

Supporting Information

Lu et al. 10.1073/pnas.1203094109

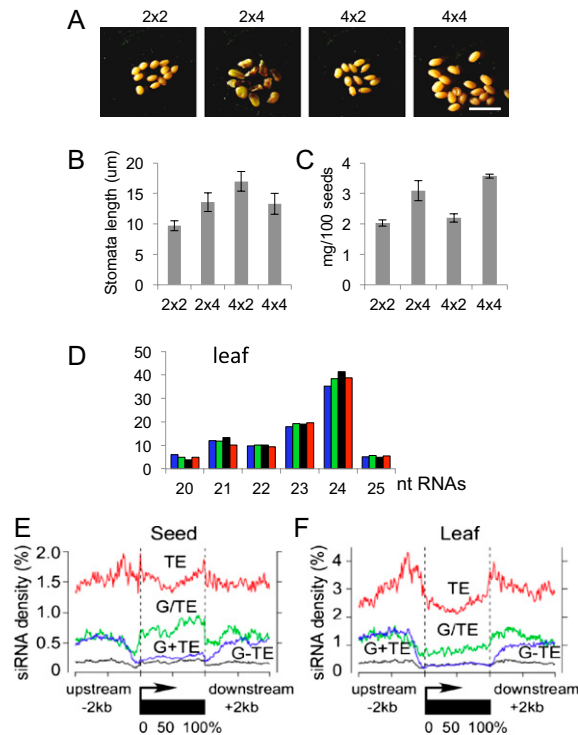


Fig. S1. Seed size variation in ploidy series (Col-0) and small RNA distribution in leaves. (A) Seed size variation in ploidy series of Col. Note that seed abortion in 2×4 triploids in Col is related to maternal expression of *TTG2* during late stage of seed development (1) and possibly other unknown genes. (Scale bar, 1 mm.) (B) Quantification of stomata sizes in ploidy series. (C) Quantification of dry seed weight in ploidy series. (D) Distribution of small RNAs in leaves. (E and F) siRNA densities in TE genes (TE, red), genes containing TEs in 5' upstream (2 kb) and 3' downstream (2 kb) regions (G+TE, blue), and TE-overlapped genes (G/TE, green), and TE-free genes (G-TE, black) in seeds (E) and leaves (F).

1. Dilkes BP, et al. (2008) The maternally expressed WRKY transcription factor *TTG2* controls lethality in interploidy crosses of *Arabidopsis*. *PLoS Biol* 6:2707–2720.

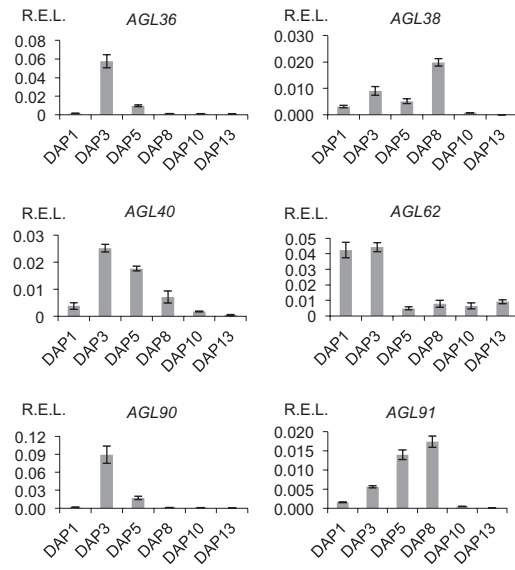


Fig. S4. qRT-PCR analysis (REL, relative expression levels) of siRNA-related *AGL* genes in different stages of developing siliques.

Table S1. Sequence reads of small RNA libraries from *Arabidopsis thaliana* reciprocal interploidy crosses and their parents (rosette leaves)

Genomic feature	2L (%)	4L (%)	2 × 4L (%)	4 × 2L (%)
All reads	8,929,273	3,165,346	11,610,370	11,886,105
Filtered	4,257,442 (48)	895,774 (28)	3,072,403 (26)	4,697,603 (40)
Gene	896,187 (10)	360,142 (11)	1,414,223 (12)	1,153,949 (10)
Transposon	1,297,503 (15)	543,352 (17)	2,606,896 (22)	2,410,799 (20)
miRNA	890,159 (10)	829,203 (26)	2,479,074 (21)	1,030,473 (9)
ta-siRNA	16,333 (0.2)	7,075 (0.2)	46,154 (0.4)	22,157 (0.2)
Intergenic regions	1,571,649 (18)	529,800 (17)	1,991,620 (17)	2,571,124 (22)
Total small RNA	4,671,831 (52)	2,269,572 (72)	8,537,967 (74)	7,188,502 (60)

