

DOI: 10.3724/SP.J.1006.2024.43012

不同 CRISPR-Cas12f 系统的编辑效率比较

黄灵芝^{1,2} 符 晓² 祁显涛² 刘昌林² 谢传晓² 吴鹏昊¹ 任姣姣^{1,*}
朱金洁^{2,*}

¹ 新疆农业大学农学院, 新疆乌鲁木齐 830052; ² 中国农业科学院作物科学研究所, 北京 100081

摘 要: 来自 Type V-F 家族的 CRISPR/Cas12f 蛋白报道仅为 Cas9 蛋白分子的 1/4 到 1/3 大小, 在病毒载体的递送上具有重要优势。然而, CRISPR/Cas12f 系统介导植物基因编辑的报道较少, 编辑活性相对较低, 限制了该系统在植物上的进一步应用。本研究分别在体外酶切、酵母以及玉米原生质体瞬时表达 3 个体系中比较了 *OsCas12f*、*SpCas12f* 及 *UnCas12f* 的编辑活性。结果表明, 基于 Cas12f/sgRNA 的体外酶切, *OsCas12f* 与 *SpCas12f* 蛋白的编辑活性相当, 未检测到 *UnCas12f* 对底物酶切活性; 在酵母突变 *eGFP* 的恢复表达试验中, *OsCas12f* 在 2 个测试位点对 *eGFP* 蛋白的表达恢复效率达到 95% 以上, 效率与 *Cas12i.3* 相当; *SpCas12f* 介导的 2 个位点 *eGFP* 蛋白表达恢复效率分别是 1.63% 与 3.20%, 效果次之; *UnCas12f* 蛋白几乎无编辑活性; 玉米原生质体瞬时表达比较 *OsCas12f* 和 *SpCas12f* 介导的玉米内源位点的编辑效率, 发现 *OsCas12f* 对 2 个位点的编辑效率分别为 2.72% 和 1.97%, 而 *SpCas12f* 仅能介导其中 1 个位点的定点编辑, 编辑效率为 1.09%。Cas12f 蛋白在靶位点处引入的突变类型以碱基的缺失为主, 缺失碱基长度在 -9~ -17 bp 之间。综上, *OsCas12f* 可以作为植物微型基因编辑器及衍生技术开发的底盘工具酶。

关键词: Cas12f; sgRNA/Cas12 核糖核蛋白复合体; 原生质体; 编辑效率

Evaluation of editing efficiency of different CRISPR-Cas12f systems

HUANG Ling-Zhi^{1,2}, FU Xiao², QI Xian-Tao², LIU Chang-Lin², XIE Chuan-Xiao², WU Peng-Hao¹, REN Jiao-Jiao^{1,*}, and ZHU Jin-Jie^{2,*}

¹ College of Agriculture, Xinjiang Agricultural University, Urumqi 830052, Xinjiang, China; ² Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081, China

Abstract: CRISPR/Cas12f proteins belonging to the Type V-F family are reported to be only 1/4 to 1/3 the size of Cas9 protein molecules, providing a significant advantage in viral vector delivery. However, the CRISPR/Cas12f system for gene editing in plants has been reported to have lower editing activity, limiting its broader application in plant research. In this study, we compared the editing activities of *OsCas12f*, *SpCas12f*, and *UnCas12f* in three different systems: *in vitro* digestion, yeast, and transient expression in maize protoplasts. The results showed that the editing activities of *OsCas12f* and *SpCas12f* proteins were comparable in terms of *in vitro* digestion of Cas12f/sgRNA complexes, while no substrate digestion activity was detected for *UnCas12f*. In the yeast mutant *eGFP* expression restoration assay, *OsCas12f* exhibited an editing efficiency of over 95% at the two tested loci, which was comparable to *Cas12i.3*. On the other hand, *SpCas12f* achieved editing efficiencies of 1.63% and 3.20% at the two sites, respectively, representing the next highest effect. However, *UnCas12f* showed minimal editing activity. Furthermore, by transiently expressing maize protoplasts, we compared the editing efficiencies of *OsCas12f* and *SpCas12f* at endogenous maize loci. It was found that *OsCas12f* successfully mediated targeted editing at two loci with editing efficiencies of 2.72% and 1.97%, respectively, while *SpCas12f* only mediated targeted editing at one locus with an editing efficiency of 1.09%. Deletion of bases was the predominant type of mutation introduced by Cas12f proteins at the target loci, with deletion lengths ranging from -9 to

本研究由国家重点研发计划项目(2023YFD1202901)资助。

This study was supported by the National Key Research and Development Program of China (2023YFD1202901).

* 通信作者(Corresponding authors): 任姣姣, E-mail: renjiaojiao789@sina.com; 朱金洁, E-mail: zhujinjie@caas.cn

第一作者联系方式: E-mail: 1993253522@qq.com

Received (收稿日期): 2024-03-15; Accepted (接受日期): 2024-06-20; Published online (网络出版日期): 2024-07-11.

URL: <https://link.cnki.net/urlid/11.1809.S.20240708.1346.014>

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

-17 base pairs. These comprehensive results indicate that OsCas12f can serve as a versatile tool for developing plant microgene editors and related technologies.

Keywords: Cas12f; sgRNA/Cas12 ribonucleoprotein complex; protoplast; editing efficiency

基因编辑是一项旨在对基因组进行定点修饰的新技术,通过人工构建工程化的序列特异性核酸酶,对基因组的目标区间序列进行特异性识别和定点剪切断裂DNA双链,细胞内源的非同源末端连接或同源重组修复在目标位点序列中产生碱基缺失、插入或替换等修饰^[1]。基因编辑技术的发展大致经历了归巢核酸内切酶、锌指蛋白核酸酶、类转录激活因子效应物核酸酶以及成簇规律间隔短回文重复序列/邻近Cas基因(clustered regularly interspaced short palindromic repeats/CRISPR associated protein, CRISPR/Cas)系统等4个阶段^[2]。前3个阶段均以人工构建工程化蛋白为导向,后者基于RNA的配对识别目标DNA,具有简便性、灵活性和成本效益低的诸多优点,成为当前最受欢迎的基因组编辑工具^[3]。

根据Cas蛋白组成和效应复合物的功能性质,CRISPR/Cas系统分为两大类,即Class1和Class2。其中,Class1系统的效应复合物由多个Cas蛋白亚基组成;而Class2系统的效应复合物则是单个多功能结构域蛋白^[4]。另外,根据剪切模块的序列、功能特征及辅助模块的构成,Class1又分为3种类型即Type I、III、IV,Class2分为Type II、V、VI,且每个类型又分为了诸多不同的亚型^[5]。目前,应用最广泛的基因编辑底盘工具酶均来自Class2系统,包括Type II的Cas9和Type V的Cas12a,两者的分子量大小、前间隔序列邻近基序(protospacer adjacent motif, PAM)及剪切特征等均有不同。其中,效应蛋白的分子量大小与基因编辑工具的递送效率息息相关,分子量越小的Cas蛋白越容易通过病毒载体等进行递送,直接形成无转基因的目标突变体。因此,寻找分子尺寸更小的CRISPR效应核酸酶变得尤为重要。

