Efficient Modification and Preparation of Circular DNA for Expression in Cell Culture

Roman Teo Oliynyk[1,2*] 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 one of the essential tools for delivery and expression of RNAs and proteins in cell culture experiments. The preparation of plasmids typically involves a laborious process of bacterial cloning, validation, and purification. While the expression plasmids can be designed and ordered from the contract manufacturers, the cost may be prohibitive when a large number of plasmids is required. We have developed an efficient fully synthetic method and protocol that enables the production of circularized DNA containing expression elements ready for transfection in as little as 3 hours, thereby eliminating the bacterial cloning steps. The protocol describes how to take a linear double-stranded DNA fragment designed with commonly used software and ordered from one of the commercial manufacturers as an input and then efficiently circularizing and purifying this DNA fragment with minimal hands-on time. As proof of the principle, we demonstrated the functionality of this method by producing DNA expressing extended prime editing guide RNA (epegRNA) for cell culture genome editing. The method is not limited to gene editing and can be used in a variety of applications requiring the expression of short RNAs and proteins. The speed, low cost, and ease of use will make this method another valuable tool in the gene editing toolkit, and due to reaction simplicity, the protocol can be easily automated.

Keywords: Expression Vector, Prime Editing, Base Editing, CRISPR, Bacterial Cloning, Gene Editing

Introduction

Many techniques exist for the preparation and delivery of RNAs and proteins for transient expression in cell cultures [1]. In this study, we present a new and efficient preparation method. As a demonstration, we will describe the use of our method in the gene editing application context. However, since the principle behind this method is generic, this method can be useful for a variety of purposes and applications that require the expression of RNAs and proteins. Although we will discuss the specific implementation for gene editing and compare our method with some existing preparation techniques, the goal of this paper is to present a new approach to preparing the expression vectors in a laboratory environment—and not a more efficient gene editing method as such.

The three most customary approaches that facilitate gene editing when delivered into cells are: (**1**) single or multiple plasmids [2] encoding Cas9 and other supplemental proteins and implementations of guide RNA suitable for traditional CRISPR/Cas9, base editing, prime editing or other methods [3–6]; (**2**) Cas9 mRNA and guide RNA [7]; (**3**) Cas9 protein complexed with guide RNAs [8]. A variety of delivery methods from the aforementioned gene editing constructs are currently being used. The choice of delivery method will depend on the specific requirements and the researcher's preferences, which can include electroporation [9, 10], lipofection [11], direct physical transfection [12], vector delivery on polymer particles and microcarriers [13], and modalities of the aforementioned methods.

The approach (1), gene editing using plasmid transfection for transient expression is commonly used due to its conceptual simplicity, ease of handling, and the stability of plasmids. However, the preparation of plasmids is labor intensive and time consuming [2]. While plasmids can be designed and ordered from contract manufacturers, the cost may be prohibitive, which compels researchers to prepare plasmids themselves. Using as an example a CRISPR/Cas9 or prime editing RNA guide expression vector, the process begins with the initial selection of a suitable plasmid, which may involve ordering it from *Addgene.org* and then modifying it to ligate into the plasmid the guide RNA coding inserts. For each planned genome edit, a researcher must design a specific DNA insert and order it as a DNA fragment from a commercial manufacturer. Then follows the time and labor-intensive process of ligating and cloning this new plasmid. This begins with the ligation of the double-stranded DNA (dsDNA) inserts into a selected plasmid vector using a cloning protocol of choice (e.g., Golden Gate assembly [14] or Gibson assembly [15, 16]). The ligation product is then used in the bacterial transformation protocol using a

competent *E. coli* [17]. Then, the transformed *E. coli* is plated onto agar plates overnight. The colonies are then sequenced and bacteria are grown in liquid culture overnight followed by the recovery of plasmid DNA from bacterial culture [18]. The plasmid concentration is then measured and validated via gel electrophoresis, PCR, or sequencing. At this point, the plasmid expressing the RNA targeting elements is ready for transfection, typically in conjunction with the Cas9-expressing plasmid—unless an all-in-one expression plasmid is used.

We developed an efficient method to circularize short (450 to 950 bp) linear dsDNA fragments, which could be also used with lower efficiency for somewhat longer fragments. We applied this method to create circular DNA for expressing the extended prime editing guide RNA (epegRNA), which includes the U6 promoter followed by coding for the expressed RNA intended for prime editing in cell culture [5, 19, 20].

For brevity, consistency, and to avoid confusion with the description of preparation steps and final product in this paper, we will refer to this circular DNA expression vector construct and method as a *Circular Vector*, capitalized and in italics. Circularized DNA is resistant to exonuclease degradation in the cytoplasm [21]. However, linear DNA lacks such stability, which is a significant barrier to utilizing it for gene delivery. Moreover, our method takes advantage of this difference in properties by using T5 exonuclease to digest unreacted or misreacted linear DNA fragments, thereby purifying the *Circular Vector* DNA.

This *Circular Vector* method can be equally useful for traditional CRISPR/Cas9 and base editing [3, 4]. Also, it can generally be applied to a wide variety of short expression vectors that are customarily delivered by plasmids.

With this method, which consists of a single tube two-step reaction followed by a typical spin column-based DNA cleanup, a researcher can produce a batch of *Circular Vectors* in 3 hours. The *Circular Vector* method simplifies making small variable targeting components that work in conjunction with an invariable large plasmid coding for Cas9 and/or other proteins and genes. This large plasmid can be cloned in large quantity in bacterial culture, extracted with one of the commercially available maxi-preps, and used for subsequent months or even years of experiments. Notably, the variable component for each editing target—the *Circular Vector*—is now a breeze to make.

For a 450-bp *Circular Vector* length, the efficiency of converting the input DNA into *Circular Vectors* was 48% at 1 hour of ligation and up to 62% at 12 hours of ligation. For a 950-bp length, this efficiency was 26% at 1 hour of ligation and 30% at 12 hours of ligation, with intermediate values for *Circular Vectors* between 450 and 950 bp in length. Longer reaction times corresponded with some improvement in *Circular Vector* yield. However, a short 1-hour ligation step (as will be described in this paper) has the advantage of rapid and efficient 3-hour processing. To demonstrate that such circular DNA constructs can work as expression vectors, we successfully edited three genomic locations that were showcased by Chen et al. [19], resulting in comparable editing efficiency.

Methods

Conceptual Introduction

Our circularization method for the CRISPR guide expression vectors requires an initial DNA sequence design, which can be created using the researcher's preferred tools, followed by ordering dsDNA fragments from a commercial manufacturer.

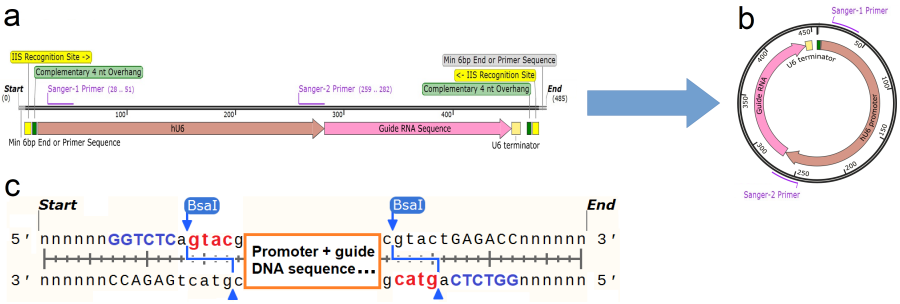

Fig. 1 Circularization of dsDNA containing an expression element. (**a**) Expression element containing type IIS restriction enzyme recognition sites with complementary overhangs on both ends of the DNA fragment. A detailed diagram is presented in Figure A1. (**b**) Cuts by the restriction enzyme and ligation of the resulting overhangs form the circular DNA expression element. T5 exonuclease is used to digest all non-circularized DNA. (**c**) Schematic view of complementary 4nt overhangs (**gtac**) created by restriction enzyme cuts (BsaI **GGTCTC**), see Figure A1 for a detailed example of prime editing epegRNA coding.

