Lawrence Berkeley National Laboratory
Joint Genome Institute

Title
Multiplex knockout of trichome-regulating MYB duplicates in hybrid poplar using a single gRNA

Authors
Bewg, William P
Harding, Scott A
Engle, Nancy L
et al.

Plant Physiology®

Multiplex knockout of trichome-regulating MYB duplicates in hybrid poplar using a single gRNA

William P. Bewg [1,2,3] Scott A. Harding [1,2,3] Nancy L. Engle [4] Brajesh N. Vaidya,[5,‡]
Ran Zhou [1,2,3] Jacob Reeves,[6,§] Thomas W. Horn [6] Nirmal Joshee [5] Jerry W. Jenkins [7,8]
Shengqiang Shu [8] Kerrie W. Barry [8] Yuko Yoshinaga [8] Jane Grimwood [7,8]
Robert J. Schmitz [2] Jeremy Schmutz [7,8] Timothy J. Tschaplinski [4] and
Chung-Jui Tsai [1,2,3,*,†]

1   School of Forestry and Natural Resources, University of Georgia, Athens, Georgia 30602, USA
2   Department of Genetics, University of Georgia, Athens, Georgia 30602, USA
3   Department of Plant Biology, University of Georgia, Athens, Georgia 30602, USA
4   Oak Ridge National Laboratory, Oak Ridge, Tennessee 37830, USA
5   Department of Plant Science, Fort Valley State University, Georgia, 31030, USA
6   Department of Computer Science, University of Georgia, Athens, Georgia 30602, USA
7   HudsonAlpha Institute for Biotechnology, Huntsville, Alabama 35806, USA
8   U.S. Department of Energy Joint Genome Institute, Berkeley, California 94720, USA

*Author for correspondence: cjtsai@uga.edu
†Senior author.
‡Present address: GreenVenus, LLC, Wimauma, Florida 33598, USA
§Present address: SoundHound Inc., Boulder, Colorado 80302, USA

Abstract

As the focus for CRISPR/Cas-edited plants moves from proof-of-concept to real-world applications, precise gene manipulation will increasingly require concurrent multiplex editing for polygenic traits. A common approach for editing across multiple sites is to design one guide RNA (gRNA) per target; however, this complicates construct assembly and increases the possibility of off-target mutations. In this study, we utilized one gRNA to target *MYB186*, a known positive trichome regulator, as well as its paralogs *MYB138* and *MYB38* at a consensus site for mutagenesis in hybrid poplar (*Populus tremula* × *P. alba* INRA 717-1B4). Unexpected duplications of *MYB186* and *MYB138* resulted in eight alleles for the three targeted genes in the hybrid poplar. Deep sequencing and polymerase chain reaction analyses confirmed editing across all eight targets in nearly all of the resultant glabrous mutants, ranging from small indels to large genomic dropouts, with no off-target activity detected at four potential sites. This highlights the effectiveness of a single gRNA targeting conserved exonic regions for multiplex editing. Additionally, cuticular wax and whole-leaf analyses showed a complete absence of triterpenes in the trichomeless mutants, hinting at a previously undescribed role for the nonglandular trichomes of poplar.

Open Access

Introduction

Clustered regularly interspaced short palindromic repeats (CRISPR) technology has been adopted for plant genome editing in an increasing number of species for both basic and applied research (Bewg et al., 2018; Chen et al., 2019; Nasti and Voytas, 2021). The power of CRISPR is due in part to its simplicity with just two core components (Jinek et al., 2012): a nuclear-localized endonuclease, such as CRISPR-associated Cas9 that works universally across all domains of life, and a synthetic guide RNA (gRNA) that is customizable and scalable for sequence-specific targeting. With its proven precision and efficiency (Li et al., 2013; Nekrasov et al., 2013; Shan et al., 2013; Endo et al., 2016) and given the polygenic nature of many agronomic traits, there is growing interest in targeting multiple loci for simultaneous CRISPR editing to aid gene function investigation and/or trait engineering (Armario Najera et al., 2019).

Multiplex editing usually involves the coexpression of multiple gRNAs. For the classic CRISPR/Cas9 system, this has been demonstrated using individual gRNA cassettes each driven by a separate RNA polymerase III (Pol III) promoter (Xing et al., 2014; Lowder et al., 2015; Ma et al., 2015). Alternatively, multiple gRNAs can be expressed in tandem with tRNAs as a single polycistronic transcript and processed into individual gRNAs using endogenous tRNA processing machinery (Xie et al., 2015). Polycistronic gRNA transcripts have also been engineered with built-in RNA cleavage sites for processing by ribozymes or the CRISPR-associated endoribonuclease Csy4 (Qi et al., 2012; Gao and Zhao, 2014; Tang et al., 2016, 2019; Čermák et al., 2017). In several cases, functional gRNAs were generated from a single transcriptional unit of Cas9 fused with an artificial gRNA array without specific flanking sequences (Mikami et al., 2017; Wang et al., 2018). It has been reported that up to eight gRNAs have been successfully deployed for multiplex editing (Ma et al., 2015; Xie et al., 2015; Čermák et al., 2017).

An understudied approach is the use of a single gRNA to target homologous sequences at discrete loci. The capability was showcased by effective inactivation of all 62 copies of porcine endogenous retroviruses in an immortalized pig cell line using two gRNAs to target a highly conserved region of the polymerase (pol) gene (Yang et al., 2015). Besides parasitic elements, a single consensus gRNA has been used to edit paralogs derived from various gene duplication events in soybean (Glycine max) and sorghum (Sorghum bicolor) (Jacobs et al., 2015; Li et al., 2018) or homoeologs in polyploid wheat (Triticum aestivum L.) and oilseed rape (Brassica napus) (Braatz et al., 2017; Zhang et al., 2017). Multigene knockout (KO) using a single gRNA has also been reported in rice (O. sativa L. cv. Nipponbare; Endo et al., 2015); however, this was accomplished by off-target mutations that are undesirable in genome editing. Multiplex targeting of duplicated genes is especially important for the investigation of functional redundancy in plant genomes that are shaped by whole-genome, segmental, tandem, and/or transposon-mediated duplications (Flagel and Wendel, 2009; Panchy et al., 2016). Depending on the duplication age and subsequent selection constraints, sequence similarity can be very high among duplicates, enabling the identification of consensus target sites for multiplex editing by a single gRNA. This approach greatly simplifies construct design and assembly, reduces off-target potential that increases with the number of gRNAs (McCarty et al., 2020), and can be bundled with other multi-gRNA editing strategies discussed above for higher order multiplex targeting of distinct gene families.

The present study explored the utility of a single gRNA for multiplex editing in hybrid poplar (Populus tremula × P. alba INRA 717-1B4, hereon referred to as 717), an outcrossing woody perennial. As an interspecific hybrid, the 717 genome is highly heterozygous which presents additional challenges to gRNA design and edit outcome determination (Xue et al., 2015). Using trichomes as visual reporter, we targeted a known positive regulator of trichome development, PtaMYB186 (Plett et al., 2010), and its paralogs PtaMYB138 and PtaMYB38 for KO. We show that a single gRNA with single-nucleotide polymorphism (SNP)-aware design is effective for multiplex KO of paralogous genes and robust against copy number variations in a hybrid genome with an unexpected tandem duplication in one of its sub-genomes. We employed multiple approaches to address the analytical challenge of discriminating among highly similar target sites to discern mutations that ranged from small indels to large genomic dropouts. Finally, analysis of the resultant trichomeless mutants revealed a complete absence of triterpenes, and implicated a role for poplar trichomes in triterpene accrual.

