International Journal of General Medicine

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Identification of Key Genes and Exploration of Therapeutic Targets for Chronic Tendon Injury Based on Bioinformatics and Machine Learning

Zishen Cheng*, Weina Ren*, Yuqing Wang, Huidong Zhao, Liangliang Jiang, Xiaopeng Pu, Yaxing Zhang, Yantao Wang, Qiangjun Kang

Department of Orthopaedics, Bethune International Peace Hospital, Shijiazhuang, 050082, People's Republic of China

*These authors contributed equally to this work

Correspondence: Qiangjun Kang, Department of Orthopaedics, Bethune International Peace Hospital, No. 398 Zhongshan Road, Qiaoxi District, Shijiazhuang, 050082, People's Republic of China, Tel +86 13933093638, Email 19933000561@163.com

Background: Chronic tendon injury (CTI) is a common musculoskeletal disorder with complex molecular mechanisms, and currently lacks effective targeted therapeutic strategies. A comprehensive analysis of its key pathogenic genes and regulatory networks is crucial for the precise diagnosis and treatment of CTI.

Methods: Differentially expressed genes (DEGs) in CTI and normal tendon tissue were identified using the GEO database, and intersected with genes derived from WGCNA to identify candidate genes. Subsequently, functional enrichment analysis was performed, and four machine learning algorithms were employed to further determine key genes. Finally, a systematic functional analysis of the key genes was performed, including assessments of diagnostic value, regulatory network construction, computational drug prediction and molecular docking.

Results: A total of 271 candidate genes were identified, which were significantly enriched in focal adhesion, ECM-receptor interaction, and p53 signaling pathway. Subsequently, three key genes (*FER, TUBA1B*, and *MICAL2*) were prioritized through machine learning analysis, and their marked upregulation in CTI samples was verified by qRT-PCR and immunohistochemical analysis. Furthermore, their expression levels were positively correlate with natural killer T cell infiltration. TF-mRNA-miRNA regulatory network revealed the predicted TFs (such as *STAT3, TFAP4, JUN, MYC*) and the miRNAs that interact with the key genes. Ultimately, drug screening and molecular docking identified several potential lead compounds and confirmed their stable binding patterns.

Conclusion: This study systematically revealed three key genes in CTI through comprehensive bioinformatics analysis. The diagnostic model, regulatory network, and predicted targeted drugs constructed based on these findings laid a solid theoretical foundation for subsequent translational medical research.

Keywords: chronic tendon injury, bioinformatic analysis, machine learning, TF-mRNA-miRNA molecular network, natural killer T cells

Introduction

Tendon injuries are common musculoskeletal disorders involving damage to the structural integrity of tendons.[1] Chronic tendon injury (CTI), also known as tendinopathy, is a degenerative condition caused by the repeated excessive load on the tendon.[2,3] Clinical evidence suggests that, in addition to traditional factors like excessive mechanical stress and aging, abnormal local anatomy plays a key role in certain tendon disorders.[4] Specifically, the accessory navicular bone is recognized as a significant local factor contributing to posterior tibial tendon dysfunction and degeneration.[5] In the United States, approximately 33 million cases of musculoskeletal injuries occur each year, of which about 50% involve tendons and ligaments.[6] A cross-sectional study of Brazilian high-performance athletes revealed that 30.6% were diagnosed with tendinopathy.[7] Tendon injuries not only cause physical pain for patients but also significantly impact their mental health, quality of life, and work capacity, leading to extensive social and economic impacts.[1,8] Currently, the

https://doi.org/10.2147/IJGM.S591312
Received: 23 December 2025
Accepted: 2 May 2026
Published: 14 May 2026

treatment of CTI primarily relies on local or systemic anti-inflammatory medications to alleviate pain. However, this approach often only provides short-term relief, and for patients with prolonged symptoms or severe conditions, the therapeutic effect may be suboptimal or even ineffective.[9] Furthermore, although progress has been made in areas such as gene therapy and stem cell therapy, these treatments have not yet been widely adopted in clinical practice.[9] Therefore, clarifying its molecular mechanism and providing key evidence for the search for core-targeted therapeutic strategies is of particular importance.

The core pathological feature of CTI is an imbalance between repair and degeneration processes. This disease is not driven by a single linear pathway but rather involves a complex regulatory network formed by the interplay of multiple biological processes, including dysregulated extracellular matrix metabolism, chronic immune responses, tenocyte dysfunction, and pathological neurovascular invasion.[10] With the widespread adoption of gene microarray and RNA sequencing technologies, the accumulation of transcriptomic data for CTI in public databases presents unprecedented opportunities to systematically dissect this complex regulatory network at the genomic level. However, the existing studies lack a comprehensive strategy that uses integrative analysis of WGCNA with multiple machine learning algorithms to screen out highly reliable key genes in CTI. To address this gap, this study constructs a comprehensive research framework that employs an integrated analytical approach, encompassing a diagnostic model, immune infiltration analysis, a TF-mRNA-miRNA regulatory network, and potential drug prediction. It aims to uncover the underlying pathophysiological mechanisms and develop innovative therapeutic strategies, which holds significant importance for improving patient outcomes and advancing sports medicine.

Methods

Data Collection and Analysis

This study leveraged publicly available gene expression databases to identify key mRNA molecules associated with CTI. The GSE26051 dataset (GPL570 platform) used for differential expression analysis comprised 46 human tendon tissue samples, including 23 CTI samples and 23 corresponding normal tendon controls. After gene annotation and probe deduplication, 15977 valid mRNAs were ultimately obtained for subsequent bioinformatics analysis.

This study further selected dataset GSE17591 (GPL17021 platform) as a validation cohort, which contained 16 mice tendon tissue samples (8 chronic injury samples and 8 normal controls) for independent validation of the screened key genes.

Screening for Differentially Expressed Genes (DEGs)

Based on the GSE26051 dataset, we used the "limma" software package to conduct differential analysis of the mRNA expression profile. The DEGs were identified with P value < 0.05 and $|\log_2 FC| > 0.5$. To further focus on the genes with the most significant changes, we extracted the top 20 genes with the greatest expression level changes and plotted a heatmap for visual display.

Construction of Co-Expression Network and Identification of Key Modules

Based on the disease state as a phenotypic value, the weighted gene co-expression network analysis (WGCNA) was conducted to systematically identify co-expressed gene modules significantly associated with CTI. By constructing a scale-free co-expression network and calculating module-phenotype association scores, the core module with the highest correlation to the disease phenotype was ultimately selected. All genes within this module were defined as a candidate gene set for CTI, serving as the basis for subsequent in-depth analysis.

Screening Candidate Gene and Enrichment Analysis

The intersection of DEGs and CTI-related genes identified by WGCNA was selected as candidate genes. Subsequently, the "clusterProfiler" software package was used to perform Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes pathway (KEGG) pathway enrichment analyses on the candidate genes, with a significance threshold set at $P < 0.05$. To further explore potential biological mechanisms, gene set enrichment analysis (GSEA) was concurrently

performed. Using "c2.cp.kegg.v7.2.symbols.gmt" as the reference gene set and the list of all genes from the differential expression analysis sorted by LogFC values, we systematically identified regulatory pathways associated with CTI.

Machine Learning Analysis for Key Genes Identification

To further refine the selection of key genes from the candidate pool, we employed integrative strategy combining four distinct machine learning algorithms: Random Forest, Boruta feature selection, XGBoost, and LASSO regression. Specifically, a random forest model was constructed using the randomForest package. The optimal mtry value was determined through parameter optimization, and the top 20 genes were selected based on feature importance scores. The Boruta package was used for comprehensive relevant feature screening, retaining all important and tentative features. LASSO regression analysis was performed using the glmnet package, with 10-fold cross-validation to determine the optimal lambda.min value, and the corresponding non-zero coefficient genes were extracted. An XGBoost model was constructed using the xgboost package, and the top 20 genes were selected based on their contribution scores. Ultimately, the genes identified as common intersections by all four methods were designated as key genes for subsequent validation and analysis.

ROC Analysis and Expression Validation of Key Genes

Based on the GSE26051 dataset, this study employed receiver operating characteristic (ROC) curves to evaluate the diagnostic efficacy of the key genes in distinguishing CTI from normal tissues. The area under the curve (AUC) was calculated to quantify the classification ability of each gene. Furthermore, the expression trends of key genes between the disease and control groups were analyzed to verify the differences in their expression.

Quantitative Reverse Transcription Polymerase Chain Reaction (qRT-PCR)

A total of 20 human tendon tissue samples were included in this study, comprising 9 degenerative posterior tibial tendon samples associated with accessory navicular bone and 11 normal control tissues. Total RNA was extracted using TRIzol reagent and reverse-transcribed into cDNA following the instructions of the FastKing cDNA First Strand Synthesis Kit (TIANGEN, China). Subsequently, PCR amplification was performed using SuperReal PreMix Plus (SYBR Green) (TIANGEN, China). Relative quantitative analysis was conducted via the $2^{-\Delta\Delta CT}$ method with a Gene-9660 fluorescence quantitative PCR instrument (BIOER, China). The primer sequences used were listed in Supplementary Table 1.

