

- Data collection

Type	WT	<i>drm12</i>	<i>cmt2</i>	<i>cmt23</i>	<i>drm12cmt2</i>	<i>drm12cmt3</i>
RNA-seq	+	+	+	+	+	+
BS-seq	+	+	+	+	+	

Type	<i>drm12cmt23</i>	<i>suvh456</i>	<i>atxr56</i>	<i>mom1</i>	<i>atxr56drm12</i>	<i>atxr56mom1</i>
RNA-seq	+	+	+	+	+	+
BS-seq	+	+	+	+		

Col-0 histone modification:

H3K4me2, H3K4me3, H3K9Ac, H3K9me2, H3K18Ac, H3K23ac, H3K27me1, H3K27me3, H3K36me2, H3K36me3.

Mutant:

suvh456 (H3k9me2, H3K23ac)

drm12cmt23 (H3k9me2, H3K23ac)

atxr5/6 (H3K27me1)

met1 (H3K4me3, H3K9me2, H3K27me3)

- RNA seq data and BS seq data quality control

Sample	SRA	Raw data		Low quality trimmed		Adapter removed		Uniq Mapped Reads
		Reads	Bases	Reads	Bases	Reads	Bases	
WT_mRNA_rep1	SRR1005385	16.13M	0.81G	14.47M(89.7%)	0.69G(85.3%)	14.46M(89.6%)	0.69G(85.1%)	13.47M(93.2%)
WT_mRNA_rep2	SRR1005386	21.47M	1.09G	19.66M(91.6%)	0.96G(87.6%)	19.61M(91.3%)	0.96G(87.2%)	18.63M(95%)
drm12_mRNA_rep1	SRR1005387	17.99M	0.9G	16.14M(89.7%)	0.77G(85.4%)	16.13M(89.7%)	0.77G(85.2%)	15.35M(95.1%)
drm12_mRNA_rep2	SRR1005388	29.6M	1.51G	27.11M(91.6%)	1.32G(87.6%)	27.07M(91.5%)	1.32G(87.3%)	25.07M(92.6%)
cmt2_mRNA_rep1	SRR1005389	20.21M	1.01G	18.16M(89.9%)	0.86G(85.5%)	18.15M(89.8%)	0.86G(85.3%)	17.22M(94.9%)
cmt2_mRNA_rep2	SRR1005390	19.03M	0.97G	17.54M(92.2%)	0.86G(88.3%)	17.48M(91.9%)	0.85G(87.8%)	16.45M(94.1%)
cmt3_mRNA_rep1	SRR1005391	17.18M	0.86G	15.39M(89.6%)	0.73G(85.2%)	15.38M(89.5%)	0.73G(84.9%)	14.7M(95.5%)
cmt3_mRNA_rep2	SRR1005392	22.18M	1.13G	20.36M(91.8%)	0.99G(87.9%)	20.28M(91.4%)	0.99G(87.3%)	19.39M(95.6%)
cmt23_mRNA_rep1	SRR1005393	17.28M	0.86G	15.52M(89.8%)	0.74G(85.5%)	15.5M(89.7%)	0.74G(85.2%)	14.71M(94.9%)
cmt23_mRNA_rep2	SRR1005394	18.65M	0.95G	17.12M(91.8%)	0.84G(87.9%)	17.08M(91.6%)	0.83G(87.5%)	16.17M(94.7%)
drm12cmt2_mRNA_rep1	SRR1005395	23.87M	1.19G	21.36M(89.5%)	1.02G(85.1%)	21.34M(89.4%)	1.01G(84.9%)	20.32M(95.2%)
drm12cmt2_mRNA_rep2	SRR1005396	29.6M	1.51G	27.23M(92%)	1.33G(88.1%)	27.2M(91.9%)	1.33G(87.8%)	25.53M(93.9%)
drm12cmt3_mRNA_rep1	SRR1005397	19.52M	0.98G	17.48M(89.5%)	0.83G(85.2%)	17.46M(89.5%)	0.83G(84.9%)	16.59M(95%)
drm12cmt3_mRNA_rep2	SRR1005398	16.63M	0.85G	14.89M(89.5%)	0.73G(85.7%)	14.82M(89.1%)	0.72G(85.1%)	14.05M(94.8%)
drm12cmt23_mRNA_rep1	SRR1005399	22.07M	1.1G	19.88M(90.1%)	0.94G(85.6%)	19.86M(90%)	0.94G(85.3%)	18.91M(95.2%)
drm12cmt23_mRNA_rep2	SRR1005400	26.83M	1.37G	24.66M(91.9%)	1.2G(88%)	24.59M(91.6%)	1.2G(87.6%)	23.16M(94.2%)
suvh456_mRNA_rep1	SRR1005401	14.63M	0.73G	13.11M(89.6%)	0.62G(85.3%)	13.1M(89.5%)	0.62G(85%)	12.31M(94%)
suvh456_mRNA_rep2	SRR1005402	31.18M	1.59G	28.66M(91.9%)	1.4G(88.1%)	28.6M(91.7%)	1.4G(87.7%)	27.15M(94.9%)
atxr56_mRNA_rep1	SRR501598	29.08M	1.45G	26.27M(90.3%)	1.25G(86.1%)	26.25M(90.3%)	1.25G(85.9%)	24.91M(94.9%)
atxr56_mRNA_rep2	SRR501599	34.52M	1.73G	31.96M(92.6%)	1.47G(85.1%)	31.93M(92.5%)	1.47G(84.9%)	27.05M(84.7%)
mom1_mRNA_rep1	SRR501610	24.85M	1.24G	22.43M(90.3%)	1.07G(86%)	22.29M(89.7%)	1.06G(85.3%)	20.98M(94.2%)
mom1_mRNA_rep2	SRR501611	18.43M	0.92G	17.57M(95.4%)	0.86G(93%)	17.56M(95.3%)	0.85G(92.8%)	16.11M(91.7%)
atxr56drm12_mRNA_rep1	SRR501614	5.64M	0.28G	5.1M(90.4%)	0.24G(86.4%)	5.1M(90.4%)	0.24G(86.2%)	4.83M(94.7%)
atxr56drm12_mRNA_rep2	SRR501615	18.85M	0.94G	18M(95.5%)	0.88G(93.2%)	18M(95.5%)	0.88G(92.9%)	16.32M(90.7%)
atxr56mom1_mRNA_rep1	SRR501616	5.84M	0.29G	4.85M(83%)	0.22G(76.9%)	4.84M(82.9%)	0.22G(76.7%)	4.52M(93.3%)
atxr56mom1_mRNA_rep2	SRR501623	22.46M	1.12G	21.38M(95.2%)	1.04G(92.8%)	21.36M(95.1%)	1.04G(92.5%)	18.08M(84.6%)

