

**Pollen-specific expression of *EcoRI* restriction endonuclease for
bioconfinement in *Panicum virgatum* L.**

**A Thesis Presented for the
Master of Science
Degree
The University of Tennessee, Knoxville**

**Jessica Nicole Stockdale
May 2024**

Copyright © 2024 by Jessica Nicole Stockdale
All rights reserved.

Acknowledgments

I would like to express my sincerest gratitude to my colleagues Rob Sears, Tyler Newton, Bryce Trull for their guidance, support, and friendship throughout the past few years. I am beyond thankful for Dr. Shamira Sultana, Dr. Alex Pfothenauer, Dr. Li Li, and Dr. Yongil Yang for offering invaluable advice and assistance throughout this research endeavor. I also would like to extend appreciation to Lance Hamilton and Nicolas Campbell for greenhouse assistance, and to Brianna Jacobs and Chandler Brown for their technical support. This research would not have been possible without their help. I would also like to acknowledge Cassie Halvorsen for exceptional lab management and assisting me with any and all needs.

Finally, I would like to express my gratitude to my committee members Dr. Neal Stewart, Dr. Feng Chen, Dr. Charlie Kwit, and Dr. Reginald Millwood for their invaluable mentorship, guidance, and patience throughout this work. I am deeply grateful for their dedication and investment in my academic development.

Abstract

The dispersal of pollen from genetically engineered (GE) crops can lead to unintended transgene flow, especially in wind-pollinated and obligate-outcrossing species such as switchgrass (*Panicum virgatum* L.). Various bioconfinement approaches have been studied for their potential to eliminate transgene flow from GE crops including selective male sterility, which aims to inhibit the germination of transgenic pollen grains. In this study, pollen-specific promoters controlling the expression of the *EcoRI* endonuclease, interrupted by a catalase intron, was evaluated for its efficacy to produce sterile pollen. The TaPSG719, PvPS1, Osg6B, OsRTS, and Zm13 promoters were assessed for pollen-specific expression patterns, none of which have previously been characterized in switchgrass. Here, I report relatively high *EcoRI* transcript abundance in pollen/anther tissue compared to the switchgrass *Ubq1* reference gene when *EcoRI* is driven by the Zm13 ($10\% \pm 0.9\%$) and PvPS1 ($6\% \pm 1.3\%$) promoters. However, the TaPSG719, Osg6B, and OsRTS promoters did not significantly differ from nontransgenic switchgrass in terms of *EcoRI* transcript abundance in pollen/anther tissue. Furthermore, the highest-expressing transgenic line of the Zm13:*EcoRI* and PvPS1:*EcoRI* constructs were used for pollen germination experiments. These lines were presumed to contain a mixture of transgenic and nontransgenic pollen and were observed to exhibit similar pollen germination percentages compared to nontransgenic switchgrass pollen. Upon further investigation, it was determined that the catalase intron was incorrectly spliced from the mRNA, likely resulting in a nonfunctional protein when translated, which may explain why similar pollen germination rates were observed between the transgenic lines and control. In addition to these findings, I observed statistically significant phenotypic differences in many transgenic switchgrass lines when compared to nontransgenic control switchgrass. The most commonly observed adverse trait is decreased plant height, followed by decreased leaf blade width, decreased tiller diameter, and increased tiller count. However, off-target *EcoRI* expression did not appear correlative with plant height ($r = -0.08$, $p = 0.4134$) or leaf blade width ($r = -0.23$, $p = 0.0133$), suggesting alternative explanations for the observed phenotypes. Overall, the effectiveness of pollen-targeted *EcoRI* expression to produce sterile pollen was not determined. However, this study identified the Zm13 and PvPS1 promoters as strong candidates for male-specific gene expression and provided valuable insights for the design and production of genetically engineered switchgrass.

Table of Contents

Chapter 1: General Introduction	1
1. Introduction.....	2
2. Thesis Organization	3
3. References	4
Chapter 2: Transgene Bioconfinement: Don't Flow There	5
Abstract.....	7
1. Introduction.....	8
2. Engineered Reproductive Sterility	10
3. Transgene Excision	17
4. Delayed Flowering	23
5. Perspectives	25
References	26
Appendix.....	33
Chapter 3: Pollen-targeted expression of <i>EcoRI</i> endonuclease in <i>Panicum virgatum</i> L.....	36
Abstract.....	37
3.1 Introduction.....	38
3.2 Materials and Methods.....	40
3.3 Results	45
3.4 Discussion.....	48
3.5 Conclusions.....	50
References	52
Appendix.....	55
Chapter 4: Conclusions and Recommendations	72
Vita	74

List of Tables

Supplementary Table S3.1 PCR Primer Sets. List of primer sequences used to amplify PCR targets and their expected band sizes	55
Supplementary Table S3.2 qRT-PCR Primer Sets. List of target genes, primer sequences, primer melting temperatures, and expected product size of qRT-PCR experiments...	55
Supplementary Table S3.3 Summary of single- and double- copy transgenic switchgrass lines. List of constructs integrated into transgenic switchgrass, and the quantity of single-copy, double-copy, and combined total of single- and double- copy transgenic lines identified by southern hybridization.....	56
Supplementary Table S3.4 Southern hybridization analysis and newly assigned line numbers. Table provides transgenic switchgrass construct name, the line numbers as referred to in the southern hybridization experiments, determined transgene integration copy number of the lines, and newly assigned line number of transgenic lines that will be further analyzed	56

List of Figures

Figure 2.1 Engineered male-sterility approaches. (a) A sterility gene, such as *Barnase*, is located between an anther-specific promoter and a terminator (T) within the left (LB) and right borders (RB) on T-DNA inserted into the plant genome. *Barnase* expression prevents pollen formation by degrading RNA in the tapetal cells, the nutritive cells surrounding microspore mother cells. (b) A sterility gene, such as *EcoRI* endonuclease, is located between a pollen-specific promoter and a terminator within the left and right borders on the T-DNA, and in this example is inserted once in the plant genome. In this theoretical scenario, *EcoRI* expression leads to the death of transgenic pollen grains and nontransgenic pollen remains viable. Pollen movement from transgenic to nontransgenic plants yields 100% nontransgenic seeds. Figure created with Biorender.com, accessed on 22 February 2023.....33

Figure 2.2 Conditional sterility via inducible system activation of sterility genes. Within the left (LB) and right borders (RB) of the T-DNA located in the plant genome, an inducible promoter is located upstream of a recombinase gene and terminator (T). An external stimulus is applied to induce recombinase gene expression. The recombinase enzyme subsequently excises the blocker sequence by cutting DNA segments at flanking recognition sites. The tissue-specific promoter (ex. seed-specific) is moved into close proximity of the sterility gene and transcription can occur. The sterility gene product leads to cell death. In this theoretical example, a homozygous tobacco line with a seed-specific promoter driving sterility gene expression leads to 100% sterile transgenic seeds. Figure created with Biorender.com, accessed on 18 January 2023.....34

Figure 2.3 Transgene excision by site-specific recombinase and CRISPR/Cas9 gene editing. (a) Recombinase expression is controlled by a pollen-specific promoter. In pollen, the recombinase recognizes sites flanking the gene of interest (GOI) or selectable marker gene (SMG) and the recombinase expression cassettes and removes the DNA between the recognition sites. One recognition site and left (LB) and right (RB) borders remains while the DNA removed is degraded within the plant cell. (b) CRISPR/Cas9 excision utilize cleavage target sites (CTSs) that flank the guideRNA

(gRNA), Cas9 enzyme, and SMG expression cassettes. The CTSs are comprised of identical DNA sequences. Expression of a single gRNA and Cas9 enzyme excises the T-DNA between the CTSs. In this example, cleaved DNA ends are repaired by non-homologous end joining, leaving behind one CTS and left and right borders and the DNA removed is degraded within the plant cell. Figure created with Biorender.com, accessed on 22 February 2023 35

Figure 3.1. *EcoRI* plasmid T-DNA maps. Left (LB) and right (RB) borders flanking the T-DNA insert are represented by blue rectangles. All vectors contain the plant selectable antibiotic hygromycin (hpt) between *Panicum virgatum Ubi2* promoter and the cauliflower mosaic virus (CaMV) 35S terminator. All vectors also contain the visual marker mOrangeER between CaMV 35S promoter and the *Agrobacterium tumefaciens* nopaline synthase terminator (Nos T). All vectors contain the *EcoRI* endonuclease gene, interrupted by a catalase intron (derived from *Ricinus communis*), between the promoter **(A)** pTaPSG719 **(B)** pPvPS1 **(C)** pOsg6B **(D)** pOsRTS or **(E)** pZm13 and ubiquitous 35S terminator 58

Figure 3.2 Promoter effect of transcript expression of *EcoRI* endonuclease transgene in pollen/anther tissue. Bar graph of *EcoRI* endonuclease transcript levels driven by TaPSG719, PvPS1, Osg6B, OsRTS, and Zm13 promoters and a nontransgenic (NT) control. Expression is relative to the switchgrass ubiquitin 1 gene. Different letters represent statistical significance as tested by a one-way ANOVA, followed by a Tukey's HSD, $\alpha = 0.05$ ($F(5, 102) = 17.4$, $p < .0001$). Error bars represent standard error 59

Figure 3.3 Graphs of transcript abundance data. Graphs of *EcoRI* transcript levels in leaf (orange), whole flower collected prior anthesis (blue), and combined pollen and anther tissue collected immediately after anthesis (green) in **(A)** pZm13: *EcoRI* transgenic lines 1 through 4 and **(B)** pPvPS1: *EcoRI* transgenic lines 1 through 7, and nontransgenic controls (NT). **(C)** Endogenous *PS1* expression levels, driven by the native pPvPS1 promoter, in the pPvPS1: *EcoRI* transgenic lines 1 through 7 and a nontransgenic control. All transcript levels are set relative to the switchgrass *Ubq1*

housekeeping gene. An asterisk (*) represents statistical significance of <i>EcoRI</i> expression in anther/pollen tissue of transgenic lines compared to the nontransgenic control, tested by a one-way ANOVA, Tukey's HSD, $\alpha=0.05$. Error bars represent standard error	60
Figure 3.4 Pollen germination analysis. Images of flowers and pollen of (A,D) pPvPS1: <i>EcoRI</i> line 6, (B,E) pZm13: <i>EcoRI</i> line 4, and (C,F) nontransgenic switchgrass. (G) Box plot graph of pollen germination percentage (%) analysis, comparing pPvPS1: <i>EcoRI</i> line 6, pZm13: <i>EcoRI</i> line 4, and a nontransgenic control. Same letters represent no statistical significance of pollen germination percentages between transgenic lines and a nontransgenic control, tested by a one-way ANOVA, Tukey's HSD, $\alpha=0.05$ ($F(2, 24)=0.18$, $p=0.84$)	61
Figure 3.5 RT-PCR analysis for presence/absence of catalase intron. PCR analysis of cDNA extracted from anther/pollen tissue of flowering transgenic lines of the following constructs: (A) pTaPSG719: <i>EcoRI</i> , (B) pPvPS1: <i>EcoRI</i> , (C) pOsg6B: <i>EcoRI</i> , (D) pOsRTS: <i>EcoRI</i> , (E) pZm13: <i>EcoRI</i> . PCR amplified products represent <i>EcoRI</i> endonuclease with the presence of the catalase intron (1066 bp) or the absence of the catalase intron (predicted ~860 bp). The observed band in the absence of the catalase intron is approximately 600 bp. Gels include the TrackIt™ 100 bp DNA Ladder (L), transgenic line numbers of their associated construct, cDNA from nontransgenic switchgrass as a negative control (N), and plasmid DNA as a positive control (P) for the presence of the intron	62
Figure 3.6 Phenotypic analysis of transgenic lines. Bar graphs of (A) plant height, (B) tiller count, (C) tiller diameter, and (D) leaf blade width of transgenic lines, grouped by promoter driving <i>EcoRI</i> expression, compared to a nontransgenic control (c). Statistical significance is denoted by an asterisk (*), as tested by a one-way ANOVA, followed by a Dunnett's test, $\alpha=0.05$. Error bar is constructed using one standard error from the mean. Scatterplots depict the relationship of <i>EcoRI</i> expression levels in leaf and (E) plant height or (F) leaf blade width using Spearman's correlation analysis.....	63

Supplementary Figure S3.1 Confirmation of qRT-PCR primer sets. Prior qRT-PCR analysis of transgenic switchgrass, the *EcoRI* and *Ubq1* primer sets were assessed by RT-PCR and observed on a 1.2% agarose gel with the All Purpose Hi-Lo DNA Marker (L). The *EcoRI* primers were confirmed using genomic DNA from transgenic switchgrass carrying the *EcoRI* transgene to test for a positive band (94 bp) observed in sample 1. The *Ubq1* and *EcoRI* primers were tested using cDNA from nontransgenic switchgrass to confirm a positive band (93 bp) for the *Ubq1* gene observed in sample 2, and absence of the *EcoRI* gene observed in sample 3 64

Supplementary Figure S3.2 Southern blot hybridizations of transgenic switchgrass. Genomic (10 µg) and plasmid (500 pg) DNA was digested with *EcoRI*-HF and *SspI*-HF enzymes. Blots were probed with the DIG labeled-*EcoRI* gene fragment and hybridized at 42°C. Lanes consist of the DNA Ladder (L), genomic DNA samples of **(A)** pTaPSG719:*EcoRI* transgenic lines 1-11 **(B)** pPvPS1:*EcoRI* transgenic lines 1-15 **(C)** pOsg6B:*EcoRI* transgenic lines 1-12 **(D)** pOsRTS:*EcoRI* transgenic lines 1-16 **(E)** pZm13:*EcoRI* transgenic lines 1-5, followed by genomic DNA of wild-type switchgrass (WT) and the relevant plasmid DNA (P) 65

Supplementary Figure S3.3 Map of *EcoRI* endonuclease gene and the expected and observed results of splicing. **(A)** Map includes the *EcoRI* endonuclease gene interrupted by the catalase intron. cDNA sequencing identified the donor GT splice site is properly recognized and spliced from pre-mRNA. The first acceptor AG splice site downstream the donor site is the other intended splice site, however sequencing confirmed that the second acceptor AG is where splicing is occurring. **(B)** Map of the *EcoRI* endonuclease gene and the predicted start codon (Met) (green rectangle) and stop codon denoted by an asterisk (*) (red rectangle) if the catalase intron were properly removed. **(C)** Map of *EcoRI* endonuclease gene and the start and stop codons when the catalase intron is not spliced out. **(D)** Map of *EcoRI* endonuclease and the start and stop codons when splicing occurs at the second acceptor AG, which is observed in this study 66

Supplementary Figure S3.4 Images of transgenic lines following eight-week growth period.

Images of nontransgenic switchgrass control (NT) compared to transgenic lines of (A) pTaPSG719:*EcoRI* (B) pPvPS1:*EcoRI* (C) pOsg6B:*EcoRI* (D) pOsRTS:*EcoRI* and (E) pZm13:*EcoRI* constructs after eight weeks of growth under greenhouse conditions.....67

Supplementary Figure S3.5 Delayed flowering phenotype. OsRTS:*EcoRI* lines 1 and 2

displayed a delayed flowering phenotype where the flowers remained thin and underdeveloped (A) into reproductive stage 3 (R3) for an extended period of time (~ 2 months) before floral tissues further developed to produce a fuller flower appearance (B), and anthesis of normal-appearing flowers (C) was first observed in the second week of September. Nontransgenic flowers transitioned from thin and underdeveloped (E) to a swollen appearance (F) within approximately one to two weeks of reaching R3, and anthesis (G) began in early July. This resulted in closed flowers (D) of the apparently delayed-flowering lines during the flowering period of the nontransgenic switchgrass (H) 70

Supplementary Figure S3.6 Cloning methodology of the *EcoRI* endonuclease gene cassette.

(A) Multiple cloning site regions (red) used to clone each of the gene cassettes, bordered by attB sites (blue) produced by golden gate cloning. (B) The *EcoRI* endonuclease gene cassette that was cloned into the pSYEN3 MCS. Between the pollen-specific promoter and the 1st exon and the *EcoRI* coding sequence is a 49 bp 71

Chapter 1: General Introduction

1.1 Introduction

Gene flow is the movement of genes from one population to another and, in plants, may occur through the dispersal of pollen, seeds, or vegetative propagules (Lu 2008). Although a natural phenomenon, gene flow can be problematic when considering genetically engineered (GE) crops due to the potential escape of transgenes into the environment. Pollen-mediated gene flow is often regarded as the route of greatest concern since pollen is not easily contained and may travel long distances (Rizwan, Hussain et al. 2019). Such an occurrence could introduce transgenes to unintended hosts such as wild populations or adjacent crop fields and pose inadvertent consequences (Stewart Jr. 2021).

To alleviate these concerns, the mitigation and elimination of transgene flow has been an area of research for several decades. Biological confinement (bioconfinement) uses genetic engineering or gene editing to alter the plant's biological processes and mechanisms through a wide array of means to prevent transgenes from entering and persisting in the environment (Liu, Yuan et al. 2013). This may be attained through various approaches including sterility of reproductive organs, excision of transgenes from reproductive tissues, and altered flowering time to prevent overlap of reproductive periods. How these bioconfinement methods are achieved and their potential to eliminate transgene flow is discussed in detail in chapter 2.

Numerous bioconfinement systems have demonstrated their potential to effectively mitigate transgene flow, and one in particular achieved total bioconfinement and may be widely applicable across many crop species. This system conferred a selective pollen-sterility phenotype via pollen-specific expression of *EcoRI* restriction endonuclease (Millwood, Moon et al. 2016). In this study, a single copy of the *EcoRI* transgene was expressed in the dicot crop tobacco (*Nicotiana tabacum* L.) which produced a 1:1 ratio of transgenic to nontransgenic pollen. Transgenic pollen that expressed *EcoRI* endonuclease was rendered sterile whereas nontransgenic pollen remained viable, attaining over 99.7% and up to 100% bioconfinement when cross-pollinated to nontransgenic tobacco. This bioconfinement system has only been studied in tobacco and will likely require modifications to be deployed in other species including monocots and non-model species such as switchgrass— a potential GE crop that could also serve as a grass model. Specifically, limitations arise from the species-specific nature of the promoter used to drive gene expression, which may not effectively regulate gene expression in dissimilar species. To address this, several pollen-specific promoters identified in other monocot crop

species were evaluated for their spatial expression in switchgrass. These included the Osg6B, OsRTS, PvPS1, TaPSG719, and Zm13 promoters. To determine its potential as a bioconfinement system in switchgrass, *EcoRI* endonuclease was assessed for its efficacy to inhibit pollen germination when driven by the Zm13 and PvPS1 promoters. These research areas collectively form the central focus of chapter 3.

1.2 Thesis Organization

This thesis is organized into four chapters to summarize the objectives of this research.

Chapter 1: General introduction introduces the topic of gene flow and transgene bioconfinement, and the rationale of promoter characterization and analysis of *EcoRI* endonuclease efficacy in switchgrass.

Chapter 2: Transgene Bioconfinement: Don't Flow There, which was published as a review paper in *Plants*.

Chapter 3: Pollen-targeted expression of *EcoRI* endonuclease in *Panicum virgatum* L.

Chapter 4: General conclusions and further research that should be conducted were discussed in this chapter.

References

Liu, W., et al. (2013). "Advanced genetic tools for plant biotechnology." Nature Reviews Genetics **14**(11): 781-793.

Lu, B. R. (2008). "Transgene escape from GM crops and potential biosafety consequences: An environmental perspective." Collection of Biosafety Reviews **4**: 66-141.

Millwood, R., et al. (2016). "Engineered selective plant male sterility through pollen-specific expression of the *EcoRI* restriction endonuclease." Plant Biotechnology Journal **14**(5): 1281-1290.

Rizwan, M., et al. (2019). "Gene flow from major genetically modified crops and strategies for containment and mitigation of transgene escape: A review." Applied Ecology and Environmental Research **17**(5): 11191-11208.

Stewart Jr., C. N. (2021). Controlling transgene flow from engineered crops to unintended hosts by molecular approaches. Gene Flow: Monitoring, Modeling and Mitigation. W. Wei and C. N. Stewart, CAB International.

Chapter 2: Transgene Bioconfinement: Don't Flow There

Title: Transgene Bioconfinement: Don't Flow There

Jessica N. Stockdale, Reginald J. Millwood*

* Correspondence: rmillwood@utk.edu (R.J.M.)

Department of Plant Sciences, University of Tennessee (UTK), Knoxville, TN, 37996, USA

Published in *Plants*, <https://doi.org/10.3390/plants12051099>

Stockdale, J.S., Millwood, R.J. (2023) Transgene Bioconfinement: Don't Flow There. *Plants*. 12(5), 1099.

Abstract

The adoption of genetically engineered (GE) crops has led to economic and environmental benefits. However, there are regulatory and environmental concerns regarding the potential movement of transgenes beyond cultivation. These concerns are greater for GE crops with high outcrossing frequencies to sexually compatible wild relatives and those grown in their native region. Newer GE crops may also confer traits that enhance fitness, and introgression of these traits could negatively impact natural populations. Transgene flow could be lessened or prevented altogether through the addition of a bioconfinement system during transgenic plant production. Several bioconfinement approaches have been designed and tested and a few show promise for transgene flow prevention. However, no system has been widely adopted despite nearly three decades of GE crop cultivation. Nonetheless, it may be necessary to implement a bioconfinement system in new GE crops or in those where the potential of transgene flow is high. Here, we survey such systems that focus on male and seed sterility, transgene excision, delayed flowering, as well as the potential of CRISPR/Cas9 to reduce or eliminate transgene flow. We discuss system utility and efficacy, as well as necessary features for commercial adoption.

Keywords

Bioconfinement, transgenic plants, GE crops, gene flow, male sterility, transgene excision, delayed flowering

1. Introduction

In the US, more than 90% of commercially cultivated crops are genetically engineered (GE) (ISAAA 2019). These crops confer well-characterized transgenic traits, primarily insect and herbicide resistance, that have been engineered in a few crop species with low potential for gene flow (ISAAA 2019). Over the last three decades, transgene movement beyond cultivation has been a concern associated with GE crops (Ellstrand 2003, Chandler and Dunwell 2008, Rizwan, Hussain et al. 2019). Transgene flow may occur via the dispersal of pollen, seed, or vegetative-propagules, although pollen-mediated gene flow is commonly described as the major route of concern (Lu 2008, Rizwan, Hussain et al. 2019). Indeed, much research has focused on potential environmental impacts of transgene movement beyond cultivation (Millwood, Nageswara-Rao et al. 2017, Chang, Snow et al. 2018, Kesoju, Kramer et al. 2021) as well as the potential for the establishment of free-living transgenic populations (Arriola and Ellstrand 1996, Cao, Xia et al. 2009, Ellstrand 2018, Bauer-Panskus, Miyazaki et al. 2020). Transgene flow has been documented for a few crops, yet few free-living transgenic populations have been identified and there have been no reports of measurable negative environmental impacts (Reichman, Watrud et al. 2006, Zapiola, Campbell et al. 2007, Schafer, Ross et al. 2011, Greene, Kesoju et al. 2015, Ellstrand and Rieseberg 2016, NASEM 2016). This may be due to the reproductive biology of current GE crops and the area of cultivation. For example, many of these crops are highly domesticated, primarily self-pollinate, and are cultivated in regions with few to no sexually compatible wild relatives (Stewart Jr. 2021).

