

·生物信息学·

# 基于生物信息学分析筛选银屑病的关键基因

任静静<sup>1</sup>, 石 瑛<sup>2</sup>, 智艳芳<sup>1</sup>, 杨 博<sup>1</sup>, 李肖甫<sup>1</sup>

(郑州大学第三附属医院输血科<sup>1</sup>, 检验科<sup>2</sup>, 河南 郑州 450052)

**摘要:**目的 利用生物信息学方法筛选银屑病的关键基因及所在信号通路,探讨其发病机制。方法 从开源性基因表达综合数据库(NCBI-GEO)中获取 3 个银屑病皮损和非皮损活检样本的相关数据集。利用 R 语言 GEO Query 和 Limma 包筛选 3 个数据集中差异基因(DEGs),通过 R 语言 clusterProfiler 包对 DEGs 进行生物学过程、信号通路分析,最后通过 Cytoscape3.6.0 软件做出蛋白-蛋白互作网络分析并筛选关键基因。结果 3 个数据集取交集共获得共计 297 个共同差异基因,其中 184 个上调 DEGs 和 113 个下调 DEGs;筛选出的 10 个关键基因,均为上调表达基因;KEGG 信号通路主要富集于 IL-17 信号通路、病毒蛋白与细胞因子/受体互作通路以及趋化因子信号通路;GOBP 主要富集在趋化因子的生物学过程中。结论 趋化因子及 IL-17 信号通路可能在银屑病的发生发展中起着重要作用,为其分子机制、诊断和生物学治疗提供指导。

**关键词:**银屑病;生物信息学;IL-17 信号通路;趋化因子

中图分类号:R758.63

文献标识码:A

DOI:10.3969/j.issn.1006-1959.2022.24.001

文章编号:1006-1959(2022)24-0001-06

## Screening of Key Genes in Psoriasis Based on Bioinformatics Analysis

REN Jing-jing<sup>1</sup>, SHI Ying<sup>2</sup>, ZHI Yan-fang<sup>1</sup>, YANG Bo<sup>1</sup>, LI Xiao-fu<sup>1</sup>

(Department of Transfusion<sup>1</sup>, Department of Clinical Laboratory<sup>2</sup>, the Third Affiliated Hospital of Zhengzhou University, Zhengzhou 450052, Henan, China)

**Abstract: Objective** To screen the key genes and signaling pathways of psoriasis by bioinformatics methods, and to explore its pathogenesis.

**Methods** The relevant data sets of 3 psoriatic lesion and non-lesion biopsy samples were obtained from the open source gene expression comprehensive database (NCBI-GEO). The R language GEO Query and Limma package were used to screen differentially expressed genes (DEGs) in three datasets. The R language clusterProfiler package was used to analyze the biological processes and signaling pathways of DEGs. Finally, the protein-protein interaction network was analyzed by Cytoscape 3.6.0 software and the key genes were screened. **Results** A total of 297 common differential genes were obtained from the intersection of the three datasets, including 184 up-regulated DEGs and 113 down-regulated DEGs; the 10 key genes screened were all up-regulated genes. The KEGG signaling pathway was mainly enriched in IL-17 signaling pathway, virus protein and cytokine/receptor interaction pathway and chemokine signaling pathway. GOBP was mainly enriched in the biological processes of chemokines.

**Conclusion** Chemokines and IL-17 signaling pathway may play an important role in the occurrence and development of psoriasis, and provide guidance for its molecular mechanism, diagnosis and biological treatment.

**Key words:** Psoriasis; Bioinformatics; IL-17 signaling pathway; Chemokine

银屑病(psoriasis)是一种常见慢性炎症性皮肤病,具有一定的复发性,主要表现为过度增殖、表皮角质形成、细胞增殖分化异常和炎性细胞浸润<sup>[1,2]</sup>。该疾病全球患病率占总人口的 2%~4%,多见于成人<sup>[3,4]</sup>。银屑病的病因和发病机制目前尚不清楚,通常具有遗传性,创伤、感染、药物、压力等环境因素以及免疫系统是其触发因素。尽管该病的现有治疗具有一定效果,但是复发频繁。本研究利用基因表达数据库(Gene Expression Omnibus, GEO)<sup>[5]</sup>和 R 语言等生物信息学分析,筛选出可能影响银屑病发生发展的关键基因,并富集其主要的生物学过程及信号通路,以期对银屑病的发病机制及治疗提供新的思路。

