Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# Metabolic Trajectories and Clinical Outcomes in Prediabetes: A 2.5-Year Study of Reprogramming Patterns

**Guowang Xu**

xugw@dicp.ac.cn

Dalian Institute of Chemical Physics, Chinese Academy of Sciences https://orcid.org/0000-0003-4298-3554

**Pengwei Guan**

Dalian Institute of Chemical Physics, Chinese Academy of Sciences

**Rong Yu**

Shanghai Sixth People's Hospital Affiliated to Shanghai Jiao Tong University School of Medicine

**Peng Dou**

Dalian Institute of Chemical Physics, Chinese Academy of Sciences

**Fengyang Jing**

Dalian Institute of Chemical Physics, Chinese Academy of Sciences

**Di Yu**

Dalian Institute of Chemical Physics, Chinese Academy of Sciences

**Huating Li**

Shanghai Sixth People's Hospital Affiliated to Shanghai Jiao Tong University School of Medicine https://orcid.org/0000-0003-0526-5545

**Yueqiong Ni**

Shanghai Sixth People's Hospital Affiliated to Shanghai Jiao Tong University School of Medicine

**Hongfeng Mao**

Shanghai Key Laboratory of Diabetes Mellitus, Department of Endocrinology and Metabolism, Shanghai Diabetes Institute, Shanghai Clinical Center for Diabetes, Shanghai Sixth People’s Hospital

**Yan Pang**

Shanghai Sixth People's Hospital Affiliated to Shanghai Jiao Tong University School of Medicine

**Xiaolin Wang**

CAS Key Laboratory of Separation Science for Analytical Chemistry, Dalian Institute of Chemical Physics, China Academy of Sciences,

**Qi Li**

Dalian Institute of Chemical Physics, Chinese Academy of Sciences

**Xuhong Hou**

Shanghai Sixth People's Hospital Affiliated to Shanghai Jiao Tong University School of Medicine

**Xinyu Liu**

Dalian Institute of Chemical Physics, Chinese Academy of Sciences https://orcid.org/0000-0002-0564-7167

**Weiping Jia**

Department of Endocrinology and Metabolism, Shanghai Sixth People's Hospital Affiliated to Shanghai Jiao Tong University School of Medicine;Institute for Proactive Healthcare of Shanghai Jiao Tong U https://orcid.org/0000-0002-6244-2168

---

**Article**

**Keywords:**

**Posted Date:** November 3rd, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-11473658/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

**Additional Declarations:**

There is **NO** Competing Interest.

Ethics oversight: Servicebio Animal Welfare NO. 2025147.

# Metabolic Trajectories and Clinical Outcomes in Prediabetes: A 2.5-Year Study of Reprogramming Patterns

Pengwei Guan[1,4,5#], Rong Yu[2#], Peng Dou[1,4#], Fengyang Jing[1,4,5#], Di Yu[1,4], Huating Li[2,3], Yueqiong Ni[2], Hongfeng Mao[2,6], Yan Pang[2], Xiaolin Wang[1,4], Qi Li[1,4], Xuhong Hou[2], Xinyu Liu*[1,4,5], Guowang Xu*[1,4,5], Weiping Jia*[2,3]

[1] Metabolomics Subcenter of the National Genomics Data Center, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian, 116023, China

[2] State Key Laboratory of Metabolic Dysregulation and Prevention and Treatment of Esophageal Cancer, Shanghai Key Laboratory of Diabetes Mellitus, Department of Endocrinology and Metabolism, Shanghai Diabetes Institute, Shanghai Clinical Center for Diabetes, Shanghai Sixth People's Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai 200233, China

[3] Institute for Proactive Healthcare, Antai College of Economics and Management, Shanghai Jiao Tong University, Shanghai 200030, China

[4] Liaoning Province Key Laboratory of Metabolomics, Dalian, China.

[5] University of Chinese Academy of Sciences, Beijing 100049, China

[6] Department of Medicine, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China

[#]: Equal contribution to this work.

[*] **Corresponding authors:**

Prof. Dr. Guowang Xu, Metabolomics Subcenter of the National Genomics Data Center Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian 116023, China. Tel.: +86-0411-84379530. E-mail: xugw@dicp.ac.cn.

Dr. Xinyu Liu, Metabolomics Subcenter of the National Genomics Data Center Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian 116023, China. Tel.: +86-0411-84379532. E-mail: liuxy2012@dicp.ac.cn.

Prof. Dr. Weiping Jia, State Key Laboratory of Metabolic Dysregulation and Prevention and Treatment of Esophageal Cancer, Shanghai Key Laboratory of Diabetes Mellitus, Department of Endocrinology and Metabolism, Shanghai Diabetes Institute, Shanghai Clinical Center for Diabetes, Shanghai Sixth People's Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai 200233, China. E-mail: wpjia@sjtu.edu.cn.

## Abstract

Prediabetes is a dynamic state in which impaired glucose regulation (IGR) can revert to normoglycaemia or progress to type 2 diabetes (T2DM), yet the metabolic basis of these divergent trajectories remains unclear. We profiled serum metabolites in 975 individuals with IGR and 128 healthy controls over 2.5 years. IGR remission involved partial normalization of amino acid metabolism and attenuated fatty acid dysregulation, whereas progression to T2DM showed persistent branched-chain amino acid perturbation, acylcarnitine accumulation and amplified lipid dysregulation. Tetradecanedioic acid emerged as the most robust outcome-associated biomarker, validated across human cohorts, mouse progression and exercise-induced remission models, and two independent multicenter datasets, with graded accumulation across worsening glycaemic states. Tissue metabolomics and $^{13}C_6$-glucose tracing linked these alterations to hepatic lipid disturbance and impaired glucose-derived TCA cycle flux in skeletal muscle, identifying fatty acid remodeling as an early and partially reversible axis distinguishing IGR remission from progression.

**Subject terms:** Type 2 diabetes mellitus, Impaired glucose regulation, Metabolic reprogramming, Lipid remodeling

Diabetes mellitus is a systemic metabolic and vascular disorder defined by chronic hyperglycaemia and progressive cellular and vascular injury. Its global prevalence has nearly doubled over the past two decades, posing a major public health challenge.[1]. Type 2 diabetes mellitus (T2DM) accounts for most cases, is increasingly diagnosed at younger ages and substantially increases the risks of multisystem complications, imposing considerable clinical and socioeconomic burdens[2-4]**.** Identifying high-risk individuals before overt diabetes is therefore essential for preventing irreversible metabolic and vascular damage.

T2DM usually develops through a prolonged prediabetic phase of impaired glucose regulation (IGR), during which insulin resistance, β-cell stress and tissue-specific metabolic dysfunction may already be present[5]. However, IGR is not a unidirectional condition with divergent outcomes. Some individuals progress to T2DM, whereas others remain metabolically stable or revert to normoglycaemia. Conventional risk assessment, based largely on glycaemic measures and related cardiometabolic factors, does not fully capture this heterogeneity. Prospective metabolomics studies have repeatedly linked circulating amino acids, acylcarnitines, lipids, glycine-related pathways and bile acid metabolism to future T2DM risk[6-8]. A recent multi-cohort study of up to 23,634 initially diabetes-free individuals identified 235 metabolites associated with incident T2DM and developed a 44-metabolite signature that improved prediction beyond conventional risk factors[9]. Nevertheless, most studies have focused on incident diabetes in broadly normoglycaemic populations or on cross-sectional comparisons involving established T2DM. How metabolic networks evolve within individuals who already have IGR, and whether those changes differ between remission and progression, remains unclear.

Several metabolic alterations are particularly relevant to this divergence. Branched-chain amino acids (BCAAs) and aromatic amino acids are reproducibly associated insulin resistance and diabetes risk[7, 10, 11]. Mechanistic work further links altered BCAA catabolism to lipid flux; for example, the valine catabolite 3-hydroxyisobutyrate promotes trans-endothelial fatty acid transport into skeletal muscle and induces insulin resistance[12]. Acylcarnitines reflect mitochondrial substrate handling, whereas fatty acids and complex lipids participate in insulin signaling, membrane remodeling, mitochondrial stress and inflammatory activation[13-15]. Lipid remodeling may therefore integrate adipose tissue lipolysis, hepatic lipid processing, skeletal muscle substrate competition and systemic oxidative stress into a measurable circulating signature. Whether these disturbances intensify continuously with worsening glycaemia or selectively reverse during IGR remission is unknown.

Evidence that established T2DM can remit has renewed interest in metabolic reversibility. International consensus defines T2DM remission as a sustained return of glycaemia below the diabetes threshold after withdrawal glucose-lowering medication[16]. Lifestyle and weight-management trials, including DiRECT, have shown that remission is achievable, although long-term durability remains challenging[17-19]. CORDIOPREV further identified metabolites associated with the likelihood of remission following dietary intervention[20]. These findings do not establish whether IGR remission follows the same metabolic path as diabetes remission. Paired comparisons of remission and progression trajectories at the prediabetic stage remain scarce, particularly analyses that account for metabolic variation associated with ageing or follow-up time in healthy individuals. Interpretation is further complicated because circulating metabolites can arise from multiple tissues and may reflect drivers, adaptive responses or consequences of disease. Crosstalk among metabolic organs is

increasingly recognized as a core feature of T2DM pathogenesis[21, 22], making tissue-level and functional follow-up important for interpreting serum biomarkers.

Here, we analysed paired serum metabolomes from the Shanghai Nicheng cohort to examine divergent IGR outcomes over 2.5 years. We first identified IGR-associated metabolic alterations using baseline and follow-up healthy controls as discovery and validation frameworks. For the participants with IGR at baseline, we then performed paired trajectory analyses to distinguish metabolites associated with reversion to normoglycaemia from those associated with progression to T2DM while excluding time-varying metabolic features observed in healthy individuals. Key markers were further examined in mouse models of T2DM progression and exercise-induced remission, tissue association analyses and $^{13}C_6$-glucose tracing in skeletal muscle cells. This trajectory-based design identifies fatty acid remodeling as a dynamic and partially reversible metabolic axis separating metabolic recovery from deterioration during the prediabetic stage. The workflow of this study is illustrated in **Fig. 1**.

## Results

### Characteristics of study participants

To characterize metabolomic trajectories associated with IGR remission and progression, we established a longitudinal metabolomics sub-cohort within the Shanghai Nicheng cohort[23, 24]. Participants were selected according to predefined criteria including available anthropometric, glycaemic and abdominal fat measurements, no history of diabetes, and absence of major comorbidities or conditions that could substantially affect metabolic profiles. The analysis was further restricted to the individuals aged 45-55 years. In total, 975 participants with IGR at baseline were included in the metabolomics analyses. In addition, 128 healthy controls matched for

sex, age, and BMI were randomly selected from the same cohort among individuals who were healthy at baseline and remained healthy throughout follow-up (**Fig. 1**). After 2.5 years of follow-up, baseline IGR participants developed into three distinct glycaemic profiles: 202 reverted to normal glucose regulation (NGR), 554 remained in the IGR state and 219 progressed to T2DM. Clinical characteristics of the study population are summarized in **Table 1**. This cohort design enabled three complementary analyses: identification of reproducible metabolic alterations associated with IGR; characterization of baseline metabolic signatures linked to future remission, persistence or progression; and paired analysis of outcome-associated biomarkers distinguishing IGR remission from progression. The overall cohort stratification and analytical workflow are shown in **Fig. 1**. Detailed information on cohort construction and participant recruitment criteria is provided in the **Supplementary Methods**.

**Metabolic alterations associated with IGR**

To define metabolic alterations associated with IGR, we analyzed serum metabolic profiles using a discovery-validation framework. Baseline samples from HC-2013 and IGR-2013 were used as the discovery set, whereas follow-up samples from HC-2016 and IGR-2016 were used as the validation set. Orthogonal partial least-squares discriminant analysis (OPLS-DA) revealed a separation between individuals with IGR and healthy controls in both datasets, indicating that the prediabetic stage is accompanied by a detectable shift in circulating metabolic profiles (**Supplementary Fig. 1**). Nonparametric testing identified 84 and 58 differential metabolites, respectively, in the discovery and validation sets ($p < 0.05$ and FDR < 0.2). Among these, 38 metabolites were consistently altered in both datasets with concordant directions of change. These reproducible IGR-associated metabolites were mainly

assigned to amino acids, carnitines, fatty acids, lysophosphatidylcholines (LPCs) and ether-linked phosphatidylcholines (PC-Os) (**Fig. 2a**, **b and Supplementary Table 1**). Pathway enrichment analysis further showed that. phenylalanine, tyrosine and tryptophan biosynthesis, phenylalanine metabolism, alanine, aspartate and glutamate metabolism, arginine and proline metabolism, and glutathione metabolism were consistently enriched in both datasets (**Fig. 2c**, **d**). Thus, amino acid metabolism and redox-related pathways are among the most reproducibly metabolic features of the IGR state. Metabolite ratio analysis provided additional evidence for early disruption of amino acid and purine-related metabolism. We found that the biological indexes including the metabolite ratio and metabolite summation, Glu/Gln, Leu+Lys and Met/Phe increased, whereas Gly/Ser decreased in individuals with IGR (**Supplementary Table 2**), suggesting altered glutamate-glutamine cycling, branched-chain and essential amino acid metabolism, and serine-glycine or one-carbon-related metabolism. The xanthine/hypoxanthine ratio also increased in IGR, consistent with enhanced xanthine oxidoreductase/xanthine oxidase activity, oxidative stress, uric acid production, insulin resistance and future T2DM risk[25]. In parallel, lipid-related alterations were prominent. IGR was associated with higher levels of medium- and long-chain unsaturated fatty acids and oxidized fatty acid products, together with reduced LPCs and PC-Os and increased PCs. These changes are consistent with lipid remodeling mechanisms involving Lp-PLA2, sPLA2-IIA and autotaxin (ENPP2), which can promote PC hydrolysis, LPC to LPA conversion and downstream acyl remodeling[26]. Macrophage-derived Lp-PLA2 can also hydrolyse oxidized phospholipids in LDL to generate oxidized fatty acids and lysophospholipids, linking lipid oxidation to inflammation and metabolic dysregulation[27]. Altered acylcarnitines

further suggested disturbed mitochondrial fatty acid handling and incomplete oxidative substrate processing in IGR.

