 2 diabetes associated with whole blood-based transcriptional risk score (wb-TRS) in testing dataset.

<table>
<thead>
<tr>
<th>Tertile 1 (-7.0 ~ -2.5)</th>
<th>Tertile 2 (-2.5 ~ -1.4)</th>
<th>Tertile 3 (-1.4 ~ -3.9)</th>
<th>P for trend</th>
<th>Each 1-unit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR (95% CI)</td>
<td>P value</td>
<td>OR (95% CI)</td>
<td>P value</td>
</tr>
<tr>
<td>Type 2 diabetes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cases, n (%)</td>
<td>8 (9.0)</td>
<td>21 (23.3)</td>
<td>47 (48.5)</td>
<td></td>
</tr>
<tr>
<td>Model 1</td>
<td>1.00 (Ref.)</td>
<td>2.57 (1.05-6.30)</td>
<td>8.36 (3.56-19.7)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Model 2</td>
<td>1.00 (Ref.)</td>
<td>2.46 (0.95-6.35)</td>
<td>8.68 (3.51-21.5)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Newly diagnosed type 2 diabetes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cases, n (%)</td>
<td>2 (2.3)</td>
<td>13 (14.4)</td>
<td>19 (19.6)</td>
<td></td>
</tr>
<tr>
<td>Model 1</td>
<td>1.00 (Ref.)</td>
<td>6.41 (1.38-29.9)</td>
<td>17.2 (3.72-79.8)</td>
<td>0.0003</td>
</tr>
<tr>
<td>Model 2</td>
<td>1.00 (Ref.)</td>
<td>6.77 (1.30-35.3)</td>
<td>20.6 (3.94-108)</td>
<td>0.0003</td>
</tr>
<tr>
<td>Type 2 diabetes with duration ≥5 years</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cases, n (%)</td>
<td>6 (6.7)</td>
<td>7 (7.8)</td>
<td>27 (27.8)</td>
<td></td>
</tr>
<tr>
<td>Model 1</td>
<td>1.00 (Ref.)</td>
<td>1.08 (0.33-3.58)</td>
<td>5.74 (2.09-15.8)</td>
<td>0.001</td>
</tr>
<tr>
<td>Model 2</td>
<td>1.00 (Ref.)</td>
<td>1.17 (0.33-4.24)</td>
<td>6.12 (2.00-18.7)</td>
<td>0.002</td>
</tr>
</tbody>
</table>

Data were odds ratio (OR), 95% confidence interval (CI) calculated from the multivariable logistic regression models. Model 1: adjusted for age, sex, body mass index, and family history of diabetes; Model 2: further adjusted for waist circumference, smoking and drinking status, education
level, physical activity, systolic and diastolic blood pressure, total cholesterol, triglycerides, HDL-C, and LDL-C. Abbreviation: HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol.
Table 2. Association of each 1-unit increment in the wb_TRS with cardiometabolic traits (n=1105)

<table>
<thead>
<tr>
<th></th>
<th>Training dataset</th>
<th>Testing dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>β (SE)</td>
<td>P value</td>
</tr>
<tr>
<td>Body mass index, kg/m^2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 1</td>
<td>0.09 (0.06)</td>
<td>0.12</td>
</tr>
<tr>
<td>Model 2</td>
<td>-0.05 (0.04)</td>
<td>0.22</td>
</tr>
<tr>
<td>Waist circumference, cm</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 1</td>
<td>0.38 (0.14)</td>
<td>0.01</td>
</tr>
<tr>
<td>Model 2</td>
<td>0.35 (0.14)</td>
<td>0.01</td>
</tr>
<tr>
<td>Systolic blood pressure, mmHg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 1</td>
<td>1.52 (0.49)</td>
<td>0.002</td>
</tr>
<tr>
<td>Model 2</td>
<td>1.00 (0.39)</td>
<td>0.01</td>
</tr>
<tr>
<td>Diastolic blood pressure, mmHg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 1</td>
<td>0.41 (0.28)</td>
<td>0.14</td>
</tr>
<tr>
<td>Model 2</td>
<td>-0.07 (0.22)</td>
<td>0.77</td>
</tr>
<tr>
<td>Total cholesterol, mmol/l</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 1</td>
<td>-0.06 (0.03)</td>
<td>0.03</td>
</tr>
<tr>
<td>Model 2</td>
<td>-0.01 (0.01)</td>
<td>0.22</td>
</tr>
<tr>
<td>Triglycerides, mmol/l</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 1</td>
<td>0.09 (0.03)</td>
<td>0.001</td>
</tr>
<tr>
<td>Model 2</td>
<td>0.09 (0.03)</td>
<td>0.001</td>
</tr>
<tr>
<td>HDL-C, mmol/l</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 1</td>
<td>-0.02 (0.01)</td>
<td>0.01</td>
</tr>
<tr>
<td>Model 2</td>
<td>-0.01 (0.01)</td>
<td>0.43</td>
</tr>
<tr>
<td>LDL-C, mmol/l</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 1</td>
<td>-0.04 (0.02)</td>
<td>0.03</td>
</tr>
<tr>
<td>Model 2</td>
<td>-0.06 (0.02)</td>
<td>0.01</td>
</tr>
<tr>
<td>Fasting glucose, mmol/l</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 1</td>
<td>0.42 (0.03)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Model 2</td>
<td>0.25 (0.03)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>2h-OGTT glucose, mmol/l</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Model 1</td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>-------------</td>
<td>---------</td>
</tr>
<tr>
<td>HbA1c, %</td>
<td>1.16 (0.08)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Model 2</td>
<td>1.01 (0.08)</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

