This is a post-peer-review, pre-copyedit version of an article published in Proceedings of the National Academy of Sciences of the USA. The final authenticated version is available online at: https://doi.org/10.1073/pnas.2013338118

Efficient CRISPR-mediated base editing in *Agrobacterium* spp.

Savio Rodrigues[a], Mansour Karimi[b,c], Lennert Impens[b,c], Els Van Lerberge[b,c], Griet Coussens[b,c], Stijn Aesaert[b,c], Debbie Rombaut[b,c], Dominique Holtappels[e], Heba M.M. Ibrahim[a], Marc Van Montagu[b,c,d,1], Jeroen Wagemans[e], Thomas B. Jacobs[b,c], Barbara De Coninck[a,1] and Laurens Pauwels[b,c,1]

[a]Division of Crop Biotechnics, Department of Biosystems, KU Leuven, 3001 Leuven, Belgium; [b]Ghent University, Department of Plant Biotechnology and Bioinformatics, 9052 Ghent, Belgium; [c]VIB-UGent Center for Plant Systems Biology, 9052 Ghent, Belgium; [d]International Plant Biotechnology Outreach, VIB, 9052 Ghent, Belgium; [e]Laboratory of Gene Technology, Department of Biosystems, Division Animal and Human Health Engineering, KU Leuven, 3001 Leuven, Belgium

ORCID ID**:** 0000-0003-1641-6310 (S.R.), 0000-0002-0246-9318 (M.K.); 0000-0002-8407-7413 (E.V.L.); 0000-0003-2285-5782 (G.C.); 0000-0001-8787-7347 (S.A.); 0000-0003-4263-3407 (D.H); 0000-0003-0603-7755 (H.I.); 0000-0003-4711-5131 (M.V.M.); 0000-0002-6330-4744 (L.I.); 0000-0002-2185-5724 (J.W.); 0000-0002-5408-492X (T.B.J); 0000-0002-9349-5086 (B.D.C.); 0000-0002-0221-9052 (L.P.)

Footnotes

Author contributions: S.R., M.K., M.V.M., J.W., B.D.C., and L.P. designed experiments; S.R, M.K., L.I., E.V.L., G.C., and S.A. performed experiments; S.R., L.I., H.I., D.H., J.W., B.D.C. and L.P. analyzed the data; S.R, D.H., B.D.C., and L.P. wrote the article with contributions from all the authors. All authors edited, proofread, and approved the final version of this manuscript.

The authors declare no competing interests.

[1]To whom correspondence may be addressed. E-mail marc.vanmontagu@ugent.be; barbara.deconinck@kuleuven.be; laurens.pauwels@psb.vib-ugent.be

This article contains supporting information online at htpps://www.pnas.org/…….

Short title: Base editing in Agrobacterium

ABSTRACT

Agrobacterium spp. are important plant pathogens that are the causative agents of crown gall or hairy root disease. Their unique infection strategy depends on the delivery of part of their DNA to plant cells. Thanks to this capacity, these phytopathogens became a powerful and indispensable tool for plant genetic engineering and agricultural biotechnology. Although *Agrobacterium* spp. are standard tools for plant molecular biologists, current laboratory strains have remained unchanged for decades and functional gene analysis of *Agrobacterium* has been hampered by time-consuming mutation strategies. Here, we developed CRISPR-mediated base editing to enable the efficient introduction of targeted point mutations into the genomes of both *Agrobacterium tumefaciens* and *A. rhizogenes*. As an example, we generated EHA105 strains with loss-of-function mutations in *recA*, that were fully functional for maize (*Zea mays*) transformation and confirmed the importance of RolB and RolC for hairy root development by *A. rhizogenes* K599. Our method is highly effective in 9 of 10 colonies after transformation, with edits in at least 80% of the cells. The genomes of EHA105 and K599 were resequenced and genome-wide off-target analysis was applied to investigate the edited strains after curing of the base editor plasmid. The off-targets present were characteristic of Cas9-independent off-targeting and point to TC motifs as activity hotspots of the cytidine deaminase used. We anticipate that CRISPR-mediated base editing is the start of 'engineering the engineer', leading to improved *Agrobacterium* strains for more efficient plant transformation and gene editing.

Keywords: *Agrobacterium* | plant transformation | CRISPR | hairy root disease | base editing

SIGNIFICANCE STATEMENT

Agrobacteria are plant-pathogenic bacteria that can deliver DNA to plant cells as part of their infection strategy. This property has been used for decades to generate transgenic plants and, more recently, to deliver gene-editing reagents to plant cells. Notwithstanding their importance for research and industry, laboratory strains have not been improved much over the years and several aspects of *Agrobacterium* biology and pathogenesis remain poorly understood. Here we developed a CRISPR-mediated base-editing approach to efficiently modify the genome of *Agrobacterium*. We show that single nucleotide changes can be introduced at targeted positions in both the *Agrobacterium tumefaciens* and *A. rhizogenes* genomes. Whole-genome analysis of edited strains revealed only a limited number of unintentional mutations.

INTRODUCTION

Agriculture is a vital sector of the global economy that is now challenged to sustainably provide food for a growing population in a changing climate. Innovative genetic strategies offer opportunities for engineering crops tailored to the modern needs and expectations of society (1, 2). Besides careful observation of phenotypes, breeders also require tools to precisely modify and insert/delete sequences. A powerful system for delivering genes or gene-editing reagents into plants cells is *Agrobacterium*-mediated transformation.

Agrobacterium spp. are Gram-negative α-proteobacteria belonging to the *Rhizobiaceae* family. Plant-pathogenic strains, such as *Agrobacterium tumefaciens* C58 and *A. rhizogenes* K599, possess tumor-inducting (Ti) or root-inducing (Ri) plasmids, of which the transferred DNA (T-DNA) is transported into the plant nucleus where it can integrate into the host genome (3, 4). The plant host-derived phenolic compound acetosyringone (AS) is recognized by the virulence A (VirA) transcription factor in *Agrobacterium* and subsequently activates VirG that then binds to the *vir* boxes in the promoter of the T-DNA transfer-mediating *vir* genes (3, 5). In the case of *A. tumefaciens*, the naturally occurring T-DNAs contain genes that trigger transformed plant cells to produce auxins and cytokinins, leading to tumor (crown gall) formation (6–8) and subsequently the production of opines that are metabolized by the bacteria (9). In contrast, transformation by *A. rhizogenes* causes extensive hairy root proliferation. The mechanism by which the T-DNA-encoded genes function, such as the *root locus B* (*rolB*), is still not completely understood (10).

Removal of the T-DNA(s) on the Ti and/or Ri plasmids results in nononcogenic strains that can be used for plant transformation by launching an artificial T-DNA from a binary plasmid (11). Nononcogenic strains, such as *A. tumefaciens* EHA105, have been essential for genetic transformation and gene editing of a wide variety of plants (12, 13). Moreover, *Agrobacterium*

spp. serve as model organisms to study bacterial cell biology, host-microbe associations, and biofilm formation (14, 15). Genetic modifications for research on gene function or improvement of transformation-related traits in *Agrobacterium* spp. are typically carried out via allelic replacement or transposon mutagenesis (16). However, these techniques are often time consuming and laborious.

The clustered regularly interspaced short palindromic repeats (CRISPR)-associated protein 9 (Cas9) technology has been successfully used for genome editing in eukaryotic organisms (17–19). Cas9 is guided to a target DNA sequence (protospacer) by association with a single guide RNA (sgRNA) molecule where it induces DNA double-stranded breaks (DSBs) (20). The location of the target in the genome is restricted by the protospacer-adjacent motif (PAM) that is essential for Cas9 (20). Eukaryotic cells typically use the nonhomologous end-joining (NHEJ) pathway to repair the break, often resulting in either an insertion or a deletion at the site that can disrupt the gene function. The NHEJ machinery is typically absent in prokaryotes (21) and CRISPR/Cas9-induced chromosomal DSBs can be lethal in the absence of an endogenous DSB repair system.

