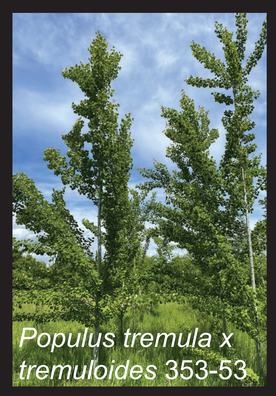
Low rates of non-target mutations in field- and greenhouse-grown CRISPR/Cas9 expressing transgenic trees

Greg S. Goralogia[1], Isabella M. Andreatta[1], Qin Xiong[1], Kelly J. Vining[2], Estefania Elorriaga[1], Cathleen Ma[1] and Steven H. Strauss[1]

[1]Department of Forest Ecosystems and Society, Oregon State University, Corvallis, OR, 97331
[2]Department of Horticulture, Oregon State University



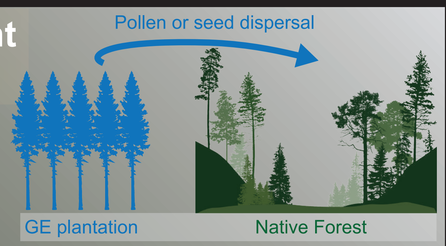
Reproductive containment in forest biotechnology

GE trees with advantageous traits might share habitat with native trees with which they can interbreed, or encroach on native forests as exotics

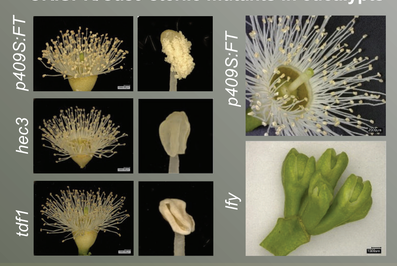
Many trees are wind pollinated or have wide seed dispersal, making containment more challenging than most other crops

We have used CRISPR/Cas9 to produce edited poplars and eucalypts with the goal of complete sterility

By targeting the floral initiation gene *LEAFY (LFY)*, no floral organs should initiate. Other edited trees in this study include those with knock-out mutations in *AGAMOUS* genes, which should produce flowers unable to produce functional male or female organs



CRISPR/Cas9 sterile mutants in eucalypts

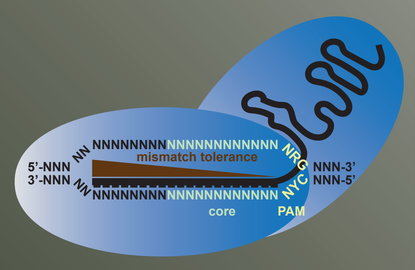


CRISPR/Cas9: a highly specific editor in plants -but what about when expressed for many years?

Cas9 has a proven record in many plant species for high activity, high specificity, and low rates of non-target mutations

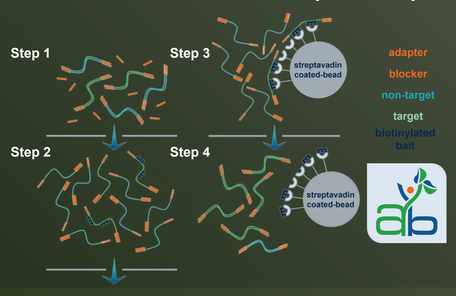
Most studies looked *in vitro* and soon segregated away the Cas9 transgene - in trees these may be in the plant for perpetuity as our other studies suggest that they appear to cause no harm

Most studies used whole genome sequencing approaches or investigated only sequences with close mismatch to target



Targeted sequencing was used to confirm CRISPR target mutations and sequence mismatched non-target sites up to 5 bases divergent

Cas-OFFinder → Bait design (20,000) → Bait capture on gDNA → Short read sequencing



Sampling timeline:

1. Two years in tissue culture/greenhouse
2. Fall 2017 planting in the field
3. Tissue collection 2019
4. Sequences obtained 2021
5. Results January 2021

This is approximately 3 or 4 years of Cas9/gRNA expression

Mutect2 - a program well suited to identify mutations in clonally propagated plants

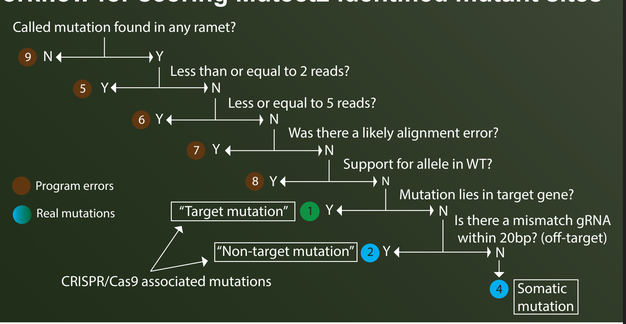
Reads (Illumina Nova-Seq) → BWA → gatk → Intervals over bait targets → Mutect2

Per construct, we obtained 20-100 sites for detailed analysis

We used a scoring key with criterion based on the strength (number of reads) for mutant allele support

For most loci we obtained 500-1,000 reads per bait, which enables high confidence characterization of chimeric, unfixed alleles

Workflow for scoring Mutect2 identified mutant sites



Non-target CRISPR mutations found in all species tested -but with only some gRNAs

Methods:

353 and 717 hybrid poplars were transformed with Cas9 only, and combinations of one or two gRNA containing constructs targeting two *AGAMOUS* homologs or the single copy *LEAFY* gene.

Two ramets were sampled from multiple events containing varied editing outcomes (homozygous/hemizygous edited, heterozygous edited, non-edited but transgenic).

26 events in poplar were surveyed which contained gRNA expression cassettes.

