

# Establishment and Application of the Differential Expression Profiles of Long Non-codingRNA Associated with Colorectal Adenomas

Baolong Jia Liqiong Wang\* Huirong Zhu

Department of Pathology, Yan'an Hospital Kunming City, Kunming, Yunnan, 650000, China

## Abstract

**Objective:** To screen the differentially expressed long non-codingRNA (LncRNA) in colorectal adenoma tissue samples, predict its target genes, and explore the biological functions and enriched signal pathways of the target genes. **Methods:** Collect fresh samples of adenomas during clinical colorectal surgery, and construct LncRNA differential expression profiles using high-throughput sequencing technology, Bbioinformatics technology screens differentially expressed LncRNA and predicts target genes, GO and KEGG enrichment analysis respectively predict the biological function of the target gene and the signal pathway of enrichment. **Results:** Compared with normal tissue samples, colorectal adenoma tissue samples were screened for 219 significantly differentially expressed LncRNAs (119 were up-regulated and 99 were down-regulated). Functional analysis showed that there were 1242 GO entries related to colorectal adenoma; 32 KEGG-enriched signal pathways. **Conclusion:** The differentially expressed LncRNA in colorectal adenoma samples may participate in the regulation of multiple biological functions and colorectal adenoma-related signal pathways through targeted regulation of its target genes, and may play an important role in the formation of colorectal adenomas.

## Keywords

colorectal adenoma; long non-codingRNA; gene

# 结直肠腺瘤相关的长链非编码 RNA 差异表达谱建立及应用

贾宝龙 王丽琼\* 朱惠荣

昆明市延安医院病理科, 中国·云南昆明 650000

## 摘要

**目的:** 筛选结直肠腺瘤组织样本中具有差异表达的长链非编码RNA (long non-codingRNA, LncRNA), 预测其靶基因, 探究靶基因的生物学功能及富集的信号通路。**方法:** 收集临床结直肠手术中腺瘤新鲜样本, 高通量测序技术构建LncRNA差异表达谱, 生物信息学技术筛选差异表达的LncRNA并进行靶基因预测, GO与KEGG富集分析分别预测靶基因的生物学功能及富集的信号通路。**结果:** 与正常组织样本相比, 结直肠腺瘤组织样本中筛选出219个显著差异表达LncRNA (119个上调, 99个下调)。功能性分析表明: 与结直肠腺瘤相关的GO条目1242条; KEGG富集的信号通路32条。**结论:** 结直肠腺瘤样本中差异表达的LncRNA可能通过靶向调控其靶基因参与调控多种生物学功能及结直肠腺瘤相关的信号通路, 在结直肠腺瘤形成过程中可能发挥重要作用。

## 关键词

结直肠腺瘤; 长链非编码RNA; 基因

## 1 引言

结直肠癌 (colorectal cancer, CRC) 作为一种常见的消化道恶性上皮肿瘤, 在全世界范围内严重威胁着人类的健康。根据 2018 年世界卫生组织发布的全球肿瘤流行病学统计数据 (GLOBOCAN 2018) 结果显示结直肠癌的发病率在最

常见的癌症中居第三位, 死亡率居第二位<sup>[1]</sup>。根据中国的国家癌症中心发表的《2015年中国癌症统计数据》结果显示, 在中国常见癌症中结直肠癌发病率排第五位, 以腺癌居多, 并呈逐年上升趋势<sup>[2]</sup>。然而, 结直肠癌大多由腺瘤发展而来。腺瘤被认为是癌前病变, 具有较高的癌变风险<sup>[3]</sup>。

最初, 许多人假设腺瘤是结直肠癌的前体, 长期以来病理学家也观察到浸润性结肠直肠癌表面和邻近组织存在腺瘤组织, 二者存在一定的连续性。到 1958 年, 腺瘤会发展成为癌这一假设被广泛接受, 尽管最初还有很多争议。1990 年 Fearon 和 Vogelstein 通过对腺瘤的大量遗传学研究

【作者简介】贾宝龙 (1985-), 男, 满族, 中国内蒙古赤峰人, 病理中级, 从事病理诊断研究。

【通讯作者】王丽琼 (1968-), 女, 白族, 中国云南昆明人, 病理正高级, 从事病理诊断研究。

到大腺瘤,再到高级别变化的腺瘤,再到浸润性癌,遗传异常的数量逐渐增加。随后,相关研究表明,完整的结肠镜检查及切除腺瘤可降低癌症的发生率<sup>[4]</sup>,而且这种筛查程序现在已成为常规。此外,结直肠癌腺瘤已被接受为结直肠癌发生过程中基因事件的端点<sup>[5]</sup>。

