

This is a PDF file of an article that is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain. The final authenticated version is available online at: <https://doi.org/10.1111/nph.20415>

For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

This work was funded by European Research Council (POPMET project 834923).

Transgene-free genome editing in poplar

Lennart Hoengenaert^{1,2}, Chantal Anders^{1,2}, Jan Van Doorselaere³, Ruben Vanholme^{1,2} and Wout Boerjan^{1,2,*}

¹Department of Plant Biotechnology and Bioinformatics, Ghent University, Ghent, Belgium

²Center for Plant Systems Biology, VIB, Ghent, Belgium

³VIVES, University of Applied Sciences, Roeselare, Belgium

* corresponding author: wout.boerjan@psb.vib-ugent.be

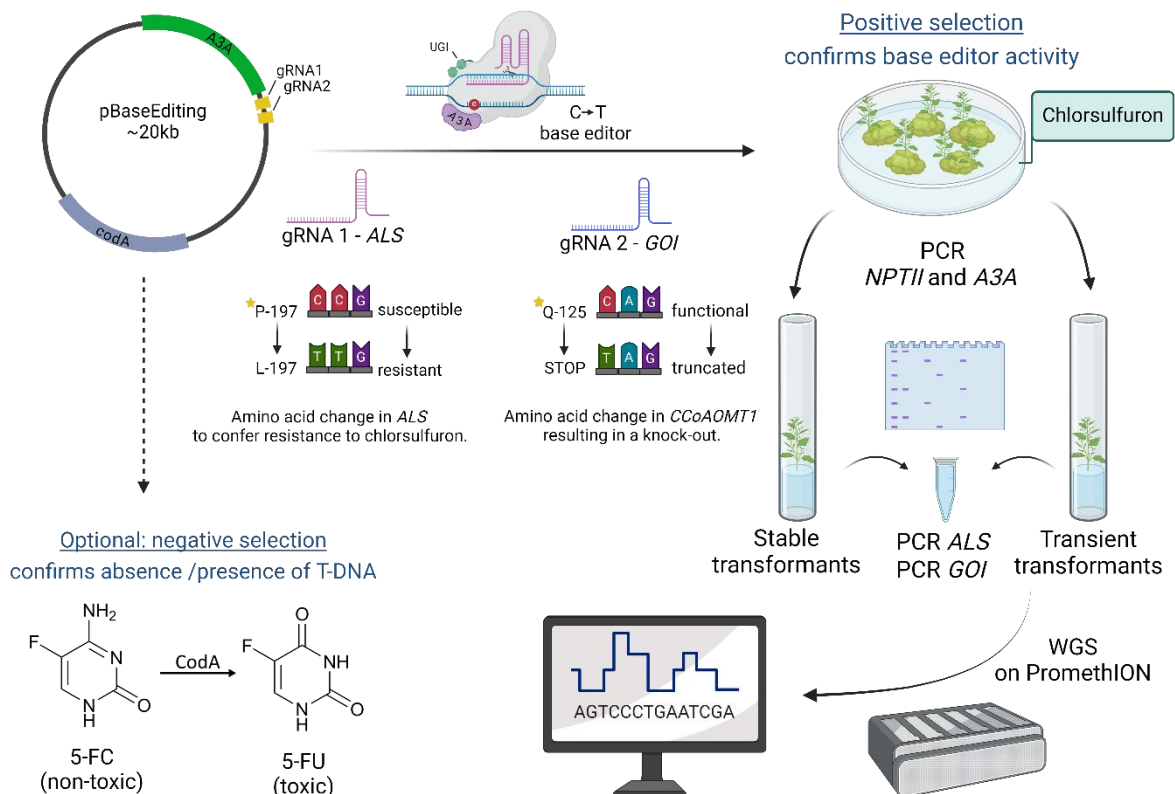
Keywords

CRISPR, gene editing, base editing, transgene-free, poplar

Abstract

Precise gene-editing methods are valuable tools to enhance genetic traits. Gene editing is commonly achieved via stable integration of a gene-editing cassette in the plant's genome. However, this technique is unfavorable for field applications, especially in vegetatively propagated plants, such as many commercial tree species, where the gene-editing cassette cannot be segregated out in subsequent generations. Here, we describe an efficient method for generating gene-edited *Populus tremula* x *P. alba* (poplar) trees without incorporating foreign DNA into the plant genome. Using *Agrobacterium tumefaciens*, we expressed a base-editing construct targeting *CCoAOMT1* along with the *ALS* genes for positive selection on a chlorsulfuron-containing medium. About 50% of the regenerated shoots were derived from transient transformation and were free of T-DNA. Overall, 7% of the chlorsulfuron-resistant shoots were T-DNA free, edited in the *CCoAOMT1* gene and non-chimeric. Long-read whole-genome sequencing confirmed the absence of any foreign DNA in the tested gene-edited lines. Additionally, we evaluated the *CodA* gene as a negative selection marker to eliminate lines that stably incorporated the T-DNA into their genome. Although the latter negative selection is not essential for selecting transgene-free, gene-edited *Populus tremula* x *P. alba* shoots, it may prove valuable for other genotypes or varieties.

Graphical abstract



Graphical abstract: Overview of the transgene-free base-editing method described in this paper. Two gRNAs, one targeting the *ACETOLACTATE SYNTHASE* gene and the other targeting the gene of interest (*GOI*) *CCoAOMT1*, were brought into poplar cells via *Agrobacterium*-mediated transformation, alongside an A3A base editor and a negative selection marker (*codA*). Effective base editing is identified by screening shoots for resistance to chlorsulfuron, which results from a point mutation in the *ALS* gene(s). Optionally, the negative selection marker can be used to eliminate gene-edited shoots that stably incorporated a copy of the T-DNA into their genome. In this paper, stable and transient transformants were screened for mutations in the *ALS* gene and the *GOI* via PCR and Sanger sequencing. Transient transformants, thus transformants containing no T-DNA, that were screened positively for mutations in the *GOI* were additionally sequenced on the PromethION to confirm the absence of T-DNA fragments.

