

DOI:10.12113/202303015

癌症中 lncRNA 调控可变剪接的生物 信息学研究进展

吴 琼,汪玉兰,赵 健*,宋晓峰*
(南京航空航天大学 自动化学院,南京 211106)

摘要:可变剪接是指基因在不同剪接模式下生成 RNA 异构体的过程,是基因转录后表达调控的重要机制。可变剪接已被证实发生在癌症发生发展中扮演着重要角色,而异常剪接被认为是肿瘤发生的重要标志。长链非编码 RNA (lncRNA) 是基因可变剪接过程中的重要参与者和调节者,并且诸多证据表明 lncRNA 可通过调控可变剪接进而影响癌症的发生发展。本文对癌症中 lncRNA 调控的可变剪接相关研究进行了综述,并对相关生物信息学工具及数据库进行了总结。

关键词:lncRNA; 可变剪接; 剪接因子; 癌症; 生信工具

中图分类号:Q811.4 **文献标志码:**A **文章编号:**1672-5565(2024)04-239-13

Bioinformatics advances of lncRNA regulated alternative splicing in cancer

WU Qiong, WANG Yulan, ZHAO Jian*, SONG Xiaofeng*

(College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China)

Abstract: Alternative splicing refers to the process that genes produce RNA isoforms under different splicing modes, and it is an important mechanism for post-transcriptional expression regulation of genes. Alternative splicing has been proved to play an important role in the occurrence and development of cancer, and abnormal splicing is considered as an important symbol of tumor occurrence. Long noncoding RNA (lncRNA) is an important participant and regulator in the process of gene alternative splicing, and many evidences show that lncRNA can affect the occurrence and development of cancer by regulating alternative splicing. In this paper, the research on alternative splicing regulated by lncRNA in cancer is reviewed, and the related bioinformatics tools and databases are summarized.

Keywords: lncRNA; Alternative splicing; Splicing factor; Cancer; Bioinformatics tools

可变剪接作为基因表达调控的重要机制之一,使得同一基因可以通过不同剪接方式产生不同的 RNA 异构体,进而扩大了蛋白质的多样性。研究表明基因的异常剪接与癌症的发生发展及预后密切相关。近来,lncRNA 被发现在基因剪接过程中扮演着重要角色,可通过以下三种方式调控可变剪接^[1]:
①lncRNA 与剪接因子直接或间接相互作用;

②lncRNA 与 pre-mRNA 分子形成 RNA-RNA 双链;
③lncRNA 影响染色质重塑。尽管目前已有部分 lncRNA 被揭示参与剪接调控过程,但由于 lncRNA 数量众多且相关分子机制复杂,lncRNA 在剪接调控中的功能与作用机制仍有待进一步探索。特别是,与癌症异常剪接调控相关 lncRNA 的识别与分子机制研究,将为癌症诊断与治疗策略提供新的见解。

收稿日期:2023-03-17;修回日期:2023-05-09;网络首发日期:2023-07-25.

网络首发地址:https://link.cnki.net/urlid/23.1513.Q.20230724.1850.010

基金项目:江苏省重点研发计划项目(No. BE2022843);国家自然科学基金项目(No. 62273175, 62003165, 61973155).

*通信作者:赵健,男,副研究员,研究方向:模式识别与生物信息学. E-mail: zhaojian@nuaa.edu.cn;

宋晓峰,男,教授,研究方向:生物信息学. E-mail: xfsong@nuaa.edu.cn.

引用格式:吴琼,汪玉兰,赵健,等.癌症中 lncRNA 调控可变剪接的生物信息学研究进展[J].生物信息学,2024,22(4):239-251.

WU Qiong, WANG Yulan, ZHAO Jian, et al. Bioinformatics advances of lncRNA regulated alternative splicing in cancer[J]. Chinese Journal of Bioinformatics, 2024, 22(4): 239-251.

1 lncRNA 概述

lncRNA 是长度超过 200 个核苷酸的非编码 RNA,序列保守性较差^[2]。根据其 与 mRNA 在基因组上的相对位置关系,lncRNA 主要可分为以下几种类型:正义型、反义型、内含子型、基因间型和增强子型^[3]。随着高通量测序技术及计算平台的发展,在过去的十多年间人们发现并注释了大量的 lncRNA。lncRNA 从一开始的“转录噪声”转变成为具有重要生物学功能的调控分子,包括充当分子支架、与 pre-mRNA 或 miRNA 形成 RNA-RNA 双链体作为转录后调节因子,转录后募集或抑制调节蛋白以及编码小肽或产生发挥生物学效应的小 RNA 分子^[2]。

lncRNA 可通过多种机制参与生物过程^[4],其调控功能通常具有以下特征:①lncRNA 的功能与其在细胞中的位置有关。细胞核内的 lncRNA 参与调节诸如转录、剪接和染色质重塑等生物过程,而细胞质中的 lncRNA 则负责调节 mRNA 运输、稳定性,翻译以及翻译后修饰等生物过程;②lncRNA 的功能取决于其相对丰度;③lncRNA 的功能与和其相互作用的 RNA 或 RNA 结合蛋白密切相关。随着研究的不断深入,越来越多的证据表明 lncRNA 在癌症等人类复杂疾病的发生发展中起着关键作用,包括肝癌、胃癌、卵巢癌、胰腺癌和结直肠癌等^[5-9]。尽管已有许多 lncRNA 的功能被发现,但鉴于其数量众多且调控机制复杂,仍有大量 lncRNA 的功能及作用机制未知。

2 lncRNA 调控可变剪接的分子机制

可变剪接是指同一 RNA 前体生成不同 RNA 剪接异构体的过程,可变剪接类型主要包括外显子跳跃,可变 5' 或 3' 剪接位点、互斥外显子和内含子保留^[10]。据估计,92%~94% 的人类基因都会经历可变剪接^[11],表明可变剪接在生命活动中扮演着极其重要的角色。影响剪接调控的因素有很多,近年来随着对 lncRNA 研究的不断深入,研究人员发现 lncRNA 是剪接过程的重要参与者和调节者。

2.1 lncRNA 与剪接因子直接或间接相互作用

催化剪接过程的剪接体复合物包括 100 多种反式作用因子,其中 hnRNP 和 SR 蛋白属于公认的剪接因子(Splicing factor, SF)家族^[12]。剪接因子对于剪接位点的识别以及剪接体的组装至关重要,而 lncRNA 可通过与剪接因子相互作用影响可变剪接(图 1)。一方面,lncRNA 与剪接因子直接相互作用,

例如 MALAT1 通过与 SR 剪接因子结合调节其表达水平和磷酸化,使 SR 剪接因子能够在正确的时间、地点和浓度下完成 pre-mRNA 的可变剪接^[13];另一方面,lncRNA 在转录或转录后水平上间接与剪接因子相互作用,例如 ZNF561-AS1 通过与 miR-26a-3p 和 miR-128-5p 相互作用来调节 SRSF6 的表达水平,而 SRSF6 作为 SR 剪接因子家族成员,其对 pre-mRNA 的剪接至关重要^[14]。此外,剪接因子调节靶基因剪接通常由转录因子(Transcription factor, TF)触发,例如 c-Myc 通过上调 PTB, hnRNPA1 和 hnRNPA2 的转录进而调节丙酮酸激酶 PKM2 亚型的剪接^[15]。lncRNA 作为转录调控分子,可与转录因子相互作用,进而影响下游剪接因子对基因的可变剪接。

2.2 lncRNA 与 pre-mRNA 分子形成 RNA-RNA 双链

lncRNA 可与 pre-mRNA 互补配对形成 RNA-RNA 双链体,进而掩盖 pre-mRNA 内部与剪接因子结合所必需的顺式作用元件,改变其剪接形式(图 2)。2008 年 Beltran M 等^[16]发现 ZEB2 的天然反义转录物(Natural antisense transcript, NAT)与 ZEB2 5'-UTR 的内含子区域互补形成 RNA-RNA 双链,从而防止了 ZEB2 5'-UTR 的剪接,进而增加了 ZEB2 蛋白的表达水平。此后,人们陆续发现许多反义 lncRNA 与 pre-mRNA 形成 RNA-RNA 双链,进而影响剪接包含或去除相关外显子,从而产生相应的致癌转录本。lncRNA Saf 通过与 Fas pre-mRNA 碱基互补配对,形成特定的双链 RNA 中间体。进一步的研究发现,二者之间最强的相互作用发生在外显子 5-6 和外显子 6-7 连接处,从而导致剪接因子 SPF45 选择性剪接 Fas 外显子 6^[17]。lncRNA PLANE 通过与 NCOR2 pre-mRNA 形成 RNA-RNA 双链,招募 hnRNPM,从而抑制剪接产物 NCOR2-202 的生成,降低其蛋白质水平,促进癌细胞的增殖^[18]。