近年来,作为基因研究的热点,长链非编码RNA(long non-codingRNA, LncRNA)在多层次调控生物体的基因表达,参与生物体内多种重要的生物学过程<sup>[6]</sup>。研究指出,LncRNA具有复杂的生物学功能与作用调控机制,通过介导基因表达、转录、翻译以及组织分化等途径调控基因表达,引起细胞功能紊乱、代谢失调,导致多种疾病和癌症的形成<sup>[7-8]</sup>。目前,有研究指出LncRNA在多种肿瘤发生过程中异常表达,参与调控肿瘤细胞的增殖、迁移、侵袭与凋亡过程,在肿瘤发生发展中发挥致癌或抑癌作用,是肿瘤发生的重要因素之一<sup>[9-10]</sup>。但是,LncRNA在结直肠腺瘤中作用机制的研究相对较少。因此,对结直肠腺瘤发病机制的深入研究至关重要。我们将高通量LncRNA测序技术与生物信息技术相结合,筛选结直肠腺瘤组织样本与正常组织样本中显著差异表达的LncRNA,对其进行靶基因预测,并对靶基因进行GO功能分析与KEGG信号通路分析,预测差异表达的LncRNA可能具有的功能,深入探讨其对结直肠腺瘤发生发展的调控机制,为结直肠腺瘤的临床研究提供数据与理论基础。

## 2 材料和方法

### 2.1 组织标本

结直肠管状腺瘤组织标本源于2019年6月—2019年12月于中国云南省昆明市延安医院诊治的且经术后病理诊断为结直肠低级别管状腺瘤(Ⅱ级)的患者,共计3例。收集患者腺瘤组织及腺瘤旁组织,取完立即放入液氮或-80℃冰箱保存。本研究已获得昆明市延安医院伦理委员会审批。所有被采集的患者均签署了知情同意书。

### 2.2 主要试剂与仪器

RNA提取试剂盒(北京天根生物科技有限公司)、普通PCR与荧光定量PCR相关试剂(宝生物工程(大连)有限公司)、Nanodrop微量分光光度计(美国Thermo Fisher公司)、Qubit荧光定量仪(美国Thermo Fisher公司)、Agilent2100生物芯片分析系统(美国Thermo Fisher公司)。

### 2.3 实验方法

#### 2.3.1 组织样本中RNA的提取

依据试剂说明书的标准操作流程进行。采用TRIzol总RNA抽提试剂提取结直肠管状腺瘤组织标本及正常黏膜组织标本的总RNA,经电泳质检合格后纯化总RNA。而后,经Nanodrop微量分光光度计及Agilent2100对其浓度进行测定,合格的RNA可用于后续测序实验。

#### 2.3.2 高通量LncRNA测序

对质检合格的RNA构建cDNA文库,按有效浓度及

目标上机数据量的需求进行Illumina测序。通过FASTQC(v0.11.5)软件对原始数据进行质量评估,使用Cutadapt(v1.10)软件去除接头序列,再使用Hisat2(v2.0.5)软件将读序比对至基因组,使用Stringtie(v1.3.3b)软件进行组装。使用R语言中DESeq2对结直肠管状腺瘤组织标本及正常黏膜组织标本的LncRNAs的差异表达进行分析。基因表达的差异用P值和差异倍数(fold change, FC)的对数(log FC)表示。设置参数 $P < 0.05$ 且 $|\log_2FC| > 1$ 为阈值,筛选差异表达的LncRNAs。对上述获得的LncRNAs进行GO(Gene Ontology)功能分析和KEGG(Kyoto Encyclopedia of Genes and Genome)通路分析。

### 2.4 统计学方法

实验数据采用Graph Pad Prism 6.0软件分析,结果以平均数±标准差( $\bar{x} \pm s$ )表示,组间差异用单因素方差分析, $P < 0.05$ 认为差异具有统计学意义。

## 3 结果

### 3.1 结直肠腺瘤组织样本与正常组织样本差异表达基因

采用高通量LncRNA测序技术对组织样本进行测序分析,结果表明与正常黏膜相比结直肠腺瘤组织样本中共筛选出218个显著差异表达的LncRNA,119个基因表达上调,99个基因表达下调。按照pval值排序列出了上、下调各Top10的差异表达LncRNA(见表1);筛选出488个显著差异表达的mRNA,234个基因表达上调,254个基因表达下调,按照pval值排序列出了上、下调各Top10的差异表达mRNA(见表2)。并绘制了显著差异表达的LncRNA的火山图(见图1)。

