Ollscoil na Gaillimhe
University of Galway

The Role of Sialylation Inhibition in Intervertebral Disc Degeneration in Humans and Preclinical Disease

A thesis submitted

by

Aung Myat Phyo

to

The Discipline of Bioinformatics,
School of Mathematics, Statistics & Applied Mathematics
University *of* Galway

in partial fulfilment of the requirements for the degree of

M.Sc. in Biomedical Genomics

February 28th 2023

Thesis Supervisor(s): Prof.Abhay Pandit & Dr.Kieran Joyce

Abstract

Intervertebral disc degeneration is still a significant health issue in the modern world. Understanding the pathophysiology of this degeneration by post-transcriptomic data analysis may reveal the potential targeted regenerative materials. Altered N-glycome is recognised in inflammation and progression of degeneration. 3 Fax-Peracetyl Neu5Ac is applied to inhibit N-glycome modulation in human nucleus pulposus cells. With cutting-edge genomic data analysis techniques, the role of sialyation inhibition in these cells of inflammation and degeneration will be identified. This review consists of updated literature on intervertebral disc degeneration, RNA-seq analysis tools and the application of various bioinformatic software tools for networking of genes in intervertebral disc degeneration. In addition, single-cell RNA-seq, competing endogenous RNA systems, microRNAs as diagnostic tools, and critical intervertebral disc degeneration genes and networks are discussed.

Contents

List of Figures

List of Tables

Chapter 1

Biological Background

1.1 Introduction of Intervertebral Disc Degeneration

Lower backache is a familiar cause of global discomfort and disability and is linked with degenerative disc disease (DDD). It has a worldwide prevalence of 38.9 % and an enormous economic burden with direct and indirect expenses [1]. Currently, using steroidal or NSAID medications to alleviate symptoms and surgery for severe intervertebral disc degeneration (IDD) with nerve involvement is still the preferred treatment option for IDD [2]. However, it has been challenging to uncover effective treatments for DDD because of its complex pathophysiology and our insufficient comprehension of the underlying IDD phenomenon. More excellent knowledge of the structure of IDD will deliver opportunities for better treatment with less invasive and personalised medicine [3]. In recent times, advanced genetic tools have helped the comprehension of the cellular fundamental of IDD, and gene involvement is recognised as a necessary element for amelioration [4].

1.2 Overview of Anatomy and Physiologic Functions of Intervertebral Disc

The intervertebral disc (IVD) is a well-built avascular fibrocartilaginous composition connecting neighbouring vertebral bodies. Regarding its anatomical and physiological features, it supports constrained movement and force diversion while persisting in the mechanical stability of the spine. It possesses two significant structures: the external annulus fibrosus (AF) and the internal gelatinous elasticity nucleus pulposus (NP), which contains an immense concentration of aggrecan and water. Although clear demarcations exist between the AF and NP regions in

fetal and juvenile discs, they decline in adulthood.

In the fetal stage, the comprehensive cell density of the disc is moderately high; however, it declines prominently with age, specifically in the furthest zones from the periphery. Damage or terminal differentiation of these cells may deliver age-associated tissue destruction. In terms of mechanical disc properties, various disc components' tensile and elastic functions permit the equilibrium between stability and flexibility. After the compression forces from the vertebral bodies alter the fluid of NP, readjust them radially to the AF. In contrast, AF is deceived by NP pressure which allows the compressed disc, and the resilience of AF concedes to return to its shape when there is no pressure.

Various risk elements of disc degeneration include genetic susceptibility, reduced nutritional supply due to thickened endplates, atypical movement, smoking and the physiological destruction of IVD with ageing [5][6]. These factors result in loss of fundamental extracellular matrix (ECM) contents, mainly collagen and aggrecan, alternation of cellular composition, and healthy disc disorganisation. In addition, the disruption of ECM components may stimulate signalling paths which result in elevated cytokines expression (for example, interleukin [IL]-1β, tumour necrosis factor [TNF]-α, and interferon [IFN]-γ). Furthermore, threatening IVD immune cells also induce the deterioration process through operating factors of inflammation. Table (1.1) describes various grades of IDD.

Table 1.1: Thompson's morphological grades of IDD

Grade	NP	AF	End plate	Vertebral body
I	Bulging gel	Discrete fibrous lamellae	Hyaline, uniformly thick	Margins rounded
II	White fibrous tissue peripherally	Mucinous material between lamellae	Thickness irregular	Margins pointed
III	Consolidated fibrous tissue	Extensive mucinous infiltration, loss of AF–NP demarcation	Focal defects in cartilage	Early chondrophytes or osteophytes at margins
IV	Horizontal clefts parallel to end plate	Focal disruptions	Fibrocartilage extending from subchondral bone with irregularity and focal sclerosis in subchondral bone	Osteophytes < 2 mm
V	Clefts extend through NP and AF	Clefts extend through NP and AF	Diffuse sclerosis	Osteophytes > 2 mm

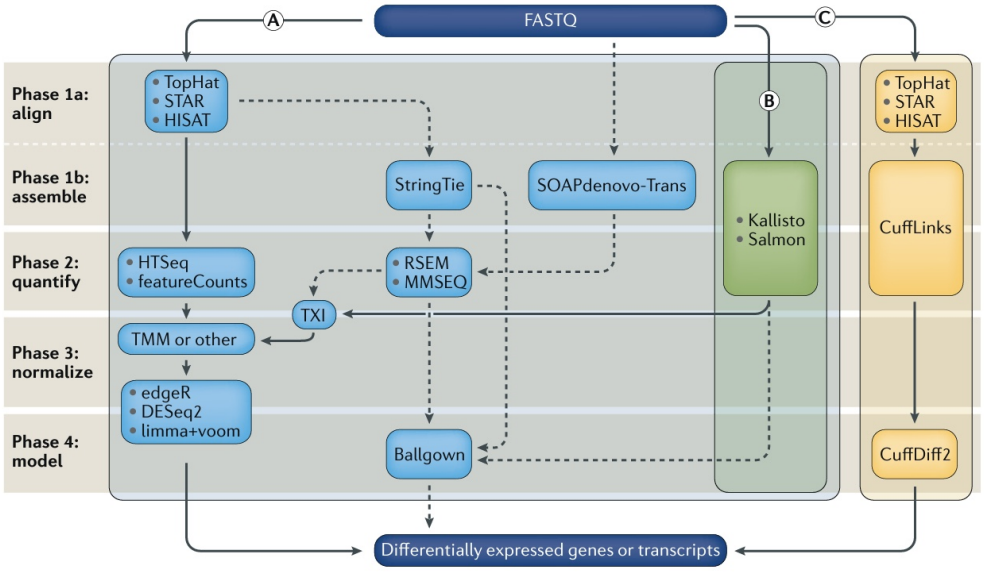
Abbreviations: AF, annulus fibrosus; IVD, intervertebral disc; NP, nucleus pulposus.

[7]

Chapter 2

RNA-seq Data Analysis

The valuable tools will vary due to the detailed biological investigation and the available computational support [8]. Despite the possibility of numerous cutoffs, our target is to explore the standard tools and techniques for assessing differentially expressed genes. There is a minimum of needed four categorical phases to accomplish (Fig.2.1).



[9]

Figure 2.1: Transcriptomic data analysis workflow

2.1 Phase 1: Reading Alignment

The beginning point of data analysis is FASTQ files [10] which consist of base-called sequencing reads. Regarding processing these files, mapping sequenced reads to recognised transcriptome or defined genome and transforming individual sequence reads to one or several genome mapping coordinates are performed. The application of definite alignment tools has achieved this phase: TopHat [11], STAR [12], or HISAT [13], which depend on a reference genome. StringTie [14] and SOAPdenovo-Trans [15] assembly tools utilise the gaps recognised in the sequencings to obtain exon boundaries and feasible splice sites. When the reference genome is probably lacking or insufficient, for example, in the compelling tumour tissue case, applying these de novo transcript assemblies is precious. Computationally intelligent alignment-free software: Sailfish [16], Kallisto [17] and Salmon [18] are established, in which the transcriptome assembly and quantification abundance are performed in a single step. Although these tools are highly demonstrative in highly abundant or long transcripts, they are limited and meticulous in low-abundant or short ones[19]. The diverse applications of mapping sequence read to transcripts influence the following expression measures[20] through the variation in how they designate a subset of the reads.

