

·研究论文·

DOI:10.16605/j.cnki.1007-7847.2020.03.006

基于蛋白质互作网络挖掘自闭症谱系障碍的功能模块与核心基因

许逸聪^a, 胡婉雪^a, 谢芹^a, 赵洪波^{a*}, 唐欣^{b*}

(昆明医科大学 a. 分子临床医学研究院暨云南省干细胞和再生医学重点实验室; b. 康复学院, 中国云南 昆明 650500)

摘要: 自闭症谱系障碍(autism spectrum disorder, ASD)是一种具有高遗传性、临床异质性和生物复杂性的神经行为障碍类疾病。为挖掘 ASD 发生发展过程中的功能模块与核心基因, 本文从自闭症谱系障碍疾病数据库获取 ASD 相关基因; 利用 STRING 数据库构建 ASD 相关基因的蛋白质互作网络; 通过 MCODE 算法对蛋白质互作网络进行模块分析并筛选核心基因; 最后对各功能模块进行 KEGG 通路分析, 根据富集到的通路类别评估功能模块之间的相互作用。结果显示, 3 个疾病基因数据库筛选出 182 个共有基因, 构建的蛋白质互作网络包含 171 个节点和 1 041 条边, 其中 *NRXN1*、*GRIN2B*、*GRIN2A*、*DLG4*、*NLGN3*、*MECP2*、*CNTNAP2*、*BDNF*、*NLGN4X*、*FMR1* 等 23 个基因具有较高的连通度(degree)。从蛋白质互作网络中分析得到 5 个功能模块, 包括 68 个核心基因。KEGG 富集分析发现功能模块参与多个生物学通路, 包括细胞黏附分子、钙离子通路、神经活性的配体-受体相互作用、多巴胺能神经突触等。分析结果提示, 挖掘的 ASD 功能模块和核心基因大多集中在神经元活动、信号分子和信号传导等, 且各模块相互作用共同影响 ASD 的发生发展。

关键词: 自闭症谱系障碍(ASD); 蛋白质互作网络; 功能模块; 核心基因

中图分类号: Q811.4, R749.94

文献标识码: A

文章编号: 1007-7847(2020)03-0214-08

Identification of Functional Modules and Hub Genes in Autism Spectrum Disorder Based on Protein-Protein Interaction Network Analysis

XU Yi-cong^a, HU Wan-xue^a, XIE Qin^a, ZHAO Hong-bo^{a*}, TANG Xin^{b*}

(a. Yunnan Key Laboratory of Stem Cell and Regenerative Medicine, Institute of Molecular and Clinical Medicine; b. School of Rehabilitation, Kunming Medical University, Kunming 650500, Yunnan, China)

Abstract: Autism spectrum disorder (ASD) is a neurobehavioral disorder with high heritability, clinical heterogeneity and biological complexity. It is necessary to identify functional modules and hub genes related to ASD. Herein, ASD related genes were obtained from disease-gene databases and imported into STRING database for protein-protein interaction (PPI) network construction. Then module analysis and hub gene selection were performed by MCODE algorithm. Subsequently, KEGG pathway enrichment analysis of each functional module was performed for evaluating the interactions between the functional modules based on the enriched pathway categories. A total of 182 related genes were screened from disease-gene databases. These overlapped genes were translated into PPI network consisting of 171 nodes and 1 041 edges, including *NRXN1*, *GRIN2B*, *GRIN2A*, *DLG4*, *NLGN3*, *MECP2*, *CNTNAP2*, *BDNF*, *NLGN4X* and *FMR1* with higher degree. Five functional modules including 68 hub genes were extracted using MCODE algorithm. KEGG en-

收稿日期: 2019-10-17; 修回日期: 2019-12-10

基金项目: 云南省科技厅-昆明医科大学联合专项面上项目(2018FE001-197); 昆明医科大学百名中青年学术技术骨干项目(60118090114)

作者简介: 许逸聪(1992-), 男, 浙江绍兴人, 硕士研究生; *通信作者: 赵洪波(1980-), 男, 四川西充人, 博士, 昆明医科大学副教授, 主要从事生物信息学研究, Tel: 0871-65922699, E-mail: zhao.hongbo@hotmail.com; 唐欣(1983-), 女, 江西龙南人, 助教, 主要从事自闭症谱系障碍康复研究, Tel: 0871-65926950, E-mail: tangxin8140730@163.com。

richment analysis of detected modules revealed that these modules are related to cell adhesion molecules, calcium signaling pathway, neuroactive ligand-receptor interaction and dopaminergic synapse. The results suggested that most of the ASD functional modules and hub genes are concentrated in neuronal activity, signal molecules and signaling transduction, and the interaction of each module affects the occurrence and development of ASD.