Results

Multiplex CRISPR/Cas9 editing of trichome-regulating MYBs

The known positive regulator of trichome initiation PtaMYB186 (Plett et al., 2010) corresponds to gene model Potri.008G089200 in the P. trichocarpa v3.1 genome. It belongs to clade 15 of the R2R3-MYB protein family tree (Wilkins et al., 2009), which is expanded in poplar and contains three additional members, MYB138, MYB38, and MYB83, with as yet unclear functions. The four clade 15 members are derived from multiple duplication events, based on whole paranome K_S (synonymous distance) distribution and gene collinearity analyses using the wgd program (Zwaenepoel and Van de Peer, 2018). These include an ancient (gamma) whole-genome duplication (MYB186 and MYB83, K_S = 3.76), a Salicoid duplication (MYB186/MYB138 and MYB38, K_S = 0.21–0.22), and a tandem duplication (MYB186 and MYB138, K_S = 0.0001; Figure 1). MYB186, MYB138, and MYB38 share higher levels (88%–96%) of amino acid sequence similarity than with MYB83 (55%–57%). To ascertain these MYB involvement in trichome development, we mined RNA-seq data from different stages of 717 leaf development. Transcript levels of MYB186, MYB138, and MYB38 were the highest in newly emerged leaves (Leaf Plastochron Index [LPI]-1) when trichome initiation occurs (Plett et al., 2010), but quickly declined
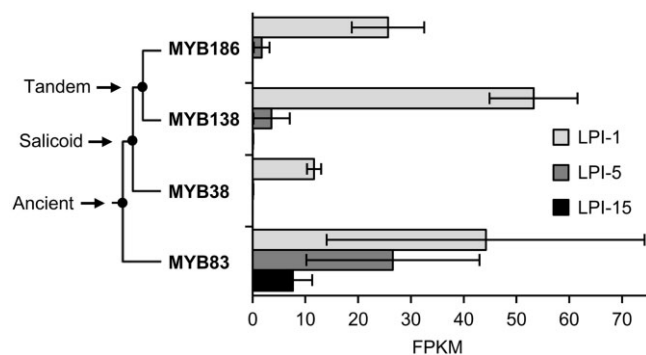
Figure 1 Expression of clade 15 *MYB* transcription factors during *Populus* leaf maturation. A simplified phylogenetic tree is shown with duplication history noted on the left. Data are mean ± SD of n = 3. LPI, leaf plastochron index; FPKM, fragments per kilobase of transcript per million mapped reads; MYB186, Potri.008G089200; MYB138, Potri.008G089700; MYB38, Potri.010G165700; and MYB83, Potri.017G086300.

thereafter in expanding (LPI-5) and mature (LPI-15) leaves (Figure 1). In contrast, *MYB83* transcripts were detected throughout leaf maturation (Figure 1), weakening support for its potential involvement in trichome development.

We designed a single gRNA to target a conserved region in exon two of *MYB186*, *MYB138*, and *MYB38* (Figure 2A) based on the *P. trichocarpa* v3.1 reference genome and cross-checked using the 717 variant database (Xue et al., 2015; Zhou et al., 2015) to assure the gRNA target sites were SNP free in 717. Two CRISPR/Cas9 constructs were generated (see "Materials and methods"); the first erroneously omitted a guanine between the gRNA spacer and the scaffold sequences (referred to as ΔG, Figure 2B), which was corrected in the second construct (Figure 2A). Both constructs were used for 717 transformation in order to learn whether ΔG would affect CRISPR/Cas9 editing. In total, 28 independent events generated from the ΔG construct were all phenotypically indistinguishable from the wild-type (WT) and Cas9-only controls (Figure 2, C–J). In contrast, 37 independent events generated from the correct KO construct were glabrous (Figure 2, N–R), unlike WT and control leaves with dense trichomes especially on the abaxial surface. The only exception is one glabrous-like event (KO-27), which had a greatly reduced trichome density across all shoot tissues (leaf, petiole, and stem) independent of age (Figure 2, K–M). Scanning electron microscopic (SEM) imaging revealed no trichome initiation or development on the abaxial surface of the glabrous mutant leaves (Figure 2Q). Epidermal cell morphology of young leaves from tissue cultured plants did not differ between control and mutant genotypes on either their abaxial (Figure 2, F and N) or adaxial surfaces (Figure 2, J and R). These results are consistent with roles for MYB186 (Plett et al., 2010) and its paralogs MYB138 and MYB38 in trichome initiation and development.

## Mutation spectrum of duplicated 717 *MYB* alleles
A random selection of 30 glabrous events, 28 ΔG events, 2 Cas9-only events, and 4 WT plants were subject to amplicon

deep-sequencing using consensus primers for *MYB186*, *MYB138*, and *MYB38*. Initial analysis by AGEseq (Xue and Tsai, 2015) showed numerous chimeric edits (mix of edited and unedited sequences at a given site) not observed in other CRISPR/Cas9-edited 717 transgenics in our experience (Zhou et al., 2015; Bewg et al., 2018; Tsai et al., 2020). *De novo* assembly of amplicon reads from control samples revealed seven distinct sequences, more than the expected six alleles of the three target genes. Blast search against the preliminary 717 genome assemblies by the Joint Genome Institute uncovered an unexpected copy number variation in 717 relative to the *P. trichocarpa* reference genome. The region containing paralogous *MYB186* and *MYB138* on Chromosome (Chr) 8 is found as a tandem duplicate in one of the 717 subgenomes (Figure 3A). This results in three alleles each for *MYB186* and *MYB138* (two on the main subgenome [Chr8m] and one on the alternative subgenome [Chr8a]) and two alleles for *MYB38* on Chr10 (Chr10m and Chr10a, Figure 3A). Two of the eight alleles were identical in the (original) amplicon region, explaining the seven distinct sequences we recovered from *de novo* assembly. Based on the 717 assemblies, we redesigned primers to ensure the amplicons span allele-specific SNP(s) to aid mutation pattern determination of the eight alleles.

Amplicon-sequencing showed no editing in the 28 ΔG events, except one (ΔG-24) with a 9-bp deletion at one of the eight target sites (Supplemental Data Set S1). This translates into a mutation rate of 0.45% (1 out of 224 potential target sites), which suggests a negative effect of the ΔG on CRISPR/Cas9 function (hereafter, the ΔG plants were treated as transformation controls). In contrast, we confirmed successful editing across the eight alleles in all glabrous mutants except KO-27 (Figure 3, B and C; Supplemental Data Set S1). This event showed six edited and two WT (unedited) alleles, consistent with trichome detection in this line (Figure 2, K–M). In aggregate, small insertions and deletions (indels) were the predominant edits at all sites (Figure 3, B–D), with frameshift deletions of 1 bp (-1), 2 bp (-2), and 4 bp (-4) accounting for over three-quarters of the indel mutations (Figure 3D). In-frame deletions (-3 or -6) accounted for 10% of indels and were detected in 14 events, including KO-27 (Figure 3, B–D). These in-frame mutations are unlikely functional because the gRNA target site is located within the third α-helix of the R2 domain critical for MYB–DNA interaction (Wang et al., 2020), and because 13 of the events with in-frame mutations are glabrous. We therefore conclude that all small indels we detected are null mutations.