2.2 Phase 2: Quantification Level

Several studies depict that the method at the quantification phase is more effective in producing conclusive results [21][22] rather than the type of aligner [23][24]. RSEM [25], CuffLinks [26], MMSeq [27] and HTSeq [28]are popular quantification tools. RSEM assigns ambiguous reads using expectation maximisation, while Kallisto also contains data in transcript count estimates which may lead to bias consequences. The tximport [29] tool is applicable for converting transcript estimates to read count equivalents. The outcomes of this phase are merged into an expression matrix with individual expression profiles as a row and each sample as a column. The values are pure sequencing reads or approximate mapping reads.

2.3 Phase 3: Filtration and Normalisation

Filtering to eliminate traits with consistently low sequencing depth is uncomplicated and meant to enhance spotting genuine differential expression [30]. However, the normalisation techniques of an expression matrix can be complicated. There are two essential premises which are dependent on most computational normalising techniques. While the first factor is that the expression levels of most genes are unchanged among the replicate groups [31], and the variance in overall mRNA levels are not revealed in different sample groups. The normalisation technique and the trimmed mean of M-values (TMM) method [32](embraced with the edgeR [33] DGE analysis package) are frequently applied for this phase. Another perspective is the manipulation of spike-in-control RNAs [34], a collection of known concentration exogenous nucleic acid to a sample before processing. It is plausible to resemble the expression levels of the transcripts from the sample because synthetic RNAs' concentration is known and directly associated with the number of reads generated. Nevertheless, controlling spike-ins at pre-pooled levels will still be challenging [35]. Moreover, as particular isoforms are expressed at significantly various concentrations within a sample, normalisation read counts of gene level are well grounded than at the transcript level [36]. Implementing spike-in controls is broadly involved in single-cell experiments [37], although it is uncommon in presented RNA-seq DGE processes.

2.4 Phase 4: Differential Expression Modelling

Some tools manipulate the read counts of gene-level expression, while alternative ones are for transcript-level estimates. The former applications (edgeR, DESeq2 [38] and limma+voom [39]) depend on sequencing depths and apply generalised linear models, which give complicated test arrangements for calculation. On the other hand, CuffDiff, MMSEQ and Ballgown [40] are software for transcript-level estimates which need greater computational power and produce more fluctuated results [41]. Before choosing these expression tools, the assembly of sequencing reads, quantifying, filtering and normalising affect the concluding results' total variance.

Chapter 3

Application of RNA-seq Analysis in the Intervertebral Disc Degeneration

Many research articles on different RNA analyses range from mRNA and circular RNA to single-cell RNA in IDD. Few studies have investigated the transcriptome of the IVD in degeneration. Below are overviews of the seminal articles that have employed unique investigative methods to characterise the IVD.

3.1 Single-cell RNA-seq (sc-RNA-seq)

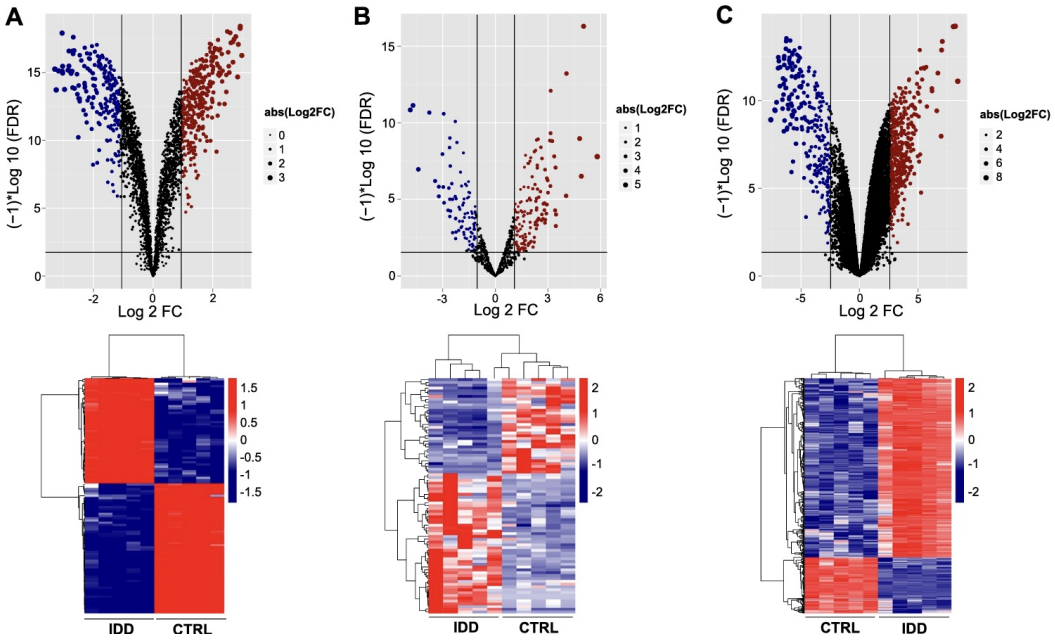
Firstly, single-cell RNA-seq (sc-RNA-seq) analysis of degeneration and non-degeneration IVD of the same individual was first characterised by revealing heterogeneity and subpopulation expression variability [42]. The authors assembled five hundred thousand NP or inner annulus fibrosus (iAF) cells of degenerative or non-degenerative discs from centrifugation and instantly carried out sc-RNA-seq analysis. Raw sequencing data of individual specimens have been transformed into matrices of expression counts through the Cell Ranger software 10X Chromium Single Cell 3. Raw BCL files from the Illumina HiSeq4000 are demultiplexed to pair-end FASTQ files by mkfastq of Cell Ranger. After that, GRCH38 serves as a human reference genome, and the following UMI count matrices are delivered to the Seurat suite as the input files [43][44]. To remove poor-quality data, they decided on a range of a minimum of 500 and a maximumly of 5000 genes detection in each cell. Canonical correlation analysis (CCA) is executed to detect shared origins of variation across datasets, succeeding SCTransform normalisation. Uniform Manifold Approximation and Projection (UMAP), an unsupervised nonlinear dimensionality-reducing method, is applied for clustering and visualis-

ing the data. The Seurat R package applies to differentially expressed genes (DEGs). Pearson correlation coefficients (PCC) are determined through the log fold-change of DEGs in each cell type. Enrichr [45] is applied with default functions regarding gene ontology (GO) analysis. The distribution pattern of GO and Kyoto Encyclopedia of Genes and Genomes (KEGG) clusters analysis is revealed via the heatmap. The prospective biomarkers of overregulation in NP cells are *C2orf40* [46], *MGP, MSMP* [47][48][49], *CHI3L1, LGALS1, ID1, ID3* and *TMED* while those in iAF cells are *MT1F* [50], *PLA2GA, EPYC, C10orf10, FGFBP2, PRELP* and *CHI3L1* [51][52]. For GO functional analysis, recognised antioxidant involvement, cytokine involvement, collagen and ECM binding, and metallopeptidase activity are upregulated common pathways in both degenerative cells. For KEGG analysis, cell-type specific signalling pathways (p53 and toll-like receptor), increased activation in programmed cell death, cytokine-cytokine receptor interaction, and Wnt signalling pathways are revealed in the degenerative samples. Although GO and KEGG analysis results are similar in both degenerative cells, there are still a few variations between them, showing the heterogeneity and their specific functionality. Finally, protein and mRNA expression of *MGST1, vimentin, SOD2* and *SYF2 (p29)* genes authenticate these sc-RNA-seq research findings.

