

引文格式:

金帆, 李奉庭, 夏霖. CRISPR/Cas 在定向进化技术中的应用 [J]. 集成技术, 2021, 10(4): 33-49.

Jin F, Li FT, Xia L. The application of CRISPR/Cas in directed evolution [J]. Journal of Integration Technology, 2021, 10(4): 33-49.

## CRISPR/Cas 在定向进化技术中的应用

金 帆<sup>1\*</sup> 李奉庭<sup>1,2</sup> 夏 霖<sup>1\*</sup>

<sup>1</sup>(中国科学院深圳先进技术研究院 深圳合成生物学创新研究院

中国科学院定量工程生物学重点实验室 深圳 518055)

<sup>2</sup>(中国科学院大学 北京 100049)

**摘 要** 定向进化作为改善生物分子目的活性的有效策略之一, 已广泛应用于食品、工业、医学等领域, 该技术依赖于高质量突变体文库的建立与高效目的产物选择与筛选。近年来, 随着 CRISPR 的飞速发展, 相继开发了各类 CRISPR 衍生工具以满足不同需求。其中, 一些新型 CRISPR 工具的出现为定向进化技术带来了更大的发展潜力, 使人们可以在广泛宿主中实现对特定基因的原位进化。同时, 借助人工或天然途径产生遗传多样性的方式给了人们更多的选择, 可以根据研究需求采取更高效的进化策略。该文首先介绍各类 CRISPR 工具, 随后总结 CRISPR 介导的突变与筛选平台, 最后讨论 CRISPR 在定向进化技术中的发展趋势与机遇。

**关键词** 基因编辑; CRISPR/Cas; 定向进化; CRISPR 介导的定向进化

中图分类号 Q 81 文献标志码 A doi: 10.12146/j.issn.2095-3135.20210427012

## The Application of CRISPR/Cas in Directed Evolution

JIN Fan<sup>1\*</sup> LI Fengting<sup>1,2</sup> XIA Lin<sup>1\*</sup>

<sup>1</sup>(CAS Key Laboratory of Quantitative Engineering Biology, Shenzhen Institute of Synthetic Biology, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China)

<sup>2</sup>(University of Chinese Academy of Sciences, Beijing 100049, China)

\*Corresponding Author: fan.jin@siat.ac.cn; lin.xia@siat.ac.cn

**Abstract** To gain the desirable activity of biomolecules, directed evolution is served as a competent technology depending on high quality mutant library, effective selection and screening, which has been widely used in food, industry and medical fields. CRISPR (clustered regularly interspaced short palindromic repeats) has developed rapidly in recent years, and various CRISPR derivatives have been developed to meet

收稿日期: 2021-04-27 修回日期: 2021-05-14

基金项目: 国家重点研发计划项目(2018YFA0902701); 中国科学院战略重点研究计划项目(XDB29050501, XDPB18); 国家自然科学基金项目(32025022, 3201101136); 深圳市科技创新委员会项目(KQTD2015033117210153, JCYJ20170818164139781)

作者简介: 金帆(通讯作者), 博士, 研究方向为合成生物学, E-mail: fan.jin@siat.ac.cn; 李奉庭(共同第一作者), 硕士研究生, 研究方向为人工基因线路的设计与开发; 夏霖(通讯作者), 博士研究生导师, 研究方向为合成生物学微生物与蛋白改造, E-mail: lin.xia@siat.ac.cn.

different demands. The emergence of some new CRISPR tools has given greater development potential for directed evolution, allowing people to evolve specific genes *in situ* in a wide range of hosts. At the same time, the way of generating genetic diversity through artificial or natural pathways has given people more choices, and more efficient evolution strategies can be adopted according to research needs. This article will first introduce the CRISPR tools, then summarize the CRISPR-mediated mutation and screening platforms, and finally discuss the development trends and opportunities of the CRISPR in the field of directed evolution.

**Keywords** gene editing; CRISPR/Cas; directed evolution; CRISPR-mediated directed evolution

**Funding** This work is supported by National Key R & D Program of China (2018YFA0902701), Strategic Priority Research Program of Chinese Academy of Sciences (XDB29050501, XDPB18), National Natural Science Foundation of China (32025022, 3201101136), Shenzhen Science Technology and Innovation (KQTD2015033117210153, JCYJ20170818164139781)

## 1 引 言

随着定向进化技术的不断发展,利用传统的物理、化学和生物诱变技术<sup>[1-2]</sup>寻找所需的生物性状往往会带来过多有害突变,已无法满足研究或应用的需求。于是人们通过先对单个基因或生物途径建立突变文库,再筛选得到具有更高活性的酶<sup>[3]</sup>、目的产物产量更高的生物体<sup>[4]</sup>。但这些方法大多需要体外建立突变文库、体内进行筛选,反复迭代,这无疑增加了得到目的突变体的难度与成本。而 CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats)技术的出现和发展,使人们可以站在新的角度,利用新的工具开发定向进化技术,以解决目前所面临的问题。

借助 CRISPR/Cas 系统,可以实现对特定基因或区域进行编辑。根据小向导 RNA (small guide RNA, sgRNA) 引导的 Cas 蛋白或 Cas 融合蛋白的特性,可以使该区域基因产生双链断裂、单链缺刻和单碱基替换<sup>[5]</sup>等变化。再由修复机制、错配机制或化学催化机制引入突变,实现突变文库的构建。结合生物大分子如转录因子、核酸适配体等,CRISPR 的功能可以扩展到分子检测领域。由此可见,CRISPR 系统的性能与定向

进化技术的需求有着很高的契合度。

另外,通过 PubMed 数据库检索 CRISPR 与定向进化的论文数量发表情况可知,定向进化论文数从 1975 年开始,到 2018 年左右出现平台期,其年增长速率仅为 2%~5%。而 CRISPR 在同期正呈蓬勃发展的趋势,年增长速率达 14%~24%。因此,借助 CRISPR 为定向进化技术提供新的方法和策略,对于定向进化技术将是一个新的机遇。

本文首先介绍 CRISPR 技术的最新进展,包括不同核酸酶及诱导系统的介绍。其次,从定向进化的两个核心出发,总结利用 CRISPR 工具产生遗传多样性的策略以及借助 CRISPR 工具进行筛选与选择的方法。最后设想 CRISPR 介导的定向进化 (CRISPR-Mediated Directed Evolution, CDE) 可能的未来发展方向。

## 2 不断丰富的 CRISPR 工具箱

CRISPR/Cas 作为适应性免疫系统,存在于许多细菌和古细菌中,由规律的间隔成簇短回文重复序列组成。在 CRISPR 序列附近,可以找到编码 Cas 蛋白的基因短簇<sup>[6]</sup>。CRISPR 系统主

要分为两类, 1 类包含 I、III 和 IV 型 CRISPR 系统, 2 类包括 II、V 和 VI 型 CRISPR 系统<sup>[7]</sup>。其中, 每个类型又具有其特殊的标签 Cas 蛋白: I、III 和 IV 型的标签蛋白分别为 Cas3、Cas10 和 Csf1<sup>[8]</sup>; II、V 和 VI 型的标签蛋白则分别为 Cas9、Cas12 和 Cas13<sup>[9]</sup>。尽管许多 CRISPR 系统被相继开发出来, 但 Cas9 和 Cas12 被人们研究得最为透彻, 故本小节主要针对 Cas9 与 Cas12 系统进行介绍。

## 2.1 CRISPR-Cas9 及其衍生变体

Cas9 属于 2 类 II 型 CRISPR 系统, 该系统最初来自于化脓性链球菌, 是一种多结构域的 DNA 核酸内切酶, 包含一个 HNH 样核酸酶结构域和一个 RuvC 样核酸酶结构域, 分别负责切割与向导 RNA (guide RNA, gRNA) 序列互补的 DNA 链 (靶向链) 和靶链互补链 (非靶向链)<sup>[10]</sup>。自 Jinek 等<sup>[11]</sup>首次利用 CRISPR-Cas9 进行体外基因组编辑, 到利用 CRISPR-Cas9 系统对哺乳动物<sup>[12-13]</sup>、植物<sup>[14-15]</sup>基因组进行编辑, 人们可

以根据需要设计出 sgRNA<sup>[11]</sup>, 通过 sgRNA 与化脓性链球菌 Cas9 (*Streptococcus pyogenes* Cas9, SpCas9) 组装并结合到靶向 DNA 3' 处的前间隔序列邻近基序 (Protospacer Adjacent Motif, PAM), 实现对目的位点 DNA 的切割 (图 1(a))。根据 Cas9 的结构, 通过向 HNH 结构域引入 H840A 或者向 RuvC 结构域引入 D10A 突变, 可以得到 Cas9 缺刻酶 (Cas9 nickase, nCas9), 仅切割 DNA 双链中的其中一条单链<sup>[16-17]</sup>。当两个突变同时被引入时, 形成无活性的 Cas9 (Catalytically Deficient Cas9, dCas9), 这时 dCas9 仅有靶向作用而无 DNA 链切割活性<sup>[18-19]</sup>。

但在 SpCas9 系统中, 存在 PAM 序列限制、SpCas9 蛋白较大和脱靶等问题, 限制了其进一步的应用。早期人们通过开发其他细菌和古细菌的 Cas9 同源蛋白来改善这些问题。比 SpCas9 更小的来自金黄色葡萄球菌的 Cas9 (SaCas9) 可以识别 5'-NNGRRT-3' PAM<sup>[20]</sup>, 且当引入 D10A 和 N580A 突变时, SaCas9 同样

