Genome Biology

RESEARCH

Open Access

Post-cleavage target residence determines asymmetry in non-homologous end joining of Cas12a-induced DNA double strand breaks

Ruo-Dan Chen[1,2,3†], Yi Yang[1,3†], Kun-Ming Liu[1,3], Jing-Zhen Hu[1,3], Yi-Li Feng[1,2,3], Chun-Yi Yang[1,3], Rui-Rui Jiang[3,4], Si-Cheng Liu[1,2], Yue Wang[1,3], Ping-An Han[5], Ru-Gang Tian[5], Yu-Long Wang[4], Shi-Ming Xu[1,3*] and An-Yong Xie[1,2,3*]

†Ruo-Dan Chen and Yi Yang contributed equally to this work.

*Correspondence: shimingxu@zju.edu.cn; anyongxie@zju.edu.cn

[1] Key Laboratory of Laparoscopic Technology of Zhejiang Province, Department of General Surgery, Sir Run-Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang 310016, People's Republic of China
Full list of author information is available at the end of the article

Abstract

Background: After Cas12a cleaves its DNA target, it generates a DNA double strand break (DSB) with two compatible 5′-staggered ends. The Cas12a-gRNA complex remains at the protospacer adjacent motif (PAM)-proximal end (PPE) while releasing the PAM-distal end (PDE). The effects of this asymmetric retention on DSB repair are currently unknown.

Results: Post-cleavage retention of *Lb*Cas12a at PPEs suppresses the recruitment of classical non-homologous end joining (c-NHEJ) core factors, leading to longer deletions at PPEs compared to PDEs. This asymmetry in c-NHEJ engagement results in approximately tenfold more accurate ligation between two compatible PDEs induced by paired *Lb*Cas12a than ligation involving a compatible PPE. Moreover, ligation to a given end of *Sp*Cas9-induced DSBs demonstrates more efficient ligation with a PDE from Cas12a-induced DSBs than with a PPE. In *Lb*Cas12a-induced NHEJ-mediated targeted integration, only two compatible PDEs from *Lb*Cas12a-induced DSBs—one from donor templates and the other from target sites—promote accurate and directional ligation. Based on these findings, we developed a strategy called Cas12a-induced PDE ligation (CIPDEL) for NHEJ-mediated efficient and precise gene correction and insertion.

Conclusions: The asymmetric retention of CRISPR-*Lb*Cas12a at DSB ends suppresses c-NHEJ at PPEs, not at PDEs. This unique repair mechanism can be utilized in the CIPDEL strategy, offering a potentially better alternative for homology-directed targeted integration.

Keywords: CRISPR-Cas12a, Asymmetric retention, PAM-distal end, Classical NHEJ, Targeted integration

Background

Like the class II type II CRISPR-Cas nuclease Cas9, the class II type V nuclease Cas12a (also known as Cpf1) has been widely used for genome editing in both plant and mammalian cells [1, 2]. Compared to *Streptococcus pyogenes* Cas9 (*Sp*Cas9), some Cas12a nucleases, such as *Lachnospiraceae bacterium* Cas12a (*Lb*Cas12a), *Acidaminococcus* sp. Cas12a (*As*Cas12a), and *Francisella novicida* Cas12a (*Fn*Cas12a), are smaller in size [3, 4]. A study has shown that the small size of Cas12a is amenable for in vivo delivery alongside guide RNA (gRNA) via a single recombinant adeno-associated virus (rAAV) vector [5]. The Cas12a gRNA consists solely of a CRISPR-related RNA (crRNA), which is significantly shorter than the Cas9 gRNA that includes both a crRNA and a trans-activating crRNA (tracrRNA) [3]. Cas12a also possesses a ribonuclease activity and can process pre-crRNA arrays into individual mature crRNAs, allowing for compact gRNA arrays in a single transcript. This enables Cas12a to mediate multiplexed genome editing in cells [4]. Furthermore, Cas12a exhibits high specificity in human cells and has a considerably lower off-target effect compared to *Sp*Cas9 [6–8]. Given these unique features, Cas12a presents advantages for certain genome editing applications.

Both CRISPR-Cas9 and CRISPR-Cas12a induce site-specific DNA double strand breaks (DSBs), which are subsequently repaired for genome editing in mammalian cells. Two major DSB repair mechanisms—homology-directed repair (HDR) and non-homologous end joining (NHEJ)—can be harnessed to improve CRISPR-Cas genome editing. HDR, which uses homologous sequences as templates, is the preferred pathway for targeted integration, allowing for accurate substitutions and insertions in CRISPR-Cas genome editing [9]. Enhancing HDR activity in cells remains an important strategy for improving targeted integration by CRISPR-Cas [9, 10]. Classical NHEJ (c-NHEJ) is the primary NHEJ pathway, requiring core factors such as DNA-PKcs, KU70/KU80, and XRCC4/DNA ligase 4. In contrast, alternative end joining (a-EJ) can ligate DSB ends if any c-NHEJ core factors are deficient or not engaged [11, 12]. Unlike a-EJ, c-NHEJ is inherently accurate for readily ligatable ends of DSBs [13]. Previous studies have demonstrated that blunt ends of SpCas9-induced DSBs and 3′-overhanging ends of I-SceI-induced DSBs are predominantly rejoined accurately by c-NHEJ [14]. Given that Cas12a-induced DSBs generate readily ligatable ends with compatible 5′-nucleotide (nt) overhangs, it is expected that NHEJ repair of Cas12a-induced DSBs would also be mostly accurate; however, this prediction has yet to be validated.

Paradoxically, while accurate repair of *Sp*Cas9-induced DSBs by c-NHEJ facilitates efficient homology-independent targeted integration (HITI) [15], Cas12a-induced NHEJ-mediated targeted integration and the ligation of compatible ends from Cas12a-induced DSBs appear to be generally inefficient or inaccurate [16, 17]. *Sp*Cas9-induced HITI is frequently employed for targeted integration, particularly in non-proliferating cells and proliferating cells lacking active HDR [15, 18]. Unlike *Sp*Cas9, which induces blunt-ended DSBs, Cas12a generates DSB ends with 5-nt 5′ overhangs [3], which may influence the direction of accurate ligation in NHEJ-mediated targeted integration based on overhang compatibility. Considering the relatively small size of Cas12a-gRNA, compact gRNA arrays in a single transcript, and minimal off-target effects, further improvements in Cas12a-induced NHEJ-mediated targeted integration are beneficial for practical applications.

Both Cas9-gRNA and Cas12a-gRNA bind their targets strongly and persistently due to several interactions, including contacts between the Cas protein and the protospacer adjacent motif (PAM) of the target, base pairing of the gRNA spacer with the target strand, and non-specific interactions between the Cas protein and target DNA [19]. Even after the cleavage of target DNA, the Cas-gRNA complex remains tightly bound to the cleaved DNA and persists for hours [8, 20–24]. However, unlike the *Sp*Cas9-gRNA complex, which binds to both DSB ends after target cleavage and dissociates from both ends simultaneously, the Cas12a-gRNA complex is retained only at the PAM-proximal end (PPE), while the PAM-distal end (PDE) is released from the Cas12a-gRNA-target DNA ternary complex upon cleavage [8, 20–24]. Our previous study demonstrated that the post-cleavage target residence of SpCas9-gRNA regulates DSB repair pathway choices in CRISPR-Cas9 genome editing [25]. It remains unclear whether the post-cleavage target residence of Cas12a-gRNA affects genome editing, and if so, how. Nevertheless, studies have shown that the rate of Cas12a-induced HDR is higher with a single-stranded oligodeoxynucleotide (ssODN) donor that is complementary to the target strand of a Cas12a target site than to the non-target strand [26, 27]. Cas12a-induced NHEJ-mediated targeted integration appears more accurate when the PDEs, rather than the PPEs, are engaged [16]. This suggests that the asymmetry in post-cleavage target residence of Cas12a-gRNA may influence the repair of Cas12a-induced DSBs. A better understanding of the underlying mechanism could facilitate further improvements and applications of these Cas12a-induced genome editing strategies.

Here, we found that in NHEJ of *Lb*Cas12a-induced DSBs (but not *Sp*Cas9-induced DSBs), the average deletion length at the PPE of *Lb*Cas12a-induced DSBs was longer than at the PDE. This deletion asymmetry was abolished in cells lacking any of the c-NHEJ core factors, including DNA-PKcs, KU80, and XRCC4. Chromatin immunoprecipitation (ChIP) demonstrated that PDEs engage KU80 more actively than PPEs, while PPEs are more occupied by *Lb*Cas12a-gRNA after DNA cleavage. This suggests that the retention of *Lb*Cas12a-gRNA at PPEs prevents the binding of c-NHEJ core factors to these ends. Using paired *Lb*Cas12a to induce a DSB with two compatible cohesive PDEs or two compatible cohesive ends involving at least one PPE at 44 target sites, we found that NHEJ of two PDEs achieved close to 50% accuracy, while ligation involving a PPE reduced the accuracy to less than 5%. This indicates that ligation between compatible PDEs of Cas12a-induced DSBs tends to be more accurate. By analyzing ligation to a given end of *Sp*Cas9-induced DSBs, we observed more efficient ligation with the PDEs of Cas12a-induced DSBs than with the PPEs. Further analysis of *Lb*Cas12a-induced NHEJ-mediated targeted integration with either compatible PDEs or compatible PPEs between a donor template and an integration site confirmed that two compatible PDEs of *Lb*Cas12a-induced DSBs—one from the donor template and the other from the integration site—promote accurate ligation. Based on this mechanistic difference between ligation involving PDEs and PPEs of Cas12a-induced DSBs, we devised a strategy termed Cas12a-induced PDE ligation (CIPDEL) for efficient and precise gene correction and insertion, accurately replacing a gene fragment of interest at three selected genomic sites.

Results

NHEJ repair of *Lb*Cas12a-induced DSBs generates asymmetric deletions at two ends

Unlike *Sp*Cas9, which retains its binding at both ends of the DSB [20, 21], Cas12a remains bound only to the PAM-proximal end (PPE) while releasing the PAM-distal end (PDE) [22–24] (Fig. 1a). We asked whether this retention asymmetry affects deletion lengths at the two DSB ends during NHEJ repair of *Lb*Cas12a-induced DSBs. We targeted *Sp*Cas9 to 15 genomic sites and *Lb*Cas12a to 16 genomic sites in mouse and human cells, analyzing the repair junctions via targeted amplicon deep sequencing. The average frequency of insertion/deletion (indel) induced was 51.4% for *Sp*Cas9 and 61.6% for *Lb*Cas12a (Fig. 1b). In NHEJ repair of *Sp*Cas9-induced DSBs, deletions were generally symmetric between the PPEs and PDEs, with an average difference in deletion length of only about −1 bp and a ratio of average deletion lengths at the PPEs to those at the PDEs close to 0.937 (Fig. 1c,d and Additional file 1: Fig. S1a,b). In contrast, NHEJ repair of *Lb*Cas12a-induced DSBs showed an average difference in deletion length of nearly 5 bp, with a ratio of average deletion lengths at the PPEs to those at the PDEs around 1.730 (Fig. 1c,d and Additional file 1: Fig. S2a,b). This indicates that *Lb*Cas12a-mediated cleavage induces asymmetric deletions, with longer deletions at the PPEs compared to

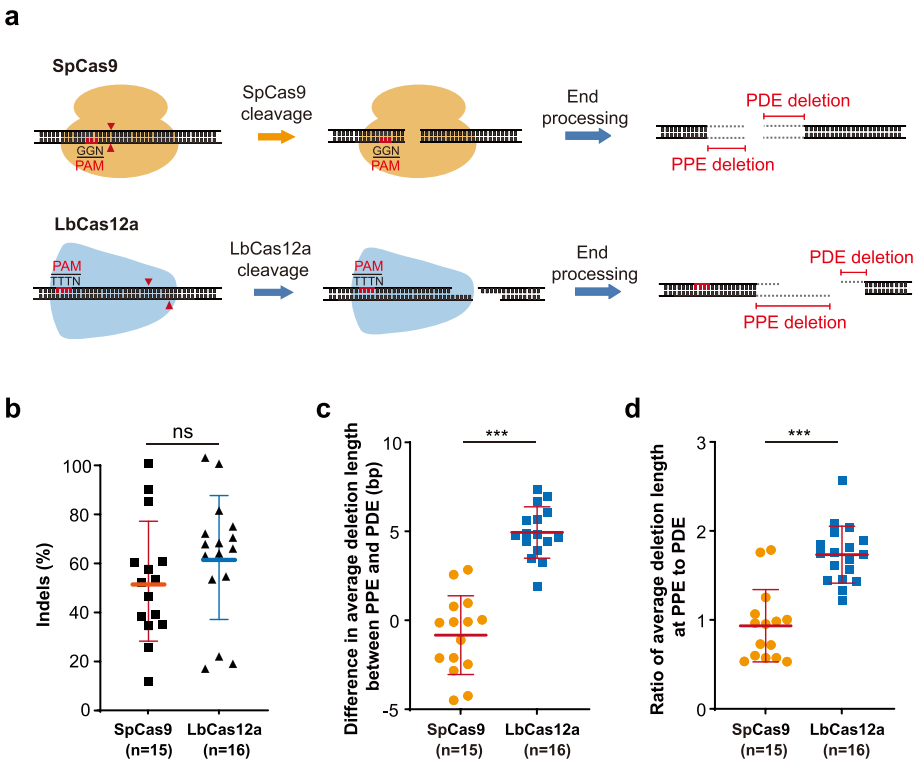

Fig. 1 NHEJ repair of *Lb*Cas12a-induced DSBs generates asymmetric deletions at both ends. **a** Schematic illustrating post-cleavage target residency of *Sp*Cas9 and *Lb*Cas12a, and their potential effects on deletion lengths at the PPE and the PDE of DSBs induced by *Sp*Cas9 and *Lb*Cas12a. **b** Frequencies of targeted genome editing by *Sp*Cas9-gRNA at 15 target sites (7 in mouse cells and 8 in human cells) and by *Lb*Cas12a-gRNA at 16 target sites (3 in mouse cells and 13 in human cells). **c** Differences in average deletion lengths between the PPE and the PDE for *Sp*Cas9-induced DSBs at 15 target sites and *Lb*Cas12a-induced DSBs at 16 target sites. **d** Ratio of average deletion length at the PPE compared to the PDE. Each symbol represents one target site in **b**–**d**. Statistical analysis was conducted using Student's two-tailed unpaired *t*-test. ns: $P > 0.05$; ***: $P < 0.001$

the PDEs. In *Lb*Cas12a-induced DSBs, "insertions only" products were found at a very low frequency. In contrast, *Sp*Cas9-mediated cleavage produced a diverse range of NHEJ products, including "deletions only," "insertions only," and "indel" mixtures (Additional file 1: Fig. S3a,b). The "insertions only" NHEJ products from *Sp*Cas9-induced DSBs were primarily 1-nt templated insertions, as previously demonstrated (Additional file 1: Fig. S3c) [14, 28–30].