2.3 lncRNA 影响染色质重塑

lncRNA 可以与染色质相互作用,然后招募蛋白质复合物来重塑染色质状态,从而调节基因表达^[19](图 3)。SWI/SNF 复合物是重要的染色质重塑因子,其利用 ATP 水解的能量来改变染色质结构^[20]。MALAT1 将染色质重塑复合物 SWI/SNF 的催化亚基 BRG1 招募到促炎介质 IL-6 和 CXCL8 的启动子区域,从而促进 NF- κ B 诱导这些炎症因子的表达^[21]。Pandey RR 等^[22]发现,Kcnq1ot1 与 PRC2 复合物相互作用,并将 H3K9 和 H3K27 甲基转移酶募集到作用部位,从而调节染色质水平。其次,研究表明,染色质重塑在可变剪接的调控中具有重要功能^[23]。Sims RJ

3rd 等^[24]发现 ATP 依赖性染色质重塑蛋白和组蛋白伴侣 CHD1 通过与 U2 snRNP 的 SF3a 亚复合物的特异性相互作用将核心剪接体组分连接到 H3K4me3。当 CHD1 和 H3K4me3 减少时,SF3a 的募集和 pre-mRNA 剪接的效率均受到干扰^[24]。

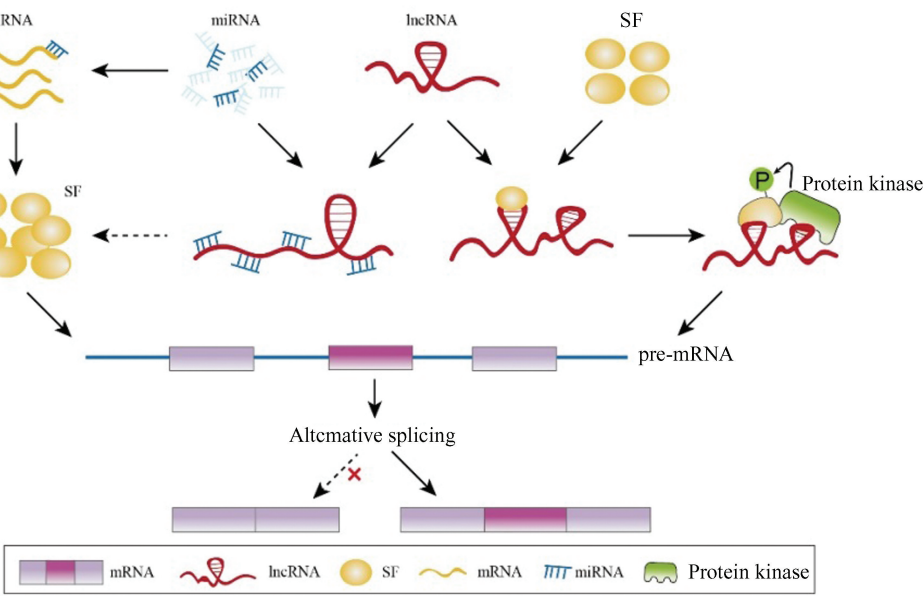


图 1 lncRNA 与剪接因子交互

Fig.1 Interaction between lncRNA and splicing factor

注:一方面,lncRNA 作为脚手架,通过介导蛋白激酶与剪接因子间的相互作用影响后者的磷酸化修饰水平,进而调节下游基因的可变剪接;另一方面,lncRNA 可海绵吸附 miRNA,进而降低 miRNA 与其靶基因间的结合,提高剪接因子的表达水平,从而影响下游基因的剪接过程。

表达增多的方向发展。Ma Y 等^[26]发现 PAXIP1-AS1 通过 miR-6744-5p 上调剪接因子 PCBP2 的表达,从而加剧卵巢癌细胞的生长和迁移。Wu J 等^[27]的研究表明 lncRNA CRNDE 通过 SRSF1/TIA1 信号通路诱导卵巢癌顺铂耐药,但其具体调控机制仍需进一步研究。

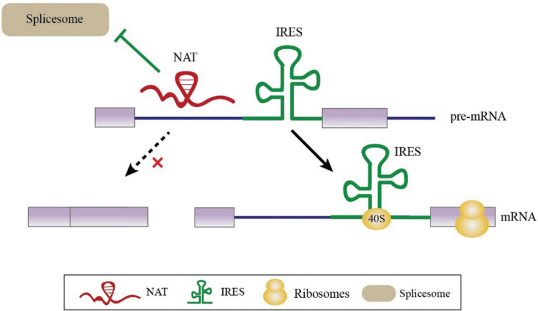


图 2 pre-mRNA 与其天然反义转录产物 NAT 形成 RNA-RNA 双链
Fig2 pre-mRNA and its natural antisense transcript (NAT) form RNA-RNA double strand

注:RNA-RNA 双链体的形成使得 pre-mRNA 的内含子在剪接过程中得以保留,此区域内的 IRES 元件可直接招募核糖体起始翻译过程。

3 lncRNA 与癌症异常剪接

3.1 卵巢癌

卵巢癌是一种异质性肿瘤,其死亡率超过任何其他类型的女性生殖系统恶性肿瘤^[25]。卵巢癌发生发展的分子机制甚为复杂,至今仍不明确。Gordon MA 等^[7]发现 lncRNA MALAT1 敲低会出现细胞增殖,侵袭和失巢细胞增多的情况,并且抑制 MALAT1 还会导致剪接因子 RBFOX2 的表达减少,进而调控可变剪接朝着 KIF1B 促凋亡亚型 KIF1B-β

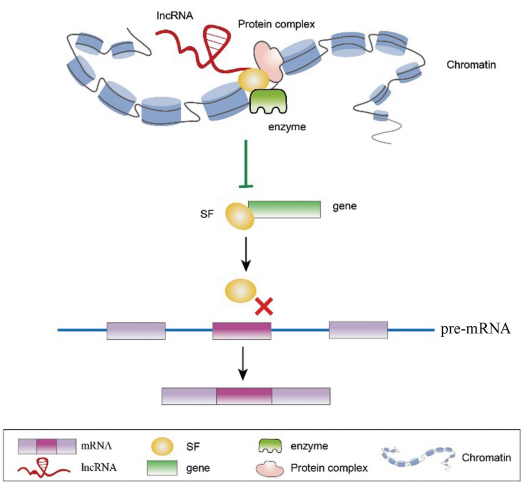


图 3 lncRNA 介导的染色质重塑

Fig.3 lncRNA mediated chromatin remodeling

注:lncRNA 将相关蛋白质复合物和酶募集到染色质,使其结构发生改变,进而抑制剪接因子与基因外显子的结合,使得该外显子在剪接过程中得以保留。

3.2 结直肠癌

结直肠癌是全球第三大常见癌症,也是第四大常见的癌症死亡原因^[28]。为确定新的治疗靶点,探索其发生发展的分子机制是十分迫切的。Lan Z 等^[29]发现 lncRNA SNHG6 可以诱导 hnRNP A1 特异性剪接 mRNA PKM,从而增加 PKM2/PKM1 的比例,这可能是增强结肠癌细胞中有氧糖酵解进而致癌的重要机制。Hu ZY 等^[30]研究表明,lncRNA MALAT1 介导的 SRPK1 过表达会导致细胞核中 SRSF1 磷酸化水平升高,而核磷酸化-SRSF1 水平的恢复可促进 AKAP-9 pre-mRNA 的可变剪接,提高 AKAP-9 蛋白的表达,进而增强结直肠癌细胞的增殖、迁移和侵袭能力。Kong J 等^[31]则发现 lncRNA LINC01133 可与剪接因子相互作用促进结直肠癌中 EMT 和转移的调节。

3.3 胃癌

胃癌是全球第五大常见癌症,由于多数胃癌患者在确诊时已处于晚期,因此死亡率极高,位于癌症相关死亡的第三位^[32]。Zhang F 等^[33]研究发现,lncRNA CRNDE 可通过与剪接因子 SRSF6 相互作用降低其蛋白质稳定性,而 SRSF6 可调节 PICALM 的可变剪接使之产生外显子 14 跳跃的剪接异构体,进而触发亚型转换,促进 PICALM 长亚型的表达,从而减弱胃癌的化学耐药性。He Q 等^[34]发现 LINC00924 调控剪接因子 hnRNPC 在 e14a 处与 Mnk2 pre-mRNA 结合,使得 Mnk2a 亚型表达减少,进而调控 p38 MAPK/PPAR α 信号通路,促进胃癌腹膜转移。Zhu W 等^[35]发现 lncRNA SNHG8 通过与 hnRNP A1 结合促进 TROY 的表达,而 TROY 已被证实会促进胃癌细胞的化学抗性。

3.4 肝癌

肝细胞癌是最常见的原发性肝癌,被列为第六大常见肿瘤,并且是肝硬化患者死亡的主要原因^[36]。Gonzalez I 等^[37]阐明了 lncRNA AS-FGFR2 通过诱导多梳抑制复合体 2(PRC2)的 SUZ12 和 EZH2 与 FGFR2 基因座结合,进而募集 H3K36 去甲基化酶 KDM2a,促进外显子 IIIb 包含。Amann T 等^[38]研究表明含有 IIIb 的 mRNA FGFR2-IIIb 会抑制肝细胞癌细胞的致癌性。Lin YH 等^[39]首次证明 lncRNA LINC01348 在肝细胞癌中显著下调并与患者生存密切相关,其通过抑制剪接因子 SF3B3 介导的 EZH2 可变剪接来调节肝癌细胞的转移,表明其作为生物标志物的效用。另外,Malakar P 等^[40]报道了 lncRNA MALAT1 在肝细胞癌中上调进而诱导剪接因子 SRSF1 调节其剪接靶点导致抗凋亡亚型的产生。另一方面,lncRNA MALAT1 通过调节 S6K1 的可变剪接激活 mTOR 途径,这可能在肝细胞癌发生中发挥着

