ALTERNATIVE SPLICING IS AFFECTED BY CHANGES IN DNA AND HISTONE METHYLATION

by

James Andrew Robertson

B.Sc.H., The University of Guelph, 2007

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

(Botany)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

April 2010

©James Andrew Robertson, 2010

ABSTRACT

Alternative splicing (AS) increases transcript and proteomic diversity of intron containing genes and has emerged as a pervasive property of eukaryotic genes. DNA methylation is a regulator of gene expression which is present in many eukaryotes. Recently, chromatin structure and histone modifications have been shown to regulate AS in humans (1), but it is unknown if this also occurs in plants. Here, I investigated the regulatory role that DNA and histone methylation have on AS in *Arabidopsis thaliana.* I utilized the SR protein gene family as a model system to examine changes in isoform abundance in response to chemical inhibition of DNA and histone methylation using the histone deacetylase inhibitor Tricostatin A (TSA) and the DNA demethylating agent 5-Azadeoxycytidine (Azad). RT-PCR revealed that 12 SR genes had isoforms that had altered abundance in response to TSA and Azad-C. Antagonistic effects were found when both drugs were applied since changes in splicing were only seen for 10 of the SR genes. AS patterns of the SR genes are known to change according to organ and developmental stage (2) and DNA methylation is also known to change over the lifespan of the plant (3). I investigated if organ type and developmental patterns of AS are disrupted in a DNA hypermethylation mutant. Splicing patterns of the SR genes displayed tissue and developmental specific changes in the DNA hypermethylation mutant. Computational analysis of three different DNA methylation mutants was performed using a publically available Illumina dataset (4). Widespread changes in AS was detected in each of the mutants across all types of AS. Changes in methylcytosine content within the coding region of the gene did not account for a large proportion of the novel AS events detected. Splice site sequence analysis of introns uniquely retained in each of the mutant genotypes uncovered sequence changes around the functionally important sequence elements. The results of my thesis indicate for the first time that AS in plants is regulated in part by changes in DNA methylation and histone modifications.

TABLE OF CONTENTS

ABSTRACT.......ii
TABLE OF CONTENTS.......iii
LIST OF TABLES.......v
LIST OF FIGURES.......vi
LIST OF ABBREVIATIONS.......vii
ACKNOWLEDGEMENTS.......viii
1. INTRODUCTION.......1
  1.1 RNA PROCESSING.......1
  1.2 FREQUENCY AND CATEGORIZATION OF AS EVENTS.......2
  1.3 REGULATION OF AS.......2
  1.4 HIGH-THROUGHPUT STUDIES OF AS.......4
  1.5 FUNCTIONS OF AS.......4
  1.6 SR PROTEINS.......5
  1.7 ESTABLISHMENT, MAINTENANCE AND FUNCTIONS OF DNA METHYLATION.......6
  1.8 HIGH-THROUGHPUT STUDIES OF DNA METHYLATION.......7
  1.9 CHROMATIN AND HISTONE MODIFICATIONS.......8
  1.10 MAJOR GOALS AND OBJECTIVES.......10
2. MATERIALS AND METHODS.......11
  2.1 PLANT GROWTH AND TREATMENTS.......11
  2.2 RNA PREPARATION.......11
  2.3 RT-PCR.......12
  2.4 TRANSCRIPT QUANTIFICATION.......12
  2.5 SPLICE VARIANT SEQUENCING AND ANALYSIS.......13
  2.6 ILLUMINA READ MAPPING.......13
  2.7 DETECTION OF AS EVENTS.......13
  2.8 SPLICE SITE SEQUENCE ANALYSIS.......14
3. RESULTS.......15
  3.1 EFFECTS OF CHEMICAL INHIBITION OF DNA AND HISTONE METHYLATION.......15
  3.2 AS OF SR GENES ACROSS DIFFERENT ORGANS AND DEVELOPMENTAL STAGES.......17
  3.3 COMPUTATIONAL ANALYSIS OF AS IN THREE DNA METHYLATION MUTANTS.......20
4. DISCUSSION.......23

4.1 AS PATTERNS OF THE SR GENES ARE AFFECTED BY DNA METHYLATION AND HISTONE HYPERACETYLATION 23
4.2 ORGAN AND DEVELOPMENTAL SPECIFIC EFFECTS ON SPLICING PATTERNS OF THE SR GENES 24
4.5 COMPUTATIONAL ANALYSIS OF AS EVENTS IN DNA METHYLATION MUTANTS 26
4.6 ASSOCIATION OF NOVEL AS WITH CIS CHANGES IN DNA METHYLATION 28
4.7 RETAINED INTRONS POSSESS ALTERED NUCLEOTIDE COMPOSITION AROUND SPLICE SITES 28
4.8 AS IN PLANTS IS AFFECTED BY CHANGES IN DNA METHYLATION AND HISTONE ACETYLATION 29
4.9 IMPLICATIONS OF EPIGENETIC REGULATION OF AS 30
5. CONCLUSIONS AND FUTURE DIRECTIONS 31
BIBLIOGRAPHY 52

LIST OF TABLES

Table 1 Gene specific primers used in this study……………………………………………………….40

Table 2 Alternative splicing ratios of drug treated seedlings……………………………………………...41

Table 3 Misregulated isoforms under each drug treatment……………………………………………..44

Table 4 Alternative splicing ratios in the *rdd* mutant across organ types and developmental stages……..45

Table 5 Misregulated isoforms in the *rdd* mutant across different organs and developmental stages……48

Table 6 Genomic locations of mapped Illumina reads…………………………………………………...49

Table 7 AS events detected in each of the genotypes…………………………………………...…50

Table 8 Associations of novel AS events with DNA methylation…………………………………..….51

LIST OF FIGURES

Figure 1 Computational detection of AS splicing events using Illumina sequence reads……………….…..52

Figure 2 RT-PCR gel of SR30 for seedlings treated with TSA, Azad-C, or both………………………...53

Figure 3 Gene structure diagrams for SR30, SR1, RSP31a for each of the identified isoforms………….54

Figure 4 Venn diagram showing isoforms that demonstrated altered abundance in response to TSA, Azad-C or both together………………………………………………………………………………..…..55

Figure 5 RT-PCR gel for wild-type and *rdd* mutant genotypes……………………………………...…56

Figure 6 Pair-wise comparisons of shared and unique splicing events between the wild-type and *met1, ddc*, and *rdd* mutants……………………………………………………………………………………57

Figure 7 Pair-wise comparisons of shared and unique splicing events between *met1, ddc,* and *rdd*…….58

Figure 8 Splice site sequence logos…………………………………………………………………......59

LIST OF ABBREVIATIONS

AltD	Alternative donor
AltA	Alternative acceptor
AltP	Alternative position
AS	Alternative splicing
Azad-C	5-Azadeoxycytidine
BRM	*BRAMA*
ES	Exon skipping
hnRNP	Heterogeneous ribonucleoprotein
IR	Intron retention
Pol II	RNA polymerase II
PTC	Premature stop codon
SR	Serine/arginine
TSA	Trichostatin A

ACKNOWLEDGEMENTS

I would like to thank Dr. Keith Adams for his help and guidance throughout the course of my degree.

This degree is dedicated to my parents who cared and supported me these many years, and to my grandma Flora who taught me a love of all things that grow.

1. INTRODUCTION

1.1 RNA PROCESSING

Gene expression involves an intimate coupling of transcription, RNA processing and translation. Transcription and RNA processing do not occur as separate steps but are now known to occur simultaneously in eukaryotes (5-14). Transcriptional regulation of gene expression involves a complex interplay between chromatin, transcription factors and RNA polymerase II (Pol II) (15). Precursor messenger RNA (pre-mRNA) must undergo the essential step of RNA splicing before the transcript is recognized as functional by the cell (16). The process of splicing involves the elimination of introns from the pre-mRNA by the massive ribonucleoprotein complex called the spliceosome, and the subsequent joining of the exons into a continuous transcript. Splicing allows for post-transcriptional regulation of gene expression through the nonsense mediated decay (NMD) pathway and has the potential to greatly expand the functional complement of the proteome (17-20).

For many years, it has been known that genes could produce more than one transcript but the prevalence and importance of this phenomenon was not appreciated until the advancement of both DNA sequencing and computer technologies allowed for rapid identification of different splicing events (21). The maturation process of pre-mRNAs allows for the assembly of multiple mature mRNAs through the differential utilization of splice sites in a process termed alternative splicing (AS) (21, 22). Numerous functional roles have been ascribed to AS such as, the functional expansion of the proteome, and post-transcriptional regulation of gene expression (18, 23, 24). Efficient and faithful recognition of degenerate splice sites involves a network of trans-acting factors that act to repress recognition of spurious sites and promote the assembly on true splice sites (18, 21, 25).

1.2 FREQUENCY AND CATEGORIZATION OF AS EVENTS

Numerous types of AS have been identified in eukaryotes and they range from simple inclusion of an intron to more elaborate and complicated splicing events (26). The major forms of AS that will be discussed are: Intron Retention (IR), Exon Skipping (ES), Alternative Donor (AltD), Alternative Acceptor (AltA), and Alternative Position (AltP). Intron Retention results in larger mRNAs that still possess the normally spliced out intron. Exon skipping encompasses a class of events that cause the final transcript to be missing one or more exons. Both AltD and AltA events involve the utilization of a different splice site at the 5' or 3' ends respectively while maintaining one consensus splice site. Alternative position occurs when both donor and acceptor sites used are different from the usual splice sites. The extent of Alternative splicing in *Homo sapiens* is estimated to be around ~95% of genes (23, 27), with plants generally thought of as having very little AS. Current estimates of the global prevalence of AS in *A. thaliana* propose that ~44% of genes can undergo AS in at least one organ type or stress conditions (21) .Variant mRNAs within higher plants contain retained introns and rarely undergo ES, whereas the situation is reversed within mammals (28). The prevalence of each type of AS varies greatly between different eukaryotes and may reflect differences in the evolution of their respective spliceosomal machineries (28). Eukaryotes belonging to animal lineages with large introns are thought to favor definition of exons to guide spliceosome assembly, whereas plant genes typically contain shorter introns and an intron definition mechanism is thought to exist (29).

1.3 REGULATION OF AS

Splice site sequences are short and highly degenerate in many higher eukaryotes and can statistically occur by chance frequently throughout the genome so tight regulation of spliceosome assembly is required to prevent undesirable splicing from occurring (30). Higher complementarity of the 5' splice site to the consensus U1 snRNP allows for more stable interactions of the snRNP (small nuclear ribonucleoprotein) to the pre-mRNA which promotes spliceosome assembly on that site (31). Proper splice site selection depends on trans-acting protein factors interacting with specific sequences in the pre-

mRNAs to promote or repress spliceosome assembly (30, 32). Serine/Arginine (SR) proteins recognize degenerate purine-rich sequences to promote splicing at the nearby splice site and are indispensable for both constitutive and alternative splicing (30, 32). Negative regulation of splice site selection is controlled in part by the heterogeneous nuclear ribonucleoproteins (hnRNPs) binding to a pre-mRNA and preventing positive regulators from binding (13). Opposing activities between the SR's and hnRNPs are thought to control splice site choice (30, 32).

Emerging evidence has highlighted that splicing occurs largely co-transcriptionally so the potential exists for the splicing process to be regulated by factors associated with transcription, and chromatin structure (7, 13, 16, 33). Splice site choice has been demonstrated to be affected by numerous factors besides the pre-mRNA sequence such as, the elongation rate of Pol II, and chromatin remodelers (34, 35). Promoter sequences have also been shown to specify the splicing patterns of their target genes and this regulation is not dependent on any part of the transcription unit but rather on the recruitment of specific transcription factors (36)

Chromatin structure has been implicated in the regulation of splicing by high-throughput CHIP sequencing in *C. elegans* which demonstrated that there are differential histone markings of introns compared to exons and this marking is weaker in alternatively spliced exons (37, 38). Histone modifications themselves have been shown directly to adjust the splicing patterns of genes (1, 39). The control of splicing by histone modifications seems to require a chromatin binding protein, as well as a splicing regulator (1, 39).The interaction of histone modifications with splicing was originally implicated by a histone deacetylase inhibitor called Trichostatin A (TSA), which causes hyperacetylation of histones and hypomethylation of DNA resulting in higher rates of transcription from genes, and the subsequent skipping of a cassette exon (40). *BRAMA* (BRM) is a histone remodeling complex that has been confirmed to be involved with splicing due to its presence in both promoters and the coding regions of some genes (41). The mechanism for BRMs control over splicing involves altering the elongation rate of Pol II over the body of a gene and stalling of the polymerase on alternative exon regions which allows

more time for weaker splice sites to be recognized (41, 42). Within the same experiment the BRM protein was demonstrated to co-purify with elements from the spliceosome and SR splicing regulatory proteins (41). Taken together, all of these results indicate that there is an epigenetic component to splicing control.

1.4 HIGH-THROUGHPUT STUDIES OF AS

Recent technological improvements have fostered a flurry of activity to inventory AS for whole genomes. Original methods for identification of AS relied on pools of EST/cDNA sequences and aligning them to genomic DNA to find different gene isoforms using strict quality control. Higher-throughput analysis of AS is possible using a combination of tilling and exon-exon junction microarrays, but these methods are limited to known AS events and cannot discover novel or more complex types of AS (43-45). Next generation sequencing has emerged as the best platform to gather massive amounts of data regarding both the expression level and splicing of both known and novel transcripts. Illumina ultra-high throughput sequencing produces millions of short sequence reads which are then mapped to a reference genome that can be read as digital counts of gene expression. The strength of this technology is that no *a priori* information regarding gene structure is required and AS events can be readily determined by analyzing splice junctions and classifying them according to the reference gene structure. Detection of known and novel AS has been successfully performed by numerous groups using Illumina sequencing(14, 21, 45-47).

1.5 FUNCTIONS OF AS

Alternate linear arrangements of exons in a processed transcript can have a broad range of effects on the coding capacity of an mRNA. AS occurring in the UTRs does not affect the final protein but can serve regulatory roles which are associated with changes in translatability and altering decay rates (48). Changes caused by AS in the coding region can have drastic consequences on the produced protein if the event shifts the reading frame. The classic example of a functional change in a protein by AS is rubisco activase where intron retention renders one of the produced proteins to be insensitive to the redox state of

the cell (49). The flowering time control gene *FCA* is known to undergo AS to produce four splice variants with only one isoform being able to promote flowering (50). Many AS events have been shown to introduce premature stop codons (PTC) which can trigger mRNA degradation through the NMD (Nonsense mediated decay) pathway (18, 51, 52). The primary role of NMD is as an RNA quality control pathway that degrades improperly spliced transcripts and prevents potentially deleterious expression of defective transcripts (20, 53). There is emerging evidence that the NMD pathway is responsible for the developmental stage specific regulation of isoform abundance (54). It has been proposed that some AS events are exploited for their ability to post-transcriptionally control the level of gene expression; however functional association between AS to NMD has been demonstrated experimentally for only a few genes (52).

1.6 SR PROTEINS

The SR protein family is characterized structurally by one or two RRM's (RNA recognition motifs) and an inherently disordered region of serine/arginine dipeptide repeats which is responsible for mediating protein-protein interactions and the target of numerous kinases (30). The importance of the SR's in both constitutive and alternative splicing has been well documented using both *in vitro* and *in vivo* methodologies (30, 32). The functions of the SR proteins are not limited solely to splicing but have expanded to regulating export, turnover and translatability of mRNA (55-58). Maintenance of mammalian genome integrity requires the functioning of *SC35* and deletion of this SR gene results in rapid and frequent DNA double-strand resulting from RNA-DNA hybrids (59-62). With 19 members present in *A. thaliana* and only 10 in humans, plants have nearly double the number of human SR genes (30). The functional importance of SR proteins within plant development is highlighted by an over expression study of *SR30* which induced a late flowering time phenotype in *A. thaliana* (56). Numerous AS variants of the SR genes are direct targets of NMD which are known to change in abundance during development and in response to numerous stressors (2, 18).

1.7 ESTABLISHMENT, MAINTENANCE AND FUNCTIONS OF DNA METHYLATION

DNA methylation is a prominent regulator of gene expression that is present in nearly all eukaryotes. Cytosine bases are the targets of DNA methyltransferases (MTases) which will covalently attach a methyl group to the fifth carbon in the purine ring (63, 64). The roles of DNA methylation within eukaryotic genomes is still a matter of debate, but it is generally accepted that DNA methylation represses the expression of retrotransposons (63, 65). The proposed roles of DNA methylation herald DNA methylation as the "caretaker" of the genome (64). Plastic DNA methylation would patterns permit for cell lineage specific tuning of gene expression throughout development (4, 66). Knowledge of DNA methylation has been greatly extended by the employment of DNA MTases mutants defective in propagation and/or establishment of proper methylation patterns (4, 67, 67-69).

