Article

Functional validation of *cas9*/guideRNA constructs for site-directed mutagenesis of Triticale *ABA8'OH1 loci*

Krzysztof Michalski[1], Christian Hertig[2], Dariusz R. Mańkowski[3], Jochen Kumlehn[2], Janusz Zimny[1] and Anna M. Linkiewicz[1,4*]

[1] Plant Biotechnology and Cytogenetics Department, GMO Controlling Laboratory, Plant Breeding and Acclimatization Institute - National Research Institute, Radzików, 05-870 Błonie, Poland; krzysztof.michalski@ihar.edu.pl, orcid.org/0000-0002-3140-6551 (K.M); j.zimny@ihar.edu.pl orcid.org/ 0000-0002-1521-0874 (J.Z); a.linkiewicz@ihar.edu.pl (A.L)

[2] Plant Reproductive Biology, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) 06466 Seeland, OT Gatersleben, Germany; kumlehn@ipk-gatersleben.de, orcid.org/0000-0001-7080-7983 (J.K); hertig@ipk-gatersleben.de, orcid.org/0000-0002-2320-7579 (C.H)

[3] Laboratory of Seed Production and Plant Breeding Economics, Department of Seed Science and Technology, Plant Breeding and Acclimatization Institute - National Research Institute, Radzików, 05-870 Błonie, Poland, ORCID: 0000-0002-7499-8016; d.mankowski@ihar.edu.pl (D.M)

[4] Institute of Biological Sciences, Faculty of Biology and Environmental Sciences, Cardinal Stefan Wyszynski University in Warsaw, Wóycickiego 1/3 Street, 01-938 Warsaw, Poland, e-mail: a.linkiewicz@uksw.edu.pl; orcid.org/0000-0002-0167-3237 (A.L)

* Correspondence: a.linkiewicz@uksw.edu.pl

Abstract: Cas endonuclease-mediated genome editing provides a long-awaited molecular biological approach to the modification of predefined genomic target sequences in living organisms. Although *cas9*/guide (g)RNA constructs are straightforward to assemble and can be customized to target virtually any site in the plant genome, the implementation of this technology can be cumbersome, especially in species like Triticale that are difficult to transform, for which only limited genome information is available and/or which carry comparatively large genomes. To cope with these challenges, we have pre-validated *cas9*/gRNA constructs (1) by frameshift restitution of a reporter gene co-introduced by ballistic DNA transfer to barley epidermis cells, and (2) via transfection in Triticale protoplasts followed by either a T7E1-based cleavage assay or by deep-sequencing of target-specific PCR amplicons. For exemplification, we addressed the Triticale *ABA 8'-HYDROXYLASE 1* gene, one of the putative determinants of pre-harvest sprouting of grains. We further show that in-del induction frequency in Triticale can be increased by TREX2 nuclease activity, which holds true for both well- and poorly performing gRNAs. The presented results constitute a sound basis for the targeted induction of heritable modifications in Triticale genes.

Keywords: genome editing, CRISPR, protoplasts, targeted mutagenesis, TREX2, construct validation, transient expression

1. Introduction

Genome editing in cereals greatly facilitates genetic improvements that would have been more cumbersome with previously existing mutagenesis tools. A particularly powerful platform for site-specific changes in plant genomes is now available, for which components of the microbial clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) endonuclease immune system have been repurposed. These RNA-guided Cas endonucleases have been employed to generate loss-of-function alleles by targeting single or multiple *loci* in bread wheat, barley, rice and other grass species but not yet in Triticale. Up to date, reports of successful genome editing using Cas9 in monocots are limited mainly to diploid species like rice, maize or barley [1-2]. Few reports of successful employment in hexaploid common wheat [3-6] suggest that low transformation efficiency for wheat as well as the large genome size makes it exceedingly difficult for the Cas9/gRNA complexes to find their target. Some authors also argue that gene

knockout via non-homologous end joining might be inefficient in polyploids due to its genetic redundancy [7-8]. Here we have designed a suite of constructs to examine the potential for Cas9-mediated editing in Triticale that is another hexaploid species. The complexity of the Triticale genome and the difficulty of monocotyledonous plant transformation present a challenge, which is reflected by limited results hitherto reported for this species. It is necessary to improve the specificity and efficiency of Cas9 on-target mutagenesis and to develop a straightforward *in vivo* validation system for cas9/gRNA constructs, since the generation of transgenic Triticale plants is not only cost and time-consuming, but has been achieved only by very few laboratories worldwide.

Despite the great potential of the new biotechnological tools, not all *cas9*/gRNA construct designs are equally successful, which is due to several factors as reviewed by [9], [10] and [11]. Since the Cas9 mode of action, besides its requirement of a protospacer-adjacent motif (PAM), relies on a specific interaction between gRNA and the target sequence, it is of great importance to implement particularly suitable sequences in *cas9*/gRNA constructs used for stable plant transformation [10]. Directing the Cas9 nuclease to a specific *locus* in the genome is mainly facilitated by the user-defined gRNA. Currently available web-based tools for gRNA design offer a variety of candidate target motifs for a given gene of interest [12]. Despite these *in silico* predictions, not every design demonstrates equal efficiency in Cas9-catalyzed cleavage. Therefore, it is advisable to pre-validate several gRNAs to identify genomic target motifs that are likely to be effectively processed. In addition, it has been demonstrated in zebrafish that some gRNAs with high *in vitro* predicted activity possess poor *in vivo* activity, suggesting the presence of factors that restrict chromatin availability and limit the mutagenesis *in vivo* [13]. Conditions reflecting chromatin availability are hard to ensure by testing gRNA solely based on a PCR product, which is why the utilization of *in planta* systems are considered a more reliable principle for gRNA validation.

In the present investigation, we report the evaluation of *cas9*/gRNA construct performance *in planta* based on two methods; first, a transient expression assay developed in barley [14] and second, by Triticale protoplast transfection. We have included the *THREE PRIME REPAIR EXONUCLEASE 2 (TREX2)* sequence in some of the constructs to evaluate its utility for increased mutagenesis efficiency in Triticale. Cas9-mediated site-directed mutagenesis was used to disrupt *TsABA8'OH1 loci* in Triticale cells. The chosen target gene *TsABA8'OH1* [15-16] controls ABA catabolic degradation in grains and is present in Triticale as a single copy gene on each of its three group 6 chromosomes. Target sites on chromosomes 6A, 6B of the wheat reference genome and on chromosome 6R of the rye reference genome were selected for the conducted analyses.

2. Results

2.1. Selection of target motifs

The availability of Triticale genomic sequence data is limited. Therefore, genomic sequences of *ABA8'OH1* of the A and B subgenomes of bread wheat (GenBank accession no. AB714574.1 and AB714575.1, respectively) and a contig containing the putative rye homolog from the USDA database (Lo7_v2_contig_2868507) were used to design common primers for the simultaneous amplification of all three homeologues of the target gene (Figure S1).

To uncover potential nucleotide polymorphisms between these data and the Triticale subgenomes, which may affect targeted mutagenesis, the first two exons of *TsABA8'OH1* from the Polish winter Triticale variety Bogo, an accession that has been used as experimental model for genetic engineering, were amplified using common primers (Table S1) and then cloned in plasmids and sequenced (Table S1). The resulting genomic comparison of *TsABA8'OH1* sequences from cv. Bogo (NCBI MW538321-MW538323) showed that the percentage in identity of these regions to the reference sequences ranged from 98.7% for the R genome to 100% for the B genome (Figure 1). By contrast, essential differences were seen within the introns of the *TsABA8'OH-1* homoeologues, which facilitated the design of subgenome-specific primers for PCR amplification (Figure S1).
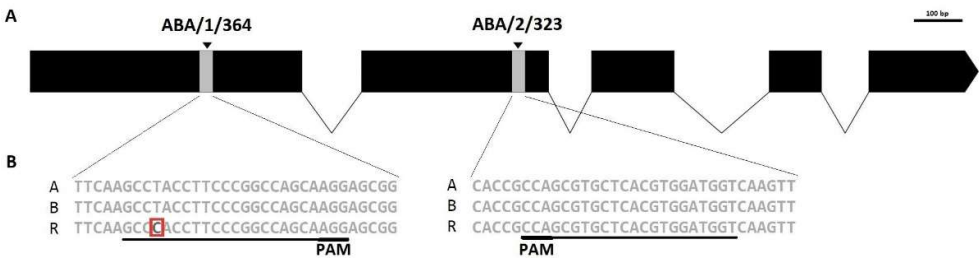

Figure 1. (A) Schematic diagram of target ABA8'OH1 regions used for selection of two gRNAs. Black boxes indicate exons, black lines indicate introns, gray boxes represent the regions in exon 1 and exon 2 designated for Cas9/gRNA-induced mutagenesis. (B) Detailed sequence context of target motifs (underlined, with PAM indicated) on A, B and R Triticale sub-genomes. Notably, one SNP is present in the target motif ABA/1/364 of the R genome.

