

综述

长链非编码 RNA 及其与肿瘤关系*

娄亮亮¹ 朱运峰^{1,2**}

(1 北京交通大学生命科学与生物工程研究院 北京 100044 2 解放军总医院肿瘤中心实验室 北京 100853)

摘要 人类基因组 DNA 信息分析显示,基因组中编码序列不超过 2%,其余为非编码序列,其中能够稳定转录的部分称之为非编码 RNA(noncoding RNA, ncRNA)基因。依据 ncRNA 长度将其分为长非编码 RNA(lncRNA, 碱基数 >200nt)和短非编码 RNA(sncRNA)。在人体内, lncRNAs 分布广泛,数目众多,并且大多数是由 RNA 聚合酶 II 转录的,且有 5'端帽子结构以及 3'端的多聚腺苷酸。就 lncRNA 的特征及其在肿瘤细胞中的功能进行阐述。

关键词 非编码 RNA 长非编码 RNA 肿瘤细胞 基因组

中图分类号 Q522

1 lncRNA 简介

对人类全基因组序列分析表明,在基因组约有 93% 的 DNA 能转录为 RNA,而只有约 2% 的 RNA 翻译为蛋白质,其余的 RNA 皆为非编码部分。研究发现生物的复杂性与非编码序列密切相关。根据非编码 RNA(ncRNA) 的长度,这些 ncRNA 被分为短 ncRNA(sncRNA)和长 ncRNA(lncRNA)。LncRNAs 能在染色质重构和组蛋白修饰上起到调控作用,并且影响肿瘤等疾病的发生。除了 lncRNAs 自身发生异常外,它们的结合蛋白异常也被认为与某些疾病发生相关,这类疾病可能是由 RNP(核酸蛋白复合物)缺陷导致。一些神经退行性疾病如脊髓小脑共济失调症(SCA),肌萎缩侧索硬化症(ALS),X 染色体易损综合症以及肿瘤等已被证明是受 RNA 结合蛋白调控^[1],错误折叠和突变蛋白的积累是这类疾病的共同特征。

1.1 lncRNA 的产生

由于 lncRNA 的种类极其多样,目前对 lncRNA 起源的具体过程仍不十分清楚,总体可归纳为以下几种情况^[2]:(1)编码蛋白基因发生框架破坏,然后与前体

编码序列一起转变为有功能的非编码 RNA。(2)通过染色体重排,两个不转录并且先前彼此分离的序列合并形成一个多外显子 ncRNA。(3)非编码基因通过逆转录子的整合而激活。(4)ncRNA 内部发生两次串联重复事件产生相邻重复序列,然后产生 lncRNA。(5)通过转座子的插入产生出有功能的 lncRNA。

1.2 lncRNA 与开放阅读框架(ORF)

LncRNA 通常定义为由超过 200 核苷酸碱基组成的且不翻译成蛋白质的 RNA 分子。但是编码 RNA 和 ncRNA 是很难区分开的,蛋白编码片段一般是有长度超过 100 氨基酸的 ORFs 的存在的,但有报道显示许多 lncRNAs 也含有这样的 ORFs。而且小于 100 长度氨基酸的 ORFs 也可能被翻译,例如果蝇中存在只有 11 氨基酸长度的有功能多肽。编码和非编码亚型的广泛重叠会混淆我们对编码和非编码转录物的划分。第一个被发现的 lncRNAs 即 SRA,能编码一种有抑制 ncRNA 功能的蛋白。另外许多 mRNAs 也能在一个 RNA 水平上发挥作用。研究发现 p53 与 Mdm2 基因之间存在相互制约关系,如 p53 mRNA 能通过结合 Mdm2 蛋白来作为一个 RNA 调节器,引起 p53 的表达。另外许多 mRNAs 的 3'UTR 能脱离其上游的 mRNA 序列独立表达,而无需编码蛋白的帮助^[3]。

为了消除由于 ORFs 的存在而引起的歧义,最近对

lncRNAs 的定义指出,lncRNA 是对初级和拼接转录物 (spliced transcripts) 发挥作用而且不属于已知小 RNA (如 miRNAs, piwi-相互作用 RNAs 和小核仁 RNAs) 以及功能性 RNAs (如 tRNAs, 小细胞核 RNAs 等) 类别的 RNA^[4]。

1.3 lncRNA 的分类和定位

相对于编码序列的位置可将 lncRNAs 划分为五类:同义和反义 lncRNA,即根据与另一转录物是否和同一链上一个或多个外显子重叠来划分;双向 lncRNA,这类 lncRNAs 的表达起始位点与其互补链上相邻编码转录物的表达起始位点十分靠近;内含子 lncRNA,来源于编码基因剪接转录物的内含子(long intronic RNAs 又分成两类:全部内含子和部分内含子 RNAs);基因间 lncRNA,位于两个编码基因的间隔中。

