



评述

位点特异性重组系统的机理和应用

张霖^①, 赵国屏^{①②③*}, 丁晓明^{①*}^① 复旦大学生命科学学院微生物学与微生物工程学系, 上海 200433;^② 中国科学院上海生命科学研究院植物生理生态研究所合成生物学重点实验室, 上海 200032;^③ 国家人类基因组南方研究中心省部共建疾病与健康基因组学重点实验室, 上海 201203

* 联系人, E-mail: gpzhao@sibs.ac.cn; xmding74@fudan.edu.cn

收稿日期: 2010-09-05; 接受日期: 2010-10-06

国家高技术研究发展计划(批准号: 2007AA021206)和国家自然科学基金(批准号: 30600009 和 30830002)资助项目

摘要 位点特异性重组酶识别特定的位点形成联会复合体, 并发生 DNA 链的切割与交换, 实现靶位点之间的整合、切离或倒位. 这一过程由位于重组酶催化活性中心的酪氨酸或丝氨酸向 DNA 磷酸骨架发起攻击, 形成共价中间体, 不需要高能量辅助因子的参与. 由于位点特异性重组系统具有高效精确的优点, 在基因工程领域得到了广泛的应用. 本文从识别位点的性质、重组酶的组成与结构及催化反应的特点三方面对位点特异性重组的机理进行了全面的阐述, 并对当前重组酶的应用研究之现状、热点及存在的问题作了深入剖析, 并讨论了未来的发展趋势.

关键词位点特异性重组
酪氨酸重组酶
丝氨酸重组酶
整合酶

近年来, 位点特异性重组酶(site-specific recombinases, SSRs)介导的重组反应在遗传工程操作中得到了广泛的应用, 由于该反应具有底物专一、快速高效、易于改造应用等特点, 获得了广泛的关注. 位点特异性重组(site-specific recombination)发生于特定的位点之间, 由重组酶催化核心的氨基酸向 DNA 骨架发起攻击并介导双链交换实现重组. 位点特异性重组涉及范围广, 存在于各类细胞, 彼此发挥着非常不同但却十分特殊的作用. 一般而言, 该过程包括: 噬菌体整合到宿主染色体与切离(典型代表为大肠杆菌噬菌体 λ 和链霉菌噬菌体 ϕ C31), 一些复合转座子相关共合体(cointegrate)的解离(如转座子 $\gamma\delta$ 和 Tn3), 某些基因表达的调节(如沙门氏菌(*Salmonella typhimurium*)鞭毛相转变)^[1]. 由于催化核心、识别位点及反应机制上的差异, 严格定义的位点特异性重组排除了免疫球蛋白基因的重排(B 淋巴细胞成熟过程中的

VDJ 重排)^[2,3]、移动内含子的归巢(intron homing)^[4]、逆转录病毒与宿主基因组的整合等情况, 虽然在某些个例中识别位点是特异的.

典型的位点特异性重组系统具备以下 3 个要素: 一组特定的识别位点(recognition site)、有简单的识别序列或具有负责不同蛋白因子识别的复杂结构; 识别 DNA 序列、介导切割重连并实现链交换的重组酶(SSRs), 有的系统中只是单一因子, 更多的情况是不同的蛋白因子协同作用以实现不同的重组结果; 位点特异性重组没有 DNA 的合成, 没有能量的获取或丢失, 则需要一套专司 DNA 断裂重连仍能保持磷酸二酯键平衡的特殊机制. 根据重组位点的序列和排列的方向性, 位点特异性重组有 3 种可能的结果, 即整合(integration)、切离(excision)和倒位(inversion). 如图 1 所示, 整合发生在当两个位点在不同的 DNA 分子上(假设其中之一为环状)且有特定的方向; 当两

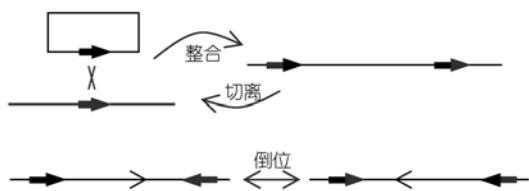


图 1 位点特异性重组的 3 种可能的结果

位点位于同一 DNA 分子上时, 不同的方向导致不同的结果: 两位点方向一致时切离, 相反时发生倒位^[1]. 以上这些特点, 为位点特异性重组系统的应用提供了良好的基础.

1 位点特异性重组系统的来源和分类

如前所述, 位点特异性重组参与的生物学过程种类繁多, 为方便鉴定分类和进行系统研究, 在排除抗体成熟过程中的 VDJ 重组等情况后, 将位点特异性重组系统分为两类, 即酪氨酸重组酶(tyrosine recombinase)和丝氨酸重组酶(serine recombinase), 分类原则主要是基于氨基酸序列的同源性和催化机制的差异. 位点特异性重组酶的来源大致包括: 噬菌体

(原噬菌体)、转座子、环状质粒以及微生物基因组编码; 这些重组酶在不同的生物过程中发挥着整合、切离、倒位、解离、转座等生物学功能(表 1).

一个位点特异性重组系统至少需要重组酶和一对相应的重组位点. 最简单的位点一般为 20~30 bp 的双链 DNA 序列, 具反向重复的两臂被二聚体重组酶识别结合. 位点中央是 DNA 断裂重连的地方, 也就是交换位点(crossing over sites). 然而实际上许多重组位点要复杂得多, 不只含有交换位点, 有的还有超过 100 bp 的附加序列. 这一复杂的位点可以和一个简单位点重组(λ integrase)^[5], 也可能和另外一个复杂位点重组($\gamma\delta$ resolvase)^[21]. 附加序列含有重组酶或者其他辅助因子(accessory host factors)的识别位点. 这些辅助因子可能参与整合酶结构变化, 也可能只是起调节作用, 或兼而有之. 它们可能起始并稳定位点配对, 或者抑制不正确的配对; 可能帮助重组酶的催化亚基靠近切割位点; 可能激活重组酶; 或者帮助控制反应的方向^[1].

酪氨酸重组酶家族也称为 λ 整合酶系(λ integrase family), 该家族数量庞大, 种类繁多, 包括一系列 λ 相关的噬菌体编码酶类和部分转座酶、一些控制原

表 1 位点特异性重组酶的分类

分类/功能	重组酶	来源	氨基酸数目(aa)	生物学作用	参考文献
酪氨酸家族	λ Int	大肠杆菌	356	噬菌体 λ 与宿主基因组的整合与切离	[5]
	P22 Int	沙门氏菌	387	噬菌体 P22 与宿主基因组的整合与切离	[6]
	L5 Int	分枝杆菌	332	噬菌体 L5 与宿主基因组的整合与切离	[7]
	XerC/D 解离酶	各种细菌	295~300	原核生物基因组二聚体的解离	[8]
	Cre 重组酶	大肠杆菌	343	P1 噬菌体二聚体质粒的解离	[9]
整合					
解离	TnpI (Tn4430, Tn5401)	苏云金芽孢杆菌	284/306	转座子复合体的解离	[10,11]
	TnpA (Tn554)	金黄色葡萄球菌	361	与 TnpB, TnpC 协同控制转座子的整合与切离	[12]
转座	Int of Tn916/Tn1545	肠球菌/链球菌	405	环状转座子的整合与切离	[13,14]
切离	XisA, XisC	鱼腥藻属	354/498	鱼腥藻氢化酶基因发育调控中的重排	[15]
	FimB, FimE	大肠杆菌	200/198	大肠杆菌纤毛相转变	[16]
	FliP	酵母	423	酵母 2 μ 质粒的倒位	[17]
倒位	Fim MrpI	变形杆菌属	205	控制奇异变形杆菌菌毛的相转变	[18,19]
丝氨酸家族	TnpR ($\gamma\delta$ /Tn3)解离酶	大肠杆菌	183/185	复合型转座子中共合体的解离	[20~24]
解离	Sin 解离酶	金黄色葡萄球菌	202	葡萄球菌二聚体质粒的消除	[25,26]
	Hin 转化酶	鼠伤寒沙门氏菌	190	沙门氏菌鞭毛的相转变	[27]
倒位	Gin/Cin 转化酶	肠杆菌	193/186	噬菌体 Mu 和 P2 尾蛋白基因的倒位	[28~30]
	ϕ C31/ ϕ BT1/TG1/R4 Int	链霉菌	605/594/619/469	链霉菌噬菌体的整合与切离	[31~34]
整合	Bxb1/ ϕ Rv1 Int	分枝杆菌	500/469	分枝杆菌噬菌体的整合与切离	[35,36]
	TP901-1 Int	乳球菌	485	噬菌体的整合与切离	[37]
	ϕ FC1 Int	肠球菌	464	噬菌体的整合与切离	[38]
转座	TnpX (Tn4451)	梭菌	560	转座子 Tn4451 的整合与切离	[39]
切离	SpoIVCA (CisA)	枯草芽孢杆菌	500	枯草芽孢杆菌发育过程中某些基因的激活	[40]

核生物鞭毛(菌毛)相位转变(phase variation)的蛋白因子, 以及少量参与发育过程中特殊蛋白的表达和染色体分配过程的重组酶, 在酪氨酸重组酶的催化结构域, 都有一个保守的酪氨酸残基负责攻击核酸骨架, 形成单链断裂切口而产生 5'-OH 末端和 3'-磷酸基-酪氨酸的切割中间体, 两个 DNA 分子交错连接, 进而形成一对新的重组位点; 该过程类似于同源重组过程中的 Holliday 模型^[1]. 对酪氨酸家族重组酶了解较为清楚的有: 大肠杆菌噬菌体 λ 编码的整合酶^[5,41]、分枝杆菌噬菌体 L5 整合酶^[7]、P1 噬菌体的 Cre 重组酶^[9]、酵母 2 μ 质粒编码的倒位酶 Flp(Flp invertase)^[17]以及细菌的 XerCD 蛋白^[42](表 1).

不同于酪氨酸重组酶, 丝氨酸重组酶有较为保守的催化结构域, 其近 N 端为活性中心, 包含一个丝氨酸. 该丝氨酸残基亲核攻击 DNA 骨架并交错切割, 形成带有 3'-OH 的双链断裂末端, 5'-磷酸基与重组酶形成共价连接中间体, 随后该联会复合物(synaptic complexes)发生翻转和重新连接完成重组反应; 该过程类似于双链断裂重组模型(double-strand breakage). 丝氨酸重组酶家族也称为倒位/解离酶系(invertase/resolvase family), 因为该家族包含大量的倒位酶与解离酶, 研究较为深入的 $\gamma\delta$ 重组酶和 Tn3 转座酶即属于解离酶亚家族(resolvase subfamily)^[20-24]; 控制鼠伤寒沙门氏菌鞭毛相转变的重组酶 Hin, 即属于倒位酶亚家族(invertase subfamily)^[27]. 然而, 随着研究的深入, 一类分子量较大、介导整合和切离的重组酶被单独划分出来, 称为大型丝氨酸重组酶(large serine recombinase)^[43]; 该亚家族研究较多的主要为细菌噬菌体编码的整合酶, 如链霉菌噬菌体编码的 ϕ C31 和 ϕ BT1 整合酶^[32,33]以及分枝杆菌噬菌体编码的 Bxb1 和 ϕ Rv1 整合酶^[35,36](表 1).

2 位点特异性重组系统的反应机理研究

关于位点特异性重组系统本身的研究由来已久, 最早的综述文献可见于 1971 年^[44]. 几十年来, 科学工作者对位点特异性重组酶(及其辅助蛋白因子)、参与识别与重组的 DNA 序列以及重组发生的机制进行了大量深入而具体的研究. 其中研究时间最长、也最为透彻的属大肠杆菌噬菌体 λ 编码的整合切离系统^[45,46]. 位点特异性重组酶的结构解析、重组机理的深入探讨以及基于结构的酶定向分子改造仍是当前核酸相关酶

类的研究热点之一.

2.1 识别位点的结构与构象

虽然两类重组酶催化反应的机制差异显著, 但其催化反应的位点却未表现出明显区分, 除了在其核心(core sequences)都具有短序列的同源区域外, 几乎无规律可循. 有的识别位点相对复杂, 包含核心序列(链交换区域)、蛋白结合序列、宿主辅助因子结合序列等; 而有的位点十分简单, 只需要一个由间隔区(spacer sequences)分开的两端反向重复序列, 仅为几十个核苷酸.

大肠杆菌噬菌体 λ -整合酶识别自身的 attP (attachment site of phage)位点和大肠杆菌染色体中的 attB (attachment site of bacteria)位点, attP 位点结构复杂, 而后者相对简单(图 2(A))^[5,46]. attP 和 attB 位点中央都包含 7 bp 的核心序列(crossover site, overlap region or spacer sequences), 核心序列的两侧为 7 bp 的反向重复序列(inverted repeats), 分别称为 C 和 C', B 和 B'. attP 位点结构更为复杂, 还包括臂型结合位点(arm-type sites), 分别称之为 P1, P2, P'1, P'2 和 P'3, 这些位点在反应过程中都与整合酶 Int 结合; 此外, 还包含 3 个宿主因子(integration host factor, IHF)结合位点 H1, H2 和 H', 3 个切离酶(excisionase, Xis)结合位点以及一个辅助因子(factor for inversion stimulation, Fis)结合位点^[47,48]. λ 整合酶正是通过如此复杂的底物结构来实现重组反应方向的精确控制. 比较而言, 大肠杆菌噬菌体 P1 编码的 Cre-loxP 系统^[9,47]和酵母 2 μ 质粒编码的 Flp-*frt* 系统的识别位点结构要简单许多^[17](图 2(B)). 在这两个系统中, 除了 8 bp 的核心序列和两侧各 13 bp 的完全对称反向重复序列外, 没有任何其他可与辅助因子结合的序列. 因为只有一对完全一样的重组位点, 所以 Cre-loxP 和 Flp-*frt* 重组系统的方向控制并不严格, 反应并未倾向某一方向, 而是趋于平衡. 一般认为, 酪氨酸重组酶系统对核心位点 DNA 序列的同源性要求十分严格. 以 λ 系统为例, 其中一位点的核心序列突变后, 反应则不能进行; 另一位点的核心做同样突变后, 反应恢复^[49,50]. 这是因为核心序列切割末端负责单链切割侵入时的同源配对. 然而, 粪肠球菌(*Enterococcus faecalis*)转座子 Tn916^[51]和肺炎链球菌(*Streptococcus pneumoniae*)转座子 Tn1545^[52]所编码整合酶的识别序列核心区并不具有同源性. 最近的研究发现, 拟杆菌属(*Bacteroides*)转座子 CtnDOT 识别位点特征与 Tn916 和 Tn1545

一致, 重组之后核心位点碱基并不配对, 需要经过进一步复制而保持遗传稳定^[53,54]; 转座子 NBU1 对识别位点的同源性要求并不在重组位点^[55]; 而噬菌体 CTX 通过单链交换机制(single strand exchange mechanism)实现与非同源核心位点的重组^[56,57]. 这都对传统的同源末端侵入形成 Holliday Junction 中间体的模型提出了挑战.