2.2. Utility of Cas9/gRNA target motifs in ABA8'OH-1 of Triticale

Three different approaches were pursued to evaluate the editing activity of Cas9/gRNA customized for two target motifs in the *ABA8'OH-1* gene of Triticale. This involved the computational prediction of genomic sequences suitable to Cas9-generated double strand breaks (DSBs) and off-target site prediction in wheat and rye genomes. In a second step, *cas9*/gRNA constructs were evaluated *in planta* as previously described by Budhagatapalli et al. [14]. In this method, the mutation activity of customized endonuclease constructs (Figure 2A) is estimated based on the restitution of a reporter gene upon ballistic DNA transfer to barley epidermis cells. In addition, Cas9/gRNA activity was validated at the sequence level via Triticale protoplast transfection followed by amplification of target regions and sequencing.
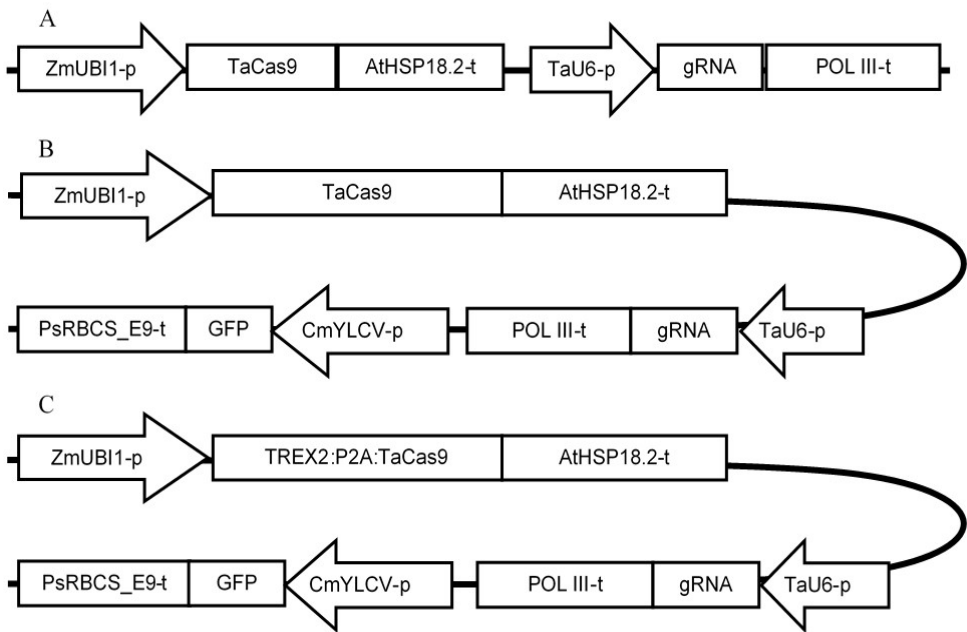
Figure 2. Schematic illustration of the *cas9* and gRNAs expression cassettes in (A) the transient expression vector used for the barley transformation experiment, and a binary vectors lacking (B) or containing (C) TREX2 nuclease sequence used for protoplast transformation. TaU6-p - *Triticum aestivum U6* promoter, POL III-t - *RNA POLYMERASE III* terminator, CmYLCV-p - strong constitutive promoter from *Cestrum yellow leaf curling virus*, GFP - sequence of *GREEN FLUORESCENT PROTEIN* gene, ZmUBI1-p - Zea mays *POLYUBIQUITIN 1* promoter, gRNA - single guide RNA scaffold, PsRBCS_E9-t - pea *RIBULOSE BISPHOSPHATE CARBOXYLASE SMALL SUBUNIT* terminator, TaCas9 - *Triticum aestivum* codon-optimized *CRISPR-associated protein 9*, AtHSP18.2-t - *Arabidopsis thaliana HEAT SHOCK PROTEIN 18.2* terminator, TREX2 - monocot codon-optimized human *THREE PRIME REPAIR EXONUCLEASE* 2, P2A - *Porcine teschovirus-1 2A* self-cleaving peptide motif.

2.3. Selection of target motifs under consideration of predicted on- and off-targets

Two gRNAs (gRNA-ABA/1/364 and gRNA-ABA/2/323) were selected for gene editing based on their *in silico* predicted on-target activity, whereby the scores for gRNAs ranged from 51.3 to 60.9 according to the Cas-Designer online tool [17]. gRNA-ABA/1/364 and gRNA-ABA/2/323 were chosen to target all Triticale homeoalleles of *TsABA8'OH1*, with the gRNA-ABA/1/364-addressed target motif containing a single mismatch 17 nt upstream of the PAM site in the R genome (Figure 1).

It is essential to ensure that the gRNA sequences match their cognate target *loci*, but do not match additional sites within the genome. Off-targets predicted by the Cas-Designer tool and a simple BLAST of the wheat and rye reference genomes featured at least two mismatches in comparison to their on-target counterparts (Table 1).

Table 1. Polymorphisms between on-target motifs and potential off-targets of the cas9/gRNA constructs used in this study. Predicted specificity of cas9/gRNA editing in Triticale based on common wheat and rye genome induced off-target events. PAM sequence is underlined, mismatches as opposed to the target are indicated by red letters.

Chromosome	Species	Sequence 5'-3'	Target

6A/6B	Triticale	GCCTACCTTCCCGGCCAGCA<u>AGG</u>	On-target motif ABA/1/364
6R	Triticale	GCCCACCTTCCCGGCCAGCA<u>AGG</u>	On-target motif ABA/1/364
6R	rye	GCCGACGTTCCCGGCCAGCA<u>AGG</u>	Potential off-target motif ABA/1/364
7B	wheat	GCATACCTTCCCG−CCAGCA<u>TGG</u>	Potential off-target motif ABA/1/364
6A/6B/6R	Triticale	GACCATCCACGTGAGCACGC<u>TGG</u>	On-target motif ABA/2/323
3A	wheat	GACCATCCACGTGAG−ACTC<u>AGG</u>	Potential off-target motif ABA/2/323
3B	wheat	GACCAGCC−CGTGAGCACGC<u>CGG</u>	Potential off-target motif ABA/2/323
4A	wheat	G−CCAGCCACGTGAGCACGC<u>CGG</u>	Potential off-target motif ABA/2/323
5A/5B/5R	wheat, rye	GACGATCCAGGTGAGCACGC<u>TGG</u>	Potential off-target motif ABA/2/323

2.4. Modeling the secondary structure of guide RNAs

In addition to the target motif-specific 3' region, the gRNAs comprise a scaffold that is required for the formation of ribonucleoprotein complexes with Cas9 endonuclease. Three stem loop structures are essential for binding with the Cas9 protein. These stem-loops result from short self-complementary inverted sequence repeats within the gRNA scaffold [18]. Depending on its sequence, the target motif-specific region of the gRNA can interfere with the formation of these stem loops by binding complementary bases. Figure 3 shows that the 2D models of both tested gRNAs feature the three functionally essential stem loops and should therefore be well-suited for Cas9-mediated mutagenesis. This analysis further revealed that the frequency of the calculated gRNA-ABA/1/364 and gRNA-ABA/2/323 structures within their ensembles of variants is 4.4% and 6.6%, respectively.

