

非编码RNA编辑技术研究进展

黄蕾 王雅洁 陈实*

(武汉大学药学院, 组合生物合成与新药发现教育部重点实验室, 武汉 430071)

摘要 非编码RNA不具有蛋白编码功能而直接以RNA的形式在机体内发挥作用, 研究发现, 它们与多种调控过程和疾病有关, 它们由发现之初被称为“垃圾”成为了研究的焦点。为了研究非编码RNA的功能, 构建其功能缺失的实验模型不可或缺, 而基因编辑技术从最初的基因打靶技术到后来的锌指核酸酶(zinc-finger nucleases, ZFN)、转录激活子样效应因子核酸酶(transcription activator-like effector nucleases, TALEN)以及CRISPR/Cas9(clustered regularly interspaced short palindromic repeat sequences/Cas9, CRISPR/Cas9)技术在应用到非编码RNA的敲除方面都与编码蛋白的基因敲除不同。该文主要对非编码RNA中的miRNA和lncRNA的研究现状以及非编码RNA的敲除策略进行探讨。

关键词 非编码RNA; 基因打靶技术; 锌指核酸酶; 转录激活子样效应因子核酸酶; CRISPR/Cas9

Progress in ncRNAs Editing

Huang Lei, Wang Yajie, Chen Shi*

(Key Laboratory of Combinatorial Biosynthesis and Drug Discovery, Ministry of Education, School of Pharmaceutical Sciences, Wuhan University, Wuhan 430071, China)

Abstract Noncoding RNAs (ncRNAs) have no protein-coding capability and they function directly in the form of RNA. Genes coding ncRNAs were called Junk DNA when they were discovered. However they are found to be related to increasingly number of diseases and become the focus of many researches. Loss-of-function models are indispensable for function digging, and there are different kinds of gene editing techniques such as gene targeting, zinc-finger nucleases (ZFN), transcription activator-like effector nucleases (TALEN) and clustered regularly interspaced short palindromic repeat sequences/Cas9 (CRISPR/Cas9). However, the editing techniques used for ncRNAs are quite different from that of protein-coding genes. ncRNAs have no open reading frames so that small deletion may not disrupt the functions. In this review, research progress in miRNA and lncRNA are summarized as well as the techniques used to delete ncRNAs.

Keywords noncoding RNAs; gene targeting; zinc-finger nucleases (ZFN); transcription activator-like effector nucleases (TALEN); clustered regularly interspaced short palindromic repeat sequences/Cas9 (CRISPR/Cas9)

1 非编码RNA

非编码RNA(noncoding RNAs, ncRNAs)是指不具有蛋白编码功能的RNA。早在发现之初, 这些非编码基因被称为“垃圾DNA”, 但是越来越多的研究

发现, 这些“垃圾DNA”是具有生物学功能的, 如不均一核RNA(heterogeneous nuclear RNA, hnRNA)、核内小RNA(small nuclear RNA, snRNA)、小核仁RNA(small nucleolar RNA, snoRNA)、miRNA和长

收稿日期: 2015-03-14 接受日期: 2015-04-15

青年千人计划(陈实)和湖北省医学领军人才培养工程(陈实)资助的课题

*通讯作者: Tel: 027-68756643, E-mail: shichen@whu.edu.cn

Received: March 14, 2015 Accepted: April 15, 2015

This work was supported by the Recruitment Program of Global Youth Experts (Chen Shi) and Hubei Province's Outstanding Medical Academic Leader Program (Chen Shi)

*Corresponding author: Tel: +86-27-68756643, E-mail: shichen@whu.edu.cn

网络出版时间: 2015-07-01 17:01

URL: <http://www.cnki.net/kcms/detail/31.2035.Q.20150701.1701.005.html>

非编码RNA(long noncoding RNA, lncRNA)等^[1]。另外, 转录组分析显示, 哺乳动物中的70%~90%的基因都被转录, 但是其中只有1%~2%能够编码蛋白质, ncRNA的量远远大于mRNA的量^[2-3], 这些占转录组中大多数的ncRNA参与了包括细胞分化、凋亡、免疫应答等在内的几乎所有生物体中所有的生理、病理过程的调控。本文主要介绍miRNA和lncRNA。

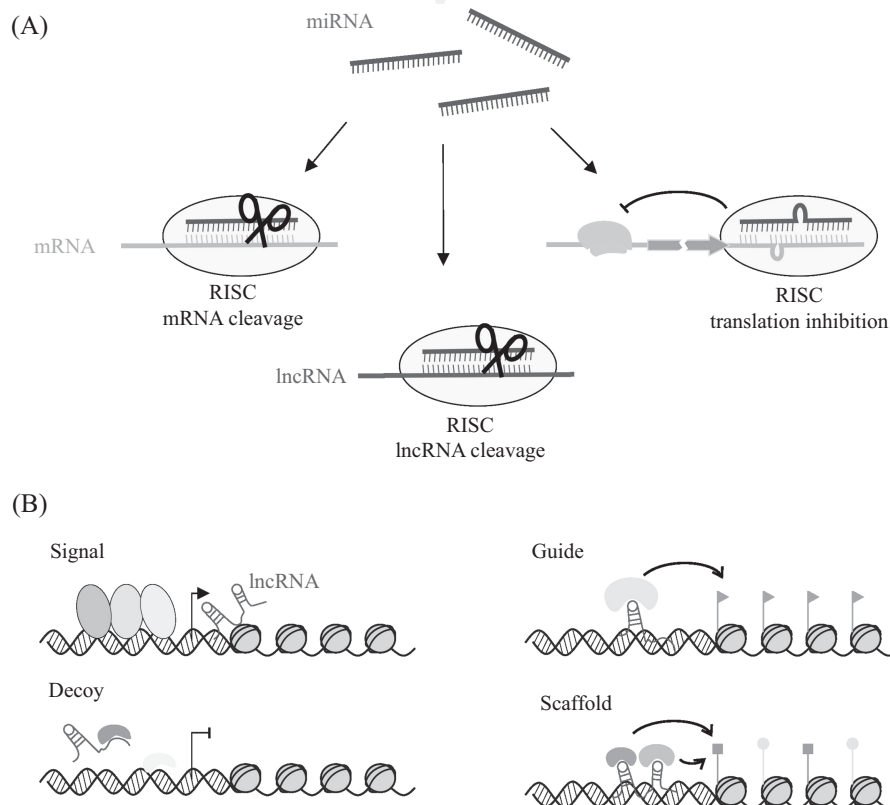
1.1 miRNA

miRNA是一类长度约为22 nt的ncRNA。成熟的miRNA的形成需要核糖核酸酶Drosha和Dicer的剪切加工。首先, 细胞核中转录形成pri-miRNA, 随后pri-miRNA由核糖核酸酶Drosha加工成pre-miRNA并被输出到细胞质, 之后pre-miRNA在核糖核酸酶Dicer的剪切作用下形成成熟的miRNA^[4]。

