Article

Construction of a Diagnostic m⁷G Regulator-Mediated Scoring Model for Identifying the Characteristics and Immune Landscapes of Osteoarthritis

Liang Hao, Xiliang Shang, Yang Wu, Jun Chen (ID) **and Shiyi Chen** *

Huashan Hospital, Fudan University, Shanghai 200040, China
* Correspondence: cshiyi@163.com; Tel.: +86-021-52888255

Abstract: With the increasingly serious burden of osteoarthritis (OA) on modern society, it is urgent to propose novel diagnostic biomarkers and differentiation models for OA. 7-methylguanosine (m⁷G), as one of the most common base modification forms in post transcriptional regulation, through which the seventh position N of guanine (G) of messenger RNA is modified by methyl under the action of methyltransferase; it has been found that it plays a crucial role in different diseases. Therefore, we explored the relationship between OA and m⁷G. Based on the expression level of 18 m⁷G-related regulators, we identified nine significant regulators. Then, via a series of methods of machine learning, such as support vector machine recursive feature elimination, random forest and lasso-cox regression analysis, a total of four significant regulators were further identified (DCP2, EIF4E2, LARP1 and SNUPN). Additionally, according to the expression level of the above four regulators, two different m⁷G-related clusters were divided via consensus cluster analysis. Furthermore, via immune infiltration, differential expression analysis and enrichment analysis, we explored the characteristic of the above two different clusters. An m⁷G-related scoring model was constructed via the PCA algorithm. Meanwhile, there was a different immune status and correlation for immune checkpoint inhibitors between the above two clusters. The expression difference of the above four regulators was verified via real-time quantitative polymerase chain reaction. Overall, a total of four biomarkers were identified and two different m⁷G-related subsets of OA with different immune microenvironment were obtained. Meanwhile, the construction of m⁷G-related Scoring model may provide some new strategies and insights for the therapy and diagnosis of OA patients.

Keywords: osteoarthritis; 7-methylguanosine; biomarkers; scoring model

1. Introduction

Being unpredictable and one of the most common chronic degenerative joint diseases, OA (OA) has a prevalence that increases with age. OA causes significant pain and disability [1–3]. Furthermore, inflammation or fibrosis of the infrapatellar fat pad is present in patients with OA, which is one of the well-established risk factors for the development of the pain caused by OA [4,5]. Moreover, pathological changes, including subchondral osteosclerosis, synovitis, fibrosis, and cartilage degeneration are closely associated with OA [6–9]. OA has a multifaceted etiology; and is caused by a combination of immune response, chronic inflammation, trauma, and biomechanical processes [10–12]. Congenital joint abnormalities, trauma, stress injury, obesity, sex, age, and knee gap narrowing [13] are all complicatedly associated with the development of OA [14].OA can have serious physical, emotional, and economic consequences; moreover, it is becoming an evolving public health issue that negatively affects the daily lives and quality of life of people [15]. With an increasing prevalence, overall, the age-standardized prevalence of OA increased by 7.5% in Northern Europe between 1990 and 2015, with an annual increase of 43% [16]. Clinical symptoms and imaging findings are required for making the standard diagnosis of

OA; however, often, when these symptoms appear, patients have already reached the late stage of the disease [17]. Moreover, patients with OA show relatively severe symptoms and have poor treatment results because no current drug treatment has been found to reverse the progression of OA in the long term [15]. Therefore, it is even more important to look for novel diagnostic modalities that can help diagnose OA as early as possible, facilitating the timely treatment of patients.

RNA modifications were not known as the "epitranscriptome" until 2015. Emerging studies on the function of these modifications have shown significant implications for human pathology [18]. Adenosine methylation is present in mRNAs and non-coding RNAs, such as circular RNA (circRNAs), microRNAs (miRNAs), and long-stranded non-coding RNAs (lncRNAs), which adjust their biogenesis and function [19]. The RNA methylation types can be divided into various modification types: m^6A, m^5C, m^7G, and so on; of which m^6A and m^7G are the two most common types [20–22]. However, so far, more extensive studies on the association between m^6A and OA are available, whereas there are few studies on m^7G [23,24]. Moreover, there is a prevalence of m^7G RNA modifications within mRNA, and their conservation, regulation, and dynamics as well as their roles in translational control have been shown. Modifications of the m^7G cap are widely seen in OA-related mRNA [25]; they play an important role in the efficient translation, splicing, and stability of related mRNA and also affect the synthesis of related proteins. Modifications of m^7G are also observed in OA-related mRNA and help in enabling the translation of OA-related mRNA. Thus, m^7G, as a transcriptional marker, is important for protein translation and can be used as the basis for making diagnostic models for OA.

Herein, we have analyzed numerous publicly available microarray datasets. Using differential analysis and algorithmic screening, we have acquired the most critical genes of m^7G-regulators for forming an intersection. We have identified four significant regulators, including DCP2, EIF4E2, SNUPN, and LARP1, in combination with machine learning. Based on the expression of these four regulators, we have divided all the OA samples into m^7G-related clusters. Next, we performed principal component analysis (PCA) to calculate the m^7G score in the above two clusters. Then, we performed differential expression analysis, enrichment analysis, and immune infiltration to explore the characteristics of the two clusters. Finally, after the intersection of the differentially expressed genes (DEGs) between the above two clusters and DEGs between the normal samples (NM) and OA samples, we constructed a diagnostic model using LASSO Cox regression. Furthermore, we verified the abovementioned four regulators with the use of real-time quantitative polymerase chain reaction (RT-qPCR).

2. Materials and Methods

2.1. Data Acquisition and Processing

By retrieving the keyword "OA" from the Gene Expression Omnibus database, datasets containing the synovial tissue samples of normal people and patients with OA were obtained [26]. Moreover, the standards for screening our datasets are as follows: (1) Homo sapiens Expression Profiling by array, (2) synovial tissue of OA from joint synovial biopsies, (3) datasets containing complete information about the samples, (4) one biopsy sample per subject was analyzed without replicates. The detailed information regarding the datasets used in our study is listed in Table 1. Then, "inSilicoMerging" [27] was used to merge and the "limma" package (v3.42.2) in R software [28] was used to remove the batch effect of these three data sets; finally, a data set containing the synovial tissue samples of 26 patients with OA and 20 normal people was obtained. Furthermore, 24 m^7G-regulators collected in the gene sets GOMF_RNA_7_METHYLGUANOSINE_CAP_BINDING, GOMF_M7G_5_PPPN _DIPHOSPHATASE_ACTIVITY, and GOMF_RNA_CAP_BINDING, were summarized, wherein only 18 regulators were annotated in our data sets [29]. In this study, the association between the expression of these 18 genes and diseases as well as that between m^7G-regulators self-expression was studied. The results were visualized using heat maps. The Wilcoxon signed-rank test was performed to select some significant regulators.

Table 1. The information of datasets included in our research. OA, osteoarthritis; NM, normal.

| Accession | Platform | Samples | | Tissue |
		NM	OA	
GSE55235	GPL96	10	10	synovium
GSE55457	GPL96	10	10	synovium
GSE55584	GPL96	0	6	synovium
GSE12021	GPL96	9	10	synovium
GSE32317	GPL570	Early OA 10	Late OA 9	synovium

2.2. Enrichment Analysis

Gene ontology (GO) analysis is a common method to annotate gene products and the functions of genes, including cellular component, biological pathway, and molecular function [30]. The Kyoto Encyclopedia of Genes and Genomes (KEGG) is a useful database for the systematic analysis of gene functions and associated high-level genomic functional information [31]. We used "clusterProfiler" (version 3.14.3) and R software packages "org.Hs.eg.db" (version 3.1.0) for the GO and KEGG pathway analysis. The maximum size gene sets were set to 5000 genes and the minimum to 5; the analysis results with a *p*-value of <0.05 were considered significant. ClueGO is an important plug-in of Cytoscape, which can be used for GO enrichment analysis. It was also used for enrichment analysis in our study [32].