Gene flow is more likely to occur in crop species with high outcrossing frequencies often associated with wind-pollinated crops and obligate outcrossing species, such as switchgrass (*Panicum virgatum* L.). The probability of interspecific hybridization is also higher if there are wild relatives native to the region of cultivation (Warwick, Beckie et al. 2009). Additionally, certain transgenes and their conferred traits may also raise biosafety concerns in regard to environmental and ecological risk if they were to escape to unmanaged wild relative populations (Stewart Jr. 2021). For example, certain transgenes may provide advantageous traits, such as improved tolerance to abiotic stress conditions including drought, salinity, or low and high temperatures (Warwick, Beckie et al. 2009, Ding, Duan et al. 2014). There is concern that GE trait introgression into wild populations could provide a competitive advantage to transgenic

hybrids under stress conditions. This may lead to potential ecological consequences such as decreased biodiversity or possibly local population extinction (Haygood, Ives et al. 2003, Pilon and Prendeville 2004).

To address these concerns several transgene confinement approaches have been proposed. The implementation of physical or biological confinement (bioconfinement) strategies could prevent transgene flow. Physical confinement approaches utilize physical barriers such as enclosures, fencing, flower removal, and geographic isolation that provide spatial separation from sexually compatible crops and wild relatives. These approaches are often not practical solutions for large-scale commercialization in terms of physical labor and facility requirements (Kausch, Hague et al. 2010). Alternatively, biological confinement approaches offer feasible and sustainable strategies to contain transgenes. Since the early 1990s, bioconfinement approaches have been studied for their potential to limit or eliminate transgene flow. To date, none are commonly used in commercial crops. The adoption of a bioconfinement system would allow new GE crops to reach the market that may not otherwise. Numerous technologies have been proposed through various approaches including maternal inheritance, genome incompatibility (Daniell 2002), apomixis (Sandhu, Blount et al. 2010), cleistogamy (Fargue, Colbach et al. 2006, Faisal, Guo et al. 2018), and total sterility, but their application is limited. In this review, we survey efficient and widely applicable bioconfinement technologies such as male and seed sterility, transgene excision, and delayed flowering. These approaches have been achieved through genetic engineering, including type II clustered regularly interspaced short palindromic repeat/ CRISPR-associated protein 9 (CRISPR/Cas9) gene editing. We also discuss the most effective bioconfinement methods developed to date and compare their potential to eliminate transgene flow. Lastly, perspectives are provided for features we believe necessary for commercial adoption of a bioconfinement system.

2. Engineered Reproductive Sterility

2.1. Male Sterility

Engineered male sterility is a bioconfinement approach that can be designed to yield pollenless flowers or flowers with sterile transgenic pollen. Generally, GE male sterile phenotypes can be produced through the tightly-controlled and targeted expression of genes that yield products cytotoxic to reproductive tissues (Daniell 2002, Ding, Duan et al. 2014, Klocko 2020). To date, genes used for engineered male sterility include *barnase*, *Diphtheria toxin A-chain*, and the endonuclease *EcoRI*. Each have been expressed in several plant species with varying results which are described below.

2.1.1. Barnase-Induced Male Sterility

The bacterial gene *barnase* (from *Bacillus amyloliquefaciens*) has been used to generate male sterile plants. Barnase is a ribonuclease that, when expressed in plant cells, leads to RNA degradation and subsequent cell death (Mariani, Beuckeleer et al. 1990). When utilized for engineered male sterility, barnase expression has been targeted to tapetal cells which form the nutritive cell layer surrounding microspore mother cells (Echlin 1971, Goldberg, Beals et al. 1993). These cells are necessary for viable pollen development (Echlin 1971). When barnase is expressed in this manner, tapetal cell death prevents microspore development and confers total pollen ablation (Figure 1a) (Mariani, Beuckeleer et al. 1990). Fertility can be restored by crossing with restoration lines genetically engineered to express *barstar* (also from *Bacillus amyloliquefaciens*) which acts as a repressor of barnase (Hartley 1988, Mariani, Gossele et al. 1992).

Experimental results in tobacco (*Nicotiana tabacum* L.) and canola (*Brassica napus* L.) under the control of the tapetum-specific promoter TA29 (derived from tobacco) revealed barnase expression conferred completely pollenless plants in 92% and 77% of transgenic events, respectively (Mariani, Beuckeleer et al. 1990). Except for the absence of pollen, these transgenic plants were indistinguishable from nontransgenic tobacco controls in vegetative and floral morphology and displayed a similar growth rate and height through all developmental phases.

Using this approach, barnase has been used for male sterile plant production in several monocot species including wheat (*Triticum aestivum* L.) (Block, Debrouwer et al. 1997), creeping bentgrass (*Agrostis stolonifera* L.) (Luo, Kausch et al. 2005), and rice (*Oryza sativa* L.) (Hague, Dellaporta et al. 2012). In wheat, tapetum-specific promoters ca55 (derived from maize (*Zea mays* L.)), and T72 and E1 (derived from rice) were used to drive barnase expression (Block, Debrouwer et al. 1997). Out of 29 transgenic events, 11 were chosen for phenotypic analysis and crossing experiments. Two events contained barnase under the control of the promoter E1, six were under the control of the T72 promoter, and the remaining three events utilized the ca55 promoter. All transgenic events displayed a completely pollenless phenotype that was stably inherited across two generations of backcrossing except for one event (T72) which produced fertile transgenic progeny in cross-pollination experiments. These transgenic plants displayed normal phenotypes and were comparable to wild-type plants in height, leaf size, and tiller count. When expressed in creeping bentgrass using the tapetum-specific promoter TAP (derived from rice), barnase expression led to significantly reduced pollen production in 90% of transgenic lines (Luo, Kausch et al. 2005). For plants that produced pollen, viability was measured and all pollen grains tested were non-viable. No off-target expression was reported, and the male sterile phenotype was stably inherited in progeny when backcrossed to nontransgenic creeping bentgrass. When barnase was expressed in rice using the pollen-specific promoter Zm13 (derived from maize) (Hague, Dellaporta et al. 2012), transgenic events were crossed to nontransgenic rice. The resulting progeny produced 50% viable nontransgenic pollen. Transgenic pollen expressing barnase was nonviable which is expected for transformants with a single T-DNA copy. The cross-pollination experiments also revealed that pollen-specific expression of barnase had no effect on seed yield, and plants displayed normal growth and development.

Barnase has also been expressed in male-specific tissues of GE trees such as pine (*Pinus radiata* D. Don) (Zhang, Norris-Caneda et al. 2012), eucalyptus (*Eucalyptus globulus* Labill.) (Zhang, Norris-Caneda et al. 2012), and hybrid poplar (*Populus tremula* × *Populus tremuloides*) (Elorriaga, Meilan et al. 2014). In pine and eucalyptus, a male cone-specific promoter PrMC2 (derived from pine) was used to express the modified barnase coding sequence barnaseH102E (Zhang, Norris-Caneda et al. 2012). BarnaseH10E2 contained a single amino acid substitution which was expected to reduce RNase activity and prevent off-target expression. When

expressed in pine, 97% of transgenic events did not produce pollen. Nine of these transgenic events were studied under field conditions, where pollenless cones were observed over a 2-year study. Similar results were observed in eucalyptus where 95% of transgenic events analyzed were pollenless. Under field conditions, 16 of 17 transgenic events produced pollenless cones over a four-year study. Cone dissection revealed that one transgenic event produced pollen, but the pollen was non-viable. Additionally, vegetative growth and development were normal for both transgenic pine and eucalyptus. Barnase was also expressed in poplar hybrids (*Populus tremula* × *Populus tremuloides*) under the control of the tapetum- specific promoter TA29 (Elorriaga, Meilan et al. 2014). Throughout a four-year field study, two of six transgenic trees were completely pollenless and four transgenic trees released approximately six pollen grains per catkin (a flower cluster found in certain tree species) during the first two years of the study. Transgenic pollen production was significantly less compared to nontransgenic trees which produced approximately 73,000 to 85,000 pollen grains per catkin during the same period. In the final two years of the study, no pollen was produced in transgenic trees whereas 775,000 to 2,700,000 pollen grains per catkin were produced in nontransgenic controls. Unfortunately, plant growth was negatively affected as determined by stem volume ($\text{height} \times \text{diameter}^2$) measurements. Of 18 transgenic poplar events, 17 were significantly smaller compared to nontransgenic controls. Because growth of transgenic poplar expressing the β -glucuronidase (GUS) reporter gene was similar to that of nontransgenic controls, this phenotype was attributed to off-target barnase expression and possibly worsened by field conditions.

In summary, the targeted expression of barnase yields pollenless flowers or sterile pollen production and may be an effective strategy to prevent transgene flow with little to no negative effects on plant growth and development. Although examples of off-target expression were observed in few transgenic events, the use of barnase for bioconfinement is applicable across many crop species, stable across several generations, and can effectively confer male sterility. Nonetheless, this bioconfinement system can be further optimized through the use of synthetic promoters that provide increased sterility gene expression to targeted tissues for enhanced efficacy in a wide array of plant species.

2.1.2. Diphtheria Toxin A-Chain-Induced Male Sterility

Another well-characterized cytotoxic gene used for engineered male sterility is *Diphtheria toxin A-chain (DTA)* (from *Corynebacterium diphtheriae*). When expressed in cells, DTA catalyzes ADP-ribosylation and inactivation of elongation factor 2 (Pappenheimer 1977). This action inhibits protein synthesis resulting in cell death. Targeted *DTA* expression has been used for cell–cell interaction and functional analysis studies (Koltunow, Truettner et al. 1990, Twell 1995). The characterization of tapetum- and pollen-specific promoters through the expression of *DTA* in transgenic plants was found to yield pollenless and pollen-sterile plants with minimal to no effects on plant growth (Koltunow, Truettner et al. 1990, Twell 1995). For this reason, targeted *DTA* expression has been adopted as a bioconfinement strategy through male sterile plant production. Male fertility can be restored by crossing with transgenic lines engineered to express a diphtheria toxin repressor (DTxR) (also from *Corynebacterium diphtheriae*) (Fourel, Phalipon et al. 1989).

Tapetum-specific expression of *DTA* by the promoter TA29 was analyzed in seven transgenic tobacco plants, all of which contained a single T-DNA copy (Koltunow, Truettner et al. 1990). From these transformants, anther development was identical to that of nontransgenic control plants except for the destruction of the tapetum and collapse of the pollen sac resulting in pollenless plants. No off-target expression or negative effects on growth were reported suggesting TA29 promoter activity is localized to tapetum cells.

Another study used the pollen-specific promoter Lat52 derived from tomato (*Solanum lycopersicum* L.) to drive the expression of *DTA* in 20 transgenic tobacco events (Twell 1995). Of the 20 events, 16 were male-sterile and displayed one of two phenotypes. Half of the events displayed a pollenless phenotype, whereas the other half had a significant reduction in pollen grains of which 50% were nonviable. Two transgenic events that produced 46.3% and 50.9% nonviable pollen were selected for further analyses, and the pollen-sterile phenotype was stably inherited through four generations of self-pollination. Additionally, two of the 20 transgenic events displayed off-target expression which resulted in decreased flower size.

Both tapetum- or pollen-DTA expressing plants are phenotypically similar to barnase expressing male-sterile plants. Tapetum- and pollen-specific expression of DTA yielded male-sterile tobacco with few to no off-target effects (Koltunow, Truettner et al. 1990, Twell 1995). However, concerns may arise in regard to the use of DTA in crops. Although a single molecule of DTA can induce cell death, diphtheria toxin B chain is required to transport DTA into cells (Pappenheimer 1977, Yamaizumi, Mekada et al. 1978). The use of DTA in this manner may raise concerns regarding potential pollen-feeding insects and human health effects even though DTA is not highly toxic alone. Despite its potential, these concerns may create roadblocks when used as a bioconfinement strategy.

2.1.3. *EcoRI* Endonuclease-Induced Male Sterility

Another gene studied for bioconfinement purposes is the *EcoRI* endonuclease (from *Escherichia coli*) which has been expressed in pollen to confer selective male sterility in tobacco (Figure 1b) (Millwood, Moon et al. 2016). *EcoRI* endonuclease is a well-characterized type II restriction endonuclease that creates a double-stranded break at the 5'-GAATTC-3' recognition sequence (Wilson 1988). The expression of *EcoRI* in pollen causes repeated double-stranded breaks in the pollen nuclear genome which induces cell death (Millwood, Moon et al. 2016).

Pollen-specific expression of *EcoRI* endonuclease by the Lat52 promoter conferred selective male sterility in tobacco, and all transformants displayed over 99% and up to 100% bioconfinement efficacy when used as the pollen donor source and crossed to nontransgenic tobacco plants under greenhouse and field conditions (Millwood, Moon et al. 2016). Additionally, bioconfinement efficacy was stable across generations and transgenic plants were indistinguishable from nontransgenic plants in regard to growth and vegetative morphology. Thus far, this approach has been characterized in tobacco, but more research is needed in other species since recognition-site frequency of restriction endonucleases and promoter activity can vary. Furthermore, this approach is advantageous for pollen-feeding insects since it permits pollen production, whereas other male-sterility approaches may be food-source limiting.

2.2. Conditional Seed Sterility

Conditional seed sterility is another promising bioconfinement approach that was originally designed to protect intellectual property in genetically engineered plants (Oliver, Quisenberry et al. 1998, Oliver, Quisenberry et al. 1999, Oliver, Quisenberry et al. 1999). Such approaches are referred to as genetic use restriction technologies (GURTs). GURTs are often designed as an inducible system that activates sterility gene expression under specific conditions (Sang, Millwood et al. 2013, Stewart Jr. 2021). One of the earliest GURTs, termed Genesafe technology, was originally developed in the 1990's to prevent the unauthorized use of transgenic seeds by inducing seed sterility in the next generation (Oliver, Quisenberry et al. 1998, Oliver, Quisenberry et al. 1999, Oliver, Quisenberry et al. 1999, FAO 2001). Although promising, Genesafe was never commercialized. This was partially due to public opposition since growers could not save and reuse seeds (Sang, Millwood et al. 2013, Lombardo 2014, Stewart Jr. 2021). However, this approach holds great potential as a bioconfinement system specifically for vegetatively or clonally propagated crops such as switchgrass, certain hemp cultivars (*Cannabis sativa* L.), and poplar.

GURT designs often require activation by an external stimulus, such as a chemical or temperature treatment, which would activate an engineered genetic circuit and inhibit seed germination in the subsequent generation (Figure 2). These GURTs are comprised of four components: (i) a genetic switch which is inducible and can activate or repress the trait switch; (ii) a trait switch which is typically a recombinase gene responsible for removing a blocker sequence; (iii) a blocker sequence flanked by recombinase recognition sites (located between the seed-specific promoter and sterility gene); and (iv) a sterility gene controlled by a seed-specific promoter (Oliver and Hake 2013, Sang, Millwood et al. 2013). Altogether, the application of a stimulus activates the genetic switch, which amplifies the trait switch and removes a blocker sequence to designs, Genesafe holds promise as a bioconfinement system that can be incorporated into bring the seed-specific promoter and sterility gene in close proximity (Sang, Millwood et al. 2013, Lombardo 2014). Seed-specific expression of the sterility gene renders transgenic seeds sterile.

The original Genesafe prototype was a repressible system and required hybridization of two parent plants (P1 and P2) to produce inactive, inducible seeds carrying the entire Genesafe system (Oliver, Quisenberry et al. 1998, Oliver and Hake 2013). P1 carried the seed-specific promoter and sterility gene divided by the blocker sequence that produces a repressor protein. P2 carried a recombinase gene that, without application of a stimulus, was repressible by the blocker sequence in P1 when hybridized. After hybridization, a stimulus treatment would release the repressor, and the recombinase would excise the blocker sequence. As a result, the sterility gene would be expressed in the next generation of seeds. This design was tested in tobacco using late embryo-specific promoters LEA4A and LEA14 derived from cotton (*Gossypium herbaceum* L.) to drive expression of the sterility gene *saporin* (from *Saponaria officinalis*), a ribosome-inactivating protein that inhibits protein synthesis when expressed in plant cells (Ferrerias, Barbieri et al. 1993). The blocker sequence was successfully excised using the recombinase system Cre/*loxP*; however, seed-sterility was not observed (Oliver, Luo et al. 2004). This was attributed to the inefficient activity of the cotton-derived promoter in tobacco, as well as insufficient *saporin* production. This design was not further tested because a new and simpler design was created which used a single parent to carry the Genesafe system. The new design was chemically inducible, where application of a stimulus would permit expression of the recombinase gene and excise the blocker sequence (Oliver, Quisenberry et al. 1998, Oliver and Hake 2013). As a result, the seed-specific promoter would express the sterility gene and was expected to produce non-germinable seeds in the following generation. This design can be altered to use other components to make up the chemically- inducible system but was described with the Cre/*loxP* recombinase system, LEA promoters, and the *saporin* gene. However, no data were provided for seed sterility.

Although Genesafe was not previously commercialized for its original purpose, it could be immensely useful for bioconfinement if used in clonally or vegetatively propagated crops. Improved tools and system components are available that can be used to optimize the system. Specifically, efficient recombinase systems can be employed such as gene-deletor technology (described in Section 3) which has displayed up to 100% transgene excision from pollen and/or seeds (Luo, Duan et al. 2007). Improved inducible- and tissue-specific synthetic promoters are available which have been optimized for strong inducible and highly specific spatiotemporal

expression (Liu and Stewart Jr 2016, Ali and Kim 2019, Huang, Kosentka et al. 2021, Yang, Chaffin et al. 2022). With optimized components and new designs, Genesafe holds promise as a bioconfinement system that can be incorporated into a wide array of crops.

3. Transgene Excision

Transgene excision has been proposed as a strategy to eliminate transgene flow. This strategy utilizes inducible- or tissue-specific promoters to express a site-specific recombinase. The DNA targeted for excision is flanked by site-specific recombinase recognition sites during design and cloning of the transformation construct. Depending on the placement of the recognition sites, this strategy can be used to remove transgenes such as selectable marker genes (SMGs) or entire T-DNA cassettes. The excised DNA is degraded by non-specific nucleases within the cell, and components that were not located between the recognition sites would remain in the genome including the left and right T-DNA borders and one recombinase recognition site (Figure 3a) (Moon, Abercrombie et al. 2011, Yau and Stewart Jr. 2013).

3.1. Transgene Excision via Cre/*loxP* Recombinase

The most studied and best characterized bidirectional recombinase systems include Cre-*loxP* (from bacteriophage P1), FLP-*FRT* (from *Saccharomyces cerevisiae*), and R-RS (from *Zygosaccharomyces rouxii*) (Araki, Jearnpipatkul et al. 1985, Senecoff, Bruckner et al. 1985, Hoess and Abremski 1990). Tyrosine recombinases that recognize relatively short sequences (34–58 base pairs) are used in these systems to mediate DNA cleavage (Araki, Jearnpipatkul et al. 1985, Senecoff, Bruckner et al. 1985, Hoess and Abremski 1990). Dependent on the orientation of the recognition sites relative to each other, integration, excision, or inversion of targeted sequences of DNA can occur (Gidoni, Srivastava et al. 2008). Although all of these outcomes have been observed, excision and degradation of the targeted DNA is observed most often since it is kinetically favored (Gidoni, Srivastava et al. 2008).

The Cre-*loxP* recombinase system has been studied in *Arabidopsis* (*Arabidopsis thaliana* L.) (Mlynárová, Conner et al. 2006) and tobacco (Mlynárová, Conner et al. 2006, Kopertekh, Schulze et al. 2010, Polóniová, Jopčík et al. 2015, Hamzeh, Motallebi et al. 2016) studies, where pollen was targeted for transgene excision to prevent pollen-mediated gene flow (Mlynárová,

Conner et al. 2006, Polóniová, Jopčík et al. 2015). In *Arabidopsis* and tobacco greenhouse experiments, the microspore-specific promoter NTM19 (derived from tobacco) was used to drive Cre expression and produce transgene-free pollen (Mlynárová, Conner et al. 2006). Transgenic *Arabidopsis* and tobacco plants were used in self- and cross-pollination experiments, followed by backcross experiments. Backcross experiments displayed 100% transgene excision in three *Arabidopsis* lines. In transgenic tobacco, over 99% transgene excision was observed in five lines. In addition to high transgene excision efficiency, plants displayed normal growth and development. Another study used the Cre/*loxP* recombinase system to excise the SMG from transgenic tobacco through the use of the pollen- and embryo-specific promoter DLL (derived from *Arabidopsis*) to drive Cre expression (Polóniová, Jopčík et al. 2015). Eight transgenic tobacco lines were self-pollinated and 99% of the progeny displayed partial or complete excision of the SMG. However, there was notable variation across transgenic lines, ranging from 3.8% to 81.5% excision efficiency. Four of these lines were used in cross-pollination experiments as pollen donors to nontransgenic tobacco, where complete or partial excision of the SMG was observed in 99% of plants. These experiments also displayed variation across transgenic lines which ranged from 40% up to 96.2% excision efficiency. High excision efficiency has been observed and supports the potential of the recombinase system for transgene excision; however, the variation among lines should be addressed by utilizing promoters that provide consistent and targeted high expression levels.

