Meta-Analyses of Multiple Gene Expression Profiles to Screen Hub Genes Related to Osteoarthritis

Xianyang Zhu Wen Guo

Department of Orthopedics, Taizhou People's Hospital, Taizhou, China

Keywords

Differentially expressed genes · Osteoarthritis · Meta-analysis · Protein-protein interaction network · Hub gene

Abstract

Background: This study aimed to screen and validate the crucial genes involved in osteoarthritis (OA) and explore its potential molecular mechanisms. **Methods:** Four expression profile datasets related to OA were downloaded from the Gene Expression Omnibus (GEO). The differentially expressed genes (DEGs) from 4 microarray patterns were identified by the meta-analysis method. The weighted gene co-expression network analysis (WGCNA) method was used to investigate stable modules most related to OA. In addition, a protein-protein interaction (PPI) network was built to explore hub genes in OA. Moreover, OA-related genes and pathways were retrieved from Comparative Toxicogenomics Database (CTD). **Results:** A total of 1,136 DEGs were identified from 4 datasets. Based on these DEGs, WGCNA further explored 370 genes included in the 3 OA-related stable modules. A total of 10 hub genes were identified in the PPI network, including *AKT1*, *CDC42*, *HLA-DQA2*, *TUBB*, *TWISTNB*, *GSK3B*, *FZD2*, *KLC1*, *GUSB*, and *RHOG*. Besides, 5 pathways in-

cluding "Lysosome," "Pathways in cancer," "Wnt signaling pathway," "ECM-receptor interaction" and "Focal adhesion" in CTD and enrichment analysis and 5 OA-related hub genes (including *GSK3B*, *CDC42*, *AKT1*, *FZD2*, and *GUSB*) were identified. **Conclusion:** In this study, the meta-analysis was used to screen the central genes associated with OA in a variety of gene expression profiles. Three OA-related modules (green, turquoise, and yellow) containing 370 genes were identified through WGCNA. It was discovered through the gene-pathway network that *GSK3B*, *CDC42*, *AKT1*, *FZD2*, and *GUSB* may be key genes related to the progress of OA and may become promising therapeutic targets.

© 2021 The Author(s).

Published by S. Karger AG, Basel

Introduction

Osteoarthritis (OA) is a chronic degenerative disease that causes joint failure, pain and disability [1]. OA currently affects about 25% of people over the age of 18 and may be the biggest cause of disability among patients over the age of 40 [2, 3]. There are many factors for the development of OA, including abnormal mechanical stress,

aging, obesity and genetic factors [4]. Articular cartilage degeneration, synovial inflammation and atypical bone formation are the main pathological features of OA [5]. Studies have shown that synovial tissue plays an important role in OA [6, 7]. Moreover, synovial lesions have been identified in many joint diseases, which may play a role in promoting the occurrence and development of the disease [8].

At present, more and more evidence indicate that the occurrence and development of OA may be mediated by many genes and signaling pathways [9]. In order to formulate clearer diagnostic criteria and more effective treatment options, the molecular mechanism of OA must be fully understood. With the development of bioinformatics, a lot of researches have been conducted on the gene expression profile of OA [10–12]. Although previous studies have used microarrays to identify some differentially expressed genes (DEGs) [13, 14], there may be some inconsistencies between the studies due to differences in sample size and quality. The meta-analysis method is a systematic method of researching and combining different public datasets, which can improve the accuracy and reliability of data analysis.

Weighted gene co-expression network analysis (WGCNA) is a novel system biology method for molecular interaction mechanism analysis and related network analysis [15–17]. Based on high-throughput microarray data, WGCNA can be used to find modules for co-expression and explore the relationship between gene networks and clinical phenotypes at the transcriptome level [18, 19], thereby providing novel insights for further OA research.

In this study, several microarray profiles of synovial tissue samples from OA patients and normal human synovial tissue samples were downloaded, and feature factors (DEGs) were screened by meta-analysis. WGCNA was further performed on DEGs to identify important modules and genes. Besides, a protein-protein interaction (PPI) network was established, and key genes that may be involved in the OA process were identified.