### 3.2 结直肠腺瘤组织样本与正常组织样本差异表达的LncRNA靶基因预测

本研究中,采用Trans靶基因预测原理根据LncRNA同mRNA的表达量相关性系数进行筛选取相关系数(correlation)绝对值大于0.95且pval小于0.01的mRNA基因作为差异LncRNA的靶向基因进行功能富集分析,进一步使用生物统计学方法筛选出5个与靶向mRNA相关性最高的差异表达的LncRNA。其中,图2是根据LncRNA和mRNA的共表达关系建立了LncRNA-mRNA共表达网络。具有最多共同表达mRNA的差异LncRNA ENSG00000233766、ENSG00000254510、ENSG00000176124、MSTRG.233559、MSTRG.43129等是网络中最重要的LncRNA,构成共表达网络的核心调控模块。推测组织样本中极显著差异表达的LncRNA可能通过与其靶基因之间相互调控发挥作用,从而调控结直肠腺瘤的发生进程。

### 3.3 结直肠腺瘤组织样本与正常组织样本差异表达LncRNA的靶基因GO分析

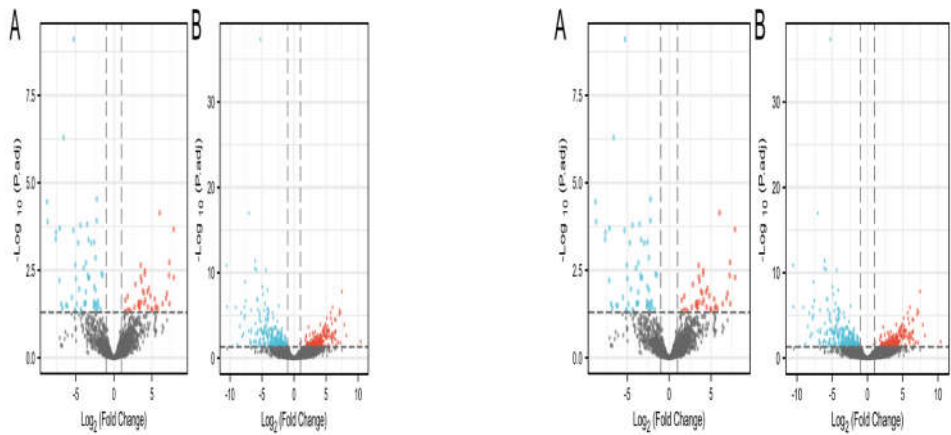
采用GO功能注释方法分析显著差异表达的LncRNA

表 1 腺瘤与正常组织相关差异表达 LncRNA ( Top10 列表, 按 pval 值排序 )

Gene	baseMean	Log <sub>2</sub> FC	pval	GeneName	Up/Down
MSTRG.197328	1056.985	7.5236	7.90E-09	—	up
MSTRG.145746	868.705	1.8487	9.96E-09	—	up
ENSG00000261616	218.593	6.0251	1.71E-07	AC036108.3	up
MSTRG.31568	307.527	6.9058	3.46E-07	—	up
MSTRG.84848	18.468	7.8124	5.27E-07	—	up
ENSG00000259478	37.440	7.8658	6.50E-07	AC024651.1	up
MSTRG.216969	205.571	4.1053	8.69E-07	—	up
MSTRG.230199	37.984	2.6797	3.11E-06	—	up
MSTRG.31527	284.113	4.0074	3.63E-06	—	up
MSTRG.185330	516.446	1.7182	4.63E-06	—	up
ENSG00000226320	237.890	-5.2736	2.38E-13	LINC01811	down
MSTRG.146827	148.095	-8.0941	1.12E-11	—	down
MSTRG.110080	135.720	-9.4315	1.41E-11	—	down
MSTRG.9372	110.165	-7.6351	9.52E-11	—	down
ENSG00000214039	108.864	-6.6196	4.03E-10	LINC02418	down
MSTRG.161726	109.605	-3.8878	3.37E-09	—	down
ENSG00000233695	324.749	-2.2259	6.14E-08	GAS6-AS1	down
ENSG00000228742	87.350	-8.7949	7.52E-08	LINC02577	down
MSTRG.47150	29.760	-6.3399	1.01E-07	—	down
ENSG00000262001	221.864	-2.2498	3.10E-07	DLGAP1-AS2	down

表 2 腺瘤与正常组织相关差异表达 mRNA ( Top10 列表, 按 pval 值排序 )