3.2 Circular RNA-Established (circ-RNA) Competing Endogenous RNA (ceRNA) System

The following article is about developing a circular RNA-anchored (circ-RNA) competing endogenous RNA network to detect predictive markers of IDD [53]. CircRNA-miRNA-mRNA regulatory interactions or ceRNA regulatory networks constructed through various bioinformatic tools perform vital parts in the progression of many diseases, including cancers [54][55] and IDD [56][57][58]. To create a specific ceRNA network of IDD, the authors downloaded the GSE67567 dataset [59][60][61](three subsets such as GSE67566: circ-RNA expression profiles, GSE63492: miRNA expression profiles and GSE63492: mRNA-lncRNA expression profiles) from NCBI GEO database. Individual datasets consist of 5 standard/control and 5 IDD patient samples. By applying R3.4.1 Limma [62], differential expression RNA in IDD and standard batch are preconceived regarding false discovery rate <0.05 and $|\log_2 FC|$ >1 as the thresholds.

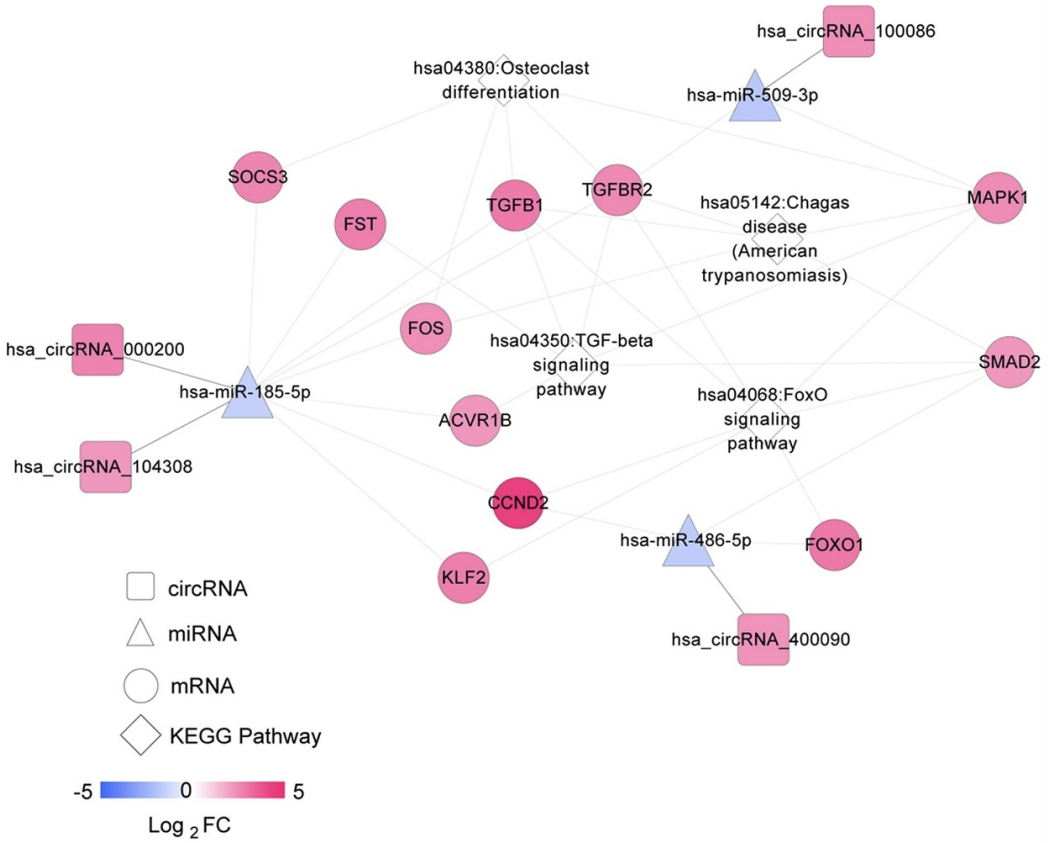
To assemble the circ-RNA and mRNA co-expression network, PCC is received by operating the 'cor.test' in R. After building the co-expression network at Cytoscape [63], mRNAs experienced GO and KEGG pathway analysis [64][65][66] by the application of DAVID software tool [67]. The connection between circRNAs and miRNAs is forecasted by miRanda [68] (Gap Open Penalty: -8, Gap Extend:-2, Score Threshold:80 % and Energy Threshold: -20). Finally, combining the circ-RNA-mRNA and miRNA-mRNA networks results in the ceRNA network. The IDD-related KEGG pathways of the Comparative Toxicogenomics Database (CTD) 2022 version [69] are manually collated with the ceRNA network pathways. After receiving the intersected RNAs in both pathways, the IDD-associated ceRNA network is obtained. GSE19943: miRNA expression profile and GSE34095: mRNA expression profile is downloaded for additional validation.



[53]

Figure 3.1: Verification of differentially expressed circ-RNAs(A), miRNAs(B) and mRNAs(C) between two groups (Patients and Healthy). Downregulation is shown as blue, and upregulation is described as red. The horizontal line shows FDR=0.05 while the perpendicular lines show $|\log_2 FC|$=1.

Regarding the results, Fig(3.1) represents the volcano plots and the heat maps. In the resulting ceRNA network, three downregulated miRNAs and four upregulated circRNAs are acquired. Moreover, four overlapping pathways (Fig.3.2) are procured, and 11 mRNAs are included. In the conclusion of this article, the authors find that not only circRNA-00200/circRNA-104308-miR-185-5p-$TGF\beta 1$ [70] play an essential role in IDD via the inflammatory mechanism but also circRNA-100086-miR-509-3p-*MAPK1* [71] axis serves as a modulator of macrophage polarisation in IDD [72]. Furthermore, circRNA-400090-miR-486-5p-*FOXO1/SMAD2* is also involved, revealing that *FoxO* is vital in the physiological process of the old IVD [73]. Recently, it can be said that *TGF-β1/SMAD2/3* pathway is the crucial part of platelet-rich plasma, one of the therapeutic options to set back IDD [74]. However, there is still a limitation, such as no suitable auxiliary validation dataset for circ-RNAs. Therefore, more advanced vivo and vitro validated experiments are needed to validate these connections in IDD.

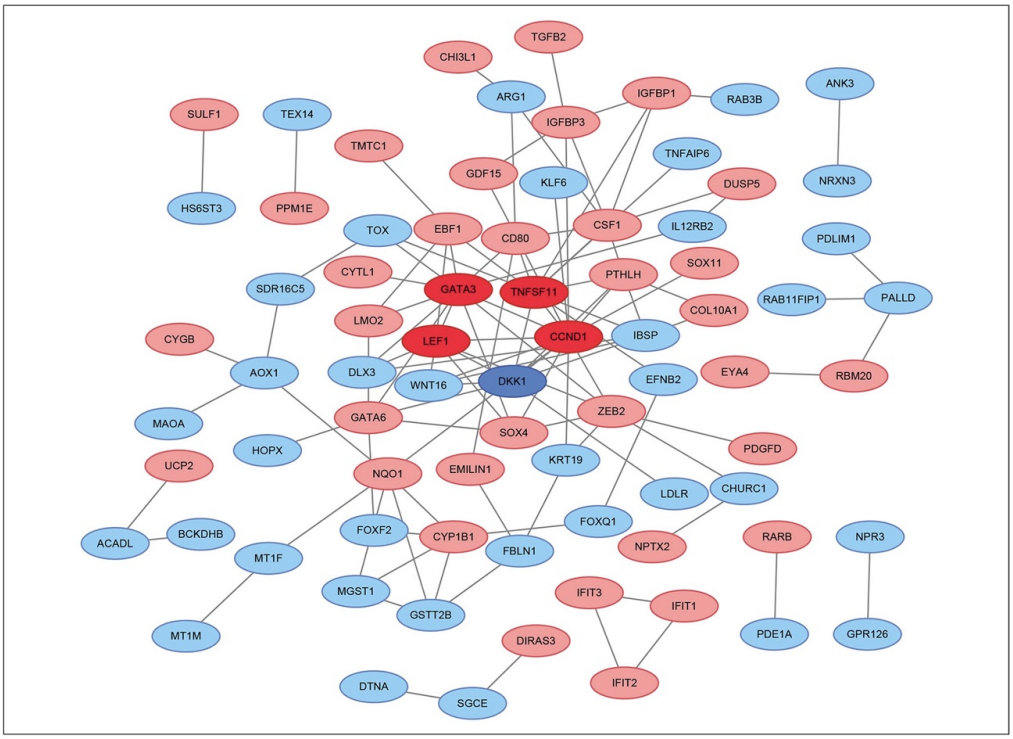


[53]

