



Chapter 2

Efficient Homologous Recombination-Mediated in Planta Gene Targeting by Egg-Cell-Specific Expression of *Staphylococcus aureus* Cas9 from Arabidopsis

Felix Wolter, Teng-Kuei Huang, and Holger Puchta

Abstract

Site-specific genome engineering approaches were greatly facilitated by the recent emergence of the CRISPR-Cas system, enabling precise induction of DNA double-strand breaks. However, up to now its application was mostly restricted to nonhomologous end-joining-mediated targeted mutagenesis. In contrast, precise genome modifications using a suitable donor sequence for homologous still pose a particular challenge in plants, as NHEJ is the dominant repair mechanism for DSBs in somatic cells. To achieve efficient HR-mediated genome modifications in plants, we recently developed the in planta gene targeting (ipGT) system, which works via the induction of DSBs by Cas9 to activate the target and the targeting vector at the same time, making it independent of high transformation efficiencies. Here, we describe an updated protocol of ipGT for the model plant Arabidopsis, taking into account our recent improvements based on egg-cell-specific expression of *Staphylococcus aureus* Cas9.

Key words CRISPR-Cas, Double-strand break, Gene targeting, Homologous recombination, Non-homologous end joining

Abbreviations

CRISPR-Cas	Clustered regularly interspaced short palindromic repeats-CRISPR associated
DSB	Double-strand break
GT	Gene targeting
ipGT	In planta GT
NHEJ	Nonhomologous end joining
PPT	Phosphinothricin

1 Introduction

Gene targeting (GT) is a genome engineering method that achieves desired genetic modifications through an endogenous DNA repair

system—homologous recombination. However, due to its low efficiency, GT was mostly demonstrated as a proof of concept in plants up to now. To stimulate HR-based GT in the target locus, the most significant and well-known method is the induction of a double-strand break (DSB) in the target locus [1]. Mainly, there are two distinctive goals in genome modification via induced DSBs: targeted mutagenesis and GT. DSBs used in genome editing result in mutagenesis mainly through the error-prone nonhomologous end-joining (NHEJ) repair pathway to obtain small deletions or insertions [2]. Differently, GT requires an additional homologous sequence to serve as template for homologous recombination, termed HR donor. It contains the desired modification in the center region and the homologous sequences in the flanking regions. Successful GT leads to transfer of sequence information from the DNA donor to the target locus in a predefined manner, enabling precise insertion, replacement, or substitution of a DNA fragment, which is not feasible through general NHEJ-based genome editing. For a more detail discussion of the development of the GT in plants, *see* [3, 4].

The development of programmable nucleases such as CRISPR-Cas (clustered regularly interspaced short palindromic repeats-CRISPR associated), which enable precise DSB induction at desired loci, has become a routine method for mutagenesis purpose in plants [5, 6]. However, despite DSB induction, GT efficiency remains too low for routine usage. The main limitations are the lack of efficient transformation systems for most crops and the general low frequency of HR in plants. To address this problem, we developed the *in planta* GT (ipGT) approach that is independent on high transformation efficiencies, since the HR donor is stably integrated into the genome. By inducing DSBs at the flanks of the HR donor simultaneously to DSB induction in the target, the donor can be released from the genome to activate it for HR (*see* Fig. 1). This way, GT efficiency could be enhanced to percentile range [7, 8]. Later on, it was demonstrated in *Arabidopsis* that *Staphylococcus aureus* Cas9 (SaCas9) possesses higher DSB-based mutagenesis activity than *Streptococcus pyogenes* Cas9 (SpCas9) [9]. Therefore, SaCas9 became a more promising molecular tool for the induction of DSB in *Arabidopsis*.

The application of an egg-cell-specific promoter-driven SaCas9 was first demonstrated in genome editing, showing the enhancement of NHEJ-mediated mutagenesis [10–12]. More recently, two studies demonstrated that using an egg-cell promoter for Cas9 expression resulted in higher GT efficiency compared to other promoters tested, which included constitutive, meristem-specific, and pollen-specific promoters [13, 14]. By combining our ipGT approach with the high-efficient SaCas9 driven by an egg-cell-specific promoter, we obtained heritable GT events with high efficiencies up to 6% of T2 *Arabidopsis* seeds [13]. These studies

