

· 论 著 ·

基于生物信息学方法筛选骨关节炎进展相关关键基因

孙伯辰, 刘 琳
(福建医科大学, 福建 福州 350004)

[摘 要] **目的** 利用生物信息学方法筛选骨关节炎进展相关的差异表达基因 (DEGs)。**方法** 在基因表达组学数据库(GEO)下载 GSE206848 数据集,运用 GEO₂R 在线工具进行分析,并筛选 DEGs;利用 R 语言绘制火山图和热图。然后对 DEGs 进行 GO 功能富集分析和 KEGG 通路富集分析;将获得的 DEGs 数据导入 STRING 数据库构建蛋白互作网络(PPI),并通过 CytoHubba 插件以 5 种算法取交集筛选关键基因(Hub gene)。**结果** 经 GEO₂R 筛选,从 GSE206848 数据集中获得 DEGs 1 592 个(545 个上调基因、1 047 个下调基因)。GO 功能富集分析发现,DEGs 的生物学过程主要富集在 DNA 代谢过程的调节、超分子纤维组织的调控、细胞器组织的负调控;细胞定位主要富集细胞边缘、肌动蛋白细胞骨架、核散斑;分子功能主要富集微管蛋白结合、蛋白丝氨酸/苏氨酸激酶活性、蛋白丝氨酸/苏氨酸/酪氨酸激酶活性。KEGG 通路富集分析结果主要涉及 PI3K-Akt 信号通路、肌动蛋白细胞骨架调控等。应用 CytoHubba 插件筛选得到 3 个关键基因(PIK3CA、EGFR、SRC)。**结论** 该研究共筛选出 3 个与骨关节炎进展相关的关键基因,分别为 EGFR、PIK3CA、SRC。

[关键词] 骨关节炎; 生物信息学方法; 差异表达基因; 关键基因; 数据库分析

DOI:10.3969/j.issn.1009-5519.2023.11.006 **中图法分类号:**R684.3

文章编号:1009-5519(2023)11-1829-05 **文献标识码:**A

Screening hub genes related to the progression of osteoarthritis based on bioinformatics method
SUN Bochen, LIU Lin
(Fujian Medical University, Fuzhou, Fujian 350004, China)

[Abstract] **Objective** To screen differentially expressed genes (DEGs) related to the progression of osteoarthritis by bioinformatics method. **Methods** The GSE206848 dataset was downloaded from the gene expression omics database (GEO), analyzed by GEO2R online tool, and screened for DEGs. R software was used to draw volcano and heat maps, then GO functional enrichment analysis and KEGG pathway enrichment analysis were performed for DEGs. The DEGs data were imported into STRING database to construct protein-protein interaction (PPI) network, the hub genes were screened through 5 algorithms for intersection selection of the CytoHubba plugin. **Results** After GEO2R screening, 1 592 DEGs (545 upregulated genes and 1 047 down-regulated genes) were obtained from the GSE206848 dataset. GO funtional enrichment analysis showed that the biological process of DEGs was mainly enriched in the regulation of DNA metabolism, the regulation of supramolecular fibrous tissue, and the negative regulation of organelle tissue. Cell localization mainly enriched cell margin, actin cytoskeleton and nuclear speckle. Molecular functions mainly enrich tubulin binding, protein serine/threonine kinase activity, and protein serine/threonine/tyrosine kinase activity. The enrichment analysis of KEGG pathway mainly involved PI3K Akt signaling pathway, actin cytoskeleton regulation, etc. After using the CytoHubba plugin to screen three hub genes (PIK3CA, EGFR, and SRC). **Conclusion** In this study, three hub genes related to the progression of osteoarthritis were screened, namely EGFR, PIK3CA and SRC.

