

# Metabolic Modeling Reveals Potential Muscle Metabolite Markers of Insulin Resistance in Newly Diagnosed Diabetic Patients

Maryam Khoshnejat

University of Tehran Institute of Biochemistry and Biophysics

Ali Mohammad Banaei-Moghaddam

University of Tehran Institute of Biochemistry and Biophysics

Kaveh Kavousi (✉ [kkavousi@ut.ac.ir](mailto:kkavousi@ut.ac.ir))

University of Tehran Institute of Biochemistry and Biophysics <https://orcid.org/0000-0002-1906-3912>

Ali Akbar Moosavi-Movahedi

University of Tehran Institute of Biochemistry and Biophysics

---

## Research

**Keywords:** Type 2 Diabetes, Skeletal muscle, metabolism, metabolic model, machine learning, marker, insulin resistance

**Posted Date:** December 15th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-124888/v1>

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

[Read Full License](#)

---

# Abstract

## Background

Type 2 diabetes mellitus (T2DM) is a challenging and globally ubiquitous metabolic disease caused by insulin resistance. Skeletal muscle is the major insulin-sensitive tissue that plays a great role in blood sugar homeostasis. Dysfunction of muscle metabolism is implicated in the disturbance of glucose hemostasis and the development of insulin resistance and T2DM. Here, we attempted to find metabolic dysregulations that are associated with the onset of T2DM. Besides, metabolite markers of T2DM were explored.

## Methods

We reconstructed a human muscle-specific metabolic model and applied it to perform metabolic analysis in newly diagnosed T2DM patients. We investigated the metabolism reprogramming by using two topology-based and constraint-based approach. Moreover, we applied a machine learning method to predict potential metabolite markers of insulin resistance in muscle.

## Results

Our results showed that metabolic alterations have occurred in carbohydrate, fatty acids, lipids, amino acids, and inositol phosphate metabolisms as well as pathways implicated in building extracellular matrix (ECM). Also, dysregulation of coenzyme Q10 metabolism was observed. Moreover, 13 exchange metabolites were predicted as the potential metabolite markers of insulin resistance in skeletal muscle. The efficiency of these markers in detecting insulin-resistant muscle was validated using a separate muscle gene expression data from another diabetes-related study.

## Conclusion

In this study, the most updated muscle-specific metabolic model was generated and successfully was validated. This model was used for the investigation of metabolic disturbances at the onset of T2DM. Our results indicated the significance of ECM metabolites in insulin resistance, and reinforce the role of coenzyme Q10 as a candidate for further research in insulin resistance and T2DM treatment. The model is freely available and can be used for other muscle metabolic studies. We also predicted metabolite markers of insulin resistance in the skeletal muscle, which can be considered for further empirical investigations.

## Background

Type 2 diabetes mellitus is a complex progressive metabolic disease with a heterogeneous etiology. The disease has a high global prevalence and is among the top 10 causes of death worldwide. According to the latest international diabetes federation report, the global prevalence of diabetes in 2019 is estimated

to be 463 million people. In addition, 374 million people are at greater risk of developing T2DM [1]. Therefore, this disease has become a major concern in the field of global health.

Insulin resistance is the main pathophysiologic feature in T2DM. The key insulin-sensitive tissues include skeletal muscle, liver, and adipose tissues. Among these tissues, skeletal muscle is responsible for the clearance of over 75% of glucose from the bloodstream thus plays a great role in lowering the blood glucose level [2]. Dysfunction of muscle metabolism is associated with disturbance in glucose hemostasis and implicated in the development of obesity, insulin resistance, and T2DM. Therefore, the study of the molecular mechanisms underlying insulin resistance in skeletal muscle will improve our understanding and treatments for T2DM.

In complex and multifactorial diseases such as diabetes, the use of systems biology provides a holistic view to study affected subsystems related to the disease in a cell. Meanwhile, genome-scale metabolic modeling serves as a context for in silico exploration of metabolism through a systems biology approach. Since there is not particularly a one-to-one relationship between transcription and translation, the study of gene expression data alone does not provide an accurate understanding of cellular metabolism while metabolic network simulation gives an in-depth insight into the molecular mechanisms involved in cellular processes. Generic metabolic networks are reconstructed based on the information encoded in the genome. This information includes all possible reactions that can occur in all types of human cells. By having gene expression data from a particular cell or disease state and integrating it into the generic model, active reactions are identified and a context-specific model can be reconstructed. Using genome-scale metabolic models (GEMs), some efforts have been made to study the molecular mechanisms contributing to muscle insulin resistance, so far. Bordbar et.al. have reconstructed a multi-tissue type genome-scale metabolic model and integrated gene expression data of obese and type 2 obese gastric bypass patients to study metabolic activity between these states [3]. Nogiec et.al. have applied flux balance analysis to model muscle insulin resistance metabolism in fasted to fed states. In addition, they tried to identify reactions that reproduce key features of insulin resistance by perturbing the muscle metabolic network [4]. Varemo et.al. have reconstructed a myocyte metabolic network and have utilized it to find reporter metabolites in type 2 diabetes mellitus [5]. However, so far there is not any study demonstrated the metabolic disturbance at the onset of diabetes. Since T2DM is a progressive disease, system-level identification of disorders in newly diagnosed T2DM patients who are at the early stages of the disease process can provide significant information about the underlying onset disorders in the disease and may somewhat clarify the causes of insulin resistance.

In this study, we attempted to find muscle metabolism dysregulations at the onset of type 2 diabetes mellitus. We used muscle gene expression data from newly diagnosed diabetic patients that their blood glucose levels were slightly higher than the diabetic diagnostic criterion, and had not taken any medications [6]. We first reconstructed the functional muscle-specific metabolic model and then employed it for our analyses. We applied a topology-based approach for the identification of reporter metabolites, which are associated with the dysregulated genes. Then using a constraint-based approach, the changes in muscle metabolic capabilities were studied. Moreover, potential metabolite markers

related to the muscle insulin resistance were predicted using a machine learning approach. Finally, we validated these markers with gene expression data from a separate study [7]. The overall workflow of this study is shown in Figure 1.

## Methods

### Data

We searched transcriptome databases for skeletal muscle data from healthy and T2DM individuals. We found the most comprehensive dataset from the dbGaP database with the accession code phs001068.v1.p1. The gene expression data were obtained from vastus lateralis muscle using high-throughput RNA-Seq technology. The data comprises of 91 healthy and 63 newly diagnosed subjects [6]. This transcriptome data were downloaded and were used for our analyses.

For validation of metabolite markers the muscle gene expression data of healthy and diabetic subjects were taken from the GEO repository with the accession numbers GSE63887 and GSE81965 [7]. This data contains RNA-Seq samples from 24 individuals in four equal following groups: normal glucose tolerant (NGT)/non-obese, NGT/obese, T2DM/non-obese, and T2DM/obese.

### Differential gene expression

The Bioconductor R package DESeq2 was used for differential gene expression analysis [8]. A pre-processing of data was applied involving pre-filtering of genes whose expression level was below a minimum cutoff level ( $< 5$  read counts in less than 25 percentages of samples). Between-samples normalization was done according to the DESeq2 manual. DEGs were found based on a negative binomial distribution. Multiple testing correction based on Benjamini–Hochberg procedure was applied and a false discovery rate lower than 0.1 was considered as the differentially expressed criterion.

### Muscle-specific metabolic model reconstruction

The muscle metabolic model was reconstructed based on the HMR2 model as the generic model and gene expression data of healthy muscles [9, 10]. The Gene-protein-reaction relationships in HMR2 were corrected using the Metabolic Atlas database [11]. For the preprocessing, low expressed genes ( $< 5$  read counts in less than 25 percentages of samples) were filtered and between-sample normalization according to the DESeq2 normalization method with gene length adjustment and  $\log_2(\text{gene expression}+1)$  transformation were applied [8]. Moreover, the myocyte biomass reaction from the Bordbar muscle metabolic model was added to our model [3]. E-Flux algorithm was employed to integrate gene-expression data to HMR2 model and to reconstruct the context-specific metabolic model [12]. To build a fully functional model we removed dead-end reactions. Figure 5 shows the workflow of this section. This model is freely available at [https://github.com/Maryamkhn/Muscle\\_MetNet\\_CBB](https://github.com/Maryamkhn/Muscle_MetNet_CBB).

### Reporter metabolite analysis

The Bioconductor R package piano was used for gene set analysis with the identification of reporter metabolites [13]. To determine the metabolite gene sets, the metabolite-reaction-gene associations of muscle metabolic model were used and gene IDs were assigned to each related metabolites. Adjusted p-values with the log2-fold changes obtained from DESeq2 were used here.

### **Personalized metabolic models and flux variability analysis**

Personalized GEMs were reconstructed by integrating gene expression data from each individual with the muscle metabolic model. Pre-processed data and E-Flux algorithm were used here. Media conditions were set by body fluid metabolites [14]. We used maximizing flux through the production of mitochondrial ATP as the objective function. In addition, to ensure having alive cell, we set the lower bound of biomass reaction to 80% of the maximum potential of the healthy model to produce biomass [15]. To obtain the minimum (min) and maximum (max) possible flux through each reaction in each personalized model, FVA was employed using Cobra Toolbox [16]. The two-sample t-test was applied on the min and max fluxes to find perturbed reactions between T2DM and healthy states. To apply multiple testing correction, the Benjamini–Hochberg procedure was used and reactions with false discovery rate less than 0.1 were regarded as significant perturbed reactions. Significant reactions were used to get perturbed pathways using flux enrichment analysis in the Cobra toolbox.

### **Classification and feature selection**

A SVM classifier with a polynomial kernel of order 2 was used for classification. The classifier was evaluated by 10-fold stratified cross-validation and analysis of accuracy, sensitivity, and specificity were reported.

For feature selection, a feature selection method, based on a hybrid of genetic algorithm and support vector machine was employed. In the GA, the feature subset is encoded as the candidate solution on a chromosome-like structure. The population is formed from a set of chromosomes that mutation and crossing over in it occur to generate the next generation. A fitness score is calculated for each chromosome representing adaptation of it to the environment and better feature subsets have more chance of reproducing the next generation. Creating the next generations will be continued until a stopping criterion is satisfied.

