

## 本刊专稿

# R2 逆转座子介导的大片段基因整合技术新进展

汪彧泽<sup>1,2</sup>, 毛邦炜<sup>1,2</sup>, 骆胜球<sup>1,2</sup>, 陈阳灿<sup>1,3</sup>, 李伟<sup>1,2,3\*</sup>

**摘要** 基因编辑是通过特定技术手段,对生物体基因组的特定目标位点进行精准修饰,进而调控其遗传信息与表型特征的技术过程。其中,大片段基因整合技术可在生物体基因组中实现大片段外源 DNA 的精准插入或替换,为重大疾病的创新治疗方案开发提供了核心技术支撑。从该技术的历史起源切入,系统梳理其发展脉络,并剖析了当前大片段基因整合技术存在的外源 DNA 依赖、递送困难、免疫风险突出等技术瓶颈。鉴于基于 R2 逆转座子的新型大片段整合技术有望成为突破该瓶颈的关键技术路径,进一步介绍了 R2 逆转座子的结构与靶标引发逆转录(target-primed reverse transcription)的整合机制,综述了其在结构解析、工程化改造、全 RNA 形式的靶向整合及重编程潜力等方面的研究进展,并总结了 R2 逆转座子介导的大片段整合工具的技术突破,以及其在体内功能基因回补、在体生成嵌合抗原受体等应用前景,为其在基因治疗领域的应用提供理论启发与实践参考。

**关键词** 基因整合;基因组工程;R2 逆转座子;RNA 技术;基因治疗

基因编辑技术是一类通过精准识别和修饰生物体基因组外源脱氧核糖核酸(deoxyribonucleic acid, DNA)序列,实现遗传信息与表型特征改变的生物技术<sup>[1]</sup>。大片段基因整合技术是基因编辑领域的重要分支,可以实现外源大片段 DNA 在基因组特定位点的稳定整合,为疾病诊疗提供创新解决方案<sup>[2]</sup>。

随着生命科学研究的深入及医学应用需求的升级,大片段整合技术历经从依赖同源重组(homologous recombination, HR)的传统模式向非 DNA 双链断裂(double-strand break, DSB)依赖的新型整合路径的转变,供体形式也从单一的 DNA 供体拓展至更适

配医学应用场景的核糖核酸(ribonucleic acid, RNA)供体。这一技术演进不仅突破了早期 HR 技术在整合效率上的局限,更推动了基因治疗从理论探索向临床转化的跨越。

本文以该技术的发展脉络为切入点,系统梳理其演变历程:从早期基于 HR 的基因整合技术<sup>[3]</sup>,到结合核酸酶实现可编程的大片段整合策略<sup>[4]</sup>,再到近年兴起的非 DSB 依赖型新技术<sup>[5]</sup>。在此基础上,着重阐述基于 R2 逆转座子的新型大片段整合技术的研究进展,解析其分子机制、技术优势及在疾病模型中的应用潜力,以期对相关领域的基础研究与转化应用提供理论参考。

1. 中国科学院动物研究所器官再生与智造全国重点实验室, 北京 100049
2. 中国科学院大学, 北京 100049
3. 北京干细胞与再生医学研究院, 北京 100101

## 1 大片段整合技术的历史与发展

大片段整合技术是一种能够将大片段 DNA 精准

收稿日期: 2025-06-04; 修回日期: 2026-05-18

基金项目: 国家重点研发计划项目(2024YFA0917300)

作者简介: 汪彧泽, 博士研究生, 研究方向为基因编辑技术的开发与应用, 电子邮箱: wangyuze22@ioz.ac.cn; 李伟(通信作者), 研究员, 研究方向为哺乳动物工程生物学技术平台、哺乳动物器官再生修复机制与技术, 重大疾病生物治疗新方法, 电子邮箱: liwei@ioz.ac.cn

引用格式: 汪彧泽, 毛邦炜, 骆胜球, 等. R2 逆转座子介导的大片段基因整合技术新进展[J]. 科技导报, 2026, 44(11): 24-32; doi:10.3981/j.issn.1000-7857.2025.06.00017

整合至基因组特定位置的基因编辑技术<sup>[6]</sup>。当前,许多疾病的发生源于基因功能异常,而大片段整合技术的应用与发展,或有望成为多种疾病治疗的通用手段<sup>[7-9]</sup>。

大片段整合技术的起源可追溯至 HR 的发现与应用。HR 通过同源序列匹配完成精准基因修复或重组,可以在哺乳动物细胞内实现基因片段整合,奠定了早期大片段基因整合的技术基石。1982 年, Folger 等<sup>[10]</sup>在哺乳动物细胞显微注射外源 DNA 体系中,证实导入的质粒分子间可自发发生 HR。1985 年, Smithies 等<sup>[11]</sup>利用电穿孔法将含人  $\beta$ -珠蛋白同源序列的质粒导入人-鼠杂交细胞,实现了外源 DNA 在哺乳动物细胞内源基因座(人  $\beta$ -珠蛋白位点)的定点整合。随后, Wong 等<sup>[12]</sup>借助显微注射技术深入探究哺乳动物细胞内外源 DNA 的同源重组特性,从分子层面阐明同源重组介导外源基因定点整合的作用机制。20 世纪 90 年代, Rouet 等<sup>[13]</sup>与 Sargent 等<sup>[14]</sup>将 I-SceI 内切酶应用于哺乳动物细胞体系,借助其特异性切割 DNA 的特性引发 DSB,进而激活细胞修复机制,显著提升 HR 效率。细胞修复 DSB 时存在 2 种路径:非同源末端连接(non-homologous end joining, NHEJ)和同源介导的双链 DNA 修复(homology directed repair, HDR)。当缺乏供体 DNA 时, NHEJ 途径对断裂 DNA 末端进行加工后可以重新连接断裂位置,该修复方式容易引发断裂处的序列插入或缺失(insertion-deletion, Indel);而当存在含同源序列的供体 DNA 时,细胞可以利用 HDR 途径精准整合外源基因<sup>[15]</sup>。I-SceI 核酸酶通过 HDR 修复机制显著提高了大片段整合效率,但其存在识别位点有限、可编程性差等诸多局限性,促使科研人员探索更为高效、灵活的核酸酶系统。

