Efficient genome editing and gene knockout in *Setaria viridis* with CRISPR/Cas9 directed gene editing by the non-homologous end-joining pathway

Marcos Fernando Basso¹,², Karoline Estefani Duarte¹,³, Thais Ribeiro Santiago¹,⁴, Wagner Rodrigo de Souza³, Bruno de Oliveira Garcia¹, Barbara Dias Brito da Cunha¹, Adilson Kenji Kobayashi¹, Hugo Bruno Correa Molinari¹,*

¹National Center for Agroenergy Research (CNPAE), Brazilian Agricultural Research Corporation (EMBRAPA), Brasília, DF, 70770-901, Brazil; ²BIOMOL/BIOTEC Laboratory, Mato Grosso Cotton Institute (IMAmt), Rondonópolis, MT, 78740-970, Brazil; ³Center for Natural and Human Sciences, Federal University of ABC (UFABC), São Bernardo do Campo, São Paulo, 09606-045, Brazil; ⁴Department of Phytopathology, Federal University of Brasília, Brasilia (UNB), Distrito Federal, 70910-900, Brazil

*E-mail: hugo.molinari@embrapa.br  Tel: +55-61-3448-2307  Fax: +55-61-34481598

Received December 1, 2020; accepted April 7, 2021 (Edited by Y. Nagano)

**Abstract**

The CRISPR/Cas9 system has been used for genome editing in several organisms, including higher plants. This system induces site-specific mutations in the genome based on the nucleotide sequence of engineered guide RNAs. The complex genomes of C4 grasses makes genome editing a challenge in key grass crops like *Zea mays*, sorghum (*Sorghum bicolor*), *Brachiaria* spp., switchgrass (*Panicum virgatum*), and sugarcane (*Saccharum* spp.). *Setaria viridis* is a diploid C4 grass widely used as a model for these C4 crop plants. Here, an optimized CRISPR/Cas9 binary vector that exploits the non-homologous end joining (NHEJ) system was used to knockout a *green fluorescent protein* (*gfp*) transgene in *S. viridis* accession A10.1. Transformation of embryogenic callus by *A. tumefaciens* generated ten glufosinate-ammonium resistant transgenic events. In the T0 generation, 60% of the events were biallelic mutants in the *gfp* transgene with no detectable accumulation of GFP protein and without insertions or deletions in predicted off-target sites. The *gfp* mutations generated by CRISPR/Cas9 were stable and displayed Mendelian segregation in the T1 generation. Altogether, the system described here is a highly efficient genome editing system for *S. viridis*, an important model plant for functional genomics studies in C4 grasses. Also, this system is a potential tool for improvement of agronomic traits in C4 crop plants with complex genomes.

**Key words:** CRISPR/Cas9, genome editing, non-homologous end-joining, panicoid grasses, *Setaria viridis*, stable genetic transformation.

**Introduction**

Genome editing is a powerful tool to introduce deletions, insertions, or sequence modifications into organismal genomes. Genome editing with the Clustered Regularly Interspaced Short Palindromic Repeat-associated endonuclease protein 9 (CRISPR/Cas9) system has been optimized for a range of organisms, replacing other site-specific nucleases (e.g., Meganucleases, ZNFs, and TALENs) due to its greater simplicity, specificity, and efficiency (Cong et al. 2013; Jinek et al. 2012; Makarova et al. 2015; Mali et al. 2013; May et al. 2013; Zhang 2013). The CRISPR/Cas9 system creates double-stranded DNA breaks (DSBs) at a specific target sequence specified by the sequence of a guide RNA (gRNA). The break is repaired by either the non-homologous end joining (NHEJ) pathway, which generates nucleotide insertions or deletions (indels) at the target site, or the homology-directed repair (HDR) pathway, where precise corrections can be made by using a homologous DNA donor.

Abbreviations: ATG, Translation initiation site; cDNA, Complementary DNA; CIM₂, Callus induction medium; CRISPR/Cas9, Clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9 associated protein 9; DNA, Deoxyribonucleic acid; DSD, Degenerate Sequence Decoding; GFP or *gfp*, Green fluorescent protein; gRNA, Guide RNA; HDR, homology-directed repair; NHEJ, Non-homologous end joining; NLS, nuclear localization signal; NOS, Nopaline synthase; OsAct₁, *Oryza sativa actin 1*; OsU6, *Oryza sativa*; PAM, protospacer adjacent motif; RNA, Ribonucleic acid; qRT-PCR, Quantitative reverse transcription PCR; S, *viridis*, *Setaria viridis*; SpCas9WT, Cas9 nuclease from *Streptococcus pyogenes*; SV40, Simian vacuolating virus 40; T-DNA, transfer DNA; ZmUBI-1, *Zea mays*.

