



OPEN ACCESS

EDITED BY

Dongwei Zhang,
Beijing University of Chinese Medicine,
China

REVIEWED BY

Dongfeng Zhao,
Shanghai University of Traditional Chinese
Medicine, China
Liu Xinhui,
Nanjing Jiangning Hospital, China
Yong Ma,
Nanjing University of Chinese Medicine,
China

*CORRESPONDENCE

Guicai Sun
✉ 13657000633@139.com

†These authors have contributed equally to
this work

RECEIVED 27 May 2023

ACCEPTED 02 August 2023

PUBLISHED 31 August 2023

CITATION

Ouyang Y, Chen S, Tu Y, Wan T, Fan H
and Sun G (2023) Exploring the potential
relationship between frozen shoulder and
Dupuytren's disease through bioinformatics
analysis and machine learning.
Front. Immunol. 14:1230027.
doi: 10.3389/fimmu.2023.1230027

COPYRIGHT

© 2023 Ouyang, Chen, Tu, Wan, Fan and
Sun. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Exploring the potential relationship between frozen shoulder and Dupuytren's disease through bioinformatics analysis and machine learning

Yulong Ouyang^{1,2†}, Shuilin Chen^{1,2†}, Yuanqing Tu^{1,2}, Ting Wan^{1,2},
Hao Fan^{1,2} and Guicai Sun^{1*}

¹Department of Orthopedics, The First Affiliated Hospital of Nanchang University, Nanchang, Jiangxi, China, ²The First Clinical Medical College, Nanchang University, Nanchang, Jiangxi, China

Background: Frozen shoulder (FS) and Dupuytren's disease (DD) are two closely related diseases, but the mechanism of their interaction is unknown. Our study sought to elucidate the molecular mechanism of these two diseases through shared gene and protein interactions.

Methods: GSE75152 and GSE140731 data were downloaded from the Gene Expression Omnibus (GEO) database, and shared genes between FS and DD were selected by using R packages. Then, we used Cytoscape software and the STRING database to produce a protein-protein interaction (PPI) network. Important interaction networks and hub genes were selected through MCODE and cytoHubba algorithms. To explore the potential mechanisms of the development of the two diseases, the hub genes were further enriched by GO and KEGG analyses. We predicted the transcription factors (TFs) of hub genes with Transcriptional Regulatory Relationships Unraveled by Sentence-based Text mining (TRRUST). Moreover, we identified candidate genes for FS with DD with cytoHubba and machine learning algorithms. Finally, we analyzed the role of immunocyte infiltration in FS and constructed the relationship between candidate genes and immunocytes in FS.

Results: We identified a total of 321 shared genes. The results of GO and KEGG enrichment of shared genes showed that extracellular matrix and collagen fibril tissue play a certain role in the occurrence and development of disease. According to the importance of genes, we constructed the key PPI network of shared genes and the top 15 hub genes for FS with DD. Then, we predicted that five TFs are related to the hub genes and are highly expressed in the FS group. Machine learning results show that the candidate genes POSTN and COL11A1 may be key for FS with DD. Finally, immune cell infiltration revealed the disorder of immunocytes in FS patients, and expression of candidate genes can affect immunocyte infiltration.

Conclusion: We identified a PPI network, 15 hub genes, and two immune-related candidate genes (POSTN and COL11A1) using bioinformatics analysis and machine learning algorithms. These genes have the potential to serve as diagnostic genes for FS in DD patients. Furthermore, our study reveals disorder of immunocytes in FS.

KEYWORDS

frozen shoulder, Dupuytren's disease, diagnosis, machine learning, immunocyte infiltration

Introduction

Frozen shoulder (FS) is a common joint disease that causes pain and stiffness (1). FS is the third most common cause of musculoskeletal disability in the United States and affects patients between the ages of 40 and 60 years, with a prevalence rate of 2% to 5% (2, 3). FS is typically classified as either primary or secondary FS (4). Primary FS usually occurs without a clear cause, whereas secondary FS is caused by trauma or immobilization (4). Although the disease is self-limited and shows some improvement after conservative treatment or surgical treatment, most patients who are followed up long term develop permanent disability (5, 6). However, the pathogenesis of FS is still unclear.

Studies have shown that FS may be associated with hyperthyroidism, hypothyroidism, diabetes, and Dupuytren's disease (DD) (7). Among these diseases, DD and FS have very similar histopathological and immunocytochemical characteristics (8, 9). DD is one of the most widespread inherited connective tissue disorders and appears to be common in northern Europeans (10). The prevalence of DD ranges from 3% to 42% worldwide (11). DD is a common benign fibroproliferative condition (12). A fibrotic nodule over the palmar fascia is the most common symptom of the disorder. These bands have the ability to contract, leading to flexion contracture of the facet joints of the hand, known as Dupuytren's contracture (DC) (13). However, the exact pathophysiology of DD is not known.

The unknown causes of FS and DD make treating patients difficult, and their high prevalence brings a great economic burden to society (14). Therefore, studying the pathogenesis is helpful and necessary for the development of disease treatments. Increasing evidence shows not only that there are some similar pathological features between DD and FS but also that there is a strong clinical correlation between them (12, 15). There may be a breakthrough in understanding the pathogenesis of these two diseases due to their similarity. However, little is known about the common characteristics of FS and DD based on gene regulatory mechanisms.

Recently, as sequencing technology and bioinformatics have developed, it has become possible to determine how diseases interact genetically (16, 17). Screening out hub genes and analyzing the correlation between them is very important for studying diseases. In this study, we aimed to uncover shared genes and clusters of coexpression between FS and DD, providing a relevant foundation for future research on the two diseases. We

constructed a protein-protein interaction (PPI) network for the shared genes and identified hub and candidate genes by using cytoHubba and machine learning algorithms. We explored the relationship between transcription factors (TFs) and these diseases on the basis of hub genes. Moreover, we analyzed the role of immunocytes in FS patients and constructed the relationship between candidate genes and immunocytes in FS.

Methods

Dataset download and processing

We used the term "frozen shoulder" or "Dupuytren's disease" to search for gene expression profiles in the Gene Expression Omnibus (GEO) (<http://www.ncbi.nlm.nih.gov/geo/>) database. GSE75152 and GSE140731 were obtained. For GSE75152, total mRNA was extracted from connective tissue in the hand and detected with the Illumina HumanHT-12 V3.0 expression beadchip (18). For GSE140731, total mRNA was extracted from tissue from the anterior capsule and rotator interval and detected using RNA sequencing (19). These data are all from humans. After downloading the two datasets, we preprocessed the data. Then, we performed log₂ transformation for gene expression profiling and matched the probes to their gene symbols using the annotation document for the appropriate platforms. Finally, a gene matrix was obtained for subsequent analyses.

Identification of shared genes

We identified differentially expressed genes (DEGs) in FS and DD ($|\log FC| > 0.585$, adjusted p -value < 0.05). There was overlap between genes in the FS and DD modules with positive correlation coefficients, as assessed by the "VennDiagram" package (20). Functional enrichment was performed to analyze common DEGs and total DEGs in DD and FS via R packages (enrichplot, RColorBrewer, ComplexHeatmap, etc.).

Physical protein interaction network

DD and FS shared genes were used to draw a physical interaction network with the STRING database (version 11.5).

The interaction network was extracted, and the extracted network data were visualized by using Cytoscape software (version 3.9.1). To analyze the extracted network data using Cytoscape software, the MCODE algorithm and the cytoHubba algorithm were employed in the analysis of the PPI network and cluster analysis. Gene clusters with the top 15 hub genes were obtained. The interacting genes and functions of these genes were predicted, and related PPIs were generated in the GeneMANIA database (<https://genemania.org/>). Further enrichment of hub clusters was achieved through analysis of GO functional terms and KEGG pathway enrichment analysis.