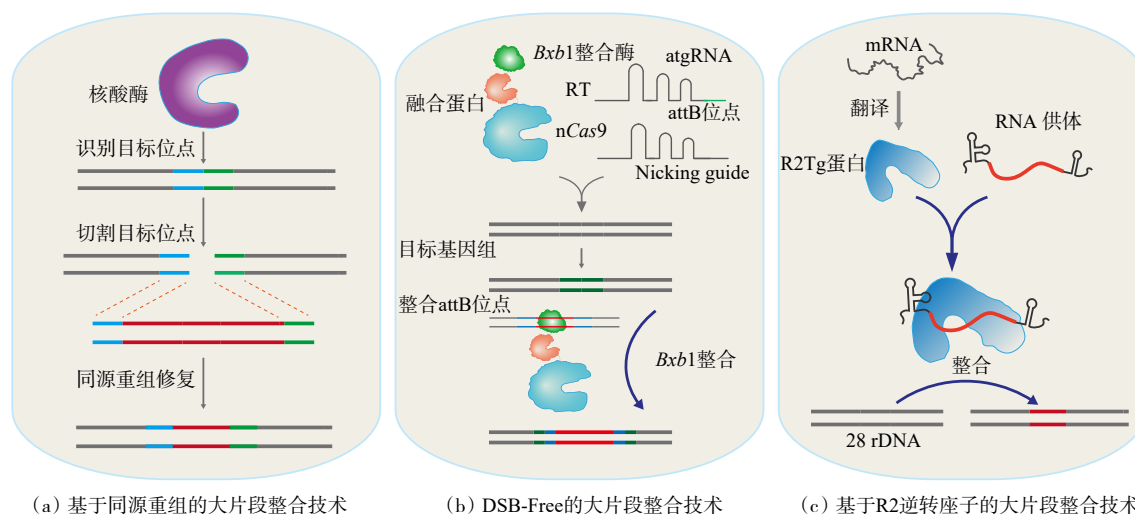
1996 年,基因编辑领域迎来新突破。科学家利用锌指蛋白对 DNA 的特异性识别能力,将多个锌指蛋白串联作为 DNA 识别模块,并与非特异性核酸内切酶 Fok I 偶联,成功构建出锌指核酸酶(zinc finger nucleases, ZFNs)<sup>[16]</sup>。这一可编程核酸内切酶的诞生,打破了传统技术在靶向切割位点上的限制,极大地拓展了 HR 技术的应用范围。此后,研究人员借鉴类似设计思路,开发出转录激活因子样效应物核酸酶(transcription activator-like effector nucleases, TALENs)<sup>[17]</sup>。

相较于 ZFNs, TALENs 在设计上更具灵活性,为基因整合的精准性提供了更有力的选择。

规律间隔成簇短回文重复序列及其相关蛋白(clustered regularly interspaced short palindromic repeats/CRISPR-associated proteins, CRISPR/Cas)系统的问世,标志着 HR 介导的大片段精准整合技术迎来又一次革命性突破。在这一系统中, precursor CRISPR RNA (pre-crRNA)经 trans-activating crRNA(tracrRNA)协助加工形成成熟的 CRISPR RNA(crRNA),后者与 tracrRNA 共同构成的结构可引导 Cas 核酸酶靶向结合 DNA 进行切割<sup>[18-21]</sup>。2012 年, Jinek 等<sup>[22]</sup>在体外解析了酿脓链球菌来源的 CRISPR/Cas9 系统(SpCas9)的功能,将 tracrRNA-crRNA 双 RNA 结构简化为单引导 RNA(sgRNA),并证实该系统可通过改变其中 20 nt 的间隔序列实现可编程基因编辑。凭借 CRISPR/Cas9 基因编辑体系的开创性研究成果, Jennifer A. Doudna 与 Emmanuelle Charpentier 共同获得 2020 年诺贝尔化学奖。随后, Cong 等<sup>[23]</sup>与 Mali 等<sup>[24]</sup>分别成功将其应用于哺乳动物细胞编辑。CRISPR/Cas9 系统具备的 RNA 引导的高度可编程特性,在整个基因编辑领域及生物学研究中都引发了深远变革。同时,该体系可在靶标位点高效产生 DSB,极大提高了 HR 的效率,降低了操作难度,革新了大片段精准整合领域(图 1(a))。

然而,基于 HR 的大片段整合技术依赖于 DSB 的形成,其产生易带来基因组损伤等风险。同时,在细胞修复进程中, NHEJ 途径难以规避,容易产生较多非预期的副产物,进而可能增加细胞损伤或癌变风险<sup>[25]</sup>。因此,开发不依赖 DSB 的新型大片段整合技术在医学应用领域具备重要的科学研究价值与临床转化意义。

为改善传统技术依赖 DSB 可能带来的弊端,研究逐步聚焦于 CRISPR/Cas9 系统的功能改造。该系统的 DNA 切割功能由 HNH 和 RuvC 双结构域协同完成,其中 HNH 结构域负责靶向链切割, RuvC 结构域介导非靶向链断裂<sup>[26]</sup>。通过对 H840A、D10A 等关键活性位点的定点突变,可获得 2 种重要功能变体:仅切割单链的切口酶 nickase Cas9(nCas9)<sup>[27]</sup>和丧失切割活性但保留靶向结合能力的 dead Cas9(dCas9)<sup>[28]</sup>。将这 2 种功能变体与效应蛋白融合,能够开发新型基



(a) 基于同源重组的大片段整合技术 (b) DSB-Free的大片段整合技术 (c) 基于R2逆转座子的大片段整合技术  
 (a) 由核酸酶识别并切割目标位点,产生DSB后,通过同源重组修复途径实现外源大片段的整合。(b) 通过nCas9/RT/Bxb1整合酶融合蛋白,先在目标基因组上生成attB位点,再由Bxb1整合酶介导外源片段的位点特异性整合,全程不引入DSB。(c) R2 mRNA翻译为R2Tg蛋白后,与RNA供体形成核糖核蛋白复合物,特异性识别28S rDNA位点,通过逆转座机制将大片段整合至该位点