Arabidopsis thaliana has been used extensively to study cytosine methylation using a variety of high throughput genomics techniques (4, 67-69). Plants are able to methylate cytosine bases outside of the canonical CpG dinucleotide context seen in mammals (4). In fact, plants have evolved very elaborate systems to methylate DNA in CpG, CpHpG and CpHpH (H = A, T, or C) contexts (67). Both passive and active mechanisms exist to establish, maintain and even remove DNA methylation from the genome (69). DNA Methyltransferase 1 (*MET1*) is the *A. thaliana* homolog of the mammalian maintenance methyltransferase (*DNMT1*) (70). In order to faithfully maintain DNA methylation after replication, *MET1* must correctly recognize hemimethylated DNA and correctly methylate the newly synthesized strand (71). An active RNA guides MTases such as Chromomethylase 3 (*CMT3*) and Domains Rearranged Methyltransferase 1 or 2(*DRM1/DRM2*) to their target DNA where they then methylate DNA complementary to the RNA (4). A class of enzymes related to DNA gylcosylases are involved in DNA repair and *ROS1* (repressor of silencing 1), *DML1*(Demeter-like 1), and *DML2*(Demeter-like 2) are responsible for excising and replacing improperly methylated cytosines with an unmodified base (64). Active demethylation has been demonstrated during endosperm development but abiotic stressors such as cold can induce loss of DNA methylation (72, 73).

The expression level of genes is strongly correlated with the presence of body methylation with moderately expressed genes tending to be methylated and genes that were expressed to either extreme demonstrating marked depletion of methylation (70, 74) . Genes possessing genic methylation have highly robust resistance to changes in their gene expression in response to numerous stress conditions (74). It has been proposed that a key function of DNA methylation within the body of genes is to buffer the expression of a wide number of genes against internal and external perturbations (74). Methylation within the promoter region of genes is associated with silencing and highly tissue specific patterns of expression (74, 75).

1.8 HIGH-THROUGHPUT STUDIES OF DNA METHYLATION

The intricacies of DNA methylation are compounded when it is considered within the context of the whole genome. Transcriptome analyses of the aforementioned mutants have revealed some extremely surprising insights (4, 75, 76). RNAi or catalytic knockout mutants of *MET1* show on mass reactivation of silenced transposons, as well as activation and repression of genic expression (70). Zilberman *et. al* 2007, found that roughly 20% of all genes in *A. thaliana* were methylated to some extent (70). They also uncovered a clear bias for DNA methylation to be absent from 5' and 3' UTRs (untranslated regions) (70). Correlation analyses also demonstrated that gene expression level was a good predictor for methylation status with moderately expressed genes having the highest probability of displaying body methylation (70). In contrast, genes with very high or low levels of expression were unlikely to show body methylation (70). Genes having tissue specific expression often showed methylation within the promoter region (70).

Illumina sequencing was performed to obtain single base pair resolution of the distribution of methylated cytosines throughout the *A. thaliana* genome using bisulfite conversion in wild-type and mutants defective in DNA methylation (4). Bisulfite treatment causes unmethylated cytosines to become uracils but does not convert methylated cytosines (4) Lister *et. al* 2008 concurrently sequenced the transcriptomes of the wild-type and mutant plants to understand functional consequences of DNA

methylation on the regulation of gene expression (4). Two DNA methylation defective mutants *met1* (*met1-3*), and *ddc* (*drm1-2, drm2-2, cmt3-11*) and one DNA demethylase mutant rdd (*ros1-3, dml2-1, dml3-1*) were examined in the Lister *et al.* 2008 study. Around 20- 30% of *A. thaliana* genes contain some genic methylation but loss of methylation in the *met1* background resulted in the upregulation of only 589 genes, which is in accordance with other studies (4). Lack of widespread changes in transcription resulting from loss of DNA methylation appears to contradict the belief that DNA methylation is critical for transcriptional control of genes.

1.9 CHROMATIN AND HISTONE MODIFICATIONS

The nucleosome represents the smallest functional unit of chromatin and it consists of a histone octamer of four different histone proteins with ~ 146bp of wrapped DNA (77). Extending from the core of the nucleosome are the basic histone tails that are the targets of numerous post-translational modifications {{176 Zilberman,D. 2008}}. Post-translational modifications of histones as well as histone variants are known to influence chromatin folding and accessibility of the associated DNA to transcription factors {171 Strahl,B.D. 2000}}. Histone tails can be decorated with several chemical groups such as methyl, acetyl, phosphor, sumo, ubiquitin, and ADP-ribose {171 Strahl,B.D. 2000}}. The flexibility and variability in histone modifications allows for high levels of information to be contained within the tails. It has been proposed that all of the modifications including their position and level on the different histone tails encompass a cellular "histone code" (77). Changing the chemical environment of the histone tails has large consequences for the interaction of the tails with negatively charged DNA as well as with other nucleosomes and additional chromatin scaffold proteins (78). Acetylation on histone tails is thought to destabilize the interactions between the coiled DNA and the histone tail and allow transcription factors to bind to their target sequences more readily (77). Methylation on histone tails is generally associated with a more compact chromatin structure since it does not destabilize the interaction between the DNA and its associated nucleosome which would restrict access to the DNA (79). A plethora of histone remodeling and modifying enzymes are present to control the ability of elongating Pol II to overcome nucleosome barriers to transcription (16, 80).

A highly simplified view of the effects of covalent histone modifications are that they alter the affinity of the histone tails for DNA, other histone tails, and accessory proteins which can create a more open chromatin state to promote gene expression or a more compact conformation and repress expression (79). Histone variants with slightly altered amino acid sequences are another potential source of regulatory control over gene expression (81, 82). The presence of the histone variant H2A.Z is enriched in the 5' region of genes where it has been proposed to attenuate transcription of its resident gene (83). Histone modifications and DNA methylation are inextricably linked with extensive crosstalk between different epigenetic marks (81, 82). Regions occupied by H2A.Z have been demonstrated to be depleted of DNA methylation and the presence of one of these modifications excludes the other (81, 82). DNA methylation and histone modifications showed nonrandom associations and have been demonstrated to have extensive crosstalk (84).

Recently the methylomes of eight diverse organisms were profiled using bisulfite conversion coupled to Illumina sequencing (85). DNA methylation has been determined to be an ancient modification that evolved in early eukaryotes due to the conserved patterns and distributions of methylation within the coding region of genes (85). Methylated genes also show distinct similarities across highly divergent species where methylated genes tended to have broad expression patterns and moderate levels of transcription (74, 85). Distributions of methylation within the coding regions of genes also showed marked enrichment in exons compared to introns (85), it will be interesting to elucidate whether this property exists to promote recognition of exons or if methylation is the result of the action of another process. Depletion of DNA methylation is consistent with a study in *C. elegans* which performed CHIP-seq and found that nucleosomes are preferentially associated with exons and markedly depleted from intronic sequences (37). Genic CpG methylation is severely reduced in the met1 but asymmetric methylation is recruited to a large number of genes in a pattern that is reflective of the original CpG methylation (4). Recruitment of non-CpG methylation in the met1 background occurs by unknown mechanism but it is tempting to speculate that nucleosomes themselves may promote methylation of the

associated DNA (85, 86). The roles of DNA methylation and nucleosomes in the recognition of exons likely overlap since they are inextricably linked with each other (78, 82, 86-88).

1.10 MAJOR GOALS AND OBJECTIVES

Regulation of AS by chromatin structure and histone modifications was recently discovered in mammals and metazoans (1, 37). Here I documented for the first time that there is an epigenetic component to AS regulation in plants.

1 - I investigated whether the AS patterns of plant genes could be changed in response to altering DNA methylation and histone acetylation state by using the SR protein family as a model to assay the impacts of altered DNA methylation and chromatin status on AS since the SR genes are known to possess considerable AS(2).

2 - Moreover, I wanted to determine if DNA methylation and histone hyperacetylation affect AS of the SR genes synergistically or antagonistically.

3 - The role of DNA methylation on organ and developmental specific RNA splicing was studied in a DNA hypermethylation mutant to see if organ-specific changes in AS of the SR genes could be affected by blocking active DNA demethylation. SR genes were chosen since there is a wealth of information available that has documented the extensive gene specific AS patterns that respond to stress, tissue and developmental state (2). The major questions of this part of the study required genes that possessed tissue specific patterns of AS.

4 - The prevalence of AS changes in three different DNA methylation mutants was estimated using Illumina sequence data to determine changes in AS in each of the mutant genotypes. The mutants studied consisted of two DNA hypomethylation and one hypermethylation mutant. The association of changes in methylcytosine content and production of novel AS was also explored.

2. MATERIALS AND METHODS

2.1 PLANT GROWTH AND TREATMENTS

Arabidopsis thaliana (Columbia ecotype) seeds were surface sterilized according to the procedures described in (76). Growing conditions for all experiments were conducted in a growth chamber set to an ambient temperature of 23°C with a 16 hour photoperiod. All chemicals used in these experiments were from Sigma Aldrich and of the highest purity available. Plants used for drug treatments were sown on sterile filter paper wetted with 0.5X MS Media and stratified at 4°C for 7 days before transferring them to the growth chamber. Plants were allowed to grow for eight days after germination before the seedlings were transferred to 15 ml of 0.5X liquid MS media supplemented with 10µM Trichostatin A(TSA) in DMSO and/or 40 µM 5-azadeoxycytidine(Azad-C) (76). Control plants and Azad-C treatments were supplemented with an equal volume of DMSO to account for any effects that DMSO might have. Treatments were applied to the seedlings for 7 days with daily media changes to maintain consistent concentrations throughout the experiment. Whole plants were harvested and flash frozen in liquid nitrogen and stored in -80°C before RNA extractions. Plants used for assaying organ type specific patterns of AS were grown in potting soil and harvested at 4 weeks when roots, leaves, flowers and siliques were present. Three biological replicates (each consisting of tissue from multiple plants) were collected for all drug treatments, and organ types. The *rdd* (*ros1-3 dml2-1 dml3-1*) demethylation mutant which was characterized in (69) was generously provided to me by Dr. Robert L. Fischer.

2.2 RNA PREPARATION

Drug treated seedlings were extracted using the RNeasy™ Plant RNA extraction kit (Qiagen) without any modifications. Organ type RNAs were extracted using the protocol described in ((89). Briefly, 100mg of tissue was ground in liquid nitrogen to a fine powder and then transferred to a 2ml tube containing 1ml of extraction buffer shaken vigorously. A 750µl volume of chloroform was added to the tube mixed thoroughly. Samples were then centrifuged at 10,000 rpm for 10 minutes at 4°C to and the resulting supernatant was transferred to a new tube and subjected to another treatment with chloroform. The supernatant was then transferred to a new 1.5ml tube and 1/3 volume of 8M LiCl was added to each

tube and mixed thoroughly. The RNAs were allowed to precipitate overnight at -20°C. Samples were pelleted by centrifugation and washed in 75% ethanol and resuspended in RNase free water. Concentration and quality of the RNA was assessed using a Nanodrop 1000 spectrophotometer and agarose gel electrophoresis.

2.3 RT-PCR

Genomic DNA contamination was removed by digestion with TurboDNAfree™ (Ambion) by treating 10μg of total RNA according to established protocol. Digestion was considered complete when 35 cycles of PCR failed to amplify a 150bp fragment of UBQ10. Reverse transcription of RNAs was performed using M-MLV (Invitrogen) with 1.5μg of total RNA and a oligo (dt) primer in a volume of 20 μl. Gene specific PCR assays used 0.5μl of the resulting cDNA in a 20μl reaction. Primers used this study were designed using web based Primer3 tools (http://frodo.wi.mit.edu/primer3/), the specific primer sequences are listed in Table 1. Optimal PCR conditions including cycle number for each primer pair were determined empirically to be 60°C for primer annealing and 30 cycles of PCR amplification. The housekeeping gene UBQ10 was amplified for 22 cycles and was selected due to multiple lines of evidence showing that its expression is unaffected under any of the conditions tested (4, 75, 76). PCR amplification was performed using Paq 5000™ (Stratagene) polymerase with denaturation at 95°C for three minutes during the first cycle and 30 seconds for all subsequent cycles, a 30 second annealing step at 60°C and a two minute extension step at 72°C. A final 15 minute extension was also conducted to ensure completed products. The PCR products were analyzed using native agarose gel electrophoresis using variable concentrations of agarose to optimize separation of products.

2.4 TRANSCRIPT QUANTIFICATION

The quantity of mRNA was normalized using UBQ10 gene expression to validate that equal amounts of RNA were present in the cDNA synthesis reaction. Analysis of mRNA isoform levels was performed using the NIH ImageJ program (http://rsbweb.nih.gov/ij/). End point PCR is not a quantitative measure of gene expression but real time PCR was not suitable for assaying the full length isoforms, so

isoform expression was measured internally in reference to the complete pool of amplified products (2). Each splice form was quantified using ImageJ and expressed as a percentage of all isoforms present in that sample in order to account for transcriptional differences between treatments (2) .

2.5 SPLICE VARIANT SEQUENCING AND ANALYSIS

Bands of interest were excised after electrophoresis size separation of PCR products. Gel extraction was performed using OmegaBioteck silica columns and re-amplified if necessary. Sequencing was performed using ABI BigDye 3.1. The resulting sequences were aligned to the genomic reference sequence using NCBI's spidey alignment software http://www.ncbi.nlm.nih.gov/spidey.

2.6 ILLUMINA READ MAPPING

Illumina reads were downloaded from the NCBI Short Read archive using the accession number SRA000284. Perl scripts were used to truncate the reads to 32 bp and filter low complexity and 'N' containing reads before aligning them to the Arabidopsis genome. TopHat v.1.0.13 (90) and Bowtie 0.12.3(91) were used to align spliced and unspliced reads to the reference Arabidopsis TAIR8 genome. Mismatching was not allowed and a minimum anchor for spliced reads was set to 8bp on either side, however up to two mismatches were allowed for unspliced reads. Maximum and minimum intron size was limited to 10kb and 50bp respectively and multi-mapping reads were suppressed. *De novo* discovery of splice junctions was performed for each genotype separately and the resulting junctions were pooled together to make a set of common junctions which were used in a secondary TopHat alignment with novel discovery turned off, which ensured consistency between all four of the tested genotypes.

2.7 DETECTION OF AS EVENTS

Custom PERL scripts were designed to associate uniquely mapped reads with a TAIR8 gene model and determine the number of reads per gene model. Alternative splicing events were grouped into five general categories consisting of alternative donor/acceptor (AltD, AltA), exon skipping (ES), alternative position (AltP) and intron retention (IR) (Fig.1). TAIR8 junctions were used as a resource for known AS events as well as to compare novel junctions to a known standard. A junction would be called

as an AltD or AltA event if the reference and the query junctions shared only one splice site in common and differed at the other, but did not begin or terminate in different exons. ES events were found if the junction began in the same exon as the reference junction and terminated in another exon. AltP events required that neither the 5' or 3' splice sites were shared between the query and the reference but the query junction must not skip an exon. *Bona fide* intron retention is difficult with 32bp reads so two different schemas were used to prevent duplication of one of the other types of events that could allow reads to map to an intron. If there was evidence for an AltD/AltA/AltP for a particular intron then an IR event required the presence of tags outside the region covered by splice junction. Introns that did not have evidence for splicing events inside them were said to be retained if at least 50% of the intron was covered by sequence tags. Events were considered real only if there were at least three sequence tags supporting it (14).

2.8 SPLICE SITE SEQUENCE ANALYSIS

Sequences corresponding to the splice sites for all AS events that met our criteria were extracted using the TAIR8 genomic reference. Ten base pairs of exon sequence up from the 5' SS and ten nucleotides into the intron were concatenated to ten bases of the intron upstream of the 3' SS and ten of the exon to produce the splice junctions. Splice sites for the entire Arabidopsis genome were collected to determine the global frequency of nucleotides at each position. Sequence logos of splice sites were produced using WebLogo(92).

3. RESULTS

3.1 EFFECTS OF CHEMICAL INHIBITION OF DNA AND HISTONE METHYLATION

Epigenetic regulation of AS has been the recent focus of several animal studies (1, 37, 38, 93, 93) but it is currently unknown what effects DNA and histone methylation have on AS in plants. The SR protein gene family was used as a model system for an RT-PCR based assay to investigate the relationship between DNA and histone methylation on AS in *A. thaliana*. Internal comparisons between isoform abundance are possible since the primers used in this study were designed to assay as many splice forms as possible in a single PCR reaction (Table 1). RT-PCR with gene specific primers detected AS in 13 out of the 18 tested SR genes but AS was not found in: SRZ21, SRZ22, RSZ22A, SCL30, or SCL28 so these genes were excluded from further analysis. Chemical demethylation of genomic DNA was performed using a cytosine analogue Azad-C which replaced a carbon for nitrogen at the fifth position in the cytosine ring (76). The effects of histone hyperacetylation on AS was tested using the histone deacetylase inhibitor TSA (76). In order to identify which splice variants are affected by the different chemical treatments the ratios of isoforms between treated and untreated plants was determined for three biological replicates (Table 2). A 20% increase or decrease in abundance of isoforms represented a visible difference in band intensities for *SR30* so this cutoff was used to quantify changes in isoform expression. Ratio changes of 50% or more are also indicated in Table 2.