Large genomic dropouts between tandem genes
The vast majority (80%) of the sequenced mutants also harbored potentially large deletions as evidenced by the dearth of mapped amplicon reads at the target sites, referred to as no-amplification (NA) alleles (Figure 3, B–D; Supplemental Data Set S1). The NA frequencies differed by chromosome position and were positively associated with copy number, being highest at the Chr8m sites (four tandem copies),
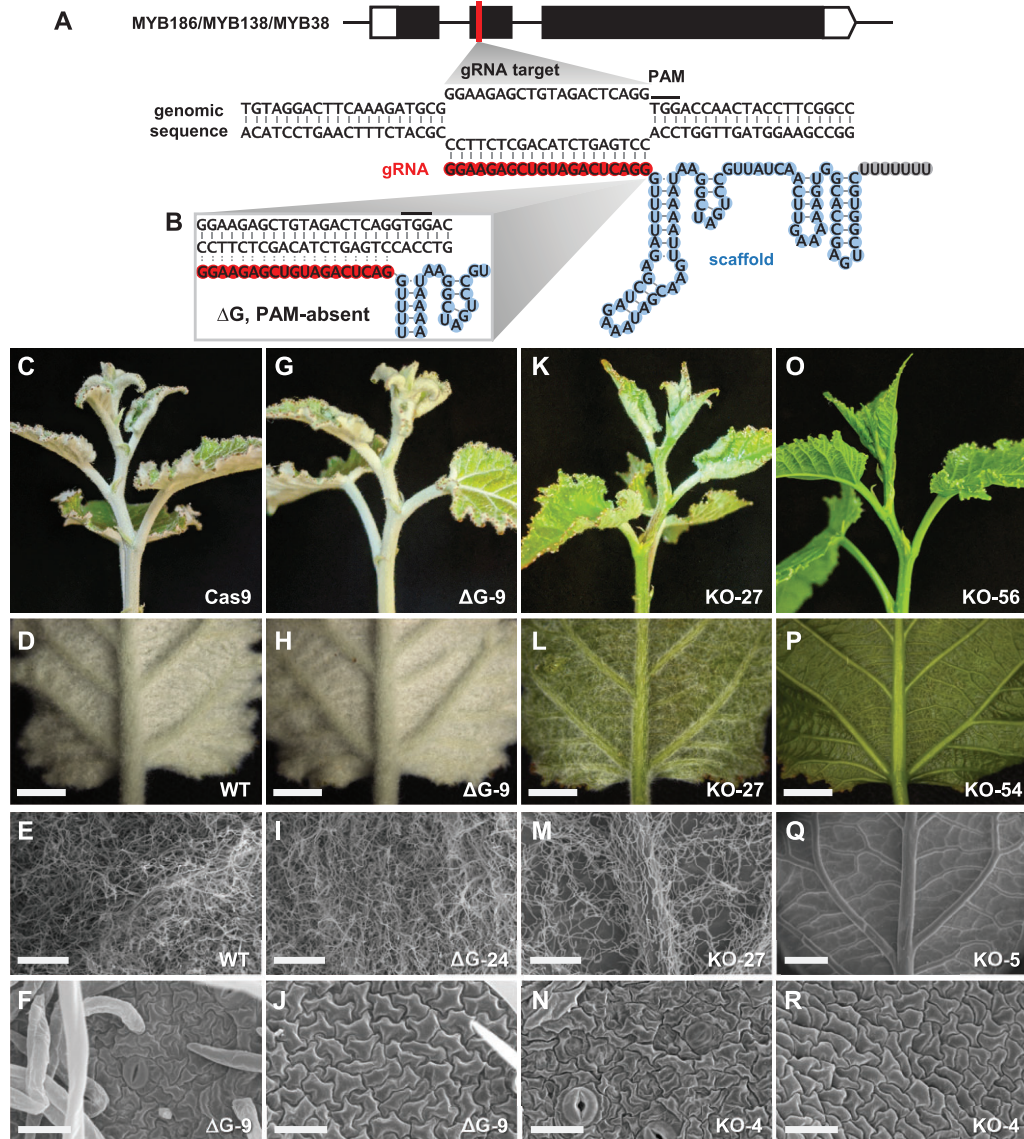
Figure 2 CRISPR/Cas9 KO of trichome-regulating *MYBs*. A, Schematic illustrations of the *MYB* gene structure, gRNA target site, and base pairing between the genomic target (black) and the gRNA spacer (red)-scaffold (blue) molecule. Black line denotes the PAM. B, Zoomed-in view of the ΔG vector configuration at the gRNA spacer-scaffold junction with a guanine omission. C–R, Representative shoot tip (C, G, K, and O) and LPI-1 abaxial (D, H, L, and P) phenotypes and SEM images (E, F, I, J, M, N, Q, and R) of soil-grown WT (D and E), Cas9 vector control (C), ΔG control (G-I), KO-27 (K–M), and null mutant (O–Q) plants, and leaf abaxial (F and N) or adaxial (J and R) images of tissue cultured ΔG (F and J) and null mutant (N and R) plants. Scale bar = 3 mm (D, H, L, and P), 500 μm (E, I, and M), 1 mm (Q), or 25 μm (F, J, N, and R).

followed by the Chr8a sites (two tandem copies) and least at the single-copy Chr10 sites (Figure 3, A and B). The NA alleles on Chr8 often spanned consecutive copies, suggesting large dropouts between two gRNA cleavage sites. To support this idea, we examined a subset of mutant lines using allele-specific primers for polymerase chain reaction (PCR) amplification of the target genes. As expected, NA alleles yielded no PCR products, whereas alleles previously detected by amplicon sequencing produced observable PCR products (Supplemental Figures S1 and S2). We next used consensus primers for PCR amplification of all six Chr8 (*MYB186* and *MYB138*) alleles, approximately 850–950 bp, from three control plants and four KO lines each with four to five NA

alleles on Chr8. These KO lines had reduced PCR band intensity when compared with controls (Figure 4, A and B). Sanger sequencing of the PCR products resulted in clean chromatograms with clear nucleotide peaks throughout the sequenced length for KO-5 and KO-69 (Figure 4C), two mutant lines with only one detectable Chr8 allele (Figure 4B). In contrast, the chromatograms for KO-63, KO-70 (both containing two detectable Chr8 alleles), and WT samples were noisy as would be expected for mixed template (Figure 4, B and C). The Sanger sequencing data of KO-5 and KO-69 not only confirmed the indel pattern (-2 in both cases) detected by amplicon sequencing, but also supported the occurrence of gene fusion between two gRNA cleavage
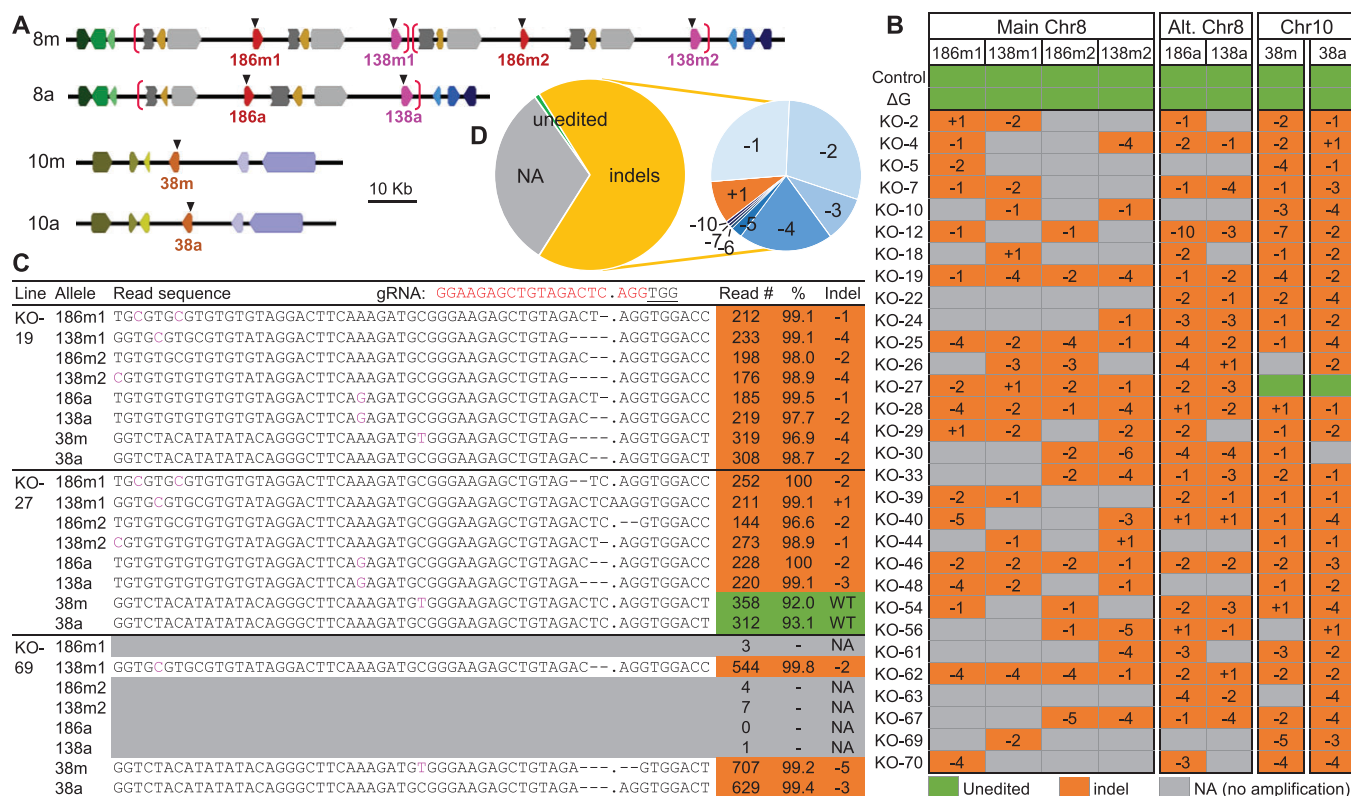
Figure 3 Mutation analysis of trichomeless mutants. A, Schematic illustration of *MYB186* and *MYB138* on Chr8 subgenomes (main and alternative, or 8m and 8a, respectively) and *MYB38* on Chr10 (10m and 10a) of the 717 genome. Neighboring genes are shown for synteny and the putative duplication block containing *MYB186* and *MYB138* on Chr8 is marked by brackets. Triangles denote the eight gRNA target sites. B, Mutation spectrum determined by amplicon sequencing. The eight alleles are arranged by genomic position for each plant line and the editing outcomes noted. NA, no amplification. C, Representative amplicon sequencing output of three mutant events. All eight alleles, their detection frequencies, and indel patterns (mapped read count and percentage with the indicated pattern) are shown. The gRNA target sequence is shown on top and protospacer adjacent motif underlined. Allele-discriminating SNPs are shown (see Supplemental Data Set S1 for the full data). D, Pie chart summary of the overall (left) and indel (right) editing patterns.

