doi: 10.13241/j.cnki.pmb.2017.04.042

·生物信息学·

椎间盘退变纤维环及髓核异常表达基因的比较及生物信息学分析

万中元[1] 宋 芳[2] 任大江[1] 单建林[1] 赵广民[1] 李 放[1Δ]

(1 北京军区总医院骨科 北京 100700;2 第二炮兵总医院口腔科 北京 100088)

摘要 目的:通过对已公开发表的基因芯片表达谱数据进行研究,探究椎间盘退变过程中纤维环与髓核组织的基因表达差异,并采用生物信息学方法对差异进行分析。**方法:** 经 GEO 数据库选取两组椎间盘退变相关的基因芯片表达谱数据 GSE23130 及 GSE67567,GSE23130 所研究标本来源于正常及退变纤维环组织,GSE67567 标本来源于正常及退变髓核组织。对上述数据系列进行质量分析,GSE23130 及 GSE67567 各有 10 例样本数据被纳入实验。采用 GeneSpring 13.0 软件对 GSE23130 正常及退变纤维环间差异表达基因及 GSE67567 正常及退变髓核间差异表达基因分别进行筛选,利用 KEGG PATHWAY 和 DAVID 功能注释簇集分析分别对 GSE23130 及 GSE67567 上调及下调基因进行生物信息学分析。**结果:**GSE23130 及 GSE67567 各筛选出差异表达基因 3182 个和 3017 个, 其中 135 个基因在上述两个基因表达谱数据中均存在差异表达。针对两组数据进行的 KEGG PATHWAY 分析发现 TGF-beta signaling pathway 和 regulation of apoptosis 等数个相同的生物学通路及 DAVID 功能注释簇集;此外,还发现了数个与 GSE23130 及 GSE67567 单独相关的 DAVID 功能注释簇集。**结论:**椎间盘退变过程中纤维环及髓核组织内基因表达情况存在差异,两种组织内发生的生物过程不尽相同。某些生物学过程在两种组织内均出现异常改变,这些生物学过程中的异常变化可能是椎间盘退变的关键环节,值得进行深入研究。

关键词:椎间盘退变;纤维环;髓核;基因芯片;生物信息学

中图分类号:R681.57 **文献标识码:**A **文章编号:**1673-6273(2017)04-761-05

Comparison and Bioinformatics Analysis of the Aberrantly Expressed Genes between Degenerated Annulus Fibrous and Nucleus Pulpous

WAN Zhong-yuan[1], SONG Fang[2], REN Da-jiang[1], SHAN Jian-lin[1], ZHAO Guang-min[1], LI Fang[1Δ]

(1 Department of orthopedics of General Hospital of Beijing Military Command, Beijing, 100700, China;

2 Department of stomatology of General Hospital of Second Artillery Corps, Beijing, 100088, China)

ABSTRACT Objective: In this paper, two released microarray datasets were employed for the comparison of the gene expression profiles between annulus fibrous and nucleus pulpous in intervertebral disc degeneration (IDD). Several bioinformatics analysis were also applied for the differentially expressed genes. **Methods:** Two IDD related microarray datasets GSE23130 and GSE67567 were selected from the Gene Expression Omnibus. Samples of GSE23130 were derived from normal and degenerated annulus fibrous. GSE67567 samples were collected from normal and degenerative nucleus pulpous tissue. Quality filtering were applied for the data series. 10 samples of GSE23130 and GSE67567 were included respectively in the current study. GeneSpring 13.0 software were applied to screen the differentially expressed genes in GSE23130 and GSE67567 respectively. KEGG PATHWAY analysis and DAVID functional annotation clustering analysis were performed for the up-regulated and down-regulated genes respectively in two datasets. **Results:** While 3,182 and 3,017 differentially expressed genes were found in GSE23130 and GSE67567 respectively, only 135 of which were both differentially expressed in two datasets. Several KEGG biological pathways such as hsa03040:Spliceosome and hsa04350:TGF-beta signaling pathway and a number of DAVID functional annotation clusters such as nuclear lumen and the regulation of apoptosis were both noted in GSE23130 and GSE67567. However, several DAVID functional annotation clusters were also found relative to GSE23130 and GSE67567 respectively alone. **Conclusions:** The current study showed differences of gene expression profile between annulus fibrous and nucleus pulpous in IDD. Subsequent bioinformatics analysis indicated that biological processed in the two tissue were not all the same. Biological processes which were aberrantly changed in both tissue might be the key links in IDD deserved further investigating.