To assess whether reproducible metabolic alterations were independently associated with IGR status, conditional binary logistic regression was also performed for 49 consistently altered metabolites and 11 biological indexes, with adjustment for age, sex and BMI. Twenty-one metabolites and four biological indexes were selected with odds ratios > 1 (**Extended Data Fig. 1a**). Stepwise logistic regression further identified an eight-marker panel with independent discriminative value, comprising CAR 16:0, glutamic acid, L-alanine, phenylalanine, FA 21:3;O, glycohyodeoxycholic acid (GHCA), PC O-42:6 and Met/Phe. This panel showed consistent performance in the discovery and validation datasets, with areas under the receiver operating characteristic curve (AUC) of 0.782 and 0.750, respectively (**Extended Data Fig. 1b**, **c**). Together, these findings indicate that IGR is characterized by reproducible alterations in amino acid metabolism, mitochondrial substrate-handling and lipid-remodeling that can be captured by a compact circulating metabolic panel.

**Metabolic trajectories in IGR progression and remission**

We first compared baseline metabolic profiles among IGR subgroups to identify pre-existing differences associated with later progression and remission. Relative to IGR-$2013_{(IGR)}$, IGR-$2013_{(DM)}$ showed predominantly elevated amino acids, whereas the IGR-$2013_{(NGR)}$ had relatively higher levels of LPCs and sphingomyelins (SMs). However, most of these metabolite differences were not significant, indicating that direct comparisons among IGR subgroups captured only limited baseline heterogeneity (**Supplementary Table 3**). We therefore further compared each subgroup with time-matched healthy controls. This revealed distinct pre-existing metabolic states by

subsequent trajectory, with OPLS-DA showing metabolic disturbances related to IGR outcomes (**Supplementary Fig. 2**). Baseline lipid signatures differed according to subsequent glycaemic trajectories. Individuals who later reverted to normoglycaemia showed elevated free fatty acid-related metabolites at baseline, whereas those who later progressed to T2DM showed decreased LPCs and increased phosphatidylcholines (PCs) at baseline (**Extended Data Fig. 2**). Future remission is characterized by baseline free fatty acid elevation and amino acid dysregulation followed by partial metabolic normalization, whereas future progression is marked by early LPC depletion and PC elevation followed by increased acylcarnitines and LPCs and persistent disturbances in branched-chain amino acid and lipid-related pathways.

In the participants who subsequently progressed to T2DM, 97 metabolites differed between IGR-2013$_{(DM)}$ and HC-2013. After progression, total of 141 metabolites significantly changed relative to HC-2016, and 64 of them were the same as those in 2013. Additional 77 differential metabolites emerged due to T2DM with 41 increased (**Supplementary Table 4**). By contrast, participants who subsequently reverted to normoglycaemia showed broad attenuation of baseline metabolic disturbance. In the comparison of IGR-2013$_{(NGR)}$ versus HC-2013, 90 metabolites exhibited significant differences. After remission, 74 of them were no longer significantly different between the NGR-2016 and HC-2016. Among the remaining 16 metabolites, 13 decreased, while 3 increased in NGR-2016 compared with HC-2016. Additionally, 8 differential metabolites newly emerged in 2016 **(Supplementary Table 5)**. In details, during the progression from IGR to T2DM, acylcarnitines and LPCs increased relative to healthy controls and deviated further from the healthy reference profile after diabetes onset (**Fig. 3a**, **b**), indicating that glycaemic deterioration is accompanied by aggravated mitochondrial fatty acid handling and lipid remodeling. By contrast, IGR remission was

associated with a partial shift of amino acid metabolism toward the healthy profile. Aspartic acid and glutamic acid were no longer significantly different from those in healthy controls after remission, suggesting partial normalization of amino-acid metabolism. Valine showed a similar trend, whereas glycine remained decreased, indicating that reversion to normoglycaemia does not fully restore amino acid homeostasis (**Fig. 3c**). Free fatty acid-related metabolites also decreased after remission compared with the pre-remission state, suggesting that attenuation of amino acid and fatty acid abnormalities is a major metabolic feature of IGR remission (**Fig. 3d**).

Pathway enrichment analysis further revealed outcome-specific metabolic dynamics. Individuals who progressed from IGR to T2DM showed persistent and extensive pathway disturbances, including phenylalanine, tyrosine and tryptophan biosynthesis, arginine and proline metabolism, phenylalanine metabolism, alanine, aspartate and glutamate metabolism, lysine degradation and arginine biosynthesis. After diabetes onset, additional enrichment emerged in glycine, serine and threonine metabolism, glycerophospholipid metabolism and valine, leucine and isoleucine biosynthesis, highlighting intensified phospholipid remodeling and branched-chain amino-acid pathway dysregulation (**Fig. 3e**). In the remission group, altered pathways included alanine, aspartate and glutamate metabolism, glutathione metabolism and primary bile acid biosynthesis.

Notably, reversion to normoglycaemia did not equate to complete metabolic normalization. In the NGR-2016 versus HC-2016 comparison, a subset of metabolites remained significantly different after remission, mainly involving bile acids, amino acids and lysophosphatidylcholines (**Supplementary Table 5**). These residual differences were consistent with the pathway-level enrichment of primary bile acid biosynthesis observed in the remission trajectory (**Fig. 3f**), suggesting that some

metabolic abnormalities were newly emergent or persisted after glycaemic normalization.

**Metabolic biomarkers of IGR progression and remission**

We next sought to identify metabolic biomarkers specifically associated with IGR progression or remission. Multivariate analysis revealed significant metabolic shifts during progression from IGR to T2DM. However, detectable metabolic changes were also observed in healthy controls between baseline and follow-up in different years (**Supplementary Fig. 3**), indicating that longitudinal variation can occur even in individuals who remain normoglycaemic.

To distinguish outcome-associated metabolic alterations from general follow-up-related fluctuations, we implemented a time-controlled, stepwise biomarker screening strategy (**Fig. 4a**). First, paired differential analyses were performed separately for individuals who reverted from IGR to NGR and those who progressed from IGR to T2DM. Metabolites shared by the remission and progression comparisons were then extracted, and only those changing in opposite directions between the two trajectories were retained as candidates of outcome-associated biomarkers. Second, paired analysis of the healthy control cohort was used to identify metabolites that varied over time in normoglycaemic individuals. Third, these time-varying metabolites among the healthy controls were excluded to prioritize biomarkers more specifically linked to divergent IGR outcomes. Total of 284 and 359 differential metabolites during remission from IGR to NGR and during progression from IGR to T2DM were respectively identified. In parallel, 248 metabolites changed between 2013 and 2016 in healthy controls, representing potential age-, time- or life style-related metabolic variation. Among the remission and progression comparisons, 25 metabolites changed in opposite directions.

After excluding metabolites that also changed in healthy controls, 19 candidate biomarkers remained (**Supplementary Table 6**). Notably, all 19 metabolites decreased during remission from IGR to NGR but increased during progression from IGR to T2DM, without synchronous changes in normoglycaemic individuals.

Conditional binary logistic regression was then performed with adjustment for major covariates, and 10 progression-associated biomarkers and 15 remission-associated biomarkers with odds ratios > 1 and $p < 0.05$ were selected, nine of the metabolites were shared between the remission and progression models (**Fig. 4b**, **c and Supplementary Table 7**). In addition, 13 of the 19 candidate biomarkers showed ordered changes across increasing glycaemic severity from HC-2016 to NGR-2016, IGR-2016 and DM-2016. Most of these graded metabolites were shared biomarkers of both remission and progression, including octadecanedioate, hexadecanedioate, tetradecanedioic acid and 9-HODE. The dicarboxylic acids octadecanedioate, hexadecanedioate and tetradecanedioic acid point to altered fatty acid ω-oxidation and downstream oxidative handling, whereas 9-HODE, an oxidized linoleic acid metabolite, links lipid peroxidation with HDL dysfunction and inflammatory lipid signalling[28-30]. The ordered changes of these markers across glycaemic states suggest that they were related to progressive metabolic remodeling across the dysglycaemia continuum rather than end-stage diabetes status alone (**Fig. 4d–g**). Other candidate biomarkers also showed graded changes across the four glycaemic states, although the magnitude and direction varied for different metabolites (**Supplementary Fig. 4**).

**Changes of metabolic biomarkers in the mouse models**

Sixteen candidate biomarkers which were significantly associated with either IGR progression or remission (**Fig. 4b**, **c**) were selected for experimental follow-up. Ten of

them could be reliably detected in mouse blood, therefore, a targeted method was established for their quantification, the analytical characteristics of the method are given in **Supplementary Table 8** and **Supplementary Fig. 5**. The concentrations of these markers were then determined in a high-fat diet- and streptozotocin-induced model of IGR progression and an exercise-induced model of IGR remission (**Fig. 5a, b**).

In IGR progression model, mice were subjected to high-fat diet feeding and streptozotocin administration. After 10 weeks of modeling, these mice showed significantly increased body weight and blood glucose levels compared with controls, supporting successful establishment of the diabetic phenotype (**Fig. 5a, c, d**). Targeted serum metabolomics showed seven candidate biomarkers that were concordantly altered between the mouse progression model and the human IGR to diabetes trajectory: FA 12:1, FA 14:1, FA 18:0, FA 18:1+O, levoglucosan, tetradecanedioic acid and 9-HODE. All of them were significantly increased in T2DM mice compared with controls. For the exercising intervening, diabetic mice were assigned to sedentary or a treadmill-training group for 4 weeks with a normal diet during the intervention (**Fig. 5b**). Exercise training significantly reduced body weight and fasting blood glucose compared with sedentary T2DM mice, indicating improved glucose metabolism (**Fig. 5e-h**). Targeted serum metabolomics further showed that four remission-related markers changed in the same direction as observed in the human remission trajectory. These results support the relevance of the exercise-induced model for evaluating biomarkers associated with metabolic recovery. Finally, venn analysis identified three markers validated across the human cohort, the IGR progression model and the exercise-induced remission model: FA 12:1, levoglucosan and tetradecanedioic acid (**Fig. 5i**). These markers increased during glycaemic deterioration and decreased during exercise-induced metabolic

improvement, supporting their bidirectional association with IGR progression and remission. The results of targeted metabolite analysis of pancreas, skeletal muscle, liver and brain showing that glucose levels increased in pancreas, skeletal muscle and brain, whereas hepatic free glucose decreased in IGR progression model (**Extended Data Fig. 3a-h**). This decrease may reflect altered hepatic glucose handling under insulin-resistant conditions, in which impaired insulin action reduces glycogen synthesis and fails to suppress gluconeogenesis and glycogenolysis, thereby promoting hepatic glucose output and systemic hyperglycaemia while lowering the intrahepatic free glucose pool. Overall, candidate markers were most consistently detected and supported in liver tissue. FA 12:1, tetradecanedioic acid and 9-HODE showed concordant hepatic changes across both the progression and remission models (**Extended Data Fig. 3c, g**). Correlation analysis between circulating and tissue metabolites further revealed strong associations with liver for FA 12:1, 9-HODE and tetradecanedioic acid in the progression model. In the remission model, FA 12:1 and 9-HODE remained strongly correlated with liver, whereas tetradecanedioic acid did not reach statistical significance. FA 12:1 and tetradecanedioic acid also showed relatively strong correlations with pancreas in the remission model. These findings suggest that circulating lipid markers exhibit tissue-associated patterns, with FA 12:1, 9-HODE and tetradecanedioic acid closely linked to hepatic metabolic disturbance (**Extended Data Fig. 4**).

**Multicenter validation of dicarboxylic acid alterations in T2DM**

To assess the reproducibility of our findings in independent external cohorts, we analyzed two publicly available plasma metabolomics datasets. From Wu *et al* [31], we extracted standardized metabolite abundances and adjusted *p* values of the Swedish

IGT cohort (n = 697: normal glucose tolerance [NGT] 220, isolated impaired fasting glucose [IFG] 185, isolated impaired glucose tolerance [IGT] 173, combined glucose intolerance [CGI] 74 and newly diagnosed, untreated T2DM 45) and the SCAPIS cohort (n = 470: NGT 201, IGT 130, CGI 84 and T2DM 55). Both untargeted cohorts and our study used the WHO classification framework, with IFG, IGT and CGI representing distinct subtypes within our IGR definition. These subgroups were therefore retained to assess the consistency of metabolic changes. The other targeted MetaboLights dataset MTBLS8644 contained 80 plasma samples including healthy, early prediabetes, late prediabetes, and T2DM groups, with 20 participants in each group [32,33]. Their late prediabetes group met the fasting glucose criterion of the IGR in our study. Accordingly, the healthy, late prediabetes and T2DM groups were included in the targeted validation (**Fig. 6a**). Data for tetradecanedioic acid and related metabolites were extracted for subsequent analyses (**Supplementary Table 9**).

In the two untargeted cohorts, tetradecanedioic acid showed higher standardized abundances in all prediabetes subgroups and the T2DM group relative to NGT **(Fig. 6b).** Analysis of raw abundance data from the targeted dataset confirmed higher mean standardized abundances in late prediabetes and T2DM than that in healthy participants and a significant overall difference across the three groups (**Fig. 6c and Supplementary Fig. 6**). To examine whether tetradecanedioic acid elevation was accompanied by consistent changes in related metabolites, we compared their abundances between T2DM and the corresponding normal control group in each of the three cohorts, it was found that tetradecanedioic acid elevation was accompanied by increases in related metabolites involved in fatty acid ω-oxidation and subsequent dicarboxylic acid β-oxidation (**Fig. 6d**). Together, these multicenter and cross-platform

results independently support the reproducible disruption of dicarboxylic acid metabolism and particularly the elevation of tetradecanedioic acid observed in our study.