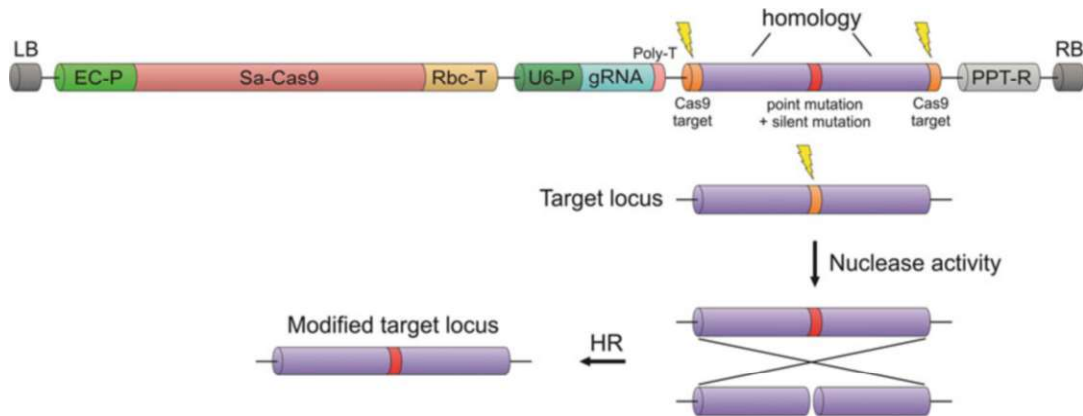


Fig. 1 Outline of the ipGT approach. The T-DNA is stably integrated into the plant genome and contains four elements: an egg-cell-specific *Staphylococcus aureus* Cas9 cassette, a gRNA cassette, a HR donor sequence flanked by gRNA recognition sites, and a plant selection marker. Upon expression of the nuclease, three DSBs are induced simultaneously, one in the target locus and two flanking the donor leading to its release. If the cells use the donor for HR-mediated repair, the desired genome modification is achieved

suggest that the egg cell of plants might have a more active HR DNA repair system and can be a valuable cell type to achieve successful GT. This was also confirmed by a recent study showing efficient in planta gene targeting using Cas12a from *Lachnospiraceae* bacterium ND2006 (LbCas12a) instead of SaCas9 [15]. In principle, our egg-cell-specific ipGT protocol is not only applicable to the model plant *Arabidopsis* but can also be utilized in other transformable crop species.

Here, we describe the procedure to perform *Staphylococcus aureus* Cas9-mediated ipGT approach in *Arabidopsis thaliana*, enabling stable heritable genome modifications, including integration of a tag or *cis*-element to a desired locus, or precise sequence modifications such as amino acid substitutions.

2 Materials

2.1 Plasmids

All plasmids are available directly from the authors and full sequence information is available at www.botanik.kit.edu/crispr. Some plasmids have also been deposited at the Arabidopsis Biological Resource Center (ABRC).

2.1.1 pDe-SaCas9-EC

Binary vector for stable plant transformation via *A. tumefaciens*, conferring plant resistance to phosphinothricin (PPT). It contains an egg-cell-specific expression system (*EC1.1/1.2* promoter) for Cas9 from *Staphylococcus aureus*. It is also a Gateway destination vector containing the respective attachment sites flanking a *ccdB* gene for insertion of the gRNA expression cassette. Confers plant

resistance to PPT. The vector's multiple cloning site also contains restriction sites suitable for insertion of the HR donor sequence.

2.1.2 *pDe-Sa-Cas9* Identical to pDe-Sa-Cas9-EC with the exception that the Cas9 expression system is constitutive (*PcUBI 4-2* promoter) and confers plant resistance to kanamycin.

2.1.3 *pEn-Sa-Chimera* Gateway entry vector designed to take up annealed oligos to create a programmed gRNA. The respective Gateway attachment sites enable transfer of the programmed gRNA into the destination vector.

2.2 Organisms

1. *Escherichia coli*, standard cloning strain for all cloning steps; *ccdB*-resistant strain for propagation of pDe-Cas9 (e.g., DB3.1 [16]; see Note 1).
2. *A. thaliana* plants, either wild-type (e.g., Col-0) or any transformable mutant.
3. *Agrobacterium tumefaciens*, any conventional transformation strain, e.g., GV3101 [17].

