Journal of Orthopaedic Translation 44 (2024) 102–113

Contents lists available at ScienceDirect

Journal of Orthopaedic Translation

journal homepage: www.journals.elsevier.com/journal-of-orthopaedic-translation

Computational screening of biomarkers and potential drugs for arthrofibrosis based on combination of sequencing and large nature language model

Xi Chen [a,b,c], Cheng Li [a], Ziyuan Wang [b,c], Yixin Zhou [a,*], Ming Chu [b,c,**]

[a] Department of Adult Joint Reconstructive Surgery, Beijing Jishuitan Hospital, Capital Medical University, 31 East Xinjiekou Street, Beijing, 100035, China
[b] Department of Immunology, School of Basic Medical Sciences, Peking University, Beijing, 100191, China
[c] NHC Key Laboratory of Medical Immunology (Peking University), Beijing, 100191, China

ARTICLE INFO

Keywords:
AF
biomarker
GPR17
macrophage
Stiff knee

ABSTRACT

Background: Arthrofibrosis (AF) is a fibrotic joint disease resulting from excessive collagen production and fibrous scar formation after total knee arthroplasty (TKA). This devastating complication may cause consistent pain and dramatically reduction of functionality. Unfortunately, the conservative treatments to prevent the AF in the early stage are largely unknown due to the lack of specific biomarkers and reliable therapeutic targets.
Methods: In this study, we extracted1782 fibrosis related genes (FRGs) from 373,461published literature based on the large natural language processing models (ChatGPT) and intersected with the 2750 differential expressed genes (DEGs) from mRNA microarray (GSE135854). A total of 311 potential AF biomarker genes (PABGs) were obtained and functional analysis were performed including gene ontology (GO) annotation and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses. Subsequently, we accomplished validation in AF animal models with immobilization of the unilateral knee joints of 16 rabbits for 1-week, 2-weeks, 3-weeks and 4-weeks. Finally, we tested the biomarkers in a retrospective cohort enrolled 35 AF patients and 35 control group patients.
Results: We identified G-protein-coupled receptor 17 (GPR17) as a reliable therapeutic biomarker for AF diagnosis with higher AUC (0.819) in the ROC curve. A total of 21 potential drugs targeted to GPR17 were screened. Among them, pranlukast and montelukast have achieved therapeutic effect in animal models. In addition, we established an online AF database for data integration (https://chenxi2023.shinyapps.io/afdbv1).
Conclusions: These results unveiling therapeutic biomarkers for AF diagnosis, and provide potential drugs for clinical treatment.
The translational potential of this article: Our study demonstrated that GPR17 holds significant promise as a potential biomarker and therapeutic target for arthrofibrosis. Moreover, pranlukast and montelukast targeted to GPR17 that could be instrumental in the treatment of AF.

1. Introduction

Arthrofibrosis (AF) is a complicated postoperative complication after total knee arthroplasty (TKA) [1]. This fibrotic joint disorder was triggered by inflammatory reaction which lead to collagen deposition and fibrous scar formation in the knee cavity [2]. And these tissues restricted the range of motion (ROM) of patients and caused significantly limitation of daily activities in driving, sitting and even walking [3]. Moreover, patients with AF also suffer with constant pain and joint swollen unrelated to effusion [4]. With the increased of the TKA surgeries worldwide, the AF surged and become a heavy burden both on patients and surgeons [5,6].

Recently studies reported that the prevalence of the AF was underestimated, and the best management remains largely unknown due to the lack of specific diagnosis biomarker and poor understanding of the AF pathogenesis [5]. Surgical treatments for AF including manipulation under anesthesia (MUA) and revision TKA are traumatic which increased the risk of the complication and hospitalization cost [1]. Even

* Corresponding author.
** Corresponding author.Department of Immunology, School of Basic Medical Sciences, Peking University, Beijing, 100191, China.
E-mail addresses: orthoyixin@yahoo.com (Y. Zhou), famous@bjmu.edu.cn (M. Chu).

https://doi.org/10.1016/j.jot.2023.11.002
Received 1 August 2023; Received in revised form 31 October 2023; Accepted 6 November 2023

Table 1
Inclusive and exclusive criteria of the AF patients.

Number	Inclusive criteria	Exclusive criteria
1	Persistent limitation of flexion ROM<90° at minimum 1 year follow-up	Prosthetic joint infection
2	Revision of TKA surgery for high degree of psychological strain and restriction of quality of life	Revision of TKA surgery for instability or loosening of the implantation
3	Inform consent of the patients	No obvious fibrous tissue under the microscope

though, these procedures are no guarantee of improved function and alleviation of pain. It was reported that 28.2 % of AF patients may still suffer with severe pain even after these therapies, and 17.4 % undergo a second reoperation [7]. Therefore, non-surgical treatments are urgently needed to prevent the AF in the early stage based on the clearly insight of the AF mechanism and specific biomarkers for staging.

As a consequence of inflammatory response in joint, AF shares common pathogenic pathways with many other fibrotic diseases such as idiopathic pulmonary fibrosis, cystic fibrosis, capsular contracture, kidney fibrosis and liver cirrhosis [6,8]. As such, major advances in fibrosis of other organs should be taken into account for better understanding of the AF mechanisms.

However, as one of the most heated topics, the literature of fibrosis disease increased sharply in recent 10 years from 155,665 (2012) to 373,461 (2022). Given that over 18,000 literature publishing in a single year, finish reading task is nearly mission impossible which exerts information overload on researchers. These requirements encourage us to apply a natural language processing (NLP) algorithm for biomarker gene information extraction from the numerous literature.

Recently, ChatGPT, a large language model that has opened a new

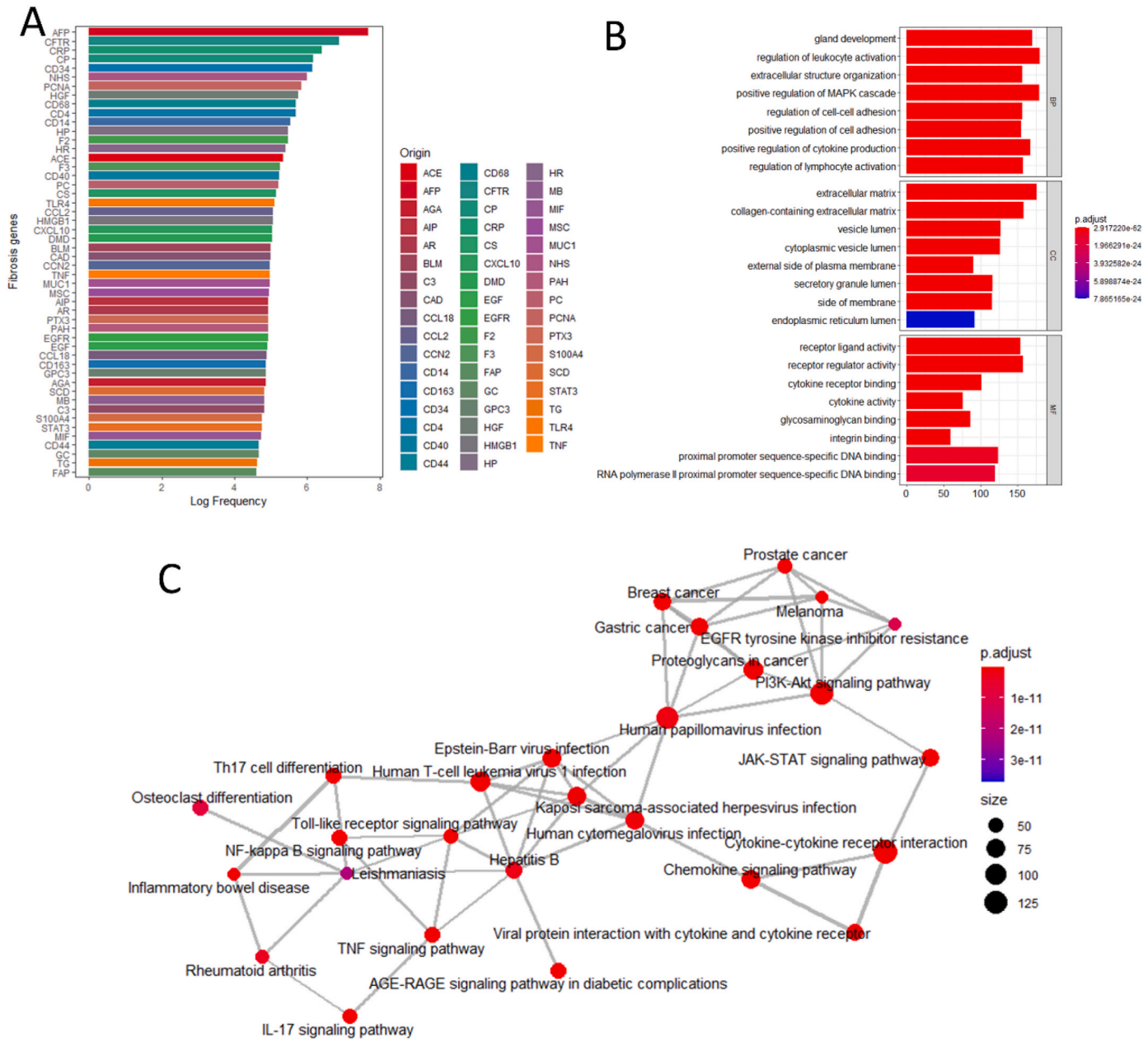

Figure 1. ChatGPT text mining results of fibrosis-related genes (FRGs) (A) Bar plot of the top 50 fibrosis-related genes ranked by frequency (B) Gene ontology analysis of fibrosis-related genes (FRGs). The color of the bar represents the adjusted p value, and the length of the bar represents the gene ratio (C) Network analysis of FRG-related pathways. The size of the point represents the gene count, while the color represents the adjusted p value.