**Functional study of human-mouse concordant lipid markers**

We next performed targeted quantification and isotope-tracing experiments to examine whether representative outcome-associated lipid biomarkers were related to cellular glucose utilization. Tetradecanedioic acid was firstly selected for quantitative and functional characterization because it was detected consistently in both human and animal analyses and showed a graded increase with worsening glycaemic status. The serum quantification result showed that its concentration was approximately three-fold higher in T2DM model mice than in controls, reaching approximately 1.0 μM (**Fig. 7c**). 9-HODE was also included in the functional assay because it is an oxidized linoleic acid metabolite related to oxidative lipid stress and inflammatory lipid signaling. Previous studies have reported impaired antioxidant and anti-inflammatory properties of HDL in patients with T2DM, with oxidized fatty acids such as 9-HODE implicated in HDL dysfunction[28]. Skeletal muscle serves as the primary site for whole-body glucose disposal and glycogen storage, playing an indispensable role in maintaining glucose homeostasis. Thus, the potential effects of theses lipid-related biomarkers on skeletal muscle glucose metabolism were assessed. The C2C12 myoblasts were differentiated into myotubes and treated with tetradecanedioic acid or 9-HODE, and then the treated myotubes were subjected to $^{13}C_6$-glucose labelling, followed by LC–MS/MS-based isotopologue analysis of TCA-cycle intermediates (**Fig. 7d**). Compared with the DMSO control, tetradecanedioic acid significantly reduced the M+2 isotopologue fractions of citrate and α-ketoglutarate, indicating decreased incorporation of glucose-derived carbon into the TCA cycle (**Fig. 7e**–**g**). 9-HODE showed a similar

direction of effect, although the statistical support was weaker. These findings suggest that tetradecanedioic acid can impair glucose-derived TCA-cycle carbon flux in skeletal-muscle cells, potentially linking circulating fatty acid remodeling to reduced glucose oxidative metabolism.

## Discussion

By integrating paired longitudinal metabolomics in a community-based IGR cohort with time-controlled biomarker screening, mouse models, tissue profiling and isotope-tracing, we identified metabolic features that distinguish reversion from IGR to NGR and progression to T2DM. Changes in amino acids, acylcarnitines, fatty acids, lysophospholipids and ether-linked phospholipids indicate coordinated disruption of amino-acid turnover, mitochondrial substrate handling and lipid remodeling. This pattern is consistent with previous metabolomics studies of early dysglycaemia[34, 35]. Higher BCAAs and lower glycine also accord with their reported associations with insulin resistance and future T2DM[10]. Altered amino-acid turnover and disposal may contribute to these circulating patterns [11], and BCAA-derived metabolites can impair β-cell glucose-stimulated insulin secretion under diabetic conditions[36]. Elevated fatty acid flux and adipose tissue insulin resistance increase the delivery of non-esterified fatty acids and glycerol to the liver, promoting hepatic glucose production and worsening hyperglycaemia[37], while ectopic lipid accumulation and lipid-derived signals may further impair insulin action[38]. In insulin-resistant states, increased fatty acid flux and BCAA-related substrate pressure may exceed mitochondrial oxidative capacity. Excess acyl-CoA intermediates can be converted into acylcarnitines through the carnitine palmitoyltransferase system and exported into the circulation, leading to increased circulating acylcarnitines[39]. The acylcarnitine elevations observed in our IGR

population are therefore compatible with lipid oversupply, incomplete fatty acid oxidation and impaired mitochondrial substrate handling at the prediabetic stage.

IGR remission and progression were metabolically asymmetric. Individuals progressing to T2DM retained amino-acid disturbances and showed stronger acylcarnitine accumulation and glycerophospholipid remodeling. Those reverting to NGR showed partial normalization of amino-acid metabolism and attenuation of free fatty acid-related abnormalities. This asymmetry is biologically important because it suggests that remission was not simply the reverse of progression. Rather, remission appears to reflect selective relief of specific metabolic pressures, whereas progression reflects expansion of the disturbance into mitochondrial fatty acid oxidation and membrane lipid metabolism. This interpretation is consistent with remission-focused intervention studies. In the DiRECT trial, weight-loss-induced diabetes remission was accompanied by broad metabolomic changes, many of them moved in the opposite direction to previously reported T2DM-associated perturbations[40]. However, reversibility was incomplete for all metabolites, in line with our finding that glycine and bile acid-related pathways did not fully return to the healthy reference profile after IGR remission. Thus, metabolic remission may involve partial repair of selected pathways while leaving residual biochemical abnormalities that could influence relapse risk.

The progression trajectory was particularly marked by acylcarnitine accumulation, suggesting unresolved mitochondrial fatty acid oxidation stress. Beyond the transport function, acylcarnitines are increasingly regarded as systemic indicators of mitochondrial fuel handling because they may accumulate when fatty acid influx exceeds downstream oxidative capacity or when β-oxidation is incompletely coupled to tricarboxylic acid cycle flux and respiratory demand[41]. In insulin-resistant states,

nutrient and lipid excess can impose mitochondrial substrate pressure, reduce metabolic flexibility and impair insulin action in skeletal muscle, the major site of insulin-stimulated glucose disposal[42, 43]. In human T2DM, reduced β-oxidative capacity together with fatty acylcarnitine accumulation has been linked to Th17-skewed inflammatory activation, suggesting a potential immunometabolic connection between incomplete lipid oxidation and chronic metabolic inflammation[44]. Consistent with these mechanistic observations, prospective metabolomics evidence supports the relevance of acylcarnitines to diabetes development: an updated meta-analysis of prospective cohorts identified several acylcarnitines, including C4-DC, C4-OH, C5, C5-OH and C8:1, as positively associated with incident T2DM risk[6], and a recent large-scale analysis highlighted carnitine-related pathways among metabolite classes associated with future T2DM and precision risk prediction[9]. In our study, the stronger acylcarnitine signal in individuals progressing to T2DM suggests that failure to resolve mitochondrial substrate overload may be a candidate metabolic feature separating progressive dysglycaemia from reversible IGR.

IGR remission should not be interpreted as biological cure. Consensus recommendations emphasize that remission requires continued surveillance, and long-term follow-up of DiRECT illustrates that remission durability is incomplete even under structured intervention[16, 45] In this study, several dysglycaemia-associated abnormalities were attenuated after reversion to normoglycaemia, residual differences in bile acids, amino acids and lysophosphatidylcholines remained detectable, suggesting incomplete pathway recovery after glucose normalization. This interpretation is biologically plausible, as bile acid synthesis and phospholipid metabolism have been linked to incident T2DM in prospective metabolomics and genetic analyses. In line with this concept, the Look AHEAD study showed that T2DM

remission was associated with substantially lower risks of chronic kidney disease and cardiovascular disease, but not with complete elimination of downstream cardiometabolic risk[46]. Together, these findings extend the remission concept to the prediabetic stage and suggest that IGR remission reflects metabolic improvement rather than complete metabolic normalization.

By excluding metabolites that changed over time in normoglycaemic individuals, we identified candidate markers that moved in opposite directions during IGR remission and progression. Many of these markers were dicarboxylic acids, oxidized fatty acid products or unsaturated fatty acid species, suggesting that fatty acid oxidation pressure, lipid peroxidation and compensatory lipid detoxification routes are dynamically linked to the direction of IGR outcomes. FA 12:1, levoglucosan and tetradecanedioic acid were validated across human serum, the progression model and the remission model, and decreased during metabolic improvement. This consistent validation across human and mouse studies reduces the likelihood that these biomarkers are cohort-specific correlates and suggests that they reflect metabolic processes related to both glycaemic deterioration and recovery. Circulating lipid-related markers that change in opposite directions during deterioration and recovery may therefore provide a biochemical readout of remission depth and relapse risk.

Several circulating markers, particularly FA 12:1, 9-HODE and tetradecanedioic acid, showed liver-associated patterns in mouse tissues. This is consistent with central role of the liver in coordinating glucose production with lipid partitioning, including de novo lipogenesis, fatty acid oxidation and VLDL-triglyceride export[47]. In T2DM, hepatic insulin resistance promotes inappropriate glucose output while altering lipid flux and lipoprotein metabolism, thereby linking hyperglycaemia to circulating lipid remodeling. The liver-associated correlations observed in our animal models therefore

support a hepatic contribution to the serum lipid signature. At the same time, the metabolic effects of fatty acid-related markers may extend beyond the liver. In skeletal muscle cells, tetradecanedioic acid reduced glucose-derived carbon entry into the TCA cycle, suggesting impaired glucose oxidation which provides a plausible functional link between circulating lipid remodeling and impaired whole-body glucose utilizaion. Tetradecanedioic acid and 9-HODE provide mechanistic clues to the nature of this lipid remodeling axis. Tetradecanedioic acid is a dicarboxylic acid generated through ω-oxidation and further processed through β-oxidation. Long-chain dicarboxylic acid metabolism is closely linked to fatty acid detoxification, peroxisomal and mitochondrial oxidation, and metabolic adaptation under lipid overload[48]. Its increase during progression and decrease during remission suggest that enhanced dicarboxylic acid production may mark a compensatory but insufficient response to excess fatty acid burden, which was further supported by multicenter validation in both the untargeted and targeted results, and the levels showed a progressive increase with worsening glycaemic status. HDL from patients with T2DM has been reported to exhibit impaired antioxidant and anti-inflammatory properties, and oxidized fatty acids such as 9-HODE have been implicated in HDL dysfunction[28]. Therefore, the combination of dicarboxylic acids, oxidized fatty acids and unsaturated fatty acid species suggests that progressive IGR is characterized by lipid oxidation pressure, altered fatty acid detoxification routes and inflammatory lipid remodeling. These pathways may not merely accompany dysglycaemia but also contribute to impaired substrate selection and reduced glucose oxidative metabolism in metabolically active tissues.

Several limitations should be acknowledged. First, the human cohort analysis is observational and cannot by itself distinguish causal mediators from correlates of metabolic deterioration. Especially the cohort had a narrow age distribution. Second,

circulating metabolites integrate influences from tissue metabolism, adiposity, diet, lipoprotein transport, renal handling and the microbiome. Third, the high-fat diet- and streptozotocin-induced model captures combined insulin resistance and β-cell injury but does not fully mimic the tempo or heterogeneity of human IGR trajectories; it is best viewed as a useful model of progression-associated metabolic stress rather than a full surrogate for human remission and progression biology. Future studies in independent cohorts with repeated sampling, quantitative targeted assays, dietary and medication annotation, and longer follow-up will be required to determine whether these lipid remodeling trajectories can improve clinical stratification beyond conventional risk factors.

# Methods

### Participants and study design

This study used a longitudinal metabolomics design within the Shanghai Nicheng community cohort. From eligible baseline participants, we selected 1,103 individuals for metabolomic profiling, including 128 healthy controls and 975 participants with impaired glucose regulation (IGR). After 2.5 years of follow-up, baseline IGR participants were classified into three trajectories: reversion to normal glucose regulation (n = 202), persistent IGR (n = 554) and progression to type 2 diabetes mellitus (n = 219). Baseline and follow-up serum samples were used to identify reproducible IGR-associated metabolic alterations, baseline signatures of future glycaemic outcomes, and paired longitudinal metabolic changes associated with IGR remission or progression. Paired healthy controls served as a time-matched reference for background metabolic variation. The study design and analytical workflow are summarized in **Fig. 1.**

**Reagents and sample preparation**

Plasma samples were stored at −80 °C at the Sixth Affiliated Hospital of Shanghai Jiao Tong University (Shanghai, China). Metabolomic analyses were centrally conducted at the Metabolomics Research Center, Dalian Institute of Chemical Physics, Chinese Academy of Sciences (Dalian, China). All laboratory personnel performing the analyses were blinded to clinical information. LC–MS-grade acetonitrile and methanol were obtained from Merck (Darmstadt, Germany), and ultrapure water was produced using a Milli-Q purification system (Merck, Darmstadt, Germany). Formic acid, ammonium bicarbonate, and ammonium hydroxide ($NH_4OH$) were purchased from J&K Scientific (Beijing, China), InnoChem (Beijing, China), and R&K, respectively. Authentic chemical standards of metabolites were supplied by Alta Scientific Co., Ltd. (Tianjin, China).

For metabolite extraction, both serum and cell samples were processed using a unified protocol with slight modifications. For serum samples, 100 μL of human plasma was mixed with 400 μL of methanol containing isotope-labeled internal standards (**Supplementary Table 10**). For cell samples, after 48 hours of culture, cells were washed with PBS, quenched with liquid nitrogen, and extracted with 80% methanol containing internal standards. The cell suspensions were scraped into Eppendorf tubes. All samples were vortexed for 1 min and centrifuged at 14,000 rpm for 15 min at 4 °C. The resulting supernatants were transferred into 1.5 mL Eppendorf tubes and lyophilized. Before LC–MS/MS analysis, the dried extracts were reconstituted in 20% acetonitrile, vortexed for 1 min, and centrifuged again under the same conditions. The final supernatants were transferred to HPLC vials for subsequent LC–MS/MS analysis.

**Animal models and in vivo experiments**

The T2DM model and the diabetes remission model were established by Cyagen Biosciences Inc. Mice in the control group were fed a standard chow diet for 10 weeks. For T2DM model induction, mice were subjected to a 10-week modeling protocol. Beginning at week 0, mice were continuously fed a 45% high fat diet. At week 8, streptozotocin (STZ) was administered. Fasting blood glucose (FBG) was measured 3 days after STZ injection at week 9 (FBG1) and again at week 10 (FBG2). Mice with both FBG measurements exceeding 11.1 mmol/L were considered successfully established as T2DM models. During model induction, an oral glucose tolerance test (OGTT) was performed at week 8, and fasting blood glucose was monitored weekly from the stage of model evaluation onward. Feeding regimens and measurement items were further adjusted as needed according to the experimental design. Mice meeting the diagnostic criterion for diabetes at both time points were defined as successfully modeled. The diabetes remission model was established through treadmill exercise intervention in T2DM mice. After successful induction of T2DM by high fat diet feeding combined with STZ injection, model mice were randomly assigned to an exercise training group (ET) and a non-exercise control group (NET), designated as the ET and NET groups, respectively. During the intervention period, mice were maintained on a normal diet for 4 weeks. At the end of the intervention, FBG1 was measured, and the experiment was terminated once a significant reduction in fasting blood glucose was observed relative to the pre intervention level.