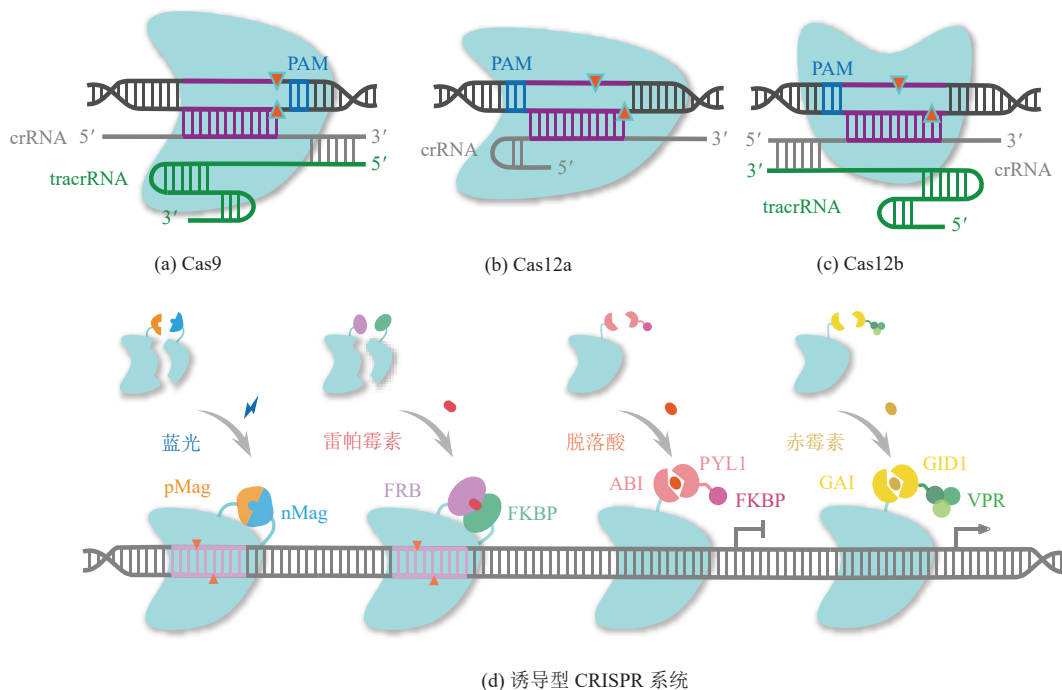


图 1 主要的 CRISPR-Cas 工具分子特征与调控

Fig. 1 Main CRISPR-Cas tools' molecule characteristic and regulation

可显示出缺刻酶活性<sup>[21]</sup>。来自脑膜炎奈瑟氏球菌 Cas9 (NmCas9) 则可以识别 5' -NNNRRT-3' 或 5' -NNNNGMTT-3' PAM 且脱靶效应降低<sup>[22]</sup>。同样还有嗜热链球菌 Cas9 (St1Cas9) 可以识别 5' -NGGNG-3' 或 5' -NNAGA AW-3'<sup>[23]</sup>，空肠弯曲杆菌 Cas9 (CjCas9) 可以识别 5' -NNNNACAC-3' 或 5' -NNNNRYAC-3'<sup>[24]</sup>，犬链球菌 (ScCas9) 可以识别 5' -NNG-3'<sup>[25]</sup>。

通过基于结构信息，借助细菌生存选择的定向进化策略与突变位点组合设计，SpCas9 VQR、EQR 和 VRER 变体可以分别识别 NGA、NGAG 和 NGCG PAM<sup>[26]</sup>。借助噬菌体辅助的连续进化，得到可以识别广泛 PAM 序列 (NG、GAA 和 GAT) 且特异性更高的 SpCas9 变体 xCas9<sup>[27]</sup>。另一种 SpCas9 变体 (SpCas9-NG) 则是通过结构设计获得识别 NG PAM 的能力<sup>[28]</sup>。SpCas9、xCas9 和 SpCas9-NG 的 PAM 均包含鸟嘌呤 (G)，但人们希望 PAM 能不受依赖 G 的限制。Miller 等<sup>[29]</sup>借助噬菌体辅助进化，得到 3 个对非 G PAM 具有活性的 SpCas9 变体，分别可以识别 NRRH、NRCH 和 NRTH PAM；利用结构导向工程策略同样得到识别非 G PAM 的 SpG (NGN PAM) 和 SpRY (NRN > NYN PAM)<sup>[30]</sup>。这些新型 SpCas9 变体的开发使得 CRISPR-Cas9 系统几乎可以靶向基因组任意位置而不再受 PAM 序列的限制。

改善 CRISPR 系统特异性和表达泄露对宿主的影响，可以对 Cas9 蛋白进行组合诱变<sup>[31]</sup>以及将 Cas9 蛋白与诱导性二聚体结构域进行融合，如光介导的 Mag 蛋白二聚化<sup>[32]</sup>、雷帕霉素诱导的 FRB 与 FKBP 互作、脱落酸及赤霉素诱导的对转录的激活或抑制<sup>[33-34]</sup>，以实现时间或空间上的表达 (图 1(d))。

## 2.2 CRISPR-Cas12 家族

Cas12a (Cpf1) 属于 2 类 V 型核酸内切酶，典型的包括氨基酸球菌 Cas12a (AsCas12a) 和毛螺菌 ND2006 Cas12a (LbCas12a)，虽然

弗朗西斯菌 Cas12a (FnCas12a) 和食肉莫拉菌 Cas12a (MbCas12a) 也被报道，但活性较不稳定<sup>[35]</sup>。与 Cas9 不同，Cas12a 仅需要 CRISPR RNA (crRNA)，而无需反式激活 crRNA (trans-activating CRISPR RNA, tracrRNA) 作为引导，识别富含胸腺嘧啶 (T) 的 PAM，并产生 5' 黏性末端的 DNA 双链断裂<sup>[36]</sup> (图 1(b))。Cas12a 也被成功地进行改造以获得更广泛的 PAM 位点，其中携带 S542R/K607R 和 K548V/N552 突变的 AsCpf1 分别识别 TYCV 和 TATV PAM，E174R/S542R/K548R 变体 (enAsCas12a) 可以针对多种 PAM，如 TTYN、VTTV 和 TRTV 等。与野生型 AsCas12a 相比，E174R/S542R 变体与 enAsCas12a 在 TTTV PAM 上的编辑效率可以提高 2 倍<sup>[37-38]</sup>。Cas12b (C2c1) 是一种含有 RuvC 结构域，不含 HNH 结构域的核酸内切酶，它需要 crRNA 和 tracrRNA 识别靶标 DNA<sup>[9,39]</sup>。来源于嗜酸热脂环酸杆菌的 Cas12b 系统 (AaCas12b) (图 1(c))，其由于比 Cas9 和 Cas12a 更小，具有更宽的耐受温度，已成功用于哺乳动物和植物的基因组工程研究<sup>[40-41]</sup>。还有如 Cas12e<sup>[42]</sup>、Cas12d<sup>[43]</sup>、Cas12g、Cas12i、Cas12c 和 Cas12h<sup>[44]</sup>，均被发现并进行了表征。

## 3 利用 CRISPR 工具实现的遗传多样性

CRISPR 工具首先作为基因编辑系统，可以与 DNA 修复机制串联，既可以对单基因进行突变，也可以在基因组水平引入广泛的变化。同时 CRISPR 作为招募平台，通过蛋白质融合或 sgRNA 蛋白相互作用，将功能域募集到特定位点，也可以在目标基因中引入突变。

### 3.1 通过 CRISPR 产生双链断裂进行突变库的建立

利用 CRISPR 系统可以对靶向 DNA 序列进行切割的特性，并激活 DNA 修复机制，可以在



基因组 DNA 中引入多样性的突变, 表 1 总结了基于该原理的定向进化平台。通过 Cas9-sgRNA 复合物对 DNA 双链的切割作用产生双链断裂 (Double-Strand Break, DSB), 而 DSB 的修复主要通过非同源末端连接 (non-Homologous End Joining, NHEJ) 途径和同源性修复 (Homology-Directed Repair, HDR) 途径完成<sup>[45]</sup> (图 2(a))。其中, NHEJ 不依赖同源模板<sup>[46]</sup>, 通常产生较小的插入或缺失, 导致不可控的随机突变, 对单核苷酸取代能力不足<sup>[47]</sup>。HDR 途径则需要外源双链或单链寡脱氧核苷酸 (double-strand/single strand DNA, dsDNA/ssDNA) 提供模板, 所以其擅长利用携带突变的供体模板进行单核苷酸的替换, 但效率一般较低<sup>[48]</sup>。由于依赖 HDR 途径可以获得更为精确的人工设计突变, 不会造成基因大片的缺失, 于是人们开展了提高 HDR 效率的研究, 如调节修复机制<sup>[49-50]</sup>、合理设计供体模板等<sup>[51-52]</sup>。

### 3.1.1 大肠杆菌底盘系统

在原核生物中, CRISPR 介导的 DNA 链切割可以杀死未被编辑过的细胞, 从而可以增加筛选效率又避免使用抗性标记<sup>[53]</sup>。Wang 等<sup>[54]</sup>借助  $\lambda$ Red 与 dsDNA/ssDNA 开发了多重自动化基因组工程 (Multiplex Automated Genome Engineering, MAGE), 该系统可同时靶向染色体多个位置。将 CRISPR-Cas9 引入该系统则可以提高较大核糖体结合位点 (Ribosome Binding Site, RBS) 的替换效率 (与 MAGE 相比提高了 56%), 并且在