许多分子量更小的Cas蛋白被陆续报道,例如SaCas9(1053个氨基酸)^[6]、CjCas9(984个氨基酸)^[7]、Cas12a(750~1373个氨基酸)^[8]、Cas12b(1129个氨基酸)^[9]、Cas12i(700~800个氨基酸)^[10]、Cas12j(700~800个氨基酸)^[11]、Cas9和Cas12a的祖先IscB(约400个氨基酸)^[12]和TnpB(约550个氨基酸)^[13]等。近期,一种微型的Cas12f蛋白被报道,其大小只有422~603个氨基酸,能够形成不对称二聚体,同sgRNA介导目标双链DNA的定点剪切^[14]。当前已经报道的Cas12f系统有来自未培养古细菌(uncultured

archaeon)的UnCas12f^[15]、来自棕榈酸互营单胞菌(*Syntrophomonas palmitatica*)的SpCas12f^[16]和来自颤螺菌属(*Oscillibacter* sp.)的OsCas12f^[17]等,可介导细菌、人类、玉米和水稻的基因编辑^[18]。然而,Cas12f蛋白对体系的离子浓度及温度的依赖性较高,编辑效率还有较大的提升空间。本研究通过PCR/RNP(Polymerase Chain Reaction/Ribonucleoprotein, PCR/RNP)体外酶切、酵母体内定向编辑和玉米原生质体瞬时表达编辑等体系,系统评估OsCas12f、SpCas12f及UnCas12f的定向编辑活性,旨在筛选出适合植物的微型高效Cas12f工具酶。

1 材料与方法

1.1 材料

将暗培养14 d的玉米品种B73的黄化苗用于制备原生质体。

1.2 菌株与质粒

大肠杆菌克隆菌株为Trans1-T1(北京全式金生物技术有限公司,CD501),原核表达菌株为BL21(DE3)(北京全式金生物技术有限公司,CD601),酵母转化菌株为AH109(北京酷莱博科技有限公司,CC317)。

原核表达载体pET-30a和酵母表达载体pGADT7和pGBKT7为本实验保存,克隆载体pEASY-Blunt 3购自全式金生物技术有限公司(CB301-01);引物均由生工生物工程(上海)股份有限公司合成,OsCas12f、SpCas12f和UnCas12f经玉米密码子优化后的全长编码序列及其sgRNA的骨架序列均由南京金斯瑞生物科技有限公司合成。

1.3 试验方法

1.3.1 Cas12f原核表达载体、酵母表达载体及玉米原生质体瞬时表达载体的构建 原核表达载体构建:以人工合成的3个目的片段OsCas12f、SpCas12f和UnCas12f作为模板,利用带同源臂序列的引物进行PCR扩增(表1),采用pET-30a载体作为目的载体,通过限制性内切酶Hind III和BamH I在37℃酶切3 h后,通过NEBuilder DNA组装试剂盒(NEB北京分公司, E2621L)将pET-30a线性载体和不同Cas12f基因片段相连,最终成功构建pET-30a-OsCas12f、pET-30a-SpCas12f和pET-30a-UnCas12f载体。

酵母表达载体构建: 根据目的载体序列来设计引物(表 1), 扩增获得 3 种 *Cas12f*、*Cas12i.3* 片段。将 pGADT7 载体作为目的载体, 通过限制性内切酶 *Kpn* I 和 *Bam* H I 进行双酶切获得线性载体。通过无缝克隆将扩增的目的片段与线性 pGADT7 载体相连, 完成中间载体 pGADT7-*Cas12i.3*、pGADT7-*OsCas12f*、pGADT7-*SpCas12f*、pGADT7-*UnCas12f* 的构建; 将中间载体利用 *Nhe* I 和 *Not* I 限制性内切酶切割, 与 *SNR52* 启动子序列和带有 T1 基因靶位点与 T2 基因靶位点的 sgRNA 序列(表 1)相连, 最终完成 7 个表达载体 pGADT7-*Cas12i.3*-sgRNA、pGADT7-*OsCas12f*-T1、pGADT7-*OsCas12f*-T2、pGADT7-*SpCas12f*-T1、pGADT7-*SpCas12f*-T2、pGADT7-*UnCas12f*-T1、pGADT7-*UnCas12f*-T2 的构建。

eGFP 报告载体构建: 将 pGBKT7 载体用 *Bsa* B I 与 *Bss* H II 进行双酶切, 与带有 T1、T2 基因靶位点和 *eGFP* 序列目的片段(表 1)相连, 完成 pGBKT7-*eGFP* 报告载体的构建。

玉米原生质体表达载体构建: 同上述方法一致, 通过 2 次酶切和 2 次连接将扩增后的片段先后克隆至瞬时表达载体 pUC19 中, 分别完成了 pUC19-*Ubi-OsCas12f-Rbcs*、pUC19-*Ubi-SpCas12f-Rbcs*、pUC19-*Ubi-OsCas12f-Rbcs*-sgRNAT1、pUC19-*Ubi-OsCas12f-Rbcs*-sgRNAT2 和 pUC19-*Ubi-SpCas12f-Rbcs*-sgRNAT1 和 pUC19-*Ubi-SpCas12f-Rbcs*-sgRNAT2 六个载体。在这 6 个表达载体中, 由玉米 *Ubiquitin1* 启动子驱动 *Cas12f* 表达, 用 *CaMV35S enhancer-CmYLCV-ZmU6* 复合型启动子驱动 sgRNA 表达^[19]。

1.3.2 不同 Cas12f 蛋白的原核表达及纯化 将不同 Cas12f 质粒转化到 BL21 (DE3) 感受态细胞中后挑取单克隆, 最终在 700 mL 含有卡那抗生素(浓度均为 50 $\mu\text{g mL}^{-1}$)的 LB 液体培养基 37℃ 进行扩大培养到 OD₆₀₀ 值在 0.70~0.75 区间, 冰浴 1 h, 按 1:100 加入 0.5 mol L⁻¹ IPTG 试剂进行诱导表达, 恒温摇床培养温度 18℃, 转速为 140 转 min⁻¹, 培养 14~16 h。4000 转 min⁻¹、4℃, 离心 20 min, 收集菌体。加入 45 mL 结合缓冲液重悬菌体。超声波破碎菌体, 12,000 转 min⁻¹、4℃, 离心 1 h, 取上清并过滤除菌。

细菌裂解液离心后按 1:1000 比例加入清洗好过的 Ni-NTA 琼脂糖介质, 4℃ 旋转孵育 60~90 min。将样品通过收集柱经重力自然流出后, 用 20 倍柱体积的清洗缓冲液清洗, 之后用洗脱缓冲液进行洗脱, 通过凝胶电泳(10% SDS-PAGE)对收集的各个组分

进行鉴定。

利用 100000-MWCO 蛋白超滤管将样品中的洗脱缓冲液替换为蛋白储存液(50 mmol L⁻¹ Tris, 600 mmol L⁻¹ NaCl, 20% glycerol, 1 mmol L⁻¹ TCEP (pH 7.5))直到样品浓缩体积为 2 mL, 并用 Pierce M BCA Protein Assay Kit (Thermo Fisher, 23227)对不同 Cas12f 蛋白进行定量。

