Plant Biotechnology Reports (2019) 13:543–551
https://doi.org/10.1007/s11816-019-00567-8

Online ISSN 1863-5474
Print ISSN 1863-5466

ORIGINAL ARTICLE

A zero-background CRISPR binary vector system for construction of sgRNA libraries in plant functional genomics applications

Jae-Young Yun[1] · Sang-Tae Kim[1] · Sang-Gyu Kim[2] · Jin-Soo Kim[1]

Received: 15 July 2019 / Accepted: 17 September 2019 / Published online: 24 September 2019

Abstract

Clustered regularly interspaced short palindromic repeats (CRISPR)-mediated genome editing is a ground-breaking biotechnology for agricultural applications such as precision breeding in crop plants. *Agrobacterium*-mediated CRISPR delivery has been successfully adapted for gene knockout applications for basic research and agricultural technology development. However, selecting proper single-guide RNA (sgRNA) for CRISPR binary constructs to induce double-strand break in certain target genes has presented difficulties mainly due to unpredictable in vivo sgRNA activities. Therefore, more than three independent CRISPR constructs, each harboring different sgRNAs, are often applied to ensure the desired CRISPR-induced knockout alleles. Here, we report a zero-background CRISPR binary vector platform, featuring *ccdB* conjugation within sgRNA cloning cassette, which is later removed by AarI endonuclease, that allows positive survival selection for bona-fide sgRNA clones and effective exclusion of uncut or self-ligated 'background' negative clones. We demonstrate the advantage of using the zero-background CRISPR binary platform in a high-throughput pooled cloning strategy of multiple different sgRNAs which produced *Agrobacteria* containing multiple sgRNAs without any background. We also tested the integrity of pooled CRISPR sgRNA construct libraries during extended bacterial culture and during the transfer between *Escherichia coli* to *Agrobacterium,* and verified that the fidelity of sgRNA species representation was faithfully maintained during library generation.

Keywords Genome editing · CRISPR/Cas9 · *ccdB* · Binary vector · sgRNA · Library screen

Introduction

Ever since clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein (Cas) was identified as an adaptive prokaryotic immune system (Barrangou et al. 2007; Marraffini 2015), CRISPR/Cas9 systems have been adopted also in plant science as tools in a wide variety of plant species to investigate gene function and to improve agricultural traits (Liu et al. 2017; Yin et al. 2017; Zhang et al. 2017, 2018). The delivery of CRISPR/Cas9 systems to plant cells has been successfully conducted by adapting three different methods: *Agrobacterium*-mediated transformation of CRISPR/Cas9 plasmids (Feng et al. 2013); the bombardment-mediated delivery of CRISPR/Cas9 systems (Shan et al. 2014); and PEG-mediated delivery into protoplasts of CRISPR/Cas9 ribonucleoproteins (Woo et al. 2015). Among those delivery methods, *Agrobacterium*-mediated transformation has been most effective for many plants and thus has been widely utilized to generate genome-edited varieties (so-called 'gedivar' comparable with 'cultivar') that have better yields and other new traits including disease and pathogen resistance (Cai et al. 2015; Du et al. 2016; Feng et al. 2016; Curtin et al. 2018; De Pater et al. 2018).

Jae-Young Yun and Sang-Tae Kim contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11816-019-00567-8) contains supplementary material, which is available to authorized users.

✉ Jae-Young Yun
jae0yun@ibs.re.kr

✉ Jin-Soo Kim
jskim01@snu.ac.kr

[1] Center for Genome Engineering, Institute for Basic Science, Daejeon 34126, South Korea

[2] Department of Biological Sciences, Korea Advanced Institute of Science and Technology, Daejeon 34141, South Korea

A sgRNA expression cassette in the CRISPR/Cas9 system contains typically 18–20 bp-long guide sequence that is matching to target sequence. The expression of sgRNA in plant binary vector systems is driven by RNA pol III-dependent promoters such as U3 or U6 promoters. To incorporate target guide sequence in the sgRNA cassette, target-adapter ligation or overlapping PCR can be used (Li et al. 2013; Mao et al. 2013; Shan et al. 2013; Xie and Yang 2013). It is very crucial to design and select effective sgRNAs for successful CRISPR application. Although numerous computational tools have been reported for identification of on-target-specific guide sequences, those in vivo sgRNA specificities and activities are often unpredictable. Therefore, it is often recommended that multiple, different guide sequences are selected from the computational prediction to target a gene of interest.

One powerful approach for functional genomics is a library screen. In animal cells, high-throughput screening using pooled CRISPR sgRNA libraries has been proven to be an effective way to identify novel protein functions by knocking out genes across a population of cells. Pooled CRISPR libraries typically consist of thousands of plasmids, each containing multiple sgRNAs for each target gene. To date, a number of CRISPR libraries for animal cells have been validated and available (Ford et al. 2019; Schuster et al. 2019). The emerging areas of research for genome editing in plants also include sgRNA library screen for high-throughput interrogation of certain gene function or regulatory elements in non-coding region. Recently, in vivo plant screens using CRISPR/Cas9 binary vector library targeting the rice *SF3B1* spliceosomal protein gene has been successfully demonstrated to interrogate or evolve functional domains of the target gene (Butt et al. 2019). In the near future, in vivo high-throughput library screen in plants, often combined with various Cas9 derivatives, may pave the way to investigate either coding or non-coding genome, facilitating research for crop improvement.