2.3. Construction and Verification of Prediction Model

To more accurately screen the m^7G-regulators associated with the occurrence of OA, significant m^7G-regulators and OA were extracted using the support vector machine recursive feature elimination (SVM-RFE) algorithm and random forest algorithm (RF) [33], and the two analysis results were intersected using the Venn diagram. Finally, to determine the best prognostic characteristic regulators, the LASSO Cox regression analysis was performed on the basis of the abovementioned results [34]. Then, cluster analysis was performed on the finally screened feature genes using the "ConsensusClusterPlus" package; accordingly, the samples were divided into two categories. PCA was used for extracting PC1 and PC2 to form signature scores. Afterward, the above scores were applied to construct the m^7G score: **m^7G score $= \Sigma (PC1_i + PC2_i)$.**

2.4. Immune Infiltration

The genes significantly associated with 28 immune cell types from the literature were collected (Table S1). Then, the expression of these immune genes were linked with the distribution of 28 types of immune cells using the single-sample gene set enrichment analysis method [35]; combined with our m^7G regulators; we finally analyzed the association of 28 immune cells with different m^7G-related clusters and immune cells with m^7G characteristic genes.

2.5. Construction of Diagnostic Model of OA

To better explore the characteristic of our two m^7G-related clusters, differential expression analysis was performed to assess the DEGs between the two m^7G-related clusters and those between the NM and OA samples with the cutoff criteria of $|\log2FC| > 1$ ($p < 0.05$). After the intersection of the two different types of DEGs, a diagnostic model was constructed using the overlapped DEGs via the LASSO Cox regression. The diagnostic score was as follows: **Diagnostic Score $= \sum_i Coefficients_i *$ Expression level of $signature_i$.**

2.6. Collection of Our Synovial Samples

In our study, synovial tissue samples of OA were obtained from patients undergoing surgery due to knee OA (n = 15). The normal synovial samples were from patients who underwent surgery for meniscus laceration of the knee joint caused by trauma (n = 15).

2.7. RT-qPCR

Trizol (Invitrogen, Waltham, MA, USA) reagent was used to extract total RNA from the synovial tissue samples of normal people and patients with OA; the Prime Script TMRT kit (Takara, RR047A) was used to reverse transcribe to obtain cDNA. Finally, the SYBR Premix Ex Taq II Kit (Takara, Japan) was used for PCR amplification according to the manufacturer's instructions. And we used Bio-Rad (CFX96) of the UK for RT-qPCR. The primer sequences are listed in Table S2.

2.8. Statistical Analysis

R software (version 4.2.1) and its related software packages were used to process and analyze data ($p < 0.05$). We used the Wilcoxon signed-rank test to assess the significance of difference between the two groups. And a *t*-test was used for the analysis of the result of RT-qPCR. Afterward, Sangerbox was used to visualize the results of the receiver operating characteristic curves and Wilcoxon signed-rank test; the "RMS" package in R software was used to visualize Nomograms.

3. Results

3.1. Identification of Significant m^{7}G-Regulators in OA

We calculated the Spearman correlation coefficient among these 18 regulators, wherein several regulators demonstrated significant correlation on the basis of the expression levels of 18 m^{7}G-regulators in our datasets which were merged with the three datasets from GEO (Figure 1A). Meanwhile, the interaction association of these 18 regulators was revealed by constructing a protein-protein interaction network (Figure 1B). Then, we used the Wilcoxon signed-rank test to identify the significant regulators in our training set. Thus, nine significant regulators ($p < 0.05$) were obtained, including DCP2, IFIT5, EIF4E2, NUDT11, NUDT3, LARP1, SNUPN, LSM1, and CYFIP1 (Figure 1C). The heat map shows the expression level of these nine significant regulators (Figure 1D).
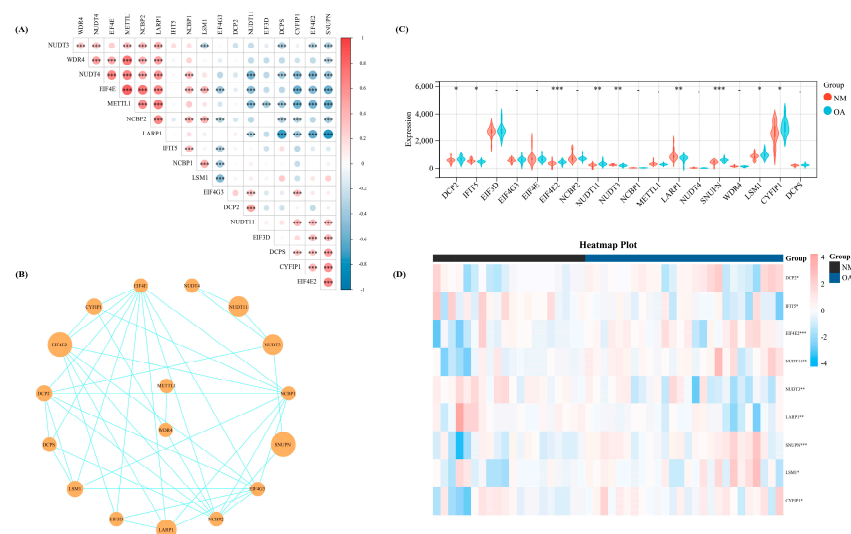

Figure 1. The basic landscape of m^{7}G-regulators in OA. (**A**) The correlation heap map of the expression level of the 18 m^{7}G-regulators. (**B**) The PPI network of the 18 m^{7}G-regulators. (**C**) Th expression difference of the 18 m^{7}G-regulators between NM and OA group. (**D**) The heap map of the nine significant regulators between NM and OA. -, $p > 0.05$; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$. NM, normal; OA, OA.

3.2. Enrichment Analysis for Our Significant m^{7}G-Regulators

The ClueGO plug-in in Cytoscape was first used to perform enrichment analysis ($p < 0.05$) for comprehensively exploring the function of the above nine significant m^{7}G-regulators in OA. Thus, the term with the largest proportion was "m^{7}G(5')pppN diphosphatase activity" (53.85%); furthermore, the rest of the terms revealed that our significant m^{7}G-regulators were almost involved in the pathways of RNA metabolism and translation progress (Figure 2A). Meanwhile, the MCODE plug-in was used to extract the important clusters of the ClueGO results. The clusters with a high score are shown in Figure 2B–D; the cluster with the highest score was the "m^{7}G(5')pppN diphosphatase activity" pathway, which was consistent with the biological function of these regulators.
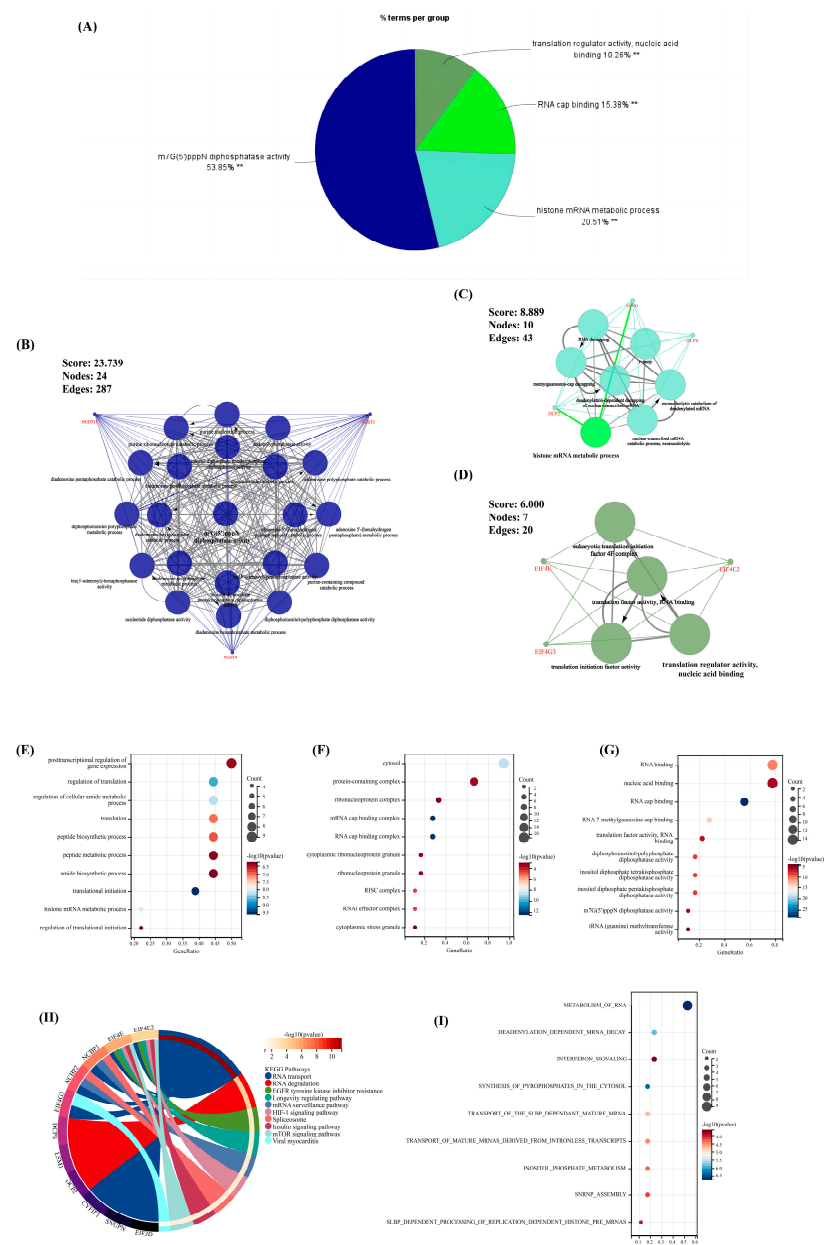