2.3 Reagents

1. Restriction enzyme BbsI, additional restriction enzymes as required (see Fig. 1 and Note 2).
2. T4 Ligase for conventional cloning steps.
3. Proofreading DNA polymerase for the generation of the donor sequence.
4. A robust Taq polymerase for *E. coli* colony PCRs and for screening of putative GT plants.
5. Gateway LR Clonase II (only available from Thermo Fisher Scientific, supplied with proteinase K).
6. LB medium (for *E. coli*): 10 g/L peptone, 5 g/L yeast extract, 10 g/L NaCl. Solid media: Add additionally 7.5 g/L agar. For selection media, add ampicillin (100 mg/L) or spectinomycin (100 mg/L).
7. YEB medium (for *A. tumefaciens*): 5 g/L beef extract and 5 g/L peptone. 1 g/L yeast extract, 5 g/L sucrose, 439 mg/L MgSO₄, and 7.5 g/L agar for solid media.
8. Germination medium (GM): 4.9 g/L Murashige & Skoog, 10 g/L sucrose, pH 5.7, and 8 g/L agar. For selection conditions, add PPT (6 mg/L) or kanamycin (30 mg/L).
9. TE buffer: 10 mM Tris-HCl and 1 mM EDTA at pH 8.

3 Methods

3.1 Experimental Design Considerations

1. Identify potential target sites for *Staphylococcus aureus* Cas9 in your target locus close to the desired modification containing the required PAM (*see Note 4*). Exclude sequences containing a stretch of five or more Ts as this terminates transcription by RNA polymerase III.
2. Design the HR donor sequence containing the desired modification. This is the most crucial step and should be given thorough planning. We recommend to use 600–800 bp homology on both flanks of the cleavage site (*see Note 5*). For excision of the donor sequence, add the gRNA target sequence including the PAM to both ends of the donor sequence (*see Note 6*). To exclude Cas9 activity on the donor, introduce silent mutations at the corresponding target site on the donor (*see Note 7*). In order to facilitate the later screening procedure, your modification or silent mutations should allow distinction between WT and GT events by PCR (*see Note 8*). Finally, suitable restriction sites on the very end of the HR donor are required for insertion of the HR donor into your T-DNA.
3. Check the feasibility of generating your HR donor sequence. Possible methods to generate the HR donor are overlap extension PCR, Gibson assembly, or gene synthesis (*see Note 9*).

3.2 Generating the Required T-DNAs

1. Order the required oligonucleotides for your target sequence. For SaCas9, the fw oligo should contain 20–21 nt upstream of the PAM with ATTG added at the 5' end. The second oligo should contain the reverse complement of the target sequence with AAAC added at the 5' end.
2. Dilute and mix your oligos in ddH₂O to a final concentration of 2 pmol/μL for each oligo in a total volume of 50 μL. Incubate for 5 min at 95 °C and put at room temperature for an additional 10 min for annealing.
3. Digest pEn-Sa-Chimera to completion with BbsI. Extract the 3 kb band by gel purification (*see Note 10*) and dilute the final concentration to 5 ng/μL.
4. Set up a ligation reaction containing 2 μL digested vector, 3 μL annealed oligos, 1 μL T4 ligase buffer, 1 μL T4 ligase, and 3 μL ddH₂O, and incubate as recommended by the vendor. Transform into *E. coli* and select for colonies on ampicillin-containing LB plates. Colonies can be screened for insert by colony PCR using your fw oligo and M13-rev as primers.
5. Purify plasmids from a small number (2–4) of correct clones and validate by sequencing with primer SS42.

6. Transfer the correct gRNA expression cassette to both pDe-Sa-Cas9 and pDe-Sa-Cas9-EC by Gateway cloning. pDe-Sa-Cas9-EC is for the actual GT construct. pDe-Sa-Cas9 is only required to check the cleavage activity at the chosen target sequence. Set up a reaction with 100 ng entry vector, 300 ng destination vector, 4 μ L TE buffer, and 1 μ L LR Clonase II in a total volume of 10 μ L, and incubate for 2–3 h at room temperature. Before transformation into *E. coli*, stop the reaction by adding 1 μ L proteinase K and incubation at 37 °C for 10 min. Select on LB plates containing spectinomycin. Screening for correct insert can be achieved by colony PCR with primers SS42/SS102 (for both Gateway reactions).
7. Isolate plasmids from 2–4 PCR-positive colonies each. A control restriction digestion is recommended to check for the integrity of the plasmid. For Gateway reactions involving pDe-Sa-Cas9, the enzymes AflII and NheI can be used generating bands at approximately 5.9, 5.0, and 3.8 kb. For Gateway reactions involving pDe-Sa-Cas9-EC, the enzymes HindIII and NheI can be used generating bands at approximately 5.9, 3.3, 1.9, 1.8, and 1.2 kb.
8. Add your GT donor sequence to the plasmid obtained after insertion of the gRNA cassette into pDe-Sa-Cas9-EC using the restriction sites included on the donor sequence. Identify correct clones by a suitable colony PCR and restriction digestion depending on your donor sequence (*see Note 10*).
9. Verify the correctness of both of your final constructs by full sequencing of the Cas9 cassette, gRNA cassette, and donor. The final constructs are the GT construct (pDe-Sa-Cas9-EC + gRNA+donor) and the construct to check cleavage efficiency (pDe-Sa-Cas9 + gRNA).
10. Transform your final verified plasmids into *A. thaliana* (e.g., [18]).