Sample	SRA	Raw data		Low quality trimmed		Adapter removed		Bsmep report	
		Reads	Bases	Reads	Bases	Reads	Bases	Mapped Reads	Average covered
WT_BS	SRR1005412	0.22G	11.1G	0.18G(83%)	8.62G(77.7%)	0.18G(82.7%)	8.42G(75.9%)	0.15G(83.3%)	27.15
cmt23_BS	SRR1005413	0.22G	11.19G	24.33M(11.1%)	1.17G(10.4%)	24.25M(11%)	1.14G(10.2%)	23.1M(95.2%)	4.9
cmt2_BS	SRR869314	0.12G	6.06G	96.46M(81.2%)	4.57G(75.5%)	95.51M(80.4%)	4.31G(71.2%)	88.0M(92.1%)	15.05
cmt3_BS	SRR534209	0.21G	10.58G	0.18G(84.8%)	7.97G(75.3%)	0.18G(84.4%)	7.77G(73.4%)	0.17G(94.4%)	27.64
ddm1_BS	SRR534215	0.15G	7.52G	0.11G(72.1%)	4.8G(63.8%)	0.11G(71.7%)	4.72G(62.8%)	0.10G(91.4%)	17.24
drm12_BS	SRR534222	0.12G	6.05G	95.27M(78.7%)	4.38G(72.4%)	94.06M(77.7%)	4.09G(67.6%)	87.0M(92.5%)	14.53
drm12cmt2_BS	SRR1005414	0.04G	2.16G	33.47M(79.2%)	1.59G(73.6%)	33.35M(78.9%)	1.56G(72.2%)	32.3M(96.8%)	6.35
drm12cmt23_BS	SRR1005415	0.2G	10.16G	0.17G(84.2%)	8.06G(79.3%)	0.17G(84%)	7.89G(77.6%)	0.17G(97.6%)	30.89
met1_BS	SRR534239	0.28G	14.14G	0.24G(85.9%)	11.25G(79.6%)	0.24G(85.7%)	11.08G(78.3%)	0.22G(91.6%)	40.52
suvh456_BS	SRR534253	0.12G	5.84G	92.05M(78.8%)	4.24G(72.6%)	91.81M(78.6%)	4.18G(71.6%)	83.1M(90.5%)	14.22
vim1/2/3_BS	SRR534266	0.25G	12.33G	0.21G(83.2%)	9.47G(76.9%)	0.2G(83.1%)	9.41G(76.3%)	0.19G(97.6%)	32.5
atxr56_BS	SRR534204	0.25G	12.55G	0.23G(91%)	10.82G(86.2%)	0.23G(90.8%)	10.63G(84.7%)	0.22G(95.6%)	41.22

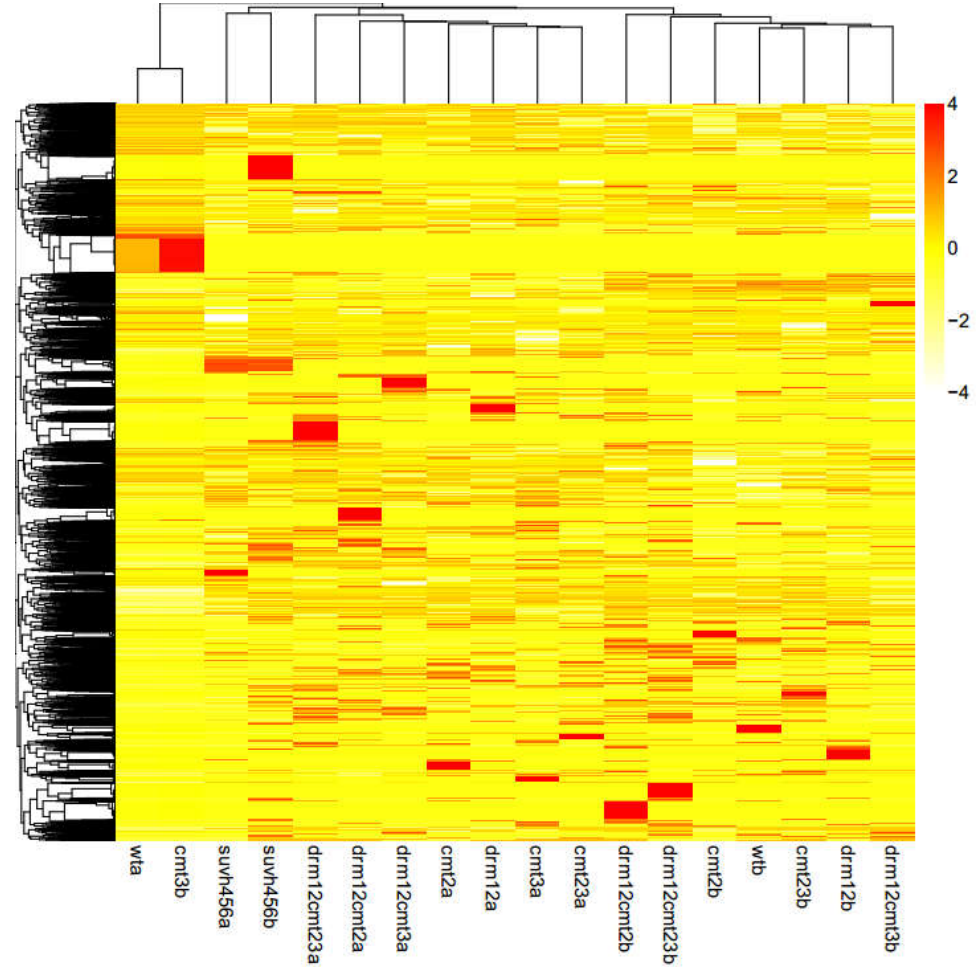
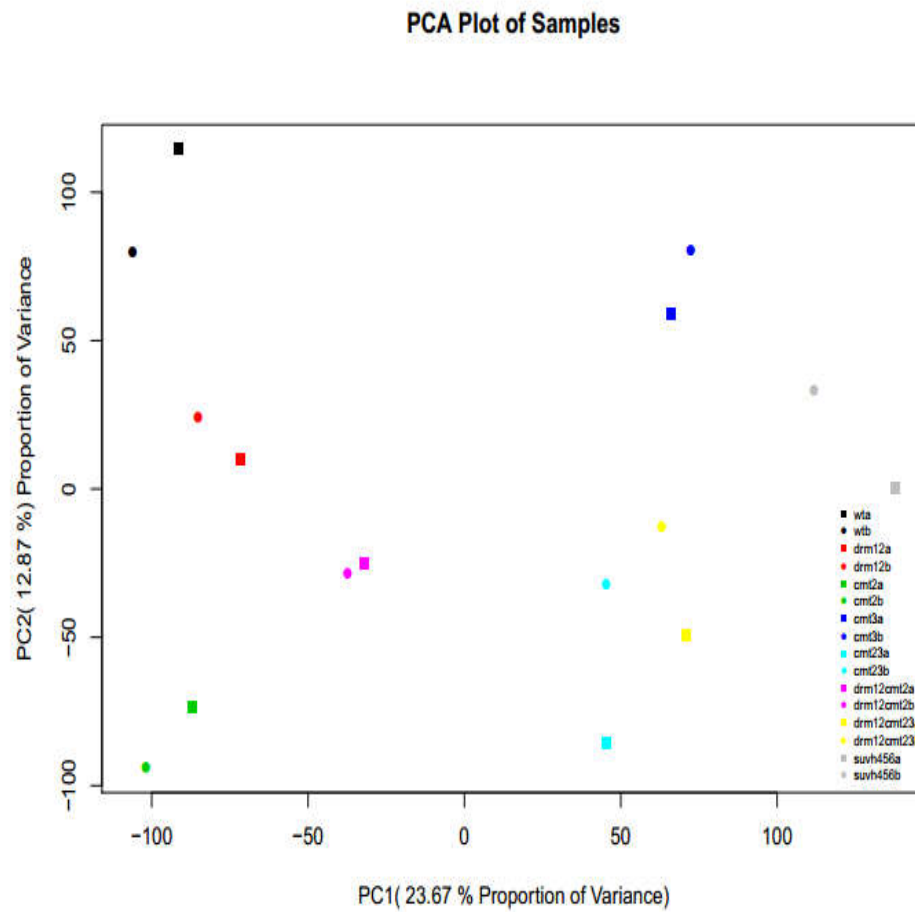
- RNA-seq Data and BS-seq Data Quality Control