The prefabricated dsDNA should—at a minimum—contain a promoter (U6, in our example), followed by a DNA segment coding for epegRNA with a terminator [22] (see Figure 1a and the detailed map in Figure A1). The epegRNA can be designed using a published tool suited to this purpose [19, 20, 23]. Two matching type IIS restriction enzyme recognition sites create complementary 4 nt overhangs near both ends of the dsDNA fragment. The *Circular Vector* can be conveniently validated with Sanger sequencing (see Figure 1b). Placing two primers, Sanger-1 and Sanger-2, near the opposite locations of the circular structure has proven to be efficient and reliable, covering most of the circular structure with the overlap that includes the opposing primer near the middle of the FASTA sequence.

The matching overhangs created by the restriction enzyme cuts are designed to provide a high-fidelity ligation [24], which is also aided by the absence of any other overhangs that would be present in a typical Golden Gate assembly (see Figure 1c). The 'nnnnnn' sequences on Figure 1c represent the minimum of six base pairs required on the back end of the restriction enzyme. Some manufacturers (e.g., Twist Bioscience) prefer to deliver their DNA fragments with end adapters that serve as well-designed PCR amplification primers and simultaneously eliminate the need to incorporate end sequences.

Step-by-Step Protocol

We aimed to make the method of *Circular Vector* production user-friendly and rapid. Thus, we primarily used reaction enzymes and buffers based on commercial kits, with only a few additional reagents. If desired, the kits may be substituted with individual reagents. If a sufficient amount of dsDNA is received from a commercial manufacturer, it may be used directly in the reaction; otherwise, it may need to be amplified by PCR to obtain the desired quantity.

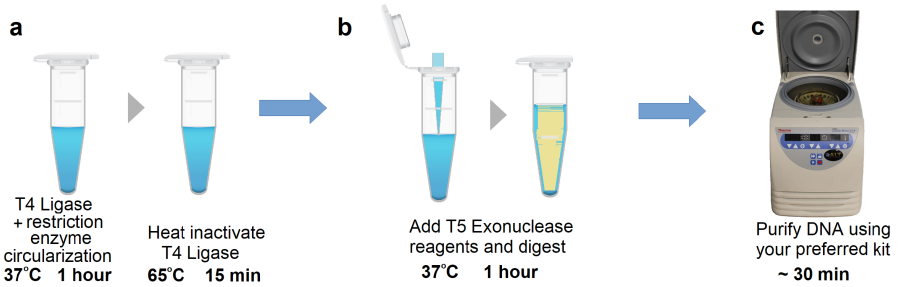

Fig. 2 Protocol steps: (**a**) Pipette dsDNA template with T4 DNA ligase and type IIS restriction enzyme reagents, then program incubation at 37°C (or longer, if desired), followed by the heat activation of T4 DNA ligase at 65°C for 15 minutes. If left unattended, set the program cycler to pause at 4°C. (**b**) Add an equal volume of T5 exonuclease reagents to the same tube and digest at 37°C for 1 hour to 1.5 hours. (**c**) Purify the circularized DNA expression construct from the reaction mix using a DNA cleanup kit.

The protocol for the preparation of the circularized DNA vectors consists of three steps:

Step 1: Circularization of the source DNA (Figure 2a).

Mix the reaction reagents listed in Table 1 by pipetting, preferably on ice, into a 200 μl PCR tube or another suitable tube with at least double the Table 1 reaction volume. The extra volume will be required for the addition of the digestion reagents in Step 2. Place the tube in a thermal cycler and run the reaction as described in Table 2.

Table 1 Circularization reaction ingredients for 6.0 μg of input dsDNA in 50 μl.

Component	Volume (μl)
dsDNA, concentration $\geq$ 180 ng/μl	volume based on concentration, up to 6.0 μg
T4 DNA Ligase Buffer (10X B0202S)	5.0
T4 DNA Ligase (NEB M0202L)	3.0
Bsai-HF2 (NEB R3733L)	3.0
ATP (10mM NEB P0756S)	5.0
Nuclease-free H_2O (Invitrogen 10977015)	as needed to 50.0 μl

A smaller quantity of DNA does not change the reaction efficiency and may be used if, for example, the DNA concentration is lower. The reaction may be also scaled proportionately if a larger yield is required.

Table 2 Circularization step: Temperature timing

Step	Temperature	Time
Circularization	37^oC	60 min or longer, if desired
T4 DNA ligase inactivation	65^oC	15 min
Pause	4^oC	indefinitely

Step 2: Digesting all remaining linear DNA (Figure 2b).

Prepare the T5 exonuclease reaction mix as listed in Table 3 and mix it in 1:1 proportion into the tube containing the reaction mix from the previous step. For example, if a 50 μl reaction was performed in Step 1, 50 μl will need to be added. Program and run the thermal cycler as listed in Table 4.

Table 3 T5 exonuclease reaction ingredients in 50 μl

Component	Volume (μl)
NEBuffer 4 (10X B7004)	5.0
T5 Exonuclease (NEB M0663L)	4.0
Tris Base 200mM (BB-2686)[1]	2.0
Nuclease-free H_2O (Invitrogen 10977015)	as needed to 50.0 μl

[1]Tris Base 200mM is a 1:4 dilution of Tris Base (1 M, pH $>$ 10.0) in water.

Table 4 T5 DNA exonuclease digestion timing

Step	Temperature	Time
Digesting linear DNA	37^oC	1 hour up to 1.5 hours
Pause	4^oC	indefinitely

Step 3: Cleanup of the circularized DNA (Figure 2c).

Use a spin column kit such as *QIAquick PCR Purification Kit*, or your preferred method.

Implementation Details

1. **The restriction enzyme digestion and ligation** reaction is followed by 15 minutes of heat inactivation to end the ligation reaction (see Figure 2a). The constant 37^oC reaction temperature was chosen to minimize mismatch ligation [25, 26]. T4 DNA ligase inactivates at 65^oC. Since the restriction enzyme was chosen with a higher inactivation temperature, the restriction enzyme will remain active and slightly assist in the following digestion step. The concentrations of ligase and restriction enzyme were experimentally determined to work better with a slightly higher concentration than is typical in Golden Gate assembly. This higher concentration serves

two purposes: it facilitates rapid processing, and, more importantly, the restriction enzyme rapidly cuts to free overhangs on both ends of a DNA fragment, thereby allowing the self-ligation of these overhangs on the same DNA fragment and limiting the proportion of multiple DNA fragments ligating on each other (see further discussion in the Results section). The concentrations of input linear dsDNA were tested in a range from 7.5 to 120 ng/μl and produced indistinguishable circularized DNA yield and composition. A concentration of 120 ng/μl results in the input of 6.0 μg DNA per 50 μl reaction, which was considered optimal from handling and reagent cost perspectives. The addition of ATP may be optional for a 1-hour ligation duration, with increasing ATP concentrations from 1 mM provided by the NEB T4 DNA ligase buffer to 2 mM showing a slight improvement in yield, particularly at longer ligation durations. This may be due to the exhaustion of ATP with a high concentration of ligated DNA, as well as ATP inactivation at longer reaction durations. The 2 mM ATP concentration is well within the optimum range for T4 DNA ligase mixes [27], and the ATP reagent is inexpensive.

The restriction enzyme substitution to a different type IIS enzyme may be required in cases where the sequence coding for epegRNA includes the BsaI recognition site (for an example of using BbsI in a slightly different scenario, see Figure A2).

2. **Digestion of non-circularized DNA** (Figure 2b). This step consists of adding an equal amount of the T5 DNA exonuclease reagents to the ligation reaction. For example, 50 μl of exonuclease reaction mix is added to 50 μl of ligation reaction output, resulting in a comfortable level of 100 μl of reaction mix in a typical 200 μl Eppendorf PCR tube thermocycler (larger tubes and dry baths, etc. may be used if desired). The concentration of T5 DNA exonuclease was validated as sufficient to completely digest all linear DNA in under 1 hour for all input DNA lengths tested in this research. We found that adding only T5 exonuclease to the ligation reaction resulted in a very slow digestion rate due to the absence of potassium ions. The role of NEBuffer 4 is to provide potassium anions that accelerate exonuclease reactions. While the concentration of potassium in the resulting mix is half that of the exclusively NEBuffer 4 reaction mix, it remains close to the optimum concentration for cleaving reactions [28]. The optimal balance between exonuclease and endonuclease activity for T5 exonuclease is at the reaction pH 7.9 – 8.0 [29]. This is the pH of NEBuffer 4, while the ligation buffer pH is 7.5. Thus, a suitable amount of Tris base was added to increase pH to this range, as presented in Table 3. Notably, it is not recommended to excessively increase pH for the T5 exonuclease reaction mix. When we tested pH increases above 8.5, all DNA—including the circularized DNA—was completely and rapidly digested.
3. **Cleanup of the circularized DNA**. This step involves using a quality DNA cleanup kit to extract *Circular Vector* DNA of sufficient concentration and purity. We used a spin column kit, which was selected as described

below. With a large number of such products available, the choice of kit is up to the researcher's preference; however, the yield may vary.

