

基于 GEO 数据库对儿童哮喘急性发作靶基因的生物信息学分析

张迪¹, 姜凡¹, 刘璐佳², 景伟超², 王有鹏^{2*}

(1. 黑龙江中医药大学 第二临床医学院, 中国黑龙江 哈尔滨 150040; 2. 黑龙江中医药大学附属第二医院, 中国黑龙江 哈尔滨 150001)

摘要: 为了通过生物信息学技术筛选并分析关键基因以探究儿童哮喘急性发作的病理机制, 从 GEO 数据库下载儿童哮喘急性发作相关基因芯片数据集(GSE103166), 用 R 软件包工具筛选出哮喘急性发作患儿与健康儿童的差异表达基因(differentially expressed genes, DEGs), 用 DAVID (version 6.8)在线工具获取 DEGs 列表的 GO 和 KEGG 注释结果, 同时通过 STRING (version 11.0)获取其蛋白质互作数据, 应用 Cytoscape 及其插件 MCODE 构建蛋白质互作网络、鉴定关键基因。结果显示, 共筛选出 78 个 DEGs, 其中上调基因共 49 个, 下调基因共 29 个。这些 DEGs 主要富集于核质转运调控、免疫系统过程调节等生物过程; 核小体、DNA 包装复合物等细胞组分; Rho GTP 酶结合、离子型谷氨酸受体结合等分子功能。KEGG 通路分析结果提示其主要富集在系统性红斑狼疮和哮喘信号通路。在这些 DEGs 中, 筛选得到一个基因模块, GO 分析显示该模块主要富集在 γ 干扰素介导的信号通路、通过 MHC II 类分子进行抗原加工和外源性抗原肽的呈递、通过 MHC II 类分子进行抗原加工和抗原肽的呈递等生物过程; MHC II 类蛋白复合物、内质网膜腔侧、内质网膜腔侧组成部分等细胞组分; MHC II 类受体活性、抗原肽结合、跨膜信号受体活性等分子功能。KEGG 通路分析提示该模块主要富集在哮喘、移植物抗宿主病、同种异体排斥反应等通路。该基因模块涵盖 5 个关键基因, 分别为 *HLA-DPB1*、*HLA-DQB1*、*HLA-DQB2*、*MT2A* 和 *KIF11*。上述分析结果表明, 文中筛选的 DEGs 和关键基因有助于加深对儿童哮喘急性发作病理机制的理解, 其中 *HLA-DPB1*、*HLA-DQB1*、*HLA-DQB2* 等关键基因已得到相关研究的验证, 未经证实的 *MT2A*、*KIF11* 关键基因, 可能是儿童哮喘急性发作研究的新靶点。

关键词: 儿童哮喘; 生物信息学; 差异表达基因(DEGs); 关键基因; 分子生物学机制

中图分类号: Q811.4, R725.6

文献标识码: A

文章编号: 1007-7847(2021)01-0070-10

Bioinformatic Analysis of Target Genes in Children with Acute Asthma Attack Based on GEO Database

ZHANG Di¹, JIANG Fan¹, LIU Lu-jia², JING Wei-chao², WANG You-peng^{2*}

(1. The Second Clinical Medicine School, Heilongjiang University of Chinese Medicine, Harbin 150040, Heilongjiang, China;
2. The Second Affiliated Hospital, Heilongjiang University of Chinese Medicine, Harbin 150001, Heilongjiang, China)

Abstract: To screen and analyze key genes by bioinformatics techniques to investigate the pathological mechanisms of acute asthma attack in children, the gene expression dataset (GSE103166) associated with acute asthma attack in children was downloaded from the GEO database. The differentially expressed genes (DEGs) between children with acute asthma attack and healthy children were screened using the R package tool, and GO analysis of DEG lists and annotation results from the KEGG were obtained using DAVID (version 6.8), which were analyzed by STRING (version 11.0). Cytoscape and its plug-in MCODE were used to construct

收稿日期: 2020-08-25; 修回日期: 2020-10-22

基金项目: 国家自然科学基金资助项目(81874485); 黑龙江省自然科学基金资助项目(LH2020H088); 黑龙江省中医药管理局项目(ZHY2020-149); 黑龙江中医药大学科研基金项目(2018bs01); 黑龙江省名中医专家传承工作室项目

作者简介: 张迪(1990—), 女, 黑龙江齐齐哈尔人, 博士研究生; 张迪和姜凡对本文的贡献相同, 为本文共同第一作者; *通信作者: 王有鹏(1964—), 男, 黑龙江哈尔滨人, 博士, 教授, 主要从事中医药防治小儿呼吸系统疾病研究, Tel: 0451-87093337, E-mail: wangyp8899@163.com。

protein interaction networks and identify key genes. The results showed that a total of 78 DEGs were identified, including 49 up-regulated and 29 down-regulated genes, which were mainly enriched in biological processes such as nucleocytoplasmic transport and immune system process regulation, cellular components such as nuclear nucleosome and DNA packaging complex, and molecular functions such as Rho GTPase binding and ionotropic glutamate receptor binding. KEGG pathway was enriched in systemic lupus erythematosus and asthma signaling pathway. Among these DEGs, a gene module was screened out, which was mainly enriched in biological processes such as interferon-gamma-mediated signaling pathway, antigen processing and presentation of exogenous peptide antigen via MHC class II, antigen processing and presentation of peptide antigen via MHC class II, cellular components such as MHC class II protein complex, lumenal side of endoplasmic reticulum membrane, integral component of lumenal side of endoplasmic reticulum membrane, and molecular functions such as MHC class II receptor activity, peptide antigen binding, transmembrane signaling receptor activity. The KEGG pathway was mainly enriched in asthma, graft-versus-host disease, and allograft rejection pathways. The gene module contains five key genes, including *HLA-DPB1*, *HLA-DQB1*, *HLA-DQB2*, *MT2A* and *KIF11*. The above results showed that DEGs and key genes contribute to a better understanding of the pathological mechanisms of children with acute asthma attack. As the correlation of the three key genes *HLA-DPB1*, *HLA-DQB1*, and *HLA-DQB2* with the disease has been validated by other studies, the two unconfirmed key genes *MT2A* and *KIF11* may be used as new targets for studying acute asthma attack in children.

