Science Advances
AAAS

advances.sciencemag.org/cgi/content/full/5/5/eaav9040/DC1

Supplementary Materials for

ALBA protein complex reads genic R-loops to maintain genome stability in *Arabidopsis*

Wei Yuan, Jincong Zhou, Jinjin Tong, Wanqing Zhuo, Lishuan Wang, Yan Li, Qianwen Sun*, Weiqiang Qian*

*Corresponding author. Email: sunqianwen@mail.tsinghua.edu.cn (Q.S.); wqqian@pku.edu.cn (W.Q.)

Published 15 May 2019, *Sci. Adv.* **5**, eaav9040 (2019)
DOI: 10.1126/sciadv.aav9040

The PDF file includes:

Other Supplementary Material for this manuscript includes the following:

(available at advances.sciencemag.org/cgi/content/full/5/5/eaav9040/DC1)

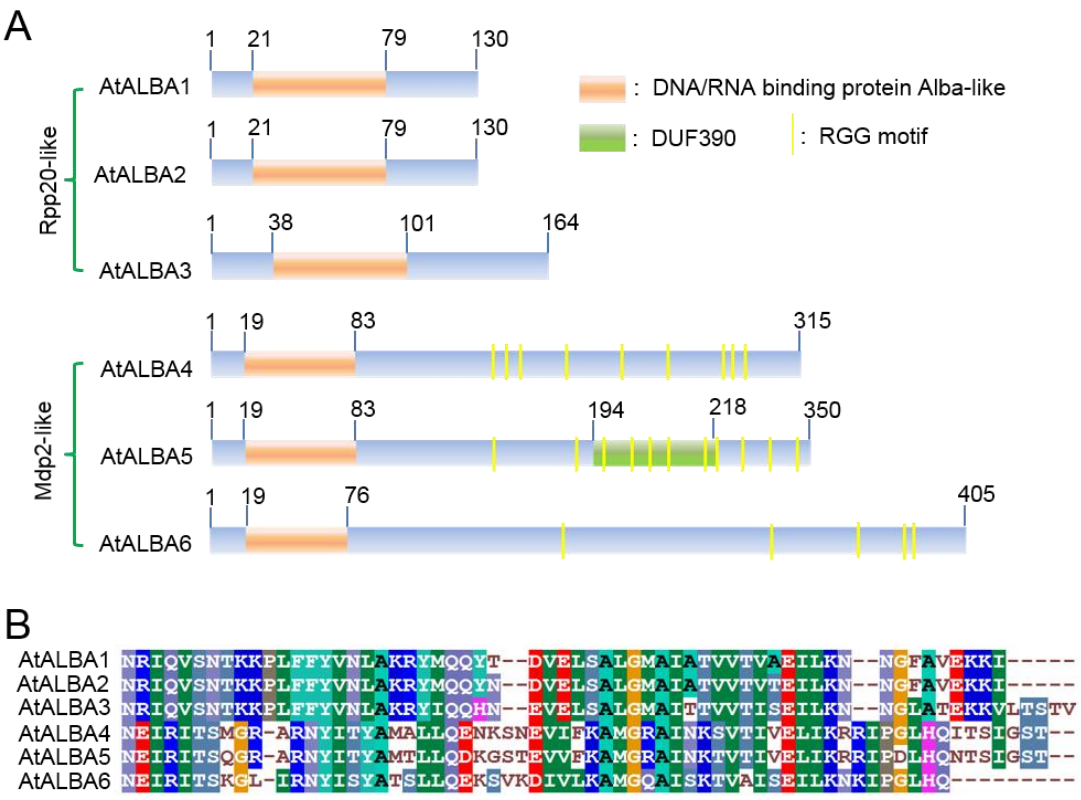


Fig. S1. Domain structure of ALBA proteins in *Arabidopsis*. (**A**) Domain structure of AtALBA proteins from the Pfam database. AtALBA1, AtALBA2 and AtALBA3 belong to the Rpp20-like subfamily and contain only an Alba-like DNA/RNA-binding domain. AtALBA4, AtALBA5 and AtALBA6 belong to the Mdb2-like subfamily. In addition to an Alba-like domain at the N-terminal, Mdb2-like subfamily members contain DUF390 and RGG motifs at the C-terminal. The numbering refers to the amino acid position. (**B**) Amino acid sequence alignment of the Alba-like DNA/RNA-binding domains of AtALBAs.