Here, a binary genetic algorithm was employed that the binary values of 1 and 0 represent the presence or absence of a specific feature at the particular chromosome, respectively. The chromosome length was set to the number of features. The population size and the maximal number of generations were set to 300 chromosomes and 100 generations, respectively. We used the accuracy of the SVM classifier as the fitness score. The genetic algorithm was stopped if the fitness score was reached an accuracy higher than 90 percent or the maximum number of generations was reproduced.

## **Results**

## Human muscle-specific metabolic model reconstruction

To reconstruct the muscle-specific metabolic network, the generic metabolic model and gene expression data of muscle were needed. For model reconstruction and analysis of metabolic disturbance in T2DM, we looked for the most comprehensive gene expression data from the muscle of healthy and diabetic individuals who had the following conditions: 1) newly diagnosed with T2DM. 2) Taking no medication for diabetes. 3) Do not have any other disease affecting the analysis. This resulted in data with the accession code phs001068.v1.p1 in the dbGaP database. Diabetic patients in this sample have been newly diagnosed as T2DM patients, and in most of them, the fasting blood glucose concentration was around 7 mmol/l, thus they were at the onset of diabetes. Table 1 shows subjects characteristics of the samples used in this study. Investigation of metabolism reprogramming in these patients who are at the onset of T2DM has significant value in identifying the underlying disorders in T2DM. We used gene expression data of healthy individuals in this sample for the reconstruction of the muscle-specific metabolic model.

Sample	Glucose (mmol/L)	Insulin ( $\mu$ u/l)	BMI	WHR
T2DM	$7.17 \pm 0.66$	$10.68 \pm 6.94$	$29.37 \pm 5.04$	$0.99 \pm 0.07$
Healthy	$5.62 \pm 0.33$	$6.87 \pm 3.34$	$26.35 \pm 3.47$	$0.92 \pm 0.08$

Table 1. Subjects characteristics of the individuals in the dataset [6]. Means  $\pm$  standard deviation values for fasting plasma glucose, fasting serum insulin, body mass index (BMI) and waist/hip ratio (WHR) in each diabetic and healthy groups are shown.

The muscle-specific metabolic model was reconstructed by integrating muscle gene expression data from healthy group into the Human Metabolic Reaction 2 (HMR2) [10] as the generic metabolic model. Muscle-specific biomass reaction was added to the model in order to have a functional and alive muscle model. The model comprises of 3585 metabolites, 5740 reactions and, 3627 genes. Table 2 shows the characteristics of the muscle-specific metabolic model in comparison with HMR2. For quality controls and validation, we first checked for the biomass production to assess that all the required reactions to have alive model are satisfied. Then, 56 metabolic tasks known to occur in all types of human cells were checked; these metabolic functions can be found in [17]. Moreover, for muscle-specific validation, like Bordbar et.al. [3], we tested the ability of the model to produce ATP from several different sources like as glucose, fatty acids, glycogen branched-chain amino acids, and ketone bodies. The model successfully passed all these tests. The reconstructed muscle-specific metabolic model was employed for further analyses.

	HMR2 model	Muscle model
<b>Reactions</b>	8181	5740
<b>Metabolites</b>	6006	3585
<b>Genes</b>	3765	3627

Table 2. Comparison of HMR2 metabolic model with our reconstructed muscle-specific metabolic model

### **Reporter metabolites associated with the onset of diabetes were identified by topology-based analysis**

In this analysis, the reconstructed muscle-specific metabolic network was employed as the scaffold to explore the effects of transcriptional reprogramming in the context of metabolism. Using this approach, it is possible to identify so-called reporter metabolites that are those metabolites associated with the most significant transcriptional changes [18]. Gene sets associated with each metabolite can be identified from a metabolic network through gene-reaction-metabolite relation. Here, the Piano R package was used to perform reporter metabolites analysis. Results demonstrated that from 3310 metabolites having associated genes, 275 ones were identified as reporter metabolites with adjusted p-value < 0.05 in at least one directional/non-directional class. Most of the reporter metabolites were affected by down-regulated genes. Also, most of them were implicated in mitochondrial functions such as the production of mitochondrial ATP, activated fatty acids, pyruvate, ubiquinol, ubiquinone, AKG, succinyl-CoA, NADH, Acetyl-CoA, lipoamide, 5,10-methenyl-THF, and 5,10-methylene-THF. In addition, some metabolites including ceramide, tyrosine, and tryptophan were affected by up-regulated genes. The reporter metabolites were mainly involved in glycolysis, tricarboxylic acid, fatty acids, and amino acids metabolism. The complete list of the reporter metabolites can be found in Additional File1 Table S1.

### **Constraint-based analysis of metabolism at the onset of T2DM reveals reprogramming of pathways related to extracellular matrix**

We attempted to capture cell metabolism at the system-level by using constraint-based metabolic modeling. In this approach, the metabolic model is converted to the mathematical format, and several constraints (e.g. adjusting lower and upper bound of flux in each reaction) were imposed on the model. Combining high-throughput transcriptional data and constraint-based models provides an opportunity to infer cellular metabolism at the relevant physiological state like diabetes [19]. Flux variability analysis (FVA) is a well-known constraint-based approach, which can provide allowable flux states in the metabolic model through linear programming based strategy [20].

In order to study metabolic capability changes in muscle between healthy and newly diagnosed diabetic patients, personalized GEMs were reconstructed and FVA was applied. Significant perturbed reactions between the two states identified using the two-sample t-test. This analysis resulted in perturbed reactions involving in several pathways such as glucose, amino acid, and lipid metabolisms. The perturbations in the glucose to glucose-6-phosphate conversion reaction and the glucose exchange

reaction were also observed. Many reactions contributing to fatty acids and lipid metabolism were dysregulated. Also, changes in fluxes of several reactions implicated in inositol phosphate, chondroitin, and heparin sulfate metabolism were observed. To find associated pathways with these perturbed reactions, flux enrichment analysis was employed. These significantly perturbed pathways are shown in Table 3.

<b>Group</b>	<b>Adjusted P-value</b>
Transport, mitochondrial	6.98E-27
Keratan sulfate biosynthesis	1.05E-23
Chondroitin / heparan sulfate biosynthesis	3.16E-22
Glycerolipid metabolism	6.86E-21
Exchange reactions	4.51E-14
Fatty acid activation (cytosolic)	5.44E-13
N-glycan metabolism	3.78E-12
Heparan sulfate degradation	4.15E-10
Nucleotide metabolism	9.91E-10
Acyl-CoA hydrolysis	3.67E-09
Carnitine shuttle	1.4E-07
Omega-6 fatty acid metabolism	1.07E-07
Arginine and proline metabolism	1.08E-07
Folate metabolism	2.8E-07
Beta oxidation of even-chain fatty acids (peroxisomal)	2.39E-06
Acylglycerides metabolism	5.77E-06
Sphingolipid metabolism	2.14E-05
Glycerophospholipid metabolism	2.59E-05
Glycine, serine and threonine metabolism	9.12E-05
Valine, leucine, and isoleucine metabolism	0.000622
Transport, nuclear	0.001187
Alanine, aspartate and glutamate metabolism	0.001398
Tricarboxylic acid cycle and glyoxylate/dicarboxylate metabolism	0.004234
Glycosphingolipid biosynthesis	0.004949
Pyrimidine metabolism	0.007809
Fatty acid elongation (even-chain)	0.017541
Amino sugar and nucleotide sugar metabolism	0.018405
Nicotinate and nicotinamide metabolism	0.018405

Pyruvate metabolism	0.018405
Cysteine and methionine metabolism	0.022241
Fatty acid elongation (odd-chain)	0.026726
Inositol phosphate metabolism	0.036758
Pentose and glucuronate interconversions	0.038008
Phenylalanine, tyrosine and tryptophan biosynthesis	0.040009
Omega-3 fatty acid metabolism	0.044516
Formation and hydrolysis of cholesterol esters	0.046091
Purine metabolism	0.048746
Pentose phosphate pathway	0.051904
Glycolysis / Gluconeogenesis	0.055204
Biopterin metabolism	0.064145
Propanoate metabolism	0.064552
Glycosphingolipid metabolism	0.083322

Table 3. Significant dysregulated pathways in T2DM obtained from flux enrichment analysis

### Potential metabolite markers prediction

The genome-scale model of human metabolism provides an opportunity to predict potential biofluid markers associated with the diseases. The overall approach involves identifying exchange metabolites that their flux range in disease state is shifted compared to healthy ones. The successful performance of this method has already been evaluated and validated [21]. Here, we applied this method to find exchange metabolite markers of insulin resistance in muscle. Two generic metabolic models of healthy and diabetic muscles were reconstructed using average gene expression data of each group and E-Flux method; Then, FVA analysis was applied to them. The exchange reactions in which the flux interval was shifted relative to the normal state were selected (Figure 2). This approach resulted in 32 exchange reactions. In order to evaluate the discriminative capability of these reactions at the individual level, a machine learning approach was employed. To do this, personalized metabolic models of healthy and diabetic individuals were reconstructed and FVA analysis was applied. The generated minimum and maximum fluxes of these reactions based on FVA analysis were used as the features for classification with support vector machine (SVM). This evaluation resulted in 72.22% accuracy, 81.11% specificity, and 60.32 sensitivity. We found that these exchange metabolites altered at the individuals' level.

To achieve a fewer and better feature subset with improvement in classification accuracy, a wrapper feature selection method was employed that comprises a combination of genetic algorithm (GA) and

SVM as the feature selector and classifier, respectively. Several subsets of features with which SVM classifier can discriminate diabetic patients from healthy ones with approximately 90 percentages accuracy were found. This GA-SVM procedure was iterated 100 times resulting in 100 features subsets; then, the observation frequency of each feature in these iterations calculated. These frequencies were considered as indicative of the importance of features and features with at least 80% frequency regarded as top-ranked features. The top-ranked features were related to 13 reactions exchanging glucose, alanine, aspartate, sodium, hyaluronate, galactose, GQ1b, acetaldehyde, deoxyribose, 4-OH-Estradiol, hypoxanthine, retinoate, and methylglyoxal. Subsequently, the SVM classification performance with the top-ranked features was assessed. Our analysis revealed that using these top-ranked features improved classification accuracy to 81.04 % with 73.02% sensitivity and 86.67% specificity. Since, the classifier performance will vary depending on which samples are assigned to the training set and which ones to the test set, we repeated the 10-fold cross-validation 100 times and evaluated the SVM performance. The boxplot of this evaluation is shown in Figure 3.

