

宿主限制长散布元件转座活性的机制*

梁志滨¹⁾ 梁 臣^{2, 3)} 耿运琪¹⁾ 乔文涛^{1)**}

¹⁾ 南开大学生命科学学院, 天津 300071; ²⁾ *Department of Medicine, McGill University, Montreal, Quebec, Canada H3A 2B4;*

³⁾ *Department of Microbiology And Immunology, McGill University, Montreal, Quebec, Canada H3A 2B4)*

摘要 转座子约占人类基因组的 45%，对基因组的结构与功能造成了重大的影响。一部分转座子现在仍然具有活性，它们的转座能引发疾病。长散布元件(long interspersed element-1, LINE-1)是现今在人类基因组中发现的唯一具有活性并能自主转座的转座子，并能介导非自主转座的元件进行转座。近年来，LINE-1 的研究有新的突破，本文简述了 LINE-1 的结构、转座机制及对基因组的影响，重点总结和分析宿主对 LINE-1 的限制机制。由于 LINE-1 的生活周期与逆转录病毒有相似之处，也希望能够为宿主抗病毒的研究提供线索。

关键词 LINE-1, 转座, 基因组, 限制

学科分类号 Q26, Q31

DOI: 10.3724/SP.J.1206.2012.00506

转座子(transposable elements, TEs), 是指能在基因组内或基因组间移动的 DNA 片段。DNA 片段的可移动性首先报道于 1948 年, 美国遗传学家 McClintock 在研究玉米杂色成因的过程中发现了这一现象^[1], 转座子后来被证实广泛存在于原核和真核生物基因组^[2]。人类基因组测序的完成明确了各类转座子占据了人类基因组的 45%^[3], 进一步揭示了转座子的重要性。

转座子曾经被认为是“自私的”寄生序列, 对基因组百害而无一利, 是“废物 DNA”。然而, 基因组在进化过程中何以积累了如此之多的“废物”? 研究表明, 转座子在基因组进化及其他生命过程中起了重要作用, 但同时也令基因组的结构和稳定性发生改变, 因而宿主进化出多种机制来控制转座子的活性, 使之成为自身服务, 实现了“变废为宝”。因此, 全面了解转座子及宿主对其限制机制对解析基因组及其生物学功能至关重要。

转座子可依据转座方式分为 DNA 转座子和 RNA 转座子。DNA 转座子占据人类基因组的 3%, 尚未发现有活性^[4]。近年来发现 RNA 转座子对哺乳动物基因组的稳定性和功能具有重要影响。其中唯一具有自主转座活性的长散布元件(LINE-1)作用

显著。本文介绍 LINE-1 转座子对基因组影响的研究进展, 着重对宿主限制 LINE-1 转座活性的机制进行综述。

1 逆转录转座子及 LINE-1

RNA 转座子又称为逆转录转座子, 通过“复制-粘贴”的方式进行转座, 先转录成 RNA, 再逆转录成 DNA 并插入到新的位置。根据能否编码用于自身转座的蛋白质, 逆转录转座子可分为自主转座的和非自主转座的。自主转座的逆转录转座子可再分为含长末端重复序列(long terminal repeats, LTR)和不含 LTR 的逆转录转座子。

含 LTR 的逆转录转座子, 又称内源逆转录病毒, 占据人类基因组的 8%, 尚未发现具有活性。然而, 它在哺乳动物中是有活性的, 如鼠的脑池内 A 型粒子(intracisternal A type particle, IAP)和鼠 D 型逆转录转座子(mouse type D retrotransposon, MusD)^[5]。

* 十二五重大专项(2012ZX10001006), 国家自然科学基金(81271812)和天津市应用基础及前沿技术研究计划(12JCQNJC06100)资助项目。

** 通讯联系人。

Tel: 022-23504547, E-mail: wentaoqiao@nankai.edu.cn

收稿日期: 2012-11-28, 接受日期: 2013-01-16

非 LTR 逆转录转座子, 即长散布元件(long interspersed element, LINE), 占人类基因组 21%^[3], 其中只有 LINE-1 是有活性的, 仍对基因组产生重大影响^[5-7]. LINE-1 在基因组中存在了至少 1.5 亿年, 在人类基因组中的拷贝超过 500 000 个, 约占人类基因组的 17%. 大部分 LINE-1 序列都因 5' 端截短、重排或突变而不能转座, 但将近 100 个拷贝仍有转座功能^[6].

非自主转座的逆转录转座子包括短散布元件(short interspersed element, SINE)和逆转录假基因(也称已加工的假基因), 能利用 LINE-1 编码的蛋白质进行转座, 占人类基因组的 13%, 仍具有活性. 其中 SINE 又可再分为 Alu 家族和 SVA(SINE-R/VNTR/Alu). Alu 在人类基因组中占 10%, 据推测其中有数千个 Alu 具有活性^[8].

1.1 LINE-1 的结构

LINE-1 全长为 6 kb, 可分为 5'UTR, 两个开放读码框 ORF1 和 ORF2, 以及 3'UTR. 5'UTR 包含一个内部的 RNA 聚合酶 II 型启动子, 能起始从元件 5' 端开始的转录; 5'UTR 中还有一个反向启动子, 并含有 SOX、YY-1 和 RUNX3 转录因子的结合位点^[9]. 3'UTR 包含一个多腺苷酸化信号并伴随着长度不等的多腺苷酸(polyA)尾(图 1a).