Sample	Raw data		Adapter removed		Uniq Mapped Reads	Coverage
	Reads	Bases	Reads	Bases		
WT_mRNA_rep1	16.13M	0.81G	14.46M(89.6%)	0.69G(85.1%)	13.47M(93.2%)	10.37
WT_mRNA_rep2	21.47M	1.09G	19.61M(91.3%)	0.96G(87.2%)	18.63M(95%)	14.71
drm12_mRNA_rep1	17.99M	0.9G	16.13M(89.7%)	0.77G(85.2%)	15.35M(95.1%)	11.81
drm12_mRNA_rep2	29.6M	1.51G	27.07M(91.5%)	1.32G(87.3%)	25.07M(92.6%)	19.71
cmt2_mRNA_rep1	20.21M	1.01G	18.15M(89.8%)	0.86G(85.3%)	17.22M(94.9%)	13.16
cmt2_mRNA_rep2	19.03M	0.97G	17.48M(91.9%)	0.85G(87.8%)	16.45M(94.1%)	12.9

• •
• •
• •
• •

- Few of splicing factor or splicing factor like genes(138) are differentially expressed

The correlation of gene expression and SI in different Samples



The number of DAS events in different mutants

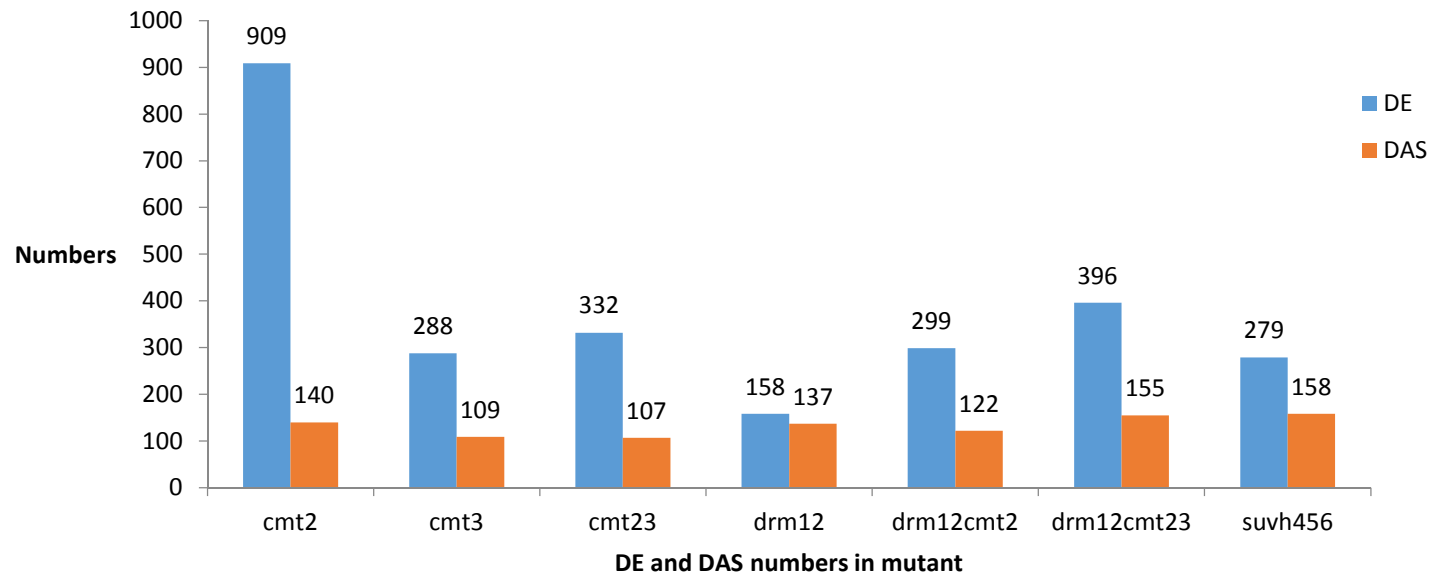
Pvalue <0.01, abs(FC)>2

Mutant	cmt2	cmt3	cmt23	drm12	drm12cmt2	drm12cmt23	suvh456
cmt2	41						
cmt3	4	28					
cmt23	2	7	31				
drm12	9	6	6	50			
drm12cmt2	6	6	6	11	39		
drm12cmt23	4	6	5	8	11	36	
suvh456	9	6	6	15	12	12	60