Key words: Intervertebral disc degeneration; Annulus fibrous; Nucleus pulpous; Microarray; Bioinformatics

Chinese Library Classification(CLC): R681.57 Document code: A

Article ID: 1673-6273(2017)04-761-05

作者简介:万中元(1983-),博士研究生,医师,主要研究方向:椎间盘退变机制研究,电话:18810638921,

E-mail:wanzhongyuan@yeah.net

Δ 通讯作者:李放,E-mail:FANGL6722@sina.com

(收稿日期:2015-10-11 接受日期:2015-10-23)

前言

椎间盘退变导致的颈椎病、腰椎间盘突出症等脊柱退行性疾病是影响中老年人生活质量的重要慢性疾病，这些疾病不仅会引起严重的躯体痛苦，还给患者带来了沉重的经济负担[1,2]。椎间盘组织由外层纤维环、中心区髓核及上、下软骨终板所构成。在正常情况下，髓核发挥了承载轴向负荷的作用，并且将其转化为外周向的张力负荷，纤维环则将这些应力吸收，并维持椎间盘的稳定[3]。纤维环及髓核在结构和功能方面的差异决定了椎间盘退变时二者的病理改变的不同，但到目前为止此差异内在的基因背景尚不明了。

基因芯片作为一种高通量的基因检测技术被广泛用于新基因的发现、基因诊断等研究过程中，利用基因芯片可以快速获得样本内所有基因同一时间点的表达信息，故其适用于差异表达基因的筛选分析[4]。基因芯片的应用过程中产生了海量的基因表达谱数据，通过生物信息学技术可以对差异表达的基因进行功能注释及预测，了解这些基因所参与的生物学过程以及发挥的具体功能。本研究分别对两组椎间盘退变基因表达谱数据中与纤维环退变及髓核退变有关的差异表达基因进行了筛选，并采用生物信息学工具对这些基因进行了比较分析和功能预测。

1 材料与方法

1.1 材料

本研究中所用两组椎间盘退变相关基因芯片表达谱数据均来源于美国国立生物技术信息中心 (The National Center for Biotechnology Information，NCBI) 基因表达数据库 (Gene Expression Omnibus，GEO)，基因系列号分别为 GSE23130 及 GSE67567，其中 GSE23130 所研究标本来源于正常及退变椎间盘纤维环组织，GSE67567 所研究标本来源于正常及退变椎间盘髓核组织。采用 GeneSpring 13.0 (Agilent Technologies，Santa Clara，CA，US)对上述系列样本数据进行质量分析，最终 GSE23130 和 GSE67567 各有 10 例样本数据被纳入实验。

1.2 方法

将上述两组系列样本数据分别导入 GeneSpring 13.0 软件，并命名为 GSE23130 及 GSE67567，随后对 GSE23130 正常及退变纤维环间差异表达基因及 GSE67567 正常及退变髓核间差异表达基因分别进行筛选，差异表达筛选条件设置为成组 Student's t 检验 $P<0.05$(双尾)，且基因表达水平相差超过两倍[5]。将差异表达上调及下调基因分别导入 KEGG PATHWAY 数据库进行生物学通路分析，经数据库默认的统计学算法计算 p 值，p 值小于 0.05 的生物学通路认定为具有统计学意义。同时，采用 DAVID(The Database for Annotation, Visualization and Integrated Discovery，DAVID)(version 6.7)数据库对筛选出的差异表达上、下调基因分别进行功能注释簇集分析，簇集分析时选择 DAVID 数据库默认的注释类别和 " 高 " 簇集分类标准选项，簇集评分大于 1.3 的簇集认定为具有显著性[6]。

