

重度抑郁症多脑区基因表达谱分析

乔婧^①, 邱江^①, 李迪康^②, 熊清^{②*}

① 认知与人格教育部重点实验室, 西南大学心理学部, 重庆 400715;

② 西南大学计算机与信息科学学院, 重庆 400715

* 联系人, E-mail: qingx@swu.edu.cn

2014-08-29 收稿, 2014-10-21 接受, 2014-12-19 网络版发表

西南大学心理学部重点学科建设项目(SWU: TR201208-3)和家蚕基因组生物学国家重点实验室开放基金(sklsb2013005)资助

摘要 重度抑郁症(major depressive disorder, MDD)是一种环境与遗传因素共同作用导致多个基因功能失调的复杂疾病, 现正在严重危害全球人类的身心健康. 全基因组基因表达分析能检测全基因组基因的时空表达模式, 正被广泛用于精神类疾病的研究. 然而, 由于MDD患者不同脑区可能存在不同的活化模式, 以及该疾病可能受多个中效或微效基因的调控, 传统的单脑区单基因分析方法难以有效揭示其发病的生物学机制. 本研究采用多脑区以及基因集合的分析模式, 搜集了前人研究中13个来源于不同脑区的、包含患者与对照组脑转录组数据的数据集, 对其进行基因集富集分析(gene set enrichment analysis). 结果显示神经元分子标记基因集在纹状体、前扣带皮层和杏仁核中显著下调, 这与目前研究者对MDD分子机理的理解较为一致. 我们的分析也显示MDD患者脑中胶质细胞的分子标记物活性出现异常, 但是其在不同脑区与同一脑区不同数据集间却呈现复杂的变化模式. 一个有趣的发现是, 星形胶质细胞分子标记基因集与少突胶质细胞分子标记基因集在MDD患者前扣带皮层脑区的活性在男女样本中呈现相反的变化趋势, 在杏仁核也存在活性变化的性别差异. 在MDD与双相情感障碍和精神分裂症的比较分析中, 我们发现与MDD类似, 神经元分子标记基因集的活性在双相情感障碍与精神分裂症患者的多个脑区也呈现一致性下调, 而在这2种疾病中胶质细胞分子标记基因集与脉络丛分子标记基因集的活性则与抑郁症存在差异. 总体来看, 本研究基于对大数据的分析, 从全基因组角度对抑郁症的分子机理进行了一些探索, 鉴别了一些在抑郁症患者中分子活性反常的脑区以及可能与这些反常相关联的群体或个体基因表达模式, 为后期的功能研究提供候选分子靶标, 为诠释基因-脑-行为的关系奠定基础.

关键词

抑郁症
基因表达分析
基因集富集分析
神经元
胶质细胞

临床上的重度抑郁症属于心境障碍(mood disorder)的一种, 是一类病因复杂, 具有情绪障碍、认知损伤、行为改变等症状的精神类疾病, 诊断标准中的典型症状有持续情绪低落伴随快感缺失、睡眠障碍、运动迟缓、注意力下降、出现自杀念头等^[1]. 重度抑郁症正在严重危害人类的身心健康. 流行病学数据揭示了MDD在西方国家极高的终身患病率(约为

16%)^[2], 而在中国, 2009年发表在Lancet上的调查数据显示其4省人群患病率也高达约6%^[3]. 重度抑郁症具有相当大的患者群体, 加之约为4%的高自杀率^[4], 治疗费用和自杀死亡引发的相关费用等带来的经济负担^[5], 以及估算37%的遗传力^[6], 因此, 了解抑郁症的病理生理学机制、探索更有效的预防和诊断治疗手段, 对于个人和社会而言都意义深远.

引用格式: 乔婧, 邱江, 李迪康, 等. 重度抑郁症多脑区基因表达谱分析. 科学通报, 2015, 60: 1010–1021

Qiao J, Qiu J, Li D K, et al. Gene expression analysis of multiple brain regions in major depressive disorder (in Chinese). Chin Sci Bull, 2015, 60: 1010–1021, doi: 10.1360/N972014-00909

尽管有相当多的研究都在尝试揭示MDD的神经生物学机制,但阐明其发病的分子机理却一直是一个难题.抑郁症是受基因与环境联合调控的复杂疾病,环境是介导抑郁发生的重要因素,环境可能导致基因的时空表达模式发生改变,这种改变可以通过微阵列(microarray)技术来检测.因此,对于受遗传因素影响较弱的精神类疾病,基于微阵列或新一代测序技术(next-generation sequencing, NGS)的全基因组基因表达分析(genome-wide gene expression analysis)在其功能基因的鉴别中具有优势^[7-9].已有一些研究团队应用全基因组基因表达分析,在MDD患者尸检脑组织或抑郁症动物模型中,发现了可能与疾病相关的基因和信号通路异常,比如与少突胶质细胞(oligodendrocyte)功能相关的基因^[10]、谷氨酸(glutamate)和GABA神经传递相关基因^[11,12]、精胺/亚精胺N¹-乙酰转移酶(spermine/spermidine N¹-acetyltransferase gene, SSAT)基因^[13]、BDNF(brain-derived neurotrophic factor)/GABA相关基因^[14,15]和*FGFR1* (fibroblast growth factor receptor 1), *NCAM1* (neural cell adhesion molecule 1), *CAMK2A* (calcium/calmodulin-dependent protein kinase II alpha)^[16]等.但是通过全基因组基因表达分析发现的MDD相关基因和通路仍然很有限,面对MDD的复杂性,此领域还有太多问题急待解决.

以往全基因组基因表达分析研究的发现多来源于单个脑区的分析,再加之可能存在一些问题,包括大脑样本的pH有个体差异、PMI (postmortem interval) 时间长短不一、患者用药史不同、样本量太少等诸多原因,都有可能影响结果的可靠度与准确性^[17,18].并且,传统的分析模式孤立地研究每个基因的生物学效应,通常基因表达谱数据的噪声很高,因而为了降低假阳性率需要采用非常严格的显著性水平,这导致大量中效与微效基因的贡献被忽略,而调控复杂疾病如抑郁症的基因中有大量的中效与微效基因.因此,为了尽可能地消除这些不利因素对研究结果的影响,本研究搜集了前人研究中共13个包含患者与对照组脑转录组数据的数据集,涉及抑郁症研究中的多个热点脑区,如前额叶、前扣带皮层、杏仁核、海马等额叶与边缘系统脑区,采用了多脑区以及基因集合的分析模式,以期能更准确地鉴别MDD患者基因表达程序中的异常,为疾病的诊断与治疗提供线索与依据.本研究的优势在于:(1) 样本量增加能

降低样本异质性带来的干扰;(2) 多脑区协同分析能更全面地鉴别与抑郁症相关联的基因表达模式,同时也能更好地揭示脑区之间的相互联系;(3) 基因集合分析能检测多个基因的叠加生物学效应,捕获中效和微效基因与抑郁症的关联,从而更全面地揭示疾病的本质.