Base editing uses either a nuclease-dead Cas9 (dCas9, D10A, and H840A mutations) or a Cas9 nickase (nCas9, D10A, or H840A) that can still bind the target DNA, but does not induce DSBs. When fused to a cytidine or adenosine deaminase, these Cas9 fusion proteins can generate C•G to T•A or A•T to G•C base pair transitions (22, 23) and are referred to as cytosine base editors (CBEs) and adenine base editors (ABEs), respectively. Depending on the deaminase used and the base editor architecture, mutations can occur in or near the protospacer, within what is referred to as a ‘deamination window’. Unintended bases within the deamination window can also be deaminated and are referred to as bystander mutations (24). Although base editing has been pioneered in eukaryotes, it has also been used in prokaryotes. A fusion protein, including nCas9

or dCas9 and a cytidine deaminase derived from sea lamprey (*Petromyzon marinus*) (PmCDA1), has been successfully used in *Escherichia coli*, *Corynebacterium glutamicum*, and *Streptomycetes* spp. with a base-editor architecture known as Target-AID (25–27). Whereas other CBE architectures have been used in prokaryotes as well, base editing has not been applied in the *Rhizobiaceae* yet (23, 28–33).

Here, we designed a curable base-editing system for *Agrobacterium* spp. with a Target-AID architecture. We validated our system in two hypervirulent *Agrobacterium* strains that are widely used for plant transformation, i.e. *A. tumefaciens* EHA105 and *A. rhizogenes* NCPPB2659 (hereafter referred to as EHA105 and K599, respectively). The targeted genes were selected based on their potential beneficial role for the plant transformation and gene editing field (*Atu1060*, *Atu4309*, and *recA* for *A. tumefaciens*) or on their involvement in hairy root disease (the cucumopine synthase-encoding gene [*cus*], *rolB*, *rolC*, and *orf13* for *A. rhizogenes*) (10, 34–37). By means of whole-genome resequencing, the off-targets were evaluated in detail. The developed base-editing system for *Agrobacterium* spp. enables us to efficiently "engineer nature's engineer" and to generate genome-edited strains for use in biotechnological applications, microbe-host interactions, and bacterial cell biology studies.

RESULTS

CRISPR/Cas9 editing in *Agrobacterium*. We first developed a CRISPR/Cas9 construct containing the *Streptococcus pyogenes* Cas9-coding sequence under the control of a constitutive promoter active in both *E. coli* and *Agrobacterium* spp. We used the promoter driving the gene encoding the aminoglycoside resistance protein *aadA* in commonly used binary vectors (38). We

designed a sgRNA for a chromosomal target, the *Atu1060* gene from *A. tumefaciens* encoding a putative diguanylate cyclase. *Atu1060*-deficient strains have been reported to be hypervirulent (35). The bacterial artificial promoter J23119 (PJ23119) was used to drive the expression of the sgRNA. As *Agrobacterium* spp. are resistant to an array of antibiotics, but not spectinomycin (39), we combined Cas9 and the sgRNA in one well-characterized binary vector with spectinomycin resistance (*SI Appendix,* Fig. S1*A*). After transformation of EHA101 and selection on the presence of the plasmid, no or very few colonies were obtained in contrast to a control without functional sgRNA (*SI Appendix*, Fig. S1*B*). Surviving colonies did not have any mutations at the targeted site (*SI Appendix,* Fig. S1*C*). These experiments confirm that CRISPR/Cas9 is effective in *A. tumefaciens* and that DSBs are not efficiently repaired, with lethality as a consequence when the encoded genes are chromosomally encoded.

Efficient base editing in *Agrobacterium* spp. The recently successful use of Target-AID in *E. coli* (25) and the lethality of the CRISPR/Cas9-mediated editing in *A. tumefaciens* prompted us to develop a base-editing system for *Agrobacterium* spp.. The Target-AID base editor consists of dCas9 fused to the *P. marinus* CDA1, a uracil DNA glycosylase inhibitor (UGI), and the LVA protein degradation tag (Fig. 1*A*) (25). When the Target-AID CBE was cloned under the control of the constitutive promoter used above, no correct clones were obtained, possibly due to lethality from the constitutive expression of the base editor in the *E. coli* cloning host. Hence, we placed the CBE under the control of the well-characterized AS-activated *A. tumefaciens virB* promoter (P*virB*) (5, 40) (Fig. 1 *A* and *B*). Based on the *virB* promoters from the plasmids pTiA6 and pTiBo542, a conserved *virB* promoter fragment was defined, including two *vir*-boxes, a -10 box, and a ribosome-binding site (RBS) (Fig. 1*B*). This fragment from pTiBo542 was cloned, leaving

an optimal 7 bp between the RBS and the ATG of the base editor (Fig. 1*B*). Again PJ23119 was used for the sgRNA expression (Fig. 1*A*).

The same functional sgRNA targeting the *Atu1060* gene from *A. tumefaciens* EHA101 was used (Fig. 1*C*). A CAG codon is present as a target for C editing 19 bp upstream from a PAM sequence (Fig. 1*D*). After transformation of the construct to the *A. tumefaciens* strain EHA101 (41), single spectinomycin-resistant colonies were selected and the region of interest was sequenced, revealing that even without AS-induced P*virB*, the target C at position -19 had already been edited into a thymidine (T) in all tested colonies (Fig. 1*D*, *SI Appendix*, Fig. S2*A*). Moreover, a bystander C at position -15 was edited as well (Fig. 1*D*). Quantification of the editing efficiency showed that 3/3 colonies had on average 91% editing of the on-target C and 25% editing of a bystander C (*SI Appendix*, Fig. S2*A*, Dataset S1), indicating that after transformation the individual colonies were in fact mixed populations, as reported also for base editing in *Corynebacterium glutamicum* and *Clostridium beijerinckii* (26, 30). Correspondingly, streak plating resulted in pure colonies with a single, clear genotype (Fig. 1*D*). An immunoblot done with antibodies against Cas9 confirmed leaky expression of the CBE under the control of P*virB* without AS treatment (*SI Appendix*, Fig. S3). As treatment with AS did not influence the efficiency of the editing or the deamination window, AS was not used in the remainder of the work (*SI Appendix*, Fig. S2*B*, Dataset S1). Base editing was validated for two other sgRNAs in EHA101 that targeted *Atu4309*. This locus is associated with the M-1 transposon mutant strain, reported to transfer its T-DNA into the nucleus of the host cell, but without stable integration (34). A similar editing efficiency was observed (*SI Appendix*, Fig. S2 *C-E*, Dataset S1).

To test whether base editing is also functional in *A. rhizogenes* K599 and to analyze whether we can base-edit two Ri plasmid-encoded genes with one sgRNA, we targeted the C at

position -15 of *cus* (B0909_24515) and its homologous gene (B0909_24450; *SI Appendix*, Fig. S4*A*) that are both encoded on the pRi2659 plasmid of K599. In *E. coli*, the deamination window of Target-AID is 16-20 bp upstream of the PAM motif, when a 20-bp spacer is selected and decrease or increase of the spacer length was reported to move the deamination window toward or away from the PAM, respectively (25). Therefore, we used a spacer of 18 bp targeting both homologous genes simultaneously. All five tested colonies showed at least 98% editing of the C at the targeted position without AS induction (*SI Appendix*, Fig. S4 *B-D*, Dataset S1). In conclusion, we constructed a Target-AID-based CBE that can be used for efficient editing of *Agrobacterium* spp.

CBE removal for strain development. To avoid continuous CBE activity, we aimed at curing the *Agrobacterium* strains of the CBE-encoding plasmid. To this end, we introduced a levansucrase-encoding gene (*sacB*) module into the construct (*SI Appendix*, Fig. S5) that has been shown to be a useful negative selection marker when *A. tumefaciens* and *A. rhizogenes* are grown on sucrose (16, 42, 43). To evaluate the base editing with this construct, we first tested two of the sgRNAs used before. Targeting *Atu1060* and *Atu4309* in *A. tumefaciens* EHA105, an EHA101 derivative, resulted in a similar efficiency with all colonies having at least 80% editing of the targeted C (*SI Appendix*, Fig. S6 *A* and *B*). Next, two sgRNAs were designed targeting the *recA* gene (*Atu1874*; Fig. 2*A*). *recA* is involved in the DNA mismatch repair pathway and *recA*-deficient mutants have been utilized in plant transformation studies to maintain the stability of vectors that contain multiple repetitive elements (44). Use of both sgRNAs separately (Fig. 2 *B* and *C*) or combined in a single construct (*SI Appendix*, Fig. S6 *C-F*) provided high on-target editing in the strain EHA105, implying that the system can also be used for multiplexing. Overall, when a range

of experiments in *A. tumefaciens* was examined, 9 out of 10 transformed colonies showed edits in at least 80% of the cells (*SI Appendix,* Dataset S1).