LINE-1 的转录本只有一个, 即全长的 LINE-1

RNA, 它是个双顺反子, 能表达出 ORF1p 和 ORF2p 两个蛋白质. ORF1p 是 40 ku 的 RNA 结合蛋白, 可分为 3 个结构域, 分别为 N 端的 Coiled Coil 结构域(CC), 中部的 RNA 识别模序(RRM)和 C 端结构域(CTD). CC 令 ORF1p 聚合成活性形式的三聚体, RRM 和 CTD 均对 RNA 结合有着重要作用^[10]. ORF2p 大小为 150 ku, 具有核酸内切酶(endonuclease, EN)和反转录酶(reverse transcriptase, RT)活性. ORF1p 和 ORF2p 具有顺式结合活性, 翻译出来后就与编码它们的 RNA 结合^[11], 并一起入核进行转座(图 1b).

1.2 LINE-1 的转座机制

LINE-1 的逆转录转座方式称为靶序列引发逆转录(target-primed reverse transcription, TPRT). 首先, LINE-1 通过 ORF2p 的 EN 活性在靶序列 5' TTAA 3' 的 T 和 A 间切开 DNA 的第一条链, 形成自由的 3' 羟基(3'-OH). LINE-1 RNA 通过 3' 端的 polyA 与靶 DNA 经切割而形成的 TTTT-OH 结合, 并以之为引物反转录成一条 DNA 链(图 1b). 接着, 靶序列 DNA 的第二条链被切开, 并以刚合成的 DNA 链为模板合成另一条 DNA 链. LINE-1 完成转座后, 5' 端通常经过不同程度地截短, 3' 端含有 polyA 尾, 插入的 LINE-1 序列两端会有长度不等(2~20 bp)的靶序列重复(TSD)^[4, 6, 9].

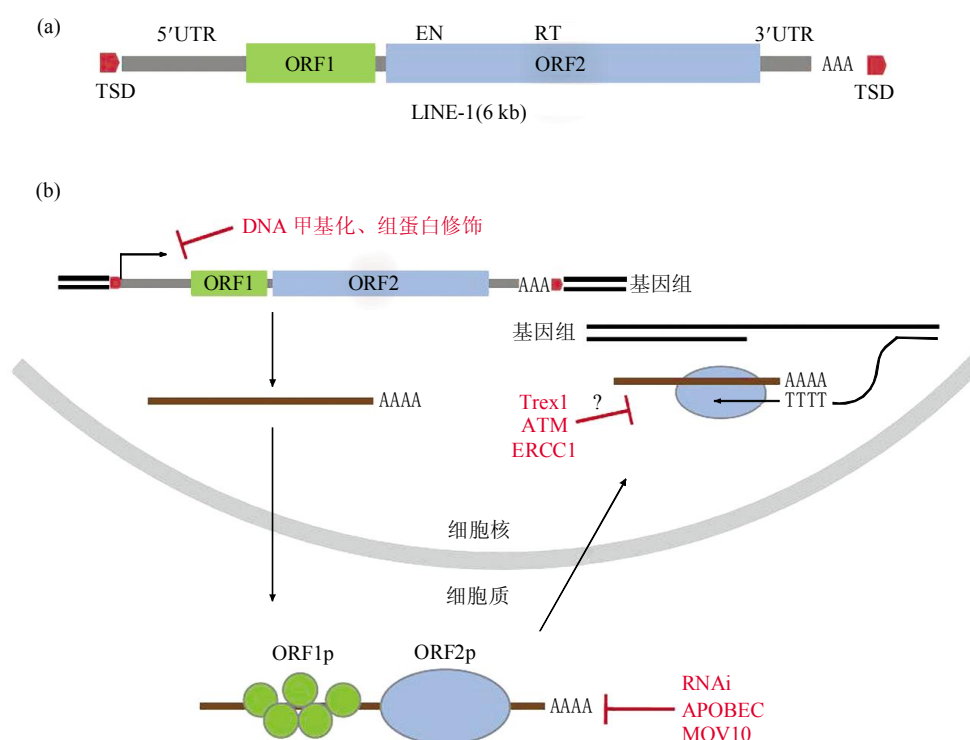


图 1 LINE-1 的结构、生活周期及细胞限制机制

(a) LINE-1 的结构. (b) LINE-1 的生活周期及细胞限制机制(用红色字体表示).

2 LINE-1 对基因组的影响

LINE-1 转座子在基因组中长期存在并持续具有活性, 已对基因组的结构、稳定性及基因表达等均造成了巨大影响。

2.1 LINE-1 转座造成序列重复

LINE-1 在转座过程除了完成自身的复制之外, 还可能携带下游的序列共同转座(称为 3'转导)。这是由于转座子 polyA 终止信号功能较弱, LINE-1 转录时, 转录复合物常可跳过较弱的终止信号将转录本延伸至转座子 3'序列的下游, 这些序列在随后转座过程中引起序列重复。LINE-1 的 3'转导已经过体外实验得到证实^[12], 经分析发现, 人类基因组中 10%~20% 的 LINE-1 序列伴随着 3'转导序列^[13-14]。

LINE-1 协助 Alu、SVA 实现转座, 令基因组中散布着许多微卫星序列^[15]。当已加工的 mRNA 利用 LINE-1 进行转座^[16], 因为缺少基因的上游调控序列, 通常形成无功能的基因拷贝, 称逆转录假基因。

