

骨关节炎患者膝关节滑膜组织的基因表达谱分析

谌虹焰^{1,2}, 张明媚¹, 陈杏¹, 王涛², 谢庆云³, 魏萌^{1,2}

(1. 西南交通大学医学院; 2. 西部战区总医院风湿免疫科; 3. 西部战区总医院骨科, 四川 成都 610083)

【摘要】目的:分析骨关节炎患者膝关节滑膜组织的基因表达谱。**方法:**从公共基因芯片数据库选取芯片数据,利用 GEO2R 筛选骨关节炎患者滑膜组织和正常滑膜组织的差异表达基因;对筛选出的基因进行 GO 功能和 KEGG 通路富集分析,利用 STRING 数据库及 Cytoscape 软件构建蛋白间相互作用网络,筛选关键基因。**结果:**纳入数据库中两组基因芯片数据,共筛选出 131 个共同差异表达基因,其中上调基因 98 个,下调基因 33 个。GO 富集分析提示,差异基因主要在细胞粘附、对成纤维细胞生长因子刺激的反应、细胞内受体信号通路等功能富集。KEGG 富集结果显示,差异表达基因参与了丝裂原活化蛋白激酶/低氧诱导因子-1、磷脂酰肌醇 3 激酶/蛋白激酶 B 等信号通路。蛋白间相互作用网络提示了 10 个关键基因。**结论:**通过对 OA 基因表达谱分析筛选差异表达基因发现 OA 的发病机制不仅与细胞增殖、病理性血管生成、炎症刺激等有关,更是多信号通路、多靶点间相互作用的结果,可为进一步研究提供参考。

【关键词】骨关节炎;差异表达基因;生物信息学;滑膜

【中图分类号】R684.3 **【文献标志码】**A

Gene expression profile of knee joint synovial tissue in patients with osteoarthritis

SHEN Hong-yan^{1,2}, ZHANG Ming-mei¹, CHEN Xing¹, WANG Tao², XIE Qing-yun³, WEI Meng^{1,2}

(1. School of Medicine, Southwest Jiaotong University; 2. Department of Rheumatology and Immunology; 3. Department of Orthopedics, the General Hospital of Western Theater Command, Chengdu 610083, Sichuan, China)

【Abstract】 Objective: To analyze the gene expression profile of knee joint synovial tissue in patients with osteoarthritis. **Methods:** The microarray data were selected from the common gene chip database, and the differentially expressed genes in synovial tissue of osteoarthritis patients and normal synovial tissue were screened by GEO2R, the Gene ontology (GO) function and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of the screened genes were carried out. The protein-protein interaction (PPI) network was constructed by STRING database and Cytoscape software, which presented the hub genes. **Results:** A total of two sets of osteoarthritis gene chip data from databases were included. 131 common DEGs were screened out, including 98 up-regulated genes and 33 down-regulated genes. GO enrichment analysis revealed that DEGs were mainly enriched in cell adhesion, response to fibroblast growth factor stimulation, and intracellular receptor signaling pathways. KEGG pathway demonstrated the inflammatory signaling pathways involved in DEGs, such as mitogen-activated protein kinase (MAPK) signaling pathway, hypoxia-inducible factor-1 (HIF-1) signaling pathway, phosphatidylinositol 3 kinase/protein kinase B (PI3K/Akt) signaling pathway. The PPI network showed 10 hub genes. **Conclusion:** Screening of differentially expressed genes by analyzing the expression profile of OA gene suggests that the pathogenesis of OA is not only related to cell proliferation, pathological angiogenesis and inflammatory stimulation, but also the result of the complex interactions between multiple signaling pathways and multiple targets, which can provide reference for further research.

【Key words】 Osteoarthritis; Differentially expressed genes; Bioinformatics; Synovium

骨关节炎(osteoarthritis, OA)是最常见的退行性关节疾病,以关节软骨损害为主要特点,可逐渐累及软骨下骨、滑膜、周围软组织等直至整个关节,给患者造成关节疼痛、关节畸形及活动受限等痛苦。我国 OA 的总患病率在 60 岁以上年龄段达 50%, 75

岁以上高达 80%^[1], 因其好发于下肢负重关节,如髋、膝关节,也成为中老年人下肢残疾的主要原因。年龄、性别、肥胖、遗传易感性、关节过度活动与压力等都是 OA 的病因^[2], 而 OA 的发病机制尚不完全清楚。从近年来不断发展的表达谱芯片入手,以基因组学来研究发病机制,对寻找潜在诊断标志物和

