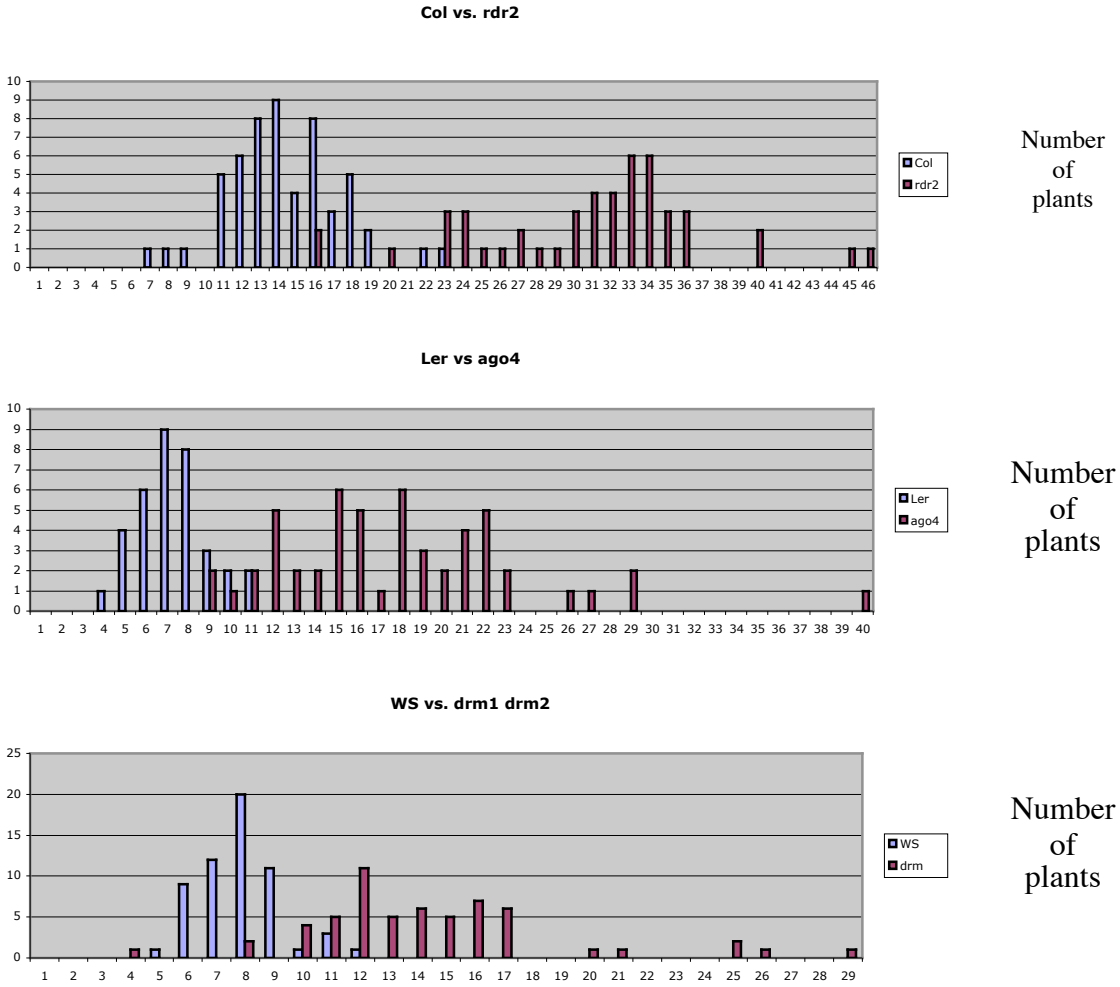
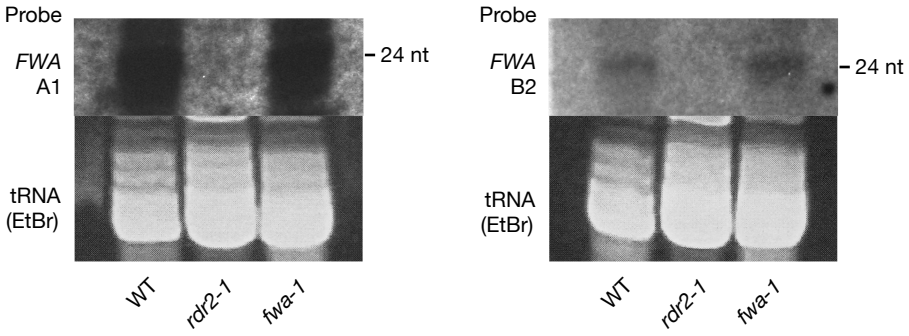
Supplementary Material for "Two-Step Recruitment of RNA-directed DNA Methylation to Tandem Repeat Sequences." by Simon W.-L. Chan, Yana Bernatavichute, Xiaoyu Zhang and Steven E. Jacobsen.