丝氨酸解离酶 $\gamma\delta$ 的 *res* 位点近 120 bp, 包含 3 个二聚体蛋白的结合位点 site I, site II 和 site III(图 2(C)). 切割的位置位于 site I, 但其他 2 个位点也是重组所必需的. 3 个位点的空间位置非常重要^[20,58]. 3 个位点的特性在二聚体结合中比较罕见, 各自的几何构象不同; 都包含一对尾对尾(tail-to-tail)的识别序列, 但是 3 位点的间隔序列长短不同, 分别为 4, 10 和 1 bp. 从侧面说明, $\gamma\delta$ 解离酶二聚体具有不寻常的柔性(也

可见于其他丝氨酸重组酶), 使得 DNA 结合域能适应不同的间距和构象. 辅助位点 Site II 和 site III 对控制反应的结果尤为重要. 倒位酶 Hin 的识别序列为 26 bp 的两段反向重复序列 *hixL* 和 *hixR*, *hin* 基因介于两位点之间^[59-61]; Hin 介导启动子的倒位来控制下游基因的表达. 反应的发生需要 Fis 的参与, 其识别位点为一段增强子(enhancer), 含有两段被 48 bp 隔开的二聚体识别序列^[62](图 2(D)). 噬菌体编码的大型丝氨酸重组酶的识别位点相对简单. 以 ϕ BT1 整合系统为例, 除其核心区含有 9 bp 的同源序列外, 两位点差异显著, 但都有 5'-GT 为核心并呈现不规则的反向重复(imperfect inverted repeats), 最小识别序列分别为 36 (*attB*)和 48 bp(*attP*)^[32,63](图 2(E)).

位点特异性重组酶作用机理类似于拓扑异构酶, 所以一般认为重组过程对底物构象有所要求^[65]. 在

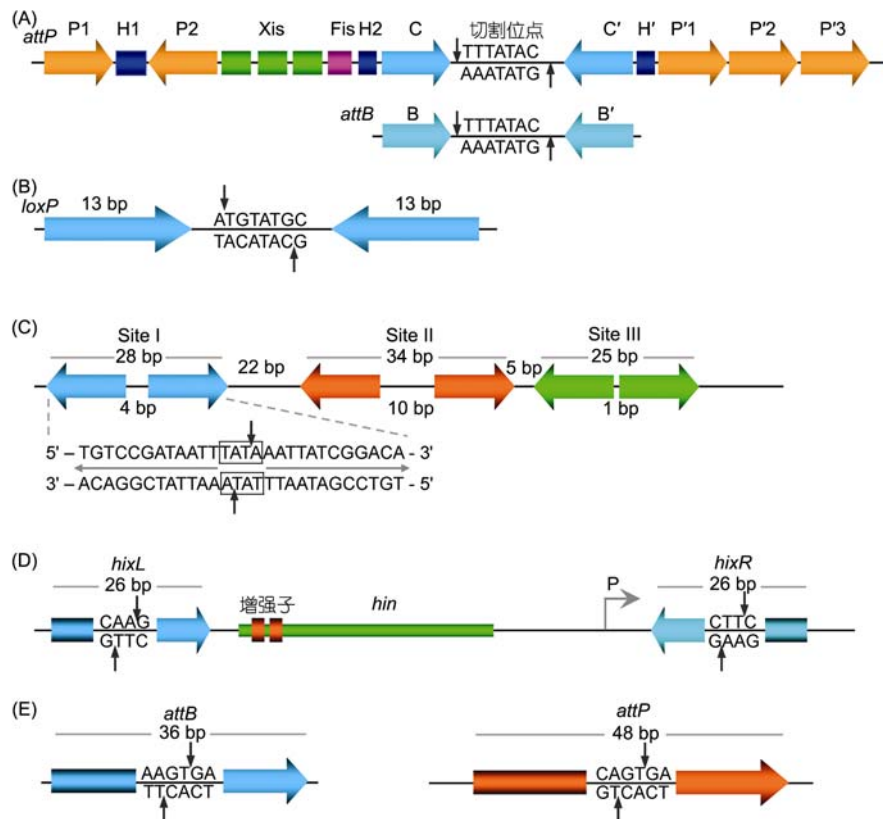


图 2 位点特异性重组酶的识别位点结构

(A) 大肠杆菌噬菌体 λ 的 *attP* 和 *attB* 位点含有 7 个核苷酸的同源序列. C, C', B 和 B' 为反向重复序列; P1, P2, P'1, P'2 和 P'3 为 Int 结合位点; 有 3 个 Xis(切离酶)结合位点, 一个 Fis 结合位点及 3 个 IHF 结合位点(H1, H2, H')^[47,64]; (B) P1 噬菌体重组酶 Cre 的识别位点为 34 bp 的反向重复序列 *loxP*^[47]; (C) 转座子 $\gamma\delta$ 解离酶的 *res* 位点结构^[20]; (D) 沙门氏菌倒位酶 Hin 介导的重组系统识别位点结构^[59-61]; (E) 链霉菌噬菌体 ϕ BT1 的最小整合底物对结构示意^[32,63]

Hin 倒位过程中, 底物的配对和倒位复合体(invertasome)的形成是分开的, 低程度的负超螺旋即可起始配对, 但是倒位复合体的形成一定需要高度的负超螺旋^[66]. 不具有超螺旋的连结(catenation)底物在 Tn3 解离酶作用下同样可以完成联会, 但是在后联会(postsynapsis)过程则需要超螺旋; 在 $\gamma\delta$ 解离中, 缺刻的连结底物(nicked catenation)反应效率约为超螺旋的 50%; 只要两位点位于超螺旋环上, Gin 倒位酶可以利用位于缺刻环上的增强子完成反应^[67]. 在 Xer 系统中, 不同程度的负超螺旋下 *cer* 位点均能重组, 但是 *mwr* 位点的重组效率随着螺旋程度的减弱而显著降低^[68]. 然而, 对噬菌体 λ , ϕ C31, ϕ BT1, Bxb1, ϕ Rv1 等整合系统的体外研究发现, 只要有合适的缓冲体系, 线性底物即能高效发生重组, 对底物的超螺旋状态完全没有要求, 虽然整合酶仍然能介导超螺旋底物的松弛^[33,35,69-71].

2.2 重组反应的分子机理

位点特异性重组都要经历 DNA 与重组酶的结合(DNA binding)(特异或非特异)、联会复合体的形成(synaptic complex formation)、链的切割(strand cleavage)、链的交换与重连(strand exchange and religation)几个主要步骤. 酪氨酸和丝氨酸重组酶利用不同的机制来完成这一过程.

(1) 酪氨酸重组酶系. 目前, 对于酪氨酸重组酶催化机理的研究经过了几十年生化水平的探索和近十几年晶体结构的分析, 对其主要过程已比较清楚^[1,47,72,73]. 重组过程起始于联会复合体中双链 DNA 在整合酶活性中心酪氨酸羟基的攻击下, 发生单链断裂, 形成共价的 DNA-蛋白质(磷酸基-酪氨酸)连接

(DNA protein phosphotyrosine linkages), 该连接位于 DNA 的 3'端; 而在 5'端留下一个游离的羟基. DNA 骨架的磷酸酯键断裂释放的能量被转移到磷酸基酪氨酸的共价连接中间体, 所以重组并不需要额外的能量参与. 随后, 游离的 5'-OH 攻击配对底物的 5'-磷酸基-酪氨酸连接, 形成 Holliday 连接(Holliday junction), 在经过异构化(isomerization)和解离最终形成重组产物. 这一过程在 Chen 等人^[72]对 Flp 的 Holliday 连接复合体的研究中及 Guo 等人^[73]对 Cre-loxP 复合物的晶体结构解析中得到了有力的验证, 但对 Holliday 中间体异构化的分子机制仍知之甚少. 2005 年, Biswas 等人^[45]解析了 λ 整合酶四聚体与臂序列和核心序列结合的不同构象的晶体结构, 为整个重组过程提供了一系列的结构信息. 在联会复合体中, 臂序列和核心序列与 λ 整合酶以不同形式结合, 形成各自特殊的排列方式, 按照严格的顺序进行 DNA 链的切割与交换. 整合酶与臂序列结合以后, 迫使 DNA 发生拓扑结构的改变, 并启动异构化形成特定构象以适应二次 DNA 链交换, 并最终完成重组过程. 这种别构调控(allosteric control)的方式成为帮助 λ 噬菌体实现特定条件下, 高效且不可逆的整合到宿主基因组的分子开关(图 3).

重组酶来源复杂, 但是对于重组结果的控制都非常精确, 控制过程本身很有意思. 简单的重组酶系统, 如 Cre-loxP 和 Flp-*frt* 系统, 只有单一一位点, 重组发生在两个完全相同的位点之间, 位点的方向即决定了是插入还是敲除; 在位点方向固定的前提下, 通过联会复合体的选择(synapsis selection)来控制反应的进一步发生, 因为错误的联会不能生成产物^[1,74]. λ 整合系统则是应用辅助蛋白因子(accessory factors)来实现

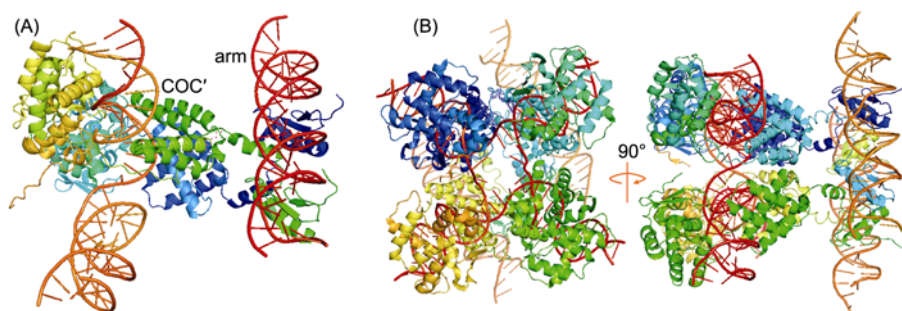


图 3 λ 整合系统对 DNA 别构调节效应^[45]

(A) 与核心序列 COC'结合的 λ 整合酶(356 aa)二聚体的晶体结构. 65~170 位的氨基酸与核心序列结合, 前 65 位与臂序列(arm site)结合, 170 位以后为催化结构域; (B) λ 整合酶四聚体与 Holliday 连接体结合的晶体结构, 包括与核心区、臂序列的结合. 晶体结构信息来源于蛋白质数据库 PDB(Protein Data Bank)(<http://www.pymol.org/>)

方向控制的典型案例. 由于 *attP* 位点结构复杂, 有不同的蛋白因子与相应的特定序列结合; 而 λ 整合酶含有与 DNA 核心序列和臂序列结合的结构域, 实现 DNA 不同构象变化以起始重组. *attB* 与 *attP* 位点在 λ 整合酶与宿主因子 IHF 存在下, 发生重组反应, 生成 *attL* 与 *attR* 位点; 当加入切离酶 Xis, 则发生逆向反应^[75]. XerCD 重组系统成员众多, 其反应结果的控制方式差异巨大, 其中也是通过辅助因子的参与. 大肠杆菌二聚体环状染色体到单体的转变需要 *dif* 位点发生 XerCD 位点特异性重组, 而 *dif* 位点的重组需要 DNA 转位酶(DNA translocase)FtsK 的激活. FtsK 识别一段 8 bp 的短序列 5'-GGGNAGGG-3', 称为 KOPS 序列; 因为 KOPS 具有方向性, FtsK 无法识别反向的 KOPS 序列. 该系统即是通过这种方式来控制其转位起始的单一方向^[76,77].

(2) 丝氨酸重组酶系. 丝氨酸家族的结构生物学研究较酪氨酸家族晚, 可能是因为丝氨酸重组酶联会迅速, 以至于联会复合体不稳定而难以检测^[1]. 已知的该家族成员催化重组的过程, 主要源自对解离酶 $\gamma\delta$, Tn3 和 Sin, 倒位酶 Hin 和 Gin 的生化及结构研究^[20-22,59,78-80]. 催化过程都是四聚体蛋白结合两个底物形成联会复合体, 在联会复合体中, 交换位点被催化域分开位于外部; 这与酪氨酸重组酶有显著区别. 联会复合体一旦形成, 4 个亚基便攻击靶位点, 形成双链断裂(double strand breakage); 丝氨酸与 5'-磷酸共价连接, 3'-OH 裸露, 并形成两个碱基的 3'突出末端. 切割是一个“反式”的偶联反应, 需要两个亚基的全力协调. 一旦重组位点被全部切割, 切开的末端即发生构象变化产生重排. 断裂末端的精确翻转方式仍然是个谜. 2005 年, Li 等人^[20]解析了共价连接的 $\gamma\delta$ 解离酶联会复合体晶体结构, 支持亚基交换的论点. 随后, 游离的 3'-OH 攻击磷酸基-丝氨酸中间体并重新连接完成重组过程.

与解离酶和倒位酶比较, 大型丝氨酸重组酶(large serine recombinase)具有更大的 C 末端结构域, 且识别位点跟前两者也有很大区别. 但目前除 TP901-1 整合酶催化结构域四聚体结构信息外^[81], 没有底物共晶或是联会复合体的结构解析. 该亚家族成员以 Bxb1 和 ϕ C31 整合酶研究最多, 都是基于生化水平研究并参考解离酶的重组模型提出^[70,82-84].