成熟的miRNA作用方式主要是通过与其的靶点进行碱基互补配对来对靶点基因进行转录和转录后调控。miRNA能结合到mRNA 3'UTR区域, 当它与

靶点序列完全互补时, RNA诱导沉默复合物(RNA-induced silencing complex, RISC)则会促使mRNA降解; 当序列与靶点序列不完全互补时, RISC则能抑制mRNA的翻译^[5-6]。另外, miRNA也能结合到lncRNA上来降低其表达量^[7](图1A)。哺乳动物中超过一半的编码蛋白质的mRNA都会受到这个调控^[8]。1993年, 有两篇文献^[9-10]同时介绍了miRNA在线虫的发育周期中所起的调控作用。

miRNA与很多疾病有关。*miR-15a-16-1*、*miR-155*、*miR-650*、*miR-34a*被发现与慢性淋巴细胞性白血病有关^[11]。一些miRNA被作为癌症预后以及癌症转移的潜在诊断标志物, 如在结肠直肠癌的癌变组织中*miR-20a*和*miR-31*表达量显著上升, *miR-143*和*miR-145*表达量明显降低^[12]。在神经退行性疾病方面, *miR-29a/b-1*和*miR-107*能够调控与阿尔兹海默症相关的A β 淀粉样沉淀物形成有关的 β -分泌酶(beta-secretase 1, BACE1)的表达^[13-14]; 在帕金森病、亨廷顿舞蹈症等疾病中, miRNA也有相关的调控作用^[15]。



A: miRNA作用机理。miRNA能募集AGO(Argonaute)蛋白等形成RNA诱导沉默复合物(RNA-induced silencing complex, RISC)来对靶序列进行加工; B: lncRNA作用机理。lncRNA能作为信号分子、诱饵、向导或支架来调控基因的表达。

A: mechanisms of miRNAs. miRNAs recruit proteins including Argonaute (AGO) to form RNA-induced silencing complex which process the targeted RNAs; B: mechanisms of lncRNAs. lncRNAs can function as signal, decoy, guide or scaffold.

图1 miRNA和lncRNA的作用机理
Fig.1 Mechanisms of miRNAs and lncRNAs

1.2 lncRNA

lncRNA是指长度大于200 nt的ncRNA,这类ncRNA保守性较差,多数由RNA聚合酶II转录,有多个外显子,大多也会有与mRNA相似的5'端帽子结构以及3'端的多聚腺苷酸化结构^[2,16]。lncRNA在细胞中的作用机理有以下几类:一是它可以作为一些转录因子或者信号通路中的信号分子,因为它不需要再翻译成蛋白质,能更快地起到调节作用;二是lncRNA能作为将某些蛋白因子或miRNA从染色体位点拉走的诱饵,从而降低该蛋白或miRNA的水平;三是lncRNA能作为向导,将蛋白招募到其作用靶点来发挥功能;最后,它能作为支架与多种蛋白质功能构成核糖核蛋白复合物,使多种不同的蛋白质协同发挥作用^[16](图1B)。

lncRNA与癌症、神经退行性疾病等相关的多个调控通路如剂量补偿^[17]、发育^[18]、基因印记^[19]及免疫应答^[20]等过程相关。X染色体失活(X inactivation, XCI)是在雌性哺乳动物中两条X染色体之一丧失性状表达的现象,这能平衡雌性和雄性哺乳动物之间的基因表达。而在这个过程中,一个被命名为*Xist*(X inactive specific transcript)的lncRNA起着非常重要的作用:在雌性哺乳动物的发育过程中,*Xist* RNA在失活的X染色体位置转录,随即覆盖在它转录位置的染色体上,从而导致染色体上该处的基因表达受到抑制^[21]。一些lncRNA能在表观遗传修饰如组蛋白修饰、DNA甲基化等过程中发挥作用,如在HOXC(homeobox C)位点转录产生的*HOTAIR* RNA能分别通过它的5'端和3'端结合PRC2(polycomb repressive complex 2)复合物和LSD1/CoREST/REST(lysine-specific demethylase 1A/RE1-silencing transcription factor corepressor/RE1-silencing transcription factor)复合物,并将这些蛋白募集到HOXD(homeobox D)位点,使该位点处的组蛋白H3第27位赖氨酸甲基化和第4位的赖氨酸去甲基化,从而调控基因的表达^[22]。

许多研究表明,在肝癌、肺癌、胰腺癌、乳腺癌、结直肠癌等癌症中都发现了一些相关的lncRNA。*H19*能与血管生成素和成纤维细胞生长因子相互作用而在肝脏中诱发肿瘤^[23],并且过表达*H19*还能增加胃癌的转移^[24]。与肝癌相关的lncRNA还有*HULC*、*TUC338*、*UCA1*、*MEG3*、*HOTTIP*和*MALAT1*等^[25]。*MALAT1*在肺和胰腺中高表达,与

癌症的复发和转移有关^[26]。lncRNA一般不只有一个通路里起作用,往往参与多种疾病,如*HOTAIR*与乳腺癌、结直肠癌、胃肠道间质瘤及原发性肝细胞瘤等疾病都有关系^[27],*NEATI* RNA在亨廷顿舞蹈症和肌萎缩侧索硬化症病人中表达量都显著升高^[28],并且它还是paraspeckle的支架^[29],在细胞面对病毒或其他压力刺激时,*NEATI* RNA的表达量以及paraspeckle的数量和大小都会增加。有许多lncRNA都参与了神经退行性疾病中的一些过程^[28,30]。

lncRNA与miRNA之间也有相互作用的关系。miRNA能结合到lncRNA上来诱导lncRNA的降解^[31],而lncRNA也能通过与miRNA结合或与mRNA竞争结合miRNA来拮抗miRNA的功能^[32-33],某些lncRNA还是miRNA形成的前体^[34]。