Materials and Methods

Processing of Microarray Data

The expression profile dataset was retrieved from the Gene Expression Omnibus (GEO) [20] (<http://www.ncbi.nlm.nih.gov/geo/>) with “osteoarthritis” and “*Homo sapiens*” as keywords (downloaded on May 3, 2020). The inclusion criteria were as follows: (i) synovial tissue in patients with OA; (ii) the samples were clearly grouped (OA vs. normal); and (iii) total number of samples ≥ 15 . Finally, a total of 4 microarray datasets were screened, including GSE55457, GSE55235, GSE82107, and GSE12021 (Table 1).

Table 1. Basic information of downloaded microarray data

GEO accession	Platform	Total sample number	OA synovial tissue samples	Normal synovial tissue sample
GSE55457	GPL96	20	10	10
GSE55235	GPL96	20	10	10
GSE82107	GPL570	17	10	7
GSE12021	GPL96	19	10	9

Screening of Feature Genes

In order to screen out more reliable genes related to OA, a meta-analysis was performed on the above 4 microarray datasets. First, MetaQC package [21] was used to control the quality of the dataset based on the following standards: (i) Internal quality control (IQC, Homogeneity test of co-expression structure between studies); (ii) external quality control (EQC, consistency of co-expression pattern with pathway database); (iii) accuracy quality control (AQCg or AQCp, the accuracy of differentially expressed gene or enriched pathway identification); (iv) consistency quality control (CQCg or CQCp, consistency of differentially expressed gene detection or enriched pathway ranking). In addition, principal component analysis (PCA) and the standardized mean rank (SMR) were used to evaluate and filter the microarray datasets. By definition, $0 < \text{SMR} \leq 1$ and large SMR represented a likely problematic study [21].

Then, MetaDE.ES in the MetaDE package [22] was used to screen for significantly DEGs. The MetaDE.ES first performs a heterogeneity test on the expression value of each gene under different sequencing platforms to obtain τ^2 , and Q value. Q value > 0.05 or $\tau^2 = 0$ indicated the expression of genes was homogenous and unbiased. Then, the differential expression of genes between groups was calculated to obtain p value, false discovery rate (FDR) and logFC (fold change). Finally, the DEGs were screened based on the following criteria: (i) the expression of genes was homogenous and unbiased ($\tau^2 = 0$ and Q value > 0.05); (ii) FDR < 0.05 ; and (iii) logFC of all dataset was consistently > 0 or consistently < 0 . Based on DAVID 6.8 [23, 24], enrichment analysis of gene ontology (GO) categories and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were performed on the selected DEGs, and $p < 0.05$ was chosen as the screening threshold.

Screening Disease-Associated Modules Using WGCNA

WGCNA is a bioinformatics algorithm that builds a co-expression network to identify modules associated with disease and screen out important pathogenic mechanisms or potential therapeutic targets [25]. First, the GSE55457 with a relatively large number of samples and clinical information was set as the training set, and the remaining 3 data sets (GSE55235, GSE12021, and GSE82107) were regarded as the validation sets. Then, the WGCNA Version 1.61 package [16] (<https://cran.r-project.org/web/packages/WGCNA/index.html>) in the R3.4.1 language was used to screen for significant stable modules related to disease across datasets. Based on clustering and dynamic pruning, highly related genes were clustered into modules. The threshold was set to the number of RNAs in the module set ≥ 30 and cut height = 0.99. Finally, the WGCNA modules related to OA were determined.

Table 2. Results of QC measures and SMRs

Study	IQC	EQC	CQCg	CQCp	AQCg	AQCp	SMR
GSE55457	6.444	6.273	123.190	195.106	201.878	74.069	3.176
GSE55235	5.984	5.538	89.436	192.364	139.358	38.622	4.339
GSE82107	5.748	4.147	84.174	192.272	34.725	126.956	3.176
GSE12021	6.312	4.239	77.756	225.054	33.465	110.723	3.753

SMR, standardized mean rank; QC, quality control.

PPI Network Construction

The STRING database (Version:10.0, <http://string-db.org/>) was used to search for the interactions between DEGs encoded proteins and those with interaction scores >0.6 were selected to construct a PPI network. The network was visualized through Cytoscape Version 3.6.1 (<http://www.cytoscape.org/>). Based on DAVID 6.8, GO functional annotation and KEGG pathway enrichment analysis were performed on DEGs-encoded proteins. $p < 0.05$ was selected as the significance threshold.

