

· 论 著 ·

绝经后骨质疏松生物标志物的筛选及免疫细胞浸润分析

肖贺方¹ 连笑宇¹ 车凌仪² 强硕¹ 郑稼^{1*}

1. 郑州大学人民医院 河南省人民医院骨科, 河南 郑州 450003

2. 郑州大学第一附属医院妇产科, 河南 郑州 450003

中图分类号: R318.04 文献标识码: A 文章编号: 1006-7108(2021) 03-0379-06

摘要: 目的 绝经后骨质疏松症 (PMOP) 是绝经后妇女的常见疾病, 但目前尚无有效的治疗方法。本研究旨在识别绝经后骨质疏松症形成过程中基因表达的重要变化, 并提供相应的临床参考。方法 从 GEO 下载的数据集 GSE56116、京都基因和基因组百科全书 (Kyoto Encyclopedia of Genes and Genomes KEGG) 扩增途径、基因本体学 (GO)、蛋白质-蛋白质相互作用 (PPI) 网络, 筛选其差异表达基因 (DEG), 并进一步分析得到了六个关键基因, 分别是 FOS、SYK、HCK、SELL、CCR1、NLRP3, 然后我们分析了免疫细胞浸润。结果 发现在骨关节炎和正常对照之间差异很大, 与正常组织相比, PMOP 组织含有较低比例的活性 CD4T 记忆细胞 ($P < 0.05$), 最后从 GSE7429 数据集验证了关键基因 NLRP3 的表达水平。结论 有助于我们进一步了解 PMOP 形成的分子机制, 并为疾病的治疗提供新的思路。

关键词: 绝经后骨质疏松症; 免疫细胞渗透; 生物信息学; GEO

Identification of key biomarkers and analysis of immune infiltration of cells in postmenopausal osteoporosis

XIAO Hefang¹, LIAN Xiaoyu¹, CHE Lingyi², QIANG Shuo¹, ZHENG Jia^{1*}

1. Department of Orthopedics, Zhengzhou University People's Hospital, Henan Provincial People's Hospital, Zhengzhou 450003, China

2. Department of Gynecology, the First Affiliated Hospital of Zhengzhou University, Zhengzhou, 450003, China

* Corresponding author: ZHENG Jia, Email: zhengjia90180@sina.com.

Abstract: Objective Postmenopausal osteoporosis (PMOP) is a usual clinical disease in postmenopausal women, but there is no effective therapy at present. This research designed to recognize the important changes in gene expression, and to provide clinical reference. **Methods** We downloaded gene expression dataset GSE56116 from GEO, and dissected of differentially expressed genes (DEG), Kyoto Gene and Genomic Encyclopedia (KEGG) expansion pathway, Gene Ontology (GO), and protein-protein interaction (PPI) cyber. In total, six hub genes, including FOS, SYK, HCK, SELL, CCR1, and NLRP3, were elected. **Results** The immune infiltration profiles varied significantly between osteoarthritis and normal controls. Compared with normal tissues, PMOP tissues contained a lower proportion of activated T cells CD4 memory ($P < 0.05$). Finally, the expression levels of NLRP3 hub genes were confirmed with GSE7429. **Conclusion** In summary, this research helps us to realize the molecular mechanism of PMOP generation and offers a new direction for the treatment of PMOP.

Key words: postmenopausal osteoporosis; infiltration of immune cells; bioinformatics; Gene Expression Comprehensive Database

绝经后骨质疏松症 (postmenopausal osteoporosis, PMOP), 是一种无症状全身渐进性疾病, 特点是骨密度低、骨脆性增加, 从而骨折的发生

率增高。疾病的关键环节是因绝经期卵泡刺激素的量增加和雌激素缺乏, 从而导致严重的骨量丢失^[1], 研究表明成骨细胞和破骨细胞的功能对骨骼的修复和重塑至关重要^[2]。目前, 常用的治疗方法大多使用利塞膦酸盐, 唑来膦酸等; 最近新增的一些生物疗法如选择性雌激素受体拮抗剂^[3]。使用双

基金项目: 科技部国家重点研发计划 (122300411149)