Asymmetric deletion in NHEJ of *Lb*Cas12a-induced DSBs involves c-NHEJ

As c-NHEJ is the primary NHEJ repair pathway for DSBs in mammalian cells, we asked whether c-NHEJ is engaged in asymmetric deletion at both ends during the NHEJ repair of *Lb*Cas12a-induced DSBs. We re-examined NHEJ at the *Rosa26* locus targeted by *Lb*Cas12a using guide RNAs gR26 - 5c, gR26 - 6 d, and gR26 - 8c in various mouse embryonic stem (ES) cell lines: isogenic wild-type (WT), *DNA-PKcs*$^{-/-}$, *Ku80*$^{-/-}$, *XRCC4*$^{+/+}$, and *XRCC4*$^{-/-}$. Loss of *DNA-PKcs* reduced the frequency of *Lb*Cas12a-induced indels only at the gR26 - 5c site. In contrast, the absence of *Ku80* affected indel frequencies at both gR26 - 5c and gR26 - 6 d sites, while loss of *XRCC4* impacted all three target sites compared to their respective WT cells (Fig. 2a).

Analysis of asymmetric end deletions in NHEJ products revealed that the average deletion length at the PPEs was consistently longer than that at the PDEs across the three

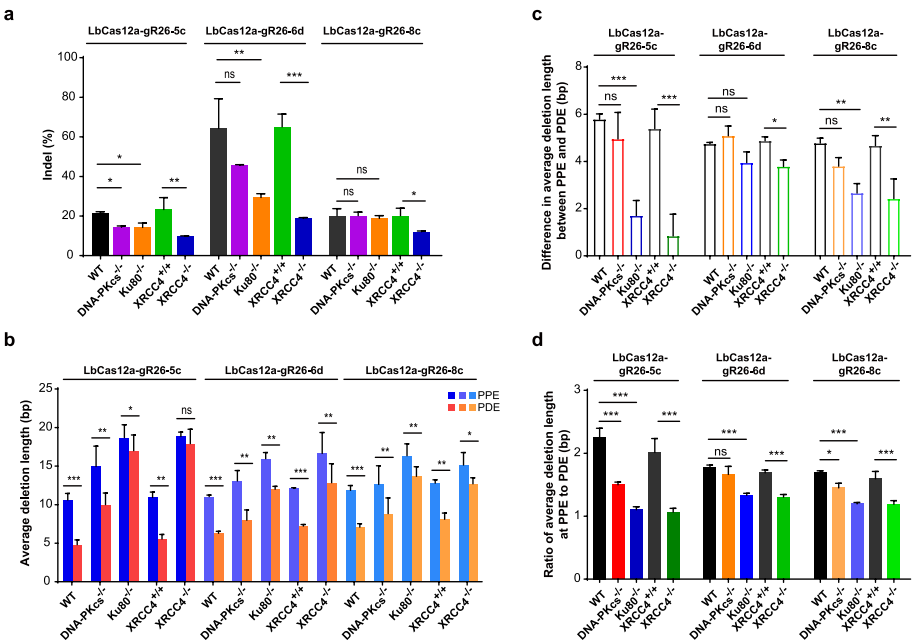

Fig. 2 Asymmetric deletion in the repair of *Lb*Cas12a-induced DSBs requires c-NHEJ factors. **a** Frequencies of indels induced by NHEJ repair of *Lb*Cas12a-induced DSBs in isogenic WT, *DNA-PKcs*$^{-/-}$ and *Ku80*$^{-/-}$ mouse ES cells, isogenic *XRCC4*$^{+/+}$ and *XRCC4*.$^{-/-}$ mouse ES cells at three *Rosa26* sites. **b** Average deletion lengths at the PPEs and PDEs in NHEJ repair of *Lb*Cas12a-induced DSBs. **c** Differences in average deletion length between the PPE and the PDE of *Lb*Cas12a-induced DSBs at the three *Rosa26* sites. **d** Ratios of average deletion lengths at the PPE to those at the PDE of *Lb*Cas12a-induced DSBs at the three *Rosa26* sites. Columns with error bars represent the mean ± S.D. of three independent experiments. Statistical analyses employed Student's one-tailed paired *t*-test (**a**), Student's two-tailed paired *t*-test (**b**), or one-way ANOVA with Tukey's multiple comparison test (**c**, **d**). ns: $P > 0.05$; *: $P < 0.05$; **: $P < 0.01$; ***: $P < 0.001$

target sites in WT cells and *XRCC4*$^{+/+}$ cells (Fig. 2b). Although the loss of *DNA-PKcs*, *Ku80*, or *XRCC4* increased deletion lengths at both ends, the difference between the PPEs and PDEs diminished in *DNA-PKcs*$^{-/-}$, *Ku80*$^{-/-}$, and *XRCC4*$^{-/-}$ cells (Fig. 2c). Specifically, the difference in average deletion length between PPEs and PDEs was approximately 5–6 bp in WT and *XRCC4*$^{+/+}$ cells but decreased in cells lacking *Ku80* or *XRCC4* (Fig. 2c). The ratios of average deletion lengths between the PPEs and PDEs dropped from 1.7–2.2 in WT and *XRCC4*$^{+/+}$ cells to nearly 1 in *Ku80*$^{-/-}$ and *XRCC4*$^{-/-}$ cells (Fig. 2d). This indicates that loss of c-NHEJ factors *Ku80* or *XRCC4* promotes symmetric deletion at both ends of *Lb*Cas12a-induced DSBs after end ligation. Loss of *DNA-PKcs* had a comparatively weaker effect (Fig. 2b–d), supporting the idea that the absence of *Ku80* or *XRCC4* more severely compromises c-NHEJ than the loss of *DNA-PKcs* [31]. These findings suggest that c-NHEJ factors are involved in asymmetric deletion during NHEJ repair of *Lb*Cas12a-induced DSBs.

Given the observed asymmetric deletion and the involvement of c-NHEJ factors, we hypothesized that their recruitment is asymmetric due to the preferential post-cleavage retention of Cas12a at the PPEs, rather than at the PDEs (Fig. 3a). We first analyzed the efficiency of target cleavage by both HA-*Lb*Cas12a-gW and HA-*Lb*Cas12a-gC and found that the cleavage was efficient in mouse ES cells (Fig. 3b,c). We then performed ChIP assays to assess the post-cleavage retention of *Lb*Cas12a at DSB ends induced by HA-*Lb*Cas12a-gW or HA-*Lb*Cas12a-gC at 12 h, 24 h, and 48 h post-transfection. The results revealed significant enrichment of *Lb*Cas12a at both the PPEs and the PPEs from 12 to 24 h in mouse ES cells, but the enrichment was reduced from 24 to 48 h (Fig. 3b,c and Additional file 1: Fig. S4). The enrichment was greater at the PPEs than at the PDEs at 12 h and 24 h while displaying little difference between the PPEs and the PDEs at 48 h. This confirms that *Lb*Cas12a-sgRNA remains more bound to the PPEs than to the PDEs upon target cleavage. In contrast, Ku80 demonstrated increased enrichment at the PDEs at both the gC and gW target from 12 to 48 h post-transfection; however, the enrichment of Ku80 at the PPEs was not elevated from 12 to 24 h, but greatly increased from 24 to 48 h (Fig. 3b,c and Additional file 1: Fig. S4). This suggests that the post-cleavage residence of Cas12a at the PPEs inhibits the recruitment of c-NHEJ factors to these sites, while permitting their recruitment to the PDEs.

Ligation between compatible cohesive PDEs of Cas12a-induced DSBs tends to be more accurate

The recruitment of c-NHEJ factors to DSB ends protects the ends from nuclease activity and facilitates a more active, less mutagenic repair process compared to a-EJ [11, 12]. Previous studies have indicated that c-NHEJ of compatible DSB ends is inherently accurate [13, 14, 32]. To confirm that compatible cohesive PDEs with 5-nt overhangs are repaired more accurately than compatible cohesive PPEs in *Lb*Cas12a-induced DSBs, we employed paired *Lb*Cas12a-sgRNAs to generate two compatible ends in close proximity for NHEJ (Fig. 4a). The paired *Lb*Cas12a-sgRNAs cleaved target DNA by four combinations of PAM orientations: W/W, W/C, C/W, and C/C, where W denotes the Watson strand and C denotes to the Crick strand (Fig. 4a). Simultaneous cutting by paired *Lb*Cas12a could remove the intervening DNA sequence between two breakage sites, leaving two compatible ends, either a

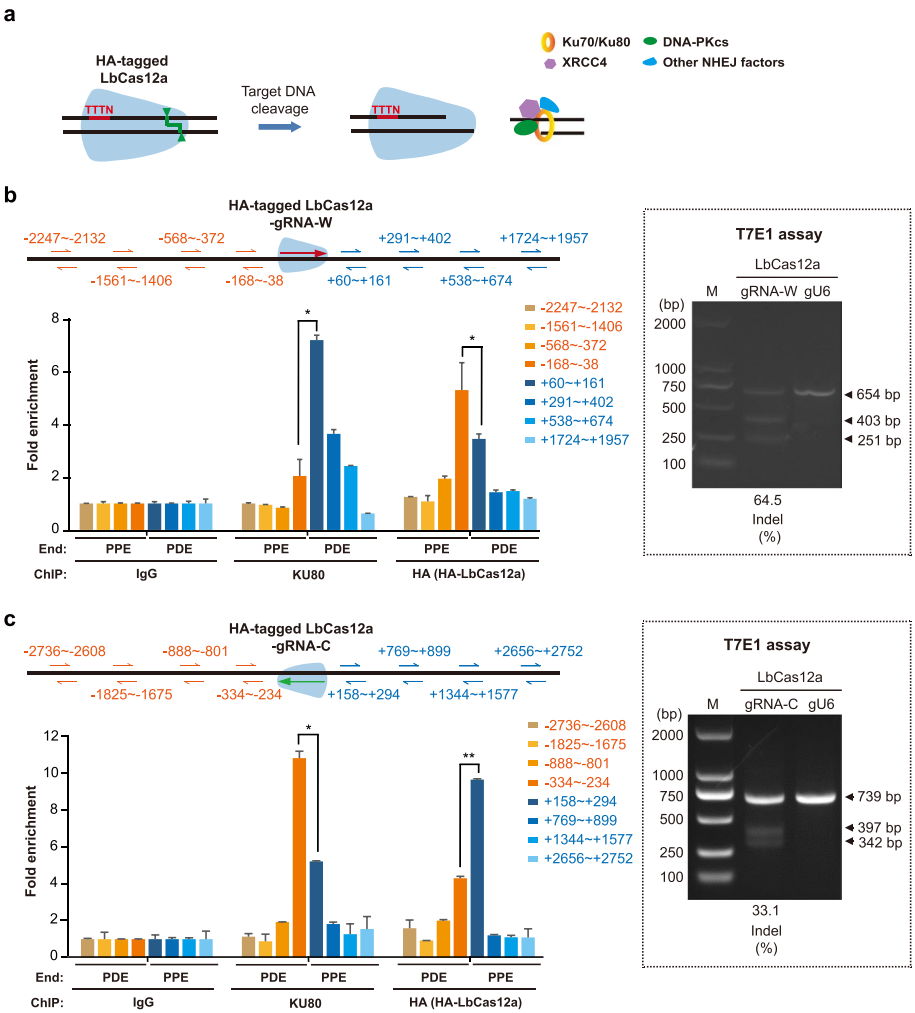

Fig. 3 Post-cleavage residency of *Lb*Cas12a at the PPEs inhibits recruitment of Ku80 to the ends. **a** Schematic depicting end-binding competition between HA-tagged *Lb*Cas12a and c-NHEJ core factors at *Lb*Cas12a-induced DSBs. Ku70/Ku80 is expected to bind the PDE of *Lb*Cas12a-induced DSBs, while *Lb*Cas12a-gRNA remains bound at the PPE. Enrichment of HA-tagged *Lb*Cas12a and Ku80 at the PPE and PDE of a DSB induced by *Lb*Cas12a-gRNA-W in the SCR-RFP reporter (**b**) and by *Lb*Cas12a-gRNA-C in the NHEJ reporter (**c**). Mouse ES cells containing either the SCR-RFP reporter or the NHEJ reporter were transfected with expression plasmids for *Lb*Cas12a and the corresponding gRNA. Inset in **b** and **c** indicates T7E1 detection of on-target cleavage by *Lb*Cas12a-gRNA-W and *Lb*Cas12a-gRNA-C at 48 h post-transfection, along with gU6 as a negative gRNA control. In these insets, uncut and cut PCR bands are indicated, along with M for DNA marker, and the indel ratios were calculated according to the band intensities with ImageJ software. ChIP analysis was performed at 24 h post-transfection using anti-IgG as a background control (left), anti-Ku80 antibody for Ku80 enrichment (middle), and anti-HA antibody for HA-*Lb*Cas12 enrichment (right). Fold enrichment of HA-*Lb*Cas12a and Ku80 was assessed by real-time qPCR amplification using primer pairs located at varying distances from the *Lb*Cas12a cleavage site, as indicated by the orange and blue arrows. The fold enrichment relative to the negative IgG control was normalized to 1. Columns with error bars represent the mean ± S.D. of three independent experiments, with statistical significance detected by Student's two-tailed paired *t*-test in **b** and **c**. *: $P < 0.05$; **: $P < 0.01$

PDE or a PPE, in the genome to be ligated by NHEJ. These two ends generated by paired *Lb*Cas12a for each combination of PAM orientations are as follows: PPE/PDE, PPE/PPE, PDE/PDE, and PDE/PPE (Fig. 4a). NHEJ of these ends leads to the deletion of intervening sequences between paired *Lb*Cas12a sites. We defined this

