In-silico Identification of Novel Drug Target for Osteoarthritisinhuman using System Network Biology Approaches

Neha Srivastava[1*], Nityendra Shukla[2], Aditya Trivedi[3], Prachi Srivastava[2], , Prahlad Kishore Seth[1]

1. Bioinformatics Centre, Biotech Park, Sector-G Jankipuram, Kursi Road, Lucknow 226021, Uttar Pradesh, India.
2. AMITY Institute of Biotechnology, AMITY University Uttar Pradesh Lucknow, UP, India-226028.
3. Department of Biochemistry, University of Lucknow, UP, India-226028.

*Email id: ns011982@gmail.com

Abstract

Osteoarthritis (OA) is the most common form of joint disability in the world affecting a large number of persons s yet the mechanisms responsible for the disease is not well n understood. And therefore there is a lack of disease-modifying treatment options. It has several risk factors from systemic (e.g. age, sex, genetics, obesity) to biochemical factors (e.g. joint injury, muscle weakness, sport). The prevalence of OA is ever increasing due to the obesity epidemic and longevity. Since OA has strong genetic predisposition, in the study we attempted system network biology approach to identify a key candidate gene in a protein-protein interaction (PPI) network of OA,which may play an important role in disease pathogenesis and help us to understand the development and progression of the disease This information will help in target specific development of new molecules which may eventually lead to curative solutions for OA in human.

Keywords: Osteoarthritis, Joint disability,Genetic predisposition, System network biology, PPI network etc.

1. INTRODUCTION

Osteoarthritis (OA) is one of the most common lifestyle diseases, and the primary cause of pain and disability in the elderly population [1]. It is multifactorial in nature with contributing factors that include sex, genes, obesity, age and past joint-injury [2]. It is a whole joint disorder characterised by the articular cartilage degradation of the synovium, alteration of the subchondral bone, synovium and other connective tissues of the joint [3].

Several Genome wide association studies (GWAS) and other experimental evidence, suggest that OA has strong genetic predisposition [4,5]. There is also a strong evidence to support that obesity is linked with OA, but due to occurrence of OA on non-weight bearing joints, a possible metabolic link has also been suggested [6,7].

The aetiology of OA as well as its pathology is not completely understood and perhaps due to this there is a lack of disease modifying treatments. The possible factors suggested are exposure of joint tissue to contributors of oxidative stress, such as dyslipidemia, reactive oxygen species (ROS) and nitric oxide (NO)which are produced by chondrocytes [8,9].Alarmins and pro-inflammatory cytokines, such as TNF, IL-1β, IL-6, IL-15, IL-17, IL-18, IL-21 also play a role in inducing low-grade inflammation [10, 11]. These could possibly serve as therapeutic targets in future. Studies of autophagy in OA have also indicated that suppression of themammalian target of rapamycin(mTOR) signalling pathway may be beneficial in improving the metabolic component of OA [12]. Growth factors such as TGF-β and TGF-α/EGFR signalling pathways also play an important role in the development of OA [13,14]. Suppression of these pathways may result in delaying of tissue degradation.

Management of osteoarthritis is currently limited to awareness about the disease, weight loss, physiotherapy and pharmacological interventions to manage and/or alleviate symptoms. Non-steroidal anti-inflammatory drugs (NSAIDs) are used as first-line agents [15]. Cyclooxygenase-II (COX-II) inhibitors such as celecoxib and rofecoxib [16], opioids or intra-articular steroids are used in the treatment protocol on failure of first-line agents[17]. There is great opportunity for developing new disease-modifying OA drugs (DMOADs). Ttwo drug classes have shown strong therapeutic effect with the bone as the target. Strontium ranelate has shown promise intwo clinical trials, exhibiting significant improvements in function, pain and stiffness at 2g/day dosage [18]. Beneficial effects in bone marrow lesions (BML) and cartilage volume loss (CVL) have been reported in a SEKOIA (SrRan Efficacy in Knee Osteoarthritis triAl) study utilizing strontium ranelate [19].Another drug, Sprifermin has shown promise as following its administration; a decrease in cartilage loss as well as increase

in cartilage thickness was observed in patients of knee OA as assessed using MRI in a post-hoc analysis [20].