Immunohistochemistry (IHC) Analysis

The expression of three key genes was analyzed using degenerative posterior tibial tendon samples (CTI) associated with accessory navicular bone and normal control tissues (Normal-1 and Normal-2) by IHC. Paraffin-embedded sections were dewaxed in xylene and rehydrated through a graded ethanol series, followed by antigen retrieval performed in 0.01 M sodium citrate buffer (pH 6.0) using a pressure cooker for 2.5 min. Endogenous peroxidase activity was blocked by incubating the sections with 3% hydrogen peroxide (BosterBio, China) at room temperature in the dark for 30 min. Non-specific binding was blocked with 10% normal goat serum at room temperature for 30 min. The sections were then incubated with FER (18W3) Mouse Monoclonal Antibody (1:150, BD-PB4099, Biodragon, China), TUBA1B (16V18) Rabbit Monoclonal Antibody (1:150, RM8134, Biodragon, China) or Anti-MICAL2 Antibody (1:150, A08683-1, BosterBio, China) at 4°C overnight. Subsequently, the sections were incubated with Primary Antibody Amplifier Quanto (TL-125-QPB, Epredia, USA) at room temperature for 10 min, followed by incubation with HRP Polymer Quanto (TL-125-QPH, Epredia, USA) at room temperature for 10 min. Diaminobenzidine (DAB) was used for chromogenic detection, followed by counterstaining with hematoxylin, bluing with a bluing solution, dehydration, and mounting. Images were acquired using a microscope (ECLIPSE Si RS, Nikon, Japan).

Nomogram Construction and Model Evaluation

To enhance the clinical applicability of key genes, a nomogram prediction model based on key genes was constructed using the R language "rms" package. This model integrates information from multiple key genes to generate a comprehensive score for individual patients and predict disease risk. Furthermore, the predictive accuracy,

discriminative ability, and clinical utility of the nomogram model were systematically evaluated by plotting calibration curves, ROC curves, and decision curve analysis (DCA).

Immune Infiltration Analysis

To explore the potential relationship between immune cell infiltration and key genes, the "ssGSEA" algorithm (https://ckg.readthedocs.io/en/latest/notebooks/recipes/ssGSEA_with_PCA.html) was employed to assess differences in immune cell infiltration levels within the immune microenvironment between the disease group and the control group.

Molecular Regulatory Networks Construction

To systematically elucidate the molecular regulatory mechanisms of key genes in CTI, this study constructed a molecular regulatory network involving interactions between transcription factors (TFs), miRNAs, and key genes. First, miRNAs associated with key genes were obtained from TargetScanHuman 8.0, miRDB, and miRTarBase databases. The intersection of these three datasets yielded high-confidence miRNA-mRNA regulatory pairs. Subsequently, potential transcription factors regulating key genes were predicted using the TRRUST online database (screening criterion: $P < 0.05$) and the ChIP-Atlas platform. The predictions from both sources were intersected to obtain reliable TF-mRNA regulatory pairs. Finally, integrating these regulatory relationships, we constructed a three-tiered molecular regulatory network (TF-mRNA-miRNA) using Cytoscape software to visually depict the transcriptional and post-transcriptional regulatory architecture of key genes.

Potential Drug Prediction and Molecular Docking Validation

The potential therapeutic drugs were selected based on the Drugs.com database (https://www.drugs.com/). Subsequently, the AutoDock software was employed to perform molecular docking simulations for the selected drugs and the corresponding proteins of key genes, analyzing their binding modes, affinities, and stabilities, and using PyMOL to visualize the results.

Results

Identification of DEG in CTI

A total of 1954 DEGs were identified between the CTI group and the normal tendon group, among which 941 genes were upregulated and 1013 genes were downregulated (Figure 1A). We constructed heatmaps to represent the expression levels of the 10 most up-regulated and 10 most down-regulated genes (Figure 1B).

Identification of Co-Expression Gene Modules

We performed WGCNA based on disease status phenotypes to systematically identify gene co-expression modules significantly associated with CTI. By setting the topological overlap matrix threshold to 0.9, the optimal soft threshold was determined to be 9 (Figure 2A), and a total of 13 gene co-expression modules were constructed (excluding the grey module) (Figure 2B and C). Module-trait relationship analysis revealed that the MEgreen module exhibited the strongest positive correlation with CTI (correlation coefficient = 0.44, $P = 0.002$) (Figure 2C). Based on this, we defined the 959 genes contained in the MEgreen module as a CTI-related gene set and included them for further in-depth analysis.

Module Genes Identification and Functional Enrichment Analysis

Intersection analysis of the 1954 DEGs and the 959 WGCNA-derived CTI-related genes identified a final set of 271 overlapping genes as candidates for further analysis (Supplementary Figure 1). These genes are implicated in the pathological processes of CTI. The GO enrichment analysis revealed biological processes were mainly enriched in pathways such as extracellular matrix (ECM) organization, extracellular structure organization and external encapsulating structure organization (Figure 3A). The cellular component analysis was mainly enriched in pathways such as endoplasmic reticulum lumen, collagen-containing ECM and outer kinetochore (Figure 3B). The molecular functions were mainly enriched in pathways such as collagen binding, ECM structural constituent and integrin binding (Figure 3C).

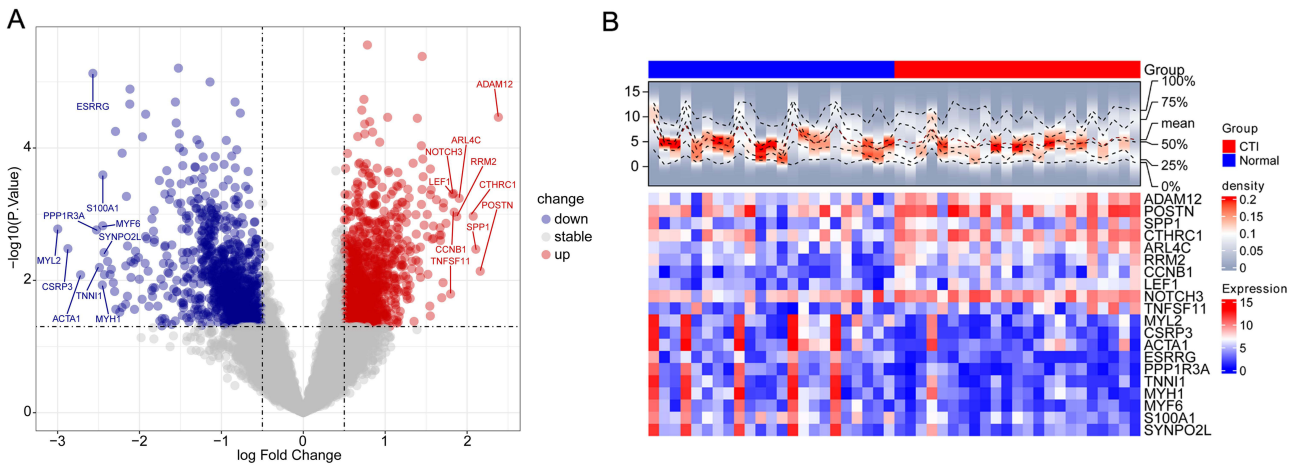

Figure 1 Identification of DEGs in the GSE26051 dataset. (**A**) Volcano plot of all DEGs. (**B**) Heatmap of the top 20 DEGs in the CTI group and normal control group.

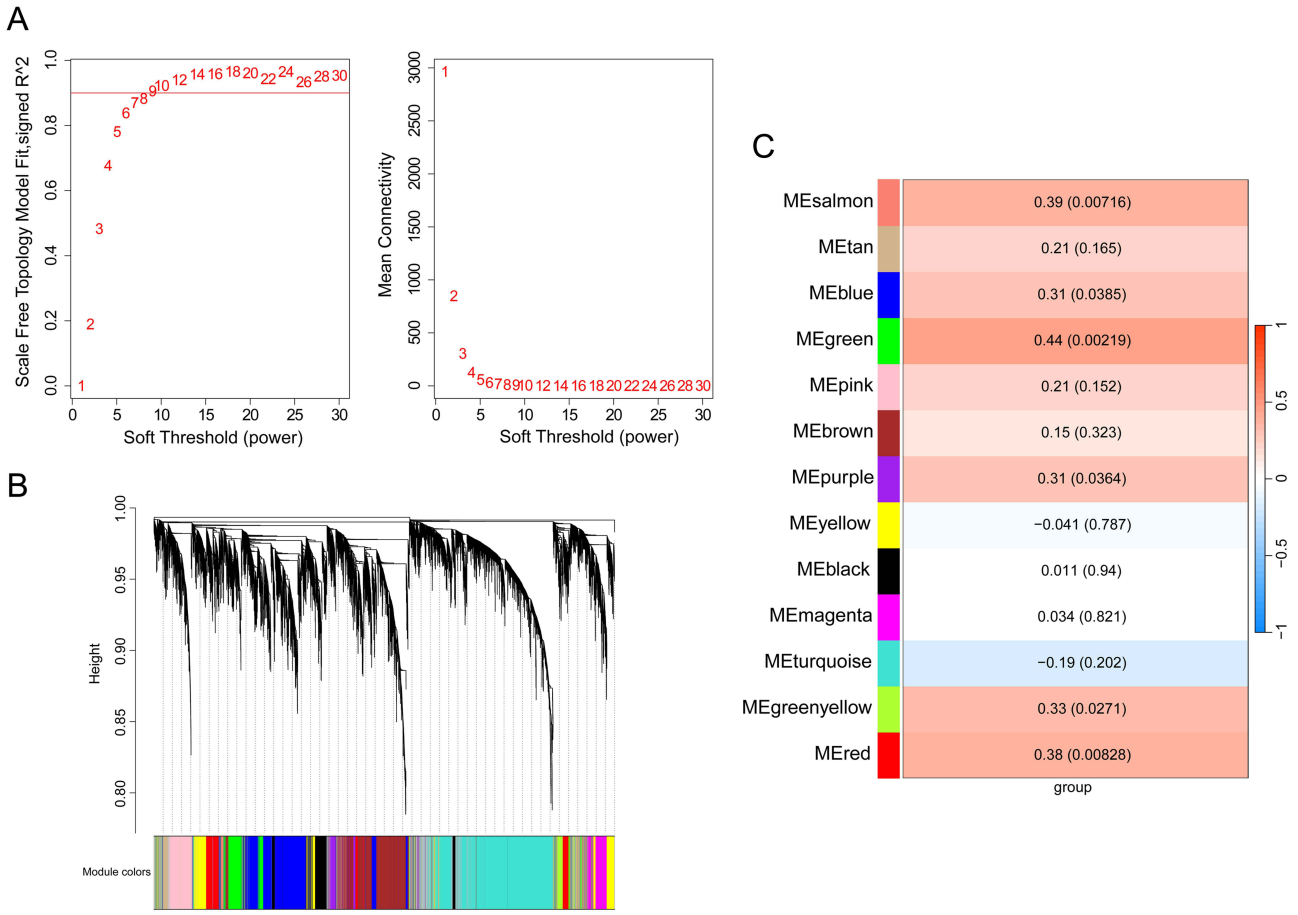

Figure 2 Construction of the weighted gene co-expression analysis network for CTI. (**A**) Soft threshold screening by Independence and mean connectivity analysis. (**B**) Cluster dendrogram of genes associated in key modules. (**C**) Module-trait relationships.

