

· 论 著 ·

婴幼儿血管瘤异常表达基因分析*

何丹华, 陈知英, 陈国菊

(暨南大学附属深圳市宝安区妇幼保健院皮肤美容科, 广东 深圳 518102)

[摘要] **目的** 通过生物信息学方法探索婴幼儿血管瘤(IH)的潜在生物标志物。**方法** 2022 年 2 月从 GEO 数据库中下载基因表达谱数据集(GSE127487)。利用 R 语言 limma 软件包分析 IH 组和健康对照组差异表达基因(DEGs)。对 DEGs 进行基因本体和京都基因与基因组百科全书富集分析。通过 STRING 分析蛋白质互作网络,利用 Cytoscape 软件根据不同算法筛选关键基因——hub 基因。计算各个 hub 基因在 IH 组和健康对照组中的表达,最后通过受试者工作特征(ROC)曲线验证 hub 基因的诊断价值。**结果** GSE127487 中有 416 个 DEGs,主要富集于血管发育、血管生成、心血管系统的发育等,参与细胞外基质结构形成、分子功能调控、生长因子结合等,富集于白细胞跨内皮的迁移、细胞黏附分子、血管平滑肌收缩等信号通路。共筛选出 PECAM1、CD34、CDH5、TEK、THY1 5 个 hub 基因,并且在 IH 组中的表达均上调,且 ROC 曲线的曲线下面积均大于 0.7。**结论** IH 的 hub 基因与疾病发生、发展相关,是 IH 潜在的生物标志物,有望成为 IH 诊治的靶点。

[关键词] 婴幼儿血管瘤; 生物信息学; 基因表达; 生物学标记
DOI:10.3969/j.issn.1009-5519.2023.02.004 **中图法分类号:**R725.4
文章编号:1009-5519(2023)02-0199-05 **文献标识码:**A

Analysis of abnormal expression gene in infantile hemangioma*

HE Danhua, CHEN Zhiying, CHEN Guoju

(Department of Cosmetic Dermatology, Shenzhen Baoan Women's and Children's Hospital, Jinan University, Shenzhen, Guangdong 518102, China)

[Abstract] **Objective** To explore the potential biomarkers of infantile hemangioma (IH) by bioinformatics methods. **Methods** The gene expression profile dataset (GSE127487) was downloaded from the GEO database in February 2022. The differentially expressed genes (DEGs) in the IH group and the healthy control group were analyzed with limma software package of R. Gene ontology and Kyoto Encyclopedia of Genes and Genomes enrichment analyses were performed on DEGs. The protein interaction network was analyzed by STRING, and hub gene, the key gene, was selected by different algorithms using Cytoscape software. The expression of each hub gene in the IH group and the healthy control group was calculated, and finally the diagnostic value of hub gene was verified by receiver operating characteristic (ROC) curve. **Results** There were 416 DEGs in GSE127487, which were mainly concentrated in vascular development, angiogenesis, development of cardiovascular system, etc, involved in the formation of extracellular matrix structure, molecular function regulation, growth factor binding, etc, and were enriched in signaling pathways such as leukocyte transendothelial migration, cell adhesion molecules, and contraction of vascular smooth muscle. A total of five hub genes, including PECAM1, CD34, CDH5, TEK and THY1, were screened, and their expressions were up-regulated in the IH group, and the area under the ROC curve was greater than 0.7. **Conclusion** The hub gene of IH is related to the occurrence and development of disease, and is a potential biomarker of IH, which is expected to become a target for diagnosis and treatment of IH.