图1 大片段整合技术实现路径

因编辑工具,从而绕过DSB修复过程,显著降低Indel频率。研究人员关注到逆转录酶(reverse transcriptase, RT)能够进行逆转录反应,以RNA为模板催化合成互补DNA。若能将其定位到基因组特定位点,或可在DNA修复阶段实现逆转录产物的定点整合。Anzalone等<sup>[29]</sup>通过整合nCas9的靶向单链切割能力与RT的逆转录功能,结合细胞内源性DNA修复机制,成功开发了先导编辑技术(prime editing, PE)。该技术能够高效实现靶向位点的片段插入、缺失及碱基替换,尤其在100 bp以内的小片段整合中表现出显著优势。然而,面对更大尺度的DNA片段,该系统无法实现有效整合<sup>[30]</sup>。为了进一步拓展这一技术,Anzalone等<sup>[31]</sup>和Pandey等<sup>[32]</sup>进一步将PE系统与位点特异性丝氨酸整合酶Bxb1偶联,研发出PASSIGE技术。该技术通过PE在基因组特定位点预安装Bxb1识别序列(如attB/attP位点),利用工程化改造的Bxb1变体识别预设位点,介导外源大片段DNA的精准插入,实现了超过10 kb DNA片段的精准整合,并通过定向进化策略显著提升了整合效率<sup>[32]</sup>(图1(b))。此外,Yarnall等<sup>[33]</sup>采用类似策略独立开发的PASTE技术,成功实现了最高36 kb DNA的大片段基因组整合,为基因治疗领域的大片段递送提供了重要技术支持。

尽管PASSIGE、PASTE等新兴大片段基因整合

技术通过蛋白工程与靶向递送策略实现了高效的外源DNA整合,但其技术框架仍依赖于外源性供体DNA的导入。基于外源DNA供体的基因整合体系虽已在基础研究中得到广泛应用,但其临床转化仍存在2大核心瓶颈。其一,游离外源DNA易被宿主天然免疫受体识别,引起强烈炎症反应与细胞毒性,显著降低体内基因治疗的安全性,缩小临床应用空间<sup>[34]</sup>;其二,外源DNA体内递送效率偏低,现阶段的递送仍高度依赖腺相关病毒(adeno-associated virus, AAV)等病毒载体<sup>[35]</sup>。而病毒载体本身存在制备成本高、质控难度大、易引发预存免疫/适应性免疫应答等固有缺陷,这进一步限制了该类技术在体内基因治疗中的规模化应用与重复给药能力<sup>[36]</sup>。

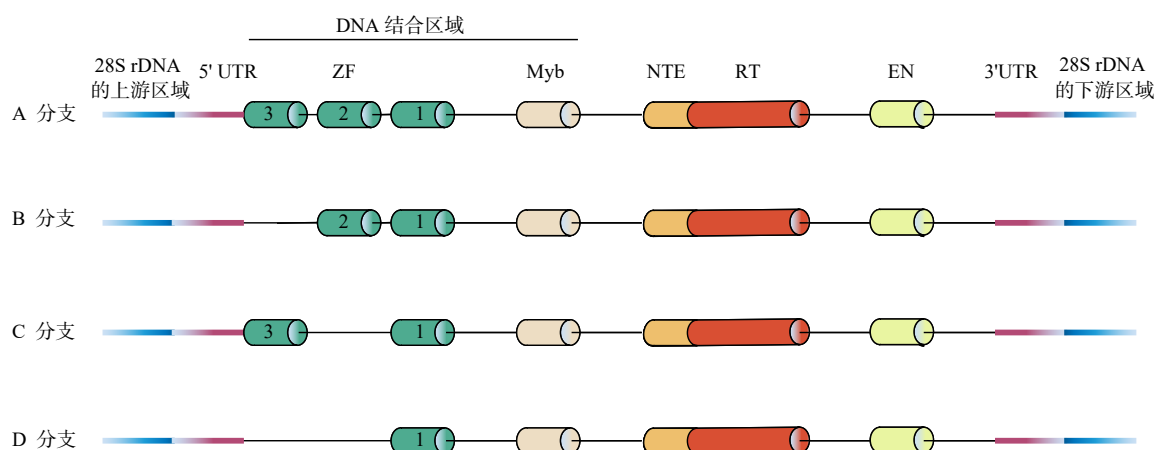
在先导编辑系统中,研究人员利用了RT以RNA模板合成DNA的特性,实现了对基因组序列的整合<sup>[37]</sup>,这为摆脱传统编辑技术对DNA供体的依赖开辟了全新思路。因此,研究者将目光转向自然界中天然存在的逆转座子元件。这类元件能通过“RNA中间体—逆转录—整合”天然路径规避外源性DNA依赖,实现自身大片段基因的整合<sup>[38]</sup>。对逆转座子功能特性与作用机制的探索,成为开发无需外源DNA供体的新型大片段整合技术的关键突破口。其中,R2逆转座子是近年来开发新型大片段整合技术的研究焦点(图1(c))。

## 2 R2 逆转座子的开发与应用

R2 逆转座子(retrotransposon)是一类可特异性整合到核糖体 DNA(rDNA)中的可移动元件,最早由 Dawid 等<sup>[39]</sup>与 Roiha 等<sup>[40]</sup>在果蝇的 28S rDNA 基因位点中鉴定。在真核生物的整个进化过程中,R2 插入位点两侧的 rDNA 基因序列非常保守,几乎没有发生过变化。除了昆虫,鱼类、鸟类等类群中也很快被发现均分布有 R2 逆转座子<sup>[41-42]</sup>,但哺乳动物中尚未发现 R2 逆转座子的存在。

