Spring 2020

Identifying RPA-Dependent Epigenetic modifications influencing Arabidopsis thaliana flowering time

Emily S. Berry
University of New Hampshire

Identifying RPA-Dependent Epigenetic modifications influencing *Arabidopsis thaliana* flowering time

Emily Berry

Mentor: Dr. Kevin Culligan

University of New Hampshire

ABSTRACT

Epigenetics is the study of heritable changes that do not directly alter the DNA sequence. Many of these inheritance mechanisms are caused by the addition of a methyl group to cytosine nucleotides in DNA that prevent gene expression. The model organism *Arabidopsis thaliana* (abbreviated *A. thaliana*) is commonly used in genetic experiments, and its usefulness is extended to research in epigenetics. Our current research (Culligan lab), in part, focuses on genes involved in DNA damage and repair, such as *brca2* and the *rpa1a/b/c/d/e* genes. Previous genetic analyses suggests that the *rpa1c/e* double-mutant typically expresses an early flowering phenotype despite no additional changes made to the nucleotide sequence or environment compared to other strains. One possible explanation to this phenomenon is the methylation of certain regions of DNA associated with flowering. The goal of this experiment is to identify epigenetic differences between the *A. thaliana* Col-O wild type and *rpa1c, rpa1e* and *rpa1c/e* mutants. From this research, different DNA methylation patterns could be identified in these *rpa1* mutants, and the results could help researchers better understand how DNA methylation and histone modifications affect the development of an organism. Early flowering can also be seen in agricultural settings and can severely affect crop yields. Comparisons of growth between mutants showed significant differences between Col-O and *rpa1e / rpa1c/e* rosette leaf count and diameter. To test whether methylation patterns play a role in the early flowering phenotype, we employed an ELISA-based assay, which uses a standard curve of positive and negative controls, to determine methylation pattern differences in the WT and mutant lines. The standard curve for the ELISA assay was not adequate to use to compare the samples, possibly due to error while performing the assay, or issues with the controls.

INTRODUCTION

Modern epigenetics is defined as the study of heritable changes in gene function not caused by the DNA sequence. In the past, epigenetics was defined as how zygotes transformed into complex multicellular organisms, and with the knowledge we have today it has become a relatively new area of research and applications in genetics. Current research in the field is focusing on identifying epigenetic markers such as DNA methylation, histone acetylation, phosphorylation and ubiquitination, among others (Felensfeld, 2014).

The epigenome of a species is much more diverse and variable than its genome. Since all cells of an individual have the same genome, the epigenetic modifications that vary between cell types help to determine cell structure and function. In 2016, researchers compiled a database of over 1000 *A. thaliana* epigenomes from various parts of the world, providing researchers with a more detailed description of the *Arabidopsis* epigenome (Lang, 2016, Kawakatsu, 2016). In the human epigenome, researchers analyzed 150 billion sequence reads in 111 cell types and found epigenetic variation in various cell types for specific human traits and diseases (Roadmap Epigenetics Consortium, 2015).

Flowering time plays a crucial role in the development and reproductive success of flowering plants. Some epigenetic processes are known to influence and regulate flowering time. For example, the methylation of floral transcription factor FWA attracts small interfering RNA (siRNA), which then represses the expression of FWA. This methylation pattern helps the plant respond to changes in the environment to enhance its survival by resulting in late flowering (Yaish, 2011; Fujimoto, 2011). Flowering time can also depend on biological factors. The brassinosteroid-insensitive 1 (BRI1) protein, when overexpressed, can lead to early flowering in *A. thaliana* (Singh, 2016). Conversely, the absence of BRI1 results in plants with a dwarf-like

phenotype and the inability to synthesize brassinosteroids (BRs). In other plants, the absence of BRI1 suggests roles in plant architecture and yield (Singh, 2016). DNA methylation can also act as a response mechanism to stressors, and the disruption of methylation patterns can change the response of the plant against biotic stressors such as pathogens (Dowen, 2012).

There are many model organisms that could be used to study epigenetics, but for this experiment, *A. thaliana* is most suitable because of its short generation time, simple care routine, and fully-sequenced genome. This would allow the use of multiple generations or replicates in a short experimental timeframe (Katagiri, F., Thilmony, R., & He, S. Y., 2002). Since the genome of *A. thaliana* is fully sequenced, it is also easy to identify the genes in question without worrying about other, possibly unknown genes that may influence the results.