Construction of the transcriptional regulatory network

To identify substantial changes at the transcriptional level and to gain an in-depth understanding of the regulatory role of DEGs, we used the hub gene to introduce Transcriptional Regulatory Relationships Unraveled by Sentence-based Text mining (TRRUST, <https://www.grnpedia.org/trrust/>). Next, enrichment of hub genes was assessed to obtain corresponding TFs. Then, Cytoscape software was used to produce a TF regulatory network. Finally, differential expression of TFs in FS patients was analyzed to identify TFs that play a critical role in the disease.

Machine learning

We used least absolute shrinkage and selection operator (LASSO) regression and support vector machine (SVM) learning to identify potential candidate genes related to the diagnosis of FS and DD. LASSO regression is a machine learning technique that combines variable selection and regularization, which can improve prediction accuracy (21). SVM learning is very powerful at recognizing subtle patterns in complex datasets (22). It aims to create a decision boundary between two classes such that labels from one or more feature vectors can be predicted (23). We used the “glmnet”, “e1071”, “kernlab”, and “caret” software packages of R software to carry out LASSO regression and SVM learning. The intersection of these two results can be used as candidate genes for diagnosis. To determine the importance of candidate genes in diagnosis, we used the “rms” R package to construct a nomogram. We further evaluated the prognostic value of the candidate genes by ROC analysis. The area under the ROC curve (AUC) and 95% confidence interval (CI) were obtained. An AUC value > 0.7 is considered to indicate good diagnostic efficacy.

Immune infiltration analysis

We used the CIBERSORT tool to evaluate immunocyte infiltration based on gene expression profiles. A bar plot was used to visualize the proportion of immunocytes, while a violin plot was used to compare the proportions of these cells between the FS and control groups. We established the relationship between candidate

genes and immunocytes by Spearman analysis and used a Lollipop plot to visualize their correlation.

Statistical analysis

Data were analyzed with R 4.3.0 (<https://www.rproject.org/>), Cytoscape software (version 3.9.1), Perl 5.32.1 (<https://www.perl.org/>), and R Bioconductor packages. All statistical *p*-values were two-sided, with *p* < 0.05 considered statistically significant.

Results

Identification of DEGs in DD and FS

Two microarray datasets were first analyzed separately. The volcano plot and heat map in **Figures 1A–D** show the genetic differences between the two diseases. Compared with the control group, there were 2,762 differentially expressed genes in the FS group, including 1,245 downregulated and 1,517 upregulated genes. There were 1,555 differentially expressed genes in the DD group compared with the control group, including 764 downregulated and 791 upregulated genes.

Common DEGs in DD and FS

There were 321 common differential genes among GSE75152 and GSE140731 DEGs, namely, 255 upregulated genes and 66 downregulated genes (**Figures 2A, B**). We analyzed GO information and KEGG pathways for these common DEGs to explore their underlying biological information. The results of GO analyses showed that in biological process (BP), genes involved in extracellular matrix structure, extracellular structure organization, encapsulating structure organization, ossification, and collagen fibril organization, among others, were especially enriched. In cellular component (CC), collagen-containing extracellular matrix, collagen trimer, and focal adhesion, among others, were represented. Molecular function (MF) items that were enriched included extracellular matrix structural constituents such as collagen binding and actin binding (**Figures 2C–E**). KEGG enrichment results showed the DEGs to be mainly enriched in protein digestion and absorption, ECM–receptor interaction, focal adhesion, and PI3K–Akt signaling pathways (**Figures 2F, G**).

PPI network construction of common DEGs

We used common DEGs to construct a PPI network in STRING to understand potential connections between proteins, with a minimum required interaction score of 0.7 and a PPI enrichment *p*-value < 1.0e-16. This score means that the connection has a high

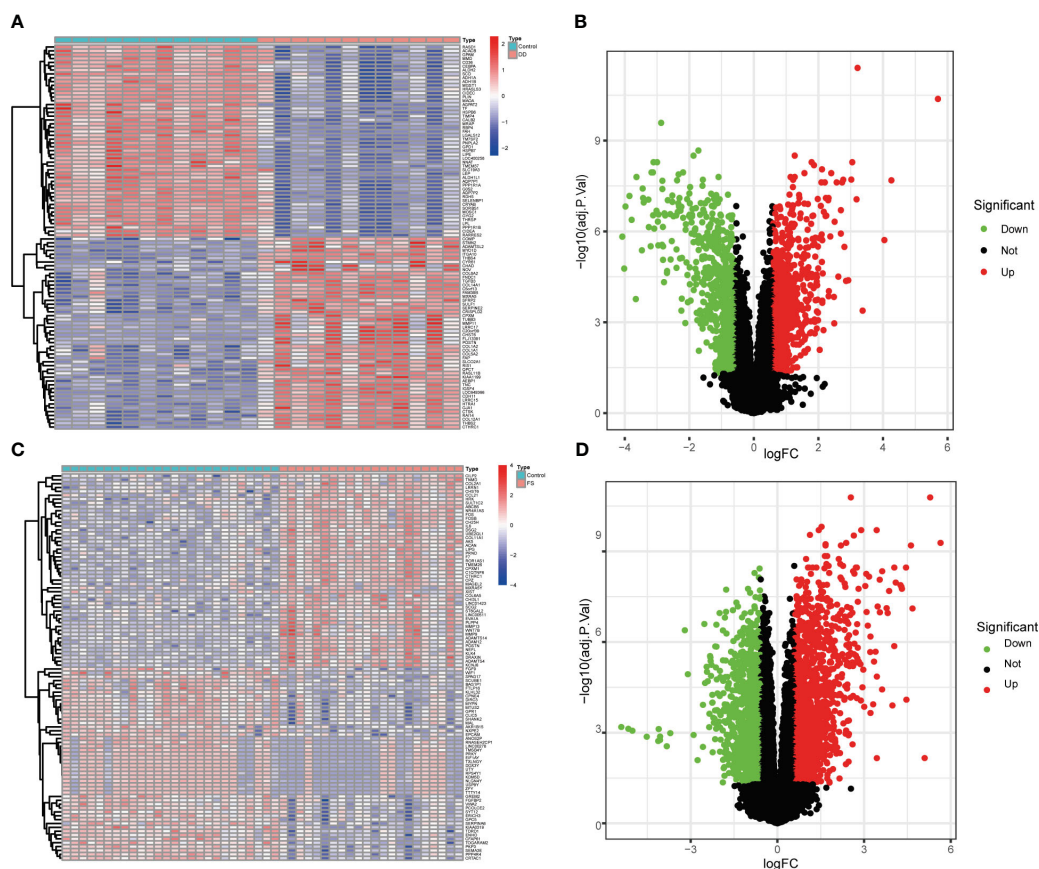


FIGURE 1

Volcano plot and heatmap show genetic differences between the two groups of diseases. (A) Heatmap of gene expression levels between the DD and control groups. (B) Volcano plot of gene expression levels between the DD and control groups. (C) Heatmap of gene expression levels between the FS and control groups. (D) Volcano plot of gene expression level differences between the DD and control groups.

degree of confidence (Figure 3A). To understand expression of interacting proteins in FS with DD, we scored protein–protein interactions and combined the PPI network with a score greater than 0.7 with whether the common DEGs were upregulated or downregulated (Figure 3B). We used the “MCODE” algorithm of Cytoscape to analyze and detect key clustering modules. The module consists of 15 nodes and 67 edges, and the cluster score (density multiplied by the number of members) was 9.571 (Figure 3C).

