

METHODOLOGY ARTICLE

Open Access



# Expansion of targetable sites for the ribonucleoprotein-based CRISPR/Cas9 system in the silkworm *Bombyx mori*

Yun-long Zou, Ai-jun Ye, Shuo Liu, Wen-tao Wu, Li-feng Xu, Fang-yin Dai and Xiao-ling Tong\*

## Abstract

**Background:** With the emergence of CRISPR/Cas9 technology, multiple gene editing procedures became available for the silkworm. Although binary transgene-based methods have been widely used to generate mutants, delivery of the CRISPR/Cas9 system via DNA-free ribonucleoproteins offers several advantages. However, the T7 promoter that is widely used in the ribonucleoprotein-based method for production of sgRNAs in vitro requires a 5' GG motif for efficient initiation. The resulting transcripts bear a 5' GG motif, which significantly constrains the number of targetable sites in the silkworm genome.

**Results:** In this study, we used the T7 promoter to add two supernumerary G residues to the 5' end of conventional (perfectly matched) 20-nucleotide sgRNA targeting sequences. We then asked if sgRNAs with this structure can generate mutations even if the genomic target does not contain corresponding GG residues. As expected, 5' GG mismatches depress the mutagenic activity of sgRNAs, and a single 5' G mismatch has a relatively minor effect. However, tests involving six sgRNAs targeting two genes show that the mismatches do not eliminate mutagenesis in vivo, and the efficiencies remain at useable levels. One sgRNA with a 5' GG mismatch at its target performed mutagenesis more efficiently than a conventional sgRNA with 5' matched GG residues at a second target within the same gene. Mutations generated by sgRNAs with 5' GG mismatches are also heritable. We successfully obtained null mutants with detectable phenotypes from sib-mated mosaics after one generation.

**Conclusions:** In summary, our method improves the utility and flexibility of the ribonucleoprotein-based CRISPR/Cas9 system in silkworm.

**Keywords:** CRISPR/Cas9, Ribonucleoprotein, Silkworm, Target repertoire, *BmBLOS2*, *BmGR66*

## Background

In the postgenomic era, genetically modified model organisms will become essential tools as researchers turn their attention to the functional dissection of the genome. The mulberry silkworm *Bombyx mori* is not only an important lepidopteran model but is also an

economically important insect in the silk industry. The genetic modification of this organism will therefore benefit basic research as well as silk production. Zinc-finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs) have been exploited for genomic engineering in *Bombyx mori* [1, 2]. However, difficulties in designing and engineering ZFNs and TALENs have greatly limited the application of gene editing in silkworms. With the development of CRISPR/Cas9, gene editing in silkworms has become far more tractable [3, 4].

Gene editing using CRISPR/Cas9 in silkworms has typically been performed with a binary transgenic

\*Correspondence: xltong@swu.edu.cn

State Key Laboratory of Silkworm Genome Biology; Key Laboratory of Sericultural Biology and Genetic Breeding, Ministry of Agriculture and Rural Affairs; College of Sericulture, Textile and Biomass Sciences, Southwest University, Chongqing 400715, People's Republic of China



system [3, 5–7]. This system requires two separate transgenic lines, one of which expresses universal Cas9 and the other a customized sgRNA. Unfortunately, new transgenic lines expressing specific sgRNAs must be established de novo for each gene locus. The procedure is time consuming and significantly restricts the flexibility of the CRISPR/Cas9 system. Even with the sgRNA and Cas9 transgenic lines in hand, at least two or three rounds of hybridization are necessary to obtain homozygous or compound heterozygous mutants [5]. Moreover, additional rounds of hybridization are required to eliminate the sgRNA and Cas9 incorporated alleles, because the random insertion of these cassettes may disrupt the expression of endogenous genes and complicate subsequent research [8].

To overcome the drawbacks of the binary transgenic system, researchers have developed two DNA-free methods using Cas9 mRNA and Cas9 protein. Protocols using the Cas9 protein are more effective than those using Cas9 mRNA because the protein acts immediately following injection without a translational delay [9, 10]. In the Cas9 ribonucleoprotein (RNP)-based method, the T7 promoter is often chosen for in vitro transcription of the sgRNA [10]. However, the T7 promoter requires a GG dinucleotide at the RNA transcriptional start site for efficient transcription. A typical 20-nucleotide targeting region in the sgRNA is therefore generated with the structure 5'GGN<sub>18</sub>, and only genomic sites containing a GG dinucleotide at the corresponding position can be targeted by the Cas9 RNP. Although the 5' GG rule can be relaxed to allow an adenine at either position [11], targetable genomic sequences are still limited. In addition, sgRNAs generated by T7 polymerase that initiate with 5'AG or 5' GA exhibit poor mutagenic activity, due in part to 5' end transcript heterogeneity [9].

To expand the number of targetable sites in the silkworm genome, we used the T7 promoter to add two supernumerary G residues to the 5' end of a 20-nucleotide sgRNA targeting sequence. The GGN<sub>20</sub> structure makes it possible to generate transcripts at high efficiency, and also provides a targeting sequence that can match a 20-nucleotide genomic target without a 5' GG motif. However, the supernumerary residues may result in a 1- or 2-nucleotide mismatch immediately upstream of the perfectly matching genomic target. We therefore assessed the impact of the supernumerary 5' GG residues on indel frequencies in vivo, and examined the transmission efficiencies for mutations generated using this system. Two well studied genes, *BmBLOS2* and *BmGR66*, were used as targets for proof-of-concept tests.