Then, the topology structure of the PPI network was analyzed. By calculating the topological structure parameters of Degree (the connection with other genes), betweenness centrality (BC), closeness centrality (CC), and Path length in the PPI network, the hub genes that play important roles in network connection were screened. Then, the descending order was sorted according to degree.

The Gene-Pathway Network Directly Correlated with OA Construction

Using “osteoarthritis” as a keyword, the KEGG pathways directly related to OA were searched from the Comparative Toxicogenomics Database Update 2019 database [26] (<http://ctd.mdibl.org/>). Then, the retrieved pathways were compared with pathways in which genes in the PPI network participate to obtain overlapping pathways. Genes involved in overlapping pathways were extracted from the PPI network to construct a gene-pathway network related to OA.

Results

Feature Gene Selection

A total of 11,812 common genes were detected on multiple platforms from the GEO. Table 2 shows the QC results for all 4 microarray datasets. The SMRs of the 4 datasets were very close, indicating that the datasets had low bias and high consistency. Further, the two-dimensional graph of PCA showed that the cumulative contribution rate of 1 and 2 principal components was >80%, and the 4 microarray patterns were all located in the parameter group, with little difference in position (Fig. 1a). These results indicated the 4 datasets were of good quality and were comparable.

Using the MetaDE software package, a total of 1,136 DEGs were identified from 4 microarray patterns. Subsequently, based on the expression values of DEGs in each data set, and bidirectional hierarchical clustering analysis was performed. The clustering analysis showed that DEGs could clearly distinguish the OA samples from the control samples (Fig. 1b), indicating that the expression of DEGs were significantly different. Besides, it can also be found from Figure 1b that the gene expression patterns of DEGs were relatively consistent in the 4 datasets.

Functional Analysis of Feature Genes

GO functional and KEGG pathway enrichment analysis was performed on DEGs. As a result, 109 significantly related GO annotations and 12 KEGG pathways were obtained. The results showed that the DEGs were mainly involved in the biological processes of “positive regulation of cell proliferation” ($p = 6.420\text{E-}06$), “regulation of cell motion” ($p = 6.460\text{E-}06$), and “regulation of locomotion” ($p = 1.610\text{E-}05$) (Fig. 2a). The most significant enrichment pathways were “pathways in cancer” ($p = 1.790\text{E-}02$), “adipocytokine signaling pathway” ($p = 3.313\text{E-}03$), and “lysosome” ($p = 1.598\text{E-}03$) (Fig. 2b).

Identification of Modules Related to OA Using WGCNA

In order to confirm whether the genes in each dataset were comparable, the gene expression levels and connectivity patterns in any 2 datasets were compared. As shown in Figure 3, the gene expression levels and connectivity patterns between each pair of datasets have a significant positive correlation ($p < 0.05$). Therefore, the data were comparable.

The dataset GSE55457 with relatively large number of samples and clinical information was used as the training set, and the remaining 3 datasets (GSE55235, GSE12021, and GSE82107) were used as the validation set. Then, the value of adjacency matrix weighting (power) was explored by using the samples in the training set of

GSE55457. The relative balance of scale independence and average connectivity of WGCNA was determined. As shown in Figure 4, power = 7 was selected when the square value of the correlation coefficient reaches 0.9 for the first time. Next, the dissimilarity coefficients of DEGs were calculated and a systematic clustering tree was obtained. According to the standard of mixed dynamic shearing trees, the least gene number of each gene network was set to 30 and the cut-height was set to 0.99. As shown in the clustering dendrograms of Figure 5a, a total of 10 modules related to OA were finally obtained in the training set. In order to assess the robustness of the modules identified in the training set, the co-expression module generated in the training dataset were reproduced in

the verification datasets (Fig. 5b–d) and the module preservation was among the 4 datasets calculated by summary preservation statistics in WGCNA. The Z-summary statistics of the 3 modules (green, turquoise, and yellow) were all >5 and $p \leq 0.05$, indicating these modules were relatively stable (Table 3). This indicates that the 3 modules have been saved between the training data set and the verification data set.