After *sacB*-based curing of the CBE with sucrose, all cells within a colony had the same genotype and contained the edited stop codon at the target loci in *recA* (Fig. 2 *D* and *E*). For each sgRNA, one clone was retained and designated EHA105-*recA(Q26*)* and EHA105-*recA(Q178*)*. As agrobacteria deficient in *recA* are hypersensitive to DNA-damaging reagents, such as methyl methanesulfonate (MMS) (45, 46), we used a drop test to check the sensitivity to 0.005% (v/v) MMS for EHA105, EHA105-*recA(Q26*)*, and EHA105-*recA(Q178*).* Whereas EHA105 was only moderately sensitive to 0.005% (v/v) MMS, both *recA* mutants exhibited severely reduced growth, providing evidence for the loss of *recA* activity in these strains by the presence of an early stop codon (Fig. 2*F*).

Next, we tested these new *recA* strains for their transformation potential in maize (*Zea mays*). The reporter construct pXBb7-SI-UBIL (38) with the maize ubiquitin-1 promoter (Zm*UBI-1*) driving the *gus* reporter gene containing the potato (*Solanum tuberosum)* PIV2 intron was transformed in both cured EHA105 *recA*-deficient strains and in EHA105 as a control. Immature embryos from the inbred line B104 were isolated and co-cultivated for 3 days with the three *Agrobacterium* genotypes. After 4 days on resting media, embryos were stained for GUS activity. The *recA*-deficient EHA105 strains behaved similarly to the EHA105 parent strain with comparable numbers of embryos showing transient expression and similar areas of the scutellum epithelium scoring positive for GUS (Fig. 2*G*, *SI Appendix*, Fig. S7 *A*-*D*). In conclusion, base editing combined with *sacB*-based curing very efficiently produced *recA* mutants in the transformation "workhorse" EHA105.

Generation of K599 mutant strains affected in hairy root formation. For K599, we designed sgRNAs to target the T-DNA-encoded genes *rolB* (B0909_24535), *rolC* (B0909_24530), and *orf13* (B0909_24525). The *Agrobacterium* phenotypic plasticity (*plast*) genes *rolB* and *rolC* were selected because of their significant contribution to the development of the hairy root phenotype (10, 47). We also included *orf13*, another *plast* gene for which the contribution to the initiation and development of hairy roots is unclear (10, 48). Similar to *A. tumefaciens*, the *A. rhizogenes* CBE-cured mutants K599-*rolB(R31*)*, K599-*rolB(W167*)*, K599-*rolC(Q40*)*, and K599-*orf13(Q29*)* were obtained (Fig. 3 *A*-*D*). To evaluate the ability of K599 and the mutant strains to initiate hairy roots, a carrot (*Daucus carota*) disk assay was carried out. Here, the percentage of carrot disks developing hairy roots was evaluated over time (Fig. 3 *E* and *F*). Additionally, the number of hairy roots per carrot disk was analyzed at the end of the experiment (*SI Appendix*, Fig. S8). Both K599-*rolB* and the K599-*rolC* mutants displayed a significantly reduced ability to induce hairy roots, whereas the K599-*orf13* mutant behaved similarly to K599 (Fig. 4 *E* and *F; SI Appendix*, Fig. S8). Our observations are in line with the previously reported crucial roles of *rolB* and *rolC* in inducing hairy root formation (10, 47, 49, 50).

***De novo* genome assembly of EHA105 and K599.** Whole-genome sequencing and *de novo* genome assembly of both EHA105 and K599 strains were carried out for the assessment of the genome-wide off-target effect of the CBE and because bacterial stocks maintained at different laboratories can acquire diverging mutations over time (51). EHA105 has an A136 background and an introduced helper plasmid pEHA105, which results from two subsequent allelic replacements to delete the T-DNA region of the type III Ti plasmid pTiBo542 (41, 52, 53).

The genome assemblies led to four and three different contigs, corresponding to the four replicons (circular chromosome, linear chromosome, and megaplasmids pAtC58 and pEHA105) for EHA105 and three replicons (circular chromosome, linear chromosome, and the megaplasmid pRi2659) for K599. A BLAST search of the EHA105 replicons against the C58 reference genome revealed a query coverage of 99.66% for the linear chromosome (NC_003063) and 100% for the circular chromosome (NC_003062) and pAtC58 (NC_003064). pEHA105 (contig 4) showed 77.35% query coverage to pTiBo542 (DQ058764), corresponding to the deletion of the 57,231-bp large T-region of the 244,978-bp pTiBo542 while retaining the outer left border (*SI Appendix,* Fig. S9) (41, 52, 53). A single-nucleotide polymorphism (SNP) analysis revealed that 65 single-nucleotide variants (SNVs) or indels were present compared to the used reference genome of the EHA105 laboratory strain (*SI Appendix,* Dataset S2), of which 43 were nonsynonymous mutations. Noteworthy, changes in pEHA105 were a frameshift in *virM* (*SI Appendix,* Dataset S2), an AS-induced gene not essential for tumorigenesis (54), and several complex mutations in *traA* (*SI Appendix,* Dataset S2), involved in DNA transfer during conjugation and linked to Ti plasmid copy control in the bacterial population (55). The sequenced K599 strain shared 100% sequence identity to the available K599 reference sequences for the circular chromosome (NZ_CP019701), the linear chromosome (NZ_CP019702), and pRi2659 (NZ_CP019703). However, we detected 300 bp extra at the end of the linear chromosome and 2 bp missing in the Ri-plasmid.

Whole-genome sequencing reveals a low level of spurious deamination. The high-quality whole-genome sequences for the EHA105 and K599 laboratory strains allowed us to evaluate potential CBE-caused off-target effects. Two base-edited and subsequently cured strains of *A. tumefaciens* [EHA105-*recA(Q26*)* and EHA105-*recA(Q178*)*] and three cured strains of *A.*

rhizogenes [K599-*rolB(R31*)*, K599-*rolC(Q40*)*, and K599-*orf13(Q29*)*] were evaluated for off-targets. SNP analysis revealed 39 [*recA(Q26**)], 34 [*recA(Q178**)], 62 [*rolB(R31*)*], 17 [*rolC(Q40*)*], and 27 [*orf13(Q29*)*] off-targets, resulting in 13, 14, 30, 7, and 9 nonsynonymous mutations, respectively (Fig. 4 *A* and *B*; *SI Appendix*, Dataset S3), in line with previous reports on CBE-caused off-targets in other prokaryotes (25, 31). In our analyses, only one common SNV was detected between the different strains. Strains K599-*rolB(R31*)* and K599-*orf13(Q29*)* shared this SNV, which is remarkably, at the very end of the linear chromosome (position: 2276818). C residues were almost exclusively converted to T residues (Fig. 4*A*), strongly indicating that these SNVs are the result of the Target-AID off-target activity. Interestingly, none of the observed off-targets were predicted by means of the software Geneious. Hence, no off-targets were similar to the spacers used, as observed for other CBEs in mouse embryos and rice (*Oryza sativa*) (56, 57).

The deaminases used in CBEs, including PmCDA1, are known to have intrinsic Cas9-independent sequence specificity, leading to activity hotspots (spurious deamination) when expressed separately or as part of a CBE (24). Analysis of all off-targets obtained here revealed a sequence preference (61.8-76.9%) for TC motifs by PmCDA1 in the Target-AID architecture (Fig. 4*C*). Moreover, a reanalysis of the off-target SNVs from *E. coli* expressing unguided dCas9-CDA-UL (23) confirmed our observations, because 55.2-61.9% of the off-targets were related to TC motifs as well (*SI Appendix*, Fig. 4*D*). In conclusion, although no obvious off-targeting by ectopic sgRNA binding could be detected in our system, the spurious deamination was probably caused by intrinsic PmCDA1 DNA binding and activity.