Figure 3.2: IDD-related integrated ceRNA regulatory network

3.3 Bioinformatics- Established on Fundamental Genes and Networks

Another review is about the bioinformatics-founded study on critical genes and networks of IDD pathogenesis by offering new emerging biomarkers of diagnosis and prevention [75]. Firstly, the data is retrieved from the GSE70362 dataset of the GEO database. The robust multiarray averaging (RMA) algorithm is applied for normalisation. The authors apply the Limma package for the screening of DEGs and clusterProfiler [76] for the GO and KEGG analysis, respectively. Furthermore, STRING(Search Tools for Retrieval of Interacting Genes) with a confidence score $\geq$0.4 [77] is implemented for protein-protein interaction (PPI) analysis, and the resulting PPI network figures are created by Cytoscape(Fig.3.3).



[75]

Figure 3.3: Protein-protein interaction network of DEG. Upregulation is described as red, downregulation is described as blue, and deepened colour shows the core genes

Regarding results, 112 DEGs, such as 60 upregulation and 52 downregulation genes, are identified. Positive regulation of leukocyte differentiation, [IFN]-α, BMP and BMP stimulus cellular response is recognised in GO functional analysis. These pathways are upregulated in inflammation, immune response and disc tissue reclamation. For KEGG analysis, the Hippo and Wnt signalling pathways are also detected. Moreover, five major proteins (bright red and blue in (Fig.3.3)) are identified in the PPI analysis. *GATA3* is the transcription factor of helper T cell (type2) differentiation[78]. However, there still needs to be more knowledge of the connection between Th-cell polarisation and IVD. Elevated *RANKL* levels converted by *TNFSF11* in IVD tissues reveal more severe human degeneration levels [79]. Anti-RANKL antibody treatment may be one of the options to relieve IVDD[80]. In addition, *H19/miR-22/LEF1* axis seems to offer a novel therapeutic approach to reverse NPC degeneration due to H2O2 [81]. Blocking the Wnt pathway [82]and specific inhibition of *MMP13* by *DKK1* overexpression alter the effect on the degeneration of the disc. Furthermore, the authors describe the literature on bioinformatic-based IVDD analysis in the previous five years. (Table.3.1)

Table 3.1: Studies of Gene Expression in IDD

Gene	Experimental Models	Expression	Function	References
CCL5, OPRL1	Human-NP cells	Upregulation	Related to IVDD caused by changes in the osmotic pressure of NP cells.	Zhao *et al.*[68]
CXCL8, ICAM1, RELB, NFKBIA BIRC3, TNFAIP3	Human-NP and AF cells	Upregulation Downregulation	Have a synergistic effect on TNF-α-induced IVDD development.	Hong *et al.*[69]
PHF6	Human-NP cells	Upregulation	Causes IVDD by promoting ECM degradation	Rui *et al.*[70]
SOD2	Human-whole blood	Upregulation	Prevents IVDD through antioxidant and anti-inflammatory mechanisms.	Wang *et al.*[71]
MMP2	Human- AF cells	Upregulation	Leads to IVDD by accelerating matrix degradation.	Xunlu *et al.*[72]
IL-6	Human-NP and AF cells	Upregulation	Accounts for the different processes of degeneration in AF and NP	Wang *et al.*[73]
BMP3, BMP4, GAS1, MMP13, ADAMTS5, NADPH, Nox4, P15	Rat-NP cells	Upregulation	Disrupt the redox dynamic balance of NP cells, accelerate the matrix metabolism of NP cells, and induce NP cell cycle arrest to retard cell growth in high oxygen tension.	Feng *et al.*[74]
GPX1, VEGFA, cyclin I		Downregulation		
CDKN1B, Smad4	Human-NP cells	Downregulation	Induce IVDD by affecting cell proliferation and extracellular matrix.	Hu *et al.*[75]

[75]

3.4 Detection and Characterisation of Peripheral Blood MicroRNAs (miRNAs)

First of all, the objective of this research article is the transcriptomic data analysis of peripheral blood collection in lumbar disc herniation (LDH) patients to detect circulating miRNAs which reveal significant biomarkers in the diagnostic process through various bioinformatic tools [83]. A total of 20 patients (half with LDH and another half without degenerative IVD) participated in this research. MRI results with the Pfirrmann Grading system are applied for the diagnosis confirmation of LDH among these participants. In terms of data processing, DEG screening, establishment and analysis of miRNA gene network, GO and KEGG pathways analysis, similar tools and applications are implemented in the previous article [75]. There are 17 upregulated miRNAs and 56 downregulation between the two groups for differentially expressed miRNAs. Distinctively, miR-766-3p, miR-6749-3p upregulation and miR-4632-5p downregulation may be selected as IVD degenerative process-related genes involving ECM disruption, programmed cell death, cellular inflammation and multiplication [84]. *PI3K/Akt, MAPK* and *Wnt* pathways are affected by the above miRNAs [85], which are essential for IVD degeneration pathology [86][87][88][89]. Regarding KEGG enrichment pathways analysis, apoptosis, axon guidance, *VEGF* pathway, ECM-receptor interaction pathways, *PI3K/Akt, mTOR* and *NF-KB* pathways are revealed. However, there are also limitations and challenges, such as small sample patient size, absent related age between samples, nonrandomised method and no verification of analysed data of deep sequencing with other techniques.

3.5 Importance of Transformed Glycosylation in Nucleus Pulposus Cells

In-house research has focused on the association of modified glycosylation in nucleus pulposus cells and human inflammation and degeneration [90]. Glycosylation structures intervene in critical intracellular processes, for example, protein assembly[91], receptor expression [92][93] and enzymatic function [94]. Sialylated glycans serve a critical role in cellular communication and signalling network [95] [96], while *sialyltransferases* (STs), one of the *glycosyltransferase* enzymes, are paramount for their synthesis. Sialic acid undergoes various conformations α-(2,3/6/8) when conjugated to galactose. The glycol-alteration, the augmentation of the sialic acid by STs to a terminal galactose at N- and O- glycans, are enhanced in IDD [97][98][99]. That is to say that transformed N-glycome, especially sialylated and fucosylated N-glycosylation motif expression, is a crucial part of the inflammatory IVD and its progression. 3Fax-Peracetyl Neu5Ac (Neu5Ac-inhib) is the fundamental structure to retard *sialyltransferase* enzymes for evaluation of the consequences of glycosylation inhibition [100]. To reduce the involvement of sialylation in IVD degeneration, the competitive inhibitor agent of the *sialyltransferase* enzyme is an unconventional method for glycosylation inhibition. Specific inhibition of sialyation can regulate the glycome to reduce cytokine-mediated signalling of NP cells. In this research, 3Fax-Peracetyl Neu5Ac is applied for sialyation inhibition in *vitro* in a human NP cellmodel of IDD (cytokine-induced) to discover the outcomes on glycosignature, ECM synthesis, cellular anabolism, catabolism and migration. This maintains the inflammatory NP cell glyco-phenotype, reduces destructive enzyme production and restores cell movement. This initial research explores the effect of glycosylation inhibitors on recovering normal glycosylation of NP cells. For biomaterial application, tissue incorporation and biomaterial absorption are considered crucial factors. Neu5Ac-inhib is an effective agent administered by a carrier system to treat degenerative IVD. This research shows potential for new personalised therapeutic materials for IVD degeneration soon.

