High incidence of homozygous mutation of floral genes in CRISPR-Cas9 transgenic poplars

Estefania Elorriaga, Amy L. Klocko, Cathleen Ma, and Steven H. Strauss

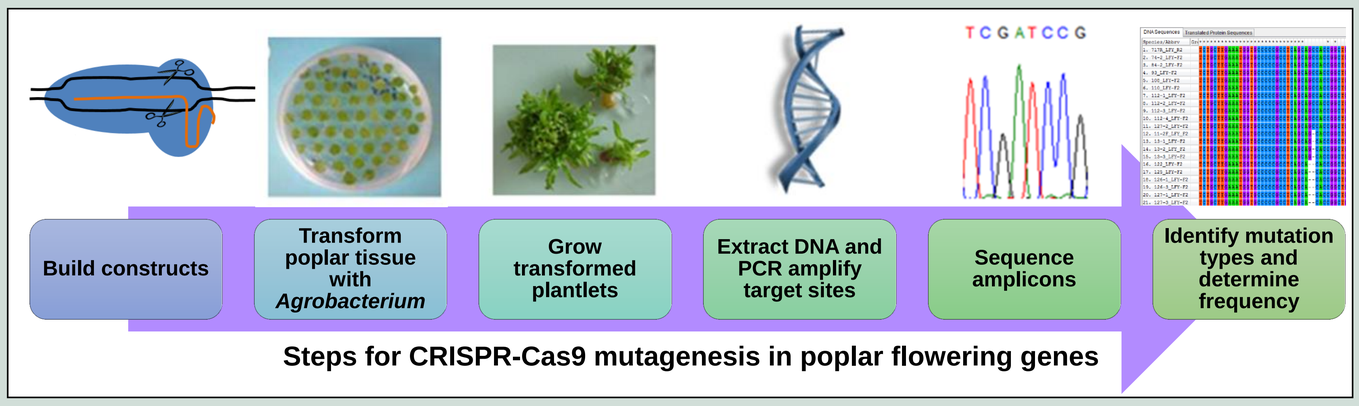
Department of Forest Ecosystems and Society, Oregon State University, Corvallis, OR 97331, USA

Overview

Gene flow from genetically engineered (GE) trees into feral or wild populations are significant obstacles to their use as a result of regulatory, public perception, and ecological concerns. Loss-of-function mutations in a number of floral transcription factor genes can lead to sterility in diverse plant species. However such mutations are rare and generally recessive, thus are very difficult to induce via conventional tree breeding. The recently rediscovered Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)-associated Cas system has proven to be a powerful directed-mutagenesis tool in many species, including trees. CRISPR-Cas induced mutations appear to be highly predictable and stable, and reversion should be extremely rare or impossible (e.g., when there are deletions of essential parts of coding regions). We are testing the mutation efficiency of four nucleases targeting two essential floral genes in *Populus*. The targets are the poplar orthologs *LEAFY* and *AGAMOUS*—well-studied genes essential for both male- and female-fertility. The nucleases have been stably transformed into hundreds of independent events that we are now analyzing for mutation rate and type. Sequencing results from 300 independent insertion events indicate that one-quarter of transgenic events have identical, loss-of-function mutations in both alleles, suggesting CRISPR-Cas9 mutagenesis is a highly efficient method for genetic containment in poplar.

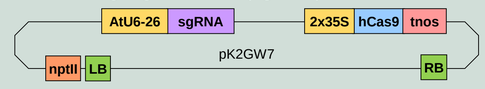
Experimental design

- Select target genes for mutagenesis
- Design and assembled CRISPR-Cas9 nuclease constructs including a Cas9 only control
- Produce stably transformed plants using *Agrobacterium tumefaciens*
- Extract DNA from rooted events
- PCR-amplify across target sites
- Sanger sequence targeted gene regions
- Align the sequences to determine mutant prevalence and type

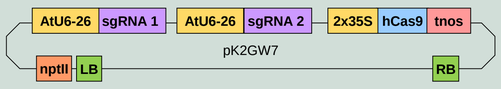


CRISPR-Cas9 vectors

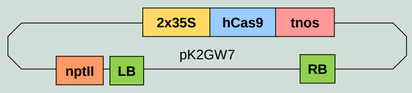
Construct with a single guide RNA (sgRNA)