* 通信作者: 郑稼, Email: zhengjia90180@sina.com

磷酸化合物治疗会产生许多不良反应,包括肌肉骨骼疼痛、下颌骨折等问题^[4]。在美国,每年因骨质疏松症,可导致约150万例骨折的发生^[5]。然而,作为一个全球性的问题,目前仍然缺乏有效的诊断和治疗。因此,研究PMOP形成的分子机制具有十分重要的作用。在之前的文章中,一些与PMOP形成相关的潜在的差异表达基因、中心基因和信号通路尚未得到确认,此外,没有利用GEO数据库,来研究此问题。

1 材料和方法

1.1 数据来源

GEO(<http://www.ncbi.nlm.nih.gov/geo/>)^[6]是一个公共基因组学数据存储库。数据集GSE56116下载于GPL4133平台,数据集有13个样本,包含3个健康的和10个疾病的样本。

1.2 差异分析

使用了R软件(版本3.6.1; <https://www.r-project.org/>)和对应的R包进行分析,在R软件中执行以下参数,包括分位数、RMA、Median polish和pmonly。再用LIMMA包进行差异分析,筛选标准: $\log_2$ fold change(FC)>1和 P value<0.05。

1.3 富集分析

通过数据库DAVID用于探索分析差异基因的功能^[7]。GO(gene ontology)数据库里面包括有描述基因基本特征的词汇或结构^[8]。KEGG(Kyoto Encyclopedia of Genes and Genomes)数据库从基因组、系统性和化学功能通路等方面综合信息^[9]。使用DAVID软件,把所有的差异基因进行功能富集分析。

1.4 构建蛋白质互作网络(PPI)和筛选关键基因

PPI网络通过寻找互作基因来进行预测(String; <http://string-db.org>) (版本12.0)^[10]。PPI网络由Cytoscape绘制,最重要的模块由MCODE插件进行确认。标准:degree cut-off=2, MCODE scores>4, Max depth=100, node score cut-off=0.3, 和k-score=2。然后,使用Hubba插件挑选中心基因。使用DAVID对网络中关键模块中的基因进行KEGG和GO分析。使用五种方法(degree、BottleNeck、Radiality centrality、Stress centrality、Closeness centrality)来排列和筛选中心基因^[11]。

1.5 用CIBERSORT算法分析免疫细胞的浸润

利用CIBERSORT算法对获得的数据进行了分

析,并得到了22种免疫细胞的比例。基于 P 值<0.05筛选,并计算样本中每个免疫细胞的百分比。通过R软件中的vioplot包,分析两组免疫细胞的不同免疫渗透水平。

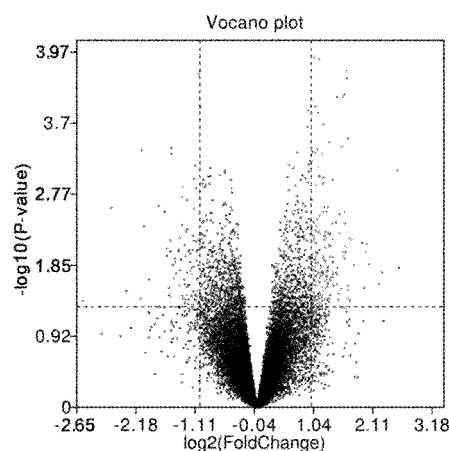
1.6 在数据库中进行验证

GSE7429数据集是从GEO数据库获得,使用此数据集验证得到的关键基因。

2 结果

2.1 确定差异基因

从数据中筛选出了308个差异基因,其中110个下调,198个上调。图1和图2分别显示了所有DEG的火山图和前50个下调基因和50个上调基因的热图。



注:红色的差异基因代表 $\log_2$ FC>1并且 P <0.05,蓝色的差异基因代表 $\log_2$ FC<-1并且 P <0.05。

图1 所有差异基因的火山图

Fig.1 The volcano plot of all DEGs

2.2 GO和KEGG的富集分析

使用DAVID对差异基因进行GO和KEGG富集分析,如图3所示,分别显示了在GO分析中的前10个富集的部分。生物过程(BP)分析结果表明,DEG在先天免疫反应(innate immune response)、炎症反应(inflammatory response)中明显富集(见图3a);分子功能(MF)分析表明,DEG在碳水化合物结合(carbohydrate binding)、受体活性(receptor activity)、转录活化活性(transcriptional activator activity)、受体结合(receptor binding)富集显著(见图3b);对于细胞组分(CC)分析表明,DEG主要富集于等离子膜(plasma membrane),等离子膜的整体成分(integral component of plasma membrane),见图3c。

