Plasmid2MC: Efficient cell-free recombination of plasmids into high-purity minicircle DNA for use in genome editing applications

Roman Teo Oliynyk[1,2*], Ahmed Mahas[1], Emil Karpinski[1] and George M. Church[1,3]

[1*]Department of Genetics, Harvard Medical School, Boston, MA, USA.
[2]Department of Computer Science, University of Auckland, Auckland, New Zealand.
[3]Wyss Institute for Biologically Inspired Engineering at Harvard University, Boston, MA, USA.

*Corresponding author(s). E-mail(s): roli573@aucklanduni.ac.nz;

Abstract

DNA plasmids are commonly used for delivering proteins and RNA in genome editing. However, their bacterial components can lead to inactivation, cell toxicity, and reduced efficiency compared to minicircle DNA (mcDNA) lacking such bacterial sequences. Existing commercial kits that recombine plasmids into mcDNA within proprietary bacterial strains are labor-intensive, yield inconsistent results, and often produce low-quality mcDNA. To address this, we developed Plasmid2MC, a cell-free method using **Φ**C31 recombination to efficiently excise the bacterial backbone from conventionally prepared plasmids, while mcDNA purification steps digest all DNA impurities and reduce endotoxin levels. We demonstrated application of mcDNA to express CRISPR-dCas9 for base editing in HEK293T cells and mouse embryonic stem cells, as well as for homology-independent targeted insertion (HITI) genome editing. The method's ease of preparation, efficiency and mcDNA purity make it a valuable alternative for applications requiring bacterial backbone-free circular DNA.

Keywords: Recombinase, Integrase, **Φ**C31, phiC31, Cloning, Endotoxins, CRISPR-Cas9, Base Editing, HITI, Genome Editing

Introduction

Plasmids, bacterial extrachromosomal self-replicating circular DNA elements, were first described by Joshua Lederberg in 1952 [1]. They were later found to facilitate antibiotic resistance sharing and bacterial genome modification [2, 3]. Cohen et al. [4] demonstrated the construction of biologically functional bacterial plasmids in vitro in 1973. Since then, plasmids have been widely used for genome editing constructs delivery [5–7], including zinc-finger nucleases, transcription activator-like effector nucleases (TALENs), and CRISPR-based methods [8–12].

However, the prokaryotic sequences within plasmids, notably antibiotic resistance genes, are redundant for eukaryotic cell transfection and have the potential to elicit adverse effects on gene expression, cellular physiology, immune response [13–15], and lead to plasmid inactivation [15]. Minicircle DNA (mcDNA), produced from plasmids by eliminating bacterial backbones, can avoid these effects. mcDNA has shown lower inactivation, better nuclear uptake, and greater protein expression compared to plasmids [15–17], making it a promising tool for non-viral vector delivery [18].

Numerous studies have demonstrated the production of mcDNA using site-specific recombination [13, 14, 19, 20], where a serine integrase performs double-stranded DNA cleavage and subunit rotation to exchange DNA strands at two recombining sites present in a plasmid—the so-called attP and attB (abbreviated from attachment-phage and attachment-bacteria, respectively) sites [21]—thus separating the intended mcDNA and a bacterial backbone into two intercoiled DNA circles (see illustration in the Results). The recombination is then followed by a restriction enzyme digestion. Both recombinase and restriction enzyme genes are designed into specialized bacterial strains and an inducible activation is used to express them, resulting in the production of mcDNA within the bacteria [13, 14, 19, 20]. We found commercial kits [22] using this methodology labor-intensive and yielding inconsistent results. Purity and toxicity are known issues for such bacterially recombined mcDNAs, and additional cleanup is also cumbersome and labor intensive [18].

No protocol or commercial kit for cell-free plasmid recombination into mcDNA for use in genome editing experiments has been available. To address this shortcoming, we present Plasmid2MC, a cell-free protocol leveraging ΦC31 integrase—one of the best-characterized serine recombinases to date [23–29]. In Plasmid2MC, ΦC31 recombination efficiently excises the bacterial backbone from conventionally prepared plasmids in a cell-free reaction, while mcDNA purification steps additionally reduce endotoxin levels. The method's ease of preparation, efficiency, and mcDNA purity position it as a valuable alternative to existing methods for genome editing applications that require bacterial backbone-free circular DNA. We demonstrated the efficacy of our mcDNA by using it to express CRISPR-dCas9 for base editing in HEK293T and mouse embryonic stem cells (mESCs), as well as for homology-independent targeted insertion (HITI) genome editing.

Results

Method overview

In this study, we developed an efficient cell-free method that takes as an input customarily cloned plasmids containing two recombining sites (attP and attB, Fig. 1a), and with the help of ΦC31 integrase recombines these plasmids into two intercoiled DNA circles, one containing bacterial backbone with all components used for bacterial amplification, and another containing the mcDNA of interest Fig. 1b,c.

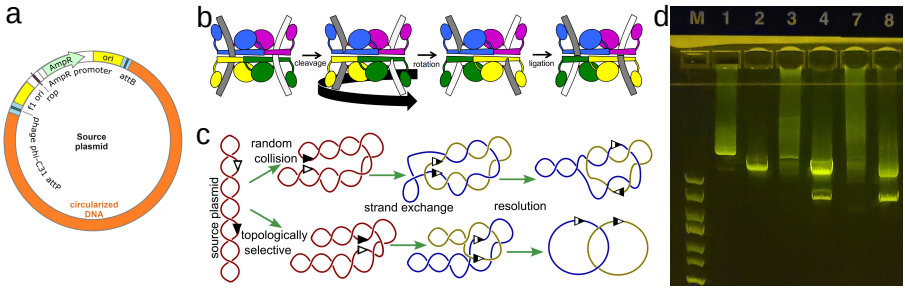

Fig. 1 ΦC31 recombination: mechanism diagram and experiment. (**a**) Diagram of a plasmid containing attP and attB recombination sites separating the bacterial backbone and the intended mcDNA. (**b**) Subunit rotation mechanism for strand exchange by serine recombinases (this subplot is adopted from Trejo et al. 2018 [30]). (**c**) Random collision and topologically selective strand exchange both lead to intercoiled structures of two DNA loops of varying complexity among the recombined molecules [21]. (**d**) Gel image showing the ΦC31 recombination of 10418bp plasmid into a 2.7kbp bacterial backbone circle and 7758bp DNA circle. Each lane contains 300ng of DNA.
L1 - pure supercoiled plasmid; L2 - same plasmid with a single cut with a restriction enzyme (NotI); linear DNA travels faster than supercoiled DNA, and L1 presents a very weak matching band showing a minuscule amount of linear plasmid;
L3 - after 80 minutes of recombination showing diffuse DNA scattering corresponding to the varying speed of travel of the miscellaneous catemers depicted in (b), and L4 same product cut with NotI, which is present in a single location on the bacterial backbone, thus allowing bacterial backbone release from the tangle, resulting in distinct bands (2.7 and 7.7kbp) and a weaker band corresponding to the non-recombined plasmid;
L7 and L8 - same as L3 and L4 after 12 hours of recombination; the original plasmid band is now nearly invisible (see the unprocessed gel photo in Supplementary Fig. 1).

We initially aimed for a single-tube protocol, where the recombination would be followed by the cutting of the bacterial backbone using restriction enzymes and digestion with T5 exonuclease within the same reaction tube, followed by the spin column cleanup, reminiscent to what we successfully achieved for the ligation reaction in Circular Vector (CV) method (see Oliynyk & Church (2022) [31]). However, preliminary experiments have shown that the cleanup step after the completion of recombination was indispensable in Plasmid2MC, even though it lowered the final mcDNA yield by 16% for 5.7kbp source plasmid size and by 31.5% for 13.9kbp source plasmid (see Methods DNA purification after the completion of recombination and Protocol steps).

The recombination buffer containing 20mM Tris, 0.1mM EDTA, and 100mM NaCl worked optimally (see Fine-tuning the recombination buffer composition in the Methods), and most experiments were performed with a recombination duration of 12 hours (see the finalized Protocol steps in the Methods).

Plasmid2MC protocol yields at varying ΦC31 integrase and source plasmid DNA concentrations

The combined recombination results are presented in Fig. 2 (see also Component concentrations in recombination experiments in the Methods). Three plasmid sizes were used in the following analysis, displayed as *mcDNA size/source plasmid size* in base pairs (bps)—3096/5758, 7758/10418, and 11282/13942bp—thus sampling the recombination protocol performance in a commonly used range of plasmids sizes and corresponding mcDNAs (see Plasmid design and assembly in the Methods, Supplementary Data and Supplementary Note 2). In Fig. 2a,b, both ΦC31 and source plasmid DNA (10418bp plasmid used) had concentrations of 100ng/μl. Recombination efficiency gradually increases between 18 and 30°C, reaching 38% at 30°C, which is the optimal temperature (see Fig. 2a). Recombination efficiency diminishes thereafter to 5% at 37°C. The reaction yield increased with longer recombination times, even showing additional gains from 33.3% at 4 hours to 37.9% at 12 hours of incubation (see Fig. 2b).

Further investigations with the three test plasmids at a constant plasmid DNA concentration of 100ng/μl and variable ΦC31 (see Fig. 2c) showed that at 25ng/μl of ΦC31, mcDNA yield was low for all three plasmids, then began plateauing at 100ng/μl. An additional investigation with fixed ΦC31 at 100ng/μl (see Fig. 2d) showed a slight yield decrease for 10418 and 13942bps plasmids when DNA concentration increased to 200ng/μl, with a steeper drop at 300ng/μl. At this ΦC31 concentration, the yield for the shortest 5758bp plasmid nearly halved from 27.2 to 17.2% when DNA concentration doubled; when DNA concentration tripled to 300ng/μl, the yield decreased to 7.6%. This suggests that 100ng/μl ΦC31 may have been insufficient for the recombination of the short plasmid at higher DNA concentrations. Increasing the ΦC31 concentration to 200ng/μl improved performance by 1–2% for all three source plasmids at a DNA concentration of 100ng/μl and mitigated the performance loss for the 5758bp plasmid at higher DNA concentrations, resulting in parallel performance patterns for all three plasmids (see Fig. 2e). This indicated that smaller plasmids, with a larger relative number of attP and/ attB sites per unit of weight, may require higher concentrations of ΦC31 recombinase. Interestingly, the mcDNA yield was similar (25–30%) for three plasmids at the most effective concentrations.

