Supporting Information Legends

Figure S1. Nuclear distribution of TE-lincRNAs and non-TE-lincRNAs in Arabidopsis (A), rice (B), and maize (C). Each chromosome was separated into 100 kb bins. The lines on top of each chromosome represent frequencies of TE-lincRNA (red) and non-TE-lincRNAs (green) in each bin. The heatmaps at the bottom of each chromosome represent the frequencies of TEs (red) and genes (blue) in each bin. **(D)** Pairwise correlations of densities of TE-lincRNAs, non-TE-lincRNAs, genes and TEs.

Figure S2. Length distribution of TE-lincRNAs and non-TE-lincRNAs identified from Arabidopsis (A), rice (B) and maize (C).

Figure S3. Exon numbers of TE-lincRNAs and non-TE-lincRNAs identified from Arabidopsis (A), rice (B) and maize (C).

Figure S4. Numbers of RNA motifs detected in TE-lincRNAs and non-TE-lincRNAs from Arabidopsis (A), rice (B) and maize (C).

Figure S5. Distribution of the distance of lincRNAs to their corresponding nearby genes in Arabidopsis (A), rice (B) and maize (C). X axis represents the relative distance of lincRNAs with respect to the coordinate of the 5' end (Sense) or 3' end (Antisense) of neighbouring genes.

Figure S6. Correlation of expression between lincRNAs and 10 closest genes and the example of lincRNA (A) and protein-coding gene co-expression network showing stress response (B and C). (A) Pearson's correlation of the expression between lincRNAs and 10 closest genes across 15 different samples. Red bar in the middle represents TE-lincRNAs, and black bar represents non-TE-lincRNAs. Grey colour in the heatmap represents non-significant correlation (*p*-value > = 0.05). (B) Expression pattern of lincRNAs and protein-coding genes in module/sub-network "paleturquoise" is shown in the top panel, and the bar plot in the bottom panel shows the eigengene values for different samples. Green represents "under-expressed" and red represents "over-expressed" in the heatmap. The "eigengene value" is defined as the first principal component of this module, so it can be considered as the representative of the gene expression profiles in this module. (C) Visualization of co-expression module/sub-network "paleturquoise". And label(s) in black colour represent protein-coding genes and label(s) in blue colour represent TE-lincRNA. The size of the node/label and edge weight is proportional to betweenness centrality.

Figure S7. Contribution of TEs to lincRNAs in Arabidopsis, rice and maize. (A) Distribution of numbers of TEs in TE-lincRNAs. **(B)** TE-lincRNAs with multiple TEs have longer sequences. **(C)** Distribution of coverage of lincRNAs covered by TEs in TE-lincRNAs. **(D)** Distribution of coverage of TEs overlapped by TE-lincRNAs. **(E)** Positional distribution of TEs in TE-lincRNAs with respect to their host TE-lincRNAs. X axis represents the relative coordinates of the 5' end of TEs compared to the 5' end of TE-lincRNAs, and the Y axis represents the relative coordinates of the 3' end of TEs compared to the 3' end of TE-lincRNAs. Therefore, each dot represents one TE-lincRNA. There is no dot located in the top left corner, because this region represents lincRNAs fully nested in TEs, which were filtered out by our pipeline.

Figure S8. Comparison of lincRNAs and protein-coding genes contributed by TEs in Arabidopsis, rice and maize. (A) Percentage of lincRNAs or protein-coding genes contributed by TEs. **(B)** Box plot showing difference of TE coverage in lincRNAs and protein-coding genes.

Figure S9. Mutated TE-lincRNA11195 alters sensitivity to ABA in seed germination and post-germination development in Arabidopsis. (A) Seed germination of WT, *lincRNA11195-1* and *lincRNA11195-2* mutants. Seeds (90 per genotype) were grown on half MS medium supplemented with different amounts of ABA. Seeds were vernalized at four degree for three days, and germinations of seeds were determined when they had grown for three days. Seeds were considered as germinated once radicle penetrated the seed coat. **(B)** Post-germination seedling development of WT, *lincRNA11195-1* and *lincRNA11195-2* under 0 or 1 μM ABA. **(C)** Primary root length and fresh shoot weight of seedlings shown in **B**. Graphs are presented as the percentage relative to growth on control half MS media. ABA assays were performed by stratifying seeds at 4 °C for 3 days, seedlings were grown for eleven days on half MS medium supplemented with or without 1 μM ABA. Error bars stand for standard deviation (n = 20).

Figure S10. Expression of Arabidopsis TE-lincRNA11195 in WT and an ABA-insensitive mutant *pyr1/pyl1/pyl4*. Transcription of an ABA responsive gene RD29A is also shown.

Figure S11. Mutated TE-lincRNA11195 alters sensitivity to salt in seed germination and post-germination seedling development in Arabidopsis. (A) Expression of lincRNA11195 and its neighbouring gene under different conditions. **(B)** Seed germination of WT, *lincRNA11195-1* and *lincRNA11195-2* mutants. Seeds (90 per genotype) were grown on half MS medium supplemented with different amounts of salt. **(C)** The percentage of green seedlings in **B**. Seedlings were grown on half MS medium with 150 mM NaCl for eleven days, and the number of seedlings with green cotyledons were measured. *: $p < 0.05$ (Two-sample Independent t test).

Figure S12. TE strengthens the ABA-responsive transcription to lincRNA11195 in Arabidopsis. The Col-0 lincRNA11195 DNA sequence with (lincRNA11195_C) or without (lincRNA11195_noTE) the LTR driven by its native promoter was introduced into *11195-1* (Right panel) and *11195-2* (Left panel), respectively. LincRNA11195 transcript levels after 12 h of 100 μM ABA treatment were analysed by qRT-PCR, and WT used in this assay was that Col-0 plants transformed empty vector. The experiment was performed with T1 plants, and it was carried out with at least three biological replicates, and each replicate contained at least three seedlings.

Figure S13. TE-lincRNA11195 is detected only in Arabidopsis genus. PCR amplification of TE-lincRNA11195 was carried out in *Arabidopsis thaliana*, *Arabidopsis lyrata* and *Capsella rubella*. Template cDNA, RNA (no RT) or DNA were used in the reactions.

Figure S14. Sequence pairwise alignment and secondary structure prediction of TE-lincRNA11195 in *A. thaliana* and *A. lyrata*. (A) Sequence pairwise alignment of lincRNA11195 between *A. thaliana* and *A. lyrata*. Secondary structures of lincRNA11195 of *A. thaliana* and *A. lyrata* were shown in (B) and (C) respectively. Colour gradient represents base-pair probabilities.

Table S1. General information about RNA sequence libraries used in this study.

Table S2. Genomic coordinates of TE-lincRNAs in three species.

Table S3. Summary of Rfam RNA motifs detected in TE-lincRNAs and non-TE-lincRNAs from three species.

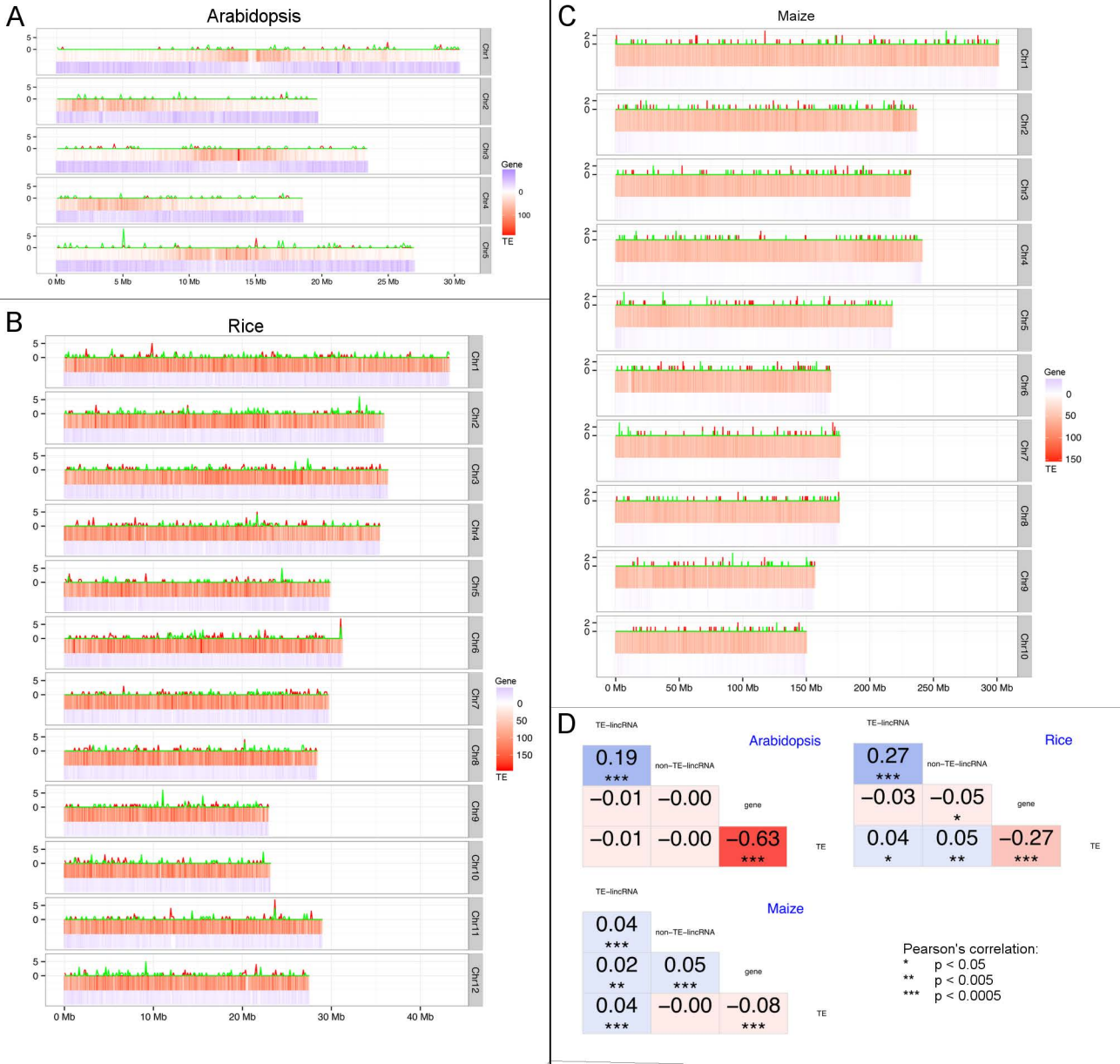
Table S4. Summary of co-expression network analysis in Arabidopsis.