|                  | Model 1     |         |         |         |
| Log Fasting insulin, uIU/ml | 0.21 (0.02) | <.0001  | 0.19 (0.04) | <.0001 |
|                  | Model 2     |         |         |         |
| Log HOMA-IR      | 0.07 (0.02) | <.0001  | 0.10 (0.03) | 0.002  |

|                  | Model 1     |         |         |         |
| Log HOMA-β       | -0.04 (0.01) | <.0001  | -0.05 (0.01) | <.0001 |
|                  | Model 2     |         |         |         |

Data were β and standard error (SE). Linear regression model was used to evaluate associations of the wb_TRS (independent) with cardiometabolic traits (dependent variables). Model 1, adjusted for age, sex, body mass index, and family history of diabetes; Model 2, further adjusted for waist circumference, smoking and drinking status, education level, physical activity, systolic and diastolic blood pressure, triglycerides, LDL-C and anti-diabetic therapy. Abbreviation: HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; OGTT, oral glucose tolerance test; HbA1c, glycated hemoglobin A1c. HOMA-IR, homeostasis model assessment of insulin resistance; and HOMA-β, homeostasis model assessment of β cell function.
Table 3. The effects of baseline wb-TRS on changes of glucose metabolic traits during the 6-month dietary intervention (n=253)

<table>
<thead>
<tr>
<th>Glucose metabolic traits</th>
<th>At 3 months (β, SE)</th>
<th>P value</th>
<th>At 6 months (β, SE)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>OGTT</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>△Fasting glucose, mmol/l</td>
<td>-0.01 (0.02)</td>
<td>0.72</td>
<td>0.01 (0.02)</td>
<td>0.64</td>
</tr>
<tr>
<td>△0.5h-OGTT glucose, mmol/l</td>
<td>0.03 (0.05)</td>
<td>0.51</td>
<td>-0.0005 (0.06)</td>
<td>0.99</td>
</tr>
<tr>
<td>△1h-OGTT glucose, mmol/l</td>
<td>0.03 (0.08)</td>
<td>0.74</td>
<td>0.12 (0.08)</td>
<td>0.15</td>
</tr>
<tr>
<td>△2h-OGTT glucose, mmol/l</td>
<td>0.04 (0.07)</td>
<td>0.59</td>
<td>0.10 (0.07)</td>
<td>0.16</td>
</tr>
<tr>
<td><strong>Continuous glucose monitoring (CGM)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>△Average glucose, mmol/l</td>
<td>-0.21 (0.45)</td>
<td>0.64</td>
<td>1.42 (0.55)</td>
<td>0.01</td>
</tr>
<tr>
<td>△Average daytime glucose, mg/dl</td>
<td>-0.26 (0.48)</td>
<td>0.59</td>
<td>1.55 (0.60)</td>
<td>0.01</td>
</tr>
<tr>
<td>△Average nighttime glucose, mg/dl</td>
<td>-0.16 (0.42)</td>
<td>0.71</td>
<td>1.09 (0.48)</td>
<td>0.03</td>
</tr>
<tr>
<td>△AUC glucose, mmol*h/l</td>
<td>0.21 (0.45)</td>
<td>0.65</td>
<td>1.41 (0.55)</td>
<td>0.01</td>
</tr>
<tr>
<td>△AUC daytime glucose, mmol*h/l</td>
<td>0.26 (0.48)</td>
<td>0.59</td>
<td>1.56 (0.60)</td>
<td>0.01</td>
</tr>
<tr>
<td>△AUC nighttime glucose, mmol*h/l</td>
<td>0.16 (0.42)</td>
<td>0.70</td>
<td>0.21 (0.45)</td>
<td>0.65</td>
</tr>
<tr>
<td>△Estimated HbA1c</td>
<td>-0.004 (0.02)</td>
<td>0.81</td>
<td>0.05 (0.02)</td>
<td>0.01</td>
</tr>
<tr>
<td>△Estimated daytime HbA1c</td>
<td>-0.01 (0.02)</td>
<td>0.55</td>
<td>0.06 (0.02)</td>
<td>0.01</td>
</tr>
<tr>
<td>△Estimated nighttime HbA1c</td>
<td>0.01 (0.01)</td>
<td>0.58</td>
<td>0.04 (0.02)</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Data were change (β) and standard error (SE). P values were from the general linear models after adjusted for sex, baseline age, body mass index, waist circumference, smoking and drinking status, education level, physical activity, systolic and diastolic blood pressure, triglycerides, LDL-C and baseline values for the respective outcome trait. Change (Δ) level was calculated as the numeric value at follow-up visits minus that at baseline. Estimated HbA1c was calculated using the formula: (average
glucose+46.7)/28.7.