Systems biologyoffers possibility of identifying genes that may play a major or an accompanying role in the development of OA. Protein protein interaction (PPI) network providesa visual view of how different proteins interact and communicate in the development of disease phenotypes [21]. Furthermore, analysis of these central proteins can provide understanding of their biological processes and roles in disease development [22], which may further result in better understanding the pathology that drives the disease, and lead into better treatment and curative solutions.

2. MATERIALS AND METHODS

2.1 Data Collection

A total of 450 genes suspected to play a role in OA were mined from the NCBI GenBank database and selected for further study.

2.2 Construction of PPI Network

We used *STRING 11.0* (Search Tool for the Retrieval of Interacting Genes/Proteins) database to construct a protein-protein interaction network, utilising 450 retrieved genes. STRING provides functional association on the interactions of input proteins, with the source of evidence being text-mining, co-expression, databases, experimental and genomic data (gene fusion, neighbourhood and co-occurrence) [23]. All of these parameters can be selected individually and analysis can be performed accordingly to the user's needs. The PPI network we constructed by input of 450 genes was set on the highest confidence score (0.9) with only experimental data being considered. The disconnected nodes were removed in order to make the network cleaner [24].

Further, the constructed network was analyzed using Cytoscape v3.7.1. A tool used for analysis of PPI networks, visualisation, data integration, interactive network construction [25]. Cytoscape v.3.7.1 uses the plug-in NetworkAnalyzer to analyze the imported or generated network. It uses two parameters, Node Degree and Betweenness Centrality (BC) as the standard for analysis. Node degree refers to the number of protein interactions with the

central protein [26]. Proteins with a large number of node degree are known as hub proteins and studies have shown that these proteins may play an important role in disease development. Betweenness Centrality is the measure of the total number of shortest paths through the node [27].

2.3 **Construction of OA-related PPI network via MCODE**

We also utilized the tool Molecular Complex Detection(MCODE) which is a clustering algorithm applied to find densely clustered regions in PPI networks that may constitute molecular complexes [28]. MCODE cluster analysis was performed using Cytoscape to identify the densest clusters based on clustering scores.

2.4 **Functional enrichment study**

This was performed using Gene Ontology (GO) database consisting of description of gene and gene product knowledge and functionality. It is primarily composed of two components; gene ontology which provides a structure of the biological processes and GO annotations, which are statements that link a gene product to an oncology term based on evidence [29]. Functional enrichment analysis is done to identify genes that are over-represented in the PPI network and may have an association with the diseased phenotypes [30].

3. RESULTS

3.1. PPI Network Analysis

The PPI network was constructed by inserting 450 genes of OA in *STRING 11.0* database. The network study reveals 65 interactions (edges) between 57 proteins (nodes) based on experimental data withthe highest confidence (0.900) score as a network parameter (Fig. **1**). Network was further analyzed in Cytoscape v3.7.1 using the *NetworkAnalyzer* plug-in. Based on network topological parameter BC and node degree with overall network cut-off value BC > 0.01 and node degree >5, CDC5L (cell division cycle 5-like) is obtained as key genes in network with highest betweeness centrality and node degree. (Fig**. (2)**) (Table **1**)

