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.

Study on the mechanism of type 2 diabetes mellitus on tendon repair based on bioinformatics analysis combined with machine learning models

Shulong Sun

dr.sun2012@hotmail.com

Henan University of Chinese Medicine Henan Province Hospital of Traditional Chinese Medicine

Hua Li

Henan University of Chinese Medicine Henan Province Hospital of Traditional Chinese Medicine

Guanghui Li

Henan University of Chinese Medicine Henan Province Hospital of Traditional Chinese Medicine

Yan Li

Henan University of Chinese Medicine Henan Province Hospital of Traditional Chinese Medicine

Liubing Yang

Henan University of Chinese Medicine Henan Province Hospital of Traditional Chinese Medicine

Yujing Cao

Henan University of Chinese Medicine Henan Province Hospital of Traditional Chinese Medicine

Ji Li

Hebei University of Chinese Medicine

Article

Keywords: tendon injury, T2DM, hub genes, mendelian randomization, immunocytes, signaling pathways, machine learning

Posted Date: August 31st, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9127287/v1

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

Additional Declarations:
No competing interests reported.

Tables 1 to 24 are available in the Supplementary Files section.

Study on the mechanism of type 2 diabetes mellitus on tendon repair based on bioinformatics analysis combined with machine learning models

Shulong Sun[1,2*], Hua Li[1], Guanghui Li[1], Yan Li[1], Liubing Yang[1], Yujing Cao[1,3*] and Ji Li[2*]

[1]Department of Emergency Trauma Center, Henan University of Chinese Medicine Henan Province Hospital of Traditional Chinese Medicine, Zhengzhou, China

[2]School of Basic Medicine, Hebei University of Chinese Medicine, Hebei, China

[3]College of Orthopedics and Traumatology, Henan University of Chinese Medicine, Zhengzhou, China

*** Correspondence:**

Shulong Sun
dr.sun2012@hotmail.com

Yujing Cao
bravecao@163.com

Ji Li
lijih1jfj@126.com

Abstract

Type 2 diabetes mellitus (T2DM) impairs tendon healing, increasing re-rupture risk. To understand the underlying mechanisms, we used Mendelian randomization (MR) to assess its causal link with tendon repair, and integrated bioinformatics with machine learning to identify the hub genes, pathways, and immune features. To evaluate the causal effect of T2DM on tendon repair, we first conducted a two-sample MR analysis using GWAS data. We subsequently downloaded T2DM (GSE15932, GSE25724, GSE29221) and tendon injury (GSE26051)

datasets from GEO. Common genes were identified via DEG analysis and WGCNA. These genes were then analyzed using GO/KEGG enrichment, followed by hub gene selection through a PPI network and seven machine learning algorithms. Finally, immune infiltration and GSEA were performed to investigate the biological functions of the hub genes and their relationship with the immune microenvironment. MR analysis confirmed T2DM as an adverse factor for tendon repair (IVW: OR = 1.00113, P = 0.00549). Integrating DEG and WGCNA analyses identified 15 common genes associated with both conditions. GO/KEGG enrichment showed these genes were involved in muscle structure/function, including cytoskeleton in muscle cells and motor proteins pathways. Cross-validation via PPI network and machine learning yielded three hub genes: MYL2, MYL3, and TTN. Immune infiltration analysis revealed significant macrophage changes (M2 and M0) in T2DM and tendon injury, with hub genes correlating with various immune cells (e.g., Tregs, resting mast cells, M0 macrophages). GSEA further validated hub gene involvement in muscle development, myofilament sliding, and striated muscle contraction. This study provides genetic evidence linking T2DM to impaired tendon repair and identifies MYL2, MYL3, and TTN as key hub genes. T2DM may disrupt tendon healing by influencing these genes, leading to cytoskeletal and immune dysfunction, especially involving macrophages. These findings reveal new molecular insights into diabetic tendinopathy and provide a theoretical basis for future therapies.

Keywords: tendon injury[1], T2DM[2], hub genes[3], mendelian randomization[4], immunocytes[3], signaling pathways[4], machine learning[5]

Introduction

The World Health Organization (WHO) defines diabetes mellitus as a metabolic disorder mediated by multiple etiological factors. It primarily involves abnormalities in insulin secretion and insulin action, leading to disturbances in the metabolism of carbohydrates, lipids, and proteins, ultimately characterized by elevated serum glucose levels[1]. Currently, the incidence of diabetes presents a concerning trend. According to the International Diabetes Federation, the global prevalence of diabetes is projected to rise from 537 million in 2021 to 783 million by 2045[2]. Among these cases, type 2 diabetes mellitus (T2DM) accounts for approximately 90% to 95%, making it the most prevalent of the three major chronic metabolic diseases[3]. Current research has elucidated several mechanisms underlying T2DM, including β-cell dysfunction, insulin resistance, and impairments in the incretin signaling pathway[4]. T2DM is one of the most common chronic metabolic diseases in clinical practice and is associated with a range of complications affecting multiple organ systems[5,6], including chronic kidney disease, cardiovascular and cerebrovascular diseases, and tendinopathy[7]. Clinical evidence has established a link between T2DM and musculoskeletal disorders, with a particular emphasis on tendon pathology[8]. A frequently observed complication in this context is impaired tissue healing, characterized by delayed or defective repair processes, which result from inadequate growth factor production, compromised angiogenesis, and deficient collagen matrix formation[9].

Tendons are predominantly composed of tenocytes and extracellular matrix rich in collagen, and serve to transmit the tensile forces generated by muscle contraction to the skeletal system, thereby enabling joint movement and locomotion[10,11]. Tendon injuries (TI) represent a prevalent clinical entity associated with substantial morbidity, including pain, functional disability, elevated healthcare expenditures, and reduced work productivity, rendering their management a persistent therapeutic challenge[12]. Acute laceration constitutes the most frequent

etiology of TI, with the highest incidence reported among individuals aged 20 to 29 years and a notable male predominance[13]. Accumulating evidence suggests that T2DM adversely affects the biomechanical properties of injured tendons—such as the Achilles tendon and rotator cuff tendons—and impairs the subsequent healing response[14-16]. The healing process following tendon injury is conventionally divided into three overlapping phases: the inflammatory phase, the proliferative phase, and the remodeling phase[17]. The inflammatory phase is initiated by hematoma formation at the injury site, accompanied by the recruitment of immune cells that clear cellular debris, necrotic tissue, and foreign material via phagocytosis[18-20]. Concurrently, increased deposition of fibronectin, glycosaminoglycans, water, and collagen type III contributes to the stabilization of the local extracellular matrix[21]. A key event during this phase is the recruitment and subsequent proliferation of tenocytes within the injured region[17,22], which facilitates the transition to the proliferative phase. The proliferative phase is characterized by the synthesis of cytokines and growth factors, primarily mediated by macrophages and myofibroblasts. Macrophages secrete growth factors that promote cellular activity, while tenocytes deposit a matrix rich in collagen type III, the production of which peaks during this stage[17,20,21]. The process then progresses into the remodeling phase, marked by a decline in collagen type III synthesis and a shift toward predominant collagen type I production[17,23,24]. This phase is associated with reduced cellularity and diminished matrix synthesis, alongside progressive organization and alignment of the extracellular matrix[18,25]. Collagen type I fibers become oriented longitudinally along the tensile axis of the tendon, thereby restoring the mechanical strength of the regenerating tissue[18]. Over time, the repaired fibrous tissue undergoes maturation into scar-like tissue[20]. In addition to these cellular events, tendon repair involves specialized progenitor populations known as tendon-derived stem cells (TDSCs). Like other stem cells, TDSCs exhibit characteristic properties of stemness, including self-renewal and multilineage differentiation capacity[26]. These cells are thought to contribute to tendon repair and regeneration and may play a role in maintaining tendon homeostasis[27,28].

However, tendons typically exhibit a prolonged healing trajectory relative to other tissues, a phenomenon largely attributable to their hypovascular nature and inherently slow regenerative capacity[29]. Although intrinsic tendon healing often yields suboptimal functional outcomes, this process is further exacerbated by T2DM, which adversely modulates the natural fibrotic healing response and thereby substantially impedes recovery[30,31]. In the context of clinically prevalent rotator cuff injuries, T2DM is associated with a markedly elevated risk of impaired healing and postoperative repair failure[32,33]. Similarly, patients with T2DM demonstrate compromised healing capacity following Achilles tendon injury[34]. Accumulating evidence indicates that the deleterious effects of T2DM on tendon repair are particularly pronounced during the early phases of healing[35]. Nevertheless, the precise molecular mechanisms through which T2DM compromises tendon injury repair remain poorly understood, and relevant mechanistic investigations are scarce.

Tendons are dense, highly ordered fibrous connective tissues composed primarily of type I collagen, along with other collagen types and glycoproteins, organized within a structurally optimized extracellular environment. Accumulating evidence indicates that diabetes mellitus adversely affects the properties of native connective tissues[36]. Specifically, diabetes-induced alterations include reduced collagen fibril diameter, morphological abnormalities, increased fibrillar disorganization and packing irregularities, as well as overall tendon thickening and calcification[37]. A high-resolution ultrasound study reported that 84.6% of patients with T2DM exhibited tendon fiber disorganization after more than ten years of follow-up since diagnosis[38], suggesting that T2DM fundamentally alters intratendinous microarchitecture. Clinical research further demonstrates that hyperglycemia suppresses collagen synthesis and promotes the accumulation of advanced

glycation end-products (AGEs), leading to ultrastructural abnormalities in collagen fibrils[39-41]. T2DM also disrupts tendon homeostasis, thereby impairing its reparative capacity. The tendon's internal milieu is primarily composed of the ECM, a collagen-rich network essential for structural integrity. Studies have shown that AGEs formed under hyperglycemic conditions pathologically disrupt ECM homeostasis and compromise the tendon healing process[42]. Additionally, T2DM induces systemic low-grade inflammation, which significantly perturbs tendon homeostasis. Tendons from diabetic individuals often exhibit fibrotic changes and exacerbated ECM disorganization[43], although the severity of involvement may vary across anatomical sites[14,44]. Beyond its impact on homeostasis, T2DM markedly impairs the tendon healing response to injury[44,45]. Following tendon injury, T2DM compromises the mechanical properties of healed tissue by reducing collagen fiber organization[14,46-48]. These impairments in the healing cascade increase the susceptibility to re-rupture following repair. Tendon gliding function is also frequently compromised due to the formation of peritendinous fibrous adhesions, which restrict joint range of motion and consequently impair overall limb function[48]. Moreover, tendons in patients with T2DM appear particularly susceptible to degenerative changes[49], with structural alterations contributing to an elevated risk of traumatic tendon rupture in approximately 40% of affected individuals[50].

Despite extensive research into the structural, compositional, biomechanical, and histopathological alterations of tendons in both clinical and animal models, effective therapeutic and rehabilitative interventions for diabetic tendinopathy remain elusive[51]. This therapeutic gap largely stems from an incomplete understanding of the pathological and molecular mechanisms through which T2DM impairs tendon repair[52,53]. Individuals with T2DM are at a notably increased risk of developing tendinopathy[36,54], with approximately 30–60% suffering from persistent consequences of impaired tendon healing—including chronic pain, adhesions, and functional limitations—that can culminate in severe disability[49]. These findings highlight a critical need to decipher the underlying cellular and molecular mechanisms to inform the development of targeted therapeutic strategies. However, the cellular milieu during tendon repair is inherently complex and dynamic, transitioning through the inflammatory, proliferative, matrix deposition, and remodeling phases. This complexity underscores that the impact of T2DM on tendon healing involves multifactorial interactions—encompassing both cellular and biomechanical determinants—thereby posing substantial challenges to effective clinical management.

To investigate the causal relationship between T2DM and tendon repair, we first conducted a two-sample Mendelian randomization (MR) analysis. By leveraging genetic variants as instrumental variables for modifiable exposures, MR provides robust estimates of the causal effect of an exposure (i.e., a potential risk factor) on an outcome, thereby offering clearer insights into etiological pathways[55]. Building on this, we sought to explore the molecular mechanisms by which T2DM influences tendon repair. Specifically, we integrated bioinformatics analysis with machine learning approaches to identify hub genes and signaling pathways involved in this process. We retrieved T2DM- and tendon injury (TI)-related datasets from the Gene Expression Omnibus (GEO) database and performed differential expression gene (DEG) analysis. To pinpoint genes associated with the T2DM and insulin resistance phenotypes, we applied weighted gene co-expression network analysis (WGCNA), a systems biology method that captures co-expression patterns and gene-phenotype relationships across samples[56]. Hub genes were further prioritized through integration of protein-protein interaction (PPI) network analysis and seven machine learning algorithms. In addition, we conducted an immune infiltration analysis to characterize the immune microenvironment associated with T2DM-impaired tendon repair. Finally, gene set enrichment analysis (GSEA) was performed to explore the biological functions and pathways

involving the identified hub genes. GSEA is a statistical method for analyzing whole-genome expression profiling microarray data. It ranks genes based on their differential expression from a predefined gene set, providing deeper insights into dysregulated biological processes under specific phenotypes or disease conditions[57]. Together, this integrative approach aims to systematically elucidate the mechanisms underlying the detrimental effects of T2DM on tendon repair and to provide a theoretical foundation for improved clinical management and therapeutic strategies.

Materials and methods

1. Mendelian randomization analysis

To explore the causal relationship between T2DM and tendon repair, we employed a two-sample mendelian randomization (MR) analysis approach. We treat T2DM as the exposure factor and use single nucleotide polymorphisms (SNPs) significantly associated with T2DM as instrumental variables (IVs), with tendon repair as the outcome variable. Data related to T2DM and tendon repair are sourced from the Genome-Wide Association Study (GWAS) database (https://gwas.mrcieu.ac.uk/)[58]. The criteria for data selection follow these principles: SNPs are highly correlated with the exposure factor, unaffected by confounding factors, and can only influence the results through the exposure factor[59]. All statistical analyses were performed in R software (version 4.2.3) using the "TwoSampleMR" package[60]. To ensure robust causal inference, we use five methods: MR-Egger, Weighted Median, IVW, Simple Mode, and Weighted Mode[61,62]. Among them, the IVW method does not require individual-level data. When horizontal pleiotropy is balanced, the IVW method can produce unbiased estimates[63]. Therefore, we adopted it as the primary approach for assessing the causal effect of exposure on the outcome. The MR-Egger test is used to assess potential pleiotropic effects. An intercept of 0 indicates no horizontal pleiotropy[64]. Additionally, a leave-one-out (LOO) analysis is performed to identify potential outliers. By removing each SNP one at a time, the overall effect is evaluated[65]. Generally, results with a P-value < 0.05 are considered statistically significant.

2. Comprehensive bioinformatics analysis based on the GEO database

2.1 Data sources and screening criteria

The datasets involved in this research are all sourced from the NCBI GEO (Gene Expression Omnibus) database (https://www.ncbi.nlm.nih.gov/gds/)[66]. For the selection of datasets related to tendon injury and T2DM, we generally adhere to the following inclusion criteria: (1) The dataset focuses on Homo sapiens as the research subject, and the data type is an array expression profile. (2) All datasets used in this study are publicly accessible[67,68].

2.2 Data preprocessing

We used R software (version 4.2.2, Team RC, R: A language and environment for statistical computing. R foundation for statistical computing, 2025, https://www.r-project.org/) to process and merge the dataset. The main steps are as follows: (1) Download a series of matrix files and corresponding platform information files from the GEO database. (2) In the R programming environment, use the 'GEOquery' package to obtain the platform annotation file[69], the 'dplyr' and 'tidyr' packages to merge gene IDs[70], and the 'limma' package to remove duplicates and NA values[71], thereby obtaining a clean matrix data file. (3) Use the 'inner_join' function from the R package 'dplyr' to merge the cleaned matrix data, resulting in the merged matrix dataset[72]. (4) Use the 'removeBatchEffect' function to remove batch effects from the dataset, and generate a box plot using the 'boxplot' function. Additionally, perform principal component analysis (PCA) using the R packages 'FactoMineR', 'Stats', and 'Factoextra' to examine the heterogeneity of the merged dataset, and construct a clustering tree of the merged dataset using the 'hclust' function[73].

2.3 DEGs screening

We used R software to filter differentially expressed genes (DEGs) from the processed T2DM and TI-related matrix files. First, we analyzed and compared the matrices using the 'limma' and 'lmtest' R packages. Then, DEGs were obtained by setting the threshold $|\log2 \text{ fold change}| \geq 0.5$ and p-value < 0.05[74]. Upregulated and downregulated genes were selected, and volcano plots and waterfall plots related to DEGs were drawn using the R packages 'ggplot2', 'ggfun', 'Seurat', 'RColorBrewer', 'patchwork', and 'aggregale'[75,76]. Finally, the top 20 genes were further filtered, and a heatmap was generated using the 'heatmap' R package[77].

2.4 WGCNA analysis and gene module construction

To further explore gene modules associated with T2DM and TI phenotypes, we used Weighted Gene Co-expression Network Analysis (WGCNA). We used the 'WGCNA' package in the R environment to identify candidate core gene modules. First, we processed the data for T2DM and TI separately to check for abnormal samples[78]. Then, by setting the fitting index of the scale-free topology model, the most appropriate soft-thresholding power was calculated using the 'pickSoftThreshold' function in the 'WGCNA' package, and the topological overlap matrix (TOM) was subsequently constructed[79,80]. Finally, hierarchical clustering is used to construct a dendrogram, and dynamic tree cutting based on the correlations in TOM is performed to identify highly co-expressed gene modules[81,82].

3. GO/KEGG analysis of common genes between T2DM and TI

First, we intersected the DEGs with the gene modules from the WGCNA analysis related to T2DM and TI obtained in the early stages to identify common genes, and present the results using a Venn diagram created with the online software Venny 2.0.2 (https://bioinfogp.cnb.csic.es/tools/venny/index2.0.2.html)[83]. Then, we used Cytoscape 3.9.1 software (https://cytoscape.org/) to construct a network diagram of the common genes[84]. Finally, we conducted Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses on the common genes[85,86]. Based on the Metascape database (https://metascape.org/gp/index.html#/main/step1), we conducted a "custom analysis" on common genes, specifying the species "Homo

sapiens"[87]. The parameters are set to minimum overlap = 3, *p*-value cutoff = 0.01, minimum enrichment = 1.5. The final results were displayed in the form of bar charts, circle charts, and sankey diagrams using the online graphing software Bioinformatics (https://www.bioinformatics.com.cn/)[88].

4. Construction of the PPI network to identify hub genes

First, we imported the common genes into the STRING database (https://cn.string-db.org/) and constructed the PPI network using the following parameter set: (1) Search type: multiple proteins; (2) Minimum required interaction score: high confidence (0.700)[89]. Then, we further optimized the PPI network using Cytoscape 3.9.1. Next, we analyzed the degree, closeness, betweenness, and MCC values of the common genes using the CytoHubba plugin[90,91]. Finally, from the common genes, we selected the top 5 genes identified by each method and further determined their intersection to obtain the hub genes.