相对于 mRNAs 的细胞质定位,lncRNAs 的亚细胞定位十分多样化。最近发现的一些 lncRNAs (包括不含 poly-A 的 ncRNAs) 参与细胞成分的加工^[5],主要见于细胞核的多种非被膜小体,包括卡哈尔体 (Cajal bodies), 早幼粒细胞白血病体 (PML bodies), paraspeckles 以及 splicing speckles^[6]。Paraspeckles 成簇出现于靠近细胞核斑点的核浆区,它是前体 mRNA 剪接因子储存,组装和修饰的地方。研究发现存留于 paraspeckles 里的两种非编码 RNA MEN β 和 MEN ϵ (nuclear-enriched autosomal transcript 1 NEAT 1),对 paraspeckles 的形成结构和形态维持具有关键的作用。除了二者之外, RNA 荧光原位杂交法显示许多 lncRNAs 如 *Evf1*^[7], *Evf2*^[8], *Kcnq1ot1*^[9] 定位于细胞核的不同区域。总之,lncRNAs 的亚细胞定位可位于细胞质以及细胞核中^[10]。目前对 lncRNAs 亚细胞定位的机制仍不清晰,推测可能是利用与 mRNAs 相同的转运途径^[11],包括在细胞骨架网络中通过识别 RNA 结合蛋白发生相互作用而实现的^[12]。

2 lncRNA 的功能

lncRNAs 起初被认为是基因组转录的“噪音”,是 RNA 聚合酶 II 转录的副产物。但是随着研究的深入,我们发现 lncRNAs 并非只是不具备生物学功能的无用片段。lncRNAs 的功能涉及到许多生物进程的调控及发生,例如基因的顺式和反式调控,指导染色质重构蛋白复合物的形成,X 染色体失活 (Xi),基因组印记,核分区,核质交换, RNA 剪接和翻译调控^[13]。

2.1 调控基因表达

研究表明 lncRNAs 能在转录前、转录和转录后水平上通过多重机制调控基因表达^[14],包括调节染色质的组蛋白修饰,控制 DNA 中转录复合物的装配,在受到外界刺激时能有效地进行调节。

许多 lncRNAs 还能与染色质重构蛋白复合物 (PRC) 结合,调节染色质组蛋白修饰来控制基因表达水平。例如 RepA (Xist) 能与 polycomb repressive complex 2 (PRC2) 相互作用并且影响它在非激活 X 染色体上的定位和活性^[15]。研究发现人体中大约占五分之一的 lncRNAs 能单独或与其他 PRC 结合后与 PRC2 组分联系起来。HOTAIR 能与 PRC2 以及 LSD1/CoREST H3K4 去甲基酶复合物(一种活化标志的清除器)相互作用,通过一系列突变体研究发现,PRC2 的结合域位于 HOTAIR 的 5'端,而 LSD1 复合物的结合域相当于 3'端,这说明它能作为二者的桥梁,通过多重机制调控基因表达。

另外有一些 lncRNAs 结合特殊蛋白分子,然后该复合体在 lncRNA 的引导下结合到特定位点,这类 lncRNAs 可采取顺式调控以及反式调控方式实现对靶基因表达的调控。其中有些 lncRNAs 能作为辅因子或配体来激活转录,另一些会通过招募转录抑制物来达到特定基因组区域的表达沉默。转录激活作用是指: DLX5 和 DLX6 的极端保守的增强子区域是被一种叫做 Evf2 的 lncRNA 所调节,通过与 Dlx2 结合将 Dlx2 转变为一个转录增强子^[8]。转录抑制作用是指:在 DNA 损伤诱导的反应中,低拷贝 CCND1 上游的 lncRNA 通过招募 RNA 结合蛋白 TLS 形成复合体,调节其与 p300 和染色质结合蛋白的结合进而改变组蛋白乙酰转移酶活性抑制 CCND1 mRNA 的表达。lncRNAs 能与转录因子形成 RNA-蛋白质复合物,为转录因子提供类似“脚手架”的作用。

一些反义 lncRNA 片段 (如: p15、p21 的反义转录片段) 能通过直接结合或招募染色质重构复合物,引起抑癌基因的表现遗传沉默^[16]。在人和小鼠体内发现的某些 lncRNAs 如 Tug1, linc-P21, PANDA, Evf-2 等能通过引导他们的蛋白伴侣到特定 DNA 位置来调控基因表达^[17-19],还有一种叫 NRON 的 lncRNA 能阻止转录因子 NEAT 的核聚集。研究发现,对 ncRNAs 进行退火处理后,发现它们能引导蛋白效应复合物结合到正向 mRNA 转录物,作用方式类似于 siRNAs 介导的 RNA 诱导的沉默复合物 (RISC) 机制。这种内源性 siRNA 来源

于互补转录片段甚至是 lncRNAs 的内部发卡状结构,这说明有些 lncRNAs 能介导 RNA 沉默机制^[20]。但它们介导调控的具体分子机制仍不清楚,推测可能与 lncRNAs 的内在序列和二级结构有关。