Key words: childhood asthma; bioinformatics; differentially expressed genes (DEGs); key genes; molecular biological mechanisms

(*Life Science Research*, 2021, 25(1): 070-079)

哮喘是儿童中常见的非传染性慢性疾病之一,以反复发作的气道阻塞和支气管痉挛为主要特征,急性发作阶段诸症状对儿童的身心影响十分严重。儿童哮喘与遗传因素和过敏性因素关系密切。随着全基因组关联研究(genome-wide association study, GWAS)和后续验证研究的不断深入,一些与儿童哮喘相关的遗传风险变体已被确定,如首项关于哮喘的 GWAS 分析发现,17q12 号染色体上的 *ORMDL3/GSDMA* 位点与儿童哮喘发病密切相关^[1]。但 GWAS 存在一定的局限性,如 Weiss 等^[2]的分析结果仅解释了哮喘风险变异的适度比例。虽然也有 DNA 甲基化与儿童哮喘的相关报道^[3],但差异表达基因(differentially expressed genes, DEGs)与儿童哮喘急性发作之间的关系尚不明确。因此,探索哮喘急性发作患儿与健康儿童之间的 DEGs,可为更好地理解易感性个体为何以及如何发展过敏性疾病提供基础,对探寻疾病的治疗新靶标具有重要意义。

本研究拟利用 R 软件包及生物信息学分析工具对哮喘急性发作患儿与健康儿童的基因表达谱数据进行分析,筛选出与儿童哮喘急性发作相关的 DEGs,探讨儿童哮喘急性发作的关键基因

及可能的病理机制,以期儿童哮喘急性发作的预防及诊疗提供新视角和新思路。

1 材料与方法

1.1 材料

本研究拟对近一年哮喘急性发作儿童的基因表达数据进行分析,故数据集筛选流程如下:1)以“asthma”和“children”为关键词在 GEO (Gene Expression Omnibus, <https://www.ncbi.nlm.nih.gov/geo/>) 数据库中进行筛选;2)限定研究类型为“expression profiling by array”,物种为“*Homo sapiens*”;3)设置样本数量>30,进一步查看筛选后的数据集。对于其他公开的儿童哮喘急性发作组患儿样本和健康儿童样本数据,由于样本类别有较大差别,为避免数据集合并后产生较大异质性,所以最终选取儿童哮喘或喘息急性发作相关的 GSE103166 基因表达谱数据集(此数据集样本为鼻拭子)。该数据集由 Anthony Bosco 提交,以 Affymetrix Human Gene 2.1 ST Array 作为研究平台,通过全转录本微阵列芯片将收集的鼻拭子样本进行基因表达分析,其中包括 56 例儿童哮喘急性发作组患儿样本(均来源于三级儿童医院急诊就诊的哮喘伴

有急性喘息的患儿)和 31 例对照组健康儿童样本(来源于入选病例的同胞或社区中随机选择的儿童)^[4]。

1.2 差异基因筛选

使用 GSE103166 数据集经 \log_2 变换后的基因表达矩阵文件进行数据分析,使用 GPL23961 平台文件进行探针注释,数据预处理和分析主要使用 R (version 4.0.2)和 limma (3.44)包^[5],应用线性模型进行差异表达评估和实验设计分析^[6]。Limma 包用于鉴定哮喘急性发作儿童和健康儿童间的 DEGs。以 $|\log_2FC| \geq 0.5$ (FC: fold change)以及校正 P 值 < 0.05 作为 DEGs 的筛选标准,其中采用 Benjamini-Hochberg 方法^[7]对 P 值进行多重比较校正,以控制假阳性。使用 ggplot2^[8]和 TBtools^[9]分别绘制 DEGs 的火山图和热图。

1.3 差异基因富集分析

基因本体论(Gene Ontology, GO)分析是指按照生物过程(biological process, BP)、细胞组分(cellular component, CC)、分子功能(molecular function, MF)对基因本体进行注释和分类。京都基因和基因组数据库(Kyoto Encyclopedia of Genes and Genomes, KEGG)分析是指通过对 DEGs 进行通路富集分析,筛选相关的生物学通路。文中运用 DAVID (<https://david.ncifcrf.gov/>, version 6.8)在线注释工具,以 $P < 0.05$ 为筛选条件,获取 DEGs 列表的 GO 和 KEGG 注释结果^[10]。

1.4 蛋白质互作网络构建和模块分析

STRING (version 11.0)是一个集已知和预测蛋白质-蛋白质互作(protein-protein interaction, PPI)功能于一体的生物网络资源数据库,数据信息主要来源于实验、计算和公共文本分析,目前涵盖来自 5 090 种生物体的 2 458 万个蛋白质^[11]。在 ST-

RING 检索工具中上传获取的 DEGs 列表,设置综合交互得分 0.400 为显著性阈值。下载蛋白质互作数据,使用 Cytoscape 的无向网络方法构建 PPI 网络。Cytoscape 软件中的 MCODE 插件是一种著名的自动化网络模块分析方法,可用于在大型 PPI 网络中寻找高度相互连接的子模块^[12]。文中采用 MCODE 对 PPI 网络中的子模块进行分析,除 $K\text{-core}=3$ 外,其余参数均为默认设置。

2 结果

2.1 差异表达基因

根据筛选条件,共获得儿童哮喘急性发作组和对照组之间的 DEGs 78 个,其中上调基因共 49 个,下调基因共 29 个(表 1,图 1)。图 1 标注出了校正 P 值排名前 10 的 DEGs,图 2 展示了 $|\log_2FC|$ 排名前 20 的上调基因和下调基因。

2.2 差异基因的 GO 和 KEGG 富集分析

对上调和下调基因开展的功能注释与通路分析如表 2、图 3 所示。在 BP 组中,上调基因主要富集于核质转运调控、胞内转运调控、胞内转运正调控;下调基因主要富集在刺激反应正向调控、免疫系统过程调节、抗原加工和外源肽抗原的呈递等过程。在 CC 组中,上调基因富集于核小体和 DNA 包装复合物等细胞组分;下调基因富集于网格蛋白包被的内吞囊泡膜、内吞囊泡和囊泡膜等细胞组分。在 MF 组中,上调基因主要富集于 Rho GTP 酶结合、离子型谷氨酸受体结合和蛋白质二聚活性等分子功能;下调基因主要富集于 MHC II 类受体活性、分子换能器活性和受体活性等分子功能。在 KEGG 通路富集分析中,上调基因富集于系统性红斑狼疮信号通路,下调基因富集于哮喘通路。