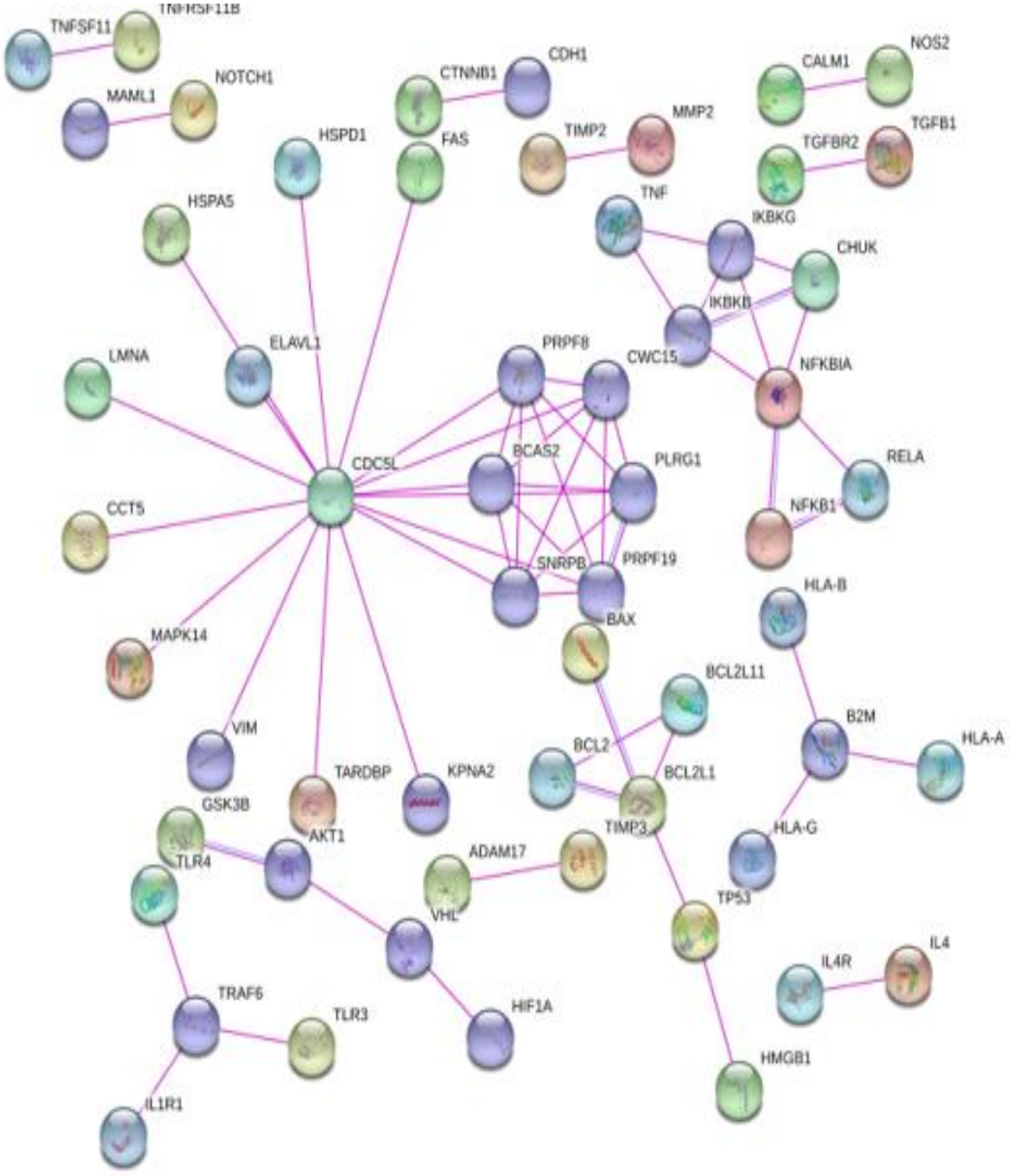

Fig. (1). Construction of PPI network of Osteoarthritis using STRING 11.0. 65 interactions between 57 proteins were observed at the highest (0.9) confidence score.

Table 1. Result ofobtained Key Genes in Network based on Topological parameter node degree and Betweenness Centrality of the top 10 genes.

Gene Name	Node Degree	Betweenness Centrality (BC)
CDC5L	17	0.58519481
IKBKG	11	0.27689703
IKBKB	11	0.2376175
AKT1	6	0.570845
PLRG1	6	0.2208991
BCAS2	6	0.012387
PRPF19	6	0.0056399
PRPF8	6	0.0036519

Table 2.MCODE cluster results and their score values.

Gene Name	MCODE_Node_Status	MCODE_Score (Double)
CDC5L	Seed	7.0
SNRPB	Clustered	7.0
PRPF19	Clustered	7.0
CWC15	Clustered	7.0
PRPF8	Clustered	7.0
BCAS2	Clustered	7.0
PLRG1	Clustered	7.0

Further the densely clustered network was constructedusing *Cytoscape_3.7.1* plug-in *MCODE*of whole network.Based on network scoring (degree cut-off value: 2 and K core: 2) parameter, CDC5L (cell division cycle 5-like) is obtained as Seed gene with MCODE score: 7.0.The topological properties of the protein interaction network via the MCODE cluster (Fig. **3**) analysis tool was studied to understand the probable biological associated processes (Table **2.**).