[Key words] Osteoarthritis; Bioinformatics method; Differentially expressed genes; Hub genes; Database analysis

骨关节炎(OA)是以关节疼痛为主要症状的退行性疾病,主要表现为关节软骨纤维化、皲裂、溃疡、脱失^[1],是全球疼痛、残疾和社会经济成本升高的主要因素,该病的流行病学是复杂和多因素的,与遗传、生物学和生物力学因素有关^[2]。OA 病因尚不明确,随着疾病的进展最终可导致患者关节功能丧失或残疾,给个人和社会带来巨大的经济负担。随着我国老龄化进程加速,OA 的患病率逐年上升,约 10%的男性和 18%的女性在 60 岁以后患有 OA 相关症状^[3]。患者早期关节结构并无明显改变,症状亦不明显,往往

作者简介:孙伯辰(1988—),硕士研究生,主治医师,主要从事关节外科创伤骨伤、运动医学等方面诊治工作。

到疾病发展后期才进行影像学检查。目前,临床还未出现可以早期诊断 OA 及判断预后的相关生物标志物。近年来,分子生物学不断发展,已有证据表明,较多细胞因子及蛋白参与了 OA 的病理变化,其与滑膜、关节软骨、软骨下骨的功能改变有关^[4]。

基因芯片,又称 DNA 微阵列(DNA microarray),属于高通量基因分析技术,运用核酸分子碱基互补配对原理和原位杂交技术,密集、固定核苷酸,在生物芯片技术中较为成熟,目前认为此技术可运用于鉴定潜在治疗靶点,但大部分用于肿瘤的研究,随着一些新的数据分析方法不断地被开发,基因芯片也可以应用于其他疾病的诊断与治疗。目前,有研究运用表达谱芯片发现了与神经损害和疼痛相关的基因和潜在的治疗靶点,阐述了加巴喷丁运用于疼痛治疗的分子机制^[5]。利用基因芯片技术,可以检测生物样本基因的表达水平,通过分析后可以获得基因的功能及相互作用关系。检测 OA 与正常骨关节之间的差异表达基因,将有助于分析 OA 发生和发展的关键基因和通路,指导临床针对性的早期治疗,从而提高患者的治疗效果,让更多患者尽早重返工作和生活。

1 材料与方法

1.1 数据来源 登录 GEO 数据库(<https://www.ncbi.nlm.nih.gov/geo/>),以“Osteoarthritis”为检索词,筛选获得 GSE206848 数据集。GSE206848 数据集是基于 GPL570 平台,采用 Affymetrix Human Genome U133 Plus 2.0 Array 阵列芯片检测。GSE206848 数据集含有 14 例组织样本,其中 OA 组织样本 7 例,正常组织样本 7 例。技术路线流程图见图 1。

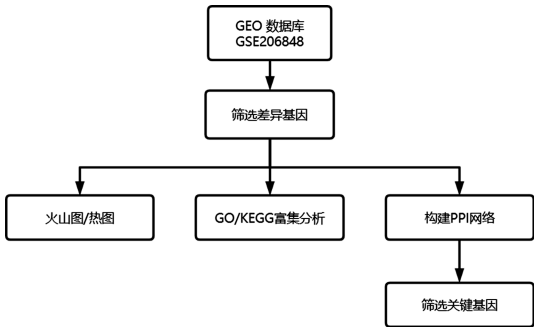


图 1 工作流程图

1.2 方法

1.2.1 DEGs 筛选 通过在线分析工具 GEO₂R 获取 GSE206848 数据集中的差异基因,以 $\text{adj-}P < 0.05$ 且 $|\log_2\text{FC}| > 1$ (FC 为差异倍数)为筛选条件进行筛选,利用 R 语言对 GSE206848 数据集中的 DEGs 进行火山图、热图绘制。