表 1 儿童哮喘急性发作组和对照组之间的 DEGs

Table 1 DEGs between children with acute asthma attack and the control group

DEGs	Gene name
Up-regulated	GMNC, CD177, SLC26A2, CD163, IFI6, IL18R1, MT2A , VTRNA2-1, SCNN1G, FKBP5, HIST1H2BH, CTTNBP2, SNORD18C, SLC7A1, GUCY1B3, PXDN, PLA2G4A, MIR614, RPL23A, ARG2, CIT, TMTG1, TP53I3, DDX10, HIST1H2BD, B3GNT6, TUBD1, NETO2, ZNF608, ARID5B, ANOS1, TPRG1-AS1, HIST1H2BK, DNAJC12, ZBTB38, ECT2, C9orf84, KIF11 , ACSL3, PTPN14, HDC, PDE4DIP, LOC344887, WASF1, NEDD4, STOM, TMEM184C, LOC101927668, ZNF562
Down-regulated	DEFB4A, GSTA2, FCER1A, LOC643802, FAM101B, MPEG1, LINC01032, LOC254896, BHLHE41, AMPD2, KLRB1, P2RY10, TGM3, DAPK2, HLA-DQB1 , NOV, CAT, CTNNA1, RPS6KA5, OGFRL1, LOC728392, HVCN1, HLA-DPB1 , DGKG, HIC1, CD207, HLA-DQB2 , DPEP3, LRRRC43

注:上调基因按照差异倍数从大到小排序,下调基因从小到大排序。

Notes: Up-regulated genes are sorted by the fold change from largest to smallest, and down-regulated genes are sorted from smallest to largest.

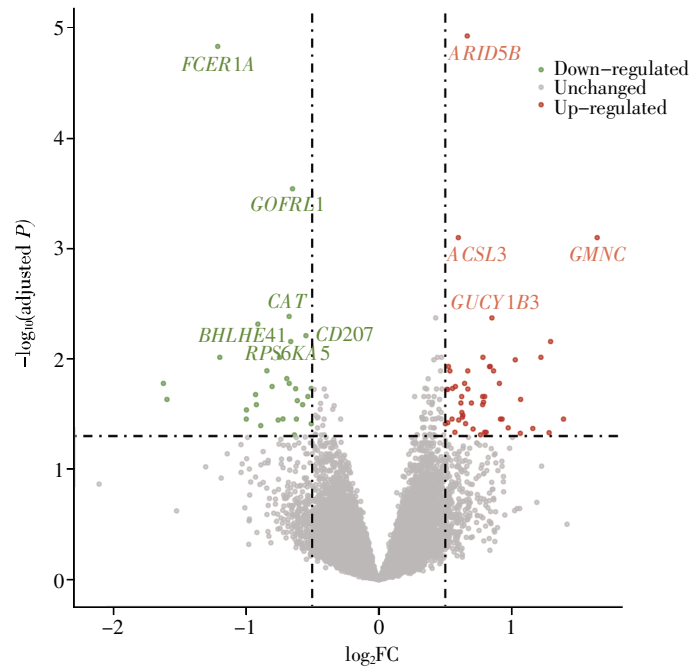


图 1 儿童哮喘急性发作组与对照组 DEGs 的火山图

横向虚线表示校正 $P=0.05$ 的阈值, 纵向虚线分别表示 $\log_2FC=-0.5$ 和 $\log_2FC=0.5$ 的阈值。图中标注的基因为校正 P 值排名前 10 的 DEGs。其中, *FCER1A* 可在变应性疾病中起到核心作用, 且与儿童哮喘具有明显的相关性^[13]; *OGFRL1* 是 *OGFR* 的重要旁系同源物, 后者与核因子 κB (nuclear factor- κB , NF- κB) 等信号通路关系密切; *CAT* 主要与氧化应激过程有关^[14]; *CD207* 主要与先天免疫系统途径相关; *BHLHE41* 已被证明是 T 细胞活化的调控器^[15]; *RPS6KA5* 可在炎症基因的调节中发挥功能^[16]; *ARID5B* 是儿童急性淋巴细胞白血病易感性和治疗结果中的重要决定因素^[17]; *ACSL3* 在脂质的生物合成和脂肪酸降解中起关键作用^[18]; *GMNC* 促进多细胞生物体中染色体 DNA 复制的启动; *GUCY1B3* 可抑制信号传导的多个亚基, 且与氧化应激关系较密切^[19]。

Fig.1 Volcano plot of DEGs in children with acute asthma attack compared with the control group

The horizontal dashed line indicates the threshold for correction $P=0.05$, the vertical dashed line represents the threshold of $\log_2FC=-0.5$ and the threshold of $\log_2FC=0.5$. The genes labeled in the figure are the top 10 DEGs in terms of corrected P value. Among them, *FCER1A* can play a central role in stress disorders and is significantly associated with childhood asthma^[13]; *OGFRL1* is an important paralog of *OGFR*, which is closely related to signaling pathways such as nuclear factor- κB (NF- κB); *CAT* is primarily associated with oxidative stress processes^[14]; *CD207* is primarily associated with innate immune system pathways; and *BHLHE41* has been shown to be a regulator of T-cell activation^[15]. *RPS6KA5* can function in the regulation of inflammatory genes^[16]; *ARID5B* is an important determinant in the susceptibility and therapeutic outcome of childhood acute lymphoblastic leukemia^[17]; *ACSL3* plays a key role in lipid biosynthesis and fatty acid degradation^[18]; *GMNC* promotes the initiation of chromosomal DNA replication in multicellular organisms; *GUCY1B3* inhibits multiple subunits of signaling, and is closely related to oxidative stress^[19].

2.3 PPI 网络和基因模块分析

将 78 个 DEGs 上传至 STRING 数据库, 最终获得由 31 个节点和 33 条边参与构建的 PPI 网络 (图 4)。使用 MCODE 插件从中鉴定出的重要基因模块由 5 个节点和 9 条边构成 (图 5), 提示我们从 DEGs 中共鉴定出 5 个关键基因: *HLA-DPB1*、*HLA-DQB1*、*HLA-DQB2*、*MT2A* 和 *KIF11*。对该模块基因进行 GO 功能和 KEGG 通路富集分析发现, 在 BP 组中, 其主要富集在 γ 干扰素介导的信号通路、通过 MHC II 类分子进行抗原加工和外源性抗原肽的呈递、通过 MHC II 类分子进行抗原加工和抗原肽的呈递; 在 CC 组中, 其主要富集在

MHC II 类蛋白复合物、内质网膜腔侧、内质网膜腔侧组成部分; 在 MF 组中, 其主要富集在 MHC II 类受体活性、抗原肽结合、跨膜信号受体活性; 在 KEGG 通路富集中, 其主要与哮喘、移植物抗宿主病、同种异体排斥反应等通路相关 (表 3)。