DISCUSSION

Currently, CRISPR-based gene editing makes it possible to efficiently introduce mutations at specific genomic locations (58). However, in plants, *Agrobacterium*-mediated delivery of CRISPR reagents to plant cell nuclei and the subsequent plant regeneration are the bottlenecks. For both processes, the efficiency depends highly on species and genotype (13). Recently, progress has been made with the use of morphogenic regulators that increase regeneration (59–61) and ternary plasmids for *A. tumefaciens* equipped with extra virulence genes (62, 63). To avoid recombination of these virulence genes with the homologous sequences on pTi, *recA*-deficient strains are typically used (63, 64). Another development warranting the utilization of *recA*-deficient strains is the ever-increasing complexity of engineered T-DNAs, often with repetitive elements, as exemplified by an array of 24 sgRNAs (65). Our base-editing approach allows the introduction of a *recA* loss-of-function mutation into a desired *A. tumefaciens* or *A. rhizogenes* strain within days.

Other desired *Agrobacterium* traits in the plant transformation and gene editing field include auxotrophic strains for biocontainment and improved tissue culture (66), inactive transposons (67), reduced vector backbone integration (63, 68, 69), increased T-DNA transfer (70), and transient-only T-DNA delivery (71, 72). Finally, agrobacteria that evade or subvert host plant factors that are detrimental to transformation efficiency, such as the plant defense system, can be critical to develop transformation systems for recalcitrant species or cultivars (73–75). To uncover unknown mutations in relation to these traits, we foresee that the high efficiency of the base-editing method will allow CRISPR library screens for functional genomics (26, 76). Indeed, our CBE can be used with multiplexing, allowing desired genotypes to be easily combined, thus “engineering the engineer”.

Here, we resequenced the nononcogenic *A. tumefaciens* strain EHA105 that is routinely used for plant transformation and the oncogenic *A. rhizogenes* strain K599. A nononcogenic K599 has been generated as well (42, 77) and successfully used for the genetic transformation of maize, *Brachypodium distachyon*, and soybean (*Glycine max*) (42, 77, 78). The CBE described here could be used for the future development of this strain as an alternative for plant transformation. Oncogenic *A. rhizogenes* strains are utilized in laboratories worldwide for generating transgenic hairy roots for research purposes (79), *in vitro* production of secondary metabolites (80), and breeding of compact plants (81). Finally, the K599 base-edited mutants will allow us to study the characterization of genes involved in virulence and symptom development of hairy root disease (10). Knowledge on how *A. rhizogenes* hijacks plant pathways could facilitate plant breeding efforts for increased resistance against hairy root disease (82).

Traditionally, targeted allelic replacement, random transposon mutagenesis, and recombineering have been used to knock out genes in agrobacteria (16, 83, 84), but CRISPR-based systems have the advantages of simplicity and efficiency. The obtained colonies can be cured of the CBE plasmid to acquire stable mutants for functional analysis or for further downstream use. CBE-free mutant strains can again be equipped with binary vectors for plant transformation or subject to further rounds of genome editing. In contrast to allelic replacement, base editing can also be done with multiple sgRNAs in parallel by means of multiplexing. Here, we have applied base editing for the introduction of stop codons (85) but it can be used for directed amino acid changes as well. However, the requirement that the targeted codons be located within the deamination window is a limiting factor. In prokaryotes, this location appears to be 16-20 bp upstream of the PAM for Target-AID (25). To alter the deamination window, the spacer length can be modulated as successfully done here (25). In addition, engineered Cas9 variants with

alternative PAM sites, such as Cas9-VQR (86) and the recently developed nearly PAM-less Cas9 (87) could easily be incorporated into our system if needed.

Although genome-wide off-target analysis did not identify obvious Cas9-dependent off-targets, limited spurious deamination was observed, typical for CBEs (56, 57). Human activation-induced deaminase (hAID) is known to have an intrinsic DNA-binding activity with a preference for the WRC motif (88) and, more specifically, the GC motif, when used in a CBE (32). In contrast, PmCDA1 prefers *in vitro* deamination of substrates with a T or a C at the -1 position (89). Our off-target analysis suggests that PmCDA1 in the Target-AID architecture preferentially binds TC motifs for C deamination *in vivo*. This native sequence preference of some deaminases can lead to poor efficiency on certain on-target residues, such as GC in APOBEC1 (90). Although a tendency for TC in the off-targets was detected, the desired on-target mutations could be obtained for all the sgRNAs tested. Nevertheless, the TC preference of PmCDA1 might be considered in the sgRNA design for more sensitive applications, such as library screens. In the future, other (engineered) CBEs, ABEs, or prime editing with potentially less Cas9-independent off-targeting could be applied as an alternative for editing agrobacteria, provided that the efficiency is not impacted (91–93).

MATERIALS AND METHODS

Strains, media, and conditions. Detailed information on *E. coli* and *Agrobacterium* growth conditions are provided in *SI Materials and Methods*. Bacterial strains used in this study are detailed in *SI Appendix*, Table S1.

Vector construction. Details on the construction of the CRISPR/Cas9 and on the base-editing constructs are provided in *SI Materials and Methods*. An overview of the vector generation by means of Golden Gate and MultiSite Gateway™ cloning is presented (*SI Appendix*, Fig.S5) and lists of the plasmids and oligonucleotides used are provided as well (*SI Appendix*, Tables S1 and S2). Plasmids are available at https://gatewayvectors.vib.be. Spacers were designed with Geneious (*SI Appendix*, Table S3).

Base editing and plasmid curing. Base-editing constructs were transferred to *A. tumefaciens* via heat shock or to *A. rhizogenes* via electroporation. For selection, cells were grown on spectinomycin-containing Luria-Bertani medium or on glucose-spectinomycin-streptomycin--containing AT minimal medium for *A. tumefaciens* and *A. rhizogenes*, respectively. The targeted locus was amplified by PCR and sequenced. The chromatograms were analyzed with EditR (94). For *sacB*-based curing, edited *Agrobacterium* strains were streak plated on sucrose-containing medium and incubated at 28°C. More details are provided in *SI Materials and Methods.*

Plant transformation and functional assays. Maize was transformed as described (95). Details on the transformation, MMS susceptibility assays, carrot disk virulence assays, molecular and statistical analyses are provided in *SI Materials and Methods.*

Next-generation sequencing and bioinformatics. Genomes of the *A. tumefaciens* EHA105 and *A. rhizogenes* K599 strains were sequenced by combining Illumina short-read and Nanopore long-read sequencing. Details on the sequencing, analysis, and determination of off-targets are given in *SI Materials and Methods.*

Data Availability. Associated genome sequencing data including raw reads are available at the EBI European Nucleotide Archives (ENA) under the project reference PRJEB38304 (http://www.ebi.ac.uk/ena/data/view). The annotated genomes and genome assemblies for the isolates of *A. rhizogenes* K599 and *A. tumefaciens* EHA105 are deposited under the accessions and GCA_903772885 and GCA_903772965, respectively. All raw sequencing reads are available under the accession numbers ERR4184086; ERR4184084; ERR4183659; ERR4183658; ERR4183657; ERR4183656; ERR4183655; ERR4183654; ERR4183653.

ACKNOWLEDGEMENTS

We thank Rudy Vanderhaeghen and Long Nguyen (VIB-Ghent University, Ghent, Belgium) and Astrid Severyns, Lut Ooms and Joram Moons (KU Leuven) for technical assistance and Keiji Nishida and Satomi Banno (Kobe University, Kobe, Japan) for providing the pScI_dCas–CDA–UL and pTAKN2_J23119-sgRNA vectors. This work was supported by Ghent University ('Bijzonder Onderzoeksfonds Methusalem project' no. BOF15/MET_V/004) and KU Leuven (Internal Funds). S.R. is supported by a predoctoral fellowship from the Research Foundation-Flanders (1S43920N). We are grateful to the KU Leuven HPC infrastructure and the Flemish Supercomputer Center (VSC) for providing the computational resources and services to perform the variant calling analysis.