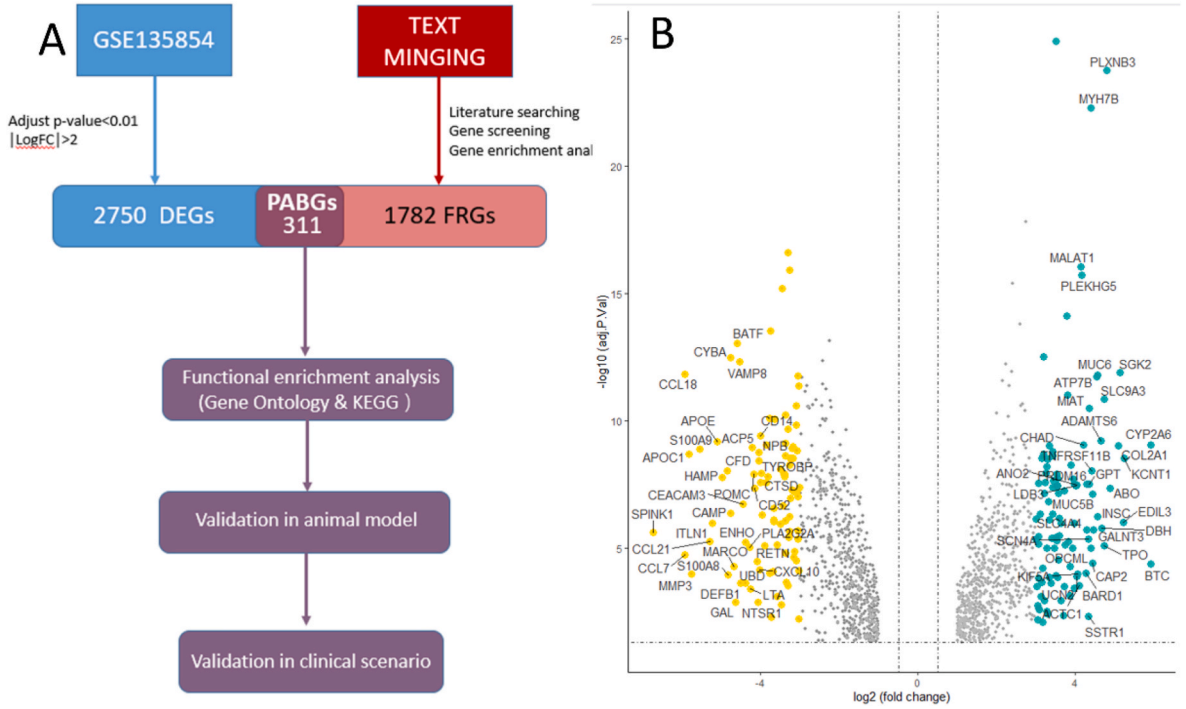

Figure 2. Identification of potential arthrofibrosis biomarker genes (PABGs) (A) Framework of PABG identification and further analysis (B) Volcano plot of PABSs. The green points and yellow points represent the significantly up- and downregulated genes (|Log2FC|>2).

era in NLP and achieved ideal results in various tasks. This model can precisely extract and capture key information or knowledge form a large scale of unstructured textual data without any pre-trained. It has been used in bioinformatics, computational biology and systems biology based on data mining, machine learning and NLP techniques. The extracted information often develops new facts and produces novel hypotheses via construction of knowledge network and linked with knowledge graphs.

In this study, we apply the ChatGPT to extract the potential biomarker genes and therapeutic targets of fibrosis diseases. Subsequently, we analyzed the GSE135854 expression array and identified DEGs. Base on the intersection of the DEGs and text mining gene set, we obtain potential biomarkers for functional analysis (Fig. 2A). Finally, a new therapeutic target was identified for further validation and potential drugs were discovered for AF treatment.

2. Methods

2.1. Data collection and preprocessing

We manage to download the gene expression chip data GSE135854 from the NCBI Gene Expression Omnibus web resource (GEO, http://www.ncbi.nlm.nih.gov/geo). This GSE135854 cohort included 4 A F patients after TKA and 4 non-AF patients after TKA [9]. After normalization of the samples expression data, the differential expression genes (DEGs) were identified based on the calculation of the FDR and the logFC between the AF-group and the Non-AF group. The remarkable DEGs ($|\log FC| \geq 2$, FDR <0.01) were selected for downstream analysis.

2.2. Functional enrichment analysis

We applied Gene ontology (GO) analysis and the Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis for functional characterization and pathway enrichment. The ClusterProfiler packages of R software was utilize and the significant level was set to FDR$<$0.05 [10].

2.3. Text mining and correlation analysis

To investigate the potential fibrosis related biomarkers, we performed a search of PubMed. Searching label was defined as “fibrosi" [All Fields] OR “fibrosing" [All Fields] OR “fibrosis" [MeSH Terms] OR “fibrosis" [All Fields] OR “fibrose" [All Fields] OR “fibroses" [All Fields] OR (“fibrotic" [All Fields] OR “fibrotics" [All Fields]) OR ((“fibrotic" [All Fields] OR “fibrotics" [All Fields]) AND (“action" [All Fields] OR “actions" [All Fields])) OR ((“fibrotic" [All Fields] OR “fibrotics" [All Fields]) AND (“change" [All Fields] OR “changed" [All Fields] OR “changes" [All Fields] OR “changing" [All Fields] OR “changings" [All Fields])). A total of 373,461 abstracts of the literature from January 1872 to January 2022 were involved in this analysis. All these abstracts were put into ChatGPT for biomarker gene information extraction and the frequency were analysis (Supplementary Table 1). The genes were extracted from these abstracts and applied for GO analysis and KEGG enrichment. Otherwise, the text mining gene set further intersected with the previously obtained DEGs for the next step of analysis.

2.4. AF animal model establishment

The animals involving in this study were in the guidance of the institutional animal research ethic and international guidelines. Sixteen male New Zealand white rabbits with the mean weight of 2.5 kg $\pm$ 0.5 kg were randomly divided into 4 groups. The right knees of each individual in these four groups were immobilized with 1.2 mm. It has been used in bioinformatics, computational biology and systems biology based on data mining, machine learning and natural language processing techniques. (K-wire) accordance with the previous protocols [11–13]. The left knees of each group were intact as the control groups.

An overdose of urethane was used humanely at the endpoint of 1 week, 2 weeks, 3 weeks and 4 weeks over these four groups. After the K-wire removed, the range of motion (ROM) of the fixation knees and the control knees were measured base on the previous studies described [14]. We used a looped wire to hook on the distal leg and exerted 5 N