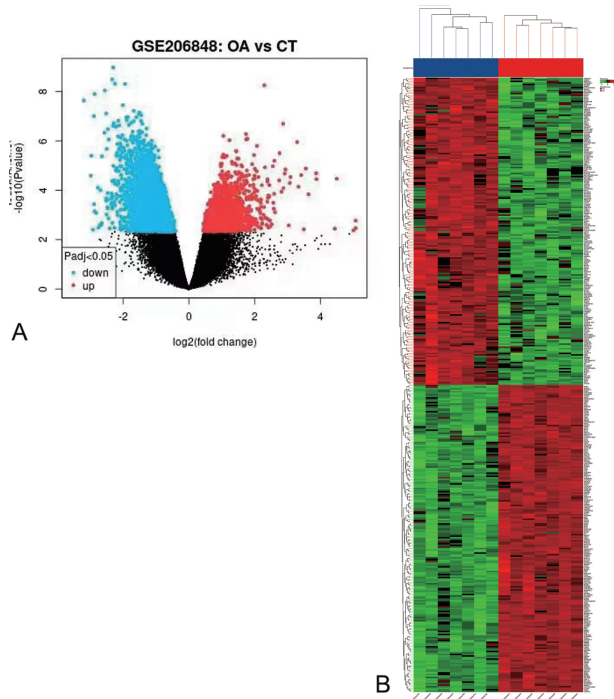
1.2.2 GO 功能富集分析和 KEGG 通路富集分析 在 R 软件中安装“org. Hs. eg. db”和“cluster profiler”R 包,进行 GO 功能富集分析和 KEGG 通路富集分析^[6-7]。GO 功能富集分析包含生物学过程(BP)、细

胞组分(CC)和分子功能(MF)。然后使用气泡图表来显示 GO 和 KEGG 分析中的前 10 个富集结果, $P < 0.05$ 为差异有统计学意义。

1.2.3 蛋白互作网络(PPI)构建与关键基因筛选 通过 STRING 数据库(<https://string-db.org>)构建 PPI 网络,随后利用 Cytoscape 软件里的 MCODE 插件筛选 PPI 网络中的关键模块,选择的标准:degree cut-off = 2, MCODE scores > 4, Max depth = 100, node score cut-off = 0.3 和 k-score = 2。最后,运用 CytoHubba 插件,以 5 种方法(MCC、MNC、Degree、EPC 和 BottleNeck)获取排名前 10 位的关键基因。

2 结果

2.1 OA DEGs 筛选 从 GSE206848 数据集中共获取表达上调基因 545 个、表达下调基因 1 047 个。得到的差异基因,利用 R 语言进行火山图可视化展示(图 2A),并筛选前 200 差异基因进行热图展示(图 2B)。



注:A. 差异基因火山图;B. 排名前 50 的差异基因表达热图。

图 2 OA DEGs 筛选

2.2 GO 功能富集分析和 KEGG 通路富集分析 GO 功能富集分析发现,在生物学过程中,DEGs 主要富集在 DNA 代谢过程的调节、超分子纤维组织的调控、细胞器组织的负调控、温度稳态、适应性发热作用等;细胞定位主要涉及细胞边缘、肌动蛋白细胞骨架、核散斑、细胞间连接、细胞与底物连接;分子功能主要涉及微管蛋白结合、蛋白丝氨酸/苏氨酸激酶活性、蛋白丝氨酸/苏氨酸/酪氨酸激酶活性(图 3A)。KEGG 通路富集主要涉及 PI3K-Akt 信号通路、癌症中的蛋白多糖、肌动蛋白细胞骨架调控、肌动蛋白结合、金属离子跨膜转运体活性等(图 3B)。