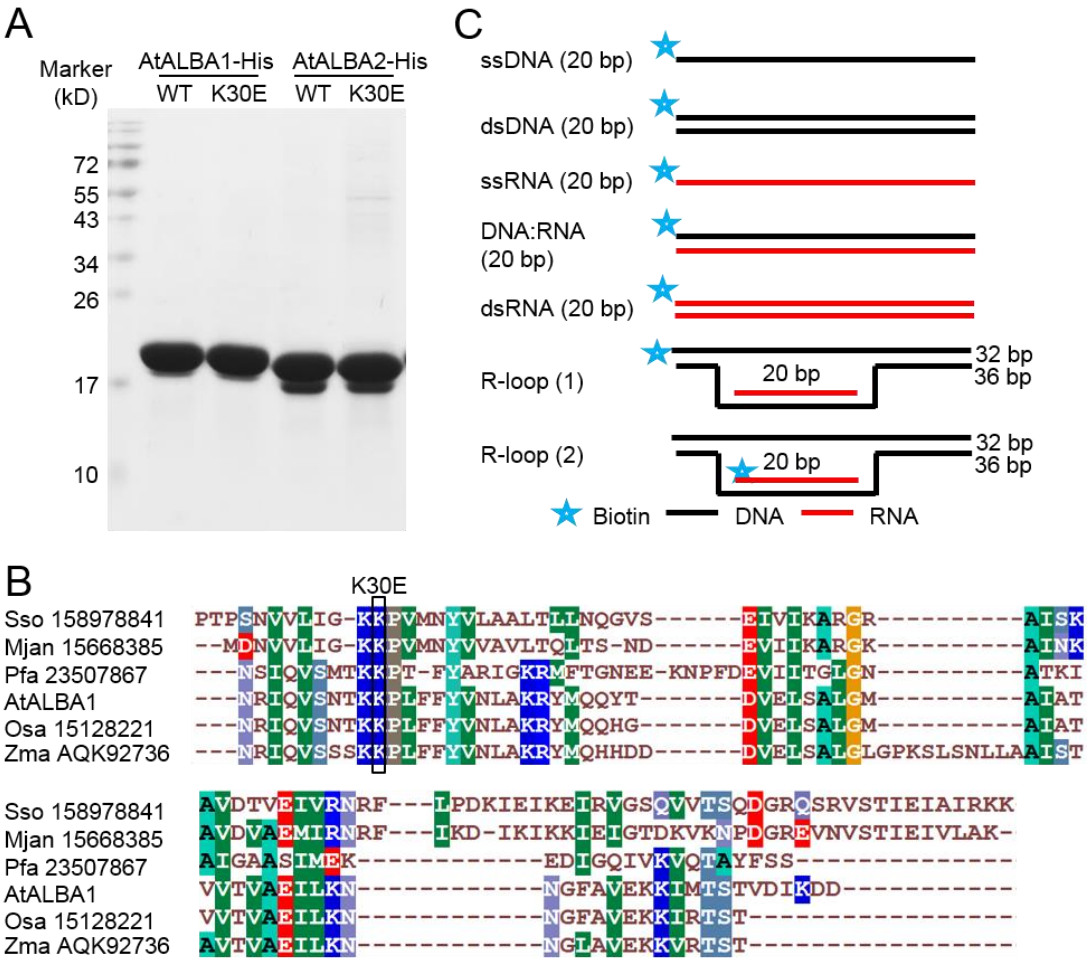
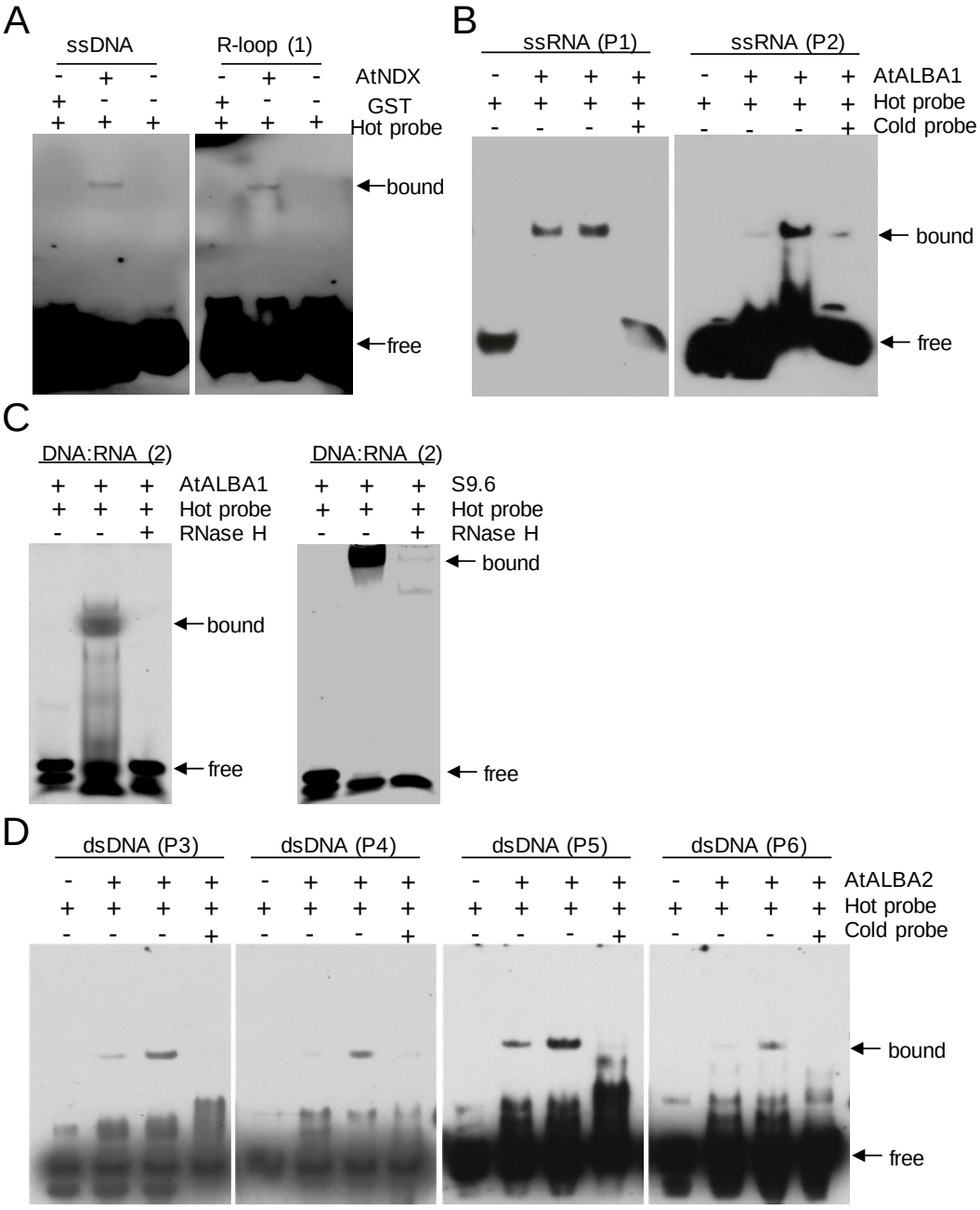


Fig. S2. Purification of AtALBA1 and AtALBA2 wild-type and mutant proteins and diagram of probes used in EMSAs. (**A**) Coomassie blue-stained gel showing purified AtALBA1 and AtALBA2 wild-type and mutant proteins. Recombinant AtALBA1-His and AtALBA2-His were purified by Ni-NTA affinity chromatography. (**B**) Amino acid sequence alignment of the Alba-like DNA/RNA-binding domains from Alba proteins in different species. The conserved lysine (K), which is critical for DNA binding, is indicated in a black box. (**C**) Diagram of different substrates used in EMSA.


A
ssDNA
R-loop (1)
AtNDX
GST
Hot probe
bound
free
B
ssRNA (P1)
ssRNA (P2)
AtALBA1
Hot probe
Cold probe
bound
free
C
DNA:RNA (2)
DNA:RNA (2)
AtALBA1
Hot probe
RNase H
S9.6
Hot probe
RNase H
bound
free
bound
free
D
dsDNA (P3)
dsDNA (P4)
dsDNA (P5)
dsDNA (P6)
AtALBA2
Hot probe
Cold probe
bound
free