**Cell culture and metabolic flux analysis**

The C2C12 cell line was purchased from Servicebio Technology Co., Ltd. (Wuhan, China). Cells were maintained in RPMI 1640 medium (Gibco) supplemented with 10%

fetal bovine serum (FBS; Gibco), 100 IU/mL penicillin, and 100 μg/mL streptomycin (Gibco). Cultures were incubated at 37 °C in a humidified atmosphere containing 5% $CO_2$. Myogenic differentiation of C2C12 myoblasts into myotubes was performed following the manufacturer's instructions for the C2C12 Differentiation and Detection Kit (Servicebio). On the first day of induction, cells were treated with 1 μg/mL tetradecanedioic acid and 9-Hode, with treatment maintained throughout the differentiation process. Following 48 h of incubation, cells from the treatment and control groups were harvested. The cells were washed with phosphate-buffered saline (PBS) and immediately quenched in liquid nitrogen. Subsequently, 1 mL of 80% methanol containing isotopic internal standards (**Supplementary Table 10**) was added to each sample. The cells were collected using a cell scraper into 5 mL Eppendorf (EP) tubes, vortexed for 20 s, and centrifuged at 14,000 rpm for 15 min at 4 °C. The resulting supernatant was collected, lyophilized, and reconstituted for subsequent chromatographic/mass spectrometric analysis.

**Untargeted LC–MS-based metabolomic profiling**

Untargeted metabolomic analysis was performed using a Q Exactive mass spectrometer (Thermo Fisher Scientific, Rockford, USA). Chromatographic separation was carried out on BEH C8 column (2.1 × 50 mm, 1.7 μm; Waters Corp., Milford, USA) and HSS T3 column (2.1 × 50 mm, 1.7 μm) in positive (ESI+) and negative ($ESI^-$) ionization modes, respectively. The column temperature was maintained at 60 °C, and the flow rate was set to 0.4 mL/min. For ESI+ and $ESI^-$ modes, mobile phase A consisted of water containing 0.1% (v/v) formic acid and 6.5 mM ammonium bicarbonate, while mobile phase B was acetonitrile with 0.1% (v/v) formic acid and 95% (v/v) methanol containing 6.5 mM ammonium bicarbonate.

In the $ESI^+$ mode, chromatographic separation began with 5% B for 0–0.5 min, gradually increased to 40% at 2 min, then ramped to 100% by 6 min, held for 2 min, and returned to the initial composition within 0.1 min, followed by a 2 min re-equilibration. The $ESI^-$ gradient started at 2% B for 0–0.5 min, increased to 40% at 2 min, reached 100% at 6 min, was held for 2 min, then returned to 2% B within 0.1 min, and re-equilibrated for 2 min. Each run lasted approximately 12 min in total. The ion source parameters were set as follows: spray voltage ±3.5 kV, capillary temperature 300 °C, sheath gas pressure 45 psi, auxiliary gas pressure 10 psi, and probe heater temperature 350 °C. Full-scan MS and MS/MS spectra were acquired across m/z 70–1050 and 50–800, respectively, with stepped normalized collision energies of 15, 30, and 45 eV.

**Targeted LC–MS/MS-based biomarkers validation**

Targeted LC–MS/MS analysis was performed on an ExionLC™ AD ultra-high-performance liquid chromatography system (AB Sciex, Framingham, USA) coupled to a Triple Quadrupole 6500 Plus mass spectrometer (AB Sciex, Framingham, USA). Multiple reaction monitoring (MRM) transitions were established based ten biomarkers associated with T2DM progression or remission. The precursor ion (Q1), product ion (Q3), declustering potential (DP), and collision energy (CE) were subsequently selected and optimized for each compound. The complete list of MRM transitions and corresponding parameters is provided in **Supplementary Table 8**. The LC gradient program was consistent with that used in the untargeted metabolomics analysis described above.

**Data preprocessing and statistical analyses**

The untargeted HRMS data were processed using MS-DIAL 4 for peak detection, alignment, and peak area integration. Metabolite annotation was performed by matching MS1 and MS/MS spectra against multiple databases, including HMDB (https://www.hmdb.ca/), MassBank of North America (https://mona.fiehnlab.ucdavis.edu/), and NIST2023 (http://www.nist.gov/srd/nist1a.cfm). In addition, an in-house database OSI-SMMS[49] was employed to achieve more accurate metabolite identification. The detailed annotation criteria are described in our previously published study[50]. Total of 545 metabolites were annotated with RSD < 30% in this study (**Supplementary Fig. 7**). For targeted metabolomics, data processing was performed using SCIEX OS (AB SCIEX, USA). Finally, all quantified peak areas were normalized to the corresponding internal standards prior to statistical analysis.

Differential analysis was performed using Multi Experiment Viewer (MeV; open-source genomic analysis software developed by the MeV team). Orthogonal partial least squares discriminant analysis (OPLS-DA) was conducted with SIMCA-P 14.1 software (Umetrics, Umea, Sweden) to obtain sample score plots and metabolite loading profiles. Additional analyses, including conditional binary logic and other visualization procedures, were implemented through in-house coding scripts.

For external validation, standardized metabolite abundances and adjusted p values of the untargeted Swedish IGT and SCAPIS cohorts were extracted directly from the published supplementary tables, retaining the original significance threshold of adjusted $p < 0.1$. In contrast, the publicly available metabolite abundance data from the targeted MTBLS8644 dataset were independently processed and analyzed using nonparametric tests and a more stringent significance threshold of $p < 0.05$. Shared differential metabolites were identified using these cohort-specific thresholds.

**Reporting summary**

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

## Data availability

The data supporting the findings from this study are available within the manuscript and its supplementary information. The raw LC−MS data files generated in this study have been deposited in the Metabolomics Subcenter of the National Genomics Data Center at the Dalian Institute of Chemical Physics, Chinese Academy of Sciences (Project No. MP000007; https://metabolomics.dicp.ac.cn/web/meta-project/detail/12493) and are publicly available. Source data are provided with this paper. The raw individual participant data are protected and are not available due to data privacy laws.

## Code availability

This study relies mainly on open-source bioinformatic tools and the analysis-specific programs are available on GitHub at https://github.com/jfy16/Metabolic-Trajectories-and-Clinical-Outcomes-in-Prediabetes.git.

## Author contributions

P.G., R.Y., P.D. and F.J. processed the data, performed statistical analyses, interpreted the results, and drafted and revised the manuscript. D.Y., X.W. and Q.L. acquired and processed metabolomic data. H.L., Y.N., H.M., Y.P. and X.H. recruited participants and collected clinical data and biospecimens. X.L. and G.X. conceived and designed the

study, coordinated research resources, supervised the metabolomic analyses and revised the manuscript. W.J. conceived and designed the study, provided clinical data and biospecimens, supervised the clinical research and critically revised the manuscript. All authors reviewed and approved the final manuscript. P.G., R.Y., P.D. and F.J. contributed equally to this work.

## Funding

This work was supported by the National Natural Science Foundation of China (22434006, 22522411, 22474138), Liaoning Province Excellent Youth Fund Program (2024JH3/10200015), and the Innovation Program (DICP I202334) of Science and Research from the DICP, CAS.

## Competing interests

The authors declare no competing interests.

**References**


1. Owens, D. R.*, et al.* IDF diabetes Atlas: A worldwide review of studies utilizing retinal photography to screen for diabetic retinopathy from 2017 to 2024 inclusive. *Diabetes research and clinical practice* **226**, 112346 https://doi.org/https://doi.org/10.1016/j.diabres.2025.112346 (2025).
2. Global, regional, and national burden of diabetes from 1990 to 2021, with projections of prevalence to 2050: a systematic analysis for the Global Burden of Disease Study 2021. *Lancet* **402**, 203–234 https://doi.org/10.1016/s0140-6736(23)01301-6 (2023).
3. Saeedi, P.*, et al.* Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9(th) edition. *Diabetes Res Clin Pract* **157**, 107843 https://doi.org/10.1016/j.diabres.2019.107843 (2019).
4. Chew, N. W. S.*, et al.* The global burden of metabolic disease: Data from 2000 to 2019. *Cell Metab* **35**, 414–428.e413 https://doi.org/10.1016/j.cmet.2023.02.003 (2023).
5. Unnikrishnan, R.*, et al.* Prediabetes. *Nature reviews Disease primers* **11**, 49 https://doi.org/10.1038/s41572-025-00635-0 (2025).
6. Morze, J.*, et al.* Metabolomics and Type 2 Diabetes Risk: An Updated Systematic Review and Meta-analysis of Prospective Cohort Studies. *Diabetes Care* **45**, 1013–1024 https://doi.org/10.2337/dc21-1705 (2022).
7. Roberts, L. D., Koulman, A. & Griffin, J. L. Towards metabolic biomarkers of insulin resistance and type 2 diabetes: progress from the metabolome. *The lancet Diabetes & endocrinology* **2**, 65–75 https://doi.org/10.1016/s2213-8587(13)70143-8 (2014).

8. Ahola-Olli, A. V.*, et al.* Circulating metabolites and the risk of type 2 diabetes: a prospective study of 11,896 young adults from four Finnish cohorts. *Diabetologia* **62**, 2298–2309 https://doi.org/10.1007/s00125-019-05001-w (2019).
9. Li, J.*, et al.* Circulating metabolites, genetics and lifestyle factors in relation to future risk of type 2 diabetes. *Nat Med* **32**, 660–670 https://doi.org/10.1038/s41591-025-04105-8 (2026).
10. Bacha, F.*, et al.* Distinct Amino Acid Profile Characterizes Youth With or at Risk for Type 2 Diabetes. *Diabetes* **73**, 628–636 https://doi.org/10.2337/db23-0375 (2024).
11. White, P. J.*, et al.* Insulin action, type 2 diabetes, and branched-chain amino acids: A two-way street. *Molecular metabolism* **52**, 101261 https://doi.org/10.1016/j.molmet.2021.101261 (2021).
12. Jang, C.*, et al.* A branched-chain amino acid metabolite drives vascular fatty acid transport and causes insulin resistance. *Nature medicine* **22**, 421–426 https://doi.org/10.1038/nm.4057 (2016).
13. Meikle, P. J. & Summers, S. A. Sphingolipids and phospholipids in insulin resistance and related metabolic disorders. *Nature reviews Endocrinology* **13**, 79–91 https://doi.org/10.1038/nrendo.2016.169 (2017).
14. Chaurasia, B. & Summers, S. A. Ceramides in Metabolism: Key Lipotoxic Players. *Annual review of physiology* **83**, 303–330 https://doi.org/10.1146/annurev-physiol-031620-093815 (2021).
15. Wittenbecher, C.*, et al.* Dihydroceramide- and ceramide-profiling provides insights into human cardiometabolic disease etiology. *Nature communications* **13**, 936 https://doi.org/10.1038/s41467-022-28496-1 (2022).
16. Riddle, M. C.*, et al.* Consensus Report: Definition and Interpretation of Remission

in Type 2 Diabetes. *Diabetes Care* **44**, 2438–2444 https://doi.org/10.2337/dci21-0034 (2021).

17. Magkos, F., Hjorth, M. F. & Astrup, A. Diet and exercise in the prevention and treatment of type 2 diabetes mellitus. *Nature reviews Endocrinology* **16**, 545–555 https://doi.org/10.1038/s41574-020-0381-5 (2020).
18. Lean, M. E.*, et al.* Primary care-led weight management for remission of type 2 diabetes (DiRECT): an open-label, cluster-randomised trial. *Lancet (London, England)* **391**, 541–551 https://doi.org/10.1016/s0140-6736(17)33102-1 (2018).
19. Lean, M. E. J.*, et al.* Durability of a primary care-led weight-management intervention for remission of type 2 diabetes: 2-year results of the DiRECT open-label, cluster-randomised trial. *The lancet Diabetes & endocrinology* **7**, 344–355 https://doi.org/10.1016/s2213-8587(19)30068-3 (2019).
20. Mora-Ortiz, M.*, et al.* Metabolomics analysis of type 2 diabetes remission identifies 12 metabolites with predictive capacity: a CORDIOPREV clinical trial study. *BMC medicine* **20**, 373 https://doi.org/10.1186/s12916-022-02566-z (2022).
21. Roden, M. & Shulman, G. I. The integrative biology of type 2 diabetes. *Nature* **576**, 51–60 https://doi.org/10.1038/s41586-019-1797-8 (2019).
22. Xourafa, G., Korbmacher, M. & Roden, M. Inter-organ crosstalk during development and progression of type 2 diabetes mellitus. *Nature reviews Endocrinology* **20**, 27–49 https://doi.org/10.1038/s41574-023-00898-1 (2024).
23. Hou, X.*, et al.* Abdominal Subcutaneous Fat: A Favorable or Nonfunctional Fat Depot for Glucose Metabolism in Chinese Adults? *Obesity (Silver Spring, Md)* **26**, 1078–1087 https://doi.org/10.1002/oby.22183 (2018).
24. Chen, P.*, et al.* Abdominal subcutaneous adipose tissue: a favorable adipose depot for diabetes? *Cardiovascular diabetology* **17**, 93 https://doi.org/10.1186/s12933-

018-0734-8 (2018).