1.3.3 RNA 体外转录与纯化 分别在玉米 *ZmGDIa* (Zm00001d010255) 和 *ZmGRF5* (Zm00001d026240) 基因上设计了靶向编辑位点 T1: 5'-TTCTGGGCACGG GGCTCAAGGAGTGCA-3' 和 T2: 5'-TTTCTTGAAC GGTTGCGGCCGCGGTGCA-3', 设计的靶位点适用于 *Cas12f* 在酵母及玉米原生质体中的活性检测。以合成的 *Cas12f* 的 sgRNA 骨架为模板, T7 sgRNA scaffold F、T1 sgRNA scaffold R、T2 sgRNA scaffold R 为引物(表 1), 通过重叠 PCR 将 T7 启动子及靶点序列分别添加到 sgRNA 序列的 5' 端和 3' 端, 扩增获得不同 Cas12f 蛋白的 sgRNA 体外转录模板, 2% 琼脂糖凝胶电泳后胶回收, 获得体外转录底物。体外转录使用 HiScribe T7 RNA 快速转录试剂盒(NEB, E2025S)。最后通过苯酚: 氯仿萃取纯化和乙醇沉淀法获得 sgRNA^[20], -80℃ 保存备用。

1.3.4 基于不同 Cas12f 蛋白的 PCR/RNP 酶切 以玉米 B73 基因组 DNA 为模板, 通过 PCR 扩增 sgRNA 靶序列两侧各 200~300 bp 序列作为 PCR/RNP 酶切体系的底物。将 2 μg 不同 Cas12f 蛋白和 2 μg sgRNA、1 μL 10×RNP 组装缓冲液混合均匀, 37℃ 孵育 30 min, 加入 0.3 μg 酶切底物, 分别在 28℃、37℃、45℃ 和 55℃ 酶切 0.5~3.0 h。加入 2 μL RNaseA, 37℃ 孵育 15 min 去除反应体系中的 sgRNA, 60℃ 孵育 20 min 终止反应, 1% 琼脂糖凝胶电泳检测酶切情况。采用 ImageJ 软件对酶切条带进行条带灰度值分析, 酶切效率计算公式为: 酶切效率(%) = $(b+c)/(a+b+c) \times 100$, 其中, a 为未消化 PCR 底物的灰度值, b 和 c 为酶切产生条带的灰度值。

1.3.5 酵母细胞转化 取 10 μL Carrier DNA, 95~100℃, 5 min, 冰浴后快速冷却, 重复 1 次。取 100 μL AH109 感受态冰上融化后依次快速加入构建完成的 Cas12f 酵母质粒 2~5 μg 、PEG/LiAc 500 μL 充分吹打混匀, 30℃ 水浴 30 min, 期间水浴至 15 min 时轻轻混匀 1 次。加入 20 μL DMSO 混匀, 45℃ 水浴 15 min (中间混匀 1 次)。5000 转 min⁻¹ 离心 40 s, 弃上清, 加入 1 mL YPD Plus, 30℃ 震荡培养 90 min,

5000 转 min^{-1} 离心 40 s, 弃上清, 加入 300 μL 0.9% NaCl 重悬, 均匀涂布在缺 Trp (Tryptophan) 和 Leu (Leucine) 的酵母培养基上, 30 $^{\circ}\text{C}$ 倒置培养 96 h。

试验设计组合方式: 阳性对照载体 pGADT7-*Cas12i.3*-sgRNA+pGBKT7-*eGFP*; *OsCas12f* 编辑活性检测对照组 pGADT7-*OsCas12f*+pGBKT7-*eGFP*, 实验组 pGADT7-*OsCas12f*-T1+pGBKT7-*eGFP*、pGADT7-*OsCas12f*-T2+pGBKT7-*eGFP*; *SpCas12f* 编辑活性检测对照组 pGADT7-*SpCas12f*+pGBKT7-*eGFP*, 实验组 pGADT7-*SpCas12f*-T1+pGBKT7-*eGFP*、pGADT7-*SpCas12f*-T2+pGBKT7-*eGFP*; *UnCas12f* 编辑活性检测对照组 pGADT7-*UnCas12f*+pGBKT7-*eGFP*, 实验组 pGADT7-*UnCas12f*-T1+pGBKT7-*eGFP*、

pGADT7-*UnCas12f*-T2+pGBKT7-*eGFP*。

酵母绿色荧光蛋白恢复表达观察: 利用 LUYOR-3415RG 便携式荧光蛋白激发光源对酵母克隆进行 *eGFP* 表达的可视化检测。酵母中 *eGFP* 恢复效率的计算公式: *eGFP* 恢复效率(%) = (绿色荧光酵母克隆数/统计酵母克隆数) $\times$ 100%。

1.3.6 玉米原生质转化及突变位点检测 玉米原生质体制备及 PEG 介导的玉米原生质体转化方法参考 Cao 的文章^[21]。分别取不同 *Cas12f* 表达质粒 20 μg , 通过 PEG 转化至玉米原生质体中, 暗培养 48 h 后收集原生质体, CTAB 方法提取原生质体基因组 DNA, 利用靶位点特异扩增引物(表 1)进行 PCR 扩增后进行 Hi-TOM 测序, 检测靶位点编辑情况。

表 1 本研究所用的引物序列

Table 1 Primer sequences used in this paper

引物名称 Primer name	引物序列 Primer sequence (5'-3')
原核表达载体构建引物 Prokaryotic expression vector construction primers	
<i>OsCas12f</i> F	CAAGGCCATGGCTGATATCGGATCCATGGACTATAAGGATCATGACG
<i>OsCas12f</i> R	TGGTGGTGCTCGAGTGC GGCCGCAAGCTTTACTTCTTCTTTTAGCTTGTC
<i>SpCas12f</i> F	CGACAAGGCCATGGCTGATATCGGATCCATGGATTATAAAGATCATGACGGTG
<i>SpCas12f</i> R	GGTGGTGGTGCTCGAGTGC GGCCGCAAGCTTCTATTTTTCTTTTTCCTTCG
<i>UnCas12f</i> F	CGACAAGGCCATGGCTGATATCGGATCCATGGATTATAAAGATCATGACGGTG
<i>UnCas12f</i> R	GGTGGTGGTGCTCGAGTGC GGCCGCAAGCTTCTATTTCTTCTTTTAGCTTGC
酵母表达载体构建引物 Yeast expression vector construction primers	
pGADT7 <i>Cas12f</i> F1	AAGAAGAGAAAGGTGCAATTGGGTACCAAGCGTCCCGCTGCAACGAAA
pGADT7 <i>Cas12f</i> R1	GGGTTCCGGATCGCGGCCCGCCGGTAGAGGTGTGGTCAATAAGAG
pGADT7 <i>Cas12f</i> F2	GAGGAGTTTACGTCCAGCCAAGCTAGCTCTTTGAAAAGATAATGTATGA
pGADT7 <i>Cas12f</i> R2	GGGTTCCGGATCGCGGCCCGCCGGTAGAGGTGTGGTCAATAAGAG
pGADT7 sgRNA T1 F	GAGAGAAAAGGTGCAATTGGGTACCAAGCGTCCCGCTGCAACGAAA
pGADT7 sgRNA T1 R	ATTAAGGGTTCCGGATCGCGGCCGCAAAAAAACTCCTTGAGCCCCGTGCCCA
pGADT7 sgRNA T2 F	AAGAGAAAAGGTGCAATTGGGTACCAAGCGTCCCGCTGCAACGAAA
pGADT7 sgRNA T2 R	GGGTTCCGGATCGCGGCCGCAAAAAAATGCACCGCGCCGCAACCGTTCAAGT
pGBKT7- <i>eGFP</i> F	ACTGTAGCCCTAGACTTGATAGCCATCATCATATCGAAAGTTTC
pGBKT7- <i>eGFP</i> R	TTAGCTTGGCTGCAAGCGCGCCTAGTACAGCTCGTCCATGCCG
植物表达载体构建引物 Plant expression vectors construct primers	
pUC19 <i>Cas12f</i> F1	CAGCTATGACCATGATTACGCCAAGCTTATCGAGCAGCTGGCTTGTGGGGAC
pUC19 <i>Cas12f</i> R1	TCACGACGTTGTAAAAACGACGGCCAGTGAATTCCTTATCTTTAATCATATTCC
pUC19 <i>Cas12f</i> sgRNA F2	ATTGATTGACAACGGATCCCCGGGTACCATGGAGTCAAAGATTCAAATAGAG
pUC19 <i>Cas12f</i> sgRNA R2	GTCACGACGTTGTAAAAACGACGGCCAGTGAATTCCTTATCTTTAATCATATTCC
向导 RNA 分子体外转录引物 Primers for transcription of RNA molecules <i>in vitro</i>	
T7 sgRNA scaffold F	ATTCTAATACGACTCACTATAGGGAGGGCCGACTTCCCGGCCCAAAATCGAGACAGTAGCC
T1 sgRNA scaffold R	CTCCTTGAGCCCCGTGCCACCTTCAAGCCGCTTTCGCGGCTCATGCACGG
T2 sgRNA scaffold R	TGCACCGCGCCGCAACCGTTCAACCTTCAAGCCGCTTTCGCGGCTCATGCACGG