To explore the clinical significance of the module, the correlation between module eigengenes (MEs) and gender, age, and status was analyzed. Figure 6 shows the heat map of the correlation between MEs and the disease factors of OA. The results showed that the gray (correlation = -0.87 , $p < 0.001$), pink (correlation = -0.49 , $p = 1E-70$),

Fig. 1. a Quality control results of the merged datasets from 4 microarray profiles (marked as 1–5) obtained via MetaQC analysis. The first principal component is presented on the x -axis, while the second principal component is shown on the y -axis. QC, quality control; IQC, internal QC; EQC, external QC; AQCg, accuracy

QC; AQCp, precision of AQCg; CQCg, consistency QC; CQCp, precision of CQCg. **b** Heat map of 1,136 significantly differentially expressed genes from all samples. Red indicates higher gene expression and blue indicates lower gene expression.

Fig. 2. The BP, CC, and MF terms, and KEGG pathways enriched by the DEGs related to OA. **a** Bubble chart showing the top 10 significant enrichment biological processes, cellular component, and molecular function analysis of DEGs. Circle, triangle, and square represent BP, CC, and MF, respectively. **b** Bubble chart showing the KEGG pathways of DEGs. The x -axis shows support/reference

numbers and FDR value of DEGs; the y -axis shows different molecular functions. BP, biological process; CC, cellular component; MF, molecular function; KEGG, Kyoto Encyclopedia of Genes and Genomes; DEG, differentially expressed gene; FDR, false discovery rate.

(For figure see next page.)

Fig. 3. a Scatter plot of gene expression correlation in any 2 datasets. **b** Scatter plot of connectivity correlation in any 2 datasets.

Fig. 4. Calculation chart of adjacency matrix weighting parameters (power). The x -axis represents weighting parameters (power). The y -axis represents quadratic of correlation index from $\log(k)$ and $\log(P[K])$.

red (correlation = -0.64 , $p = 4E-132$), brown (correlation = -0.52 , $p = 3E-79$), and turquoise (correlation = -0.7 , $p = 8E-171$) modules were negatively correlated with disease status, while magenta (correlation = 0.63 , $p = 6E-127$), blue (correlation = 0.62 , $p = 6E-123$), green (correlation = 0.7 , $p = 6E-165$), black (correlation = 0.59 , $p = 3E-109$), and yellow (correlation = 0.6 , $p = 2E-112$) were positively correlated with disease status. These 3 modules, which contained a total of 370 genes, were selected as clinically important modules for further analysis.

Generation and Annotation of OA-Associated Module Network

The STRING database was used to search the interactions among the DEG-encoded proteins contained in the 3 modules, and proteins with an interaction score >0.6 were selected to construct a PPI network (Fig. 7). In addition, the top 10 hub genes were identified according to ascending order of Degree, including *AKT1*, *CDC42*, *HLA-DQA2*, *TUBB*, *TWISTNB*, *GSK3B*, *FZD2*, *KLC1*, *GUSB*, and *RHOG* (Table 4).

Table 3. Module preservation statistics table

ID	Color	Module size	Preservation information	
			Z-score	p value
Module 1	Black	64	4.2160	$9.60E-02$
Module 2	Blue	163	4.5112	$3.00E-02$
Module 3	Brown	150	3.9468	$1.90E-04$
Module 4	Green	87	5.8471	$4.10E-02$
Module 5	Gray	238	13.2585	$9.10E-62$
Module 6	Magenta	30	1.2542	$2.00E-01$
Module 7	Pink	36	2.8457	$7.50E-02$
Module 8	Red	85	3.8927	$3.20E-01$
Module 9	Turquoise	182	5.2342	$3.10E-02$
Module 10	Yellow	101	7.3072	$1.60E-02$

All these proteins have undergone enrichment analysis of GO function and KEGG pathway, and a total of 48 significantly related GO annotations and 6 KEGG pathways were obtained. Figure 8a shows the top 10 GO function entries sorted in ascending order of FDR value. The

Fig. 5. Clustering dendrograms of GSE55457 (a), GSE55235 (b), GSE12021 (c), and GSE82107 (d). Ten kinds of color present 10 modules.

results showed that DEGs were mainly involved in the biological processes of “wnt receptor signaling pathway” ($p = 5.160\text{E-}06$), “protein amino acid phosphorylation” ($p = 1.305\text{E-}03$) and “cell surface receptor linked signal transduction” ($p = 2.831\text{E-}03$). As for the cellular component, the DEGs were mainly enriched in “vacuole” ($p = 9.500\text{E-}04$), “endoplasmicreticulum” ($p = 1.147\text{E-}03$) and “endomembrane system” ($p = 2.705\text{E-}03$). Regard-

ing molecular function, the DEGs were mainly enriched in “purine ribonucleotide binding” ($p = 2.100\text{E-}04$), “ribonucleotide binding” ($p = 2.100\text{E-}04$), and “purine nucleotide binding” ($p = 5.150\text{E-}04$). Moreover, KEGG pathway enrichment analysis suggests that “lysosome” ($p = 7.930\text{E-}04$), “glycosaminoglycan degradation” ($p = 5.149\text{E-}03$), and “pathways in cancer” ($p = 9.377\text{E-}03$) were the most critical pathway in OA (Fig. 8b).