R2 逆转座子由蛋白组分与 RNA 组分 2 部分组

成。其中蛋白组分是一个多功能 R2 蛋白,其 N 端是一个包含有锌指(ZF)与 Myb 样(Myb-like)模块的 DNA 结合结构域,C 端则是一个核酸内切酶结构域(endonuclease, EN),而发挥逆转录功能的 RT 结构域位于蛋白中央,此外 RT 的 N 端还有一个 RNA 结合结构域,命名为 NTE-1。而 RNA 组分则是一条可翻译出 R2 蛋白的 mRNA,其除了蛋白表达框,还包含有两端的非编码区(untranslated regions, UTR)。R2 逆转座子在进化谱系中可进一步划分为 R2-A、R2-B、R2-C、R2-D 4 个分支,各分支之间的核心差异主要体现在 N 端 ZF 基序的数量上(图 2)。



R2 逆转座子可划分为 A、B、C、D 4 个进化分支,各分支均特异性插入宿主 28S rDNA 特定位点,其结构从 5'到 3'依次包含:28S rDNA 上游序列、5'UTR、DNA 结合区域(含 ZF 基序与 Myb 结构域)、NTE、RT 结构域、内切核酸酶结构域及 3'UTR。各分支的差异体现在 N 端 ZF 基序的拷贝数和排列上:A 分支含 3 个 ZF 基序(ZF1、ZF2 和 ZF3),B 分支含 2 个 ZF 基序(ZF1 和 ZF2),C 分支含 2 个 ZF 基序(ZF1 和 ZF3),D 分支仅含 1 个 ZF 基序(ZF1),改绘自文献<sup>[43]</sup>

图 2 R2 逆转座子的组成

逆转座子可利用 RNA 中间体实现在基因组中的复制,意味着其可以完全以 RNA 形式递送并实现基因整合。而 R2 逆转座子特异性整合到 rDNA 的特性避免了随机整合带来的风险,因此十分有利于开发成为可全 RNA 递送的靶向基因整合工具。

### 2.1 R2 逆转座子的整合机制

在逆转座过程中,R2 蛋白的多个功能结构域协同发挥功能,并与 R2 RNA 形成复合物,通过有序的多步反应实现自身序列的精准整合。R2 整合机制与分子过程的阐述主要来自来源于家蚕(*Bombyx mori*)的 R2 逆转座子(R2Bm)为研究模型。

整合开始前,R2 蛋白的 NTE-1 结构域参与识别并结合 R2 RNA 的 3' UTR,从而形成核糖核蛋白复合

物。其中 R2 RNA 的 3' UTR 的识别特异性主要依赖于其二级结构而非序列特征<sup>[44-45]</sup>。随后该复合物特异性识别 rDNA 插入位点附近的 2 段保守 DNA 基序。第一段保守基序(RUM 或 Drr)<sup>[46-48]</sup>位于逆转座子插入位置上游被 R2 的 DNA 结合结构域的 ZF 基序、Myb 基序与 RT 结构域共同识别。R2 逆转座子 DNA 结合域内 ZF 基序的排布位置与其逆转座功能密切相关。研究表明,ZF1-Myb 结构域是介导元件精准插入宿主 rDNA 位点完成逆转座整合的核心必需元件;ZF2 与 ZF3 可调控整合活性及 5'端连接序列的形成精度,但并非逆转座过程必备结构<sup>[47]</sup>。另一段保守基序(RASIN 或 Dcr)则涉及插入位点的切割<sup>[48]</sup>,其在整合过程中靠近 R2 蛋白的核酸内切酶结构域并被解

旋,随后由 R2 蛋白的内切酶活性切割第一条靶 DNA 链,暴露出 3'羟基。

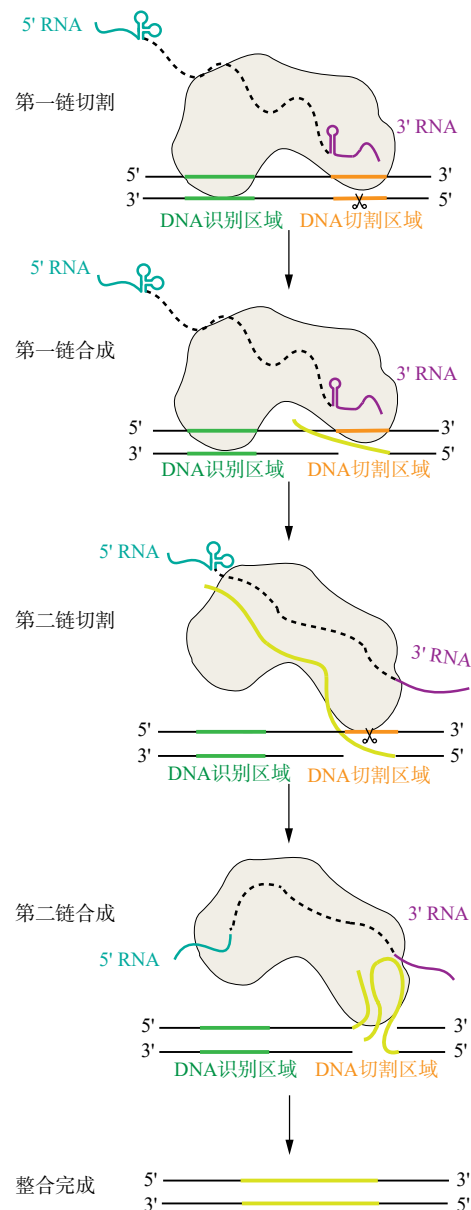
第一条链切割后, R2 蛋白的 RT 启动逆转录过程,利用切口处暴露的 DNA 的 3'羟基与 R2 RNA 模板开始第一条 cDNA 链的合成,这一过程被称为靶标引发的逆转录(target-primed reverse transcription, TPRT)。R2 的 RT 具有很多特殊的特性,首先, R2 的 RT 可利用 DNA 暴露出来的 3'羟基起始逆转录过程<sup>[49]</sup>,并且可以使用非连续的 RNA 模板,实现模板跳跃<sup>[50-51]</sup>。即 R2 的 RT 可以从一个 RNA 模板跳到另一个模板,在非连续模板上进行连续的 cDNA 合成,并且这个过程不需要 2 个 RNA 模板之间的序列具有相似性<sup>[51]</sup>。其次, R2 的 RT 具有较高的持续合成能力,其合成的 cDNA 长度是逆转录病毒 RT 合成长度的 2~5 倍<sup>[52]</sup>。最后, R2 的 RT 缺乏 RNase H 结构域和降解 RNA 模板的相应活性<sup>[52]</sup>,因此逆转座过程中 RNA 模板的去 除不同于常规逆转录过程中 RNA 模板的降解。