While the above results reported the pure mcDNA yield, the successful recombination reaction also contains a circular bacterial backbone of constant size. Fig. 2f–h are equivalent to Fig. 2c–e, showing full recombination efficiency while accounting for the digested bacterial backbone. Maximum recombination performances were 54, 40, and 32% for the shortest, medium, and longest

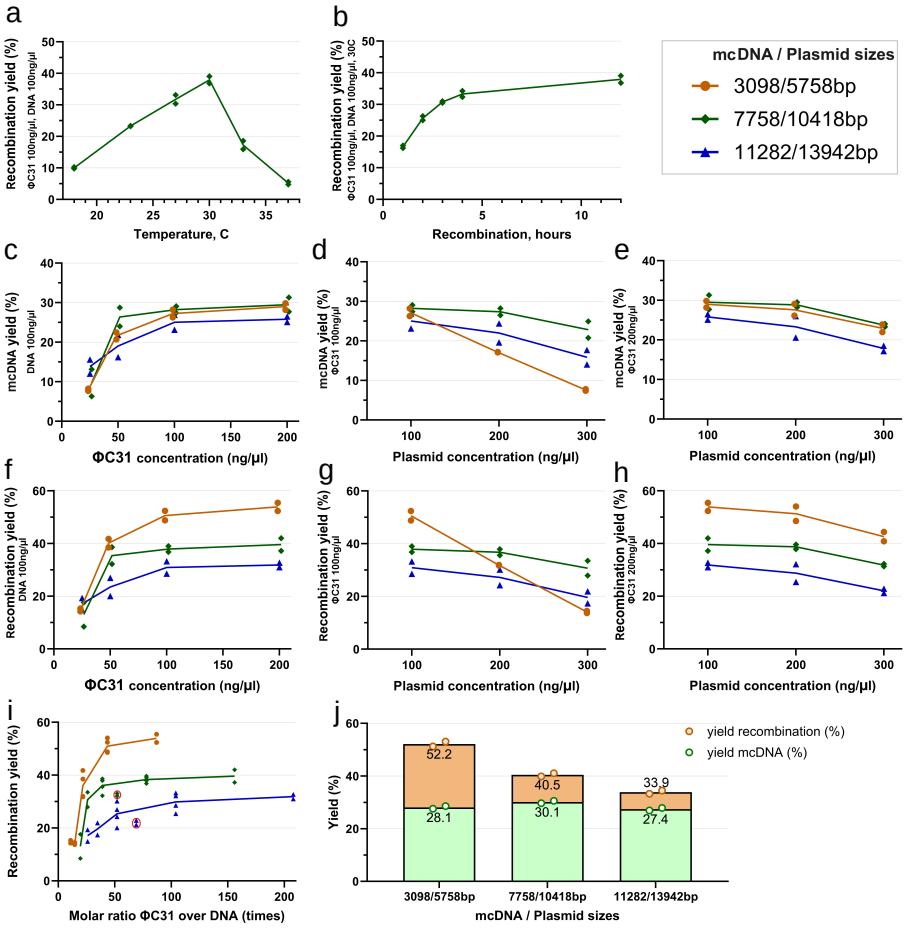

Fig. 2 Recombination reaction efficiency and mcDNA yields. (**a**) Temperature curve of ΦC31 recombination efficiency, showing the distinct maximum yield at 30°C. (**b**) Recombination yield by time at the optimal temperature of 30°C. (**c**) mcDNA yield at ΦC31 concentrations of 25–200ng/μl and a constant DNA concentration of 100ng/μl. (**d**) mcDNA yield at 100ng/μl ΦC31 and DNA concentrations of 100–300ng/μl. (**e**) mcDNA yield at 200ng/μl ΦC31 and DNA concentrations of 100–300ng/μl. The following plots (f)–(h) account for the bacterial backbone DNA circle for complete recombination efficiency (including length of bacterial backbone, and discounting processing losses): (**f**) Recombination yield at ΦC31 concentrations of 25–200ng/μl and a constant DNA concentration of 100ng/μl. (**g**) Recombination yield at 100ng/μl ΦC31 and DNA concentrations of 100–300ng/μl. (**h**) Recombination yield at 200ng/μl ΦC31 and DNA concentrations of 100–300ng/μl. (**i**) Combining the data points shown above presented a molar ratio of ΦC31 to source plasmid DNA molecules. (**j**) Yield test at optimum concentrations and 20μg of input DNA at a concentration of 100ng/μl and ΦC31 concentrations in between those used in plots (e) and (h), and larger elution volume; maximum performance is higher in plots (e) and (h) at 200ng/μl ΦC31 for 5758 and 10418bp plasmids. Note: Tests for each data point were performed in duplicate. In cases when error bars or data points are not visible, they overlap within a data point.

plasmids, respectively. One explanation for this performance difference is that the cleanup column efficiency decreases with increasing DNA length, with column losses from two cleanups exceeding 50Additional DNA nicking or sheering may occur during initial plasmid extraction or intermediate cleanup, and these molecules will be digested by T5 exonuclease during digestion stage.

Recalculating the molar concentration of ΦC31 and dividing it by molar concentrations of the three test plasmids from the available data points (see Fig. 2i, and Component concentrations in recombination experiments in the Methods) demonstrated that the recombination reaction yield is relatively low when ΦC31 molar excess is at or below 25X, and begins plateauing for all three plasmids at approximately 50X ΦC31 molar excess, with only fractional improvement at higher molar excesses. This figure is a sampling of clearly non-linear correspondence between concentrations and the molar excesses of these two variables. For example, at higher concentrations (DNA at 300ng/μl and ΦC31 at 200ng/μl), the relative performance is notably below the trend line for lower molar ratios (see circled red data points for 10418bp and 13942bp plasmids in Fig. 2i), showing that at such a high DNA concentration, the efficiency of recombination is diminishing.

The experiments described above were performed with elution volumes of 22–25μl, which aimed to achieve higher concentrations of DNA, thereby allowing for accurate measurements. Based on the observations in Fig. 2i, we decided to verify the yields of the middle ground concentrations of ΦC31, equal to 130ng/μl for 13942 and 10418bp plasmids and 180ng/μl for the 5758bp plasmid, with a reaction volume of 200μl and double the source plasmid weight of 20μg (DNA concentration of 100ng/μl) using an elution volume of 54μl. After incubating the columns at 55°C for 3–4 minutes before elution, centrifugation yielded 51–52μl of mcDNA elute. The results presented in Fig. 2j show that even with higher elution volume, the recombination yield was essentially in between the results for ΦC31 concentrations of 100ng/μl in Fig. 2e and 200ng/μl in Fig. 2h. Only the largest 13942bp source plasmid showed a 2% improvement, thus reconciling the maximum recombination efficiency to 34% and mcDNA yield to 27% at this source plasmid size.

This outcome aligns with our experience that heating the QIAquick cleanup columns at 55°C for 3–4 minutes before centrifugation results in an effective QIAquick column elution rate when using elution volumes of 20–25μl, with a repeated elution from the same column yielding only a few additional ng/μl. This shows that the use of low elution volumes, necessary for achieving high mcDNA concentration required for the genome editing experiments, will not result in significant mcDNA losses with QIAquick spin columns.

Optimal concentrations of source DNA and ΦC31

Having consolidated the above data, we can formulate the consideration for the optimal concentrations and quantities of ΦC31 and source plasmids. Plasmid midi preps typically yield 400–600μg of plasmid per 45ml of *E. coli* culture, providing abundant source plasmid DNA. To economize the use of ΦC31

recombinase, it may be optimal to use 100ng/ul of ΦC31 and 200–300ng/μl of DNA for larger plasmids (e.g., see Fig. 2d,g) and a somewhat higher ΦC31 concentration for smaller plasmids (see the 5758bp plasmid in Fig. 2e,h).

We used QIAquick cleanup column kits for ease of handling; however, researchers may choose their preferred efficient cleanup methods. Larger quantities of mcDNA can be prepared more economically using isopropanol precipitation, although this method involves more hands-on steps and processing time [32] (see specifics in Protocol steps). For instance, using 100.0μg of the 13.9 kbp source plasmid, we obtained 28.0μg of mcDNA (28.0 ± 0.3% yield).

The mcDNA use in CRISPR-dCas9 base editing in HEK293T and mESCs

Having optimized the protocol using ΦC31 recombinase/integrase to effectively produce mcDNA, it would be customary to demonstrate the performance of the resulting DNA constructs. For this validation, we used the plasmid *p-att-ef1a-ABE8e-dCas9-BSD* (see the Methods) and corresponding mcDNA produced using our protocol. This plasmid is an implementation of TadA for the base editing version ABE8e introduced by Richter et al. [11], assembled into our plasmid containing ef1α promoter, attP and attB sites, blasticidin selection resistance gene, and dCas9—which is used in our ongoing research (rather than nCas9 used by Richter et al. [11]; see Fig. 3a). Applying the Plasmid2MC protocol to the source plasmid *p-att-ef1a-ABE8e-dCas9-BSD* we produced 27μg of mcDNA at a concentration of 588ng/μl (further vacuum-concentrated when needed) for use in these validation experiments. The resulting mcDNA was primarily represented by the single circular DNA, confirming the known ΦC31 preference for intra-molecular recombination [21], with a small proportion of double-sized intermolecular concatemer, marked "dimer" on the sequencing histogram from Primordium Labs (see Fig. 3b and Supplementary Note 2). The sequencing showed the perfect matching of the expected and achieved mcDNA (sequencing files are available in the Supplementary Data).

We used the Circular Vector (CV) protocol to prepare the five gRNA expression vectors used for five base editing experiments (see Supplementary Fig. 3). Briefly, the CV protocol takes a linear double-stranded DNA fragment and ligates/purifies it into short, circularized DNA containing gRNA expression element illustrated in Fig. 3c, which can be completed and ready for transfection in as little as 3 hours (for implementation details, see Oliynyk & Church [31]). This protocol is particularly useful for making multiple targeting gRNA expression vectors with minimal hands-on time, in conjunction with larger proteins expressing plasmids (and now mcDNAs), which are more labor intensive but can be reused for multiple editing targets.