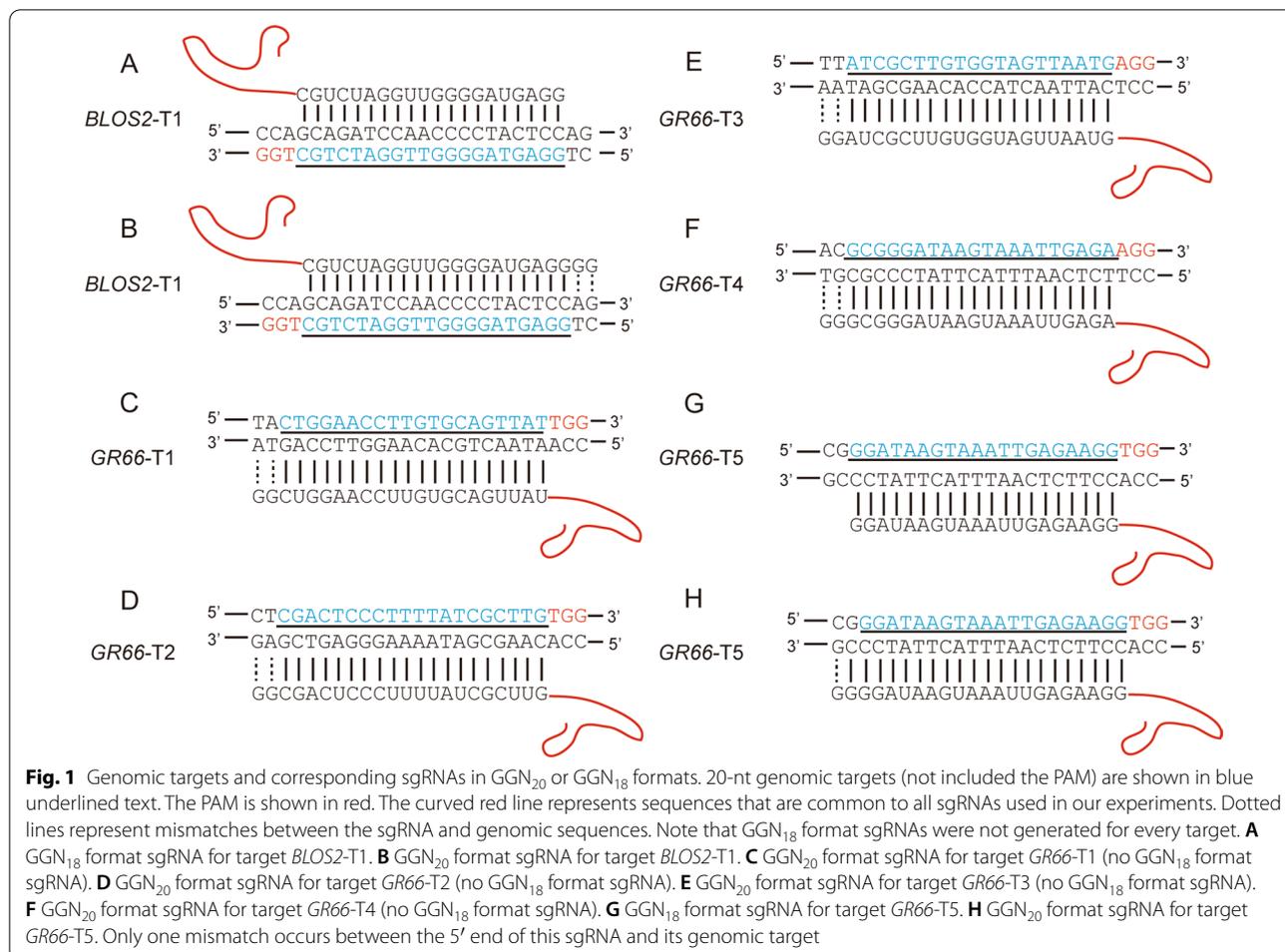
## Results

### Introduction of GG residues to the 5' end of an sgRNA does not eliminate cleavage

To determine whether it is feasible to expand the range of sgRNAs synthesized in vitro using the T7 promoter, we designed sgRNAs containing two GG residues at their 5' ends and used them to target genomic sequences that did or did not contain 5' GG. Two sgRNA variations were tested. In the first case, the sgRNA target region had a GGN<sub>18</sub> structure consisting of two GG residues and 18 additional nucleotides. GGN<sub>18</sub> sgRNAs were used to target genomic sites that contained a perfect match to the sgRNA targeting sequence, including the two 5' G residues. In the second case, the GG residues were added to a targeting sequence 20 nucleotides long, creating a GGN<sub>20</sub> structure. We refer to the GG residues on GGN<sub>20</sub> sgRNAs as “supernumerary”. GGN<sub>20</sub> sgRNAs were tested at genomic sites that did not contain a 5' GG sequence, but otherwise perfectly matched the 20 nucleotides in the target region.

Two genes were targeted to evaluate the performance of the GGN<sub>18</sub> and GGN<sub>20</sub> sgRNAs. *BmBLOS2* is located on the Z chromosome, and is responsible for the biosynthesis of urate granules, which accumulate in epidermal cells and make the larval integument opaque. Because mutations in this gene result in an easily detected oily skin phenotype, *BmBLOS2* is often used to confirm the efficacy of gene editing methods [1, 2, 4]. *GR66* is located on the third chromosome and has recently been identified as a key gustatory receptor responsible for the mulberry-specific feeding preference of the silkworm [7]. These two genes, for which null mutations generate detectable phenotypes, were used to test the efficacy of our strategy.

Two targets were selected for initial tests in *BmBLOS2* and *BmGR66* (designated *BLOS2-T1* and *GR66-T5*, respectively; Fig. 1). For both targets, sgRNAs in the GGN<sub>18</sub> and GGN<sub>20</sub> formats were synthesized using T7 RNA polymerase-mediated in vitro transcription. Note that the GGN<sub>18</sub> sgRNAs match their respective targets in these genes perfectly (Fig. 1A, G), while the supernumerary GG residues in the GGN<sub>20</sub> sgRNAs have two mismatches at *BLOS2-T1* (5' CT) and one mismatch at *GR66-T5* (5' CG) (Fig. 1B, H). Each sgRNA was separately mixed with Cas9 protein and incubated at room temperature to form a Cas9 protein/sgRNA complex, which was then injected into pre-blastoderm embryos. To examine indel frequencies, pools of 60 randomly selected injected embryos were harvested for genomic DNA extraction. Regions surrounding the sgRNA targets in *BmBLOS2* and *BmGR66* were PCR amplified and T7EN1 assays were performed to detect cleavage efficiencies. The results demonstrated that the GGN<sub>18</sub> sgRNAs

