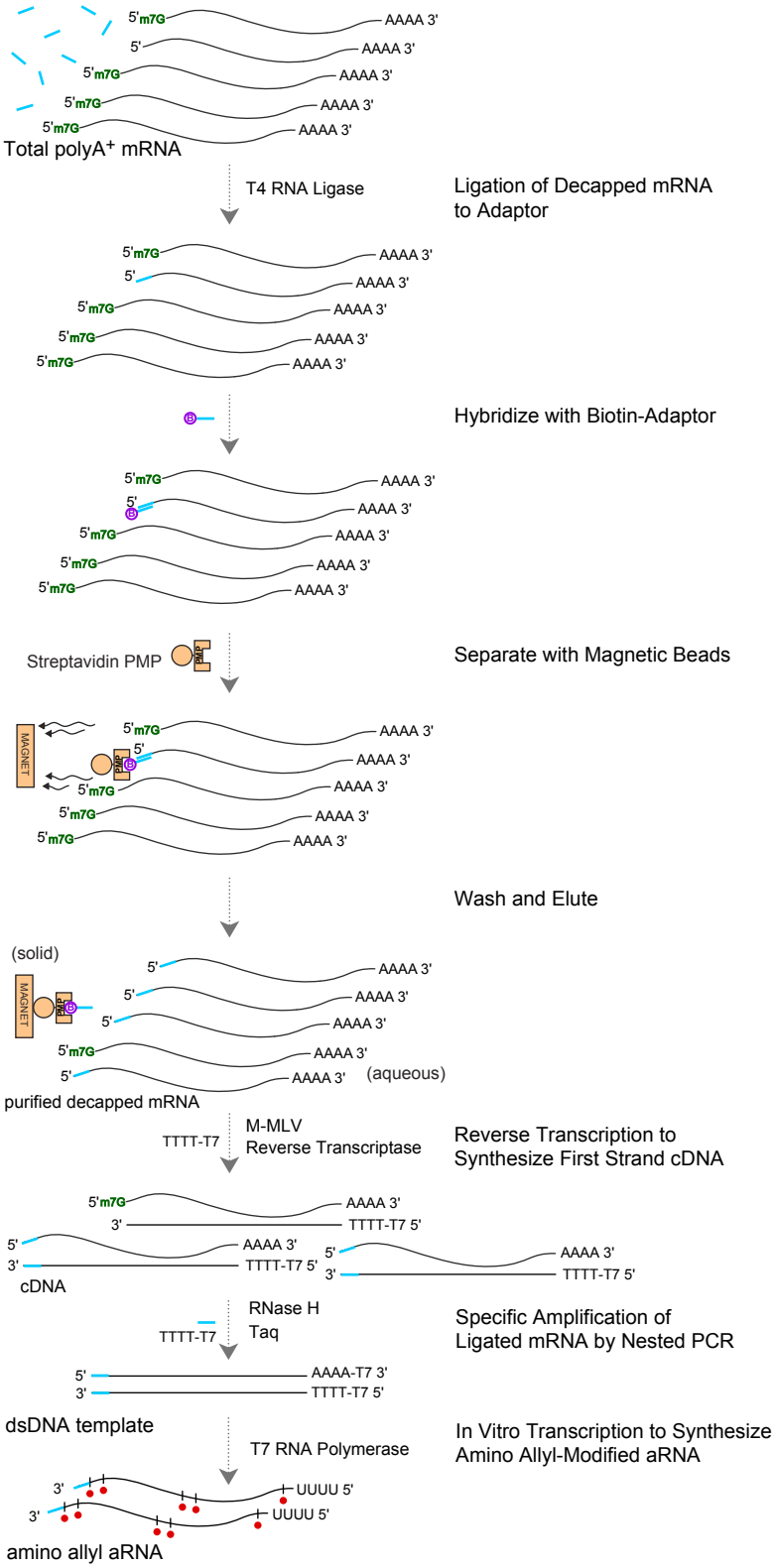
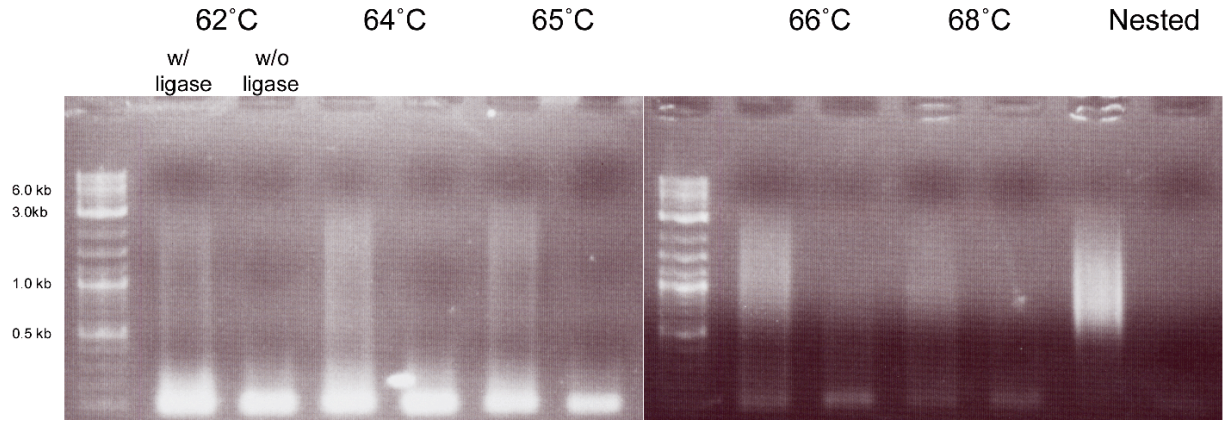
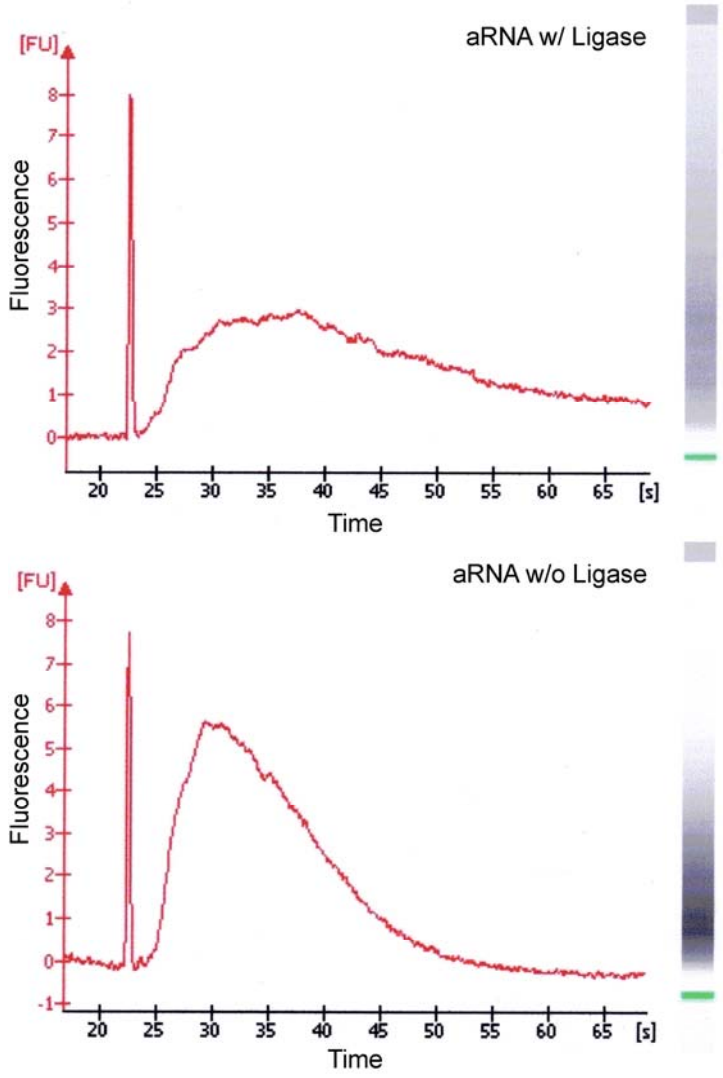
Supplemental Figure 1. Overview of the T4 RNA ligase-mediated method for the isolation of uncapped mRNA.