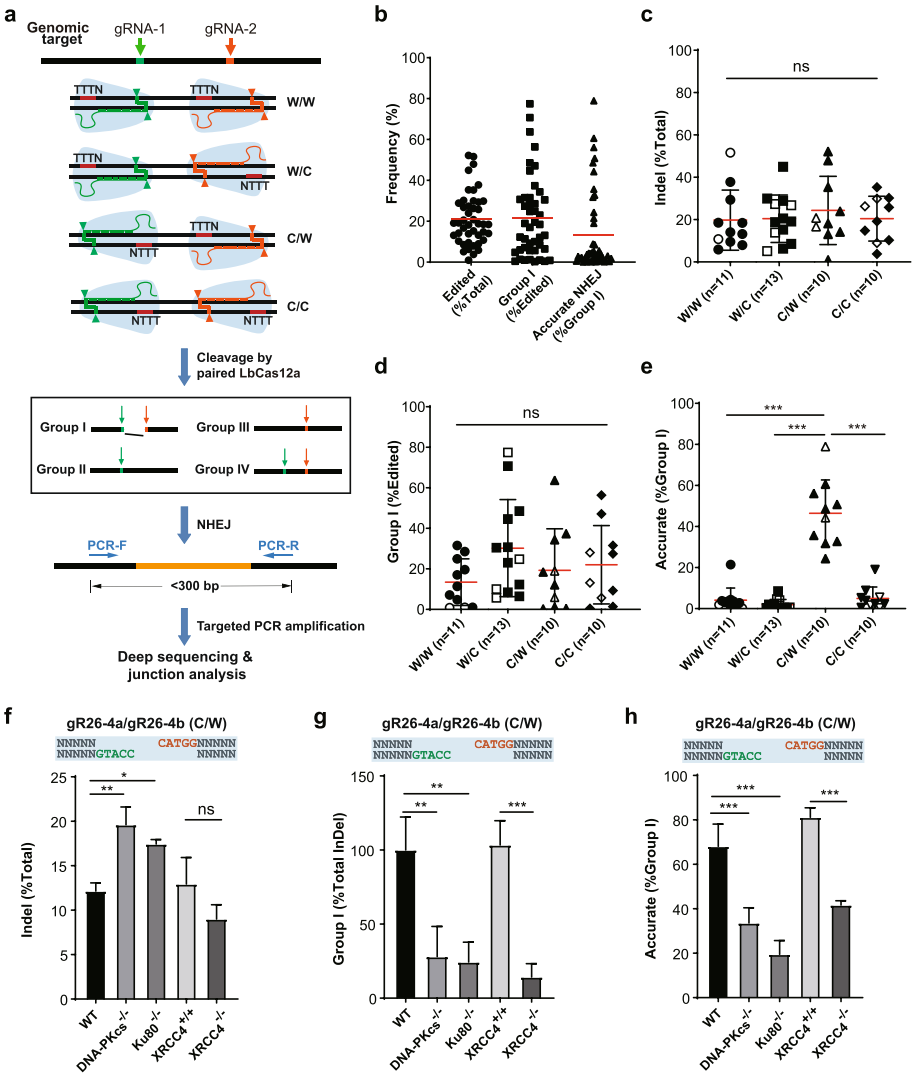

Fig. 4 Ligation between compatible cohesive ends at the PDEs of *Lb*Cas12a-induced DSBs tends to be more accurate. **a** Schematic representation of paired *Lb*Cas12a-gRNAs with four combinations of PAM orientations (W/W, W/C, C/W, and C/C) to generate two compatible cohesive ends in close proximity for NHEJ repair. **b** The frequency of genome editing induced by paired *Lb*Cas12a that generate compatible cohesive ends at 44 endogenous genome sites (left), the frequency of group I events among edited events (middle), and the accuracy of events within group I (right). Editing frequency (**c**), group I frequency among edited events (**d**), and accurate NHEJ rates in group I events (**e**) for each of the four paired PAM orientations at 44 endogenous genome sites. Among 11 pairs of targets in mouse ES cells (indicated with hollow symbols), there are 2 pairs of W/W, 4 of W/C, 2 of C/W, and 3 of C/C. Among 33 pairs of targets in human 293 cells, there are 9 pairs of W/W, 9 of W/C, 8 of C/W, and 7 of C/C. Indel frequency (**f**), group I frequency among edited events (**g**), and accurate NHEJ rates in group I events (**h**) induced by *Lb*Cas12a together with the gR26 - 4a (C)/gR26 - 4b (W) pair in isogenic WT, *DNA-PKcs*$^{-/-}$, and *Ku80*$^{-/-}$ mouse ES cells, as well as isogenic *XRCC4*$^{+/+}$ and *XRCC4.*$^{-/-}$ mouse ES cells. Each symbol in **b**–**e** represents the average value from three independent experiments for one sgRNA pair, with red lines and error bars indicating the mean ± S.D. for all gRNA pairs. Columns with error bars in **f**–**h** represent the mean ± S.D. from three independent experiments. Statistical analysis was performed using one-way ANOVA with Tukey's multiple comparison test in **b**–**e** and Student's two-tailed paired *t*-test in **f**–**h**. *: $P < 0.05$; **: $P < 0.01$; ***: $P < 0.001$

type of NHEJ products as "group I events," consistent with our previous study [14]. Group I events were thus used to determine accurate joining of DSB ends induced by paired *Lb*Cas12a. Groups II, III, and IV represent mutagenic rejoining of DSB

ends induced by a single *Lb*Cas12a without intervening DNA deletion as did previously with *Sp*Cas9 [14].

After normalizing for transfection efficiencies, the levels of induced group I events and their accuracy varied significantly across 44 target pairs, comprising 11 pairs in mouse ES cells and 33 pairs in human 293 cells (Fig. 4b). To assess whether the type of end (PDE or PPE) influenced the efficiency and accuracy of group I events, we categorized the 44 pairs into 11 W/W, 13 W/C, 10 C/W, and 10 C/C pairs. The efficiency and accuracy of group I events showed minimal differences among these PAM combinations (Fig. 4c,d). However, end joining accuracy for group I events in C/W pairs ranged from 20.0 to 80.0%, averaging 46.7%, significantly higher than the approximately 3.6% average for W/W, W/C, and C/C pairs (Fig. 4e). The loss of *DNA-PKcs*, *Ku80*, or *XRCC4* did not reduce the level of indels induced by C/W-paired *Lb*Cas12a at the mRosa #4 site but significantly decreased the frequency of group I events among total indels and the accuracy of events within group I (Fig. 4f–h). Since paired *Lb*Cas12a in the C/W orientation produces two PDEs, while W/W, W/C, and C/C produce at least one PPE, these findings suggest that accurate joining of compatible cohesive ends is more effective when both ends are accessible to c-NHEJ core factors.

In accurate NHEJ for group I events, only the intervening sequences between paired *Lb*Cas12a sites are deleted without additional indels. In contrast, mutagenic group I events result in extra deletions due to end processing, with the average length of these additional deletions reflecting the overall deletion length. Analysis indicated that the average deletion length in mutagenic group I events was shorter for the 10 C/W pairs compared to the 11 W/W pairs, 13 W/C pairs, and 10 C/C pairs (Additional file 1: Fig. S5a). Additionally, loss of *Ku80* or *XRCC4* increased the average deletion length in group I products induced at the m*Rosa26* loci by C/W-paired *Lb*Cas12a with gR26 -4a/gR26 -4b, while loss of *DNA-PKcs* had minimal impact (Additional file 1: Fig. S5b). These results confirm that, unlike PPEs in *Lb*Cas12a-induced DSBs, PDEs preferentially engage c-NHEJ, resulting in shorter deletion lengths and a higher frequency of accurate ligation between two PDEs.

PDEs of Cas12a-induced DSBs are more efficiently ligated than PPEs

Given the greater recruitment of c-NHEJ core factors to the PDEs of *Lb*Cas12a-induced DSBs compared to the PPEs, we predicted that ligation to the PDEs would be more efficient than to the PPEs in the presence of c-NHEJ. To test this hypothesis, we used *Sp*Cas9-g10 to induce a site-specific DSB with two blunt ends at the short arm of chromosome 3 in human 293 T cells, facilitating ligation with either the PDE or the PPE of the DSB induced by *Lb*Cas12a-g2 or *Lb*Cas12a-g2r at the long arm of chromosome 8 (Fig. 5a). With target cleavage by *Sp*Cas9-g10, *Lb*Cas12a-g2, and *Lb*Cas12a-g2r respectively at 37.5%, 54.2%, and 62.3% (Additional file 1: Fig. S6a), four types of translocations could be induced. We constructed 4 plasmids, each containing a respective translocation template for one translocation type, and used them with respective primer pairs to generate qRT-PCR standard curves for quantification of each translocation type (Additional file 1: Fig. S6b,c). As expected, translocations were minimally detected by qRT-PCR with primer pairs F1/R2, F2/R1, F1/F2, and R1/R2 for *Lb*Cas12a-g2 and F1/R2, F2/R1, F1/F2, and R1/R2 for *Lb*Cas12a-g2r in the absence of DSBs. However, a significant level of

translocations were detected by F2/R1 and F1/F2 for *Lb*Cas12a-g2 and by F1/R2 and R1/R2 for *Lb*Cas12a-g2r, not by F1/R2 and R1/R2 for *Lb*Cas12a-g2 or by F2/R1 and F1/F2 for *Lb*Cas12a-g2r, in the presence of *Sp*Cas9- and *Lb*Cas12a-induced DSBs (Fig. 5b,c). The detection with F2/R1 and F1/F2 for *Lb*Cas12a-g2 and with F2/R1 and F1/F2 for *Lb*Cas12a-g2r represent the translocation of the PDE from the *Lb*Cas12a-induced DSB to either end of the *Sp*Cas9-induced DSB, while the detection with F1/F2 and R1/R2 for *Lb*Cas12a-g2 and with F1/R2 and R1/R2 for *Lb*Cas12a-g2r represent the translocation of the PPE from the *Lb*Cas12a-induced DSB. This suggests that ligation of a PDE to a given end is more efficient than ligation of a PPE in the presence of c-NHEJ.

Further, we used *Sp*Cas9-g5 to target chromosome 3 and either *Lb*Cas12a-gC or *Lb*Cas12a-gCr to target the *AAVS1* locus on the long arm of chromosome 19 in 293 T cells (Fig. 5d). Target cleavage by *Sp*Cas9-g5, *Lb*Cas12a-gC, and *Lb*Cas12a-gCr at 52.2%, 27.8%, and 36.4%, respectively (Additional file 1: Fig. S6 d). Four types of translocations were generated and measured by qRT-PCR assisted with standard curves (Additional file 1: Fig. S6e,f). As expected, the absence of DSBs with *Sp*Cas9-gU6 and *Lb*Cas12a-gU6 induced a trace level of translocations measured by qRT-PCR (Fig. 5e,f). Consistently, in the presence of DSBs induced by *Sp*Cas9-g5, *Lb*Cas12a-gC, and *Lb*Cas12a-gCr, translocations between the PDEs of the *Lb*Cas12a-induced DSBs and the ends of the *Sp*Cas9-induced DSBs were more frequent than those between the PPEs of the *Lb*Cas12a-induced DSBs and the ends of the *Sp*Cas9-induced DSBs (Fig. 5e,f).

We also examined four types of translocations induced by *Lb*Cas12a-g2 and *Sp*Cas9-g10, as well as NHEJ of the two local ends, using PCR with the respective primer pair combinations followed by agarose gel electrophoresis (Additional file 1: Fig. S7a,b). The PCR products of the translocations were subcloned into a sequencing vector and confirmed by Sanger sequencing (Additional file 1: Fig. S7c–f). As expected for individual target sites, significant PCR products were generated with primer pairs F1/R1 and F2/R2, regardless of whether DSBs were induced by *Sp*Cas9 or *Lb*Cas12a (EV vs. *Sp*Cas9-g10 or EV vs. *Lb*Cas12a-g2 in Additional file 1: Fig. S7a). No translocation was detected

(See figure on next page.)

Fig. 5 *Lb*Cas12a-sgRNA induces translocations more efficiently with the PDEs than with the PPEs. **a** Schematic representation of chromosomal translocations involving a *Sp*Cas9-induced DSB at the *Sp*Cas9-g10 target site on chromosome 3 and an *Lb*Cas12a-induced DSB at the *Lb*Cas12a-g2 or *Lb*Cas12a-g2r target site on chromosome 8. Primer pairs are indicated for qRT-PCR of their respective products: F1/R2 (1), F2/R1 (2), F1/F2 (3), and R1/R2 (4) for the reciprocal translocation between chromosomes 3 and 8. **b**, **c** Translocation products were detected by qRT-PCR with gDNA from cells transfected by *Sp*Cas9-g10 and *Lb*Cas12a-g2 or *Lb*Cas12a-g2r, along with gU6 as a negative gRNA control, respectively. The amounts of translocation products (pmol) in gDNA (ng) were calculated using standard curves. The PDEs of long and short chromosomal fragments are indicated by "long" and "short" in parentheses. **d** Schematic for chromosomal translocations involving a *Sp*Cas9-induced DSB at the *Sp*Cas9-g5 target site on chromosome 3 and an *Lb*Cas12a-induced DSB at the *Lb*Cas12a-gC or *Lb*Cas12a-gCr target site on chromosome 19. Primer pairs are indicated for qRT-PCR: F3/R4 (1), F4/R3 (2), F3/F4 (3), and R3/R4 (4) for the reciprocal translocation. **e**, **f** Translocation products were detected by qRT-PCR with gDNA from cells transfected by *Sp*Cas9-g5 and *Lb*Cas12a-gC or *Lb*Cas12a-gCr, along with gU6 as a negative gRNA control, respectively. The amounts of translocation products (pmol) in gDNA (ng) were calculated using standard curves. The PDEs of long and short chromosomal fragments are indicated by "long" and "short" in parentheses. Columns with error bars represent the mean ± S.D. of three independent experiments in **b**, **c**, **e**, and **f**. Statistical analysis was performed using Student's two-tailed paired *t*-test, with significance indicated as ns: $P > 0.05$; *: $P < 0.05$; **: $P < 0.01$; ***: $P < 0.001$

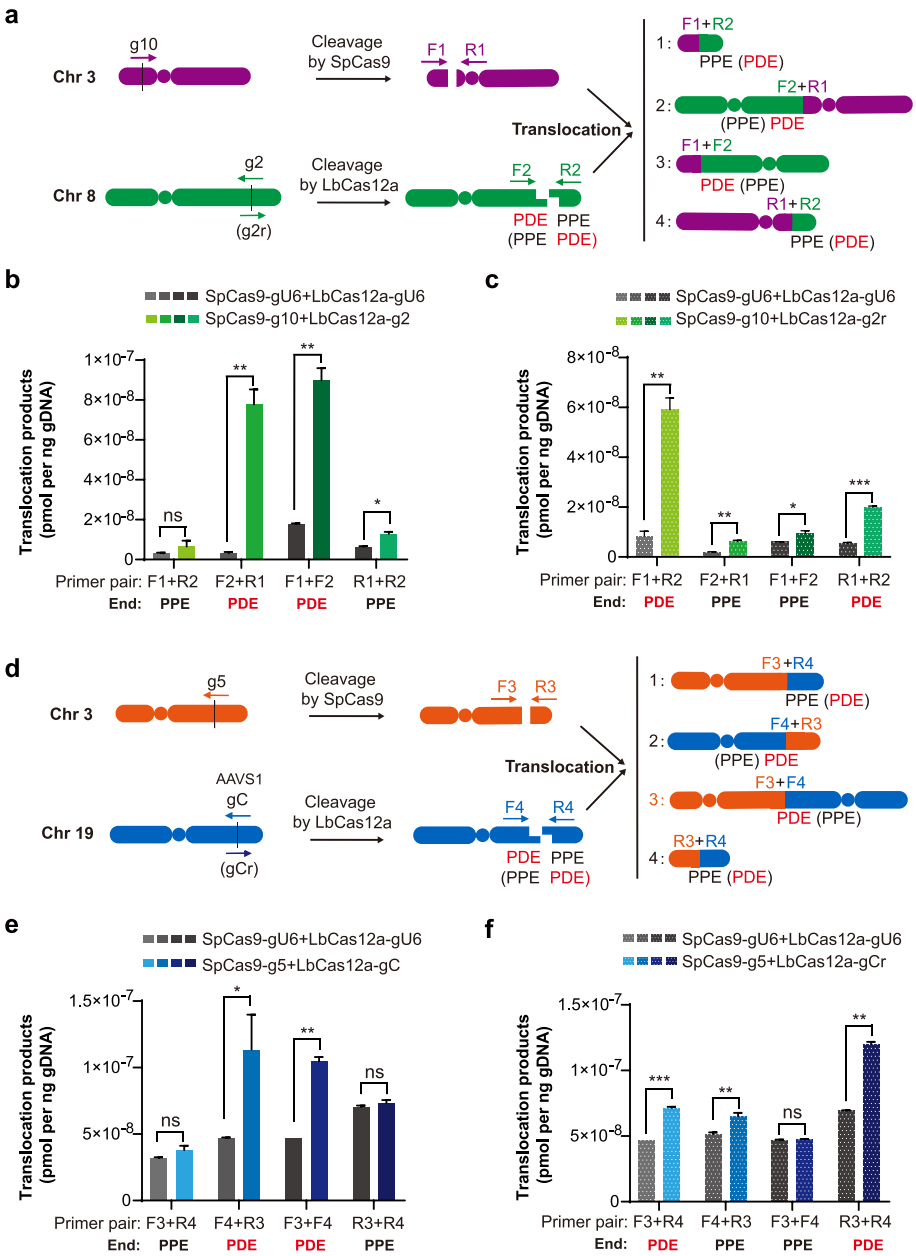

Fig. 5 (See legend on previous page.)

by PCR with primer pairs F1/F2, F2/R1, F1/R2, and R1/R2 in the absence of DSBs. However, in the presence of *Sp*Cas9- and *Lb*Cas12a-induced DSBs, we observed translocations, with more detected using F1/F2 and F2/R1 for the PDEs of *Lb*Cas12a-induced DSBs than with F1/R2 and R1/R2 for the PPEs of *Lb*Cas12a-induced DSBs (Additional file 1: Fig. S7a).