This article can be found at http://www.jspcmb.jp/

Published online June 12, 2021
directed repair (HDR) pathway, which can generate single and double knock-ins of a desired target gene or promoter sequence (Li et al. 2016; Sun et al. 2016; Zhao et al. 2016). The NHEJ pathway is more efficient than the HDR pathway (Li et al. 2016; Liang et al. 2017; Sun et al. 2016; Tang et al. 2017; Woo et al. 2015; Zhang et al. 2016). However, the efficiency of either CRISPR/Cas9 system varies between phyllogenetically distant plant species. Thus, it is necessary to optimize the elements of the CRISPR/Cas9 system for the target genome (Basso et al. 2020; Li et al. 2016; Yin et al. 2017).

Several features of the C4 grass *Setaria viridis* make it an experimental model plant for C4 grass crops, including a short life cycle, high seed production, and established protocols for genetic transformation and plant regeneration (Brutnell et al. 2010; Li and Brutnell 2011; Martins et al. 2015), as well as a small diploid genome that is fully sequenced (Bennetzen et al. 2012). This model plant is widely used to study C4 grass functional genomics and for proof-of-concept experiments for panicle bioenergy feedstocks and food crops with highly complex genomes like maize (*Zea mays*), sorghum (*Sorghum bicolor*), Brachiaria spp., switchgrass (*Panicum virgatum*), and sugarcane (*Saccharum spp.*).

Huang et al. (2019) edited the genome of *S. viridis* with a CRISPR/Cas9 genome system to introduce mutations in the *SvLes1* (Sevir.5G085400) gene. More recently, Weiss et al. (2020) demonstrated efficient genome editing in *S. viridis* with a combination of a CRISPR/Cas9 and expression of exonuclease T.Scroll2 (CRISPR/Cas9_Trex2), which promoted repair by either NHEJ or microhomology-mediated end joining (MMEJ). These authors showed that this multiplex CRISPR/Cas9_Trex2 system induced targeted indels in both the *svDrmla* and *svDrmlb* genes of T0 transgenic events at a frequency of 73% to 100%. These indels were transmitted to at least 60% of the transgene-free T1 plants, with 33% of them containing biallelic or homozygous indels in both genes.

Here, we optimized a CRISPR/Cas9 binary vector that promotes the activity of the NHEJ pathway and this vector was used to edit a transgenic copy of the *green fluorescent protein* (*gfp*) gene in *S. viridis* accession A10.1. This study generated ten biallelic mutants (homozygous) at the T0 generation. These indels in the *gfp* transgene were stable and showed Mendelian segregation in the T1 generation. These lines carried no off-target indels in predicted off-target sites, indicating highly site-specific targeting by this combination of CRISPR/Cas9 binary vector and gRNA.

**Materials and methods**

**Construction of a CRISPR/Cas9 binary vector to enhance genome editing through NHEJ pathway**

The T-DNA of the binary vector is composed of (i) selection marker gene (*bar* gene from *Streptomyces hygroscopicus* for resistance to glufosinate-ammonium herbicide), with an intron sequence, controlled by *Oryza sativa* constitutive promoter *actin 1* (*OsAct-1*), and 3′*rbcsE9* transcription terminator; (ii) Cas9 nuclease from *Streptococcus pyogenes* (*SpCas9WT*), with codon-usage optimized for monocots and fused in-frame with two SV40 nuclear localization signal (NLS) at the 5′ and 3′ end of the gene. This fragment was engineered under control of the constitutive ubiquitin-1 promoter from *Zea mays* (*ZmUBI-1*), with an intron sequence and linked with the nopaline synthase (NOS) transcription terminator sequence. The Kozak sequence CCG AA (optimized Kozak sequence for monocots) was introduced in front of the translation initiation site (ATG); (iii) DNA sequence that corresponds to the gRNA sequence under control of the constitutive U6-2 small nuclear RNA promoter (coding for RNA polymerase III) from *Oryza sativa* (*OsU6*), with extra guanine (G) nucleotide at 3′ end fused in-frame at the 3′ end of a synthetic RNA sequence (scaffold RNA used for Cas9 coupling). The dyad symmetry sequence was added after the transcription stop signal (poly A7). The gRNA was designed for targeting the 5′ end of the *gfp* (*green fluorescent protein*) gene and direct the cleavage by Cas9 from the sense strand of target DNA using NGG sequence as the Protospacer Adjacent Motif (PAM). Thus, the gfp transgene was used as target DNA to optimize the CRISPR/Cas9 NHEJ system in *S. viridis* and the GFP protein as a reporter of the knockout efficiency. The absence of off-target mutations was analyzed by scanning the Phytozome 12 *S. viridis* genome (*Setaria viridis* v2.1, DOE-JGI, (http://phytozome.jgi.doe.gov/) using GT-Scan (O’Brien and Bailey 2014) and CCTop (Stemmer et al. 2015) software. The binary vector also contained the CoIE1 and pVS1 origin of replication from *Escherichia coli* and *Agrobacterium tumefaciens*, respectively, and resistance marker gene to streptomycin/spectinomycin (*Smr/Sp*). The CRISPR/Cas9 binary vector was synthesized, and assembled by DNA Cloning Service (www.dna-cloning.com; Hamburg, Germany).