For further validation, we assessed the performance of these top-ranked features with data obtained from another study [7]. This data comprises of RNA-Seq samples from vastus lateralis muscle of 24 participants, which is divided into four following groups: normal glucose tolerant (NGT)/non-obese, NGT/obese, T2DM/non-obese, and T2DM/obese. Preprocessing of the read counts was applied. Then personalized metabolic models were reconstructed and FVA analysis was performed. Associated fluxes of top-ranked features were obtained and were used as the features for SVM classifier.

We divided these people into NGT and diabetic groups regardless of the obesity state and the generated fluxes of marker reactions in these 24 participants were used for validation with SVM. This analysis resulted in 49.46% accuracy, 42.25% specificity, and 56.67% sensitivity, which is remarkably low, probably due to the presence of normoglycemic obese people in the healthy group. Although the blood glucose level in these individuals is in the normal range, high levels of fasting blood insulin level ( $60 \pm 33$  pmol/l in NGT/obese vs  $25 \pm 7$  pmol/l in NGT/non-obese) indicates insulin resistance in this group, which is compensated by more insulin secretion. Thus, we decided to remove data from NGT/obese individuals and re-evaluated markers in this set. The result significantly altered with the 78.50% accuracy, 70.17% specificity, and 82.67% sensitivity. Therefore, the metabolites markers are related to insulin-resistant metabolism and can successfully discriminate insulin-resistant individuals from healthy ones.

## Discussion

Here, we attempted to unravel the metabolism reprogramming in skeletal muscle of newly diagnosed diabetic patients. Since diabetes is a metabolic disease, our goal was to identify the metabolic disorders that occur at the onset of the disease. We used the gene expression data from the most comprehensive human skeletal muscle transcriptome study to date. The diabetic sample consists of participants who have newly been diagnosed with diabetes (with an average blood glucose concentration of  $\sim 7.2$  mmol/L) and had not taken any medications. For understanding metabolism alterations in T2DM, we first reconstructed our functional muscle-specific metabolic model because we could not utilize Bordbar

[3] or Nogiec [4] muscle metabolic models due to the being small-scale. Varemo model [5] has been the most complete muscle metabolic model to date, but because of the wrong gene-protein-reaction (GPR) association and lack of biomass production reaction, it is not suitable for simulation. Thus, we reconstructed the most updated skeletal muscle metabolic model. The model quality was successfully validated for several pathways, and metabolites productions as well as biomass production. This model is freely available and can be used for other muscle metabolic analysis.

We employed the reconstructed model for investigation of dysmetabolism in T2DM by using topology-based and constraint-based analyses. By topology-based analysis, we determined those metabolites that are involved in reactions in which the associated genes are significantly dysregulated. We also applied constraint-based analysis to quantitatively compare metabolic capabilities between healthy and newly diagnosed diabetic patients. In these analyses ubiquinone and ubiquinol the oxidized and reduced forms of coenzyme Q10, respectively, were identified as reporter metabolites. These metabolites also participate in some perturbed reactions. Coenzyme Q10, a critical component of oxidative phosphorylation, is produced or consumed in the electron transport of the mitochondrial respiratory chain and possesses antioxidant and anti-inflammatory properties. Pieces of evidence have revealed that supplementations of coenzyme Q10 in T2DM patients, preserve mitochondrial function, reduce oxidative stress, and improve glucose tolerance [22, 23]. In addition, the administration of coenzyme Q10 in patients with prediabetes has alleviated the progression from prediabetes to diabetes [24]. Therefore, this metabolite should receive more attention in the treatment of diabetes.

We also found that metabolic alterations occurred in carbohydrates, fatty acids, lipids, and amino acids metabolisms. Dysregulation of Branched-chain amino acids (BCAAs) metabolism results in the serine phosphorylation of insulin receptor substrates and subsequent uncoupling of insulin signaling [25]. Moreover, perturbations in metabolism of inositol phosphate, keratin, chondroitin, and heparan sulfate were observed. Inositol mediates insulin signal transduction, associated with glucose uptake and plays an important role in oxidative stress and inflammation. Administration of inositol supplements improves glucose metabolism and insulin resistance [26]. Chondroitin sulfate, keratin sulfate, heparan sulfate, and hyaluronic acid are glycosaminoglycans (GAGs). GAGs play a vital role in cell physiology including cell signaling, proliferation, and cell adhesion. In T2DM, alteration in GAGs structures and functions can occur [27]. The insulin-sensitizing and anti-diabetic impacts of some GAGs have been reported [28, 29]. Perturbations in GAGs related pathways and sphingolipid metabolism imply the role of the ECM in insulin resistance, which involved in the regulation of insulin action. Our analyses suggested that it seems muscle dysmetabolism disrupts the abundance of metabolites involved in the process of sensing insulin and transmitting insulin signal into the cell. Decreased insulin sensitivity results in lower expression of insulin-responsive genes, reduced glucose uptake, and consequently changes in energy, glucose, lipids, and amino acids metabolisms (Figure 4).

As the final analysis, potential metabolite markers were predicted. For this purpose, we first identified exchange metabolites that their flux interval was shifted in comparison with healthy ones. Then, a wrapper feature selection method applied to find potential metabolite markers. This approach led to the

identification of 13 exchange reactions that could discriminate healthy individuals from T2DM patients with 81% accuracy. We validated these markers using a separate gene expression data from another study that has investigated the gene expression pattern in the muscle of obese and non-obese of healthy and diabetic individuals. Their result has shown that transcriptional reprogramming in obesity is similar to that occur in T2DM [7]. Here, we found that using all data from this study, including normoglycemic obese individuals for validation, the accuracy of our proposed marker was notably low (~ 50%). We thought that this low accuracy may be due to the metabolic similarities between obese and diabetic individuals. In fact, as noted in the original data article, obese and diabetic individuals have shown similar gene expression patterns [7] that can lead to similar metabolisms. Moreover, obese individuals had high levels of fasting insulin levels, which demonstrate the insulin resistance in this group. To test this issue, we removed obese healthy individuals from the validation set and examined the classification result, which leads to the improvement in accuracy to 78.50%. Therefore, this analysis confirmed both the original data article claim about the similarity of the gene expression pattern between obese and diabetic individuals and the appropriate efficiency of our proposed markers in identifying insulin-resistant individuals. These markers represent some of the insulin-resistance associated abnormalities represented in exchanging metabolites. Important metabolites such as methylglyoxal, hyaluronan, retinoic acid (vitamin A), sodium, alanine, and aspartate are present in the predicted markers. Notably, this method also successfully identified glucose as one of the markers. Methylglyoxal, a glycolytic by-product, is a toxic and highly reactive compound involved in cellular dysregulations. This compound modifies nucleotides, proteins, and lipids producing advanced glycation end products (AGEs) which also contribute to the diabetes complications. Methylglyoxal is associated with oxidative stress, cellular inflammation, and age-related disease such as diabetes [30]. Several studies have revealed the impact of methylglyoxal on insulin signaling pathways and insulin resistance [31, 32] and recently this metabolite has introduced as an emerging marker for T2DM diagnosis [30]. Hyaluronan is an anionic GAG metabolite implicated in several functions like as cell signaling, proliferation and migration, and angiogenesis. This metabolite also contributes to the inflammation and pathogenesis of T2DM [33, 34]. Studies have demonstrated that hyaluronan increases in the serum and skeletal muscle of T2DM subjects [35, 36]. Several analyses have shown the possible roles of vitamin A in glucose metabolism, and the progression of insulin resistance [37-39]. Association of purine metabolites such as xanthine and hypoxanthine with the risk of T2DM incidence and complications has been reported [40, 41]. Also, change in serum concentration of amino acids [42], and sodium [43] is associated with obesity and T2DM. In addition, the implication of gangliosides in insulin resistance has been shown [44-46]. Here, GQ1b ganglioside was predicted in the top-ranked metabolites list that can be considered for future analysis. We also checked metabolomics-based studies of T2DM and the Human Metabolome Database for these metabolites [47]. We found that glucose, hypoxanthine, alanine, aspartate, galactose, hyaluronate, and methylglyoxal levels have been reported to be associated with T2DM [40, 47, 48]. These metabolite markers can be used for further empirical investigation to verify their prognostic and diagnostic values in insulin resistance.

## Conclusion

In the present study, the dysmetabolism at the onset of T2DM was studied. We found that in addition to the glucose, lipid, and amino acid metabolism, the inositol phosphate and metabolisms related to GAGs e.g. chondroitin, keratan, and heparin sulfate metabolisms also were perturbed. These results indicate the importance of ECM metabolites in insulin signaling and subsequent insulin resistance. Coenzyme Q10, which observed both in the list of reporter metabolites and in metabolites of perturbed reactions, could be a candidate for further investigation in order to find its role in insulin resistance and glycemic control. Moreover, by employing a machine-learning approach, 13 exchange metabolites were predicted as a potential marker of insulin-resistant skeletal muscle, which can be regarded for further clinical studies in prognosis and diagnosis of T2DM. Besides, the reconstructed muscle metabolic model in this study is freely available at [https://github.com/Maryamkhn/Muscle\\_MetNet\\_CBB](https://github.com/Maryamkhn/Muscle_MetNet_CBB), which can be used for other muscle metabolic studies.

## Abbreviations

Type 2 diabetes mellitus	T2DM
Normal glucose tolerant	NGT
Support vector machine	SVM
Genetic algorithm	GA
Genome-scale metabolic models	GEMs
Flux variability analysis	FVA
Differentially expressed genes	DEGs
Extracellular matrix	ECM
Body mass index	BMI
waist/hip ratio	WHR
Branched-chain amino acids	BCAAs
Advanced glycation end products	AGEs
Glycosaminoglycans	GAGs
Gene-protein-reaction	GPR
Human Metabolic Reaction 2	HMR2
Maximum	Max
Minimum	Min

## Declarations

## **Ethics approval and consent to participate**

Not applicable.

## **Consent for publication**

Not applicable.