2.2 介导 X 染色体失活

迄今为止关于 lncRNAs 研究最透彻的便是 X 染色体特异性失活转录物(Xist),Xist 首先发现于 1991 年,它能引发并导致雌性体细胞内 X 染色体失活^[21]。在 Xist 上有一段称为 RepA 的重复区域,它的最初转录是通过 X 染色体和 Xist 的反义物 Tsix 共同作用^[15]。同时发现,另外一种叫 Tsix 的 lncRNA 能防止 RepA 与 X 染色体结合。后期细胞分化中,RepA 能与染色质重构蛋白复合物 PRC2 结合,从位于 X 染色体失活中心部位的一条 X 染色体上启动 Xi^[22]。在顺式调控中,Tsix 通过阻止 Xist 转录来维持一个活性染色质态。但女性中有多达 20% 的基因能逃离 Xi^[23],其作用机制尚不清楚。除了 Xist 和 Tsix,还有其他一些通过 X 失活中心转录的 lncRNAs 能对 Xi 起作用,如 E3 泛素活化酶 Rnf12 是一种 Xist 的催化剂,它可能直接或间接介导 Xist 活化作用。最近报道,两种 Xic-编码的 lncRNAs, Jpx 和 Ftx,能通过 XCI 调控转录和积聚,也被认为是 Xist 的催化剂。而一旦 Xi 开始,不活跃的 X 染色体就会浓缩成异染色质结构而呈现出浓缩的圆体(巴氏小体),它通常位于核心周围^[24]。

2.3 维持全能性

胚胎干细胞主要通过外源性信号分子和某些重要的内源性转录因子的共同作用来达到其维持多能性的目的。与全能性相关的 lncRNAs 最先发现于小鼠胚胎干细胞(ESC)中,它们是 ESC 调节回路中的组成成员^[25]。某些 lncRNAs 与多能性标记分子的表达有关,超过 100 个 lncRNAs 的启动子区被核心干细胞转录因子所调控,例如 OCT4 和 NANOG。需要指出的是,这些受核心转录因子调控的 lncRNAs 自身也可调控多能性干细胞的发展进程。这类受干细胞因子调控的 lncRNAs 的过表达或下调将影响 ESC 全能性的维持并且调控 ESC 细胞系的特殊分化^[26]。

报道显示有 26 种 lncRNAs 是对维持细胞全能性所必需的,其中任何一种 lncRNA 的去除都能导致全能性的丧失。一些 lncRNAs 与染色质重构蛋白复合物相互结合后,能为一些维持全能性所必需的复合物起到基因支架的作用,还有一些 lncRNAs 能根据环境中的变化维持干细胞全能性,起到“环境传感器”的作用。

比如最近发现一种叫做 ANCR(反变异非编码 RNA)的 lncRNA 在上皮细胞中对维持细胞未分化态是必需的^[27]。总之,lncRNAs 在对 ESC 维持全能性和促进细胞分化上作用非常重要。除遗传学因素之外,化学抑制剂,信号分子和 miRNAs 等因素也对 ES 细胞全能性的维持起到重要作用。

2.4 调控 microRNA 以及小 RNAs

2007 年,一项关于拟南芥的研究表明 lncRNA IPS1 能与 microRNA miR-399 结合并限制其调控 PHO2 mRNA 的能力^[28]。一些哺乳动物的 lncRNAs 也能通过结合 miRNA 来防止其结合到靶 mRNA 上。在 2011 年的一项研究发现,一种叫做 linc-MD1 的 lncRNA,能作为两种 miRNA 的共同作用物来调控肌肉分化中关键的转录因子^[29],起到一种“海绵”作用,这表明 lncRNAs 能通过协同作用调控基因表达。另外,和 lncRNAs 相关的异常蛋白伴侣也能在空间和时间水平上对 lncRNAs 和 miRNAs 间相互作用起到重要影响。

关于 lncRNAs 作为小 RNA 前体所占的比例还不是很清楚,已经证实存在许多长非编码蛋白转录片段为类似 miRNAs 的小 RNAs 的前体^[30]。

最初认为多数小 RNAs 来自于基因间区域,但是大部分 ncRNAs 坐落于内含子区域(如 miRNAs 和 snoRNAs),并且通过 RNA 聚合酶 II 转录而作为它们的转录单元。长内含子 ncRNAs 具备多种生物学功能,而且包含一些功能性调节因子(包括 miRNAs)。与内含子 snoRNAs 相比,研究发现脊椎动物中的一些内含子 miRNAs 会在剪切催化作用发挥之前在相对于它们宿主基因转录的反方向上加工而成,这个过程是在非拼接(unspliced)内含子区域进行的,但它们也可能在正方向上转录形成 pre-mRNA 的基因结构,并进一步加工形成成熟 miRNAs。这些证据表明长内含子 lncRNAs 同样具有重要的功能^[31]。我们熟知的两种双胞胎 lncRNAs Xist 和 Tsix 在 X 染色体失活中具有重要的调控作用,同时它们也能加工产生一些小 RNAs。报道显示,在 Xist 和 Tsix 链上检测到 24~42nt 长度片段的小 RNAs^[20],它们可能是通过活性 X 染色体上的 Dicer 依赖方式,组成一种二连单位然后加工生成小 RNAs。