REFERENCES

1. S. S.-e.-A. Zaidi *et al.*, New plant breeding technologies for food security. *Science* **363**, 1390-1391 (2019).

2. J. Bailey-Serres, J. E. Parker, E. A. Ainsworth, G. E. D. Oldroyd, J. I. Schroeder, Genetic strategies for improving crop yields. *Nature* **575**, 109-118 (2019).

3. B. Lacroix, V. Citovsky, Pathways of DNA transfer to plants from *Agrobacterium tumefaciens* and related bacterial species . *Annu. Rev. Phytopathol.* **57**, 231–251 (2019).

4. W. Ream, *Agrobacterium tumefaciens* and *A. rhizogenes* use different proteins to transport bacterial DNA into the plant cell nucleus. *Microb. Biotechnol.* **2**, 416–427 (2009).

5. S. E. Stachel, P. C. Zambryski, *virA* and *virG* control the plant-induced activation of the T-DNA transfer process of *A. tumefaciens*. *Cell* **46**, 325–333 (1986).

6. A. Follin *et al.*, Genetic evidence that the tryptophan 2-monooxygenase gene of *Pseudomonas savastanoi* is functionally equivalent to one of the T-DNA genes involved in plant tumour formation by *Agrobacterium tumefaciens*. *Mol. Gen. Genet*. **201**, 178-185 (1985).

7. L. S. Thomashow, S. Reeves, M. F. Thomashow, Crown gall oncogenesis: evidence that a T-DNA gene from the *Agrobacterium* Ti plasmid pTiA6 encodes an enzyme that catalyzes synthesis of indoleacetic acid. *Proc. Natl. Acad. Sci. USA* **81**, 5071-5075 (1984).

8. G. F. Barry, S. G. Rogers, R. T. Fraley, L. Brand, Identification of a cloned cytokinin biosynthetic gene. *Proc. Natl. Acad. Sci. USA* **81**, 4776-4780 (1984).

9. J. Schell *et al.*, Interactions and DNA transfer between *Agrobacterium tumefaciens*, the Ti-plasmid and the plant host. *Proc. R. Soc. Lond. B.* **204**, 251–266 (1979).

10. L. Otten, The *Agrobacterium* phenotypic plasticity (*Plast*) genes. *Curr. Top. Microbiol. Immunol.* **418**, 375–419 (2018).

11. T. Komari *et al.*, Binary vectors and super-binary vectors. *Methods Mol. Biol.* **343**, 15–41 (2006).

12. S. B. Gelvin, *Agrobacterium*-mediated plant transformation: the biology behind the "gene-jockeying" tool. *Microbiol. Mol. Biol. Rev.* **67**, 16–37 (2003).

13. F. Altpeter *et al.*, Advancing crop transformation in the era of genome editing. *Plant Cell* **28**, 1510–1520 (2016).

14. N. Feirer *et al.*, The *Agrobacterium tumefaciens* CheY-like protein ClaR regulates biofilm formation. *Microbiology* **163**, 1680–1691 (2017).

15. J. R. Zupan, R. Grangeon, J. S. Robalino-Espinosa, N. Garnica, P. Zambryski, GROWTH POLE RING protein forms a 200-nm-diameter ring structure essential for polar growth and rod shape in *Agrobacterium tumefaciens*. *Proc. Natl. Acad. Sci. U.S.A.* **166**, 10962–10967 (2019).

16. E. R. Morton, C. Fuqua, Genetic manipulation of *Agrobacterium*. *Curr. Protoc. Microbiol.* **25**, 3D.2.1-3.D.2.15 (2012).

17. L. Cong *et al.*, Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013).

18. P. Mali *et al.*, RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826 (2013).

19. J. E. DiCarlo *et al.*, Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Res.* **41**, 4336–4343 (2013).

20. M. Jinek *et al.*, A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).

21. T. E. Wilson, L. M. Topper, P. L. Palmbos, Non-homologous end-joining: Bacteria join the chromosome breakdance. *Trends Biochem. Sci.* **28**, 62–66 (2003).

22. A. C. Komor, Y. B. Kim, M. S. Packer, J. A. Zuris, D. R. Liu, Programmable editing of a

target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**, 420–424 (2016).

23. N. M. Gaudelli *et al.*, Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature* **551**, 464–471 (2017).

24. H. A. Rees, D. R. Liu, Base editing: precision chemistry on the genome and transcriptome of living cells. *Nat. Rev. Genet.* **19**, 770–788 (2018).

25. S. Banno, K. Nishida, T. Arazoe, H. Mitsunobu, A. Kondo, Deaminase-mediated multiplex genome editing in *Escherichia coli*. *Nat. Microbiol.* **3**, 423–429 (2018).

26. Y. Wang *et al.*, MACBETH: Multiplex automated Corynebacterium glutamicum base editing method. *Metab. Eng.* **47**, 200–210 (2018).

27. Y. Zhao *et al.*, Multiplex genome editing using a dCas9-cytidine deaminase fusion in *Streptomyces*. *Sci. China Life Sci*, in press (2020) (doi: 10.1007/s11427-019-1559-y).

28. T. Gu *et al.*, Highly efficient base editing in *Staphylococcus aureus* using an engineered CRISPR RNA-guided cytidine deaminase. *Chem. Sci*. **9**, 3248–3253 (2018).

29. K. Zheng *et al.*, Highly efficient base editing in bacteria using a Cas9-cytidine deaminase fusion. *Commun. Biol.* **1**, 32 (2018).

30. Q. Li *et al.*, CRISPR–Cas9^{D10A} nickase-assisted base editing in the solvent producer *Clostridium beijerinckii*. *Biotechnol. Bioeng.* **116**, 1475–1483 (2019).

31. Y. Tong *et al.*, Highly efficient DSB-free base editing for streptomycetes with CRISPR-BEST. *Proc. Natl. Acad. Sci. USA* **116**, 20366–20375 (2019).

32. J. L. Doman, A. Raguram, G. A. Newby, D. R. Liu, Evaluation and minimization of Cas9-independent off-target DNA editing by cytosine base editors. *Nat. Biotechnol.* **38**, 620–628 (2020).

33. Z. Zhong *et al.*, Base editing in *Streptomyces* with Cas9-deaminase fusions. *bioRxiv* **630137** (2019).

34. P. Majumder, H. Shioiri, M. Nozue, M. Kojima, Isolation and characterization of a new virulence gene (*abvA*) of *Agrobacterium tumefaciens*. *J. Gen. Plant Pathol.* **67**, 124-133 (2001).

35. S. K. Farrand, D. M. Barnhart, Hypervirulent mutant of *Agrobacterium tumefaciens*. United States Patent Appl. US20100950833 (University of Illinois, United States) (2010).

36. M. Gupta, S. Bennett, S. Kumar, D. Merlo, N. Sardesai, Methods and compositions for recombination A gene-deficient strains of *Agrobacterium tumefaciens*. United States Patent Appl. 20160068851 (DOW Agrosciences LLC, United States) (2016).

37. K. Chen, L. Otten, Natural *Agrobacterium* transformants: Recent results and some theoretical considerations. *Front. Plant Sci.* **8**, 1600 (2017).

38. M. Karimi, A. Depicker, P. Hilson, Recombinational cloning with plant Gateway vectors. *Plant Physiol.* **145**, 1144–1154 (2007).

39. L.-Y. Lee, S. B. Gelvin, T-DNA binary vectors and systems. *Plant Physiol.* **146**, 325–332 (2008).

40. S. E. Stachel, E. W. Nester, The genetic and transcriptional organization of the *vir* region of the A6 Ti plasmid of *tumefaciens*. *EMBO J.* **5**, 1445–1454 (1986).

41. E. E. Hood, G. L. Helmer, R. T. Fraley, M.-D. Chilton, The hypervirulence of *Agrobacterium tumefaciens* A281 is encoded in a region of pTiBo542 outside of T-DNA. *J. Bacteriol.* **168**, 1291–1301 (1986).