治疗靶点有重要意义。本研究旨在通过挖掘基因表达综合数据库 (gene expression omnibus, GEO) 中膝 OA 滑膜组织与正常滑膜相比的基因芯片数据,对差异表达基因进行生物信息学分析,为找到新的治疗靶点和揭示 OA 的发病机制提供更多的线索和依据。

1 资料与方法

1.1 一般资料

从美国国家生物信息技术中心 (national center for biotechnology information, NCBI) 的 GEO 数据库中寻找 OA 滑膜组织相关的基因表达谱数据。以“Osteoarthritis”、“Synovial tissue”为关键词,在人类组织标本的分类下,搜索近 5 年的数据,选取在同一研究平台使用相同芯片进行分析的数据集,尽量减少数据集之间比较的差异,保证数据分析的可靠性,数据集 GSE55457 和 GSE55235 符合要求。这两组数据集均来自 GPL96 平台,所用芯片类型为 Affymetrix Human Genome U133A Array (HG-U133A),每组数据中均包含健康对照样本、OA 患者样本和 RA 患者样本,本研究选择其中的 OA 患者样本和健康对照样本作为研究对象,共纳入 40 例样本 (20 例 OA 样本、20 例健康对照样本) 进行综合分析。

1.2 方法

1.2.1 差异表达基因的筛选 利用 GEO2R 在线软件 (<http://www.ncbi.nlm.nih.gov/geo/geo2r/>) 分别对两组数据集进行分析,从 GEO 数据库中获取特定基因的表达值,利用 R 语言中的 limma 工具包确定 OA 和健康对照滑膜组织之间的差异表达基因,通过 Benjamini-Hochberg 方法纠正错误发现率 (FDR),以 $FDR-P < 0.05$, $|\log_2 FC| \geq 1$, (差异表达倍数为两倍) 为条件筛选差异表达基因,并对筛选的两组差异表达基因求交集,获得共同表达上调和下调的基因,构建维恩图 (venn diagram),绘制基因热图。

1.2.2 差异表达基因的功能富集及注释 将上调和下调的基因分别上传至 DAVID 6.8 在线分析软件 (<https://david.ncifcrf.gov/summary.jsp>),对差异表达基因进行 GO (gene ontology) 富集分析,包括生物过程、细胞组成和分子功能三个方面,以及 KEGG (kyoto encyclopedia of genes and genomes) 通路富集分析,以 $P < 0.05$ 为筛选条件,得到富集分析结果,从功能层面对差异表达基因进行解读。

1.2.3 蛋白间相互作用网络分析 将差异表达基因导入 STRING 数据库 (<https://string-db.org/>) 获得基因所编码的蛋白间相互作用 (protein-protein in-

teraction, PPI) 网络图。再利用 Cytoscape 3.8.0 软件进行可视化,通过 cytoHubba 插件计算各个节点间相互作用程度“degree”,根据 degree 值筛选出关键基因。

1.3 统计学分析

采用 SPSS 20.00 软件对数据进行分析与处理。采用 t 检验分析 OA 和正常滑膜组织之间的基因表达。 $P < 0.05$ 为差异具有统计学意义。

2 结果

2.1 差异表达基因

通过 GEO2R 筛选出数据集 GSE55457 和 GSE55235 满足调整后 $P < 0.05$, $|\log_2 FC| \geq 1$ 的差异表达基因,其中 GSE55457 组 788 个 (上调基因 673 个、下调基因 115 个)、GSE55235 组 1048 个 (上调基因 405 个、下调基因 643 个)。取交集并构建维恩图,可见两组共有 131 个差异表达基因,其中有 98 个上调基因,33 个下调基因。见图 1、图 2。