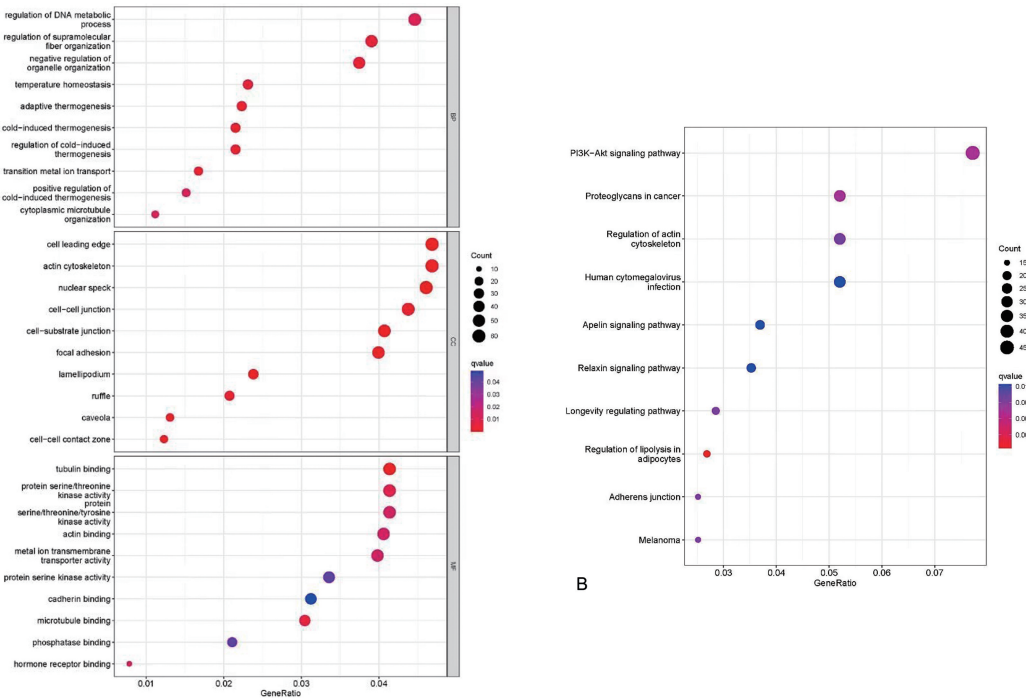


图 3 GO 功能富集分析和 KEGG 通路富集分析

2.3 PPI 构建与关键基因筛选结果 通过 STRING 数据库构建 PPI,见图 4。随后采用 Cytoscape 软件里的 MCODE 插件确认 PPI 中的重要模块,得到 2 个关键模块(图 5A、B)。随后利用 CytoHubba 插件运用 5 种算法进行筛选。筛选 PPI 中相互作用排名前 10 位的 DEGs 作为关键基因,见表 1。随后将 5 种算法所得的前 10 基因取交集,绘制韦恩图得到 3 个真正的关键基因(图 6),分别为 PIK3CA、表皮生长因子受体(EGFR)、SRC。