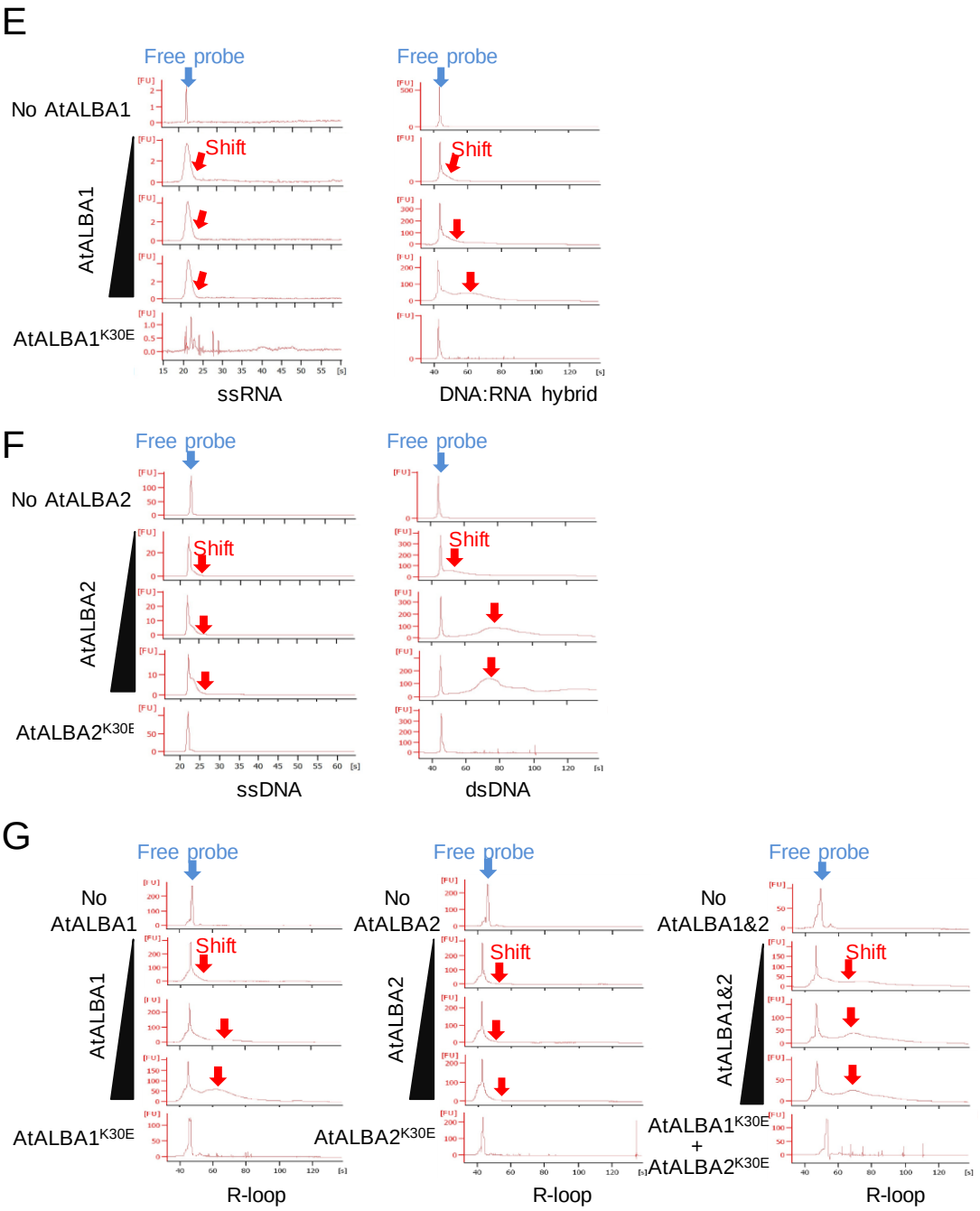


Fig. S3. Characterization of the nucleic acid binding properties of AtALBA1 and AtALBA2. (**A**) EMSA gel showing AtNDX (16 nM) binding to single-stranded DNA probes (5 nM) and artificial R-loops with 5'-biotin-labeled DNA (1) (5 nM). 80 nM of GST proteins were used for negative control. (**B**) EMSA gel showing AtALBA1 (25, 50, 75 nM) binding to single-stranded RNA probes (5 nM) and competition by unlabeled RNA probes (2 μM). Two randomly selected sequences were used in this assay. (**C**) EMSA gel showing AtALBA1 (75 nM) binding to the DNA-RNA hybrid probe (5 nM). The DNA-RNA hybrid probes were incubated with RNase H1 for 0 min or 10 min. S9.6 antibody was used as a positive control. (**D**) EMSA gel showing AtALBA2 (25, 50, 75 nM) binding to dsDNA probes (5 nM) and competition by unlabeled DNA probes (2

μM). Four randomly selected sequences were used in this assay. (**E-G**) Quantification of the affinities of proteins (25, 50, 75 nM) towards different types of nucleic acids (2 μM) by Agilent 2100 BioAnalyzer. Sharp peaks (blue arrow) indicate free probe and broad peaks (red arrow) indicate formation of nucleic acids/protein complexes. 75 nM of K30E mutant proteins were used as negative controls. For the right lane in S3G, AtALBA1 and AtALBA2 were mixed in the mole ratio 1:1.