25. Li, X.*, et al.* Elevated Serum Xanthine Oxidase Activity Is Associated With the Development of Type 2 Diabetes: A Prospective Cohort Study. *Diabetes Care* **41**, 884–890 https://doi.org/10.2337/dc17-1434 (2018).
26. Nelson, T. L.*, et al.* Lipoprotein-associated phospholipase A2 (Lp-PLA2) and future risk of type 2 diabetes: results from the Cardiovascular Health Study. *J Clin Endocrinol Metab* **97**, 1695–1701 https://doi.org/10.1210/jc.2011-3026 (2012).
27. Wu, Z., Wu, S., Liang, T. & Wang, L. Lipoprotein-Associated Phospholipase A2 Is a Risk Factor for Patients With Parkinson's Disease. *Frontiers in neuroscience* **15**, 633022 https://doi.org/10.3389/fnins.2021.633022 (2021).
28. Morgantini, C.*, et al.* Anti-inflammatory and antioxidant properties of HDLs are impaired in type 2 diabetes. *Diabetes* **60**, 2617–2623 https://doi.org/10.2337/db11-0378 (2011).
29. Feng, J.*, et al.* High levels of oxidized fatty acids in HDL impair the antioxidant function of HDL in patients with diabetes. *Frontiers in endocrinology* **13**, 993193 https://doi.org/10.3389/fendo.2022.993193 (2022).
30. Kwon, S. Y.*, et al.* Oxidised metabolites of the omega-6 fatty acid linoleic acid activate dFOXO. *Life science alliance* **3**, https://doi.org/10.26508/lsa.201900356 (2020).
31. Wu, H.*, et al.* Microbiome-metabolome dynamics associated with impaired glucose control and responses to lifestyle changes. *Nat Med* **31**, 2222–2231 https://doi.org/10.1038/s41591-025-03642-6 (2025).
32. Yurekten, O.*, et al.* MetaboLights: open data repository for metabolomics. *Nucleic Acids Res* **52**, D640–d646 https://doi.org/10.1093/nar/gkad1045 (2024).
33. Ji, W.*, et al.* Metabolomic approaches to dissect dysregulated metabolism in the

progression of pre-diabetes to T2DM. *Mol Omics* **20**, 333–347 https://doi.org/10.1039/d3mo00130j (2024).

34. Shi, L.*, et al.* Plasma metabolites associated with type 2 diabetes in a Swedish population: a case-control study nested in a prospective cohort. *Diabetologia* **61**, 849–861 https://doi.org/10.1007/s00125-017-4521-y (2018).
35. Jung, Y., Ahn, E., Park, T. & Hwang, G. S. Longitudinal study investigating serum metabolites and their association with type 2 diabetes risk in a Korean population. *Diabetes Obes Metab* **25**, 2120–2130 https://doi.org/10.1111/dom.15084 (2023).
36. Lin, H.*, et al.* Branched-chain α-keto acids impair glucose-stimulated insulin secretion in pancreatic β-cells under diabetes by reactivating the LDHA-lactate axis. *Nature communications* **17**, https://doi.org/10.1038/s41467-026-70004-2 (2026).
37. Sekizkardes, H.*, et al.* Free fatty acid processing diverges in human pathologic insulin resistance conditions. *J Clin Invest* **130**, 3592–3602 https://doi.org/10.1172/jci135431 (2020).
38. Turpin-Nolan, S. M. & Brüning, J. C. The role of ceramides in metabolic disorders: when size and localization matters. *Nat Rev Endocrinol* **16**, 224–233 https://doi.org/10.1038/s41574-020-0320-5 (2020).
39. Bene, J., Hadzsiev, K. & Melegh, B. Role of carnitine and its derivatives in the development and management of type 2 diabetes. *Nutr Diabetes* **8**, 8 https://doi.org/10.1038/s41387-018-0017-1 (2018).
40. Corbin, L. J.*, et al.* The metabolomic signature of weight loss and remission in the Diabetes Remission Clinical Trial (DiRECT). *Diabetologia* **67**, 74–87 https://doi.org/10.1007/s00125-023-06019-x (2024).
41. Fritzen, A. M., Lundsgaard, A. M. & Kiens, B. Tuning fatty acid oxidation in

skeletal muscle with dietary fat and exercise. *Nat Rev Endocrinol* **16**, 683–696 https://doi.org/10.1038/s41574-020-0405-1 (2020).

42. Petersen, M. C. & Shulman, G. I. Mechanisms of Insulin Action and Insulin Resistance. *Physiological reviews* **98**, 2133–2223 https://doi.org/10.1152/physrev.00063.2017 (2018).
43. Samuel, V. T. & Shulman, G. I. The pathogenesis of insulin resistance: integrating signaling pathways and substrate flux. *The Journal of clinical investigation* **126**, 12–22 https://doi.org/10.1172/jci77812 (2016).
44. Nicholas, D. A.*, et al.* Fatty Acid Metabolites Combine with Reduced β Oxidation to Activate Th17 Inflammation in Human Type 2 Diabetes. *Cell Metab* **30**, 447–461.e445 https://doi.org/10.1016/j.cmet.2019.07.004 (2019).
45. Lean, M. E.*, et al.* 5-year follow-up of the randomised Diabetes Remission Clinical Trial (DiRECT) of continued support for weight loss maintenance in the UK: an extension study. *The lancet Diabetes & endocrinology* **12**, 233–246 https://doi.org/10.1016/s2213-8587(23)00385-6 (2024).
46. Gregg, E. W.*, et al.* Impact of remission from type 2 diabetes on long-term health outcomes: findings from the Look AHEAD study. *Diabetologia* **67**, 459–469 https://doi.org/10.1007/s00125-023-06048-6 (2024).
47. Roumans, K. H. M.*, et al.* Hepatic saturated fatty acid fraction is associated with de novo lipogenesis and hepatic insulin resistance. *Nature communications* **11**, 1891 https://doi.org/10.1038/s41467-020-15684-0 (2020).
48. Ranea-Robles, P. & Houten, S. M. The biochemistry and physiology of long-chain dicarboxylic acid metabolism. *The Biochemical journal* **480**, 607–627 https://doi.org/10.1042/bcj20230041 (2023).
49. Zhao, X.*, et al.* Comprehensive Strategy to Construct In-House Database for

Accurate and Batch Identification of Small Molecular Metabolites. *Analytical Chemistry* **90**, 7635–7643 https://doi.org/10.1021/acs.analchem.8b01482 (2018).

50. Guan, P.*, et al.* Novel Method for Simultaneously Untargeted Metabolome and Targeted Exposome Analysis in One Injection. *Anal Chem* **97**, 3996–4004 https://doi.org/10.1021/acs.analchem.4c05565 (2025).

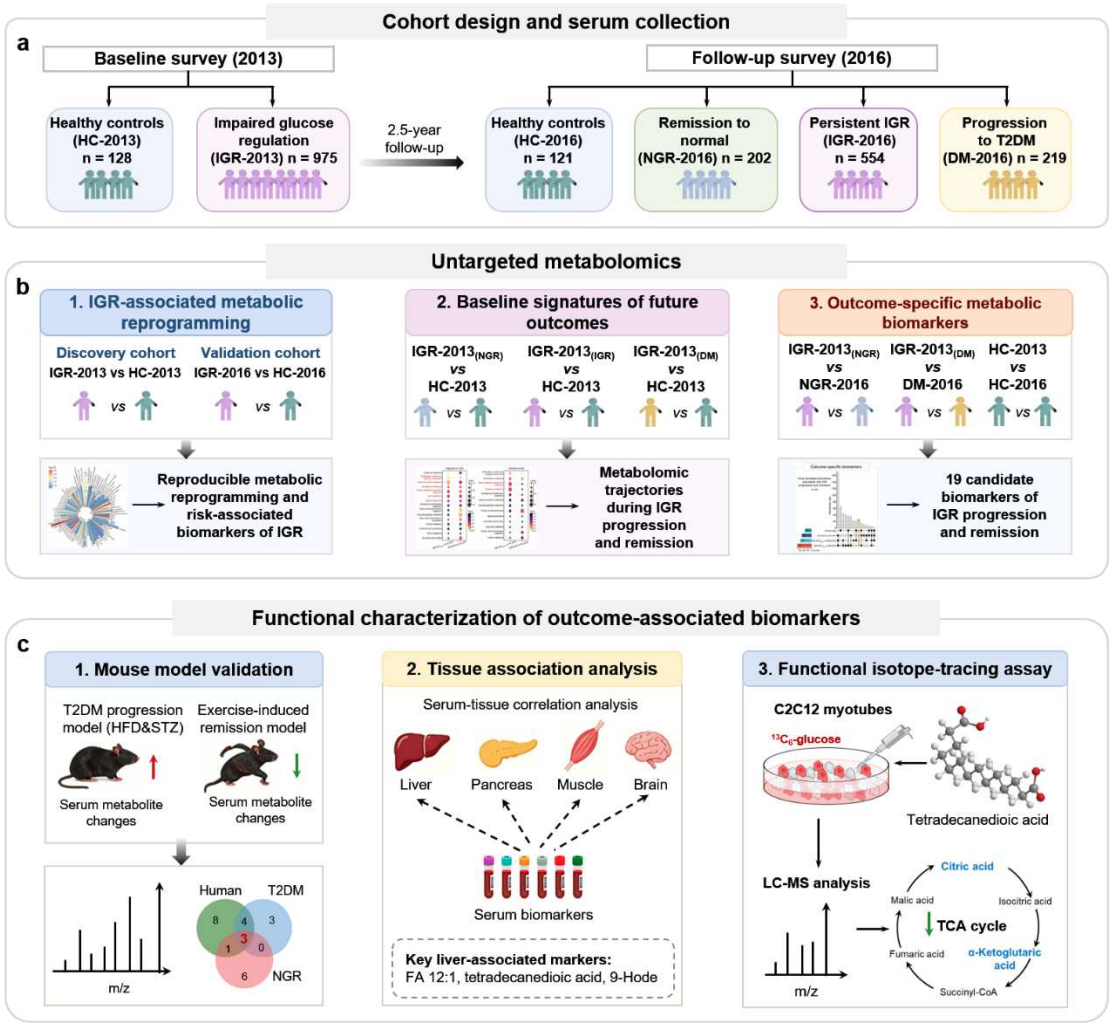



**Fig. 1 Study design and analytical workflow. a,** Construction of the longitudinal metabolomics sub-cohort from the Shanghai Nicheng cohort. A total of 1,103 participants were included, comprising 975 individuals with impaired glucose regulation (IGR) and 128 healthy controls (HC) at baseline. After 2.5 years of follow-up, individuals with IGR were classified into three outcome groups: reversion to normal glucose regulation (NGR), persistent IGR, or progression to type 2 diabetes mellitus (T2DM). Time-matched HC samples were used as a reference for longitudinal comparisons. **b,** Untargeted serum metabolomics workflow for defining reproducible IGR-associated metabolic reprogramming, baseline metabolic signatures linked to future IGR outcomes, and outcome-specific metabolic trajectories during remission and progression. These analyses were used to identify candidate biomarkers associated with divergent IGR outcomes. **c,** Downstream analytical and validation framework, including targeted metabolomics validation in IGR progression and exercise-induced remission mouse models, serum-tissue association analysis, and functional characterization using $^{13}C_6$-glucose isotope-tracing assays.

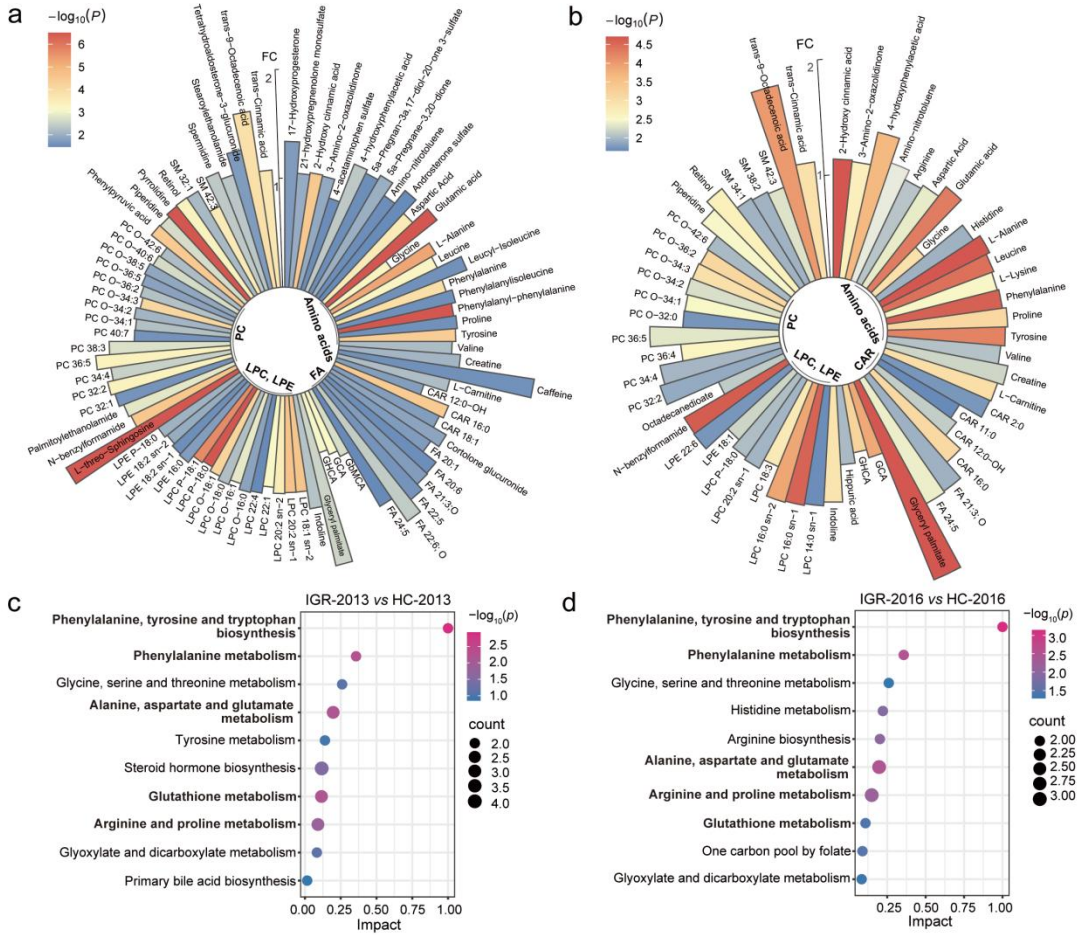



**Fig. 2 Serum metabolomics reveals reproducible metabolic reprogramming. a and b,** Circular bar plots showing differential metabolites between individuals with IGR and healthy controls in the discovery set (IGR-2013 versus HC-2013) and validation set (IGR-2016 versus HC-2016), respectively. Bar length indicates fold change, and bar color indicates -log$_{10}$($p$). Metabolites are grouped by chemical class. **c and d**, Pathway enrichment analysis of differential metabolites identified in the discovery and validation sets. Dot size represents the number of matched metabolites, and dot color indicates -log$_{10}$($p$).