sites, based on SNP patterns upstream and downstream of the gRNA target (Figure 4, B and C). KO-5 harbors a fusion junction between *MYB186m1* and *MYB138m1* with a ~29-kb genomic dropout, whereas KO-69 contains a fusion of *MYB138m1* and *MYB138m2* with a ~62-kb genomic dropout (Figure 4, B and C; Supplemental Figure S3). Both events likely contain additional large deletions or genomic fusions, as allele(s) downstream (KO-5) or upstream (KO-69) of the respective fusion point could not be PCR amplified (Figure 4B). Regardless, our findings show that a single gRNA is highly effective for multiplex KO of tandem duplicates via either small indels or large deletions.

Assessment of off-target activity in mutants

A combination of computational prediction and experimental verification was used to assess off-target effects. Potential off-target sites of the gRNA were predicted by CCTop (Stemmer et al., 2015) using the *P. trichocarpa* v3.1 reference genome as well as the two SNP-substituted Pta717 v2 (*P. alba* and *P. tremula*) subgenomes (Xue et al., 2015). The same four exonic locations were ranked among the top potential off-target sites (excluding intergenic or unassembled scaffold sequences) across the three genomes, each having three mismatches with the gRNA sequence. We designed three sets of primers to examine potential editing at the four off-target sites; OT1 (Potri.004G115600 and Potri.004G118000), OT2 (Potri.004G138000), and OT3 (Potri.014G024400). Amplicon sequencing of 20 trichomeless mutants found no off-target activity across these four sites (Supplemental Data Set S2).

Absence of triterpenes in trichomeless leaves

Trichomes as epidermal outgrowths are covered with waxy cuticles like other epidermis cells (Hegebarth et al., 2016). The striking glabrous phenotype of the mutants prompted us to compare leaf wax composition between control and trichomeless plants. Total wax load of mature leaves (extractable wax from leaf surface) did not change significantly between genotypes (Figure 5A). Alkanes were the most abundant class of leaf cuticular waxes detected in 717 and differed little between control and trichomeless plants (Figure 5B). In contrast, levels of fatty alcohols and β-sitosterol were significantly reduced in the mutants (Figure 5B), with two primary alcohols, 1-octacosanol (C28)
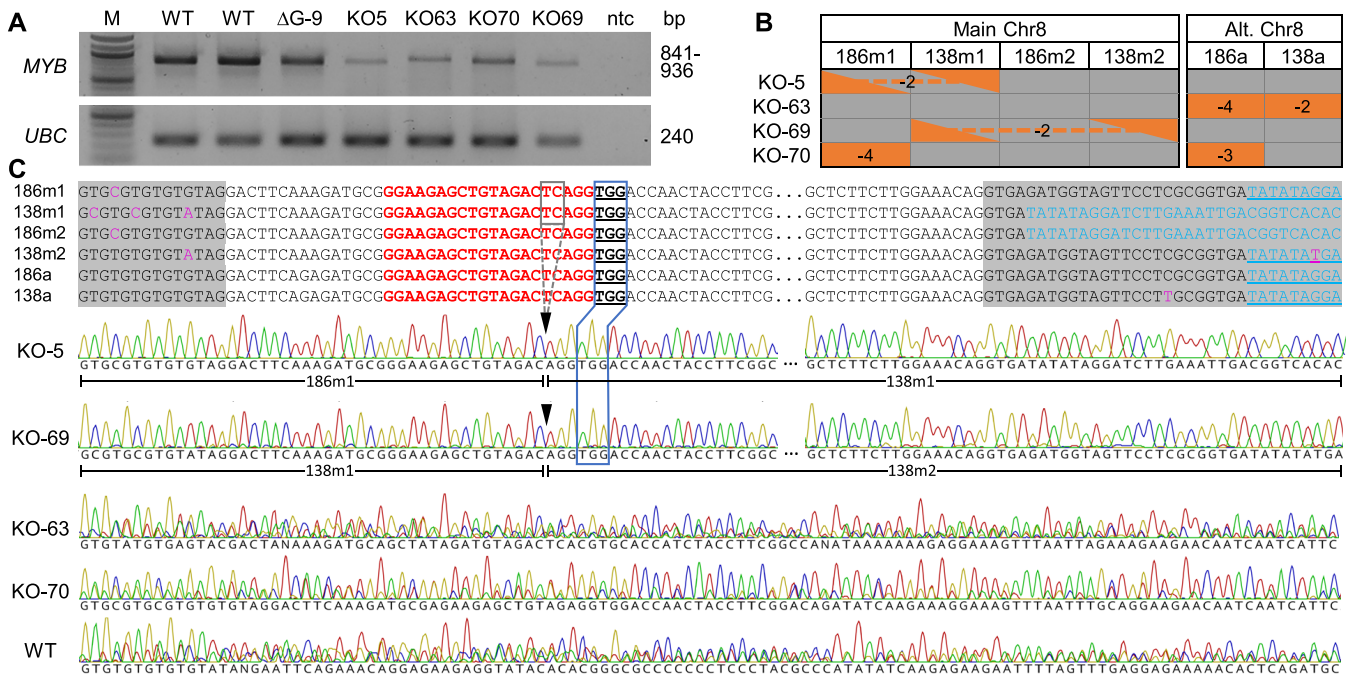
Figure 4 PCR analysis of selected mutant lines. A, PCR amplification of the six *MYB* alleles on Chr8 from two WT, one ΔG and four KO lines. The four KO lines were selected to represent one (KO-5 and KO-69) or two (KO-63 and KO-70) remaining Chr08 alleles (see Supplemental Figure S2 for uncropped gel images). *UBC* (ubiquitin-conjugating enzyme) was included as loading control. M, molecular weight marker; ntc, no-template control. B, Mutation patterns of Chr08 alleles in the selected mutant lines as shown in Figure 3B. Triangles connected by dashes represent fusion junctions shown in (C). C, Sanger sequencing of PCR products from A. Sequence alignment of the six alleles flanking the gRNA target site is shown on top and chromatograms of the same region are shown below. Shaded alignments are introns, with allele-discriminating SNPs and homologous intron 2 sequences (shifted upstream by 21 bp in *138m1* and *186m2* due to gaps) highlighted. PAM is underlined and boxed for correspondence with the sequence traces below. Triangles denote the Cas9 cleavage site and dashed box corresponds to the 2-bp deletion (-2) detected in KO-5 and KO-69. The two fusion alleles as determined by SNPs are marked below the KO-5 and KO-69 traces (see Supplemental Figure S3 for the full sequence alignment).