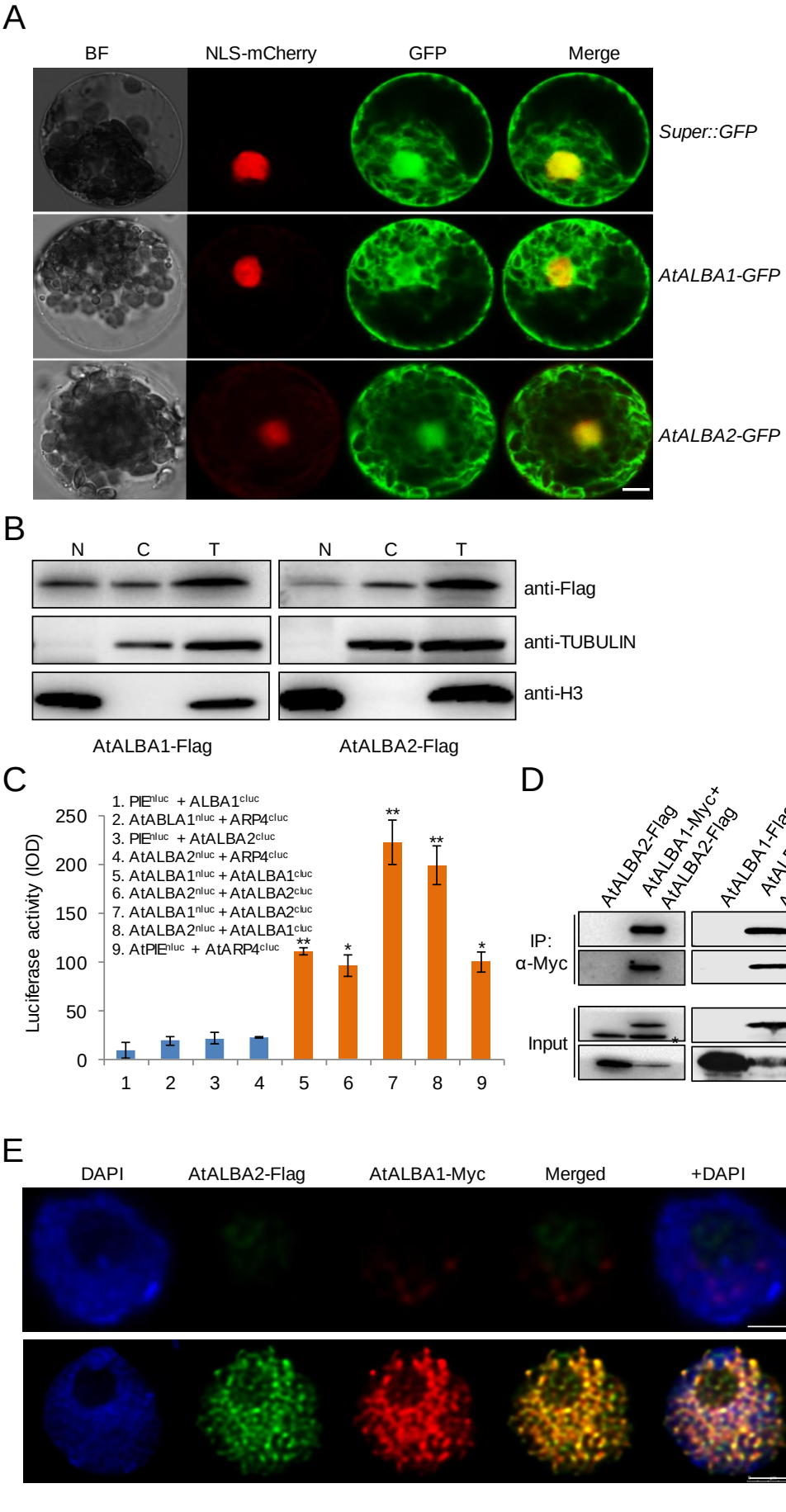

A
BF
NLS-mCherry
GFP
Merge
Super::GFP
AtALBA1-GFP
AtALBA2-GFP
B
N
C
T
N
C
T
anti-Flag
anti-TUBULIN
anti-H3
AtALBA1-Flag
AtALBA2-Flag
C
1. PIE^nluc + ALBA1^cluc
2. AtABLA1^nluc + ARP4^cluc
3. PIE^nluc + AtALBA2^cluc
4. AtALBA2^nluc + ARP4^cluc
5. AtALBA1^nluc + AtALBA1^cluc
6. AtALBA2^nluc + AtALBA2^cluc
7. AtALBA1^nluc + AtALBA2^cluc
8. AtALBA2^nluc + AtALBA1^cluc
9. AtPIE^nluc + AtARP4^cluc
Luciferase activity (IOD)
250
200
150
100
50
0
1
2
3
4
5
6
7
8
9
D
AtALBA2-Flag
AtALBA1-Myc+
AtALBA2-Flag
AtALBA1-Flag
AtALBA1-Flag+
AtALBA1-Myc
IP:
α-Myc
α-Myc
α-Flag
Input
α-Myc
α-Flag
E
DAPI
AtALBA2-Flag
AtALBA1-Myc
Merged
+DAPI
100%
n=180
Col-0
92%
n=101
F1

Fig. S4. Subcellular localization and interaction of AtALBA1 and AtALBA2. (**A**) Localization of AtALBA1-GFP and AtALBA2-GFP in *Arabidopsis* protoplasts. NLS-mCherry (red) was used to label the nuclei. Bars = 10 μm. (**B**) Western blot analysis of AtALBA1-Flag and AtALBA2-Flag in total protein extract, nuclear and cytoplasmic fractions. Transgenic *AtALBA1-Flag/alba1-1* and *AtALBA2-Flag/alba2-1* plants were used. *AtALBA1-Flag* and *AtALBA2-Flag* expression were under the control of their respective native promoters. Histone H3 and Tubulin were used as nuclear and cytoplasmic markers, respectively. T: total protein extract, N: nuclear fraction, C: cytoplasmic fraction. (**C**) Split luciferase complementation results showing the interaction between AtALBA1 and AtALBA2 in *Nicotiana benthamiana* leaves. Standard errors were calculated from three biological