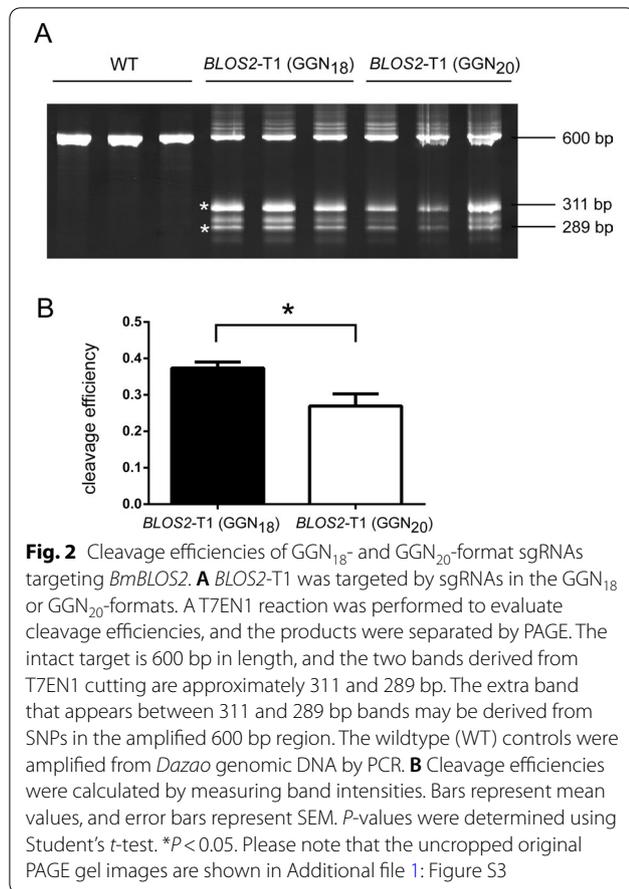


generated indels at higher efficiency ( $37.4\% \pm 1.6\%$ ) than the GGN<sub>20</sub> sgRNAs ( $26.9\% \pm 3.3\%$ ) at the sgRNA *BLOS2*-T1 target ( $P < 0.05$ ) (Fig. 2A, B). In contrast, no significant differences in indel frequencies were detected between GGN<sub>18</sub> ( $31.6\% \pm 6.8\%$ ) and GGN<sub>20</sub> ( $30.6\% \pm 6.5\%$ ) at *GR66*-T5 ( $P = 0.95$ ) (Fig. 3A, B). The result at *GR66*-T5 suggests that a single 5' G mismatch has a relatively minor effect on sgRNA cleavage efficiency compared with a 5' GG mismatch.

To further test the method, we selected four additional genomic targets in *BmGR66*, none of which have native (G/A) (G/A) residues at their 5' ends (*GR66*-T1, T2, T3, T4). Thus, at all four targets, sgRNAs in the GGN<sub>20</sub> format will mismatch the genomic sequences at both supernumerary G residues (Fig. 1C–F). After in vitro synthesis from a T7 promoter, the sgRNAs were mixed with Cas9 protein, and the complexes were injected into pre-blastoderm silkworm embryos. DNA was extracted from pools of 60 randomly selected injected embryos for each sgRNA and subjected to a T7EN1 assay. *GR66*-T1, T2, T3, and T4 generated

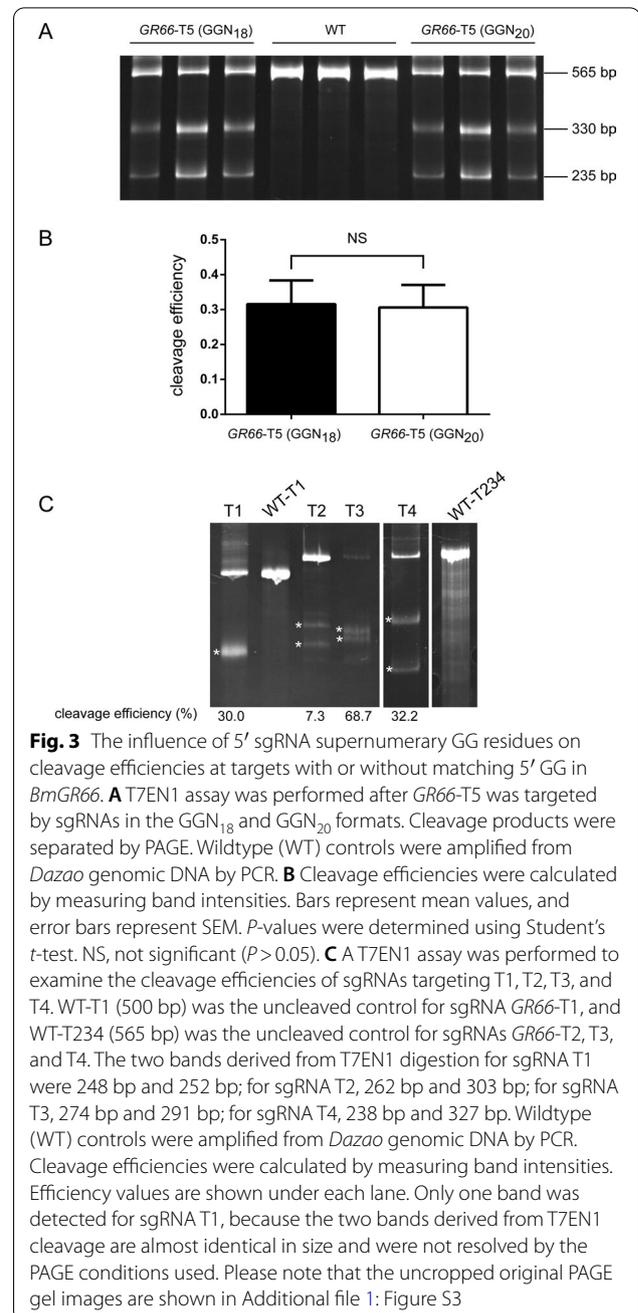
indels at efficiencies of 30.0%, 7.3%, 68.7%, and 32.2%, respectively (Fig. 3C). Importantly, the GGN<sub>20</sub> sgRNA at *GR66*-T3 produced indels much more efficiently than did the GGN<sub>18</sub> sgRNA at *GR66*-T5 (68.7% vs 31.6%). This result demonstrates that an sgRNA containing 5' GG mismatches may outperform an sgRNA with 5' GG matches within the same gene, suggesting that high knock-out efficiencies can sometimes be obtained without the requirement for 5' GG matching.

In summary, the introduction of 5' GG mismatches reduces the cutting efficiency of a GGN<sub>20</sub> sgRNA, although a single 5' G mismatch has a relatively minor effect. Nevertheless, the presence of one or two 5' mismatching G residues does not eliminate cleavage. Therefore, the addition of supernumerary 5' GG residues via the T7 promoter permits the synthesis of sgRNAs at high efficiency and also makes it possible to expand the target repertoire of the Cas9 RNP-based method. Indeed, in some cases it is possible to achieve superior gene knock-out efficiency at genomic targets that do not contain matching 5' GG residues.