RT-PCR using gene specific primers for *SR30* produced three isoforms which responded to one or more of the chemical treatments (Fig.2, Table 2). Isoform 1 differs from the fully spliced mRNA by the inclusion of the intron between exons 3-4 and an AltA event in the intron between exons 10 and 11(Fig. 3). There was a marked decrease in the abundance of isoform 1 in response to TSA treatment but Azad-C resulted in an increase in the presence of the splice variant. Isoform 2 lacks the IR event between exon 3-4 but possesses the same AltA event as isoform 1(Fig. 3). Treatments with Azad-C singly and in combination with TSA saw an increase in the level of isoform 2 but no effect was seen in TSA singly (Fig.1, Table 2). Isoform 4 is the fully spliced mRNA with 12 exons and the longest open reading frame

(Fig. 3). Isoform 4 decreased in response to treatment with Azad-C but was unaffected in treatments that included TSA.

SR1 was not affected by treatment with TSA but isoform 1 increased in response to Azad-C and both drugs applied together (Table 2). The first isoform of *SR1* possesses an AltP event present in the intron between exons 10 and 11 (Fig. 3). No effect was observed for forms 2 and 3 of *SR1* under any of the tested conditions. Opposing responses to drug treatments were seen in *RS31a* which saw an increase in isoform 1 by application of TSA but no net effect was seen under either of the other conditions (Table 2). The composition of the first isoform of *RS31a* differs from the reference by the retention of 4 out of 5 introns (Fig. 3). Isoform 2 and 3 increased in TSA and Azad-C singly but when both drugs were applied there was not a difference in their abundance compared to the control plants (Table 2). Isoform 2 differs from isoform 1 of *RS31a* in that it does not retain intron 4 (Fig. 3). Introns 3 and 5 are retained in both isoforms 2 and 3 but isoform 3 underwent an AltP event in intron 2 where isoform 2 retained the entire intron (Fig. 3). Azad-C resulted in a decrease in isoform 4 of *RS31a* but this effect was not seen in TSA or when both drugs are applied (Table 2). The fourth isoform of *RS31a* is the reference splice product and consists of 6 properly spliced exons (Fig. 3). In a final example, all three drug treatments for *RSZ32* resulted in an increase in abundance of isoforms 2 and 5 but drug specific effects on isoform levels were also observed (Table 2). Isoform 1 of *RSZ32* only increased in response to TSA treatment whereas isoform 4 decreased under the other two conditions (Table 2).

There is a large amount of transcript diversity present in the tested SR genes since 13 genes produced a total of 77 isoforms. With a cutoff of a 20% difference in splicing ratios I found that treatment with TSA and Azad-C modified splicing patterns for 12 of the 13 genes with AS, but only 10 were affected when the drugs were combined (Table 2). The number of genes affected under each treatment is very similar but there are considerable differences in how many isoforms display altered abundance. TSA caused 37 isoforms to change in abundance while Azad-C affected only 25, and both drugs together altered 27 (Table 2). Both drugs when applied together produced splice patterns that mirrored the

untreated control in the cases of *SR34a*, *RS31a*, and *SR41* (Table 2). TSA affected a larger number of isoforms than Azad-C even though both treatments resulted in 12 genes having measurable splicing differences. It is clear from Figure 3 that there are isoforms which are uniquely affected under TSA and Azad-C applied singly. The application of both drugs does not produce an additive impact on the splice patterns of the SR genes (Fig. 3). The data presented in Figure 3 also shows that Azad-C when applied in conjunction with TSA will produce splicing patterns more similar to Azad-C than to TSA. Differences observed for the splicing patterns were largely quantitative in nature but a few qualitative examples of presence/absence were found. Qualitative differences in splicing were observed in TSA and Azad-C applied singly as well as in combination for *RSZ32* (Table 2). *RSZ32* possesses a splice form that is not detectable in control plants but is present in all drug treatments based on the sensitivity of the assay. These results highlight that there are distinct effects of inhibiting histone and DNA methylation which do not function additively in splicing regulation.

3.2 AS OF SR GENES ACROSS DIFFERENT ORGANS AND DEVELOPMENTAL STAGES

Multiple SR proteins are known to possess organ and developmental stage specific splicing profiles (2) however it is not known how these patterns would be affected in an epigenetic mutant. To determine the effects of DNA methylation on the splicing patterns of the SR genes I utilized an RT-PCR based assay to amplify mostly complete isoforms in different tissue types with an *rdd* mutant which is defective in removal of methylated cytosines. I extracted RNA from roots, one and two week old whole seedlings, four week old rosette leaves, unopened flower buds, mature open flowers, as well as siliques. The organ types used for these experiments represent a diverse selection of reproductive as well as vegetative tissues.

SR30 has well documented tissue and developmental specific RNA splicing patterns in wild-type plants (2) and the results presented in Figure 4 indicate that regulation over splicing patterns across different organs and developmental stages remains intact in *rdd*. *SR30* is a dramatic example of tissue

specific AS in a hypermethylation mutant background, since there is no alteration of splicing ratios in roots, as well as in 1 and 2 week old seedlings. However, I did find that there were distinct changes in relative isoform abundance in 4 week rosette leaves, immature flower buds, mature flowers as well as siliques (Figure 2, Table 4). Rosette leaves of the rdd mutant resulted in isoform one decreasing in abundance with a concomitant increase in isoform 4 (Table 4). Sampled floral organs demonstrated an opposite effect on isoform 1 abundance (Table 4). Developmental stage specific effects on isoform 2 were observed between immature and mature flowers (Table 4). Splicing patterns obtained from sampling siliques produced an opposite effect on isoform 2 with an increased presence in the *rdd* mutant compared to the wild-type (Table 4). Splicing data obtained for *SR30* highlights the necessity for a broad sampling of different tissues types in order to gain in depth insights into the effects of DNA methylation on splicing.

SR1 splicing patterns change across the different tissues and developmental stages in both mutant and wild-type plants (Fig.3,Table 4). Reciprocal effects on isoform 1 were seen for *SR1* in response to developmental state since there was a decrease in mature flowers but an increase in flower buds in the *rdd* mutant (Fig.3,Table 4). However in one and two week old seedlings there is a consistent increase in the abundance of isoform 1 which indicates that the observed developmental stage effects are dependent on the sampled organ (Fig.3,Table 4). Root tissues displayed a reduction in isoforms 1 and 3 but an increase in form 2 (Fig.3, Table 4). Siliques and roots represent the only sampled tissues where all observed isoforms of *SR1* displayed changes in the *rdd* mutant (Fig.3, Table 4). Four week old rosette leaves were the only tissue sampled that displayed an increase in isoforms 1 and 2.

A total of eight spliced products were detected for *RS40* which indicates that this gene undergoes extensive AS (Fig.3). There were numerous changes in isoforms found in the *rdd* compared to the wild-type (Table 4). Siliques had the most profound effect on splicing of *RS40* with a total of 5 isoforms demonstrating altered levels (Fig.3, Table 5). Assaying roots uncovered a total of three isoforms upregulated in the *rdd* background (Fig.3, Table4). Again developmental specific and organ specific

effects were found between 1 and 2 week seedlings and between mature and immature flowers. No effect on splicing was observed in 2 week old seedlings but there was an increased level of isoforms 7 and 8 in 1 week old seedlings. Isoform 7 was upregulated in flower buds but was not detectable in the mature flowers of either wild-type or *rdd* (Table 4). Down regulation of isoform 4 in both flower buds and mature flowers is a case of organ specific effects on splicing which are not sensitive to developmental stage (Table 4).

Analysis of isoform splicing ratios for 13 *SR* genes revealed that there are a large number of splice products affected in the *rdd* mutant. The *rdd* triple mutant displays strong organ specific effects on the splicing patterns of 13 SR genes (Table3). The overall effects on AS in the *rdd* mutant was highly gene specific with *SR1* and *RSZ33* being the only two genes to be affected in all sampled tissues. Frequently it was observed that multiple isoforms would change in abundance in a single tissue and this is indicative that the observed changes are not the result of higher transcription level from a target locus but rather a quantitative shift in the splice products produced. The majority of splicing differences were quantitative in nature but a few cases of presence/absence were observed. Both gains and losses of isoforms were discovered in *rdd* mutant compared to the wild-type. Only nine occurrences of qualitative presence/absence of isoforms were found with the *rdd* mutant possessing 7 isoforms that were not detectable in wild-type and 2 isoforms in the wild-type that were not present in the *rdd. SR33* represents a qualitative example where isoform 2 is absent in wild-type roots but clearly evident in the rdd mutant (Table 2). Isoform 2 of *SR33* is not exclusive to the *rdd* mutant since it is present in wild-type: 2 week seedlings, 4 week rosettes, as well as mature flowers and siliques (Table 2). The SR genes demonstrated diverse changes in splicing in the *rdd* mutant but the production of “unnatural” splice products was not observed. The presence of “novel” splice forms was merely ectopic production of a “natural” splice form that would otherwise be restricted to a different tissue or developmental stage. Developmental specific differences between two vegetative seedling stages and different levels of maturity illustrate that a relationship exists between AS, DNA methylation and senescence.

3.3 COMPUTATIONAL ANALYSIS OF AS IN THREE DNA METHYLATION MUTANTS

The RT-PCR results from the SR genes demonstrates that AS patterns can change in response to DNA hypomethylation and histone hyperacetylation but this information cannot be used to gain an estimate of the prevalence of this phenomena. A publically available Illumina data from the NCBI short read archive (4) was employed to understand the effects of altered DNA methylation on AS on a large scale . Short read alignments were performed using TopHat v1.0.13 (90) and Bowtie v0.12.3(91) and the resulting output was parsed to identify AS events using custom perl scripts. The number of aligned reads for each of the sampled genotypes showed variability in the total number of mapped reads as well as where each of the reads localized (Table 7). The *ddc* mutant defective in CHH and CHG methylation had the highest number of aligned reads followed closely by the wild-type and the *rdd* hypermethylation mutant (Table7). The *met1* defective in CG methylation had the lowest number of mapped reads of any of the tested samples (Table 7). As expected the vast majority of RNA-seq reads aligned to annotated genes for all four samples but the wild-type had the highest number of intergenic reads which fell outside annotated transcripts (Table 7). The *ddc* mutant possessed the highest number of reads mapping to exons, exon-exon junctions, as well as introns(Table 7) This is not a general feature of an increased number of aligned reads since the wild-type had the second highest number of exon mapping reads but the lowest number of intronic reads (Table7). Junction spanning reads accounted for considerably less of the total pool of reads with less than 4 percent of reads spanning exon-exon boundaries (Table 7).

The results from my analyses indicate that there is more than a 3 fold increase in the number of AS events in mutant genotypes compared to the wild-type (Table 8). All types of splicing events increased with IR events representing the largest difference with more than 10 fold increase in the mutants compared to the wild-type (Table 6). All types of AS events are higher in the mutant compared to the wild-type which demonstrates that these mutations have broad reaching affects on the splicing of a large number of genes. Due to the large difference in the numbers of detected AS in the mutants compared to the wild-type I grouped both constitutive and alternative splicing together and determined

the number of shared and unique events present in each genotype which are presented in Figure 5. The results of pair-wise comparisons indicate that only ~43% of splicing events were common to the *met1* and wild-type. Similar results were obtained for both the *ddc* and *rdd* mutants with ~43 percent of splicing events shared between the mutants and wild-type (Fig. 5). Pair-wise comparisons also demonstrate that there is not simply less splicing present in the wild-type since unique events were detected in the wild-type that were absent from the mutants. Overall the results obtained signify that there is remarkable consistency between shared and unique events across the considered mutants. Mutant splicing events were compared to determine if the same RNA splicing was affected in each mutant. The results obtained from this analysis show that the mutants share a greater percentage of events compared to the wild-type but a significant fraction of splicing events are unique to each mutant (Fig.6). When comparing the *ddc* and *rdd* mutants it became apparent that they are more similar to each other than to the met1 which is consistent with the much more drastic changes in genome methylation present in the met1 (4). It is worth noting that the *ddc* mutant possessed more unique AS events in all pair-wise comparisons performed (Fig.6).

DNA methylation is a critical regulator of the genome and mutations in enzymes responsible for proper establishment, propagation and removal of genomic patterns of methylcytosine have far reaching implications for the expression of a wide number of genes. Genome-wide information for methylcytosine content of genes was generously provided by Ryan Lister (4) and was used to associate changes in body methylation status with differences in AS (Table7). Decreased DNA methylation accounted for more changes in splicing compared to increased methylation. Only a small proportion of splicing changes can be associated with altered DNA methylation within the coding region of the gene. In *met1* there is an association of DNA methylation with AS for only 14.6% of genes and in the *ddc* and *rdd* there was 17.1% and 8.2%. Changes in body methylation are associated only with a small percentage of AS changes. I would like to make it clear that association with DNA methylation and AS does not indicate causation since the exact effects of local DNA methylation on splicing has not been demonstrated experimentally.

Since only a small percentage of all AS changes could be associated with a cis change in genic DNA methylation I wanted to determine whether there are intron features associated with IR events in each of the mutant genotypes. Unique IR events present in only one genotype were used for splice site sequence analysis. The produced sequence logos are presented in Figure 7 and represent 20bp surrounding the 5' and 3' splice sites. Within the 5' exon there is a marked difference in sequence composition in all of the mutant backgrounds (Fig. 7). Functionally important nucleotides for base pairing to U1 snRNP do not deviate from the genome-wide average. Since splice sites function as pairs, I also analyzed the 3' splice site to determine if there are any significant differences in the introns compared to the genome-wide average. Taken together these results indicate that there are changes in sequence compositions around splice sites for introns that are uniquely retained in each of the mutants.

4. DISCUSSION

4.1 AS PATTERNS OF THE SR GENES ARE AFFECTED BY DNA METHYLATION AND HISTONE HYPERACETYLATION

AS has emerged as a pervasive property of plant genes but epigenetic regulation of RNA splicing is an unexplored avenue of investigation. DNA methylation is a well characterized mechanism of transcriptional gene regulation and several genome-wide studies have been conducted in *A. thaliana*(4, 67, 70, 75, 94). Numerous studies in animals and fungi have demonstrated that pre-mRNA splicing is a highly regulated process that is affected by Pol II elongation rate, promoter structure, and recently histone modifications themselves have been shown to directly influence splice site choice(1, 6, 15, 16, 36-38, 95). In order to determine whether AS is affected by changes in DNA and histone methylation I treated *A. thaliana* seedlings with TSA and Azad-C separately and in combination with each other to investigate if the splicing patterns of the SR genes are altered.

Inhibition of histone deacetylation can alter splicing decisions *in vivo* since treatment with TSA adjusted the splicing ratios of 12 SR genes (Table 3). Application of Azad-C to *A. thaliana* seedlings caused differential splicing ratios of 12 SR genes (Table 3). Combined treatment with TSA and Azad-C produced splicing patterns that were not present in plants treated singly as well as restored splicing patterns to that of untreated plants in the case of three genes (Table 3); consistent with a transcript profiling study which found that Azad-C and TSA can function antagonistically on gene expression(68). Histone and DNA methylation have highly gene specific effects on splicing patterns and both modifications can have opposing impacts on splicing decisions (Table 2). Affecting histone modifications directly had a greater overall effect on the splicing of the SR genes which suggests that DNA methylation my function primarily as a stable mark for the maintenance of histone markings (Table 3) (96). My results reveal that RNA splicing patterns of the SR genes diversely respond to genome-wide DNA demethylation and histone hyperacetylation.

The observed changes in splicing could be the direct effect of changes in chromatin structure influencing splicing outcomes or it is also possible that changes in overall transcription could induce splicing changes due to negative feedback loops (19, 97). Splice variants like isoform 3 of *RS31* which responded to all treatments in the same way, likely reflect splicing events which are sensitive to an increased elongation rate of Pol II, rather than specific chromatin signatures (Fig. 2). *SR33* represents an interesting case where simultaneous application of both drugs resulted in an increased number of misregulated isoforms (Table 3) which may be the result from the release of a "double lock" of epigenetic modifications over splicing (84). My pharmacological approaches represent a brute force methodology which results in transcriptional changes of a large number of genes and gaining mechanistic insights into the factors directly causing changes in AS are extraordinarily challenging. Targeting exon and intron portions of genes with siRNAs has been demonstrated to affect the splicing of the EDI exon in humans by inducing a closed chromatin state which slowed Pol II elongation (98). Treatment with both TSA and Azad-C singly or in conjunction with each other was shown to ameliorate the effects of the siRNA silencing by increasing the rate of transcript synthesis (98). The data obtained for the SR genes indicates that many of the observed changes in splicing cannot be explained solely by a faster Pol II, since treatment with drugs associated with higher transcriptional rates from genes failed to produce consistent effects on AS (Figure 2).