1 材料和方法

1.1 数据集

GEO (gene expression omnibus, <http://www.ncbi.nlm.nih.gov/geo/>) 是目前最大且可公开获取的基因表达数据库,其中包含丰富的关于人类疾病的基因表达谱信息.我们用“depression”和“depressive disorder”作为关键词搜索了2014年7月以前GEO数据库中所有跟抑郁症有关的基因表达数据集,从中去除掉非人类、非脑组织、非病例-对照实验设计以及非重度抑郁症的数据集,最后得到13个人类重度抑郁症的病例-对照脑基因表达数据集.这13个数据集集中的2个(GSE12654和GSE53987)来源于与双相情感障碍(bipolar disorder)与精神分裂症(schizophrenia)的对比研究,所以我们的分析也包含了这些对应的双相情感障碍与精神分裂症的数据集.此外, GSE53987包含来源于海马(hpc),纹状体(str)和背外侧前额叶(pfc)3个脑区的数据,所以将其细分成了3个子数据集: GSE53987_hpc, GSE53987_str和GSE53987_pfc.因此,本研究共包含15个重度抑郁症数据集、4个双相情感障碍数据集和4个精神分裂症数据集,这些数据集所对应的GEO网址和参考文献如表1所示.

1.2 基因集富集分析

基因集富集分析是指通过对全基因组基因表达谱数据的分析,挖掘与复杂疾病相关的生物学通路或基因集合,是解析复杂疾病分子调控机制的主要方法之一.复杂疾病如抑郁症通常受少数生物学效应较大的主效基因与众多效应较小的中效与微效基因的控制^[23].传统的单基因分析模式存在的一个主要缺陷是,该方法仅能检测一部分主效基因,且这些主效基因只能解释一小部分疾病表型变异^[24];通路关联分析能够捕获由多个中效与微效基因产生的联合生物学效应,因而成为解析复杂疾病调控通路与挖掘功能基因的理想方法.

表 1 基因表达数据集信息^{a)}

Table 1 Information of data sets

数据集编号	数据集	GEO链接	参考文献
D1	GSE12654_mdd	http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE12654	[19]
D2	GSE53987_hpc_mdd	http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE53987	—
D3	GSE53987_str_mdd	http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE53987	—
D4	GSE54562	http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE54562	[20]
D5	GSE54563	http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE54563	[20]
D6	GSE54571	http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE54571	[20]
D7	GSE54572	http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE54572	[20]
D8	GSE54565	http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE54565	[21]
D9	GSE54564	http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE54564	[20]
D10	GSE54566	http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE54566	[21]
D11	GSE53987_pfc_mdd	http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE53987	—
D12	GSE54567	http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE54567	[20]
D13	GSE54568	http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE54568	[20]
D14	GSE54570	http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE54570	[22]
D15	GSE54575	http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE54575	[22]
D16	GSE12654_bipolar	http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE12654	[19]
D17	GSE53987_hpc_bipolar	http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE53987	—
D18	GSE53987_str_bipolar	http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE53987	—
D19	GSE53987_pfc_bipolar	http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE53987	—
D20	GSE12654_schiz	http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE12654	[19]
D21	GSE53987_hpc_schiz	http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE53987	—
D22	GSE53987_str_schiz	http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE53987	—
D23	GSE53987_pfc_schiz	http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE53987	—

a) “—”表示文章信息暂无

本研究采用GSEA^[25](gene set enrichment analysis: <http://www.broadinstitute.org/gsea/index.jsp>)程序进行基因集富集分析, GSEA已经被广泛用于microarray基因表达数据的基因集富集分析研究中^[26]. GSEA先对病例组与对照组基因表达谱数据进行单基因差异表达分析, 并根据差异表达程度对表达谱数据集中的基因进行排序, 然后用Kolmogorov-Smirnov统计测试检验基因集中的基因是否富集在这个排序的顶端或底端. 本研究所使用的2个表型类分别为病例与对照, 病例为第一表型类而对照为第二表型类. 因此, 如果基因集中的基因富集于排序的顶端或底端, 则表明这个基因集在疾病组中的上调(富集于顶端)或下调(富集于底端)与疾病的产生或发展有关联. 这里所使用的基因集是预先定义好的, 通常来源于先前的实验结果或基因功能注释, 每个基因集代表一个生物学通路或由一群具有相关生物学效应的基因组

成. 通过这种方式, 我们可以并行测试大量的基因集与疾病之间的关联效应, 从而挖掘在疾病的产生或发展中起作用的基因或基因集合.

GSEA提供了多个统计量用于单个基因差异表达程度的计算, 本研究使用最为通用的Signal2Noise来计算基因的差异表达分值. Signal2Noise统计量除了能够度量差异表达的强度外, 还能度量其方向, 如果基因的差异表达分值为正值则表示该基因的表达与疾病正相关; 反之则负相关. 本研究所使用的基因集来源于MSigDB(molecular signatures database: <http://www.broadinstitute.org/gsea/msigdb/index.jsp>)的c2数据集——c2.all.v4.0.symbols.gmt. c2中的基因集来源于通路数据库、PubMed中的文献以及领域专家的知识, 其4.0版本共包含4722个精选的基因集, 我们在分析中去掉了那些在被分析的表达谱数据中低于10个或大于100个基因的基因集, 以避免这些基因集所

代表的生物学功能过于宽泛或狭窄. Gene permutation test用于评估基因集富集的显著性(significance), 本研究中的显著性阈值设置为GSEA的推荐值0.25, 即FDR(false discovery rate)小于0.25的基因集被视为与抑郁症有显著关联.