## 1 资料与方法

**1.1 数据获取、预处理及差异表达分析** 从开源性基因表达数据库 NCBI-GEO (<http://www.ncbi.nlm.nih.gov/geo>) 中检索到与银屑病相关的三组数据集分别为 GSE13355、GSE30999 和 GSE50790。三组基因芯片数据均为基于 GPL570 平台的 Affymetrix Human Genome U133 Plus 2.0 Array 所获得的人全基因转录本生物信息。利用 R 语言 Bioconductor 包分析原始数据,采用 GEOQuery 和 Limma 包<sup>[6]</sup>鉴定皮损组织和对应非皮损组织的差异表达基因(differentially expressed genes, DEGs)。DEGs 的筛选条件设定为  $P < 0.05$ , 标准化差异倍数 ( $|\log_2 FC| > 1$ ); 并利用 ggplot2 包绘制芯片原始的火山图。利用 Venn 在线工具 ([bioinformatics.psb.ugent.be/webtools/Venn/](http://bioinformatics.psb.ugent.be/webtools/Venn/)) 生成韦恩图筛选出在 GSE13355、GSE30999 和 GSE50790 中同时上调或下调的 DEGs 用于进一步的筛选分析,以降低假阳性率,保证结果准确性。

**1.2 差异表达基因的 GO 富集分析和 KEGG 通路分析** 利用 R 语言 clusterProfiler 包对筛选出的共同 DEGs 分别进行基因本体生物学过程(gene ontology

基金项目:河南省高等学校重点科研计划项目(编号:19A320055)

作者简介:任静静(1991.5-),女,河南商丘人,硕士,主管技师,主要从事肿瘤与生物信息学研究

通讯作者:李肖甫(1964.7-),男,河南淇县人,硕士,主任技师,教授,硕士生导师,主要从事宫颈癌临床及分子机制研究

biological process, GOBP) 富集分析和 KEGG 信号通路分析, 并将结果可视化处理。

1.3 构建差异表达基因的蛋白互作网络 (protein-protein interaction network, PPI) 及模块化分析 利用 String 数据库 (<https://string-db.org/>) 检索分子间相互作用并预测蛋白质互作关系。对银屑病芯片中 297 个 DEGs 进行 PPI 网络构建, 导入 Cytoscape3.6.0 软件<sup>[7]</sup>进行可视化基因-疾病网络分析, 利用 MCODE 插件对 DEGs 进行聚类功能模块构建, 关键模块筛选标准为 MCODE score > 5, degree cut-off = 2, node score cut-off = 0.2, Max depth = 100, k-score = 2。

1.4 获取关键基因及富集分析 利用 Cytoscape 软件中 Cytohubba 插件对 297 个 DEGs 进行关键基因分析, 并利用 R 语言 clusterProfiler 包对关键基因进行 GOBP 富集分析和 KEGG 通路富集分析。