4.2 ORGAN AND DEVELOPMENTAL SPECIFIC EFFECTS ON SPLICING PATTERNS OF THE SR GENES

There is a spatial and temporal component to AS regulation which tailors gene expression and proteomic diversity to each cell's specific requirements (27, 55, 93, 99, 100). I investigated the possibility that epigenetic modifications contribute to specifying organ specific patterns of RNA splicing in a mutant background where the ability to remove DNA methylation has been abolished (69). Loss of DNA methylation is known to induce widespread transcriptional changes in a large number of genes but in the hypermethylation background there are relatively few genes with measurable differences in expression (4, 69). The low number of misregulated genes makes the *rdd* mutant an attractive system to

study the effects of DNA methylation on splicing. The *rdd* mutant maintained the ability to undergo organ and developmental specific splicing but distinct changes in isoform abundance were observed in different tissues (Table 4). Epigenetic modifications likely contribute to establishing splicing patterns but numerous other mechanisms exist to regulate and specify AS patterns since these plants were still able to produce distinct splicing patterns across the sampled tissue types (Table 5).

Many of the SR genes displayed tissue specific AS differences in the *rdd* hypermethylation mutant compared to the wild-type. There were more isoforms with altered abundance in roots and mature flowers compared to all other sampled tissues (Table 5). The two developmental stages of seedlings showed the fewest number of changes in isoforms compared to the other sampled tissues (Table 5); this may in part be due to the accumulation of DNA methylation changes over the lifespan of the plant. Dynamic DNA methylation has been shown to be a regulator of a myriad of developmental processes (94, 101, 102) and my results attribute a new function for DNA methylation in plants. Moreover, the data collected from the *rdd* mutant demonstrates that disrupting the ability of plants to modulate the distribution of DNA methylation induced changes in AS.

Broad sampling of different tissues and organ types uncovered a highly complex relationship between AS and loss of demethylation activity. The total number of genes and isoforms affected varied between the sampled tissues and opposite effects on isoforms were observed between different samples. Sampling based methodologies must include a broad array of tissues in order to gain a more accurate view of the effects of epigenetic modifications on splicing. Splicing outcomes are decided by the interplay between transcription and activities of antagonistic splicing factors that exist in different concentrations across tissues (57, 59, 103, 104). Tissue dependent effects on isoforms could be due to the cis DNA methylation changes or changes in the expression of positive and negative splicing regulators.

Chemical treatments were associated with widespread changes in AS in response to global reductions in DNA methylation and histone hyperacetylation. Genome-wide distributions of

methylcytosines in the *rdd* mutant have been the subject of several investigations which uncovered that there are roughly ~180 discrete loci targeted for hypermethylation in the triple mutant but there were also regions that are normally densely methylated in the wild-type but dropped in methylation level in the mutant (4, 67, 69). Therefore, the *rdd* mutant does not represent a strict hypermethylation phenotype but methylation density for some genes also decreased. DNA methylation is known to change over the course of development and between different organs (94, 105) so the observed changes in splicing within the *rdd* mutant may reflect specific alterations in methylation status which are prevented in the mutant genotype.

4.5 COMPUTATIONAL ANALYSIS OF AS EVENTS IN DNA METHYLATION MUTANTS

The results obtained from the SR genes represent a small sample size so I utilized a genome-wide analysis of AS using ultra-high throughput sequencing data from the NCBI short read archive for three DNA methylation mutants. Comparative analysis between the mutants and the wild-type uncovered a large increase in AS in each of the mutant genotypes (Table 7). Increased detection of AS in the mutants compared to the wild-type is not due to an increased number of aligned reads since the wild-type yielded the second highest total number of mapped reads (Table 6). All of the five general types of AS events increased in the mutants but IR events saw the largest increase. Previous computational studies of AS in plants showed that the most prevalent type of AS is IR (26, 106, 107) but using stringent cutoffs for IR I found that AltD events were the most abundant in wild-type plants (Table 7). This result is likely an artifact produced by the need for high coverage of introns in order to reliably identify IR events and biases against lowly expressed genes and minor splicing events. ES represents a rare event in all sampled genotypes which is in good concordance with its expected frequency from EST based studies (26).

Determination of *bona fide* IR events from 32bp Illumina reads represents an extraordinary challenge since a single read cannot unambiguously specify the exact AS event that occurred in that region. My estimates of IR events are likely an underestimate due to highly stringent cutoffs required for

calling an intron as retained. The most prevalent AS event in plants is IR so it is reasonable to assume that if DNA methylation influenced AS then there would be a significant increase in IR events. Increased IR in the mutants does not represent a defect in expression of genes associated with debranching or degradation of intron lariats since RT-PCR assays from the *rdd* mutant with mostly complete isoforms revealed altered AS patterns. Computational analysis of genes misregulated in the mutant backgrounds did not reveal any genes associated with intron lariat degradation or debranching enzymes. Furthermore, an overall increase in both constitutive splicing and other AS events argues against catalytic splicing defects accounting for the differences identified in these mutants. Together these results imply that affecting DNA methylation changes splicing decisions rather than inducing a defect in overall splicing efficiency.

Comparing both constitutive and alternative splicing between the mutants and the wild-type uncovered a strikingly similar number of shared events between the mutants and the wild-type (Fig.3). Roughly 43% of identified splicing events are present in both mutant and wild-type backgrounds so there are a large proportion of unique spicing events (Fig.3). Gains and losses of splicing occurred in the mutants since the wild-type possessed splicing events that were absent in the mutants (Fig.3). The similarity of shared and unique events in the mutant backgrounds was surprising considering the unique effects of the mutations on genome-wide distribution of DNA methylation. In order to ascertain if AS profiles in the mutants were shared across the different genotypes I performed pair-wise comparisons between each of the mutants (Fig.4). Mutants possess a greater number of shared events between each other than the wild-type but unique splicing events were found in each genotype (Fig.4). The results of the pair-wise comparisons show that the underlying mutations affect a broad spectrum of genes but have specific effects on splicing that are unique to each background. A closer examination of the precise changes in methylation would help to understand the exact associations of DNA methylation with AS.

4.6 ASSOCIATION OF NOVEL AS WITH CIS CHANGES IN DNA METHYLATION

Reporting the number of changes in AS in each of the mutants is an interesting result but each of the backgrounds has well documented misregulation of genes which could account for the alterations in splicing. An association analysis between methylation status and AS was performed in each of the mutants compared to the wild-type in order to see whether cis changes in genic DNA methylation are correlated with changes in AS (Table 7). Both gains and losses of DNA methylation were associated with novel AS events however lowered DNA methylation had a more profound effect on AS than accumulation of methylcytosines (Table 7). Novel AS events associated with DNA methylation changes accounted for a small proportion of all detected events in each of the mutants. DNA methylation distribution in *met1* is starkly different from the wild-type (4, 67) so it was unexpected that such a small proportion of AS changes would be associated with changes in DNA methylation content. The distribution of DNA methylation within genes in the *ddc* mutant was largely intact but transposons and intergenic regions displayed distinct changes in presence of methylcytosine {183 Lister,R. 2008}. Genic DNA methylation in the coding region is unlikely to play a predominant role in specifying AS patterns since *ddc* mutant with largely intact CpG methylation in genes displayed the largest number of AS associated with DNA methylation status changes (Table 7). Genome-wide patterns of methylcytosines in the *rdd* mutant were largely reflective of the wild-type patterns {183 Lister,R. 2008}which is consistent with the low association of cis genic DNA methylation to ectoptic AS. Methylation in the coding regions of genes may serve to promote the efficient recognition of exons however examination of all three mutants indicates that this is not the case for a large number of genes due to the rather low association of methylation within the coding region and altered splicing.

4.7 RETAINED INTRONS POSSESS ALTERED NUCLEOTIDE COMPOSITION AROUND SPLICE SITES

My analysis of the association between cis changes in DNA methylation and AS demonstrated that only a small proportion of changes in AS could be directly associated with altered DNA methylation. It is known that splicing can be impacted by the presence of splicing enhancer/silencer sequences in the

pre-mRNA as well as influenced by splice site strength (108-112). Intron retention results from a failure of the splicing machinery to recognize and assemble on the 5' and 3' splice sites. I examined the sequences around the 5' and 3' splice sites for introns that were uniquely retained in each genotype (Fig. 8). Analysis at the 5' splice site indicated that there is high conservation of the nucleotides functionally important with U1 snRNP base pairing but there are distinct changes in frequency of nucleotides in the upstream portion of the exon for the met1 and *rdd* mutants compared to the genome-wide average (Fig. 8). The *ddc* mutant does not demonstrate a high degree of variability in the 5' exon immediately prior to the exon with the exception of A->C switch in dominance at the -5 position (Fig.8). The composition of nucleotides in the 3' exon also demonstrates changes in frequency across all three mutants in comparison to the global average for all introns (Fig. 8). Based on these results there is not a unifying change at the 5' splice site in each of the mutants but the 3' exon showed changes in nucleotide frequencies in all of the mutants. It is tempting to speculate that weaker splice sites are more susceptible to changes in chromatin compared to stronger splice sites.

4.8 AS IN PLANTS IS AFFECTED BY CHANGES IN DNA METHYLATION AND HISTONE ACETYLATION

From the presented work it is clear that there AS is affected by conditions that alter DNA methylation and histone acetylation. This work also demonstrated that the contributions of DNA methylation to the regulation of AS are highly complex and largely functions in a supportive role to promote or repress splicing at a particular site. Co-ordination of RNA splicing across organ types and developmental stages has an epigenetic component to fine-tune the splicing of genes since there is not a complete breakdown in partition of splice patterns (Table 4). The results from both the drug treatments and bioinformatics demonstrate that DNA methylation likely functions indirectly in the regulation of splicing by specifying specific histone marking rather than directly recruiting splicing factors. The observed changes in splicing across the numerous treatments and mutants cannot be explained by increased Pol II elongation rate since genic DNA methylation is not consistently altered in the genes that produced altered AS patterns. It was found in humans that histone methylation directly regulates the

splicing of nascent pre-mRNA (1) and the combined results presented in this thesis are consistent with this mode of AS regulation. Determining what factors directly affect AS is exceptionally challenging when epigenetic regulators are affected on a genome-wide scale.

4.9 IMPLICATIONS OF EPIGENETIC REGULATION OF AS

Directly inhibiting DNA methylation and histone deacetylation through drug treatments and broad spectrum DNA methylation mutants represent an extreme phenotype but plants are known to experience comparable situations in nature. During cold stress it is known that there is genome-wide demethylation (73, 113) and it is known that extensive changes in AS occur during low temperature stress(114). Epigenetic changes may provide the ability for the cell to alter the splicing patterns of a wide number of genes simultaneously. Polyploidy and hybridization are also known to induce changes in DNA methylation patterns immediately and over successive generations in a species dependant manner (84, 84, 84, 84, 115, 116). Epigenetic changes associated with hybridization may be partially responsible for novel AS events seen in *Populus* hybrids (117).

5. CONCLUSIONS AND FUTURE DIRECTIONS

In summary, the experiments performed in this thesis work represent the first demonstration that DNA methylation can cause AS changes as well as the first description of plants possessing epigenetic regulation of AS. The data revealed that depleting DNA methylation and inducing hyperacetylation of histones affects the splicing ratios of numerous isoforms of the SR genes. Moreover, there are strong developmental and organ specific effects on splicing in response to removing demethylation activities in *A. thaliana*. Genome-wide profiling of publically available bisulfate sequence data and Illumina sequence data revealed that cis changes in DNA methylation were associated with novel AS events in each of the mutants. The association between changes in methylcytosine composition within the coding regions of genes and novel AS events was rather weak. I found that there were marked changes around the splice sites of uniquely retained introns in each of the mutants which may indicate that epigenetic regulation of splicing affects genes with weaker splice sites. Taken together my results implicate that DNA methylation may function to directly affect splicing decision but the overall contributions are difficult to ascertain given the fact that DNA methylation is a prominent regulator of gene expression and affecting DNA methylation has consequences for numerous genes. Also, future experiments should try to elucidate mechanistic insights into whether DNA methylation is a direct regulator of RNA splicing or whether it functions to specify other chromatin markings which will directly interact with splicing factors.

Table 1: Gene specific primers used to amplify mostly completed transcripts of the SR genes.

Name	AGI	Fwd Seq	Rev Seq	Genomic(bp)
SR30	AT1G09140	CGCAAGTGTGAGGTTGAAGA	ATGCAGCCGAGACAGAGTTT	3080
SR1	AT1G02840	TCGACGACCAACAGAATGAG	GCTAGGGCTCTTGCTTCCTT	2632
SR34a	AT3G49430	TCCGAGTATTGTTGGCTTCA	ATGGTCACGGAAACATCCAT	3362
SR34b	AT4G02430	TCGCCTTTCTTCTTCTCTTCC	CTCGATGGGTCCATTTTGAT	2364
RS40	AT4G25500	ACTACGCCTGCCAAAATCAT	TGCAAAATTCCTCCTGATCC	1788
RS31a	AT2G46610	GCAAGTTCGGGAGAGTGAAG	GCTCTCTTGTGCGAATAGGG	1043
SR41	AT5G52040	GTGATCTTGAGCGGCTTTTC	GCGATTTCGAATGGAGTCAT	2359
SRZ21	AT1G23860	GCTGCTCAGGACCAAGAATC	AAAAGGCGCCACAGAGTAGA	1261
SRZ22	AT4G31580	CCCTTGAGTGCTTTCAGCTC	CATAGCTTGGGCTCCTTCTG	947
RSZ22a	AT2G24590	AAAGAGCCAAAGCCGTTTCT	CCTCTTGACGTCACAAGCAA	1249
RSZ32	AT3G53500	TGCACTCTCCTTCTTCGTCTC	TGTCCTCCACGCTTTTCTCT	2249
RSZ33	AT2G37340	CCTTCCGAAATTAGGGCTTC	GAAGCTTTGGTGATGGTGGT	2132
SC35	AT5G64200	GGAAAACCATGTCGCACTTC	CTCAACATGGTTGTGCCATC	2067
SR33	AT1G55310	CCAGTCTTTTGGTTCGCAAT	TCCCATCATATCGCTCTTCC	1605
SCL30a	AT3G13570	TTCCCCTGTGTTTTTCTTCG	CTTTGGCTCCTTGCTTGTTC	2409
SCL30	AT3G55460	CTCCTCGACGTGGATATGGT	ACCTTCATAGCCAGGGGAGT	2158
RS31	AT3G61860	GGAACGGTTGTTCGACAAGT	GGACTTGGACGCCTACGATA	1491
SCL28	AT5G18810	GTTTGCTTATCCGTAATCTC	ATAGATCTTGACCTTCTTCG	1802

Table2: Alternative splicing ratios of SR gene isoforms resulting from treatment with TSA and Azad-C applied singly or in conjunction with each other. All values represent an isoform ratio between treated and control seedlings. Shaded boxes indicate greater than 20% difference and bolded numbers indicate greater than 50%. Percent standard deviation of ratios across replicates is indicated on the right. Isoforms that were not detected in both treatments are marked N/A. Boxes marked T indicate an isoform where it is present in TSA but absent in the control plants. Boxes marked with A indicate isoforms that were only present in the Azad-C treatment and not in the control plants. Boxes marked with both indicate that the Azad and TSA combined treatment had an isoform that was not present in the control. "C" marked boxed indicate a form that is detectable only in the control plant.