Main text

Gene editing is revolutionizing plant breeding by enabling precise enhancement of genetic traits. Typically, the development of gene-edited crops involves the integration of a T-DNA, comprising genes encoding a Cas nuclease and one or more guide RNA (gRNA) sequences, into the plant's genome. However, growing gene-edited plants in the open field requires the removal of nuclease-encoding T-DNA sequences from the genome due to regulatory and public concerns¹. In seed-propagated annual crops, the removal of the T-DNA can be readily achieved through crossing the gene-edited, T-DNA-containing line with a transgene-free line followed by Mendelian segregation, allowing the uncoupling of the gene-editing machinery from the edited gene in the off-spring. However, for perennial crops and vegetatively propagated species, such as poplar, this approach poses significant challenges. Most tree species require years to reach sexual maturity, are outcrossing, and have a high degree of heterozygosity, resulting in the loss of the genetic constitution of elite genotypes upon sexual reproduction. To advance the genetic improvement of these species, it is essential to develop strategies that immediately allow transgene-free gene editing within the T0 generation^{2,3}.

To tackle these concerns, various groups have attempted to eliminate the T-DNA from the plant genome through site-specific recombination using the FLIPPASE (FLP)/FLP RECOGNITION TARGET SITE (FRT) system or by introducing Cas9-mediated DNA breaks near the T-DNA borders^{4,5}. However, trimming of T-DNA border sequences, before or during integration, leads to loss or incompleteness of these recombination and Cas9-target sites, rendering those methods inefficient and unpredictable⁴. Alternative approaches rely on either the delivery of ribonucleoproteins or transient expression of Cas-nucleases and gRNA sequence(s) in protoplasts. While protoplast transformation can be accomplished at high frequency, the subsequent regeneration process is tedious and restricted to just a few species or genotypes⁶⁻¹². Additionally, the culture and regeneration of protoplasts is known to induce somaclonal variation, resulting in unintended changes in the plant's genome, potentially yielding adverse effects^{13,14}. Another approach relies on the physical delivery of CRISPR DNA/RNA or ribonucleoproteins into undifferentiated callus cells through microprojectile bombardment^{15,16}. We previously investigated the feasibility of delivering microparticles coated with Cas9 and *NEOMYCIN PHOSPHOTRANSFERASE II (NPTII)*-encoding plasmid DNA into undifferentiated poplar callus. We then attempted to regenerate transgene-free, gene-edited trees using a short selection with kanamycin to enrich in cells that had been (transiently) transformed with the plasmid and expressed the included transgenes¹⁷. While the method was successful, its efficiency was low. Additionally, the method resulted in the frequent insertion of vector DNA fragments at the Cas9-targeted site, posing challenges for future commercialization and deregulation of the gene-edited plants. Recently, a method was described employing a co-editing strategy through *Agrobacterium tumefaciens*-mediated transient expression of a three-component system in tomato, tobacco, potato and pomelo. This system comprises a cytosine base editor to modify the *ACETOLACTATE SYNTHASE (ALS)* genes conferring herbicide resistance, a Cas12a nuclease to edit target genes, and a green fluorescent protein to screen for transgene-free transformants¹⁸.

Here, we developed a strategy to gene edit forest trees without transgene insertion (Fig. S1). The poplar hybrid *Populus tremula* x *P. alba* (*Pta*) cv. 717-1B4 has been a common model for molecular

genetic research in trees since it was first genetically transformed¹⁹. Utilizing *A. tumefaciens*, we transfected this poplar hybrid with a T-DNA carrying, next to the two gRNAs, also three genes encoding a base editor, a negative selection marker and the *NPTII* gene conferring resistance to kanamycin. The A3A base editor comprises an intron variant of the APOBEC3A (A3A) cytosine deaminase fused to a nickase Cas9 (nCas9)²⁰ (Fig. 1A). We designed one gRNA targeting the endogenous *ALS* genes in poplar (*ALS1* and *ALS2*) and another gRNA targeting the lignin biosynthesis gene *CAFFEYOYL-CoA O-METHYLTRANSFERASE1* (*CCoAOMT1*). *ALS* encodes a crucial enzyme involved in the biosynthesis of branched-chain amino acids. Many studies have demonstrated that specific point mutations in this gene confer resistance to the herbicide chlorsulfuron²¹, allowing the selection of plant cells that have or had an active base editing machinery (Fig. 1B). The deamination region of *CCoAOMT1* is targeted to introduce an early stop codon in the gene, resulting in a knock-out mutation. The negative selection marker on the T-DNA is an amino acid-modified cytosine deaminase (*CodA*) from *E. coli*²². *CodA* enzymatic activity converts non-cytotoxic 5-fluorocytosine (5-FC) to cytotoxic 5-fluorouracil (5-FU)²³. Using this strategy, we envisaged positive selection for base-editing activity with selection on chlorsulfuron, while allowing simultaneous negative selection of plantlets that stably expressed the T-DNA with selection on 5-FC. The kanamycin resistance conferred by the *NPTII* gene was only used to create lines with a stable integration of the T-DNA into the genome to evaluate the base-editing efficiency.

We first evaluated the base-editing efficiency for the different target genes by selecting transgenic lines that stably integrated the T-DNA. Following plant transformation and selection on kanamycin, we obtained 25 transgenic calli, from which one or multiple shoots were regenerated. Sanger sequencing of the three target genes, *ALS1*, *ALS2* and *CCoAOMT1*, in each regenerated shoot showed that 33 independent transgenic genotypes were generated; shoots from different calli are by definition independent genotypes and, additionally, shoots from the same callus with different mutations are also considered as independent (Fig. 1C, Extended data 1). By analyzing the sequencing chromatograms, we assessed both the efficiency of base editing and the overall editing efficiency at each nucleotide of the protospacer sequence. After all, A3A base editors possess both base-editing activity, conferred by the A3A cytosine deaminase, and endonuclease activity, conferred by the nCas9. Although base editing encompasses on-target C-to-T transitions as its primary function, it also allows C-to-G and C-to-A transitions, albeit at lower frequency, along with other off-target nucleotide transitions such as G-to-A, T-to-G, etc²⁴. In addition, gene editing can occur through the action of nCas9, creating a single-stranded nick resulting in insertion or deletion (INDEL) mutations²⁴. Sequence analyses of our stably transformed lines revealed base-editing efficiencies (C-to-T) up to 22%, 17% and 32% for *ALS1*, *ALS2* and *CCoAOMT1*, respectively (Fig. 1D). Gene-editing efficiencies (including INDELs) were 30%, 24% and 47% for *ALS1*, *ALS2* and *CCoAOMT1*, respectively (Fig. S2A, Table S1). To further explore the various mutation profiles within the *CCoAOMT1* gene at the allele level, we cloned and sequenced the *CCoAOMT1* PCR amplicons. Our data revealed that chimeric editing frequently occurs in the transgenic lines (Fig. 1C, Extended data 1, Table S2), suggesting that the sustained expression of the base editor allows the continuous editing of multiple cytosine residues within the base-editing window, even after multiple rounds of cell division. To investigate potential off-target base editing, we

