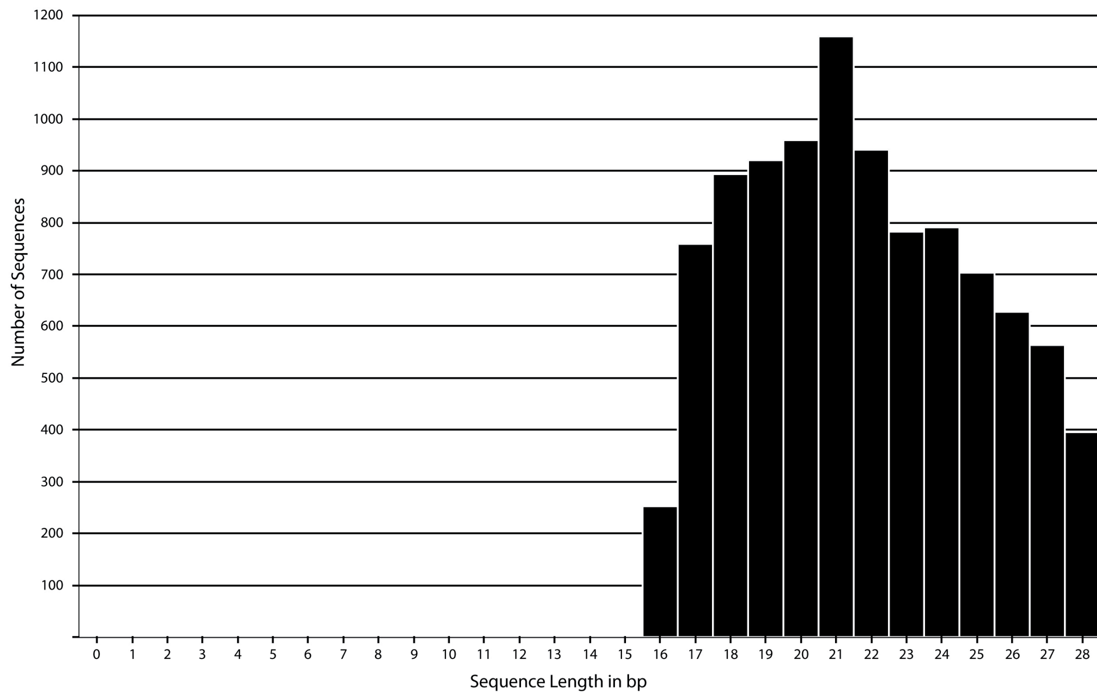
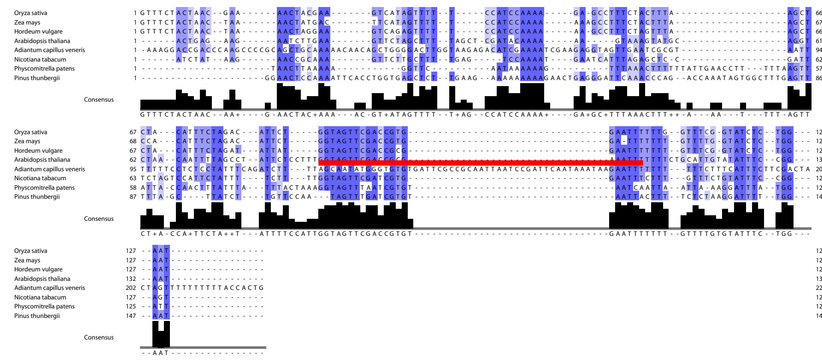
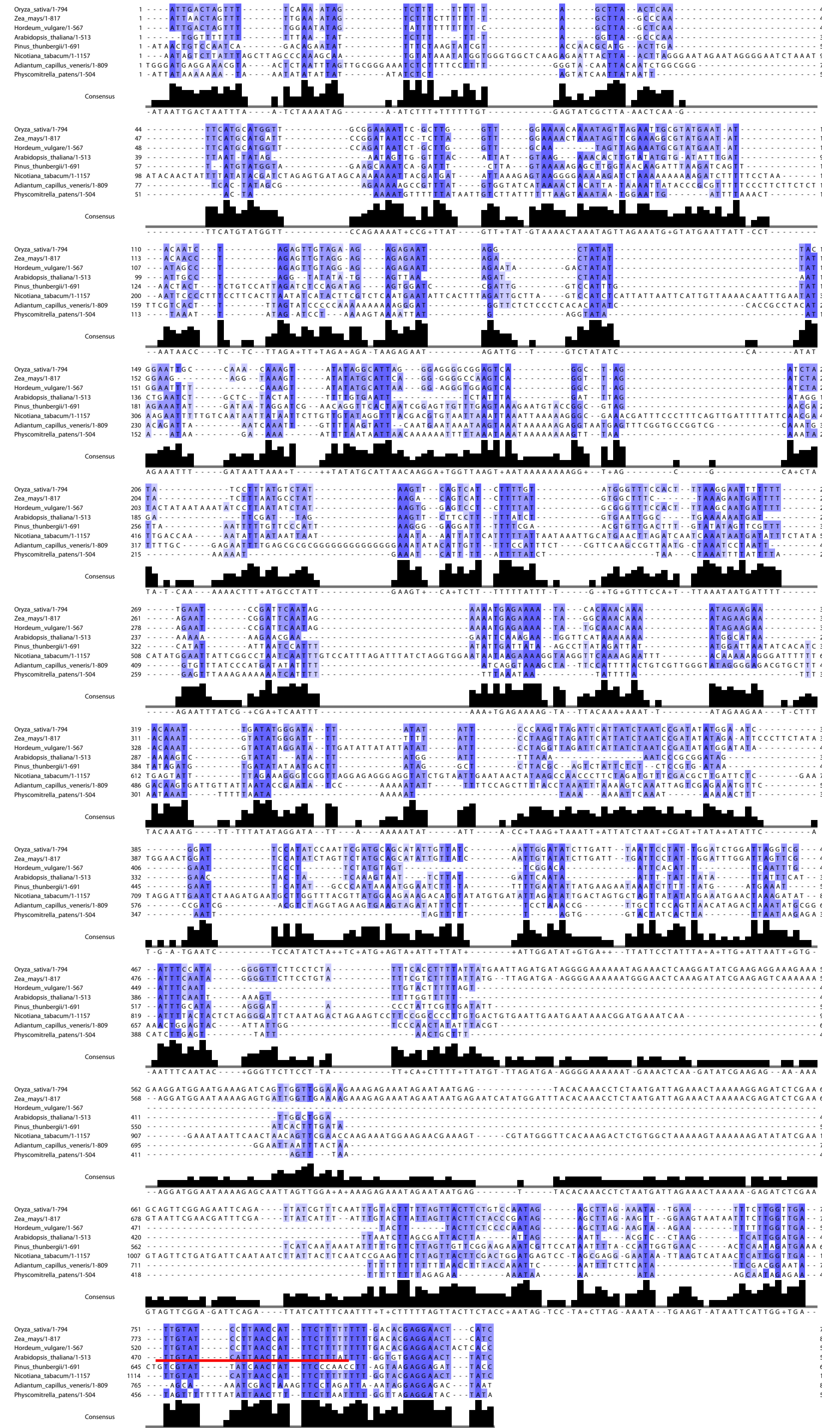