Pvalue <0.05, abs(FC)>1.5

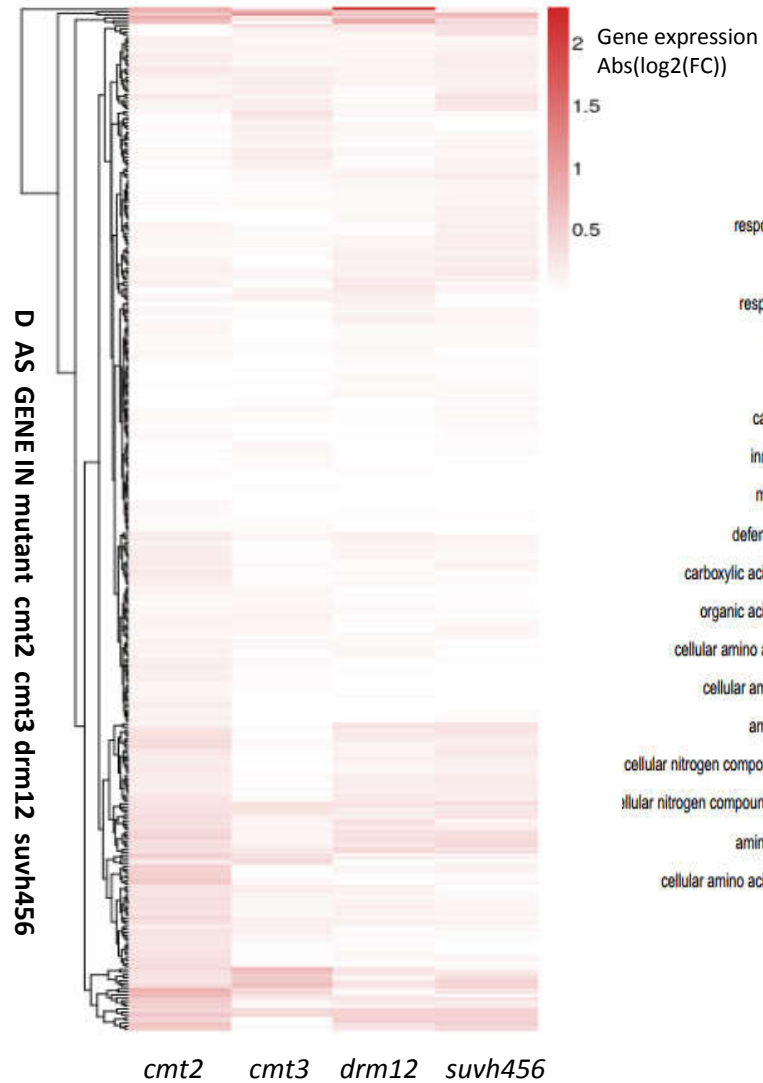
Mutant	cmt2	cmt3	cmt23	drm12	drm12cmt2	drm12cmt23	suvh456
cmt2	154						
cmt3	24	118					
cmt23	30	25	116				
drm12	42	26	33	150			
drm12cmt2	38	30	28	32	134		
drm12cmt23	37	31	26	40	32	166	
suvh456	41	31	40	55	44	47	179

DE and DAS genes are separately regulated

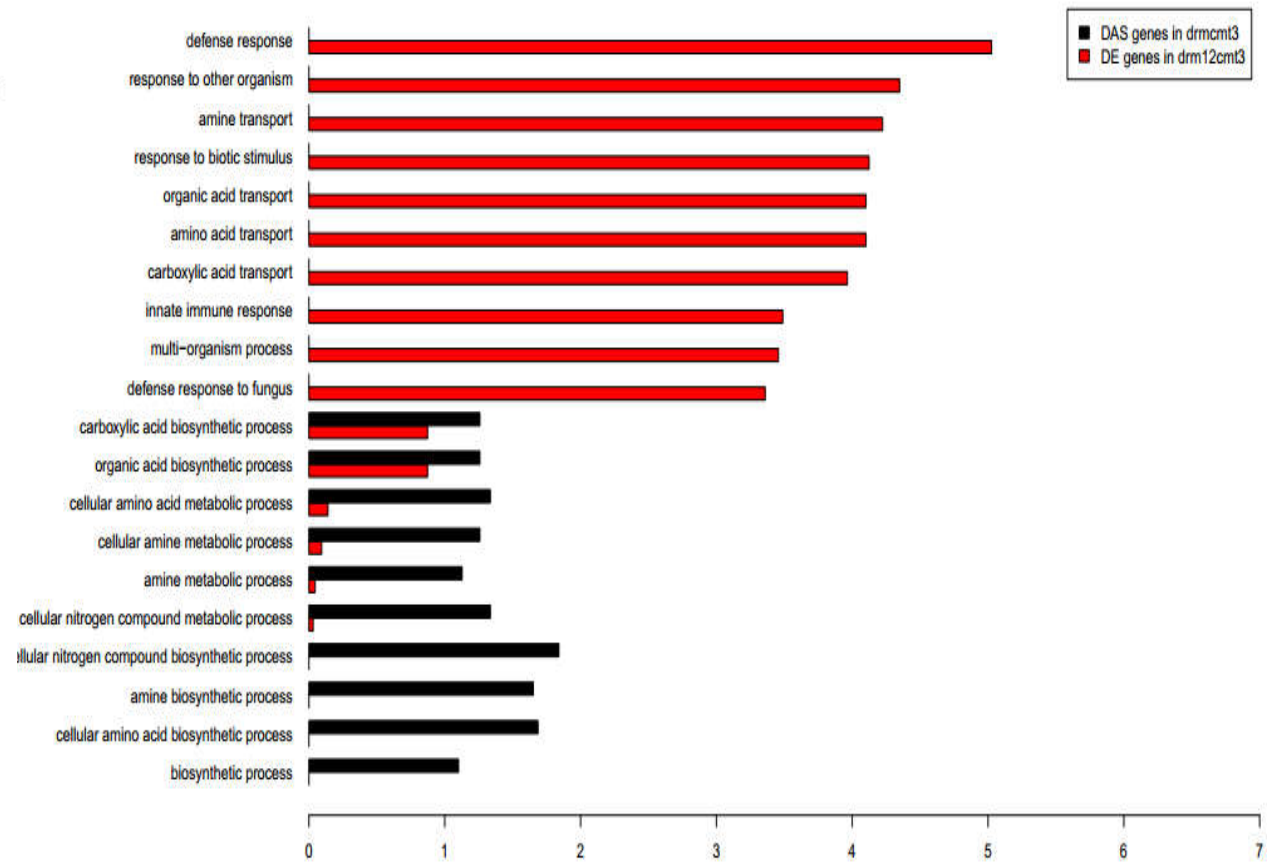


Mutant	DE	DAS	Overlap
cmt2	909	140	4
cmt3	288	109	2
cmt23	332	107	1
drm12	158	137	0
drm12cmt2	299	122	1
drm12cmt23	396	155	2
suvh456	279	158	1