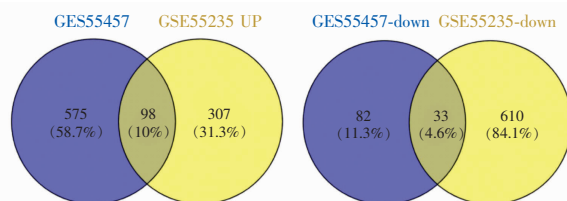


图 1 共同上调及下调基因

蓝色代表数据集 GSE55457,黄色代表数据集 GSE55235,交叉部分即两组数据集中共同上调或下调的基因个数。

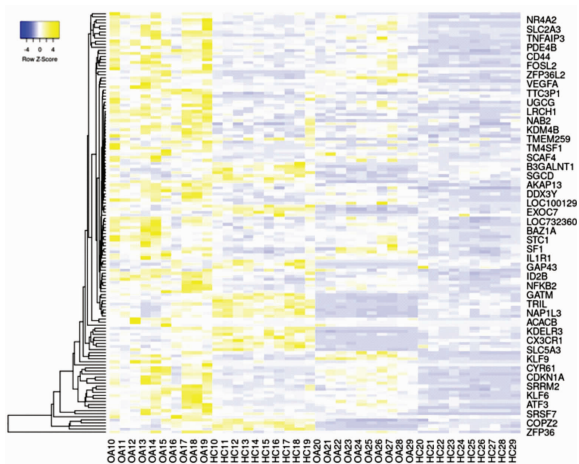


图 2 共同差异表达基因水平

蓝色代表下调,黄色代表上调。

2.2 差异表达基因的 GO 功能和 KEGG 通路富集分析

GO 富集分析从生物过程、细胞组成和分子功能三个方面诠释了差异表达基因的功能。上调的差异表达基因在生物过程方面富集丰富,如对成纤维

细胞生长因子刺激的反应、细胞内受体信号通路等,下调的差异表达基因主要在细胞粘附、运输囊泡等功能富集。气泡图显示,差异表达基因在信号通路富集明显,例如丝裂原活化蛋白激酶(mitogen-activated protein kinase,MAPK)信号通路、磷脂酰肌醇3激酶(phosphatidylinositol 3 kinase,PI-3K)/蛋白激酶B(protein kinase B,PKB)-PI3K-Akt信号通路、低氧诱导因子-1(hypoxia inducible factor-1,HIF-1)信号通路等。见表1、表2、图3。

表1 上调差异表达基因的GO富集分析

类别	富集基因数	P值
生物过程		
细胞对成纤维细胞生长因子刺激的反应	5	<0.001
RNA聚合酶II启动子转录的正调控	17	<0.001
昼夜节律	5	<0.001
细胞内受体信号通路	4	<0.001
凋亡过程的负调控	9	0.002
细胞迁移参与发芽血管生成	3	0.002
成纤维细胞增殖的正调控	4	0.003
细胞组成		
核质	32	<0.001
核	47	<0.001
细胞质	38	0.004
分子功能		
核苷酸结合	9	<0.001
Poly(A)RNA结合	16	<0.001
蛋白质结合	61	<0.001

表2 下调差异表达基因的GO富集分析

类别	富集基因数	P值
生物过程		
细胞粘附	5	0.007
神经细胞发育	4	0.013
运输囊泡	3	0.009
细胞外空间	7	0.017
细胞组成		
肌营养不良蛋白相关糖蛋白复合物	2	0.025
细胞连接	4	0.036
内质网腔	3	0.037
分子功能		
肝素结合	3	0.023

2.3 PPI网络分析

PPI网络图显示,有83个节点蛋白,248条相互作用关系。通过degree值筛选出排名前10的核心节点蛋白发现,编码这些蛋白的基因均表达上调,分别为JUN、VEGFA、ATF3、MYC、DUSP1、NR4A1、CDKN1A、BTG2、SOCS3、GADD45B。见图4、图5。