关键作用。

3.5 胰腺癌

胰腺癌是全球癌症死亡的主要原因,随着全球人口老龄化,胰腺癌的发病率日益升高^[41]。Li J 等^[8]研究发现 lncRNA DKFZp434J0226 的表达增加与胰腺导管腺癌的预后不良呈正相关,进一步的内源性 RNA pull-down 实验与质谱分析揭示了 DKFZp434J0226 与剪接因子 SF3B6 相互作用从而促进其磷酸化,进而影响 MDM4 基因的可变剪接,使得致癌转录本 MDM4-S 的表达升高。类似的,Meng LD 等^[42]研究表明 LINC01232 通过上调剪接因子 HNRNP A2B1 参与 A-Raf 的可变剪接,使得 A-Raf 全长亚型的 mRNA 水平增加和 A-Raf 短亚型的 mRNA 水平下调,并通过激活 MAPK/ERK 信号通路促进胰腺癌细胞转移。Zhang Y 等^[43]研究证实 LINC00857 通过剪接因子 SRSF1 调节 CLDN12 的可变剪接,增加其蛋白产物表达水平,进而促进胰腺癌细胞的上皮-间质转化。

3.6 其他癌症

除以上癌症外,lncRNA 在其他癌症的异常剪接调控中也扮演着重要角色。Xu S 等^[44]发现 lncRNA EGOT 通过 RNA-蛋白质互作招募剪接因子 hnRNPH1,促进 ITPR1 的选择性剪接并提升其表达水平,进而增加乳腺癌细胞自噬水平,提高了肿瘤对紫杉醇化疗的敏感性。Singh R 等^[45]在乳腺癌中发现 BC200 可招募剪接因子 hnRNP A2/B1 改变 Bcl-x 基因的剪接模式,使其蛋白产物由促凋亡亚型 Bcl-xS 转为抗凋亡亚型 Bcl-xL,进而起到致癌作用。Duan Y 等^[46]发现 lncRNA DGCR5 在食管鳞状细胞癌中与 SRSF1 相结合继而调节 mRNA Mcl-1 的可变剪接,促使其抗凋亡亚型 Mcl-1L 的表达增加。Zhang Z 等^[47]对前列腺特异性 lncRNA PCGEM1 进行了研究,发现其可与包括 hnRNP A1 和 U2AF65 在内的剪接因子相互作用,调节 AR 基因的可变剪接模式从而产生不同的功能后果。

3.7 小结

可变剪接异常往往在肿瘤发生发展中起着主导作用^[48],这些异常剪接事件的发生会导致一些关键基因,诸如 TP53^[49],RB1^[50]和 PTEN^[51]等的失调,使其功能丧失。基因剪接异常会增加肿瘤细胞的异质性^[52],对肿瘤细胞和微环境的代谢重编程、增殖、转移及耐药性均有显著影响,并以此驱动肿瘤的发生发展^[53-54]。目前,剪接调控相关研究主要集中在剪接因子^[55-57],lncRNA 可与剪接因子的相互作用,进而已被认为是调控剪接事件的重要参与者^[58-59]。特别是,lncRNA 可作为 miRNA 的分子海绵,而后者

可靶向调控剪接因子表达^[60–62]。此外,lncRNA 还可通过与 pre-mRNA 分子形成 RNA-RNA 双链或影响染色质重塑两种机制,调控可变剪接事件。鉴于其调控过程复杂且研究手段有限,目前仍需更多的研究来深入探索其分子机制。lncRNA 在肿瘤异常剪接中的功能作用研究,或可为肿瘤的诊断、治疗及预防提供新的思路和研究方向,有助于深入理解肿瘤

发生发展的分子机制。

4 生物信息学工具及数据库

4.1 生物信息学工具

lncRNA 调控可变剪接研究相关的生物信息学工具见表 1。

表 1 lncRNA 调控可变剪接研究相关生信工具			
Table 1 Bioinformatics tools for the study of alternative splicing regulated by lncRNA			
用途	工具名称	模型	网址
可变剪接分析	rMATS ^[63]	层次模型、似然比检验	https://rmats.sourceforge.io
	TCGA SpliceSeq ^[67]	RNASeq,SpliceSeq	http://bioinformatics.mdanderson.org/TCGASpliceSeq
	LeafCutter ^[68]	狄利克雷多项式-广义线性模型	https://github.com/davidaknowles/leafcutter
	SUPPA2 ^[69]	转录本定量	https://github.com/comprna/SUPPA
	FLAIR ^[70]	基于三代测序	https://github.com/BrooksLabUCSC/flair
	TargetScan ^[71]	规范定位定量模型	www.targetscan.org
lncRNA-RNA 互作分析	RNAhybrid ^[72]	RNA 二级结构预测	http://bibiserv.techfak.uni-bielefeld.de/rnahybrid
	LncTar ^[73]	自由能最小化	http://www.cuilab.cn/lncstar
	RIBlast ^[74]	RNA 二级结构能量	https://github.com/fukunagatsu/RIBlast
	RIsearch2 ^[75]	GUUGle, RNAplfold, siRNA 脱靶预测	https://rth.dk/resources/risearch
lncRNA-蛋白质互作分析	miRWalk ^[76]	随机森林	http://mirwalk.umm.uni-heidelberg.de
	CatRAPID ^[78]	二级结构、氢键、范德华作用	http://service.tartagliab.com/page/catrapid_omics2_group
	LncADeep ^[79]	深度学习	https://github.com/cyang235/LncADeep
	Capsule-LPI ^[80]	深度学习	www.csbj-jlu.site/lpc/predict
lncRNA-染色质重塑分析	ChIRP-seq ^[81]	免疫共沉淀	PMID: 21963238
	CHART-seq ^[82]	RNA 杂交	PMID: 34856126

4.1.1 差异可变剪接分析

rMATS 是目前最常用的针对 RNA-seq 数据的差异可变剪接分析工具,由 Shen S 等^[63]于 2014 年开发。与先前其它工具^[64–65]相比,rMATS 是首个可处理生物学重复样本 RNA-seq 数据的工具。rMATS 采用层次框架来同时考虑单个重复样本的采样不确定性和重复样本间的变异性对计算结果的影响。rMATS 不仅可以处理非配对生物学重复样本 RNA-seq 数据,还可计算配对生物学重复样本 RNA-seq 数据的差异剪接事件。

SpliceSeq 是 Ryan MC 等^[66]开发的用于探索 TCGA 数据库中肿瘤组织可变剪接模式的标准工具,分析结果收录于 TCGA SpliceSeq^[67]数据库中。TCGA SpliceSeq 包含来自 33 种不同类型的肿瘤样本(包括可用的相邻正常样本)上的剪接事件的剪接百分比(Percent spliced in, PSI),研究人员可以通过输入基因名来寻找所选肿瘤类型之中表现出最高剪接比的基因,或者探索肿瘤和相邻正常样本之间的剪接模式变化。TCGA SpliceSeq 还提供剪接模式,

reads 计数及各种统计结果的直观图形。

LeafCutter 是 Li YI 等^[68]在 2018 年开发的从短读长 RNA-seq 数据中识别可变剪接事件的软件,其无需转录本注释信息。LeafCutter 采用基于计数的统计模型,使用狄利克雷多项式-广义线性模型来检验整个样本集中的内含子使用变化,而不是分别在不同条件或基因型下检验内含子的差异切除。与其他方法不同,LeafCutter 通过测序数据中拼接的读段来识别和量化可变剪接事件,可发现高复杂性的事件。此外,LeafCutter 还可用于绘制剪接数量性状基因座(sQTL)。

SUPPA2 是由 Trincado JL 等^[69]开发并与 LeafCutter 同年发表的可变剪接分析工具,其主要基于转录本定量来识别可变剪接事件,具有计算快和存储要求低的优点。SUPPA2 可根据转录本注释生成剪接事件,定量剪接事件或转录本的剪接水平(PSI),识别不同条件下的差异剪接事件,以及根据 PSI 值对剪接事件进行聚类。SUPPA2 同样可以处理具有生物学重复的 RNA-seq 数据。

FLAIR 是 Tang AD 等^[70]开发的一套适用于纳米孔 cDNA、native RNA 及 PacBio 测序数据,分析差异可变剪接事件或剪接异构体的工具。FLAIR 可选择与短读长测序数据一起运行,以提高长读长测序 reads 的剪接位置识别精度。FLAIR 流程主要包括比对、校正、折叠、量化,差异表达及差异剪接分析,使用多个比对步骤和剪接位点过滤方法提高结果置信度。

4.1.2 lncRNA-RNA 互作分析

TargetScan 是 Lewis BP 等^[71]于 2003 年开发的软件,用于识别 miRNA 的靶标,最后一次更新时间为 2018 年。TargetScan 通过搜索 RNA 序列中与 miRNA 种子区域配对的保守 8mer, 7mer 和 6mer 位点来预测 miRNA-RNA 相互作用,采用位点的累积加权分值对靶点可能性进行评估。对于 lncRNA 和 miRNA 的相互作用预测,用户仅需提供 lncRNA 的原始序列并下载 miRNA 种子序列文件,然后设置相应参数即可进行预测。