3 讨论

3.1 哮喘急性发作患儿和健康儿童之间存在 DEGs

哮喘是一种复杂的遗传疾病, 几十年来一直是基因组研究的目标^[20]。儿童哮喘的表型异质性显著阻碍了我们对病原和决定因素的理解。随着研究技术的革新, 我们已在导致哮喘早期发病

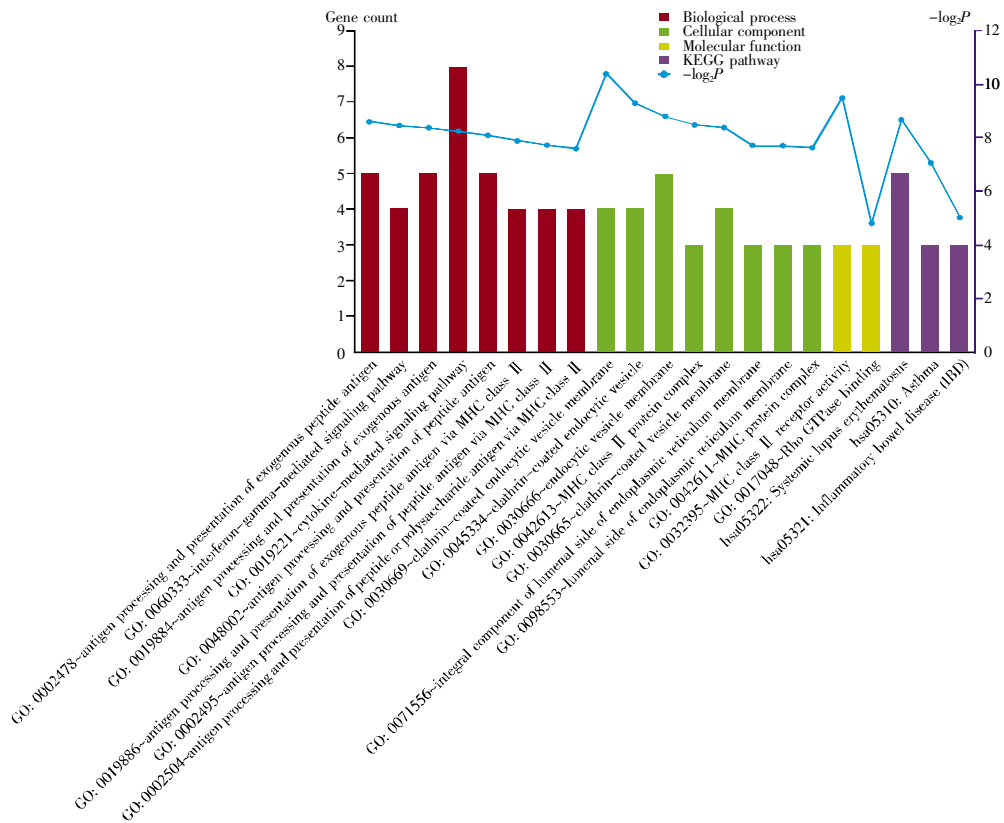


图 3 DEGs 的 GO 和 KEGG 富集分析

所有富集到的具体条目均展示在相应的 x 轴上, 每个条目所对应的 P 值由蓝色折线表示。

Fig.3 GO and KEGG enrichment analyses of DEGs

The x-axis represents the enriched terms, and the corresponding P value of each entry is represented by a blue broken line.

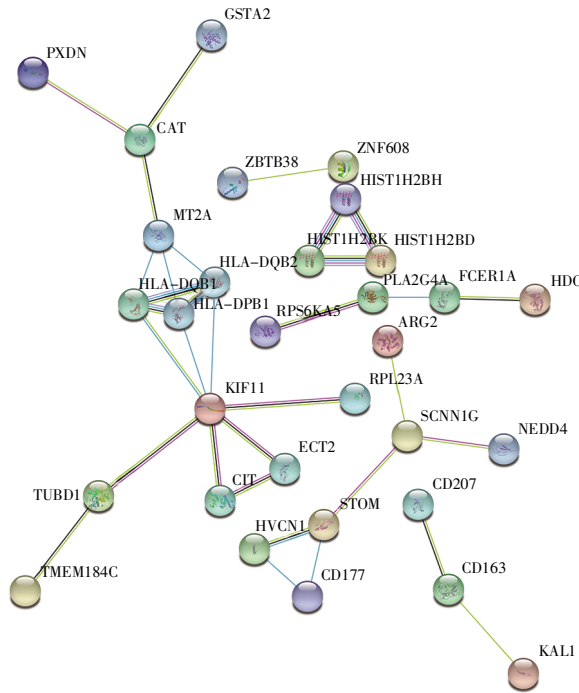


图 4 DEGs 的 PPI 网络

每个节点代表各基因对应的蛋白质, 连接各节点的边代表蛋白质之间的互动关系, 节点连边越多代表其互动关系越密切。

Fig.4 PPI network of DEGs

Each node represents the protein corresponding to each gene, and the edge connecting each node represents the interaction between proteins. The more edges the nodes are connected with, the closer the interactions between them are.

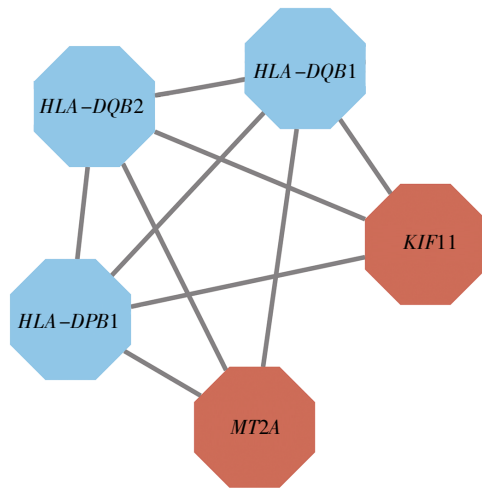


图5 PPI网络中最显著的基因模块

橙色是上调基因,蓝色是下调基因。

Fig.5 The most significant gene module from the PPI network

Orange represents up-regulated genes, and blue represents down-regulated genes.