Supplementary Figure S1. Flowering time distribution for *rdr2-1*, *ago4-1*, and *drm1-1 drm2-1* compared to their respective wild-type ecotypes.

Histograms show rosette leaf number on the X axis and number of plants on the Y axis. Late flowering was defined as 23 or more leaves for *rdr2-1/RDR2*, 10 or more leaves for *ago4-1/AGO4*, and 12 or more leaves for *drm1-1/DRM1 drm2-1/DRM2*.



Supplementary Figure S2



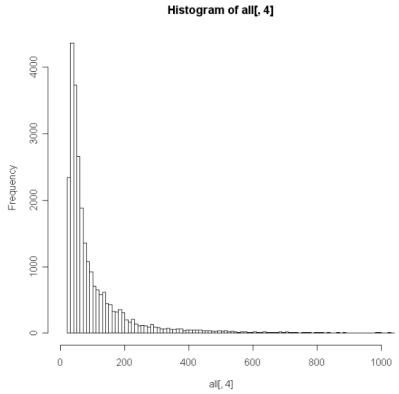
siRNAs from the indicated genotypes were probed with LNA probes corresponding to the *FWA* tandem repeats.

The *FWA* tandem repeats are base pairs -1063 to -570. The probe used in Figures 4 and 5 is antisense -889 to -936 (within first large repeat).

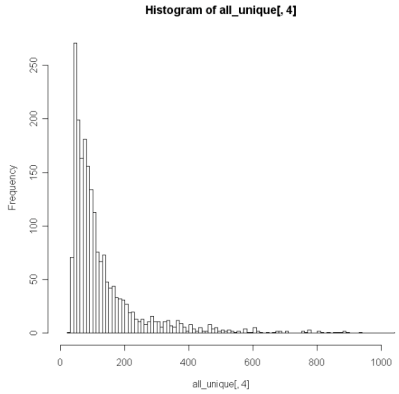
Probe A1 is antisense -1026 to -1063 (within small repeat).
Probe B2 is antisense -602 to -643 (within second large repeat).

Supplemental Figure 3. Comparisons of the repeat characteristics of all TRs, unique TRs and unique TRs associated with siRNAs

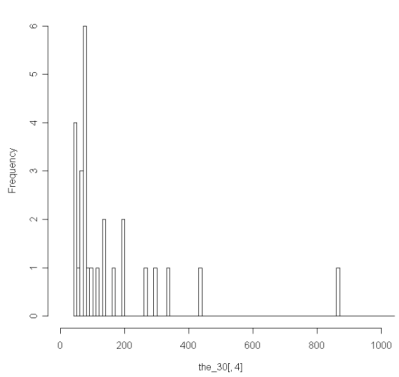
all TRs



all unique TRs



unique TRs with siRNAs



Supplemental Figure 3A. Comparison of the length of all TRs (top), unique TRs (middle) and unique TRs associated with siRNAs (bottom).