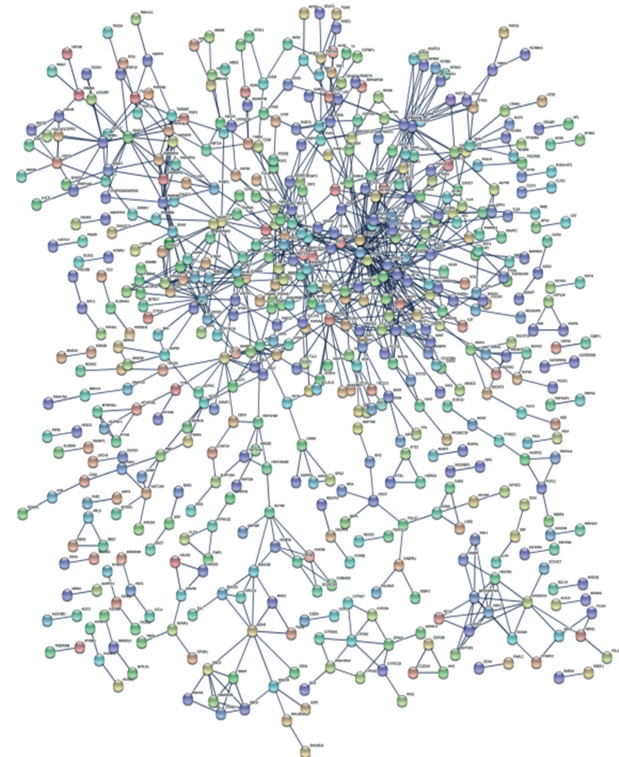


图 4 蛋白质-蛋白质 PPI

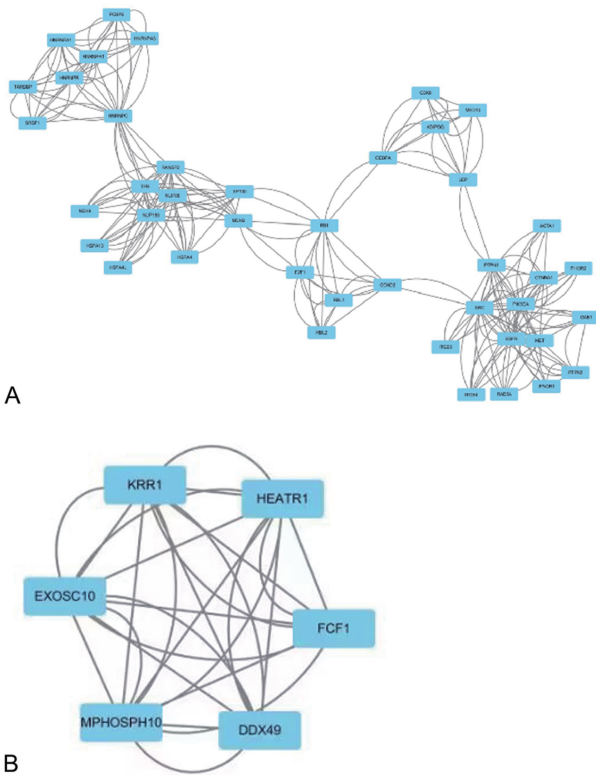


图 5 MCODE 得到的 2 个关键模块

表 1 CytoHubba 5 种算法所得前 10 位的关键基因

MCC	MNC	Degree	EPC	BottleNeck
SRC	SRC	SRC	SRC	SRC
EGFR	PIK3CA	PIK3CA	PIK3CA	RB1
PTPN11	PTPN11	PTPN11	EGFR	MDM2
PIK3CA	EGFR	EGFR	PTPN11	CCND2
MET	PRKACB	PRKACB	GAB1	ITGB3
GAB1	PRKACG	PRKACG	MET	HIST1H1E
CTNNA1	JAK1	JAK1	JAK1	EGFR

续表 1 CytoHubba 5 种算法所得前 10 位的关键基因				
MCC	MNC	Degree	EPC	BottleNeck
HNRNPC	GAB1	GAB1	PTPN2	HSPA4
PTPN2	MET	HSPA4	PTPN1	PIK3CA
HNRNPA1	PTPN2	MET	CD247	MAPK8

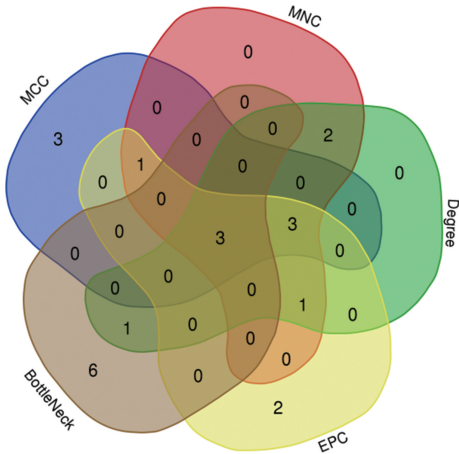


图 6 5 种算法的韦恩图

3 讨 论

OA 是骨科中最常见的慢性关节疾病,其发病率随着年龄的增长而增加,并影响大多数 65 岁以上的个体。关节软骨由浅表(切线)层、中间(移行)层、深(放射)层和钙化软骨层构成,关节软骨在预防 OA 发生过程中至关重要,在该病发生后,软骨细胞数量减少,排列紊乱、成簇,关节镜下可见软骨表面不光滑、皲裂,甚至出现软骨剥脱。目前,OA 主要的临床症状包括慢性疼痛、关节不稳、僵硬、关节畸形和影像学上关节间隙狭窄,其治疗方法包括减轻疼痛、缓解僵硬、维持功能,从而提高患者的生活质量^[17]。其治疗方法还包括低冲击有氧运动、减肥、针灸、服用氨基葡萄糖及手术。虽然各界已经提出了一系列与 OA 相关的危险因素,包括遗传易感性、衰老、肥胖和关节排列不良,但 OA 的发病机制目前仍不清楚。

本研究基于 GEO 数据库,利用生物信息学生物工具筛选出上调基因 545 个、下调基因 1 047 个。GO 分析发现,在生物学过程中,DEGs 主要富集在 DNA 代谢过程的调节、超分子纤维组织的调控、细胞器组织的负调控等;细胞定位主要富集在细胞边缘、肌动蛋白细胞骨架、核散斑等;分子功能主要富集在微管蛋白结合、蛋白丝氨酸/苏氨酸激酶活性、蛋白丝氨酸/苏氨酸/酪氨酸激酶活性等。KEGG 通路富集主要涉及 PI3K-Akt 信号通路、肌动蛋白细胞骨架调控等。通过 Cytoscape 软件,共筛选出 2 个关键模块。然后通过 CytoHubba 插件,以 5 种算法取交集后共筛选出 OA 发生相关的 3 个关键基因,分别为 EGFR、PIK3CA、SRC。