Supplemental Figure 1: Read distribution of chloroplast RNA seq reads.
In line with the experimental strategy applied by Ragagopalan et al. (2006), which aimed at identifying miRNAs, only reads between 16 and 28 nucleotides were obtained and further analysed here.

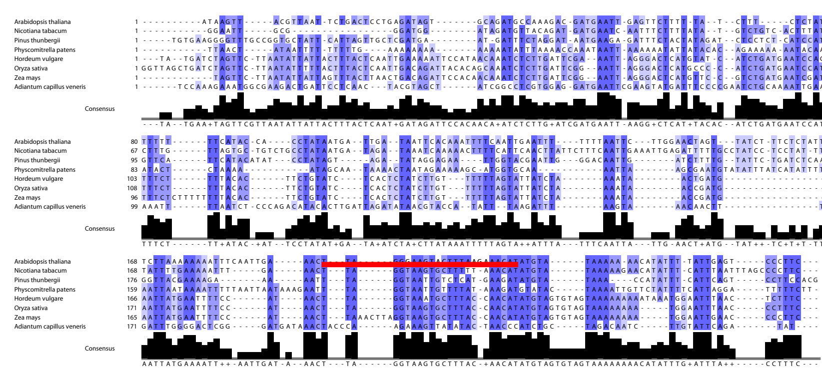
A) *psbH-petB* intergenic spacer