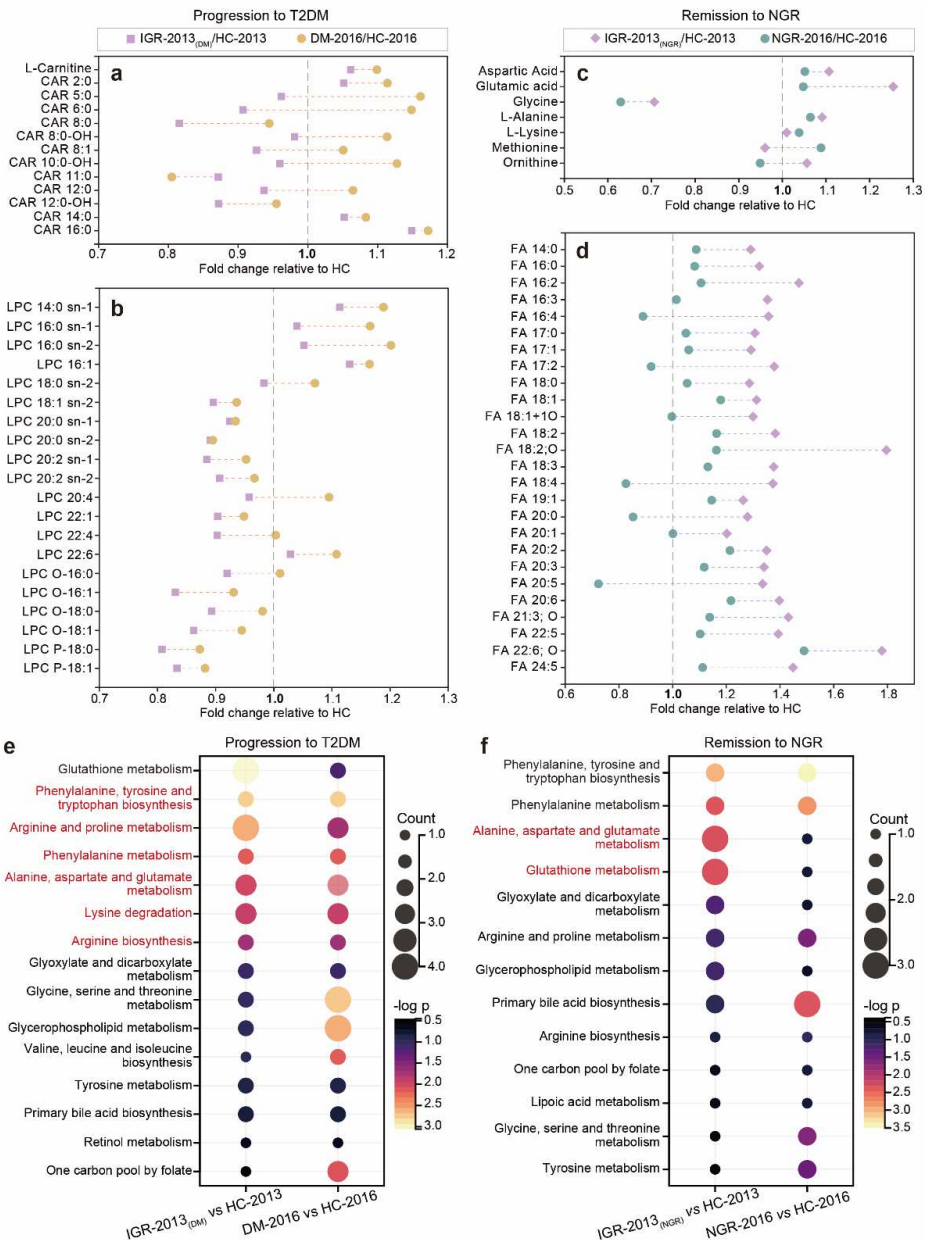



**Fig. 3 Divergent metabolomic trajectories during IGR progression and remission.** **a and b**, Longitudinal changes in acylcarnitines and lysophosphatidylcholines during progression from IGR to T2DM, shown as fold changes relative to time-matched healthy controls. **c and d**, Longitudinal changes in amino acids and free fatty acid-related metabolites during remission from IGR to NGR, shown as fold changes relative to time-matched healthy controls. Connected points indicate the same metabolites before and after follow-up, and the dashed vertical line indicates the time-matched

healthy reference level. **e**, Pathway-level trajectories during progression to T2DM, showing persistent amino-acid dysregulation together with intensified glycerophospholipid remodeling and branched-chain amino-acid pathway disturbance after diabetes onset. **f**, Pathway-level trajectories during IGR remission, highlighting partial normalization of alanine, aspartate and glutamate metabolism and glutathione metabolism, with persistent alteration of primary bile acid biosynthesis. Dot size represents the number of matched metabolites, and dot color indicates $-\log_{10}(p)$.

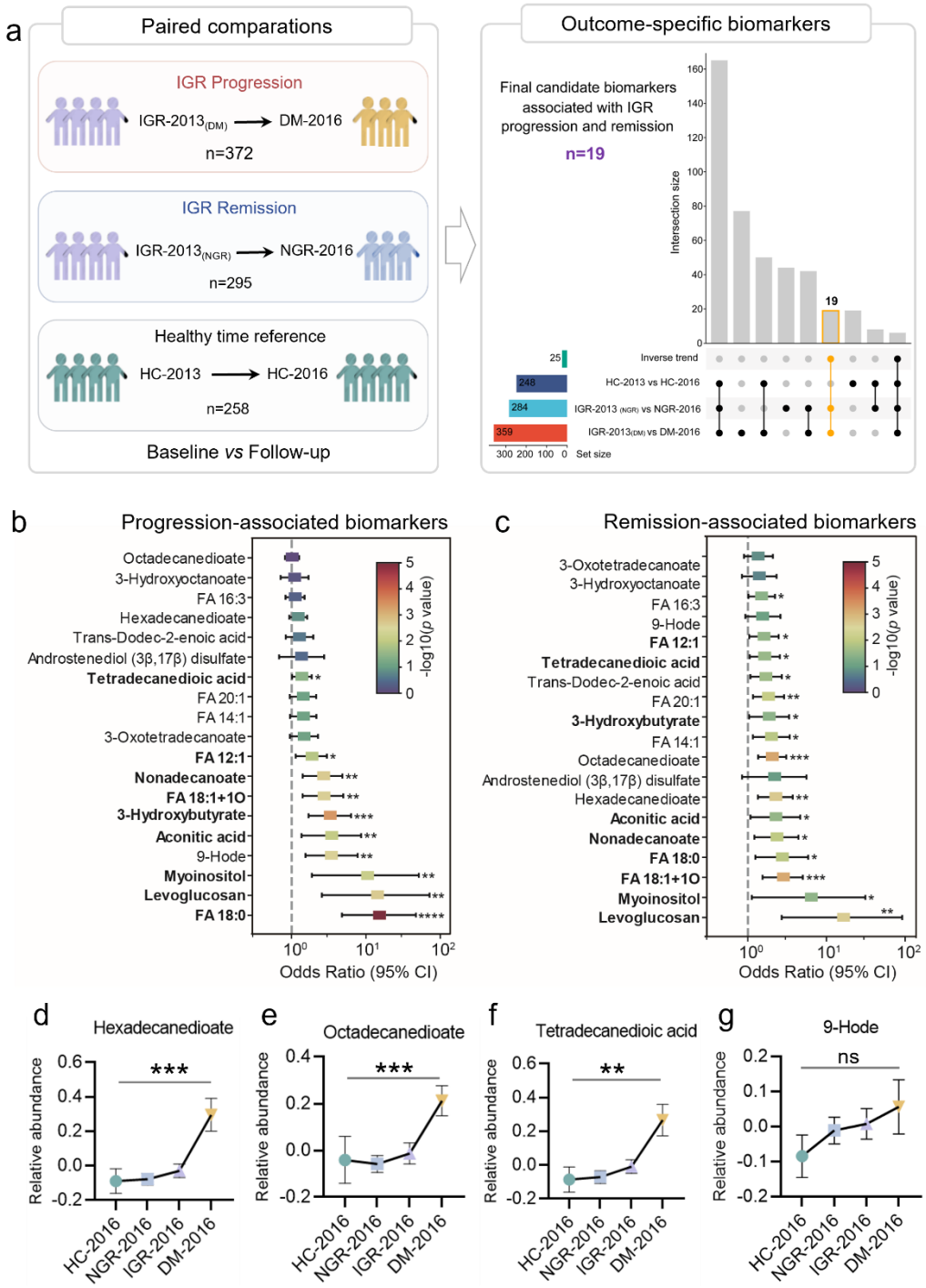



**Fig. 4 Time-controlled identification of outcome-associated biomarkers linked to IGR progression and remission. a**, Screening strategy for identifying biomarkers associated with divergent IGR outcomes. Paired comparisons were performed in individuals who progressed from IGR to T2DM, individuals who reverted from IGR to NGR, and healthy controls during follow-up. Metabolites showing opposite directions of change between progression and remission were retained, whereas time-varying

metabolic features observed in healthy controls were excluded. The UpSet plot summarizes the intersections among the paired comparisons and identifies 19 candidate biomarkers associated with IGR progression and remission. **b and c**, Odds ratios and 95% confidence intervals for candidate biomarkers associated with progression to T2DM and remission to NGR, respectively, estimated using conditional logistic regression adjusted for age, sex and BMI. In the progression model, odds ratios greater than 1 indicate higher odds of progression to T2DM. In the remission model, odds ratios greater than 1 indicate higher odds of reversion to NGR. Point color indicates $-\log_{10}(p)$, horizontal bars indicate 95% confidence intervals, and the dashed vertical line indicates an odds ratio of 1. Bold labels indicate biomarkers shared by the progression and remission models. **d-g**, Representative candidate biomarkers showing graded changes across glycaemic states from HC-2016 to NGR-2016, IGR-2016 and DM-2016. Data are shown as mean±SEM. Statistical significance was assessed using nonparametric tests; $*p < 0.05$, $**p < 0.01$, $***p < 0.001$; ns, not significant.

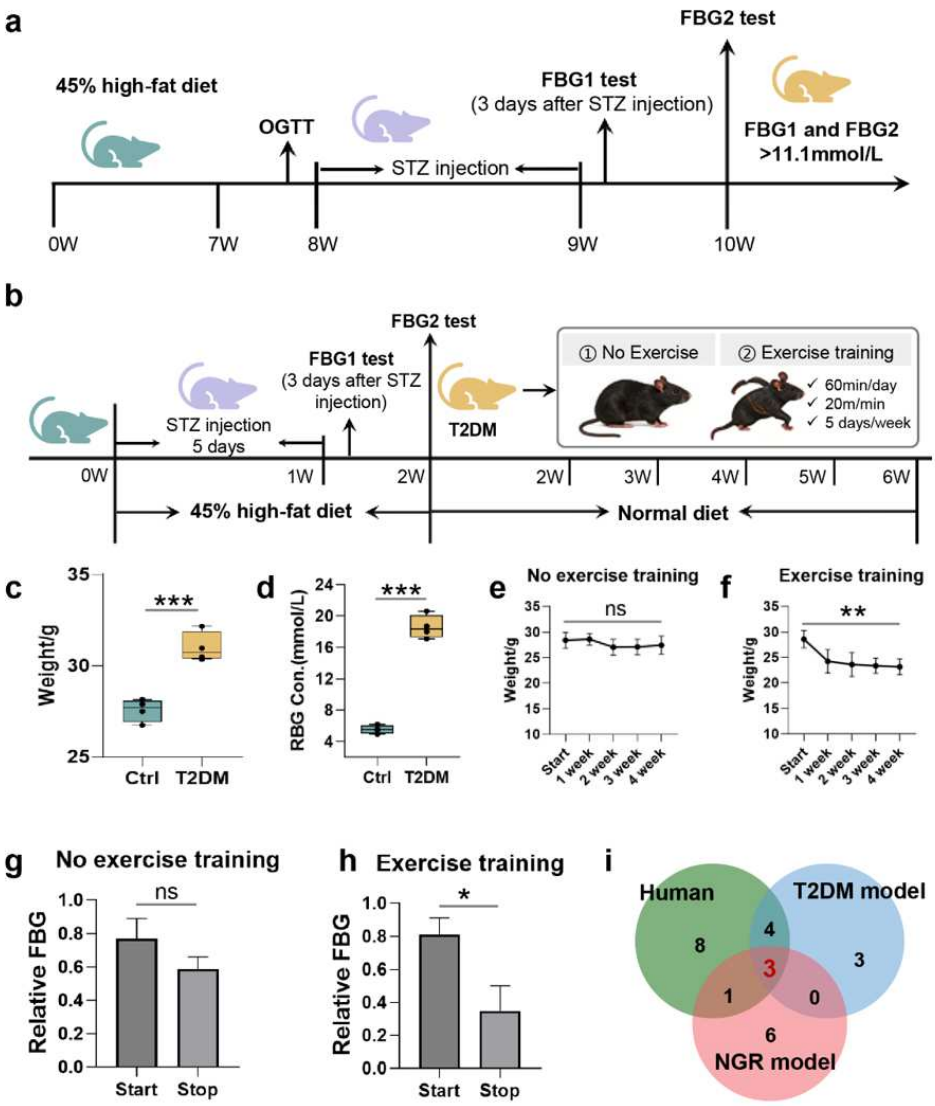



**Fig. 5 Cross-species validation of outcome-associated biomarkers in T2DM progression and remission mouse models. a**, Experimental design of the high-fat diet- and streptozotocin-induced T2DM progression mouse model. Mice were fed a 45% high-fat diet, subjected to oral glucose tolerance testing, and then administered streptozotocin. Fasting blood glucose was measured at two time points after streptozotocin administration, and mice with both FBG1 and FBG2 values greater than 11.1 mM were considered successfully established T2DM models. b, Experimental design of the exercise-induced remission model. After T2DM induction, mice were assigned to sedentary T2DM control or treadmill exercise groups and maintained on a normal diet for 4 weeks. **c-d**, Body weight and fasting blood glucose levels in control

and T2DM progression model mice. **e-h**, Changes in body weight and fasting blood glucose in sedentary and exercise-trained T2DM mice during the intervention period. i, Venn diagram identifying biomarkers consistently validated in human serum, the T2DM progression model and the exercise-induced remission model. Statistical significance was assessed using nonparametric tests; *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$; ns, not significant.

