

影响人脑胶质瘤细胞发生、增殖关键lncRNA、miRNA的筛选与分析

李明扬, 王仕超, 凌宇, 刘春霞, 曹俊伟, 张焱如*

内蒙古农业大学内蒙古自治区生物制造重点实验室, 内蒙古 呼和浩特
Email: yanru1964@163.com

收稿日期: 2020年12月20日; 录用日期: 2021年1月20日; 发布日期: 2021年1月27日

摘要

脑胶质瘤是一种极度常见颅腔内的原发性恶性肿瘤, 患病人群占据所有颅内肿瘤患者的60%。IV级胶质瘤为最高恶性程度, 也称为胶质母细胞瘤(胶质母细胞瘤, GBM), 约占全部胶质瘤的55%, 且诊断为GBM的患者5年生存率低于6%。许多研究表明, 脑胶质瘤中异常表达的lncRNA不仅调控着与癌细胞增殖、凋亡有关的信号转导通路, 还与病人的预后有关。利用转录组高通量测序方法, 筛选影响脑胶质瘤发生、发展及复发过程中关键通路中的重要基因, 对脑胶质瘤的分子诊断、靶向药物的研发及新治疗方案的选择具有重要的价值。本研究通过RNA-seq和生物信息学分析, 筛选出在脑胶质瘤组织中差异表达的lncRNA 106种, 其中上调表达52种, 下调表达54种。microRNA共计15个, 其中上调表达8种; 下调表达7种。这些差异表达的ncRNA主要参与有机循环化合物代谢过程、胞内运输、细胞蛋白质定位、蛋白质运输、有丝分裂细胞周期、细胞周期调节、DNA代谢过程的正调控、水解酶活性的调节等生物过程、生物调节、应激反应、多细胞有机体过程、细胞成分的组织或生物发生、细胞定位、信号途径、发育过程、生物过程的正调控与负调控。经lncRNA-miRNA-mRNA关联分析, 建立了ceRNAs调控网络, 明确了lncRNA、miRNA与靶标mRNA之间的对应关系。这些关键lncRNA的筛选, 为脑胶质瘤的诊断及治疗提供了新的思路与方向。

关键词

胶质瘤, lncRNA, miRNA, ceRNA网络, 联合分析

Screening and Analysis of Key lncRNA and miRNA Affecting the Occurrence and Proliferation of Human Glioma Cells

Mingyang Li, Shichao Wang, Yu Ling, Chunxia Liu, Junwei Cao, Yanru Zhang*

*通讯作者。

文章引用: 李明扬, 王仕超, 凌宇, 刘春霞, 曹俊伟, 张焱如. 影响人脑胶质瘤细胞发生、增殖关键 lncRNA、miRNA 的筛选与分析[J]. 亚洲肿瘤科病例研究, 2021, 10(1): 1-23. DOI: 10.12677/acrpo.2021.101001

Key Laboratory of Biomanufacturing of Inner Mongolia Autonomous Region, Inner Mongolia Agricultural University, Hohhot Inner Mongolia
Email: yanru1964@163.com

Received: Dec. 20th, 2020; accepted: Jan. 20th, 2021; published: Jan. 27th, 2021

Abstract

Glioma is an extremely common primary malignant tumor in the cranial cavity, which accounts for 60% of all patients with intracranial tumors. Grade IV glioma is the highest degree of malignancy, also called glioblastoma (glioblastoma, GBM), accounting for about 55% of all gliomas, and the 5-year survival rate of patients diagnosed with GBM is lower than 6%. Many studies have shown that the abnormally expressed lncRNA in brain gliomas not only regulates the signal transduction pathways related to cancer cell proliferation and apoptosis, but is also related to the patient's prognosis. Using transcriptome high-throughput sequencing to screen important genes in key pathways that affect the development, development and recurrence of glioma is of great value for molecular diagnosis of glioma, research and development of targeted drugs and selection of new treatment options. In this study, through RNA-seq and bioinformatics analysis, 106 types of lncRNAs that are differentially expressed in glioma tissues were screened, of which 52 were up-regulated and 54 were down-regulated. There are a total of 15 microRNAs, of which 8 are up-regulated and 7 are down-regulated. These differentially expressed ncRNAs are mainly involved in the metabolic process of organic circulating compounds, intracellular transport, cell protein localization, protein transport, mitotic cell cycle, cell cycle regulation, positive regulation of DNA metabolism, regulation of hydrolase activity and other biological processes and biological regulation, Stress response, multicellular organism process, tissue or biogenesis of cell components, cell location, signal pathway, development process, positive and negative regulation of biological processes. After lncRNA-miRNA-mRNA association analysis, a ceRNAs regulatory network was established, and the correspondence between lncRNA, miRNA and target mRNA was clarified. The screening of these key lncRNAs provides new ideas and directions for the diagnosis and treatment of glioma.

Keywords

Glioma, lncRNA, miRNA, ceRNA Network, Conjoint Analysis

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

脑胶质瘤是一种极度常见颅腔内的原发性恶性肿瘤，患病人群占据所有颅内肿瘤患者的 60%。常因神经突触中的实质细胞或间质细胞发生畸变所致[1]。浸润性生长、极高的原位复发率和死亡率是其最明显的三大特征。神经胶质瘤可以导致颅内压增高以及部分神经功能丧失，引起患者语言困难，行动能力异常，癫痫抽搐并伴随持续强烈疼痛感，重症患者将伴随脑疝甚至死亡。依据世界人民卫生服务组织 (World Health Organization, WHO) 公布的分类管理标准[2]，将胶质瘤按组织学类型分为 I~IV 级，其中 I、II 级为低级别系统胶质瘤，III、IV 级为高级别神经胶质瘤。IV 级胶质瘤为最高恶性程度，也称为胶质母

细胞瘤(胶质母细胞瘤, GBM), 约占全部胶质瘤的 55% [3] [4]。目前, GBM 尚无治疗方案, 尽管有深入的治疗研究, 但诊断为 GBM 的患者生存率仍然极低, 5 年生存率低于 6% [5]。近几年的研究表明, 一些相关信号通路网络中关键基因突变是肿瘤的发生原因。异常的信号通路影响到身体组织器官的功能, 进而产生异常的细胞群, 导致肿瘤形成[6] [7]。随着脑胶质瘤分子机制研究的不断深入, 在脑胶质瘤的发生、恶化和复发过程中靶向阻断主要信号通路的关键因子已成为一种具有巨大潜力的新方法[5] [6] [7]。

lncRNA 是一类长度大于 200 bp 的 RNA, 缺少有意义的开放阅读框, 但能以 RNA 的形式在表观遗传修饰、转录水平、转录后水平调控肿瘤的增殖、诱导凋亡、血管生成、侵袭和转移。lncRNA 可通过以下两种形式调控信号转导通路: Cis-lncRNA 一般通过转录干扰或募集染色质修饰复合物进行染色质修饰来调控与其临近的上下游的蛋白编码基因; trans-lncRNA 可以独立于序列互补, 靶向基因位点发挥作用, 除了染色质修饰复合物外, 还可以与转录延伸因子 RNA 聚合酶结合, 从而影响转录。一些具有结合 miRNA 位点的 lncRNA, 可作为竞争性内源性 RNA (competing endogenous RNA, ceRNA)与 miRNA 靶基因竞争性结合 miRNA, 进而调控 miRNA 及 miRNA 靶基因的表达, 从而影响肿瘤增殖、凋亡过程[8]。但 lncRNA 的表达并不稳定, 其表达量会随着肿瘤发展阶段的变化而出现上调或下调表, 而这些异常表达的分子往往与肿瘤的发生与恶化有着密切的联系。大量研究表明, 一些 lncRNA 够调节胶质瘤干细胞特性, 进而导致肿瘤复发和不良预后。

二代测序(Next-generation sequencing, NGS), 也称为大规模平行测序[9], 它可以同时测序数百万个 DNA 片段。与传统的测序方法[10]相比, 具有短时、高效、成本节约的优点。基于二代测序技术的全转录组测序是监测细胞所能转录出来的所有类型 RNA 的总和, 包括 mRNA 和非编码 RNA (non-coding RNA, ncRNA)。目前针对非编码 RNA 的研究主要集中在微小 RNA (microRNAs, miRNA)、长链非编码 RNA (Long non-coding RNA, lncRNA)和环形 RNA (circularRNA, circRNA) [11]。CeRNA (competing endogenous RNAs, 内源竞争 RNA)是一种调控机制, 是对 mRNA 表达调控机制的一种补充。在细胞内, lncRNA, mRNA, circRNA 都能够和 miRNA 进行结合。当细胞内 lncRNA 的表达量下降时, 更多的 miRNA 与 mRNA 结合, 由于 miRNA 对 mRNA 起负控作用, 导致 mRNA 的表达量下降[12]。从 ceRNA 的调控原理可以看出, 在这个调控关系中有两个核心: 1) 两个 RNA 都能够和 miRNA 结合; 2) 两个 RNA 的表达量呈现正相关[13]。因此, 筛选差异表达的 lncRNA 和差异表达的 mRNA, 首先分别分析 lncRNA-mRNA 和 mRNA-miRNA 之间的靶标关系, 再以 miRNA 为桥梁构建 lncRNA-miRNA-mRNA 关系对, 并筛选 lncRNA 和 mRNA 表达量呈现正相关性的关系对, 获得 lncRNA-miRNA-mRNA 的 ceRNA 网络。

综上, 本研究在脑胶质瘤病例中, 取脑胶质瘤组织为样本, 通过 RNA-Seq [14] (Chu *et al.*, 2012, Wang *et al.*, 2009)筛选脑胶质瘤中差异表达的 lncRNA、miRNA 及其靶基因, 以期获得具体的 lncRNA-miRNA-mRNA 调控网络[15], 并对其靶基因进行 GO 和 KEGG 富集, 对其所在通路及所调控的生物过程进行分析, 探究 lncRNA、miRNA 对脑胶质瘤发生发展调控的影响, 为靶向 ncRNA 的脑胶质瘤治疗及预后提供分子理论依据。

2. 材料与方法

2.1. 样本采集

2017 年 11 月至 2018 年 2 月共采集四例符合《中国中枢神经系统胶质瘤诊断与治疗指南(2015)》[16] 诊断标准的患者的肿瘤组织及癌旁组织, 病理学诊断分期为 IV 级神经胶质瘤。所有组织标本的取材均需事先向患者及/或家属详细解释, 征得患者及/或家属的同意后签署《知情同意书》。标本采集在手术室进行, 并得到相关管理人员的许可, 同时争取得到相关人员的积极配合以保证标本的及时获得。选取的标本包括肿瘤组织、正常组织和(或)癌旁组织。标本切取的方法与时机: 首先进行大体标本的观察或拍照,

确认肿瘤的部位、范围，注意与周围组织及坏死组织的鉴别，在标本离体后应在最短的时间内切取标本(<30 min)，将癌组织及癌旁组织切取成多块直径为 0.5 cm 左右的组织块，分别装入已编号的 1.5 ml 灭菌冻存管中，迅速放入液氮转移罐并做好长期保存准备。记录样本编号、姓名、住院号、肿瘤类型、取材数量(分别用 T = 肿瘤，C = 正常，T 及 C 后面的数字表示取材份数)、离体时间(分钟)；四位患者的肿瘤样本分别编号为 AT、BT、CT、DT；相应的对照组织为 AC、BC、CC、DC。四组混测的肿瘤组织为 Q1，癌旁组织为 Q2 (表 1)。

Table 1. Sample grouping record form

表 1. 样本分组情况记录表

编号	性别	年龄	癌组织编号	癌旁组织编号	离体时间(分钟)
A	女	35	AT	AC	<5
B	女	57	BT	BC	<5
C	男	60	CT	CC	<5
D	男	43	DT	DC	<5
Q (四组混测)	-	-	Q1	Q2	--

2.2. RNA 的分离、纯化和质量控制

RNA 的分离、纯化和质量控制使用 Qiagen RNeasy Mini Kit (Qiagen, 德国)从脑胶质瘤样本组织中提取总 RNA。再对提取得到的 RNA 使用 Nano Photometer 分光光度计(IMPLEN, 美国)、含 Qubit RNA Assay Kit (Life Technologies, 美国)的 Qubit 2.0 Fluorometer 和含 RNA Nano 6000 Assay Kit (Agilent Technologies, 美国)的 Bioanalyzer 2100 系统来检测 RNA 的纯度，浓度和完整性。