The RPA large-subunit (*rpa1*) protein in *A. thaliana* has been well studied. This protein is translated from 5 genes; *rpa1a*, *rpa1b, rpa1c*, *rpa1d* and *rpa1e*. These genes code for the 70 kDa subunit of the three subunits (RPA70, RPA32 and RPA14) that comprise the Replication Protein A holoenzyme. The RPA protein binds to single-stranded DNA (ssDNA) and preserves it from degradation during replication (Aklilu, 2013). If one or more of these genes are mutated, homologous recombination cannot be used to repair DNA. Double mutants have been found to generate phenotypes different from single mutants. The *rpa1a/c* double mutation results in sterile individuals, while the *rpa1c/e* double mutation results in hypersensitivity to radiation and other DNA damaging agents (Aklilu, 2013).

The RPA protein is also found in humans, with identical functions as the ortholog found in *A. thaliana.* RPA is involved largely in DNA repair, and also plays a role in recombination and replication (Zou, 2006). In humans, the RPA protein is also coded by three RPA subunits found in *A. thaliana* (RPA70, RPA32 and RPA14). This, in addition to the ubiquitous nature of

RPA, suggests that this protein existed before the divergence of many species, and that it has seldom changed since then (Zou, 2006).

In addition to mutations, it is also possible that DNA methylation or histone modifications could alter the expression of these genes. One well-known silencing mechanism of DNA is methylation of the C^5 position of cytosine residues, which is highly mutagenic. This methylation is typically found in the promoter regions of DNA and prevents affected genes from being transcribed and expressed in the cell (Egger, et al., 2004). These methylated regions can be analyzed by bisulfite sequencing or an ELISA-based assay.

The goal of this research is to examine any potential epigenetic factors influencing the expression of two of these five *rpa* genes (*rpa1c* and *rpa1e*) and determine how they may influence flowering time. The findings from this research project could enhance our understanding of how gene expression is regulated, and potentially lead to other, similar studies focusing on other genes found in the *A. thaliana* genome.

STATEMENT OF HYPOTHESIS

This research project addressed the following questions;

1. *What is the average difference in flowering time (in days) and/or growth rate between the Col-O control and rpa1c, rpa1e and rpa1c/e mutants?*
2. *Are there any significant differences in methylation between Col-O and the various RPA mutants? Are these changes specific to certain mutants?*

METHODS

Growth and maintenance of plants. *rpa1c, rpa1e, rpa1c/e*, and wild-type (Col-O) *A. Thaliana* seeds were sown on Phyto-agar in a sterile environment. Seeds were cleaned with bleach, rinsed

with sterile water, and suspended in 1% agarose for sowing. For each mutant in a replicate, approximately 20 seeds were sown, 12 planted, and 10 used for the experiment. Seeds were stratified at 4°C for 72 hours. The seedlings were moved to a growth chamber until the roots reached a length of about 2 cm (5-7 days). Seedlings were then planted in separate containers and carefully monitored to ensure no contact between other plants in the greenhouse. Plants were grown for 31 days. Rosette diameter, number of rosette leaves, length of stem and number of flowers and flower buds were recorded every seven days starting at day 1. Replicates were done with new, naïve seeds, not from seeds from the previous replicate plants.

DNA Extraction and Genetic Analysis. Immediately after the final data collection on day 31, two rosette leaves and two 1cm sections from the stem were cut from each plant and frozen in liquid nitrogen for future analysis the same day. DNA was extracted using isopropanol. Mutants were confirmed using qPCR. Once the genomic DNA was extracted and the mutants confirmed, the initial samples were used for an enzyme-linked immunosorbent assay (ELISA), using the MethylFlash Global DNA Methylation (5-mc) ELISA Easy kit (colorimetric) by EpiGentek. The assay was used to get a general sense of methylation areas in each mutant. This assay typically takes about 2 hours to complete according to the manufacturer. Depending on the outcome of the ELISA, bisulfite pyrosequencing would then be run on new samples to obtain more detailed information about methylation locations along the nucleotide sequence. Bisulfite conversion changes any unmethylated Cytosine to Uracil, so that when the sequencing readout is being analyzed, all ‘C’s in the data will be identified as methylated.

Data Analysis using Microsoft Excel, JMP and R. Results from this experiment were obtained by sampling the plant tissue and recording the flowering time. Data from plant growth rates, agarose gels and ELISA assays were used to analyze and identify any epigenetic markers

in the genes of interest. Microsoft Excel was used to record and sort all raw data. Data from plant growth was then transferred to JMP Pro 13 to compare plant size to other mutants (Figures 1 and 2). Data from the ELISA assay was graphed using Microsoft Excel.