We targeted three sites from Richter et al. [11], demonstrating A→G conversion in HEK293T cells (see Fig. 3d). Fig. 3e also demonstrates successful A→G conversion in two locations in mESCs using the same plasmid and

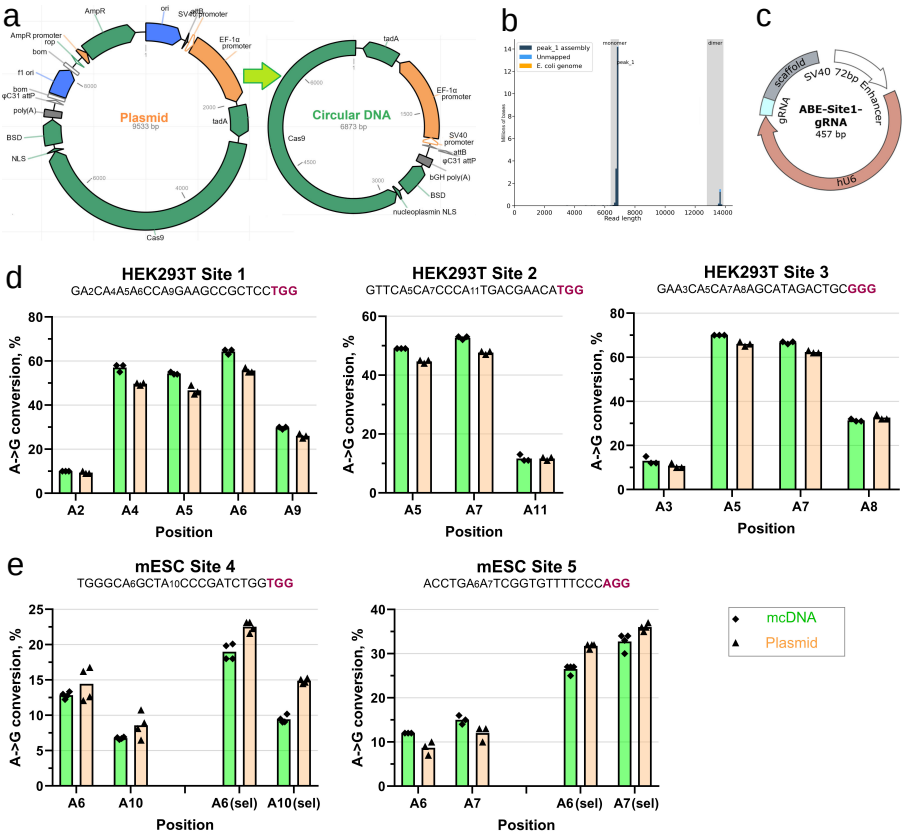

Fig. 3 Validation of mcDNA performance in ABE8e dCas9 base editing. (**a**) Diagram of the source plasmid and the resulting mcDNA with excluded bacterial backbone, used in this validation edit. (**b**) Primordium Labs sequencing histogram showing the high quality of the mcDNA product. (**c**) A circular vector for the expression of gRNA for base editing targeting, using the method developed earlier in our laboratory [31]. (**d**) Base editing (dCas9 ABE8e) A→G conversion rates in HEK293T cells for the first three sites from [11], using mcDNA (green) and the original plasmid (orange), with blasticidin selection. (**e**) dCas9 ABE8e editing A→G conversion rates in two mESC sites using mcDNA (green) and the original plasmid (orange), without selection and following blasticidin selection marked by (sel). All data points are based on two physical and one or two technical samples. Guide sequences for the sites are followed by PAM (in bold and red).

mcDNA. The editing of HEK293T cells was followed by antibiotic (blasticidin) selection, and mESC editing was performed with and without selection. The editing performance was close in magnitude and differed fractionally for both the plasmid and the mcDNA. Although the transfection concentrations could be further optimized, this is not necessary for the present validation. These experiments showed comparable base editing performance when using the plasmid and mcDNA in both cell types.

The mcDNA use in homology-independent targeted integration (HITI)

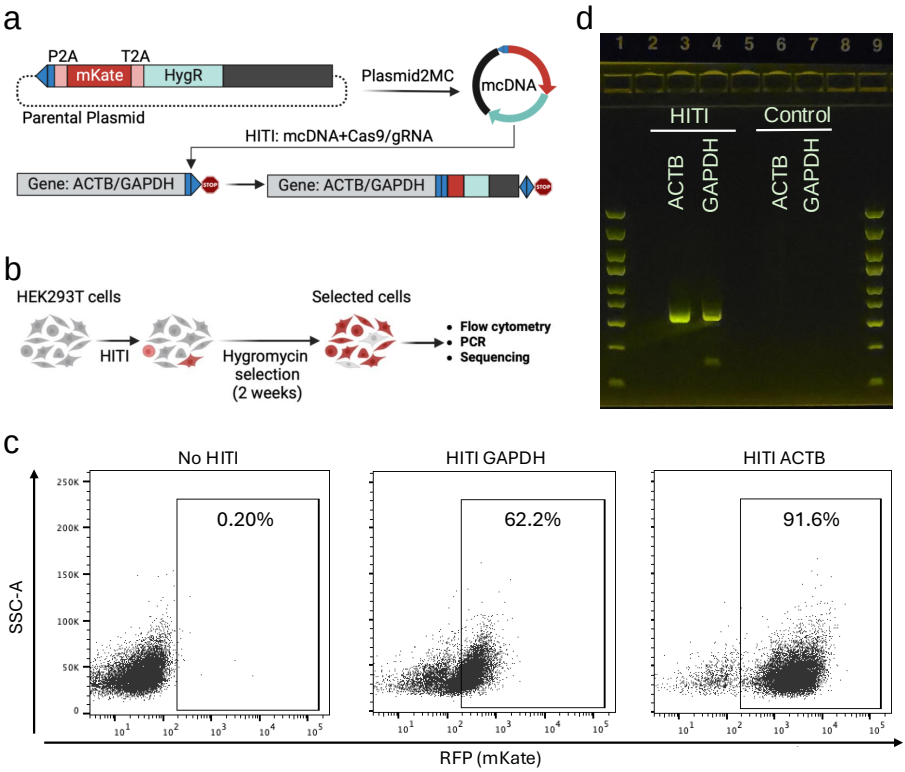

Fig. 4 HITI editing results. (**a**) Schematic showing the design of HITI constructs. Blue pentagon, Cas9/gRNA target sequence in the host genome and HITI mcDNA. HygR: Hygromycin-resistant gene. Black square: non-functional random DNA sequence. (**b**) Schematic of HITI experiment with antibiotic selection to enrich cells with the correct HITI-mediated integration before subsequent analysis. (**c**) Flow cytometry analysis of cells treated with HITI or control (untransfected cells). Percentage of positive population is shown within the gates. (**d**) Gel electrophoresis of PCR overlapping the mKate gene and either the ACTB or GAPDH genes. HITI - shows bands for edited samples, while no bands are visible for unedited Controls.
See Supplementary Fig. 9 for additional analysis and unprocessed gel electrophoresis photos.

mcDNAs have gained significant interest for their utility in site-specific genome integration applications. HITI utilizes mcDNA for the insertion of exogenous DNA into targeted genomic sites without the need for homologous recombination [33, 34]. Here, we further demonstrated the utility of Plasmid2MC to produce HITI-compatible mcDNAs. We designed mcDNA constructs encoding a fluorescent protein cassette (mKate) linked to an antibiotic selection marker (Hygromycin) via a T2A peptide sequence, which allows in-frame insertion of the fluorescent protein and selection marker into

housekeeping genes, including ACTB and GAPDH (see Fig. 4a). To showcase the preparation of larger HITI templates, we added a non-functional random DNA sequence downstream of the hygromycin resistance gene, bringing the total size of the HITI template to approximately 3 kbp. The integration sites were strategically chosen just upstream of the stop codons of these housekeeping genes. A P2A sequence was incorporated immediately upstream of the donor template sequence to ensure that the insertion of the HITI constructs occurred in-frame, facilitating the translation of the downstream HITI elements as part of the same polypeptide as the endogenous proteins (see Fig. 4a). We employed the CRISPR-Cas9 system with gRNA CVs specifically designed to target these integration sites alongside our mcDNA constructs. Following transfection, cells were cultured under hygromycin selection for two weeks to ensure the survival of only those cells that successfully integrated the mcDNA constructs (see Fig. 4b).

Flow cytometry analysis revealed a significant population of cells expressing mKate, with 91.6% of ACTB-targeted cells and 62.2% of GAPDH-targeted cells displaying the red fluorescent protein, indicating successful integration and expression (see Fig. 4c). To confirm the integration of the HITI mcDNA at the genomic level, PCR was conducted using primers specific to the mKate gene and either the ACTB or GAPDH genes. The expected amplicons were observed only in the HITI-treated cells, with no amplification in control groups (see Fig. 4d). Subsequent Sanger sequencing of these amplicons verified the precise integration of the HITI constructs at the targeted genomic loci, confirming the correct junctions between the mcDNA and the host genome. These findings substantiate the efficacy of Plasmid2MC method for producing large HITI-compatible mcDNA donor templates.

Reduction of endotoxin levels in mcDNA compared to source plasmids

GenScript's testing revealed that source plasmids prepared with QIAGEN Midiprep Plus contained 0.26598 EU per μg of DNA, while resultant mcDNA had 0.08986 EU per μg. Although the protocol involves two cleanups with wash stages, spin column cleanups alone are ineffective at removing endotoxins. Our hypothesis is that endotoxin elimination occurs during the digestion stage, where T5 exonuclease and SspI restriction enzyme may bind endotoxins, forming endotoxin–protein complexes [35, 36], which are subsequently washed away. To validate our hypothesis, we introduced Proteinase K after the digestion step and observed increased endotoxin levels. Additionally, commercial T5 exonuclease and restriction enzymes are likely to contain some level of bound endotoxins, which can be released through proteinase digestion. Therefore, using Proteinase K after the recombination reaction is crucial (see in Methods), but proteinase digestion should not follow T5 exonuclease and restriction enzyme digestion.

In conclusion, following the Plasmid2MC protocol demonstrated three-fold reduction in endotoxin levels in mcDNA compared to the source plasmid prepared using QIAGEN Midiprep Plus.