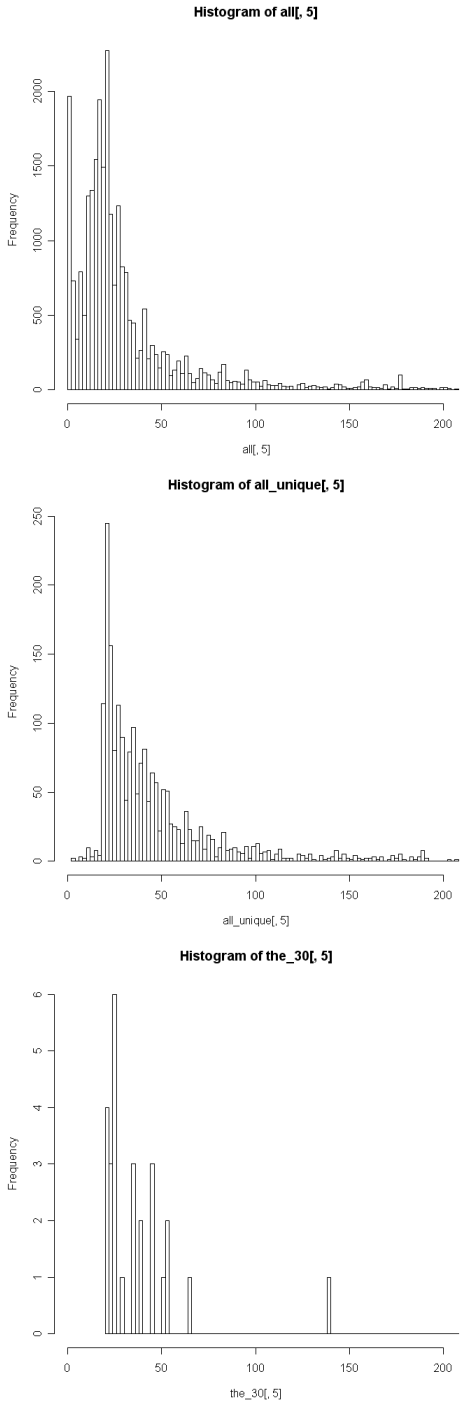
x-axis: length of each TR

y-axis: frequency

all TRs

all unique TRs

unique TRs with siRNAs

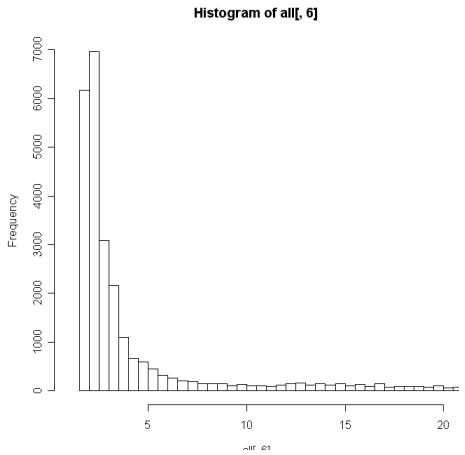


Supplemental Figure 3B. Comparison of the repeat unit length of all TRs (top), unique TRs (middle) and unique TRs associated with siRNAs (bottom).

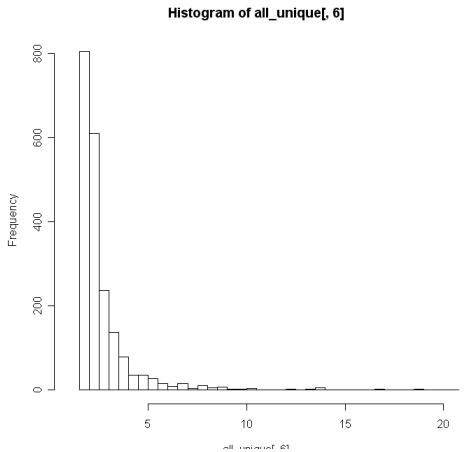
x-axis: repeat unit length length of each TR

y-axis: frequency

all TRs

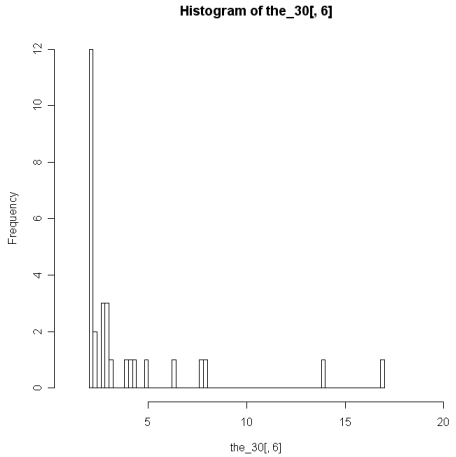


all unique TRs



unique TRs

with siRNAs



Supplemental Figure 3C. Comparison of the number of repeat units within each TR, for all TRs (top), unique TRs (middle) and unique TRs associated with siRNAs (bottom).