Fig. 6. Module-trait associations were evaluated by correlations between MEs and clinical traits. Each row corresponds to an ME, column to a trait. Each cell contains the corresponding correlation (first line) and *p* value (second line). The table is color-coded by correlation according to the color legend. The change of color from blue to red indicates the process of changing from negative to positive correlation with characterization. ME, module eigengene.

Fig. 7. The PPI networks of the top 10 downregulated DEGs. All the nodes are proteins encoded by DEGs and the blue borders represent proteins encoded by top 10 downregulated DEGs; purple

borders represent proteins encoded by other genes interacted with proteins encoded by these DEGs. PPI, protein-protein interaction; DEG, differentially expressed gene.

Fig. 8. The GO annotation node (a) and the KEGG signal path histogram (b) are significantly related to genes in the interaction network. The horizontal axis represents the number of genes, and the vertical axis represents the name of the entry. GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.

The Gene-Pathway Network Directly Correlated with OA

A total of 171 KEGG pathways directly related to OA were obtained from the CTD database. A total of 5 overlapping pathways including “lysosome,” “pathways in

cancer,” “Wnt signaling pathway,” “ECM-receptor interaction,” and “focal adhesion” were obtained by comparing the retrieved pathways with those enriched from the PPI network. Based on overlapping pathways, an interaction network directly related to OA was extracted (Fig. 9).

Fig. 9. The interaction network directly correlated with OA. Circles represent genes. Squares represent the pathways correlated with OA. Red and green separately represent up-regulation and down-regulation. OA, osteoarthritis; FC, fold change.

Table 4. Topological parameter information of the top 10 gene nodes in the interaction network

ID	Average shortest path length	Betweenness centrality	Closeness centrality	Degree
AKT1	3.423	0.561	0.292	18
CDC42	3.648	0.282	0.274	16
HLA-DQA2	4.007	0.206	0.250	10
TUBB	5.648	0.042	0.177	8
TWISTNB	4.592	0.175	0.218	8
GSK3B	3.958	0.212	0.253	8
FZD2	4.866	0.033	0.205	8
KLC1	4.092	0.048	0.244	7
GUSB	5.641	0.055	0.177	7
RHOG	3.873	0.042	0.258	7

In addition, 5 hub genes related to OA (*GSK3B*, *CDC42*, *AKT1*, *FZD2*, and *GUSB*) were also included in this network. Among them, the *GSK3B* was simultaneously involved in multiple pathways, suggesting its crucial role in the network as well as the development of OA.

Discussion

In this study, a total of 1,136 DEGs were identified from 4 databases through MetaDE. Based on these DEGs, 3 OA-related modules (green, turquoise, and yellow)

containing 370 genes were identified by WGCNA. The STRING database was used to search the interactions among the DEG-encoded proteins contained in the 3 modules, and proteins with an interaction score >0.6 were selected to construct a PPI network. Ten hub genes were detected in the PPI network, including *AKT1*, *CDC42*, *HLA-DQA2*, *TUBB*, *TWISTNB*, *GSK3B*, *FZD2*, *KLC1*, *GUSB*, and *RHOG*. The DEGs in the PPI network were enriched with 48 significantly related GO annotations and 6 KEGG pathways. Five pathways in the gene-pathway network, including “lysosome,” “pathways in cancer,” “Wnt signaling pathway,” “ECM-receptor interaction,” and “focal adhesion,” comprised 5 hub genes (*GSK3B*, *CDC42*, *AKT1*, *FZD2*, and *GUSB*) that were related to OA.