Seed-specific expression of the Cre/*loxP* recombinase system was also studied to prevent seed-mediated transgene flow. In tobacco (*Nicotiana benthamiana* L.), the seed-specific promoter napin (derived from canola) was used to drive Cre expression and determine the efficiency of SMG excision from seeds (Kopertekh, Schulze et al. 2010). Tobacco was co-transformed with two plasmids, the first of which contained the SMG flanked by the *loxP* recognition sites and the second construct contained Cre driven by the napin promoter. A total of 10 lines were co-transformed, self-pollinated, and the progeny were subjected to analysis. Excision of the SMG occurred in all 10 lines, eight of which displayed 100% excision efficiency and two lines displayed 76.8% and 93.1% excision efficiency. Additionally, all transgenic lines were normal in regard to vegetative and reproductive morphology. The expression of the Cre/*loxP* recombinase system was also studied in tobacco (*Nicotiana tabacum* L.) under the control of the seed-specific

promoter BcNA1 (derived from canola) to remove the SMG (Hamzeh, Motallebi et al. 2016). Eight transgenic lines were obtained and self-pollinated, and progeny were subjected to a germination assay on selection media to determine whether excision had occurred. In four of the transgenic lines, 100% excision was observed. In the remaining transgenic lines, excision occurred at a rate of 25.4%, 29.9%, 59.6%, and 71.8%. Additionally, all transgenic lines appeared normal in regard to plant growth and fertility. In transgenic canola, Cre expression was driven by the embryo-specific promoter USP88 (derived from *Vicia faba* L.) to excise the SMG and produce SMG-free progeny (Ye, Vaghchhipawala et al. 2022). Progeny from eight transgenic lines were analyzed and an average of 32.7% excision efficiency was observed, none of which were reported to have adverse morphological features. Although these lines were able to produce SMG-free progeny, the average excision rates were relatively low compared to other studies. To improve this system, stronger and more tightly-controlled promoters should be selected to drive Cre expression.

Both pollen- and seed-specific expression of the Cre/*loxP* recombinase system have displayed up to 100% transgene excision in multiple transgenic lines. However, significantly lower efficiency has been observed in several lines. To reduce variability among lines, it is apparent that design changes, specifically for promoters that drive recombinase expression, as well as further testing and optimization are needed to ensure high and consistent levels of transgene excision in all transgenic lines. If consistency is achieved, the Cre/*loxP* recombinase system may be a useful bioconfinement tool.

3.2. Transgene Excision from Pollen via CinH-RS2

Other well-characterized recombinase systems are CinH-RS2 (from *Acinetobacter* plasmids pKLH2 and pKLH204) and ParA-MRS (from the broad host range plasmid RK2), which are unidirectional recombinase systems and use serine recombinases to irreversibly excise targeted sequences of DNA (Sobecky, Easter et al. 1996, Kholodii, Mindlin et al. 2004, Liu, Yuan et al. 2013). Recombinases of the CinH-RS2 and ParA-MRS recombinase systems recognize 113 and 133 base pair sequences, respectively, which are much longer than the tyrosine-derived recombinase systems previously described and are expected to decrease potential off-target DNA cleavage (Thomson and Ow 2006, Shao, Blechl et al. 2017).

Transgene excision from pollen has been studied in tobacco using the CinH-RS2 recombinase system (Moon, Abercrombie et al. 2011). In transgenic tobacco, the pollen specific Lat52 promoter (derived from tomato) was used to drive expression of CinH recombinase. The RS2 recognition sites flanked all transgenes in the construct and were within the T-DNA borders. A green fluorescent protein (GFP) visual marker gene was located in this region and expression was driven by the Lat59 pollen-specific promoter (derived from tomato). In this design, successful excision would prevent expression of GFP in pollen. Pollen grains from five transgenic lines with a single T-DNA copy were collected and analyzed, three of which displayed less than 1% GFP-expressing pollen and the other two lines displayed 1.94% and 17.48% GFP-expressing pollen. Furthermore, transgenic plants were phenotypically indistinguishable from nontransgenic controls in regard to plant growth and reproduction. Thus, the authors concluded that highly efficient transgene excision is possible, and this approach holds promise for transgene bioconfinement.

Based on these results, the CinH-RS2 recombinase system can effectively excise trans- genes from pollen without negatively impacting plant growth. The pollen analysis showed promising results. However, cross-pollination experiments should be performed to confirm that transgenes are not present in the progeny. Further research is also needed to evaluate design and system components for improvements to achieve a 100% excision efficiency as observed in other designs.

3.3. Transgene Excision from Pollen and/or Seeds via Fused loxP-FRT Recombinase System

Gene-deletor technology is a term used to describe a site-specific recombinase system that uses fused *loxP-FRT* recognition sites on either side of the transgene (s), along with expression of either Cre or FLP recombinase genes (Luo, Duan et al. 2007). Cre and/or FLP recombinases were driven by pollen-specific promoters Lat52 (derived from tomato) or BGP1 (derived from *Brassica campestris*), or pollen- and seed-specific promoter PAB5 (derived from *Ara- bidopsis*) to remove transgenes from pollen and/or seeds of GE tobacco. Gene-deletor technology displayed 100% transgene excision in pollen and seeds from multiple transgenic tobacco events. Specifically, analysis of progeny produced from self- and cross-pollination experiments revealed that constructs with fused *loxP-FRT* recognition sequences and pro- moter:recombinase

combinations BGP1:FLP, PAB5:FLP, and Lat52:Cre displayed 100% excision efficiency in multiple lines. Additionally, no deleterious effects on vegetative or reproductive growth were reported. In the same study, the promoter:recombinase combination BGP1:FLP was used with only the *FRT* recognition site and displayed 99% excision efficiency, a slightly lower rate than when both *loxP-FRT* recognition sites were present (100%). The study also examined how the simultaneous expression of Cre and FLP recombinases driven by Lat52 affected transgene excision in the presence of fused *loxP-FRT* recognition sites. This design produced up to 45% transgene excision, a significantly lower rate from Cre expression alone.

These studies suggest gene-deletor technology increased excision efficiency and can effectively eliminate transgene flow in tobacco without negatively impacting plant growth and development. The promoter:recombinase combinations with high performance should be tested in other crop species to determine if efficacy is comparable.

3.4. CRISPR/Cas9-Mediated Transgene Excision

In recent years, type II clustered regularly interspaced short palindromic repeat/CRISPR-associated protein 9 (CRISPR/Cas9) has been utilized for transgene excision. It requires two Cas9 cleavage target sites (CTS) on either side of the target DNA and a guide RNA (gRNA)-Cas9 assembly that recognizes and cleaves the CTSs. The excised DNA is degraded by non-specific nucleases in the cell. Plasmid construct components that were not located between the CTSs would remain in the genome (Figure 3b) (He, Mudgett et al. 2022).

In one study, CRISPR/Cas9 transgene excision efficacy was examined for the excision of SMGs in transgenic rice (Srivastava, Underwood et al. 2017). To do so, callus of transgenic rice that constitutively expressed GUS was transformed with a CRISPR/Cas9 construct. The cleavage target sites were located at the ends of the 1.6 kb *GUS* gene. Since the *GUS* gene is expected to be removed from the genome, successful excision would be determined by lack of GUS expression. In this experiment, a total of 34 plants were produced from 12 transgenic events. Three of these events displayed 100% transgene excision. Despite high excision efficiency, the authors report that normal vegetative growth was observed in four of 34 transgenic plants and only two plants produced a small quantity of seeds. No specific details were provided for the

adverse phenotypes and off-target Cas9 activity. While this study displayed 100% transgene excision from multiple transgenic events, reports of no phenotypic and off-target effects are necessary to eliminate concerns of potential negative outcomes.

Another study also aimed to excise SMGs from rice through CRISPR/Cas9-mediated transgene excision (Tan, Wang et al. 2022). The promoter Pssi (derived from rice) was expected to express strongly in stem, shoot tip, and inflorescence tissue, and was therefore selected to drive Cas9 nuclease expression. Transgenic rice was generated to constitutively express GUS; however, the *GUS* gene was separated into two segments denoted as ‘GU’ and ‘US’, interrupted by a CRISPR/Cas9 cassette and a selectable marker gene that acted as a blocker sequence flanked by CTSs. Additionally, both ‘GU’ and ‘US’ carried a 1027 bp ‘U’ sequence as a homologous region, inducing CRISPR/Cas9-mediated homology-directed repair following excision of the blocker sequence resulting in a functional *GUS* reporter gene if excision was successful. Transgene excision occurred in 11 of 15 transgenic lines. In lines where excision was observed, T₁ lines were produced and three were randomly selected for further analysis. In one line, 6.7% homozygous excision and 66.7% heterozygous excision was observed. Another displayed 73.3% homozygous excision and 13.3% heterozygous excision, and the last line displayed 86.7% homozygous excision. In contrast to the study previously described, no adverse effects were observed in regard to plant growth and reproduction.

CRISPR/Cas9-mediated transgene excision displayed high excision efficiency in rice. However, there are limited studies that have used CRISPR/Cas9 for transgene excision, one of which resulted detrimental effects on plant morphology and reproduction. The use of tissue-specific promoters to drive Cas9 expression, such as those that target pollen or seeds, may help reduce potential off-target effects in vegetative tissues. We believe CRISPR/Cas9-mediated transgene excision can be a useful bioconfinement system, but further optimization is needed.

4. Delayed Flowering

The transition from vegetative to flowering phase is a complex regulatory system controlled by microRNAs (miRNAs), transcription factor (TF) genes, and floral identity genes. miRNAs are a class of small regulatory RNAs that regulate gene expression and RNA silencing through RNA degradation or translational inhibition (Brodersen, Sakvarelidze-Achard et al. 2008, Jonas and Izaurralde 2015). The overexpression of certain miRNAs can silence floral identity genes that may result in shifts in flowering time or reduced flower production. This approach has the potential to shift flowering time to prevent complete or partial overlap with the flowering time of sexually compatible plants, thus preventing or limiting transgene flow (Brookes, Barfoot et al. 2004, Ortega-Molina 2004).

4.1 Overexpression of miR156 Delayed Flowering

miRNA 156 (miR156) is one of several miRNAs responsible for regulating the floral transition in the shoot apical meristem (Luo, Guo et al. 2013). miR156 expression is regulated by the age-dependent pathway, and expression is high in seedlings and declines as the plant matures. It has been characterized to repress transcripts of the *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPL)* TF family that upregulate floral identity genes *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1 (SOC1)* and *LEAFY (LFY)* (Luo, Guo et al. 2013). Thus, the overexpression of miR156 is expected to prevent the upregulation of *SOC1* and *LFY* as the plant matures, which consequently inhibits/delays the transition from the vegetative to flowering phase. For this reason, the overexpression of miR156 has been studied in switchgrass under greenhouse and field conditions to determine its potential for transgene bioconfinement via delayed flowering (Fu, Sunkar et al. 2012, Johnson, Millwood et al. 2017).

In a greenhouse study, switchgrass plants engineered to overexpress miR156 (from rice) were categorized as low, moderate, and high based on transgene expression levels (Fu, Sunkar et al. 2012). Transgenic switchgrass with low expression of miR156 flowered at the same time and displayed normal reproductive growth compared to nontransgenic controls. These plants also displayed an increase in tiller number and vegetative biomass. Plants considered to have moderate expression of miR156 did not flower after 360 days and were also observed to have an

increase in tiller number and vegetative biomass. However, these lines had several notable defects when compared to nontransgenic controls, including reduced plant height, decreased internode diameter, leaf blade width, and leaf sheath length. Similar to the plants categorized as moderate expression, transgenic switchgrass with high expression of miR156 did not flower after 360 days. Additionally, high expression of miR156 led to an increase in tiller number and severe defects including a reduction in plant height and biomass, as well as significantly decreased internode diameter and length, leaf sheath length, and leaf blade length and width compared to nontransgenic switchgrass (Fu, Sunkar et al. 2012).

Subsequently, a two-year field study was conducted using low and moderate miR156 overexpression lines to determine flowering time and biomass production under field conditions (Johnson, Millwood et al. 2017). Compared to nontransgenic controls, low overexpression lines were highly variable in flowering time during the first year of establishment post-transplanting. Once these lines were established, there was no delay in flowering during the second year. Reduced biomass was observed in some plants, and fewer flowers and seeds were observed in all low expression lines. However, when flowering was monitored for moderate overexpression lines, many plants did not flower in either year of the field study. Of the plants that did flower, timing was delayed by 12 weeks in the first year and two weeks in the second year. Also, some plants displayed reduced biomass production whereas others were unaffected. Additionally, decreased flower and seed production was observed in all moderate overexpression plants.

The overexpression of miR156 in switchgrass results in delayed and decreased flowering. Several phenotypes observed in some of the low and moderate overexpression lines are desirable, including significantly delayed flowering. To optimize this system, miR156 expression levels may be fine-tuned through promoter selection that yields expression levels high enough to delay flowering without a negative impact on plant fitness. Delayed or decreased flowering could also be achieved by downregulating the *SPL* gene (s) that are targets of miR156, including *SPL 2/10* and *SPL 3/4/5* (Johnson, Millwood et al. 2017). Temporal downregulation of these genes individually or in combination prior to floral transition could be achieved through RNA interference (RNAi) or CRISPR interference (CRISPRi) and result in delayed flowering without negatively affecting plant biomass production. Additionally, expression of other miRNAs known to regulate floral development can be altered to confer delayed flowering (Trumbo, Zhang et al.

2015). For example, miR172 expression is negatively correlated with miR156 expression, and overexpression of miR172 has been observed to accelerate flowering time in *Arabidopsis* (Aukerman and Sakai 2003). Downregulation of miR172 could also confer similar phenotypes to those of miR156 overexpression and be useful for bioconfinement purposes.

5. Perspectives

GE crops that have been approved for commercial release are considered low risk of transgene flow and generally safe (NASEM 2016). It is likely that regulatory bodies will require new GE crops with a high risk of gene flow or those that contain novel transgenic traits be put through lengthy and costly biosafety assessments before nonregulated status is approved. These crops would benefit from the inclusion of a bioconfinement system that significantly limits or eliminates transgene flow. In this review, we have highlighted the most prospective bioconfinement designs and many of these approaches have been tested in several plant species. Several designs have shown promise with nearly 100% bioconfinement efficacy. Should a system be adopted, the design must meet several key criteria. First, the system should be highly reliable, consistently reaching near 100% efficiency to reduce transgene exposure to unintended hosts. In addition, system components should not pose biosafety concerns regarding ecological risks, human health, or non-target organisms that may be exposed such as pollen-feeding insects. Lastly, it is vital that the bioconfinement system have no negative impact on crop yield, whether that is vegetative biomass, seed, or fruit yield. Many of the strategies reviewed here meet most of these criteria yet further improvements are needed for more robust and targeted gene expression. Indeed, the future is bright for this area of research and we expect advances in synthetic biology to be leveraged so new tools, such as inducible synthetic promoters, can be fine-tuned for robust gene expression in target tissue as well as improved efficiency of recombinases and CRISPR/Cas9-based methods. Although no technology is without risks, we are optimistic that several bioconfinement systems will soon be adopted that reduce biosafety concerns of stakeholders, regulators, and the public.

References

- Ali, S. and W. C. Kim (2019). "A fruitful decade using synthetic promoters in the improvement of transgenic plants." Frontiers in Plant Science **10**: 1433.
- Araki, H., et al. (1985). "Molecular and functional organization of yeast plasmid pSR1." Journal of Molecular Biology **182**(2): 191-203.
- Arriola, P. E. and N. C. Ellstrand (1996). "Crop-to-weed gene flow in the genus *Sorghum* (Poaceae): Spontaneous interspecific hybridization between johnsongrass, *Sorghum halepense*, and crop sorghum, *S. bicolor*." American Journal of Botany **83**(9): 1153-1159.
- Aukerman, M. J. and H. Sakai (2003). "Regulation of flowering time and floral organ identity by a microRNA and its *APETALA2*-like target genes." The Plant Cell **15**(11): 2730-2741.
- Bauer-Panskus, A., et al. (2020). "Risk assessment of genetically engineered plants that can persist and propagate in the environment." Environmental Sciences Europe **32**(1): 1-15.
- Block, M. D., et al. (1997). "The development of a nuclear male sterility system in wheat. Expression of the barnase gene under the control of tapetum-specific promoters." Theoretical and Applied Genetics **95**(1): 125-131.
- Brodersen, P., et al. (2008). "Widespread translational inhibition by plant miRNAs and siRNAs." Science **320**(5880): 1185-1190.
- Brookes, G., et al. (2004). "Genetically modified maize: Pollen movement and crop co-existence." PG Economics: 3-20.
- Cao, Q. J., et al. (2009). "Performance of hybrids between weedy rice and insect-resistant transgenic rice under field experiments: Implication for environmental biosafety assessment." Journal of Integrative Plant Biology **51**(12): 1138-1148.
- Chandler, S. and J. M. Dunwell (2008). "Gene flow, risk assessment and the environmental release of transgenic plants." Critical Reviews in Plant Sciences **27**(1): 25-49.
- Chang, H., et al. (2018). "Extent of pollen-mediated gene flow and seed longevity in switchgrass (*Panicum virgatum* L.): Implications for biosafety procedures." Biomass and Bioenergy **109**: 114-124.
- Daniell, H. (2002). "Molecular strategies for gene containment in transgenic crops." Nature Biotechnology **20**(6): 581-586.
- Ding, J., et al. (2014). "Molecular strategies for addressing gene flow problems and their potential applications in abiotic stress tolerant transgenic plants." Critical Reviews in Plant Sciences **33**(2-3): 190-204.

- Echlin, P. (1971). The role of the tapetum during microsporogenesis of angiosperms. Pollen, Elsevier Ltd: 41-61.
- Ellstrand, N. C. (2003). "Current knowledge of gene flow in plants: Implications for transgene flow." Philosophical Transactions: Biological Sciences **358**(1434): 1163-1170.
- Ellstrand, N. C. (2018). "'Born to run'? Not necessarily: Species and trait bias in persistent free-living transgenic plants." Frontiers in Bioengineering and Biotechnology **6**: 88.
- Ellstrand, N. C. and L. H. Rieseberg (2016). "When gene flow really matters: Gene flow in applied evolutionary biology." Evolutionary Applications **9**(7): 833-836.
- Elorriaga, E., et al. (2014). "A tapetal ablation transgene induces stable male sterility and slows field growth in *Populus*." Tree Genetics & Genomes **10**(6): 1583-1593.
- Faisal, S., et al. (2018). "Morphological and genetic analysis of a cleistogamous mutant in rapeseed (*Brassica napus* L.)." Genetic Resources and Crop Evolution **65**(2): 397-403.
- FAO (2001). "Potential Impacts of Genetic Use Restriction Technologies (GURTs) on Agricultural Biodiversity and Agricultural Production Systems."
- Fargue, A., et al. (2006). "Predictive study of the advantages of cleistogamy in oilseed rape in limiting unwanted gene flow." Euphytica **151**(1): 1-13.
- Ferreras, J., et al. (1993). "Distribution and properties of major ribosome-inactivating proteins (28 S rRNA N-glycosidases) of the plant *Saponaria officinalis* L. (Caryophyllaceae)." Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression **1216**(1): 31-42.
- Fourel, G., et al. (1989). "Evidence for direct regulation of diphtheria toxin gene transcription by an Fe²⁺-dependent DNA-binding repressor, *DtoxR*, in *Corynebacterium diphtheriae*." Infection and Immunity **57**(10): 3221-3225.
- Fu, C., et al. (2012). "Overexpression of miR156 in switchgrass (*Panicum virgatum* L.) results in various morphological alterations and leads to improved biomass production." Plant Biotechnology Journal **10**(4): 443-452.
- Gidoni, D., et al. (2008). "Site-specific excisional recombination strategies for elimination of undesirable transgenes from crop plants." In Vitro Cellular & Developmental Biology- Plant **44**(6): 457-467.
- Goldberg, R. B., et al. (1993). "Anther development: Basic principles and practical applications." The Plant Cell **5**(10): 1217-1229.
- Greene, S. L., et al. (2015). "Occurrence of transgenic feral alfalfa (*Medicago sativa* subsp. *sativa* L.) in alfalfa seed production areas in the United States." PLoS One **10**(12): e0143296.

Hague, J. P., et al. (2012). "Pollen sterility—A promising approach to gene confinement and breeding for genetically modified bioenergy crops." Agriculture **2**(4): 295-315.

Hamzeh, S., et al. (2016). "Efficient seed-specifically regulated autoexcision of marker gene (*nptII*) with inducible expression of interest gene in transgenic *Nicotiana tabacum*." Turkish Journal of Biology **40**(1): 1-11.

Hartley, R. W. (1988). "Barnase and barstar: Expression of its cloned inhibitor permits expression of a cloned ribonuclease." Journal of Molecular Biology **202**(4): 913-915.

Haygood, R., et al. (2003). "Consequences of recurrent gene flow from crops to wild relatives." Proceedings of the Royal Society of London. B: Biological Sciences **270**(1527): 1879-1886.

He, Y., et al. (2022). "Advances in gene editing without residual transgenes in plants." Plant Physiology **188**(4): 1757-1768.

Hoess, R. H. and K. Abremski (1990). The Cre-*lox* recombination system. Nucleic Acids and Molecular Biology. F. Eckstein and D. M. J. Lilley, Springer, Berlin, Heidelberg. **4**: 99-109.

Huang, D., et al. (2021). "Synthetic biology approaches in regulation of targeted gene expression." Current Opinion in Plant Biology **63**: 102036.

ISAAA (2019). "Global Status of Commercialized Biotech/GM Crops in 2019: Biotech Crops Drive SocioEconomic Development and Sustainable Environment in the New Frontier." ISAAA Brief No. 55. Retrieved 5 September 2022, from <https://www.isaaa.org/resources/publications/briefs/55/executivesummary/default.asp>.

Johnson, C. R., et al. (2017). "Field-grown miR156 transgenic switchgrass reproduction, yield, global gene expression analysis, and bioconfinement." Biotechnology for Biofuels **10**(1): 1-12.

Jonas, S. and E. Izaurralde (2015). "Towards a molecular understanding of microRNA-mediated gene silencing." Nature Reviews Genetics **16**(7): 421-433.

Kausch, A. P., et al. (2010). "Transgenic perennial biofuel feedstocks and strategies for bioconfinement." Biofuels **1**(1): 163-176.

Kesoju, S. R., et al. (2021). "Gene flow in commercial alfalfa (*Medicago sativa* subsp. *sativa* L.) seed production fields: Distance is the primary but not the sole influence on adventitious presence." PLoS One **16**(3): e0248746.

Kholodii, G., et al. (2004). "Translocation of transposition-deficient (TndPKLH2-like) transposons in the natural environment: Mechanistic insights from the study of adjacent DNA sequences." Microbiology **150**(4): 979-992.

Klocko, A. L. (2020). "Strategies to facilitate containment of genetically engineered crops." CABI Reviews.

Koltunow, A. M., et al. (1990). "Different temporal and spatial gene expression patterns occur during anther development." The Plant Cell **2**(12): 1201-1224.

Kopertekh, L., et al. (2010). "Cre-mediated seed-specific transgene excision in tobacco." Plant Molecular Biology **72**(6): 597-605.

Liu, W. and C. N. Stewart Jr (2016). "Plant synthetic promoters and transcription factors." Current Opinion in Biotechnology **37**: 36-44.