### Generation of somatic and heritable mutations in *BmBLOS2* using a GGN<sub>20</sub> format sgRNA and Cas9

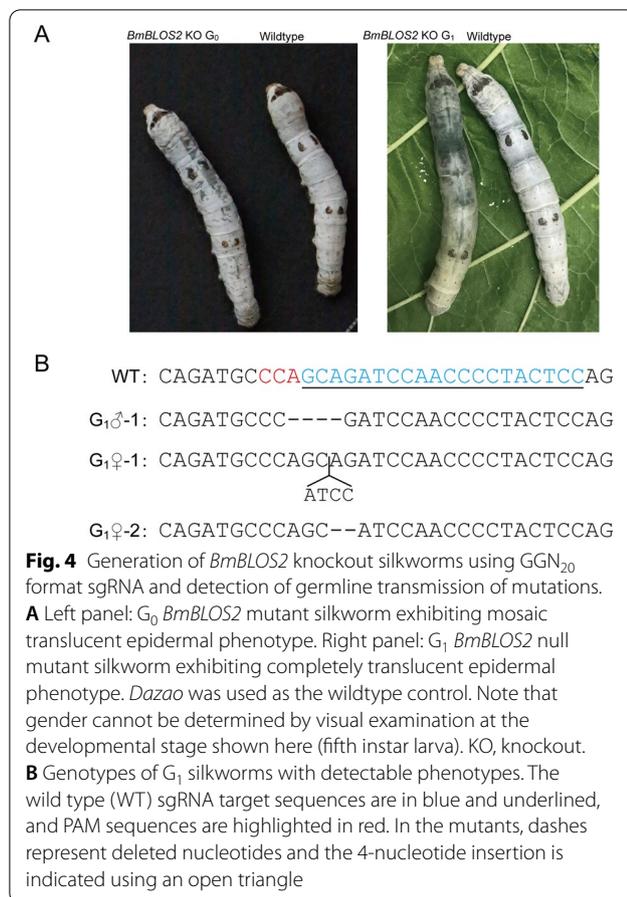
To test the germline transmission efficiency of mutations introduced by GGN<sub>20</sub> sgRNAs, we injected 560 pre-blastoderm embryos with a complex containing the Cas9 protein and an sgRNA targeting *BmBLOS2* (*BLOS2*-T1). We detected a mosaic translucent epidermal phenotype in 42 fifth instar larvae out of a total of 81 in the injected generation (*G*<sub>0</sub>), yielding an efficiency of 51.9% (42/81) (Fig. 4A). *G*<sub>0</sub> mosaic mutants with pronounced phenotypes were sib-mated. To assess heritability of the mutation, we randomly chose hatched larvae from three *G*<sub>1</sub> egg batches (around 100 individuals per batch), and detected at least one larva in each batch with a completely translucent epidermal phenotype, indicative of homozygosity (Fig. 4A). A total of five larvae with completely translucent epidermal phenotypes were found in the three batches. One phenotypic larva from each batch (one male and two females) was chosen randomly to determine the genotype. The analysis showed that all three harbored mutations in the sgRNA target regions, and no wild type alleles were detected (Fig. 4B).



These results demonstrate that the transmission efficiency of mutations generated by GGN<sub>20</sub> format sgRNAs is high enough to obtain phenotypic male (ZZ) homozygous mutants and female (ZW) hemizygous mutants in the *BmBLOS2* gene after only one breeding.

### Generation of heritable *BmGR66* mutants by co-injection of four sgRNAs with 5' (G/A) (G/A) target mismatches

To examine the germline transmission efficiency of mutations introduced by multiple sgRNAs with 5'GG



mismatches, we prepared GGN<sub>20</sub> sgRNAs targeting T1, T2, T3, T4 in *BmGR66*, complexed them with Cas9 protein, and co-injected them into 80 preblastoderm embryos. Six embryos hatched, and genomic DNA was extracted from their wings after eclosion. Regions surrounding the four sgRNA targeting sites were amplified by PCR and then subcloned. Three to five randomly selected subclones generated from each silkworm were sequenced. Nucleotide substitutions, indels (small insertions or deletions), large fragment deletions, and inversions were detected surrounding all four sgRNA targeting sites (Additional file 1: Figure S1 and S2). No wildtype sequences were detected from any silkworm, and at least three mutant alleles were identified (Additional file 1: Figure S1 and S2), demonstrating that the silkworms were mosaic in the injected generation (G<sub>0</sub>). Similar outcomes have been reported previously in silkworms [4], zebrafish [12], and mouse [13, 14].

To test whether the mutations can be transmitted through the germline, we randomly selected one G<sub>0</sub> male silkworm (#2) and one female G<sub>0</sub> silkworm (#4), both with somatic mutations (Additional file 1: Figure S1 and S2). They were mated to obtain offspring,

and newly moulted fifth-instar larvae in generation 1 (G<sub>1</sub>) were tested to determine their feeding preferences. 21 out of 79 ate cabbage leaves continuously, which are not normally consumed by wildtype silkworms, demonstrating an altered feeding preference (Additional file 2: Video S1). To correlate phenotype with genotype, 17 of the 21 phenotypic silkworms were randomly selected for analysis, and regions surrounding the sgRNA targets were amplified by PCR, subcloned, and sequenced. A total of 11 genotypes were recovered, and all harbored homozygous or compound heterozygous mutations. It is noteworthy that 7 of the 11 genotypes contained large (> 150 bp) insertions or deletions, which are readily obtained by simultaneous injection of multiple sgRNAs, and 10 of the 11 genotypes encoded truncated and presumably functionless proteins (Fig. 5 and Table 1). The changed feeding preferences resulting from the loss of *BmGR66* function are consistent with a previous report [7].

Taken together, these results demonstrate that co-injection of multiple sgRNAs with mismatched 5'GG residues can efficiently generate somatic mutations at targeted regions in injected individuals. Furthermore, mutant alleles can be transmitted through the germline, and null mutants can be readily obtained by sib-mating G<sub>0</sub> mosaics.

## Discussion

The delivery of CRISPR/Cas9 system into silkworms has been accomplished using the binary transgenic method [3, 5–7] and the Cas9 mRNA- or protein-based methods [4, 10]. In the binary method, the RNA polymerase type III promoter U6 is widely exploited to express sgRNA in vivo. The U6 promoter requires a guanosine nucleotide to initiate transcription, and thus genomic targets must contain a GN<sub>19</sub>NGG motif [15]. Strategies have been developed to circumvent this limitation. For example, a recent study demonstrated that replacing the U6 promoter with the H1 promoter expands the human genome target repertoire to include AN<sub>19</sub>NGG and GN<sub>19</sub>NGG [16]. In contrast, the silkworm U6 promoter effectively expresses sgRNAs initiated with any nucleotide [17], providing additional versatility. Unfortunately, even with this improvement, the transgene-based method is time consuming and difficult to implement [3, 5].