replicates, *p<0.05, **p<0.01, two-tailed Student's *t*-test. (**D**) Co-immunoprecipitation (co-IP) results showing the interaction between AtALBA1 and AtALBA2. Transgenic plants expressing *AtALBA1-Flag*, *AtALBA2-Flag* and *AtALBA1-Myc* under their native promoters and their F1 offspring were used for co-IP. The asterisk indicates a non-specific band. (**E**) Co-localization of AtALBA1 and AtALBA2 in the nucleus. AtALBA1 and AtALBA2 were stained using anti-Myc (red) and anti-Flag (green), respectively, in Col-0 plants (upper panel) or F1 heterozygous plants expressing AtALBA1-Myc and AtALBA2-Flag under their native promoters (lower panel). DNA was stained with DAPI (blue). The frequency of nuclei displaying each interphase pattern is shown on the right. Bars = 2.5 μm.

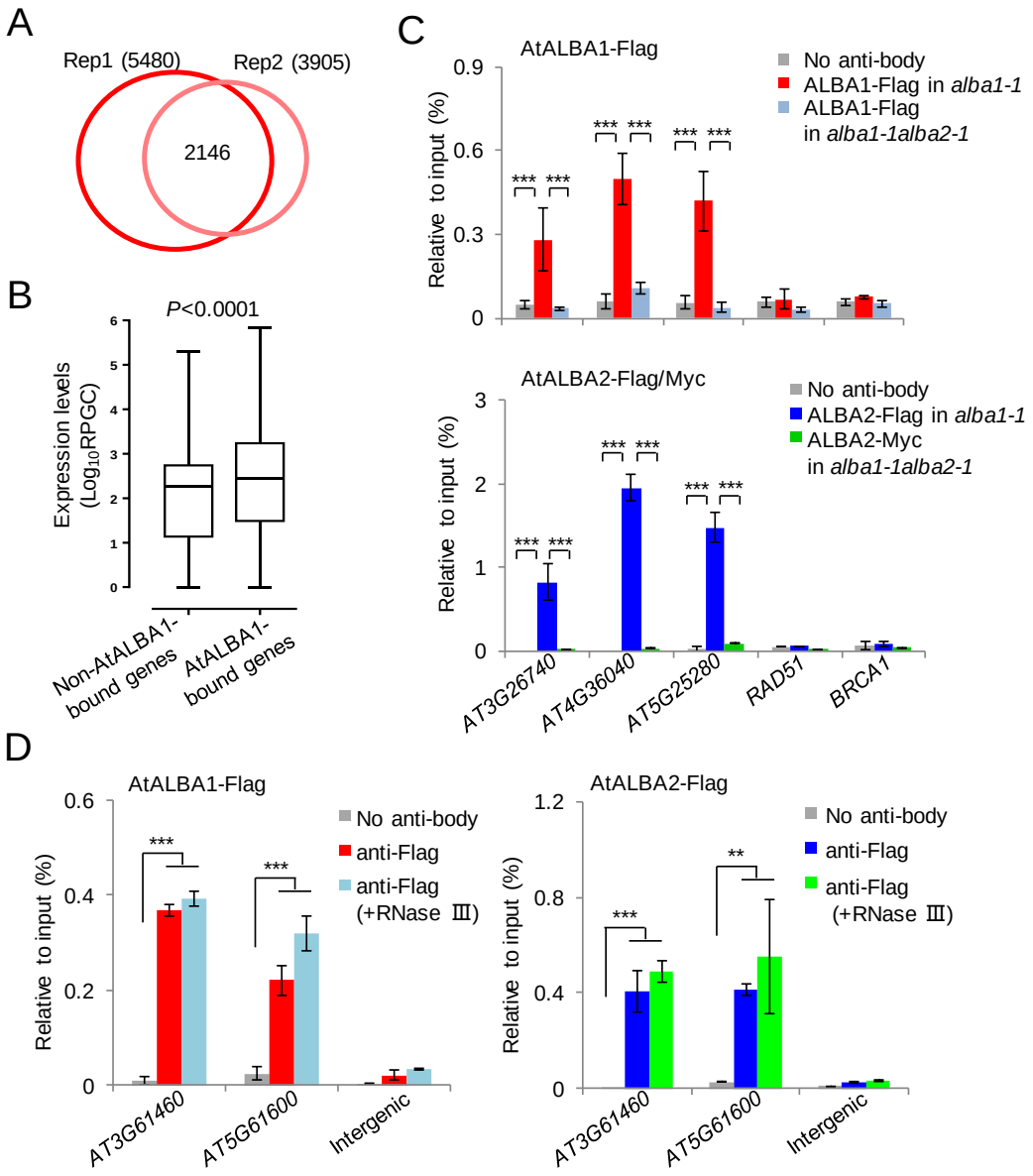


Fig. S5. Characterization of AtALBA1-bound loci. (**A**) Venn diagram showing the overlap of AtALBA1 ChIP-seq peaks in two biological replicates. (**B**) Box plots showing expression levels of non-AtALBA1-bound genes and AtALBA1-bound genes. RPGC: reads per genomic content (1 x normalization). (**C**) Association of AtALBA1 and AtALBA2 with specific genic regions. Transgenic *AtALBA1-Flag/alba1-1*, *AtALBA1-Flag/alba1-1alba2-1*, *AtALBA2-Flag/alba2-1* and *AtALBA2-Myc/alba1-1alba2-1* plants were used for ChIP assays. Expression of *AtALBA1-Flag* and *AtALBA2-Flag/Myc* were under the control of their respective native promoters. ChIP experiments were performed with anti-Flag or anti-Myc. Two biological replicates were performed,