2 结果

2.1 不同脑细胞类型分子标记物基因集活性比较分析

Lein等人^[27]绘制了成年小鼠的全脑基因表达谱. 他们把脑细胞分为4个类别: 神经元(neuron)、星形胶质细胞(astrocyte)、少突胶质细胞(oligodendrocytes)和脉络丛(choroid plexus), 并选取*Snrpn*, *Gja1*, *Cnp1*,

*Col8a2*这4个基因分别为这些细胞类别的种子基因. 通过对Allen brain atlas数据集中的其他基因与这4个种子基因进行关联分析以及专家审定, 鉴别出这4种脑细胞类型特异性的分子标记物(biomarker). 在c2基因集数据集中, 这些小鼠脑细胞特定类别的分子标记物对应的人类基因分别由如下4个基因集代表: 神经元分子标记基因集(LEIN_NEURON_MARKERS); 星形胶质细胞分子标记基因集(LEIN_ASTROCYTE_MARKERS); 少突胶质细胞分子标记基因集(LEIN_OLIGODENDROCYTE_MARKERS); 脉络丛分子标记基因集(LEIN_CHOROID_PLEXUS_MARKERS). 这4个基因集在全基因组基因集富集分析中的结果如表2及表S1~S15所示.

通过对抑郁症组与对照组的基因表达进行比较,

表 2 4 种脑细胞类别分子标记物基因集的基因集富集分析结果^{a)}
Table 2 Results of gene set enrichment analysis for gene sets of markers of four brain cell types

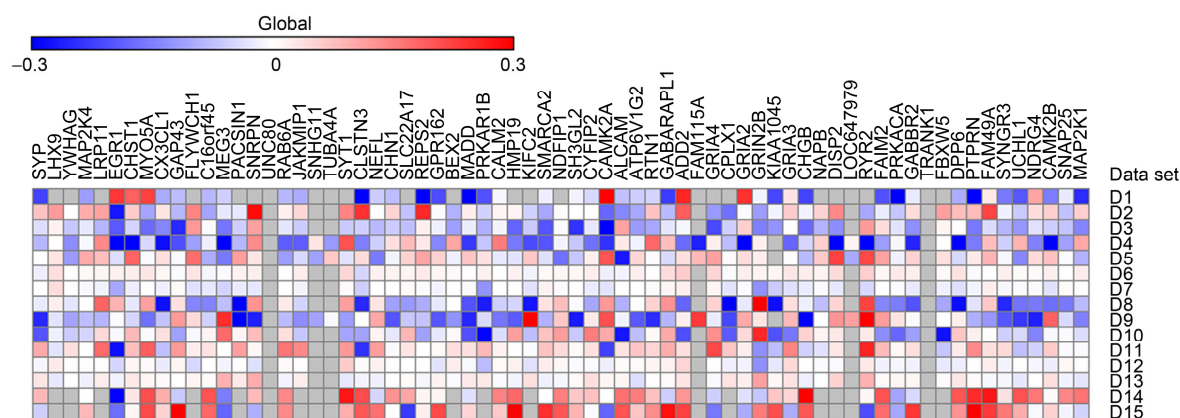
数据集 编号	脑区	实验组 性别	对照组 性别	LEIN_NEURO N_MARKERS	LEIN_ASTRO CYTE_MARK ERS	LEIN_OLIGOD ENDROCYTE_ MARKERS	LEIN_CHOROI D_PLEXUS_M ARKERS	显著上调 基因总数	显著下调 基因总数
D1	前额叶 (BA10)	6M:5F	9M:6F	—	↑	—	—	170	3
D2	海马 (hippocampus)	9M:8F	9M:9F	—	↓	—	—	0	340
D3	纹状体 (striatum)	10M:6F	10M:8F	↓	—	↑	—	218	26
D4	前扣带皮层 (ACC)	10M	10M	↓	↑	↑	—	1130	322
D5	前扣带皮层 (ACC)	25F	25F	—	—	↓	—	248	353
D6	前扣带皮层 (ACC)	13F	13F	—	↓	↓	↓	0	900
D7	前扣带皮 (ACC)	12M	12M	—	↑	↑	—	34	0
D8	前扣带皮层 (ACC)	16M	16M	—	—	—	—	8	11
D9	杏仁核 (amygdala)	21F	21F	↓	↑	↑	↑	890	334
D10	杏仁核 (amygdala)	14M	14M	—	↑	↓	—	6	5
D11	背外侧前额叶 (BA46)	9M:8F	10M:9F	—	↓	—	—	0	75
D12	背外侧前额叶 (BA9)	14M	14M	—	—	—	—	0	0
D13	背外侧前额叶 (BA9)	15F	15F	—	—	↓	—	2	386
D14	背外侧前额叶 (BA9)	未知	未知	↑	↓	—	—	213	692
D15	腹眶额叶皮层 (BA47)	未知	未知	—	↓	↓	—	0	784

a) ↑, 显著上调; ↓, 显著下调; —, 不显著 (FDR>0.25)

发现神经元分子标记基因集在纹状体(D3)、前扣带皮层(D4)和杏仁核(D9)中显著下调,其FDR值分别为0.05, 0.001和0.06. 对这个基因集中的基因做进一步分析,图1显示神经元分子标记基因集中的基因在15个数据集中的差异表达分值热图(heatmap),热图用GENE-E (<http://www.broadinstitute.org/cancer/software/GENE-E/>)创建;这些基因在每个数据集中具体的差异表达分值参见表S16. 在我们的数据中,一些和神经元结构或功能相关的基因在抑郁症患者中出现了多脑区的下调,比如*MADD*(MAP-kinase activating death domain)、*EGR1* (early growth response 1)和*PRKAR1B* (protein kinase, cAMP-dependent, regulatory, type I, beta). *MADD*能激活MAPK信号级联(MAPK signaling cascade)^[28], MAPK信号级联是调节神经元可塑性、功能与存活的关键通路,它的一个负调节子能够在小鼠与大鼠动物模型中诱发抑郁症状^[29]. *EGR1*与神经元可塑性和短时记忆向长期记忆的转化有关^[30,31],也有报道显示应激诱发的糖皮质激素受体激活会导致MAPK信号级联基因与*EGR1*表达量增加和活性增强,从而可能引起应激有关的恐惧记忆改变^[32];此外,我们发现*RYR2* (ryanodine receptor 2)在抑郁症患者的多个脑区显著上调,*RYR2*是主要富集于心肌与脑中的钙释放通道^[33],*RYR2*的敲除会导致记忆能力受损^[34].