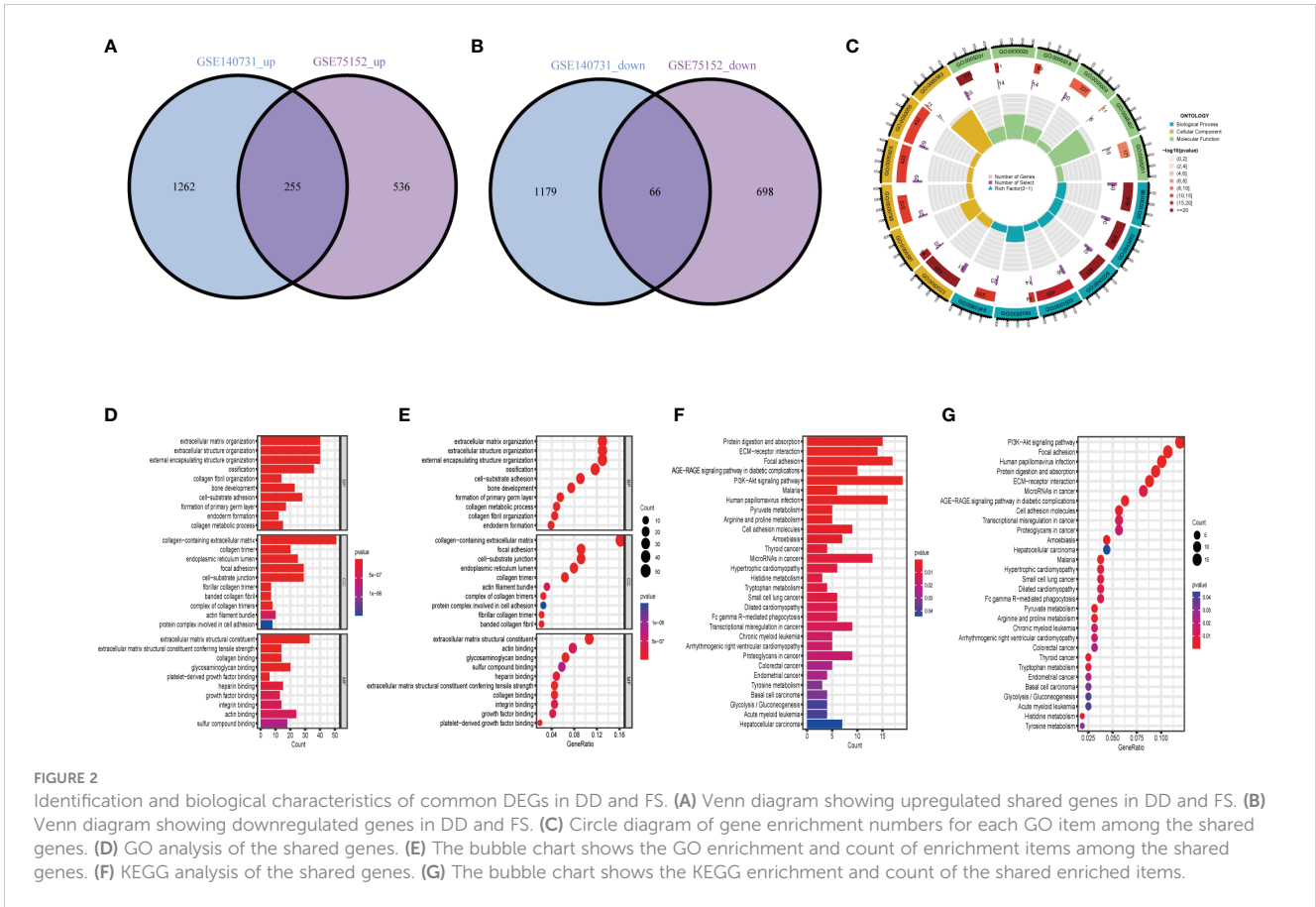
Identification of hub genes and PPI network construction

To understand interaction between hub genes, we used the “cytoHubba” algorithm to identify the top 15 hub genes. Then, the interaction of these genes was visualized, whereby a darker color indicates a more important gene (Figure 4A). GeneMANIA was used with the PPI to evaluate 15 hub genes and 20 interacting genes to predict relationships between coexpression, shared protein domain, colocation, and pathway aspects (Figure 4B). The outer circle is the predictive genes, and the inner circle is the hub genes.

Network analysis showed that these genes are associated with extracellular matrix structural constituents, extracellular matrix organization, collagen trimer, collagen trimer complex, and fibrillar collagen trimer.

GO and KEGG pathway analyses of hub genes

To further understand the function of hub genes in FS with DD, we enriched hub genes. According to GO analysis, collagen fibril organization, extracellular matrix organization, and extracellular structure organization, among others, were enriched in BP. Collagen-containing extracellular matrix, fibrillar collagen trimer, and banded collagen fibril, among others, were enriched in CC. Extracellular matrix structural constituents conferring tensile strength, extracellular matrix structural constituents, and platelet-derived growth factor binding, among others, were enriched in MF (Figures 5A–C). KEGG analysis showed that the hub genes were enriched in protein digestion and absorption, ECM–receptor interaction, and focal adhesion, among others (Figures 5D–F).

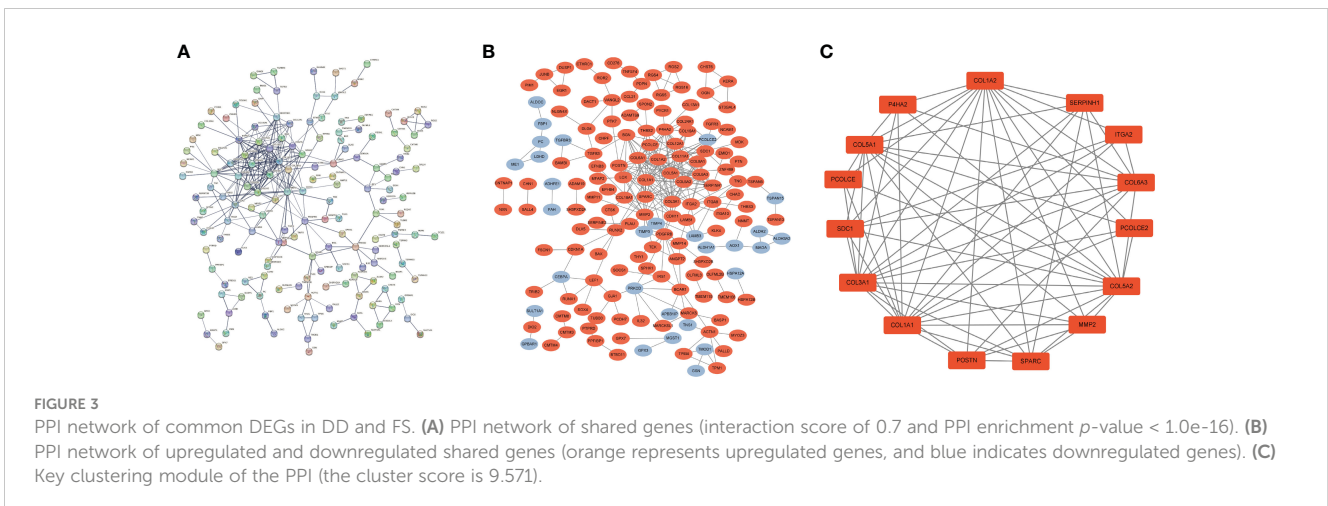


Association between hub genes and TFs

TFs are involved in gene regulation. To explore the role of TFs, we used TRRUST to predict key TFs that affect these two diseases through hub genes. Analysis of interactions between TFs and common DEGs revealed that 11 TFs coordinated 8 common DEGs, indicating a high level of cooperation (Figure 6A). The top TFs were ranked according to their *p*-values and included RELA,

NFKB1, CEBPZ, SP3, TWIST2, CIITA, STAT6, MYB, TFAP2A, YY1, and SP1 (Table S1). These findings reveal a significant relationship between common DEGs and TFs.