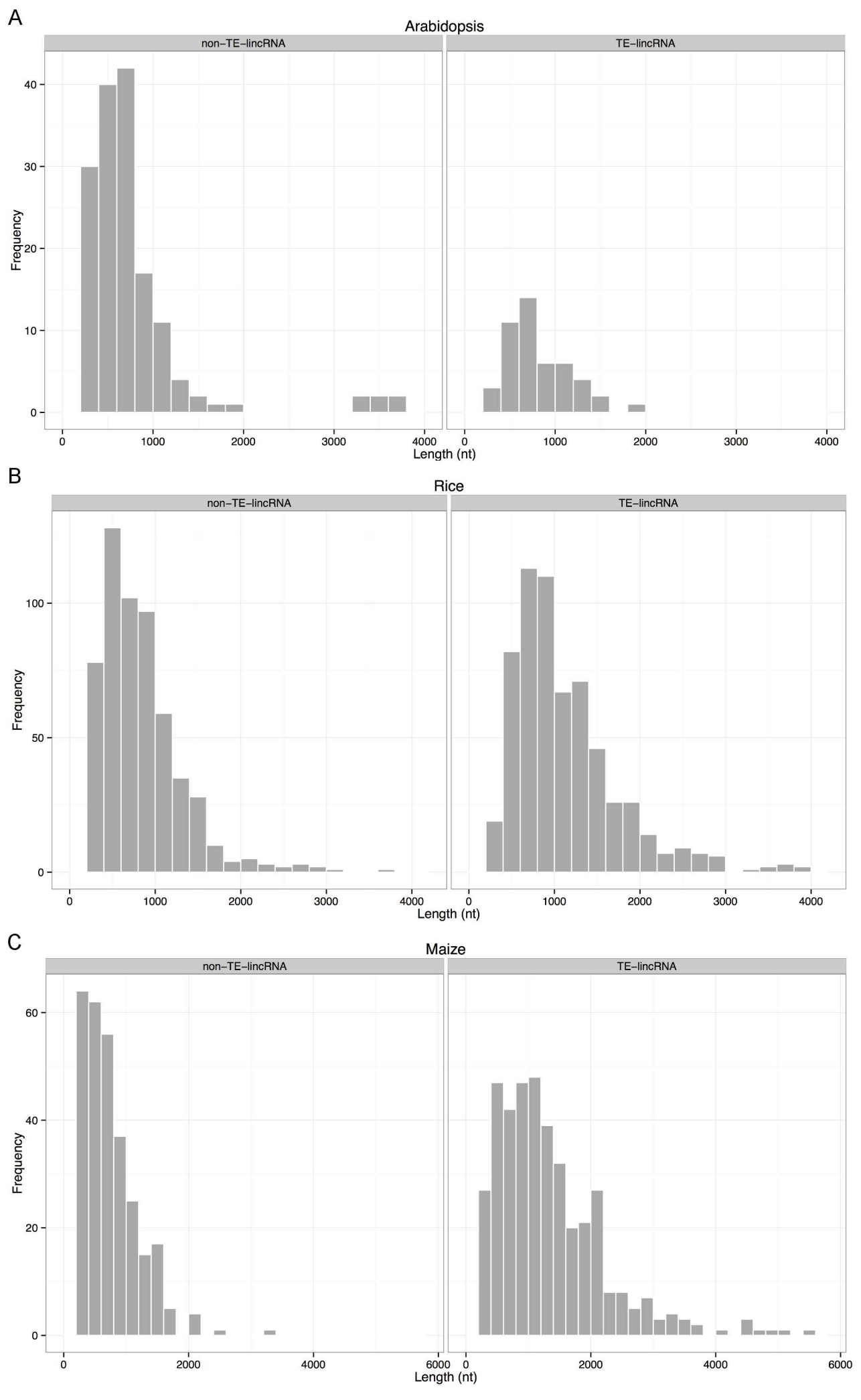
Table S5. Statistics of different TE families contributing to lincRNAs in three species.

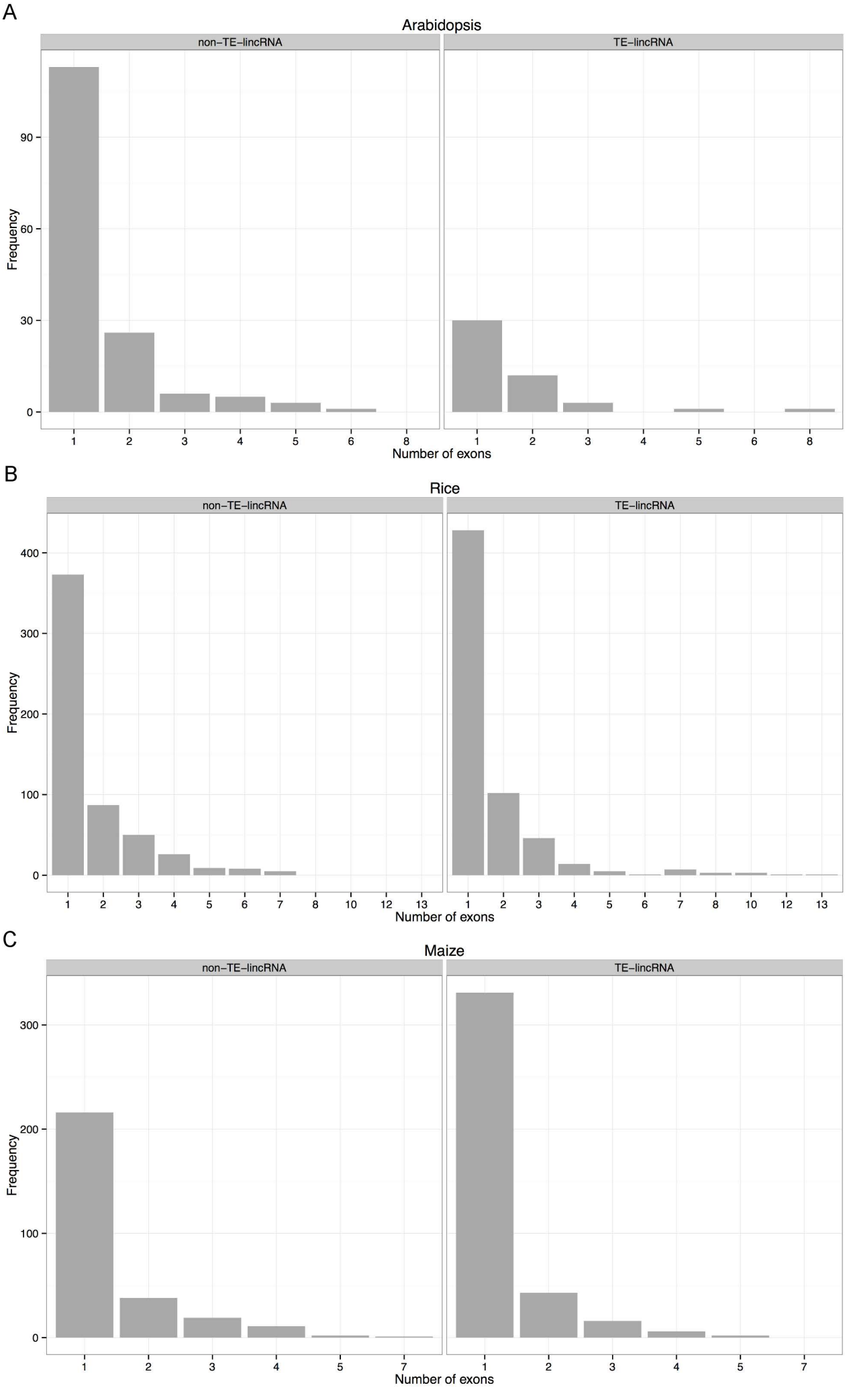
Table S6. Summary of differential gene expression between wild type and mutant *TE-linc11195-2*.