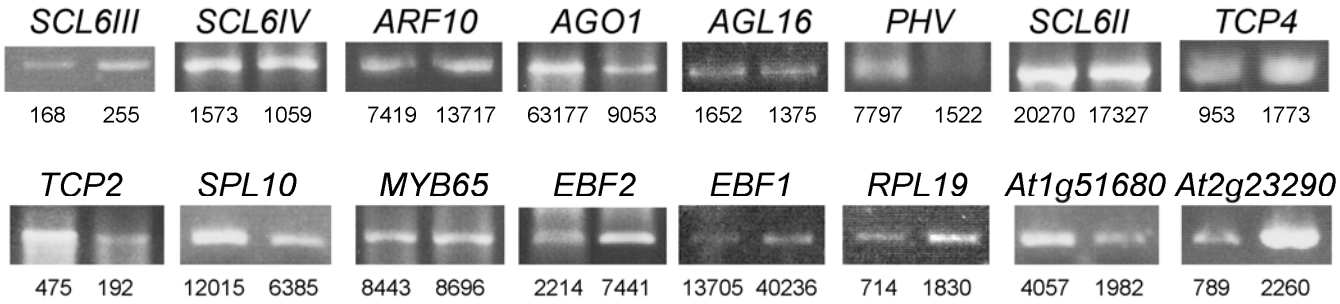
Supplemental Figure 2. Comparison of PCR conditions for second strand synthesis and amplification. The first five pairs were PCR with one primer pair using different annealing temperatures. The last pair was nested PCR performed using the 68 °C annealing temperature. Negative controls were performed without T4 RNA ligase in the initial adaptor ligation step. An extra 25 cycles were added under the same conditions at the end of PCR for each pair for visualization on gels.



