Paramutation at the *Pl1-CML52* haplotype

Honors Research Thesis

Presented in partial fulfillment of the requirements for graduation *with honors research distinction* in Molecular Genetics in the undergraduate colleges of The Ohio State University

by

Lauren Masters

The Ohio State University

April 2023

Project Advisor: Dr. Jay Hollick, Department of Molecular Genetics

Table of Contents

Abstract

Paramutation involves meiotically heritable changes in gene regulation that do not follow simple dominant-recessive, Mendelian genetics. In *Zea mays* (maize), this process is exhibited among specific alleles of the *pl1* (*purple plant 1*) gene, which encodes a transcription factor that regulates synthesis of anthocyanins. As an example, the *Pl1-Rhoades* (*Pl1-Rh*) allele can exist in different epigenetic states reflected by visible pigments. In the reference state (*Pl-Rh*), dark purple anther color is observed due to the formation of anthocyanins. In the transcriptionally and post-transcriptionally repressed state (*Pl′*), yellow anther color is observed, and anthocyanins are not produced. As a clear example of paramutation, the *Pl-Rh* state is invariable converted to *Pl′* state in *Pl-Rh / Pl′* heterozygotes with the presence of the *Pl′* state and is meiotically inherited at *Pl′* to future generations. A key regulatory feature of the paramutation behavior involves a set of tandem repeats found downstream of the *Pl1-Rhoades* gene model. (1-2)

Among 24 founder lines of a nested association mapping population, CML52 strain is the only one harboring a *pl* allele where paramutation also occurs. In the *Pl1-Rhoades* haplotype, only one copy of the *pl1* gene is present. However, in the *Pl1-CML52* haplotype of maize there are two copies of the *pl1* gene, denoted *Pl1-CML52a* and *Pl1-CML52b.* Given the structural similarities and difference between the *Pl1-Rhaodes* and *Pl1-CML52* haplotypes it was investigated whether one or both of *Pl1-CML52* genes were subjected to silencing in the presence of *Pl′.* Also, it was investigated whether molecular differences in cytosine methylation at putative regulatory features were associated with this particular example of paramutation.

Introduction

The understanding of genetics and inheritance patterns first originated from Gregor Mendel and his laws. However, exceptions to these laws started becoming increasingly common and broadened the understanding of gene regulation, one example being paramutation. Paramutation is a non-Mendelian inheritance pattern defined as a meiotically heritable alteration of gene regulation based on allelic interactions (3). In maize, the *Pl1-Rhoades (Pl1-Rh)* allele of the *purple plant1* (*pl1*) gene exists in two functionally distinct epigenetic states. The reference *Pl-Rh* state has in dark purple anthers while anthers in the epigenetically repressed *Pl′* state relatively few cells express purple pigmentation (4). (Figure 1) The *Pl′* state is transcriptionally and post-transcriptionally more repressed than *Pl-Rh* state. (5-6) As a clear example of paramutation, *Pl-Rh* is invariably changed to *Pl′* when exposed to *Pl′* in a heterozygous condition and only *Pl′* states allele are sexually transmitted (7).

Through forward mutation screens, genetic factors *required to maintain repression* (*rmr*) of the *Pl′* state have been identified including components that are involved in RNA polymerase IV (Pol IV)-dependent 24 nucleotide (nt) small RNA (sRNA) production that, in *Arabidopsis* thaliana help target *de novo* cytosine methylation through a pathway known as RNA-directed DNA methylation (RdDM) (8-15). (Figure 2) Pol IV-dependent 24nt sRNAs map to five tandem repeats (penta-repeat) ~16 kb downstream of the *Pl1-Rh* coding sequence which is important for *pl1* paramutation (16). *Pl-Rh / Pl-Rh* homozygotes are highly unmethylated whereas *Pl′ / Pl′* plants heterozygotes are highly methylated at the unique subregion (USR) within each repeat unit (Figure 3).

Pl1-Rhoades, a well-studied haplotype in this lab, has only one *pl1* gene (17), the inbred line CML52 contains two *pl1* gene copies (18); one of which (*Pl1-CML52a*) is identical in sequence to that found in *Pl1-Rh* and the other (*Pl1-CML52b*) has multiple polymorphisms with *Pl1-Rh* (19). Downstream of *Pl1-CML52b*, a single repeat unit precedes the *Pl1-CML52a* gene followed by two additional repeats found in the same position as that of the penta-repeat in *Pl1-Rh* (Figure 4).

The aim of this project was to develop a method to distinguish *Pl1-CML52a* and *Pl1-CML52b* mRNA levels in different Pl1-CML52 and *Pl′* backcrosses, allowing us to understand the relationship between *Pl1-CML52a* and *Pl1-CML52b* expression in the CML52 inbred line. The *Pl1-CML52* haplotype contains three repeats, which differs from the *Pl1-Rhoades* haplotype which contains five tandem repeats serving as a putative enhancer downstream of *Pl1-Rh* coding region. A unique sub-repeat (USR) region, which is unique in the maize genome, other than the penta-repeat was determined in the repeat sequences to be a good location to investigate cytosine methylation. (Figure 5) Another goal of this project was to determine if the methylation patterns differ from *Pl1-CML52* and in*Pl1*-CML52′. Prior experiments indicated an increase in methylations in *Pl1-Rh* when *Pl′* is present (20). Other genes that exhibit paramutation in maize, *B* gene and *R* gene, as the number of repeats increase, the strength of paramutation increases. With these experiments we are determining if a similar behavior is present in the *pl1* gene. *Pl1-CML52* is less paramutagenic than *Pl1-Rhoades*, if there is a possible connection to the number of repeats and pl1 genes and the strength of paramutation. (2)

Figure 1 | Two epigenetic states.
Genetically identical siblings exhibiting contrasting *Pl1-Rh* phenotypes.