本实验室近年来从事原核生物噬菌体编码的大型丝氨酸重组酶催化反应机理及体外重组系统的应

用研究. 在对 ϕ BT1 整合酶酶学性质研究的基础上, 建立了体外位点特异性重组系统^[32], 并进一步在生化水平上对其催化反应的精细过程进行研究. 一般认为, DNA 链的切割发生在联会以后, 双链的切割需要在联会复合体内部被激活并起始. 然而, 分别对 ϕ BT1 和 ϕ C31 整合酶识别的 4 个位点(*attB*, *attP*, *attL* 和 *attR*)进行单底物切割实验发现, 在没有配对底物存在下, 链的切割同样发生, 虽然单底物切割效率很低; 这一现象与解离酶 Sin 不需要联会即可发挥催化活性的报道相吻合^[26]. 实验同时发现, ϕ BT1 整合酶可以催化单整合位点(*attB* 或 *attP*)的拓扑结构改变, 这也从另外一个侧面证明了单底物切割的发生. 另外, 在单底物切割实验中加入配对底物时, 切割效率大幅提高; 在进一步证明单一底物不能联会后, 推测链的起始切割并不需要形成联会复合体, 但是联会的形成会极大地促进双链的切割. 此外, 本实验发现高 ϕ BT1 整合酶浓度下反应被抑制的现象. 由于目前的模型中重组酶都是以四聚体形式作用, 推测该体系中大量游离四聚体与单一底物的结合限制了联会的形成. 蛋白质的体外交联检测到溶液中四聚体蛋白的存在, DNA 结合实验也证明随着酶浓度的提高, 整合酶以四聚体形式与单一底物结合的形式增加. 在 λ 系统中, 底物 *attP* 可与四聚体蛋白结合形成 intasome, 再进一步捕获 *attB* 而联会^[85]. 在高酶浓度下(150, 300 和 750 nmol/L), 对 ϕ BT1 整合系统中 *att*: Int-tetramer 结合配对底物完成重组的能力进行了测定, 发现 *attP*: Int-tetramer 可以进一步结合游离的 *attB* 进行反应, 而 *attB*: Int-tetramer 则不能. 这跟 λ 系统相吻合, 即联会的发生可以是二聚体结合单一底物随后靠拢, 也可以是 *attP* 先形成 intasome 再捕获另一底物的方式^[69]. 基于以上研究, 本实验室提出了大型丝氨酸重组酶介导位点特异性重组过程的改进模型(图 4).

丝氨酸解离酶通过识别位点的复杂性和底物的负超螺旋要求来控制反应的结果^[58,86]. 然而, 倒位酶 Hin 和 Gin 的底物结构相对简单(图 2), 反应只发生在顺式的反向重复底物间, 分子间反应检测不到, 正向重组底物也几乎不发生反应^[60]. 如此严格的控制是通过增强子(enhancer)的顺势作用来实现的, 增强子与 Fis 结合, 然后通过与倒位酶的相互作用来激活双链的切割^[87,88]. 噬菌体整合酶识别一对不同的位点, 且最小位点要求不一, 这很可能导致蛋白质与之结合方式的差别. 二聚体蛋白以不同构象分别与 4 种底物

结合(*attB*, *attP*, *attL* 和 *attR*), 而只有与 *attB* 和 *attP* 结合才能形成稳定的联会复合体^[69,84,89]. 在 Bxb1 和 ϕ Rv1 系统中, 切离酶 Xis 的参与能使 *attL/attR* 之间形成稳定联会并抑制 *attB/attP* 之间联会^[35,90,91], 从而实现反应方向的严格控制.

2.3 重组酶的结构域与杂合酶改造

作过对 105 种酪氨酸家族重组酶的氨基酸序列比对^[92]和一些突变分析, 在催化结构域鉴定出 5 个高度保守的活性位点 RKHRH, 均位于催化酪氨酸之前, 一直被认为是催化反应所需关键位点^[1,42](图 5(A)). 然而, 新近对于 Cre 重组酶高度保守位点的突变研究发现, 大部分高度保守位点被替换后仍然具有很高的重组活性, 包括之前被认为是催化高效重组所必需的一个精氨酸残基(Arg), 而这些位点在拓扑异构酶 TopIB 家族成员中都是高度保守且必需的^[93].

大型丝氨酸重组酶系统十分特殊, 位点简单却高度特异, 不论是在机理研究或是系统开发应用上都具有很高的价值. 目前, 已有大量的新基因被注释为大型丝氨酸重组酶, 但完成功能鉴定的却只有少数几个. 对 81 种已被研究和注释为大型丝氨酸重组酶的全长氨基酸序列进行了系统发育分析(图 6(A)). 按来源和功能可大致分为三类: 噬菌体编码的整合系统(如 ϕ C31, Bxb1, ϕ FC1 和 TP901-1)、转座子(如 Tn4451 编码的 TnpX^[39]和 Tn5397 编码的 TndX^[94])以及基因组编码相关基因(如孢子发育过程中介导基因重排的 *SpoIVCA*^[95]). 鉴于大型丝氨酸重组酶全长序

列的同源性较低, 而 N 端的催化结构域相对较高, 进一步对其催化结构域做了系统发育分析(图 6(B)). 总体的分类没有显著差别, 但之间的发育关系则有所区分. 比较了 81 种大型丝氨酸重组酶的氨基酸序列(450~620 aa), 发现其活性中心有 3 个保守的 YRS; 催化结构域另外几个高度保守的氨基酸为 RPRR; 在其 DNA 结合结构域, 高度保守的氨基酸为 GGYLGCCG. 目前, 除催化中心外, 未见对这些高度保守位点突变分析的报道(图 5(A)). Rowley 和 Smith^[96]对 ϕ C31 整合酶进行突变分析后发现, N 端的 V129 对联会的形成及联会后切割所需正确构象的形成至关重要, 位于 C 端的一个 coiled-coil 区域负责控制反应方向, 该区域的 E449 突变为赖氨酸 K 后, 该酶完全丧失了对反应方向的严格控制^[97]. 需要更多位点的突变分析, 才能了解大型丝氨酸重组酶各个关键位点在整个催化过程中的确切功能.

位点特异性重组酶家族庞大, 不同系统的识别位点和各自对反应方向控制的方式多样而互有优劣, 所以构建整合不同酶系统优点的杂合酶成为该类酶分子改造的重要方向. 近期, 锌指核酸酶(Zinc finger-nucleases)在基因打靶操作中显示出巨大的优势^[98,99], 是核酸酶的催化结构域和锌指 DNA 结合结构域的杂合酶. 已有不少工作利用丝氨酸重组酶(Hin, Gin 和 Tn3)的催化结构域做为筛选对象, 且已在人类细胞中得到了很好地应用^[100]. Cre 重组酶应用广泛, 但是反应方向不能控制. Warren 等人^[48]利用 λ 整合酶能够严格实现定向的特点, 构建了杂合的 Cre 重组酶和整

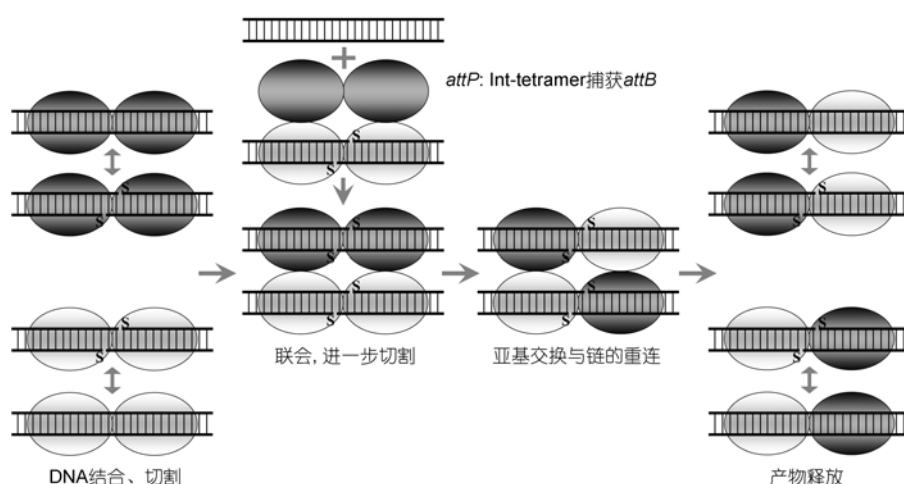


图4 大型丝氨酸重组酶介导重组反应的改进模型

合位点. 该系统赋予了 Cre 精确可控的特性, 为进一步利用 Cre 重组酶打下了良好的基础(图 5(B)). 为实现 Cre 重组酶的可控表达, Metzger 等人^[101]构建了杂合的 Cre 重组酶, 在其 C 端融合了一个雌激素受体 (estrogen receptor, ER) 的配体结合结构域 (ligand-binding domain, LBD), 通过添加雌激素拮抗剂来诱导活化 Cre 重组酶. 采用类似方式, 还实现了在小鼠肝脏、脂肪细胞等的特异激活^[101-104].

3 位点特异性重组系统的应用

作为遗传操作中一种极为有效的工具酶, 一些位点特异性重组 (SSR) 及其衍生系统受到了很大的关注, 其中使用最为广泛的是 P1 噬菌体 Cre-loxP 系统、酵母 2 μ 质粒 Flp-*frt* 系统和链霉菌噬菌体 ϕ C31 整合系统. 这些系统各有优劣, 并在不断地完善中, 同时新的系统也在不断地开发及应用 (如 Bxb1 和 ϕ BT1 整合系统). 随着大量关于位点特异性重组文献的涌现, 乔治亚大学 Urbanski 和 Condie^[105]开发了一个用于挖掘位点特异性重组 (SSR) 文献资料的在线服务器 (<http://ssrc.genetics.uga.edu/>). 该服务器可集中搜索目前超过 9000 篇关于 SSR 系统的结构功能研究以及在细菌、果蝇、斑马鱼、干细胞及拟南芥中的应用研究实例.

3.1 位点特异性重组酶做为工具酶的体外应用

位点特异性重组酶的生化研究显示, λ 整合系统和 Cre-loxP 系统在体外操作简便, 只需要纯化的重组酶 (及相关辅助因子) 和适当的缓冲条件即可高效催化反应, 对底物构象也没有要求. 各大公司争相开发

了基于 λ 整合和 Cre-loxP 重组的克隆及表达系统. 主要有 Invitrogen 公司 Gateway 系统^[71]、Clontech 公司 Creator 系统^[106]以及 Stephen Elledge 实验室开发的 Univector 系统 (也称 Echo 系统, 属 Invitrogen 公司)^[107].

Gateway 系统基于 λ 整合, 在 Int 和 IHF 的作用下 *attB* 和 *attP* 发生重组, 再加入 Xis 即发生 *attL* 和 *attR* 之间的逆向重组. 利用这一原理和一对突变的位点, Hartley 等人^[71]设计了通过两步克隆高效快速构建表达载体的 Gateway 系统. 第一步构建入门克隆 (Entry Clone), PCR 扩增目的基因在两侧分别带上 25 bp 的 *attB*₁ 和 *attB*₂, 通过正向反应定位克隆到入门载体 (Entry Vector); 第二步以逆向反应将目的基因重组到系列目的载体 (Destination Vector). 该系统设计巧妙, 载体上都带有自杀基因 *ccdB*, 宿主只有正确重组克隆转入后才能生长, 方便筛选; Invitrogen 更是整合其在表达系统方面的优势, 为 Gateway 系统设计了一系列的配套表达载体. 很多研究工作也不断报道新的改装载体以方便与 Gateway 系统结合使用^[108-112]. Creator 系统跟 Gateway 系统类似, 区别在于它使用的是 Cre-loxP 重组系统. Clontech 公司为该系统设计了几十种配套表达载体, 还将其引入著名的 SMART 建库试剂盒中, 这都为该系统的广泛应用大开绿灯. 这几种技术的开发为科研工作者进行高通量的基因功能研究提供了更多的选择^[113].

本实验室在建立 ϕ BT1 体外重组系统和筛选突变底物对的基础上^[32,69], 开发了用于重组克隆与快速基因打靶载体构建的新方法. 前者类似于 Gateway 系统, 可方便地实现目的基因的定向克隆; 后者在一系列载体系统的基础上, 可在体外将上下游同源臂片

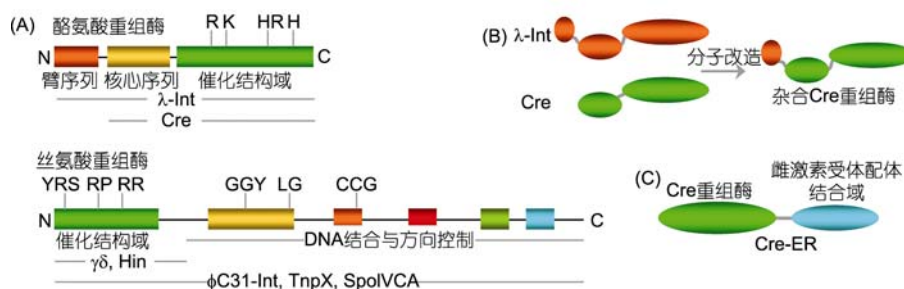


图 5 位点重组酶的结构域与杂合酶分子改造

(A) 酪氨酸和丝氨酸重组酶的结构组成, 保守氨基酸已标出; (B) 杂合酶改造示例, 加入 λ -Int 臂序列结合结构域, Cre 获得了方向控制的特性; (C) 融合了雌激素受体 ER 配体结合域后, Cre 重组酶活性在哺乳动物细胞可调控

段、筛选标记基因和骨架载体经一步反应串联重组, 精确高效地构建基因打靶载体. 相关技术已申请专利.