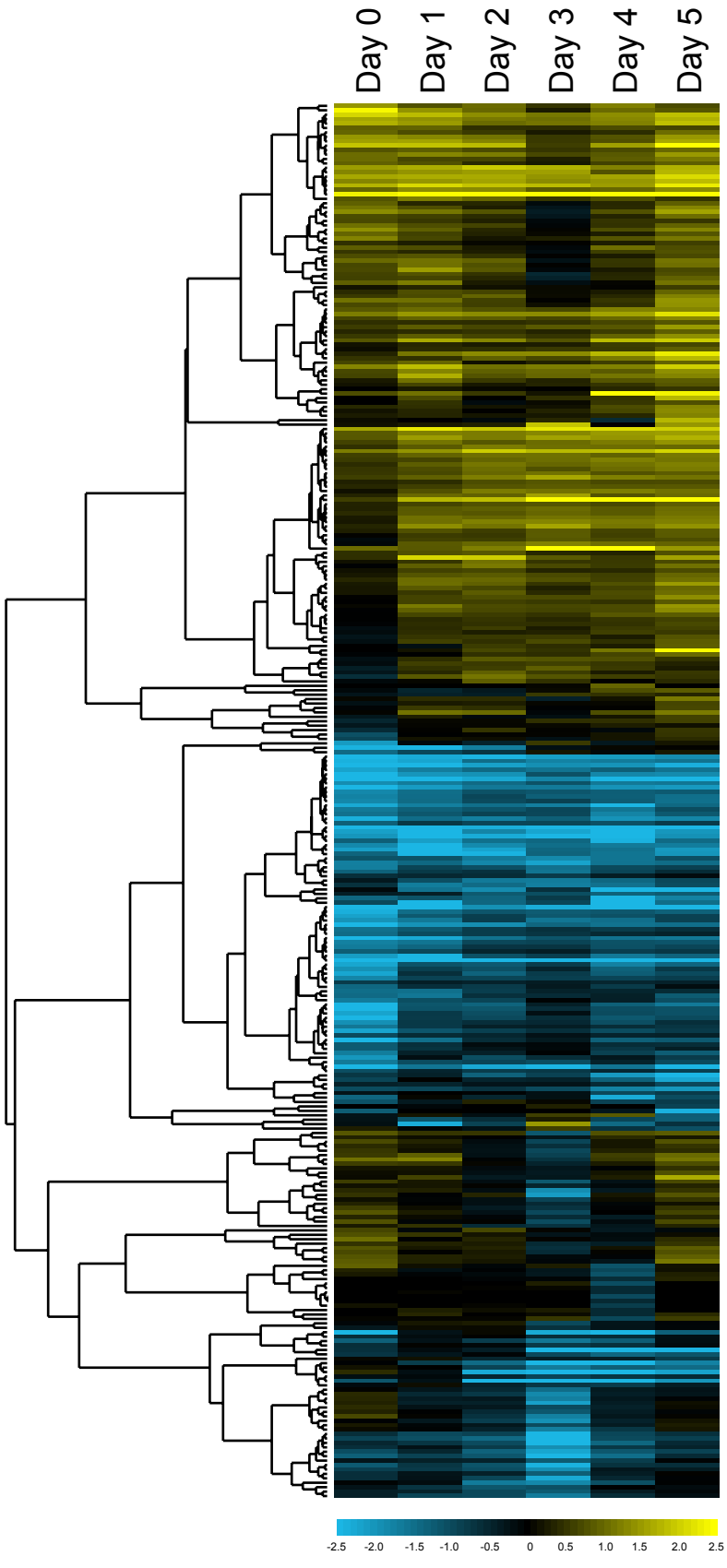
Supplemental Figure 3. Bioanalyzer electropherograms of amplified aRNA samples. Total poly(A)$^+$ mRNA (500 ng) was used for uncapped mRNA isolation and subsequent aRNA synthesis resulting in 112 µg aRNA. A negative control was performed without T4 RNA ligase in the initial adaptor ligation step, which resulted in 2.7 µg aRNA. Both aRNA populations (1 µg each) were analyzed on an Agilent 2100 bioanalyzer. The resulting electropherogram from a positive experiment (upper panel) shows the standard output of aRNA sample. In contrast, the resulting electropherogram from a negative control experiment (lower panel) shows dominating short aRNA.