There are 5 top 10 hub genes in the gene-pathway network, namely *GSK3B*, *CDC42*, *AKT1*, *FZD2*, and *GUSB*. *GSK3B* is an essential regulator of articular cartilage destruction and plays multiple functions in the pathological process of osteoarthritis [27]. Studies have shown that *GSK3B* can regulate various cell activities including cell growth, cell survival, metabolism, and apoptosis gene expression [28, 29]. Wu et al. [30] showed that inhibition of *GSK3B* might prevent cartilage formation, and our results also indicate that *GSK3B* is down-regulated in OA. Therefore, *GSK3B* may play an important role in the cartilage formation of OA. *CDC42* belongs to the small GTP-binding protein in the Ras superfamily [31]. Hu et al. [32] found that the genetic destruction of *CDC42*, knocking

down *CDC42* expression or inhibiting *CDC42* activity can significantly restore the increase in the number of mesenchymal stem cells, osteoprogenitor cells, osteoblasts, osteoclasts, and new blood vessels forming. In this study, *CDC42* was significantly down-regulated, which indicates *CDC42* may play a role in the growth and development of OA cells.

AKT1 is a serine/threonine-specific protein kinase that mainly controls cell growth and survival [33]. In human cells, *AKT1* activation occurs in response to growth factor stimulation. *AKT1* can promote the cartilage calcification by inhibiting the accumulation of inorganic pyrophosphate in chondrocytes, thereby controlling the bone growth and osteophyte formation in OA [34]. *FZD2* is a member of the Fzd receptor family of cell membranes on the Wnt pathway [35], which is involved in the activation and repair of skeletal muscle satellite cells [36, 37]. *GUSB* is a lysosomal enzyme that is important in the normal gradual degradation of glycosaminoglycans. Studies have shown that *GUSB* deficiency causes lysosomal storage disease mucopolysaccharide storage disease VII (MPS VII), and MPS VII can cause mental retardation, hepatomegaly and splenomegaly, and skeletal dysplasia [27]. *AKT1*, *FZD2*, and *GUSB* were all upregulated in OA, which may play an important role in the occurrence and development of OA.

The Wnt signaling pathway is a key regulator of endochondral ossification, which is the process of endochondral formation. The Wnt signaling pathway is normally inhibited, but it will promote the development of OA when it is activated [38]. In addition, activation of the Wnt signaling pathway leads to increased chondrocyte catabolism, hypertrophy-like changes, and cartilage degradation [39]. Therefore, Wnt signal pathways may provide a new key to control the cartilage formation of OA to understand the conduction mechanism. Focal adhesion is a transmembrane protein complex, which can maintain multiple links with intracellular organelles through the cytoskeleton, and can attach chondrocytes to the surrounding extracellular matrix at the same time [40]. Studies have shown that the ECM-receptor interaction pathway and focal adhesion pathway are involved in the pathogenesis of OA [41]. Jang et al. [40] found that most acute chondrocyte death occurs in a focal adhesion kinase-dependent and Src family kinase-dependent manner, so inhibition of adhesion plaques can prevent cartilage cell death. Therefore, ECM-receptor interaction and focal adhesion pathway may help to inhibit chondrocyte death in OA. In addition, in this study, 4 top 10 hub genes (*GSK3B*, *CDC42*, *AKT1*, and *FZD2*) were involved in

“pathways in cancer,” which indicates that the progression of OA is significantly related to the “pathways in cancer.” The limitation of this study is that the selected genes and pathways have not been experimentally verified, which may become the focus of future research.

Conclusion

In this study, the meta-analysis was used to screen the central genes associated with OA in a variety of gene expression profiles. Three OA-related modules (green, turquoise, and yellow) containing 370 genes were identified through WGCNA. In addition, 10 hub genes were identified in the PPI network, namely *AKT1*, *CDC42*, *HLA-DQA2*, *TUBB*, *TWISTNB*, *GSK3B*, *FZD2*, *KLC1*, *GUSB*, and *RHOG*. Finally, it was discovered through the gene-pathway network that *GSK3B*, *CDC42*, *AKT1*, *FZD2*, and *GUSB* may be key genes related to the progression of OA and may become promising therapeutic targets.

Acknowledgements

We are very grateful to the individuals who participated in this study.

Statement of Ethics

Not applicable because this study did not involve any in vitro and in vivo experiments.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Funding Sources

The authors did not receive any funding.

Author Contributions

X.Z. performed the literature review, created the figures, and wrote the review article; W.G. contributed content expertise and performed critical revisions necessary for the final version of the review article.