3.2 位点特异性重组酶在体内遗传操作中的应用

按照作用方式, 位点特异性重组系统在体内(*in*

vivo)遗传操作中的应用可分为三类: 外源基因的定点整合; 特定基因的删除、置换或倒位; 染色体工程中大片段 DNA 的缺失、倒位或易位. 相比同源重组, SSR 系统最大的优势是高效率; 对随机的转座系统而言, 其最大的优势是靶位点的专一. 每种方式都有其不可替代性, 目前很多研究工作则是综合几种系

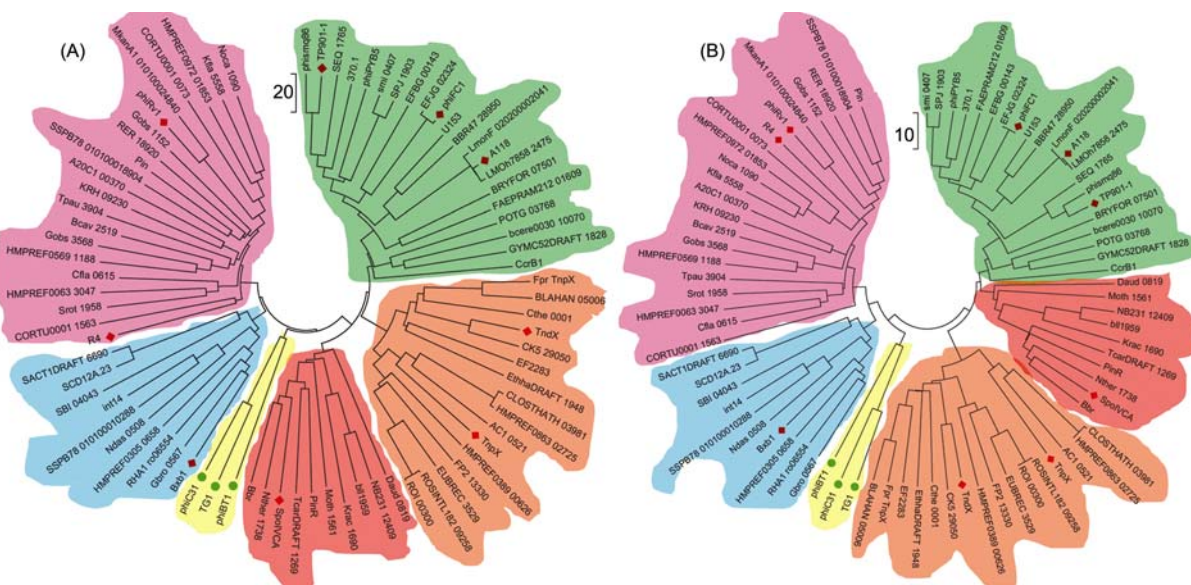


图 6 对 81 种大型丝氨酸重组酶的全长(A)和催化结构域(B)氨基酸序列的系统发育分析

系列比对使用 CLUSTALW 软件, NJ 树的绘制使用 MEGA 软件(www.megasoftware.net). 使用蛋白质序列的来源和 NCBI 登录号依次为: *Actinomyces sp.* (HMPREF0972_01853, ZP_06163058.1); *Aeromicrobium marinum* (HMPREF0063_3047, ZP_06607663.1); *Bacillus cereus* (bcere0030_10070, ZP_04173374.1); *Bacillus subtilis* (SpoIVCA, YP_002771411); *Beutenbergia cavernae* (Bcav_2519, YP_003410527); *Blautia hansenii* (BLAHAN_05006, ZP_05853856); *Bradyrhizobium japonicum* (bll1959, NP_768599); *Brevibacillus brevis* (Bbr CisA and BR47_28950, YP_002771411.1 和 YP_002772376.1); *Bryantella formatexigens* (BRYFOR_07501, ZP_05346707.1); *Carnobacterium sp.* (370.1, ZP_02184702.1); *Cellulomonas flavigena* (Cfla_0615, YP_003635726.1); *Clostridium difficile* transposons Tn4451 (TnpX, ZP_05616494.1) 和 Tn5397 (TndX, AAF35174.1); *Clostridium hathewayi* (CLOTHATH_03981, ZP_06115669); *Clostridium perfringens* (AC1_0521, ZP_02634597); *Clostridium thermocellum* (Cthe_0001, YP_001036436); *Corynebacterium pseudogenitalium* (HMPREF0305_0658, ZP_03920469); *Corynebacterium tuberculostearicum* (CORTU0001_0073 和 CORTU0001_1563, ZP_05366632 和 ZP_05366267); *Desulfurudis audaxviator* (Daud_0819, YP_001716966); *Enterococcus faecalis* (EFJG_02324, EFBG_00143 和 phage ϕ FC1, ZP_05575004, ZP_05425497 和 AF124258_1); *Erysipelotrichaceae bacterium* (HMPREF0863_02725, ZP_06646584); *Ethanoligenens harbinense* (EthhaDRAFT_1948, ZP_06471870); *Eubacterium rectale* (EUBREC_3529, YP_002939389); *Faecalibacterium prausnitzii* (FAEPRAM212_01609 和 FP2_13330, ZP_02091337.1 和 CBK98844.1); *Filifactor alocis* (HMPREF0389_00626, ZP_06572821); *Geobacillus sp.* (GYMC52DRAFT_1828, ZP_04392678); *Geodermaphilus obscurus* (Gobs_1152 和 Gobs_3568, YP_003408281.1 和 YP_003410527.1); *Gordonia bronchialis* (Gbro_0567, YP_003271790); *Kocuria rhizophila* (KRH_09230, YP_001854776); *Kribbella flavida* (Kfla_5558, YP_003383363); *Ktedonobacter racemifer* (Krac_1690, EFH81025.1); *Lactococcal phage TP901-1* (CAA59475.1); *Listeria monocytogenes* (LmonF_020200002041, LMOh7858_2475, phage A118 和 U153, ZP_05289043, ZP_00230221, NP_463492.1 和 CAD10283.1); *Marine actinobacterium* (A20C1_00370, ZP_01131007); *Micrococcus luteus* (HMPREF0569_1188, ZP_06501013); *Moorella thermoacetica* (Moth_1561, YP_430413.1); *Mycobacterium phage Bxb1* and prophage ϕ Rv1 (AAG59740.1 和 NP_216102.1); *Mycobacterium kansasii* (MkanA1_010100024840, ZP_04751223.1); *Natranaerobius thermophilus* (Nther_1738, YP_001917900.1); *Nitrococcus mobilis* (NB231_12409, ZP_01126794.1); *Nocardioideis sp.* (Noca_1090, YP_922294.1); *Nocardioopsis dassonvillei* (Ndas_0508, YP_003678461.1); *Paenibacillus sp.* (POTG_03768, ZP_04854347.1); *Pelotomaculum thermopropionicum* (PinR, YP_001212789.1); *Rhodococcus erythropolis* (RER_18920, YP_002765339.1); *Rhodococcus jostii* (RHA1_ro06554, YP_706485.1); *Roseburia intestinalis* (ROI_00300 和 ROSINTL182_09258, CBL07393.1 和 ZP_04745837.1); *Ruminococcus obeum* (CK5_29050, CBL24160.1); *Segniliparus rotundus* (Srot_1958, YP_003659246.1); *Staphylococcus hominis* (CcrB1, BAB83487.2); *Streptococcus equi* (SEQ_1765, YP_002747001.1); *Streptococcus mitis* (smi_0407, YP_003445547.1); *Streptococcus pneumoniae* (CGSSp14BS69_06672 和 SPJ_1903, ZP_01828805.1 和 YP_002736920.1); *Streptomyces avermitilis* (int14, NP_826732.1); *Streptomyces bingchengensis* (SBI_04043, ADI07164.1); *Streptomyces coelicolor* (SCD12A.23, NP_628511.1); *Streptomyces sp.* (SSPB78_010100010288, SSPB78_010100018904 和 SACT1DRAFT_6690, ZP_05487022.1, ZP_05488708.1 和 ZP_06282665.1); *Streptomyces sp.* phage ϕ BT1, ϕ C31, TG1 和 R4 (NP_813744.2, NP_047974.1, BAF03600.1 和 BAA07372.1); *Streptosporangium roseum* (Pin invertase, ZP_05487022); *Thermosinus carboxydivorans* (TcarDRAFT_1269, ZP_01666637.1); *Tsukamurella paurometabola* (Tpau_3904, YP_003648815.1)

统的优点来实现期望的遗传操作^[114,115].

(1) 微生物的遗传改造. 大多数位点特异性重组系统来源于微生物, 相应的系统在微生物的遗传改造中应用也非常广泛. 基于 ϕ C31 和 ϕ BT1 系统构建的整合载体 pSET152 和 pRT802 已成为链霉菌遗传互补研究的常用载体系统^[63]. Watson 等人^[116]将基于 Cre 重组酶和突变底物的基因标记(gene tagging)与重组酶介导的盒式交换(recombinase-mediated cassette exchange, RMCE) 系统用于裂殖酵母 (*Schizosaccharomyces pombe*). 在基因组改造方面, 日本北里大学北里生命科学研究所于 2003 年测定了除虫链霉菌(*Streptomyces avermitilis*)全基因组^[117]. 其基因组大小为 9.02 Mb, 通过与天蓝色链霉菌(*S. coelicolor*)和灰色链霉菌(*S. griseus*)的比较基因组分析显示, 约 6.28~6.5 Mb 的基因组区域保守, 包括生长代谢必需基因. 为了构建用于异源表达次生代谢产物生物合成基因簇的通用菌株, 他们拟敲除约 1.4 Mb 大小的非保守区. 该大片基因组 DNA 的敲除采用了 Cre 介导的位点特异性重组, 先以同源重组在其两侧放置一对同向的 *loxP* 位点, 后表达 Cre 重组酶即可敲除该 1.4 Mb 的大片段. 该方法非常高效, 挑选的 24 个克隆均实现了成功敲除(图 7(A))^[114].

(2) 植物基因工程. 在植物遗传转化过程中出

现的一个重要问题是转基因沉默(transgene silencing)和位置效应(position effect), 导致外源基因在不同植株中的表达水平有很大差异. 筛选位于转录活跃区的整合位点以此为基础建立相应的定点整合系统, 不仅有利于消除位置效应, 而且可以提高转基因的安全性. 此外, 在植物转基因操作中, 同源重组发生概率很低, 所以需要借助筛选标记来提高实验效率, 而筛选标记本身造成植物代谢负担的同时, 也增加了转基因的风险; 基于位点特异性重组的筛选标记移除系统很好地解决了这个问题. 第一例批准商业应用的转基因玉米 LY038 即是通过 Cre 重组酶去除了卡那霉素抗性基因(*npt II*)^[118].

Hu 等人^[119]评估了 FLP 重组酶用于转基因水稻的可行性. 首先构建一株带有卡那霉素抗性基因(*neomycin phosphotransferase II*, *npt II*)的突变株, 抗性基因两侧带有 *frt* 位点, 该区段把谷类泛素启动子和 *gusA* 基因分开而关闭报告基因的活性. 将该植株与表达 FLP 重组酶的植株杂交, 然后检测 *npt II* 的丢失和 *gusA* 的表达. 在杂交系中检测出 *GUS* 基因表达, 而自交系未检测到 *GUS* 活性. 重组后的区段遗传稳定, 说明 FLP-*frt* 系统可用于水稻的性状改良研究. Fladung 和 Becker^[120]测试了 Cre-*loxP* 和 FLP-*frt* 系统在杨树遗传改良研究中进行筛选标记移除和目的基

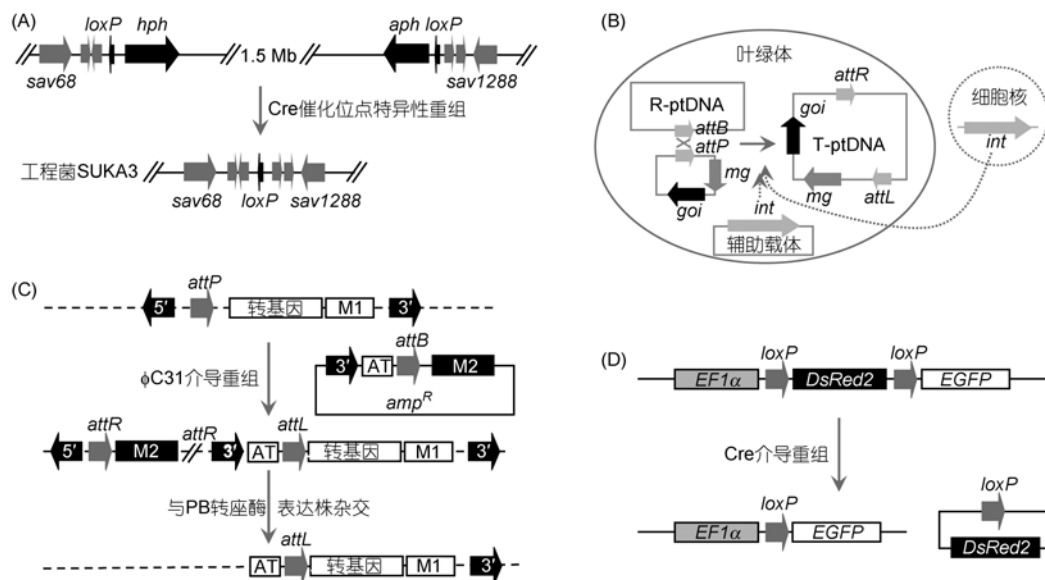


图 7 位点特异性重组系统在转基因中的应用示例^[64,114,115,131]

(A) 利用 Cre 重组酶对 *S. avermitilis* 进行基因组大片段删除; (B) ϕ C31 整合系统用于植物叶绿体转化; (C) 结合 piggyBac 随机转座和 ϕ C31 整合系统, 通过三步法获得稳定遗传的转基因果蝇突变株; (D) 他莫昔芬(TAM)激活 Cre 重组酶介导斑马鱼体内基因的高效删除

因转移中的可行性. 将带有重组酶基因和重组位点的质粒转到欧洲山杨(*Populus tremula*)和北美颤杨(*Populus tremuloides*)杂交系, 重组酶的表达受热激活启动子 HSP(来源于大豆)控制; 热激活后, 重组酶表达而将位点之间区域切除, 致使 GUS 表达. 实验证明, Cre-loxP 和 FLP-*frt* 系统可以在杨树中介导高效的筛选标记移除和进行精确的定点基因转移. Rubtsova 等人^[121]把 ϕ C31 整合酶应用于小麦(*Triticum aestivum*)的转基因研究. 他们通过 T-DNA 把 ϕ C31 整合酶基因转移到小麦中, 然后以基于植物病毒的系统来监视重组酶在体内的表达情况, 顺利的建立了几个双单倍体(DH)自交系. 该自交系组成型表达整合酶基因, 而对于植株的生长和发育没有显著的不利影响. 目前, 可用于转基因植物的位点特异性重组系统很多, 所实现的目的更是多种多样. Birchler 等人^[122]完成了利用 Cre 重组酶对植物微型染色体和人工染色体进行改造.

叶绿体转化由于表达效率高、遗传稳定以及可有效控制基因污染等优势, 成为核转化之后一项非常有效的植物基因工程操作技术. 目前, 对叶绿体基因改造主要通过同源重组来实现, Lutz 等人^[64]将 ϕ C31 整合系统引入进来作为可选策略. 先以同源重组在叶绿体基因组中插入一个 *attB* 位点, 以此位点作为转基因靶点; 然后将含有 *attP* 位点和外源基因的质粒转化到叶绿体. 而 ϕ C31 整合酶的表达或通过核染色体编码或共转化一个游离质粒到叶绿体来实现(图 7(B))^[64,123,124]. 另外, 叶绿体转化经抗性筛选后, 出于安全性和减少因为高拷贝数带来的代谢负担考虑, Kittiwongwattana 等人^[124]设计了基于 ϕ C31 的筛选标记切除系统.

