

RESEARCH ARTICLE

Screening of key biomarkers of tendinopathy based on bioinformatics and machine learning algorithms

Ya xi Zhu^{1,2} , Jia qiang Huang¹ , Yu yang Ming^{2,3}, Zhao Zhuang⁴, Hong Xia^{3*} 

1 District 1, Department of Orthopedics, Xiangtan Central Hospital, Yuhu District, Xiangtan City, Hunan Province, China, **2** Nanhua University, Hengyang City, Hunan Province, China, **3** Department of Orthopedics, Xiangtan Central Hospital, Yuhu District, Xiangtan City, Hunan Province, China, **4** Academy of Anesthesiology, Weifang Medical University, Weifang, China

 These authors contributed equally to this work.

* xhxi1002@outlook.com



OPEN ACCESS

Citation: Zhu Yx, Huang Jq, Ming Yy, Zhuang Z, Xia H (2021) Screening of key biomarkers of tendinopathy based on bioinformatics and machine learning algorithms. PLoS ONE 16(10): e0259475. <https://doi.org/10.1371/journal.pone.0259475>

Editor: Qi Zhao, University of Science and Technology Liaoning, CHINA

Received: August 14, 2021

Accepted: October 19, 2021

Published: October 29, 2021

Copyright: © 2021 Zhu et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE106292> <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE26051> <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE167226>.

Funding: This work was supported by Xiangtan Science and Technology Planning Project (Project No.: SF-YB20181006) and Xiangtan Arthroscopy Minimally Invasive Diagnosis and Treatment Clinical Medical Technology Demonstration Base Funding Project (Project No. SF-LCYL20191003).

Abstract

Tendinopathy is a complex multifaceted tendinopathy often associated with overuse and with its high prevalence resulting in significant health care costs. At present, the pathogenesis and effective treatment of tendinopathy are still not sufficiently elucidated. The purpose of this research is to intensely explore the genes, functional pathways, and immune infiltration characteristics of the occurrence and development of tendinopathy. The gene expression profile of GSE106292, GSE26051 and GSE167226 are downloaded from GEO (NCBI comprehensive gene expression database) and analyzed by WGCNA software bag using R software, GSE26051, GSE167226 data set is combined to screen the differential gene analysis. We subsequently performed gene enrichment analysis of Gene Ontology (GO) and "Kyoto Encyclopedia of Genes and Genomes" (KEGG), and immune cell infiltration analysis. By constructing the LASSO regression model, Support vector machine (SVM-REF) and Gaussian mixture model (GMMs) algorithms are used to screen, to identify early diagnostic genes. We have obtained a total of 171 DEGs through WGCNA analysis and differentially expressed genes (DEGs) screening. By GO and KEGG enrichment analysis, it is found that these dysregulated genes were related to mTOR, HIF-1, MAPK, NF-κB and VEGF signaling pathways. Immune infiltration analysis showed that M1 macrophages, activated mast cells and activated NK cells had infiltration significance. After analysis of THE LASSO SVM-REF and GMMs algorithms, we found that the gene *MACROD1* may be a gene for early diagnosis. We identified the potential of tendon disease early diagnosis way and immune gene regulation *MACROD1* key infiltration characteristics based on comprehensive bioinformatics analysis. These hub genes and functional pathways may as early biomarkers of tendon injuries and molecular therapy level target is used to guide drug and basic research.

Competing interests: The authors have declared that no competing interests exist.

Background

Tendinopathy is usually described as pathological changes in injured and diseased tendons, which in turn lead to limb pain and functional decline. It is characterized by abnormalities in the molecular structure, composition and cell matrix of the tendon [1]. In recent years, the prevalence of tendinopathy has gradually increased, and some tendinopathy patients have long-term or permanent limb dysfunction and loss [2]. Tendinopathy is more common in limbs and accounts for 30%-50% of muscle skeletal system and locomotor system diseases [3,4]. It is well known that the causes of tendinopathy include internal and external factors. External factors mainly cause acute tendon injury, while chronic tendinopathy is the result of both external and internal factors [5]. Previous studies have suggested that the imbalance of normal homeostasis in tendon tissue, including immune cell infiltration, stromal cell dysfunction, cell apoptosis, oxidative stress, and stromal dysfunction, and these comprehensive conjunct factors lead to the early pathological changes of tendon [1]. Unfortunately, the current diagnosis of tendinopathy has reached the middle and late stage, but the treatment plans are mostly targeted at the later stage of the disease, including centrifugal exercise [6,7] drug injection and surgical treatment [8], etc. Studies have shown that the efficacy and evidence of most treatment plans for tendinopathy are still insufficient. Therefore, early diagnosis and treatment are the key to complete recovery of the disease. At present, the hub genes and pathways in the early stage of tendinopathy are blank studied, and understanding the key pathways that affect the regulation and dynamic homeostasis of the extracellular matrix in the early stage is critical for future targeted therapies of tendinopathy. Therefore, more in-depth studies are urgently needed to elucidate the hub genes.

With the development of high-throughput omics, bioinformatics analysis has become an important tool for identifying potentially hub genes and signaling pathways in a variety of diseases [9,10]. At present, the research progress of bioinformatics analysis of tendinopathy gene expression has not been insufficient, and only a small amount of lncRNA and mRNA related to tendinopathy has been studied. Among them, studies have analyzed the related characteristics of lncRNAs and mRNAs in rotator cuff tendinopathy, including NONHSAT209114.1, ENST00000577806, NONHSAT168464.1, PLK2, TMEM214 and IGF2 [11,12]. Unfortunately, the use of bioinformatics to analyze DEGs related to tendinopathy still needs further research. In this study, we downloaded the original data from the NCBI Gene Expression Synthesis Database (GEO) and constructed a co-expression network through the data set GSE106292 and determined the correlation gene module. By processing the data set GSE26051 and GSE167226 carry on incorporative and differential gene screening. We take the intersection of differential genes and gene modules to obtain common differential genes, and then performed gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis. Then, we used the whole gene expression data of tendon tissues for immune cell infiltration analysis. By establishing LASSO regression model [13], Machine learning algorithm Support vector Machine (SVM-REF) [14], Clustering algorithm Gaussian Mixture model(GMMs) [15], Screening of genes for early diagnosis. The objective of this study was to identify hub genes and pathways that contribute to the occurrence and development of tendinopathy at the molecular level, and to identify candidate genes for early diagnosis and treatment targets.

Materials and method

Data collection and download

GEO (Gene Expression Omnibus) database (<http://www.ncbi.nlm.nih.gov/geo/>) is an international public database, used to store and provide free microarray, second-generation

sequencing and high-throughput functional genome data sets [16]. We searched and downloaded the GSE106292, the GSE26051, the GSE167226 data set using the R software GEO database [17–19]. The GSE106292 data sets included the gene expression profiles of 35 cases of tendon, bone, muscle, cartilage and ligament. The GSE26051 data sets included gene expression profiles from 23 patients with chronic tendonopathy and 23 normal tendons; the GSE167226 data sets included gene expression profiles from 19 patients with tendonopathy. Institutional Review Board approval was not required because the study was based on a public database and did not involve animal or human studies.

Data preprocessing and DEG screening

The original data of GSE106292, GSE26051 and GSE167226 datasets were corrected and normalized by using log₂ of the R software. Bioconductor platform (<http://www.bioconductor.org/>) gene comments file is used to probe the matrix [20]. At the same time, the expression matrix of GSE26051 and GSE167226 datasets was merged, and the differences between batches were eliminated using the Limma software package. The differences of gene expression profiles between GSE26051 and GSE167226 datasets were analyzed by using the Limma function and the Remove Batch Effect function using Limma [21]. By deleting the genes with too low expression value, the expression profile value was converted to log₂-counts per million (logCPM), and linear regression was performed to construct the comparison matrix. Where each row represents the gene name and each column represents the sample name of this study. Based on Bayesian calculation of T-values, F-values and log-odds, the eligible differential genes were screened for $|\log_2(\text{FC})| > 1$ and PValue < 0.05, and the data was visualized by plotting volcano plots using the ggplot2 program package [22].

Gene WGCNA analysis

The expression matrix of GSE106292 data sets was selected to construct the expression profile of potential related genes in tendinopathy. The WGCNA software package is an open source and widely used method for identifying co-expression networks in R software [23]. The Pearson correlation coefficient is calculated to construct the correlation matrix, and the soft threshold function is used to transform the correlation matrix into a weighted adjacency matrix. In order to obtain a balanced co-expressed network between scale independence and mean connectivity, a soft connectivity algorithm is used to calculate the scale independence and mean connectivity with different powers. We transform the adjacency matrix into topological overlap matrix (TOM). According to 1-tom as distance measurement, we classified gene clustering as co-expression modules, with a depth split value of 2, a minimum size cutoff value of 20, and a maximum module size of 5,000. To determine the association between co-expressed modules and clinical features, we cut the tree into different modules (minimum number of genes in a module is 30) using a dynamic shearing method by setting the soft threshold power to 1–10 and $\beta = 16$ (scale-free $R^2 = 0.83$) and cut the height to 0.3, using 0.25 as the merging threshold (shearing height), under which modules will be merged (correlation (modules with a correlation higher than 0.75 will be merged).

GO and KEGG enrichment analysis

Gene ontology (GO) analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis were calculated using the R package Clusterprofiler program [24]. It includes annotation mining the function and path of the module, and identifying the dysfunctional module with function and path. Pvalue < 0.05 was considered significant, and the identified significant analyses were classified by gene counting.

The analysis of immune cell infiltration

Cibersort is an online analysis tool for immunocell infiltration for a variety of diseases, including osteoarthritis, lupus nephritis, atopic dermatitis, acne and rosacea dermatitis [25–28]. At present, few studies have used the Cibersort method to study the characteristics of immune cell infiltration in tendinopathy. We assessed the proportion of immune cells in patients and the normal population by Cibersort analysis based on the genome-wide level of tendinopathy [29]. This revealed the characteristics of tendinopathy immune cell infiltration by heat mapping using ggplot2.

Screening and identification of gene prediction model for early diagnosis

Lasso logistic regression is a machine learning method, which determines the variable by finding the λ value with the smallest classification error [30]. By processing the data of GSE26051 and GSE167226 processed, 75% of the samples in the data set are selected as the test set, 25% of the samples are used as the validation set, and the glmnet package of R is used as the binomial LASSO model of the training set. We plotted the operating characteristic curve of the recipient and determined its AUC. The diagnostic value of hub genes was evaluated using the pROC software package in R [31]. SVM-RFE is a machine learning method based on support vector machine, which finds the best variable by subtracting the feature vector generated by svm. We run the e1071 package to eliminate the recursive features of the difference genes obtained, and use the svmRFE function for data calculation. We set up folding by wrapping the entire feature selection and generalization error estimation process in the top loop of external cross-validation. Then, we use the Lapply function to perform feature ordering on all training sets. Finally, we obtain all folded top features through the top.features function, and use the loop function to estimate the generalization error of different numbers of top features using the 10x CV standard, so that the error rate reaches the lowest point, and finally the best gene signature is obtained. We used the above two algorithms to screen the hub genes of tendinopathy at the same time and obtained the same hub genes. Finally, the same hub genes were classified and screened by the Gaussian mixture model (GMMS). The Gaussian mixture model (GMMS) is a feasible screening method with better clustering performance [32]. Through repeated training of Gaussian mixture model algorithm, we screened the highest cluster in the optimal cluster in the figure, screened the optimal AUC value in the fifth category as 0.95, and identified the final candidate hub genes.