2.5 介导基因的可变剪接

lncRNAs 能加工产生一些小 RNAs,也可通过调节剪切小 RNAs 的能力来影响 miRNAs 的分布及剪接方式。

前体 mRNAs 的可变剪接增加了细胞基因表达的

复杂性,报道显示,一种 lncRNA, MALAT 1 (肺腺癌转移关联转录产物 1, 主要集中于细胞核) 在前体 mRNA 的可变剪接过程中具有重要的作用。它的核酸蛋白复合物中含有一些涉及可变剪接的蛋白质, 能调节 SR 蛋白 (一类丝氨酸 / 精氨酸富集的能调节并选择前体 mRNAs 剪接位点的蛋白质) 的磷酸化作用和细胞内活性水平。MALAT 1 能与 SR 剪接因子相互作用, 并通过调整相互作用物在核小点上的分布情况调控许多前体 mRNAs 的剪接^[32]。随着 RNA 免疫共沉淀以及 RNA 深度测序等先进技术的应用, 会有更多具有类似功能的 lncRNAs 被鉴定出来, 例如发现有自然反义 lncRNA 能参与调控 Zeb 2 蛋白的可变剪接^[33]。

2.6 参与形态发生的调控

lncRNAs 同样影响许多形态发生进程, 该过程涉及突触发生和组织重排等, 在最基本的形态发生过程中, ncRNAs 在上皮细胞到间质细胞转变进程 (EMT) 以及突触可塑性上所起的作用已被证实。

EMT 是一个表型发生退化和基因表达谱发生改变的上皮细胞, 变成具有迁移能力的间质细胞的进程。这种形态学变化伴随有特定上皮细胞基因表达的下调和间质细胞基因表达的上调, 它们通过 Snail, Zeb1 和 Zeb2 的锌指转录抑制剂同源物调控^[34], 这些抑制剂能作为 EMT 发生的诱导物。最近, miR-200 家族成员被认为是肿瘤细胞中基因表达重要的调控者^[35]。与 miR-200 相比, Zeb2 自然反义转录物 (Zeb2-NAT) 能作为剪接调控者来增强 Zeb2 蛋白的表达, 然后在间质细胞中抑制钙粘蛋白表达。Snail 激活 Zeb2-NAT 转录继而抑制钙粘蛋白表达, 两种过程都是直接结合到钙粘蛋白启动区, 剪接促进 Zeb2 翻译^[33]。lncRNAs 不仅能以剪接调节器形式来调控 EMT 发生, 还会在染色质形态上通过结合 PRC2 来起作用, 在完整 PRC2 存在的情况下, HOTAIR 能促进初级乳腺上皮组织中癌细胞的侵袭性和转移性, 另外据报道称许多由 HOTAIR 产生的基因都是 EMT 发生的顺式调控者, 如 Snail 基因。

在对神经中枢连接性和突触可塑性的调节过程中, Evf2 会对邻近编码序列招募转录因子起到促进作用。在 *Evf2*/TS 功能异常的小鼠中突触活动中显示出异常现象, 这是由于其海马体和齿状回中氨基丁酸中间神经元的畸变导致^[36]。但是, 人工培养的海马神经元中的 Malat1/Neat2 功能的丧失和增加能导致突触密度的降低和升高, 在 Malat1/Neat2 高度表达的神经元中, Malat1/Neat2 会通过调节对 SR 蛋白的招募作用,

控制一系列涉及细胞核和突触功能的基因的表达, 从而导致突触发生^[37]。这些研究表明 lncRNAs 会通过转录调控和转录后调控来维持神经中枢的连接性以及突触可塑性的保真度。

2.7 作为假基因成员

假基因一直以来被视为功能基因的“弃儿”, 但近年来其相关研究越来越引起人们的关注。例如, Makorin 1 的假基因通过生成的一种缩短了的非编码片段来增加同源基因转录的 mRNA 的稳定性。另外发现, 假基因也可通过反义序列的转录来影响同源基因的表达, 如假基因 asOct4-ps5 能与 Oct4 mRNA 相互作用, 继而生成交叉沉默复合物结合到 Oct4 启动子区来调节 Oct4 表达。PTENP1 的 3'UTR 是一种肿瘤抑制假基因, 它能结合到 miRNAs 调控区域也能结合到抑癌基因 *PTEN* 上, 既能导致 *PTEN* mRNA 的下调, 也能使它翻译为肿瘤抑制蛋白 PTEN, 有报道显示 PTENP1 在人类肿瘤中存在选择性的丢失。这类发现说明了假基因能通过非编码 RNA 参与对同源基因的调控。