A significant advantage of the CRISPR/Cas9 system over ZFNs and TALENs is that mutagenesis can be directed to diverse genomic targets by simply exchanging the corresponding sgRNA, eliminating the labor that is required to reengineer the Cas9 enzyme [15, 18]. However, when the binary transgene-based method is used to deliver CRISPR/Cas9, every new genomic target requires a new transgenic line to express specific sgRNAs, limiting

genotype 1	<p>ACAAGACGTCCTACCACCAGTAAACATATATTTTTCTCATT<b>GTGGTAGTTGAGGATGAGGATAG</b>(196 bp deletion)AATGAGGATCCATGTAATAGTTTATTAATGTAACCGCGGATAAGT-----AAGGTGGCCATAATAATGCACTTCGAACGACACAAAGGCGTGCATGGGAAGCAAATATATAACTGGTTTGTATTAGAATCTGAAGTGAATTCGGCACAAATAAACACCAAAGCTATCAAACATTTATAATGCAAGTCTGAATTACAAGTTTACTGGAACCTTGTGCAG<b>C</b>-ATTGGTCTGCATCCGATGCAAATATCAA</p> <p>23 bp substitution, 196 bp deletion, 8 bp deletion, 1 bp substitution and 1 bp deletion</p>
genotype 2	<p>ACAATACAATAAGACATTCAGAACATGAGGATTA<b>AGGTGAGTGTGTG</b>(678 bp deletion)AGGTGGCCATAATAATGCACTTCGAACGACACAAAGGCGTGCATGGGAAGCAAATATATAACTGGTTTGTATTAGAATCTGAAGTGAATTCGGCACAAATAAACACCAAAGCTATCAAACATTTATAATGCAAGTCTGAATTACAAGTTTACTGGAACCTT<b>GTCTGGTTGGT</b>TATTGGTCTGCATCCG</p> <p>8 bp substitution, 678 bp deletion, 3 bp substitution and 4 bp insertion</p>
genotype 3	<p>GTACACTCGACTCCCTTTTA(352 bp sequences replaced by 174 bp unknown sequences)TCAAACCTCATGATAGATAA</p>
genotype 4	<p>AGTTTATTAATGTAACCGCGGATAAGTAAATTG---AGGTGGCCATAATAATGCACTTCGAACGACACAAAGGCGTGCATGGGAACGACAAATAACTGGTTTGTATTAGAATCTGAAGTGAATTCGGCACAAATAAACACCAAAGCTATCAAACATTTATAATGCAAGTCTGAATTACAAGTTTACTGGAACCTTGTGCAG<b>CT</b>ATTGGTCTGCATCCGATGCAAATATCAAACCTTACAGAGCTCT</p> <p>3 bp deletion and 1 bp insertion</p>
genotype 5	<p>CCCTTTTATCGCTTGTGGTAGTT---GAGGATCCATGTAATA(428 bp deletion)TAATAGGCAACCTATTAACGCCTAGAAG</p> <p>3 bp deletion and 428 bp deletion</p>
genotype 6	<p>GCCGCCACCCACCTTAATACATGAGTCTAAGG(230 bp deletion)TCTACGATAATATAATAAAGATCCATAAAATTTTTTCATTTGACAGTTCGATTCCTTATTATAATATTTAAATAAAATAACATGTTTCAGTACACTCGACTCCCTTTTATCGCTTGTGGTAG<b>C</b>--ATGAGGATCCATGTAATAGTTTATTAATGTAACCGCGGATAAGTAAATGAGAG<b>TGCCATAATATTG</b>AGGTGGCCAT</p> <p>230 bp deletion, 1 bp substitution, 2 bp deletion, 14 bp insertion</p>
genotype 7	<p>TCAGTACACTCGACTCCCTTTTATCGCTTGTGGTAGT(231 bp deletion)TCTGCATCCGATGCAAATATCAAACCTTACAGAGC</p>
genotype 8	<p>CGCTTGTGGTAGT<b>TATGAGG</b>TAAATGAGGATCCATGTAATAGTTTATTAATGTAACCGCGGATAAGTAAATGAG-AGGTGGCCATAATAATGCACTTCGAACGACACAAAGGCGTGCATGGGAAGCAAATATATAACTGGTTTGTATTAGAATCTGAAGTGAATTCGGCACAAATAAACACCAAAGCTATCAAACATTTATAATGCAAGTCTGAATTACAAGTTTACTGGAACCTTGTGCAG<b>C</b>-ATT</p> <p>7 bp insertion, 1 bp deletion, 1 bp substitution, 1 bp deletion</p>
genotype 9	<p>TACATAGATGGCGAGCTCACGGCCACC(435 bp sequences replaced by 2110 bp unknown sequences)GAGGATCCATGTAATAGTTTATTAATGTAACCGCGGATAAGTAAAT</p>
genotype 10	<p>TACACTCGACTCCCTTTT<b>TGTGGTAGT</b>TGTGGTAGT<b>GAT</b>GAGGATCCATGTAATAGTTTATTAATGTAACCGCGGATAAGTAAATGAGAA<b>AAAGTAATTA</b>AGGTGGCCATAATAATGCACTTCGAACGACACAAAGGCGTGCATGGGAAGCAAATATATAACTG</p> <p>5 bp substitution, 4 bp insertion, 11 bp insertion</p>
genotype 11	<p>TACACTCGACTCCCTTTTATCGCTTGTGGTAGT<b>GTGGTAGTTGAT</b>GAGGATCCATGTAATAGTTTATTAATGTAACCGCGGATAAGTAAATGAGAA<b>AAAGTAATTA</b>AGGTGGCCATAATAATGCACTTCGAACGACACAAAGGCGTGCATGGGAAGCAAATATAT</p> <p>1 bp substitution, 9 bp insertion, 11 bp insertion</p>

**Fig. 5** Genotypes recovered in G<sub>1</sub> *BmGR66* knockout silkworms obtained by sib-mating of G<sub>0</sub> mosaics. Mutations were detected in 17 G<sub>1</sub> silkworms by PCR-amplification of regions surrounding the four sgRNA targeting sites. Amplified products were then subcloned and sequenced. Nucleotides shown in red are base substitutions; boldface represents insertions; dashes represent deletions. For large deletions or insertions, the sizes of the deleted or replaced regions are shown in parentheses

the flexibility of the system. Therefore, the DNA-free RNP-based CRISPR/Cas9 system has become increasingly popular in silkworms [10]. In this method, sgRNAs are synthesized in vitro, typically from a T7 promoter because of its high efficiency. The efficiency advantage of the T7 promoter is partially offset by its requirement for two guanosine nucleotides to initiate transcription, which significantly limits the number of targetable sites in the silkworm and other organisms. Although a Csy4-based gRNA cleavage strategy can be used to expand targetable regions for sgRNAs transcribed from a T7

promoter in vitro, it requires an additional step of purification and increases the risk of sgRNA degradation [19].