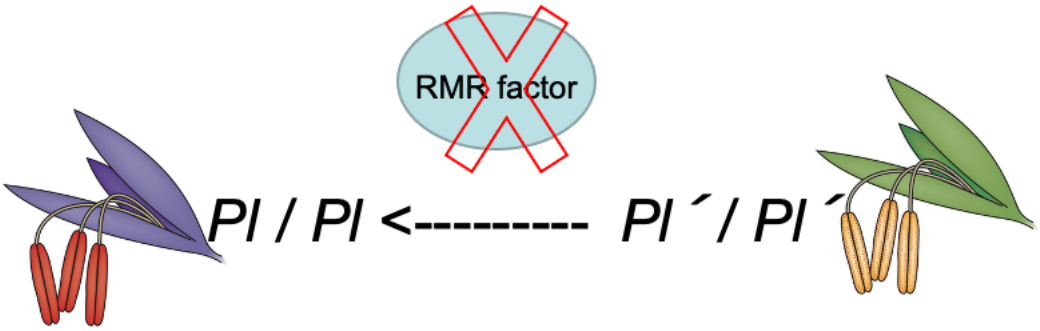

Figure 2 | Reversion back to *Pl* state.
When RMR factors are not present, *Pl′* reverts to the *Pl* state with purple anther color, shown by arrow.

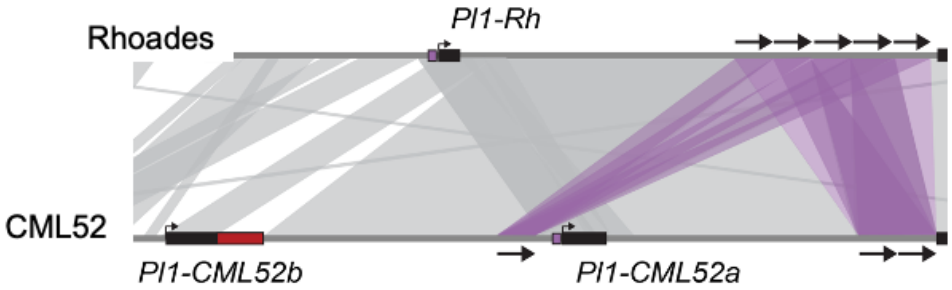

Figure 3 | Comparison of haplotypes.
Pl1-CML52 and *Pl1-Rhoades* gene and repeat region comparison. Repeat regions are shown with large black arrows and purple rays indicate the same sequence. Small black arrows indicate transcriptional start sites.

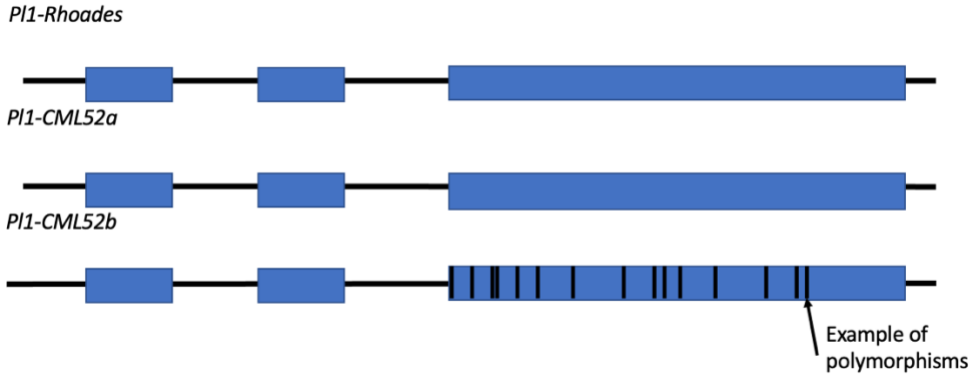

Figure 4 | Exon comparison of haplotypes.
Pl1-Rh, *Pl1-CML52a*, and *Pl1-CML52b* gene model. Blue boxes indicate exons. Black lines on *Pl1-CML52b* indicates the abundance of polymorphisms at relative positions in the exon 3 region.

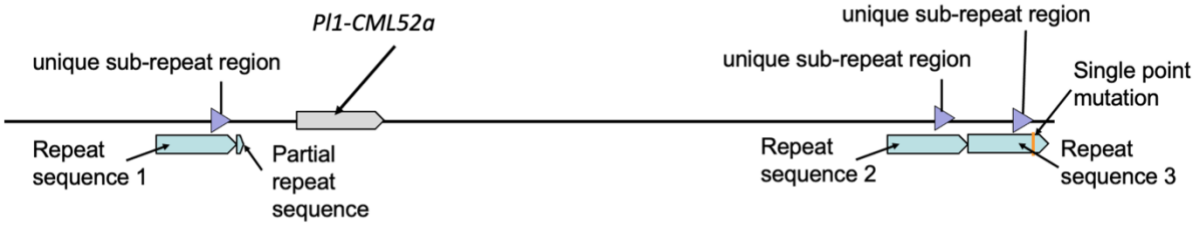

Figure 5 | Diagram of repeats and *Pl1-CML52a*
This diagram shows the location of the three (3) repeat sequence candidate enhancers, relative to *Pl1-CML52a*. Repeat sequence 1 is located up stream with a partial repeat sequence directly downstream. ~16kb downstream of *Pl1-CML52a* is repeat sequence 2 directly followed by repeat sequence 3, which contains one single point mutation.