Key words: autism spectrum disorder (ASD); protein-protein interaction (PPI) network; functional module; hub genes

(*Life Science Research*, 2020, 24(3): 214-221)

自闭症谱系障碍(autism spectrum disorder, ASD)是一类常见的神经发育疾病,诊断标准包括持续的社会交流与互动缺失、语言障碍、刻板重复行为及狭隘的兴趣。除上述症状外,患者可能同时伴随有癫痫、睡眠障碍、抑郁、焦虑以及胃肠道活动紊乱等症状,表现出明显的异质性和多样性^[1],这些症状在儿童早期出现并严重影响患儿的日常生活^[2]。作为一种具有高遗传性和生物复杂性的神经行为障碍类疾病,ASD是由遗传、环境等复杂因素相互作用形成的,其中遗传变异导致的发病率相对较高^[3]。2016年发表的一项基于双胞胎研究的荟萃分析指出,64%~91%的ASD风险是由遗传因素引起的^[4]。

尽管目前已发现上百种风险基因与ASD相关,表明ASD具有明显的遗传异质性,但遗传背景有差异的ASD个体往往表现出相似的行为特征^[5]。越来越多研究也表明,与ASD密切相关的基因可能汇聚于一些共同的生物学过程^[6-7],在ASD发生过程中起关键作用,但尚未有研究系统筛选这些功能模块和核心基因。因此,本研究利用生物信息学方法,通过疾病基因数据库中报道的ASD相关基因信息,整合蛋白质互作(protein-protein interaction, PPI)网络先验知识,挖掘ASD的功能模块和核心基因,并对每个模块进行通路富集分析,为ASD的遗传研究提供参考。

1 材料与方法

1.1 基因集数据获取

本研究从3个疾病数据库收集ASD相关基因。AutismKB 2.0 (http://db.cbi.pku.edu.cn/autismkb_v2/index.php)是ASD遗传证据的知识库^[8],当前版本包含ASD相关的1379个基因(99个综合征自闭症相关基因和1280个非综合征自闭症相关基因)、5420个拷贝数变异(copy number variation, CNV)/结构变异(structural variation, SV)、11669个

单核苷酸变异(single nucleotide variation, SNV)/插入和缺失(insertions and deletions, InDels)以及与ASD相关的172个连锁区域。SFARI Gene数据库(<https://gene.sfari.org/>)收录了1089个ASD基因、2290个CNV基因座和2296个动物模型^[9]。DisGeNET是一个综合性疾病基因数据库^[10],整合了来自专家库、全基因组关联分析(genome-wide association study, GWAS)、动物模型和相关学术文献的数据,通过“Autism Spectrum Disorder”进行检索,得到571个相关基因。为获得稳健的基因集,筛选出在3个数据库中都出现的基因用于后续分析。

1.2 蛋白质互作网络构建

交互基因检索工具(Search Tool for the Retrieval of Interacting Genes, STRING)是一种用于评价蛋白质相互作用的工具^[11]。将筛选出的ASD基因集导入STRING在线工具(<http://string-db.org>),选择物种为*Homo sapiens*,设置置信度阈值大于0.4,得到基因集表达的蛋白质互作网络(PPI network)。通过Cytoscape 3.7.2软件进行可视化和连通度分析,将筛选出的连通度大于25的节点作为PPI网络的重要基因。

1.3 功能模块及核心基因识别

利用MCODE (Molecular Complex Detection)算法对PPI网络中的节点进行密度聚类^[12]。首先排除连通度小于3的节点,计算纳入节点的核聚类系数(core-clustering coefficient)并将其作为节点的权重,之后以当前节点权重百分率(vertex weight percentage, VWP)为阈值(本研究设VWP为0.2),筛选阈值范围内的相关节点构建模块,并将各模块中所包含的基因作为核心基因。上述过程通过Cytoscape软件中的MCODE插件进行,参数设置:degree cutoff ≥ 3 , K-score ≥ 3 ,其余选择默认值。

1.4 通路富集分析

采用R软件包clusterProfiler^[13]对得到的各功能模块进行KEGG (kyoto encyclopedia of genes and

genomes)通路富集分析,通过超几何分布计算各通路富集结果的显著性,并使用 Bonferroni 方法进行多重假设检验校正,以错误发现率(false discovery rate, FDR)小于 0.05 为标准,对富集结果进行筛选,得到显著富集的 KEGG 通路。将富集到的 KEGG 通路输入 KEGG 数据库,识别其类别,进而评估功能模块之间的相互作用,利用 Cytoscape 3.7.2 软件构建功能模块与通路分类的关系。

2 结果

2.1 ASD 相关基因筛选结果

为获得稳健的基因集,本研究共纳入了 AutismKB 数据库中的 1 379 个基因、SFARI Gene 数据库中的 1 089 个基因以及 DisGeNET 数据库中的 571 个基因。通过取交集,最终筛选出共有的 182 个基因作为 ASD 相关基因集用于后续分析(图 1)。

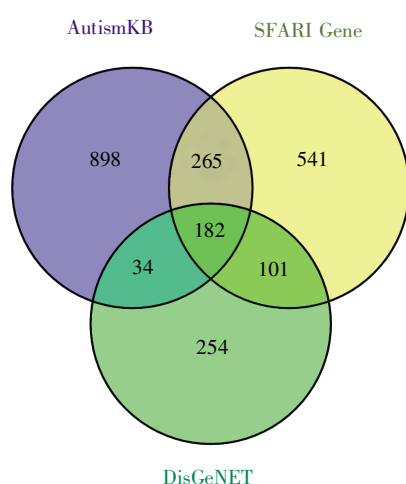


图 1 基于 3 个数据库获得 ASD 共同基因的 Venn 图
Fig.1 Venn diagram of common ASD genes from three databases