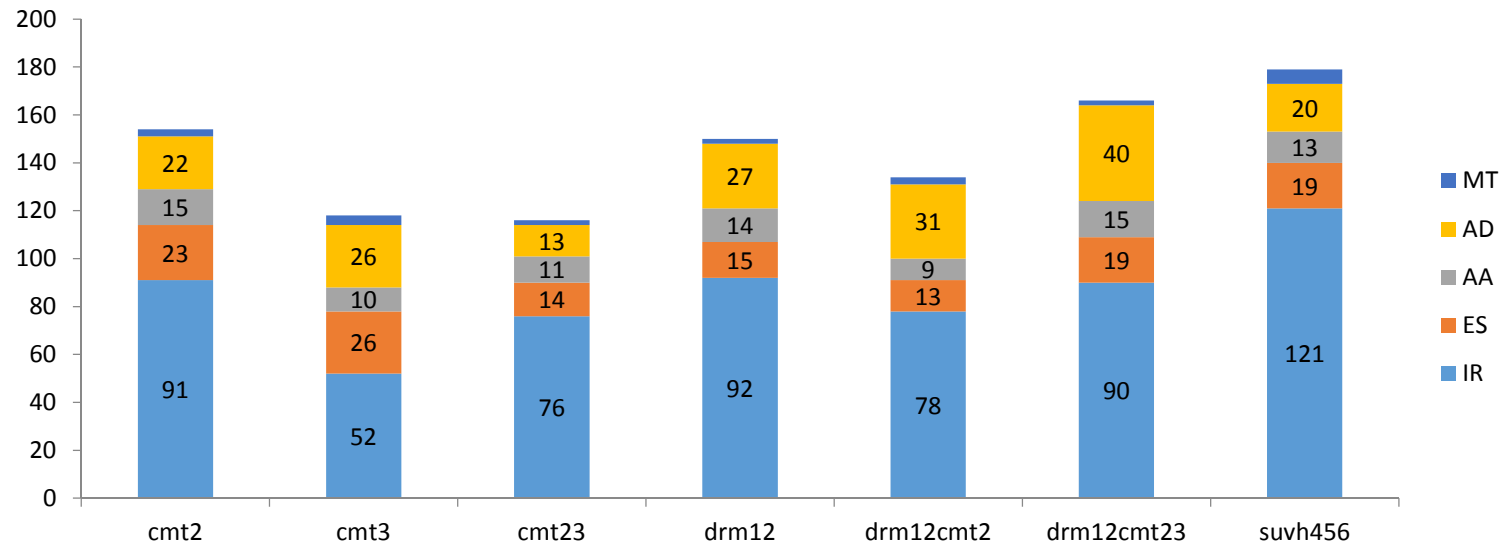
DAS gene expression level



DAS gene and DE gene GO enrichment GO



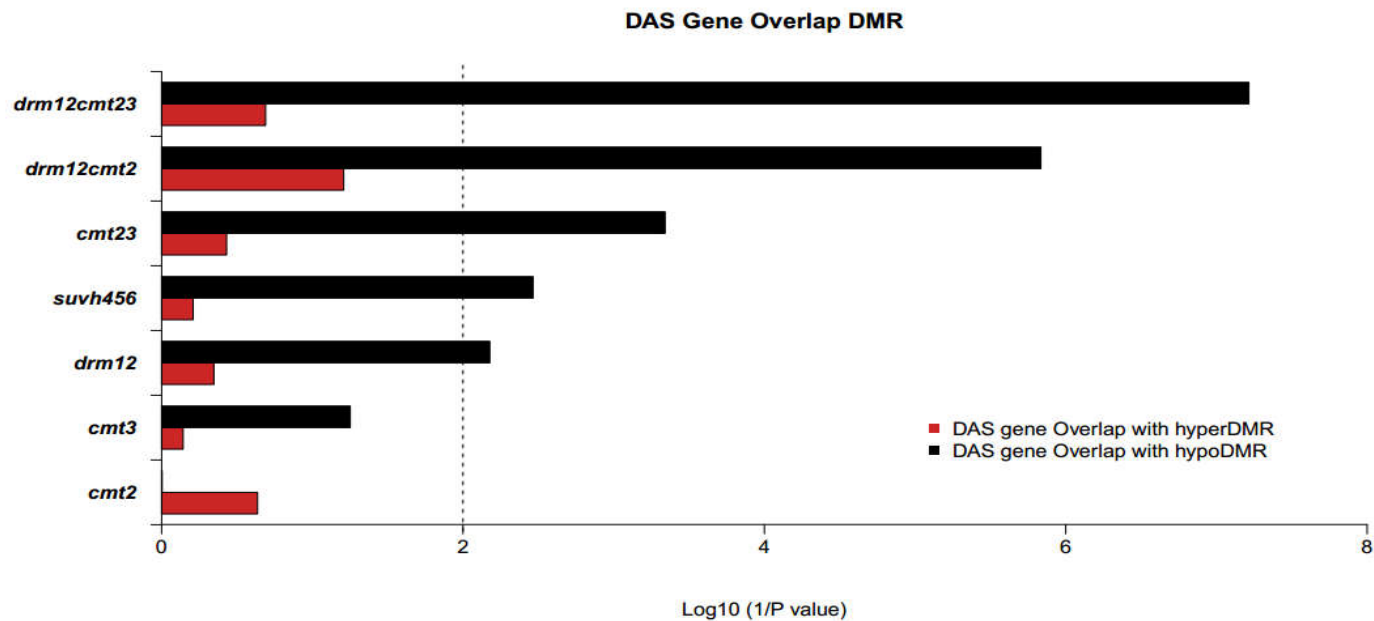
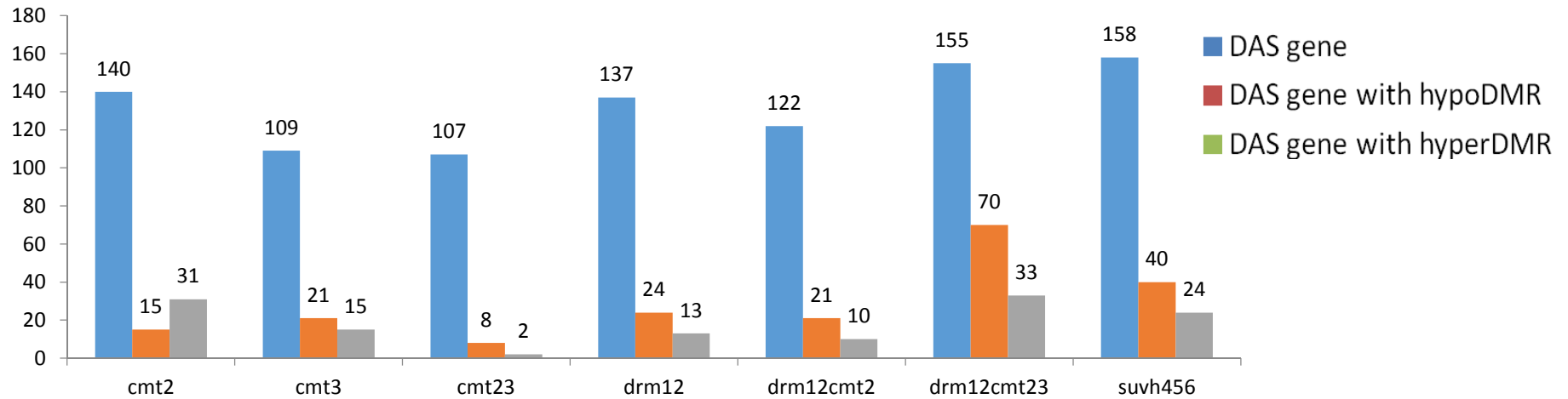
- The proportion of five type DAS events in mutants