Furthermore, the KEGG pathway enrichment analysis revealed a significant enrichment in focal adhesion, ECM-receptor interaction, and p53 signaling pathway (Figure 3D). Subsequently, the GSEA results showed that ECM-receptor interaction and p53 signaling pathway were significantly enriched in the CTI group (Figure 3E).

https://doi.org/10.2147/IJGM.S591312

Feature Selection Through Various Machine Learning Methods

Four machine learning algorithms identified 20 (Random Forest), 34 (Boruta), 20 (XGBoost) and 17 (LASSO) feature genes, respectively (Figure 4A–E). The Venn diagram revealed that *FER, TUBA1B*, and *MICAL2* genes were common to all four methods (Figure 4F), thus designated as key genes for subsequent analysis.

ROC Curve and Expression Verification of the Key Genes

To validate the diagnostic efficacy of key genes, we constructed ROC curves based on the GSE26051 dataset and calculated the AUC. As shown in Figure 5A, the AUC values for *FER, TUBA1B*, and *MICAL2* reached 0.87, 0.88, and 0.85, respectively, indicating that all three possessed outstanding disease state discrimination capabilities. Further analysis revealed that the expression levels of these genes were significantly upregulated in the CTI group

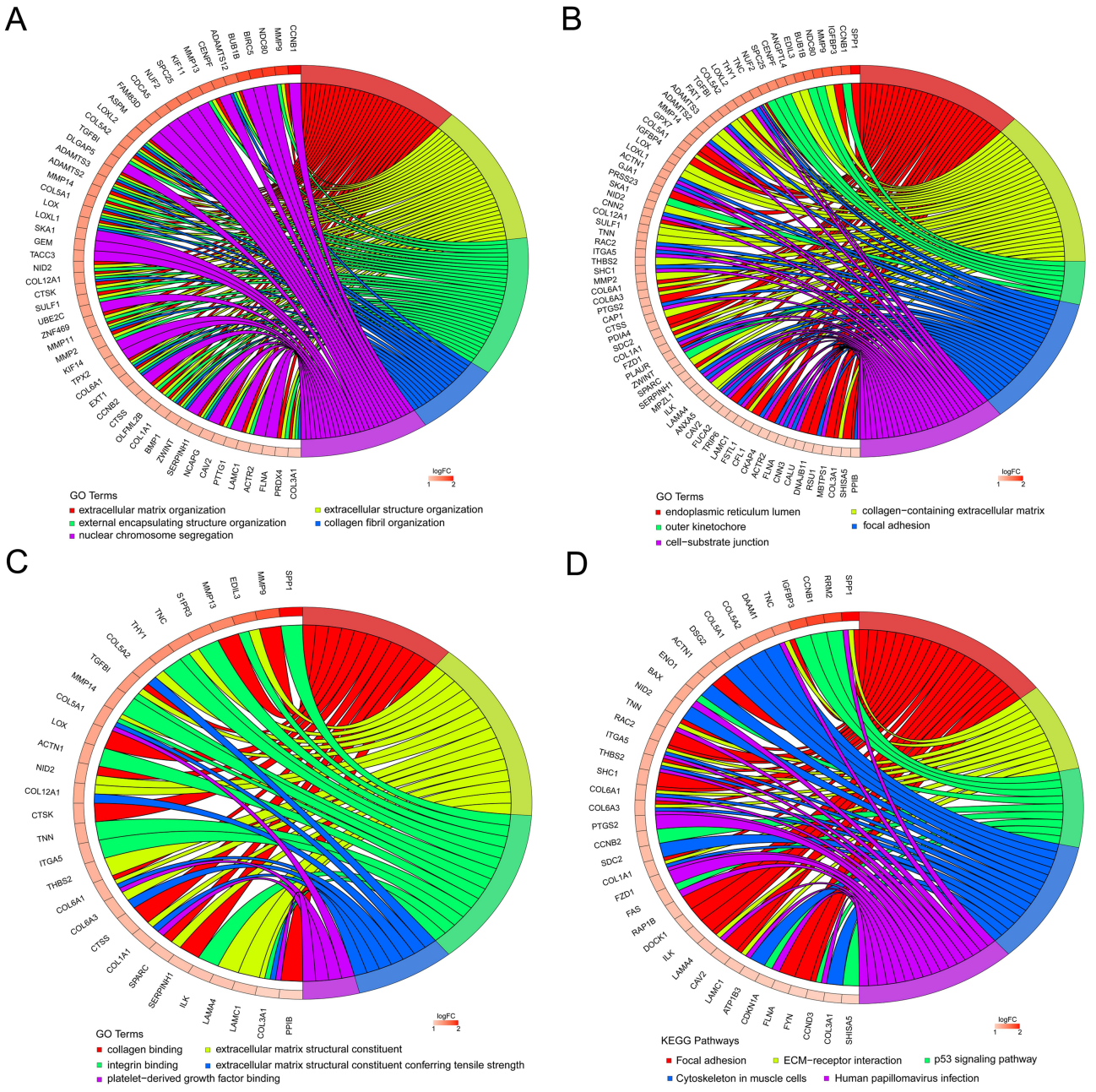

Figure 3 continued.

https://doi.org/10.2147/IJGM.S591312

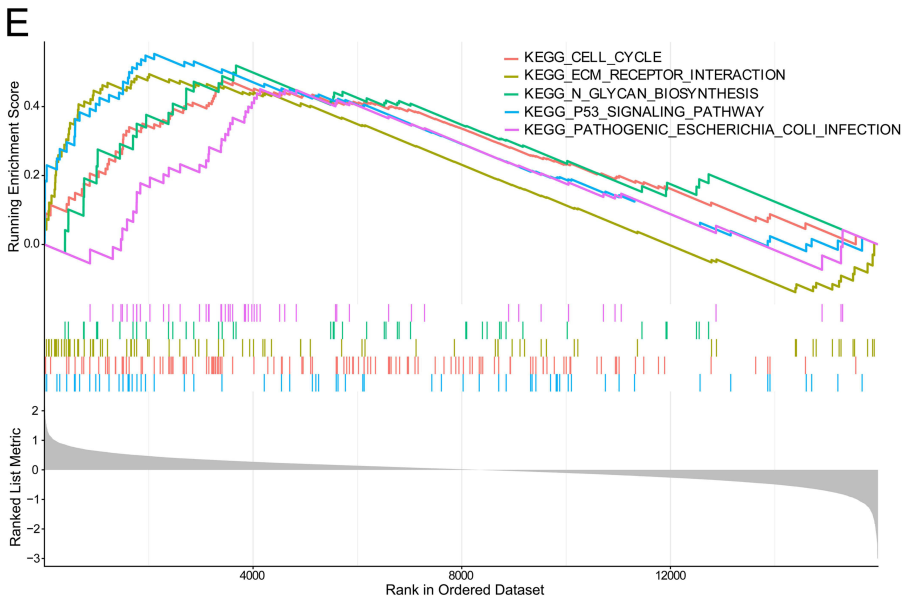

Figure 3 Functional enrichment analysis. (**A**) GO biological processes analysis. (**B**) GO cellular component analysis. (**C**) GO molecular functions analysis. (**D**) KEGG pathway analysis. (**E**) Top 5 of GSEA.

(Figure 5B). To confirm the universality of this expression pattern, we validated the findings in an independent validation cohort (GSE17591), where *FER, TUBA1B*, and *MICAL2* were consistently observed to be significantly overexpressed in the CTI group (Figure 5C). In addition, the expression of key genes in tendon injuries and normal tissues of CTI patients was evaluated using qRT-PCR. The results also confirmed the significant high expression of *FER, TUBA1B*, and *MICAL2* in CTI tissues (Figure 5D). IHC analysis further confirmed the same trend (Figure 5E).

Construction of Nomogram

To enhance the clinical translational potential of the model, we integrated key genes to construct a nomogram model (Figure 6A). The ROC results demonstrated excellent discriminative ability of the model with an AUC of 0.977 (Figure 6B). The calibration curve confirmed a high consistency between the predicted probabilities and the actual observed probabilities (P = 0.78) (Figure 6C). Decision curve analysis (DCA) further revealed that the nomogram provides a favorable clinical net benefit across a wide range of threshold probabilities, demonstrating its remarkable accuracy and stability for clinical application (Figure 6D).

Immune Characteristic Assessment and Correlation Analysis

Immune infiltration analysis was performed on GSE26051. A total of six immune cell types with significant infiltration level differences were identified, including activated $CD8^+$ T cells, effector memory $CD4^+$ T cells, type 17 helper T cells, activated B cells, memory B cells, and natural killer T (NKT) cells (Figure 7A). Further, we analyzed the correlation between the expression levels of key genes and the above immune cell scores (Figure 7B). By screening the gene-immune cell pairs with strong correlation ($|cor| > 0.5$, $P < 0.01$) and visualizing them, it was found that NKT cells were significantly positively correlated with all three key genes (Figure 7C). This result suggests that these genes may collaborate with NKT cells to participate in pathological processes such as immune dysregulation, abnormal cell migration, and disordered tissue remodeling.