Materials and Methods

Genetic materials:

The waxy gene is pseudo-linked to the *Pl1-Rh* allele and is used to track the *Pl′* allele. Waxy acts in a dominant-recessive pattern, therefore when there are two waxy alleles (wx/wx) the endosperm cells of the kernel are yellow and opaque. When a nonwaxy allele (Wx) and waxy allele (wx) are present, the endosperm cells are orange and partially translucent. To create the *Pl′ / Pl1-CML52* progeny, 201056, material this pseudo-linkage was exploited. The *Pl1-CML52* haplotype was tracked with the nonwaxy (Wx) allele and the *Pl′* state *Pl1-Rh* haplotype was tracked with the waxy (wx) allele. *Pl′* wx / (*Pl1-CML52* Wx/ *Pl* wx) was crossed and half of the kernels were *Pl′* wx/ *Pl1-CML52* Wx and half were *Pl′* wx / *Pl* wx. They were then separated based on the waxy gene phenotype of waxy *Pl′* wx / *Pl* wx or nonwaxy *Pl′* wx/ *Pl1-CML52* Wx. Nomenclature was simplified in this paper without the waxy gene, nonwaxy is therefore *Pl′ / Pl1-CML52* and waxy is *Pl'.* (Figure 6) *Pl′* wx / *Pl* wx was used as a positive control.

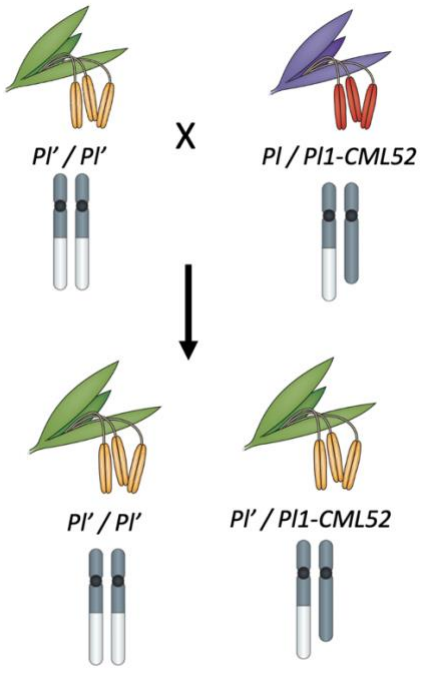

Figure 6 | *Pl′ / Pl1-CML52* Cross. Cross of (*Pl′* wx / *Pl′* wx) / (*Pl1-CML52* Wx/ *Pl* wx) and the progeny *Pl′* / *Pl1-CML52.*

To create *Pl1*-CML52′, *Pl1-CML52* must first be crossed with *Pl′* to add paramutation. *Pl1-CML52* was tracked with Wx (*Pl1-CML52* Wx) and *Pl′* was tracked with wx (*Pl′* wx*).* The F1 progeny were then self-crossed. The nonwaxy plants (*Pl1-CML52′*Wx) fully fertile, plump pollen plants were selected so *Pl'* is not present. Heterozygous plants for the interchange pair will have sterile pollen which are empty cell anthers. The (*Pl1*-CML52′) progeny were then crossed with (*Pl′* wx) again to continue the exposure to the paramutant state. This was then backcrossed 4 times. The final (*Pl1-CML52′*) progeny, 211751, now is in the paramutant state. (Figure 7)

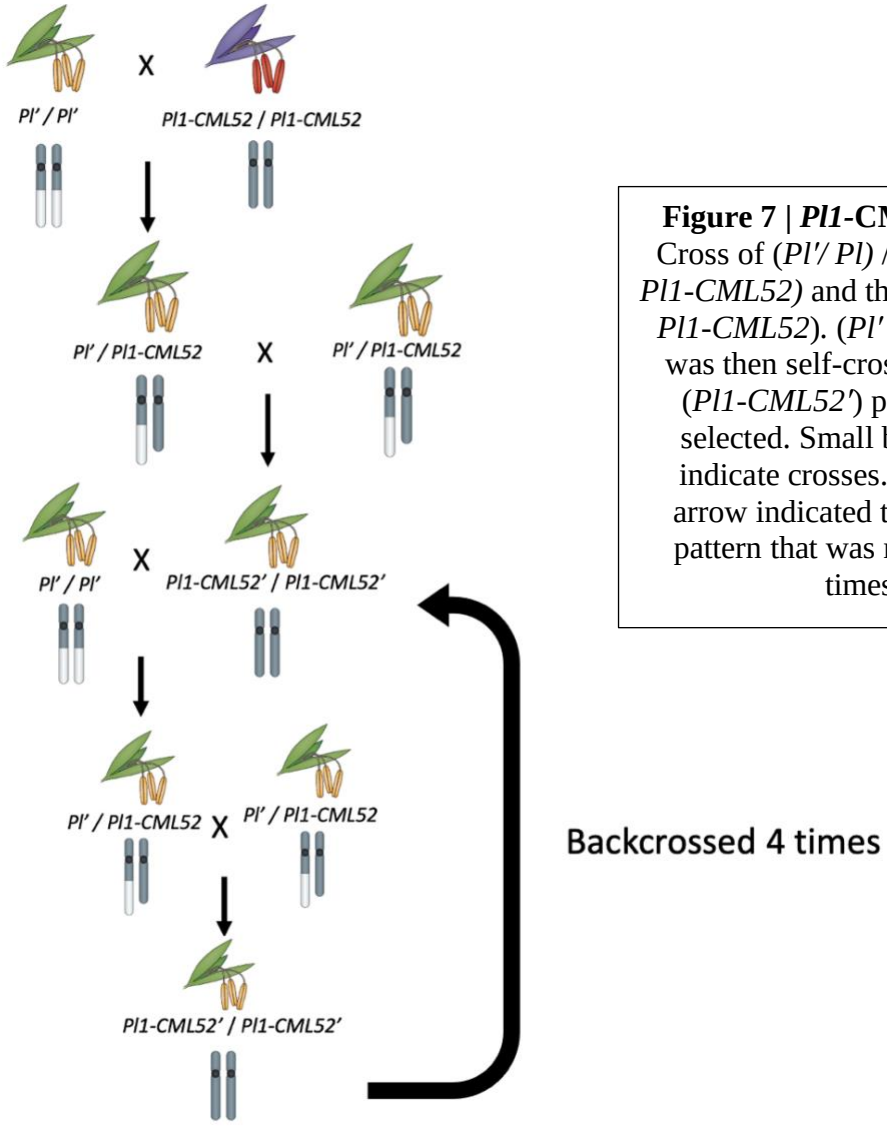

Figure 7 | *Pl1*-CML52′ Cross. Cross of (*Pl′/ Pl)* / (*Pl1-CML52/ Pl1-CML52)* and the progeny (*Pl′* / *Pl1-CML52*). (*Pl′* / *Pl1-CML52*) was then self-crossed. Then F2 (*Pl1-CML52′*) progeny were selected. Small black arrows indicate crosses. Large black arrow indicated the backcross pattern that was repeated four times.