**Stable transformation and regeneration of *Setaria viridis* gene editing events from embryogenic callus**

*S. viridis* (accession A10.1) was co-transformed by a high-efficiency *Agrobacterium*-mediated transformation protocol (Martins et al. 2015). After removing the lemmas and paleas, mature seeds were disinfested, blotted onto sterile filter paper, and transferred to callus induction medium (CIM2) for embryogenic callus induction. After 22 days of incubation in the dark at 25 ± 2°C, embryogenic calli were isolated and transferred onto a fresh CIM2 medium. After a week, selected embryogenic calli were transformed by *A. tumefaciens* EHA105. Embryogenic calli from a homozygous
transgenic *S. viridis* plants constitutively expressing GFP and resistant to hygromycin B (30 mg/l) (Martins et al. 2015) were co-transformed with the CRISPR/Cas9 NHEJ binary vector. Selection with glufosinate-ammonium herbicide (3 mg/l; Liberty<sup>®</sup>-Bayer CropScience) identified ten independent CRISPR/Cas9 NHEJ events after. Transgenic seedlings were planted into pots containing 250 g of latosol, substrate (Plantmax<sup>®</sup>), and vermiculite (Agrifloc, Brasil Minérios) mixture (3:1:0.5; w/w/w). The plants were maintained in a growth chamber under 16/8-h photoperiod at 500 µmol m<sup>−2</sup> s<sup>−1</sup> light intensity, 26±2°C temperature, and 65% relative humidity until seed establishment (seeds from T1 generation).

**Genotyping of the gene editing events**

Genomic DNA was isolated from co-transforming T0 lines and control (GFP positive, but not co-transformed with our CRISPR/Cas9 binary vector) plants using the CTAB method (Doyle and Doyle 1987). PCR assays were carried out using specific primers targeting *bar* gene and ZmUBI-1:GFP to confirm the insertion of ZmUBI-1:GFP and CRISPR/Cas9 transgene in the regenerated plants from selection with glufosinate-ammonium herbicide. PCR-positive events were also confirmed for BAR protein accumulation with a quick test strip (QuickStix™ Kit for PAT/bar, EnviroLogix, Inc., USA), according to the manufacturer’s instructions. The ZmUBI-1:GFP PCR products (1.23 kb length) from ten co-transforming events and two control plants were directly purified and dual sequenced (forward and reverse) using the Sanger method by Macrogen Service (Geumcheon-gu, Seoul, South Korea). Chromatograms were analyzed using Geneious R10 software (Kearse et al. 2012). Mono (heterozygous) and biallelic (homozygous) *indels* in the DNA sequence targeted by the gRNA were decoded using standard degenerate symbols (IUPAC/IUB), based on the color peaks of the chromatograms, according to the Degenerate Sequence Decoding (DSD) method (Ma et al. 2015). Chromatograms from both forward and reverse sequencing confirmed the *indels* and their exact positions in the target sequence. The inheritance and stability of *indels* induced by the CRISPR/Cas9 system were confirmed by PCR, and subsequent Sanger sequencing of the ZmUBI-1:GFP transgene from EC1 to EC3, EC5, and EC7 events from T1 generation. Wild-type plants (ECWT) were used as control. The T-DNA copy number in each event was estimated using seeds from T1 co-transforming events and two control plants were directly purified and dual sequenced (forward and reverse) using the Sanger method (Kearse et al. 2012). The relative expression was calculated using the 2<sup>−∆∆Ct</sup> method (Schmittgen and Livak 2008) using *Sv* (endogenous reference gene). All primers used in the present work are listed in Supplementary Table S2.