虽然已有许多研究发现,ncRNA与许多疾病或其他生物学过程有关,但是我们对于它们在其中所承担的具体功能的认识还是远远不够的。

2 非编码RNA编辑技术

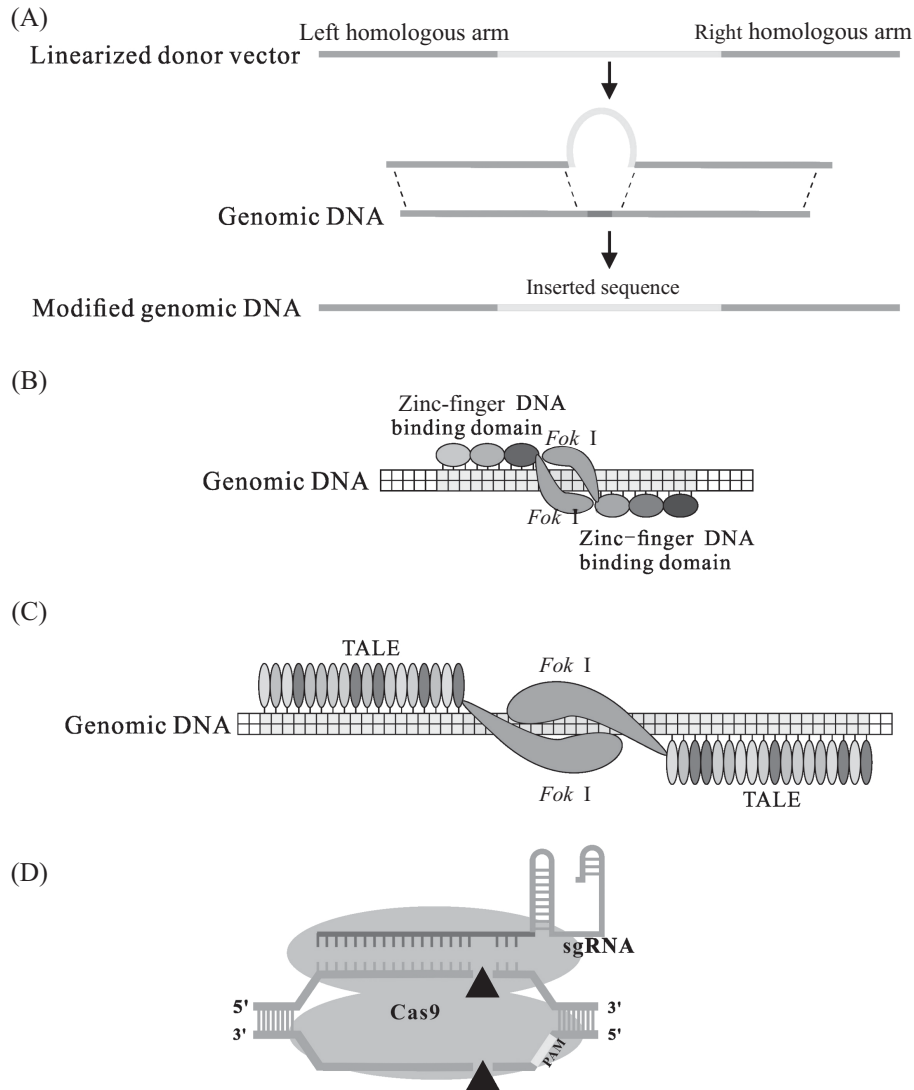
构建基因功能缺失的细胞或动物模型是研究基因生物学功能的不可或缺的手段。在基因功能缺失方面应用最为广泛的是RNA干扰(RNA interference, RNAi)技术,这项技术利用小干扰RNA(small interfering RNA, siRNA)来对某个目的基因进行敲减^[35],但是由于siRNA特异性和沉默效率方面的问题使这项技术有很大的局限性^[36]。另外,在基因组编辑方面使用的技术还有基因打靶技术、锌指核酸酶(zinc-finger nucleases, ZFN)、转录激活因子样效应因子核酸酶(transcription activator-like effector nuclease, TALEN)以及CRISPR/Cas9(clustered regularly interspaced short palindromic repeats/CRISPR-associated protein)。其中,由于基因打靶技术的操作过程复杂并且效率很低^[37],使用的局限性很大。而ZFN、TALEN和CRISPR/Cas9这三种技术都是结合了DNA识别结构和能够使DNA双链发生断裂的蛋白的技术,利用能够识别基因组上的特定序列的模块来引导核酸酶切割基因组DNA,使其发生双链DNA的断裂,诱导同源重组(homologous recombination, HR)和非同源末端连接(non-homologous end joining, NHEJ)来修复断裂的基因组DNA。在将这三种技术用于敲除编码蛋白的基因时,由于非同源末端连接机制在使基因组发生修

复时容易引起DNA的插入或缺失突变, 造成移码突变, 从而导致其编码的氨基酸发生改变, 最终引起相应蛋白的功能消失或异常。然而在面对敲除ncRNA的情况时, 这一方法却行不通, 因为ncRNA不具有开放阅读框(open reading frame, ORF), 不存在移码突变的情况, 基因组上的小缺失或插入可能并不影

响其行使正常功能, 而且由于一些ncRNA特别是lncRNA的功能域无法预测, 这些都使得构建ncRNA功能缺失的突变株成为了一个难题。本文主要介绍除RNAi等敲减技术以外的基因敲除技术。

2.1 基因打靶技术

基因打靶技术是指利用减数分裂时期的同源



A: 基因打靶技术。基因打靶技术需要一个供体质粒来诱导细胞利用同源重组将外源DNA整合到目的基因组上, 其中使用的供体质粒上应包含有两类序列, 一是一部分靶基因序列作为左右同源臂, 二是左右同源臂之间的想要整合到基因组上的序列; B: 锌指核酸酶(zinc-finger nucleases, ZFN)作用机制。两个锌指DNA结合域分别识别靶基因的正义链和反义链, 核酸内切酶Fok I二聚产生DNA双链切割活性, 从而使靶基因DNA发生双链断裂; C: 转录激活子样效应因子核酸酶(transcription activator-like effector nuclease, TALEN)作用机制。TALEN与ZFN作用机制相似, 都融合了DNA结合域和核酸内切酶Fok I, TALEN的DNA结合域是转录激活子样效应因子(transcription activator-like effector, TALE); D: CRISPR/Cas9技术作用机制。CRISPR/Cas9系统由单链向导序列(single guide RNA, sgRNA)和Cas9蛋白两部分构成。sgRNA利用碱基互补配对来识别靶基因序列, 并募集Cas9蛋白对靶基因进行切割。