mRNA expression analyses:

The genes *Pl1-CML52a* and *Pl1-CML52b* genes are both homologs to the single gene in *Pl1-Rhoades*. Gene-specific primers were used to differentiate and determine the mRNA expression levels of these genes. *Pl1-CML52b* has multiple single polymorphism sites when comparing *Pl1-Rhoades* and were exploited to create gene-specific primers. (Figure 7) The sequence of both *Pl1-CML52a* and *Pl1-Rhoades* transcription units are identical, therefore primers could not be made to differentiate the transcript levels from these two alleles. However, primers were created over a polymorphism in *Pl1-CML52b*, so the primer did not amplify transcripts from *Pl1-CML52b.* (Figure 8) PCR was used to confirm that the primers amplified each gene from cDNA and only one, strong band was seen. The PCR product was visualized using gel electrophoresis with an ethidium bromide stained agarose gel (1% agarose, 150 V, 25 min) (Figure 8).

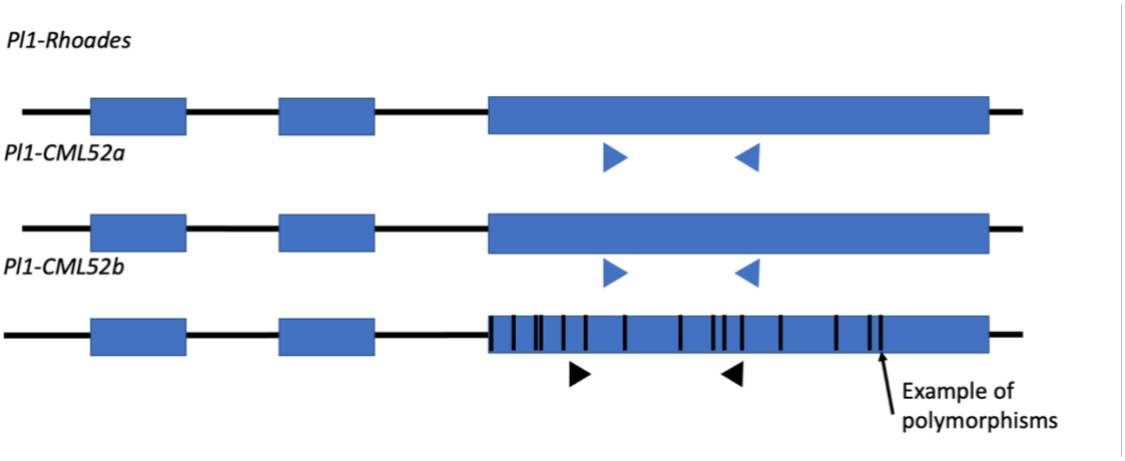

Figure 8 | Gene-specific primers.
Arrows indicate the location of the primers used in qRT-PCR analyses. Closed blue primers were used to amplify in *Pl1-Rh* and *Pl1-CML52a* and the reverse primer was created over a polymorphic region in *Pl1-CML52b.* Closed black primers indicate where *Pl1-CML52b* specific primers were created, the forward primer is in a polymorphic region.

The of *Pl1-CML52a* and *Pl1-CML52b* transcript levels were compared to one another using quantitative reverse transcription PCR (qRT-PCR). RNA was first isolated from 2-3 anthers using TRIzol reagent with three biological replicates for *Pl1-CML52* and *Pl′ / Pl1-CML52*, then reverse transcribed to cDNA using Protoscript II (NEB) and oligo(dT) primers. Real time qRT-PCR assay was then set up, using *gapdh (glyceraldehyde 3-phosphate dehydrogenase)* as a control. *gapdh* is a housekeeping gene in the maize genome, whose expression should be constant and is what the biological replicates were normalized to. A real time qRT-PCR with Luna Universal qPCR Master Mix (NEB) was then set up on an Eppendorf Mastercycler EP Gradient S thermocycler for the three biological replicates of *Pl1-CML52* as well as three technical replicates of each biological replicate and negative controls for both *Pl1-CML52a* and *Pl1-CML52b* primers. (Table S1)

Cytosine methylation profiling:

Specific primer design of each repeat sequence was attempted, however was unsuccessful in amplifying one specific region of the genome. Instead, only one primer set was used to determine the methylation of the three repeat sequences, and they could not be differentiated. (Table S2)

gDNA was extracted from day 8 seedlings. Enzymatic Methyl-Seq (NEBNext) first converts 5-methylcytosine (5mC) into 5-carboxycytosine (5caC) using TET2, which protects the 5mC from deamination. Then cytosines are enzymatically converted into uracils using APOBEC (apolipoprotein B mRNA editing enzyme, catalytic polypeptide). The uracils were then converted to thymines after PCR, therefore the only cytosines left are representative of methylated cytosines. A general primer was created over the USR region of the repeat candidate

enhancer and PCR was used to amplify each repeat. The sequence was then ligated with pGEM-T easy vector and transformed into NEB 5-alpha competent *E. coli* cells. A white and blue assay was preformed, plasmid DNA was extracted, and Sanger sequencing was utilized. (Figure S2)