**GFP protein detection in gene editing events**

GFP-specific fluorescence in leaves of the co-transforming events was visualized in Leica M205 FA stereomicroscope equipped with either a long pass filter with a 395–455 nm excitation filter and a 480 nm emission filter or a GFP filter. Detection of GFP protein in the crude protein extracts isolated from leaf tissue of adult co-transforming and transgenic plants was carried out using indirect ELISA (Clark 1981). Initially, the protein amount in the total protein extracts was quantified using Bradford’s method (Bradford 1976). Serological assays were conducted using the monoclonal IgG1 Anti-GFP (G1546, Sigma-Aldrich) as a primary antibody produced in mice, which were immunized with a synthetic peptide corresponding to amino acids 132–144 of GFP. Polyclonal anti-mouse IgG produced in goat and alkaline phosphatase-conjugated (A3688,
Sigma-Aldrich) was used as a secondary antibody. Samples consisting of leaves from S. viridis A10.1 non-transgenic plants (wild-type plant used as a negative control for GFP in the indirect ELISA), transgenic ECWT1 and ECWT2 plants (GFP-positive control for the indirect ELISA), and co-transforming EC1 to EC10 events were grounded to a fine powder with pestle and mortar in liquid nitrogen and homogenized in 1:2 (w/v) extraction buffer (10 mM Na$_2$CO$_3$, 10 mM NaHCO$_3$, 15 mM NaN$_3$, 5 mM Na-DIECA, 0.2% bovine serum albumin, 2% PVP40, pH 9.6). Plant samples were considered GFP-positive when absorbance at 405 nm was at least twice the average value of the negative control.

**Results**

We constructed a new CRISPR/Cas9 binary vector engineered with several genetic elements to enhance the efficiency and specificity of genome editing in S. viridis, as described in Material and Methods and shown in Figure 1a. We also designed a gRNA to target the 5’-

![Diagram](image)