3 与肿瘤发生之间的关系

lncRNAs 的作用会涉及一系列疾病的发生, 例如癌症, 心脏病, 阿尔茨海默症, 牛皮癣, 脊髓小脑共济失调症 8 和 X 染色体易损综合症^[38]。这里重点介绍其与肿瘤的关系。

大量由临床观察和实验研究所得的结果表明, lncRNAs 的异常表达会导致很多疾病的发生, 这种机制在癌生物学中尤为重要。尽管涉及肿瘤的大多数 lncRNAs 的作用机制还没得到准确阐明, 但就其对部分已知 lncRNA 的研究发现, 其参与染色质重构和组蛋白修饰的调控, 特别是通过表观遗传的修饰调控 (例如基因印迹和组蛋白修饰) 会影响肿瘤的发生和迁移。例如, HOTAIR 能通过调控染色质位点重排促进肿瘤转移, 而 linc-P21 通过 p53 途径将 RNA / DNA 结合蛋白 hnRNPK 结合到染色质上实现其抑制剂作用^[18]。lncRNA 序列的基因突变 (如碱基增添, 缺失和序列突变), 也是肿瘤发生的原因之一。例如在线粒体 RNA 处理内切核糖核酸酶的 RNA 组分的序列突变会导致软骨皮毛发育不全综合症, 同样能增加一些肿瘤的发生风险^[39]。近代基因组学研究发现, lncRNAs 与编码序列存在相互作用关系, 但是这种关系并不涉及编码序列本身发生变化^[40], 这个现象值得引起人们注意, 它预示着 lncRNAs 能参与蛋白编码基因的表观遗传调

控,能为治疗肿瘤等疾病提供一种新的思路。另外某些 lncRNA 基因具有癌基因和抑癌基因的双重功能,例如第一个被发现的非编码 RNA 基因 *H19*,在结肠癌中起到癌基因的作用,而在畸胎瘤、肝癌等模型的研究中发现,缺少 *H19* 能导致肿瘤的发生,表明它起到了抑癌基因的作用^[41]。对于 *H19* 基因的调控机制我们尚不明瞭,但随着相关研究的不断深入,我们对于复杂基因表达调控和生物学功能的认知必将进入一个新的阶段。

LncRNAs 可以作为某些肿瘤发生的抑制剂。母系表达基因 3 (*MEG3*)是首次由 Miyoshi 等在 2000 年发现的小鼠母系表达基因,它是基因捕获基因座 2 (gene-trap locus 2, Gtl2)的人类同系物,定位于染色体 14q32.3,长度约为 1.6 kb。*MEG3* 能编码出一条在许多正常组织中表达的 lncRNA,但是它在初级肿瘤患者和肿瘤细胞系中是缺失的(包括基因缺失,启动子高度甲基化和基因间位点的甲基化等多重机制是造成 *MEG3* 基因缺失表达的主要因素)。研究发现,*MEG3* 的重新表达能促进细胞凋亡,从而抑制肿瘤细胞增殖。*MEG3* 还能

造成 p53 的积累,刺激 p53 介导的转录激活并且选择性地调控 p53 靶基因表达。例如在小鼠中 *MEG3* 的去除能导致小鼠骨骼肌缺陷和围产期死亡^[42],而 *MEG3* 的失活能导致大脑中促血管生成基因表达增加和微脉管形成,这些证据都表明 *MEG3* 能作为一个 lncRNA 肿瘤抑制剂。

另外 lncRNAs 的表达与肿瘤发生的组织学阶段相联系,可能会是肿瘤检测的一种有效的标志分子。在非小细胞肺癌 (NSCLC) 里,转移性肿瘤中 *MALAT1* 的表达量是非转移性肿瘤的三倍,它的高表达能引起病人的不良预后^[43]。在转移性乳腺癌中,*HOTAIR* 的表达量是普通乳腺癌上皮细胞的 125 倍^[44],它在初级乳腺癌患者中的过表达是癌症转移的预兆。另一种 lncRNA 基因 *PCA3* 被认为是一种与前例腺癌发生相关的标志物,通过对血浆、尿液以及其他临床样本的 *PCA3* 定量分析结果表明其对于前列腺癌的早期诊断和分子分期具有重要的指示作用^[45]。