2.2 ASD 基因集的蛋白质互作网络

将 182 个 ASD 相关基因导入 STRING 数据库,进行 PPI 网络分析。结果显示,整个网络包含 171 个节点和 1 041 条边,其中每个节点代表基因对应的蛋白质,每条边代表两个蛋白质之间的互作关系(图 2)。根据 PPI 网络中每个节点的连通度,筛选得到 PPI 网络中核心蛋白质的编码基因 23 个,分别为 *NRXN1*、*GRIN2B*、*GRIN2A*、*DLG4*、*NLGN3*、*MECP2*、*CNTNAP2*、*BDNF*、*NLGN4X*、*FM-R1*、*SHANK2*、*RELN*、*CHD8*、*SNAP25*、*CACNA1C*、*GAD1*、*PTEN*、*NRXN3*、*GABRB3*、*SYN1*、*SHANK1*、

SCN2A 和 *UBE3A*。

2.3 ASD 疾病网络中的功能模块与核心基因

通过 MCODE 算法对 PPI 网络进行分析,最终获得 5 个功能模块,包含 68 个核心基因(表 1 和图 3)。节点颜色越深表明该节点在其模块网络中的连通度越高,即其所代表的编码基因在模块网络中的重要性越高。评分最高的模块 1 (M1)包含 20 个基因及 146 条边,平均每个节点含有 7.3 条边,说明该模块中的基因之间关系比较密切,其中排名靠前的基因有 *GRIN2B*、*GRIN2A*、*DLG4*、*MECP2*、*NLGN3*、*CNTNAP2*、*NLGN4X*、*FMR1*、*SHANK2* 等。模块 2 (M2)由 20 个基因和 54 条边组成,平均每个节点含有 2.7 条边,其中 *BDNF* 处于模块核心位置。模块 3 (M3)有 16 个核心基因,模块 4 (M4)含有的核心基因最少,仅 4 个。

2.4 各功能模块的通路富集分析

对筛选出的各个模块进行 KEGG 通路富集分析,结果显示:模块 M1 显著富集于谷氨酸能突触、细胞黏附分子和物质依赖通路;模块 M2 主要富集于突触囊泡循环、多巴胺能神经突触、神经活性的配体-受体相互作用等通路;模块 M3 富集的通路有促性腺激素分泌、昼夜节律、MAPK 信号通路、钙离子通路等;模块 M4 主要富集在免疫相关通路;模块 M5 涉及到钙离子通路和神经活性的配体-受体相互作用(表 2)。分析模块之间的共同通路发现, M1 和 M2 都包括可卡因成瘾通路, M1 和 M4 都包括细胞黏附分子通路, M2 和 M5 都涉及神经活性的配体-受体相互作用通路, M3 和 M5 都包括钙离子信号通路。

为了展示各模块之间的关系,根据各模块富集到的 KEGG 通路类别评估 5 个功能模块之间的相互作用,构建模块和通路分类的关系图(图 4)。图中的功能类是 KEGG 通路的分类,包括信号分子和相互作用(signaling molecules and interaction)、神经系统(nervous system)、物质依赖(substance dependence)、内分泌代谢病(endocrine and metabolic disease)、信号转导(signal transduction)、转运和降解(transport and catabolism)、环境适应(environmental adaptation)、免疫系统(immune system)、免疫疾病(immune disease)、细胞生长和死亡(cell growth and death)、内分泌系统(endocrine system)、传染病(infectious disease)、循环系统(circulatory system)、抗药性(drug resistance)、心血管疾病(cardiovascular disease)和肿瘤(cancer)。其中信号分子和相互