Discussion

The Plasmid2MC method and protocol presented here elucidate the preparation of mcDNA, originating with conventionally amplified or cloned plasmids containing ΦC31 attP and attB sites at the junctions between the bacterial backbone and the intended mcDNA, facilitated by ΦC31 recombinase. In this protocol, ΦC31 recombination efficiently excises the bacterial backbone, followed by purification steps digesting the separated bacterial backbone and reducing endotoxin levels (see Methods). Our approach provides an alternative to costly proprietary kits reliant on specialized bacterial strains, and yields highly purified mcDNA with low levels of bacterial endotoxins. Notably, the intrinsic recombination reaction efficiency approaches 100%, with DNA losses primarily attributed to these purification steps, resulting in an overall protocol efficiency ranging from 55% to 34% for plasmid sizes ranging from 5.8 to 13.9 kbp, respectively. Using a 2.7kbp bacterial backbone, the pure mcDNA conversion efficiency stands at impressive 27% to 30% after deducting the backbone size. While on the face of it, a conversion yield of 55–34% may appear low, the fact is that we are using conventionally produced, cheap, and abundant plasmids as the source material. The efficiency lies in the simplicity and ease of production of high-quality purified mcDNA in quantities needed for genome editing experiments.

Presently, ΦC31 protein remains unavailable off-the-shelf, yet one could foresee its availability emerging with demand, similarly to numerous commercially available proteins and restriction enzymes. Our pricing for a single custom project with initial validation totaled $4932.71, with optional nucleic acid and nuclease removal priced at $1125 from GenScript, yielding 18.96mg of ΦC31. At a recombination yield of 30%, this quantity of ΦC31 suffices for producing at least 5.6 mg of mcDNA and up to double this amount when using a higher excess of source plasmid to preserve ΦC31. Hypothetically, a kit containing 1/100 of this amount of ΦC31 would cost $60.58 and would be capable of producing 56–112μg of mcDNA. Assuming that production of ΦC31 at industrial volumes would cost a fraction of the amount we paid, and with the recombination buffer not containing any expensive or difficult-to-handle components, such a kit may be priced only slightly above this. Even at double the aforementioned price estimate, such a kit would be comparable to many commercially available restriction enzyme and protein kits.

There are fundamental differences between the Plasmid2MC method and our previously published CV method [31] that is primarily aimed at efficient fully synthetic substitution for the plasmid-based guide RNA expression in CRISPR, base, and prime gene editing in cell cultures. The CV method offers rapid cell-free preparation of short circular DNA ready for transcription within

3 hours after receiving dsDNA from a commercial supplier, with minimal hands-on time [31]. In Oliynyk & Church 2022 [31] we noted that error rates in commercially delivered dsDNA fragments were ∼1/6,000 per nucleotide for various commercial products [37–39]. This error rate is excellent for short CRISPR-Cas9 guide RNA sequences but may lead to a significant proportion of defective proteins when extending the circular DNA length to thousands of base pairs [31] in case of mcDNA. Additionally, the CV method performance is optimal at a CV size of 450bp, where it reaches up to 62% efficiency in converting input linear double-stranded DNA into the circular expression vectors, diminishing to 30% efficiency at a 950bp CV size and further to 5–8% efficiency at 1800bp. This led us to define 450–950bp as the practical range of the CV protocol [31]. The principal reason for this is that at longer linear source DNA sizes at practical concentrations, the CV ligation reaction favors polymerization into infinite chains of concatemers, with a vanishingly low proportion of circular DNA units [40, 41].

The Plasmid2MC method addresses the aforementioned limitations of the CV method by leveraging *E. coli* amplification of a single plasmid clone, virtually eliminating DNA replication errors [31, 42, 43]. Our sequencing of recombined and purified mcDNA invariably showed a perfect match of the expected and achieved mcDNA sequences. Consequently, mcDNA produced by this method proves applicable not only for protein-expressing genome editing constructs but also for HITI templates and other large circular DNA constructs.

Both methods can be used complementarily, as demonstrated by utilizing CVs for guide gRNA expression alongside Cas9 and other proteins expressed by mcDNA in base editing validations in the present study. Here, the Plasmid2MC method facilitates the production of larger volumes of mcDNA, reusable across multiple experiments, while multiple guide RNAs expressing CVs are promptly produced by the CV protocol as needed. In our lab we continue to explore genome editing combining these two methods.

In most scenarios, the use of kits like QIAGEN Midiprep Plus ensures the removal of proteins and RNA during the preparation of the source plasmids, while the Plasmid2MC protocol additionally lowers endotoxin levels and completely digests all DNA other than the intended mcDNA. In a recent analysis, Almeida et al. [18] highlighted a common issue with specialized bacterial kits utilizing ΦC31 and I-SceI: incomplete recombination and digestion, resulting in mcDNA tainted by remnants of parental plasmids and recombined empty bacterial backbone plasmids, which are unacceptable for therapeutic applications, posing significant challenges for downstream purification [44–46]. In contrast, the Plasmid2MC cell-free recombination approach guarantees the production of highly purified mcDNA.

In conclusion, the method's ease of preparation, efficiency, and mcDNA purity make Plasmid2MC a valuable new tool for CRISPR-Cas9 and other applications that require bacterial backbone-free circular DNA.

Methods

Plasmid2MC protocol

Preliminary steps

Most of the reaction enzymes and buffers were based on commercially available kits and commonly used equipment. The exceptions include obtaining ΦC31 integrase and preparing the 2X reaction buffer as described below, and a small modification to the design of the plasmids.

1. **Obtain ΦC31 integrase protein from a reliable vendor**. We subcontracted GenScript and received 18.96mg of ΦC31 protein at a concentration of 4.74mg/ml (see Reagents and Materials).
2. **Dilute the ΦC31 with pure glycerol to make a working buffer**. Add glycerol to achieve a 50% vol/vol concentration suitable for storage at -20°C in liquid form. In our case, the ΦC31 concentration in the working buffer became 2.6μg/μl.
3. **Prepare the 2X recombination buffer to contain:** 200mM NaCl, 40mM Tris pH, 0.2mM EDTA pH 7.8–8.0.
   For example, 50ml of 2X buffer can be prepared as follows:
   a) 10ml of 10mM-Tris, 1mM-EDTA (pH 8.0; Fisher Scientific Cat#BP2473) and a negligible addition of Tris to produce 2X of 0.1Mm EDTA.
   b) 2ml of 1M Tris (pH 8.0; Corning Cat#46-031-CM).
   c) 584mg of NaCl.
   d) 38ml of nuclease-free H_2O (Invitrogen 10977015).
   Note: The literature lists a variety of buffer compositions and additives, such as TPP, BSA, and spermidine [21, 23, 47]. Our testing demonstrated that customarily used ingredients other than those listed above resulted in no effect—or even lower recombination yield—and appeared to be unnecessary. As demonstrated in the Results, we achieved complete recombination using our 2X buffer overnight, while ΦC31 protein was sufficiently stable to provide yield improvements over a 12-hour reaction versus a 4-hour reaction. Buffers with NaCl concentration in the range of 100–150mM worked equally well, and the small increase of NaCl from the working buffer remained well within this range, even at the highest concentration of ΦC31 (the Φ31 storage buffer contained NaCl, see Reagents and materials).
4. **Make alterations to the design of source plasmids to include ΦC31 attP and attB sequences and unique restriction enzyme sites in the bacterial backbone area.** The sequences are attB 5'-GTGCCAGGGCGTGCCCTTGGGCTCCCCGGGCGCG and attP 5'-GCCCCAACTGGGGTAACCTTTGAGTTCTCTCAGTTGGGGG, with the desired circularizing DNA sequence in between. Position them in a manner that allows the separation of the desired mcDNA and bacterial backbone elements into two circular units. The recombination occurs near

the middle of the above sequences, with circularized DNA containing 5'-GCCCCAACTGGGGTAACCTTTGGGCTCCCCGGGCGCG, which may need to be considered when designing the mcDNA for a specific purpose.
The bacterial backbone must contain a unique restriction enzyme site, which will be used to cut and digest the bacterial backbone while leaving the desired mcDNA intact. It is preferred to use an efficient restriction enzyme. As only three examples out of many possibilities, we prefer to use New England Biolabs (NEB) SspI-HF, NotI-HF, or BsaI-HFv2 since these are highly efficient and inexpensive enzymes. In this study, the bacterial backbone contained two sites for SspI and a single site for NotI.
Clone the plasmids as usual [48–50], amplify using a competent *E. coli*, and extract DNA using a suitable kit for the required quantity of plasmid DNA. You can use our plasmid (deposited at Addgene as dCas9-ABE8e-Blasticidin-MC-Plasmid #XXXXXXX) for a bacterial backbone containing convenient restriction enzyme sites around attP and attB, as well as a unique NotI restriction enzyme site and double SspI restriction enzyme sites within the backbone. The sequence is also available in the Supplementary Data (*dCas9-ABE8e-Blasticidin-MC-Plasmid.dna*).

Protocol steps

1. **Recombine the source plasmid.** Prepare the recombination reaction

Table 1 Recombination reaction example for 100ng/μg (10.0 μg) of input of both plasmid DNA at 1.0μg/μl and ΦC31 at 2.6μg/μl in volume **V** = 100 μl

Reagents		
Component	Volume (μl)	
Plasmid DNA, volume X to 10μg	10.0 (**X**)	
Nuclease-free H_2O, to half-volume	40.0 (**V/2 - X**)	
ΦC31, volume Y to 10μg	3.85 (**Y**)	
2X recombination buffer, to half-volume	46.15 (**V/2 - Y**)	
Cycler schedule		
Step	Temperature	Time
Recombination	30°C	12 hours
Pause	4°C	indefinitely
Add 1.0 μl Proteinase K, digest ΦC31	55°C	45–60 minutes
Pause	4°C	indefinitely

Note: Recommended concentration of DNA and ΦC31 is 100–200ng/μl. It is preferable to use more plasmid DNA as a ubiquitous component and the optimal amount of ΦC31 as an expensive component. For example, 200ng/μl of DNA and 100ng/μl of ΦC31 (or 150ng/ul of ΦC31 for smaller plasmids). Incubate at 30°C for at least 4 hours, or preferably overnight. Notably, 30C is the optimal recombination temperature, and the reaction practically stops at 37°C (see Fig. 2a). Be certain not to exceed 30°C.

as listed in Table 1 below, with ΦC31 and 2X buffer being half the reaction volume and plasmid DNA in water the other half. Recombine for at least

4 hours, preferably 12 hours or overnight. It is necessary to add proteinase K and digest ΦC31 after the completion of recombination because both intact and degraded ΦC31 negatively affect the column's efficiency.

