DNA methylation can alter CRISPR/Cas9 editing frequency and DNA repair outcome in a target-specific manner

Adéla Pribylová1,2, Lukáš Fischer2, Douglas E. Pyott3, Andrew Bassett4 and Attila Molnar1

1Institute of Molecular Plant Sciences, The University of Edinburgh, Edinburgh, EH9 3BF, UK; 2Faculty of Science, Charles University, Prague 128 44, Czech Republic; 3The Wellcome Trust Center for Cell Biology, Institute of Cell Biology, The University of Edinburgh, Edinburgh, EH9 3BF, UK; 4Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, CB10 1SA, UK

Summary

- The impact of epigenetic modifications on the efficacy of CRISPR/Cas9-mediated double-stranded DNA breaks and subsequent DNA repair is poorly understood, especially in plants.
- In this study, we investigated the effect of the level of cytosine methylation on the outcome of CRISPR/Cas9-induced mutations at multiple Cas9 target sites in Nicotiana benthamiana leaf cells using next-generation sequencing.
- We found that high levels of promoter methylation, but not gene-body methylation, decreased the frequency of Cas9-mediated mutations. DNA methylation also influenced the ratio of insertions and deletions and potentially the type of Cas9 cleavage in a target-specific manner. In addition, we detected an over-representation of deletion events governed by a single 5'-terminal nucleotide at Cas9-induced DNA breaks.
- Our findings suggest that DNA methylation can indirectly impair Cas9 activity and subsequent DNA repair, probably through changes in the local chromatin structure. In addition to the well described Cas9-induced blunt-end double-stranded DNA breaks, we provide evidence for Cas9-mediated staggered DNA cuts in plant cells. Both types of cut may direct microhomology-mediated DNA repair by a novel, as yet undescribed, mechanism.

Introduction

CRISPR/Cas9 has been extensively used in gene editing in eukaryotes, and many of the mechanistic details of Cas9-induced double-stranded breaks (DSBs) have been well described. Cas9-mediated cleavage mostly occurs as a blunt DSB three nucleotide upstream of the protospacer adjacent motif (PAM) (Jinek et al., 2012; Josephs et al., 2015). However, animal and yeast studies indicate that the Cas9 can also induce staggered DSBs with single 5' nucleotide overhangs (Li et al., 2015; Lemos et al., 2018). Double-stranded breaks can be repaired by several endogenous DNA repair mechanisms. In eukaryotes, the predominant mechanism in differentiated cells is classical nonhomologous end joining (c-NHEJ), in which blunt DNA ends are directly re-ligated. Alternatively, short overhangs, resulting for example from CRISPR/Cas9 staggered cuts, can be filled in and then ligated forming short insertions (IN) (Beying et al., 2021). Double-stranded breaks can also be repaired through homology-directed repair. When the repair relies on extensive nucleotide sequence similarity/homology (such as 20–100 bp) at 3' DNA ends it comprises two main pathways: single-strand annealing and homologous recombination (Roth et al., 2012; Bothmer et al., 2017; Beying et al., 2021). The remaining dominant pathway is microhomology-mediated end joining (MMEJ), also known as alternative end joining (Alt-EJ), which relies on annealing of short complementary regions referred to as microhomologies present on both 3' DNA ends close to the break, which can result in short deletions (DEL) (Chiruvella et al., 2013; Beying et al., 2021; Vu et al., 2021).

Successful CRISPR/Cas9-induced mutagenesis is conditioned by the selection of an appropriate target sequence. Several bioinformatics tools have been developed to aid the selection of optimal sequences to be targeted by Cas9. Many of these algorithms provide predicted mutagenic efficiencies (on-target scores) and, if a whole-genome sequence is available, some programs will also provide predictions of off-target sequences (Liu et al., 2017; Cui et al., 2018). While these algorithms facilitate the design of gRNAs, the predicted on-target efficiencies can often be inaccurate (Naim et al., 2020), indicating that our understanding of the principles underlying Cas9 binding and cleavage of target DNA and subsequent repair of DSB in vivo is far from complete. A possible reason for the discrepancies between the in silico predictions and experimental results could be that most of the gRNA design tools do not consider the epigenetic status of the target locus. It is likely that the chromatin structure of a locus (in addition to the underlying base sequence) will be an important determinant for the efficacy of Cas9 binding, cleavage and subsequent repair. Indeed, there have been reports that the frequency of Cas9-induced mutagenesis is lower for heterochromatin compared with euchromatin (Chen et al., 2016; Daer et al., 2017, 2018). However, other reports have suggested that Cas9-induced
mutation frequencies are comparable for heterochromatin and euchromatin (Yu et al., 2013; Feng et al., 2016; Kallimasioti-Pazi et al., 2018). This indicates that epigenetic factors influencing the CRISPR/Cas outcomes are more complex than just considering the broad heterochromatin/euchromatin dichotomy (Schep et al., 2021). For instance, it has been shown that the epigenetic availability of the PAM sequence is crucial to the action of Cas9, such that PAMS associated with nucleosomes cannot be located by Cas9 (Hinz et al., 2015; Horlbeck et al., 2016; Yarrington et al., 2018). However another epigenetic effect that could influence the ability of Cas9 to find and cleave DNA is the methylation status of the target locus. While DNA methylation at the PAM was found to have no effect on Cas9 activity in human cell lines (Hsu et al., 2013; Fujita et al., 2016), this phenomenon is yet to be tested in other organisms.

Here we present a detailed study on *Nicotiana benthamiana* in which we use Tobacco Rattle Virus (TRV) as a molecular switch to change the chromatin state of a reporter gene from an actively transcribed to a transcriptionally silenced state. Unlike previously used systems, our unique approach enabled us to interrogate different chromatin states of the same locus with the same set of CRISPR/Cas9 genome editing reagents and systematically describe the effect of chromatin state on the frequency and type of mutations induced at various Cas9 targets in a huge set of independently edited cells.

### Materials and Methods

**Generation of CRISPR/Cas9 gene-editing constructs**

Cas9 target sites were identified by GENEIOUS v.11.1.5 (http://www.geneious.com) (Kearse et al., 2012) and potential off-targets were analysed by Sol Genomics Network BLAST (https://www.solgenomics.net/tools/blast) (Fernandez-Pozo et al., 2015) and CRISPR RGEN CAS-OFFINDER (http://www.rgenome.net/cas-offinder) (Bae et al., 2014). The corresponding oligonucleotides with *BbsI* compatible overhangs (Supporting Information Table S1) were annealed into a dsDNA duplex, phosphorylated with T4 Polynucleotide Kinase (New England Biolabs, Hitchin, UK) and subsequently cloned into *BbsI*-digested entry vector *pEn_Chimera* (Fauser et al., 2014). The resulting gRNA expression cassettes (*pEn_Chimera_x*, *x* = *ut*, *ub*, *pt*, *pb*, *mt*, *mb*, *dt*, *db*, *urt*, *mht*, *db*; Table S2) were verified by Sanger sequencing. The BASTA resistance cassette (*CaMV35S::bar::nos*) was deleted from the original destination binary vector *pDe_CAS9* (Fauser et al., 2014) by HindIII restriction enzyme digestion and religation to avoid interference with the 35S sequence. The gRNA expression cassettes were then recombined into the modified destination vector (*pDe_CAS9_no_basta*) by Gateway cloning (LR reaction) according to the manufacturer’s instruction (Life Technologies, Carlsbad, CA, USA), which yielded the *pDE_CAS9_no_basta_x* binary vectors (Table S2) for plant transformation.