[Key words] Infantile hemangioma; Bioinformatics; Gene expression; Biological markers

婴幼儿血管瘤(IH)是婴幼儿最常见的良性肿瘤, 人、低体重早产儿、多胞胎、大龄婴儿更易患病^[2]。约发病率为 3%~10%^[1]。有研究发现,女性、高加索 10%的 IH 患儿在 1 年内可自行缓解,而其余 90%的

* 基金项目:广东省深圳市宝安区医疗卫生基础研究项目(2018JD173)。
作者简介:何丹华(1986—),硕士研究生,主治医师,主要从事儿童皮肤病诊治及皮肤激光美容的研究和临床工作。

IH 患儿需 3~10 年或更长时间才能好转^[3],部分患儿会留下瘢痕而影响美观^[4],甚至造成永久性毁容。IH 并发症包括溃疡、出血、视力障碍、气道阻塞、充血性心力衰竭等,在某些情况下 IH 甚至可能导致死亡^[5]。如能在 IH 早期积极干预、促进自发消退过程、减少血管瘤对周围组织的侵袭则能改善患儿预后。因此,有必要深入探讨 IH 的调控过程,以更好地理解 IH 的发病机制,对研发相关药物具有重要意义。本研究采用生物信息学方法筛选 IH 的差异表达基因(DEGs)及其功能富集通路,筛选靶基因,旨在为后续研究提供新思路。

1 资料与方法

1.1 研究对象 2022 年 2 月通过公开的 GEO 数据库(<https://www.ncbi.nlm.nih.gov/geo/>)检索与 IH 相关文献,获取基因表达谱数据集(GSE127487)进行分析。该基因表达谱种属为人类。样本来源于皮肤组织。芯片平台:GPL10558 Illumina HumanHT-12 V4.0 expression beadchip。选取该数据集中 IH 患儿 18 例,健康对照组研究对象 5 例。

1.2 方法

1.2.1 DEGs 的识别 通过 R 4.1.0 的 GEOquery 包从 GEO 数据库中下载 GSE127487,去除一个探针对应多个分子的探针,当遇到对应同一个分子的探针时仅保留信号值最大的探针。然后通过箱式图查看样本标准化的情况,利用 R 语言 limma 软件包进行 IH 组和健康对照组 DEGs 分析。统计学上显著的 DEGs 定义为 $|\log FC| \geq 2, P < 0.05$ 。

1.2.2 DEGs 功能和通路富集分析 采用 WebGestalt^[6](<http://www.webgestalt.org/>)在线工具对 DEGs 进行功能富集分析,即生物过程、细胞成分、分子功能及京都基因与基因组百科全书(KEGG)富集分析。

1.2.3 构建蛋白质互作网络(PPI) 通过在线分析工具 STRING 构建 PPI^[7]。由 Cytoscape 软件(3.8.1)对 PPI 进行可视化^[8]。利用 Cytohubba 筛选构成 PPI 稳定结构的 hub 基因,通过 MCC、NNC、Degree、EPC、BottleNeck 5 种算法,挑选出排名前 10 的 hub 基因,然后通过在线工具 InteractiVenn^[9]取交集后得到 hub 基因,并将其用于后续研究。

1.2.4 hub 基因表达量分析 通过 Graphpad prism 9 分析 hub 基因表达差异,以箱式图显示。

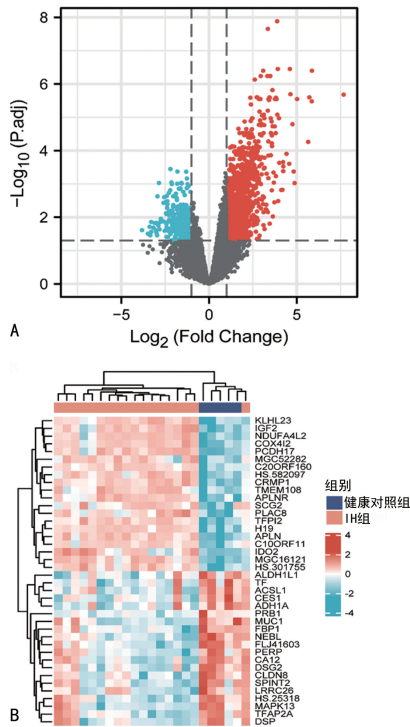
1.2.5 hub 基因的诊断效能评估 通过 R 4.1.0 进行诊断效能评估,pROC 包用于统计分析,ggplot2 进行可视化分析。最终绘制 5 个 hub 基因的受试者工作特征(ROC)曲线。