2 结果

GSE23130 及 GSE67567 各筛选出差异表达基因 3182 个及 3017 个。其中，135 个基因在上述两个基因表达谱数据系列中均存在差异表达。GSE23130 中表达上调的基因有 1735 个，表达下调的有 1447 个，其中 116 个基因表达上调超过 10 倍，28 个表达下调超过 10 倍。GSE67567 中表达上调的基因有 2208 个，809 个表达下调，其中表达上调超过 10 倍的基因有 140 个，表达下调超过 10 倍的有 125 个。

针对上述两个基因表达谱数据系列进行的 KEGG PATHWAY 生物学通路分析结果显示如下(见表 1、表 2)。

针对上述两个基因表达谱数据系列进行的 DAVID 功能注释簇集分析结果显示如下(见表 3、表 4)。

3 讨论

既往的研究发现退变椎间盘内存在细胞凋亡增加、基质降解加速、基质合成改变等一系列病理改变[7]，随后针对退变椎间盘进行的基因芯片研究发现大量蛋白编码基因的异常表达[8,9]，基因的异常表达是造成上述病理改变的内在原因。纤维环和髓核作为椎间盘的主要组成部分，二者的胚胎来源不同[10]，然而，纤维环和髓核退变时的病理改变却表现出某些共同特征 [3]，这提示二者在基因表达上可能存在某些共性。为了比较椎间盘退变过程中纤维环和髓核基因表达谱的差异，本研究对椎间盘退变相关的两组基因表达谱数据进行了分析。结果显示不仅退变纤维环与正常纤维环间存在基因表达差异，退变髓核与正常髓核组织间也存在差异；退变时两种组织的差异表达基因谱不尽相同，135 个基因在两者间同时存在差异表达。这些结果提示差异表达基因谱的部分一致是导致纤维环和髓核病理变化存在部分共性，而又各有特点的主要原因。

为了进一步了解椎间盘退变时纤维环和髓核内生物学过程的差异，本研究采用 KEGG PATHWAY 和 DAVID 功能注释簇集对两组差异表达基因进行了分析。KEGG PATHWAY 分析结果显示两组中表达上调的基因均参与了 hsa03040:Spliceosome 等数个生物学通路，与此类似 DAVID 功能注释簇集分析结果显示两组中表达上调的基因均与 regulation of apoptosis 等部分生物学过程有关。两组中表达上调的基因均参与了 TGF-beta signaling pathway，过去的研究指出 TGF-beta 信号通路的激活在椎间盘细胞的生长、增殖、基质产生等方面发挥重要作用[11,12]。GSE23130 中表达上调的基因参与了 Wnt signaling pathway，椎间盘内 Wnt 通路的激活被认为与椎间盘细胞外基质的破坏有关[13]，而 TGF-beta 具有对抗 Wnt 信号通路激活效应的作用 [14]。由于椎间盘中 TGF-beta 信号通路的异常改变，TGF-beta 失去对 Wnt 的拮抗作用，导致退变椎间盘难以进行自身修复，加速了退变的发生。

正常情况下，仅外层纤维环有少量神经、血管分布，退变时神经、血管沿外层纤维环裂隙深入椎间盘内[15,16]。在本研究中，GSE23130 中表达上调的基因与 regulation of axonogenesis、regulation of neurogenesis 等生物学过程有关，这些基因还同时参与了 VEGF signaling pathway。此外，GSE67567 中表达上调的基因与 blood vessel morphogenesis 等功能注释簇集有关，上述基因可能诱导了退变时血管、神经向椎间盘内的生长。

