Comprehensive Analysis of Key Genes, Signaling Pathways and miRNAs in Human Knee Osteoarthritis: Based on Bioinformatics

Liang Chang1,2, Hao Yao1,2, Zhi Yao1,2, Kevin Ki-Wai Ho1, Michael Tim-Yun Ong1, Bingyang Dai1, Wenxue Tong1, Jiankun Xu1,2* and Ling Qin1,2*

1Musculoskeletal Research Laboratory, Department of Orthopedics and Traumatology, The Chinese University of Hong Kong, Hong Kong, Hong Kong, SAR China, 2Innovative Orthopaedic Biomaterial and Drug Translational Research Laboratory, Li Ka Shing Institute of Health Sciences, The Chinese University of Hong Kong, Hong Kong, Hong Kong, SAR China

Background: Osteoarthritis (OA) is one of the main causes of disability in the elderly population, accompanied by a series of underlying pathologic changes, such as cartilage degradation, synovitis, subchondral bone sclerosis, and meniscus injury. The present study aimed to identify key genes, signaling pathways, and miRNAs in knee OA associated with the entire joint components, and to explain the potential mechanisms using computational analysis.

Methods: The differentially expressed genes (DEGs) in cartilage, synovium, subchondral bone, and meniscus were identified using the Gene Expression Omnibus 2R (GEO2R) analysis based on dataset from GSE43923, GSE12021, GSE98918, and GSE51588, respectively and visualized in Volcano Plot. Venn diagram analyses were performed to identify the overlapping DEGs (overlapping DEGs) that expressed in at least two types of tissues mentioned above. Gene Ontology (GO) enrichment analysis, Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis, protein-protein interaction (PPI) analysis, and module analysis were conducted. Furthermore, qRT-PCR was performed to validate above results using our clinical specimens.

Results: As a result, a total of 236 overlapping DEGs were identified, of which 160 were upregulated and 76 were downregulated. Through enrichment analysis and constructing the PPI network and miRNA-mRNA network, knee OA-related key genes, such as HEY1, AHR, VEGFA, MYC, and CXCL12 were identified. Clinical validation by qRT-PCR experiments further supported above computational results. In addition, knee OA-related key miRNAs such as miR-101, miR-181a, miR-29, miR-9, and miR-221, and pathways such as Wnt signaling, HIF-1 signaling, PI3K-Akt signaling, and axon guidance pathways were also identified. Among above identified knee OA-related key genes, pathways and miRNAs, genes such as AHR, HEY1, MYC, GAP43, and PTN, pathways like axon guidance, and miRNAs such as miR-17, miR-21, miR-155, miR-185, and miR-1 are lack of research and worthy for future investigation.
**INTRODUCTION**

Osteoarthritis (OA) is the most common joint disease, mainly manifesting as pain, limited joint movement, and joint deformity. The risk factors of OA include trauma, aging, obesity, and heredity (Sandell, 2012; Chen et al., 2017a; Chang et al., 2021). During OA development, the entire joint are affected and undergo articular cartilage degeneration, osteophyte formation, subchondral sclerosis, synovitis, and meniscus degeneration, respectively, indicating the complicated and interactive OA pathogenic mechanisms (Chen et al., 2017a). The efficacies of the current treatments for OA in our clinics are limited. In recent years, exploration of disease-modifying osteoarthrits drugs (DMOADs) aiming at alleviating OA symptoms and/or prevent structural progression have drawn much attention. However, the DMOADs under research and development (R&D) and/or clinical trials mainly focus on one of the OA symptoms, such as cartilage degeneration, subchondral bone remodeling, local inflammation, or joint pain, and their potential downstream targets (Latourte et al., 2020). The possibility that newly explored-drug targets have heterogeneous expression profiles in different joint components raises uncertainty of the drug effectiveness. Besides, the R&D of joint component-specific drugs are also limited so far (Latourte et al., 2020).

Based on the rapid development of high-throughput genomics technologies, such as microarray and next-generation sequencing, bioinformatic analysis has been widely used to identify key genes, signaling pathways, and microRNAs (miRNAs) in various musculoskeletal disorders (Kang et al., 2021; Li et al., 2021; Umeno et al., 2021). So far, many studies have identified and explored potential therapeutic targets of OA based on bioinformatic screening. For example, upregulated arginase 2 (ARG2) in OA cartilage was screened out by microarray and further validated to facilitate cartilage destruction via upregulating matrix metalloproteinases (MMPs) (Choi et al., 2019). Activated osteochondral turnover, neurogenesis and inflammation in OA bone marrow lesions (BML) were also identified by microarray bioinformatically (Kuttapitiya et al., 2017). Besides, miRNA candidates that have potential as biomarkers and therapeutic targets in OA were identified and validated via comprehensively paired miRNA-messenger RNA (mRNA) analysis and functional enrichment analysis (Kung et al., 2018). In addition to identification of potential therapeutic targets, bioinformatics-based bulk sample analysis further helps classify potential OA subtypes for more precise diagnosis and personalized treatments. In recently years, several studies have stepped forward substantially in classifying potential OA subtypes based on bioinformatics (Soul et al., 2018; Yuan et al., 2020). Their surprising discoveries undoubtedly deepen our understandings on knee OA and will facilitate personalized treatments in the future. However, since previous bioinformatic studies mainly focus on one type of joint components as well, investigations on the overlapping DEGs (overlapping DEGs) in different joint components during OA development are still lacking. In 2016, about 5% overlapping DEGs were observed between the DEGs of the synovium and cartilage, while no further analysis was performed on these identified overlapping DEGs (Park and Ji, 2016). Recently, another study observed about 10% overlapping DEGs in OA cartilage and subchondral bone. They identified *IL11* and *CHADL* as two potential therapeutic targets of OA by comparing their identified cartilage-subchondral bone overlapping DEGs with previously identified OA risk genes (Styrkarsdottir et al., 2018; Tuerlings et al., 2021).