2.2 LINE-1 引起基因组的不稳定

转座子插入到基因的蛋白质编码区或调控区中能引起基因功能的丧失, 目前已发现有 65 例人类遗传病与 LINE-1 及其介导的 Alu 和 SVA 插入相关, 其中引起疾病的 LINE-1 序列大部分插入在 X 染色体中。据推测, 大约 0.3% 的人类变异可归因于 LINE-1、Alu 和 SVA 的插入^[1, 4]。

LINE-1 和 Alu 的插入还能造成靶位点的序列缺失。这个现象最早在体外实验中观察到^[17], 并随后在人类及黑猩猩的基因组中得到证实^[18]。在自然状态下, 大约 2% 的 LINE-1 插入和 0.3% 的 Alu 插入会伴随靶序列的缺失^[19]。已有报道, LINE-1 插入造成的基因片段缺失会引发疾病。据估计, 在灵长类的进化中, 总共有 45 000 例转座子插入引发的缺失, 可能删除了 > 30Mb 的基因组序列^[4]。

由于 LINE-1 与 Alu 在基因组中存在大量的拷贝, 转座子序列之间的异常重组能改变基因组的结构, 并引发疾病。LINE-1 的 ORF2p 内切酶活性会造成 DNA 双链断裂, 令基因组变得不稳定, 易发生突变和异常重组^[1, 4]。

2.3 LINE-1 影响基因表达

LINE-1 转座子能通过多种机制影响附近基因的表达。LINE-1 的启动子和反向启动子能启动两侧基因的转录^[20]; LINE-1 的 polyA 会造成转录的提前终止, 形成不成熟的转录本^[21]; 内含子中的

LINE-1 能影响 RNA 的正常剪接; LINE-1 还能通过改变靶序列的表观遗传学修饰而改变附近的染色体状态, 从而影响附近基因的转录。雌性哺乳动物其中一条 X 染色体是没有转录活性的, 称为 X 染色体失活, 据推测 X 染色体中的 LINE-1 能传递和 / 或保持沉默信号^[22]。

3 细胞对 LINE-1 转座活性的限制

转座子能通过多种途径对基因组造成影响, 机体也必须进化出一系列防御机制来控制转座, 方可幸存至今。目前文献主要涉及宿主通过 DNA 甲基化、改变核小体的修饰状态、RNA 干扰以及作用于核酸的功能蛋白质来拮抗 LINE-1 的转座活性(图 1b), 维持细胞稳态。

3.1 DNA 甲基化修饰

DNA 甲基化修饰是机体进化出来的广谱而有效的转录调控机制。在哺乳动物中, DNA 甲基化位点主要是 CpG 双核苷酸中胞嘧啶的 C5 位。高甲基化启动子区的 CpG 造成转录抑制, 相对地, 低甲基化则处于转录激活状态。正常细胞中大部分甲基化的 CpG 都位于转座子序列中, 因此, 有学者提出假设, DNA 甲基化是宿主为对抗转座子而进化出来的机制^[23]。抑制 DNA 甲基化后, LINE-1 的转录水平大幅上升^[24]。在敲除了 DNA 甲基转移酶(Dnmt)的细胞中发现 LINE-1 与 IAP 的转座活性上升^[25]。肿瘤细胞中 DNA 甲基化状态是异常的, 最近的研究表明 LINE-1 的低甲基化与肿瘤的发生有一定相关性^[26-28]。

甲基化的 DNA 序列通过募集组蛋白去乙酰化酶(HDAC)来改变染色质结构, 从而造成转录抑制。MBD(methyl-CpG-binding domain)蛋白家族能与甲基化的 CpG 序列结合, 并募集 HDAC 在靶位点形成转录抑制的异染色质。Yu 等^[29]发现 MBD 家族中的 MeCP2 蛋白能抑制 LINE-1 的转录。Muotri 等^[30]发现, 敲除 MeCP2 的小鼠中, LINE-1 在神经细胞中的转录及转座水平均有上升, 并在来源于 Rett 综合征(由 MeCP2 基因突变引起)病人体内的细胞中发现 LINE-1 ORF2 的序列明显多于对照组, 在这些细胞中 LINE-1 的转座活性更高。

3.2 组蛋白修饰

组蛋白修饰能影响染色质的状态, 调控基因的转录, 宿主同样能通过组蛋白修饰来限制转座子的转录。组蛋白 H3 Lys9 的三甲基化(H3K9me3)能引起异染色质的形成, 对转座子的抑制起着重要作用。

用. Martens 等^[31]在 Alu 和 IAP 序列中发现 H3K9me3 的富集. Maze 等^[32]发现可卡因(cocaine)能降低 LINE-1 序列的 H3K9me3 富集程度, 并引起 LINE-1 的表达水平上升.