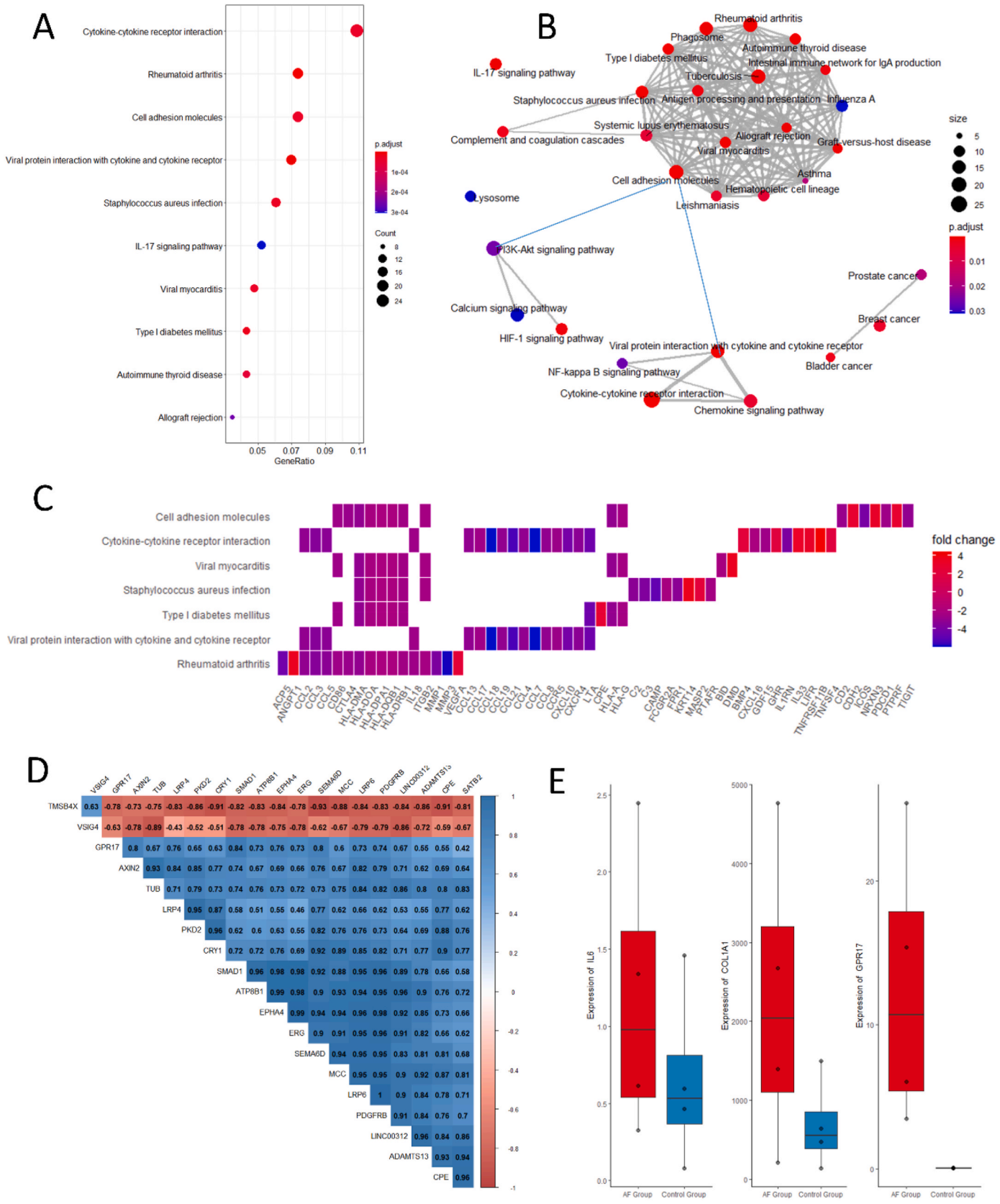

Figure 3. Functional analysis of potential arthrofibrosis biomarker genes (PABGs) (A) Gene ontology analysis of PABGs. The color of the bar represents the adjusted p value, and the length of the bar represents the gene ratio (B)Network analysis of PABG-related pathways. The size of the point represents the gene count, while the color represents the adjusted p value (C)Heatmap of key genes related to the extracellular matrix. The color of the bar represents the adjusted p value (D)Correlation coefficient matrix of PABGs (E)Expression of IL-6, COL1A1 and GPR17 in the arthrofibrosis group and control group.

Table 2
The top 20 potential arthrofibrosis biomarker genes (PABGs).

Gene	Genes Name	Freq	log2FC	FDR	DE
GPR17	G protein-coupled receptor 17	2	6.88	7.00E-24	Up
SEMA6D	semaphoring 6D	5	3.427	4.21E-08	Up
EPHA4	EPH receptor A4	1	2.931	1.04E-05	Up
SATB2	SATB homeobox 2	3	2.875	1.44E-06	Up
CPE	carboxypeptidase E	4	2.679	2.18E-11	Up
LRP4	LDL receptor related protein 4	1	2.656	6.17E-06	Up
AXIN2	axin 2	2	2.5	0.000475	Up
TUB	TUB bipartite transcription factor	1	2.445	8.41E-06	Up
CRY1	cryptochrome circadian regulator 1	2	2.313	1.68E-05	Up
ERG	ETS transcription factor ERG	11	2.303	6.05E-05	Up
LINC00312	long intergenic non-protein coding RNA 312	5	2.24	1.97E-05	Up
PKD2	polycystic 2, transient receptor potential cation channel	4	2.228	1.49E-06	Up
ATP8B1	ATPase phospholipid transporting 8B1	7	2.227	1.57E-06	Up
MCC	MCC regulator of WNT signaling pathway	10	2.209	6.47E-06	Up
LRP6	LDL receptor related protein 6	7	2.165	2.09E-05	Up
SMAD1	SMAD family member 1	3	2.112	0.000299	Up
ADAMTS13	ADAM metallopeptidase with thrombospondin type 1 motif 13	33	2.051	2.27E-07	Up
PDGFRB	platelet derived growth factor receptor beta	6	2.05	0.000266	Up
TMSB4X	thymosin beta 4 X-linked	3	−2.54	1.68E-08	Down
VSIG4	V-set and immunoglobulin domain containing 4	6	−3.098	3.11E-05	Down

consistent force at the point 8 cm distal from the proximal tibia joint surface. The ROM of this model was defined as the angle of femur and tibia at that condition.

2.5. Drug treatment of AF animal models

Subsequently, we carried out random grouping for another 21 New Zealand white rabbits. Three were designated as the control group, while nine were assigned to the montelukast treatment group, and another nine to the pranlukast treatment group. Each treatment group was further divided into three subgroups, receiving Intraarticular injections of the drug at low dose (1 mg/kg), medium dose (5 mg/kg), or high dose (10 mg/kg) into the joint cavity after surgery. Different dosage of montelukast (TargetMol,T1677) and pranlukast (TargetMol, T0694) was diluted into 1 ml of normal saline based on animal weight. All of these procedures were exclusively performed on the right leg, with the left leg serving as the control group.

2.6. Hydroxyproline content determination and immunohistochemistry evaluation

The synovial membrane in the joint cavity were carefully removed from both fixation knees and the control knees. In accordance with the previous study, we used 20 mg (wet weight) of the synovial tissue for hydroxyproline content determination [13]. The rest of the synovial tissue were fixed in the 10 % buffered formalin solution and subsequently embedded in paraffin. Slices of 4-μm thick were stained with hematoxylin/eosin (HE) to determine the severity of AF.

For immunohistochemical, we used the automated immunostainer (Autostainer 720, Labvision) based on the standard heat-induced epitope retrieval and the avidin–biotin–peroxidase complex method. The synovial tissue were incubated with one of these primary antibodies: anti-α-SMA antibody (1:500,ab32575, Abcam), anti-CD68 antibody (ab125212, Abcam) or anti-GPR17 antibody (1:50,ab279382, Abcam). The slices were analysis under the optical microscope from Olympus and evaluated with the digital scanner (KF-PRO-005 Magscanner).

The IHC score were assess by two pathologists independently as the previous studies mentioned about. The target field were set as 100 ± 5 positive stained fibroblasts field and graded from the lowest (0 score) to the highest (3 score).

2.7. Quantitative reverse-transcription polymerase chain reaction analysis

Total RNA was isolated from the synovial membrane of each knee with the TRIzol® Reagent (Plant RNA Purification Reagentfor plant tissue). We used the RevertAid First Strand cDNA Synthesis Kit (Thermo Scientific, San Jose, CA) to get the reverse transcription based on the standard instructions. Specific primers (GAPDH: F—TCA CCA TCT TCC AGG AGC GA and R—CAC AAT GCC GAA GTG GTC GT; ACTA2(a-SMA) F—GACCGAATGCAGAAGGAG R—CGGTGGACAATGGAAGG; GPR17 F—CACCCTGTCAAGTCCCTCAAG R—GTGGGCTGACTAGCAGTGG) were selected as the previous studies mentioned [15].

2.8. Patients and samples collection

673 Patients' records were carefully inspected who underwent revision TKA surgery from January 2010 to January 2020. The baseline data including: gender, age, body mass index (BMI) and range of motion (ROM) were collected. A total 35 patients with AF were enrolled based on the inclusive criteria and exclusive criteria (Table 1) [16–18]. Another 35 patients were selected as the control group matching with their age and gender. We collected the synovial tissue during the revision surgery for further testing with patients' informed consent. All these samples were kept in the JST hospital pathology sample bank after fixation.

2.9. Homology modeling and drug screening

The protein sequence of the G-protein-coupled receptor 17 (GPR17) was obtain for Gene Card database. C-X-C motif chemokine ligand 10 (CXCL10) was selected as the template for modeling, and crystal structure was downloaded from Protein Data Bank (PDB, http://www.rcsb.org/). Based on the homology modeling protocol provided by the Discovery Studio 2017R2 (DS; BIOVIA-Dassault Systèmes, San Diego, CA, USA), the 3D structure of GPR17 was constructed [19]. Active sites of GPR17 were also searched by the DS software with the binding active cavity searching protocol. The bioactivity of binding drugs was screened via the ChEMBL database (https://www.ebi.ac.uk/ChEMBLdb) [20].