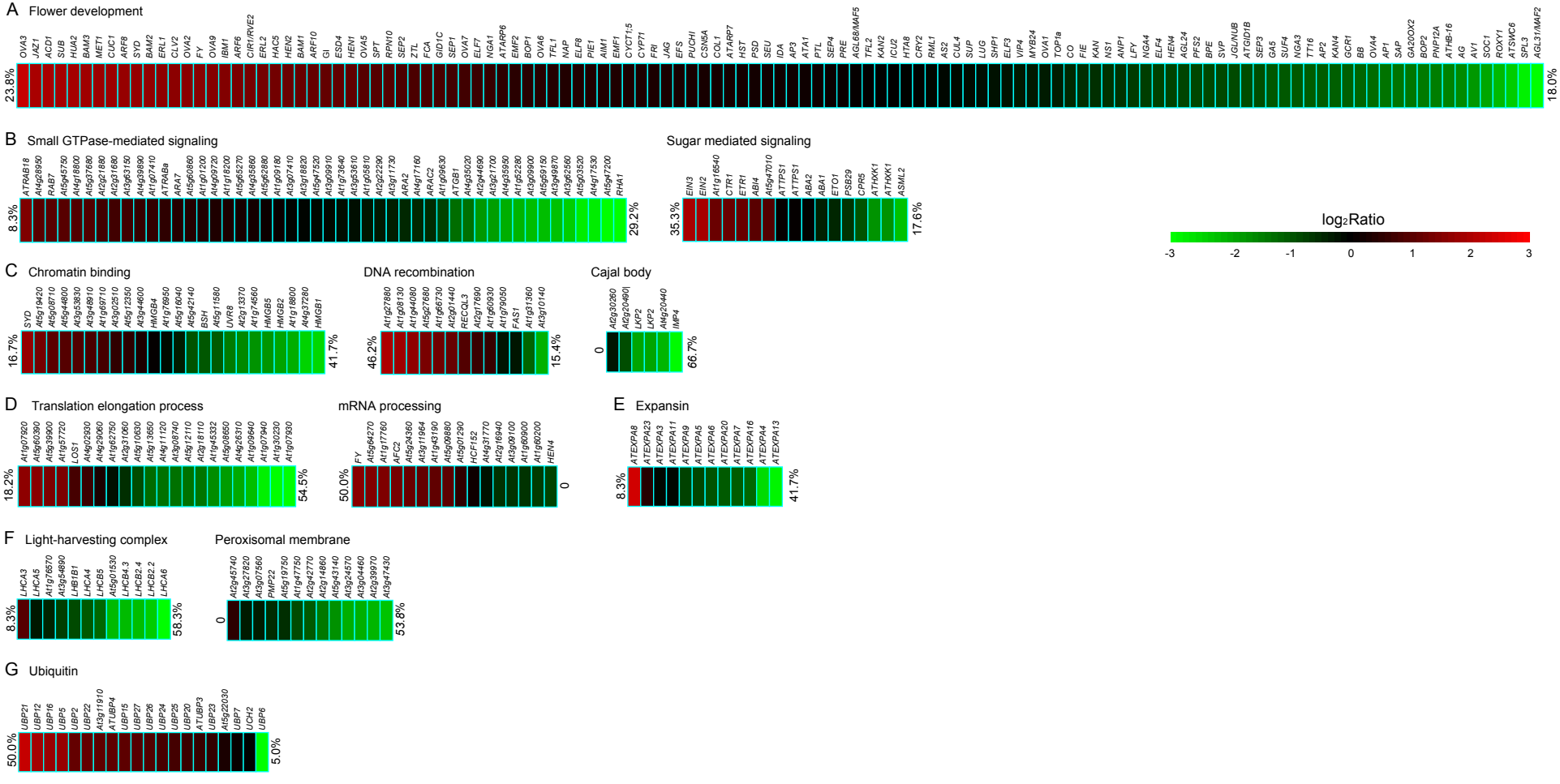
Supplemental Figure 4. RLM-RACE PCR verification of uncapping profiles of sample genes. Uncapped mRNA abundance of selected genes was compared between wild type (left) and *xrn4-1* (right) flower tissues using PCR with gene-specific primers and microarray. PCR products were analyzed using agrose gel and stained with ethidium bromide. Corresponding normalized microarray signal values are shown below PCR gel images.