组蛋白的乙酰化使组蛋白与 DNA 的结合能力下降, 促进转录; 相对应地, 组蛋白的去乙酰化则能抑制基因转录. LINE-1 的启动子上有两个 SOX 家族转录因子的结合位点, Muotri 等^[33]发现, Sox2 在神经前体细胞中能抑制 LINE-1 的转录, 并发现 Sox2 与组蛋白去乙酰化酶 HDAC1 聚集在被沉默的 LINE-1 启动子上. 转录抑制因子 Rb 能与转录因子家族 E2F 相互作用, 并募集 HDAC 对转录造成抑制. Montoya-Durango 等^[34]报道, Rb/E2F 复合物能与鼠源及人源的 LINE-1 启动子相结合, 并募集 HDAC1 和 HDAC2 抑制 LINE-1 转录. Garcia-Perez 等^[35]发现, LINE-1 报告基因随 LINE-1 整合进入胚胎癌细胞基因组后, 被迅速地沉默掉, 并能通过 HDAC 的抑制剂 TSA 处理而恢复. 此沉默恢复与 LINE-1 整合位点的组蛋白乙酰化水平上升, H3K9 的甲基化水平下降相关.

3.3 RNA 干扰

小干扰 RNA (siRNA) 及其引发的 RNA 干扰 (RNAi) 途径对保护机体起着至关重要的作用. 外界入侵的病毒, 机体内反向重复序列和异常转录产物等所产生的双链 RNA (dsRNA) 由核酸内切酶 Dicer 加工成 21 bp 左右的 siRNA, 并与 RNA 沉默复合物(RISC)中的 Argonaute 蛋白相结合. siRNA 靶向于与其同源的 RNA, 并通过 RISC 将靶 RNA 降解^[36]. Dicer 对 siRNA 的形成十分重要, 在人及鼠的细胞中将 Dicer 敲除之后, 发现 LINE-1 及 IAP 的转录水平均上升^[37], 说明体内的 RNAi 对限制转座子起到重要作用. LINE-1 的 5'UTR 上含有正向及反向启动子, 在两个方向上均能转录出 RNA 并形成 dsRNA, 进而被加工成靶向于 5'UTR 的 siRNA^[38].

Argonaute 蛋白家族可根据序列同源性分为 Argonaute 蛋白和 Piwi 蛋白两支, 与 Piwi 相互作用的非编码 RNA 称为 piRNA. piRNA 长度为 26~30 bp, 来源于单链 RNA 并且不需要经过 Dicer 加工. 近年来的研究表明, piRNA 在生殖细胞中对转座子的沉默至关重要. 小鼠中的 Piwi 蛋白包括 MILI、MIWI 和 MIWI2^[39]. MILI 和 MIWI2 的功能丧失均会引起 LINE-1 的转座活性上升^[39-40], 它们主要是通过建立 DNA 甲基化来抑制转座^[41-43]. 最新研究表明 MIWI 也能抑制 LINE-1 活性, 它能

对转座子的 mRNA 直接进行切割^[44].

转座子在生殖细胞引起的突变将会传递到下一代, 因此在生殖细胞中沉默转座子尤为重要. 事实上, 生殖细胞中 siRNA 和 piRNA 都对转座子起到沉默的作用. siRNA 的作用很可能在于快速应对新转座子的转座, 就像对付外源的病毒一样, 而 piRNA 则可能倾向于更持久有效地抑制转座子^[36].

3.4 APOBEC 家族蛋白

载脂蛋白 B mRNA 编辑酶 (apolipoprotein B mRNA-editing enzyme, APOBEC) 家族是一类胞嘧啶脱氨酶, 能造成 RNA 和 / 或 DNA 由 C 到 U 的突变, 因 APOBEC3G 能显著抑制 *vif* 基因缺陷的人免疫缺陷病毒 (human immunodeficiency virus, HIV) 而备受关注. APOBEC3 在啮齿类中只有一个基因, 而在灵长类中则扩充为 7 个, 分别简称为 A3A-H^[45]. 已发现 A3A、A3B、A3C 和 A3F 均能对 LINE-1 转座造成抑制, A3G 虽然并不抑制 LINE-1, 却能明显下调 Alu 的转座. 除了对非 LTR 转座子有抑制, APOBEC3 对 LTR 转座子 (如鼠源的 IAP 和 MusD) 也有明显的抑制作用^[45-47]. 灵长类中逆转录转座子的活性比啮齿类要低, 与其 APOBEC3 的基因扩充相一致, 表明了 APOBEC3 在限制转录转座子中起到重要作用.

最近还有研究表明, APOBEC 家族的另外一支, APOBEC1 对 LINE-1、IAP 和 MusD 也有抑制作用, 虽然人源的 APOBEC1 抑制作用并非十分显著, 但啮齿类的 APOBEC1 抑制效果十分明显^[48].

3.5 RNA 解旋酶 MOV10

RNA 解旋酶 MOV10 (moloney leukemia virus 10 homolog), 是 RNA 沉默复合物 (RISC) 的组成部分, 近年的研究表明其具有抗病毒功能. MOV10 能与 HIV 的 RNA 及 Gag 蛋白结合, 抑制逆转录^[49]. 最近的研究指出, MOV10 对转座子 LINE-1、Alu 及 IAP 也具有抑制作用^[50-52]. MOV10 能结合 LINE-1 的 RNA 及蛋白, 还与 PRC1 (polycomb repressive complex 1) 的相关蛋白有相互作用^[51]. 然而 MOV10 抑制转座的机制仍不明确, 有待进一步研究.