Construction of Molecular Regulatory Networks

To further investigate the molecular regulatory mechanisms of key genes in CTI, we systematically constructed a regulatory network involving TFs and miRNAs. Among the high-confidence miRNA-mRNA pairs identified, *FER,*

https://doi.org/10.2147/IJGM.S591312

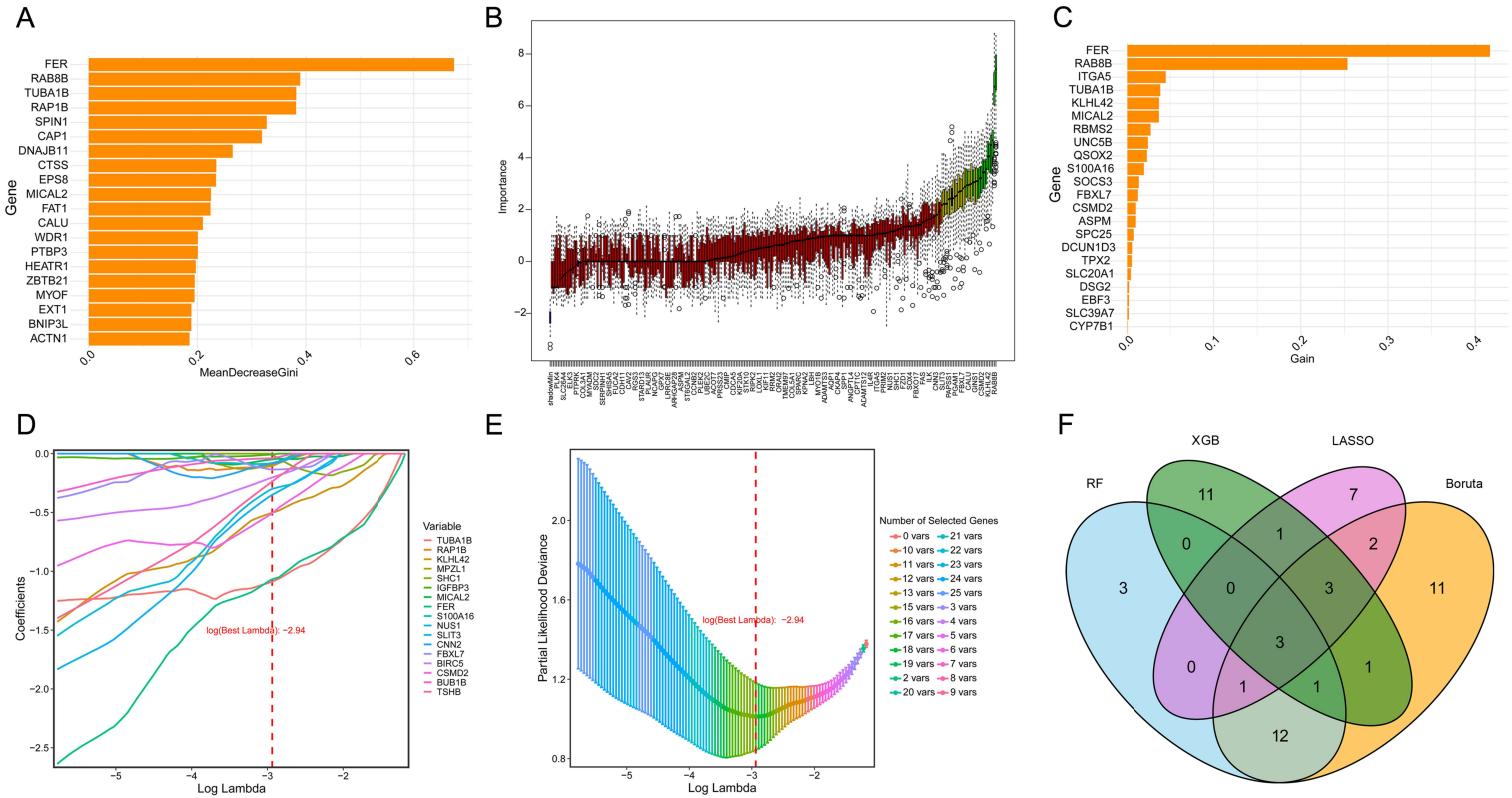

Figure 4 Identify key genes using four machine learning methods. (**A**) The top 20 feature genes selected by the random forest model. (**B**) Boruta feature importance ranking (green, yellow and red indicate important, tentative and rejected variables, respectively). (**C**) Bar chart of feature importance for genes identified by the XGBoost method. (**D** and **E**) LASSO regression model feature selection. (**F**) Venn diagram of the intersected genes of four machine learning algorithms.

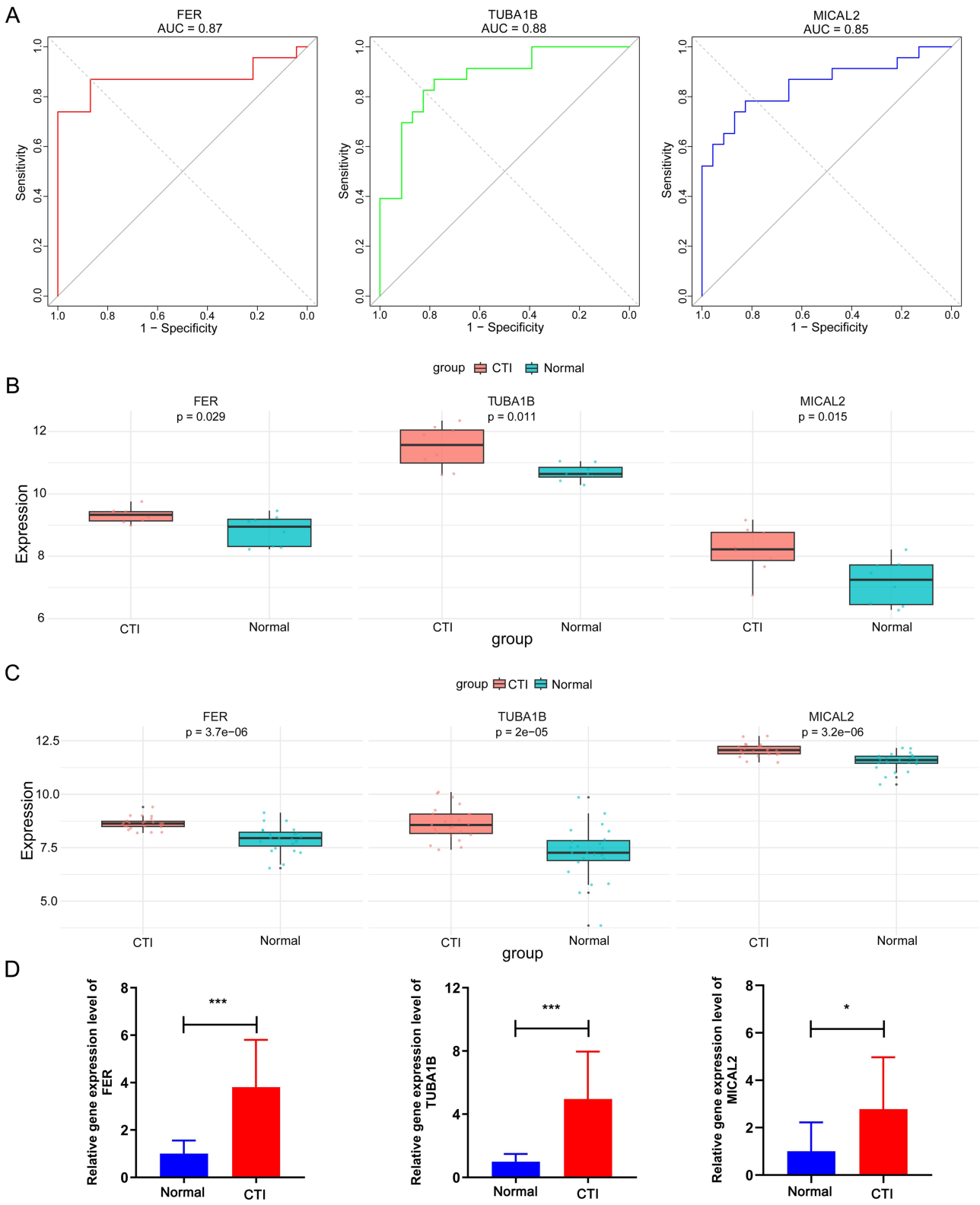

Figure 5 continued.