Genome editing of the plant model *Arabidopsis* generally uses *Agrobacterium*-mediated floral dipping transformation (Clough and Bent 1998). Since the *Agrobacterium*-mediated transformation is the most effective method for the genome editing in many plants, CRISPR binary vectors have been developed in many ways for effective expression of both the *Cas9* and sgRNA cassettes, ensuring generating germline-transmissible edited alleles (Wang et al. 2015; Yan et al. 2015; Tsutsui and Higashiyama 2017). Among developed CRISPR binary vector systems for *Arabidopsis*, the system with *RIBOSOMAL PROTEIN S5A* (*RPS5A*) promoter-driven Cas9 has been reported to be one of the most effective CRISPR binary vectors (Kang et al. 2018). Here, we present a newly developed pZeroBG-CRISPR binary vector systems, based on the modification of *RPS5A* promoter-driven Cas9 binary vector via conjugation of cytotoxic *ccdB* gene within its sgRNA cloning cassette, which ensures absolute zero-background colonies by enabling positive survival selection of guide-positive colonies over guide-negative colonies after *Escherichia coli* transformation (Bernard and Couturier 1992; Bernard et al. 1994). Our data suggest that pooled cloning strategies using the pZeroBG-CRISPR system may allow efficient and convenient construction of the libraries harboring multiple sgRNAs, reliably applicable to plant library-screening approaches.

Materials and methods

pZeroBG-RpCas9 vector construction

All primers used in this study are listed in Table 1. PCR-amplified fragment from pDONOR^{TM}221 (Thermo Fisher Scientific, Cat#12,536,017), containing both *ccdB* and chloramphenicol-resistant marker expression cassette, was transferred into pJY-RpCas9 via using In-Fusion® HD cloning kit (Takara, Cat#639,650) to generate pZeroBG-RpCas9 vector. pJY-RpCas9 is a modified form of pKI1.1R (Tsutsui and Higashiyama 2017), removing seed-fluorescence marker, and was tested for its *in planta* activities of Cas9-induced indels (Kang et al. 2018; deposited to Korean Collection for Type Cultures, BP#1429736). The primer pair ccdB-F and ccdB-R, consisting of 15 bp homology sequence to destined vector (pJY-RpCas9) and target sequence, was used to amplify the *ccdB* fragment from pDONOR^{TM}221. The PCR amplicon and the AarI-treated pJY-RpCas9 vector were purified and were subject to In-Fusion® HD cloning reaction. Resulting material was used to transform chemically competent *E. coli* cells of strain DB3.1 to suppress *ccdB* toxicity (Bernard and Couturier 1992; Bernard et al. 1994). Transformed colonies from selective media containing spectinomycin (plasmid marker) and chloramphenicol (*ccdB* cassette marker) were validated for the conjugation of the *ccdB* cassette.

Cloning of guide RNA sequences into CRISPR binary vectors

The CRISPR vector pZeroBG-RpCas9 and pJY-RpCas9 were linearized by restriction enzyme AarI for 3 h to produce 4 bp overhangs. The linearized vectors were purified using Expin™ PCR SV mini kit (GeneAll®, Cat# 103-202). Forward and reverse oligos (FT-F & FT-R for *FT* and PDS3-F & PDS3-R for *PDS3*) for each guide sequences were mixed together to make 10 µM concentration each in the presence of T4 DNA ligase 10 × buffer (New England BioLabs®, Cat# B0202S), and then were subject to oligo annealing conditions using thermocycler programmed for: predenaturation for 1 min at 95 °C; 95–25 °C for – 1 °C/min;

Table 1 The primers used in this study

Name	Sequence (5′-3′)
ccdB-F	TCGAAGTAGTGATTGAAATGCAGGTGCGGAATCGTCGTATCCAGCC
ccdB-R	TTCTAGCTCTAAAACTTATGCAGGTGCGCACTTTGCGCCGAATA
FT-F	ATTGGAGATATTCTCGGAGGTGA
FT-R	AAACTCACCTCCGAGAATATCTC
PDS3-F	ATTGCCTCCAGATAGCTGCATGGA
PDS3-R	AAACTCCATGCAGCTATCTGGAGG
Guide1-F	ATTGGAATACGGTTTCTACGCCT
Guide1-R	AAACAGGCGTAGAAACCGTATTC
Guide2-F	ATTGGTAACGTCCTTCCTCGTCA
Guide2-R	AAACTGACGAGGAAGGACGTTAC
Guide3-F	ATTGGGTACAGATGCGTTTCAAG
Guide3-R	AAACCTTGAAACGCATCTGTACC
Guide4-F	ATTGATCAACCAAAACAGGTCCA
Guide4-R	AAACTGGACCTGTTTTGGTTGAT
Guide5-F	ATTGGTGTATGCAAATTACGCTG
Guide5-R	AAACCAGCGTAATTTGCATACAC
Guide6-F	ATTGGTTTGATGATCGTGTCACG
Guide6-R	AAACCGTGACACGATCATCAAAC
Guide7-F	ATTGACCATGCCAAGATGCTGGT
Guide7-R	AAACATTGACCAGCATCTTGGCATGGT
Guide8-F	CAAGAACATGTGTTGCCGA
Guide8-R	AAACATTGTCGGCAACACATGTTCTTG
NGS-F	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTCTTCAAAAGTCCCACATCG
NGS-R	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGGTGCCACTTTTTCAAGTTG

hold at 10 °C. The linearized vectors and annealed oligos for guide sequences were ligated and the ligation products were transformed to *E. coli* competent cells as previously described (Kim et al. 2016). Depending on the experimental conditions, the vector was treated again for 3-h AarI digestion after purification or gel-purified using Expin™ Gel SV kit (GeneAll®, Cat# 102-102). *E. coli* transformants were selected on the media containing spectinomycin, and those isolated plasmid DNA using Exprep™ plasmid SV mini kit (GeneAll®, Cat# 101-102) were validated for Sanger sequencing.