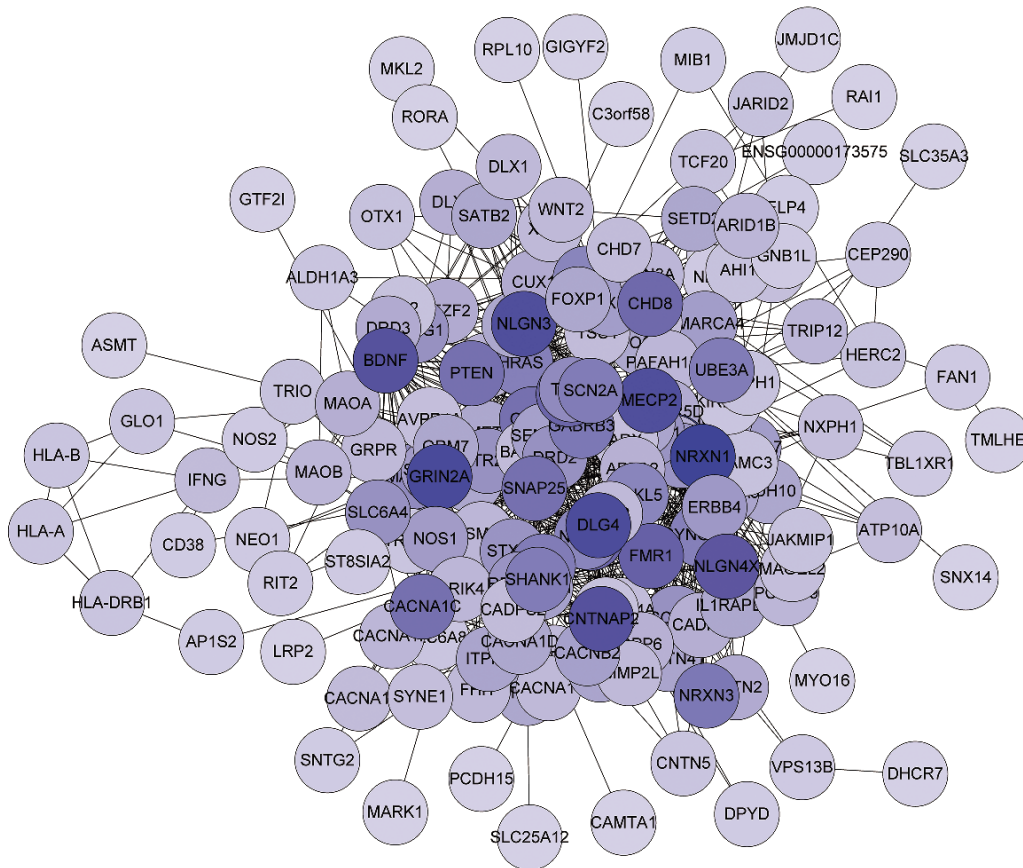


图 2 ASD 相关基因的 PPI 网络图

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相互作用。节点连边越多表示其连通度越高,在网络中的位置越趋于中心,且颜色也越深。

Fig.2 The PPI network of ASD related genes

Each dot represents a protein corresponding to each gene, and the edge represents the interaction between the proteins. The center nodes with deeper color represent higher degree.

表 1 各模块的基本信息

Table 1 Basic information of each module

Module	Score	Node number	Edge number	Hub genes
M1	15.368	20	146	<i>GRIN2B, GRIN2A, DLG4, MECP2, NLGN3, CNTNAP2, NLGN4X, FMR1, SHANK2, CHD8, GABRB3, NRXN3, SHANK1, UBE3A, SYN1, SYNGAP1, CDKL5, NRXN2, CASK, HOMER2</i>
M2	5.684	20	54	<i>BDNF, SNAP25, PAX6, TBR1, HTR2A, FOXG1, DRD2, STX1A, SLC6A3, SEMA5A, CNTN4, GRM7, CDH10, SATB2, ASTN2, DRD3, FEZF2, DLX2, APBA2, DLX1</i>
M3	4.400	16	33	<i>GAD1, PTEN, CACNA1C, HRAS, MAPK1, SLC6A4, NOS1, FOXP2, MET, FOXP1, ATP2B2, CACNA1D, CUX1, CACNA1G, TSC1, CACNA1I</i>
M4	4.000	4	6	<i>IFNG, HLA-A, HLA-B, HLA-DRB1</i>
M5	3.143	8	11	<i>OXTR, CACNB2, EIF4E, WNT1, PPP2R5D, GRPR, AVPR1A, TRIO</i>

作用(M1、M2、M4 和 M5)与神经系统(M1、M2 和 M3)这两个通路分类的连通度最高。

3 讨论

ASD 是一种复杂的遗传性神经行为障碍类疾病,目前已有的一些 ASD 相关的分子机制的研究^[14-15],

但各独立研究之间缺乏一个明确的共识。ASD 的高异质性则表明,其发生发展涉及多种生物学过程的相互作用,而这些互作关系所涉及的具体机制尚不明确。本研究结合 ASD 疾病基因数据库与蛋白质互作网络先验知识,构建了 ASD 风险基因网络,通过网络分解,挖掘出紧密联系的功能模