RESULTS

Differences in rosette diameter and rosette leaf count between strains

Plants were measured every seven days, starting with the day they were removed from the phyto agar plates and planted in soil (day 1). Initial growth measurements included rosette leaf count, rosette diameter and when the plants bolted stem length, bud count, and flower count were also recorded. The results from the replicates were similar and showed no significant differences between the replicate strains. The data from Figure 1 illustrates the number of rosette leaves counted each day. Rosette leaves that were under 5mm in length were excluded, as well as leaves that were dead or had fallen off of the plant. Dead leaves were not common enough to affect the overall data. The data shows a steady increase in rosette leaves over time, as expected. From days 7-21, the strains have relatively similar leaf counts, with no significant differences between any of the strains. Data from days 28-31 shows a significant increase in rosette leaf abundance for the *rpa1e* and *rpa1c/e* mutants, with the average leaf count for *rpa1e* being 84 on day 28 and 106.4 on day 31, and *rpa1c/e* being 81.1 on day 28 and 90.7 on day 31. The leaf counts for Col-O and *rpa1c* were not significantly different from one another on day 28, but *rpa1c* mutants had a slightly lower average leaf count on day 31 (12.7) compared to Col-O (16.7).

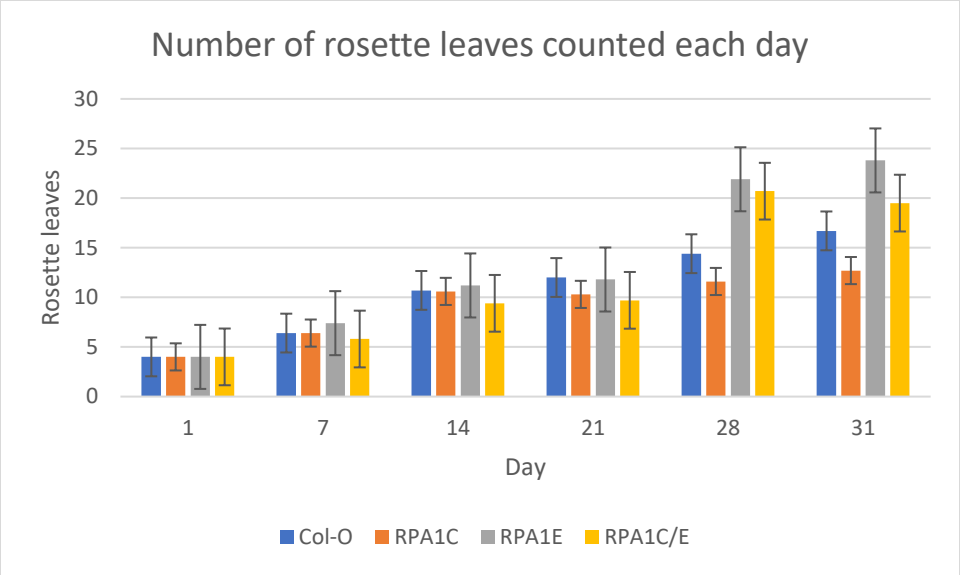

Figure 1. Rosette leaves were counted every seven days. There appears to be a significant difference in number of leaves between Col-O and *rpa1e*. n=10 per strain for each replicate.

Rosette diameter was measured in millimeters and taken the same days as the rosette leaf counts. The diameter was measured at the widest point of the rosette, typically from the tips of the two longest, parallel rosette leaves. The data collected from the two replicates suggests a significant difference in rosette diameter between the Col-O and *rpa1c* from the *rpa1e* and *rpa1c/e* mutants at day 31. In table 2 below, the average rosette diameter for *rpa1e* and *rpa1c/e* appears to increase at a faster rate than Col-O and *rpa1c* between days 21 and 28, and after day 31. The Col-O and *rpa1c* strains did not change much in diameter from days 28 to 31, compared to *rpa1e* and *rpa1c/e.*