1.3 统计学处理 应用 Graphpad prism 9 软件进行

数据分析,计量资料以 $\bar{x} \pm s$ 表示,采用独立样本 t 检验;计数资料以率或构成比表示,采用 χ^2 检验。 $P < 0.05$ 表示差异有统计学意义。

2 结果

2.1 IH 的 DEGs GSE127487 中有 416 个 DEGs。其中 350 个基因表达上调,66 个基因表达下调。见图 1。



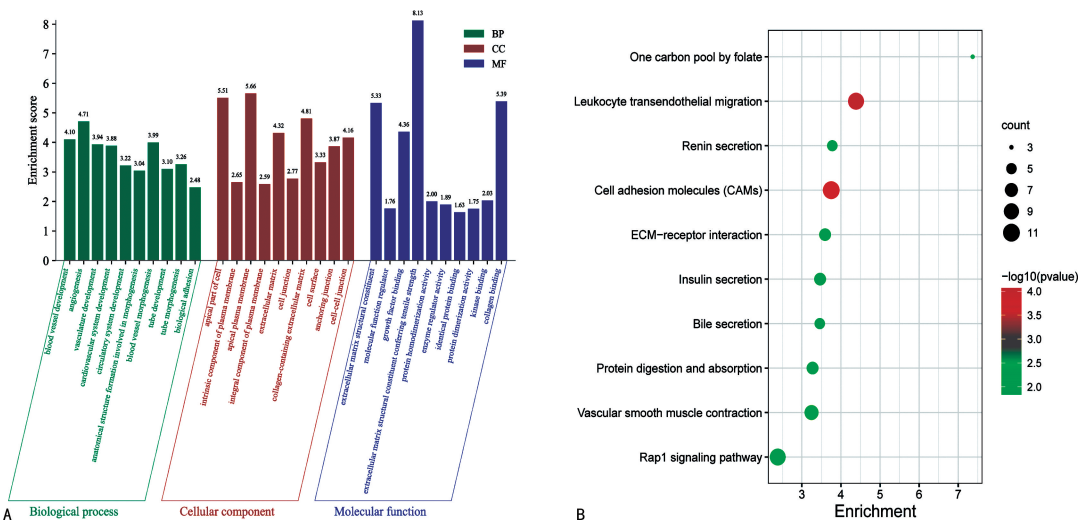
注:A.火山图;B.热图。

图 1 DEGs 火山图和热图

2.2 基因本体(GO)和 KEGG 富集分析 416 个 DEGs 均用 WebGestalt 在线工具进行分析。在生物过程中 DEGs 富集于血管发育、血管生成、心血管系统的发育、循环系统发育、参与形态发生的解剖结构形成等。在细胞外基质结构形成、分子功能调控、生长因子结合、具有抗拉强度的细胞外基质结构成分、蛋白二聚体活性等方面显著富集。DEGs 显著富集于细胞的顶端部分、质膜的固有成分、顶端质膜、质膜的组成部分、细胞外基质。主要富集在白细胞跨内皮的迁移、细胞黏附分子、血管平滑肌收缩、细胞外基质受体相互作用等信号通路。见图 2。

2.3 PPI 构建和 hub 基因筛选 共得到 5 个 hub 基因,分别为 PECAM1、CD34、CDH5、TEK、THY1。见图 3、4。

2.4 hub 基因表达量 与健康对照组比较,IH 组 PECAM1、CD34、CDH5、TEK、THY1 mRNA 表达均明显升高,差异均有统计学意义($P < 0.05$)。见图 5。