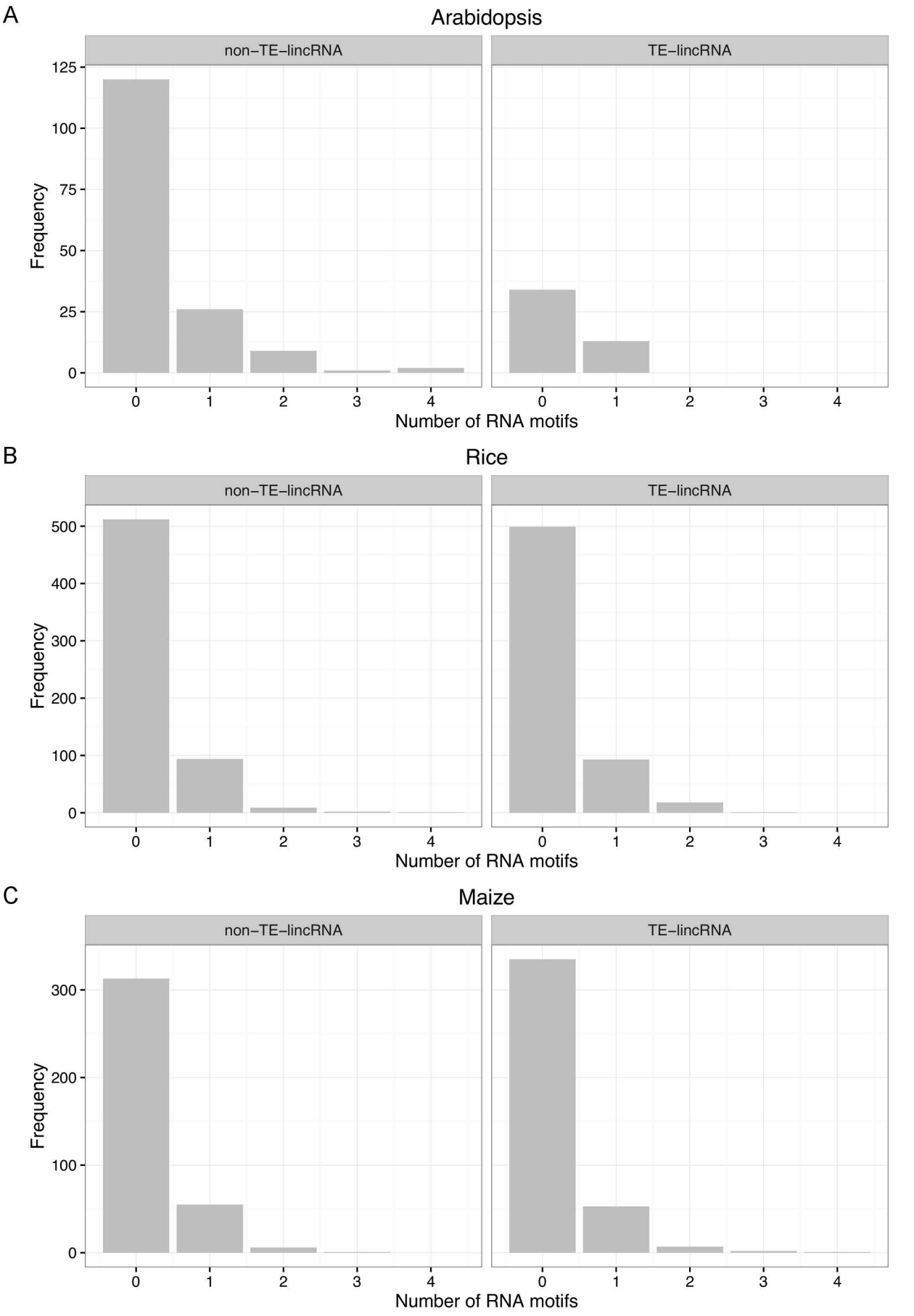
Table S7. Non-Mendelian inheritance of *ddm1* induced lincRNAs.

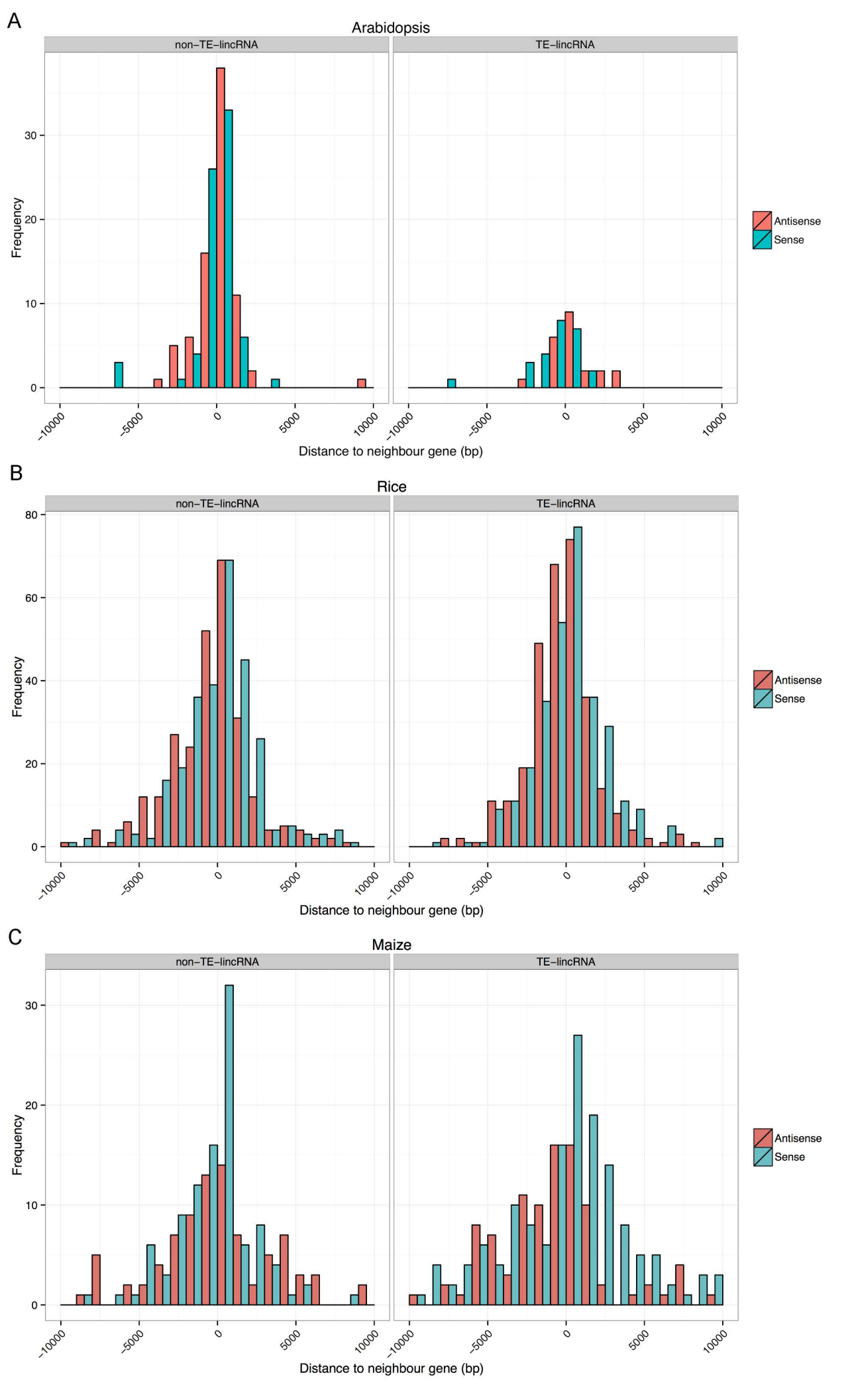
Table S8. Primers used in this study.

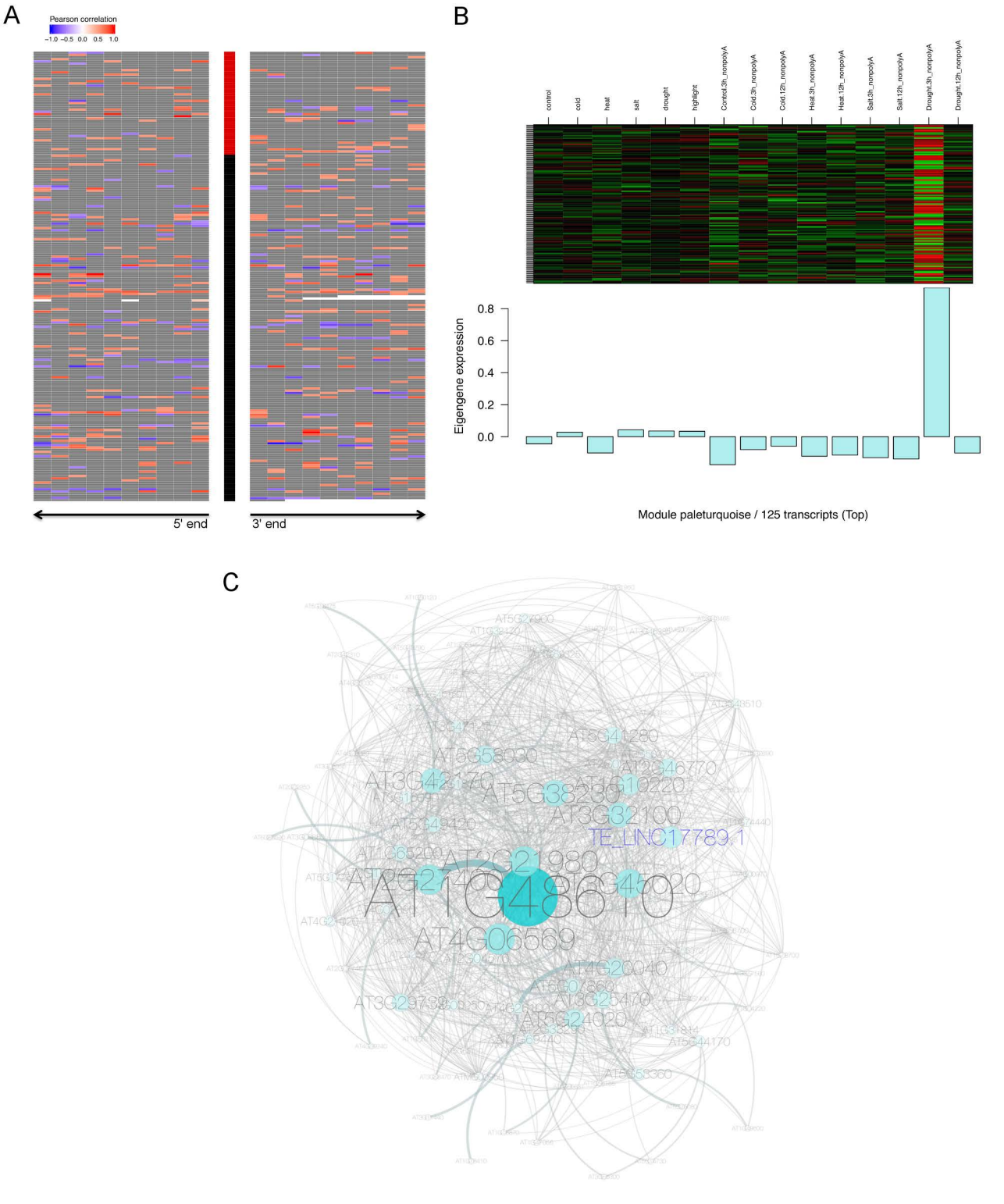

A
Arabidopsis
Chr1
Chr2
Chr3
Chr4
Chr5
0 Mb
5 Mb
10 Mb
15 Mb
20 Mb
25 Mb
30 Mb
Gene
0
100
TE
B
Rice
Chr1
Chr2
Chr3
Chr4
Chr5
Chr6
Chr7
Chr8
Chr9
Chr10
Chr11
Chr12
0 Mb
10 Mb
20 Mb
30 Mb
40 Mb
Gene
0
50
100
150
TE
C
Maize
Chr1
Chr2
Chr3
Chr4
Chr5
Chr6
Chr7
Chr8
Chr9
Chr10
0 Mb
50 Mb
100 Mb
150 Mb
200 Mb
250 Mb
300 Mb
Gene
0
50
100
150
TE
D
TE−lincRNA
Arabidopsis
0.19
***
non−TE−lincRNA
−0.01
−0.00
gene
−0.01
−0.00
−0.63
***
TE
TE−lincRNA
Rice
0.27
***
non−TE−lincRNA
−0.03
−0.05
*
gene
0.04
*
0.05
**
−0.27
***
TE
TE−lincRNA
Maize
0.04
***
non−TE−lincRNA
0.02
**
0.05
***
gene
0.04
***
−0.00
−0.08
***
TE
Pearson's correlation:
* p < 0.05
** p < 0.005
*** p < 0.0005