Gene	baseMean	Log <sub>2</sub> FC	pval	GeneName	Up/Down
ENSG00000181072	470.093	7.4623	7.40E-12	CHRM2	up
ENSG00000130176	5774.454	6.0103	1.11E-09	CNN1	up
ENSG00000175084	22544.565	7.0886	2.78E-09	DES	up
ENSG00000163017	19116.860	5.7144	6.19E-09	ACTG2	up
ENSG0000080224	111.390	7.2394	6.38E-09	EPHA6	up
ENSG00000034971	58.586	7.0468	1.22E-08	MYOC	up
ENSG00000158246	211.246	5.1297	3.74E-08	TENT5B	up
ENSG00000133392	77964.572	5.7723	4.48E-08	MYH11	up
ENSG00000075073	711.287	6.3742	5.51E-08	TACR2	up
ENSG00000089250	212.826	6.5082	8.38E-08	NOS1	up
ENSG00000003249	189.620	-3.0735	9.51E-07	DBNDD1	down
ENSG00000103888	1113.944	-5.2664	1.68E-42	CEMIP	down
ENSG00000167767	191.319	-7.0340	7.87E-22	KRT80	down
ENSG00000062038	648.690	-6.0703	4.80E-16	CDH3	down
ENSG00000137673	540.289	-10.5303	2.11E-15	MMP7	down
ENSG00000149968	684.769	-6.1223	5.57E-15	MMP3	down
ENSG00000164379	144.819	-4.3702	1.21E-14	FOXQ1	down
ENSG00000178773	264.229	-5.8666	1.27E-14	CPNE7	down
ENSG00000165376	1121.813	-7.6839	4.16E-13	CLDN2	down
ENSG00000163347	734.814	-4.9649	1.71E-12	CLDN1	down

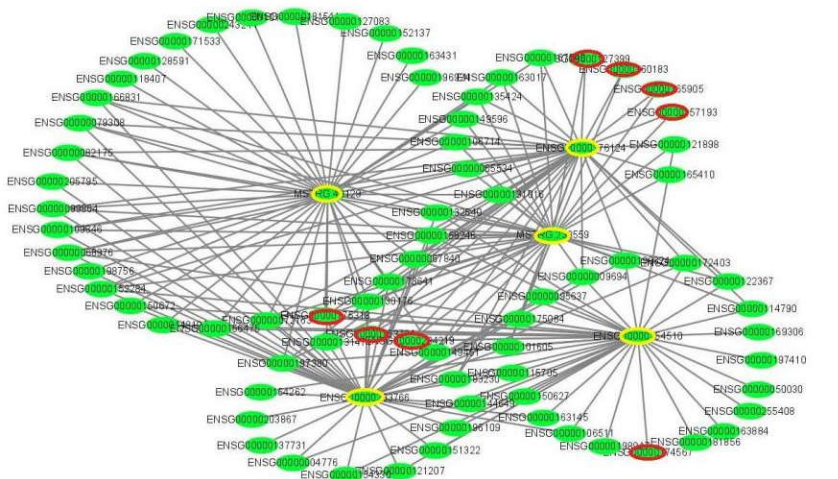


(A) 腺瘤组与正常组中差异表达 lncRNAs

(B) 腺瘤组与正常组中差异表达 mRNAs

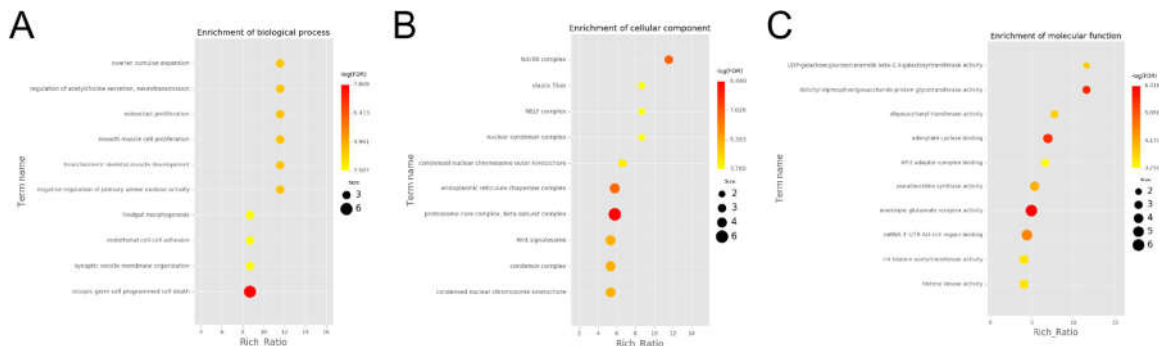
注: X轴代表差异表达倍数的 log<sub>2</sub> 值, 纵坐标代表各个基因的经过标准化后的 P 值。显著上调的基因 (|log<sub>2</sub>(差异倍数)| > 1) 为红色标识, 显著下调的基因为蓝色标识, 表达量没有变化的基因为灰色标识。