Gene	Locus	Isoform	TSA	Azad	Both
SR30	AT1G09140	1	0.54±0.08	**1.76±0.15**	1.03±0.06
		2	0.92±0.16	1.40±0.05	1.41±0.04
		3	N/A	N/A	N/A
		4	0.89±0.16	0.67±0.02	0.86±0.02
SR1	AT1G02840	1	0.80±0.09	1.35±0.07	1.32±0.11
		2	0.99±0.08	1.04±0.06	0.93±0.07
		3	1.05±0.04	0.90±0.03	0.94±0.05
SR34a	AT3G49430	1	1.34±0.07	1.19±0.12	0.83±0.18
		2	0.82±0.06	1.09±0.07	0.99±0.11
		3	1.00±0.01	0.98±0.01	1.01±0.02
SR34b	AT4G02430	1	1.40±0.05	0.92±0.09	1.46±0.20
		2	**2.94±0.25**	1.44±0.57	**2.13±0.30**
		3	0.59±0.02	0.87±0.01	0.98±0.02
		4	1.12±0.29	0.99±0.14	0.75±0.14
		5	0.55±0.21	0.75±0.027	0.79±0.02
RS40	AT4G25500	1	**0.43±0.04**	0.71±0.01	0.68±0.01
		2	1.41±0.07	0.80±0.03	0.91±0.01
		3	0.87±0.07	0.80±0.07	0.86±0.04
		4	**0.30±0.13**	0.82±0.17	1.04±0.13
		5	**2.59±0.13**	1.33±0.05	1.12±0.02
		6	**1.80±0.13**	1.36±0.10	1.38±0.07
		7	**6.98±1.45**	1.03±0.80	0.95±0.26
		8	**0.21±0.01**	1.00±0.02	1.03±0.02
RS31a	AT2G46610	1	**1.50±0.42**	0.96±0.02	1.01±0.04
		2	**2.68±0.72**	**1.76±0.24**	1.12±0.07
		3	1.31±0.33	1.24±0.05	0.83±0.01
		4	0.87±0.12	**0.46±0.04**	1.19±0.05

Table 2. (Continued)

Gene	Locus	Isoform	TSA	Azad	Both
RSZ32	AT3G53500	1	**1.56±0.68**	1.23±0.55	**1.51±0.77**
		2	0.57±0.67	0.93±0.99	0.67±0.54
		3	1.34±0.64	**1.96±1.67**	**2.47±2.53**
		4	**1.52±0.83**	1.4±1.21	**2.00±2.06**
		5	1.31±0.50	1.38±0.09	**1.64±0.36**
		6	T	A	Both
		7	0.98±0.05	0.98±0.05	0.96±0.10
RSZ33	AT2G37340	1	1.29±0.07	1.18±0.10	1.02±0.07
		2	0.99±0.04	0.81±0.07	0.70±0.15
		3	1.43±0.11	1.27±0.26	1.28±0.12
		4	0.93±0.13	0.73±0.15	0.55±0.06
		5	1.42±0.10	1.45±0.19	**1.53±0.07**
		6	N/A	N/A	N/A
		7	0.85±0.01	1.12±0.10	1.14±0.05
		8	0.82±0.02	0.84±0.03	0.88±0.03
SC35	AT5G64200	1	N/A	N/A	N/A
		2	N/A	N/A	N/A
		3	1.24±0.22	0.49±0.06	C
		4	1.20±0.01	0.93±0.08	1.49±0.21
		5	0.97±0.01	0.99±0.08	1.05±0.02
SR33	AT1G55310	1	1.04±0.07	0.75±0.19	**2.73±0.18**
		2	**1.74±0.06**	0.96±0.20	1.24±0.32
		3	1.04±0.09	0.92±0.17	0.66±0.13
		4	1.04±0.07	0.96±0.08	0.93±0.03
		5	0.89±0.02	0.94±0.13	**0.14±0.02**
SCL30a	AT3G13570	1	1.40±0.06	0.80±0.02	**0.40±0.02**
		2	1.18±0.10	0.84±0.06	0.63±0.02
		3	1.31±0.20	0.72±0.16	1.18±0.8
		4	N/A	N/A	N/A
		5	1.47±0.15	1.44±0.05	**1.5±0.19**
		6	1.33±0.19	0.68±0.19	**1.54±0.26**
		7	0.88±0.03	1.02±0.03	0.98±0.03
RS31	AT3G61860	1	1.29±0.05	0.96±0.25	1.13±0.26
		2	1.27±0.04	1.45±0.18	1.01±0.05
		3	**1.63±0.18**	1.40±0.20	1.33±0.17
		4	0.64±0.01	1.21±0.13	0.96±0.05

Table 2. (Continued)

Gene	Locus	Isoform	TSA	Azad	Both
SR41	AT5G52040	1	1.28±0.69	**2.35±1.12**	1.16±0.27
		2	N/A	N/A	N/A
		3	N/A	N/A	N/A
		4	N/A	N/A	N/A
		5	N/A	N/A	N/A
		6	0.98±0.07	0.76±0.22	0.85±0.08
		7	1.00±0.01	1.00±0.01	1.01±0.01

Table 3. The number of isoforms with altered abundance ratios in TSA and Azad-C applied singly and in combination.

Gene	TSA	Azad	TSA+Azad
SR30	1	3	1
SR1	0	1	1
SR34a	1	0	0
SR34b	4	1	3
RS40	7	3	2
RS31a	3	2	0
SR41	1	2	0
RSZ32	6	3	6
RSZ33	3	3	4
SC35	1	1	2
SR33	1	1	4
SCL30a	4	3	4
RS31	4	2	1

Table 4. Alternative splicing ratios of SR gene isoforms in an *rdd* mutant compared to wild-type across vegetative and reproductive organs and developmental stages. All values represent an isoform ratio between wild-type and *rdd* plants. Shaded boxes indicate greater than 20% difference and bolded numbers indicate greater than 50%. Percent standard deviation of ratios across replicates is indicated on the right. Isoforms that were not detected in the wild-type or *rdd* are marked N/A. Boxes marked "R" represent the presence of an isoform that is not detectable in the wild-type. Conversely, a box marked W indicates an isoform that is present only in the wild-type.

Locus	Isoform	Root	1 Week Seedling	2Week Seedling	4 Week Rosette	Flower Buds	Mature Flowers	Siliques
SR30	1	0.80±0.01	0.80±0.01	0.92±0.01	0.73±0.01	1.26±0.03	1.26±0.03	0.87±0.02
	2	1.03±0.02	1.03±0.02	1.08±0.03	1.23±0.02	**0.39±0.07**	**0.39±0.07**	**1.61±0.16**
	3	0.91±0.05	0.91±0.05	N/A	N/A	N/A	N/A	W
	4	1.16±0.01	1.16±0.01	1.05±0.03	1.44±0.02	1.15±0.00	1.15±0.00	1.07±0.01
SR1	1	0.52±0.26	0.62±0.30	0.76±0.16	1.22±0.02	1.21±0.12	0.53±0.1	0.77±0.01
	2	**3.38±2.16**	0.91±0.08	1.13±0.08	1.31±0.22	1.01±0.06	0.42±0.25	1.40±0.14
	3	**0.36±0,50**	1.17±0.26	0.95±0.02	0.99±0.01	0.99±0.01	1.08±0.03	1.41±0.03
SR34a	1	**0.41±0.02**	0.96±0.18	0.99±0.15	0.88±0.02	0.97±0.04	**0.25±0.13**	0.95±0.14
	2	**0.31±0.01**	0.75±0.19	1.14±0.25	1.00±0.01	0.66±0.21	0.97±1.01	0.59±0.12
	3	1.22±0.01	1.06±0.02	1.00±0.02	0.68±0.22	1.02±0.02	1.02±0.02	1.06±0.04
SR34b	1	1.21±0.36	0.91±0.01	0.81±0.2	0.98±0.28	1.21±0.14	1.21±0.14	**0.45±0.27**
	2	0.94±0.07	0.87±0.03	0.87±0.03	1.29±0.11	1.12±0.11	1.12±0.04	**1.75±0.16**
	3	1.13±0.17	0.98±0.03	1.06±0.03	0.85±0.15	0.85±0.15	1.09±0.15	0.60±0.16
	4	1.05±0.09	1.15±0.02	1.14±0.03	0.69±0.09	0.94±0.01	0.94±0.01	0.94±0.23
RS40	1	1.41±0.44	0.92±0.03	0.87±0.09	1.08±0.00	1.12±0.00	0.79±0.27	**1.88±0.02**
	2	**1.93±0.75**	1.07±0.06	0.98±0.08	1.44±0.09	0.88±0.01	0.77±0.19	1.11±0.02
	3	1.34±0.75	1.14±0.05	0.90±0.06	1.18±0.19	0.89±0.21	**3.13±0.82**	**2.00±0.04**
	4	1.19±0.24	1.17±0.05	1.02±0.17	1.09±0.05	0.78±0.13	0.64±0.44	1.44±0.52
	5	1.01±0.08	1.12±0.06	1.07±0.02	1.24±0.02	0.84±0.02	0.88±0.11	0.80±0.03
	6	1.06±0.01	1.40±0.08	1.03±0.04	1.15±0.01	0.98±0.01	0.66±0.42	2.14±0.05
	7	N/A	1.23±0.11	0.81±0.12	1.20±0.25	N/A	0.70±0.12	0.51±0.14
	8	0.93±0.03	0.90±0.02	1.02±0.00	0.87±0.01	1.18±0.02	1.23±0.25	0.73±0.01

Table 4 (continued)

Locus	Isoform	Root	1 Week Seedling	2Week Seedling	4 Week Rosette	Flower Buds	Mature Flowers	Siliques
RS31a	1	1.17±0.33	0.85±0.05	0.94±0.09	1.12±0.02	0.54±0.01	0.54±0.01	**1.57±0.07**
	2	0.74±0.24	0.98±0.27	**2.98±0.23**	1.12±0.10	1.27±0.17	1.27±0.17	0.97±0.04
	3	0.93±0.05	**2.24±0.88**	1.18±0.92	1.19±0.39	**0.47±0.06**	**0.47±0.06**	1.22±0.46
	4	**1.63±0.42**	0.97±0.07	0.96±0.12	0.68±0.02	1.44±0.05	1.44±0.05	0.53±0.02
SR41	1	**1.75±1.13**	0.91±0.11	0.66±0.32	1.00±0.07	**1.55±0.38**	**2.63±1.57**	1.23±0.01
	2	1.00±0.85	N/A	W	N/A	2.67±1.00	N/A	N/A
	3	**1.56±0.91**	1.12±0.20	**0.48±0.45**	0.75±0.07	1.15±0.12	0.58±0.27	1.08±0.02
	4	0.97±0.37	0.92±0.13	0.75±0.28	N/A	0.71±0.01	N/A	0.80±0.01
	5	N/A	N/A	0.97±0.21	0.95±0.03	0.70±0.11	0.95±0.48	R
	6	0.94±0.56	N/A	0.63±0.41	1.48±0.20	0.78±0.13	**0.12±0.04**	0.63±0.08
	7	1.00±0.04	1.00±0.01	1.14±0.15	0.99±0.01	1.02±0.01	0.77±0.35	0.92±0.01
SR33	1	**1.74±0.05**	0.72±0.10	**2.06±0.29**	**1.76±0.62**	1.00±0.29	**3.76±1.16**	1.28±0.23
	2	**0.36±0.02**	0.66±0.02	**3.75±1.02**	1.24±0.42	0.59±0.12	**3.77±2.39**	0.89±0.01
	3	**0.46±0.03**	0.94±0.11	0.88±0.13	1.01±0.12	0.63±0.29	**0.31±0.52**	1.05±0.09
	4	N/A	1.02±0.12	0.61±0.48	1.00±0.15	0.67±0.26	**4.44±1.54**	0.91±0.03
	5	**0.38±0.01**	0.92±0.10	**2.12±1.15**	1.26±0.17	**2.54±2.40**	**6.07±1.05**	1.30±0.23
	6	1.14±0.21	1.18±0.01	1.19±0.35	1.04±0.06	1.02±0.15	**3.67±1.41**	0.74±0.17
	7	N/A	1.00±0.08	1.29±0.36	1.28±0.03	**1.71±0.20**	**3.62±1.32**	1.15±0.17
	8	1.33±0.01	1.01±0.00	1.09±0.10	0.94±0.05	1.00±0.01	**6.32±5.46**	1.06±0.03
RS40	1	0.96±0.06	0.98±0.04	0.73±0.03	0.85±0.04	1.47±0.01	1.06±0.19	**2.16±0.06**
	2	**1.54±0.12**	0.69±0.00	0.92±0.19	**1.69±0.01**	1.37±0.09	1.08±0.64	0.80±0.05
	3	1.48±0.12	0.97±0.20	0.76±0.23	1.20±0.18	1.41±0.04	N/A±N/A	1.04±0.15
	4	**0.39±0.11**	1.02±0.02	1.03±0.05	0.98±0.03	**1.64±0.09**	1.04±0.06	1.33±0.18
	5	0.94±0.09	0.81±0.11	1.15±0.14	1.20±0.01	**1.57±0.27**	**1.95±0.94**	N/A±N/A
	6	**1.64±0.17**	0.91±0.04	0.81±0.01	1.12±0.06	1.18±0.05	**1.53±0.99**	0.56±0.03
	7	0.73±0.12	1.02±0.07	1.14±0.04	**1.64±0.06**	0.79±0.02	1.41±0.58	N/A±N/A
	8	0.87±0.00	1.05±0.02	1.12±0.03	0.88±0.03	0.82±0.00	0.72±0.25	0.52±0.06
SC35	1	0.63±0.26	**2.03±0.87**	N/A	N/A	N/A	N/A	N/A
	2	0.70±0.06	0.76±0.29	N/A	1.10±0.18	0.98±0.31	N/A±N/A	1.05±0.52
	3	1.40±0.06	0.60±0.17	0.80±0.15	1.02±0.11	0.86±0.11	1.04±0.08	0.96±0.11
	4	**0.50±0.04**	1.49±0.10	0.63±0.07	1.00±0.15	0.56±0.03	1.01±0.07	0.76±0.13
	5	1.17±0.02	0.99±0.01	1.04±0.02	1.00±0.00	1.05±0.02	1.00±0.01	1.11±0.07

Table 4 (continued)

Locus	Isoform	Root	1 Week Seedling	2Week Seedling	4 Week Rosette	Flower Buds	Mature Flowers	Siliques
SR33	1	0.66±0.09	N/A	N/A	N/A	N/A	1.00±0.12	1.06±0.07
	3	0.90±0.16	1.08±0.34	1.05±0.04	0.98±0.05	0.78±0.11	1.02±0.07	0.54±0.02
	4	1.06±0.03	0.78±0.15	1.06±0.07	0.93±0.12	0.89±0.21	1.00±0.05	0.80±0.02
	5	1.06±0.02	1.02±0.02	0.99±0.01	1.00±0.02	1.01±0.02	0.99±0.01	1.06±0.02
SCL30a	1	0.90±0.20	0.67±0.11	**1.98±0.19**	**1.50±0.03**	0.84±0.04	**0.34±0.08**	**2.29±1.89**
	2	0.92±0.13	1.55±1.02	R	1.29±0.13	**1.96±0.24**	**0.28±0.09**	N/A±N/A
	3	N/A	1.01±0.20	1.07±0.32	**2.86±0.67**	0.84±0.16	N/A	0.41±0.57
	4	N/A	1.10±0.19	1.40±0.21	**1.75±0.41**	0.83±0.03	N/A	R
	5	1.10±0.90	0.99±0.07	1.02±0.25	**1.56±0.10**	0.65±0.09	0.54±0.14	**1.93±1.36**
	6	N/A	1.15±0.12	1.00±0.31	1.47±0.11	0.68±0.17	**0.44±0.18**	1.05±1.60
	7	0.99±0.06	0.99±0.01	0.98±0.02	0.89±0.00	1.05±0.01	0.94±0.28	**2.01±1.31**
RS31	1	**2.07±0.81**	N/A	0.91±0.83	0.41±0.10	N/A	N/A	R
	2	1.35±0.75	1.23±0.04	1.01±0.01	0.68±0.01	0.68±0.00	**1.59±0.08**	**2.41±0.26**
	3	**2.38±1.91**	**2.37±0.98**	1.09±0.17	0.98±0.16	0.54±0.14	R	R
	4	1.18±0.10	**1.65±0.13**	1.09±0.03	1.26±0.11	0.69±0.06	R	0.72±0.03
	5	1.00±1.11	0.96±0.01	0.99±0.01	1.08±0.01	1.15±0.02	0.81±0.01	0.82±0.01

Table 5: The number of isoforms with altered abundance ratios in the *rdd* mutant compared to the wild-type across the multiple organ types and developmental stages.

Locus	8 Day Root	8 Day Seedling	2 Week Seedling	4 Week Aerial	Flower Organ	Silique
SR30	0	0	0	3	2	2
SR1	3	1	1	2	2	2
SR34a	3	0	0	1	0	1
SR34b	0	0	0	2	0	0
RS40	3	2	0	2	1	5
RS31a	2	1	0	1	4	4
SR41	2	0	5	2	5	4
RSZ32	5	2	5	3	5	0
RSZ33	5	1	1	2	5	3
SC35	4	4	1	0	1	0
SR33	2	1	1	1	1	0
SCL30a	0	3	2	6	3	4
RS31	3	2	x	3	3	2

Table 6. Illumina sequence reads were aligned to TAIR8 genomic reference sequences using TopHat v1.011 and Bowtie 1.0 and the total number of mapable reads are summarized in the above table. The total number of reads aligning to annotated genes was broken down into whether the reads localized to an exon, exon-exon junction or an intron. Intergenic reads are defined as any read that did not fall within the boundaries of an annotated gene.

SRA Identifier	Genotype	Exon Reads	Intron Reads	Junction Reads	Intergenic Reads	Total Aligned Reads
SRX002554	Wild-Type	9861345	453851	393042	757531	11465769
SRX002555	*met1*	8660583	486776	340512	105811	9593682
SRX002556	*ddc*	11060857	608466	451542	178829	12299694
SRX002557	*rdd*	9197366	512891	367369	61078	10138704

Table 7. AS events detected using aligned Illumina sequence data and custom perl scripts for wild-type, *met1, ddc*, and *rdd* backgrounds.