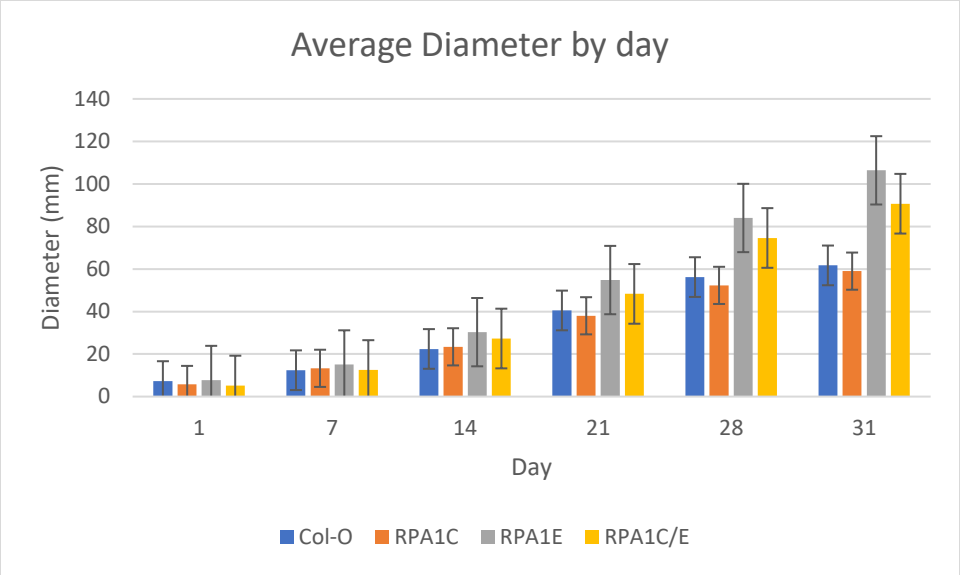

Figure 2. Average diameter of rosettes (mm). Diameters were taken at the widest point of the rosette. n=10 per strain for each replicate.

No significant differences in stem growth or flower formation

Additional data was taken on stem length as well as bud and flower counts for the second replicate. Data was taken on the same days as rosette data, and measurements were recorded as each plant bolted. Samples size varies by day, depending on how many plants bolted and had stems taller than 1cm. Although there was a slight increase in stem length observed in *rpa1e* and *rpa1c/e*, there were no significant differences in stem length between any of the strains, as shown in Figure 3.

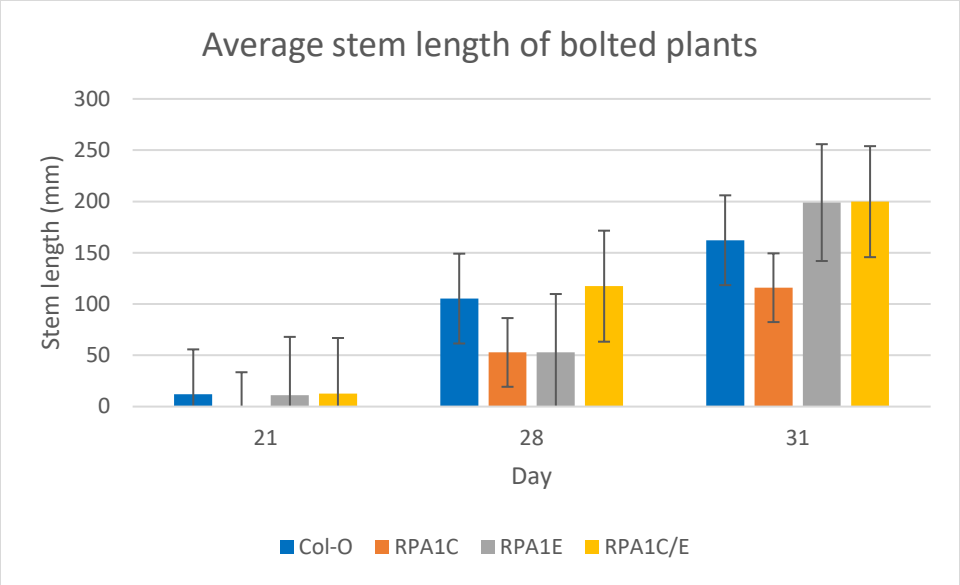

Figure 3. Stem length (in mm) of bolted plants. Samples size varies by day and by strain, as data was collected as plants bolted.

The sample sizes varied by day when taking stem data. Plants that had stems greater than 1cm were counted in the data. Figure 4 below summarizes the sample size for each strain by each day. Although the sample sizes were relatively equal for each day, there were no bolted *rpa1c* plants on day 21. One Col-O plant did not bolt during the data collection period, despite having no significant difference in rosette growth compared to other Col-O individuals in the replicate. From the stem length data, it was determined that there was not as much significance in growth rate between the strains as the rosettes. In addition to stem length, flower and flower bud counts were also recorded. Flower buds appeared a few days after bolting and are recorded in Figure 5.