2.10. Molecular docking

The molecular docking was performed with Discovery Studio 2021 (DS; BIOVIA-Dassault Systèmes). Firstly, using tools for active site identification in DS, we determined the active site of GPR17. Subsequently, we prepared all the drug molecules, generating multiple conformations for each drug. Lastly, we performed molecular docking between the prepared drug files and the predicted files of GPR17 using the Libdock method, and computed the docking scores (Table 5).

2.11. Online database establishment

This online database (AFDB version 1.0) was created by the shiny (https://CRAN.R-project.org/package=shiny). The small molecular candidate drugs were visualized with the r3dmol (https://github.com/swsoyee/r3dmol).

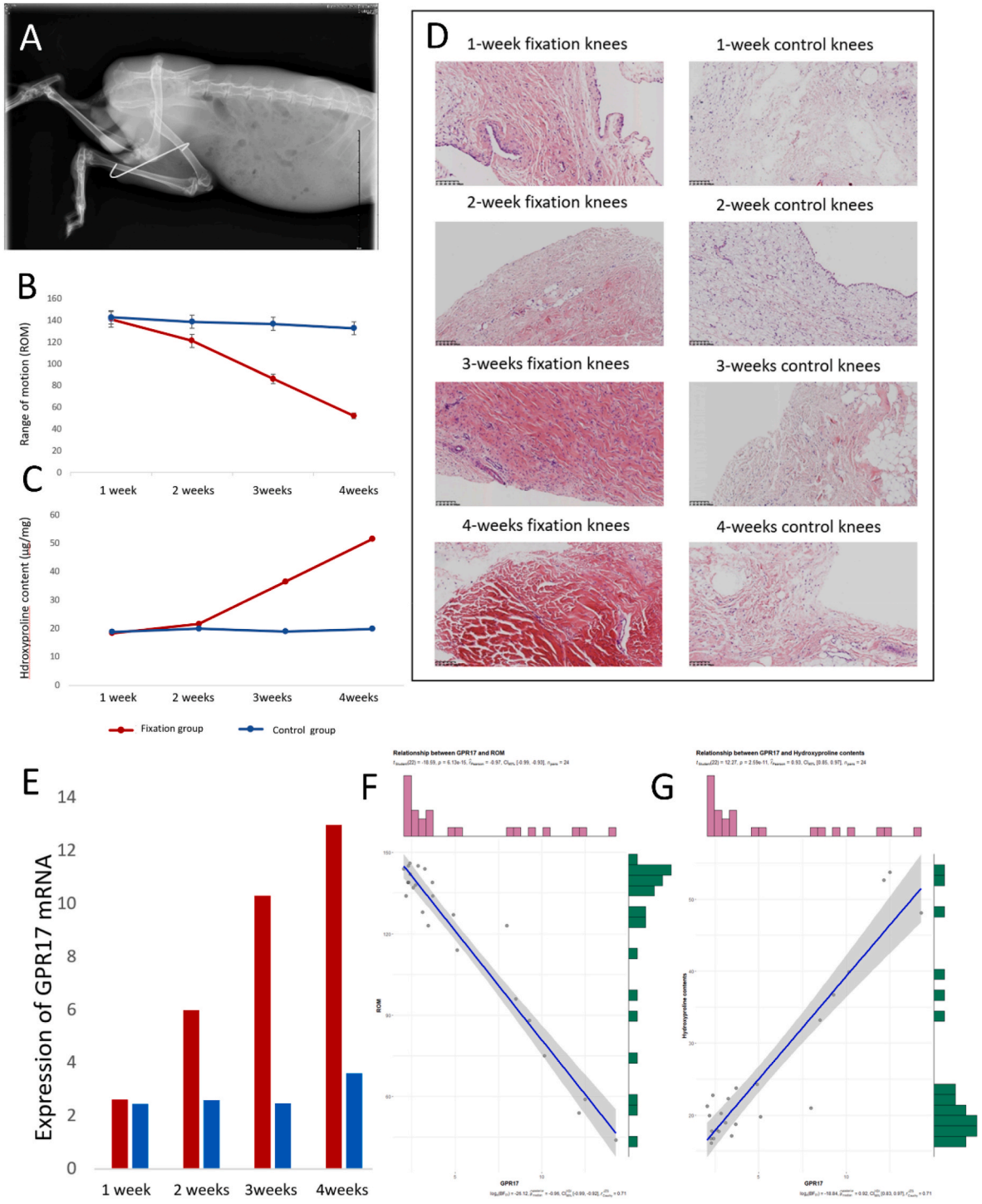

Figure 4. Analysis of GPR17 expression in an animal model (A) Representative X-ray photo of the rabbit fixation knee (lower leg) and control knee (upper leg)(B) Comparison of the range of motion (ROM) in the fixation and control groups(C) Comparison of the hydroxyproline contents in each group (D) Representative histological slices of rabbit knees in the control and fixation groups at 1 week, 2 weeks, 3 weeks and 4 weeks (E) Expression of GPR17 mRNA in the fixation and control groups (F) Correlation analysis between GPR17 expression and ROM (G) Correlation analysis between GPR17 expression and hydroxyproline content;

2.12. Statistical analysis

The quantitative data were reported with a mean and standard deviation calculated by the R software. The chi-squared test or *t*-test were performed to compare the differences between two groups of categorical variables and continuous variables separately. Otherwise, the Kruskal–Wallis test was used to determine the difference of the DEGs between these groups. The correlation analysis was performed with the Corrplot packages [21]. The comparison between the ROC curve were calculated with the DeLong's test by the pROC package of R programming software [22]. For all the analysis, the p-value$<$0.05 indicated significant difference in this study.

Table 3
Characteristic of baseline information of enrolled patients.

Parameter	Arthrofibrosis group	Non-arthrofibrosis group	P-value
Number of patients	35	35	—
Age (year)	66.94 ± 2.32	69.25 ± 1.48	0.4045
Gender (male/female)	14/21	15/20	0.8116
BMI(kg/㎡)	25.95 ± 0.62	24.67 ± 0.59	0.147
affection side (left/right)	18/17	20/15	0.6372
Baseline ROM	70.57 ± 3.13	132.10 ± 2.03	2.2E-16
Average score of GPR17	2.31 ± 0.11	1.28 ± 0.31	9.49E-07
Average score of SMA	1.91 ± 0.13	1.45 ± 0.31	0.02824
Average score of CD68	1.25 ± 0.17	1.97 ± 0.19	0.00488

3. Results

3.1. Text mining analysis of FRGs

In this study, we enrolled 373,461abstracts of the fibrosis related diseases from January 1872 to January 2022. A total of 1782 genes were extracted from the literature by the text mining approach (Supplementary Table 1). Top 50 genes including CD68 molecule, CF transmembrane conductance regulator (CFTR) CXCL10, C–C motif chemokine ligand 2 (CCL2), and epidermal growth factor receptor (EGFR) were show in Fig. 1A ranking by the gene frequency.

To establish the potential role of these fibrosis related genes (FRGs), we performed GO term analysis on the 1782 genes. The results indicated that these FRGs were enriched for regulation of the cell adhesion (BP), extracellular matrix (CC) and cytokine receptor bind (MF) (Fig. 1B). While the KEGG analysis revealed 111 significantly enriched pathways (Supplementary Table 2). The network of the KEGG manifest a tight connection and widely cross talk between these pathways (Fig. 1C). PI3K-Akt signaling pathway has been interwoven into the FRGs network of cell signal and plays important role in the arthrofibrosis pathogenesis [23]. These results may contribute to the understanding of the fibrosis related disease and the hub gene identification.

3.2. Potential AF biomarker genes (PABGs) identification

Based on the analysis of GSE135854, we identified 2750 differential expressed genes (DEGs). The criteria were set as $|\log FC| \geq 2$ and FDR <0.01. All these genes including 1550 upregulated genes and 1200 down-regulated genes were listed in the Supplementary Table 3. Then the DEGs were intersected with the FRGs from text mining, and 311 potential AF biomarker genes (PABGs) were obtained (Fig. 2A). The PABGs were depicted by volcano plot in Fig. 2B and listed in Supplementary Table 4.

3.3. Functional enrichment analysis of PABGs

Subsequently, we carried out the functional enrichment analysis on the 311 PABGs. GO term analysis of PABGs were very similar to the FRGs, which mainly enriched for regulation of the immunology cell adhesion (BP), extracellular matrix (CC) and cytokine receptor bind (MF) (Fig. 3A). This result indicated that these PABGs were highly correlated with the fibrosis pathogenesis. The only difference is the PABGs involving immune response including leukocyte cell adhesion and T cell activation. This may indicated that the AF are more relative to the inflammatory. The KEGG network analysis also showed two important pathway in AF: TGF-β signaling pathway and PI3K-Akt signaling pathway. These two pathways are interwoven into the vast disease-pathway network and intervene with many other diseases including rheumatoid arthritis (RA) and *Staphylococcus aureus* infection.