Figure 2. The results of enrichment analysis for the 18 m^{7}G-regulators. (**A**) The pie chart of the results for ClueGO enrichment analysis. (**B–D**) Three significant clusters identified via MCODE plug-in for the results of ClueGO enrichment analysis. (**E–G**) The bubble chart of the results for GO enrichment analysis, (**E**) biological pathway (BP), (**F**) cellular component (CC), (**G**) molecule function (MF). (**H,I**) The result of the (**H**) KEGG and (**I**) Reactome enrichment analysis. **, $p < 0.05$.

Moreover, the GO, KEGG, and Reactome analyses were performed to ensure the preciseness of our research ($p < 0.05$) (Figure 2E–I). Correspondingly, almost all results of different methods of enrichment analysis revealed that these regulators focused on RNA metabolism. Interestingly, pathways associated with the immune system and other pathways, including the viral myocarditis, interferon signaling, and HIF-1 signaling pathways, were found, indicating that these m^{7}G-regulators played a significant role in RNA modification as well as immune and other fields.

3.3. Selection of Significant m^{7}G-Regulators via Machine Learning

In line with the above analysis, we explored the important and key role of OA by using several methods of machine learning to further identify some significant regulators in OA. First, SVM-RFE was performed to evaluate the diagnostic effectiveness of these regulators. Thus, seven regulators were obtained. Meanwhile, another method, RF, was used to calculate the importance of these regulators (Figure 3A,B). With a score of >2, six regulators were selected (Figure 3C,D). Then, we intersected the results of SVM-RFE and RF, through which four regulators (EIF4E2, DCP2, SNUPN, and LARP1) were considered the final crucial regulators (Figure 3E).
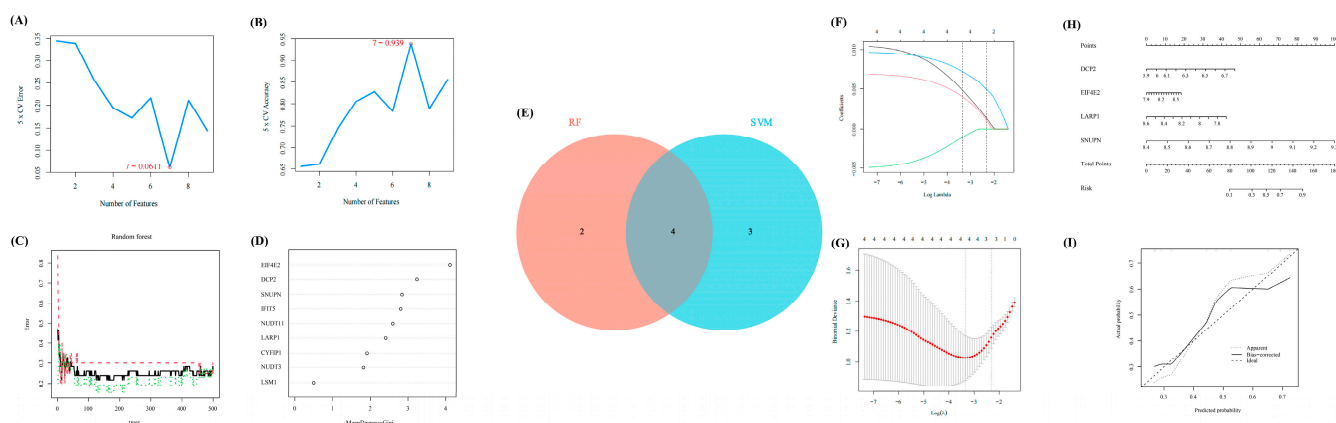

Figure 3. The further selection of crucial m^{7}G-regulators for OA via machine learning. (**A**,**B**) The results of SVM-RFE identified seven significant features. (**C**,**D**) The results of RF identified six significant features. (**E**) The Venn diagram to extract the overlapped features between RF and SVM-RFE. (**F**,**G**) The results of lasso-cox regression analysis. (**H**) The nomogram of the four significant regulators to distinguish early- and end-stage of OA confirmed via machine learning. (**I**) The calibration curve of our nomogram.

Furthermore, the LASSO Cox regression was performed to verify the diagnostic effectiveness of these four regulators (Figure 3F,G). Thus, all four regulators were regarded as significantly diagnostic signatures.

Moreover, based on the expression matrix of GSE32317, the nomogram was further exhibited to reveal the efficiency of the above four m^{7}G-regulators in distinguishing early- and end-stage OA, and the calibration curve revealed the accuracy of our model (Figure 3H,I).

3.4. Identification of Two Different m^{7}G-Related Clusters

According to the expression levels of the four key regulators selected via machine learning, we divided our OA samples in the training set into two m^{7}G-related clusters with the most appropriate K value (K = 2) via consensus cluster analysis (Figure 4A–C). Furthermore, the PCA diagram revealed a significant difference between clusters A and B (Figure 4D). Meanwhile, all four regulators showed a significant statistical difference between m^{7}G-related clusters A and B (Figure 4E,F). Based on the PCA algorithm, an m^{7}G score module was calculated to distinguish the above two clusters ($p < 0.05$), which was higher in m^{7}G-related cluster B and lower in m^{7}G-related cluster A (Figure 4G).
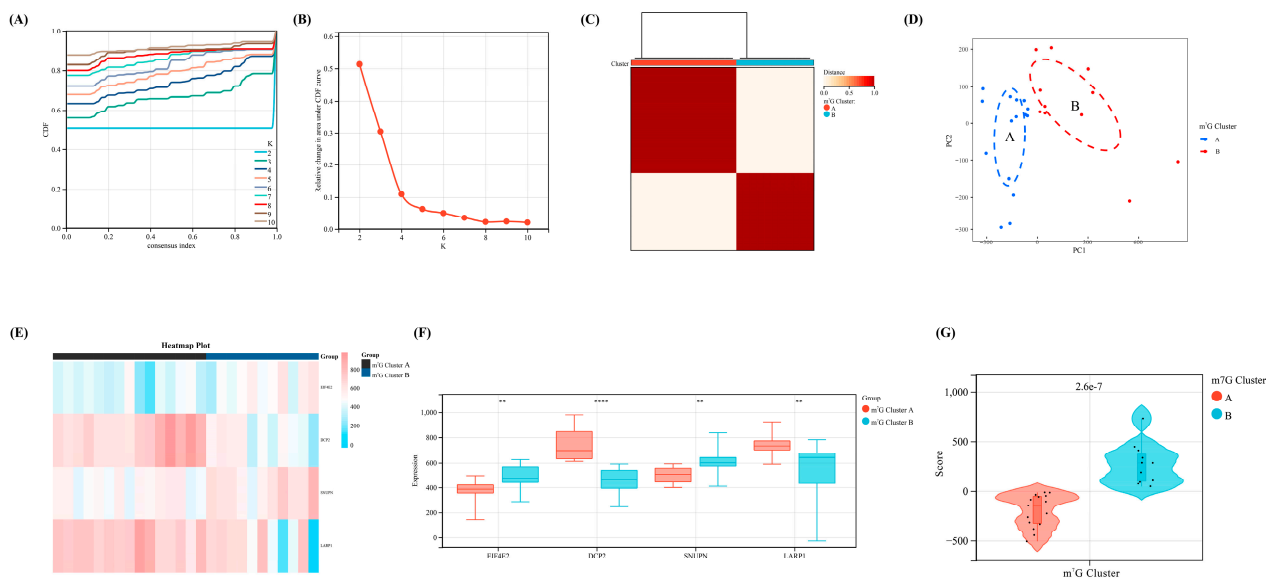
Figure 4. The identification of m^{7}G-related clusters and construction of m^{7}G scoring model. (**A**) The cumulative distribution curve of the result for consensus cluster analysis. (**B**) The area under the curve of the result for consensus cluster analysis. (**C**) The heat map of two different m^{7}G-related clusters. (**D**) The PCA map of the above two different clusters. (**E**) The heap map of our four crucial regulators revealed the expression level between two clusters. (**F**) The expression difference of the four regulators between two clusters. (**G**) The difference of m^{7}G score between two clusters. **, $p < 0.01$; ****, $p < 0.0001$.