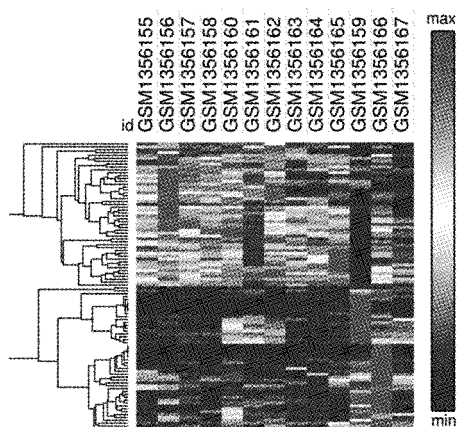


图 2 红色表示热图中前 50 上调的差异基因, 蓝色表示前 50 个下调的差异基因

Fig.2 The heat map of the top 50 upgrade genes in red and 50 downgrade genes in blue

2.3 KEGG 通路分析

KEGG 通路富集分析表明, DEG 主要在破骨细胞分化 (osteoclast differentiation)、造血细胞系 (hematopoietic cell lineage)、细胞粘附分子 (cell adhesion molecules CAMs)、甲状腺素合成 (thyroid hormone synthesis) 途径中富集显著。在这些通路中破骨细胞分化的通路占主要部分。KEGG 通路富集分析结果见图 3 d。

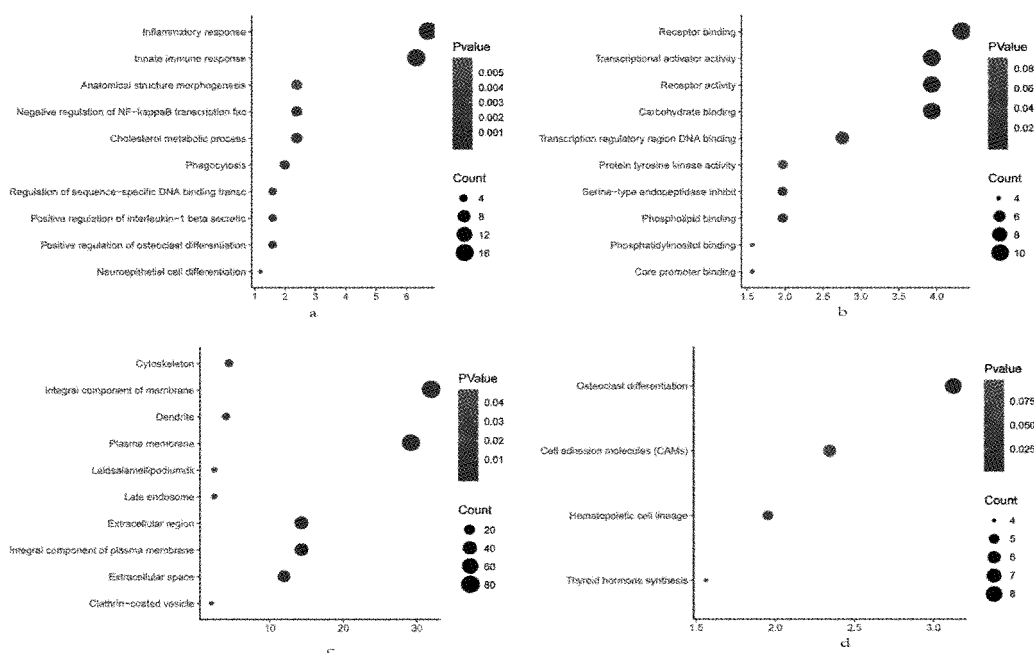
2.4 构建 PPI 网络

节点表示 DEG 和边缘表示 DEG 之间的交互连接, 根据 STRING 数据库中的信息, 我们使用 Cytoscape 中的插件 MCODE 在 PPI 网络中查找 165 个节点和 339 个边 (见图 4a) 和得到了最显著的模块 (见图 4b)。

使用 DAVID 对最显著模块中基因进行了富集分析。结果表明, 该模块中的基因主要在 GO 的炎症反应 (inflammatory response)、先天免疫反应 (innate immune response) 和血浆膜 (plasma membrane) 中富集显著, 而在 KEGG 中主要在 Fc 伽马 R 介导的噬菌体 (Fc gamma R-mediated phagocytosis)、骨细胞分化 (osteoclast differentiation) 和化学因子信号 (chemokine signaling) 等通路中富集, 见表 1。