环境和遗传因素方面取得了重要进展。目前,通过GWAS分析人们已鉴定出超过40个与哮喘相关的基因变异^[21],例如:通过 *ORMDL3* 和 *GSDMB* 的差异调节,染色体17q21上的单核苷酸多态性(single nucleotide polymorphism, SNP)可导致哮喘^[22]; *IL1RL1* 和 *IL18R1* 相关的2号染色体上的基因座与哮喘有关,9号染色体上 *IL33* 侧翼的SNP也与哮喘有关联^[23]。此外,Gu等^[24]研究发现,在中国人群中 *FCRL3* 和 *FCRL5* 基因中新鉴定的SNP与哮喘合并过敏性鼻炎的风险相关,如 *FCRL5* 中的

rs6692977 CT基因型和T等位片段显著升高,CC基因型和C等位片段显著降低;*FCRL3* 中的rs7528684 A等位片段和rs10489678 G等位片段升高,提示遗传变异可能在哮喘患儿表型的发展中起作用。Guo等^[25]针对102名哮喘儿童和80名健康儿童的研究结果显示,*LT α* 和 *NQO1* 基因的多态性与儿童哮喘有关。

哮喘是一个动态变化的过程,哮喘急性发作时机体处于高敏反应状态,患儿的痛苦程度尤为明显,因此知晓该时期基因变化模式是探寻缓解患儿疾患手段的关键。但目前针对儿童哮喘急性发作方面的研究十分有限,故探索哮喘急性发作患儿和健康儿童之间存在的DEGs,分析这些DEGs的功能信息,对深入了解该病的作用因素,探寻靶向性干预方法十分必要。

本研究基于GEO数据库对儿童哮喘急性发作基因表达谱数据集GSE103166进行了分析,确定了哮喘急性发作患儿与健康儿童之间存在的78个DEGs,其中上调基因49个,下调基因29个。这些基因可能与儿童哮喘急性发作有关,为了更好地理解这些DEGs的重要性及功能作用,我们对所得DEGs展开了GO和KEGG富集分析,并做了进一步鉴定,最终获取一个涵盖5个关键基因的重要基因模块,这些关键基因可富集到 γ 干扰素介导的信号通路、哮喘通路等。对关键基因分子生物学机制的靶向分析,有利于我们进一步明确这些关键基因与儿童哮喘急性发作之间是否存在某些已知关联性,或者是否具有某些未被证

表3 基因模块的GO功能和KEGG通路富集分析

Table 3 Functional and pathway enrichment analyses of gene module

Category	Term	Count	P value	Genes
BP	GO: 0060333~interferon-gamma-mediated signaling pathway	4	4.26E-07	<i>HLA-DQB1</i> , <i>HLA-DQB2</i> , <i>MT2A</i> , <i>HLA-DPB1</i>
BP	GO: 0019886~antigen processing and presentation of exogenous peptide antigen via MHC class II	4	6.50E-07	<i>HLA-DQB1</i> , <i>HLA-DQB2</i> , <i>KIF11</i> , <i>HLA-DPB1</i>
BP	GO: 0002495~antigen processing and presentation of peptide antigen via MHC class II	4	7.40E-07	<i>HLA-DQB1</i> , <i>HLA-DQB2</i> , <i>KIF11</i> , <i>HLA-DPB1</i>
CC	GO: 0042613~MHC class II protein complex	3	1.31E-05	<i>HLA-DQB1</i> , <i>HLA-DQB2</i> , <i>HLA-DPB1</i>
CC	GO: 0098553~luminal side of endoplasmic reticulum membrane	3	2.30E-05	<i>HLA-DQB1</i> , <i>HLA-DQB2</i> , <i>HLA-DPB1</i>
CC	GO: 0071556~integral component of luminal side of endoplasmic reticulum membrane	3	2.30E-05	<i>HLA-DQB1</i> , <i>HLA-DQB2</i> , <i>HLA-DPB1</i>
MF	GO: 0032395~MHC class II receptor activity	3	5.25E-06	<i>HLA-DQB1</i> , <i>HLA-DQB2</i> , <i>HLA-DPB1</i>
MF	GO: 0042605~peptide antigen binding	2	0.007 21	<i>HLA-DQB1</i> , <i>HLA-DPB1</i>
MF	GO: 0004888~transmembrane signaling receptor activity	3	0.038 20	<i>HLA-DQB1</i> , <i>HLA-DQB2</i> , <i>HLA-DPB1</i>
KEGG	hsa05310: Asthma	2	0.008 70	<i>HLA-DQB1</i> , <i>HLA-DPB1</i>
KEGG	hsa05332: Graft-versus-host disease	2	0.009 57	<i>HLA-DQB1</i> , <i>HLA-DPB1</i>
KEGG	hsa05330: Allograft rejection	2	0.010 72	<i>HLA-DQB1</i> , <i>HLA-DPB1</i>

实的潜在研究价值。

3.2 儿童哮喘急性发作关键基因的分子生物学机制

成人哮喘与儿童哮喘在遗传关联方面有所不同。有研究认为,在儿童中评估的遗传关联可以说是与哮喘诊断的因果关联,在成人中评估的遗传关联则可以说是哮喘的持续性,而引起儿童哮喘的基因可能与引起哮喘持续性的基因不同^[26]。一项多国研究的分层分析发现,17q 风险 SNP rs7216389-T 的关联仅限于哮喘早期发作病例,特别是在幼儿期(0~5 岁)和青少年期(14~17 岁),而在成人病例中则没有关联^[27]。其他证据表明,哮喘发作的年龄可能受遗传因素影响,早期发作更有可能暗示该疾病的遗传原因^[21]。在本研究中,通过筛选最终得到以下 5 个与儿童哮喘急性发作密切相关的关键基因,即 *HLA-DPB1*、*HLA-DQB1*、*HLA-DQB2*、*MT2A* 和 *KIF11*。

3.2.1 关键基因 *HLA-DPB1*、*HLA-DQB1* 及 *HLA-DQB2*

人类白细胞抗原(human leucocyte antigen, HLA)在人类免疫性疾病、移植、宿主抵抗感染以及所有已知的哮喘危险因素中起着至关重要的作用^[28]。HLA II 类分子(含 DP、DQ 等亚区)可呈递外源性抗原肽,其通过与 CD4⁺ T 辅助细胞结合^[29],对免疫应答起到作用。相关研究已证实,一些 HLA II 类等位基因可能导致小儿哮喘的易感性或保护性^[30]。

HLA-DPB1 是 HLA II 类分子 DP 区的基因。针对亚洲人群的 GWAS 发现,*HLA-DP* 是小儿哮喘的易感基因,其中,*HLA-DPB1*0901* 与小儿哮喘有关^[31]。Caraballo 等^[32]报道,在混血儿人群(欧洲和非洲血统的混血人群)中,*HLA-DPB1*0401* 在过敏性哮喘患者中显著降低。