RNAhybrid 是 Krüger J 等^[72]在 2006 年发表的可简单、快速和灵活地预测 miRNA 靶标的工具。RNAhybrid 提供在线版本,用户只需上传 miRNA 和候选靶序列即可进行预测,并且有多个参数可供选择。RNAhybrid 的算法核心是 RNA 二级结构预测的变体。与其他二级结构预测软件相比, RNAhybrid 的算法专门为 RNA 杂交而设计,计算速度显著提升。

LncTar 是 Li J 等^[73]在 2015 年开发的一种通过自由能最小化来预测 lncRNA-RNA 相互作用的软件。与其他工具相比, LncTar 具有以下优势:首先, LncTar 对 RNA 大小没有限制,可以处理当前 RNA 分子的所有长度。值得注意的是, LncTar 提供了一种基于自由能的定量标准,可以自动确定两个 RNA 分子是否相互作用。其次, LncTar 运行速度快,因此可用于 lncRNA 的 RNA 靶标的大规模鉴定。最后, LncTar 具有很高的预测精度。

Riblast 是 Fukunaga T 等^[74]开发的一种基于种子区域扩展方法的 RNA-RNA 相互作用预测算法。其主要使用后缀阵列发现种子区域,随后根据 RNA 二级结构能量模型扩展种子区域。Riblast 由数据库构建和 RNA 相互作用搜索两个步骤组成。虽然 Riblast 预测精度与现有方法相似,但其速度比现有方法快 64 倍以上,适用于大型 lncRNA 数据集,全面分析 lncRNA-RNA 相互作用。

RIsearch2 是 Alkan F 等^[75]在 2017 年开发的大规模 RNA-RNA 相互作用预测工具,能够快速定位查询序列和目标序列之间潜在相互作用位点。RIsearch2 使用基于后缀数组的种子扩展框架,采用修正后的 Smith-Waterman-Gotoh 算法来识别 RNA-

RNA 相互作用。另外, RIsearch2 中还提出了一种 siRNA 脱靶识别方法,可预测脱靶转录本,同时还可计算给定 siRNA 的脱靶潜力。

miRWalk 是 Sticht C 等^[76]在 2018 年开发的综合性的 miRNA 靶基因数据库,收录有人类、小鼠、大鼠、狗和牛等多个物种的 miRNA 靶基因信息。miRWalk 收录有 TargetScan, RNAhybrid, TarPmiR 等多个 miRNA 靶标预测工具的结果,收录有包括 5'-UTR, CDS 和 3'-UTR 在内的完整转录序列上的 miRNA 靶位点信息。miRWalk 可以从 miRNA 出发,检索其靶基因,也可以从基因出发,检索可靶向结合的 miRNA。除可以表格形式导出检索结果, miRWalk 还提供 miRNA 和靶基因之间调控网络可视化及基因集富集分析功能。

4.1.3 lncRNA-蛋白互作分析

CatRAPID 是 Agostini F 等^[77]在 2013 年开发的用于大规模计算蛋白质-RNA 相互作用的工具,并由 Armaos A 等^[78]于 2021 年更新至第二版。CatRAPID 可在八种模式生物的全转录组水平或全蛋白质组水平计算蛋白质-RNA 相互作用倾向。用户可以根据已知蛋白质序列预测与其结合的 RNA 序列,根据已知 RNA 序列预测与其结合的蛋白质,或预测一对特定 RNA 与蛋白质之间的作用倾向。预测结果会以图及表格的形式展示,其中表格部分包含蛋白质与 RNA 间的交互倾向、RNA 及蛋白质结合区域等共计十项内容。

LncADeep 是 Yang C 等^[79]开发的基于深度学习的 lncRNA 识别、lncRNA-蛋白质互作预测以及功能注释的生物信息学工具。作者基于 lncRNA-蛋白质相互作用对的序列和结构特征,使用 DNN 构建深度学习模型,用于预测 lncRNA 的蛋白质靶标。此外, LncADeep 还可根据 KEGG, Reactome 数据库中功能注释,从预测的相互作用蛋白推断 lncRNA 的生物学功能。

Capsule-LPI 是 Li Y 等^[80]开发的一种基于胶囊网络的 lncRNA-蛋白质相互作用预测工具。Capsule-LPI 集成了四组多模态特征,包括序列特征、基序信息、理化性质和二级结构特征。其多通道框架不仅可以考虑每个特征的预测趋势,还可以考虑其他交互信息和不同特征之间的关系,从而更好地集成和学习多个特征。与现有的 lncRNA-蛋白质相互作用预测工具相比, Capsule-LPI 集成了更多的特征信息,如物理化学特征和基序信息,并且预测精度也有所提升。

4.1.4 lncRNA-染色质互作分析

ChIRP-seq 是 Chu C 等^[81]开发的可绘制全基因组范围内 RNA 结合位点的高通量测序技术。生物素或链霉亲和素探针的互补寡核苷酸与目标 RNA

杂交后,使用磁性链霉亲和素珠分离并纯化与其作用的染色质片断。洗脱后进行高通量测序,鉴定和定量分析结合区域。亲和探针的设计仅基于 RNA 序列,不需要事先了解 RNA 的结构或功能域,因此方便快捷。ChIRP-seq 可为 lncRNA-染色质互作的功能机制研究提供新的技术支撑。

CHART-seq 是 Simon MD 等^[82]开发的基于杂交的高通量测定内源性非编码 RNA 结合靶标的新技术。该方法使用互补寡核苷酸从可逆交联提取物中

纯化 RNA 及其靶标,可用于绘制内源性 RNA 的全基因组结合位点。作者使用 CHART-seq 分析了 Jpx 结合的染色质位点,发现其倾向于结合启动子、转录起始临近区域或转录终止区等,尤其与转录活跃基因显示出高相关性。鉴于 lncRNA 在染色质结构和基因表达调控中的重要作用,CHART 提供了一种有价值的工具来鉴定其在染色质上的结合位点。

4.2 数据库

lncRNA 与可变剪接相关数据库见表 2。

表 2 lncRNA 与可变剪接相关数据库
Table 2 Long non coding RNA and alternative splicing related database

类别	数据库名称	数据来源	数据库特点	网址
lncRNA-SF	spliceAid-F ^[83]	文本挖掘、数据库筛选	71 种可与 RNA 结合的人类剪接因子相关实验数据;剪接因子的功能域及其蛋白质和化学相互作用;	http://srv00.recas.ba.infn.it/SpliceAidF
	NPInter ^[84]	文本挖掘,CLIP-seq, ChIRP-seq,RISE Database	650 180 个 RNA-蛋白质作用对;2 579 283 个 RNA 结合互作和 138 868 个 RNA 调控数据;	http://bigdata.ibp.ac.cn/npinter5
	starBase ^[86]	CLIP-seq, RNA-seq, miRNA-seq	~117 000 个 RBP-ncRNA 作用对;366 种疾病相关的 3 236 000 个突变;大量 lncRNA 与剪接因子互作数据	http://starbase.sysu.edu.cn
	ChIPBase ^[87]	ChIP-seq, RNA-seq 等	~151 187 000 个转录因子与基因间的调控关系;41 796 000 个 lncRNA-转录因子调控关系;	http://rnasysu.com/chipbase3
lncRNA-TF-SF	lncMap ^[89]	TANRIC, ChIPBase, RNA-seq	20 种癌症类型; lncRNA-TF-基因; 药物-lncRNA;生存分析;TF-基因;	http://bio-bigdata.hrbmu.edu.cn/LncMAP
	Harmonizome ^[90]	125 个数据集、72 个开放在线资源	71 597 788 个基因-生物学功能关联;TF-SF 相互作用;	http://amp.pharm.mssm.edu/Harmonizome
	Cistrome Cancer ^[91]	TCGA, GEO, ENCODE, Roadmap Epigenomics	318 个 TF 和 20 多种癌症;每种癌症类型中的 TF-靶基因互作关系;	http://cistrome.org/CistromeCancer
	GTRD ^[92]	ChIP-exo, ChIP-nexus, ChIP-seq, FAIRE-seq 等	最全面的人类转录调控实验数据资源;不同类型细胞或组织中的 TF-靶基因作用对;	http://gtrd.biouml.org
lncRNA-miRNA-SF	miRTarBase ^[94]	miRBase, HMDD, TransmiR 等	4 475 477 个 miRNA-target 互作对;可检索与生物学通路、癌症或复杂疾病相关的 miRNA;	http://miRTarBase.cuhk.edu.cn
	DIANA-LncBase ^[97]	文本挖掘、CLIP-seq	~240 000 个 miRNA-lncRNA 相互作用对;>9 000 个 miRNA 靶位点的已知短变异信息;	http://www.microrna.gr/LncBase
lncRNA-chromatin	LnChrom ^[98]	文献挖掘、GEO 数据集、TCGA 数据库等	人和小鼠的382 743 个lncRNA-染色质相互作用;lncRNA-染色质相互作用验证实验数据;	http://biocc.hrbmu.edu.cn/LnChrom

注:starBase 横跨多个类别,包括 lncRNA-SF, lncRNA-TF-SF 及 lncRNA-miRNA-SF.