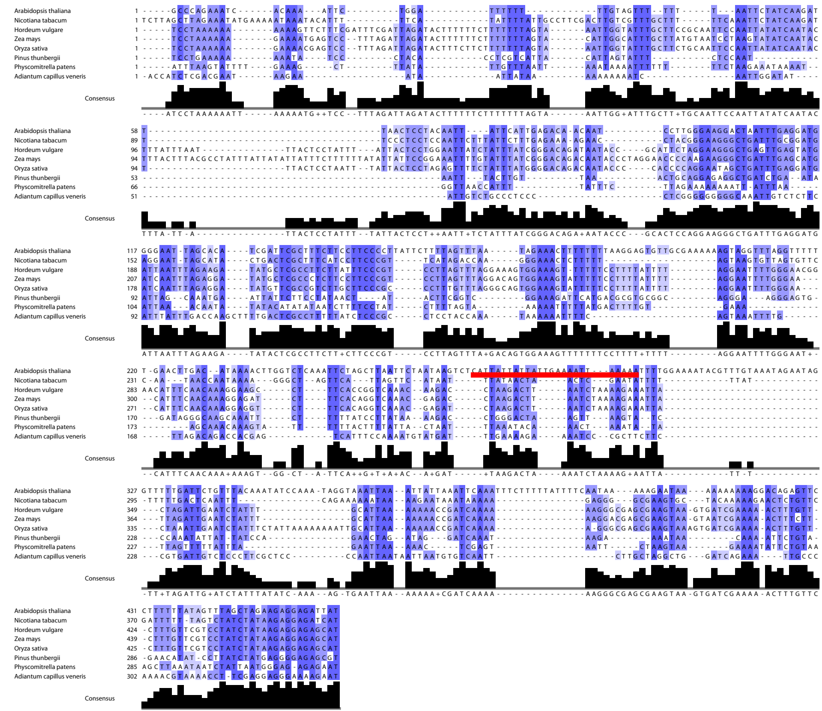
B) *atpI-atpH* intergenic spacer



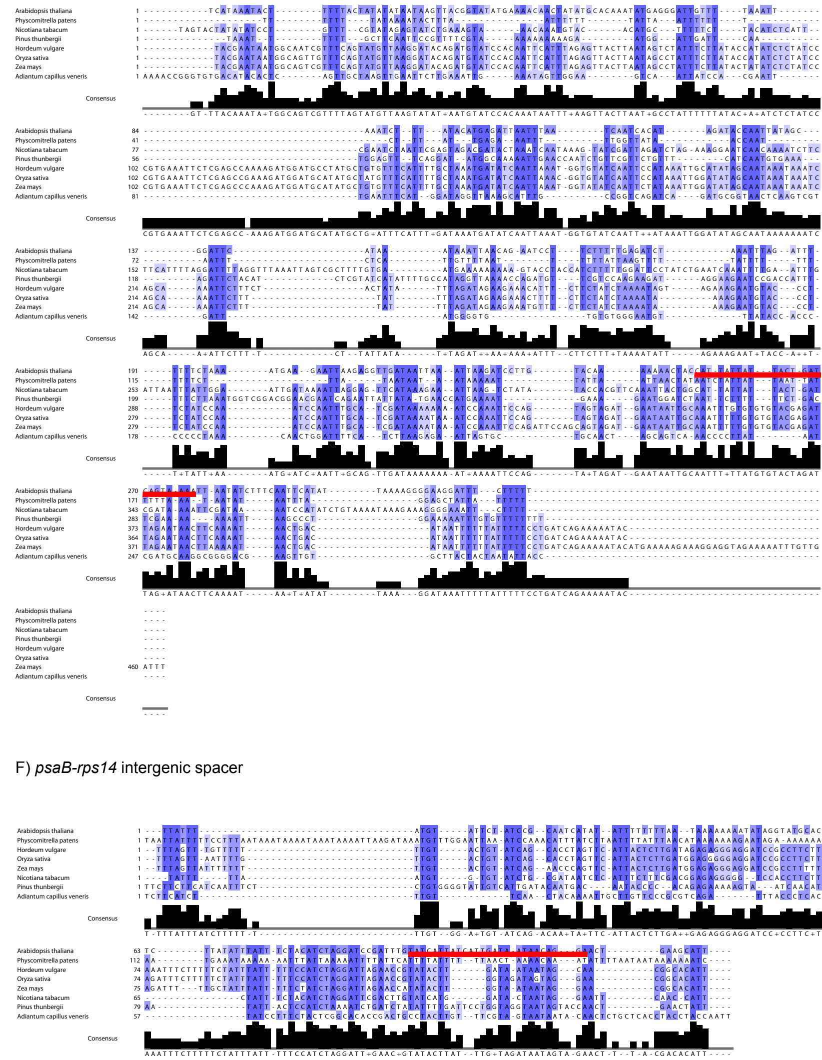
C) *petL* 5'-area



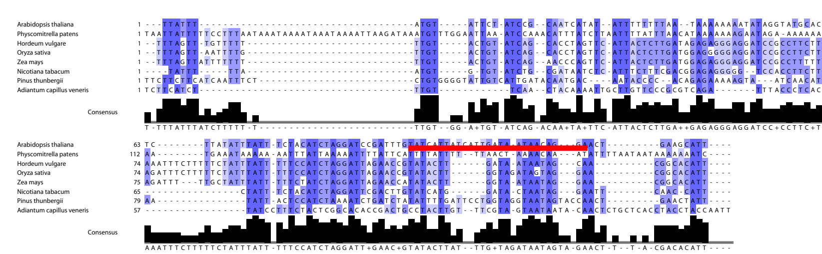
D) *atpH-atpF* intergenic spacer



E) *ycf1-rps15* intergenic spacer

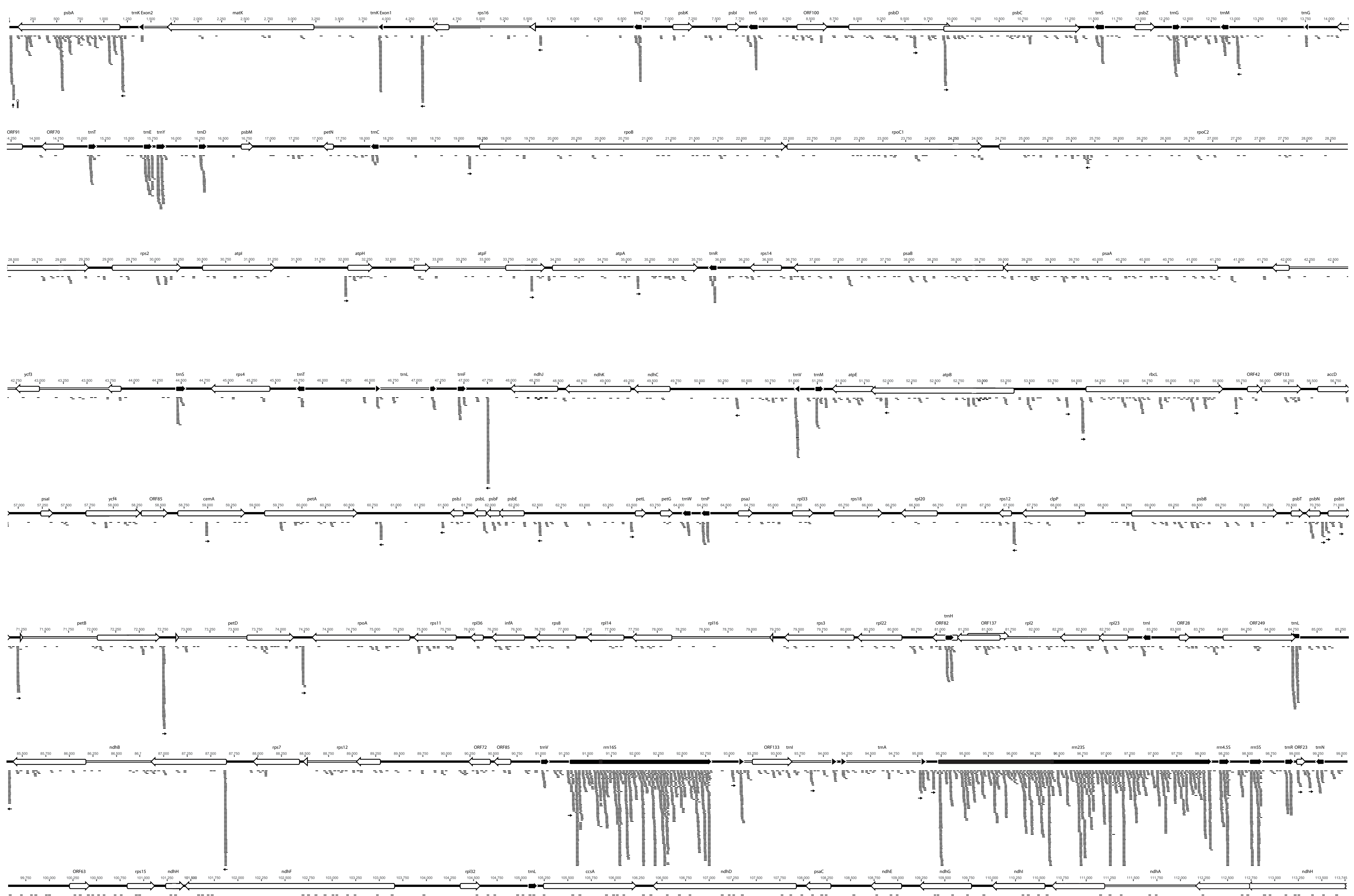


F) *psaB-rps14* intergenic spacer