图 1 差异表达基因火山图



注: 有黄色标记的为 lncRNA, 有红色标记的为下调的 mRNA, 纯绿色无标记的为上调的 mRNA。

图 2 差异 lncRNA 与其靶向 mRNA 基因的共表达网络



注: 横坐标表示富集程度, 纵坐标表示 GO 条目。

图 3 差异 lncRNA 靶基因的 GO 功能富集条形图

的靶基因富集的生物学功能。GO 注释结果按照  $p\text{-val} < 0.05$  的标准，共得到显著性富集的 1242 个 GO 条目，其中生物过程 (Biological process, BP) 条目 955 个，细胞成分 (Cellular component, CC) 条目 177 个，分子功能 (Molecular function, MF) 条目 110 个。分别可视化了各个条目中，GO 项目的富集因子 (rich factor) 和富集显著性 (见图 3)。推测以上生物学功能的紊乱或异常可能引起结直肠腺瘤的发生。

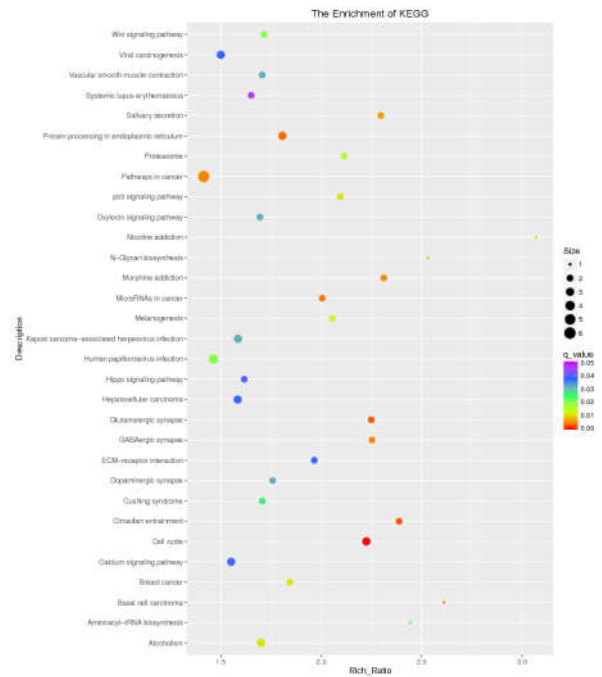
### 3.4 结直肠腺瘤组织样本与正常组织样本差异表达 LncRNA 的靶基因 KEGG 分析

通过 KEGG Pathway 富集分析以  $q\text{-value} < 0.05$  为富集显著性，将差异表达 LncRNA 的靶向进行 KEGG 通路的富集，以明确 LncRNA 调控的生物代谢和信号转导途径。差异表达的 LncRNA 的靶基因共富集了 32 条信号通路 (见图 4、表 3)，其中信号通路 map04110、map04713、map04724、map04141、map05206、map05032、map05200、map05217、map04727、map04970，这些通路大多与结直肠腺瘤的发生密切相关。推测在结直肠腺瘤中这些信号通路在不同程度上发挥调控作用，影响结直肠腺瘤的发生发展。

### 3.5 差异 LncRNA 的靶基因的蛋白互作分析

以差异 LncRNA 的靶向 mRNA 基因在 STRING 数据库中进行蛋白互作分析，结果得到的 PPI 网络中共包含 312 个节点 (node) 和 509 条连线 (edge)，以最高置信度 (0.9) 评估蛋白质之间的相互作用，去掉周边的散点，以不同颜色的线和线的条数代表不同节点之间的相互作用类别，与某个

节点相连的节点和边越多，节点的价值也就越大，这些节点在网络中可能处于较为核心的位置 (见图 5)。



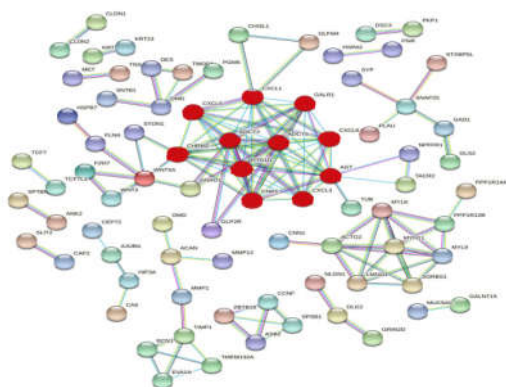
注：横轴表示样本分组名称，纵轴表示 KEGG 生物学通路。每个点表示该 KEGG 条目的富集程度，颜色越趋近于红色表示富集程度越高。每个点的大小表示富集到该 KEGG 条目的基因的个数，点越大表示富集到该 KEGG 条目的基因越多，反之则越少。