Results

Data preprocessing and the screening of differential genes

The flow chart of this study is shown in Fig 1. After gene annotation and standardization of the data, GSE26051 data set contains 46 samples (2088 genes), and GSE167226 data set contains 19 samples (19402 genes). We showed by performing a whole-gene expression profile analysis of the combined data, Fig 2A shows the DEG expression heat map. Subsequent differential analysis and volcano mapping of counts genes (20,188 genes) in patient tissue and normal tissue samples were noted. A total of 2995 DEGs were identified, including 2729 up-regulated genes and 266 down-regulated genes (Fig 2B).

Construction of weighted co-expression network and identification of key modules

We performed WGCNA analysis on 20,226 genes that data set GSE106292 has been annotated. We preprocessed the sample expression values, and screened 15034 genes by screening

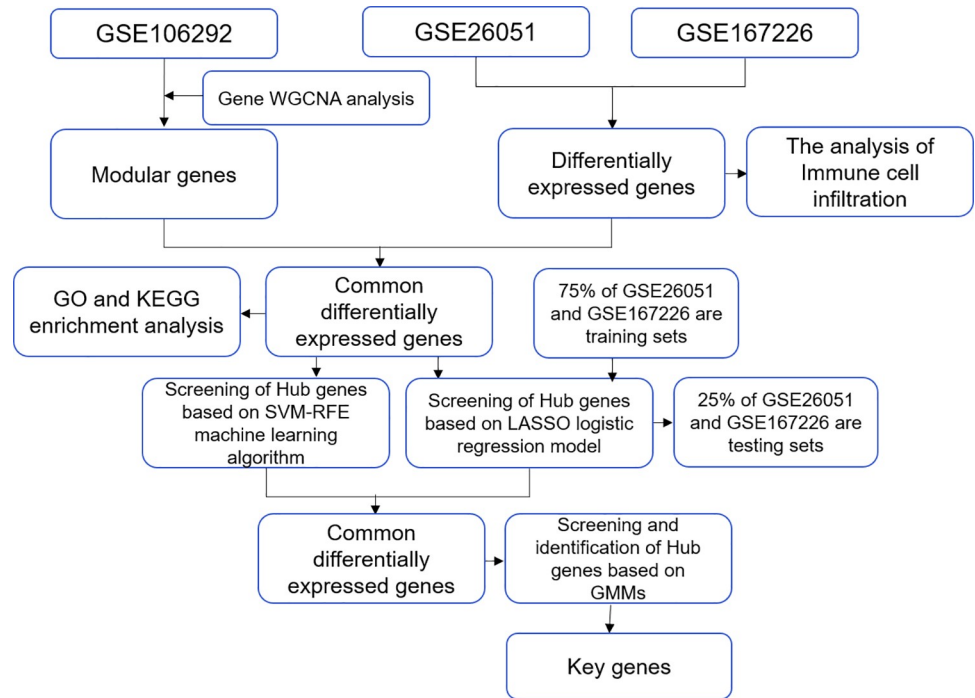


Fig 1. Flowchart of this study. The following datasets were used for the identification of potential diagnostic genes and mechanisms associated with the development of sepsis: GSE106292, GSE26051, GSE167226.

<https://doi.org/10.1371/journal.pone.0259475.g001>

standard deviation > 0.5. Based on the proximity matrix $\beta = 16$, we made the gene distribution conform to the scale-free network, sample tree and soft threshold estimation according to the connectivity degree. We set the vertical axis as the scale-free topology fitting index R^2 (the values in the SFT.R.Sq column in the statistical results) (Fig 3A) and the average connectivity of the network (Fig 3B). According to the negative correlation between k and $p(k)$ (correlation

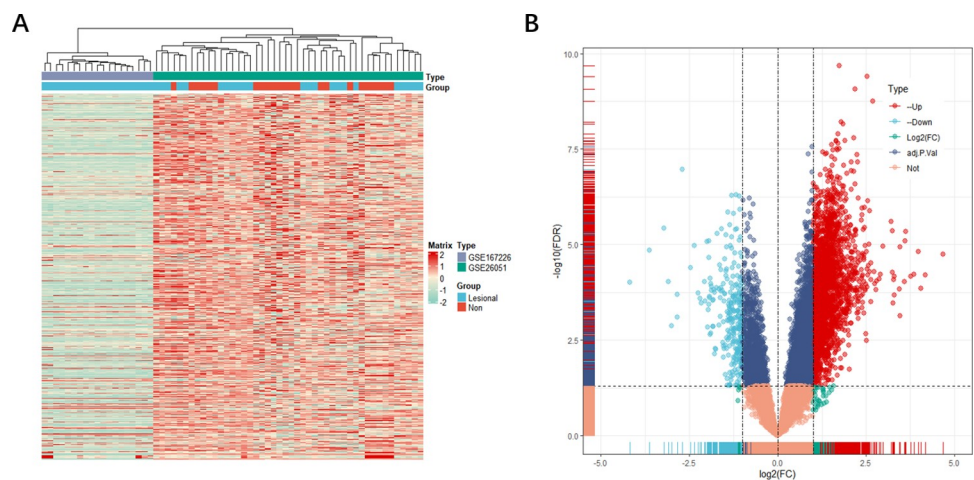


Fig 2. The gene differential expression analysis of GSE26051 and GSE167226 data set. (A) Whole gene expression heat map: Whole gene expression heat map of tendon tissue, with high expression in red and low expression in blue (B) The DEG Volcano map shows upregulated genes in red and down-regulated genes in blue.

<https://doi.org/10.1371/journal.pone.0259475.g002>

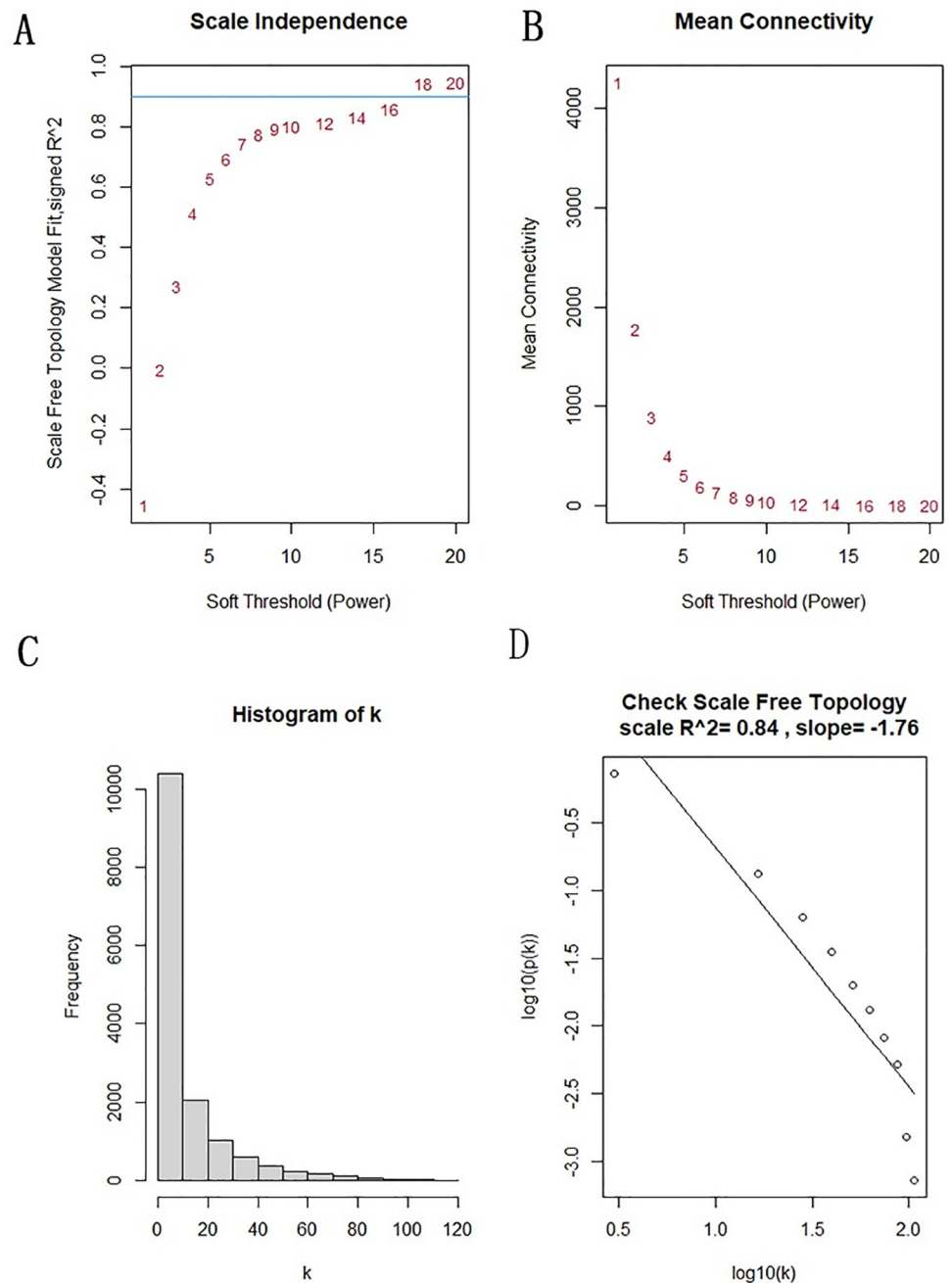


Fig 3. The screening criteria of WGCNA. (A) Soft Threshold (Power) represents the weight, and the vertical axis shows the scale-free topology fitting index R^2 (B) Soft Threshold (Power) represents the weight, and the vertical axis shows the average connectivity of the network (C) Distribution of node connectivity K (D) Correlation graph of K and $P(K)$.

<https://doi.org/10.1371/journal.pone.0259475.g003>

coefficient 0.84), it can be shown that the selected $\beta = 16$ meets the standard of establishing the gene scale-free network (Fig 3C and 3D).

We set the soft threshold power to 1–10, $\beta = 16$ (without scale $R^2 = 0.83$) to obtain the final template as shown in Fig 4A. In the results of the respruning of the cluster tree, we identified 20 modules in the heat map of correlation between the tree diagram of gene expression and module features (Fig 4B). We selected the two modules with the highest correlation as the

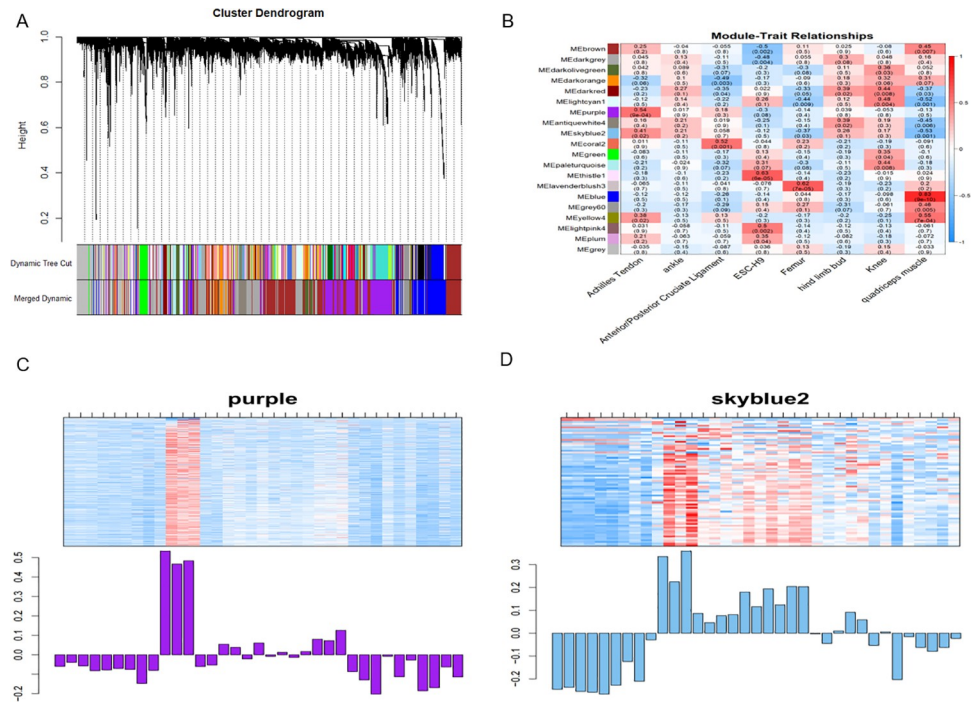


Fig 4. The WGCNA analysis of GSE106292 data set. (A) Tree of all gene expressions based on the Difference Measure (1-TOM) cluster (B) The heat maps of correlations between modules feature genes and samples, with each cell containing correlation coefficients and P values (C) The expression calorimetry and feature vector histogram of PURPRE module (D) The expression calorimetry and feature vector histogram of Skyblue2 module.