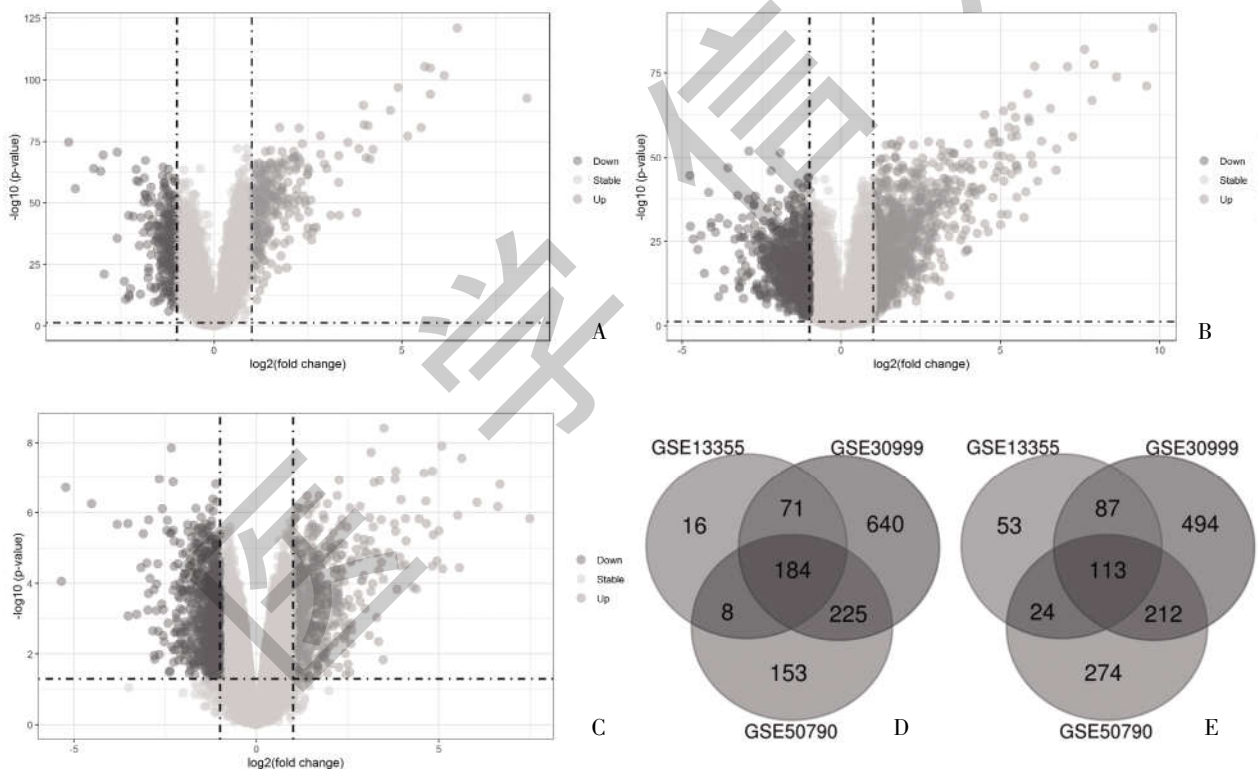
## 2 结果

### 2.1 基因芯片数据预处理及差异表达分析结果

利用 R 语言 Limma 包对 GSE13355、GSE30999 和 GSE50790 三组数据集进行分析, 以  $|\log_2 \text{FC}| > 1$  且  $P < 0.05$  为分组依据, 见图 1。数据集基本资料及筛选信息见表 1。韦恩图结果显示, 共计 297 个共同差异基因, 其中 184 个上调 DEGs 和 113 个下调 DEGs。

2.2 银屑病共同差异基因 GOBP 和 KEGG 分析 共同差异基因的本体生物学过程 GOBP 主要集中在白细胞 (粒细胞和淋巴细胞) 趋化性和迁移能力, 干扰素反应以及对病毒的防御反应, 见图 2。KEGG 通路主要富集在细胞因子/受体相互作用信号通路、病毒蛋白与细胞因子/受体相互作用信号通路及 NLRs 炎症信号通路中, 见图 3。

2.3 银屑病共同差异基因 DEGs 的蛋白互作网络 297 个共同差异基因经可视化处理后, 去除孤立节点后获得 240 个节点基因, 共 976 个相互作用关系, 见图 4。在本研究中, 上调基因在银屑病案例的网络调控中起到更为关键的作用。



注: A、B、C: GSE13355、GSE30999、GSE50790 数据集的火山图; D、E: 共同上调、下调差异表达基因的韦恩图

图 1 数据集火山图和韦恩图

表 1 数据集资料

GEO 数据集	GPL 平台	研究对象	DEGs	上调 DEGs	下调 DEGs
GSE13355	GPL570	PP(58) vs. PN(58)	556	279	277
GSE30999	GPL570	PP(85) vs. PN(85)	2026	1120	906
GSE50790	GPL570	PP(4) vs. PN(4)	1193	570	623

注: PP: 银屑病患者皮损组织 (involved skin from psoriasis); PN: 银屑病患者非皮损组织 (uninvolved skin from psoriasis)