4.2.1 lncRNA-SF 相互作用

SpliceAid-F 是 Giulietti M 等^[83] 在 2013 年开发的剪接因子数据库,并于 2014 年进行了更新。研究人员通过人工筛选文献和数据库,收集整理了 71 种可与 RNA 结合的人类剪接调控蛋白相关实验数据。数据库还收录了实验验证的 RNA-SF 相互作用及其相关信息,包括结合位点序列、位点所在的基因、结合位点坐标、参考文献等。此外,数据库中还记录了剪接因子的功能结构域及其相互作用蛋白。该数据库中收集的 RNA-剪接因子互作可为 lncRNA 调控可变剪接机制研究提供借鉴。

NPInter 是 Wu T 等^[84] 创建的非编码 RNA 和生物大分子相互作用数据库。2022 年 Zheng Y 等^[85] 将其更新至第五版,收录了 200 多万条非编码 RNA 功能相互作用信息。其中,新增添的 RNA 结合蛋白数据包含 277 个 RNA 结合蛋白,并提供细胞定位, RNA 结合区域及生物学功能等信息。依据功能信息,用户可方便地检索到剪接调控相关 RNA 结合蛋白,并下载与其互作的 lncRNA 相关信息,也可从数据库中直接下载 lncRNA-蛋白质互作数据。

starBase v2.0 是 Li JH 等^[86] 基于 CLIP-seq 等高通量数据构建的 RNA 互作数据库,包含 410 万个 miRNA-ncRNA、290 万个 miRNA-mRNA、410 万个 RBP-RNA 和 150 万个 RNA-RNA 相互作用,并且提供数据集可视化,分析和下载功能,极大地扩展了对非编码 RNA 功能及其调控网络的理解。其中,从 lncRNA-RBP 数据集中可以筛选得到 lncRNA-剪接因子相互作用对。

4.2.2 lncRNA-TF-SF 互作数据库

ChIPBase 是 Yang JH 等^[87] 开发的基于 ChIP-Seq 数据的转录因子数据库,用于全面注释和发现 lncRNA 和 miRNA 与转录因子的互作调控关系。目前,Huang J 等^[88] 已将数据库更新到第三版,利用收集的来自 CistromeDB, ReMap2022, ENCODE 的 ChIP-seq 数据,来自 ENCODE, TCGA, GTEx, CCLE, ICGC 的 RNA-seq 数据以及其他来源的 SNV, SNP 和基因注释数据,绘制了 mRNA 和 ncRNA 的转录调控关系图谱。作者整合了在不同样本中基因-转录因子相互作用,构建了以基因为中心的调控网络。用户仅需输入官方基因名称,即可获得与查询基因结合的转录调节因子。

LncMAP 数据库是 Li Y 等^[89] 构建的包含 20 种癌症中 lncRNA 介导的转录扰动资源数据库,主要包括 lncRNA-TF-Gene、药物-lncRNA、生存分析等功能模块。用户可以根据癌症、lncRNA、转录因子或基因检索数据库,获取目标癌症的 lncRNA-TF-Gene 互作

关系,也可下载相关数据自行筛选。该数据库提供了比 ChIPBase 增加了 lncRNA-转录因子作用对的靶基因信息,可提供更详细的 lncRNA-TF-gene 互作信息。

Harmonizome 是 Rouillard AD 等^[90] 于 2016 年创建的数据库,旨在整合基因、蛋白质及其关联信息,以实现更好的知识提取和数据重用。该数据库收集了来自 72 个公开资源的 125 个独特数据集,涵盖了关于哺乳动物基因和蛋白质的六大类信息:疾病和表型关联、基因组图谱、物理相互作用、蛋白质组图谱,结构或功能注释和转录组图谱。在搜索界面输入“Transcription factor targets”即可获得转录因子靶标数据集,通过对靶基因的筛选可获得转录因子与剪接因子的互作关系。

Cistrome Cancer 是 Mei S 等^[91] 为了深入了解调节癌症异常基因表达模式的转录和表观遗传因素开发的数据库。癌症相关数据资源来自 TCGA 数据库,包括转录组、DNA 甲基化、拷贝数变异和临床生存信息。另外,作者也收集整理了 GEO, ENCODE 和 Roadmap 中表观基因组学 ChIP-seq 和染色质可及性数据。用户可在“Transcription Factor Targets in Cancer”功能页面左侧列表中,输入或选择目标转录因子获取转录因子靶标信息。

GTRD 是 Yevshin I 等^[92] 于 2017 年构建的数据库,2021 年 Kolmykov S 等^[93] 对数据库进行了更新。更新后的版本是目前真核生物转录调控数据的最大集成资源,包含与基因转录相关的测序数据:ChIP-seq, ChIP-exo, DNase-seq, MNase-seq, ATAC-seq 和 RNA-seq。用户通过高级浏览与搜索功能,可快速准确获取转录因子靶标信息,包括基因 ID、基因名、转录本 ID 等。

4.2.3 miRNA-target 相互作用

miRTarBase 数据库由 Hsu SD 等^[94] 于 2011 年创建,主要提供 miRNA-靶标相互作用,历经五次修订,最新版本由 Huang HY 等^[95] 于 2022 年更新丰富,集成了数十种数据集,添加了 miRNA 表达谱及其生物学数据。目前,miRTarBase 共收录了 5 173 924 个 miRNA-靶标互作数据对,包含 5 818 个 miRNA 和 33 071 个靶标基因。用户可通过搜索功能,直接检索与指定基因(lncRNA, 转录因子或剪接因子)结合的 miRNA,且其互作信息提供证据来源,可筛选实验验证的互作关系对。

DIANA-LncBase 是 Paraskevopoulou MD 等^[96] 于 2013 年构建的用于探索 miRNA 与 lncRNA 之间相互作用的在线数据库,在近期进行了更新^[97]。该数据库通过整合多个公共数据源中的 RNA-RNA 相互作

用信息,包含约 24 万个具有实验数据支撑的人类和小鼠组织,细胞类型特异性相互作用 miRNA-lncRNA 对。DIANA-LncBase 允许用户采用不同的筛选标准组合进行搜索与浏览,并提供 miRNA-lncRNA 相互作用的详细信息,例如 lncRNA 表达谱、亚细胞定位、转录本基因组位置的图形绘制以及外部数据库信息链接等。

4.2.4 lncRNA-chromatin 相互作用

LnChrom 是 Yu F 等^[98]从文献中人工收集整理实验验证的 lncRNA-染色质相互作用信息构建的数据库,总共包含了 263 种人类和小鼠细胞类型或组织中的 382 743 个 lncRNA-染色质相互作用,涉及 34 345 个靶基因和 2 390 个 lncRNA。该数据库除提供 lncRNA-染色质相互作用图谱外,还提供“Detailed Annotation”,“Genome Brower”,“TF Co-occupancy”及“Cancer Exploration”等功能模块,用于提供详细数据和分析结果。其中,“Cancer Exploration”功能模块提供了来自 TCGA 的 13 种癌症的 4 351 个肿瘤样本和 556 个正常样本中 lncRNA 及其靶基因的表达,有助于阐述 lncRNA-染色质相互作用的功能意义。

5 总 结

剪接失调被认为是驱动癌症发生发展的关键因素,而近来越来越多的证据表明 lncRNA 在可变剪接过程中发挥着重要的调控作用。研究发现,lncRNA 可通过与剪接因子互作,与 pre-mRNA 形成双链结构以及染色体重塑等方式调节可变剪接,其表达失调与癌症中基因异常剪接事件的发生密切相关。目前,癌症相关研究大多聚焦于 lncRNA 与剪接因子间的相互作用,对于其他类型的剪接调控机制报道有限。此外,鉴于 lncRNA 调控机制的复杂性和多样性,其剪接调控模式也有待进一步的挖掘。尽管已有诸多生信工具和数据库可用于剪接调控相关 lncRNA 的识别及作用机制分析,但受限于当前高通量测序技术的固有缺陷及数据质量,预测结果的准确性及可靠性仍有待进一步提升。

参考文献(References)

[1] LEWIS B P, SHIH I H, JONES-RHOADES M W, et al. Prediction of mammalian microRNA targets[J]. Cell, 2003, 115(7):787-798. DOI:10.1016/s0092-8674(03)01018-3.

[2] LIU Yunze, LIU Xin, LIN Changwei, et al. Noncoding RNAs regulate alternative splicing in Cancer[J]. Journal of Experimental & Clinical Cancer Research, 2021, 40(1):

11. DOI:10.1186/s13046-020-01798-2.

[3] CAO Pengbo, JIN Qian, FENG Lan, et al. Emerging roles and potential clinical applications of noncoding RNAs in hepatocellular carcinoma[J]. Seminars in Cancer Biology, 2021, 75:136-152. DOI:10.1016/j.semcancer.2020.09.003.

[4] HERMAN A B, TSITSIPATIS D, GOROSPE M. Integrated lncRNA function upon genomic and epigenomic regulation [J]. Molecular Cell, 2022, 82(12): 2252-2266. DOI:10.1016/j.molcel.2022.05.027.