We aimed to select a simple spin column-based extraction kit that would provide a consistently high yield, a low level of impurities, and a sufficiently high DNA concentration for accurate and consistent measurements. This purity was important for the direct use of the *Circular Vectors* in cell culture transfection and for the accuracy of concentration and yield measurements reported by this study. A number of kits were reviewed, and three kits were then tested and evaluated for their efficiency in DNA extraction and low chaotropic salt contamination: the Takara Bio *NucleoSpin Gel and PCR Cleanup* 740609.50 [30], QIAGEN *QIAquick PCR Purification Kit* 28106 [31], and NEB *Monarch PCR & DNA Cleanup Kit* T1030S [32].

Based on a validation test with the same set of PCR products, the QIAquick kit was chosen based on it having the best recovery of input to output DNA when compared to the other two kits (20% better than NucleoSpin and 30% better than NEB) and exhibiting a low amount of captured chaotropic salts when testing a blank PCR product cleanup without any input DNA. Additionally, even though the QIAquick kit rates 30 μl as a minimal elution volume, it performs with a good compromise of yield and higher DNA concentrations and shows excellent purity when eluting with volumes below 20 μl, which was preferable for our prime editing validation experiments. The QIAquick PCR Purification Kit has proven very reliable and consistent, with only one outlier that had approximately half the yield observed in the whole set of experiments. This outlier result may have been due to a defect in the spin column; thus, the sample was discarded.

The QIAquick kit includes an optional pH indicator that is recommended to add to the binding buffer due to its importance in maximizing DNA yield. The indicator reagent allows us to verify whether the mix pH is optimal for binding. Notably, we determined that pH adjustment was required via the addition of one or more 10 μl of 3M pH=5.2 sodium acetate, as recommended by QIAGEN [31].

Special attention was given to minimizing pipetting losses. The amount of input DNA was adjusted to result in output product DNA in the range of 2–4 μg, which provided a sufficient amount of the final DNA product for measurement and testing. In extreme cases, the amount of input dsDNA for the 1782-bp *Circular Vector* had to be scaled up to 27 μg, which still resulted in only a 1.3–1.7 μg yield, while an input DNA amount of 6–9 μg was sufficient in 450–950 bp range. Such low yield, which would require high DNA input and the use of large quantities of reagents to produce sufficient product, indicates that long *Circular Vectors* are impractical to synthesize. We preferably aimed to achieve a concentration of output DNA well over 80 ng/μl, which would result in high-purity DNA based on 260/230 and 260/280 ratios (see sample sets in Supplementary Information). This was additionally aided by performing an extra wash step when using the purification kits, which

resulted in consistently pure DNA output with a minimal potential sacrifice in yield. The DNA concentration measurements were primarily performed on a NanoDrop Eight Spectrophotometer (ThermoFisher Scientific) and repeated on a DS-11 Series Spectrophotometer/Fluorometer (DeNovix). Notably, the measured concentrations were closely matched between these two devices.

The reagents and materials used in this study are listed in Appendix A.3.

Results

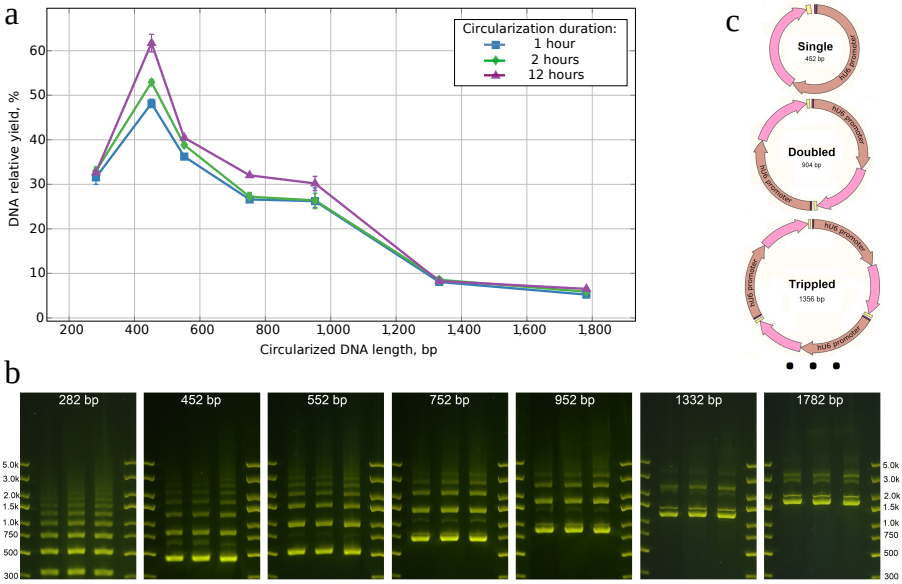

Fig. 3 Circularization yield: (**a**) Relative yield at a range of input DNA lengths at 1, 2 and 12 hours of circularization. All points are the averages of two to three independent reaction samples per data point. Error bars show minimum and maximum yield (when not visible, fitting within the data point marks). (**b**) Gel electrophoresis combination image. In each subplot, Lane 1 = 1 hour, Lane 2 = 2 hours, and Lane 3 = 12 hours of circularization. The quantity of DNA was normalized to 300 ng per lane in subplots 282 to 952 bp. To avoid overload due to the small number of bands, DNA quantity was normalized to 150 ng per lane for the 1332 and 1782 bp subplots. (**c**) Schematic diagram of self-circularization and the joining of repeated multiple fragments.

In every instance Sanger sequencing confirmed the perfect match between the designs and the constructed *Circular Vectors*.

Circularization of the prime editing guides was performed following the protocol described in the Methods section. The goal was to test the relative yield of the reaction, as defined by the ratio of the amount of circularized DNA after purification to the theoretical maximum possible yield when excluding the restriction enzyme sites and padding or PCR adapters that we retained for convenience (see Equation A1). The typical length of the circularized PEmax

epegRNA expression construct would be between 450 and 500 bp, including the promoter and the sequence coding for pegDNA with the end RNA loop, as described in Chen et al. [19]. For this research, we chose a 452-bp *Circular Vector* coding for a single nucleotide substitution in HBB gene [19]. We designed a set of dsDNA with varying padding lengths composed of randomized dsDNA sequences to test how the yield depends on the input dsDNA length and circularization (ligation) step duration. Adding such DNA padding of 100, 300, 500, 880, and 1330 bp to a bare bones 452-bp *Circular Vector* resulted in 552, 752, 952, 1332, and 1782 bp *Circular Vectors* (see Figure A1 and Supplementary Information). Additionally, we tested a shortened 282-bp circularized structure that was too short to contain a functional expression vector in our implementation but was interesting from the perspective of verifying the behavior of potentially shorter constructs.

The relative yield of the method is presented in Figure 3a. Notably, the yield was highest for the circular DNA length range coinciding with the lengths of our prime editing vectors (at or above 450 bp). Excluding this high yield peak, the yield remained relatively constant and dropped off rapidly for lengths above 1000 bp.

The gel electrophoresis results show a picture containing exclusively circularized DNA in multiple bands (see Figure 3b). As expected, in addition to single-unit circularization, some ligations resulted in duplicate, triplicate, and higher-order structures, which are schematically depicted in Figure 3c. This hypothesis was confirmed by cutting out, purifying and sequencing individual bands, as well as re-running gel electrophoresis on DNA purified out of individual bands. The bulk of the circularization occurred within the first hour, exceeding 48% at a size of 452 bp and decreasing to 26% at 952 bp, thereby making it a practically useful method throughout this length range. Unsurprisingly, while yield increased even further with increased circularization/ligation duration, the proportion of higher-order combinations also increased with ligation duration (see gel images in Figure 3b).