Results

Pl1- CML52a and *Pl1-CML52b* were distinguished from *Pl1-Rhoades.*

Extracted RNA was used to confirm that mRNA specific primers only amplify one gene transcript, allowing to distinguish the different transcript levels of *Pl1-CML52b* from *Pl1-CML52a*. The gel electrophoresis and EtBr-staining of RT-PCR amplicons showed that *Pl1-CML52b* specific primers strongly amplify one band of the correct size. The bands on the gel confirm that primers created over the exact homolog sequence between *Pl1-CML52a,* and *Pl-Rhoades* was amplified. Through these mRNA specific primers, mRNA transcript levels could be distinguished between *Pl1-CML52b* from *Pl1- CML52a* and *Pl1-Rhoades.* (Figure 9)

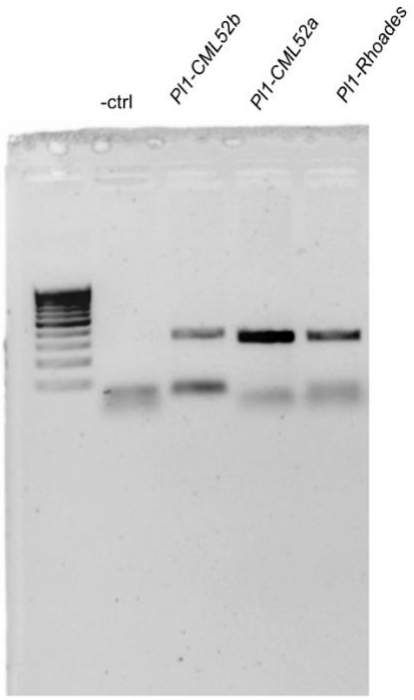

Figure 9 | Gel image of gene-specific primers.
Lane 1 contains a 100 bp (NEB Biolabs) ladder. Lane 2 holds the negative control. RT-PCR gel illustrating *Pl1-CML52b* primer amplification was confirmed in Lane 3. Lanes 4-5 indicate primer amplification that is shared

Pl1-CML52a transcript levels are significantly higher than *Pl1-CML52b.*

To determine if the *pl1* gene expression in *Pl1-CML52* is impacted by the presence of *Pl′*, the transcript levels were compared between *Pl1-CML52* and *Pl′* / *Pl1-CML52*. qRT-PCR was used; first the *gapdh* control was used to set a normalization of transcript levels in the *Pl1-CML52* and *Pl′* / *Pl1-CML52* biological replicates. Then *Pl1-CML52a* levels were compared in the biological replicates, each having three technical replicates as well, using *Pl1-CML52a* specific primers. *Pl1-CML52a* was used because it is expressed at a much higher level than *Pl1-CML52b* (Figure 6), giving a more accurate reading. First, the mRNA transcript levels of *Pl1-CML52b* and *Pl1-CML52a* were measured to determine if both were important to the paramutation expression in CML52 haplotype. Relevant comparisons were made, and Figure 10

shows the mRNA transcript levels of the two genes. It can be determined that *Pl1-CML52a* transcript levels are significantly higher than *Pl1-CML52b*, with a p-value value of 0.0001.

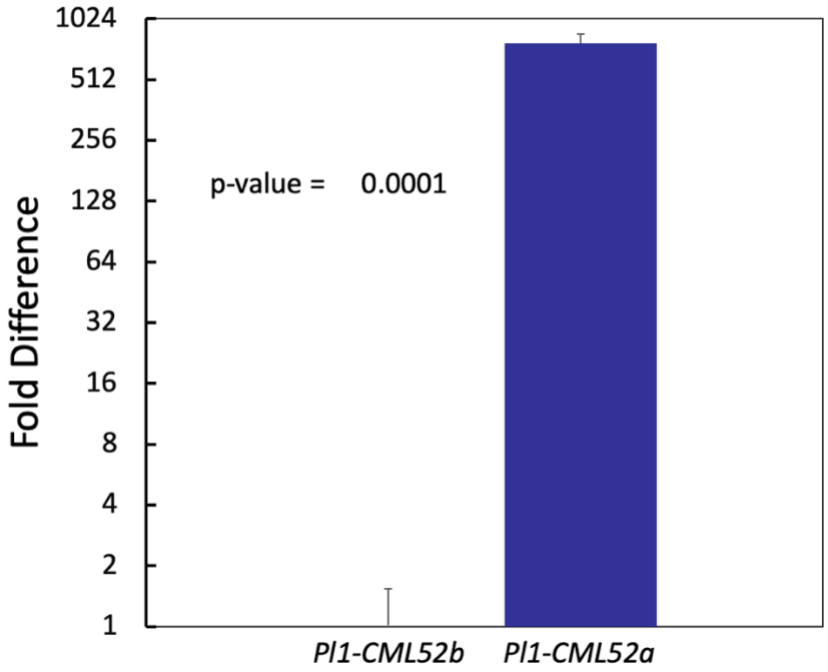

Figure 10 | Pl1-CML52b and Pl1-CMl52a transcript level fold difference. qRT-PCR comparison between *Pl1-CML52a* and *Pl1-CML52b* with T-test and standard error of the mean (SEM) bars for 3 biological replicates. Y-axis is in log2 fold change.

Pl' may reduce *Pl1-CML52a* transcript levels.

Pl1-CML52a was used to compare the transcript level of *Pl1- CML52* and *Pl1- CML52* in the presence of *Pl′* (*Pl′* / *Pl1- CML52*). Relevant comparisons were made and while there is about a 6-fold difference and the standard error of the mean (SEM) bars do not overlap, the T-test value is not significant. The values for *Pl′ / Pl1- CML52* were quite variable which led to the p-value to be greater than 0.05. (Figure 11) While the difference between the samples is not significant, it does suggest an overall repressed transcription level in *Pl′ / Pl1- CML52* material.