2.3. 文库构建及转录组测序

测序工作由天津诺禾致源基因测序公司完成，miRNA、lncRNA 独立建库，即去除样本中的核糖体 RNA，然后对其进行测序。

样品检测合格后，使用 Small RNA Sample Pre Kit 构建文库，利用 Small RNA 的 3'及 5'端特殊结构(5'端有完整的磷酸基团，3'端有羟基)，以 total RNA 为起始样品，直接将 Small RNA 两端加上接头，然后反转录合成 cDNA。随后经过 PCR 扩增，PAGE 胶电泳分离目标 DNA 片段，切胶回收得到的即为 cDNA 文库。

文库构建完成后，先使用 Qubit2.0 进行初步定量，稀释文库至 1 ng/ul，随后使用 Agilent 2100 对文库的 insert size 进行检测，insert size 符合预期后，使用 Q-PCR 方法对文库的有效浓度进行准确定量(文库有效浓度 > 2 nM)，以保证文库质量。

库检合格后，把不同文库按照有效浓度及目标下机数据量的需求 pooling 后进行 HiSeq/MiSeq 测序。

2.4. 测序数据的处理与分析

原始测序数据(raw data)经去除接头和低质量 reads 等处理后得到有效数据(clean data)。使用 Tophat2 [17]进行参考基因组比对，再通过 cufflinks [18]完成转录本的拼接。以是否具有编码潜能作为判断标准，利用 CNCI v2 分析[19]、CPC v9.4 分析[20]、Pfam26.0 蛋白结构域分析[21]和 PhyloCSF 分析[22]筛选出 lncRNA、miRNA 和 mRNA，再使用 cuffdiff (<http://cole-trapnell-lab.github.io/cufflinks/cuffdiff/index.html>)软件进行 RNA 表达水平分析。

根据顺式 cis 作用原理: lncRNA 临近位置(上下游 100 kb)的蛋白编码基因作为其靶基因; 反式 trans 作用原理: 样本间 lncRNA 与蛋白编码基因的表达量采用 Pearson 相关系数法分析(相关系数 > 0.95)来预测 lncRNA 的靶基因。对差异表达的 lncRNA、miRNA 和 mRNA 使用 GSeq [23]、UniProt (<https://www.uniprot.org/>)、KOBAS (2.0) [24]和 Enrichr (<http://amp.pharm.mssm.edu/Enrichr/>)进行基因本体论(Gene Ontology, GO)富集和京都基因与基因组百科全书(Kyoto Encyclopedia of Genes and Genomes, KEGG)通路富集分析。

3. 结果与分析

在生物体内, 不同基因相互协调行使其生物学功能, 通过 Pathway 显著性富集能确定候选靶基因参与的最主要生化代谢途径和信号转导途径。KEGG (Kyoto Encyclopedia of Genes and Genomes)是有关 Pathway 的主要公共数据库[25]。Pathway 显著性富集分析以 KEGG Pathway 为单位, 应用超几何检验, 找出与整个基因组背景相比, 在候选靶基因中显著性富集的 Pathway [25]。该分析的计算公式:

$$p = 1 - \sum_{i=0}^{m-1} \frac{\binom{M}{i} \binom{N-M}{n-i}}{\binom{N}{n}}$$

其中, N 为所有基因中具有 Pathway 注释的基因数目; n 为 N 中候选靶基因的数目; M 为所有基因中注释为某特定 Pathway 的基因数目; m 为注释为某特定 Pathway 的候选靶基因数目。用 BH 的方法对 p-value 进行校正, 得到的校正后的 P-value 值越小代表越显著。这里将值小于 0.05 的 Pathway 定义为在候选靶基因中显著富集的 Pathway。

3.1. 差异 lncRNA 的筛选及富集分析

测序结果利用韦恩图显示四组间共有的差异 lncRNA 数目为零, 如图 1 所示。因此以混测组 (GLQ1_vs_GLQ2)所得差异表达的 lncRNA 为样本, 对其共表达的基因进行 GO 富集, 在混测组 (GLQ1_vs_GLQ2)中差异 lncRNA 共计 106 个, 其中上调 52 个包括 lnc_000266、lnc_000276、lnc_000417、

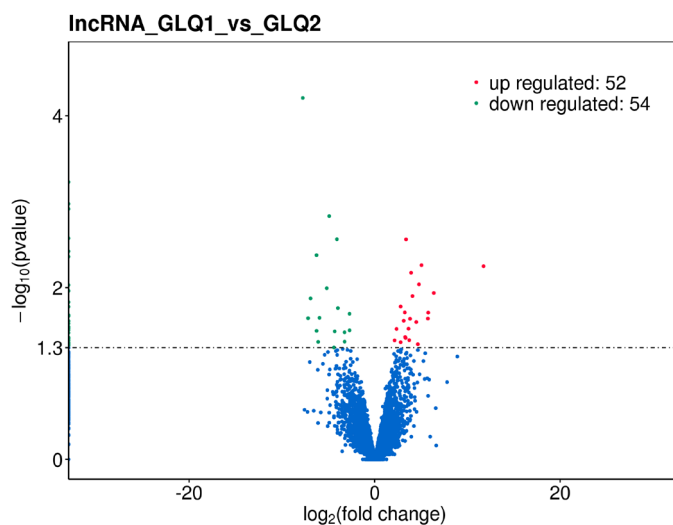


Figure 1. Differentially expressed lncRNA volcano map

图 1. 差异表达 lncRNA 火山图

lnc_000510、lnc_000565、lnc_000662、lnc_000853、lnc_001155、lnc_001315、lnc_001370、lnc_001788、lnc_001827 等; 下调 54 个, 包括 ENST00000572151、lnc_000009、lnc_000021、lnc_000036、lnc_000115、lnc_000410、lnc_000467 等(见表 2); 如图 2 所示。在 GO 富集结果分类统计图中, 从左往右依次为生物学过程、细胞组分、分子功能包括, 如图 3 所示。结果显示差异 lncRNA 主要参与有机循环化合物代谢过程、胞内运输、细胞蛋白质定位、蛋白质的运输、有丝分裂细胞周期、细胞周期调节、DNA 代谢过程的正调控、水解酶活性的调节等生物过程。如图 4 所示。差异 lncRNA 共表达 mRNA 的 KEGG 富集通路散点图如图 4, 主要富集的细胞通路有血小板激活、基础转录因子及补体和凝血级联等相关途径。q-value 的取值范围为[0, 1], 越接近于零, 表示富集结果的可信度越高。差异表达转录本(lncRNA)所对应的靶基因进行了 GO 富集分析, 主要富集在有机循环化合物代谢过程、胞内运输、细胞蛋白质定位、蛋白质的运输、有丝分裂细胞周期、细胞周期调节、DNA 代谢过程的正调控、水解酶活性的调节等生物过程。

Table 2. Mixed test group (GLQ1_vs_GLQ2) differential lncRNA screening results
表 2. 混测组(GLQ1_vs_GLQ2)差异 lncRNA 筛选结果

转录本 ID transcript_id	基因 ID gene_id	基因名 gene_name	癌组织表达量 GLQ1_FPKM	癌旁组织表达量 GLQ2_FPKM	log2. foldchange.	pval	padj
ENST00000329743.3	ENSG00000185837.3	CECR5-AS1	1.321697	0	Inf	0.005694	0.456532
ENST00000412495.5	ENSG00000174403.15	C20orf166-AS1	0.055641	0.361857	-2.7012	0.031411	1
ENST00000424587.6	ENSG00000228463.9	AP006222.2	0.197026	0	Inf	0.017627	0.821515
ENST00000425112.1	ENSG00000235954.6	TTC28-AS1	0.777166	0.078821	3.30158	0.03762	1
ENST00000425271.1	ENSG00000229588.1	RP11-479J7.2	0	0.253397	Inf	0.043235	1
ENST00000426715.5	ENSG00000237737.5	DCTN1-AS1	0.094955	1.591254	-4.06678	0.002733	0.297451
ENST00000435067.2	ENSG00000235366.2	LINC01055	0.499809	0.009094	5.780278	0.019621	0.874801
ENST00000444939.1	ENSG00000235897.1	TM4SF19-AS1	0.203001	0	Inf	0.03037	1
ENST00000451837.5	ENSG00000228315.11	GUSBP11	1.596795	0.152712	3.386296	0.002744	0.297989
ENST00000463143.5	ENSG00000243701.5	DUBR	0.058796	1.161521	-4.30416	0.03242	1
ENST00000477513.1	ENSG00000205930.8	C21orf62-AS1	0	0.291069	Inf	0.03869	1
ENST00000488584.1	ENSG00000240875.5	LINC00886	0.241863	0	Inf	0.039771	1
ENST00000496389.5	ENSG00000240057.5	RP11-572M11.4	0.491534	0.051418	3.256939	0.01952	0.871556
ENST00000505149.1	ENSG00000249307.5	LINC01088	2.266866	0.068304	5.052585	0.005493	0.449092
ENST00000505844.1	ENSG00000248455.5	RP11-321E2.3	0.015737	0.329017	-4.38593	0.049939	1
ENST00000507761.1	ENSG00000249307.5	LINC01088	4.470149	0.870082	2.361099	0.030368	1
ENST00000509088.5	ENSG00000249307.5	LINC01088	10.85965	1.550521	2.808152	0.016659	0.800687
ENST00000510087.5	ENSG00000247828.7	TMEM161B-AS1	0.542471	3.541659	-2.70681	0.020203	0.891566
ENST00000521404.5	ENSG00000214548.14	MEG3	0.265479	0	Inf	0.020493	0.897167
ENST00000522747.5	ENSG00000245146.6	LINC01024	0	0.138021	Inf	0.031289	1
ENST00000540882.5	ENSG00000256193.5	LINC00507	0.015798	0.974333	-5.94665	0.022486	0.936989
ENST00000568966.2	ENSG00000260391.2	RP11-71H17.7	1.517343	0.099538	3.930152	0.006726	0.496316
ENST00000570301.5	ENSG00000262223.6	RP11-1055B8.3	0	0.872976	Inf	0.003783	0.359047
ENST00000572151.1	ENSG00000263069.5	CTD-2047H16.4	0	0.277937	Inf	0.028468	1

Continued

ENST00000580045.1	ENSG00000263766.5	RP11-580I16.2	0.034864	0.539885	-3.95284	0.017324	0.815143
ENST00000609442.5	ENSG00000186594.12	MIR22HG	0	0.744879	Inf	0.001059	0.187946
ENST00000623070.3	ENSG00000228794.8	LINC01128	0.010925	0.392027	-5.16525	0.010188	0.617517
LNC_000009	XLOC_000367	.	0.003548	0.272429	-6.26283	0.031884	1
LNC_000021	XLOC_000434	.	0	0.873833	Inf	0.038025	1
LNC_000036	XLOC_000839	.	0	0.526067	Inf	0.01673	0.800687
LNC_000115	XLOC_007755	.	0	0.258049	Inf	0.021498	0.914279
LNC_000266	XLOC_019040	.	0.259798	0	Inf	0.010122	0.61593
LNC_000269	XLOC_019813	.	0.319856	0	Inf	0.027765	1
LNC_000276	XLOC_020592	.	0.149726	0	Inf	0.023653	0.963619
LNC_000410	XLOC_032710	.	0	0.184696	Inf	0.021326	0.911866
LNC_000417	XLOC_033405	.	0.171251	0	Inf	0.027822	1
LNC_000467	XLOC_037210	.	0	0.197732	Inf	0.045953	1
LNC_000479	XLOC_037967	.	0	0.154129	Inf	0.04384	1
LNC_000494	XLOC_039296	.	0.004717	0.563647	-6.90085	0.013361	0.708158
LNC_000510	XLOC_040050	.	0.393638	0.038834	3.341487	0.038722	1
LNC_000521	XLOC_040623	.	0.042489	0.400015	-3.2349	0.042694	1
LNC_000565	XLOC_045043	.	0.264409	0	Inf	0.017004	0.807935
LNC_000570	XLOC_045448	.	0	0.177139	Inf	0.047729	1
LNC_000662	XLOC_056010	.	0.395915	0	Inf	0.006839	0.501673
LNC_000694	XLOC_060487	.	0	0.273759	Inf	0.004371	0.393991
LNC_000798	XLOC_073979	.	0	0.217494	Inf	0.045961	1
LNC_000845	XLOC_078789	.	0.465307	0.053043	3.132943	0.024325	0.974234
LNC_000853	XLOC_080147	.	0.588241	0.084005	2.80786	0.043285	1
LNC_001155	XLOC_115160	.	0.181953	0	Inf	0.03034	1
LNC_001173	XLOC_116066	.	0.045814	0.434029	-3.24394	0.033192	1
LNC_001254	XLOC_122620	.	0.005249	0.754748	-7.16795	0.022717	0.943285
LNC_001300	XLOC_124776	.	0	0.362518	Inf	0.001216	0.20293
LNC_001315	XLOC_125784	.	0.335061	0.023615	3.826649	0.023139	0.952305
LNC_001359	XLOC_129378	.	0	0.153206	Inf	0.029025	1
LNC_001370	XLOC_131364	.	0.186167	0.013987	3.734489	0.040925	1
LNC_001521	XLOC_141586	.	0	0.418985	Inf	0.00059	0.140597
LNC_001763	XLOC_161705	.	0.329112	0.006127	5.747195	0.022992	0.949315
LNC_001787	XLOC_164705	.	5.637431	0.001648	11.74011	0.005645	0.456448
LNC_001788	XLOC_164705	.	2.519544	0	Inf	0.010249	0.619682
LNC_001789	XLOC_164705	.	0	9.485185	Inf	0.002668	0.293497