x-axis: number of repeat units within each TR

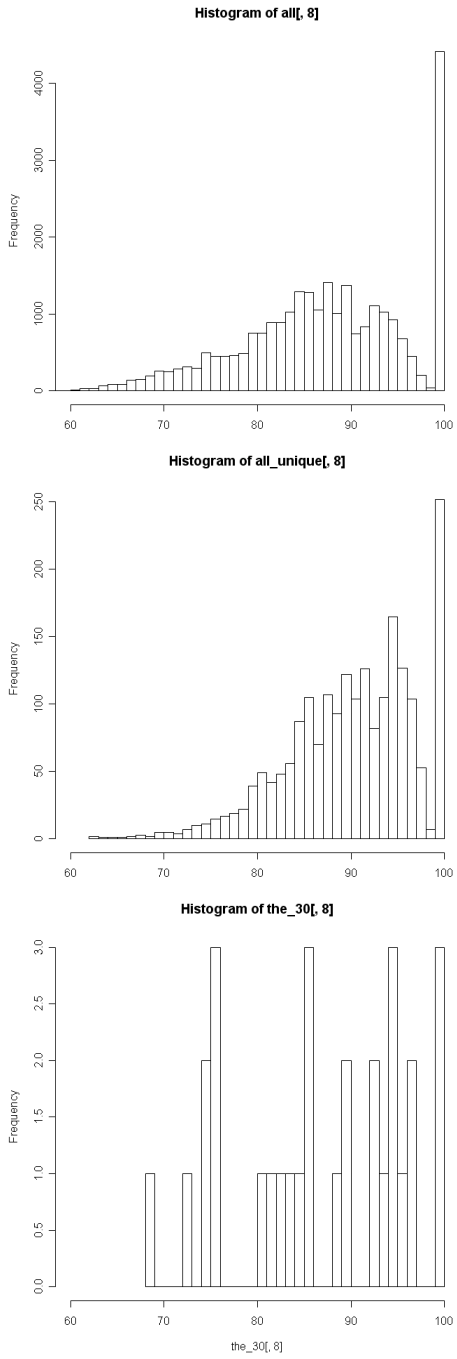
y-axis: frequency

all TRs

all unique TRs

unique TRs

with siRNAs



Supplemental Figure 3D. Comparison of the average identity among the repeat units within each TR, for all TRs (top), unique TRs (middle) and unique TRs associated with siRNAs (bottom).

x-axis: average identity among the repeat units within each TR

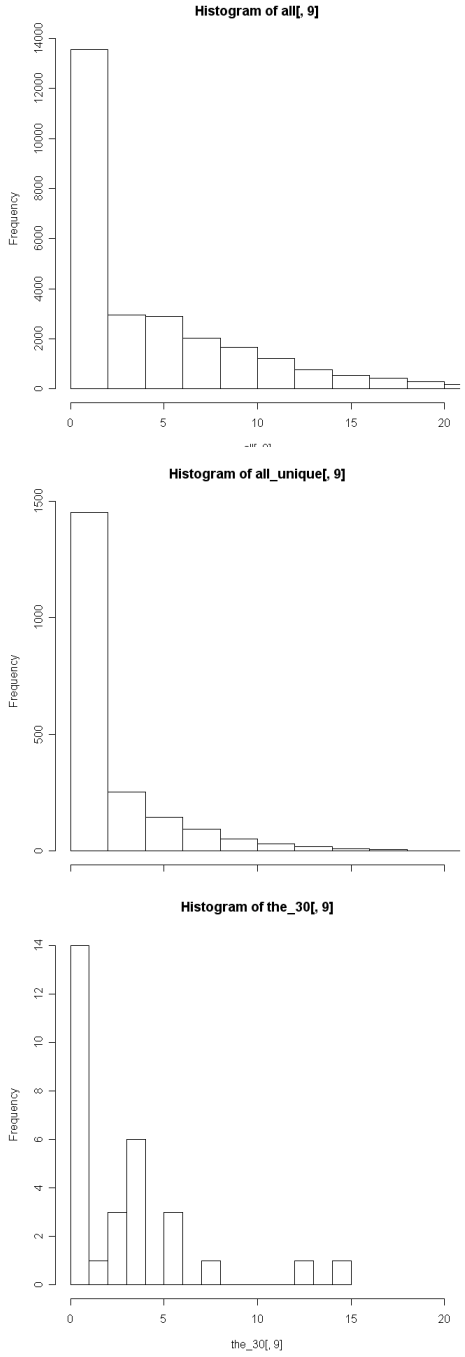
y-axis: frequency

all TRs

all unique TRs

unique TRs

with siRNAs

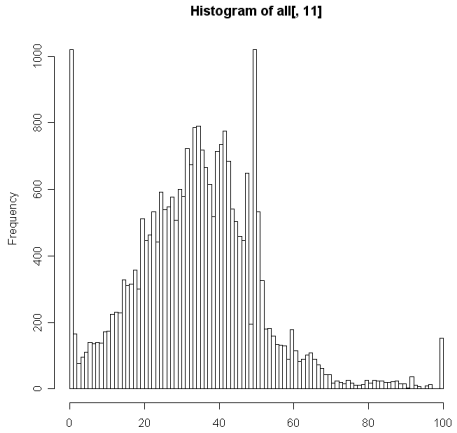


Supplemental Figure 3E. Comparison of the percentage of insertions and deletions (indels) among the repeat units within each TR, for all TRs (top), unique TRs (middle) and unique TRs associated with siRNAs (bottom).