5. Construction of machine learning models for screening hub genes

In the R programming environment, we used R packages (mainly including 'glmnet', 'DALEX', 'randomForest', and 'e1071') to build seven machine learning models for further screening of common genes to identify hub genes. The key steps in model construction and optimization are as follows: (1) In the LASSO model, we used the R package 'glmnet' to optimize the lambda (λ) value through 10-fold cross-validation, selecting the optimal λ and identifying the candidate hub genes from coefficients[92]. (2) We developed a Random Forest (RF) model using the 'randomForest' R package, constructing 500 trees by minimizing the overall error rate of the dataset[93]. We ranked the importance of genes using the Gini coefficient and selected the top 5 genes[94]. (3) The Support Vector Machine (SVM) model identifies feature genes through linear prediction. These are ranked based on the ratio of the mean squared error to the square of the standard deviation. The model is built using the 'e1071' R package, and to enhance robustness, we employ k = 10 subsampling to select feature genes[95]. (4) The K-Nearest Neighbors (KNN) algorithm is a non-parametric classification algorithm based on supervised learning[96]. We use the R package 'class' to find the nearest neighbors (K = 5) to the target point for classification decisions[97]. (5) We constructed a neural network (NNET) model using the 'nnet' R package, employing the backpropagation algorithm for model training to obtain the core genes[98,99]. (6) A generalized linear model (GLM) was constructed using the R package 'lme4' to screen for hub genes[100]. (7) Decision Tree (DT) is a classification algorithm in supervised learning. The model is built using the 'rpart' package[101], with the 'caret' R package used to create a training set (70%) and a test set (30%). The model is then fitted to the training set to identify hub genes[102]. To gain a deeper understanding of the performance of these models, we use the 'DALEX' package to visualize the residual distributions of the seven methods. Through this analysis, we identify the most effective machine learning models by prioritizing the minimization of residuals[103-105]. The genes from machine learning models are crossbred to obtain the hub genes and create a Venn diagram[106].

6. Immune infiltration analysis of T2DM and TI related expressed genes

Using the CIBERSORT (v.1.0) package developed in the R language, we analyzed the infiltration levels of 22 types of immune cells in the expressed genes related to T2DM and TI[107]. To assess the differences in immune cell proportions, the Wilcoxon rank-sum test was applied, with $P < 0.05$ set as the threshold for statistical significance[108]. First, the data were processed using the R package 'dplyr'. Then, the packages 'ggplot2', 'ggpubr', and 'ggthemes' were used to create stacked bar charts and heatmaps to display the differences in immune cell proportions across different groups in T2DM and TI[109]. In order to provide a more intuitive display of the infiltration levels of different immune cells, we used the 'RColorBrewer', 'reshape2', and 'doParallel' packages to construct box plots ranking the levels of immune cell infiltration associated with T2DM and TI, respectively[67]. Finally, a correlation analysis was conducted using the 'tidyr' and 'ggcorrplot' packages to explore the relationship between hub genes and 22 immune cell types, and a heatmap was generated[110]. Additionally, butterfly plots of correlations were created using 'linkET', 'ggplot2', and 'vegan'[111].

7. GSEA analysis of hub genes

To further validate the biological role of the selected hub genes, we performed a GSEA analysis on the genes using the 'clusterProfiler (version 4.0)' package in the R environment[112]. We first downloaded the dataset related to GO enrichment analysis from the GSEA (https://www.gsea-msigdb.org/gsea/index.jsp) database, specifically C5. It includes a total of 10,480 gene sets across the BP, CC, and MF categories (https://www.gsea-msigdb.org/gsea/msigdb/human/collections.jsp#C5). The genes were ranked based on their differential expression compared to predefined gene sets. Finally, we set the parameter pvalueCutoff = 0.05, and then used the R packages 'ggplot2' and 'enrichplot' to further visualize the results obtained[113].

Results

1. Mendelian randomization

We obtained the T2DM-related dataset bbj-a-153 (https://opengwas.io/datasets/bbj-a-153#) and the tendon repair-related dataset ukb-d-M13_SYNOTEND (https://opengwas.io/datasets/ukb-d-M13_SYNOTEND) from the GWAS database. The bbj-a-153 dataset contains 210,865 samples and a total of 8,885,694 SNPs. On the other hand, the ukb-d-M13_SYNOTEND dataset includes 361,194 samples and 12,330,288 SNPs. In the R language environment, we filtered genetic variation loci strongly associated with the diseases by setting a threshold of $P < 0.05$ and plotted linear Manhattan diagrams (Figure 1a and b). The results show that the vertical axis (-log10 *P*-value) represents the logarithmic transformation values with statistical significance. Each point represents a SNP, with its position and significance level labeled on the graph. SNPs above the red line indicate significant associations. We further performed clustering by setting the thresholds p1 = 5e-08, $r^2 = 0.001$, and window size = 1Mb to eliminate linkage disequilibrium, identifying 114 SNPs associated with T2DM and 110 SNPs associated with tendon repair (Figures 1c and 1d).

All selected SNPs were confirmed as robust instruments and used as IVs for subsequent analyses (Tables 1 and 2). Next, we used these SNPs to analyze the causal relationship between T2DM and tendon repair using five analytical methods. The results showed that IVW (OR

= 1.00113, 95% CI: 1.00034–1.001939, *P* = 0.00549); weighted median (OR = 1.001647, 95% CI: 1.000351–1.002945, *P* = 0.012759); MR-Egger (OR = 1.00123, 95% CI: 0.999512–1.002951, *P* = 0.163632) (Table 3). The forest plot depicts the risk value of tendon repair mediated by a single SNP. The results show that, overall, both red dotted lines are to the right of 0, indicating that both methods (IVW and Egger) concluded that T2DM is a detrimental factor for tendon repair (Figure 1e). The funnel plot shows that the causal effects obtained from each IV are symmetrically distributed, indicating that the results are unlikely to be affected by potential bias and exhibit strong stability (Figure 1f). LOO analysis shows that after removing each SNP, all the error lines lie to the right of 0, suggesting that the causal relationship remains statistically significant, suggesting that the results are reliable (Figure 1g). This finding further validates the reliability of the results. The scatter plot displays the MR analysis results between exposure factor and outcome factor (Figure 1h). Each point in the figure represents an IV, and the line on each point reflects the 95% CI. The horizontal axis represents the effect of the SNP on the exposure factor (T2DM), while the vertical axis represents the effect of the SNP on the outcome factor (tendon repair). The ratio of these two effects, that is, the effect of the exposure on the outcome, corresponds to the slope of the colored lines in the figure. Different colored lines indicate different algorithms. The results show that, overall, the lines of different algorithms slope upwards, which suggests that T2DM is an adverse factor for tendon repair.

2. Bioinformatics Analysis

2.1 Dataset download and preprocessing

The MeSH strategy was used to determine the disease names as tendon injury and type 2 diabetes mellitus for data retrieval[114,115]. According to the selection criteria, we included the T2DM-related datasets GSE15932, GSE25724, and GSE29221 for this study. Among them, the GSE25724 dataset contains islet expression levels from 6 patients with T2DM and 7 healthy individuals[116]. The GSE29221 dataset comprises gene expression profiles of skeletal muscle from 12 patients with T2DM and 12 healthy individuals[117]. As for the GSE15932 dataset, it includes samples from patients with pancreatic cancer combined with diabetes as well as those with pancreatic cancer alone[118]. Therefore, we excluded the unrelated samples and ultimately obtained data from 8 patients with T2DM and 8 healthy individuals. Similarly, based on the selection criteria, we included the tendon injury-related dataset GSE26051 in this study. This dataset contains gene expression profile data from 23 patients with tendon injuries and 23 healthy tendon tissue samples[119]. We downloaded the matrix files of the dataset and the corresponding platform information files from the GEO database, including annotations related to T2DM for GPL570, GPL96, and GPL6947, as well as annotations related to TI for GPL570. Data preprocessing and integration were performed using R software (version 4.2.2) to obtain a clean matrix file (Tables 4-8). We normalized the dataset, conducted heterogeneity detection, and removed batch effects from the matrix data. These procedures are illustrated by the boxplot, clustering dendrogram, and PCA plot. Specifically, the handling of T2DM-related data and the results before and after merging indicate that no batch effects were observed after data integration, demonstrating comparability between groups in the dataset (Figure 2a, b, and c). Similarly, the processing of TI-related data before and after integration showed no batch effects, indicating comparability between groups in the dataset (Figure 2d, e, and f).

2.2 Identification of DEGs associated with TI and T2DM

First, we separately screened the differentially expressed genes (DEGs) from the gene expression matrix data of the T2DM and TI-related samples. The analysis showed that in the T2DM data, there were a total of 57 differentially expressed genes (42 upregulated, 15 downregulated) compared to the normal group (Figure 3a and Table 9). The TI-related analysis revealed that, compared to the normal group, the TI group had 721 differentially expressed genes (460 upregulated, 261 downregulated) (Figure 3d and Table 10). Then, a waterfall plot was used to further present the top 10 ranked genes, including the upregulated genes CXCL10, CCL5, PYGM, PPP1R3B and MYL2, and the downregulated genes CCDC89, GAB3, MARK2, KIAA0363 and LOC644615 in T2DM samples (Figure 3b). Similarly, in TI samples, the upregulated genes MYL2, CSRP3, ACTA1, PPP1R3A and MYL3, and the downregulated genes RP11-803D5.4, CTHRC1, SPP1, TNFSF11, and TREM1 were also identified (Figure 3e). Finally, we ranked the differentially expressed genes (DEGs) based on gene expression levels and visualized the top 20 genes from each sample in the T2DM dataset using a heatmap (Figure 3c). Similarly, the top 20 genes were extracted from the TI dataset and visualized using a heatmap (Figure 3f).

2.3 WGCNA analysis and gene module construction related to T2DM

First, we processed the merged T2DM dataset using the 'WGCNA' package in R, and no outlier samples were found (Figure 4a). Subsequently, the 'pickSoftThreshold' function was used to analyze the network topology under different soft-threshold powers. The results revealed that when the soft threshold was set to 9, the network approximated a scale-free topology, and its average connectivity tended to stabilize (Figure 4b). Then, based on the TOM (topological overlap matrix) dissimilarity metric, all candidate genes were clustered into similar modules. Using the dynamic tree cutting method (minModuleSize = 30 and mergeCutHeight = 0.25), 6,107 candidate genes were divided into 12 modules (Figure 4c and d). By calculating the dissimilarity between feature genes and module features, similar modules are clustered. Using the 'merge' function, 400 genes were selected, and a gene clustering network heatmap based on TOM similarity was plotted (Figure 4e). The darker the color, the stronger the connection between genes. Finally, we selected two modules with high correlation coefficients from the 12 modules, including the blue module (coefficient = 0.57, *P*-value = 0.004) and the turquoise module (coefficient = 0.46, *P*-value = 0.02) (Figure 4f and Tables 11-12). We combined the two gene modules, resulting in a total of 2124 genes.

2.4 WGCNA analysis and gene module construction related to TI

We used the 'WGCNA' package to process the TI dataset and found no abnormal samples (Figure 5a). The 'pickSoftThreshold' function was used to select the optimal soft threshold. It was found that when the soft threshold = 4, the network approximated a scale-free topology, with its average connectivity stabilizing (Figure 5b). Based on TOM dissimilarity metrics, all candidate genes are clustered into similar modules. Similarly, using the dynamic tree cutting method (minModuleSize = 30 and mergeCutHeight = 0.25), 23,323 candidate genes are divided into 24 modules (Figures 5c and d). We clustered similar modules, selected 400 genes, and created a gene clustering network heatmap (Figure 5e). The darker the color, the stronger the connection between the genes. We selected 5 modules with high correlation coefficients from the 24 modules, including the brown module (coefficient = 0.42, *P*-value = 0.004), the magenta module (coefficient = 0.39, *P*-value = 0.008), the pink module (coefficient = 0.48, *P*-value = 7e-04), the blue module (coefficient = 0.37, *P*-value = 0.01), and the grey module (coefficient = 0.38, *P*-value = 0.009) (Figure 5f and Tables 13-17). We combined the 5 gene modules, resulting in a total of 2617 genes.

3. GO/KEGG analysis

By intersecting the differentially expressed genes (DEGs) related to T2DM and TI with the genes obtained from the weighted gene co-expression network analysis (WGCNA) related to T2DM and TI, a total of 15 common genes were identified. The results are presented in Venn diagrams and network plots (Figure 6a and b). Subsequently, we performed GO analysis on 15 common genes and found that these genes are primarily distributed in the myofibril region of muscle fibers, as well as in the sarcomere and the I and A band cellular components. These genes cover the stability of muscle structure, the binding of actin and myosin, and the composition of the cytoskeletal structure. Moreover, these genes are involved in myofibril assembly, muscle organ development, regulation of striated muscle contraction, and skeletal muscle organ development (Table 18). The main results of each GO analysis are presented in the form of circular and bar charts (Figure 6c, d). Through KEGG enrichment analysis, three signaling pathways were identified. These primarily include the cytoskeleton in muscle cells, the hypertrophic cardiomyopathy, and the motor proteins signaling pathways (Table 19). The results are displayed using a Sankey diagram to illustrate the relationships between pathways and targets (Figure 6e).

4. Identification of hub genes via PPI network

We imported the 15 common genes obtained from the initial DEGs and WGCNA screening into the STRING database to construct the PPI network. The results showed 15 nodes and 45 edges, with an average node degree = 6, a clustering coefficient = 0.728, and a PPI enrichment *p*-value < 1.0e-16 (Figure 7a). Through the PPI network, we were able to visualize the relationships between the genes. Next, using the CytoHubba plugin in Cytoscape, we calculated the degree, MCC, closeness, and betweenness values for each common gene (Figure 7b). The top 5 genes were selected for further analysis to obtain hub genes, which included: MYL2, MYL3, TTN and ACTN2 (Figure 7c).

5. Identification of hub genes via machine learning

First, the performance of the seven models is evaluated by comparing their residual distributions. Overall, based on the reverse cumulative distribution of the residuals, we found that the performance of the seven models is similar (Figure 8a). The GLM model shows the highest proportion of minimum residual values in the left tail of the residual distribution. In the right tail of the residual distribution, the SVM model shows the lowest proportion of maximum residual values. However, when considering the impact of median absolute and root mean square of residuals on model performance, we found that the RF, SVM, and LASSO models outperformed the others (Figure 8b). At the same time, we observed that the residuals of the RF, SVM, and LASSO models were more stable, demonstrating more robust performance (Figure 8c). We also evaluated the feature importance of the seven models to understand which predictive factors have a greater impact on the outcome variable. Overall, the RF and SVM models have the lowest root mean square error (RMSE) (Figure 8d). Through the LASSO regression algorithm, we identified 8 central genes based on the optimal λ value (Figure 8e and Table 20). Using the SVM method, we determined the top 5 most significant genes based on the lowest classification error rate and average ranking (Figure 8f and Table 21). Next, we constructed 500 random forests to rank the genes and extracted the top 7 most important genes (Figures 8g, h). Finally, we identified the hub genes (MYL2, MYL3 and TTN) through the intersection of genes selected by the three machine models (Figure 8i).

6. Assessment of immunocytes infiltration

First, we analyzed the immune microenvironment throughout the TI-related process and determined the precise infiltration proportions of 22 types of immune cells in each sample, as well as the degree of immune cell infiltration across different groups (Figure 9a). Additionally, we conducted a systematic evaluation of the infiltration rates of the 22 immune cell types (Figure 9b). The results revealed significant differences in the infiltration proportions of Eosinophils, M0.macrophages, and T.cells.CD8, indicating their key roles in the TI-related immune microenvironment (Figure 9c). Building upon the hub genes identified in the earlier stages, we further analyzed the relationship between these genes and 22 types of immune cells, as well as the interactions between the immune cells themselves (Figure 9d and e). The correlation analysis between 22 types of immune cells shows that mast.cells.activated are positively correlated with T.cells.follicular.helper, and macrophages.M0 are positively correlated with T.cells.CD8. However, B.cells.naive are negatively correlated with B.cells.memory, and macrophages.M0 are negatively correlated with NK.cells.activated. Infiltration studies conducted on hub genes and 22 types of immune cells revealed that MYL3 is positively correlated with regulatory T cells (Tregs), resting dendritic cells, and resting mast cells. MYL2 shows a positive correlation with resting dendritic cells and resting mast cells, while exhibiting a negative correlation with resting memory CD4 T cells and M0 macrophages. TTN is positively correlated with naïve CD4 T cells, regulatory T cells (Tregs), resting dendritic cells, and resting mast cells, but negatively correlated with M0 macrophages. Then, based on the same analytical strategy, we analyzed the immune microenvironment associated with T2DM, and the results showed significant differences in the infiltration proportions of M2 macrophages, M0 macrophages, and T.cells.CD4.memory.resting (Figure 9f, g and h). The relationship between immune cells revealed that NK.cells.resting is positively correlated with Plasma.cells, T.cells.CD4.naive is positively correlated with Eosinophils, NK.cells.activated is positively correlated with T.cells.follicular.helper, and Dendritic.cells.resting is positively correlated with T.cells.gamma.delta. The relationship between hub genes and immune cells revealed that MYL3 has a positive correlation with Macrophages.M0 and a negative correlation with T.cells.CD8. MYL2 is positively correlated with Macrophages.M0 and Eosinophils, and negatively correlated with Dendritic.cells.resting. TTN shows a positive correlation with T.cells.regulatory..Tregs. and Eosinophils, and a negative correlation with NK.cells.resting (Figure 9i and j). Finally, based on the above results, we found that the immune cell Macrophages.M0 shows significant expression in both TI and T2DM.

7. GSEA analysis of the hub genes TTN, MYL2, and MYL3

We performed a GSEA analysis on the core genes MYL2, MYL3, and TTN selected in the earlier phase, and the results revealed that TTN is primarily involved in biological processes related to muscle organ development (P = 1e−10) (Figure 10a). Furthermore, it is also associated with biological processes such as actin myosin filament sliding, skeletal muscle adaptive regulation, and skeletal myofibril assembly (Figure 10b, c and Table 22). The analysis results of MYL2 show that it primarily involves the biological process of striated muscle contraction (P = 1e−10) (Figure 10d). Furthermore, it also involves the sliding of actin myosin filaments, the assembly of skeletal muscle myofibrils, aerobic respiration, and sarcomere organization (Figure 10e, f and Table 23). The results of MYL3 show that it primarily involves the development of skeletal muscle organs (P = 1e−10) (Figure 10g). At the same time, it also involves the sliding of actin myosin filaments and the adaptive regulation of skeletal muscle (Figure 10h, i and Table 24).

Discussion

Tendons, while histologically classified as dense connective tissues, are frequently implicated in sports-related injuries owing to their critical role in mediating force transmission between muscle and bone[120]. The escalating global prevalence of type 2 diabetes mellitus (T2DM)[121] has led to a growing cohort of patients presenting with both T2DM and tendon injuries (TI), many of whom exhibit unfavorable outcomes following tendon repair. A well-established association exists between the metabolic dysregulation characteristic of T2DM and various musculoskeletal disorders. Notably, T2DM compromises both the homeostatic environment and the regenerative capacity of tendons after acute injury[122-124]. Despite this recognition, the precise mechanisms by which T2DM adversely affects tendon healing remain poorly understood. To address this knowledge gap, the present study integrated Mendelian randomization with bioinformatic analyses and machine learning-based modeling to systematically investigate the causal relationship and underlying molecular mechanisms linking T2DM to tendon injury and repair. Our findings provide converging evidence at both the genetic-epidemiological and molecular levels, indicating that T2DM serves as a detrimental factor for tendon repair. Furthermore, we identify hub genes and characterize potential immune microenvironment features that may mediate this association.