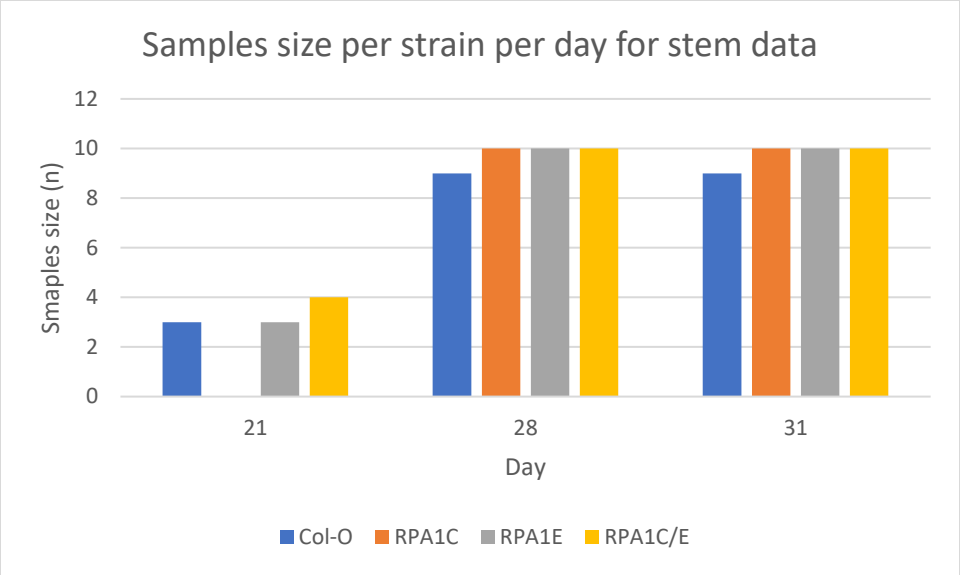

Figure 4. Samples size of each strain for stem length data. Data was taken from the second replicate of plants.

The number of flowers and flower buds was also recorded as they were observed and is described in Figure 5. There were no significant differences on days 21 and 30 between any of the strains, however there was some variation between Col-O and *rpa1c* on day 28. The Col-O day 28 plants had an average of 6.3 flowers/flower buds, while *rpa1e* had an average of 20.5.

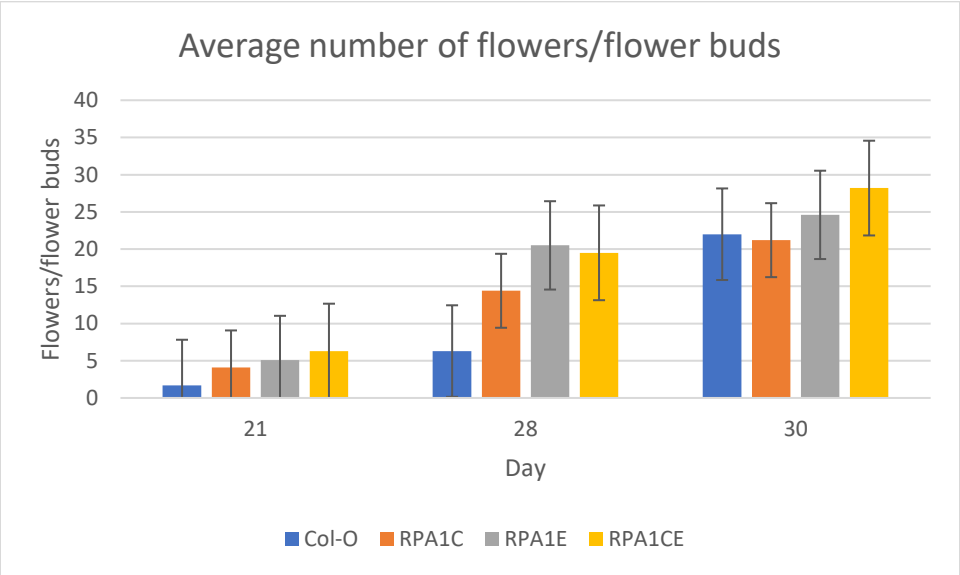

Figure 5. Flower and flower bud counts per strain. Sample size also varied by strain and by day.

ELISA standard curve

The ELISA assay was performed using the EpiGentek MethylFlash Global DNA methylation (5-mC) colorimetric kit (cat. Number P-1030). The assay was performed according to the specifications and protocol listed in the manual, and each item (controls or samples) were done in duplicate to obtain an average. Samples were diluted to 50ng/uL, and since the protocol recommended using no less than 2uL of DNA, a total of 100ng/uL was added to each well. Positive and negative controls were included with the kit for the purpose of generating a standard curve on which to analyze the samples. According to the manual, the standard curve should look similar to the curve illustrated in Figure 6.