3.5. GSEA, Immune Infiltration, and Immune Checkpoint Characteristics in m^{7}G–Related Clusters

To better describe the characteristics and functions in the abovementioned m^{7}G-related clusters, the GSEA analysis was performed. We identified three pathways with a *p*-value of <0.05, including TGF_BETA_SIGNALING, ALLOGRAFT_REJECTION, and ESTROGEN_RESPONSE_EARLY, which indicated that the m^{7}G score was mainly associated with the metabolism and immune system (Figure 5A). Therefore, we performed the immune infiltration analysis and the mantel test to demonstrate the association between the four significant regulators and the infiltration score of the 28 immune cells. Interestingly, a stronger association was noted in cluster A, indicating that the lower m^{7}G score suggested prominently elevated infiltration of immune cells in patients with OA (Figure 5B,C).

Furthermore, the Pearson correlation coefficient between the expressions of a series of immune checkpoint-related genes in the above two clusters and the m^{7}G score was calculated for more comprehensively exploring the immune signature between these two clusters. For the immune checkpoint inhibitors, the m^{7}G score in m^{7}G-related cluster A was positively associated with most of the inhibitors (Figure 5D). On the contrary, the m^{7}G score in m^{7}G-related cluster B was negatively correlated with most inhibitors (Figure 5E).
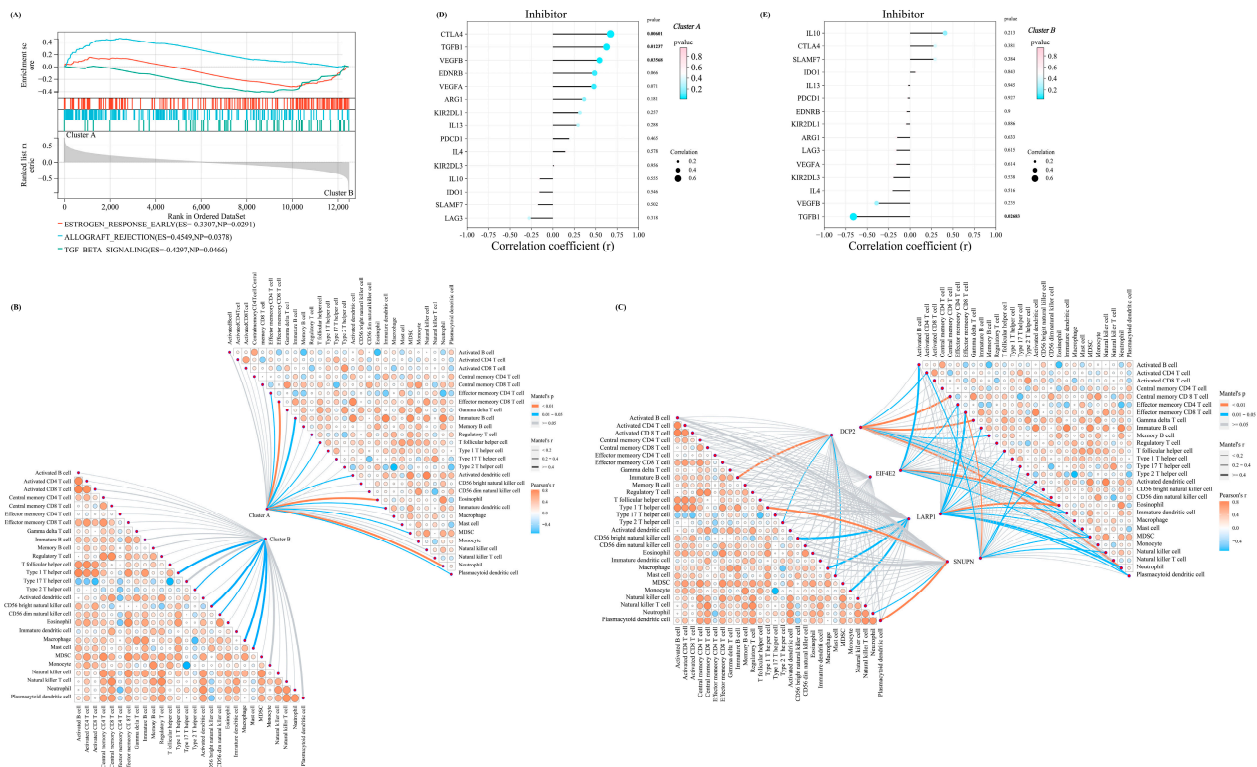
Figure 5. The distinctive characteristic of the two m⁷G-related clusters. (**A**) The results of the GSEA enrichment analysis in two clusters. (**B**) The mantel correlation heat map between the m⁷G score and 28 different kinds of immune cells in two different clusters. (**C**) The mantel correlation heat map between the expression level of four significant regulators and 28 different kinds of immune cells in two different clusters. (**D,E**) The lollipop chart revealed the results of Pearson correlation analysis between m⁷G score and the expression level of several immune checkpoint inhibitors in (**D**) cluster A and (**E**) cluster B.

3.6. Exploration of Difference between the above Two Clusters and Construction of a Diagnostic Model

To further emphasize the significance of the m⁷G-related clusters, differential expression analysis was performed with $|\log2FC| > 1$ ($p < 0.05$) as the cutoff. Thus, 113 DEGs were identified, including 46 upregulated and 67 downregulated, which were visualized using the volcano and heat maps (Figure 6A,B). Moreover, the DEGs between NM and OA and between the above two clusters intersected. Finally, eight overlapped DEGs were obtained (CRYBB1, N6AMT1, SNORA21, HAUS2, P2RX3, RRN3P1, CC2D1A, and FKBP5), which were considered the candidate factors extracted for the LASSO Cox regression (Figure 6C,D). We regarded lambda-min: 0.0469 as the optimal value after running the 10-fold cross-validation. Thus, five factors were selected to construct our diagnostic model for OA: Diagnostic value = $(-0.00806730919977491 \times$ SNORA21$)$ + $(-0.00181179794976226 \times$ HAUS2$)$ + $(-0.0134814157908044 \times$ CC2D1A$)$ + $(-0.00123114088948981 \times$ FKBP5$)$ + $(0.00983579963789009 \times$ N6AMT1$)$ (Figure 6E,F). Furthermore, all samples were randomly divided into two different subsets with a ratio of 1:1: a verification set and a training set. According to the above formula of the diagnostic score, the Wilcoxon signed-rank test was performed to explore the statistical difference between the NM and OA samples. Finally, the diagnostic score of both sets demonstrated a significant difference ($p < 0.05$) (Figure 6G,I). In addition, the area under the receiver operating characteristic curve of the diagnostic model in the verification and training set was 85.4701 and 93.0070, respectively, which further indicated the excellent effectiveness of our diagnostic model (Figure 6H,J). Moreover, to further verify the accuracy of the diagnostic model, another external dataset GSE12021 was selected, wherein a significant difference was observed in the diagnostic value between NM and OA samples ($p < 0.05$) (Figure 6K).