Construct with two guide RNAs to induce large deletions



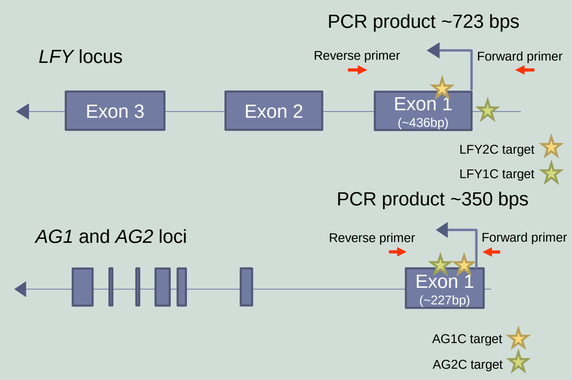
Cas9 control construct



CRISPR-Cas9 transgenic populations studied

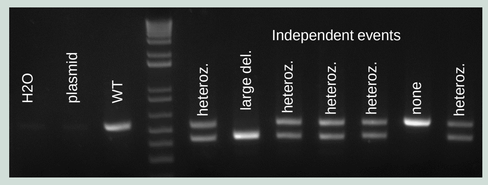
Construct	Target gene(s)	No. of regenerated shoots	No. of putative transgenic events	No. of events analyzed molecularly	Percent of events analyzed
AG1C	*PtAGs*	1,149	294	54	18%
AG2C	*PtAGs*	1,376	125	12	10%
AG1-AG2C	*PtAGs*	1,953	196	79	40%
LFY1C	*PtLFY*	1,307	153	131	86%
LFY-C	*PtLFY*	1,205	131	46	35%
LFY1-LFY2C	*PtLFY*	1,673	107	60	56%
Total		**8,663**	**1,006**	**382**	**38%**
Cas9	N/A	970	95	14	15%

PCR assays for both target genes



We designed the assay with primers far enough away from the target sites so that PCR amplification should not, except for major deletions, be affected.

Large mutations were easily visualized on gels



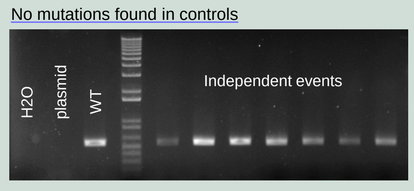
Example of a large deletion and heterozygous mutants

Gel image to the left shows seven LFY1-LFY2C independent events with different mutation types across the two alleles: five heterozygous mutants (deletion and/or insertion in one allele or both), one biallelic large deletion, and one non-mutant. Heteroz.; heterozygous, del; deletion.

Example of a large insertion

Gel image of a large (391 bp) insertion in two different events with LFY1C.

No mutations found in controls



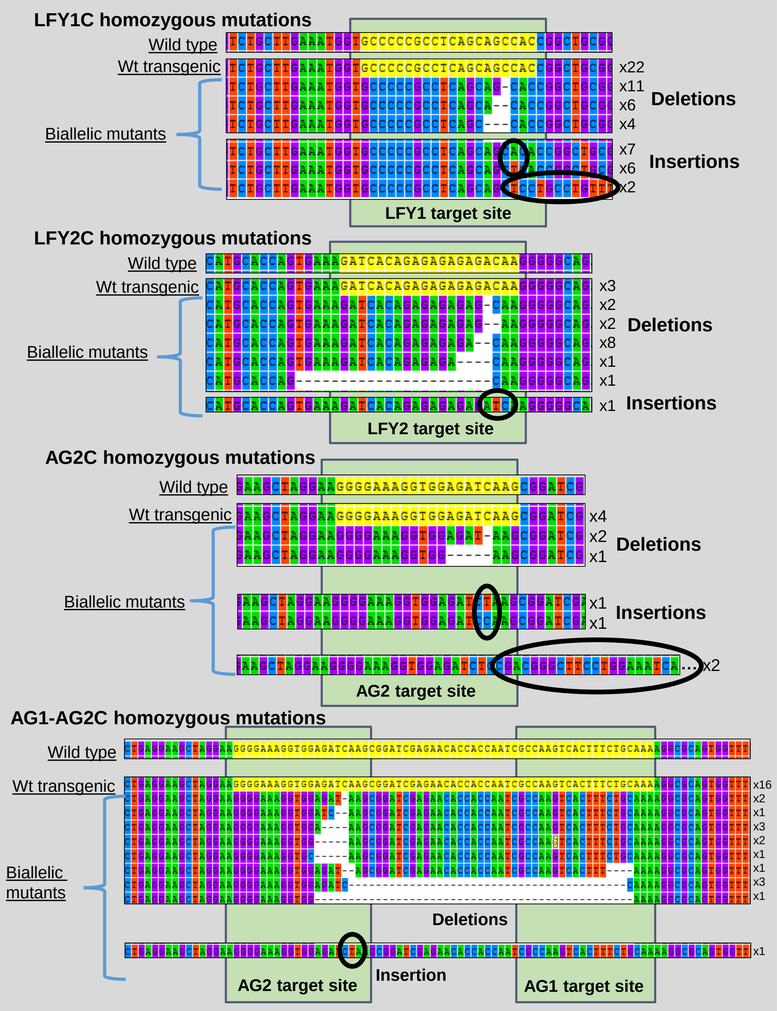
Gel image showing six Cas9 only (control vector) independent events. No obvious mutations were observed based on band size.