注:A. GO 分析;B. KEGG 分析。

图 2 DEGs 的 GO 分析和 KEGG 分析

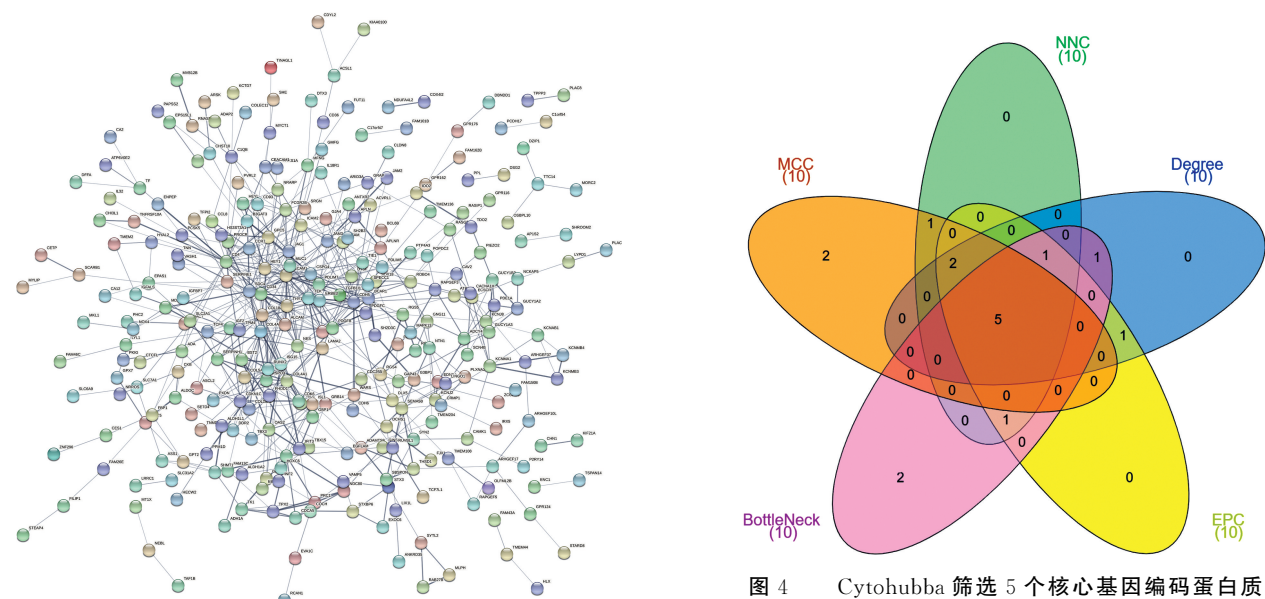
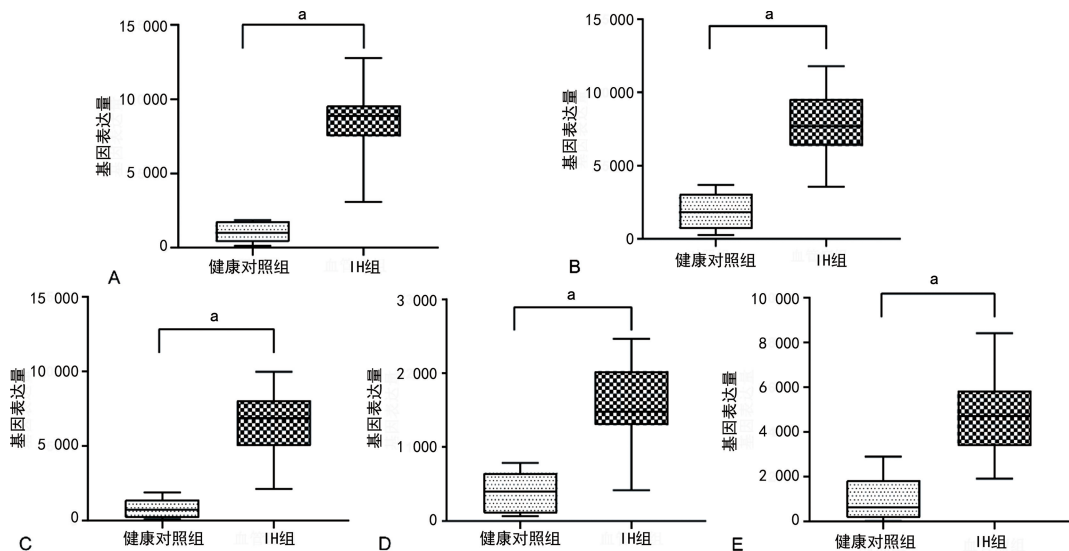