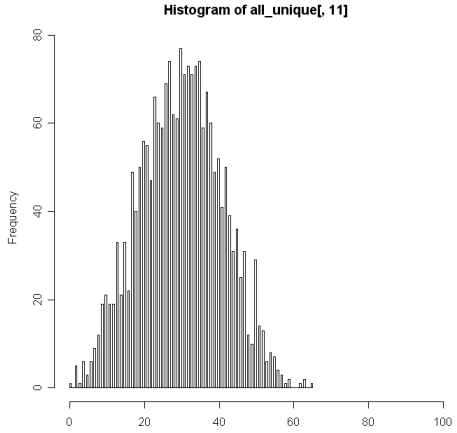
x-axis: percentage of indels among the repeat units within each TR

y-axis: frequency

all TRs

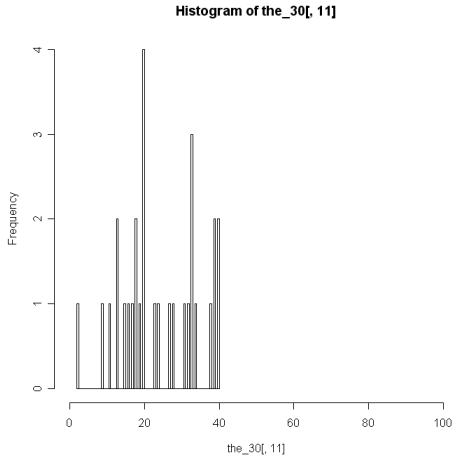


all unique TRs



unique TRs

with siRNAs

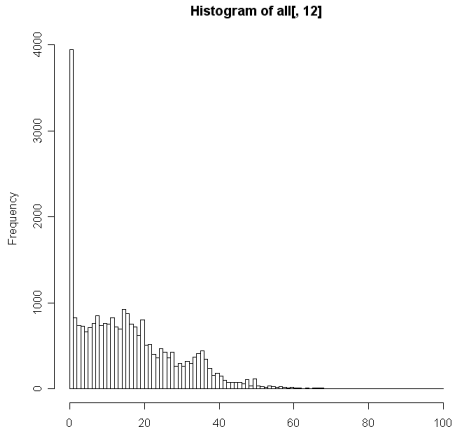


Supplemental Figure 3F. Comparison of the sequence content (% of A) of all TRs (top), unique TRs (middle) and unique TRs associated with siRNAs (bottom).

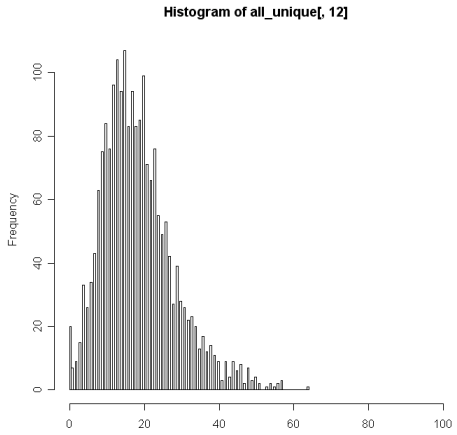
x-axis: percentage of indels among the repeat units within each TR

y-axis: frequency

all TRs

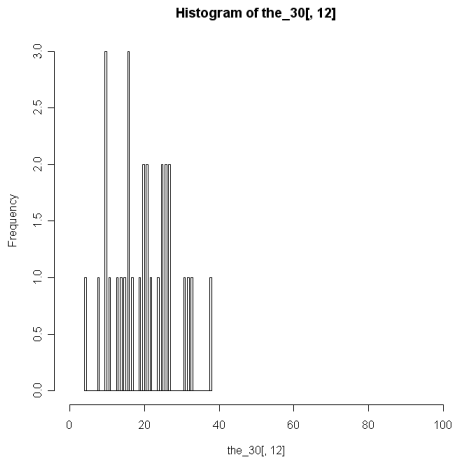


all unique TRs



unique TRs

with siRNAs

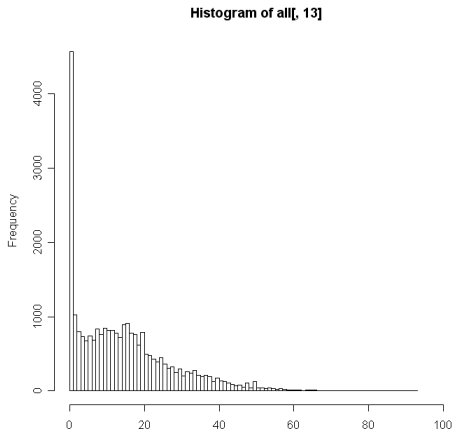


Supplemental Figure 3G. Comparison of the sequence content (% of C) of all TRs (top), unique TRs (middle) and unique TRs associated with siRNAs (bottom).