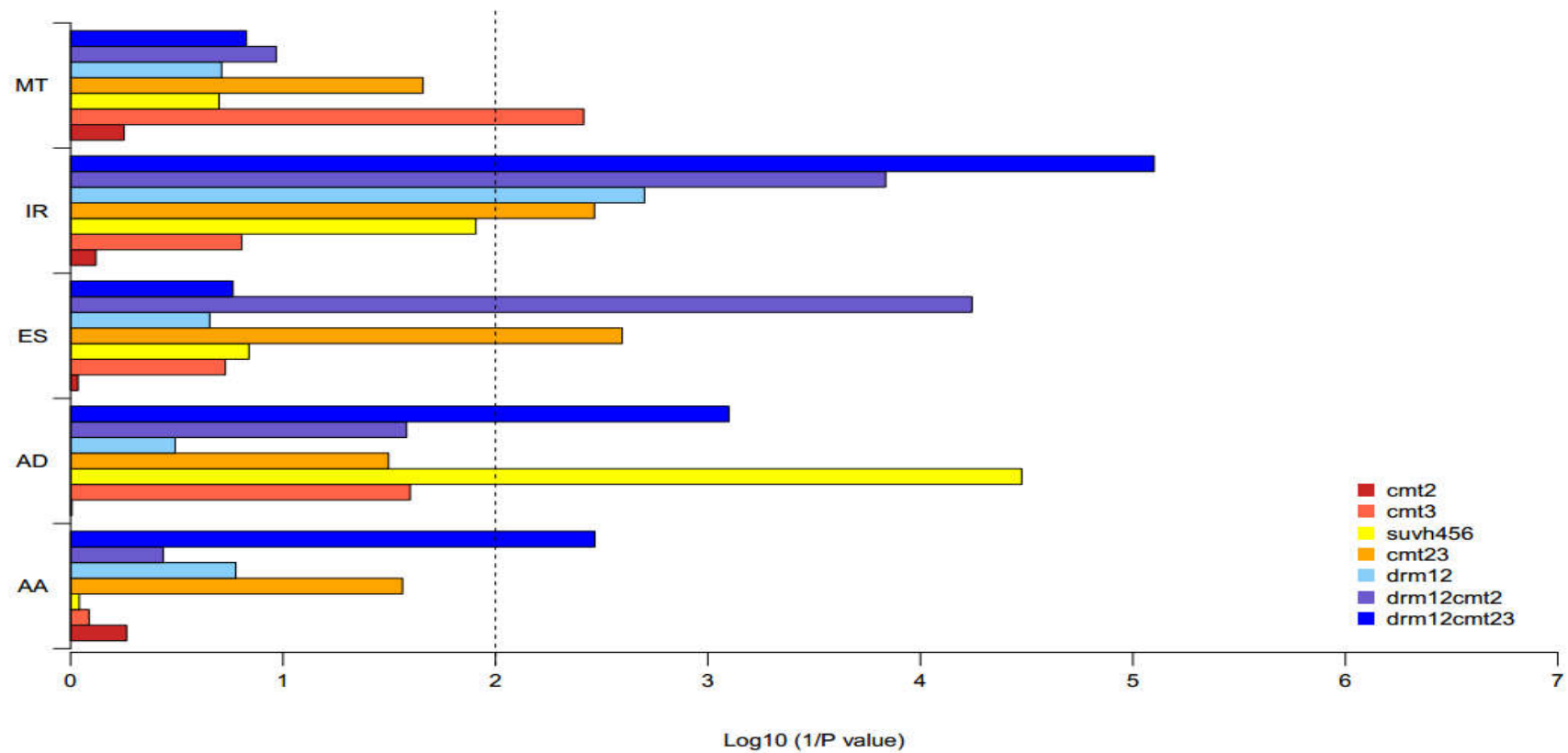
- The splice signal of DAS events in mutants

Splice signals	IR	ES	AA	AD	MT
GTAG	336	86	52	119	14
GCAG	14	2	7	1	2
CTAC	19	1	0	0	0
else	4	2	0	0	0