Collectively, we believe it is meaningful to comprehensively analyze the key genes that are differentially expressed during OA development in the different joint components, including articular cartilage, subchondral bone, synovium, and meniscus, and their relevant pathways and miRNAs. Such approaches may provide clues to develop adequate treatments for OA by targeting at overlapping differentially expressed genes (DEGs) in different joint components and their relevant miRNAs and signaling pathways. The present study aims at identifying key genes, signaling pathways, and miRNAs in human knee OA by comparing the preexisting gene expression profiles derived from different joint components, including articular cartilage, synovium, subchondral bone, and meniscus. Specifically, the gene expression profiles (GSE) were obtained from the public available Gene Expression Omnibus database (GEO, http://www.ncbi.nlm.nih.gov/geo/). Gene Expression Omnibus 2R (GEO2R) was performed to identify the overlapping DEGs and followed by qRT-PCR validation. Furthermore, functional enrichment analysis, protein-protein interaction (PPI) analysis, and miRNA-mRNA interaction analysis were carried out to identify relevant signaling pathways and interactive miRNAs. This study may shed light on completer and undiscovered pathogenic mechanisms of knee OA development and pave the way toward the identification of new therapeutic targets for further R&D of effective therapies and clinical translation.

**MATERIALS AND METHODS**

**Gene Expression Profiles in Human Knee OA joint Tissues**

The gene expression profiling in cartilage, synovial membrane, subchondral bone, and meniscus tissues was obtained from GEO...
datasets GSE43923 (Klinger et al., 2013), GSE12021 (Huber et al., 2008), GSE51588 (Chou et al., 2013), and GSE98918 (Brophy et al., 2018), respectively. Three degenerated and three intact cartilage samples were retrieved from a human dataset using the Affymetrix Human Genome U133 Plus 2.0 Array platform (GSE43923). Nine normal and ten OA synovial tissue samples were retrieved from a human dataset using the Affymetrix Human Genome U133 Array platform (GSE12021). Twenty OA and five normal medial tibial subchondral bone samples were retrieved from a human study using the Agilent-026652 Whole Human Genome Microarray 4×44K v2 platform (GSE51588). Twelve OA and twelve normal meniscus samples were retrieved from a human dataset using the Agilent-072363 SurePrint G3 Human GE v3 8×60K Microarray 039494 platform (GSE98918).

**Identifying DEGs**

The original gene expression profiles were analyzed by GEO2R (GEO2R, RRID:SCR_016569; https://www.ncbi.nlm.nih.gov/geo/geo2r/?acc=GSE43923, https://www.ncbi.nlm.nih.gov/geo/geo2r/?acc=GSE12021, https://www.ncbi.nlm.nih.gov/geo/geo2r/?acc=GSE51588, https://www.ncbi.nlm.nih.gov/geo/geo2r/?acc=GSE98918) to identify the upregulated and downregulated DEGs in OA joint tissues, respectively. The criteria for a DEG were log2FC>1 and adjusted P-value<0.05. The results were visualized in volcano plots.

**Identification of Overlapping DEGs in Human KOA Joint Tissues**

Venn diagram (http://bioinformatics.psb.ugent.be/webtools/Venn/; VennDiagram, RRID: SCR_002414) was used to identify upregulated and downregulated overlapping DEGs in the integral joint tissues including cartilage, synovial membrane, subchondral bone, and meniscus. A specific DEG was identified as overlapping DEG when it appeared at least in two of the joint tissues. All the DEGs were identified by comparing gene expression profiles between osteoarthritic and relatively healthy joint tissues.

**Gene Ontology Enrichment and Kyoto Encyclopedia of Genes and Genomes Pathway Analysis**

GO enrichment analysis and KEGG pathway analysis were performed on Metascape platform (https://metascape.org/gp/index.html#main/step1; Metascape, RRID: SCR_016620) (Zhou et al., 2019). Upregulated or downregulated overlapping DEGs were listed and followed by “Custom Analysis.” GO enrichment analysis and KEGG pathway analysis were performed with the thresholds of P-value<0.05 and enrichment gene count ≥2.

**Construction of the Protein-Protein Interaction Network**

The Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database (https://string-db.org/; STRING, RRID: SCR_005223) was used to construct the PPI network (Szklarczyk et al., 2021). The overlapping DEGs were mapped to STRING list to perform multiple proteins search and get a PPI network with interaction scores >0.4. Cytoscape V.3.7.2 (Cytoscape, RRID: SCR_003032) was used to visualize the results from the PPI network and perform module analysis. Genes with connectivity degree ≥10 were identified as hub genes (Shannon et al., 2003).

**Module Analysis**

Module analysis was performed using the molecular complex detection (MCODE) plugin on Cytoscape platform (MCODE, RRID: SCR_015828; Cytoscape V.3.7.2, RRID: SCR_003032). The parameters set to identify enriched functional modules were as follows: Degree Cutoff = 2, Node Score Cutoff = 0.2, K-Core = 2 and Maxium. Depth = 100. Modules with the MCODE score ≥4 were identified as significant modules and were further evaluated for GO enrichment analysis and KEGG pathway analysis with the thresholds of P-value<0.05 and enrichment gene count >2.

**Construction of the miRNA-mRNA Network**

Experimentally validated key gene-related miRNAs were screened out based on key genes identified above by using miRTarBase 8.0 (http://miRTarBase.cuhk.edu.cn/; miRTarBase, RRID: SCR_017355) with strong evidence (Huang et al., 2020). Those miRNAs targeting at least two key genes were identified as key miRNAs and visualized on Cytoscape V.3.7.2 by constructing the miRNA-mRNA network.