A: gene targeting. A donor vector which contains homologous sequences and sequence for insertion is needed to induce a foreign sequence recombined into genome of cells; B: mechanism of ZFN. The two zinc-finger DNA binding domains recognize opposite strands respectively which allows endonuclease Fok I to dimerize and cleave target DNA; C: mechanism of TALEN. TALEN is much similar to ZFN which was fused DNA binding domain and endonuclease Fok I. But the binding domain of TALEN is TALE; D: mechanism of CRISPR/Cas9. This system consists of sgRNA and Cas9 protein. SgRNA recognizes the target sequence based on base pairing and recruit Cas9 protein to cut DNA.

图2 基因编辑技术

Fig.2 Gene editing

重组来将导入的外源供体DNA上的目的序列整合到基因组上,从而达到基因组编辑目的的技术(图2A),这项技术在体外培养胚胎干细胞技术发展起来以后得到广泛应用,它使利用同源重组得到基因完全敲除的动物模型成为可能^[38],但是效率极低,在 $10^4\sim 10^7$ 个细胞当中才有一个阳性的突变细胞^[39]。

BCI RNA是在神经元树突处调控蛋白翻译的ncRNA, Boris等^[40]构建了一个载体,这个载体上带有*BCI*基因两侧的序列作为同源臂,左右两侧同源臂之间插入一段新霉素抗性基因序列,他们利用这个载体在减数分裂期同源重组时诱导外源序列整合到细胞基因组上,从而得到了新霉素抗性基因完全替换了*BCI*基因的小鼠。许多课题组都用类似的方法拿到了基因敲除的小鼠或其他动物模型^[41]。

由于基因打靶技术效率低、周期长、花费大,应用起来局限很大,但是将基因打靶技术与能使基因组双链发生断裂的技术结合起来,能极大地提高同源重组的概率^[42-43],这部分将在下文中的ZFN、TALEN和CRISPR/Cas9部分介绍。

2.2 ZFN

ZFN技术是融合了锌指DNA结合域和核酸内切酶*Fok I*的能使基因组双链DNA发生断裂(double-strand break, DSB)的技术。ZFN的一个锌指结构识别三个碱基,将三个锌指结构串联起来就构成一个锌指DNA结合域,通过设计合适位点的两个锌指DNA结合域使非特异的*Fok I*核酸内切酶发生二聚,从而激活*Fok I*切割DNA双链的活性,最终导致靶向序列发生DNA双链的断裂(图2B)。随后,细胞通过同源重组或非同源末端连接的方式对基因组进行修复^[44-45]。ZFN自发现之后,在果蝇、斑马鱼、小鼠、大鼠等许多不同的物种中都得到了广泛的应用^[39]。为了实现ncRNA的敲除, Tony等^[46]通过结合ZFN与同源重组的方法成功地在A549细胞中对一个约8 000 nt的lncRNA实现了敲除。他们使用了敲除位点在*MALAT1*这个lncRNA的转录起始位点之前的ZFN,同时也构建了一个供体质粒,这个供体质粒上带有ZFN敲除位点两侧的序列作为左右同源臂,左右同源臂之间是一段CMV-GFP-RDE序列。将ZFN与这个供体质粒共转到细胞,带有靶基因同源臂序列的外源供体质粒能在ZFN诱导的双链断裂处促使细胞以同源重组的方式进行修复,由此实现在ZFN靶点处插入外源序列(CMV-GFP-RDE)(图3A)。

他们使用的RDE(RNA destabilizing element)序列是多腺苷酸化信号序列(polyadenylation signals, polyA signals),能使其插入位点下游的转录本发生降解,从而达到沉默这个基因的目的。

有课题组在HEK293T和HEK293细胞中分别用两对ZFN实现了174 bp的小片段和15 Kb的大片段缺失^[47-48],但他们敲除的并不是ncRNA。之后这一思路被用在了TALEN和Cas9技术上。

2.3 TALEN

TALEN与ZFN的作用原理类似,都是利用蛋白质-DNA相互作用来识别特定的DNA序列,再由*Fok I*核酸内切酶使基因组DNA双链发生断裂,从而达到基因组编辑的目的(图2C)。TALEN的靶点特异性依赖于转录激活子样效应因子(transcription activator-like effector, TALE)上的能够结合DNA的串联重复区域,这个串联重复区域由17~18个重复单元构成,而每个重复单元由34个氨基酸构成,位于第12和13位的两个氨基酸决定这个重复单元所能特异识别的碱基,这两个氨基酸称为重复可变双残基(repeat variable di-residue, RVD), RVD决定了每个重复单元所识别的碱基,如HD识别C、NG识别T、NI识别A、NN识别A和G^[49-50]。在TALE的两端分别加上核定位信号和*Fok I*核酸内切酶即构成TALEN。TALEN与ZFN一样,也需要一对TALE来共同发挥作用,但是与ZFN相比,TALEN的每个重复单元能识别单个碱基,所以其设计更加灵活,构建方法也更为简便^[51]。

敲除单个miRNA的方法相对简单,因为miRNA序列较短,只需要删除几个碱基,就能使pre-miRNA的发卡结构发生改变,从而导致miRNA无法正常合成。Liu等^[52]就利用这个原理在斑马鱼中显微注射TALEN mRNA,成功地敲除了*miR-1-1*和*miR-1-2*。

对于miRNA基因簇或者lncRNA,由一对TALEN的编辑所造成的几个碱基的变化无法达到使基因功能发生缺失的目的。Liu等^[52]尝试通过两对TALEN来完成较大片段的缺失(图3B)。他们针对一个基因簇或一个基因同时设计了两对TALEN,这两对TALEN在基因组上的靶序列之间的序列则是想要缺失的序列,将这两对TALEN的mRNA同时导入到单细胞时期的斑马鱼胚胎中,将胚胎培养两天之后对其基因组DNA进行PCR检测并测序,证实两对TALEN位点之间的基因组片段发生了缺失,即在基因组上实现较大片段的删除。由此,他们成