[5] YUAN Kefei, LAN Jiang, XU Lin, et al. Long noncoding RNA TLNC1 promotes the growth and metastasis of liver cancer via inhibition of p53 signaling[J]. Molecular Cancer, 2022, 21(1): 105. DOI:10.1186/s12943-022-01578-w.

[6] ZHU Yiran, ZHOU Bingluo, HU Xinyang, et al. lncRNA LINC00942 promotes chemoresistance in gastric cancer by suppressing MSI2 degradation to enhance c-Myc mRNA stability[J]. Clinical and Translational Medicine, 2022, 12(1): e703. DOI:10.1002/ctm2.703.

[7] GORDON M A, BABBS B, COCHRANE D R, et al. The long non-coding RNA MALAT1 promotes ovarian cancer progression by regulating RBFOX2-mediated alternative splicing [J]. Molecular Carcinogenesis, 2019, 58(2): 196-205. DOI:10.1002/mc.22919.

[8] LI Jinglei, TONG Hanxing, LI Dongping, et al. The long non-coding RNA DKFZp434J0226 regulates the alternative splicing process through phosphorylation of SF3B6 in PDAC [J]. Molecular Medicine, 2021, 27(1): 95. DOI:10.1186/s10020-021-00347-7.

[9] LI Bosheng, KANG Houyi, XIAO Yufeng, et al. lncRNA GAL promotes colorectal cancer liver metastasis through stabilizing GLUT1 [J]. Oncogene, 2022, 41(13): 1882-1894. DOI:10.1038/s41388-022-02230-z.

[10] BARALLE F E, GIUDICE J. Alternative splicing as a regulator of development and tissue identity[J]. Nature Reviews Molecular Cell Biology, 2017, 18(7): 437-451. DOI:10.1038/nrm.2017.27.

[11] WANG E T, SANDBERG R, LUO Shujun, et al. Alternative isoform regulation in human tissue transcriptomes[J]. Nature, 2008, 456(7221): 470-476. DOI: 10.1038/nature07509.

[12] DU Junxian, ZHU Guiqin, CAI Jialiang, et al. Splicing factors: Insights into their regulatory network in alternative splicing in cancer[J]. Cancer Letters, 2021, 501: 83-104. DOI:10.1016/j.canlet.2020.11.043.

[13] TRIPATHI V, ELLIS J D, SHEN Zhen, et al. The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation [J]. Molecular Cell, 2010, 39(6): 925-938. DOI:10.1016/j.molcel.2010.08.011.

[14] SI Zizhen, YU Lei, JING Haoyu, et al. Oncogenic lncRNA ZNF561-AS1 is essential for colorectal cancer proliferation

- and survival through regulation of miR-26a-3p/miR-128-5p-SRSF6 axis[J]. *Journal of Experimental & Clinical Cancer Research*, 2021, 40(1): 78. DOI:10.1186/s13046-021-01882-1.
- [15] DAVID C J, CHEN Mo, ASSANAH M, et al. HnRNP proteins controlled by c-Myc deregulate pyruvate kinase mRNA splicing in cancer[J]. *Nature*, 2010, 463(7279): 364-368. DOI:10.1038/nature08697.
- [16] BELTRAN M, PUIG I, PENA C, et al. A natural antisense transcript regulates Zeb2/Sip1 gene expression during Snail1-induced epithelial-mesenchymal transition [J]. *Genes & Development*, 2008, 22(6): 756-769. DOI:10.1101/gad.455708.
- [17] VILLAMIZAR O, CHAMBERS C B, RIBERDY J M, et al. Long noncoding RNA Saf and splicing factor 45 increase soluble Fas and resistance to apoptosis [J]. *Oncotarget*, 2016, 7(12): 13810-13826. DOI:10.18632/oncotarget.7329.
- [18] TENG Liu, FENG Yuchen, GUO Sutang, et al. The pan-cancer lncRNA PLANE regulates an alternative splicing program to promote cancer pathogenesis[J]. *Nature Communications*, 2021, 12(1): 3734. DOI:10.1038/s41467-021-24099-4.
- [19] ZHANG Guanxiong, LAN Yujia, XIE Aimi, et al. Comprehensive analysis of long noncoding RNA (lncRNA)-chromatin interactions reveals lncRNA functions dependent on binding diverse regulatory elements[J]. *Journal of Biological Chemistry*, 2019, 294(43): 15613-15622. DOI:10.1074/jbc.RA119.008732.
- [20] TANG Yanyan, WANG Jinpeng, LIAN Yu, et al. Linking long non-coding RNAs and SWI/SNF complexes to chromatin remodeling in cancer[J]. *Molecular Cancer*, 2017, 16(1): 42. DOI:10.1186/s12943-017-0612-0.
- [21] HUANG Mingyan, WANG Huamin, HU Xiang, et al. lncRNA MALAT1 binds chromatin remodeling subunit BRG1 to epigenetically promote inflammation-related hepatocellular carcinoma progression [J]. *Oncoimmunology*, 2019, 8(1): e1518628. DOI:10.1080/2162402X.2018.1518628.
- [22] PANDEY R R, MONDAL T, MOHAMMAD F, et al. Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation[J]. *Molecular Cell*, 2008, 32(2): 232-246. DOI:10.1016/j.molcel.2008.08.022.
- [23] LUCO R F, ALLO M, SCHOR I E, et al. Epigenetics in alternative pre-mRNA splicing[J]. *Cell*, 2011, 144(1): 16-26. DOI:10.1016/j.cell.2010.11.056.
- [24] SIMS R R, MILLHOUSE S, CHEN C F, et al. Recognition of trimethylated histone H3 lysine 4 facilitates the recruitment of transcription postinitiation factors and pre-mRNA splicing[J]. *Molecular Cell*, 2007, 28(4): 665-676. DOI:10.1016/j.molcel.2007.11.010.
- [25] CHO K R, SHIH I. Ovarian cancer[J]. *Annual Review of Pathology-Mechanisms of Disease*, 2009, 4: 287-313. DOI:10.1146/annurev.pathol.4.110807.092246.
- [26] SHOU Jian, ZHANG Chuanling, ZHENG Xiaoyu, et al. LncRNA HCP5 Facilitates the Progression of Ovarian Cancer by Interacting with the PTBP1 Protein [J]. *Biochemical Genetics*, 2024, 62(4): 3136-3154. DOI:10.1007/s10528-023-10558-8.
- [27] WU Jianfa, NI Xiaohong, YU Zhen, et al. CRNDE inducing cisplatin resistance through SRSF1/TIA1 signaling pathway in ovarian cancer [J]. *Pathology Research and Practice*, 2022, 235: 153957. DOI:10.1016/j.prp.2022.153957.
- [28] WEITZ J, KOCH M, DEBUS J, et al. Colorectal cancer [J]. *Lancet*, 2005, 365(9454): 153-165. DOI:10.1016/S0140-6736(05)17706-X.
- [29] LAN Zhixian, YAO Xiang, SUN Kangyue, et al. The interaction between lncrna snhg6 and hnnpa1 contributes to the growth of colorectal cancer by enhancing aerobic glycolysis through the regulation of alternative splicing of PKM [J]. *Frontiers in Oncology*, 2020, 10: 363. DOI:10.3389/fonc.2020.00363.
- [30] HU Zhiyan, WANG Xiaoyan, GUO Wenbin, et al. Long non-coding RNA MALAT1 increases AKAP-9 expression by promoting SRPK1-catalyzed SRSF1 phosphorylation in colorectal cancer cells[J]. *Oncotarget*, 2016, 7(10): 11733-11743. DOI:10.18632/oncotarget.7367.
- [31] KONG Jianlu, SUN Wenjie, LI Chen, et al. Long non-coding RNA LINC01133 inhibits epithelial-mesenchymal transition and metastasis in colorectal cancer by interacting with SRSF6 [J]. *Cancer Letters*, 2016, 380(2): 476-484. DOI:10.1016/j.canlet.2016.07.015.
- [32] SMYTH E C, NILSSON M, GRABSCH H I, et al. Gastric cancer[J]. *Lancet*, 2020, 396(10251): 635-648. DOI:10.1016/S0140-6736(20)31288-5.
- [33] ZHANG Feifei, WANG Hui, YU Jiang, et al. LncRNA CRNDE attenuates chemoresistance in gastric cancer via SRSF6-regulated alternative splicing of PICALM [J]. *Molecular Cancer*, 2021, 20(1): 6. DOI:10.1186/s12943-020-01299-y.
- [34] HE Qiuming, YANG Chaogang, XIANG Zhenxian, et al. LINC00924-induced fatty acid metabolic reprogramming facilitates gastric cancer peritoneal metastasis via hnRNPC-regulated alternative splicing of Mnk2 [J]. *Cell Death & Disease*, 2022, 13(11): 987. DOI:10.1038/s41419-022-05436-x.
- [35] ZHU Wenzhong, TAN Lulu, MA Tiantian, et al. Long non-coding RNA SNHG8 promotes chemoresistance in gastric cancer via binding with hnRNPA1 and stabilizing TROY expression[J]. *Digestive and Liver Disease*, 2022, 54(11): 1573-1582. DOI:10.1016/j.dld.2022.02.011.