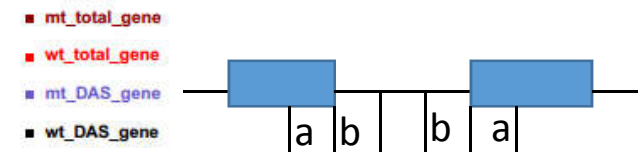
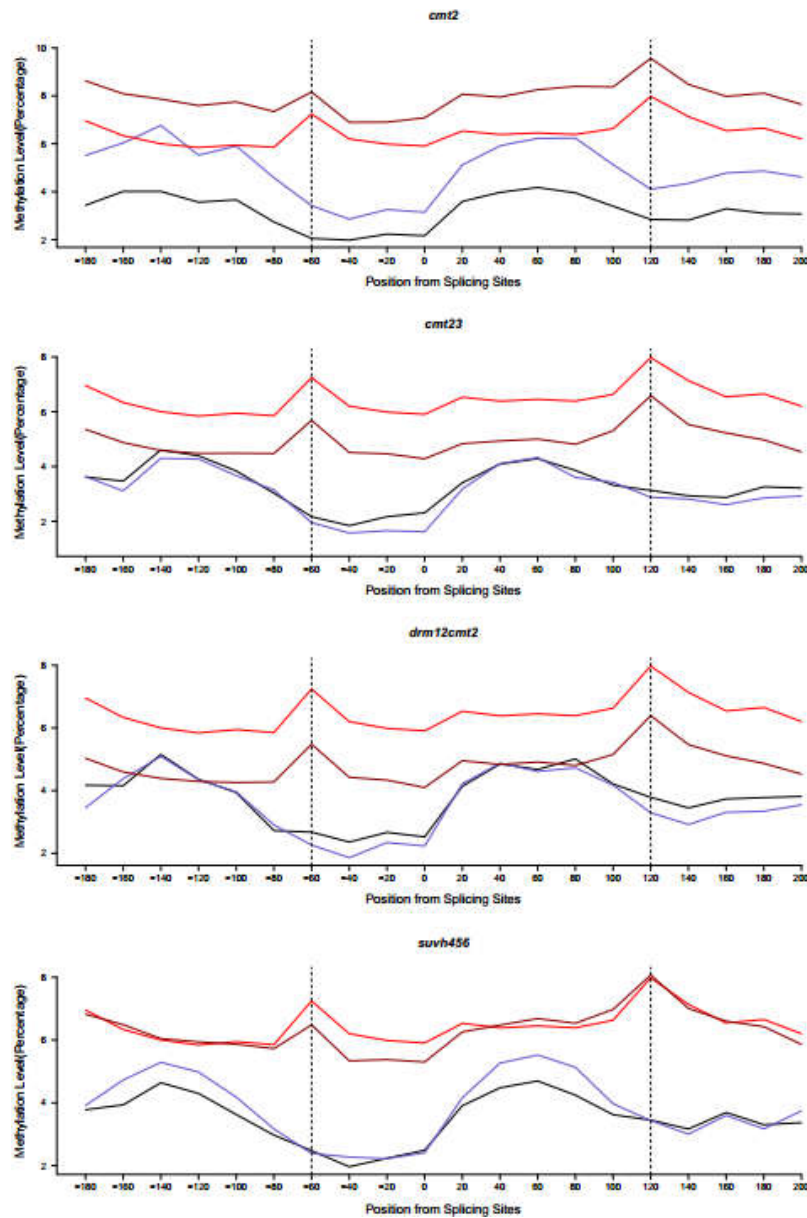
- DAS events have cooccurrence with the DNA methylation level change



Different type of DAS Gene Overlap HypoDMR



- DAS genes have distinct DNA methylation level in exon-intron boundary



a:0-200
b:-180-0

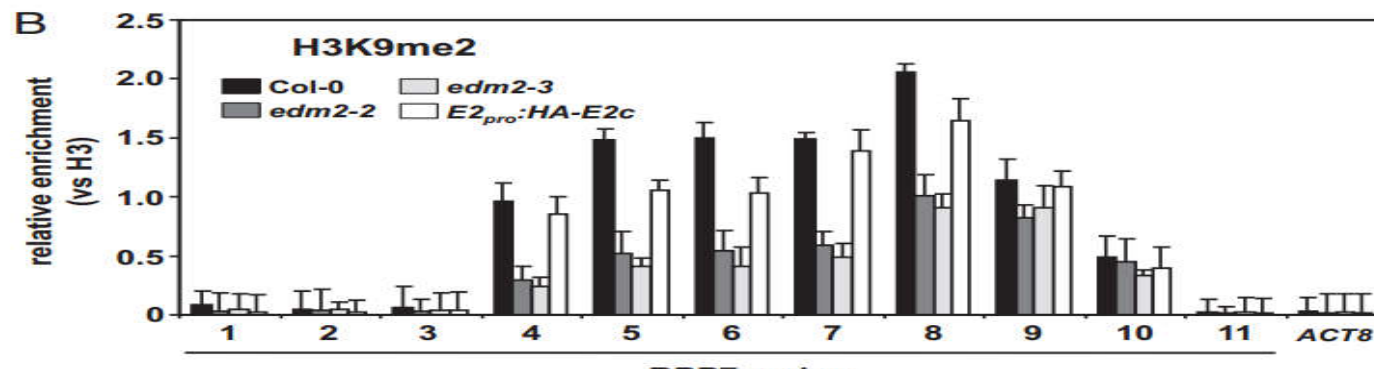
- The DNA methylation and H3K9me2 modification of LTR-transposon in AT1G58602 first intron can regulate the AT1G58602 (RPP7) splicing

An alternative polyadenylation mechanism coopted to the *Arabidopsis RPP7* gene through intronic retrotransposon domestication