## **Availability of data and materials**

The dataset analyzed during the current study is available in the dbGaP database with the accession code phs001068.v1.p1. This data is accessible through the repository's data access request procedures (<https://dbgap.ncbi.nlm.nih.gov/>). The validation datasets are available at GEO repository with the accession numbers GSE63887 and GSE81965 (<https://www.ncbi.nlm.nih.gov/geo/>). The muscle metabolic model reconstructed in this article is available at [https://github.com/Maryamkhn/Muscle\\_MetNet\\_CBB](https://github.com/Maryamkhn/Muscle_MetNet_CBB)

## **Competing interests**

The authors declare no conflicts of interest.

## **Funding**

Not applicable.

## **Author's Contributions**

MK: Conceptualization, design, data gathering, implementation, analysis, interpretation, writing. AMB-M: Supervision, design, interpretation of data, review & editing. KK: Supervision, interpretation of data, review & editing. AM: Supervision, review. All authors have read and approved the final manuscript.

## **Acknowledgments**

Not applicable.

## **Additional material**

Additional file 1(Additional\_file1.xlsx): The complete list of the reporter metabolites could be found in Table S1.

## **References**

1. Saeedi P, Petersohn I, Salpea P, Malanda B, Karuranga S, Unwin N, Colagiuri S, Guariguata L, Motala AA, Ogurtsova K: Global and regional diabetes prevalence estimates for 2019 and projections for

- 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas. *Diabetes research and clinical practice* 2019, 157:107843.
2. Stump CS, Henriksen EJ, Wei Y, Sowers JR: The metabolic syndrome: role of skeletal muscle metabolism. *Annals of medicine* 2006, 38(6):389-402.
  3. Bordbar A, Feist AM, Usaite-Black R, Woodcock J, Palsson BO, Famili I: A multi-tissue type genome-scale metabolic network for analysis of whole-body systems physiology. *BMC systems biology* 2011, 5(1):180.
  4. Nogiec C, Burkart A, Dreyfuss JM, Lerin C, Kasif S, Patti M-E: Metabolic modeling of muscle metabolism identifies key reactions linked to insulin resistance phenotypes. *Molecular metabolism* 2015, 4(3):151-163.
  5. Våremo L, Scheele C, Broholm C, Mardinoglu A, Kampf C, Asplund A, Nookaew I, Uhlén M, Pedersen BK, Nielsen J: Proteome-and transcriptome-driven reconstruction of the human myocyte metabolic network and its use for identification of markers for diabetes. *Cell reports* 2015, 11(6):921-933.
  6. Scott LJ, Erdos MR, Huyghe JR, Welch RP, Beck AT, Wolford BN, Chines PS, Didion JP, Narisu N, Stringham HM: The genetic regulatory signature of type 2 diabetes in human skeletal muscle. *Nature communications* 2016, 7.
  7. Våremo L, Henriksen TI, Scheele C, Broholm C, Pedersen M, Uhlén M, Pedersen BK, Nielsen J: Type 2 diabetes and obesity induce similar transcriptional reprogramming in human myocytes. *Genome medicine* 2017, 9(1):47.
  8. Love MI, Huber W, Anders S: Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology* 2014, 15(12):550.
  9. Mardinoglu A, Agren R, Kampf C, Asplund A, Nookaew I, Jacobson P, Walley AJ, Froguel P, Carlsson LM, Uhlen M: Integration of clinical data with a genome-scale metabolic model of the human adipocyte. *Molecular systems biology* 2013, 9(1).
  10. Mardinoglu A, Agren R, Kampf C, Asplund A, Uhlen M, Nielsen J: Genome-scale metabolic modelling of hepatocytes reveals serine deficiency in patients with non-alcoholic fatty liver disease. *Nature communications* 2014, 5:3083.
  11. Robinson JL, Kocabaş P, Wang H, Cholley P-E, Cook D, Nilsson A, Anton M, Ferreira R, Domenzain I, Billa V: An atlas of human metabolism. *Science Signaling* 2020, 13(624).
  12. Colijn C, Brandes A, Zucker J, Lun DS, Weiner B, Farhat MR, Cheng T-Y, Moody DB, Murray M, Galagan JE: Interpreting expression data with metabolic flux models: predicting Mycobacterium tuberculosis mycolic acid production. *PLoS Comput Biol* 2009, 5(8):e1000489.
  13. Våremo L, Nielsen J, Nookaew I: Enriching the gene set analysis of genome-wide data by incorporating directionality of gene expression and combining statistical hypotheses and methods. *Nucleic acids research* 2013, 41(8):4378-4391.
  14. Hadi M, Marashi S-A: Reconstruction of a generic metabolic network model of cancer cells. *Molecular BioSystems* 2014, 10(11):3014-3021.

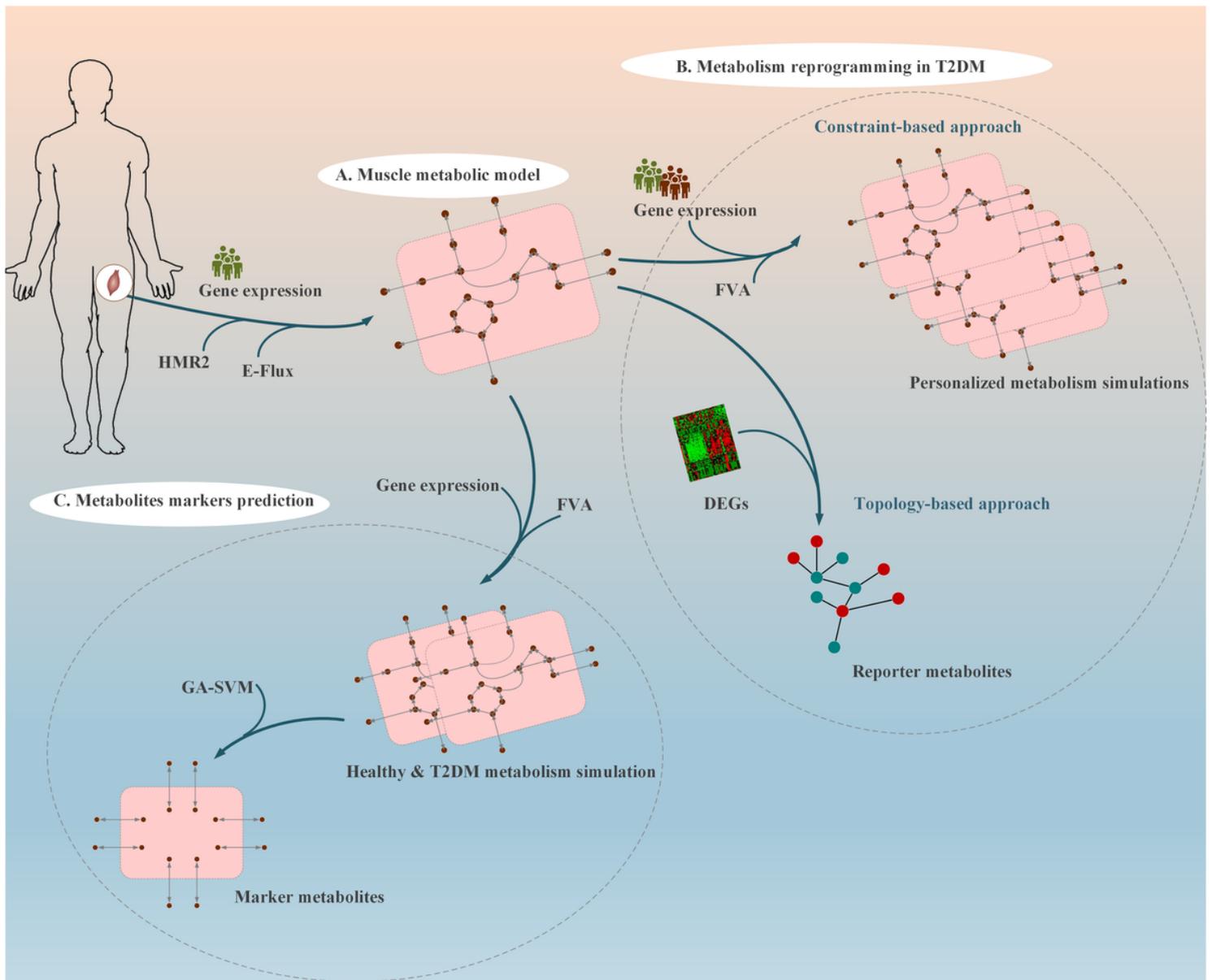
15. Chénard T, Guénard F, Vohl M-C, Carpentier A, Tchernof A, Najmanovich RJ: Remodeling adipose tissue through in silico modulation of fat storage for the prevention of type 2 diabetes. *BMC systems biology* 2017, 11(1):60.
16. Heirendt L, Arreckx S, Pfau T, Mendoza SN, Richelle A, Heinken A, Haraldsdottir HS, Keating SM, Vlasov V, Wachowiak J: Creation and analysis of biochemical constraint-based models: the COBRA Toolbox v3. 0. *arXiv preprint arXiv:171004038* 2017.
17. He Y, Wang Y, Zhang B, Li Y, Diao L, Lu L, Yao J, Liu Z, Li D, He F: Revealing the metabolic characteristics of human embryonic stem cells by genome-scale metabolic modeling. *FEBS letters* 2018, 592(22):3670-3682.
18. Patil KR, Nielsen J: Uncovering transcriptional regulation of metabolism by using metabolic network topology. *Proceedings of the national academy of sciences* 2005, 102(8):2685-2689.
19. Bordbar A, Monk JM, King ZA, Palsson BO: Constraint-based models predict metabolic and associated cellular functions. *Nature Reviews Genetics* 2014, 15(2):107.
20. Mahadevan R, Schilling C: The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. *Metabolic engineering* 2003, 5(4):264-276.
21. Shlomi T, Cabili MN, Ruppin E: Predicting metabolic biomarkers of human inborn errors of metabolism. *Molecular systems biology* 2009, 5(1).
22. Yen C-H, Chu Y-J, Lee B-J, Lin Y-C, Lin P-T: Effect of liquid ubiquinol supplementation on glucose, lipids and antioxidant capacity in type 2 diabetes patients: a double-blind, randomised, placebo-controlled trial. *British Journal of Nutrition* 2018, 120(1):57-63.
23. Shen Q, Pierce JD: Supplementation of coenzyme Q10 among patients with type 2 diabetes mellitus. In: *Healthcare: 2015*: Multidisciplinary Digital Publishing Institute; 2015: 296-309.
24. Yoo J-Y, Yum K-S: Effect of Coenzyme Q10 on Insulin Resistance in Korean Patients with Prediabetes: A Pilot Single-Center, Randomized, Double-Blind, Placebo-Controlled Study. *BioMed research international* 2018, 2018.
25. Lynch CJ, Adams SH: Branched-chain amino acids in metabolic signalling and insulin resistance. *Nature Reviews Endocrinology* 2014, 10(12):723.
26. Bevilacqua A, Bizzarri M: Inositols in insulin signaling and glucose metabolism. *International journal of endocrinology* 2018, 2018.
27. Gowd V, Gurukar A, Chilkunda ND: Glycosaminoglycan remodeling during diabetes and the role of dietary factors in their modulation. *World journal of diabetes* 2016, 7(4):67.
28. Reynés B, Serrano A, Petrov PD, Ribot J, Chetrit C, Martínez-Puig D, Bonet ML, Palou A: Anti-obesity and insulin-sensitising effects of a glycosaminoglycan mix. *Journal of Functional Foods* 2016, 26:350-362.
29. Moto M, Takamizawa N, Shibuya T, Nakamura A, Katsuraya K, Iwasaki K, Tanaka K, Murota A: Anti-diabetic effects of chondroitin sulfate on normal and type 2 diabetic mice. *Journal of Functional Foods* 2018, 40:336-340.