and 1-hexacosanol (C26), depleted by >50% (Figure 5C), and β-sitosterol, by 42% (Figure 5D). Strikingly, the wax of mutant leaves was devoid of any triterpenes (Figure 5B), including α-amyrin, β-amyrin, β-amyrone, and lupenone (Figure 5E). To further investigate the absence of triterpenes in the mutant wax, whole leaf tissues were also profiled for compounds that were significantly reduced in cuticular wax. Again, triterpenes were not detected in the leaves of trichomeless mutants (Figure 5F), whereas 1-octacosanol, 1-hexacosanol, and β-sitosterol were detected at levels comparable with controls (Figure 5, C and D). The data support a previously unsuspected link between triterpene accrual and nonglandular trichomes in poplar.

Discussion

The present study demonstrates that a single gRNA targeting conserved genomic sites is highly effective for multiplex editing in poplar. The 30 independent KO lines experienced an average of 5.4 CRISPR/Cas9-mediated cleavages per line based on indel alleles, which is likely an underestimate because many NA alleles may also result from CRISPR/Cas9 cleavages as shown for KO-5 and KO-69 (Figures 3 and 4). The unexpected genomic complexity in the hybrid 717 highlights the importance of ensuring SNP-free targets for gRNA

design (Zhou et al., 2015), as well as the challenge of decoding multiplexed edits among highly similar gene duplicates.

The negligible editing by the ΔG construct (Supplementary Data Set S1) provides insight into scaffold structure and stability. The ΔG configuration can lead to two hypothetical outcomes: either the guanine is omitted from the scaffold and the gRNA remains intact and capable of base pairing to the target sites for Cas9 cleavage, or the guanine is sequestered for secondary structure folding of the scaffold, resulting in a 3′-truncated gRNA no longer protospacer adjacent motif (PAM)-adjacent at the target sites (Figure 2B). The lack of mutations in ΔG transformants supports the latter scenario and is consistent with transcription and folding of gRNA molecules preceding their base-pairing with genomic targets. Our finding suggests that 3′-truncated gRNA could serve as an alternative approach for generating transgenic controls.

A number of methods are commonly used for decoding CRISPR-mediated mutation patterns, including restriction digestion, endonuclease-based mismatch detection, gene-/allele-specific PCR sometimes in conjunction with cloning and/or Sanger sequencing (Figure 4), or amplicon deep sequencing (Figure 3). The pros and cons of these methods have been discussed elsewhere (e.g. Germini et al., 2018).
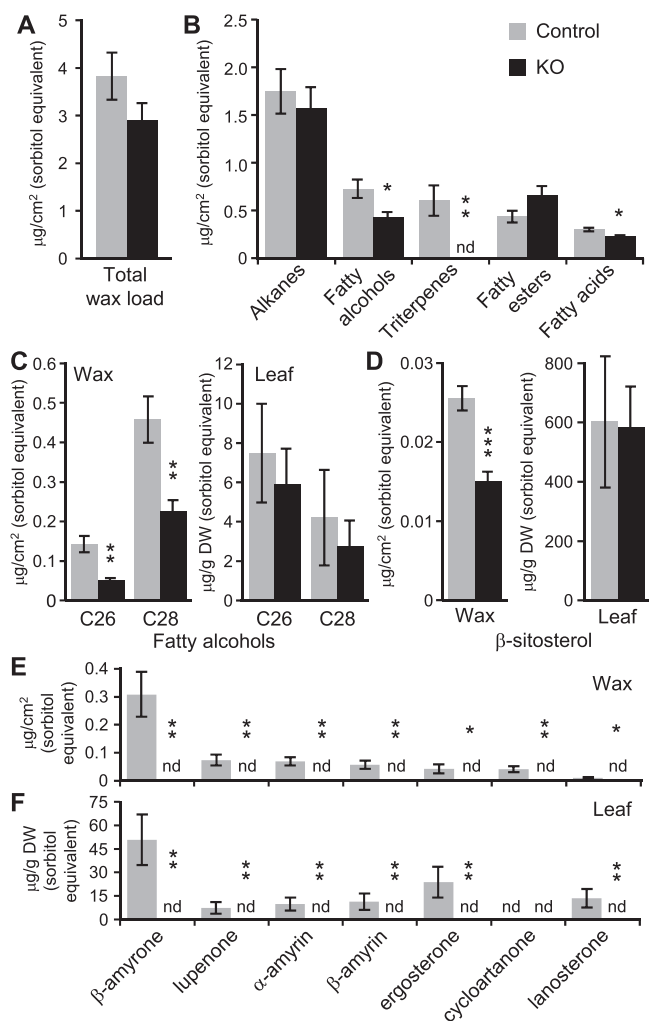
Figure 5 Cuticular wax composition of trichomeless and control leaves. A, Total wax load. B, Major classes of cuticular wax. C, Fatty alcohols (C26, 1-hexacosanol; C28, 1-octacosanol) in wax (left) or whole leaves (right). D, β-sitosterol detected in wax (left) or whole leaves (right). E and F, Triterpenes detected in wax (E) or whole leaves (F). Ergosterone, 14,24-dimethyl-ergosta-8,25-dien-3-one; cycloartanone, 24-methylene cycloartan-3-one; lanosterone, lanosta-8,24-dien-3-one. Data are mean±SD of *n* =5–6. All concentration estimates were based on sorbitol equivalent. Statistical significance was determined by two-tailed *t* test (*P < 0.05, **P < 0.01, ***P < 0.001). ND, not detected.

Analysis of genome editing across multiple target sites poses additional challenges over mono-targeted experiments, especially when highly homologous target and flanking sequences are encountered. These multiplex scenarios generally exceed the resolution of most methods or may require additional assays (e.g. allele-specific PCR) to determine editing outcomes. In the case exemplified here, deep sequencing of a pooled library of amplicons obtained with consensus primers for all eight target alleles was highly effective for decoding multiplexed edits. The use of consensus primers provides built-in controls for each PCR, allowing for high confidence calling of NA alleles (see Supplemental Data Set S1) which are otherwise difficult to distinguish from failed

PCR in individual reactions. As another advantage, the amplicon deep sequencing data can be used for *de novo* assembly which in our case led to the discovery of unexpected copy number variations of *MYB186* and *MYB138* in the experimental poplar 717 genotype. Although technical limitations remain in short-read mapping to homologous sites, inclusion of allele-specific SNPs within the amplicon region and adoption of bioinformatic programs with parameter tuning capabilities (e.g. AGEseq) are key to multiplexed mutation pattern determination.