就目前已经鉴别过的常见 lncRNA 及其相关的肿瘤总结见表 1。

表 1 常见的与肿瘤相关的 lncRNA

Table1 lncRNAs that have been or might be linked to cancer

LncRNA	分子机制 Molecular mechanism	肿瘤类型 Tumor	基因组中位置 Genome position	长度 Size(nt)
aHIF	信使 RNA 衰退	多种肿瘤	hg19 chr14:61,283,843-61,285,036	1 577
ak023948	NA(Not Available)	乳突状甲状腺癌	hg18 chr8:134136386-134139194	2 807
anril	反义转录调控	前列腺癌	hg18 chr9:21,984,790-22,111,091	3 857
bc200	蛋白结合	多种肿瘤	hg19 chr2:47,562,454-47,562,653	200
gas5	糖皮质激素受体被诱	乳腺癌	hg18 chr1:172,099,662-172,103,748	651
h19	转录控制	多种肿瘤	hg18 chr11:1,972,982-1,975,641	2 322
HOTAIR	表观遗传调控	多种肿瘤	hg18 chr12:52,642,363-52,648,782	2 377
HULC	microRNA 被诱	多种肿瘤	hg18 chr6:8,597,441-8,599,080	500
kraspl	microRNA 被诱	前列腺癌	hg18 chr6:54,743,128-54,743,996	6 801
KvlQT1-AS (Kenq1ot1)	DNMT1 相互作用和转录基因沉默	结肠癌	hg19 chr11:2,465,330-2,870,445	59 461
LEU2	microRNA 沉默	慢性淋巴细胞肿瘤	hg19 chr13:50,556,688-50,699,677	3 035,3 177
LOC285194	NA	骨肉瘤	hg18 chr3:117,911,325-117,918,575	961
MALAT-1 (NEAT2)	RNA 剪接,小 RNA 加工和蛋白间作用	多种肿瘤	hg18 chr11:65,021,809-65,030,513	8 708
MEG3	NA	多种肿瘤	hg18 chr14:100,362,198-100,397,121	1 595-1 855
MER11C	RNA 蛋白间相互作用及基因表达调控	细胞系	hg18 chr11:50,410,308-50,411,367	1 065
NDM29	NA	成神经细胞瘤	hg18 chr11:8,917,158-8,917,288	351
p53 mRNA	RNA 蛋白质结合	多种肿瘤	hg19 chr17:7571720-7590863	2 953
PCA3/DD3	NA	前列腺癌	hg18 chr9:78,569,172-78,592,305	3 735,3 923
PCGEM1	NA	前列腺癌	hg18 chr2:193,322,816-193,349,870	1 603
PTENP1	microRNA 被诱,抑制细胞生长	前列腺癌	hg18 chr9:33,663,502-33,667,418	3 917

续表 1

LncRNA	分子机制 Molecular mechanism	肿瘤类型 Tumor	基因组中位置 Genome position	长度 Size(nt)
RMRP	线粒体 RNA 加工核酸内切酶, hTERT 依赖的小干扰 RNA 途径	白血病和淋巴瘤	hg19 chr9:35,657,750-35,658,014	348
RPS6KA2 反义转录片段	NA	细胞系	未知	未知
SRA RNA	RNA-蛋白结合转录因子共催化剂	乳腺癌	hg19 chr5:139,930,090-139,937,036	875
TERC	端粒模板	多种肿瘤	hg19 chr3:169,481,881-169,483,646	1 765
terra	端粒酶调控	多种肿瘤	未知	未知
ucl	NA	膀胱癌	hg19 chr19:15,939,757-15,946,226	1 441
xist	X 染色体失活	多种肿瘤	hg19 chrX:73,043,280-73,072,588	12 250,17 918
Zfas1	NA	乳腺癌	hg19 chr20:47,894,715-47,905,797	516-1 020

4 展 望

lncRNAs 作为一种新型肿瘤标志物,它在肿瘤发生、发展的调控过程中起着非常重要的作用,这种机制能为人们治疗肿瘤提供新的思路。目前人们对绝大多数 lncRNAs 的生物和分子特性所知甚少,但是随着生物信息学发展,二级结构以及三级结构分析会为 lncRNAs 的研究提供重要的手段,特别是其与编码序列之间的相互作用关系,这一方向的发展将会拓展出全新的研究领域。

lncRNAs 在疾病的分级检测、分子分型以及疗效评估中同样会展现出巨大的应用前景,它作为药物作用的靶点对新型靶向药物的开发具有重要的指导意义,如通过 RNA 干扰介导的基因沉默疗法选择性地沉默 lncRNAs,可用于某些肿瘤的治疗,另外 lncRNAs 基因的遗传和表观遗传上的改变同样有助于其病理学机制研究。尽管它们在肿瘤中的角色才刚刚被阐述,但关于 lncRNAs 的研究必会成为除 RNA 依赖治疗疗法外的另一种新的生物诊断标志。随着药物递送系统以及 RNA 干扰靶效应技术的改善,其发展已经展现出美好的应用前景,值得我们期待和做进一步研究。

参考文献

[1] Tian D, Sun S, Lee J T. The long noncoding RNA, Jpx, is a molecular switch for X chromosome inactivation. *Cell*, 2010, 143 (3): 390-403.

[2] Ponting C P, Oliver P L, Reik W. Evolution and functions of long noncoding RNAs. *Cell*, 2009, 136(4): 629-641.

[3] Mercer T R, Dinger M E, Mattick J S. Long non-coding RNAs: insights into functions. *Nat Rev Genet*, 2009, 10(3): 155-159.