First, the MR analysis provided genetic evidence supporting a causal relationship between T2DM and tendon healing. Using large-scale GWAS data, we selected independent genetic variants significantly associated with T2DM as instrumental variables. The results from multiple MR methods showed that both the IVW and weighted median approaches supported a significant positive causal effect of T2DM on tendon healing, suggesting that T2DM may increase the risk of tendon injury or adversely affect the repair process. Although the MR-Egger method yielded a slightly wider confidence interval and did not reach statistical significance, its effect direction was consistent with that of IVW. Sensitivity analyses (leave-one-out) and funnel plots indicated that the findings were robust and not substantially influenced by potential horizontal pleiotropy. These results align with clinical observations that T2DM patients often present with tendinopathy, characterized by tendon thickening, impaired mechanical properties, and reduced healing capacity[37,125]. Moreover, clinical evidence shows that T2DM patients exhibit decreased tendon mobility and a significantly higher risk of re-tear or rupture following repair[126]. Research indicates that T2DM interferes with tendon healing by disrupting tendon homeostasis, a pathological condition termed diabetic tendinopathy[127-129]. Its typical manifestations include negative effects on tendon function, such as restricted range of motion, and impaired structural remodeling, as evidenced by disorganized collagen architecture leading to tendon thickening and calcification[7,130,131].

To investigate the molecular mechanisms underlying the effect of T2DM on tendon repair, we integrated transcriptomic data related to T2DM and tendon injury from the GEO database and performed WGCNA. By intersecting the DEGs with the key module genes identified through WGCNA, we selected 15 core genes commonly involved in both T2DM and tendon injury. GO enrichment analysis revealed that these genes are primarily localized in cellular components such as myofibrils and sarcomeres, and are implicated in biological processes including myofibril assembly, maintenance of muscle structure stability, regulation of striated muscle contraction, actin-myosin binding, skeletal muscle organ development, and cytoskeleton organization. As tendons are dense connective tissues that link muscles to bones, their functional integrity is closely associated with muscle health. Metabolic disturbances in muscles may compromise tendon homeostasis through mechanical transduction or paracrine signaling[132]. KEGG enrichment analysis identified three signaling pathways: cytoskeleton in

muscle cells, hypertrophic cardiomyopathy, and motor proteins. These findings suggest that T2DM may impair tendons through two interrelated mechanisms: directly by disrupting tendon tissue stability (affecting the extracellular matrix and cytoskeleton), and indirectly by altering the fundamental units of muscle contraction (modifying muscle fiber composition and function), thereby compromising tendon health and ultimately impeding the repair process. The cytoskeleton in muscle cells pathway involves a network of collagen, glycoproteins, proteoglycans, and elastin[133]. This network forms a scaffold in the extracellular matrix (ECM), creating a stable environment called tendon homeostasis that supports tendon repair and remodeling[134]. The ECM contains type I collagen and tenocytes (mature tendon cells), which extend into the matrix to maintain this stability—essential for regeneration[135,136]. Tendon regeneration depends on satellite cells located between the muscle fiber membrane and basement membrane. Tendon homeostasis acts as a dynamic environment that transmits signals to activate and guide these stem cells[137]. Upon injury, satellite cells activate, become fibroblasts, and form mature fibers—a process involving transcription factors where the cytoskeleton pathway plays a role[138]. Specifically, this pathway, through components like fibronectin and collagen VI, helps balance stem cell self-renewal and differentiation to maintain regeneration[139-141]. Tendons start as cell clusters that organize into structured tissues connecting muscle to bone for movement. Regeneration involves both tenocytes and their environment. When this balance is disrupted, it causes collagen disorganization, reduced blood supply, ECM issues, inflammation, and cell death[142]. T2DM disrupts tendon homeostasis by altering collagen fiber size and structure[143]. T2DM creates high glucose in the ECM, which impairs tenocyte function and increases collagen glycosylation and cross-linking, leading to structural changes[144,145]. Studies show high glucose reduces type I collagen and tenocytes while increasing apoptosis[146]. In T2DM mice, injuries caused more ECM disruption, limited cell migration, and weaker biomechanics[47]. Healing also showed more fibrosis, reducing motion and increasing resistance[147]. Ultimately, T2DM creates high glucose in the ECM, which—through the cytoskeleton pathway—alters collagen structure, disrupts tendon homeostasis, and impairs repair. Long-term insulin resistance also reduces collagen density and organization[45].

We also identified the involvement of the motor proteins signaling pathway. Motor proteins play an indispensable role in numerous cellular functions, particularly in regulating cell motility, as well as maintaining cell shape and polarity. The interaction between filamentous actin and myosin constitutes the molecular basis of muscle contraction[148]. Biologically, the mechanical load generated by muscle contraction promotes the formation of the tendon ECM[149,150] and contributes to the recovery of its mechanical properties[151,152]. The cytoskeleton-ECM interface has been shown to provide critical biomechanical support and transmit contractile (tensile) forces within muscle, thereby serving as a key regulator of muscle mechanical homeostasis[153]. Notably, mechanical loading induced by muscle stretching has been shown to promote tendon healing following injury, in contrast to conventional cast immobilization[154]. A study on rotator cuff repair in mice further demonstrated that eccentric contraction is more effective than other contraction modalities in promoting tendon healing[155]. Collectively, these findings suggest that the motor proteins signaling pathway may facilitate tendon repair through mechanical forces generated by actin-myosin interactions. However, in the context of T2DM, tendon pathology is characterized by increased thickness and an altered ratio of type I to type III collagen[156]. Structural abnormalities, including disorganized collagen fibers, microtears, tendon thickening, and calcification, have also been documented[44,157,158]. Ahmed et al. observed significantly reduced collagen organization in injured tendons of T2DM rats[48]. These findings indicate that T2DM modifies the compositional and structural properties of tendons, thereby impairing the mechanical loading capacity mediated by the motor proteins pathway and compromising the healing process. Type I collagen, the predominant component of the tendon ECM, is highly susceptible to glucose oxidation. Glycogen accumulation during this process further alters ECM

integrity[159,160]. Additionally, glycation of type I collagen and other matrix proteins is prevalent in diabetic tendons[161], which undermines ECM remodeling capacity[162]. In line with prolonged scar maturation, the fibrotic repair response in T2DM tendons is characterized by excessive deposition of both type III and type I collagen. Type III collagen, a major constituent of scar tissue, is associated with impaired biomechanical function. Such aberrant matrix accumulation is also observed in other diabetes-related complications, including retinopathy and cardiomyopathy[163]. Although the underlying mechanisms may vary across tissues, chronic hyperglycemia and its associated tissue responses are widely regarded as drivers of fibrosis, with abnormal type III collagen deposition frequently noted in various diabetic pathologies. Persistent hyperglycemia may disrupt tenocyte and ECM homeostasis, promoting the overexpression of pro-inflammatory and pro-fibrotic mediators[164]. Despite its clinical relevance, the impact of T2DM on tendon homeostasis remains largely underexplored in basic research, possibly due to limited recognition of the severity of associated tendon complications. Nevertheless, tendon homeostasis is a critical determinant of successful tendon repair and subsequent functional recovery. Impaired tendon healing contributes to reduced patient mobility and diminished quality of life, imposing considerable clinical and societal burdens[165-167].

To further identify hub genes associated with both T2DM and tendon injury, we performed PPI network analysis integrated with cross-validation from seven machine learning models. This approach enabled the identification of three hub genes: MYL2, MYL3, and TTN. Notably, all three genes were significantly enriched in the cytoskeleton in muscle cells signaling pathway. Both MYL2 and MYL3 are myosin proteins—actin-activated mechanoenzymes that hydrolyze ATP to generate mechanical force and motion along actin filaments[168]. Specifically, MYL2 (myosin regulatory light chain 2) is a contractile protein critical for cardiac development and function[169]. MYL3 exhibits differential expression during tendon repair phases[67], while MYL2 is markedly upregulated during tendon regeneration and angiogenesis[170] and shows elevated expression in degenerative tendon tears[171]. TTN (titin) is the largest known human protein and plays an indispensable role in the assembly and function of vertebrate striated muscle. Functioning as a molecular spring within the sarcomere, TTN maintains resting tension in muscle fibers and participates in mechanosignaling[172]. By establishing filament-level connections, it helps preserve the mechanical balance across the sarcomere. Given that cross-bridge properties largely determine sarcomere extensibility[173], TTN is essential for maintaining the internal mechanical environment of tendons, thereby supporting appropriate muscle–tendon tension. Emerging evidence suggests that healthy tendons modulate the timing of muscle energy dissipation during movement, a function in which TTN may play a key role in energy absorption and muscle protection[174,175]. Elastic mechanisms identified to date involve actin and myosin filaments, TTN, and the extracellular matrix (ECM) scaffold[176]. TTN exhibits spring-like behavior in tendons, with an elastic deformation range exceeding that of actin and myosin filaments[177]. In addition, TTN contributes to the stability of the ECM as a parallel elastic element within the tendon environment[178,179]. A study examining the TTN/ECM passive tension ratio revealed that T2DM may increase passive stiffness in muscle fibers, potentially due to excessive ECM accumulation driven by insulin resistance[180]. Insulin resistance has also been linked to increased collagen deposition in skeletal muscle, a precursor to mitochondrial dysfunction and a feature of ECM abnormalities[181]. These findings suggest that in T2DM, insulin resistance impairs mitochondrial function and disrupts tendon ECM stability via TTN. Moreover, cytoskeletal and ECM dysregulation, accompanied by actin filament and collagen deposition, has been observed under diabetic conditions. With respect to contractile proteins, slow-type myosins MYL2 and MYL3 are significantly altered in T2DM[182]. Slow-twitch fibers, rich in mitochondria and characterized by oxidative metabolism, exhibit high insulin sensitivity[183]. This indicates that T2DM may reshape muscle fiber composition through MYL2 and MYL3. Prolonged glycolytic metabolism may also influence tenocyte function, tendon

homeostasis, and healing via these myosins[184], suggesting that disrupted glucose metabolism in T2DM compromises tendon homeostasis through MYL2 and MYL3. Interestingly, although these genes are traditionally associated with cardiac and skeletal muscle function, our findings reveal their critical involvement in tendon injury. This suggests that T2DM may alter the biomechanical properties of the muscle–tendon unit by modulating the expression or function of these contractile proteins, rendering tendons more susceptible to abnormal stress during movement and impairing repair capacity. Gene set enrichment analysis (GSEA) further supports this notion, showing that high expression of MYL2, MYL3, and TTN is strongly associated with biological processes such as skeletal muscle organ development, actin–myosin filament sliding, and striated muscle contraction.

In addition, immune infiltration analysis revealed a shared immune microenvironment characteristic between T2DM and tendon injury. We found that during tendon injury, the infiltration proportions of eosinophils, M0 macrophages, and CD8+ T cells changed significantly; whereas in T2DM, macrophages (M2 and M0) and resting CD4+ memory T cells exhibited significant differences. Notably, M0 and M2 macrophages showed significant infiltration differences in both disease states. M0 macrophages are undifferentiated macrophages whose polarization towards M1 (pro-inflammatory) or M2 (anti-inflammatory/reparative) phenotypes is critical for tissue repair and inflammation resolution[185]. In diabetic conditions, imbalanced macrophage polarization (e.g., altered M1/M2 ratio) is considered a major cause of poor diabetic wound healing[186]. Among the immune cells involved in tendon repair, macrophages represent a highly significant dynamic population that, depending on their phenotype, can modulate and participate in all healing phases. M1 macrophages are involved in apoptosis, ECM catabolism, and inflammatory responses. In contrast, M2 macrophages promote ECM deposition and restoration of native tissue structure[31]. Tendon injury typically involves an inflammatory phase lasting about one week. During this phase, mediated by cytokines and chemokines, vascular permeability increases, allowing immune cells to infiltrate the injury site.

Subsequently, macrophages and fibroblasts are recruited and proliferate. During the proliferative and remodeling phases, fibroblasts further proliferate and begin producing, depositing, aligning, and cross-linking fibrillar collagen[187]. M2 macrophages are primarily involved in and promote fibroblast proliferation and new tissue deposition, which is crucial for tendon tissue regeneration and remodeling[188]. In the later stages of tendon healing, the concentration of M2 macrophages significantly increases in the tendon ECM region[189]. Another study found that fresh bone marrow promotes M2 macrophage polarization, an intervention leading to the formation of more organized connective tissue[190]. Dysfunction of M2 macrophages can lead to an imbalance in the inflammatory response, thereby disrupting the local immune environment of the injured tendon and ultimately hindering the healing process[191]. Therefore, the immunomodulatory role of M2 macrophages is essential for effective tendon repair. Since macrophages are both direct contributors to and orchestrators of inflammation and repair processes, a deeper understanding of their role in tendon repair is crucial for better comprehending this process. At least two studies have demonstrated temporal changes in macrophage populations during Achilles tendon healing[192], further supporting the role of macrophages in both early inflammatory events and subsequent tissue repair/remodeling. Given the higher inflammatory profile in T2DM patients[193], T2DM leads to altered macrophage responses, which may underlie the impact of T2DM on tendon repair. Unlike the phenotypic changes in macrophages during normal healing, some authors suggest T2DM may affect macrophage metabolism, disrupting the balance between M1 and M2 phenotypes[194]. Macrophages participate in both inflammatory and matrix deposition processes during tendon healing. In tendon repair of T2DM patients, the expression of markers for M2 macrophages involved in ECM deposition is elevated and prolonged.

Studies have found increased scar formation and enhanced M2 polarization during tendon healing in mice; therefore, the activity of M2 macrophages and matrix synthesis represent potential mechanisms through which T2DM influences the fibrotic healing phenotype in tendons[147]. T2DM disrupts tendon homeostasis by altering the expression of genes related to tendon ECM, consistent with previous research, primarily manifested as an imbalanced ratio of collagen fiber structures and increased production of proteoglycans and glycosaminoglycans. Excessive ECM deposition may exacerbate tendon scarring[195]. Furthermore, the regulation of M2 macrophage function might be a key mechanism underlying altered tendon healing in T2DM, and a thorough understanding of this immune environment is crucial for improving patient outcomes. Moreover, our study also found that the hub genes MYL2, MYL3, and TTN were significantly correlated with various immune cells (particularly M0 macrophages and regulatory T cells). This suggests that T2DM may interfere with the local immune microenvironment, especially macrophage polarization and function, by affecting the expression of these core genes, thereby hindering the tendon repair process. For example, the positive correlation between TTN and regulatory T cells might imply its role in modulating local inflammatory responses. In summary, T2DM disrupts the immune environment during the tendon repair phase by interfering with macrophages, leading to impaired tendon healing. The role of macrophages is multifaceted, encompassing their involvement in mediating inflammation, cell proliferation, tissue deposition, and scar formation. There is also some variability based on cell functional phenotype and the stage of tendon healing, which necessitates further animal experiments and clinical evidence to enhance our understanding of macrophage function and the cellular and molecular mechanisms of tendon healing. Further research in these areas may provide new insights for the development of targeted therapies for tendon injuries.

In addition, postoperative rehabilitation following tendon healing is also crucial for functional recovery. Aerobic, stretching, and strengthening exercises help prevent and alleviate tissue stiffness. However, due to the high risk of tendon re-injury in diabetic patients, rehabilitation protocols should be individually tailored to avoid intense or prolonged exercise. Regarding pharmacotherapy, insulin, platelet-rich plasma, aprotinin, corticosteroids, and anti-inflammatory agents may offer some benefits for tendon healing in patients with T2DM, though relevant clinical evidence remains limited[196]. Several limitations of this study should be acknowledged. First, the MR analysis was conducted using GWAS data derived from European populations; therefore, the generalizability of the findings to other ancestral groups requires further validation using diverse population datasets. Second, while the transcriptomic analyses revealed significant associations, they were primarily based on bulk RNA sequencing, which lacks the resolution to capture cell type-specific expression dynamics. Future investigations should employ single-cell sequencing technologies to enable a more precise delineation of transcriptomic changes across distinct cell populations—including tenocytes, fibroblasts, muscle cells, and immune cells—within the muscle–tendon unit under diabetic conditions. Lastly, the hub genes and pathways identified in this study necessitate functional validation through in vitro cellular assays and in vivo animal models to clarify their mechanistic roles in the pathogenesis of diabetic tendinopathy.

Conclusion

In summary, this study provides genetic evidence through MR analysis for the adverse effects of T2DM on tendon repair, and identifies related hub genes (MYL2, MYL3, TTN) and major signaling pathways (cytoskeleton in muscle cells and motor proteins) via multi-omics integration analysis. T2DM, driven by insulin resistance and glucose metabolism disorders, leads to mitochondrial dysfunction, which,

mediated by TTN, MYL2, and MYL3, results in excessive deposition of actin filaments and collagen fibers within the tendon ECM. Concurrently, the hyperglycemic environment influences collagen fiber morphology through the cytoskeleton in muscle cells pathway, and alters the ratio of filamentous actin to myosin via the motor proteins pathway, thereby affecting the mechanical load capacity of the repaired tendon. Furthermore, T2DM may also exert its effects by impacting muscle fiber structure and the immune microenvironment, particularly involving macrophages. These findings offer new insights into the pathogenesis of diabetic tendinopathy and lay a theoretical foundation for the development of targeted preventive and therapeutic strategies.

Declarations

Clinical trial number

not applicable.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The datasets analyzed in the study are available in the Gene Expression Omnibus database (https://ncbi.nlm.nih.gov/geo/) with accession codes GSE26051, GSE15932, GSE25724 and GSE29221. The data related to TI, and T2DM can be accessed in the GWAS database (https://gwas.mrcieu.ac.uk/) using the search term “Tendon injury” and “T2DM”. All data generated or analysed during this study are included in this published article and its table files.

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding statement

This study was supported by the Special Fund for Key Scientific Research Projects of Higher Education Institutions in Henan Province in 2025 (Project number: 25A360023).

Authors' contributions

S.S. wrote the main manuscript text, S.S., J.L. and Y.C. conducted overall quality control of the article, H.L., S.S., Y.L., G.L. and L.Y. performed early data screening and processing, and S.S., Y.C. and J.L. optimized the images. All authors reviewed the manuscript.

Acknowledgments

Much appreciation goes to the team of Professor Yujing Cao from Henan University of Chinese Medicine for their technical assistance. We also thank our colleagues in the Emergency Department of Henan Provincial Hospital of Traditional Chinese Medicine for their help in the data study.

References

1 Timasheva, Y. *et al.* Integrating Common Risk Factors with Polygenic Scores Improves the Prediction of Type 2 Diabetes. *Int J Mol Sci* **24**, doi:10.3390/ijms24020984 (2023).

2 Mohammed, A. A. *et al.* Allelic Discrimination of Vitamin D Receptor Polymorphisms and Risk of Type 2 Diabetes Mellitus: A Case-Controlled Study. *Healthcare (Basel)* **11**, doi:10.3390/healthcare11040485 (2023).

3 Wu, L. *et al.* Dissecting the Association of Apolipoprotein E Gene Polymorphisms With Type 2 Diabetes Mellitus and Coronary Artery Disease. *Front Endocrinol (Lausanne)* **13**, 838547, doi:10.3389/fendo.2022.838547 (2022).

4 Azarova, I., Polonikov, A. & Klyosova, E. Molecular Genetics of Abnormal Redox Homeostasis in Type 2 Diabetes Mellitus. *Int J Mol Sci* **24**, doi:10.3390/ijms24054738 (2023).

5 Qiu, H. *et al.* The value of radiomics to predict abnormal bone mass in type 2 diabetes mellitus patients based on CT imaging for paravertebral muscles. *Front Endocrinol (Lausanne)* **13**, 963246, doi:10.3389/fendo.2022.963246 (2022).