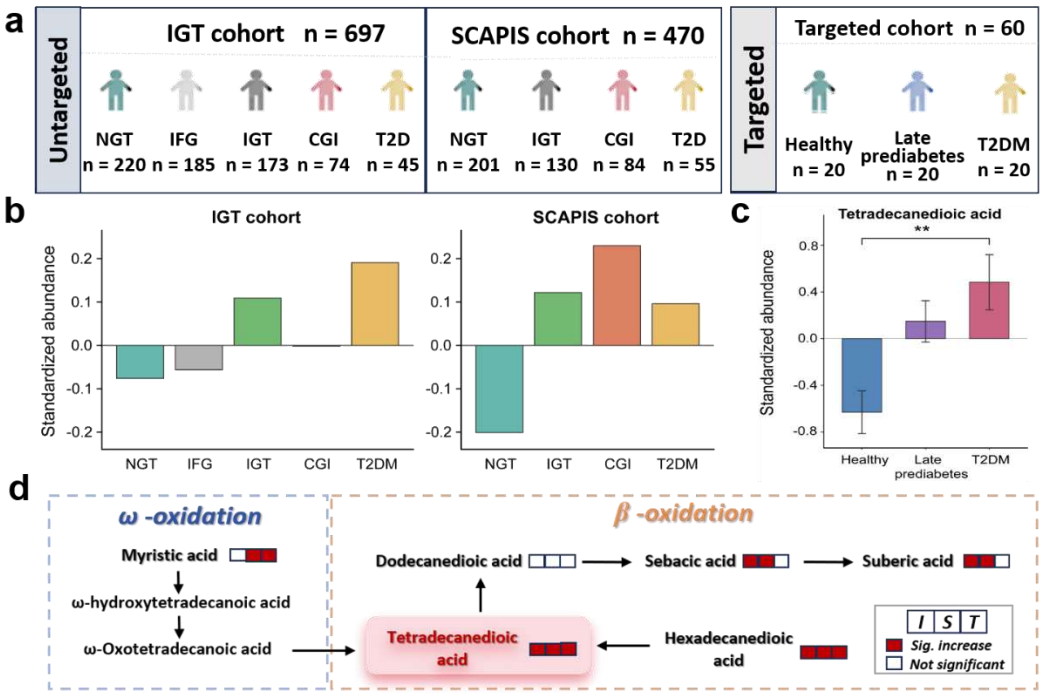



**Fig. 6 Multicenter validation reveals the progressive accumulation of dicarboxylic acids across stages of dysglycaemia. a,** Overview of the untargeted cohorts, and targeted cohort used for external validation of tetradecanedioic acid metabolic alterations. **b,** Standardized tetradecanedioic acid abundances across glycaemic subgroups in the two untargeted cohorts. **c,** Targeted validation of tetradecanedioic acid elevation in the healthy, late prediabetes and T2DM groups. **d,** Schematic summary of alterations in fatty acid ω-oxidation and β-oxidation in T2DM compared with the corresponding control groups. For each metabolite, the three squares represent the IGT (I), SCAPIS (S) and targeted cohorts (T), from left to right. Red color indicates significantly higher metabolite abundances in T2DM, whereas white color indicates no significant difference. Standardized abundances represent z-scores of $\log_{10}$-transformed metabolite abundances. Statistical significance was defined as adjusted $p < 0.1$ for untargeted data (consistent with the original literature) and as the more stringent $p < 0.05$ for targeted data, given its smaller metabolite panel and relative quantification. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$; ****$p < 0.0001$, ns, not significant.

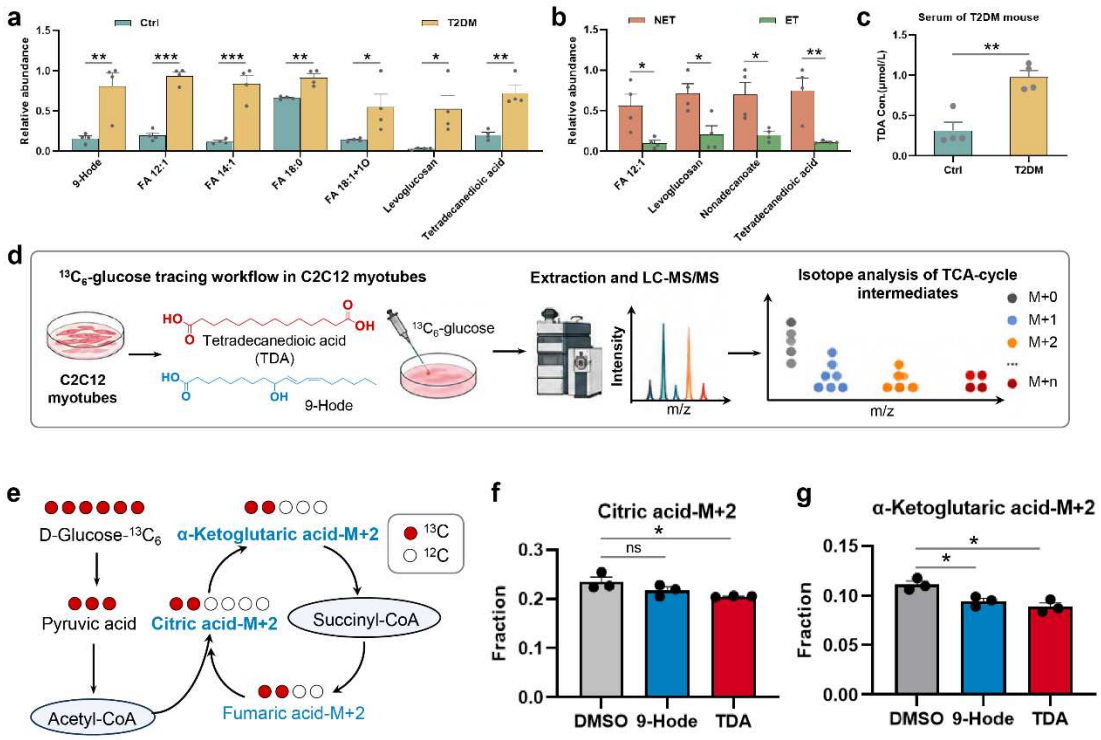
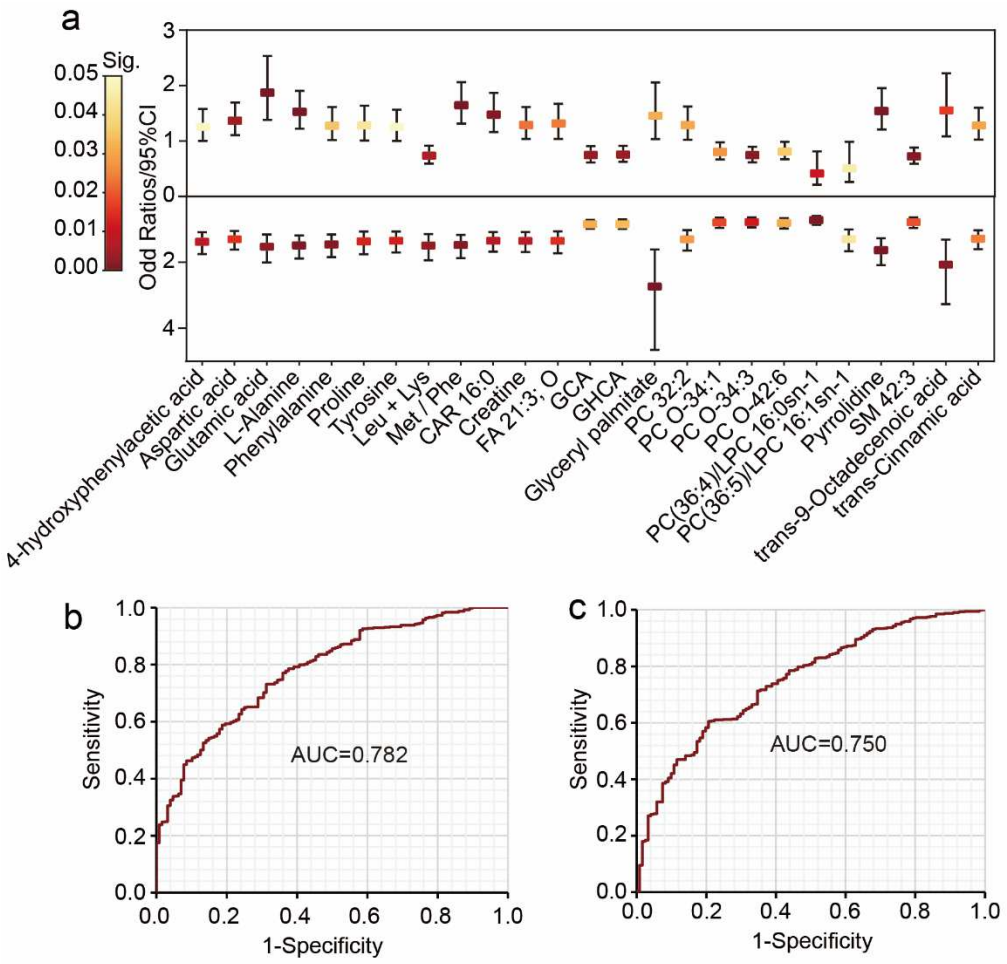



**Fig. 7 Functional isotope-tracing analysis links outcome-associated lipid biomarkers to impaired skeletal-muscle glucose oxidation**. **a**, Relative abundance of the seven cross-species validated biomarkers in control and T2DM progression model mice, measured by targeted LC–MRM analysis. **b**, Relative abundance of shared validated biomarkers during exercise-associated metabolic improvement. **c**, Targeted quantification of serum tetradecanedioic acid (TDA) in control and T2DM model mice. **d**, Experimental workflow for functional isotope-tracing analysis. Differentiated C2C12 myotubes were treated with 9-HODE or tetradecanedioic acid, followed by $^{13}C_6$-glucose labelling and LC–MS/MS-based isotopologue analysis of glucose-derived TCA-cycle intermediates. **e**, Schematic representation of glucose-derived carbon entry into the TCA cycle, highlighting citrate and α-ketoglutarate labelling. **f-g**, M+2 isotopologue fractions of citrate and α-ketoglutarate in C2C12 myotubes after treatment with DMSO, 9-HODE or tetradecanedioic acid. Data are shown as mean±SEM. Statistical significance was assessed using nonparametric tests; *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$; ns, not significant.