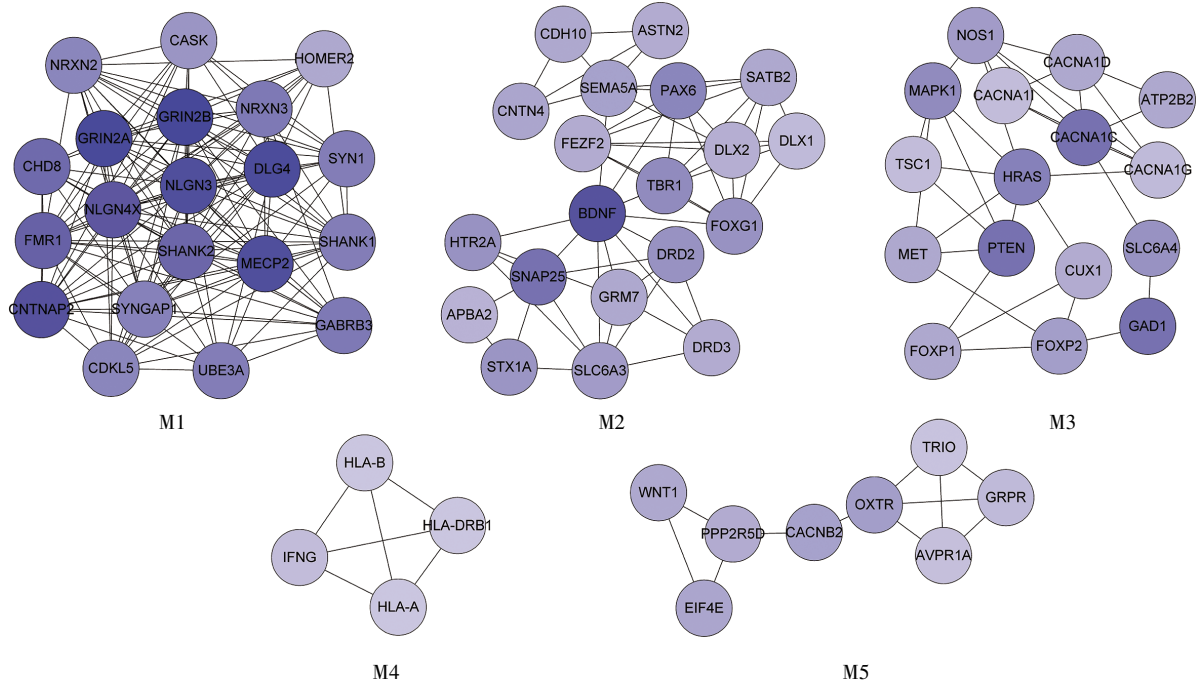


图3 各功能模块网络图

各节点代表功能模块的核心基因，颜色越深表示该基因的连通度越大。

Fig.3 Network of each functional module

The nodes represent the hub genes in functional module, and the deeper colors represent genes with higher degree.

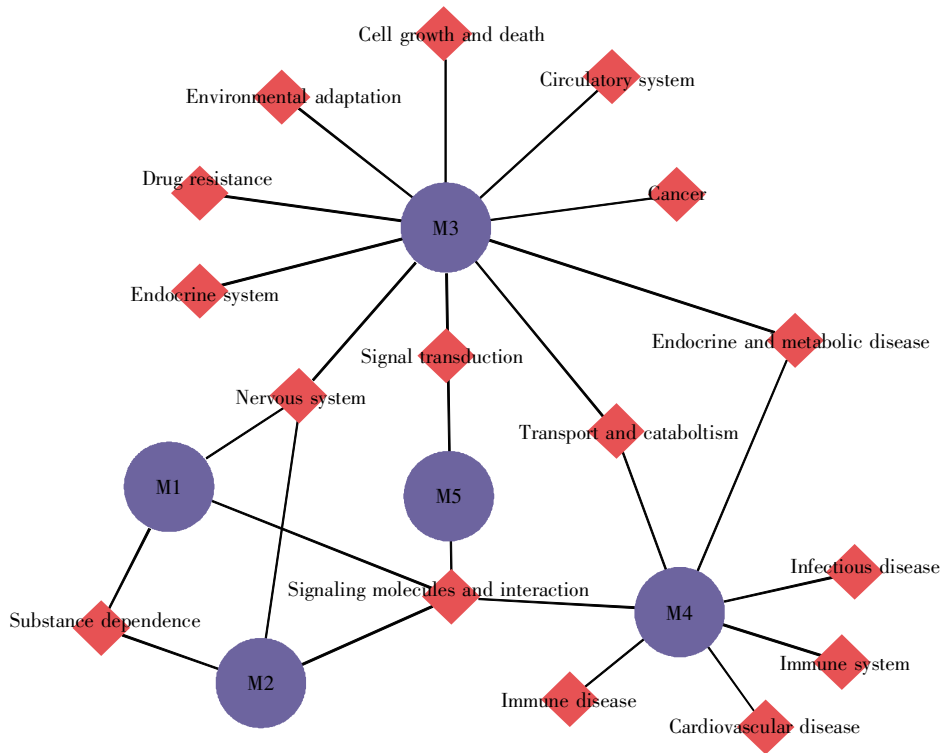


图4 功能模块与通路分类关系图

蓝色表示模块，红色表示通路分类。

Fig.4 Diagram of functional modules and pathway classification

Blue indicates the module and red indicates the pathway classification.

表 2 各功能模块的 KEGG 通路富集分析结果
Table 2 The results of KEGG pathway enrichment analysis in each functional module