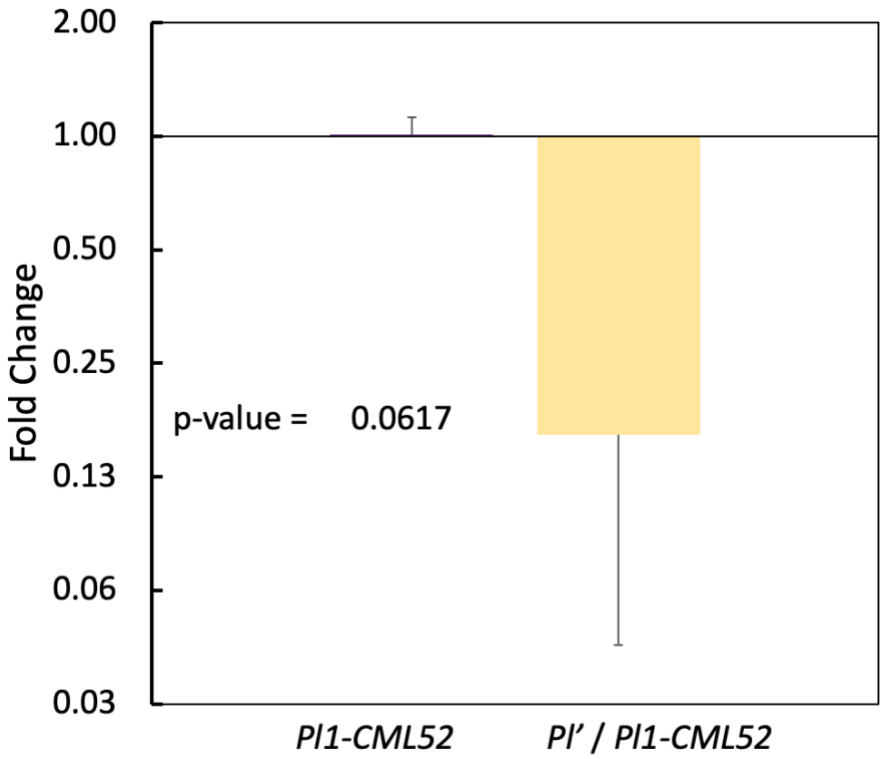

Figure 11 | Fold Change between Pl1-CML52 and *Pl' / Pl1- CML52.* qRT-PCR of *Pl1-CML52* levels in CML52 and *Pl' / Pl1- CML52* with standard error of the mean (SEM) bars for 3 biological replicates. Y-axis is in log2 fold change.

Pl1-CML52' cytosine methylation patterns are consistent with epigenetic repression.

To determine the methylation patterns of the three repeat candidate enhancers, Sanger sequencing was utilized, and the results are shown in a dotplot. When *Pl1-CML52* is paramutagenic (*Pl1-CML52'*), there is more cytosine methylation produced at the USR. In the *Pl1-CML52* samples, no methylation is seen at any methylation sites in any of the samples. However, in *Pl1-CML52'* there is an increased amount of CG and CHG methylation.

Methylation is also relatively consistent across all samples. (Figure 12) H is an IUPAC degenerate base symbol which represents A, C, and T as possible nucleic acids in that position.

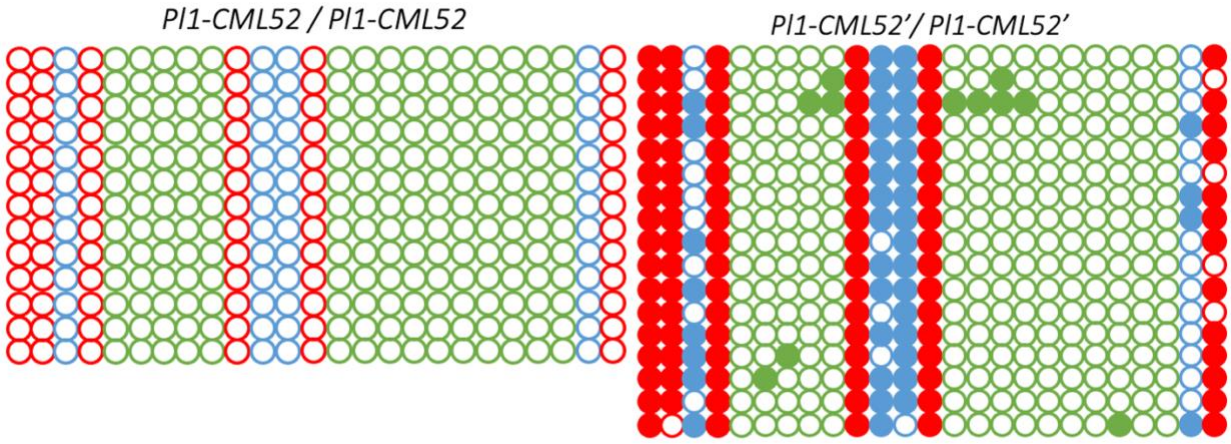

Figure 12 | *Pl1-CML52* and *Pl1-CML52′* cytosine methylation patterns. Dotplot of methylation patterns in unique sub-repeat region of candidate enhancers. Each column represents a different methylation site in the unique sub-repeat region of the candidate enhancer. Each row represents a different sample. Empty circles denote unmethylated cytosines and filled circles denote methylated cytosines. Red circles represent CG methylation, blue is CHG methylation, and green is CHH methylation.