功地在斑马鱼中缺失了*miR-17-92*(1.2 Kb)和*miR-430*(79.8 Kb)这两个miRNA基因簇以及一个lncRNA *MALAT1*(9 Kb)。Xiao等^[53]用同样的方法在斑马鱼中缺失了*sema3fb*这个蛋白基因(43 Kb)、*miR-126a*和*miR-126b*这两个miRNA、包含*miR-17a-1*基因簇到*miR-92a-1*基因簇序列在内的约1.6 Kb的区域以及包含*miR-17a-2*基因簇到*miR-92a-2*基因簇序列在内的约1.5 Kb的区域。

另外, Xiao等^[53]的实验发现, 在斑马鱼中同时显微注射两对TALEN用于敲除时, 这两对TALEN靶点之间的DNA序列可能会出现两种突变形式, 一种是这段序列被删除, 另一种是这段序列可能会发生翻转。Ankit等^[54]也利用一对TALEN在斑马鱼中成功缺失了3.2 Kb的*linc-birc6*。当两对TALEN在基因组上同时切割造成两个DNA双链断裂位点时, 基因组上的两个双链断裂位点可能会由于非同源末端连接修复而被连接到一起, 造成断裂位点中间的大片段从基因组上丢失, 另外, 这个大片段也有可能依然连接在断裂位点处, 但是序列发生翻转、片段发生缺失的概率要高于片段翻转的概率^[53-54]。所以这种产生两个双链断裂位点的方法也能用于ncRNA的敲除。

2.4 CRISPR/Cas9

CRISPR是指成簇的、规律间隔的短回文重复序列(clustered regularly interspaced short palindromic repeats), 它是在细菌和古细菌中发现的由21~37个核苷酸组成的重复序列和长度相似的间隔序列交错间隔排列而成的一段序列, 形成“重复序列-间隔序列-重复序列”的模式。后来人们发现, 这段序列中的间隔序列与病毒中的部分序列相同, 当面对病毒入侵时, 细菌和古细菌中的这段CRISPR序列就能够引导CRISPR相关(CRISPR-associated, Cas)蛋白作用使外源DNA降解, 从而起到免疫保护的作用^[55-59], 这一系统被称为CRISPR/Cas系统。在CRISPR/Cas系统介导的免疫过程中, 细菌和古细菌首先将外源入侵的病毒上的原型间隔序列(protospacer)整合到自己CRISPR序列的一端, 然后宿主上的CRISPR序列转录形成CRISPR RNA前体(pre-cursor CRISPR RNA, pre-crRNA), 随后通过相关酶的剪切成短的crRNA, 由于crRNA与原型间隔序列互补, 随之招募而来的Cas蛋白复合物则会将入侵的核酸序列降解^[60]。

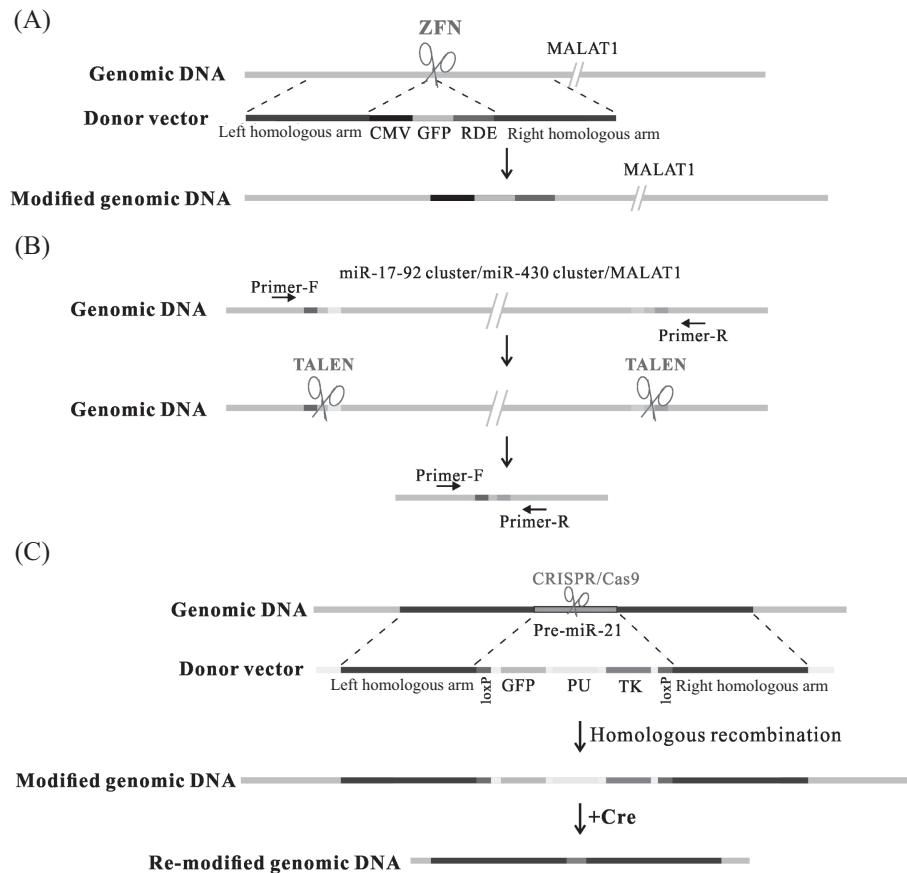
CRISPR/Cas系统分为三种类型^[61], I型和III型

系统需要特定的Cas核酸内切酶来加工pre-crRNA形成成熟的能与入侵核酸序列互补的crRNA, crRNA再与多种Cas蛋白组成复合物, 而这种复合物既能识别靶点又能实现对靶点进行切割; II型系统只需要Cas9蛋白复合物即可发挥作用, 目前广泛应用的即为II型系统。Cas9蛋白复合物由反式激活crRNA(trans-activating crRNA, tracrRNA)、crRNA以及Cas9蛋白共同组成, Cas9蛋白作用依赖于tracrRNA与crRNA形成的碱基互补配对结构, 其切割位点的特异性由crRNA与原型间隔序列形成的互补配对结构以及原型间隔序列邻近基序(protospacer adjacent motif, PAM) NGG共同决定^[60-63], 并且tracrRNA与含有靶点序列的crRNA被改造成一条单链向导序列(single guide RNA, sgRNA), 其中sgRNA上的靶点序列长度约20 bp, 这使得Cas9在体外的应用更为简便(图2D)^[60], 我们只需要合成约20 bp的靶点序列并将它连接到已改造好的sgRNA表达载体上即可。这项应用一开始只用在原核生物中^[60,62-65], 但经过改造, II型CRISPR/Cas系统也被成功地应用到了哺乳细胞中^[66-67]。