蛋白聚糖及胶原纤维是椎间盘细胞外基质的主要成分，在椎间盘退变过程中有大量的细胞外基质被降解[17]。本研究结果

表 1 GSE23130 的 KEGG PATHWAY 分析结果

Table 1 Results of KEGG PATHWAY analysis for GSE23130

KEGG_PATHWAY Term	%*	P-value
Pathways of up-regulated genes		
hsa03040:Spliceosome#	1.976048	1.32E-07
hsa03010:Ribosome#	1.556886	2.57E-07
hsa04520:Adherens junction#	1.197605	7.56E-05
hsa00190:Oxidative phosphorylation	1.616766	1.80E-04
hsa04110:Cell cycle	1.317365	0.006841
hsa04350:TGF-beta signaling pathway#	0.958084	0.016435
hsa04510:Focal adhesion#	1.736527	0.02491
hsa04120:Ubiquitin mediated proteolysis#	1.257485	0.034467
hsa00900:Terpenoid backbone biosynthesis	0.299401	0.047704
hsa04310:Wnt signaling pathway	1.317365	0.048571
hsa04370:VEGF signaling pathway	0.778443	0.049961
hsa04662:B cell receptor signaling pathway	0.778443	0.049961
Pathways of down-regulated genes		
hsa00601:Glycosphingolipid biosynthesis	0.528634	0.006745
hsa04512:ECM-receptor interaction	0.881057	0.022289
hsa04650:Natural killer cell mediated cytotoxicity	1.145374	0.031218

* 差异表达基因数目占该生物学通路总基因数目的百分比

* Percentage of the number of differentially expressed genes over the total genes count of the pathway

在 GSE23130 和 GSE67567 的 KEGG PATHWAY 分析结果中均出现的簇集条目

Pathway which were presented in both KEGG PATHWAY results of GSE23130 and GSE67567

表 2 GSE67567 的 KEGG PATHWAY 分析结果

Table 2 Results of KEGG PATHWAY analysis for GSE67567

KEGG_PATHWAY Term	%*	P-Value
Pathways of up-regulated genes		
hsa03010:Ribosome#	1.085884	0.003145
hsa03040:Spliceosome#	1.332675	0.010293
hsa00511:Other glycan degradation	0.345508	0.011997
hsa00600:Sphingolipid metabolism	0.542942	0.024296
hsa04142:Lysosome	1.1846	0.026248
hsa04350:TGF-beta signaling pathway#	0.937808	0.028975
hsa04520:Adherens junction#	0.839092	0.036758
hsa04510:Focal adhesion#	1.7769	0.041189
hsa04120:Ubiquitin mediated proteolysis#	1.283317	0.047756
Pathways of down-regulated genes		
hsa04623:Cytosolic DNA-sensing pathway	0.948509	0.039959
hsa03010:Ribosome	1.219512	0.046644

* 差异表达基因数目占该生物学通路总基因数目的百分比

* Percentage of the number of differentially expressed gene over the total gene count of the pathway

在 GSE23130 和 GSE67567 的 KEGG PATHWAY 分析结果中均出现的簇集条目

Pathway which were presented in both KEGG PATHWAY results of GSE23130 and GSE67567

显示 GSE67567 中上调表达的基因不仅参与了 Other glycan degradation 生物学通路，还与 lysosome 生物学过程有关，而表达下调基因也参与了 negative regulation of macromolecule biosynthetic process、protein complex biogenesis 等生物学过程，

表 3 GSE23130 及 GSE67567 各自表达上调基因包含的簇集

Table 3 DAVID functional annotation clusters for up-regulated genes of GSE23103 and GSE67567 respectively