3.6 DNA 外切酶 Trex1

病毒的核酸会引发机体的干扰素通路, 机体将通过固有免疫反应降解外源的核酸. 机体内源过量堆积的核酸, 如转座子的 RNA 及 DNA, 是否也能引发免疫反应? Stetson 等^[53]发现, Trex1, 细胞中含量最多的 3' DNA 外切酶, 能够识别并切割转座

子的逆转录单链 DNA, 防止其在细胞内的积累并引发自身免疫疾病. 在 *Trex1* 的基因敲除小鼠中, 其心脏细胞大量积累转座子的单链线性 DNA (ssDNA), 其中主要是 LINE-1 和 IAP. 过表达 *Trex1* 能有效降低 LINE-1 和 IAP 的转座.

3.7 DNA 损伤修复通路

丝氨酸 / 苏氨酸激酶 ATM(*ataxia Telangiectasia Mutated*)能检测 DNA 双链断裂, 并通过磷酸化它的底物来传递信号, 进而启动 DNA 损伤修复通路^[7]. Coufal 等^[54]发现, 在 ATM 缺陷的细胞中 LINE-1 的转座能力得到提高, 并在 ATM 基因缺陷的病人脑组织样本中发现 LINE-1 的拷贝数比正常人要多.

ERCC1/XPF 异二聚体作为 DNA 内切酶在核苷酸切除修复中起到重要作用. Gasior 等^[55]发现在 ERCC1 功能缺陷的细胞中 LINE-1 的转座活性明显上升. 上述两个研究表明 DNA 损伤修复通路中存在着 LINE-1 的限制因子.

4 宿主与逆转录病毒之间的博弈

逆转录病毒的生活周期与逆转录转座子有相似之处, 下面以 HIV-1 为例, 进行简单对比. HIV-1 生活周期主要包括如下几个阶段: 通过膜融合进入细胞, 脱衣壳, 反转录, 整合进基因组; 之后通常会经历一个潜伏期, 然后转录出病毒 RNA 并合成病毒蛋白; 最后是病毒颗粒的装配、释放和成熟. 其中, 病毒基因组 RNA 必须经过逆转录成为 cDNA, 并整合入宿主基因组, 病毒基因才能够完成转录、翻译等过程, 与 LINE-1 的逆转录转座过程极为相似. 研究表明, 针对 HIV-1 生活周期各个阶段, 宿主也具有相应对抗机制. 如抑制脱衣壳过程的 Trim5 蛋白^[56], 抑制逆转录过程的 SAMHD1 蛋白^[57]以及抑制病毒释放的 Tetherin 蛋白^[58]等.

DNA 甲基化是宿主对抗逆转录病毒的重要机制. 最近的研究表明 DNA 甲基化能抑制潜伏期的 HIV-1 的活化, MBD2 和 HDAC2 募集于甲基化的 HIV-1 前病毒序列抑制转录^[59]. 宿主也通过修饰组蛋白来抑制 HIV-1 的转录, 如 H3K9me3 同样也能沉默 HIV-1^[60], 多个转录因子能募集去乙酰化酶来抑制 HIV-1 的转录^[61]. 宿主的 RNA 干扰机制具有重要的抗病毒功能, 目前已发现, 宿主能产生源于细胞的 miRNA 和源于病毒的 viRNA 来靶向于病毒 HIV-1 的 RNA^[62]. 而一些抑制 LINE-1 转座活性的蛋白质, 如 APOBEC3G 和 MOV10 等能有效抑制

Vif 缺陷或野生型的 HIV-1 的复制. 由此可见, 宿主对抗逆转录病毒复制的能力部分得益于宿主固有的限制 LINE-1 转座活性的机制.

面对宿主的限制机制, LINE-1 没有进化出相应的对抗机制, 因而被宿主有效控制. 对于相似的宿主限制, HIV-1 则进化出一系列的拮抗机制逃逸宿主的限制. 如 HIV-1 的反式激活因子 (trans-activator, Tat) 能募集组蛋白乙酰转移酶 (HAT) 来乙酰化组蛋白, 促进病毒的转录. Tat 还能抑制 Dicer 对 dsRNA 的加工, TAR (trans-activation response element) RNA 能与 Dicer 的辅助因子 TRBP 结合, 抑制 miRNA 的产生, 因而 HIV-1 能逃逸宿主的 RNA 干扰^[63]. HIV-1 的 Vif 蛋白能募集 E3 泛素连接酶复合物对 APOBEC3G 进行泛素化, 进而通过蛋白酶体将其降解. *Trex1* 能抑制 LINE-1 的转座, HIV-1 却能利用 *Trex1* 来降低自身 ssDNA 的量^[63], 从而避免引发宿主的干扰素通路.

5 结语和展望

1988 年, 在两个没有血缘关系的甲型血友病人中发现 LINE-1 插入在凝血因子 VIII 的第 14 个外显子中^[64], 宣布了 LINE-1 在人类基因组中是有活性的. 此后 20 多年的研究令我们对 LINE-1 有了一定的了解. 然而, 现今的技术手段尚不足以使我们全面了解 LINE-1 的转座机制. 例如, 虽然 LINE-1 转座报告质粒系统的建立与发展使得我们能在体外培养细胞系中检测 LINE-1 的转座活性, 但尚未能建立无细胞体系检测 ORF2p 的逆转录和整合活性. 同时, 我们尚需要了解生殖细胞和胚胎干细胞如何限制 LINE-1 活性, 因为在其发育过程中存在的一次性大规模 DNA 去甲基化^[65]会使 LINE-1 失去重要的制约因素, 此时 LINE-1 转座如何精确控制? 此外, LINE-1 在不同体细胞中的活性如何? LINE-1 与肿瘤有多大的相关性? 基因组计划的完成使我们了解到 LINE-1 在不同物种基因组中所占比例、它的分类、在基因组中的分布以及它与宿主的进化关系. 但是, LINE-1 的起源仍然是个谜. 这个寄生的基因元件究竟从何而来? 是由于外来病原体的入侵而整合进基因组的? 同时, 细胞抑制 LINE-1 转座活性的新机制也在不断被揭示, 本实验室与合作者即已发现 MOV10 对 LINE-1 转座活性的抑制与 RNAi 途径关联. 相信随着对 LINE-1 及其他转座子生物学本质研究的不断深入, 我们对人类基因组本身复杂的起源和进化过程会有更加深

