Genome-wide suppression of aberrant mRNA-like noncoding RNAs by NMD in *Arabidopsis*

Yukio Kurihara[a], Akihiro Matsui[a], Kousuke Hanada[b], Makiko Kawashima[a], Junko Ishida[a], Taeko Morosawa[a], Maho Tanaka[a], Eli Kaminuma[c], Yoshiki Mochizuki[c], Akihiro Matsushima[c], Tetsuro Toyoda[c], Kazuo Shinozaki[b], and Motoaki Seki[a,d,1]

[a]Plant Genomic Network Research Team, Plant Functional Genomics Research Group, [b]Gene Discovery Research Group, RIKEN Plant Science Center, and [c]Bioinformatics and Systems Engineering Division, RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan; and [d]Kihara Institute for Biological Research, Yokohama City University, 641-12 Maioka-cho, Totsuka-ku, Yokohama 244-0813, Japan

The nonsense-mediated mRNA decay (NMD) pathway is a well-known eukaryotic surveillance mechanism that eliminates aberrant mRNAs that contain a premature termination codon (PTC). The UP-Frameshift (UPF) proteins, UPF1, UPF2, and UPF3, are essential for normal NMD function. Several NMD substrates have been identified, but detailed information on NMD substrates is lacking. Here, we noticed that, in *Arabidopsis*, most of the mRNA-like nonprotein-coding RNAs (ncRNAs) have the features of an NMD substrate. We examined the expression profiles of 2 *Arabidopsis* mutants, *upf1-1* and *upf3-1*, using a whole-genome tiling array. The results showed that expression of not only protein-coding transcripts but also many mRNA-like ncRNAs (mlncRNAs), including natural antisense transcript RNAs (nat-RNAs) transcribed from the opposite strands of the coding strands, were up-regulated in both mutants. The percentage of the up-regulated mlncRNAs to all expressed mlncRNAs was much higher than that of the up-regulated protein-coding transcripts to all expressed protein-coding transcripts. This finding demonstrates that one of the most important roles of NMD is the genome-wide suppression of the aberrant mlncRNAs including nat-RNAs.

ncRNA | tiling array | UPF1 | UPF3

In *Arabidopsis thaliana*, hundreds of nonprotein-coding RNA (ncRNA) transcripts have been discovered based on the cloning of full-length cDNAs, whole-genome tiling array, and deep-sequencing analysis (1–8). They include microRNA (miRNA) precursors, transacting siRNA (ta-siRNA) precursors, and mRNA-like ncRNAs (mlncRNA). The mlncRNAs were divided into 2 types—natural antisense transcript RNAs (nat-RNAs) that arise from the strands opposite the coding strands and other mlncRNAs.

Nonsense-mediated mRNA decay (NMD) is a eukaryotic mRNA quality-control mechanism that eliminates aberrant mRNAs containing a premature termination codon (PTC) from cells to avoid the production of truncated proteins (9–13). Such transcripts can arise by genomic frameshifts, nonsense mutations, inefficiently spliced premRNAs, and so on. Several proteins involved in NMD have already been discovered (10). Among them, the UP-Frameshift (UPF) proteins, UPF1, UPF2, and UPF3, are core components of mRNA surveillance complexes and are essential for normal NMD function. In yeasts and mammals, aberrant transcripts with nonsense mutations are recognized by the UPF complex and then degraded from the 5′ and 3′ ends by recruiting-decapping and 5′→3′ exonuclease activities and deadenylating and 3′→5′ exonuclease activities (14–16).

Plants also have a sophisticated NMD system (10). The *Arabidopsis* genome encodes homologues of 3 UPF genes (*UPF1*, *UPF2*, and *UPF3*). Some *UPF1* and *UPF3* mutants were obtained in *Arabidopsis* and analyzed to investigate the NMD system in plants and to identify several mRNA substrates for NMD by using microarrays (17–20). For example, the aberrant mRNAs

containing PTC were overaccumulated in the *upf1* and *upf3* mutants (17–19).

In plants, aberrant mRNAs with termination codons located distant ($\geq$300 nt) from the 3′ termini of mRNAs or $\geq$50 nt upstream of the last exon–exon junction tend to be recognized as substrates for NMD (21–23) (Fig. 1A). Furthermore, NMD in plants also targets mRNAs that do not possess any introns as in yeast, which is different from NMD in animals (24).

We hypothesized that many mlncRNAs are degraded by the NMD system, because most of the mlncRNAs have the features of NMD substrates and the potential to be recognized as aberrant transcripts by the UPF complex in *Arabidopsis*. Here, we performed whole-genome tiling arrays for *upf1-1* and *upf3-1* mutant plants. The results showed that NMD suppresses not only protein-coding transcripts but also many mlncRNAs. Identification of NMD substrates including many mlncRNAs will aid in understanding the regulatory mechanisms of eukaryotic transcriptomes and elucidating the roles of mlncRNAs.