**Clinical Specimens Sampling**

Clinical specimens of preserved and degenerated cartilage were harvested from osteoarthritis patients undergoing total knee arthroplasty (TKA) surgery. The included patients had no history of chronic diseases, tumors, autoimmune diseases, and viral chronic infections (hepatitis B virus, hepatitis C virus, human immunodeficiency virus). All patients provided informed consent, and this study was approved by the Joint Chinese University of Hong Kong-New Territories East Cluster Clinical Research Ethics Committee (CREC Ref. No: 2013.248). The estimated sample size equaled to three based on a pilot study. A total of three donors (Age: 70.17±3.66; gender: one male and two females; K-L: grade III) were included. The cartilage extracted from hypertrophic and the severely destructed region was classified into degenerated cartilage (DC) group, and the cartilage extracted from the relatively smooth region was classified into preserved cartilage (PC) group. About 0.5–1 g DC and PC samples were collected from each donor respectively. All specimens were stored at −80°C with 1 ml Trizol (Invitrogen, United States) after grinding and homogenizing with liquid nitrogen. TRIzol™ Plus RNA Purification Kit was used for RNA extraction. Briefly, homogenized tissues were followed by phase separation, RNA precipitation, RNA wash, and RNA redissolving according to experimental protocol. During precipitation step, a 1/10 volume of 3M Sodium acetate (Cat No. AM9740, Invitrogen, United States) was added additionally to help RNA precipitation.
Quantitative Real-Time Polymerase Chain Reaction

The cDNA was synthesized from total RNA by using PrimeScript RT Master Mix (Perfect Real Time) Kit (Takara, Japan). Quantitative real-time PCR (qRT-PCR) was performed in triplicate on a QuantStudio™ 7 Flex Real-Time PCR System (Life Technologies QuantStudio 7 Real Time PCR System, RRID: SCR_020245, United States) by using TB Green Premix Ex Taq II (Tli RNase H Plus) Kit (Takara, Japan). The primers (5′-3′) were ordered from Tech Dragon Ltd. (Hong Kong) and listed in Table 1. The relative expression of each gene was normalized to GAPDH and presented in heatmap after normalization (log10 transformation).

Statistical Analysis

All data were analyzed using SPSS Statistics 23.0 software (IBM SPSS Statistics, RRID: SCR_019096, Chicago, United States). Two groups (PC and DC) with paired data were assessed by the paired sample t-test. A P-value less than 0.05 (p < 0.05) was considered statistically significant and P-values were presented numerically.

RESULTS

Identification of DEGs in Human Knee OA Joint Tissues

For GSE43923 dataset, a total of 542 genes were identified by GEO2R analysis, of which 466 were upregulated and 76 were downregulated. For GSE12021 dataset, a total of 807 genes were identified by GEO2R analysis, of which 122 were upregulated and 685 were downregulated. For GSE51588, a total of 2,584 genes were identified by GEO2R analysis, of which 160 were upregulated (Figure 1F). No overlapping DEG was found in all the OA cartilage, synovial membrane, subchondral bone, and meniscus tissues. Those genes that appeared the most (at least three times) were identified as the most overlapping DEGs and listed in Table 2. A total of 13 most overlapping DEGs were identified. Among them, AHR, HEY1, CXCL12, MPP9, OLFLM2A, SLITRK6, RHBDL2 were highly expressed in cartilage, meniscus, and subchondral bone. COL8A1, GAP43, and PTN were highly expressed in cartilage, synovial membrane, and subchondral bone. In addition, RUNX1, ARL4C, and PIM1 were lower expressed in the meniscus, synovial membrane, and subchondral bone.

GO Enrichment Analysis

The GO enrichment analysis results were presented in Figure 3. For upregulated overlapping DEGs, the most enriched GO Molecular Functions were identified as “proteoglycan binding,” “extracellular matrix structural constituent,” “lipid binding,” “collagen binding,” and “Wnt-protein binding” (Figure 2A). The most enriched GO Biological Processes mainly included “blood vessel development,” “ossification,” “cell morphogenesis involved in differentiation,” “cellular response to growth factor stimulus,” “response to mechanical stimulus” and “extracellular structure organization.” In addition, the most enriched GO Cellular Components were “extracellular matrix,” “cell-cell junction,” “dystrophin-associated glycoprotein complex,” “filopodium,” and “distal axon,” etc. For downregulated overlapping DEGs, the most enriched GO Molecular Functions mainly included “glucose transport activity,” “protein homodimerization activity,” “signaling adaptor activity,” “transcription factor binding,” and “cytokine activity” (Figure 2B). The GO Biological Processes were enriched in “activation of protein kinase activity,” “glucose transport activity,” “cellular response to leptin stimulus,” “SMAD protein signal transduction,” “response to interleukin-6,” and “response to toxic substance.” In addition, the most enriched GO Cellular Components were identified as “secretory granule lumen,” “apical plasma membrane,” “specific granule,” “adherent junction,” and “perinuclear region of cytoplasm.”

KEGG Pathway Analysis

KEGG pathway analysis by Metascape indicated that total overlapping DEGs were enriched in 17 pathways including...
“Wnt signaling pathway,” “Fluid shear stress,” “Axon guidance,” “Nicotinate and nicotinamide metabolism,” “PI3K-Akt signaling pathway,” “HIF-1 signaling pathway,” “MAPK signaling pathway,” “Cytokine-cytokine receptor interaction,” “PPAR signaling pathway,” “NOD-like receptor signaling pathway,” and “TGF-beta signaling pathway,” etc. (Figure 2C). Enriched genes locating in corresponding pathways were summarized in Table 3. MYC, VEGFA, IL2RB, MAPK14, IL6R, MMP9, HLA-DQB1, and CXCL12 were most enriched in these identified KEGG pathways.
In addition, 25 hub genes were obtained and presented in Module one was enriched in nine functions such as stem cell proliferation, response to growth factor, and response to mechanical stimulus, etc. For pathway analysis, Module one was significantly enriched in pathways such as Wnt signaling pathway and Cell adhesion molecules (CAMs) (Figure 4A). Module two was composed of 10 genes, of which five were upregulated and five were downregulated. Module two was enriched in six functions such as regulation of interleukin-1 beta production, response to chemokine, anatomical structure homeostasis, etc. In addition, Rheumatoid arthritis and Cytokine-cytokine receptor interaction pathways were enriched by genes within Module 2 (Figure 4B). Module three included four genes and all of them were upregulated. Enrichment analysis showed that Module three was enriched in Wnt-protein binding and Wnt signaling pathway (Figure 4C). Module five included three genes, all of them were upregulated. Enrichment analysis suggested that Module five was enriched in collagen trimer pathway (Figure 4D).