42. S. L. Mankin *et al.*, Disarming and sequencing of *Agrobacterium rhizogenes* strain K599 (NCPPB2659) plasmid pRi2659. *In Vitro Cell. Dev. Biol.-Plant* **43**, 521–535 (2007).

43. S. M. Traore, B. Zhao, A novel Gateway®-compatible binary vector allows direct selection of recombinant clones in *Agrobacterium tumefaciens*. *Plant Methods* **7**, 42 (2011).

44. R. Collier, J. G. Thomson, R. Thilmony, A versatile and robust Agrobacterium-based gene stacking system generates high-quality transgenic Arabidopsis plants. *Plant J.* **95**, 573-583 (2018).

45. S. K. Farrand, S. P. O'Morchoe, J. McCutchan, Construction of an *Agrobacterium tumefaciens* C58 *recA* mutant. *J. Bacteriol.* **171**, 5314–5321 (1989).

46. G. R. Lazo, P. A. Stein, R. A. Ludwig, A DNA transformation–competent *Arabidopsis* genomic library in *Agrobacterium*. *Bio/Technology* **9**, 963–967 (1991).

47. J.-H. Wang *et al.*, Transcriptomic analysis reveals that reactive oxygen species and genes encoding lipid transfer protein are associated with tobacco hairy root growth and branch development. *Mol. Plant-Microbe Interact.* **27**, 678-687 (2014).

48. P. A. Stieger, *et al.*, The *orf13* T-DNA gene of *Agrobacterium rhizogenes* confers meristematic competence to differentiated cells. *Plant Physiol.* **135**, 1798–1808 (2004).

49. M. C. Failla, F. Maimone, A. De Paolis, P. Costantino, M. Cardarelli, The non-conserved region of cucumopine-type *Agrobacterium rhizogenes* T-DNA is responsible for hairy root induction. *Plant Mol. Biol.* **15**, 747–753 (1990).

50. J. Brevet, D. Borowski, J. Tempé, Identification of the region encoding opine synthesis and of a region involved in hairy root induction on the T-DNA of cucumber-type Ri plasmid. *Mol. Plant-Microbe Interact.* **1**, 75–79 (1988).

51. T. R. Ioerger *et al.*, Variation among genome sequences of H37Rv strains of *Mycobacterium tuberculosis* from multiple laboratories. *J. Bacteriol.* **192**, 3645–3653 (2010).

52. E. E. Hood, S. B. Gelvin, L. S. Melchers, A. Hoekema, New *Agrobacterium* helper plasmids

for gene transfer to plants. *Transgenic Res.* **2**, 208–218 (1993).

53. A. J. Weisberg *et al.*, Unexpected conservation and global transmission of agrobacterial virulence plasmids. *Science* **368**, eaba5256 (2020).

54. V. S. Kalogeraki, S. C. Winans, Wound-released chemical signals may elicit multiple responses from an *Agrobacterium tumefaciens* strain containing an octopine-type Ti plasmid. *J. Bacteriol.* **180**, 5660–5667 (1998).

55. K. M. Pappas, Cell-cell signaling and the *Agrobacterium tumefaciens* Ti plasmid copy number fluctuations. *Plasmid* **60**, 89–107 (2008).

56. E. Zuo *et al.*, Cytosine base editor generates substantial off-target single-nucleotide variants in mouse embryos. *Science* **364**, 289–292 (2019).

57. S. Jin *et al.*, Cytosine, but not adenine, base editors induce genome-wide off-target mutations in rice. *Science* **364**, 292–295 (2019).

58. K. Chen, Y. Wang, R. Zhang, H. Zhang, C. Gao, CRISPR/Cas genome editing and precision plant breeding in agriculture. *Annu. Rev. Plant Biol.* **70**, 667–697 (2019).

59. K. Lowe *et al.*, Morphogenic regulators *Baby boom* and *Wuschel* improve monocot transformation. *Plant Cell* **28**, 1998–2015 (2016).

60. B. Gordon-Kamm *et al.*, Using morphogenic genes to improve recovery and regeneration of transgenic plants. *Plants* **8**, 38 (2019).

61. M. F. Maher *et al.*, Plant gene editing through de novo induction of meristems. *Nat. Biotechnol.* **38**, 84–89 (2019).

62. A. Anand *et al.*, An improved ternary vector system for *Agrobacterium*-mediated rapid maize transformation. *Plant Mol. Biol.* **97**, 187–200 (2018).

63. N. Sardesai *et al.*, Coexpression of octopine and succinamopine *Agrobacterium* virulence

genes to generate high quality transgenic events in maize by reducing vector backbone integration. *Transgenic Res.* **27**, 539–550 (2018).

64. M. Mookkan, K. Nelson-Vasilchik, J. Hague, Z. J. Zhang, A. P. Kausch, Selectable marker independent transformation of recalcitrant maize inbred B73 and sorghum P898012 mediated by morphogenic regulators *BABY BOOM* and *WUSCHEL2*. *Plant Cell Rep.* **36**, 1477-1491 (2017).

65. K. Barthel *et al.*, One-shot generation of duodecuple (12x) mutant Arabidopsis: Highly efficient routine editing in model species. *bioRxiv* **18671** (2020).

66. J. I. Collens, D. R. Lee, A. M. Seeman, W. R. Curtis, Development of auxotrophic *Agrobacterium tumefaciens* for gene transfer in plant tissue culture. *Biotechnol. Prog.* **20**, 890–896 (2004).

67. S.-R. Kim, G. An, Bacterial transposons are co-transferred with T-DNA to rice chromosomes during *Agrobacterium*-mediated transformation. *Mol. Cells* **33**, 583–589 (2012).

68. B. Ülker *et al.*, T-DNA–mediated transfer of *Agrobacterium tumefaciens* chromosomal DNA into plants. *Nat. Biotechnol.* **26**, 1015–1017 (2008).

69. F. Jupe *et al.*, The complex architecture and epigenomic impact of plant T-DNA insertions. *PLoS Genet.* **15**, e1007819 (2019).

70. S. Nonaka, T. Someya, Y. Kadota, K. Nakamura, H. Ezura, Super-*Agrobacterium* ver. 4: improving the transformation frequencies and genetic engineering possibilities for crop plants. *Front. Plant Sci.* **10**, 1204 (2019).

71. F. Veillet *et al.*, Transgene-*free genome editing in tomato and potato plants using Agrobacterium*-mediated delivery of a CRISPR/Cas9 cytidine base editor. *Int. J. Mol. Sci.*

20, 402 (2019).

72. Y. Wang *et al.*, VirD5 is required for efficient *Agrobacterium* infection and interacts with Arabidopsis VIP2. *New Phytol.* **217**, 726–738 (2018).

73. N. Sardesai *et al.*, Cytokinins secreted by *Agrobacterium* promote transformation by repressing a plant Myb transcription factor. *Sci. Signal.* **6**, ra100 (2013).

74. E. E. Hwang, M. B. Wang, J. E. Bravo, L. M. Banta, Unmasking host and microbial strategies in the *Agrobacterium*-plant defense tango. *Front. Plant Sci.* **6**, 200 (2015).

75. A. Pitzschke, *Agrobacterium* infection and plant defense---transformation success hangs by a thread. *Front. Plant Sci.* **4**, 519 (2013).

76. O. Shalem, N. E. Sanjana, F. Zhang, High-throughput functional genomics using CRISPR-Cas9. *Nat. Rev. Genet.* **16**, 299-311 (2015).

77. R. Collier, J. Bragg, B. T. Hernandez, J. P. Vogel, R. Thilmony, Use of *Agrobacterium rhizogenes* strain 18r12v and paromomycin selection for transformation of *Brachypodium distachyon* and *Brachypodium sylvaticum*. *Front. Plant Sci.* **7**, 716 (2016).

78. P. M. Olhoft *et al.*, A novel *Agrobacterium rhizogenes*-mediated transformation method of soybean [*Glycine max* (L.) Merrill] using primary-node explants from seedlings. *In Vitro Cell. Dev. Biol.-Plant* **43**, 536–549 (2007).

79. M. Ron *et al.*, Hairy root transformation using *Agrobacterium rhizogenes* as a tool for exploring cell type-specific gene expression and function using tomato as a model. *Plant Physiol.* **166**, 455–469 (2014).