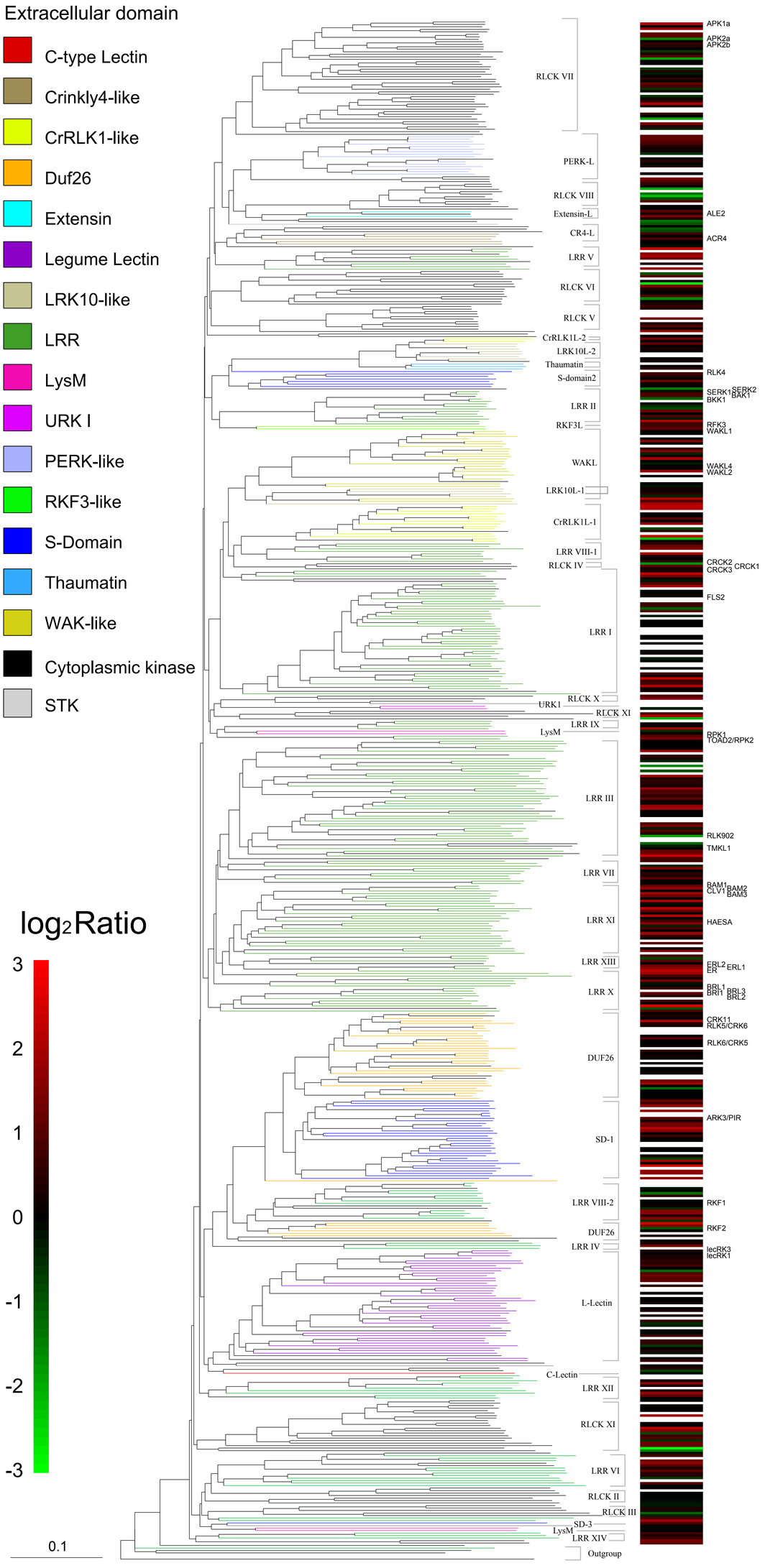
Supplemental Figure 5. Overview of genes with developmentally regulated transcript uncapping levels. Complete linkage hierarchical clustering was performed on transcription ratios of uncapped versus total mRNAs. Each lane represent one time point as labled. Only those transcripts that exhibited differential uncapping levels among six time points (time course $P < 0.05$) were included. A total of 315 transcripts were included in this cluster analysis. Yellow indicates high levels of uncapping and blue indicates low levels of uncapping. The dendrogram on the left shows the relationship among transcripts uncapping levels.