Continued

LNC_001790	XLOC_164705	.	0.022131	1.497452	-6.08031	0.042932	1
LNC_001827	XLOC_168953	.	0.51866	0	Inf	0.036278	1
LNC_001934	XLOC_168978	.	0.600014	0	Inf	0.01441	0.740057
LNC_002091	XLOC_169062	.	0.386222	0.015235	4.664016	0.045734	1
LNC_002171	XLOC_169093	.	0	0.28152	Inf	0.024453	0.975955
LNC_002196	XLOC_169112	.	0.239854	0	Inf	0.02842	1
LNC_002218	XLOC_169575	.	0	0.625311	Inf	0.020923	0.904419
LNC_002436	XLOC_189417	.	0	0.134111	Inf	0.026241	1
LNC_002514	XLOC_196104	.	0	0.178863	Inf	0.045259	1
LNC_002523	XLOC_196947	.	0	0.16073	Inf	0.040958	1
LNC_002545	XLOC_198663	.	0.22451	0	Inf	0.000301	0.101765
LNC_002576	XLOC_202277	.	0.185142	0	Inf	0.037008	1
LNC_002660	XLOC_210683	.	0.175563	0	Inf	0.037318	1
LNC_002670	XLOC_211305	.	0.00305	0.235054	-6.26792	0.004193	0.383829
LNC_002738	XLOC_219310	.	1.96684	0.442293	2.152807	0.041188	1
LNC_002765	XLOC_223733	.	0	0.170617	Inf	0.047406	1
LNC_002816	XLOC_228769	.	0	0.275173	Inf	0.010921	0.642293
LNC_002819	XLOC_228781	.	0.410449	0	Inf	2.33E-05	0.021745
LNC_002830	XLOC_229641	.	0	0.354426	Inf	0.014756	0.748978
LNC_002831	XLOC_229642	.	0	0.498238	Inf	0.041203	1
LNC_002841	XLOC_230383	.	0	0.441705	Inf	0.038664	1
LNC_002927	XLOC_240626	.	0	0.342247	Inf	0.009353	0.595621
LNC_002940	XLOC_243325	.	0	0.187294	Inf	0.021126	0.909446
LNC_002965	XLOC_247874	.	0.293177	0.023328	3.651653	0.030167	1
LNC_002996	XLOC_249865	.	0.17017	0	Inf	0.044745	1
LNC_003083	XLOC_259353	.	0.301593	0	Inf	0.018348	0.839064
LNC_003092	XLOC_260368	.	0	0.181184	Inf	0.033685	1
LNC_003160	XLOC_265610	.	0.529069	0.019271	4.778952	0.009191	0.589476
LNC_003191	XLOC_268753	.	0.168433	0	Inf	0.009208	0.589747
LNC_003225	XLOC_272726	.	0.268701	0.011997	4.485256	0.025239	0.992031
LNC_003259	XLOC_276571	.	0.342624	0	Inf	0.014933	0.755263
LNC_003277	XLOC_277968	.	0.212491	0	Inf	0.01983	0.88007
LNC_003306	XLOC_282244	.	0	0.381837	Inf	0.039356	1
LNC_003371	XLOC_288937	.	0.12854	0	Inf	0.024828	0.981586
LNC_003426	XLOC_296063	.	0	0.330483	Inf	0.010965	0.642886
LNC_003482	XLOC_302093	.	0.31455	0	Inf	0.003595	0.34929
LNC_003495	XLOC_303175	.	0.174044	0	Inf	0.003122	0.321776
LNC_003498	XLOC_303273	.	0	0.080993	Inf	0.044795	1
LNC_003557	XLOC_311294	.	0.203297	0	Inf	0.03206	1

Continued

LNC_003613	XLOC_317217	.	0.849733	0.049924	4.089205	0.012542	0.684973
LNC_003671	XLOC_325487	.	0	0.125253	Inf	0.029909	1
LNC_003672	XLOC_325487	.	0.155251	0	Inf	0.032466	1
LNC_003745	XLOC_332734	.	0.344085	0.004141	6.376727	0.011558	0.655778
LNC_003874	XLOC_345209	.	0.010717	0.320129	-4.90075	0.001472	0.220468
LNC_003875	XLOC_345209	.	0.003259	0.698022	-7.74281	6.20E-05	0.040318
LNC_003988	XLOC_357378	.	0	0.386403	Inf	0.024658	0.980174

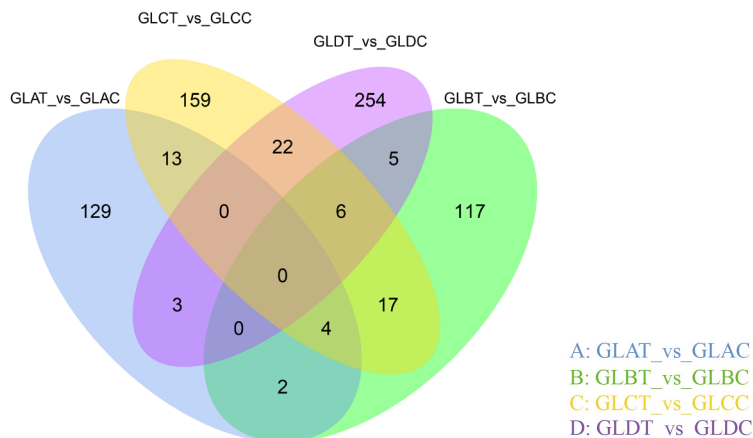


Figure 2. Venn diagram of lncRNA differential expression between all groups
图 2. 组间差异表达 lncRNA 韦恩图

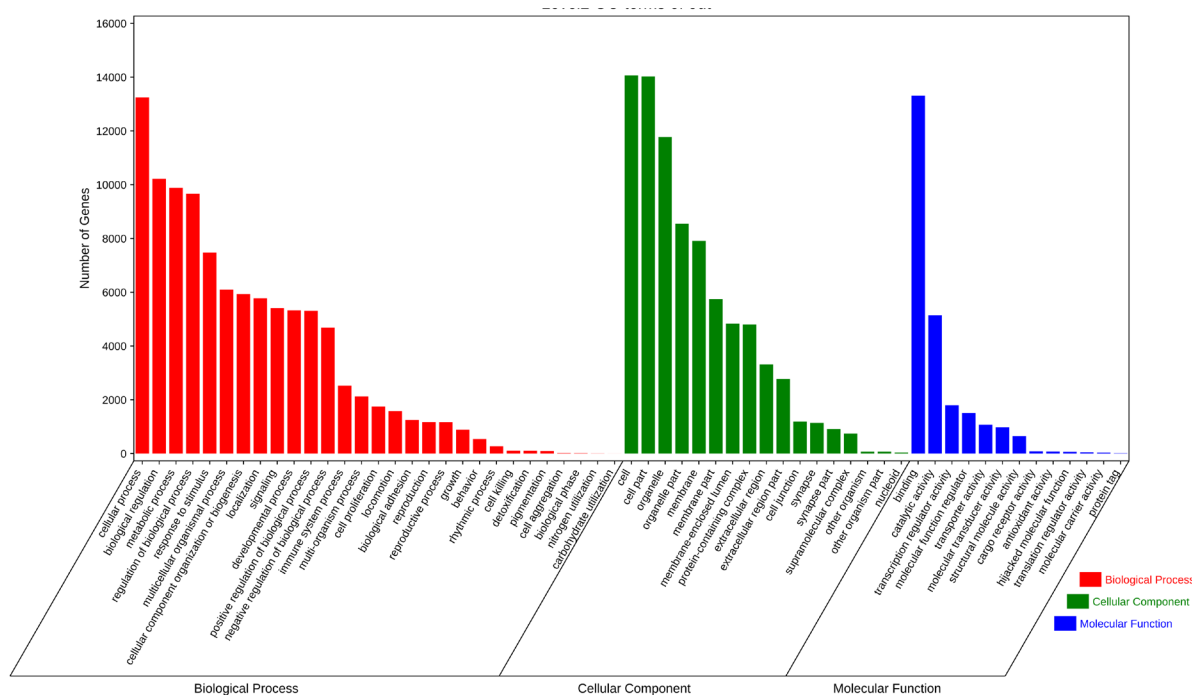


Figure 3. Mixed test group (GLQ1_vs_GLQ2) difference lncRNA co-expression gene GO enrichment classification statistics
图 3. 混测组(GLQ1_vs_GLQ2)差异 lncRNA 共表达基因 GO 富集分类统计图

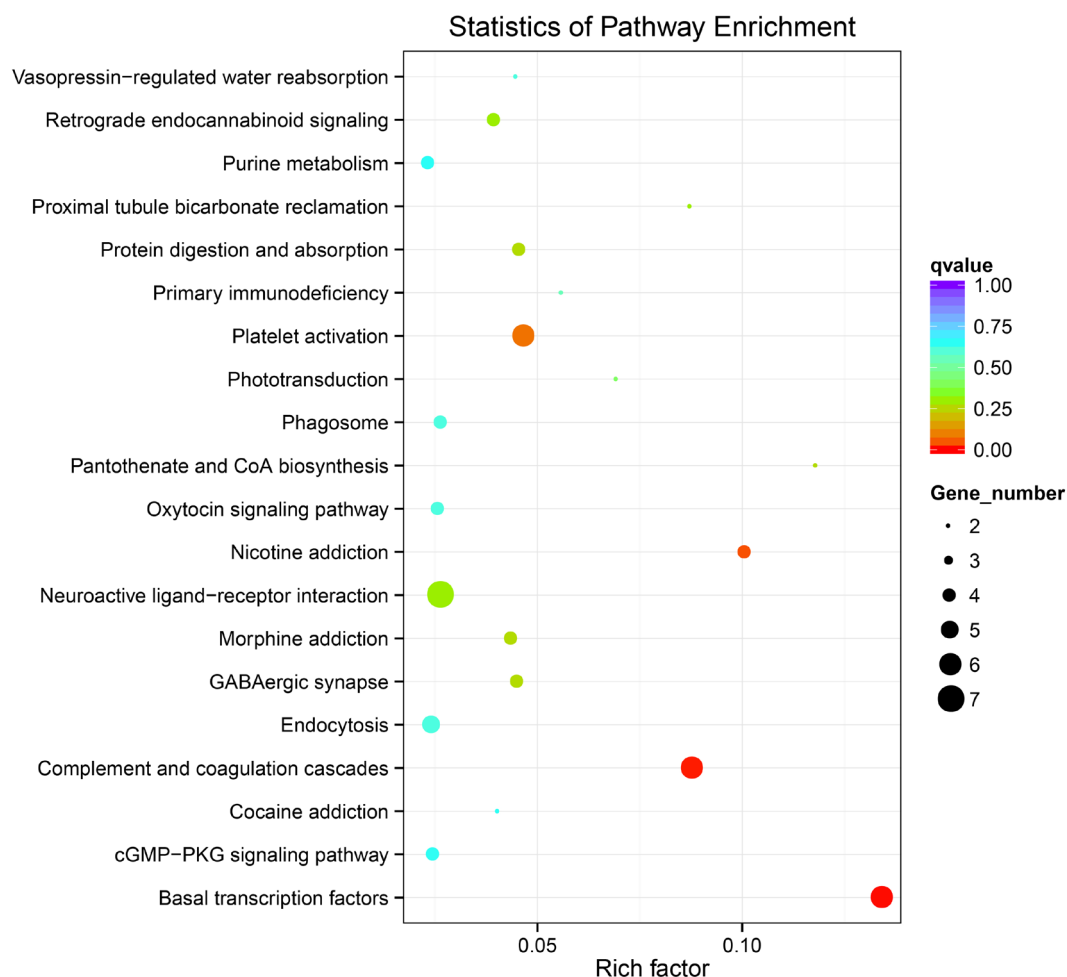


Figure 4. Mixed test group (GLQ1_vs_GLQ2) differential lncRNA co-expression gene KEGG pathway enrichment scatter diagram