Analysis of the gel electrophoresis images using the National Institutes of Health's ImageJ software [33] showed that the higher band representation by DNA weight slightly increased with circularization duration (see Figure 4a). This is particularly noticeable when comparing Band 1 (single circular DNA, blue line) versus the cumulative value of Bands 3 and higher (yellow line). Only at a *Circular Vector* size of 282 bp does simplex Band 1 represent less than Band 2 and less than the combined value of Bands 3 and higher of the vector units by weight in the reaction output.

The number of molar units is always relatively high for Band 1 since the molar representation is proportionate to band DNA weight *divided* by band number (see Figure 4b). For a 1-hour ligation duration in a *Circular Vector* between 450 and 950 bp in length, Band 1 makes a molar fraction of approximately 70%, with Band 2 representing another 20%, and the higher-order bands combined representing the remaining 10–15%. At 12 hours of ligation, the Band 1 share shifts lower by approximately 6%, while Band 2

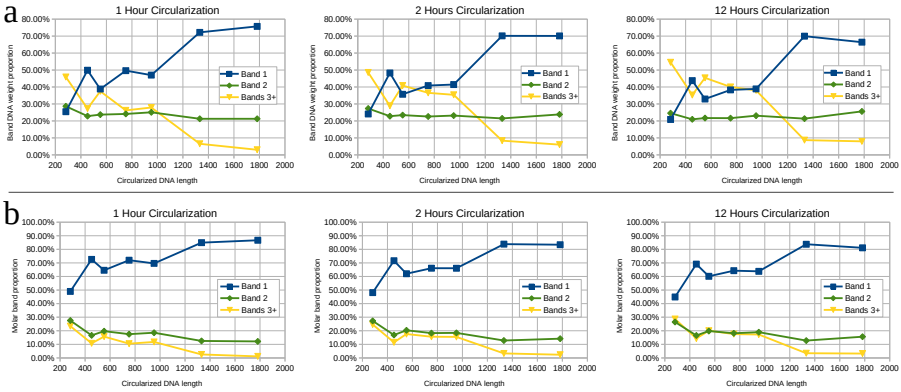

Fig. 4 Characterization of the circularization product by ImageJ gel electrophoresis band brightness analysis. We show the first, second, and the sum of third and higher bands corresponding to single, doubled, and the sum of the higher replicant vectors. (**a**) Relative DNA weight is equal to relative band brightness. (**b**) Relative band molar count is counted as relative band brightness divided by band number (see also Equation A2).

remains rather constant, and the combined higher bands' share increased by approximately 6%. This represents a noticeable increase in higher-order bands with minimal, if any, additional yield for the single-unit *Circular Vectors*.

Inversely, above 1000 bp, the higher bands combine to below 5% molar fraction. However, due to their low yield of 5 to 8%, we consider these longer *Circular Vectors* as not practically useful in our method. We hypothesize that these changes in input DNA length for the yield and fraction distribution of joint *Circular Vector* multiples—including the apparent spike in yield around 450 bp—are likely caused by a combination of DNA length and folding patterns affected by the reaction mix pH, ion influences, and the reaction temperature.

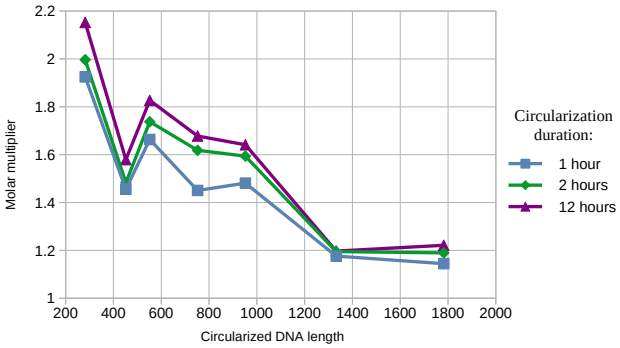

Fig. 5 Molar multiplier by length and circularization duration. Between 452 bp and 952 bp, a simple rule of thumb based on the distribution is a 1.5X multiplier after 1 hour of circularization, and 1.7X for 2 hours of circularization or longer. For example, if the Cas9 plasmid is 13,000 bp long and single circular vector is 452 bp long, then 1.0 μg of the main plasmid at molar ratio 1:1 will require $1.0 * 452/13000 = 0.035$ μg = 35 ng. Multiplying by 1.5 will result in 52 ng of a vector rather than 35 ng for a 1:1 molar ratio based on a 1-hour circularization reaction.

While the majority of circular vectors by number are comprised of single and double units, the higher-order combinations add disproportionately more weight. For transfection, it is important to use a molar proportion of the Cas9-carrying plasmid and guide plasmid. Thus, we calculate a "molar multiplier" that indicates how much more *Circular Vector* by weight must be used when compared to the case of a pure single-unit *Circular Vector* using Equation A2.

The molar multiplier values for 1 hour and 12 hours of ligation in Figure 5 imply that a longer ligation duration results in a molar ratio increase due to a predominant increase in the higher band fraction. When the *Circular Vector* is used for transfection, this higher molar ratio practically negates the nominal higher yield that comes with the longer ligation times. This makes a 1-hour ligation optimal for both the processing time and the final *Circular Vector* product composition.

Verification of the circular vector functionality was performed by editing single base substitutions at three genomic locations (i.e., HBB, PRPN, and CDKL5) that were used by Chen et al. [19] as a demonstration of the PEmax (see methods in Appendix A.4). We coded the guide design from a study by Nelson et al. [20], in conjunction with the Addgene plasmid #174828 [19, 34], which was modified with the addition of blasticidin resistance gene to facilitate an antibiotic selection (see Supplementary Information).

Table 5 Comparison of prime editing efficiency with that of Chen et al. [19] for matching HBB, CDKL5, and PRPN locations.

	Chen et al.[1]		4X ratio		8X ratio	
Genomic location	PE2max	PE4max	minimal	+500bp	minimal	+500bp
HBB +5 G to A	21%	35%	15%	17%	16%	18%
CDKL5 +1 C to T	11%	25%	14%	15%	12%	13%
PRNP +6 G to T	25%	34%	24%	26%	24%	28%

[1]Prime editing efficiency estimated from Figure 7D in Chen et al. [19]. Column entries for our results stating 'minimal' are for minimal length PE4max *Circular Vectors* with lengths of approximately 452 bp. +500 denotes added padding to lengthen guide expression vectors to 952 bp. 4X and 8X ratios are the molar excesses of the *Circular Vector* to the Cas9 plasmid.

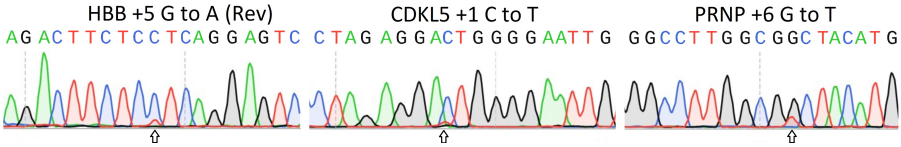

Fig. 6 Validation of the *Circular Vector* by prime editing HBB (shown in reverse), CDKL5, and PRPN.

The analysis performed with EditR [35, 36] is summarized in Table 5 and illustrated in Figure 6. It is evident that the *Circular Vectors* are functional for epegRNA expression when used as a component of prime editing. The editing efficiency was slightly below PE4max and on par with PE2max. Since the purpose of our test was to validate the functional use of *Circular Vectors*, we did not proceed with further fine-tuning of the editing and cell maintenance protocol to maximize the prime editing efficiency, which was finetuned over the years by Chen et al. [19], where the primary goal was to reach maximum editing efficiency. Notably, longer *Circular Vectors* appear to display 1 to 2% greater editing efficiency; however, the difference is minimal. While the use of an 8X molar ratio of guide to Cas9 plasmid may be excessive, we believed that it was worth testing. However, it showed negligible, if any, difference in editing efficiency when compared to the 4X ratio (i.e., the customary molar excesses). These results prove that either length of *Circular Vector* is adequate for use as a epegRNA expression vector in prime editing.