[4] Spizzo R, Almeida M I, Colombatti A, et al. Long non-coding

RNAs and cancer: a new frontier of translational research?. *Oncogene*, 2012.

[5] Clemson C M, Hutchinson J N, Sara S A, et al. An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. *Mol Cell*, 2009, 33(6): 717-726.

[6] Spector D L. SnapShot: Cellular bodies. *Cell*, 2006, 127(5): 1071.

[7] Kohtz J D, Fishell G. Developmental regulation of EVF-1, a novel non-coding RNA transcribed upstream of the mouse Dlx6 gene. *Gene Expr Patterns*, 2004, 4(4): 407-412.

[8] Feng J, Bi C, Clark B S, et al. The Evf-2 noncoding RNA is transcribed from the Dlx-5/6 ultraconserved region and functions as a Dlx-2 transcriptional coactivator. *Genes Dev*, 2006, 20 (11): 1470-1484.

[9] Redrup L, Branco M R, Perdeaux E R, et al. The long noncoding RNA Kcnq1ot1 organises a lineage-specific nuclear domain for epigenetic gene silencing. *Development*, 2009, 136 (4): 525-530.

[10] Imamura T, Yamamoto S, Ohgane J, et al. Non-coding RNA directed DNA demethylation of Sphk1 CpG island. *Biochem Biophys Res Commun*, 2004, 322(2): 593-600.

[11] Kloc M, Zearfoss N R, Etkin L D. Mechanisms of subcellular mRNA localization. *Cell*, 2002, 108(4): 533-544.

[12] Martin K C, Ephrussi A. mRNA localization; gene expression in the spatial dimension. *Cell*, 2009, 136(4): 719-730.

[13] Wang K C, Chang H Y. Molecular mechanisms of long noncoding RNAs. *Mol Cell*, 2011, 43(6): 904-914.

[14] Clark M B, Mattick J S. Long noncoding RNAs in cell biology. *Semin Cell Dev Biol*, 2011, 22(4): 366-376.

[15] Augui S, Nora E P, Heard E. Regulation of X-chromosome inactivation by the X-inactivation centre. *Nat Rev Genet*, 2011, 12(6): 429-442.

[16] Khalil A M, Guttman M, Huarte M, et al. Many human large intergenic noncoding RNAs associate with chromatin-modifying