High-throughput pool cloning and pool transformation to bacterial cells

For high-throughput sgRNA pool cloning, we further optimized the procedures based on the single-sgRNA cloning protocol described above. Forward and reverse oligo pairs for guide 1 to guide 8 (Table 1) were mixed individually for 15 μM concentration each to make total 20-μl volume and were annealed in the programmed thermocycler as mentioned above. For pool ligation reaction of eight guides, each 3 μl of annealed products were mixed with 240 ng of linearized vector (1 guide for 30 ng of linearized vector) to make total 64-μl volume (1 guide for 8-μl volume) of ligation products. 16 μl of ligation products was used to transform 150 μl of competent DH5α cell suspension to obtain more than 50 colonies on selective media in our optimized condition. All the colonies in the plate were collected and inoculated in 5-ml liquid LB media for overnight growth at 37 °C shaking incubator. The next day, pooled plasmid DNA was isolated from the overnight *E. coli* culture using Exprep™ plasmid SV mini kit (GeneAll®, Cat# 101-102) for either next-generation sequencing (NGS) validation for each guides' representation or *Agrobacterium* transformation. 50 ng of pooled plasmid DNA (representing approximately 2.3×10^9 copies) was used to transform 50 μl of *Agrobacterium* competent cells (strain GV3101) by electroporation to get more than thousands of colonies in a selection media in our optimized condition. Ten scoops of colonies were inoculated for overnight seed culture in 25 ml of liquid YEP media containing spectinomycin at 28 °C shaking incubator. The next day, 5 ml of the overnight culture was subject to miniprep of plasmid DNA for sequencing validation and the rest 20 ml was used in 200 ml of subculture without antibiotics for 4 h at 28 °C shaking incubator.

Next-generation sequencing (NGS) validation for guide representation

To construct NGS libraries, target amplicon spanning sgRNA expression cassette were PCR-amplified using primer NGS-F & NGS-R with Phusion™ polymerase (Thermo Fisher Scientific, Cat# F530N). Then we performed the index PCR to carry out NGS using MiniSeq m System (Illumina®). All target amplicon sequencing reads were joined to get joined reads with pair-end sequenced reads for each indexed samples using 'fastq-join' implemented in 'ea-utils (https://github.com/ExpressionAnalysis/ea-utils). With joined reads we identified and counted the number of reads representing each sgRNA bearing construct by matching the 30 bp-long sequences including sgRNA sequences inserted using bash scripts. Similarly, constructs that might have no sgRNA insertion were identified by matching the 30-bp-long sequences in the original backbone construct from joined reads.

Results and discussion

Construction of a zero-background CRISPR-guide cloning vector

We sought to generate a cloning vector that would allow for all *E. coli* transformants to be guide-positive ones from the CRISPR sgRNA cloning process which would require a strategy to eliminate uncut or self-ligated CRISPR vectors. Most currently available CRISPR binary vectors for plant transformation utilize endonuclease type IIS-dependent cloning strategies for the incorporation sgRNA expression cassettes. Typically, the treatment of endonuclease type IIS such as AarI to binary CRISPR vectors generate a linearized vector with 4 bp of 5′ overhangs at both ends, which later serve as 'sticky ends' when hybridized with the annealed oligos harboring guide sequences for sgRNA expression (Fig. 1). However, incomplete vector digestion and/or removal of small cut fragments (20–30 bp length) containing nuclease recognition sites results in undesirable 'background' *E. coli* transformants containing the original vector form. These 'background' *E. coli* transformants