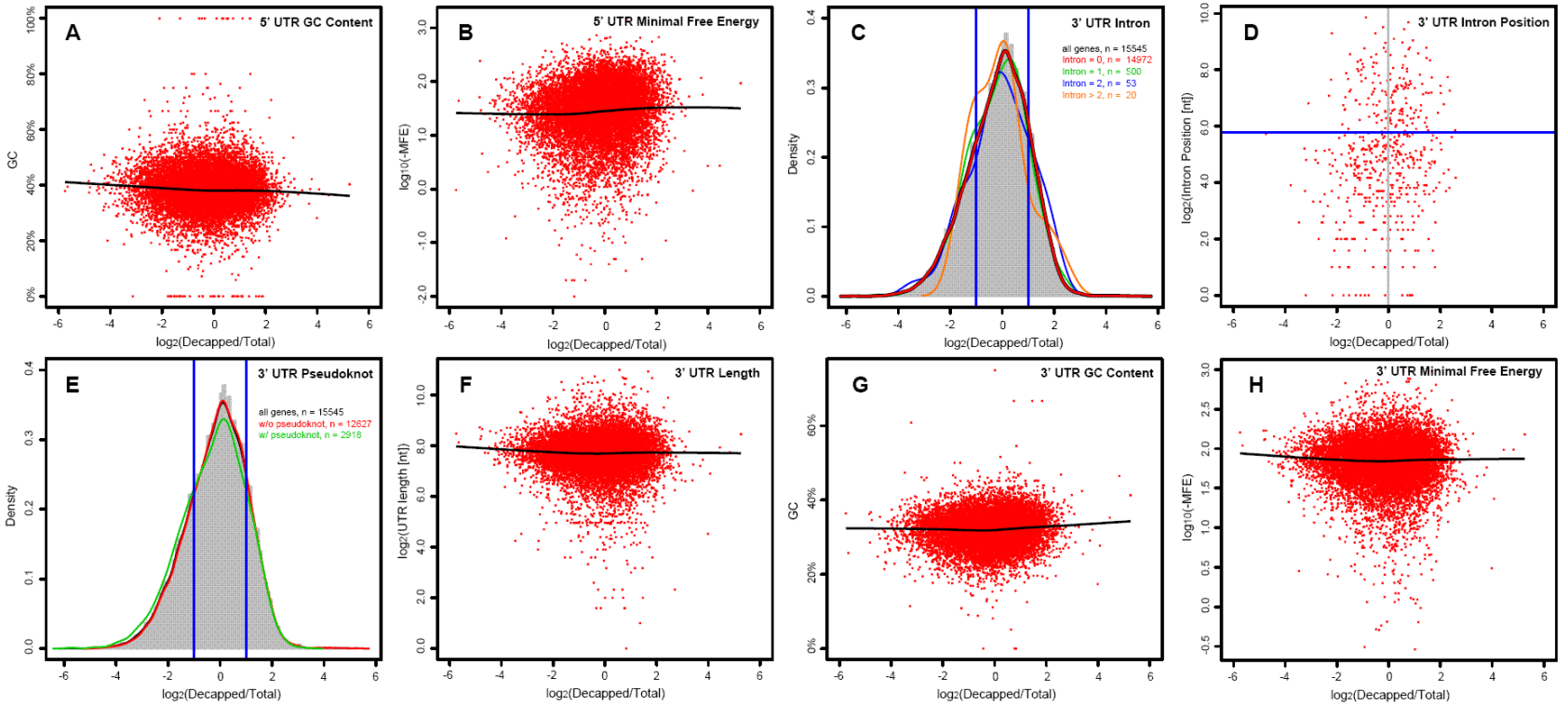
Supplemental Figure 6. Representative patterns of mRNA uncapping in various functional groups and pathways. (A) Development; (B) Signaling; (C) Chromatin; (D) Translation; (E) Cell wall; (F) Cellular components; and (G) Ubiquitin. For each pathway, rectangular blocks represent the measured ratios of relative uncapped mRNA abundance versus total mRNA abundance. Percentages of genes considered as enriched (red) and depleted (green) in the uncapped form were labeled at the right and left sides of each group. Chi-square test P values for all groups were smaller than 0.001.