Module	KEGG ID	Pathway name	FDR	Pathway classification	
M1	hsa04724	Glutamatergic synapse	1.617E-06	Nervous system	
	hsa04514	Cell adhesion molecules (CAMs)	1.073E-04	Signaling molecules and interaction	
	hsa05033	Nicotine addiction	6.188E-04	Substance dependence	
	hsa05030	Cocaine addiction	8.558E-04	Substance dependence	
M2	hsa05030	Cocaine addiction	9.243E-04	Substance dependence	
	hsa04721	Synaptic vesicle cycle	1.867E-03	Nervous system	
	hsa04728	Dopaminergic synapse	5.646E-03	Nervous system	
	hsa04080	Neuroactive ligand-receptor interaction	5.646E-03	Signaling molecules and interaction	
	hsa05034	Alcoholism	9.245E-03	Substance dependence	
M3	hsa04929	GnRH secretion	9.742E-08	Endocrine system	
	hsa04713	Circadian entrainment	6.238E-07	Environmental adaptation	
	hsa04010	MAPK signaling pathway	1.319E-05	Signal transduction	
	hsa04925	Aldosterone synthesis and secretion	1.540E-05	Endocrine system	
	hsa04020	Calcium signaling pathway	1.540E-05	Signal transduction	
	hsa04930	Type II diabetes mellitus	2.371E-05	Endocrine and metabolic disease	
	hsa04726	Serotonergic synapse	2.371E-05	Nervous system	
	hsa04927	Cortisol synthesis and secretion	7.558E-05	Endocrine system	
	hsa04934	Cushing syndrome	7.602E-05	Endocrine and metabolic disease	
	hsa05230	Central carbon metabolism in cancer	7.602E-05	Cancer	
	hsa04218	Cellular senescence	7.602E-05	Cell growth and death	
	hsa05218	Melanoma	7.602E-05	Cancer	
	hsa01521	EGFR tyrosine kinase inhibitor resistance	1.018E-04	Drug resistance	
	hsa04912	GnRH signaling pathway	1.810E-04	Drug resistance	
	hsa04725	Cholinergic synapse	3.525E-04	Nervous system	
	hsa04140	Autophagy-animal	7.284E-04	Transport and catabolism	
	hsa04261	Adrenergic signaling in cardiomyocytes	9.431E-04	Circulatory system	
	hsa04150	mTOR signaling pathway	9.431E-04	Signal transduction	
	hsa04921	Oxytocin signaling pathway	9.431E-04	Endocrine system	
	hsa05213	Endometrial cancer	9.855E-04	Cancer	
	hsa05206	MicroRNAs in cancer	9.855E-04	Cancer	
	M4	hsa05330	Allograft rejection	1.234E-08	Immune disease
		hsa05332	Graft-versus-host disease	1.234E-08	Immune disease
hsa04940		Type I diabetes mellitus	1.234E-08	Endocrine and metabolic disease	
hsa04612		Antigen processing and presentation	1.015E-07	Immune system	
hsa05320		Autoimmune thyroid disease	1.107E-05	Immune system	
hsa05416		Viral myocarditis	1.346E-05	Cardiovascular disease	
hsa05168		Herpes simplex virus 1 infection	1.025E-04	Infectious disease	
hsa04650		Natural killer cell mediated cytotoxicity	1.073E-04	Immune system	
hsa04514		Cell adhesion molecules (CAMs)	1.322E-04	Signaling molecules and interaction	
hsa04145		Phagosome	1.343E-04	Transport and catabolism	
hsa05169		Epstein-Barr virus infection	2.823E-04	Infectious disease	
hsa05166		Human T-cell leukemia virus 1 infection	3.345E-04	Infectious disease	
M5		hsa04020	Calcium signaling pathway	1.627E-02	Signal transduction
	hsa04080	Neuroactive ligand-receptor interaction	4.231E-02	Signaling molecules and interaction	

块和核心基因,进而分析了各模块的功能。为得到稳健的候选基因集,我们从 3 个数据库获取基因交集,增加了结果的可信度。分析结果显示筛选得到 68 个核心基因,并提取了 ASD 相关的 5 个功能模块。

为阐明模块的生物学功能,本研究对每个模块中的核心基因进行了基于 KEGG 信号通路的富集分析。评分最高的模块 M1 主要与谷氨酸能

神经元突触形成及可塑性等过程相关。其中, *NLGN3*、*NRXN3* 等基因编码的神经连接蛋白(neureligns, NLGNs)家族和神经轴突蛋白(neurexins, NRXNs)家族均为单次跨膜 I 型蛋白家族。在谷氨酸能或 γ 氨基丁酸(γ -aminobutyric acid, GABA)能神经元中, NLGNs 蛋白家族位于突触后膜,而 NRXNs 蛋白家族则位于突触前膜。两者可以相互结合并与 SHANK3 蛋白绑定形成复合物,在

突触的形成以及突触间的信息传递等过程中起到核心调控作用^[16]。*GRIN2A* 及 *GRIN2B* 基因分别编码 N-甲基-D-天冬氨酸(N-methyl-D-aspartate, NMDA)受体的 2a 及 2b 亚基,其在神经发育早期阶段有着丰富的表达,并在发育后期的表达量逐渐降低,表明其主要参与神经系统的早期发育过程;此外,*GRIN2B* 也在神经元的增殖迁移、轴突树突形成、学习记忆及突触可塑性等多种神经活动相关过程中起调控作用^[17]。*DLG4* 基因编码 PSD-95 蛋白,该蛋白质是突触后区域中一个特征明确的支架蛋白质,可与各种其他功能蛋白质相结合^[18]。在突触后神经元中,SHANK 蛋白与 PSD-95 构成复合物,并与 NMDA 受体交联,参与谷氨酸能神经活动调节的突触可塑性^[19]。*FMR1* 基因编码脆性 X 智力迟钝蛋白 (fragile X mental retardation protein, FMRP),FMRP 定位于神经元的胞体和树突,通过与 mRNA 结合调节蛋白质合成,主要起转录抑制因子的作用。而 SHANK 蛋白家族则作为 FMRP 的下游靶标被调控^[20]。*CASK* 基因编码钙/钙调蛋白依赖性丝氨酸蛋白激酶,在突触前神经元中,通过与 NRXNs 结合构成复合物,从而连接细胞骨架,参与突触可塑性的相关调节^[21]。通过文献挖掘和模块网络图可以发现, M1 中各核心基因之间关系紧密,任何部分的改变均可能导致谷氨酸能神经元突触形成及可塑性过程的异常。比如:通过编辑小鼠的 NLGNs 蛋白家族相关基因发现,无论是这些基因的缺失还是过度表达,都可能导致小鼠出现 ASD 样的行为表现和中枢兴奋性与抑制性网络环路的异常^[22];*GRIN2B* 区域内的多个单核苷酸多态性 (single nucleotide polymorphism, SNP)与 ASD 的发病密切相关,不同的突变表型会影响 Mg^{2+} 对 Ca^{2+} 渗透性阳离子通道的电压依赖性阻滞作用,从而导致异常的神经电活动^[23-24];*CASK* 作为发育障碍相关的基因之一,其突变也在 ASD 患者中被发现^[25]。此外,小鼠 *Fmr1* 的缺失导致功能性皮质内神经元连接的发育缺陷^[26],兴奋与抑制的失衡^[27],躯体感觉皮层神经元网络活动的同步性增加^[28],以及视觉皮层中神经元电活动依赖的可塑性异常^[29]。