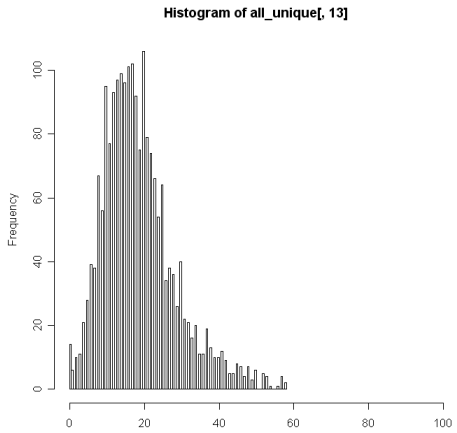
x-axis: percentage of indels among the repeat units within each TR

y-axis: frequency

all TRs

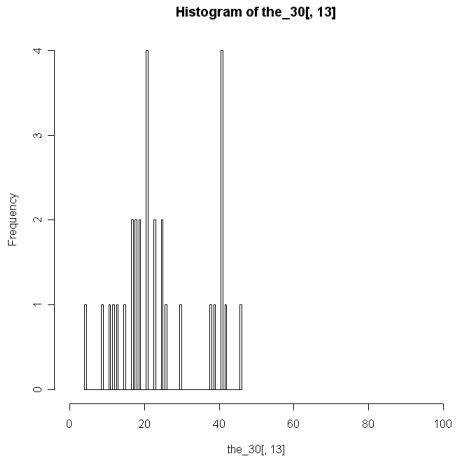


all unique TRs



unique TRs

with siRNAs



Supplemental Figure 3H. Comparison of the sequence content (% of G) of all TRs (top), unique TRs (middle) and unique TRs associated with siRNAs (bottom).

x-axis: percentage of indels among the repeat units within each TR

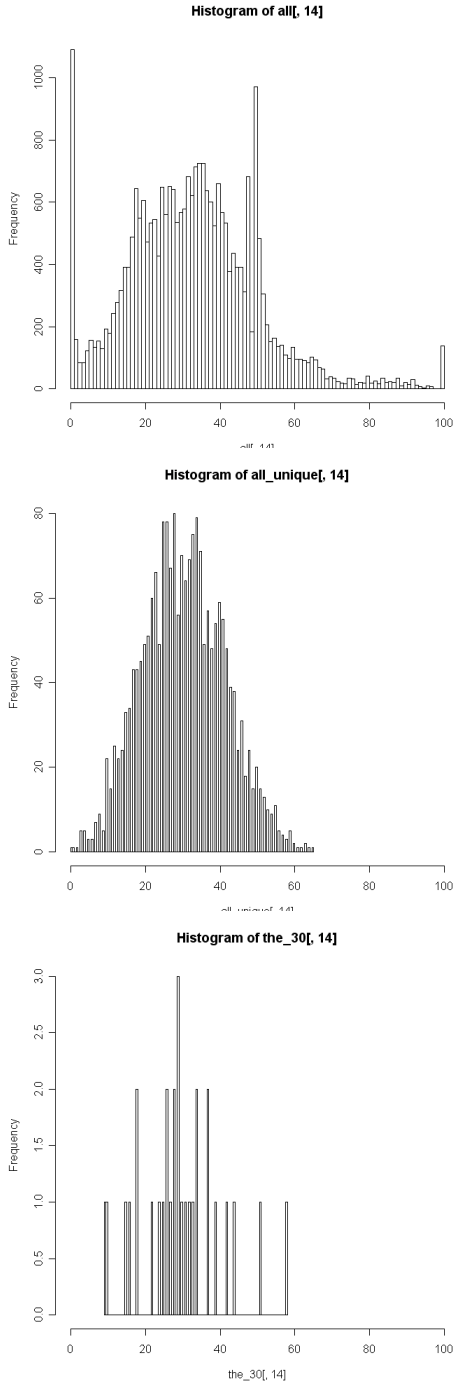
y-axis: frequency

all TRs

all unique TRs

unique TRs

with siRNAs



Supplemental Figure 3I. Comparison of the sequence content (% of T) of all TRs (top), unique TRs (middle) and unique TRs associated with siRNAs (bottom).