https://doi.org/10.2147/IJGM.S591312

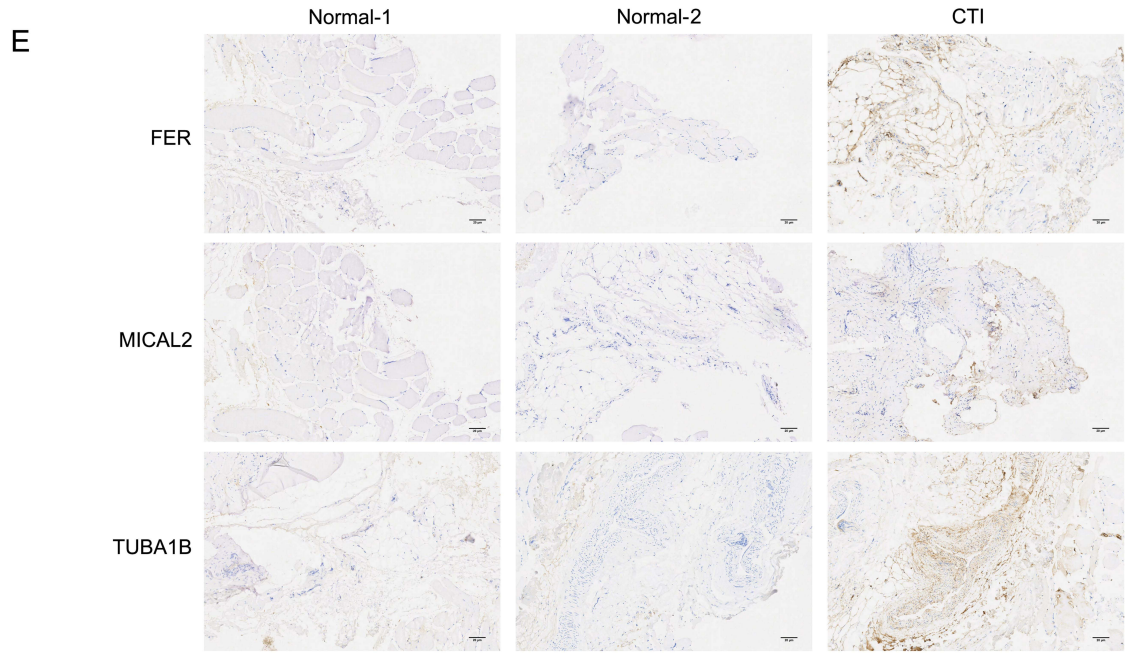

Figure 5 ROC curve and expression verification of the key genes. (**A**) ROC curves of selected genes evaluated on GSE26051 dataset. (**B**) Expression levels of the selected genes between normal and CTI groups based on GSE26051 dataset. (**C**) Expression verification of the selected genes between normal and CTI groups based on GSE17591 dataset. (**D**) Relative expression of three selected genes based on qRT-PCR. **P* < 0.05, ****P* < 0.01. (**E**) The protein level of three selected genes based on IHC analysis. Bar=20 μm. Normal-1 and Normal-2 depict specimens from two different patients with healthy tendon tissue.

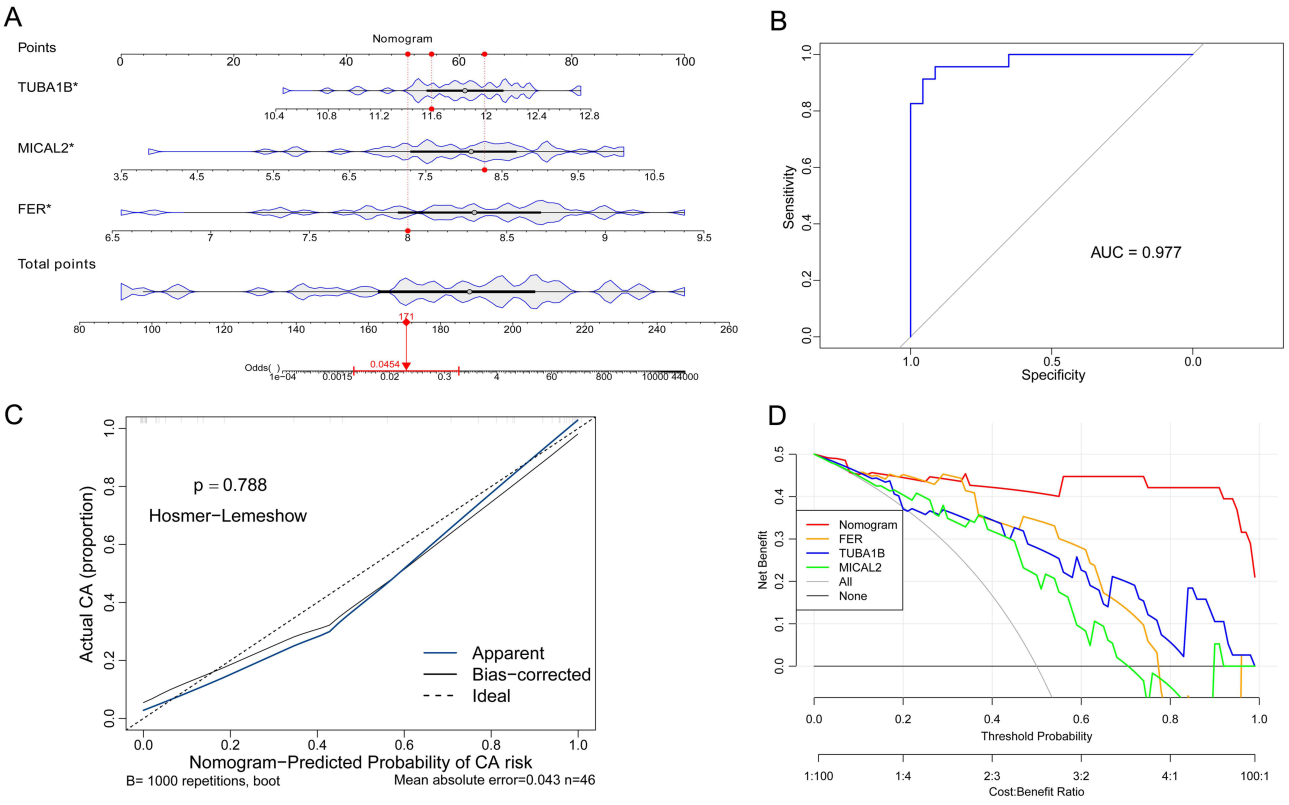

Figure 6 Construction of nomogram and validation. (**A**) Nomogram for predicting individual risk of CTI based on expression levels of key genes. Red line indicates the scoring path for a single example patient. **P* < 0.05 (**B**) ROC curve evaluating the overall diagnostic performance. (**C**) Calibration curve assessing the agreement between predicted and observed outcomes. (**D**) Decision curve evaluating the clinical net benefit.

TUBA1B, and *MICAL2* were predicted to be targeted by 14, 4, and 12 miRNAs, respectively. Subsequently, at the TFs level, by integrating prediction results from TRRUST and ChIP-Atlas, we found that *FER, TUBA1B*, and *MICAL2* may be regulated by 23, 21, and 34 TFs, respectively. Ultimately, we integrated all regulatory relationships to construct a comprehensive TF-mRNA-miRNA molecular network (Figure 8). Notably, *STAT3, TFAP4, JUN*, and *MYC* were predicted to be common TFs regulating all three genes.

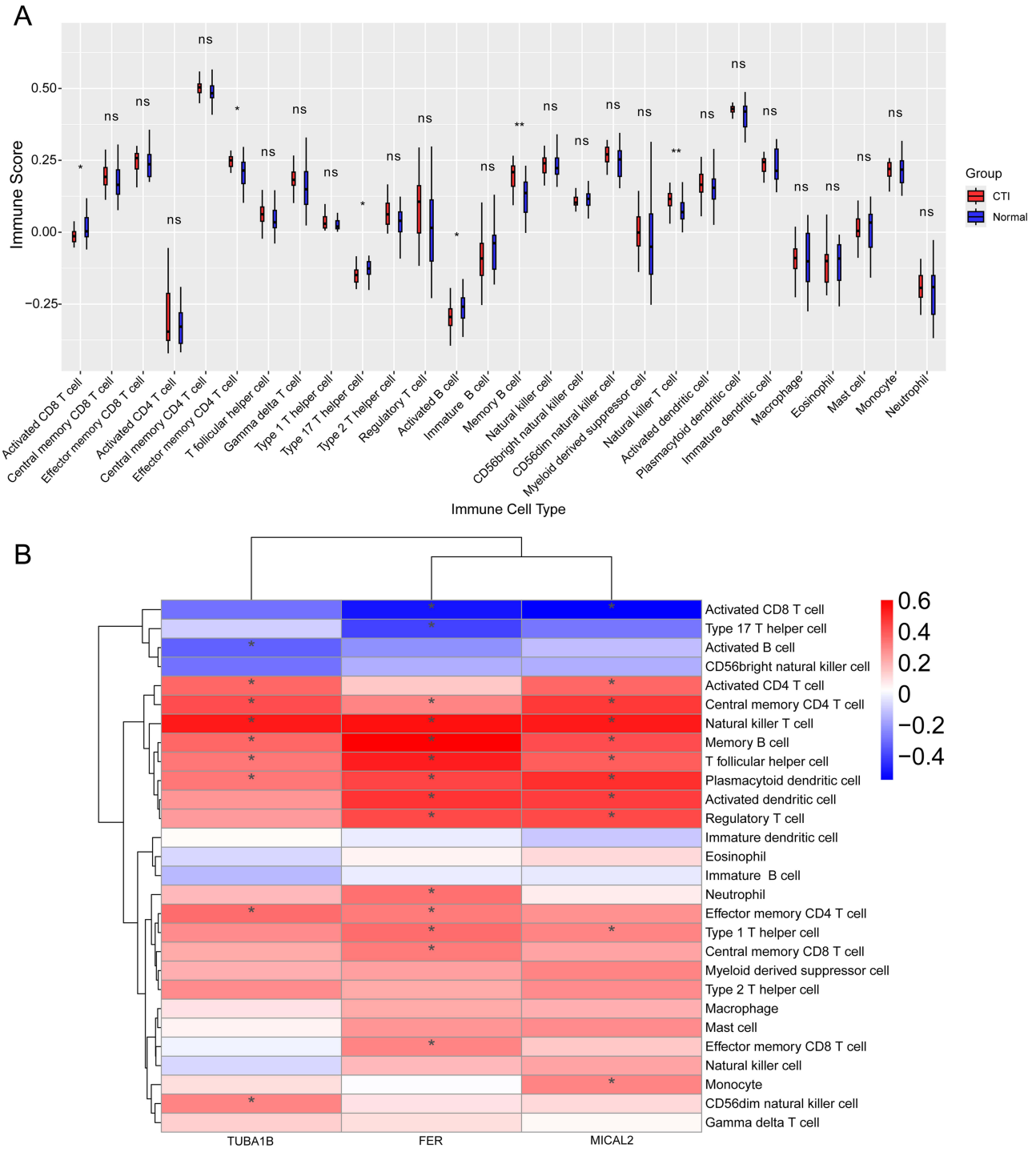

Figure 7 continued.