Rosette leaves in diploid (2L); tetraploids (4L), paternal-excess triploids (2 × 4L), and maternal-excess triploids (4 × 2L); all reads, reads perfectly match the genome; filtered, number of reads that match chloroplast, mitochondrial genomes, and structural noncoding RNAs, including rRNA, tRNA, snoRNA, and snRNA.

Table S2. Sequence reads of small RNA libraries from *Arabidopsis thaliana* reciprocal interploidy crosses and their parents (seeds)

Genomic feature	2S (%)	4S (%)	2 × 4S (%)	4 × 2S (%)
All reads	12,082,171	11,642,856	12,705,829	11,516,592
Filtered	778,465 (6)	1,274,150 (11)	2,637,149 (20)	1,742,706 (15)
Gene	2,330,912 (19)	2,039,493 (18)	1,729,226 (13)	2,029,347 (18)
Transposon	2,069,358 (17)	1,740,812 (15)	1,398,336 (11)	1,814,304 (16)
miRNA	1,227,444 (10)	1,647,975 (14)	2,421,478 (19)	668,594 (6)
ta-siRNA	469 (0.004)	550 (0.004)	395 (0.003)	688 (0.006)
Intergenic regions	5,675,523 (47)	4,939,876 (42)	4,519,245 (36)	5,260,953 (46)
Total small RNA	11,303,706 (94)	10,368,706 (89)	10,068,680 (79)	9,773,886 (85)

Dissected seeds at 6 d after pollination (DAP); seeds in diploid (2S), tetraploids (4S), paternal-excess triploid (2 × 4S), and maternal-excess triploid (4 × 2S); reads classifications are as in Table S1.

Table S3. A list of MADS box genes that possess more than 10 reads from transcribed regions in at least one line

Locus	Gene name	2 × 2	4 × 4	2 × 4	4 × 2	SS, %	P value	TE	Type
AGL91	AT3G66656	15029	17843	10688	22982	89	0.0079	3	I
AGL40	AT4G36590	7941	9395	4834	11665	44	0.0092	4	I
AGL33	AT2G26320	580	797	452	1076	67	0.0540	6	I
AGL36	AT5G26650	980	886	510	1034	30	0.0009	6	I
AGL86	AT1G31630	1000	950	1052	835	40	0.0200	2	I
AGL62	AT5G60440	1121	530	548	820	46	0.0393	5	I
AGL90	AT5G27960	252	252	122	311	36	0.0495	5	I
AGL34	AT5G26580	82	75	39	98	66	0.0849	5	I
AGL87	AT1G22590	117	100	5	84	100	0.2453	5	I
AGL92	AT1G31640	72	43	73	61	54	0.0712	5	I
AGL23	AT1G65360	84	55	40	54	95	0.2088	3	I
AGL28	AT1G01530	28	12	12	23	13	0.0893	3	I
AGL84	AT5G49420	11	9	7	15	83	0.0305	1	I
AGL42	AT5G62165	22	22	42	34	29	0.1165	0	II
AGL66	AT1G77980	32	12	10	13	26	0.0456	3	II
AGL14	AT4G11880	7	29	4	12	68	0.3756	1	II
AGL12	AT1G71692	20	11	24	10	16	0.2452	0	II
AGL11	AT4G09960	10	16	13	6	23	0.3756	4	II

The 24-nt reads were normalized per 10 million; locus, the unique *AGI* identifier; %, percentage of reads from the sense strand; P values, probability of a locus that generates secondary siRNAs by chance; TE, number of transposons in the locus and ±2-kb regions.