80. N. Gutierrez-Valdes *et al.*, Hairy root cultures–A versatile tool with multiple applications. *Front. Plant Sci.* **11**, 33 (2020).

81. B. Christensen, S. Sriskandarajah, M. Serek, R. Müller, Transformation of *Kalanchoe*

blossfeldiana with *rol*-genes is useful in molecular breeding towards compact growth. *Plant Cell Rep.* **27**, 1485–1495 (2008).

82. L. Bosmans *et al.*, Rhizogenic agrobacteria in hydroponic crops: epidemics, diagnostics and control. *Plant Pathol.* **66**, 1043–1053 (2017).

83. J.-P. Hernalsteens, H. De Greve, M. Van Montagu, J. Schell, Mutagenesis by insertion of the drug resistance transposon Tn7 applied to the Ti-plasmid of *Agrobacterium tumefaciens*. *Plasmid* **1**, 218–225 (1978).

84. S. Hu *et al.*, Genome engineering of *Agrobacterium tumefaciens* using the lambda Red recombination system. *Appl. Microbiol. Biotechnol.* **98**, 2165–2172 (2014).

85. C. Kuscu *et al.*, CRISPR-STOP: Gene silencing through base-editing-induced nonsense mutations. *Nat. Methods* **14**, 710–712 (2017).

86. B. P. Kleinstiver *et al.*, Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature* **523**, 481–485 (2015).

87. R. T. Walton, K. A. Christie, M. N. Whittaker, B. P. Kleinstiver, Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. *Science* **368**, 290–296 (2020).

88. P. Pham *et al.*, Activation-induced deoxycytidine deaminase: Structural basis for favoring WRC hot motif specificities unique among APOBEC family members. *DNA Repair* **54**, 8–12 (2017).

89. E. M. Quinlan, J. J. King, C. T. Amemiya, E. Hsu, M. Larijani, Biochemical regulatory features of activation-induced cytidine deaminase remain conserved from lampreys to humans. *Mol. Cell. Biol.* **37**, e00077-17 (2017).

90. B. W. Thuronyi *et al.*, Continuous evolution of base editors with expanded target

compatibility and improved activity. *Nat. Biotechnol.* **37**, 1070–1079 (2019) [Erratum *Nat. Biotechnol*. **37**, 1091].

91. N. M. Gaudelli *et al.*, Directed evolution of adenine base editors with increased activity and therapeutic application. *Nat. Biotechnol.*, in press (2020) https:/doi.org/10.1038/s41587-020-0491-6.

92. A. V. Anzalone, *et al.*, Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* **576**, 149–157 (2019).

93. Y. Yu *et al.*, Cytosine base editors with minimized unguided DNA and RNA off-target events and high on-target activity. *Nat. Commun.* **11**, 2052 (2020).

94. M. G. Kluesner *et al.*, EditR: A method to quantify base editing from Sanger sequencing. *CRISPR J.* **1**, 239–250 (2018).

95. G. Coussens *et al.*, *Brachypodium distachyon* promoters as efficient building blocks for transgenic research in maize. *J. Exp. Bot.* **63**, 4263–4273 (2012).

Figures

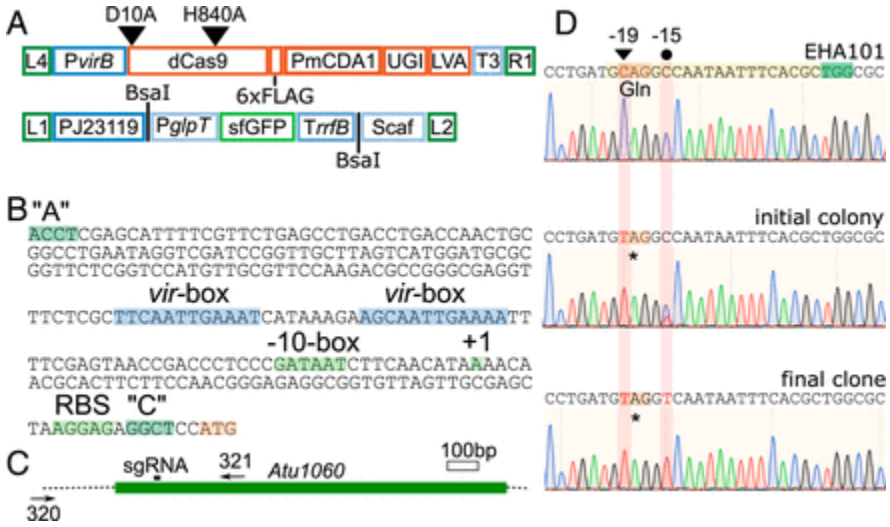

Fig. 1. Efficient base editing in *Agrobacterium tumefaciens*. *(A)* Schematic overview of the *Agrobacterium* Target-AID CBE and sgRNA expression cassettes. P*virB*, promoter fragment of the *A. tumefaciens* pTiBo542 *virB* gene; dCas9, a nuclease-dead Cas9 with mutations indicated; PmCDA1, *Petromyzon marinus* cytidine deaminase; UGI, uracil DNA glycosylase inhibitor; LVA, protein degradation tag; T3, terminator of the RNA polymerase III of bacteriophage T3; PJ23119, artificial promoter; P*glpT*, *E. coli glpT* promoter; sfGFP, superfolder GFP; T*rrfB*, *E. coli rrfB* terminator; Scaf, sgRNA scaffold; L1, R1, L4, and L2, Gateway recombination sites. *(B)* Overview of the *virB* promoter used. "A" and "C", GreenGate overhangs; *vir*-box, VirG recognition site; RBS, ribosome-binding site. The start codon of the base editor is highlighted in orange. *(C)* The *Atu1060* genomic locus. Locations of the protospacer used for the sgRNA are indicated together with the position of the primers used for genotyping. *(D)* Base-editing outcomes. *Top*, sequence obtained from EHA101 with the PAM (green highlighted), protospacer (yellow), and target codon (orange). The targeted C is indicated with a triangle, together with the position relative to the PAM. The solid circle denotes a bystander C. *Middle*, representative sequences from the colony obtained directly after transformation. *Bottom*, representative sequences obtained after streak plating without any prior AS induction (final clone). Edited bases are highlighted in red, the relevant codons are translated, and an asterisk indicates a stop codon. Gln, glutamine.

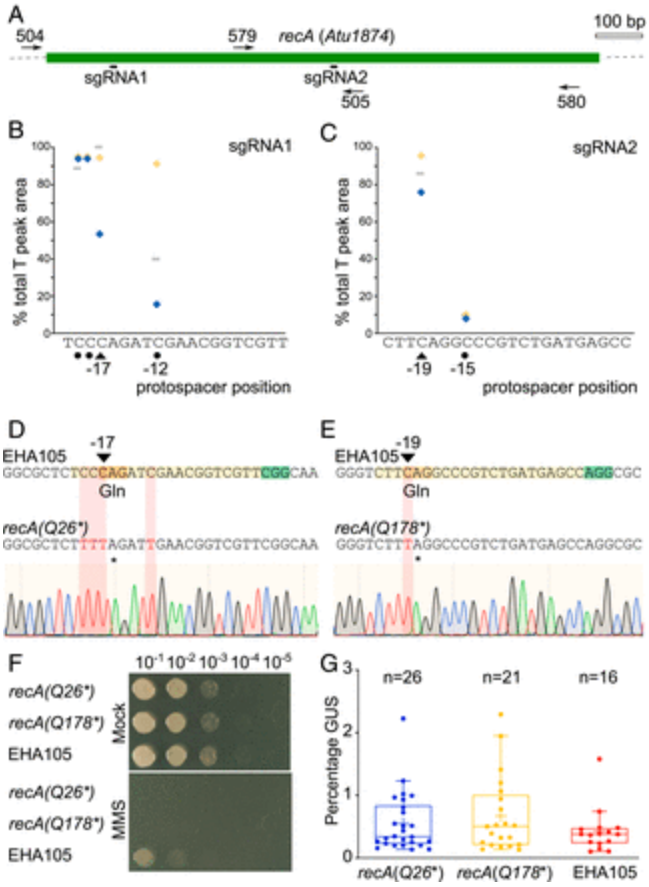

Fig. 2. Development of *A. tumefaciens* EHA105 *recA* mutants for plant transformation. *(A)* The *recA* genomic locus. Locations of the protospacers are indicated together with the position of the primers used for genotyping. (*B)* and (*C*) Quantification of the C→T editing before curing for sgRNA1 and sgRNA2, respectively. The percentage of the total peak area of the Sanger sequencing reads is shown at the different protospacer positions when an edit differed significantly from noise. Each series of symbols represents an independent colony after transformation (n = 3). Solid triangles and solid circles denote targeted C and bystander C, respectively, with the numbering indicating the positions relative to the PAM.