https://doi.org/10.2147/IJGM.S591312

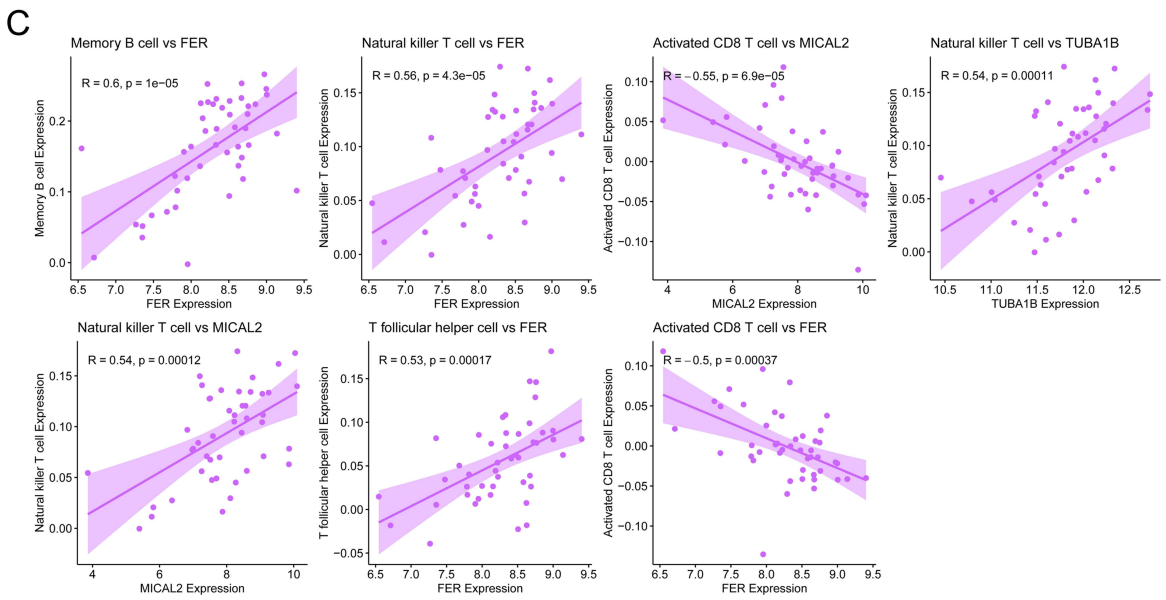

Figure 7 Immune infiltration analysis between CTI and normal groups. (**A**) Box plot showing differences in immune infiltration between CTI and normal samples. $^{ns}P > 0.05$, $^{*}P < 0.05$, $^{**}P < 0.01$. (**B**) Correlation between immune cells and expression of key genes. $^{*}P < 0.05$. (**C**) Screening the gene-immune cell pairs with strong correlation.

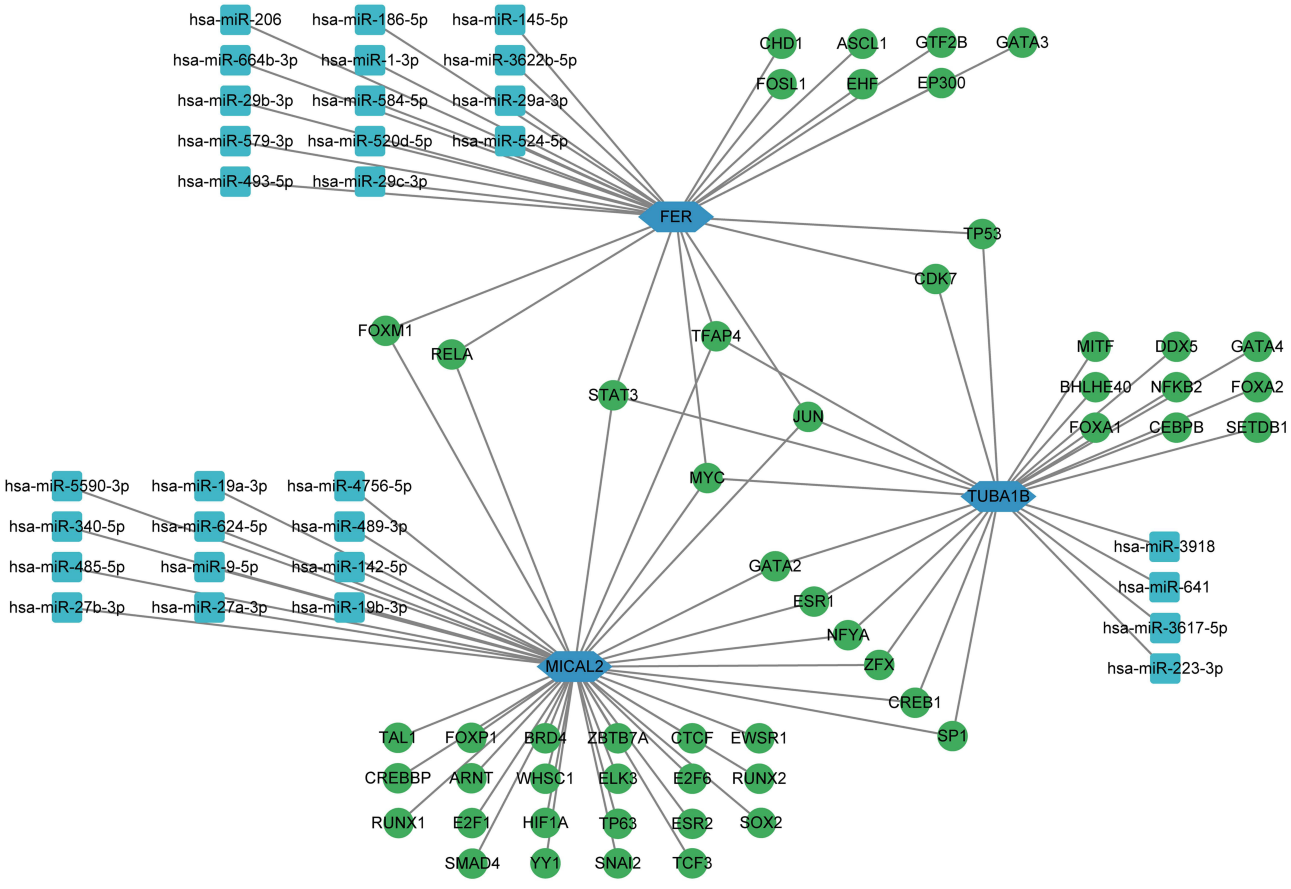

Figure 8 Visualization of TF-mRNA-miRNA molecular network.

https://doi.org/10.2147/IJGM.S591312

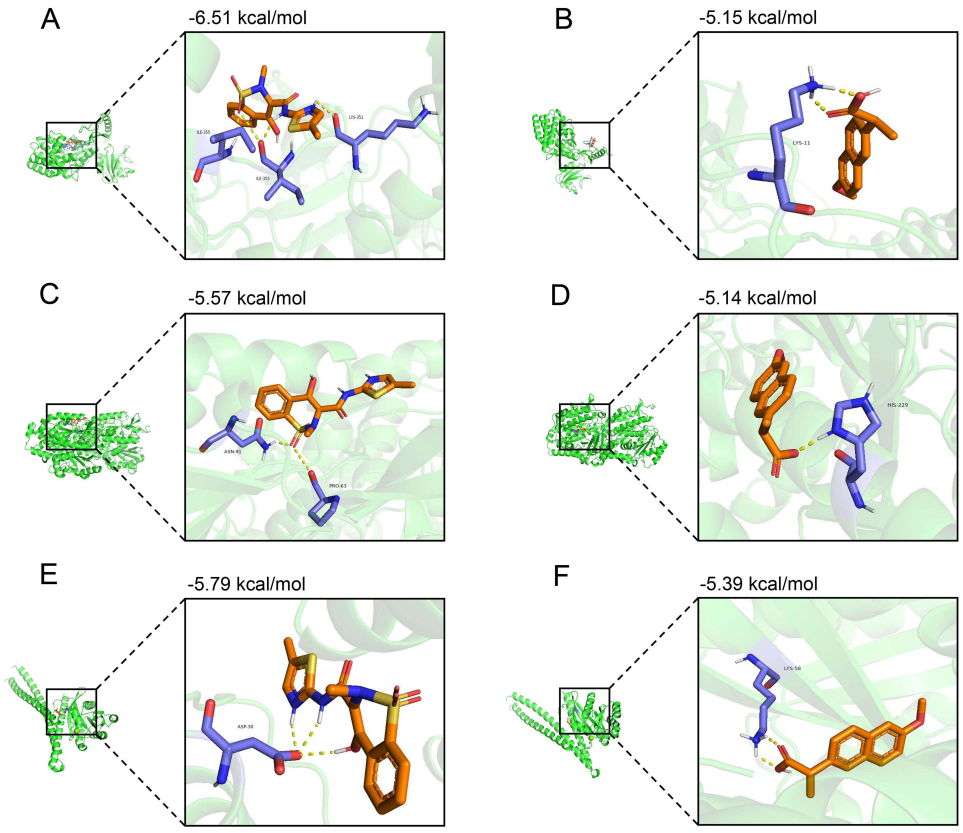

Figure 9 Three-dimensional schematic diagram of the molecular docking model. (**A** and **B**) Binding modes of FER with (**A**) meloxicam and (**B**) naproxen. (**C** and **D**) Binding modes of TUBA1B with (**C**) meloxicam and (**D**) naproxen. (**E** and **F**) Binding modes of MICAL2 with (**E**) meloxicam and (**F**) naproxen.