Table S4. 5' and 3' adapters

Adapter	Sample	Sequence
5' adapter	2 × 2 seed/leaf	5'-GUUCAGAGUUCUACAGUCCGACGAUCA-3'
	4 × 4 seed/leaf	5'-GUUCAGAGUUCUACAGUCCGACGAU <u>C</u> T-3'
	2 × 4 seed/leaf	5'-GUUCAGAGUUCUACAGUCCGACGAU <u>C</u> C-3'
	4 × 2 seed/leaf	5'-GUUCAGAGUUCUACAGUCCGACGAU <u>C</u> G-3'
3' adapter		5' P-UCGUAUGCCGUCUUCUGCUUGIdT

The bases that were used as a barcode in each 5' adaptor are underlined.

Table S5. qPCR primers

Gene	Forward primer sequence	Reverse primer sequence
AGL23	TGACCACTTTCGAGGGTGTGTG	TTACTCCACCACTCCTCAGCGTTT
AGL28	ACTTGACCACTCTTTAGGGCGTG	ACCACCTCTCAGCGTCTTGTTC
AGL33	TTGTTTCTCCACCGAGAAGCCT	TCACACCTCTCTCCCGTCTTGT
AGL34	GATGCGAATGCAACTGCGGTAAC	ACAAGGTGTTGAAACGGTCGATGC
AGL36	ATGAATCATGTTGGAGGGCG	GCAGTTTCCGTCCACGAAAG
AGL40	GCCATTGTCCATCATCAGAACA	CGGACGGGTTTCAACAAGTT
AGL42	GATCGAACGCTACCGCAAGT	TCGTGATTGCTGGTTTCATGA
AGL62	AATTGGTGGGAAGATCCCGT	CCCCTCGAGTTGAGATAACGC
AGL84	ACCAAACGCCGTGAAGGTCTCTA	AGACAACGGCATCAACAGAGGAG
AGL86	TGAGCATCAGTATGGTGCAAGGG	TGAAGCGGAGGAGGAAAGGAAGA
AGL87	ACCGTTCTATGTGGTTTACCCGCT	TGCTTCTACCGGAAGTTCGCTT
AGL90	ACCAGCCGTTGATCTTGCTT	ACATCGGTGGTTGAAGGCAT
AGL91	AACAAACCGTATTCCTTCGGG	CCGTTCTGCAATCACATCAAA
AGL92	TCATGAGACTTCTTGCGGGACC	TGAAGGTCCCTTGACCATACTGT
FIS2	TCTTGCCCATTTTGCTTGATT	AAGTTGCAAGCCCTCGTGAC
FWA	AGCCTGGTGAGCTAACTGGG	GCCAAACAGAAGTGGATGCAC
MEA	GTTTGGATGATCTGGTCGTGC	CCACTTCGAGGTA
PHE1	GTGGTGTGACGCATGTGC	CCTGGATCGAGTTGTACGGG
GRF1	GGATTAGGCGTCAACACCGA	GTTATTCGCTTTTTCCCGGG
GRF3	TGAGGCCCTTCTTTGACGAT	TGTCAGCTTCTTGGAGCGAA
GRF4	CACCAACCTTCTTGGTATTGGG	CCCTGGCTCAGGATCCATT
NRPD1A	GCGGGTGAGCTGTACTTGA	CTTTTGCCCGATCTCCATA

Table S6. Probe sequences

Gene	Sequence
miR166	GGG GAA TGA AGC CTG GTC CGA
miR172	ATG CAG CAT CAT CAA GAT TCT
miR396	AGTTCAAGAAAGCTGTGGAA
miR832	TTTCGATTCCCGATCCAGCA
AGL91	TTGCCTCTACTATAGCCTGAT; ATTGCCTCTACTATAGCCTGA; TCTTAAACCGTTCTGCAATCACAT; GTTGAGGCGTTTACATATCTTCT; CTTAAACCGTTCTGCAATCACATC; CTATTGCCTCTACTATAGCCTGAT; TTGCCTCTACTATAGCCTGATGTT
AGL40	ATACCAGTTTCTACTTGTCTCT; ATGCCAATTCTATGTCTATAACTT; TTCTGCCATTGTCCATCATCA; AACCATAATTCTGCCATTGTCCAT; TTCATCGCTTTTTCGAATCCTA; TGACGACCTTGGTACTTCT
U6	GCTAATCTTCTGTATCGTTCC

Other Supporting Information Files

[Dataset S1 \(XLS\)](#)

[Dataset S2 \(XLS\)](#)

[Dataset S3 \(XLS\)](#)

[Dataset S4 \(XLS\)](#)