(*D*) and (*E*) Base editing outcomes after curing for sgRNA1 and sgRNA2, respectively. *Top,* sequence obtained from EHA105 is shown with the PAM (green), protospacer (yellow), and target codon (orange). The targeted C is indicated with a triangle, together with the position relative to the PAM. *Bottom,* sequence of the cured strain is shown together with the chromatogram. Edited bases are highlighted in red, the relevant codons are translated, and an asterisk indicates a stop codon. Gln, glutamine. *(F)* Drop test of the two EHA105 *recA* mutants obtained and the EHA105 parent strain. Cells were grown to OD_{600} of 0.5 and dropped in serial dilutions on YEP media (*Top*) or on YEP media with 0.005% (v/v) MMS (*Bottom*). *(G)* Percentage of the embryo surface scoring positive for GUS. Plotted are values for embryos from a single ear infected with the different strains. No statistically significant differences were observed between the

EHA105 parent strain and the EHA105 *recA* mutants (Kruskal-Wallis test and Dunn's multiple comparisons test, $P>0.05$). The box limits are the 25th and 75th percentile; the middle line in the individual boxes represents the median; the cross in the individual boxes marks the mean; whiskers extend to 1.5-fold the interquartile range from the 25th and 75th percentiles; the solid circles highlight data points.

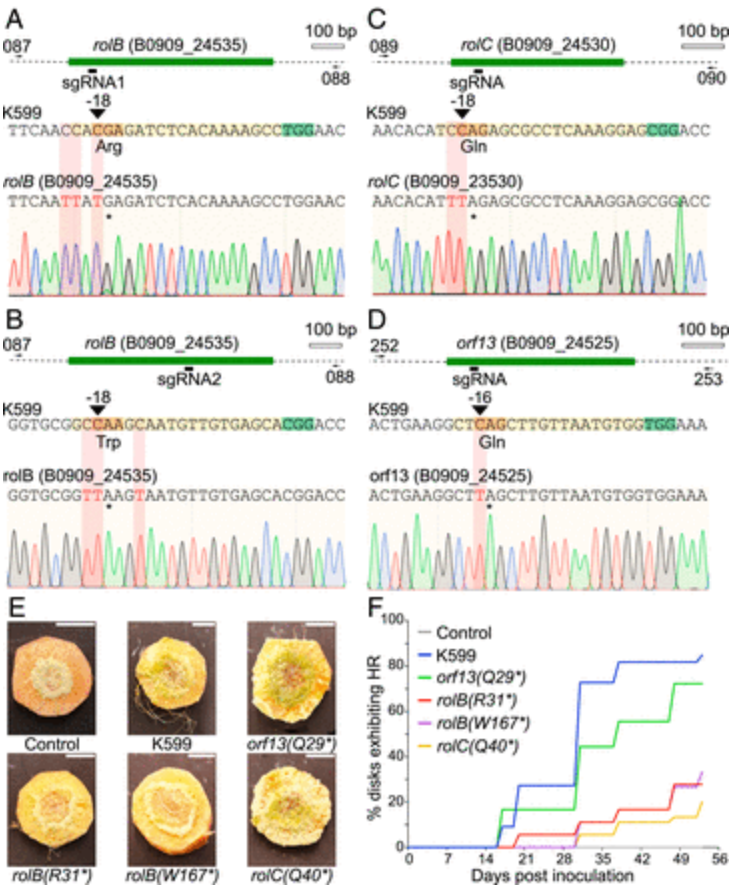

Fig. 3. *A. rhizogenes* K599 T-DNA mutants are affected in hairy root formation. (*A*-*D*) Base editing outcomes after curing. *Top*, sequence obtained from K599 is shown with the PAM (green), protospacer (yellow), and target codon (orange). The targeted C is indicated with a triangle, together with the position relative to the PAM. *Bottom*, sequence of the cured strain is shown together with the chromatogram. Edited bases are highlighted in red, the relevant codons are translated, and an asterisk indicates a stop codon. Arg, arginine; Gln, glutamine; Trp, tryptophan. *(E)* Representative pictures of carrot disks inoculated with K599 and various K599 mutants. Scale bars, 1 cm. *(F)* Percentage of carrot disks exhibiting hairy roots different days post inoculation with K599 and K599 mutants (n = 18). HR, hairy roots.

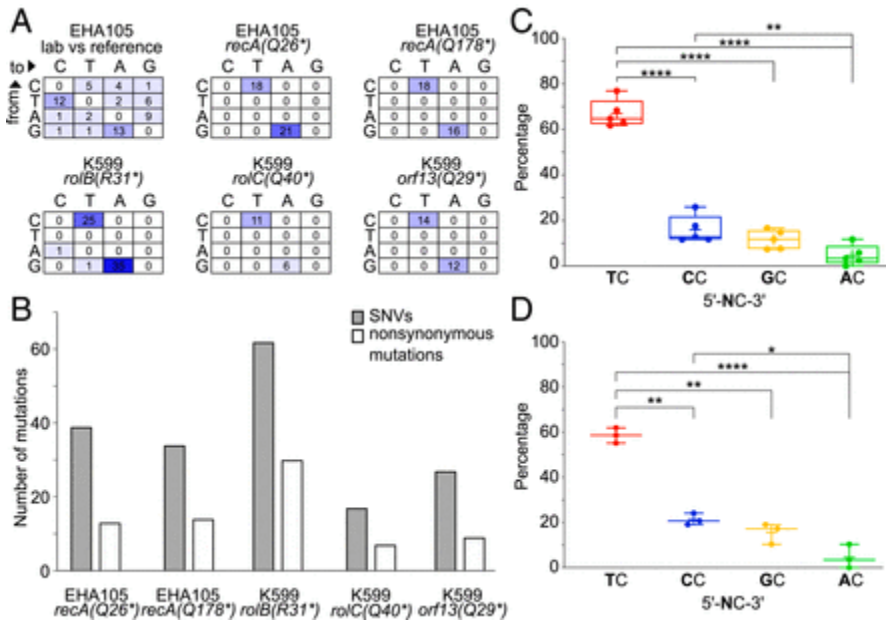

Fig. 4. Low level of spurious deamination revealed by whole-genome sequencing. *(A)* Total number of SNPs recorded for EHA105 against the reference sequence from *A. tumefaciens* C58 (NC_003062, NC_003063, NC_003064) and pTiBo542 (DQ058764) and total number of SNVs recorded between the base-edited strains against the sequenced laboratory strains EHA105 and K599. *(B)* Graph of the total number of recorded SNVs and of the synonymous/ nonsynonymous mutations recorded and compared as in (*A*). *(C)* and *(D)* Preference for 5'-TC-3' motifs demonstrated by guided Target-AID in *Agrobacterium* spp. and unguided Target-AID deamination in *E. coli* BW25113 (25), respectively. Boxplots highlight the percentage of deaminated C residues in off-targets in combination with the nucleotide located upstream of the C residue. Bold characters indicate the nucleotide preceding the deaminated C; the box limits are the 25th and 75th percentile; the middle line in the individual boxes represents the median; the cross in the individual boxes marks the mean; whiskers extend to 1.5 fold the interquartile range from the 25th and 75th percentiles; solid circles highlight data points; asterisks mark significant differences with * $P<0.05$, ** $P<0.01$, *** $P<0.001$ and **** $P<0.0001$ (One-way ANOVA and Tukey's multiple comparisons test); C, cytosine; T, thymine; G, guanine; A, adenine.