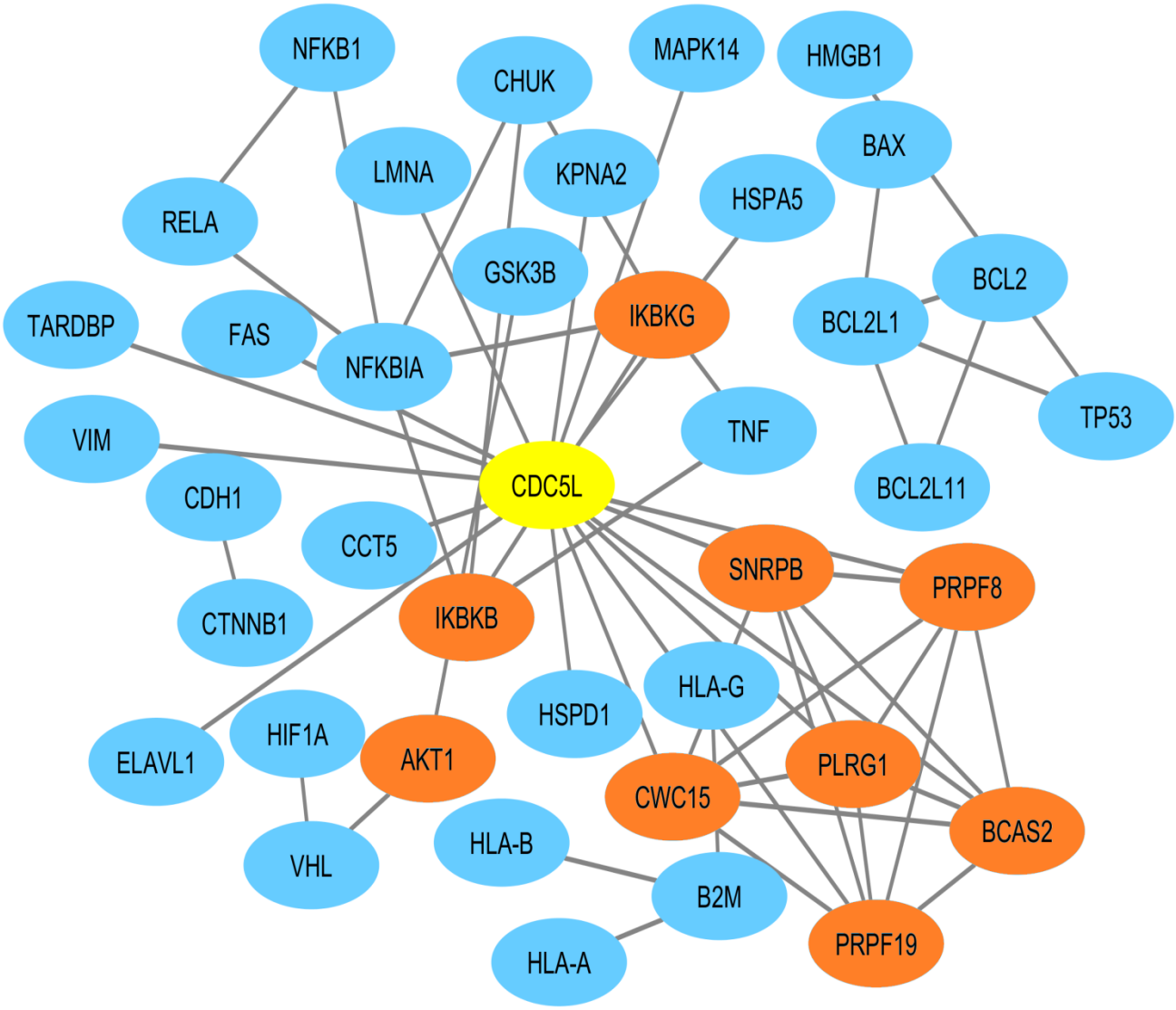

Fig. (2). Overview of PPI Network analyzed using Cytoscape_3.7.1 plug-in *Network Analyzer*. Network includes 65 edges (interaction) between 57 nodes respectively. The node

with yellow and orange color represents the key genes in the network with overall network cut off value BC > 0.01 and node degree >5. Among key genes, the node with yellow color represents the superhub gene with highest betweeness centrality and node degree.