Meanwhile, in GSE12021, the area under the receiver operating characteristic curve of the diagnostic model was 97.7778, demonstrating the accuracy of our model (Figure 6L).
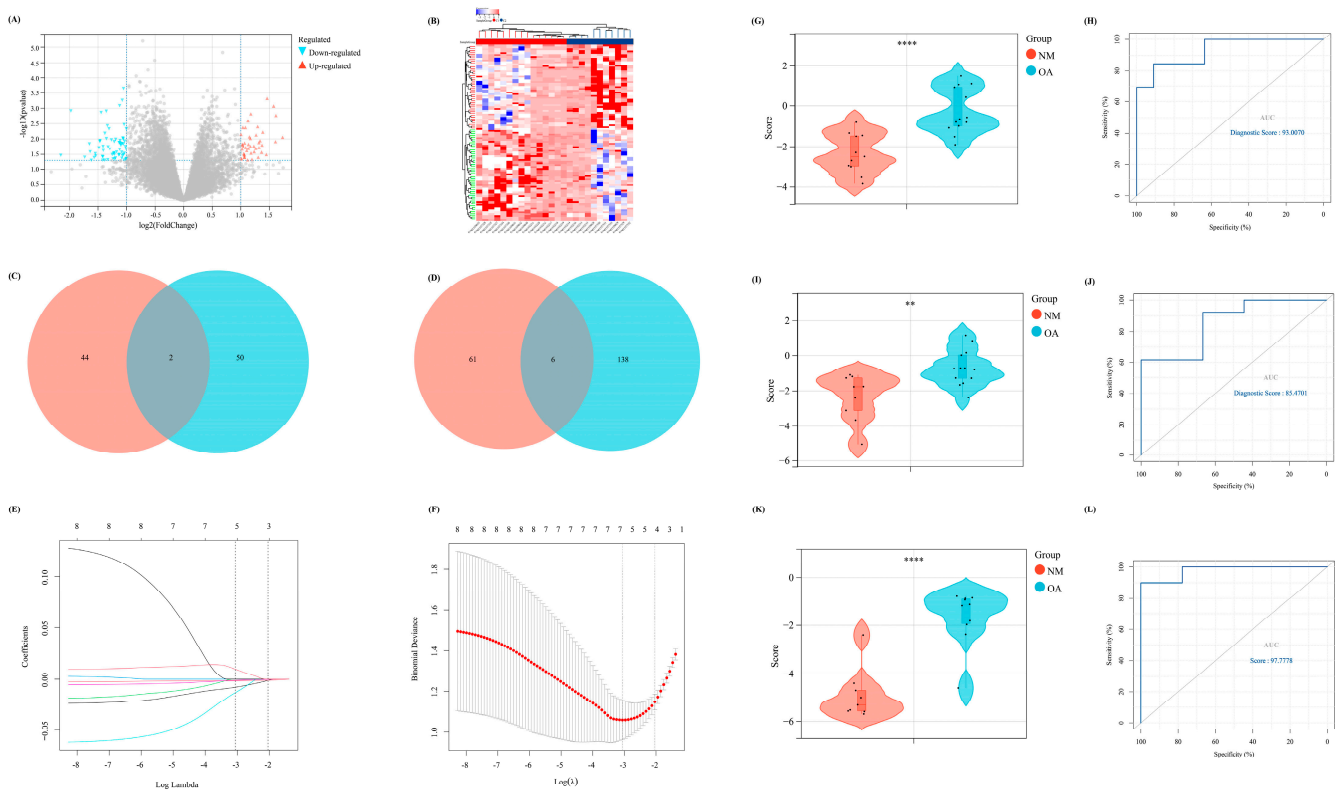
Figure 6. The construction of a diagnostic model related to m^{7}G. (**A**) The volcano map of the DEGs between two different m^{7}G clusters. (**B**) The heat map revealed the expression level of DEGs in two clusters. (**C,D**) The Venn diagram to intersect the overlapped (**C**) up-regulated and (**D**) down-regulated DEGs between the DEGs in two different clusters and DEGs between NM and OA group. (**E,F**) The results of lasso-cox regression analysis. (**G,I,K**) The difference of diagnostic score between NM and OA in our (**G**) training set, (**I**) internal verification set, and (**K**) external verification set. (**H,J,L**) The ROC curve of the diagnostic score in our (**H**) training set, (**J**) internal verification set, and (**L**) external verification set.

3.7. Validation of the Four Significant m^{7}G-Regulators in the Synovial Tissue of Patients with OA

To explore the abovementioned four significant m^{7}G-regulators (DCP2, EIF4E2, SNUPN, and LARP1) in OA, the synovial tissues of normal people and patients with OA were collected from the Second Affiliated Hospital of Nanchang University and the demographic data of the patients included in our study is listed in Table 2. Then, the expression difference in the RNA level was verified using qRT-PCR. Thus, except for LARP1, the other three regulators exhibited a significant high-expression level in OA, which was consistent with our above analysis (Figure 7A–D).

Table 2. The demographic data of patients included in our study.

| OA Samples				Normal Samples			
Age	Gender	Height (cm)	Weight (kg)	Age	Gender	Height (cm)	Weight (kg)
75	Female	154	60	43	Male	156	60
66	Male	168	66	53	Male	158	61
60	Female	155	55	53	Male	172	86
65	Male	163	70	68	Female	150	50
75	Female	150	50	54	Female	157	56
58	Female	158	65	55	Female	165	80
90	Male	160	61	54	Female	155	48
64	Female	163	63	57	Female	158	64
69	Female	155	58	56	Female	156	60
71	Female	150	55	55	Male	177	75
73	Male	178	90	47	Male	160	65
70	Female	160	45	60	Female	153	50
53	Male	171	71	61	Male	185	73
69	Female	150	50	54	Female	156	56
53	Female	150	46	53	Male	171	75
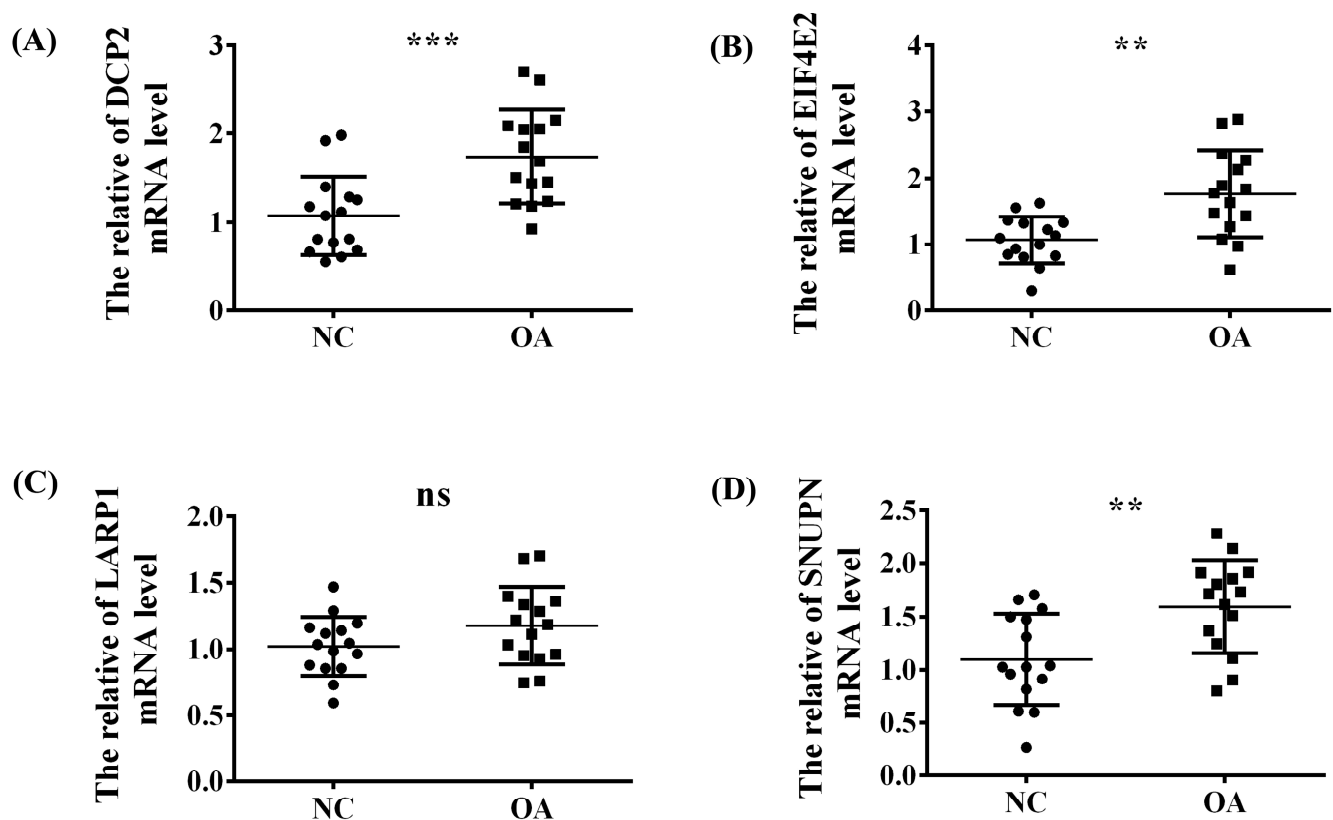
Figure 7. The validation experiment and *t*-test used to verify the expression trend of the four significant regulators (**A**) DCP2, (**B**) EIF4E2, (**C**) LARP1 and (**D**) SNUPN between the NC and OA group in the synovial tissue of our OA patients via RT-qPCR. ns, $p > 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