<https://doi.org/10.1371/journal.pone.0259475.g004>

research target module, among which ModuleTraitCor in Purple = 0.54, ModuleTraitPvalue = 9E-04 and ModuleTraitCor in Skyblue2 = 0.41, ModuleTraitPvalue = 0.02 meets the criteria (Fig 4C and 4D).

GO and KEGG enrichment analysis of differential genes

A total of 171 genes were obtained from the intersection of genes and differential genes in the module, including 150 up-regulated genes and 21 down-regulated genes. GO and KEGG enrichment analysis were used for enrichment analysis of the obtained DEG using the cluster-Profiler package of the R language. The results showed that the biological processes (BP) of the up-regulated genes were mainly related to the regulation of mitochondrion organization, the regulation of apoptotic signaling pathway, regulation of reactive oxygen species biosynthetic process, the regulation of tumor necrosis factor-mediated signaling pathway and other pathways. In terms of cell composition (CC), it is mainly related to peroxisome, microbody and ubiquitin ligase complex. In terms of molecular function (MF), it is mainly related to guanyl-nucleotide exchange factor activity, RNA helicase activity, GTPase regulator activity and phosphatidylinositol binding (Fig 5A and 5B). Among the down-regulated genes, biological processes (BP) are mainly related to protein acylation, positive regulation of protein-containing complex assembly, and I-kappaB kinase/NF-kappaB signaling. In terms of cell composition (CC), it is mainly related to catalytic step 2 spliceosome and a cluster of actin-based cell projections. In terms of molecular function (MF), it is mainly related to the binding of small GTPase binding and Ras GTPase binding (Fig 5A and 5B). The upregulated genes were mainly concentrated in mTOR, HIF-1, MAPK, NF-κB, NOD-like receptor and VEGF signaling pathways in the KEGG pathway rich concentration. The relatively down-regulated genes were mainly

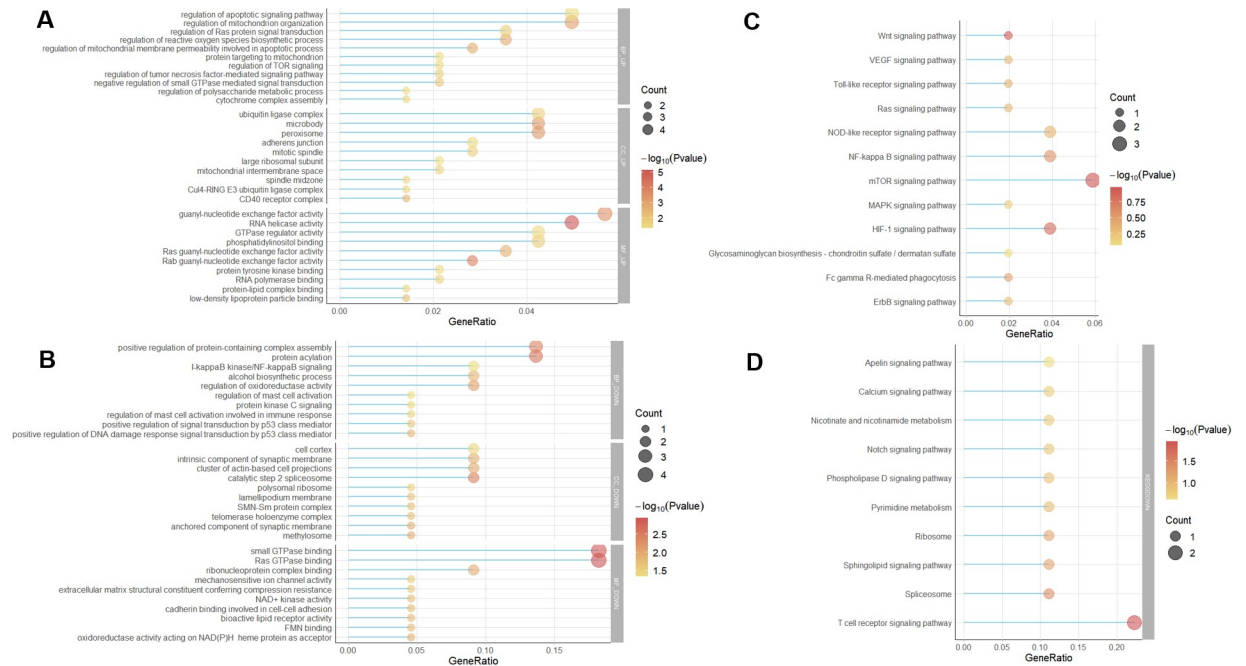


Fig 5. Baseball figure of differential gene enrichment analysis. The horizontal axis represents the proportion of differential genes in GO and KEGG enrichment analysis, and the vertical axis represents the enrichment category. (A) Up-regulated GO enrichment distribution of differentially expressed genes (B) Down-regulated GO enrichment distribution of differentially expressed genes (C) Up-regulated differential gene KEGG enrichment distribution (D) Down-regulated differential gene KEGG enrichment distribution.

<https://doi.org/10.1371/journal.pone.0259475.g005>

concentrated in T cell receptor signaling pathway, spliceosome, and sphingolipid signaling pathway (Fig 5C and 5D).

The analysis of immune cell infiltration

We obtained the CIBERSORT absolute score for the analysis of immune cell infiltration based on the CIBERSORT algorithm, which can reflect the absolute content of 22 immune cells in each sample. According to CIBERSORT absolute score, we calculated the proportion of each infiltrating cell when the total infiltration rate was 100%. We mapped the overall proportion pattern of 22 subgroups of immune cells in tendinopathy and healthy controls. The results showed that M2 macrophages, resting mast cells, neutrophils, activated NK cells and regulatory T cells (Tregs) were the highest infiltrating cells in all samples (Fig 6A). Heat maps of immune cell infiltration between the tendinopathy group and the control group showed significant enrichment of M1 macrophages, activated mast cells, γ - δ T cells, regulatory T cells (Tregs), neutrophils, and M0 macrophages (Fig 6B). We found that M2 macrophages, resting memory CD4 + T cells, memory B cells, resting mast cells, CD8 + T cells, activated NK cells and regulatory T cells (Tregs) were highly infiltrated in all samples (Fig 6C). The differential expression between the two groups showed that the infiltration of regulatory T cells (Tregs), activated NK cells and naive B cells was higher than that of the control group ($p < 0.05$), while the infiltration of plasma cells and memory B cells in the tendon disease group was lower ($p < 0.05$) (Fig 6D).

Screening and identification of hub genes in tendinopathy

First, we combine GSE26051 and GSE167226 data sets, and randomly divide all samples into a training set (75%) and verification set (25%). In order to prevent data snooping errors and

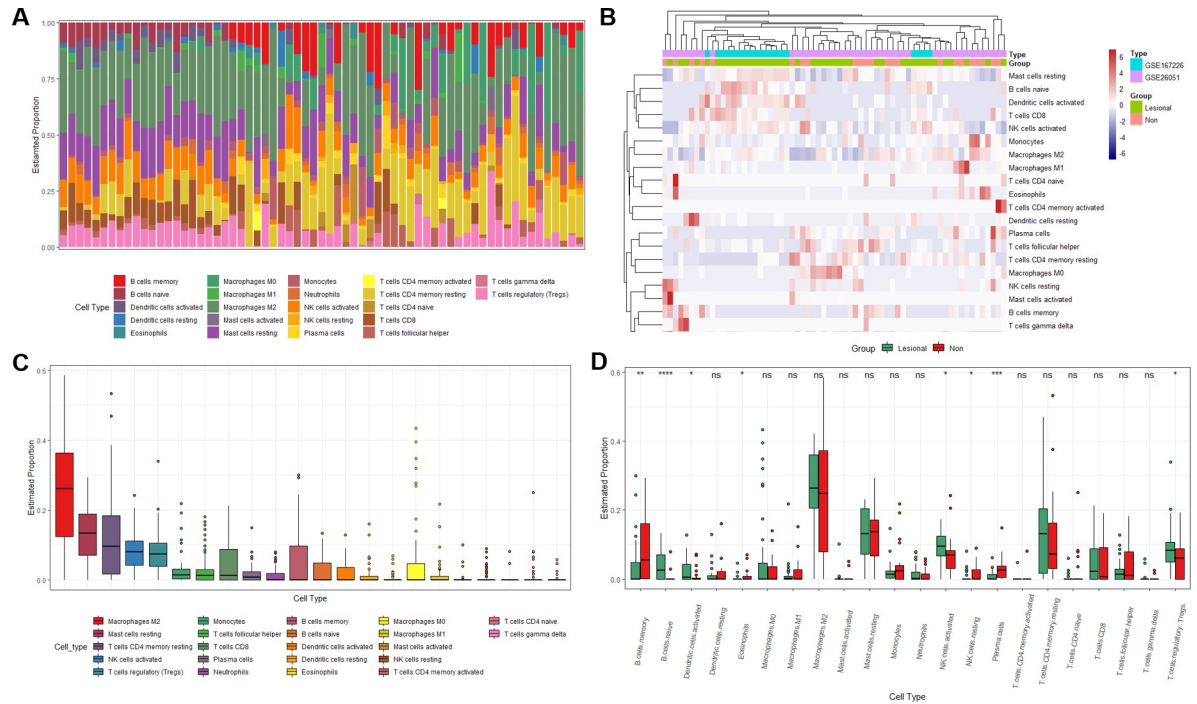


Fig 6. Infiltration patterns of immune cells in different groups. (A)Relative percentage of 22 immune cell subsets in tendon disease samples (B) Heat map of immune cell infiltration between tendinopathy group and control group, green represents tendinopathy group, red line represents control group (C) Infiltration degree of 22 immune cell subsets in tendon disease samples (D) Box Diagram of Immune Infiltration Difference between Tendon Disease Group and Control Group, Green as Tendon Disease Group, Red as Control Group.