sequenced the most likely off-target site, i.e. *CCoAOMT2*, in 14 lines stably expressing the gene editing construct. No off-target gene editing was observed in any of the lines tested (Extended data 1).

Once we validated the efficiency of our base-editing construct, we aimed to obtain gene-edited plants that had only transiently expressed the genes on the T-DNA. Following co-cultivation with *Agrobacterium*, calli and plantlets were regenerated on medium supplemented with chlorsulfuron, allowing the selection of plantlets that had the envisioned mutation in at least one of the four *ALS* alleles. These plantlets expressed the genes from the T-DNA that was either transiently present in the cell or had integrated stably into the plant genome. As such, we obtained 37 chlorsulfuron-resistant calli from which one or more shoots were regenerated (Fig. 2A). Sequencing of both *ALS* genes confirmed the effectiveness of our selection procedure, as nearly all calli, except for two, yielded shoots that contained edits in at least one allele of either *ALS1* or *ALS2*. Base-editing efficiencies of up to 72.3% were achieved in the *ALS* target cytosines C13 and C14 (Fig. 2B, Table S1, Extended data 2). Interestingly, base editing at C20 in the *ALS1* and *ALS2* protospacer would yield early stop codons, making the respective genes non-functional. As the synthesis of branched chain amino acids is essential for plant development²⁵⁻²⁷, we out-selected plants with non-functional *ALS* activity (i.e. two non-functional alleles for both *ALS* genes, obtained through either base-editing activity or off-target INDEL formation via nCas9 activity). Within the *CCoAOMT1* gene, base-editing activity in the target cytosine C8 was observed up to 31%, while the overall editing efficiency (including INDELs) reached up to 52% (Fig. S2B, Table S1). We consider the occurrence of INDEL mutations in the gene of interest as an advantage, because it increases the possibility to retrieve knock-out alleles, beyond the introduction of stop-codons via a C-to-T transition. Similar to the lines that stably expressed the gene-editing construct, we further investigated the mutation profile within the *CCoAOMT1* gene at the allele level using amplicon cloning and sequencing. Notably, the frequency of chimerism for the *CCoAOMT1* gene tended to be lower as compared to the lines that stably expressed the gene editing construct (Table S2). This observation suggests that the temporary presence of the base-editing machinery could be beneficial to reduce chimerism.

To further investigate whether shoots regenerated on chlorsulfuron-containing medium were derived from transiently expressed or stably integrated gene-editing construct(s), we evaluated the presence of the *A3A* base editor sequence and the *NPTII* selectable marker gene via PCR. Strikingly, approximately half (49%) of all regenerated shoots were found to be free of *A3A* and *NPTII* sequences, suggesting they had only transiently expressed the genes on the T-DNA construct (Fig. 2A, 2C). From those lines that were resistant to chlorsulfuron and were devoid of the tested *A3A* and *NPTII* sequences, we identified ten lines (11.1, 11.2, 14.1, 14.2, 14.3, 14.4, 51, 59.1, 59.2, 59.3) harboring edits in the lignin biosynthesis *CCoAOMT1* gene, five of which were non-chimeric (Fig. 2A, 2C). To further demonstrate that we obtained lines that were completely free of transgenic DNA, we performed long-read whole-genome sequencing on non-chimeric lines 11.1 and 14.4. Analysis of the obtained raw reads indicated that no reads were found matching the T-DNA sequence. In fact, none of the reads matched with any part of the entire binary vector, indicating that no parts of the T-DNA or the backbone were integrated in the poplar genome. Taken together, using chlorsulfuron selection, we achieved transgene-free, non-chimeric plants carrying gene edits in our gene of interest at a

frequency of 7% (Fig. 2C), demonstrating the robustness of our approach. Notably, our transgene-free gene-editing efficiency aligns with reported efficiencies in literature. For instance, Huang *et al.* reported transgene-free gene-editing efficiencies ranging from as high as 42.1% in tomato (*Solanum lycopersicum*) and as low as 1.9% in pomelo (*Citrus maxima*)¹⁸. In our previous study, using particle bombardment, we achieved a maximum transgene-free gene-editing efficiency of only 3%, signifying the importance of this advancement compared to our previous results¹⁷. The base editing efficiency could be further improved by using alternative promoters like *AtU3* to drive gRNA expression. This promoter has been shown to significantly increase base editing efficiency in poplar as compared to the *AtU6* promoter we used here²⁸. Additionally, employing more advanced versions of cytosine base editors that generate lower rates of INDEL mutations provides further opportunities to enhance this system²⁹.

Given that a large fraction (49%) of the lines that survived on chlorsulfuron were transgene-free, there is no absolute need for counterselection against stable transformants. Nonetheless, we evaluated *CodA* as a negative selection marker against T-DNA integration by assessing callus development at various 5-FC concentrations. Given that the ratio between T-DNA integration and transient expression is likely genotype-dependent¹⁸, this approach could be valuable for other, more recalcitrant genotypes or tree species. We found that 5-FC concentrations above 0.5 g/L inhibited callus development in both wild-type (WT) and T-DNA-expressing poplars. However, 0.25 g/L 5-FC effectively inhibited the growth of stable T-DNA-expressing lines while allowing the development of transgene-free lines (Fig. S3).