CRISPR-Cas9 system was highly mutagenic

Construct	GE events sequenced	Type of mutation	# of events (%)
Single LFY1C	131	Homozygous	40 (30%)
		Heterozygous	69 (53%)
		None	22 (17%)
Single LFY2C	46	Homozygous	15 (32%)
		Heterozygous	28(61%)
		None	3 (7%)
Double LFY1-LFY2C	60	Homozygous	11 (18%)
		Heterozygous	44 (74%)
		None	5 (8%)
Single AG1C	54	Homozygous	0 (0%)
		Heterozygous	7 (74%)
		None	47 (8%)
Single AG2C	12	Homozygous	7 (58%)
		Heterozygous	1 (8%)
		None	4 (34%)
Double AG1-AG2C	79	Homozygous	19 (24%)
		Heterozygous	45 (57%)
		None	15 (19%)
Cas9 (control vector)	14	None	14 (100%)
Total (w/out control)	**382**	**Homozygous**	**92 (24%)**
		Heterozygous	**194 (51%)**
		None	**96 (25%)**

The AG1C construct performed very poorly. If we removed it from the summary, we found that the rate of homozygous mutations increased to 28%, the rate of heterozygous mutations increased to 57%, and the rate of events with no mutations was reduced to 15%.

Small indels, large indels and inversions were observed



The events with inversions are not shown because of their size. They all happened in the LYF1C-LFY2C population and were 120 bps long.

SNPS in the poplar *LFY* gene sequence helped to identify alleles

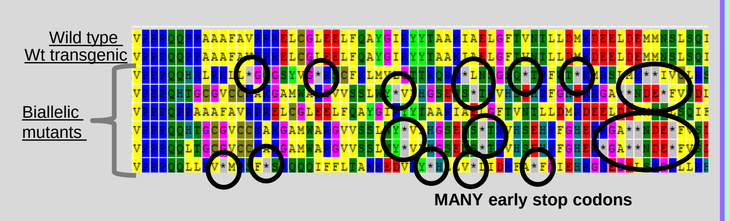
CCTGTTAAGGGCAGTTTTGGTATA(C/T)AAATAAAACAAGAAGCTCACTTGTCTTTATATATCTACCAAATCCAAGACATGCACCAGTGAAAGATCACAGAGAGAGAGACAAGGGGGCAGATAGATATGGATCCGGAGGCTTTCACGGCGAGTTTGTTCAAATGGGACACGAGAGCAATGGTGCCACATCCTAACCGTCTGCTTGAAATGGTGCCCCCGCCTCAGCAGCCACCGGCTGCGGCGTTTGCTGTAAGGCCAAGGGAGCTATGTGGGCTAGAGGAGTTGTTTCAAGCTTATGGTATTAGGTACTACACGGCAGCGAAAATAGCTGAACTCGGGTTCACAGTGAACACCCTTTTGGACATGAAAGA(T/C)GAGGAGCTTGATGAAATGATGAATAGTTTGTCTCAGATCTTTAGGTGGGATCTTCTTGTTGGTGAGAGGTATGGTATTAAAGCTGCTGTTAGAGCTGAAAGAAGAAGGCTTGATGAGGAGGATCCTAGGCGTAGGCAATTGCTCTCTGGTGATAATAATACAAATACTCTTGATGCTCTCTCCCAAGAAGGTTTGGTTAGCATTGATTCTACCTTTTAGTGTAATTAAGCTAAGCTCATACTATTACTAGCTATAGGAG(T/G)CCATGGCCGTTTT(G/A)TATTTCCCTTTTGGTTGTAGTT TTATAGAGTAAATTAATTCTATGTATACTTGGATAAGATAATTAGCTTATTATAAGTTGTTACTTGCCAGCT

Sequences matching sgRNAs; PCR primers; SNPs; First exon

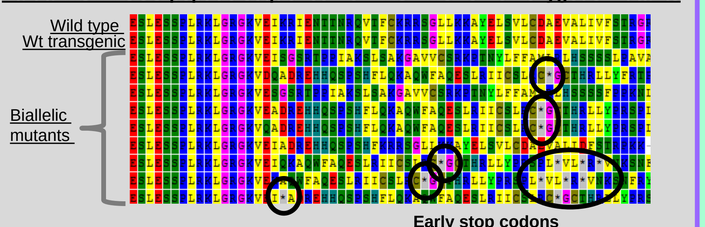
Haplotype one: C…T…T…G Haplotype two: T…C…G…A