图 3 PPI

图 4 Cytohubba 筛选 5 个核心基因编码蛋白质



注:A. PECAM1;B. CD34;C. CDH5;D. TEK;E. THY1。与健康对照组比较,^a $P<0.05$ 。

图 5 hub 基因表达水平箱式图

2.5 诊断效能 PECAM1、CD34、CDH5、TEK、THY1 对 IH 均具有较好的诊断效能(ROC 曲线下面积分别为 1.000、0.991、1.000、0.965、0.965)。见图 6。

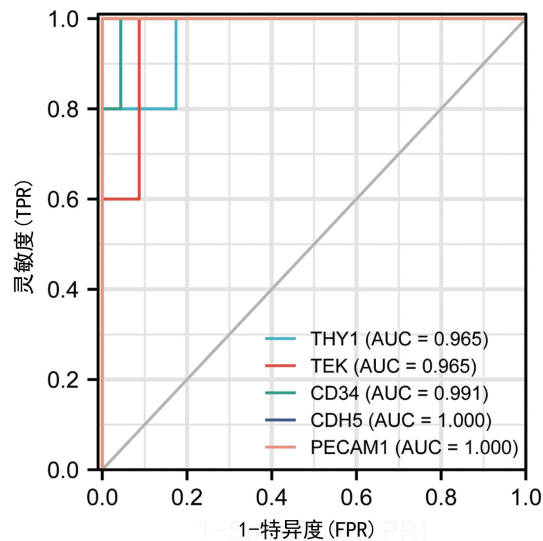


图 6 hub 基因诊断 IH 的 ROC 曲线

3 讨 论

IH 是婴幼儿最常见良性肿瘤之一,在正常情况下并不会出现,多在 3~12 个月内快速生长,部分患儿在 3~7 年后自然消失。在新生儿中发病率为 10%~12%,在早产儿中发病率为 20%,以女婴居多^[10]。尽管对 IH 进行了大量的研究,但 IH 发生、发展机制尚不明确。目前,对 IH 的药物及激光治疗研究较多,但从分子水平研究血管瘤发生、发展的重要基因和途径较少见。因此,IH 的生物信息学研究对阐明其发病的分子机制具有重要参考意义。

本研究通过对 IH 的微阵列数据集进行差异表达分析,分别通过火山图和热图展示了表达上调和下调的基因。对 DEGs 进行 GO 和 KEGG 分析发现,主要富集于血管发育、血管生成、细胞外基质结构形成、分子功能调控、生长因子结合等,主要参与白细胞跨内皮的迁移、细胞黏附分子、血管平滑肌收缩、细胞外基质受体相互作用等 KEGG 信号通路,符合血管瘤生长、发育过程。生物信息学方法能够将与每例患儿相关的结果和模式与所有其他患儿的结果和模式联系起来。通过生物信息学工具整合了来自分子、基因组和生物链接研究的数据。揭示了生物过程的复杂性和细节。这些生物系统的新概念将有助于理解和发现疾病的发病机制,设计高度准确的诊断方法,并开发新的治疗和预防策略^[11]。