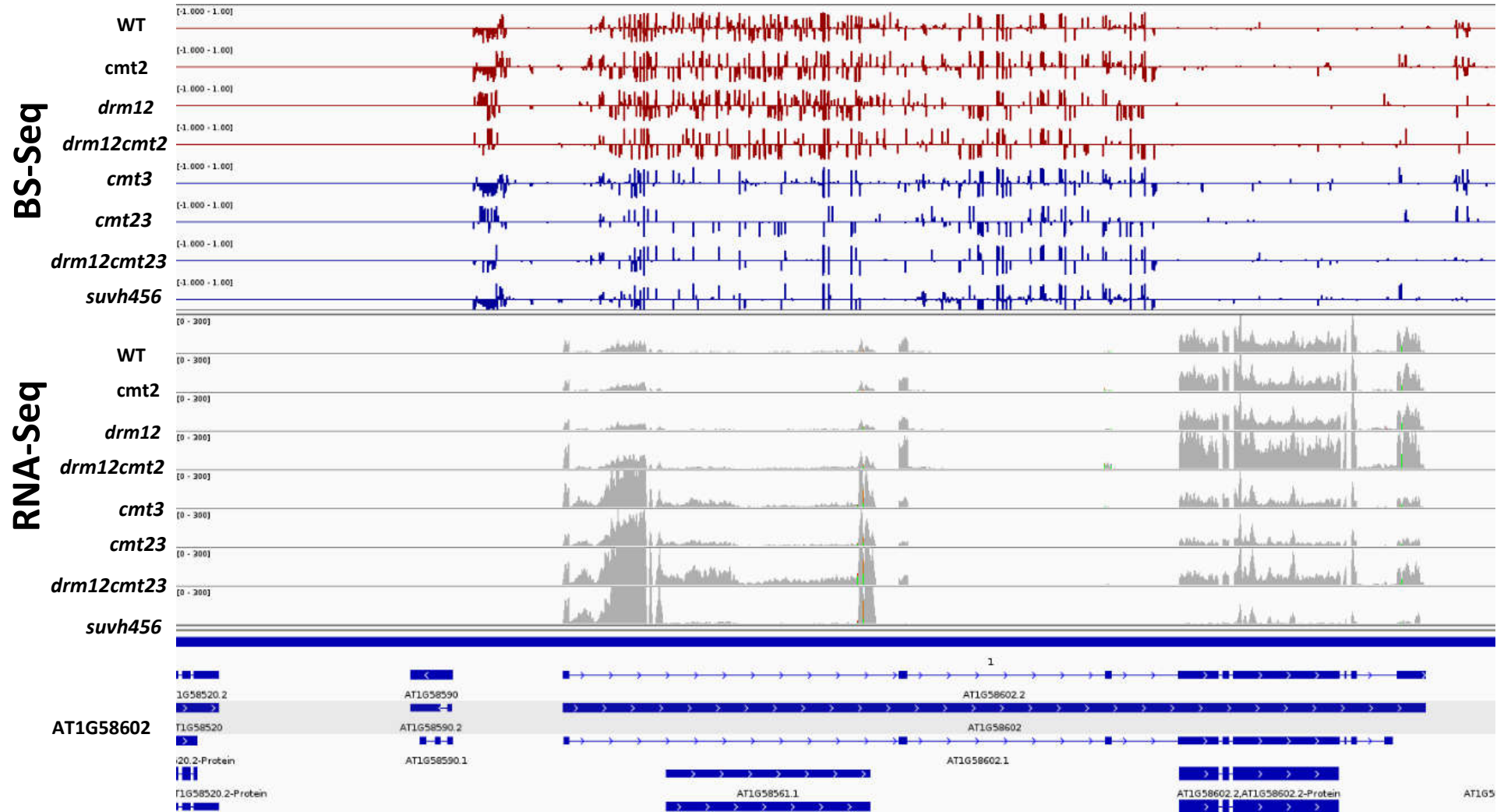
Tokuji Tsuchiya and Thomas Eulgem¹

Center for Plant Cell Biology, Institute for Integrative Genome Biology, Department of Botany and Plant Sciences, University of California, Riverside, CA 92521

Edited by Susan R. Wessler, University of California, Riverside, CA, and approved July 23, 2013 (received for review July 2, 2013)



(Tokuji Tsuchiya et al., 2013, PNAS)



- Three couples of genes and TE which the TE is inserted into gene intron with the same strand

AT1G58602 and AT1G58561

AT1G58602 defense response

AT1G58561 copia-like retro-transposon family

AT3G32920 and AT3G32925

AT3G32920 recA DNA recombination family protein

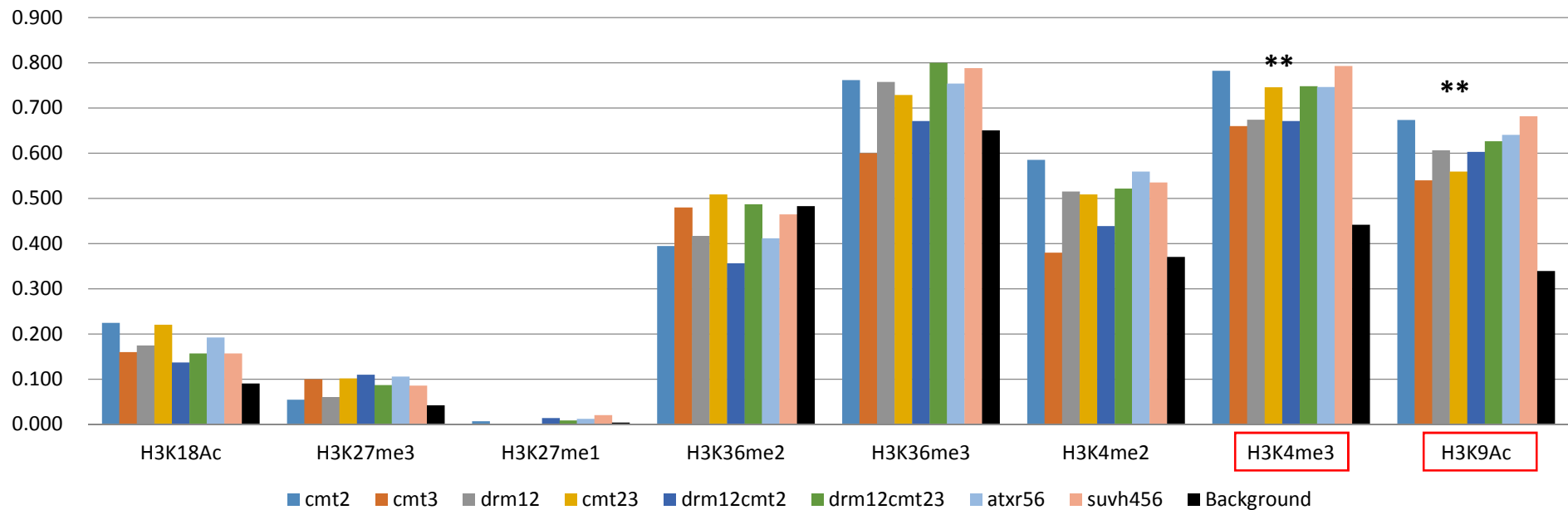
AT3G32925 Non-LTR retro-transposon family

AT3G52700 and AT3G52705

AT3G52700 unknown protein

AT3G52705 copia-like retro-transposon family

- The AS events with the H3K4me3/H3K9Ac histone modification may be more vulnerable to be influenced



** P value <0.01