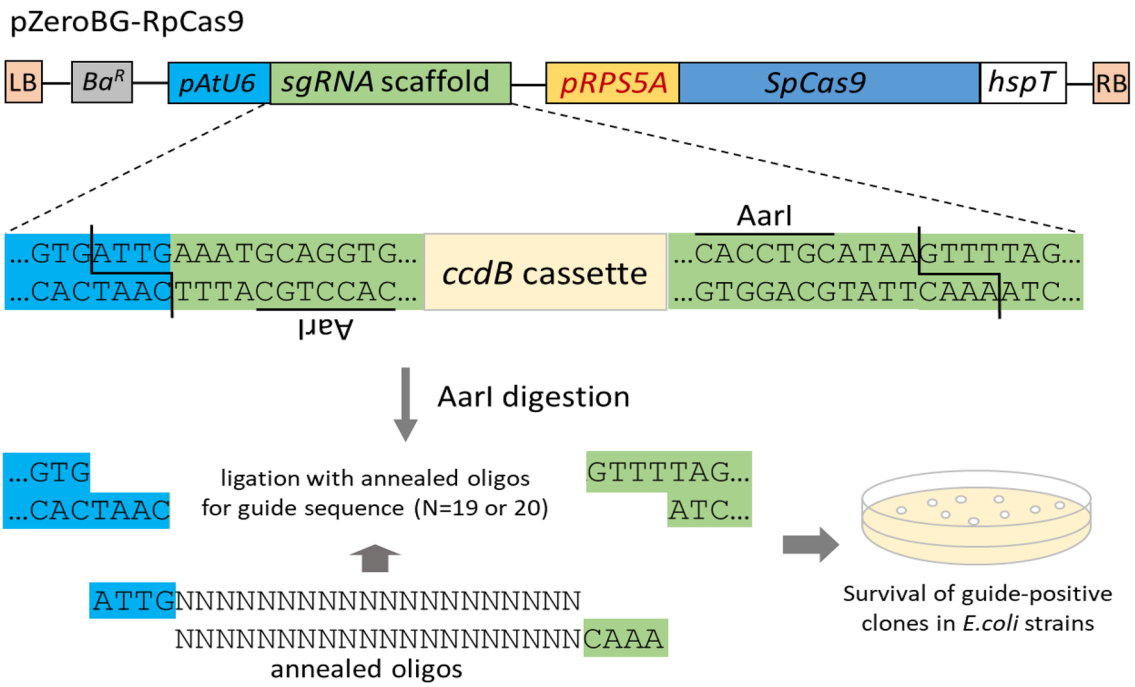

Fig. 1 Schematic representation of pZeroBG-RpCas9 vector and cloning of guide sequence. pZeroBG-RpCas9 binary vector system is characterized by inserting the *ccdB* expression cassette between AarI recognition sites within the sgRNA scaffold (represented in top panel). The treatment of endonuclease AarI generates a linearized vector with 4 bp of 5′ overhangs at both ends, which later serve as 'sticky ends' when hybridized with the annealed oligos harboring guide sequences for sgRNA expression (represented in bottom panel). The cytotoxicity from the *ccdB* cassette genetically excludes undesirable self-ligation products during *E. coli* transformation. LB and RB represent left border and right border of T-DNA, respectively. Ba^R represents Basta resistance marker. *pAtU6* and *pRPS5A* denotes Arabidopsis U6 and RPS5A promoters, respectively. *hspT* indicates heat shock protein terminator

hinder a 'hassle-free' cloning experience and necessitate additional validation steps to the cloning process.

To eliminate unwanted intact vector species during *E. coli* transformation, we developed a pZeroBG-CRISPR binary vector system that genetically excludes undesirable ligation products during *E. coli* transformation by inserting the *ccdB* expression cassette between AarI recognition sites within the sgRNA scaffold (Fig. 1). *ccdB* gene produces a cytotoxic protein that is a potent poison of gyrase and therefore lethal to most *E. coli* strains including DH5α (Bernard and Couturier 1992; Bernard et al. 1994). During the ligation process, annealed guide oligos with complementary 5′ overhangs to sticky ends of linearized AarI-treated pZeroBG vector can replace the cut fragment containing *ccdB* gene. Thus, only CRISPR-guides-positive plasmids that have lost the *ccdB* gene are able to survive in the transformed *E. coli* cells (Fig. 1).

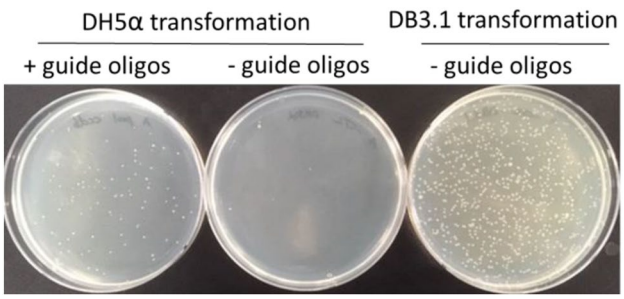

Fig. 2 Cytotoxic *ccdB* effects on colony formation in the cloning process. Colonies formed from either DH5α or DB3.1 transformation of self-ligated products in the absence of guide oligos after insufficient AarI digestion of pZeroBG-RpCas9 vector are shown in the right two plates. Predicted guide-positive colonies from the DH5α transformation of ligation products in the presence of guide oligos are shown in the left plate

Testing the cloning efficiency of binary vector pZeroBG-RpCas9

First, to test the function of cytotoxic *ccdB* effects of pZeroBG-RpCas9 in the cloning process, the vector was cut by AarI (~1 h) and subject to a ligation reaction either in the absence or the presence of annealed oligos with compatible overhangs at both ends. *E. coli* transformation of the self-ligated products without oligos gave rise to nearly zero colonies from DH5α transformation (Fig. 2), while a number of colonies were formed from DB3.1 strain transformation (Fig. 2) where the cytotoxic effects of *ccdB* are tolerated by the gyrA462 mutation (Bernard et al. 1994). These data indicate that the *ccdB* cassette in pZeroBG vector can effectively eliminate colonies that were transformed with vectors lacking sgRNA oligo cassettes that would have otherwise survived when using canonical CRISPR binary vectors. Meanwhile, the ligation reaction with guide oligos produced a significant number of colonies upon DH5α transformation (Fig. 2), indicating that the *ccdb* cassette was readily replaced by guide oligos during the ligation reaction.