In this study, we developed a simple method to expand the target repertoire for sgRNAs in the Cas9 RNP system. Briefly, conventional sgRNAs are designed to contain a targeting sequence of 20 nucleotides to obtain efficient mutagenesis. When the T7 promoter is used to synthesize an sgRNA, the first two residues in the product are 5'GG, yielding a targeting sequence with the structure GGN<sub>18</sub>. Since the two G residues must match

**Table 1** Genotypes of 17 examined G<sub>1</sub> *BmGR66* knockout silkworms

Silkworm ID	Genotype	
	Genotype 4	Homozygous
GR66 G <sub>1</sub> ♀-1		
GR66 G <sub>1</sub> ♀-2	Genotype 1 and 2	Compound heterozygous
GR66 G <sub>1</sub> ♀-3	Genotype 4 and 2	Compound heterozygous
GR66 G <sub>1</sub> ♀-4	Genotype 10 and 11	Compound heterozygous
GR66 G <sub>1</sub> ♀-5	Genotype 4 and 2	Compound heterozygous
GR66 G <sub>1</sub> ♀-7	Genotype 1 and 2	Compound heterozygous
GR66 G <sub>1</sub> ♀-9	Genotype 4	Homozygous
GR66 G <sub>1</sub> ♀-10	Genotype 3	Homozygous
GR66 G <sub>1</sub> ♀-12	Genotype 2	Homozygous
GR66 G <sub>1</sub> ♀-13	Genotype 9	Homozygous
GR66 G <sub>1</sub> ♂-1	Genotype 6	Homozygous
GR66 G <sub>1</sub> ♂-2	Genotype 5	Homozygous
GR66 G <sub>1</sub> ♂-4	Genotype 1 and 2	Compound heterozygous
GR66 G <sub>1</sub> ♂-5	Genotype 2	Homozygous
GR66 G <sub>1</sub> ♂-6	Genotype 7	Homozygous
GR66 G <sub>1</sub> ♂-8	Genotype 8	Homozygous
GR66 G <sub>1</sub> ♂-9	Genotype 1	Homozygous

The corresponding genotypes are shown in detail in Fig. 5

a corresponding 5' GG motif at the genomic target, the range of possible targets is constrained. Because the 5' GG residues perform unrelated roles in sgRNA synthesis and genomic targeting, we hypothesized that these functions could be satisfied independently with a slightly longer sgRNA structure. Specifically, we artificially introduced two supernumerary G residues to the 5' end of a conventional 20-nt sgRNA, and then asked if the resulting GGN<sub>20</sub> sgRNA generates mutations even if the supernumerary nucleotides have no corresponding matches in the genome. The remaining 20 nucleotides (N<sub>20</sub>) in the GGN<sub>20</sub> sgRNA match the target perfectly. This strategy eliminates the need to select genomic targets containing 5' GG residues, while preserving the ability to leverage the T7 promoter for efficient sgRNA synthesis.

GGN<sub>20</sub> sgRNAs were tested against genomic targets that do not contain a native 5' (G/A) (G/A) motif. We first examined the effect of the unpaired 5' supernumerary guanines on mutagenesis efficiency in vivo in pooled injected embryos and individuals. As expected, the mutagenesis efficiencies of control sgRNAs that perfectly match their genomic targets (GGN<sub>18</sub> format) equaled or surpassed those of sgRNAs containing two unpaired 5' supernumerary guanines (GGN<sub>20</sub> format), possibly because the GG mismatches at the 5' end of sgRNA impair cleavage. These results are consistent with previous reports in zebrafish [20]. We also found that a single G mismatch at the 5' end of the target has a relatively minor effect on cleavage efficiency. Nevertheless,

the supernumerary guanines at the 5' end do not abolish cleavage, which remains at levels high enough to generate somatic mutations and germline-transmissible mutations that can be recovered in succeeding generations. Homozygous or compound heterozygous mutants are readily obtained after only one breeding of mosaic G<sub>0</sub> silkworms. Furthermore, we found that one GGN<sub>20</sub> format sgRNA (GR66-T3) is much more efficient than the GGN<sub>18</sub> format sgRNA GR66-T5 (68.7% vs 31.6%), demonstrating that sgRNAs with 5' matched GG are not necessarily more efficient than sgRNAs with 5' GG mismatches for targets within the same gene. Taken together, the results show that our improved method provides researchers with an expanded target repertoire, and can reach some targets at high efficiency even if matching 5' GG residues are unavailable. Our method further increases the advantages of the Cas9 ribonucleoprotein (RNP)-based system over the widely used binary transgenic system [3, 5].

We also co-injected four different GGN<sub>20</sub> sgRNAs into larvae and detected multiple mutant alleles in the G<sub>0</sub> recipients. Homozygous and compound heterozygous G<sub>1</sub> silkworms that harbored a wide range of mutations were easily generated. Almost all of the mutations (10 of 11) that were examined presumably encode truncated protein products, and 7 of the 11 contained large (> 150 bp) deletions or insertions. We hypothesize that the simultaneous activity of multiple sgRNAs not only increases the diversity of mutant alleles, but also favors the formation of large fragment insertions or deletions, in contrast to the smaller indels that are often generated when a single sgRNA is used. A previous study in zebrafish demonstrated that DNA repair machinery corrects double-stranded-breaks induced by Cas9 in a stereotypical and target-specific fashion, resulting in reduced mutant allele diversity [9]. Some somatic alleles were also over-represented in the germline, and at certain genomic targets the predominant alleles harboring indels were not frameshifts, making it difficult to obtain offspring with heritable null mutations [9]. Therefore, when the generation of null mutants is desired, injection of several different sgRNAs is much more effective than injection of a single sgRNA species. In addition, cleavage by multiple sgRNAs enhances indel formation efficiency and facilitates the production of complete gene knockouts in the injected generation [21, 22].