模块 M2 中包括的核心基因则与神经递质的囊泡转运等过程相关。*APBA2* 编码的蛋白质又称 MINT2,是神经元衔接蛋白家族成员之一,其通过与 *CASK* 蛋白结合,进而与 NRXNs 家族蛋白交联,并构成一种转运蛋白复合体^[30]。该转运蛋白复

合体通过与 *CASK* 相连接的细胞骨架将 Munc18-1 等突触囊泡相关蛋白质募集到 NRXNs 家族蛋白周围,而 Munc18-1 是一种 sec1 样蛋白,参与胞吐作用等神经递质释放相关的过程^[31]。*SNAP25* 与 *STX1A* 编码的蛋白质共同参与构成可溶性 N-乙基马来酰亚胺敏感因子附着蛋白受体(soluble N-ethylmaleimide-sensitive factor attachment protein receptor, SNARE)复合物^[32],这种蛋白质复合物在 Munc18-1 的拉链作用促进下黏附谷氨酸或 GABA 递质囊泡,并通过囊泡的依赖性膜融合开始突触囊泡循环过程^[33]。模块 M3 的核心基因主要涉及突触后膜电活性调节等相关过程。其中 *NOS1* 编码神经元型一氧化氮合酶(neuronal nitric oxide synthase, nNOS),nNOS 通过 PSD-95 与 NMDA 受体交联,调节 NMDA 受体相关下游过程的激活,包括 *CACNA1C*、*CACNA1D* 等编码的电压依赖性钙通道等,从而对突触后区域的活性及可塑性进行调节^[34]。

通过对功能模块内核心基因的分析,我们发现 NRXNs 家族、PSD-95、*CASK* 等以功能蛋白质复合物的形式参与 ASD 相关生物学过程。这些核心基因的异常将会影响模块的功能,进而导致疾病的发生发展。根据富集到的 KEGG 通路类别,我们得到了更为深入的功能联系,说明了模块与模块之间不是孤立的,而是相互作用,共同参与 ASD 演化的核心病理机制。

综上所述,本研究通过数据库挖掘及蛋白质互作网络分析,筛选出了 ASD 相关的 5 个功能模块,得到 68 个核心基因,它们汇聚于一些共同的生物学过程,如突触传递、细胞间相互作用及细胞内信号传递的分子通路等。这些模块和基因异常均会导致整体网络的失衡,使得神经发育中的某些过程发生延迟或中断,从而导致感知觉与学习记忆能力异常、昼夜节律障碍等 ASD 相关症状的发生。本研究对 ASD 基因信息的挖掘和整合,有助于进一步地了解 ASD 的分子机制,可为 ASD 的基础研究提供新的参考。

参考文献(References):

- [1] ANAGNOSTOU E. Clinical trials in autism spectrum disorder: evidence, challenges and future directions[J]. *Current Opinion in Neurology*, 2018, 31(2): 119-125.
- [2] ELSABBAGH M, DIVAN G, KOH Y J, *et al.* Global prevalence of autism and other pervasive developmental disorders[J]. *Autism Research*, 2012, 5(3): 160-179.