Discussion

Unique primer design over mRNA sequence specific regions allowed for *Pl1-CML52b* to be distinguished from *Pl1-CML52a* and *Pl1-Rh* in a qRT-PCR assay. The significantly higher level of *Pl1-CML52a* mRNA transcript levels shows that only one gene is likely responsible for paramutation, similar to the B73 strain with *Pl1-Rh.* This result therefore influenced the next experiment and is the reason *Pl1-CML52a* levels were compared between *Pl1- CML52* and *Pl1- CML52* in the presence of *Pl′* (*Pl′* / *Pl1- CML52*). The difference between *Pl1- CML52* and *Pl′* /

Pl1- CML52 was not significantly different, even though the SEM bars do not overlap, and the graph shows a fold change of 6 difference. This lack of significance is due to the variability in results from the *Pl′* / *Pl1- CML52*. Possible future experiments could increase the number of samples from three to possibly decrease the variability present. It is acknowledged that *Pl1-CML52* has two copies of the gene *Pl1-CML52a* which could impact mRNA levels. Also, *Pl′* / *Pl1- CML52* contains a copy of the *Pl1-Rhoades* allele, that also could have been amplified. It is known from past experiments that *Pl1-Rhoades* transcript levels decrease when in the *Pl′* state. Even though *Pl1-CML52* has two copies of *Pl1-CML52a* and *Pl′* / *Pl1- CML52* has one copy of *Pl1-CML52a* and *Pl1-Rh′,* a 6-fold difference is still seen, suggesting that *Pl1- CML52* transcript level is decreased with the presence of *Pl′.* However, the project moved forward, and methylation patterns of enhancer regions were examined.

The attempt to create primers that differentiate the candidate enhancer repeats was not successful. The results were varying, and due to the translocation nature of the maize genome it was difficult to isolate a single band. Instead, the same primer set was utilized for all three repeats in the unique sub-repeat region. Whichever repeat was amplified and inserted into the plasmid would be random and the samples would be a mixture of the three repeats. Therefore, if all the repeats are methylated the same, we would expect to see one methylation pattern in all the samples. However, if the methylation patterns of the three repeats were different, we would expect to see varying results. In the *Pl1-CML52* dotplot results, we see no methylation in any sample. This is consistent with the idea that *Pl1-CML52* (specifically *Pl1-CML52a*) is transcribed, and the enhancers are in an active state. However, when *Pl1-CML52* is in a paramutant state, methylation increases and *Pl1-CML52* is no longer transcribed. Methylation at the putative enhancer happens at the CG and CHG positions; this is because CHH methylation in

maize does not have a maintenance methyltransferase to keep the methylation pattern maintained. (19) There are also no differences in methylation patterns between the fragments from *Pl1-CML52′* that I sequenced; therefore, all the repeat candidate enhancers are methylated the same. weaker paramutagenic but we see the same methylation profile across all the repeats and the number of repeats and not clear the connection between gene silencing.

Conclusion

Pl1-CML52 is an interesting haplotype of the *pl1* locus because it is the only known haplotype in maize, other than Rhoades, in maize that expresses paramutation. Also, all 24 founder sequences were sequenced and *Pl1-CML52* is the only haplotype that contains two copies of the *pl1* locus. However, based on the mRNA transcript levels of *Pl1-CML52a* and *Pl1-CML52b*, it is likely only *Pl1-CML52a* is expressed at a high enough level to be the primary driver to have the reduced pigment and is consistent with theWhile a statistically significant difference between *Pl1-CML52a* mRNA transcript levels in the presence of *Pl'* could not be concluded, a decrease in *Pl1-CML52a* mRNA pattern is present. This in consistent with other *pl1* locus repression findings when paramutation is present. *Pl1-CML52′* would make for a better test line because it has the same number of *Pl1-CML52a* alleles and no *Pl1-Rhoades* alleles. Also with more biological replicate samples, it could be determined if this decrease in transcript levels is significant or not.

Pl1-CML52 methylation patterns are intriguing because of the location and quantity of the candidate enhancers. Compared to *Pl1-Rhoades*, *Pl1-CML52* does not contain the penta-repeat enhancer region, however all the enhancer repeats are methylated in both haplotypes

which exhibit paramutation. Therefore, a penta-repeat or contiguous enhancer region are not necessary to methylate all the enhancers. There also isn't a specific number of methylated repeats necessary for paramutation. More comparisons between these two haplotypes methylation and the others which don't exhibit paramutation, would be necessary to determine if an increase in methylation patterns is a determining factor of paramutation.

Through these *Pl1-CML52* haplotype experiments, more information about the mRNA transcript levels of the *pl1* locus genes and methylation patterns of paramutation were discovered. The presence of candidate enhancer methylation in all repeats is seemingly important to the repression of the *pl1* locus and paramutation. The continuation of studying the methylation process is crucial to our understanding of paramutation and transgenerational gene regulation.

References

1. Natalie C. Deans, Joy-El R.B. Talbot, Mowei Li, Cristian Sáez-González, Iris Hövel, Chris Watkins, Darren Heavens, Maike Stam and Jay B. Hollick. Paramutation at the maize *pl1* locus is associated with RNA polymerase IV action at distal tandem repeats. (In submission).
2. Karl F. Erhard, Jr, Susan E. Parkinson, Stephen M. Gross, Joy-El R. Barbour, Jana P. Lim, Jay B. Hollick, Maize RNA Polymerase IV Defines *trans*-Generational Epigenetic

Variation, *The Plant Cell*, Volume 25, Issue 3, March 2013, Pages 808–819, https://doi.org/10.1105/tpc.112.107680

3. Brink, R. A. Paramutation at the R locus in maize. *Cold Spring Harb Symp Quant Biol*. **23**, 379-91 (1958).

4. Hollick, J. B., Patterson, G., Coe, E. H. Jr., Cone, K. C. & Chandler, V. L. Allelic interactions heritably alter the activity of a metastable maize *pl* allele. *Genetics* **141**, 709–719 (1995).