分析显示抑郁症患者脑中胶质细胞的分子标记物活性出现异常,但是其不同脑区与同一脑区不

同数据集间却呈现复杂的变化模式. 一个有趣的发现是,星形胶质细胞与少突胶质细胞分子标记基因集在前扣带皮层脑区的活性在男女样本中呈现相反的变化趋势. D4, D7, D8是来自男性的样本;而D5, D6是来自女性的样本. 星形胶质细胞与少突胶质细胞分子标记基因集在D4与D7中均显著上调;星形胶质细胞分子标记基因集在D6,少突胶质细胞分子标记基因集在D5与D6中显著下调;这2个基因集在D8中均不显著. 对杏仁核而言,星形胶质细胞分子标记基因集在男女样本中均上调(D9, D10),而少突胶质细胞分子标记基因集在女性样本中上调(D9),男性样本中下调(D10). 此外,星形胶质细胞分子标记基因集在海马(D2)、背外侧前额叶(D11, D14)与腹眶额叶皮层(D15)中均下调,在纹状体(D3)中不显著;少突胶质细胞分子标记基因集在海马(D2)中不显著,在背外侧前额叶(D13)与腹眶额叶皮层(D15)中均下调,在纹状体(D3)中上调. 星形胶质细胞是中枢神经系统中数量最多的胶质细胞,参与形成突触的三联复合体结构(tripartite synapse),调节谷氨酸/GABA平衡和神经元的兴奋程度,负责约90%的谷氨酸再摄取和代谢^[35]. 最新的研究揭示,星形胶质细胞中谷氨酸摄取的封阻能诱发抑郁症的核心症状快感缺失(anhedonia)^[36];另一项研究则表明,星形胶质细胞内源性ATP是调节小鼠抑郁行为的一个关键因子^[37]. 少突胶质细胞异常,如少突胶质细胞特异性基因表达量降低或少突胶质细胞细胞数减少,已被报道与



多种精神疾病包括抑郁症有关联^[38,39]。

我们进一步分析了2种胶质细胞在各个数据集中的基因差异表达模式,结果如图2与表S17, S18所示。我们发现很多基因在前扣带皮层与杏仁核脑区的胶质细胞尤其是少突胶质细胞中的表达呈现性别差异,如*GPR37*(G protein-coupled receptor 37), *UGT8* (UDP glycosyltransferase 8)和*MAG* (myelin associated glycoprotein)。

2.2 不同数据集基因差异表达的比较分析

为了检测基因的差异表达在数据集间的相似度,我们提取了每个数据集中差异表达程度最高的200个在病例组上调的基因,然后对所有数据集的这200个基因进行两两比较,计算相同基因的个数,结果如图3所示。从图3中可以看出:总的来说,数据集之间差异表达的基因存在较大的差异,相似度最高的2个数据集间也只有36个相同基因,占18%,有些数据集之间甚至一个相同基因也没有。这些结果可能由如下

因素导致:(1) 抑郁症受一些微效基因的影响,用单基因分析方法很难检测到其效应;(2) 多种基因的表达失调可能导致同样的抑郁表型;(3) 数据集本身的异质性。

2.3 抑郁症候选基因活性比较分析

全基因组关联研究(genome-wide association study, GWAS)能通过对2组样本之间基因型的比对找出与疾病相关联的SNP(single nucleotide polymorphism)位点,已被广泛用于疾病遗传机理的研究。近年来,研究人员已经开展了多个大规模的重度抑郁症GWAS项目,鉴别了大约200个抑郁症候选基因^[40]。由于这些GWAS结果之间通常存在较大差异,研究人员试图通过元分析来消除这些不一致,Flint和Kenneth等人^[40]列出了7个在元分析中呈显著相关($P<0.05$)的候选基因:*SLC6A4*(solute carrier family 6(neurotransmitter transporter), member 4), *APOE* (apolipoprotein E), *DRD4* (dopamine receptor D4), *GNB3* (guanine nucleotide