单轮循环中引入单突变和双突变的效率比 MAGE 分别高 93% 和 90%<sup>[55]</sup>。Li 等<sup>[56]</sup>报道了一种基于 CRISPR-Cas9 的迭代基因组编辑系统。该系统以 dsDNA 作为编辑模板, 实现基因缺失、插入和替换的效率可接近 100%, 随后针对  $\beta$ -胡萝卜素, 进行了甲基赤藓糖醇磷酸和中心代谢途径的组合优化, 筛选得到的最优菌株在发酵中可产出 2.0 g/L 的  $\beta$ -胡萝卜素。

随着 DNA 合成与测序成本的降低, 人们可以借助高通量方法来在定向进化过程中产生突变并分析表型 (图 2(b))。结合大规模并行寡核苷酸合成和测序, Garst 等<sup>[57]</sup>开发了可在全基因组范围内产生能追踪的编辑平台 (CRISPR-Enabled Trackable Genome Engineering, CREATE)。该平台的核心设计是将每个 gRNA 连接到同源修复盒上, 该盒既可以编辑基因位点, 又可以作为条形码来追踪基因型与表型关系, 实现引导切割的 gRNA 和同源修复模板序列的偶联。借助 CREATE 可以对大肠杆菌耐热和耐药性的相关基因绘制超过 50 000 个全基因组突变, 并鉴定其赋予目标表型的突变。利用迭代 CREATE (iCREATE) 系统<sup>[58]</sup>, 研究人员为得到高葡萄糖和木糖利用率且抗水解产物抑制的大肠杆菌菌株以用于 3-羟基丙酸 (3-Hydroxypropionic Acid, 3HP) 生产, 针对 30 个基因中约 40 000 个突变设计文库, 得到的突变株可以在高水解产物浓度下达到亲本株 3HP 产量的 7~8 倍。同时 Liu 等<sup>[59]</sup>又针对 115 个基因中约 162 000 个突变

表 1 基于双链断裂的 CDE 策略

Table 1 CDE strategy based on double-strand break

系统	底盘	供体	描述	参考文献
CRMAGE	大肠杆菌	ssDNA	单点突变效率可达到 95% 以上, RBS 替换效率为 66%, 同时引入双突变效率为 90% 以上	[55]
CREATE	大肠杆菌	dsDNA	可构建 $10^4 \sim 10^5$ 大小的文库, 单点突变效率为 70% 以上, 种群中可追踪的基因座数量为 50 000 个	[57]
CRISPRm	酿酒酵母	dsDNA	顺式多重化效率 100%, 反式多重化效率双位点为 49%, 三位点为 19%	[60]
CasPER	酿酒酵母	dsDNA	针对 300~600 bp 的大片段, 单基因整合效率达到 98%~99%, 双基因整合效率为 98%~99%	[61]
CHAnGE	酿酒酵母	dsDNA	针对 7 459 个开放阅读框建立了 24 765 大小的文库, 靶序列编辑效率为 82%	[62]
CRISPRres	肿瘤细胞	无	针对 9 个基因设计了包含 2 209 个 sgRNA 的文库	[63]

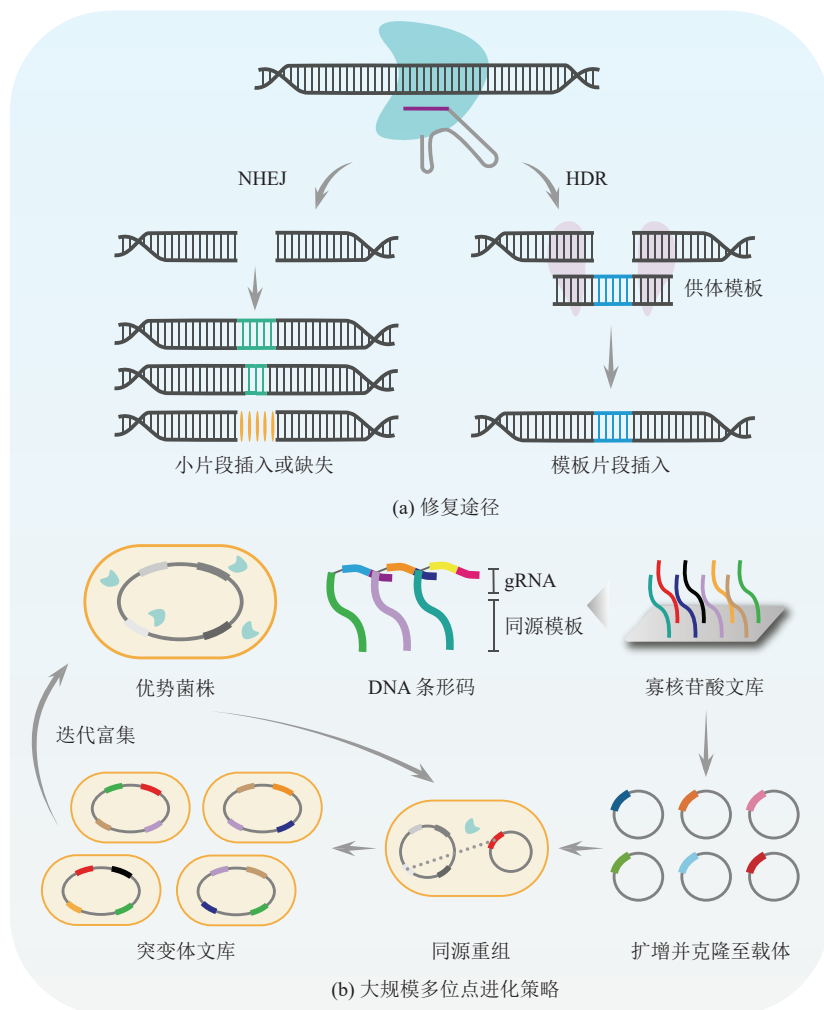


图 2 DNA 修复途径介导的定向进化策略

Fig. 2 DNA repair pathway mediated directed evolution strategy

设计文库，优化了 3HP 合成代谢相关途径，使突变株 3HP 产量高出野生型菌株 60 倍。这种系统的局限性是在追踪基因型与表型的关系时，以质粒为载体的条形码稳定性较低，但可以采取基因组整合的办法解决该问题。

### 3.1.2 酵母底盘系统

酵母的代谢途径工程研究对于生物燃料、化学药品的生成具有重要意义，故人们开展了酵母中的基因组多样性研究。早期人们使用的突变文库多为可以在酵母细胞内自我复制的环状质粒，这种着丝粒质粒在单细胞水平拷贝数差异很大，相当于在表型与基因型之间增加了拷贝数的干扰

项，这会导致定向进化的效率大大降低。

Ryan 等<sup>[60]</sup>在酵母中建立了 CRISPRm (Multiplex CRISPR System)，该系统以 dsDNA 为模板，可在二倍体或多倍体酿酒酵母基因组中创建多个 DSB。另外，为增加 CRISPRm 的编辑效率，将 sgRNA 与自切割肺炎三角洲病毒核酶 3' 端融合以保护 sgRNA 免受核酸外切酶的影响。最后在单轮定向进化中筛选出纤维糊精转运蛋白变体，将纤维二糖到酒精的发酵速率提高了 10 倍。

对于多基因相关代谢途径的优化，可以利用 CRISPR 系统对酿酒酵母基因组中甲羟戊酸代谢相关的 5 个基因进行组合编辑<sup>[64]</sup>，得到 31 个突

变株, 其中产量最高的突变菌株比野生型菌株高 41 倍。CasPER (Cas9-Mediated Protein Evolution Reaction)<sup>[61]</sup> 是一种 CRISPR-Cas9 介导的大序列定向进化方法, 可以整合 300~600 bp 的供体模板, 其首先通过易错 PCR 得到甲羟戊酸途径两个关键代谢酶的突变文库, 再利用 CRISPR 系统在靶标处引入 DSB 后发生的 HDR, 使突变文库整合到靶向位点, 最后得到了异戊二烯产量增加 11 倍的突变体。

借助高通量寡核苷酸合成与测序技术同样可以实现酵母基因组范围的定向进化研究。Guo 等<sup>[65]</sup>介绍了一种基于 CRISPR-Cas9 的策略用于在酵母细胞中同时创建数百个无选择标记的遗传变异, 并针对 315 个特性不明的小型开放阅读框进行评估, 研究这些小型开放阅读框对环境适应性生长的必要性, 讨论了该策略用于定向进化的可能性。CRISPR-Cas9 与 HDR 辅助的基因组规模工程 (CRISPR-Cas9 and Homology-Directed-Repair-Assisted Genome-Scale Engineering, CHAnGE)<sup>[62]</sup>可以实现单核苷酸精度的全基因组突变, 进化出对生长抑制剂 (如糠醛、乙酸) 具有耐受性的菌株。在酵母中也可以使用类似大肠杆菌中 CREATE 的策略, 利用条形码对每个引导-供体对序列进行标记, 实现对突变文库的高通量测序和功能分析<sup>[66]</sup>。同时, 该实验还发现利用融合蛋白 LexA-Fkh1p 将供体 DNA 招募到断裂位点, 可以将编辑效率提高 5 倍以上<sup>[66]</sup>。

在基因组水平对基因表达水平进行单一类型的调节 (上调、下调或敲除) 是不全面的, 因为其无法研究具有不同表达水平基因间的相互作用对复杂表型的影响。多功能基因组水平 CRISPR 系统 (Multi-Functional Genome-Wide CRISPR, MAGIC)<sup>[67]</sup>可以有效解决这一问题。MAGIC 使用 3 个正交的 Cas 蛋白开发了三功能 CRISPR 系统, 将基因激活、干扰和缺失整合在一起。与 CHAnGE 相比, 在相同筛选条件下 MAGIC 可以