4. Discussion

OA is usually assumed to be caused by non-inflammatory factors; that is, a series of mechanical stresses that destroys the cartilage. However, recently, some associated inflammatory factors have also been shown to contribute to OA development, allowing inflammatory cells to infiltrate the synovium [36]. Moreover, inflammatory cytokines play an essential role in the progression of OA by stimulating the production of matrix

metalloproteinases and thus increasing matrix degradation [37]. An increasing number of studies have focused on the effects of nucleic acid site changes on the cell function and even body activities, wherein RNA modification plays a critical role. This can be seen in the methylation of m^{7}G, which is located in the inner part of tRNA and rRNA and plays a significant role in coordinating numerous functions during the mRNA lifecycle. This is largely accredited to the protein factors that particularly bind to the cap structures, the cap-binding complex in the nucleus, and eIF4E in the cytoplasm. m^{7}G is accountable for mRNA processing and nuclear export in the nucleus and is needed for effective pre-mRNA splicing. In vivo, mRNA interacts with protein factors throughout its life cycle and also plays a role in transcription termination and exosome degradation [38]. In the present study, we have briefly discussed the association between m^{7}G and OA. Several research gaps still exist despite a growing interest in this field. Thus, we have attempted to create a novel prospect for the clinical diagnosis of OA.

In this study, three datasets were downloaded from the Gene Expression Omnibus database, namely GSE55235, GSE55457, and GSE55584. After unified treatment, 18 regulatory factors associated with m^{7}G were identified and analyzed from the data of 26 patients with OA and 20 normal people. Nine statistically significant m^{7}G regulators were obtained; their association was strong enough. Second, a protein-protein interaction network was constructed to enrich and analyze these regulators using GO and KEGG. The results showed that these genes were mainly involved in no MTG-related dephosphorylation, RNA metabolism, RNA modification, and other processes and were enriched in hypoxia-related GFR signaling, insulin metabolism-related, and virus-related pathways. Next, two machine learning algorithms (RF and SVM-RFE) were used to further screen the nine genes, obtain the intersection, and further obtain four key genes, EIF4E2, DCP2, SNUPN, and LARP1. Based on these four genes, the diagnostic effectiveness of these four genes was further verified using the LASSO Cox regression. Furthermore, based on the expression levels of the abovementioned four hub genes, the 26 OA samples in the training set were divided into two different m^{7}G-related clusters; the PCA algorithm was used to further calculate the m^{7}G score to differentiate the two subtypes. Immune infiltration analysis demonstrated that cluster A was more closely associated with the immune system. Moreover, to better exhibit the characteristic of our m^{7}G-related clusters, a diagnostic model was constructed for calculating the diagnostic score using the differential expression analysis and LASSO Cox regression. Finally, the differential expression of the abovementioned four genes in the synovial tissues of patients with OA was verified using an external validation set and RT-qPCR. Although unilaterally LARP1 exhibited a non-statistically significant difference between NM and OA in RT-qPCR, the m^{7}G-related Score to distinguish two different m^{7}G-related clusters of OA exhibited an accurate efficiency and the diagnostic model constructed via these four m^{7}G-regulators played an extraordinary role in the diagnosis of OA.

EIF4E2 is a protein-coding gene located in the p-body and belongs to part of the mRNA cap-binding active complex [39,40]. In the early stage of initiation, EIF4E2 recognizes and binds m^{7}G mRNA cap, activates ubiquitin protein ligase-binding activity, and participates in miRNA-mediated translational inhibition. Contrary to EIF4E, EIF4E2 is unable to bind EIF4G (EIF4G1, EIF4G2, or EIF4G3), signifying that it assembles EIF4F by competing with EIF4E and blocking it in the cap region [41]. EEIF4E2-related diseases include casket-Siris syndrome 2, melanoma, cancer, viruses, and so on. The related pathways include the PI3K-Akt signaling and innate immune system pathways. RNA-binding and ubiquitin protein ligase binding are the GO annotations associated with this gene. Shaohong Chen et al. proposed that TNRC6 competes with EIF4E1 to recruit EIF4E2 for targeting mRNA, thus blocking translation initiation. Moreover, EIF4E2 mainly inhibits the expression of genes at the translational level but does not significantly affect the level of coding mRNA [42]. Moreover, Mir-29b is associated with EIF4E2. However, Mir-29b can silence premature AID expression in naive B cells, thus reducing the probability of inappropriate and potentially dangerous deamination activity [43]. Cadherin-22, a cell-cell adhesion molecule, is upregulated by promoter eIF4E2-mediated mTORC1-independent translational

control during hypoxia; the novel function of cadherin-22 acts as a hypoxia-specific cell surface molecule and is involved in cancer cell migration, invasion, and adhesion [44]. However, no previous studies have noted and discussed the association between EIF4E2 and OA or the role of EIF4E2 in OA, which may become a potential topic for studies in the future.

The protein encoded by DCP2 is a key component of the mRNA uncoating complex required for mRNA degradation. It removes the 7-methylguanine cap structure from mRNA and then degrades from the 5′ end. The involved pathways are an unfolded protein response and the regulation of activated PAK-2P34 by proteasome-mediated degradation. The GO annotations associated with this gene include RNA binding and manganese ion binding. Moreover, T cell intracellular antigen-1 (TIA-1)-induced transformation silencing promotes the decay of selected mRNAs, and TIA-1-mediated decay is inhibited by small interfering RNAs targeting the 5′-3′ (e.g., DCP2) or 3′-5′ (e.g., exosomal component Rrp46) decay pathway, suggesting that TIA-1 sensitizes mRNA to both major decay pathways [45]. Interestingly, there is no study on the association between DCP2 and OA. This is also the novelty of our study, and it is worthy of more detailed research.

SNUPN is also a protein-coding gene, acts as a U snRNP-specific nuclear import adapter, and participates in the trimethylguanosine cap-dependent nuclear import of U snRNPs. SNUPN is associated with chronic lymphocytic leukemia, cancer, and so on. The related pathways include the translocation of pre-mRNA containing introns and SLBP-independent mature mRNA. The GO annotations associated with this gene include outdated protein transporter activity and RNA cap binding. Moreover, XPO1 binding to various proteins is mediated by the recognition of leucine-rich nuclear export signals at the N-terminus of SNUPN, thus transporting proteins. Overexpression or dysfunction of XPO1 has been reported in different cancers [46]. As the role of SNUPN in OA has not been explored by previous studies, our study is the first one to demonstrate a link between the two. To understand the role of SNUPN in OA in detail, further deeper research is necessary.

LARP1 is a class of RNA-binding proteins that regulates the translation of specific target mRNAs downstream of the mTORC1 complex and plays a role in growth signaling and nutrient availability while regulating cell growth and proliferation [47]. The diseases associated with LARP1 are dengue virus and hepatocellular carcinoma. The pathways involved include disease and SARS-CoV-2 infection. The GO annotations associated with this gene include RNA binding and translation initiation factor binding. Furthermore, RNMT selectively regulates the LARP1 target (TOP mRNA in the terminal polypyrimidine tract) expression. Increased ribosome abundance leads to the upregulation of RNMT for coordinating mRNA capping and processing and increasing translational capacity during T-cell activation [48]. Meanwhile, the association between LARP1 and OA is still not studied. Our study is the first one to show that LARP1 may have an impact on the pathogenesis of OA, but it still needs further exploration.

Furthermore, up to now, the diagnosis of OA has become clearer and accurate, which is based on X-rays and clinical symptoms [17,49,50]. However, with OA development, the joint pain caused by OA and its effect on the daily life and exercise capacity of patients becomes more and more severe. Thus, early and timely diagnosis of OA is urgent and necessary. In our study, the nomogram constructed based on the GSE32317 further revealed that our four regulators (DCP2, EIF4E2, SNUPN, and LARP1) also showed excellent accuracy to distinguish early- and end-stage OA, which provided effective insight to the early diagnosis of OA.