AS Type	**Wild-type**	***met1***	***ddc***	***rdd***
AltD	682	1039	1299	1181
AltA	502	806	939	880
AltP	143	243	268	259
ES	62	125	134	127
IR	370	4836	5333	6147
Total	**1759**	**7049**	**7973**	**8594**

Table 8. Novel AS events associated with changes in DNA methylation within the coding regions of genes between the wild-type and the *met1, ddc,* and *rdd* mutants.

Δ	***met1***			***ddc***			***rdd***		
	No As Change	**Lose AS**	**Gain AS**	**No As Change**	**Lose AS**	**Gain AS**	**No As Change**	**Lose AS**	**Gain AS**
No Meth Change	3602	91	257	2082	67	268	17880	645	2102
Meth Up	2082	67	268	2160	91	389	4193	192	809
Meth down	11921	645	2102	10297	429	2159	8987	390	1655

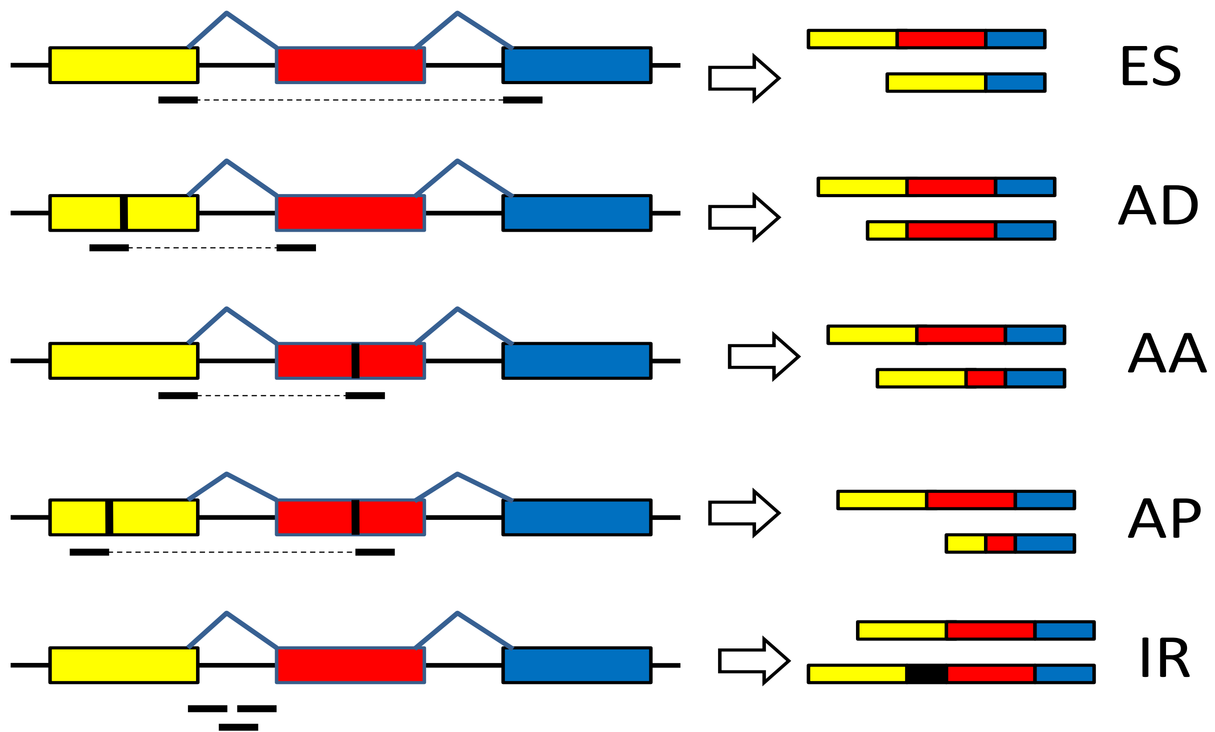


Figure 1. Computational detection of AS splicing events using Illumina sequence reads. Types of events include exon skipping (ES), alternative donor (AltD), alternative acceptor (AltA), alternative position (AltP), and intron retention (IR). Exons are drawn as boxes and introns are shown as a solid horizontal line. Splicing events are indicated by raised lines connecting exons. Bent blue lines indicate the constitutively spliced form. Illumina reads are shown as black bars and junctions spanning reads are indicated by dotted lines. Perl scripts were designed to detect each of these AS types based on spliced reads or intron spanning reads.

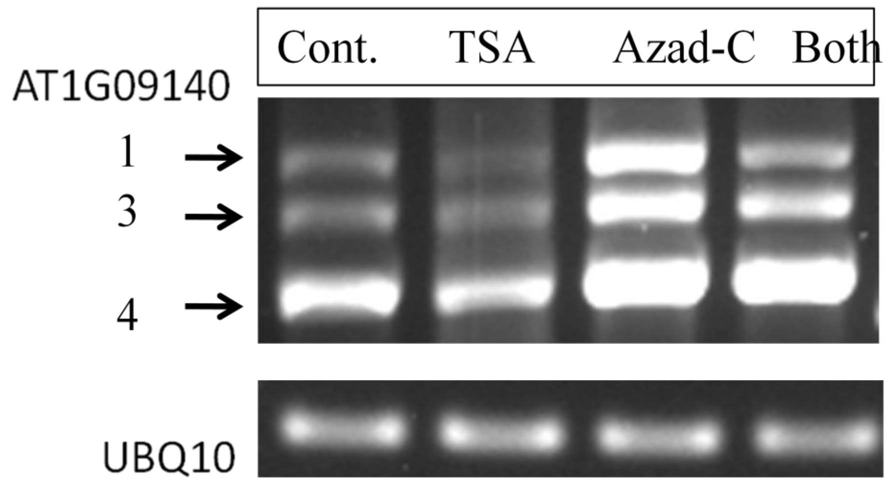


Figure 2: RT-PCR gel of *SR30* for seedlings treated with TSA, Azad-C, or both. Three isoforms were identified under these conditions. The expression level of *UBQ10* is shown for reference to indicate cDNA amounts.

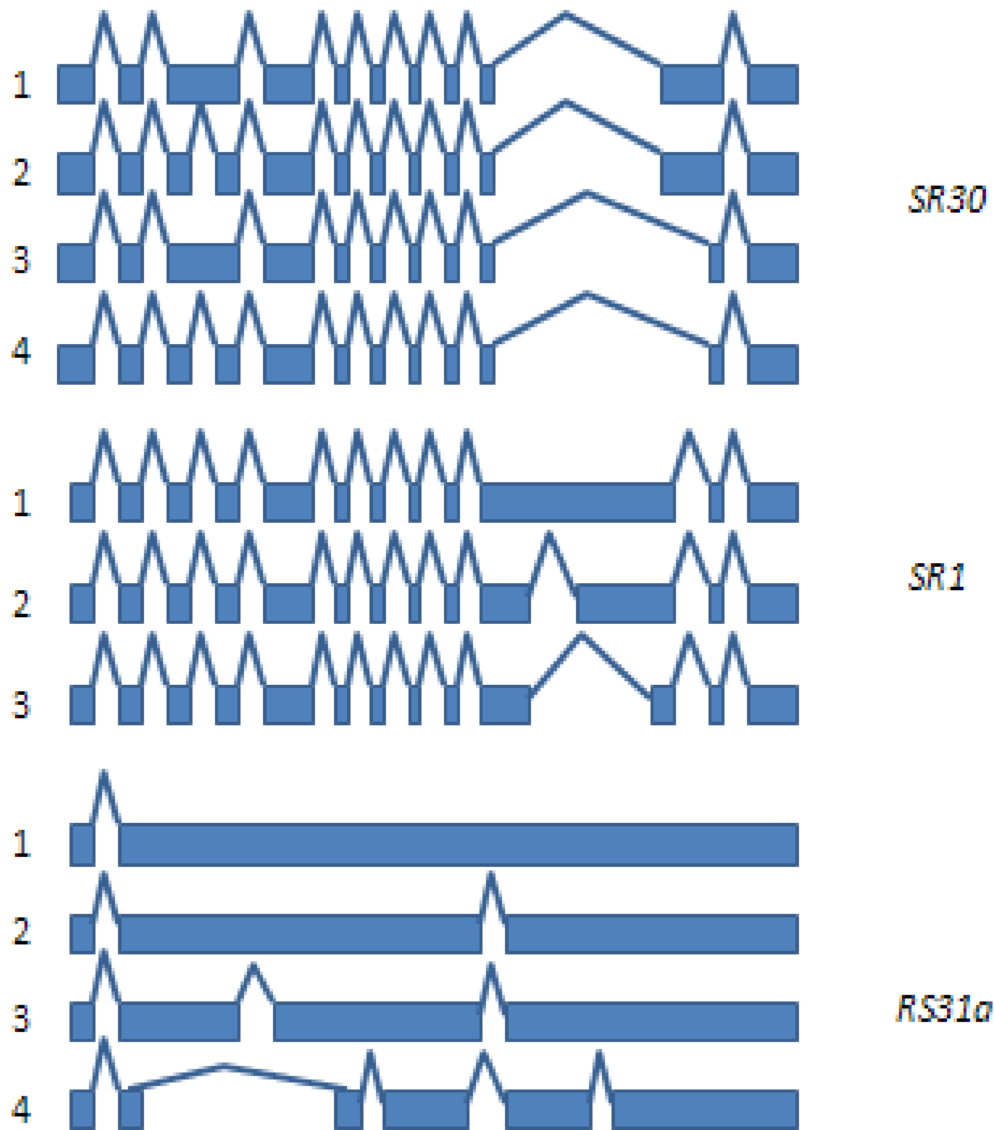


Figure 3: Gene structure diagrams for *SR30, SR1, RS31a* for each of the identified isoforms. Exons are drawn as boxes and introns are represented as bent lines.

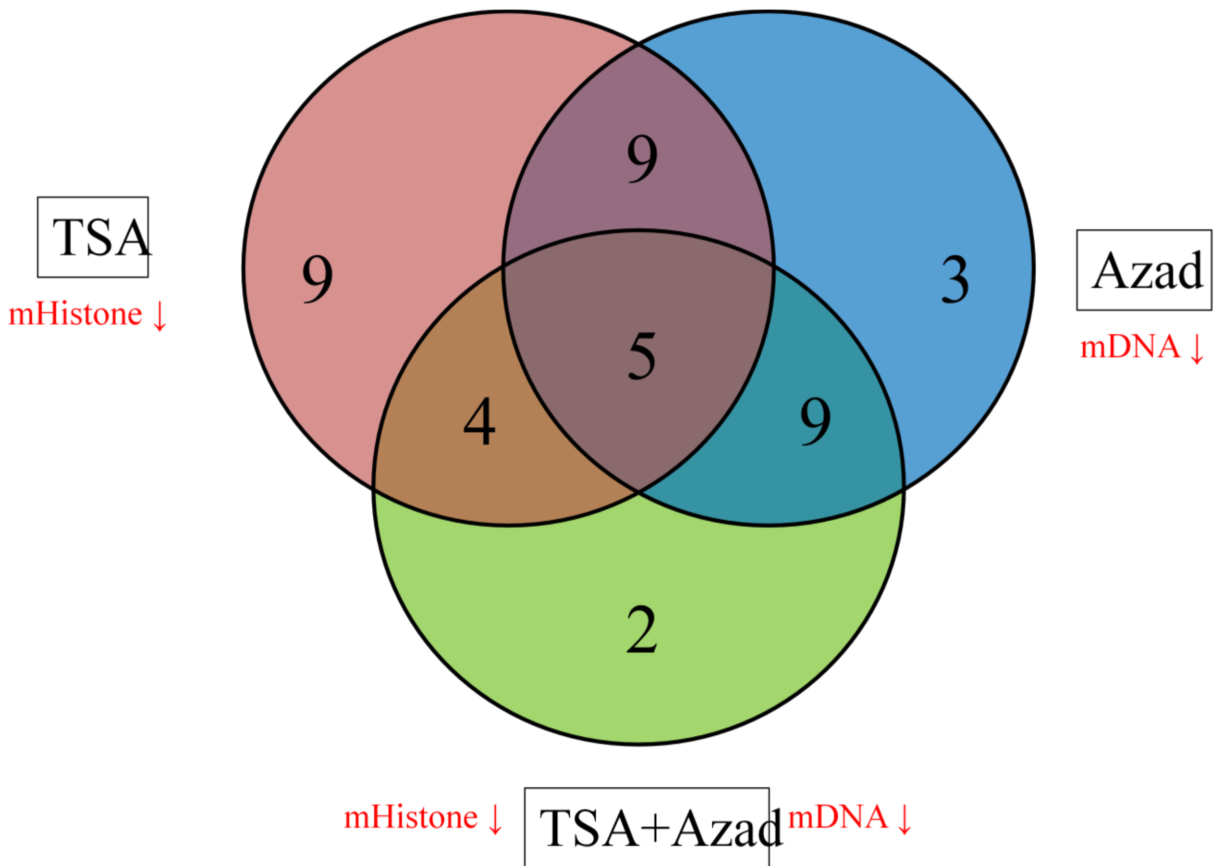


Figure 4: Venn diagram showing isoforms that demonstrated altered abundance in response to TSA, Azad-C or both together. Overlapping sections indicate isoforms that were affected in the same way between two or more treatments.

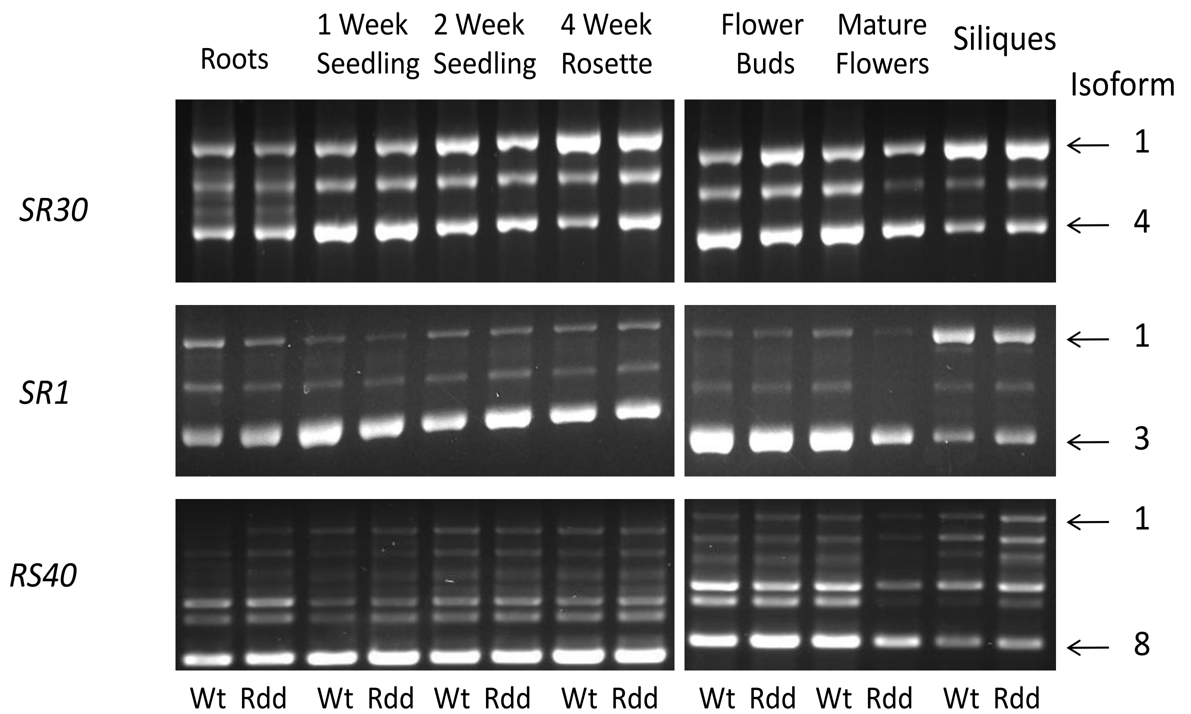


Figure 5: RT-PCR gel for wild-type and *rdd* mutant genotypes. Sampled tissues included Roots, 1 and 2 week old whole seedlings, 4 week old rosette leaves, immature flower buds, mature open flowers, and siliques.