发现更多糠醛耐受性相关基因间的相互作用, 进化出糠醛耐受性菌株的效率更高。

### 3.1.3 植物与动物底盘系统

尽管在酵母中 HDR 途径非常活跃<sup>[68]</sup>, 但是在植物和哺乳动物产生 DSB 后首个反应是 NHEJ。主要原因是动植物中 NHEJ 途径发生的概率要高于 HDR, 且难以传递足够的供体 DNA (修复模板)。目前利用 NHEJ 途径可以实现广泛的动植物基因缺失文库的建立<sup>[69-73]</sup>。尽管多数条件下 NHEJ 途径会导致基因的插入和缺失, 使开放阅读框出现移码突变, 但借助 NHEJ 开发定向进化平台仍然是可行的。例如, 在水稻中开发的 CRISPR-Cas9 定向进化平台<sup>[74]</sup>, 可以进化 SF3B1 剪接体蛋白, 以抵抗剪接抑制剂。其 sgRNA 文库靶向必需基因 *SF3B1*, 并从 15 000 个愈伤组织中得到 6 个抗剪接抑制剂 (GEX1A) 的 SF3B1 突变体。

在哺乳动物底盘中, NHEJ 产生的基因多样性可以用于细胞耐药性研究。CRISPRres (CRISPR-Induced Resistance in Essential Genes) 系统<sup>[63]</sup>可以利用大规模的 sgRNA 文库进行癌细胞抗药性的遗传筛选, 快速获得和鉴定必需基因中的耐药性突变。为验证系统的有效性, 结合蛋白晶体结构, 解析了烟酰胺磷酸核糖基转移酶突变前后与抗癌药 KPT-9274 的结合机制。Ipsaro 等<sup>[75]</sup>对赖氨酸甲基转移酶 (DOT1L) 的 KMT 结构域设计了包含 73 个 sgRNA 的文库, 并在 EPZ-5676 抑制剂的选择作用下得到具有 VVEL293MM (两个蛋氨酸替换为两个缬氨酸, 同时缺失谷氨酸和赖氨酸残基) 突变的 DOT1L 变体, 生化分析发现该突变可以使酶活增加, 同时降低对 EPZ-5676 抑制剂的敏感性。

对于抗体工程, NHEJ 无法满足精确位点突变的需求, 所以人们开发了利用 HDR 获得高效抗体的研究。早期, 研究人员利用 CRISPR-Cas9, 针对 BRCA1 第 18 个外显子中 6 个碱基<sup>[76]</sup>,

克隆了包含随机六聚体的 HDR 文库,证明了利用借助 HDR 对基因组区域进行饱和编辑的可能性。Mason 等<sup>[77]</sup>利用单链寡核苷酸作为供体模板,将文库引入抗鸡蛋溶菌酶抗体(HEL3)重链可变域(VH)的互补决定区 3(CDRH3),使用定向进化和高通量筛选方法获得了新的 CDRH3 序列。将 DNA 供体形式由 ssDNA 变为可自线性化质粒,可以增加 HDR 的效率,再借助易错 PCR 对 HEL3 的 VH 区生成随机突变文库,筛选出亲和力范围在皮摩尔范围内的变体<sup>[78]</sup>。

### 3.2 CRISPR 系统与具有诱变功能的蛋白联用

借助 CRISPR 系统对基因组产生 DSB 后修复途径,往往会对宿主本身造成损伤,并且受到需要提供 HDR 供体模板的限制,而利用 Cas9 的引导功能,引导诱变功能蛋白可以解决这些问题<sup>[79]</sup>。碱基编辑器是由 DNA 靶向模块(dCas9 或 nCas9)与催化结构域(胞嘧啶或腺苷脱氨酶<sup>[80-81]</sup>)组成的融合蛋白,其中靶向模块可以结合到特定的 DNA 位点,随后 sgRNA 与靶 DNA 结合形成“R-环”<sup>[82]</sup>,导致一小段单链 DNA 移位,这一小段单链 DNA 作为催化窗口被催化结构域进行碱基替换。由于碱基编辑的 DNA 靶向模块使用的是 dCas9 或 nCas9,所以并不会导致 DSB 的产生。借助目前开发的各类 CRISPR 融合蛋白平台,可以实现单碱基的变化、局部序列的多样化<sup>[83-84]</sup>,用于生成新的蛋白质突变体或作物改

良<sup>[85-86]</sup>等。表 2 为相应平台的汇总描述。

#### 3.2.1 碱基编辑工具的开发

胞嘧啶碱基编辑器(Cytosine Base Editor, CBE)可以令胞嘧啶(C)脱去氨基,使其转化为尿嘧啶(U),从而将 C·G 碱基对转化为 T·A 碱基对(图 3(a))。第一代碱基编辑器由来自大鼠的胞嘧啶脱氨酶 APOBEC1,通过连接肽与 dCas9 连接。以 PAM 的远端作为位置 1,其有效脱氨窗口在距 PAM 的 4~8 位。由于尿嘧啶糖基化酶(Uracil DNA Glycosylase, UDG)经常会除去尿嘧啶,导致编辑效率较低(体外脱氨效率为 25%~40%,细胞内脱氨效率为 0.9%~7.7%),所以在 dCas9 蛋白的 C 末端偶联了 UDG 的抑制剂 UGI(Uracil DNA Glycosylase Inhibitor),开发了第二代碱基编辑器,使体内编辑效率与第一代碱基编辑器相比提高了 3 倍。而如果将 dCas9 替换为 nCas9(D10A),则可将第二代碱基编辑器效率进一步提升 2~6 倍<sup>[80]</sup>。借助丰富的 CRISPR 工具箱,人们开发了诸多针对不同 PAM 或者编辑窗口的碱基编辑器。如利用 SpCas9 变体(VQR-BE3、VRER-BE3)分别识别 NGAN 和 NGCG PAM,编辑窗口分别为 4~11 和 3~10 位<sup>[90]</sup>;利用不同来源的 Cas 蛋白(Sa-BE3、SaKKH-BE3 和 Cas12a-BE)分别识别 NNGRRT、NNNRRT 和 TTTV<sup>[90-91]</sup>。

腺嘌呤碱基编辑器(Adenine Base Editor,

表 2 基于与诱变功能蛋白联用 CDE 策略

Table 2 CDE strategy based on coupling protein with mutation

系统	Cas 类型	效应蛋白	描述	参考文献
CRISPR-X	dCas9	MS2-AID*Δ	PAM 上游-50~+50 bp 为编辑窗口,诱变率约 $1 \times 10^{-3}$ /碱基/代	[83]
TAM	dCas9	AIDx	PAM 上游-16 和-12 bp 处有 U 最高的核苷酸取代频率,诱变率约为 $4 \times 10^{-4}$ /碱基/代	[84]
CRISPR-BEST	nCas9	APOBEC1, TadA	以 TC>CC>AC>GC 的优先级接受其脱氨底物,编辑窗口为 PAM 上游-17~-11 bp,在-16、-15、-14 bp 处编辑效率最高	[87]
STEME	nCas9	APOBEC3A-ecTadA-ecTadA*	编辑窗口 C>T 效率为 0.10%~61.61%,可覆盖 C1~C17 区域,A>G 效率为 0.69%~15.50%,可覆盖 A4~A8 区域。同时发生 C>T 和 A>G 效率为 0.49%~15.10%	[88]
EvolvR	enCas9	Poll3M	编辑窗口长度最多为 350 bp,靶向位点突变率约 $5 \times 10^{-6}$ /碱基/代	[89]



ABE)则是可以将 T·A 转化为 C·G (图 3(b))。与 CBE 类似, 人们开展了一系列 ABE 的优化及应用<sup>[92-94]</sup>。例如, Wang 等<sup>[95]</sup>通过对 ABE 进行核心模块组合优化(包括 sgRNA、nCas9、脱氨酶与核定位信号肽), 获得了具有更精确的编辑窗口、更高编辑效率的 sABE 系统。

### 3.2.2 基于碱基编辑的定向进化技术

定向进化的实现需要借助单碱基取代和局部序列多样化, 创建遗传突变库<sup>[96]</sup>。CRISPR-X<sup>[83]</sup>和 TAM (Targeted AID-mediated Mutagenesis)<sup>[84]</sup>是两个碱基编辑介导的产生局部序列多样化的平台。其中, CRISPR-X 使用的 sgRNA 融合有两个 MS2 结合适体, 可以招募 MS2-AID (Activation Induced Cytidine Deaminase, AID) 融合蛋白实现对局部区域的诱变(图 3(c))。Hess 等<sup>[83]</sup>对野生型绿色荧光蛋白进行诱变, 得到 S65T 和 Q80H 的绿色荧光蛋白变体(加强型绿色荧光蛋白)。借助 CRISPR-X 同样可以进行耐药机制的研究: PSMB5 是蛋白酶抑制剂硼替佐米的靶标, 作者通过迭代进化发现了 5 个对硼替佐米强抗性突变与 3 个适度抗性突变。在 TAM 系统中(图 3(d)), 研究人员则在慢性髓性白血病细胞中用 dCas9-AID-P182X (AIDx) 靶标 *BCR-ABL* 基因, 有效鉴定了赋予细胞伊马替尼抗性的已知突变和新突变<sup>[84]</sup>。除了耐药性研究, 将碱基编辑工具与哺乳动物细胞展示系统结合使用, 可以在多轮突变和选择下得到更高亲和力的抗体<sup>[97]</sup>。