Next, to further evaluate the cloning efficiencies of pZeroBG-CRISPR vector, we attempted to clone two different guide oligos into pZeroBG-RpCas9 as well as pJY-RpCas9, a prototype CRISPR vector prior to the *ccdB* conjugation. All the cloning processes using two different guide sequences, targeting *FT* and *PDS3*, respectively, were carried out under the same conditions for oligo annealing, ligation, and DH5α transformation. Plasmid DNA was isolated from cultures of randomly picked colonies generated from DH5α transformation, and were subject to Sanger sequencing analyses to test if the guide sequences are present within their sgRNA cassettes. All twenty-four colonies tested from the guides cloning using pZeroBG vector, regardless of target genes, showed absolutely no guide-negative clones but all guide-positive clones, indicating 100% cloning efficiency (Table 2). In cases of guide cloning using pJY-RpCas9, the sequencing results revealed that a significant

Table 2 Number of guide-positive and -negative clones after DH5α transformation

gRNA targets	Binary vector used	No. of tested clones	No. of guide-positive clones	No. of guide-negative clones	The ratios of positive clones
FT	pZeroBG-RpCas9	24	24	n/a	24/24 (100%)
FT	pJY-RpCas9	24	12	12	12/24 (50%)
FT	pJY-RpCas9 (EC)	24	18	6	18/24 (75%)
FT	pJY-RpCas9 (GP)	12	10	2	20/24 (83%)
PDS3	pZeroBG-RpCas9	24	24	n/a	24/24 (100%)
PDS3	pJY-RpCas9	24	14	10	14/24 (58%)
PDS3	pJY-RpCas9 (EC)	24	20	4	20/24 (83%)
PDS3	pJY-RpCas9 (GP)	10	7	3	21/24 (70%)

EC extended cut, *GP* gel purified

number of randomly picked colonies were guide-negative clones, which were verified by their intact original vector sequences, resulting in cloning efficiencies of *FT*-targeting guide and *PDS3*-targeting guide of 50% and 58%, respectively (Table 2). We also made an effort to increase the cloning efficiencies of pJY-RpCas9 vector by allowing longer AarI digestion [extended cut (EC)] or gel purification of linearized vector, but found only a slight or no increase of efficiencies (Table 2). Notably, cloning via gel purification of the linearized vector produced much fewer colonies (12 colonies for FT and 10 colonies for *PDS3*) compared with more than 30 colonies in other cases, suggesting that the gel purification step might compromise the quality of the linearized vector in our tested condition. Taken together, the results suggest that pZeroBG vector ensures zero background and thus enables highly efficient cloning of guide sequences in the binary vector.

pZeroBG system allows high-efficiency construction of pooled CRISPR vectors containing sgRNA libraries

The advantages of the pZeroBG vector prompted us to test a pooled cloning approach of multiple guide oligos in a single-tube reaction. Due to high cloning efficiencies which carry zero background, we reasoned that the pZeroBG vector would be particularly useful for high-throughput cloning of binary CRISPR vectors harboring multiple different CRISPR guides. Annealed multiple oligo pairs for each different guide were subjected to ligation with AarI-treated pZeroBG vector (Fig. 3a). The pooled ligation reaction was transformed to DH5a and all resulting colonies were scraped and pooled into a single culture tube for overnight growth with suitable antibiotics (Fig. 3a). Plasmid DNA was isolated from the overnight-grown pooled culture for NGS validation to evaluate representation of the multiple guide species before moving on to the next step including *Agrobacterium* transformation (Fig. 3a).

In a pilot experiment, we conducted pooled cloning using only two guide oligo pairs of *FT* and *PDS3* targets with AarI-treated pZeroBG vector in the same tube. Twenty colonies randomly chosen from the pooled transformation were individually cultured overnight. Plasmid DNA from each culture was isolated and subjected to Sanger sequencing analysis to evaluate guide representation. Nine out of twenty colonies represent *PDS3*-guide sequences in their sgRNA expression cassettes whereas eleven colonies display *FT*-guide sequences (Fig. S1), suggesting no significant cloning bias exists a given guide and that guide representation is maintained in the pooled cloning process Then we carried out pool cloning (Fig. 3a) including eight different guide sequences with either pZeroBG-RpCas9 or pJY-RpCas9 vector. NGS analyses performed on the plasmid libraries, which were isolated from pool cultures, revealed that both vector platforms generated all eight guide reads (Fig. 3b). pZeroBG-RpCas9 appeared to be better than pJY-RpCas9 for even representation as gauged by the differences between the numbers of transformants with the most and least represented guide sequences for pZeroBG and pJY-RpCas9 are 2.87 × and 3.55 ×, respectively (Fig. 3b). More importantly, pZeroBG vector did not produce any no-guide reads in the pool (Fig. 3b), consistent with the previous observation (Table 2), while more than 10% of reads represented background for pooled cloning using pJY-RpCas9 (Fig. 3b). These results demonstrate that pZeroBG vector is superior to canonical CRISPR binary vectors that allow for higher efficiencies for pooled cloning strategies to make sgRNA libraries.