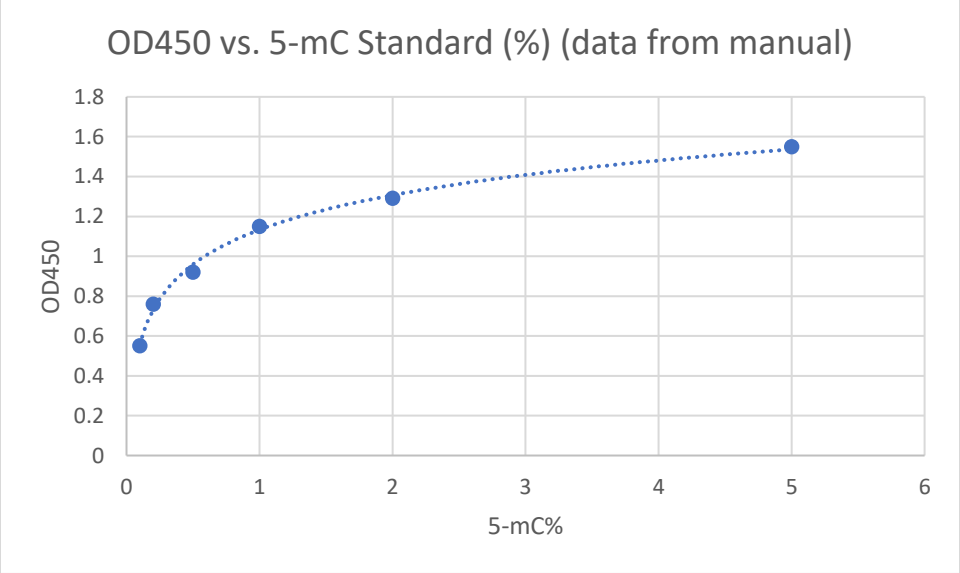

Figure 6. Expected standard curve, obtained using the sample data provided in the ELISA assay manual.

The actual curve generated from this experiment was very different from the expected curve. The data in Figure 7 shows the results obtained from this assay, which resemble a line

more than a curve, as indicated by the trendline. There is little to no fit between the points and the trendline, as indicated by the R^2 value of 0.0014. The data produced no standard curve as expected, and therefore the data was deemed unusable. There were enough reagents and controls left over to redo the curve, but similar results were obtained. Raw data is included in table 1 of the Appendix.

The ELISA kit can also be loosely interpreted by the naked eye prior to analysis. During one of the last incubation periods, heavy methylation (5% or greater) would result in dark blue wells, while little to no methylation would result in light blue or clear wells (0-0.1%), with a gradient in between. From visual observation of the first test in this experiment, the methylation appeared very randomized, as some of the replicates had a wide range of color. For example, the negative control, which was expected to have no methylation and therefore result in a clear well, had one well that was dark blue, and the other had roughly the same shade of blue as a majority of the other wells. The positive controls, which were expected to form a gradient from light to dark blue in color, looked similar to the negative control, with lots of variation between the replicates. This, along with the data obtained from the plate reader, confirms that something went wrong during the assay.