The combined intensity of the PCR bands in each lane of the agarose gel was measured using ImageJ, and the relative intensity to *GAPDH* was calculated as the ratio of the intensity of PCR bands for the *Sp*Cas9 target site, the *Lb*Cas12a target sites, or the translocation products, to the intensity of the internal *GAPDH* control (Additional file 1: Fig. S7b). The relative intensity of the PCR bands for the target sites was apparent, as

expected (Additional file 1: Fig. S7b). In the absence of DSBs, the relative intensity of the PCR bands for the translocations was negligible. In contrast, in the presence of *Sp*Cas9- and *Lb*Cas12a-induced DSBs, the relative intensities with F1/F2, F2/R1, F1/R2, and R1/R2 were strong, but stronger with F1/F2 and F2/R1 than with F1/R2 and R1/R2 (Additional file 1: Fig. S7b). Despite less quantitative, the results from this approach are consistent with those obtained by qRT-PCR described above.

Translocations induced by *Lb*Cas12a-gC and *Sp*Cas9-g5 were also similarly detected by PCR and confirmed by Sanger sequencing after subcloning into a sequencing vector (Additional file 1: Fig. S8a–f). Using primer pairs F3/R3 and F4/R4, we readily detected PCR products from target sites without translocations (Additional file 1: Fig. S8a,b). While no translocation was observed in the absence of DSBs induced by *Sp*Cas9 and *Lb*Cas12a, a significant level of translocation was detected by PCR when DSBs were induced by both *Sp*Cas9 and *Lb*Cas12a. Notably, translocations were more frequently identified with primer pairs F4/R3 and F3/F4 for the PDEs of *Lb*Cas12a-induced DSBs compared to F3/R4 and R3/R4 for the PPEs (Additional file 1: Fig. S8a,b).

Additionally, we used *Sp*Cas9-g5 and *Lb*Cas12a-g2 to induce a site-specific DSB at the long arm of chromosome 3 and 8 in human 293 T cells, respectively (Additional file 1: Fig. S9a). This allowed for translocation between either the PDEs or the PPEs of the *Lb*Cas12a-induced DSBs and the ends of *Sp*Cas9-induced DSBs. As expected, PCR products from the target sites using primer pairs F3/R3 and F2/R2 were significant, regardless of DSB induction by *Sp*Cas9 or *Lb*Cas12a (Additional file 1: Fig. S9b,c). We detected more translocations via PCR with primer pairs F2/F3 and F2/R3 compared to F3/R2 and R2/R3 in the presence of both *Sp*Cas9- and *Lb*Cas12a-induced DSBs, with no translocation observed in the absence of DSBs induced by *Sp*Cas9 and *Lb*Cas12a (Additional file 1: Fig. S9b,c). While no PCR products for translocation were detected with R3/R2, products were observed with primer pairs F3/R2, F2/F3, and F2/R3, which were confirmed by subcloning followed by Sanger sequencing (Additional file 1: Fig. S9 d–f). Collectively, these results suggest that ligation with the PDE of a DSB induced by *Lb*Cas12a is more efficient than with the PPE, and that PDEs of *Lb*Cas12-induced DSBs are more prone to translocation than PPEs.

PDEs of *Lb*Cas12a-induced DSBs promote accurate ligation of donor templates into targeted sites

Given that compatible cohesive PDEs of *Lb*Cas12a-induced DSBs favor efficient and accurate ligation by c-NHEJ in mammalian cells, we investigated whether these ends also enhance accurate ligation between a donor template and a target site during targeted integration. We selected targeted integration sites at the *AAVS1*, *CLTA*, and *HNRNPA1* loci in 293 T cells and designed a donor template containing a *pPGK-RFP-pA* cassette flanked by two artificial *Lb*Cas12a target sites with either W/W, W/C, C/W, or C/C-paired PAM orientations (Fig. 6a). A single *Lb*Cas12a-sgRNA was used to cleave the target site, generating two compatible ends with a 5-nt 5′ overhang—one PDE and one PPE (Fig. 6a). Cleavage of the two artificial *Lb*Cas12a target sites flanking the *pPGK-RFP-pA* cassette by paired *Lb*Cas12a-sgRNAs produced a linear donor template with either a PDE or a PPE on each side, complementary to the respective ends at the targeted integration site (Fig. 6a).

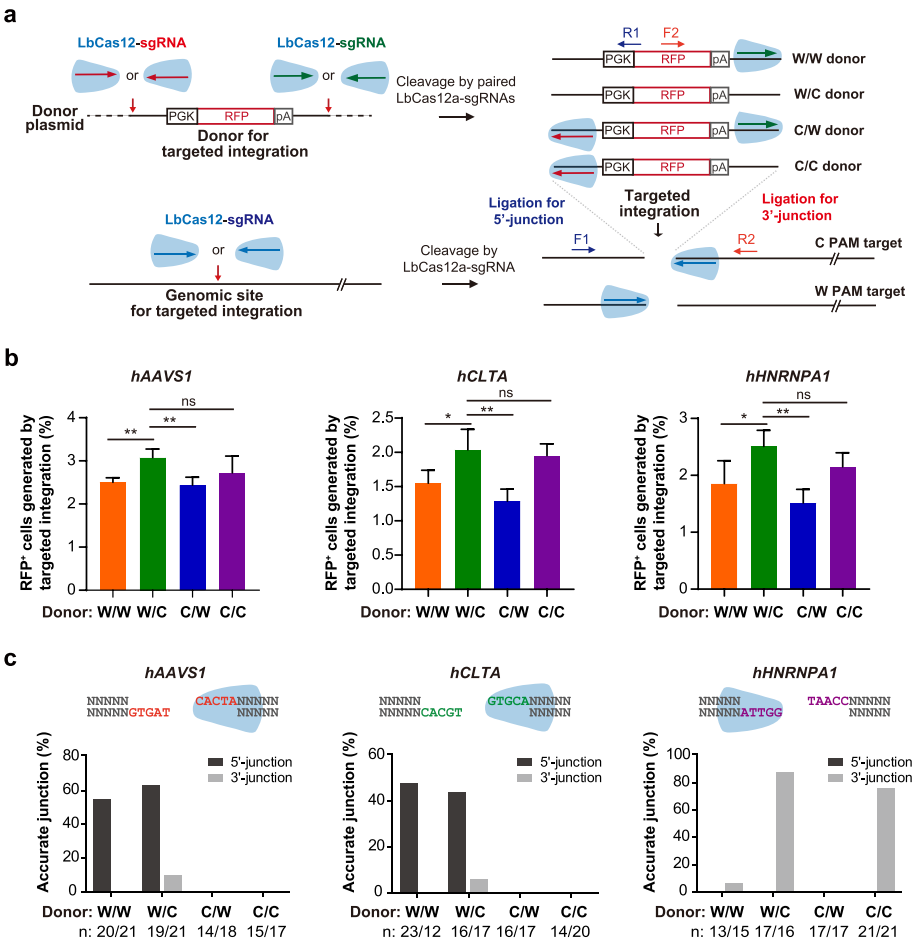

Fig. 6 The PDEs of *Lb*Cas12a-induced DSBs promote accurate ligation of donor templates into targeted sites. **a** Schematic representation of *Lb*Cas12a-induced NHEJ-based targeted integration. The targeted integration sites include both C PAM and W PAM targets. The donor plasmid contains a *pPGK-RFP-pA* cassette flanked by two artificial *Lb*Cas12a target sites with paired PAM orientations of W/W, W/C, C/W, or C/C. The 5′-staggered ends generated at the two artificial target sites upon *Lb*Cas12a-mediated cleavage are designed to be complementary to the 5′-staggered ends produced by *Lb*Cas12a at the targeted integration site in the genome. Primers F1 and R2 at the original targeted integration sites and R1 and F2 at the donors are indicated. **b** Frequency of RFP^+ cells generated by targeted integration using donor ends induced by paired *Lb*Cas12a with W/W, W/C, C/W, or C/C PAM orientations at the *AAVS1* (left), *CLTA* (middle), or *HNRNPA1* (right) loci in 293 T cells. The frequency of $RFP.^+$ cells was calculated by subtracting the frequency of random integration control (using the gU6 empty vector) from the targeted integration group with paired *Lb*Cas12a-sgRNAs. **c** Frequency of accurate 5′- or 3′-junctions in targeted integration using donor ends induced by paired *Lb*Cas12a with W/W, W/C, C/W, or C/C PAM orientations at the *AAVS1* (left), *CLTA* (middle), or *HNRNPA1* (right) loci in 293 T cells. PCR products of the targeted integration junctions were subcloned into a TA vector and analyzed by Sanger sequencing. The number of subclones analyzed for each junction is indicated by *n*. Columns with error bars represent the mean ± S.D. of three independent experiments in **b**, with significance determined by one-way ANOVA with Tukey's multiple comparison test. ns: $P > 0.05$; *: $P < 0.05$; **: $P < 0.01$

Targeted integration of *pPGK-RFP-pA* was assessed by flow cytometry for RFP^+ cells and PCR using primer pairs F1/R1 for the 5′-junction and F2/R2 for the 3′-junction, along with flow cytometry to measure the frequency of stable RFP^+ cells (Fig. 6a). While approximately 1% of RFP^+ cells could arise from random integration, we observed an increased frequency of RFP^+ cells in the presence of *Lb*Cas12a-sgRNAs

compared to random integration (Additional file 1: Fig. S10a–c). We derived the frequency of RFP^+ cells generated by targeted integration by subtracting the frequency of random integration from that induced by *Lb*Cas12a-sgRNAs. This frequency was generally highest with two PDEs from the W/C donor and lowest with two PPEs from the C/W donor across the three targeted integration sites (Fig. 6b).

PCR followed by agarose gel electrophoresis confirmed the occurrence of targeted integration (Additional file 1: Fig. S11a–c). Sanger sequencing of PCR products revealed that the 5′-junctions of targeted integration into the *AAVS1* and *CLTA* loci were mostly accurate (up to 70.0%) with the W/W and W/C donors, while the 3′-junctions into the *HNRNPA1* locus were accurate (up to 80.0%) with the W/C and C/C donors (Fig. 6c and Additional file 1: Fig. S11a–c). These findings indicate that ligation between compatible PDEs—one from the donor and one from the target site—tends to be more accurate. Although donors with all four PAM orientations generated significant levels of RFP^+ cells, the majority exhibited accurate 5′-junctions at the *AAVS1* and *CLTA* loci with the W/W and W/C donors and accurate 3′-junctions at the *HNRNPA1* locus with the W/C and C/C donors. In contrast, few or no RFP^+ cells contained accurate junctions at the targeted integration sites in other cases. These results suggest that compatible cohesive PDEs of *Lb*Cas12a-induced DSBs, rather than PPEs, promote accurate ligation of donor templates into targeted sites.

Aside from oriented integration (targeted integration in the correct orientation), we also considered the possibilities of reverse integration (integration in the reverse orientation) and duplicated integration (integration of two donor copies) (Additional file 1: Fig. S12a). To analyze these occurrences, we sorted RFP^+ cells generated by targeted integration using four types of donors (W/W, W/C, C/W, and C/C) at the *AAVS1* locus and examined the 5′- and 3′-junctions of single RFP^+ clones by PCR followed by Sanger sequencing (Additional file 1: Fig. S12a). Oriented integration was assessed using primer pairs F1/R2 and F2/R1, reverse integration with F1/F2 and R2/R1, and duplicated integration with F2/R2 (Additional file 1: Fig. S12a).

Among 24 RFP^+ single clones for each donor type, we found 19 oriented and 5 reverse integrations for the W/W donor, 14 oriented and 10 reverse for the W/C donor, 12 oriented and 12 reverse for the C/W donor, and 2 oriented and 22 reverse for the C/C donor (Additional file 1: Fig. S12b–f). Oriented integration with the W/W donor and reverse integration with the C/C donor occurred much more frequently than reverse integration with the W/W donor and oriented integration with the C/C donor, respectively (Additional file 1: Fig. S12b,e,f). Both oriented integration with the W/W donor and reverse integration with the C/C donor were mediated by ligation between two PDEs—one from the cleaved donor and the other from the cleaved target site—at the 5′ junction. Conversely, reverse integration with the W/W donor and oriented integration with the C/C donor were generated by ligation between one PDE and one PPE at both 5′ and 3′ junctions. This supports the notion that ligation between two PDEs of Cas12a-induced DSBs is more efficient than ligation involving at least one PPE, suggesting that more efficient ligation between two PDEs may help direct the orientation of end ligation between the donor and target site. Additionally, we found rare duplicated integration events: 1 with the W/W donor, 2 with the W/C

donor, 1 with the C/W donor, and 2 with the C/C donor among 24 RFP^+ single clones for each donor type (Additional file 1: Fig. S12b–e).

Among the oriented integrations directed by two compatible ends, only the 5′ junctions with the W/W and W/C donors were generated by ligation of two PDEs. The other 5′ junctions and all 3′ junctions were associated with at least one PPE. Accurate 5′ junctions were found in 13 out of 19 oriented integration events with the W/W donor and 11 out of 14 with the W/C donor (Additional file 1: Fig. S12 g). This further confirms that ligation between two compatible PDEs—one from the cleaved donor and the other from the cleaved target site—tends to be more accurate (Fig. 6c and Additional file 1: Fig. S11a–c).