**Plant material**

Wild-type *N. benthamiana* plants, *N. benthamiana* 16c line expressing the green fluorescent protein (GFP) reporter gene under the constitutive *CaMV 35S* promoter (Ruiz et al., 1998) and its transcriptionally silenced progeny (*S1*: the virus-free progeny of TRV-35S-infected 16c; Fig. 1) (Fei et al., 2021) were grown in SANYO/Panasonic growth chambers with 16 h : 8 h, light : dark photoperiods at 21°C under 150 µmol m⁻² s⁻¹ light.

**Viruses and virus sap**

Wild-type TRV, recombinant TRV-35S and TRV-GFP, and virus sap collection from *N. benthamiana* plants have been described in Fei et al. (2021).

**Transient expression of gene-editing reagents in planta**

*Agrobacterium tumefaciens* AGL1 strain was transformed with *pDe_Cas9_no_basta_x* vectors (Table S2) by electroporation and selected on solidified LB plates using 100 µg ml⁻¹ spectinomycin and 50 µg ml⁻¹ rifampicin. For each construct, 200 µl of overnight starter culture was added to 10 ml LB supplemented with the above antibiotics and incubated for 24 h at 28°C and 11 g. The bacterial culture was then centrifuged at 3000 g for 10 min, and the pellet was resuspended in infiltration medium (IM, 10 mM MES (pH 5.6), 300 µM acetosyringone, 10 mM MgCl₂) and incubated for 3–4 h at room temperature. The optical density was adjusted to 1.5 and 0.15 OD₆₀₀ by adding a fresh IM buffer. An *Agrobacterium* suspension was infiltrated with a 1-ml syringe into the abaxial side
of three leaves (4th–6th youngest leaves) of 5-wk-old N. benthamiana. For agro-infiltrating virus-infected plants, the first and the second leaf of 3-wk-old N. benthamiana were rub-inoculated with the corresponding virus sap as previously described (Fei et al., 2021) and the plants were kept under standard conditions (described in ‘Plant material’ in the Materials and Methods section) for another 2 wk before agro-infiltration. All experiments were independently repeated three to five times.
Imaging GFP fluorescence

GFP expression was regularly monitored under UV light using a handheld mercury UV lamp (UVP Blak-Ray™ B-100AP High-Intensity UV Inspection Lamp, Upland, CA, USA). Photographs were taken using a Canon G16 camera as described previously (Fei et al., 2021).

Sample collection

A leaf disc (8 mm in diameter) was cut out from each infiltrated leaf (4th, 5th and 6th) and combined into a single biological sample 5 d after post infiltration (dpi). Three samples were collected from each plant, frozen in liquid nitrogen and stored at −80°C until further use.

Mutation analysis

DNA was isolated using the DNeasy Plant Mini Kit (Qiagen). A T7 endonuclease assay was carried out as previously described (Pyott et al., 2016). After Turbo DNase treatment (Ambion), cDNAs were made from 0.5 μg RNA using random hexamers primers and SuperScript III (Life Technologies). Quantitative RT-PCR was performed using the SYBR Green I Master Mix on a LightCycler480 (Roche) using gene-specific primers (Table S1). Data were converted with LC480 software (v.2014.1), and Ct values were calculated using LinRegPCR (v.2012.0) (Ramakers et al., 2003), both software are available at https://www.medischebiologie.nl/files/. The analysis was carried out in MS EXCEL 365 (Microsoft, Redmond, WA, USA).

RNA analysis

RNA was isolated using the RNeasy Plant Mini Kit (Qiagen). After Turbo DNase treatment (Ambion), cDNAs were made from 0.5 μg RNA using random hexamers primers and SuperScript III (Life Technologies). Quantitative RT-PCR was performed using the SYBR Green I Master Mix on a LightCycler480 (Roche) using gene-specific primers (Table S1). Data were converted with LC480 software (v.2014.1), and Ct values were calculated using LinRegPCR (v.2012.0) (Ramakers et al., 2003), both software are available at https://www.medischebiologie.nl/files/. The analysis was carried out in MS EXCEL 365 (Microsoft). Three technical replicates were performed for each biological replicate.

DNA methylation analysis

Genomic DNA was converted using the EZ DNA Methylation-Gold Kit (Zymo Research, Irvine, CA, USA). Target loci were amplified using bisulfite PCR primers (Table S1) and EpiMark® Hot Start Taq DNA Polymerase (NEB). The resulting PCR products were ligated into the pGEM®-T Vector (Promega), which were then used to transform Escherichia coli (DH5α) cells. In total, 8–10 clones from each sample were sequenced using BigDye 3.1, as described previously. DNA methylation patterns were analysed using GENEIOUS v.11.1.5 (http://www.geneious.com) (Kearse et al., 2012).

Statistical analysis

Statistical analyses were performed using the NCSS 9 and SPSS 28 statistical software. Significance level α was always set to 0.05. All variables were checked for outliers, and assessed for normality using the Shapiro–Wilk test and Levene’s test for equal variance. A one-tailed or two-tailed two-sample independent t-test was used in which the two independent groups were compared, based on expectations from the groups being compared. To determine whether the independent variables had an interaction, two-way analysis of variance (ANOVA) tests were performed, followed by Tukey’s post hoc test if possible.

Results

Virus-induced gene silencing can be used as an ‘epigenetic switch’ to alter the cytosine methylation status of a target locus

To test how chromatin state affects the efficacy of CRISPR/Cas9 mutagenesis in planta, we developed a system that allowed us to change the chromatin state of a single locus and subsequently analyse CRISPR/Cas9-induced mutations at that locus. In this system, we made use of a recombinant RNA virus to induce DNA methylation by triggering the plant’s RNA-dependent DNA methylation (RdDM) pathway at the target locus. The RdDM pathway can be triggered by some viruses when short interfering RNAs (siRNAs) are produced by the plant’s antiviral RNA silencing response. Certain classes of these virus-derived siRNAs can direct plant-encoded DNA methyl transferases to methylate cytosine residues of DNA that bear sequence complementarity to the siRNA. Plant viruses can be turned into virus-induced gene silencing (VIGS) vectors if sequences present in the plant’s genome are cloned into the viral genome. Infection with such viruses will result in the production of siRNAs that are complementary to the region of the plant’s genome that was inserted into the virus. If a promoter sequence is inserted into the VIGS vector, the siRNA-mediated RdDM can lead to cytosine methylation of that promoter and consequently result in transcriptional gene silencing (TGS) (Jones et al., 2001; Pribylova et al., 2019).

For our experimental system, we utilised a previously described and well characterised transgenic N. benthamiana line (named 16c) that contained a single-copy genomic insertion of the GFP reporter gene under the constitutive 35S promoter (Fig. 1a) (Ruiz et al., 1998). In the 16c line, the 35S::GFP cassette is highly expressed, implying that the transgenic cassette is integrated into an epigenetically active region of the plant’s genome. In a
previous study (Fei et al., 2021), we adapted a TRV VIGS vector (Jones et al., 2001) to induce RdDM of the 35S promoter by targeting a short segment (120 bp) upstream of the transcription start site and therefore transcriptionally silence the expression of GFP in 16c plants. We named this vector TRV-35S (Fig. 1b). To control for possible effects of the virus infection itself (unrelated to the VIGS-induced RdDM of the target) we used a TRV vector without the 120-bp section of the 35S promoter inserted and named this vector TRV-WT (Fig. 1b). Furthermore, we collected seeds from the TRV-35S-infected plants to allow us to evaluate the effects of cytosine methylation at the target locus in plants containing no viral RNA. The basis of this is that TRV transmission to progeny seedlings was extremely low (<1%) (Jones et al., 2001) but cytosine methylation that is in a symmetrical sequence context (i.e. CG or CHG, where H represents A, C or T) can be maintained through both mitotic and meiotic cell divisions and exhibits transgenerational inheritance (Jones et al., 2001). We use the term S1 to refer to these plants grown from the seeds collected from TRV-35S-infected plants as they are the first-generation progeny of the transcriptionally silenced plants (Fig. 1c). Therefore, our system allows for the comparison of two different epigenotypes (high/low cytosine methylation) at the 35S promoter in the context of virus-infected and nonvirus-infected tissue (Fig. 1d).