2.5 确定关键基因

使用 Cytoscape 中的 cytoHubba 插件来确定关键基因。根据 cytoHubba 的 5 种分类方法筛选中心基因, 表 2 显示了分别用这些方法选取的 15 个中心基因。最后, 把 5 种方法筛选出的基因进行重叠得到了 6 个关键基因 (图 5)。这 6 个基因都存在于这 5 个方法中, 分别是: Fos、SYK、HCK、SELL、CCR1、NLRP3。



注: a 生物过程; b 分子功能; c 细胞组分; d 上、下调的差异基因的 KEGG 通路分析。

图 3 差异基因的 GO 和 KEGG 的富集分析

Fig.3 The Gene Ontology (GO) analysis of differentially expressed genes

表1 对关键模块的差异基因进行GO和KEGG富集分析的前10项富集结果

Table 1 The top 10 Gene Ontology (GO) functions and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways enriched for the genes involved in significant module

项目	P 值	数量	包含的基因
GO:0006954 inflammatory response (BP)	2.36×10^{-6}	6	SIGLEC1, CCR1, HCK, NLRP3, SYK, S100A12
GO:0045087 innate immune response (BP)	0.002360	4	NCF1, NLRP3, SYK, S100A12
GO:0005886 plasma membrane (CC)	0.004144	8	SIGLEC1, LILRB2, CD93, CD33, CCR1, SYK, CD163, S100A12
GO:0050764 regulation of phagocytosis (BP)	0.005881	2	HCK, SYK
GO:0030246 carbohydrate binding (MF)	0.006885	3	SIGLEC1, CD93, CD33
GO:0004872 receptor activity (MF)	0.008381	3	LILRB2, CD93, CD33
GO:0007267 cell-cell signaling (BP)	0.011450	3	LILRB2, CD33, CCR1
GO:0016032 viral process (BP)	0.015630	3	CD93, HCK, SYK
GO:0051090 regulation of sequence-specific DNA binding transcription factor activity (BP)	0.016260	2	HCK, SYK
GO:0038083 ~peptidyl-tyrosine autophosphorylation (BP)	0.025900	2	HCK, SYK
hsa04666: Fc gamma R-mediated phagocytosis	0.003933	3	NCF1, HCK, SYK
hsa04380: Osteoclast differentiation	0.009348	3	LILRB2, NCF1, SYK
hsa04062: Chemokine signaling pathway	0.018290	3	NCF1, CCR1, HCK

表2 用 cytoHubba 筛选的 15 个中心基因

Table 2 The top 15 hub genes rank in cytoHubba

Radiality	Degree	Closeness	Stress	BottleNeck
FOS	LILRB2	LILRB2	FOS	FOS
SYK	CYBB	CYBB	SYK	SYK
CYBB	FOS	FOS	LILRB2	HCK
LILRB2	HCK	SYK	CTSS	NLRP3
HCK	CD163	HCK	CYBB	NTF3
SELL	SYK	CD33	HCK	CTSS
CD33	HK3	SELL	NLRP3	FGF8
CEBPA	SELL	CCR1	SELL	SELL
CCR1	CCR1	NLRP3	CD163	CCR1
KLF4	CD33	CD163	CD33	SGK1
PLAUR	CD93	CEBPA	CCR1	GPRI1
NLRP3	NLRP3	HK3	SGK1	AP5B1
NCF1	SLC11A1	CD93	FGF8	SGIP1
GPRI1	CTSS	PLAUR	KLF4	KLF4
CD1D	CD14	CD14	HK3	POLR2A

2.6 免疫细胞浸润分析

利用 CIBERSORT 算法,研究了 PMOP 的 22 个免疫细胞集与正常组织免疫渗透的差异。图 6a 总结了 3 例正常对照组和 10 例 PMOP 患者的结果。与正常组织相比,PMOP 组织通常含有较低比例的活性 CD4T 记忆细胞 (activated T cells CD4 memory),见图 6b。