In conclusion, we established a method for the efficient generation of gene-edited, transgene-free poplar trees by combining transient T-DNA expression and positive selection for base-editing activity. Through whole-genome sequencing, we demonstrated the absence of foreign DNA in the genome of the gene-edited plant, facilitating the public acceptance and deregulation of such gene-edited trees. In addition, we have tested a selection strategy to eliminate lines that stably incorporated the T-DNA into their genome. Although this strategy is not essential for selecting transgene-free, gene-edited lines of the model poplar (*P. tremula* x *P. alba* cv. 717-1B4), it may prove to be valuable for other genotypes or varieties.

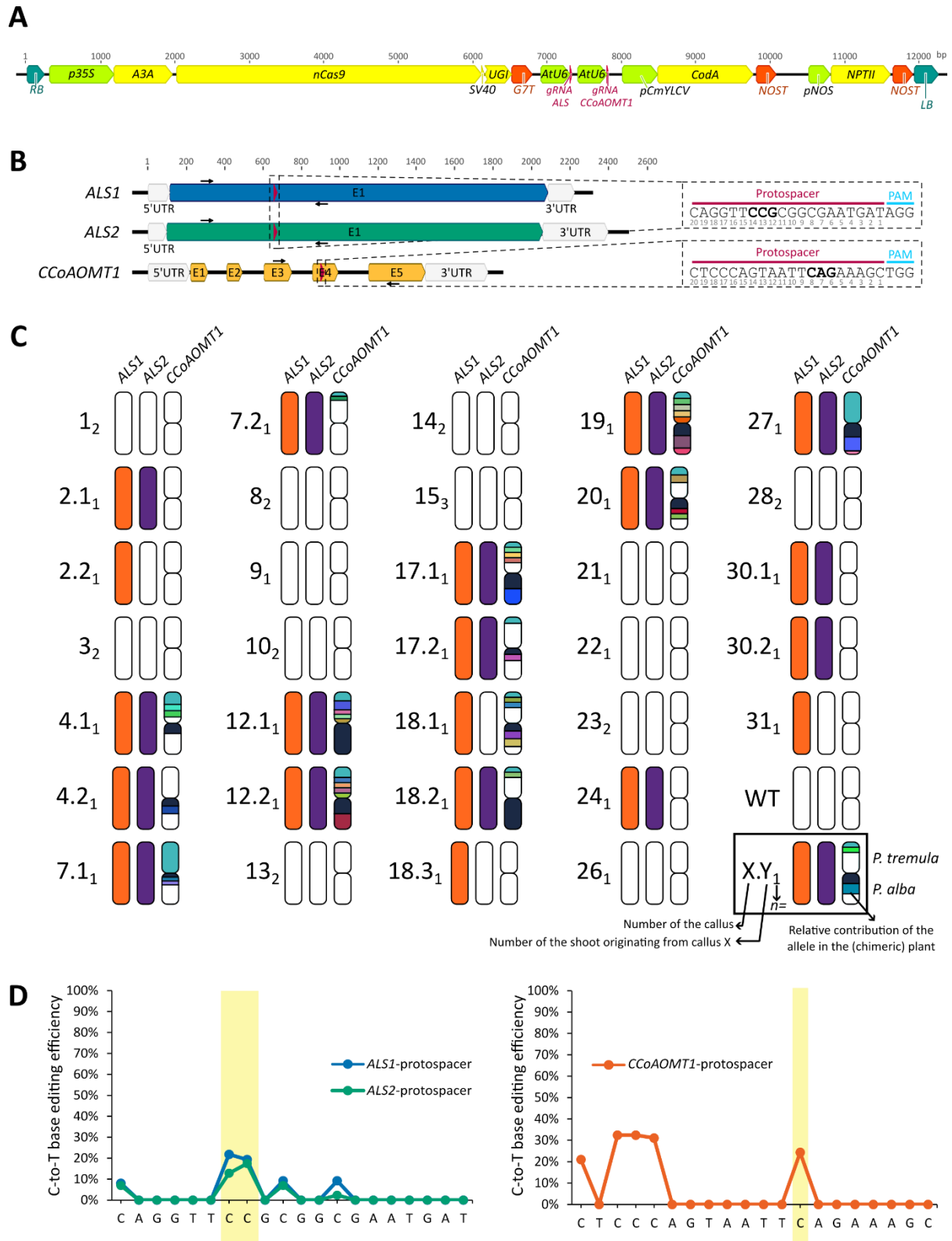
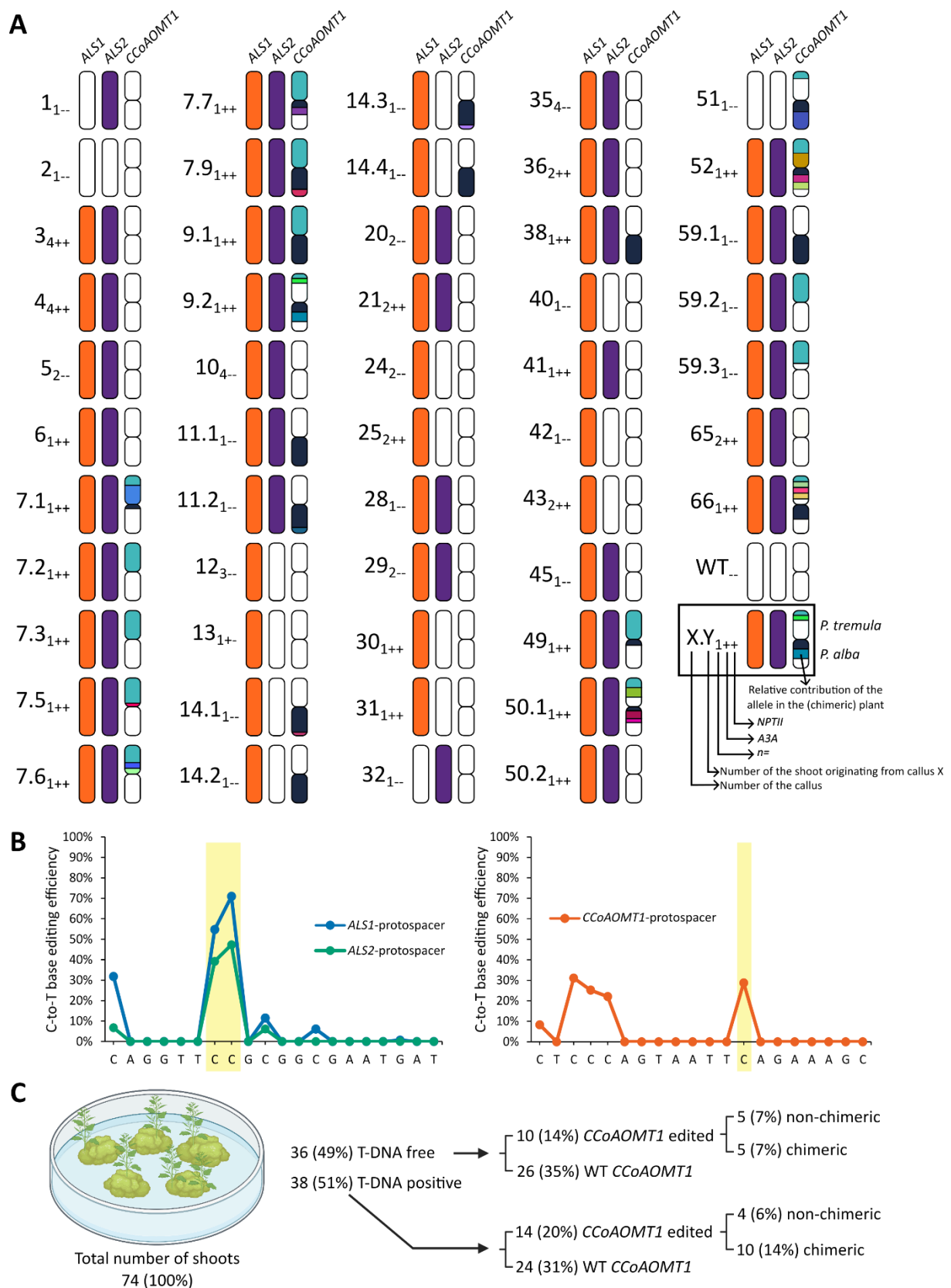