HLA-DQB1 和 *HLA-DQB2* 是 HLA II 类分子 DQ 区的基因。关于 *HLA-DQB1* 与儿童哮喘的研究,目前已有一些报道。针对中国人群的一项研究表明,*HLA-DQB1*0201* 与哮喘呈正相关,*HLA-DQB1*0301* 等位基因与哮喘呈负相关^[33]。Mishra 等^[30]针对 103 名支气管哮喘儿童和 152 名健康人的研究结果显示,*HLA-DQB1*03* 和 *HLA-DQB1*02:01* 等位基因与哮喘有显著的正相关性。Movahedi 等^[34]研究表明,在哮喘患儿中,*HLA-DQB1*0603* 和 *0604* 等位基因的水平显著高于正常对照组儿童,而 *HLA-DQB1*0501* 和 *0602* 的水平则明显低于

正常对照组儿童。需要指出的是,当前针对 *HLA-DQB2* 与儿童哮喘的研究尚未查询到相关报道,这可能与 *HLA-DQB1* 和 *HLA-DQB2* 互为彼此的重要旁系同源物,大家择一研究有关。

由于 HLA II 抗原在免疫应答与调节方面起重要作用,故 *HLA-DPB1*、*HLA-DQB1* 以及 *HLA-DQB2* 与免疫应答密切相关,推断其可能是通过免疫调节对儿童哮喘急性发作产生影响。

3.2.2 关键基因 *MT2A* 与 *KIF11*

金属硫蛋白 2A (metallothionein 2A, MT2A)是金属硫蛋白家族的成员。该家族蛋白质可起抗氧化剂的作用,并在重金属的解毒中起作用。迄今为止,对于 *MT2A* 的研究主要着眼于肝癌,尚未有该基因与儿童哮喘或成人哮喘相关的研究报道,但经过文献梳理,我们发现 *MT2A* 或许可以通过以下途径对儿童哮喘起到调节作用。1) 氧化应激反应。相关研究报道,髓过氧化物酶(myeloperoxidase, MPO)相关的哮喘过敏和总抗氧化能力(total antioxidant capacity, TAC)相关的肺功能参数与儿童哮喘的发生风险密切相关^[35];空气污染中常见的细颗粒物经吸入后会深层渗透到呼吸道中,直接诱发炎症和氧化应激,这可能会进一步加重哮喘的发展和进程^[36]。根据 *MT2A* 基因的抗氧化作用,我们认为它或许可通过氧化应激反应对儿童哮喘急性发作产生作用;2) 免疫反应。有研究表明,抗氧化基因(*MT1* 和 *MT2*)的药理或遗传诱导可影响白介素-17 (interleukin-17, IL-17)的水平^[37]。在哮喘儿童中有研究检测到高水平的血清 IL-17,该研究认为 IL-17⁺ T 细胞与儿童哮喘的严重程度有关^[38]。哮喘的风险是一种遗传特征,部分由改变的树突状细胞(dendritic cell, DC)表观基因组介导。DC 从哮喘母亲的新生儿转移到非哮喘母亲的新生儿,可增加接受者的哮喘易感性^[39]。DC 是先天和适应性免疫反应的关键因素,在其成熟过程中可以显著上调 *MT2A*^[40]。在本研究的哮喘儿童中,*MT2A* 也呈上调趋势,由此我们认为,*MT2A* 可能是从免疫调节角度出发对儿童哮喘急性发作产生影响,推测其可用作判断儿童哮喘发作的一项新指标;3) 哮喘相关信号通路。在敲除 *MT2A* 的小鼠中, NF- κ B 信号通路表现出明显的免疫反应性,表明 *MT2A* 可以调节免疫反应介导的细胞炎症反应^[41],不仅如此,MT 在细胞系中沉默也会影响肿瘤坏死因子- α (tumor necrosis factor- α , TNF- α)的表达^[42];另外,有报道发现 *MT2A* 与 PI3K/

Akt 信号通路直接相关^[43]。众所周知,在哮喘疾病中, TNF- α /NF- κ B 信号通路与哮喘炎症相关,而 PI3K/Akt 信号通路与哮喘气道重塑关系密切,故推测本研究鉴定出的关键基因 *MT2A* 或可通过 PI3K/Akt、TNF- α /NF- κ B 信号通路对儿童哮喘急性发作起到调控作用。

KIF11 (kinesin family member 11, 另一常用名为 Eg5)是驱动蛋白超家族成员。该蛋白质家族的成员参与各种纺锤体动力学,属于有丝分裂驱动蛋白。目前,对于 *KIF11* 的研究多体现在癌症方面,尚未见有关于儿童哮喘/成人哮喘的相关报道,但根据 *KIF11* 蛋白的上述属性,推测它或许与儿童哮喘气道重塑或哮喘相关信号通路有关。

1) 气道重塑。气道平滑肌层增厚是哮喘重要病理表现之一,而气道平滑肌层厚度的增加主要归因于细胞分裂的增强, *KIF11* 通常作用于微管,直接参与细胞分裂过程,本研究中核心基因 *KIF11* 呈上调趋势,提示其可能与儿童哮喘急性发作的气道重塑相关; 2) 哮喘相关信号通路。有研究发现, *KIF11* 的表达与 Wnt 激活的基因特征呈正相关,同时 *KIF11* 可增强 β -catenin 的核易位,从而激活经典的 Wnt 信号通路^[44]。在哮喘发病过程中, Wnt/ β -catenin 的含量会增高,故推测 *KIF11* 可能通过 Wnt/ β -catenin 通路影响儿童哮喘急性发作。在探索与 *KIF11* 相关通路过程中,我们发现一篇关于 p38 α MAPK 影响 *KIF11* 定位的报道^[45],但该文侧重于阐述 *KIF11* 过表达对受 p38 α MAPK 影响的纺锤体的作用,关于二者之间作用机制的研究较少,目前仅查阅到二者可能是通过 ETS 样蛋白 1 (ETS like-1 protein, ELK-1)等连接因子相互作用来调节纺锤体的组装和有丝分裂过程的假说^[46]。总的来讲, *KIF11* 在儿童哮喘急性发作中的作用或许和 ERK/p38 MAPK 通路有联系,但有待进一步证实。

综上所述,本研究共筛选出 5 个与儿童哮喘急性发作相关的关键基因。综合大量文献报道,这些关键基因中有 3 个基因(*HLA-DPB1*、*HLA-DQB1* 和 *HLA-DQB2*)的诸多位点已被证实与儿童哮喘存在强相关性,有两个基因(*MT2A* 和 *KIF11*)尚未有与哮喘相关的直接报道。结合 *MT2A* 和 *KIF11* 已报道的相关功能,我们认为,作为儿童哮喘急性发作的关键节点基因, *MT2A* 和 *KIF11* 在儿童哮喘急性发作的发生、发展过程中扮演重要角色,或许可以成为儿童哮喘病理机制及防治研究的新