<https://doi.org/10.1371/journal.pone.0259475.g006>

avoid sample characteristics in the test set, we use the sample function in the R language to randomly use 80% of the data set as the training set and 20% as the test set. We implement model prediction through the predict function, that is, the fitCV object is the construction model, and the evaluation is performed in the two data sets of train and test. By constructing the LASSO model in the training set and selecting $\lambda = 18$ to determine the hub genes that can predict early tendinopathy accurately (Fig 7A). Based on the optimal λ value of 18, we obtained the LASSO coefficient spectrum of differentially expressed genes (Fig 7B). Normally, the training set is used to train the model, and the test set is used to evaluate the performance of the model as a whole. In this study, the AUC of the training set is $0.981 > 0.7$, indicating that the built model has good real performance, and the AUC of the test set is $0.807 > 0.7$, indicating that the built model has good generalization performance and excellent verification performance (Fig 7C). Finally, we obtained 18 genes with nonzero coefficients, including *MACROD1*, *CES2*, *GFER*, *MRPL52*, *SKIV2L*, *B3GNT4*, *LYNX1*, *C19orf57*, *SAFB2*, *NOM1*, *C7orf43*, *FRMPD1*, *MLPH*, *MFSD10*, *PIEZO1*, *FAM222A*, *PRG4*, *POU2AF1*.

At the same time, we used multiple support vector machine recursive feature elimination (mSVM-RFE) algorithm to screen 171 different genes, and obtained 167 hub genes. Build the SVM model by selecting the first 171 variables and check that the model error rate is 167–0.139, and the accuracy rate is 167–0.861. The position of the red circle is the lowest point of error rate (Fig 8A). The position of the red circle is the point with the highest accuracy (Fig 8B). We obtained common hub genes from the above two algorithms, including *MACROD1*, *CES2*, *GFER*, *MRPL52*, *SKIV2L*, *B3GNT4*, *LYNX1*, *C19orf57*, *SAFB2*, *NOM1*, *C7orf43*, *FRMPD1*, *MLPH*, *MFSD10*, *PIEZO1*, *FAM222A*, *PRG4*, *POU2AF1*(Fig 8C).

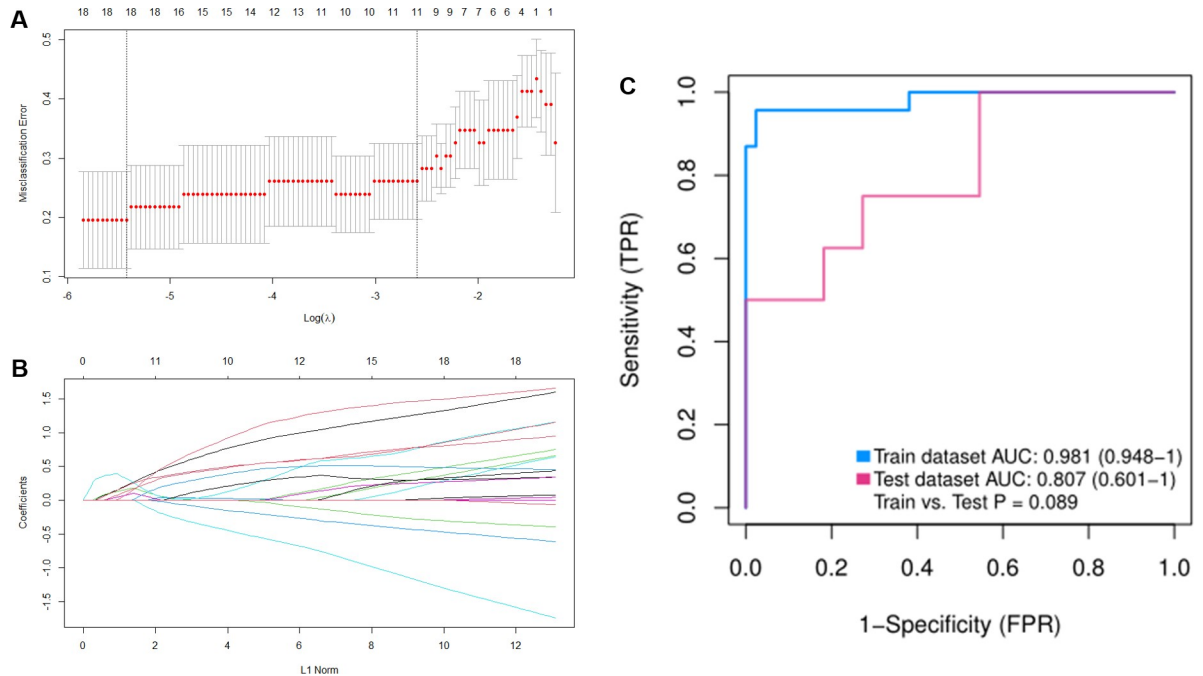


Fig 7. The potential key genes of tendinopathy were screened by LASSO regression model. In Fig 7A and 7B, the ordinate is the value of the coefficient, the lower abscissa is $\log(\lambda)$, and the upper abscissa is the number of non-zero coefficients in the model at this time. (A) Selection of the best parameter (number of non-zero coefficients in the model at this time) (B) LASSO coefficient spectrum of 18 differentially expressed genes selected by optimal (s timeti) (C) Comparison of ROC curves between training set and validation set for gene signature.

<https://doi.org/10.1371/journal.pone.0259475.g007>

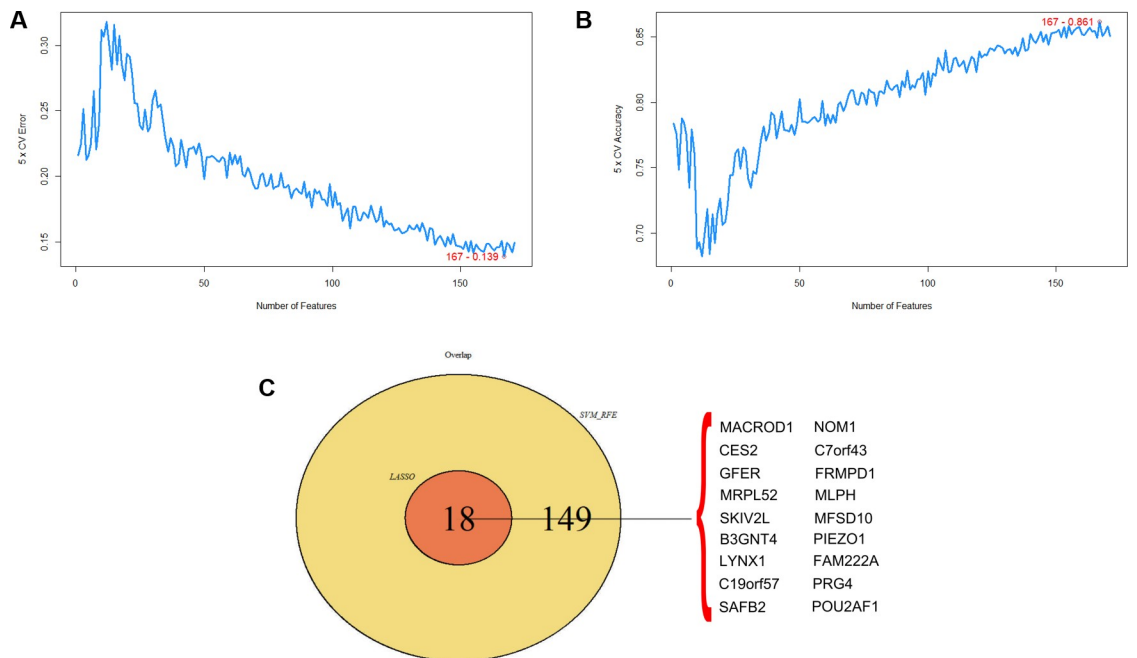


Fig 8. MSVM-RFE algorithm for screening key genes. (A) Shows the error rate of the SVM model (B) Shows the accuracy of the SVM model (C) The Venn diagram shows the same key genes obtained by the two algorithms.

<https://doi.org/10.1371/journal.pone.0259475.g008>

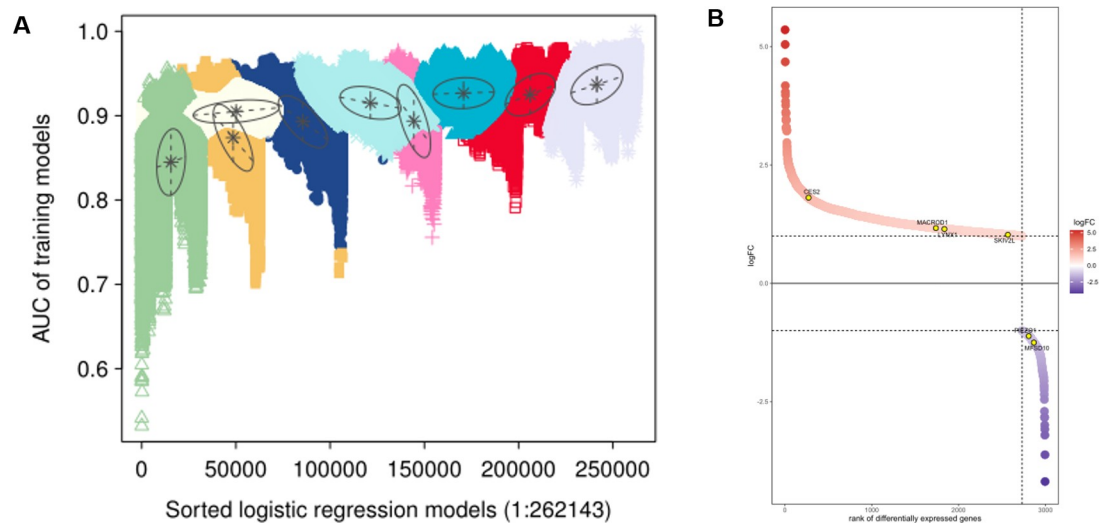


Fig 9. Displays the patterns of AUC and 262143 logistic regression model based on Gaussian finite mixture model. (A) The pattern of the logistic regression model is related to the AUC score and is determined by Gaussian mixture (B) The waterfall diagrams of 6 key genes in different genes.

<https://doi.org/10.1371/journal.pone.0259475.g009>

In order to more accurately predict genes for the early diagnosis of tendinopathy. Based on the Gaussian finite mixture model, we use the model-based hierarchical agglomerative clustering method for classification. We use a Gaussian Mixture Model (GMM) to classify miRNA clusters, and use Logistic regression analysis to establish a combined model for predicting recurrence. At the same time, we constructed a receiver operating characteristic (ROC) curve, and calculated AUC to evaluate the predictive value of the model, using a predictive miRNA signature model. We used the Gaussian mixture model algorithm to calculate and verify the best hub genes from 18 candidate hub genes by constructing 262143 AUC models. Finally, we screened 6 potential hub genes, including *MACROD1*, *CES2*, *SKIV2L*, *LYNX1*, *MFSD10* and *PIEZO1* (Fig 9A). The FoldChange and p-value populations were used to display differential gene waterfall plots, and the predictive mRNA signature model was described using Gaussian finite mixture model markers (Fig 9B). The overall situation of the 6 potential hub genes is listed in Table 1.