In summary, in this study, the collected data sets were unified and merged into a whole for analysis to avoid homogeneity. The number of samples used was also significant. Simultaneously, we used two different machine learning methods to avoid the one-sidedness of screening methods and make the results more convincing. Furthermore, the scoring model of m^7G was constructed for the first time, which is progressive. However, not enough synovial tissue samples were collected. Moreover, the role of our four m^7G-regulators in the occurrence and development of OA was not explored with detailed experiments, which

needs to be explored in the future. Moreover, the five-gene diagnostic model obtained via our m^{7}G-related cluster was verified only via external and internal datasets without experiments; thus, its diagnostic accuracy needs to be further validated. In conclusion, we constructed an m^{7}G-related scoring model, which can significantly differentiate patients with OA, and correlated it with different statuses of the immune microenvironment, based on which we constructed a diagnostic model to diagnose patients with OA.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3 390/biom13030539/s1. Table S1: The gene sets of 28 kinds of immune cells utilized for single sample gene set enrichment analysis; Table S2: The primer pairs utilized in Real-Time Quantitative PCR.

Author Contributions: Conceptualization, L.H. and S.C.; methodology, Y.W. and X.S.; validation, J.C.; formal analysis, L.H.; investigation, X.S.; data curation, Y.W. and J.C.; writing—original draft preparation, L.H., Y.W. and J.C.; writing—review and editing, S.C. and X.S.; visualization, X.S., Y.W. and J.C.; supervision, S.C.; project administration, S.C.; funding acquisition, L.H. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Second Affiliated Hospital of Nanchang University Medical Research Ethics Committee (No. Review [2022] No. (031) and 23 March 2022).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study. Written informed consent has been obtained from the patient(s) to publish this paper.

Data Availability Statement: The original data used in this project can be downloaded in the public database GEO (https://www.ncbi.nlm.nih.gov/geo/ accessed on 31 August 2022).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Loeser, R.F.; Goldring, S.R.; Scanzello, C.R.; Goldring, M.B. Osteoarthritis: A disease of the joint as an organ. *Arthritis. Rheum.* **2012**, *64*, 1697–1707. [CrossRef]
2. Geneen, L.J.; Moore, R.A.; Clarke, C.; Martin, D.; Colvin, L.A.; Smith, B.H. Physical activity and exercise for chronic pain in adults: An overview of Cochrane Reviews. *Cochrane Database Syst. Rev.* **2017**, *4*, CD011279. [CrossRef]
3. Singh, J.A.; Noorbaloochi, S.; MacDonald, R.; Maxwell, L.J. Chondroitin for osteoarthritis. *Cochrane Database Syst. Rev.* **2015**, *1*, CD005614. [CrossRef] [PubMed]
4. Yoon, K.H.; Tak, D.H.; Ko, T.S.; Park, S.E.; Nam, J.; Lee, S.H. Association of fibrosis in the infrapatellar fat pad and degenerative cartilage change of patellofemoral joint after anterior cruciate ligament reconstruction. *Knee* **2017**, *24*, 310–318. [CrossRef] [PubMed]
5. Culvenor, A.G.; Cook, J.L.; Collins, N.J.; Crossley, K.M. Is patellofemoral joint osteoarthritis an under-recognised outcome of anterior cruciate ligament reconstruction? A narrative literature review. *Br. J. Sports Med.* **2013**, *47*, 66–70. [CrossRef]
6. Hugle, T.; Geurts, J. What drives osteoarthritis?—Synovial versus subchondral bone pathology. *Rheumatology* **2017**, *56*, 1461–1471. [CrossRef]
7. Sellam, J.; Berenbaum, F. The role of synovitis in pathophysiology and clinical symptoms of osteoarthritis. *Nat. Rev. Rheumatol.* **2010**, *6*, 625–635. [CrossRef] [PubMed]
8. Zhang, L.; Li, M.; Li, X.; Liao, T.; Ma, Z.; Zhang, L.; Xing, R.; Wang, P.; Mao, J. Characteristics of sensory innervation in synovium of rats within different knee osteoarthritis models and the correlation between synovial fibrosis and hyperalgesia. *J. Adv. Res.* **2022**, *35*, 141–151. [CrossRef]
9. Schulze-Tanzil, G. Intraarticular Ligament Degeneration Is Interrelated with Cartilage and Bone Destruction in Osteoarthritis. *Cells* **2019**, *8*, 990. [CrossRef]
10. Wang, Q.; Zhu, W.; Cheng, Y.; Wu, X. Comparative study of female Chinese mitten crabs based on their sizes and weights. *J. Food Sci. Technol.* **2022**, *59*, 2572–2583. [CrossRef]
11. Zargar, H.; Zargar-Shoshtari, K.; Lotan, Y.; Shah, J.B.; van Rhijn, B.W.; Daneshmand, S.; Spiess, P.E.; Black, P.; Collaborators. Final Pathological Stage after Neoadjuvant Chemotherapy and Radical Cystectomy for Bladder Cancer-Does pT0 Predict Better Survival than pTa/Tis/T1? *J. Urol.* **2016**, *195*, 886–893. [CrossRef] [PubMed]

12. Batushansky, A.; Zhu, S.; Komaravolu, R.K.; South, S.; Mehta-D'souza, P.; Griffin, T.M. Fundamentals of OA. An initiative of Osteoarthritis and Cartilage. Obesity and metabolic factors in OA. *Osteoarthr. Cartil.* **2022**, *30*, 501–515. [CrossRef]

13. Glyn-Jones, S.; Palmer, A.J.; Agricola, R.; Price, A.J.; Vincent, T.L.; Weinans, H.; Carr, A.J. Osteoarthritis. *Lancet* **2015**, *386*, 376–387. [CrossRef]

14. Mobasheri, A. The future of osteoarthritis therapeutics: Emerging biological therapy. *Curr. Rheumatol. Rep.* **2013**, *15*, 385. [CrossRef] [PubMed]

15. Hawker, G.A. Osteoarthritis is a serious disease. *Clin. Exp. Rheumatol.* **2019**, *37* (Suppl. S120), 3–6. [PubMed]

16. Kiadaliri, A.A.; Lohmander, L.S.; Moradi-Lakeh, M.; Petersson, I.F.; Englund, M. High and rising burden of hip and knee osteoarthritis in the Nordic region, 1990–2015. *Acta Orthop.* **2018**, *89*, 177–183. [CrossRef]

17. Braun, H.J.; Gold, G.E. Diagnosis of osteoarthritis: Imaging. *Bone* **2012**, *51*, 278–288. [CrossRef]

18. Jonkhout, N.; Tran, J.; Smith, M.A.; Schonrock, N.; Mattick, J.S.; Novoa, E.M. The RNA modification landscape in human disease. *RNA* **2017**, *23*, 1754–1769. [CrossRef]

19. Yang, D.; Qiao, J.; Wang, G.; Lan, Y.; Li, G.; Guo, X.; Xi, J.; Ye, D.; Zhu, S.; Chen, W.; et al. N6-Methyladenosine modification of lincRNA 1281 is critically required for mESC differentiation potential. *Nucleic Acids Res.* **2018**, *46*, 3906–3920. [CrossRef]

20. Zhang, H.; Shi, X.; Huang, T.; Zhao, X.; Chen, W.; Gu, N.; Zhang, R. Dynamic landscape and evolution of m6A methylation in human. *Nucleic Acids Res.* **2020**, *48*, 6251–6264. [CrossRef]

21. Yang, X.; Yang, Y.; Sun, B.F.; Chen, Y.S.; Xu, J.W.; Lai, W.Y.; Li, A.; Wang, X.; Bhattarai, D.P.; Xiao, W.; et al. 5-methylcytosine promotes mRNA export-NSUN2 as the methyltransferase and ALYREF as an m(5)C reader. *Cell Res.* **2017**, *27*, 606–625. [CrossRef]

22. Luo, Y.; Yao, Y.; Wu, P.; Zi, X.; Sun, N.; He, J. The potential role of N(7)-methylguanosine (m7G) in cancer. *J. Hematol. Oncol.* **2022**, *15*, 63. [CrossRef] [PubMed]

23. Chen, X.; Gong, W.; Shao, X.; Shi, T.; Zhang, L.; Dong, J.; Shi, Y.; Shen, S.; Qin, J.; Jiang, Q.; et al. METTL3-mediated m(6)A modification of ATG7 regulates autophagy-GATA4 axis to promote cellular senescence and osteoarthritis progression. *Ann. Rheum. Dis.* **2022**, *81*, 87–99. [CrossRef] [PubMed]