Clusters of GSE23130*	Clusters of GSE6756*
nuclear lumen#	nuclear lumen#
ribosomal protein#	nucleosome assembly
nuclear mRNA splicing, via spliceosome#	ribosomal protein#
RNA recognition motif, RNP-1#	Histone H2A
Cadherin	lysosome
protein transport#	protein transport#
mitochondrial respiratory chain complex I#	heparin binding
mitochondrial inner membrane	positive regulation of protein kinase activity
regulation of axonogenesis	intracellular protein transport#
regulation of apoptosis#	blood vessel morphogenesis
modification-dependent protein catabolic process	negative regulation of cell growth
Immunoglobulin C1-set	nuclear mRNA splicing, via spliceosome#
Ubiquitin	RNA elongation from RNA polymerase II promoter
intracellular protein transport#	regulation of apoptosis#
MHC class I protein complex	mitochondrial respiratory chain complex I#
regulation of neurogenesis	peroxidase activity
antigen processing and presentation of peptide antigen via MHC class I	mitochondrial outer membrane
Nuclear factor of activated T cells (NFAT)	Fibrillar collagen, C-terminal
Ubiquitin carboxyl-terminal hydrolase, N-terminal region 2	RNA recognition motif, RNP-1#
cytoplasmic membrane-bounded vesicle#	histone H2B
	cytoplasmic membrane-bounded vesicle#

* 簇集顺序按照 DAVID 功能注释簇集分析簇集评分排列

* Order of the clusters was sequenced by DAVID functional annotation clustering enrichment score

在 GSE23130 和 GSE67567 的 DAVID 功能注释簇集分析结果中均出现的簇集条目

Clusters which were presented in both DAVID functional annotation clustering results of GSE23130 and GSE67567

表 4 GSE23130 及 GSE67567 各自表达下调基因包含的簇集

Table 4 DAVID functional annotation clusters for down-regulated genes of GSE23103 and GSE67567 respectively

Clusters of GSE23130*	Clusters of GSE6756*
cell adhesion	focal adhesion
metal ion binding	negative regulation of macromolecule biosynthetic process
metallopeptidase activity	cytoplasmic membrane-bounded vesicle
extracellular matrix#	protein complex biogenesis
antigen binding	regulation of cell migration
	extracellular matrix#
	negative regulation of gene-specific transcription

* 簇集顺序按照 DAVID 功能注释簇集分析簇集评分排列

* Order of the clusters was sequenced by DAVID functional annotation clustering enrichment score

在 GSE23130 和 GSE67567 的 DAVID 功能注释簇集分析结果中均出现的簇集条目

Clusters which were presented in both DAVID functional annotation clustering results of GSE23130 and GSE67567

与此类似 GSE23130 上调表达基因也与 modification-dependent protein catabolic process 等簇集有关。此外，GSE23130 及 GSE67567 上调表达基因还与 Ubiquitin mediated proteolysis 等生物通路有关，这些基因的表达改变可能是椎间盘退变时细胞外基质产生减少、降解增加的重要原因，而 Ubiquitin mediated proteolysis 通路很可能充当了基质降解的主要途径[18]。

自体免疫损伤是导致椎间盘退变的一个重要因素[19]。本研究结果发现 GSE23130 中表达上调的基因与 DAVID 功能注释簇集 antigen processing and presentation of peptide antigen via MHC class I、Nuclear factor of activated T cells (NFAT)、MHC class I protein complex、Immunoglobulin C1-set 有关，还同 B cell receptor signaling pathway 生物学通路有关。此外，表达下调的基因也参与了 Natural killer cell mediated cytotoxicity、antigen binding，这些结果再次证实椎间盘的自身免疫损伤过程中，不仅有 T 细胞的参与，B 细胞及 NK 细胞均扮演了重要角色[20]。

总之，本研究结果显示椎间盘退变过程中纤维环及髓核组织内基因表达情况存在差异，生物学过程不尽相同，同时发现了数个在两种组织内均出现异常改变的生物学过程，这些生物学过程可能在椎间盘退变过程中的有着重要作用，针对这些生物学过程背后的异常表达基因进行研究将有助于我们加深对椎间盘退变机制的理解。

参 考 文 献(References)

[1] Hoy D, March L, Brooks P, et al. The global burden of low back pain: estimates from the Global Burden of Disease 2010 study [J]. Ann Rheum Dis, 2014, 73(6): 968-974

[2] Manchikanti L, Singh V, Falco FJ, et al. Epidemiology of low back pain in adults[J]. Neuromodulation, 2014, 17 (Suppl 2): 3-10