The glabrous mutants (Figure 2) provide strong support for an essential role of PtaMYB186/138/38 in the initiation of trichome development in 717. Additionally, the low trichome density of KO-27 suggests that MYB38 plays a redundant but minor role in leaf/stem trichome initiation (Figures 2 and 3). Follow-up research, including allele-specific KOs, is needed to dissect the functional redundancy and allele-dose response of clade 15 MYB members more fully. The unedited (WT) *MYB38* alleles in KO-27 appear stable during vegetative propagation as this event has maintained a low trichome density for over 2 years in both tissue culture and greenhouse environments. This adds to previously reported stability of CRISPR editing outcomes in clonally propagated poplar (Bewg et al., 2018)

The loss of trichomes did not significantly affect the total epidermal wax load but led to altered wax composition and a complete absence of triterpenes in the mutant leaves. Reduced accumulation of certain wax components in the mutant leaves may simply reflect a greatly reduced surface area due to loss of trichomes. This is likely the case for fatty alcohols and β-sitosterol as their levels were depleted only in cuticular wax extracted from leaf surface, but not when whole leaves were analyzed (Figure 5, C and D). However, triterpenes were completely absent both in cuticular wax and whole leaves of the mutants, a finding that cannot be explained by reduced surface area of glabrous leaves. It is unlikely that MYB186/138/38 have an additional role in triterpene biosynthesis (i.e. lack of triterpenes as a direct KO effect) because of their recent duplication history (Figure 1) and because a recent report implicated phylogenetically distinct MYBs in triterpene regulation (Falginella et al., 2021). We interpret the absence of triterpenes in trichomeless leaves as suggesting a role for nonglandular trichomes in triterpene accrual in poplar. While glandular trichomes are well known for their roles in biosynthesis and storage of terpenes (Lange and Turner, 2013), the presence of terpenes in nonglandular trichomes has only been reported recently (Santos Tozin et al., 2016; Dmitruk et al., 2019). The genetic evidence presented herein provides strong support for a functional link between triterpenes and nonglandular trichomes that warrants further investigation.

The glabrous phenotype of the null mutants we obtained highlights the potential utility of trichomes as a visual reporter. Assessments of CRISPR/Cas functionality often target the chlorophyll biosynthetic enzyme phytoene desaturase (PDS; Norris et al., 1995), as mutations result in an albino

phenotype (Shan et al., 2013; Ma et al., 2015; Xie et al., 2015). Whilst phenotypically obvious, *PDS* mutations are lethal for the regenerated plant, thus limiting follow-up investigations. Alternatively, the glabrous phenotype achieved by KO of trichome-regulating *MYBs* is nonlethal and no inhibition to plant growth was detected. This allows stacked mutagenesis of these mutants, including reparative transformations to restore trichome initiation. The use of trichomes as a visual reporter for CRISPR/Cas9 mutation or repair of a defective allele has been established in Arabidopsis (*Arabidopsis thaliana*) (Hahn et al., 2017, 2018) which provides support for further developing this system in poplar.

Materials and methods

Generation of KO mutants

The ΔG and KO constructs in p201N-Cas9 (Jacobs et al., 2015) were prepared by Gibson assembly. PCR was used to amplify the p201N-Cas9 binary vector following *Swa*I (New England BioLabs) digestion, and the *Medicago truncatula MtU6.6* promoter and scaffold fragments from HindIII and EcoRI (New England BioLabs) digested pUC-gRNA shuttle vector (Jacobs et al., 2015), with Q5 High-Fidelity DNA Polymerase (New England BioLabs) and primers (Sigma) listed in Supplemental Table S1. The p201N-Cas9 (Addgene 59175) and pUC-gRNA (Addgene 47024) plasmids were both gifts from Wayne Parrott. Two pairs of oligos (Sigma) corresponding to the consensus gRNA target site in exon two of *MYB186* (Potri.008G089200), *MYB138* (Potri.008G089700), and *MYB38* (Potri.010G165700) were assembled with p201N-Cas9. The NEBuilder HiFi DNA Assembly Cloning Kit (New England Biolobs) was used to assemble p201N-Cas9, *MtU6.6* promoter and scaffold fragments with a pair of oligos containing the gRNA target sequence (Supplemental Table S1). Following transformation into DH5α *Escherichia coli* (Zymo Research Mix & Go! Competent Cells), PCR-positive colonies were used for plasmid purification before Sanger sequencing (Eurofins) confirmation. Plasmids were then heat shocked into *Agrobacterium tumefaciens* strain C58/GV3101 (pMP90; Koncz and Schell, 1986) and confirmed by colony PCR.

Poplar (*Populus tremula* x *alba* IRNA 717-1B4) transformation and regeneration was performed as outlined in Meilan and Ma (2006), except 0.05-mg·L^{-1} 6-benzylaminopurine was used in shoot elongation media, and 200-mg·L^{-1} L-glutamine was added to all media, with 3-g·L^{-1} gellan gum (PhytoTechnology Lab) as a gelling agent. Following a 2-d agrobacterial cocultivation, leaf discs were washed in sterile water followed by washing in 200-mg·L^{-1} cefotaxime and 300-mg·L^{-1} timentin with shaking for 1.5 h. Transformants were selected on media supplemented with 100-mg·L^{-1} kanamycin, 200-mg·L^{-1} cefotaxime, and 300-mg·L^{-1} timentin for callus induction and shoot regeneration and with kanamycin and timentin for shoot elongation and rooting. All cultures were grown and maintained at 22°C under a 16-h

light/8-h dark photoperiod with Growlite FPV24 LED (Barron Lighting Group) at ∼150 μmol·m^{-2}s^{-1}.

RNA-seq analysis

For developmental profiling, LPI-1, LPI-5, and LPI-15 were collected from three greenhouse-grown WT plants (∼5 ft in height) for RNA extraction using Direct-zol RNA MiniPrep kit (Zymo Research) with Plant RNA Purification Reagent (Invitrogen). RNA-seq library preparation and Illumina NextSeq 500 sequencing was performed at the Georgia Genomics and Bioinformatics Core. We obtained 10.8-13.3 PE75 reads per sample. After pre-processing to remove adapter and rRNA sequences, reads were mapped to the 717 SNP-substituted genome sPta717 v2 (Xue et al., 2015) using STAR v2.5.3a (Dobin and Gingeras, 2015). Transcript abundance in FPKM (fragments per kilobase of transcript per million mapped reads) was estimated by featureCounts v1.5.2 (Liao et al., 2014).

Amplicon sequencing determination of mutation spectrums

Newly emerged leaves were excised from individual events in tissue culture for genomic DNA extraction (Dellaporta et al., 1983). The DNA pellet was resuspended in water with RNase A (10 μg·mL^{-1}) for amplicon library preparation using GoTaq G2 Green Master Mix (Promega) and primers (Supplemental Table S1) spanning the gRNA target site (between 264 bp and 280 bp). Samples were then barcoded with Illumina amplicon indexing primers and pooled for Illumina MiSeq nano PE150 sequencing performed at the University of Georgia's Georgia Genomics and Bioinformatics Core. Demultiplexed sequence reads were analyzed by the AGEseq (Analysis of Genome Editing by Sequencing) program (Xue and Tsai, 2015), with mismatch allowance set at 1%, followed by manual curation.