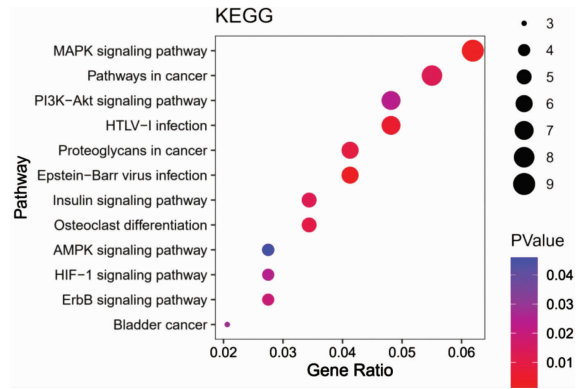


图3 差异表达基因的KEGG通路富集

X轴显示基因比例,Y轴显示所富集的KEGG通路;气泡大小表示在KEGG通路中富集的基因数量,气泡颜色代表调整后的P值。

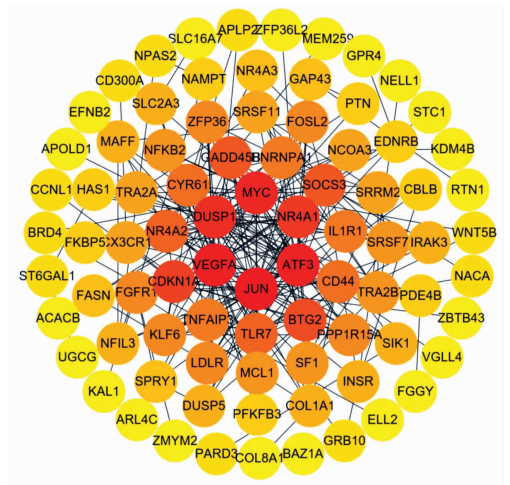


图4 差异表达基因的蛋白间相互作用网络

节点代表蛋白,被一条线连接的两个节点表示这两个蛋白的作用是相互的,节点连线的多少即是该节点的degree值,颜色表示节点degree值的排名,颜色越深,排名越靠前,与该节点相互作用的蛋白越多。

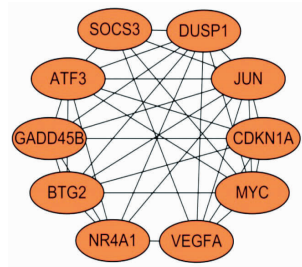


图5 基因编码的核心蛋白相互作用关系

3 讨论

本研究综合两组高通量的基因芯片数据,筛选出OA滑膜组织中131个共同差异表达基因,通过分析其功能探讨OA发病的分子机制。从GEO数据库中可以看到,随着技术的完善与发展,基因表达谱芯片的类型繁多,各种芯片能检测的基因数及检测效力各不相同。即便是使用同种芯片多次对同类

型样本进行检测,由于检测人员的操作以及分析过程带来的不可避免的批次差异,原始数据处理方式、质量控制标准和分析方法的不同,都会使检测结果不完全一样。因此,选择相同检测平台的数据集进行综合分析,所选出的共同差异表达基因在一定程度上排除了未知混杂因素的干扰,通过这些差异表达基因来研究疾病的发病机制可靠性更高。

从 GO 富集结果来看,上调的差异表达基因在细胞核和细胞质的组成成分方面富集,主要参与多种核苷酸结合以及蛋白质结合等分子功能,调节多种生物过程,其中昼夜节律、细胞内受体信号通路、细胞迁移参与发芽血管生成、成纤维细胞增殖的正调控、对机械刺激的反应、炎症反应的调节等都与 OA 的发病相关^[3]。而下调的差异表达基因主要位于细胞器、细胞外基质的组成成分,调节细胞粘附、神经功能发育以及肝素结合等过程。

KEGG 富集结果则提示了上调的差异表达基因通过 MAPK、PI3K-Akt、HIF-1 等通路参与 OA 的炎症过程及滑膜纤维化^[4]。滑膜组织中的成纤维样滑膜细胞和巨噬细胞均在炎症的发生发展中发挥作用,促进炎症因子的产生与释放,导致关节的疼痛和硬化。MAPK 与 PI3K-Akt 信号通路是经典的炎症信号通路,已被发现涉及多种疾病,它们对于调节 OA 的病理生理至关重要,其中的 ERK、JNK 信号传导途径还参与 OA 软骨细胞的增殖与凋亡以及病理性降解。在动物实验中发现,抑制 PI3K-Akt 信号传导可以减弱炎症反应,还能够阻止异常骨形成;而低氧诱导因子-1 其亚基 HIF-1 α 作为转录因子,可以调节促炎症细胞因子, HIF-1 α 的表达也与 OA 滑膜的增殖程度正相关。