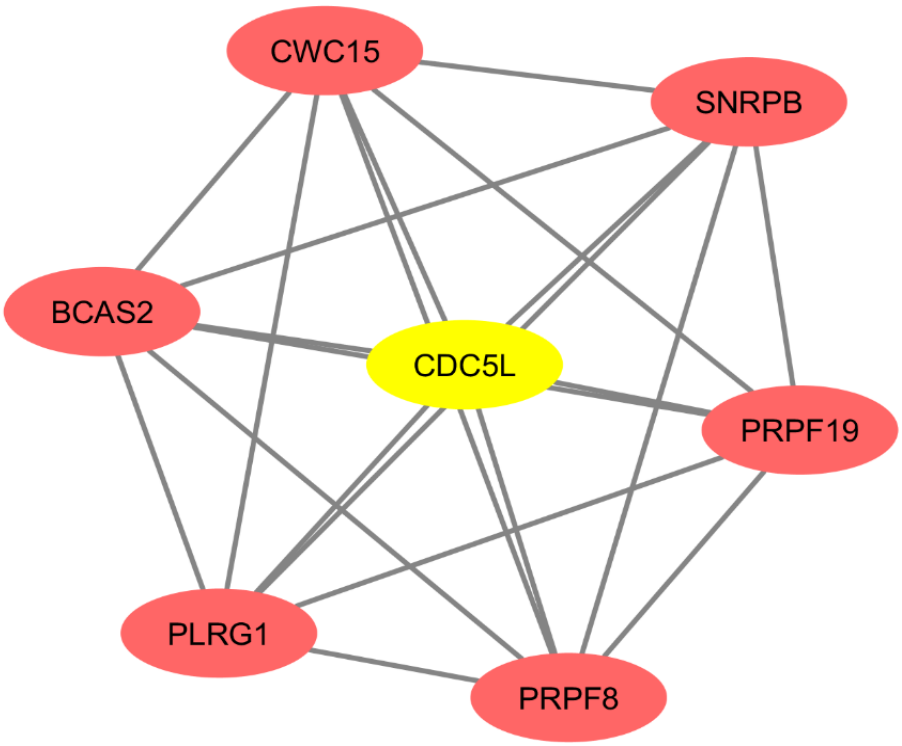

Fig. (3). The cluster network generated using *Cytoscape_3.7.1* plug-in *MCODE* tool. The node showing yellow color represent seed gene with score: 7.0.

Functional Enrichment Analysis

Functional enrichment analysis was done using the GO Enrichment Analysis tool linked to the PANTHER classification systemto detect the presence of the top 10 highest scoring genes on the basis of their Node Degree and BC in various biological processes and their roles, and how they play a part in the OA pathology. The results (Table **3**.) showed the involvement of CDC5L in almost all vital biological processes such as nitrogen metabolism (GO: 0034641) and gene expression (GO: 0010467) indicating that CDC5L may be involved in OA development.

4. DISCUSSION & CONCLUSION

In spite of numerous studies, the possible key genes responsible for OA are not known. Using in-silico tools we have identified one candidate gene, CDC5L as a key gene which may play promising role in the development of OA. CDC5L is a protein that acts as a positive regulator during the G2/M transition phase in the cell cycle. It is involved in DNA binding and is a component of the PRP19/CDC5L complex. It is also involved in major pathways such as mRNA splicing and gene expression which could indicate an error in transcription or translation in OA patients. Interestingly, a GWAS study reported an SNP, rs10948172 residing between the loci of the CDC5L and SUPT3H (suppressor of Ty3 homolog) genes in European males patients suffering with hip and knee OA [31] indicating a possible role of the gene in OA pathogenesis.

Table 3. Functional enrichment analysis result, along with their P-value and the genes associated in the mentioned various biological processes.

GO Accession Number	**P-Value**	**Associated Genes**
GO:0034641	5.81E-03	[CDC5L;CWC15;AKT1;PRPF8;PLRG1; SNRPB; PRPF19; BCAS2; IKBKB]
GO: 0006396	1.90E-04	[CDC5L;CWC15;PRPF8;PLRG1;SNRPB; PRPF19; BCAS2]
GO: 0010467	2.26E-03	[CDC5L;CWC15; AKT1; PRPF8; PLRG1; SNRPB; PRPF19; BCAS2]
GO: 0006397	2.84E-06	[CDC5L;CWC15;PRPF8;PLRG1;SNRPB; PRPF19; BCAS2]
GO: 0000398	1.00E-07	[CDC5L;CWC15;PRPF8;PLRG1;SNRPB; PRPF19; BCAS2]

Presence of another SNP, rs10948155 involved in the enhancer regions which may possibly regulate the gene expression of the RUNX2 (Runt-related transcription factor) gene has also been reported [32].The RUNX2 gene is located 500kb away from the CDC5L/SUPT3H genes and is a vital transcription factor being involved in chondrocyte hypertrophy and osteoblast differentiation [33, 34]. This further indicates role of CDC5L in OA pathogenesis. Based on these, our in-silico study suggests a possible role of CDC5L gene in OA

pathogenesis and its under-expression may causedown regulation of the RUNX2 activity leading to degradation of bone and cartilage density in OA patients. Since OA has strong genetic predisposition, further studies in wider populations are needed.