By analyzing expression of these TFs in GSE140731, we detected significant differences in SP3, YY1, SP1, RELA, and CEBPZ between FS patients and healthy patients (Figures 6B–F). This indicates that TFs influence the development of these diseases and their occurrence.



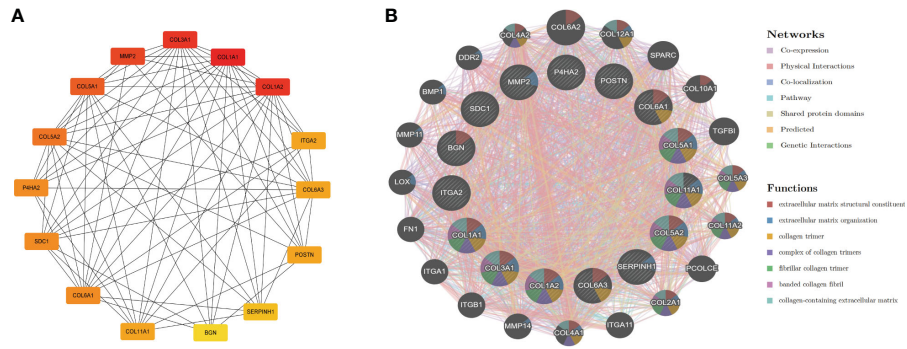


FIGURE 4 PPI network of hub genes. **(A)** PPI network of 15 hub genes (the darker the color is, the more important the gene is). **(B)** GeneMANIA predicted relationships between coexpression, shared protein domain, colocation, and pathway.

Identification of candidate genes via machine learning

LASSO regression and SVM algorithms were utilized to identify potential candidate genes of FS combined with DD. LASSO regression analysis and SVM algorithms identified 23 and 37 genes closely related to FS with DD, respectively (Figures 7A, B). The intersection of the 37 genes from SVM and 23 genes from

LASSO identified six genes (COL11A1, POSTN, LBH, CH25H, SMOC2, and PALLD). COL11A1 and POSTN were also the top 15 hub genes (Figure 7C). The AUC and 95% CI of these genes were calculated by constructing an ROC curve to evaluate the diagnostic effect, as shown in Figures 7D, E. The results were as follows: POSTN (AUC: 0.979, CI: 0.950–0.997) and COL11A1 (AUC: 0.994, CI: 0.981–1.000). We constructed a nomogram with 2 candidate diagnostic genes (Figure 7F).

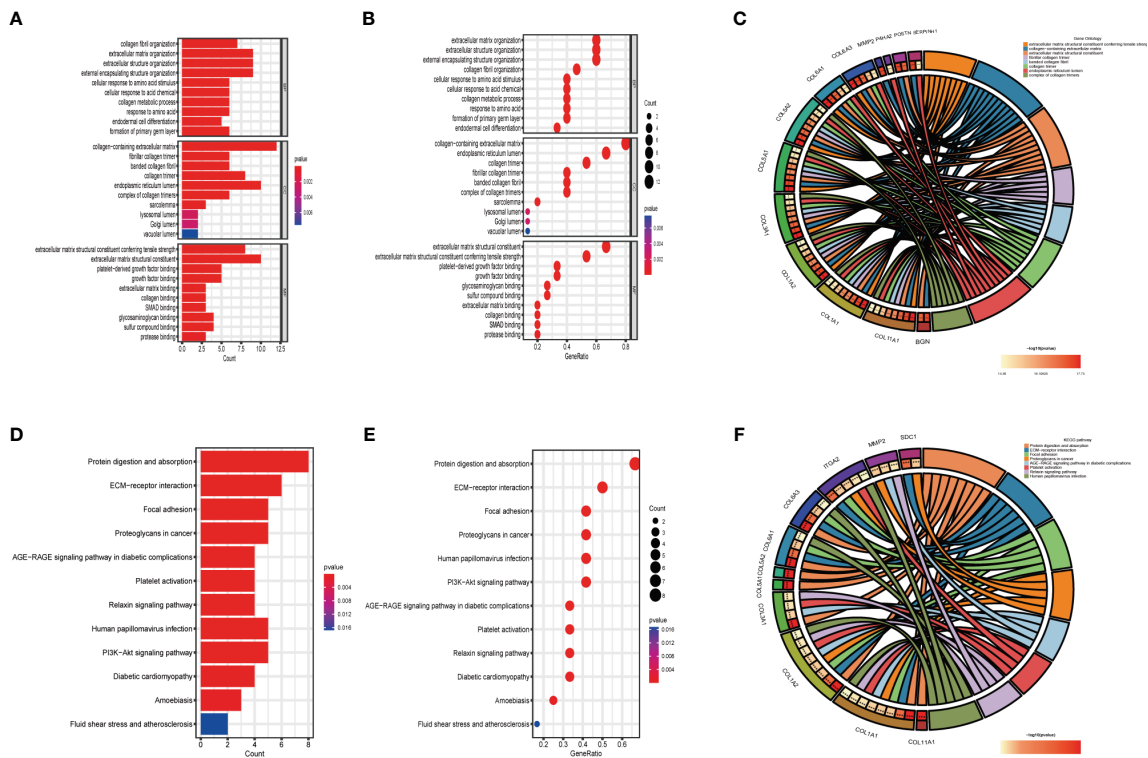


FIGURE 5 GO and KEGG analyses of hub genes. **(A)** GO analysis of hub genes. **(B)** The bubble chart shows the GO enrichment and count of enrichment items in hub genes. **(C)** The circle diagram shows the relationship between key genes and GO enrichment items. **(D)** KEGG analysis of hub genes. **(E)** The bubble chart shows the KEGG enrichment and count of enrichment items in hub genes. **(F)** The circle diagram shows the relationship between hub genes and KEGG enrichment items.