Figure 1: Multiplex base editing in poplar upon stable integration of the T-DNA. **A)** Schematic representation of the T-DNA construct created to achieve transgene-free base editing in poplar. The construct consists of a gene encoding an A3A cytosine deaminase fused to nCas9 tagged with the nuclear localization signal from Simian virus 40 (SV40) and an uracil glycosylase inhibitor (UGI), driven by a double *Cauliflower Mosaic Virus* 35S promoter (*p35S*) and terminated by the G7 terminator (*G7T*). Guide RNA (gRNA) expression is driven by the *Arabidopsis thaliana* U6 promoter (*AtU6*) and terminated by a poly-T sequence. Counterselection against T-DNA integration is achieved through the gene encoding *CodA*, which is able to convert non-toxic 5-FC to cytotoxic 5-FU. The gene is driven by the *Cestrum Yellow Leaf Curling Virus* promoter (*pCmYLCV*) and its expression is terminated by the *NOS* terminator (*NOST*). To evaluate the efficiency of the base-editing construct, an additional *NPTII*

selectable marker gene, driven by the *NOS* promoter (*pNOS*) and terminated by the *NOST*, was incorporated in the construct to allow for the selection of stable transformants on kanamycin. The complete construct was cloned between the left (LB) and right border (RB) sequence into a binary vector. **B)** Gene structure of the target genes, *ACETOLACTATE SYNTHASE 1* (*ALS1*) and *2* (*ALS2*), and *CAFFEYL-CoA O-METHYLTRANSFERASE1* (*CCoAOMT1*). The protospacer is highlighted in magenta, the protospacer adjacent motif (PAM) in light blue. The target codon is represented in bold. Sequencing primers are represented by black arrows. Exons are indicated by the letter E. **C)** Schematic overview of the regenerated transgenic population, consisting of 43 shoots in total. The numbers aside the bar represent the callus number (X) from which a particular shoot (Y) was derived. The number of identical shoots (*n*) is represented as a subscript. Rounded rectangles represent the different target genes, *ALS1*, *ALS2* and *CCoAOMT1*, respectively. Base editing in *ALS1* and *ALS2* was evaluated by sequencing both alleles simultaneously, with colored bars indicating successful editing in at least one allele. For *CCoAOMT1*, we investigated base editing at the allelic level through the cloning of PCR amplicons followed by sequencing, allowing us to discriminate between multiple allelic variants in the *P. tremula* and *P. alba* alleles. The size of the colored bars corresponds to the relative contribution of each detected allelic variant, while white bars represent the non-edited WT sequence. Different colors represent different allelic variants within a given genotype. **D)** C-to-T base editing efficiency along the protospacer sequence of poplar lines stably expressing the T-DNA (as shown in panel C). A summary of all editing efficiencies including C-to-D (D = A, G or T) and INDEL mutations is shown in Table S1. Target cytosines (C) are highlighted in yellow.



allowing to discriminate between multiple allelic variants in the *P. tremula* and *P. alba* allele. The size of the colored bars corresponds to the relative contribution of each detected allelic variant, while white bars represent the non-edited WT sequence. Different colors represent different allelic variants within a given genotype. **B)** Base-editing efficiency along the protospacer sequence of lines selected on chlorsulfuron (as shown in panel A). The efficiency of C-to-T transitions is depicted by a line graph, while the efficiency of C-to-D (D = A, G or T) editing frequencies and INDEL mutations are given in Table S1. Target cytosines (C) are highlighted in yellow. **C)** Overall efficiency of T-DNA-free editing in *CCoAOMT1* following shoot regeneration on medium supplemented with chlorsulfuron. Percentages between brackets are relative to the total number of shoots harvested.