(续表 1)

引物名称	引物序列
Primer name	Primer sequence (5'-3')
PCR/RNP 体外酶切底物扩增引物 Poole/Enpp in vitro digestion substrate amplification primers	
T1F	CTTGACTACTACCTCCCACC
T1R	TGGTAACTCCCCGGACCAA
T2F	CCATTCCGGTATCGTTGCT
T2R	GACCCATCTGACACCGATCA
玉米原生质体靶位点扩增引物 Primers for amplification of maize protoplasts target sites	
T1F1	GGAGTGAGTACGGTGTGCTTCAACTCGACTCCAGCA
T1R1	GAGTTGGATGCTGGATGGGTCGGTGCTCACCTTGAGGC
T2F1	GGAGTGAGTACGGTGTGCCCCGGACTCCAAGTACTGC
T2R1	GAGTTGGATGCTGGATGGGTACAGCGAGTGGTTCTGGA

表格中下画线为构建载体同源臂序列。
The underlined sequences represent homology arms used for constructing the vector.

2 结果与分析

2.1 OsCas12f、SpCas12f、UnCas12f 蛋白的原核表达纯化

基于原核表达骨架载体 pET-30a, 构建了 OsCas12f、

SpCas12f和 UnCas12f蛋白的原核表达载体(图 1-A)。分别在蛋白两端融合了 6×His 标签, 融合标签后 OsCas12f、SpCas12f 及 UnCas12f 蛋白分子量大小预测分别为 62.93 kD、70.41 kD 和 75.12 kD。研究分别在 28℃和 37℃下进行 IPTG 诱导蛋白表达, 通过

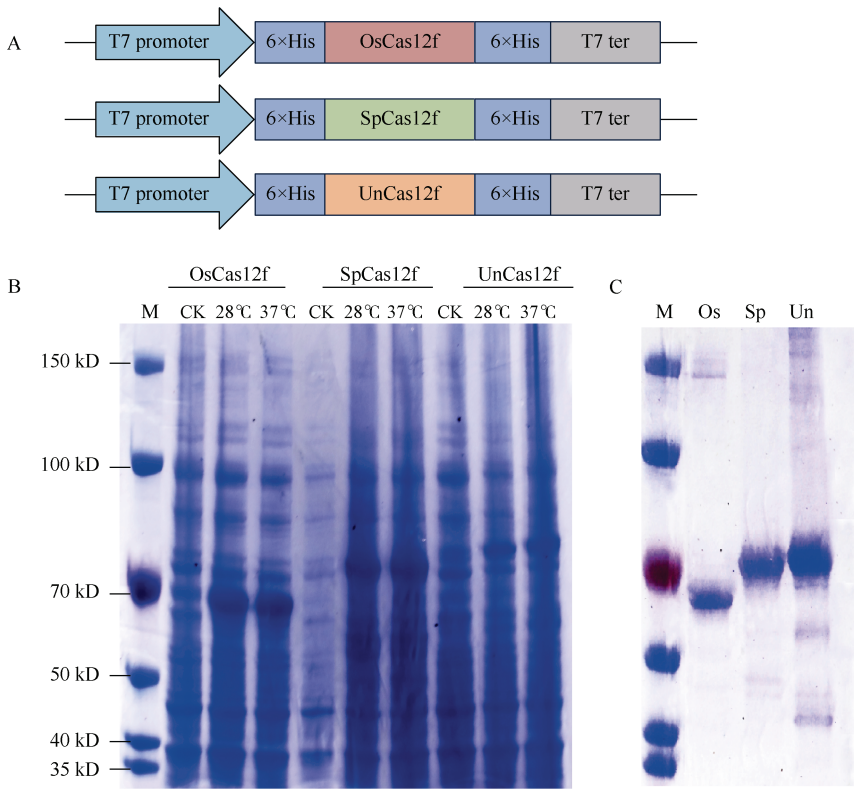


图 1 不同 CRISPR-Cas12f 蛋白的原核表达及蛋白纯化
Fig. 1 Prokaryotic expression and protein purification of different CRISPR-Cas12f proteins
A: CRISPR-Cas12f 原核表达载体结构示意图; B: 3 种 CRISPR-Cas12f 蛋白在 28℃和 37℃下的诱导表达; C: 3 种 CRISPR-Cas12f 蛋白 Ni-NTA 纯化后 12% SDS-PAGE 凝胶电泳图。
A: schematic diagram of the prokaryotic expression vector for CRISPR-Cas12f; B: induced expression of three types of CRISPR-Cas12f proteins at 28℃ and 37℃; C: 12% SDS-PAGE gel electrophoresis image of three CRISPR-Cas12f proteins purified by Ni-NTA chromatography.

SDS-PAGE 蛋白凝胶电泳检测表明, Cas12f 蛋白均有表达, 且表达蛋白大小符合预期(图 1-B)。研究进一步扩大诱导表达体系, 为保持蛋白活性, 蛋白诱导表达温度调整为 18℃, 利用 Ni-NTA 亲和层析柱进行蛋白纯化, 获得较高纯度的 Cas12f 蛋白(图 1-C), 用于后续体外 PCR-RNP 活性检测。

2.2 Cas12f/sgRNA 介导的体外底物酶切

将 OsCas12f、SpCas12f 和 UnCas12f 纯化蛋白分别与 sgRNA-T1、sgRNA-T2 组装为 RNP 复合体, 对 2 种 PCR 产物进行酶切, 结束后进行琼脂糖凝胶电泳检测, 以验证不同 Cas12f 蛋白的底物酶切活性。研究比较了反应温度为 28℃、37℃、45℃和 55℃时不同

Cas12f 蛋白对底物的酶切效率。结果显示, 28℃时 OsCas12f 和 SpCas12 对靶点 T1 酶切效率分别为 45%和 44%; 37℃时 OsCas12f 和 SpCas12 对靶点 T1 酶切效率分别为 73%和 87%, 对靶点 T2 酶切效率分别为 20%和 10%; 45℃时 OsCas12f 和 SpCas12 对靶点 T1 酶切效率分别为 81%和 85%, 对靶点 T2 酶切效率分别为 86%和 46%; 55℃时 OsCas12f 和 SpCas12 对靶点 T1 酶切效率分别为 49%和 96%, 对靶点 T2 酶切效率分别为 45%和 78%。其中, OsCas12f 的最佳温度为 45℃, SpCas12f 的最佳温度为 55℃。28℃时 OsCas12f 和 SpCas12f 对靶点 T2 序列无剪切活性。所有测试温度下 UnCas12f 均不能对底物进行剪切(图 2)。