First, we tested the efficacy of our TRV-based ‘epigenetic switch’. When imaged under UV light, 16c plants that had been rub-inoculated with TRV-35S glowed red due to the autofluorescence of chlorophyll and the green glow of GFP was completely absent. By contrast, for both the noninfected and TRV-WT-infected 16c plants a green GFP signal was discernable and was particularly visible in the upper leaves and petioles. A weak GFP signal could be seen in the S1 plants, indicative of only a partial silencing of GFP (Figs 1e, S1). This silencing of the GFP expression was confirmed in three randomly selected independent biological replicates by RT-qPCR. TRV-TT-infected plants had a significantly higher level of GFP mRNA compared with TRV-35S-infected plants (P = 0.006, two-tailed t-test), which had negligible levels of GFP mRNA. The noninfected 16c plants had more than twice the level of GFP mRNA than S1 plants (P = 0.011, two-tailed t-test), but there was still a weak expression of GFP mRNA in the S1 plants (Fig. 1f). Moreover, we did not observe any difference in the abundance of TRV RNA between the TRV-35S and TRV-WT infections (P = 0.441, two-tailed t-test).

To confirm that the basis of the observed GFP silencing was a result of VIGS-induced RdDM, we performed bisulphite sequencing. DNA samples taken from TRV-WT-infected plants showed almost no methylation in the 120-bp target region, and weak methylation was observed in the upstream promoter region, mainly in the CG context. By contrast, DNA samples from TRV-35S-infected plants showed strong methylation in the target region with a weak spreading of methylation towards 3’ and 5’ ends of the 35S promoter, independent of the C sequence context. In the upstream part of the promoter, we observed increased methylation in the symmetrical CG and CHG context, mainly in the region where even TRV-WT-infected plants showed some CG methylation. As expected, noninfected plants showed a similar methylation pattern as TRV-WT-infected plants: no methylation in the target region and limited methylation in the upstream region of the promoter. S1 plants inherited an increased level of mostly symmetric CG and CHG methylation within the target region of the 35S promoter, but only at some C positions (Fig. 1g,h). Taken together, our data show that infection with TRV-35S can switch the 35S::GFP transgene of 16c plants from an epigenetically active to an epigenetically inactive (silenced) state due to virus-induced methylation of the 35S promoter and that this silenced epigenotype is partially inherited to the virus-free progeny.

The efficacy of CRISPR/Cas9 editing is inversely correlated with the level of cytosine methylation for targets within a promoter region

Having demonstrated that we could alter the cytosine methylation status of a promoter by VIGS, we sought to use this as a tool to determine how DNA methylation at a given locus might influence the frequency of CRISPR/Cas9-induced mutations (Fig. 2a). To this end, we designed CRISPR gRNAs that would direct Cas9 to the region of the 35S promoter that we had targeted for RdDM using TRV-35S (Fig. 2b). Additionally, we designed gRNAs to direct Cas9 to a region of the 35S promoter that was upstream of the TRV-35S targeted region but still showed an increase in VIGS-induced cytosine methylation (Figs 1b, 2b). Within the TRV-35S targeted region of the promoter, we designed gRNAs to direct Cas9 to three positions which we referred to as proximal, middle and distal (based on their position relative to the start of the promoter). For each of the four regions of the 35S promoter (namely: upstream, proximal, middle and distal) we designed pairs of gRNAs that would anneal to each of the two complementary strands of the target DNA, which for convenience we referred to as the top and bottom strands. As far as possible we designed these pairs of gRNAs so that they would overlap, although we could not obtain a 100% overlap of the gRNAs directed to the top and bottom DNA strands due to the constraints of PAM availability. For ease of reference, we named each of the gRNAs using acronyms to describe both their position in the 35S promoter (upstream/proximal/middle/distal) and their cognate DNA strand (top/bottom). Therefore, we named the eight gRNAs ut, ub, pt, pb, mt, mb, dt, db (Fig. 2b).

To test the in planta activity of these gRNAs, we cloned them individually into the pDe-Cas9 binary vector under the control of the Arabidopsis thaliana U6 promoter. This binary vector also contained a Cas9 expression cassette driven by the constitutive Ubiquitin4–2 promoter from Petroselinum crispum (PcUbi4-2) (Fauser et al., 2014). Syringe infiltration of Arabidopsis thaliana U6 promoter. This binary vector also contained a Cas9 expression cassette driven by the constitutive Ubiquitin4–2 promoter from Petroselinum crispum (PcUbi4-2) (Fauser et al., 2014). Syringe infiltration of Arabidopsis thaliana U6 promoter.
extracted from the infiltrated leaves and a T7 assay was performed to semiquantitatively assess the mutagenic capacity of each of the gRNA-containing constructs. Encouragingly, the target ampli-
cons digested with T7 endonuclease produced the expected cleav-
age products for all of the gRNA constructs, indicating that all of
the gRNAs are capable of directing Cas9 to the respective target
loci to induce mutations using both the high and the low bac-
terium densities (Fig. S2A).

Next, we used our VIGS system to induce methylation at the
35S promoter to determine any effects of this induced

Fig. 2 The impact of epigenotype on CRISPR/Cas9-induced mutation frequency. (a) Summary of the experimental design. Here, 3-wk-old 16c plants were
infected with the Tobacco Rattle Virus (TRV) constructs. Noninfected plants were used as controls. At 2 wk later, the newly emerging leaves were infiltrated
with Agrobacterium tumefaciens carrying the corresponding CRISPR/Cas9 constructs. Samples were collected 5 d post infiltration (dpi) and were subse-
quently analysed for the efficacy of CRISPR/Cas9-mediated gene editing. (b) Schematic diagram of the CRISPR/Cas9 target sites at the P35S::GFP::NosT
locus. gRNAs were named according to the position of the protospacer adjacent motif (PAM) sequence (circled). t, top; b, bottom; u, upstream; p, prox-
imal; m, middle; d, distal. The region of P35S targeted for induced cytosine methylation is indicated with a black line. (c) CRISPR/Cas9-induced mutation
frequencies in plants infected with TRV-WT or TRV-35S (left panel), and in noninfected 16c and TRV-35S-noninfected progeny (S1) plants (right panel).
Individual points (o, x, □) represent next-generation sequencing (NGS) data (OD600 = 0.15) from three independent biological replicates. Lower range,
box, upper range and a horizontal line represents the first quartile, second and third quartile, fourth quartile and median, respective. Two-way ANOVA was
used to examine the effect of the target position and epigenotype on mutation frequency and showed a statistically significant interaction between the tar-
get position and the epigenotypes (F(7, 32) = 3.017, P = 0.015) in infected plants. By contrast, the interaction was not significant for noninfected plants (F
(7, 32) = 0.960, P = 0.477). For both groups, infected and noninfected, the simple and main effect of the target site was significant (P < 0.001). To examine
the effect of the epigenotype in the individual target site, a one-tailed t-test was used, as individual target sites had only two independent variables. * and
ns indicate significant difference for P < 0.05 and not significant, respectively.
methylation on the mutagenesis frequency of each of the gRNAs. Here, 16c plants were infected with either TRV-35S or TRV-WT and were incubated for 2 wk to allow the infection to become systemic and initiate RdDM. Newly formed leaves of these systemically infected plants were then infiltrated with the different gRNA-containing CRISPR/Cas9 constructs (Fig. 2a).