Acknowledgments: The financial support provided by DBT, G.O.I to Bioinformatics centre is gratefully acknowledged. PKS thanks National Academy of Sciences India (NASI), Allahabad, for financial support in the form of NASI Senior Scientist Platinum Jubilee Fellowship.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No animals or humans were used for this study and are not the basis of this research.

REFERENCES

1. Picavet H, Hazes J. “Prevalence of self-reported musculoskeletal diseases is high”. Ann Rheum Dis. 2003 Jul; 62(7): 644-650.
2. Bijlsma JW, Berenbaum F, Lafeber FP. “Osteoarthritis: an update with relevance for clinical practice”. Lancet. 2011 Jun; 377(9783): 2115-26.
3. Uhalte EC, Wilkinson JM, Southam L, Zeggini E. “Pathways to understanding the genomic aetiology of osteoarthritis”. Hum Mol Genet. 2017 Oct; 26(R2): R193-R201.
4. Zeggini E, Panoutsopoulou K, Southam L, *et. al.* “Identification of new susceptibility loci for osteoarthritis (arcOGEN): a genome-wide association study”. Lancet. 2012 Sep; 380(9844): 815-823.
5. Styrkarsdottir U,Helgason H, Sigurdsson A,*et. al.* “Whole genome sequencing identifies rare genotypes in COMP and CHADL associated with high risk of hip osteoarthritis”. Nat. Genet. 2017 May; 49(5): 801-805.
6. de Sousa EB, Junior, dos Santos GC, Duarte MEL, Neto, Moura V, Aguiar DP. “Metabolomics as a promising tool for early osteoarthritis diagnosis”. Braz J Med Biol Res. 2017; 50(11): e6485.
7. Zhuo Q, Yang W, Chen J, Wang Y. “Metabolic syndrome meets osteoarthritis”. Nat Rev Rheumatol. 2012 Dec; 8(12): 729-37.

8. Rathakrishnan C, Tiku K, Raghavan A, Tiku ML. “Release of oxygen radicals by articular chondrocytes: a study of luminol-dependent chemiluminescence and hydrogen peroxide secretion”. J Bone Miner Res. 1992 Oct; 7(10): 1139-48.
9. Henrotin Y, Deby-Dupont G, Deby C, De Bruyn M, Lamy M, Franchimont P. “Production of active oxygen species by isolated human chondrocytes”. Br J Rheumatol. 1993 Jul;32(7):562-7.
10. Goldring MB, Otero M. “Inflammation in osteoarthritis”. CurrOpinRheumatol. 2011 Sep; 23(5): 471-478.
11. Kapoor M[1], Martel-Pelletier J, Lajeunesse D, Pelletier JP, Fahmi H. “Role of proinflammatory cytokines in the pathophysiology of osteoarthritis”. Nat Rev Rheumatol. 2011 Jan; 7(1): 33-42.
12. Zhang Y, *et al.* “Cartilage-specific deletion of mTOR upregulates autophagy and protects mice from osteoarthritis”. Ann Rheum Dis. 2015 Jul; 74(7): 1432-40.
13. Shen J, Li S, Chen D. “TGF-β signaling and the development of osteoarthritis”. Bone Res. 2014; 2: 14002.
14. Appleton CT[1], Pitelka V, Henry J, Beier F. “Global analyses of gene expression in early experimental osteoarthritis”. Arthritis Rheuma. 2007 Jun; 56(6):1854-68.
15. Conaghan PG, Dickson J, Grant RL. “Care and management of osteoarthritis in adults: summary of NICE guidance”. BMJ. 2008 Mar; 336(7642): 502-503.
16. Hawkey CJ. “COX-1 and COX-2 inhibitors”. Best Pract Res Gastroenterol. 2001 Oct; 15(5): 801-20.
17. Smith SR, Deshpande BR, Collins JE, Katz JN, Losina E. “Comparative pain reduction of oral non-steroidal anti-inflammatory drugs and opioids for knee osteoarthritis: systematic analytic review”. Osteoarthritis Cartilage. 2016 Jun; 24(6): 962-72.
18. Bruyère O, Reginster JY, Bellamy N,*et. al.* “Clinically meaningful effect of strontium ranelate on symptoms in knee osteoarthritis: a responder analysis.” Rheumatology (Oxford). 2014 Aug; 53(8): 1457-64.
19. Pelletier JP, Roubille C,Raynauld JP,*et. al.* “Disease-modifying effect of strontium ranelate in a subset of patients from the Phase III knee osteoarthritis study SEKOIA using quantitative MRI: reduction in bone marrow lesions protects against cartilage loss”. Ann Rheum Dis. 2015 Feb; 74(2):422-9.
20. Eckstein F, Wirth W, Guermazi A, Maschek S, Aydemir A. “Brief report: intraarticularsprifermin not only increases cartilage thickness, but also reduces