Methods

Plant material, transformation, and regeneration

Poplar hybrid (*Populus tremula* x *P. alba* cv. INRA 717-B4) transformation was performed as previously reported¹⁹ with slight modifications. Following co-cultivation and callus formation, shoots were regenerated on shoot-inducing medium [M3,¹⁹]. For the stable transformation, M3 medium was supplemented with 100 mg/L kanamycin sulfate. To facilitate the regeneration of gene-edited shoots that either stably or transiently expressed the T-DNA, the M3 medium was supplemented with 40 µg/L chlorsulfuron. In both experiments shoot regeneration continued up to 6 months, with the medium being refreshed every 3 weeks to circumvent degradation of the antibiotic or herbicide. For the regeneration of shoots on medium containing either kanamycin or chlorsulfuron, 960 and 1570 poplar internodes were used for transformation, respectively.

5-FC counterselection

Callus formation was induced in lines that stably expressed the T-DNA (lines 3, 4.1, and 7.1), lines that had transiently expressed the T-DNA (lines 10, 11.1, 12 and 14.1), and WT by cultivating wounded internodes on callus inducing medium [M1,¹⁹], either supplemented with 5-FC or not. To investigate the efficiency of the *CodA* negative selection marker, callus development was followed for 4 weeks on M1 medium that was either supplemented with the following concentrations 5-FC (0 g/L, 0.25 g/L, 0.50 g/L, 0.75 g/L, and 1 g/L).

Construct design

The gRNA sequences were designed to induce a specific point mutation in the *ALS* genes conferring herbicide resistance, and to induce an early stop codon in *CCoAOMT1* gene. Putative off-targets were evaluated via the CRISPOR online tool³⁰. Apart for a sequence in the exon of *CCoAOMT2* that had 3 mismatches with the *CCoAOMT1* gRNA, other sequences in the *Populus tremula* x *alba* genome had at least four mismatches with the designed gRNAs, which makes it unlikely that these could act as off-target sites. The gRNA Golden Gate gRNA entry vectors were constructed as described in Decaestecker et al.³¹. In summary, DNA oligos were ordered from Integrated DNA Technologies (IDT) and subsequently annealed and inserted into gRNA entry vectors, containing an Arabidopsis U6-26 Pol III promoter, a *ccdB*-CmR selection cassette, a Cas9 gRNA scaffold, and a poly-T sequence as termination signal, through Golden Gate (GG) reaction using BbsI-HF (New England Biolabs). Three destination vectors, all containing *ccdB*-CmR selection cassettes, were assembled, using Golden Gate reactions with BsaI-HFv2 (NEB). All three destination vectors come with classic Golden Gate A-G cloning sites, but also in-house 'golden Gibson' specific cloning sites. In destination vector 1, using basic destination vector 12-73 of the PSB vector collection (<https://vectorvault.vib.be/>), the following elements were assembled between the GG A-G sites: 2x35S promoter (vector 11_78 from the PSB vector collection), APOBEC3a [intron variant²⁰], Cas9D10³², UGI-NLS²⁰ and the G7 terminator. In destination vector 2, assembled into the basic destination vector 12-74 of the PSB vector collection, two main elements were included between the GG A-G sites: the two gRNAs with the scaffold and their respective promoter and terminator sequences. Destination vector 3, using basic destination vector 12-75 of the PSB vector collection, compiled the CmYLCV (*Cestrum Yellow Leaf Curling Virus*) promoter, the *CodA*

gene, which was synthesized and ordered from GenScript, and the *NOS* terminator. The final expression vector was assembled with the in-house 'golden Gibson' protocol, which uses equimolar concentrations (100 ng/μl) of the vectors, mixed with 1 μL *I-SceI* restriction enzyme (New England Biolabs) and 1 μL CutSmart buffer (10x) (NEB) in a total volume of 7 μL. This mixture was incubated for 2 h at 37°C and then for 25 minutes at 65°C. Subsequently, 5 μL of the reaction was mixed with 5 μL 2x NEBuilder Hifi DNA Assembly Mix and incubated for 1 h at 50°C.

All cloning products were transformed into *ccdB*-sensitive DH5α *Escherichia coli* via heat-shock. *E. coli* was cultured on lysogeny broth medium containing 100 μg/mL carbenicillin or 50 μg/mL kanamycin sulphate. Colonies were verified through colony-touch PCR, restriction digestion, Sanger sequencing (Eurofins Genomics, Germany), and/or whole plasmid sequencing (Plasmidsaurus, USA). Colony-touch PCRs were carried out with ALL-in RedTaq Mastermix 2x (highQu GmbH), using 10 μM primer solutions. For PCRs on the gRNA entry vector, primer 1 and primer 2 were employed. For destination vector 1, primer 3 and primer 4 were used. Destination vector 2 was verified using primer 5 and primer 6 (Table S3) and destination vector 3 was verified with primer 7 and 8. The final vector was confirmed with primers 3 to 8. PCR products were evaluated on a 1.2% agarose gel (100 V, 30 min) with a molecular weight marker (Promega, Benchtop 1kb ladder). Restriction digests were carried out using NEB restriction enzymes and their respective buffers. The reactions were carried out in a total volume of 20 μL, with 10 μL of 100 ng/μL plasmid, following the NEB protocol. The results were evaluated on a 2% agarose gel (100 V, 45-60 min).