刻的理解, 可以使我們更全面地認識機體維持穩態的機制, 為醫學等相關學科提供堅實的理论支持。

参 考 文 献

- [1] Goodier J, Kazazian Jr H. Retrotransposons revisited: The restraint and rehabilitation of parasites. *Cell*, 2008, **135**(1): 23–35
- [2] Beauregard A, Curcio M J, Belfort M. The take and give between retrotransposable elements and their hosts. *Annu Rev Genet*, 2008, **42**: 42587–42617
- [3] Lander E S, Linton L M, Birren B, *et al.* Initial sequencing and analysis of the human genome. *Nature*, 2001, **409**(6822): 860–921
- [4] Cordaux R, Batzer M A. The impact of retrotransposons on human genome evolution. *Nature Reviews Genetics*, 2009, **10**(10): 691–703
- [5] Baillie J K, Barnett M W, Upton K R, *et al.* Somatic retrotransposition alters the genetic landscape of the human brain. *Nature*, 2011, **479**(7374): 534–537
- [6] Beck C R, Garcia-Perez J L, Badge R M, *et al.* LINE-1 elements in structural variation and disease. *Annu Rev Genomics Hum Genet*, 2011, **12**: 12187–12215
- [7] Thomas C A, Paquola A C, Muotri A R. LINE-1 retrotransposition in the nervous system. *Annual Review of Cell and Developmental Biology*, 2012, **28**: 28555–28573
- [8] Belancio V P, Hedges D J, Deininger P. Mammalian non-LTR retrotransposons: for better or worse, in sickness and in health. *Genome Res*, 2008, **18**(3): 343–358
- [9] Babushok D V, Kazazian H H. Progress in understanding the biology of the human mutagen LINE-1. *Human Mutation*, 2007, **28**(6): 527–539
- [10] Khazina E, Truffault V, Buttner R, *et al.* Trimeric structure and flexibility of the L1ORF1 protein in human L1 retrotransposition. *Nat Struct Mol Biol*, 2011, **18**(9): 1006–1014
- [11] Kulpa D A, Moran J V. Cis-preferential LINE-1 reverse transcriptase activity in ribonucleoprotein particles. *Nat Struct Mol Biol*, 2006, **13**(7): 655–660
- [12] Moran J V, DeBerardinis R J, Kazazian H H, Jr. Exon shuffling by L1 retrotransposition. *Science*, 1999, **283**(5407): 1530–1534
- [13] Goodier J L, Ostertag E M, Kazazian H H, Jr. Transduction of 3'-flanking sequences is common in L1 retrotransposition. *Hum Mol Genet*, 2000, **9**(4): 653–657
- [14] Xing J, Wang H, Belancio V P, *et al.* Emergence of primate genes by retrotransposon-mediated sequence transduction. *Proc Natl Acad Sci USA*, 2006, **103**(47): 17608–17613
- [15] Arcot S S, Wang Z, Weber J L, *et al.* Alu repeats: a source for the genesis of primate microsatellites. *Genomics*, 1995, **29**(1): 136–144
- [16] Esnault C, Maestre J, Heidmann T. Human LINE retrotransposons generate processed pseudogenes. *Nat Genet*, 2000, **24**(4): 363–367
- [17] Gilbert N, Lutz-Prigge S, Moran J V. Genomic deletions created upon LINE-1 retrotransposition. *Cell*, 2002, **110**(3): 315–325
- [18] Morisada N, Rendtorff N D, Nozu K, *et al.* Branchio-oto-renal syndrome caused by partial EYA1 deletion due to LINE-1 insertion. *Pediatr Nephrol*, 2010, **25**(7): 1343–1348
- [19] Callinan P A, Wang J, Herke S W, *et al.* Alu retrotransposition-mediated deletion. *J Mol Biol*, 2005, **348**(4): 791–800
- [20] Alexandrova E A, Olovnikov I A, Malakhova G V, *et al.* Sense transcripts originated from an internal part of the human retrotransposon LINE-1 5' UTR. *Gene*, 2012, **511**(1): 46–53
- [21] Han J S, Szak S T, Boeke J D. Transcriptional disruption by the L1 retrotransposon and implications for mammalian transcriptomes. *Nature*, 2004, **429**(6989): 268–274
- [22] Chow J, Heard E. X inactivation and the complexities of silencing a sex chromosome. *Curr Opin Cell Biol*, 2009, **21**(3): 359–366
- [23] Yoder J A, Walsh C P, Bestor T H. Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet*, 1997, **13**(8): 335–340
- [24] Woodcock D M, Lawler C B, Linsenmeyer M E, *et al.* Asymmetric methylation in the hypermethylated CpG promoter region of the human L1 retrotransposon. *J Biol Chem*, 1997, **272**(12): 7810–7816
- [25] Kato Y, Kaneda M, Hata K, *et al.* Role of the Dnmt3 family in *de novo* methylation of imprinted and repetitive sequences during male germ cell development in the mouse. *Hum Mol Genet*, 2007, **16**(19): 2272–2280
- [26] Antelo M, Balaguer F, Shia J, *et al.* A high degree of LINE-1 hypomethylation is a unique feature of early-onset colorectal cancer. *PloS one*, 2012, **7**(9): e45357
- [27] Stricker I, Tzivras D, Nambiar S, *et al.* Site- and grade-specific diversity of LINE1 methylation pattern in gastroenteropancreatic neuroendocrine tumours. *Anticancer Research*, 2012, **32**(9): 3699–3706
- [28] Ushida H, Kawakami T, Minami K, *et al.* Methylation profile of DNA repetitive elements in human testicular germ cell tumor. *Molecular Carcinogenesis*, 2012, **51**(9): 711–722
- [29] Yu F, Zingler N, Schumann G, *et al.* Methyl-CpG-binding protein 2 represses LINE-1 expression and retrotransposition but not Alu transcription. *Nucl Acid Res*, 2001, **29**(21): 4493–4501
- [30] Muotri A R, Marchetto M C, Coufal N G, *et al.* L1 retrotransposition in neurons is modulated by MeCP2. *Nature*, 2010, **468**(7322): 443–446
- [31] Martens J H A, O'Sullivan R J, Braunschweig U, *et al.* The profile of repeat-associated histone lysine methylation states in the mouse epigenome. *Embo J*, 2005, **24**(4): 800–812
- [32] Maze I, Feng J, Wilkinson M B, *et al.* Cocaine dynamically regulates heterochromatin and repetitive element unsilencing in nucleus accumbens. *Proc Natl Acad Sci USA*, 2011, **108**(7): 3035–3040
- [33] Muotri A R, Chu V T, Marchetto M C, *et al.* Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. *Nature*, 2005, **435**(7044): 903–910
- [34] Montoya-Durango D E, Liu Y, Teneng I, *et al.* Epigenetic control of mammalian LINE-1 retrotransposon by retinoblastoma proteins. *Mutat Res*, 2009, **665**(1–2): 20–28
- [35] Garcia-Perez J L, Morell M, Scheys J O, *et al.* Epigenetic silencing of engineered L1 retrotransposition events in human embryonic