6 11. Microvascular Complications and Foot Care: Standards of Medical Care in Diabetes-2020. *Diabetes Care* **43**, S135-s151, doi:10.2337/dc20-S011 (2020).

7 Ranger, T. A., Wong, A. M., Cook, J. L. & Gaida, J. E. Is there an association between tendinopathy and diabetes mellitus? A systematic review with meta-analysis. *Br J Sports Med* **50**, 982-989, doi:10.1136/bjsports-2015-094735 (2016).

8 Tan, H. Y. *et al.* Development of a novel in vitro insulin resistance model in primary human tenocytes for diabetic tendinopathy research. *PeerJ* **8**, e8740, doi:10.7717/peerj.8740 (2020).

9 Bernardi, S., Severini, G. M., Zauli, G. & Secchiero, P. Cell-based therapies for diabetic complications. *Exp Diabetes Res* **2012**, 872504, doi:10.1155/2012/872504 (2012).

10 Giannopoulos, A. *et al.* Cellular homeostatic tension and force transmission measured in human engineered tendon. *J Biomech* **78**, 161-165, doi:10.1016/j.jbiomech.2018.07.032 (2018).

11 Bi, Y. *et al.* Identification of tendon stem/progenitor cells and the role of the extracellular matrix in their niche. *Nat Med* **13**, 1219-1227, doi:10.1038/nm1630 (2007).

12 Thomopoulos, S., Parks, W. C., Rifkin, D. B. & Derwin, K. A. Mechanisms of tendon injury and repair. *J Orthop Res* **33**, 832-839, doi:10.1002/jor.22806 (2015).

13 Johanna P, d. J. *et al.* The incidence of acute traumatic tendon injuries in the hand and wrist: a 10-year population-based study. *Clin Orthop Surg* **6**, doi:10.4055/cios.2014.6.2.196 (2014).

14 Bedi, A. *et al.* Diabetes mellitus impairs tendon-bone healing after rotator cuff repair. *J Shoulder Elbow Surg* **19**, 978-988, doi:10.1016/j.jse.2009.11.045 (2010).

15 Ramchurn, N. *et al.* Upper limb musculoskeletal abnormalities and poor metabolic control in diabetes. *Eur J Intern Med* **20**, 718-721, doi:10.1016/j.ejim.2009.08.001 (2009).

16 Gaida, J. E., Cook, J. L. & Bass, S. L. Adiposity and tendinopathy. *Disabil Rehabil* **30**, 1555-1562, doi:10.1080/09638280701786864 (2008).

17 Voleti, P. B., Buckley, M. R. & Soslowsky, L. J. Tendon healing: repair and regeneration. *Annu Rev Biomed Eng* **14**, 47-71, doi:10.1146/annurev-bioeng-071811-150122 (2012).

18 James, R., Kesturu, G., Balian, G. & Chhabra, A. B. Tendon: biology, biomechanics, repair, growth factors, and evolving treatment options. *J Hand Surg Am* **33**, 102-112, doi:10.1016/j.jhsa.2007.09.007 (2008).

19 Wojciak, B. & Crossan, J. F. The accumulation of inflammatory cells in synovial sheath and epitenon during adhesion formation in healing rat flexor tendons. *Clin Exp Immunol* **93**, 108-114, doi:10.1111/j.1365-2249.1993.tb06505.x (1993).

20 Sharma, P. & Maffulli, N. Tendon injury and tendinopathy: healing and repair. *J Bone Joint Surg Am* **87**, 187-202, doi:10.2106/jbjs.D.01850 (2005).

21 Palmes, D. *et al.* Achilles tendon healing: long-term biomechanical effects of postoperative mobilization and immobilization in a new mouse model. *J Orthop Res* **20**, 939-946, doi:10.1016/s0736-0266(02)00032-3 (2002).

22 Branford, O. A., Mudera, V., Brown, R. A., McGrouther, D. A. & Grobbelaar, A. O. A novel biomimetic material for engineering postsurgical adhesion using the injured digital flexor tendon-synovial complex as an in vivo model. *Plast Reconstr Surg* **121**, 781-793, doi:10.1097/01.prs.0000299373.25294.65 (2008).

23 Garner, W. L., McDonald, J. A., Koo, M., Kuhn, C., 3rd & Weeks, P. M. Identification of the collagen-producing cells in healing flexor tendons. *Plast Reconstr Surg* **83**, 875-879, doi:10.1097/00006534-198905000-00018 (1989).

24 Chuen, F. S. *et al.* Immunohistochemical characterization of cells in adult human patellar tendons. *J Histochem Cytochem* **52**, 1151-1157, doi:10.1369/jhc.3A6232.2004 (2004).

25 Loiselle, A. E. *et al.* Remodeling of murine intrasynovial tendon adhesions following injury: MMP and neotendon gene expression. *J Orthop Res* **27**, 833-840, doi:10.1002/jor.20769 (2009).

26 Rui, Y. F. *et al.* Isolation and characterization of multipotent rat tendon-derived stem cells. *Tissue Eng Part A* **16**, 1549-1558, doi:10.1089/ten.TEA.2009.0529 (2010).

27 Ni, M. *et al.* Tendon-derived stem cells (TDSCs) promote tendon repair in a rat patellar tendon window defect model. *J Orthop Res* **30**, 613-619, doi:10.1002/jor.21559 (2012).

28 Lui, P. P., Rui, Y. F., Ni, M. & Chan, K. M. Tenogenic differentiation of stem cells for tendon repair-what is the current evidence? *J Tissue Eng Regen Med* **5**, e144-163, doi:10.1002/term.424 (2011).

29 Sarıkaya, B. *et al.* Comparison of the effects of human recombinant epidermal growth factor and platelet-rich plasma on healing of rabbit patellar tendon. *Eklem hastaliklari ve cerrahisi = Joint diseases & related surgery* **28**, 92-99, doi:10.5606/ehc.2017.55396 (2017).

30 Galatz, L. M., Ball, C. M., Teefey, S. A., Middleton, W. D. & Yamaguchi, K. The outcome and repair integrity of completely arthroscopically repaired large and massive rotator cuff tears. *J Bone Joint Surg Am* **86**, 219-224, doi:10.2106/00004623-200402000-00002 (2004).

31 Aydin, A. *et al.* [Single-stage flexor tendoplasty in the treatment of flexor tendon injuries]. *Acta Orthop Traumatol Turc* **38**, 54-59 (2004).

32 Mall, N. A., Tanaka, M. J., Choi, L. S. & Paletta, G. A., Jr. Factors affecting rotator cuff healing. *J Bone Joint Surg Am* **96**, 778-788, doi:10.2106/jbjs.M.00583 (2014).

33 Cho, N. S., Moon, S. C., Jeon, J. W. & Rhee, Y. G. The influence of diabetes mellitus on clinical and structural outcomes after arthroscopic rotator cuff repair. *Am J Sports Med* **43**, 991-997, doi:10.1177/0363546514565097 (2015).

34 Maffulli, N., Longo, U. G., Maffulli, G. D., Khanna, A. & Denaro, V. Achilles tendon ruptures in diabetic patients. *Arch Orthop Trauma Surg* **131**, 33-38, doi:10.1007/s00402-010-1097-0 (2011).

35 Clement, N. D., Hallett, A., MacDonald, D., Howie, C. & McBirnie, J. Does diabetes affect outcome after arthroscopic repair of the rotator cuff? *J Bone Joint Surg Br* **92**, 1112-1117, doi:10.1302/0301-620x.92b8.23571 (2010).

36 Rosenbloom, A. L. & Silverstein, J. H. Connective tissue and joint disease in diabetes mellitus. *Endocrinol Metab Clin North Am* **25**, 473-483, doi:10.1016/s0889-8529(05)70335-2 (1996).

37 Giha, H. A., Sater, M. S. & Alamin, O. A. O. Diabetes mellitus tendino-myopathy: epidemiology, clinical features, diagnosis and management of an overlooked diabetic complication. *Acta Diabetol* **59**, 871-883, doi:10.1007/s00592-022-01860-9 (2022).

38 Afolabi, B. I., Idowu, B. M. & Onigbinde, S. O. Achilles tendon degeneration on ultrasound in type 2 diabetic patients. *J Ultrason* **20**, e291-e299, doi:10.15557/JoU.2020.0051 (2021).

39 DeGroot, J. *et al.* Accumulation of advanced glycation end products decreases collagen turnover by bovine chondrocytes. *Exp Cell Res* **266**, 303-310, doi:10.1006/excr.2001.5224 (2001).

40 Odetti, P. *et al.* Scanning force microscopy reveals structural alterations in diabetic rat collagen fibrils: role of protein glycation. *Diabetes Metab Res Rev* **16**, 74-81, doi:10.1002/(sici)1520-7560(200003/04)16:2<74::aid-dmrr80>3.0.co;2-1 (2000).

41 Spanheimer, R. G. Correlation between decreased collagen production in diabetic animals and in cells exposed to diabetic serum: response to insulin. *Matrix* **12**, 101-107, doi:10.1016/s0934-8832(11)80051-x (1992).

42 Snedeker, J. G. How High Glucose Levels Affect Tendon Homeostasis. *Adv Exp Med Biol* **920**, 191-198, doi:10.1007/978-3-319-33943-6_18 (2016).

43 Fitzgibbons, P. G. & Weiss, A. P. Hand manifestations of diabetes mellitus. *J Hand Surg Am* **33**, 771-775, doi:10.1016/j.jhsa.2008.01.038 (2008).

44 Studentsova, V., Mora, K. M., Glasner, M. F., Buckley, M. R. & Loiselle, A. E. Obesity/Type II Diabetes Promotes Function-limiting Changes in Murine Tendons that are not reversed by Restoring Normal Metabolic Function. *Sci Rep* **8**, 9218, doi:10.1038/s41598-018-27634-4 (2018).

45 Connizzo, B. K., Bhatt, P. R., Liechty, K. W. & Soslowsky, L. J. Diabetes alters mechanical properties and collagen fiber re-alignment in multiple mouse tendons. *Ann Biomed Eng* **42**, 1880-1888, doi:10.1007/s10439-014-1031-7 (2014).

46 Abate, M., Schiavone, C., Salini, V. & Andia, I. Occurrence of tendon pathologies in metabolic disorders. *Rheumatology (Oxford)* **52**, 599-608, doi:10.1093/rheumatology/kes395 (2013).

47 David, M. A. *et al.* Tendon repair is compromised in a high fat diet-induced mouse model of obesity and type 2 diabetes. *PLoS One* **9**, e91234, doi:10.1371/journal.pone.0091234 (2014).

48 Ahmed, A. S. *et al.* Type 2 diabetes impairs tendon repair after injury in a rat model. *J Appl Physiol (1985)* **113**, 1784-1791, doi:10.1152/japplphysiol.00767.2012 (2012).

49 Sukenik, S. *et al.* Limited joint mobility and other rheumatological manifestations in diabetic patients. *Diabete Metab* **13**, 187-192 (1987).

50 Fox, A. J. *et al.* Diabetes mellitus alters the mechanical properties of the native tendon in an experimental rat model. *J Orthop Res* **29**, 880-885, doi:10.1002/jor.21327 (2011).

51 Guney, A. *et al.* Biomechanical properties of Achilles tendon in diabetic vs. non-diabetic patients. *Exp Clin Endocrinol Diabetes* **123**, 428-432, doi:10.1055/s-0035-1549889 (2015).

52 Lu, P. P. *et al.* Understanding cellular and molecular mechanisms of pathogenesis of diabetic tendinopathy. *World J Stem Cells* **12**, 1255-1275, doi:10.4252/wjsc.v12.i11.1255 (2020).

53 Ackermann, P. W. *et al.* Neuronal pathways in tendon healing and tendinopathy--update. *Front Biosci (Landmark Ed)* **19**, 1251-1278, doi:10.2741/4280 (2014).

54 Cagliero, E., Apruzzese, W., Perlmutter, G. S. & Nathan, D. M. Musculoskeletal disorders of the hand and shoulder in patients with diabetes mellitus. *Am J Med* **112**, 487-490, doi:10.1016/s0002-9343(02)01045-8 (2002).

55 Richmond, R. C. & Davey Smith, G. Mendelian Randomization: Concepts and Scope. *Cold Spring Harb Perspect Med* **12**, doi:10.1101/cshperspect.a040501 (2022).

56 Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC bioinformatics* **9**, 559, doi:10.1186/1471-2105-9-559 (2008).

57 Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-15550, doi:10.1073/pnas.0506580102 (2005).

58 Beck, T., Rowlands, T., Shorter, T. & Brookes, A. J. GWAS Central: an expanding resource for finding and visualising genotype and phenotype data from genome-wide association studies. *Nucleic Acids Res* **51**, D986-d993, doi:10.1093/nar/gkac1017 (2023).

59 Lv, X., Wang, L., Yao, J. & Huang, Y. Investigating the Gene Relation Between Cervical Spondylosis and Depression: Bidirectional Mendelian Randomization Study. *J Pain Res* **17**, 4343-4355, doi:10.2147/jpr.S488082 (2024).

60 Gu, X., Dou, M., Cao, B., Jiang, Z. & Chen, Y. Peripheral level of CD33 and Alzheimer's disease: a bidirectional two-sample Mendelian randomization study. *Transl Psychiatry* **12**, 427, doi:10.1038/s41398-022-02205-4 (2022).

61 Dudbridge, F. Polygenic Mendelian Randomization. *Cold Spring Harbor perspectives in medicine* **11**, doi:10.1101/cshperspect.a039586 (2021).

62 Emdin, C., Khera, A. & Kathiresan, S. Mendelian Randomization. *JAMA* **318**, 1925-1926, doi:10.1001/jama.2017.17219 (2017).

63 Burgess, S., Small, D. S. & Thompson, S. G. A review of instrumental variable estimators for Mendelian randomization. *Stat Methods Med Res* **26**, 2333-2355, doi:10.1177/0962280215597579 (2017).

64 Burgess, S. & Thompson, S. G. Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur J Epidemiol* **32**, 377-389, doi:10.1007/s10654-017-0255-x (2017).

65 Yu, H. & Song, X. The relationship between Alzheimer disease and thyroiditis: A two-sample Mendelian randomization study. *Medicine (Baltimore)* **102**, e35712, doi:10.1097/md.0000000000035712 (2023).

66 Clough, E. *et al.* NCBI GEO: archive for gene expression and epigenomics data sets: 23-year update. *Nucleic Acids Res* **52**, D138-d144, doi:10.1093/nar/gkad965 (2024).

67 Sun, S. *et al.* Screening of hub genes and immunocytes related to tendon injury based on bioinformatics and machine learning models. *Sci Rep* **15**, 31432, doi:10.1038/s41598-025-15210-6 (2025).

68 Shen, Q. *et al.* Identification and validation of the diagnostic biomarker MFAP5 for CAVD with type 2 diabetes by bioinformatics analysis. *Front Immunol* **15**, 1506663, doi:10.3389/fimmu.2024.1506663 (2024).

69 Wang, S. *et al.* Bioinformatics analysis of oxidative phosphorylation-related differentially expressed genes in osteoporosis. *Eur J Med Res* **30**, 294, doi:10.1186/s40001-025-02568-6 (2025).

70 Kuang, Y. *et al.* Weighted gene coexpression network analysis reveals negative regulation of hypertrophic cardiomyopathy by carboxylesterase 1 and cathepsin C. *Gen Physiol Biophys* **42**, 361-372, doi:10.4149/gpb_2023009 (2023).

71 Xin, P., Pei, S., Ma, N. & Xiao, L. PCK1 as a potential hub gene in distinguishing lactate metabolism between rheumatoid arthritis and osteoarthritis. *PeerJ* **13**, e19661, doi:10.7717/peerj.19661 (2025).

72 Zhu, Y. X., Huang, J. Q., Ming, Y. Y., Zhuang, Z. & Xia, H. Screening of key biomarkers of tendinopathy based on bioinformatics and machine learning algorithms. *PLoS One* **16**, e0259475, doi:10.1371/journal.pone.0259475 (2021).

73 Haoru, D. *et al.* Integrated bioinformatic analysis reveals the underlying molecular mechanism of and potential drugs for pulmonary arterial hypertension. *Aging (Albany NY)* **13**, doi:10.18632/aging.203040 (2021).

74 Haichao, G. *et al.* Identification of key genes and molecular mechanisms of chronic urticaria based on bioinformatics. *Skin Res Technol* **30**, doi:10.1111/srt.13624 (2024).

75 Ast, T. *et al.* METTL17 is an Fe-S cluster checkpoint for mitochondrial translation. *Mol Cell* **84**, 359-374.e358, doi:10.1016/j.molcel.2023.12.016 (2024).

76 Li, N. *et al.* Potent prophylactic cancer vaccines harnessing surface antigens shared by tumour cells and induced pluripotent stem cells. *Nat Biomed Eng* **9**, 215-233, doi:10.1038/s41551-024-01309-0 (2025).

77 Emil K, G., David, Z., Regina H, R., Sonia, G.-R. & Mina, R. ggtranscript: an R package for the visualization and interpretation of transcript isoforms using ggplot2. *Bioinformatics* **38**, doi:10.1093/bioinformatics/btac409 (2022).

78 Xiaoqian, L. *et al.* Identification of hub genes and key pathways associated with angioimmunoblastic T-cell lymphoma using weighted gene co-expression network analysis. *Cancer Manag Res* **11**, doi:10.2147/cmar.S185030 (2019).

79 Zijian, L., Juan, F. & Zhengwei, Y. Key Modules and Hub Genes Identified by Coexpression Network Analysis for Revealing Novel Biomarkers for Spina Bifida. *Front Genet* **11**, doi:10.3389/fgene.2020.583316 (2020).

80 Zhang, B. & Horvath, S. A general framework for weighted gene co-expression network analysis. *Stat Appl Genet Mol Biol* **4**, Article17, doi:10.2202/1544-6115.1128 (2005).

81 Horvath, S. & Dong, J. Geometric interpretation of gene coexpression network analysis. *PLoS Comput Biol* **4**, e1000117, doi:10.1371/journal.pcbi.1000117 (2008).

82 Zhang, F. *et al.* Identification of Potential Therapeutic Targets of Alzheimer's Disease By Weighted Gene Co-Expression Network Analysis. *Chin Med Sci J* **35**, 330-341, doi:10.24920/003695 (2020).

83 Bardou, P., Mariette, J., Escudié, F., Djemiel, C. & Klopp, C. jvenn: an interactive Venn diagram viewer. *BMC Bioinformatics* **15**, 293, doi:10.1186/1471-2105-15-293 (2014).

84 Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498-2504, doi:10.1101/gr.1239303 (2003).

85 Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **28**, 27-30, doi:10.1093/nar/28.1.27 (2000).

86 Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. & Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res* **44**, D457-462, doi:10.1093/nar/gkv1070 (2016).

87 Zhou, Y. *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nature communications* **10**, 1523, doi:10.1038/s41467-019-09234-6 (2019).

88 Simon J, W., Maryam A, A. & Daniel J, H. A Bioinformatics Toolkit: In Silico Tools and Online Resources for Investigating Genetic Variation. *Semin Thromb Hemost* **45**, doi:10.1055/s-0039-1692978 (2019).

89 Damian, S. *et al.* The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res* **51**, doi:10.1093/nar/gkac1000 (2022).