**Construction of the PPI Network**

A total of 168 interactions were obtained with interaction scores>0.4 by using STRING database. The PPI network was then constructed and presented at Cytoscape platform (Figure 3). In addition, 25 hub genes were obtained and presented in Table 4. The top 10 hub genes included VEGFA, MYC, MMP9, RUNX2, PTPRC, CXCL12, COL1A1, MAPK14, PECAM1, and CD34.

**Module Analysis**

A total of eight modules and four significant modules (Module 1, 2, 3, and 5) were obtained through MCODE analysis (Figure 3). Among significant modules, Module 1 included a total of 13 genes, of which 11 were upregulated and two were downregulated. GO Enrichment analysis showed that Module one was enriched in nine functions such as stem cell proliferation, response to growth factor, and response to mechanical stimulus, etc. For pathway analysis, Module one was significantly enriched in pathways such as PI3K-Akt signaling pathway and Cell adhesion molecules (CAMs) (Figure 4A). Module two was composed of 10 genes, of which five were upregulated and five were downregulated. Module two was enriched in six functions such as regulation of interleukin-1 beta production, response to chemokine, anatomical structure homeostasis, etc. In addition, Rheumatoid arthritis and Cytokine-cytokine receptor interaction pathways were enriched by genes within Module 2 (Figure 4B). Module three included four genes and all of them were upregulated. Enrichment analysis showed that Module three was enriched in Wnt-protein binding and Wnt signaling pathway (Figure 4C). Module five included three genes, all of them were upregulated. Enrichment analysis suggested that Module five was enriched in collagen trimer pathway (Figure 4D).

**Construction of the miRNA-mRNA Network**

The above identified most overlapping DEGs and hub genes were regarded as knee OA-related key genes. By constructing the PPI network of these key genes, we then screened out the experimentally validated key gene-related miRNAs targeting at above OA-related key genes by using miRTarBase (Table 5). As a result, a total of 57 key miRNAs were obtained and visualized in Figure 5. The top 10 key miRNAs included miR-29, miR-101, miR-17, miR-181a, miR-124, miR-8, miR-9, miR-21, miR-155, and miR-185. Key miRNAs were further compared with OA-related miRNAs that were obtained from the Human microRNA Disease Database (HMDD v3.2, http://www.cuilab.cn/hmdd) to check the reliability of our analyses and screen out miRNAs that lack researches so far (Huang et al., 2019).

**Clinical Validation**

Quantitatively Real-time PCR assay (qRT-PCR) was performed to evaluate the relative expression of putative knee OA-related key genes in osteoarthritic cartilage specimens. As a result, significantly upregulated AHR, CYP1A1, and HEY1 were observed in degenerated cartilage compared to the intact cartilage. Besides, MYC and CXCL12 were also upregulated, while no significant difference was observed (Figures 6A–F).

**DISCUSSION**

The present study aims to screen out key DEGs, their relevant signaling pathways, and interactive miRNAs in human knee OA based on bioinformatic analysis. The gene expression profiles derived from different joint tissues are covered and followed by identifying overlapping DEGs. To the best of our knowledge, this is the first time that overlapping DEGs in knee OA are identified from different OA joint tissues, including articular cartilage, synovial membrane, subchondral bone, and meniscus. Gene expression profiles are all obtained from GEO database.

After identification of DEGs in different joint tissues, overlapping DEGs are then identified via Venn Diagram. As a result, 236 overlapping DEGs are identified, of which 160 are upregulated and 76 are downregulated. Those DEGs that are differentially expressed in at least three joint tissues are identified as the most overlapping DEGs. As a result, a total of 13 most overlapping DEGs are identified in our study, include AHR, HEY1, CXCL12, MMP9, COL1A1, GAP43, PTN, RUNX1, RUNX1 family transcription factor 1, PIM1, proto-oncogene serine/threonine-protein kinase Pim-1, ARL4C, ADP ribosylation factor like GTPase 4C.

**Table 2 | List of the most overlapping DEGs in OA joint tissues.**

| Gene  | Locations                          | Expression |
|-------|-----------------------------------|------------|
| AHR   | Cartilage, Meniscus, Subchondral Bone | ↑          |
| HEY1  | Cartilage, Meniscus, Subchondral Bone | ↑          |
| CXCL12/SDF1 | Cartilage, Meniscus, Subchondral Bone | ↑          |
| MMP9  | Cartilage, Meniscus, Subchondral Bone | ↑          |
| OLFML2A | Cartilage, Meniscus, Subchondral Bone | ↑          |
| SLITRK6 | Cartilage, Meniscus, Subchondral Bone | ↑          |
| RHBDL2 | Cartilage, Meniscus, Subchondral Bone | ↑          |
| COL8A1 | Cartilage, Synovium, Subchondral Bone | ↑          |
| GAP43  | Cartilage, Synovium, Subchondral Bone | ↑          |
| PTN   | Cartilage, Synovium, Subchondral Bone | ↑          |
| RUNX1  | Synovium, Subchondral Bone, Meniscus | ↑          |
| PIM1   | Synovium, Subchondral Bone, Meniscus | ↑          |
| ARL4C  | Synovium, Subchondral Bone, Meniscus | ↑          |