2.5. Assessment of gene editing frequency via frameshift restitution in a transiently expressed reporter gene

We have cloned synthetic target regions of *ABA8'OH-1* into *Bam*HI and *Eco*RI sites of pNB1 (pTARGET) vector [14]. The performed transient expression test indicates cleavage
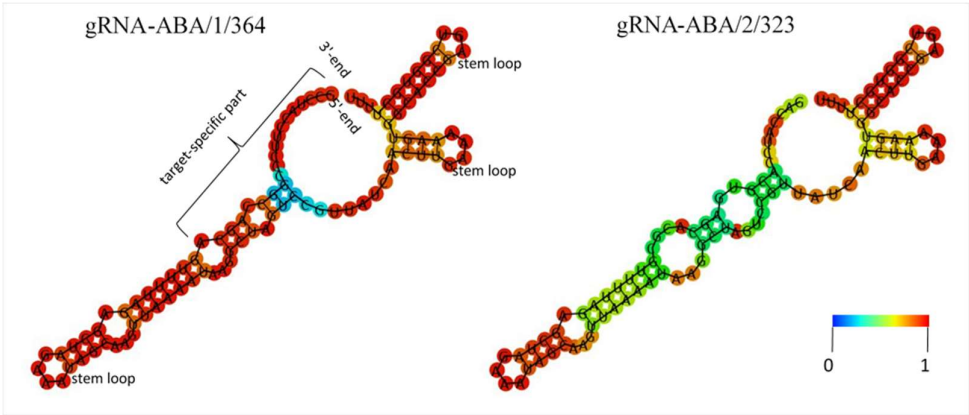
Figure 3. Secondary structure models of guide RNAs. Two-dimensional models of the minimum free energy structures of gRNA-ABA/1/364 and gRNA-ABA/2/323 as generated by the mfold online platform (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi)[19]. RNA ends, target-specific part and essential stem loops are indicated only for gRNA-ABA/1/364 but apply likewise for the other as well. The color code represents the base-pairing probabilities of individual nucleobases. For unpaired bases, the colors denote the probabilities of being unpaired.

activity via frameshift restitution of a *YFP* reporter gene. Barley leaves were co-bombarded with ABA8'OH-1_pTARGET vectors, pNB2 (*mCHERRY*) and vectors encoding specific gRNAs and Cas9. Construct activity was calculated as the proportion of cells with restored YFP functionality among all transformed cells showing mCHERRY signal [14]. Both gRNA-ABA/1/364 and gRNA-ABA/2/323 showed detectable activity and restored YFP functionality when the *cas9*/gRNA constructs were co-transfected with the *mCHERRY* reporter construct (Figure S3). While gRNA-ABA/1/364 yielded only 5.6% of YFP-positive cells among all mCHERRY expressing cells in barley, gRNA-ABA/2/323 led to *YFP*-frameshift repair cells in 24.3% of transfected cells on average (Table 2).

Table 2. Relative cleavage activity upon transient expression of cas9/gRNA and YFP/mCHERRY test constructs in barley leaf epidermis. Mean editing efficiency values ± SD indicated in bold letters.

	Number of red fluorescing cells	Number of yellow fluorescing cells	Proportion of yellow out of total fluorescing cells
	33	2	6.1%
gRNA-ABA/1/364	110	6	5.5%
	38	2	5.3%
			5.6 ±0.42%
	62	14	22.6%
gRNA-ABA/2/323	101	27	26.7%
	72	17	23.6%
			24.3 ±0.42%

2.6. Assessment of gene editing frequency in Triticale protoplasts using T7E1 assay

Here we have adopted a protocol of wheat protoplast transfection [20] for Triticale and optimized the process of protoplast isolation and transfection in the latter species. In order to validate the functionality of assembled gRNA expression units (Figure 2B, C) and to assess their efficiency in targeted mutagenesis, Triticale protoplasts were transfected with customized *cas9*/gRNA constructs using PEG-mediated DNA transfer. Transfection efficiency was evaluated based on *GREEN FLUORESCENT PROTEIN* (*GFP*) expression in the protoplasts (Figure S2). The highest transient expression efficiency (ranging from 40%

to 60%) was achieved with a plasmid concentration of 25 μg per 100,000 cells and the density of protoplasts being adjusted to 2.5 x 10⁵ cells/mL prior to adding PEG.

To see if we can enhance the efficiency of mutagenesis for gRNA-ABA/1/364 and gRNA-ABA/2/323, *TREX2*-enhanced vectors (+TREX2) [21] were used for comparison with the conventional constructs (Table 3, Table S3A).

A *T7 endonuclease I* (T7E1) assay [22] was used to detect mutation events in the target motifs of transfected versus untreated protoplasts. T7E1 results confirmed that the used constructs were capable of inducing mutations in their cognate target motifs *in vivo*. We observed that the addition of the *TREX2* sequence in the constructs generally increased the editing efficiency. The TREX2-mediated improvement in editing efficiency was significant in the case of the Triticale A genome addressed by gRNA-ABA/1/364, where the increase was about 3.3 times (p<0.0001). Even higher differences were observed for gRNA-ABA/2/323 in the comparison of no-*TREX2* and TREX2+, where the editing efficiency in the R genome jumped from an undetectable level to 28.1% (Table 3). Regarding no-*TREX2* constructs, the gRNA-ABA/1/364 was observed to be the most active on the R genome locus, with the mean indel frequency being 24.5%, followed by the B (15.6 %) and A genome loci (6.1%) (p=0.01). A significantly lower number of edited sequences on A, B and R genomes were observed for no-*TREX2* gRNA-ABA/2/323 construct (p<0.0001)(Table 3).

Table 3. Gene editing efficiency for gRNA-ABA/1/364 and gRNA-ABA/2/323 vectors containing or lacking the TREX2 component, analyzed in transfected Triticale protoplasts using T7E1 assay. Mean editing efficiency values ± SD indicated in bold letters. ND - not determined.

	Editing efficiency [%]		
	genome A	genome B	genome R
gRNA-ABA/1/364	6.3	17.1	27.3
	6.8	11.1	13.9
	5.3	18.6	32.3
	6.2±0.8ᵃ	*15.6±4.0ᵃ,ᵇ*	*24.4±9.5ᵇ*
gRNA-ABA/1/364 +TREX2	14.9	35.2	32.7
	27.5	22.0	22.5
	17.8	38.9	14.1
	20.1±6.6ᵃ	*32.0±8.9ᵃ*	*23.1±9.3ᵃ*
gRNA-ABA/2/323	0.3	1.9	ND
	ND	1.4	ND
	0.4	1.1	ND
	0.2±0.2ᵃ	*1.4±0.4ᵇ*	*ND*
gRNA-ABA/2/323 +TREX2	26.2	35.4	22.2
	15.6	19.2	34.7
	27.5	19.0	27.3
	23.1±6.5ᵃ	*24.5±9.4ᵃ*	*28.1±6.3ᵃ*

2.7. Assessment of targeted mutagenesis frequency in Triticale protoplasts using deep-sequencing of short reads

To assess the quantity and pattern of induced mutations with or without *TREX2* in the constructs and to crosscheck the results from T7E1 analysis, PCR products amplified from the target regions of transfected protoplast-derived genomic DNA were subjected to deep-sequencing. The analysis of reads was conducted using Geneious Prime Software by which we were able to correctly map the reads to the respective reference sequences of A, B or R genomes. On average, about 368,000 high quality reads were generated for each sample. The total number of reads assigned to target motifs 1 and 2 on A, B and R genomes ranged from about 15,000 to 135,000.

We detected modifications in the target regions in all samples (Table 4), except for the negative controls, where modifications were not detectable beyond the level of sequencing errors. Deletions were the predominant type of occurring mutations, while some insertions and substitutions were also observed. The most abundant modifications in regions targeted by gRNA-ABA/1/364 and gRNA-ABA/2/323 vectors equipped with *TREX2* or not, are summarized in Table S2. We observed high variability in mutagenesis efficiency and generated mutation types between the two target motifs addressed, with mainly 11-99 bp or 2-10 bp deletions being detected in case of target ABA/1/364, and 2-10 bp deletions together with 1 to 99 bp insertions typically seen in target ABA/2/323 (Figure 4). In the experiment where no *TREX2* vectors were applied, the total efficiencies of identified modifications ranged from 2.0% for gRNA-ABA/2/323 to 30.4% in case of gRNA-ABA/1/364, with both of these values derived from the respective target motifs of the A genome.