(3) 寄生虫研究. 寄生虫感染一直是人类健康的重大威胁, 它与儿童发育不良、多系统疾病与病毒感染等关系密切. 首要措施在于预防, 而对于寄生虫本身的基因功能研究能为人们了解其发病及流行机制、更好地进行预防和治疗提供科学依据. 有限的可用药物筛选标记已成为布鲁斯锥虫(*Trypanosoma brucei*)遗传操作的重要限制因素. Scahill 等人^[125]利用 Cre 重组系统开发了一种可去除抗性的方法. 更昔洛韦(ganciclovir)在进行转基因筛选后, 以四环素诱导的启动子表达 Cre 重组酶而将其切除. 研究还发现, 携带 *cre* 基因的质粒瞬时表达即可高效去除筛选基因. 他们用这种方法敲除了嘧啶合成途径中的两个关

酶, 构建了一个具有更昔洛韦抗性的细胞株. Davis 等人^[126]利用红色荧光蛋白和绿色荧光蛋白双报告系统评估了在秀丽隐杆线虫(*C. elegans*)中进行基因敲除的可行性. Nkrumah 等人^[127]在恶性疟原虫(*Plasmodium falciparum*)*cg6* 基因中引入了一个分枝杆菌噬菌体 Bxb1 识别的 *attB* 位点, 该突变的引入对疟原虫无性血内阶段生长(asexual blood stage growth)是无害的. 随后共转染含有 *attP* 位点和瞬时表达 Bxb1 整合酶的质粒, 转染后的质粒以单拷贝插入, 在观察的一个月时间内, 其基因型稳定, 疟原虫群体表型均一, 除去抗性压力后仍遗传稳定. 该系统可用于恶性疟原虫生活史各阶段快速基因整合插入和互补研究, 首次显示了基于 Bxb1 的整合系统可用于真核生物的胞内研究. Combe 等人^[128]报道了一种在伯氏疟原虫(*Plasmodium berghei*)中利用 FLP-*frt* 系统对基因进行条件性失活的方法. 他们在 *MSP1* 基因(该基因对于疟原虫裂殖子入侵红细胞十分关键)两侧插入了 *frt* 位点, 并以在蚊中肠孢子小体和唾液腺孢子小体中特异表达的启动子控制 Flp 蛋白和其热敏感突变体 FlpL 的表达, 结果成功敲除了该基因, 随后在肝脏中发现裂殖子的形成受损.

(4) 昆虫转基因. 昆虫的转基因主要基于随机的插入转座来进行, 一旦某个特定位置的插入被证明能很好地用于外源基因的表达, 该位点即可用于后续的转基因研究. 家蚕作为鳞翅目昆虫(农林业中的主要害虫)的模式生物, 是理想的生物反应器, 具有重要的理论研究和经济价值. Nakayama 等人^[129]把 ϕ C31 整合系统应用到家蚕 BmN4 细胞系中. 他们首先以 Gateway 系统构建了 3 个整合酶表达辅助质粒, 即分别含有野生型和在其 N 或 C 端融合了核定位信号序列(NLS)的 ϕ C31-*int* 基因; 然后设计了一套大肠杆菌质粒融合检测系统(extrachromosomal gene exchange). 在分析的 144 个克隆中, 有 140 个发生了重组, 并且核定位信号序列的添加对 ϕ C31 整合酶活性无显著影响. 将此系统应用到家蚕 BmN4 细胞, 得到一致的结果, 说明 ϕ C31 整合酶在家蚕细胞中能充分发挥活性. 随后, 他们将一对 *attB* 位点通过 PiggyBac 转座子插入到家蚕基因组, 随后将 DsRED 报告基因两侧含有一对 *attP* 位点的整合质粒转染入细胞. 整合发生后, 报告基因表达. 结果显示, ϕ C31 整合酶能高效介导外源基因与基因组的整合, 且 N 端添加 NLS 后, 重组效率显著提高. 虽然之前 Flp 系

统已在家蚕细胞中得到应用,但是 ϕ C31 系统的单向性无疑使其成为鳞翅目研究更好的选择. 同样, Schetelig 等人^[115]开发了用于地中海实蝇(*Ceratitis capitata*)的结合 *piggyBac* 随机转座和 ϕ C31 整合系统获得稳定遗传突变株的一种拓展策略. 通过 *piggyBac* 随机转座将第一个目的基因、报告基因和 ϕ C31 的 *attP* 位点插入到基因组;然后将带有第 2 个目的基因、报告基因、 ϕ C31 的 *attB* 位点、PB 转座酶识别的 3' 序列和 ϕ C31 整合酶的 mRNA 共转染到前一突变株;随后,将该突变株与编码转座酶的果蝇交配,即可将对应部分切除而获得稳定的转基因突变株. 该策略可以在一个特定的位点插入多个基因并获得稳定遗传,极大地方便了对果蝇突变和基因功能的研究(图 7(C)).

(5) 斑马鱼. 斑马鱼(Zebra fish)作为多种系统发育、功能和疾病研究的模式生物,很早就有利用 Cre 重组酶对其进行基因操作的报道,然而这些工作中,Cre 介导重组效率相对较低^[130]. 为解决这一问题, Hans 等人^[131]通过 Tol2 转座子将报告系统导入斑马鱼,该报告系统由红色荧光蛋白 DsRed2(两侧带有 *loxP* 位点)和绿色荧光蛋白 EGFP 组成,由非洲爪蟾延伸因子 1 α (*EF1 α*)启动子控制,正常情况下 DsRed2 表达,而当 Cre 发挥作用时, DsRed2 基因被切除而表达 EGFP. 他们先将该品系与 *Tg(hsp70: EGFP-Cre)* 品系(受温度诱导启动子 *hsp70* 控制)杂交,发现该系统不仅重组效率低且启动子不受严格调控. 随后他们构建了两个品系, *Tg(pax2a: CreER^{T2})^{#19}* 和 *Tg(pax2a: CreER^{T2})^{#45}*, 前一品系杂合重组酶 CreER^{T2} 在发育中前脑的间脑表达,后一品系在后脑部的原基菱脑原节 3 和 5 表达,且都受斑马鱼 *pax2a* 启动子控制. 将这两个品系与报告基因品系杂交,后添加他莫昔芬(tamoxifen, TAM)或 4-羟基-他莫昔芬(4-hydroxy-tamoxifen, 4-OHT),结果发现在发育中的斑马鱼胚胎中配体介导的重组不仅高效而且异常迅速(添加 TAM 4 h, 4-OHT 2 h 后即可检测到). 另外,由于他莫昔芬的剂量和添加时间很容易调节,所以该系统可以完成精确的预期调控(图 7(D)). 与该系统相似, Lister^[132]把 ϕ C31 整合系统应用到斑马鱼胚胎的分子内重组. 结果表明, ϕ C31 整合酶在斑马鱼细胞和胚胎中具有功能;跟 Cre-*loxP* 系统一样, ϕ C31 系统可以切除两端带有 *attB* 和 *attP* 位点的基因片段;当融合有雌激素受体(ER)的配体结合域后,该杂合酶仍

然具有依赖于 4-OHT 的催化活性,虽然活性略低于野生型.

(6) 哺乳动物与基因治疗载体. 自 20 世纪 80 年代末,科学家便开始利用 Cre 重组酶对基因组大片段进行切除或者倒位. 发展到现在, Cre-*loxP* 系统已经成为染色体改造的利器. 例如,在小鼠中实现可诱导的染色体易位(chromosome translocation)和体细胞性改变(somatic alterations),结合干细胞培养和同源重组技术,科学家可以在任意需要的位置进行基因组修饰^[133].

Tiscornia 等人^[134]开发了一种基于慢病毒载体受 Cre 重组酶诱导的可控 RNA 干扰系统. 发夹靶标序列以小鼠 U6 启动子诱导,但是被一段随机的填充序列分开,并带有一对 *loxP* 位点,只有在 Cre 重组酶表达的情况下,发夹靶标序列才得以转录从而起始 RNA 干扰. Wu 等人^[135]报道了一种结合了 Cre-*loxP* 和 *piggyBac* 转座系统的对小鼠进行基因组范围突变的方法,该系统简便快捷,可以系统性地对编码区和非编码区进行基因功能分析. 以 *piggyBac* 转座子将 *loxP* 位点随机分布于基因组产生突变,然后通过杂交和以 Cre 重组酶的作用实现基因组大片段的缺失或重复(频率为 5.5%~21.6%),甚至可以产生非同源染色体的易位(频率超过 1%). 为克服转座的随机插入在人胚胎干细胞转基因操作中的缺陷, Sakurai 等人^[136]报道了结合同源重组和 Cre 位点特异性重组的定点整合策略. 以同源重组在人 ES 细胞株 KhES-1 的次黄嘌呤磷酸核糖转移酶 1(HPRT)基因座插入抗性基因并携带 *loxP* 位点,随后通过 Cre 重组酶把携带靶基因的质粒整合到 *loxP* 位点. 报告基因显示,该系统特异且高效,转基因的 ES 细胞仍然保持全能性. 虽然目前用于哺乳动物细胞的重组系统较多,但是 Cre 最成熟,效率最高. Raymond 和 Soriano^[137]按照哺乳动物细胞偏好优化了 FLP 和 ϕ C31 整合酶的密码子,命名为 FLPo 和 ϕ C31o,并利用重组酶依赖性表达半乳糖苷酶 ROSA26 报告系统检测其在哺乳动物干细胞中的活性,发现优化后的 FLPo 和 ϕ C31o 具有与 Cre 重组酶对等的重组效率,极大地拓展了 FLP-*frt* 和 ϕ C31 系统在哺乳动物细胞中的应用.

安全性是基因治疗首要的考虑因素. 基因治疗载体有整合型和非整合型两类. 整合型载体有逆转录病毒载体、腺相关病毒(AAV)载体和转座子载体,这类载体都具有高效率且持续表达等优点,然而这

三种载体的整合都是相对随机事件, 这就大大增加了安全隐患; 腺病毒载体为非整合载体, 其优点是可携带大片段外源基因且相对安全, 但是其免疫原性较强, 细胞毒性大. 基于噬菌体位点特异性重组酶的基因治疗载体可携带大片段 DNA 单向定点整合到基因组, 可持续表达且遗传稳定, 可能是基因治疗的理想载体.

斯坦福大学医学院(Stanford University School of Medicine)Michele P. Calos 实验室在利用 ϕ C31 整合系统进行基因治疗方面做出了系列开创性的工作. 2000 年, 他们首次将 ϕ C31 整合酶用于人细胞并获得成功^[138]. 2003 年, 借助 ϕ C31 整合酶, 他们将人凝血因子 IX(human Factor IX, hFIX)基因整合到小鼠基因组, 而 hFIX 的缺失导致血友病 B. 带有 hFIX 基因和野生型 *attB* 位点的质粒通过高压尾静脉注射并整合于小鼠肝脏, 当与整合酶表达质粒共同注射时, 血清中 hFIX 的水平增加了 10 倍, 约 4 $\mu\text{g/mL}$, 与正常水平相当, 并且在实验检测的 8 个月里保持稳定. 部分切除肝脏后, hFIX 水平无变化, 说明携带 hFIX 基因的质粒确实整合到基因组. 进一步分析整合位点和频率, 发现整合发生于两个假 *attP* 位点(pseudo-*attP*), 且其中之一占主导. 该项工作做为 ϕ C31 系统用于动物模型的首例报道引起了广泛的关注^[139]. 紧接着, 他们与同校皮肤病学 Paul Khavari 实验室合作, 利用该系统治疗隐性营养不良性大疱性表皮松解症(recessive dystrophic epidermolysis bullosa, RDEB). RDEB 由于 *COL7A1* 基因突变引起, *COL7A1* 全长 32 kb, 转录并加工后产生 8.9 kb 的 mRNA, 编码 2944 aa 的 VII 型胶原蛋白, 该基因突变后引发水疱. ϕ C31 整合酶将 *COL7A1* cDNA 导入 4 个独立 RDEB 病例的原初表皮祖细胞稳定整合, 利用这些细胞进行皮肤再生后, 典型的 RDEB 症状得到修复, 包括 VII 型胶原蛋白的表达、固定纤维的形成和表皮真皮的内聚. 这为利用非病毒载体治疗人类遗传紊乱提供了一个可实践的方法^[140]. 随后, 他们将 ϕ C31 系统用于尝试治疗交界型大疱性表皮松解症(junctional epidermolysis bullosa, JEB)^[141,142]和进行性肌营养不良症(duchenne muscular dystrophy, DMD)^[143,144].

3.3 多重重组及多种重组酶的联合应用

基于不兼容底物对筛选的多重重组与多个重组酶系统的联合应用, 使得多基因的重组成为可能. 多

位点的 Gateway 重组克隆系统可以将多个 DNA 片段按照特定的顺序、方向和框架排列起来. 为了利用这套系统研究植物中的某些功能基因, Karimi 等人^[145]构建了 36 个入门克隆, 这些克隆包括了启动子、终止子、报告基因及 LhG4/LhGR 双组份调控原件. 该系统可以和常规的克隆系统配合使用, 同时还可以将位于不同表达模块上的 2~3 个基因同时克隆到一个载体上. 他们用一系列报告基因验证了该系统的可行性.

果蝇作为研究基因调控和发育过程的模式生物, 对其遗传操作手段的快速更新将极大促进相关研究的进行. Huang 等人^[146]结合 Cre 重组酶和 ϕ C31 整合酶系统, 发展了对果蝇基因组高效打靶的“两步法”. 首先, 通过同源重组将 ϕ C31 的 *attB* 位点插入到靶序列, 并在抗性基因两侧带有 *loxP* 位点, 使用 Cre 即可去除抗性并留下一个 *loxP* 位点; 第二步, 利用 ϕ C31 系统将靶基因定点整合入 *attB* 位点, 并带有另外一个 *loxP* 位点, 即可再次去除抗性基因. 通过该方法, 已经成功改造了 6 个基因并在相应的座位上整合了 70 个独立的等位基因. Cre-*loxP* 系统的反应方向和专一性可以通过突变的 *loxP* 位点来实现. Kameyama 等人^[147]设计了一套可多次插入靶基因的系统. 他们首先在体外证明了不同突变位点反应的可能性, 随后在中国仓鼠卵巢细胞(CHO)实现了多个基因的连续插入.

孟山都(Monsanto)公司的高赖氨酸抗虫转基因玉米品种 LY038, 作为第一例高营养转基因作物在很多国家被允许商业应用. 虽然转基因作物一直饱受争议, 但是 LY038 却是位点特异性重组系统(Cre)应用的一个著名案例^[148]. 在区永祥(David W. Ow)的一篇综述^[118]中, 介绍了从 LY038 出发, 联合利用多种位点特异性重组系统进行改造的策略. 在去除卡那霉素抗性基因(*npt II*)后, LY038 高赖氨酸基因 *cordaA* 一侧留下 *loxP* 位点, 可进行第二轮转基因; 第二轮转基因时, 带上 ϕ C31 整合位点, 并以此位点进行第三轮转基因; 第三轮又引入 ParA 的整合位点, 如此反复, 第四轮再引入其他整合系统的位点(如 Bxb1 整合酶). 这样即可反复进行多基因的插入, 通过巧妙的设计而不留多余序列; 另外, 在第一轮利用 Cre 系统时外侧还引入一个 *loxP* 位点, 这样结合初始的 *loxP* 位点即可方便地把后来的转基因都去除^[118].

3.4 位点特异性重组系统体内应用存在的一些问题

位点特异性重组系统为很多物种的遗传操作扫除了大量障碍,但随着该系统的广泛应用,越来越多的问题被发现.一方面,由于众多假位点的存在,使得基因组异常重组的风险增加;另一方面,得益于该类酶系统的高活性,使得在大量的基因组序列背景下,位点特异性重组似乎不再那么“特异”,出现类似于体外高浓度限制性内切酶反应时一样的“星号活性”,给该类酶的高安全性打上了一个问号.