2.7 对差异基因的验证

NLRP3 是我们得到的关键基因,基于 GSE7429 数据集对 NLRP3 的表达进行验证(图 7)。这一结果与我们综合分析的结果基本一致。

3 讨论

GO 分析表明,DEG 主要参与生物过程,基因主要在先天免疫反应中富集,可能是因为缺乏维生素

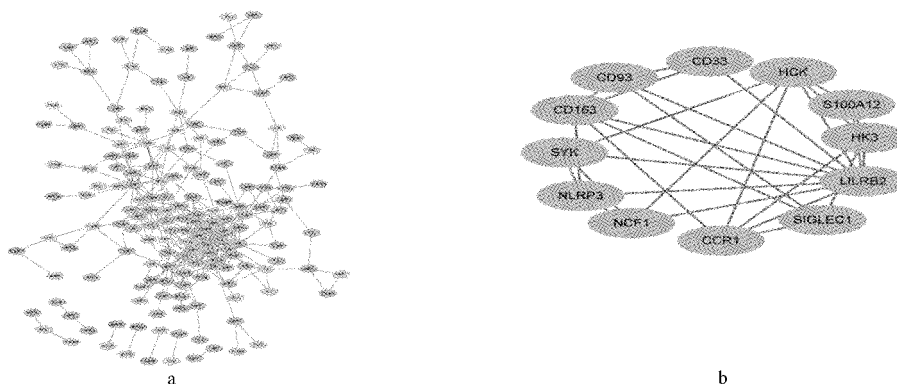


图4 a 使用 Cytoscape 构建的差异基因的蛋白质互作网络; b 从蛋白质互作网络选取的最重要的模块

Fig.4 (a) The PPI network of DEGs was constructed using Cytoscape. (b) The most significant module was obtained from PPI network

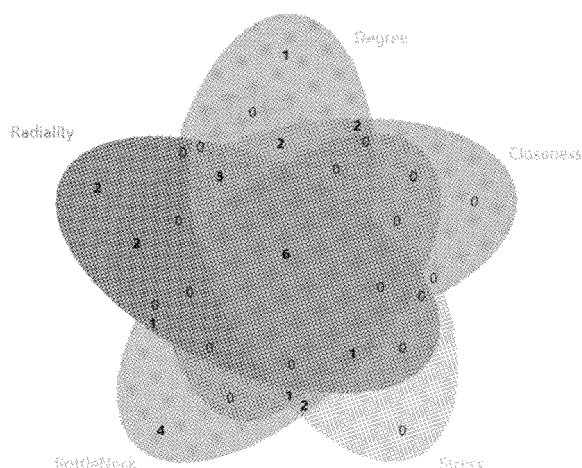


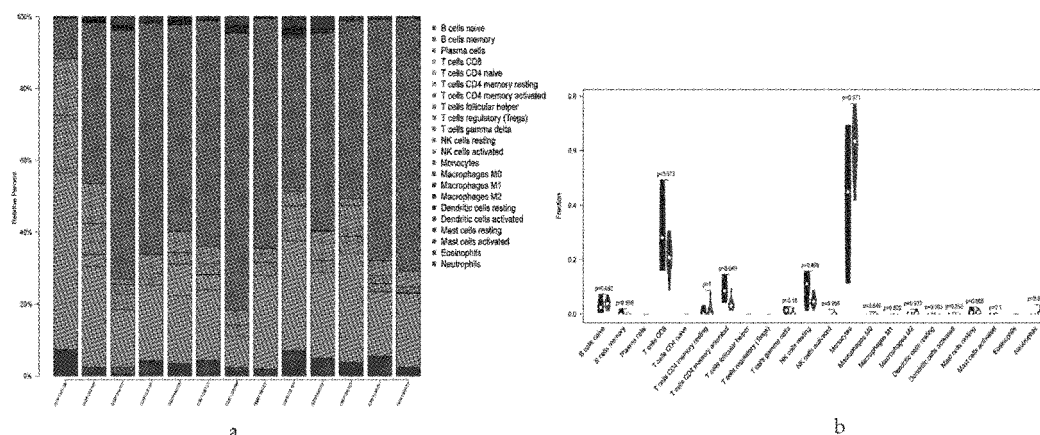
图5 使用重叠方法对 cytoHubba 中筛选出的 15 个基因进行重叠确定了 6 个关键基因

Fig.5 Six hub genes were identified by overlapping the first 15 genes in the cytoHubba