2. **Use the spin column kit to clean up the recombined DNA.** Use a spin column purification kit to perform intermediate purification of the recombination reaction output. We used QIAGEN's *QIAquick PCR Purification Kit* [51]. Although the kit advertises use for up to 10kbp DNA, we found that it still performed well and resulted in cleaner DNA at 13.9kbp when compared to other kits. Make sure to perform proteinase digestion of the integrase because this interferes with DNA binding to the purification columns. For better elution yield, heat the spin column per the manufacturer's recommendations. We warmed up the purification columns to 55°C for 3 minutes in a heat block before elution. For maximum elution efficiency, we used ≥60 μl of water. It is not necessary to measure the concentration of the eluted DNA in this step (unless desired). Typical elution yield is approximately 90% for short DNA fragments, and much less for larger plasmid sizes (see the note below). Use multiple columns if needed, based on the column binding capacity.

   Note: We validated the efficiency of the purification kits used for a similar purpose in Oliynyk & Church [31], where we preferred QIAGEN's *QIAquick PCR Purification Kit* (QIAquick) [51]. In this research, we additionally tested the yield of QIAquick and Takara's *NucleoSpin Gel and PCR Cleanup* (NucleoSpin) [52] with a 13.9kbp plasmid. We found that with the input of 10μg of the largest (13.9kbp) plasmid used in this research, we retrieved 7.46μg (74.6%) with NucleoSpin and 6.85μg (68.5%) with QIAquick, using an elution volume of 55μl in both cases. However, we found that when we used both kits for two-stage purification, the yield of NucleoSpin was lower, while the NucleoSpin DNA quality with preheating recommended by the vendor was consistently showing salt carry-over, with 260/230 quality typically in the 1.5–1.8 range, while the production runs of QIAquick exhibited excellent DNA quality. We achieved an even higher efficiency of >85% large DNA retrieval rate using the QIAGEN QIAEX II Gel Extraction Kit; however, the DNA produced by this kit appeared to carry over resin particle pollution, and contamination indicated by an absorption peak at 225 nm with an amplitude exceeding the 260 nm DNA peak. QIAGEN does not recommend using DNA produced by QIAEX II kit for transfection, but recommends using the QIAquick kit instead, and we concur. Thus, we opted to use a QIAquick kit with better mcDNA final yield and quality. The two cleanups required have an estimated combined retrieval of the plasmid and mcDNA of 0.685 x 0.685 = 0.469 (46.9%)—i.e., 53.1% is lost in the column cleanup for this DNA size. This implies that with the maximum reported reaction yield of 33.9% for 13942bp source plasmids—and assuming that recombination reaction

efficiency is close to 100%—the additional pipetting losses and digestion of nicked and broken DNA constitute the remaining 13.0% difference.

3. **Cut the bacterial backbone DNA circle and digest it and all other non-circular or nicked DNA.** Perform the reaction (as listed in Table 2) using a restriction enzyme unique to the bacterial backbone.

Table 2 Example of the digestion/purification recombined DNA; reaction volume $\mathbf{V} = 100\ \mu$l

Reagents		
Component	Volume (μl)	
DNA, volume X as eluted from Step 2	(**X**)	
T5 Exonuclease (NEB M0663L)	4.0	
Restriction Enzyme, e.g. SspI-HF (NEB R3132L)	2.0	
rCutSmart 10x buffer (NEB B6004S)	10.0	
Nuclease-free H_2O (Invitrogen 10977015)	as needed to 100.0 μl	
Cycler schedule		
Step	Temperature	Time
Digestion	37°C	60–90 minutes
Pause	4°C	indefinitely

Note: Perform a digestion test before performing this step for the first time to validate the time required for the complete digestion of byproducts with your DNA concentration and the chosen restriction enzyme cutting efficiency. The easiest method involves taking the quantity of the original plasmid equal to that used in a recombination reaction (much like in the example in this table), and incubating while taking samples for gel electrophoresis from the reaction volume every 20 minutes. When the gel shows complete digestion, purify the remaining reaction volume and verify on a gel that all of the DNA was digested. In our example using SspI or NotI, the complete digestion was achieved within 40 minutes for 10μg of plasmid DNA for all plasmid sizes. The digestion takes longer—but not an excessive amount of time—for higher DNA concentrations. For example, for a 300ng/μl digestion test reaction (30μg in 100μl reaction), almost complete digestion of a 13.9kbp plasmid was achieved in 60 minutes, with no traces of DNA remaining at 90 minutes of digestion (see Supplementary Fig. 4).

4. **Use spin column kit to perform final clean up similar to Step 2 above.** Note that the expected mcDNA yield is approximately 30% of the input DNA, and a lower column capacity will suffice. Adjust the number of used QIAquick spin columns accordingly.

For researchers using this protocol for the first time, we recommend using at least 40μg of source plasmid for the initial experiment and opting for a low

elution volume of 20–25μl. This will allow to become familiar with the protocol, and if necessary, adjust the production scale accordingly.

The spin column purification steps described above are designed for the easy preparation of mcDNA in quantities ranging from 2 to about 30μg. If larger quantities are required, scaling up the number of QIAquick columns may become impractical. In such cases, the cleanup steps can be performed more economically using isopropanol precipitation [32]. We recommend using two 70% ethanol washes following isopropanol precipitation, resulting in excellent mcDNA purity. However, lower DNA extraction efficiency should be expected with methods other than modern spin columns. In a comparison study [53], isopropanol precipitation reported DNA recovery of 52.47 ± 19.69% (see also [54]), centrifugal dialysis 12.58 ± 7.15%, and magnetic bead extraction 9.92 ± 3.89%, all notably lower than the above spin column efficiencies.

Fine-tuning the recombination buffer composition

The mcDNA yield depends on the reaction buffer used, the source plasmid size, and a combination of DNA and ΦC31 proportions and concentrations, recombination reaction temperatures, and recombination time. We explored this multivariate space while focusing on determining the optimal conditions for practical mcDNA yield.

The work began with the validation of the buffers facilitating recombination. We tested the buffers used with ΦC31 in other studies [21, 23, 47], as well as a number of buffers customarily used for restriction enzyme reactions sourced from NEB. We found that all NEB buffers containing Mg ions were not suitable for the recombination reactions, resulting in all DNA being completely digested in our proposed protocol. The buffers containing only NaCl, Tris, and EDTA performed best, resulting in visually complete recombination of the DNA product in 12 hours (see Fig. 1d).

We tested whether customarily used additives improve recombination efficiency. Formulations of the buffers with 1 to 5% glycerol had no noticeable difference in the recombination speed. The final recombination reaction mix ultimately contained 1% or more glycerol from the ΦC31 working buffer. NaCl was most effective in the 100–150mM range, with diminished performance at lower or higher concentrations. We also tested the addition of 100μg/ml BSA and 5 mM spermidine—as used in [21, 47]—and found that while complicating the handling, these additives had a slight negative effect on the recombination efficiency and thus were not used. Therefore, the final reaction concentrations were 20mM Tris, 0.1mM EDTA, 100mM NaCl (used in the form of the 2X buffer described in the Methods), and quantities of ΦC31 and plasmids DNA as required per experiment. This minimal buffer achieved complete recombination of the DNA product shown in Fig. 1d, thus it makes sense that additives beyond this would have no effect at best, and at worst interfere with the reaction.

DNA purification after the completion of recombination

Initially, we expected to develop a single-tube protocol where the recombination would be followed by the cutting of the bacterial backbone using restriction enzymes and digestion with T5 exonuclease within the same reaction tube, similar to what we successfully achieved for the ligation reaction of CVs (see Oliynyk & Church [31]). This would be particularly helpful because DNA purification spin column yields diminish with increasing DNA sizes. However, we found that the presence of ΦC31 integrase consistently reduced the on-target efficiency of the restriction enzymes we used (NotI, SspI, and PsiI). Simultaneously, all of these enzymes exhibited star activity, leading to the deterioration of well-recombined mcDNA. We hypothesized that this behavior may be associated with ΦC31 molecules binding with similar affinity to attP and attB, as well as to attL and attR [24, 25], while normally not effecting the reversal of recombination. It is possible that ΦC31 intermittently binds in other locations, or co-binds in the presence of restriction enzymes targeting specific DNA sequences and modifies their activity. Using the thermal inactivation of ΦC31 at 80°C has not abolished this behavior, and the use of thermolabile proteinase K [55] with subsequent heat inactivation has not completely resolved the issue, with a fraction of partially digested DNA fragments consistently present in Primordium Labs' sequencing. After the above testing, it was determined that we could not dispense with the cleanup step after the completion of recombination. Additionally, the proteinase K digestion of ΦC31 after recombination reaction was necessary before the column cleanup, which was also rendered less effective by ΦC31, with spin columns exhibiting clogging—particularly with higher ΦC31 concentrations. As a result, the protocol included ΦC31 digestion after the completion of recombination using proteinase K, followed by spin column cleanup of the DNA. This was followed by the cutting of the bacterial backbone with a suitable restriction enzyme and the digestion of all non-circular or nicked DNA using T5 exonuclease, followed by final spin column cleanup.

DNA purification was performed using QIAGEN's *QIAquick PCR Purification Kit* (QIAGEN Cat#28106) on a QIAGEN vacuum manifold and *Eppendorf 5425 R* microcentrifuge, per the manufacturer's instructions.

Component concentrations in recombination experiments

As shown in the Results, the mcDNA yield depends on the source plasmid sizes, as well as a combination of DNA and ΦC31 proportions and concentrations, recombination reaction temperatures, and times. A thorough mapping of such multivariate space would take a large number of experiments. Our task was simplified by focusing on finding optimal practical concentrations useful for mcDNA production, which allowed us to rapidly pinpoint the useful reaction proportions. All experiments in Fig. 2 were performed using the Protocol steps described earlier, with the following particulars. The concentrations used were (numbers following components are in ng/μl): Fig. 2a,b: ΦC31 100, DNA

100; Fig. 2c,f: ΦC31 25–200, DNA 100; Fig. 2d,g: ΦC31 100, DNA 100–300; Fig. 2e,h: ΦC31 100, DNA 100; Fig. 2j: ΦC31 180, 130, 130 (in order of columns), DNA 100.