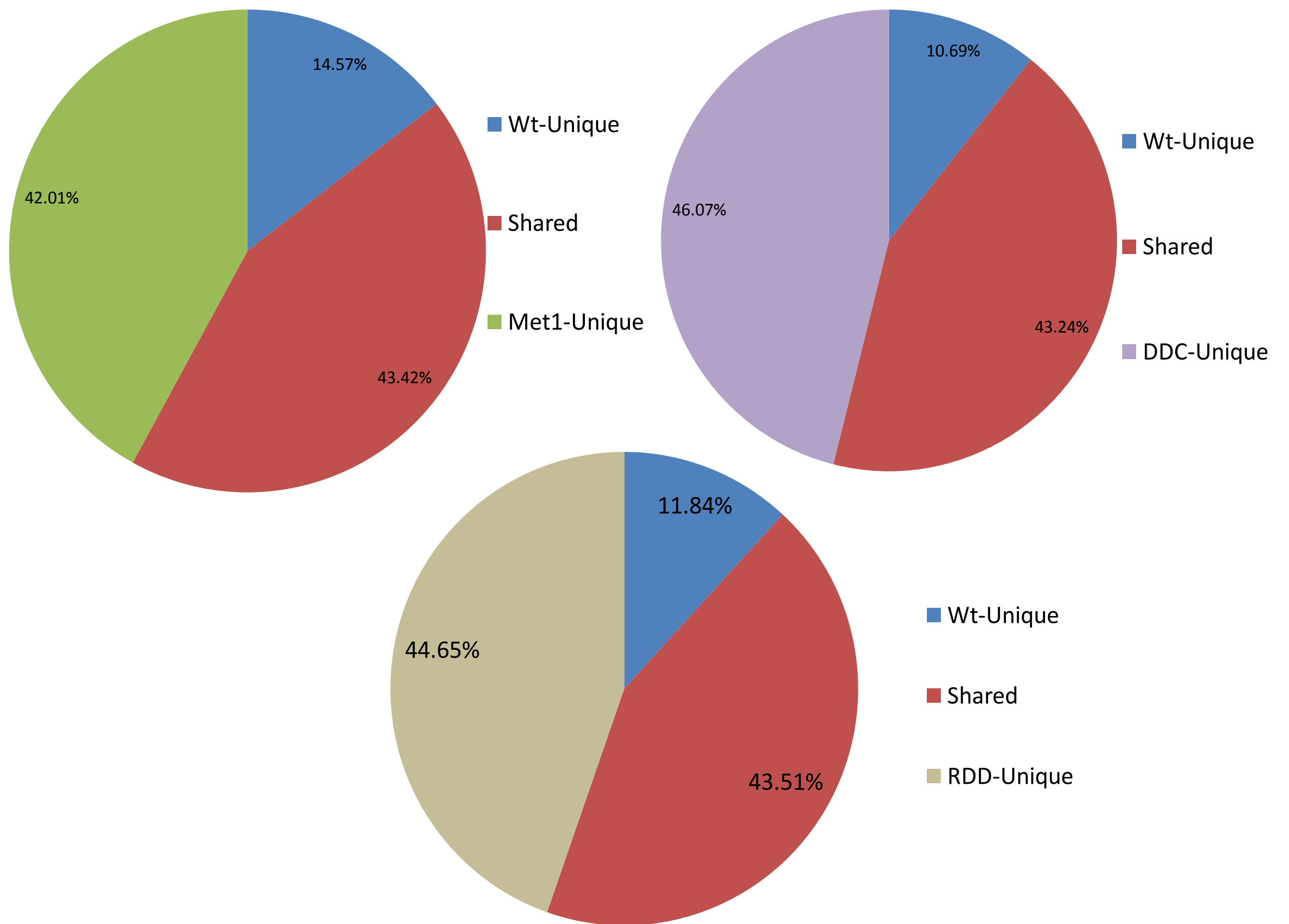


Figure 6: Pair-wise comparisons of shared and unique splicing events between the wild-type and *met1*, *ddc*, and *rdd* mutants.

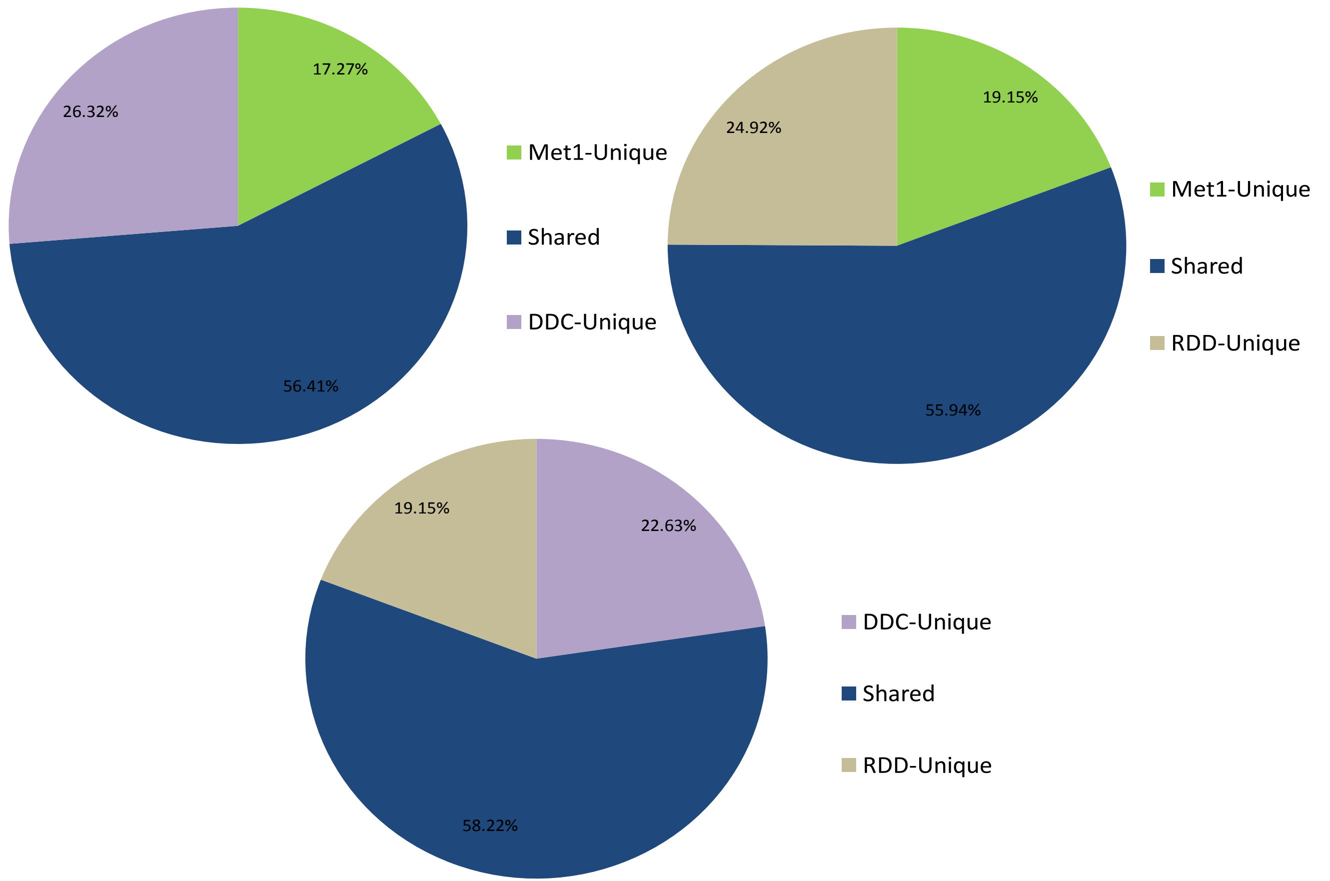
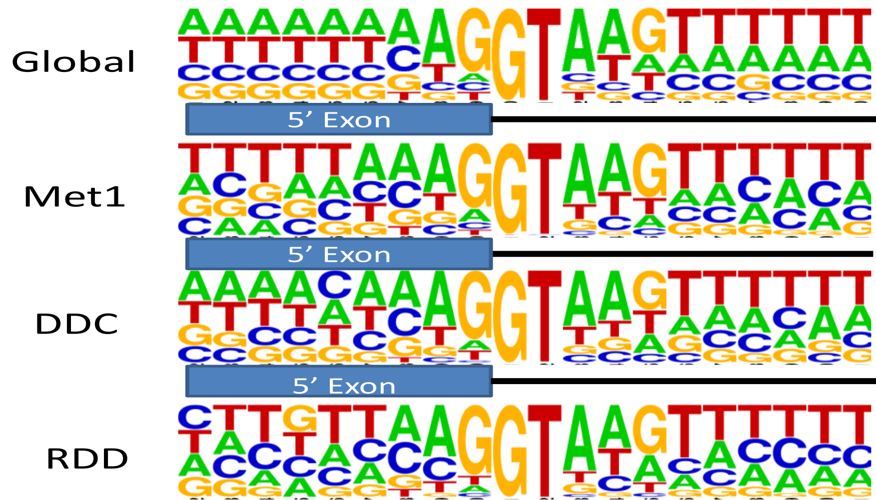


Figure 7: Pair-wise comparisons of shared and unique splicing events between *met1, ddc*, and *rdd*.

5’ SS



3’ SS

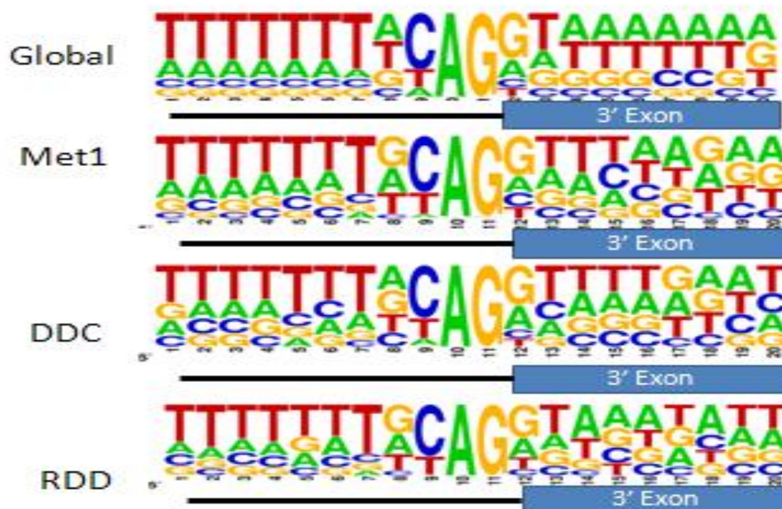


Figure 8: Splice site sequence logos. Nucleotides surrounding the 5’ and 3’ splice sites of uniquely retained introns in each mutant genotype.

BIBLIOGRAPHY

1. Luco RF, Pan Q, Tominaga K, Blencowe BJ, Pereira-Smith OM, Misteli T. Regulation of alternative splicing by histone modifications. Science. 2010 Feb 19;327(5968):996-1000.

2. Palusa SG, Ali GS, Reddy AS. Alternative splicing of pre-mRNAs of arabidopsis serine/arginine-rich proteins: Regulation by hormones and stresses. Plant J. 2007 Mar;49(6):1091-107.

3. Ruiz-Garcia L, Cervera MT, Martinez-Zapater JM. DNA methylation increases throughout arabidopsis development. Planta. 2005 Oct;222(2):301-6.

4. Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH, et al. Highly integrated single-base resolution maps of the epigenome in arabidopsis. Cell. 2008 May 2;133(3):523-36.

5. Aguilera A. Cotranscriptional mRNP assembly: From the DNA to the nuclear pore. Curr Opin Cell Biol. 2005 Jun;17(3):242-50.

6. Bird G, Zorio DA, Bentley DL. RNA polymerase II carboxy-terminal domain phosphorylation is required for cotranscriptional pre-mRNA splicing and 3'-end formation. Mol Cell Biol. 2004 Oct;24(20):8963-9.

7. Bittencourt D, Dutertre M, Sanchez G, Barbier J, Gratadou L, Auboeuf D. Cotranscriptional splicing potentiates the mRNA production from a subset of estradiol-stimulated genes. Mol Cell Biol. 2008 Sep;28(18):5811-24.

8. de la Mata M, Kornblihtt AR. RNA polymerase II C-terminal domain mediates regulation of alternative splicing by SRp20. Nat Struct Mol Biol. 2006 Nov;13(11):973-80.

9. Gornemann J, Kotovic KM, Hujer K, Neugebauer KM. Cotranscriptional spliceosome assembly occurs in a stepwise fashion and requires the cap binding complex. Mol Cell. 2005 Jul 1;19(1):53-63.

10. Gromak N, Talotti G, Proudfoot NJ, Pagani F. Modulating alternative splicing by cotranscriptional cleavage of nascent intronic RNA. RNA. 2008 Feb;14(2):359-66.

11. Kotovic KM, Lockshon D, Boric L, Neugebauer KM. Cotranscriptional recruitment of the U1 snRNP to intron-containing genes in yeast. Mol Cell Biol. 2003 Aug;23(16):5768-79.

12. Lacadie SA, Rosbash M. Cotranscriptional spliceosome assembly dynamics and the role of U1 snRNA:5'ss base pairing in yeast. Mol Cell. 2005 Jul 1;19(1):65-75.

13. Listerman I, Sapra AK, Neugebauer KM. Cotranscriptional coupling of splicing factor recruitment and precursor messenger RNA splicing in mammalian cells. Nat Struct Mol Biol. 2006 Sep;13(9):815-22.

14. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-seq. Nat Methods. 2008 Jul;5(7):621-8.

15. Kornblihtt AR. Coupling transcription and alternative splicing. Adv Exp Med Biol. 2007;623:175-89.

16. Allemand E, Batsche E, Muchardt C. Splicing, transcription, and chromatin: A menage a trois. Curr Opin Genet Dev. 2008 Apr;18(2):145-51.

17. Farlow A, Meduri E, Dolezal M, Hua L, Schlotterer C. Nonsense-mediated decay enables intron gain in drosophila. PLoS Genet. 2010 Jan 22;6(1):e1000819.

18. Palusa SG, Reddy AS. Extensive coupling of alternative splicing of pre-mRNAs of serine/arginine (SR) genes with nonsense-mediated decay. New Phytol. 2010 Jan;185(1):83-9.

19. Schoning JC, Streitner C, Meyer IM, Gao Y, Staiger D. Reciprocal regulation of glycine-rich RNA-binding proteins via an interlocked feedback loop coupling alternative splicing to nonsense-mediated decay in arabidopsis. Nucleic Acids Res. 2008 Dec;36(22):6977-87.

20. Mendell JT, Sharifi NA, Meyers JL, Martinez-Murillo F, Dietz HC. Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. Nat Genet. 2004 Oct;36(10):1073-8.

21. Filichkin SA, Priest HD, Givan SA, Shen R, Bryant DW, Fox SE, et al. Genome-wide mapping of alternative splicing in arabidopsis thaliana. Genome Res. 2010 Jan;20(1):45-58.

22. Blekhman R, Marioni JC, Zumbo P, Stephens M, Gilad Y. Sex-specific and lineage-specific alternative splicing in primates. Genome Res. 2010 Feb;20(2):180-9.

23. Sultan M, Schulz MH, Richard H, Magen A, Klingenhoff A, Scherf M, et al. A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. Science. 2008 Aug 15;321(5891):956-60.

24. Telonis-Scott M, Kopp A, Wayne ML, Nuzhdin SV, McIntyre LM. Sex-specific splicing in drosophila: Widespread occurrence, tissue specificity and evolutionary conservation. Genetics. 2009 Feb;181(2):421-34.

25. Barbazuk WB, Fu Y, McGinnis KM. Genome-wide analyses of alternative splicing in plants: Opportunities and challenges. Genome Res. 2008 Sep;18(9):1381-92.

26. Wang BB, Brendel V. Genomewide comparative analysis of alternative splicing in plants. Proc Natl Acad Sci U S A. 2006 May 2;103(18):7175-80.

27. Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nat Genet. 2008 Dec;40(12):1413-5.

28. Bhasi A, Philip P, Sreedharan VT, Senapathy P. AspAlt: A tool for inter-database, inter-genomic and user-specific comparative analysis of alternative transcription and alternative splicing in 46 eukaryotes. Genomics. 2009 Jul;94(1):48-54.

29. Keren H, Lev-Maor G, Ast G. Alternative splicing and evolution: Diversification, exon definition and function. Nat Rev Genet. 2010 May;11(5):345-55.

30. Barta A, Kalyna M, Lorkovic ZJ. Plant SR proteins and their functions. Curr Top Microbiol Immunol. 2008;326:83-102.

31. Lacadie SA, Rosbash M. Cotranscriptional spliceosome assembly dynamics and the role of U1 snRNA:5'ss base pairing in yeast. Mol Cell. 2005 Jul 1;19(1):65-75.

32. Das R, Yu J, Zhang Z, Gygi MP, Krainer AR, Gygi SP, et al. SR proteins function in coupling RNAP II transcription to pre-mRNA splicing. Mol Cell. 2007 Jun 22;26(6):867-81.

33. Perales R, Bentley D. "Cotranscriptionality": The transcription elongation complex as a nexus for nuclear transactions. Mol Cell. 2009 Oct 23;36(2):178-91.

34. Listerman I, Sapra AK, Neugebauer KM. Cotranscriptional coupling of splicing factor recruitment and precursor messenger RNA splicing in mammalian cells. Nat Struct Mol Biol. 2006 Sep;13(9):815-22.

35. Lynch KW. Cotranscriptional splicing regulation: It's not just about speed. Nat Struct Mol Biol. 2006 Nov;13(11):952-3.