图 4 差异 LncRNA 靶基因的 KEGG pathway 气泡图

表 3 显著差异 KEGG pathway 统计表

Name (通路名称)	Map (序列号)	p-val	UpCount	DownCount
Cell cycle	map04110	6.57994E-08	37	12
Circadian entrainment	map04713	1.06154E-05	22	6
Glutamatergic synapse	map04724	1.80509E-05	23	7
Protein processing in endoplasmic reticulum	map04141	3.20177E-05	39	10
MicroRNAs in cancer	map05206	4.96518E-05	26	9
Morphine addiction	map05032	7.67265E-05	19	5
Pathways in cancer	map05200	9.89338E-05	76	34
Basal cell carcinoma	map05217	0.000116362	9	9
GABAergic synapse	map04727	0.000116952	16	8
Salivary secretion	map04970	0.000165305	16	6

注：Name: KEGG 通路的描述信息；Map: KEGG 数据库中通路唯一的编号信息；p-val: p 值；UpCount: 上调的富集基因数；DownCount: 下调的富集基因数。



注：标记为红色的节点在网络中处于核心的位置。

图5 差异 LncRNA 靶基因的蛋白互作分析网络

## 4 讨论

结直肠腺瘤是最常见的结直肠腺癌的癌前病变，约有 90% 的结直肠癌由腺瘤发展而来<sup>[11]</sup>。但是，结直肠腺瘤的发病机制目前仍不完全清楚。因此，进一步研究结直肠腺瘤的分子机制具有重要意义。LncRNA 在结构上与 mRNA 相似，是长度大于 200 bp 的非编码 RNA 不具有编码蛋白质的功能<sup>[12-13]</sup>。其可以在基因水平、转录水平、转录后水平等参与基因表达的调控，与各种疾病的发生发展密切相关<sup>[14]</sup>。长链非编码 RNA，不编码蛋白质，在生物体生长发育、生殖分化过程中发挥重要作用，是目前生物学与医学的热点研究领域<sup>[15]</sup>。研究报道，LncRNA 在表观遗传与基因表达过程中扮演重要角色，参与细胞增殖、分化、凋亡进程<sup>[16-17]</sup>。

LncRNA 与其靶 mRNA 的 3' 非翻译区部分序列通过互补配对的方式结合，调控靶 mRNA 的表达，导致多种肿瘤的发生，在一定程度上影响肿瘤细胞增殖，参与调控细胞病变与癌症恶化过程，在肿瘤增生及患者生存期不同的病理生理特点中发挥作用，表明 LncRNA 可能作为潜在的癌症诊断标记物<sup>[18]</sup>。事实上，LncRNA 在细胞内的功能还不完全清楚，通过高通量测序技术对大量 LncRNA 基因组信息进行分析，研究 LncRNA 和共表达 mRNA 在肿瘤发生发展中的作用。预测潜在的癌症相关 LncRNA 有利于加速新的肿瘤标记物的产生，推动肿瘤的临床研究与治疗。

有研究显示，与正常粘膜相比，PGM5 在腺瘤和 / 或腺癌中下调。此外，他们还表明 PGM5 是筛查大肠腺瘤的一个指标<sup>[19]</sup>。有学者通过代表性差别分析法发现长链非编码 RNA 转录因子 CCAT-1 (结直肠癌相关转录因子 1) 在结直肠腺瘤、结直肠原位癌、结直肠癌侵袭边缘组织、结直肠癌淋巴转移瘤、结直肠癌肝转移瘤中呈逐级增强表达，而在正常结肠、小肠组织内均无表达，表明 CCAT-1 在结直肠腺瘤