Because initial amplicon data analysis revealed lower editing efficiencies (<90%) than we typically observed in 717 (Zhou et al., 2015; Bewg et al., 2018) at several target sites, we performed *de novo* assembly of WT amplicon reads using Geneious v11.1.2 (Biomatters), and recovered seven distinct alleles. We then searched the JGI draft 717 genome assembly v1.0 with the *P. trichocarpa* Nisqually-1 v3.1 (Phytozome v12) *MYB186*, *MYB138*, and *MYB38* gene models and extracted the surrounding 50- to 150-kb regions from Chr8 and Chr10 for manual annotation against the *P. trichocarpa* Nisqually-1 reference (Figure 3A). The matching *MYB* gene sequences were extracted for error correction using 717 resequencing data (Xue et al., 2015). Curated sequences were used for new (amplicon and allele-/gene-specific) primer design and as references in amplicon data analysis. In the case of WT and transgenic controls with no editing, erroneous read assignments—and hence indel calls—still remained because the amplicon region between some alleles differs only in the number of intronic dinucleotide (GT) repeats (Supplemental Data Set 1). Misassigned reads led to erroneous indel calls of −2, +2 or their multiples outside of the gRNA target site. For this reason, WT

and control samples were processed by ustacks from Stacks 2.3 (Catchen et al., 2011). Parameters were adjusted to avoid collapsing reads with SNPs and/or Indels from paralogous alleles into the same tag group and gapped alignments were disabled. Tags from the output were then used for allele assignment.

Determination of leaf and cuticle wax compositions

Leaf punches (25-mm diameter) were taken from mature leaves of similar size (between LPI-10 and LPI-15) of soil-grown control ($n = 5$ from two WT and three ΔG lines) and KO ($n = 5$ lines) plants in a growth chamber and washed in 4 mL of methylene chloride for 30 s. The washes were dried under a continuous N_2 stream before resuspension in 400-μL chloroform. A 200-μL aliquot was subsequently dried under vacuum and the residues shipped to the Oak Ridge National Laboratory for analysis. Sorbitol (1 mg·mL^{-1}) was added to the residues as an internal standard and re-dried under N_2. For whole leaf analysis, liquid nitrogen-ground and freeze-dried powders from LPI-5 (25 mg) of control ($n = 6$ from three WT and three ΔG lines) and KO ($n = 5$ lines) plants were extracted by 80% (v/v) ethanol to which sorbitol (1 mg·mL^{-1}) was added and dried under N_2. The samples were derivatized prior to analysis on an Agilent Technologies 7890A GC coupled to a 5975C inert XL MS fitted with an Rtx-5MS capillary column with a 5m Integra-Guard column (Restek) as described in Holwerda et al. (2014). Compound identification was based on mass spectral fragmentation patterns against the NIST14 database and an in-house library built with authentic standards. Two-tailed Student's t test was determined using JMP Pro Version 15.0.0 (SAS).

Tissue imaging and SEM analysis

Images of poplar were taken with either a Google Pixel 3a running Android v11, or a Leica M165 FC dissection microscope attached to a Leica DFC500 camera running Leica Application Suite software v3.8.0. SEM observations were obtained using Hitachi 3400 NII (Hitachi High Technologies America) microscope following optimized protocols at the Center for Ultrastructural Research at the Fort Valley State University. LPI-1 from growth chamber plants or young leaves of tissue culture plants were processed for primary fixation at 25°C in 2% glutaraldehyde (Electron Microscopy Sciences, EMS) prepared with Sorensen's Phosphate buffer, pH 7.2 (EMS) for one hour and then washed three times for 15 min each with the same buffer before secondary fixation in 1% osmium tetroxide (EMS) prepared in Sorensen's Phosphate buffer, pH 7.2 for 1 h at 25°C. After three washes with dH$_2$O for 15 min each, fixed tissues were dehydrated with ethanol series passing through 25%, 50%, 75%, and 95% for 15 min each, followed by three changes of 100% ethanol for 15 min each. Critical point drying of fixed samples was conducted using a critical point dryer (Leica) and then samples were placed on Hitachi M4 aluminum specimen mounts (Ted Pella) using double sided carbon adhesive tabs (EMS) for coating. Gold coating of 50 Å thickness was done

for 60 s using sputter coater (Denton Desk V) under a vacuum pressure of 0.05 Torr. Image acquisition in various magnifications was done at accelerating voltage of 5 kV.

Accession numbers

The RNA-seq data have been deposited to the National Center for Biotechnology Information's Sequence Read Archive under accession no. PRJNA753499. The poplar *MYB* gene models are *MYB186*, Potri.008G089200; *MYB138*, Potri.008G089700; *MYB38*, Potri.010G165700; and *MYB83*, Potri.017G086300.

Supplemental data

The following materials are available in the online version of this article.

Supplemental Table S1. Primers used in this study.

Supplemental Figure S1. PCR confirmation of NA alleles using allele-specific primers.

Supplemental Figure S2. Original gel images.

Supplemental Figure S3. Sequence alignment of wild-type and fusion *MYB* alleles from KO-5 and KO-69.

Supplemental Data Set S1. CRISPR/Cas9 mutation patterns of the eight target *MYB* alleles in ΔG and KO lines.

Supplemental Data Set S2. Assessment of off-target activity in trichomeless mutants.

References

Armario Najera V, Twyman RM, Christou P, Zhu C (2019) Applications of multiplex genome editing in higher plants. Curr Opin Biotechnol **59**: 93–102

Bewg WP, Ci D, Tsai C-J (2018) Genome editing in trees: from multiple repair pathways to long-term stability. Front Plant Sci 9: 1732

Braatz J, Harloff H-J, Mascher M, Stein N, Himmelbach A, Jung C (2017) CRISPR-Cas9 targeted mutagenesis leads to simultaneous modification of different homoeologous gene copies in polyploid oilseed rape (Brassica napus). Plant Physiol 174: 935–942

Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH (2011) Stacks: building and genotyping loci de novo from short-read sequences. Genes Genomes Genet 1: 171–182

Čermák T, Curtin SJ, Gil-Humanes J, Čegan R, Kono TJY, Konečná E, Belanto JJ, Starker CG, Mathre JW, Greenstein RL, et al. (2017) A multipurpose toolkit to enable advanced genome engineering in plants. Plant Cell 29: 1196–1217

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

Dellaporta SL, Wood J, Hicks JB (1983) A plant DNA minipreparation: version II. Plant Mol Biol Rep 1: 19–21

Dmitruk M, Sulborska A, Żuraw B, Stawiarz E, Weryszko-Chmielewska E (2019) Sites of secretion of bioactive compounds in leaves of Dracocephalum moldavica L.: anatomical, histochemical, and essential oil study. Braz J Bot 42: 701–715

Dobin A, Gingeras TR (2015) Mapping RNA-seq reads with STAR. Curr Protoc Bioinformatics 51: 11.14.11–11.14.19

Endo A, Masafumi M, Kaya H, Toki S (2016) Efficient targeted mutagenesis of rice and tobacco genomes using Cpf1 from Francisella novicida. Scient Rep 6: 38169

Endo M, Mikami M, Toki S (2015) Multigene knockout utilizing off-target mutations of the CRISPR/Cas9 system in Rice. Plant Cell Physiol 56: 41–47

Falginella L, Andre CM, Legay S, Lin-Wang K, Dare AP, Deng C, Rebstock R, Plunkett BJ, Guo L, Cipriani G, et al. (2021) Differential regulation of triterpene biosynthesis induced by an early failure in cuticle formation in apple. Horticult Res 8: 75

Flagel LE, Wendel JF (2009) Gene duplication and evolutionary novelty in plants. New Phytol 183: 557–564

Gao Y, Zhao Y (2014) Self-processing of ribozyme-flanked RNAs into guide RNAs in vitro and in vivo for CRISPR-mediated genome editing. J Integr Plant Biol 56: 343–349

Germini D, Tsfasman T, Zakharova VV, Sjakste N, Lipinski M, Vassetzky Y (2018) A comparison of techniques to evaluate the effectiveness of genome editing. Trends Biotechnol 36: 147–159

Hahn F, Eisenhut M, Mantegazza O, Weber APM (2018) Homology-directed repair of a defective Glabrous gene in Arabidopsis with Cas9-based gene targeting. Front Plant Sci 9: 424

Hahn F, Mantegazza O, Greiner A, Hegemann P, Eisenhut M, Weber AP (2017) An efficient visual screen for CRISPR/Cas9 activity in Arabidopsis thaliana. Front Plant Sci 8: 39