Chapter 4

Project Aims

IDD is the frequent long-term pathophysiological status felt by 80 % of people at least once in their lifetime and is a challenging disability and discomfort with a tremendous burden with high socioeconomic cost. Unfortunately, although updated genetic tools and pathologic characterisation may help understand IDD's molecular process, but no disease-modifying remedies exist.

The primary intention of the project is to explore the effect of sialylation inhibition in intervertebral disc degeneration by delivering 3Fax-Peracetyl Neu5Ac in models of IVDD using generated RNA-seq datasets. The data is received from the lab of Professor Abhay Pandit and Dr Kieran Joyce at CÚRAM, SFI Research Centre for Medical devices at the University of Galway. While the first dataset is Homo Sapiens, 30x, 90Gb raw data on Illumina PE150, the preclinical canine data set was 6Gb of data per sample on a Novse1 6000, resulting in 20M paired-end 150 bp reads. First, TopHat or STAR tools will represent alignment reads and assembly, while the reference genome will be received from NCBI or Ensembl genome browser. To complete the quantification step, HTSeq or Cufflinks software will be employed. In favour of more awareness of compared transcriptomic data across all samples and groups, principal component analysis (PCA) will be operated. DESeq2 built with negative binomial distribution or EdgeR with p-value 0.05 and $|\log_2 FC| \geq 1$ will be used to analyse DEGs in each dataset. For robust DESeq2 results, apeglm will be applied. These potential results are constructed via the volcano plot and heat maps. ClusterProfiler will leverage GO functional enrichment and the KEGG pathway. The resulting genes and mRNAs across experimental groups will be applied to building networks using various bioinformatic tools. Time permitting, proteomics analysis with the STRING tool and proteomic data visualisation with Cytoscape will be done in the human data set.

This cross-species RNA-seq data analysis will help understand sialylation inhibi-

tion's role in IDD of preclinical canine models and humans. Finally, it may bring another step of new personalised regenerative biomaterials for IDD.

Bibliography

[1] D. Hoy, C. Bain, G. Williams, L. March, P. Brooks, F. Blyth, A. Woolf, T. Vos, and R. Buchbinder, "A systematic review of the global prevalence of low back pain," *Arthritis & rheumatism*, vol. 64, no. 6, pp. 2028–2037, 2012.

[2] K. T. Weber, T. D. Jacobsen, R. Maidhof, J. Virojanapa, C. Overby, O. Bloom, S. Quraishi, M. Levine, and N. O. Chahine, "Developments in intervertebral disc disease research: pathophysiology, mechanobiology, and therapeutics," *Current reviews in musculoskeletal medicine*, vol. 8, pp. 18–31, 2015.

[3] S. Wang, Y. Rui, J. Lu, and C. Wang, "Cell and molecular biology of intervertebral disc degeneration: current understanding and implications for potential therapeutic strategies," *Cell proliferation*, vol. 47, no. 5, pp. 381–390, 2014.

[4] C. Wang, W.-J. Wang, Y.-G. Yan, Y.-X. Xiang, J. Zhang, Z.-H. Tang, and Z.-S. Jiang, "Microrna: New players in intervertebral disc degeneration," *Clinica chimica acta*, vol. 450, pp. 333–341, 2015.

[5] J. E. Mayer, J. C. Iatridis, D. Chan, S. A. Qureshi, O. Gottesman, and A. C. Hecht, "Genetic polymorphisms associated with intervertebral disc degeneration," *The Spine Journal*, vol. 13, no. 3, pp. 299–317, 2013.

[6] F. M. Williams, A. T. Bansal, J. B. van Meurs, J. T. Bell, I. Meulenbelt, P. Suri, F. Rivadeneira, P. N. Sambrook, A. Hofman, S. Bierma-Zeinstra, *et al.*, "Novel genetic variants associated with lumbar disc degeneration in northern europeans: a meta-analysis of 4600 subjects," *Annals of the rheumatic diseases*, vol. 72, no. 7, pp. 1141–1148, 2013.

[7] J. T. Gravel, F. H. Abduljabbar, J. Ouellet, and L. Haglund, "Chapter 1: The human spinal disc: relevant anatomy and physiology," *Biological Approaches to Spinal Disc Repair and Regeneration for Clinicians. New York: Thieme Medical Publishers*, pp. 2–10, 2017.

[8] S. M. E. Sahraeian, M. Mohiyuddin, R. Sebra, H. Tilgner, P. T. Afshar, K. F. Au, N. Bani Asadi, M. B. Gerstein, W. H. Wong, M. P. Snyder, *et al.*, "Gaining comprehensive biological insight into the transcriptome by performing a broad-spectrum rna-seq analysis," *Nature communications*, vol. 8, no. 1, p. 59, 2017.

[9] R. Stark, M. Grzelak, and J. Hadfield, "Rna sequencing: the teenage years," *Nature Reviews Genetics*, vol. 20, no. 11, pp. 631–656, 2019.

[10] P. J. Cock, C. J. Fields, N. Goto, M. L. Heuer, and P. M. Rice, "The sanger fastq file format for sequences with quality scores, and the solexa/illumina fastq variants," *Nucleic acids research*, vol. 38, no. 6, pp. 1767–1771, 2010.

[11] D. Kim, G. Pertea, C. Trapnell, H. Pimentel, R. Kelley, and S. L. Salzberg, "Tophat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions," *Genome biology*, vol. 14, no. 4, pp. 1–13, 2013.

[12] A. Dobin, C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, and T. R. Gingeras, "Star: ultrafast universal rna-seq aligner," *Bioinformatics*, vol. 29, no. 1, pp. 15–21, 2013.

[13] D. Kim, B. Langmead, and S. L. Salzberg, "Hisat: a fast spliced aligner with low memory requirements," *Nature methods*, vol. 12, no. 4, pp. 357–360, 2015.

[14] M. Pertea, D. Kim, G. M. Pertea, J. T. Leek, and S. L. Salzberg, "Transcript-level expression analysis of rna-seq experiments with hisat, stringtie and ballgown," *Nature protocols*, vol. 11, no. 9, pp. 1650–1667, 2016.

[15] Y. Xie, G. Wu, J. Tang, R. Luo, J. Patterson, S. Liu, W. Huang, G. He, S. Gu, S. Li, *et al.*, "Soapdenovo-trans: de novo transcriptome assembly with short rna-seq reads," *Bioinformatics*, vol. 30, no. 12, pp. 1660–1666, 2014.

[16] R. Patro, S. M. Mount, and C. Kingsford, "Sailfish enables alignment-free isoform quantification from rna-seq reads using lightweight algorithms," *Nature biotechnology*, vol. 32, no. 5, pp. 462–464, 2014.

[17] N. L. Bray, H. Pimentel, P. Melsted, and L. Pachter, "Near-optimal probabilistic rna-seq quantification," *Nature biotechnology*, vol. 34, no. 5, pp. 525–527, 2016.

[18] R. Patro, G. Duggal, M. I. Love, R. A. Irizarry, and C. Kingsford, "Salmon provides fast and bias-aware quantification of transcript expression," *Nature methods*, vol. 14, no. 4, pp. 417–419, 2017.

[19] D. C. Wu, J. Yao, K. S. Ho, A. M. Lambowitz, and C. O. Wilke, "Limitations of alignment-free tools in total rna-seq quantification," *BMC genomics*, vol. 19, pp. 1–14, 2018.

[20] C. Yang, P.-Y. Wu, L. Tong, J. Phan, and M. Wang, "The impact of rna-seq aligners on gene expression estimation," in *Proceedings of the 6th ACM Conference on Bioinformatics, Computational Biology and Health Informatics*, pp. 462–471, 2015.

[21] C. Robert and M. Watson, "Errors in rna-seq quantification affect genes of relevance to human disease," *Genome biology*, vol. 16, pp. 1–16, 2015.