D,影响骨骼的代谢所产生的结果,维生素 D 受体由大多数免疫细胞表达,包括 B 和 T 淋巴细胞和树突状细胞,其次,免疫细胞具有活跃的维生素 D 代谢功能,可以将 $25(\text{OH})\text{D}_3$ 局部转移到 $1,25(\text{OH})_2\text{D}_3$ 中^[12],表明维生素 D 在调节免疫功能方面起着关键作用。此外,之前研究表明先天性免疫系统和肠道微生物群之间的通信障碍可能导致复杂的疾病^[13],而且肠道微生物群的变化与老年人骨矿物质密度的降低有关^[14],这些研究也侧面的证实了我们实验的富集功能结果,我们推测,这些基因是通过先天免疫系统影响肠道微生物群,进而影响雌激素对破骨细胞的影响,最终导致骨质疏松症的发生^[15]。

KEGG 通路分析表明,差异基因主要参与甲状腺素合成等过程。骨骼是 T3 作用的关键目标组织^[16],甲状腺素短缺和过量都与骨折风险增高有关^[17]。因此,了解 T3 的分子作用机制可以指导骨质疏松症的控制和治疗。在关键模块中,我们发现,该模块的差异基因在 GO 富集分析中炎症反应、先天免疫反应处富集最为显著,这与我们之前的功能富集分析结果一致。此外,我们确定了 6 个关键基因: FOS、SYK、HCK、SELL、CCR1、NLRP3。其中 NLRP3 是最重要的一個关键基因,此基因在骨侵蚀水平中起着重要作用, NLRP3 的功能表达可能是导致感染部位骨质丢失的关键因素^[18]。先前的研究发现将 SYK 募集至磷酸化的 ITAM 对破骨细胞的形成至关重要^[19],SYK 是嗜中性粒细胞中整合信号传导的重要组成部分^[20]。HCK 和 SRC 在成骨细胞中具有部分重叠的功能,而 HCK 在成骨的表达中可改善其功能缺陷^[21]。CCR1 可以改变成骨细胞和破骨细胞的功能和分化,以及改变成骨细胞和脂肪形成之间的不平衡^[22],关节炎患者的破骨细胞前体细胞高表达 CCR1^[23]。因此,对这些基因的深入了解有助于阐明 PMOP 形成的关键机制。然而,我们研究的局限性是,没有进一步的临床试验来验证它,但它仍然会对以后的研究产生很大的帮助。

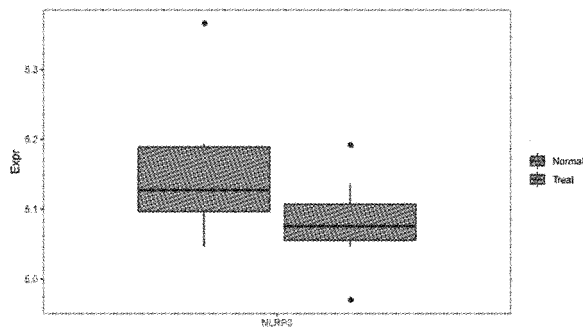
综上,我们确定了 308 个差异基因、6 个关键基因和一个重要网络模块,富集分析了差异基因的信号通路,通过将可靠的反卷积算法与大规模基因组数据相结合,发现了 PMOP 和正常对照之间的免疫渗透差异。这为临床诊断和治疗提供了可靠的参考价值。



注: a GSE56116 数据集中 13 个样本中 22 个免疫细胞亚群的相对百分比; b PMOP 与正常对照组免疫浸润的差异。正常对照组为蓝色, PMOP 组为红色, $P < 0.05$ 。