- [36] FORNER A, REIG M, BRUIX J. Hepatocellular carcinoma [J]. *Lancet*, 2018, 391 (10127): 1301–1314. DOI: 10.1016/S0140-6736(18)30010-2.
- [37] GONZALEZ I, MUNITA R, AGIRRE E, et al. A lncRNA regulates alternative splicing via establishment of a splicing-specific chromatin signature [J]. *Nature Structural & Molecular Biology*, 2015, 22 (5): 370–376. DOI: 10.1038/nsmb.3005.
- [38] AMANN T, BATAILLE F, SPRUSS T, et al. Reduced expression of fibroblast growth factor receptor 2IIIb in hepatocellular carcinoma induces a more aggressive growth [J]. *American Journal of Pathology*, 2010, 176 (3): 1433–1442. DOI: 10.2353/ajpath.2010.090356.
- [39] LIN Y H, WU M H, LIU Y C, et al. LINC01348 suppresses hepatocellular carcinoma metastasis through inhibition of SF3B3-mediated EZH2 pre-mRNA splicing [J]. *Oncogene*, 2021, 40 (28): 4675–4685. DOI: 10.1038/s41388-021-01905-3.
- [40] MALAKAR P, SHILO A, MOGILEVSKY A, et al. Long Noncoding RNA MALAT1 Promotes Hepatocellular Carcinoma Development by SRSF1 Upregulation and mTOR Activation [J]. *Cancer Research*, 2017, 77 (5): 1155–1167. DOI: 10.1158/0008-5472.CAN-16-1508.
- [41] KLEIN A P. Pancreatic cancer epidemiology: Understanding the role of lifestyle and inherited risk factors [J]. *Nature Reviews Gastroenterology & Hepatology*, 2021, 18 (7): 493–502. DOI: 10.1038/s41575-021-00457-x.
- [42] MENG L D, SHI G D, GE W L, et al. Linc01232 promotes the metastasis of pancreatic cancer by suppressing the ubiquitin-mediated degradation of HNRNP2B1 and activating the A-Raf-induced MAPK/ERK signaling pathway [J]. *Cancer Letters*, 2020, 494: 107–120. DOI: 10.1016/j.canlet.2020.08.001.
- [43] ZHANG Yong, FANG Yuan, MA Lijie, et al. LINC00857 regulated by ZNF460 enhances the expression of CLDN12 by sponging miR-150-5p and recruiting SRSF1 for alternative splicing to promote epithelial-mesenchymal transformation of pancreatic adenocarcinoma cells [J]. *RNA Biology*, 2022, 19 (1): 548–559. DOI: 10.1080/15476286.2021.1992995.
- [44] XU Shouping, WANG Peiyuan, ZHANG Jian, et al. Ai-lncRNA EGOT enhancing autophagy sensitizes paclitaxel cytotoxicity via upregulation of ITPR1 expression by RNA-RNA and RNA-protein interactions in human cancer [J]. *Molecular Cancer*, 2019, 18 (1): 89. DOI: 10.1186/s12943-019-1017-z.
- [45] SINGH R, GUPTA S C, PENG W X, et al. Regulation of alternative splicing of Bcl-x by BC200 contributes to breast cancer pathogenesis [J]. *Cell Death & Disease*, 2016, 7 (6): e2262. DOI: 10.1038/cddis.2016.168.
- [46] DUAN Yuqing, JIA Yunlong, WANG Jiali, et al. Long noncoding RNA DGCR5 involves in tumorigenesis of esophageal squamous cell carcinoma via SRSF1-mediated alternative splicing of Mcl-1 [J]. *Cell Death & Disease*, 2021, 12 (6): 587. DOI: 10.1038/s41419-021-03858-7.
- [47] ZHANG Ziqiang, ZHOU Nanjiang, HUANG Jianguo, et al. Regulation of androgen receptor splice variant AR3 by PC-GEM1 [J]. *Oncotarget*, 2016, 7 (13): 15481–15491. DOI: 10.18632/oncotarget.7139.
- [48] BERNARD A, BOIDOT R, VEGRAN F. Alternative splicing in cancer and immune cells [J]. *Cancers*, 2022, 14 (7): 1726. DOI: 10.3390/cancers14071726.
- [49] AUBREY B J, STRASSER A, KELLY G L. Tumor-suppressor functions of the TP53 pathway [J]. *Cold Spring Harbor Perspectives in Medicine*, 2016, 6: a026062. DOI: 10.1101/cshperspect.a026062.
- [50] YAO Yiran, GU Xiang, XU Xiaofang, et al. Novel insights into RB1 mutation [J]. *Cancer Letters*, 2022, 547: 215870. DOI: 10.1016/j.canlet.2022.215870.
- [51] CHEN C Y, CHEN Jingyu, HE Lina, et al. PTEN: Tumor suppressor and metabolic regulator [J]. *Frontiers in Endocrinology*, 2018, 9: 338. DOI: 10.3389/fendo.2018.00338.
- [52] CALABRESE C, DAVIDSON N R, DEMIRCIOGLU D, et al. Genomic basis for RNA alterations in cancer [J]. *Nature*, 2020, 578 (7793): 129–136. DOI: 10.1038/s41586-020-1970-0.
- [53] BAO Guangyao, LI Tian, GUAN Xiaojiao, et al. Development of a prognostic alternative splicing signature associated with tumor microenvironment immune profiles in lung adenocarcinoma [J]. *Frontiers in Oncology*, 2022, 12: 880478. DOI: 10.3389/fonc.2022.880478.
- [54] GUO Shuang, WANG Xinyue, ZHOU Hanxiao, et al. Identification and characterization of immunogene-related alternative splicing patterns and tumor microenvironment infiltration patterns in breast cancer [J]. *Cancers*, 2022, 14 (3): 595. DOI: 10.3390/cancers14030595.
- [55] LIU Weizhen, LI Dongfang, LU Ting, et al. Comprehensive analysis of RNA-binding protein SRSF2-dependent alternative splicing signature in malignant proliferation of colorectal carcinoma [J]. *Journal of Biological Chemistry*, 2023, 299 (2): 102876. DOI: 10.1016/j.jbc.2023.102876.
- [56] FUENTES-FAYOS A C, PEREZ-GOMEZ J M, G-GARCIA M E, et al. SF3B1 inhibition disrupts malignancy and prolongs survival in glioblastoma patients through BCL2L1 splicing and mTOR/ss-catenin pathways imbalances [J]. *Journal of Experimental & Clinical Cancer Research*, 2022, 41 (1): 39. DOI: 10.1186/s13046-022-02241-4.
- [57] GUO Jing, LI Chuzhong, FANG Qiuyue, et al. The SF3B1 (R625H) mutation promotes prolactinoma tumor progression through aberrant splicing of DLG1 [J]. *Journal of Experimental & Clinical Cancer Research*, 2022, 41 (1): 26. DOI: 10.1186/s13046-022-02245-0.