单纯依靠 CBE 或 ABE 可能无法产生足够丰富的突变类型, 而 CBE 和 ABE 的联合使用可以有效解决这一问题。Liu 等<sup>[98]</sup>通过对编码 *OsACC* 羧基转移酶结构域基因设计 141 条 sgRNA, 结合两种 CBEV (eBE3 和 eCDA), 及一种 ABE (eABE), 分别构建了 3 个碱基编辑文库。从 eABE 文库中获得了未报道过的 I1879V 突变, 赋予水稻除草剂抗性, 而另外两个文库未获得新的突变。碱基编辑器介导的基因进化技术(a Base-

Editing-Mediated Gene Evolution, BEMGE)<sup>[99]</sup>则是将胞嘧啶碱基编辑器(rBE9)和腺嘌呤碱基编辑器(rBE14)同时转入水稻细胞, 对水稻除草剂靶标基因 *OsALS1* 进行靶向诱变, 获得 4 个自然界中未曾被发现的、对除草剂双嘧菌钠具有抗性的蛋白变体。类似的策略同样被用于在放线菌的单碱基编辑。CRISPR-BEST (CRISPR-base Editing System)<sup>[87]</sup>包含两个独立的碱基编辑器(CRISPR-cBEST 和 CRISPR-aBEST), 并在非模式链霉菌中通过引入终止密码子失活了奇霉素合成途径中两个拷贝的 *kirN* 基因, 同时该研究借助 Csy4 (1 类 IF 型核糖核酸内切酶) 设计了 sgRNA 多重化系统, 可以实现仅依靠一对启动子和终止子表达多个 sgRNA, 且每个 sgRNA 表达强度相同, 显示出强大的用于代谢工程类定向进化的潜

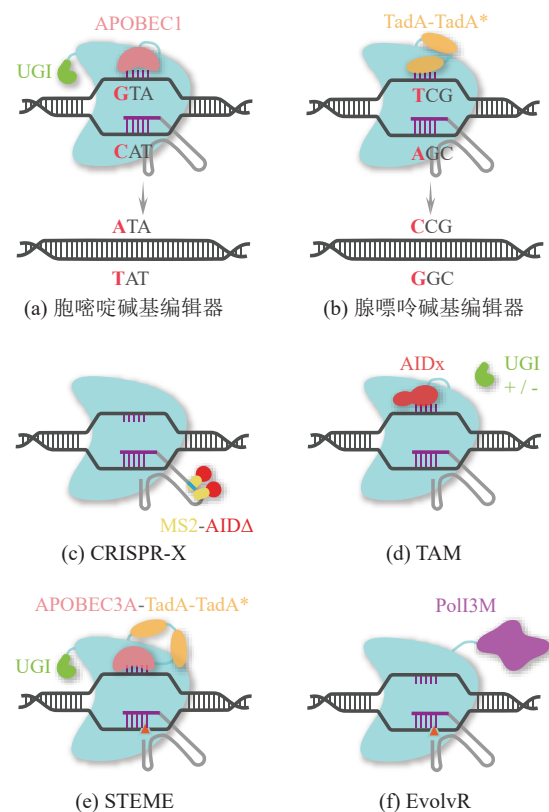


图 3 Cas9 与诱变蛋白联用介导的定向进化

Fig. 3 Cas9 and protein with mutation coupling mediated directed evolution

力。饱和靶向内源基因突变碱基编辑器 (Saturated Targeted Endogenous Mutagenesis Editors, STEME)<sup>[88]</sup>则将 CBE 与 ABE 系统进行合并融合 (图 3(e)), 通过将胞嘧啶脱氨酶 (APOBEC3A) 和腺嘌呤脱氨酶 (ecTadA-ecTadA7.10) 同时融合在 nCas9 的 N 端, 并将 UGI 融合至 nCas9 (D10A) 的 C 端, 实现只在一个 sgRNA 引导下就可以诱导靶位点 C>T 和 A>G 的同时突变, 显著增加靶基因碱基突变的饱和度与多样性, 最后验证 STEME 在植物中的定向进化能力, 获得了 OsACC 羧基转移酶结构域 3 个对氟吡甲禾灵具有抗性的新突变位点。

### 3.2.3 Cas-DNA 聚合酶定向进化技术

除了利用碱基编辑器进行单碱基的转换, 借助 CRISPR 系统来招募其他诱变蛋白同样可以产生局部序列多样化。EvolvR 系统<sup>[89]</sup>将 nCas9 与易错 DNA 聚合酶 I (PolI3M) 进行融合 (图 3(f)), 利用 nCas9 创建的 DNA 缺刻作为 DNA 聚合酶起始位点, 诱变窗口长度、突变率和取代偏差则由聚合酶变体本身性质决定。该系统可以产生高于野生型细胞 770 多万倍的突变率, 编辑 350 个核苷酸。最后使用 EvolvR 鉴定对抗生素壮观霉素耐药的新型核糖体突变位点。借助类似的原理, Tou 等<sup>[100]</sup>在酿酒酵母中也开发了 EvolvR 系统。

## 4 CRISPR 介导的筛选与选择

对多样化突变文库进行筛选与选择是定向进化中另一个关键部分。筛选可以对每种突变体的特性进行评估, 而选择则只会产生理想的突变体<sup>[101]</sup>。筛选技术可以减少错过所需突变体的机会, 但降低通量; 选择技术则可以用于更大的文库, 但需要对突变文库施加选择压力来消除非目的突变体<sup>[102]</sup>。CRISPR 系统除了具有出色的基因编辑能力外, 还已应用于分子检测。而筛选和选择方法恰恰依赖分子检测等检测技

术的开发, 因为大多数生物分子与可直接观察的表型无关<sup>[96]</sup>, 需要荧光、比色或其他报告分子。例如, 借助气相色谱检测产物的形成, 获得具有高活性细胞色素 P450 酶变体<sup>[103]</sup>, 通过一连串酶促反应来产生与纤维素酶活相对应的荧光<sup>[104]</sup>等。目前, CRISPR 系统被广泛地应用于核酸分子的检测<sup>[105-106]</sup>, 在 2019 年也出现了对非核酸靶标的检测系统, 所以本节重点对 CRISPR 介导的筛选检测技术进行阐述, 在未来展望部分则对 CRISPR 介导的选择进行讨论。

Liang 等<sup>[107]</sup>开发了 CRISPR-Cas12a 和细菌变构转录因子 (allosteric Transcription Factors, aTFs) 介导的小分子检测平台 CAT-SMellor (CRISPR-Cas12a- and aTF-mediated Small Molecule Detector), 借助 CRISPR-Cas12a 的单链 DNA 切割能力与 aTFs 对小分子和双链 DNA 的竞争结合活性, 当目的小分子存在时, aTF 的构象发生改变, 导致 dsDNA 从 aTF 结合域解离, dsDNA 激活 Cas12a 对携带有荧光团-淬灭剂 (Fluorophore Quencher, FQ) 标记的 ssDNA 的切割, 检测荧光信号发生变化, 实现对尿酸和对羟基苯甲酸的纳摩尔水平的高通量检测。另一种 CRISPR 介导的传感器则是借助功能性 DNA (functional DNA, fDNA)。当 fDNA 的靶标出现时, 单链 DNA 激活剂与 fDNA 解聚, 触发 Cas12a 对标记有 FQ 的单链 DNA 进行裂解, 进而出现可检测的荧光信号, 该策略也成功地用于检测 ATP 和  $\text{Na}^+$ <sup>[108]</sup>。类似地, Li 等<sup>[109]</sup>同样借助 fDNA 实现了对  $\text{Pb}^{2+}$ 、鲍曼不动杆菌和微小 RNA 的检测, 再次扩充了 CRISPR 系统对非核酸分子的检测范围。

## 5 总结与展望

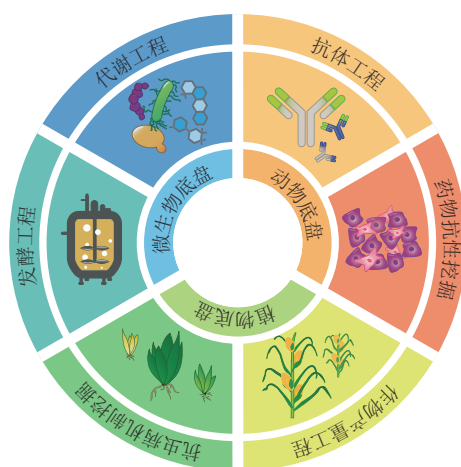
CRISPR 技术的出现和发展已经彻底改变人们对基因进行编辑、表达调控、分子检测的能力,

出现了一系列基于 CRISPR 技术的拓展工具, 这也促进定向进化技术的发展, 并已应用于代谢工程、抗体工程、植物育种等领域(图 4(a))。不同来源的 CRISPR 系统及其相关变体, 已经可以较为有效地解决如 PAM 位点限制<sup>[110]</sup>、脱靶<sup>[111]</sup>等问题, 为进一步多基因间协同进化(如人造基因回路或代谢通路)提供可能。利用大规模 gRNA 文库及高通量测序, 结合 DNA 条形码技术, 可以更高效地追踪与获得基因组层面上与突变相关的遗传信息, 提高定向进化效率。同时通过机器学习与建立模型<sup>[112]</sup>, 有希望从获得的原始数据对进化模型进行完善, 或通过逻辑线路设计实现对基因突变(如重复序列导致的缺失)的追踪与定量<sup>[113]</sup>, 得到普适性的系统突变库与目的性状的定量关系, 帮助人们更好地借助 CRISPR 工具箱对定向进化进行操纵和预测。