方向,具有潜在的研究价值与研究意义。

参考文献(References):

- [1] MOFFATT M F, KABESCH M, LIANG L, *et al.* Genetic variants regulating *ORMDL3* expression contribute to the risk of childhood asthma[J]. *Nature*, 2007, 448(7152): 470–473.
- [2] WEISS S T, SILVERMAN E K. Pro: genome-wide association studies (GWAS) in asthma[J]. *American Journal of Respiratory and Critical Care Medicine*, 2011, 184(6): 631–633.
- [3] REESE S E, XU C J, DEN DEKKER H T, *et al.* Epigenome-wide meta-analysis of DNA methylation and childhood asthma[J]. *Journal of Allergy and Clinical Immunology*, 2019, 143(6): 2062–2074.
- [4] KHOO S K, READ J, FRANKS K, *et al.* Upper airway cell transcriptomics identify a major new immunological phenotype with strong clinical correlates in young children with acute wheezing[J]. *The Journal of Immunology*, 2019, 202(6): 1845–1858.
- [5] RITCHIE M E, PHIPSON B, WU D, *et al.* limma powers differential expression analyses for RNA-seq and microarray studies[J]. *Nucleic Acids Research*, 2015, 43(7): e47.
- [6] SMYTH G K. Linear models and empirical bayes methods for assessing differential expression in microarray experiments[J]. *Statistical Applications in Genetics and Molecular Biology*, 2004, 3(1): Article3.
- [7] BENJAMINI Y, HOCHBERG Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing[J]. *Journal of the Royal Statistical Society: Series B (Methodological)*, 1995, 57(1): 289–300.
- [8] WICKHAM H. ggplot2: Elegant Graphics for Data Analysis[M]. 2nd ed. Houston: Springer, 2016.
- [9] CHEN C, CHEN H, ZHANG Y, *et al.* TBtools: an integrative toolkit developed for interactive analyses of big biological data[J]. *Molecular Plant*, 2020, 13(8): 1194–1202.
- [10] HUANG D W, SHERMAN B T, LEMPICKI R A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists[J]. *Nucleic Acids Research*, 2009, 37(1): 1–13.
- [11] SZKLARCZYK D, GABLE A L, LYON D, *et al.* STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets[J]. *Nucleic Acids Research*, 2019, 47(D1): D607–D613.
- [12] BADER G D, HOGUE C W V. An automated method for finding molecular complexes in large protein interaction networks[J]. *BMC Bioinformatics*, 2003, 4: 2.
- [13] TRIPATHI P, HONG X, CARUSO D, *et al.* Genetic determinants in the development of sensitization to environmental allergens in early childhood[J]. *Immunity, Inflammation and Disease*, 2014, 2(3): 193–204.
- [14] BABUSIKOVA E, JESENAK M, EVINOVA A, *et al.* Frequency of polymorphism -262 *ct* in catalase gene and oxidative damage in Slovak children with bronchial asthma[J]. *Archivos de Bronconeumologia*, 2013, 49(12): 507–512.
- [15] OW J R, TAN Y H, JIN Y, *et al.* Stra13 and Sharp-1, the non-grouchy regulators of development and disease[J]. *Current Topics in Developmental Biology*, 2014, 110: 317–338.
- [16] SILLESEN M, LI Y, ALAM H B. Transfusion strategies are associated with epigenetic changes following blunt trauma[J]. *Shock*, 2018, 50(1): 24–30.
- [17] XU H, CHENG C, DEVIDAS M, *et al.* *ARID5B* genetic polymorphisms contribute to racial disparities in the incidence and treatment outcome of childhood acute lymphoblastic leukemia[J]. *Journal of Clinical Oncology*, 2012, 30(7): 751–757.
- [18] CHANG Y S, TSAI C T, HUANGFU C A, *et al.* ACSL3 and GSK-3 β are essential for lipid upregulation induced by endoplasmic reticulum stress in liver cells[J]. *Journal of Cellular Biochemistry*, 2011, 112(3): 881–893.
- [19] MULLIGAN M K, WANG X, ADLER A L, *et al.* Complex control of GABA(A) receptor subunit mRNA expression: variation, covariation, and genetic regulation[J]. *PLoS One*, 2012, 7(4): e34586.