Results

Features of mlncRNAs as NMD Substrates. Of the mlncRNAs with *Arabidopsis* Genome Initiative (AGI) codes in *Arabidopsis* The *Arabidopsis* Information Resource 8 (TAIR8) genome version, 74.6% (293/393) have one or more introns. Of the mlncRNAs with introns, 39.9% (117/293) have termination codons that are located $\geq$50 nt upstream of the last exon–exon junction when the first AUGs near the 5′ termini are recognized as start codons (Fig. 1B). Notice that the majority of the mlncRNAs with intron(s) have a termination codon located upstream of the last exon–exon junction (Fig. 1B, black bars). However, only 17.0% (3,463/20,374) of the protein-coding mRNAs with intron(s) have the 5′-end-closest termination codons that are located $\geq$50 nt upstream of the last exon–exon junction (Fig. 1B, white bars). They include either the termination codons of the protein-coding ORFs or the termination codons of the upstream ORFs (uORFs) that sometimes exist in front of the major ORFs. Furthermore, of the mlncRNAs with ORFs, 87.3% (331/379) have termination codons that are located >300 nt from the 3′
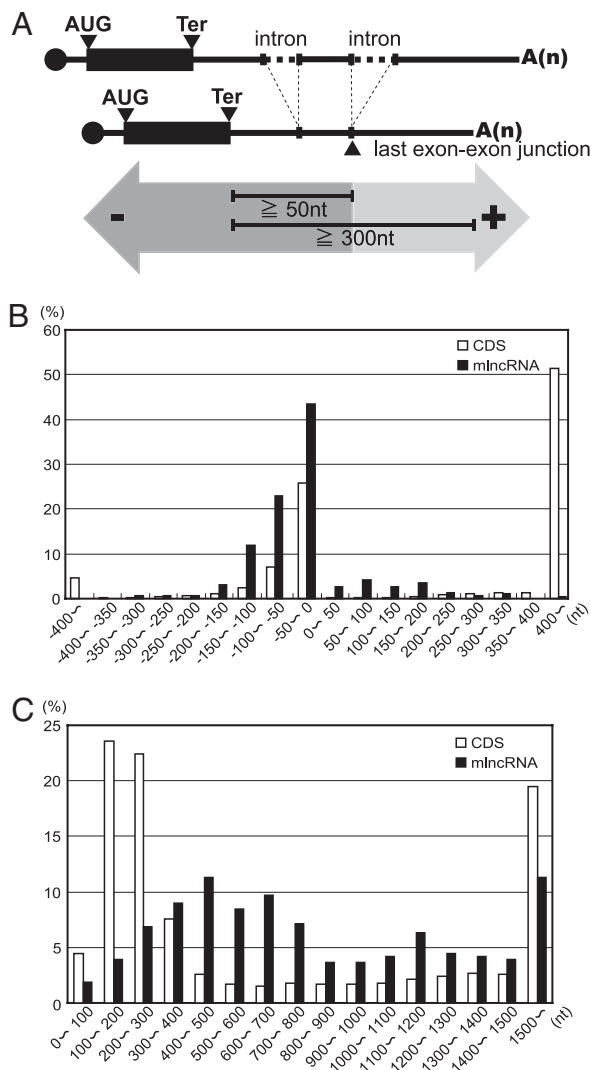
PLANT BIOLOGY

Fig. 1. The features of mRNA-like ncRNAs. (*A*) Illustration of the proposed consensus for the NMD target. The mRNAs with termination codons located distant (≧300 nt) from the 3′ termini of mRNAs or ≧50 nt upstream of the last exon–exon junction tend to be recognized as substrates for NMD. Arrows (−) and (+) indicate upstream and downstream from the last exon–exon junction, respectively. (*B*) Distribution of the distances from the last exon–exon junctions to the 5′-end-closest termination codons of protein-coding mRNAs or mRNA-like ncRNAs (mlncRNAs). (*C*) Distribution of the distances from the 3′ termini of RNAs to the 5′-end-closest termination codons of protein-coding mRNAs or mlncRNAs. Ter, termination codon. CDS, coding sequence.

termini of the mRNA, although, only 49.5% (11,655/23,438) of protein-coding mRNAs have such termination codons (Fig. 1*C*). Thus, many mlncRNAs have the features of a plant NMD target as described above and may be suppressed by NMD.