图 4. 混测组(GLQ1_vs_GLQ2)差异 lncRNA 共表达基因 KEGG 途径富集散点图

3.2. 差异 microRNA 的筛选及富集分析

将多组比较的差异 miRNA 个数进行统计, 绘制维恩图, 直观展现出各个比较组共有及特有的差异 miRNA 数目(见图 5)。如图 6 所示, 四组建交集 miRNA 数目为零, 因此以混测组(GLQ1_vs_GLQ2)所得结果进行后期分析。混测组(GLQ1_vs_GLQ2)差异 microRNA 共计 15 个(见表 3), 如图 6 所示。其中上调 8 个包 hsa-miR-3157-3p、hsa-miR-6761-5p、hsa-miR-3124-5p、hsa-miR-5009-5p、hsa-miR-216b-3p、hsa-miR-296-3p、novel_912、novel_328; 下调 7 个, hsa-miR-135b-5p、hsa-miR-377-3p、hsa-miR-136-5p、hsa-miR-381-3p、hsa-miR-153-5p、hsa-miR-299-3p、hsa-miR-4473。富集的 GO 相关的所有差异 miRNA 靶基因(见表 4, 无确定靶基因则未列出)的分类统计图, 从左往右依次为生物学过程、细胞组分、分子功能包括, 如图 7 所示。结果显示差异 miRNA 主要参与细胞过程、生物调节、代谢过程、应激反应、多细胞有机体过程、细胞成分的组织或生物发生、细胞定位、信号途径、发育过程、生物过程的正调控、生物过程的负调控。如图 7 所示。差异 miRNA 靶基因的 KEGG 富集通路散点图如图 8, 主要富集的细胞通路维生素的消化吸收、肿瘤坏死因子信号通路、突触囊泡循环、唾液分泌、癌症中的蛋白多糖、蛋白质输出、癌症相关通路、胰腺癌等相关途径。(q-value 的取值范围为[0, 1], 越接近于零, 表示富集结果的可信度越高)。

Table 3. Mixed test group (GLQ1_vs_GLQ2) differential miRNA screening results
表 3. 混测组(GLQ1_vs_GLQ2)差异 miRNA 筛选结果

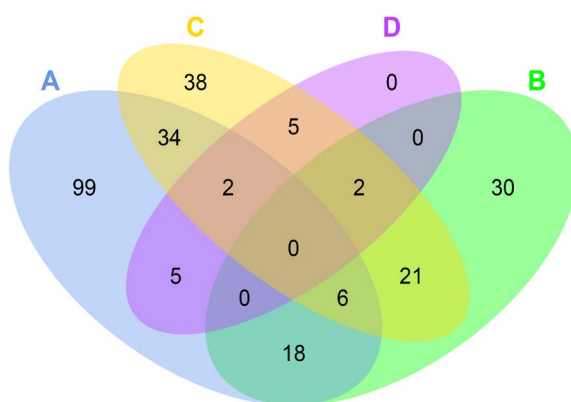
miRNA 编号 miRNA-ID	癌组织表达量 GLQ1_readcount	癌旁组织表达量 GLQ2_readcount	log2FoldChange	pval	padj
hsa-miR-135b-5p	88.99762	163.8794	-0.77662	0.004145	0.99975
hsa-miR-377-3p	75.37529	170.7388	-0.92773	0.006783	0.99975
hsa-miR-3157-3p	9.345945	3.034668	1.0028	0.013002	0.99975
hsa-miR-136-5p	632.8087	1442.705	-0.8977	0.01311	0.99975
hsa-miR-381-3p	8668.745	17737.01	-0.8269	0.01399	0.99975
novel_912	2.42883	0	0.88379	0.016828	0.99975
hsa-miR-6761-5p	7.562805	2.212444	0.9635	0.021225	0.99975
hsa-miR-3124-5p	15.3221	7.072264	0.82849	0.024967	0.99975
hsa-miR-5009-5p	1.709722	0	0.8047	0.026246	0.99975
novel_328	7.026624	2.158375	0.89549	0.032114	0.99975
hsa-miR-153-5p	19.4879	33.3998	-0.64796	0.039772	0.99975
hsa-miR-216b-3p	2.787696	0.292098	0.78371	0.041112	0.99975
hsa-miR-299-3p	79.26836	197.027	-0.81686	0.045115	0.99975
hsa-miR-4473	20.53004	44.69347	-0.7704	0.047291	0.99975
hsa-miR-296-3p	27.00353	9.479225	0.82828	0.04798	0.99975

Table 4. Mixed test group (GLQ1_vs_GLQ2) differential miRNA and its target gene comparison table
表 4. 混测组(GLQ1_vs_GLQ2)差异 miRNA 及其靶基因对照

miRNA 编号 miRNA-ID	靶基因编号	靶基因名称	描述
hsa-miR-3157-3p	ENSG00000141738	GRB7	生长因子受体结合蛋白 7
	ENSG00000151353	TMEM18	跨膜蛋白 18
	ENSG00000161011	SQSTM1	整合体 1
	ENSG00000165458	INPPL1	肌醇多磷酸磷酸酶类似物 1
	ENSG00000168748	CA7	碳酸酐酶 7
hsa-miR-381-3p	ENSG00000183837	PNMA3	PNMA 家庭成员 3
	ENSG00000145191	EIF2B5	真核翻译起始因子 2B 亚基 ϵ
	ENSG00000015153	YAF2	YY1 相关因子 2
novel_912	ENSG00000170949	ZNF160	锌指蛋白 160
	ENSG00000179921	GPBAR1	G 蛋白偶联的胆汁酸受体 1
	ENSG00000188938	FAM120AOS	序列相似性为 120A 反向链的家族
	ENSG00000281348	AC120114.4	智人 16 号染色体克隆 CTD-2574D22 完整序列
hsa-miR-6761-5p	ENSG00000078070	MCCC1	甲基巴豆酰辅酶 A 羧化酶 1
	ENSG00000100578	KIAA0586	KIAA0586
	ENSG00000109339	MAPK10	丝裂原激活的蛋白激酶 10
	ENSG00000196209	SIRPB2	信号调节蛋白 β 2

Continued

hsa-miR-3124-5p	ENSG00000105737	GRIK5	谷氨酸离子受体的海藻酸酯型亚基 5
	ENSG00000107742	SPOCK2	骨结合蛋白 cwcvc 和 kazal 样结构域蛋白聚糖 2
novel_328	ENSG00000109917	ZPR1	ZPR1 锌指
	ENSG00000090447	TFAP4	转录因子 AP-4
	ENSG00000146576	C7orf26	7 号染色体开放阅读框 26
hsa-miR-299-3p	ENSG00000167920	TMEM99	跨膜蛋白 99 (可能)
	ENSG00000173638	SLC19A1	溶质载体家族 19 成员 1
	ENSG00000196345	ZKSCAN7	具有 KRAB 和 SCAN 域的锌指 7
	ENSG00000006025	OSBPL7	氧固醇结合蛋白类似物 7
	ENSG00000007001	UPP2	尿苷磷酸化酶 2
	ENSG00000023191	RNH1	核糖核酸酶/血管生成素抑制剂 1
	ENSG00000068903	SIRT2	sirtuin 2
	ENSG00000089041	P2RX7	嘌呤能受体 P2X 7
	ENSG00000089060	SLC8B1	溶质载体家族 8 成员 B1
hsa-miR-296-3p	ENSG00000092096	SLC22A17	溶质载体家族 22 成员 17
	ENSG00000099812	MISP	有丝分裂纺锤体定位
	ENSG00000100401	RANGAP1	Ran GTPase 激活蛋白 1
	ENSG00000100985	MMP9	基质金属蛋白酶 9



A: GLATvsGLAC
 B: GLBTvsGLBC
 C: GLCTvsGLCC
 D: GLDTvsGLDC

Figure 5. Venn diagram of microRNA differential expression between all groups
 图 5. 组间差异表达 microRNA 韦恩图

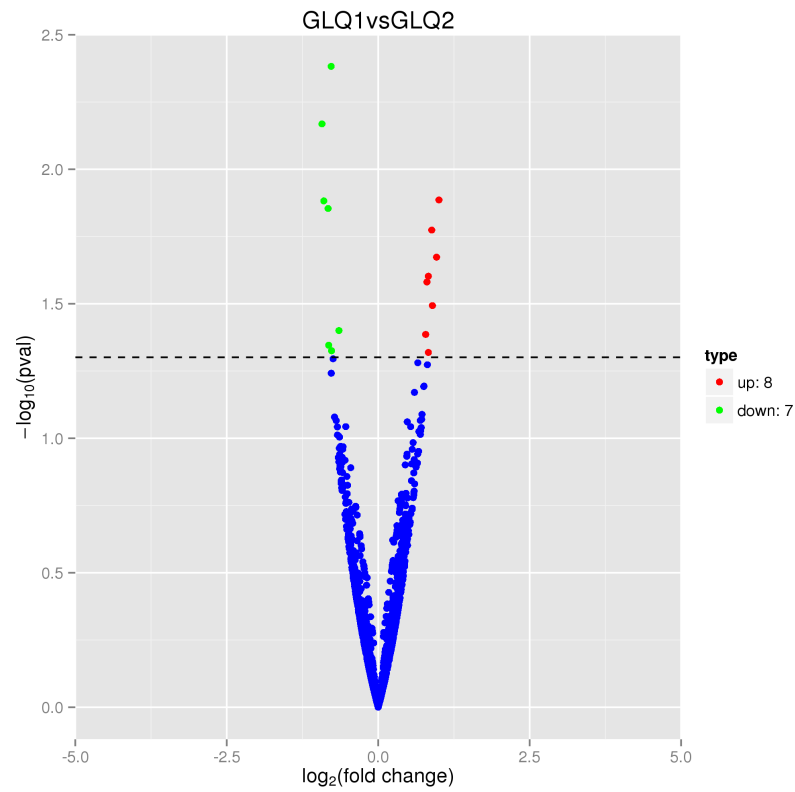


Figure 6. The volcano map of mixed test group (GLQ1_vs_GLQ2) differential microRNA