Nine SP7 eucalypt events were surveyed which contained gRNA expression cassettes. For eucalypts, only gRNAs targeting the *LFY* gene were tested, in single and double gRNA architectures.

No non-target mutations were observed in the *LFY* targeting constructs in the 717 or 353 poplar backgrounds.

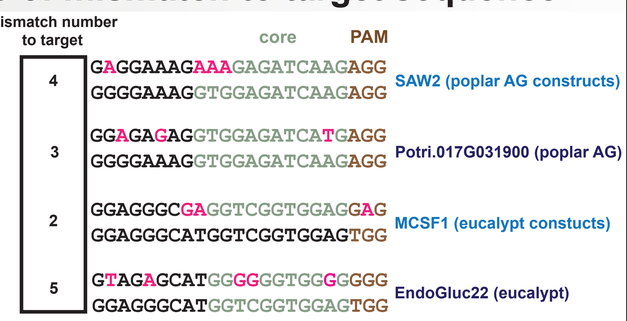
Results:

Somatic mutations identified tended to be shared between events or shared within all transgenic events, showing a common clonal lineage. Non-targets showed event specificity in mutational outcome.

Non-target mutation sites had two to five bases of mismatch to target sequence

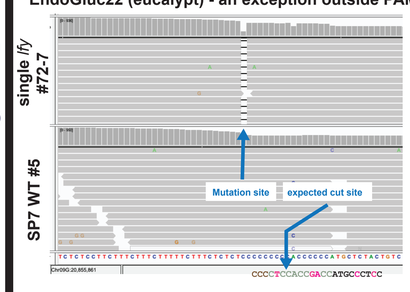
Non-target mutations had more mismatches to target sequences than was expected.

Though non-target sites tended to have more mismatches outside of the core spacer, mismatches were also present in those locations.

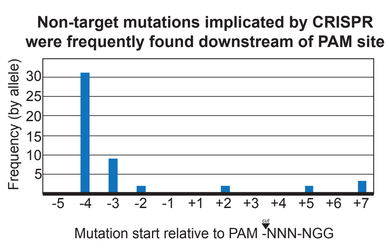


Mutations at non-target sites were at predicted locations relative to the PAM

EndoGluc22 (eucalypt) - an exception outside PAM

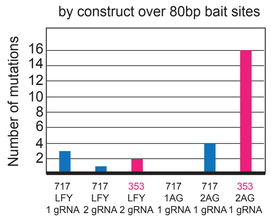


Non-target mutations implicated by CRISPR were frequently found downstream of PAM site



Very low rates of non-target mutations compared to somatic mutations

Accumulated novel somatic mutations by construct over 80bp bait sites



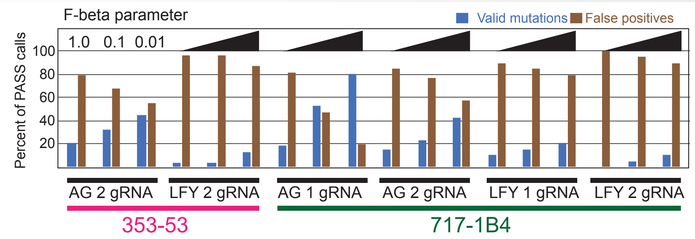
Overall somatic and non-target mutation rates were low over surveyed DNA

Construct	Somatic mutation rate[1]	Non-target mutation rate	Mutation frequency (somatic/non-target)	Mutation ratio (somatic to non-target)
All	9.59E-07	7.14E-07	(0.12/0.38)	1.34
1LFY	5.74E-07	0.00	(0.13/0.00)	
2LFY	2.87E-07	0.00	(0.10/0.00)	
1AG	0.00	1.53E-07	(0.00.0.17)	
2AG	1.62E-06	8.31E-07	(0.09/0.45)	1.95

[1] (# mutation sites /bp of genome surveyed) * frequency among ramets

We compared novel somatic mutations against non-target mutations within the bait windows surveyed.

High false-positive rates for passing sites in Mutect2



At the suggestion of the developers of Mutect2, we varied the F-beta parameter score of the model and queried whether false-positive rates decreased with increasing stringency. Unfortunately increased stringency also increased false-negatives, so manual curation at F=1.0 was used.

Conclusions and next steps

- Mutations due to CRISPR/Cas9 activity occur and at target sites that differ substantially from sgRNA sequences (up to 5 bp). We found more off-target mutations than Young et al., 2019 *Sci Rep*, or Wang et al., 2021 *Hort Res*, in maize or grape, respectively
- Better models of gRNA binding affinity will reduce chances of designing gRNAs with unintented mutagenic potential - still, at least in eucalypts (Elorriaga et al. 2021, *Plant Biotech J*) (TBD in poplar), there were no vegetative impacts on growth in non-target mutated events even when KOs were found
- Shared non-target mutations between ramets suggests that non-target edits accumulate during transformation and tissue culture, not due to prolonged expression of Cas9 - future studies will examine the relationship of Cas9/gRNA expression to the rate of non-target editing
- The edited poplars will likely begin forming catkins and flowering next year or the year after, when we will be studying whether we achieved complete sterility and if there are vegetative consequences related to target or non-target edits

Acknowledgments

We thank Estefania Elorriaga, Amy Klocko, and Cathleen Ma for their work generating and planting the CRISPR poplars and eucalypts

We thank the USDA Biotechnology Risk Assessment Grant Program, grant # 2017-33522-27098, for financial support

We also thank the Genetic Research on Engineering and Advanced Transformation of Trees (GREAT Trees) cooperative for their support

We also thank the numerous undergrads and staff who helped maintain the CRISPR/Cas9 Poplar transgenic field trial