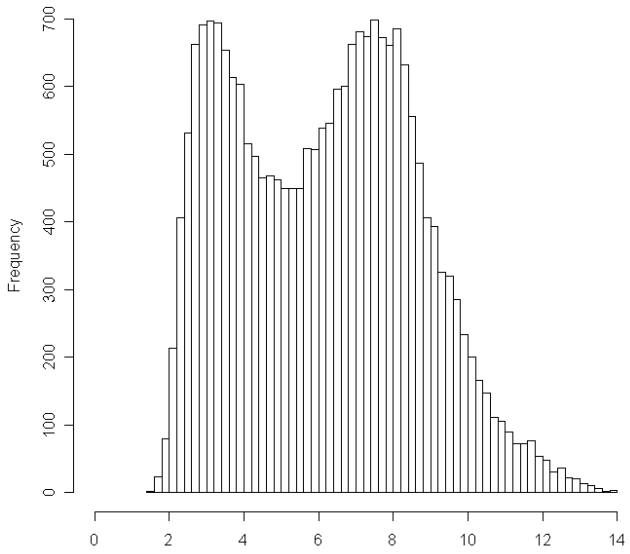
x-axis: percentage of indels among the repeat units within each TR

y-axis: frequency

Supplementary Figure S4

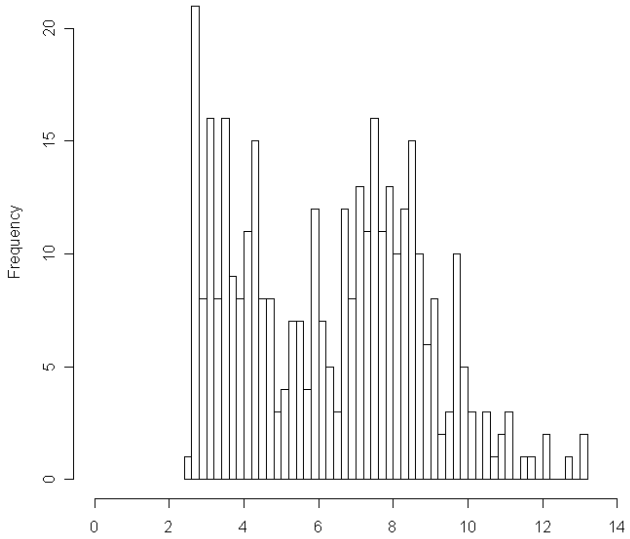
Expression level of genes with and without unique tandem repeats in their promoter regions.

The distribution of expression levels of A) all genes, B) genes containing unique tandem repeats (TRs, see text for a detailed definition of unique tandem repeats) within their promoter regions (defined as 1 kb upstream of transcription start site), and C) genes containing unique tandem repeats within their promoter regions that are also associated with siRNAs. The expression level of each gene is the average over 79 tissues and conditions (Schmid et al. 2005 Nat Genet 37: 501-506) shown in log2 scale.



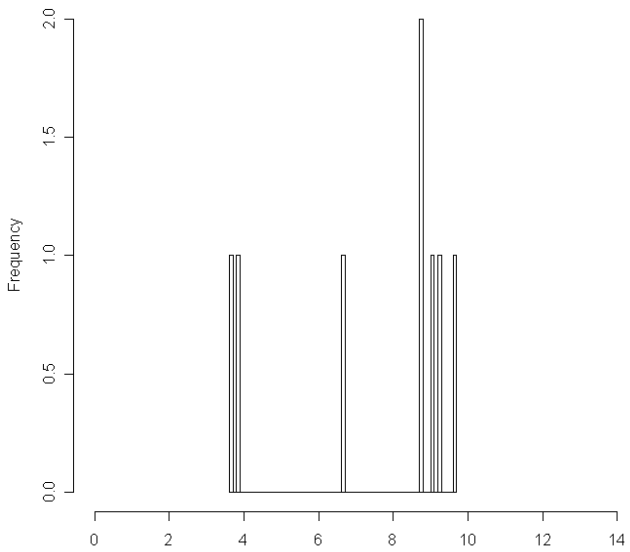
A) All genes

- x: intensity (log2 of RMA mean)
- y: frequency



B) Genes with TRs in promoters (1 kb)

- x: intensity (log2 of RMA mean)
- y: frequency



C) Genes with siRNA-associated TRs in promoters (1 kb)

- x: intensity (log2 of RMA mean)
- y: frequency

Supplementary Material for "Two-Step Recruitment of RNA-directed DNA Methylation to Tandem Repeat Sequences." by Simon W.-L. Chan, Yana Bernatavichute, Xiaoyu Zhang and Steven E. Jacobsen.