90 Zhengfei, M., Ping, Z., Peidong, Y. & Zhongwu, S. Identification of immune-related molecular markers in intracranial aneurysm (IA) based on machine learning and cytoscape-cytohubba plug-in. *BMC Genom Data* **24**, doi:10.1186/s12863-023-01121-w (2023).

91 Ali, A., Mohan, J., Nadaf, T. A. A., Ravishankar, H. & Deepa, K. R. Bioinformatics-Driven Discovery of Signaling Pathways and Genes Influencing Cervical Cancer. *SN Computer Science* **5**, 989, doi:10.1007/s42979-024-03347-6 (2024).

92 Gaofeng, H. *et al.* Identification of thyroid cancer biomarkers using WGCNA and machine learning. *Eur J Med Res* **30**, doi:10.1186/s40001-025-02466-x (2025).

93 Breiman, L. Random Forests. *Machine Learning* **45**, 5-32, doi:10.1023/A:1010933404324 (2001).

94 Shi, X. *et al.* Integrated bioinformatics and experiment validation reveal cuproptosis-related biomarkers and therapeutic targets in sepsis-induced myocardial dysfunction. *BMC Infect Dis* **25**, 445, doi:10.1186/s12879-025-10822-9 (2025).

95 Huang, S. *et al.* Applications of Support Vector Machine (SVM) Learning in Cancer Genomics. *Cancer Genomics Proteomics* **15**, 41-51, doi:10.21873/cgp.20063 (2018).

96 Ukey, N. *et al.* Survey on Exact kNN Queries over High-Dimensional Data Space. *Sensors (Basel)* **23**, doi:10.3390/s23020629 (2023).

97 Yao, Z. & Ruzzo, W. L. A regression-based K nearest neighbor algorithm for gene function prediction from heterogeneous data. *BMC Bioinformatics* **7 Suppl 1**, S11, doi:10.1186/1471-2105-7-s1-s11 (2006).

98 Wu, Y., Cai, X., Hu, M., Cao, R. & Wang, Y. Oxidative stress-related genes in uveal melanoma: the role of CALM1 in modulating oxidative stress and apoptosis and its prognostic significance. *Front Oncol* **15**, 1618601, doi:10.3389/fonc.2025.1618601 (2025).

99 Liu, Z. *et al.* Identification of palmitoylated biomarkers in non-alcoholic fatty liver disease via integrated bioinformatics analysis and machine learning. *Sci Rep* **15**, 28177, doi:10.1038/s41598-025-13477-3 (2025).

100 Zhao, X. *et al.* Comprehensive analysis of gene expression profiles of annulus fibrosus subtypes and hub genes in intervertebral disc degeneration. *Aging (Albany NY)* **16**, 5370-5386, doi:10.18632/aging.205653 (2024).

101 Yang, C. *et al.* Identification of gene biomarkers in patients with postmenopausal osteoporosis. *Mol Med Rep* **19**, 1065-1073, doi:10.3892/mmr.2018.9752 (2019).

102 Song, S. *et al.* Molecular Indicator for Distinguishing Multi-drug-Resistant Tuberculosis from Drug Sensitivity Tuberculosis and Potential Medications for Treatment. *Mol Biotechnol* **67**, 3946-3959, doi:10.1007/s12033-024-01299-z (2025).

103 Willi, S., Patrick, R. & Harald, B. Selection of important variables and determination of functional form for continuous predictors in multivariable model building. *Stat Med* **26**, doi:10.1002/sim.3148 (2007).

104 Saejoon, K. Margin-maximised redundancy-minimised SVM-RFE for diagnostic classification of mammograms. *Int J Data Min Bioinform* **10**, doi:10.1504/ijdmb.2014.064889 (2014).

105 Jan, Y. *et al.* Machine learning analysis of motor evoked potential time series to predict disability progression in multiple sclerosis. *BMC Neurol* **20**, doi:10.1186/s12883-020-01672-w (2020).

106 Ling, G. & Chong-En, X. Integrated bioinformatics and machine learning algorithms reveal the critical cellular senescence-associated genes and immune infiltration in heart failure due to ischemic cardiomyopathy. *Front Immunol* **14**, doi:10.3389/fimmu.2023.1150304 (2023).

107 Hu, X., Ni, S., Zhao, K., Qian, J. & Duan, Y. Bioinformatics-Led Discovery of Osteoarthritis Biomarkers and Inflammatory Infiltrates. *Frontiers in Immunology* **13**, doi:10.3389/fimmu.2022.871008 (2022).

108 Li, A. L. *et al.* Exploration of the Immune-Related Signatures and Immune Infiltration Analysis in Melanoma. *Anal Cell Pathol (Amst)* **2021**, 4743971, doi:10.1155/2021/4743971 (2021).

109 Weitao, S. *et al.* Sangerbox: A comprehensive, interaction-friendly clinical bioinformatics analysis platform. *Imeta* **1**, doi:10.1002/imt2.36 (2022).

110 Zhou, Q. *et al.* Identification of ubiquitination-related hub genes in chronic myeloid leukemia cell by bioinformatics analysis. *J Cancer* **15**, 3750-3759, doi:10.7150/jca.96405 (2024).

111 An, W. *et al.* Identification of crosstalk genes and immune characteristics between Alzheimer's disease and atherosclerosis. *Front Immunol* **15**, 1443464, doi:10.3389/fimmu.2024.1443464 (2024).

112 Wu, T. *et al.* clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (Camb)* **2**, 100141, doi:10.1016/j.xinn.2021.100141 (2021).

113 Zhang, C. *et al.* Genome-wide mutation profiling and related risk signature for prognosis of papillary renal cell carcinoma. *Ann Transl Med* **7**, 427, doi:10.21037/atm.2019.08.113 (2019).

114 Richter, R. R. & Austin, T. M. Using MeSH (medical subject headings) to enhance PubMed search strategies for evidence-based practice in physical therapy. *Phys Ther* **92**, 124-132, doi:10.2522/ptj.20100178 (2012).

115 DeMars, M. & Perruso, C. MeSH and text-word search strategies: precision, recall, and their implications for library instruction. *Journal of the Medical Library Association : JMLA* **110**, 23-33, doi:10.5195/jmla.2022.1283 (2022).

116 Dominguez, V. *et al.* Class II phosphoinositide 3-kinase regulates exocytosis of insulin granules in pancreatic beta cells. *J Biol Chem* **286**, 4216-4225, doi:10.1074/jbc.M110.200295 (2011).

117 Jain, P. *et al.* Systems biology approach reveals genome to phenome correlation in type 2 diabetes. *PLoS One* **8**, e53522, doi:10.1371/journal.pone.0053522 (2013).

118 Huang, H. *et al.* Novel blood biomarkers of pancreatic cancer-associated diabetes mellitus identified by peripheral blood-based gene expression profiles. *Am J Gastroenterol* **105**, 1661-1669, doi:10.1038/ajg.2010.32 (2010).

119 Jelinsky, S. A. *et al.* Regulation of gene expression in human tendinopathy. *BMC Musculoskelet Disord* **12**, 86, doi:10.1186/1471-2474-12-86 (2011).

120 Cheng, X. *et al.* Hypoxia-Mimicking Microenvironment Scaffold for Enhanced Tendon Regeneration. *ACS Appl Mater Interfaces* **17**, 8937-8948, doi:10.1021/acsami.4c18082 (2025).

121 Lin, J. *et al.* Projection of the future diabetes burden in the United States through 2060. *Popul Health Metr* **16**, 9, doi:10.1186/s12963-018-0166-4 (2018).

122 Nichols, A. E. C., Oh, I. & Loiselle, A. E. Effects of Type II Diabetes Mellitus on Tendon Homeostasis and Healing. *J Orthop Res* **38**, 13-22, doi:10.1002/jor.24388 (2020).

123 Schett, G. *et al.* Diabetes is an independent predictor for severe osteoarthritis: results from a longitudinal cohort study. *Diabetes Care* **36**, 403-409, doi:10.2337/dc12-0924 (2013).

124 Mooney, R. A., Sampson, E. R., Lerea, J., Rosier, R. N. & Zuscik, M. J. High-fat diet accelerates progression of osteoarthritis after meniscal/ligamentous injury. *Arthritis Res Ther* **13**, R198, doi:10.1186/ar3529 (2011).

125 Fuentes-Murguia, M. *et al.* Types of Injuries and the Severity of Shoulder Dysfunction Associated with Diabetes Mellitus in Patients with Functional Impairment: A Case-Control Study. *Biomedicines* **12**, doi:10.3390/biomedicines12112634 (2024).

126 Rőszer, T. Understanding the Mysterious M2 Macrophage through Activation Markers and Effector Mechanisms. *Mediators Inflamm* **2015**, 816460, doi:10.1155/2015/816460 (2015).

127 Rosenbloom, A. L. Limitation of finger joint mobility in diabetes mellitus. *J Diabet Complications* **3**, 77-87, doi:10.1016/0891-6632(89)90016-0 (1989).

128 Balci, N., Balci, M. K. & Tüzüner, S. Shoulder adhesive capsulitis and shoulder range of motion in type II diabetes mellitus: association with diabetic complications. *J Diabetes Complications* **13**, 135-140, doi:10.1016/s1056-8727(99)00037-9 (1999).

129 D'Ambrogi, E., Giacomozzi, C., Macellari, V. & Uccioli, L. Abnormal foot function in diabetic patients: the altered onset of Windlass mechanism. *Diabet Med* **22**, 1713-1719, doi:10.1111/j.1464-5491.2005.01699.x (2005).

130 Batista, F. *et al.* Achilles tendinopathy in diabetes mellitus. *Foot Ankle Int* **29**, 498-501, doi:10.3113/fai-2008-0498 (2008).

131 Akturk, M., Karaahmetoglu, S., Kacar, M. & Muftuoglu, O. Thickness of the supraspinatus and biceps tendons in diabetic patients. *Diabetes Care* **25**, 408, doi:10.2337/diacare.25.2.408 (2002).

132 Kääriäinen, M. *et al.* Integrin and dystrophin associated adhesion protein complexes during regeneration of shearing-type muscle injury. *Neuromuscul Disord* **10**, 121-132, doi:10.1016/s0960-8966(99)00077-2 (2000).

133 Takala, T. E. & Virtanen, P. Biochemical composition of muscle extracellular matrix: the effect of loading. *Scand J Med Sci Sports* **10**, 321-325, doi:10.1034/j.1600-0838.2000.010006321.x (2000).

134 Mukund, K. & Subramaniam, S. Skeletal muscle: A review of molecular structure and function, in health and disease. *Wiley Interdiscip Rev Syst Biol Med* **12**, e1462, doi:10.1002/wsbm.1462 (2020).

135 Kalson, N. S. *et al.* A structure-based extracellular matrix expansion mechanism of fibrous tissue growth. *Elife* **4**, doi:10.7554/eLife.05958 (2015).

136 Kannus, P. Structure of the tendon connective tissue. *Scand J Med Sci Sports* **10**, 312-320, doi:10.1034/j.1600-0838.2000.010006312.x (2000).

137 Csapo, R., Gumpenberger, M. & Wessner, B. Skeletal Muscle Extracellular Matrix - What Do We Know About Its Composition, Regulation, and Physiological Roles? A Narrative Review. *Front Physiol* **11**, 253, doi:10.3389/fphys.2020.00253 (2020).

138 Almada, A. E. & Wagers, A. J. Molecular circuitry of stem cell fate in skeletal muscle regeneration, ageing and disease. *Nat Rev Mol Cell Biol* **17**, 267-279, doi:10.1038/nrm.2016.7 (2016).

139 Bentzinger, C. F. *et al.* Fibronectin regulates Wnt7a signaling and satellite cell expansion. *Cell Stem Cell* **12**, 75-87, doi:10.1016/j.stem.2012.09.015 (2013).

140 Urciuolo, A. *et al.* Collagen VI regulates satellite cell self-renewal and muscle regeneration. *Nat Commun* **4**, 1964, doi:10.1038/ncomms2964 (2013).

141 Brack, A. S., Conboy, I. M., Conboy, M. J., Shen, J. & Rando, T. A. A temporal switch from notch to Wnt signaling in muscle stem cells is necessary for normal adult myogenesis. *Cell Stem Cell* **2**, 50-59, doi:10.1016/j.stem.2007.10.006 (2008).

142 Millar, N. L. *et al.* Tendinopathy. *Nat Rev Dis Primers* **7**, 1, doi:10.1038/s41572-020-00234-1 (2021).

143 Grant, W. P. *et al.* Electron microscopic investigation of the effects of diabetes mellitus on the Achilles tendon. *J Foot Ankle Surg* **36**, 272-278; discussion 330, doi:10.1016/s1067-2516(97)80072-5 (1997).

144 Beason, D. P., Abboud, J. A., Kuntz, A. F., Bassora, R. & Soslowsky, L. J. Cumulative effects of hypercholesterolemia on tendon biomechanics in a mouse model. *J Orthop Res* **29**, 380-383, doi:10.1002/jor.21255 (2011).

145 Kent, M. J., Light, N. D. & Bailey, A. J. Evidence for glucose-mediated covalent cross-linking of collagen after glycosylation in vitro. *Biochem J* **225**, 745-752, doi:10.1042/bj2250745 (1985).

146 Lin, Y. C. *et al.* The effects of high glucose on tendon-derived stem cells: implications of the pathogenesis of diabetic tendon disorders. *Oncotarget* **8**, 17518-17528, doi:10.18632/oncotarget.15418 (2017).

147 Ackerman, J. E., Geary, M. B., Orner, C. A., Bawany, F. & Loiselle, A. E. Obesity/Type II diabetes alters macrophage polarization resulting in a fibrotic tendon healing response. *PLoS One* **12**, e0181127, doi:10.1371/journal.pone.0181127 (2017).

148 Dominguez, R. & Holmes, K. C. Actin structure and function. *Annu Rev Biophys* **40**, 169-186, doi:10.1146/annurev-biophys-042910-155359 (2011).

149 Kayama, T. *et al.* Gtf2ird1-Dependent Mohawk Expression Regulates Mechanosensing Properties of the Tendon. *Mol Cell Biol* **36**, 1297-1309, doi:10.1128/mcb.00950-15 (2016).

150 Subramanian, A., Kanzaki, L. F., Galloway, J. L. & Schilling, T. F. Mechanical force regulates tendon extracellular matrix organization and tenocyte morphogenesis through TGFbeta signaling. *Elife* **7**, doi:10.7554/eLife.38069 (2018).

151 Freedman, B. R. *et al.* Dynamic Loading and Tendon Healing Affect Multiscale Tendon Properties and ECM Stress Transmission. *Sci Rep* **8**, 10854, doi:10.1038/s41598-018-29060-y (2018).

152 Eliasson, P., Andersson, T. & Aspenberg, P. Rat Achilles tendon healing: mechanical loading and gene expression. *J Appl Physiol (1985)* **107**, 399-407, doi:10.1152/japplphysiol.91563.2008 (2009).

153 Humphrey, J. D., Dufresne, E. R. & Schwartz, M. A. Mechanotransduction and extracellular matrix homeostasis. *Nat Rev Mol Cell Biol* **15**, 802-812, doi:10.1038/nrm3896 (2014).

154 Yoneno, M. *et al.* Muscle Contraction Is Essential for Tendon Healing and Muscle Function Recovery After Achilles Tendon Rupture and Surgical Repair. *J Orthop Res* **43**, 746-755, doi:10.1002/jor.26044 (2025).

155 Diao, L. *et al.* Eccentric Contraction Enhances Healing of the Bone-Tendon Interface After Rotator Cuff Repair in Mice. *Am J Sports Med* **51**, 3835-3844, doi:10.1177/03635465231202901 (2023).

156 Filgueiras, J. R. *et al.* Morphological and functional changes in skeletal muscle in type 2 diabetes mellitus: A systematic review and meta-analysis. *Physiother Theory Pract* **39**, 1789-1815, doi:10.1080/09593985.2022.2057375 (2023).

157 Rui, Y. F., Lui, P. P., Wong, Y. M., Tan, Q. & Chan, K. M. Altered fate of tendon-derived stem cells isolated from a failed tendon-healing animal model of tendinopathy. *Stem Cells Dev* **22**, 1076-1085, doi:10.1089/scd.2012.0555 (2013).

158 Voog, J. & Jones, D. L. Stem cells and the niche: a dynamic duo. *Cell Stem Cell* **6**, 103-115, doi:10.1016/j.stem.2010.01.011 (2010).

159 Gautieri, A. *et al.* Advanced glycation end-products: Mechanics of aged collagen from molecule to tissue. *Matrix Biol* **59**, 95-108, doi:10.1016/j.matbio.2016.09.001 (2017).

160 Shi, L., Lu, P. P., Dai, G. C., Li, Y. J. & Rui, Y. F. Advanced glycation end productions and tendon stem/progenitor cells in pathogenesis of diabetic tendinopathy. *World J Stem Cells* **13**, 1338-1348, doi:10.4252/wjsc.v13.i9.1338 (2021).

161 Xiao, H., Cai, G. & Liu, M. Fe2+-catalyzed non-enzymatic glycosylation alters collagen conformation during AGE-collagen formation in vitro. *Arch Biochem Biophys* **468**, 183-192, doi:10.1016/j.abb.2007.08.035 (2007).

162 Liang, C. *et al.* Rosiglitazone via upregulation of Akt/eNOS pathways attenuates dysfunction of endothelial progenitor cells, induced by advanced glycation end products. *Br J Pharmacol* **158**, 1865-1873, doi:10.1111/j.1476-5381.2009.00450.x (2009).

163 Strickland, J. W. Development of flexor tendon surgery: twenty-five years of progress. *J Hand Surg Am* **25**, 214-235, doi:10.1053/jhsu.2000.jhsu25a0214 (2000).

164 Xu, K. *et al.* Evaluating the role of type 2 diabetes mellitus in rotator cuff tendinopathy: Development and analysis of a novel rat model. *Front Endocrinol (Lausanne)* **13**, 1042878, doi:10.3389/fendo.2022.1042878 (2022).

165 Dean, B. J. F., Dakin, S. G., Millar, N. L. & Carr, A. J. Review: Emerging concepts in the pathogenesis of tendinopathy. *Surgeon* **15**, 349-354, doi:10.1016/j.surge.2017.05.005 (2017).

166 Hopkins, C. *et al.* Critical review on the socio-economic impact of tendinopathy. *Asia Pac J Sports Med Arthrosc Rehabil Technol* **4**, 9-20, doi:10.1016/j.asmart.2016.01.002 (2016).

167 Ceravolo, M. L., Gaida, J. E. & Keegan, R. J. Quality-of-Life in Achilles Tendinopathy: An Exploratory Study. *Clin J Sport Med* **30**, 495-502, doi:10.1097/jsm.0000000000000636 (2020).

168 Cao, H. *et al.* MYL3 protects chondrocytes from senescence by inhibiting clathrin-mediated endocytosis and activating of Notch signaling. *Nat Commun* **14**, 6190, doi:10.1038/s41467-023-41858-7 (2023).

169 Manivannan, S. N. *et al.* Novel frameshift variant in MYL2 reveals molecular differences between dominant and recessive forms of hypertrophic cardiomyopathy. *PLoS Genet* **16**, e1008639, doi:10.1371/journal.pgen.1008639 (2020).

170 Rai, M. F., Cai, L., Tycksen, E. D., Keener, J. & Chamberlain, A. RNA-Seq reveals distinct transcriptomic differences in rotator cuff tendon based on tear etiology and patient sex. *J Orthop Res* **40**, 2728-2742, doi:10.1002/jor.25296 (2022).