Fig. 2i used all of the same data points displayed in Fig. 2c–h, presenting them as molar ratios of ΦC31 molar concentrations relative to DNA molar concentrations. In this figure, ΦC31 concentration was adjusted to account for 85% purity, as reported by GenScript quality analysis. The molar ratios were calculated based on ΦC31 integrase containing 613 residues and a corresponding molecular mass of 68 kDa. The molecular masses of corresponding plasmids were calculated based on plasmid size [56], and the molar ratios were calculated as multiples of ΦC31 molecules per DNA molecule for corresponding ΦC31 and DNA concentrations.

CRISPR-Cas9 base editing: transfection, editing, handling, and sequencing of edited cells

For all base editing validation experiments, we used plasmid *p-att-ef1a-ABE8e-dCas9-BSD* (see also Plasmid design and assembly further below) and mcDNA produced from this plasmid with our proposed protocol. For gRNA expression, we used the CV method described in Oliynyk & Church [31], with a minimal protocol simplification. Specifically, the NEB Ligation Buffer in the ligation step was substituted by NEB rCutSmart buffer (Cat#B6004), in the identical proportion. To compensate for NEB rCutSmart buffer not containing ATP, the ATP amount in the reaction was doubled to achieve the CV protocol original ATP concentration. This simplified the reagents for the digestion step of the CV protocol, allowing, instead of preparing a four-component digestion buffer, just to add 4μl of T5 Exonuclease and optionally 0.4μl of additional NEB rCutSmart buffer per 100μl of CV ligation reaction, with identical final yield. The CV sequences are available in the Supplementary Data. When needed, mcDNA was further vacuum-concentrated using Thermo Fisher Scientific Savant^{TM} SpeedVac^{TM}.

Both the HEK293T and mESC lines were transfected using a Neon Transfection System (Thermo Fisher Scientific) using 10μl tips (Cat#MPK1096K). Cells were resuspended at 40–70% confluency, centrifuged at 300G, and resuspended in Neon Buffer R to accommodate 40k cells per edit—similar for both cell lines. At this point, the editing steps differed as follows:

- **Editing and selection of HEK293T cells.** We used 1000ng of dCas9-ABE8e mcDNA (size 6833bp). To match molar amounts due to the larger weight of the plasmid (size: 9533bp), we used 1400ng of dCas9-ABE8e plasmid. In both cases, we used 400ng of the gRNA CVs to achieve a 4X molar excess of guide expression vector to a protein expression vector, following [31]. The transfection was performed using a Neon protocol with two pulses at 1,150V and a duration of 20ms [57].

Edited HEK293T cells were seeded on 24-well plates (Corning) in DMEM supplemented with 10% FBS. Then, 14–16 hours after transfection, 10 $ng/\mu l$ of blasticidin (Thermo Fisher Scientific) was used to select cells containing the base editor (based on a recommendation by Xiong et al. [58]). Selection continued for 3 days, noting that the control, non-transfected cells died 48 hours into the selection process. Genomic DNA was extracted 96 hours after the beginning of selection using a Zymo Research Quick-DNA Microprep Kit (Cat#D3020). Genomic DNA processing and sequencing were performed similarly for both cell lines, as described further.

- **Editing and selection of mESCs.** The initial transfection was performed with the same amount of DNA as above for HEK cells, with more cell mortality than expected. Thus, the DNA concentration was decreased by approximately four-fold, using 270ng of dCas9-ABE8e mcDNA or 380ng of plasmid with 100ng of gRNA CV. This concentration resulted in the typical cell survival. The transfection was performed using a Neon protocol with two pulses at 1,200V and a duration of 20ms [59]. Edited mESCs were seeded on 24-well plates (Corning) treated with Matrigel (Corning) in ESGRO-2i Medium (Millipore). For mESCs, we tested editing in duplicate with and without blasticidin antibiotic selection, as presented in the Results. The blasticidin selection was performed 14–16 hours after transfection; however, based on our experience with faster-dividing mESCs, the selection duration was 24 hours, followed by a media change. Genomic DNA was collected a day later in the same manner as for HEK293T cells.
- **Genomic DNA processing and sequencing.** Genomic areas of interest were amplified from 6ng of whole genome DNA using HiFi Hot Start - Readymix (Kapa) at 26–28 cycles using the master mix protocol (see Supplementary Table 1 for the list of the primers). This was followed by DNA separation by electrophoresis on 1% E-Gel EX Agarose Gels (Invitrogen) and the use of a Monarch DNA Gel Extraction Kit (T1020S) with one extra washing step. Sanger sequencing was performed using Azenta Genewiz with our PCR primers, which are listed in Supplementary Table 1.
  The editing efficiency analysis performed using EditR [60, 61] is summarized in Fig. 3 and illustrated in Supplementary Figs. 5–7.

HITI editing: transfection, editing, handling, and sequencing of edited cells

- **Transfection for HITI experiments.** HEK293T cells were seeded into 24-well plates at a density of 120K per well 20-24 hours before transfection. At approximately 70% confluency, cells were transfected using Lipofectamine 3000 (Thermo Fisher Scientific) following the manufacturer's instructions. Briefly, 250ng of Cas9 plasmid (pCAG-1BPNLS-Cas9-1BPNLS, Addgene #87108), 200ng of gRNA CVs, and 300ng of HITI mcDNA template were mixed in 50μl of Opti-MEM with 1μl of 3000 Reagent and 1.5μl of Lipofectamine 3000 Reagent per well. 48 hours later, cells were harvested by removing the media, washed one time with PBS, and then trypsinized

using TrypLE Express (Thermo Fisher Scientific, Cat#12604013) for 3 minutes. Resuspended cells were pelleted via centrifugation for 5 minutes at 500g. After centrifugation, cells were resuspended in media containing hygromycin (Thermo Fisher Scientific, Cat#10687010) at a concentration of 50μg/ml and seeded in 6-well plates. Cells were allowed to grow for 3 days, then the hygromycin concentration was increased to 200μg/ml. Continuous hygromycin treatment was carried out for 14 days until all unedited cells were eradicated, as measured by observing a control well with untransfected cells and treatment with the same hygromycin concentration.

- **Flow cytometry analysis.** Hygromycin-selected cells and untransfected control cells were harvested as described above. After centrifugation, cell pellets were resuspended into PBS with 2% FBS and 1μg/ml of 4',6-diamidino-2-phenylindol (DAPI) (Thermo Fisher Scientific, Cat#D1306) and transferred into 5 ml round bottom polystyrene flow tubes with cell strainer (Corning Life Sciences, 352235). Cells were analyzed by flow cytometry (Symphony A1, BD Biosciences) and data was processed with FlowJoTM 10 software. An overview of gating and analysis is given in Supplementary Fig. 8.
- **Genomic DNA extraction and PCR analysis.** To confirm the correct integration of HITI templates, genomic DNA was extracted from HITI-treated cells and untransfected cells (as negative controls) using a Zymo Research Quick-DNA Microprep Kit (Cat#D3020). Genomic DNA processing and sequencing were performed as described above with slight modifications. ACTB or GAPDH-specific forward primers were used with a common mKate reverse primer to amplify the genomic DNA region spanning the junction of the HITI-mediated integration. Q5 Hot Start High-Fidelity 2X Master Mix (NEB., Cat#M0494) was used to amplify the genomic region using 30 cycles of amplification with 20ng of total genomic DNA per reaction. PCR amplicons were cleaned up using QIAGEN's QIAquick PCR Purification Kit (QIAquick) and separated by electrophoresis on 1% E-Gel EX Agarose Gels (Invitrogen). For sequencing, the Monarch DNA Gel Extraction Kit (T1020S) was used to purify the excised DNA band from the gel and sent for Sanger sequencing. Sanger sequencing was performed using Azenta Genewiz with a primer specific to mKate insert. All primers are listed in Supplementary Table 1.

Reagents and materials

ΦC31 protein was subcontracted from GenScript (see the protein amino acid sequence in Supplementary Note 9). We received 18.96mg of protein at the cost of $4932.71 with the addition of optional nucleic acid & nuclease removal at $1125. The product concentration was 4.74mg/ml(by Bradford) A260/A280: 0.557, ≥85% purity, endotoxin level 0.338 EU/mg (or equivalently 0.000338 EU/μg - effectively endotoxin free). The GenScript ΦC31 storage buffer composition was: 50mM Tris-HCl, 500mM NaCl, 10% Glycerol, pH 8.0. GenScript conducted endotoxin level testing on source plasmids and mcDNA.

Table 3 Reagents and resources

REAGENT or RESOURCE	SOURCE	IDENTIFIER
SspI-HF	New England BioLabs	Cat#R3132L
NotI-HF	New England BioLabs	Cat#R3189L
T5 Exonuclease	New England BioLabs	Cat#M0663L
Thermolabile Proteinase K	New England BioLabs	Cat#P8111S
Proteinase K	New England BioLabs	Cat#P8107S
TrypLE	Thermo Fisher Scientific	Cat#12605010
Puromycin Dihydrochloride	Thermo Fisher Scientific	Cat#A1113803
Blasticidin S HCl	Thermo Fisher Scientific	Cat#A1113903
Penicillin-Streptomycin	Thermo Fisher Scientific	Cat#15070063
Gibco DMEM, high glucose, pyruvate	Thermo Fisher Scientific	Cat#11995040
ESGRO-2i Medium	Millipore	Cat#SF016-100
Embryonic Stem Cell FBS	Thermo Fisher Scientific	Cat#16141002
Nuclease-free H2O	Invitrogen	Cat#10977015
HiFi Hot Start - Readymix	KAPA Biosystems	Cat#KK2602
Q5 High-Fidelity Master Mix	New England BioLabs	Cat#M0492L
NEBNext Ultra II Q5 Master Mix	New England BioLabs	Cat#M0544L
QIAquick PCR Purification Kit	QIAGEN	Cat#28106
Sodium Acetate Solution (3 M), pH 5.2	Thermo Fisher Scientific	Cat#R1181
Monarch PCR & DNA CleanUp Kit	New England BioLabs	Cat#T1030S
Monarch DNA Gel Extraction Kit	New England BioLabs	Cat#T1020S
NucleoSpin Gel and PCR Cleanup	Macherey-Nagel	Cat#74609.50
QIAGEN Plasmid Plus Midi Kit	QIAGEN	Cat#12943
QIAGEN EndoFree Plasmid Maxi Kit	QIAGEN	Cat#12362
Invitrogen 1% E-GELS EX	Invitrogen Corporation	Cat#G402021
NEB 5-alpha F'Iq Competent *E. coli*	New England BioLabs	Cat#C2992H
HEK293T	ATCC	Cat#CRL-3216
ESGRO Adapted C57/BL6 Mouse ESC Line	Millipore	Cat#SF-CMTI-2
p-dCas9-SSAP-MS2-BB	Addgene	Cat#183826
NEBuilder HiFi DNA Assembly Master Mix	New England BioLabs	Cat#E2621
Tris-EDTA, pH 8.0	Fisher Scientific	Cat#BP2473
1M Tris-HCl, pH 8.0	Corning	Cat#46-031-CM
Matrigel	Corning	Cat#354277
Lipofectamine 3000	Thermo Fisher Scientific	Cat#L3000015
Opti-MEM™ Reduced Serum Medium	Thermo Fisher Scientific	Cat#31985062
PBS, pH 7.4	Thermo Fisher Scientific	Cat#10010023
Hygromycin B (50 mg/mL)	Thermo Fisher Scientific	Cat#10687010
FBS, Premium, heat-inactivated	Thermo Fisher Scientific	Cat#A5670801
DAPI	Thermo Fisher Scientific	Cat#D1306
TrypLE Express	Thermo Fisher Scientific	Cat#12604013
Q5 Hot Start High-Fidelity 2X Master Mix	New England BioLabs	Cat#M0494
5 ml round bottom polystyrene flow tubes	Corning Life Sciences	Cat#352235
pCAG-1BPNLS-Cas9-1BPNLS	Addgene	Plasmid#87108
Quick-DNA Microprep Kit	Zymo Research	Cat#D3020