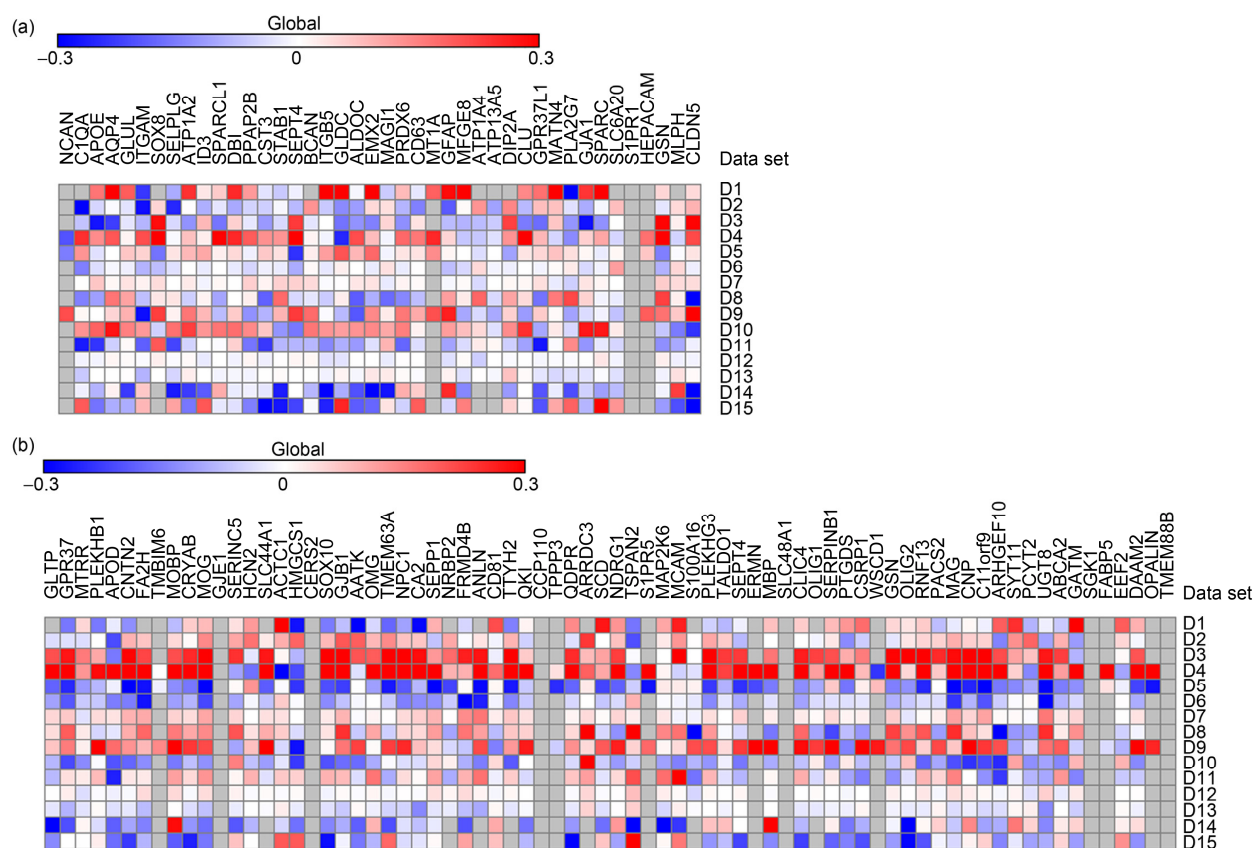


图2 2种胶质细胞分子标记基因集中的基因在15个数据集上的差异表达热图。(a) 星形胶质细胞; (b) 少突胶质细胞

Figure 2 Differential expression heatmap of genes in the two glial cell markers gene sets across 15 data sets. (a) Astrocyte; (b) oligodendrocyte

	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12	D13	D14
D2	3													
D3	2	6												
D4	3	0	36											
D5	4	2	0	2										
D6	1	5	0	1	5									
D7	2	5	17	19	1	8								
D8	5	3	5	2	3	3	1							
D9	5	1	13	14	6	0	1	4						
D10	5	4	0	2	2	2	2	12	1					
D11	2	5	9	0	4	10	4	4	1	3				
D12	2	6	6	5	3	6	21	14	0	2	5			
D13	7	6	4	1	7	14	9	5	5	7	11	10		
D14	2	1	2	3	2	3	2	4	4	3	1	5	4	
D15	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12	D13	D14

图3 15个数据集200个差异表达程度最高的基因之间的相同基因数

Figure 3 The number of common genes in top 200 most differentially expressed genes between 15 data sets

binding protein (G protein), beta polypeptide 3), *HTR1A* (5-hydroxytryptamine (serotonin) receptor 1A, G protein-coupled), *MTHFR* (methylenetetrahydrofolate reductase (NAD(P)H)) 和 *SLC6A3* (solute carrier family 6 (neurotransmitter transporter), member 3). 我们提取了这些基因在各个数据集中的差异表达分值, 如表S19所示, 基于这些分值绘制的热图如图4所示, 这些基因在不同脑区或同一脑区不同来源的样本中呈现复杂的差异表达模式, 如 *APOE* 在纹状体 (D3)、背外侧前额叶 (D11, D14) 和腹眶额叶皮层 (D15) 中下调, 在杏仁核 (D10) 中上调; *MTHFR* 在女性杏仁核样本 (D9) 中上调, 在男性杏仁核样本 (D10) 中下调. 胡海岚研究组的 Li 等人^[41] 的最新研究表明 *CAMK2B* (calcium/calmodulin-dependent protein kinase II beta) 在外侧缰核中介导抑郁症的核心症状. 外侧缰核 (LHb) 被认为是涉及抑郁症病理的关键脑区, 神经影像学以及抑郁动物模型均表明在抑郁状态下外侧缰核过度兴奋^[42,43], 胡海岚研究组^[41] 以雄性大鼠与雄性小鼠为实验对象, 首次确定了 *CAMK2B* 是导致这些动物外侧缰核过度兴奋和抑郁表型的关键分子. 我们也分析了 *CAMK2B* 在15个人类数据集中的差异表达情况, 发现在5个只包含男性样本的数据集 (D4, D7, D8, D10, D12) 中, *CAMK2B* 基因的表达水平均有不同程度的下调, 而在4个只包含女性样本的数据集 (D5, D6, D9, D13) 中 *CAMK2B* 均有不同程度的上调, 这表明 *CAMK2B* 基因在这些MDD患者前扣带皮层、杏仁核和背外侧前额叶脑区的表达具有性别差异, 同时, 我们发现 *MTHFR* 基因的表达也存在类似的性别差异.

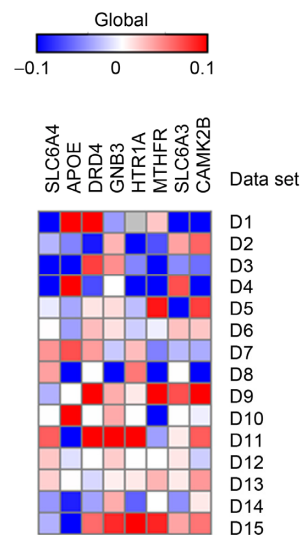


图4 抑郁症候选基因在15个数据集中的差异表达热图

Figure 4 Differential expression heatmap of candidate genes in depressive disorder across 15 data sets

2.4 抑郁症与其他精神疾病之间的比较分析

本研究包含4个双相情感障碍数据集和4个精神分裂症数据集, 检测了神经元分子标记基因集、星形胶质细胞分子标记基因集、少突胶质细胞分子标记基因集和脉络丛分子标记基因集在这些数据集中的活性情况, 结果如表3与表S20~S27所示. 通过与对应的重度抑郁症数据比较, 我们发现神经元分子标记基因集的活性在3种疾病的多个脑区中一致下调, 脉络丛分子标记基因集则一致上调, 相比而言胶质细胞则呈现比较复杂的情况. 星形胶质细胞分子标记基因集在精神分裂症的海马、背外侧前额叶与纹状体

表 3 4 种脑细胞类别分子标记物基因集活性在 3 种精神疾病中的比较分析^{a)}
Table 3 Results of gene set enrichment analysis for gene sets of markers of four brain cell types in three mental disorders