Hegebarth D, Buschhaus C, Wu M, Bird D, Jetter R (2016) The composition of surface wax on trichomes of Arabidopsis thaliana differs from wax on other epidermal cells. Plant J 88: 762–774

Holwerda EK, Thorne PG, Olson DG, Amador-Noguez D, Engle NL, Tschaplinski TJ, van Dijken JP, Lynd LR (2014) The exometabolome of Clostridium thermocellum reveals overflow metabolism at high cellulose loading. Biotechnol Biofuels 7: 155

Jacobs TB, LaFayette PR, Schmitz RJ, Parrott WA (2015) Targeted genome modifications in soybean with CRISPR/Cas9. BMC Biotechnol 15: 16

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337: 816–821

Koncz C, Schell J (1986) The promoter of TL-DNA gene 5 controls the tissue-specific expression of chimaeric genes carried by a novel type of Agrobacterium binary vector. Mol Gen Genet 204: 383–396

Lange BM, Turner GW (2013) Terpenoid biosynthesis in trichomes—current status and future opportunities. Plant Biotechnol J 11: 2–22

Li A, Jia S, Yobi A, Ge Z, Sato SJ, Zhang C, Angelovici R, Clemente TE, Holding DR (2018) Editing of an alpha-kafirin gene family increases, digestibility and protein quality in sorghum. Plant Physiol 177: 1425–1438

Li J-F, Norville JE, Aach J, McCormack M, Zhang D, Bush J, Church GM, Sheen J (2013) Multiplex and homologous recombination–mediated genome editing in Arabidopsis and Nicotiana benthamiana using guide RNA and Cas9. Nat Biotechnol 31: 688–691

Liao Y, Smyth GK, Shi W (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30: 923–930

Lowder LG, Zhang D, Baltes NJ, Paul JW, Tang X, Zheng X, Voytas DF, Hsieh T-F, Zhang Y, Qi Y (2015) A CRISPR/Cas9 toolbox for multiplexed plant genome editing and transcriptional regulation. Plant Physiol 169: 971–985

Ma X, Zhang Q, Zhu Q, Liu W, Chen Y, Qiu R, Wang B, Yang Z, Li H, Lin Y, et al. (2015) A robust CRISPR/Cas9 system for convenient, high-efficiency multiplex genome editing in monocot and dicot plants. Mol Plant 8: 1274–1284

McCarty NS, Graham AE, Studená L, Ledesma-Amaro R (2020) Multiplexed CRISPR technologies for gene editing and transcriptional regulation. Nat Commun 11: 1281

Meilan R, Ma C (2006) Poplar (Populus spp.). In K Wang, ed, Agrobacterium Protocols Volume 2. Methods in Molecular Biology, Vol 344. Humana Press, Totowa, NJ

Mikami M, Toki S, Endo M (2017) In planta processing of the SpCas9–gRNA complex. Plant Cell Physiol 58: 1857–1867

Nasti RA, Voytas DF (2021) Attaining the promise of plant gene editing at scale. Proc Natl Acad Sci USA 118: e2004846117

Nekrasov V, Staskawicz B, Weigel D, Jones JDG, Kamoun S (2013) Targeted mutagenesis in the model plant Nicotiana benthamiana using Cas9 RNA-guided endonuclease. Nat Biotechnol 31: 691–693

Norris SR, Barrette TR, DellaPenna D (1995) Genetic dissection of carotenoid synthesis in Arabidopsis defines plastoquinone as an essential component of phytoene desaturase. Plant Cell 7: 2139–2149

Panchy N, Lehti-Shiu M, Shiu S-H (2016) Evolution of gene duplication in plants. Plant Physiol 171: 2294–2316

Plett JM, Wilkins O, Campbell MM, Ralph SG, Regan S (2010) Endogenous overexpression of Populus MYB186 increases trichome density, improves insect pest resistance, and impacts plant growth. Plant J 64: 419–432

Qi L, Haurwitz RE, Shao W, Doudna JA, Arkin AP (2012) RNA processing enables predictable programming of gene expression. Nat Biotechnol 30: 1002–1006

Santos Tozin LRd, de Melo Silva SC, Rodrigues TM (2016) Non-glandular trichomes in Lamiaceae and Verbenaceae species: morphological and histochemical features indicate more than physical protection. NZ J Bot 54: 446–457

Shan Q, Wang Y, Li J, Zhang Y, Chen K, Liang Z, Zhang K, Liu J, Xi JJ, Qiu J-L, et al. (2013) Targeted genome modification of crop plants using a CRISPR-Cas system. Nat Biotechnol 31: 686–688

Stemmer M, Thumberger T, del Sol Keyer M, Wittbrodt J, Mateo JL (2015) CCTop: an intuitive, flexible and reliable CRISPR/Cas9 target prediction tool. PLoS One 10: e0124633

Tang X, Ren Q, Yang L, Bao Y, Zhong Z, He Y, Liu S, Qi C, Liu B, Wang Y, et al. (2019) Single transcript unit CRISPR 2.0 systems for robust Cas9 and Cas12a mediated plant genome editing. Plant Biotechnol J 17: 1431–1445

Tang X, Zheng X, Qi Y, Zhang D, Cheng Y, Tang A, Voytas Daniel F, Zhang Y (2016) A single transcript CRISPR-Cas9 system for efficient genome editing in plants. Mol Plant 9: 1088–1091

Tsai C-J, Xu P, Xue L-J, Hu H, Nyamdari B, Naran R, Zhou X, Goeminne G, Gao R, Gjersing E, et al. (2020) Compensatory guaiacyl lignin biosynthesis at the expense of syringyl lignin in 4CL1-knockout poplar. Plant Physiol 183: 123–136

Wang B, Luo Q, Li Y, Yin L, Zhou N, Li X, Gan J, Dong A (2020) Structural insights into target DNA recognition by R2R3-MYB transcription factors. Nucleic Acids Res 48: 460–471

Wang M, Mao Y, Lu Y, Wang Z, Tao X, Zhu J-K (2018) Multiplex gene editing in rice with simplified CRISPR-Cpf1 and CRISPR-Cas9 systems. J Integr Plant Biol **60**: 626–631

Wilkins O, Nahal H, Foong J, Provart NJ, Campbell MM (2009) Expansion and diversification of the *Populus* R2R3-MYB family of transcription factors. Plant Physiol **149**: 981–993

Xie K, Minkenberg B, Yang Y (2015) Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. Proc Natl Acad Sci USA **112**: 3570–3575

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

Xue L-J, Alabady MS, Mohebbi M, Tsai C-J (2015) Exploiting genome variation to improve next-generation sequencing data analysis and genome editing efficiency in Populus tremula x alba 717-1B4. Tree Genet Genomes **11**: 82

Xue LJ, Tsai CJ (2015) AGEseq: analysis of genome editing by sequencing. Mol Plant **8**: 1428–1430

Yang L, Güell M, Niu D, George H, Lesha E, Grishin D, Aach J, Shrock E, Xu W, Poci J, et al. (2015) Genome-wide inactivation of porcine endogenous retroviruses (PERVs). Science **350**: 1101–1104

Zhang Y, Bai Y, Wu G, Zou S, Chen Y, Gao C, Tang D (2017) Simultaneous modification of three homoeologs of TaEDR1 by genome editing enhances powdery mildew resistance in wheat. Plant J **91**: 714–724

Zhou X, Jacobs TB, Xue L-J, Harding SA, Tsai C-J (2015) Exploiting SNPs for biallelic CRISPR mutations in the outcrossing woody perennial *Populus* reveals 4-coumarate:CoA ligase specificity and redundancy. New Phytol **208**: 298–301

Zwaenepoel A, Van de Peer Y (2018) wgd—simple command line tools for the analysis of ancient whole-genome duplications. Bioinformatics **35**: 2153–2155