Figure 1. Genome editing and gene knock-out strategy using the CRISPR/Cas9 system exploring the non-homologous end joining (NHEJ) pathway in S. viridis accession A10.1. (a) Optimized CRISPR/Cas9 NHEJ binary vector for genome editing in S. viridis; (b) schematic diagram of the guide RNA targeting the gfp gene. (c) Schematic T-DNA region of the binary vector used to transform Setaria viridis for constitutive expression of GFP protein. The generated Setaria viridis events were edited in co-transforming events by CRISPR/Cas9 system.
<table>
<thead>
<tr>
<th>Events</th>
<th>Decoding of mono and biallelic mutations</th>
<th>Decoding results</th>
<th>gfp mRNA accumulation</th>
<th>GFP protein accumulation</th>
<th>Segregation ratio in T1</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECWT1</td>
<td>GGATGGTGATGGAAGGCCCAAC</td>
<td>Non-edited control lineage 1</td>
<td>Positive</td>
<td>Positive</td>
<td>—</td>
</tr>
<tr>
<td>ECWT2</td>
<td>GGATGGTGATGGAAGGCCCAAC</td>
<td>Non-edited control lineage 2</td>
<td>Positive</td>
<td>Positive</td>
<td>—</td>
</tr>
<tr>
<td>EC1</td>
<td>Allele 1: deletion of G nucleotide; Allele 2: no indels; Monoallelic mutant.</td>
<td>Positive</td>
<td>Positive</td>
<td>3:1</td>
<td></td>
</tr>
<tr>
<td>EC2</td>
<td>Allele 1: insertion of A nucleotide; Allele 2: deletion of GA nucleotides; Biallelic mutant.</td>
<td>Positive</td>
<td>Negative</td>
<td>3:1</td>
<td></td>
</tr>
<tr>
<td>EC3</td>
<td>Allele 1: deletions of A and T nucleotides and insertions of C/A, T and G nucleotides; Allele 2: insertions of AA, A, G, and G nucleotides and two deletions of CC nucleotides; Biallelic mutant.</td>
<td>Positive</td>
<td>Negative</td>
<td>3:1</td>
<td></td>
</tr>
<tr>
<td>EC4</td>
<td>Allele 1: deletion of G nucleotide; Allele 2: insertion of A nucleotide; Biallelic mutant.</td>
<td>Positive</td>
<td>Negative</td>
<td>3:1</td>
<td></td>
</tr>
<tr>
<td>EC5</td>
<td>Allele 1: deletion of G nucleotide; Allele 2: deletions of GT nucleotides; Biallelic mutant.</td>
<td>Positive</td>
<td>Negative</td>
<td>3:1</td>
<td></td>
</tr>
<tr>
<td>Events</td>
<td>Decoding of mono and biallelic mutations</td>
<td>Decoding results</td>
<td>GFP mRNA accumulation</td>
<td>GFP protein accumulation</td>
<td>Segregation ratio in T1</td>
</tr>
<tr>
<td>--------</td>
<td>------------------------------------------</td>
<td>------------------</td>
<td>----------------------</td>
<td>-------------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>EC6</td>
<td>EC6 line GGAACCTGAGTGTGAGATGTGAACGGGCACA</td>
<td>Allele 1: deletion of TG nucleotides; Allele 2: one insertion of T nucleotide; Biallelic mutant.</td>
<td>Positive</td>
<td>Negative</td>
<td>3:1</td>
</tr>
<tr>
<td></td>
<td>Allele 1 GGAACCTGAGTGTGAGATGTGAACGGGCACA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Allele 2 GGAACCTGAGTGTGAGATGTGAACGGGCACA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EC7</td>
<td>EC7 line GGAACCTGAGTGTGAGATGTGAACGGGCACA</td>
<td>Allele 1: no indels; Allele 2: deletion of TG nucleotides; Monoallelic mutant.</td>
<td>Positive</td>
<td>Positive</td>
<td>3:1</td>
</tr>
<tr>
<td></td>
<td>Allele 1 GGAACCTGAGTGTGAGATGTGAACGGGCACA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Allele 2 GGAACCTGAGTGTGAGATGTGAACGGGCACA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EC8</td>
<td>EC8 line GGAACCTGAGTGTGAGATGTGAACGGGCACA</td>
<td>Allele 1: deletion of T nucleotide; Allele 2: no indels; Monoallelic mutant.</td>
<td>Positive</td>
<td>Positive</td>
<td>3:1</td>
</tr>
<tr>
<td></td>
<td>Allele 1 GGAACCTGAGTGTGAGATGTGAACGGGCACA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Allele 2 GGAACCTGAGTGTGAGATGTGAACGGGCACA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EC9</td>
<td>EC9 line GGAACCTGAGTGTGAGATGTGAACGGGCACA</td>
<td>Allele 1: no indels; Allele 2: deletion of TG nucleotides; Monoallelic mutant.</td>
<td>Positive</td>
<td>Positive</td>
<td>3:1</td>
</tr>
<tr>
<td></td>
<td>Allele 1 GGAACCTGAGTGTGAGATGTGAACGGGCACA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Allele 2 GGAACCTGAGTGTGAGATGTGAACGGGCACA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EC10</td>
<td>EC10 line GGAACCTGAGTGTGAGATGTGAACGGGCACA</td>
<td>Allele 1: deletions of TG nucleotides; Allele 2: deletion of T nucleotide; Biallelic mutant.</td>
<td>Positive</td>
<td>Negative</td>
<td>3:1</td>
</tr>
<tr>
<td></td>
<td>Allele 1 GGAACCTGAGTGTGAGATGTGAACGGGCACA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Allele 2 GGAACCTGAGTGTGAGATGTGAACGGGCACA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
end of a *gfp* transgene (Figure 1b) previously inserted into the genome of *S. viridis* (Figure 1c), directing Cas9 cleavage of the sense strand of the target site using NGG nucleotides as the PAM sequence.

**Stable genetic transformation and genotyping of recovered genome editing events**

Embryogenic calli from transgenic plants overexpressing *gfp* were transformed with the engineered CRISPR/Cas9 binary vector to generate *indels* that disrupt the *gfp* transgene (Figure 2a). Ten independent transgenic events obtained following *Agrobacterium*-mediated transformation were sequentially named as EC1 to EC10. PCR analysis confirmed these lines carried the CRISPR/Cas9 minimal expression cassette (Figure 2b and 2c), which was confirmed by the presence of the BAR protein from the *bar* gene (Table 2). Plants from these events were compared to two independent *gfp* overexpressing

![Figure 2](image)

**Table 2.** Results of indirect ELISA assays (absorbance values at 405 nm) and quick test strip to serological detection, respectively, of GFP and BAR proteins in leaf samples from co-transforming events of *Setaria viridis*. Primary antibody against the C-terminal portion was used for GFP detection, while for BAR protein detection was used the QuickStix™ Kit for PAT/bar (EnviroLogix, Inc., USA).