and very similar results were obtained. Standard errors were calculated from three technical replicates, *p<0.05, **p<0.01, ***p<0.001 (two tailed Student *t*-Test). (**D**) Association of AtALBA1 and AtALBA2 with specific genic regions. Transgenic *AtALBA1-Flag/alba1-1* and *AtALBA2-Flag/alba2-1* plants were used. Expression of *AtALBA1-Flag* and *AtALBA2-Flag* was under the control of their respective native promoters. ChIP experiments were performed with anti-Flag antibody. The RNase III treatment was performed before cross-linking. Two biological replicates yielded very similar results. Standard errors were calculated from three technical replicates, *p<0.05, **p<0.01, ***p<0.001, two-tailed Student's *t*-test.

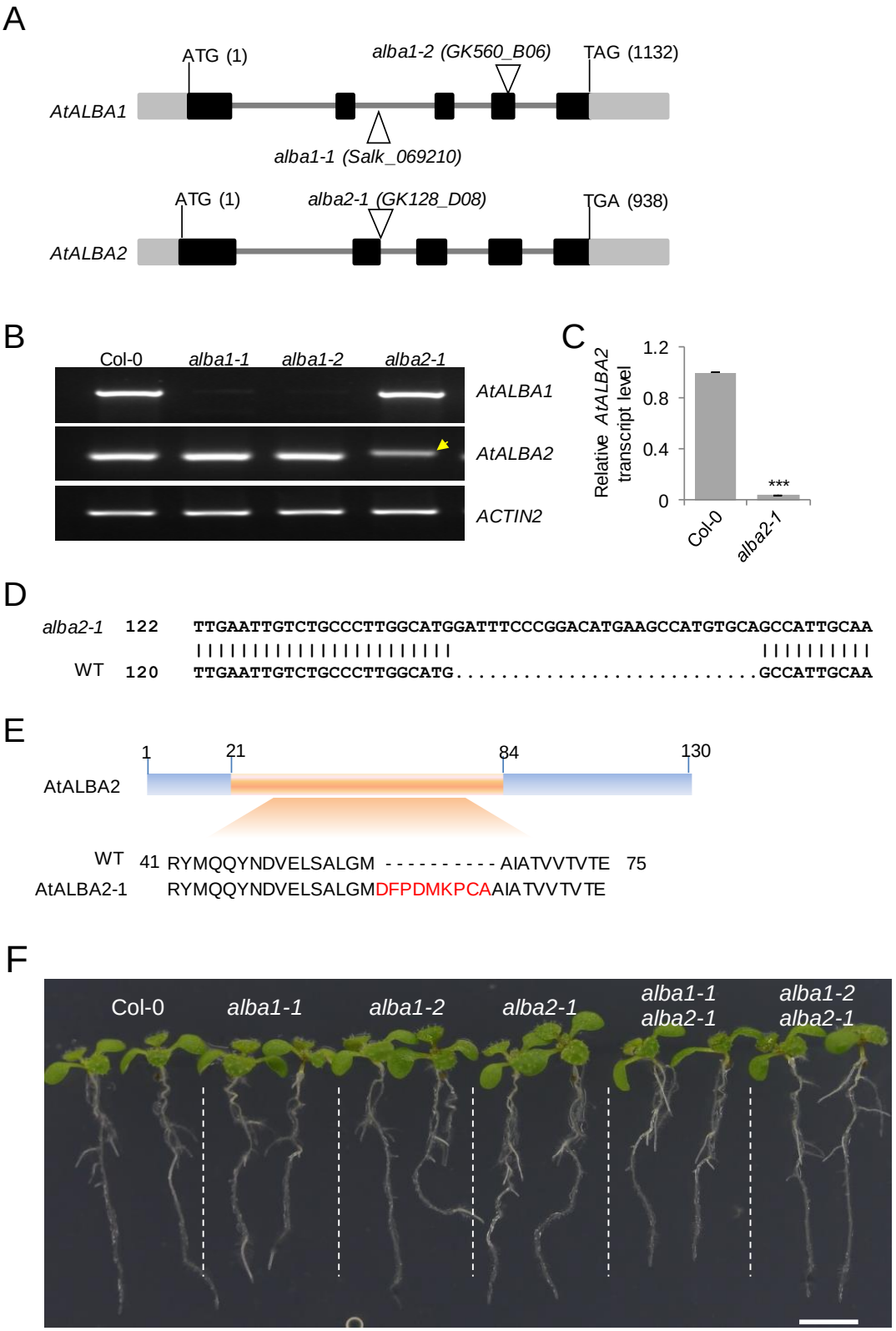
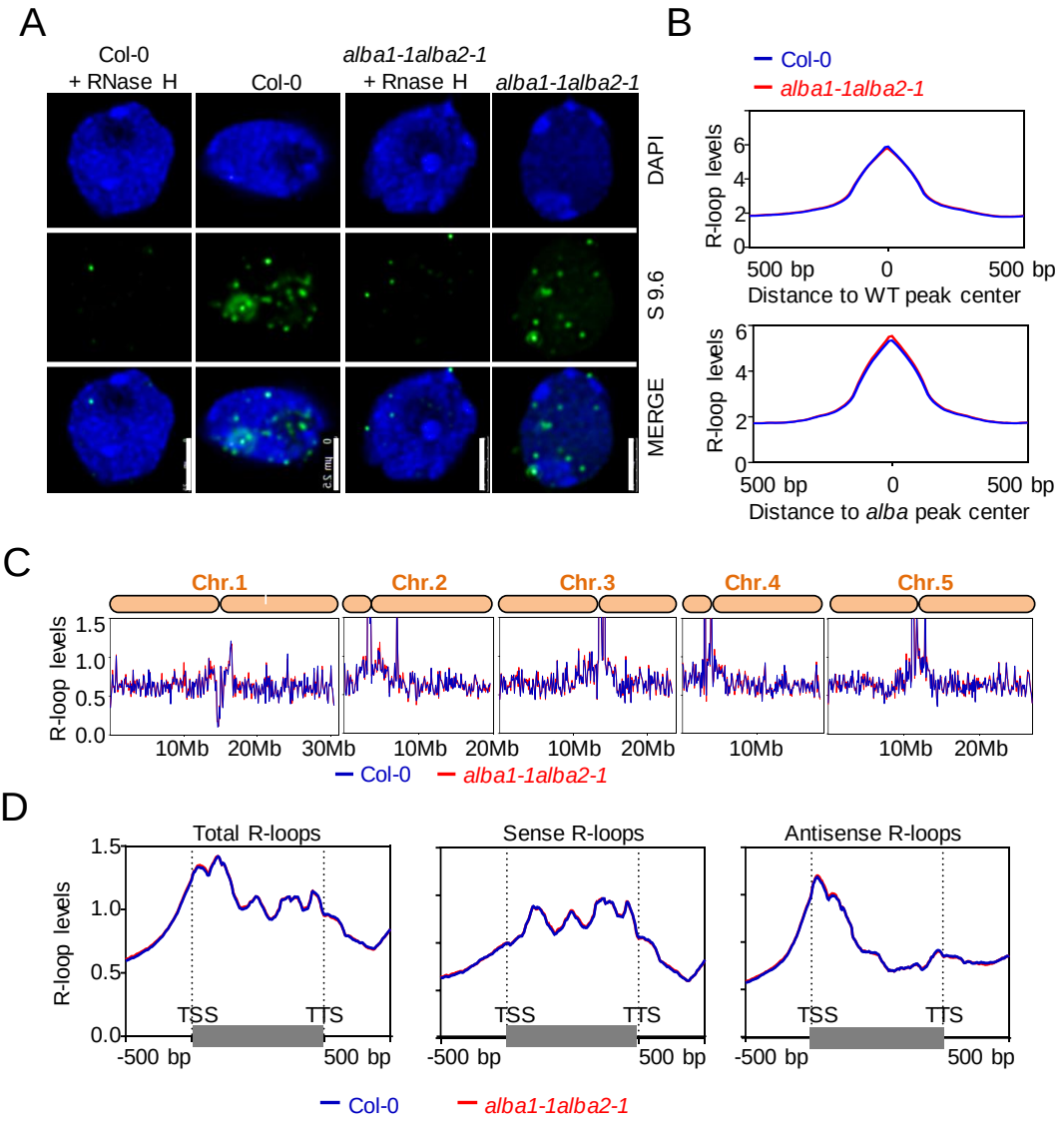