We have noticed an increased mutation frequency in case of both gRNAs when the *cas9* gene was supplemented with the *TREX2* sequence, which resulted in nearly 1.7-fold increase of total mutagenesis efficiency for gRNA-ABA/1/364 and over 19-fold for gRNA-ABA/2/323. Notably, TREX2 activity significantly reduced the proportion of single nucleotide deletions for gRNA-ABA/2/323 and small (2-10 nt) deletions for gRNA-ABA/1/364 constructs, simultaneously increasing frequencies of medium (11-99 nt) and large (>100 nt) deletions in both gRNAs tested. Additionally, TREX2 significantly reduced the occurrence of medium-size insertions (Table S2, and Table S3B).

Regardless of the constructs used, we did not observed statistically significant differences in mutagenesis frequencies between the analyzed sub-genomes (Table 4 and Table S3C).

Table 4. Gene editing efficiency for gRNA-ABA/1/364 and gRNA-ABA/2/323 vectors containing or lacking the TREX2 component, analyzed in transfected Triticale protoplasts using amplicon deep-sequencing. Mean editing efficiency values ± SD indicated in bold letters.

	Editing efficiency [%]		
	genome A	genome B	genome R
gRNA-ABA/1/364	25.6	33.2	32.0
	27.6	20.7	23.6
	26.8	37.3	31.7
	26.7±1.0	*30.4±8.6*	*29.1±4.8*
gRNA-ABA/1/364 +TREX2	50.4	36.7	38.6
	63.4	48.4	62.9
	46.7	43.5	44.8
	53.5±8.7	*42.9±5.9*	*48.8±12.6*
gRNA-ABA/2/323	0.4	0.2	1.7
	2.4	2.4	5.5
	3.3	4.4	2.6
	2.0±1.5	*2.3±2.1*	*3.3±2.0*
gRNA-ABA/2/323 +TREX2	31.2	50.7	43.2
	15.2	33.0	22.8
	45.9	48.9	49.8
	30.7±15.3	*44.2±9.7*	*38.6±14.1*

```
gRNA-ABA/1/364
A             GTCACTCAGGCGCACCTCTTCAAGCCTACCTTCCCGGCCAGCAAGGAGCGGATGCTGGGC
[-17]  9.5%   GTCACTCAGGCGCACCTCTTCAA----------------GCAAGGAGCGGATGCTGGGC
[-12]  9.3%   GTCACTCAGGCGCACCTCTTCAAGCCTAC-----------CAAGGAGCGGATGCTGGGC
[-8]   6.3%   GTCACTCAGGCGCACCTCTTCAAGCCTACCT--------AGCAAGGAGCGGATGCTGGGC
[-18]  5.6%   GTCACTCAGGCGCACCTCTTC----------------AGCAAGGAGCGGATGCTGGGC
[-16]  5.0%   GTCACTCAGGCGCACCTCTTCAAGCC---------------AAGGAGCGGATGCTGGGC

B             GTTACTCAGGCGCACCTCTTCAAGCCTACCTTCCCGGCCAGCAAGGAGCGGATGCTGGGT
[-12]  8.2%   GTTACTCAGGCGCACCTCTTCAAGCCT-----------AGCAAGGAGCGGATGCTGGGT
[-18]  7.2%   GTTACTCAGGCGCACCTCTT-------------------CAGCAAGGAGCGGATGCTGGGT
[-13]  6.1%   GTTACTCAGGCGCACCTCTTCAAGCCTA-----------CAAGGAGCGGATGCTGGGT
[-14]  5.9%   GTTACTCAGGCGCACCTCTTCAAG-------------CAGCAAGGAGCGGATGCTGGGT
[-9]   5.6%   GTTACTCAGGCGCACCTCTTCAAGCCTACC--------AGCAAGGAGCGGATGCTGGGT

R             GTCACGCAGGCGCACCTCTTCAAGCCCACCTTCCCGGCCAGCAAGGAGCGGATGCTGGGC
[-12] 11.5%   GTCACGCAGGCGCACCTCTTCAAGCCCA-----------GCAAGGAGCGGATGCTGGGC
[-18]  5.9%   GTCACGCAGGCGCACCTCTTCA-----------------GCAAGGAGCGGATGCTGGGC
[-17]  5.4%   GTCACGCAGGCGCACCTCTTCAA----------------GCAAGGAGCGGATGCTGGGC
[-7]   4.5%   GTCACGCAGGCGCACCTCTTCAAGCCCACCTT-------AGCAAGGAGCGGATGCTGGGC
[-13]  4.3%   GTCACGCAGGCGCACCTCTTCAAGCC-------------AGCAAGGAGCGGATGCTGGGC

gRNA-ABA/1/364+TREX2
A             GTCACTCAGGCGCACCTCTTCAAGCCTACCTTCCCGGCCAGCAAGGAGCGGATGCTGGGC
[-12]  8.9%   GTCACTCAGGCGCACCTCTTCAAGCCT-----------AGCAAGGAGCGGATGCTGGGC
[-17]  6.8%   GTCACTCAGGCGCACCTCTTCA-----------------AGCAAGGAGCGGATGCTGGGC
[-21]  5.3%   GTCACTCAGGCGCACCTCTTCA--------------------AGGAGCGGATGCTGGGC
[-7]   4.2%   GTCACTCAGGCGCACCTCTTCAAGCCTACCTTCCC-------AAGGAGCGGATGCTGGGC
[-9]   3.8%   GTCACTCAGGCGCACCTCTTCAAGCCTACC--------AGCAAGGAGCGGATGCTGGGC

B             GTTACTCAGGCGCACCTCTTCAAGCCTACCTTCCCGGCCAGCAAGGAGCGGATGCTGGGT
[-21]  7.3%   GTTACTCAGGCGCACCTCTT-------------------CAAGGAGCGGATGCTGGGT
[-18]  5.9%   GTTACTCAGGCGCACCTCTTCA-----------------GCAAGGAGCGGATGCTGGGT
[-12]  5.0%   GTTACTCAGGCGCACCTCTTCAAGCCT-----------AGCAAGGAGCGGATGCTGGGT
[-17]  5.0%   GTTACTCAGGCGCACCTCTTCAAGC----------------AAGGAGCGGATGCTGGGT
[-14]  3.9%   GTTACTCAGGCGCACCTCTTCAAGCCTACCTTA-------------GCGGATGCTGGGT

R             GTCACGCAGGCGCACCTCTTCAAGCCCACCTTCCCGGCCAGCAAGGAGCGGATGCTGGGC
[-17]  7.4%   GTCACGCAGGCGCACCTCTTCAAGC----------------AAGGAGCGGATGCTGGGC
[-15]  5.2%   GTCACGCAGGCGCACCTCTTCAAGCCCA---------------AGGAGCGGATGCTGGGC
[-12]  5.1%   GTCACGCAGGCGCACCTCTTCAAGC-----------CCAGCAAGGAGCGGATGCTGGGC
[-21]  4.3%   GTCACGCAGGCGCACCTCTTC--------------------AAGGAGCGGATGCTGGGC
[-13]  3.5%   GTCACGCAGGCGCACCTCTTCAAGC-------------CAGCAAGGAGCGGATGCTGGGC

gRNA-ABA/2/323
A             GTCATCTTCGCCGCGCGCGACACCACCGCCAGCG TGCTCACGTGGATGGTCAAGTTCCT
[-1]  44.3%   GTCATCTTCGCCGCGCGCGACACCACCGCCAGCG -GCTCACGTGGATGGTCAAGTTCCT
[-2]  13.9%   GTCATCTTCGCCGCGCGCGACACCACCGCCAGCG --CTCACGTGGATGGTCAAGTTCCT
[+1]  12.9%   GTCATCTTCGCCGCGCGCGACACCACCGCCAGCGTTGCTCACGTGGATGGTCAAGTTCCT
[-5]   8.7%   GTCATCTTCGCCGCGCGCGACACCACCGCCAGCG -----ACGTGGATGGTCAAGTTCCT
[-4]   7.3%   GTCATCTTCGCCGCGCGCGACACCACCGCCAGCG ----CACGTGGATGGTCAAGTTCCT

B             GTCATCTTCGCCGCGCGCGACACCACCGCCAGCG TGCTCACGTGGATGGTCAAGTTCCT
[+1]  31.8%   GTCATCTTCGCCGCGCGCGACACCACCGCCAGCGATGCTCACGTGGATGGTCAAGTTCCT
[-1]  30.5%   GTCATCTTCGCCGCGCGCGACACCACCGCCAGCG -GCTCACGTGGATGGTCAAGTTCCT
[-3]   9.9%   GTCATCTTCGCCGCGCGCGACACCACCGCCAGCG ---TCACGTGGATGGTCAAGTTCCT
[-2]   7.7%   GTCATCTTCGCCGCGCGCGACACCACCGCCAGCG --CTCACGTGGATGGTCAAGTTCCT
[-5]   6.0%   GTCATCTTCGCCGCGCGCGACACCACCGCCAGCG -----ACGTGGATGGTCAAGTTCCT

R             GTCATCTTCGCCGCGCGGGACACCACCGCCAGCG TGCTCACGTGGATGGTCAAGTTCCT
[-1]  30.3%   GTCATCTTCGCCGCGCGGGACACCACCGCCAGCG -GCTCACGTGGATGGTCAAGTTCCT
[+1]  24.0%   GTCATCTTCGCCGCGCGGGACACCACCGCCAGCGTTGCTCACGTGGATGGTCAAGTTCCT
[-2]  12.8%   GTCATCTTCGCCGCGCGGGACACCACCGCCAGCG --CTCACGTGGATGGTCAAGTTCCT
[-3]  12.3%   GTCATCTTCGCCGCGCGGGACACCACCGCCAGCG ---TCACGTGGATGGTCAAGTTCCT
[-4]   5.1%   GTCATCTTCGCCGCGCGGGACACCACCGCCAGC- ---TCACGTGGATGGTCAAGTTCCT

gRNA-ABA/2/323+TREX2
A             GCCGCGCGCGACACCACCGCCAGCGTGCTCACGTGGATGGTCAAGTTCCTCGGCGACAAC
[-78]  8.7%   ------------------------------------------------TCGCCGACAAC
[-8]   6.3%   GCCGCGCGCGACACCACCGCCAG--------CGTGGATGGTCAAGTTCCTCGGCGACAAC
[-39]  5.9%   ----------------------------------------GGTCAAGTTCCTCGGCGACAAC
[-76]  4.9%   -----------------------------------------------------------C
[-10]  4.2%   GCCGCGCGCGACACCACCGCCAGC----------GGATGGTCAAGTTCCTCGGCGACAAC

B             GCGCGCGACACCACCGCCAGCGTGCTCACGTGGATGGTCAAGTTCCTCGGCGACAACCCC
[-10]  9.9%   GCGCGCGACACCACCGCCAG---------TGGATGGTCAAGTTCCTCGGCGACAACCCC
[-95]  9.8%   ---------------------------------------------------------CC
[-30]  6.2%   GCGCGCGACACCACCGCCAG------------------------------CGACAACCCC
[-8]   5.4%   GCGCGCGACACCACCGCCAG--------CGTGGATGGTCAAGTTCCTCGGCGACAACCCC
[-43]  5.3%   ------------------------------------------AAGTTCCTCGGCGACAACCCC

R             GTCATCTTCGCCGCGCGGGACACCACCGCCAGCGTGCTCACGTGGATGGTCAAGTTCCTC
[-78]  8.6%   ----------------------------------------------------------TC
[-119] 7.0%   ---------------------------------------TCACGTGGATGGTCAAGTTCCTC
[-28]  5.2%   GTCATCTTCGCCGCGCGGGACA-----------------------CAAGTTCCTC
[-24]  4.8%   GTCATCTTCGCCGCGCGGGACACCACCG----------------------AGTTCCTC
[-91]  4.8%   GTCATCTTC---------------------------------------------------
```