30. Bhat LR, Vedantham S, Krishnan UM, Rayappan JBB: Methylglyoxal–an emerging biomarker for diabetes mellitus diagnosis and its detection methods. *Biosensors and bioelectronics* 2019.
31. Riboulet-Chavey A, Pierron A, Durand I, Murdaca J, Giudicelli J, Van Obberghen E: Methylglyoxal impairs the insulin signaling pathways independently of the formation of intracellular reactive oxygen species. *Diabetes* 2006, 55(5):1289-1299.
32. Maessen DE, Stehouwer CD, Schalkwijk CG: The role of methylglyoxal and the glyoxalase system in diabetes and other age-related diseases. *Clinical Science* 2015, 128(12):839-861.
33. Litwiniuk M, Krejner A, Speyrer M, Gauto A, Grzela T: Hyaluronic acid in inflammation and tissue regeneration. *Wounds* 2016, 28(3):78-88.
34. Sainio A, Takabe P, Oikari S, Salomäki-Myftari H, Koulu M, Söderström M, Pasonen-Seppänen S, Järveläinen H: Metformin decreases hyaluronan synthesis by vascular smooth muscle cells. *Journal of Investigative Medicine* 2020, 68(2):383-391.
35. Kang L, Lantier L, Kennedy A, Bonner JS, Mayes WH, Bracy DP, Bookbinder LH, Hasty AH, Thompson CB, Wasserman DH: Hyaluronan accumulates with high-fat feeding and contributes to insulin resistance. *Diabetes* 2013, 62(6):1888-1896.
36. Nagy N, Sunkari VG, Kaber G, Hasbun S, Lam DN, Speake C, Sanda S, McLaughlin TL, Wight TN, Long SR: Hyaluronan levels are increased systemically in human type 2 but not type 1 diabetes independently of glycemic control. *Matrix Biology* 2019, 80:46-58.
37. Beydoun MA, Chen X, Jha K, Beydoun HA, Zonderman AB, Canas JA: Carotenoids, vitamin A, and their association with the metabolic syndrome: A systematic review and meta-analysis. *Nutrition reviews* 2019, 77(1):32-45.
38. Blaner WS: Vitamin A signaling and homeostasis in obesity, diabetes, and metabolic disorders. *Pharmacology & therapeutics* 2019.
39. Kuang H, Wei C-h, Wang T, Eastep J, Li Y, Chen G: Vitamin A status affects weight gain and hepatic glucose metabolism in rats fed a high-fat diet. *Biochemistry and Cell Biology* 2019, 97(5):545-553.
40. Papandreou C, Li J, Liang L, Bulló M, Zheng Y, Ruiz-Canela M, Yu E, Guasch-Ferré M, Razquin C, Clish C: Metabolites related to purine catabolism and risk of type 2 diabetes incidence; modifying effects of the TCF7L2-rs7903146 polymorphism. *Scientific reports* 2019, 9(1):1-11.
41. Varadaiah YGC, Sivanesan S, Nayak SB, Thirumalarao KR: Purine metabolites can indicate diabetes progression. *Archives of Physiology and Biochemistry* 2019:1-5.
42. Gar C, Rottenkolber M, Prehn C, Adamski J, Seissler J, Lechner A: Serum and plasma amino acids as markers of prediabetes, insulin resistance, and incident diabetes. *Critical reviews in clinical laboratory sciences* 2018, 55(1):21-32.
43. Soltani S, Kolahdouz Mohammadi R, Shab-Bidar S, Vafa M, Salehi-Abargouei A: Sodium status and the metabolic syndrome: A systematic review and meta-analysis of observational studies. *Critical reviews in food science and nutrition* 2019, 59(2):196-206.
44. Yamashita T, Hashiramoto A, Haluzik M, Mizukami H, Beck S, Norton A, Kono M, Tsuji S, Daniotti JL, Werth N: Enhanced insulin sensitivity in mice lacking ganglioside GM3. *Proceedings of the National*

*Academy of Sciences* 2003, 100(6):3445-3449.

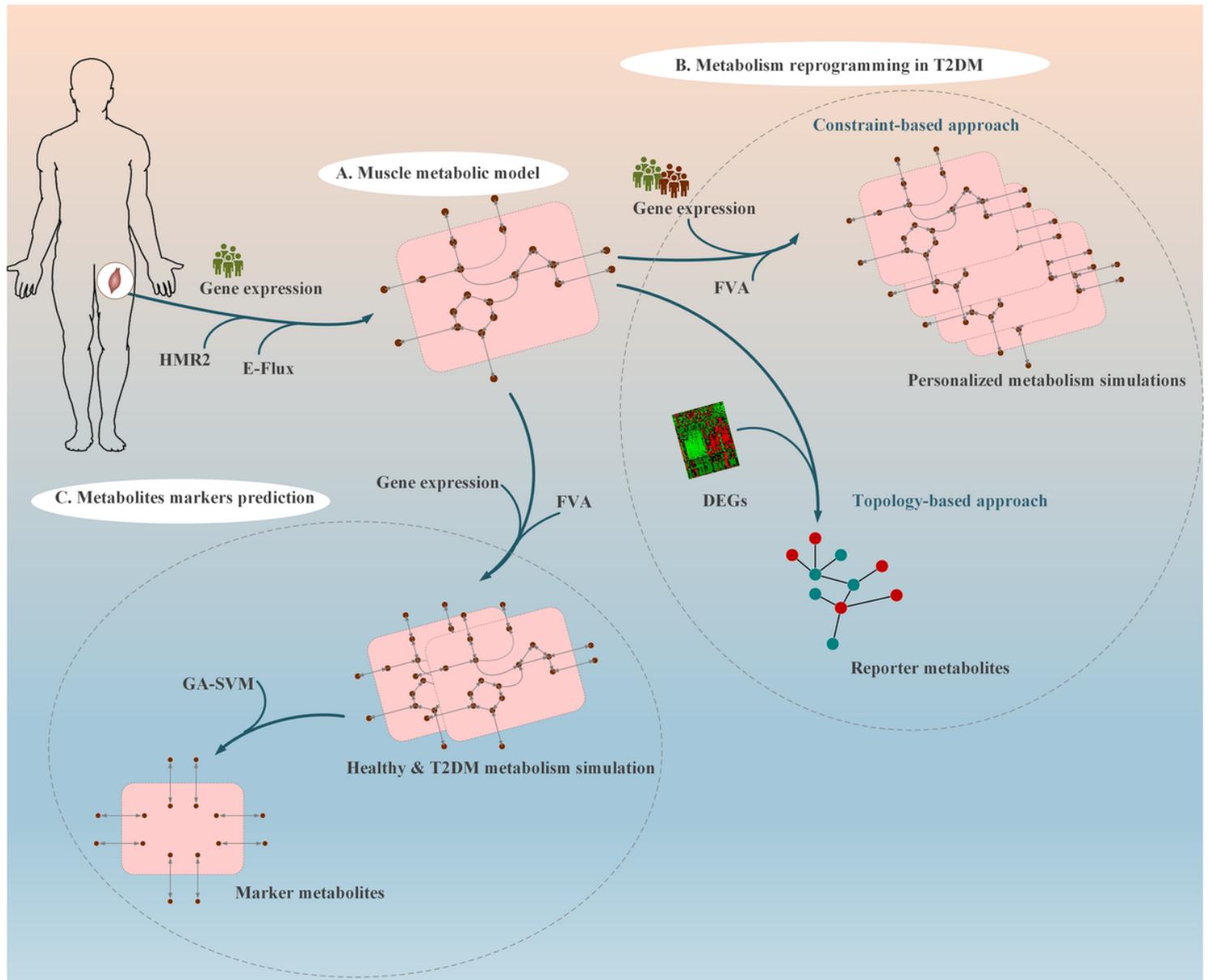
45. Inokuchi J-i, Kabayama K, Sato T, Igarashi Y: A New Pathological Feature of Insulin Resistance and Type 2 Diabetes: Involvement of Ganglioside GM3 and Membrane Microdomains. In: *Sphingolipid Biology*. Springer; 2006: 273-284.
46. Sasaki N, Itakura Y, Toyoda M: Gangliosides contribute to vascular insulin resistance. *International journal of molecular sciences* 2019, 20(8):1819.
47. Wishart DS, Feunang YD, Marcu A, Guo AC, Liang K, Vázquez-Fresno R, Sajed T, Johnson D, Li C, Karu N: HMDB 4.0: the human metabolome database for 2018. *Nucleic acids research* 2018, 46(D1):D608-D617.
48. Satheesh G, Ramachandran S, Jaleel A: Metabolomics-Based Prospective Studies and Prediction of Type 2 Diabetes Mellitus Risks. *Metabolic Syndrome and Related Disorders* 2020, 18(1):1-9.