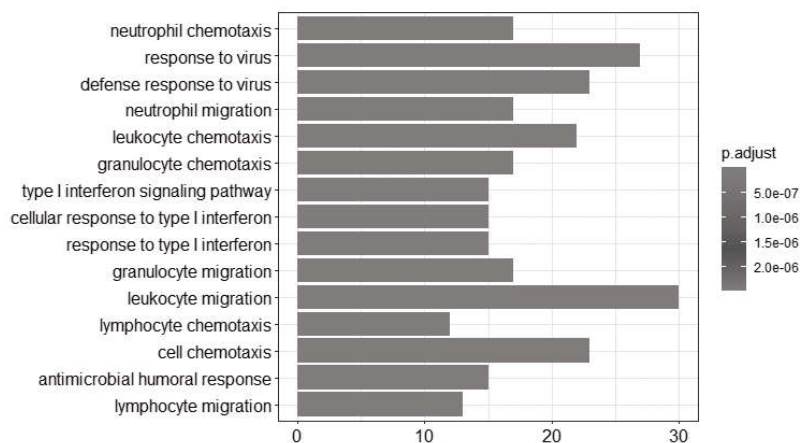


图 2 DEGs 生物学过程 GOBP 富集结果

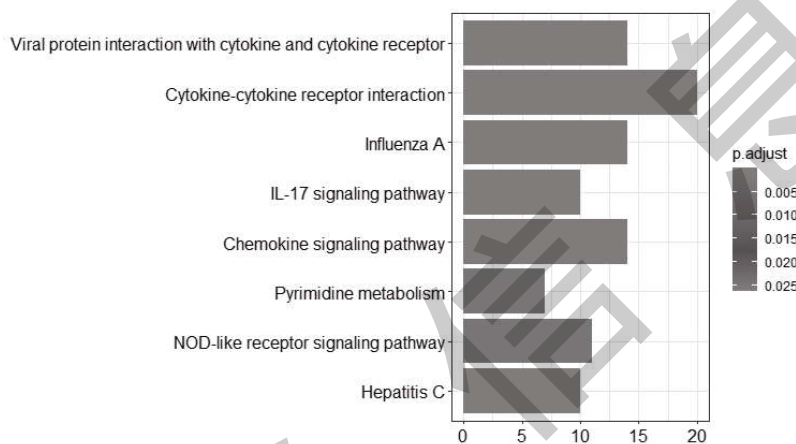
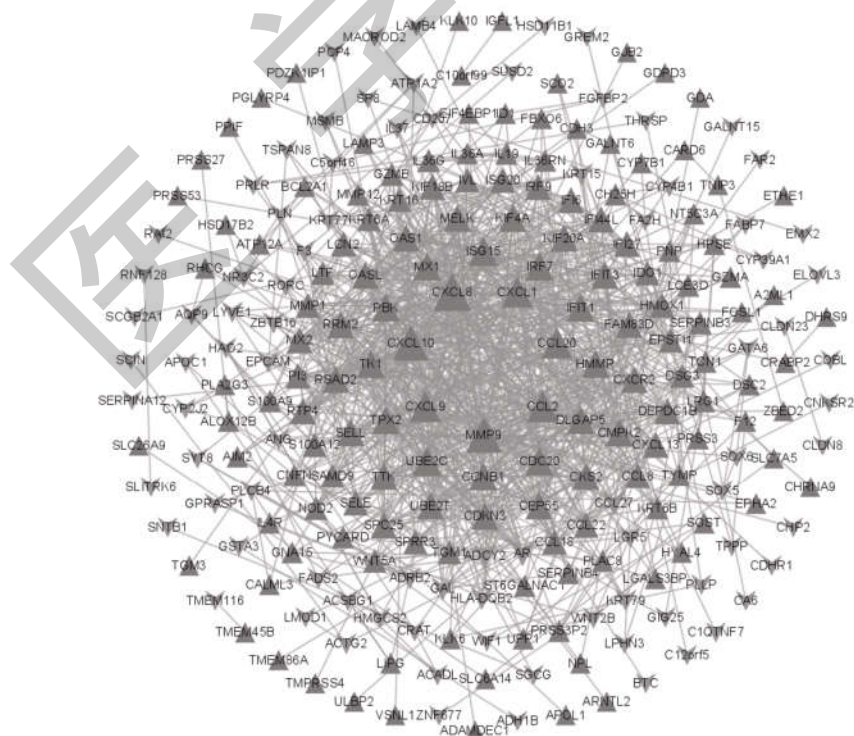