- [3] MASI A, DEMAYO M M, GLOZIER N, *et al.* An overview of autism spectrum disorder, heterogeneity and treatment options[J]. *Neuroscience Bulletin*, 2017, 33(2): 183–193.
- [4] TICK B, BOLTON P, HAPPE F, *et al.* Heritability of autism spectrum disorders: a meta-analysis of twin studies[J]. *Journal of Child Psychology and Psychiatry*, 2016, 57(5): 585–595.
- [5] GESCHWIND D H, STATE M W. Gene hunting in autism spectrum disorder: on the path to precision medicine[J]. *The Lancet Neurology*, 2015, 14(11): 1109–1120.
- [6] PINTO D, DELABY E, MERICO D, *et al.* Convergence of genes and cellular pathways dysregulated in autism spectrum disorders[J]. *American Journal of Human Genetics*, 2014, 94(5): 677–694.
- [7] KRISHNAN A, ZHANG R, YAO V, *et al.* Genome-wide prediction and functional characterization of the genetic basis of autism spectrum disorder[J]. *Nature Neuroscience*, 2016, 19(11): 1454–1462.
- [8] YANG C, LI J, WU Q, *et al.* AutismKB 2.0: a knowledgebase for the genetic evidence of autism spectrum disorder[J]. *Database*, 2018, 2018: bay106.
- [9] ABRAHAMS B S, ARKING D E, CAMPBELL D B, *et al.* SFARI Gene 2.0: a community-driven knowledgebase for the autism spectrum disorders (ASDs)[J]. *Molecular Autism*, 2013, 4: 36.
- [10] PINERO J, BRAVO A, QUERALT-ROSINACH N, *et al.* DisGeNET: a comprehensive platform integrating information on human disease-associated genes and variants[J]. *Nucleic Acids Research*, 2017, 45(D1): D833–D839.
- [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. An automated method for finding molecular complexes in large protein interaction networks[J]. *BioMed Central Bioinformatics*, 2003, 4: 2.
- [13] YU G C, WANG L G, HAN Y Y, *et al.* clusterProfiler: an R package for comparing biological themes among gene clusters[J]. *OMICS*, 2012, 16(5): 284–287.
- [14] DOAN R N, LIM E T, DE RUBEIS S, *et al.* Recessive gene disruptions in autism spectrum disorder[J]. *Nature Genetics*, 2019, 51(7): 1092–1098.
- [15] GROVE J, RIPKE S, ALS T D, *et al.* Identification of common genetic risk variants for autism spectrum disorder[J]. *Nature Genetics*, 2019, 51(3): 431–444.
- [16] CAO X, TABUCHI K. Functions of synapse adhesion molecules neurexin/neuroligins and neurodevelopmental disorders[J]. *Neuroscience Research*, 2017, 116: 3–9.
- [17] ENDELE S, ROSENBERGER G, GEIDER K, *et al.* Mutations in *GRIN2A* and *GRIN2B* encoding regulatory subunits of NMDA receptors cause variable neurodevelopmental phenotypes[J]. *Nature Genetics*, 2010, 42(11): 1021–1026.
- [18] FRANK R A, GRANT S G. Supramolecular organization of NMDA receptors and the postsynaptic density[J]. *Current Opinion in Neurobiology*, 2017, 45: 139–147.
- [19] NAISBITT S, KIM E, TU J C, *et al.* Shank, a novel family of postsynaptic density proteins that binds to the NMDA receptor/PSD-95/GKAP complex and cortactin[J]. *Neuron*, 1999, 23(3): 569–582.
- [20] DARNELL J C, VAN DRIESCHE S J, ZHANG C, *et al.* FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism[J]. *Cell*, 2011, 146(2): 247–261.
- [21] HATA Y, BUTZ S, SÜDHOF T C. CASK: a novel dlg/PSD95 homolog with an N-terminal calmodulin-dependent protein kinase domain identified by interaction with neurexins[J]. *The Journal of Neuroscience*, 1996, 16(8): 2488–2494.
- [22] VAROQUEAUX F, ARAMUNI G, RAWSON R L, *et al.* Neuroligins determine synapse maturation and function[J]. *Neuron*, 2006, 51(6): 741–754.
- [23] YOO H J, CHO I H, PARK M, *et al.* Family based association of *GRIN2A* and *GRIN2B* with Korean autism spectrum disorders[J]. *Neuroscience Letters*, 2012, 512(2): 89–93.
- [24] PAN Y, CHEN J, GUO H, *et al.* Association of genetic variants of *GRIN2B* with autism[J]. *Scientific Reports*, 2015, 5: 8296.
- [25] HUANG T N, HSUEH Y P. Calcium/calmodulin-dependent serine protein kinase (CASK), a protein implicated in mental retardation and autism-spectrum disorders, interacts with T-Brain-1 (TBR1) to control extinction of associative memory in male mice[J]. *Journal of Psychiatry and Neuroscience*, 2017, 42(1): 37–47.
- [26] BUREAU I, SHEPHERD G M G, SVOBODA K. Circuit and plasticity defects in the developing somatosensory cortex of *Fmr1* knock-out mice[J]. *The Journal of Neuroscience*, 2008, 28(20): 5178–5188.
- [27] GIBSON J R, BARTLEY A F, HAYS S A, *et al.* Imbalance of neocortical excitation and inhibition and altered UP states reflect network hyperexcitability in the mouse model of fragile X syndrome[J]. *Journal of Neurophysiology*, 2008, 100(5): 2615–2626.
- [28] GONÇALVES J T, ANSTEY J E, GOLSHANI P, *et al.* Circuit level defects in the developing neocortex of fragile X mice[J]. *Nature Neuroscience*, 2013, 16(7): 903–909.
- [29] DÖLEN G, OSTERWEIL E, RAO B S, *et al.* Correction of fragile X syndrome in mice[J]. *Neuron*, 2007, 56(6): 955–962.
- [30] BIEDERER T, SÜDHOF T C. Mints as adaptors. Direct binding to neurexins and recruitment of Munc18[J]. *The Journal of Biological Chemistry*, 2000, 275(51): 39803–39806.
- [31] VERHAGE M, MAIA A S, PLOMP J J, *et al.* Synaptic assembly of the brain in the absence of neurotransmitter secretion[J]. *Science*, 2000, 287(5454): 864–869.
- [32] FAN H P, FAN F J, BAO L, *et al.* SNAP-25/syntaxin 1A complex functionally modulates neurotransmitter gamma-aminobutyric acid reuptake[J]. *The Journal of Biological Chemistry*, 2006, 281(38): 28174–28184.
- [33] SHEN C, RATHORE S S, YU H, *et al.* The trans-SNARE-regulating function of Munc18-1 is essential to synaptic exocytosis[J]. *Nature Communications*, 2015, 6: 8852.
- [34] ZHOU L, LI F, XU H B, *et al.* Treatment of cerebral ischemia by disrupting ischemia-induced interaction of nNOS with PSD-95[J]. *Nature Medicine*, 2010, 16(12): 1439–1443.