- [20] CHAKRABORTY S, DAKLE P, SINHA A, *et al.* Genetic variations in olfactory receptor gene *OR2AG2* in a large multigenerational family with asthma[J]. *Scientific Reports*, 2019, 9: 19029.
- [21] FORNO E, LASKY-SU J, HIMES B, *et al.* Genome-wide association study of the age of onset of childhood asthma[J]. *Journal of Allergy and Clinical Immunology*, 2012, 130(1): 83-90.e4.
- [22] MOFFATT M F, GUT I G, DEMENAI S, *et al.* A large-scale, consortium-based genome-wide association study of asthma[J]. *The New England Journal of Medicine*, 2010, 363(13): 1211-1221.
- [23] GUDBJARTSSON D F, BJORNSDOTTIR U S, HALAPI E, *et al.* Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction[J]. *Nature Genetics*, 2009, 41(3): 342-347.
- [24] GU Z, SHEN Y, TANG X Y, *et al.* Genetic risk of *FCRL3* and *FCRL5* polymorphisms in children with asthma and allergic rhinitis in a Chinese Han population[J]. *International Journal of Pediatric Otorhinolaryngology*, 2019, 120: 58-63.
- [25] GUO S L, LIU F, REN C J, *et al.* Correlations of *LT α* and *NQO1* gene polymorphisms with childhood asthma[J]. *European Review for Medical and Pharmacological Sciences*, 2019, 23(17): 7557-7562.
- [26] LASKY-SU J, HIMES B E, RABY B A, *et al.* HLA-DQ strikes again: genome-wide association study further confirms *HLA-DQ* in the diagnosis of asthma among adults[J]. *Clinical and Experimental Allergy*, 2012, 42(12): 1724-1733.
- [27] HALAPI E, GUDBJARTSSON D F, JONSDOTTIR G M, *et al.* A sequence variant on 17q21 is associated with age at onset and severity of asthma[J]. *European Journal of Human Genetics*, 2010, 18(8): 902-908.
- [28] SVEJGAARD A, PLATZ P, RYDER L P, *et al.* HL-A and disease associations- a survey[J]. *Transplantation Reviews*, 1975, 22: 3-43.
- [29] SHIINA T, HOSOMICHI K, INOKO H, *et al.* The HLA genomic loci map: expression, interaction, diversity and disease[J]. *Journal of Human Genetics*, 2009, 54(1): 15-39.
- [30] MISHRA M N, DUDEJA P, GUPTA R K. Association of HLA-Class II and IgE serum levels in pediatric asthma[J]. *Iranian Journal of Immunology*, 2014, 11(1): 21-28.
- [31] NOGUCHI E, SAKAMOTO H, HIROTA T, *et al.* Genome-wide association study identifies *HLA-DP* as a susceptibility gene for pediatric asthma in Asian populations[J]. *PLoS Genetics*, 2011, 7(7): e1002170.
- [32] CARABALLO L, MARRUGO J, JIMENEZ S, *et al.* Frequency of *DPB1*0401* is significantly decreased in patients with allergic asthma in a Mulatto population[J]. *Human Immunology*, 1991, 32(3): 157-161.
- [33] GAO J, LIN Y, QIU C, *et al.* Association between *HLA-DQA1*, *-DQB1* gene polymorphisms and susceptibility to asthma in northern Chinese subjects[J]. *Chinese Medical Journal*, 2003, 116(7): 1078-1082.
- [34] MOVAHEDI M, MOIN M, GHARAGOZLOU M, *et al.* Association of HLA class II alleles with childhood asthma and total IgE levels[J]. *Iranian Journal of Allergy, Asthma and Immunology*, 2008, 7(4): 215-220.
- [35] WEI CHOO C Y, YEH K W, HUANG J L, *et al.* Oxidative stress is associated with atopic indices in relation to childhood rhinitis and asthma[J]. *Journal of Microbiology, Immunology and Infection*, 2020. DOI: 10.1016/j.jmii.2020.01.009.
- [36] DE CASTRO MENDES F, PACIÊNCIA I, CAVALEIRO RUFO J, *et al.* The inflammatory potential of diet impacts the association between air pollution and childhood asthma[J]. *Pediatric Allergy and Immunology*, 2020, 31(3): 290-296.
- [37] KOZELA E, JUKNAT A, GAO F, *et al.* Pathways and gene networks mediating the regulatory effects of cannabidiol, a nonpsychoactive cannabinoid, in autoimmune T cells[J]. *Journal of Neuroinflammation*, 2016, 13: 136.
- [38] LICARI A, MANTI S, CASTAGNOLI R, *et al.* Immunomodulation in pediatric asthma[J]. *Frontiers in Pediatrics*, 2019, 7: 289.
- [39] BOUKHALED G M, CORRADO M, GUAQ H, *et al.* Chromatin architecture as an essential determinant of dendritic cell function[J]. *Frontiers in Immunology*, 2019, 10: 1119.
- [40] JIN P, HAN T H, REN J, *et al.* Molecular signatures of maturing dendritic cells: implications for testing the quality of dendritic cell therapies[J]. *Journal of Translational Medicine*, 2010, 8: 4.
- [41] WU M, LENG W, PAN H, *et al.* The reduced expression of *EO-LA1* may be related to refractory diabetic foot ulcer[J]. *Mediators of Inflammation*, 2019, 2019: 6705424.
- [42] AQUIME J R H S, ZAMPIERI L C D P, KATAOKA M S D S, *et al.* Metallothionein expression and its influence on the *in vitro* biological behavior of mucoepidermoid carcinoma[J]. *Cells*, 2020, 9(1): 157.
- [43] BOSSLER F, KUHN B J, GÜNTHER T, *et al.* Repression of human papillomavirus oncogene expression under hypoxia is mediated by PI3K/mTORC2/AKT signaling[J]. *mBio*, 2019, 10(1): e02323-18.
- [44] PEI Y Y, LI G C, RAN J, *et al.* Kinesin family member 11 enhances the self-renewal ability of breast cancer cells by participating in the Wnt/ β -catenin pathway[J]. *Journal of Breast Cancer*, 2019, 22(4): 522-532.
- [45] OU X H, LI S, XU B Z, *et al.* p38 α MAPK is a MTOC-associated protein regulating spindle assembly, spindle length and accurate chromosome segregation during mouse oocyte meiotic maturation[J]. *Cell Cycle*, 2010, 9(20): 4130-4143.
- [46] LIANG Y J, YANG W X. Kinesins in MAPK cascade: how kinesin motors are involved in the MAPK pathway?[J]. *Gene*, 2019, 684: 1-9.

(上接第 57 页)

- [27] PELVIG D P, PAKKENBERG H, REGEUR L, *et al.* Neocortical glial cell numbers in Alzheimer's disease. A stereological study[J]. *Dementia and Geriatric Cognitive Disorders*, 2003, 16: 212-219.
- [28] OLABARRIA M, NORISTANI H N, VERKHRATSKY A, *et al.* Concomitant astroglial atrophy and astrogliosis in a triple transgenic animal model of Alzheimer's disease[J]. *Glia*, 2010, 58(7): 831-838.
- [29] KATSSOURI L, BIRCH A M, RENZIEHAUSEN A W J, *et al.* Ablation of reactive astrocytes exacerbates disease pathology in a model of Alzheimer's disease[J]. *Glia*, 2020, 68(5): 1017-1030.
- [30] OZAWA S, KAMIYA H, TSUZUKI K. Glutamate receptors in the mammalian central nervous system[J]. *Progress in Neurobiology*, 1998, 54(5): 581-618.
- [31] BRITO-MOREIRA J, PAULA-LIMA A C, BOMFIM T R, *et al.* *A β* oligomers induce glutamate release from hippocampal neurons[J]. *Current Alzheimer Research*, 2011, 8(5): 552-562.
- [32] WANG Y, QIN Z H. Molecular and cellular mechanisms of excitotoxic neuronal death[J]. *Apoptosis*, 2010, 15(11): 1382-1402.
- [33] HU N W, NICOLL A J, ZHANG D, *et al.* mGlu5 receptors and cellular prion protein mediate amyloid β -facilitated synaptic long-term depression *in vivo*[J]. *Nature Communications*, 2014, 5: 3374.
- [34] ZHANG D, QI Y, KLYUBIN I, *et al.* Targeting glutamatergic and cellular prion protein mechanisms of amyloid β -mediated persistent synaptic plasticity disruption: longitudinal studies[J]. *Neuropharmacology*, 2017, 121: 231-246.
- [35] BORTOLOTTO Z A, COLLETT V J, CONQUET F, *et al.* The regulation of hippocampal LTP by the molecular switch, a form of metaplasticity, requires mGlu5 receptors[J]. *Neuropharmacology*, 2005, 49(Suppl. 1): 13-25.