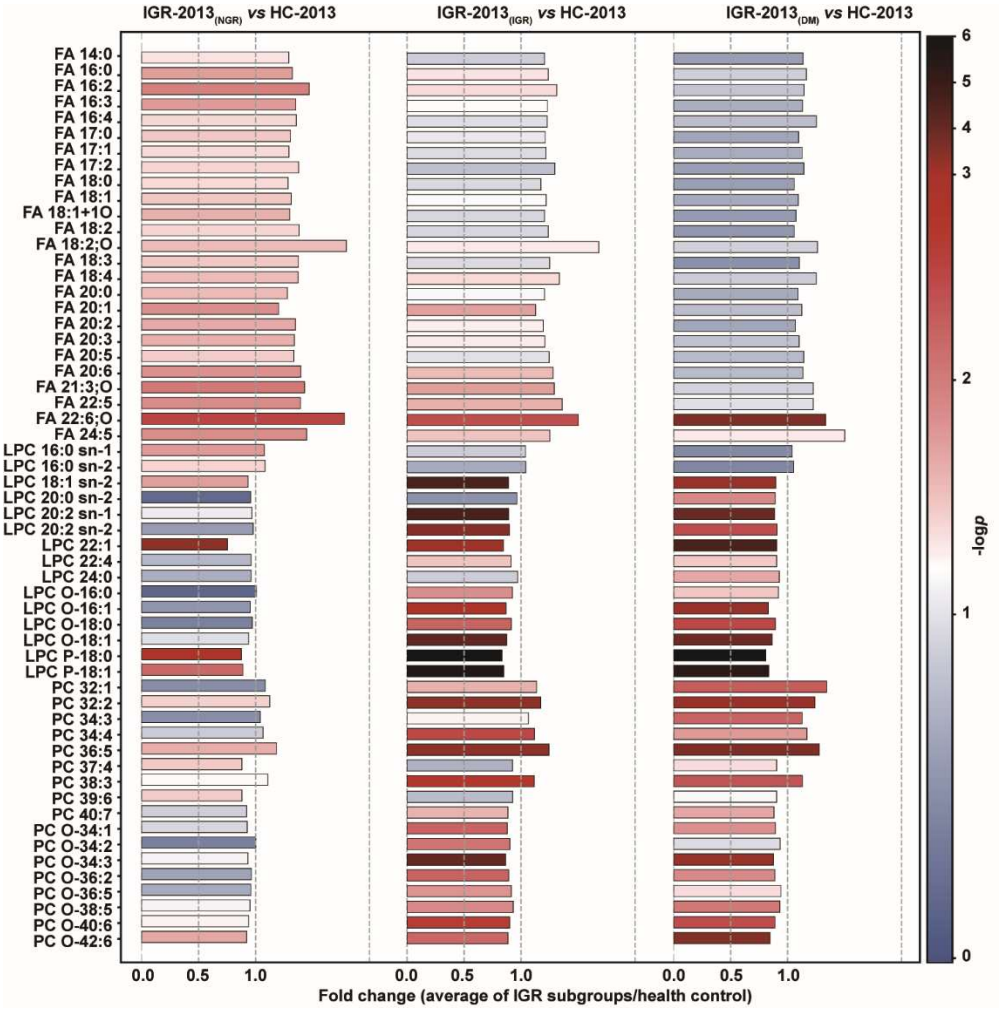
**Table 1 Basic information of the cohort in this study**

| Variables | Total | Healthy control | IGR-2013(NGR) | IGR-2013(IGR) | IGR-2013(DM) | p value |
|---|---|---|---|---|---|---|
| Age, year | 50.8 (48.5-53.0) | 51.0 (48.0-53.1) | 50.5 (48.4-52.8) | 50.7 (48.5-53.0) | 51.0 (48.8-53.1) | 0.373 |
| Female, % | 755 (68.4) | 91 (71.1) | 130 (64.4) | 382 (69.0) | 152 (69.4) | 0.544 |
| BMI, kg/m2 | 25.0 (22.9-27.2) | 23.2 (21.5-25.5) | 24.7 (23.0-26.6) | 25.1 (23.0-27.4) | 25.8 (23.8-27.8) | <0.001 |
| SBP, mmHg | 129.0 (120.0-139.0) | 123.7 (120.0-130.0) | 127.0 (120.0-136.0) | 129.0 (121.0-139.0) | 131.0 (122.0-142.0) | <0.001 |
| DBP, mmHg | 82.0 (79.0-88.0) | 80.0 (78.0-84.0) | 82.0 (79.0-87.0) | 82.0 (79.0-89.0) | 83.5 (80.0-90.0) | <0.001 |
| FPG, mmol/L | 5.9 (5.4-6.3) | 5.3 (5.2-5.5) | 5.7 (5.4-6.2) | 6.0 (5.6-6.3) | 6.2 (5.8-6.5) | <0.001 |
| 2hPG, mmol/L | 8.2 (7.0-9.1) | 5.9 (5.1-6.6) | 8.2 (7.3-8.8) | 8.2 (7.7-9.1) | 9.0 (7.9-10.0) | <0.001 |
| HbA1c, % | 5.5 (5.3-5.8) | 5.2 (5.1-5.4) | 5.5 (5.2-5.7) | 5.6 (5.4-5.8) | 5.8 (5.5-6.0) | <0.001 |
| FINS, mU/mL | 7.5 (5.5-10.6) | 5.4 (4.1-6.7) | 7.4 (5.4-10.2) | 7.9 (6.0-10.9) | 8.3 (6.2-12.0) | <0.001 |
| TC, mmol/L | 5.0 (4.5-5.6) | 4.7 (4.3-5.4) | 4.8 (4.4-5.6) | 5.0 (4.5-5.7) | 5.2 (4.6-5.8) | <0.001 |
| LDLC, mmol/L | 3.1 (2.6-3.6) | 2.8 (2.4-3.3) | 3.0 (2.5-3.5) | 3.1 (2.7-3.6) | 3.2 (2.7-3.7) | <0.001 |
| HDLC, mmol/L | 1.3 (1.1-1.5) | 1.4 (1.2-1.6) | 1.2 (1.1-1.5) | 1.3 (1.1-1.5) | 1.2 (1.0-1.4) | <0.001 |
| TG, mmol/L | 1.4 (1.0-2.1) | 1.2 (0.8-1.6) | 1.3 (0.9-1.9) | 1.5 (1.0-2.2) | 1.5 (1.0-2.4) | <0.001 |
| Level of education, % | | | | | | 0.214 |
| Primary school or less | 288 (26.1) | 38 (29.7) | 46 (22.8) | 144 (26.0) | 60 (27.4) | |

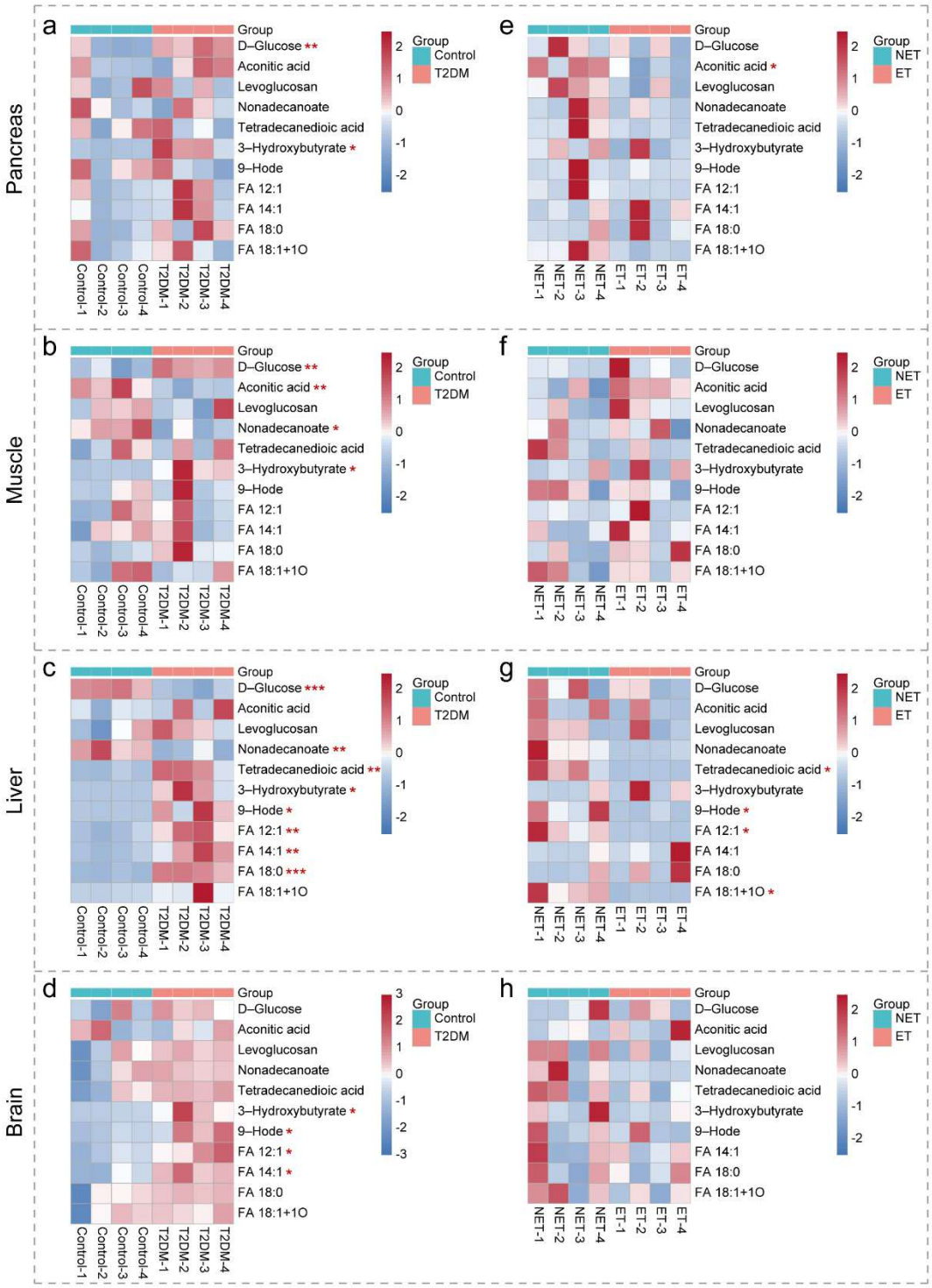
| | | | | | | |
|---|---|---|---|---|---|---|
| Middle school | 691 (62.6) | 78 (60.9) | 140 (69.3) | 336 (60.6) | 137 (62.6) | |
| High school or more | 124 (11.2) | 12 (9.4) | 16 (7.9) | 74 (13.4) | 22 (10.0) | |
| Smoking status, % | | | | | | 0.253 |
| Never or ex-smoker | 934 (84.7) | 108 (84.4) | 162 (80.2) | 477 (86.1) | 187 (85.4) | |
| Current smoker | 169 (15.3) | 20 (15.6) | 40 (19.8) | 77 (13.9) | 32 (14.6) | |
| Drinking status, % | | | | | | 0.66 |
| Never or ex-drinker | 1030 (93.4) | 125 (97.7) | 183 (90.6) | 518 (93.5) | 204 (93.2) | |
| Current drinker | 73 (6.6) | 3 (2.3) | 19 (9.4) | 36 (6.5) | 15 (6.8) | |
| Leisure-time physical activity, % | | | | | | 0.887 |
| 0 min/day | 1059 (96.0) | 125 (97.7) | 190 (94.1) | 536 (96.8) | 208 (95.0) | |
| 1–29 min/day | 19 (1.7) | 2 (1.6) | 3 (1.5) | 10 (1.8) | 4 (1.8) | |
| ≥ 30 min/day | 25 (2.3) | 1 (0.8) | 9 (4.5) | 8 (1.4) | 7 (3.2) | |
| Family history of diabetes, % | | | | | | 0.017 |
| No | 925 (83.9) | 116 (90.6) | 175 (86.6) | 462 (83.4) | 172 (78.5) | |
| Yes | 177 (16.0) | 12 (9.4) | 27 (13.4) | 91 (16.4) | 47 (21.5) | |



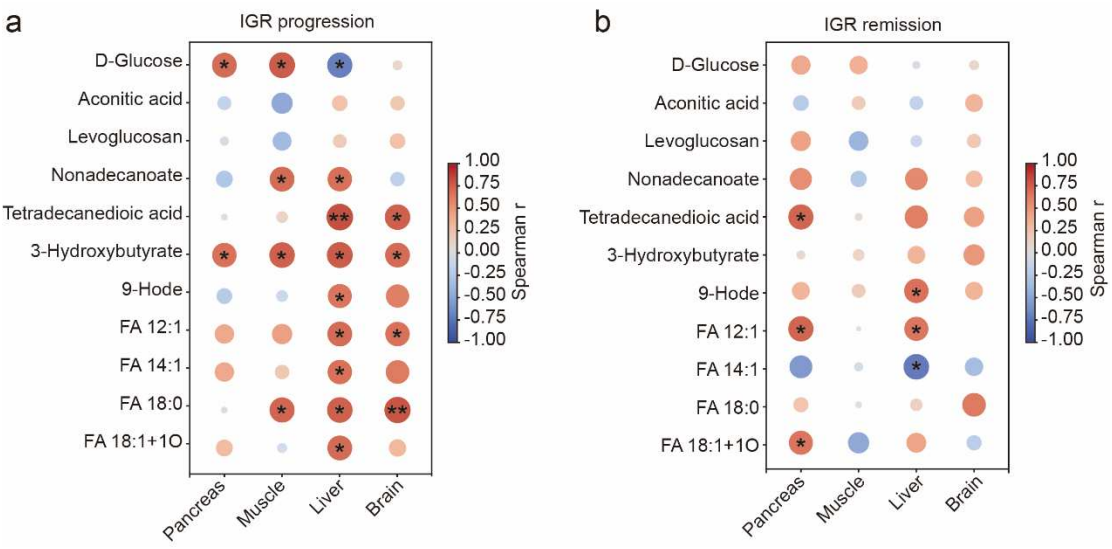

**Extended Data Fig. 1 Reproducible IGR-associated metabolic markers and eight-marker panel performance. a**, Odds ratios and 95% confidence intervals for reproducible IGR-associated metabolites estimated using logistic regression adjusted for age, sex and BMI. **b-c**, Receiver operating characteristic curves evaluating the discriminative performance of the eight-marker metabolic panel comprising CAR 16:0, glutamic acid, L-alanine, phenylalanine, FA 21:3;O, glycohyodeoxycholic acid (GHCA), PC O-42:6 and Met/Phe in the discovery and validation sets.




**Extended Data Fig. 2 The metabolic alterations in different IGR subgroups compared with matched healthy controls presented as fold changes.** Three comparisons were performed: IGR-2013$_{(NGR)}$ versus HC-2013, IGR-2013$_{(IGR)}$ versus HC-2013, and IGR-2013$_{(DM)}$ versus HC-2013, representing baseline lipid alterations in individuals who subsequently reverted to NGR, remained in the IGR state, or progressed to diabetes, respectively. The FC value was calculated as the ratio of the mean metabolite abundance in each IGR subgroup to that in the corresponding healthy control group.




**Extended Data Fig. 3 Tissue distribution of metabolic biomarkers in the pancreas, muscle, liver, and brain. a-d**, Heatmaps showing the abundance profiles of biomarkers across different organs in the IGR progression model. **e-h**, Heatmaps showing the abundance profiles of biomarkers across different organs in the IGR remission model. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$; ns, not significant.




**Extended Data Fig. 4** Correlation analysis between serum metabolic biomarkers and tissue metabolic features in the progression (a) and reversal (b) models. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$; ns, not significant.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTables.xlsx
- SupportInformation.docx