Discussion

The circularization method and protocol presented here describes how to prepare *Circular Vectors* that can be used as an efficient replacement for the guide plasmids in CRISPR, base, and prime gene editing in cell cultures. The main advantage of this method is that after designing, ordering, and receiving dsDNA from a commercial supplier, the product ready for transcription can be prepared for use within 3 hours, with little hands-on time. Moreover, the majority of hands-on time consists of the basic spin column kit DNA cleanup. If the amount of DNA received is low and PCR amplification is required, the process could take an extra hour to amplify the necessary quantity of input DNA.

There are two main limitations or concerns that we wish to discuss:

1. The amount of *Circular Vector* produced by a synthetic reaction. Typically, a single 50 μl reaction will produce 2.0–2.5 μg of product. Although the reaction can be scaled up with more reagents, in situations when a large amount of vectors is needed, the ability of bacterial cloning to produce milligrams of product in one maxi-prep will make it a better choice. In most situations, only a few edits are made with each guide, thus simplicity and speed of preparation favor our circularization method.
2. A more significant concern is the error rate inherent in the manufacturing process of DNA fragments. High fidelity is a strong suite of the bacterial cloning method. The *E. coli* DNA replication error rate as it clones a plasmid is estimated at $5 \cdot 10^{-10}$ per base pair [37]. Once an *E. coli* colony that contains a perfect plasmid is identified by sequencing and cloned, the chances of any plasmid containing an error are vanishingly small, provided that precautions are taken [38].

   On the other hand, as presented in Table 6, the error rate advertised by commercial DNA fragment manufacturers is quite high.

Table 6 Error rates of DNA fragments under 800 bp in length produced by three manufacturers

Manufacturer	Twist Bioscience	Thermo Fisher Scientific	Integrated DNA Technologies
Error rate	1/6,253 bp [39]	1/6,757 bp [40]	1/5,000 bp[1] [41]

[1]Note: both Twist Bioscience [39] and Thermo Fisher Scientific [40] claim that the true error rate for Integrated DNA Technologies gBlocks is much higher (1/2,705 and 1/1,329 respectively).

Let us consider the example of the Twist Bioscience manufacturing process—with an error rate of 1/6,253 bp—and our 452 bp *Circular Vector*. The chance of any *Circular Vector* having an error is approximately 1 in 14. Assuming that an error occurs within the 264 bp promoter area (see Figure A1), with a 1 in 24 chance, it will either leave it functional or render it non-functional. If the promoter is non-functional, it would have no effect on the accuracy of the edit. Moreover, *Circular Vectors* are typically supplied in excess of the Cas9 plasmid, while non-functional *Circular Vectors* will be backed up by functional ones. Similarly, if the error falls in the RNA scaffolding area, it will most likely render the epegRNA non-functional. Furthermore, errors in the linker, extension, or end loop may somewhat lower the editing efficiency. In the case of prime editing, a template error may lead to an error substitution or render the epegRNA non-functional altogether. The chance of an error occurring in the spacer is approximately $20 \div 6,253 \approx 1/312$ for *Circular Vectors*. More likely than not, they will not find any target, which may lead to off-target edit attempts in a small proportion of cases.

It is difficult to estimate the exact potential for an off-target or mis-edit. It appears that one out of a few hundred performed edits could be mis-edited due to a corrupted RT area, with perhaps a many-fold lower probability of erroneous targeting by a corrupted spacer that finds its target in some area of the genome. Simultaneously, prime editing—and to an even larger extent, CRISPR/Cas9 and base editing [3–5, 19]—exhibit a much more significant level of off-target and mis-edits. The presence of off-target edits continues in the newly developed techniques [42]. The Sanger trace analysis conducted with EditR for our three edited genomic locations did not indicate noticeable off-target edits, which is consistent with the aforementioned reasoning that off-target edits, if any, will be orders of magnitude less frequent than successful edits.

There are many scenarios in which this method will be applicable since some off-target fraction is an unavoidable side effect of most approaches, with potential additional off-target edits being a few orders of magnitude lower. Even if the gene-editing approach is inherently precise and free of off-target edits, the *Circular Vector* method may be applied for the rapid validation and selection of a sometimes large set of guides offered by design

tools (e.g., pegFinder by Chow et al.[23]). In situations when the initial evaluation of guides and genomic targets is only performed for initial validation with the intention of using plasmid-based editing for the subsequent precise edits, researchers may simply reuse the input dsDNA fragments used for the production of the *Circular Vector* as an insert into a customary guide plasmid via Golden Gate assembly. Such plasmids can then be cloned and used in a traditional manner.

A few-fold improvement in error rate of input dsDNA would be welcome for the unconditional usability of this method. We hope that within a relatively short time, the precision of current dsDNA suppliers in synthesizing DNA fragments will improve.

There are certain situations where two or more RNAs must be expressed. For example, the prime editing methods PE3 or PE5max [19] require a second RNA guide to produce an extra nick. Even more complex scenarios exist [6, 43]. In such cases, multiple elements may be combined within one *Circular Vector*. Most manufacturers cannot make DNA fragments with long repeats (e.g., repeating promoters or pegRNA scaffolds). However, dsDNA can be ordered as separate fragments and performing the easy additional step of linear ligation can prepare a combined linear dsDNA that can be subsequently circularized by the *Circular Vector* method (see Figure A2 and associated comments). Additionally, in circumstances where a pure single unit *Circular Vector* is required, it can be extracted via gel electrophoresis.

Another important advantage of this method is its low cost. Our cost calculation example in Appendix A.5 shows that the fixed cost of the linear dsDNA fragment ordered from a commercial supplier is \$32.90 for a 452-bp *Circular Vector*. The reaction costs only \$9.30. If PCR amplification is required, this adds \$4.76 to the total and eliminates the need to reorder the input dsDNA fragments from a supplier. The protocol produces 2.0–2.5 μg of *Circular Vector*, which is sufficient for more than 10 edits on a 24-well plate (based on the quantities used in Appendix A.4) and approximately 40 edits on a 96-well plate. This compares favorably to bacterial cloning, where the complete cost from design to a transfection-ready plasmid is typically severalfold higher (in addition to much longer hands-on and overall time). It is even more expensive to order ready-made plasmids.

Conclusion. In this research, we presented a circularization method and protocol for the preparation of *Circular Vectors* that can be used in a variety of applications requiring the expression of short RNAs and proteins. As a proof of the principle, we demonstrated the functionality of the *Circular Vector* method by producing DNA expressing epegRNA for cell culture genome editing. Its speed, low cost, and ease of use will make this method another valuable tool in the gene-editing toolkit, while the reaction simplicity makes the protocol suitable for automated liquid-handling systems [44].

Supplementary Information. The Supplementary Information includes the following files:

1. *Circularization-Experimental-Data-and-Analysis.xlsx* contains circularization quantification data and spreadsheet calculations.
2. *ImageJ-Band-Analysis-Result.xlsx* contains ImageJ gel band analysis results.
3. *Gel-Images.zip* contains an extended set of gel photos.
4. *dsDNA-and-Primers.zip* contains the *Circular Vector* designs and primers used for *Circular Vector* validation and prime edit efficiency PCR and Sanger sequencing, as well as the PE4max Addgene #174828 with a blasticidin resistance insert.