Mutations destroy protein function

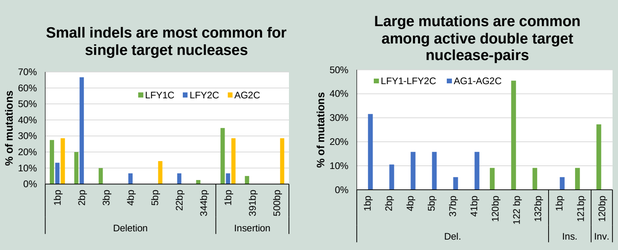
Partial LEAFY peptide sequence from LFY1C homozygous mutants



Partial AGAMOUS peptide sequence from AG1-AG2C homozygous mutants



Mutation prevalence varied among nuclease targets



The single nucleases (having only one sgRNA) tended to have mostly small indels (1 to 5bp), while the double nuclease with two highly active sgRNAs (LFY1-LFY2C) tended to have mainly large deletions. The double nuclease with only one highly active sgRNA (AG1-AG2C) had mainly small deletions just like the single nucleases. For this table we used the homozygous mutation data. The total number of mutations for each nuclease can be seen in the table above. Del.: deletion, Ins.: insertion, Inv.: inversion.

Determining mutation frequency from heterozygous events

Rationale: To determine per allele mutation frequency (in the absence of possible double-strand break stimulated gene conversion that may occur with homozygous mutants), we sequenced both alleles of a subset of our heterozygous, mutated events.

Methods: We selected 12 heterozygous events at random from a total of 69 events in the LFY1C population. We amplified the sequence across target sites using PCR, ligated PCR products into a TOPO vector, transformed competent *E. coli*, and sequenced individual colonies. The amplification of each allele was verified by four natural SNPs present in the 717 hybrid genotype. We also sequenced both alleles of 5 homozygous wild type and mutant events.

Heterozygotes	# Events	Genotype frequency
Wt/Mt1	4	0.33
Mt1/Mt2	7	0.59
Mt1/Mt2/Mt3	1	0.08
Totals	12	1.00

Homozygotes	# Events	Genotype frequency
Mt1/?	2	0.40
Mt1/Mt1	3	0.60
Totals	5	1.00

Mt: mutated allele; Wt: wildtype allele; Mt1/Mt1: both alleles have identical mutations; Mt1/Mt2: both alleles are mutated but have different mutations; Mt1/Mt2/Mt3: chimeric mutant with different mutations in the same allele; Mt1/?: one allele is mutated and the other has not been amplified via PCR yet.

Among the 12 heterozygous events, we found that 7 had different mutations in both alleles, 4 had a wildtype allele and a mutated allele, and 1 was chimeric (where all alleles were mutated but one allele had two different mutations). Among the 5 homozygous events, 3 had the same mutation in both alleles, and in the other 2 events we only were able to sequence one allele after PCR amplification. We suspect that the other allele had a large deletion that removed one of the primer binding sites, which are now verifying. With our current sample, we were unable to determine if the homozygous mutation rate was above that expected due to chance alone (i.e., if there was evidence for gene conversion, aka “asexual gene drive”).

Summary

- CRISPR-Cas9 nucleases are highly effective at inducing site-directed mutations at the target loci in poplar.
- The majority of the mutations occur 3bp upstream from the protospacer adjacent motif (PAM) site as expected.
- Constructs with a single guide RNA generally lead to small indel mutations (1-5 bp).
- Constructs with two active guide RNAs targeting the same locus induced large deletions (6-121 bp).
- Transgenic controls with only the Cas9 gene showed no mutations.
- CRISPR mutagenesis appears to be a highly effective means for making permanent gene knock-outs for genetic containment in poplar.

Acknowledgements

We thank the industrial members of the Tree Biosafety and Genomics Research Cooperative (TBGRC) at Oregon State University, the United States Department of Agriculture (USDA award 2011-68005-30407, System For Advanced Biofuels Production From Woody Biomass In The Pacific Northwest), the USDA Biotechnology Risk Assessment (grants 2011-68005-30407 and 2010-33522-21736), NSF I/UCRC Center for Advanced Forestry (grant 0736283), USDA-IFAS (grant OREZ-FS-671-R), and the J. Frank Schmidt Charitable Foundation. We thank Dr. Jian-Kang Zhu at Purdue University and Dr. Yanfei Mao at Shanghai Center for Plant Stress Biology for providing the CRISPR-Cas9 constructs. We thank the OSU Graduate School and PAG for travel funding. We thank Sarah Robertson, high school student Clark Embleton, and undergrad workers Analeslie Martinez and Melissa Meyhoff for their good work in the laboratory.