GO analysis also provide expression level of clear annotated extracellular matrix genes (Fig. 3C). These genes are used as the indicators for correlation analysis. The top 20 PABGs with high correlation (R square >0.6) of these indicators were define as the identified biomarker genes (Table 2). The correlation of these identified biomarker genes show that these genes are highly related (Fig. 3D). The GPR17 ranking highest in the table was finally selected for further analysis. The expression level of IL-6, COL1A1 and GPR17 also show a sharp difference between the AF group and the Non-AF group (Fig. 3E).

3.4. Analysis of GPR17 expression in animal models

We have divided the 16 rabbits into four groups and fixed the right knee for 1 week, 2 weeks, 3 weeks and 4weeks with the K-wires (Fig. 4A). The ROM of the knee decrease sharply the hydroxyproline contents accumulated over time (Fig. 4B and C) Meanwhile, the samples of synovial membrane were observed in the microscope where the fibrotic/hyalinized tissue were exhibit (Fig. 4D). Interesting, macrophages are easily detected in the sample around the fibroblasts.

For quantitative assessment, the mRNA level of GPR17 was measured by RT-PCR (Fig. 4E). Based on the correlation analysis, GPR17 show highly relative with ROM and hydroxyproline contents (Fig. 4F and G). These results indicated that GPR17 increased over time which perform to be a reliable indicator for AF.

3.5. Validation of GPR17 in AF patients

We enrolled 35 patients who underwent revision TKA surgery for AF based on the criteria mentioned in Table 1. 35 patients underwent revision TKA without AF were match by age and gender. The baseline information such as age, gender, BMI and ROM are showed in Table 3. Except the baseline ROM, there was no significant difference between the two group. We obtained the synovial membrane tissue intraoperatively with the patients' informed consent. The slices of these samples were stained by HE and IHC. Immunohistochemical scores of GPR17, a-SMA and CD68 were calculated and showed in Table 3. There were significant differences between AF groups and Non-AF groups which were depicted in Fig. 5A-C. ROC curve manifest that GPR17 have a higher AUC area (0.819) over a-SMA (0.657) (Fig. 5D). Table 4 reported the best cut off point, sensitivity and specificity. The GPR17 diagnosis model achieved 0.6 sensitivity and 0.91 specificity at the threshold of 1.5.

Moreover, GPR17 also performed a higher correlation coefficient with ROM (0.88) comparing with a-SMA (0.68) (Fig. 5E–F). These results revealed that the GPR17 was a better indicator for AF diagnosis. Interestingly, the GPR17 also have high correlation with CD68, which is the biomarker of M1 macrophages (Fig. 5G). This may contribute to the underlying mechanism of CD68 mediated inflammation in the AF.

3.6. Drug screening and validation

Since the GPR17 is not available, we built the 3D structure of GPR17 by homology modeling (Fig. S1A). The CXC chemokine receptor 4 (CXCR4) was selected as the template because of the phylogenetically and structurally similarity to GPR17 [24]. Our construction of the GPR17 was consistent with the previous studies which consists of seven transmembrane domains region (TM1-TM7), eight part of amphipathic helices, and a specific portion of the sites for N-glycosylation [25]. The blue sphere was the most likely binding site of the GPR17 by our calculation. This cavity was consisted by seven alpha helixes located closer to the extracellular surface.

Then we searched in the CHEMBL database and obtained 23 small compounds which have inhibitory bioactivity on GPR17 (Table 5). Binding efficiency index (BEI) and surface efficiency index (SEI) various varied in Fig. S1B. Five of these drugs including Pranlukast, Montelukast, Zafirlukast, Cangrelor and Ticagrelor had been approved by

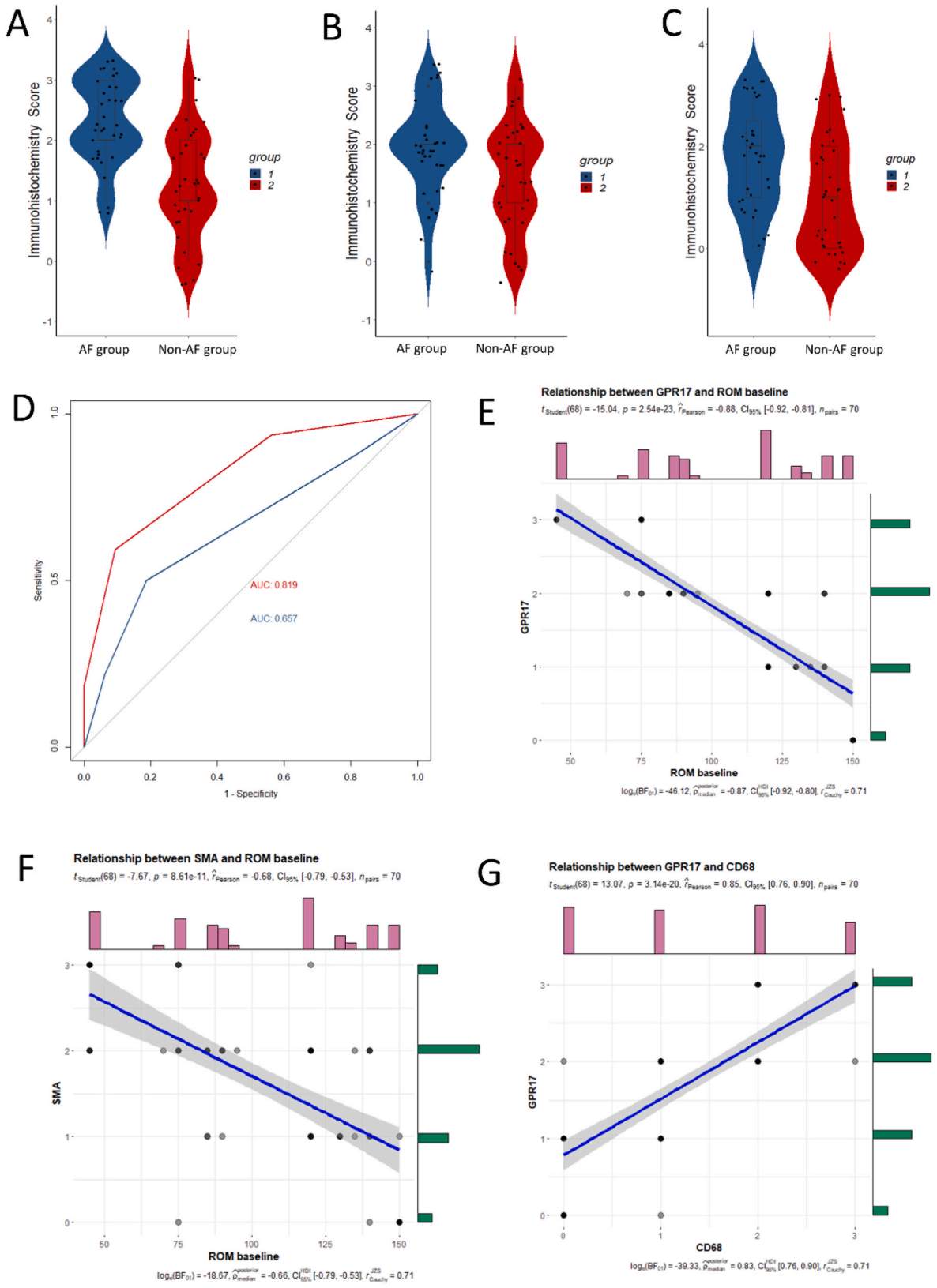

Figure 5. Analysis of GPR17 in clinical samples (A, B, C) Violin plot of the GPR17, α-SMA and CD68 scores in the arthrofibrosis group (colored red) and the control group (colored blue). The scatter points represent every single data point of these two groups (D) ROC curve of the GPR17 scoring model (colored red). ROC curve of the α-SMA scoring model (colored blue) (E, F, G) Correlation analysis between E) GPR17 score and ROM baseline; F) α-SMA score and G) GPR17 score and CD68.

Table 4
ROC curve of the diagnosis test.

Indicators	Parameters	Values
GPR17	AUC (95%CI)	0.82 [0.73–0.91]
	Sensitivity	0.6
	Specificity	0.91
	Cutoff point	1.5
a-SMA	AUC (95%CI)	0.66 [0.53–0.78]
	Sensitivity	0.5
	Specificity	0.812
	Cutoff point	2.5

Table 5
Drugs target to GPR17.