Discussion

Tendinopathy is not only a very common chronic disease, but also a disease that lacks real effective treatment [33]. So far, tendinopathy is still a major challenge in musculoskeletal diseases due to the widespread disease population, low cure rate, and huge medical expenditure. However, the mechanism of the occurrence and development of tendinopathy is not yet fully

Table 1. 6 potential hub gene display.

	logFC	AveExpr	t	P.Value	Type
MACROD1	1.167552501	5.712454732	5.41405282	8.56E-07	Up
CES2	1.809870576	7.075108564	5.355859588	1.07E-06	Up
SKIV2L	1.026460579	6.111613386	4.671443796	1.44E-05	Up
LYNX1	1.148012232	6.163685681	4.321936756	5.13E-05	Up
MFSD10	-1.250111926	5.106939598	-3.626082073	0.000548936	Down
PIEZO1	-1.112748319	6.687175452	-3.52261718	0.000765663	Down

<https://doi.org/10.1371/journal.pone.0259475.t001>

understood. At present, there are many hypotheses about the etiology of tendinopathy, including Biomechanical theory [34], inflammation theory [33], apoptosis theory [35], vascular or neurogenic theory [36], etc. Although these theoretical models closely link the basic science of tendinopathy with clinical applications, none of the theoretical theories can fully clarify the pathological mechanism of tendinopathy and the complex relationship between tendon pain and function. Although the current research on tendinopathy has involved many aspects, the hub genes of early tendinopathy are rarely studied. We believe that early diagnosis and treatment of tendinopathy is essential to prevent the further progression of tendinopathy to avoid the subsequent pathological cascade of tendinopathy. In this study, we finally screened out 6 key genes, including *MACROD1*, *CES2*, *SKIV2L*, *LYNX1*, *MFS10*, and *PIEZO1*. Among them, the gene with the smallest PValue is *MACROD1*. We speculate that this gene may be important in the occurrence and development of tendinopathy.

The *MACROD1* gene is still rare in current research and its specific function is unclear. A few studies have proved the functions of *MACROD1* in the nucleus and the cytoplasm of the body, including *MACROD1* binding and regulating transcription factors ER α and NF- κ B related proteins [37,38]. Recent studies have shown that endogenous *MACROD1* protein is highly enriched in mitochondria and is highly expressed in human and mouse skeletal muscle [39]. As we all know, mitochondria, as a kind of organelle, have abundant biological functions in cells [40]. Mitochondria are not only factories that produce ATP but also participate in many biological processes, such as steroid biosynthesis [41], metal ion homeostasis in the human body [42,43], immune cell activation and regulation [44], cell signaling [45], apoptosis [46] and inflammation [47]. At the same time, mitochondrial dysfunction can cause many diseases, such as atherosclerosis [48], Alzheimer's disease [49]. Mitochondria can produce reactive oxygen species (ROS) in the process of oxidative phosphorylation. When the accumulation of ROS exceeds the cellular antioxidant defense system, the accumulation of ROS levels will cause oxidative stress. Oxidative stress can cause ROS-mediated damage to molecular substances such as proteins, nucleic acids, and lipids [50]. And some studies believe that oxidative stress is closely related to vascular disease [51], neurodegeneration [52], and cancer [53]. Similarly, ROS can also regulate the cascade reaction in the MAPK signaling pathway by activating apoptosis signal regulator 1 (ASK1) to induce apoptosis [54,55]. Therefore, based on the high enrichment of *MacroD1* gene in mitochondria, high expression in skeletal muscle, and the biological function of mitochondria, we speculate that the occurrence of *MacroD1* gene in early tendinopathy is related to the comprehensive cascade reaction of hypoxic microenvironment, inflammatory response, apoptosis and so on.

The *MACROD1* gene may trigger the hypoxic microenvironment-mediated tendon inflammation through mitochondrial dysfunction. As we all know, hypoxic cell injury has been considered as the basic mechanism of tendinopathy [56]. Research suggests that hypoxia may be a potential cause of early tendinopathy. Under the action of mechanical stimulation and injury, hypoxia promotes the release of inflammatory cytokines in human tendon cells, the expression of key apoptosis mediators, the formation of blood vessels [57], and significantly affects the synthesis of collagen matrix [58]. Previous research suggests that chronic tendinopathy is caused by a degenerative process without inflammation. In a study of biopsy components of tendinopathy rupture tissue, no obvious inflammation was observed, and more than 85% of the biopsy specimens had almost no inflammatory cells. On the contrary, there was a marked increase in tissue degeneration, including thinning and disorientation of collagen fibers, myxoid degeneration, hyaluronic acid degeneration, chondroid metaplasia, tissue calcification, angiogenesis, and fatty infiltration [59,60]. Interestingly, some recent studies have reached the opposite conclusion that there is an early and important inflammatory response in the process of tendinopathy. Molloy et al found that the expression of the inflammatory cell receptor and

immunoglobulin was up-regulated in the rat supraspinatus tendon disease model by microarray analysis [61]. Matthews et al. showed that a small-area tear is more obvious in the inflammatory infiltration of macrophages and mast cells, reflecting more degenerative changes through a biopsy sample of tearing tendon tissue [62]. Interestingly, in our study, the KEGG enrichment pathway was significantly up-regulated by the NF- κ B signaling pathway and the MAPK signaling pathway. Studies have shown that these two pathways are related to inflammation, apoptosis and ossification. For example, ERK regulates apoptosis, while P38 mediates apoptosis and inflammation [63]. And there are also studies suggesting that the activation of NF- κ B can induce the activation of the MAPK pathway [64]. Another study showed that TNF- α can induce TDSC inflammation and apoptosis, and promote the development of tendinopathy by up-regulating the activation of MAPK and NF- κ B pathways [65]. Our analysis of the difference in immune cell infiltration of the whole gene of the samples showed that the infiltration of M1 macrophages, activated mast cells, activated NK cells, and regulatory T cells (Tregs) in tendinopathy samples increased, while the infiltration of plasma cells and memory B cells decreased. Through a cross-sectional and case-control study, Maja et al. found that in chronic tendinopathy tissues, compared with healthy tendons, most (52%-96%) biopsy specimens were observed in macrophages, T lymphocytes, and hypertrophy Cells and natural killer cells [66]. This indicates that M1 macrophages, activated mast cells, and activated NK cells play an important role in the progression and treatment of tendinopathy. It can be considered that the *MACROD1* gene may induce inflammatory cell infiltration by mediating the hypoxic microenvironment.

Disregulation of apoptosis is thought to be one of the causes of tendinopathy [35]. Apoptosis is a kind of programmed cell death, which plays a key role in tissue homeostasis. Apoptosis causes many diseases, such as autoimmune diseases and skeletal muscle degeneration [67]. Current research believes that cell apoptosis are crucial in the occurrence and development of tendinopathy [35]. The reasons include: mechanical overuse of tendons [68], hypoxic microenvironment [69], and oxidative stress [70]. In our study, the differential gene enrichment pathway was significantly up-regulated on the HIF-1 signaling pathway. However, hypoxia-inducible factor 1 (HIF-1), as a transcriptional activator sensitive to oxygen [71], is a key regulator in the process of cell apoptosis [72]. In hypoxia, HIF-1 α can initiate cell apoptosis by inducing high concentrations of pro-apoptotic proteins. Vascular endothelial growth factor (VEGF) is a glycosylated protein of about 45 kDa, composed of two subunits connected by disulfide bonds, which can mediate angiogenesis. Previous studies have suggested that VEGF is at a high level of expression in degenerative tendinopathy, while its expression is almost completely down-regulated in healthy Achilles tendons [73]. There are many factors that cause high expression of VEGF in tendon cells, including hypoxia, inflammatory factors and mechanical stress load [74]. As for the mechanism of action of VEGF in tendinopathy, some studies have shown that VEGF activates the binding of VEGF and its receptor VEGFR-2 to promote angiogenesis in tendon tissue by up-regulating the expression of matrix metalloproteinases (MMPs) and down-regulating the expression of metalloproteinase-3 (TIMP-3) in tendon cells [74–77]. Dakin et al. studied the tendon biopsy in symptomatic patients with tendinopathy or rupture and found that the superposition of inflammatory infiltration and neovascularization promotes tendon rupture [78]. Interestingly, in our study, based on KEGG enrichment, the VEGF signaling pathway is significantly up-regulated, which may lead us to speculate that VEGF plays an important role in the early pathogenesis of tendinopathy under hypoxic factors.

In the development of tendinopathy, the wrong differentiation of tendon cells may promote the occurrence of tendinopathy [79]. Hypoxia and inflammation are related to the occurrence of heterotopic endochondral ossification (HEO), but the specific molecular mechanism is

unclear. It has been proven that hypoxic environment can stimulate the differentiation of progenitor cells into cartilage during the development of the skeletal system [80]. Wang et al. believed that cell hypoxia promoted heterotopic ossification by amplifying BMP signal transduction [81]. Shailesh et al. in the rat Achilles tendon, muscle ossification model and FOP mouse model found that HIF-1 α was significantly up-regulated in the chondrogenic differentiation stage [82]. Tendon-derived stem cells (TDSC) have the potential to differentiate into tendon cells, osteoblasts, chondrocytes and fibroblasts [83]. TDSC plays an important role in the healing of tendon injuries. However, if the tendon is not properly healed, it will cause tendon ossification and promote the formation of tendinopathy [84–86]. Studies have shown that TNF- α can induce apoptosis of TDSC [87], and chronic tendinopathy is closely related to the up-regulation of TNF- α [88]. This is closely related to our study and screening of different genes in the regulation of tumor necrosis factor-mediated signaling pathways on GO enrichment. Some studies believe that Wnt pathway mediators are expressed in chondrocyte-like cells and ossification deposits, and are related to endochondral ossification [89]. Differentiation of tendon stem cells into non-tendon cells, such as osteoblasts, may reduce the total amount of tendon stem cells used for tendon repair and lead to failure of tendon healing. Studies have shown that activation of the Wnt pathway can promote calcification of related tissues, such as cardiovascular calcification [90,91]. In our study, the Wnt pathway was significantly up-regulated. We believe that the Wnt pathway is activated and at a high level of expression in tendinopathy injuries. And Wnt induces the osteogenic differentiation of tendon stem cells, and promotes the differentiation of tendon cells in the wrong direction, which ultimately leads to tendinopathy [92]. Studies have shown that in the rat Achilles tendon injury model, the key regulators of the Wnt pathway and the Notch pathway are activated at the wound [93]. Therefore, it can be considered that the *MACROD1* gene may promote the misdifferentiation of tendon cells by mediating the hypoxic microenvironment of the cells and ultimately lead to tendinopathy.

Based on the above research, we speculate that the up-regulation of *MACROD1* gene may cause early tendinopathy hypoxia microenvironment and oxidative stress response through tendinocyte mitochondrial dysfunction, and then activate multiple signal pathways under the combined action of inflammatory cytokines and angiogenic factors Lead to apoptosis of tendon cells. These cascades ultimately lead to the development of chronic tendinopathy. On the other hand, we used the online tool CIBERSORT algorithm to analyze the immune cell infiltration to get the difference in the content of immune cell infiltration in 22. This may have important implications for tendinopathy in the study of immune cells. In short, we speculate that the *MACROD1* gene may be a potential hub gene for early tendinopathy. It is necessary for us to study clearly the function of *MACROD1* gene in mitochondria and the possible specific molecular mechanism of the early occurrence of tendinopathy. By exploring the mechanism of *MACROD1* gene in mitochondria and the characteristics of immune cell infiltration, we can find new therapeutic targets in molecular pathways, which may be a promising treatment method for tendinopathy.