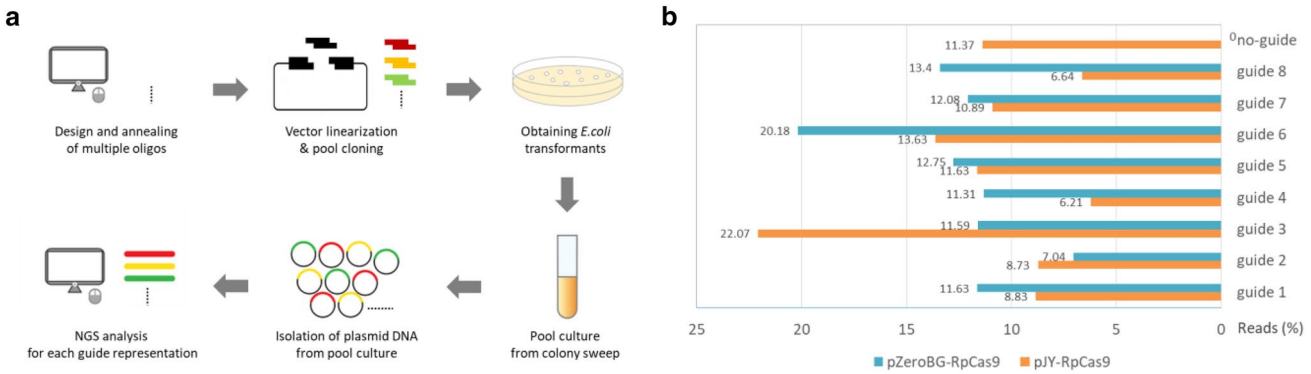

Fig. 3 NGS validations for pool cloning using either pZeroBG-RpCas9 or pJY-RpCas9 vector. **a** Schematic representation of pool cloning and NGS validation. **b** The results of NGS validation for each guide representation of pool cloning depending on using either pZeroBG-RpCas9 or pJY-RpCas9. *X* axis represents the ratios (%) of target amplicon sequencing reads out of total reads as described in method section. *Y* axis represents each guide sequences analyzed including intact vector sequence (denoted by 'no-guide'). The % values of each guide representation are indicated next to the bar

Testing the integrity of the pZeroBG-dependent library during Agrobacteria transfer and subculture for plant transformation

The small binary vector library representing multiple guides established using pZeroBG are required to be transferred into *Agrobacterium* host cells to finally be able to infect plants. To test the integrity of the library in the following steps, we simulated the conditions that the library is exposed to before infecting plants. The plasmid DNA isolated from *E. coli* pool culture were transformed to GV3101 strain to generate *Agrobacteria* transformants (Fig. 4a). Dozens of scoops were subject to a single seed culture for overnight growth (Fig. 4a). On the following day, the resulting stationary-growth phase ($OD_{600} = 2.0$) culture was diluted to 10 × subculture volume and was allowed to grow for 4–5 h until the *Agrobacteria* culture suspension reached log-growth phase ($OD_{600} = 0.8$) (Fig. 4a). To test the integrity of the sgRNA library during the transfer between host bacterial cells or subsequent subculture, plasmid DNA was isolated and subjected to NGS analysis from cells harvested at the stationary phase of seed culture and at the log-growth phase of subculture (Fig. 4a).

Tested plasmid DNA harboring eight different guide sequences from *E. coli* cells were pool transformed to *Agrobacteria* cells. NGS analyses for each guide representation were carried out on the seed culture (Agro 1) and the subsequent subculture (Agro 2). All guide sequences were evenly represented according to bacterial cell transfer and the continued growth (Fig. 4b). The pattern of the each guide's reads including those most represented and least represented guides was also maintained throughout the subsequent subculture (Fig. 4b). A new batch of the pool cloning using the same eight different guide oligos and pZeroBG vector were conducted to examine the reproducibility of the library integrity. Likewise, all eight different guides were represented evenly and the read pattern was maintained throughout the following steps in the new batch (Fig. 4c). However, the pattern of the each guide's reads was different from the previous batch. For example, the second batch results show that the guide 2 and the guide 1 are the most represented and the least represented one, respectively (Fig. 4c), which are not consistent with the first batch results. We have tested pooled cloning of up to twenty-four guides at a time and did not find any missing guides or significant distortion of guide representation (Fig. S2). The data suggest that constructing libraries representing multiple guides using pZeroBG vector is feasible and that the resulting library is appropriate for the following procedures for functional screen via plant genome editing.

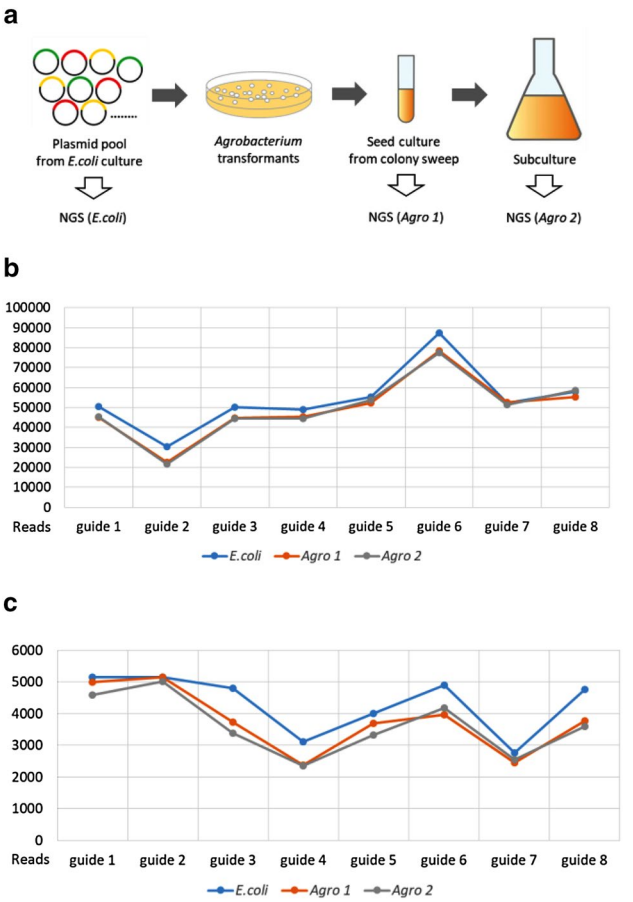

Fig. 4 NGS validation for the integrity of the sgRNA library. **a** Schematic representation of the library transfer between host bacterial cells, and subsequent subculture. Plasmid DNA pool was isolated for NGS analyses from three checkpoints: *E. coli* denotes *E. coli* culture; *Agro 1* denotes *Agrobacteria* seed culture from colony sweep; *Agro 2* denotes *Agrobacteria* subculture thereafter. **b** The results of NGS validation for each guide representation from the 1st batch of experiment from 8 guides-pooled cloning. **c** The results of NGS validation for each guide representation from the 2nd batch of the same experiment. *X* axis represents each guide sequences analyzed. *Y* axis represents number of target amplicon sequencing reads. The values of each guide representation are indicated by dots