References

- Barnett R. Osteoarthritis. *Lancet*. 2018 May 19;391(10134):1985.
- Thomas E, Peat G, Croft P. Defining and mapping the person with osteoarthritis for population studies and public health. *Rheumatology*. 2014 Feb;53(2):338–45.
- Sacitharan PK. Ageing and osteoarthritis. *Subcell Biochem*. 2019;91:123–59.
- Bijlsma JW, Berenbaum F, Lafeber FP. Osteoarthritis: an update with relevance for clinical practice. *Lancet*. 2011 Jun 18;377(9783):2115–26.
- Sulzbacher I. Osteoarthritis: histology and pathogenesis. *Wien Med Wochenschr*. 2013 May;163(9–10):212–9.
- Kleine SA, Budberg SC. Synovial membrane receptors as therapeutic targets: a review of receptor localization, structure, and function. *J Orthop Res*. 2017 Aug;35(8):1589–605.
- Wang X, Hunter DJ, Jin X, Ding C. The importance of synovial inflammation in osteoarthritis: current evidence from imaging assessments and clinical trials. *Osteoarthr Cartil*. 2018 Feb;26(2):165–74.
- Scanzello CR, Goldring SR. The role of synovitis in osteoarthritis pathogenesis. *Bone*. 2012 Aug;51(2):249–57.
- Chen D, Shen J, Zhao W, Wang T, Han L, Hamilton JL, et al. Osteoarthritis: toward a comprehensive understanding of pathological mechanism. *Bone Res*. 2017;5:16044.
- Li Z, Wang Q, Chen G, Li X, Yang Q, Du Z, et al. Integration of gene expression profile data to screen and verify hub genes involved in osteoarthritis. *Biomed Res Int*. 2018;2018:9482726.
- Lin J, Wu G, Zhao Z, Huang Y, Chen J, Fu C, et al. Bioinformatics analysis to identify key genes and pathways influencing synovial inflammation in osteoarthritis. *Mol Med Rep*. 2018 Dec;18(6):5594–602.
- Zhu N, Hou J, Wu Y, Li G, Liu J, Ma G, et al. Identification of key genes in rheumatoid arthritis and osteoarthritis based on bioinformatics analysis. *Medicine*. 2018 Jun;97(22):e10997.
- Li M, Zhi L, Zhang Z, Bian W, Qiu Y. Identification of potential target genes associated with the pathogenesis of osteoarthritis using microarray based analysis. *Mol Med Rep*. 2017 Sep;16(3):2799–806.
- Li HZ, Lu HD. Transcriptome analyses identify key genes and potential mechanisms in a rat model of osteoarthritis. *J Orthop Surg Res*. 2018 Dec 14;13(1):319.
- Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008 Dec 29;9:559–9.
- Presson AP, Sobel EM, Papp JC, Suarez CJ, Whistler T, Rajeevan MS, et al. Integrated weighted gene co-expression network analysis with an application to chronic fatigue syndrome. *BMC Syst Biol*. 2008 Nov 6;2:95.
- Zhou Q, Su X, Jing G, Ning K. Meta-QC-Chain: comprehensive and fast quality control method for metagenomic data. *Genom Proteom Bioinf*. 2014 Feb;12(1):52–6.
- Udyavar AR, Hoeksema MD, Clark JE, Zou Y, Tang Z, Li Z, et al. Co-expression network analysis identifies Spleen Tyrosine Kinase (SYK) as a candidate oncogenic driver in a subset of small-cell lung cancer. *BMC Syst Biol*. 2013;7(Suppl 5):S1.
- Pei G, Chen L, Zhang W. WGCNA application to proteomic and metabolomic data analysis. *Methods Enzymol*. 2017;585:135–58.
- Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: archive for functional genomics data sets: update. *Nucleic Acids Res*. 2013 Jan;41(Database issue):D991–5.
- Kang DD, Sibille E, Kaminski N, Tseng GC. MetaQC: objective quality control and inclusion/exclusion criteria for genomic meta-analysis. *Nucleic Acids Res*. 2012 Jan;40(2):e15.
- Wang X, Kang DD, Shen K, Song C, Lu S, Chang LC, et al. An R package suite for microarray meta-analysis in quality control, differentially expressed gene analysis and pathway enrichment detection. *Bioinformatics*. 2012 Oct 1;28(19):2534–6.
- Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009;4(1):44–57.
- Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res*. 2009 Jan;37(1):1–13.
- Chen R, Mias GI, Li-Pook-Than J, Jiang L, Lam HY, Chen R, et al. Personal omics profiling reveals dynamic molecular and medical phenotypes. *Cell*. 2012 Mar 16;148(6):1293–307.
- Davis AP, Grondin CJ, Johnson RJ, Sciaky D, McMorran R, Wiegiers J, et al. The comparative toxicogenomics database: update 2019. *Nucleic Acids Res*. 2019 Jan 8;47(D1):D948–54.
- Vogler C, Galvin N, Levy B, Grubb J, Jiang J, Zhou XY, et al. Transgene produces massive overexpression of human beta-glucuronidase in mice, lysosomal storage of enzyme, and strain-dependent tumors. *Proc Natl Acad Sci U S A*. 2003 Mar 4;100(5):2669–73.
- Sanchez JF, Sniderhan LF, Williamson AL, Fan S, Chakraborty-Sett S, Maggirwar SB. Glycogen synthase kinase 3beta-mediated apoptosis of primary cortical astrocytes involves inhibition of nuclear factor kappaB signaling. *Mol Cell Biol*. 2003 Jul;23(13):4649–62.
- Martin M, Rehani K, Jope RS, Michalek SM. Toll-like receptor-mediated cytokine production is differentially regulated by glycogen synthase kinase 3. *Nat Immunol*. 2005 Aug;6(8):777–84.
- Wu Q, Huang JH, Sampson ER, Kim KO, Zuscik MJ, O'Keefe RJ, et al. Smurf2 induces degradation of GSK-3beta and upregulates beta-catenin in chondrocytes: a potential mechanism for Smurf2-induced degeneration of articular cartilage. *Exp Cell Res*. 2009 Aug 15;315(14):2386–98.
- Woods A, Wang G, Beier F. Regulation of chondrocyte differentiation by the actin cytoskeleton and adhesive interactions. *J Cell Physiol*. 2007 Oct;213(1):1–8.
- Hu X, Ji X, Yang M, Fan S, Wang J, Lu M, et al. Cdc42 is essential for both articular cartilage degeneration and subchondral bone deterioration in experimental osteoarthritis. *J Bone Miner Res*. 2018 May;33(5):945–58.
- Manning BD, Tokar A. AKT/PKB signaling: navigating the network. *Cell*. 2017 Apr 20;169(3):381–405.
- Fukai A, Kawamura N, Saito T, Oshima Y, Ikeda T, Kugimiya F, et al. Akt1 in murine chondrocytes controls cartilage calcification during endochondral ossification under physiologic and pathologic conditions. *Arthritis Rheum*. 2010 Mar;62(3):826–36.
- Dijksterhuis JP, Baljinnayam B, Stanger K, Serican HO, Ji Y, Andres O, et al. Systematic mapping of WNT-FZD protein interactions reveals functional selectivity by distinct WNT-FZD pairs. *J Biol Chem*. 2015 Mar 13;290(11):6789–98.
- Whyte JL, Smith AA, Helms JA. Wnt signaling and injury repair. *Cold Spring Harb Perspect Biol*. 2012 Aug 1;4(8):a008078.
- Fujimaki S, Hidaka R, Asashima M, Takemasa T, Kuwabara T. Wnt protein-mediated satellite cell conversion in adult and aged mice following voluntary wheel running. *J Biol Chem*. 2014 Mar 14;289(11):7399–412.
- Corr M. Wnt-beta-catenin signaling in the pathogenesis of osteoarthritis. *Nat Clin Pract Rheumatol*. 2008 Oct;4(10):550–6.
- Burr DB, Gallant MA. Bone remodelling in osteoarthritis. *Nat Rev Rheumatol*. 2012 Nov;8(11):665–73.
- Jang KW, Buckwalter JA, Martin JA. Inhibition of cell-matrix adhesions prevents cartilage chondrocyte death following impact injury. *J Orthop Res*. 2014 Mar;32(3):448–54.
- Wang W, Liu Y, Hao J, Zheng S, Wen Y, Xiao X, et al. Comparative analysis of gene expression profiles of hip articular cartilage between non-traumatic necrosis and osteoarthritis. *Gene*. 2016 Oct 10;591(1):43–7.