图 3 DEGs KEGG 信号通路富集结果



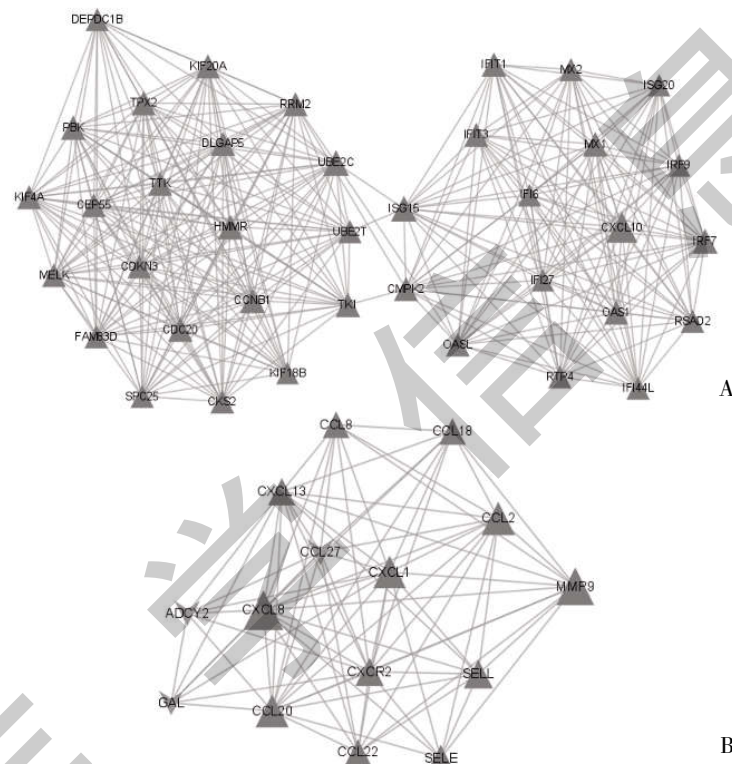
注:图中蛋白节点面积越大,度值越高,即显著性越大,在网络调控中起到的作用越大;其中▲标记节点表示上调基因,▼标记节点表示下调基因

图 4 共同差异基因 PPI 网络图

2.4 关键模块分析 利用 MCODE 插件对 PPI 进行关键模块筛选,共筛选出的 2 个最具差异的模块,其中模块一的分值为 18.432 分,包含 38 个节点基因,341 个相互作用关系,该模块均为上调基因,见图 5A;模块二的分值为 11.429 分,包含 15 个节点基因,80 个相互作用关系,见图 5B。

2.5 Hub 基因的筛选和分析 利用 CytoHubba 插件选择前 10 个节点的差异表达基因作为银屑病的 Hub 基因,见表 2,可能在银屑病的发生发展中其关键作用,均为表达上调基因。对前 10 个 Hub 基因进行 GO 富集分析见图 6,KEGG 信号通路分析见图 7。结果显示,银屑病患者中的关键基因 GOBP 主要

富集在趋化因子介导的信号通路(chemokine-mediated signaling pathway)、对趋化因子反应(response to chemokine)、细胞内趋化因子反应 (cellular response to chemokine)、中性粒细胞趋化(neutrophil chemotaxis)、中性粒细胞迁移(neutrophil migration)等生物学过程。KEGG 富集信号通路主要在 IL-17 信号通路(IL-17 signaling pathway)、病毒蛋白与细胞因子/受体相互作用信号通路(viral protein interaction with cytokine and cytokine receptor)、趋化因子信号通路(chemokine signaling pathway)、TNF 信号通路(TNF signaling pathway)、细胞因子/受体相互作用信号通路(cytokine-cytokine receptor interaction)等。