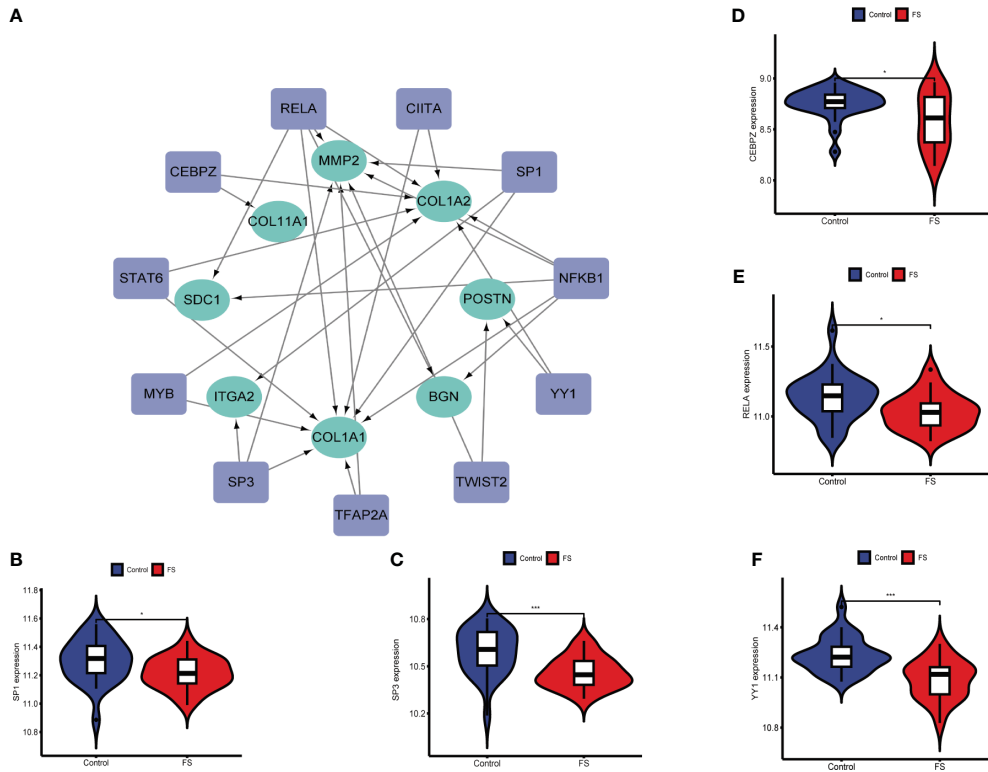


FIGURE 6 Association between hub genes and TFs. **(A)** Prediction of TF genes and their interaction network with hub genes. **(B)** Differences in expression of SP1 in the FS and control groups. **(C)** Differences in expression of SP3 in the FS and control groups. **(D)** Differences in expression of CEBPZ in the FS and control groups. **(E)** Differences in expression of RELA in the FS and control groups. **(F)** Differential expression of YY1 in the FS and control groups (* $p < 0.05$, *** $p < 0.001$).

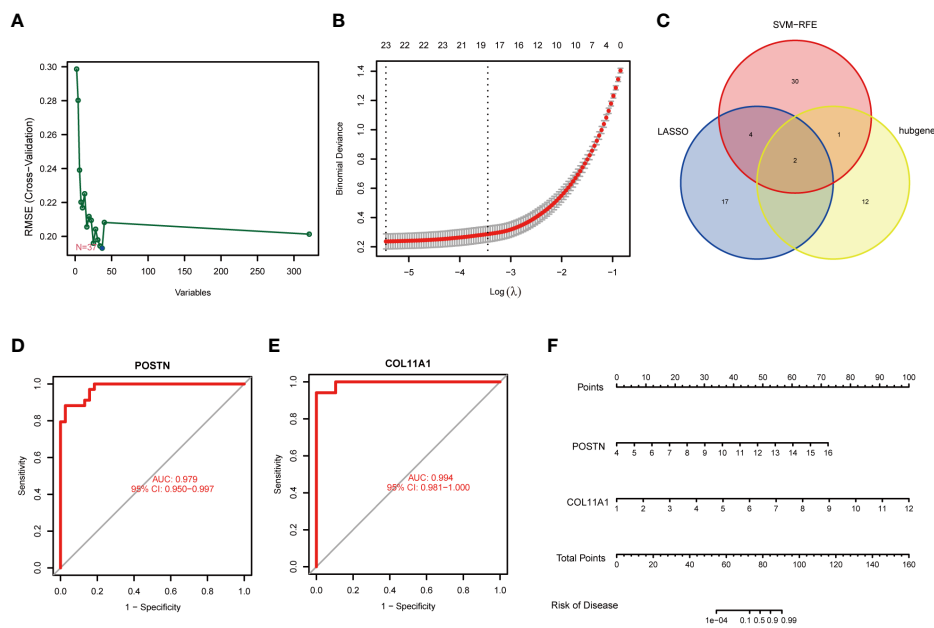
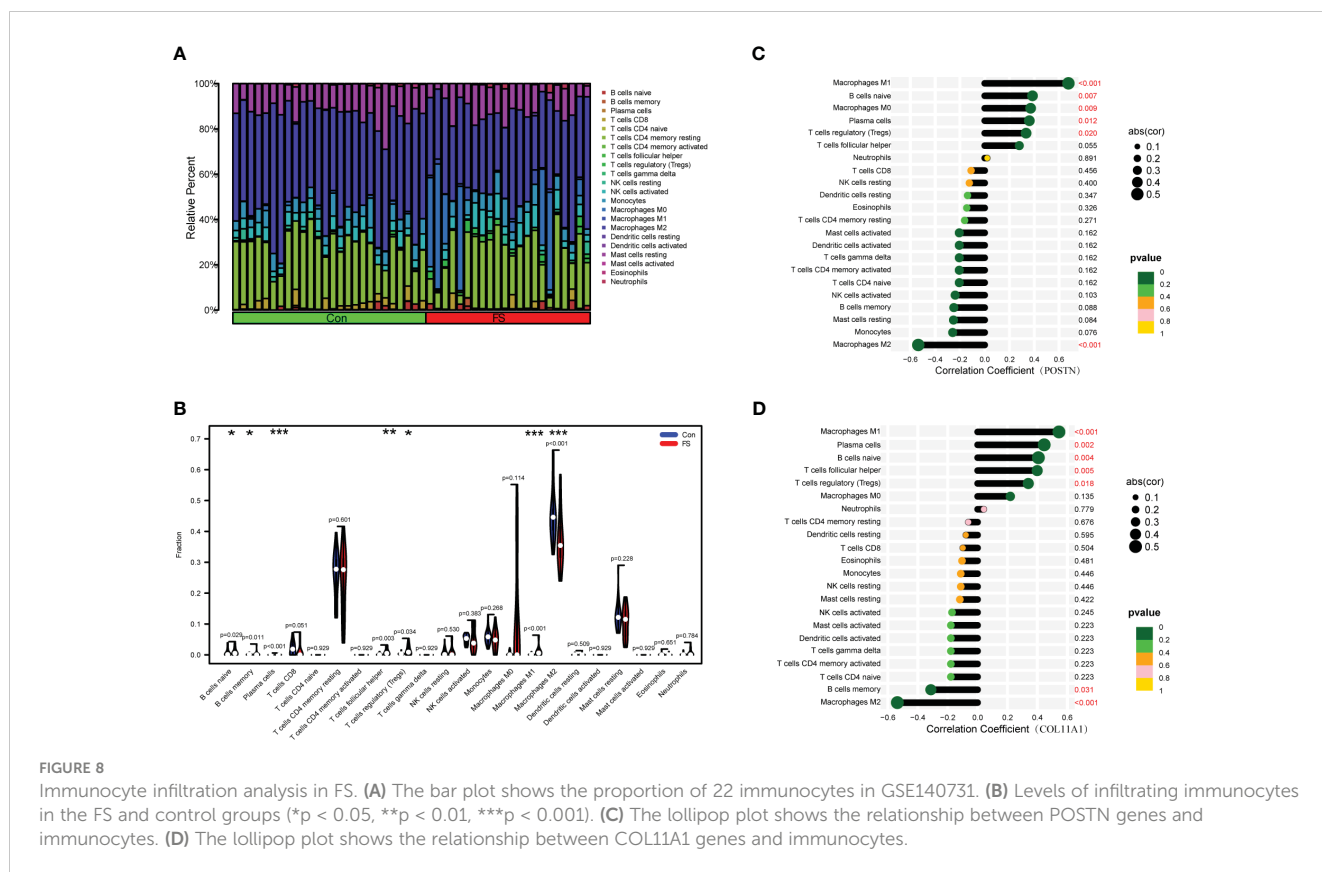


FIGURE 7 Identification of candidate genes. **(A)** Screening of key genes by LASSO regression. **(B)** Screening of key genes by SVM learning. **(C)** Candidate genes identified in the LASSO model, SVM learning, and cytoHubba algorithm. **(D)** ROC curve of the POSTN gene. **(E)** ROC curve of the COL11A1 gene. **(F)** Nomogram for diagnosis FS with DD.