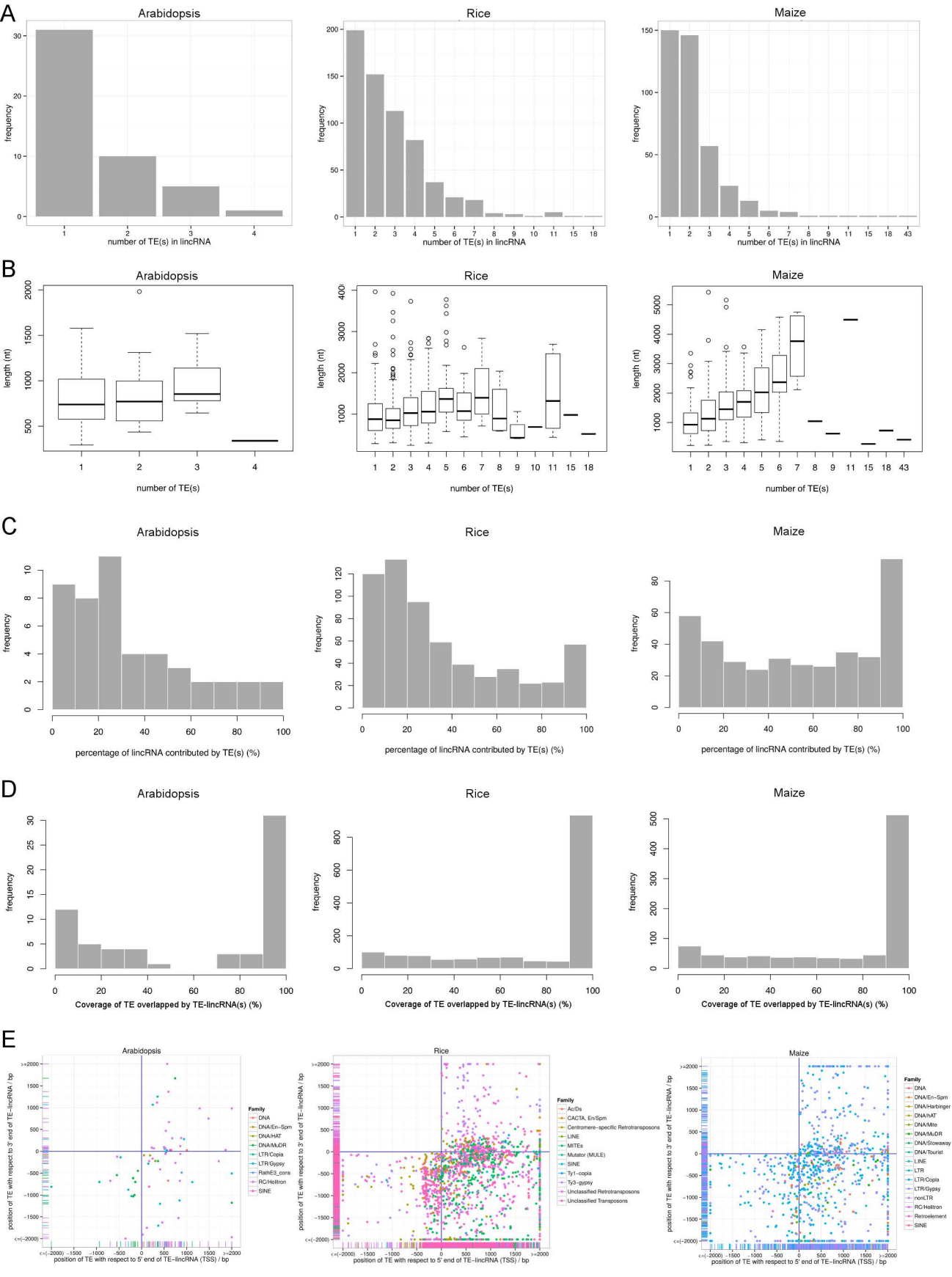

A
Arabidopsis
non−TE−lincRNA
TE−lincRNA
Frequency
Length (nt)
B
Rice
non−TE−lincRNA
TE−lincRNA
Frequency
Length (nt)
C
Maize
non−TE−lincRNA
TE−lincRNA
Frequency
Length (nt)

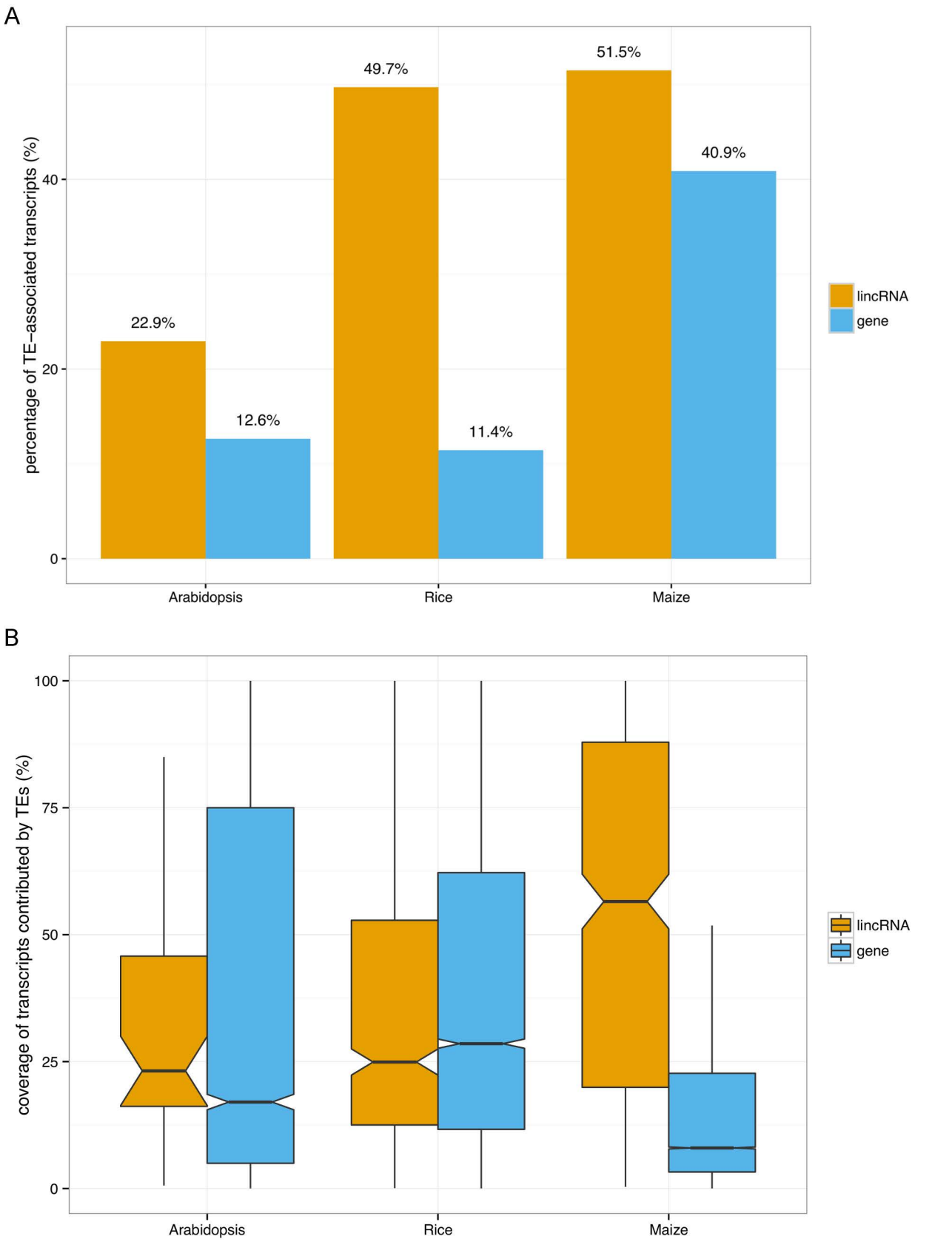

A
Arabidopsis
non−TE−lincRNA
TE−lincRNA
Frequency
0
30
60
90
1
2
3
4
5
6
8
Number of exons
B
Rice
non−TE−lincRNA
TE−lincRNA
Frequency
0
100
200
300
400
1
2
3
4
5
6
7
8
10
12
13
Number of exons
C
Maize
non−TE−lincRNA
TE−lincRNA
Frequency
0
100
200
300
1
2
3
4
5
7
Number of exons


A
Arabidopsis
non−TE−lincRNA
TE−lincRNA
Frequency
125
100
75
50
25
0
0
1
2
3
4
Number of RNA motifs
B
Rice
non−TE−lincRNA
TE−lincRNA
Frequency
500
400
300
200
100
0
0
1
2
3
4
Number of RNA motifs
C
Maize
non−TE−lincRNA
TE−lincRNA
Frequency
300
200
100
0
0
1
2
3
4
Number of RNA motifs