Molecular analyses

Genomic DNA was isolated from poplar leaves using the Wizard® Genomic DNA Purification Kit (Promega) according to manufacturing procedures. Purified DNA samples were used for PCR amplification using ALLin™ Taq DNA Polymerase (highQu), 10 μM primer solution, and 1 μL DNA template in MQ-H₂O in a T100 Thermal Cycler (Bio-Rad). The following gene-specific primer combinations were used: primer 9 and primer 10 for *ALS1*, primer 11 and primer 12 for *ALS2*, primer 13 and primer 14 for *CCoAOMT1*, primer 15 and primer 16 for *NPTII* and primer 17 and primer 18 for *A3A* (Table S3). The presence of the PCR amplicons was evaluated on a 1.2% agarose gel (100 V, 30 min) with a molecular weight marker (Promega, Benchtop 1kb ladder). *ALS1*, *ALS2*, and *CCoAOMT1* amplicons were purified using the HighPrep™ (MAGBIO AC-60050) post-PCR clean-up prior to sequence analyses. Purified amplicons were sent for Sanger sequencing (Eurofins Genomics, Germany) and the obtained sequencing chromatograms were analyzed for gene editing using TIDE³³, ICE (ICE Analysis. 2019. v3.0. Synthego; 2023), EditR³⁴, DECODR³⁵, and Geneious Prime 2024.0.5. Edited *CCoAOMT1* amplicons were further analyzed to investigate editing at the allelic level through amplicon cloning using pGEM®-T Vector System (Promega) according to manufacturing procedures. Cloned amplicons were transfected into *E.coli* during a 1 min heat shock and the transfected plasmids were harvested the consecutive day using the GeneJET Plasmid Miniprep Kit (Thermo Scientific™). Up to ten cloned amplicons were sent for Sanger sequencing (Eurofins Genomics, Germany) using the *pUC/M13* forward sequencing primer 19 and the purified plasmid as a template. Base editing and general gene editing efficiencies (i.e., including INDELs) were calculated in Excel as the number of mutated alleles compared to the number of WT alleles in the population (e.g., for 100 heterozygous plant with a

mutated *P. tremula* allele and a WT *P. alba* allele, the editing frequency is 50%. Similarly, for a population of 100 plants where 50 plants are biallelically mutated, and 50 plants have two WT alleles, the editing frequency is 50%).

Whole genome sequencing and data analysis

Two grams of leaf material was harvested from the first fully developed leaf from 2-m tall poplars. The midvein was removed before flash-freezing the material in liquid nitrogen. High-molecular weight genomic DNA was extracted according to the following procedure: frozen leaf material was ground into a fine powder using a mortar and pestle in liquid nitrogen. Ground material was transferred and homogenized in 10 mL Carson lysis buffer (100 mM Tris-HCl at pH 9.5, 2% CTAB, 1.4 M NaCl, 1% PEG 8000, 20 mM EDTA) supplemented with 70 μ L Proteinase K and 25 μ L β -mercaptoethanol and briefly vortexed before incubation at 74°C for 1 h in a hot water bath. The samples were briefly vortexed every 5 minutes for 1 second. Upon incubation, the samples were allowed to cool down to room temperature before adding 7 μ L RNaseA and incubated for 20 minutes at 37°C. Subsequently, one volume of chloroform/isoamylalcohol (24:1) was added and the samples were placed in a tube rotator for 15 minutes to precipitate the DNA before centrifugation at 2 755 x *g* for 15 minutes at 18°C. The upper aqueous phase was transferred to a fresh 50 mL screw-cap tube before adding one volume of distilled water and adjusting the pH to 7.0 with HCl. Subsequently, the DNA was further purified using the QIAGEN Genomic-tip 100/G procedure described in the QIAGEN Genomic DNA Handbook (06/2015). The extracted high-molecular weight genomic DNA was sent to the VIB-Neuromics Support Facility for library preparation, sequencing and base calling. All collected raw reads were deposited at NCBI SRA with the accession number: PRJNA1200679. An average read depth of 31X and 30X was obtained for line 11.1 and 14.4, respectively. All raw reads were collected into a BLAST database using the nblast module (Nucleotide-Nucleotide BLAST 2.6.0+)³⁶. Integration of DNA fragments in the plant's genomic DNA was investigated through BLAST analysis by using the complete sequence of the base editor plasmid as a query against the pre-built target database. Additionally, raw reads were aligned to the base editor plasmid sequence using Minimap2.

Acknowledgments

We would like to thank Tuur Neyrinck and Jonathan Smet for their help with experiments, Mansour Karimi, Christophe Gaillochet, and Thomas Jacobs for their advice on the cloning strategy, and Annick Bleys for critically reading the manuscript and preparing it for submission. The Graphical abstract, Fig 2C, and Fig. S1 were created using BioRender.

Funding

This work was supported by the Energy Transition Fund (project ADV_BIO for a predoctoral fellowship to C.A. and project AD-LIBIO to L.H. for a postdoctoral fellowship). L.H. is also indebted to the Research Foundation – Flanders (FWO project number G011620N), and W.B. to the interuniversity iBOF project NextBioRef and the Advanced ERC grant POPMET.

Author contributions

L.H., C.A., J.V.D., R.V., and W.B. conceptualized and designed the experiments. L.H., C.A., and J.V.D. performed experiments. L.H. performed bioinformatics analysis. L.H. wrote the manuscript with input from all co-authors.