本研究通过 STRING 在线工具及 Cytoscape 对 DEGs 进行 PPI 构建及可视化,并且筛选了 hub 基

因,结果显示,THY1、TEK、CD34、CDH5、PECAM1 在血管瘤患儿中其表达均明显升高。由于是取多个结果交集找到的 hub 基因,因此,具有较高的准确性。本研究结果显示,IH 组 hub 基因表达明显高于健康对照组,可作为 IH 的重要标志物。本研究判断了 hub 基因的诊断效能,通过 ROC 曲线发现,上述基因 ROC 曲线下面积均大于 0.7,可见其具有较强的诊断效能。

PECAM1 是一种跨膜糖蛋白,属于免疫球蛋白家族,在血小板、单核细胞、中性粒细胞表面表达。PECAM1 分子通过所谓的嗜酸性连接相互结合,导致内皮形成一种特定的保护屏障。PECAM1 表达下调导致内皮屏障完整性破坏和通透性增加,从而导致动脉内膜炎症过程的激活,影响血管粥样硬化病变的形成和冠状动脉疾病的发展。因此,PECAM1 可能是 IH 的一种新的候选分子,为改善 IH 治疗策略提供了新的视角。

CD34 是一种高度糖基化的跨膜蛋白。首先在髓系白血病细胞系中被发现,抗 CD34 抗体的免疫组织学反应性也在非血淋巴肿瘤中被发现,包括血管肉瘤、孤立性纤维性肿瘤、上皮样肉瘤、梭形细胞脂肪瘤、隆突皮肤纤维肉瘤、肌成纤维细胞瘤等。因此,CD34 可能被认为是 IH 的一种新的生物标志物。

CDH5 定位于内皮细胞的细胞间连接,在控制血管完整性和通透性方面发挥重要作用,并有助于内皮细胞在管状结构中的整合^[12-13]。CDH5 是血管内皮生长因子在内皮细胞上的生存信号^[14]。CDH5 在肿瘤血管生成中也很重要,其在乳腺癌血管系统中表达上调,也被认为是乳腺癌的转移标志物^[15-16]。因此,CDH5 可能是 IH 进展和靶向治疗的重要分子。

TEK 是酪氨酸激酶受体亚家族成员,主要表达于内皮细胞中,是血管生成素配体 ANGPT1、ANGPT2、ANGPT4 的酪氨酸激酶受体^[17]。TEK 在调节血管生成、血管重塑、成熟和完整性方面发挥重要作用^[18]。提示 TEK 可能参与了 IH 的发病机制,对 TEK 的进一步研究可能为开发治疗 IH 及相关疾病的药物提供了新靶点。

THY1 也称为 CD90,是一种以糖基磷脂酰肌醇为基础的糖蛋白,主要表达于血干细胞、活化的微血管内皮细胞和成纤维细胞上,是胸腺细胞祖细胞的标志物。有研究发现,THY1 能促进肝纤维化发展,是肝纤维化驱动基因^[19]。THY1 可能与 IH 的发生、发展密切相关。

综上所述,本研究通过对 IH 与正常样本的 DEGs 进行一系列生物信息学分析,找出了一些与 IH

密切相关的候选基因和通路调控网络。结果提示, PECAM1、CD34、CDH5、TEK、THY1 参与了 IH 的发病过程,有望成为 IH 诊治的靶点,为进一步诊治 IH 提供了依据。然而,本研究的局限性是缺乏实验验证。这些通过生物信息学分析得到的预测结果还需要实时荧光定量-聚合酶链反应和免疫印迹试验等实验研究进一步证实。

参考文献

[1] TIEMANN L, HEIN S. Infantile hemangioma: A review of current pharmacotherapy treatment and practice pearls[J]. J Pediatr Pharmacol Ther, 2020, 25(7): 586-599.

[2] HAGGSTROM A N, DROLET B A, BASELGA E, et al. Prospective study of infantile hemangiomas: Demographic, prenatal, and perinatal characteristics[J]. J Pediatr, 2007, 150(3): 291-294.

[3] CHEN T S, EICHENFIELD L F, FRIEDLANDER S F. Infantile hemangiomas: An update on pathogenesis and therapy [J]. Pediatrics, 2013, 131(1): 99-108.