171 Rai, M. F., Cai, L., Tycksen, E. D., Chamberlain, A. & Keener, J. RNA-Seq analysis reveals sex-dependent transcriptomic profiles of human subacromial bursa stratified by tear etiology. *J Orthop Res* **40**, 2713-2727, doi:10.1002/jor.25316 (2022).

172 Tong, S. *et al.* A Comprehensive Review of Muscle-Tendon Junction: Structure, Function, Injury and Repair. *Biomedicines* **12**, doi:10.3390/biomedicines12020423 (2024).

173 Mayans, O. *et al.* Structural basis for activation of the titin kinase domain during myofibrillogenesis. *Nature* **395**, 863-869, doi:10.1038/27603 (1998).

174 Konow, N., Azizi, E. & Roberts, T. J. Muscle power attenuation by tendon during energy dissipation. *Proc Biol Sci* **279**, 1108-1113, doi:10.1098/rspb.2011.1435 (2012).

175 Konow, N. & Roberts, T. J. The series elastic shock absorber: tendon elasticity modulates energy dissipation by muscle during burst deceleration. *Proc Biol Sci* **282**, 20142800, doi:10.1098/rspb.2014.2800 (2015).

176 Roberts, T. J. Contribution of elastic tissues to the mechanics and energetics of muscle function during movement. *J Exp Biol* **219**, 266-275, doi:10.1242/jeb.124446 (2016).

177 Kellermayer, M. S., Smith, S. B., Granzier, H. L. & Bustamante, C. Folding-unfolding transitions in single titin molecules characterized with laser tweezers. *Science* **276**, 1112-1116, doi:10.1126/science.276.5315.1112 (1997).

178 Gajdosik, R. L. Passive extensibility of skeletal muscle: review of the literature with clinical implications. *Clin Biomech (Bristol)* **16**, 87-101, doi:10.1016/s0268-0033(00)00061-9 (2001).

179 Lieber, R. L. & Meyer, G. Structure-Function relationships in the skeletal muscle extracellular matrix. *J Biomech* **152**, 111593, doi:10.1016/j.jbiomech.2023.111593 (2023).

180 Lee, E. J., Jang, H. C., Koo, K. H., Kim, H. Y. & Lim, J. Y. Mechanical Properties of Single Muscle Fibers: Understanding Poor Muscle Quality in Older Adults with Diabetes. *Ann Geriatr Med Res* **24**, 267-273, doi:10.4235/agmr.20.0078 (2020).

181 Berria, R. *et al.* Increased collagen content in insulin-resistant skeletal muscle. *Am J Physiol Endocrinol Metab* **290**, E560-565, doi:10.1152/ajpendo.00202.2005 (2006).

182 Moriggi, M. *et al.* Skeletal Muscle Proteomic Profile Revealed Gender-Related Metabolic Responses in a Diet-Induced Obesity Animal Model. *Int J Mol Sci* **22**, doi:10.3390/ijms22094680 (2021).

183 Duan, Y., Li, F., Tan, B., Yao, K. & Yin, Y. Metabolic control of myofibers: promising therapeutic target for obesity and type 2 diabetes. *Obes Rev* **18**, 647-659, doi:10.1111/obr.12530 (2017).

184 Lin, Y. *et al.* Multi-Proteomic Analysis Reveals the Effect of Protein Lactylation on Matrix and Cholesterol Metabolism in Tendinopathy. *J Proteome Res* **22**, 1712-1722, doi:10.1021/acs.jproteome.2c00756 (2023).

185 Peng, X. *et al.* miR-146a promotes M2 macrophage polarization and accelerates diabetic wound healing by inhibiting the TLR4/NF-κB axis. *J Mol Endocrinol* **69**, 315-327, doi:10.1530/jme-21-0019 (2022).

186 Hou, Y. *et al.* Downregulation of nutrition sensor GCN2 in macrophages contributes to poor wound healing in diabetes. *Cell Rep* **43**, 113658, doi:10.1016/j.celrep.2023.113658 (2024).

187 Pramod B, V., Mark R, B. & Louis J, S. Tendon healing: repair and regeneration. *Annu Rev Biomed Eng* **14**, doi:10.1146/annurev-bioeng-071811-150122 (2012).

188 Hu, J., Liu, S. & Fan, C. Applications of functionally-adapted hydrogels in tendon repair. *Front Bioeng Biotechnol* **11**, 1135090, doi:10.3389/fbioe.2023.1135090 (2023).

189 Nichols, A. E. C., Best, K. T. & Loiselle, A. E. The cellular basis of fibrotic tendon healing: challenges and opportunities. *Transl Res* **209**, 156-168, doi:10.1016/j.trsl.2019.02.002 (2019).

190 Lu, J. *et al.* Tendon-to-Bone Healing in a Rat Extra-articular Bone Tunnel Model: A Comparison of Fresh Autologous Bone Marrow and Bone Marrow-Derived Mesenchymal Stem Cells. *Am J Sports Med* **47**, 2729-2736, doi:10.1177/0363546519862284 (2019).

191 Dakin, S. G., Dudhia, J. & Smith, R. K. Resolving an inflammatory concept: the importance of inflammation and resolution in tendinopathy. *Vet Immunol Immunopathol* **158**, 121-127, doi:10.1016/j.vetimm.2014.01.007 (2014).

192 Chavey, C. *et al.* Matrix metalloproteinases are differentially expressed in adipose tissue during obesity and modulate adipocyte differentiation. *J Biol Chem* **278**, 11888-11896, doi:10.1074/jbc.M209196200 (2003).

193 Novak, M. L. & Koh, T. J. Phenotypic transitions of macrophages orchestrate tissue repair. *Am J Pathol* **183**, 1352-1363, doi:10.1016/j.ajpath.2013.06.034 (2013).

194 Hassanshahi, A. *et al.* Macrophage-Mediated Inflammation in Skin Wound Healing. *Cells* **11**, doi:10.3390/cells11192953 (2022).

195 Riley, G. Tendinopathy--from basic science to treatment. *Nat Clin Pract Rheumatol* **4**, 82-89, doi:10.1038/ncprheum0700 (2008).

196 Coombes, B. K., Bisset, L. & Vicenzino, B. Efficacy and safety of corticosteroid injections and other injections for management of tendinopathy: a systematic review of randomised controlled trials. *Lancet* **376**, 1751-1767, doi:10.1016/s0140-6736(10)61160-9 (2010).

Figures

Figure 1 The results of Mendelian randomization analysis

a. The linear Manhattan plot of SNPs in T2DM

b. The linear Manhattan plot of SNPs in tendon repair

c. The density plot of SNPs within 1Mb window size related to T2DM

d. The density plot of SNPs within 1Mb window size related to tendon repair

e. The forest plot of the causal effect of T2DM on tendon repair

f. The funnel plot of the causal effect of T2DM on tendon repair

g. The leave-one-out plot of the causal effect of T2DM on tendon repair

h. The scatter plot of the causal effect of T2DM on tendon repair

Figure 2 Data preprocessing and batch effect elimination results for the dataset of T2DM and TI

a. Clustering dendrogram of T2DM datasets (GSE15932, GSE25724, GSE29221) before batch effect removal

b. Clustering dendrogram of T2DM datasets after batch effect removal

c. PCA plot of the merged T2DM datasets (data1, data2, data3)

d. Clustering dendrogram of the TI dataset (GSE26051) before batch effect removal

e. Clustering dendrogram of the TI dataset after batch effect removal

f. PCA plot of the processed TI dataset

Figure 3 Identification of DEGs in T2DM and TI Samples

a. Volcano plot of DEGs in T2DM dataset

b. Waterfall plot of top-ranked DEGs in T2DM dataset

c. Heatmap of top 20 DEGs in T2DM samples

d. Volcano plot of DEGs in TI dataset

e. Waterfall plot of top-ranked DEGs in TI dataset

f. Heatmap of top 20 DEGs in TI samples

Figure 4 WGCNA of merged T2DM dataset

a. Sample dendrogram and trait heatmap

b. The results for selecting the soft threshold power based on the scale-free topological fit index and the average connectivity

c. Cluster dendrogram of co-expression modules

d. Heatmap of the correlation between module feature genes and T2DM status

e. TOM-based network heatmap of selected genes

f. Scatter plot of selected modules (blue and turquoise) in T2DM

Figure 5 WGCNA of TI dataset

a. Sample dendrogram and trait heatmap

b. The results for selecting the soft threshold power based on the scale-free topological fit index and the average connectivity

c. Cluster dendrogram of co-expression modules

d. Heatmap of the correlation between module feature genes and TI status

e. TOM-based network heatmap of selected genes

f. Scatter plot of selected modules (brown, magenta, pink, blue and grey) in TI

Figure 6 GO/KEGG analysis

a. Venn diagram of common genes derived from DEGs and WGCNA modules related to T2DM and TI

b. Network of common genes

c. Circular diagram of GO analysis

d. Bar graph of GO analysis

e. Sankey diagram of KEGG analysis

Figure 7 Hub genes identified through PPI network analysis

a. PPI network of the 15 common genes

b. Network of the 15 common genes based on four algorithms (Degree, MCC, Closeness and Betweenness)

c. Venn diagram from four algorithms (Degree, MCC, Closeness and Betweenness)

Figure 8 Identification of hub genes based on machine learning models

a. Reserve cumulative distribution of residual plot for machine learning models

b. Boxplots of residual for machine learning models

c. Residual histograms of machine learning models

d. Feature importance comparison across machine learning models

e. Plots of LASSO regression coefficient

f. Plots of SVM

g. Important genes map based on RF model

h. Plot of RF

i. Venn diagram based on three machine learning models (LASSO, SVM and RF)

Figure 9 Immune infiltration analysis

a. Stacked plot of immunocytes proportions in the TI dataset

b. Heatmap of immunocytes infiltration in each group of the TI dataset

c. Boxplot of the infiltration level rankings of 22 immunocyte types in the TI dataset

d. Correlation butterfly chart between hub genes and 22 types of immunocytes

e. Heatmap of immunocyte infiltration levels in hub genes

f. Stacked plot of immunocytes proportions in the T2DM dataset

g. Heatmap of immunocytes infiltration in each group of the T2DM dataset

h. Boxplot of the infiltration level rankings of 22 immunocyte types in the T2DM dataset

i. Correlation butterfly chart between hub genes and 22 types of immunocytes

j. Heatmap of immunocyte infiltration levels in hub genes

Figure 10 GSEA analysis

a. Enrichment plots of TTN in muscle organ development

b. Ridge plot of the top 10 for TTN enrichment rankings

c. Enrichment plots of the top 10 for TTN enrichment rankings

d. Enrichment plots of MYL2 in striated muscle contraction

e. Enrichment plots of the top 10 for MYL2 enrichment rankings

f. Ridge plot of the top 10 for MYL2 enrichment rankings

g. Enrichment plots of MYL3 in skeletal muscle organ development

h. Enrichment plots of the top 10 for MYL3 enrichment rankings

i. Ridge plot of the top 10 for MYL3 enrichment rankings

Tables

Table 1 The SNPs of exposure factors (T2DM) from GWAS database

Table 2 The SNPs of outcome factors (TI) from GWAS database

Table 3 MR analysis results

Table 4 GSE15932 series matrix file

Table 5 GSE25724 series matrix file

Table 6 GSE29221 series matrix file

Table 7 GSE26051 series matrix file

Table 8 Matrix file after merging and processing the GSE15932, GSE25724, and GSE29221 datasets

Table 9 The results of the DEGs in T2DM

Table 10 The results of the DEGs in TI

Table 11 WGCNA results of the TI dataset (genesmodule_blue)

Table 12 WGCNA results of the TI dataset (genesmodule_turquoise)

Table 13 WGCNA results of the merged T2DM dataset (genesmodule_blue)

Table 14 WGCNA results of the merged T2DM dataset (genesmodule_brown)

Table 15 WGCNA results of the merged T2DM dataset (genesmodule_grey)

Table 16 WGCNA results of the merged T2DM dataset (genesmodule_magenta)

Table 17 WGCNA results of the merged T2DM dataset (genesmodule_pink)

Table 18 GO enrichment analysis results

Table 19 KEGG enrichment analysis results

Table 20 Prediction results of the LASSO model

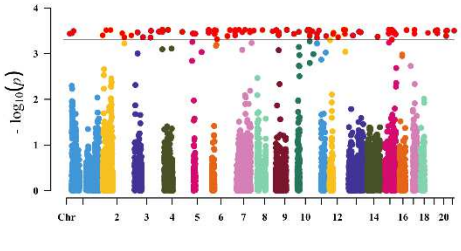
Table 21 Prediction results of the SVM model

Table 22 The GSEA enrichment analysis results of TTN

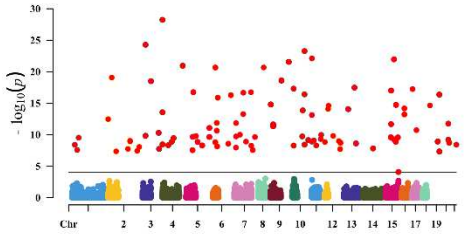
Table 23 The GSEA enrichment analysis results of MYL2

Table 24 The GSEA enrichment analysis results of MYL3

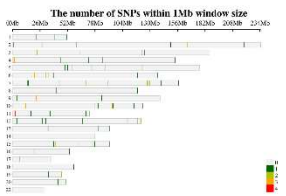
Figure 1



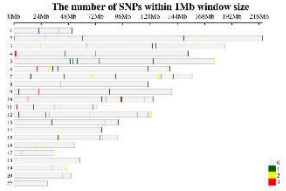
a



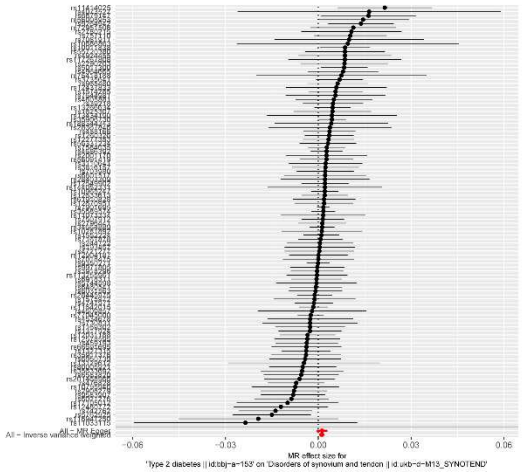
b



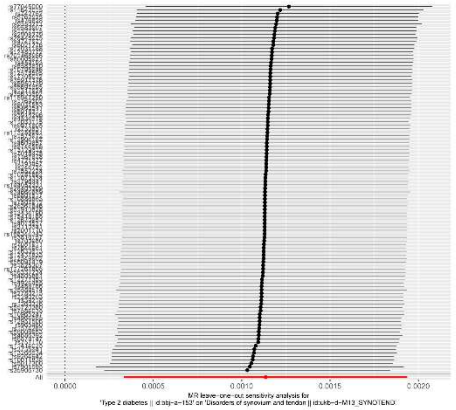
c



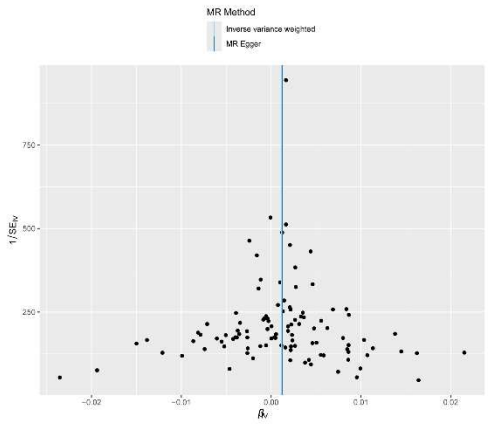
d



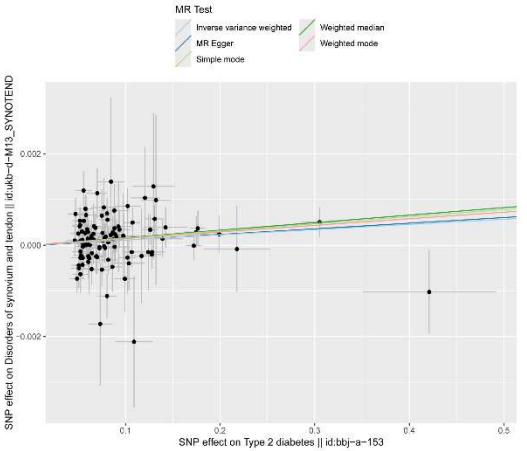
e



g



f



h

Figure 2

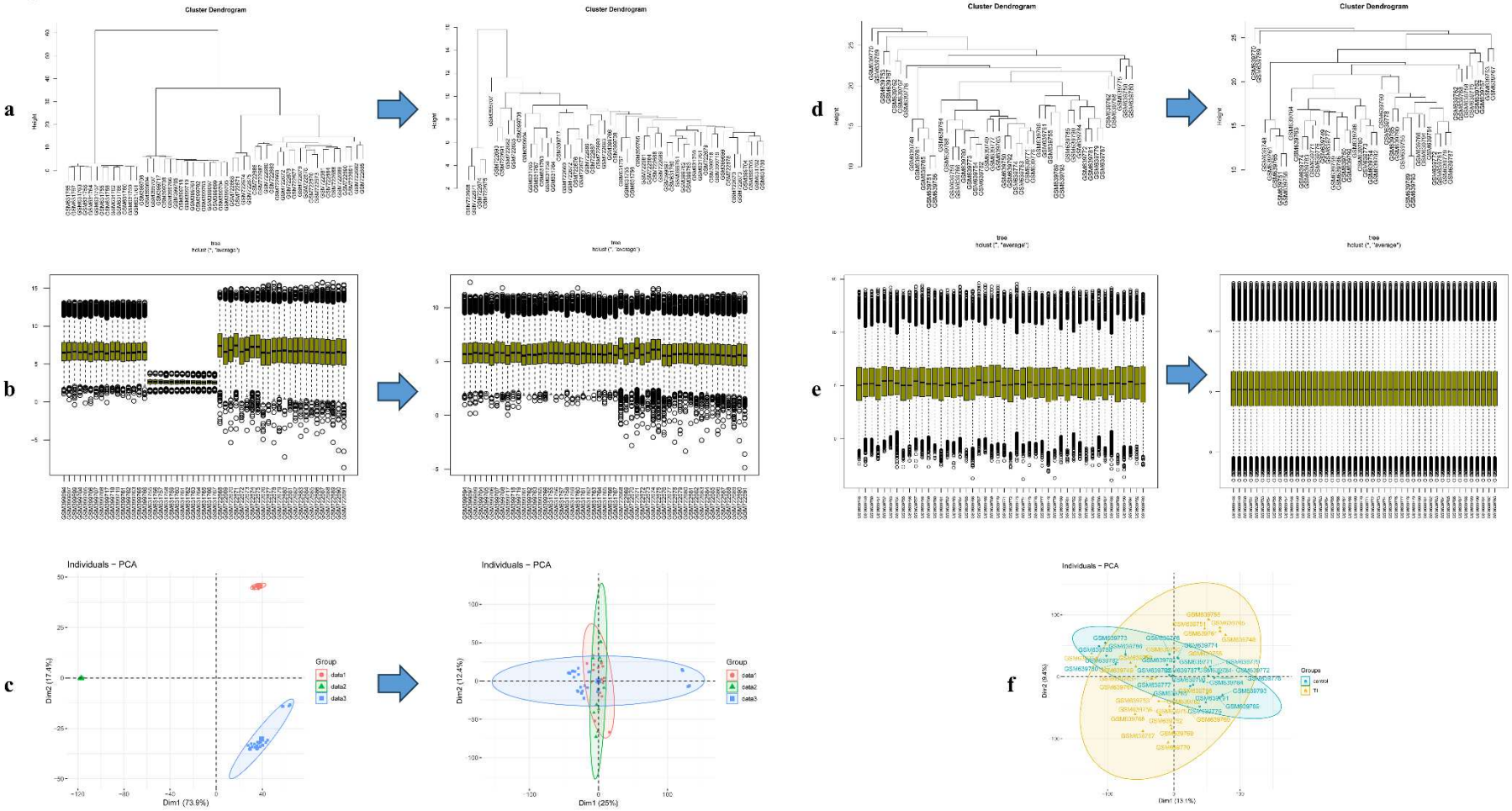

a
Cluster Dendrogram
Cluster Dendrogram
d
Cluster Dendrogram
Cluster Dendrogram
b
e
c
Individuals – PCA
Group
data1
data2
data3
Dim1 (73.9%)
Dim2 (17.4%)
Individuals – PCA
Dim1 (25%)
Dim2 (12.4%)
f
Individuals – PCA
Groups
control
TI
Dim1 (13.1%)
Dim2 (9.4%)