<table>
<thead>
<tr>
<th>Sample</th>
<th>Absorbance values at 405 nm</th>
<th>ELISA for GFP</th>
<th>Quick test strip for BAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buffer</td>
<td>0.101</td>
<td>Negative</td>
<td>Negative</td>
</tr>
<tr>
<td>SvA10.1 (negative control; wild-type plants)*</td>
<td>0.332</td>
<td>Negative</td>
<td>Negative</td>
</tr>
<tr>
<td>ECWT1 (GFP-positive control 1)*</td>
<td>1.372</td>
<td>Positive</td>
<td>Negative</td>
</tr>
<tr>
<td>ECWT2 (GFP positive control 2)*</td>
<td>1.624</td>
<td>Positive</td>
<td>Negative</td>
</tr>
<tr>
<td>EC1</td>
<td>0.762</td>
<td>Positive</td>
<td>Positive</td>
</tr>
<tr>
<td>EC2</td>
<td>0.412</td>
<td>Negative</td>
<td>Positive</td>
</tr>
<tr>
<td>EC3</td>
<td>0.358</td>
<td>Negative</td>
<td>Positive</td>
</tr>
<tr>
<td>EC4</td>
<td>0.421</td>
<td>Negative</td>
<td>Positive</td>
</tr>
<tr>
<td>EC5</td>
<td>0.456</td>
<td>Negative</td>
<td>Positive</td>
</tr>
<tr>
<td>EC6</td>
<td>0.309</td>
<td>Negative</td>
<td>Positive</td>
</tr>
<tr>
<td>EC7</td>
<td>0.695</td>
<td>Positive</td>
<td>Positive</td>
</tr>
<tr>
<td>EC8</td>
<td>0.869</td>
<td>Positive</td>
<td>Positive</td>
</tr>
<tr>
<td>EC9</td>
<td>0.987</td>
<td>Positive</td>
<td>Positive</td>
</tr>
<tr>
<td>EC10</td>
<td>0.379</td>
<td>Negative</td>
<td>Positive</td>
</tr>
</tbody>
</table>

*a* = wild-type *Setaria viridis* A10.1 (non-transgenic; negative control for GFP in the indirect ELISA); *b* and *c* = transgenic plants positive control for *ZmUBI-1:GFP* and negative control for CRISPR/Cas9 NHEJ (used as positive controls for GFP protein in the indirect ELISA); EC1 to EC10 = co-transforming events PCR-positive to *ZmUBI-1:GFP* and CRISPR/Cas9 NHEJ minimal expression cassette.
Gene knockout in *Setaria viridis* with CRISPR/Cas9 and NEHJ

plants (GFP-positive control plants, named ECWT1 and ECWT2) and two non-transgenic wild-type plants (GFP-negative control plants). GFP accumulation in leaves of young gene edited plants, detected by fluorescence microscopy, presented three accumulation patterns: (i) absence, (ii) low GFP accumulation, and (iii) GFP accumulation equivalent to ECWT1 or ECWT2 control plants (Figure 2c). The results indicate different levels of allele disruption in each event. The morphology of adult T0 gene edited plants appeared similar to control plants, indicating transformation with the CRISPR/Cas9 binary vector did not cause pleiotropic alterations of growth and development (Figure 2d). The CRISPR/Cas9 minimal expression cassette segregated at a 3 : 1 ratio in the T1 progeny of events EC1 to EC10, indicating a single copy insertion for each event, as expected for *A. tumefaciens*-mediated transformation.

To identify the nature of the mutation introduced by gene editing in each event, purified PCR products amplified from the targeted genomic region in the T0 EC1-EC10 and GFP-positive control plants (ECWT1 or ECWT2) were sequenced by Sanger sequencing (Figure 3a). Sequencing revealed the presence of indels at the gRNA-targeted site in the *gfp* transgene of plants from each of the ten events and these changes appeared as either mono and biallelic indels (Table 1, Figure 3B).

Finally, the effect of these mutations on *gfp* expression and GFP protein accumulation in the adult T0 plants was determined with RT-qPCR and indirect ELISA assays, respectively. mRNA from the *gfp* transgene was observed in all plants from all events (Figure 3C), while GFP protein levels were different amongst the lines, ranging from low accumulation in plants with monoallelic mutations to no accumulation in plants with biallelic mutations (Table 2). These results indicate the mutations introduced by gene editing modified the mRNA sequence to encode either a truncated protein (resulting from a premature stop codon) or a protein with an amino acid sequence different from GFP (resulting from a frameshift mutation) (Supplementary File S1).