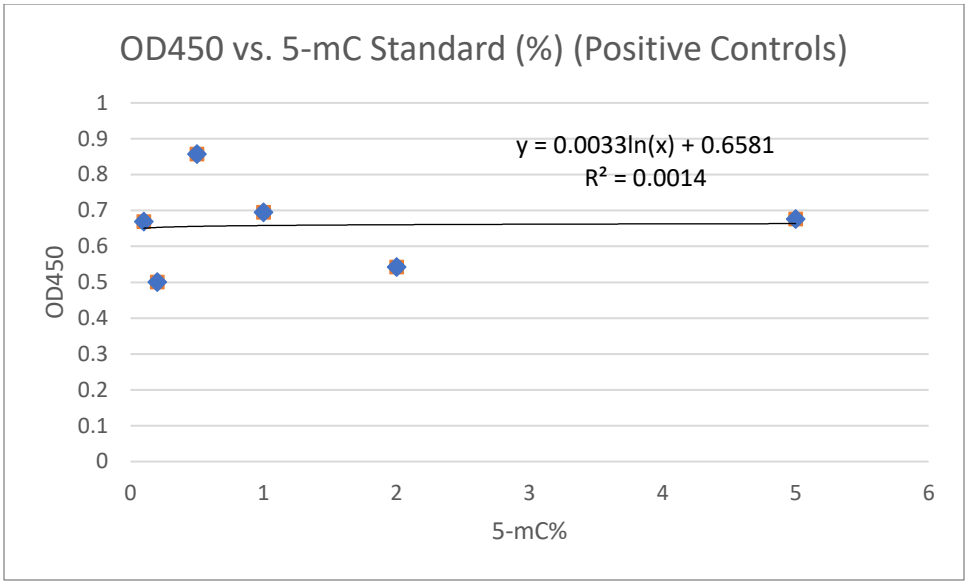

Figure 7. Standard curve obtained from the controls provided in the assay.

DISCUSSION

This research project addressed the following questions;

1. *What is the average difference in flowering time (in days) and/or growth rate between the Col-O control and rpa1c, rpa1e and rpa1c/e mutants?*
2. *Are there any significant differences in methylation between Col-O and the mutants?*
3. *Does the lack of an RPA function contribute to methylation patterns in A. thaliana?*

Average difference in flowering time between mutant strains

The plant growth data suggests a difference in rosette growth rate between the *rpa1e* and the *rpa1c/e* mutants compared to Col-O or *rpa1c*. There appeared to be differences specifically in rosette diameter and rosette leaf counts in the latter half of the 31-day measurement period. Other data, such as stem length and flower count, had no significant differences. Although a difference was observed between the average flower/flower bud counts between *rpa1e* and Col-

O, more replicates should be done with more plants to confirm this, as stem growth was much more varied than rosette growth.

The change in rosette growth suggests that the *rpa1e* region may be attributed with this change, as both the single mutant and double mutant shared this trait. The *rpa1c* mutants had no significant changes in either the rosette or the stem and flowers compared to Col-O.

Differences in methylation

Due to complications with the ELISA kit, no useable data was obtained to determine difference in methylation between any of the strains. From physical observations taken during the assay in addition to the data from the plate reader, there was a lot of variation in the duplicate well that should not be seen in an assay like the one in Figure 6. As mentioned previously, there may have been an error while performing the assay, or there may have been contamination in some of the wells. From these inferences, it is clear that the assay is inconclusive and would need to be repeated with new materials to determine what may have caused these anomalies.

Troubleshooting

No errors or setbacks were encountered during plant growth, data collection, or DNA extraction. The plants grew as expected, with no apparent contamination during sowing or incubation. All plants grew at relatively the same rate, as indicated by Figures 1 and 2. Bolting did seem to occur at different times, depending on the strain, as seen in Figure 4. DNA extraction from the rosette leaves was performed with no issues, and the samples concentrations were all relatively normal and within the parameters suggested for further study (DNA concentration >100ng/uL, 260/280>1.7).

The biggest setback during this experiment was from the ELISA assay. As the controls did not produce a workable standard curve, and there were not enough materials from the kit to

repeat both the control and samples, we were not able to redo the assay due to these reasons and time limitations. Possible explanations for this complication may be attributed to contamination, either at the bench or in the samples and controls themselves. If this test were to be redone, the results may improve if the assay was conducted in a sterile environment.

Future Directions

Additional replicates and data for plant growth should be performed and compared to the current findings. The growth data could also be extended to more than 31 days, to see if the Col-O and *rpa1c* mutants 'catch-up' to the size of the *rpa1e* and *rpa1c/e* mutants seen on days 28 and 31. More data should also be collected for stem length, as the first replicate had a small sample size, and only one replicate was used for data collection. Other forms of measurement could also be used, such as taking images of the plants each day and calculating the surface area of the rosette leaves.

In addition to additional growth data and samples, it would be advised to find either another ELISA-based kit, one that is more suited for plant genomic DNA, or another form of methylation analysis like bisulfite sequencing. While conducting this research, I was also assisting another lab member with histone extractions. This could be another potential route for analysis, but further investigation into histone extraction would be needed before it would be added to the experiment.

Conclusion and personal outcome

Although the first part of the research project suggested a change in rosette growth rate, the methylation analysis part of the experiment was inconclusive. It is likely that there was possible contamination or an error while preparing the assay, and that another assay would likely have to be run to determine what may have gone wrong. This research experience has helped me

learn how to troubleshoot and break down each component of an experiment when something goes wrong. Despite the second part of the research being inconclusive, I think this experience has helped me prepare for graduate school and future research. My short-term future goals are to improve my troubleshooting skills by learning more about other research techniques I have not used yet but might use during graduate school.