ChEMBL ID	Molecular Weight	Compound Name	Bioactivity	Docking score
CHEMBL21333	481.51	pranlukast	Ki = 4060 nM	99.8395
CHEMBL787	586.2	montelukast	Ki = 6540 nM	99.1312
CHEMBL3264006	310.15	PSB-12150	Kd = 1256 nM	97.0718
CHEMBL3968577	309.32	SAL024	EC50 = 724.44 nM	94.065
CHEMBL3897541	261.28	RA–III–55	EC50 = 13803.84 nM	94.0288
CHEMBL4206497	333.38	PSB-1867	Ki = 1500 nM	91.3115
CHEMBL3915176	385.42	SAL016	EC50 = 3019.95 nM	88.1505
CHEMBL3132880	391.02	KL126	EC50 = 13.49 nM	87.1496
CHEMBL3921330	374.18	RA–III–40	EC50 = 186.21 nM	86.9839
CHEMBL3132887	359.12	SAL009-1	EC50 = 707.95 nM	84.417
CHEMBL3132881	485.02	SAL019	EC50 = 4786.3 nM	82.2135
CHEMBL398435	522.58	Ticagrelor	Unknown	81.1375
CHEMBL334966	776.37	Cangrelor	Unknown	80.6045
CHEMBL3132885	312.12	SAL006-1	EC50 = 3162.28 nM	75.5553
CHEMBL603	575.69	Zafirlukast	Unknown	74.6532
CHEMBL53914	269.2	KL28	EC50 = 7413.1 nM	74.3483
CHEMBL4215280	367.83	PSB-18183	Ki = 13100 nM	68.6407
CHEMBL44793	375.21	GV150526A	Ki = 1630 nM	66.3742
CHEMBL295718	267.67	KL16-1	EC50 = 3715.35 nM	52.0039
CHEMBL31344	302.11	RA–II–150	EC50 = 812.83 nM	2.5496

FDA, and 3D structures are show in Fig. S1C to Fig. S1G. Subsequently, we performed molecular docking for all the drugs, and selected the top two drugs with the highest docking scores (Pranlukast and Montelukast) for subsequent validation in animal experiments (Table 5, Fig. 6A–D). The results of the animal experiments confirmed that both Pranlukast and Montelukast, in comparison to the control group, exhibited dose-dependent reduction of hyaline cartilage in the synovium and increased joint mobility (Fig. 6E and F).

4. Discussion

In this study, we identified 311 PABGs by intersection of 2750 DEGs and 1782 FRGs. We then accomplished functional analysis and constructed the KEGG pathway network. Base on the correlation analysis, total of 20 reliable biomarker were obtained (Table 2) and the GPR17 was selected as the key biomarker genes and therapeutic target for validation and drug screening.

GPR17, also known as the P2Y-like receptor and R12, belong to the G protein-coupled receptors (GPCRs) superfamily. Our results manifest that the GPR17 is time-dependent upregulated in the fixation animal knee. It was significantly correlated with the ROM and the hydroxyproline contents accumulation. Study in the retrospective cohort also proved that GPR17 take advantage over a-SMA for AF diagnosis due to the higher AUC area (0.819 over 0.675) and higher correlation coefficient with the ROM (0.88 over 0.68). Several studies have proposed the GPR17 un-regulation in pulmonary fibrosis and myocardial fibrosis, but there is no relative research on the AF field. Our study is the first time report the GPR17 as the potential biomarker and underlying therapeutic target for AF.

Previous studies have elucidated that the GPR17 was highly expressed in organs undergoing ischemic damage. The GPR17 upregulation had been proved to contribute to the myocardial fibrosis due to the heart ischemia. It also mediated the pulmonary fibrosis due to the macrophage related pulmonary inflammation. Our results also showed the up-regulation of the GPR17 in the process of the AF which is probably due to the ischemia in the TKA surgery. This may help to understand the lower inflammation markers and better ROM without using the tourniquet in previous studies [26].

Moreover, we have also implicated the relationship between the GPR17 and the macrophage. The macrophage is known as a master regulator of inflammation and plays a key role in all stage of AF [27]. We have observed the correlation between GPR17 and the CD68 expression. Previous studies have been proved that the GPR17 can activate macrophage in multiple organs including CNS, kidney, heart and lung. Zhan and his college have found that the GPR17 mediated the pulmonary fibrosis by activate the macrophage related pulmonary inflammation [28]. Heike Franke and college uncover the GPR17 expression in the CNS and activate the microglia and macrophages which lead to the irreversible death of the neurons [29]. As such, we inferred the GPR17 act as the same role in the AF pathogenesis.

In addition, previous studies have elucidated that the GPR17 was highly expressed in organs undergoing ischemic damage. Simona Cosentino reported that the GPR17 upregulation had been proved to contribute to the myocardial fibrosis due to the heart ischemia [30]. Our results support the hypothesis that the ischemia caused by tourniquet usage during TKA surgery may contribute to the up-regulation of GPR17 and leading to the macrophage activation. The activated macrophage produced profibrotic mediators such as TGF-b that promote collagen synthesis in myofibroblasts. Numerous studies have revealed that the various macrophage was key producers and dominant resources of TGF-b1 [31–33]. Recent study suggested that the only TGF-b produced by the macrophage can promote fibrosis while the TGF-b derived from regulatory T cells suppresses the inflammation. These finding help to explain the lack of critical progress with classical TGF-b signaling inhibitors for fibrosis treatment. With this, the identification of GPR17 provide a promising therapeutic target for suppression of the profibrotic macrophage which is more rational for ameliorate AF.

Indeed, numerous studies have reported that GPR17 targeted drugs have showed significant improvement on fibrosis [34]. As such, we proposed the potential binding site and 23 underlying drugs targeting to the GPR17. CysLT receptor antagonists such as Pranlukast, Montelukast and Zafirlukast can reduce the renal fibrosis, pulmonary fibrosis and post-operative capsular contracture formation [35–38]. Cangrelor and Ticagrelor are a well-known anti-platelet agent which have the bioactivity of blocking GPR 17 in macrophage. Previous studies show they can alleviate pulmonary fibrosis and myocardial fibrosis by inhibiting the GPR-17 induced inflammation in mice [28,39]. Some studies revealed that Cytl1 may appears to activate the TGF-beta-SMAD signaling pathway and plays a significant role in cardiac fibrosis [1,40]. It is noteworthy that early in 2014, Will Efird and his colleagues also reported that the montelukast provided a significantly improvement on the arthrofibrosis after trauma. With a dose of 3.75 mg/kg/day oral montelukast, the rats performed 12°improvement of ROM. Our research has demonstrated that intra-articular administration of montelukast may yields enhanced efficacy of AF treatment.