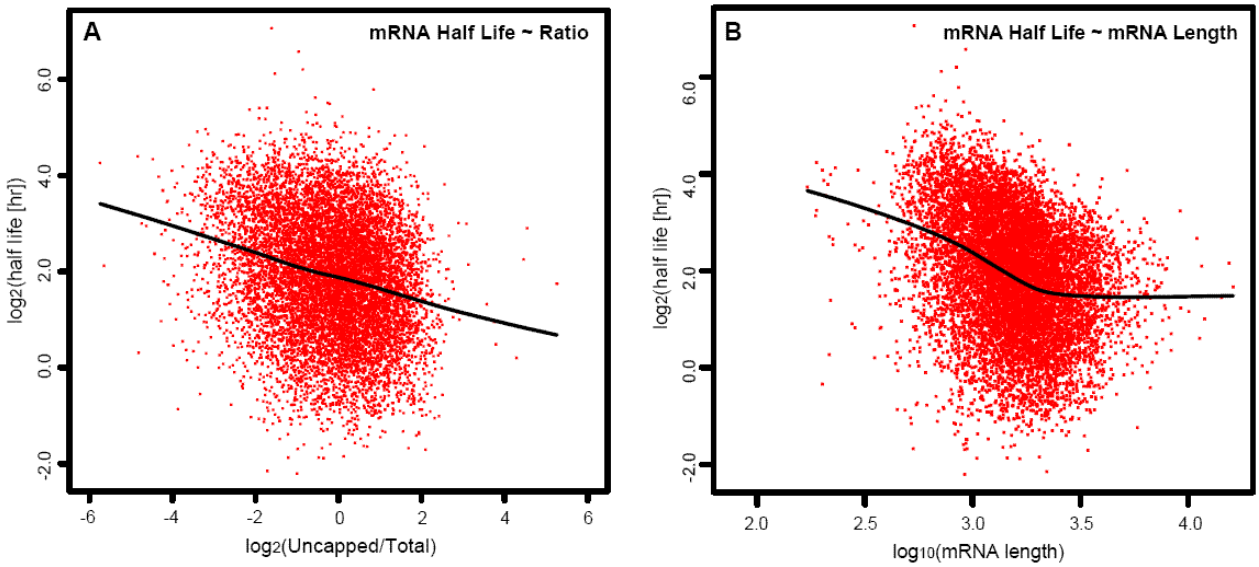
Supplemental Figure 7. Degrees of lineage-specific uncapping in the RLK family. The phylogenetic relationship and subfamily classification were adapted from Shiu et al., 2001. Boxes on the right represented the measured ratios of relative uncapped mRNA abundance versus total mRNA abundance. Red indicates high levels of uncapping and green indicates low levels of uncapping.