In recent decades, computational models have become an important tool for the identification of novel MicroRNAs, LncRNAs and CircRNAs in association with diseases. Meanwhile, a large number of experimental methods and computational models have been designed and implemented to identify novel MicroRNA, LncRNA, and CircRNA associations with complex diseases, which contribute to the understanding of human complex disease mechanisms, biomarker detection, and disease diagnosis, treatment, prognosis, and prevention at the molecular level [94–97]. And, there are also studies on LncRNA-MicroRNA interaction prediction by network distance analysis model, which has an important role in the screening of therapeutic targets and diagnostic biomarkers for a variety of human diseases [97–99]. On the other hand,

logistic matrix factorization with neighborhood regularized (LMFNRLMI) is a new matrix factorization model for predicting the interaction of lncRNA-miRNA. Research shows that through comparison with several other network algorithms and various similarity tests, the model is superior and has higher performance in predicting the association of lncRNA-miRNA [100]. In the future, we can use LMFNRLMI to predict potential lncRNA-miRNA association studies in tendinopathy. Unfortunately, association studies at the transcriptome level with tendinopathies using bioinformatics approaches have been inadequate in recent years. Therefore, more research is needed in the future on top of computational modeling studies of tendinopathy gene biomarker identification and association studies at the transcriptome level.

However, our research still has certain limitations. First of all, the data set we are studying contains different populations of tendon tear patients and the control group, which may affect the results of the study. In addition, there is a slight difference between the immune cell infiltration condition obtained by the whole genetic immunoassay and the immune cell infiltration condition obtained from the experimental study, which may be caused by the difference in different stages of the disease. Finally, based on our research data comes from public databases, it is necessary to conduct molecular cell and animal experiments to verify the results of this research.

Conclusion

In conclusion, based on the comprehensive bioinformatics analysis method, we identified the potential early hub genes, key regulatory pathways and immune infiltration characteristics of tendon disease. This will help to provide new insights into the future drug and molecular mechanism of tendon disease.

Author Contributions

Conceptualization: Ya xi Zhu.

Data curation: Ya xi Zhu, Jia qiang Huang.

Formal analysis: Yu yang Ming.

Funding acquisition: Hong Xia.

Resources: Jia qiang Huang.

Software: Zhao Zhuang.

Supervision: Hong Xia.

Writing – original draft: Ya xi Zhu.

Writing – review & editing: Hong Xia.

References

1. Millar NL, Silbernagel KG, Thorborg K, Kirwan PD, Galatz LM, Abrams GD, et al. Tendinopathy. *Nature reviews Disease primers*. 2021; 7(1):1. Epub 2021/01/09. <https://doi.org/10.1038/s41572-020-00234-1> PMID: 33414454
2. Hopkins C, Fu SC, Chua E, Hu X, Rolf C, Mattila VM, et al. Critical review on the socio-economic impact of tendinopathy. *Asia-Pacific journal of sports medicine, arthroscopy, rehabilitation and technology*. 2016; 4:9–20. Epub 2016/04/22. <https://doi.org/10.1016/j.asmart.2016.01.002> PMID: 29264258

3. Riley G. Chronic tendon pathology: molecular basis and therapeutic implications. *Expert reviews in molecular medicine*. 2005; 7(5):1–25. Epub 2005/03/31. <https://doi.org/10.1017/S1462399405008963> PMID: 15796783
4. Lui PP, Maffulli N, Rolf C, Smith RK. What are the validated animal models for tendinopathy? *Scandinavian journal of medicine & science in sports*. 2011; 21(1):3–17. Epub 2010/08/03. <https://doi.org/10.1111/j.1600-0838.2010.01164.x> PMID: 20673247
5. Sharma P, Maffulli N. Basic biology of tendon injury and healing. *The surgeon: journal of the Royal Colleges of Surgeons of Edinburgh and Ireland*. 2005; 3(5):309–16. Epub 2005/10/26. [https://doi.org/10.1016/s1479-666x\(05\)80109-x](https://doi.org/10.1016/s1479-666x(05)80109-x) PMID: 16245649
6. Alfredson H, Pietilä T, Jonsson P, Lorentzon R. Heavy-load eccentric calf muscle training for the treatment of chronic Achilles tendinosis. *The American journal of sports medicine*. 1998; 26(3):360–6. Epub 1998/06/09. <https://doi.org/10.1177/03635465980260030301> PMID: 9617396
7. Camargo PR, Albuquerque-Sendín F, Salvini TF. Eccentric training as a new approach for rotator cuff tendinopathy: Review and perspectives. *World journal of orthopedics*. 2014; 5(5):634–44. Epub 2014/11/19. <https://doi.org/10.5312/wjo.v5.i5.634> PMID: 25405092
8. Irby A, Gutierrez J, Chamberlin C, Thomas SJ, Rosen AB. Clinical management of tendinopathy: A systematic review of systematic reviews evaluating the effectiveness of tendinopathy treatments. *Scandinavian journal of medicine & science in sports*. 2020; 30(10):1810–26. Epub 2020/06/03. <https://doi.org/10.1111/sms.13734> PMID: 32484976
9. Huang X, Liu S, Wu L, Jiang M, Hou Y. High Throughput Single Cell RNA Sequencing, Bioinformatics Analysis and Applications. *Advances in experimental medicine and biology*. 2018; 1068:33–43. Epub 2018/06/27. https://doi.org/10.1007/978-981-13-0502-3_4 PMID: 29943294
10. Wang T, Zheng X, Li R, Liu X, Wu J, Zhong X, et al. Integrated bioinformatic analysis reveals YWHAB as a novel diagnostic biomarker for idiopathic pulmonary arterial hypertension. *Journal of cellular physiology*. 2019; 234(5):6449–62. Epub 2018/10/15. <https://doi.org/10.1002/jcp.27381> PMID: 30317584
11. Ge Z, Tang H, Lyu J, Zhou B, Yang M, Tang K, et al. Conjoint analysis of lncRNA and mRNA expression in rotator cuff tendinopathy. *Annals of translational medicine*. 2020; 8(6):335. Epub 2020/05/02. <https://doi.org/10.21037/atm.2020.02.149> PMID: 32355779
12. Zhang Q, Ge H, Jiang Y, Cheng B, Zhou D, Xu N. Microarray profiling analysis of long non-coding RNAs expression in tendinopathy: identification for potential biomarkers and mechanisms. *International journal of experimental pathology*. 2015; 96(6):387–94. Epub 2016/01/15. <https://doi.org/10.1111/iep.12158> PMID: 26764085
13. Li Z, Sillanpää MJ. Overview of LASSO-related penalized regression methods for quantitative trait mapping and genomic selection. *TAG Theoretical and applied genetics Theoretische und angewandte Genetik*. 2012; 125(3):419–35. Epub 2012/05/25. <https://doi.org/10.1007/s00122-012-1892-9> PMID: 22622521
14. Sanz H, Valim C, Vegas E, Oller JM, Reverter F. SVM-RFE: selection and visualization of the most relevant features through non-linear kernels. *BMC bioinformatics*. 2018; 19(1):432. Epub 2018/11/21. <https://doi.org/10.1186/s12859-018-2451-4> PMID: 30453885
15. Zhao Y, Shrivastava AK, Tsui KL. Regularized Gaussian Mixture Model for High-Dimensional Clustering. *IEEE transactions on cybernetics*. 2019; 49(10):3677–88. Epub 2018/07/12. <https://doi.org/10.1109/TCYB.2018.2846404> PMID: 29994696
16. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: archive for functional genomics data sets—update. *Nucleic acids research*. 2013; 41(Database issue):D991–5. Epub 2012/11/30. <https://doi.org/10.1093/nar/gks1193> PMID: 23193258
17. Hicks MR, Hiserodt J, Paras K, Fujiwara W, Eskin A, Jan M, et al. ERBB3 and NGFR mark a distinct skeletal muscle progenitor cell in human development and hPSCs. *Nature cell biology*. 2018; 20(1):46–57. Epub 2017/12/20. <https://doi.org/10.1038/s41556-017-0010-2> PMID: 29255171
18. Ferguson GB, Van Handel B, Bay M, Fiziev P, Org T, Lee S, et al. Mapping molecular landmarks of human skeletal ontogeny and pluripotent stem cell-derived articular chondrocytes. *Nature communications*. 2018; 9(1):3634. Epub 2018/09/09. <https://doi.org/10.1038/s41467-018-05573-y> PMID: 30194383
19. Jelinsky SA, Rodeo SA, Li J, Gulotta LV, Archambault JM, Seeherman HJ. Regulation of gene expression in human tendinopathy. *BMC musculoskeletal disorders*. 2011; 12:86. Epub 2011/05/05. <https://doi.org/10.1186/1471-2474-12-86> PMID: 21539748
20. Gautier L, Cope L, Bolstad BM, Irizarry RA. affy—analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics (Oxford, England)*. 2004; 20(3):307–15. Epub 2004/02/13. <https://doi.org/10.1093/bioinformatics/btg405> PMID: 14960456
21. Parker HS, Leek JT, Favorov AV, Conside M, Xia X, Chavan S, et al. Preserving biological heterogeneity with a permuted surrogate variable analysis for genomics batch correction. *Bioinformatics*

- (Oxford, England). 2014; 30(19):2757–63. Epub 2014/06/08. <https://doi.org/10.1093/bioinformatics/btu375> PMID: 24907368
22. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic acids research*. 2015; 43(7):e47. Epub 2015/01/22. <https://doi.org/10.1093/nar/gkv007> PMID: 25605792
 23. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC bioinformatics*. 2008; 9:559. Epub 2008/12/31. <https://doi.org/10.1186/1471-2105-9-559> PMID: 19114008
 24. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics: a journal of integrative biology*. 2012; 16(5):284–7. Epub 2012/03/30. <https://doi.org/10.1089/omi.2011.0118> PMID: 22455463
 25. Deng YJ, Ren EH, Yuan WH, Zhang GZ, Wu ZL, Xie QQ. GRB10 and E2F3 as Diagnostic Markers of Osteoarthritis and Their Correlation with Immune Infiltration. *Diagnostics (Basel, Switzerland)*. 2020; 10(3). Epub 2020/04/03. <https://doi.org/10.3390/ani10030530> PMID: 32235700
 26. Cao Y, Tang W, Tang W. Immune cell infiltration characteristics and related core genes in lupus nephritis: results from bioinformatic analysis. *BMC immunology*. 2019; 20(1):37. Epub 2019/10/23. <https://doi.org/10.1186/s12865-019-0316-x> PMID: 31638917
 27. Félix Garza ZC, Lenz M, Liebmann J, Ertaylan G, Born M, Arts ICW, et al. Characterization of disease-specific cellular abundance profiles of chronic inflammatory skin conditions from deconvolution of biopsy samples. *BMC medical genomics*. 2019; 12(1):121. Epub 2019/08/20. <https://doi.org/10.1186/s12920-019-0567-7> PMID: 31420038
 28. Yang L, Shou YH, Yang YS, Xu JH. Elucidating the immune infiltration in acne and its comparison with rosacea by integrated bioinformatics analysis. *PloS one*. 2021; 16(3):e0248650. Epub 2021/03/25. <https://doi.org/10.1371/journal.pone.0248650> PMID: 33760854
 29. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nature methods*. 2015; 12(5):453–7. Epub 2015/03/31. <https://doi.org/10.1038/nmeth.3337> PMID: 25822800
 30. Antonacci Y, Toppi J, Mattia D, Pietrabissa A, Astolfi L. Single-trial Connectivity Estimation through the Least Absolute Shrinkage and Selection Operator. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Annual International Conference*. 2019; 2019:6422–5. Epub 2020/01/18. <https://doi.org/10.1109/EMBC.2019.8857909> PMID: 31947312
 31. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC bioinformatics*. 2011; 12:77. Epub 2011/03/19. <https://doi.org/10.1186/1471-2105-12-77> PMID: 21414208
 32. Ficklin SP, Dunwoodie LJ, Poehlman WL, Watson C, Roche KE, Feltus FA. Discovering Condition-Specific Gene Co-Expression Patterns Using Gaussian Mixture Models: A Cancer Case Study. *Scientific reports*. 2017; 7(1):8617. Epub 2017/08/19. <https://doi.org/10.1038/s41598-017-09094-4> PMID: 28819158
 33. Rees JD, Stride M, Scott A. Tendons—time to revisit inflammation. *British journal of sports medicine*. 2014; 48(21):1553–7. Epub 2013/03/12. <https://doi.org/10.1136/bjsports-2012-091957> PMID: 23476034
 34. Almekinders LC, Weinhold PS, Maffulli N. Compression etiology in tendinopathy. *Clinics in sports medicine*. 2003; 22(4):703–10. Epub 2003/10/17. [https://doi.org/10.1016/s0278-5919\(03\)0067-x](https://doi.org/10.1016/s0278-5919(03)0067-x) PMID: 14560542
 35. Xu Y, Murrell GA. The basic science of tendinopathy. *Clinical orthopaedics and related research*. 2008; 466(7):1528–38. Epub 2008/05/15. <https://doi.org/10.1007/s11999-008-0286-4> PMID: 18478310
 36. Riley G. The pathogenesis of tendinopathy. A molecular perspective. *Rheumatology (Oxford, England)*. 2004; 43(2):131–42. Epub 2003/07/18. <https://doi.org/10.1093/rheumatology/keg448> PMID: 12867575
 37. Han WD, Zhao YL, Meng YG, Zang L, Wu ZQ, Li Q, et al. Estrogenically regulated LRP16 interacts with estrogen receptor alpha and enhances the receptor's transcriptional activity. *Endocrine-related cancer*. 2007; 14(3):741–53. Epub 2007/10/05. <https://doi.org/10.1677/ERC-06-0082> PMID: 17914104
 38. Wu Z, Li Y, Li X, Ti D, Zhao Y, Si Y, et al. LRP16 integrates into NF- κ B transcriptional complex and is required for its functional activation. *PloS one*. 2011; 6(3):e18157. Epub 2011/04/13. <https://doi.org/10.1371/journal.pone.0018157> PMID: 21483817