的表达具有高度特异性，可作为结直肠腺瘤特异性标志物之一<sup>[20]</sup>。CCAT-1 在结直肠腺瘤的表达，可作为结直肠腺瘤筛查指标，具有重要临床意义。CCAT-1 在结直肠腺瘤患者外周血中高灵敏表达，可作为结直肠腺瘤无创筛查指标之一。另有研究发现，CCAT-1 与结直肠腺瘤的发生发展密切相关，可能参与调控结直肠腺瘤 - 癌序列<sup>[21]</sup>，但需进一步研究明确。Xu 等<sup>[22]</sup>发现 MALAT-1 (人肺腺瘤转移相关转录因子 1) 的 3' 端在结直肠癌细胞增殖、迁移和入侵的生物过程中扮演了一个关键的角色，并可能参与调控结直肠腺瘤 - 癌序列。Graham 等<sup>[23]</sup>发现 CNRDE (结直肠癌形成差异表达基因) 在结直肠腺瘤早期就被激活，呈高度表达，具有高度敏感性和特异性，在结直肠腺瘤 - 癌序列中发挥重要作用。

LncRNAs 作为一类特殊的 RNA 分子以其广泛而复杂的生物学作用参与调控肿瘤的发生发展。因此，阐明新发现的 LncRNA 与已知的 LncRNA 在肿瘤形成过程中具有的生物学功能，将有助于肿瘤的早期诊断和预后治疗。本研究选用高通量 LncRNA 测序技术对结直肠腺瘤组织样本与正常结直肠黏膜样本进行测序分析，分析发现了 LncRNA 在两种组织样本中的差异表达，成功建立了结直肠腺瘤相关的长链非编码 RNA 差异表达谱。采用聚类分析的方法对实验组与对照组差异表达基因进行研究，共筛选出 218 个显著差异表达的 LncRNA，119 个基因表达上调，99 个基因表达下调。对结直肠腺瘤组织样本与正常组织样本差异表达的 LncRNA 靶基因预测，并建立 LncRNA-mRNA 共表达网络。发现前 5 个具有最多共同靶向 mRNA 的差异 LncRNA ENSG00000233766、ENSG00000254510、ENSG00000176124、MSTRG.233559、MSTRG.43129 等是网络中最重要 LncRNA，构成共表达网络的核心调控模块。推测组织样本中极显著差异表达的 LncRNA 可能通过与其靶基因之间相互调控发挥作用。对显著差异表达 LncRNA 的靶基因进行功能性分析发现，结直肠腺瘤相关的 GO 条目有 1242 条，其中生物过程 (Biological process, BP) 955 条，细胞成分 (Cellular component, CC) 177 条，分子功能 (Molecular function, MF) 110 条。差异表达 LncRNA 的靶基因主要参与蛋白酶体核心复合物， $\beta$  亚基复合物、wnt 信号、离子型谷氨酸受体活性、mRNA3' - UTR 富含 AU 的区域绑定、腺苷酸环化酶、异常生殖细胞程序性死亡、内质网伴侣复合物等功能。KEGG 信号通路富集分析结果指出，差异表达的 LncRNA 的靶基因共富集了 32 条信号通路，其中信号通路 map04110、map04713、map04724、map04141、map05206、map05032、map05200、map05217、map04727、map04970 这些通路大多与结直肠腺瘤的发生密切相关，其中包括细胞周期、谷氨酸能突触、

内质网蛋白加工、癌症通路、肿瘤中的 microRNA 等信号通路。推测在结直肠腺瘤中这些信号通路在不同程度上发挥调控作用，影响结直肠腺瘤的发生发展。这些信号通路在其他癌症发生发展中的作用机制已有报道，但其在结直肠腺瘤中的报道相对较少。

因此，深入研究结直肠腺瘤发生进程中的生物学功能变化与相关信号通路的作用机制，有利于更好的揭示结直肠腺瘤的发病机制。本实验以结直肠腺瘤为材料通过高通量测序基于生物信息学技术对结直肠腺瘤中差异表达 LncRNA 进行筛选，对其靶基因的功能进行预测，并未进行大量的实验验证，在接下来的研究中还有许多问题有待解决。

此外，在下一步的研究中，将深入探讨测序结果筛选出的极显著差异表达 LncRNA 参与调控结直肠腺瘤发生发展的具体调控机制，阐明其与靶基因的相互关系以及在结直肠腺瘤相关的信号通路中发挥的作用，以期能为结直肠腺瘤的早期筛查与预后治疗提供新的肿瘤标记物。