**Inheritance of introduced indels and the absence of off-target mutations in recovered events**

Previous studies in other plant species show that the mutations induced by the CRISPR/Cas9 system are stable and heritable as classical Mendelian alleles. Similarly, the mutations present in the T0 edited plants where heritable in the T1 generation and segregated as expected (Supplementary Figure S1A). Because CRISPR/Cas9 can target off-target sequences, we evaluated the specificity of our CRISPR/Cas9 system by sequencing the two genome locations that best matched the gRNA sequence.
according to in silico predictions (Supplementary Table S1). Sanger sequencing of PCR fragments amplified from these genomic regions showed no mutations in each region, demonstrating the specificity of gRNA targeting to the gfp gene sequence within transgene (Supplementary Figure S1B).

Discussion

The development of genome editing technology over the last 20 years has employed several types of nucleases (Boch et al. 2009; Christian et al. 2010; Durai et al. 2005; Joung and Sander 2013; Kim et al. 1996; Paques and Duchateau 2007; Smith et al. 2006; Streubel et al. 2012), but CRISPR/Cas9 has revolutionized genome editing because of its simplicity, versatility, efficiency, and specificity (Gil-Humanes et al. 2017; Li et al. 2016; Liang et al. 2017; Shan et al. 2014; Yin et al. 2017; Zhang et al. 2016). CRISPR/Cas9 has been used successfully in numerous organisms, including several plant species (Chang et al. 2016; Gao et al. 2015; Gerasimova et al. 2017; Jacobs et al. 2013, 2014; Li et al. 2016; Lowder et al. 2015; Michno et al. 2015; Yin et al. 2015; Zhang et al. 2016). However, C4 grasses like maize, sorghum, Brachiaria, switchgrass, and sugarcane have highly complex genomes, which contributes to low efficiency for genome editing (Mohan 2016). Despite the economic importance of C4 crop plants as food crops or biomass for biofuel production, there are still few studies reporting successful genome editing by CRISPR/Cas9 in these plant species (Gerasimova et al. 2017). Also, genetic manipulation of these crops is laborious and time-consuming, as long periods are usually necessary to advance generations, making functional genomic studies difficult (Basso et al. 2019; Basso et al. 2020). The use of model plants is a common alternative for functional genomics studies, as work in these translates to the target C4 crop species (Nguyen et al. 2020; Santos et al. 2020).

In this context, S. viridis serves as a model plant for important phylogenetically related C4 crops (Brutnell et al. 2015). Therefore, new resources and investigative tool development for S. viridis represent important advances. Huang et al. (2019) showed that CRISPR/Cas9 efficiently induced mutations in S. viridis to create new mutant alleles.

Here we describe a successful pipeline for S. viridis genome editing. We engineered a CRISPR/Cas9 binary vector to promote NHEJ. This vector coupled with a high-efficiency transformation protocol (Martins et al. 2015) generated ten independent gene editing events, all containing indels in the gfp target gene, in which 60% were biallelic. The efficiency achieved with this system was higher than reported with other approaches like transient expression of CRISPR/Cas9 components (Zhang et al. 2016), viral vectors (Yin et al. 2015), or ribonucleoprotein complexes (Liang et al. 2017; Woo et al. 2015). Also, the efficiency was similar to or higher than DNA-integration-based genome-editing methods used in other monocots (Svitashev et al. 2016). Further, the pipeline describe here was simple, accurate, and highly efficient with indels equivalent to the CRISPR/Cas9_Trex2 system (Weiss et al. 2020). A contributor to the efficiency of editing observed here may be the use of strong constitutive promoters driving Cas9 nuclelease and the gRNA, which improves the frequency of DSB due to high expression in transgenic cells at the early stages of dedifferentiation and differentiation (Gil-Humanes et al. 2017; Yang et al. 2017).