3.3 Obtaining Heritable GT Plants (See also Fig. 2)

1. To generate primary transformants, sow seeds of the transformation on germination medium containing PPT for your GT construct and kanamycin for your construct to check cleavage efficiency (*see Note 12*).
2. For the GT construct, pick at least 40 T1 plants for further cultivation in soil (*see Note 13*). Extract DNA from a leaf and verify the presence of the correct construct by PCR.
3. To check for cleavage activity at your chosen target site, extract DNA from 20 primary transformants carrying the cleavage efficiency construct after 3 weeks. Analyze mutagenesis efficiency by, e.g., TIDE, RE assay, T7E1, HRM, or NGS (described in detail elsewhere: [19, 20]).

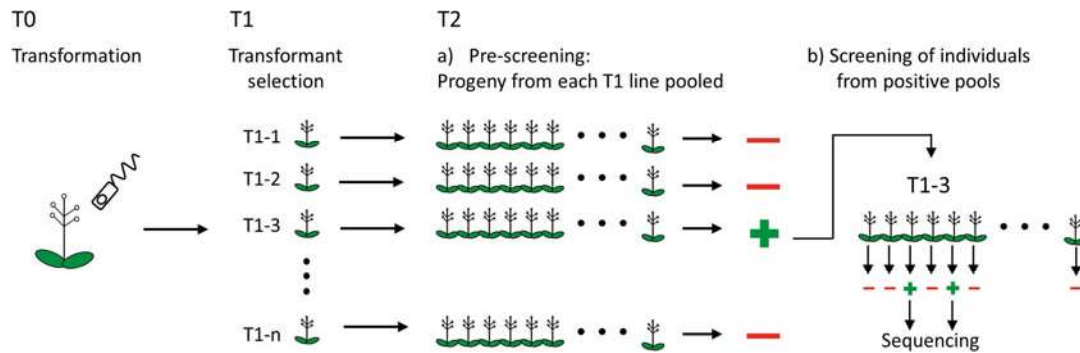


Fig. 2 Overview of the two-step screening procedure to identify GT-positive plants

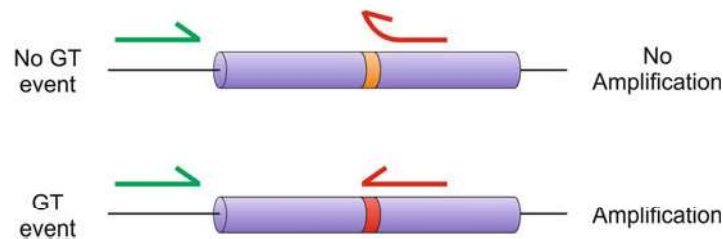


Fig. 3 PCR-based identification of GT events. One primer should be placed outside the homologous sequence region, the other primer inside the homology in such a way that the specific sequence modification is recognized (see also **Note 14**)

4. Harvest the seeds for each of your T1 plants carrying the GT construct separately, and sow a small amount of seeds (~100 per line) on GM medium.
5. After 14 days of growth, isolate DNA from leaf material of 100 plants per line as a pool (one big leaf per plant). Screen for a positive GT signal via PCR (see Fig. 3 and **Note 14**).
6. From those T1 pools identified in **step 5**, isolate DNA from each plant separately after another week of growth. Now, analyze individual T2 plants for heritable GT by PCR. Transfer GT-positive T2 plants to soil and cultivate to maturity.
7. Harvest seeds from each GT-positive T2 line separately and screen for loss of the T-DNA in the T3 generation. Confirm successful GT again to assure stable inheritance and screen for loss of the transgene.