In parallel, noninfected plants and S1 plants were infiltrated with the same CRISPR/Cas9 constructs to analyse the active/silenced epigenotypes in the absence of TRV infection. DNA was then extracted from the infiltrated leaves to determine the rate of mutagenesis under each of the different conditions. To obtain more detailed and quantitative data than would be possible using a T7 assay, we analysed the samples by Sanger sequencing. First, we plotted the mutation frequencies for all eight gRNAs infiltrated at either OD$_{600}$ = 1.5 or OD$_{600}$ = 0.15 for each of the four different plant variants, from this point forwards referred to as epigenotypes (TRV-WT-infected, TRV-35S-infected, noninfected and S1). In general, regardless of the epigenotype, higher bacterial densities resulted in higher mutation frequencies ($P<0.001$, two-tailed t-test). However, differences in mutation rate between the epigenotypes could be seen mainly with the low bacterial density and we assumed that they were partially masked when the plants were infiltrated with the higher bacterial density (Fig. S2B).

Interestingly, TRV WT-infected plants exhibited equally high mutation frequencies for both bacterial densities ($P=0.575$, two-tailed t-test) and the mutation frequencies for the TRV WT-infected plants were higher than in any of the other plant variants, regardless of the bacterial density used for delivering the CRISPR/Cas9 components. When comparing noninfected and TRV WT-infected 16c plants, which were infiltrated at OD$_{600}$ = 0.15, there was a significant difference in mutation frequency ($P=0.001$, two-tailed t-test). This suggests that the TRV infection somehow predisposes the plant to a lower fidelity of DNA repair, which we speculate could arise from the plant’s stress response to the virus infection or even interference of TRV RNA with the DNA repair pathways. This result highlighted the importance of our experimental controls to accurately interpret our data. Due to the observed effect of TRV infection on mutation frequency, it is only meaningful to make pairwise comparisons between the TRV WT-infected and TRV 35S-infected plants or between the noninfected and S1 plants when assessing the effect of chromatin state.

Based on the global analyses outlined above, we chose to analyse the mutation frequencies for each gRNA individually for the four different plant variants using the lower bacterial density (OD$_{600}$ = 0.15). To accurately discern the effects of promoter methylation on the efficacy of CRISPR/Cas9-induced gene editing for each locus, we analysed the mutation frequency of the corresponding amplicons using TIDE-coupled Sanger sequencing (Fig. S2C) and subsequently by Illumina NGS (Fig. 2c) using at least three independent biological replicates. In agreement with previous findings (Sentmanat et al., 2018), there was a strong correlation between the Sanger and NGS datasets for all epigenotypes and targets, although mutation frequencies determined by NGS were, on average, 1.14 times higher than the estimates determined by Sanger sequencing (Fig. S3). Comparing the TRV WT-infected and TRV 35S-infected plants, the mutation frequency was clearly higher for the promoter with active chromatin compared with the silenced, methylated promoter and this trend was consistent across the majority of targets, with pb, mt, dt and db targets showing statistically significant differences ($P<0.05$, one-tailed t-test; Figs 2c, S2C, left panels). Interestingly, the difference in mutation frequency became significant for ut when the number of biological repeats was increased from three (NGS) to five (Sanger; Fig. S2C).

When we compared the mutation frequencies for gRNAs targeted to either the top or bottom strand for a given locus, we found no global correlation between the orientation (top or bottom) of the target and the mutation frequency (data not shown). As the impact of cytosine methylation on the mutation frequency in the promoter region was significant, when comparing the nonmethylated TRV WT with heavily methylated TRV 35S plants ($P=0.009$, two-tailed t-test; Figs S5, S6), we were interested to also assess the impact of cytosine methylation of transcribed DNA on the efficacy of CRISPR/Cas9-mediated gene editing. To this end, we induced RdDM at the coding region of GFP with a recombinant TRV VIGS vector (TRV-GFP) and subsequently analysed the frequency of Cas9-induced mutations by amplicon sequencing (Notes S1; Fig. S4). We observed no statistically significant difference in mutation frequency between the TRV-GFP and TRV WT (control) samples for individual target sites (Fig. S4) and no statistically significant effect of the cytosine methylation on the mutation frequency ($P=0.640$, two-tailed t-test; Figs S5, S6), implying that the TRV-GFP-induced cytosine methylation has little or no effect on the frequency of CRISPR/Cas9-induced mutations for three GFP targets including dg (Figs 2b top panel, S4G,H, S5, S6). Taken together, our data showed that the frequency of CRISPR/ Cas9-induced mutations at the promoter DNA was generally inversely correlated with its level of cytosine methylation, but this phenomenon did not apply to coding regions of DNA, suggesting that the effect of DNA methylation on the efficacy of gene editing was rather indirect (see the second paragraph of Discussion section).

The mutational outcome of CRISPR/Cas9 editing is affected by DNA methylation

Next, we were interested to see if DNA cytosine methylation could affect not only the frequency of CRISPR/Cas9-induced mutations but also the mutational outcome. To this end, we used the output of CRISPResso2 software (Clement et al., 2019, p. 20) to analyse our NGS data to resolve the specific features of insertion (IN), deletion (DEL) and substitution (SUB) events within our dataset. SUB events occurred along the whole analysed DNA regions without any specific accumulations near the predicted Cas9 cut sites, including negative controls that had been infiltrated with a CRISPR/Cas9 cassette with no gRNA, and accounted for c. 2% of reads. All these data indicated that SUB events can be considered as sequencing or PCR errors, therefore they were omitted from subsequent analyses.

Detailed analysis of total reads of IN and DEL events in individual target sites showed that each target site had a specific pattern of IN:DEL ratios. At some target sites IN dominated over
DEL and vice versa (Figs 3, S7A). Interestingly, for three sites, the IN:DEL ratio was obviously shifted in connection with RdDM, either in favor of insertions (mt and db site) or deletions (dtg site; Fig. S7B). Moreover, we found that RdDM had a significant effect on the frequency of insertions and deletions at several individual target sites (pb and dt sites: Fig. 3a, left panel; and pb, mt, dt, db and dtg sites: Fig. 3b, left panel).

When we analysed IN mutations, c. 98% of them represented single-nucleotide insertions (Fig. S8) which were located just next to the predicted Cas9 cut site. We observed a strong bias within the inserted nucleotides towards the 4th nucleotide upstream of the PAM at most target sites, which is consistent with a previous report from a nonplant model that Cas9 can produce either blunt ends or staggered ends with single-nucleotide (the 4th nucleotide from PAM) 5’ overhangs (Figs 4, S9) (Li et al., 2015). The only exceptions were pb and mb, for which nontemplated T and A nucleotides were the most prevalent single-nucleotide insertions over the templated Gs (Figs 4, S9). Interestingly, the identity of the inserted nucleotide also appeared to be affected by the epigenotype to varying degrees for different target loci, with the biggest effect seen for the site targeted by the dt gRNA (Figs 4, S9).