AHR, aryl hydrocarbon receptor; HEY1, hairy/enhancer-of-split related with YRPW motif protein 1; CXCL12/SDF1, stromal cell-derived factor 1; MMP9, matrix metalloproteinase 9; OLFML2A, olfactomedin Like 2A; SLITRK6, SLIT and NTRK-like protein 6; RHBDL2, rhomboid Like 2; COL8A1, collagen type VIII alpha 1 chain; GAP43, growth-associated protein 43; PTN, pleiotrophin; RUNX1, RUNX family transcription factor 1; PIM1, proto-oncogene serine/threonine-protein kinase Pim-1; ARL4C, ADP ribosylation factor like GTPase 4C.
metabolism, cartilage homeostasis, and pro-inflammatory responses (De Klerck et al., 2005; Garcia-Cuesta et al., 2019). Upregulation of CXCL12 in OA cartilage, subchondral bone, synoviocytes, and synovial fluid has been reported by previous studies (Chen et al., 2017b; Bragg et al., 2019). Furthermore, inhibition of CXCL12/CXCR4 signaling was shown to prevent

**FIGURE 2** | GO enrichment and KEGG pathway analyses. (A) GO molecular functions, biological processes, and cellular components enrichment analysis on upregulated overlapping DEGs. (B) GO molecular functions, biological processes, and cellular components enrichment analysis on downregulated overlapping DEGs. (C) Enriched KEGG pathways based on total overlapping DEGs.
subchondral bone loss and significantly attenuate cartilage degradation (Dong et al., 2016; Chen et al., 2017b). GAP43 is a nervous tissue-specific cytoplasmic protein related to nerve regeneration. Previous study showed that the expression of GAP43 in pain-related sensory innervation of dorsal-root ganglia (DRG) was upregulated during OA progression (Bachmann and Moroy, 2005). However, there is no study reporting their relationships so far.

The rest of most overlapping DEGs, including HEY1, AHR, RUNX1, and HEY1, are all transcription factors. HEY1 is a basic helix-loop-helix protein (bHLH) transcription factor that belongs to HES/HEY family. Besides, HEY1 is also a direct target of canonical Notch signaling. Previous studies indicated that inhibition of Notch1 exacerbated experimental OA, while increased levels or activity of Notch2 contributed to the progression of OA (Hosaka et al., 2013; Lin et al., 2016). In addition, HEY1 is not only transcriptionally induced by Notch ligands, but also induced by BMP/TGF-β axis to exert as a transcription repressor. Its functions on repressing osteogenic differentiation, neuronal differentiation, and pro-inflammatory production have been reported before (Weber et al., 2014). Hes1, another transcription factor that belongs to HES/HEY family like HEY1, has been reported to be upregulated in OA cartilage and accelerate cartilage destruction via promoting MMP3, a disintegrin and metalloproteinase with thrombospondin motifs 5 (ADAMTS5), and interleukin-6 (IL-6) transcription (Sugita et al., 2015). Nonetheless, the specific role of HEY1 in OA progression remain to be investigated. AHR is a bHLH transcription factor that plays important role in development, immune system, and toxic response. Previous studies have demonstrated that AHR signaling activation significantly alleviated progression of rheumatoid arthritis (RA) through repressing C-reactive protein (CRP), NLR family pyrin domain containing 3 (NLRP3), tumor necrosis factor-alpha (TNF-α), and IL-6 expression (Jin et al., 2011; Huai et al., 2014; Liang et al., 2019; Piper et al., 2019), and enhancing nuclear factor erythroid 2-related factor 2 (NRF2) and IL-10 expression in B cells.
macrophages, or hepatocytes (Tsuji et al., 2012; Piper et al., 2019; Rosser et al., 2020). However, several reports also indicated that AHR signaling activation exacerbated RA inflammation through activating cytokine-mediated inflammatory signaling in primary human fibroblast-like synoviocytes (Adachi et al., 2013; Lahoti et al., 2013). Ogando J et al. reported that the AHR signaling pathway was significantly more active in OA synovial tissues than in RA synovial tissues (Ogando et al., 2016). Furthermore, compared to resting chondrocytes, significantly upregulated AHR was also observed in hypertrophic chondrocytes (Cedervall et al., 2015). Nevertheless, the specific effects of AHR on OA are still in debate and remain to be investigated. RUNX1 was reported to be significantly downregulated in OA cartilage compared with normal control. Furthermore, its anabolic effect on chondrocytes contributes to the maintenance of cartilage homeostasis during OA development.

**FIGURE 3** Construction of protein-protein interaction (PPI) network. PPI network for all the overlapping DEGs is constructed and followed by module analysis on Cytoscape platform. Genes with yellow border are most overlapping DEGs and genes with blue gene symbol are hub genes. Modules with blue border are significant modules. The size of circles reflects the degree of connectivity. The shades of color reflect MCODE score.
and that has also been recognized as a disease-modifying target of OA (Yano et al., 2013; Aini et al., 2016; Yano et al., 2019). In addition, the anti-angiogenic effect of RUNX2 in repressing vascular endothelial growth factor (VEGF) expression has also been reported (Ter Elst et al., 2011).

According to KEGG pathway analysis based on overlapping DEGs, we observe that overlapping DEGs are enriched in “PI3K-Akt signaling pathway,” “Wnt signaling pathway,” “Fluid shear stress,” “Nicotinate and nicotinamide metabolism,” “HIF-1 signaling pathway,” “MAPK signaling pathway,” “Cytokine-cytokine receptor interaction,” “PPAR signaling pathway,” “NOD-like receptor signaling pathway,” “Axon guidance,” and “TGF-β signaling pathway.” The above pathways are well consistent with existing research findings. For example, PI3K-Akt signaling and MAPK signaling are closely involved in inflammatory response to induce the production of catabolic markers such as MMPs, Admats, IL-1β, and TNF-α in OA (Herrero-Beaumont et al., 2019; Chow and Chin, 2020).