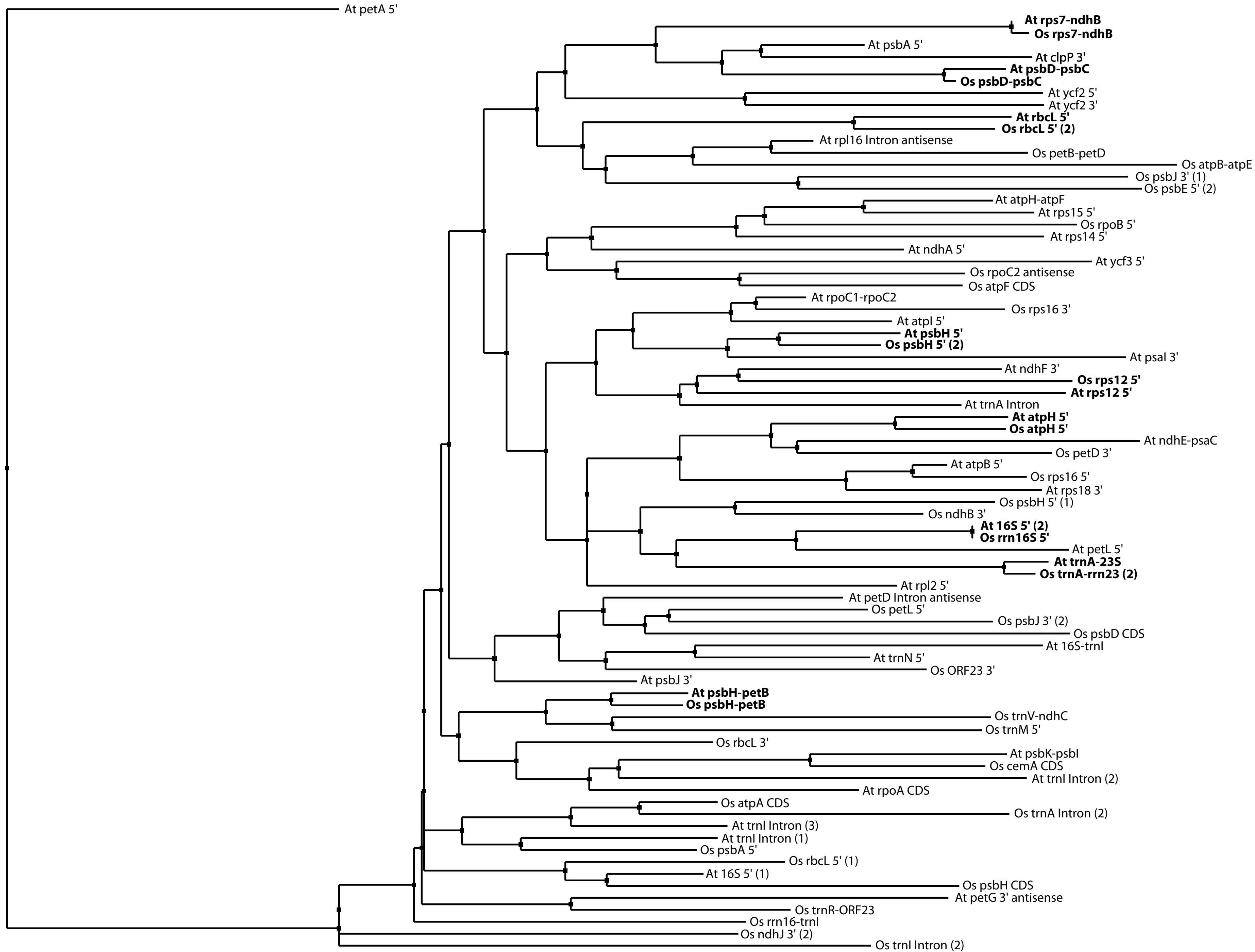


Supplemental Figure 2: Alignment of randomly selected intergenic regions containing sRNAs.

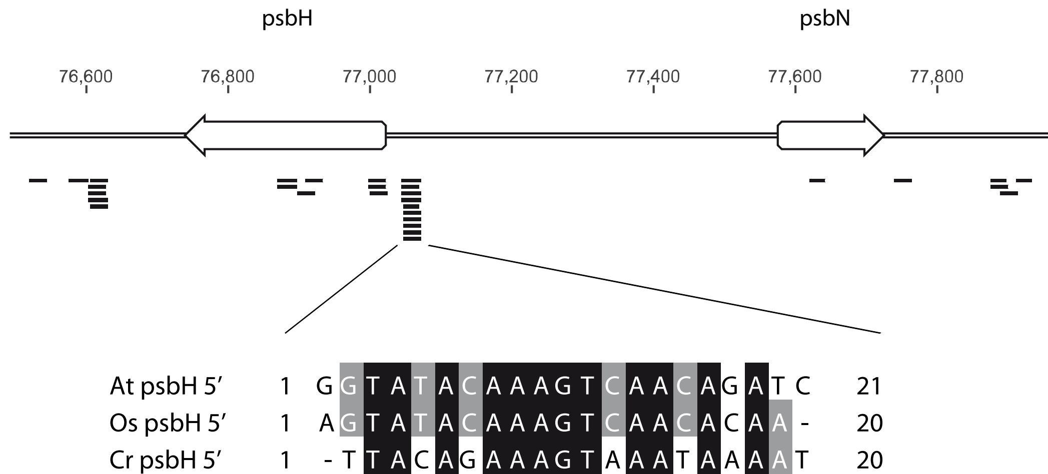
The alignment was prepared using ClustalW2 (10) and the Jalview software package for graphical output (11). The Arabidopsis sRNAs in this intergenic spacer are indicated by solid red bars. A graphical and a sequence consensus are given at the bottom of the alignments. Shading refers to different levels of sequence conservation: dark blue = 7 or 8 of 8 sequences share the same base; intermediate blue = 5 or 6 of 8 share the same base; light blue = 4 of 8 share the same base.