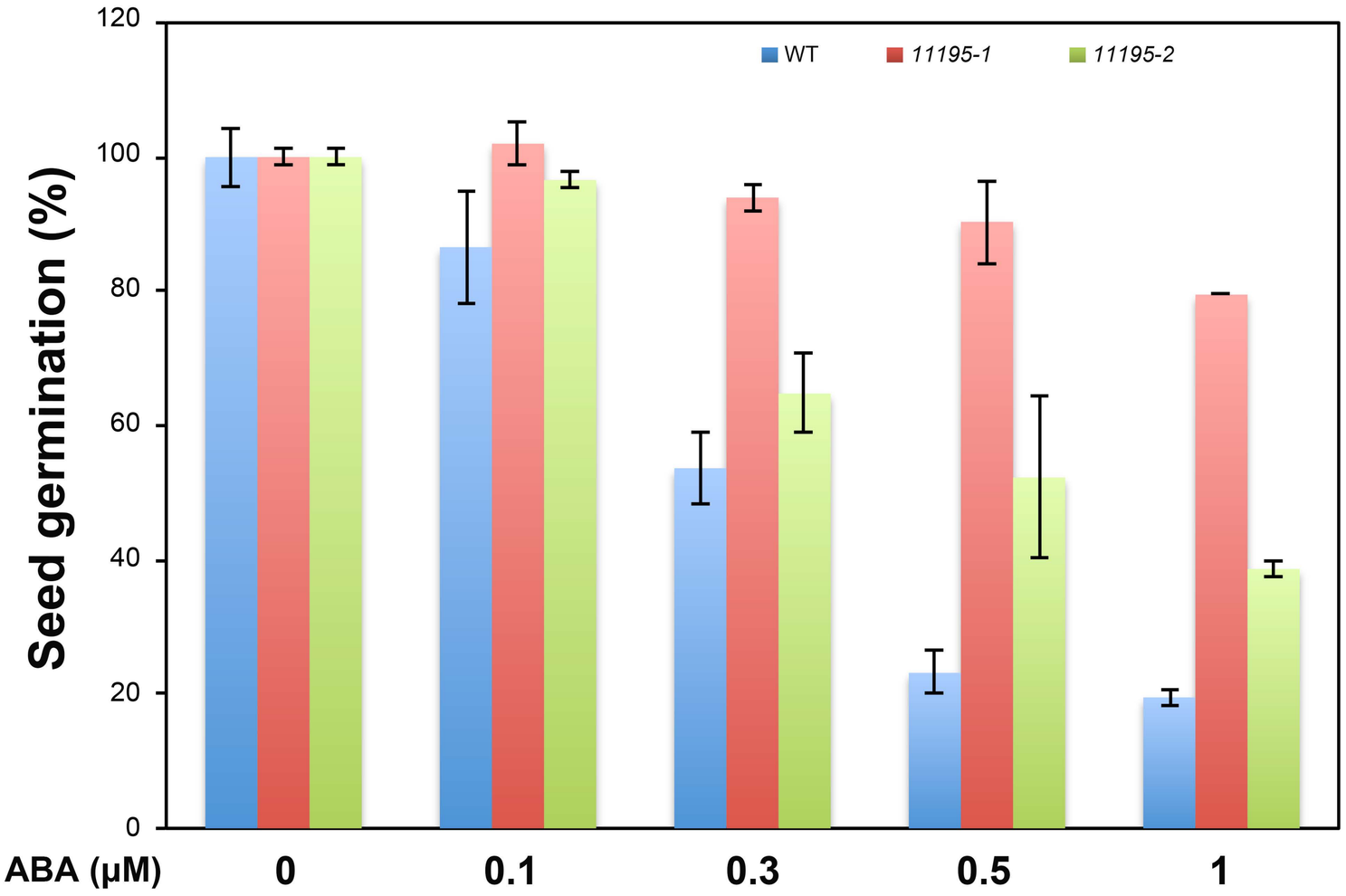

A
Arabidopsis
non−TE−lincRNA
TE−lincRNA
Frequency
Distance to neighbour gene (bp)
Antisense
Sense
B
Rice
non−TE−lincRNA
TE−lincRNA
Frequency
Distance to neighbour gene (bp)
Antisense
Sense
C
Maize
non−TE−lincRNA
TE−lincRNA
Frequency
Distance to neighbour gene (bp)
Antisense
Sense


A
Pearson correlation
−1.0 −0.5 0.0 0.5 1.0
5' end
3' end
B
control
cold
heat
salt
drought
highlight
Control.3h_nonpolyA
Cold.3h_nonpolyA
Cold.12h_nonpolyA
Heat.3h_nonpolyA
Heat.12h_nonpolyA
Salt.3h_nonpolyA
Salt.12h_nonpolyA
Drought.3h_nonpolyA
Drought.12h_nonpolyA
Eigengene expression
0.8
0.6
0.4
0.2
0.0
Module paleturquoise / 125 transcripts (Top)
C
AT1G48610
AT4G06569
AT4G20040
AT3G25470
AT5G24020
AT5G44170
AT5G53360
AT3G42170
AT5G38230
AT3G32100
TE_LINC17789.1
AT2G21980
AT5G41280
AT2G46770

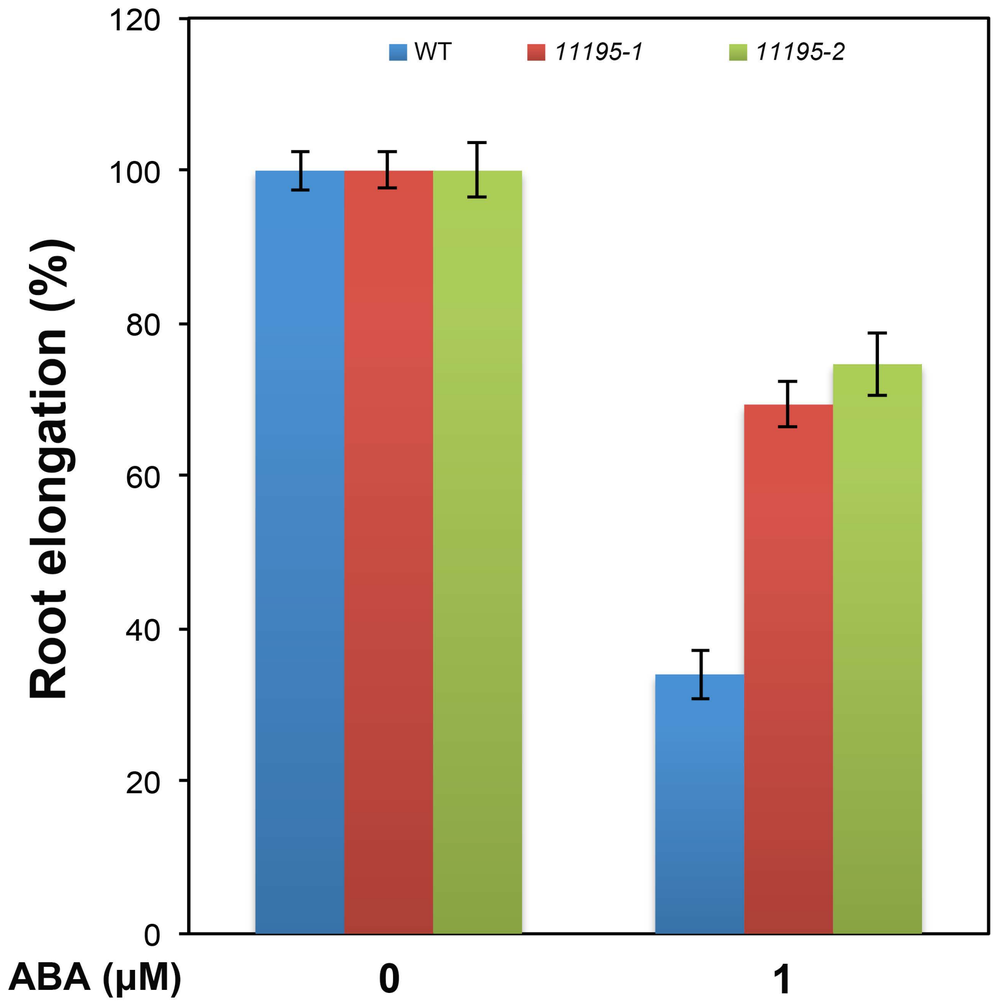

A
Arabidopsis
Rice
Maize
frequency
number of TE(s) in lincRNA
B
Arabidopsis
Rice
Maize
length (nt)
number of TE(s)
C
Arabidopsis
Rice
Maize
frequency
percentage of lincRNA contributed by TE(s) (%)
D
Arabidopsis
Rice
Maize
frequency
Coverage of TE overlapped by TE-lincRNA(s) (%)
E
Arabidopsis
Rice
Maize
position of TE with respect to 3' end of TE-lincRNA / bp
position of TE with respect to 5' end of TE-lincRNA (TSS) / bp
Family
DNA
DNA/En-Spm
DNA/HAT
DNA/MuDR
LTR/Copia
LTR/Gypsy
RathE3_cons
RC/Helitron
SINE
Ac/Ds
CACTA, En/Spm
Centromere-specific Retrotransposons
LINE
MITEs
Mutator (MULE)
Ty1-copia
Ty3-gypsy
Unclassified Retrotransposons
Unclassified Transposons
DNA/Harbinger
DNA/hAT
DNA/Mite
DNA/Slowaway
DNA/Tourist
LTR
nonLTR
Retroelement


A
percentage of TE−associated transcripts (%)
0
20
40
22.9%
12.6%
49.7%
11.4%
51.5%
40.9%
Arabidopsis
Rice
Maize
lincRNA
gene
B
coverage of transcripts contributed by TEs (%)
0
25
50
75
100
Arabidopsis
Rice
Maize
lincRNA
gene