EGFR 是一种酪氨酸激酶受体,最近研究证明了 EGFR 及其同源配体之一:转化生长因子-α(TGF-α),

与软骨内骨化中的软骨基质降解有关。EGFR 在关节软骨发育过程中维持浅表软骨细胞中起着至关重要的作用^[8],是细胞增殖、分化、分裂、存活和癌症发展的关键调节因子^[9],可保护软骨细胞免受 OA 诱导的细胞凋亡^[10]。马小远^[11]通过研究认为,EGFR 通路与保持表层细胞活性有关,比如保持细胞活性、抑制软骨外基质,可以防止骨性关节炎的发生和进展。此外,ZHANG 等^[10]的研究发现,EGFR 通过调节软骨细胞存活和软骨降解,在 OA 进展过程中发挥了主要的保护作用。

本研究中另一个关键基因为 PIK3CA,目前,大部分相关研究证实 PIK3CA 基因的改变与实体恶性肿瘤预后不良有关,有研究中的可用数据强调了 PIK3CA 突变通过增加 PI3K/Akt/mTOR 生存途径的信号传导介导肿瘤发生的重要作用^[12],但关于 PIK3CA 与 OA 的报道较少。WANG 等^[13]报道 PI3K/Akt/mTOR 的遗传变异与中国汉族人群中 OA 易感性有关。有研究还发现,Akt 信号通路的激活促进软骨细胞合成代谢,而 Akt 信号通路活性受损则抑制软骨细胞合成代谢。SUN 等^[14]认为,根据现有证据,PI3K/Akt/mTOR 信号通路对关节组织的正常代谢至关重要,但也参与 OA 的发展,强调了 PI3K/AKT/mTOR 信号传导在软骨退化、软骨下骨功能障碍和滑膜炎中的作用,所以 PIK3CA 的改变是否通过此生物途径的信号传导而导致 OA 的发生,抑制该通路活性是否对于防治 OA 有所帮助,还有待于进一步研究。

本研究获得的最后一个关键基因为 SRC,其是一种酪氨酸蛋白激酶,可调节细胞代谢,存活和增殖。研究表明,SRC 在巨噬细胞的功能激活中起着关键作用,而巨噬细胞参与各种免疫反应和炎症性疾病,包括类风湿性关节炎,癌症、骨质疏松症等^[15]。SUSVA 等^[16]通过研究发现,SRC 在骨代谢中的作用在 SRC 缺乏的小鼠中变得明显。有研究表明,SRC 抑制剂可能在抑制骨吸收、缓解 OA 进程中发挥作用。此外,SRC 抑制剂能使骨软骨外植体在 24 h 内减少软骨细胞死亡^[17]。

综上所述,本研究利用 GEO 数据库,运用生物信息学方法共筛选出与 OA 进展相关的关键基因共 3 个,分别为 PIK3CA、EGFR、SRC,这些关键基因的差异表达可能与 OA 的发生有着紧密关系,为该病防治的潜在分子靶点提供新的见解。因此,本研究结果可为 OA 的治疗和后续的研究提供进一步的参考。

参考文献

[1] 中华医学会骨科学分会关节外科学组,中国医师协会骨科医师分会骨关节炎学组,国家老年病临床医学研究中心(湘雅医院),等.中国骨关节

炎诊疗指南(2021 年版)[J]. 中华骨科杂志, 2021,41(18):1291-1314.

[2] GLYN-JONES S, PALMER A J, AGRICOLA R, et al. Osteoarthritis [J]. Lancet, 2015, 386 (9991):376-387.

[3] 曾文波, 兰生辉, 蔡贤华, 刘曦明. EGFR 信号通路在骨性关节炎研究中进展[J/CD]. 中华老年骨科与康复电子杂志, 2019, 5(2):109-113.