[4] MASNARI O, SCHIESTL C, ROSSLER J, et al. Stigmatization predicts psychological adjustment and quality of life in children and adolescents with a facial difference[J]. J Pediatr Psychol, 2013, 38(2): 162-172.

[5] DROLET B A, FROMMELT P C, CHAMLIN S L, et al. Initiation and use of propranolol for infantile hemangioma: Report of a consensus conference [J]. Pediatrics, 2013, 131(1): 128-140.

[6] LIAO Y, WANG J, JAEHNIG E J, et al. Web-Gestalt 2019: Gene set analysis toolkit with revamped UIs and APIs[J]. Nucleic Acids Res, 2019, 47(W1): W199-W205.

[7] SZKLARCZYK D, GABLE A L, NASTOU K C, et al. The STRING database in 2021: Customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets[J]. Nucleic Acids Res, 2021, 49(D1): D605-D612.

[8] SHANNON P, MARKIEL A, OZIER O, et al. Cytoscape: A software environment for inte-

grated models of biomolecular interaction networks[J]. Genome Res, 2003, 13(11): 2498-2504.

[9] HEBERLE H, MEIRELLES G V, DA S F, et al. InteractiVenn: A web-based tool for the analysis of sets through Venn diagrams [J]. BMC Bioinformatics, 2015, 16(1): 169.

[10] SUN X, LIU X, LU N, et al. Short-term curative effect and safety of propranolol combined with laser in the treatment of infantile hemangiomas [J]. Oncol Lett, 2018, 16(5): 6561-6565.

[11] ZACHARIOU M, MINADAKIS G, OULAS A, et al. Integrating multi-source information on a single network to detect disease-related clusters of molecular mechanisms [J]. J Proteomics, 2018, 188: 15-29.

[12] MAO X G, XUE X Y, WANG L, et al. CDH5 is specifically activated in glioblastoma stemlike cells and contributes to vasculogenic mimicry induced by hypoxia [J]. Neuro Oncol, 2013, 15(7): 865-879.

[13] LAGENDIJK A K, HOGAN B M. VE-cadherin in vascular development: A coordinator of cell signaling and tissue morphogenesis [J]. Curr Top Dev Biol, 2015, 112: 325-352.

[14] CARMELIET P, COLLEN D. Molecular basis of angiogenesis. Role of VEGF and VE-cadherin [J]. Ann N Y Acad Sci, 2000, 902: 249-264.

[15] LABELLE M, SCHNITTTLER H J, AUST D E, et al. Vascular endothelial cadherin promotes breast cancer progression via transforming growth factor beta signaling [J]. Cancer Res, 2008, 68(5): 1388-1397.

[16] FRY S A, ROBERTSON C E, SWANN R, et al. Cadherin-5: A biomarker for metastatic breast cancer with optimum efficacy in oestrogen receptor-positive breast cancers with vascular invasion [J]. Br J Cancer, 2016, 114(9): 1019-1026.

[17] SANTHOSH D, HUANG Z. A Tie2-driven BAC-TRAP transgenic line for in vivo endothelial gene profiling [J]. Genesis, 2016, 54(3): 136-145.

(下转第 210 页)

GHEGAN J C, et al. Osteocyte control of bone formation via sclerostin, a novel BMP antagonist[J]. *EMBO J*, 2003, 22(23):6267-6276.

[16] O'BRIEN C A, JIA D, PLOTKIN L I, et al. Glucocorticoids act directly on osteoblasts and osteocytes to induce their apoptosis and reduce bone formation and strength[J]. *Endocrinology*, 2004, 145(4):1835-1841.

[17] LIU W, QI M, KONERMANN A, et al. The p53/miR-17/Smurfl pathway mediates skeletal deformities in an age related model via inhibiting the function of mesenchymal stem cells [J]. *Aging*, 2015, 7(3):205-218.