Fig. S6. Characterization of the T-DNA insertion mutants for *AtALBA1* and *AtALBA2*. (**A**) Schematic diagram showing the positions of T-DNA insertion mutations at the *AtALBA1* and *AtALBA2* loci. Black rectangles represent exons. Gray rectangles represent UTRs. (**B**) RT-PCR analysis of *AtALBA1* and

AtALBA2 mRNA levels in the *alba1 and alba2* mutants, respectively. The yellow arrow indicates a mutant transcript of *AtALBA2* in *alba2-1*. (**C**) Relative expression levels of AtALBA2 in *alba2-1*. Standard divisions were calculated from three technical replicates, ***p<0.001 (two tailed Student *t*-Test). (**D**) Sequence alignment of the mutant transcript of *AtALBA2* in *alba2-1* with wild-type *AtALBA2* CDS. The nucleotide insertion carried by the mutant transcript is from the T-DNA vector backbone. (**E**) Amino acid sequence alignment showing a 9-amino acid insertion in the Alba-like DNA/RNA-binding domain of *AtALBA2* in *alba2-1*. (F) Developmental phenotype ofCol-0 and the indicated mutants. Bars = 0.5 cm.



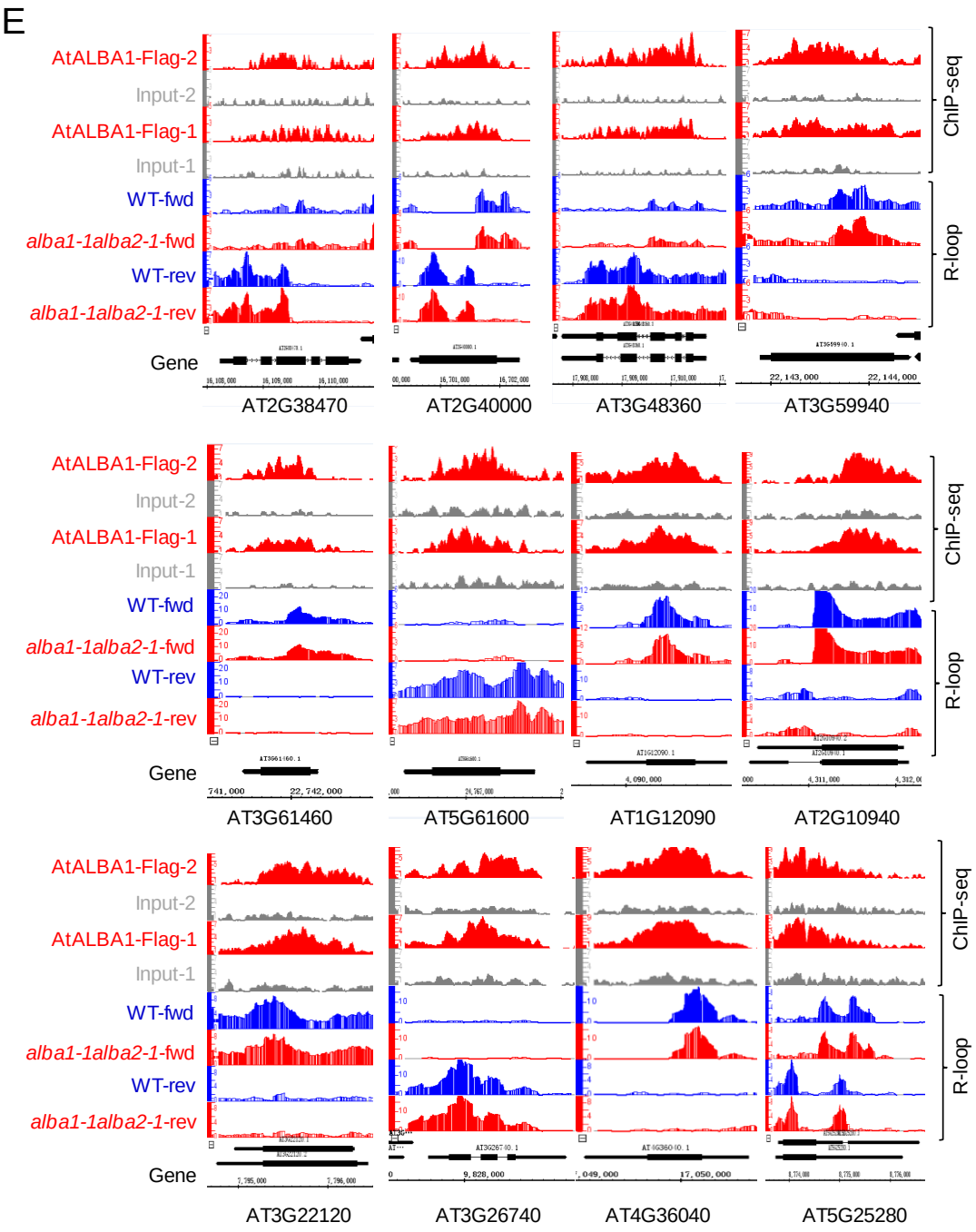
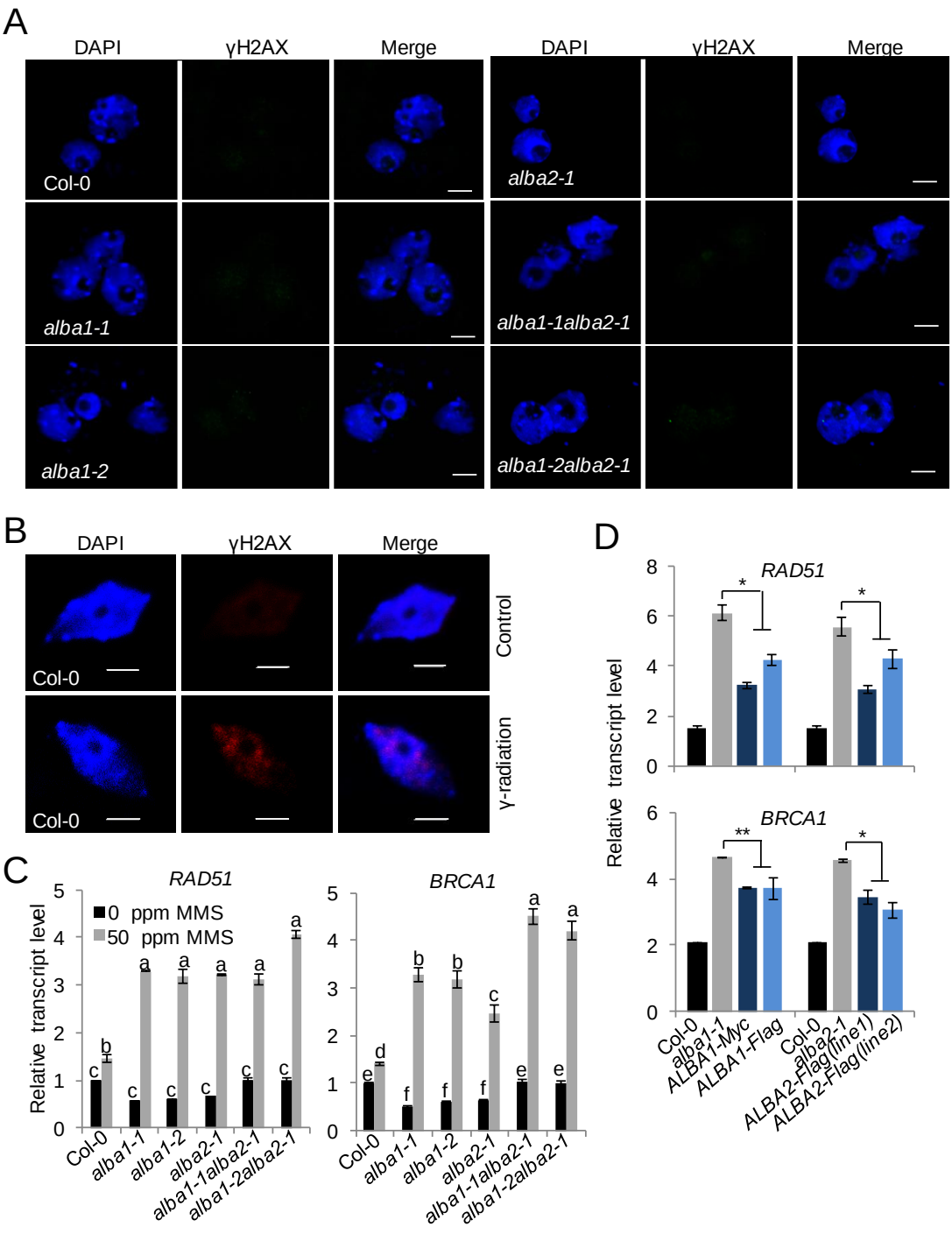
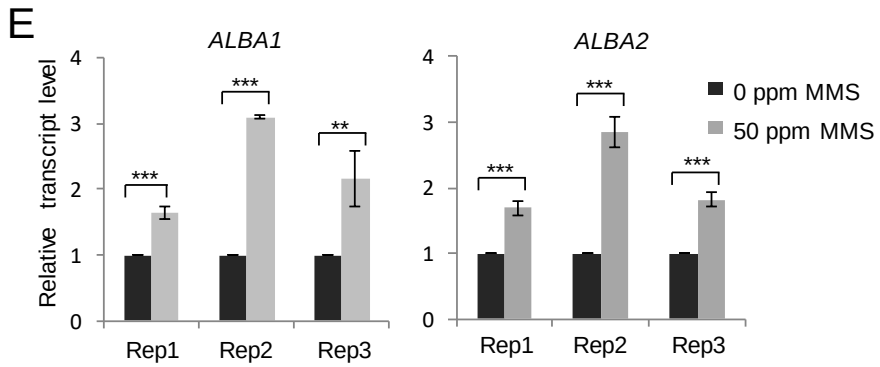
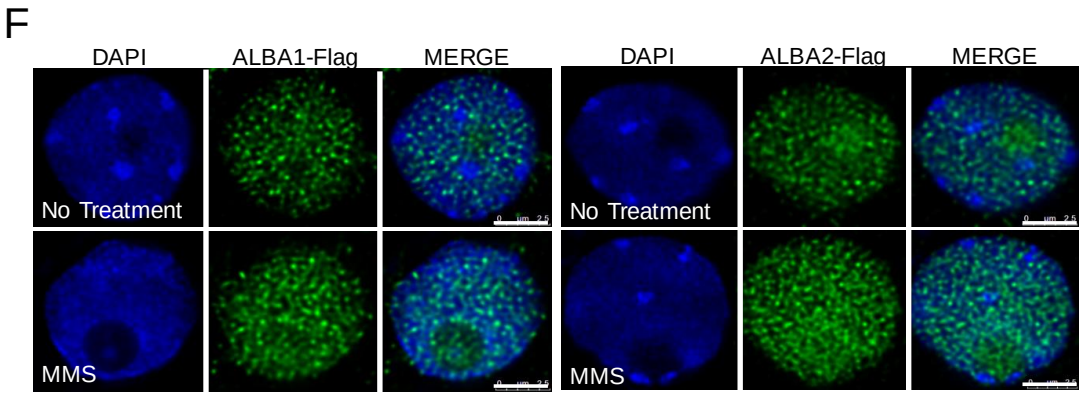