39. Agnew T, Munnur D, Crawford K, Palazzo L, Mikoč A, Ahel I. MacroD1 Is a Promiscuous ADP-Ribosyl Hydrolase Localized to Mitochondria. *Frontiers in microbiology*. 2018; 9:20. Epub 2018/02/08. <https://doi.org/10.3389/fmicb.2018.00020> PMID: 29410655
40. McBride HM, Neuspiel M, Wasiak S. Mitochondria: more than just a powerhouse. *Current biology: CB*. 2006; 16(14):R551–60. Epub 2006/07/25. <https://doi.org/10.1016/j.cub.2006.06.054> PMID: 16860735
41. Chien Y, Rosal K, Chung BC. Function of CYP11A1 in the mitochondria. *Molecular and cellular endocrinology*. 2017; 441:55–61. Epub 2016/11/07. <https://doi.org/10.1016/j.mce.2016.10.030> PMID: 27815210
42. Bravo-Sagua R, Parra V, López-Crisosto C, Díaz P, Quest AF, Lavandero S. Calcium Transport and Signaling in Mitochondria. *Comprehensive Physiology*. 2017; 7(2):623–34. Epub 2017/03/24. <https://doi.org/10.1002/cphy.c160013> PMID: 28333383
43. Paul BT, Manz DH, Torti FM, Torti SV. Mitochondria and Iron: current questions. *Expert review of hematology*. 2017; 10(1):65–79. Epub 2016/12/03. <https://doi.org/10.1080/17474086.2016.1268047> PMID: 27911100
44. Liu PS, Ho PC. Mitochondria: A master regulator in macrophage and T cell immunity. *Mitochondrion*. 2018; 41:45–50. Epub 2017/11/18. <https://doi.org/10.1016/j.mito.2017.11.002> PMID: 29146487
45. Blajszczak C, Bonini MG. Mitochondria targeting by environmental stressors: Implications for redox cellular signaling. *Toxicology*. 2017; 391:84–9. Epub 2017/07/29. <https://doi.org/10.1016/j.tox.2017.07.013> PMID: 28750850
46. Jeong SY, Seol DW. The role of mitochondria in apoptosis. *BMB reports*. 2008; 41(1):11–22. Epub 2008/02/29. <https://doi.org/10.5483/bmbrep.2008.41.1.011> PMID: 18304445
47. Kolmychkova KI, Zhelankin AV, Karagodin VP, Orekhov AN. Mitochondria and inflammation. *Patologicheskaja fiziologija i eksperimental'naia terapija*. 2016; 60(4):114–21. Epub 2016/10/01. PMID: 29244932
48. Suárez-Rivero JM, Pastor-Maldonado CJ, Povea-Cabello S, Álvarez-Córdoba M, Villalón-García I, Talaverón-Rey M, et al. From Mitochondria to Atherosclerosis: The Inflammation Path. *Biomedicines*. 2021; 9(3). Epub 2021/04/04. <https://doi.org/10.3390/biomedicines9030258> PMID: 33807807
49. Kerr JS, Adriaanse BA, Greig NH, Mattson MP, Cader MZ, Bohr VA, et al. Mitophagy and Alzheimer's Disease: Cellular and Molecular Mechanisms. *Trends in neurosciences*. 2017; 40(3):151–66. Epub 2017/02/14. <https://doi.org/10.1016/j.tins.2017.01.002> PMID: 28190529
50. Ray PD, Huang BW, Tsuji Y. Reactive oxygen species (ROS) homeostasis and redox regulation in cellular signaling. *Cellular signalling*. 2012; 24(5):981–90. Epub 2012/01/31. <https://doi.org/10.1016/j.cellsig.2012.01.008> PMID: 22286106
51. Paravicini TM, Touyz RM. Redox signaling in hypertension. *Cardiovascular research*. 2006; 71(2):247–58. Epub 2006/06/13. <https://doi.org/10.1016/j.cardiores.2006.05.001> PMID: 16765337
52. Shukla V, Mishra SK, Pant HC. Oxidative stress in neurodegeneration. *Advances in pharmacological sciences*. 2011; 2011:572634. Epub 2011/09/24. <https://doi.org/10.1155/2011/572634> PMID: 21941533
53. Trachootham D, Alexandre J, Huang P. Targeting cancer cells by ROS-mediated mechanisms: a radical therapeutic approach? *Nature reviews Drug discovery*. 2009; 8(7):579–91. Epub 2009/05/30. <https://doi.org/10.1038/nrd2803> PMID: 19478820
54. Tobiume K, Matsuzawa A, Takahashi T, Nishitoh H, Morita K, Takeda K, et al. ASK1 is required for sustained activations of JNK/p38 MAP kinases and apoptosis. *EMBO reports*. 2001; 2(3):222–8. Epub 2001/03/27. <https://doi.org/10.1093/embo-reports/kve046> PMID: 11266364
55. Ichijo H, Nishida E, Irie K, ten Dijke P, Saitoh M, Moriguchi T, et al. Induction of apoptosis by ASK1, a mammalian MAPKKK that activates SAPK/JNK and p38 signaling pathways. *Science (New York, NY)*. 1997; 275(5296):90–4. Epub 1997/01/03. <https://doi.org/10.1126/science.275.5296.90> PMID: 8974401
56. Kannus P. Etiology and pathophysiology of chronic tendon disorders in sports. *Scandinavian journal of medicine & science in sports*. 1997; 7(2):78–85. Epub 1997/04/01. <https://doi.org/10.1111/j.1600-0838.1997.tb00123.x> PMID: 9211608
57. Järvinen TA. Neovascularisation in tendinopathy: from eradication to stabilisation? *British journal of sports medicine*. 2020; 54(1):1–2. Epub 2019/10/09. <https://doi.org/10.1136/bjsports-2019-100608> PMID: 31594793
58. Millar NL, Reilly JH, Kerr SC, Campbell AL, Little KJ, Leach WJ, et al. Hypoxia: a critical regulator of early human tendinopathy. *Annals of the rheumatic diseases*. 2012; 71(2):302–10. Epub 2011/10/06. <https://doi.org/10.1136/ard.2011.154229> PMID: 21972243