Wang等^[68]同时向小鼠胚胎干细胞中注射了Cas9蛋白的mRNA以及五个分别靶向五个不同基因的sgRNA, 最终他们得到了有五个基因同时发生突变的小鼠。这让构建多个基因同时突变的模型成为可能。在水稻中, Zhou等^[69]构建了一个可同时连接多个sgRNA进行表达的载体对多个大片段缺失的敲除, 最大到260 Kb。另外, Patrick等^[70]在一个人类近单倍体细胞系HAP1中实现了一段长达30 Mb的序列缺失。

在敲除ncRNA方面, Han等^[71]设计了两个靶点在*Rian*上且相距约23 Kb的sgRNA, 并且将这两个sgRNA和Cas9蛋白在体外转录得到的mRNA显微注射到单细胞阶段的小鼠胚胎中, 由此成功得到了在基因组上相应基因位置缺失23 Kb的小鼠。为了提高效率, 他们还同时显微注射了两对sgRNA以及Cas9蛋白的mRNA——将小鼠基因组中发生缺失的概率从16%提高到了33.3%, 但同时也增大了脱靶(off-target)的风险。Xiao等^[53]在斑马鱼中注射Cas9蛋白mRNA和靶点分别位于*dre-mir-126a*两侧的sgRNA或者Cas9蛋白mRNA和一个包括有六个miRNA的基因簇两侧的两个sgRNA, 也能将miRNA或miRNA基因簇敲除, 但是与TALEN相比效率更低。

Ho等^[72]在人类细胞系HEK293、HEK293T、



A: 利用ZFN在基因组上插入RNA不稳定元件(RNA destabilizing element, RDE)来敲除*MALAT1*。绿色荧光蛋白(green fluorescent protein, GFP)序列是为了方便筛选出成功插入外源序列的突变细胞株, RDE序列能够使其插入位点下游的转录产物发生降解, 从而达到沉默*MALAT1* RNA的目的; B: 利用双位点TALEN敲除基因组上的大片段DNA。图中基因组序列上的四条短线指两对TALEN的四个TALE在基因组上的结合位点。利用PCR(引物为Primer-F和Primer-R)检测基因组上的缺失情况, 大片段缺失的基因组能扩增出较小片段, 而没有发生片段缺失的基因组则会因为片段太大而无法有效扩增; C: 利用CRISPR/Cas9和供体质粒敲除miR-21。PU: 嘌呤霉素抗性基因序列; TK: 指胸苷激酶基因。Cre重组酶能够删除loxP位点之间的序列。

A: delete *MALAT1* through integration of RDE using ZFN. GFP is used to screen mutant cells in which effective integration happened. RDE here is polyadenylation signals which silence downstream sequences; B: delete large DNA fragment with dual TALENs. The binding sites of TALENs are shown in short lines. The effective deletion gives rise to a shorter PCR product with Primer-F and Primer-R; C: delete miR-21 through homologous recombination induced by CRISPR/Cas9 and a donor vector. PU: puromycin resistance gene; TK: thymidine kinase gene. loxP: recognition site of Cre. Cre: a kind of recombinase which deletes the sequence between loxP site.

图3 用ZFN、TALEN以及CRISPR/Cas9技术敲除ncRNA

Fig.3 Knock out ncRNAs with the methods of ZFN, TALEN and CRISPR/Cas9

HCT-116、MCF-7和LNCaP细胞中验证了利用Cas9和供体质粒敲除ncRNA的可行性。他们构建了一个能同时表达sgRNA和Cas9蛋白的新敲除载体以及在左右同源臂序列之间为loxP-GFP-PU-TK-loxP序列的供体质粒, 将这两个质粒在细胞中共转后利用绿色荧光蛋白(green fluorescent protein, GFP)和嘌呤霉素抗性(puromycin resistance gene, PU)这两个筛选标记对细胞进行筛选, 成功得到了基因组上pre-miR-21基因被替换为loxP-GFP-PU-TK-loxP序列的突变细胞株。另外, 在成功插入了外源DNA序列的突变细胞株中表达Cre重组酶还能将插入的两段loxP序

列之间的序列从基因组上删除。利用GFP和嘌呤霉素大大增加了筛选到基因组上插入外源序列的突变细胞株的概率, 而胸苷激酶基因(thymidine kinase gene, TK)序列则是用于筛选在表达Cre重组酶之后成功删除掉重组序列的突变细胞株(图3C)。他们通过这种方法完成了miR-21的敲除。对于3个lncRNA *UCA1*、*lncRNA-21A*和*AK023948*, 他们采用了双sgRNA与供体质粒共转的方法来实现了敲除。

3 展望

越来越多的研究发现, miRNA和lncRNA参与了

多种结构、代谢和调控等功能, 从而在癌症和神经退行性疾病等疾病中都起到了促进或抑制的作用。然而我们对于它们的具体作用机制的研究还处在初步阶段, 真正了解具体功能的ncRNA非常少。基因敲除正是研究ncRNA功能非常重要的环节, 利用两对TALEN或两个sgRNA的方式对ncRNA进行敲除已成为可能, 并且由于CRISPR/Cas9在ncRNA敲除方面体现出的比基因打靶技术、ZFN和TALEN设计更灵活、构建更简单更快捷、成本花费更低等优点, 它将成为用于ncRNA敲除的主要技术手段, 帮助我们揭示更多ncRNA的作用机制, 以填补大多数ncRNA功能未知这片空白。

参考文献 (References)