Schmidt 等人^[149]制造了以 *Protamine 1 (Prm1)* 启动子诱导在精细胞减数分裂后期表达 Cre 重组酶的转基因小鼠品系.所有携带 Cre 重组酶基因的雄性小鼠以及转基因雌性小鼠的雄性后代都是不育的;雌性小鼠未见影响.然而精子计数、精子的游动能力以及形态均正常,转基因雄性小鼠的交配行为以及交配后两细胞胚胎的形成也未见异常.以突变后的 Cre 重组酶基因作为对照,雄性小鼠的生殖能力正常.对携带 Cre 重组酶基因的雄性小鼠和野生型雌性小鼠交配后的胚胎进行分析,发现由于精细胞中 Cre 的高表达引起的染色体重排导致了外显率高达 100% 的妊娠异常(*abortive pregnancies*).他们同时认为,独立于 *loxP* 的基因组修饰也发生在表达 Cre 的体细胞组织中,但是由于检测困难而难以发现;所以建议在使用完 Cre 重组酶后,应尽快将其失活. Gromley 等人^[150]为研究肿瘤抑制因子 Arf 在生殖细胞有丝分裂和减数分裂中的作用,构建了依赖于 Cre(受 Arf 启动子控制)表达的 YFP(黄色荧光蛋白) knock-in 小鼠.肿瘤抑制因子 Arf 在雄性小鼠生殖细胞和眼部细胞发育过程中瞬时表达,该因子的抑制会导致精子发生异常并伴随出生后眼部玻璃质细胞的异常增殖.然而,当他们构建该系统进一步研究 Arf 的功能时,发现 Cre 重组酶的表达会导致雄性小鼠不育,相对于雌性小鼠,其胚胎移植存活率显著降低.雌性 Arf-Cre 小鼠和“floxed”雄性 Arf 小鼠(即 Arf 基因两侧带有 *loxP* 位点)杂交后产生 Arf^{Cre/FL} 后代,该后代表现为不同程度的视觉缺陷甚至完全失明. Loonstra 等人^[151]报道, Cre 重组酶的表达会严重影响细胞增殖,染色体分析发现 Cre 重组酶的表达导致大量的染色体畸变(*chromosomal aberrations*)并增加姐妹染色单体的交换.通过对可诱导表达的杂合重组酶 CreER 的分析发现,该细胞毒性随着 Cre 表达量的升高而增加. Buerger 等

人^[152]发现,在小鼠心肌中特异高表达 Cre 重组酶会导致器官功能障碍,具体为扩张型心肌病. Forni 等人^[153]研究了在构建巢蛋白 Cre 转基因品系小鼠胚胎过程中高表达 Cre 重组酶对大脑发育的影响,发现神经元的增值显著减缓,染色体倍型紊乱和细胞死亡比率增加,室管膜缺陷导致脑过水和脑积水. Naiche 等人^[154]发现, Cre 的活性在胚胎发育过程中能导致大范围的细胞凋亡和致命贫血症.同时发现, Cre 重组酶能诱导严重的发育缺陷,如丧失干细胞造血活性和不同干细胞组织严重的细胞凋亡.这些表型会极大干扰遗传分析,对之前利用 Cre 系统但未设置相应对照的研究工作提出了质疑. Higashi 等人^[155]发现,在成年小鼠和胚胎中系统性的激活杂合重组酶 CreER^{T2} 后,会导致非常直接的血液毒性和异常的染色体重组.转基因小鼠品系 *B6.Cg-Tg(Ins2-cre)25 Mgn/J* 通常称为 RIP-Cre, 该品系中 Cre 重组酶的表达受到 *insulin II* 基因启动子的诱导,且被用于至少 17 个基因的功能研究中.这些研究发现,相应基因的缺失会伴随葡萄糖耐受不良或糖尿病.然而, Lee 等人^[156]的研究发现, RIP-Cre 小鼠本身就表现为葡萄糖耐受不良.同样,有报道显示, Cre 重组酶的表达能导致植物的表型异常. Coppoolse 等人^[157]在番茄、牵牛花和烟草中用 CaMV 35S 启动子控制 Cre 重组酶表达时,发现重组酶能够稳定地导入和表达,但是都导致了不同程度的生长缓慢和显著的叶片萎黄病.虽然没有直接的证据表明表型和 Cre 重组酶表达之间存在联系,但是表型异常的植株都有 Cre 的表达,显示出了极强的关联性.然而当使用组织特异的启动子,如菜豆球蛋白启动子来控制 *cre* 基因时,并没有发生表型异常.

Ehrhardt 等人^[158]调查了人细胞中潜在的 ϕ C31 识别位点是否会导致染色体重组.在 69 个独立的重组事件中,有 6 个(8.7%)整合到 19 号染色体(19q13.31), 10 个(14.49%)整合到 12 号染色体(12q22);最重要的是,通过 PCR 检测分析和单细胞克隆的基因型确定,他们发现, 15% 的整合事件导致了染色体的易位. Liu 等人^[159]在原代人纤维母细胞中稳定表达了 ϕ C31 整合酶.对照组细胞表型及基因型正常,但是表达了 ϕ C31 整合酶的细胞出现了大量的染色体畸变,包括不同程度的易位,该酶表现出严重的诱发突变倾向.为进一步研究其机理,他们分析了 H2AX 的磷酸化,该组蛋白磷酸化为 DNA 双链断裂的重要指示.细胞

中转入 ϕ C31 整合酶基因后, 通过免疫荧光检测到了 H2AX 的磷酸化, Western 杂交分析证实了 H2AX 磷酸化的上调, 然而未发现细胞凋亡相关基因被诱导(p53 和 p21). 睡美人转座酶(Sleeping Beauty transposase)对对照组表型和基因型正常, H2AX 的磷酸化也较弱^[160]. Chen 等人^[161]通过酵母双杂交发现, ϕ C31 整合酶能够与死亡结构域相关蛋白(the death domain-associated protein, DAXX)发生相互作用从而抑制重组; Wang 等人^[162]发现, ϕ C31 整合酶能够与 TTRAP(TRAFA and TNF receptor associated protein)相互作用抑制 NF κ B 细胞信号通路的激活.

问题主要体现在两方面: 一方面是由于重组酶的存在发生了染色体异常, 使得科研工作者不得不重新审视以前的一些研究工作, 单基因突变研究中导致的表型或者基因型变化真正的罪魁祸首可能是重组酶引起的细胞毒性. 那么如何正确地评估实验结果, 如何严格设置对照, 避免由于重组酶的存在导致的结果误读; 另一方面, 即是相关转基因作物和基因治疗载体的安全性. 目前的应对策略主要为严格控制重组酶的表达和提高整合酶的位点特异性以尽量减少非预期重组的发生. Coppoolse 等人^[157]建议在植物中使用 Cre 重组系统时, 最好能够实现组织特异表达, 并且该酶的表达控制在很短的时间内. Silver 和 Livingston^[163]设计了一个含有 Cre 重组酶可以自我切除的反转录病毒载体, 该系统在保证有效切除靶序列的同时切除 Cre 重组酶基因, 有效地降低了可检测到的细胞毒性. Keravala 等人^[164]将目标定位于筛选更高效率和位点专一性的突变 ϕ C31 整合酶, 幸运的是, 突变 P2 整合效率是野生型的 2 倍; 突变酶 P3 介导的体内重组事件有 44% 发生在同一位置, 极大地提高了靶位点的特异性.

4 展望

两类位点特异性重组酶在生物学功能上没有明显的划分, 然而却采用完全不同的催化机制, 这种差异体现在 DNA 结合、联会复合体的形成、链的切割

与交换等一系列连续的精细过程中^[11]. 较之酪氨酸重组酶, 丝氨酸重组酶的催化机制不甚明了, 虽然目前已有对解离酶 $\gamma\delta$, Tn3 和 Sin 的溶液结构或晶体结构的解析^[20,78,79], 但对于丝氨酸家族众多成员来说, 想要了解 DNA 链的交换是如何通过亚基翻转来实现的, 以及执行不同功能的重组酶其催化方式是否存在特定的差异, 而这些差异又是如何展现的等问题, 仅仅是一个开始. 目前, 仍然没有丝氨酸类倒位酶和大型丝氨酸重组酶完整晶体结构的解析, 尤其是大型丝氨酸重组酶具有一个巨大的未知功能的 C 末端, 其作用方式必然不同于丝氨酸类解离酶. 巨大的 C 末端是否仅参与反应方向的控制, 还是跟 DNA 结合紧密相关? 甚至是对其催化活性的发挥也不可或缺? 虽然已有一些氨基酸突变工作提供了一些信息^[97], 但是仍然缺乏全局性的保守氨基酸位点的系统突变分析, 以了解其可变性及在催化中的作用; 而大型丝氨酸重组酶做为目前遗传操作中颇具潜力的工具酶, 更多的系统分析将更有利这类酶的开发利用.

噬菌体整合酶介导的位点特异性重组最大的优势在于能够高效专一地实现特定位点之间的重组, 且其重组 DNA 能力不受片段大小限制. 虽然目前也发现存在一定非特异性重组之类的问题, 但是系统的优化和改进以避免或降低问题出现的风险仍是可行的方向. 今后, 位点特异性重组酶的应用研究大致会集中在以下几个方面: 基于保守氨基酸突变分析和晶体结构解析的酶分子改造, 包括杂合酶的开发, 基于位点的突变筛选等, 首要的目标是增加在不同宿主细胞中的整合特异性; 新的位点特异性重组系统的发掘与应用, 目前基因组测序数据不断出现, 其中有大量被注释为噬菌体(原噬菌体)序列, 可作为新系统筛选和优化的重要来源; 多系统的综合循环应用, 既可有效地增加转基因的数量, 也能更加方便在体外进行多基因操作. 位点特异性重组系统开发和优化的总体目标是安全(专一)、高效、简单和可控, 相信随着研究的深入, 该类系统将在遗传工程中发挥更加重要的作用.

致谢 感谢复旦大学生命科学学院微生物学与微生物工程学系白洁、张渤和戴瑞雪对本文校对提供的帮助.

参考文献

- 1 Grindley N D F, Whiteson K L, Rice P A. Mechanisms of site-specific recombination. *Annu Rev Biochem*, 2006, 75: 567—605
- 2 Sadofsky M J. The RAG proteins in V(D)J recombination: more than just a nuclease. *Nucleic Acids Res*, 2001, 29: 1399—1409
- 3 Lieber M R. Site-specific recombination in the immune system. *FASEB J*, 1991, 5: 2934—2944
- 4 Haugen P, Simon D M, Bhattacharya D. The natural history of group I introns. *Trends Genet*, 2005, 21: 111—119
- 5 Hsu P L, Ross W, Landy A. The λ -phage *att* site: functional limits and interaction with Int protein. *Nature*, 1980, 285: 85—91
- 6 Leong J M, Nunesduby S, Lesser C F, et al. The ϕ 80 and P22 attachment sites: primary structure and interaction with *Escherichia coli* integration host factor. *J Biol Chem*, 1985, 260: 4468—4477
- 7 Lewis J A, Hatfull G F. Control of directionality in L5 integrase-mediated site-specific recombination. *J Mol Biol*, 2003, 326: 805—821
- 8 Colloms S D, Sykora P, Szatmari G, et al. Recombination at ColE1 *cer* requires the *Escherichia coli xerC* gene product, a member of the λ -integrase family of site-specific recombinases. *J Bacteriol*, 1990, 172: 6973—6980
- 9 Abremski K, Hoess R. Bacteriophage P1 site-specific recombination: purification and properties of the Cre recombinase protein. *J Biol Chem*, 1984, 259: 1509—1514
- 10 Baum J A. Tn5401, a new class II transposable element from *Bacillus thuringiensis*. *J Bacteriol*, 1994, 176: 2835—2845
- 11 Mahillon J, Lereclus D. Structural and functional-analysis of Tn4430-identification of an integrase-like protein involved in the co-integrate-resolution process. *EMBO J*, 1988, 7: 1515—1526
- 12 Bastos M D D, Murphy E. Transposon Tn554 encodes three products required for transposition. *EMBO J*, 1988, 7: 2935—2941
- 13 Poyartsalmeron C, Trieu cuot P, Carlier C, et al. Molecular characterization of two proteins involved in the excision of the conjugative transposon Tn1545: homologies with other site-specific recombinases. *EMBO J*, 1989, 8: 2425—2433
- 14 Su Y A, Clewell D B. Characterization of the left 4 Kb of conjugative transposon Tn916: determinants involved in excision. *Plasmid*, 1993, 30: 234—250
- 15 Carrasco C D, Buettner J A, Golden J W. Programmed DNA rearrangement of a cyanobacterial *hupL* Gene in heterocysts. *Proc Natl Acad Sci USA*, 1995, 92: 791—795
- 16 Klemm P. Two regulatory *fim* Genes, *fimB* and *fimE*, control the phase variation of type I fimbriae in *Escherichia coli*. *EMBO J*, 1986, 5: 1389—1393
- 17 Hartley J L, Donelson J E. Nucleotide-sequence of the yeast plasmid. *Nature*, 1980, 286: 860—865
- 18 Massad G, Mobley H L T. Genetic organization and complete sequence of the *Proteus mirabilis pmf* fimbrial operon. *Gene*, 1994, 150: 101—104
- 19 Li X, Lockatell C V, Johnson D E, et al. Identification of MrpI as the sole recombinase that regulates the phase variation of MR/P fimbria, a bladder colonization factor of uropathogenic *Proteus mirabilis*. *Mol Microbiol*, 2002, 45: 865—874
- 20 Li W K, Kamtekar S, Xiong Y, et al. Structure of a synaptic $\gamma\delta$ resolvase tetramer covalently linked to two cleaved DNAs. *Science*, 2005, 309: 1210—1215
- 21 Yang W, Steitz T A. Crystal-structure of the site-specific secombinase $\gamma\delta$ resolvase complexed with a 34-bp cleavage site. *Cell*, 1995, 82: 193—207
- 22 Olorunniji F J, He J Y, Wenwieser S V C T, et al. Synapsis and catalysis by activated Tn3 resolvase mutants. *Nucleic Acids Res*, 2008, 36: 7181—7191
- 23 Nollmann M, Byron O, Stark W M. Behavior of Tn3 resolvase in solution and its interaction with *res*. *Biophys J*, 2005, 89: 1920—1931
- 24 Arnold P H, Blake D G, Grindley N D F, et al. Mutants of Tn3 resolvase which do not require accessory binding sites for recombination activity. *EMBO J*, 1999, 18: 1407—1414
- 25 Rowland S J, Boocock M R, McPherson A L, et al. Regulatory mutations in Sin recombinase support a structure-based model of the synaptosome. *Mol Microbiol*, 2009, 74: 282—298
- 26 Rowland S J, Stark W M, Boocock M R. Sin recombinase from *Staphylococcus aureus*: synaptic complex architecture and transposon targeting. *Mol Microbiol*, 2002, 44: 607—619
- 27 Dhar G, Heiss J K, Johnson R C. Mechanical constraints on Hin subunit rotation imposed by the Fis/enhancer system and DNA supercoiling during site-specific recombination. *Mol Cell*, 2009, 34: 746—759
- 28 Mertens G, Klippel A, Fuss H, et al. Site-specific recombination in bacteriophage Mu: characterization of binding-sites for the DNA invertase Gin. *EMBO J*, 1988, 7: 1219—1227
- 29 Klippel A, Mertens G, Patschinsky T, et al. The DNA invertase Gin of phage Mu: formation of a covalent complex with DNA via a phosphoserine at amino-acid position-9. *EMBO J*, 1988, 7: 1229—1237