4 Notes

1. All cloning steps require a conventional *E. coli* strain, e.g., DH5 α . The *ccdB*-resistant strain is only required to propagate pDe-CAS9 before removing the *ccdB* gene in a gateway reaction.

2. Best choose a restriction site in the multiple cloning site, and take care that the restriction site is not present in your HR donor sequence.
3. We routinely use Q5 DNA polymerase (NEB) for proofreading purposes and DreamTaq polymerase (Thermo Fisher Scientific) for colony and plant screening PCRs.
4. We highly recommend to use NNGRRRT as PAM of SaCas9, even though NNGRRV can also be cleaved. Furthermore, avoid target sites containing too low GC content. There are many software tools available to aid in target choice, reviewed in detail in [21].
5. The actual size can vary greatly depending on your experiment. Generally, longer homologies (0.8–1 kb) should improve HR frequencies. However, if the homology contains parts of a promoter region, one has to be aware of potential expression from the T-DNA itself.
6. Using a vector set which is capable of expressing more than one sgRNA [8], it is also possible to have different Cas9 target sites for the target locus and for releasing the donor sequence.
7. A silent mutation disrupting the PAM is ideal. Consult Ran et al. [22] which PAMs are still cleavable by SaCas9. If disrupting the PAM is not possible, two to three silent mutations in the seed region are sufficient. If you require a confirmation of your GT events via Southern blot, include a restriction site in the silent mutations that is not present in the genomic sequence or vice versa.
8. Three mismatches immediately at the 3' end should be sufficient to prevent amplification from WT alleles.
9. Having the donor sequence synthesized is the easiest, yet most expensive, method. However, overlap PCR or Gibson assembly may be challenging and time-consuming due to reoccurring sequence elements flanking the construct.
10. To save time and effort, it is also possible to purify the BbsI digest by a PCR purification kit, as the excised fragment is very small. However, cloning efficiency is higher using gel purification, because all of the unwanted fragments are removed.
11. The orientation of the donor construct within the T-DNA is irrelevant. Thus, a single restriction enzyme can be used for this cloning step. However, this requires dephosphorylation which is not required when two enzymes are used.

12. When floral dipping is used to transform *Arabidopsis*, cefotaxime should be added to the germination medium to limit *Agrobacterium* growth.
13. All numbers given in this paragraph are based on experience. It might be necessary to scale them up to detect a GT event.
14. Proper functionality of the PCR assay is crucial. Thus, optimization of primers and PCR conditions will be necessary. For details on primer design and PCR conditions, *see* also [23]. Note that when analyzing the pools, the GT signal is strongly diluted so a large number of cycles (at least 40) is required.