图 6. 混测组(GLQ1_vs_GLQ2)差异 microRNA 火山图

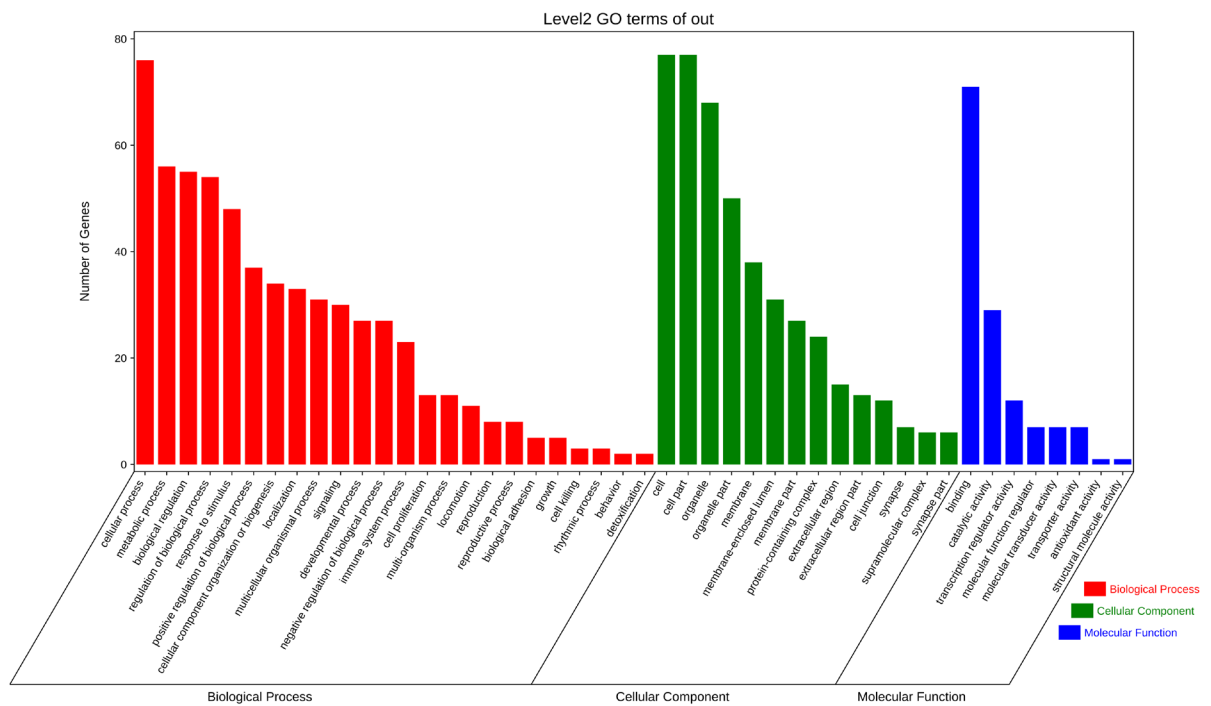


Figure 7. Mixed test group (GLQ1_vs_GLQ2) difference microRNA target gene GO enrichment classification statistics

图 7. 混测组(GLQ1_vs_GLQ2)差异 microRNA 靶基因 GO 富集分类统计图

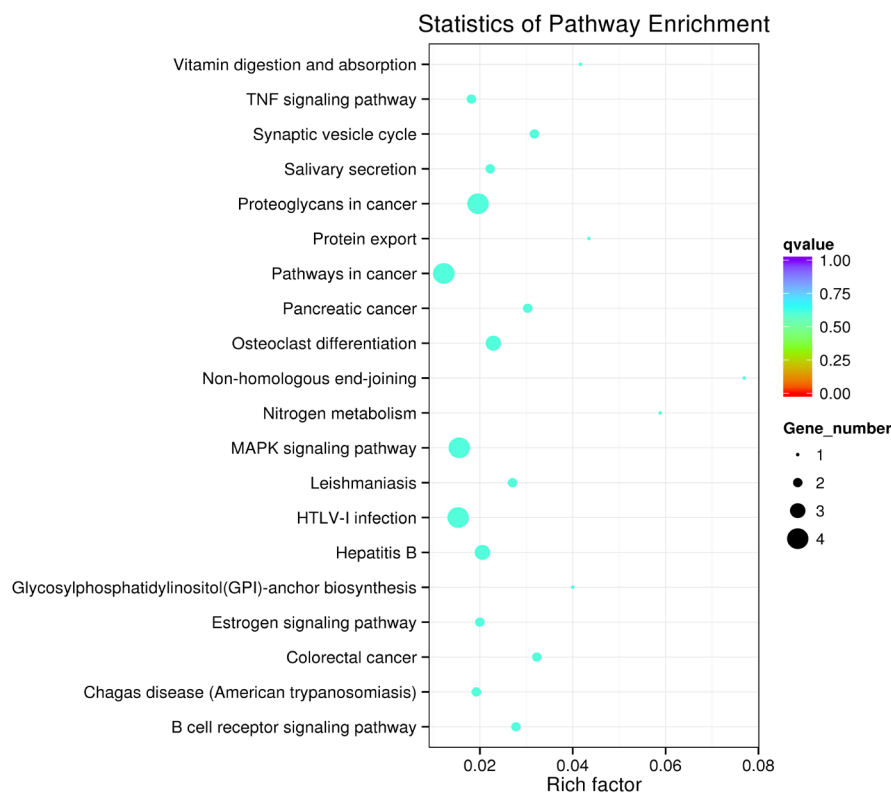


Figure 8. Mixed test group (GLQ1_vs_GLQ2) differential microRNA KEGG pathway enrichment scatter diagram

图 8. 混测组(GLQ1_vs_GLQ2)差异 microRNA KEGG 途径富集散点图

3.3. lncRNA-miRNA-mRNA 的关联分析

lncRNA 具有广泛的调控作用,不仅能够直接调控 DNA 的结构、RNA 的转录和翻译,而且具有 miRNA 结合位点,可以竞争性地结合 miRNA,抑制 miRNA 对靶基因的调节作用,从而间接地调控基因表达。根据 ceRNAs 理论,筛选具有相同 miRNA 结合位点的 lncRNA-gene 关系对,构建以 lncRNA 为起始、miRNA 为核心、mRNA 为靶标的 lncRNA-miRNA-gene 组合来构建 ceRNA 调控网络(图 8~10)。在全转录组层面,通过 ceRNA 调控网络,揭示非编码 RNA 调控基因表达的机制。

统计筛选得到的 lncRNA-miRNA-gene 组合中上下调 RNA 的数目,可以直观展现不同组合中相关基因数目的变化。不同的形状代表不同的 RNA 种类,不同颜色代表上下调关系,不同的节点(node)的大小与此节点在通路中所处位置的重要程度(degree)成正比,即与此节点相连的边越多,它的重要程度越大,节点也就越大,这些节点在网络中可能处于较为核心的位置,与所研究的生物学问题更相关。

根据筛选所得数据,分别建立 lncRNA (down)-miRNA (up)-mRNA (down)和 lncRNA (up)-miRNA (down)-mRNA (up)三级互作关系网络(图 9~11)。在 lncRNA (down)-miRNA (up)-mRNA (down)网络中,以 hsa-miR-296-3p 为中间节点,上游 lncRNA: LNC_002514、LNC_001173 下调表达,使得 hsa-miR-296-3p 表达量增加,进而下游基因: MRAS (ENSG00000158186)、GABBR1 (ENSG00000204681)、AIFM3 (ENSG00000183773)、TRIM26 (ENSG00000234127)、ZNF875 (ENSG00000181666)、HNRNPUL1 (ENSG00000105323)、FOSL1 (ENSG00000175592)、INO80E (ENSG00000169592)、PDGFRB (ENSG00000113721)的表达量也随之下调(见表 5),对以上 9 种下调表达的基因进行 KEGG 富集,共到 6 条通路,分别为 HTLV-I 感染、磷脂酶 D 信号通路、肌动蛋白细胞骨架的调节、Rap1 信号通路、Ras 信

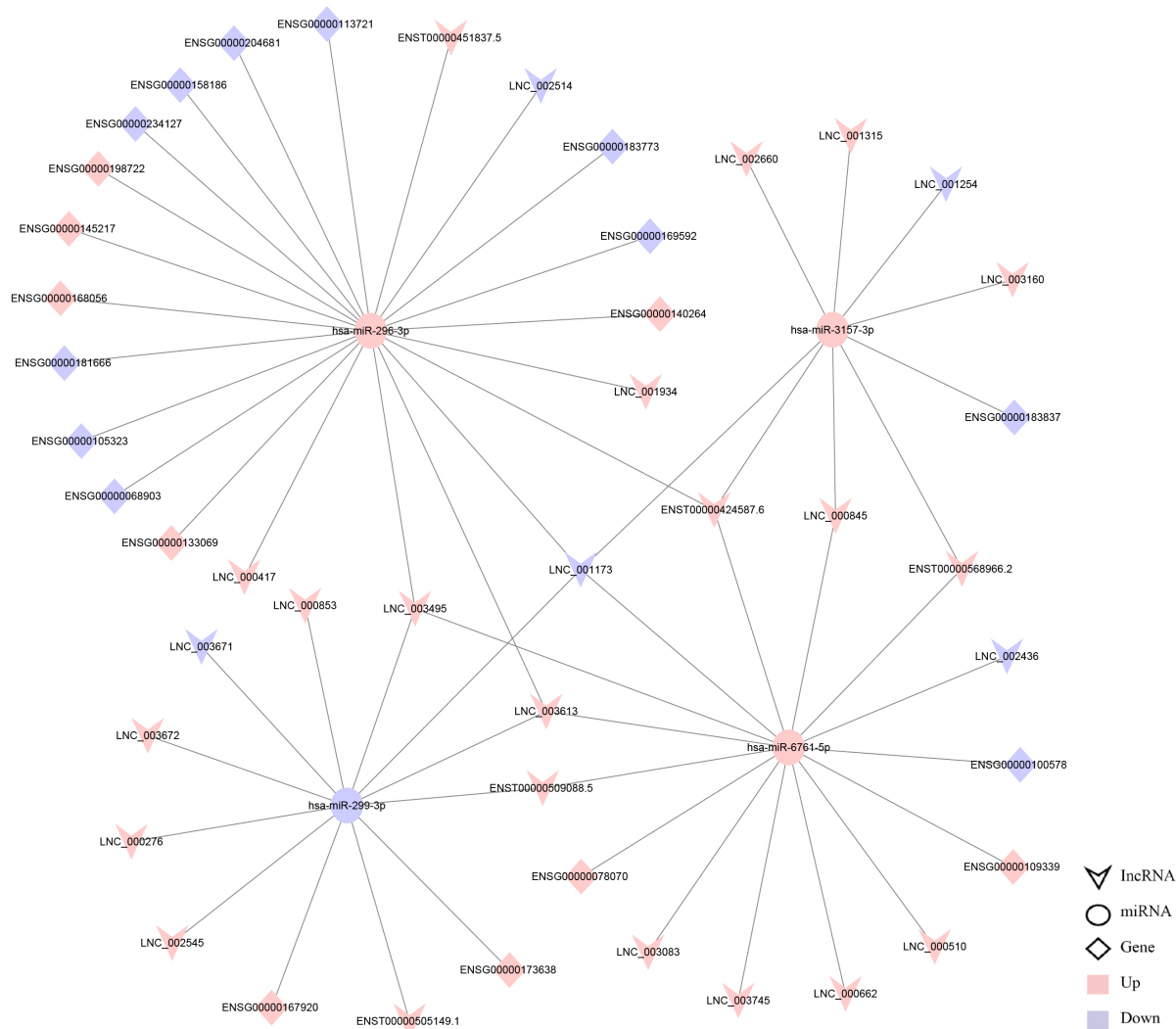


Figure 9. Interaction map of all differential lncRNA-miRNA-mRNA

图 9. 全部差异 lncRNA-miRNA-mRNA 互动图

号通路、MAPK 信号通路。其中下调表达的 lncRNA: LNC_001173 与 LNC_002436、LNC_001254 分别调控两种 miRNA: hsa-miR-6761-5p 和 hsa-miR-3157-3p 的表达量上调, 而相应的 hsa-miR-6761-5p 的靶基因 KIAA0586 (ENSG00000100578)、hsa-miR-3157-3p 的靶基因 PNMA3 (ENSG00000183837) 的表达量也随之下调。在 lncRNA (up)-miRNA (down)-mRNA (up) 互作网络中, 8 种上调表达的 lncRNA: LNC_003613、LNC_003495、LNC_000853、LNC_000276、LNC_003672、LNC_002545、ENST00000509088、ENST00000505149 可调控 miRNA: hsa-miR-299-3p 使之表达减少, 并使下游基因 SLC19A1 (ENSG00000173638) 的表达量同步增加。

在以上数据筛选基础上, 将上述结果与 Targetscan 数据库中已有的 miRNA 与靶基因的互作结合位点进行比对, 在上述 lncRNA (down)-miRNA (up)-mRNA (down) 网络中只有 FOSL1 (ENSG00000175592) 与 hsa-miR-296-3p 有两个确定的互作位点, 上游下调的两个: LNC_002514、LNC_001173 均无文献数据支持, 但 LNC_001173 在本次筛选结果中涉及内容过多, 验证难度较大, 因此拟建立由 LNC_002514-hsa-miR-296-3p-FOSL1 组成的轴向调控通路(图 12), 以备后期研究验证。

Table 5. Gene ID and its Chinese and English gene name comparison table
表 5. 基因 ID 及其中英文基因名对照表