脑区	LEIN_NEURON _MARKERS	LEIN_ASTROCYTE_ MARKERS	LEIN_OLIGODENDRO CYTE_MARKERS	LEIN_CHOROID_PLEX US_MARKERS	显著上调 基因总数	显著下调 基因总数
重度抑郁症						
D1: 前额叶	—	↑	—	—	170	3
D2: 海马	—	↓	—	—	0	340
D3: 纹状体	↓	—	↑	—	218	26
D11: 背外侧前额叶	—	↓	—	—	0	75
双相情感障碍						
D16: 前额叶	↓	↑	—	—	217	3
D17: 海马	↓	—	—	↑	508	274
D18: 纹状体	—	—	—	—	0	0
D19: 背外侧前额叶	—	—	↑	—	915	115
精神分裂症						
D20: 前额叶	—	—	↓	—	4	4
D21: 海马	↓	↑	—	↑	637	474
D22: 纹状体	↓	↑	↓	↑	852	439
D23: 背外侧前额叶	↓	↑	↑	↑	1310	178

a) ↑, 显著上调; ↓, 显著下调; —, 不显著(FDR>0.25)

中活性上调, 在双相情感障碍与重度抑郁症的前额叶中上调, 在重度抑郁症的海马与背外侧前额叶中下调. 少突胶质细胞分子标记基因集在精神分裂症与双相情感障碍的背外侧前额叶中上调, 在精神分裂症的前额叶与纹状体中下调.

3 讨论

本研究分析比较了MDD患者和健康对照组前额叶和边缘系统多个脑区的基因差异表达模式, 这些脑区在以往研究中都认为与MDD的神经生物学机制是相关的. 分析结果揭示, 神经元分子标记基因集在MDD患者的前扣带皮层、杏仁核和纹状体中显著下调. 前扣带皮层和杏仁核作为大脑边缘系统的重要组成部分, 与情绪加工有着密切关系. 我们的结果中MDD患者前扣带皮层一些与神经元结构或功能相关基因的下调, 可能预示着神经元可塑性、存活等发生改变, 从而影响其情绪控制与加工, 与前人研究中发现抑郁症患者的前扣带皮层白质纤维完整性降低^[44]、额叶神经元形态变小^[45]、额叶-纹状体环路(fronto-striatal)正性情绪维持异常^[46]这些结果相似.

相比较而言, 胶质细胞分子标记基因集则呈现更为复杂的变化模式, 一个有趣的发现是, 星形胶质细胞分子标记基因集与少突胶质细胞分子标记基因

集在MDD患者前扣带皮层脑区的活性在男女样本中呈现相反的变化趋势, 在杏仁核也存在活性变化的性别差异. 此外, 神经元分子标记基因集的*CAMK2B*能在外侧缰核中介导抑郁症的核心症状^[41], 其表达模式在我们的数据中也呈现明显的性别差异, 这些结果表明未来的试验设计与分析或许要将性别作为一个重要考虑的因素.

在MDD与双相情感障碍和精神分裂症的比较分析中, 我们发现神经元分子标记基因集在3种精神疾病患者的多个脑区中存在一致性的活性下调, 显示神经元功能的失调, 这与以往在人类精神类疾病中发现的神经元突触缺陷和转录异常的研究结果相类似^[47]. 脉络丛分子标记基因集在精神分裂症的多脑区和双相情感障碍的海马均表现出一致性上调. 脉络丛为产生脑脊髓液(cerebrospinal fluid)的主要结构^[48]. 已有研究显示, 神经发育中脉络丛基因表达失调导致的脑脊髓液中组分浓度的改变可能是导致精神分裂症以及其他一些神经退行性疾病的风险因子^[49]. 本研究中发现的脉络丛分子标记基因集活性的显著增加也预示着脉络丛基因表达的异常可能与严重的精神类疾病的发生相关. 此外, 胶质细胞的活性似乎在精神分裂症特定脑区中存在与MDD相反的模式, 这可能与2类疾病本身的特异性有关, 还需要

后期进一步验证.

除了来源于尸检脑组织的MDD数据集外, GEO中还包含一些来自于血液样本的MDD数据集. MDD是一种心境障碍, 其主要病因源于脑功能的失调, 再加上基因表达是组织特异性的, 脑组织与血液组织具有不同的基因表达谱, 因此本研究中我们仅选取了来源于脑组织的数据, 以期能更有效地鉴别与MDD相关联的基因表达模式. 此外, 在本研究中, 我们没有将来自同一脑区的数据集进行整合分析, 而是先单独分

析这些数据集然后再进行整合理解, 以避免数据异质性与样本异质性给分析带来的不利影响, 因为这些数据可能来自于不同的芯片类型或不同的样本类型.

总的来说, 本研究基于对大数据的分析, 从全基因组角度对抑郁症的分子机理进行了一些探索, 鉴别了一些在抑郁症患者中分子活性反常的脑区, 以及可能与这些反常相关联的群体或个体基因表达模式, 为后期的功能研究提供候选分子靶标, 为诠释基因-脑-行为的关系奠定基础.

致谢 感谢西南大学心理学部的张庆林教授、李海江、杜雪, 以及美国马里兰州NIMH/NIH的张喜淋博士后对于本工作的大力支持.