Supplemental Figure 8. Relationship between gene transcript features and relative mRNA uncapping levels. (A) 5' UTR GC content; (B) 5'UTR minimal free energy for secondary structure; (C) Number of 3' UTR introns; (D) 3' UTR first intron position; (E) Number of 3' UTR pseudoknots; (F) 3' UTR length; (G) 3' UTR GC content; and (H) 3' UTR minimal free energy for secondary structure. In (C) and (E), the distribution of mRNA uncapping levels of all expressed genes is shown by the histogram and the black line, and colored lines show the distribution of other mRNA groups as labeled. Blue vertical lines show uncapped/total ratios of 0.5 and 2, which were used as cutoffs for depletion and enrichment, respectively. In (A), (B), (D), (F), (G) and (H), each measurement of a sequence feature was plotted against $\log_2$-transformed ratios of relative uncapped versus total mRNA abundance and the black line of each panel is the locally weighted scatterplot smoothing (lowess) fit of the data. In (D), the blue line represents 55 nt after the stop codon.



Supplemental Figure 9. Comparison of mRNA half-lives with uncapping levels and mRNA length. Scatterplot of (A) $\log_2$-transformed mRNA half-life in cultured cells (Narsai et al., 2007) and $\log_2$-transformed uncapping levels; and (B) $\log_2$-transformed mRNA half-life in cultured cells (Narsai et al., 2007) and $\log_{10}$-transformed mRNA length. The black line of each panel is the lowess scatterplot smoothing fit of the data.



Supplemental Methods

Ligation-Mediated Isolation of Uncapped mRNA

Total RNA was extracted from synchronized early developing flower tissues using a Qiagen RNeasy mini kit (Valencia, CA). Poly(A)$^+$ mRNA was purified using FastTrack oligodT cellulose (Invitrogen, Carlsbad, CA) and ligated (500 ng) to an RNA oligo adapter (5' – GCUGAUGGCGAUGAAUGAACACUGCGUUUGCUGGCUUUGAUGAAA- 3', 300 ng) using T4 RNA ligase (10 U, Ambion, Austin, TX) at 37 °C for 90 minutes. Ligation products were then annealed to a biotinylated complementary DNA adapter (synthesized by Integrated DNA Technologies, Coralville, IA) at room temperature for 10 min and purified using Streptavidin MagneSphere Paramagnetic Particles (SA-PMPs, Promega, Madison, MI) according to supplier's instructions. For microarray hybridization, a T7 Oligo(dT) primer was used for first stand cDNA synthesis using an Ambion MessageAmp II aRNA amplification kit. A 10-cycle two-step nested PCR was used for specific second strand synthesis and amplification of uncapped mRNA using outer primer 5' - CGACTGGAGCACGAGGACACTGA - 3' and inner primer 5' - GGACACTGACATGGACTGAAGGAGTA - 3' paired with T7 Oligo(dT) primer. The amplified cDNA was column purified (using Zymoclean columns, Zymo Research, Orange, CA) and was used as template in *in vitro* transcription reactions using the MessageAmp II aRNA amplification kit, which results in production of amino-allyl-tagged, antisense RNA (aRNA) that is labeled with Cy3 and Cy5. The quality of aRNA was checked using a 2100 Bioanalyzer (Agilent, Santa Clara, CA).

To evaluate the possible contamination of the final aRNA population with sequences derived from intact mRNAs with 5' cap structures, mock experiments without T4 RNA ligase were carried out. In such mock experiments, we found that RNA amplification was poor (~2% of the amount obtained in the reactions with ligase), as expected, and resulted in a reduced size

range in the corresponding aRNA population (Figure S3). When applied to the microarray at an arbitrarily balanced input amount, such mock samples gave detectable signals in only a small number of microarray spots, which were mostly non-detectable when samples that had been processed normally were used. This result indicated that contamination from intact mRNAs with a 5' cap was generally below the detectable level in the positive control.

Motif Search

Motif searches were carried out in the UTRs of the 500 genes with the highest relative mRNA uncapping levels and in the UTRs of the 500 genes with the lowest levels. An enumerative algorithm implemented in Sift (Hudson and Quail, 2003) was modified to examine the frequencies of all possible 2- through 12-mer motifs in the 3' or 5' UTRs of target gene transcripts. The number of occurrences of each motif was compared with the frequency of the same motif in UTRs of all microarray covered genes. A chi-squared test (after a Bonferroni correction for multiple testing) with one degree of freedom was used to define significant enrichment or depletion. Each positive candidate identified by raw counting ($P <$ 0.001) was screened to identify those also over-represented on a per-UTR basis (binomial distribution, $P <$ 1E-5). Motifs identified through this pipeline were aligned using CLUSTALW.