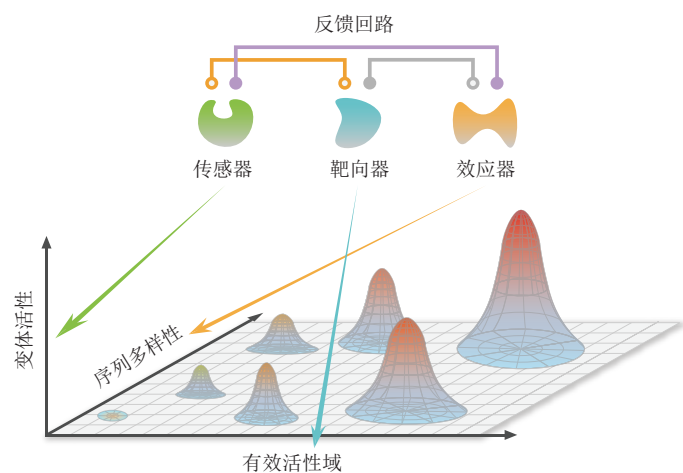
借助 CRISPR 系统对 DNA 的切割活性或靶向活性, 可以实现多样化突变文库的建立。在 CRISPR 介导的修复途径的诱变中, NHEJ 和 HDR 途径产生的突变文库各有利弊, 选择哪种途径要取决于宿主本身性质和研究需求。因此对修复途径进行更深入的研究, 规避不同修复途

径的劣势, 对于 CRISPR 介导的定向进化非常重要。而对于 Cas 蛋白将诱变功能蛋白引入靶向位点的策略, 可以避免出现双链断裂对宿主造成不利影响, 实现对特定窗口核苷酸进行突变。另外, 自 2017 年以来, Cas13 系统也展示出巨大的应用潜力, 其 RNA 的靶向能力将给研究者在 RNA 水平开发定向进化系统提供可能。与 DNA 文库相比, RNA 文库多样性更强, 但是如何将活性 RNA 突变体映射到 DNA 上并迭代遗传是其开发难点。

基于 CRISPR 的检测技术已经拓展到了多种小分子, 这为开发新型筛选与选择系统提供了可能。如在选择方法的展示技术中, 已提出的 CRISPR 展示系统可以构建结合有蛋白质、人工适配体、随机序列库或长非编码 RNA 的 Cas9 蛋白复合物<sup>[114]</sup>, 在定向进化的高通量选择方法中具有广阔的应用前景。同时, 借助 Cas9 对 DNA 的切割宿主致死性, 可以进一步开发基于生存和生长的选择方法, 而不是仅依靠药物或抗生素选择导致的淘汰, 拓宽可研究表型。目前已有研究利用该性质进行微生物种群水平的选择性调控<sup>[115]</sup>, 以及结合致死基因, 设计基于 CRISPR



(a) CRISPR 介导的定向进化平台应用



(b) 定向进化活性空间与 CRISPR 的全局关系

图 4 CRISPR 介导定向进化的应用与机遇

Fig. 4 Application and opportunities of CRISPR-mediated directed evolution



的正负选择策略用于 Cas9 蛋白的定向进化<sup>[116]</sup>。由此可见, CRISPR 技术在定向进化中筛选与选择策略上仍有很大的发展空间。

CRISPR 的多用途性, 能更好地实现基因型-表型的联系, 实现将建立突变文库和选择与筛选的偶联, 并帮助已有定向进化技术进行改良。Cas 蛋白作为模块化程度极高的核酸靶向平台, 可以通过串联多种功能蛋白来满足定向进化的多种需求, 实现更高的序列多样性、有效活性域可及性。Cas 蛋白介导的有效活性域可及性将提高突变位点的精准性, 致突变功能蛋白(如易错 DNA 聚合酶、脱氨酶等)将提高序列的多样性。二者共同决定了待进化分子的进化潜力。利用恰当的传感器对变体活性进行筛选, 并开发基于传感器-靶向器-效应器之间反馈基因回路的新颖可迭代定向进化技术(如当筛选系统检测到高活性变体在种群中占比较高时, 线路会产生反馈, 进而调节效应器与靶向器的效率, 使得高活性变体在迭代过程中具有更小突变压力的同时, 靶向器可对更广泛的位点进行探索), 将是未来的发展方向之一(图 4(b))。

未来, CRISPR 技术将加速定向进化相关领域的发展, CRISPR 工具箱将为人们提供多元化工具以解决定向进化中的问题并使定向进化技术设计进入全新的阶段。值得一提的是, 也可以借助定向进化手段对 CRISPR 系统进行优化, 而 CRISPR 技术又能改善定向进化策略, 显示出二者潜在的紧密联系。

### 参 考 文 献

- [1] Li X, Song YJ, Century K, et al. A fast neutron deletion mutagenesis-based reverse genetics system for plants [J]. *The Plant Journal*, 2001, 27(3): 235-242.
- [2] Mittelsten Scheid O, Afsar K, Paszkowski J. Release of epigenetic gene silencing by trans-acting mutations in *Arabidopsis* [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 1998, 95(2): 632-637.
- [3] Zhao HM, Chockalingam K, Chen ZL. Directed evolution of enzymes and pathways for industrial biocatalysis [J]. *Current Opinion in Biotechnology*, 2002, 13(2): 104-110.
- [4] Johannes TW, Zhao HM. Directed evolution of enzymes and biosynthetic pathways [J]. *Current Opinion in Microbiology*, 2006, 9(3): 261-267.
- [5] Butt H, Zaidi SEA, Hassan N, et al. CRISPR-based directed evolution for crop improvement [J]. *Trends in Biotechnology*, 2020, 38(3): 236-240.
- [6] Makarova KS, Haft DH, Barrangou R, et al. Evolution and classification of the CRISPR-Cas systems [J]. *Nature Reviews Microbiology*, 2011, 9(6): 467-477.
- [7] Koonin EV, Makarova KS, Zhang F. Diversity, classification and evolution of CRISPR-Cas systems [J]. *Current Opinion in Microbiology*, 2017, 37: 67-78.
- [8] Makarova KS, Wolf YI, Alkhnbashi OS, et al. An updated evolutionary classification of CRISPR-Cas systems [J]. *Nature Reviews Microbiology*, 2015, 13(11): 722-736.
- [9] Shmakov S, Abudayyeh OO, Makarova KS, et al. Discovery and functional characterization of diverse class 2 CRISPR-Cas systems [J]. *Molecular Cell*, 2015, 60(3): 385-397.
- [10] Jiang FG, Doudna JA. CRISPR-Cas9 structures and mechanisms [J]. *Annual Review of Biophysics*, 2017, 46(1): 505-529.
- [11] 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.
- [12] Cong L, Ran FA, Cox D, et al. Multiplex genome engineering using CRISPR/Cas systems [J]. *Science*, 2013, 339(6121): 819-823.
- [13] Mali P, Yang LH, Esvelt KM, et al. RNA-guided human genome engineering via Cas9 [J]. *Science*, 2013, 339(6121): 823-826.
- [14] Li JF, Norville JE, Aach J, et al. Multiplex and homologous recombination-mediated genome



- editing in *Arabidopsis* and *Nicotiana benthamiana* using guide RNA and Cas9 [J]. *Nature Biotechnology*, 2013, 31(8): 688-691.
- [15] Nekrasov V, Staskawicz B, Weigel D, et al. Targeted mutagenesis in the model plant *Nicotiana benthamiana* using Cas9 RNA-guided endonuclease [J]. *Nature Biotechnology*, 2013, 31(8): 691-693.
- [16] Ran FA, Hsu PD, Lin CY, et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity [J]. *Cell*, 2013, 154(6): 1380-1389.
- [17] Mali P, Aach J, Stranges PB, et al. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering [J]. *Nature Biotechnology*, 2013, 31(9): 833-838.
- [18] Gilbert LA, Larson MH, Morsut L, et al. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes [J]. *Cell*, 2013, 154(2): 442-451.
- [19] Dominguez AA, Lim WA, Qi LS. Beyond editing: repurposing CRISPR-Cas9 for precision genome regulation and interrogation [J]. *Nature Reviews Molecular Cell Biology*, 2016, 17(1): 5-15.
- [20] Nishimasu H, Cong L, Yan WX, et al. Crystal structure of *Staphylococcus aureus* Cas9 [J]. *Cell*, 2015, 162(5): 1113-1126.
- [21] Friedland AE, Baral R, Singhal P, et al. Characterization of *Staphylococcus aureus* Cas9: a smaller Cas9 for all-in-one adeno-associated virus delivery and paired nickase applications [J]. *Genome Biology*, 2015, 16(1): 257.
- [22] Hou ZG, Zhang Y, Propson NE, et al. Efficient genome engineering in human pluripotent stem cells using Cas9 from *Neisseria meningitidis* [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2013, 110(39): 15644-15649.
- [23] Karvelis T, Gasiunas G, Miksys A, et al. crRNA and tracrRNA guide Cas9-mediated DNA interference in *Streptococcus thermophilus* [J]. *RNA Biology*, 2013, 10(5): 841-851.
- [24] Kim E, Koo T, Park SW, et al. *In vivo* genome editing with a small Cas9 orthologue derived from *Campylobacter jejuni* [J]. *Nature Communications*, 2017, 8(1): 14500.
- [25] Chatterjee P, Jakimo N, Jacobson JM. Minimal PAM specificity of a highly similar SpCas9 ortholog [J]. *Science Advances*, 2018, 4(10): eaau0766.
- [26] Kleinstiver BP, Prew MS, Tsai SQ, et al. Engineered CRISPR-Cas9 nucleases with altered PAM specificities [J]. *Nature*, 2015, 523(7561): 481-485.
- [27] Hu JH, Miller SM, Geurts MH, et al. Evolved Cas9 variants with broad PAM compatibility and high DNA specificity [J]. *Nature*, 2018, 556(7699): 57-63.
- [28] Nishimasu H, Shi X, Ishiguro S, et al. Engineered CRISPR-Cas9 nuclease with expanded targeting space [J]. *Science*, 2018, 361(6408): 1259-1262.
- [29] Miller SM, Wang T, Randolph PB, et al. Continuous evolution of SpCas9 variants compatible with non-G PAMs [J]. *Nature Biotechnology*, 2020, 38(4): 471-481.
- [30] Walton RT, Christie KA, Whittaker MN, et al. Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants [J]. *Science*, 2020, 368(6488): 290-296.
- [31] Slaymaker IM, Gao LY, Zetsche B, et al. Rationally engineered Cas9 nucleases with improved specificity [J]. *Science*, 2016, 351(6268): 84-88.
- [32] Nihongaki Y, Kawano F, Nakajima T, et al. Photoactivatable CRISPR-Cas9 for optogenetic genome editing [J]. *Nature Biotechnology*, 2015, 33(7): 755-760.
- [33] Zetsche B, Volz SE, Zhang F. A split-Cas9 architecture for inducible genome editing and transcription modulation [J]. *Nature Biotechnology*, 2015, 33(2): 139-142.
- [34] Gao YC, Xiong X, Wong S, et al. Complex transcriptional modulation with orthogonal and inducible dCas9 regulators [J]. *Nature Methods*, 2016, 13(12): 1043-1049.
- [35] Kim D, Kim J, Hur JK, et al. Genome-wide analysis reveals specificities of Cpf1 endonucleases in