[22] C. R. Williams, A. Baccarella, J. Z. Parrish, and C. C. Kim, "Empirical assessment of analysis workflows for differential expression analysis of human samples using rna-seq," *BMC bioinformatics*, vol. 18, pp. 1–12, 2017.

[23] T. P. Quinn, T. M. Crowley, and M. F. Richardson, "Benchmarking differential expression analysis tools for rna-seq: normalization-based vs. log-ratio transformation-based methods," *BMC bioinformatics*, vol. 19, no. 1, pp. 1–15, 2018.

[24] M. Teng, M. I. Love, C. A. Davis, S. Djebali, A. Dobin, B. R. Graveley, S. Li, C. E. Mason, S. Olson, D. Pervouchine, *et al.*, "A benchmark for rna-seq quantification pipelines," *Genome biology*, vol. 17, no. 1, pp. 1–12, 2016.

[25] B. Li and C. N. Dewey, "Rsem: accurate transcript quantification from rna-seq data with or without a reference genome," *BMC bioinformatics*, vol. 12, pp. 1–16, 2011.

[26] C. Trapnell, A. Roberts, L. Goff, G. Pertea, D. Kim, D. R. Kelley, H. Pimentel, S. L. Salzberg, J. L. Rinn, and L. Pachter, "Differential gene and transcript expression analysis of rna-seq experiments with tophat and cufflinks," *Nature protocols*, vol. 7, no. 3, pp. 562–578, 2012.

[27] E. Turro, S.-Y. Su, Â. Gonçalves, L. J. Coin, S. Richardson, and A. Lewin, "Haplotype and isoform specific expression estimation using multi-mapping rna-seq reads," *Genome biology*, vol. 12, pp. 1–15, 2011.

[28] S. Anders, P. T. Pyl, and W. Huber, "Htseqâa python framework to work with high-throughput sequencing data," *bioinformatics*, vol. 31, no. 2, pp. 166–169, 2015.

[29] C. Soneson, M. I. Love, and M. D. Robinson, "Differential analyses for rna-seq: transcript-level estimates improve gene-level inferences," *F1000Research*, vol. 4, 2015.

[30] R. Bourgon, R. Gentleman, and W. Huber, "Independent filtering increases detection power for high-throughput experiments," *Proceedings of the National Academy of Sciences*, vol. 107, no. 21, pp. 9546–9551, 2010.

[31] D. Risso, J. Ngai, T. P. Speed, and S. Dudoit, "Normalization of rna-seq data using factor analysis of control genes or samples," *Nature biotechnology*, vol. 32, no. 9, pp. 896–902, 2014.

[32] M. D. Robinson and A. Oshlack, "A scaling normalization method for differential expression analysis of rna-seq data," *Genome biology*, vol. 11, no. 3, pp. 1–9, 2010.

[33] M. D. Robinson, D. J. McCarthy, and G. K. Smyth, "edger: a bioconductor package for differential expression analysis of digital gene expression data," *bioinformatics*, vol. 26, no. 1, pp. 139–140, 2010.

[34] T. Qing, Y. Yu, T. Du, and L. Shi, "mrna enrichment protocols determine the quantification characteristics of external rna spike-in controls in rna-seq studies," *Science China life sciences*, vol. 56, pp. 134–142, 2013.

[35] K. Chen, Z. Hu, Z. Xia, D. Zhao, W. Li, and J. K. Tyler, "The overlooked fact: fundamental need for spike-in control for virtually all genome-wide analyses," *Molecular and cellular biology*, vol. 36, no. 5, pp. 662–667, 2016.

[36] D. Leshkowitz, E. Feldmesser, G. Friedlander, G. Jona, E. Ainbinder, Y. Parmet, and S. Horn-Saban, "Using synthetic mouse spike-in transcripts to evaluate rna-seq analysis tools," *PloS one*, vol. 11, no. 4, p. e0153782, 2016.

[37] A. T. Lun, F. J. Calero-Nieto, L. Haim-Vilmovsky, B. Göttgens, and J. C. Marioni, "Assessing the reliability of spike-in normalization for analyses of single-cell rna sequencing data," *Genome research*, vol. 27, no. 11, pp. 1795–1806, 2017.

[38] M. I. Love, W. Huber, and S. Anders, "Moderated estimation of fold change and dispersion for rna-seq data with deseq2," *Genome biology*, vol. 15, no. 12, pp. 1–21, 2014.

[39] C. W. Law, Y. Chen, W. Shi, and G. K. Smyth, "voom: Precision weights unlock linear model analysis tools for rna-seq read counts," *Genome biology*, vol. 15, no. 2, pp. 1–17, 2014.

[40] A. C. Frazee, G. Pertea, A. E. Jaffe, B. Langmead, S. L. Salzberg, and J. T. Leek, "Ballgown bridges the gap between transcriptome assembly and expression analysis," *Nature biotechnology*, vol. 33, no. 3, pp. 243–246, 2015.

[41] F. Rapaport, R. Khanin, Y. Liang, M. Pirun, A. Krek, P. Zumbo, C. E. Mason, N. D. Socci, and D. Betel, "Comprehensive evaluation of differential gene expression analysis methods for rna-seq data," *Genome biology*, vol. 14, no. 9, pp. 1–13, 2013.

[42] H. Cherif, M. Mannarino, A. S. Pacis, J. Ragoussis, O. Rabau, J. A. Ouellet, and L. Haglund, "Single-cell rna-seq analysis of cells from degenerating and non-degenerating intervertebral discs from the same individual reveals new biomarkers for intervertebral disc degeneration," *International Journal of Molecular Sciences*, vol. 23, no. 7, p. 3993, 2022.

[43] A. Butler, P. Hoffman, P. Smibert, E. Papalexi, and R. Satija, "Integrating single-cell transcriptomic data across different conditions, technologies, and species," *Nature biotechnology*, vol. 36, no. 5, pp. 411–420, 2018.

[44] R. Satija, J. A. Farrell, D. Gennert, A. F. Schier, and A. Regev, "Spatial reconstruction of single-cell gene expression data," *Nature biotechnology*, vol. 33, no. 5, pp. 495–502, 2015.

[45] M. V. Kuleshov, M. R. Jones, A. D. Rouillard, N. F. Fernandez, Q. Duan, Z. Wang, S. Koplev, S. L. Jenkins, K. M. Jagodnik, A. Lachmann, *et al.*, "Enrichr: a comprehensive gene set enrichment analysis web server 2016 update," *Nucleic acids research*, vol. 44, no. W1, pp. W90–W97, 2016.

[46] Y. H. Huh, J.-H. Ryu, S. Shin, D.-U. Lee, S. Yang, K.-S. Oh, C.-H. Chun, J.-K. Choi, W. K. Song, and J.-S. Chun, "Esophageal cancer related gene 4 (ecrg4) is a marker of articular chondrocyte differentiation and cartilage destruction," *Gene*, vol. 448, no. 1, pp. 7–15, 2009.

[47] T. Mitamura, S. Pradeep, M. McGuire, S. Wu, S. Ma, H. Hatakeyama, Y. A. Lyons, T. Hisamatsu, K. Noh, A. Villar-Prados, *et al.*, "Induction of anti-vegf therapy resistance by upregulated expression of microseminoprotein (msmp)," *Oncogene*, vol. 37, no. 6, pp. 722–731, 2018.

[48] H.-J. Park, M. S. Shin, M. Kim, J. B. Bilsborrow, S. Mohanty, R. R. Montgomery, A. C. Shaw, S. You, and I. Kang, "Transcriptomic analysis of human

il-7 receptor alpha low and high effector memory cd8+ t cells reveals an age-associated signature linked to influenza vaccine response in older adults," *Aging Cell*, vol. 18, no. 4, p. e12960, 2019.