## Figures



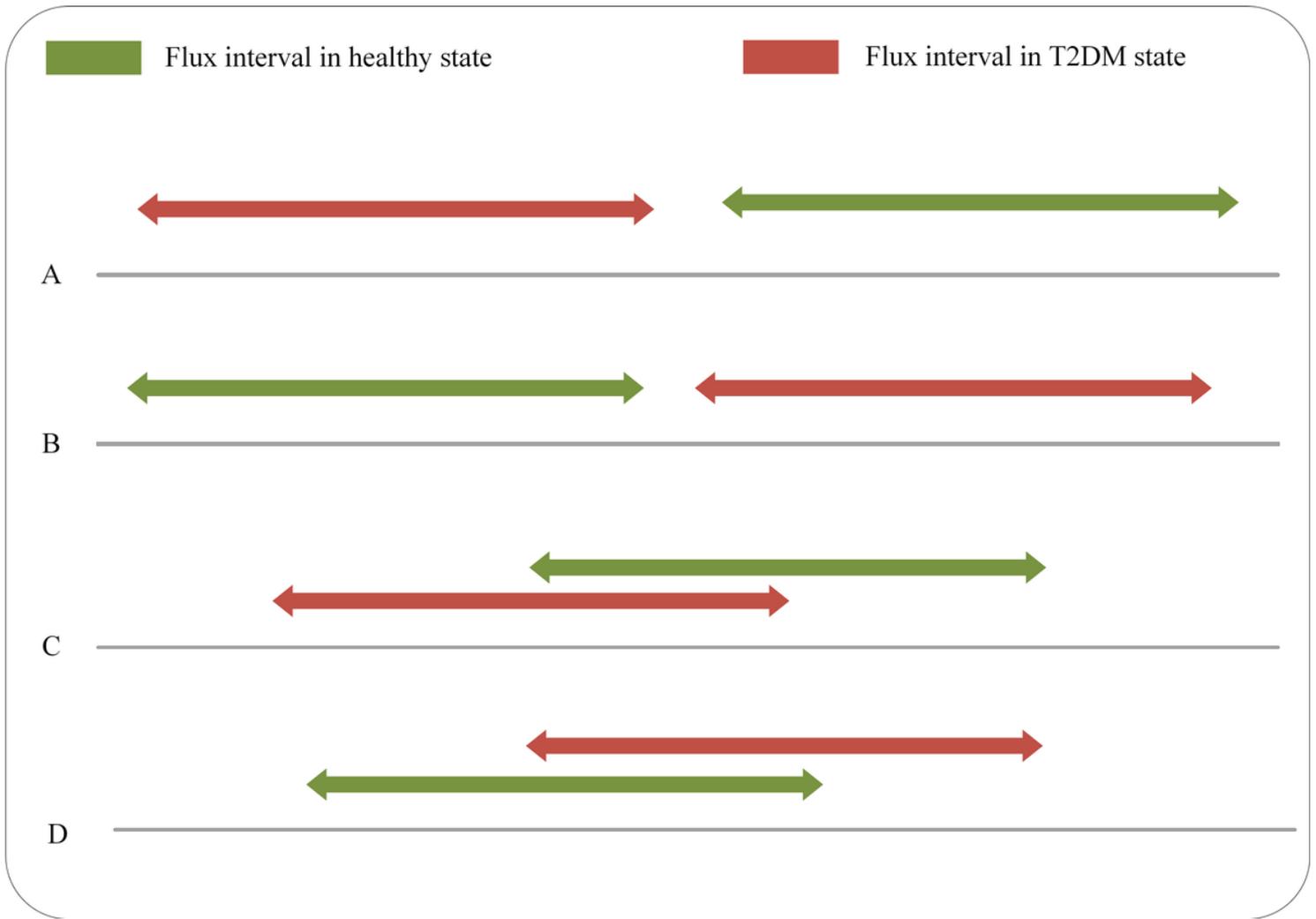
**Figure 1**

The workflow of study design. (A) The muscle-specific metabolic model was reconstructed and was employed for our following analyses: (B) Study of metabolism reprogramming in newly diagnosed T2DM using the two constraint-based and topology-based approaches. (C) Potential metabolite markers prediction of insulin resistance in muscle by applying a machine-learning technique.



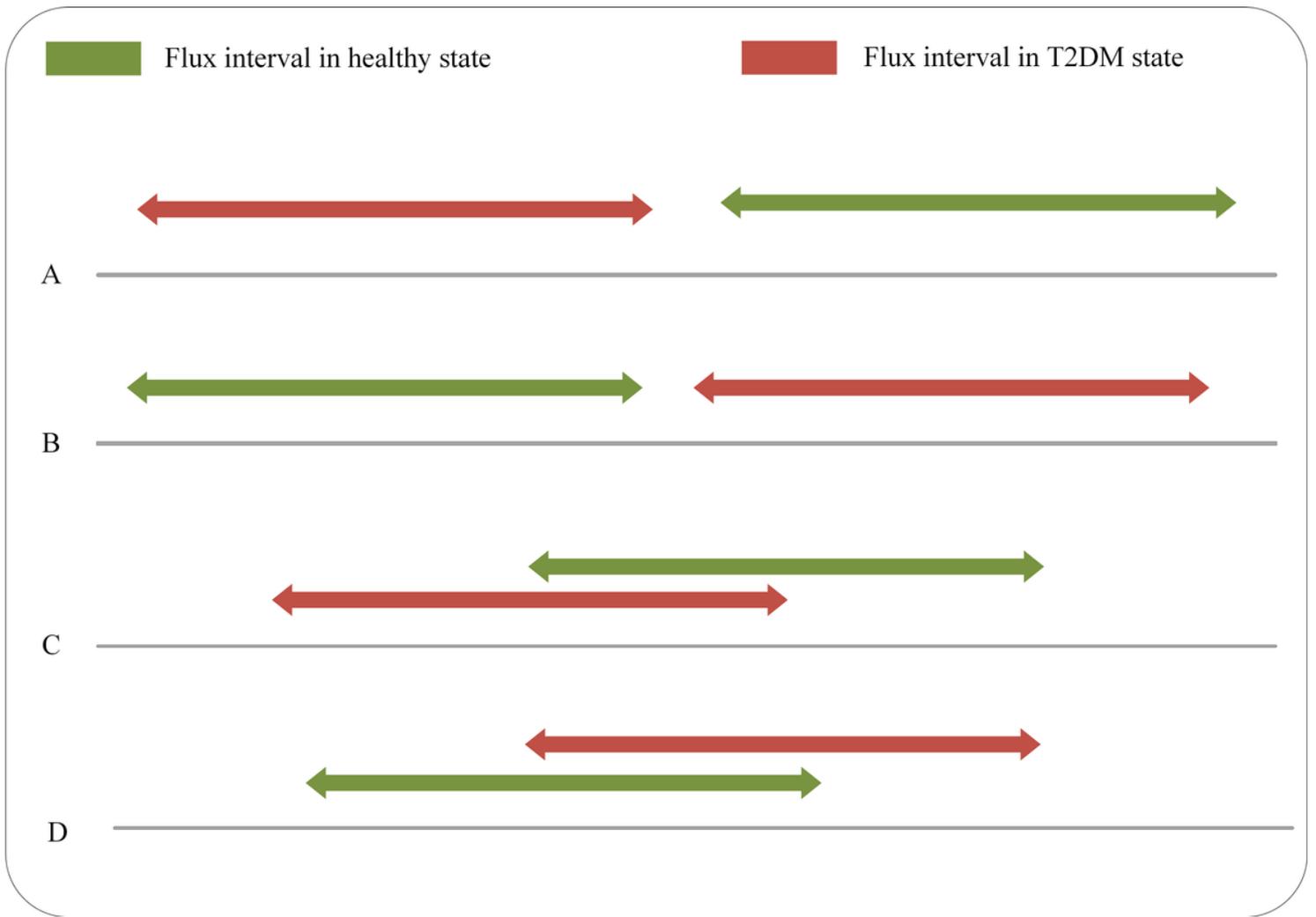
**Figure 1**

The workflow of study design. (A) The muscle-specific metabolic model was reconstructed and was employed for our following analyses: (B) Study of metabolism reprogramming in newly diagnosed T2DM using the two constraint-based and topology-based approaches. (C) Potential metabolite markers prediction of insulin resistance in muscle by applying a machine-learning technique.



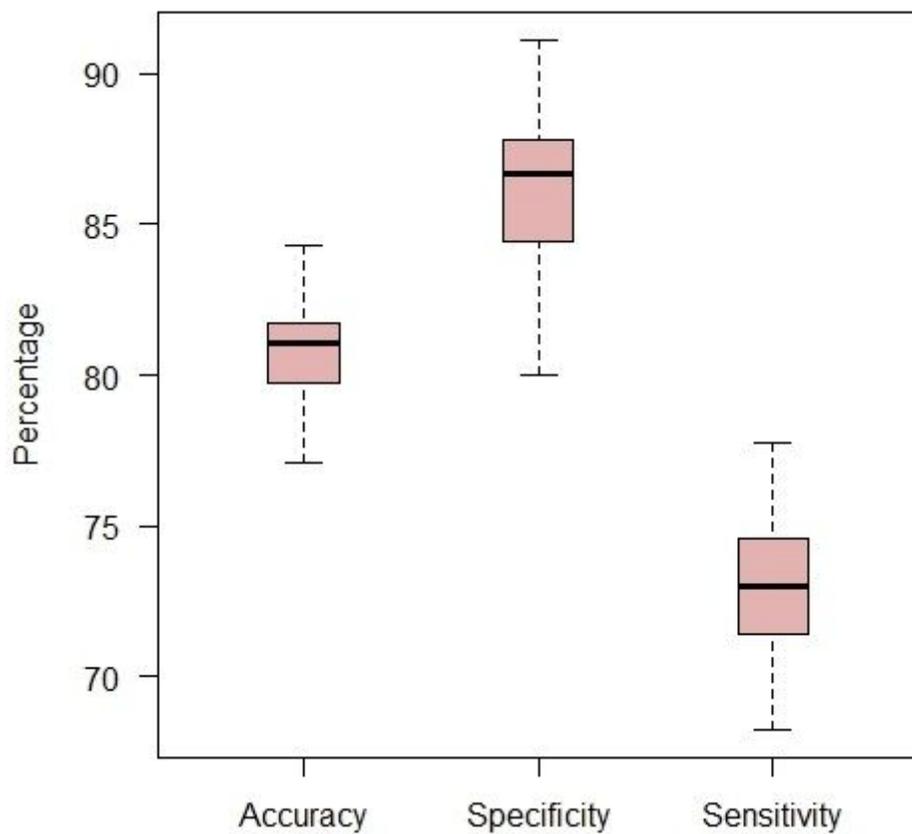
**Figure 2**

Major possibilities of flux interval shift in T2DM (red) versus healthy state (green). Those reactions in which flux interval shifts were grouped into one of A to D were selected as the primary markers.