Potential Drug Prediction and Molecular Docking

To identify potential therapeutic drugs for CTI with higher clinical translational potential, we screened the Drugs.com database for clinically relevant non-steroidal anti-inflammatory drugs (NSAIDs) and selected two drugs, meloxicam and naproxen. Notably, the candidate drugs demonstrated favorable binding characteristics with the target protein in subsequent molecular docking simulations (*FER*-meloxicam/naproxen, *TUBA1B*-meloxicam/naproxen, *MICAL2*-meloxicam/naproxen), suggesting its potential for further development as a therapeutic strategy for tendon repair (Figure 9A–F).

Discussion

The pathological mechanisms of CTI are highly complex, making it difficult for traditional single differential expression analyses to reveal its systemic regulatory core. To address this challenge, this study successfully identified *FER, TUBA1B*, and *MICAL2* as key genes in CTI by integrating differential transcriptome analysis, WGCNA, and machine learning algorithms. Through in-depth analysis of these genes, this study revealed their enriched core pathways, constructed a diagnostic model, and predicted upstream regulatory networks and targeted drugs. In addition, the immune infiltration characteristics of CTI were characterized. These integrated findings lay a solid foundation for advancing precision medicine in CTI.

Following differential mRNA analysis and WGCNA, we identified 271 CTI-related candidate genes and performed functional enrichment analysis. GO analysis revealed that these candidate genes were highly enriched in molecular functions such as ECM structural constituent, collagen binding, and integrin binding. This indicates that the imbalance of ECM homeostasis may be a key link in the disease process. As a highly specialized connective tissue, tendons rely heavily on a structurally intact and well-organized ECM to maintain their biomechanical function.[11] Our findings indicate that under chronic injury conditions, tendon cells may initiate abnormal ECM remodeling processes, leading to disrupted

https://doi.org/10.2147/IJGM.S591312

collagen metabolism. This, in turn, compromises tendon structural integrity and weakens their mechanical properties. KEGG pathway enrichment analysis was subsequently performed, which revealed that the candidate genes were significantly enriched in focal adhesion, ECM-receptor interaction, and p53 signaling pathway. The co-enrichment of ECM-receptor interactions and focal adhesion suggests abnormal communication between cells and the extracellular microenvironment.[12] In chronic injuries, tendon cells perceive abnormal mechanical stress and biochemical signals through receptors such as integrins, primarily transducing these signals into the cell via the focal adhesion pathway.[12,13] In particular, the significant enrichment of the p53 signaling pathway provides a crucial perspective for elucidating the repair failure mechanism in CTI. As a central regulatory hub of cellular stress responses,[14] the p53 pathway may be activated by abnormal focal adhesion signaling or direct mechanical stress. Once initiated, this pathway directly impairs tissue regenerative capacity by inducing senescence or apoptosis in tendon cells.[15] In summary, the functional enrichment results highlight the concurrent involvement of key biological processes in CTI, including focal adhesion, ECM-receptor interaction, and p53 signaling pathway. The co-enrichment of these processes warrants further investigation into their potential causal relationships and roles in driving CTI.

A core finding of this study lies in the identification of three key CTI-related genes (*FER, TUBA1B*, and *MICAL2*) through the application of multiple machine learning algorithms. FER, as a non-receptor tyrosine kinase, is known to play a central role in cell adhesion, cytoskeletal reorganization, and growth factor signaling.[16,17] Tubulin alpha 1b (TUBA1B) is one of the key tubulin subunits and a core component of the cytoskeleton.[18] The cytoskeleton not only serves as the foundation for maintaining cell shape and withstanding mechanical forces, but also functions as a vital platform for intracellular material transport and signal transduction.[19] MICAL2 is a class of NADPH-dependent oxidoreductases that can disassemble the cytoskeleton by oxidatively cleaving tubulin, thereby directly participating in the dynamic remodeling of the cytoskeleton.[20] In summary, these three key genes play a crucial role in regulating the mechanical sensing and structural stability of tendon cells. Given that FER, TUBA1B, and MICAL2 are all critically involved in cytoskeletal organization, these genes may cooperatively govern cytoskeletal remodeling in tendon cells. As mentioned earlier, FER modulates cytoskeletal dynamics and cell adhesion, TUBA1B maintains the structural integrity of microtubules, and MICAL2 regulates microtubule disassembly. Together, they shape the mechanical properties, mechanosensing, and physiological function of tendon cells. Dysregulation of this axis may disrupt cytoskeletal homeostasis, impair tendon cell adaptation to mechanical stress, and thereby contribute to the onset and progression of CTI. Furthermore, we hypothesize that dysregulation of *FER, TUBA1B*, and *MICAL2*-related cytoskeletal regulatory networks may represent a common molecular hub underlying CTI across diverse etiologies. Multiple studies have demonstrated that the presence of an accessory navicular alters the insertion and force vector of the posterior tibial tendon, creating a chronic attrition zone which predisposes to tendinosis and eventual tear.[21,22] To validate and explore its significance in specific clinical contexts, we performed qRT-PCR, which revealed significantly upregulated key genes in injured tendons derived from the accessory navicular bone, fully consistent with bioinformatics analysis findings. The protein level of three key genes also showed a similar trend based on IHC analysis. These findings suggest that while the accessory navicular bone serves as a unique anatomical driver, the molecular mechanisms underlying the degenerative changes it induces in posterior tibial tendons may converge upon pathways shared with other CTI. However, the causal relationship between the upregulation of these genes and the progression of CTI requires further experimental validation.

A significant finding in the immune characterization assessment was that the expression levels of the three key genes *FER, TUBA1B*, and *MICAL2* were all significantly positively correlated with the infiltration level of NKT cell. As a bridge connecting innate and adaptive immunity, NKT cells possess potent immunoregulatory and cytokine secretion functions.[23,24] This strong association suggests that there might be a synergistic mechanism between key genes and NKT cells, jointly regulating the tendon repair process. Additionally, to translate the discovery of key genes into clinical therapies, we performed a clinically oriented drug screening and molecular docking analysis. Based on the Drugs.com database, we selected two NSAIDs, Meloxicam and Naproxen. Molecular docking simulations revealed that both drugs exhibited favorable binding with FER, TUBA1B, and MICAL2 proteins, with Meloxicam showing consistently stronger binding affinities across all three targets. These findings provide a computational basis for understanding the molecular interactions underlying NSAID therapy in CTI. Nevertheless, direct binding assays and functional experiments are required to validate these predictions and establish causal relationships between target engagement and therapeutic effects.

To thoroughly elucidate the upstream regulatory mechanisms of *FER, TUBA1B*, and *MICAL2* in tendon injury, we systematically constructed a comprehensive regulatory network involving TFs and miRNAs. Notably, our analysis predicted that *STAT3, TFAP4, JUN*, and *MYC* are common TFs capable of simultaneously regulating all three key genes. *STAT3* is a central mediator in cytokine signaling, playing a crucial role in inflammatory responses, cell proliferation, and survival.[25] During tendon injury repair, *STAT3* could mediate anti-inflammatory effects, promoting environmental stability during the early stages of repair.[26] *JUN* and *MYC*, as immediate-early genes, are rapidly expressed upon cellular stimulation by growth factors, cytokines, or stress signals, regulating the transcription of downstream target genes. In tendon injury repair, their timely expression signifies the successful activation of pro-healing signaling pathways. By modulating key biological processes such as cell proliferation and differentiation, they synergistically promote the formation of repaired tissue with enhanced structural integrity and superior mechanical properties.[27] Furthermore, although *TFAP4* has been studied less in tendon repair, existing evidence indicates its role in accelerating cell cycle progression and promoting cell proliferation.[28] These findings suggested that the coordinated action of these four TFs might be crucial for maintaining the normal process of tendon repair, but further verification is still needed. In addition to upstream TFs regulation, our analysis also predicts that key genes are regulated by a series of miRNAs at the post-transcriptional level. The miRNAs targeting *FER* enriched in members closely related to matrix metabolism and myogenic regulation. For instance, the miR-29 family is recognized as a core anti-fibrotic factor, exerting protective effects in fibrosis across multiple organs including the liver, lungs, and heart by targeting ECM proteins such as collagen and fibronectin.[29,30] Furthermore, miR-206 is a prototypical myogenic miRNA and an indispensable regulator of skeletal muscle development, differentiation, and regeneration.[31] The miRNAs targeting *MICAL2* and *TUBA1B* are associated with inflammation, cell proliferation and migration. For example, miR-142-5p is a key regulator of macrophage polarization and immune response,[32] while the miR-27 family is widely involved in the regulation of cell migration.[33] MiR-223-3p also plays a role in inflammation and immune differentiation.[34] Overall, our study provides important clues for elucidating the molecular mechanisms underlying the development and progression of CTI.

When interpreting the validation results from the mouse dataset (GSE17591), potential interspecies differences should be considered. Although mice are widely used in tendon research, differences in tendon anatomy, mechanical loading, and genetic background between humans and mice may affect gene expression comparability. Nonetheless, the consistent expression trends observed across species support the potential involvement of these genes in conserved CTI-related pathways. Direct extrapolation to human pathophysiology requires caution, and future validation in human samples or human-derived models is needed.

Despite these findings, it's critical to acknowledge the study's limitations. Firstly, the GSE26051 dataset relied upon for analysis has a limited sample size and originates from a single source, which may restrict the generalizability of our findings. Future validation in larger, multicenter independent cohorts is warranted. Secondly, the validation of key genes is based solely on mice tendon samples (GSE17591) as no appropriate human CTI dataset was publicly available, and thus their relevance to human pathophysiology requires further confirmation. Thirdly, although our bioinformatic analysis included tendons from multiple sites, validation was limited to posterior tibial tendon samples. Further studies are needed to confirm these genes' expression patterns in CTI affecting other anatomical sites. Fourthly, although IHC confirmed increased protein expression of the three genes in degenerative tendons, definitive cellular localization data are currently unavailable due to limitations in sample preservation and antibody performance. Future immunofluorescence studies on fresh samples are needed to clarify their cellular distribution. Finally, it is important to emphasize that our findings establish associative, rather than functional or causative, links between the identified genes and CTI. In order to clarify the specific functions of these key genes in the biological behavior of tendon cells, future research should conduct functional experiments both in vitro and in vivo.