[3] Kepler CK, Ponnappan RK, Tannoury CA, et al. The molecular basis of intervertebral disc degeneration[J]. Spine J, 2013, 13(3): 318-330

[4] Yan B, Wang ZH, Guo JT. The research strategies for probing the function of long noncoding RNAs[J]. Genomics, 2012, 99(2): 76-80

[5] Ziats MN, Rennert OM. Aberrant expression of long noncoding RNAs in autistic brain[J]. J Mol Neurosci, 2013, 49(3): 589-593

[6] Gillett A, Maratou K, Fewings C, et al. Alternative splicing and transcriptome profiling of experimental autoimmune encephalomyelitis using genome-wide exon arrays [J]. PLoS One, 2009, 4(11): e7773

[7] Rahul J, Mike YC. Regenerative Biology of the Spine and Spinal Cord [M]. Landes Bioscience and Springer Science+Business Media, 2012, CHAPTER 8: 114-133

[8] Chen K, Wu D, Zhu X, et al. Gene expression profile analysis of human intervertebral disc degeneration [J]. Genet Mol Biol, 2013, 36 (3): 448-454

[9] Tang Y, Wang S, Liu Y, et al. Microarray analysis of genes and gene functions in disc degeneration[J]. Exp Ther Med, 2014, 7(2): 343-348

[10] 肖应权, 孙凤, 牛祥科, 等. 椎间盘退变与再生:胚胎发育中的经验[J]. 中华临床医师杂志(电子版), 2013, 7(1): 301-303
Xiao Ying-quan, Sun Feng, Niu Xiang-ke, et al. Degeneration and regeneration of intervertebral disc: experience in embryonic development [J]. Chin J Clinicians (Electronic Edition), 2013, 7(1): 301-303

[11] Jin H, Shen J, Wang B, et al. TGF-β signaling plays an essential role in the growth and maintenance of intervertebral disc tissue [J]. FEBS Lett, 2011, 585(8): 1209-1215

[12] Yang H, Cao C, Wu C, et al. TGF-βl Suppresses Inflammation in Cell Therapy for Intervertebral Disc Degeneration [J]. Sci Rep, 2015, 5: 13254

[13] Hiyama A, Sakai D, Risbud MV, et al. Enhancement of intervertebral disc cell senescence by WNT/ β-catenin signaling-induced matrix metalloproteinase expression [J]. Arthritis Rheum, 2010, 62 (10): 3036-3047

[14] Hiyama A, Sakai D, Tanaka M, et al. The relationship between the Wnt/β-catenin and TGF-β/BMP signals in the intervertebral disc cell [J]. J Cell Physiol, 2011, 226(5): 1139-1148

[15] Freemont AJ, Peacock TE, Goupille P, et al. Nerve ingrowth into diseased intervertebral disc in chronic back pain[J]. Lancet, 1997, 350 (9072): 178-181

[16] Binch AL, Cole AA, Breakwell LM, et al. Class 3 semaphorins expression and association with innervations and angiogenesis within the degenerate human intervertebral disc[J]. Oncotarget, 2015, 6(21): 18338-18354

[17] Wang WJ, Yu XH, Wang C, et al. MMPs and ADAMTSs in intervertebral disc degeneration [J]. Clin Chim Acta, 2015, 448: 238-246

[18] Mayer JE, Iatridis JC, Chan D, et al. Genetic polymorphisms associated with intervertebral disc degeneration [J]. Spine J, 2013, 13 (3): 299-317

[19] Stich S, Stolk M, Girod PP, et al. Regenerative and immunogenic characteristics of cultured nucleus pulposus cells from human cervical intervertebral discs[J]. PLoS One, 2015, 10(5): e0126954

[20] Sun Z, Zhang M, Zhao XH, et al. Immune cascades in human intervertebral disc: the pros and cons [J]. Int J Clin Exp Pathol, 2013, 6(6): 1009-1014