Supplemental Figure 3: Map of RNA seq data onto the chloroplast genome of rice.
Individual RNA seq reads were extracted from the small RNA dataset found in the Cereals Small RNA Database for rice (9) and aligned with the chloroplast chromosome of rice (acc. no. NC_001320) using the Geneious software package. Only sequences matching 100% with the genome were considered. The bars below the genome map indicate groups of identical reads from a single library. Thus, several bars can represent identical sequence stretches but are from different RNA seq libraries. In several cases, e.g. for the rRNA operon, the number of reads piled up at certain genomic positions exceeds the space between two rows of the genome map and are thus cut off. Other symbols and figure details as in figure 1.

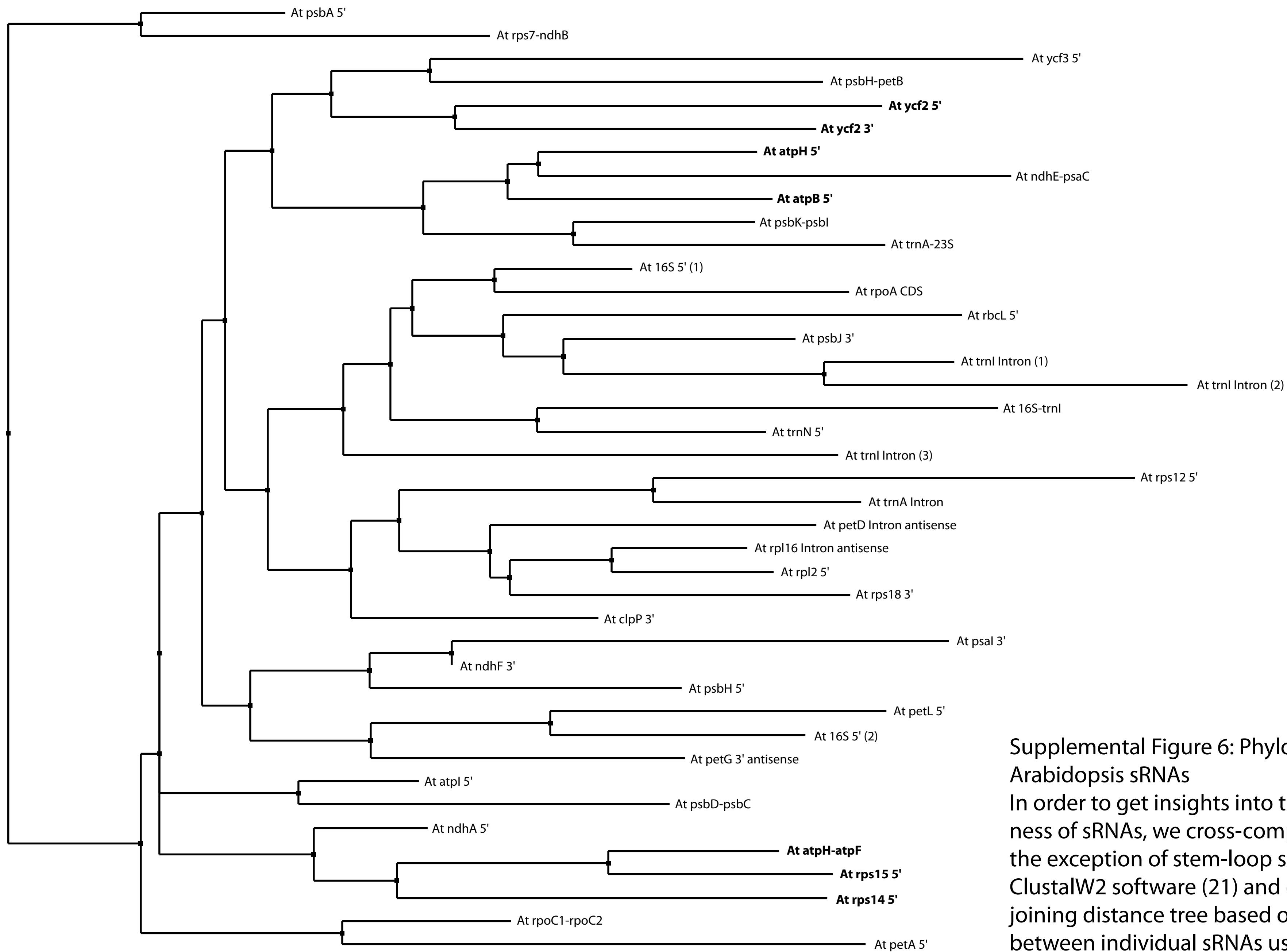