#### 参考文献

- [1] Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: OBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. *CA Cancer J Clin*, 2018(68): 4-424.
- [2] Chen W, Zheng R, Baade P D, et al. Cancer statistics in China[J]. *Ca A Cancer Journal for Clinicians*, 2016, 66(2):115-132.
- [3] Bjerrum A, Lindebjerg J, Andersen O, et al. Long-term risk of colorectal cancer after creen-detected adenoma: experiences from a Danish gFOBT-positive screening cohort[J]. *Cancer*, 2020(147): 940-947.
- [4] Citarda F, Tomaselli G, Capocaccia R, et al. Efficacy in standard clinical practice of colonoscopic polypectomy in reducing colorectal cancer incidence[J]. *Gut*, 2001(48):812-815.
- [5] Kelloff GJ, Schilsky RL, Alberts DS, et al. Colorectal adenomas: a prototype for the use of surrogate end points in the development of cancer prevention drugs[J]. *Clin Cancer Res*, 2004(10):3908-3918.
- [6] Mercer TR, Dinger ME, Mattick JS. Long non-coding RNAs: insights into functions[J]. *Nature Reviews Genetics*, 2009,10(3):155-159.
- [7] Clark MB, Mattick JS. Long noncoding RNAs in cell biology[J]. *Seminars in Cell & Developmental Biology*, 2011,22(4):366-376.
- [8] Wilusz JE, Sunwoo H, Spector DL. Long noncoding RNAs: functional surprises from the RNA world[J]. *Genes & Development*, 2009,23(13):1494-1504.
- [9] Thomson DW, Dinger M E. Endogenous microRNA sponges: evidence and controversy[J]. *Nature Reviews Genetics*, 2016,17(5):272-287.
- [10] 杨倩,黄宏斌,龚朝建,等.LncRNAs参与基因表达调控机制的研究进展[J]. *中南大学学报(医学版)*, 2014,39(1):91-95.
- [11] Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis.[J]. *Cell*, 1990,61(5):759-767.
- [12] Khalil, M A , Azhar, et al. Aetiology, maternal and foetal outcome in 60 cases of obstetrical acute renal failure[J]. *Ayub Medical College, Abbottabad*, 2009(21):46-49.
- [13] Wilusz J E, Sunwoo H, Spector, et al. Long noncoding RNAs: functional surprises from the RNA world[J]. *Genes Dev*, 2009(23):1494-1504.
- [14] Kurokawa R, Rosenfeld MG, Glass CK. Transcriptional regulation through noncoding RNAs and epigenetic modifications[J]. *Rna Biology*, 2009,6(3):233-236.
- [15] Yin Q, Wu A, Liu M . Plasma Long Non-Coding RNA (lncRNA) GAS5 is a New Biomarker for Coronary Artery Disease[J]. *Medical Science Monitor International Medical Journal of Experimental & Clinical Research*, 2017(23):6042-6048.
- [16] Kapranov P, Cheng J, Dike S, et al. RNA Maps Reveal New RNA Classes and a Possible Function for Pervasive Transcription[J]. *Science*, 2007,316(5830):1484-1488.
- [17] Hirata H, Hinoda Y, Shahryari V, et al. Long Noncoding RNA MALAT1 Promotes Aggressive Renal Cell Carcinoma through Ezh2 and Interacts with miR-205[J]. *Cancer Research*, 2015, 75(7):1223-1225.
- [18] Koyo N, Reo M, Takeshi N, et al. Screening for long noncoding RNAs associated with oral squamous cell carcinoma reveals the potentially oncogenic actions of DLEU1[J]. *Cell Death & Disease*, 2018, 9(8):826-828.
- [19] Christiana A , Uzozie, Nathalie, et al. Targeted Proteomics for Multiplexed Verification of Markers of Colorectal Tumorigenesis. [J]. *Molecular & cellular proteomics : MCP*, 2017(3):3112-3117.
- [20] Nissan A, Stojadinovic A, Mitrani-Rosenbaum S, et al. Colon cancer associated transcript-1: a novel RNA expressed in malignant and pre-malignant human tissues[J]. *Cancer*, 2012,130(7):1598-1606.
- [21] Sonja I, Berndt John D, Potter Aditi Hazra Meredith Yeager Gilles Thomas Karen W, et al. Pooled analysis of genetic variation at chromosome 8q24 and colorectal neoplasia risk[J]. *Human Molecular Genetics*, 2008,17(17):2665.
- [22] Li, Zuguo. MALAT-1: a long non-coding RNA and its important 3' end functional motif in colorectal cancer metastasis.[J]. *International Journal of Oncology*, 2011(6):132-137.
- [23] Graham LD, Pedersen SK, Brown GS, et al. Colorectal Neoplasia Differentially Expressed (CRNDE), a Novel Gene with Elevated Expression in Colorectal Adenomas and Adenocarcinomas[J]. *Genes, Cancer*, 2011,2(8):434-436.