- human cells [J]. *Nature Biotechnology*, 2016, 34(8): 863-868.
- [36] Zetsche B, Gootenberg JS, Abudayyeh OO, et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system [J]. *Cell*, 2015, 163(3): 759-771.
- [37] Gao LY, Cox DBT, Yan WX, et al. Engineered Cpf1 variants with altered PAM specificities [J]. *Nature Biotechnology*, 2017, 35(8): 789-792.
- [38] Kleinstiver BP, Sousa AA, Walton RT, et al. Engineered CRISPR-Cas12a variants with increased activities and improved targeting ranges for gene, epigenetic and base editing [J]. *Nature Biotechnology*, 2019, 37(3): 276-282.
- [39] Yang H, Gao P, Rajashankar KR, et al. PAM-dependent target DNA recognition and cleavage by C2c1 CRISPR-Cas endonuclease [J]. *Cell*, 2016, 167(7): 1814-1828.e12.
- [40] Teng F, Cui TT, Feng GH, et al. Repurposing CRISPR-Cas12b for mammalian genome engineering [J]. *Cell Discovery*, 2018, 4(1): 63.
- [41] Ming ML, Ren QR, Pan CT, et al. CRISPR-Cas12b enables efficient plant genome engineering [J]. *Nature Plants*, 2020, 6(3): 202-208.
- [42] Liu JJ, Orlova N, Oakes BL, et al. CasX enzymes comprise a distinct family of RNA-guided genome editors [J]. *Nature*, 2019, 566(7743): 218-223.
- [43] Burstein D, Harrington LB, Strutt SC, et al. New CRISPR-Cas systems from uncultivated microbes [J]. *Nature*, 2017, 542(7640): 237-241.
- [44] Yan WX, Hunnewell P, Alfonse LE, et al. Functionally diverse type V CRISPR-Cas systems [J]. *Science*, 2019, 363(6422): 88-91.
- [45] Kakarougkas A, Jeggo PA. DNA DSB repair pathway choice: an orchestrated handover mechanism [J]. *The British Journal of Radiology*, 2014, 87(1035): 20130685.
- [46] Hefferin ML, Tomkinson AE. Mechanism of DNA double-strand break repair by non-homologous end joining [J]. *DNA Repair*, 2005, 4(6): 639-648.
- [47] Bibikova M, Golic M, Golic KG, et al. Targeted chromosomal cleavage and mutagenesis in *Drosophila* using zinc-finger nucleases [J]. *Genetics*, 2002, 161(3): 1169-1175.
- [48] Capecchi MR. Altering the genome by homologous recombination [J]. *Science*, 1989, 244(4910): 1288-1292.
- [49] Chu VT, Weber T, Wefers B, et al. Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells [J]. *Nature Biotechnology*, 2015, 33(5): 543-548.
- [50] Nambiar TS, Billon P, Diedenhofen G, et al. Stimulation of CRISPR-mediated homology-directed repair by an engineered RAD18 variant [J]. *Nature Communications*, 2019, 10(1): 3395.
- [51] Baker O, Tsurkan S, Fu J, et al. The contribution of homology arms to nuclease-assisted genome engineering [J]. *Nucleic Acids Research*, 2017, 45(13): 8105-8115.
- [52] Li H, Beckman KA, Pessino V, et al. Design and specificity of long ssDNA donors for CRISPR-based knock-in [Z/OL]. *bioRxiv*, 2017, doi: <https://doi.org/10.1101/178905>.
- [53] Jiang WY, Bikard D, Cox D, et al. RNA-guided editing of bacterial genomes using CRISPR-Cas systems [J]. *Nature Biotechnology*, 2013, 31(3): 233-239.
- [54] Wang HH, Isaacs FJ, Carr PA, et al. Programming cells by multiplex genome engineering and accelerated evolution [J]. *Nature*, 2009, 460(7257): 894-898.
- [55] Ronda C, Pedersen LE, Sommer MOA, et al. CRIMAGE: CRISPR optimized MAGE recombineering [J]. *Scientific Reports*, 2016, 6(1): 19452.
- [56] Li YF, Lin ZQ, Huang C, et al. Metabolic engineering of *Escherichia coli* using CRISPR-Cas9 mediated genome editing [J]. *Metabolic Engineering*, 2015, 31: 13-21.
- [57] Garst AD, Bassalo MC, Pines G, et al. Genome-wide mapping of mutations at single-nucleotide resolution for protein, metabolic and genome engineering [J]. *Nature Biotechnology*, 2017, 35(1): 48-55.
- [58] Liu RM, Liang LY, Garst AD, et al. Directed

- combinatorial mutagenesis of *Escherichia coli* for complex phenotype engineering [J]. *Metabolic Engineering*, 2018, 47: 10-20.
- [59] Liu RM, Liang LY, Choudhury A, et al. Iterative genome editing of *Escherichia coli* for 3-hydroxypropionic acid production [J]. *Metabolic Engineering*, 2018, 47: 303-313.
- [60] Ryan OW, Skerker JM, Maurer MJ, et al. Selection of chromosomal DNA libraries using a multiplex CRISPR system [J]. *eLife*, 2014, 3: e03703.
- [61] Jakočiūnas T, Pedersen LE, Lis AV, et al. CasPER, a method for directed evolution in genomic contexts using mutagenesis and CRISPR/Cas9 [J]. *Metabolic Engineering*, 2018, 48: 288-296.
- [62] Bao ZH, Hamedirad M, Xue P, et al. Genome-scale engineering of *Saccharomyces cerevisiae* with single-nucleotide precision [J]. *Nature Biotechnology*, 2018, 36(6): 505-508.
- [63] Neggers JE, Kwanten B, Dierckx T, et al. Target identification of small molecules using large-scale CRISPR-Cas mutagenesis scanning of essential genes [J]. *Nature Communications*, 2018, 9(1): 502.
- [64] Jakočinas T, Bonde I, Herrgård M, et al. Multiplex metabolic pathway engineering using CRISPR/Cas9 in *Saccharomyces cerevisiae* [J]. *Metabolic Engineering*, 2015, 28: 213-222.
- [65] Guo XG, Chavez A, Tung A, et al. High-throughput creation and functional profiling of DNA sequence variant libraries using CRISPR-Cas9 in yeast [J]. *Nature Biotechnology*, 2018, 36(6): 540-546.
- [66] Roy KR, Smith JD, Vonesch SC, et al. Multiplexed precision genome editing with trackable genomic barcodes in yeast [J]. *Nature Biotechnology*, 2018, 36(6): 512-520.
- [67] Lian JZ, Schultz C, Cao MF, et al. Multi-functional genome-wide CRISPR system for high throughput genotype-phenotype mapping [J]. *Nature Communications*, 2019, 10(1): 5794.
- [68] Dicarlo JE, Norville JE, Mali P, et al. Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems [J]. *Nucleic Acids Research*, 2013, 41(7): 4336-4343.
- [69] Kim BH, Zhang GJ. Generating stable knockout Zebrafish lines by deleting large chromosomal fragments using multiple gRNAs [J]. *G3*, 2020, 10(3): 1029-1037.
- [70] Lee H, Kim JI, Park JS, et al. CRISPR/Cas9-mediated generation of a Plac8 knockout mouse model [J]. *Laboratory Animal Research*, 2018, 34(4): 279-287.
- [71] Shalem O, Sanjana NE, Hartenian E, et al. Genome-scale CRISPR-Cas9 knockout screening in human cells [J]. *Science*, 2014, 343(6166): 84-87.
- [72] Hahn F, Mantegazza O, Greiner A, et al. An efficient visual screen for CRISPR/CAS9 activity in *Arabidopsis thaliana* [J]. *Frontiers in Plant Science*, 2017, 8: 39.
- [73] Zhou HB, Liu B, Weeks DP, et al. Large chromosomal deletions and heritable small genetic changes induced by CRISPR/Cas9 in rice [J]. *Nucleic Acids Research*, 2014, 42(17): 10903-10914.
- [74] Butt H, Eid A, Momin AA, et al. CRISPR directed evolution of the spliceosome for resistance to splicing inhibitors [J]. *Genome Biology*, 2019, 20(1): 73.
- [75] Ipsaro JJ, Shen C, Arai E, et al. Rapid generation of drug-resistance alleles at endogenous loci using CRISPR-Cas9 indel mutagenesis [J]. *PLoS One*, 2017, 12(2): e0172177.
- [76] Findlay GM, Boyle EA, Hause RJ, et al. Saturation editing of genomic regions by multiplex homology-directed repair [J]. *Nature*, 2014, 513(7516): 120-123.
- [77] Mason DM, Weber CR, Parola C, et al. High-throughput antibody engineering in mammalian cells by CRISPR/Cas9-mediated homology-directed mutagenesis [J]. *Nucleic Acids Research*, 2018, 46(14): 7436-7449.
- [78] Parola C, Neumeier D, Friedensohn S, et al. Antibody discovery and engineering by enhanced CRISPR-Cas9 integration of variable gene cassette libraries in mammalian cells [J]. *mAbs*, 2019, 11(8): 1367-1380.
- [79] Hess GT, Tycko J, Yao D, et al. Methods and applications of CRISPR-mediated base editing