## Conclusions

In conclusion, we present a simple strategy to expand the targeting range of Cas9 RNP-based mutagenesis. The method leverages the highly efficient T7 promoter to produce sgRNAs with supernumerary 5' GG

residues that do not match corresponding nucleotides at the genomic target. Although the mismatches diminish the efficiency of mutagenesis, the efficiency is sufficiently high to generate somatic and heritable mutations. Homozygous and compound heterozygous mutants were readily obtained in a single generation by sib-breeding mosaics. The new method significantly expands the number of sites that can be targeted by the CRISPR/Cas9 system via the injection of in vitro transcribed sgRNAs in silkworm.

## Methods

### sgRNA design and synthesis

sgRNAs were designed using the CHOP-CHOP online utility (<http://chopchop.cbu.uib.no/>). sgRNA targeting sites are shown in Fig. 1. As described in a recent publication [23], the DNA template for T7 promoter used to drive in vitro transcription was constructed by PCR. Briefly, a customized oligonucleotide containing the T7 promoter and the sgRNA target sequence (N<sub>20</sub> or N<sub>18</sub>) was designed as a forward primer with the sequence 5'-TAATACGACTCACTATAGG(N<sub>20</sub> or N<sub>18</sub>)GTTTTAGAGCTAGAAATAGC. The T7 promoter sequences are underlined. The reverse primer was 5'-AAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAAC-3'. sgRNA synthesis was performed using a RiboMax large scale RNA production system—T7 kit (Promega, cat. P1300), following the manufacturer's instructions.

### Silkworm embryo microinjection

The bivoltine silkworm strain *Dazao* was obtained from the Silkworm Gene Bank of Southwest University (Chongqing, China) and was used in all experiments in this study. To collect non-diapaused eggs, silkworm eggs were incubated at 15°C until hatching, and the larvae were reared at 25°C and fed with fresh mulberry leaves until the wandering stage. Adult moths then oviposited non-diapaused eggs, which were used for microinjection. A mixture of sgRNA and Cas9 protein (Thermo fisher, cat. A36496) was incubated at room temperature for 15 min and microinjected into preblastoderm embryos within 5 h after oviposition. All silkworm embryo microinjection experiments in the study were conducted using this protocol.

Injected embryos were incubated at 25°C and 80% humidity for 48 h. 60 silkworm eggs were then collected to extract genomic DNA for T7EN1 assays.

For experiments requiring hatched larvae, injected embryos were incubated at 25°C and 80% humidity for

approximately 11 days until hatching. Larvae were maintained at 25°C and fed fresh mulberry leaves.

### T7EN1 assay and calculation of cleavage efficiency

Genomic DNA was extracted from pools containing 60 silkworm eggs, and PCR was performed to amplify sequences surrounding the sgRNA targeting site. To amplify the region surrounding the *GR66* T1 target, the primer was F: 5'-TCCCTTTTATCGCTTGTGGT-3', and R: 5'-CTTCTAGGCGTTAATAGGTTGC-3'. To amplify the region surrounding the *GR66* T2, T3, T4 and T5 targets, the primer was F: 5'-TGATTTCGGACTCACAAGACG-3', and R: 5'-GGAAGAGAATGCGCCTGTAT-3'. To amplify the region surrounding the *BLOS2* T1 target, the primer was F: 5'-TGAGATGCTTTATGAGACAAGTCC-3', and R: 5'-ATTTTCGAACCCGACAATGA-3'.

To determine sgRNA cleavage efficiencies, after the PCR products were annealed, they were subjected to T7EN1 (NEB, cat. M0302S) digestion at 37 °C for 1 h, and fragments were then separated by polyacrylamide gel electrophoresis (PAGE). The detailed protocol is available in our previous publication [24]. Band intensities were measured after PAGE using Image J [25]. Using a previously described approach [26], cleavage efficiencies were calculated by the formula  $100 \times (1 - (1 - (b + c)/(a + b + c))^{1/2})$ , where *a* is the integrated intensity of the undigested PCR product, and *b* and *c* are the integrated intensities of each cleavage product.

### Genotype analysis

Genomic DNA was extracted from moth wings, and regions surrounding the sgRNA targeting sites were amplified by PCR. The PCR primers for genotyping *BmBLOS2* were F: 5'-TGAGATGCTTTATGAGACAAGTCC-3', and R: 5'-ATTTTCGAACCCGACAATGA-3'. The PCR primers for genotyping *BmGR66* were F: 5'-CCCATCCTTCAAACCTGAAA-3', and R: 5'-ACATTTGTTCAACCCCAAGC-3'. PCR products were subcloned into pEASY-blunt-zero vectors (TransGen, cat. CB501-01) and sequenced.

### Statistical analysis

All values are presented as means ± SEM. Student's *t*-test was used to compare means. Differences were defined to be statistically significant at *P* < 0.05.

### Abbreviations

CRISPR/Cas9: Clustered regularly interspersed short palindromic repeat/CRISPR-associated protein 9; gRNA: Guide RNA; PAM: Protospacer adjacent motif; WT: Wildtype; KO: Knockout; RNP: Ribonucleoprotein; ZFNs: Zinc-finger nucleases; TALENs: Transcription activator-like effector nucleases.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12896-021-00714-6>.

**Additional file 1: Figure S1.** Genotypes of three G<sub>0</sub> *BmGR66* knock-out silkworms generated by co-injection of sgRNAs T1, T2, T3, and T4 complexed with Cas9 protein. **Figure S2.** Genotypes of two G<sub>0</sub> *BmGR66* knockout silkworms generated by co-injection of sgRNA T1, T2, T3 and T4 with Cas9 protein. **Figure S3.** Uncropped original PAGE gel image.

**Additional file 2. Video S1.** A compound heterozygous mutant of *BmGR66* eating cabbage leaves.

### Acknowledgements

We wish to thank our group members Chunyan Fang and Tingting Gai for their technical assistance.

### Authors' contributions

XT, YZ and FD planned and designed the experiments. YZ, AY, SL, WW, LX performed the experiments and analyzed the data. YZ and XT wrote the paper. All authors read and approved the final manuscript.

### Funding

This work was supported by the National Natural Science Foundation of China [Grant No.: U20A2058] and Fundamental Research Funds for the Central Universities [Grant No.: SWU118082]. The funding body did not play any roles in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

### Availability of data and materials

All data generated or analysed during this study are included in this published article and its supplementary information files.

### Declarations

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

Not applicable.