[4] 王景红, 夏坤, 张志千, 等. 骨关节炎相关细胞因子及生物标志物的研究进展[J]. 中国实验方剂学杂志, 2015, 21(10):225-230.

[5] 孙伟, 孙珺, 薛骋, 等. 基于基因芯片筛选恶性周围神经鞘瘤的核心基因和通路[J]. 徐州医科大学学报, 2019, 39(11):836-842.

[6] The Gene Ontology C. The gene ontology resource:20 years and still going strong[J]. Nucleic Acids Res, 2019, 47(D1):D330-D338.

[7] KANEHISA M, SATO Y, FURUMICHI M, et al. New approach for understanding genome variations in KEGG [J]. Nucleic Acids Res, 2019, 47(D1):D590-D595.

[8] WEI Y, MA X, SUN H, et al. EGFR signaling is required for maintaining adult cartilage homeostasis and attenuating osteoarthritis progression[J]. J Bone Miner Res, 2022, 37 (5): 1012-1023.

[9] SABBAH D A, HAJJO R, SWEIDAN K. Review on epidermal growth factor receptor (EGFR) structure, signaling pathways, interactions, and recent updates of EGFR inhibitors [J]. Curr Top Med Chem, 2020, 20 (10): 815-834.

[10] ZHANG X, ZHU J, LIU F, et al. Reduced EGFR signaling enhances cartilage destruction in a mouse osteoarthritis model [J]. Bone Res, 2014, 2:14015.

[11] 马小远. EGFR 通路在骨关节炎中的作用研究[D]. 济南:山东大学, 2017.

[12] ALQAHTANI A, AYESH H S K, HALAWANI H. PIK3CA gene mutations in solid malignancies: association with clinicopathological parameters and prognosis [J]. Cancers (Basel), 2019, 12(1):93.

[13] WANG K, CHU M, WANG F, et al. Putative functional variants of PI3K/AKT/mTOR pathway are associated with knee osteoarthritis susceptibility [J]. J Clin Lab Anal, 2020, 34 (6):e23240.

[14] SUN K, LUO J, GUO J, et al. The PI3K/AKT/mTOR signaling pathway in osteoarthritis: a narrative review [J]. Osteoarth Cartil, 2020, 28 (4):400-409.

[15] BYEON S E, YI Y S, OH J, et al. The role of SRC kinase in macrophage-mediated inflammatory responses [J]. Med Inflamm, 2012, 2012: 512926.

[16] SUSVA M, MISSBACH M, GREEN J. SRC inhibitors: Drugs for the treatment of osteoporosis, cancer or both? [J]. Trends Pharmacol Sci, 2000, 21(12):489-495.

[17] JANG K W, BUCKWALTER J A, MARTIN J A. Inhibition of cell-matrix adhesions prevents cartilage chondrocyte death following impact injury [J]. J Orthop Res, 2014, 32(3):448-454.

(收稿日期:2022-10-20 修回日期:2023-02-01)

(上接第 1828 页)

[12] XYLOURGIDIS N, MIN K, AHANGARI F, et al. Role of dual-specificity protein phosphatase DUSP10/MKP-5 in pulmonary fibrosis [J]. Am J Physiol Lung Cell Mol Physiol, 2019, 317(5): L678-689.

[13] HAN J, HE Y, ZHAO H, et al. Hypoxia inducible factor-1 promotes liver fibrosis in nonalcoholic fatty liver disease by activating PTEN/p65 signaling pathway [J]. J Cell biochemistry, 2019, 120(9):14735-14744.

[14] PATERNOSTRO R, HEINISCH B B, REIBERG-ER T, et al. Dysbalanced sex hormone status is an independent predictor of decompensation and mortality in patients with liver cirrhosis [J]. Hepatol Res, 2019, 49(2):201-211.

[15] ZHANG M, SUI W, XING Y, et al. Angiotensin IV attenuates diabetic cardiomyopathy via suppressing FoxO1-induced excessive autophagy, apoptosis and fibrosis [J]. Theranostics, 2021, 11(18):8624-8639.

(收稿日期:2022-10-12 修回日期:2023-01-11)