- [58] WANG Xianghu, SUN Mingjun, GAO Zhikui, et al. N-nitrosamines-mediated downregulation of LncRNA-UCA1 induces carcinogenesis of esophageal squamous by regulating the alternative splicing of FGFR2[J]. *Science of the Total Environment*, 2023, 855: 158918. DOI:10.1016/j.scitotenv.2022.158918.
- [59] HUAN Lin, GUO Tianan, WU Yangjun, et al. Hypoxia induced LUCAT1/PTBP1 axis modulates cancer cell viability and chemotherapy response[J]. *Molecular Cancer*, 2020, 19(1): 11. DOI:10.1186/s12943-019-1122-z.
- [60] FU Rong, YANG Peng, AMIN S, et al. A novel miR-206/hnRPA1/PKM2 axis reshapes the Warburg effect to suppress colon cancer growth[J]. *Biochemical and Biophysical Research Communications*, 2020, 531(4): 465–471. DOI:10.1016/j.bbrc.2020.08.019.
- [61] WU Haili, CUI Mimi, LI Chenglu, et al. Kaempferol reverses aerobic glycolysis via miR-339-5p-Mediated PKM alternative splicing in colon cancer cells[J]. *Journal of Agricultural and Food Chemistry*, 2021, 69(10): 3060–3068. DOI:10.1021/acs.jafc.0c07640.
- [62] LIU Zhiyong, LI Wei, PANG Yanan, et al. SF3B4 is regulated by microRNA-133b and promotes cell proliferation and metastasis in hepatocellular carcinoma[J]. *Ebiomedicine*, 2018, 38: 57–68. DOI:10.1016/j.ebiom.2018.10.067.
- [63] SHEN Shihao, PARK J W, LU Zhixiang, et al. rMATS: Robust and flexible detection of differential alternative splicing from replicate RNA-Seq data[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2014, 111(51): E5593–E5601. DOI: 10.1073/pnas.1419161111.
- [64] WU Jie, AKERMAN M, SUN S, et al. SpliceTrap: A method to quantify alternative splicing under single cellular conditions[J]. *Bioinformatics*, 2011, 27(21): 3010–3016. DOI:10.1093/bioinformatics/btr508.
- [65] SHI Yang, JIANG Hui. rSeqDiff: Detecting differential isoform expression from RNA-Seq data using hierarchical likelihood ratio test[J]. *Plos One*, 2013, 8(11): e79448. DOI:10.1371/journal.pone.0079448.
- [66] RYAN M C, CLELAND J, KIM R, et al. SpliceSeq: A resource for analysis and visualization of RNA-Seq data on alternative splicing and its functional impacts[J]. *Bioinformatics*, 2012, 28(18): 2385–2387. DOI:10.1093/bioinformatics/bts452.
- [67] RYAN M, WONG W C, BROWN R, et al. TCGASpliceSeq a compendium of alternative mRNA splicing in cancer [J]. *Nucleic Acids Research*, 2016, 44(D1): D1018–D1022. DOI:10.1093/nar/gkv1288.
- [68] LI Y I, KNOWLES D A, HUMPHREY J, et al. Annotation-free quantification of RNA splicing using LeafCutter [J]. *Nature Genetics*, 2018, 50(1): 151–158. DOI:10.1038/s41588-017-0004-9.
- [69] TRINCADO J L, ENTIZNE J C, HYSENAJ G, et al. SUPPA2: Fast, accurate, and uncertainty-aware differential splicing analysis across multiple conditions[J]. *Genome Biology*, 2018, 19(1): 40. DOI:10.1186/s13059-018-1417-1.
- [70] TANG A D, SOULETTE C M, VAN BAREN M J, et al. Full-length transcript characterization of SF3B1 mutation in chronic lymphocytic leukemia reveals downregulation of retained introns[J]. *Nature Communications*, 2020, 11(1): 1438. DOI:10.1038/s41467-020-15171-6.
- [71] LEWIS B P, SHIH I H, JONES-RHOADES M W, et al. Prediction of mammalian microRNA targets [J]. *Cell*, 2003, 115(7): 787–798. DOI:10.1016/s0092-8674(03)01018-3.
- [72] KRUGER J, REHMSMEIER M. RNAhybrid: microRNA target prediction easy, fast and flexible[J]. *Nucleic Acids Research*, 2006, 34(Web server issue): W451–W454. DOI:10.1093/nar/gkl243.
- [73] LI Jianwei, MA Wei, ZENG Pan, et al. LncTar: a tool for predicting the RNA targets of long noncoding RNAs [J]. *Briefings in Bioinformatics*, 2015, 16(5): 806–812. DOI: 10.1093/bib/bbu048.
- [74] FUKUNAGA T, HAMADA M. Ribblast: An ultrafast RNA-RNA interaction prediction system based on a seed-and-extension approach[J]. *Bioinformatics*, 2017, 33(17): 2666–2674. DOI:10.1093/bioinformatics/btx287.
- [75] ALKAN F, WENZEL A, PALASCA O, et al. RIssearch2: Suffix array-based large-scale prediction of RNA-RNA interactions and siRNA off-targets[J]. *Nucleic Acids Research*, 2017, 45(8): e60. DOI:10.1093/nar/gkw1325.
- [76] STICHT C, DE LA TORRE C, PARVEEN A, et al. miR-Walk: An online resource for prediction of microRNA binding sites[J]. *Plos One*, 2018, 13(10): e206239. DOI:10.1371/journal.pone.0206239.
- [77] AGOSTINI F, ZANZONI A, KLUS P, et al. catRAPID omics: A web server for large-scale prediction of protein-RNA interactions [J]. *Bioinformatics*, 2013, 29(22): 2928–2930. DOI:10.1093/bioinformatics/btt495.
- [78] ARMAOS A, COLANTONI A, PROIETTI G, et al. catRAPID omics v2.0: Going deeper and wider in the prediction of protein-RNA interactions [J]. *Nucleic Acids Research*, 2021, 49(W1): W72–W79. DOI:10.1093/nar/gkab393.
- [79] YANG Cheng, YANG Longshu, ZHOU Man, et al. LncADeep: An ab initio lncRNA identification and functional annotation tool based on deep learning[J]. *Bioinformatics*, 2018, 34(22): 3825–3834. DOI:10.1093/bioinformatics/bty428.
- [80] LI Ying, SUN Hang, FENG Shiyao, et al. Capsule-LPI: A LncRNA-protein interaction predicting tool based on a capsule network[J]. *Bmc Bioinformatics*, 2021, 22(1): 246.

- DOI:10.1186/s12859-021-04171-y.
- [81] CHU Ci, QU Kun, ZHONG F L, et al. Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions[J]. *Molecular Cell*, 2011, 44(4): 667-678. DOI:10.1016/j.molcel.2011.08.027.
- [82] SIMON M D, WANG C I, KHARCHENKO P V, et al. The genomic binding sites of a noncoding RNA[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2011, 108(51): 20497-20502. DOI:10.1073/pnas.1113536108.
- [83] GIULIETTI M, PIVA F, D'ANTONIO M, et al. SpliceAid-F: A database of human splicing factors and their RNA-binding sites[J]. *Nucleic Acids Research*, 2013, 41(Database issue): D125-D131. DOI:10.1093/nar/gks997.
- [84] WU Tao, WANG Jie, LIU Changning, et al. NPInter: The noncoding RNAs and protein related biomacromolecules interaction database[J]. *Nucleic Acids Research*, 2006, 34(Database issue): D150-D152. DOI:10.1093/nar/gkj025.
- [85] ZHENG Yu, LUO Huaxia, TENG Xueyi, et al. NPInter v5.0: ncRNA interaction database in a new era[J]. *Nucleic Acids Research*, 2023, 51(D1): D232-D239. DOI:10.1093/nar/gkac1002.
- [86] LI Junhao, LIU Shun, ZHOU Hui, et al. starBase v2.0: Decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data[J]. *Nucleic Acids Research*, 2014, 42(Database issue): D92-D97. DOI:10.1093/nar/gkt1248.
- [87] YANG Jianhua, LI Junhao, JIANG Shan, et al. ChIPBase: A database for decoding the transcriptional regulation of long non-coding RNA and microRNA genes from ChIP-Seq data[J]. *Nucleic Acids Research*, 2013, 41(Database issue): D177-D187. DOI:10.1093/nar/gks1060.
- [88] HUANG Junhong, ZHENG Wujian, ZHANG Ping, et al. ChIPBase v3.0: The encyclopedia of transcriptional regulations of non-coding RNAs and protein-coding genes[J]. *Nucleic Acids Research*, 2023, 51(D1): D46-D56. DOI:10.1093/nar/gkac1067.
- [89] LI Yongsheng, LI Lili, WANG Zishan, et al. LncMAP: Pan-cancer atlas of long noncoding RNA-mediated transcriptional network perturbations[J]. *Nucleic Acids Research*, 2018, 46(3): 1113-1123. DOI:10.1093/nar/gkx1311.
- [90] ROUILLARD A D, GUNDERSEN G W, FERNANDEZ N F, et al. The harmonizome: A collection of processed datasets gathered to serve and mine knowledge about genes and proteins[J]. *Database-the Journal of Biological Databases and Curation*, 2016, 2016: baw100. DOI:10.1093/database/baw100.
- [91] MEI Shenglin, MEYER C A, ZHENG Rongbin, et al. Cis-trome cancer: A web resource for integrative gene regulation modeling in cancer[J]. *Cancer Research*, 2017, 77(21): e19-e22. DOI:10.1158/0008-5472.CAN-17-0327.
- [92] YEVSHIN I, SHARIPOV R, VALEEV T, et al. GTRD: A database of transcription factor binding sites identified by ChIP-seq experiments[J]. *Nucleic Acids Research*, 2017, 45(D1): D61-D67. DOI:10.1093/nar/gkw951.
- [93] KOLMYKOV S, YEVSHIN I, KULYASHOV M, et al. GTRD: An integrated view of transcription regulation[J]. *Nucleic Acids Research*, 2021, 49(D1): D104-D111. DOI:10.1093/nar/gkaa1057.
- [94] HSU S D, LIN Fengmao, WU Weiyun, et al. miRTarBase: A database curates experimentally validated microRNA-target interactions[J]. *Nucleic Acids Research*, 2011, 39(Database issue): D163-D169. DOI:10.1093/nar/gkq1107.
- [95] HUANG H Y, LIN Y C, LI Jing, et al. miRTarBase 2020: Updates to the experimentally validated microRNA-target interaction database[J]. *Nucleic Acids Research*, 2020, 48(D1): D148-D154. DOI:10.1093/nar/gkz896.
- [96] PARASKEVOPOULOU M D, GEORGAKILAS G, KOSTOULAS N, et al. DIANA-LncBase: Experimentally verified and computationally predicted microRNA targets on long non-coding RNAs[J]. *Nucleic Acids Research*, 2013, 41(Database issue): D239-D245. DOI:10.1093/nar/gks1246.
- [97] KARAGKOUNI D, PARASKEVOPOULOU M D, TASTSO-GLOU S, et al. DIANA-LncBase v3: Indexing experimentally supported miRNA targets on non-coding transcripts[J]. *Nucleic Acids Research*, 2020, 48(D1): D101-D110. DOI:10.1093/nar/gkz1036.
- [98] YU Fulong, ZHANG Guanxiong, SHI Aiai, et al. LnChrom: A resource of experimentally validated lncRNA-chromatin interactions in human and mouse[J]. *Database-the Journal of Biological Databases and Curation*, 2018, 2018: bay039. DOI:10.1093/database/bay039.