[49] M. J. VandeKopple, J. Wu, E. N. Auer, A. J. Giaccia, N. C. Denko, and I. Papandreou, "Hilpda regulates lipid metabolism, lipid droplet abundance, and response to microenvironmental stress in solid tumors," *Molecular Cancer Research*, vol. 17, no. 10, pp. 2089–2101, 2019.

[50] R. Liu, Q. Liu, K. Wang, X. Dang, and F. Zhang, "Comparative analysis of gene expression profiles in normal hip human cartilage and cartilage from patients with necrosis of the femoral head," *Arthritis Research & Therapy*, vol. 18, pp. 1–8, 2016.

[51] M. Di Rosa, M. A. Szychlinska, D. Tibullo, L. Malaguarnera, and G. Musumeci, "Expression of chi3l1 and chit1 in osteoarthritic rat cartilage model. a morphological study," *European journal of histochemistry: EJH*, vol. 58, no. 3, 2014.

[52] T. Zhao, Z. Su, Y. Li, X. Zhang, and Q. You, "Chitinase-3 like-protein-1 function and its role in diseases," *Signal transduction and targeted therapy*, vol. 5, no. 1, p. 201, 2020.

[53] B. Yu, Z. Zhu, T. Hu, J. Lu, B. Shen, T. Wu, K. Guo, S. K. Chaudhary, H. Feng, W. Zhao, *et al.*, "Construction of a circular rna-based competing endogenous rna network to screen biomarkers related to intervertebral disc degeneration," *BMC Musculoskeletal Disorders*, vol. 23, no. 1, p. 675, 2022.

[54] S. Bai, Y. Wu, Y. Yan, S. Shao, J. Zhang, J. Liu, B. Hui, R. Liu, H. Ma, X. Zhang, *et al.*, "Construct a circrna/mirna/mrna regulatory network to explore potential pathogenesis and therapy options of clear cell renal cell carcinoma," *Scientific reports*, vol. 10, no. 1, pp. 1–15, 2020.

[55] Z. Dong, Z. Liu, M. Liang, J. Pan, M. Lin, H. Lin, Y. Luo, X. Zhou, and W. Yao, "Identification of circrna–mirna–mrna networks contributes to explore underlying pathogenesis and therapy strategy of gastric cancer," *Journal of Translational Medicine*, vol. 19, no. 1, pp. 1–18, 2021.

[56] S. Cao, J. Li, K. Yang, and H. Li, "Major cerna regulation and key metabolic signature analysis of intervertebral disc degeneration," *BMC Musculoskeletal Disorders*, vol. 22, no. 1, pp. 1–11, 2021.

[57] Z. Guo, Y. Liu, Y. Gao, X. Guan, H. Li, and M. Cheng, "Circ-rna expression pattern and circ-rna-mirna-mrna network in the pathogenesis of

human intervertebral disc degeneration," *Cell Journal (Yakhteh)*, vol. 23, no. 2, p. 218, 2021.

[58] Z. Huo, H. Li, L. Tian, J. Li, K. Zhang, Z. Li, G. Li, L. Du, H. Xu, and B. Xu, "Construction of a potentially functional circrna-mirna-mrna network in intervertebral disc degeneration by bioinformatics analysis," *BioMed Research International*, vol. 2021, 2021.

[59] P.-H. Lan, Z.-H. Liu, Y.-J. Pei, Z.-G. Wu, Y. Yu, Y.-F. Yang, X. Liu, L. Che, C.-J. Ma, Y.-K. Xie, *et al.*, "Landscape of rnas in human lumbar disc degeneration," *Oncotarget*, vol. 7, no. 39, p. 63166, 2016.

[60] L.-Y. Fan and H.-Q. Wang, "Noncoding rnas in human intervertebral disc degeneration: An integrated microarray study," *Genomics data*, vol. 5, pp. 80–81, 2015.

[61] H.-Q. Wang, X.-D. Yu, Z.-H. Liu, X. Cheng, D. Samartzis, L.-T. Jia, S.-X. Wu, J. Huang, J. Chen, and Z.-J. Luo, "Deregulated mir-155 promotes fas-mediated apoptosis in human intervertebral disc degeneration by targeting fadd and caspase-3," *The Journal of pathology*, vol. 225, no. 2, pp. 232–242, 2011.

[62] M. E. Ritchie, B. Phipson, D. Wu, Y. Hu, C. W. Law, W. Shi, and G. K. Smyth, "limma powers differential expression analyses for rna-sequencing and microarray studies," *Nucleic acids research*, vol. 43, no. 7, pp. e47–e47, 2015.

[63] P. Shannon, A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski, and T. Ideker, "Cytoscape: a software environment for integrated models of biomolecular interaction networks," *Genome research*, vol. 13, no. 11, pp. 2498–2504, 2003.

[64] M. Kanehisa, "Toward understanding the origin and evolution of cellular organisms," *Protein Science*, vol. 28, no. 11, pp. 1947–1951, 2019.

[65] M. Kanehisa, M. Furumichi, Y. Sato, M. Ishiguro-Watanabe, and M. Tanabe, "Kegg: integrating viruses and cellular organisms," *Nucleic acids research*, vol. 49, no. D1, pp. D545–D551, 2021.

[66] M. Kanehisa and S. Goto, "Kegg: kyoto encyclopedia of genes and genomes," *Nucleic acids research*, vol. 28, no. 1, pp. 27–30, 2000.

[67] D. W. Huang, B. T. Sherman, and R. A. Lempicki, "Systematic and integrative analysis of large gene lists using david bioinformatics resources," *Nature protocols*, vol. 4, no. 1, pp. 44–57, 2009.

[68] D. Betel, A. Koppal, P. Agius, C. Sander, and C. Leslie, "Comprehensive modeling of microrna targets predicts functional non-conserved and non-canonical sites," *Genome biology*, vol. 11, pp. 1–14, 2010.

[69] A. P. Davis, C. J. Grondin, R. J. Johnson, D. Sciaky, J. Wiegers, T. C. Wiegers, and C. J. Mattingly, "Comparative toxicogenomics database (ctd): update 2021," *Nucleic acids research*, vol. 49, no. D1, pp. D1138–D1143, 2021.

[70] Z. Qu, Z. Quan, Q. Zhang, Z. Wang, Q. Song, X. Zhuang, C. Fu, F. Xu, Y. Liu, Y. Wang, *et al.*, "Comprehensive evaluation of differential lncrna and gene expression in patients with intervertebral disc degeneration," *Molecular Medicine Reports*, vol. 18, no. 2, pp. 1504–1512, 2018.

[71] Z. Li, X. Wang, H. Pan, H. Yang, X. Li, K. Zhang, H. Wang, Z. Zheng, H. Liu, and J. Wang, "Resistin promotes ccl4 expression through toll-like receptor-4 and activation of the p38-mapk and nf-κb signaling pathways: implications for intervertebral disc degeneration," *Osteoarthritis and cartilage*, vol. 25, no. 2, pp. 341–350, 2017.

[72] C. Yang, P. Cao, Y. Gao, M. Wu, Y. Lin, Y. Tian, and W. Yuan, "Differential expression of p38 mapk α, β, γ, δ isoforms in nucleus pulposus modulates macrophage polarization in intervertebral disc degeneration," *Scientific reports*, vol. 6, no. 1, pp. 1–12, 2016.

[73] O. Alvarez-Garcia, T. Matsuzaki, M. Olmer, K. Miyata, S. Mokuda, D. Sakai, K. Masuda, H. Asahara, and M. K. Lotz, "Foxo are required for intervertebral disk homeostasis during aging and their deficiency promotes disk degeneration," *Aging Cell*, vol. 17, no. 5, p. e12800, 2018.

[74] H. Yang, C. Yuan, C. Wu, J. Qian, Q. Shi, X. Li, X. Zhu, and J. Zou, "The role of tgf-β1/smad2/3 pathway in platelet-rich plasma in retarding intervertebral disc degeneration," *Journal of cellular and molecular medicine*, vol. 20, no. 8, pp. 1542–1549, 2016.