Nicotinate and nicotinamide metabolism play important roles in redox reaction. So far, several studies have reported the important role of nicotinate and nicotinamide metabolism in OA development (Yang et al., 2015; Junker et al., 2017). In 2015, Yang et al. (2015) demonstrated that nicotinamide phosphoribosyltransferase (NAMPT), a rate-limiting enzyme in the Nicotinamide adenine dinucleotide (NAD+) salvage pathway, acted as a crucial catabolic regulator of osteoarthritic cartilage destruction. In addition, aberrantly activated Wnt signaling or TGF-β signaling contributes to cartilage degradation, osteophyte formation, and formation of subchondral bone marrow osteoid islets (Zhen et al., 2013; van der Kraan 2017). Therapies targeting Wnt signaling have been undergoing clinical trials (ClinicalTrials.gov Identifier: NCT03928184). HIF-1α and HIF-2α were also reported to exert anabolic and catabolic effects on chondrocytes, respectively (Zhang et al., 2015). PPAR signaling, in particular PPARγ, which is significantly downregulated in OA cartilage, has been demonstrated to maintain articular cartilage homeostasis via regulating the mTOR pathway (Vasheghani et al., 2015; Zhu et al., 2019). NOD-like receptor signaling, which mediates innate immunity and participates in regulating inflammatory and apoptotic responses, mainly includes two subfamilies, NODs and NLRPs (Platnich and Muruve, 2019). Both NOD-dependent pathway and NLRP-dependent inflammasome pathway were validated to mediate OA development under external stimulus (Jin et al., 2011; Xu et al., 2015). Fluid shear stress (FSS) was known as one of the pathogenic mechanisms of OA and has also been used in the construction of an osteoarthritic cell model for a long time (Yang et al., 2020). In addition to above well-validated OA related-pathways, axon guidance pathway is lack of research up to now. Only several in vitro studies reported the effects of an axon guidance molecule, Semaphorin 3A (Sem3A), on osteoarthritic chondrocytes (Okubo et al., 2011; Sumi et al., 2018).

Through constructing the PPI network and miRNA-mRNA network in Cytoscape, a total of 25 hub genes, four significant modules, and 57 key miRNAs are identified. Among hub genes, MAPK14, IL2RB, and IL6R are all involved in cytokine-mediated inflammatory response during OA (Boileau et al., 2005, Liang et al., 2018, Shkhyan et al., 2018). A genome-wide association (GWAS) study has identified a single nucleotide polymorphism (SNP) of HLA-DQB1, rs7775228, associating with knee OA in Asian population (Valdes and Spector, 2011). Furthermore, a small molecule gp130 modulator (RGCD 423) was proven to improve chondrocyte proliferation and inhibit cartilage degradation via upregulating transcription factor MYC and suppressing IL-6-mediated inflammatory response (Shkhyan et al., 2018). For growth factor VEGFA, it is closely related to angiogenesis and inflammatory response (Gao et al., 2013). To validate above computational results, we performed qRT-PCR to validate the relative expression of identified key genes in OA cartilage clinically. Our results show that AHR and its downstream target CYP1A1, HEY1, MYC, and CXCL12 are indeed upregulated in degenerated cartilage compared to preserved cartilage (Figure 6).

According to module analysis, majority of hub genes are located in module one and module 2. In addition, MMP9, CXCL12, and HEY1 are not only the most overlapping DEGs, but also hub genes. Module analysis shows that Modules 1, 2, 3, and 5 are significant modules with the MCODE score ≥4. For Module 1, all of the genes within Module 1, except ANGPT2, are hub genes. MMP9 is located in Module 1. Enrichment analysis suggests that Module one plays important role in PI3K-Akt signaling pathway and cell adhesion molecules (CAMs) pathway. Regarding genes within Module 2, all of them are hub genes except AC5 and FUT4. Genes in Module two play important roles in rheumatoid arthritis and cytokine-cytokine

---

**Table 4 | List of hub genes and corresponding connectivity degree.**

| Gene  | Degree | Gene  | Degree | Gene  | Degree | Gene  | Degree | Gene  | Degree |
|-------|--------|-------|--------|-------|--------|-------|--------|-------|--------|
| VEGFA | 54     | CXCL12| 26     | COL1A2| 17     | WWF   | 14     | DUSP1 | 11     |
| MYC   | 39     | COL1A1| 23     | WNT5A | 17     | NLRP3 | 13     | HEY1  | 11     |
| MMP9  | 32     | MAPK14| 21     | SPT7  | 17     | LOX   | 13     | CXCL2 | 10     |
| RUNX2 | 29     | PECAM1| 21     | FGF13 | 17     | CD163 | 13     | ACTA2 | 10     |
| TPTRC | 29     | CD24  | 20     | TLR7  | 16     | CXCR1 | 12     | MSX1  | 10     |