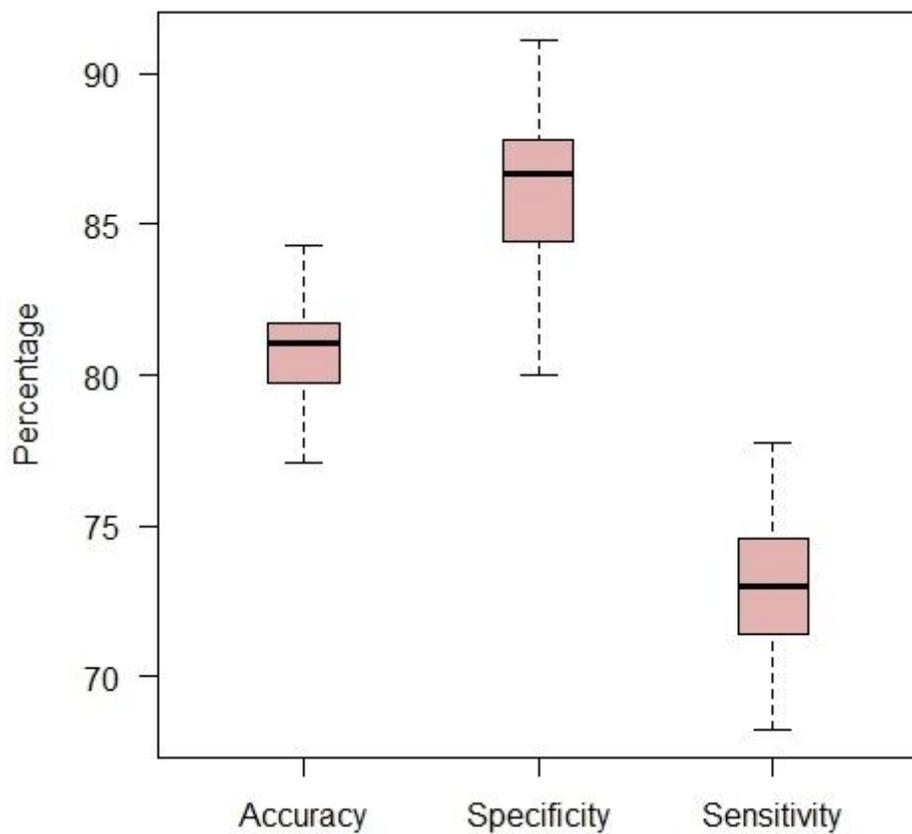
**Figure 2**

Major possibilities of flux interval shift in T2DM (red) versus healthy state (green). Those reactions in which flux interval shifts were grouped into one of A to D were selected as the primary markers.



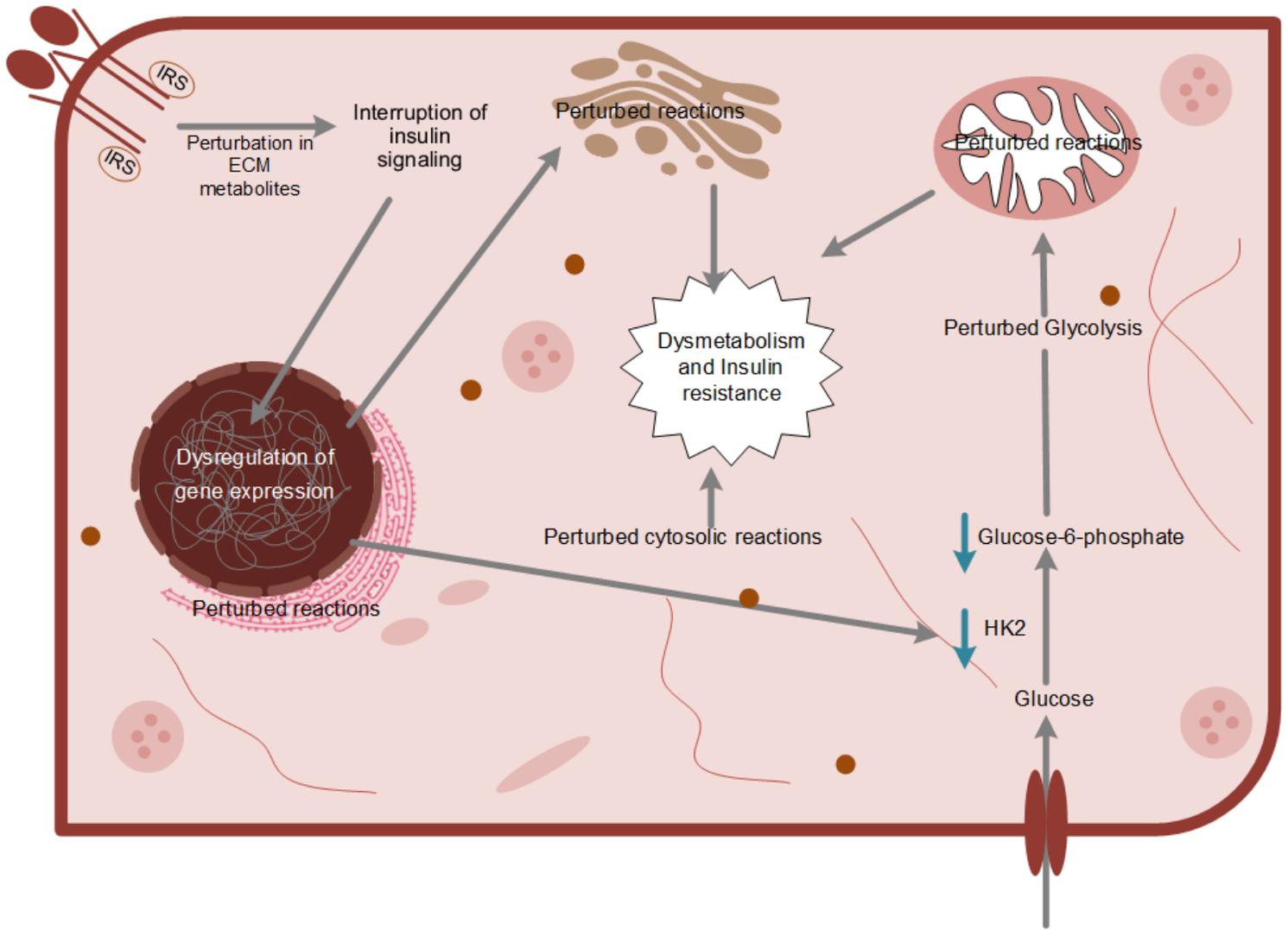
**Figure 3**

Boxplot of SVM evaluation using 13 metabolite markers as features. 10-fold cross validation was repeated 100 times and the SVM performance was evaluated.



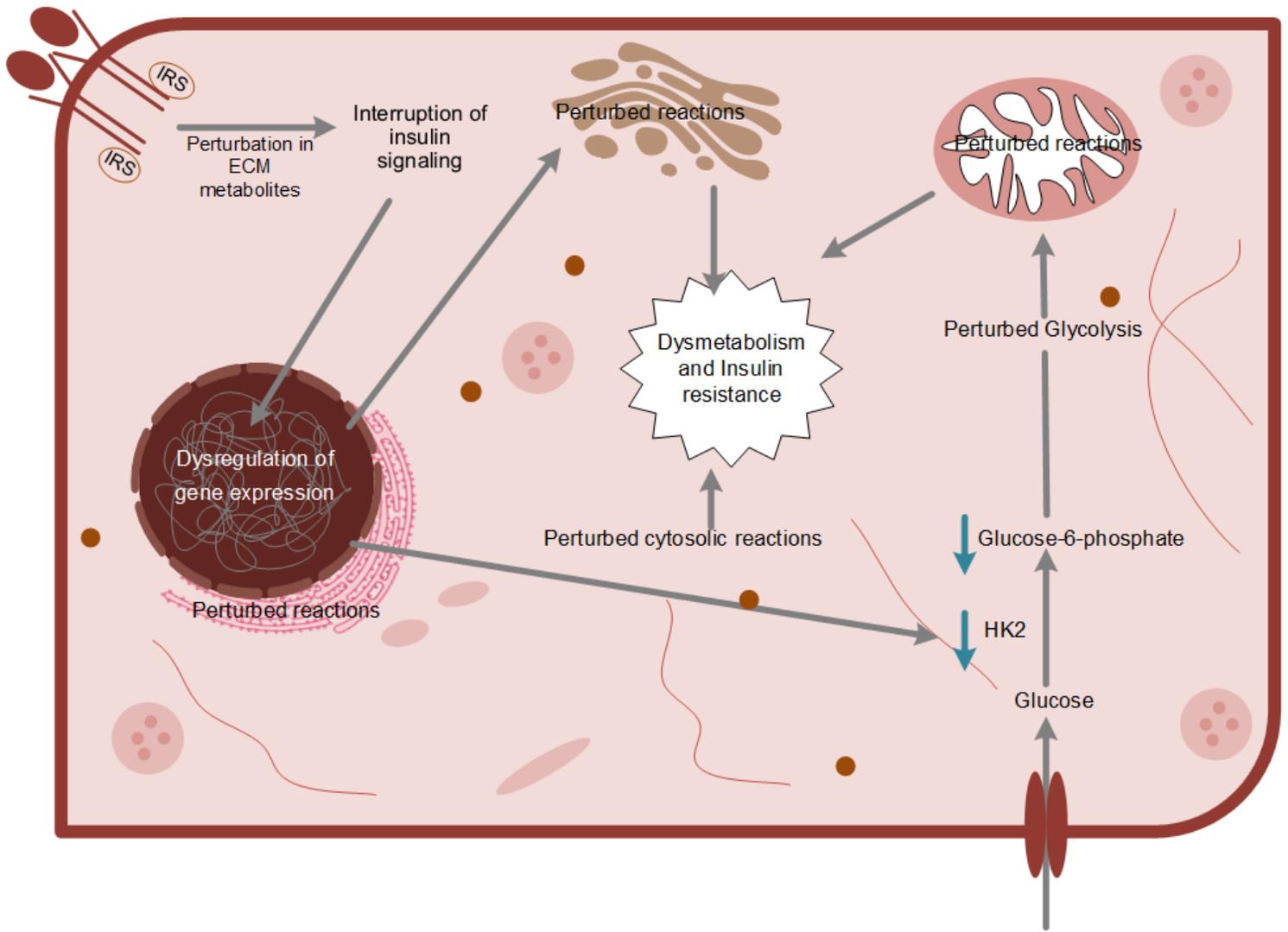
**Figure 3**

Boxplot of SVM evaluation using 13 metabolite markers as features. 10-fold cross validation was repeated 100 times and the SVM performance was evaluated.