Supplemental Figure 4: Phylogenetic analysis of rice and Arabidopsis sRNAs
In order to get insights into the interspecific relatedness of sRNAs beyond a mere co-occurrence in the same intergenic region, we cross-compared all sRNAs from rice and Arabidopsis (with the exception of stem-loop sRNAs) using the ClustalW2 software (10) and calculated a neighbor joining distance tree based on percent identity between individual sRNAs using Jalview Software (11). The tree was rooted against the sRNA with the least similarity to any other sRNA in the set. In this tree, we marked those pairs of sRNAs that share the same chromosomal region (in bold) as being likely orthologous.

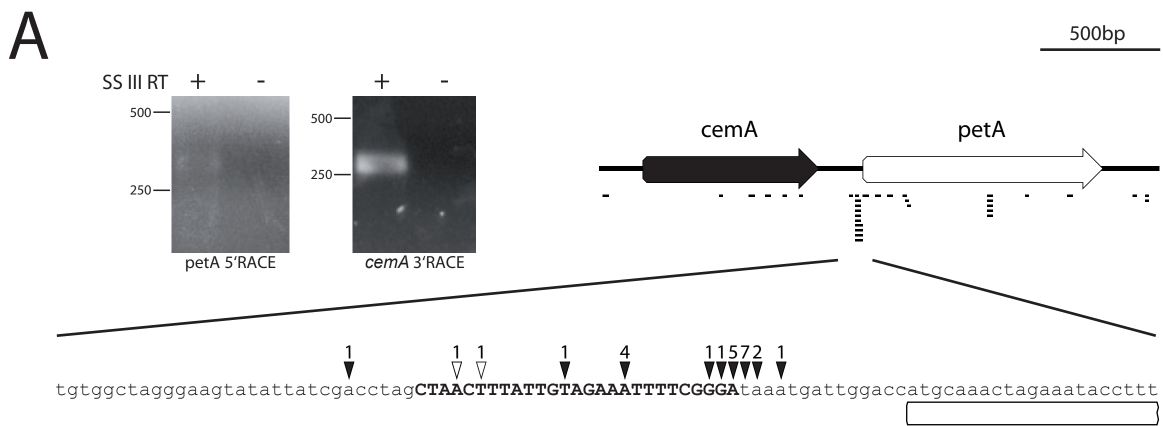


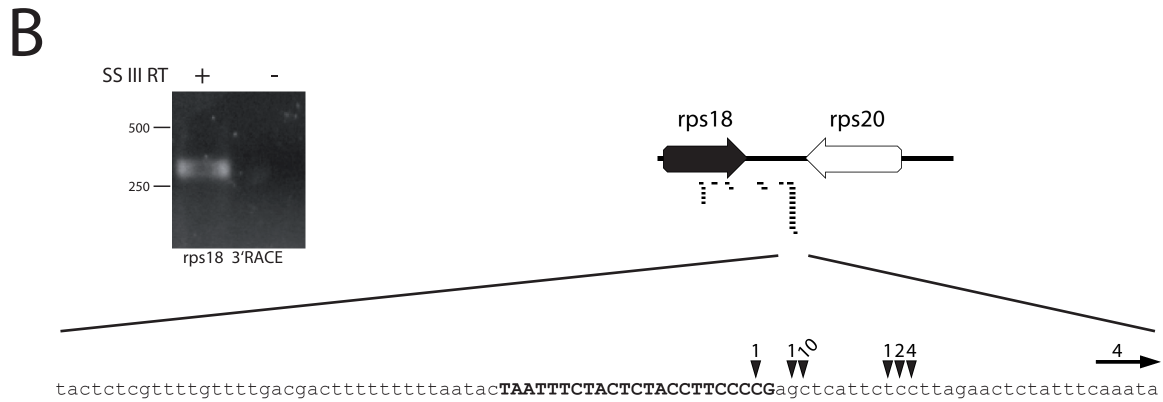
Supplemental Figure 5: A SCAR conserved between Arabidopsis and Chlamydomonas