36. Moldon A, Malapeira J, Gabrielli N, Gogol M, Gomez-Escoda B, Ivanova T, et al. Promoter-driven splicing regulation in fission yeast. Nature. 2008 Oct 16;455(7215):997-1000.

37. Andersson R, Enroth S, Rada-Iglesias A, Wadelius C, Komorowski J. Nucleosomes are well positioned in exons and carry characteristic histone modifications. Genome Res. 2009 Oct;19(10):1732-41.

38. Kolasinska-Zwierz P, Down T, Latorre I, Liu T, Liu XS, Ahringer J. Differential chromatin marking of introns and expressed exons by H3K36me3. Nat Genet. 2009 Mar;41(3):376-81.

39. Schor IE, Rascovan N, Pelisch F, Allo M, Kornblihtt AR. Neuronal cell depolarization induces intragenic chromatin modifications affecting NCAM alternative splicing. Proc Natl Acad Sci U S A. 2009 Mar 17;106(11):4325-30.

40. Nogues G, Kadener S, Cramer P, Bentley D, Kornblihtt AR. Transcriptional activators differ in their abilities to control alternative splicing. J Biol Chem. 2002 Nov 8;277(45):43110-4.

41. Tyagi A, Ryme J, Brodin D, Ostlund Farrants AK, Visa N. SWI/SNF associates with nascent pre-mRNPs and regulates alternative pre-mRNA processing. PLoS Genet. 2009 May;5(5):e1000470.

42. Batsche E, Yaniv M, Muchardt C. The human SWI/SNF subunit brm is a regulator of alternative splicing. Nat Struct Mol Biol. 2006 Jan;13(1):22-9.

43. Ameur A, Wetterbom A, Feuk L, Gyllensten U. Global and unbiased detection of splice junctions from RNA-seq data. Genome Biol. 2010 Mar 17;11(3):R34.

44. Marguerat S, Bahler J. RNA-seq: From technology to biology. Cell Mol Life Sci. 2010 Feb;67(4):569-79.

45. Nagalakshmi U, Waern K, Snyder M. RNA-seq: A method for comprehensive transcriptome analysis. Curr Protoc Mol Biol. 2010 Jan;Chapter 4:Unit 4.11.1-13.

46. Wang L, Xi Y, Yu J, Dong L, Yen L, Li W. A statistical method for the detection of alternative splicing using RNA-seq. PLoS One. 2010 Jan 8;5(1):e8529.

47. Levin JZ, Berger MF, Adiconis X, Rogov P, Melnikov A, Fennell T, et al. Targeted next-generation sequencing of a cancer transcriptome enhances detection of sequence variants and novel fusion transcripts. Genome Biol. 2009;10(10):R115.

48. Kertesz S, Kerenyi Z, Merai Z, Bartos I, Palfy T, Barta E, et al. Both introns and long 3'-UTRs operate as cis-acting elements to trigger nonsense-mediated decay in plants. Nucleic Acids Res. 2006;34(21):6147-57.

49. Werneke JM, Chatfield JM, Ogren WL. Alternative mRNA splicing generates the two ribulosebisphosphate carboxylase/oxygenase activase polypeptides in spinach and arabidopsis. Plant Cell. 1989 Aug;1(8):815-25.

50. Macknight R, Duroux M, Laurie R, Dijkwel P, Simpson G, Dean C. Functional significance of the alternative transcript processing of the arabidopsis floral promoter FCA. Plant Cell. 2002 Apr;14(4):877-88.

51. Gardner LB. Nonsense-mediated RNA decay regulation by cellular stress: Implications for tumorigenesis. Mol Cancer Res. 2010 Mar;8(3):295-308.

52. Severing EI, van Dijk AD, Stiekema WJ, van Ham RC. Comparative analysis indicates that alternative splicing in plants has a limited role in functional expansion of the proteome. BMC Genomics. 2009 Apr 9;10:154.

53. Isken O, Maquat LE. The multiple lives of NMD factors: Balancing roles in gene and genome regulation. Nat Rev Genet. 2008 Aug 5.

54. Barberan-Soler S, Lambert NJ, Zahler AM. Global analysis of alternative splicing uncovers developmental regulation of nonsense-mediated decay in C. elegans. RNA. 2009 Sep;15(9):1652-60.

55. Zhang XN, Mount SM. Two alternatively spliced isoforms of the arabidopsis SR45 protein have distinct roles during normal plant development. Plant Physiol. 2009 Jul;150(3):1450-8.

56. Kalyna M, Lopato S, Barta A. Ectopic expression of atRSZ33 reveals its function in splicing and causes pleiotropic changes in development. Mol Biol Cell. 2003 Sep;14(9):3565-77.

57. Shepard PJ, Hertel KJ. The SR protein family. Genome Biol. 2009;10(10):242.

58. Buvoli M, Buvoli A, Leinwand LA. Interplay between exonic splicing enhancers, mRNA processing, and mRNA surveillance in the dystrophic mdx mouse. PLoS One. 2007 May 9;2(5):e427.

59. Zhong XY, Wang P, Han J, Rosenfeld MG, Fu XD. SR proteins in vertical integration of gene expression from transcription to RNA processing to translation. Mol Cell. 2009 Jul 10;35(1):1-10.

60. Li X, Niu T, Manley JL. The RNA binding protein RNPS1 alleviates ASF/SF2 depletion-induced genomic instability. RNA. 2007 Dec;13(12):2108-15.

61. Xiao R, Sun Y, Ding JH, Lin S, Rose DW, Rosenfeld MG, et al. Splicing regulator SC35 is essential for genomic stability and cell proliferation during mammalian organogenesis. Mol Cell Biol. 2007 Aug;27(15):5393-402.

62. Li X, Manley JL. Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. Cell. 2005 Aug 12;122(3):365-78.

63. Komashko VM, Farnham PJ. 5-azacytidine treatment reorganizes genomic histone modification patterns. Epigenetics. 2010 Apr 4;5(3).

64. Ikeda Y, Kinoshita T. DNA demethylation: A lesson from the garden. Chromosoma. 2009 Feb;118(1):37-41.

65. Gehring M, Bubb KL, Henikoff S. Extensive demethylation of repetitive elements during seed development underlies gene imprinting. Science. 2009 Jun 12;324(5933):1447-51.

66. Zhang M, Kimatu JN, Xu K, Liu B. DNA cytosine methylation in plant development. J Genet Genomics. 2010 Jan;37(1):1-12.

67. Cokus SJ, Feng S, Zhang X, Chen Z, Merriman B, Haudenschild CD, et al. Shotgun bisulphite sequencing of the arabidopsis genome reveals DNA methylation patterning. Nature. 2008 Mar 13;452(7184):215-9.

68. Chang S, Pikaard CS. Transcript profiling in arabidopsis reveals complex responses to global inhibition of DNA methylation and histone deacetylation. J Biol Chem. 2005 Jan 7;280(1):796-804.

69. Penterman J, Zilberman D, Huh JH, Ballinger T, Henikoff S, Fischer RL. DNA demethylation in the arabidopsis genome. Proc Natl Acad Sci U S A. 2007 Apr 17;104(16):6752-7.

70. Zilberman D, Gehring M, Tran RK, Ballinger T, Henikoff S. Genome-wide analysis of arabidopsis thaliana DNA methylation uncovers an interdependence between methylation and transcription. Nat Genet. 2007 Jan;39(1):61-9.

71. Zilberman D. The evolving functions of DNA methylation. Curr Opin Plant Biol. 2008 Oct;11(5):554-9.

72. Causevic A, Gentil MV, Delaunay A, El-Soud WA, Garcia Z, Pannetier C, et al. Relationship between DNA methylation and histone acetylation levels, cell redox and cell differentiation states in sugarbeet lines. Planta. 2006 Sep;224(4):812-27.

73. Steward N, Ito M, Yamaguchi Y, Koizumi N, Sano H. Periodic DNA methylation in maize nucleosomes and demethylation by environmental stress. J Biol Chem. 2002 Oct 4;277(40):37741-6.

74. Aceituno FF, Moseyko N, Rhee SY, Gutierrez RA. The rules of gene expression in plants: Organ identity and gene body methylation are key factors for regulation of gene expression in arabidopsis thaliana. BMC Genomics. 2008 Sep 23;9:438.

75. Zhang X, Yazaki J, Sundaresan A, Cokus S, Chan SW, Chen H, et al. Genome-wide high-resolution mapping and functional analysis of DNA methylation in arabidopsis. Cell. 2006 Sep 22;126(6):1189-201.

76. Chang S, Pikaard CS. Transcript profiling in arabidopsis reveals complex responses to global inhibition of DNA methylation and histone deacetylation. J Biol Chem. 2005 Jan 7;280(1):796-804.

77. Strahl BD, Allis CD. The language of covalent histone modifications. Nature. 2000 Jan 6;403(6765):41-5.

78. Vaissiere T, Herceg Z. Histone code in the cross-talk during DNA damage signaling. Cell Res. 2010 Feb;20(2):113-5.

79. Yang W, Jiang D, Jiang J, He Y. A plant-specific histone H3 lysine-4 demethylase represses the floral transition in arabidopsis. Plant J. 2010 Feb 24.

80. Bonisch C, Nieratschker SM, Orfanos NK, Hake SB. Chromatin proteomics and epigenetic regulatory circuits. Expert Rev Proteomics. 2008 Feb;5(1):105-19.

81. Talbert PB, Henikoff S. Histone variants - ancient wrap artists of the epigenome. Nat Rev Mol Cell Biol. 2010 Apr;11(4):264-75.

82. Kobor MS, Lorincz MC. H2A.Z and DNA methylation: Irreconcilable differences. Trends Biochem Sci. 2009 Apr;34(4):158-61.

83. Zilberman D, Coleman-Derr D, Ballinger T, Henikoff S. Histone H2A.Z and DNA methylation are mutually antagonistic chromatin marks. Nature. 2008 Nov 6;456(7218):125-9.

84. Eckardt NA. A double lock on polyploidy-associated epigenetic gene silencing. Plant Cell. 2010 Jan;22(1):3.

85. Feng S, Cokus SJ, Zhang X, Chen PY, Bostick M, Goll MG, et al. Conservation and divergence of methylation patterning in plants and animals. Proc Natl Acad Sci U S A. 2010 Apr 15.

86. Choi JK, Bae JB, Lyu J, Kim TY, Kim YJ. Nucleosome deposition and DNA methylation at coding region boundaries. Genome Biol. 2009;10(9):R89.

87. Kumar SV, Wigge PA. H2A.Z-containing nucleosomes mediate the thermosensory response in arabidopsis. Cell. 2010 Jan 8;140(1):136-47.

88. Zhang X, Bernatavichute YV, Cokus S, Pellegrini M, Jacobsen SE. Genome-wide analysis of mono-, di- and trimethylation of histone H3 lysine 4 in arabidopsis thaliana. Genome Biol. 2009;10(6):R62.

89. Fu X, Den S, Su G, Zeng Q, Shi S. **Isolating high-quality RNA from mangroves without liquid nitrogen**. PMBR. 2004;22, Number(2):1572-9818.

90. Trapnell C, Pachter L, Salzberg SL. TopHat: Discovering splice junctions with RNA-seq. Bioinformatics. 2009 May 1;25(9):1105-11.

91. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009;10(3):R25.

92. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: A sequence logo generator. Genome Res. 2004 Jun;14(6):1188-90.

93. Schor IE, Kornblihtt AR. Playing inside the genes: Intragenic histone acetylation after membrane depolarization of neural cells opens a path for alternative splicing regulation. Commun Integr Biol. 2009 Jul;2(4):341-3.

94. Hsieh TF, Ibarra CA, Silva P, Zemach A, Eshed-Williams L, Fischer RL, et al. Genome-wide demethylation of arabidopsis endosperm. Science. 2009 Jun 12;324(5933):1451-4.

95. Kornblihtt AR, de la Mata M, Fededa JP, Munoz MJ, Nogues G. Multiple links between transcription and splicing. RNA. 2004 Oct;10(10):1489-98.

96. Mathieu O, Reinders J, Caikovski M, Smathajitt C, Paszkowski J. Transgenerational stability of the arabidopsis epigenome is coordinated by CG methylation. Cell. 2007 Sep 7;130(5):851-62.

97. Staiger D, Zecca L, Wieczorek Kirk DA, Apel K, Eckstein L. The circadian clock regulated RNA-binding protein AtGRP7 autoregulates its expression by influencing alternative splicing of its own pre-mRNA. Plant J. 2003 Jan;33(2):361-71.

98. Allo M, Buggiano V, Fededa JP, Petrillo E, Schor I, de la Mata M, et al. Control of alternative splicing through siRNA-mediated transcriptional gene silencing. Nat Struct Mol Biol. 2009 Jul;16(7):717-24.

99. Zibetti C, Adamo A, Binda C, Forneris F, Toffolo E, Verpelli C, et al. Alternative splicing of the histone demethylase LSD1/KDM1 contributes to the modulation of neurite morphogenesis in the mammalian nervous system. J Neurosci. 2010 Feb 17;30(7):2521-32.

100. Jablonski JA, Caputi M. Role of cellular RNA processing factors in human immunodeficiency virus type 1 mRNA metabolism, replication, and infectivity. J Virol. 2009 Jan;83(2):981-92.

101. Baubec T, Pecinka A, Rozhon W, Mittelsten Scheid O. Effective, homogeneous and transient interference with cytosine methylation in plant genomic DNA by zebularine. Plant J. 2009 Feb;57(3):542-54.

102. Yaish MW, Peng M, Rothstein SJ. AtMBD9 modulates arabidopsis development through the dual epigenetic pathways of DNA methylation and histone acetylation. Plant J. 2009 Jul;59(1):123-35.

103. Lenasi T, Barboric M. P-TEFb stimulates transcription elongation and pre-mRNA splicing through multilateral mechanisms. RNA Biol. 2010 Mar 29;7(2).

104. Tyagi A, Ryme J, Brodin D, Ostlund Farrants AK, Visa N. SWI/SNF associates with nascent pre-mRNPs and regulates alternative pre-mRNA processing. PLoS Genet. 2009 May;5(5):e1000470.

105. Ruiz-Garcia L, Cervera MT, Martinez-Zapater JM. DNA methylation increases throughout arabidopsis development. Planta. 2005 Oct;222(2):301-6.

106. Iida K, Fukami-Kobayashi K, Toyoda A, Sakaki Y, Kobayashi M, Seki M, et al. Analysis of multiple occurrences of alternative splicing events in arabidopsis thaliana using novel sequenced full-length cDNAs. DNA Res. 2009 Jun;16(3):155-64.

107. Xiao YL, Smith SR, Ishmael N, Redman JC, Kumar N, Monaghan EL, et al. Analysis of the cDNAs of hypothetical genes on arabidopsis chromosome 2 reveals numerous transcript variants. Plant Physiol. 2005 Nov;139(3):1323-37.

108. Crotti LB, Horowitz DS. Exon sequences at the splice junctions affect splicing fidelity and alternative splicing. Proc Natl Acad Sci U S A. 2009 Nov 10;106(45):18954-9.

109. Yang JO, Kim WY, Bhak J. ssSNPTarget: Genome-wide splice-site single nucleotide polymorphism database. Hum Mutat. 2009 Dec;30(12):E1010-20.

110. Low KH, Lim C, Ko HW, Edery I. Natural variation in the splice site strength of a clock gene and species-specific thermal adaptation. Neuron. 2008 Dec 26;60(6):1054-67.

111. Shepard PJ, Hertel KJ. Conserved RNA secondary structures promote alternative splicing. RNA. 2008 Aug;14(8):1463-9.

112. Yang W, Li QZ. One parameter to describe the mechanism of splice sites competition. Biochem Biophys Res Commun. 2008 Apr 4;368(2):379-81.

113. Causevic A, Delaunay A, Ounnar S, Righezza M, Delmotte F, Brignolas F, et al. DNA methylating and demethylating treatments modify phenotype and cell wall differentiation state in sugarbeet cell lines. Planta. 2005;43(7):681.

114. Iida K, Seki M, Sakurai T, Satou M, Akiyama K, Toyoda T, et al. Genome-wide analysis of alternative pre-mRNA splicing in arabidopsis thaliana based on full-length cDNA sequences. Nucleic Acids Res. 2004 Sep 27;32(17):5096-103.

115. Wang J, Tian L, Madlung A, Lee HS, Chen M, Lee JJ, et al. Stochastic and epigenetic changes of gene expression in arabidopsis polyploids. Genetics. 2004 Aug;167(4):1961-73.

116. Zhang X, Shiu SH, Cal A, Borevitz JO. Global analysis of genetic, epigenetic and transcriptional polymorphisms in arabidopsis thaliana using whole genome tiling arrays. PLoS Genet. 2008 Mar 21;4(3):e1000032.

117. Scascitelli M, Cognet M, Adams KL. An interspecific plant hybrid shows novel changes in parental splice forms of genes for splicing factors. Genetics. 2010 Jan 25.