Immune infiltration analysis

Immunocytes may be involved in the pathogenesis of DD. However, as the relationship between immunocytes and the pathogenesis of FS is rarely reported, we conducted immune infiltration analysis to explore the effect of immunity on FS. The proportion of immunocytes in the FS group and control group is shown in Figure 8A. Compared with the control group, there were significant differences in levels of naive B cells, memory B cells, plasma cells, follicular helper T cells, regulatory T cells (Tregs), M1 macrophages, and M2 macrophages in the FS group (Figure 8B). Then, we established the relationship between candidate genes and immunocytes and observed the effects of different candidate genes on immunocytes (Figures 8C, D, S1). The results showed that expression of the POSTN gene was closely related to M1 macrophages, M2 macrophages, naive B cells, M0 macrophages, plasma cells, and regulatory T cells (Tregs) (Figure 8C). Expression of the COL11A1 gene was closely related to M1 macrophages, M2 macrophages, plasma cells, naive B cells, follicular helper T cells, regulatory T cells (Tregs), and memory B cells (Figure 8D).

Discussion

Before 1872, FS was thought to be an inflammation of the shoulder and was labeled periartthritis by Duplay. Then, in 1934, Codman first proposed the term frozen shoulder (24). The change in terms seems to indicate that the pathological changes associated

with FS have gradually shifted from inflammation to fibrosis. As the most common hereditary disease of connective tissue, DD is mainly characterized by pathological fibrosis of connective tissue (25). In 1936, Schaers (26) first reported the close relationship between FS and DD. Subsequently, increasing clinical evidence has proven the correlation between them (8, 27, 28). Furthermore, an increasing number of basic studies are also attempting to reveal the pathogenesis and potential relationship between FS and DD. Lundberg (29) reported histological similarities linking the two pathologies as early as 1969. Bunker (9) also found that there are histological similarities between FS tissue and DD tissue, confirming that both types of tissue constitute a combination of synovial inflammation and capsular fibrosis. These studies show that DD and FS share a common mechanism, but the molecular mechanism of their interaction and pathogenesis is still unclear.

Few studies have explored the susceptibility factors of DD in FS at the genetic level. Exploring the interaction of two related diseases at the genetic level may provide some inspiration for research on the two diseases. In this study, bioinformatics analysis through a public database was used to explore the mechanism of FS and DD at the gene level. First, we analyzed genetic differences in the high-throughput data of patients with FS and DD and found significant differences between patients with the two diseases and the healthy group. Subsequently, shared genes with common upregulation and downregulation were found in the two diseases, and GO and KEGG enrichment analyses were performed. These analyses showed a certain association between the two diseases at the genetic level. Third, we used the shared genes to build a PPI

model and selected key clustering modules from it. The results showed an interaction at the protein level. Next, we selected the top 15 genes according to the algorithm and constructed an interactive network; similar to the key PPI model, interaction of these genes plays an important role in the pathogenesis of the two diseases. Some of the 15 hub genes have been reported to be closely related to the two diseases, such as COL1A1 and MMP2 (30–32). However, many genes have not been reported in the pathogenesis of DD or FS. Next, we examined the GO and KEGG pathways that were enriched among the hub genes, and the results showed extracellular matrix deposition, cell adhesion, and fibrosis to be closely related to these two diseases. This view is supported by previous studies (33, 34). We found some pathways that may be related to the two diseases, such as the PI3K-AKT pathway, Wnt pathway, and TGF-β pathway (Tables S2, S3). Rui Yang showed that IL-6 is upregulated in FS synovium and can promote the formation of FS fibrosis through the PI3K-AKT pathway (35). Chen reported that TGF-β1 can induce shoulder fibrosis in a rat model and that a PPAR-γ agonist can be used as a treatment (36). Similarly, some studies have suggested that the TGF-β pathway and Wnt pathway are key pathways in DD (37, 38). These studies prove the reliability of our enrichment. However, compared with other diseases, studies on the mechanism of these two diseases are scarce. Interestingly, we identified the AGE-RAGE signaling pathway, which is closely related to diabetes, in our enrichment results. Many studies have found that diabetes is a high-risk factor for FS and DD (39–42). However, the mechanism is unclear, though the AGE-RAGE signaling pathway may be a key pathway linking these diseases. Moreover, according to hub genes, we predicted TFs related to the

diseases and found that some differ between the diseases. TFs play an important role in the regulation of various diseases. Our results prove that these TFs may provide an important basis for the study of FS with DD.

We identified the candidate genes POSTN and COL11A1 by machine learning. They are involved in skeletal system development, ossification, or osteoblast differentiation (43). They may play a role in FS with DD disease progression. Linda's studies have proven that POSTN is upregulated in DD and that it can promote the transformation of fibroblasts into myofibroblasts, thus promoting progression of the disease (44). This study proves the accuracy of our bioinformatics analysis. However, studies on the effect of POSTN on FS have not been reported.

Since DD can induce FS, we wanted to determine the effect of candidate genes on FS immune cells. Then, we found that immunocytes are involved in the pathogenesis of FS. Although there are few studies on the mechanism of the effect of immunocytes on FS, the role of inflammation in the two diseases has been proven; thus, immunocytes are very important for progression of the disease. Exploring the mechanism of immune cells in FS may be a breakthrough in the study of disease mechanisms. The important role of imbalance between M1 and M2 macrophages in fibrotic diseases has been demonstrated by many studies (45, 46). Our study also found that imbalance between M1 and M2 macrophages is involved in the occurrence and development of FS. Finally, to further explore the important role of immune cell infiltration and candidate genes in FS, we assessed the correlation between them. Our results show that the candidate genes can affect disease progression by regulating immunocytes.