Liu, W., et al. (2013). "Advanced genetic tools for plant biotechnology." Nature Reviews Genetics **14**(11): 781-793.

Lombardo, L. (2014). "Genetic use restriction technologies: A review." Plant Biotechnology Journal **12**(8): 995-1005.

Lu, B. R. (2008). "Transgene escape from GM crops and potential biosafety consequences: An environmental perspective." Collection of Biosafety Reviews **4**: 66-141.

Luo, H., et al. (2005). "Controlling transgene escape in GM creeping bentgrass." Molecular Breeding **16**(2): 185-188.

Luo, K., et al. (2007). "'GM-gene-deletor': Fused *loxP-FRT* recognition sequences dramatically improve the efficiency of FLP or CRE recombinase on transgene excision from pollen and seed of tobacco plants." Plant Biotechnology Journal **5**(2): 263-374.

Luo, Y., et al. (2013). "Evolutionary conservation of microRNA regulatory programs in plant flower development." Developmental Biology **380**(2): 133-144.

Mariani, C., et al. (1990). "Induction of male sterility in plants by a chimaeric ribonuclease gene." Nature **347**(6295): 737-741.

Mariani, C., et al. (1992). "A chimaeric ribonuclease-inhibitor gene restores fertility to male sterile plants." Nature **357**(6377): 384-387.

Millwood, R., et al. (2016). "Engineered selective plant male sterility through pollen-specific expression of the *EcoRI* restriction endonuclease." Plant Biotechnology Journal **14**(5): 1281-1290.

Millwood, R., et al. (2017). "Pollen-mediated gene flow from transgenic to non-transgenic switchgrass (*Panicum virgatum* L.) in the field." BMC Biotechnology **17**(1): 1-10.

Mlynárová, L., et al. (2006). "Directed microspore-specific recombination of transgenic alleles to prevent pollen-mediated transmission of transgenes." Plant Biotechnology Journal **4**(4): 445-452.

- Moon, H. S., et al. (2011). "Transgene excision in pollen using a codon optimized serine resolvase CinH-RS2 site-specific recombination system." Plant Molecular Biology **75**(6): 621-631.
- NASEM (2016). Genetically engineered crops: Experiences and prospects. National Academies of Sciences, Engineering, and Medicine. Washington, DC.
- Oliver, M. J. and K. Hake (2013). Seed-based containment strategies. Plant Gene Containment. M. J. Oliver and Y. Li, John Wiley & Sons, Inc.: 113-124.
- Oliver, M. J., et al. (2004). "Seed-based strategies for transgene containment." Proceedings of the 8th International Symposium on the Biosafety of Genetically Modified Organisms: 154-161.
- Oliver, M. J., et al. (1998). Control of plant gene expression. US.
- Oliver, M. J., et al. (1999). Control of gene expression. 5,925,808. US.
- Oliver, M. J., et al. (1999). Control of gene expression. 5,977,441. US.
- Ortega-Molina, J. (2004). Results of the studies into the coexistence of genetically modified and conventional maize. COPA-COGECA colloquy on the co-existence and thresholds of adventitious presence on GMOs in conventional seeds.
- Pappenheimer, A. M., Jr (1977). "Diphtheria toxin." Annual Review of Biochemistry **46**(1): 69-94.
- Pilson, D. and H. R. Prendeville (2004). "Ecological effects of transgenic crops and the escape of transgenes into wild populations." Annual Review of Ecology, Evolution, and Systematics **35**: 149-174.
- Polóniová, Z., et al. (2015). "The pollen- and embryo- specific Arabidopsis DLL promoter bears good potential for application in marker-free Cre/loxP self-excision strategy." Plant Cell Reports **34**(3): 469-481.
- Reichman, J. R., et al. (2006). "Establishment of transgenic herbicide-resistant creeping bentgrass (*Agrostis stolonifera* L.) in nonagronomic habitats." Molecular Ecology **15**(13): 4243-4255.
- Rizwan, M., et al. (2019). "Gene flow from major genetically modified crops and strategies for containment and mitigation of transgene escape: A review." Applied Ecology and Environmental Research **17**(5): 11191-11208.
- Sandhu, S., et al. (2010). "Apomixis and ploidy barrier suppress pollen-mediated gene flow in field grown transgenic turf and forage grass (*Paspalum notatum* Flüggé)." Theoretical and Applied Genetics **121**(5): 919-929.

- Sang, Y., et al. (2013). "Gene use restriction technologies for transgenic plant bioconfinement." Plant Biotechnology Journal **11**(6): 649-658.
- Schafer, M. G., et al. (2011). "The establishment of genetically engineered canola populations in the US." PLoS One **6**(10): e25736.
- Senecoff, J. F., et al. (1985). "The FLP recombinase of the yeast 2-micron plasmid: Characterization of its recombination site." Proceedings of the National Academy of Sciences **82**(21): 7270-7274.
- Shao, M., et al. (2017). "Small serine recombination systems ParA-MRS and CinH-RS2 perform precise excision of plastid DNA." Plant Biotechnology Journal **15**(12): 1577-1589.
- Sobecky, P. A., et al. (1996). "Characterization of the stable maintenance properties of the par region of broad-host-range plasmid RK2." Journal of Bacteriology **178**(7): 2086-1093.
- Srivastava, V., et al. (2017). "Dual-targeting by CRISPR/Cas9 for precise excision of transgenes from rice genome." Plant Cell, Tissue and Organ Culture **129**(1): 153-160.
- Stewart Jr., C. N. (2021). Controlling transgene flow from engineered crops to unintended hosts by molecular approaches. Gene Flow: Monitoring, Modeling and Mitigation. W. Wei and C. N. Stewart, CAB International.
- Tan, J., et al. (2022). "An efficient marker gene excision strategy based on CRISPR/Cas9-mediated homology-directed repair in rice." International Journal of Molecular Sciences **23**(3): 1588.
- Thomson, J. G. and D. W. Ow (2006). "Site-specific recombination systems for the genetic manipulation of eukaryotic genomes." Genesis **44**(10): 465-476.
- Trumbo, J. L., et al. (2015). "Manipulating microRNAs for improved biomass and biofuels from plant feedstocks." Plant Biotechnology Journal **13**(3): 337-354.
- Twell, D. (1995). "Diphtheria toxin-mediated cell ablation in developing pollen: Vegetative cell ablation blocks generative cell migration." Protoplasma **187**(1): 144-154.
- Warwick, S. I., et al. (2009). "Gene flow, invasiveness, and ecological impact of genetically modified crops." Annals of the New York Academy of Sciences **1168**(1): 72-99.
- Wilson, G. G. (1988). "Type II restriction—Modification systems." Trends in Genetics **4**(11): 314-318.
- Yamaizumi, M., et al. (1978). "One molecule of diphtheria toxin fragment A introduced into a cell can kill the cell." Cell **15**(1): 245-250.

Yang, Y., et al. (2022). "Plant synthetic biology innovations for biofuels and bioproducts." Trends in Biotechnology **40**(12): 1454-1468.

Yau, Y. Y. and C. N. Stewart Jr. (2013). "Less is more: Strategies to remove marker genes from transgenic plants." BMC Biotechnology **13**(1): 1-23.

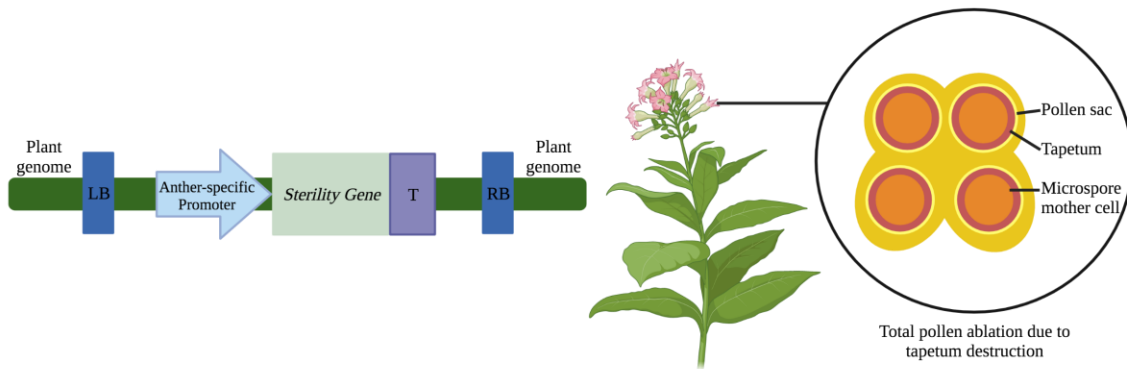
Ye, X., et al. (2022). "Cre-mediated autoexcision of selectable marker genes in soybean, cotton, canola and maize transgenic plants." Plant Cell Reports **42**: 45-55.

Zapiola, M. L., et al. (2007). "Escape and establishment of transgenic glyphosate-resistant creeping bentgrass *Agrostis stolonifera* in Oregon, USA: A 4-year study." Journal of Applied Ecology **45**(2): 486-494.

Zhang, C., et al. (2012). "Control of pollen-mediated gene flow in transgenic trees." Plant Physiology **159**(4): 1319-1334.

Appendix

(a)



(b)

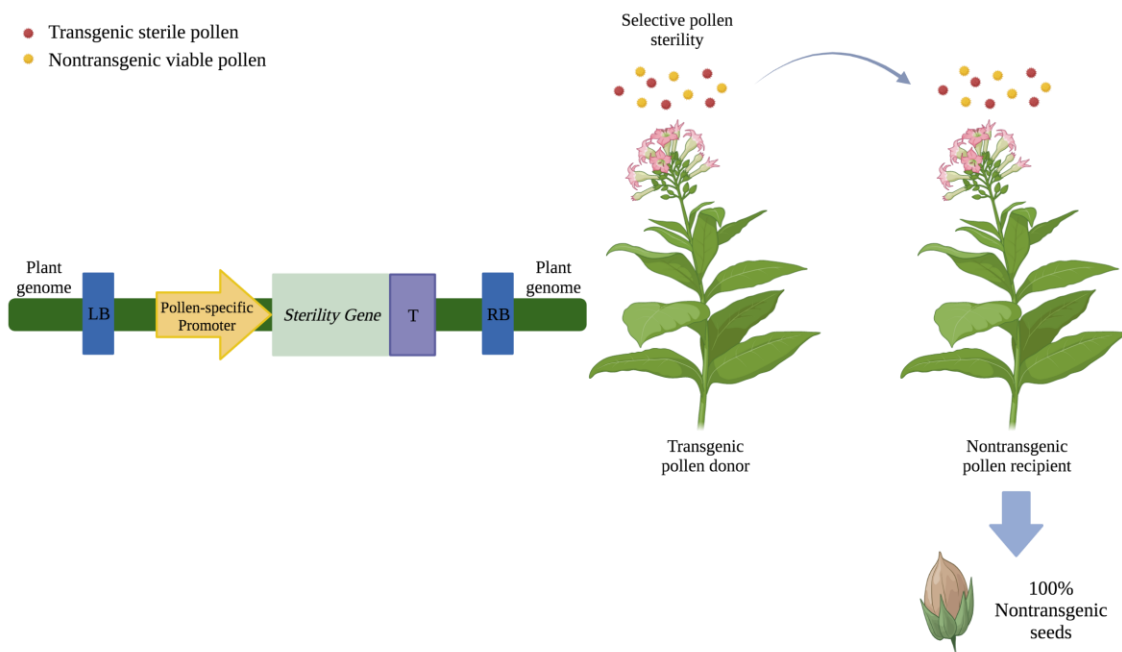


Figure 2.1 Engineered male-sterility approaches. **(a)** A sterility gene, such as *Barnase*, is located between an anther-specific promoter and a terminator (T) within the left (LB) and right borders (RB) on T-DNA inserted into the plant genome. *Barnase* expression prevents pollen formation by degrading RNA in the tapetal cells, the nutritive cells surrounding microspore mother cells. **(b)** A sterility gene, such as *EcoRI* endonuclease, is located between a pollen-specific promoter and a terminator within the left and right borders on the T-DNA, and in this example is inserted once in the plant genome. In this theoretical scenario, *EcoRI* expression leads to the death of transgenic pollen grains and nontransgenic pollen remains viable. Pollen movement from transgenic to nontransgenic plants yields 100% nontransgenic seeds. Figure created with Biorender.com, accessed on 22 February 2023.

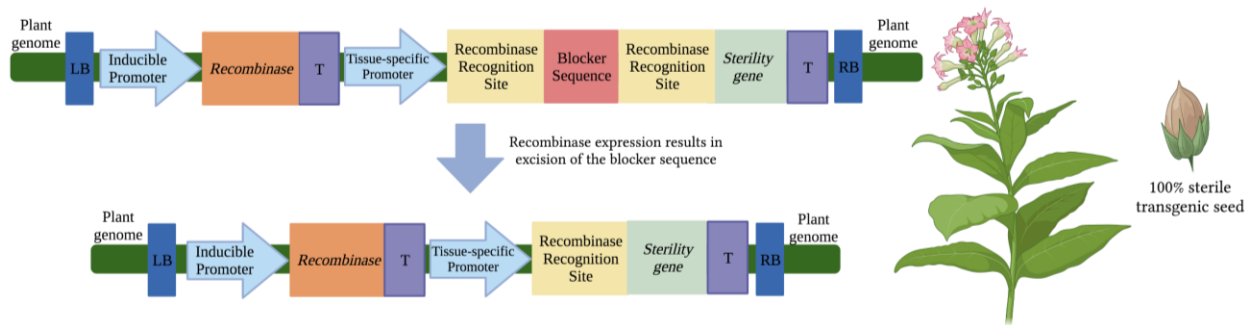


Figure 2.2 Conditional sterility via inducible system activation of sterility genes. Within the left (LB) and right borders (RB) of the T-DNA located in the plant genome, an inducible promoter is located upstream of a recombينase gene and terminator (T). An external stimulus is applied to induce recombينase gene expression. The recombينase enzyme subsequently excises the blocker sequence by cutting DNA segments at flanking recognition sites. The tissue-specific promoter (ex. seed-specific) is moved into close proximity of the sterility gene and transcription can occur. The sterility gene product leads to cell death. In this theoretical example, a homozygous tobacco line with a seed-specific promoter driving sterility gene expression leads to 100% sterile transgenic seeds. Figure created with Biorender.com, accessed on 18 January 2023.

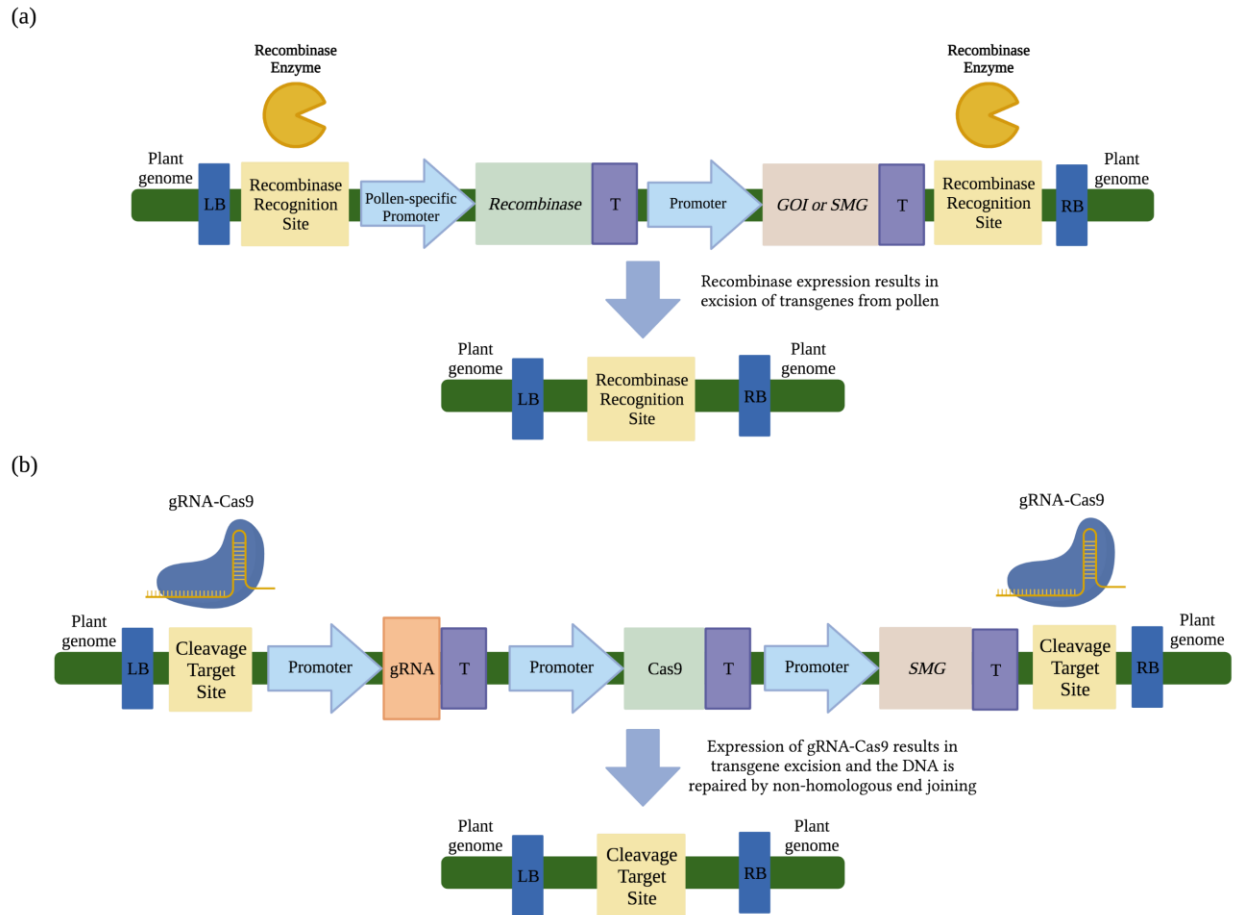


Figure 2.3 Transgene excision by site-specific recombinase and CRISPR/Cas9 gene editing. **(a)** Recombinase expression is controlled by a pollen-specific promoter. In pollen, the recombinase recognizes sites flanking the gene of interest (GOI) or selectable marker gene (SMG) and the recombinase expression cassettes and removes the DNA between the recognition sites. One recognition site and left (LB) and right (RB) borders remains while the DNA removed is degraded within the plant cell. **(b)** CRISPR/Cas9 excision utilize cleavage target sites (CTSs) that flank the guideRNA (gRNA), Cas9 enzyme, and SMG expression cassettes. The CTSs are comprised of identical DNA sequences. Expression of a single gRNA and Cas9 enzyme excises the T-DNA between the CTSs. In this example, cleaved DNA ends are repaired by non-homologous end joining, leaving behind one CTS and left and right borders and the DNA removed is degraded within the plant cell. Figure created with Biorender.com, accessed on 22 February 2023.

Chapter 3: Pollen-targeted expression of *EcoRI* endonuclease in *Panicum virgatum* L.

Abstract

Pollen-targeted expression of the *EcoRI* restriction endonuclease in switchgrass (*Panicum virgatum* L.) holds promise as a valuable tool for preventing transgene flow to unintended hosts. This study evaluates the potential of *EcoRI* endonuclease, with the inclusion of a catalase intron, to induce sterility of transgenic switchgrass pollen when driven by monocot-derived, putative pollen-active promoters. Of the TaPSG719, PvPS1, Osg6B, OsRTS, and Zm13 promoters tested, the Zm13 and the PvPS1 promoters exhibited the highest expression levels in transgenic anther and pollen tissues, averaging 10% and 6% transcript abundance, respectively, compared to the switchgrass ubiquitin 1 gene used as a reference. Despite relatively high transcript expression, pollen germination (%) of the highest expressing transgenic lines of the Zm13:*EcoRI* and PvPS1:*EcoRI* constructs were not significantly different from the pollen germination (%) of nontransgenic switchgrass. The pollen tested was, presumably, a mixture of transgenic and nontransgenic pollen from the hemizygous transgenic T₀ clones. Further investigation revealed that cryptic splicing occurred as an unintended result of the inclusion of the catalase intron in the *EcoRI* coding sequence. The cryptic splicing event resulted in the generation of a truncated mRNA product that apparently was translated to a nonfunctional protein. Thus, the constructs designed for bioconfinement were not fit for purpose.

Introduction

Over 90% of crops cultivated in the US are genetically engineered (GE) and primarily confer few well-characterized traits, such as insect and herbicide resistance, that are deployed in a limited number of crop species (ISAAA 2019). These species typically have little potential for inadvertent gene flow, often a result of their reproductive biology, and no wild relatives native to the region of cultivation (Stewart Jr. 2021). Species with an increased potential of gene flow raise concerns of biosafety and ecological consequences, thereby hindering the commercial adoption of innovative GE crops. Pollen-mediated gene flow from transgenic crops to wild relatives is more likely to occur in species that are wind-pollinated and obligate-outcrossers, such as switchgrass (*Panicum virgatum* L.) (Rognli, Nilsson et al. 2000, Haygood, Ives et al. 2003, Stewart Jr., Halfhill et al. 2003). Implementation of a biological confinement (bioconfinement) method could confine transgenes to the GE crop and resolve biosafety concerns.

A proposed bioconfinement method that is applicable across a wide range of crop species is male-sterility via the overexpression of restriction endonucleases in targeted pollen cells. There are three types (I, II, and III) of restriction endonucleases, type II being the most studied and understood. *EcoRI* is a well-known type II restriction endonuclease that creates a double-stranded break with “sticky ends” at the nucleotide sequence 5’-GAATTC-3’ (Wilson 1988). Targeted overexpression of *EcoRI* endonuclease in transgenic pollen tissue is predicted to prevent transgene flow from GE crops by inducing extensive DNA damage that renders transgenic pollen cells sterile. *EcoRI* endonuclease has previously been overexpressed in pollen of the dicot crop tobacco (*Nicotiana tabacum* L.) and displayed up to 100% bioconfinement with minimal to no phenotypic disadvantage (Millwood et al., 2016). Translation of the *EcoRI* endonuclease bioconfinement system into monocot crops including switchgrass will require appropriate monocot-promoter selection for *EcoRI* endonuclease expression. This will be critical to ensure the bioconfinement system is tightly controlled to limit off-target effects that may negatively impact plant productivity, e.g. vegetative biomass production or floral development.