- 30 Mertens G, Fuss H, Kahmann R. Purification and properties of the DNA invertase Gin encoded by bacteriophage Mu. *J Biol Chem*, 1986, 261: 5668—5672
- 31 Morita K, Yamamoto T, Fusada N, et al. *In vitro* characterization of the site-specific recombination system based on actinophage TG1 integrase. *Mol Genet Genomics*, 2009, 282: 607—616
- 32 Zhang L, Ou X J, Zhao G P, et al. Highly efficient *in vitro* site-specific recombination system based on *Streptomyces* phage ϕ BT1 integrase. *J Bacteriol*, 2008, 190: 6392—6397
- 33 Thorpe H M, Smith M C M. *In vitro* site-specific integration of bacteriophage DNA catalyzed by a recombinase of the resolvase/invertase family. *Proc Natl Acad Sci USA*, 1998, 95: 5505—5510
- 34 Olivares E C, Hollis R P, Calos M P. Phage R4 integrase mediates site-specific integration in human cells. *Gene*, 2001, 278: 167—176
- 35 Bibb L A, Hancox M I, Hatfull G F. Integration and excision by the large serine recombinase ϕ Rv1 integrase. *Mol Microbiol*, 2005, 55: 1896—1910
- 36 Kim A I, Ghosh P, Aaron M A, et al. Mycobacteriophage Bxb1 integrates into the *Mycobacterium smegmatis* *groEL1* gene. *Mol Microbiol*, 2003, 50: 463—473
- 37 Breuner A, Brondsted L, Hammer K. Novel organization of genes involved in prophage excision identified in the temperate lactococcal bacteriophage TP901-1. *J Bacteriol*, 1999, 181: 7291—7297
- 38 Park M O, Lim K H, Kim T H, et al. Characterization of site-specific recombination by the integrase MJ1 from enterococcal bacteriophage ϕ FC1. *J Microbiol Biotechnol*, 2007, 17: 342—347
- 39 Lyras D, Adams V, Lucet I, et al. The large resolvase TnpX is the only transposon-encoded protein required for transposition of the Tn4451/3 family of integrative mobilizable elements. *Mol Microbiol*, 2004, 51: 1787—1800
- 40 Boocock M R, Zhu X W, Grindley N D F. Catalytic residues of $\gamma\delta$ resolvase act *in cis*. *EMBO J*, 1995, 14: 5129—5140
- 41 Abremski K, Gottesman S. Site-specific recombination: Xis-independent excisive recombination of bacteriophage λ . *J Mol Biol*, 1981, 153: 67—78
- 42 Groth A C, Calos M P. Phage integrases: biology and applications. *J Mol Biol*, 2004, 335: 667—678
- 43 Smith M C M, Thorpe H M. Diversity in the serine recombinases. *Mol Microbiol*, 2002, 44: 299—307
- 44 Echols H. Lysogeny: viral repression and site-specific recombination. *Annu Rev Biochem*, 1971, 40: 827—854
- 45 Biswas T, Aihara H, Radman-Livaja M, et al. A structural basis for allosteric control of DNA recombination by λ integrase. *Nature*, 2005, 435: 1059—1066
- 46 Landy A, Ross W. Viral integration and excision: structure of λ *att* sites. *Science*, 1977, 197: 1147—1160
- 47 Rajeev L, Malanowska K, Gardner J F. Challenging a paradigm: the role of DNA homology in tyrosine recombinase reactions. *Microbiol Mol Biol R*, 2009, 73: 300—309
- 48 Warren D, Laxmikanthan G, Landy A. A chimeric Cre recombinase with regulated directionality. *Proc Natl Acad Sci USA*, 2008, 105: 18278—18283
- 49 Bauer C E, Gardner J F, Gumpert R I. Extent of sequence homology required for bacteriophage λ site-specific recombination. *J Mol Biol*, 1985, 181: 187—197
- 50 Weisberg R A, Enquist L W, Foeller C, et al. Role for DNA homology in site-specific recombination: the isolation and characterization of a site affinity mutant of coliphage λ . *J Mol Biol*, 1983, 170: 319—342
- 51 Caparon M G, Scott J R. Excision and insertion of the conjugative transposon Tn916 involves a novel recombination mechanism. *Cell*, 1989, 59: 1027—1034
- 52 Ouaisi A, Vergnes B, Borges M, et al. Identification and molecular characterization of two novel *Trypanosoma cruzi* genes encoding polypeptides sharing sequence motifs found in proteins involved in RNA editing reactions. *Gene*, 2000, 253: 271—280
- 53 Malanowska K, Salyers A A, Gardner J F. Characterization of a conjugative transposon integrase, IntDOT. *Mol Microbiol*, 2006, 60: 1228—1240
- 54 Cheng Q I, Paszkiet B J, Shoemaker N B, et al. Integration and excision of a bacteroides conjugative transposon, CTnDOT. *J Bacteriol*, 2000, 182: 4035—4043
- 55 Schmidt J W, Rajeev L, Salyers A A, et al. NBU1 integrase: evidence for an altered recombination mechanism. *Mol Microbiol*, 2006, 60: 152—164
- 56 Bouvier M, Demarre G, Mazel D. Integron cassette insertion: a recombination process involving a folded single strand substrate. *EMBO J*, 2005, 24: 4356—4367
- 57 Val M E, Bouvier M, Campos J, et al. The single-stranded genome of phage CTX is the form used for integration into the genome of *Vibrio*

- cholerae*. Mol Cell, 2005, 19: 559—566
- 58 Salvo J J, Grindley N D F. The $\gamma\delta$ resolvase bends the *res* site into a recombinogenic complex. EMBO J, 1988, 7: 3609—3616
 - 59 Dhar G, McLean M M, Heiss J K, et al. The Hin recombinase assembles a tetrameric protein swivel that exchanges DNA strands. Nucleic Acids Res, 2009, 37: 4743—4756
 - 60 Moskowitz I P G, Heichman K A, Johnson R C. Alignment of recombination sites in Hin-mediated site-specific DNA recombination. Gene Dev, 1991, 5: 1635—1645
 - 61 Hughes K T, Youderian P, Simon M I. Phase variation in *Salmonella*: analysis of Hin recombinase and *hix* recombination site interaction *in vivo*. Gene Dev, 1988, 2: 937—948
 - 62 Johnson R C, Simon M I. Hin-mediated site-specific recombination requires two 26-bp recombination sites and a 60-bp recombinational enhancer. Cell, 1985, 41: 781—791
 - 63 Gregory M A, Till R, Smith M C M. Integration site for *Streptomyces* phage ϕ BT1 and development of site-specific integrating vectors. J Bacteriol, 2003, 185: 5320—5323
 - 64 Lutz K A, Corneille S, Azhagiri A K, et al. A novel approach to plastid transformation utilizes the ϕ C31 phage integrase. Plant J, 2004, 37: 906—913
 - 65 Forterre P, Gribaldo S, Gadelle D, et al. Origin and evolution of DNA topoisomerases. Biochimie, 2007, 89: 427—446
 - 66 Lim H M, Simon M I. The role of negative supercoiling in Hin-mediated site-specific recombination. J Biol Chem, 1992, 267: 11176—11182
 - 67 Benjamin K R, Abola A P, Kanaar R, et al. Contributions of supercoiling to Tn3 resolvase and phage Mu Gin site-specific recombination. J Mol Biol, 1996, 256: 50—65
 - 68 Trigueros S, Tran T, Sorto N, et al. mwr Xer site-specific recombination is hypersensitive to DNA supercoiling. Nucleic Acids Res, 2009, 37: 3580—3587
 - 69 Zhang L, Wang L, Wang J, et al. DNA cleavage is independent of synapsis during *Streptomyces* phage ϕ BT1 integrase-mediated site-specific recombination. J Mol Cell Biol, 2010, 2: 264—275
 - 70 Ghosh P, Kim A I, Hatfull G F. The orientation of mycobacteriophage Bxb1 integration is solely dependent on the central dinucleotide of *attP* and *attB*. Mol Cell, 2003, 12: 1101—1111
 - 71 Hartley J L, Temple G F, Brasch M A. DNA cloning using *in vitro* site-specific recombination. Genome Res, 2000, 10: 1788—1795
 - 72 Chen Y, Narendra U, Iype L E, et al. Crystal structure of a Flp recombinase-Holliday junction complex: assembly of an active oligomer by helix swapping. Mol Cell, 2000, 6: 885—897
 - 73 Guo F, Gopaul D N, VanDuyne G D. Structure of Cre recombinase complexed with DNA in a site-specific recombination synapse. Nature, 1997, 389: 40—46
 - 74 Ghosh K, Guo F, Van Duyne G D. Synapsis of *loxP* sites by Cre recombinase. J Biol Chem, 2007, 282: 24004—24016
 - 75 Craig N L, Nash H A. The Mechanism of phage λ site-specific recombination: site-specific breakage of DNA by Int topoisomerase. Cell, 1983, 35: 795—803
 - 76 Graham J E, Sivanathan V, Sherratt D J, et al. FtsK translocation on DNA stops at XerCD-*dif*. Nucleic Acids Res, 2010, 38: 72—81
 - 77 Lowe J, Ellonen A, Allen M D, et al. Molecular mechanism of sequence-directed DNA loading and translocation by FtsK. Mol Cell, 2008, 31: 498—509
 - 78 Nollmann M, He J Y, Byron O, et al. Solution structure of the Tn3 resolvase-crossover site synaptic complex. Mol Cell, 2004, 16: 127—137
 - 79 Mouw K W, Rowland S J, Gajjar M M, et al. Architecture of a serine recombinase-DNA regulatory complex. Mol Cell, 2008, 30: 145—155
 - 80 Spaenydekking L, Vanhemert M, Vandeputte P, et al. Gin Invertase of bacteriophage Mu is a dimer in solution, with the domain for dimerization in the N-terminal part of the protein. Biochemistry-US, 1995, 34: 1779—1786
 - 81 Yuan P, Gupta K, Van Duynel G D. Tetrameric structure of a serine integrase catalytic domain. Structure, 2008, 16: 1275—1286
 - 82 Smith M C M, Brown W R A, McEwan A R, et al. Site-specific recombination by ϕ C31 integrase and other large serine recombinases. Biochem Soc T, 2010, 38: 388—394
 - 83 Gupta M, Till R, Smith M C M. Sequences in *attB* that affect the ability of the ϕ C31 integrase to synapse and to activate DNA cleavage. Nucleic Acids Res, 2007, 35: 3407—3419
 - 84 Ghosh P, Pannunzio N R, Hatfull G F. Synapsis in phage Bxb1 integration: selection mechanism for the correct pair of recombination sites. J Mol Biol, 2005, 349: 331—348
 - 85 Richet E, Abcarian P, Nash H A. Synapsis of attachment sites during λ integrative recombination involves capture of a naked DNA by a protein-DNA complex. Cell, 1988, 52: 9—17

- 86 Grindley N D F, Lauth M R, Wells R G, et al. Transposon-mediated site-specific recombination: identification of three binding sites for resolvase at the *res* sites of *gd* and *Tn3*. *Cell*, 1982, 30: 19—27
- 87 Crisona N J, Kanaar R, Gonzalez T N, et al. Processive recombination by wild-type *Gin* and an enhancer-independent mutant: insight into the mechanisms of recombination selectivity and strand exchange. *J Mol Biol*, 1994, 243: 437—457
- 88 Kanaar R, Klippel A, Shekhtman E, et al. Processive recombination by the phage *Mu* *Gin* system: implications for the mechanisms of DNA strand exchange, DNA site alignment, and enhancer action. *Cell*, 1990, 62: 353—366
- 89 Smith M C A, Till R, Brady K, et al. Synapsis and DNA cleavage in ϕ C31 integrase-mediated site-specific recombination. *Nucleic Acids Res*, 2004, 32: 2607—2617
- 90 Ghosh P, Wasil L R, Hatfull G F. Control of phage *Bxb1* excision by a novel recombination directionality factor. *PLoS Biol*, 2006, 4: 964—974
- 91 Ghosh P, Bibb L A, Hatfull G F. Two-step site selection for serine-integrase-mediated excision: DNA-directed integrase conformation and central dinucleotide proofreading. *Proc Natl Acad Sci USA*, 2008, 105: 3238—3243
- 92 Nunes-Duby S E, Kwon H J, Tirumalai R S, et al. Similarities and differences among one hundred and five members of the *Int* family of site-specific recombinases. *Nucleic Acids Res*, 1998, 26: 391—406
- 93 Gibb B, Gupta K, Ghosh K, et al. Requirements for catalysis in the *Cre* recombinase active site. *Nucleic Acids Res*, 2010, doi: 10.1093/nar/gkq384
- 94 Wang H M, Smith M C M, Mullany P. The conjugative transposon *Tn5397* has a strong preference for integration into its *clostridium difficile* target site. *J Bacteriol*, 2006, 188: 4871—4878
- 95 Sato T, Harada K, Kobayasi Y. Analysis of suppressor mutations of *spoIVCA* mutations: occurrence of DNA rearrangement in the absence of a site-specific DNA recombinase *SpoIVCA* in *Bacillus subtilis*. *J Bacteriol*, 1996, 178: 3380—3383
- 96 Rowley P A, Smith M C M. Role of the N-terminal domain of ϕ C31 integrase in *attB-attP* synapsis. *J Bacteriol*, 2008, 190: 6918—6921
- 97 Rowley P A, Smith M C A, Younger E, et al. A motif in the C-terminal domain of ϕ C31 integrase controls the directionality of recombination. *Nucleic Acids Res*, 2008, 36: 3879—3891
- 98 Urnov F D, Miller J C, Lee Y L, et al. Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature*, 2005, 435: 646—651
- 99 Lombardo A, Genovese P, Beausejour C M, et al. Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nat Biotechnol*, 2007, 25: 1298—1306
- 100 Gordley R M, Smith J D, Graslund T, et al. Evolution of programmable zinc finger recombinases with activity in human cells. *J Mol Biol*, 2007, 367: 802—813
- 101 Metzger D, Clifford J, Chiba H, et al. Conditional site-specific recombination in mammalian cells using a ligand-dependent chimeric *Cre* recombinase. *Proc Natl Acad Sci USA*, 1995, 92: 6991—6995
- 102 Tannour-Louet M, Porteu A, Vaulont S, et al. A tamoxifen-inducible chimeric *Cre* recombinase specifically effective in the fetal and adult mouse liver. *Hepatology*, 2002, 35: 1072—1081
- 103 Imai T, Jiang M, Chambon P, et al. Impaired adipogenesis and lipolysis in the mouse upon selective ablation of the retinoid X receptor α mediated by a tamoxifen-inducible chimeric *Cre* recombinase (*Cre-ER^{T2}*) in adipocytes. *Proc Natl Acad Sci USA*, 2001, 98: 224—228
- 104 Brocard J, Feil R, Chambon P, et al. A chimeric *Cre* recombinase inducible by synthetic, but not by natural ligands of the glucocorticoid receptor. *Nucleic Acids Res*, 1998, 26: 4086—4090
- 105 Urbanski W M, Condie B G. Textpresso site-specific recombinases: a text-mining server for the recombinase literature including *Cre* mice and conditional alleles. *Genesis*, 2009, 47: 842—846
- 106 Colwill K, Wells C D, Elder K, et al. Modification of the Creator recombination system for proteomics applications: improved expression by addition of splice sites. *BMC Biotechnol*, 2006, 6: 13
- 107 Liu Q H, Li M Z, Leibham D, et al. The univector plasmid-fusion system, a method for rapid construction of recombinant DNA without restriction enzymes. *Curr Biol*, 1998, 8: 1300—1309
- 108 Rual J F, Hill D E, Vidal M. ORFeome projects: Gateway between genomics and omics. *Curr Opin Chem Biol*, 2004, 8: 20—25
- 109 Sasaki Y, Sone T, Yoshida S, et al. Evidence for high specificity and efficiency of multiple recombination signals in mixed DNA cloning by the multisite Gateway system. *J Biotechnol*, 2004, 107: 233—243
- 110 Curtis M D, Grossniklaus U. A Gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiol*, 2003, 133: 462—469
- 111 Magnani E, Bartling L, Hake S. From Gateway to multisite Gateway in one recombination event. *BMC Mol Biol*, 2006, 7