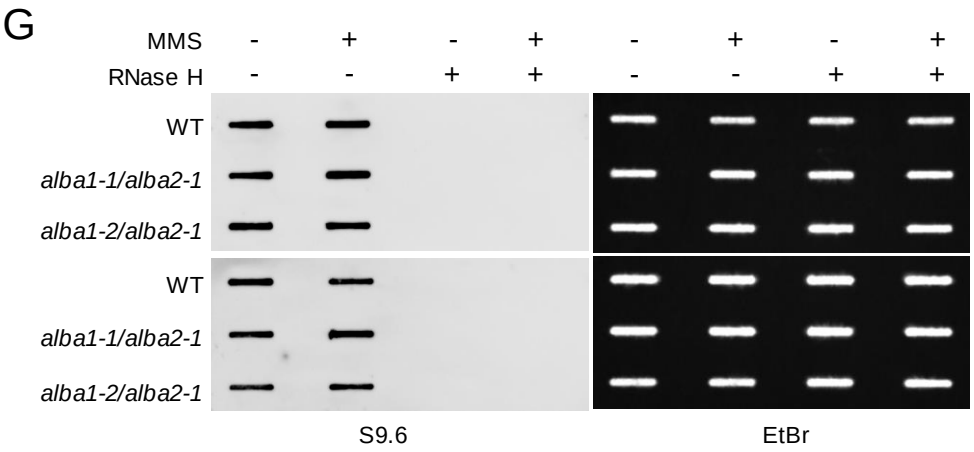
Fig. S7. Detection of R-loop levels in Col-0 and *alba1-1alba2-1* by immunostaining and ssDRIP-seq. (**A**) Immunostaining detection of R-loops in Col-0 and *alba1-1alba2-1*. R-loops in individual cells were stained with S9.6 antibodies (green). DNA was stained with DAPI (blue). At least 30 cells were stained for each sample. Bars = 2.5 μm. (**B**) Metaplot of R-loop levels in Col-0 and *alba1-1alba2-1* as determined by ssDRIP-seq. 0 on the X-axis indicates the R-loop peak center. The Y-axis indicates ssDRIP-seq read density. (**C**) R-loop levels on five chromosomes in Col-0 and *alba1-1alba2-1*. Chr. 1, Chr. 2, Chr. 3, Chr. 4, and Chr. 5 represent chromosomes 1, 2, 3, 4 and 5, respectively. (**D**) Metaplot of R-loop levels on AtALBA1-bound genes in Col-0 and

alba1-1alba2-1. The gray bar on the X-axis indicates the genes. -500 bp and +500 bp represent 500 bp upstream of the TSS and 500 bp downstream of the TTS, respectively. (**E**) Snapshots of genomic regions with AtALBA1 enrichment and R-loops.




E
ALBA1
ALBA2
Relative transcript level
0 ppm MMS
50 ppm MMS
Rep1
Rep2
Rep3


F
DAPI
ALBA1-Flag
MERGE
DAPI
ALBA2-Flag
MERGE
No Treatment
MMS


G
MMS
RNase H
WT
alba1-1/alba2-1
alba1-2/alba2-1
S9.6
EtBr

H

samples	Clean Reads	Unique Map Reads	Map Rate	Average length	Average quality
Col-0_γH2AX_ChIP_seq1	43938734	28428898	0.65	127	37.8
Col-0_γH2AX_ChIP_seq2	47405847	30849069	0.65	126	37.7
*alba1-1alba2-1*_γH2AX_ChIP_seq1	31057693	19861579	0.64	126	37.8
*alba1-1alba2-1*_γH2AX_ChIP_seq2	28078164	18233775	0.65	127	37.6
Col-0_H3_ChIP_seq1	34603561	23151857	0.67	127	37.7
Col-0_H3_ChIP_seq2	40691246	27127269	0.67	127	37.8
*alba1-1alba2-1*_H3_ChIP_seq1	37555884	24183416	0.64	126	37.7
*alba1-1alba2-1*_H3_ChIP_seq2	34698837	22893684	0.66	126	37.7

I

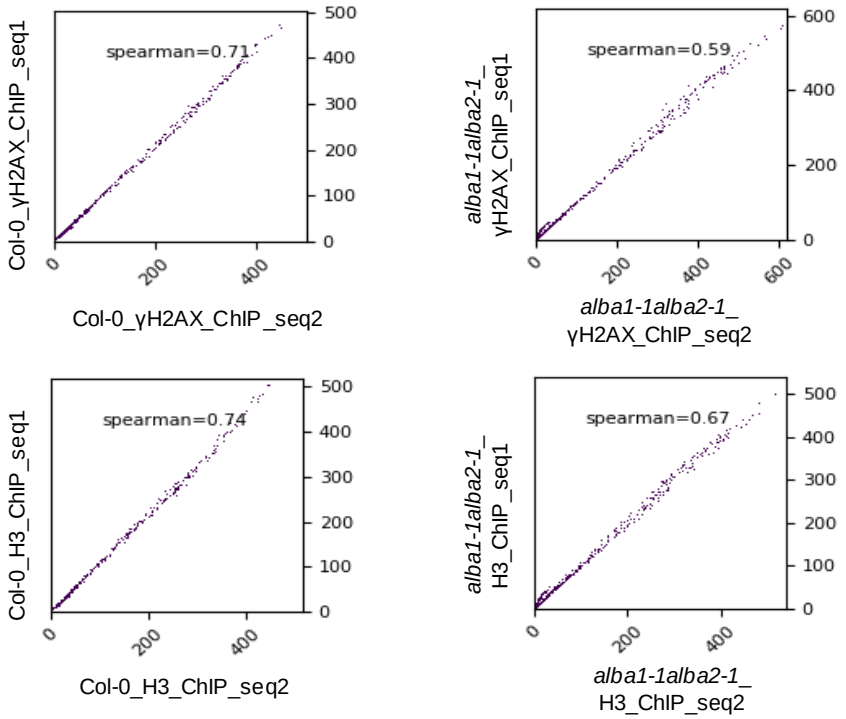


Fig. S8. Molecular phenotypes of Col-0 and alba1-1alba2-1 without and with MMS treatment. (**A**) Representative microscopic images showing γH2AX foci formation (green) in Col-0, *alba1-1*, *alba1-2*, *alba2-1*, *alba1-1alba2-1,* and *alba1-2alba2-1* plants under normal growth conditions. γH2AX foci were detected by immunostaining using anti-γH2AX antibody. Nuclei were stained with DAPI (blue). Bars = 5 μm. (**B**) Immunofluorescence showing elevated γH2AX signal in Col-0 after 5 min of γ-irradiation (100 Gy). Bars = 2 μm. **(C and D)** Relative expression levels of *RAD51* and *BRCA1* in Col-0 and the indicated mutants. After growth under normal conditions for 5 d, plants were treated with 50 ppm MMS for 48 h. Two biological replicates were performed, and very similar results were obtained. The multiple comparison in S8C was calculated with Tukey. The level by alpha default is 0.05. Standard errors in S8D were calculated from three technical replicates, *p<0.05, **p<0.01, ***p<0.001,

two-tailed Student's *t*-test. **(E)** Relative expression levels of *AtALBA1* and *AtALBA2* in Col-0 plants grown on 1/2 MS medium supplemented with 0 or 50 ppm MMS. Three biological replicates were performed. Standard errors were calculated from three technical replicates, **p<0.01, ***p<0.001, two-tailed Student's *t*-test. (**F**) Localization pattern of AtALBA1 and AtALBA2 in the nucleus without and with MMS treatment. AtALBA1 and AtALBA2 were stained using anti-Flag (green), under normal conditions (upper panel) or with 50 ppm MMS treatment (lower panel). DNA was stained with DAPI (blue). Bars = 2.5 µm. (**G**) Slot blot showing total R-loop levels in Col-0 and the indicated mutants without and with MMS treatment. 100 ng of genomic DNA were used for blotting. Results from two replicates were presented. (**H**) Information for γH2AX ChIP-seq data, including number of reads after adapter cut, unique mapped reads and rate, average length and quality of reads. **(I)** Scatterplot of replicates using bigwig files and deepTools. Spearman's correlation coefficients are presented.