Figure 4. Relative frequency of the most abundant mutations among changes induced on each subgenome of Triticale protoplast by transient expression of cas9/gRNA-ABA/1/364 and cas9/gRNA-ABA/2/323 constructs, lacking or containing TREX2 enhancer. PAM motif bolded; target motif underlined. SNP present in the target motif ABA/1/364 of the R genome is marked in red.

3. Discussion

Estimation of Cas9/gRNA activity by computational methods may not be accurate for all applications as the criteria are derived from context-specific data [23]. Available methods for the analysis of on-target and off-target mutations have some limitations and some benefits that were comprehensively discussed in Zischewski et al. [24]. The ideal methods should take into account the particular genome edited, anticipated size and type of mutation as well as cost of the method. Thus an experimental validation of any given gRNA activity is advisable. It is highly beneficial to determine which constructs have the highest potential for successful genome editing, before starting with the cumbersome methods of genetic engineering at the whole-plant level. To date, quick and efficient transformation systems are still not available for some plant species, including Triticale. Therefore, protoplast transfection or transient transgene expression in other cell systems are useful validation tools for multiple mutagenesis parameters.

cas9/gRNA constructs can exhibit a broad range of efficiencies. Testing multiple gRNAs increases the chance of identifying a construct that is specific and highly active [25-26]. Several vector sets have been developed for genome editing in plants [27-29], [21], [30]. We have decided to use multiplasmid, flexible system developed by Čermák et al. [21] available from plasmid repository ADDGENE, that enables targeted, specific modification of monocot genomes, because its utility has been demonstrated by several authors [11, 12, 31]. To verify whether multiple homeoalleles of the *ABA8'OH-1* gene can be simultaneously addressed, we were taking into account the nature of individual target sites, gRNAs design, possibility of off-target activity and two different Cas9 variants.

Cas9 has been shown to preferentially binding DNA targets with purines in the four PAM-proximal bases of the targeting sequence, whereas pyrimidines and, especially, thymines were disfavored [32-33]. Working with human cells, Graf et al. [34] discovered that TT-motif and GCC-motifs within the four PAM-proximal bases of the target motif negatively impact the efficiency of gRNAs resulting in a about 10-fold reduction in editing frequency. The gRNAs used in the present investigation do not contain any TT-motifs. On the other hand, the four base PAM-proximal region of gRNA-ABA/1/364 contains three purines, whereas in gRNA-ABA/2/323 there are two. To find out whether these features have an impact on the formation and function of Cas9/gRNA complexes is beyond the scope of this study.

3.1. Off-target and on-target SNP evaluation

cas9/gRNA constructs typically include a 20-bp gRNA 5' end that directs the Cas9 nuclease to the target site by complementary base pairing. It is well established that the gRNA sequence plays a pivotal role in determining the on- and off-target activities of Cas9 [35-36]. Moreover, *in silico* analysis of gRNA sequences are only approximations to the experimental data. In general, off-target sites are not cleaved as efficiently in particular when mismatches occur near the PAM, so gRNAs with no homology or those with mismatches close to the PAM sequence will have the highest specificity [37]. Our results show that the closest sequence to the targets of gRNA-ABA/1/364 and gRNA-ABA/2/323 has at least two mismatches within the 20 bp region bound by the cognate gRNA. However, potential off-target matches on the 6R for gRNA-ABA/1/364 and on the 5A/5B/5R for gRNA-ABA/2/323 should be evaluated more carefully than the others, as the SNPs are not located in proximity to the PAM.

We have found common target motifs on A, B and R homeologous Triticale genomes in exon 1 and 2 of *ABA8'OH-1*, having only one mismatch for the gRNA-ABA/1/364 that might influence the processing of the R genome. However, the R genome was as efficiently mutated as those derived from wheat. A possible explanation is that the SNP position is outside the core region of the gRNA-ABA/1/364 target sequence [37].