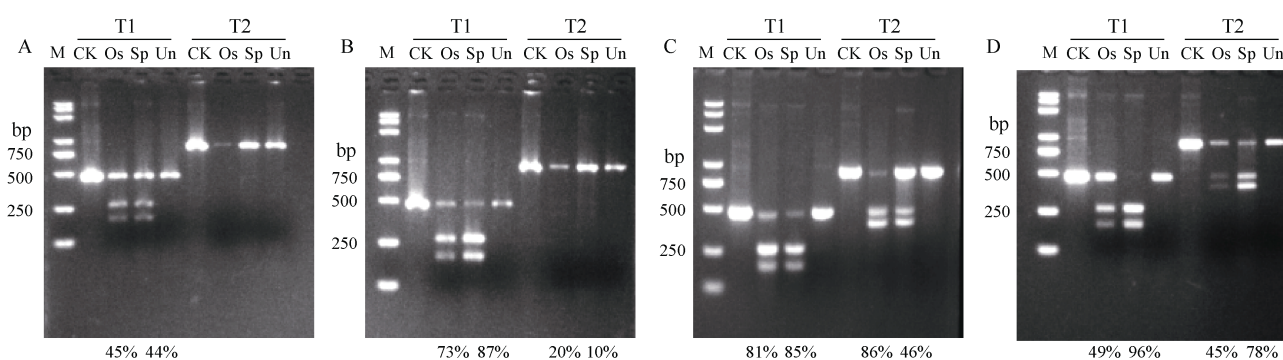


图 2 3 种 CRISPR-Cas12f 蛋白在不同温度下的底物酶切效率

Fig. 2 Substrate cleavage efficiency of three CRISPR-Cas12f proteins at different temperatures

A: 28℃时 Cas12f 的 PCR/RNP 酶切效率; B: 37℃时 Cas12f 的 PCR/RNP 酶切效率; C: 45℃时 Cas12f 的 PCR/RNP 酶切效率; D: 55℃时 Cas12f 的 PCR/RNP 酶切效率。

A: PCR/RNP digestion efficiency of Cas12f at 28℃; B: PCR/RNP digestion efficiency of Cas12f at 37℃; C: PCR/RNP digestion efficiency of Cas12f at 45℃; D: PCR/RNP digestion efficiency of Cas12f at 55℃.

研究进一步测试了反应温度为 45℃, 不同反应时间 Cas12f 蛋白对底物的剪切效率。结果表明, 0.5 h 时 OsCas12f 和 SpCas12 对靶点 T1 剪切效率分别为 49%和 47%, 对靶点 T2 剪切效率分别为 79%和 18%; 1 h 时 OsCas12f 和 SpCas12 对靶点 T1 剪切效率分别为 70%和 61%, 对靶点 T2 剪切效率分别为 81%和 82%; 2 h 时 OsCas12f 和 SpCas12 对靶点 T1 剪切效率分别为 72%和 78%, 对靶点 T2 剪切效率分别为 81%和 34%; 3 h 时 OsCas12f 和 SpCas12 对靶点 T1 剪切效率分别为 64%和 84%, 对靶点 T2 剪切效率分别为 85%和 49%。其中, 酶切 1 h, OsCas12f 对靶点 T1 和 T2 的酶切效率最大, 而 3 h 时 SpCas12f 对靶点 T1 和 T2 的酶切效率最大。UnCas12f 在延长酶切时间的条件下仍没有酶切活性(图 3)。

2.3 CRISPR/Cas12 蛋白在酵母体系中活性检测

为比较不同 CRISPR-Cas12f 系统的真核生物基因编辑活性, 构建了基于荧光蛋白表达恢复的 Cas12f 蛋白编辑活性报告系统。然后分别构建仅 Cas12f 蛋

白表达和 Cas12f 与向导 sgRNA 分子同时表达的编辑载体(图 4-A)。将报告载体与编辑载体共转化酵母后, Cas12f 蛋白在靶位点剪切产生 DNA 双链断裂, 通过同源重组修复可恢复 eGFP 蛋白的表达(图 4-B)。

将 OsCas12f、SpCas12f 及 UnCas12f 携带或不携带靶向 T1 或 T2 sgRNA 表达载体同报告载体共转化酵母涂 SD/-Leu/-Trp 双缺板, 28℃培养 48 h 后统计 eGFP 蛋白表达恢复率。结果显示, Cas12i.3 系统能够高效恢复 eGFP 的表达, 荧光蛋白的恢复效率达 97.28%。OsCas12f 同 sgRNA-T1、sgRNA-T2 对 eGFP 的表达恢复效率分别是 97.60%与 95.67%; SpCas12f 同 sgRNA-T1、sgRNA-T2 对 eGFP 的表达恢复效率分别是 1.63%与 3.20%; UnCas12f 组合对 eGFP 的表达恢复效率分别 0.24%与 0.16%, 基本为背景表达水平(图 4-C 和图 5)。研究结果表明, OsCas12f 在酵母细胞内的编辑活性跟 Cas12i.3 相当, SpCas12f 在酵母细胞内的编辑活性较低, 而 UnCas12f 几乎没有编辑活性。

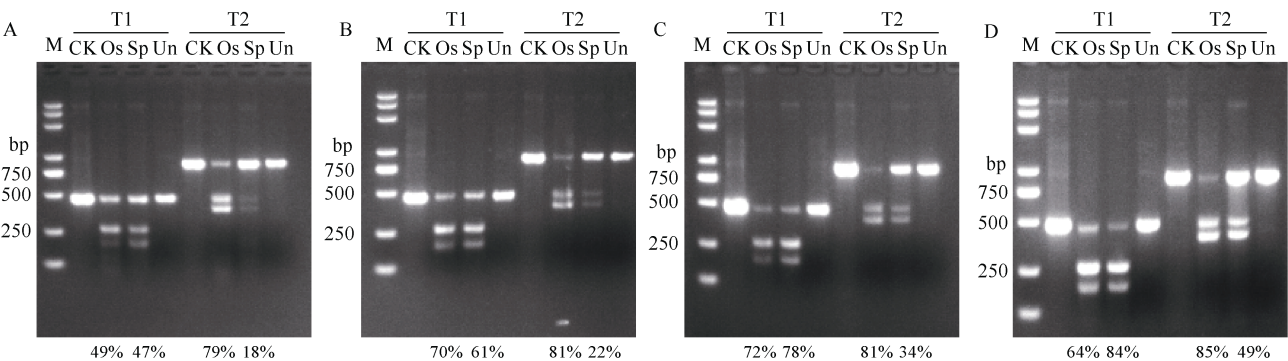


图 3 3 种 CRISPR-Cas12f 蛋白在不同时间下底物酶切效率

Fig. 3 Substrate cleavage efficiency of three CRISPR-Cas12f proteins at different time points

A: Cas12f 在 0.5 h 时 PCR/RNP 酶切效率; B: Cas12f 在 1 h 时 PCR/RNP 酶切效率; C: Cas12f 在 2 h 时 PCR/RNP 酶切效率; D: Cas12f 在 3 h 时 PCR/RNP 酶切效率。

A: PCR/RNP digestion efficiency of Cas12f at 0.5 h; B: PCR/RNP digestion efficiency of Cas12f at 1 h; C: PCR/RNP digestion efficiency of Cas12f at 2 h; D: PCR/RNP digestion efficiency of Cas12f at 3 h.