基因编号 Gene stable ID	基因名 Gene name	描述 Description
ENSG00000105323	HNRNPUL1	异核核糖核蛋白 U 样 1
ENSG00000113721	PDGFRB	血小板衍生的生长因子受体 β
ENSG00000158186	MRAS	肌肉 RAS 癌基因同源物
ENSG00000169592	INO80E	INO80 复合亚基 E
ENSG00000175592	FOSL1	FOS 类似 1, AP-1 转录因子亚基
ENSG00000181666	ZNF875	锌指蛋白 875
ENSG00000183773	AIFM3	凋亡诱导因子线粒体相关 3
ENSG00000204681	GABBR1	γ -氨基丁酸 B 型受体亚基 1
ENSG00000234127	TRIM26	Tripartite 基序蛋白 26
ENSG00000183837	PNMA3	副肿瘤抗原样 3 基因 3
ENSG00000100578	KIAA0586	KIAA0586
ENSG00000173638	SLC19A1	溶质载体家族 19 成员 1

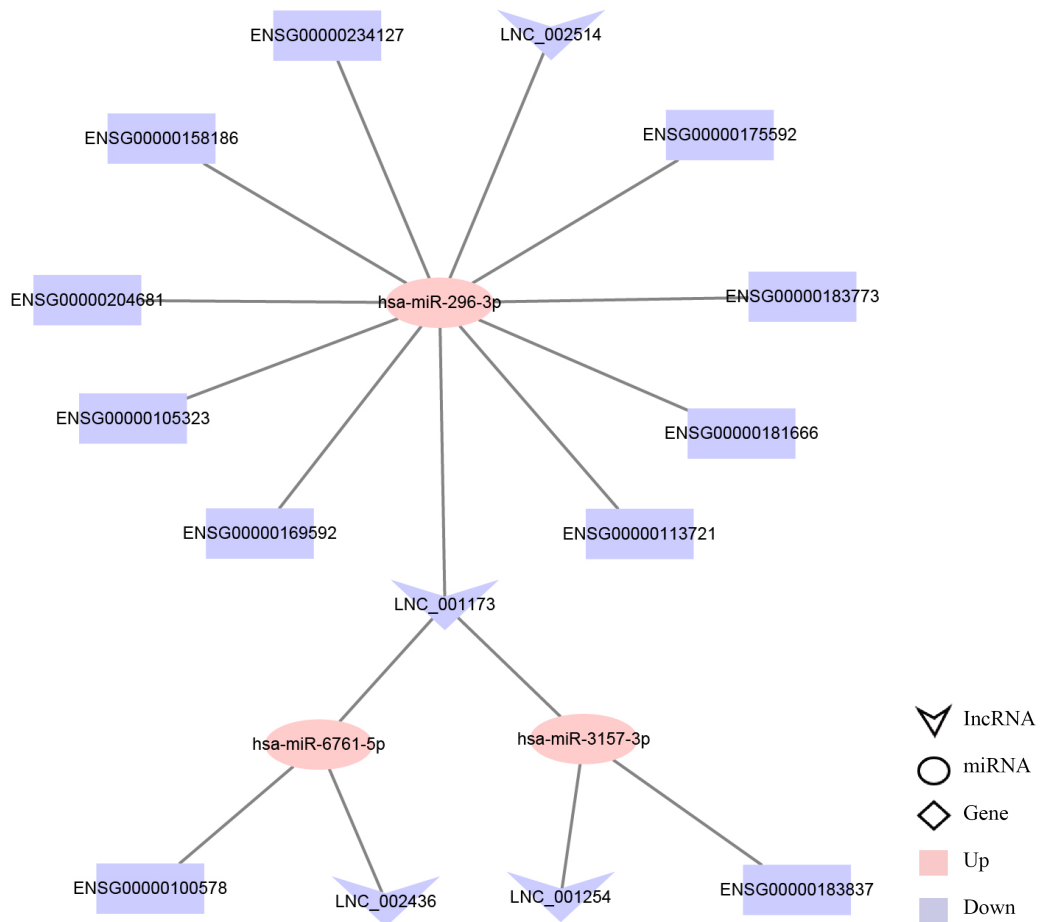


Figure 10. Differential lncRNA (down)-miRNA (up)-mRNA (down) interaction network diagram
图 10. 差异 lncRNA (down)-miRNA (up)-mRNA (down) 交互网络图

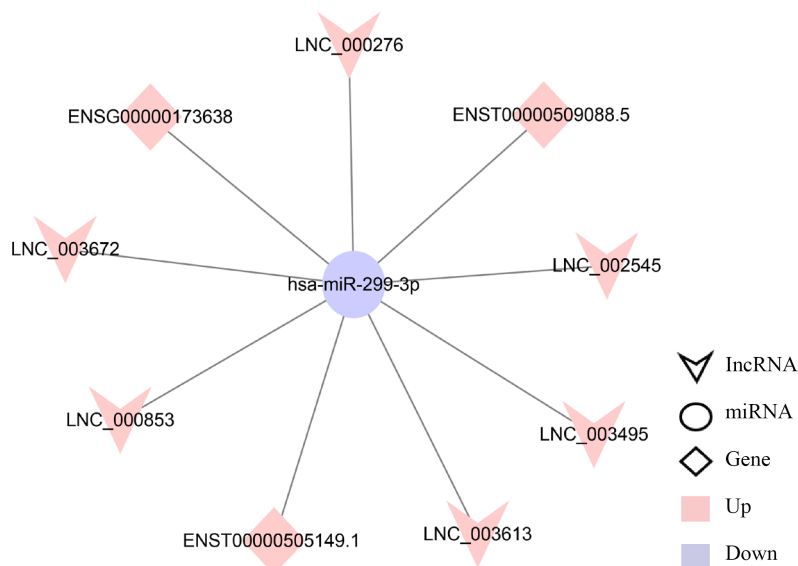


Figure 11. Differential lncRNA (up)-miRNA (down)-mRNA (up) interaction network diagram

图 11. 差异 lncRNA (up)-miRNA (down)-mRNA (up) 交互网络图

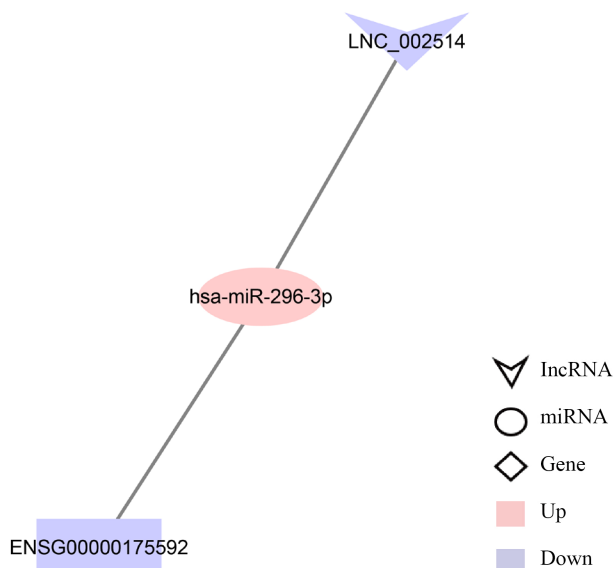


Figure 12. LNC_002514-hsa-miR-296-3p-FOSL1 axial regulation pathway diagram

图 12. LNC_002514-hsa-miR-296-3p-FOSL1 轴向调控通路图

4. 讨论

通过肿瘤 RNA 测序(RNA-Seq)监控基因和转录组的表达变化可以协助了解肿瘤的分类和发展, 全转录组测序不仅可以得到每种 RNA 的全套分析结果, 还可通过二重、三重关联分析, 对目标基因的范围进行有效的缩小, 进而得到准确的关键基因。本研究基于二代测序技术的全转录组测序研究, 同时分析样本中差异表达的 mRNA, lncRNA, miRNA, 并且通过两两关联分析、三元关联分析, 建立可能的 ceRNA 调控网络, 使研究内容更加系统化, 致力于深入挖掘生命现象背后的转录调控问题。

4.1. 差异 lncRNA 的分析

在本次测序结果中,已知有注释的差异 lncRNA 四种:HOXD-AS2 [26]、LINC01088 [27]、DNMBP-AS1 [28]、Metazoa_SRP [29]。其中 HOXD-AS2 与多种癌症密切相关,在其与胶质瘤相关研究中,研究人员对胶质瘤细胞系和正常星形胶质细胞系进行了针对 lncRNA 表达谱高通量的芯片检测,结果表明:HOXD-AS2 在神经胶质瘤细胞中升高,与神经胶质瘤的分级和预后不良有关;HOXD-AS2 的丢失可通过诱导体外细胞周期 G1 阻滞来抑制神经胶质瘤细胞的生长,进而胶质瘤的增殖受到抑制。除胶质瘤外还与肺癌细胞的耐药性、胃癌、卵巢癌、乳腺癌和皮肤鳞状细胞癌的发生及恶化密切相关;DNMBP-AS1 已有明确研究证实其可作为肝癌患者预后的重要生物标记。Metazoa_SRP 已有研究表明其与肺纤维化[30]、骨肉瘤[31]的发生有关,同时在乳腺癌、结直肠癌表型的易感性标记有极高的关联性[32],并且在人脑的发育及相关精神疾病[29]的研究中具有深远影响;在非小细胞肺癌(NSCLC)的研究中,linc01088 和 EZH2 在 NSCLC 组织和细胞系中均高表达:抑制 linc01088 就抑制了 NSCLC 细胞的增殖,并延长了 G1 期,同时缩短了 S 和 G2-M 期;RIP 分析揭示了 linc01088 与 EZH2 之间的直接结合关系;抑制 linc01088 或 EZH2 诱导 p21 表达的上调,随后抑制了肿瘤的生长;证明了 linc01088 可以通过与 EZH2 结合抑制 p21 来促进细胞增殖,从而加剧 NSCLC 的肿瘤发生。因此,linc01088 可能是潜在的癌基因和新型抗肿瘤治疗的靶标[33]。在卵巢癌中,linc01088 可通过靶向 miR-24-1-5p 抑制卵巢上皮细胞的肿瘤发生[34],并且与胃癌、肝癌[35]和子宫内膜癌[34]发生发展的多个重要通路相关。而 CTC-458G6.2 和 CTB-1I21.1 在目前的研究中尚未有明确机制表明其与疾病发生或预后的影响及意义,但现有数据及分析可知,这两个 lncRNA 在大脑中特异性表达,并且与多个已知与胶质瘤相关基因具有互作关系,可作为日后研究的重要方向。这些 lncRNA 与相关文献报道表达水平相一致,目前在胶质瘤发病机制研究领域也有多篇报道

本次测序也发现多个未注释的差异 lncRNA,但其对胶质瘤的调控机制还有待进一步实验验证。而 lncRNA: LNC_002514、LNC_001173、LNC_002436 和 LNC_001254 目前在人类疾病方面尚未研究,因此作为潜在生物标记,为今后的研究验证工作提供方向。

4.2. 差异 miRNA 的分析

在本次测序结果中,差异表达的 miRNA: hsa-miR-296-3p 在人脐带静脉内皮细胞(HUVEC)中的 miRNA 和 mRNA 表达谱相关研究中差异调节涉及细胞粘附,血管生成,细胞周期,JAK-STAT 信号传导,MAPK 信号传导,一氧化氮信号传导,VEGF 信号传导和伤口愈合途径的基因。此外,基于 q-PCR 的 miRNA 和 mRNA 上调和下调实验研究表明,hsa-miR-296-3p 等 miRNA 有效地影响斜管恒温器刺激微重力(MG)条件下 HUVEC 的细胞增殖和血管功能。对相互作用组结果的共识表明,暴露于短期 MG 的 HUVEC 转录组的波动受限,这可能导致较高水平的内皮功能,如血管生成和血管形成[36]。hsa-miR-6761-5p、hsa-miR-3157-3p 在最新研究中被列为颅内动脉瘤发展的潜在靶标[37],hsa-miR-3157-3p 在大肠癌相关的 ceRNA 调控网络中重要节点[38]。与 hsa-miR-299-3p 相关的研究较为广泛,与黑色素瘤[39]、多形性胶质母细胞瘤[40]、透明肾细胞癌[41]等多种常见癌症相关,并且被列为喉癌[42]、结肠癌[43]、颅内动脉瘤[37]等疾病的独立预后因素。