24. Yang, J.; Zhang, M.; Yang, D.; Ma, Y.; Tang, Y.; Xing, M.; Li, L.; Chen, L.; Jin, Y.; Ma, C. m(6)A-mediated upregulation of AC008 promotes osteoarthritis progression through the miR-328-3p–AQP1/ANKH axis. *Exp. Mol. Med.* **2021**, *53*, 1723–1734. [CrossRef] [PubMed]

25. Malbec, L.; Zhang, T.; Chen, Y.S.; Zhang, Y.; Sun, B.F.; Shi, B.Y.; Zhao, Y.L.; Yang, Y.; Yang, Y.G. Dynamic methylome of internal mRNA N(7)-methylguanosine and its regulatory role in translation. *Cell Res.* **2019**, *29*, 927–941. [CrossRef]

26. Edgar, R.; Domrachev, M.; Lash, A.E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* **2002**, *30*, 207–210. [CrossRef] [PubMed]

27. Taminau, J.; Meganck, S.; Lazar, C.; Steenhoff, D.; Coletta, A.; Molter, C.; Duque, R.; de Schaetzen, V.; Weiss Solis, D.Y.; Bersini, H.; et al. Unlocking the potential of publicly available microarray data using inSilicoDb and inSilicoMerging R/Bioconductor packages. *BMC Bioinform.* **2012**, *13*, 335. [CrossRef] [PubMed]

28. Ritchie, M.E.; Phipson, B.; Wu, D.; Hu, Y.; Law, C.W.; Shi, W.; Smyth, G.K. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **2015**, *43*, e47. [CrossRef] [PubMed]

29. Tomikawa, C. 7-Methylguanosine Modifications in Transfer RNA (tRNA). *Int. J. Mol. Sci.* **2018**, *19*, 4080. [CrossRef]

30. Wilkerson, M.D.; Hayes, D.N. ConsensusClusterPlus: A class discovery tool with confidence assessments and item tracking. *Bioinformatics* **2010**, *26*, 1572–1573. [CrossRef]

31. Kanehisa, M.; Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **2000**, *28*, 27–30. [CrossRef]

32. Bindea, G.; Mlecnik, B.; Hackl, H.; Charoentong, P.; Tosolini, M.; Kirilovsky, A.; Fridman, W.H.; Pages, F.; Trajanoski, Z.; Galon, J. ClueGO: A Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* **2009**, *25*, 1091–1093. [CrossRef]

33. Zhao, C.; Roy, S. Special issue: Cilia and ciliopathies. *J. Cell Physiol.* **2022**, *237*, 2611–2612. [CrossRef]

34. Zhang, G.Z.; Wu, Z.L.; Li, C.Y.; Ren, E.H.; Yuan, W.H.; Deng, Y.J.; Xie, Q.Q. Development of a Machine Learning-Based Autophagy-Related lncRNA Signature to Improve Prognosis Prediction in Osteosarcoma Patients. *Front. Mol. Biosci.* **2021**, *8*, 615084. [CrossRef] [PubMed]

35. Barbie, D.A.; Tamayo, P.; Boehm, J.S.; Kim, S.Y.; Moody, S.E.; Dunn, I.F.; Schinzel, A.C.; Sandy, P.; Meylan, E.; Scholl, C.; et al. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature* **2009**, *462*, 108–112. [CrossRef] [PubMed]

36. Nedunchezhiyan, U.; Varughese, I.; Sun, A.R.; Wu, X.; Crawford, R.; Prasadam, I. Obesity, Inflammation, and Immune System in Osteoarthritis. *Front. Immunol.* **2022**, *13*, 907750. [CrossRef] [PubMed]

37. Woodell-May, J.E.; Sommerfeld, S.D. Role of Inflammation and the Immune System in the Progression of Osteoarthritis. *J. Orthop. Res.* **2020**, *38*, 253–257. [CrossRef]

38. Ramanathan, A.; Robb, G.B.; Chan, S.H. mRNA capping: Biological functions and applications. *Nucleic Acids Res.* **2016**, *44*, 7511–7526. [CrossRef]

39. Amaya Ramirez, C.C.; Hubbe, P.; Mandel, N.; Bethune, J. 4EHP-independent repression of endogenous mRNAs by the RNA-binding protein GIGYF2. *Nucleic Acids Res.* **2018**, *46*, 5792–5808. [CrossRef]

40. Peter, D.; Weber, R.; Sandmeir, F.; Wohlbold, L.; Helms, S.; Bawankar, P.; Valkov, E.; Igreja, C.; Izaurralde, E. GIGYF1/2 proteins use auxiliary sequences to selectively bind to 4EHP and repress target mRNA expression. *Genes Dev.* **2017**, *31*, 1147–1161. [CrossRef]

41. Rosettani, P.; Knapp, S.; Vismara, M.G.; Rusconi, L.; Cameron, A.D. Structures of the human eIF4E homologous protein, h4EHP, in its m7GTP-bound and unliganded forms. *J. Mol. Biol.* **2007**, *368*, 691–705. [CrossRef]

42. Chen, S.; Gao, G. MicroRNAs recruit eIF4E2 to repress translation of target mRNAs. *Protein Cell* **2017**, *8*, 750–761. [CrossRef]

43. Recaldin, T.; Hobson, P.S.; Mann, E.H.; Ramadani, F.; Cousins, D.J.; Lavender, P.; Fear, D.J. miR-29b directly targets activation-induced cytidine deaminase in human B cells and can limit its inappropriate expression in naive B cells. *Mol. Immunol.* **2018**, *101*, 419–428. [CrossRef]

44. Kelly, N.J.; Varga, J.F.A.; Specker, E.J.; Romeo, C.M.; Coomber, B.L.; Uniacke, J. Hypoxia activates cadherin-22 synthesis via eIF4E2 to drive cancer cell migration, invasion and adhesion. *Oncogene* **2018**, *37*, 651–662. [CrossRef]

45. Yamasaki, S.; Stoecklin, G.; Kedersha, N.; Simarro, M.; Anderson, P. T-cell intracellular antigen-1 (TIA-1)-induced translational silencing promotes the decay of selected mRNAs. *J. Biol. Chem.* **2007**, *282*, 30070–30077. [CrossRef]

46. Jain, P.; Kanagal-Shamanna, R.; Wierda, W.; Keating, M.; Sarwari, N.; Rozovski, U.; Thompson, P.; Burger, J.; Kantarjian, H.; Patel, K.P.; et al. Clinical and molecular characteristics of XPO1 mutations in patients with chronic lymphocytic leukemia. *Am. J. Hematol.* **2016**, *91*, E478–E479. [CrossRef] [PubMed]

47. Burrows, C.; Abd Latip, N.; Lam, S.J.; Carpenter, L.; Sawicka, K.; Tzolovsky, G.; Gabra, H.; Bushell, M.; Glover, D.M.; Willis, A.E.; et al. The RNA binding protein Larp1 regulates cell division, apoptosis and cell migration. *Nucleic Acids Res.* **2010**, *38*, 5542–5553. [CrossRef] [PubMed]

48. Galloway, A.; Kaskar, A.; Ditsova, D.; Atrih, A.; Yoshikawa, H.; Gomez-Moreira, C.; Suska, O.; Warminski, M.; Grzela, R.; Lamond, A.I.; et al. Upregulation of RNA cap methyltransferase RNMT drives ribosome biogenesis during T cell activation. *Nucleic Acids Res.* **2021**, *49*, 6722–6738. [CrossRef]

49. Tack, A.; Preim, B.; Zachow, S. Fully automated Assessment of Knee Alignment from Full-Leg X-Rays employing a "YOLOv4 And Resnet Landmark regression Algorithm" (YARLA): Data from the Osteoarthritis Initiative. *Comput. Methods Programs Biomed.* **2021**, *205*, 106080. [CrossRef] [PubMed]

50. Prieto-Alhambra, D.; Judge, A.; Javaid, M.K.; Cooper, C.; Diez-Perez, A.; Arden, N.K. Incidence and risk factors for clinically diagnosed knee, hip and hand osteoarthritis: Influences of age, gender and osteoarthritis affecting other joints. *Ann. Rheum. Dis.* **2014**, *73*, 1659–1664. [CrossRef] [PubMed]