- carcinoma cells. *Nature*, 2010, **466**(7307): 769–773
- [36] Ghildiyal M, Zamore P D. Small silencing RNAs: an expanding universe. *Nat Rev Genet*, 2009, **10**(2): 94–108
- [37] Yang N, Kazazian H H. L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells. *Nat Struct Mol Biol*, 2006, **13**(9): 763–771
- [38] Soifer H S, Zaragoza A, Peyvan M, *et al.* A potential role for RNA interference in controlling the activity of the human LINE-1 retrotransposon. *Nucl Acid Res*, 2005, **33**(3): 846–856
- [39] Aravin A A, Sachidanandam R, Girard A, *et al.* Developmentally regulated piRNA clusters implicate MILI in transposon control. *Science*, 2007, **316**(5825): 744–747
- [40] Carmell M A, Girard A, van de Kant H J, *et al.* MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. *Developmental Cell*, 2007, **12**(4): 503–514
- [41] Aravin A A, Sachidanandam R, Bourc'his D, *et al.* A piRNA pathway primed by individual transposons is linked to *de novo* DNA methylation in mice. *Molecular Cell*, 2008, **31**(6): 785–799
- [42] Kuramochi-Miyagawa S, Watanabe T, Gotoh K, *et al.* DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. *Genes & Development*, 2008, **22**(7): 908–917
- [43] De Fazio S, Bartonicek N, Di Giacomo M, *et al.* The endonuclease activity of Mili fuels piRNA amplification that silences LINE1 elements. *Nature*, 2011, **480**(7376): 259–263
- [44] Reuter M, Berninger P, Chuma S, *et al.* Miwi catalysis is required for piRNA amplification-independent LINE1 transposon silencing. *Nature*, 2011, **480**(7376): 264–267
- [45] Chiu Y L, Greene W C. The APOBEC3 cytidine deaminases: an innate defensive network opposing exogenous retroviruses and endogenous retroelements. *Annu Rev Immunol*, 2008, **26**: 2631–26353
- [46] Koito A, Ikeda T. Intrinsic restriction activity by AID/APOBEC family of enzymes against the mobility of retroelements. *Mobile Genetic Elements*, 2011, **1**(3): 197–202
- [47] Wissing S, Montano M, Garcia-Perez J L, *et al.* Endogenous APOBEC3B restricts LINE-1 retrotransposition in transformed cells and human embryonic stem cells. *J Biol Chem*, 2011, **286**(42): 36427–36437
- [48] Ikeda T, Abd El Galil K H, Tokunaga K, *et al.* Intrinsic restriction activity by apolipoprotein B mRNA editing enzyme APOBEC1 against the mobility of autonomous retrotransposons. *Nucl Acid Res*, 2011, **39**(13): 5538–5554
- [49] Lorgeoux R P, Guo F, Liang C. From promoting to inhibiting: diverse roles of helicases in HIV-1 Replication. *Retrovirology*, 2012, **9**: 79
- [50] Arjan-Odedra S, Swanson C M, Sherer N M, *et al.* Endogenous MOV10 inhibits the retrotransposition of endogenous retroelements but not the replication of exogenous retroviruses. *Retrovirology*, 2012, **9**: 53
- [51] Goodier J L, Cheung L E, Kazazian H H, Jr. MOV10 RNA helicase is a potent inhibitor of retrotransposition in cells. *PLoS Genetics*, 2012, **8**(10): e1002941
- [52] Lu C, Luo Z, Jager S, *et al.* Moloney leukemia virus type 10 inhibits reverse transcription and retrotransposition of intracisternal particles. *J Virology*, 2012, **86**(19): 10517–10523
- [53] Stetson D B, Ko J S, Heidmann T, *et al.* Trex1 prevents cell-intrinsic initiation of autoimmunity. *Cell*, 2008, **134**(4): 587–598
- [54] Coufal N G, Garcia-Perez J L, Peng G E, *et al.* Ataxia telangiectasia mutated (ATM) modulates long interspersed element-1 (L1) retrotransposition in human neural stem cells. *Proc Natl Acad Sci USA*, 2011, **108**(51): 20382–20387
- [55] Gasior S L, Roy-Engel A M, Deininger P L. ERCC1/XPF limits L1 retrotransposition. *DNA Repair (Amst)*, 2008, **7**(6): 983–989
- [56] Stremlau M, Owens C M, Perron M J, *et al.* The cytoplasmic body component TRIM5alpha restricts HIV-1 infection in old world monkeys. *Nature*, 2004, **427**(6977): 848–853
- [57] Lahouassa H, Daddacha W, Hofmann H, *et al.* SAMHD1 restricts the replication of human immunodeficiency virus type 1 by depleting the intracellular pool of deoxynucleoside triphosphates. *Nat Immunol*, 2012, **13**(3): 223–228
- [58] Neil S J, Zang T, Bieniasz P D. Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature*, 2008, **451**(7177): 425–430
- [59] Kauder S E, Bosque A, Lindqvist A, *et al.* Epigenetic regulation of HIV-1 latency by cytosine methylation. *PLoS Pathog*, 2009, **5**(6): e1000495
- [60] du Chene I, Basyuk E, Lin Y L, *et al.* Suv39H1 and HP1gamma are responsible for chromatin-mediated HIV-1 transcriptional silencing and post-integration latency. *EMBO J*, 2007, **26**(2): 424–435
- [61] Coiras M, Lopez-Huertas M R, Perez-Olmeda M, *et al.* Understanding HIV-1 latency provides clues for the eradication of long-term reservoirs. *Nat Rev Microbiol*, 2009, **7**(11): 798–812
- [62] Huang J, Wang F, Argyris E, *et al.* Cellular microRNAs contribute to HIV-1 latency in resting primary CD4+ T lymphocytes. *Nat Med*, 2007, **13**(10): 1241–1247
- [63] Yan N, Regalado-Magdos A D, Stiggelbout B, *et al.* The cytosolic exonuclease TREX1 inhibits the innate immune response to human immunodeficiency virus type 1. *Nat Immunol*, 2010, **11**(11): 1005–1013
- [64] Kazazian H H, Jr., Wong C, Youssoufian H, *et al.* Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. *Nature*, 1988, **332**(6160): 164–166
- [65] He X J, Chen T, Zhu J K. Regulation and function of DNA methylation in plants and animals. *Cell Research*, 2011, **21** (3): 442–465