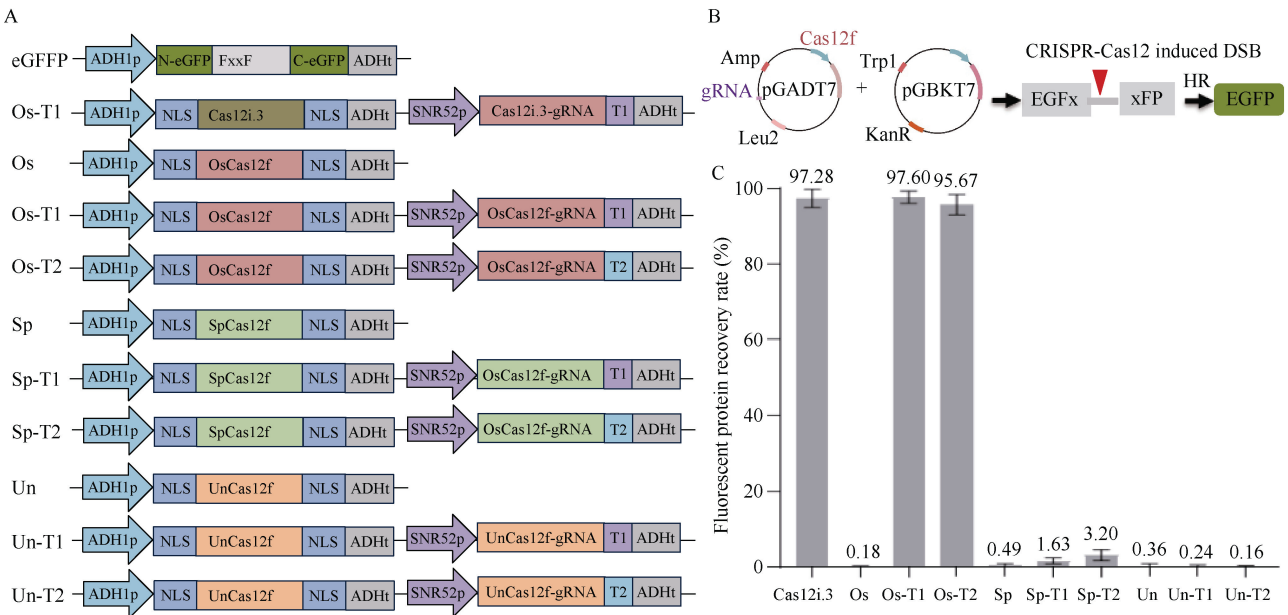
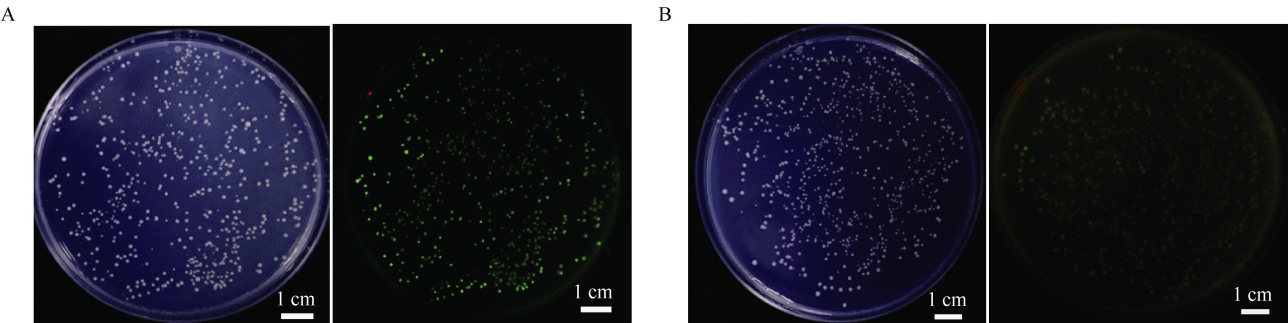


图 4 3 种 CRISPR-Cas12f 系统的酵母编辑效率

Fig. 4 Editing efficiency of the three CRISPR-Cas12f systems in yeast

A: 酵母转化载体示意图; B: 通过荧光蛋白恢复率评估 CRISPR-Cas12f 系统编辑效率原理图; C: 不同 CRISPR-Cas12f 介导的酵母 eGFP 恢复效率统计。

A: schematic diagram of yeast transformation vector; B: schematic diagram of the editing efficiency of CRISPR-Cas12f system assessed by fluorescent protein recovery rate; C: statistics of yeast editing efficiency with different CRISPR-Cas12f.



(图 5)

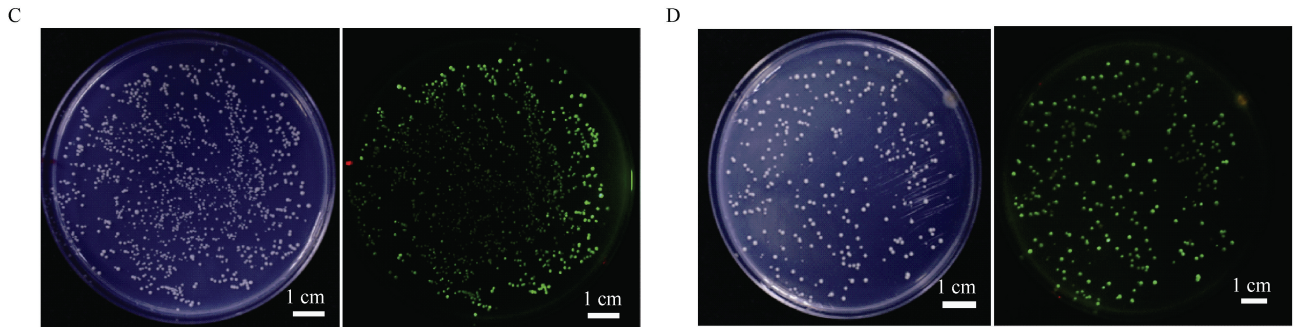


图 5 *OsCas12f* 系统介导的 *eGFP* 表达恢复

Fig. 5 *eGFP* expression recovery mediated by *OsCas12f* system

A: *Cas12i.3* 介导的 *eGFP* 表达恢复, 为阳性对照; B: 无 sgRNA 不能介导 *OsCas12f* 系统恢复 *eGFP*; C: sgRNA-T1 介导 *OsCas12f* 系统恢复 *eGFP*; D: sgRNA-T2 介导 *OsCas12f* 系统恢复 *eGFP*。

A: *Cas12i.3*-mediated restoration of *eGFP* expression, a positive control; B: no sgRNA fails to mediate restoration of *eGFP* in the *OsCas12f* system; C: sgRNA-T1 mediates restoration of *eGFP* using *OsCas12f* system; D: sgRNA-T2 mediates restoration of *eGFP* using *OsCas12f* system.