**Notes:**

| VEGFA, vascular endothelial growth factor A; MYC, c-myc; MMP9, matrix metallopeptidase 9; RUNX2, runt-related transcription factor 2; TPTRC, protein tyrosine phosphatase receptor type C; CXCL12/SDF1, stromal cell-derived factor 1; COL1A1, collagen type I alpha 1 chain; MAPK14, mitogen-activated protein kinase 14; PECAM1, platelet endothelial cell adhesion molecule 1; COL1A2, collagen type I alpha 2 chain; WNT5A, Wnt-5a; SPT7, Sp7; FGF13, fibroblast growth factor 13; TLR7, Toll-like receptor 7; VWF, Von Willebrand factor; NLRP3, NLR family Pyrin domain containing 3; LOX, lipoxygenase; CXCR1, C-X-C motif chemokine receptor 1; DUSP1, dual specificity phosphatase 1; HEY1, Hes related family BHLH transcription factor with YRPW motif 1; CXCL2, C-X-C motif chemokine ligand 2; ACTA2, actin alpha 2; MSX1, Msh homeobox 1.
FIGURE 4 | Module analysis. (A) Module one is comprised of 13 genes and enriched in PI3K-Akt signaling pathway and cell adhesion molecules (CAMs). (B) Module two is comprised of 10 genes and enriched in rheumatoid arthritis pathway and cytokine-cytokine receptor interaction pathway. (C) Module three is comprised of four genes and enriched in the Wnt signaling pathway. (D) Module five is comprised of three genes and enriched in collagen trimer pathway. Upregulated and downregulated genes are stained with red and green, respectively.
receptor interaction pathways. Furthermore, CXCL12 is located in Module 2. Module three is a group of genes associated with Wnt signaling pathway. No most overlapping DEGs, hub genes, or pathways have been demonstrated to be aberrantly expressed in degraded cartilage (Karlsson et al., 2010; Feng et al., 2019). A recent study reported that Lgr5+/Col22a1+ stem cells play important roles in differentiation toward articular chondrocytes and Col22a1-expressing cartilage superficial layer contributed to repair of cartilage defect (Feng et al., 2019).

Collectively, based on above identified overlapping DEGs, hub genes, pathways, and functional modules, the present study depicted the potential OA mechanisms covering the entire knee joint. We believe that our identified knee OA-related key genes, such as AHR, HEY1, MYC, GAP43, and PTN, and their relevant signaling pathways, including AHR signaling, Notch signaling and TGF-β signaling, C-MYC signaling, and axon guidance pathway may play important roles in knee OA development. Our predicted genes are worthy of being explored as novel targets of DMOADs in the future.

By comparing key miRNAs with OA-related miRNAs included in HMDD v3.2 database, 41 out of 57 miRNAs (about 71.9%) have been reported to be associated with OA, suggesting the considerable reliability of our miRNA-mRNA interactome predictions. For example, miR-29 family, including miR-29a, miR-29b, and miR-29c, was differentially expressed in OA cartilage and negatively regulated Smad, NF-κB, and canonical Wnt signaling (Le et al., 2016). In addition, upregulation of miR-101 significantly facilitated cartilage degradation and chondrocyte apoptosis (Dai et al., 2015; Lü et al., 2020). For miR-181a, Akihiro Nakamura et al. demonstrated that intra-articular injections of locked nucleic acid (LNA) miR-181a antisense oligonucleotides (ASO)

### TABLE 5 | List of miRNAs targeting at knee OA-related key genes.

| Gene | miRNAs |
|------|---------|
| AHR  | miR-124-3p, miR-181a-5p, miR-29a-3p |
| HEY1 | miR-410-3p |
| CXCL12 | miR-23a-3p, miR-31-5p, miR-886-3p, miR-126-5p, miR-146a-5p, miR-221-3p, miR-454-3p, miR-137, miR-1-5p, miR-1-3p, miR-448 |
| MMP9 | miR-651a, miR-491-5p, miR-338-3p, miR-204-5p, miR-21-5p, miR-9-5p, miR-211-5p, let-7e-5p, miR-133b, miR-29b-3p, miR-9-3p, miR-524-5p, miR-302a-5p, miR-132-3p, miR-15b-5p, miR-942-3p, miR-203a-5p, miR-133a-5p, miR-143-3p |
| RUNX1 | miR-17-5p, miR-20a-5p, miR-106a-5p, miR-675-5p, miR-221-3p, miR-27b-3p, miR-18a-5p, miR-215-5p, miR-9-5p, miR-101-3p, miR-144-5p, miR-181a-5p, miR-378a-3p |
| GAP43 | miR-363-3p |
| PTN  | miR-155-5p |
| PIM1 | miR-210-3p, miR-1-3p, miR-192-5p, miR-16-5p, miR-33a-5p, miR-33b-5p, miR-214-3p, miR-124-3p, miR-542-3p, miR-101-3p, miR-486-5p |
| VEGFA | miR-373-3p, miR-302d-3p, miR-126-3p, miR-147a, miR-134-5p, miR-140-5p, miR-29b-3p, miR-107, miR-16-5p, miR-90-5p, miR-17-5p, miR-150-5p, miR-195-5p, miR-15b-5p, miR-15a-5p, miR-520g-3p, miR-378a-3p, miR-230-3p, miR-385-5p, miR-125a-5p, miR-381-5p, miR-20a-5p, miR-20b-5p, miR-504-5p, miR-520h, miR-372-3p, miR-101-3p, miR-106b-5p |
| MYC  | miR-24-3p, let-7a-5p, let-7g-5p, miR-34a-5p, miR-98-5p, let-7c-5p, miR-26a-5p, miR-145-5p, miR-21-5p, miR-34b-5p, miR-34c-5p, miR-18a-5p, miR-17-5p, miR-20a-5p, miR-34b-3p, miR-378a-3p, miR-371a-3p, miR-373-3p, miR-33b-5p, miR-135a-5p, miR-449c-5p, miR-429, miR-335-5p, let-7i-5p, miR-320b, miR-744-5p, miR-320a, miR-148a-3p, miR-212-3p, miR-491-5p, miR-155-5p, miR-15a-5p, miR-449a, miR-487b-3p, miR-7-5p, miR-90-5p, miR-324-3p, miR-184, miR-126-5p, miR-25-5p, miR-92a-2-5p, miR-92a-1-5p, miR-19b-2-5p, miR-19b-1-5p, miR-106b-5p, miR-295-5p, miR-130a-3p, miR-106a-5p, miR-185-5p, miR-29a-3p, miR-561-3p, miR-34a-3p, miR-599, miR-29b-3p, miR-129-2-3p |
| RUNX2 | miR-135b-5p, miR-155-5p, miR-335-5p, miR-200a-3p, miR-497-5p, miR-195-5p, miR-204-5p, miR-433-3p, miR-330d-5p, miR-30a-5p, miR-30b-5p, miR-218-5p, miR-23b-3p, miR-34c-5p, miR-30c-5p, miR-205-5p, miR-320a, miR-135a-5p, miR-222-3p, miR-221-5p, miR-628-3p, miR-103a-3p, miR-30a-3p, miR-376c-3p |
| COL1A1 | miR-29b-3p, miR-29c-3p, miR-143-3p, miR-133a-3p, miR-124-3p, miR-29b-1-5p, miR-185-5p, miR-129-5p |
| WNT5A | miR-374a-5p, miR-487b-3p, miR-516a-3p, miR-154-5p, miR-154-3p |
| SP7  | miR-135b-5p, miR-637, miR-7-5p, miR-93-5p, miR-31-5p, miR-145-5p |
| TLR7 | miR-758-3p, miR-17-5p, miR-19a-3p |
| NLPR3 | miR-223-3p |
| LOX  | miR-29a-3p, miR-29b-3p, miR-29c-3p, miR-767-5p, miR-200b-3p, miR-30a-5p |
| CXCR1 | miR-396-3p |
| DUSP1 | miR-101-3p, miR-200c-3p, miR-940, miR-146a-5p |
| CXCL2 | miR-223-3p, miR-532-5p |
| MAPK14 | miR-124-3p, miR-24-3p, miR-199a-3p, miR-200a-3p, miR-141-3p, miR-125b-5p, miR-214-3p, miR-155-5p, miR-17-5p, miR-106a-5p, miR-128-3p, miR-27a-3p, miR-125a-5p |
| CD34 | miR-125a-5p, miR-9-5p, miR-24-3p, miR-377-3p |
| COL1A2 | let-7g-5p, miR-29c-3p, miR-26b-5p, miR-25-3p, miR-29a-3p |
significantly attenuated cartilage destruction in facet and knee joints in vivo (Nakamura et al., 2019). Another study reported that miR-9 promoted IL-6 expression and exacerbated cartilage degradation by targeting MCPIP1 expression (Makki et al., 2015). In 2019, a hydrogel-based drug delivery system equipped with locked nucleic acid (LNA) miR-221 inhibitor was constructed and enhanced cartilage regeneration significantly (Lolli et al., 2019). miR-199a* was also shown to inhibit IL-1β-induced Cyclooxygenase 2 (COX-2) expression as a direct regulator (Akhtar and Haqqi, 2012). Notably, a RNA sequencing-based miRNA-mRNA interactome study which confirmed a OA-specific miRNAs array showed considerable overlap with our identified key miRNAs as well, including miR-143, miR-155, let-7g, miR-7, miR-101, miR-145, miR-16, miR-30a, miR-29, miR-1, miR-133a-3p, miR-20a, miR-320a, miR-31, miR-93, miR-221, and miR-335 (Coutinho de Almeida et al., 2019). Among our identified top 10 key miRNAs, in addition to those that have been reported before, miR-17, miR-21, miR-155, miR-185, and miR-1 are still lack of research and remain to be investigated in the future.