- **Fig. 3** The impact of epigenotype on CRISPR/Cas9-induced mutations. (a) Frequencies of insertion and (b) deletion events at each target site and epigenotype. Individual points (○, □, △) represent next-generation sequencing data (OD<sub>200</sub> = 0.15) from three independent biological replicates. Lower range, box, upper range and a horizontal line represents the first quartile, second and third quartile, fourth quartile and median, respective. Two-way ANOVA was used to examine the effect of the target position and epigenotype on the frequency of insertions and deletions and showed statistically significant interaction between the target position and the epigenotypes (F(8, 36) = 5.097, P < 0.001 and F(8, 36) = 7.347, P < 0.001, respective) in infected plants. The interaction was not significant for noninfected plants nor for insertions or deletions (F(8, 36) = 1.094, P = 0.390 or F(8, 36) = 1.087, P = 0.394). For both groups, infected and noninfected, the simple and main effect of the target site was significant (P < 0.001). To examine the effect of the epigenotype in the individual target site a one-tailed t-test was used, as individual target sites had only two independent variables. * and ns indicate significant difference for P < 0.05 and not significant, respectively. Terms TRV-WT, TRV-35S, noninfected and S1 are explained in Fig. 1; colour-coded CRISPR/Cas9 target sites ut, ub, pt, pb, mt, mb, dt, db are explained in Fig. 2 and dtg in Supporting Information Fig. S4.

- **Fig. 4** Detailed analysis of single-nucleotide insertion events detected next to the Cas9 cut site. 1-bp insertion events were split into four categories, one each for every nucleotide (A, C, G, T). The top bar shows which nucleotide was in the original sequence on the 4th position upstream from the protospacer adjacent motif (PAM) sequence. Each data point represents the mean of next-generation sequencing data (OD<sub>200</sub> = 0.15) from three independent biological replicates. Terms TRV-WT, TRV-35S, noninfected and S1 are explained in Fig. 1; colour-coded CRISPR/Cas9 target sites ut, ub, pt, pb, mt, mb, dt, db are explained in Fig. 2 and dtg in Supporting Information Fig. S4.
We then evaluated DEL events based on the length and position of deletions relative to the predicted CRISPR/Cas9 cut site and the PAM. We found that only a small fraction of reads (3%) contained mutations outside a 20-nt window around the Cas9-induced DSB (Group 5; Fig. 5a). Close inspection of DEL events within this 20-nt window resulted in the identification of four groups. DEL events in group 1 spanned the cut site in both a PAM-proximal and PAM-distal direction. Group 2 contained DEL events that were exclusively in the PAM-distal region adjacent to the Cas9 cut site. Group 3 contained DEL events that could be mostly mapped on the PAM distal end where terminal nucleotides can be mapped to either side of the Cas9 cut site. Finally, group 4 contained DEL events that were exclusively in the PAM-proximal region adjacent to the Cas9 cut site (Fig. 5a,b). We found the distribution of mutations within each of these four groups to be variable for the different gRNAs with no noticeable effect of the epigenotype. Groups 1, 2, 3 and 4 represented 1%, 61%, 27% and 7% of the total reads, respectively, indicating that most deletions occurred in the PAM-distal region adjacent to the Cas9 cut site (Figs 5c, S10). Intriguingly, we found that in many cases, especially in group 2, the deletion was insulated from the predicted Cas9 cleavage site by at least one additional nucleotide, which was almost exclusively aligned to the 4th nucleotide upstream of the PAM. This observation is consistent with Cas9-induced staggered ends with single-nucleotide 5’ overhang (see Fig. 4), which mostly did not vary between epigenotypes (Figs 5d, S11).

DNA ends with single-stranded overhangs in nondividing cells are prone to be repaired by the MMEJ pathway if an identical sequence of 2–25 nucleotides is present at both 3’ DNA ends near the cut site (Allen et al., 2019; Beying et al., 2021). DNA repair by this MMEJ mechanism results in a deletion (in the simplest scenario that does not involve insertions) that spans one of the two matching microhomology regions and the sequence in between them. Consequently, the preserved copy of the microhomology region can be mapped to either side of the deletion, and this mutational signature can be used to infer likely cases of MMEJ-directed repair. In our dataset, we observed increased occurrence of such deletions only for ub and dt targets, in which trinucleotides present at the terminus of the PAM strand were repeated on the PAM-distal DNA end near the Cas9 cut site (Fig. 6a). This indicated that the MMEJ repair pathway did not play a significant role in our experimental system, which is consistent with its predominant activity in dividing cells. However, the surprisingly high occurrence of 1-nt microhomologies in our dataset (Fig. 6a) implied that a significant number of the deletions were formed by a single-nucleotide-dependent microhomology repair distinct from the classical MMEJ pathway. Moreover, when we analysed the 1-nt microhomologies in more detail for individual gRNA targets, we noted that they nearly always aligned to the nucleotide that was either 3rd or 4th nt upstream of the PAM (Figs 6b, S12).

Fig. 5 Detailed analysis of deletion events detected next to the Cas9 cut site. (a) Schematic of mutation categories: (1) deletions on both DNA ends, (2) only on the protospacer adjacent motif (PAM) distal end, (3) mostly on the PAM distal end where terminal nucleotides can be mapped to either side of Cas9 cut DNA, (4) only on proximal DNA end, and (5) not directly adjacent to the cut site. (b) Schematic of mutations at nucleotide resolution for categories 1–4 using representative sequences from the ut next-generation sequencing (NGS) data as an example. (c) Type and frequency of deletion mutations in CRISPR/Cas9 cut sites. (d) Detailed analysis of group 2 and its first adjacent position next to the cut site on the PAM distal DNA end. Position can contain plain deletion or a mutation with a remaining single nucleotide (A, C, G or T) adjacent to the predicted double-stranded break. The top bar represents which nucleotide was in the original sequence on the 4th position upstream from the PAM sequence. The top panel shows a detailed example of ‘ut’ gRNA target for a TRV-WT-infected sample. Each value represents the mean of NGS data (OD600 = 0.15) from three independent biological replicates. Terms TRV-WT, TRV-3SS, noninfected and S1 are explained in Fig. 1; colour-coded CRISPR/Cas9 target sites ut, ub, pt, pb, mt, mb, dt, db are explained in Fig. 2 and dtg in Supporting Information Fig. S1.
This again was consistent with our conclusion that Cas9 could cut either in front of the 3<sup>rd</sup> or at the 4<sup>th</sup> nt upstream of the PAM on the PAM-containing strand. When we considered all deletion mutations in our dataset, we observed an enrichment of deletion sizes corresponding to the distance between the predicted Cas9 blunt/staggered cleavage site and the position of a single-nucleotide microhomology at the PAM-distal end of DSB (Figs 6c, S13). Together with the fact that the majority of DELs occurred on PAM-distal DNA strands (Figs 5c, S10, group 2) our data suggested that the nonrandom mutations were produced by a previously unidentified 5′ DNA single-nucleotide microhomology-mediated DNA repair. This highlighted the importance of considering the potential for triggering single-nucleotide microhomology-mediated DNA repair when designing gRNAs.