第一条 cDNA 链的合成完成后, R2 蛋白的核酸内切酶结构域启动第二条 DNA 链的切割。在 R2Bm 逆转座子模型中,通常认为 2 个 R2 蛋白-RNA 复合物分别于上下游靶 DNA 链结合,并通过 5' RNA(指 R2 蛋白表达框的 5'端的 RNA 片段,与 5' UTR 并不相同)与 3' UTR 来调节 2 个复合物的整合过程。最新的结构研究也支持这一模型:在 3' UTR 存在的情况下, R2 蛋白与上游 DNA 的结合效率提高,同时 5' RNA 由于与 3' UTR 的竞争而从 R2 蛋白上解离,从而抑制第二条链的切割<sup>[48]</sup>。逆转录完成后, 3' UTR 被移除, 5' RNA 与蛋白形成的复合物开始激活第二条链的切割。随后第二条 cDNA 链起始合成,最后通过 DNA 修复完成整个整合(图 3)。

尽管已经有不少针对 R2 逆转座子整合机制的研究,但仍存在一些细节有待解析。例如,第二条 cDNA 链的合成细节尚不清晰,以及逆转座过程是否涉及内源性细胞因子有待阐明。对 R2 逆转座子的独特整合机制的进一步理解为大片段基因整合工具开发提供了新的视野。

## 2.2 R2 逆转座子的研究进展

R2 逆转座子整合的特异性与形成 RNA 中间体的特性,使其成为开发全 RNA 形式的精准靶向基因整合工具的有力候选系统。2019 年, Kuroki-Kami 等<sup>[54]</sup>



R2 逆转座子的 TPRT 机制,展示了其 RNA 识别、逆转录合成 cDNA 链与分步合成的整合过程。改绘自文献<sup>[53]</sup>

图 3 R2 逆转座子的 TPRT 机制示意

就尝试将来源于青鳞鱼的 R2OI 逆转座子用于在斑马鱼中实现位点特异性整合,并且在斑马鱼生殖细胞中实现高效的 EGFP 整合(高达 95%),但 R2OI 逆转座子在哺乳动物细胞上 R2OI 逆转座子的表现并不理想<sup>[55]</sup>。2023 年, Wilkinson 等<sup>[46]</sup>针对 R2Bm 做了结构解析,阐释了 R2Bm 逆转座过程中 TPRT 起始的分子机制,同时在体外证明了 R2Bm 可通过 Cas9 的重编程性质实现对非 28S rDNA 位点的靶向。同期, Luan 等<sup>[49]</sup>鉴定并解析了 R2Bm 逆转座过程中 R2 RNA 组

分中 5' RNA 与 3' UTR 的调控功能。这些工作为深入理解 R2 的逆转座过程与开发逆转座子工具提供了全新的分子层面的见解。

2024 年, Zhang 等<sup>[56]</sup>与 Chen 等<sup>[57]</sup>各自独立开发了可在哺乳动物细胞上实现全 RNA 介导的靶向基因整合的 R2 工具。其中, Zhang 等开发了精确 RNA 介导的转基因插入(precise RNA-mediated insertion of transgenes, PRINT)技术。该技术具备极高的特异性, 3'端整合正确率高达 99%, 5'端整合存在一些重组和截短。在人 hTERT RPE-1 细胞系上, 以超过 50% 的效率将约 2 kb 的转基因成功插入位点特异性 rDNA 位点。而 Chen 等<sup>[57]</sup>通过系统的工程化改造开发的 en-R2Tg 工具, 具有极高的靶向整合特异性(高达 99%)和极好的整合精确度, 约 90% 的 5'端被精确整合到 28S rDNA 预定位点。在 Huh7 细胞与小鼠胚胎中分别实现 25% 与 60% 的整合效率。这 2 种工具设计上略有差异, 但都采用一条 mRNA 表达 R2 蛋白, 另一条 RNA 作为供体的构造, 由于供体 RNA 两端序列在整合过程发挥重要功能, 通过保留 UTR, 将含启动子的目的基因替换原有的 R2 蛋白阅读框, 从而可以实现功能基因的整合与整合后的基因表达。2025 年, Fell 等<sup>[58]</sup>利用 SpCas9<sup>H840A</sup> 与 R2Tocc 逆转座子在哺乳动物细胞中实现了一定程度的 R2 的重编程基因插入。随后, Edmonds 等<sup>[59]</sup>解析鸟类 R2 逆转座子(R2Tg)的结构与生化机制, 开发全 RNA 介导、高效基因整合系统, 在多种人源细胞系中实现 80% 以上整合效率, 可完成高精度度的靶向基因插入。2026 年, McIntyre 等<sup>[60]</sup>进一步阐明 R2 逆转录子介导基因插入的分子机制, 以 2024 年的 PRINT 技术体系为基础, 系统揭示 ATR-Pol $\theta$ 、53BP1-Shieldin-CST-Pol $\alpha$  及 CtIP-MRN 3 条宿主修复通路分别调控微同源连接、截短插入与全长精准整合。在人 hTERT RPE-1 细胞中, PRINT 可实现>50% 效率的 rDNA 位点靶向插入, 通过抑制 ATR-Pol $\theta$  与 53BP1-Shieldin-CST-Pol $\alpha$  通路, 全长精准插入比例可达 95% 以上, 从而调控宿主修复机制显著提升 R2 工具的保真度。