Precise gene fragment replacement by ligation with PDEs of *Lb*Cas12a-induced DSBs

Since the compatible cohesive PAM-distal ends (PDEs) of *Lb*Cas12a-induced DSBs promote accurate ligation between a donor template and its target site, we investigated whether this could be harnessed to efficiently and precisely replace a gene fragment in mammalian cells. We focused on the *CANX*, *PCNA*, and *H4 C16* genes, aiming to add a C-terminal RFP tag in 293 T cells by replacing specific gene fragments using compatible PDEs of *Lb*Cas12a (Fig. 7a). We identified W/C-paired *Lb*Cas12a target sites flanking the normal stop codon in the last exons of these genes—exon 21 of *CANX*, exon 7 of *PCNA*, and exon 1 of *H4 C16*. Expression of paired *Lb*Cas12a-gRNAs would delete the segment containing the stop codon, generating two PDEs for ligation with an exogenous DNA fragment (Fig. 7a). We designed a targeted integration donor cassette flanked by two artificial *Lb*Cas12a target sites with W/C-paired PAM orientation (Fig. 7a). This cassette included a linker-RFP sequence surrounded by the deleted gene fragment, which also lacked the original stop codon (Fig. 7a). Cleavage of both target sites and the donor plasmid generated compatible cohesive PDEs between the cleaved target sites and the

(See figure on next page.)

Fig. 7 The PDEs of *Lb*Cas12a-induced DSBs are utilized to precisely replace gene fragments for C-terminal tagging via NHEJ-based targeted integration. **a** Schematic representation of paired *Lb*Cas12a-induced NHEJ-based replacement of a gene fragment. The targeted integration sites at the *CANX*, *PCNA*, and *H4 C16* loci in 293 T cells are targeted by *Lb*Cas12a with C/W-paired PAM orientation, generating two PDEs for targeted integration. The donor plasmid contains an *exon-linker-RFP* replacement cassette flanked by two artificial *Lb*Cas12a target sites with W/C-paired PAM orientation. The 5′-staggered ends generated at the artificial target sites upon *Lb*Cas12a cleavage are designed to be complementary to the 5′-staggered ends produced by paired *Lb*Cas12a at the targeted integration site in the genome. Exon 21 of *CANX*, exon 7 of *PCNA*, and exon 1 of *H4 C16* were partially replaced with *exon-linker-RFP* without losing any sequence from *CANX*, *PCNA*, or *H4 C16*, respectively. Primers for the original targeted integration sites and donors are indicated. **b** Frequency of accurate 5′- or 3′-junctions in gene replacement at the targeted integration sites of *CANX*, *PCNA*, or *H4 C16* in 293 T cells. PCR products of the targeted integration junctions were subcloned into a TA vector and analyzed by Sanger sequencing. The number of subclones analyzed for each junction is indicated by *n*. **c** Frequency of RFP^+ cells at 7 and 10 days post-transfection for intended gene replacement at *CANX*, *PCNA*, or *H4 C16* in 293 T cells. Paired *Lb*Cas12a with gC/gW for cleavage of the targeted integration sites is indicated, along with gU6 as a negative gRNA control. Immunofluorescence (**d**) and western blot (**e**) analysis of sorted RFP^+ cells from **c**. $RFP.^+$ cells were sorted 10 days post-transfection for intended gene replacement at *CANX*, *PCNA*, or *H4 C16*. The cells were then fixed for fluorescence detection of RFP-fusion proteins (**d**) or lysed for western blotting of CANX, PCNA, and H4 C16 proteins, with β-actin or GAPDH as loading controls (**e**). Ctrl represents targeted integration-negative cells generated by the negative control *Lb*Cas12a-gU6, which does not cleave the targeted integration sites. Columns with error bars represent the mean ± S.D. of three independent experiments in **c**, with statistical analysis performed using Student's two-tailed paired *t*-test. *: $P < 0.05$; **: $P < 0.01$; ***: $P < 0.001$

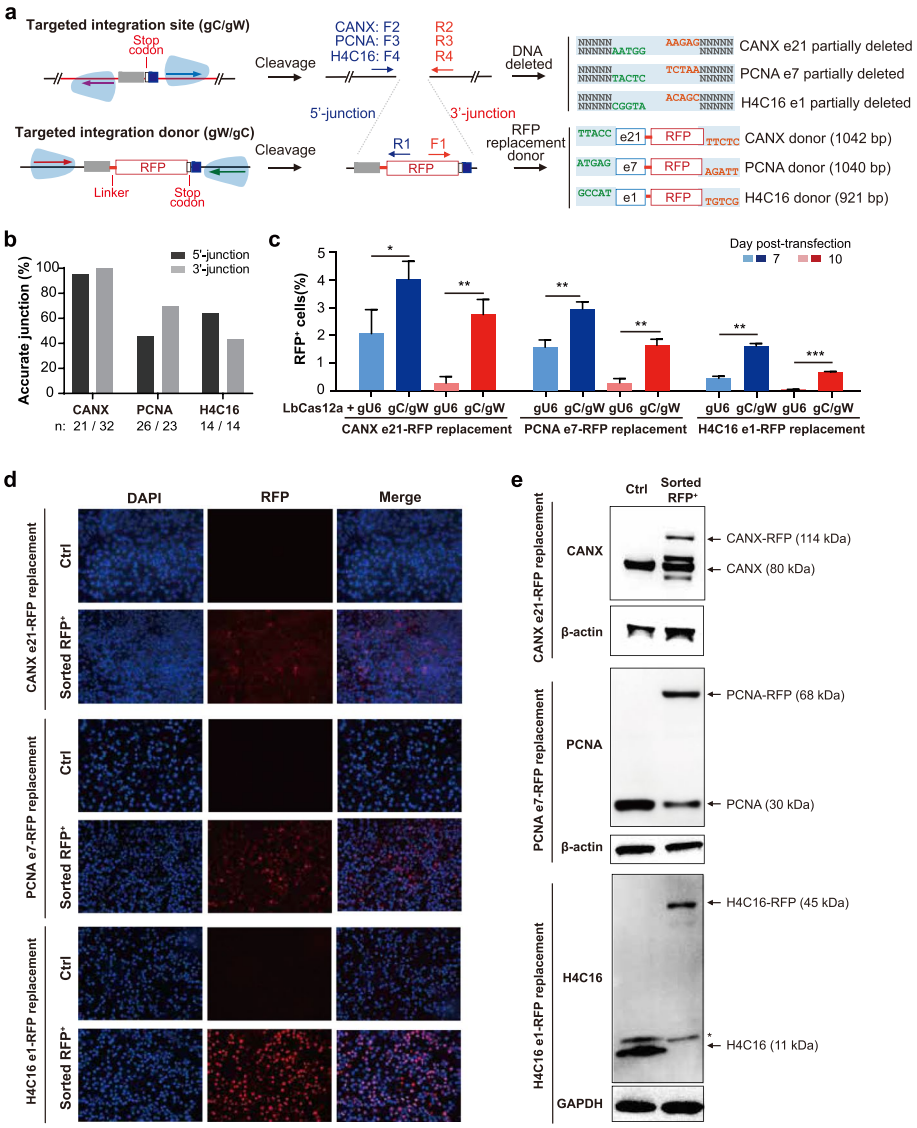

Fig. 7 (See legend on previous page.)

donor, enabling the ligation of these ends and resulting in the replacement of the gene fragment along with the addition of the C-terminal tag (Fig. 7a).

Targeted integration products were assessed by PCR using specific primer pairs for the 5′ and 3′ junctions. Flow cytometry measured the frequency of stable RFP^+ cells (Fig. 7a). PCR followed by agarose gel electrophoresis confirmed successful C-terminal RFP tagging of the three genes (Additional file 1: Fig. S13a–c). Sanger sequencing of the PCR products showed that both junctions of the C-terminal tagging were mostly accurate, achieving up to 100% accuracy (Fig. 7b and Additional file 1: Fig. S13 d–f). The frequency of background RFP^+ cells remained high until 7 days (d) post-transfection but decreased to negligible levels by 10 d (Fig. 7c and Additional file 1: Fig. S13 g). Transfection with expression plasmids for paired *Lb*Cas12a-sgRNAs resulted in significantly higher frequencies of RFP^+ cells: approximately 4.0% for *CANX*, 2.9% for *PCNA*, and 1.6% for *H4 C16* at 7 d post-transfection, stabilizing at

about 2.7%, 1.6%, and 0.67%, respectively, by 10 d (Fig. 7c). After sorting RFP^+ cells at 10 d post-transfection, we examined RFP localization and found that RFP was cytoplasmic for the cytoplasmic protein CANX and nuclear for the nuclear proteins PCNA and H4 C16 (Fig. 7d). Western blot analysis of protein extracts from sorted RFP^+ cells at 10 d post-transfection confirmed the successful C-terminal RFP tagging of CANX, PCNA, and H4 C16 at the predicted sizes of the fusion proteins (Fig. 7e). As expected, C-terminal RFP tagging for CANX or PCNA was less efficient in the presence of DNA-PKcs inhibitors or when Ku80 or XRCC4 were depleted (Additional file 1: Fig. S14a–c). These results indicate that *Lb*Cas12a-induced DSBs can be effectively harnessed to accurately replace a gene fragment via c-NHEJ, a strategy we term Cas12a-induced PDE ligation (CIPDEL) for NHEJ-mediated targeted integration.

To compare the efficiency of CIPDEL with HDR-mediated targeted integration induced by *Lb*Cas12a and *Sp*Cas9 and with NHEJ-mediated targeted integration induced by *Sp*Cas9 (i.e., HITI) [15, 18], we performed experiments in 293 T cells and hTERT-immortalized RPE- 1 cells. While both cells displayed the preference for NHEJ over HDR in repair of *Lb*Cas12a-induced DSBs, the NHEJ proportion in RPE- 1 cells was greater than that in 293 T cells, likely suggesting a stronger NHEJ activity in RPE-1 cells (Additional file 1: Fig. S15a–c). Consistently, following targeted integration with both the W/C donor and an HDR donor containing a *pPGK-RFP-pA* cassette at the *AAVS1* locus (Additional file 1: Fig. S15 d), we found that CIPDEL appeared less efficient than HDR in 293 T cells but more efficient in RPE- 1 cells (Additional file 1: Fig. S15e,f).

In CANX C-terminal tagging experiments, where the last exon of *CANX* was replaced with an *e21-RFP* tag via either CIPDEL, *Lb*Cas12a-induced HDR, HITI, or *Sp*Cas9-induced HDR (Additional file 1: Fig. S16a), we selected *Lb*Cas12a-gC, *Lb*Cas12a-gW, *Sp*Cas9-gC, and *Sp*Cas9-gW, all of which cleaved their targets in a similar efficiency, for targeted integration (Additional file 1: Fig. S16b). FACS analysis indicated that CIPDEL and *Lb*Cas12a-induced HDR were more efficient than HITI and *Sp*Cas9-induced HDR for targeted integration in both 293 T cells and RPE- 1 cells, respectively (Additional file 1: Fig. S16c). NHEJ-mediated targeted integration, such as CIPDEL and HITI, was much more active than HDR in RPE- 1 cells, as compared to 293 T cells (Additional file 1: Fig. S16c). Additionally, CIPDEL resulted in a higher level of accurate 5′- and 3′-junctions than HITI (Additional file 1: Fig. S16 d).

We also analyzed off-target effect of *Lb*Cas12a-gC, *Lb*Cas12a-gW, *Sp*Cas9-gC, and *Sp*Cas9-gW. Total 1, 3, 201, and 157 off-target sites were predicted for *Lb*Cas12a-gC, *Lb*Cas12a-gW, *Sp*Cas9-gC, and *Sp*Cas9-gW, respectively. Indel frequencies induced by *Lb*Cas12a-gC and *Lb*Cas12a-gW at their respective off-target sites in 293 T cells were much smaller than those induced by *Sp*Cas9-gC and *Sp*Cas9-gW at their top 3 off-target sites (Additional file 1: Fig. S16e). This is consistent with previous findings that *Lb*Cas12a has a considerably lower off-target effect than *Sp*Cas9 [6–8]. Collectively, these results suggest that CIPDEL may offer advantages in targeted integration and be particularly suited for targeted integration in certain cell types, including slow-cycling and non-dividing cells where HDR activity is lower or inactive. A step-by-step protocol for CIPDEL-mediated targeted integration is presented in Additional file 1: Fig. S17.

Discussion

In this work, we first examined the junctions of NHEJ repair for *Lb*Cas12a-induced DSBs at 16 sites and observed that the average deletion length was longer from the PPEs than from the PDEs. Further investigation revealed that the post-cleavage retention of Cas12a-sgRNA at the PPEs—rather than the PDEs, which are free from Cas12a-sgRNA—prevents access by the c-NHEJ core factors, leading to asymmetric deletions in the NHEJ repair of *Lb*Cas12a-induced DSBs. More importantly, the easier access to c-NHEJ core factors at the PDEs promotes more efficient and accurate end ligation compared to the PPEs. With this mechanistic insight, we devised the CIPDEL approach to harness the free 5-nt 5′-overhanging PDEs of *Lb*Cas12a-induced DSBs for NHEJ-mediated targeted integration with improved efficiency and accuracy.

A well-documented yet unique feature of CRISPR-Cas systems is that both CRISPR-Cas9 and CRISPR-Cas12a remain bound to their targets for hours, even after cleavage [8, 20–24]. This post-cleavage target residency may influence the exposure of CRISPR-induced DNA breaks to specific stages of the cell cycle and certain DNA damage response (DDR) proteins, which determine the choice of DNA break repair pathways. Indeed, previous studies have indicated that upon target cleavage, *Sp*Cas9-sgRNA is released from the cleaved DNA either spontaneously or through forces from transcription, DNA replication, or chromatin remodeling, in a manner partly dependent on the binding affinity of *Sp*Cas9-sgRNA to its target [25, 33–37]. This can elicit different DDR responses and, by extension, result in heterogeneity in genome editing products [19, 25].

Unlike the *Sp*Cas9-gRNA complex, which binds to both ends of SpCas9-induced DSBs, Cas12a-gRNA is retained at the PPE, leaving the PDE free from the post-cleavage Cas12a-gRNA-DNA ternary complex [8, 22, 23]. The impact of this retention on repair activity at the PPE was unclear. However, it is expected that free DSB ends, such as the PDE, readily bind to the KU70/KU80 dimer, one of the most abundant nuclear proteins in mammalian cells [38]. KU70/KU80 not only protects DSB ends from nucleolytic attack but also recruits other c-NHEJ core factors, such as DNA-PKcs and XRCC4/DNA ligase 4, facilitating NHEJ repair [39]. Despite the apparent asymmetry in target residency between the PPEs and PDEs of Cas12-induced DSBs, previous profiling of CRISPR-Cas12a genome editing did not examine potential repair differences between these ends, focusing instead on overhang polarity's role in generating various mutation signatures [6, 7]. For instance, due to distinct overhangs, *Lb*Cas12a and *As*Cas12a rarely induce insertions, while *Sp*Cas9 frequently generates 1-nt insertions [6, 7]. Prior analyses of *As*Cas12a-induced targeted integration suggested a higher mutagenic tendency of the PPEs than the PDEs in NHEJ, attributing this to re-cleavage by *As*Cas12a [16]. Not until our current study was the asymmetric effect of post-cleavage target residency on the repair of Cas12a-induced DSBs identified.