- 112 Chen Q J, Zhou H M, Chen J, et al. A Gateway-based platform for multigene plant transformation. *Plant Mol Biol*, 2006, 62: 927—936
- 113 Marsischky G, LaBaer J. Many paths to many clones: a comparative look at high-throughput cloning methods. *Genome Res*, 2004, 14: 2020—2028
- 114 Komatsua M, Uchiyama T, Omura S, et al. Genome-minimized *Streptomyces* host for the heterologous expression of secondary metabolism. *Proc Natl Acad Sci USA*, 2010, 107: 2646—2651
- 115 Schetelig M F, Scolari F, Handler A M, et al. Site-specific recombination for the modification of transgenic strains of the Mediterranean fruit fly *Ceratitis capitata*. *Proc Natl Acad Sci USA*, 2009, 106: 18171—18176
- 116 Watson A T, Garcia V, Bone N, et al. Gene tagging and gene replacement using recombinase-mediated cassette exchange in *Schizosaccharomyces pombe*. *Gene*, 2008, 407: 63—74
- 117 Ikeda H, Ishikawa J, Hanamoto A, et al. Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*. *Nat Biotechnol*, 2003, 21: 526—531
- 118 Ow D W. GM maize from site-specific recombination technology, what next? *Curr Opin Biotechnol*, 2007, 18: 115—120
- 119 Hu Q, Kononowicz-Hodges H, Nelson-Vasilchik K, et al. FLP recombinase-mediated site-specific recombination in rice. *Plant Biotechnol J*, 2008, 6: 176—188
- 120 Fladung M, Becker D. Targeted integration and removal of transgenes in hybrid aspen (*Populus tremula* L. x *P-tremuloides* Michx.) using site-specific recombination systems. *Plant Biol*, 2010, 12: 334—340
- 121 Rubtsova M, Kempe K, Gils A, et al. Expression of active *Streptomyces* phage ϕ C31 integrase in transgenic wheat plants. *Plant Cell Rep*, 2008, 27: 1821—1831
- 122 Birchler J A, Yu W, Han F. Plant engineered minichromosomes and artificial chromosome platforms. *Cytogenet Genome Res*, 2008, 120: 228—232
- 123 Khan M S, Khalid A M, Malik K A. Phage ϕ C31 integrase: a new tool in plastid genome engineering. *Trends Plant Sci*, 2005, 10: 1—3
- 124 Kittiwongwattana C, Lutz K, Clark M, et al. Plastid marker gene excision by the ϕ C31 phage site-specific recombinase. *Plant Mol Biol*, 2007, 64: 137—143
- 125 Seahill M D, Pastar I, Cross G A M. Cre recombinase-based positive-negative selection systems for genetic manipulation in *Trypanosoma brucei*. *Mol Biochem Parasit*, 2008, 157: 73—82
- 126 Davis M W, Morton J J, Carroll D, et al. Gene activation using FLP recombinase in *C. elegans*. *PLoS Genet*, 2008, 4: e1000028, doi: 10.1371/journal.pgen.1000028
- 127 Nkrumah L J, Muhle R A, Moura P A et al. Efficient site-specific integration in *Plasmodium falciparum* chromosomes mediated by mycobacteriophage Bxb1 integrase. *Nat Methods*, 2006, 3: 615—621
- 128 Combe A, Giovannini D, Carvalho T G, et al. Clonal conditional mutagenesis in *Malaria parasites*. *Cell Host Microbe*, 2009, 5: 386—396
- 129 Nakayama G, Kawaguchi Y, Koga K, et al. Site-specific gene integration in cultured silkworm cells mediated by ϕ C31 integrase. *Mol Genet Genomics*, 2006, 275: 1—8
- 130 Le X N, Langenau D M, Keefe M D, et al. Heat shock-inducible Cre/*lox* approaches to induce diverse types of tumors and hyperplasia in transgenic zebrafish. *Proc Natl Acad Sci USA*, 2007, 104: 9410—9415
- 131 Hans S, Kaslin J, Freudenreich D, et al. Temporally-controlled site-specific recombination in zebrafish. *PLoS ONE*, 2009, 4: e4640, doi: 10.1371/journal.pone.0004640
- 132 Lister J A. Transgene excision in zebrafish using the ϕ C31 integrase. *Genesis*, 2010, 48: 137—143
- 133 Adams D J, van der Weyden L. Are we Creating problems? Negative effects of Cre recombinase. *Genesis*, 2001, 29: 115
- 134 Tiscornia G, Tergaonkar V, Galimi F, et al. Cre recombinase-inducible RNA interference mediated by lentiviral vectors. *Proc Natl Acad Sci USA*, 2004, 101: 7347—7351
- 135 Wu S, Ying G, Wu Q, et al. Toward simpler and faster genome-wide mutagenesis in mice. *Nature Genet*, 2007, 39: 922—930
- 136 Sakurai K, Shimoji M, Tahimic C G T, et al. Efficient integration of transgenes into a defined locus in human embryonic stem cells. *Nucleic Acids Res*, 2010, 38: e96, doi: 10.1093/nar/gkp1234
- 137 Raymond C S, Soriano P. High-efficiency FLP and ϕ C31 site-specific recombination in mammalian cells. *PLoS ONE*, 2007, 2: e162, doi: 10.1371/journal.pone.0000162
- 138 Groth A C, Olivares E C, Thyagarajan B, et al. A phage integrase directs efficient site-specific integration in human cells. *Proc Natl Acad Sci USA*, 2000, 97: 5995—6000
- 139 Olivares E C, Hollis R P, Chalberg T W, et al. Site-specific genomic integration produces therapeutic factor IX levels in mice. *Nat Biotechnol*, 2002, 20: 1124—1128

- 140 Ortiz-Urda S, Thyagarajan B, Keene D R, et al. Stable nonviral genetic correction of inherited human skin disease. *Nat Med*, 2002, 8: 1166—1170
- 141 Ortiz-Urda S, Thyagarajan B, Keene D R et al. ϕ C31 integrase-mediated nonviral genetic correction of junctional epidermolysis bullosa. *Hum Gene Ther*, 2003, 14: 923—928
- 142 Ortiz-Urda S, Lin Q, Yant S R, et al. Sustainable correction of junctional epidermolysis bullosa *via* transposon-mediated nonviral gene transfer. *Gene Ther*, 2003, 10: 1099—1104
- 143 Bertoni C, Jarrahan S, Wheeler T M, et al. Enhancement of plasmid-mediated gene therapy for muscular dystrophy by directed plasmid integration. *Proc Natl Acad Sci USA*, 2006, 103: 419—424
- 144 Quenneville S P, Chapdelaine P, Rousseau J, et al. Nucleofection of muscle-derived stem cells and myoblasts with ϕ C31 integrase: stable expression of a full-length-dystrophin fusion gene by human myoblasts. *Mol Ther*, 2004, 10: 679—687
- 145 Karimi M, Bleys A, Vanderhaeghen R, et al. Building blocks for plant gene assembly. *Plant Physiol*, 2007, 145: 1183—1191
- 146 Huang J, Zhou W K, Dong W, et al. Directed, efficient, and versatile modifications of the *Drosophila* genome by genomic engineering. *Proc Natl Acad Sci USA*, 2009, 106: 8284—8289
- 147 Kameyama Y, Kawabe Y, Ito A, et al. An accumulative site-specific gene integration system using Cre recombinase-mediated cassette exchange. *Biotechnol Bioeng*, 2010, 105: 1106—1114
- 148 Gilbertson L. Cre-lox recombination: Cre-ative tools for plant biotechnology. *Trends Biotechnol*, 2003, 21: 550—555
- 149 Schmidt E E, Taylor D S, Prigge J R, et al. Illegitimate Cre-dependent chromosome rearrangements in transgenic mouse spermatids. *Proc Natl Acad Sci USA*, 2000, 97: 13702—13707
- 150 Gromley A, Churchman M L, Zindy F, et al. Transient expression of the Arf tumor suppressor during male germ cell and eye development in Arf-Cre reporter mice. *Proc Natl Acad Sci USA*, 2009, 106: 6285—6290
- 151 Loonstra A, Vooijs M, Beverloo H B, et al. Growth inhibition and DNA damage induced by Cre recombinase in mammalian cells. *Proc Natl Acad Sci USA*, 2001, 98: 9209—9214
- 152 Buerger A, Rozhitskaya O, Sherwood M C, et al. Dilated cardiomyopathy resulting from high-level myocardial expression of Cre-recombinase. *J Card Failure*, 2006, 12: 392—398
- 153 Forni P E, Scuoppo C, Imayoshi I, et al. High levels of Cre expression in neuronal progenitors cause defects in brain development leading to microencephaly and hydrocephaly. *J Neurosci*, 2006, 26: 9593—9602
- 154 Naiche L A, Papaioannou V E. Cre activity causes widespread apoptosis and lethal anemia during embryonic development. *Genesis*, 2007, 45: 768—775
- 155 Higashi A Y, Ikawa T, Muramatsu M, et al. Direct hematological toxicity and illegitimate chromosomal recombination caused by the systemic activation of CreER^{T2}. *J Immunol*, 2009, 182: 5633—5640
- 156 Lee J Y, Ristow M, Lin X Y, et al. RIP-Cre revisited, evidence for impairments of pancreatic beta-cell function. *J Biol Chem*, 2006, 281: 2649—2653
- 157 Coppoolse E R, de Vroomen M J, Roelofs D, et al. Cre recombinase expression can result in phenotypic aberrations in plants. *Plant Mol Biol*, 2003, 51: 263—279
- 158 Ehrhardt A, Engler J A, Xu H, et al. Molecular analysis of chromosomal rearrangements in mammalian cells after ϕ C31-mediated integration. *Hum Gene Ther*, 2006, 17: 1077—1094
- 159 Liu J, Jeppesen I, Nielsen K, et al. ϕ C31 integrase induces chromosomal aberrations in primary human fibroblasts. *Gene Ther*, 2006, 13: 1188—1190
- 160 Liu J S T, Gjetting T, Jensen T G. PhiC31 integrase induces a DNA damage response and chromosomal rearrangements in human adult fibroblasts. *BMC Biotechnol*, 2009, 9: 31
- 161 Chen J Z, Ji C N, Xu G L, et al. DAXX interacts with phage ϕ C31 integrase and inhibits recombination. *Nucleic Acids Res*, 2006, 34: 6298—6304
- 162 Wang B Y, Xu G L, Zhou C H, et al. PhiC31 integrase interacts with TTRAP and inhibits NF- κ B activation. *Mol Biol Rep*, 2010, 37: 2809—2816
- 163 Silver D P, Livingston D M. Self-excising retroviral vectors encoding the Cre recombinase overcome Cre-mediated cellular toxicity. *Mol Cell*, 2001, 8: 233—243
- 164 Keravala A, Lee S, Thyagarajan B, et al. Mutational derivatives of ϕ C31 integrase with increased efficiency and specificity. *Mol Ther*, 2009, 17: 112—120

Site-Specific Recombination Systems: Mechanisms and Applications

ZHANG Lin¹, ZHAO GuoPing^{1,2,3} & DING XiaoMing¹

1 Department of Microbiology and Microbial Engineering, School of Life Sciences, Fudan University, Shanghai 200433, China;

2 Key Laboratory of Synthetic biology, Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200032, China;

3 Shanghai-MOST Key Laboratory of Health and Disease Genomics, Chinese National Human Genome Center at Shanghai, Shanghai 201203, China

Site-specific recombinases mediate integration, excision and inversion of target sites by recognizing the defined attachment sites, forming synaptic complexes followed by DNA cleavage and rejoining. Tyrosine or serine within the active center of recombinases attacks the sugar-phosphate backbone to form covalently linked intermediates, without the accessory of high-energy cofactors. Site-specific recombination systems have been widely used in genetic engineering attributed to the high efficiency and specificity of the reaction. The mechanisms of site-specific recombination are described based on the properties of recognition sites, the compositions and structures of recombinases, and the features of catalytic processes. The current status, research hot spots and problems of developing site-specific recombination systems for applications are extensively reviewed, and the future trends are discussed.

site-specific recombination, tyrosine recombinase, serine recombinase, integrase

doi: 10.1360/052010-719