3.2. Protoplast transfection as a convenient test for Cas9/gRNA functionality

In the present study, we assessed and verified the activity of two gRNAs targeting the *ABA8'OH-1* gene using protoplast transfection and transient expression in the barley

leaf epidermis. The major advantage of protoplast transfection followed by amplicon sequencing as compared to the expression of test constructs in epidermis cells is that the Cas9/gRNA target motifs are processed in their chromosomal context. Protoplasts from model species such as *Arabidopsis* and tobacco as well as crops like rapeseed, rice, wheat and maize have been used to evaluate reagents of Cas9-based systems [38-39]; [20]; [40-41]. However, this method requires experience in terms of handling delicate cell cultures and a careful optimization of conditions depending on the species used. Here, we used a protoplast isolation and transfection protocol previously available for wheat, adapted it for Triticale and eventually achieved Cas9/gRNA-triggered mutagenesis with useful efficiency.

Methods like deep-sequencing of PCR amplicons or the T7E1 cleavage assay are available to assess targeted mutagenesis. In the present study, both approaches were applied and compared. The T7E1 assay is widely used, cost-effective, technically straightforward, and easy to interpret. However, Vouillot et al. [22] revealed that 1 bp indels are not recognized by this method, which is why it is likely to result in underestimations of the total mutagenesis efficiency. In line with this, we were not able to detect any modification in the R genome and only very few on the A and B sub-genomes after transfection of the construct harboring the no-*TREX2* gRNA-ABA/2/323. By contrast, the deep-sequencing data revealed 2.0-3.3% total mutagenesis efficiency when using the same construct, coming mostly from 1 bp and 2-10 bp deletions and small insertions. No such bias between T7E1 and amplicon sequencing was observed when TREX2+ constructs were used.

While most of the observed mutations in protoplasts were deletions and single nucleotide substitutions, insertions were only rarely found. The mutation efficiency demonstrated by deep-sequencing was as high as 48.8%. This result suggests that our best-performing construct may also be useful at the whole-plant level.

3.3. Gene editing frequency in plants can be estimated via transient expression

A biolistic-based transient expression test for *cas9*/gRNA constructs using leaf epidermis was previously established and used in barley [42-43]. Here, we were able to employ this method for evaluation of no-*TREX2* constructs through co-expression with Triticale-derived targets in barley leaves. Both, gRNA-ABA/1/364 and gRNA-ABA/2/323 showed acceptable efficiency in this test, which is contradictory to the results obtained in Triticale protoplasts by deep-sequencing, where gRNA-ABA/2/323 no-*TREX2* construct showed very poor activity. Possible reasons for such inconsistent results might be that not only the different species barley and Triticale, but also different cell types were used, in which DNA repair underlie specific circumstances. Moreover, the non-genomic nature of the target plasmids co-introduced into barley epidermis cells is likely to be of some importance. Despite the observed discrepancies, the deep-sequencing analysis strongly suggests that this approach is more suited and accurate for the detection of small indels. It is within reason to suspect that smaller insertions and deletions are more likely to repair the *YFP* reading frame as it happened in case of gRNA-ABA/2/323. Meanwhile, the larger indels, even when correcting the frame-shift, might cause some unwanted changes to YFP protein activity.

3.4. The effect of TREX2 co-expression on Cas9-mediated gene editing

Some additional strategies can be pursued to further improving the editing efficiency of *cas9*/gRNA constructs. To increase the mutagenesis efficiency achieved by Cas9, we have tested the TREX2 enhancer for its utility as a translationally fused domain [44-45]. TREX2 was reported to increase frequencies of heritable mutations in tomato and barley protoplasts by 2.5-fold [21]. Similarly, targeted mutagenesis efficiency was further improved by TREX2 in *Setaria viridis* by an average of 1.4-fold [31]. The *TREX2* gene encodes a nuclear protein with 3' to 5' exonuclease activity. The encoded protein participates in double-stranded DNA break repair [46]. TREX2 does increase mutation rates across both on- and off-target sites to varying degrees. We have found that the TREX2 significantly

increased editing efficiency in Triticale even for the low performing gRNA-ABA/2/323, which was equally clear in both T7E1 and deep-sequencing assays. In the latter, it was observed that *TREX2*-including vectors caused an increased occurrence of >10 bp deletions, while reducing small (1-10 bp) deletion and insertions. This observation is in accord with the results of Čermak et al. [21].

4. Materials and Methods

4.1. Plant material and ABA8'OH-1 identification in Triticale

Triticale cv. Bogo (EGISET, Accession Number: 66000), a highly regenerable hexaploid Triticale cultivar, was used as a material source in all described experiments. If not stated otherwise, 7 days-old plants were vernalized for 8 weeks at 4 °C and then cultivated under controlled conditions in a growth chamber with 16 h photoperiod and 21 °C temperature.

Genomic sequences of *ABA8'OH-1* genes corresponding to both A and B genomes of bread wheat were verified and downloaded from the National Center for Biotechnology Information (NCBI). Wheat sequences were used to find scaffolds containing rye homologs in the USDA database. Homologous sequences of A, B and R genomes were aligned and conserved regions in exon 1 and exon 3 were used to design common primer pairs (Figure S1) [47, 48].

DNA isolated from fresh leaves was used as a template in PCRs. Reaction products were cloned directly into pJET2.1 (CloneJET PCR Cloning Kit; Thermo Fisher Scientific) vector and transformed into E. coli strain DH5α with CaCl$_2$/heat shock method. Twenty single colonies were used for plasmid isolation with Qiagen MiniPrep Kit and sent for Sanger sequencing.

4.2. gRNA design, 2D modeling and vectors construction

Cas-Designer on-line tool [17] was used to design gRNA candidates based on wheat A genome sequence and to predict the specificity of target to B and R genomes based on bread wheat genome (assembly: IWGSC1.0, 2014) and rye contig (Lo7_v2_contig_2868507). Only targets adjacent to an appropriate protospacer adjacent motif (PAM) for Cas9 (NGG) were considered. For the purpose of the study and based on the specificity to each and every Triticale genomes, two gRNA were selected for the gene and named gRNA-ABA/1/364 and gRNA-ABA/2/323, thereby indicating the addressed exons (1 or 2) and the nucleotide positions within the gene. To assess the formation of the secondary structures of the gRNAs, the two preselected gRNAs were subjected to 2-dimensional modeling, for which the online tool mfold (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi) was used [19]. This tool determines and displays the minimum free energy structural variant of a given RNA sequence, i.e., the one among the many possible structures of an RNA that has the highest stability and thus is present with the highest abundance.

Plasmids developed by the Voytas Lab [21] were used for two-step, GoldenGate-based construction of expression vectors. This plasmid series provided starting material to generate a RNA-guided Cas9 tool for gene editing in plants. At first, synthetic oligonucleotides, ordered externally, were annealed and cloned into pMOD_B vectors. In the second step, Cas9 or Cas9+TREX2-coding pMOD_A plasmids were mixed with pMOD_B and pMOD_C plasmids and cloned into pTRANS expression vectors. Different pMOD_A, pMOD_C and pTRANS plasmids were used, depending on downstream application.

4.3. Transient expression method

Synthetic oligos coding target sequences were introduced into pNB1 vector causing non-sense frameshift in *YELLOW FLUORESCENCE PROTEIN* (*YFP*) gene [14]. Thus acquired pTARGET vectors were mixed with mCHERRY-coding pNB2 and *cas9*/gRNA vectors and coated onto 1 μm gold particles. BioRad PDS-1000 system was used for bombardment of leaves taken from 10-days old barley cv. Golden Promise seedlings. Plasmids used

for this assay contained only *cas9* and gRNA cassettes with no *TREX2* and GFP coding sequences. Mixes of empty (with no frameshift in *YFP*) pNB1 and pNB2, as well as pNB2 and one of pTARGET vectors were used as positive and negative controls, respectively. After 48 h incubation, leaf fragments were fixed to glass slides and subjected to observation in Zeiss LSM780 confocal laser microscope. Cas9/gRNA activity was calculated as the ratio between the number of cells with restored YFP signal compared to cells with mCHERRY signal.