4.3. 下游靶基因分析

受差异 miRNA 调控的下游靶基因共 12 种:PDGFRB [44]、MRAS [45]、INO80E [46]、hnRNPUL1 [47]、FOSL1 [48]、ZNF875 [49]、AIFM3 [50]、GABBR1 [51]、TRIM26 [52]、PNMA3 [53]、KIAA0586 [54]、SLC19A1 [55]。

血小板衍生的生长因子(PDGF)信号通路驱动高度神经胶质瘤(HGG)进程,并且 PDGF 受体(PDGFRs)

的表达增强是胶质母细胞瘤(GBMs)子集中的一种普遍的突变。PDGFRA 在神经胶质瘤细胞中表达,而 PDGFRB 主要限于与神经胶质瘤相关的基质。相关研究还显示,小胶质细胞(TAMMs)的空间位置与已经获得 PDGFRB 在小鼠和人类低级神经胶质瘤和 HCGs 中表达的肿瘤细胞亚群的扩展相关。此外, M2 极化的小胶质细胞诱导神经胶质瘤细胞中 PDGFRB 的表达并刺激其迁移能力[56]。

MRAS 是相对于经典 RAS 癌蛋白而言最接近的,并且共享大多数调节和效应子相互作用。但是它也具有独特的功能,包括与 SHOC2 和蛋白磷酸酶 1 (PP1)结合时,具有磷酸酶调节亚基的功能。这种磷酸酶复合物调节 RAF 激酶激活周期中的关键步骤,并提供有效的 ERK 途径激活和 RAS 转化所需的关键坐标输入。MRAS 突变很少在癌症中发生,但在某些情况下表达失调可能在肿瘤发生中起作用。在 RAS 病 Noonan 综合征中确实发生了 MRAS 的激活突变(以及 SHOC2 和 PP1),这强调了 RAS-ERK 途径中 MRAS 的关键作用。MRAS 在细胞迁移和分化中也具有独特作用,其性质与细胞极性调节中的关键作用相一致[45]。

INO80E 在评估异基因造血干细胞移植(HSCT)在胰腺导管腺癌治疗中的免疫学作用中,是其重要的免疫应答指标。并能够对临床和生物学相关细胞免疫组库的建立产生影响[46]。

2020 年 4 月 *FOSL1* 在低级神经胶质瘤(LGG)的研究中,通过比较不同类型肿瘤之间的 *FOSL1* 表达,探索 TCGA-LGG 和亚组中 *FOSL1* 表达与总体生存(OS)之间的关联。结果表明:LGG 中的 *FOSL1* 表达在 21 种不同的癌症中排名第一。*FOSL1* 表达较低的 LGG 的 OS 较长($P < 0.001$)。星形细胞瘤组的 *FOSL1* 表达最高,OS 最短,其次是少星形细胞瘤和少突胶质细胞瘤($P < 0.05$)。1p19q 共缺失或 *IDH* 突变亚组具有较低的 *FOSL1* 表达和较长的 OS ($P < 0.001$)。与相应的组相比, *FOSL1* 表达较低的 LGG 的 OS 比以下组更长,因此 *FOSL1* 是 LGG 和亚组的预后标志物[47]。

4.4. LNC_002514-hsa-miR-296-3p-FOSL1 轴向调控的分析

在 lncRNA (down)-miRNA (up)-mRNA (down)网络中,经过生物信息学分析后发现 LNC_002514 表达量下降,释放 hsa-miR-296-3p 作用于下游基因: *FOSL1*。

在 2019 年 12 月发表的关于冠状动脉疾病(CAD)的生物标记文章中,研究人员建立一个用于诊断 CAD 的 circRNA-miRNA-mRNA 三重网络。使用接收器工作特性(ROC)曲线评估了 hsa-miR-21-3p 和 hsa-miR-296-3p、circYOD1 的预后价值。circ-YOD1 的 ROC 曲线下面积(AUC)为 0.824, hsa-miR-21-3p 的 AUC 为 0.731, hsa-miR-296-3p 为 0.776。因而在对 cad 中 circRNA-miRNA-mRNA 的三重调控网络的研究表明, hsa-miR-296-3p 在潜在生物标记的研究过程中有着及其重要作用[57]。

FOSL1 (Fos 样抗原 1, 也称为 Fra-1)是 FOS 家族的成员,它由四个基因成员 c-FOS, FOSB, *FOSL1* 和 *FOSL2* 组成,所有这些基因编码包含亮氨酸拉链的蛋白质。与 Jun (c-jun, junB, junD), 激活转录因子(ATF)和肌腱膜纤维肉瘤(MAF)结合的 FOS 是激活蛋白 1 (AP-1)复合物的成员。*FOSL1* 是调节正常细胞生长,分化和凋亡的重要核转录因子。在宫颈癌发病机制的最新研究中,研究人员发现 Fra-1 抑制宫颈癌细胞的增殖,同时还促进细胞凋亡并影响细胞周期分布。此外, Fra-1 上调子宫颈癌细胞中的 STAT1 表达并调节 p53 信号通路活性。Fra-1 的过表达通过改变 HeLa 细胞中的 sirtuin 1 (SIRT1)表达来抑制细胞衰老,而 Fra-1 的过表达恢复了线粒体疾病并抑制了 HeLa 细胞中的代谢重编程。STAT1 沉默会削弱 Fra-1 对宫颈癌细胞生长的抑制作用,而 STAT1 的敲低则逆转了由 Fra-1 引起的 HeLa 细胞对细胞衰老和线粒体功能障碍的影响。STAT1 沉默也恢复了子宫颈癌细胞的代谢重编程。其研究结果表明,Fra-1 通过 STAT1 介导的 p53 信号通路调节来抑制子宫颈癌细胞的生长和 Warburg 效应[48]。此外,已发现它在包括肺癌在内的许多癌症中具有重要作用。在胶质瘤的相关研究中,当 *FOSL1* 沉默时,改变了恶性神经胶质瘤细胞的形态,锚定非依赖性生长和体内肿瘤形成。它的过表达赋予非致瘤性神经胶质瘤细胞系致瘤性[56]。

LNC_002514 在目前的研究中尚未有明确的研究结果,通过生物信息学分析结果可知 LNC_002514 与 hsa-miR-296-3p 及其下游基因 FOSL1 存在可能的调控关系,因而可作为潜在的研究方向。

在本次检测结果中, FOSL1 的表达量与预期结果相反,因而有如下可能: 1. 在组织切取过程中因癌组织与癌旁组织边界不明显、难以准确区分,所以所取得的样品可能不完全是胶质瘤组织,故而检测结果与理论依据不符。2. 本次测序样本容量较小,因此可能无法准确展现 FOSL1 在恶性胶质瘤中的表达特点。

5. 结论

本研究通过 RNA-seq 和生物信息学分析,筛选出了在四例脑胶质瘤样本中差异表达的 lncRNA 106 种,上调表达 52 种,下调表达 54 种;差异表达 miRNA 15 种,上调表达 8 种,下调表达 7 种;这些非编码 RNA 主要参与有机循环化合物代谢过程、胞内运输、细胞蛋白质定位、蛋白质的运输、有丝分裂细胞周期、细胞周期调节、DNA 代谢过程的正调控、水解酶活性的调节等生物过程。并且根据 lncRNA、miRNA、mRNA 的靶向关系建立了 ceRNAs 调控网络,在此基础上提出了 LNC_002514-hsa-miR-296-3p-FOSL1 轴向调控的可能性,并对其作出简要分析。上述 nc-RNA 的筛选及调控轴向的提出将为以后脑胶质瘤相关研究提供新的方向。

致 谢

感谢王仕超师姐提供的数据支持,感谢导师张焱如教授在各个方面无微不至的照顾与帮助,感谢舍友为我提供了安静的写作环境,感谢父母多年来的养育与经济和精神支持,感谢男友的支持与安慰,谢谢大家!

基金项目

内蒙古自治区自然科学基金项目(2017MS0307)。

参考文献

- [1] Rees, J.H. (2011) Diagnosis and Treatment in Neuro-Oncology: An Oncological Perspective. *The British Journal of Radiology*, **84**, 82-89. <https://doi.org/10.1259/bjr/18061999>
- [2] 中华医学会病理学分会脑神经病理学组. 2016 WHO 中枢神经系统肿瘤分类第 4 版修订版概述及胶质瘤部分介绍[J]. *中华病理学杂志*, 2016, 45(11): 745-747.
- [3] Muir, C.S., Storm, H.H. and Polednak, A. (1994) Brain and Other Nervous System Tumours. *Cancer Survival*, **19-20**, 369-392.
- [4] 中国中枢神经系统胶质瘤诊断与治疗指南编写组. *中华医学杂志*[J]. 中华医学会, 2016, 96(7): 485-509.
- [5] Strom, Q.T., Gittleman, H., Liao, P., et al. (2016) CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2009-2013. *Neuro-Oncology*, **18**, v1-v75. <https://doi.org/10.1093/neuonc/nov207>
- [6] Ng, H.K. and Lam, P.Y. (1998) The Molecular Genetics of Central Nervous System Tumors. *Pathology*, **30**, 196-202. <https://doi.org/10.1080/00313029800169236>
- [7] Schlessinger, J. (2000) Cell Signaling by Receptor Tyrosine Kinases. *Cell*, **103**, 211-225. [https://doi.org/10.1016/S0092-8674\(00\)00114-8](https://doi.org/10.1016/S0092-8674(00)00114-8)
- [8] 王曦, 陈定. LncRNAs 作为 miRNA 的靶模拟物调节 miRNA [J]. *科技创新导报*, 2016, 13(28): 176-177.
- [9] Marziali, A. and Akesson, M. (2001) New DNA Sequencing Methods. *Annual Review of Biomedical Engineering*, **3**, 195-223. <https://doi.org/10.1146/annurev.bioeng.3.1.195>
- [10] Goodwin, S., McPherson, J.D. and McCombie, W.R. (2016) Coming of Age: Ten Years of Next-Generation Sequencing Technologies. *Nature Reviews Genetics*, **17**, 333-351. <https://doi.org/10.1038/nrg.2016.49>
- [11] Niedzicka, M., Fijarczyk, A., Dudek, K., et al. (2016) Molecular Inversion Probes for Targeted Resequencing in