**Discussion**

The primary aim of this study was to investigate the effects of an epigenetic modification, namely DNA cytosine methylation, on the efficacy of generating targeted mutations using CRISPR/Cas9. Previous studies have suggested that the frequency of CRISPR/Cas9-induced mutations may be affected by the epigenetic status of the target DNA, using genomewide analyses to look for associations between the frequency of mutations and the epigenetic status (euchromatin vs heterochromatin) of the target, for multiple endogenous loci (Chen et al., 2016; Daer et al., 2017, 2018). To our knowledge, no study has to date compared the effect of DNA methylation on CRISPR/Cas9-induced mutations by comparing the mutation frequency at the same locus containing either a high or low cytosine methylation content. In
this work, we developed an ‘epigenetic switch’ that enabled us to control the methylation status of the genomic loci, therefore allowing us to compare the efficacy of CRISPR/Cas9-induced mutations for a given locus containing either a high or low level of cytosine methylation.

Our results showed that the level of cytosine methylation can affect the efficacy of CRISPR/Cas9 mutagenesis in a target site-specific manner (Figs 2c, S2C). The RdDM pathway, which we used to trigger local chromatin modifications, is known to result in transcriptional silencing if a promoter sequence is targeted used to trigger local chromatin modifications, is known to result in transcriptional silencing if a promoter sequence is targeted by the infection, which might have caused a higher error rate in the cells’ DNA repair mechanisms. Indeed, a previous report has linked heat stress to higher rates of CRISPR/Cas9 mutagenesis (LeBlanc et al., 2018). Additionally, we observed that the density of Agrobacterium used for infiltrating plants with the CRISPR/Cas9 components had a significant effect on the rate of mutagenesis. The higher bacterial density of OD_{600} = 1.5 resulted in consistently higher rates of mutagenesis compared with the lower bacterial density of OD_{600} = 0.15 (Fig. S2B). From this, we concluded that the level of expression of CRISPR/Cas9 components or the number of cells successfully transformed within the infiltrated patch was likely to be a limiting factor for the mutagenesis frequency, over this range of OD_{600} values. Testing a wider range of Agrobacterium densities would be required to determine the optimum density of Agrobacterium suspension to achieve maximal mutagenesis frequencies. Importantly, our results highlight the need to carefully determine the most appropriate Agrobacterium density to use for studies aimed at understanding specific mechanisms of CRISPR/Cas9 editing. In our case, we chose to use a lower bacterial density (OD_{600} = 0.15) for the subsequent analyses, because the effect of cytosine methylation on CRISPR/Cas9-induced mutagenesis frequency was partially masked when a higher bacterial density was used (OD_{600} = 1.5; Fig. S2B). This might be due to chromatin/histone changes that are related to quickly repeating cycles of Cas9 cleavage.

As well as observing that extrinsic factors (virus infection and bacterial density) significantly and consistently affected the rates of mutagenesis for different target loci, we noted a strong intrinsic effect of the sequence itself for the different target loci (Fig. 7a). In other words, some target loci appear to be more prone to mutagenesis than others. Previous studies in human and yeast cells have come to similar conclusions that the efficacy of CRISPR/Cas9 mutagenesis is dependent on the sequence of the target locus (van Overbeek et al., 2016; Lemos et al., 2018). It is worth bearing in mind that the introduction of a mutation at a target locus relies not only on Cas9 recognition and subsequent DSB formation but also on the erroneous repair of that DSB. The sequence of a target locus could potentially affect all three of these steps (recognition, cleavage, and erroneous repair) that underlie mutagenesis. Parsing these effects remains a significant challenge to understanding how the underlying sequence of a target locus can predetermine the efficacy of CRISPR/Cas9 mutagenesis. Interestingly, even when we targeted the very same locus findings (Daer et al., 2017). Alternatively, DSBs might be more accurately repaired (resulting in lower rates of mutagenesis) by the RNA-templated DSB repair mechanism, in which RNA polymerase II transcripts can serve as the DNA repair template (Keskin et al., 2014). Further experiments are required to determine which of the above mechanisms may underpin the effect of DNA methylation on CRISPR/Cas9 activity in plants.
but on complementary DNA strands, we observed differences in the rates of mutagenesis. This effect was most visible for the mb positions in all epigenotypes. The mutation frequency induced by the mb gRNA in TRV-WT-infected plants was approximately four times higher than that of the mb gRNA, which targeted the same locus on the complementary strand (Figs 2c, S2C).

According to the original and widely accepted model, Cas9 nuclease produces blunt-ended DSBs (Jinek et al., 2012; Ahmad et al., 2020). More recent studies on animal and yeast models showed that, in vivo, the Cas9 endonuclease can also form staggered DSBs with single-nucleotide overhangs (Li et al., 2015; Lemos et al., 2018). These results were also supported by molecular dynamics simulations (Zuo & Liu, 2016). Our analyses of insertion and deletion events clearly showed that a high proportion of these mutations can be most simply explained as repairs of staggered DSBs with single-nucleotide 5′ overhangs that are first filled in and then ligated presumably by c-NHEJ (Beying et al., 2021), which is the most prevalent DNA repair pathway in differentiated nonproliferating cells such as the leaf tissue in our study. We reviewed data from previous studies describing CRISPR/Cas9 mutations in several other plant species (including Arabidopsis, maize and cotton) and found that staggered-end formation can also explain their most frequently inserted nucleotides (Li et al., 2013, 2019; Nekrasov et al., 2013; Lee et al., 2019). These observations suggested that Cas9 can form both blunt and staggered ends in planta. Interestingly, our data also revealed that the frequency of blunt and staggered-end cuts is predetermined by the CRISPR/Cas9 target site (Figs 4, 5, S9, S11), and also partially influenced by the epigenotype (e.g. mt, dt and db). To our knowledge, this is the first description of such an effect in any model system. To explain this phenomenon, we speculate that the process of Cas9 unwinding the DNA duplex at a target locus is sensitive to the physical tension of DNA in the locus, such that relaxed DNA duplexes are more easily unwound than DNA under tension. The tensioned conformation of DNA could therefore affect the position of the RuvC nuclease domain of Cas9 relative to the nontargeted strand, causing a shift of the cut by one nucleotide further from the PAM and resulting in a staggered DSB (Fig. 7b). Therefore, the relaxed DNA might favour the formation of blunt DSBs, whereas DNA under tension may favour the formation of a staggered DSB. In agreement with our model, Cas9 in vitro assays with DNA in a relaxed conformation produced blunt cuts (Sternberg et al., 2014), whereas when we analysed data from in vivo studies, for which DNA is more likely to be under tension at certain loci, we inferred that Cas9 frequently produces staggered DSBs (Li et al., 2013, 2019; Nekrasov et al., 2013; Lee et al., 2019). DNA tension could be affected by the local chromatin environment, mainly by DNA-interacting proteins such as histones, transcription factors and polymerases. Their binding to DNA can be affected (or even directed) by cytosine methylation, which could result in a highly localised and strand-specific increase or decrease in DNA tension.

In our system, we also observed that the epigenetic status of gRNA target sites could influence the frequency of CRISPR/Cas9-induced deletions (Fig. 3b). Interestingly, the vast majority of deletion events happened on the PAM-distal DNA end (Figs 5c, S10, group 2). We also found similar deletion patterns in other plant studies (Li et al., 2013, 2019; Nekrasov et al., 2013; Kumar et al., 2018; Lee et al., 2019; Fellenberg et al., 2020), but not in animal model systems (van Overbeek et al., 2016). Based on these findings, we assumed that in plants the PAM-distal cleavage product is released first from the CRISPR/Cas9 complex, and not just the nontargeted PAM-distal DNA strand, as suggested by an animal in vitro study (Richardson et al., 2016). In such a release mechanism, the PAM-distal DNA end is more exposed to exonucleases, leading to deletions occurring predominantly at this side of the cut, whereas the PAM-proximal DNA end remains protected and mostly unchanged (Fig. 7b). Moreover, our analysis suggests that a significant proportion of such deletions may be directed by a 5′ single nucleotide at the PAM-proximal end of the Cas9-induced DSB.