- Wright MW, Bruford EA. Naming 'junk': Human non-protein coding RNA (ncRNA) gene nomenclature. *Hum Genomics* 2011; 5(2): 90-8.
- Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, *et al.* Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* 2009; 458(7235): 223-7.
- Carninci P, Kasukawa T, Katayama S, Gough J, Frith MC, Maeda N, *et al.* The transcriptional landscape of the mammalian genome. *Science* 2005; 309(5740): 1559-63.
- Ha M, Kim VN. Regulation of microRNA biogenesis. *Nat Rev Mol Cell Biol* 2014; 15(8): 509-24.
- Stark A, Brennecke J, Bushati N, Russell RB, Cohen SM. Animal microRNAs confer robustness to gene expression and have a significant impact on 3' UTR evolution. *Cell* 2005; 123(6): 1133-46.
- Lim LP, Lau NC, Garrett-Engele P, Grimson A, Schelter JM, Castle J, *et al.* Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* 2005; 433(7027): 769-73.
- Yoon JH, Abdelmohsen K, Gorospe M. Functional interactions among microRNAs and long noncoding RNAs. *Semin Cell Dev Biol* 2014; 34: 9-14.
- Friedman RC, Farh KK, Burge CB, Bartel DP. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res* 2009; 19(1): 92-105.
- Lee RC, Feinbaum RL, Ambros V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* 1993; 75(5): 843-54.
- Wightman B, Ha I, Ruvkun G. Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*. *Cell* 1993; 75(5): 855-62.
- Mraz M, Pospisilova S. MicroRNAs in chronic lymphocytic leukemia: From causality to associations and back. *Expert Rev Hematol* 2012; 5(6): 579-81.
- Luo X, Burwinkel B, Tao S, Brenner H. MicroRNA signatures: Novel biomarker for colorectal cancer? *Cancer Epidemiol Biomarkers Prev* 2011; 20(7): 1272-86.
- Hebert SS, Horre K, Nicolai L, Papadopoulou AS, Mandemakers W, Silaharoglu AN, *et al.* Loss of microRNA cluster miR-29a/b-1 in sporadic Alzheimer's disease correlates with increased BACE1/beta-secretase expression. *Proc Natl Acad Sci USA* 2008; 105(17): 6415-20.
- Wang WX, Rajeev BW, Stromberg AJ, Ren N, Tang G, Huang Q, *et al.* The expression of microRNA miR-107 decreases early in Alzheimer's disease and may accelerate disease progression through regulation of beta-site amyloid precursor protein-cleaving enzyme 1. *J Neurosci* 2008; 28(5): 1213-23.
- Salta E, de Strooper B. Non-coding RNAs with essential roles in neurodegenerative disorders. *Lancet Neurol* 2012; 11(2): 189-200.
- Wang KC, Chang HY. Molecular mechanisms of long noncoding RNAs. *Mol Cell* 2011; 43(6): 904-14.
- Gendrel AV, Heard E. Noncoding RNAs and epigenetic mechanisms during X-chromosome inactivation. *Annu Rev Cell Dev Biol* 2014; 30: 561-80.
- Flynn RA, Chang HY. Long noncoding RNAs in cell-fate programming and reprogramming. *Cell Stem Cell* 2014; 14(6): 752-61.
- da Rocha ST, Edwards CA, Ito M, Ogata T, Ferguson-Smith AC. Genomic imprinting at the mammalian Dlk1-Dio3 domain. *Trends Genet* 2008; 24(6): 306-16.
- Peng X, Gralinski L, Armour CD, Ferris MT, Thomas MJ, Proll S, *et al.* Unique signatures of long noncoding RNA expression in response to virus infection and altered innate immune signaling. *MBio* 2010; 1(5): e00206-10.
- Pontier DB, Gribnau J. Xist regulation and function explored. *Hum Genet* 2011; 130(2): 223-36.
- Tsai MC, Manor O, Wan Y, Mosammamaparast N, Wang JK, Lan F, *et al.* Long noncoding RNA as modular scaffold of histone modification complexes. *Science* 2010; 329(5992): 689-93.
- Court F, Baniol M, Hagege H, Petit JS, Lelay-Taha MN, Carbonell F, *et al.* Long-range chromatin interactions at the mouse *Igf2/H19* locus reveal a novel paternally expressed long non-coding RNA. *Nucleic Acids Res* 2011; 39(14): 5893-906.
- Li H, Yu B, Li J, Su L, Yan M, Zhu Z, *et al.* Overexpression of lncRNA H19 enhances carcinogenesis and metastasis of gastric cancer. *Oncotarget* 2014; 5(8): 2318-29.
- Takahashi K, Yan I, Haga H, Patel T. Long noncoding RNA in liver diseases. *Hepatology* 2014; 60(2): 744-53.
- Gutschner T, Hammerle M, Eissmann M, Hsu J, Kim Y, Hung G, *et al.* The noncoding RNA MALAT1 is a critical regulator of the metastasis phenotype of lung cancer cells. *Cancer Res* 2013; 73(3): 1180-9.
- Wapinski O, Chang HY. Long noncoding RNAs and human disease. *Trends Cell Biol* 2011; 21(6): 354-61.
- Johnson R. Long non-coding RNAs in Huntington's disease neurodegeneration. *Neurobiol Dis* 2012; 46(2): 245-54.
- Clemson CM, Hutchinson JN, Sara SA, Ensminger AW, Fox AH, Chess 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): 717-26.
- Fenoglio C, Ridolfi E, Galimberti D, Scarpini E. An emerging role for long non-coding RNA dysregulation in neurological disorders. *Int J Mol Sci* 2013; 14(10): 20427-42.
- Yoon JH, Abdelmohsen K, Srikantan S, Yang X, Martindale JL,