Figure 3

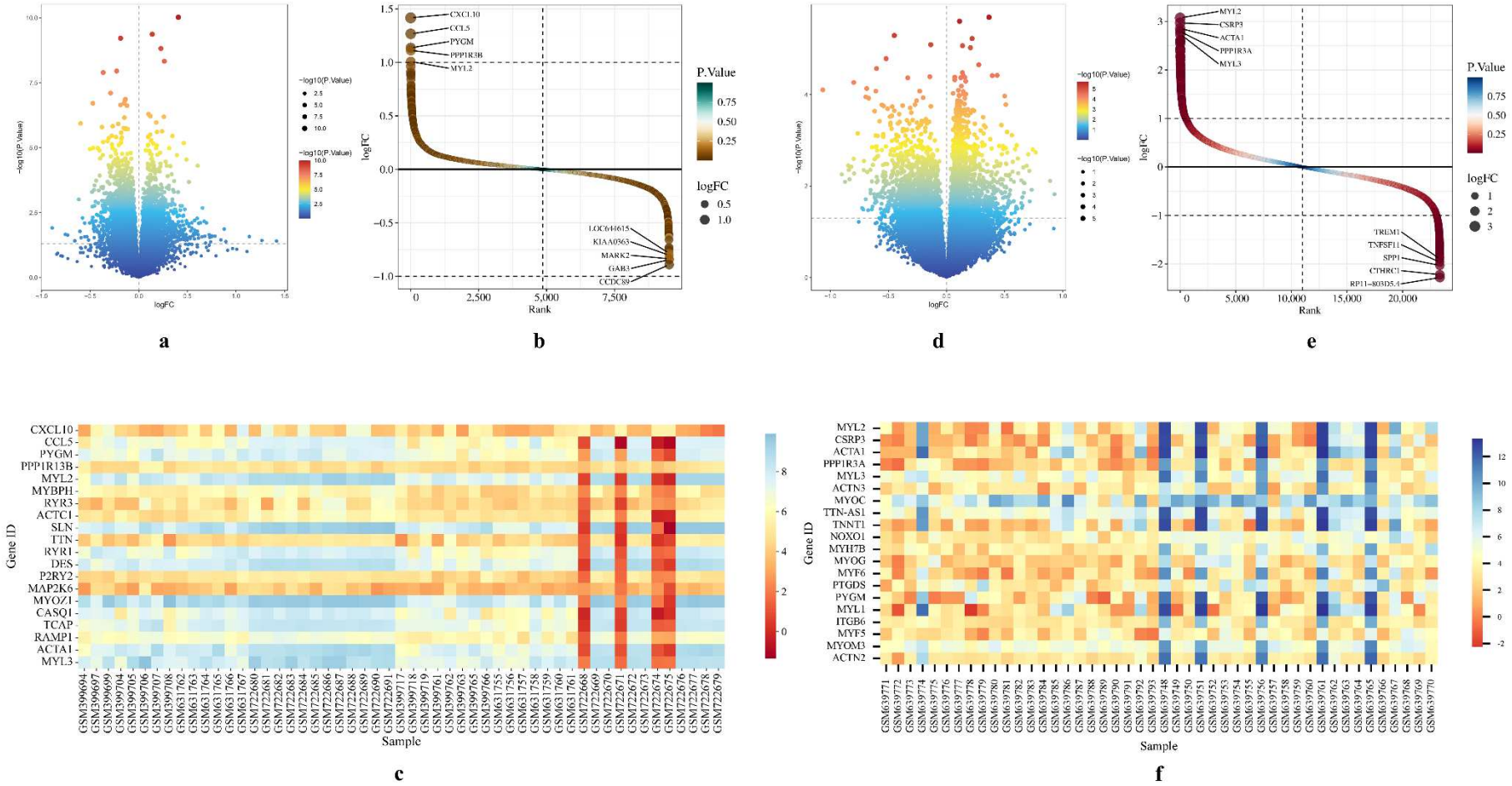

a
b
d
e
c
f
CXCL10
CCL5
PYGM
PPP1R13B
MYL2
MYBPH
RYR3
ACTC1
SLN
TTN
RYR1
DES
P2RY2
MAP2K6
MYOZ1
CASQ1
TCAP
RAMP1
ACTA1
MYL3
MYL2
CSRP3
ACTA1
PPP1R3A
MYL3
ACTN3
MYOC
TTN-AS1
TNNT1
NOXO1
MYH7B
MYOG
MYF6
PTGDS
PYGM
MYL1
ITGB6
MYF5
MYOM3
ACTN2
Gene ID
Sample
logFC
Rank
P.Value
−log10(P.Value)

Figure 4

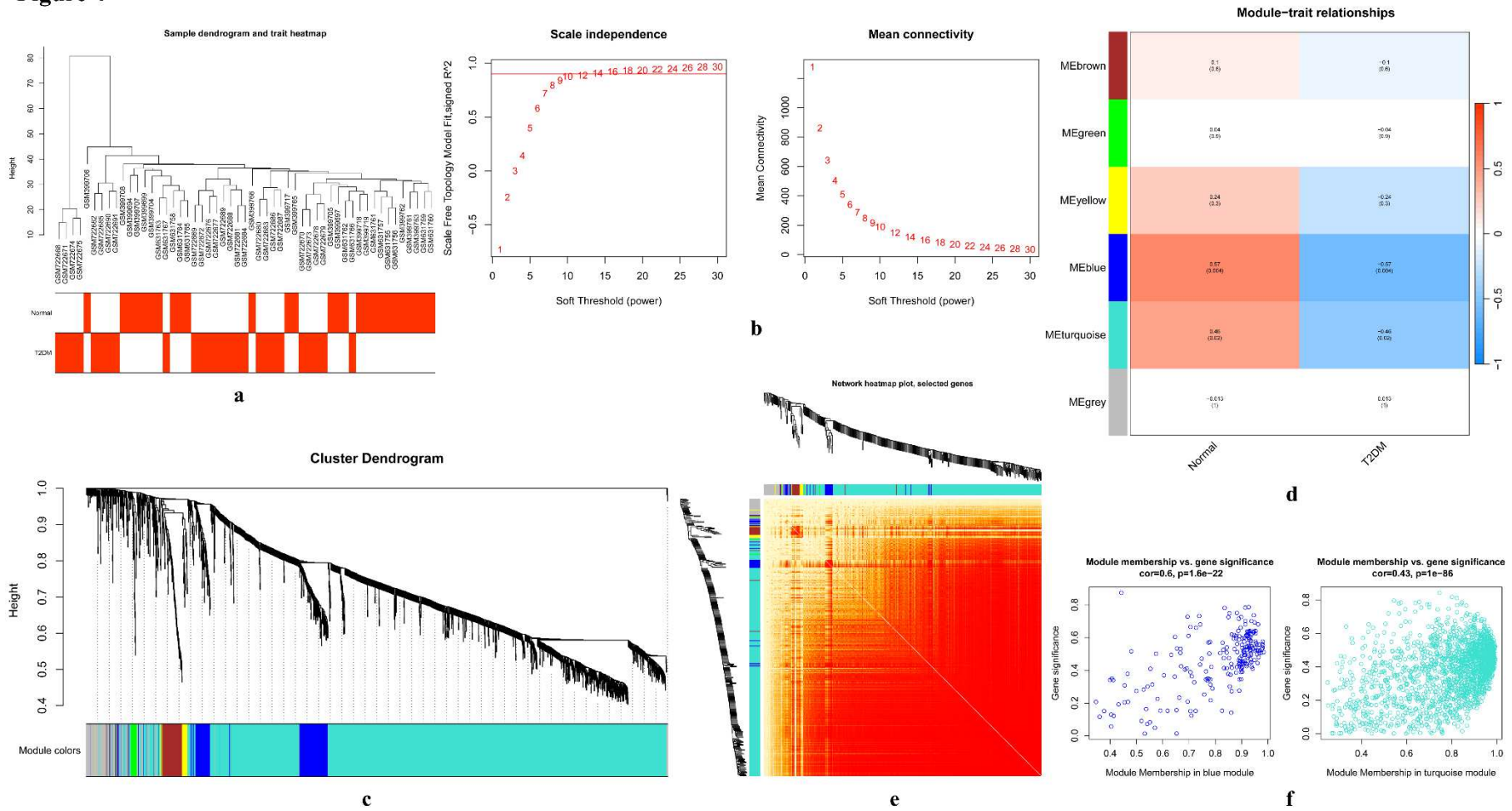

Sample dendrogram and trait heatmap
Height
Normal
T2DM
a
Scale independence
Scale Free Topology Model Fit,signed R^2
Soft Threshold (power)
Mean connectivity
Mean Connectivity
Soft Threshold (power)
b
Module−trait relationships
MEbrown
MEgreen
MEyellow
MEblue
MEturquoise
MEgrey
Normal
T2DM
d
Cluster Dendrogram
Height
Module colors
c
Network heatmap plot, selected genes
e
Module membership vs. gene significance
cor=0.6, p=1.6e−22
Gene significance
Module Membership in blue module
Module membership vs. gene significance
cor=0.43, p=1e−86
Gene significance
Module Membership in turquoise module
f

Figure 5

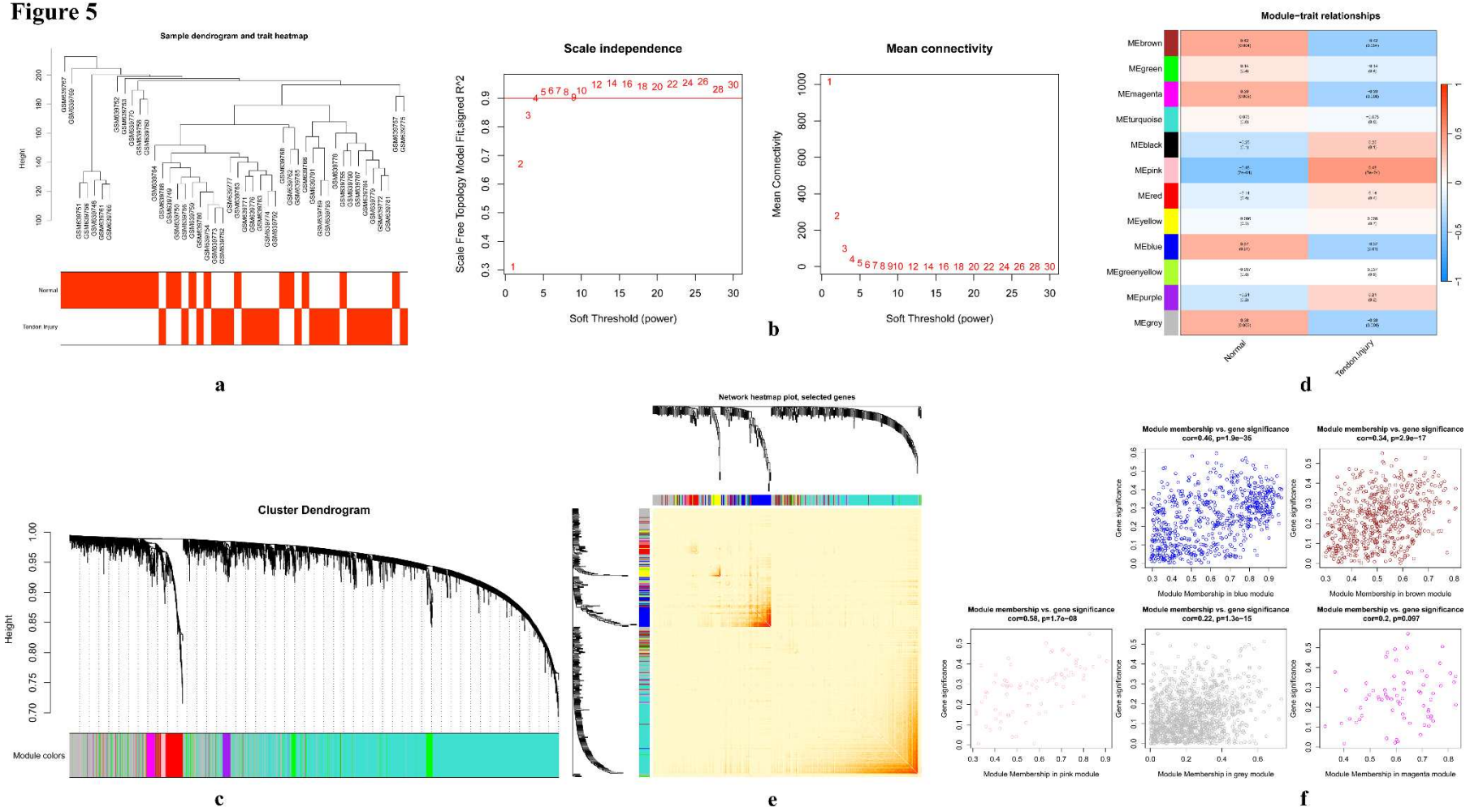

Sample dendrogram and trait heatmap
Height
Normal
Tendon Injury
a
Scale independence
Scale Free Topology Model Fit,signed R^2
Soft Threshold (power)
Mean connectivity
Mean Connectivity
Soft Threshold (power)
b
Module−trait relationships
MEbrown
MEgreen
MEmagenta
MEturquoise
MEblack
MEpink
MEred
MEyellow
MEblue
MEgreenyellow
MEpurple
MEgrey
Normal
Tendon Injury
d
Cluster Dendrogram
Height
Module colors
c
Network heatmap plot, selected genes
e
Module membership vs. gene significance
cor=0.46, p=1.9e−35
Gene significance
Module Membership in blue module
Module membership vs. gene significance
cor=0.34, p=2.9e−17
Module Membership in brown module
Module membership vs. gene significance
cor=0.58, p=1.7e−08
Module Membership in pink module
Module membership vs. gene significance
cor=0.22, p=1.3e−15
Module Membership in grey module
Module membership vs. gene significance
cor=0.2, p=0.097
Module Membership in magenta module
f

Figure 6

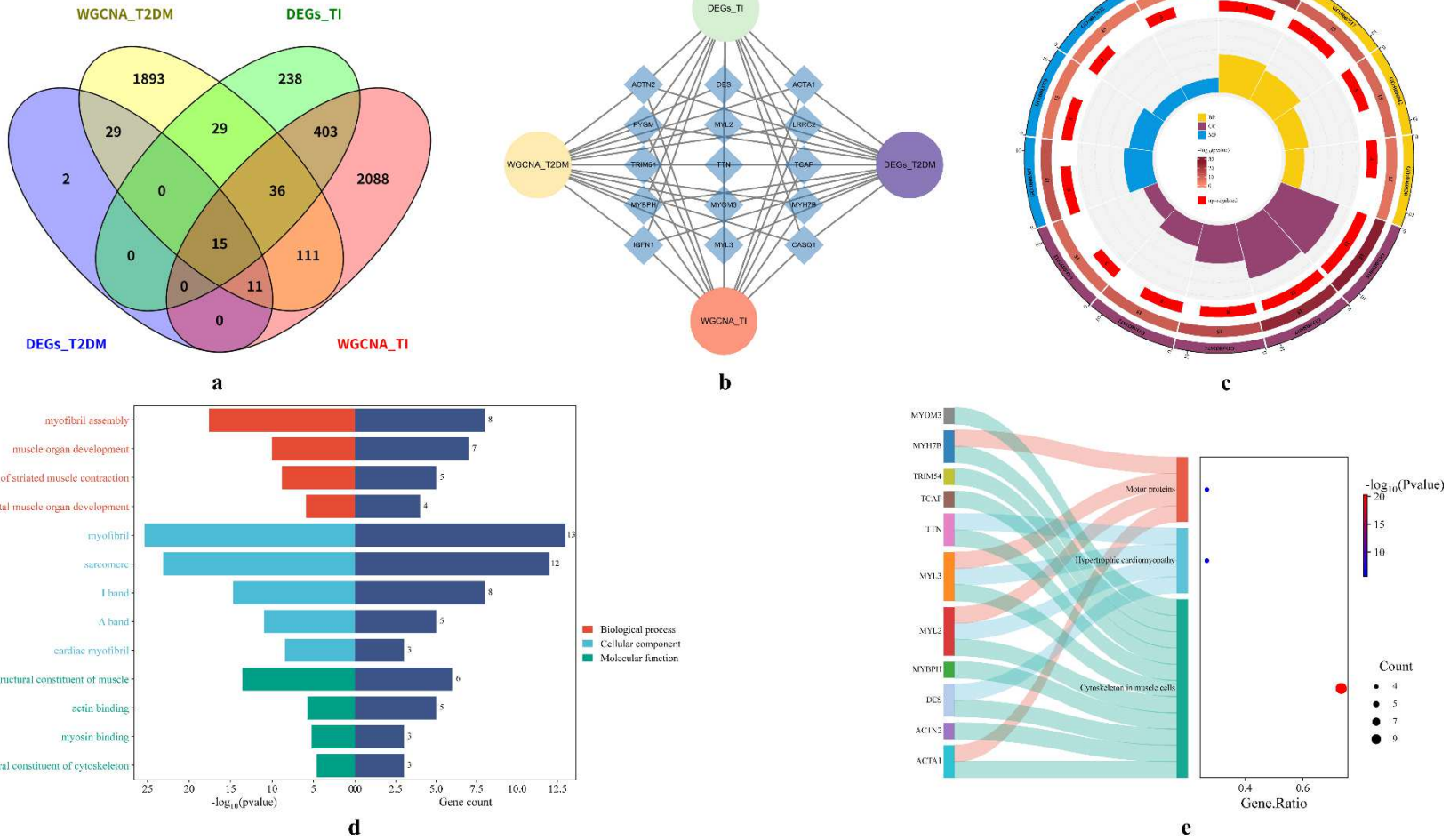

WGCNA_T2DM
DEGs_TI
1893
238
29
29
403
2
0
36
2088
0
15
111
0
11
0
DEGs_T2DM
WGCNA_TI
a
DEGs_TI
ACTN2
DES
ACTA1
PYGM
MYL2
LRRC2
WGCNA_T2DM
TRIM63
TTN
TCAP
DEGs_T2DM
MYBPH
MYOM3
MYH7B
IGFN1
MYL3
CASQ1
WGCNA_TI
b
c
myofibril assembly
muscle organ development
regulation of striated muscle contraction
skeletal muscle organ development
myofibril
sarcomere
I band
A band
cardiac myofibril
structural constituent of muscle
actin binding
myosin binding
structural constituent of cytoskeleton
Biological process
Cellular component
Molecular function
-log10(pvalue)
Gene count
d
MYOM3
MYH7B
TRIM54
TCAP
TTN
MYL3
MYL2
MYBPH
DES
ACTN2
ACTA1
Motor proteins
Hypertrophic cardiomyopathy
Cytoskeleton in muscle cells
-log10(Pvalue)
Count
Gene.Ratio
e