Table S1. Bisulfite genomic sequencing data.

Number of cytosines methylated in different sequence contexts in cloned PCR products from bisulfite-treated DNA.

FWA T1 transgene. 20 CG sites, 10 CNG sites, 61 asymmetric sites.

FWA "single copy" T1 transgene. 11 CG sites, 5 CNG sites, 26 asymmetric sites.

	CG methylated	CNG methylated	CHH (asymmetric) methylated
Col + *FWA*	117/240 (48.8%)	32/120 (26.7%)	84/732 (11.5%)
Col + *FWA* single copy	0/286 (0%)	0/130 (0%)	1/676 (0.15%)
fwa-1 + *FWA* late flowering T1	0/420 (0%)	0/210 (0%)	5/1281 (0.4%)

FWA endogene. 20 CG sites, 10 CNG sites, 61 asymmetric sites.

	CG methylated	CNG methylated	CHH (asymmetric) methylated
L*er*	335/400 (84%)	31/200 (16%)	83/1220 (6.8%)
rdr2-1	352/400 (88%)	0/200 (0%)	6/1220 (0.5%)
nrpd1a-1	347/400 (87%)	7/200 (3.5%)	10/1220 (0.8%)
F1 *rdr2-1 x nrpd1a-1*	181/240 (75%)	13/120 (11 %)	53/752 (7%)
kyp-2	375/440 (85.2%)	20/242 (8.3%)	51/1342 (3.8%)
F1 *kyp2 rdr2-1 x kyp2 nrpd1a-1*	297/340 (87.3%)	21/170 (12.4%)	94/1037 (9.1%)

Supplementary Table S2
Chan et al "Two-step recruitment of RNA-directed DNA methylation to tandem repeats"

PCR primer sequence	Purpose
TGACTGACAGCTGAAAATGGGATGTGGAT	ago4-1 genotyping (CAPS marker with AvaII)
GCCACTCCCTAGAACTCACCACCTAAGTT	ago4-1 genotyping
tag cat ctg aat ttc ata acc aat ctc gat aca c	rdr2-1 SAIL T-DNA left border primer
aca cat tag gac taa caa att tac c	rdr2-1 genotyping
atg gtg tca gag acg acg acg aac cga tca	rdr2-1 genotyping
ggttttatattaatattaaagagttatgggtygaagttt	bisulfite sequencing of FWA
caaaatactttacacataaacraaaaacaaacaaatcraa	bisulfite sequencing of FWA
caaaatactttacacataaacraaaaacaaacaaatcraa	bisulfite PCR/ClaI assay of FWA methylation
GTttAAGTGtTATTTGGTTGTTTAAGGTTGtTTTTAGtAt	bisulfite PCR/ClaI assay of FWA methylation
ag+caa+cct+taa+aca+acc+aaa+tag+cac+ttg+gac+caa+tgg+cga+a	FWA LNA probe from Figure 4A
AT+ATG+AGA+TTC+TCG+ACG+GAA+AGA+TGT+ATG+GGC+TTCG	FWA LNA probe A1 (Supplementary Figure S3)
CTG+ATT+GTC+AGT+ATCCT+ACA+AATC+GAT+AAATC+GTA+AGA+AG	FWA LNA probe B2 (Supplementary Figure S3) (+ indicates LNA at following position)
CCTTCGGGATTTTCGATAGTGCCA	Taqman probe for FWA real-time RT-qPCR
CCCACCAAGATCTGAAGTCC	FWA RT-qPCR
CAGGTGCAATGGTGGTGTAT	FWA RT-qPCR
Ttt tcc cta gtt gag atg gga att gaa	Taqman probe for ACT7 real-time RT-qPCR
Cagcatcatcacaagcatcc	ACT7 RT-qPCR
Tcgtggtggtgagtttgttac	ACT7 RT-qPCR
aaa gtg gtt gta gtt tat gaa agg ttt tat	bisulfite PCR/BamHI assay of MEA-ISR methylation
CTTAAAAAATTTTCAACTCATTTTTAAAAAA	bisulfite PCR/BamHI assay of MEA-ISR methylation