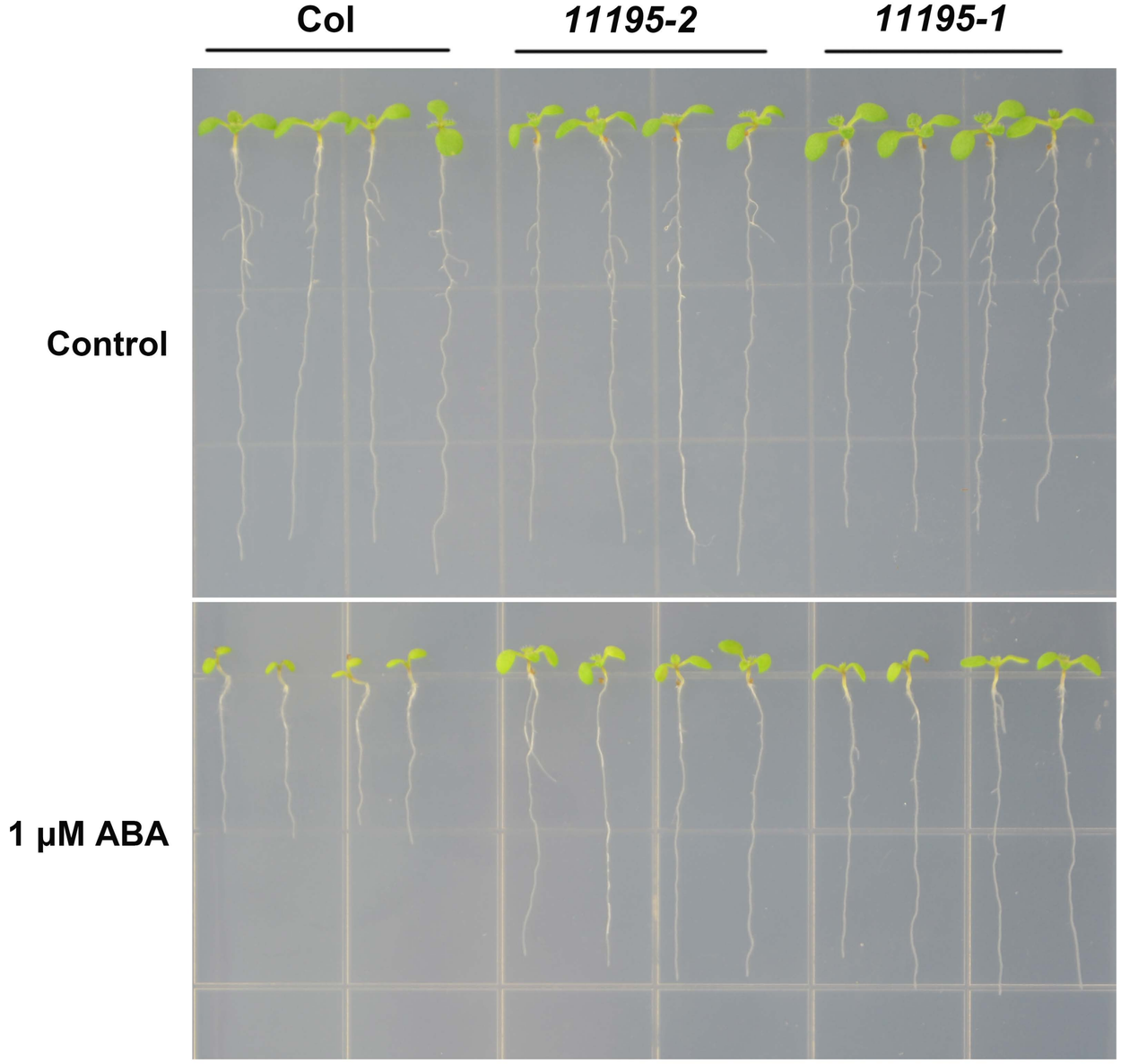
A

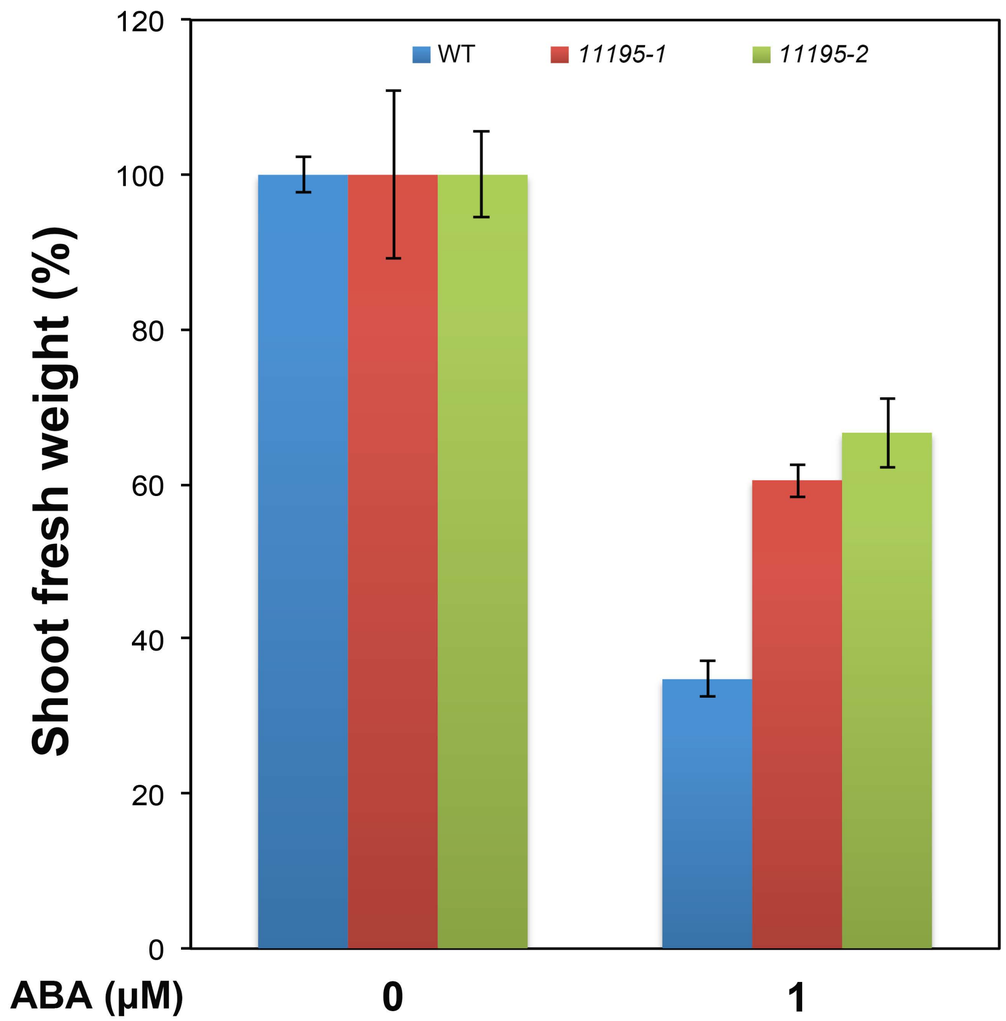

WT
11195-1
11195-2
Seed germination (%)
120
100
80
60
40
20
0
ABA (μM)
0
0.1
0.3
0.5
1