Figure 7

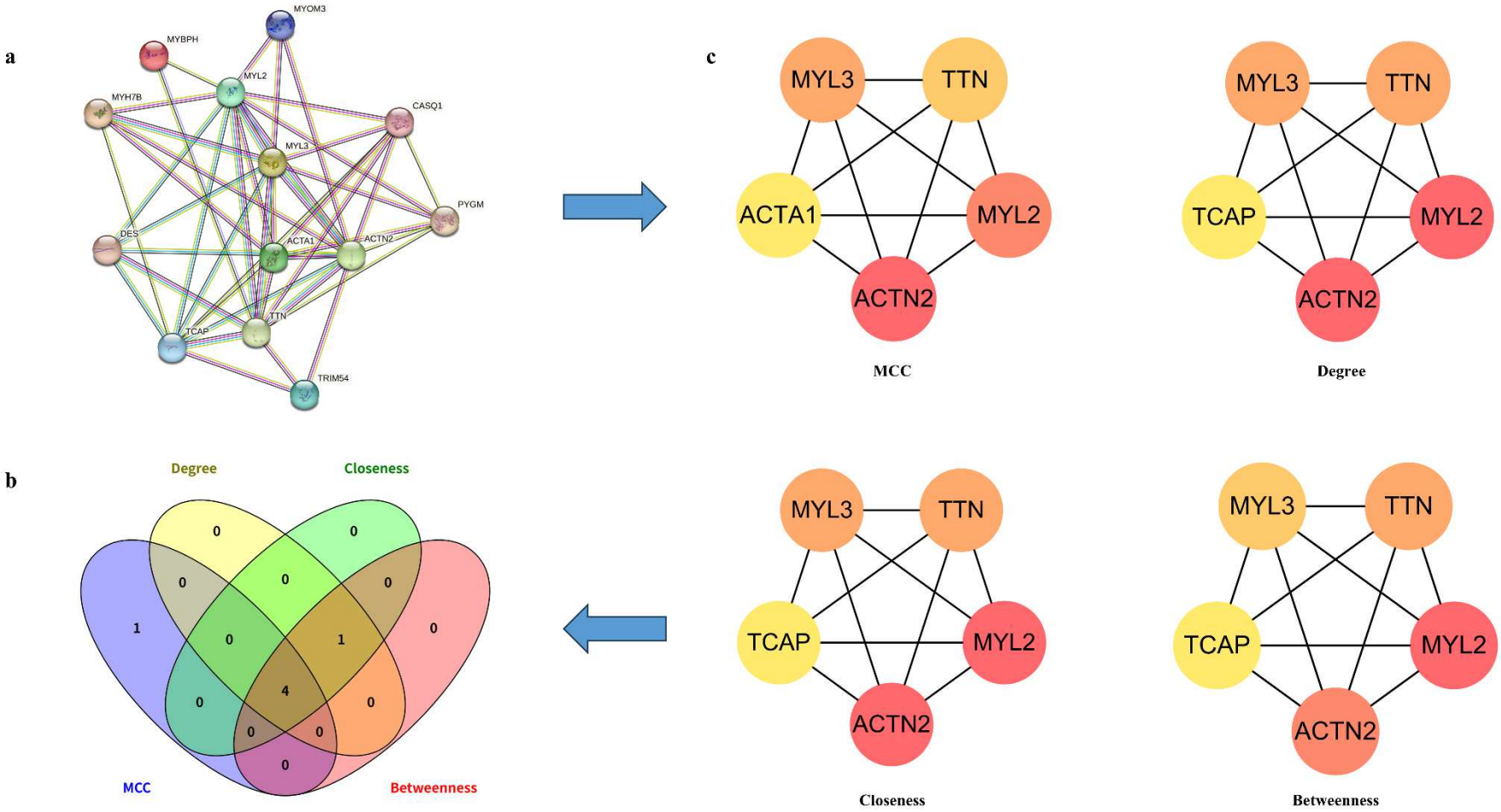

a
MYOM3
MYBPH
MYL2
MYH7B
CASQ1
MYL3
PYGM
DES
ACTA1
ACTN2
TTN
TCAP
TRIM54
b
Degree
Closeness
MCC
Betweenness
0
1
4
c
MYL3
TTN
ACTA1
MYL2
ACTN2
MCC
TCAP
Degree
Closeness
Betweenness

Figure 8

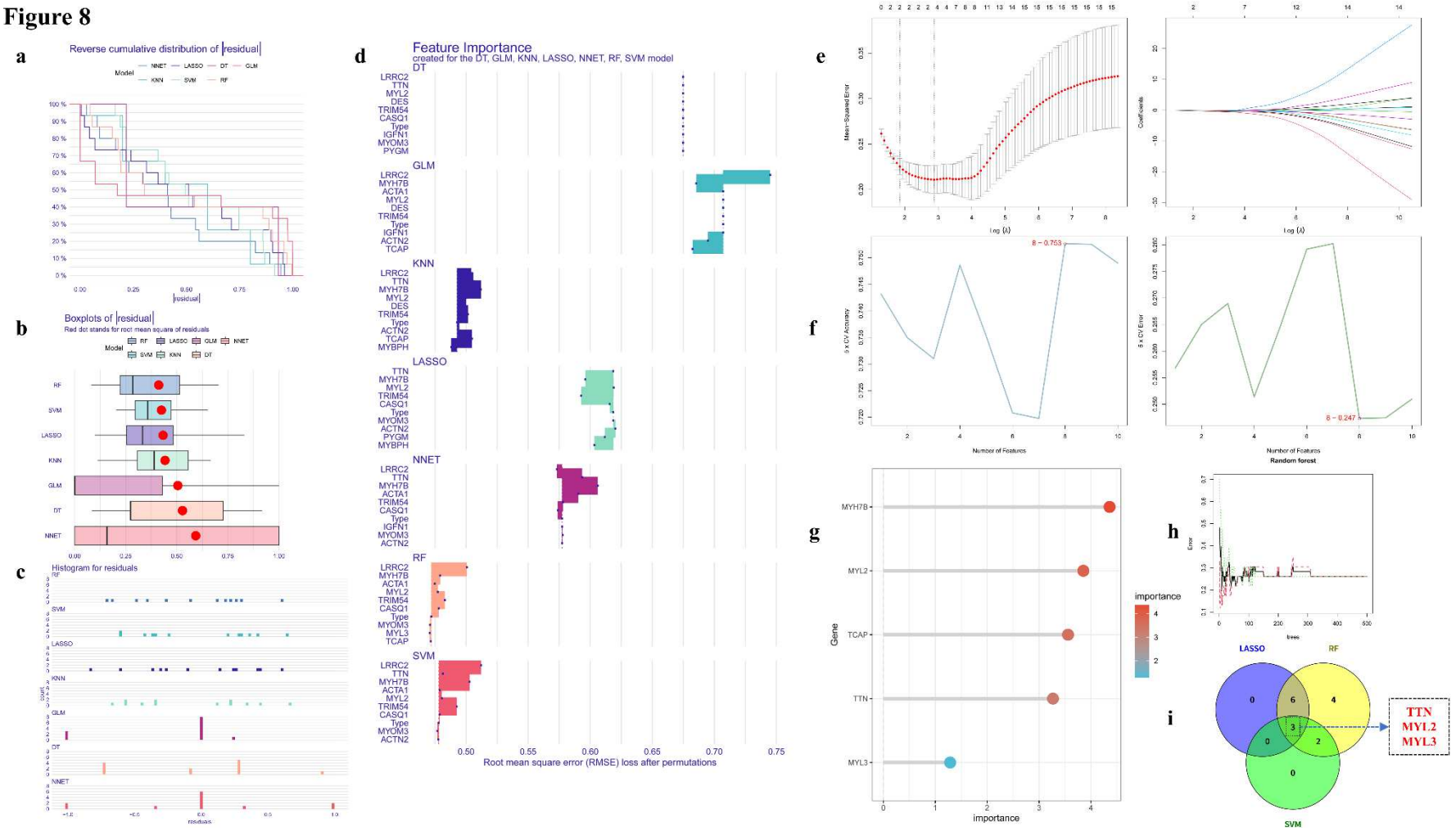

a
Reverse cumulative distribution of |residual|
Model
NNET
LASSO
DT
GLM
KNN
SVM
RF
100 %
90 %
80 %
70 %
60 %
50 %
40 %
30 %
20 %
10 %
0 %
0.00
0.25
0.50
0.75
1.00
|residual|
b
Boxplots of |residual|
Red dot stands for root mean square of residuals
Model
RF
LASSO
GLM
NNET
SVM
KNN
DT
c
Histogram for residuals
RF
SVM
LASSO
KNN
GLM
DT
NNET
count
residuals
d
Feature Importance
created for the DT, GLM, KNN, LASSO, NNET, RF, SVM model
DT
LRRC2
TTN
MYL2
DES
TRIM54
CASQ1
Type
IGFN1
MYOM3
PYGM
GLM
LRRC2
MYH7B
ACTA1
MYL2
DES
TRIM54
Type
IGFN1
ACTN2
TCAP
KNN
LRRC2
TTN
MYH7B
MYL2
DES
TRIM54
Type
ACTN2
TCAP
MYBPH
LASSO
TTN
MYH7B
MYL2
TRIM54
CASQ1
Type
MYOM3
ACTN2
PYGM
MYBPH
NNET
LRRC2
TTN
MYH7B
ACTA1
TRIM54
CASQ1
Type
IGFN1
MYOM3
ACTN2
RF
LRRC2
MYH7B
ACTA1
MYL2
TRIM54
CASQ1
Type
MYOM3
MYL3
TCAP
SVM
LRRC2
TTN
MYH7B
ACTA1
MYL2
TRIM54
CASQ1
Type
MYOM3
ACTN2
0.50
0.55
0.60
0.65
0.70
0.75
Root mean square error (RMSE) loss after permutations
e
Mean−Squared Error
Log (λ)
Coef.Gamlr
Log (λ)
f
5 x CV Accuracy
8 − 0.753
Number of Features
5 x CV Error
8 − 0.247
Number of Features
Random forest
g
MYH7B
MYL2
TCAP
TTN
MYL3
Gene
importance
importance
h
Error
trees
i
LASSO
RF
SVM
0
6
4
3
0
2
0
TTN
MYL2
MYL3

Figure 9

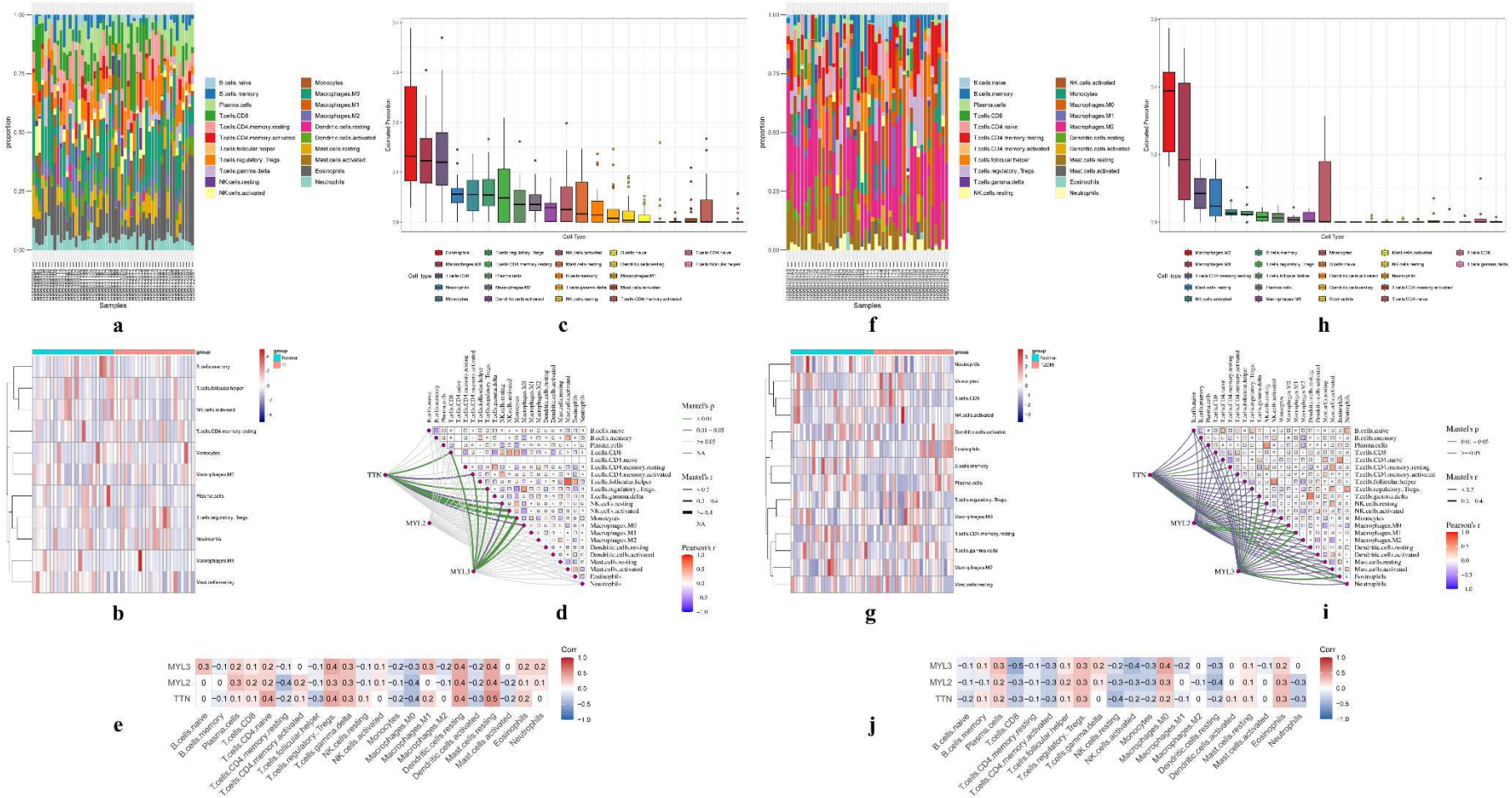
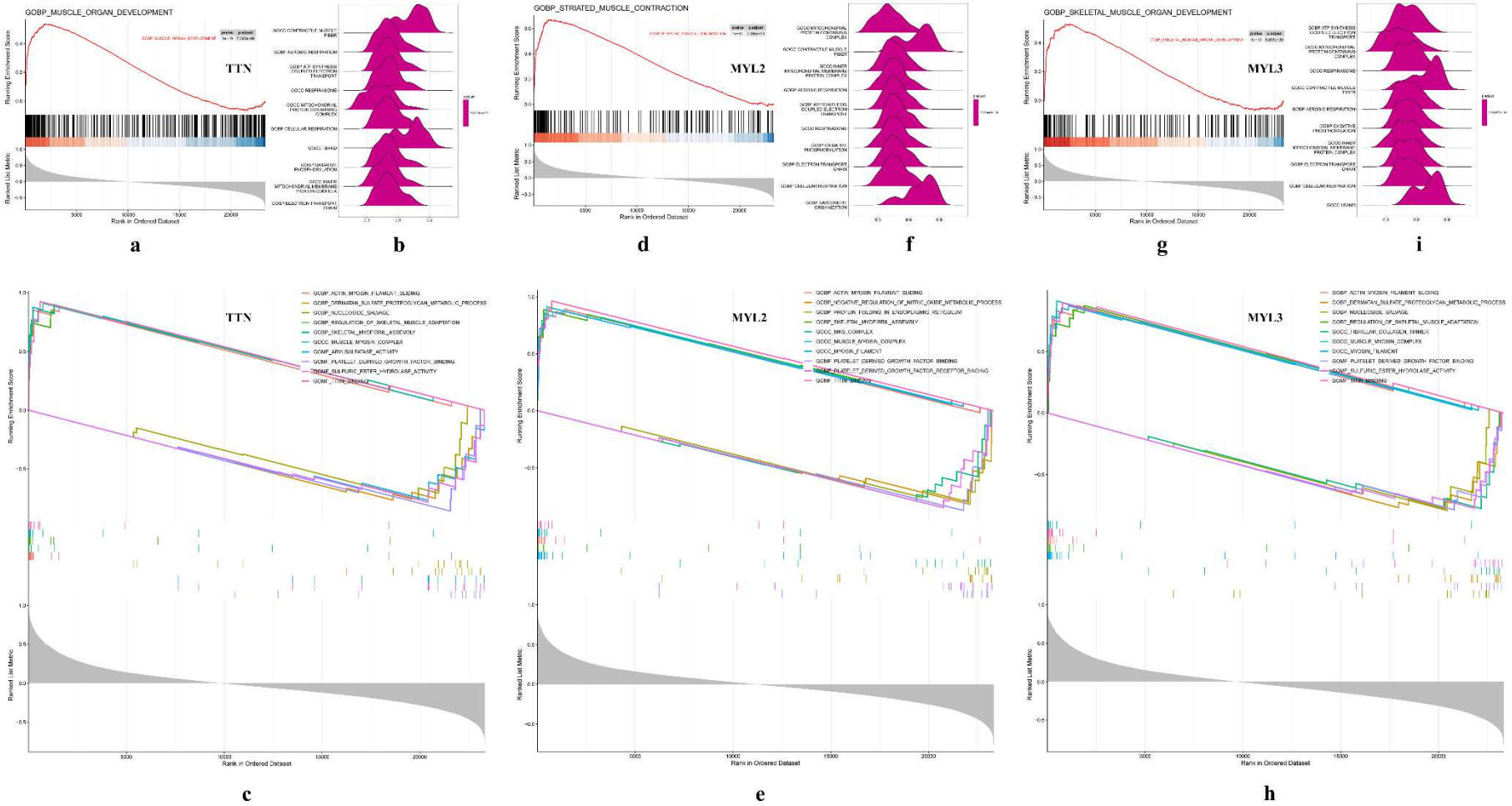


Figure 10


GOBP_MUSCLE_ORGAN_DEVELOPMENT
TTN
a
b
GOBP_STRIATED_MUSCLE_CONTRACTION
MYL2
d
f
GOBP_SKELETAL_MUSCLE_ORGAN_DEVELOPMENT
MYL3
g
i
TTN
c
MYL2
e
MYL3
h
Running Enrichment Score
Ranked List Metric
Rank in Ordered Dataset

Supplementary Files

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

- Table1.csv
- Table2.csv
- Table3.csv
- Table5.csv
- Table4.csv
- Table6.csv
- Table9.csv
- Table8.csv
- Table10.csv
- Table11.csv
- Table7.csv
- Table12.csv
- Table13.csv
- Table14.csv
- Table15.csv
- Table16.csv
- Table17.csv
- Table18.xlsx
- Table19.xlsx
- Table20.csv
- Table22.csv
- Table23.csv
- Table24.csv
- Table21.csv
- Tablecaptions.docx