CRISPR-Cas genome editing strategies primarily include the deletion or disruption of DNA fragments, targeted integration for precise gene correction and DNA insertion, as well as base editing and prime editing—both of which do not require the induction of DSBs [40]. Targeted integration for precise gene correction and DNA insertion has demonstrated therapeutic potential in treating monogenic diseases and has also been used to engineer immune cells for cancer therapy [41]. While HDR is currently the primary pathway for targeted integration, HDR-based strategies are often inefficient and

not applicable in non-dividing cells, where HDR is inactive [9, 10, 42, 43]. The inclusion of homology arms needed for HDR increases the size of the targeting vector, which may limit both the efficiency of targeted integration and the size of the DNA insert. Particularly, when the donor template is delivered via AAV, the size of the DNA insert can become a limiting factor for packaging into conventional AAV vectors [44–46]. Additionally, *Sp*Cas9 may induce p53-dependent DDR in primary cells and human stem cells, arresting these cells in the G1 phase and preventing efficient HDR-based targeted integration [47, 48]. Consequently, NHEJ-based targeted integration, such as *Sp*Cas9-induced HITI, has been developed to effectively insert a DNA sequence of interest into targeted genomic loci, potentially addressing some of the challenges encountered with HDR-based strategies [15, 49, 50].

Due to the blunt ends of *Sp*Cas9-induced DSBs, inverted integration of donor DNA could theoretically occur in up to half of NHEJ-mediated targeted integration events in *Sp*Cas9-induced HITI [15, 18]. Additional design considerations are necessary to prevent this inverted integration. Moreover, the lower accuracy of NHEJ repair involving the PPEs limits the applicability of *Sp*Cas9-induced HITI for precise gene correction and insertion [14]. Furthermore, *Sp*Cas9 may induce significant off-target effects, necessitating thorough investigation and validation for safety prior to clinical use [51]. In contrast, CRISPR-Cas12a offers several inherent advantages, including a smaller size, minimal off-target cutting, staggered ends with 5-nt 5′-overhangs from Cas12a-induced DSBs, and the ability to process its own crRNA array. The CIPDEL approach leverages these benefits, providing a potentially safer and more precise alternative to *Sp*Cas9-based HITI for targeted integration.

Given that the repair of Cas12a-induced DSBs is influenced by the asymmetry in post-cleavage target residency of Cas12a-sgRNA, a mechanistic understanding of this regulation could provide further opportunities to enhance Cas12-induced genome editing. However, several obstacles must be addressed to broaden the application of the CIPDEL strategy and maximize its therapeutic potential. Neither *Lb*Cas12a, *As*Cas12a, nor *Fn*Cas12a can cleave DNA as efficiently as *Sp*Cas9 in general settings, although they may perform comparably at specific sites [1, 3, 4, 6, 7, 52]. Additionally, the T-rich PAM restricts the target range for *As*Cas12a, *Lb*Cas12a, and *Fn*Cas12a, likely limiting target choices for CIPDEL and highlighting the need for further expansion of PAM specificities [3, 53–56]. In the CIPDEL approach, reverse integration and duplicated integration could still occur, similar to what is seen in *Sp*Cas9-mediated HITI [15], necessitating efforts to minimize these occurrences.

While the CIPDEL strategy is designed to save space by eliminating the need for two homology arms in targeted integration, we have yet to determine the maximum size of DNA fragments that can be efficiently knocked into a specific genomic site using this strategy. It also remains unclear whether CIPDEL-based multiplexed targeted integration can be achieved with a compact crRNA array, which is commonly used to guide Cas12a-based multiplexed genome editing [4, 57, 58]. Particularly, the delivery of single or multiple donor templates and multiplexed Cas12a-gRNAs via AAV for CIPDEL should be examined to apply this strategy for targeted integration in vivo. Importantly, since HDR is inactive and NHEJ is the dominant pathway for DSB repair in non-dividing cells such as neurons and cardiomyocytes [15, 43], CIPDEL has a distinct advantage

over the classical HDR approach for directing targeted integration in these cell types, although this advantage has yet to be confirmed. A systematic head-to-head comparison with *Sp*Cas9-mediated HITI is also necessary to evaluate whether CIPDEL offers superior efficiency, accuracy, and reduced off-target effects in targeted integration.

Conclusions

Collectively, this study revealed that the asymmetric retention of CRISPR-*Lb*Cas12a at DSB ends suppresses c-NHEJ at the PPEs of *Lb*Cas12a-induced DSBs, not at the PDEs. This unique repair mechanism has been harnessed for CIPDEL-directed targeted integration. However, practical applications of CIPDEL in genome editing still require further optimization of this NHEJ-mediated targeted integration strategy.

Methods

Plasmids

The expression plasmids for *Sp*Cas9 (Cat# 42,230) and *Lb*Cas12a (Cat# 69,988) were originally obtained from Addgene. The expression plasmids for gRNAs were constructed as described previously [59, 60]. The donor plasmids for NHEJ-based targeted integration of the *pPGK-RFP-pA* cassette at the *AAVS1*, *CLTA*, or *HNRNPA1* locus of human 293 T cells were generated by placing a *Lb*Cas12a-gRNA target site to either side of *pPGK-RFP-pA*. The donor plasmids for HDR-based targeted integration of the *pPGK-RFP-pA* cassette at the *AAVS1* locus of human 293 T cells were generated by a 5′- and 3′-homology arm, each 500 bp, to either side of *pPGK-RFP-pA*. The donor plasmids for NHEJ-based targeted integration of *exon-(GGGGS)3-linker-RFP* to replace the respective exon sequences of *CANX*, *PCNA*, and *H4 C16* in human 293 T cells were constructed by placing a *Lb*Cas12a-gRNA target site to either side of *exon-(GGGGS)3-linker-RFP*. The donor plasmids for HDR-based targeted integration of *exon-(GGGGS)3-linker-RFP* to replace the respective exon sequences of *CANX* in human 293 T cells were constructed by placing a 5′- and 3′-homology arm, each 500 bp, to either side of *exon-(GGGGS)3-linker-RFP*. The donor sequences for targeted integration along with the sequences of targeted integration sites are listed in Additional file 2: Table S1. Newly constructed plasmids were confirmed by Sanger sequencing.

Cell lines and cell culture

Human embryonic kidney cells 293 and 293 T and hTERT-immortalized retinal pigment epithelial cells RPE- 1 were obtained from the American Type Culture Collection (ATCC) and cultured in Dulbecco's modified Eagle's medium (DMEM) containing 10% fetal bovine serum, 1% penicillin–streptomycin, and 2 mM L-glutamine. Mouse ES cells containing the sGEJ reporter were previously established and were cultured in DMEM medium (Biological Industries) supplied with 20% fetal bovine serum (Gibco), 1% penicillin–streptomycin (Gibco), 2 mM L-glutamine (Gibco), 0.1 mM β-mercaptoethanol (Sigma), 0.1 mM non-essential amino acid (Gibco), 1 mM sodium pyruvate (Gibco), and 1000 U/mL leukemia inhibitory factor (Millipore) on either mouse embryonic fibroblast (MEF) feeders or gelatinized plates [25]. Isogenic wild-type, *DNA-PKcs*$^{-/-}$ and *Ku80*$^{-/-}$ mouse ES cells clones and isogenic *XRCC4*$^{+/+}$ and *XRCC4*$^{-/-}$ mouse ES cells

were generated previously [25]. All the cell lines for the study were tested negative for mycoplasma.

Transfection

Transfection of mouse ES cells were performed with Lipofectamine 2000 (Invitrogen) in 24-well plates as previously described [61, 62]. Briefly, total 1.0×10^5 mouse ES cells in 200 μL culture medium were transfected with 0.5 μg expression plasmids for *Sp*Cas9-gRNA or *Lb*Cas12a-gRNA. For 293 cells and 293 T cells, 1.4×10^5 cells were transfected with the expression plasmids (0.5 μg) for *Sp*Cas9-gRNA or *Lb*Cas12a-gRNA by Hieff Trans (Yeasen, #40805ES). The plasmid ratio of *Sp*Cas9 or *Lb*Cas12a to gRNA in transfection amount is 1:1. Transfection efficiencies were determined by parallel transfection of pcDNA3β-GFP (0.1 μg) together with the empty vector pcDNA3β (0.4 μg). For RPE- 1 cells, 1.4×10^5 cells were transfected with 0.25 μg expression plasmids for *Lb*Cas12a, 0.125 μg expression plasmids for gRNAs, and 0.125 μg NHEJ donor plasmids or HDR donor plasmids by Lipofectamine® LTX & PLUS™ Reagent (Invitrogen). Transfection efficiencies were measured by parallel transfection of 0.25 μg pcDNA3β-GFP together with 0.25 μg the empty vector pcDNA3β. Experiments with transfection efficiencies below 50% were excluded. Transfection efficiencies were determined at 72 h post-transfection by fluorescence-activated cell sorting (FACS) using the Beckman Coulter CytoFLEX flow cytometer. Cells were analyzed at 72 h post-transfection for genome editing by targeted PCR amplification and Illumina deep sequencing. The gRNA target sequences along with primers for targeted PCR amplification are listed in Additional file 3: Table S2 and Additional file 4: Table S3. FACS data were analyzed using the CytExpert 2.0 software.

T7 endonuclease I (T7E1) assay

After cells were transfected with plasmids expressing *Lb*Cas12a-gRNAs or *Sp*Cas9-gRNAs, along with gU6 as a negative gRNA control, genomic DNA (gDNA) was isolated from cells at 48 h or 72 h post-transfection using a gDNA purification kit (Vazyme). PCR amplification of targeted sites was performed using 100 ng gDNA with primers listed in Additional file 5: Table S4. PCR products were purified with PCR Clean-up kit (Vazyme) and 200 ng of purified PCR products was mixed with 1 μL 10 × T7E1 buffer (Vazyme, EN303) and ultrapure water to a final volume of 10 μL, and subjected to a re-annealing process to enable heteroduplex formation: 95 °C for 3 min, 95 °C for 30 s, 90 °C for 30 s, 85 °C for 30 s, 80 °C for 30 s, 75 °C for 30 s, 70 °C for 30 s, 65 °C for 30 s, 60 °C for 30 s, 55 °C for 30 s, 50 °C for 30 s, 45 °C for 30 s, 40 °C for 30 s, 35 °C for 30 s, 30 °C for 30 s, and 25 °C for 1 min. After re-annealing, products were treated with T7 endonuclease I (Vazyme, EN303) for 15 min at 37 °C. The reaction mixtures were separated by 2% agarose gel electrophoresis and the indel ratios were calculated according to the band intensities by ImageJ software using the following equation: indel ratios = the intensities of cut bands/(the intensities of cut bands + the intensity of uncut band).

Detection of the HDR and NHEJ activity in 293 T and RPE- 1 cells

The ssODN donor containing a barcode ("GGATCC") in the middle and 50-bp homologous arms at 5′ and 3′ ends was synthesized by TsingKe Biological Technology as the

HDR template (Additional file 2: Table S1). 293 T cells and RPE- 1 cells were co-transfected with 0.25 μg ssODNs and the expression plasmids for *Lb*Cas12a (0.15 μg)-gC (0.1 μg). At 72 h post-transfection, gDNA was isolated from these cells using a gDNA purification kit (Vazyme). The targeted regions of less than 300 bp in gDNA were amplified by PCR with primers listed in Additional file 2: Table S1 and Illumina deep sequencing was performed by Novogene Co. Ltd (Beijing). Sequences were analyzed to determine the insertion frequency of the barcode at the cleavage site for the HDR efficiency and the indel frequency at the cleavage site for the NHEJ efficiency. Both frequencies were normalized with transfection efficiencies.

Targeted PCR amplification and Illumina deep sequencing

The edited sequences by targeted genome editing at endogenous genome loci were analyzed by targeted PCR amplification and Illumina deep sequencing as described previously [59]. Briefly, cells transfected with *Sp*Cas9-gRNA and *Lb*Cas12a-sgRNA for targeted genome editing were harvested at 72 h post-transfection and gDNA was isolated from these cells using a gDNA purification kit (Vazyme). The targeted regions of less than 300 bp in gDNA were amplified by PCR with primers listed in Additional file 3: Table S2 for single *Sp*Cas9-gRNA or single *Lb*Cas12a-gRNA and with primers listed in Additional file 4: Table S3 for paired *Lb*Cas12a-gRNAs. Illumina deep sequencing was performed at Novogene Co. Ltd (Beijing). PCR products were purified with PCR Clean-up kit (Vazyme), end-repaired, adenylated at 3′ ends, ligated with adapters, purified, and amplified by the second round of PCR to incorporate the P7 and P5 Illumina adapters according to the manufacturer's protocols (Yeasen, Hieff NGS Ultima DNA Library Prep Kit for Illumina). Sequences were analyzed to determine the frequency of genome editing and identify the indel pattern at repair junctions of edited sites using DBS-Aligner as described previously [59]. The frequency of editing was calculated and corrected with background readings and normalized with transfection efficiencies.

Targeted integration assay

For NHEJ- and HDR-based targeted integration, 1.4×10^5 293 T cells were transfected with 0.4 μg of the donor plasmids in 0.8 μg of total DNA with expression plasmids for *Lb*Cas12a (0.2 μg) or *Sp*Cas9 (0.2 μg)-gRNAs (0.2 μg), and 1.4×10^5 RPE- 1 cells were transfected with 0.125 μg of the donor plasmids in 0.5 μg of total DNA with expression plasmids for *Lb*Cas12a (0.25 μg) or *Sp*Cas9 (0.25 μg)-gRNAs (0.125 μg). 293 T cells were treated with the DNA-PKcs inhibitor NU7441 (2.0 μM; TopScience Cat# T6276) or M3814 (2.0 μM; Selleck Cat# S8586) at 6 h post-transfection for 72 h as needed. For siRNA-mediated depletion assays, 293 T cells were transfected with 20 pmol mixture of three siRNAs for each gene (TSINGKE Co.) together with the expression plasmids for *Lb*Cas9-sgRNAs and the donor plasmids. SiRNAs sequences were listed in Additional file 3: Table S2. RNA depletion was determined by quantitative reverse transcriptase PCR on qPCR CFX 96 Thermocycler (Bio-Rad). Parallel transfection of 0.16 μg pcDNA3β-GFP expression vector together with 0.64 μg pcDNA3β (the empty vector) was performed to determine transfection efficiencies, which were measured at 72 h post-transfection by FACS using Beckmann Coulter CytoFLEX. The frequencies of RFP^+ cells induced by targeted integration were measured at 7 d and 10 d post-transfection by

FACS using Beckmann Coulter CytoFLEX and corrected with background readings and transfection efficiencies. FACS data were analyzed using the CytExpert 2.0 software.