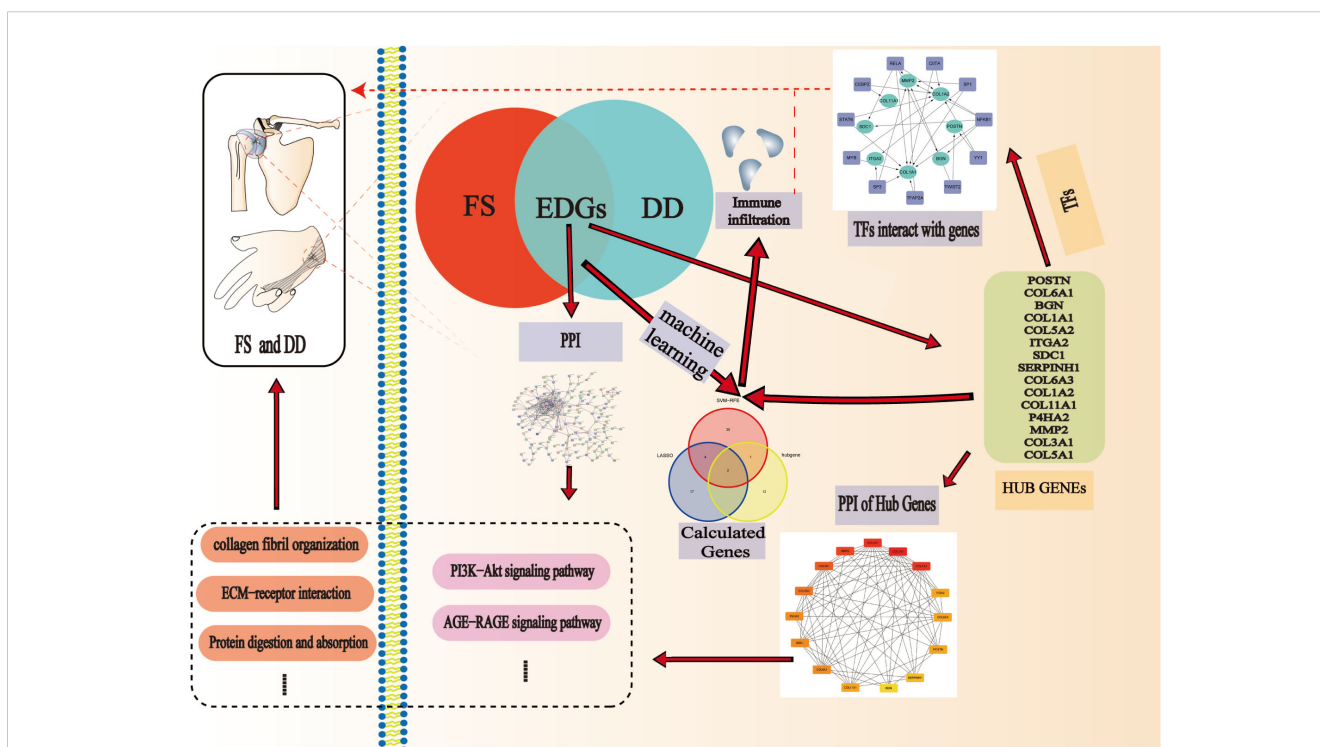


FIGURE 9 Pathogenesis of the two diseases at the genetic level.

Limitations of the study

Through the interactive study of FS and DD, we explored the pathogenesis of the two diseases at the genetic level (Figure 9). The results suggest new implications for the study of the pathogenesis of the two diseases. However, there are still some shortcomings in our study. First, the key genes and pathways found were not verified by experiments. However, some genes and related pathways have been confirmed in previous studies, which proves the reliability of this study. The focus of our subsequent studies will be to demonstrate the effects of these pathways and genes on FS. Due to the lack of high-throughput FS and DD datasets, the results cannot be verified by other datasets. More high-throughput sequencing data are needed to verify our results. Finally, the larger the sample size is, the more accurate the results of bioinformatics analysis is. It is hoped that there will be larger sample size datasets in the future.

Conclusion

We identified the PPI network, 15 hub genes, and two immune-related candidate genes (POSTN and COL11A1) using bioinformatics analysis and machine learning algorithms. These genes have the potential to serve as diagnostic candidate genes for FS in DD patients. Furthermore, our study reveals disorder of immunocytes in FS. We clarified the potential mechanism of the relationship between FS and DD at the genetic level, providing a new idea for future research on DD and FS.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: <https://www.ncbi.nlm.nih.gov/geo/> under the accession numbers GSE75152 and GSE140731. The GSE75152 and GSE140731 datasets were obtained from the GEO database.

Author contributions

GS and YO designed this study. SC and TW conducted the literature searches and management. YO, YT, and HF were responsible for the data management and statistical analysis. SC and YO interpreted the findings. YO wrote the manuscript. The

manuscript was revised by all authors, and the final version was approved by all.

Funding

This work was supported by the National Natural Science Foundation of China (No. 81960881) and the Jiangxi Province Graduated Innovation Foundation (YC2021-B050). The funders were not involved in the study design, data collection, statistical analysis, or writing of the manuscript.

Conflict of interest

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

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2023.1230027/full#supplementary-material>

SUPPLEMENTARY FIGURE 1
(A–M) Correlation between candidate genes and immunocytes.

SUPPLEMENTARY TABLE 1
Key TFs predicted according to hub genes.

SUPPLEMENTARY TABLE 2
KEGG analyses of hub genes.

SUPPLEMENTARY TABLE 3
GO analyses of hub genes.

References

- Brue S, Valentin A, Forssblad M, Werner S, Mikkelsen C, Cerulli G. Idiopathic adhesive capsulitis of the shoulder: a review. *Knee surgery sports traumatology arthroscopy* (2007) 15:1048–54. doi: 10.1007/s00167-007-0291-2
- Cher JZB, Akbar M, Kitson S, Crowe LAN, Garcia-Melchor E, Hannah SC, et al. Alarmins in frozen shoulder: A molecular association between inflammation and pain. *Am J sports Med* (2018) 46:671–8. doi: 10.1177/0363546517741127