59. Maffulli N, Barrass V, Ewen SW. Light microscopic histology of achilles tendon ruptures. A comparison with unruptured tendons. *The American journal of sports medicine*. 2000; 28(6):857–63. Epub 2000/12/02. <https://doi.org/10.1177/03635465000280061401> PMID: 11101109
60. Hashimoto T, Nobuhara K, Hamada T. Pathologic evidence of degeneration as a primary cause of rotator cuff tear. *Clinical orthopaedics and related research*. 2003;(415):111–20. Epub 2003/11/13. <https://doi.org/10.1097/01.blo.0000092974.12414.22> PMID: 14612637
61. Molloy TJ, Kemp MW, Wang Y, Murrell GA. Microarray analysis of the tendinopathic rat supraspinatus tendon: glutamate signaling and its potential role in tendon degeneration. *Journal of applied physiology* (Bethesda, Md: 1985). 2006; 101(6):1702–9. Epub 2006/08/05. <https://doi.org/10.1152/jappphysiol.00386.2006> PMID: 16888051
62. Matthews TJ, Hand GC, Rees JL, Athanasou NA, Carr AJ. Pathology of the torn rotator cuff tendon. Reduction in potential for repair as tear size increases. *The Journal of bone and joint surgery British volume*. 2006; 88(4):489–95. Epub 2006/03/29. <https://doi.org/10.1302/0301-620X.88B4.16845> PMID: 16567784
63. Zhou J, Du T, Li B, Rong Y, Verkhatsky A, Peng L. Crosstalk Between MAPK/ERK and PI3K/AKT Signal Pathways During Brain Ischemia/Reperfusion. *ASN neuro*. 2015; 7(5). Epub 2015/10/08. <https://doi.org/10.1177/1759091415602463> PMID: 26442853
64. Wang L, Lee W, Cui YR, Ahn G, Jeon YJ. Protective effect of green tea catechin against urban fine dust particle-induced skin aging by regulation of NF- κ B, AP-1, and MAPKs signaling pathways. *Environmental pollution (Barking, Essex: 1987)*. 2019; 252(Pt B):1318–24. Epub 2019/06/30. <https://doi.org/10.1016/j.envpol.2019.06.029> PMID: 31252129
65. Moqbel SAA, Xu K, Chen Z, Xu L, He Y, Wu Z, et al. Tectorigenin Alleviates Inflammation, Apoptosis, and Ossification in Rat Tendon-Derived Stem Cells via Modulating NF-Kappa B and MAPK Pathways. *Frontiers in cell and developmental biology*. 2020; 8:568894. Epub 2020/11/17. <https://doi.org/10.3389/fcell.2020.568894> PMID: 33195199
66. Kraggsnaes MS, Fredberg U, Stribolt K, Kjaer SG, Bendix K, Ellingsen T. Stereological quantification of immune-competent cells in baseline biopsy specimens from achilles tendons: results from patients with chronic tendinopathy followed for more than 4 years. *The American journal of sports medicine*. 2014; 42(10):2435–45. Epub 2014/08/02. <https://doi.org/10.1177/0363546514542329> PMID: 25081311
67. Elmore S. Apoptosis: a review of programmed cell death. *Toxicologic pathology*. 2007; 35(4):495–516. Epub 2007/06/15. <https://doi.org/10.1080/01926230701320337> PMID: 17562483
68. Egerbacher M, Arnoczky SP, Caballero O, Lavagnino M, Gardner KL. Loss of homeostatic tension induces apoptosis in tendon cells: an in vitro study. *Clinical orthopaedics and related research*. 2008; 466(7):1562–8. Epub 2008/05/07. <https://doi.org/10.1007/s11999-008-0274-8> PMID: 18459026
69. Kannus P, Natri A. Etiology and pathophysiology of tendon ruptures in sports. *Scandinavian journal of medicine & science in sports*. 1997; 7(2):107–12. Epub 1997/04/01. <https://doi.org/10.1111/j.1600-0838.1997.tb00126.x> PMID: 9211611
70. Murrell GA, Szabo C, Hannafin JA, Jang D, Dolan MM, Deng XH, et al. Modulation of tendon healing by nitric oxide. *Inflammation research: official journal of the European Histamine Research Society [et al]*. 1997; 46(1):19–27. Epub 1997/01/01. <https://doi.org/10.1007/s000110050027> PMID: 9117513
71. Ke Q, Costa M. Hypoxia-inducible factor-1 (HIF-1). *Molecular pharmacology*. 2006; 70(5):1469–80. Epub 2006/08/05. <https://doi.org/10.1124/mol.106.027029> PMID: 16887934
72. Greijer AE, van der Wall E. The role of hypoxia inducible factor 1 (HIF-1) in hypoxia induced apoptosis. *Journal of clinical pathology*. 2004; 57(10):1009–14. Epub 2004/09/29. <https://doi.org/10.1136/jcp.2003.015032> PMID: 15452150
73. Pufe T, Petersen W, Tillmann B, Mentlein R. The angiogenic peptide vascular endothelial growth factor is expressed in foetal and ruptured tendons. *Virchows Archiv: an international journal of pathology*. 2001; 439(4):579–85. Epub 2001/11/17. <https://doi.org/10.1007/s004280100422> PMID: 11710646
74. Pufe T, Petersen WJ, Mentlein R, Tillmann BN. The role of vasculature and angiogenesis for the pathogenesis of degenerative tendons disease. *Scandinavian journal of medicine & science in sports*. 2005; 15(4):211–22. Epub 2005/07/07. <https://doi.org/10.1111/j.1600-0838.2005.00465.x> PMID: 15998338
75. Wang H, Keiser JA. Vascular endothelial growth factor upregulates the expression of matrix metalloproteinases in vascular smooth muscle cells: role of flt-1. *Circulation research*. 1998; 83(8):832–40. Epub 1998/10/20. <https://doi.org/10.1161/01.res.83.8.832> PMID: 9776730
76. Qi JH, Ebrahim Q, Moore N, Murphy G, Claesson-Welsh L, Bond M, et al. A novel function for tissue inhibitor of metalloproteinases-3 (TIMP3): inhibition of angiogenesis by blockage of VEGF binding to VEGF receptor-2. *Nature medicine*. 2003; 9(4):407–15. Epub 2003/03/26. <https://doi.org/10.1038/nm846> PMID: 12652295

77. Pufe T, Lemke A, Kurz B, Petersen W, Tillmann B, Grodzinsky AJ, et al. Mechanical overload induces VEGF in cartilage discs via hypoxia-inducible factor. *The American journal of pathology*. 2004; 164(1):185–92. Epub 2003/12/26. [https://doi.org/10.1016/S0002-9440\(10\)63109-4](https://doi.org/10.1016/S0002-9440(10)63109-4) PMID: 14695332
78. Dakin SG, Newton J, Martinez FO, Hedley R, Gwilym S, Jones N, et al. Chronic inflammation is a feature of Achilles tendinopathy and rupture. *British journal of sports medicine*. 2018; 52(6):359–67. Epub 2017/11/10. <https://doi.org/10.1136/bjsports-2017-098161> PMID: 29118051
79. Rui YF, Lui PP, Chan LS, Chan KM, Fu SC, Li G. Does erroneous differentiation of tendon-derived stem cells contribute to the pathogenesis of calcifying tendinopathy? *Chinese medical journal*. 2011; 124(4):606–10. Epub 2011/03/03. PMID: 21362289
80. Mennan C, Garcia J, McCarthy H, Owen S, Perry J, Wright K, et al. Human Articular Chondrocytes Retain Their Phenotype in Sustained Hypoxia While Normoxia Promotes Their Immunomodulatory Potential. *Cartilage*. 2019; 10(4):467–79. Epub 2018/04/20. <https://doi.org/10.1177/1947603518769714> PMID: 29671342
81. Wang H, Lindborg C, Lounev V, Kim JH, McCarrick-Walmsley R, Xu M, et al. Cellular Hypoxia Promotes Heterotopic Ossification by Amplifying BMP Signaling. *Journal of bone and mineral research: the official journal of the American Society for Bone and Mineral Research*. 2016; 31(9):1652–65. Epub 2016/03/31. <https://doi.org/10.1002/jbmr.2848> PMID: 27027798
82. Agarwal S, Loder S, Brownley C, Cholok D, Mangiavini L, Li J, et al. Inhibition of Hif1 α prevents both trauma-induced and genetic heterotopic ossification. *Proceedings of the National Academy of Sciences of the United States of America*. 2016; 113(3):E338–47. Epub 2016/01/02. <https://doi.org/10.1073/pnas.1515397113> PMID: 26721400
83. Lui PP. A practical guide for the isolation and maintenance of stem cells from tendon. *Methods in molecular biology (Clifton, NJ)*. 2015; 1212:127–40. Epub 2014/07/21. https://doi.org/10.1007/7651_2014_92 PMID: 25038747
84. Chaudhury S. Mesenchymal stem cell applications to tendon healing. *Muscles, ligaments and tendons journal*. 2012; 2(3):222–9. Epub 2013/06/06. PMID: 23738300
85. Yee Lui PP, Wong YM, Rui YF, Lee YW, Chan LS, Chan KM. Expression of chondro-osteogenic BMPs in ossified failed tendon healing model of tendinopathy. *Journal of orthopaedic research: official publication of the Orthopaedic Research Society*. 2011; 29(6):816–21. Epub 2011/04/27. <https://doi.org/10.1002/jor.21313> PMID: 21520255
86. Zhang Q, Zhou D, Wang H, Tan J. Heterotopic ossification of tendon and ligament. *Journal of cellular and molecular medicine*. 2020; 24(10):5428–37. Epub 2020/04/16. <https://doi.org/10.1111/jcmm.15240> PMID: 32293797
87. Han P, Cui Q, Yang S, Wang H, Gao P, Li Z. Tumor necrosis factor- α and transforming growth factor- β 1 facilitate differentiation and proliferation of tendon-derived stem cells in vitro. *Biotechnology letters*. 2017; 39(5):711–9. Epub 2017/02/06. <https://doi.org/10.1007/s10529-017-2296-3> PMID: 28155178
88. Rath PC, Aggarwal BB. TNF-induced signaling in apoptosis. *Journal of clinical immunology*. 1999; 19(6):350–64. Epub 2000/01/14. <https://doi.org/10.1023/a:1020546615229> PMID: 10634209
89. Kitagaki J, Iwamoto M, Liu JG, Tamamura Y, Pacifci M, Enomoto-Iwamoto M. Activation of beta-catenin-LEF/TCF signal pathway in chondrocytes stimulates ectopic endochondral ossification. *Osteoarthritis and cartilage*. 2003; 11(1):36–43. Epub 2002/12/31. <https://doi.org/10.1053/joca.2002.0863> PMID: 12505485
90. Shao JS, Cheng SL, Pingsterhaus JM, Charlton-Kachigian N, Loewy AP, Towler DA. Msx2 promotes cardiovascular calcification by activating paracrine Wnt signals. *The Journal of clinical investigation*. 2005; 115(5):1210–20. Epub 2005/04/21. <https://doi.org/10.1172/JCI24140> PMID: 15841209
91. Al-Aly Z, Shao JS, Lai CF, Huang E, Cai J, Behrmann A, et al. Aortic Msx2-Wnt calcification cascade is regulated by TNF-alpha-dependent signals in diabetic Ldlr-/- mice. *Arteriosclerosis, thrombosis, and vascular biology*. 2007; 27(12):2589–96. Epub 2007/10/13. <https://doi.org/10.1161/ATVBAHA.107.153668> PMID: 17932314
92. Lui PP, Chan KM. Tendon-derived stem cells (TDSCs): from basic science to potential roles in tendon pathology and tissue engineering applications. *Stem cell reviews and reports*. 2011; 7(4):883–97. Epub 2011/05/26. <https://doi.org/10.1007/s12015-011-9276-0> PMID: 21611803
93. Molloy TJ, Wang Y, Horner A, Skerry TM, Murrell GA. Microarray analysis of healing rat Achilles tendon: evidence for glutamate signaling mechanisms and embryonic gene expression in healing tendon tissue. *Journal of orthopaedic research: official publication of the Orthopaedic Research Society*. 2006; 24(4):842–55. Epub 2006/03/04. <https://doi.org/10.1002/jor.20093> PMID: 16514666
94. Chen X, Wang L, Qu J, Guan NN, Li JQ. Predicting miRNA-disease association based on inductive matrix completion. *Bioinformatics (Oxford, England)*. 2018; 34(24):4256–65. Epub 2018/06/26. <https://doi.org/10.1093/bioinformatics/bty503> PMID: 29939227

95. Chen X, Xie D, Zhao Q, You ZH. MicroRNAs and complex diseases: from experimental results to computational models. *Briefings in bioinformatics*. 2019; 20(2):515–39. Epub 2017/10/19. <https://doi.org/10.1093/bib/bbx130> PMID: 29045685
96. Chen X, Yan CC, Zhang X, You ZH. Long non-coding RNAs and complex diseases: from experimental results to computational models. *Briefings in bioinformatics*. 2017; 18(4):558–76. Epub 2016/06/28. <https://doi.org/10.1093/bib/bbw060> PMID: 27345524
97. Wang CC, Han CD, Zhao Q, Chen X. Circular RNAs and complex diseases: from experimental results to computational models. *Briefings in bioinformatics*. 2021. Epub 2021/07/31. <https://doi.org/10.1093/bib/bbab286> PMID: 34329377
98. Zhang L, Liu T, Chen H, Zhao Q, Liu H. Predicting lncRNA-miRNA interactions based on interactome network and graphlet interaction. *Genomics*. 2021; 113(3):874–880. Epub 2021/02/16. <https://doi.org/10.1016/j.ygeno.2021.02.002> PMID: 33588070
99. Zhang L, Yang P, Feng H, Zhao Q, Liu H. Using Network Distance Analysis to Predict lncRNA-miRNA Interactions. *Interdisciplinary sciences, computational life sciences*. 2021; 13(3):535–545. Epub 2021/07/08. <https://doi.org/10.1007/s12539-021-00458-z> PMID: 34232474
100. Liu H, Ren G, Chen H, Liu Q, Yang Y, Zhao Q. Predicting lncRNA-miRNA interactions based on logistic matrix factorization with neighborhood regularized. *Knowledge-Based Systems*. 2020; 191:105261. <https://doi.org/https%3A//doi.org/10.1016/j.knosys.2019.105261>