[75] Q. Yan, Q. Xiao, J. Ge, C. Wu, Y. Wang, H. Yu, H. Yang, and J. Zou, "Bioinformatics-based research on key genes and pathways of intervertebral disc degeneration," *Cartilage*, vol. 13, no. 2_suppl, pp. 582S–591S, 2021.

[76] G. Yu, L.-G. Wang, Y. Han, and Q.-Y. He, "clusterprofiler: an r package for comparing biological themes among gene clusters," *Omics: a journal of integrative biology*, vol. 16, no. 5, pp. 284–287, 2012.

[77] D. Szklarczyk, A. L. Gable, D. Lyon, A. Junge, S. Wyder, J. Huerta-Cepas, M. Simonovic, N. T. Doncheva, J. H. Morris, P. Bork, *et al.*, "String v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets," *Nucleic acids research*, vol. 47, no. D1, pp. D607–D613, 2019.

[78] J. Zhu, "Transcriptional regulation of th2 cell differentiation," *Immunology and cell biology*, vol. 88, no. 3, pp. 244–249, 2010.

[79] N. Takegami, K. Akeda, J. Yamada, T. Sano, K. Murata, J. Huang, K. Masuda, and A. Sudo, "Rank/rankl/opg system in the intervertebral disc," *Arthritis research & therapy*, vol. 19, no. 1, pp. 1–10, 2017.

[80] M. Sato, K. Inage, Y. Sakuma, J. Sato, S. Orita, K. Yamauchi, Y. Eguchi, N. Ochiai, K. Kuniyoshi, Y. Aoki, *et al.*, "Anti-rankl antibodies decrease cgrp expression in dorsal root ganglion neurons innervating injured lumbar intervertebral discs in rats," *European Spine Journal*, vol. 24, pp. 2017–2022, 2015.

[81] X. Wang, M. Zou, J. Li, B. Wang, Q. Zhang, F. Liu, and G. Lü, "Lncrna h19 targets mir-22 to modulate h2o2-induced deregulation in nucleus pulposus cell senescence, proliferation, and ecm synthesis through wnt signaling," *Journal of cellular biochemistry*, vol. 119, no. 6, pp. 4990–5002, 2018.

[82] X. Huang, L. Zhong, E. van Helvoort, F. Lafeber, S. Mastbergen, J. Hendriks, J. N. Post, and M. Karperien, "The expressions of dickkopf-related protein 1 and frizzled-related protein are negatively correlated to local inflammation and osteoarthritis severity," *Cartilage*, vol. 12, no. 4, pp. 496–504, 2021.

[83] S. Cui, Z. Zhou, X. Liu, R. G. Richards, M. Alini, S. Peng, S. Liu, X. Zou, Z. Li, and S. Grad, "Identification and characterization of serum micrornas as biomarkers for human disc degeneration: An rna sequencing analysis," *Diagnostics*, vol. 10, no. 12, p. 1063, 2020.

[84] X. Zhou, L. Chen, S. Grad, M. Alini, H. Pan, D. Yang, W. Zhen, Z. Li, S. Huang, and S. Peng, "The roles and perspectives of micrornas as biomarkers for intervertebral disc degeneration," *Journal of Tissue Engineering and Regenerative Medicine*, vol. 11, no. 12, pp. 3481–3487, 2017.

[85] B. Zhao, Q. Yu, H. Li, X. Guo, and X. He, "Characterization of microrna expression profiles in patients with intervertebral disc degeneration," *International journal of molecular medicine*, vol. 33, no. 1, pp. 43–50, 2014.

[86] A. Hiyama, K. Yokoyama, T. Nukaga, D. Sakai, and J. Mochida, "A complex interaction between wnt signaling and tnf-α in nucleus pulposus cells," *Arthritis research & therapy*, vol. 15, no. 6, pp. 1–11, 2013.

[87] A. Hiyama, K. Yokoyama, T. Nukaga, D. Sakai, and J. Mochida, "Response to tumor necrosis factor-α mediated inflammation involving activation of prostaglandin e2 and wnt signaling in nucleus pulposus cells," *Journal of Orthopaedic Research*, vol. 33, no. 12, pp. 1756–1768, 2015.

[88] D. Miao and L. Zhang, "Leptin modulates the expression of catabolic genes in rat nucleus pulposus cells through the mitogen-activated protein kinase and janus kinase 2/signal transducer and activator of transcription 3 pathways," *Molecular medicine reports*, vol. 12, no. 2, pp. 1761–1768, 2015.

[89] X. Zhang, Z. Hu, J. Hao, and J. Shen, "Low intensity pulsed ultrasound promotes the extracellular matrix synthesis of degenerative human nucleus pulposus cells through fak/pi3k/akt pathway," *Spine*, vol. 41, no. 5, pp. E248–E254, 2016.

[90] K. Joyce, I. Mohd Isa, A. Krouwels, L. Creemers, A. Devitt, and A. Pandit, "The role of altered glycosylation in human nucleus pulposus cells in inflammation and degeneration," *Eur Cell Mater*, vol. 41, pp. 401–20, 2021.

[91] H. Komekado, H. Yamamoto, T. Chiba, and A. Kikuchi, "Glycosylation and palmitoylation of wnt-3a are coupled to produce an active form of wnt-3a," *Genes to Cells*, vol. 12, no. 4, pp. 521–534, 2007.

[92] J. da Silva Correia and R. J. Ulevitch, "Md-2 and tlr4 n-linked glycosylations are important for a functional lipopolysaccharide receptor," *Journal of Biological Chemistry*, vol. 277, no. 3, pp. 1845–1854, 2002.

[93] Y.-C. Liu, H.-Y. Yen, C.-Y. Chen, C.-H. Chen, P.-F. Cheng, Y.-H. Juan, C.-H. Chen, K.-H. Khoo, C.-J. Yu, P.-C. Yang, *et al.*, "Sialylation and fucosylation of epidermal growth factor receptor suppress its dimerization and activation in lung cancer cells," *Proceedings of the National Academy of Sciences*, vol. 108, no. 28, pp. 11332–11337, 2011.

[94] Y. J. Kim and A. Varki, "Perspectives on the significance of altered glycosylation of glycoproteins in cancer," *Glycoconjugate journal*, vol. 14, pp. 569–576, 1997.

[95] A. Varki, "Glycan-based interactions involving vertebrate sialic-acid-recognizing proteins," *Nature*, vol. 446, no. 7139, pp. 1023–1029, 2007.

[96] X. Chen and A. Varki, "Advances in the biology and chemistry of sialic acids," *ACS chemical biology*, vol. 5, no. 2, pp. 163–176, 2010.

[97] E. Collin, M. Kilcoyne, S. White, S. Grad, M. Alini, L. Joshi, and A. Pandit, "Unique glycosignature for intervertebral disc and articular cartilage cells and tissues in immaturity and maturity," *Scientific reports*, vol. 6, no. 1, pp. 1–12, 2016.

[98] Z. Kazezian, Z. Li, M. Alini, S. Grad, and A. Pandit, "Injectable hyaluronic acid down-regulates interferon signaling molecules, igfbp3 and ifit3 in the bovine intervertebral disc," *Acta biomaterialia*, vol. 52, pp. 118–129, 2017.

[99] I. L. Mohd Isa, S. A. Abbah, M. Kilcoyne, D. Sakai, P. Dockery, D. P. Finn, and A. Pandit, "Implantation of hyaluronic acid hydrogel prevents the pain phenotype in a rat model of intervertebral disc injury," *Science Advances*, vol. 4, no. 4, p. eaaq0597, 2018.

[100] C. D. Rillahan, A. Antonopoulos, C. T. Lefort, R. Sonon, P. Azadi, K. Ley, A. Dell, S. M. Haslam, and J. C. Paulson, "Global metabolic inhibitors of sialyl-and fucosyltransferases remodel the glycome," *Nature chemical biology*, vol. 8, no. 7, pp. 661–668, 2012.