### 3 结论与展望

相较于 CRISPR 同源重组, PE 结合重组酶等 DNA

依赖型大片段基因整合工具, R2 逆转座子以全 RNA 递送模式为核心, 摆脱了外源 DNA 导入带来的细胞毒性与免疫原性问题, 具有易递送、定点整合安全性高等独特优势, 适用于原代免疫细胞改造与体内基因编辑研究。然而, 目前这类工具仍存在整合位点固定、可编程性不足、大片段整合易形成截短插入产物、整体整合效率偏低等缺陷。值得说明的是, 该体系并非旨在取代传统大片段基因整合体系, 而是作为现有技术的重要补充, 进一步完善大片段基因整合技术工具箱, 为细胞治疗、基因功能研究提供全新技术选择。

R2 逆转座子工具有望开发在体功能基因回补、在体生成嵌合抗原受体 T 细胞(chimeric antigen receptor T cell, CAR-T)细胞等全新的疾病治疗方法, 首先, 对于许多单基因遗传病而言, 导致疾病发生的致病点突变往往并不唯一, 相同表型可能存在不止一个致病点突变。对于这类疾病, 采用序列纠正的基因疗法, 往往需要针对患者定制治疗方案<sup>[61-62]</sup>。相较而言, 利用 R2 逆转座子编辑工具回补正常功能性基因, 有望构建通用基因治疗方案。此外, CAR-T 细胞疗法, 目前正在治疗血液系统的恶性肿瘤中展现出显著的疗效。目前, 传统的 CAR-T 细胞疗法主要依赖体外病毒感染来制造 CAR-T 细胞, 该技术体系发展成熟、转导效率高, 能够稳定获得具备高效肿瘤杀伤活性的工程化 T 细胞。但该制备体系过程复杂、成本高, 容易引发随机整合, 病毒组分还易诱发机体免疫排斥反应并且有潜在的致癌风险。而 R2 逆转座子的全 RNA 形式, 可通过非病毒载体进行递送, 无需体外细胞改造与病毒包装流程, 有望实现在体内直接生产 CAR-T 细胞。并且相较于单纯 mRNA 瞬时递送<sup>[63]</sup>, R2 逆转座子的靶向 DNA 整合可实现 CAR 分子的持续表达。并且全 RNA 递送模式无外源 DNA 与病毒骨架掺入, 显著降低免疫原性, 制备流程精简且成本可控。不过 R2 逆转座子技术现阶段仍存在明显局限性, 其靶向整合位点较为固定, 在人原代 T 细胞中的整体整合效率仍不及传统病毒转导体系, 大片段外源基因导入时易出现截短型插入产物, 且该技术仍需大量动物实验与临床前研究进一步验证优化。R2 逆转座子作为一类全新的基因整合工具, 通过全 RNA 递送可实现精确的基因整合, 从而为开发全新的疗法提

供了一个可扩展的生物技术平台。现有的 R2 逆转座子工具在哺乳动物细胞上的基因整合效率仍有很大的提升空间,而随着未来对 R2 逆转录转座子结构、整合机制的深入解析,有望利用这种独特的可移动元件设计出更高效的、全 RNA 递送与可编程的基因整合工具。

### 参考文献 (References)

- [1] Pál C, Papp B, Pósfai G. The dawn of evolutionary genome engineering[J]. *Nature Reviews Genetics*, 2014, 15(7): 504–512.
- [2] Li B S, Sun C, Li J Y, et al. Targeted genome-modification tools and their advanced applications in crop breeding[J]. *Nature Reviews Genetics*, 2024, 25(9): 603–622.
- [3] Capecchi M R. Altering the genome by homologous recombination[J]. *Science*, 1989, 244(4910): 1288–1292.
- [4] Sommer D, Peters A E, Wirtz T, et al. Efficient genome engineering by targeted homologous recombination in mouse embryos using transcription activator-like effector nucleases[J]. *Nature Communications*, 2014, 5: 3045.
- [5] Abdullah, Jiang Z Z, Hong X L, et al. CRISPR base editing and prime editing: DSB and template-free editing systems for bacteria and plants[J]. *Synthetic and Systems Biotechnology*, 2020, 5(4): 277–292.
- [6] Villiger L, Joung J, Koblan L, et al. CRISPR technologies for genome, epigenome and transcriptome editing[J]. *Nature Reviews Molecular Cell Biology*, 2024, 25(6): 464–487.
- [7] Landrum M J, Lee J M, Riley G R, et al. ClinVar: Public archive of relationships among sequence variation and human phenotype[J]. *Nucleic Acids Research*, 2014, 42(1): 980–985.
- [8] Weischenfeldt J, Symmons O, Spitz F, et al. Phenotypic impact of genomic structural variation: Insights from and for human disease[J]. *Nature Reviews Genetics*, 2013, 14(2): 125–138.
- [9] Yang H, Schloss J, Dunn M, et al. A global reference for human genetic variation[J]. *Nature*, 2015, 526(7571): 68–74.
- [10] Folger K R, Wong E A, Wahl G, et al. Patterns of integration of DNA microinjected into cultured mammalian cells: Evidence for homologous recombination between injected plasmid DNA molecules[J]. *Molecular and Cellular Biology*, 1982, 2(11): 1372–1387.
- [11] Smithies O, Gregg R G, Boggs S S, et al. Insertion of DNA sequences into the human chromosomal  $\beta$ -globin locus by homologous recombination[J]. *Nature*, 1985, 317(6034): 230–234.
- [12] Wong E A, Capecchi M R. Analysis of homologous recombination in cultured mammalian cells in transient expression and stable transformation assays[J]. *Somatic Cell and Molecular Genetics*, 1986, 12(1): 63–72.
- [13] Rouet P, Smih F, Jasin M. Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 1994, 91(13): 6064–6068.
- [14] Sargent R G, Brenneman M A, Wilson J H. Repair of site-specific double-strand breaks in a mammalian chromosome by homologous and illegitimate recombination[J]. *Molecular and Cellular Biology*, 1997, 17(1): 267–277.
- [15] Sonoda E, Hohegger H, Saberi A, et al. Differential usage of non-homologous end-joining and homologous recombination in double strand break repair[J]. *DNA Repair*, 2006, 5(9/10): 1021–1029.
- [16] Kim Y G, Cha J, Chandrasegaran S. Hybrid restriction enzymes: Zinc finger fusions to Fok I cleavage domain[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 1996, 93(3): 1156–1160.
- [17] Boch J, Scholze H, Schornack S, et al. Breaking the code of DNA binding specificity of TAL-type III effectors[J]. *Science*, 2009, 326(5959): 1509–1512.
- [18] Brouns S J J, Jore M M, Lundgren M, et al. Small CRISPR RNAs guide antiviral defense in prokaryotes[J]. *Science*, 2008, 321(5891): 960–964.
- [19] Marraffini L A, Sontheimer E J. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA[J]. *Science*, 2008, 322(5909): 1843–1845.
- [20] Garneau J E, Dupuis M, Villion M, et al. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA[J]. *Nature*, 2010, 468(7320): 67–71.
- [21] Deltcheva E, Chylinski K, Sharma C M, et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III[J]. *Nature*, 2011, 471(7340): 602–607.
- [22] Jinek M, Chylinski K, Fonfara I, et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity[J]. *Science*, 2012, 337(6096): 816–821.
- [23] Cong L, Ran F A, Cox D, et al. Multiplex genome engineering using CRISPR/cas systems[J]. *Science*, 2013, 339(6121): 819–823.
- [24] Mali P, Yang L H, Esvelt K M, et al. RNA-guided human genome engineering via Cas9[J]. *Science*, 2013, 339(6121): 823–826.
- [25] Song B, Yang S, Hwang G H, et al. Analysis of NHEJ-based DNA repair after CRISPR-mediated DNA cleavage [J]. *International Journal of Molecular Sciences*, 2021, 22(12): 6397.
- [26] Gasiunas G, Barrangou R, Horvath P, et al. Cas9-crRNA

- ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2012, 109(39): 2579–2586.
- [27] Ran F A, Hsu P D, Lin C Y, et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity[J]. *Cell*, 2013, 154(6): 1380–1389.
- [28] Qi L S, Larson M H, Gilbert L A, et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression[J]. *Cell*, 2013, 152(5): 1173–1183.
- [29] Anzalone A V, Randolph P B, Davis J R, et al. Search-and-replace genome editing without double-strand breaks or donor DNA[J]. *Nature*, 2019, 576(7785): 149–157.
- [30] Chen P J, Liu D R. Prime editing for precise and highly versatile genome manipulation[J]. *Nature Reviews Genetics*, 2023, 24(3): 161–177.
- [31] Anzalone A V, Gao X D, Podracky C J, et al. Programmable deletion, replacement, integration and inversion of large DNA sequences with twin prime editing[J]. *Nature Biotechnology*, 2022, 40(5): 731–740.
- [32] Pandey S, Gao X D, Krasnow N A, et al. Efficient site-specific integration of large genes in mammalian cells *via* continuously evolved recombinases and prime editing[J]. *Nature Biomedical Engineering*, 2025, 9(1): 22–39.
- [33] Yarnall M T N, Ioannidi E I, Schmitt-Ulms C, et al. Drag-and-drop genome insertion of large sequences without double-strand DNA cleavage using CRISPR-directed integrases[J]. *Nature Biotechnology*, 2023, 41(4): 500–512.
- [34] Briard B, Place D E, Kanneganti T D. DNA sensing in the innate immune response[J]. *Physiology*, 2020, 35(2): 112–124.
- [35] Madigan V, Zhang F, Dahlman J E. Drug delivery systems for CRISPR-based genome editors[J]. *Nature Reviews Drug Discovery*, 2023, 22(11): 875–894.
- [36] Ronzitti G, Gross D A, Mingozzi F. Human immune responses to adeno-associated virus (AAV) vectors[J]. *Frontiers in Immunology*, 2020, 11: 670.
- [37] Baltimore D. Viral RNA-dependent DNA polymerase: RNA-dependent DNA polymerase in virions of RNA tumour viruses[J]. *Nature*, 1970, 226(5252): 1209–1211.
- [38] Wicker T, Sabot F, Hua-Van A, et al. A unified classification system for eukaryotic transposable elements[J]. *Nature Reviews Genetics*, 2007, 8(12): 973–982.
- [39] Dawid I B, Rebert M L. Nucleotide sequences at the boundaries between gene and insertion regions in the rDNA of *Drosophila melanogaster*[J]. *Nucleic Acids Research*, 1981, 9(19): 5011–5020.
- [40] Roiha H, Miller J R, Woods L C, et al. Arrangements and rearrangements of sequences flanking the two types of rDNA insertion in *D. Melanogaster*[J]. *Nature*, 1981, 290(5809): 749–754.
- [41] Fujiwara H, Ogura T, Takada N, et al. Introns and their flanking sequences of *Bombyx mori* rDNA[J]. *Nucleic Acids Research*, 1984, 12(17): 6861–6869.
- [42] Smith V L, Beckingham K. The intron boundaries and flanking rRNA coding sequences of *Calliphora erythrocephala* rDNA[J]. *Nucleic Acids Research*, 1984, 12(3): 1707–1724.
- [43] Thompson B K, Christensen S M. Independently derived targeting of 28S rDNA by A- and D-clade R2 retrotransposons: Plasticity of integration mechanism[J]. *Mobile Genetic Elements*, 2011, 1(1): 29–37.
- [44] Mathews D H, Banerjee A R, Luan D D, et al. Secondary structure model of the RNA recognized by the reverse transcriptase from the R2 retrotransposable element[J]. *RNA*, 1997, 3(1): 1–16.
- [45] Ruschak A M, Mathews D H, Bibillo A, et al. Secondary structure models of the 3' untranslated regions of diverse R2 RNAs[J]. *RNA*, 2004, 10(6): 978–987.
- [46] Wilkinson M E, Frangieh C J, MacRae R K, et al. Structure of the R2 non-LTR retrotransposon initiating target-primed reverse transcription[J]. *Science*, 2023, 380(6642): 301–308.
- [47] Lee R J, Horton C A, Van Treeck B, et al. Conserved and divergent DNA recognition specificities and functions of R2 retrotransposon N-terminal domains[J]. *Cell Reports*, 2024, 43(5): 114239.
- [48] Deng P J, Tan S Q, Yang Q Y, et al. Structural RNA components supervise the sequential DNA cleavage in R2 retrotransposon[J]. *Cell*, 2023, 186(13): 2865–2879. e20.
- [49] Luan D D, Eickbush T H. RNA template requirements for target DNA-primed reverse transcription by the R2 retrotransposable element[J]. *Molecular and Cellular Biology*, 1995, 15(7): 3882–3891.
- [50] Bibillo A, Eickbush T H. The reverse transcriptase of the R2 non-LTR retrotransposon: Continuous synthesis of cDNA on non-continuous RNA templates[J]. *Journal of Molecular Biology*, 2002, 316(3): 459–473.
- [51] Bibillo A, Eickbush T H. End-to-end template jumping by the reverse transcriptase encoded by the R2 retrotransposon[J]. *Journal of Biological Chemistry*, 2004, 279(15): 14945–14953.
- [52] Kurzynska-Kokorniak A, Jamburuthugoda V K, Bibillo A, et al. DNA-directed DNA polymerase and strand displacement activity of the reverse transcriptase encoded by the R2 retrotransposon[J]. *Journal of Molecular Biology*, 2007, 374(2): 322–333.
- [53] Luo S Q, Chen Q C, Chen Y C, et al. Biology and utilization of R2 retrotransposons[J]. *RNA Biology*, 2025, 22(1):