---

Fig. 7 Overview of CRISPR/Cas9 mutagenesis outcomes. (a) A table summarising the main conclusions of our study. Checkmarks and crosses indicate statistically significant and nonsignificant effects, respectively. Mixed marks designate statistical significance for some, but not all target sites. (b) Proposed model for how DNA tension may influence the outcome of Cas9-induced double-stranded breaks (DSBs) to produce either a blunt or a staggered DSB and 5′ single-nucleotide mediated DNA repair. On relaxed DNA, Cas9 forms a blunt DSB three nucleotides upstream from protospacer adjacent motif (PAM) due to the RuvC and HNH domains of Cas9 cleaving between the 3′rd and 4th nucleotides upstream from the PAM position on the PAM strand and non-PAM strand, respectively (left panel). DNA under tension could stretch the PAM strand, causing the RuvC domain of Cas9 to cleave between the 4th and 5th nucleotides upstream of the PAM. This, together with HNH-mediated cleavage between the 3′rd and 4th nucleotides upstream of the PAM position on the non-PAM strand would result in a staggered DSB with a 1-nt 5′ overhang on the PAM strand (right panel). After cleavage by Cas9, the PAM-distal DNA strand is probably released first because most of the deletions happened on the PAM-distal DNA (see Fig. 5). Moreover, the nucleotide adjacent to the cut on the PAM-containing proximal strand serves as a single-nucleotide searching for complementarity to allow specific 5′ microhomology-mediated DNA repair (bottom panels).

---

Table 1: Summary of CRISPR/Cas9 mutagenesis outcomes

| Variable                  | Epigenotype | Target site | CRISPR/Cas9 concentration | Virus infection |
|---------------------------|-------------|-------------|---------------------------|-----------------|
| Mutagenesis efficacy      | ✔           | ✔           | ✔                         | ✔               |
| Ratio of blunt/staggered  | ✔           | ✔           | ✔                         | ✔               |
| Ratio of IN : DEL         | ✔           | ✔           | ✔                         | ✔               |
| Size of DEL               | ✔           | ✔           | ✔                         | ✔               |
| Size of IN                | ✔           | ✔           | ✔                         | ✔               |
(Figs 7b, S13), which we also observed in other plant data sets (data not shown). Intriguingly, the well characterised microhomology-dependent DNA repair pathway (MMEJ/alt-EJ) is associated with 3’ DNA overhangs and proliferating cells (Steier & Symington, 2015). Therefore, it is plausible that the 5’ single-nucleotide microhomology-mediated DNA repair that we described in this study may be governed by NHEJ or novel genetic factors. However, more data including genetic evidence is required to identify the underpinning molecular mechanism and consequently assign the corresponding DNA repair pathway.

Taken together, our findings (Fig. 7a) have demonstrated that DNA methylation can influence the mutation frequency and mutational outcomes in the context of CRISPR/Cas9 genome editing, probably indirectly through changes in chromatin state. Our findings including the Cas9 staggered-end-mediated insertions and deletions and 5’ DNA single-nucleotide mediated DNA repair (Fig. 7b) may contribute to better target selection and gRNA designs for diverse future applications in plants and beyond.

Acknowledgements

We are grateful to Professor David Baulcombe for 16c N. benthamiana seeds, Dr Holger Puchta for pEn_Chimera and pDe_CAS9 plasmids, Muriel Monteiro and Aron Ferenczi for fruitful discussions, and to Professor Andrew Hudson for critically reading the manuscript. This work was supported by the Royal Society International Collaboration Award IC170320, Wellcome grant (206194) and AP was supported by the Charles University Society International Collaboration Award IC170320, Wellcome Trust. The NGS data have been deposited in the ArrayExpress database (E-MTAB-10256).

Competing interests

None declared.

Author contributions

AM performed conceptualisation; AM and AP performed methodology; AP performed software; AP, AM, LF and AB investigated the research; AP, LF and AM wrote original draft; DEP and AB reviewed and edited; AP visualised; and AM supervised the study.

ORCID

Andrew Bassett https://orcid.org/0000-0003-1632-9137
Lukáš Fischer https://orcid.org/0000-0003-1852-0461
Attila Molnar https://orcid.org/0000-0002-1044-6327
Adela Pribylová https://orcid.org/0000-0001-8675-5065
Douglas E. Pyott https://orcid.org/0000-0001-6340-9266

Data availability

The NGS data have been deposited in the ArrayExpress database at EMBL-EBI (www.ebi.ac.uk/arrayexpress) under accession number E-MTAB-10256.

References

Ahmad S, Wei X, Sheng Z, Hu P, Tang S. 2020. CRISPR/Cas9 for development of disease resistance in plants: recent progress, limitations and future prospects. Briefings in Functional Genomics 19: 26–39.
Allen F, Crespaldi L, Alsinet C, Strong AJ, Klechkevichov K, De Angelis P, Páleníková P, Khodak A, Kiselev V, Kosicki M et al. 2019. Predicting the mutations generated by repair of Cas9-induced double-strand breaks. Nature Biotechnology 37: 64–72.
Bae S, Park J, Kim J-S. 2014. CAS-OFFINDER: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. Bioinformatics 30: 1473–1475.
Beying N, Schmidt C, Puchta H. 2021. Double strand break (DSB) repair pathways in plants and their application in genome engineering. In: Willmann MR, ed. Genome editing for precision crop breeding. London, UK: Burleigh Dodds Science Publishing: 27–62.
Bothmer A, Phadke T, Herrera LA, Margules CM, Lee CS, Buquicchio F, Moss S, Adhikari HS, Selleck W, Jayaram H et al. 2017. Characterization of the interplay between DNA repair and CRISPR/Cas9-induced DNA lesions at an endogenous locus. Nature Communications 8: 13905.
Brinkman EK, Chen T, Amendola M, van Steensel B. 2014. Easy quantitative assessment of genome editing by sequence trace decomposition. Nucleic Acids Research 42: e168.
Chen X, Rinjma M, Janssen JM, Liu J, Maggio I, Goncaves MAFV. 2016. Probing the impact of chromatin conformation on genome editing tools. Nucleic Acids Research 44: 6482–6492.
Chiruvella KK, Liang Z, Wilson TE. 2013. Repair of double-strand breaks by end joining. Cold Spring Harbor Perspectives in Biology 5: a012757.
Clement K, Rees H, Canver MC, Gehrke JM, Farouni R, Hsu JY, Cole MA, Liu DR, Joung J, Bauer DE et al. 2019. CRISPResso2 provides accurate and rapid genome editing sequence analysis. Nature Biotechnology 37: 224–226.
Coleman-Derr D, Zilberman D. 2012. Deposition of histone variant H2AZ within gene bodies regulates responsive genes. PLoS Genetics 8: e1002988.
Cuí Y, Xu J, Cheng M, Liao X, Peng S. 2018. Review of CRISPR/Cas9 sgRNA design tools. Interdisciplinary Sciences: Computational Life Sciences 10: 455–465.
Daer R, Barrett CM, Haynes KA. 2018. Manipulation of chromatin to enhance CRISPR activity. BioRxiv. doi: 10.1101/228601.
Daer RM, Cutts JP, Braffman DA, Haynes KA. 2017. The impact of chromatin dynamics on Cas9-mediated genome editing in human cells. ACS Synthetic Biology 6: 428–438.
Farnung L, Ochmann M, Engesholm M, Cramer P. 2021. Structural basis of nucleosome transcription mediated by Chd1 and FACT. Nature Structural & Molecular Biology 28: 382–387.
Fauser F, Schiml S, Puchta H. 2014. Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in Arabidopsis thaliana. The Plant Journal 79: 348–359.
Fei Y, Nikó T, Molnár A. 2021. Non-perfectly matching small RNAs can induce stable and heritable epigenetic modifications and can be used as molecular markers to trace the origin and fate of silencing RNAs. Nucleic Acids Research 49: 1900–1913.
Fellenberg C, Corea O, Yan L-H, Archinuk F, Piirtola E-M, Gordon H, Reichelt M, Brandt W, Wulff J, Ehling J et al. 2020. Discovery of salicyl benzoate UDP-glycosyltransferase, a central enzyme in poplar salicinoid phenolic glycoside biosynthesis. The Plant Journal 102: 99–115.
Feng C, Yuan J, Wang R, Liu Y, Birchler JA, Han F. 2016. Efficient targeted genome modification in maize using CRISPR/Cas9 system. Journal of Genetics and Genomics 43: 37–43.
Fernandez-Pozo N, Menda N, Edwards JD, Saha S, Teclle FY, Stricker SR, Bombarely A, Fisher-York T, Pajar A, Foerster H et al. 2015. The Sol Genomics Network (SGN)—from genotype to phenotype to breeding. Nucleic Acids Research 43: D1036–D1041.
Fujita T, Yuno M, Fujii H. 2016. Allele-specific locus binding and genome editing by CRISPR at the p16INK4a locus. Scientific Reports 6: 30485.
Hinz JM, Laughery MF, Wyrick JJ. 2015. Nucleosomes inhibit Cas9 endonuclease activity in vitro. Biochemistry 54: 7063–7066.