注:图示中蛋白节点面积越大,度值越高,即显著性越大,在网络调控中起到的作用越大;其中▲标记节点表示上调基因,▼标记节点表示下调基因

图 5 关键模块图

表 2 前 10 个 hub 基因及其生物学功能

基因符号	别称	中文名称	生物学功能
CXCL1	GRO-α	趋化因子生长相关癌基因-α	促进血管生成
CXCL8	IL-8	白介素-8	介于炎症反应,趋化中性粒细胞
CXCL9	MIG	单核因子	拮抗血管生成
CXCL10	IP10	诱导蛋白-10	拮抗血管生成
CCL2	MCP-1	单核细胞趋化蛋白-1	参与血管生成,趋化细胞迁移
CCL20	MIP-3α	趋化因子巨噬细胞炎症蛋白-3α	促进细胞定向迁移
MMP9	CLG4B	血清基质金属蛋白酶 9	降解基底膜和细胞外基质成分,促进细胞侵袭
UBE2C	UBH10	泛素偶联酶 E2C	参与调控纺锤体装配检验点
ISG15	GIP2	干扰素刺激基因 15	参与中性粒细胞趋化,抗病毒活性
CCNB1	CCNB	细胞周期蛋白 B1	参与调控 G <sub>2</sub> /M 转化