The Osg6B, OsRTS, PvPS1, TaPSG719, and Zm13 promoters have been identified as anther- or pollen- specific in other monocot species and have been selected as candidates to drive transgene expression in switchgrass pollen. Tapetum-targeted expression of *EcoRI* endonuclease may disrupt pollen development and prevent transgene flow from GE crops since fit tapetal cells are necessary for supporting the development of mature, viable pollen grains (Echlin 1971,

Goldberg, Beals et al. 1993, McCormick 1993). The Osg6B promoter from rice (*Oryza sativa* L.) has been observed to drive gene expression in tapetal cells of both monocot and dicot species and its activity is no longer detectable by the early binucleate pollen stage (Tsuchiya, Toriyama et al. 1994, Yokoi, Tsuchiya et al. 1997, Kumar, Kaur et al. 2017). The rice-derived OsRTS promoter has also been previously identified as tapetum-specific in transgenic rice, *Arabidopsis* (*Arabidopsis thaliana*), and creeping bentgrass (*Agrostis stolonifera*), and gene expression under control of the OsRTS promoter is no longer detected at anthesis (Luo, Lee et al. 2006). Pollen-targeted *EcoRI* endonuclease expression may prevent gene-flow by producing non-viable, nongerminable pollen. To my knowledge, the switchgrass-derived PvPS1 promoter has not been previously characterized. However, it is homologous to the rice-derived OsPS1 promoter which has been observed to drive gene expression in mature pollen of transgenic rice, tobacco, and switchgrass (Zou, Zhan et al. 1994, Zhan, Wu et al. 1996, Somleva, Xu et al. 2014). The pollen-specific TaPSG719 promoter is from wheat (*Triticum aestivum* L.) and, in transgenic tobacco, was documented to drive gene expression primarily during middle to late pollen development and decreases during pollen maturation (Chen, Tu et al. 2010). The final candidate is the Zm13 promoter from maize (*Zea mays* L.), which has been identified as pollen-specific, with transcript levels greatest in mature pollen grains just prior anthesis (Hamilton, Bashe et al. 1989, Hanson, Hamilton et al. 1989). Additionally, the activity of the Zm13 promoter has been characterized in engineered plants from various monocot and dicot species (Hamilton, Schwarz et al. 1998, Hamilton, Schwarz et al. 2000, Hague, Dellaporta et al. 2012). Although most of these promoter candidates have been characterized in other species, spatial expression is undetermined in switchgrass.

My objective of the thesis research was to determine whether these pollen/anther-active promoters driving the expression of an *EcoRI* endonuclease gene conferred selective male sterility in engineered switchgrass.

2. Materials and Methods

2.1 Generation of stably transgenic switchgrass

Plasmid constructs with putative pollen- and anther-specific promoters each driving the *EcoRI* coding sequence were unable to be transformed into *Agrobacterium tumefaciens* due to leaky expression, therefore the gene included a catalase intron (derived from *Ricinus communis*) (Figure 1) that was constructed by a previous lab member, and I characterized the plasmids by DNA sequencing. The constructs were transformed into *A. tumefaciens* strain EHA105 using a freeze-thaw method (Weigel and Glazebrook 2006). *Agrobacterium*-mediated transformation of seed-derived switchgrass embryogenic callus was performed using a protocol modified from (Li and Qu 2011).

Embryogenic callus was derived from switchgrass (*P. virgatum* ‘Performer’) seeds obtained from Ernst Conservation Seeds, Inc. Seed coats were removed using 65% sulfuric acid for 30 minutes, then sterilized using 50% Clorox bleach and 1% Tween solution for 30 min. In a sterile laminar flow hood, the seeds were washed 3 to 4 times with sterile de-ionized (DI) water for approximately 2 min per wash. Seeds were dried on a sterile paper towel for 30 min, then plated on callus induction media and placed in the dark at 20°C for six weeks. The callus induction media was prepared by dissolving 4.43 g of Murashige & Skoog basal (MS) medium with vitamins (PhytoTech Labs), 1 mg/L 6-benzaminopurine (BAP), 5 mg/L 2,4-dichlorophenoxyacetic acid (2,4-D), 2 grams of L-proline, 3% D-(+)-maltose monohydrate, pH=5.8, 3 g of phytagel up to 1 L in DI water. After this period, callus was sub-cultured every two weeks until yellow, friable, type II callus was present. Once type II callus was identified, only this type of callus was further sub-cultured since it is essential to receive fast growing, transgenic callus with high regeneration efficiency (Burris, Mann et al. 2009).

A. tumefaciens strain EHA105 containing the vector of interest was grown from a single colony in 50 mL of yeast extract peptone (YEP) with 50 mg/L rifampicin and 50 mg/L kanamycin in a shaker overnight at 28°C with shaking at 225 rpm. Once the culture reached an $OD_{600}=1.0$, the *Agrobacterium* culture was centrifuged at $4,000 \times G$ for 10 min and resuspended in 25 mL of liquid MP at an adjusted $OD_{600}=0.8$ and supplemented with 100 μM acetosyringone (AS) to induce *vir* genes. The liquid MP medium was prepared by dissolving 4.43 g of MS medium with vitamins, 1 mg/L BAP, 5 mg/L 2,4-D, 2 gr of L-proline, 3% D-(+)-maltose monohydrate, pH=5.8 up to 1 L in DI water. Thirty pieces of callus were added to a sterile 50

mL falcon tube and immersed in 25 mL of the *Agrobacterium* suspension. A vacuum was applied at -80 kPa for 10 min, followed by 20 min of moderate agitation on a vortex shaker. Callus pieces were placed on sterile paper towels to dry for 30 min, then transferred to petri dishes with sterile Whatman No. 1 filter paper and sealed with micropore tape. The petri dishes containing the callus pieces were placed in the dark for a two-day desiccation period. Following desiccation, the callus pieces were transferred to S1 selection media and placed in the dark for 2 weeks. S1 media was prepared by dissolving 4.43 g of MS medium with vitamins, 1 mg/L BAP, 5 mg/L 2,4-D, 2 g of L-proline, 3% D-(+)-maltose monohydrate, 500 mg/L timentin, 100 mg/L hygromycin B, pH=5.8, 3 g of phytigel up to 1 L in DI water. After 2 weeks, the callus pieces were moved to S2 selection media for 4 weeks in the dark, transferred every 2 weeks. S2 media was prepared by dissolving S1 media with an increased concentration of 200 mg/L hygromycin B. Putative transgenic callus were identified by the mOrangeER fluorescent reporter gene, constitutively expressed by the cauliflower mosaic virus (CaMV) 35S promoter, that was included in the T-DNA. Callus that appeared to fluoresce were transferred to regeneration selection media for 4 weeks, transferred every 2 weeks. Regeneration selection media was prepared by dissolving 4.43 g of MS medium with vitamins, 1 mg/L BAP, 0.20 mg/L of 1-naphthaleneacetic acid (NAA), 0.50 mg/L gibberellic acid (GA3), 250 mg/L timentin, 100 mg/L hygromycin B, 3% D-(+)-maltose monohydrate, pH=5.8, 3 g of phytigel up to 1 L in DI water. Regeneration occurred in an environmental-controlled growth room at 16/8 hour day/night. Shoots were transferred to a rooting selection media for 4-6 weeks, transferred every two weeks. Rooting selection media was prepared by dissolving 2.215 g of MS medium with vitamins, 150 mg/L timentin, 50 mg/L hygromycin B, 3% D-(+)-maltose monohydrate, pH=5.8, 3 g of phytigel up to 1 L in DI water. After root formation, plantlets were transferred to potting mix and grown in growth chambers at 16/8 hour day/night.

2.2 Assessment of transgenicity: PCR analysis

Genomic DNA was extracted from leaf tissue samples of switchgrass plants using the CTAB extraction method (Stewart and Via 1993). PCR was performed in a 25 µL reaction using DreamTaq Green PCR Master Mix (2X) (Thermo Fisher Scientific) and carried out in accordance with the manufacturer's protocol. Primer sets that spanned a fragment of the relevant promoter into the *EcoRI* gene were used to confirm presence of the expected promoter and gene

of interest. PCR conditions for the promoters into the *EcoRI* gene were as follows: 95°C for 2 min followed by 35 cycles of 95°C for 30 sec, 56°C for 30 sec, 72°C for 1 min, and a final extension at 72°C for 5 min. PCR was also performed to amplify the length of the *EcoRI* gene to confirm its full insertion. PCR conditions for the *EcoRI* gene were as follows: 95°C for 2 min followed by 35 cycles of 95°C for 30 sec, 60°C for 30 sec, 72°C for 1 min, and a final extension at 72°C for 5 min. Amplification of the *Agrobacterium* gene *chvA* (chromosomal virulence gene A) was used as a control for *Agrobacterium* contamination and the PCR conditions were as follows: 95°C for 2 min followed by 35 cycles of 95°C for 30 sec, 63°C for 30 sec, 72°C for 45 sec, and a final extension at 72°C for 5 min. PCR products were visualized on a 1.0% agarose gel with ethidium bromide. All primer sets are listed in Supplementary Table S3.1.

2.3 Assessment of transgenic: Southern blot analysis

Southern blot analysis was performed to quantify copy number of transgene integration of 11 pTaPSG719:*EcoRI*, 15 pPvPS1:*EcoRI*, 12 pOsg6B:*EcoRI*, 16 pOsRTS:*EcoRI*, and 5 pZm13:*EcoRI* transgenic switchgrass lines. Genomic DNA extraction from leaf tissue was carried out using the CTAB method previously described, cleaned with a Genomic DNA Clean & Concentrator™ kit (Zymo Research), and 10 µg of genomic DNA was used in a 30 µL restriction digestion reaction using enzymes *EcoRI*-HF® and *SspI*-HF® (New England BioLabs). Gel electrophoresis and blotting were carried out through an established protocol (Mellars and Gomez 2011). The probe targeted the *EcoRI* gene and was generated using primers forward, 5'- CGTAGCAGCACAAAAGAGATGCG and reverse, 5'- CCATCTCGGGGAAAGTTCACG. Each probe was DIG labeled (Roche Applied Science) and hybridizations were conducted at 42°C. Blots were visualized using ChemiDoc Imaging System (Bio-Rad).

2.4 RNA extraction and expression analysis by qRT-PCR

RNA extraction was performed using TRI reagent (Zymo Research) from leaf tissue, flowers (collected prior anthesis), and combined anthers/pollen (collected immediately after anthesis) of 9 pTaPSG719:*EcoRI*, 7 pPvPS1:*EcoRI*, 7 pOsg6B:*EcoRI*, 12 pOsRTS:*EcoRI*, and 4 pZm13:*EcoRI* transgenic lines and a nontransgenic control. Extractions were performed according to the manufacturer's protocol, treated with Dnase I, and cleaned with an RNA Clean

& Concentrator™ kit (Zymo Research). RNA quantification was determined using a nanodrop spectrophotometer ND-1000 (Thermo Fisher Scientific). cDNA of leaf and flower tissue was synthesized using 2 µg of RNA per sample, and cDNA of anther/pollen tissue was synthesized using 500 ng of RNA per sample with Applied Biosystems™ High Capacity cDNA Reverse Transcription Kit (Thermo Fisher Scientific) following the manufacturer's protocol. qPCR was performed using PowerUp™ SYBR™ Green 2X Master Mix (Thermo Fisher Scientific) in a 5 µL reaction, using 5 ng of cDNA and 1 µL (0.5 µM) of each forward and reverse primers. The qPCR was performed using QuantStudio™ 6 Flex qPCR System (Applied Biosystems), and the settings for all primer sets were as follows: 50°C for 2 min, 95°C for 10 min, then 40 cycles of 95°C for 15 sec, descends 1.6°C per second until 60°C, 60°C for 1 min, then ascends to 95°C at 1.6°C per second. The qRT-PCR results were analyzed by the $2^{-\Delta C_T}$ method, with *EcoRI* and *PS1* C_T values set relative to *Ubq1* C_T values, and further analyzed by a one-way ANOVA, followed by a Tukey's HSD, $\alpha=0.05$.

Primer sets used for *EcoRI*, *PS1*, and the reference gene switchgrass ubiquitin 1 (*Ubq1*) (GeneBank accession number: [FL899020](#)) can be found in Supplementary Table S3.2, and gel confirmation of the *EcoRI* and *Ubq1* primer sets can be found Supplementary Figure S3.1.

2.5 Pollen germination analysis

Pollen was collected from recently dehiscent anthers between hours 1500 and 1530, and germinated on media containing 0.5 M sucrose, 1 mM boric acid, 2 mM calcium chloride, and 0.3% (w/v) gelzan (Fei and Nelson 2003). Pollen was collected from transgenic switchgrass lines pPvPS1:*EcoRI* line 6 and pZm13:*EcoRI* line 4), which were determined to have the greatest *EcoRI* transcript abundance in pollen/ anther tissue. Transgenic plants were presumed to produce a mixture of transgenic and nontransgenic pollen based on a T-DNA insertion. Three pollen samples were collected per clonal propagate, three clones per transgenic line. Nontransgenic pollen was also collected as a control. Germination was defined by the observation of pollen tubes that protruded at least the diameter of the pollen grain, and germination percentage was determined by counting the number of pollen grains germinated per at least 100 grains observed (Tuinstra and Wedel 2000). Pollen grains that appeared to have burst were not counted. Pollen

was observed and imaged using EVOS™ M7000 compound microscope (Thermo Fisher). These results were analyzed by a one-way ANOVA, followed by a Tukey's HSD, $\alpha = 0.05$.

2.6 RT-PCR analysis

RT-PCR was conducted to determine whether the catalase intron was properly spliced from the transcript. RT-PCR was performed in a 20 μ L reaction using DreamTaq Green PCR Master Mix (2X) (Thermo Fisher Scientific), 20 ng of cDNA template (generated from the anther/ pollen RNA samples as described in section 2.5), and 1 μ L (0.5 μ M) of each forward 5'-GCGATCGATATCATGGTAGATCTGAGG and reverse 5'-TTACTTCTTAGATGTAAGCTGTTCAAACAAGTCAC primers that amplified the full length of the *EcoRI* coding sequence. PCR conditions were as follows: 95°C for 2 min followed by 35 cycles of 95°C for 30 sec, 60°C for 30 sec, 72°C for 30 sec, and a final extension at 72°C for 5 min. RT-PCR products were visualized on a 1.2% agarose gel and product sizes were analyzed using the TrackIt™ 100 bp DNA Ladder (Thermo Fischer Scientific).

2.7 Phenotypic analysis

To assess off-target effects of *EcoRI* expression that may affect plant productivity and phenotype, all single- and double-copy T_0 transgenic lines and one nontransgenic control line that was generated from seed were grown in the greenhouse and assessed after eight weeks. Three clonal propagates per line were each potted in a 2.37 L pot, totaling 120 potted plants. Plants were trimmed to a height of 0.50 meters above the base of the plant and randomly organized (Excel Ablebits Tools). All plants were grown simultaneously between July and August under long day photoperiods and watered 3 times per week. Every other week one water application was supplemented with Peter's Professional 20-20-20 General Purpose Fertilizer at a concentration of 25 ppm N. After eight weeks of growth, measurements of plant height, tiller count, leaf blade width, and tiller diameter were recorded. Plant height (cm) was measured from base of the plant to the apex of the tallest extended leaf blade. Tiller production was assessed by tallying total tillers per plant. The leaf blade width (mm) of three of the most recently expanded leaves per clonal propagate was measured using digital vernier calipers (Maxwell). Digital vernier calipers were also used to measure the diameter (mm) of three tillers approximately 10 cm above the plant's base. Flowers were visually assessed for phenotypic differences. To

determine if these measured traits of the transgenic lines differed from the nontransgenic control, results were analyzed by a one-way ANOVA, followed by a Dunnett's test, $\alpha=0.05$. A correlation analysis was also performed using Spearman's correlation to determine the relationship between *EcoRI* expression in leaf tissue and plant height and leaf blade width.

2.8 Statistical analysis

A total of 9 pTaPSG719:*EcoRI*, 7 pPvPS1:*EcoRI*, 7 pOsg6B:*EcoRI*, 12 pOsRTS:*EcoRI*, and 4 pZm13:*EcoRI* transgenic switchgrass lines and 1 nontransgenic control line, with three biological replicates per line, were organized in a completely randomized design.. Statistical analysis was performed using the JMP pro 17 software (SAS, Cary, NC).

3. Results

3.1 Molecular characterization of transgenic switchgrass

A total of 10 pTaPSG719:*EcoRI*, 15 pPvPS1:*EcoRI*, 12 pOsg6B:*EcoRI*, 16 pOsRTS:*EcoRI*, and 5 pZm13:*EcoRI* independent stable T₀ transgenic switchgrass lines were PCR confirmed for their respective promoters and the *EcoRI* endonuclease transgene. Southern hybridization of these lines identified copy number of *EcoRI* endonuclease transgene integration (Supplementary Figure S3.2), and all single- and double- copy lines were selected for further analysis (except for few lines described below). This included 9 pTaPSG719:*EcoRI*, 7 pPvPS1:*EcoRI*, 7 pOsg6B:*EcoRI*, 12 pOsRTS:*EcoRI*, and 4 pZm13:*EcoRI* transgenic lines, summarized in Supplementary Table S3.3. Transgenic lines in Southern blot lanes 4 and 9 of pTaPSG719:*EcoRI* appeared to be clones, thus line 9 was excluded from further analysis. Additionally, pOsRTS:*EcoRI* line 10 and pPvPS1:*EcoRI* line 1 showed early signs of severely stunted growth and were also excluded from further analysis. For simplicity, transgenic lines further characterized have been reassigned line names described in Supplementary Table S3.4.

3.2 Transcript abundance analysis in transgenic switchgrass

Quantitative reverse transcription-polymerase chain reaction (qRT-PCR) was used to determine transcript levels of the *EcoRI* endonuclease coding sequence in various tissues. In pollen/anther tissues (collected immediately following anthesis), the *EcoRI* endonuclease gene transcript abundance was highest when driven by the Zm13 and PvPS1 promoters with a mean of 10.48%

and 5.52%, respectively, relative to switchgrass ubiquitin 1 (*PvUbq1*) (Figure 3.2). Of the transgenic lines with these promoters, greater *EcoRI* expression levels were commonly observed in flower (collected prior anthesis) and leaf tissue compared to the pollen and anther tissue (Figure 3.3A,B), which may be indicative the promoter is not tightly controlled.. In the pPvPS1:*EcoRI* transgenic lines, qRT-PCR was also used to compare expression patterns of the *EcoRI* transgene to the endogenous *PS1* gene (both under control of the PvPS1 promoter). In pollen/ anther tissue, I found the endogenous PvPS1 gene was expressed from 20-fold to 162-fold *PvUbq1* with little expression in leaf tissue, some expression in flowers (Figure 3.3C).

3.3 Pollen germination analysis

To determine if pollen-targeted *EcoRI* endonuclease expression produced a male-sterile phenotype, pollen germination experiments were performed in the greenhouse immediately after anthesis between hours 1500 and 1530. Pollen germination of the greatest pollen/anther-expressing transgenic lines, pPvPS1:*EcoRI* line 6 and pZm13:*EcoRI* line 4 (determined by qRT-PCR) and a nontransgenic control were analyzed. Pollen samples of pPvPS1:*EcoRI* line 6 contained a single copy insertion of the T-DNA and were predicted to have a 1:1 ratio of transgenic to nontransgenic pollen. Therefore, I expected to observe a pollen germination rate approximately half that of nontransgenic pollen. In the pZm13:*EcoRI* line 4, two copies of the T-DNA were integrated into the genome, and I predicted to have either a 1:1 ratio of transgenic to nontransgenic pollen or all transgenic pollen, depending the chromosome locations of the insertions. Thus, I expected the pollen germination rate to be at least half that of the nontransgenic pollen, and potentially observe no pollen germination. From these experiments, I determined the mean pollen germination of pPvPS1:*EcoRI* line 6, pZm13:*EcoRI* line 4, and the nontransgenic control was 49.0%, 47.6%, and 47.9%, respectively (Figure 3.4). There are no significant differences between the transgenic lines and the nontransgenic control determined by a one-way ANOVA, followed by a Tukey's HSD, $\alpha = 0.05$ ($F(2, 24) = 0.18$, $p = 0.84$).

3.4 Analysis of intron removal through RT-PCR

RT-PCR of cDNA from pollen/anther tissue revealed amplification of two bands in many transgenic lines (Figure 3.5), however were predominantly observed in pPvPS1:*EcoRI* lines 1-6 and pZm13:*EcoRI* line 4. The larger band was approximately 1,066 bp, the expected size if the

intron was present in the mature mRNA. The smaller band was expected to be ~860 bp, but was observed to be 590 bp. Sanger sequencing (Eurofins Genomics) of the 590 bp band revealed cryptic splicing at an AG nucleotide sequence of Serine87, and the larger band included the intron as predicted. The final mRNA products were the truncated *EcoRI* endonuclease gene and the *EcoRI* endonuclease gene with the interrupting intron that both produce repeated, premature stop codons (Supplementary Figure S3.3).

3.5 Plant phenotype analysis

Many lines were indistinguishable from nontransgenic switchgrass, however numerous transgenic plants displayed a significantly different phenotype from nontransgenic switchgrass (Supplementary Figure S3.4) in each category measured (Figure 3.6A-D) (one-way ANOVA, followed by Dunnett's test, $\alpha = 0.05$). Reduced plant height was the most commonly observed phenotypic difference, where the majority of transgenic lines were significantly shorter than the nontransgenic control. Decreased tiller diameter was also observed in several transgenic lines. No transgenic lines displayed a significant decrease in tiller count, however few lines exhibited an increase in tiller count. Additionally, several lines displayed a significant reduction of leaf blade width. Correlation analysis was performed using Spearman's correlation to determine the relationship between *EcoRI* expression in leaf tissue and plant height or leaf blade width (Figure 3.6E,F). Results indicated a weak negative correlation that was statistically significant between *EcoRI* expression in leaf tissue and leaf blade width ($r = -0.23$, $p < 0.05$), indicating some association between increased *EcoRI* expression in leaf tissue and decreased leaf blade width. A weak negative, and not statistically significant, correlation was found between *EcoRI* expression in leaf tissue and plant height ($r = -0.08$, $p = 0.4134$). Another intriguing observation pertained to two of the pOsRTS:*EcoRI* lines, which displayed a delayed flowering phenotype (Supplementary Figure S3.5). These lines had flowers that remained thin and underdeveloped into reproductive stage 3 for an extended period of time (approximately two months). This was followed by maturation of floral tissues, which produced a full flower with a normal appearance, and anthesis was first observed in the second week of September, months after the nontransgenic controls began to flower.