Conclusion

This study systematically revealed *FER, TUBA1B*, and *MICAL2* as core key genes in CTI through comprehensive bioinformatics analysis, with their differential expression further validated in human tendon tissue samples. Collectively, the diagnostic model, regulatory network, and candidate targeted drugs derived from these key genes provide a systematic theoretical framework to facilitate subsequent translational research and the development of novel CTI therapeutic strategies.

Data Sharing Statement

The datasets generated and/or analysed during the current study are available in the NCBI GEO database (https://www.ncbi.nlm.nih.gov/geo/).

Ethics Approval and Informed Consent

This study was performed according to the Declaration of Helsinki and approved by the Ethics Committee of Bethune International Peace Hospital (NO. 2023-KY-187). Written informed consent has been obtained from all participants for this study.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This study was funded by General Program Project of Bethune International Peace Hospital (2025FYMS21).

Disclosure

Zishen Cheng and Weina Ren are co-first authors for this study. The authors declare no competing interests in this work.

References

1. Li ZJ, Yang QQ, Zhou YL. Basic research on tendon repair: strategies, evaluation, and development. *Front Med Lausanne*. 2021;8:664909. doi:10.3389/fmed.2021.664909
2. Loiacono C, Palermi S, Massa B, et al. Tendinopathy: pathophysiology, therapeutic options, and role of nutraceutics. a narrative literature review. *Medicina*. 2019;55(8). doi:10.3390/medicina55080447
3. 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. *Scand J Med Sci Sports*. 2020;30(10):1810–1826. doi:10.1111/sms.13734
4. Tarantino D, Mottola R, Resta G, et al. Achilles tendinopathy pathogenesis and management: a narrative review. *Int J Environ Res Public Health*. 2023;20(17). doi:10.3390/ijerph20176681
5. Taskin Senol G, Ayanoglu T, Arikan E, Gökkus H, Ray A. Anatomical description of the structures that cause tibialis posterior tendon dysfunction. *Cureus*. 2025;17(3):e80196. doi:10.7759/cureus.80196
6. Wu F, Nerlich M, Docheva D. Tendon injuries: basic science and new repair proposals. *EFORT Open Rev*. 2017;2(7):332–342. doi:10.1302/2058-5241.2.160075
7. Lopes LR, Machado DE, Goes RA, et al. Implications of tendinopathy on the sports career: epidemiological and clinical profile of high-performance athletes. *Phys Ther Sport*. 2025;72:59–68. doi:10.1016/j.ptsp.2025.02.003
8. da LodiCMT, Salerno M, Merli G, Brama P, Jenner F, Filardo G. Tendinopathy: sex bias starts from the preclinical development of tendon treatments. A systematic review. *Biol Sex Differ*. 2022;13(1):44. doi:10.1186/s13293-022-00453-z
9. Liang W, Zhou C, Deng Y, et al. The current status of various preclinical therapeutic approaches for tendon repair. *Ann Med*. 2024;56(1):2337871. doi:10.1080/07853890.2024.2337871
10. Lu J, Li H, Zhang Z, Xu R, Wang J, Jin H. Platelet-rich plasma in the pathologic processes of tendinopathy: a review of basic science studies. *Front Bioeng Biotechnol*. 2023;11:1187974. doi:10.3389/fbioe.2023.1187974
11. Stowe EJ, Keller MR, Connizzo BK. Cellular senescence impairs tendon extracellular matrix remodeling in response to mechanical unloading. *Aging Cell*. 2024;23(11):e14278. doi:10.1111/acel.14278
12. Katoh K. Signal transduction mechanisms of focal adhesions: src and FAK-mediated cell response. *Front Biosci*. 2024;29(11):392. doi:10.31083/j.fbl2911392
13. Bachir AI, Horwitz AR, Nelson WJ, Bianchini JM. Actin-based adhesion modules mediate cell interactions with the extracellular matrix and neighboring cells. *Cold Spring Harb Perspect Biol*. 2017;9(7). doi:10.1101/cshperspect.a023234
14. Hernández Borrero LJ, El-Deiry WS. Tumor suppressor p53: biology, signaling pathways, and therapeutic targeting. *Biochim Biophys Acta Rev Cancer*. 2021;1876(1):188556. doi:10.1016/j.bbcan.2021.188556
15. Guo H, Cao H, Lu Q, Gu Z, Feng G. TNF-α induces premature senescence in tendon stem cells via the NF-κB and p53/p21/cyclin E/CDK2 signaling pathways. *Int J Mol Med*. 2025;56(3). doi:10.3892/ijmm.2025.5581
16. Ivanova IA, Vermeulen JF, Ercan C, et al. FER kinase promotes breast cancer metastasis by regulating α6- and β1-integrin-dependent cell adhesion and anoikis resistance. *Oncogene*. 2013;32(50):5582–5592. doi:10.1038/onc.2013.277
17. Haughton PD, Enserink L, Tavares S, et al. FER kinase governs invasive growth of head and neck squamous cell carcinoma through dynamic control of growth factor receptor activity. *Neoplasia*. 2025;70:101241. doi:10.1016/j.neo.2025.101241

18. Qi J, Zhou M, Yang N, et al. TUBA1B as a novel prognostic biomarker correlated with immunosuppressive tumor microenvironment and immunotherapy response. *Front Pharmacol*. 2025;16:1517887. doi:10.3389/fphar.2025.1517887
19. Kumar S, Maxwell IZ, Heisterkamp A, et al. Viscoelastic retraction of single living stress fibers and its impact on cell shape, cytoskeletal organization, and extracellular matrix mechanics. *Biophys J*. 2006;90(10):3762–3773. doi:10.1529/biophysj.105.071506
20. Kim J, Lee H, Roh YJ, et al. Structural and kinetic insights into flavin-containing monooxygenase and calponin-homology domains in human MICAL3. *IUCrJ*. 2020;7(Pt 1):90–99. doi:10.1107/s2052252519015409
21. Alhashimi I, Al-Taie A, Alam S, Ibrahem Adam R. Congenital absence of the navicular bone: a rare cause of adult-acquired flatfoot and posterior tibial tendon dysfunction. *Cureus*. 2025;17(5):e85151. doi:10.7759/cureus.85151
22. Chen YJ, Hsu RW, Liang SC. Degeneration of the accessory navicular synchondrosis presenting as rupture of the posterior tibial tendon. *J Bone Joint Surg Am*. 1997;79(12):1791–1798. doi:10.2106/00004623-199712000-00004
23. Torina A, Guggino G, La Manna MP, Sireci G. The Janus face of NKT cell function in autoimmunity and infectious diseases. *Int J Mol Sci*. 2018;19(2). doi:10.3390/ijms19020440
24. Kuno Y, Ishikawa H, Nagashima R, et al. NKT cell deficiency exacerbates adenine-induced renal fibrosis through enhanced Treg infiltration and TGF-β expression. *Am J Physiol Cell Physiol*. 2025;329(2):C471–c479. doi:10.1152/ajpcell.00373.2025
25. Cortés-Ballinas L, López-Pérez TV, Rocha-Zavaleta L. STAT3 and the STAT3-regulated inhibitor of apoptosis protein survivin as potential therapeutic targets in colorectal cancer (Review). *Biomed Rep*. 2024;21(6):175. doi:10.3892/br.2024.1863
26. Tarafder S, Chen E, Jun Y, et al. Tendon stem/progenitor cells regulate inflammation in tendon healing via JNK and STAT3 signaling. *FASEB J*. 2017;31(9):3991–3998. doi:10.1096/fj.201700071R
27. Wang R, Xu B, Xu HG. Up-regulation of TGF-β promotes tendon-to-bone healing after anterior cruciate ligament reconstruction using bone marrow-derived mesenchymal stem cells through the TGF-β/MAPK signaling pathway in a new zealand white rabbit model. *Cell Physiol Biochem*. 2017;41(1):213–226. doi:10.1159/000456046
28. Yuan YW, Yue ZQ, Zhou Q, et al. TFAP4 regulation of MCM5 activates the PI3K/AKT pathway to promote invasion and metastasis of gastric cancer. *Dig Dis Sci*. 2025;70(4):1411–1427. doi:10.1007/s10620-025-08897-0
29. Harmanci D, Erkan EP, Kocak A, Akdogan GG. Role of the microRNA-29 family in fibrotic skin diseases. *Biomed Rep*. 2017;6(6):599–604. doi:10.3892/br.2017.900
30. Leal MF, Caires Dos Santos L, Martins de Oliveira A, et al. Epigenetic regulation of metalloproteinases and their inhibitors in rotator cuff tears. *PLoS One*. 2017;12(9):e0184141. doi:10.1371/journal.pone.0184141
31. Ma G, Wang Y, Li Y, et al. MiR-206, a key modulator of skeletal muscle development and disease. *Int J Biol Sci*. 2015;11(3):345–352. doi:10.7150/ijbs.10921
32. Wei Y, Schober A. MicroRNA regulation of macrophages in human pathologies. *Cell Mol Life Sci*. 2016;73(18):3473–3495. doi:10.1007/s00018-016-2254-6
33. Song H, Qu HC, Liu S, Shen XX. MiR-27 inhibits cell migration and invasion by targeting CREB1 to influence MAPK/ERK signaling pathway in breast cancer. *J Biol Regul Homeost Agents*. 2021;35(Special Issue on Internal Medicine n.1). doi:10.23812/21-si1-11
34. Barbagallo D, Ponti D, Bassani B, et al. MiR-223-3p in cancer development and cancer drug resistance: same coin, different faces. *Int J Mol Sci*. 2024;25(15). doi:10.3390/ijms25158191