参考文献

- 1 Sullivan P F, Cichon S, Coventry W L, et al. Genome-wide association for major depressive disorder: A possible role for the presynaptic protein piccolo. *Mol Psychiatry*, 2009, 14: 359–375
- 2 Kessler R C, Berglund P, Demler O, et al. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry*, 2005, 62: 593–602
- 3 Phillips M R, Zhang J, Shi Q, et al. Prevalence, treatment, and associated disability of mental disorders in four provinces in China during 2001–05: An epidemiological survey. *Lancet*, 2009, 373: 2041–2053
- 4 Meng C, Brandl F, Tahmasian M, et al. Aberrant topology of striatum's connectivity is associated with the number of episodes in depression. *Brain*, 2014, 137: 598–609
- 5 Greenberg P E, Kessler R C, Birnbaum H G, et al. The economic burden of depression in the United States: How did it change between 1990 and 2000? *J Clin Psychiatry*, 2003, 64: 1465–1475
- 6 Sullivan P F, Neale M C, Kendler K S. Genetic epidemiology of major depression: Review and meta-analysis. *Am J Psychiatry*, 2000, 157: 1552–1562
- 7 Mirnics K A R, Middleton F A, Lewis D A, et al. Analysis of complex brain disorders with gene expression microarrays: Schizophrenia as a disease of the synapse. *Trends Neurosci*, 2001, 24: 479–486
- 8 Bunney W E, Bunney B G, Vawter M P, et al. Microarray technology: A review of new strategies to discover candidate vulnerability genes in psychiatric disorders. *Am J Psych*, 2003, 160: 657–666
- 9 Iwamoto K, Tadaumi K. Gene expression profiling in schizophrenia and related mental disorders. *Neuroscientist*, 2006, 12: 349–361
- 10 Aston C, Jiang L, Sokolov B P. Transcriptional profiling reveals evidence for signaling and oligodendroglial abnormalities in the temporal cortex from patients with major depressive disorder. *Mol Psychiatry*, 2005, 10: 309–322
- 11 Choudary P V, Molnar M, Evans S J, et al. Altered cortical glutamatergic and GABAergic signal transmission with glial involvement in depression. *Proc Natl Acad Sci USA*, 2005, 102: 15653–15658
- 12 Klempan T A, Sequeira A, Canetti L, et al. Altered expression of genes involved in ATP biosynthesis and GABAergic neurotransmission in the ventral prefrontal cortex of suicides with and without major depression. *Mol Psychiatry*, 2009, 14: 175–189
- 13 Sequeira A, Gwady F G, Canetti L, et al. Implication of SSAT by gene expression and genetic variation in suicide and major depression. *Arch Gen Psychiatry*, 2006, 63: 35–48
- 14 Tripp A, Oh H, Guilloix J, et al. Brain-derived neurotrophic factor signaling and subgenual anterior cingulate cortex dysfunction in major depressive disorder. *Am J Psych*, 2012, 169: 1194–1202
- 15 Guilloix J P, Douillard-Guilloix G, Kota R, et al. Molecular evidence for BDNF- and GABA-related dysfunctions in the amygdala of female subjects with major depression. *Mol Psychiatry*, 2012, 17: 1130–1142
- 16 Tochigi M, Iwamoto K, Bundo M, et al. Gene expression profiling of major depression and suicide in the prefrontal cortex of postmortem brains. *Neurosci Res*, 2008, 60: 184–191

- 17 Lewis D A. The human brain revisited: Opportunities and challenges in postmortem studies of psychiatric disorders. *Neuropsychopharmacology*, 2002, 26: 143–154
- 18 Mirnics K A R, Levitt P, Lewis D A. DNA microarray analysis of postmortem brain tissue. *Int Rev Neurobiol*, 2004, 60: 153–181
- 19 Iwamoto K, Kakiuchi C, Bundo M, et al. Molecular characterization of bipolar disorder by comparing gene expression profiles of post-mortem brains of major mental disorders. *Mol Psychiatry*, 2004, 9: 406–416
- 20 Chang L, Jamain S, Lin C, et al. A conserved BDNF, glutamate- and GABA-enriched gene module related to human depression identified by coexpression meta-analysis and DNA variant genome-wide association studies. *PLoS One*, 2014, 9: e90980
- 21 Sibille E, Wang Y, Joeyen-Waldorf J, et al. A molecular signature of depression in the amygdala. *Am J Psych*, 2009, 166: 1011–1024
- 22 Sibille E, Arango V, Galfalvy H C, et al. Gene expression profiling of depression and suicide in human prefrontal cortex. *Neuropsychopharmacology*, 2004, 29: 351–361
- 23 Wang W Y, Barratt B J, Clayton D G, et al. Genome-wide association studies: Theoretical and practical concerns. *Nat Rev Genet*, 2005, 6: 109–118
- 24 Maher B. Personal genomes: The case of the missing heritability. *Nature*, 2008, 456: 18–21
- 25 Subramanian A, Tamayo P, Mootha V K, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA*, 2005, 102: 15545–15550
- 26 Hung J H, Yang T H, Hu Z, et al. Gene set enrichment analysis: Performance evaluation and usage guidelines. *Brief Bioinform*, 2012, 13: 281–291
- 27 Lein E S, Hawrylycz M J, Ao N, et al. Genome-wide atlas of gene expression in the adult mouse brain. *Nature*, 2007, 445: 168–176
- 28 Schievella A R, Chen J H, Graham J R, et al. MADD, a novel death domain protein that interacts with the type 1 tumor necrosis factor receptor and activates mitogen-activated protein kinase. *J Biol Chem*, 1997, 272: 12069–12075
- 29 Duric V, Banasr M, Licznarski P, et al. A negative regulator of MAP kinase causes depressive behavior. *Nat Med*, 2010, 16: 1328–1332
- 30 Lee J L, Everitt B J, Thomas K L. Independent cellular processes for hippocampal memory consolidation and reconsolidation. *Science*, 2004, 304: 839–843
- 31 Bozon B, Davis S, Laroche S. A requirement for the immediate early gene *zif268* in reconsolidation of recognition memory after retrieval. *Neuron*, 2003, 40: 695–701
- 32 Revest J M, Di Blasi F, Kitchener P, et al. The MAPK pathway and Egr-1 mediate stress-related behavioral effects of glucocorticoids. *Nat Neurosci*, 2005, 8: 664–672
- 33 Rossi D, Sorrentino V. The ryanodine receptors gene family: Expression and functional meaning. *Basic Appl Myol*, 2004, 14: 323–343
- 34 Galeotti N, Quattrone A, Vivoli E, et al. Different involvement of type 1, 2, and 3 ryanodine receptors in memory processes. *Learn Mem*, 2008, 15: 315–323
- 35 Smialowska M, Szewczyk B, Wozniak M, et al. Glial degeneration as a model of depression. *Pharmacol Rep*, 2013, 65: 1572–1579
- 36 John C S, Smith K L, Van't Veer A, et al. Blockade of astrocytic glutamate uptake in the prefrontal cortex induces anhedonia. *Neuropsychopharmacology*, 2012, 37: 2467–2475
- 37 Cao X, Li L P, Wang Q, et al. Astrocyte-derived ATP modulates depressive-like behaviors. *Nat Med*, 2013, 19: 773–777
- 38 Edgar N, Sibille E. A putative functional role for oligodendrocytes in mood regulation. *Transl Psychiatry*, 2012, 2: e109
- 39 Tkachev D, Bahn S, Mimmack M L, et al. Oligodendrocyte dysfunction in schizophrenia and bipolar disorder. *Lancet*, 2003, 362: 798–805
- 40 Flint J, Kenneth K. The genetics of major depression. *Neuron*, 2014, 81: 484
- 41 Li K, Zhou T, Liao L, et al. β CaMKII in lateral habenula mediates core symptoms of depression. *Science*, 2013, 341: 1016–1020
- 42 Caldecott-Hazard S, Mazziotta J, Phelps M. Cerebral correlates of depressed behavior in rats, visualized using ^{14}C -2-deoxyglucose autoradiography. *J Neurosci*, 1988, 8: 1951–1961
- 43 Li B, Piriz J, Mirrione M, et al. Synaptic potentiation onto habenula neurons in the learned helplessness model of depression. *Nature*, 2011, 470: 535–539
- 44 Bae J N, Macfall J R, Krishnan K R, et al. Dorsolateral prefrontal cortex and anterior cingulate cortex white matter alterations in late-life depression. *Biol Psychiatry*, 2006, 60: 1356–1363
- 45 Cotter D, Mackay D, Landau S, et al. Reduced glial cell density and neuronal size in the anterior cingulate cortex in major depressive disorder. *Arch Gen Psychiatry*, 2001, 58: 545–553
- 46 Heller A S, Johnstone T, Shackman A J, et al. Reduced capacity to sustain positive emotion in major depression reflects diminished maintenance of fronto-striatal brain activation. *Proc Natl Acad Sci USA*, 2009, 106: 22445–22450
- 47 Wen Z, Nguyen H N, Guo Z, et al. Synaptic dysregulation in a human iPSC cell model of mental disorders. *Nature*, 2014, 515: 414–418
- 48 Damkier H H, Brown P D, Praetorius J. Cerebrospinal fluid secretion by the choroid plexus. *Physiol Rev*, 2013, 93: 1847–1892