4. Discussion

Implementation of a bioconfinement system in genetically engineered crops, especially those with increased risk of gene flow, may be imperative to avert both foreseen and unforeseen consequences. Although numerous approaches have been researched (Stewart Jr. 2021), they often come with associated drawbacks, and few have managed to achieve close to 100% transgene confinement. A promising solution lies in the adoption of a selective pollen-sterility approach, which could have broad utility across various species and few drawbacks, making it a valuable bioconfinement system. Prior research has achieved complete transgene confinement in the dicot model species tobacco, through pollen-targeted expression of *EcoRI* endonuclease (Millwood et al. 2016). However, translation of this system in monocot species remains unresearched. Given that many monocot crops including switchgrass are outcrossing species and have high gene flow frequencies, integrating a similar bioconfinement system in monocots would be advantageous.

Plasmid constructs were developed by a previous lab member using five putative pollen- or anther- specific promoters from four monocot species, each driving expression of the *EcoRI* endonuclease coding sequence interrupted by a modified intron (mCBC) to prevent leaky expression in *Agrobacterium* (Ohta, Mita et al. 1990). These plasmids were employed to generate transgenic switchgrass. Downstream analysis of the transgenic lines revealed various challenges related to plasmid design, including the use of the catalase intron.

Quantitative reverse transcription-polymerase chain reaction (qRT-PCR) identified the Zm13 and PvPS1 promoters as the primary candidates to drive transgene expression in anther and pollen tissues. Interestingly, *EcoRI* endonuclease expression was also detected in leaf and flower tissues, often surpassing levels observed in pollen and anther tissues. This finding raises several concerns, as these promoters were initially anticipated to predominantly drive gene expression in mature pollen grains since the Zm13 promoter has been characterized to express in mature pollen of multiple species (Hamilton, Bashe et al. 1989, Hanson, Hamilton et al. 1989, Hamilton, Schwarz et al. 1998, Hamilton, Schwarz et al. 2000, Hague, Dellaporta et al. 2012) and we could expect the PvPS1 promoter to have similar spatiotemporal expression as the rice-derived OsPS1 promoter, which has been previously characterized to express in mature pollen of several species including switchgrass (Zou, Zhan et al. 1994, Zhan, Wu et al. 1996, Somleva, Xu et al. 2014). Several factors may contribute to this unexpected expression pattern. The genomic

location of transgene integration, along with close proximity to a constitutive 35S promoter on the T-DNA, may account for the elevated expression in leaves. Additionally, it is possible that, in switchgrass pollen, the Zm13 promoter expresses at a different developmental stage than predicted and observed in other species, therefore resulting in heightened expression in flowers, which encompass developing anther and pollen tissues.

Furthermore, a substantial disparity in expression levels was observed between the endogenous *PS1* gene and the *EcoRI* transgene, both driven by the switchgrass *PS1* promoter in the pPv*PS1*:*EcoRI* transgenic lines. Specifically, the endogenous *PS1* gene exhibited expression levels in pollen and anther tissues that ranged from 20-fold to 162-fold that of the reference gene. This variation may be attributed to the rapid dehiscence of the anthers and lifetime of the pollen at time of collection, with studies stating the switchgrass pollen half-life ranges from as little as 3.7 to 14.1 minutes under greenhouse conditions (Ge, Fu et al. 2011). In contrast, the highest expression of the *EcoRI* gene driven by the Pv*PS1* promoter was merely 20% of the reference gene. Although this may be considered relatively high, it is a notably lower level than that of the native *PS1* gene driven by the Pv*PS1* promoter in the same transgenic line. This disparity between *PS1* and *EcoRI* expression may be a result of using the Pv*PS1* promoter in a location of the plant's genome other than its native position, or possibly the design of the *EcoRI* plasmid. Situated between the promoter and *EcoRI* transgene in each of the five *EcoRI* plasmids is a 49 bp “cloning scar” (Supplementary Figure S3.6), a result of the traditional cloning methodology used in this case, and a common issue that has been observed to significantly diminish transcript expression when as little as four bp scars are present (Pfotenhauer, Occhialini et al. 2022). Nonetheless, the observed *EcoRI* transcript levels may have been suitable to induce a male-sterile phenotype if *EcoRI* was functional and efficient, but further work must be conducted to conclude these low *EcoRI* transcript levels are sufficient.

I also conducted pollen germination experiments using the top-expressing line of the leading constructs, pZm13:*EcoRI* and pPv*PS1*:*EcoRI*, but did not observe differences in pollen germination when compared to the nontransgenic control. This prompted me to perform RT-PCR to confirm the splicing of the catalase intron from the *EcoRI* endonuclease mRNA. This analysis revealed a significant shortcoming in the project- the intron splicing was inefficient and inadequate. The RT-PCR produced two bands: one of the expected size if the intron had not been spliced from the RNA, and another smaller than the predicted size if the intron had undergone

splicing. Sequencing of both bands confirmed the intron was not efficiently spliced from all mRNA. When it was spliced, it occurred at a cryptic splice site 257 bp into the *EcoRI* coding sequence. In both instances, these intron-related issues resulted in repeated premature stop codons, which were not conducive to achieving the intended phenotype. Other groups have also reported the mCBC catalase intron produced varying results of inadequate splicing (Tanaka, Mita et al. 1990, Ma, Lin et al. 2011) and its use should be avoided when possible.

Prior to determining the *EcoRI* endonuclease protein was likely nonfunctional, I began an eight-week growth study to assess the effects of off-target *EcoRI* endonuclease expression. During this period, I observed a wide range of phenotypic variation in the transgenic lines compared to the nontransgenic control. These variations included reduced plant height in the majority of lines, with few showing increased tiller count, decreased tiller diameter, and decreased leaf blade width. Additionally, two lines carrying the pOsRTS:*EcoRI* construct flowered in early September, months after the other transgenic lines and nontransgenic control flowered. However, I hesitate to attribute these phenotypic variations to *EcoRI* endonuclease expression, which did not appear correlative with plant height or leaf blade width ($r = -0.08$, $p = 0.4134$ and $r = -0.23$, $p = 0.0133$, respectively). Instead, I postulate other variables may be at play, such as the potential effects of random site-integration of the T-DNA, an inherent outcome of basic *Agrobacterium*-mediated transformation, that may disrupt native genes. Additionally, switchgrass's substantial genetic diversity may contribute to phenotypic variation observed in the switchgrass lines.

5. Conclusions

In this study, I engineered transgenic switchgrass lines carrying plasmid constructs with the *EcoRI* endonuclease coding sequence, with the inclusion of a catalase intron, under control of the putative pollen- and anther- specific TapSG719, PvPS1, Osg6B, OsRTS, and Zm13 promoters. Our findings report the Zm13 and the PvPS1 promoters produced the greatest transcript expression in anther/pollen tissue with a mean expression level of 10% and 6%, respectively, relative to the switchgrass ubiquitin 1 gene. These promoters produced similar expression levels, and lead to significantly greater expression levels than the other promoters tested and the control group. Interestingly, I found a notably significant difference in expression between the endogenous *PS1* gene and the *EcoRI* transgene, both being driven by the PvPS1 promoter. I

attribute this disparity to potential issues in plasmid design, which might result in reduced transcription or gene silencing. This observation may have implications for transgenic lines in which *EcoRI* endonuclease is driven by other promoters used in this study, although we could not determine this as the TaPSG719, Osg6B, OsRTS, and Zm13 genes and promoters are derived from different species. Furthermore, no effect on pollen germination was observed likely because the catalase intron was inadequately spliced from the pre-mRNA, leading to the generation of a truncated transcript product. This, in turn, caused a frameshift and repeated, premature stop codons. Additionally, this study revealed phenotypic variations in many of the transgenic lines including decreased plant height, increased tiller count, decreased tiller diameter, decreased leaf blade width. In some instances, delayed flowering was also noted. Importantly, these effects are not attributed to off-target *EcoRI* expression. In conclusion, this research provides significant insights into expression patterns, potential plasmid design challenges, and significant variation in switchgrass phenotype. These findings offer valuable insights for the continued development and optimization of a male-sterility system in switchgrass, as well as GE switchgrass which may be produced for various other applications.

References

- Burris, J. N., et al. (2009). "An improved tissue culture system for embryogenic callus production and plant regeneration in switchgrass (*Panicum virgatum* L.)." Bioenergy Research **2**: 267-274.
- Chen, L., et al. (2010). "Isolation and heterologous transformation analysis of a pollen-specific promoter from wheat (*Triticum aestivum* L.)." Molecular Biology Reports **37**: 737-744.
- Echlin, P. (1971). The role of the tapetum during microsporogenesis of angiosperms. Pollen, Elsevier Ltd: 41-61.
- Engler, C., et al. (2008). "A one pot, one step, precision cloning method with high throughput capability." PLoS ONE **3**(11): 3647.
- Fei, S. and E. Nelson (2003). "Estimation of pollen viability, shedding pattern, and longevity of creeping bentgrass on artificial media." Crop Science **43**(6): 2177-2181.
- Ge, Y., et al. (2011). "Pollen viability and longevity of switchgrass (*Panicum virgatum* L.)." Crop Science **51**(6): 2698-2705.
- Goldberg, R. B., et al. (1993). "Anther development: Basic principles and practical applications." The Plant Cell **5**(10): 1217-1229.
- Hague, J. P., et al. (2012). "Pollen sterility—A promising approach to gene confinement and breeding for genetically modified bioenergy crops." Agriculture **2**(4): 295-315.
- Hamilton, D. A., et al. (1989). "Characterization of a pollen-specific genomic clone from maize." Sexual Plant Reproduction **2**: 208-212.
- Hamilton, D. A., et al. (1998). "A monocot pollen-specific promoter contains separable pollen-specific and quantitative elements." Plant Molecular Biology **38**(4): 663-669.
- Hamilton, D. A., et al. (2000). "Comparison of transient and stable expression by a pollen-specific promoter: The transformation results do not always agree." Sexual Plant Reproduction **12**: 292-295.
- Hanson, D. D., et al. (1989). "Characterization of a pollen-specific cDNA clone from *Zea mays* and its expression." The Plant Cell **1**(2): 173-179.
- Haygood, R., et al. (2003). "Consequences of recurrent gene flow from crops to wild relatives." Proceedings of the Royal Society of London. B: Biological Sciences **270**(1527): 1879-1886.
- ISAAA (2019). "Global Status of Commercialized Biotech/GM Crops in 2019: Biotech Crops Drive SocioEconomic Development and Sustainable Environment in the New Frontier." ISAAA Brief No. 55. Retrieved 5 September 2022, from <https://www.isaaa.org/resources/publications/briefs/55/executivesummary/default.asp>.

- Kumar, P., et al. (2017). "Development of male sterile transgenic lines in rice by tapetum-specific expression of *barnase* gene." Biotechnology **44**(4): 364-371.
- Li, R. and R. Qu (2011). "High throughput *Agrobacterium*-mediated switchgrass transformation." Biomass and Bioenergy **35**(3): 1046-1054.
- Luo, H., et al. (2006). "RTS, a rice anther-specific gene is required for male fertility and its promoter sequence directs tissue-specific gene expression in different plant species." Plant Molecular Biology **62**: 397-408.
- Ma, H., et al. (2011). "The modified castor bean catalase intron is incompletely spliced in tobacco and Arabidopsis." Plant Science **181**(2): 188-194.
- McCormick, S. (1993). "Male gametophyte development." The Plant Cell **5**(10): 1265-1275.
- Mellars, G. and K. Gomez (2011). Mutation detection by Southern blotting. PCR Mutation Detection Protocols. B. Theophilus and R. Rapley, Humana Press. **688**: 281-291.
- Millwood, R., et al. (2016). "Engineered selective plant male sterility through pollen-specific expression of the *EcoRI* restriction endonuclease." Plant Biotechnology Journal **14**(5): 1281-1290.
- Ohta, S., et al. (1990). "Construction and expression in tobacco of a β -glucuronidase (GUS) reporter gene containing an intron within the coding sequence." Plant and Cell Physiology **31**(6): 805-813.
- Pfotenhauer, A. C., et al. (2022). "Building the plant SynBio toolbox through combinatorial analysis of DNA regulatory elements." ACS Synthetic Biology **11**(8): 2741-2755.
- Rognli, O., et al. (2000). "Effects of distance and pollen competition on gene flow in the wind-pollinated grass *Festuca pratensis* Huds." Heredity **85**(6): 550-560.
- Somleva, M. N., et al. (2014). "Transgene autoexcision in switchgrass pollen mediated by the Bxb1 recombinase." BMC Biotechnology **14**(1): 1-13.
- Stewart Jr, C. N. and L. E. Via (1993). "A rapid CTAB DNA isolation technique useful for RAPD fingerprinting and other PCR applications." Biotechniques **15**(5): 748-751.
- Stewart Jr., C. N. (2021). Controlling transgene flow from engineered crops to unintended hosts by molecular approaches. Gene Flow: Monitoring, Modeling and Mitigation. W. Wei and C. N. Stewart, CAB International.
- Stewart Jr., C. N., et al. (2003). "Transgene introgression from genetically modified crops to their wild relatives." Nature Reviews Genetics **4**(10): 806-817.

Tanaka, A., et al. (1990). "Enhancement of foreign gene expression by a dicot intron in rice but not in tobacco is correlated with an increased level of mRNA and an efficient splicing of the intron." Nucleic Acids Research **18**(23): 6767-6770.

Tsuchiya, T., et al. (1994). "Molecular characterization of rice genes specifically expressed in the anther tapetum." Plant Molecular Biology **26**: 1737-1746.

Tuinstra, M. R. and J. Wedel (2000). "Estimation of pollen viability in grain sorghum." Crop Science **40**(4): 968-970.

Weigel, D. and J. Glazebrook (2006). "Transformation of *agrobacterium* using the freeze-thaw method." Cold Spring Harbor Protocols **7**.

Wilson, G. G. (1988). "Type II restriction—Modification systems." Trends in Genetics **4**(11): 314-318.

Yokoi, S., et al. (1997). "Tapetum-specific expression of the Osg6B promoter- β -glucuronidase gene in transgenic rice." Plant Cell Reports **16**: 363-367.

Zhan, X. Y., et al. (1996). "Nuclear male sterility induced by pollen-specific expression of a ribonuclease." Sexual Plant Reproduction **9**: 35-43.

Zou, J. T., et al. (1994). "Characterization of a rice pollen-specific gene and its expression." American Journal of Botany **81**(5): 552-561.

Appendix

Supplementary Table S3.1 PCR Primer Sets. List of primer sequences used to amplify PCR targets and their expected band sizes.

PCR Target	Primer Sequences	Product Size (bp)
pPvPS1 into <i>EcoRI</i>	F: 5'-gcgtgcactagaagggaca-3' R: 5'-gcgagatcatgagcttttgc-3'	845
pZm13 into <i>EcoRI</i>	F: 5'-gcgttttgttgagcgcg-3' R: 5'-gcgagatcatgagcttttgc-3'	670
pOsRTS into <i>EcoRI</i>	F: 5'-cgtagcagcacaaaagagatgcg-3' R: 5'-gcgagatcatgagcttttgc-3'	959
pOsg6B into <i>EcoRI</i>	F: 5'-cggcgtgtatgcttccattg-3' R: 5'-gcgagatcatgagcttttgc-3'	1,011
pTaPSG719 into <i>EcoRI</i>	F: 5'-ggcaaggcacggaataggatg-3' R: 5'-gcgagatcatgagcttttgc-3'	709
<i>EcoRI</i>	F: 5'-cgtagcagcacaaaagagatgcg-3' R: 5'-ccatctcggggaaagttcacg-3'	744
<i>ChvA</i>	F: 5'-gttcagcaggccggcatcctgg-3' R: 5'-cgaaacgctgttcggcctgtgg-3'	890

Supplementary Table S3.2 qRT-PCR Primer Sets. List of target genes, primer sequences, primer melting temperatures, and the expected product size of qRT-PCR experiments.

Gene	Primer Sequences	T _m	Product Size (bp)
<i>EcoRI</i>	F: 5'-caaggagatgggagggagtg-3' R: 5'-acaagtcacgccccaacac-3'	59°C 60°C	94
<i>Ubq1</i>	F: 5'-agaaggagtccaccctccat-3' R: 5'-accagacgaccaggtacag-3'	59°C 60°C	93
<i>PvPS1</i>	F: 5'-gctgtccttgagctcgatct-3' R: 5'-ttaggctggagtgcaagcac-3'	59°C 60°C	101

Supplementary Table S3.3 Summary of single- and double- copy transgenic switchgrass lines. List of constructs integrated into transgenic switchgrass, and the quantity of single-copy, double-copy, and combined total of single- and double- copy transgenic lines identified by Southern hybridization.

Construct	Single-copy lines	Double-copy lines	Total lines
pTaPSG719: <i>EcoRI</i>	4	5	9
pPvPS1: <i>EcoRI</i>	3	4	7
pOsg6B: <i>EcoRI</i>	4	2	7
pOsRTS: <i>EcoRI</i>	6	6	12
pZm13 <i>EcoRI</i>	1	3	4

Supplementary Table S3.4 Southern hybridization analysis and newly assigned line numbers. Table provides transgenic switchgrass construct name, the line numbers as referred to in the Southern hybridization experiments, determined transgene integration copy number of the lines, and newly assigned line number of transgenic lines that will be further analyzed.

Construct	Southern Hybridization Sample Number	Transgene Integration Copy Number	Transgenic Line Name
pTaPSG719: <i>EcoRI</i>	1	2	1
	2	1	2
	3	2	3
	4	1	4
	5	1	5
	6	3	n/a
	7	2	6
	8	2	7
	9	1	n/a
	10	1	8
	11	2	9
pPvPS1: <i>EcoRI</i>	1	1	n/a
	2	~7	n/a
	3	2	1
	4	5	n/a
	5	2+	n/a
	6	3	n/a
	7	2	2
	8	1-2	3

Supplementary Table S3.4 continued

	9	1	4
	10	3	n/a
	11	2	5
	12	3	n/a
	13	1	6
	14	3	n/a
	15	2	7
pOsg6B: <i>EcoRI</i>	1	3	n/a
	2	3	n/a
	3	3	n/a
	4	3	n/a
	5	2	1
	6	1	2
	7	2	3
	8	1	4
	9	1-2	5
	10	1	6
	11	4	n/a
	12	1	7
pOsRTS: <i>EcoRI</i>	1	3	n/a
	2	1	1
	3	2	2
	4	2	3
	5	1	4
	6	4	n/a
	7	2	5
	8	2	6
	9	1-2	7
	10	1	n/a
	11	3	n/a
	12	1	8
	13	1	9
	14	1	10
	15	2	11
	16	2	12
pZm13: <i>EcoRI</i>	1	2	1
	2	3	n/a
	3	1	2
	4	2	3
	5	2	4

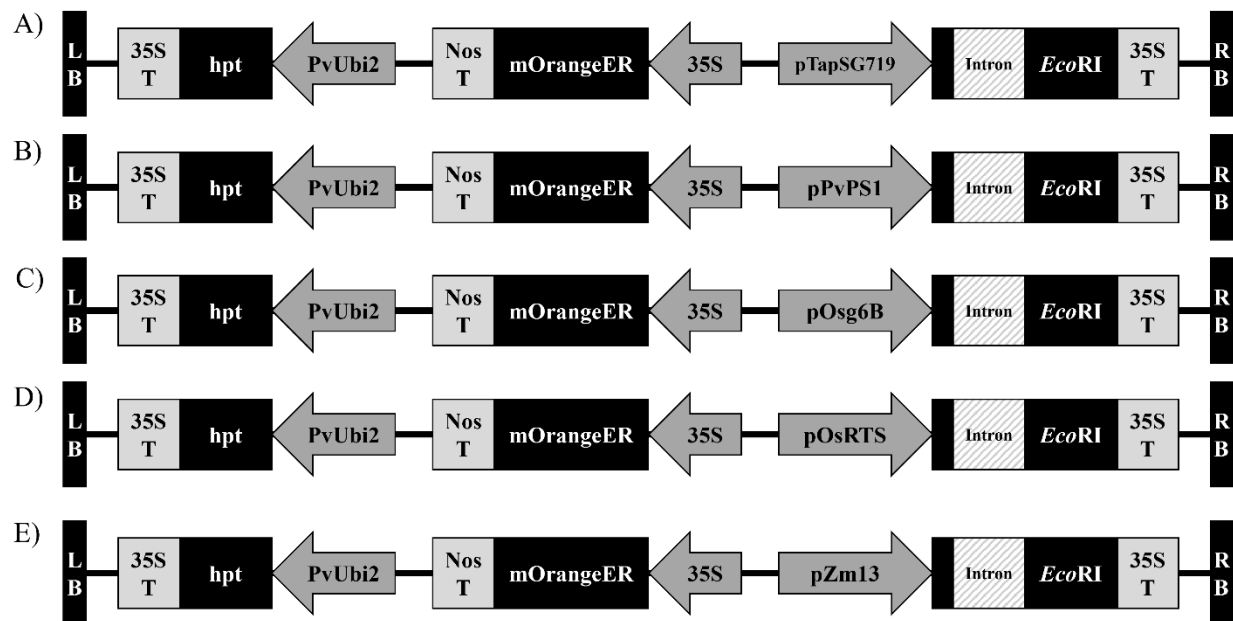


Figure 3.1. *EcoRI* plasmid T-DNA maps. Left (LB) and right (RB) borders flanking the T-DNA insert are represented by blue rectangles. All vectors contain the plant selectable antibiotic hygromycin (hpt) between *Panicum virgatum Ubi2* promoter and the cauliflower mosaic virus (CaMV) 35S terminator. All vectors also contain the visual marker mOrangeER between CaMV 35S promoter and the *Agrobacterium tumefaciens* nopaline synthase terminator (Nos T). All vectors contain the *EcoRI* endonuclease gene, interrupted by a catalase intron (derived from *Ricinus communis*), between the promoter (A) pTaPSG719 (B) pPvPS1 (C) pOsg6B (D) pOsRTS or (E) pZm13 and CaMV 35S terminator.