蛋白相互作用网络则进一步解释了差异表达基因所编码蛋白的功能,基因编码的核心蛋白也参与 OA 的发生发展。OA 患者的滑膜组织中血管内皮生长因子 (vascular endothelial growth factor, VEGF) 水平升高,与 OA 的严重程度和疼痛相关^[5]。VEGFA 作为其家族成员,在诱导生理性和病理性血管生成方面至关重要。滑膜中血管生成的增加能刺激炎症反应,分泌炎症细胞因子,导致滑膜增生。靶向 VEGF 抑制其信号可以减缓 OA 进展,减少疼痛^[6]。活化转录因子 3 (activating transcription factor 3, ATF3) 参与细胞增殖和分化的调控。研究^[7]提示,软骨细胞中表达的 ATF3 可能通过修饰 NF- κ B 通路,调控炎症细胞因子诱导自身的表达及分泌,增强炎症细胞因子的产生,参与 OA 的发展,该基因可能成为治疗 OA 的新靶点。MYC 基因是癌基因,包括 C-myc、N-myc 和 L-myc。目前对 MYC 基因在 OA 中的作用研究较少,有一项大鼠实验^[8]表示,用 mi-

croRNA-24 靶向 C-myc 基因,下调 C-myc 可以抑制细胞凋亡并促进软骨细胞的增殖,从而通过失活 MAPK 信号通路来防止 OA 的发生发展。MYC 基因在滑膜组织中的作用还有待研究。

双特异性磷酸酶 1 (dual specificity phosphatase 1, DUSP1) 被发现在 OA 的成纤维样滑膜细胞中具有保护作用,其过表达抑制了 MAPK 途径的激活以及 OA 相关介质的表达^[9],这可能成为治疗 OA 的潜在靶点。而细胞信号传导因子 3 (suppressor of cytokine signaling 3, SOCS3) 表达水平的升高在 OA 的发展中产生矛盾的影响^[10]。SOCS3 一方面可能通过抗炎作用以及抑制过度增生的软骨组织修复来减少炎症及骨赘的形成;另一方面有可能会延迟 OA 中软骨下松质骨损伤的修复。

综上所述,通过分析 GSE55457 和 GSE55235 这两组数据的共同差异表达基因,发现 OA 的发病机制不仅仅与细胞增殖、病理性血管生成、炎症刺激等方面有关,更是多信号通路、多靶点间相互作用的结果。这些差异表达基因 (如 VEGFA、DUSP1) 可为深入解读 OA 的致病机理提供参考。

参考文献

- [1] Woolf AD, Pfleger B. Burden of major musculoskeletal conditions [J]. Bull World Health Organ, 2003, 81(9): 646-656.
- [2] Pereira D, Ramos E, Branco J. Osteoarthritis [J]. Acta Med Port, 2015, 28(1): 99-106.
- [3] Bekki H, Duffy T, Okubo N, et al. Suppression of Circadian Clock Protein Cryptochrome 2 Promotes Osteoarthritis [J]. Osteoarthritis Cartilage, 2020, 28(7): 966-976.
- [4] Chow YY, Chin KY. The Role of Inflammation in the Pathogenesis of Osteoarthritis [J]. Mediators Inflamm, 2020, Doi: 10.1155/2020/8293921.
- [5] Takano S, Uchida K, Inoue G, et al. Vascular endothelial growth factor expression and their action in the synovial membranes of patients with painful knee osteoarthritis [J]. BMC Musculoskelet Disord, 2018, 19(1): 204.
- [6] Hamilton JL, Nagao M, Levine BR, et al. Targeting VEGF and Its Receptors for the Treatment of Osteoarthritis and Associated Pain [J]. J Bone Miner Res, 2016, 31(5): 911-924.
- [7] Iezaki T, Ozaki K, Fukasawa K, et al. ATF3 deficiency in chondrocytes alleviates osteoarthritis development [J]. J Pathol, 2016, 239(4): 426-437.
- [8] Wu YH, Liu W, Zhang L, et al. Effects of microRNA-24 targeting C-myc on apoptosis, proliferation, and cytokine expressions in chondrocytes of rats with osteoarthritis via MAPK signaling pathway [J]. J Cell Biochem, 2018, 119(10): 7944-7958.
- [9] Peng HZ, Yun Z, Wang W, et al. Dual specificity phosphatase 1 has a protective role in osteoarthritis fibroblastlike synoviocytes via inhibition of the MAPK signaling pathway [J]. Mol Med Rep, 2017, 16(6): 8441-8447.
- [10] Gui T, He BS, Gan Q, et al. Enhanced SOCS3 in osteoarthritis may limit both proliferation and inflammation [J]. Biotech Histochem, 2017, 92(2): 107-114.

(收稿日期: 2020-05-10

修回日期: 2020-07-07)