49 Palha J A, Santos N C, Marques F, et al. Do genes and environment meet to regulate cerebrospinal fluid dynamics? Relevance for schizophrenia. *Front Cell Neurosci*, 2012, 6: 31

Gene expression analysis of multiple brain regions in major depressive disorder

QIAO Jing¹, QIU Jiang¹, LI DiKang² & XIONG Qing²

¹ Key Laboratory of Cognition and Personality (Southwest University), Ministry of Education, School of Psychology, Southwest University, Chongqing 400715, China;

² School of Computer and Information Science, Southwest University, Chongqing 400715, China

Major depressive disorder (MDD), a polygenic disease resulting from complicated interactions between environmental and genetic factors, is becoming an increasingly serious threat to global public health. Genome-wide gene expression analysis is widely used for psychiatric disorders because of its ability to detect spatio-temporal patterns of gene expression. However, it is difficult to elucidate the molecular mechanism of MDD by methods that analyze brain regions and/or genes individually because activation patterns may differ between brain regions and MDD may be regulated by multiple genes with moderate or weak effects. In this study, we used a strategy that leverages information across multiple brain regions and genes. We performed gene set enrichment analysis of transcriptomics data from different brain regions of MDD patients and health controls in 13 datasets from the Gene Expression Omnibus. Our results show that the neuron markers gene set was significantly down-regulated in the associative striatum, anterior cingulate cortex (ACC), and amygdala, which is consistent with current MDD knowledge. The gene expression of glial cell markers was also altered in MDD patients, but there were no consistent patterns across brain regions or datasets from the same region. Moreover, gender-specific activation patterns of the astrocyte markers and oligodendrocyte markers gene sets were observed in the ACC and amygdala. In a comparative analysis of psychiatric disorders, the neuron markers gene set were also significantly down-regulated in multiple brain regions of patients with bipolar disorder and schizophrenia, although the activation patterns of two glial cell markers gene sets and the choroid plexus markers gene set in these two disorders differed from that of MDD. Our results reveal potential candidate brain regions, gene sets, and genes that are altered in MDD.

major depressive disorder, gene expression analysis, gene set enrichment analysis, neuron, glial cell

doi: 10.1360/N972014-00909

补充材料

表 S1 数据集 D1 的基因集富集分析结果

表 S2 数据集 D2 的基因集富集分析结果

表 S3 数据集 D3 的基因集富集分析结果

表 S4 数据集 D4 的基因集富集分析结果

表 S5 数据集 D5 的基因集富集分析结果

表 S6 数据集 D6 的基因集富集分析结果

表 S7 数据集 D7 的基因集富集分析结果

表 S8 数据集 D8 的基因集富集分析结果

表 S9 数据集 D9 的基因集富集分析结果

表 S10 数据集 D10 的基因集富集分析结果

表 S11 数据集 D11 的基因集富集分析结果

表 S12 数据集 D12 的基因集富集分析结果

表 S13 数据集 D13 的基因集富集分析结果

- 表 S14 数据集 D14 的基因集富集分析结果
- 表 S15 数据集 D15 的基因集富集分析结果
- 表 S16 神经元分子标记基因集基因在 15 个数据集集中的差异表达分值
- 表 S17 星形胶质细胞分子标记基因集基因在 15 个数据集集中的差异表达分值
- 表 S18 少突胶质细胞分子标记基因集基因在 15 个数据集集中的差异表达分值
- 表 S19 抑郁症候选基因在 15 个数据集集中的差异表达分值
- 表 S20 数据集 D16 的基因集富集分析结果
- 表 S21 数据集 D17 的基因集富集分析结果
- 表 S22 数据集 D18 的基因集富集分析结果
- 表 S23 数据集 D19 的基因集富集分析结果
- 表 S24 数据集 D20 的基因集富集分析结果
- 表 S25 数据集 D21 的基因集富集分析结果
- 表 S26 数据集 D22 的基因集富集分析结果
- 表 S27 数据集 D23 的基因集富集分析结果

本文以上补充材料见网络版 csb.scichina.com. 补充材料为作者提供的原始数据，作者对其学术质量和内容负责.