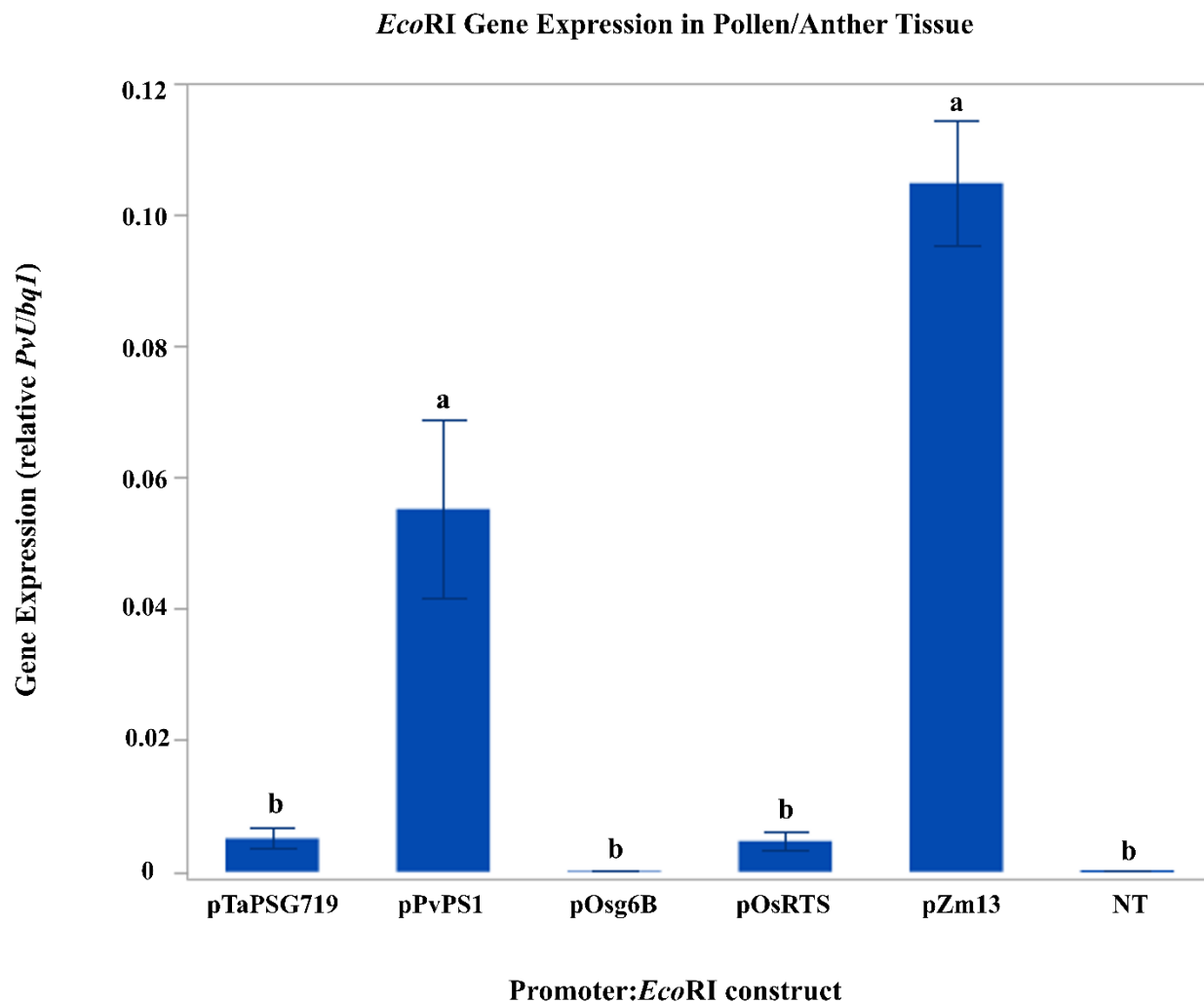
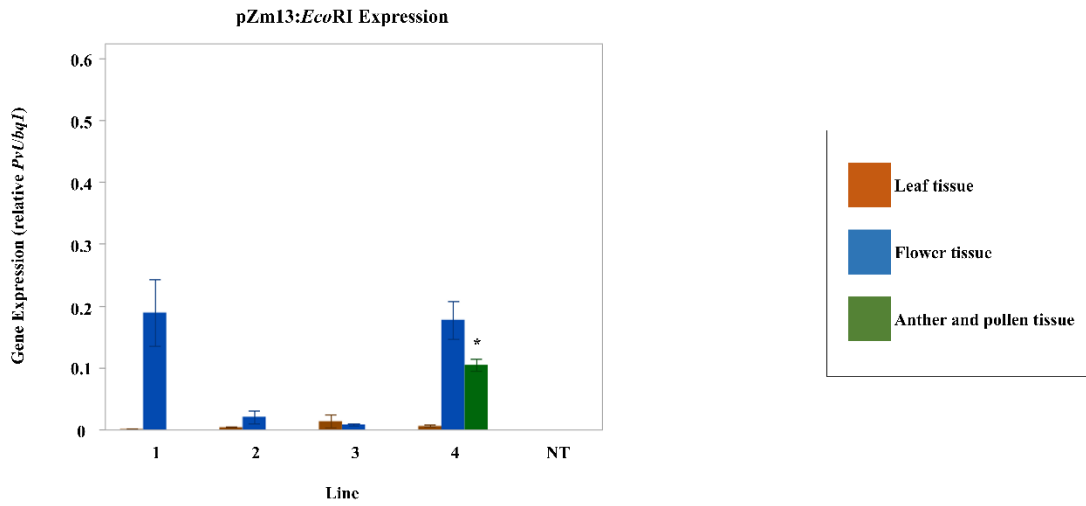
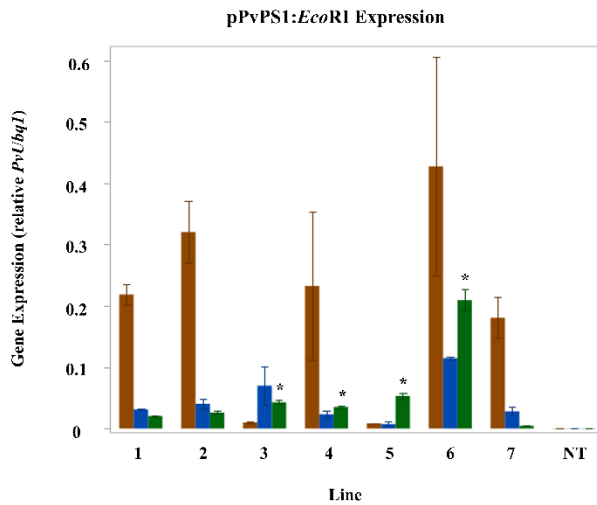


Figure 3.2 Promoter effect of transcript expression of *EcoRI* endonuclease transgene in pollen/anther tissue. Bar graph of *EcoRI* endonuclease transcript levels driven by TaPSG719, PvPS1, Osg6B, OsRTS, and Zm13 promoters and a nontransgenic (NT) control. Expression is relative to the switchgrass ubiquitin 1 gene. Different letters represent statistical significance as tested by a one-way ANOVA, followed by a Tukey's HSD, $\alpha = 0.05$ ($F(5, 102) = 17.4$, $p < .0001$). Error bars represent standard error.

A)



B)



C)

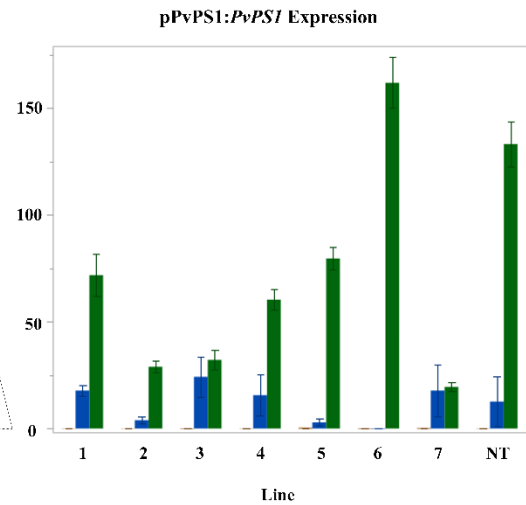


Figure 3.3 Graphs of transcript abundance data. Graphs of *EcoRI* transcript levels in leaf (orange), whole flower collected prior anthesis (blue), and combined pollen and anther tissue collected immediately after anthesis (green) in **(A)** pZm13: *EcoRI* transgenic lines 1 through 4 and **(B)** pPvPS1:*EcoRI* transgenic lines 1 through 7, and nontransgenic controls (NT). **(C)** Endogenous *PS1* expression levels, driven by the native pPvPS1 promoter, in the pPvPS1:*EcoRI* transgenic lines 1 through 7 and a nontransgenic control. All transcript levels are set relative to the switchgrass *Ubiq1* housekeeping gene. An asterisk (*) represents statistical significance of *EcoRI* expression in anther/pollen tissue of transgenic lines compared to the nontransgenic control, tested by a one-way ANOVA, Tukey's HSD, $\alpha = 0.05$. Error bars represent standard error.

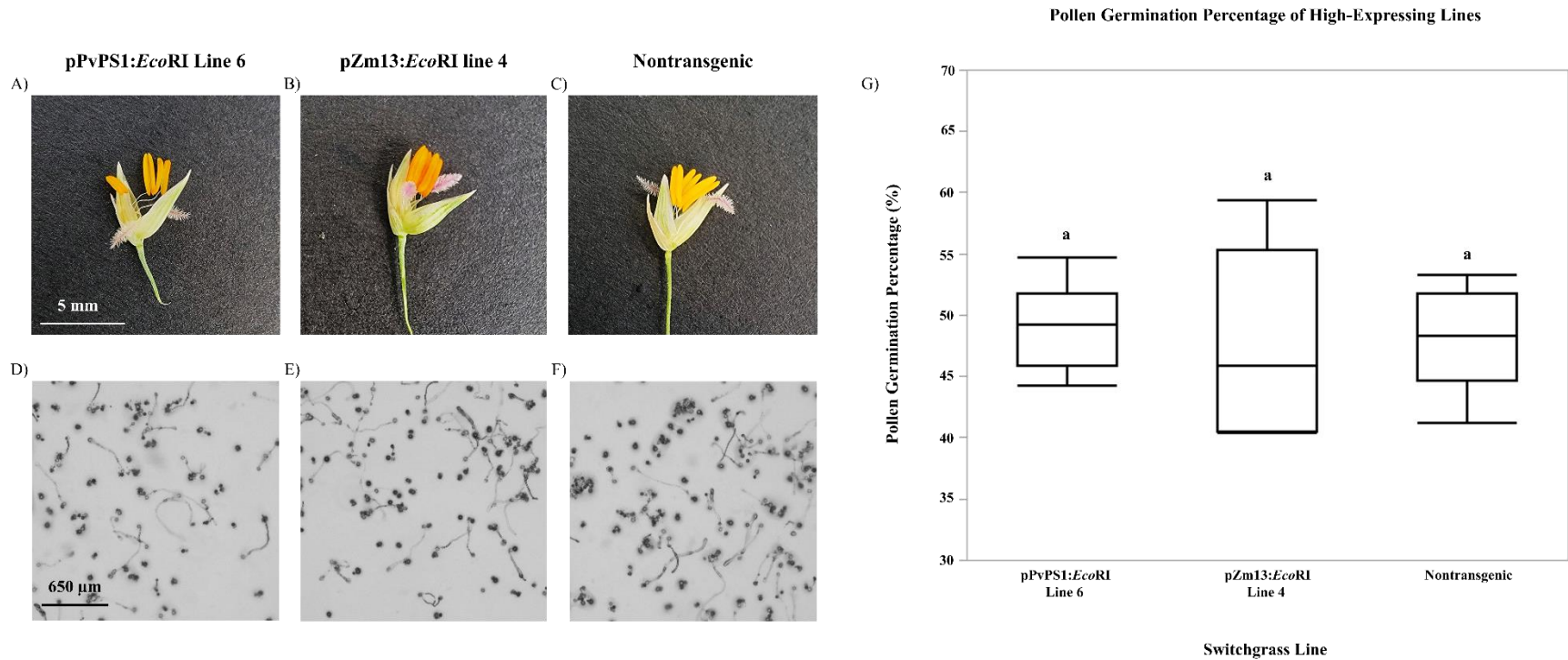


Figure 3.4 Pollen germination analysis. Images of flowers and pollen of (A,D) pPvPS1:EcoRI line 6, (B,E) pZm13:EcoRI line 4, and (C,F) nontransgenic switchgrass. (G) Box plot graph of pollen germination percentage (%) analysis, comparing pPvPS1:EcoRI line 6, pZm13:EcoRI line 4, and a nontransgenic control. Same letters represent no statistical significance of pollen germination percentages between transgenic lines and a nontransgenic control, tested by a one-way ANOVA, Tukey's HSD, $\alpha = 0.05$ ($F(2, 24) = 0.18$, $p = 0.84$).

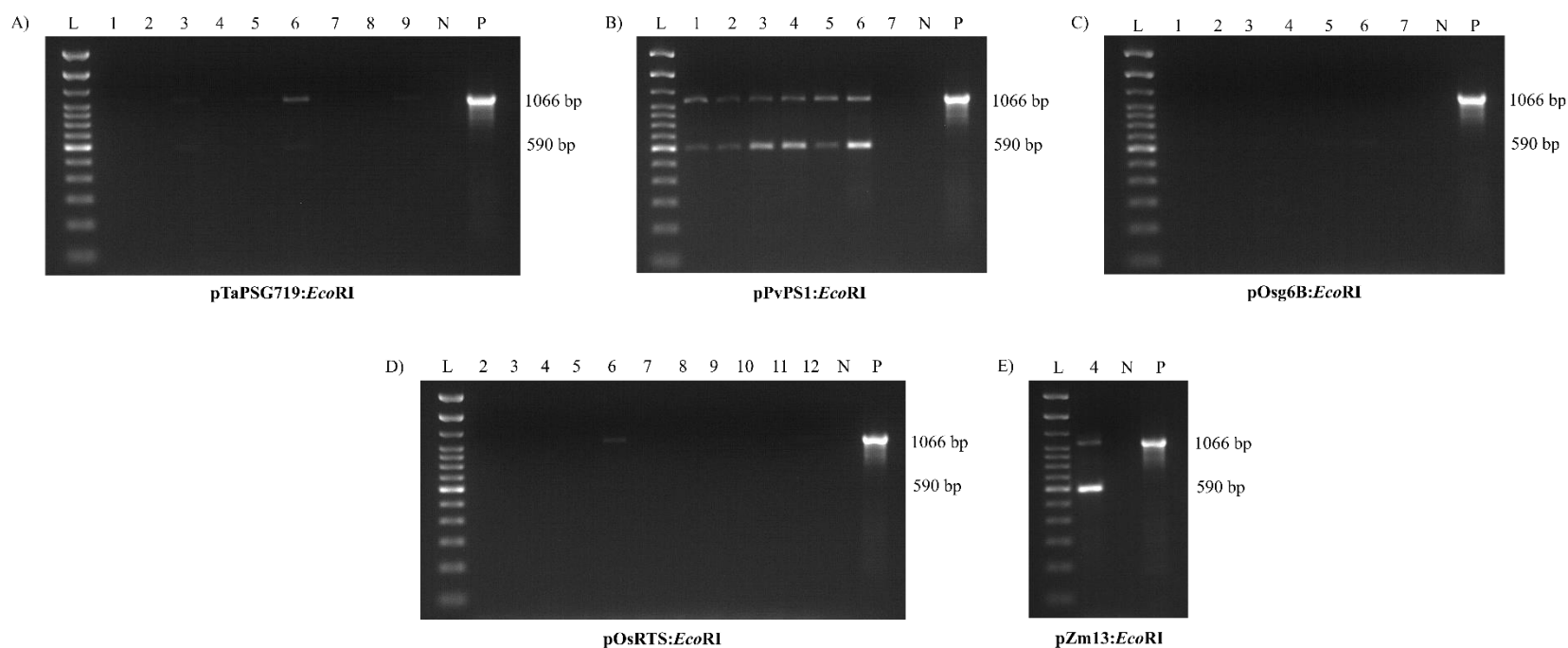


Figure 3.5 RT-PCR analysis for presence/absence of catalase intron. PCR analysis of cDNA extracted from anther/pollen tissue of flowering transgenic lines of the following constructs: **(A)** pTaPSG719:*EcoRI*, **(B)** pPvPS1:*EcoRI*, **(C)** pOsg6B:*EcoRI*, **(D)** pOsRTS:*EcoRI*, **(E)** pZm13:*EcoRI*. PCR amplified products represent *EcoRI* endonuclease with the presence of the catalase intron (1066 bp) or the absence of the catalase intron (predicted ~860 bp). The observed band in the absence of the catalase intron is 590 bp. Gels include the TrackIt™ 100 bp DNA Ladder (L), transgenic line numbers of their associated construct, cDNA from nontransgenic switchgrass as a negative control (N), and plasmid DNA as a positive control (P) for the presence of the intron.

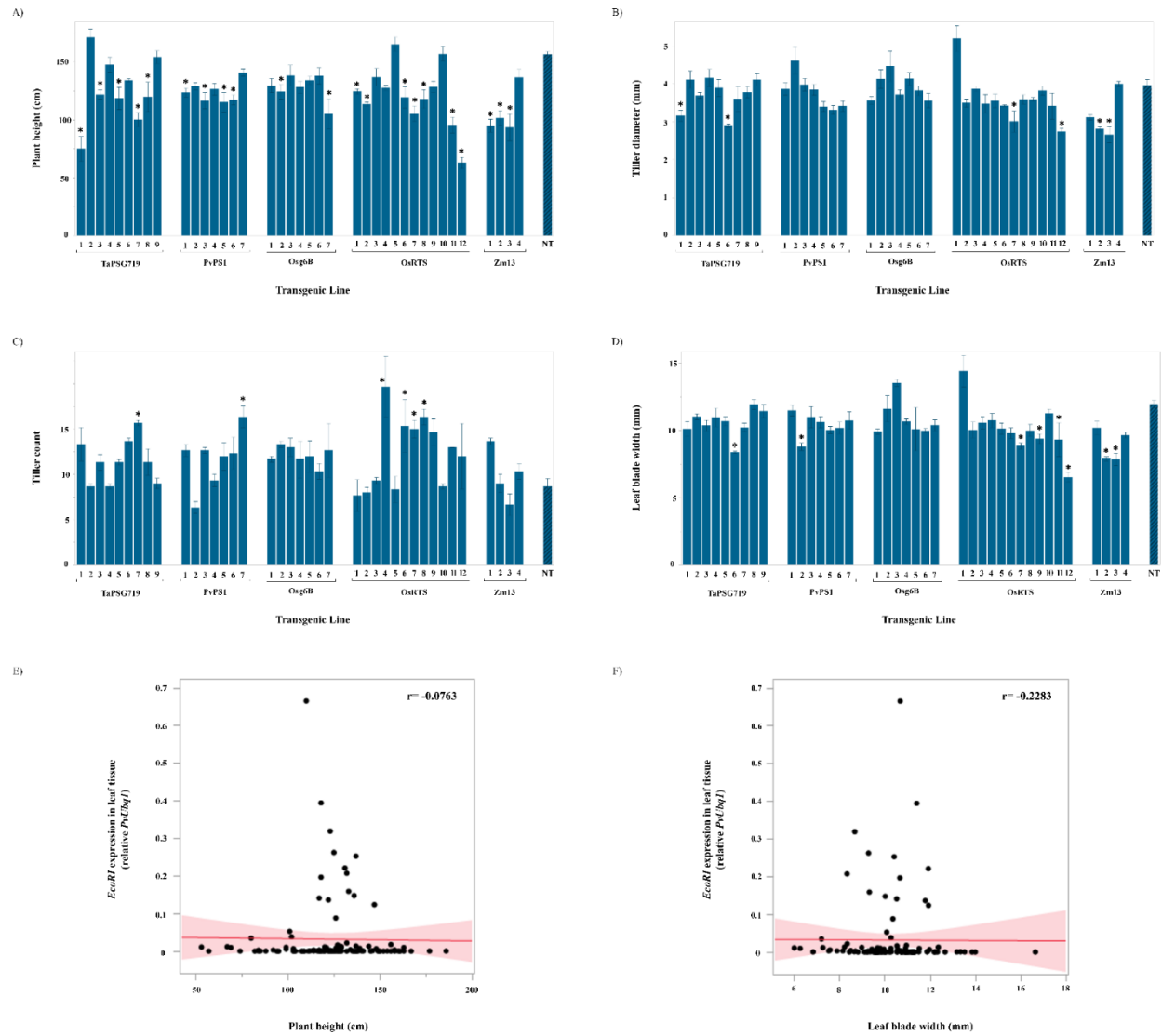
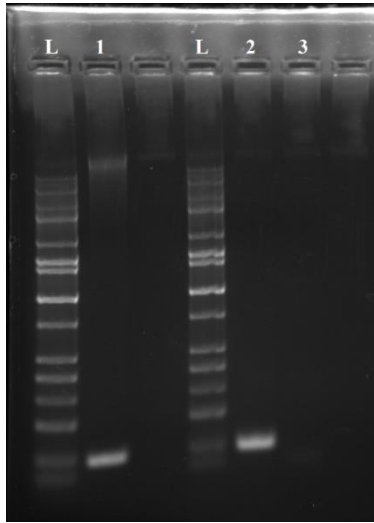
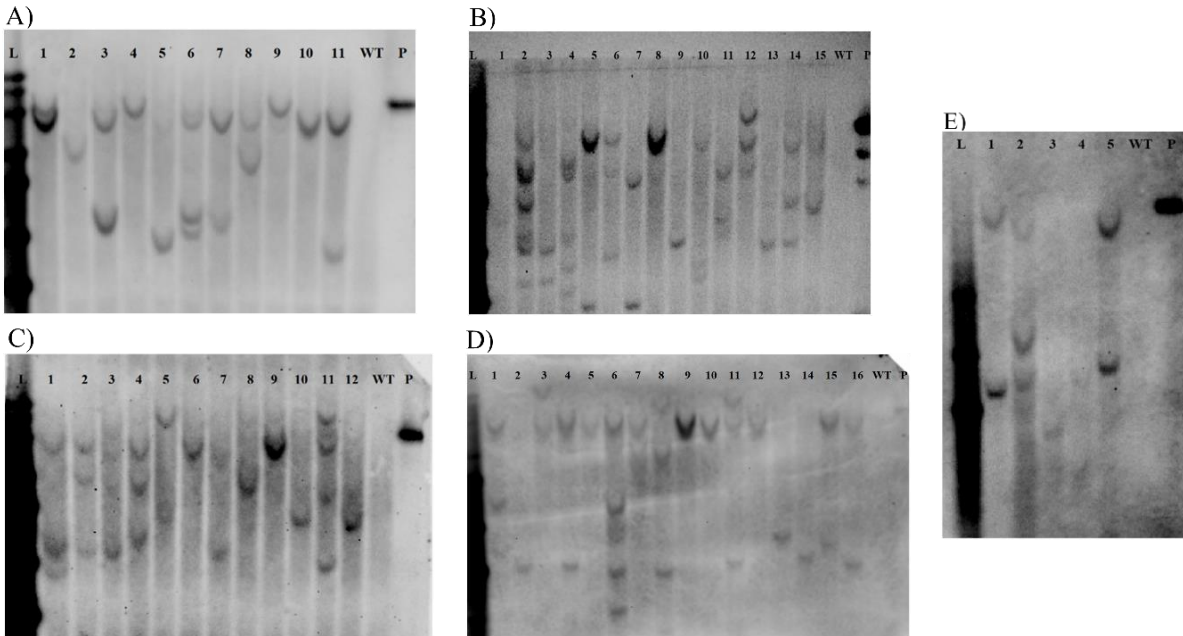


Figure 3.6 Phenotypic analysis of transgenic lines. Bar graphs of (A) plant height, (B) tiller count, (C) tiller diameter, and (D) leaf blade width of transgenic lines, grouped by promoter driving *EcoRI* expression, compared to a nontransgenic control (c). Statistical significance is denoted by an asterisk (*), as tested by a one-way ANOVA, followed by a Dunnett's test, $\alpha = 0.05$. Error bar is constructed using one standard error from the mean. Scatterplots depict the relationship of *EcoRI* expression levels in leaf and (E) plant height or (F) leaf blade width using Spearman's correlation analysis.