- De S, *et al.* LincRNA-p21 suppresses target mRNA translation. *Mol Cell* 2012; 47(4): 648-55.
- 32 Wang Y, Xu Z, Jiang J, Xu C, Kang J, Xiao L, *et al.* Endogenous miRNA sponge lincRNA-RoR regulates Oct4, Nanog, and Sox2 in human embryonic stem cell self-renewal. *Dev Cell* 2013; 25(1): 69-80.
- 33 Faghihi MA, Zhang M, Huang J, Modarresi F, van der Brug MP, Nalls MA, *et al.* Evidence for natural antisense transcript-mediated inhibition of microRNA function. *Genome Biol* 2010; 11(5): R56.
- 34 Keniry A, Oxley D, Monnier P, Kyba M, Dandolo L, Smits G, *et al.* The H19 lincRNA is a developmental reservoir of miR-675 that suppresses growth and Igf1r. *Nat Cell Biol* 2012; 14(7): 659-65.
- 35 Elbashir SM, Harborth J, Lendeckel W, Yalcin A, Weber K, Tuschl T. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* 2001; 411(6836): 494-8.
- 36 Jackson AL, Linsley PS. Recognizing and avoiding siRNA off-target effects for target identification and therapeutic application. *Nat Rev Drug Discov* 2010; 9(1): 57-67.
- 37 Capecchi MR. Altering the genome by homologous recombination. *Science* 1989; 244(4910): 1288-92.
- 38 Thomas KR, Folger KR, Capecchi MR. High frequency targeting of genes to specific sites in the mammalian genome. *Cell* 1986; 44(3): 419-28.
- 39 Carroll D. Genome engineering with zinc-finger nucleases. *Genetics* 2011; 188(4): 773-82.
- 40 Skryabin BV, Sukonina V, Jordan U, Lewejohann L, Sachser N, Muslimov I, *et al.* Neuronal untranslated BCI RNA: Targeted gene elimination in mice. *Mol Cell Biol* 2003; 23(18): 6435-41.
- 41 McCreath KJ, Howcroft J, Campbell KH, Colman A, Schnieke AE, Kind AJ. Production of gene-targeted sheep by nuclear transfer from cultured somatic cells. *Nature* 2000; 405(6790): 1066-9.
- 42 Rouet P, Smih F, Jasin M. Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. *Proc Natl Acad Sci USA* 1994; 91(13): 6064-8.
- 43 Carlson DF, Tan W, Lillico SG, Stverakova D, Proudfoot C, Christian M, *et al.* Efficient TALEN-mediated gene knockout in livestock. *Proc Natl Acad Sci USA* 2012; 109(43): 17382-7.
- 44 Cathomen T, Joung JK. Zinc-finger nucleases: the next generation emerges. *Mol Ther* 2008; 16(7): 1200-7.
- 45 Smith J, Bibikova M, Whitby FG, Reddy AR, Chandrasegaran S, Carroll D. Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains. *Nucleic Acids Res* 2000; 28(17): 3361-9.
- 46 Gutschner T, Baas M, Djedrichs S. Noncoding RNA gene silencing through genomic integration of RNA destabilizing elements using zinc finger nucleases. *Genome Res* 2011; 21(11): 1944-54.
- 47 Lee HJ, Kim E, Kim JS. Targeted chromosomal deletions in human cells using zinc finger nucleases. *Genome Res* 2010; 20(1): 81-9.
- 48 Sollu C, Pars K, Cornu TI, Thibodeau-Beganny S, Maeder ML, Joung JK, *et al.* Autonomous zinc-finger nuclease pairs for targeted chromosomal deletion. *Nucleic Acids Res* 2010; 38(22): 8269-76.
- 49 Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, *et al.* Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* 2009; 326(5959): 1509-12.
- 50 Moscou MJ, Bogdanove AJ. A simple cipher governs DNA recognition by TAL effectors. *Science* 2009; 326(5959): 1501.
- 51 Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, *et al.* Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res* 2011; 39(12): e82.
- 52 Liu Y, Luo D, Zhao H, Zhu Z, Hu W, Cheng CH. Inheritable and precise large genomic deletions of non-coding RNA genes in zebrafish using TALENs. *PLoS One* 2013; 8(10): e76387.
- 53 Xiao A, Wang Z, Hu Y, Wu Y, Luo Z, Yang Z, *et al.* Chromosomal deletions and inversions mediated by TALENs and CRISPR/Cas in zebrafish. *Nucleic Acids Res* 2013; 41(14): e141.
- 54 Gupta A, Hall VL, Kok FO, Shin M, McNulty JC, Lawson ND, *et al.* Targeted chromosomal deletions and inversions in zebrafish. *Genome Res* 2013; 23(6): 1008-17.
- 55 Wiedenheft B, Sternberg SH, Doudna JA. RNA-guided genetic silencing systems in bacteria and archaea. *Nature* 2012; 482(7385): 331-8.
- 56 Bolotin A, Quinquis B, Sorokin A, Ehrlich SD. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* 2005; 151(Pt 8): 2551-61.
- 57 Mojica FJ, Diez-Villasenor C, Garcia-Martinez J, Soria E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J Mol Evol* 2005; 60(2): 174-82.
- 58 Pourcel C, Salvignol G, Vergnaud G. CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* 2005; 151(Pt 3): 653-63.
- 59 Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 2007; 315(5819): 1709-12.
- 60 Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 2012; 337(6096): 816-21.
- 61 Makarova KS, Haft DH, Barrangou R, Brouns SJ, Charpentier E, Horvath P, *et al.* Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol* 2011; 9(6): 467-77.
- 62 Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, *et al.* CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 2011; 471(7340): 602-7.
- 63 Garneau JE, Dupuis ME, Villion M, Romero DA, Barrangou R, Boyaval P, *et al.* The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 2010; 468(7320): 67-71.
- 64 Sapranaukas R, Gasiunas G, Fremaux C, Barrangou R, Horvath P, Siksnys V. The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res* 2011; 39(21): 9275-82.
- 65 Gasiunas G, Barrangou R, Horvath P, Siksnys V. Cas9-crRNA

- ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci USA* 2012; 109(39): E2579-86.
- 66 Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* 2013; 339(6121): 819-23.
- 67 Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, *et al.* RNA-guided human genome engineering via Cas9. *Science* 2013; 339(6121): 823-6.
- 68 Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F, *et al.* One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* 2013; 153(4): 910-8.
- 69 Zhou H, Liu B, Weeks DP, Spalding MH, Yang B. Large chromosomal deletions and heritable small genetic changes induced by CRISPR/Cas9 in rice. *Nucleic Acids Res* 2014; 42(17): 10903-14.
- 70 Essletzbichler P, Konopka T, Santoro F, Chen D, Gapp BV, Kralovics R, *et al.* Megabase-scale deletion using CRISPR/Cas9 to generate a fully haploid human cell line. *Genome Res* 2014; 24(12): 2059-65.
- 71 Han J, Zhang J, Chen L, Shen B, Zhou J, Hu B, *et al.* Efficient in vivo deletion of a large imprinted lncRNA by CRISPR/Cas9. *RNA Biol* 2014; 11(7): 829-35.
- 72 Ho TT, Zhou N, Huang J, Koirala P, Xu M, Fung R, *et al.* Targeting non-coding RNAs with the CRISPR/Cas9 system in human cell lines. *Nucleic Acids Res* 2015; 43(3): e17.