4.4. Protoplast transfection assay

Methodology previously optimized for wheat protoplasts transfection [20] was adapted for Triticale. Triticale protoplasts were transfected with preassembled vectors via polyethylene glycol (PEG)-mediated delivery. Seven days-old, etiolated seedlings of Triticale cv. Bogo were used for protoplast isolation. The central part of the second leaves were chopped with new razor blades and incubated for 5 h in W5 containing cellulase R-10 (Duchefa) and macerozyme R-10 (Duchefa) at 25 °C. Protoplasts were then collected and immediately transfected with 25µg of plasmid vector per 100 thousand cells. The four vectors were transfected using PEG treatment into Triticale protoplasts. After 48h incubation in darkness, at room temperature, without shaking, transfection efficiency was evaluated based on *GREEN FLUORESCENT PROTEIN* (GFP) expression. Genomic DNA was extracted from transfected protoplasts and used for PCR. Fragments of the *ABA'OH-1* targeted sequences were amplified using PCR with the specific primers (Table S1), cleaned-up, re-annealed and digested with T7 endonuclease. Digestion products were visualized with BioAnalyzer 2100 using High Sensitivity DNA chips. Cas9/gRNA activity was calculated based on the ratio of digested and undigested products.

4.5. Deep-sequencing of amplicons

Genomic DNA isolated previously from transfected protoplast was used to amplify target regions with primers containing adaptors for subsequent NGS library synthesis. Library synthesis and Illumina HiSeq 2000 sequencing were conducted by a third party company. Thus acquired reads were then trimmed and mapped to reference sequences with Geneious Prime (v2020.2.4) software [49]. Finally, a build-in tool was used to count Cas9-induced modification frequencies in sequences annotated to particular genomes

4.6. Statistical analysis

In order to assess the differences in editing efficiency between constructs, a series of one-way analyses of variance were performed together with the Tukey method for multiple comparison procedure. Analyses were performed for all data combined and divided by each genome. For the variance analysis Bliss transformation was applied on the data and Statistica program vol. 13.3 (TIBCO Software, Inc., 2017) was used [50].

Supplementary Materials: The following are available online at www.mdpi.com/xxx/s1,

Figure S1: TsABA8'OH partial sequences of Triticale cv. Bogo generated in the study.

Figure S2: Population of Triticale mesophyll protoplasts transfected with CRISPR/Cas9 coding plasmid after 48h of incubation.

Figure S3. Barley leaf transient expression assay.

Table S1: Sequence of primers used for PCR and sequencing.

Table S2. Summarized deep-sequencing data showing the share of distinct mutation types in the A, B and R Triticale sub-genomes generated through protoplast transfection with four tested constructs.

Table S3. Editing efficiency of constructs, based on one-way analyses of variance performed together with the Tukey method for multiple comparison.

References

1. Ma X., Liu Y.G. CRISPR/Cas9-Based Multiplex Genome Editing in Monocot and Dicot Plants. *Curr Protoc Mol Biol*. **2016** Jul 1; *115*:31.6.1-31.6.21. doi: 10.1002/cpmb.10.

2. Hisano H., Abe F., Hoffie R.E. and Kumlehn J. Targeted genome modification in cereals. *Breeding Science* **2021,** in press

3. Sánchez-León S., Gil-Humanes J., Ozuna C.V., Giménez M.J., Sousa C., Voytas D.F., et al. Low-gluten, nontransgenic wheat engineered with CRISPR/Cas9. *Plant Biotechnol J*. **2018**.*16*: 902–10. doi: 10.1111/pbi.12837

4. Abe F., Haque E., Hisano H., Tanaka T., Kamiya Y., Mikami M., Kawaura K., Endo M., Onishi K., Hayashi T., et al. Genome-Edited Triple-Recessive Mutation Alters Seed Dormancy in Wheat. *Cell Rep*. **2019**, *28* pp. 1362-1369.e1364.

5. Li, J., Jiao, G., Sun, Y., Chen, J., Zhong, Y., Yan, L., Jiang, D., Ma, Y. and Xia, L. Modification of starch composition, structure and properties through editing of *TaSBEIIa* in both winter and spring wheat varieties by CRISPR/Cas9. *Plant Biotechnol. J.* **2020**. pp.1-15, https://doi.org/10.1111/pbi.13519.

6. Budhagatapalli N., Halbach T., Hiekel S., Büchner H., Müller A.E., Kumlehn J.. Site-directed mutagenesis in bread and durum wheat via pollination by cas9/guide RNA-transgenic maize used as haploidy inducer. *Plant Biotechnol J*. **2020,** 18(12):2376-2378. doi: 10.1111/pbi.13415.

7. Kumar, R., Kaur, A., Pandey, A. et al. CRISPR-based genome editing in wheat: a comprehensive review and future prospects. *Mol Biol Rep* **2019**, *46*, 3557–3569. https://doi.org/10.1007/s11033-019-04761-3

8. Arndell, T., Sharma, N., Langridge, P. *et al.* gRNA validation for wheat genome editing with the CRISPR-Cas9 system. *BMC Biotechnol* **2019,** *19*, 71. https://doi.org/10.1186/s12896-019-0565-z

9. Paul J.W. 3rd, Qi Y. CRISPR/Cas9 for plant genome editing: accomplishments, problems and prospects. *Plant Cell Rep*. **2016**; *35*(7):1417-1427. doi:10.1007/s00299-016-1985-z

10. Kumlehn J., Pietralla J., Hensel G., Pacher M., Puchta H. The CRISPR/Cas revolution continues: From efficient gene editing for crop breeding to plant synthetic biology. *J Integr Plant Biol*. **2018,** *60*(12):1127-1153. doi:10.1111/jipb.12734

11. Shockey J. Gene editing in plants: assessing the variables through a simplified case study. *Plant Mol Biol*. **2020**; *103*(1-2):75-89. doi:10.1007/s11103-020-00976-2

12. Koeppel I., Hertig C., Hoffie R. and Kumlehn J. Cas endonuclease technology - a quantum leap in the advancement of barley and wheat genetic engineering. *International Journal of Molecular Sciences* **2019,** *20*, 2647

13. Uusi-Mäkelä M.I.E., Barker H.R., Bäuerlein C.A., Häkkinen T., Nykter M., et al. Chromatin accessibility is associated with CRISPR-Cas9 efficiency in the zebrafish (*Danio rerio*). *PLOS ONE* **2018,** *13*(4): e0196238. https://doi.org/10.1371/journal.pone.0196238

14. Budhagatapalli N., Schedel S., Gurushidze M., Pencs S., Hiekel S. , Rutten T., Kusch S., Morbitzer R., Lahaye T., Panstruga R., Kumlehn J and Hensel G. A simple test for the cleavage activity of customized endonucleases in plants. *Plant Methods* **2016,** *12*:18 DOI 10.1186/s13007-016-0118-6

15. Okamoto M., Kuwahara A, Seo A., Kushiro T., Asami A., Hira Ni., Yuji Kamiy Y., Koshiba T., Nambara E. CYP707A1 and CYP707A2, Which Encode Abscisic Acid 8'-Hydroxylases, Are Indispensable for Proper Control of Seed Dormancy and Germination in Arabidopsis. *Plant Physiology* **2006**, *141* (1) 97-107; DOI**:**10.1104/pp.106.079475

16. Fidler J., Grabowska A., Prabucka B., Więsyk A., Góra-Sochacka A., Bielawski W., Pojmaj M., Zdunek-Zastocka E. The varied ability of grains to synthesize and catabolize ABA is one of the factors affecting dormancy and its release by afterripening in imbibed triticale grains of cultivars with different pre-harvest sprouting susceptibilities. *Journal of Plant Physiology* **2018,** *226*:48-55.

17. Park J., Bae S., and Kim J.-S. Cas-Designer: A web-based tool for choice of CRISPR-Cas9 target sites. *Bioinformatics* **2015,** *31*, 4014-4016 ().

18. Wong, N.; Liu, W.; Wang, X. WU-CRISPR: Characteristics of functional guide RNAs for the CRISPR/Cas9 system. *Genome Biol*. **2015**, *16*, 218.

19. RNAfold web server. Available online: http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi (accessed on 25.04 2021).

20. Shan, Q.W., Wang, Y.P., Li, J. and Gao, C.X. Genome editing in rice and wheat using the CRISPR/Cas system. *Nat. Protoc*. **2014**, *9*, 2395–2410.