References

[1] Liu, C., Zhang, L., Liu, H., Cheng, K.: Delivery strategies of the crispr-cas9 gene-editing system for therapeutic applications. Journal of controlled release **266**, 17–26 (2017). https://doi.org/10.1016/j.jconrel.2017.09.012

[2] Casali, N., Preston, A.: E. Coli Plasmid Vectors: Methods and Applications vol. 235. Humana Press, Totowa, N.J. (2003)

[3] Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., Zhang, F.: Genome engineering using the crispr-cas9 system. Nature protocols **8**(11), 2281–2308 (2013). https://doi.org/10.1038/nprot.2013.143

[4] 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

[5] Anzalone, A.V., Randolph, P.B., Davis, J.R., Sousa, A.A., Koblan, L.W., Levy, J.M., Chen, P.J., Wilson, C., Newby, G.A., Raguram, A., Liu, D.R.: 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

[6] Anzalone, A.V., Gao, X.D., Podracky, C.J., Nelson, A.T., Koblan, L.W., Raguram, A., Levy, J.M., Mercer, J.A.M., Liu, D.R.: Programmable deletion, replacement, integration and inversion of large dna sequences with twin prime editing. Nature Biotechnology **40**(5), 731–740 (2022). https://doi.org/10.1038/s41587-021-01133-w

[7] Niu, Y., Shen, B., Cui, Y., Chen, Y., Wang, J., Wang, L., Kang, Y., Zhao, X., Si, W., Li, W., Xiang, A.P., Zhou, J., Guo, X., Bi, Y., Si, C., Hu, B., Dong, G., Wang, H., Zhou, Z., Li, T., Tan, T., Pu, X., Wang, F., Ji, S., Zhou, Q., Huang, X., Ji, W., Sha, J.: Generation of gene-modified cynomolgus monkey via cas9/rna-mediated gene targeting

in one-cell embryos. Cell **156**(4), 836–843 (2014)

[8] Zuris, J.A., Thompson, D.B., Shu, Y., Guilinger, J.P., Bessen, J.L., Hu, J.H., Maeder, M.L., Joung, J.K., Chen, Z.-Y., Liu, D.R.: Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. Nature Biotechnology **33**(1), 73–80 (2015). https://doi.org/10.1038/nbt.3081

[9] Kumar, P., Nagarajan, A., Uchil, P.D.: Dna transfection by electroporation. Cold Spring Harbor Protocols **2019**(7), 095471 (2019)

[10] Geng, B.-c., Choi, K.-H., Wang, S.-z., Chen, P., Pan, X.-d., Dong, N.-g., Ko, J.-K., Zhu, H.: A simple, quick, and efficient crispr/cas9 genome editing method for human induced pluripotent stem cells. Acta Pharmacologica Sinica **41**(11), 1427–1432 (2020). https://doi.org/10.1038/s41401-020-0452-0

[11] Kumar, P., Nagarajan, A., Uchil, P.D.: Lipofection. Cold Spring Harbor Protocols **2019**(3), 096248 (2019)

[12] 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

[13] Timin, A.S., Muslimov, A.R., Lepik, K.V., Epifanovskaya, O.S., Shakirova, A.I., Mock, U., Riecken, K., Okilova, M.V., Sergeev, V.S., Afanasyev, B.V., Fehse, B., Sukhorukov, G.B.: Efficient gene editing via non-viral delivery of crispr-cas9 system using polymeric and hybrid microcarriers. Nanomedicine: Nanotechnology, Biology and Medicine **14**(1), 97–108 (2018)

[14] Engler, C., Kandzia, R., Marillonnet, S.: A one pot, one step, precision cloning method with high throughput capability. PloS one **3**(11), 3647–3647 (2008). https://doi.org/10.1371/journal.pone.0003647

[15] Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A., Smith, H.O.: Enzymatic assembly of dna molecules up to several hundred kilobases. Nature Methods **6**(5), 343–345 (2009). https://doi.org/10.1038/nmeth.1318

[16] Addgene: Gibson Assembly Cloning (accessed August 3, 2022). (2022). Addgene. https://www.addgene.org/protocols/gibson-assembly/

[17] Addgene: Bacterial Transformation (accessed August 3, 2022). (2022). Addgene. https://www.addgene.org/protocols/bacterial-transformation/

[18] Addgene: Recovering Plasmid DNA from Bacterial Culture (accessed August 3, 2022). (2022). Addgene. https://www.addgene.org/protocols/purify-plasmid-dna/

[19] Chen, P.J., Hussmann, J.A., Yan, J., Knipping, F., Ravisankar, P., Chen, P.-F., Chen, C., Nelson, J.W., Newby, G.A., Sahin, M., Osborn, M.J., Weissman, J.S., Adamson, B., Liu, D.R.: Enhanced prime editing systems by manipulating cellular determinants of editing outcomes. Cell **184**(22), 5635–565229 (2021). https://doi.org/10.1016/j.cell.2021.09.018

[20] Nelson, J.W., Randolph, P.B., Shen, S.P., Everette, K.A., Chen, P.J., Anzalone, A.V., An, M., Newby, G.A., Chen, J.C., Hsu, A., Liu, D.R.: Engineered pegrnas improve prime editing efficiency. Nature Biotechnology **40**(3), 402–410 (2022). https://doi.org/10.1038/s41587-021-01039-7

[21] Sasaki, A., Kinjo, M.: Monitoring intracellular degradation of exogenous dna using diffusion properties. Journal of controlled release : official journal of the Controlled Release Society **143**, 104–11 (2010). https://doi.org/10.1016/j.jconrel.2009.12.013

[22] Gao, Z., Herrera-Carrillo, E., Berkhout, B.: Delineation of the exact transcription termination signal for type 3 polymerase iii. Molecular Therapy - Nucleic Acids **10**, 36–44 (2018). https://doi.org/10.1016/j.omtn.2017.11.006

[23] Chow, R.D., Chen, J.S., Shen, J., Chen, S.: A web tool for the design of prime-editing guide rnas. Nature Biomedical Engineering **5**(2), 190–194 (2021). https://doi.org/10.1038/s41551-020-00622-8

[24] New England BioLabs: NEBridge Ligase Fidelity Viewer (accessed August 3, 2022). (2022). New England BioLabs. https://ligasefidelity.neb.com/viewset/run.cgi

[25] Potapov, V., Ong, J.L., Kucera, R.B., Langhorst, B.W., Bilotti, K., Pryor, J.M., Cantor, E.J., Canton, B., Knight, T.F., Evans, T.C., Lohman, G.J.S.: Comprehensive profiling of four base overhang ligation fidelity by t4 dna ligase and application to dna assembly. ACS Synthetic Biology **7**(11), 2665–2674 (2018). https://doi.org/10.1021/acssynbio.8b00333

[26] Osman, E.A., Alladin-Mustan, B.S., Hales, S.C., Matharu, G.K., Gibbs, J.M.: Enhanced mismatch selectivity of t4 dna ligase far above the probe: Target duplex dissociation temperature. Biopolymers **112**(1), 23393 (2021). https://doi.org/10.1002/bip.23393

[27] Cherepanov, A.V., de Vries, S.: Kinetics and thermodynamics of nick sealing by t4 dna ligase. European Journal of Biochemistry **270**(21),

4315–4325 (2003). https://doi.org/10.1046/j.1432-1033.2003.03824.x

[28] Zhao, G., Tang, S., Li, J., Hu, T., Guan, Y.: Effects of cations on small fragment of dna polymerase i using a novel fret assay. Acta Biochimica et Biophysica Sinica **46**(8), 659–667 (2014). https://doi.org/10.1093/abbs/gmu050

[29] J., G.S., A., C.T., Dietrich, S., R., S.J.: Mutagenesis of conserved lysine residues in bacteriophage t5 exonuclease suggests separate mechanisms of endoand exonucleolytic cleavage. Proceedings of the National Academy of Sciences **96**(1), 38–43 (1999). https://doi.org/10.1073/pnas.96.1.38

[30] Takara Bio: NucleoSpin Gel and PCR Clean-Up User Manual (accessed August 3, 2022). (2022). Takara Bio. https://www.takarabio.com/documents/User%20Manual/NucleoSpin%20Gel%20and%20PCR%20Clean/NucleoSpin%20Gel%20and%20PCR%20Clean-up%20User%20Manual_Rev_04.pdf

[31] QIAGEN: QIAquick® PCR Purification Kit (accessed August 3, 2022). (2022). QIAGEN. https://www.qiagen.com/us/products/discovery-and-translational-research/dna-rna-purification/dna-purification/dna-clean-up/qiaquick-pcr-purification-kit/

[32] New England BioLabs: Monarch® PCR & DNA Cleanup Kit (accessed August 3, 2022). (2022). New England BioLabs. https://www.neb.com/products/t1030-monarch-pcr-dna-cleanup-kit-5-ug