- Non-Model Organisms. *Scientific Reports*, **6**, Article No. 24051. <https://doi.org/10.1038/srep24051>
- [12] Xu, J., Gao, C., Zhang, F., *et al.* (2016) Differentially Expressed lncRNAs and mRNAs Identified by Microarray Analysis in GBS Patients vs Healthy Controls. *Scientific Reports*, **6**, Article No. 21819. <https://doi.org/10.1038/srep21819>
- [13] Xu, J., Zhang, F., Gao, C., *et al.* (2017) Microarray Analysis of lncRNA and mRNA Expression Profiles in Patients with Neuromyelitis Optica. *Molecular Neurobiology*, **54**, 2201-2208. <https://doi.org/10.1007/s12035-016-9754-0>
- [14] Wang, Z., Gerstein, M. and Snyder, M. (2009) RNA-Seq: A Revolutionary Tool for Transcriptomics. *Nature Reviews Genetics*, **10**, 57-63. <https://doi.org/10.1038/nrg2484>
- [15] Yang, Y., *et al.* (2018) Analyzing the Interactions of mRNAs, miRNAs, lncRNAs and circRNAs to Predict Competing Endogenous RNA Networks in Glioblastoma. *Journal of Neuro-Oncology*, **137**, 493-502. <https://doi.org/10.1007/s11060-018-2757-0>
- [16] 《中国中枢神经系统胶质瘤诊断与治疗指南》编写组. 中华医学杂志[J]. 中华医学会, 2016, 96(7): 485-509.
- [17] Kim, D., Pertea, G., Trapnell, C., *et al.* (2013) TopHat2: Accurate Alignment of Transcriptomes in the Presence of Insertions, Deletions and Gene Fusions. *Genome Biology*, **14**, R36. <https://doi.org/10.1186/gb-2013-14-4-r36>
- [18] Trapnell, C., Williams, B.A., Pertea, G., *et al.* (2010) Transcript Assembly and Quantification by RNA-Seq Reveals Unannotated Transcripts and Isoform Switching during Cell Differentiation. *Nature Biotechnology*, **28**, 511-515. <https://doi.org/10.1038/nbt.1621>
- [19] Liang, S., Luo, H., Bu, D., *et al.* (2013) Utilizing Sequence Intrinsic Composition to Classify Protein-Coding and Long Noncoding Transcripts. *Nucleic Acids Research*, **41**, e166. <https://doi.org/10.1093/nar/gkt646>
- [20] Kong, L., Zhang, Y., Ye, Z.Q., *et al.* (2007) CPC: Assess the Protein-Coding Potential of Transcripts Using Sequence Features and Support Vector Machine. *Nucleic Acids Research*, **35**, W345-W349. <https://doi.org/10.1093/nar/gkm391>
- [21] Punta, M., Coggill, P.C., Eberhardt, R.Y., *et al.* (2011) The Pfam Protein Families Database. *Nucleic Acids Research*, **40**, D290-D301. <https://doi.org/10.1093/nar/gkr1065>
- [22] Lin, M.F., Jungreis, I. and Kellis, M. (2011) PhyloCSF: A Comparative Genomics Method to Distinguish Protein Coding and Non-Coding Regions. *Bioinformatics*, **27**, i275-i282. <https://doi.org/10.1093/bioinformatics/btr209>
- [23] Young, M.D., Wakefield, M.J., Smyth, G.K., *et al.* (2010) Gene Ontology Analysis for RNA-seq: Accounting for Selection Bias. *Genome Biology*, **11**, R14. <https://doi.org/10.1186/gb-2010-11-2-r14>
- [24] Mao, X., Cai, T., Olyarchuk, J.G., *et al.* (2005) Automated Genome Annotation and Pathway Identification Using the KEGG Orthology (KO) as a Controlled Vocabulary. *Bioinformatics*, **21**, 3787-3793. <https://doi.org/10.1093/bioinformatics/bti430>
- [25] Kanehisa, M., Araki, M., Goto, S., *et al.* (2008) KEGG for Linking Genomes to Life and the Environment. *Nucleic Acids Research*, **36**, D480-D484. <https://doi.org/10.1093/nar/gkm882>
- [26] Yao, L., Ye, P.C., Tan, W., *et al.* (2020) Decreased Expression of the Long Non-Coding RNA HOXD-AS2 Promotes Gastric Cancer Progression by Targeting HOXD8 and Activating PI3K/Akt Signaling Pathway. *World Journal of Gastrointestinal Oncology*, **12**, 1237-1254. <https://doi.org/10.4251/wjgo.v12.i11.1237>
- [27] Zhang, W., Fei, J., Yu, S., *et al.* (2018) LINC01088 Inhibits Tumorigenesis of Ovarian Epithelial Cells by Targeting miR-24-1-5p. *Scientific Reports*, **8**, Article No. 2876. <https://doi.org/10.1038/s41598-018-21164-9>
- [28] Xue, C., Zhao, Y., Jiang, J. and Li, L. (2020) Expression Levels of lncRNAs Are Prognostic for Hepatocellular Carcinoma Overall Survival. *American Journal of Translational Research*, **12**, 1873-1883.
- [29] O'Brien, H.E., Hannon, E., Hill, M.J., *et al.* (2018) Expression Quantitative Trait Loci in the Developing Human Brain and Their Enrichment in Neuropsychiatric Disorders. *Genome Biology*, **19**, 194. <https://doi.org/10.1186/s13059-018-1567-1>
- [30] Deist, M.S., Gallardo, R.A., Bunn, D.A., Dekkers, J.C.M., Zhou, H. and Lamont, S.J. (2017) Resistant and Susceptible Chicken Lines Show Distinctive Responses to Newcastle Disease Virus Infection in the Lung Transcriptome. *BMC Genomics*, **18**, 989. <https://doi.org/10.1186/s12864-017-4380-4>
- [31] Rathe, S.K., Popescu, F.E., Johnson, J.E., *et al.* (2019) Identification of Candidate Neoantigens Produced by Fusion Transcripts in Human Osteosarcomas. *Scientific Reports*, **9**, Article No. 358. <https://doi.org/10.1038/s41598-018-36840-z>
- [32] Pande, M., Joon, A., Brewster, A.M., *et al.* (2018) Genetic Susceptibility Markers for a Breast-Colorectal Cancer Phenotype: Exploratory Results from Genome-Wide Association Studies. *PLoS ONE*, **13**, e0196245. <https://doi.org/10.1371/journal.pone.0196245>
- [33] Liu, J.Q., Feng, Y.H., Zeng, S. and Zhong, M.Z. (2020) linc01088 Promotes Cell Proliferation by Scaffolding EZH2 and Repressing p21 in Human Non-Small Cell Lung Cancer. *Life Sciences*, **241**, Article ID: 117134.

- <https://doi.org/10.1016/j.lfs.2019.117134>
- [34] Ai, H., Xie, W., Xiu, A.H., *et al.* (2018) The Down-Regulation of Long Non-Coding RNA LINC01088 Is Associated with the Poor Prognosis of Epithelial Ovarian Cancer Patients. *European Review for Medical and Pharmacological Sciences*, **22**, 5836-5841.
- [35] Liu, J., Yao, Y., Hu, Z., Zhou, H. and Zhong, M. (2019) Transcriptional Profiling of Long-Intergenic Noncoding RNAs in Lung Squamous Cell Carcinoma and Its Value in Diagnosis and Prognosis. *Molecular Genetics & Genomic Medicine*, **7**, e994. <https://doi.org/10.1002/mgg3.994>
- [36] Kasiviswanathan, D., Chinnasamy Perumal, R., Bhuvaneshwari, S., *et al.* (2020) Interactome of miRNAs and Transcriptome of Human Umbilical Cord Endothelial Cells Exposed to Short-Term Simulated Microgravity. *NPJ Microgravity*, **6**, 18. <https://doi.org/10.1038/s41526-020-00108-6>
- [37] Liao, B., Zhou, M.X., Zhou, F.K., *et al.* (2020) Exosome-Derived MiRNAs as Biomarkers of the Development and Progression of Intracranial Aneurysms. *Journal of Atherosclerosis and Thrombosis*, **27**, 545-610. <https://doi.org/10.5551/jat.51102>
- [38] Zhao, H., *et al.* (2020) Construction of ceRNA Coexpression Network and Screening of Molecular Targets in Colorectal Cancer. *Disease Markers*, **2020**, Article ID: 2860582. <https://doi.org/10.1155/2020/2860582>
- [39] Motti, M.L., Minopoli, M., Di Carluccio, G., Ascierio, P.A. and Carriero, M.V. (2020) MicroRNAs as Key Players in Melanoma Cell Resistance to MAPK and Immune Checkpoint Inhibitors. *International Journal of Molecular Sciences*, **21**, 4544. <https://doi.org/10.3390/ijms21124544>
- [40] Henriksen, M., Johnsen, K.B., Andersen, H.H., Pilgaard, L. and Duroux, M. (2014) MicroRNA Expression Signatures Determine Prognosis and Survival in Glioblastoma Multiforme—A Systematic Overview. *Molecular Neurobiology*, **50**, 896-913. <https://doi.org/10.1007/s12035-014-8668-y>
- [41] He, H., Wang, L., Zhou, W., *et al.* (2015) MicroRNA Expression Profiling in Clear Cell Renal Cell Carcinoma: Identification and Functional Validation of Key miRNAs. *PLoS ONE*, **10**, e0125672. <https://doi.org/10.1371/journal.pone.0125672>
- [42] Liu, H., Sun, Y., Tian, H., *et al.* (2019) Characterization of Long Non-Coding RNA and Messenger RNA Profiles in Laryngeal Cancer by Weighted Gene Co-Expression Network Analysis. *Aging (Albany NY)*, **11**, 10074-10099. <https://doi.org/10.18632/aging.102419>
- [43] Roman-Canal, B., Tarragona, J., Moiola, C.P., *et al.* (2019) EV-Associated miRNAs from Peritoneal Lavage as Potential Diagnostic Biomarkers in Colorectal Cancer. *Journal of Translational Medicine*, **17**, 208. <https://doi.org/10.1186/s12967-019-1954-8>
- [44] Gilbertson, R.J. and Clifford, S.C. (2003) PDGFRB Is Overexpressed in Metastatic Medulloblastoma. *Nature Genetics*, **35**, 197-198. <https://doi.org/10.1038/ng1103-197>
- [45] Young, L.C. and Rodriguez-Viciana, P. (2018) MRAS: A Close but Understudied Member of the RAS Family. *Cold Spring Harbor Perspectives in Medicine*, **8**, a033621. <https://doi.org/10.1101/cshperspect.a033621>
- [46] Omazic, B., Ayoglu, B., Löhr, M., *et al.* (2017) A Preliminary Report: Radical Surgery and Stem Cell Transplantation for the Treatment of Patients with Pancreatic Cancer. *Journal of Immunotherapy*, **40**, 132-139. <https://doi.org/10.1097/CJI.0000000000000164>
- [47] Zhu, J., Zhao, Y.P. and Zhang, Y.Q. (2020) Low Expression of FOSL1 Is Associated with Favorable Prognosis and Sensitivity to Radiation/Pharmaceutical Therapy in Lower Grade Glioma. *Neurological Research*, **42**, 522-527. <https://doi.org/10.1080/01616412.2020.1748323>
- [48] Zhang, M., Liang, L., He, J., *et al.* (2020) Fra-1 Inhibits Cell Growth and the Warburg Effect in Cervical Cancer Cells via STAT1 Regulation of the p53 Signaling Pathway. *Frontiers in Cell and Developmental Biology*, **8**, Article ID: 579629. <https://doi.org/10.3389/fcell.2020.579629>
- [49] Oguri, T., Katoh, O., Takahashi, T., *et al.* (1998) The Krüppel-Type Zinc Finger Family Gene, HKR1, Is Induced in Lung Cancer by Exposure to Platinum Drugs. *Gene*, **222**, 61-67. [https://doi.org/10.1016/S0378-1119\(98\)00464-8](https://doi.org/10.1016/S0378-1119(98)00464-8)
- [50] Chua-On, D., Prongvitaya, T., Tummanatsakun, D., *et al.* (2020) Apoptosis-Inducing Factor, Mitochondrion-Associated 3 (AIFM3) Protein Level in the Sera as a Prognostic Marker of Cholangiocarcinoma Patients. *Biomolecules*, **10**, 1021. <https://doi.org/10.3390/biom10071021>
- [51] Li, J., Zhang, C., Yuan, X., Ren, Z. and Yu, Z. (2020) Correlations between Stemness Indices for Hepatocellular Carcinoma, Clinical Characteristics, and Prognosis. *American Journal of Translational Research*, **12**, 5496-5510.
- [52] Wang, K., Chai, L., Qiu, Z., Zhang, Y., Gao, H. and Zhang, X. (2019) Overexpression of TRIM26 Suppresses the Proliferation, Metastasis, and Glycolysis in Papillary Thyroid Carcinoma Cells. *Journal of Cellular Physiology*, **234**, 19019-19027. <https://doi.org/10.1002/jcp.28541>
- [53] Takaji, M., Komatsu, Y., Watakabe, A., Hashikawa, T. and Yamamori, T. (2009) Paraneoplastic Antigen-Like 5 Gene

-
- (PNMA5) Is Preferentially Expressed in the Association Areas in a Primate Specific Manner. *Cerebral Cortex*, **19**, 2865-2879. <https://doi.org/10.1093/cercor/bhp062>
- [54] Wang, T., Xuan, Z., Dou, Y., *et al.* (2019) Identification of Novel Mutations in Preaxial Polydactyly Patients through Whole-Exome Sequencing. *Molecular Genetics & Genomic Medicine*, **7**, e690. <https://doi.org/10.1002/mgg3.690>
- [55] Li, C., Yuan, B., Yu, X., *et al.* (2020) SLC19A1 May Serve as a Potential Biomarker for Diagnosis and Prognosis in Osteosarcoma. *Clinical Laboratory*, **66**. <https://doi.org/10.7754/Clin.Lab.2020.200246>
- [56] Wallmann, T., Zhang, X.M., Wallerius, M., *et al.* (2018) Microglia Induce PDGFRB Expression in Glioma Cells to Enhance Their Migratory Capacity. *iScience*, **9**, 71-83. <https://doi.org/10.1016/j.isci.2018.10.011>
- [57] Miao, L., Yin, R.X., Zhang, Q.H., *et al.* (2019) A Novel circRNA-miRNA-mRNA Network Identifies circ-YOD1 as a Biomarker for Coronary Artery Disease. *Scientific Reports*, **9**, Article No. 18314. <https://doi.org/10.1038/s41598-019-54603-2>