5. Dorweiler, Jane & Carey, Charles & Kubo, Kenneth & Hollick, Jay & Kermicle, Jerry & Chandler, Vicki. (2000). mediator of paramutation1 Is Required for Establishment and Maintenance of Paramutation at Multiple Maize Loci. The Plant cell. 12. 2101-18. 10.2307/3871108.

6. Hale CJ, Stonaker JL, Gross SM, Hollick JB (2007) A Novel Snf2 Protein Maintains *trans*-Generational Regulatory States Established by Paramutation in Maize.

7. Hollick, J. B. Paramutation and related phenomena in diverse species. *Nat. Rev. Genet.* **18**, 5-23 (2017).

8. Alleman, M. *et al*. An RNA-dependent RNA polymerase is required for paramutation in maize. Nature **442**, 295–298 (2006).

9. Nobuta, K. *et al*. Distinct size distribution of endogenous siRNAs in maize: evidence from the deep sequencing in the *mop1-1* mutant. *Proc. Natl Acad. Sci. USA* **105**, 14958–14963 (2008).

10. Erhard, K. J. *et al*. RNA polymerase IV functions in paramutation in *Zea mays*. *Science* **323**, 1201–1205 (2009).

11. Hale, C. J., Stonaker, J. L., Gross, S. M., & Hollick J. B. A novel Snf2 protein maintains *trans*-generational regulatory states established by paramutation in maize. *PLoS Biology* **5**, e275 (2007).

12. Sidorenko, L. *et al.* A dominant mutation in *mediator of paramutation2*, one of three second-largest subunits of a plant-specific RNA polymerase, disrupts multiple siRNA silencing processes. *PLoS Genet.* **5**, e1000725 (2009).

13. Stonaker, J. L., Lim, J. P., Erhard, K. F. Jr & Hollick, J. B. Diversity of Pol IV functions is defined by mutations at the maize *rmr7* locus. *PLoS Genet.* **5**, e1000706 (2009).

14. Barbour, J. R. *et al*. required to maintain repression2 is a novel protein that facilitates locus-specific paramutation in maize. *Plant Cell* **24**, 1761–1775 (2012).

15. Narain, A. S. *et al*. A dicer-like3 protein is essential for paramutations in *Zea mays*. Unpublished.

16. Talbot, J. R. B. (2015) Maize RNA polymerase IV complexes and their control of gene function. UC Berkeley Dissertation.

17. Jiao, Y. *et al*. Improved maize reference genome with single-molecule technologies. *Nature*. **546**, 524-527 (2017).

18. Hufford, M. B. *et al*. De novo assembly, annotation, and comparative analysis of 26 diverse maize genomes. Science. **373**, 655-662 (2021).

19. Natalie C. Deans & Zoe E. Reinstein, Unpublished

20. Zemach, A., Kim, M. Y., Hsieh, P.-H., Coleman-Derr, D., Eshed-Williams, L., Thao, K., Harmer, S. L., & Zilberman, D. (2013). The Arabidopsis nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. *Cell*, *153*(1), 193–205. https://doi.org/10.1016/j.cell.2013.02.033

Supplemental Materials

Primer	Sequence
pl1B_spec_F	**5'-GGGGAGCCGCGATATGAC-3'**
nonspec_2R	**5'-CTCACGTCGTCCATCCAG-3'**
nonspec_3F	**5'-GGCTCTTCTTCCACCGGC-3'**
pl1A_spec_2R	**5'-CGTGTACTTGTTGTCTTGTCTACA-3'**
gapdh_F	**5'-CCTGCTTCTCATGGATGGTT-3'**
gapdh_R	**5'-TGGTAGCAGGAAGGGAAACA-3'**

Table S1: mRNA expression analyses primers.
Primers were used to determine the presence of the exon extension region and utilized in qRT-PCT to determine mRNA transcript levels. pl1B_spec_F and nonspec_2R were used to amplify *Pl1-CML52b.* nonspec_3F and pl1A_spec_2R were used as a control to amplify *Pl1-CML52a* and *Pl1-Rh.* gapdh_F and gapdh_R were used as a control in qRT-PCR assay and amplified *gapdh*, a housekeeping gene.

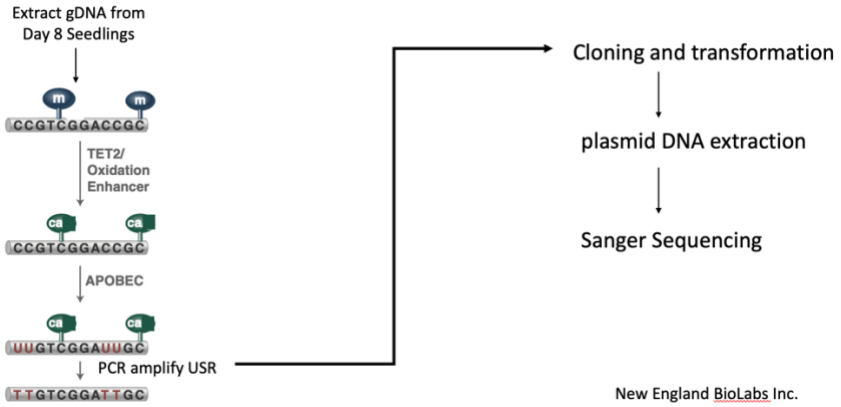

Figure S2 | NEBNext Enzymatic Methyl-Seq workflow
Blue circle with "m" indicated 5-methylcytosine. Oxidation Enhancer is utilized protect 5-methylcytosine from deamination by converting it to 5-carboxycytosine. APOBEC deaminates all cytosines to uracils.

Primer	Sequence
unqmeF1	**5'-AGAATGATGTTTGTTATATYGGTT -3'**
unqmeR1	**5'-CTCCAACATRTTCCAAAARCAACAC -3'**

Table S2: Cytosine methylation primers.
Primers were used to amplify unique sub-repeat regions in candidate enhancers. unqmeR1 was also used in Sanger Sequencing.