To analyze the junction of NHEJ-based targeted integration in a correct orientation in unsorted cells, gDNA was isolated from unsorted cells using a gDNA purification kit (Vazyme) at 10 d post-transfection. Junction of targeted integration in a correct orientation was amplified by PCR with primers listed in Additional file 2: Table S1 and determined by subcloning PCR products into a TA cloning vector (TSINGKE TSV- 007S pClone007 Simple Vector Kit) followed by Sanger sequencing of individual PCR products subcloned.

To analyze the junction of NHEJ-based targeted integration in individual RFP^+ cells, RFP^+ cells were sorted by FACS using Beckman Moflo Astrios EQ at 10 d post-transfection, plated and grown. Single RFP^+ clones were picked and gDNA was isolated from these individual clones using a gDNA purification kit (Vazyme). 5′ and 3′ junction of targeted integration for each clone was amplified by PCR with primers listed in Additional file 2: Table S1 and determined by Sanger sequencing.

Analysis of oriented integration, reverse integration, and duplication integration

To determine the frequencies of oriented integration, reverse integration, and duplication integration, the *pPGK-RFP-pA* cassette as the donor plasmids and the expression plasmids for *Lb*Cas12a-gC targeting the *AAVS1* locus in the human genome were transfected into 293 T cells, and RFP^+ cells were sorted by FACS using Beckman Moflo Astrios EQ at 10 d post-transfection, plated and grown. Single RFP^+ clones were picked and gDNA was isolated from these individual clones using a gDNA purification kit (Vazyme). RFP^+ clones with oriented integration, reverse integration, and duplication integration were identified by amplification of 5′- and 3′-junction for oriented integration, reverse integration, and duplication integration by PCR with primers listed in Additional file 2: Table S1 and determined by Sanger sequencing. The frequencies of oriented integration, reverse integration, and duplication integration were derived by the percentages of clones with oriented integration, reverse integration, and duplication integration.

Chromatin immunoprecipitation (ChIP)

ChIP assay was performed using SimpleChIP® Plus Enzymatic Chromatin IP Kit (#9003, CST) following the manufacturer's instructions as described previously [60]. Briefly, 2×10^5 mouse ES cells containing single-copy DSB reporter, either the SCR-RFP reporter or the NHEJ reporter [32, 63, 64], were transfected with expression plasmids for HA-tagged *Lb*Cas12a and respective gRNAs in each well of a 24-well plate. The gRNA target sequences are listed in Additional file 5: Table S4. Total 1×10^7 cells were collected at 12 h, 24 h, and 48 h post-transfection, fixed with 1% formaldehyde at 37 °C for 10 min and quenched with 0.125 M glycine for 5 min at room temperature. Nuclei were separated from cell pellets, digested with 0.5 μL micrococcal nuclease and sonicated to generate chromatin supernatant. Chromatin fractions containing Ku80 and HA-tagged *Lb*Cas12a were precipitated with antibodies against Ku80 (#2753, CST) and HA tag (sc-7392, Santa Cruz), as well as nonspecific IgG antibody as a negative control. After chromatin immunoprecipitates were de-crosslinked, DNA were purified with spin columns

from the ChIP Kit (#9003, CST). Enrichment of DNA was detected by qRT-PCR on CFX 96 Thermocycler (Bio-Rad) using the primer pairs listed in Additional file 5: Table S4. The fold enrichment of Ku80 and HA-*Lb*Cas12a at each genomic position relative to the negative IgG background was determined using the following equation: Fold enrichment $= 2^{-(\text{Ct genomic fragment with antibody of interest} - \text{Ct genomic fragment with IgG})}$.

Induction and analysis of translocation events

Translocation was similarly induced and analyzed as described previously [65]. Briefly, cells were transfected with expression plasmids for *Sp*Cas9-sgRNA and *Lb*Cas12a-sgRNA to induce translocation of a PPE or a PDE of *Lb*Cas12a-induced DSBs to a given end of *Sp*Cas9-induced DSBs. After 72 h, cells were harvested and gDNA was isolated.

To generate a standard curve for quantification of translocations, each type of translocation products in gDNA generated by *Lb*Cas12a-g2 and *Sp*Cas9-g10 and by *Lb*Cas12a-gC and *Sp*Cas9-g5 were amplified from gDNA by PCR with their respective primer pairs listed in Additional file 6: Table S5. About 100–200-bp PCR products of translocations were subcloned into the TA cloning vector (TSINGKE TSV- 007S pClone007 Simple Vector Kit). The plasmids containing single translocation templates for each translocation type were confirmed by PCR and Sanger sequencing and selected for generation of the standard curve. Sequences of single translocation templates in the plasmids for the standard curve were listed in Additional file 6: Table S5. We mixed 3 μg of gDNA from wild type cells with individual standard curve plasmids containing single translocation templates for each translocation type in serial dilution in 100 μL of standard solution, in which the concentrations of the plasmids were 2.53×10^{-6}, 2.53×10^{-7}, 2.53×10^{-8}, and 2.53×10^{-9} pmol per ng gDNA, respectively. We added 1 μL of standard solution containing 30 ng of gDNA into 10 μL qRT-PCR mixture using ChamQ Universal SYBR qPCR Master Mix (Vazyme) for qRT-PCR on CFX 96 Thermocycler (Bio-Rad). The primer pairs for qRT-PCR were also listed in Additional file 6: Table S5. The standard curves were created using the values from 3 independent experiments, triplicates in each experiment, with the *GAPDH* gene of gDNA as an internal control. The values of the X- and Y-axis indicate the amount of translocation products in pmol per ng gDNA for qRT-PCR and the corresponding $2^{-(\text{Ct standard curve plasmid} - \text{Ct GAPDH})}$ (i.e., $2^{-\Delta Ct}$) from qRT-PCR, respectively. The correlation coefficient (R^2) was calculated by GraphPad Prism 8.0.

To determine the level of translocations, translocations in gDNA were analyzed by qRT-PCR with 30 ng of gDNA containing translocations induced by *Lb*Cas12a-g2 and *Sp*Cas9-g10, by *Lb*Cas12a-g2r and *Sp*Cas9-g10, by *Lb*Cas12a-gC and *Sp*Cas9-g5, and by *Lb*Cas12a-gCr and *Sp*Cas9-g5. With *GAPDH* as internal control for qRT-PCR, the Y-axis values from qRT-PCR were calculated by $2^{-(\text{Ct translocation} - \text{Ct GAPDH})}$. The levels of translocation products were calculated according to their respective standard curves.

To examine each translocation type by PCR, translocation products were amplified from gDNA by PCR with primers listed in Additional file 6: Table S5, together with PCR products of the *GAPDH* region from each sample as internal PCR control. The intensity of translocation bands and *GAPDH* band was calculated with ImageJ software. The relative frequency of translocation was calculated as the ratio of translocation intensity to *GAPDH* intensity. PCR products of translocation were subcloned into the TA cloning

vector (TSINGKE TSV- 007S pClone007 Simple Vector Kit) and translocation junctions were determined by Sanger sequencing of individual PCR products subcloned.

Western blot and immunofluorescence

For western blot, RFP^+ cells from targeted integration experiments were sorted by FACS using Beckman Moflo Astrios EQ at 10 d post-transfection and lysed with RIPA buffer for 30 min. Cell extractions were separated by SDS-PAGE electrophoresis and analyzed by western blot with primary antibodies including anti-calnexin from ABclonal (A15631, 1:1000), anti-PCNA from Santa Cruz (sc- 56, 1:500), anti-H4 C16 from ABclonal (A19815, 1:1000), anti-β-actin from Vazyme (Ab101 -01/02/03, 1:1000), and anti-GAPDH from HuaBio (EM1101, 1:1000).

For immunofluorescence assays, sorted RFP^+ cells were grown on glass coverslips in a 24-well plate and fixed with 4% paraformaldehyde. Cells were stained with DAPI and imaged by a fluorescence microscope (Leica DM4000).

Off-target analysis

Potential off-target sites were identified using the latest version of the CRISPR Off-Target prediction website (http://crispor.tefor.net/). All potential sites were ranked by an off-target hit score, and high-ranked potential sites were selected. Off-target sites were amplified by PCR with primers listed in Additional file 7: Table S6 after gDNA extraction from cells transfected with *Sp*Cas9-sgRNAs or *Lb*Cas12a-sgRNAs at 3 d post-transfection in 293 T cells. Off-target editing efficiency was determined by Illumina deep sequencing and calculated as the ratio of edited reads to total reads normalized by transfection efficiency.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13059-025-03567-w.

Additional file 1: Fig. S1. Nearly symmetric deletion at two ends in NHEJ repair of *Sp*Cas9-induced DSBs. Fig. S2. Asymmetric deletion at two ends in NHEJ repair of *Lb*Cas12a-induced DSBs. Fig. S3. Insertions and deletions at *Sp*Cas9-induced DSBs at 15 target sites and *Lb*Cas12a-induced DSBs at 16 target sites. Fig. S4. Post-cleavage residency of LbCas12a at the PPEs inhibits recruitment of Ku80 to the ends. Fig. S5. Ligation between compatible PDEs of *Lb*Cas12a-induced DSBs is inherently accurate. Fig. S6. Standard curves for qRT-PCR quantification of translocations. Fig. S7. *Lb*Cas12a-g2 induces translocations more efficiently with the PDEs than the PPEs to *Sp*Cas9-induced DSBs. Fig. S8. *Lb*Cas12a-gC induces translocations more efficiently with the PDEs than the PPEs to *Sp*Cas9-induced DSBs. Fig. S9. *Lb*Cas12a induces translocations more efficiently with the PDEs than with the PPEs. Fig. S10. Frequency of RFP^+ cells generated by targeted integration with the donor cleaved by paired *Lb*Cas12a with either W/W-, W/C-, C/W-, or C/C-paired PAM orientation at the *AAVS1*, *CLTA*, or *HNRNPA1*locus in 293 T cells as indicated. Fig. S11. DNA agarose gel electrophoresis of PCR products for 5′and 3′ junctionof targeted integration with the donor cleaved by *Lb*Cas12a with either W/W-, W/C-, C/W-, or C/C-paired PAM orientation as indicated at *AAVS1*, *CLTA*, and *HNRNPA1*in 293 T cells. Fig. S12. Analysis of oriented integration, reverse integration, and duplicated integration of the donor induced by paired *Lb*Cas12a with either W/W-, W/C-, C/W-, or C/C-paired PAM orientation at the *AAVS1* locus in 293 T cells. Fig. S13. DNA agarose gel electrophoresis and Sanger sequencing of PCR products for gene replacement. Fig. S14. Effect of c-NHEJ deficiency on CIPDEL-mediated gene replacement. Fig. S15. Comparison between CIPDEL- and *Lb*Cas12-induced HDR-mediated targeted integration in 293 T cells and RPE- 1 cells. Fig. S16. Comparison between *Lb*Cas12a- and *Sp*Cas9-induced NHEJ- or HDR-mediated targeted integration in 293 T cells and RPE- 1 cells. Fig. S17. Flowchart of the CIPDEL protocol for N-terminal tagging, two-ended targeted integration and C-terminal tagging that requires precise ligation of compatible 5′-overhangs.

Additional file 2: Table S.1 Sequence information on targeted integration

Additional file 3: Table S2. Sequence information on gRNAs, primers, and single *Sp*Cas9 or *Lb*Cas12 target sites.

Additional file 4: Table S3. Sequence information on paired gRNAs, primers, and paired *Lb*Cas12a target sites.

Additional file 5: Table S4. Sequence information on ChIP gRNAs, primers, T7E1 primers, and target sequences

Additional file 6: Table S5. Sequence information on translocation

Additional file 7: Table S6. Sequence information on off-target

Additional file 8: Table S7. Source data

Acknowledgements
We are grateful to helpful discussion with members of the Xie lab. We thank Bi Chao and Hong Xiaoli from the Core Facilities, Zhejiang University School of Medicine for their technical support.

Peer review information
Albert Cheng and Claudia Feng were the primary editors of this article and managed its editorial process and peer review in collaboration with the rest of the editorial team. The peer-review history is available in the online version of this article.

Authors' contributions
A.-Y.X. conceived the project. A.-Y.X. and S.-M.X. supervised the study. R.-D.C., Y.Y., Q.-M.L., Y.-L.F., S.-C.L. and Y.W. generated DNA constructs and cell lines, and R.-D.C., Y.Y., Q.-M.L., Y.-L.F., J.-Z.H., C.-Y.Y., S.-C.L. and R.-R.J performed experiments and statistical analyses. P.-N.H., R.-G.T., and Y.-L.W provided assistance in experiments. A.-Y.X., R.-D.C., Y.Y., Q.-M.L. and Y.-L.F. analyzed the data. R.-D.C., Y.Y. and A.-Y.X. made figures and tables. A.-Y.X., R.-D.C. and Y.Y. wrote the manuscript. All authors read and approved the final manuscript.

Funding
This work is funded by the National Science and Technology Major Project of China (2023ZD0500501 to A.-Y.X.), the National Natural Science Foundation of China (No. 32371348 to A.-Y.X. and No. 32071439 to Y.-L.F.), the Department of Science and Technology of Hangzhou (202204 A05 to A.-Y.X. and 202204B08 to Y.-L.F.), and STI2030—Major Projects (2023ZD04048/2023ZD0404803 - 01 to R.-G.T.).

Data availability
All data supporting the findings of this study are available in the article or in the supplementary information files. Source data are provided in Additional file 8: Table S7. Deep sequencing raw data were deposited to SRA BioProject (accession number PRJNA1068007) [66]. Raw data of flow cytometry, Western blots, and immunofluorescence imaging were deposited to Zenodo, directly accessible at https://doi.org/10.5281/zenodo.10437601 [67] and https://doi.org/10.5281/zenodo.15074832 [68]. No custom source code was used in this study.

Declarations

Declarations
Ethics approval and consent to participate.
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Key Laboratory of Laparoscopic Technology of Zhejiang Province, Department of General Surgery, Sir Run-Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang 310016, People's Republic of China. [2]Hangzhou Qiantang Hospital, Hangzhou, Zhejiang 310018, People's Republic of China. [3]Institute of Translational Medicine, Zhejiang University School of Medicine, and Zhejiang University Cancer Center, Hangzhou, Zhejiang 310029, People's Republic of China. [4]Zhejiang Key Laboratory of Multiomics and Molecular Enzymology, Yangtze Delta Region Institute of Tsinghua University, Jiaxing, Zhejiang 314006, People's Republic of China. [5]Institute of Animal Husbandry and Inner Mongolia Key Laboratory of Sugarbeet Genetics and Germplasm Enhancement, Inner Mongolia Academy of Agricultural and Animal Husbandry Sciences, Hohhot, Inner Mongolia 010031, People's Republic of China.