- in eukaryotic genomes [J]. *Molecular Cell*, 2017, 68(1): 26-43.
- [80] Komor AC, Kim YB, Packer MS, et al. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage [J]. *Nature*, 2016, 533(7603): 420-424.
- [81] Gaudelli NM, Komor AC, Rees HA, et al. Programmable base editing of A·T to G·C in genomic DNA without DNA cleavage [J]. *Nature*, 2017, 551(7681): 464-471.
- [82] Nishimasu H, Ran FA, Hsu PD, et al. Crystal structure of Cas9 in complex with guide RNA and target DNA [J]. *Cell*, 2014, 156(5): 935-949.
- [83] Hess GT, Frésard L, Han K, et al. Directed evolution using dCas9-targeted somatic hypermutation in mammalian cells [J]. *Nature Methods*, 2016, 13(12): 1036-1042.
- [84] Ma YQ, Zhang JY, Yin WJ, et al. Targeted AID-mediated mutagenesis (TAM) enables efficient genomic diversification in mammalian cells [J]. *Nature Methods*, 2016, 13(12): 1029-1035.
- [85] Shimatani Z, Kashojiya S, Takayama M, et al. Targeted base editing in rice and tomato using a CRISPR-Cas9 cytidine deaminase fusion [J]. *Nature Biotechnology*, 2017, 35(5): 441-443.
- [86] Zong Y, Wang YP, Li C, et al. Precise base editing in rice, wheat and maize with a Cas9-cytidine deaminase fusion [J]. *Nature Biotechnology*, 2017, 35(5): 438-440.
- [87] Tong YJ, Whitford CM, Robertsen HL, et al. Highly efficient DSB-free base editing for streptomycetes with CRISPR-BEST [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2019, 116(41): 20366-20375.
- [88] Li C, Zhang R, Meng XB, et al. Targeted, random mutagenesis of plant genes with dual cytosine and adenine base editors [J]. *Nature Biotechnology*, 2020, 38(7): 875-882.
- [89] Halperin SO, Tou CJ, Wong EB, et al. CRISPR-guided DNA polymerases enable diversification of all nucleotides in a tunable window [J]. *Nature*, 2018, 560(7717): 248-252.
- [90] Kim YB, Komor AC, Levy JM, et al. Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions [J]. *Nature Biotechnology*, 2017, 35(4): 371-376.
- [91] Li XS, Wang Y, Liu YJ, et al. Base editing with a Cpf1-cytidine deaminase fusion [J]. *Nature Biotechnology*, 2018, 36(4): 324-327.
- [92] Yan F, Kuang Y, Ren B, et al. Highly efficient A·T to G·C base editing by Cas9n-guided tRNA adenosine deaminase in rice [J]. *Molecular Plant*, 2018, 11(4): 631-634.
- [93] Kang BC, Yun JY, Kim ST, et al. Precision genome engineering through adenine base editing in plants [J]. *Nature Plants*, 2018, 4(7): 427-431.
- [94] Li C, Zong Y, Wang YP, et al. Expanded base editing in rice and wheat using a Cas9-adenosine deaminase fusion [J]. *Genome Biology*, 2018, 19(1): 59.
- [95] Wang FP, Zhang CW, Xu W, et al. Developing high-efficiency base editors by combining optimized synergistic core components with new types of nuclear localization signal peptide [J]. *The Crop Journal*, 2020, 8(3): 408-417.
- [96] Packer MS, Liu DR. Methods for the directed evolution of proteins [J]. *Nature Reviews Genetics*, 2015, 16(7): 379-394.
- [97] Liu LD, Huang M, Dai PF, et al. Intrinsic nucleotide preference of diversifying base editors guides antibody *ex vivo* affinity maturation [J]. *Cell Reports*, 2018, 25(4): 884-892.e3.
- [98] Liu XS, Qin RY, Li J, et al. A CRISPR-Cas9-mediated domain-specific base-editing screen enables functional assessment of ACCase variants in rice [J]. *Plant Biotechnology Journal*, 2020, 18(9): 1845-1847.
- [99] Kuang Y, Li S, Ren B, et al. Base-editing-mediated artificial evolution of *OsALS1 In Planta* to develop novel herbicide-tolerant rice germplasms [J]. *Molecular Plant*, 2020, 13(4): 565-572.
- [100] Tou CJ, Schaffer DV, Dueber JE. Targeted diversification in the *S. cerevisiae* genome with CRISPR-guided DNA polymerase I [J]. *ACS Synthetic Biology*, 2020, 9(7): 1911-1916.



- [101] Leemhuis H, Kelly RM, Dijkhuizen L. Directed evolution of enzymes: library screening strategies [J]. *IUBMB Life*, 2009, 61(3): 222-228.
- [102] Xiao H, Bao ZH, Zhao HM. High throughput screening and selection methods for directed enzyme evolution [J]. *Industrial and Engineering Chemistry Research*, 2015, 54(16): 4011-4020.
- [103] Coelho PS, Brustad EM, Kannan A, et al. Olefin cyclopropanation via carbene transfer catalyzed by engineered cytochrome P450 enzymes [J]. *Science*, 2013, 339(6117): 307-310.
- [104] Ostafe R, Prodanovic R, Ung WL, et al. A high-throughput cellulase screening system based on droplet microfluidics [J]. *Biomicrofluidics*, 2014, 8(4): 041102.
- [105] Kellner MJ, Koob JG, Gootenberg JS, et al. SHERLOCK: nucleic acid detection with CRISPR nucleases [J]. *Nature Protocols*, 2019, 14(10): 2986-3012.
- [106] Chen JS, Ma E, Harrington LB, et al. CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity [J]. *Science*, 2018, 360(6387): 436-439.
- [107] Liang MD, Li ZL, Wang WS, et al. A CRISPR-Cas12a-derived biosensing platform for the highly sensitive detection of diverse small molecules [J]. *Nature Communications*, 2019, 10(1): 3672.
- [108] Xiong Y, Zhang JJ, Yang ZL, et al. Functional DNA regulated CRISPR-Cas12a sensors for point-of-care diagnostics of non-nucleic-acid targets [J]. *Journal of the American Chemical Society*, 2020, 142(1): 207-213.
- [109] Li JJ, Yang SS, Zuo C, et al. Applying CRISPR-Cas12a as a signal amplifier to construct biosensors for non-DNA targets in ultralow concentrations [J]. *ACS Sensors*, 2020, 5(4): 970-977.
- [110] Ma DC, Xu ZM, Zhang ZY, et al. Engineer chimeric Cas9 to expand PAM recognition based on evolutionary information [J]. *Nature Communications*, 2019, 10(1): 560.
- [111] Liu Q, He D, Xie L. Prediction of off-target specificity and cells-specific fitness of CRISPR-Cas system using attention boosted deep learning and network-based gene feature [J]. *PLoS Computational Biology*, 2019, 15(10): e1007480.
- [112] Beneš D, Rodríguez-Patón A, Sosik P. Directed evolution of biocircuits using conjugative plasmids and CRISPR-Cas9: design and in silico experiments [J]. *Natural Computing*, 2017, 16(3): 497-505.
- [113] Huang YJ, Yang S, Chen WH, et al. A synthetic genetic circuit enables precise quantification of direct repeat deletion in bacteria [J]. *ACS Synthetic Biology*, 2020, 9(5): 1041-1050.
- [114] Shechner DM, Hacısuleyman E, Younger ST, et al. Multiplexable, locus-specific targeting of long RNAs with CRISPR-display [J]. *Nature Methods*, 2015, 12(7): 664-670.
- [115] Hamilton TA, Pellegrino GM, Therrien JA, et al. Efficient inter-species conjugative transfer of a CRISPR nuclease for targeted bacterial killing [J]. *Nature Communications*, 2019, 10(1): 4544.
- [116] Lee JK, Jeong E, Lee J, et al. Directed evolution of CRISPR-Cas9 to increase its specificity [J]. *Nature Communications*, 2018, 9(1): 3048.