The Mechanisms of Host Restriction on LINE-1 Element*

LIANG Zhi-Bin¹⁾, LIANG Chen^{2, 3)}, GENG Yun-Qi¹⁾, QIAO Wen-Tao^{1)**}

¹⁾ College of Life Sciences, Nankai University, Tianjin 300071, China;

²⁾ Department of Medicine, McGill University, Montreal, Quebec, Canada H3A 2B4;

³⁾ Department of Microbiology and Immunology, McGill University, Montreal, Quebec, Canada H3A 2B4)

Abstract Approximately 45% of the human genome is occupied by transposable elements. Transposable elements have had an important impact on the structure and function of human genomes. Some of these elements are still capable of transposing, and their movements often cause diseases. Long interspersed nuclear element 1 (LINE-1) is the only active autonomous transposon in humans, and it mediates the mobilization of nonautonomous elements. Recently, great progress has been made in understanding the biology of LINE-1. Here, we briefly review the structure of LINE-1, the mechanism of its mobilization, and its impact on human genome and human health. We also highlight the mechanisms that host cell has evolved to control LINE-1 replication. Given the similarity of the life cycle between LINE-1 and retroviruses, the LINE-1 research may also shed new light on the biology of retroviruses.

Key words LINE-1, transposition, genome, restriction

DOI: 10.3724/SP.J.1206.2012.00506

* This work was supported by grants from Chinese Ministry of Health (2012ZX10001006), The National Natural Science Foundation of China (81271812), and Tianjin Research Program of Application Foundation and Advanced Technology (12JCQNJC06100).

**Corresponding author.

Tel: 86-22-23504547, E-mail: wentaoqiao@nankai.edu.cn

Received: November 28, 2012 Accepted: January 16, 2013