References Cited

Aklilu, B. (2013). Genetic analysis of the Replication Protein A large subunit family in Arabidopsis reveals unique and overlapping roles in DNA repair, meiosis and DNA replication. *Nucleic Acids Research, 42* (5).

Carey, N. (2012). *The Epigenetics Revolution: How Modern Biology is rewriting our understanding of Genetics, Disease, and Inheritance.* New York City: Columbia University Press.

Collotta, M., P.A.Bertazzi and V. Bollati (2013). Epigenetic and pesticides. *Toxicology, 307,* 35-41.

Craufurd, P.Q. and T.R. Wheeler. (2009). Climate Change and the flowering time of annual crops. *Journal of Experimental Biology, 60* (9), 2529-2539.

Dowen, R et al (2012). Widespread dynamic DNA methylation in response to biotic stress. *PNAS, 109 (32),* 2183-2191.

Egger, G. et al. (2004). Epigenetics in human disease and prospects for epigenetic therapy. *Nature Publishing Group, 429,* 457-463.

Felensfeld, G. (2014). A Brief History of Epigenetics. *Cold Spring Harbor Perspectives in Biology, 6* (1).

Fujimoto, R. et al. (2011). Epigenetic variation in the *FWA* gene within the genus *Arabidopsis*. *Society for experimental biology*. *66.* 831-843

Katagiri, F., Thilmony, R., & He, S. Y. (2002). The Arabidopsis Thaliana-Pseudomonas Syringae Interaction. *The Arabidopsis Book,* (1).

Kawakatsu, T., et al. (2016). Epigenomic Diversity in a Global Collection of *Arabidopsis thaliana* Accessions. *Cell, 166* (2), 492-505.

Lang, Z., et al. (2016) The 1001 Arabidopsis DNA Methylomes: An Important Resource for Studying Natural Genetic, Epigenetic, and Phenotypic Variation. *Trends in Plant Science, 21,* 906–908.

Mansuy, I. (2014, December 04). Retrieved February 15, 2018, from https://www.youtube.com/watch?v=L0v0uAloptQ&list=LL-5EUk4HrOSmS5zS3Td6lg&index=5&t=0s

Markus, C., A. Pecinka, R. Karan, J. Barney and A. Merotto Jr. (2017). Epigenetic Regulation – contribution to herbicide resistance in weeds? *Pest Management Science, 74* (2), 275-281.

Roadmap Epigenetics Consortium, et al. (2015). Integrative analysis of 111 reference human epigenomes. *Nature. 518.* 317-330.

Shutoh, Y. et al (2009). Low dose effects of dichlorodiphenyltrichloroethane (DDT) on gene transcription and DNA methylation in the hypothalamus of young male rats: Implication of hormesis like effects. *Journal of Toxicological Sciences, 34* (5), 469-482.

Singh, A, P. Breja, J. Khurana and P. Khurana (2016). Wheat Brassinosteroid-Insensitive1 (TaBRI1) Interacts with Members of TaSERK Gene Family and Cause Early Flowering and Seed Yield Enhancement in Arabidopsis. *PLOS one, 11* (6).

Yaish, M. et al. (2011). The role of epigenetic processes in controlling flowering time in plants exposed to stress. *Journal of Experimental Biology, 62 (11)*. 3727-3735.

Zou, Y, Y. Liu, X. Wu and S. M. Shell (2006). Functions of Human Replication Protein A (RPA): From DNA replication to DNA damage and stress responses. *Journal of Cell Physiology, 208 (2).* 267-273.

Appendix

Table 1: Data from the ELISA assay (left) and key for samples in each well (right).

<>	1	2	3		<>	1	2	3
A	1.0221	0.6273	0.5069		A	NC	1.0%PC	Sample 2
B	0.5807	0.7632	0.4549		B	NC	1.0%PC	Sample 2
C	0.8834	0.6123	0.5661		C	0.1%PC	2.0%PC	Sample 3
D	0.4549	0.4726	2.1943		D	0.1%PC	2.0%PC	Sample 3
E	0.4243	0.7635	0.645		E	0.2%PC	5.0%PC	Sample 4
F	0.5773	0.5881	0.5454		F	0.2%PC	5.0%PC	Sample 4
G	0.8924	0.4617	0.6408		G	0.5%PC	Sample 1	Sample 5
H	0.8224	0.7714	0.7754		H	0.5%PC	Sample 1	Sample 5