- complexes and affect gene expression. *Proc Natl Acad Sci U S A*, 2009, 106(28): 11667-11672.
- [17] Guttman M, Donaghey J, Carey B W, et al. lincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature*, 2011, 477(7364): 295-300.
- [18] Huarte M, Guttman M, Feldser D, et al. A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell*, 2010, 142(3): 409-419.
- [19] Hung T, Wang Y, Lin M F, et al. Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. *Nat Genet*, 2011, 43(7): 621-629.
- [20] Ogawa Y, Sun B K, Lee J T. Intersection of the RNA interference and X-inactivation pathways. *Science*, 2008, 320(5881): 1336-1341.
- [21] Pontier D B, Gribnau J. Xist regulation and function explored. *Hum Genet*, 2011, 130(2): 223-236.
- [22] Zhao J, Sun B K, Erwin J A, et al. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science*, 2008, 322(5902): 750-756.
- [23] Khalil A M, Driscoll D J. Trimethylation of histone H3 lysine 4 is an epigenetic mark at regions escaping mammalian X inactivation. *Epigenetics*, 2007, 2(2): 114-118.
- [24] Barr M L, Bertram E G. A morphological distinction between neurones of the male and female, and the behaviour of the nucleolar satellite during accelerated nucleoprotein synthesis. *Nature*, 1949, 163(4148): 676.
- [25] Guttman M, Amit I, Garber M, et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature*, 2009, 458(7235): 223-227.
- [26] Sheik Mohamed J, Gaughwin P M, Lim B, et al. Conserved long noncoding RNAs transcriptionally regulated by Oct4 and Nanog modulate pluripotency in mouse embryonic stem cells. *RNA*, 2010, 16(2): 324-337.
- [27] Kretz M, Webster D E, Flockhart R J, et al. Suppression of progenitor differentiation requires the long noncoding RNA ANCR. *Genes Dev*, 2012, 26(4): 338-343.
- [28] Franco-Zorrilla J M, Valli A, Todesco M, et al. Target mimicry provides a new mechanism for regulation of microRNA activity. *Nat Genet*, 2007, 39(8): 1033-1037.
- [29] Cesana M, Cacchiarelli D, Legnini I, et al. A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell*, 2011, 147(2): 358-369.
- [30] Cai X, Cullen B R. The imprinted H19 noncoding RNA is a primary microRNA precursor. *RNA*, 2007, 13(3): 313-316.
- [31] Nott A, Meislin S H, Moore M J. A quantitative analysis of intron effects on mammalian gene expression. *RNA*, 2003, 9(5): 607-617.
- [32] Tripathi V, Ellis J D, Shen Z, et al. The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol Cell*, 2010, 39(6): 925-938.
- [33] Beltran M, Puig I, Pena C, et al. A natural antisense transcript regulates Zeb2/Sip1 gene expression during Snail1-induced epithelial-mesenchymal transition. *Genes Dev*, 2008, 22(6): 756-769.
- [34] Cano A, Nieto M A. Non-coding RNAs take centre stage in epithelial-to-mesenchymal transition. *Trends Cell Biol*, 2008, 18(8): 357-359.
- [35] Chang C J, Chao C H, Xia W, et al. p53 regulates epithelial-mesenchymal transition and stem cell properties through modulating miRNAs. *Nat Cell Biol*, 2011, 13(3): 317-323.
- [36] Bond A M, Vangompel M J, Sametsky E A, et al. Balanced gene regulation by an embryonic brain ncRNA is critical for adult hippocampal GABA circuitry. *Nat Neurosci*, 2009, 12(8): 1020-1027.
- [37] Bernard D, Prasanth K V, Tripathi V, et al. A long nuclear-retained non-coding RNA regulates synaptogenesis by modulating gene expression. *EMBO J*, 2010, 29(18): 3082-3093.
- [38] Taft R J, Pang K C, Mercer T R, et al. Non-coding RNAs: regulators of disease. *J Pathol*, 2010, 220(2): 126-139.
- [39] Wang F, Li X, Xie X, et al. UCA1, a non-protein-coding RNA up-regulated in bladder carcinoma and embryo, influencing cell growth and promoting invasion. *FEBS Lett*, 2008, 582(13): 1919-1927.
- [40] Caley D P, Pink R C, Trujillano D, et al. Long noncoding RNAs, chromatin, and development. *ScientificWorldJournal*, 2010, 10:90-102.
- [41] Smits G, Mungall A J, Griffiths-Jones S, et al. Conservation of the H19 noncoding RNA and H19-IGF2 imprinting mechanism in therians. *Nat Genet*, 2008, 40(8): 971-976.
- [42] Zhou Y, Zhang X, Klibanski A. MEG3 noncoding RNA: a tumor suppressor. *J Mol Endocrinol*, 2012, 48(3): R45-R53.
- [43] Ji P, Diederichs S, Wang W, et al. MALAT-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer. *Oncogene*, 2003, 22(39): 8031-8041.
- [44] Yang Z, Zhou L, Wu L M, et al. Overexpression of long non-coding RNA HOTAIR predicts tumor recurrence in hepatocellular carcinoma patients following liver transplantation. *Ann Surg Oncol*, 2011, 18(5): 1243-1250.
- [45] Hessels D, Klein Gunnewiek J M, Van Oort I, et al. DD3 (PCA3)-based molecular urine analysis for the diagnosis of prostate cancer. *Eur Urol*, 2003, 44(1): 8-15; discussion -6.

Long Noncoding RNA and Its Relationship with Cancer

LOU Liang-liang¹ ZHU Yun-feng^{1,2}

(1 The Institute of Life Science and Bio-Engineering, Beijing Jiaotong University, Beijing 100044, China)
(2 The Key Laboratory in Tumor Center of PLA General Hospital, Beijing 100853, China)

Abstract The data from human genome indicated that coding sequences are less than 2% and rest of them are considered to be non-coding sequences. The transcripts derived from non-coding sequences are termed as non-coding RNA. Based on the length of non-coding RNAs (ncRNAs), ncRNAs are categorized into long non-coding RNA(lncRNA, greater than 200 nt)and short non-coding RNA(sncRNA). Most lncRNAs are transcribed by RNA polymerase II, which make them have some characters of mRNA, eg: 3' polyadenylation tail and 5' CAP structure. In human,lncRNAs are identified to have a wide distribution in genome and some of which may contribute to tumorigenesis and development. The feature of lncRNAs and their functions in tumor cells were focused in this review.

Key words Non-coding RNA LncRNA Tumor cells Genome

VICAM 新型全球真菌毒素法规在线数据库 可即时查询全球法规数据

沃特世公司近日宣布 Waters® 旗下业务 VICAM® 重磅推出独家在线数据库工具,通过此工具可访问简便易用的综合性全球真菌毒素法规数据库(Global Mycotoxin Regulations Tool™)。

全球真菌毒素法规在线数据库(commodityregs.com)应对纷繁的各国食品进口要求开创了全新的解决方案,可确保出口货物符合全球食品安全标准,满足客户面对不断变化的全球市场的需求,逐渐降低真菌毒素对人体和动物健康的影响,从而对解决全球范围内食品安全问题产生极大的推动作用。该数据库以络或移动式应用为基础,通过区域、商品名称或食品类型划分,可在几秒钟内帮助客户快速获得全球真菌毒素限制要求。食品制造商和出口商无需进行大量网页搜索即可获得关键的法规数据,从而有效节约时间和资源。此外,用户还可以使用原料和成品图片搜索产品,简化搜索过程的同时确保全球食品和农业市场相关人员获取清晰且可追溯的结果。