21. Čermák T., Curtin S.J., Gil-Humanes J., Čegan R., Kono T.J.Y., Konecna E., Belanto J.J., Starker C.G., Mathre J.W., Greenstein R.L., Voytas D.F. A multi-purpose toolkit to enable advanced genome engineering in plants. *Plant Cell*. **2017** *18*. pii: tpc.00922.2016. doi: 10.1105/tpc.16.00922.

22. Vouillot L., Thelie A., Pollet N. Comparison of T7E1 and surveyor mismatch cleavage assays to detect mutations triggered by engineered nucleases. *G3 (Bethesda)* **2015**, *5*:407–415.

23. Gagnon J.A. Valen E. Thyme S.B. Huang P. Ahkmetova L. Pauli A. Montague T.G. Zimmerman S. Richter C. Schier A.F. Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs. *PLoS One* **2014**, *9*(5): e98186. https://doi.org/10.1371/journal.pone.0098186

24. Zischewski, J., Fischer, R., & Bortesi, L. Detection of on-target and off-target mutations generated by CRISPR/Cas9 and other sequence-specific nucleases. *Biotechnology advances*, **2017**, *35*, 95-104. doi: 10.1016/j.biotechadv.2016.12.003

25. Arora L., Narula A. Gene Editing and Crop Improvement Using CRISPR-Cas9 System. *Front Plant Sci*. **2017**; *8*:1932. Published 2017 Nov 8. doi:10.3389/fpls.2017.01932

26. Liang, G., Zhang, H., Lou, D. *et al.* Selection of highly efficient sgRNAs for CRISPR/Cas9-based plant genome editing. *Sci Rep* **2016**, *6*, 21451. https://doi.org/10.1038/srep21451

27. Xing H.L., Dong L., Wang Z.P., Zhang H.Y., Han C.Y., Liu B., Wang X.C., Chen Q.J. A CRISPR/Cas9 toolkit for multiplex genome editing in plants. *BMC Plant Biol*. **2014.** Nov *29*; 14(1):327.

28. Fauser F., Schiml S., Puchta H. Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in Arabidopsis thaliana. *Plant J*. **2014,** Jul; *79*(2):348-59. doi: 10.1111/tpj.12554. Epub 2014 Jun 17.

29. Lowder L.G., Paul J.W., Baltes N.J., Voytas D.F., Zhang Y., Zhang D., Tang X., Zheng X., Hsieh T.F., Qi Y. A CRISPR/Cas9 toolbox for multiplexed plant genome editing and transcriptional regulation. *Plant Physiol.* **2015**, Aug *21. pii: pp.00636.2015.*

30. Zong Y., Wang Y., Li C., Zhang R., Chen K., Ran Y., Qiu J.L., Wang D., Gao C. Precise base editing in rice, wheat and maize with a Cas9-cytidine deaminase fusion. *Nat Biotechnol*. **2017**, May; *35*(5):438-440. doi: 10.1038/nbt.3811. Epub 2017 Feb 27.

31. Weiss T., Wang Ch., Kang X., Zhao H., Gamo M.E., Starker C.G., Crisp P.A., Zhou P., Springer N.M., Voytas D.F., Zhang F. Optimization of multiplexed CRISPR/Cas9 system for highly efficient genome editing in *Setaria viridis bioRxiv* **2020**.04.11.037572; doi: https://doi.org/10.1101/2020.04.11.037572

32. Wang T., Wei J.J., Sabatini D.M., Lander E.S. Genetic screens in human cells using the CRISPR-Cas9 system. *Science*, **2014**, *343*, pp. 80-84.

33. Moreno-Mateos, M. A., Vejnar, C. E., Beaudoin, J. D., Fernandez, J. P., Mis, E. K., Khokha, M. K., et al. CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting *in vivo*. *Nat. Methods*, **2015**, *12*, 982–988. doi: 10.1038/nmeth.3543

34. Graf R., Li X., Chu V.T., Rajewsky K. sgRNA Sequence Motifs Blocking Efficient CRISPR/Cas9-Mediated Gene Editing. *Cell Rep*. **2019**;*26*(5):1098-1103.e3. doi:10.1016/j.celrep.2019.01.024

35. Kleinstiver B.P., Pattanayak V., Prew M.S., Tsai S.Q., Nguyen N.T., Zheng Z., Joung J.K. High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature*, **2016**, Jan *28*; 529(7587):490-5. doi: 10.1038/nature16526.

36. Kocak D.D., Josephs E.A., Bhandarkar V., Adkar S.S., Kwon J.B., Gersbach C.A. Increasing the specificity of CRISPR systems with engineered RNA secondary structures. *Nat Biotechnol*. **2019**; *37*(6):657-666. doi:10.1038/s41587-019-0095-1

37. Zheng, T., Hou, Y., Zhang, P. *et al.* Profiling single-guide RNA specificity reveals a mismatch sensitive core sequence. *Sci Rep,* **2017**, 7, 40638. https://doi.org/10.1038/srep40638

38. Cermak, T., Baltes, N.J., Cegan, R., Zhang, Y. and Voytas, D.F. High frequency, precise modification of the tomato genome. *Genome Biol*. **2015**, *16*, 232.

39. Liang, Z., Zhang, K., Chen, K.L. and Gao, C.X. Targeted mutagenesis in Zea mays using TALENs and the CRISPR/Cas system. *J. Genet. Genomics*, **2014,** *41*, 63–68.

40. Woo, J.W., Kim, J., Kwon, S.I., Corvalan, C., Cho, S.W., Kim, H., Kim, S.G. et al. DNA-free genome editing in plants with preassembled CRISPRCas9 ribonucleoproteins. *Nat. Biotechnol*. **2015,** *33*, 1162–1164.

41. Murovec J., Guček K., Bohanec B., Avbelj M., Jerala R. DNA-Free Genome Editing of Brassica oleracea and B. rapa Protoplasts Using CRISPR-Cas9 Ribonucleoprotein Complexes. *Frontiers in Plant Science*, **2018**, *9*; 159 DOI=10.3389/fpls.2018.01594

42. Budhagatapalli N., Schedel S., Gurushidze M., Pencs S., Hiekel S. Rutten T. Kusch S., Morbitzer R., Lahaye, T., Panstruga R., Kumlehn J., Hensel G. A simple test for the cleavage activity of customized endonucleases in plants. *Plant Methods*. **2016**, 12. 10.1186/s13007-016-0118-6.

43. Gerasimova, S.V., Hertig, C., Korotkova, A.M. et al. Conversion of hulled into naked barley by Cas endonuclease-mediated knockout of the NUD gene. *BMC Plant Biol*. **2020**, *20*, 255. https://doi.org/10.1186/s12870-020-02454-9

44. Mazur, D. J., Perrino, F. W. Identification and expression of the TREX1 and TREX2 cDNA sequences encoding mammalian 3-prime-to-5-prime exonucleases. *J. Biol. Chem*. **1999**, *274*: 19655-19660, PubMed: 10391904.

45. Chari R., Mali P., Moosburner M., Church G.M. Unraveling CRISPR-Cas9 genome engineering parameters via a library-on-library approach. *Nature Methods*, **2015**, *12*(9):823-6. PMID: 26167643

46. Sanjana N.E., Shalem O., Zhang F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat Methods*. **2014** Aug; *11*(8):783-784. doi: 10.1038/nmeth.3047.

47. Sievers F., Wilm A., Dineen D., Gibson T.J., Karplus K., Li W., Lopez R., McWilliam H., Remmert M., Söding J., Thompson J.D. and Higgins D.G. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **2011.** *7*:539; DOI: 10.1038/msb.2011.75, PMID: 21988835

48. Brown N.P., Leroy C. and Sander C. MView: a web-compatible database search or multiple alignment viewer. *Bioinformatics* **1998**. *14*(4):380-1. DOI: 10.1093/bioinformatics/14.4.380, PubMed: 9632837

49. Geneious Prime (v2020.2.4) software. Available online: https://www.geneious.com/ (accessed on 03.04 2021).

50. Statistica (data analysis software system), version 13. TIBCO Software Inc. 2017 http://statistica.io. (accessed on 03.04 2021).