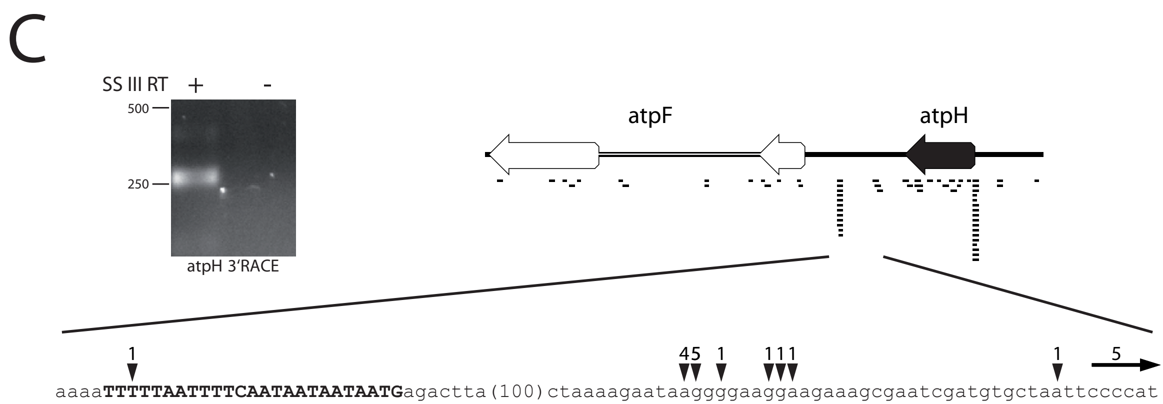
SCARs identified in Chlamydomonas deep sequencing data were compared with SCARs in Arabidopsis. If the genomic position relative to adjacent genes was conserved for a SCAR between the two species, sequence alignments were prepared as well. This identified only one SCAR with a sequence conserved between Arabidopsis, Chlamydomonas, and also in rice. This SCAR upstream of psbH is likely a footprint of the R-TPR protein HCF107 / Mbb1.

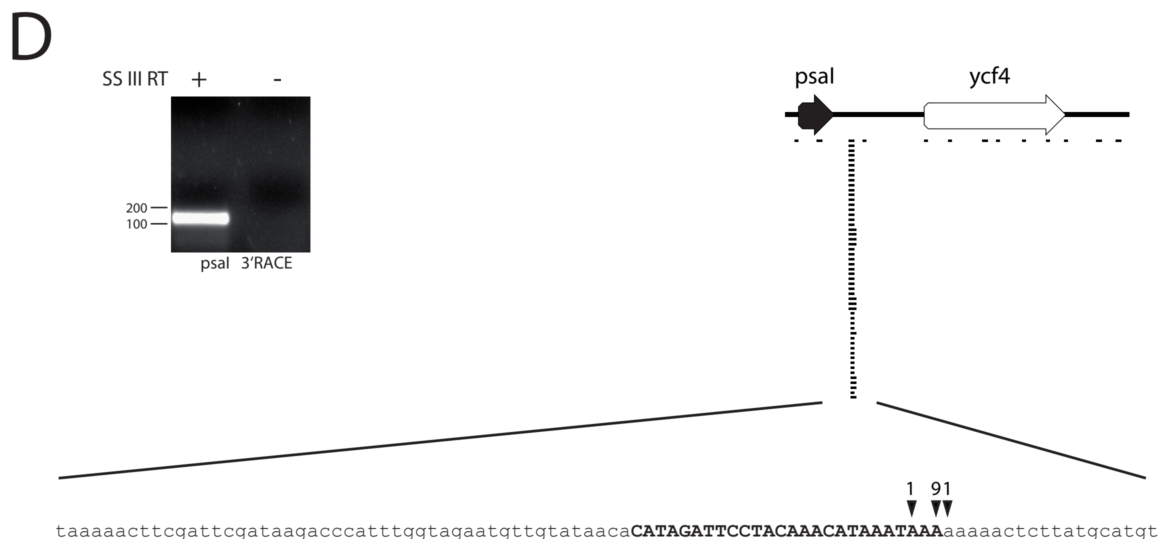


Supplemental Figure 6: Phylogenetic analysis of Arabidopsis sRNAs
In order to get insights into the intraspecific relatedness of sRNAs, we cross-compared all sRNAs (with the exception of stem-loop sRNAs) using the ClustalW2 software (21) and calculated a neighbor joining distance tree based on percent identity between individual sRNAs using Jalview Software (22). The tree was rooted against the sRNA with the least similarity to any other sRNA in the set.









Supplemental Figure 7: Mapping of transcript ends in the vicinity of selected SCARs
Transcript end mapping in four different intergenic spacers by RACE analysis. All symbols and numbers are explained in figure 1 and figure 5.