2.4 CRISPR/Cas12 蛋白在玉米原生质体中的活性检测

前人研究证明 *SpCas12f* 在玉米和水稻中具有编辑活性^[14], 本研究发现 *OsCas12f* 在体外及酵母胞内具有优于 *SpCas12f* 的靶标编辑活性。因此, 研究进一步评估 *OsCas12f* 同 *SpCas12f* 在玉米原生质体瞬时表达的编辑活性。研究分别构建了靶向玉米内源 T1、T2 位点的 *OsCas12f* 或 *SpCas12f* 植物表达载体, 利用转录增强子串联玉米的 *Ubiquitin1* 启动子驱动 *OsCas12f* 或 *SpCas12f* 的表达和 *CaMV35s enhancer-CmYLCV-ZmU6* 复合型启动子驱动向导 RNA 的表达。瞬时转化玉米原生质体后置于 37℃

培养 48 h, 提取基因组 DNA, 扩增靶位点序列进行二代测序。

Hi-TOM 测序结果分析表明, *OsCas12f* 在靶向位点 T1 和 T2 均可以进行定点编辑引入缺失突变, 在内源位点 T1 的编辑效率为 2.72%, 在靶位点远离 PAM 端发生 9 bp 到 14 bp 不等的碱基缺失, 在内源位点 T2 的编辑效率为 1.97%, 在远离 PAM 端的靶位点处发生 11 bp 的碱基缺失。*SpCas12f* 在 T1 靶位点没有检测到突变, 在 T2 靶位点处检测到 17 bp 的插入缺失突变, 编辑效率仅为 1.09% (表 2)。结果表明, *OsCas12f* 蛋白的玉米原生质体瞬时表达编辑活性高于 *SpCas12f*。

表 2 *OsCas12f* 及 *SpCas12f* 瞬时表达介导的玉米内源位点编辑

Table 2 Editing activity of maize endogenous loci mediated by transient expression of *OsCas12f* and *SpCas12f*

系统 System	靶位点突变 Target mutation (5'–3')	突变类型 Mutation type	突变数量 Mutation number	比例 Percentage (%)
OsCas12f-T1	ATCGT <u>T</u> CTGGGCACGGGGCTCAAGGAGTGCATCCT	WT	6734	91.91
	ATCGT <u>T</u> CTGGGCACGGGGc-----ATCCT	–11 bp	199	2.72
	ATCGT <u>T</u> CTGGGCACGGGGct-----ATCCT	–10 bp	176	2.40
	ATCGT <u>T</u> CTGGGCACGGGg-----CCT	–14 bp	136	1.86
	ATCGT <u>T</u> CTGGGCACGGGGct-----CATCCT	–9 bp	82	1.12
OsCas12f-T2	GCCACATGCACCGCGCCGCAACCGTTCAAGAAAGCCTGTGGAA GCGCA	WT	7655	98.03
	GCCACATGCa-----ACCGTTCAAGAAAGCCTGTGGAAAGCGCA	–11 bp	154	1.97
SpCas12f-T2	GCCACATGCACCGCGCCGCAACCGTTCAAGAAAGCCTGTGGAA GCGCA	WT	7506	98.91
	GCCACATGCACCGCGCCGCA---GTGGAA-a-----CGCA	–17 bp	83	1.09

下画线处碱基为靶位点序列, 双下画线处碱基为 *Cas12f* 蛋白的 PAM 序列, “-”为缺失碱基。

The underlined base is the target site sequence, the double-underlined base is the PAM sequence of *Cas12f* protein, and “-” is the missing base.

3 讨论

由于 Cas12f 蛋白分子量小, V-F 型的 CRISPR-Cas12f 系统在编辑工具的递送环节具有较大的优势, 是基因编辑应用领域的重要潜力工具。本研究检测了 *OsCas12f*、*SpCas12f* 和 *UnCas12f* 分别在体外酶切、酵母以及玉米原生质体瞬时表达等对靶位点的编辑活性。结果表明, *OsCas12f* 和 *SpCas12f* 在体外对底物 PCR-RNP 的酶切能力基本相当, 但 *OsCas12f* 的最佳适宜温度要求相对 *SpCas12f* 低, *OsCas12f* 的最佳适宜温度为 45℃, *SpCas12f* 的最佳适宜温度为 55℃; 酶切反应速度上, *OsCas12f* 对底物实现最大酶切大约 1 h, 而 *SpCas12f* 基本需要 3 h; 总体上, *OsCas12f* 优于 *SpCas12f*。由此, 更好的解释了酵母体内靶位点的编辑活性表现。*OsCas12f* 介导的荧光蛋白表达恢复效率远高于 *SpCas12f* 和 *UnCas12f*, 其对靶位点的编辑效率与 *Cas12i.3* 相当。在本研究中的体外酶切和酵母体系中均未检测 *UnCas12f* 的底物编辑活性, Karvelis 等^[22]研究比较了不同 Cas12f 核酸酶介导的大肠杆菌中质粒干扰活性, 比较发现 *Un1Cas12f1*、*Un2Cas12f2* 和 *AuCas12f2* 等均不具备质粒干扰活性, *SpCas12f1* 具有较强的质粒干扰活性。

前人的研究表明 *SpCas12f* 的最适温度较高, 需要热激活或者较高的培养温度才能获得 *SpCas12f* 介导的玉米或水稻稳定表达株系靶位点的定向编辑^[21,23-24]。本研究中瞬时表达 Cas12f 蛋白介导的玉米内源靶位点编辑活性比较证明, *OsCas12f* 介导了两个靶位点的定向编辑, 造成远离 PAM 端的不同类型的多碱基缺失, 突变类型符合该家族蛋白的编辑特征; 而 *SpCas12f* 蛋白仅在其中的 1 个靶位点检测到 1 种突变类型。本研究系统的比较了 3 种 Cas12f 的靶向编辑活性, 证明了 *OsCas12f* 较 *SpCas12f* 和 *UnCas12f* 具有更高效的编辑活性, 未来其在稳定转化株系中的编辑活性可能不需要热激活等额外操作。此外, *OsCas12f* 酶的小尺寸和自二聚化^[25]将为基于病毒的植物转基因递送提供可能。

4 结论

本研究通过体外酶切、酵母以及玉米原生质体 3 种方式综合评估了 *OsCas12f*、*SpCas12f* 和 *UnCas12f* 的编辑活性, 发现 *OsCas12f* 和 *SpCas12f* 在体外具有同等的底物酶切能力, 尤其是 *OsCas12f*, 其最适反应温度较 *SpCas12f* 低, 且酶切反应速度较快。在酵母中, *OsCas12f* 对底物的编辑能力与已报道

的 *Cas12i.3* 蛋白相当。最后, 在玉米原生质体的瞬时表达编辑活性上, *OsCas12f* 的编辑效率高于 *SpCas12f*。因此, 微型 *OsCas12f* 系统在植物基因编辑中具有较大的应用潜力。