Conclusions

We have established pZeroBG-CRISPR binary vector system by inserting *ccdB* gene between AarI recognition sites within its guide sequence cloning cassette in conventional CRISPR binary vector for plant transformation. The advantage of using the pZeroBG system over the conventional vector is obvious, ensuring highly efficient guide cloning in that there is no 'background' bacterial clones after *E. coli* transformation of the ligation products, which would have been generated in case of using conventional vectors.

The absolute zero-background feature of the pZeroBG vector prompted us to design a pooled cloning strategy

of multiple guide oligos in a single-tube reaction. As expected with zero background, pZeroBG vector was shown to be particularly useful for constructing binary vector libraries representing multiple sgRNAs. Those constructed CRISPR libraries we tested, aiming to infect plants, seems intact, representing all the guide sequences we intended to incorporate in the libraries. The integrity of the libraries appears to be well maintained during the *Agrobacteria* transformation and the subsequent subculture for proper plant transformation.

Depending on size of the pools for the library construction in a single-tube reaction, all-in-one digestion-ligation reaction (Weber et al. 2011) could be also applied in combination with pZeroBG-dependent cloning for further enhanced high-throughput library construction. Once the agrobacterium pool harboring sgRNA libraries is delivered to plants, each guide representation in resulting transgenic plants would most likely conform to the original representation in bacterial libraries (Depicker et al. 1985). However, it would be interesting, although it seems rare, to test if there occur any alterations of the integrity of guide representation possibly from the co-transformation events of different T-DNAs in a single plant cell as previously reported (De Buck et al. 2009; Meng et al. 2017).

Constructing CRISPR library aided by pZeroBG platform might be quite useful in certain cases for plant functional genomics, for example: when expected phenotypes are apparent but it is not clear which guide sequences or which target genes among candidates are responsible for the phenotype; and when functional CRISPR screening is required for unknown elements, unknown domain mapping, or unknown amino-acid changes, etc. According to Butt et al. (2019) their CRISPR library for plant screen contained 119 sgRNAs, targeting all possible PAM-adjacent sites in the whole coding sequence of the target protein to interrogate or evolve its molecular function. However, constructing the library would have been much easier if pZeroBG vector system were adapted for the study because those 119 guide sequences had been individually cloned using conventional CRISPR binary vector and validated one by one by Sanger sequencing before pooling each plasmids (Butt et al. 2019).

CRISPR screening using binary vector libraries in plant and agricultural science has great potential and gaining more promising utility with the development of CRISPR base editors, and relaxed-PAM derivatives such as xCas9 (Hu et al. 2018) and Cas9-NG (Nishimasu et al. 2018). Our results suggest that the pZeroBG vector system will greatly enhance the applications of CRISPR screen approaches when applied to relevant library construction in combination with such CRISPR derivatives.

Acknowledgements We thank all members of Center for Genome Engineering for support. We thank J. Suh for comments on the manuscript. This work was supported by the research fund from Institute for Basic Science, Republic of Korea (IBS-R021-D1).

Author contributions J-YY and J-SK conceived and supervised the project; J-YY designed the experiments; J-YY generated all the constructs; J-YY and S-TK performed the library construction and analyzed NGS data; S-GK contributed to the cloning of a main construct; J-YY wrote the main part of manuscript; All the authors analyzed the results and commented on the manuscript.

Compliance with ethical standards

Conflict of interest The authors declare no competing interests.

References

Barrangou R, Fremaux C, Deveau H et al (2007) CRISPR provides acquired resistance against viruses in prokaryotes. Science. https://doi.org/10.1126/science.1138140

Bernard P, Couturier M (1992) Cell killing by the F plasmid CcdB protein involves poisoning of DNA-topoisomerase II complexes. J Mol Biol 226:735–745

Bernard P, Gabant P, Bahassi EM, Couturier M (1994) Positive-selection vectors using the F plasmid ccdB killer gene. Gene 148:71–74

Butt H, Eid A, Momin AA et al (2019) CRISPR directed evolution of the spliceosome for resistance to splicing inhibitors. Genome Biol. https://doi.org/10.1186/s13059-019-1680-9

Cai Y, Chen L, Liu X et al (2015) CRISPR/Cas9-mediated genome editing in soybean hairy roots. PLoS ONE. https://doi.org/10.1371/journal.pone.0136064

Clough SJ, Bent AF (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. Plant J 16:735–743

Curtin SJ, Xiong Y, Michno J-M et al (2018) CRISPR/Cas9 and TALENs generate heritable mutations for genes involved in small RNA processing of Glycine max and Medicago truncatula. Plant Biotechnol J 16:1125–1137. https://doi.org/10.1111/pbi.12857