[18] 马崢. 绝经后骨质疏松患病率及相关危险因素[J]. *临床研究*, 2019, 27(2):5-6.

[19] ABDI F, MOBEDI H, BAYAT F, et al. The effects of transdermal estrogen delivery on bone mineral density in postmenopausal women: A meta-analysis [J]. *Iran J Pharm Res*, 2017, 16(1):380-389.

[20] WANG X Y, ZHANG X Z, LI F, et al. MiR-128-3p accelerates cardiovascular calcification and insulin resistance through ISL1-dependent Wnt pathway in type 2 diabetes mellitus rats [J]. *J Cell Physiol*, 2019, 234(4):4997-5010.

[21] AMJADI-MOHEB F, AKHAVAN-NIAKI H. Wnt signaling pathway in osteoporosis: Epigenetic regulation, interaction with other signaling pathways, and therapeutic promises[J]. *J Cell Physiol*, 2019.

[22] YAO Y, SUN S, WANG J, et al. Canonical Wnt signaling remodels lipid metabolism in zebrafish hepatocytes following ras oncogenic insult[J]. *Cancer Res*, 2018, 78(19):5548-5560.

[23] RIANCHO J A, OLMOS J M, PINEDA B, et al. Wnt receptors, bone mass, and fractures: Gene-wide association analysis of LRP5 and LRP6 polymorphisms with replication[J]. *Eur J Endocrinol*, 2011, 164(1):123-131.

[24] LI C, HUANG Q, YANG R, et al. Targeted next generation sequencing of nine osteoporosis-related genes in the Wnt signaling pathway among Chinese postmenopausal women [J]. *Endocrine*, 2020, 68(3):669-678.

[25] STÜRZNICKEL J, ROLVIEN T, DELSMA NN A, et al. Clinical Phenotype and Relevance of LRP5 and LRP6 Variants in Patients With Early-Onset Osteoporosis (EOOP) [J]. *J Bone Miner Res*, 2021, 36(2):271-282.

[26] SEBASTIAN A, HUM N R, MURUGESH D K, et al. Wnt co-receptors Lrp5 and Lrp6 differentially mediate Wnt3a signaling in osteoblasts[J]. *PLoS One*, 2017, 12(11):e0188264.

[27] 张奇, 蔡菁, 靳勇, 等. 中老年女性血清 25 羟维生素 D 水平与骨密度、体质量指数的调查研究[J]. *中国骨质疏松杂志*, 2020, 26(3):328-332.

[28] 杨红英, 李虹, 刘璐, 等. 绝经后女性骨密度与维生素 D 和高密度脂蛋白水平相关性研究[J]. *中国骨质疏松杂志*, 2019, 25(2):193-197.

[29] XU Y, CHEN D, LIN X X, et al. The LRP6 functional mutation rs2302685 contributes to individual susceptibility to alcoholic liver injury related to the Wnt/beta-catenin-TCF1-CYP2E1 signaling pathway[J]. *Arch Toxicol*, 2019, 93(6):1679-1695.

[30] 王振敏, 罗建权, 张伟. LRP6 基因多态性与代谢综合征的相关性研究进展[J]. *中国药理学通报*, 2016, 32(5):603-607.

(收稿日期:2022-08-05 修回日期:2022-10-18)

(上接第 203 页)

[18] DEKKER N, VAN LEEUWEN A, VAN ST RIEN W, et al. Microcirculatory perfusion disturbances following cardiac surgery with cardiopulmonary bypass are associated with in vitro endothelial hyperpermeability and increased angiopoietin-2 levels [J]. *Crit Care*, 2019, 23(1):117.

[19] AI L, WANG Q, CHENG K. Key genes in the liver fibrosis process are mined based on single-cell transcriptomics[J]. *Biochem Biophys Res Commun*, 2022, 598:131-137.

(收稿日期:2022-05-06 修回日期:2022-09-10)