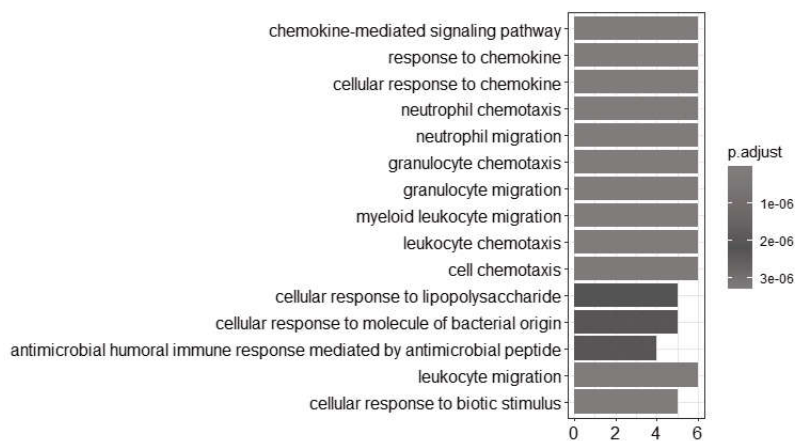


图 6 Hub 基因生物学过程 GOBP 富集结果

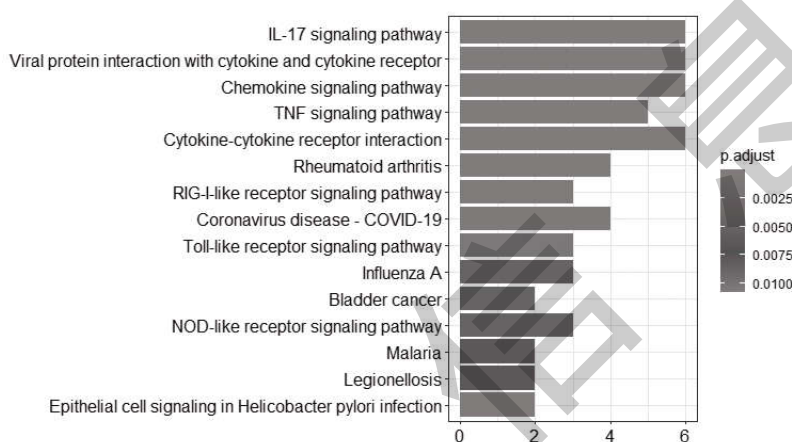


图 7 Hub 基因 KEGG 信号通路富集结果

### 3 讨论

银屑病是慢性炎症性皮肤病,主要发生于青壮年,病程较长,有复发倾向,治疗效果较差导致患者生活和身心健康受到很大影响<sup>[8,9]</sup>。本研究通过公共数据库收集 3 个数据集,利用 R 语言处理得到 297 个共同 DEGs,其中 184 个上调基因,113 个下调基因,主要富集于白细胞的趋化和迁移、干扰素反应以及病毒的防御反应等生物学过程中,同时 KEGG 富集信号通路主要体现在细胞因子/受体相关作用信号通路、病毒蛋白与细胞因子/受体相互作用信号通路以及 NLRs 炎症信号通路中。前 10 个 hub 基因中包括 6 个趋化因子。趋化因子<sup>[10-12]</sup>是小分子细胞因子,具有能趋化细胞定向移动的生物学特性,与趋化因子受体作用参与细胞的生长、发育、分化、凋亡和分布等多种生理功能,并在炎症反应、病原体感染、创伤修复及肿瘤形成和转移等病理过程中发挥作用。根据其 N 端半胱氨酸残基的位置和数目可分为以下 4 个亚族:C、CC、CXC、CX3C,其中 CC 和 CXC 作为趋化因子的主要成员。其中,CXC 趋化因子亚家族<sup>[13,14]</sup>根据其第 1 个半胱氨酸前有无谷氨酸、亮氨酸、精氨酸序列(Glu-Len-Arg)进一步

分为 ELR+和 ELR-两类,CXCL1(GRO- $\alpha$ )和 CXCL8(IL-8)是 ELR+趋化因子,能够促进血管新生;而 CXCL9(MIG)和 CXCL10(IP10)是 ELR-趋化因子,能够拮抗 ELR+趋化因子诱发的血管生成作用。CCL2(MCP-1)<sup>[15]</sup>是第一个发现的趋化因子,其参与血管生成,且趋化单核细胞进入病变部位,促进血管平滑肌细胞增殖和迁移,进而参与炎症反应。CCL-2 与活化的 T 细胞孵育后导致 Th2 细胞活化,促使 IL-4 升高,而 IFN- $\gamma$  降低。CCL20(MIP-3 $\alpha$ )与 CCR6 配体结合,促进树突状细胞的定向迁移,在器官特异性表达的黏附分子的协助下,聚集于上皮组织。在银屑病患者群体中,CCL20/CCR6 是第 1 个与记忆性 T 细胞募集相关的配体/受体对。

本研究中网络互作分析的前 10 个关键基因均在银屑病患者皮损组织中表达上调,分别为: CXCL1、CXCL8、CXCL9、CXCL10、CCL2、CCL20、MMP9、UBE2C、ISG15 和 CCB1。且其中有 6 个位于 IL-17 信号通路中,包括趋化因子 CXCL1、8、10 和 CCL2、20,以及组织重塑因子 MMP9。Th17 细胞分泌 IL-17A 和 IL-17F,IL-17 信号通路<sup>[16]</sup>的主要靶细胞是角质形成细胞、内皮细胞和固有免疫细胞,其下游趋化因

子和组织重塑因子的上调,表明 IL-17 信号通路在银屑病的发生发展过程中有着重要的作用。

趋化因子 CXCL1、8、9、10 和 CCL2、20 基因均在银屑病皮损组织中上调表达,表明血管新生和血管新生拮抗活动中在同时发生;能选择性地诱导白细胞表达整合素,增加它们的结合能力,从而促进它们与内皮细胞结合并迁移到炎症部位;树突状细胞、单核细胞经趋化因子作用后,进入上皮组织病变部位,参与炎症反应的发生发展。CCL20 具有免疫监视作用,其表达上调表明皮肤屏障功能减退以及免疫监视功能过强,且表达量与病损严重程度相关。MMP9 是一种明胶酶<sup>[17]</sup>,能够降解基底膜和细胞基质成分,其上调表达促进细胞迁移及间质血管新生,参与组织重塑。UBE2C 和 CCNB1 在细胞周期中具有重要调控作用,参与蛋白泛素化以及细胞内蛋白修饰<sup>[18]</sup>。ISG15 也称干扰素刺激基因 15<sup>[19,20]</sup>,是泛素样蛋白家族的一员,能与靶蛋白结合并 ISG 化,参与中性粒细胞趋化,具有抗病毒活性。

综上所述,本研究利用生物信息学技术对 3 个数据集进行数据分析,筛选出 297 个差异基因,利用这些基因进行 GOBP 和 KEGG 信号通路富集分析,揭示了其生物学过程和主要涉及的信号通路。利用蛋白互作网络分析筛选出 10 个关键基因,均上调表达,且其中 6 个位于 IL-17 信号通路中,此信号通路已有研究证实其与银屑病机制,这一发现可为银屑病潜在分子标志物的筛选以及进一步研究银屑病分子机制、诊断、治疗和预后提供参考。