[33] National Institutes of Health (NIH): ImageJ (accessed August 3, 2022). (2022). National Institutes of Health (NIH). https://imagej.nih.gov/ij/

[34] Addgene: pCMV-PEmax-P2A-hMLH1dn Was a Gift from David Liu (Addgene Plasmid 174828). (2021). Addgene. https://www.addgene.org/174828/

[35] Kluesner, M.G., Nedveck, D.A., Lahr, W.S., Garbe, J.R., Abrahante, J.E., Webber, B.R., Moriarity, B.S.: 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

[36] Moriarity Lab - Medical School - University of Minnesota: EditR (accessed August 3, 2022). (2022). Moriarity Lab - Medical School - University of Minnesota. https://moriaritylab.shinyapps.io/editr_v10/

[37] 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

[38] Tomoiaga, D., Bubnell, J., Herndon, L., Feinstein, P.: High rates of plasmid cotransformation in e. coli overturn the clonality myth and reveal colony development. bioRxiv (2021)

[39] Twist Bioscience: Twist Gene Fragments (accessed August 3, 2022). (2022). Twist Bioscience. https://www.twistbioscience.com/sites/default/files/resources/2020-12/ProductSheet_Genes_GeneFragments_9DEC20_Rev5.pdf

[40] Thermo Fisher Scientific: Evaluation of Linear Synthetic DNA Fragments from Separate Suppliers (accessed August 3, 2022). (2022). Thermo Fisher Scientific. https://assets.thermofisher.com/TFS-Assets/LSG/brochures/geneart-strings-synthetic-dna-white-paper.pdf

[41] Integrated DNA Technologies, Inc: gBlocks Gene Fragments (accessed August 3, 2022). (2022). Integrated DNA Technologies, Inc. https://www.idtdna.com/pages/products/genes-and-gene-fragments/double-stranded-dna-fragments/gblocks-gene-fragments

[42] Choi, J., Chen, W., Suiter, C.C., Lee, C., Chardon, F.M., Yang, W., Leith, A., Daza, R.M., Martin, B., Shendure, J.: Precise genomic deletions using paired prime editing. Nature Biotechnology **40**(2), 218–226 (2022). https://doi.org/10.1038/s41587-021-01025-z

[43] Feng, Y., Liu, S., Mo, Q., Xiao, X., Liu, P., Ma, H.: Enhancing prime editing efficiency and flexibility with tethered and split pegrnas. bioRxiv, 2022–0405487236 (2022). https://doi.org/10.1101/2022.04.05.487236

[44] Tegally, H., San, J.E., Giandhari, J., de Oliveira, T.: Unlocking the efficiency of genomics laboratories with robotic liquid-handling. BMC Genomics **21**(1), 729 (2020). https://doi.org/10.1186/s12864-020-07137-1

[45] Xiong, K., la Cour Karottki, K.J., Hefzi, H., Li, S., Grav, L.M., Li, S., Spahn, P., Lee, J.S., Ventina, I., Lee, G.M., Lewis, N.E., Kildegaard, H.F., Pedersen, L.E.: 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

Appendix A

Appendix A.1 Equations

1. **The maximum possible yield ratio of the circularization reaction** is calculated by dividing the circularized DNA length, as determined by the restriction enzyme cut sites based on input dsDNA length. Typically, a type IIS restriction enzyme requires six base pairs or more between it and the end of the dsDNA strand to ensure efficient cutting. The restriction enzyme recognition site itself, an offset between the recognition site and the cut location, and one length of the overhang are also discarded. For our handling, it was convenient to retain the Twist Bioscience end adapters, which are 22 bp in length and can be used with Twist Bioscience primers for the PCR amplification of these DNA fragments. In this case, the maximum possible yield ratio denominator is the length of the dsDNA ordered from Twist Bioscience plus 44 bp. All of the elements can be formally accounted for by the following equation:

$$R = \frac{L_c}{L_c + 2 * L_{enz_site} + 2 * L_{enz_offset} + L_{cut_length} + 2 * L_{end_pad}}, \quad \text{(A1)}$$

   where R is the maximum yield ratio, L_c is the length of the circularized DNA, L_{enz_site} is the recognition enzyme site, L_{enz_offset} is the offset between the enzyme recognition site and the beginning of the cut, L_{cut_length} is the length of the cut (typically 4 bp), and L_{end_pad} is either padding or end adapters (as in our case), assuming that they are of equal length on either end. As an example, for dsDNA delivered by Twist Bioscience with adapters for making a 452-bp *Circular Vector*, the maximum possible yield ratio is 0.88. Thus, if 6 μg of DNA is used as an input, the maximum possible yield will be $6 * 0.88 = 5.28$ μg. This equation can be used as a guide to calculations of efficiency performed in the yield formulas presented in the Supplementary Information.
2. We define **the molar multiplier** of the circularization reaction product as the ratio of the combined relative weight produced by circularization to the molar relative fraction of a single circular guide. We assess the relative band brightnesses from the gel electrophoresis image, which was processed using the NIH's ImageJ. Since we know the relative brightness, we must sum all band relative brightnesses, which, by definition, will be = 1 divided by the sum of each band's relative brightness (I_i) divided by that band's number i. Thus, the molar multiplier M is calculated as:

$$M = \frac{\sum_{i=1}^{n} I_i}{\sum_{i=1}^{n} I_i/i} = \frac{1}{\sum_{i=1}^{n} I_i/i}. \quad \text{(A2)}$$

Appendix A.2 Figures

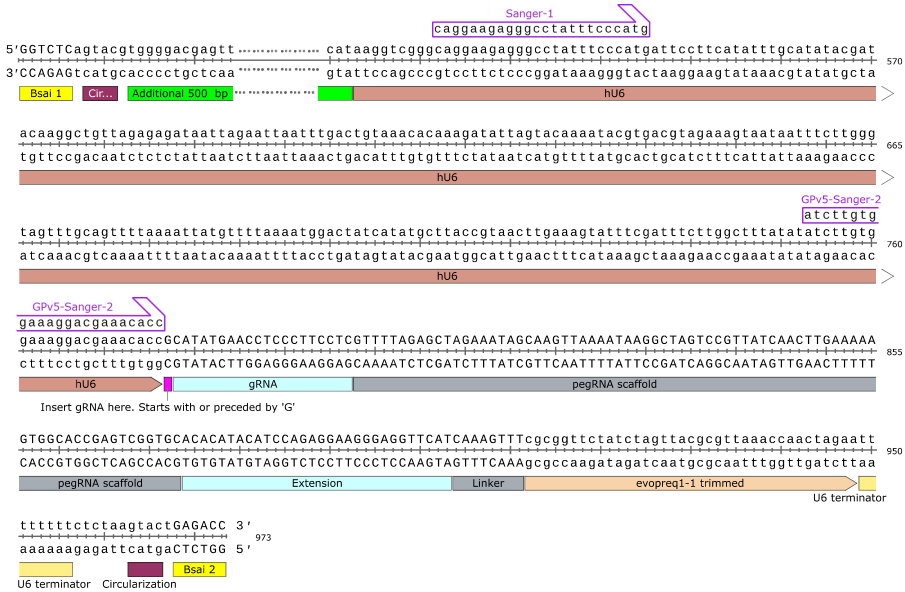

Fig. A1 Full example of a *Circular Vector* containing all elements required for PE2max and PE4max. This type of design applies to any short expression vector—and thus not exclusively RNA guides. The end padding is not included because dsDNA was manufactured by Twist Bioscience, with the option to keep the end adapters; otherwise, at least 6 bp of padding would be required on each end. The green interrupted DNA block denotes an optional space for additional padding, which is filled with random well-balanced DNA sequences (exemplified by 500 bp in this case). In this study, this was used to implement the 100, 300, 500, 880, and 1330 bp length extensions.