Plasmid design and assembly

We used *p-dCas9-SSAP-MS2-BB* (Addgene Cat#183826) [62] as a convenient source of bacterial backbone and dCas9 for Gibson assembly. The insert of the BSD gene is based on our earlier publication [31], while the 10482bp plasmid/mcDNA (named *p-att-ef1a-MS2-RecT-dCas9-BSD*) is used in our other ongoing project, and our dCas9-ABE8e plasmid (named *p-att-ef1a-ABE8e-dCas9-BSD*) is a modification of this plasmid with TADa for ABE base editing derived from [11] and cloned using Gibson assembly NEBuilder HiFi DNA Assembly Master Mix (NEB Cat#E2621) and amplified using NEB 5-alpha F'Iq Competent *E. coli* (NEB Cat#C2992H). The necessary dsDNA fragments were ordered from Twist Bioscience.

The plasmid *p-att-ef1a-MS2-RecT-dCas9-BSD*, 10482bp in length, is a functional plasmid that was used as a basis for the recombination experiments (plasmid available at Addgene). Two modifications of this plasmid—a truncation to create a 5758bp plasmid *p-att-ef1a-MS2-RecT-dCas9-BSD-truncated* and an extension to create a 13942bp plasmid *p-att-ef1a-MS2-RecT-dCas9-BSD-extended*—were used for recombination experiments but are not intended to be functional for genome editing. The sequences for all aforementioned plasmids are available in the *Plasmids* folder in the Supplementary Data.

Statistics and reproducibility

All protocol reactions were performed in independent reaction duplicates (n=2) and yield percentage values were averaged.

The editing efficiency analysis was performed from Sanger sequencing using EditR [60, 61] from independent samples in triplicates (n=3) or quadruplicates (n=4). The P values for editing efficiency calculated by EditR are listed in the Supplementary Data. The editing efficiency values were averaged.

Data availability

The raw data for the results presented in this study can be found in a single SupplementaryData.ZIP file, composed of the following items:

1. *Recombination Yields Table.xlsx.*
2. *Base Editing Efficiency Sanger-EditR Results.xlsx.*
3. Base Editing plasmid, mcDNA and gDNA CV sequences.
4. HITI plasmid, mcDNA and gDNA CV sequences.
5. Sequenced PrimordiumLabs Recombined MC DNA.

Acknowledgments

A.M. was supported by the Ibn Rushd Postdoctoral Fellowship from King Abdullah University of Science and Technology. E.K. is funded through an SRA with Colossal Biosciences.

Author contributions

R.O. designed this research and performed Plasmid2MC and base editing experiments; R.O. and A.M. collaborated on conceptual foundations; E.K. defined mouse genomic targets for mESC base editing; A.M. conceived and performed HITI experiments and analysis; R.O., A.M. and G.C. analyzed results and wrote the manuscript.

Competing interests

GMC disclosures: https://arep.med.harvard.edu/gmc/tech.html

References

[1] Lederberg, J. Cell genetics and hereditary symbiosis. *Physiological reviews* **32** (4), 403–430 (1952). https://doi.org/10.1152/physrev.1952.32.4.403 .

[2] Datta, N. Transmissible drug resistance in an epidemic strain of salmonella typhimurium. *Epidemiology & Infection* **60** (3), 301–310 (1962). https://doi.org/10.1017/s0022172400020416 .

[3] Watanabe, T., Ogata, C. & Sato, S. Episome-mediated transfer of drug resistance in enterobacteriaceae viii: Six-drug-resistance r factor. *Journal of bacteriology* **88** (4), 922–928 (1964). https://doi.org/10.1128/jb.88.4.922-928.1964 .

[4] Cohen, S. N., Chang, A. C., Boyer, H. W. & Helling, R. B. Construction of biologically functional bacterial plasmids in vitro. *Proceedings of the National Academy of Sciences* **70** (11), 3240–3244 (1973). https://doi.org/10.1073/pnas.70.11.3240 .

[5] Byrne, S. M., Mali, P. & Church, G. M. in *Chapter six - genome editing in human stem cells* (eds Doudna, J. A. & Sontheimer, E. J.) *The Use of CRISPR/Cas9, ZFNs, and TALENs in Generating Site-Specific Genome Alterations*, Vol. 546 of *Methods in Enzymology* 119–138 (Academic Press, Cambridge, 2014). URL https://doi.org/10.1016/B978-0-12-801185-0.00006-4.

[6] Doghaither, H. A. & Gull, M. in *Plasmids as genetic tools and their applications in ecology and evolution* (ed.Gull, M.) *Plasmid* Ch. 5 (IntechOpen, Rijeka, 2019). URL https://doi.org/10.5772/intechopen.85705.

[7] Fajrial, A. K., He, Q. Q., Wirusanti, N. I., Slansky, J. E. & Ding, X. A review of emerging physical transfection methods for crispr/cas9-mediated gene editing. *Theranostics* **10** (12), 5532–5549 (2020). https://doi.org/10.7150/thno.43465 .

[8] Gaj, T., Sirk, S. J., Shui, S.-l. & Liu, J. Genome-editing technologies: principles and applications. *Cold Spring Harbor perspectives in biology* **8** (12), a023754 (2016). https://doi.org/10.5772/intechopen.85705 .

[9] Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A. & Liu, D. R. Programmable editing of a target base in genomic dna without double-stranded dna cleavage. *Nature* **533** (7603), 420–424 (2016). https:

//doi.org/10.1038/nature17946 .

[10] Anzalone, A. V. *et al.* Search-and-replace genome editing without double-strand breaks or donor dna. *Nature* **576** (7785), 149–157 (2019). https://doi.org/10.1038/s41586-019-1711-4 .

[11] Richter, M. F. *et al.* Phage-assisted evolution of an adenine base editor with improved cas domain compatibility and activity. *Nature Biotechnology* **38** (7), 883–891 (2020). https://doi.org/10.1038/s41587-020-0453-z .

[12] Chen, P. J. *et al.* Enhanced prime editing systems by manipulating cellular determinants of editing outcomes. *Cell* **184** (22), 5635–5652.e29 (2021). https://doi.org/10.1016/j.cell.2021.09.018 .

[13] Darquet, A., Cameron, B., Wils, P., Scherman, D. & Crouzet, J. A new dna vehicle for nonviral gene delivery: supercoiled minicircle. *Gene therapy* **4** (12), 1341–1349 (1997). https://doi.org/10.1038/sj.gt.3300540 .

[14] Darquet, A.-M. *et al.* Minicircle: an improved dna molecule for in vitro and in vivo gene transfer. *Gene Therapy* **6** (2), 209–218 (1999). https://doi.org/10.1038/sj.gt.3300816 .

[15] Eusébio, D. *et al.* The performance of minicircle dna versus parental plasmid in p53 gene delivery into hpv-18-infected cervical cancer cells. *Nucleic Acid Therapeutics* **31** (1), 82–91 (2021). https://doi.org/10.1089/nat.2020.0904 .

[16] Maniar, L. E. G. *et al.* Minicircle dna vectors achieve sustained expression reflected by active chromatin and transcriptional level. *Molecular Therapy* **21** (1), 131–138 (2013). https://doi.org/10.1038/mt.2012.244 .

[17] Munye, M. M. *et al.* Minicircle dna provides enhanced and prolonged transgene expression following airway gene transfer. *Scientific reports* **6** (1), 23125 (2016). https://doi.org/10.1038/srep23125 .

[18] Almeida, A. M., Queiroz, J. A., Sousa, F. & Sousa, Ã. Minicircle dna: The future for dna-based vectors? *Trends in Biotechnology* **38** (10), 1047–1051 (2020). https://doi.org/10.1016/j.tibtech.2020.04.008 .

[19] Kay, M. A., He, C.-Y. & Chen, Z.-Y. A robust system for production of minicircle dna vectors. *Nature biotechnology* **28** (12), 1287–1289 (2010). https://doi.org/10.1038/nbt.1708 .

[20] Gaspar, V. M. *et al.* Improved minicircle dna biosynthesis for gene therapy applications. *Human Gene Therapy Methods* **25** (2), 93–105 (2014). https://doi.org/10.1089/hgtb.2013.020 .

[21] Olorunniji, F. J. *et al.* Gated rotation mechanism of site-specific recombination by ϕc31 integrase. *Proceedings of the National Academy of Sciences* **109** (48), 19661–19666 (2012). https://doi.org/10.1073/pnas.1210964109 .

[22] System Biosciences. *MC-Easy™ Minicircle DNA Production Kit (accessed May 3, 2024)* (2024). URL https://www.systembio.com/mc-easy-minicircle-dna-production-kit-with-competent-cells.