- 1–8.
- [54] Kuroki–Kami A, Nichuguti N, Yatabe H, et al. Targeted gene knockin in zebrafish using the 28S rDNA–specific non–LTR–retrotransposon R2OI[J]. *Mobile DNA*, 2019, 10(1): 23.
- [55] Su Y T, Nichuguti N, Kuroki–Kami A, et al. Sequence–specific retrotransposition of 28S rDNA–specific LINE R2OI in human cells[J]. *RNA*, 2019, 25(11): 1432–1438.
- [56] Zhang X Z, Van Treeck B, Horton C A, et al. Harnessing eukaryotic retroelement proteins for transgene insertion into human safe–harbor loci[J]. *Nature Biotechnology*, 2025, 43(1): 42–51.
- [57] Chen Y C, Luo S Q, Hu Y P, et al. All–RNA–mediated targeted gene integration in mammalian cells with rationally engineered R2 retrotransposons[J]. *Cell*, 2024, 187(17): 4674–4689. e18.
- [58] Fell C W, Villiger L, Lim J, et al. Reprogramming site–specific retrotransposon activity to new DNA sites[J]. *Nature*, 2025, 642(8069): 1080–1089.
- [59] Edmonds K K, Wilkinson M E, Strebinger D, et al. Structure and biochemistry–guided engineering of an all–RNA system for DNA insertion with R2 retrotransposons[J]. *Nature Communications*, 2025, 16: 6079.
- [60] McIntyre J J R, Horton C A, Collins K. Different DNA repair pathways support intact or truncated insertions by R2 retrotransposon protein[J]. *Science*, 2026, 391(6788): eadz3121.
- [61] Bansal V, Gassenhuber J, Phillips T, et al. Spectrum of mutations in monogenic diabetes genes identified from high–throughput DNA sequencing of 6888 individuals[J]. *BMC Medicine*, 2017, 15(1): 213.
- [62] Aksentijevich I, Schnappauf O. Molecular mechanisms of phenotypic variability in monogenic autoinflammatory diseases[J]. *Nature Reviews Rheumatology*, 2021, 17(7): 405–425.
- [63] Wang Y, Yang M, Zhang B Y, et al. Leveraging T cell–specific fusogenicity of HIV for *in vivo* mRNA delivery to produce human CAR–T cells[J]. *Cell Biomaterials*, 2025, 1(2): 100011.

## Recent advances in R2 retrotransposon–mediated large–fragment gene integration technology

WANG Yuze<sup>1,2</sup>, MAO Bangwei<sup>1,2</sup>, LUO Shengqiu<sup>1,2</sup>, CHEN Yangcan<sup>1,3</sup>, LI Wei<sup>1,2,3\*</sup>

1. State Key Laboratory of Organ Regeneration and Reconstruction, Institute of Zoology, Chinese Academy of Sciences, Beijing 100049, China
2. University of Chinese Academy of Sciences, Beijing 100049, China
3. Beijing Institute for Stem Cell and Regenerative Medicine, Beijing 100101, China

**Abstract** Gene editing refers to the process of specifically modifying organism's genome using specific technical approaches to regulate its genetic information and phenotypic characteristics. Among them, large–fragment gene integration technology enables the accurate insertion or replacement of large exogenous DNA fragments in the organismal genome, which provides core technical support for the development of innovative therapeutic strategies against severe diseases. Starting from the historical origin of this technology, this paper systematically sorts out its developmental trajectory and analyzes the current technical bottlenecks of large–fragment gene integration technology, including exogenous DNA dependence, delivery challenges and prominent immune risks. Given that the novel large–fragment integration techniques based on R2 retrotransposons are expected to become a key solution to break through these limitations, the study further introduces the structural characteristics and retrotransposition mechanisms of R2 retrotransposons, and reviews the research advances in its structural analysis and engineering modification, and summarizes the technical breakthroughs and application prospects of R2 retrotransposon–mediated large–fragment integration tools, which provides theoretical insights and practical references for the application of R2 retrotransposon tools in gene therapy.

**Keywords** gene integration; genome engineering; R2 retrotransposon; RNA technology; gene therapy ●



(责任编辑 徐丽娇)