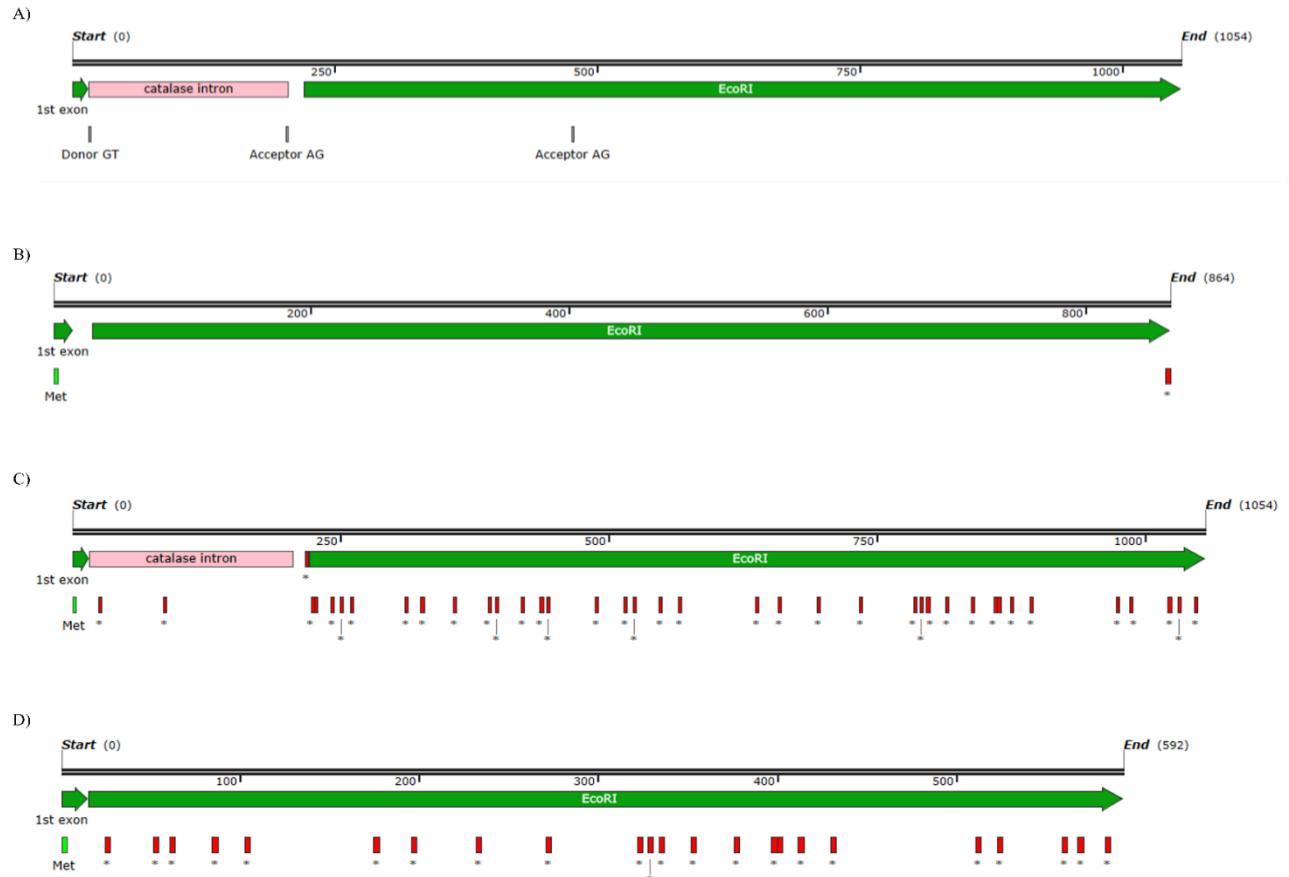


Supplementary Figure S3.1 Confirmation of qRT-PCR primer sets. Prior qRT-PCR analysis of transgenic switchgrass, the *EcoRI* and *Ubq1* primer sets were assessed by RT-PCR and observed on a 1.2% agarose gel with the All Purpose Hi-Lo DNA Marker (L). The *EcoRI* primers were confirmed using genomic DNA from transgenic switchgrass carrying the *EcoRI* transgene to test for a positive band (94 bp) observed in sample 1. The *Ubq1* and *EcoRI* primers were tested using cDNA from nontransgenic switchgrass to confirm a positive band (93 bp) for the *Ubq1* gene observed in sample 2, and the absence of the *EcoRI* gene observed in sample 3.



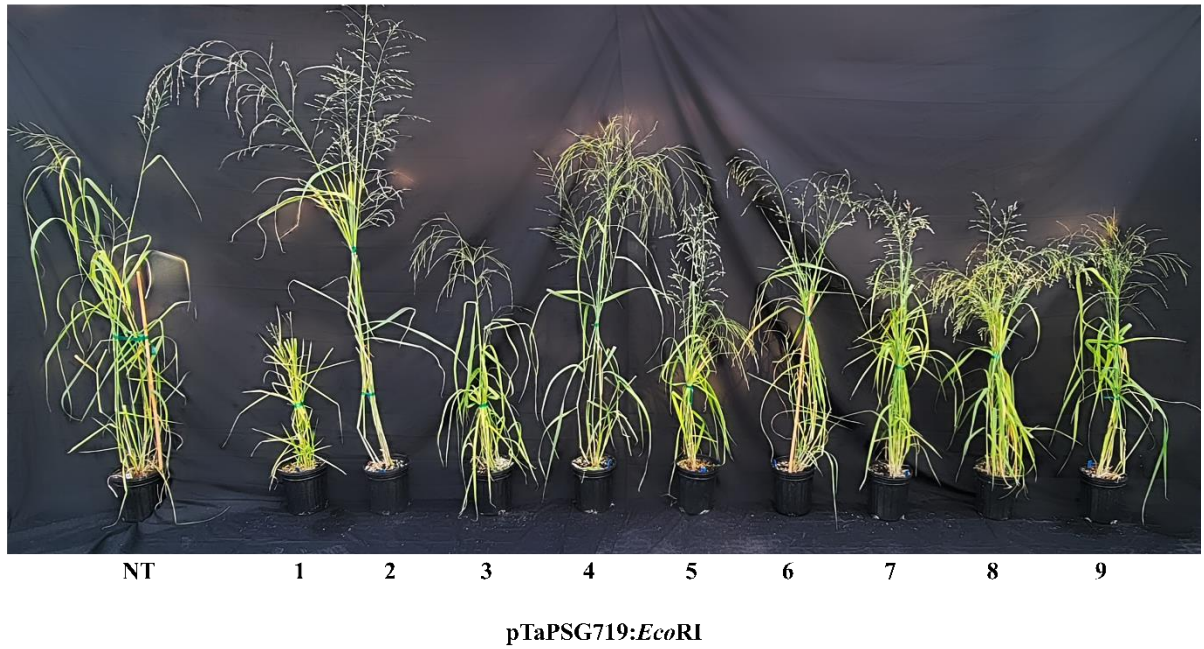
Supplementary Figure S3.2 Southern blot hybridizations of transgenic switchgrass.

Genomic (10 µg) and plasmid (500 pg) DNA was digested with *EcoRI*-HF and *SspI*-HF enzymes. Blots were probed with the DIG labeled-*EcoRI* gene fragment and hybridized at 42°C. Lanes consist of the DNA Ladder (L), genomic DNA samples of **(A)** pTaPSG719:*EcoRI* transgenic lines 1-11 **(B)** pPvPS1:*EcoRI* transgenic lines 1-15 **(C)** pOsg6B:*EcoRI* transgenic lines 1-12 **(D)** pOsRTS:*EcoRI* transgenic lines 1-16 **(E)** pZm13:*EcoRI* transgenic lines 1-5, followed by genomic DNA of wild-type switchgrass (WT) and the relevant plasmid DNA (P).

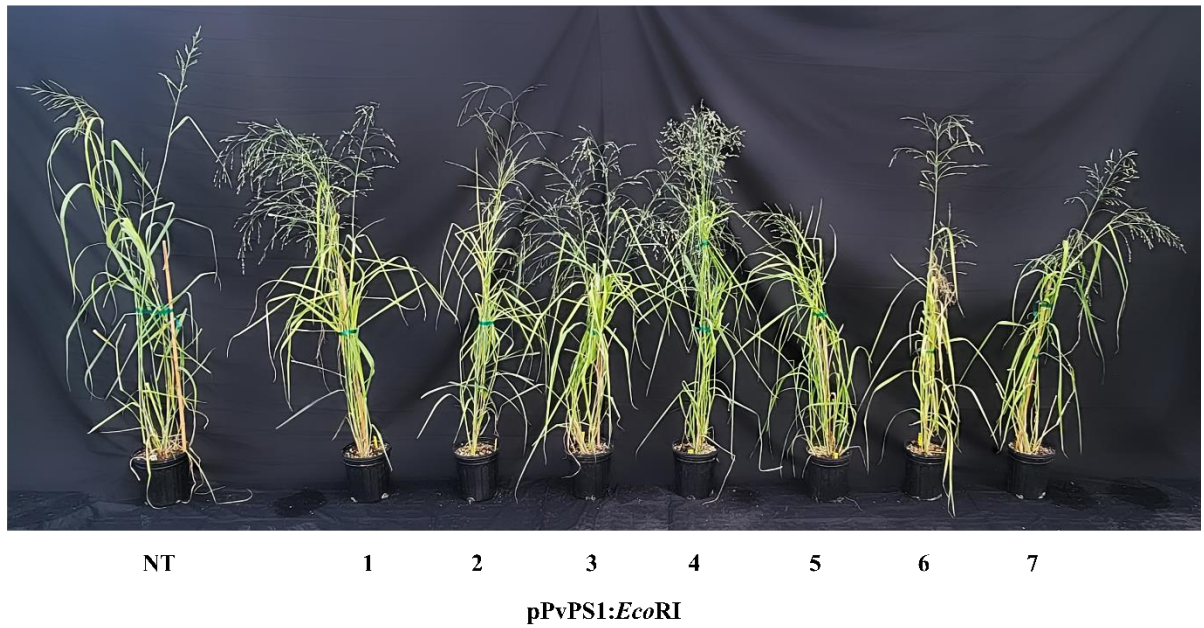


Supplementary Figure S3.3 Map of *EcoRI* endonuclease gene and the expected and observed results of splicing. (A) Map includes the *EcoRI* endonuclease gene interrupted by the catalase intron. cDNA sequencing identified the donor GT splice site is properly recognized and spliced from pre-mRNA. The first acceptor AG splice site downstream the donor site is the other intended splice site, however sequencing confirmed that the second acceptor AG is where splicing is occurring. (B) Map of the *EcoRI* endonuclease gene and the predicted start codon (Met) (green rectangle) and stop codon denoted by an asterisk (*) (red rectangle) if the catalase intron were properly removed. (C) Map of *EcoRI* endonuclease gene and the start and stop codons when the catalase intron is not spliced out. (D) Map of *EcoRI* endonuclease and the start and stop codons when splicing occurs at the second acceptor AG, which is observed in this study.

A)



B)



Supplementary Figure S3.4 Images of transgenic lines following eight-week growth period.

Images of nontransgenic switchgrass control (NT) compared to transgenic lines of (A) pTaPSG719:*EcoRI* (B) pPvPS1:*EcoRI* (C) pOsg6B:*EcoRI* (D) pOsRTS:*EcoRI* and (E) pZm13:*EcoRI* constructs after eight weeks of growth under greenhouse condition.

C)



NT

1

2

3

4

5

6

7

pOsg6B:*EcoRI*

D)



NT

1

2

3

4

5

6

7

8

9

10

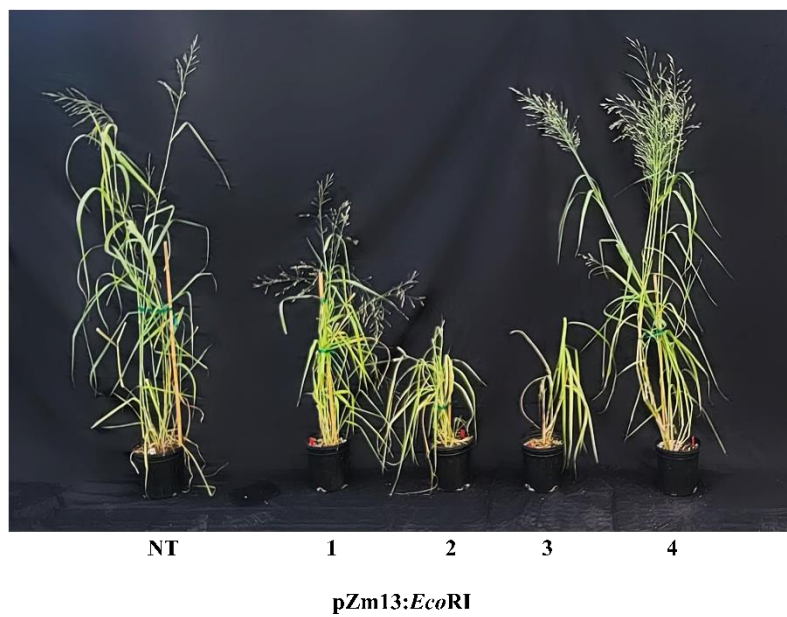
11

12

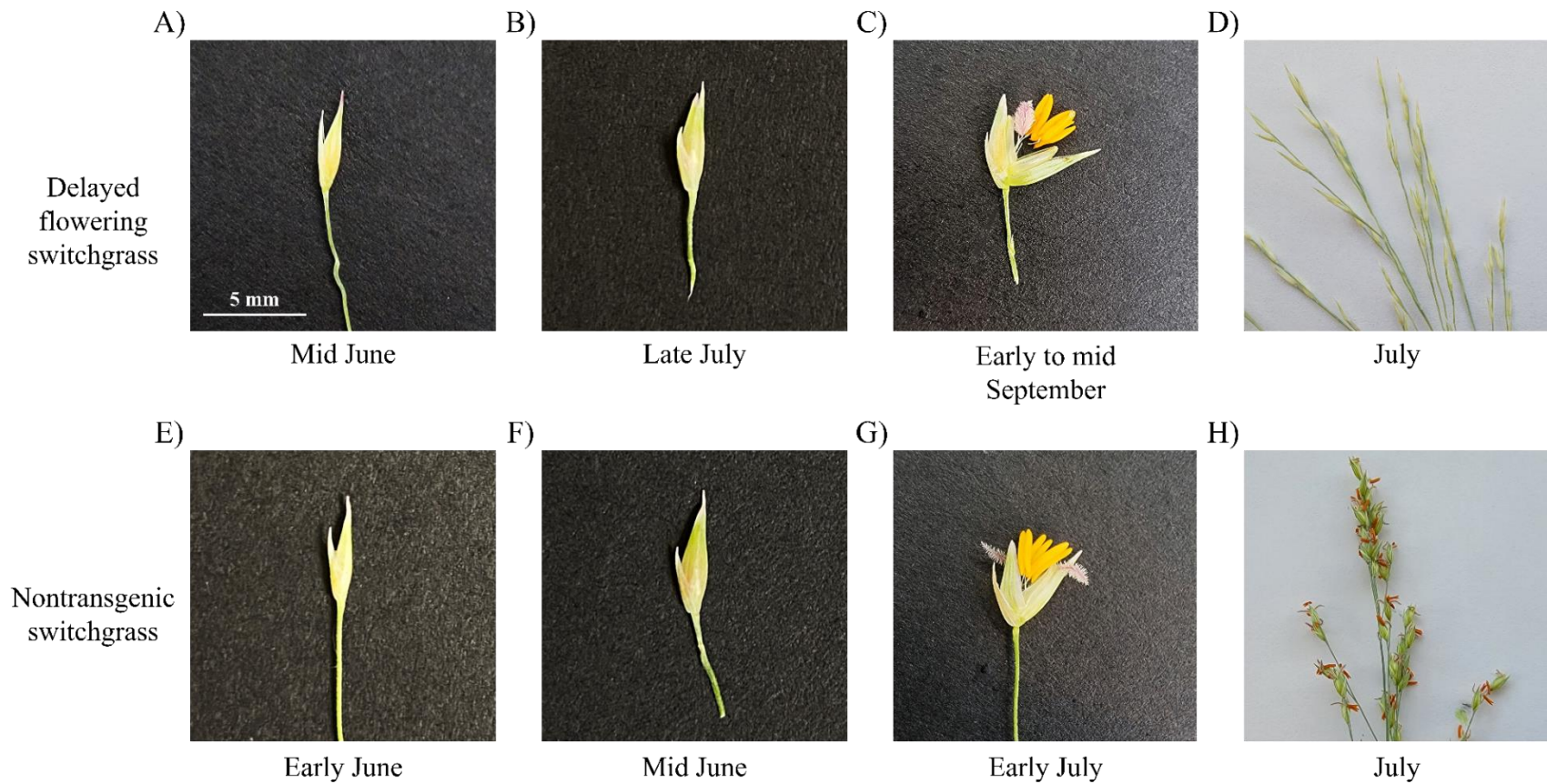
pOsRTS:*EcoRI*

Supplementary Figure S3.4 continued

E)

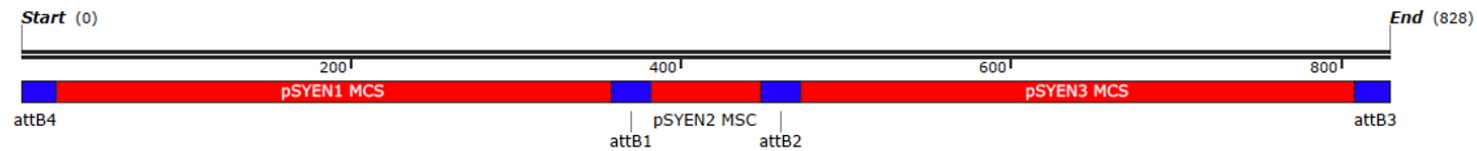


Supplementary Figure S3.4 continued

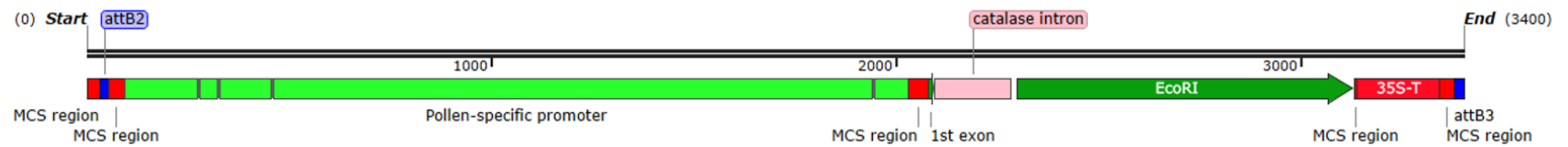


Supplementary Figure S3.5 Delayed flowering phenotype. OsRTS:*EcoRI* lines 1 and 2 displayed a delayed flowering phenotype where the (A) flowers remained thin and underdeveloped into reproductive stage 3 (R3) for an extended period of time (~ 2 months) before (B) floral tissues further developed to produce a fuller flower appearance, and (C) anthesis of normal-appearing flowers was first observed in the second week of September. Nontransgenic flowers transitioned from (E) thin and underdeveloped to a (F) swollen appearance within approximately one to two weeks of reaching R3, and (G) anthesis began in early July. This resulted in (D) closed flowers of the apparently delayed-flowering lines during the (H) flowering period of the nontransgenic switchgrass.

A)



B)



Supplementary Figure S3.6 Cloning methodology of the *EcoRI* endonuclease gene cassette. (A) Multiple cloning site regions (red) used to clone each of the gene cassettes, bordered by attB sites (blue) produced by Golden Gate cloning. (B) The *EcoRI* endonuclease gene cassette that was cloned into the pSYEN3 MCS. Between the pollen-specific promoter and the 1st exon and the *EcoRI* coding sequence is a 49 bp cloning “scar”, a possible cause for low transcript expression.

Chapter 4: Conclusions and Recommendations

I believe the pollen-targeted expression of *EcoRI* endonuclease holds the potential to effectively induce male-sterility for bioconfinement purposes; however, improvements are necessary, beginning with a more robust design of the plasmid construct. While the catalase intron was chosen as a potential solution to leaky *EcoRI* expression in *Agrobacterium* under some of the anther- and pollen- specific promoters, it failed to undergo proper splicing from mRNA when expressed in switchgrass. The intron should not be included in the constructs where it is nonessential. For those where leaky expression was observed in *Agrobacterium*, an alternative route should be taken to prevent the leaky expression. Furthermore, equipping the transgene with 3' and 5' untranslated regions could stabilize mRNA and enhance gene expression. Additionally, utilizing a nuclear localization signal would ensure the *EcoRI* endonuclease is effectively trafficked to the nucleus, increasing the potential of *EcoRI* endonuclease to induce genomic DNA damage and cellular death. I also emphasize the importance of using characterized tissue-culture switchgrass lines for the production of genetically engineered switchgrass which is vital to minimize phenotypic variation arising from the high genetic diversity of switchgrass and to either produce a large selection of transgenic lines or utilize a site-targeted integration approach to reduce the random, potentially deleterious nature of T-DNA insertion, respectively. By incorporating these strategic considerations, the full potential of *EcoRI* endonuclease might be harnessed to achieve bioconfinement in switchgrass.

Vita

Jessica Stockdale grew up in Chattanooga, TN. She graduated from the University of Tennessee Knoxville in 2020 with a Bachelor of Science in Biological Sciences, with a concentration in biochemistry and cellular and molecular biology. In August of 2020, she began pursuing her Master of Science degree in Plant Sciences with a concentration in plant molecular genetics at the University of Tennessee Knoxville under the mentorship of Dr. Neal Stewart. The objectives of her work were to characterize putative pollen-specific promoters in switchgrass, as well as determine the efficacy of pollen-specific *EcoRI* endonuclease expression to confer pollen-sterility for bioconfinement purposes.