In this study, we proposed novel hypothesis of AF which may

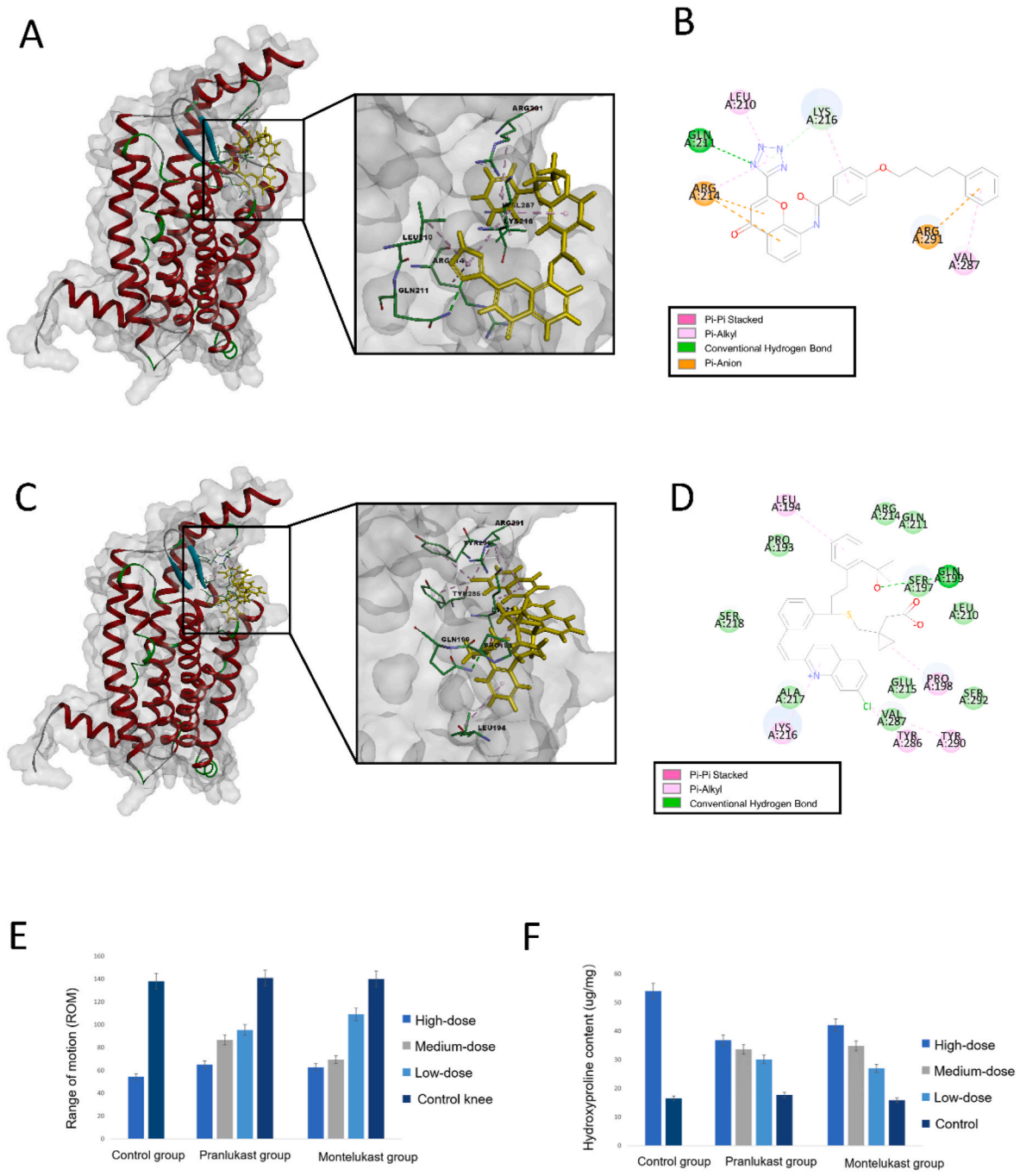

Figure 6. Validation of drugs (A, C) The 3-dimensional complex of GPR17 protein and pranlukast (A) and montelukast (C) after molecular docking. GPR17 protein secondary structural elements are depicted as ribbons (coils, α-helices; arrows, β-sheets). Color is based on secondary structures (α-helices, red; β-sheets, sky-blue; loops, green). Pranlukast and montelukast was shown as sticks highlighted with yellow color. The image on the right was magnified to display the key amino acid residues that interact with pranlukast/montelukast. These key residues were shown as sticks with carbon, oxygen, and nitrogen colored gray, red, and blue, respectively. Interaction between pranlukast/montelukast and GPR17 was shown as dashed lines with π-π, π-alkyl, and hydrogen bonds colored purple, pink, and green, respectively (B, D) 2-dimensional depiction of pranlukast (B) and montelukast (D) in the active site of the GPR17. Residues involved in hydrogen bonding or polar interactions are represented by dark green circles, and residues involved in hydrophobic interactions are shown by magenta-colored circles. Other residues in the binding site are represented by light green circles. Interaction between drugs and GPR17 was shown as dashed lines with π-π, π-alkyl, and hydrogen bonds colored purple, pink, and green, respectively (E) Comparison of the range of motion (ROM) among control groups and different dose of pranlukast and montelukast treatment groups (low dose, 1 mg/kg; medium dose, 5 mg/kg; and high dose, 10 mg/kg) (D) Comparison of the hydroxyproline contents among control groups and different dose of pranlukast and montelukast treatment groups (low dose, 1 mg/kg; medium dose, 5 mg/kg; and high dose, 10 mg/kg).

contribute to understand these encouraging results. TKA causes ischemia and induced the pro-inflammatory cytokines. This leads to a stimulation of macrophage and up-regulation of GPR17. The activated macrophage produced TGF-b and other inflammatory mediators which stimulated the fibroblasts to differentiate into myofibroblasts. Adhesion complexes on the surface of myofibroblasts attached to ECM and other cells which form fibrous scar and reduce the movement of the joints over time. Montelukast and Pranlukast may inhibited GPR17 activated of the macrophage and alleviated the inflammation in the knee joint after TKA surgery (Fig. 7) which contribute to the AF prevention.

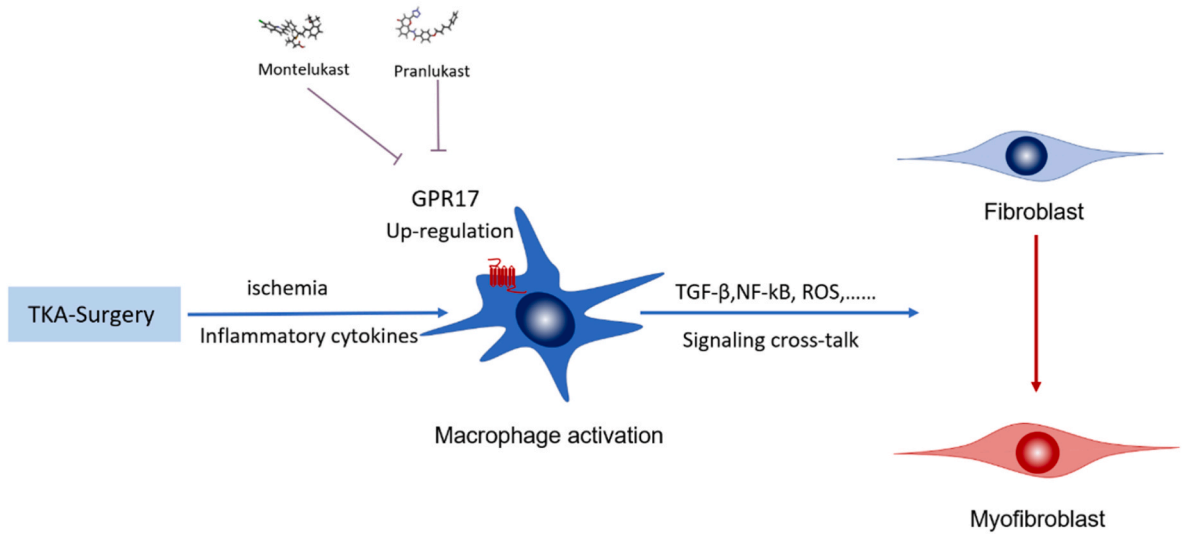

Figure 7. Schematic illustration of the proposed mechanism of AF. Schematic representation of the proposed mechanism of GPR17-targeted drugs and their contribution to arthrofibrosis. Total Knee Arthroplasty (TKA) induces ischemia and triggers pro-inflammatory cytokines, leading to macrophage stimulation and up-regulation of GPR17. The activated macrophages then produce TGF-b and other inflammatory mediators, stimulating fibroblasts to differentiate into myofibroblasts. Adhesion complexes on the surface of myofibroblasts attach to ECM and other cells, forming fibrous scars that progressively restrict joint movement over time. Montelukast and Pranlukast may inhibit GPR17 activation in macrophages and alleviate inflammation in the knee joint following TKA surgery.

There are still some limitations to this study. First, only the abstracts of the literature were included, and the language is limited to English. More information would be extracted if full texts or other language literature were involved in this study. Second, the number of patients enrolled in the RNA sequencing and validation cohort was relatively small. Moreover, further experiment of GPR17 regulation mechanism of the AF pathogenesis were still needed to explore. Future work will be conducted from this perspective.

5. Conclusion

Based on the combination of sequencing and large nature language model (ChatGPT), we identified GPR17 as an important regulator and a potential therapeutic target of AF. Pharmaceutical agents targeting GPR17, such as Montelukast and Pranlukast, have demonstrated a prophylactic effect against fibrosis in animal models. These results may contribute to the diagnosis and prevention for the AF.

Author contributions

M.C. and Y.X.Z. designed the experiments; X.C. and C.L. performed the experiments; X.C., C.L., Z.Y.W. and S.Y.W. analyzed the data; X.C. wrote the manuscript; M.C. and Y.X.Z. revised the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

The protocol of the animal study was approved by the Ethic Committee of Jishuitan hospital (No.201811–09) in accordance with the registration. The clinical part of this study was also conducted with the permission of the Ethics Committee of the Jishuitan hospital (No.202011–02). Informed consent was obtained from all of the patients enroll in this study.

Funding

This work was supported by the Significant Science and Technology Project of Beijing Life Science Academy [grant number 2023000CA0040]; the National Natural Science Foundation of China [grant number 81603119]; the Natural Science Foundation of Beijing Municipality [grant number 7174316]; the Peking University Medicine Seed Fund for Interdisciplinary Research supported by "the Fundamental Research Funds for the Central Universities" [grant number No. BMU2022MX017, No. 284 BMU2022MX003].

Declaration of competing interest

These authors declare no competing interests.

Acknowledgments

This work was supported by the Significant Science and Technology Project of Beijing Life Science Academy [grant number 2023000CA0040]; the National Natural Science Foundation of China [grant number 81603119]; the Natural Science Foundation of Beijing Municipality [grant number 7174316]; the Peking University Medicine Seed Fund for Interdisciplinary Research supported by "the Fundamental Research Funds for the Central Universities" [grant number No. BMU2022MX017, No. 284 BMU2022MX003].

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jot.2023.11.002.

References

[1] Cheuy VA, Foran JRH, Paxton RJ, Bade MJ, Zeni JA, Stevens-Lapsley JE. Arthrofibrosis associated with total knee arthroplasty. J Arthroplasty 2017;32(8):2604–11.
[2] Usher KM, Zhu S, Mavropalias G, Carrino JA, Zhao J, Xu J. Pathological mechanisms and therapeutic outlooks for arthrofibrosis. Bone Res 2019;7:9.
[3] Morup-Petersen A, Holm PM, Holm CE, Klausen TW, Skou ST, Krogsgaard MR, et al. Knee osteoarthritis patients can provide useful estimates of passive knee range of motion: development and validation of the Copenhagen knee ROM scale. J Arthroplasty 2018;33(9):2875–28783 e3.
[4] Chen AF, Lee YS, Seidl AJ, Abboud JA. Arthrofibrosis and large joint scarring. Connect Tissue Res 2019;60(1):21–8.
[5] Tibbo ME, Limberg AK, Salib CG, Ahmed AT, van Wijnen AJ, Berry DJ, et al. Acquired idiopathic stiffness after total knee arthroplasty: a systematic review and meta-analysis. J Bone Joint Surg Am 2019;101(14):1320–30.
[6] Nanthakumar CB, Hatley RJ, Lemma S, Gauldie J, Marshall RP, Macdonald SJ. Dissecting fibrosis: therapeutic insights from the small-molecule toolbox. Nat Rev Drug Discov 2015;14(10):693–720.