Received: 4 November 2024 Accepted: 3 April 2025
Published online: 14 April 2025

References

1. Swarts DC, Jinek M. Cas9 versus Cas12a/Cpf1: structure-function comparisons and implications for genome editing. Wiley Interdiscip Rev RNA. 2018;9: e1481.
2. Kim H, Kim S-T, Ryu J, Kang B-C, Kim J-S, Kim S-G. CRISPR/Cpf1-mediated DNA-free plant genome editing. Nat Commun. 2017;8: 14406.
3. Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell. 2015;163:759–71.
4. Zetsche B, Heidenreich M, Mohanraju P, Fedorova I, Kneppers J, DeGennaro EM, et al. Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. Nat Biotechnol. 2017;35:31–4.
5. Koo T, Park SW, Jo DH, Kim D, Kim JH, Cho H-Y, et al. CRISPR-LbCpf1 prevents choroidal neovascularization in a mouse model of age-related macular degeneration. Nat Commun. 2018;9:1855.
6. Kim D, Kim J, Hur JK, Been KW, Yoon S-H, Kim J-S. Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells. Nat Biotechnol. 2016;34:863–8.

7. Kleinstiver BP, Tsai SQ, Prew MS, Nguyen NT, Welch MM, Lopez JM, et al. Genome-wide specificities of CRISPR-Cas Cpf1 nucleases in human cells. Nat Biotechnol. 2016;34:869–74.
8. Strohkendl I, Saifuddin FA, Rybarski JR, Finkelstein IJ, Russell R. Kinetic basis for DNA target specificity of CRISPR-Cas12a. Mol Cell. 2018;71:816-824.e3.
9. Nambiar TS, Baudrier L, Billon P, Ciccia A. CRISPR-based genome editing through the lens of DNA repair. Mol Cell. 2022;82:348–88.
10. Yang H, Ren S, Yu S, Pan H, Li T, Ge S, et al. Methods favoring homology-directed repair choice in response to CRISPR/Cas9 induced-double strand breaks. Int J Mol Sci. 2020;21:E6461.
11. Ramsden DA, Carvajal-Garcia J, Gupta GP. Mechanism, cellular functions and cancer roles of polymerase-theta-mediated DNA end joining. Nat Rev Mol Cell Biol. 2022;23:125–40.
12. Stinson BM, Loparo JJ. Repair of DNA double-strand breaks by the nonhomologous end joining pathway. Annu Rev Biochem. 2021;90:137–64.
13. Bétermier M, Bertrand P, Lopez BS. Is non-homologous end-joining really an inherently error-prone process? PLoS Genet. 2014;10: e1004086.
14. Guo T, Feng Y-L, Xiao J-J, Liu Q, Sun X-N, Xiang J-F, et al. Harnessing accurate non-homologous end joining for efficient precise deletion in CRISPR/Cas9-mediated genome editing. Genome Biol. 2018;19:170.
15. Suzuki K, Tsunekawa Y, Hernandez-Benitez R, Wu J, Zhu J, Kim EJ, et al. In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. Nature. 2016;540:144–9.
16. Li P, Zhang L, Li Z, Xu C, Du X, Wu S. Cas12a mediates efficient and precise endogenous gene tagging via MITI: microhomology-dependent targeted integrations. Cell Mol Life Sci. 2020;77:3875–84.
17. Zhao Z, Shang P, Sage F, Geijsen N. Ligation-assisted homologous recombination enables precise genome editing by deploying both MMEJ and HDR. Nucleic Acids Res. 2022;50: e62.
18. Suzuki K, Izpisua Belmonte JC. In vivo genome editing via the HITI method as a tool for gene therapy. J Hum Genet. 2018;63:157–64.
19. Feng Y, Liu S, Chen R, Xie A. Target binding and residence: a new determinant of DNA double-strand break repair pathway choice in CRISPR/Cas9 genome editing. J Zhejiang Univ Sci B. 2021;22:73–86.
20. Sternberg SH, Redding S, Jinek M, Greene EC, Doudna JA. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. Nature. 2014;507:62–7.
21. Richardson CD, Ray GJ, DeWitt MA, Curie GL, Corn JE. Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. Nat Biotechnol. 2016;34:339–44.
22. Singh D, Mallon J, Poddar A, Wang Y, Tippana R, Yang O, et al. Real-time observation of DNA target interrogation and product release by the RNA-guided endonuclease CRISPR Cpf1 (Cas12a). Proc Natl Acad Sci USA. 2018;115:5444–9.
23. Swarts DC, Jinek M. Mechanistic insights into the cis- and trans-acting DNase activities of Cas12a. Mol Cell. 2019;73:589-600.e4.
24. Jeon Y, Choi YH, Jang Y, Yu J, Goo J, Lee G, et al. Direct observation of DNA target searching and cleavage by CRISPR-Cas12a. Nat Commun. 2018;9:2777.
25. Liu S-C, Feng Y-L, Sun X-N, Chen R-D, Liu Q, Xiao J-J, et al. Target residence of Cas9-sgRNA influences DNA double-strand break repair pathway choices in CRISPR/Cas9 genome editing. Genome Biol. 2022;23:165.
26. Moreno-Mateos MA, Fernandez JP, Rouet R, Vejnar CE, Lane MA, Mis E, et al. CRISPR-Cpf1 mediates efficient homology-directed repair and temperature-controlled genome editing. Nat Commun. 2017;8:2024.
27. Wang Y, Liu KI, Sutrisnoh N-AB, Srinivasan H, Zhang J, Li J, et al. Systematic evaluation of CRISPR-Cas systems reveals design principles for genome editing in human cells. Genome Biol. 2018;19:62.
28. Shou J, Li J, Liu Y, Wu Q. Precise and predictable CRISPR chromosomal rearrangements reveal principles of Cas9-mediated nucleotide insertion. Mol Cell. 2018;71:498-509.e4.
29. Lemos BR, Kaplan AC, Bae JE, Ferrazzoli AE, Kuo J, Anand RP, et al. CRISPR/Cas9 cleavages in budding yeast reveal templated insertions and strand-specific insertion/deletion profiles. Proc Natl Acad Sci USA. 2018;115:E2040–7.
30. van Overbeek M, Capurso D, Carter MM, Thompson MS, Frias E, Russ C, et al. DNA repair profiling reveals nonrandom outcomes at Cas9-mediated breaks. Mol Cell. 2016;63:633–46.
31. Kuhfittig-Kulle S, Feldmann E, Odersky A, Kuliczkowska A, Goedecke W, Eggert A, et al. The mutagenic potential of non-homologous end joining in the absence of the NHEJ core factors Ku70/80, DNA-PKcs and XRCC4-LigIV. Mutagenesis. 2007;22:217–33.
32. Xie A, Kwok A, Scully R. Role of mammalian Mre11 in classical and alternative nonhomologous end joining. Nat Struct Mol Biol. 2009;16:814–8.
33. Verkuijl SA, Rots MG. The influence of eukaryotic chromatin state on CRISPR-Cas9 editing efficiencies. Curr Opin Biotechnol. 2019;55:68–73.
34. Wang AS, Chen LC, Wu RA, Hao Y, McSwiggen DT, Heckert AB, et al. The histone chaperone FACT induces Cas9 multi-turnover behavior and modifies genome manipulation in human cells. Mol Cell. 2020;79:221-233.e5.
35. Clarke R, Heler R, MacDougall MS, Yeo NC, Chavez A, Regan M, et al. Enhanced bacterial immunity and mammalian genome editing via RNA-polymerase-mediated dislodging of Cas9 from double-strand DNA breaks. Mol Cell. 2018;71:42-55.e8.
36. Strohkendl I, Saifuddin FA, Gibson BA, Rosen MK, Russell R, Finkelstein IJ. Inhibition of CRISPR-Cas12a DNA targeting by nucleosomes and chromatin. Sci Adv. 2021;7: eabd6030.
37. Reginato G, Dello Stritto MR, Wang Y, Hao J, Pavani R, Schmitz M, et al. HLTF disrupts Cas9-DNA post-cleavage complexes to allow DNA break processing. Nat Commun. 2024;15:5789.
38. Mimori T, Hardin JA, Steitz JA. Characterization of the DNA-binding protein antigen Ku recognized by autoantibodies from patients with rheumatic disorders. J Biol Chem. 1986;261:2274–8.
39. Grundy GJ, Moulding HA, Caldecott KW, Rulten SL. One ring to bring them all—the role of Ku in mammalian non-homologous end joining. DNA Repair. 2014;17:30–8.
40. Anzalone AV, Koblan LW, Liu DR. Genome editing with CRISPR-Cas nucleases, base editors, transposases and prime editors. Nat Biotechnol. 2020;38:824–44.
41. Doudna JA. The promise and challenge of therapeutic genome editing. Nature. 2020;578:229–36.

42. Orthwein A, Noordermeer SM, Wilson MD, Landry S, Enchev RI, Sherker A, et al. A mechanism for the suppression of homologous recombination in G1 cells. Nature. 2015;528:422–6.
43. Nami F, Basiri M, Satarian L, Curtiss C, Baharvand H, Verfaillie C. Strategies for in vivo genome editing in nondividing cells. Trends Biotechnol. 2018;36:770–86.
44. Wu Z, Yang H, Colosi P. Effect of genome size on AAV vector packaging. Mol Ther. 2010;18:80–6.
45. Samulski RJ, Muzyczka N. AAV-mediated gene therapy for research and therapeutic purposes. Annu Rev Virol. 2014;1:427–51.
46. Wang D, Zhang F, Gao G. CRISPR-based therapeutic genome editing: strategies and in vivo delivery by AAV vectors. Cell. 2020;181:136–50.
47. Ihry RJ, Worringer KA, Salick MR, Frias E, Ho D, Theriault K, et al. p53 inhibits CRISPR-Cas9 engineering in human pluripotent stem cells. Nat Med. 2018;24:939–46.
48. Haapaniemi E, Botla S, Persson J, Schmierer B, Taipale J. CRISPR-Cas9 genome editing induces a p53-mediated DNA damage response. Nat Med. 2018;24:927–30.
49. Pickar-Oliver A, Gough V, Bohning JD, Liu S, Robinson-Hamm JN, Daniels H, et al. Full-length dystrophin restoration via targeted exon integration by AAV-CRISPR in a humanized mouse model of Duchenne muscular dystrophy. Mol Ther. 2021;29:3243–57.
50. Stephenson AA, Nicolau S, Vetter TA, Dufresne GP, Frair EC, Sarff JE, et al. CRISPR-Cas9 homology-independent targeted integration of exons 1–19 restores full-length dystrophin in mice. Mol Ther Methods Clin Dev. 2023;30:486–99.
51. Kim D, Luk K, Wolfe SA, Kim J-S. Evaluating and enhancing target specificity of gene-editing nucleases and deaminases. Annu Rev Biochem. 2019;88:191–220.
52. Swartjes T, Staals RHJ, van der Oost J. Editor's cut: DNA cleavage by CRISPR RNA-guided nucleases Cas9 and Cas12a. Biochem Soc Trans. 2020;48:207–19.
53. Tóth E, Varga É, Kulcsár PI, Kocsis-Jutka V, Krausz SL, Nyeste A, et al. Improved LbCas12a variants with altered PAM specificities further broaden the genome targeting range of Cas12a nucleases. Nucleic Acids Res. 2020;48:3722–33.
54. Kleinstiver BP, Sousa AA, Walton RT, Tak YE, Hsu JY, Clement K, et al. Engineered CRISPR-Cas12a variants with increased activities and improved targeting ranges for gene, epigenetic and base editing. Nat Biotechnol. 2019;37:276–82.
55. Gao L, Cox DBT, Yan WX, Manteiga JC, Schneider MW, Yamano T, et al. Engineered Cpf1 variants with altered PAM specificities. Nat Biotechnol. 2017;35:789–92.
56. Nishimasu H, Yamano T, Gao L, Zhang F, Ishitani R, Nureki O. Structural basis for the altered PAM recognition by engineered CRISPR-Cpf1. Mol Cell. 2017;67:139-147.e2.
57. Dai X, Park JJ, Du Y, Na Z, Lam SZ, Chow RD, et al. Massively parallel knock-in engineering of human T cells. Nat Biotechnol. 2023;41:1239–55.
58. Campa CC, Weisbach NR, Santinha AJ, Incarnato D, Platt RJ. Multiplexed genome engineering by Cas12a and CRISPR arrays encoded on single transcripts. Nat Methods. 2019;16:887–93.
59. Feng Y-L, Xiang J-F, Liu S-C, Guo T, Yan G-F, Feng Y, et al. H2AX facilitates classical non-homologous end joining at the expense of limited nucleotide loss at repair junctions. Nucleic Acids Res. 2017;45:10614–33.
60. Feng Y-L, Liu S-C, Chen R-D, Sun X-N, Xiao J-J, Xiang J-F, et al. Proximal binding of dCas9 at a DNA double strand break stimulates homology-directed repair as a local inhibitor of classical non-homologous end joining. Nucleic Acids Res. 2023;51:2740–58.
61. Xie A, Puget N, Shim I, Odate S, Jarzyna I, Bassing CH, et al. Control of sister chromatid recombination by histone H2AX. Mol Cell. 2004;16:1017–25.
62. Willis NA, Scully R. Measurement of homologous recombination at stalled mammalian replication forks. Methods Mol Biol. 2021;2153:329–53.
63. Rass E, Chandramouly G, Zha S, Alt FW, Xie A. Ataxia telangiectasia mutated (ATM) is dispensable for endonuclease I-SceI-induced homologous recombination in mouse embryonic stem cells. J Biol Chem. 2013;288:7086–95.
64. Chandramouly G, Kwok A, Huang B, Willis NA, Xie A, Scully R. BRCA1 and CtIP suppress long-tract gene conversion between sister chromatids. Nat Commun. 2013;4:2404.
65. Feng Y-L, Liu Q, Chen R-D, Liu S-C, Huang Z-C, Liu K-M, et al. DNA nicks induce mutational signatures associated with BRCA1 deficiency. Nat Commun. 2022;13:4285.
66. Chen RD, Yang Y, Liu KM, Hu JZ, Feng YL, Yang CY, Jiang RR, Liu SC, Wang Y, Han PA, Tian RG, Wang YL, Xu SM, Xie AY. Post-cleavage target residence determines asymmetry in non-homologous end joining of Cas12a-induced DNA double strand breaks. Datasets. SRA. https://www.ncbi.nlm.nih.gov/sra/PRJNA1068007.
67. Chen RD, Yang Y, Liu KM, Hu JZ, Feng YL, Yang CY, Jiang RR, Liu SC, Wang Y, Han PA, Tian RG, Wang YL, Xu SM, Xie AY. Post-cleavage target residence determines asymmetry in non-homologous end joining of Cas12a-induced DNA double strand breaks. Datasets. Zenodo. 2024. https://doi.org/10.5281/zenodo.10437601.
68. Chen RD, Yang Y, Liu KM, Hu JZ, Feng YL, Yang CY, Jiang RR, Liu SC, Wang Y, Han PA, Tian RG, Wang YL, Xu SM, Xie AY. Post-cleavage target residence determines asymmetry in non-homologous end joining of Cas12a-induced DNA double strand breaks. Datasets. Zenodo. 2025. https://doi.org/10.5281/zenodo.15074832.

Publisher's Note