图6 PMOP 患者与正常对照组的免疫浸润情况

Fig.6 The landscape of immune infiltration between PMOP and normal controls



注: X轴表示基因, Y轴表示基因表达水平。

图7 验证在 GSE7429 数据 P 集的关键基因的表达水平

Fig.7 Validation of the expression levels of selected based on GSE7429

【参 考 文 献】

- [1] Faienza MF, Ventura A, Marzano F, et al. Postmenopausal osteoporosis: the role of immune system cells[J]. Clin Develop Immunol, 2013, 2013:575936.
- [2] 罗丽梅,李杰,杨帆,等.骨质疏松发病机制及治疗药物研究进展[J].中国骨质疏松杂志,2020,26(4):610-614,624
- [3] Reginster JY. Antifracture efficacy of currently available therapies for postmenopausal osteoporosis[J]. Drugs, 2011, 71:65-78.
- [4] Bone HG, McClung MR, Roux C, et al. Odanacatib, a cathepsin-K inhibitor for osteoporosis: a two-year study in postmenopausal women with low bone density[J]. J Bone Miner Res, 2010, 25:937-947.
- [5] Savic Pavicic I, Dumancic J, Jukic T, et al. Digital orthopantomograms in osteoporosis detection: mandibular density and mandibular radiographic indices as skeletal BMD predictors[J]. Dentomaxillofac Radiol, 2014, 43:20130366.
- [6] Edgar R, Domrachev M, Lash AE. Gene expression omnibus: NCBI gene expression and hybridization array data repository[J]. Nucleic Acids Res, 2002, 30:207-210.
- [7] Huang DW, Sherman BT, Tan Q, et al. DAVID bioinformatics resources: expanded annotation database and novel algorithms to better extract biology from large gene lists[J]. Nucleic Acids Res, 2007, 35(Web Server issue):W169-75.
- [8] Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology. The gene ontology consortium[J]. Nat Genet, 2000, 25:25-29.
- [9] Kanehisa M, Araki M, Goto S, et al. KEGG for linking genomes to life and the environment[J]. Nucleic Acids Res, 2008, 36: D480-484.
- [10] Franceschini A, Szklarczyk D, Frankild S, et al. STRING v9.1: protein-protein interaction networks, with increased coverage and integration[J]. Nucleic Acids Res, 2013, 41:D808-815.
- [11] Chin CH, Chen SH, Wu HH, et al. CytoHubba: identifying hub objects and sub-networks from complex interactome[J]. BMC Syst Biol, 2014, 8 Suppl 4:S11.
- [12] Sassi F, Tamone C, D'Amelio P. Vitamin D: Nutrient, Hormone, and Immunomodulator[J]. Nutrients, 2018, 10(11):1656.
- [13] Thaiss CA, Zmora N, Levy M, et al. The microbiome and innate immunity[J]. Nature, 2016, 535:65-74.
- [14] Das M, Cronin O, Keohane DM, et al. Gut microbiota alterations associated with reduced bone mineral density in older adults[J]. Rheumatology (Oxford), 2019, 58(12):2295-2304.
- [15] Li JY, Chassaing B, Tyagi AM, et al. Sex steroid deficiency-associated bone loss is microbiota dependent and prevented by probiotics[J]. J Clinical Invest, 2016, 126:2049-2063.
- [16] Williams GR. Thyroid hormone actions in cartilage and bone[J]. Eur Thyroid J, 2013, 22:3-13.
- [17] Wojcicka A, Bassett JH, Williams GR. Mechanisms of action of thyroid hormones in the skeleton[J]. Biochim Biophys Acta, 2013, 1830:3979-3986.
- [18] McCall SH, Sahraei M, Young AB, et al. Osteoblasts express NLRP3, a nucleotide-binding domain and leucine-rich repeat region containing receptor implicated in bacterially induced cell death[J]. J Bone Miner Res, 2008, 23:30-40.
- [19] Mócsai A, Humphrey MB, Van Ziffle JA, et al. The immunomodulatory adapter proteins DAP12 and Fc receptor gamma-chain (FcRgamma) regulate development of functional osteoclasts through the Syk tyrosine kinase[J]. Proceed Natl Acad Sci U S A, 2004, 101:6158-6163.
- [20] Mócsai A, Zhou M, Meng F, et al. Syk is required for integrin signaling in neutrophils[J]. Immunity, 2002, 16:547-558.
- [21] Lowell CA, Niwa M, Soriano P, et al. Deficiency of the Hck and Src tyrosine kinases results in extreme levels of extramedullary hematopoiesis[J]. Blood, 1996, 87:1780-1792.
- [22] Hopwood B, Tsykin A, Findlay DM, et al. Gene expression profile of the bone microenvironment in human fragility fracture bone[J]. Bone, 2009, 44:87-101.
- [23] Sucur A, Jajic Z, Artukovic M, et al. Chemokine signals are crucial for enhanced homing and differentiation of circulating osteoclast progenitor cells[J]. Arthritis Res Ther, 2017, 19:142.

(收稿日期: 2020-04-28;修回日期: 2020-07-01)