Competing interest

The authors declare that they have no competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Boerjan, W. & Strauss, S. H. Social and biological innovations are essential to deliver transformative forest biotechnologies. *New Phytol.* **243**, 526-536 (2024).
2. Anders, C., Hoengenaert, L. & Boerjan, W. Accelerating wood domestication in forest trees through genome editing: Advances and prospects. *Curr. Opin. Plant Biol.* **71**, 102329 (2023).
3. Goralogia, G. S., Redick, T. P. & Strauss, S. H. Gene editing in tree and clonal crops: progress and challenges. *In Vitro Cell. Dev. Biol. - Plant* **57**, 683-699 (2021).
4. Dalla Costa, L. *et al.* Strategies to produce T-DNA free CRISPRed fruit trees via *Agrobacterium tumefaciens* stable gene transfer. *Sci. Rep.* **10**, 20155 (2020).
5. Pompili, V., Dalla Costa, L., Piazza, S., Pindo, M. & Malnoy, M. Reduced fire blight susceptibility in apple cultivars using a high-efficiency CRISPR/Cas9-FLP/FRT-based gene editing system. *Plant Biotechnol. J.* **18**, 845-858 (2020).
6. Charrier, A. *et al.* Efficient targeted mutagenesis in apple and first time edition of pear using the CRISPR-Cas9 system. *Front. Plant Sci.* **10**, 40 (2019).
7. Choi, S. H. *et al.* TSA promotes CRISPR/Cas9 editing efficiency and expression of cell division-related genes from plant protoplasts. *Int. J. Mol. Sci.* **22**, 7817 (2021).
8. Andersson, M. *et al.* Efficient targeted multiallelic mutagenesis in tetraploid potato (*Solanum tuberosum*) by transient CRISPR-Cas9 expression in protoplasts. *Plant Cell Rep.* **36**, 117-128 (2017).
9. Lin, C. S. *et al.* Application of protoplast technology to CRISPR/Cas9 mutagenesis: from single-cell mutation detection to mutant plant regeneration. *Plant Biotechnol. J.* **16**, 1295-1310 (2018).
10. Malnoy, M. *et al.* DNA-free genetically edited grapevine and apple protoplast using CRISPR/Cas9 ribonucleoproteins. *Front. Plant Sci.* **7**, 1904 (2016).
11. Murovec, J., Gucek, K., Bohanec, B., Avbelj, M. & Jerala, R. DNA-free genome editing of *Brassica oleracea* and *B. rapa* protoplasts using CRISPR-Cas9 ribonucleoprotein complexes. *Front. Plant Sci.* **9**, 1594 (2018).
12. Woo, J. W. *et al.* DNA-free genome editing in plants with preassembled CRISPR-Cas9 ribonucleoproteins. *Nat. Biotechnol.* **33**, 1162-1164 (2015).
13. Fossi, M., Amundson, K., Kuppu, S., Britt, A. & Comai, L. Regeneration of *Solanum tuberosum* plants from protoplasts induces widespread genome instability. *Plant Physiol.* **180**, 78-86 (2019).
14. Serres, R., Ostry, M., McCown, B. & Skilling, D. Somaclonal variation in *Populus* hybrids regenerated from protoplast culture. in *Woody Plant Biotechnology NATO ASI Series* 59-61 (Springer, New York, 1991).
15. Zhang, Y. *et al.* Efficient and transgene-free genome editing in wheat through transient expression of CRISPR/Cas9 DNA or RNA. *Nat. Commun.* **7**, 12617 (2016).
16. Svitashv, S., Schwartz, C., Lenderts, B., Young, J. K. & Mark Cigan, A. Genome editing in maize directed by CRISPR–Cas9 ribonucleoprotein complexes. *Nat. Commun.* **7**, 13274 (2016).
17. Hoengenaert, L., Van Doorselaere, J., Vanholme, R. & Boerjan, W. Microparticle-mediated CRISPR DNA delivery for genome editing in poplar. *Front. Plant Sci.* **14**, 1286663 (2023).
18. Huang, X. *et al.* Transgene-free genome editing of vegetatively propagated and perennial plant species in the T0 generation via a co-editing strategy. *Nat. Plants* **9**, 1591-1597 (2023).
19. Leplé, J. C., Brasileiro, A. C. M., Michel, M. F., Delmotte, F. & Jouanin, L. Transgenic poplars: expression of chimeric genes using four different constructs. *Plant Cell Rep.* **11**, 137-141 (1992).
20. Gaillochet, C. *et al.* Systematic optimization of Cas12a base editors in wheat and maize using the ITER platform. *Genome Biol.* **24**, 6 (2023).
21. Tranel, P. J. & Wright, T. R. Resistance of weeds to ALS-inhibiting herbicides: what have we learned? *Weed Sci.* **50**, 700-712 (2002).

22. Osakabe, K., Nishizawa-Yokoi, A., Ohtsuki, N., Osakabe, Y. & Toki, S. A mutated cytosine deaminase gene, *codA* (D314A), as an efficient negative selection marker for gene targeting in rice. *Plant Cell Physiol.* **55**, 658-665 (2014).
23. Stougaard, J. Substrate-dependent negative selection in plants using a bacterial cytosine deaminase gene. *Plant J.* **3**, 755-761 (1993).
24. Gehrke, J. M. *et al.* An APOBEC3A-Cas9 base editor with minimized bystander and off-target activities. *Nat. Biotechnol.* **36**, 977-982 (2018).
25. Sathasivan, K., Haughn, G. W. & Murai, N. Molecular basis of imidazolinone herbicide resistance in *Arabidopsis thaliana* var Columbia. *Plant Physiol.* **97**, 1044-1050 (1991).
26. Zabalza, A., Zulet, A., Gil-Monreal, M., Igal, M. & Royuela, M. Branched-chain amino acid biosynthesis inhibitors: herbicide efficacy is associated with an induced carbon-nitrogen imbalance. *J. Plant Physiol.* **170**, 814-821 (2013).
27. Dezfulian, M. H. *et al.* Acetolactate synthase regulatory subunits play divergent and overlapping roles in branched-chain amino acid synthesis and Arabidopsis development. *BMC Plant Biol.* **17**, 71 (2017).
28. Li, G., Sretenovic, S., Eisenstein, E., Coleman, G. & Qi, Y. Highly efficient C-to-T and A-to-G base editing in a *Populus* hybrid. *Plant Biotechnol Journal* **19**, 1086-1088 (2021).
29. Ren, Q. *et al.* Improved plant cytosine base editors with high editing activity, purity, and specificity. *Plant Biotechnol Journal* **19**, 2052-2068 (2021).
30. Concordet, J. P. & Haeussler, M. CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Res* **46**, W242-W245 (2018).
31. Decaestecker, W. *et al.* CRISPR-TSKO: a technique for efficient mutagenesis in specific cell types, tissues, or organs in Arabidopsis. *Plant Cell* **31**, 2868-2887 (2019).
32. Hu, Z. *et al.* Genome editing-based engineering of CESA3 dual cellulose-inhibitor-resistant plants. *Plant Physiol.* **180**, 827-836 (2019).
33. Brinkman, E. K., Chen, T., Amendola, M. & van Steensel, B. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res.* **42**, e168 (2014).
34. Kluesner, M. G. *et al.* EditR: A method to quantify base editing from Sanger sequencing. *CRISPR J.* **1**, 239-250 (2018).
35. Bloh, K. *et al.* Deconvolution of Complex DNA Repair (DECODR): Establishing a novel deconvolution algorithm for comprehensive analysis of CRISPR-edited Sanger sequencing data. *CRISPR J.* **4**, 120-131 (2021).
36. Zhang, Z., Schwartz, S., Wagner, L. & Miller, W. A greedy algorithm for aligning DNA sequences. *J. Comput. Biol.* **7**, 203-214 (2000).