[7] Rutherford RW, Jennings JM, Levy DL, Parisi TJ, Martin JR, Dennis DA. Revision total knee arthroplasty for arthrofibrosis. J Arthroplasty 2018;33(7S):S177–81.
[8] Rockey DC, Bell PD, Hill JA. Fibrosis–a common pathway to organ injury and failure. N Engl J Med 2015;372(12):1138–49 [eng].
[9] Bayram B, Limberg AK, Salib CG, Bettencourt JW, Trousdale WH, Lewallen EA, et al. Molecular pathology of human knee arthrofibrosis defined by RNA sequencing. Genomics 2020;112(4):2703–12.
[10] Guangchuang Yu L-GW, Han Yanyan, He Qing-Yu. clusterProfiler: an R package for comparing biological themes among gene clusters. J. Integr. Biol. 2012;16(5):284–7.
[11] Wang J, Yan L, Sun Y, Wang D, Dai S, Yu T, et al. A comparative study of the preventive effects of mitomycin C and chitosan on intraarticular adhesion after knee surgery in rabbits. Cell Biochem Biophys 2012;62(1):101–5.
[12] Huang PP, Zhang QB, Zhou Y, Liu AY, Wang F, Xu QY, et al. Effect of radial extracorporeal shock wave combined with ultrashort wave diathermy on fibrosis and contracture of muscle. Am J Phys Med Rehabil 2021;100(7):643–50. https://doi.org/10.1097/PHM.0000000000001599.
[13] Sun Y, Liang Y, Hu J, Wang J, Wang D, Li X, et al. Reduction of intraarticular adhesion by topical application of colchicine following knee surgery in rabbits. Sci Rep 2014;4:6405.
[14] Efird W, Kellam P, Yeazell S, Weinhold P, Dahners LE. An evaluation of prophylactic treatments to prevent post traumatic joint stiffness. J Orthop Res 2014;32(11):1520–4.
[15] Chen K, Yu Z, Yang J, Li H. Expression of cysteinyl leukotriene receptor GPR17 in eosinophilic and non-eosinophilic chronic rhinosinusitis with nasal polyps. Asian Pac J Allergy Immunol 2018;36(2):93–100. https://doi.org/10.12932/AP-030417-0063.
[16] Pitta M, Esposito CI, Li Z, Lee YY, Wright TM, Padgett DE. Failure after modern total knee arthroplasty: a prospective study of 18,065 knees. J Arthroplasty 2018;33(2):407–14.
[17] Gandhi R, de Beer J, Leone J, Petruccelli D, Winemaker M, Adili A. Predictive risk factors for stiff knees in total knee arthroplasty. J Arthroplasty 2006;21(1):46–52.
[18] Boldt JG, Stiehl JB, Hodler J, Zanetti M, Munzinger U. Femoral component rotation and arthrofibrosis following mobile-bearing total knee arthroplasty. Int Orthop 2006;30(5):420–5.
[19] Chu M, Chen X, Wang J, Guo L, Wang Q, Gao Z, et al. Polypharmacology of berberine based on multi-target binding motifs. Front Pharmacol 2018;9:801.
[20] Gaulton A, Bellis LJ, Bento AP, Chambers J, Davies M, Hersey A, et al. ChEMBL: a large-scale bioactivity database for drug discovery. Nucleic Acids Res 2012;40(Database issue):D1100–7.
[21] Simko TWaV. R package "corrplot": Visualization of a Correlation Matrix (Version 0.84) 2017. https://githubcom/taiyun/corrplot.
[22] Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinf 2011;12(1):77.
[23] Sun Y, Dai J, Jiao R, Jiang Q, Wang J. Homoharringtonine inhibits fibroblasts proliferation, extracellular matrix production and reduces surgery-induced knee arthrofibrosis via PI3K/AKT/mTOR pathway-mediated apoptosis. J Orthop Surg Res 2021;16(1):9.
[24] Wu B, Chien EYT, Mol CD, Fenalti G, Liu W, Katritch V, et al. Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. Science 2010;330(6007):1066–71 [eng].
[25] Marucci G, Dal Ben D, Lambertucci C, Santinelli C, Spinaci A, Thomas A, et al. The G protein-coupled receptor GPR17: overview and update. ChemMedChem 2016;11(23):2567–74.
[26] Zhao HY, Yeersheng R, Kang XW, Xia YY, Kang PD, Wang WJ. The effect of tourniquet uses on total blood loss, early function, and pain after primary total knee arthroplasty: a prospective, randomized controlled trial. Bone Joint Res 2020;9(6):322–32.
[27] Wynn TA, Barron L. Macrophages: master regulators of inflammation and fibrosis. Semin Liver Dis 2010;30(3):245–57.
[28] Zhan T-W, Tian Y-X, Wang Q, Wu Z-X, Zhang W-P, Lu Y-B, et al. Cangrelor alleviates pulmonary fibrosis by inhibiting GPR17-mediated inflammation in mice. Int Immunopharm 2018;62:261–9.
[29] Abbracchio Hfcpdlerzchkbmfprllnsm-Gdsmwmp. Changes of the GPR17 receptor, a new target for neurorepair, in neurons and glial cells in patients with traumatic brain injury. Purinergic Signal 2013;2013(9):451–62.
[30] Cosentino S, Castiglioni L, Colazzo F, Nobili E, Tremoli E, Rosa P, et al. Expression of dual nucleotides/cysteinyl-leukotrienes receptor GPR17 in early trafficking of cardiac stromal cells after myocardial infarction. J Cell Mol Med 2014;18(9):1785–96.
[31] Lin SL, Castano AP, Nowlin BT, Lupher Jr ML, Duffield JS. Bone marrow Ly6Chigh monocytes are selectively recruited to injured kidney and differentiate into functionally distinct populations. J Immunol 2009;183(10):6733–43.
[32] Kitani A, Fuss I, Nakamura K, Kumaki F, Usui T, Strober W. Transforming growth factor (TGF)-beta1-producing regulatory T cells induce Smad-mediated interleukin 10 secretion that facilitates coordinated immunoregulatory activity and amelioration of TGF-beta1-mediated fibrosis. J Exp Med 2003;198(8):1179–88 [eng].
[33] Xing Z, Tremblay GM, Sime PJ, Gauldie J. Overexpression of granulocyte-macrophage colony-stimulating factor induces pulmonary granulation tissue formation and fibrosis by induction of transforming growth factor-beta 1 and myofibroblast accumulation. Am J Pathol 1997;150(1):59–66 [eng].
[34] Ide M, Kuwamura M, Kotani T, Sawamoto O, Yamate J. Effects of gadolinium chloride (GdCl(3)) on the appearance of macrophage populations and fibrogenesis in thioacetamide-induced rat hepatic lesions. J Comp Pathol 2005;133(2–3):92–102.
[35] Huang CK, Handel N. Effects of Singulair (montelukast) treatment for capsular contracture. Aesthetic Surg J 2010;30(3):404–8 [eng].
[36] Hur J, Kang JY, Rhee CK, Kim YK, Lee SY. The leukotriene receptor antagonist pranlukast attenuates airway remodeling by suppressing TGF-β signaling. Pulm Pharmacol Therapeut 2018;48:5–14 [eng].
[37] Miyaguchi S, Oda M, Saito H, Ishii H. Novel therapeutic approach to primary biliary cirrhosis patients: anti-eosinophil strategy. Hepato-Gastroenterology 1998;45(23):1457–61 [eng].
[38] Shimbori C, Shiota N, Okunishi H. Pranlukast, a cysteinyl leukotriene type 1 receptor antagonist, attenuates the progression but not the onset of silica-induced pulmonary fibrosis in mice. Int Arch Allergy Immunol 2012;158(3):241–51 [eng].
[39] Webster CM, Hokari M, McManus A, Tang XN, Ma H, Kacimi R, et al. Microglial P2Y12 deficiency/inhibition protects against brain ischemia. PLoS One 2013;8(8):e70927 [eng].
[40] Zhu S, Kuek V, Bennett S, Xu H, Rosen V, Xu J. Protein Cytl1: its role in chondrogenesis, cartilage homeostasis, and disease. Cell Mol Life Sci 2019;76:3515–23. https://doi.org/10.1007/s00018-019-03137-x.