#### 参考文献:

- [1]Armstrong AW,Read C.Pathophysiology, Clinical Presentation, and Treatment of Psoriasis: A Review[J].JAMA,2020,323(19):1945.
- [2]Griffiths CEM,Armstrong AW,Gudjonsson JE,et al.Psoriasis[J].The Lancet,2021,397(10281):1301–1315.
- [3]Parisi R,Iskandar IYK,Kontopantelis E,et al.National, regional, and worldwide epidemiology of psoriasis: systematic analysis and modelling study[J].BMJ,2020,369:m1590.
- [4]Korman NJ.Management of psoriasis as a systemic disease: what is the evidence?[J].British Journal of Dermatology,2020,182(4):840–848.
- [5]Barrett T,Wilhite SE,Ledoux P,et al.NCBI GEO: archive for functional genomics data sets—update[J].Nucleic Acids Research,2012,41(D1):D991–D995.
- [6]Ritchie ME,Phipson B,Wu D,et al.limma powers differential expression analyses for RNA—sequencing and microarray studies[J].Nucleic Acids Research,2015,43(7):e47.
- [7]Saito R,Smoot ME,Ono K,et al.A travel guide to Cytoscape plugins[J].Nature Methods,2012,9(11):1069–1076.
- [8]Tsakok T,Wilson N,Dand N,et al.Association of Serum Ustekinumab Levels With Clinical Response in Psoriasis [J].JAMA Dermatology,2019,155(11):1235–1243.
- [9]Sommer R,Augustin M,Hilbring C,et al.Significance of chronic pruritus for intrapersonal burden and interpersonal experiences of stigmatization and sexuality in patients with psoriasis [J].Journal of the European Academy of Dermatology and Venereology,2021,35(7):1553–1561.
- [10]Mahalingam S,Clark K,Matthaei KI,et al.Antiviral potential of chemokines[J].Bioessays,2001,23(5):428–435.
- [11]Coelho AL,Hogaboam CM,Kunkel SL.Chemokines provide the sustained inflammatory bridge between innate and acquired immunity [J].Cytokine & Growth Factor Reviews,2005,16(6):553–560.
- [12]Yoshie O,Imai T,Nomiyama H.Chemokines in Immunity[J].Advances in Immunology,2001,78:57–110.
- [13]Liu K,Wu L,Yuan S,et al.Structural basis of CXC chemokine receptor 2 activation and signalling [J].Nature,2020,585(7823):135–140.
- [14]Ngo T,Stephens BS,Gustavsson M,et al.Crosslinking—guided geometry of a complete CXC receptor—chemokine complex and the basis of chemokine subfamily selectivity [J].PLoS Biology,2020,18(4):e3000656.
- [15]Ji W,Shi H,Feng T,et al.Majie Cataplasma Promotes Th1 Response to Fight against Asthmatic Th2 Inflammation through NKs [J].Evidence Based Complementary and Alternative Medicine,2022,2022:6745420.
- [16]Ghoreschi K,Balato A,Enerbäck C,et al.Therapeutics targeting the IL-23 and IL-17 pathway in psoriasis [J].The Lancet,2021,397(10275):754–766.
- [17]Lund IK,Nielsen BS,Almholt K,et al.Concomitant lack of MMP9 and uPA disturbs physiological tissue remodeling[J].Developmental Biology,2011,358(1):56–67.
- [18]Wang Y,Wang J,Tang Q,et al.Identification of UBE2C as hub gene in driving prostate cancer by integrated bioinformatics analysis[J].PLoS One,2021,16(2):e0247827.
- [19]Pitha—Rowe IF,Pitha PM.Viral defense, carcinogenesis and ISG15: Novel roles for an old ISG[J].Cytokine & Growth Factor Reviews,2007,18(5–6):409–417.
- [20]Andersen JB,Hassel BA.The interferon regulated ubiquitin—like protein, ISG15, in tumorigenesis: Friend or foe?[J].Cytokine & Growth Factor Reviews,2006,17(6):411–421.

收稿日期:2022-07-18;修回日期:2022-08-25

编辑/肖婷婷