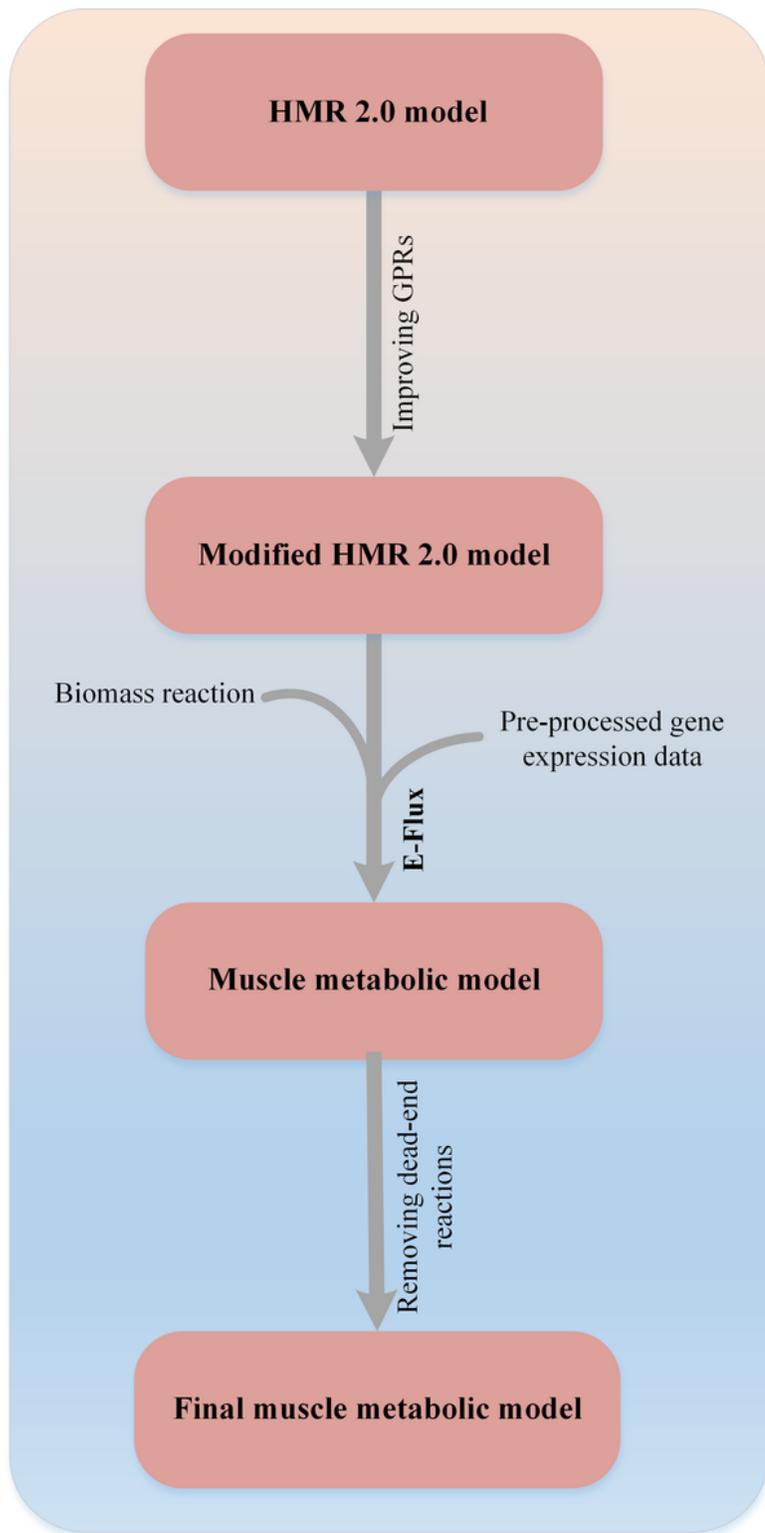
**Figure 4**

Schematic illustration of metabolism dysregulation in T2DM. Our analysis demonstrated that dysmetabolism of those metabolites implicated in extracellular matrix and insulin signaling possibly leads to a lower sense of insulin and subsequent lower expression of insulin-responsive genes such as HK2. The lower sense of insulin, the lower glucose entrance, and its lower phosphorylation, subsequently disrupt glucose, fatty acids, lipid, and amino acids metabolisms.



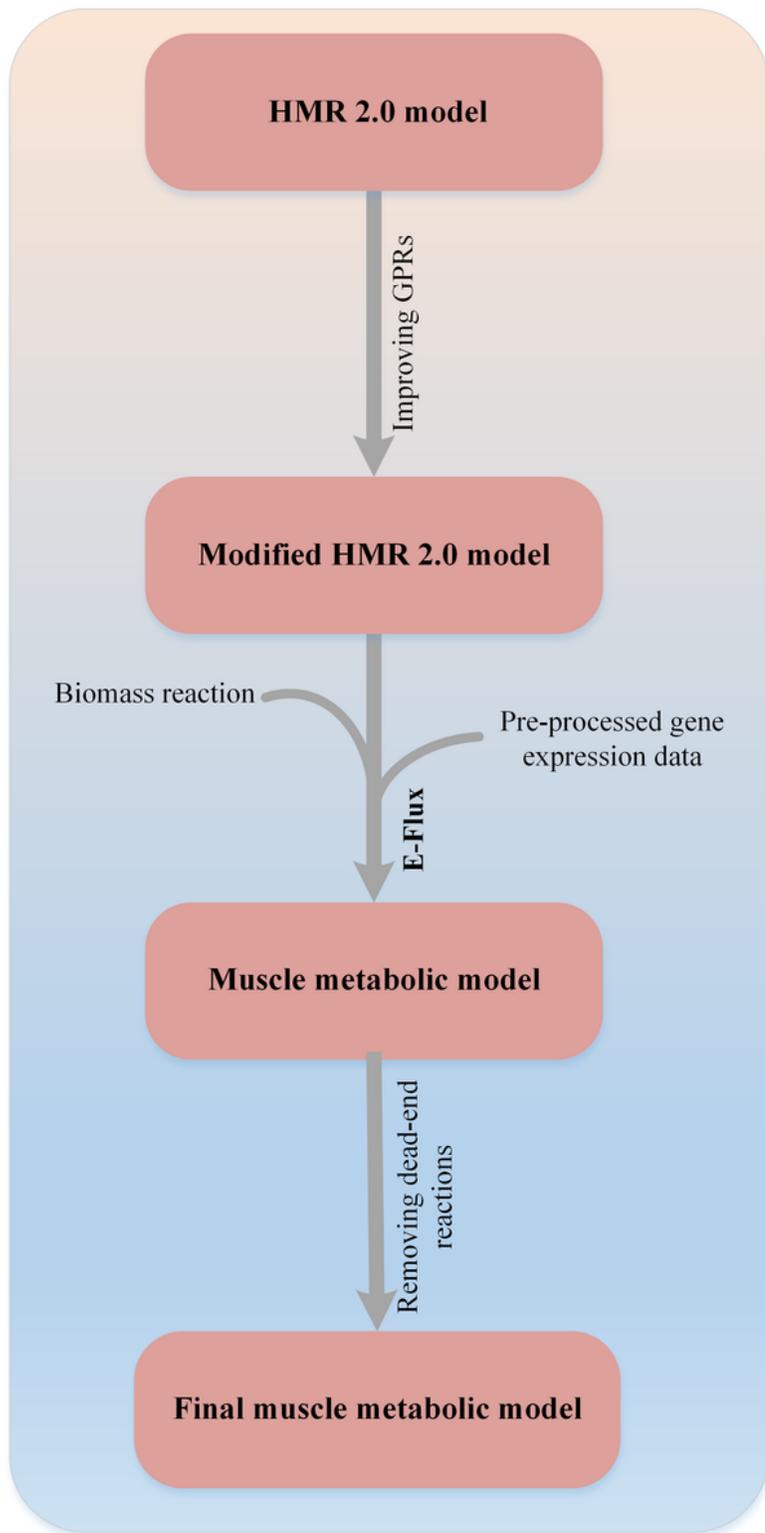
**Figure 4**

Schematic illustration of metabolism dysregulation in T2DM. Our analysis demonstrated that dysmetabolism of those metabolites implicated in extracellular matrix and insulin signaling possibly leads to a lower sense of insulin and subsequent lower expression of insulin-responsive genes such as HK2. The lower sense of insulin, the lower glucose entrance, and its lower phosphorylation, subsequently disrupt glucose, fatty acids, lipid, and amino acids metabolisms.



**Figure 5**

The workflow of muscle metabolic model reconstruction. Using gene expression data and the E-Flux algorithm the muscle metabolic model was reconstructed. To have a functional model, biomass reaction was added and dead-end reactions were removed.



**Figure 5**

The workflow of muscle metabolic model reconstruction. Using gene expression data and the E-Flux algorithm the muscle metabolic model was reconstructed. To have a functional model, biomass reaction was added and dead-end reactions were removed.

## Supplementary Files

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

- [AdditionalFile1.xlsx](#)
- [AdditionalFile1.xlsx](#)