3. Paul A, Rajkumar JS, Peter S, Lambert L. Effectiveness of sustained stretching of the inferior capsule in the management of a frozen shoulder. *Clin orthopaedics related Res* (2014) 472:2262–8. doi: 10.1007/s11999-014-3581-2
4. Zuckerman JD, Rokito A. Frozen shoulder: a consensus definition. *J shoulder elbow Surg* (2011) 20:322–5. doi: 10.1016/j.jse.2010.07.008
5. Chan HBY, Pua PY, How CH. Physical therapy in the management of frozen shoulder. *Singapore Med J* (2017) 58:685–9. doi: 10.11622/smedj.2017107
6. Hand C, Clipsham K, Rees JL, Carr AJ. Long-term outcome of frozen shoulder. *J shoulder elbow Surg* (2008) 17:231–6. doi: 10.1016/j.jse.2007.05.009
7. Xu Q, Gai PY, Lv HL, Li GR, Liu XY. Association of MMP3 genotype with susceptibility to frozen shoulder: a case-control study in a Chinese Han population. *Genet Mol Res GMR* (2016) 15. doi: 10.4238/gmr.15017228
8. Smith SP, Devaraj VS, Bunker TD. The association between frozen shoulder and Dupuytren's disease. *J shoulder elbow Surg* (2001) 10:149–51. doi: 10.1067/mse.2001.112883
9. Bunker TD, Anthony PP. The pathology of frozen shoulder. *A Dupuytren-like disease J Bone Joint surgery Br volume* (1995) 77:677–83. doi: 10.1302/0301-620X.77B5.7559688
10. Bayat A, McGrouther DA. Management of Dupuytren's disease—clear advice for an elusive condition. *Ann R Coll Surgeons Engl* (2006) 88:3–8. doi: 10.1308/003588406X83104
11. Ross DC. Epidemiology of dupuytren's disease. *Handb Clinics* (1999) 15:53–62. doi: 10.1016/S0749-0712(21)00439-X
12. Van Nuffel M, Poelmans A, Kellens S, Vandenneucker H, Degreef I. The prevalence of frozen shoulder in patients with limited Dupuytren disease. *Acta orthopaedica Belgica* (2022) 88:387–91. doi: 10.52628/88.2.9933
13. Feldman G, Rozen N, Rubin G. Dupuytren's contracture: Current treatment methods. *Israel Med Assoc J IMAJ* (2017) 19:648–50.
14. Bouaicha S, Wieser K, Kriechling P, Scholz-Odermatt SM. A large-scale assessment of the healthcare burden of adhesive capsulitis of the shoulder joint. *Swiss Med weekly* (2020) 150:w20188. doi: 10.4414/SMW.2020.20188
15. Green HD, Jones A, Evans JP, Wood AR, Beaumont RN, Tyrrell J, et al. A genome-wide association study identifies 5 loci associated with frozen shoulder and implicates diabetes as a causal risk factor. *PLoS Genet* (2021) 17:e1009577. doi: 10.1371/journal.pgen.1009577
16. Wang S, Tian J, Wang J, Liu S, Ke L, Shang C, et al. Identification of the biomarkers and pathological process of heterotopic ossification: Weighted gene co-expression network analysis. *Front Endocrinol* (2020) 11:581768. doi: 10.3389/fendo.2020.581768
17. Hu Y, Zeng N, Ge Y, Wang D, Qin X, Zhang W, et al. Identification of the shared gene signatures and biological mechanism in type 2 diabetes and pancreatic cancer. *Front Endocrinol* (2022) 13:847760. doi: 10.3389/fendo.2022.847760
18. Becker K, Siebert S, Toliat MR, Du J, Casper R, Dolmans GH, et al. Meta-analysis of genome-wide association studies and network analysis-based integration with gene expression data identify new suggestive loci and unravel a wnt-centric network associated with dupuytren's disease. *PLoS One* (2016) 11:e0158101. doi: 10.1371/journal.pone.0158101
19. Kamal N, McGee SL, Eng K, Brown G, Beattie S, Collier F, et al. Transcriptomic analysis of adhesive capsulitis of the shoulder. *J orthopaedic Res* (2020) 38:2280–9. doi: 10.1002/jor.24686
20. Jia A, Xu L, Wang Y. Venn diagrams in bioinformatics. *Briefings Bioinf* (2021) 22:bbab108. doi: 10.1093/bib/bbab108
21. Yang C, Delcher C, Shenkman E, Ranka S. Machine learning approaches for predicting high cost high need patient expenditures in health care. *Biomed Eng Online* (2018) 17:131. doi: 10.1186/s12938-018-0568-3
22. Aruna S, Rajagopalan D. A novel SVM based CSSFFS feature selection algorithm for detecting breast cancer. *Int J Comput Appl* (2013).
23. Noble WS. What is a support vector machine? *Nat Biotechnol* (2006) 24:1565–7. doi: 10.1038/nbt1206-1565
24. Dias R, Cutts S, Massoud S. Frozen shoulder. *BMJ* (2005) 331:1453–6. doi: 10.1136/bmj.331.7530
25. Michou L, Lermusiaux JL, Teyssedou JP, Bardin T, Beaudreuil J, Petit-Teixeira E. Genetics of dupuytren's disease. *Joint Bone Spine* (2012) 79:7–12. doi: 10.1016/j.jbspin.2011.05.027
26. Schaefer H. Die periarthritus humeroscapularis. In: Payr E, Kirschnner M (eds) *Ergebnisse der Chirurgie und Orthopädie*. (Berlin, Heidelberg: Springer) (1936). doi: 10.1007/978-3-642-91091-3_4
27. Lund M. Dupuytren's contracture and epilepsy. *Acta Psychiatrica Scandinavica* (1941) 16:465–92. doi: 10.1111/j.1600-0447.1941.tb07416.x
28. Meulengracht E, Schwartz M. The course and prognosis of periarthrosis humeroscapularis with special regard to cases with general symptoms. *Acta Med Scandinavica* (1952) 143:350–60. doi: 10.1111/j.0954-6820.1952.tb14270.x
29. Lundberg J. The frozen shoulder. Clinical and radiographical observations. The effect of manipulation under general anesthesia. Structure and glycosaminoglycan content of the joint capsule. Local bone metabolism. *Acta orthopaedica Scandinavica* (1969) Suppl 119:1–59. doi: 10.3109/ort.1969.40.suppl-119.01
30. Wilkinson JM, Davidson RK, Swingle TE, Jones ER, Corps AN, Johnston P, et al. MMP-14 and MMP-2 are key metalloproteases in Dupuytren's disease fibroblast-mediated contraction. *Biochim Biophys Acta* (2012) 1822:897–905. doi: 10.1016/j.bbdis.2012.02.001
31. Cho CH, Lho YM, Hwang I, Kim DH. Role of matrix metalloproteinases 2 and 9 in the development of frozen shoulder: human data and experimental analysis in a rat contracture model. *J shoulder elbow Surg* (2019) 28:1265–72. doi: 10.1016/j.jse.2018.11.072
32. Yang R, Deng H, Hou J, Li W, Zhang C, Yu M, et al. Investigation of salmon calcitonin in regulating fibrosis-related molecule production and cell-substrate adhesion in frozen shoulder synovial/capsular fibroblasts. *J orthopaedic Res* (2020) 38:1375–85. doi: 10.1002/jor.24571
33. Akbar M, Crowe LAN, McLean M, Garcia-Melchor E, MacDonald L, Carter K, et al. Translational targeting of inflammation and fibrosis in frozen shoulder: Molecular dissection of the T cell/IL-17A axis. *Proc Natl Acad Sci USA*. (2021) 118(39): e2102715118. doi: 10.1073/pnas.2102715118
34. Tripković I, Ogorevc M, Vuković D, Saraga-Babić M, Mardešić S. Fibrosis-associated signaling molecules are differentially expressed in palmar connective tissues of patients with carpal tunnel syndrome and dupuytren's disease. *Biomedicines* (2022) 10:3214. doi: 10.3390/biomedicines10123214
35. Yang R, Tang Y, Hou J, Yu M, Long Y, Yamuhanmode A, et al. Fibrosis in frozen shoulder: Activation of IL-6 through PI3K-Akt signaling pathway in synovial fibroblast. *Mol Immunol* (2022) 150:29–38. doi: 10.1016/j.molimm.2022.07.007
36. Chen J, Zhu J, Zhu T, Cui J, Deng Z, Chen K, et al. Pathological changes of frozen shoulder in rat model and the therapeutic effect of PPAR- γ agonist. *J orthopaedic Res* (2021) 39:891–901. doi: 10.1002/jor.24920
37. Krause C, Kloen P, Ten Dijke P. Elevated transforming growth factor β and mitogen-activated protein kinase pathways mediate fibrotic traits of Dupuytren's disease fibroblasts. *Fibrogenesis Tissue Repair* (2011) 4:14. doi: 10.1186/1755-1536-4-14
38. Dolmans GH, Werker PM, Hennies HC, Furniss D, Festen EA, Franke L, et al. Wnt signaling and Dupuytren's disease. *New Engl J Med* (2011) 365:307–17. doi: 10.1056/NEJMoa1101029
39. Ruettermann M, Hermann RM, Khatib-Chahidi K, Werker PMN. Dupuytren's disease-etiology and treatment. *Deutsches Arzteblatt Int* (2021) 118:781–8. doi: 10.3238/arztbl.m2021.0325
40. Broekstra DC, Groen H, Molenkamp S, Werker PMN, van den Heuvel ER. A Systematic Review and meta-analysis on the strength and consistency of the associations between dupuytren disease and diabetes mellitus, liver disease, and epilepsy. *Plast reconstructive Surg* (2018) 141:367e–79e. doi: 10.1097/PRS.00000000000004120
41. Whelton C, Peach CA. Review of diabetic frozen shoulder. *Eur J orthopaedic Surg traumatology* (2018) 28:363–71. doi: 10.1007/s00590-017-2068-8
42. Dyer BP, Burton C, Rathod-Mistry T, Blagojevic-Bucknall M, van der Windt DA. Diabetes as a prognostic factor in frozen shoulder: A systematic review. *Arch Rehabil Res Clin translation* (2021) 3:100141. doi: 10.1016/j.rrct.2021.100141
43. Deng CC, Hu YF, Zhu DH, Cheng Q, Gu JJ, Feng QL, et al. Single-cell RNA-seq reveals fibroblast heterogeneity and increased mesenchymal fibroblasts in human fibrotic skin diseases. *Nat Commun* (2021) 12:3709. doi: 10.1038/s41467-021-24110-y
44. Landim AV, Peres MCR, Costa HHA, Silveira RMF, Costa AC, Parente MOM, et al. Feeding restriction in the pre and postpartum period of hair ewes raised in the semi-arid region: implications on performance and carcass traits of the progeny. *Trop Anim Health production* (2022) 54:303. doi: 10.1007/s11250-022-03279-4
45. Cheng P, Li S, Chen H. Macrophages in lung injury, repair, and fibrosis. *Cells* (2021) 10:436. doi: 10.3390/cells10020436
46. Tang PM, Nikolic-Paterson DJ, Lan HY. Macrophages: versatile players in renal inflammation and fibrosis. *Nat Rev Nephrol* (2019) 15:144–58. doi: 10.1038/s41581-019-0110-2