[23] Thorpe, H. M. & Smith, M. C. In vitro site-specific integration of bacteriophage dna catalyzed by a recombinase of the resolvase/invertase family. *Proceedings of the National Academy of Sciences* **95** (10), 5505–5510 (1998). https://doi.org/10.1073/pnas.95.10.5505 .

[24] Thorpe, H. M., Wilson, S. E. & Smith, M. C. Control of directionality in the site-specific recombination system of the streptomyces phage φc31. *Molecular microbiology* **38** (2), 232–241 (2000). https://doi.org/10.1046/j.1365-2958.2000.02142.x .

[25] Smith, M. C. & Thorpe, H. M. Diversity in the serine recombinases. *Molecular microbiology* **44** (2), 299–307 (2002). https://doi.org/10.1046/j.1365-2958.2002.02891.x .

[26] Brown, W. R., Lee, N. C., Xu, Z. & Smith, M. C. Serine recombinases as tools for genome engineering. *Methods* **53** (4), 372–379 (2011). https://doi.org/10.1016/j.ymeth.2010.12.031 .

[27] Smith, M. C. Phage-encoded serine integrases and other large serine recombinases. *Mobile DNA iii* 253–272 (2015). https://doi.org/10.1128/microbiolspec.MDNA3-0059-2014 .

[28] Stark, W. M. Making serine integrases work for us. *Current Opinion in Microbiology* **38**, 130–136 (2017). https://doi.org/10.1016/j.mib.2017.04.006 .

[29] Durrant, M. G. *et al.* Systematic discovery of recombinases for efficient integration of large dna sequences into the human genome. *Nature Biotechnology* **41** (4), 488–499 (2023). https://doi.org/10.1038/s41587-022-01494-w .

[30] Trejo, C. S., Rock, R. S., Stark, W. M., Boocock, M. R. & Rice, P. A. Snapshots of a molecular swivel in action. *Nucleic Acids Research* **46** (10), 5286–5296 (2018). https://doi.org/10.1093/nar/gkx1309 .

[31] Oliynyk, R. T. & Church, G. M. Efficient modification and preparation of circular dna for expression in cell culture. *Communications Biology* **5** (1), 1393 (2022). https://doi.org/10.1038/s42003-022-04363-z .

[32] QIAGEN. *Isopropanol precipitation of DNA (accessed May 10, 2024)* (2024). URL https://www.qiagen.com/us/knowledge-and-support/knowledge-hub/bench-guide/dna/handling-dna/isopropanol-precipitation-of-dna.

[33] Suzuki, K. *et al.* In vivo genome editing via crispr/cas9 mediated homology-independent targeted integration. *Nature* **540** (7631), 144–149 (2016). https://doi.org/10.1038/nature20565 .

[34] Suzuki, K. & Izpisua Belmonte, J. C. In vivo genome editing via the hiti method as a tool for gene therapy. *Journal of Human Genetics* **63** (2), 157–164 (2018). https://doi.org/10.1038/s10038-017-0352-4 .

[35] Petsch, D., Deckwer, W.-D. & Anspach, F. B. Proteinase k digestion of proteins improves detection of bacterial endotoxins by thelimulusamebocyte lysate assay: Application for endotoxin removal from cationic proteins. *Analytical biochemistry* **259** (1), 42–47 (1998). https://doi.org/10.1006/abio.1998.2655 .

[36] Fux, A. C. *et al.* Heterogeneity of lipopolysaccharide as source of variability in bioassays and lps-binding proteins as remedy. *International Journal of Molecular Sciences* **24** (9) (2023). https://doi.org/10.3390/ijms24098395 .

[37] Twist Bioscience. *Twist Gene Fragments (accessed May 3, 2024)* (2024). URL https://www.twistbioscience.com/products/genes/gene-synthesis?tab=fragment.

[38] Thermo Fisher Scientific. *Evaluation of linear synthetic DNA fragments from separate suppliers (accessed May 3, 2024)* (2024). URL https://assets.thermofisher.com/TFS-Assets/LSG/brochures/geneart-strings-synthetic-dna-white-paper.pdf.

[39] Integrated DNA Technologies, Inc. *gBlocks Gene Fragments (accessed May 3, 2024)* (2024). URL https://www.idtdna.com/pages/products/genes-and-gene-fragments/double-stranded-dna-fragments/gblocks-gene-fragments.

[40] Dugaiczyk, A., Boyer, H. W. & Goodman, H. M. Ligation of ecori endonuclease-generated dna fragments into linear and circular structures. *Journal of Molecular Biology* **96** (1), 171–184 (1975). https://doi.org/https://doi.org/10.1016/0022-2836(75)90189-8 .

[41] Panoukidou, M. *et al.* Runaway transition in irreversible polymer condensation with cyclization. *Phys. Rev. Res.* **6**, 023189 (2024). https://doi.org/10.1103/PhysRevResearch.6.023189 .

[42] Fijalkowska, I. J., Schaaper, R. M. & Jonczyk, P. Dna replication fidelity in escherichia coli: a multi-dna polymerase affair. *FEMS microbiology reviews* **36** (6), 1105–1121 (2012). https://doi.org/10.1111/j.1574-6976.2012.00338.x .

[43] Tomoiaga, D., Bubnell, J., Herndon, L. & Feinstein, P. High rates of plasmid cotransformation in e. coli overturn the clonality myth and reveal colony development. *Scientific Reports* **12** (1), 11515 (2022). https://doi.org/10.1038/s41598-022-14598-9 .

[44] Hou, X. H., Guo, X. Y., Chen, Y., He, C.-Y. & Chen, Z.-Y. Increasing the minicircle dna purity using an enhanced triplex dna technology to eliminate dna contaminants. *Molecular Therapy Methods & Clinical Development* **2** (2015). https://doi.org/10.1038/mtm.2014.62 .

[45] Alves, C. P., Šimčíková, M., Brito, L., Monteiro, G. A. & Prazeres, D. M. F. Development of a nicking endonuclease-assisted method for the purification of minicircles. *Journal of Chromatography A* **1443**, 136–144 (2016). https://doi.org/10.1016/j.chroma.2016.03.035 .

[46] Almeida, A., Queiroz, J., Sousa, F. & Sousa, A. Minicircle dna purification: Performance of chromatographic monoliths bearing lysine and cadaverine ligands. *Journal of Chromatography B* **1118-1119**, 7–16 (2019). https://doi.org/10.1016/j.jchromb.2019.04.024 .

[47] Abioye, J. *et al.* High fidelity one-pot dna assembly using orthogonal serine integrases. *Biotechnology Journal* **18** (3), 2200411 (2023). https://doi.org/10.1002/biot.202200411 .

[48] Addgene. *Gibson Assembly Cloning (accessed May 9, 2024)* (2024). URL https://www.addgene.org/protocols/gibson-assembly/.

[49] Addgene. *Bacterial Transformation (accessed May 9, 2024)* (2024). URL https://www.addgene.org/protocols/bacterial-transformation/.

[50] Addgene. *Recovering Plasmid DNA from Bacterial Culture (accessed May 9, 2024)* (2024). URL https://www.addgene.org/protocols/purify-plasmid-dna/.

[51] QIAGEN. *QIAquick PCR Purification Kit (accessed May 9, 2024)* (2024). URL https://www.qiagen.com/us/products/discovery-and-translational-research/dna-rna-purification/dna-purification/dna-clean-up/qiaquick-pcr-purification-kit/.

[52] Takara Bio. *NucleoSpin Gel and PCR Clean-Up User Manual (accessed May 9, 2024)* (2024). URL https://www.takarabio.com/documents/User%20Manual/NucleoSpin%20Gel%20and%20PCR%

20Clean/NucleoSpin%20Gel%20and%20PCR%20Clean-up%20User%20Manual_Rev_04.pdf.

[53] Kirtane, A. & Deiner, K. Improving the recovery for dissolved edna state: A comparative analysis of isopropanol precipitation, magnetic bead extraction, and centrifugal dialysis. *Environmental DNA* **6** (3), e572 (2024). https://doi.org/10.1002/edn3.572 .

[54] He, S. *et al.* Dna precipitation revisited: A quantitative analysis. *Nano Select* **3** (3), 617–626 (2022). https://doi.org/10.1002/nano.202100152 .

[55] New England BioLabs. *Monarch PCR & DNA Cleanup Kit (accessed May 9, 2024)* (2024). URL https://www.neb.com/en-us/products/p8111-thermolabile-proteinase-k.

[56] Thermo Fisher Scientific. *DNA and RNA Molecular Weights and Conversions (accessed May 12, 2024)* (2024). URL https://www.thermofisher.com/us/en/home/references/ambion-tech-support/rna-tools-and-calculators/dna-and-rna-molecular-weights-and-conversions.html.

[57] Thermo Fisher Scientific. *HEK293, Neon Cell Protocols (accessed May 1, 2024)* (2023). URL https://www.thermofisher.com/content/dam/LifeTech/migration/en/filelibrary/cell-culture/neon-protocols.par.61129.file.dat/hek%20293-kidney.pdf.

[58] Xiong, K. *et al.* An optimized genome-wide, virus-free crispr screen for mammalian cells. *Cell Reports Methods* **1** (4), 100062 (2021). https://doi.org/10.1016/j.crmeth.2021.100062 .

[59] Thermo Fisher Scientific. *Mouse Embryonic Stem Cells, Neon Cell Protocols (accessed May 1, 2024)* (2023). URL https://www.thermofisher.com/content/dam/LifeTech/migration/en/filelibrary/cell-culture/neon-protocols.par.27415.file.dat/mouse%20embryonic%20stem%20cells-blastocyst.pdf.

[60] Kluesner, M. G. *et al.* EditR: A method to quantify base editing from sanger sequencing. *The CRISPR Journal* **1** (3), 239–250 (2018). https://doi.org/10.1089/crispr.2018.0014 .

[61] Moriarity Lab - Medical School - University of Minnesota. *EditR (accessed May 14, 2024)* (2024). URL https://moriaritylab.shinyapps.io/editr_v10/.

[62] Wang, C. *et al.* dCas9-based gene editing for cleavage-free genomic knock-in of long sequences. *Nature cell biology* **24** (2), 268–278 (2022). https://doi.org/10.1038/s41556-021-00836-1 .