References

- [1] Chen K, Wang Y, Zhang R, Zhang H, Gao C. CRISPR/Cas genome editing and precision plant breeding in agriculture. *Annu Rev Plant Biol*, 2019, 70: 667–697.
- [2] Geurts A M, Cost G J, Freyvert Y, Zeitler B, Miller J C, Choi V M, Jenkins S S, Wood A, Cui X, Meng X, Vincent A, Lam S, Michalkiewicz M, Schilling R, Foeckler J, Kalloway S, Weiler H, Ménoret S, Anegón I, Davis G D, Zhang L, Rebar E J, Gregory P D, Urnov F D, Jacob H J, Buelow R. Knockout rats via embryo microinjection of zinc-finger nucleases. *Science*, 2009, 325: 433.
- [3] An Y, Geng Y, Yao J, Fu C, Lu M, Wang C, Du J. Efficient genome editing in populus using CRISPR/Cas12a. *Front Plant Sci*, 2020, 11: 593938.
- [4] Makarova K S, Wolf Y I, Alkhnbashi O S, Costa F, Shah S A, Saunders S J, Barrangou R, Brouns S J, Charpentier E, Haft D H, Horvath P, Moineau S, Mojica F J, Terns R M, Terns M P, White M F, Yakunin A F, Garrett R A, van der Oost J, Backofen R, Koonin E V. An updated evolutionary classification of CRISPR-Cas systems. *Nat Rev Microbiol*, 2015, 13: 722–736.
- [5] Makarova K S, Wolf Y I, Iranzo J, Shmakov S A, Alkhnbashi O S, Brouns S J, Charpentier E, Cheng D, Haft D H, Horvath P, Moineau S, Mojica F J M, Scott D, Shah S A, Siksnys V, Terns M P, Venclovas Č, White M F, Yakunin A F, Yan W, Zhang F, Garrett R A, Backofen R, van der Oost J, Barrangou R, Koonin E V. Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants. *Nat Rev Microbiol*, 2020, 18: 67–83.
- [6] Nishimasu H, Cong L, Yan W X, Ran F A, Zetsche B, Li Y, Kurabayashi A, Ishitani R, Zhang F, Nureki O. Crystal Structure of *Staphylococcus aureus* Cas9. *Cell*, 2015, 162: 1113–1126.
- [7] Kim E, Koo T, Park S W, Kim D, Kim K, Cho H Y, Song D W, Lee K J, Jung M H, Kim S, Kim J H, Kim J H, Kim J S. *In vivo* genome editing with a small Cas9 orthologue derived from *Campylobacter jejuni*. *Nat Commun*, 2017, 8: 14500.
- [8] Zetsche B, Gootenberg J S, Abudayyeh O O, Slaymaker I M, Makarova K S, Essletzbichler P, Volz S E, Joung J, van der Oost J, Regev A, Koonin E V, Zhang F. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell*, 2015, 163: 759–771.
- [9] Teng F, Cui T, Gao Q, Guo L, Zhou Q, Li W. Artificial sgRNAs engineered for genome editing with new Cas12b orthologs. *Cell Discov*, 2019, 5: 23.
- [10] McGaw C, Garrity A J, Munoz G Z, Haswell J R, Sengupta S, Keston-Smith E, Hunnewell P, Ornstein A, Bose M, Wessells Q, Jakimo N, Yan P, Zhang H, Alfonse L E, Ziblat R, Carte J M, Lu W C, Cerchione D, Hilbert B, Sothivelvam S, Yan W X, Cheng D R, Scott D A, DiTommaso T, Chong S. Engineered Cas12i2 is a versatile high-efficiency platform for therapeutic genome editing. *Nat Commun*, 2022, 13: 2833.
- [11] Wang Y, Qi T, Liu J, Yang Y, Wang Z, Wang Y, Wang T, Li M, Li

- M, Lu D, Chang A C Y, Yang L, Gao S, Wang Y, Lan F. A highly specific CRISPR-Cas12j nuclease enables allele-specific genome editing. *Sci Adv*, 2023, 9: eabo6405.
- [12] Hirano S, Kappel K, Altae-Tran H, Faure G, Wilkinson M E, Kannan S, Demircioglu F E, Yan R, Shiozaki M, Yu Z, Makarova K S, Koonin E V, Macrae R K, Zhang F. Structure of the OMEGA nickase IsrB in complex with ω RNA and target DNA. *Nature*, 2022, 610: 575–581.
- [13] Altae-Tran H, Kannan S, Demircioglu F E, Oshiro R, Nety S P, McKay L J, Dlakić M, Inskip W P, Makarova K S, Macrae R K, Koonin E V, Zhang F. The widespread IS200/IS605 transposon family encodes diverse programmable RNA-guided endonucleases. *Science*, 2021, 374: 57–65.
- [14] Kim D Y, Lee J M, Moon S B, Chin H J, Park S, Lim Y, Kim D, Koo T, Ko J H, Kim Y S. Efficient CRISPR editing with a hypercompact Cas12f1 and engineered guide RNAs delivered by adeno-associated virus. *Nat Biotechnol*, 2022, 40: 94–102.
- [15] Xin C, Yin J, Yuan S, Ou L, Liu M, Zhang W, Hu J. Comprehensive assessment of miniature CRISPR-Cas12f nucleases for gene disruption. *Nat Commun*, 2022, 13: 5623.
- [16] Wang Y, Wang Y, Pan D, Yu H, Zhang Y, Chen W, Li F, Wu Z, Ji Q. Guide RNA engineering enables efficient CRISPR editing with a miniature *Syntrophomonas palmitatica* Cas12f1 nuclease. *Cell Rep*, 2022, 40: 111418.
- [17] Kong X, Zhang H, Li G, Wang Z, Kong X, Wang L, Xue M, Zhang W, Wang Y, Lin J, Zhou J, Shen X, Wei Y, Zhong N, Bai W, Yuan Y, Shi L, Zhou Y, Yang H. Engineered CRISPR-OsCas12f1 and RhCas12f1 with robust activities and expanded target range for genome editing. *Nat Commun*, 2023, 14: 2046.
- [18] Sukegawa S, Nureki O, Toki S, Saika H. Genome editing in rice mediated by miniature size Cas nuclease SpCas12f. *Front Genome Ed*, 2023, 5: 1138843.
- [19] Jiang Y Y, Chai Y P, Lu M H, Han X L, Lin Q, Zhang Y, Zhang Q, Zhou Y, Wang X C, Gao C, Chen Q J. Prime editing efficiently generates W542L and S621I double mutations in two ALS genes in maize. *Genome Biol*, 2020, 21: 257.
- [20] Bähler J, Wise J A. Preparation of total RNA from fission yeast. *Cold Spring Harb Protoc*, 2017, 2017: pdb. prot091629.
- [21] Cao J, Yao D, Lin F, Jiang M Y. PEG-mediated transient gene expression and silencing system in maize mesophyll protoplasts: a valuable tool for signal transduction study in maize. *Acta Physiol Plant*, 2014, 36: 1271–1281.
- [22] Karvelis T, Bigelyte G, Young J K, Hou Z, Zedaveinyte R, Budre K, Paulraj S, Djukanovic V, Gasior S, Silanskas A, Venclovas Č, Siksnys V. PAM recognition by miniature CRISPR-Cas12f nucleases triggers programmable double-stranded DNA target cleavage. *Nucleic Acids Res*, 2020, 48: 5016–5023.
- [23] Xiao R, Li Z, Wang S, Han R, Chang L. Structural basis for substrate recognition and cleavage by the dimerization-dependent CRISPR-Cas12f nuclease. *Nucleic Acids Res*, 2022, 49: 4120–4128.
- [24] Bigelyte G, Young J K, Karvelis T. Miniature type V-F CRISPR-Cas nucleases enable targeted DNA modification in cells. *Nat Commun*, 2021, 12: 6191.
- [25] Hill Z B, Martinko A J, Nguyen D P, Wells J A. Human antibody-based chemically induced dimerizers for cell therapeutic applications. *Nat Chem Biol*, 2018, 14: 112–117.