References

1. Puchta H, Dujon B, Hohn B (1996) Two different but related mechanisms are used in plants for the repair of genomic double-strand breaks by homologous recombination. *Proc Natl Acad Sci U S A* 93:5055–5060. <https://doi.org/10.1073/pnas.93.10.5055>
2. Puchta H (2005) The repair of double-strand breaks in plants: mechanisms and consequences for genome evolution. *J Exp Bot* 56:1–14. <https://doi.org/10.1093/jxb/eri025>
3. Huang T-K, Puchta H (2019) CRISPR/Cas-mediated gene targeting in plants: finally a turn for the better for homologous recombination. *Plant Cell Rep* 38:443–453. <https://doi.org/10.1007/s00299-019-02379-0>
4. Puchta H, Fauser F (2013) Gene targeting in plants: 25 years later. *Int J Dev Biol* 57:629–637. <https://doi.org/10.1387/ijdb.130194hp>
5. Kumlehn J, Pietralla J, Hensel G, Pacher M, Puchta H (2018) The CRISPR/Cas revolution continues: from efficient gene editing for crop breeding to plant synthetic biology. *J Integr Plant Biol* 60:1127–1153. <https://doi.org/10.1111/jipb.12734>
6. Scheben A, Wolter F, Batley J, Puchta H, Edwards D (2017) Towards CRISPR/Cas crops - bringing together genomics and genome editing. *New Phytol* 216:682–698. <https://doi.org/10.1111/nph.14702>
7. Fauser F, Roth N, Pacher M, Ilg G, Sanchez-Fernandez R, Biesgen C, Puchta H (2012) In planta gene targeting. *Proc Natl Acad Sci U S A* 109:7535–7540. <https://doi.org/10.1073/pnas.1202191109>
8. Schiml S, Fauser F, Puchta H (2014) The CRISPR/Cas system can be used as nuclease for in planta gene targeting and as paired nickases for directed mutagenesis in *Arabidopsis* resulting in heritable progeny. *Plant J* 80:1139–1150. <https://doi.org/10.1111/tpj.12704>
9. Steinert J, Schiml S, Fauser F, Puchta H (2015) Highly efficient heritable plant genome engineering using Cas9 orthologues from *Streptococcus thermophilus* and *Staphylococcus aureus*. *Plant J* 84:1295–1305. <https://doi.org/10.1111/tpj.13078>
10. Mao Y, Zhang Z, Feng Z, Wei P, Zhang H, Botella JR, Zhu J-K (2015) Development of germ-line-specific CRISPR-Cas9 systems to improve the production of heritable gene modifications in *Arabidopsis*. *Plant Biotechnol J* 14:519–532. <https://doi.org/10.1111/pbi.12468>
11. Wang Z-P, Xing H-L, Dong L, Zhang H-Y, Han C-Y, Wang X-C, Chen Q-J (2015) Egg cell-specific promoter-controlled CRISPR/Cas9 efficiently generates homozygous mutants for multiple target genes in *Arabidopsis* in a single generation. *Genome Biol* 16:144. <https://doi.org/10.1186/s13059-015-0715-0>
12. Yan L, Wei S, Wu Y, Hu R, Li H, Yang W, Xie Q (2015) High-efficiency genome editing in *Arabidopsis* using YAO promoter-driven CRISPR/Cas9 system. *Mol Plant* 8:1820–1823. <https://doi.org/10.1016/j.molp.2015.10.004>
13. Wolter F, Klemm J, Puchta H (2018) Efficient in planta gene targeting in *Arabidopsis* using egg cell-specific expression of the Cas9 nuclease of *Staphylococcus aureus*. *Plant J* 94:735–746. <https://doi.org/10.1111/tpj.13893>
14. Miki D, Zhang W, Zeng W, Feng Z, Zhu J-K (2018) CRISPR/Cas9-mediated gene targeting in *Arabidopsis* using sequential transformation. *Nat Commun* 9:1967. <https://doi.org/10.1038/s41467-018-04416-0>

15. Wolter F, Puchta H (2019) In planta gene targeting can be enhanced by the use of CRISPR /Cas12a. *Plant J* 100:1083–1094. <https://doi.org/10.1111/tbj.14488>.16
16. 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. [https://doi.org/10.1016/0022-2836\(92\)90629-X](https://doi.org/10.1016/0022-2836(92)90629-X)
17. Koncz C, Kreuzaler F, Kalman Z, Schell J (1984) A simple method to transfer, integrate and study expression of foreign genes, such as chicken ovalbumin and alpha-actin in plant tumors. *EMBO J* 3:1029–1037. <https://doi.org/10.1002/j.1460-2075.1984.tb01923.x>
18. Clough SJ, Bent AF (1998) Floral dip: a simplified method for agrobacterium-mediated transformation of *Arabidopsis thaliana*. *Plant J* 16:735–743. <https://doi.org/10.1046/j.1365-313x.1998.00343.x>
19. Brinkman EK, Chen T, Amendola M, van Steensel B (2014) Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res* 42:e168. <https://doi.org/10.1093/nar/gku936>
20. Zischewski J, Fischer R, Bortesi L (2017) Detection of on-target and off-target mutations generated by CRISPR/Cas9 and other sequence-specific nucleases. *Biotechnol Adv* 35:95–104. <https://doi.org/10.1016/j.biotechadv.2016.12.003>
21. Cui Y, Xu J, Cheng M, Liao X, Peng S (2018) Review of CRISPR/Cas9 sgRNA design tools. *Interdiscip Sci* 10:455–465. <https://doi.org/10.1007/s12539-018-0298-z>
22. Ran FA, Cong L, Yan WX, Scott DA, Gootenberg JS, Kriz AJ et al (2015) In vivo genome editing using *Staphylococcus aureus* Cas9. *Nature* 520:186–191. <https://doi.org/10.1038/nature14299>
23. Medrano RFV, de Oliveira CA (2014) Guidelines for the tetra-primer ARMS-PCR technique development. *Mol Biotechnol* 56:599–608. <https://doi.org/10.1007/s12033-014-9734-4>