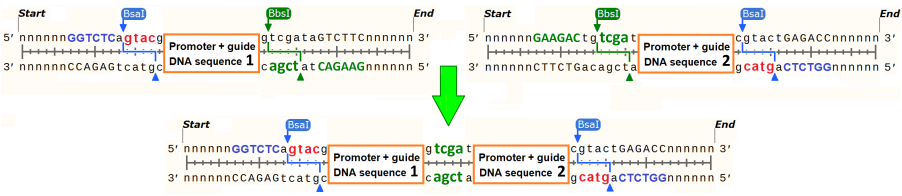

Fig. A2 Linear joining of dsDNA fragments as a preliminary step to circularizing dsDNA fragments with long repeats.

Joining two or more vectors using Golden Gate protocol reagents can be used as a rapid and fully synthetic preliminary step to join dsDNA fragments with matching overhangs. The majority of commercial suppliers of DNA fragments fail to make dsDNA sequences that contain large repeats. Thus, they would not accept an order for such combined fragments. For

example, prime editing PE3 and PE5 [19] require two U6 promoters that allow the transcription of two RNA guides. The corresponding constructs are customarily inserted into plasmid vectors using Golden Gate or Gibson assembly, followed by bacterial cloning. In an even more complex scenario, many stages of assembly and bacterial cloning involve approximately 1 week of work and handling [43]. The joining step requires designing and ordering fragments that can be cut with an intermediate type IIS restriction enzyme (i.e., BbsI in the example presented in Figure A2). The joining can be accomplished in a couple of hours, involving 1 hour of a constant 37°C reaction using the reagents listed in Table 1 while substituting BbsI instead of BsaI. This must be followed by DNA cleanup to discard short DNA end cut-offs by using products such as diluted 1:1 Takara Bio *NucleoSpin® Gel and PCR Cleanup* [30]. Such linear ligation allows us to join linear dsDNA with nearly 100% efficiency, with losses primarily being due to the cleanup using a spin column kit. This preliminary step can then be followed by circularization using the *Circular Vector* protocol, resulting in the entire process being completed many times faster than the many stages of bacterial cloning.

Appendix A.3 Reagents and Materials

Table A1 KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
BsaI-HFv2	New England BioLabs	Cat#R3733L
T4 DNA Ligase	New England BioLabs	Cat#M0202L
Adenosine-5 Triphosphate (ATP)	New England BioLabs	Cat#P0756S
Lipofectamine 2000	Thermo Fisher Scientific	Cat#11668019
T5 Exonuclease	New England BioLabs	Cat#M0663L
Tris Base (1 M, pH >10.0)	Boston Bioproducts	Cat#BB-2686
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
Opti-MEM I Reduced Serum Medium	Thermo Fisher Scientific	Cat#31985062
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 2 × 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
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
pCMV–PEmax–P2A–MLH1dn	Addgene	Cat#174828

Appendix A.4 Transfection, Handling, and Sequencing of Edited HEK293T Cells

HEK293T cells were seeded on 24-well plates (Corning) at $25 - 30 \cdot 10^3$ cells per well in DMEM supplemented with 10% FBS. Approximately 24 hours after seeding, the cells were transfected at 20% confluency with 2.0 mL of Lipofectamine 2000 in Optimem (Thermo Fisher Scientific) according to the manufacturer's protocol and 800 ng of Cas9 prime editor plasmid, with molar ratios of 4:1 (165 ng for a 450 bp length and 350 ng for an 850 bp length) and 8:1 (330 ng for 450 bp length and 700 ng for 850 bp length) of epegRNA *Circular Vectors* for each of the three edit locations.

Then, 14–16 hours after transfection, media was replaced with fresh DMEM plus 10% FBS and 10 $ng/\mu l$ of blasticidin (Thermo Fisher Scientific) to select for cells containing prime editor. This was followed 24 hours later by 5 $ng/\mu l$. Then, a media change was performed after 24 hours (based on a recommendation by Xiong et al. [45]). Genomic DNA was extracted 48 hours later using a Zymo Research Quick-DNA Microprep Kit (D3020).

Genomic areas of interest were amplified from 4 ng of whole genome DNA using Q5 High-Fidelity 2x Master Mix (New England BioLabs) at 24–26 cycles using the master mix protocol (see Supplementary Information files for a 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 PCR & DNA Cleanup Kit (New England Biolabs) with one extra wash step. Sanger sequencing was performed by Azenta Genewiz using our PCR primers.

Appendix A.5 Reagent Costs for a 50 μl Reaction

The bulk of the expense inherent in *Circular Vectors* is the cost of linear dsDNA ordered from a commercial supplier. We will present the costs for the example of our 452-bp *Circular Vector*. The prices we paid for the reagents were typical for retail university buyers, without any extra discounts. With a circularized length of 452 bp, a 470-bp dsDNA fragment was ordered from Twist Bioscience at the list price of \$0.07 per base pair, resulting in an input DNA cost of \$32.90. Usually, 1.5–2.0 μg of dsDNA fragment was delivered by Twist Bioscience. The costs of the reagents for a 50 μl protocol are presented in Table A2 (Step 1, circularization), Table A3 (Step 2, digestion), and Table A4 (Step 3, DNA cleanup), as described in the Methods section. The planned output is 2.0–2.5 μg of *Circular Vector*, which requires an input of 6.0 μg of dsDNA. Such yield is sufficient for more than 10 edits on a 24-well plate and approximately 40 edits on a 96-well plate. The cost of this *Circular Vector* protocol is \$9.30 (see Table A5). If your supplier can deliver a sufficient quantity of DNA fragments, this is all your expense.

In our study, we had to run a PCR amplification step (priced in Table A6), which added \$4.76 to the cost. If only a small amount of *Circular Vectors* is needed, the reaction can be scaled down to a smaller quantity of available DNA, with a lower use of reagents and lower resulting cost.

Table A2 Circularization reaction reagent costs for 6.0 μg of input dsDNA in 50 μl.

Component and price	Volume (μl)	Cost
T4 DNA Ligase Buffer (10X B0202S included in M0202L)	5.0	\$0 L
T4 DNA Ligase (NEB M0202L 0.250 ml, \$202.75)	3.0	\$2.43
Bsai-HF2 (NEB R3733L 0.25 ml, \$236.00)	3.0	\$2.83
ATP (10mM NEB P0756S 1.0 ml, \$30.04)	5.0	\$0.15
Nuclease-free H_2O (Invitrogen 10977015 500 ml, \$27.00)	as needed to 50.0 μl	\$0.001
Circularization cost		**\$5.41**

Table A3 T5 exonuclease digestion reaction reagent costs for a 50 μl reaction.

Component and price	Volume (μl)	Cost
NEBuffer 4 (10X B7004 included in M0663L)	5.0	\$0
T5 Exonuclease (NEB M0663L 0.5 ml, \$222.40)	4.0	\$1.78
Tris Base 200mM (BB-2686 500ml, \$40.43)[1]	2.0	\$0.000
Nuclease-free H2O (Invitrogen 10977015 500 ml, \$27.00)	as needed to 50.0 μl	\$0.002
T5 exonuclease digestion cost		**\$1.78**

[1]Tris Base 200mM is a 1:4 dilution of Tris Base (1 M, pH $>$ 10.0) in water.

Table A4 QIAquick kit DNA purification for a 50 μl reaction.

Component and price	Number	Cost
QIAquick PCR Purification Kit (QIAGEN 28106, 250 reactions \$527.50)	1	\$2.11
QIAquick kit DNA purification cost		**\$2.11**

Table A5 Sum total of the *Circular Vector* protocol for a 50 μl reaction (2.0–2.5 μg yield).

Step	Cost
Circularization - Table A2	\$5.41
Digestion - Table A3	\$1.78
QIAquick kit DNA purification - Table A4	\$2.11
Cost for the *Circular Vector* protocol	**\$9.30**

Table A6 PCR amplification, 75 μl reactions yielding above 6.0 μg of input dsDNA.

Component and price	Units	Cost
HiFi Hot Start–Readymix (KAPA Biosystems KK2602 6.25 ml, \$434.28)	37.5 μl	\$2.61
Nuclease-free H2O (Invitrogen 10977015 500 ml, \$27.00)	30 μl	\$0.002
Two primers combined (Approx 1000 μl, \$3.00)	6 μl	\$0.015
QIAquick PCR Purification Kit (QIAGEN 28106, 250 reactions \$527.50)	1	\$2.11
PCR amplification cost		**\$4.76**