De Buck S, Podevin N, Nolf J et al (2009) The T-DNA integration pattern in Arabidopsis transformants is highly determined by the transformed target cell. Plant J 60:134–145. https://doi.org/10.1111/j.1365-313X.2009.03942.x

De Pater S, Klemann BJPM, Hooykaas PJJ (2018) True gene-targeting events by CRISPR/Cas-induced DSB repair of the PPO locus with an ectopically integrated repair template. Sci Rep. https://doi.org/10.1038/s41598-018-21697-z

Depicker A, Herman L, Jacobs A et al (1985) Frequencies of simultaneous transformation with different T-DNAs and their relevance to the Agrobacterium/plant cell interaction. MGG Mol Gen Genet 201:477–484. https://doi.org/10.1007/BF00331342

Du H, Zeng X, Zhao M et al (2016) Efficient targeted mutagenesis in soybean by TALENs and CRISPR/Cas9. J Biotechnol. https://doi.org/10.1016/j.jbiotec.2015.11.005

Feng C, Yuan J, Wang R et al (2016) Efficient targeted genome modification in maize using CRISPR/Cas9 system. J Genet Genom. https://doi.org/10.1016/j.jgg.2015.10.002

Feng Z, Zhang B, Ding W et al (2013) Efficient genome editing in plants using a CRISPR/Cas system. Cell Res. https://doi.org/10.1038/cr.2013.114

Ford K, McDonald D, Mali P (2019) Functional genomics via CRISPR–Cas. J Mol Biol 43:48–65

Hu JH, Miller SM, Geurts MH et al (2018) Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. Nature. https://doi.org/10.1038/nature26155

Kang BC, Yun JY, Kim ST et al (2018) Precision genome engineering through adenine base editing in plants. Nat Plants 4:427–431. https://doi.org/10.1038/s41477-018-0178-x

Kim H, Kim ST, Ryu J et al (2016) A simple, flexible and high-throughput cloning system for plant genome editing via CRISPR-Cas system. J Integr Plant Biol. https://doi.org/10.1111/jipb.12474

Li J-F, Norville JE, Aach J et al (2013) Multiplex and homologous recombination–mediated genome editing in Arabidopsis and *Nicotiana benthamiana* using guide RNA and Cas9. Nat Biotechnol. https://doi.org/10.1038/nbt.2654

Liu X, Xie C, Si H, Yang J (2017) CRISPR/Cas9-mediated genome editing in plants. Methods 121:94–102

Mao Y, Zhang H, Xu N et al (2013) Application of the CRISPR–Cas system for efficient genome engineering in plants. Mol Plant. https://doi.org/10.1093/mp/sst121

Marraffini LA (2015) CRISPR-Cas immunity in prokaryotes. Nature 526:55s

Meng X, Yu H, Zhang Y et al (2017) Construction of a genome-wide mutant library in rice using CRISPR/Cas9. Mol Plant 10:1238–1241

Nishimasu H, Shi X, Ishiguro S et al (2018) Engineered CRISPR-Cas9 nuclease with expanded targeting space. Science. https://doi.org/10.1126/science.aas9129s

Schuster A, Erasimus H, Fritah S et al (2019) RNAi/CRISPR screens: from a pool to a valid hit. Trends Biotechnol 37:38–65

Shan Q, Wang Y, Li J et al (2013) Targeted genome modification of crop plants using a CRISPR-Cas system. Nat Biotechnol. https://doi.org/10.1038/nbt.2650

Shan Q, Wang Y, Li J, Gao C (2014) Genome editing in rice and wheat using the CRISPR/Cas system—supplementary tables. Nat Protoc. https://doi.org/10.1038/nprot.2014.157

Tsutsui H, Higashiyama T (2017) PKAMA-ITACHI vectors for highly efficient CRISPR/Cas9-mediated gene knockout in Arabidopsis thaliana. Plant Cell Physiol 58:46–56. https://doi.org/10.1093/pcp/pcw191

Wang ZP, Xing HL, Dong L et al (2015) Egg cell-specific promoter-controlled CRISPR/Cas9 efficiently generates homozygous mutants for multiple target genes in Arabidopsis in a single generation. Genome Biol. https://doi.org/10.1186/s13059-015-0715-0

Weber E, Engler C, Gruetzner R et al (2011) A modular cloning system for standardized assembly of multigene constructs. PLoS ONE. https://doi.org/10.1371/journal.pone.0016765

Woo JW, Kim J, Il Kwon S et al (2015) DNA-free genome editing in plants with preassembled CRISPR-Cas9 ribonucleoproteins. Nat Biotechnol 33:1162–1164. https://doi.org/10.1038/nbt.3389

Xie K, Yang Y (2013) RNA-Guided genome editing in plants using a CRISPR-Cas system. Mol Plant. https://doi.org/10.1093/mp/sst119

Yan L, Wei S, Wu Y et al (2015) High-efficiency genome editing in Arabidopsis using YAO promoter-driven CRISPR/Cas9 system. Mol Plant. https://doi.org/10.1016/j.molp.2015.10.004

Yin K, Gao C, Qiu JL (2017) Progress and prospects in plant genome editing. Nat Plants 3:17107

Zhang Y, Massel K, Godwin ID, Gao C (2018) Applications and potential of genome editing in crop improvement. Genome Biol 19:210. https://doi.org/10.1186/s13059-018-1586-y

Zhang Y, Xie X, Liu YG et al (2017) CRISPR/Cas9-based genome editing in plants. In: Progress in molecular biology and translational science. Academic Press, New York