© 2022 The Authors
New Phytologist © 2022 New Phytologist Foundation
www.newphytologist.com
Horbeck MA, Witkowski LB, Guglielmi B, Reploge JM, Gilbert LA, Villalta JE, Torigoe SE, Tjian R, Weissman JS. 2016. Nucleosomes impede Cas9 access to DNA in vivo and in vitro. eLife 5: e12677.

Hsu PD, Scott DA, Weinstein JA, Caparso D, Carter M, Thompson M, Frias E, Russ C, Reece-Hoyes J, Nye C, Gradi S, Vital B et al. 2016. DNA repair profiling reveals nonrandom outcomes at Cas9-mediated breaks. Molecular Cell 63: 633–646.

Hoyes J, Nye C, Gradia S, Vidal B, Fischer L. 2019. Detailed insight into the dynamics of the initial phases of de novo RNA-directed DNA methylation in plant cells. Epigenetics & Chromatin 12: 54.

Hoyes J, Nye C, Gradia S, Vidal B, Fischer L. 2019. Detailed insight into the dynamics of the initial phases of de novo RNA-directed DNA methylation in plant cells. Epigenetics & Chromatin 12: 54.

Ikejiri Y, Ohashi K, Itai Y, Itoh H, Itoh M, Iwasaki N, Ito S, Kikuchi T, Kishida H, Kiyosawa K et al. 2018. siRNA-directed genome editing enhanced by guide RNA and Cas9. Nature 552: 515–519.

Imamura Y, Yamamoto M, Nagata K, Nakamura K, Suzuki T, Yamauchi T, Kato T, Kurokawa K, Yamamoto H, Oda H et al. 2018. Transcription-RNA-templated DNA recombination and repair. Nature 555: 343–347.

Inoue T, Nishimura K, Yoshimura M, Nakao Y, Endo J, Metodiev P, Chom发生在 CRISPR-Cas9 selects and activates specific DNA targets. Nature 561: 200–205.

Inoue J, Nishimura K, Yoshimura M, Nakao Y, Endo J, Metodiev P, Chom发生在 CRISPR-Cas9 selects and activates specific DNA targets. Nature 561: 200–205.

Ishibashi K, Ishii K, Ueno M, Suzuki Y, Sato K, Ueda H, Kuroiwa A, Yagi K, Sato H, Tsuchida H et al. 2018. CRISPR-P 2.0: an improved CRISPR-Cas9 tool for genome editing in plants. Molecular Plant 10: 530–532.

Ikenmeyer CE, Bornemann AS, Petkova M, Hagerman D, Tang X, Sun D, Li J, Liu W, Zhang X, Fang J et al. 2018. Enhancing genome editing in Arabidopsis by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. Nature Biotechnology 36: 761–769.

Ishii K, Ishibashi K, Ueno M, Suzuki Y, Sato K, Ueda H, Kuroiwa A, Yagi K, Sato H, Tsuchida H et al. 2018. CRISPR-P 2.0: an improved CRISPR-Cas9 tool for genome editing in plants. Molecular Plant 10: 530–532.

Ikenmeyer CE, Bornemann AS, Petkova M, Hagerman D, Tang X, Sun D, Li J, Liu W, Zhang X, Fang J et al. 2018. Enhancing genome editing in Arabidopsis by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. Nature Biotechnology 36: 761–769.

Ikenmeyer CE, Bornemann AS, Petkova M, Hagerman D, Tang X, Sun D, Li J, Liu W, Zhang X, Fang J et al. 2018. Enhancing genome editing in Arabidopsis by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. Nature Biotechnology 36: 761–769.

Ikenmeyer CE, Bornemann AS, Petkova M, Hagerman D, Tang X, Sun D, Li J, Liu W, Zhang X, Fang J et al. 2018. Enhancing genome editing in Arabidopsis by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. Nature Biotechnology 36: 761–769.

Ikenmeyer CE, Bornemann AS, Petkova M, Hagerman D, Tang X, Sun D, Li J, Liu W, Zhang X, Fang J et al. 2018. Enhancing genome editing in Arabidopsis by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. Nature Biotechnology 36: 761–769.
Fig. S3 Comparison of the overall frequency of mutations induced by CRISPR/Cas9 in P35S.

Fig. S4 The impact of gene-body methylation on CRISPR/Cas9-induced mutations.

Fig. S5 Analysis of DNA (cytosine) methylation at CRISPR/Cas9 gRNA target sites in individual epigenotypes.

Fig. S6 Correlation analysis between the mutation frequencies and the percentage of methylated cytosines in the target locus in the compared epigenotypes.

Fig. S7 Comparison of insertion (IN) and deletion (DEL) events in individual CRISPR/Cas9 target sites.

Fig. S8 Insertion events analysis in P35S and GFP coding region.

Fig. S9 Detailed analysis of single-nucleotide insertion events detected next to the Cas9 cut site in individual epigenotypes.

Fig. S10 Detailed analysis of the type and frequency of deletion mutations in individual epigenotypes.

Fig. S11 Detailed analysis of group 2 and its first adjacent position next to the cut site on the PAM-distal DNA end in individual epigenotypes.

Fig. S12 Detailed analysis of the normalised frequency of deletion reads surrounded by no (0 nt) or 1 nucleotide microhomology in individual target sites and epigenotypes.

Fig. S13 The pattern of the deletion sizes in individual target sites.

Notes S1 DNA cytosine methylation within coding sequences does not alter the frequency of CRISPR/Cas9-induced mutations.

Table S1 List of primers used in this study.

Table S2 List of vectors used in this study.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the New Phytologist Central Office.

See also the Commentary on this article by Raffan et al., 235: 2146–2148.