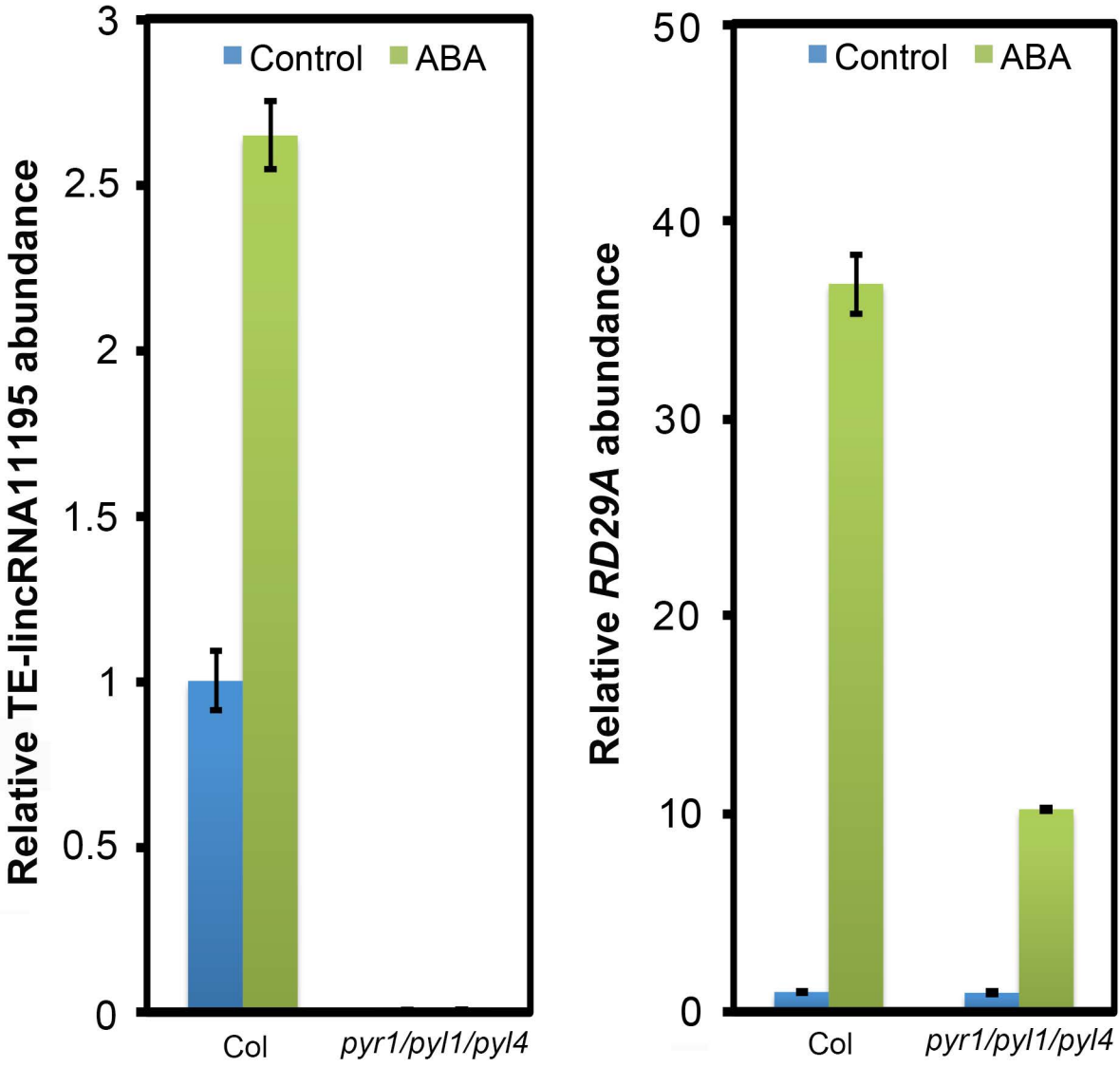
C

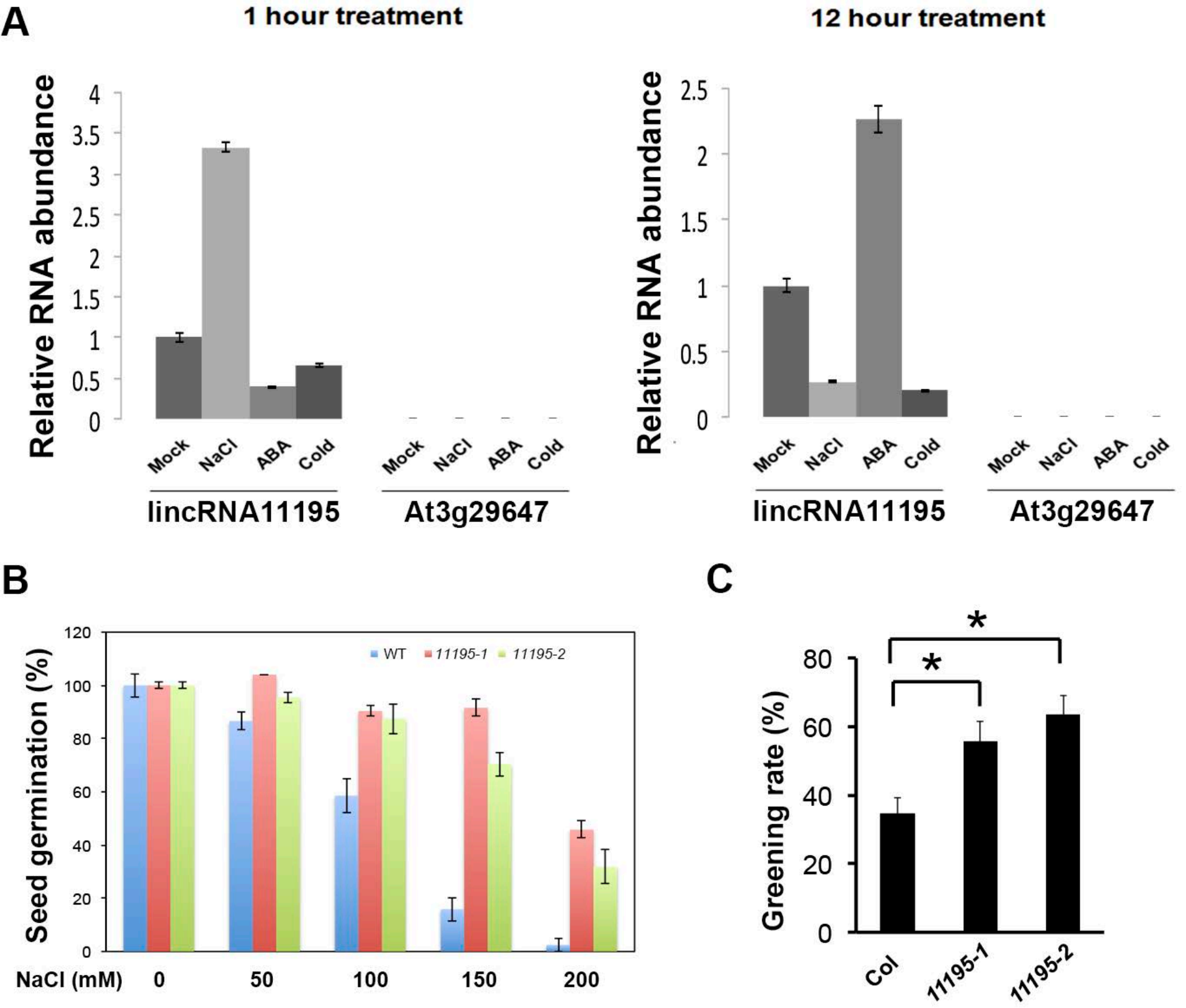

WT
11195-1
11195-2
Root elongation (%)
120
100
80
60
40
20
0
ABA (μM)
0
1

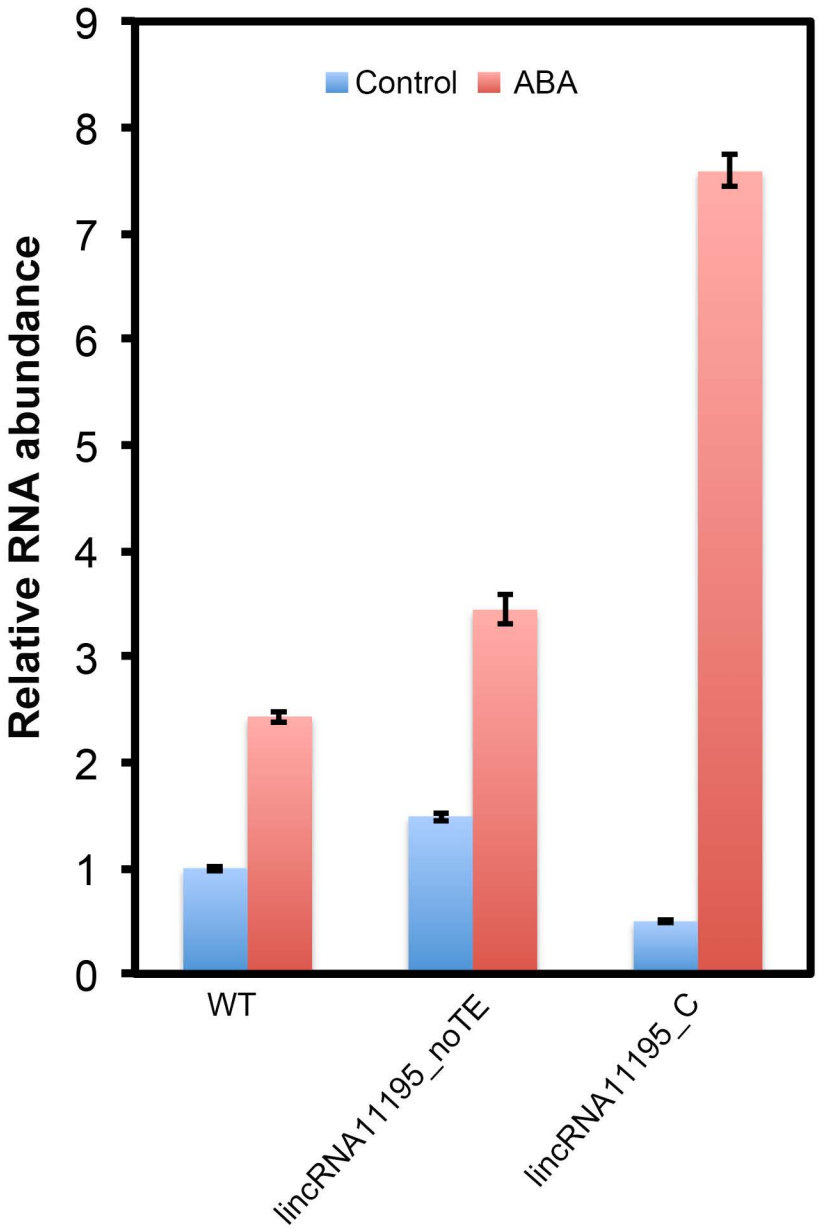
B

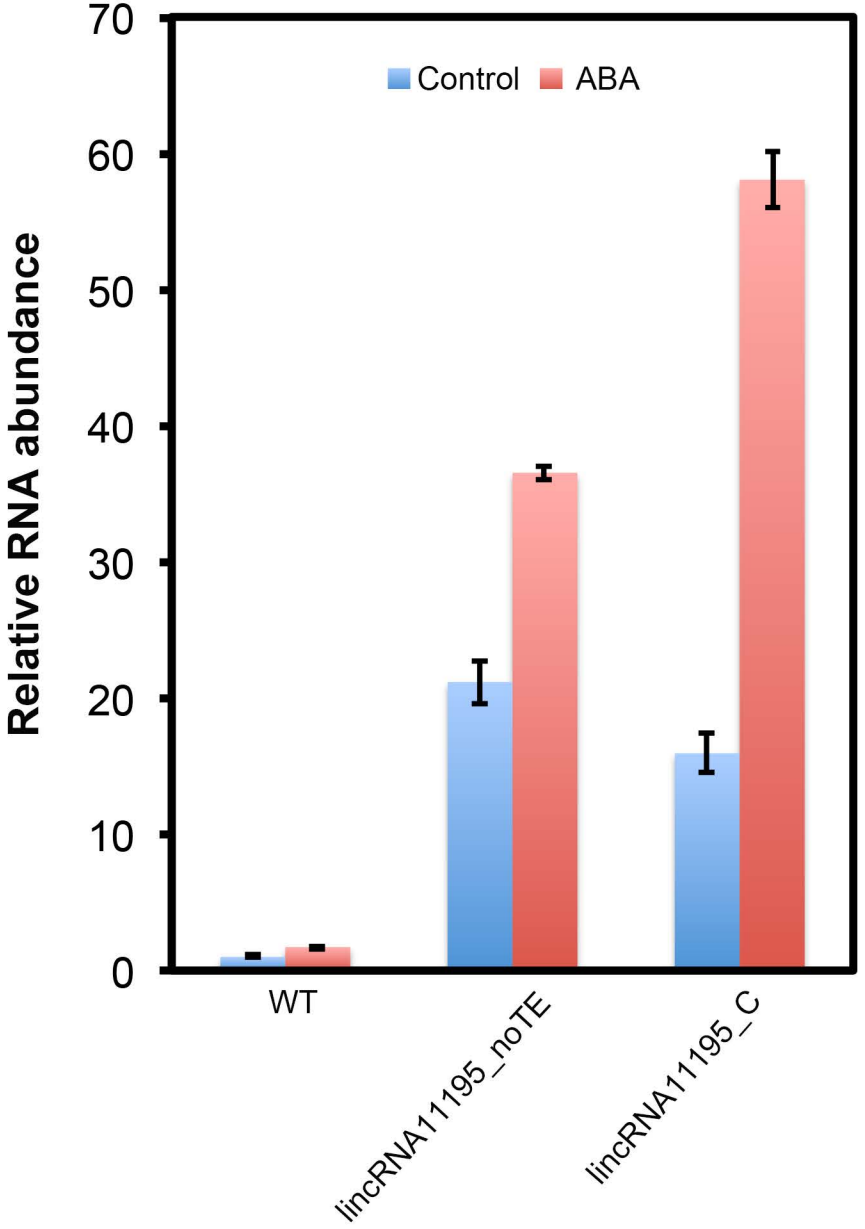

Col
11195-2
11195-1
Control
1 μM ABA

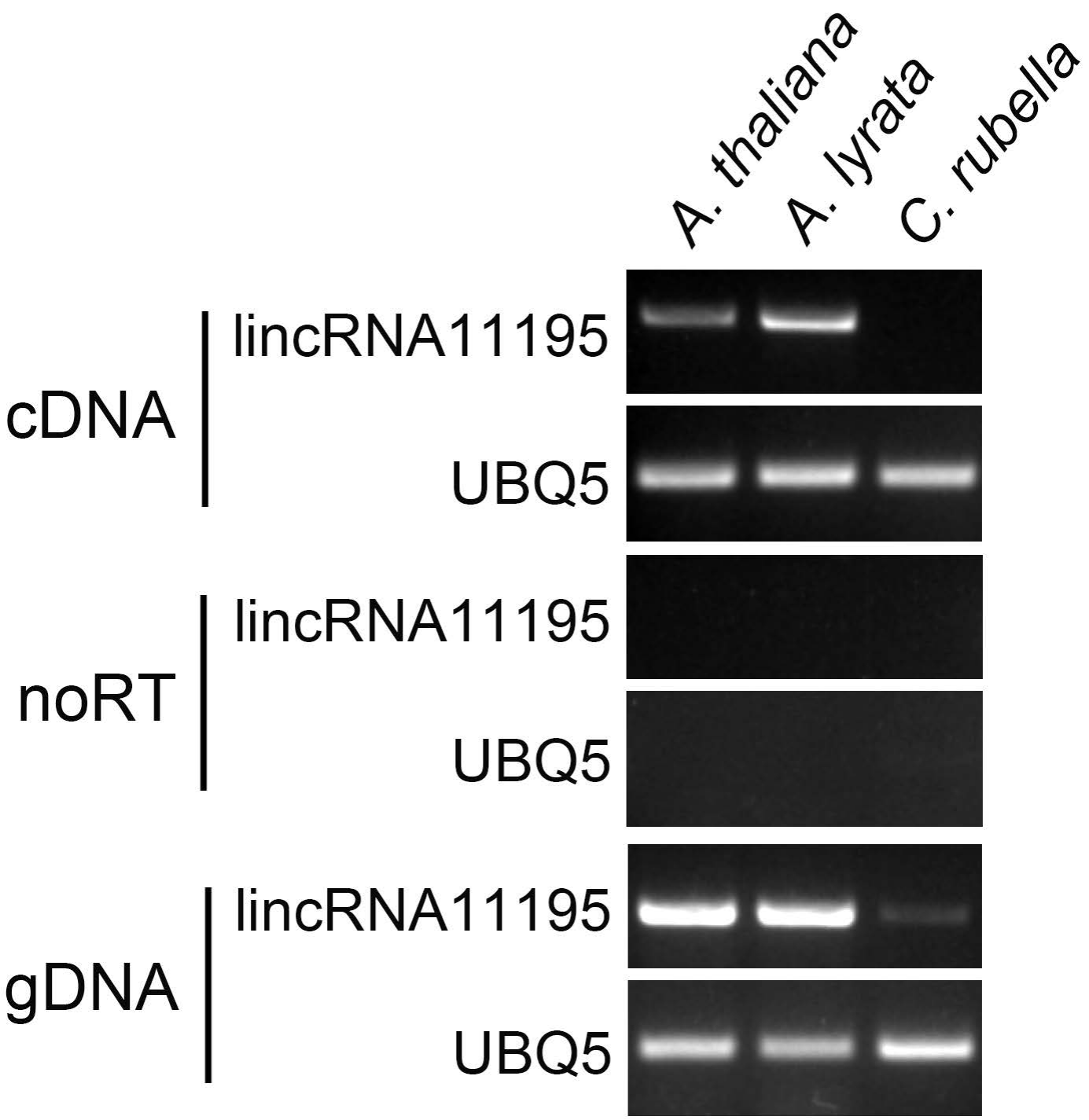

WT
11195-1
11195-2
Shoot fresh weight (%)
120
100
80
60
40
20
0
ABA (μM)
0
1

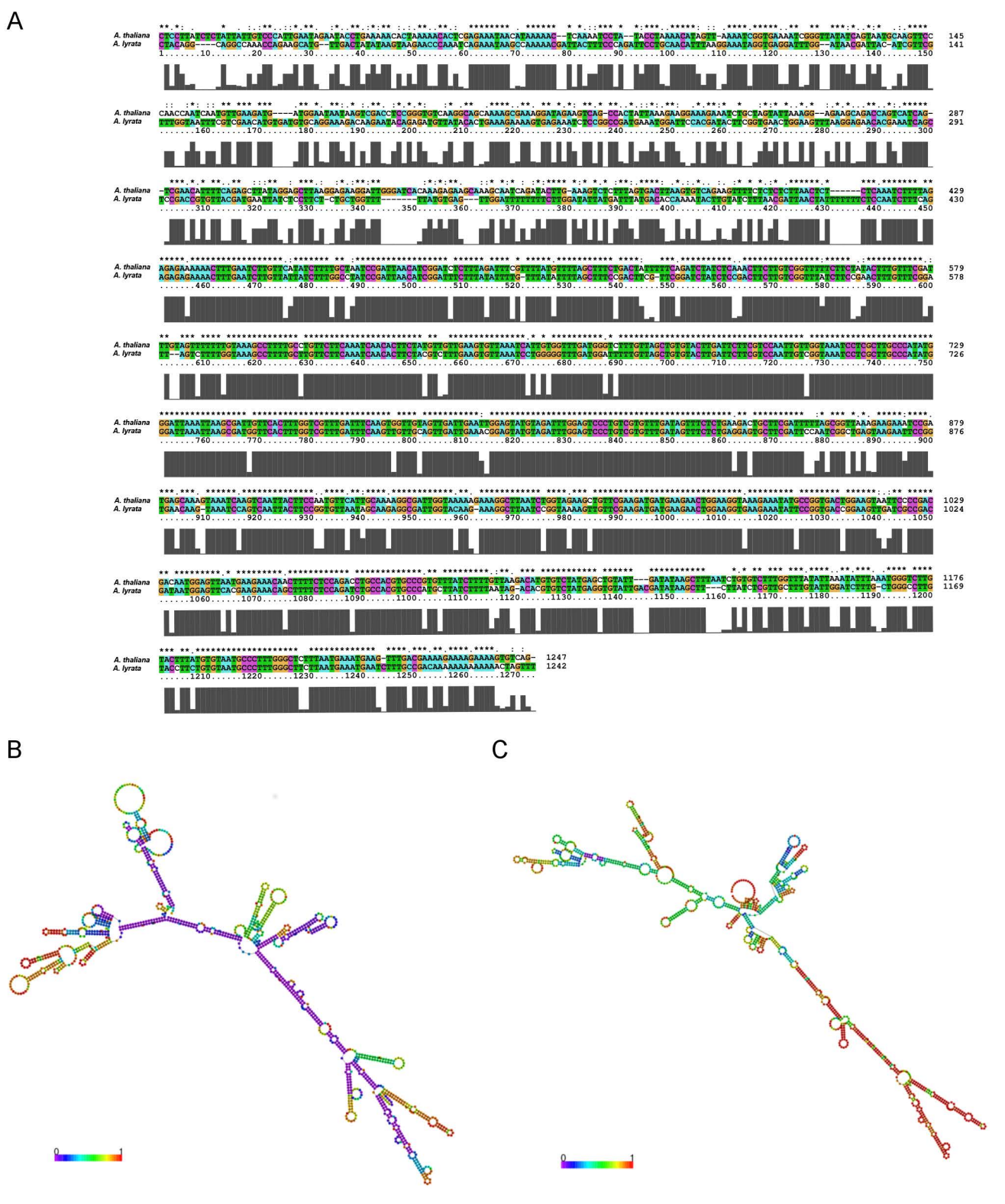

Relative TE-lincRNA11195 abundance
3
2.5
2
1.5
1
0.5
0
Control
ABA
Col
pyr1/pyl1/pyl4
Relative RD29A abundance
50
40
30
20
10
0
Control
ABA
Col
pyr1/pyl1/pyl4


A
1 hour treatment
12 hour treatment
Relative RNA abundance
Mock
NaCl
ABA
Cold
lincRNA11195
At3g29647
B
Seed germination (%)
WT
11195-1
11195-2
NaCl (mM)
0
50
100
150
200
C
Greening rate (%)
Col
11195-1
11195-2
*


Control
ABA
Relative RNA abundance
0
1
2
3
4
5
6
7
8
9
WT
lincRNA11195_noTE
lincRNA11195_C


Control
ABA
Relative RNA abundance
0
10
20
30
40
50
60
70
WT
lincRNA11195_noTE
lincRNA11195_C


A. thaliana
A. lyrata
C. rubella
cDNA
lincRNA11195
UBQ5
noRT
lincRNA11195
UBQ5
gDNA
lincRNA11195
UBQ5


A
A. thaliana
A. lyrata
B
C
0
1