cartilage loss: location-independent post hoc analysis using magnetic resonance imaging". Arthritis Rheumatol. 2015 Nov; 67(11): 2916-22.

21. Safari-Alighiarloo N, Taghizadeh M, Rezaei-Tavirani M, Goliaei B, Peyvandi AA. "Protein protein interaction networks (PPI) and complex diseases". Gastroenterol. Hepatol Bed Bench. 2014 Winter; 7(1): 17-31.
22. Franz M, Lopes CT, Huck G, Dong Y, Sumer O, Bader GD. "Cytoscape.js: a graph theory library for visualisation and analysis". Bioinformatics. 2016 Jan; 32(2): 309-311.
23. Szklarczyk D, Gable AL, Lyon D,*et. al.* "STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets." Nucleic Acids Res. 2019 Jan; 47 (Database issue): D607-D613.
24. Srivastava N, Mishra BN, Srivastava P. "In-Silico Identification of Drug Lead Molecule Against Pesticide Exposed-neurodevelopmental Disorders Through Network-based Computational Model Approach". Current Bioinformatics. 2019; 14, 460-467.
25. Saito R, Smoot E, Ono K, *et. al.* "A travel guide to Cytoscape plugins". Nat Methods. 2012 Nov; 9(11): 1069-1076.
26. Azuaje F, Devaux Y, Wagner DR. "Coordinated modular functionality and prognostic potential of a heart failure biomarker-driven interaction network". BMC Syst Biol. 2010; 4: 60.
27. Bader GD, Hogue CWV. "An automated method for finding molecular complexes in large protein interaction networks". BMC Bioinformatics. 2003; 4: 2.
28. Carbon S. Mungall CJ, Munoz-Torres MC, *et. al.* "Expansion of the Gene Ontology knowledgebase and resources". Nucleic Acids Res. 2017 Jan; 45 (Database issue): D331-D338.
29. Mi H, Huang X, Muruganujan A, *et. al.* "PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements". Nucleic Acids Res. 2017 Jan; 45 (Database issue): D183-D189.
30. Marchler-Bauer A,*et al*. "CDD/SPARCLE: functional classification of proteins via subfamily domain architectures.". Nucleic Acids Res. 2017; 45(D)200-3.
31. Zeggini E, Panoutsopoulou K, Southam L, *et. al.* "Identification of new susceptibility loci for osteoarthritis (arcOGEN): a genome-wide association study". Lancet. 2012 Sep; 380(9844): 815-823.

32. Castaño-Betancourt MC,Evans DS, Ramos YFM, *et. al.* “Novel Genetic Variants for Cartilage Thickness and Hip Osteoarthritis”. PLoS Genet. 2016 Oct; 12(10): e1006260.
33. Enomoto H, Enomoto-Iwamoto M, Iwamoto M, *et. al.* “Cbfa1 is a positive regulatory factor in chondrocyte maturation”. J Biol Chem. 2000 Mar; 275 (12): 8695-702.
34. Stein GS, Lian JB, van Wijnen AJ, *et. al.* “Runx2 control of organization, assembly and activity of the regulatory machinery for skeletal gene expression”. Oncogene. 2004 May; 23(24): 4315-29.