#### Competing interests

We declare no competing financial interests.

Received: 24 February 2021 Accepted: 8 September 2021

Published online: 20 September 2021

### References

1. Takasu Y, Kobayashi I, Beumer K, Uchino K, Sezutsu H, Sajwan S, Carroll D, Tamura T, Zurovec M. Targeted mutagenesis in the silkworm *Bombyx mori* using zinc finger nuclease mRNA injection. *Insect Biochem Mol Biol*. 2010;40(10):759–65.
2. Wang Y, Tan A, Xu J, Li Z, Zeng B, Ling L, You L, Chen Y, James AA, Huang Y. Site-specific, TALENs-mediated transformation of *Bombyx mori*. *Insect Biochem Mol Biol*. 2014;55:26–30.
3. Gai T, Tong X, Han M, Li C, Fang C, Zou Y, Hu H, Xiang H, Xiang Z, Lu C, et al. Cocoonase is indispensable for Lepidoptera insects breaking the sealed cocoon. *PLoS Genet*. 2020;16(9):e1009004.
4. Wang Y, Li Z, Xu J, Zeng B, Ling L, You L, Chen Y, Huang Y, Tan A. The CRISPR/Cas system mediates efficient genome engineering in *Bombyx mori*. *Cell Res*. 2013;23(12):1414–6.
5. Liu Q, Liu W, Zeng B, Wang G, Hao D, Huang Y. Deletion of the *Bombyx mori* odorant receptor co-receptor (*BmOrco*) impairs olfactory sensitivity in silkworms. *Insect Biochem Mol Biol*. 2017;86:58–67.
6. Xu J, Chen RM, Chen SQ, Chen K, Tang LM, Yang DH, Yang X, Zhang Y, Song HS, Huang YP. Identification of a germline-expression promoter for genome editing in *Bombyx mori*. *Insect Sci*. 2019;26(6):991–9.
7. Zhang ZJ, Zhang SS, Niu BL, Ji DF, Liu XJ, Li MW, Bai H, Palli SR, Wang CZ, Tan AJ. A determining factor for insect feeding preference in the silkworm, *Bombyx mori*. *PLoS Biol*. 2019;17(2):e3000162.
8. Wilson C, Bellen HJ, Gehring WJ. Position effects on eukaryotic gene expression. *Annu Rev Cell Biol*. 1990;6:679–714.
9. Gagnon JA, Valen E, Thyme SB, Huang P, Akhmetova L, Pauli A, Montague TG, Zimmerman S, Richter C, Schier AF. Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs. *PLoS ONE*. 2014;9(5):e98186.
10. Jia L, Ma SY, Zhang T, Xing WQ, Liu Y, Li YF, Chen XX, Xia QY. Enhanced *Bombyx* genome editing via Cas9 ribonucleoprotein injection. *Insect Sci*. 2019;26(6):1059–62.
11. Hwang WY, Fu Y, Reyon D, Maeder ML, Tsai SQ, Sander JD, Peterson RT, Yeh JR, Joung JK. Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat Biotechnol*. 2013;31(3):227–9.
12. Jao LE, Wentz SR, Chen W. Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. *Proc Natl Acad Sci USA*. 2013;110(34):13904–9.
13. Shen B, Zhang J, Wu H, Wang J, Ma K, Li Z, Zhang X, Zhang P, Huang X. Generation of gene-modified mice via Cas9/RNA-mediated gene targeting. *Cell Res*. 2013;23(5):720–3.
14. Yen ST, Zhang M, Deng JM, Usman SJ, Smith CN, Parker-Thornburg J, Swinton PG, Martin JF, Behringer RR. Somatic mosaicism and allele complexity induced by CRISPR/Cas9 RNA injections in mouse zygotes. *Dev Biol*. 2014;393(1):3–9.
15. Mali P, Yang L, Svelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM. RNA-guided human genome engineering via Cas9. *Science*. 2013;339(6121):823–6.
16. Ranganathan V, Wahlin K, Maruotti J, Zack DJ. Expansion of the CRISPR-Cas9 genome targeting space through the use of H1 promoter-expressed guide RNAs. *Nat Commun*. 2014;5:4516.
17. Zeng B, Zhan S, Wang Y, Huang Y, Xu J, Liu Q, Li Z, Huang Y, Tan A. Expansion of CRISPR targeting sites in *Bombyx mori*. *Insect Biochem Mol Biol*. 2016;72:31–40.
18. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science*. 2013;339(6121):819–23.
19. Qin W, Liang F, Feng Y, Bai H, Yan R, Li S, Lin S. Expansion of CRISPR/Cas9 genome targeting sites in zebrafish by Csy4-based RNA processing. *Cell Res*. 2015;25(9):1074–7.
20. Hoshijima K, Jurynek MJ, Klatt Shaw D, Jacobi AM, Behlke MA, Grunwald DJ. Highly efficient CRISPR-Cas9-based methods for generating deletion mutations and F0 embryos that lack gene function in zebrafish. *Dev Cell*. 2019;51(5):645–657.e644.
21. Zuo E, Cai YJ, Li K, Wei Y, Wang BA, Sun Y, Liu Z, Liu J, Hu X, Wei W, et al. One-step generation of complete gene knockout mice and monkeys by CRISPR/Cas9-mediated gene editing with multiple sgRNAs. *Cell Res*. 2017;27(7):933–45.
22. Wu RS, Lam II, Clay H, Duong DN, Deo RC, Coughlin SR. A rapid method for directed gene knockout for screening in G0 zebrafish. *Dev Cell*. 2018;46(1):112–125.e114.
23. Varshney GK, Pei W, LaFave MC, Idol J, Xu L, Gallardo V, Carrington B, Bishop K, Jones M, Li M, et al. High-throughput gene targeting and phenotyping in zebrafish using CRISPR/Cas9. *Genome Res*. 2015;25(7):1030–42.
24. Zou YL, Li ZY, Zou YJ, Hao HY, Hu JX, Li N, Li QY. Generation of pigs with a Belgian Blue mutation in *MSTN* using CRISPR/Cpf1-assisted ssODN-mediated homologous recombination. *J Integr Agric*. 2019;18(6):1329–36.
25. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods*. 2012;9(7):671–5.
26. Han H, Ma Y, Wang T, Lian L, Tian X, Hu R, Deng S, Li K, Wang F, Li N, et al. One-step generation of myostatin gene knockout sheep via the CRISPR/Cas9 system. *Front Agric Sci Eng*. 2014;1(1):2–5.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.