The present study also has several limitations. For example, although we originally intended to include as many data sets as possible, only four expression profiles met the criteria and were included into the data analysis. In addition, due to difficulties in obtaining appropriate specimens, only OA articular cartilage specimens were used to carry out clinical validation experiments. The relative expression of our identified knee OA-related key genes in other joint components need to be further validated.

In conclusion, the present study provides a comprehensive bioinformatics analysis of key genes, signaling pathways, and miRNAs in different joint tissues of knee OA patients. A total of 35 knee OA-related key genes and 57 key miRNAs were identified. Among them, key genes such as AHR, HEY1, MYC, GAP43, and PTN, and key miRNAs such as miR-17, miR-21, miR-155, miR-185, and miR-1 are lack of research so far. For key
genes identified in the present study, their downstream mechanisms and specific effects on the different joint components also need to be explored, respectively. Through enrichment analysis, a number of OA-related pathways were identified, including PI3K-Akt signaling, Wnt signaling, fluid shear stress, nicotinate and nicotinamide metabolism, HIF-1 signaling, MAPK signaling, cytokine-cytokine receptor interaction, PPAR signaling, NOD-like receptor signaling, TGF-β signaling, and axon guidance pathways. Among them, many pathways have been well investigated and even under clinical trials, except for axon guidance pathway which is implicated in nerve innervation and axon guidance while still lack of research so far. Our study provides insight for the first time in identification of potential therapeutic targets of knee OA by comprehensively analyzing the overlapping genes differentially expressed in multiple joint components based on bioinformatics.

**DATA AVAILABILITY STATEMENT**

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding authors.

**ETHICS STATEMENT**

For clinical specimens sampling, all donors have provided informed consent, and this study was approved by the Joint Chinese University of Hong Kong-New Territories East Cluster Clinical Research Ethics Committee (CREC Ref. No: 2013.248).

**AUTHOR CONTRIBUTIONS**

LQ, JX, and LC designed the study; LC carried out the experiments; HY, ZY, WT, and BD helped perform the experiments; KH, and MO helped collect clinical specimens; LC performed computational analysis and analyzed the data; LC wrote the manuscript; JX and LQ supervised the project and helped amend the manuscript. All authors read and approved the final manuscript.

**FUNDING**

This work was fully supported by Theme-based Research Scheme from RGC-Hong Kong (No. T13-402/17N) and partially supported by Health and Medical Research Fund (18190481).

**ACKNOWLEDGMENTS**

The authors would like to thank the support of Tech Dragon Ltd. (Hong Kong, China) for primers preparation. Furthermore, we would like to appreciate all reviewers for their insightful comments and constructive suggestions to polish this paper in high quality.