No indels occurred in potential off-target sites in the T0 events obtained here, confirming the specificity of the Cas9-gRNA combination, which is in part due to the low level of off-target cleavage when gRNAs are designed to be specific to target sequences (Young et al. 2019). The occurrence of indels in off-target sites and toxicity of strong and prolonged Cas9-gRNA expression in plant and bacteria are major concerns (Johnson et al. 2015; Lowder et al. 2015, 2016; Peng et al. 2016). However, highly target-specific gRNAs might eliminate the need for strong Cas9-gRNA expression (Osakabe et al. 2016). In addition, long term Cas9 cytotoxic effects have not been observed in most plant species. The use of the CRISPR/Cas9_Trex2 system to enhance indels in human cells indicated that Trex2 exonuclease increases frequency of off-target mutations (Chari et al. 2015). Weiss et al. (2020) demonstrated that CRISPR/Cas9_Trex2 system improves indels in S. viridis compared to the canonical CRISPR/Cas9 system, along with enhanced off-target mutations.

The results here showed codon optimized Cas9 nuclease was not cytotoxic for E. coli, A. tumefaciens, and S. viridis cells. Furthermore, indels introduced in S. viridis by this CRISPR/Cas9 system at single or multiple sites were stably inherited from the T0 to T1 generations, indicating Mendelian inheritance similar to that observed in other plant species (Yang et al. 2017; Zhang et al. 2016; Zhao et al. 2016) or even in S. viridis (Weiss et al. 2020). Transgene-free mutant plants can be obtained by self-crossing or backcrossing in the T1 generation from events carrying either homozygous or heterozygous for the CRISPR/Cas9 T-DNA (Zhang et al. 2014). Recently, transgene-free plants were obtained in the T0 generation with new strategies based on transient expression of CRISPR/Cas9 DNA or RNA (Zhang et al. 2016), viral vectors (Gil-Humanes et al. 2017; Yin et al. 2015), and ribonucleoprotein complexes (Liang et al. 2017; Svitashev et al. 2016; Woo et al. 2015).

In summary, we developed an optimized CRISPR/Cas9 binary vector, and this study confirms the efficiency of a pipeline for site-directed editing using this CRISPR/Cas9 system in S. viridis, similar to that shown by
Weiss et al. (2020). This work shows that promoting NHEJ pathway activity leads to accurate gene editing in *S. viridis*, with 60% of the transgenic events in a biallelic state at the T0 generation. This demonstrates the CRISPR/Cas9 NHEJ system is highly efficient in *S. viridis*. Furthermore, our results showed that a binary vector containing Cas9 nuclease codon-optimized for monocots and flanked by two NLS driven by strong constitutive promoters, and an optimized gRNA sequence driven by specific promoters were highly efficient in generating site-specific and biallelic indels in the *S. viridis* genome. Currently, our group is using this pipeline for genome editing in the sugarcane genome and we expect it to be effective in other important C4 grass and monocot crops.

**Conclusion**

In this work, we have successfully established a pipeline using the CRISPR/Cas9 NHEJ system for efficient editing and knock-out of target genes in *S. viridis*, which utilizes a new CRISPR/Cas9 binary vector, which exploits the NHEJ pathway, that is optimized and engineered for C4 monocots. This pipeline is expected to be useful for reverse genetics studies or to introduce desirable agronomic traits into economically important C4 crops such as maize, Brachiaria, wheat, rice, sorghum, and sugarcane.

**Acknowledgements**

We are grateful to Embrapa and FAPDF for their financial support. HBCM is grateful to FAPDF from Edital Universal 03/2016, Embrapa from Macroprograma number 03/2017 (13.16.05.026.00.00), and number 02/2017 (02.16.05.022.00.00). MFB is grateful to CNPq for the postdoctoral research fellowship (PD) 150936/2018-4. WRS is thankful to Fapesp (2019/04878-7) for the financial support. We also thank Dr. Frank G. Harmon to help improve the English writing of the revised manuscript.

**Authors’ contributions**

MFB, TRS, and HBCM planned and designed CRISPR/Cas9 NHEJ strategy for *gfp* transgene editing and knockout in *S. viridis* accession A10.1; KED and BDBC carried out the *S. viridis* *gfp* transformation; MFB carried out the co-transformation of *S. viridis* using the CRISPR/Cas9 NHEJ binary vector, conducted biological and molecular analyses, wrote the manuscript, and prepared the figures and tables; WRS, BOG, and AKK provided inputs in the manuscript. All authors revised and approved the final version of the manuscript.

**Competing interests**

All authors declare that the research was carried out in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Consent for publication or Ethics approval and consent to participate**

Not applicable.

**References**


Optimization of multiplexed CRISPR/Cas9 system for highly efficient genome editing in *Setaria viridis*. *Plant J* 104: 828–838