Overaccumulation of Many mlncRNAs in *upf1-1* and *upf3-1*. To examine whether these mlncRNAs are suppressed by NMD, we performed whole-genome-tiling array analysis by using 2 *upf* mutants: *upf1-1* and *upf3-1*. We focused on the up-regulated transcripts, because the target RNAs should be released and up-regulated when the NMD mechanism is inhibited in the *upf* mutants. In this study, we used the gene annotations of TAIR8 for gene classification. However, we excluded the AGI codes for conserved uORFs from the analyses described throughout this article, because they represent uORFs in the 5′ untranslated regions (UTRs) of the protein-coding mRNAs but not independent transcripts themselves.

Compared with the wild-type accumulation of 237 and 167, AGI-annotated transcripts were increased ≧1.8-fold in *upf1-1* and *upf3-1*, respectively (*P* initial ≦10^{-8}, FDR α = 0.05, Fig. 2*A* and supporting information (SI) Table S1 and Table S2). Of the increased AGI-annotated transcripts, 198 (83.5%) and 138 (82.6%) were protein-coding genes in *upf1-1* and *upf3-1*, respectively (Fig. 2*A* and Table S1 and Table S2). We estimated that 55.6% and 60.1% of those in *upf1-1* and *upf3-1*, respectively, were genes whose transcripts have uORFs in their 5′UTRs, whereas only 31.8% and 32.0% of the protein-coding genes that were not up-regulated in *upf1-1* and *upf3-1*, respectively, had uORFs. This result is consistent with previous reports that many transcripts with uORFs were regulated by the NMD system in mammals and yeasts (16, 25, 26).

Of the increased AGI-annotated transcripts, 31 (13.1%) and 25 (15.0%) transcripts encoded mlncRNAs in *upf1-1* and *upf3-1*, respectively. Furthermore, of the mlncRNAs that were up-regulated, 25 (80.6%) and 20 (80.0%) were classified as nat-RNAs, and 6 (19.4%) and 5 (20.0%) were classified as other mlncRNAs in *upf1-1* and *upf3-1* (Fig. 2*B*). Of the mlncRNAs that were up-regulated in each mutant, 20 were found in both mutants (Fig. 2*C*). Eighteen (90%) of them were nat-RNAs, and two (10%) were other mlncRNAs. For example, accumulation of a nat-RNA from the gene At3g56408 (Fig. 2*D*) and an mlncRNA transcribed from the gene At3g26612 (Fig. 2*E*) are up-regulated in both mutants, but accumulation of the transcripts from the sense gene At3g56410, encoding an unknown protein, is not up-regulated (Fig. 2*D*).

Importantly, all termination codons of 36 AGI-annotated mlncRNAs up-regulated in *upf1-1* and *upf3-1* were located >400 nt upstream (average 1,250 nt upstream) of the 3′ termini. Furthermore, 30 of the up-regulated mlncRNAs have exon–exon junction(s). Their termination codons are located upstream (average 59 nt upstream) of the last exon–exon junction. These features of the up-regulated mlncRNAs in the *upf* mutants almost meet the previously proposed consensus of aberrant RNA molecules targeted by the NMD system (21–23). Therefore, our tiling-array analysis indicated that the mlncRNAs are recognized as the aberrant transcripts and suppressed by the NMD system.

Of all expressed protein-coding AGI-annotated transcripts (16,268 in *upf1* and 15,771 in *upf3*, *P* initial ≦10^{-8}), the transcripts with ≧1.8-fold increase (FDR α = 0.05) were 1.2% and 0.9%, respectively; however, of all the expressed mlncRNAs with AGI codes (148 in *upf1* and 149 in *upf3*, *P* initial ≦10^{-8}), the mlncRNAs with ≧1.8-fold increase (FDR α = 0.05) were 20.9% and 16.8% (Fig. 2*F*). These percentages were much higher than those of the protein-coding transcripts.

We also classified the expressed mlncRNAs (153 in the wild type or *upf1-1*, and 158 in the wild type or *upf3-1*) or the protein-coding transcripts (17,004 in the wild type or *upf1-1*, and 16,800 in the wild type or *upf3-1*) with AGI codes into some groups based on the fold changes (Fig. S1 *A* and *B*, Table S3 and Table S4). The mlncRNAs showed an apparent tendency toward increased accumulation in both *upf1-1* and *upf3-1* mutants (Fig. S1*A*), whereas the protein-coding transcripts relatively did not (Fig. S1*B*). These findings suggested that one of the most important roles of NMD is the genome-wide suppression of aberrant mlncRNAs including nat-RNAs.

Prediction and Characterization of Non-AGI Transcriptional Units (TUs). The ARTADE program (4, 27) to detect expressed genes and unannotated (non-AGI) TUs from tiling-array data predicted 1,752, 1,707, and 1,894 non-AGI TUs in the wild type, *upf1-1* and *upf3-1*, respectively (*P* initial ≦10^{-8}). We classified these TUs into 2,980 nonredundant groups. The non-AGI TUs that are located on the same strand of the genome and overlapped each other by more than one base were classified into the
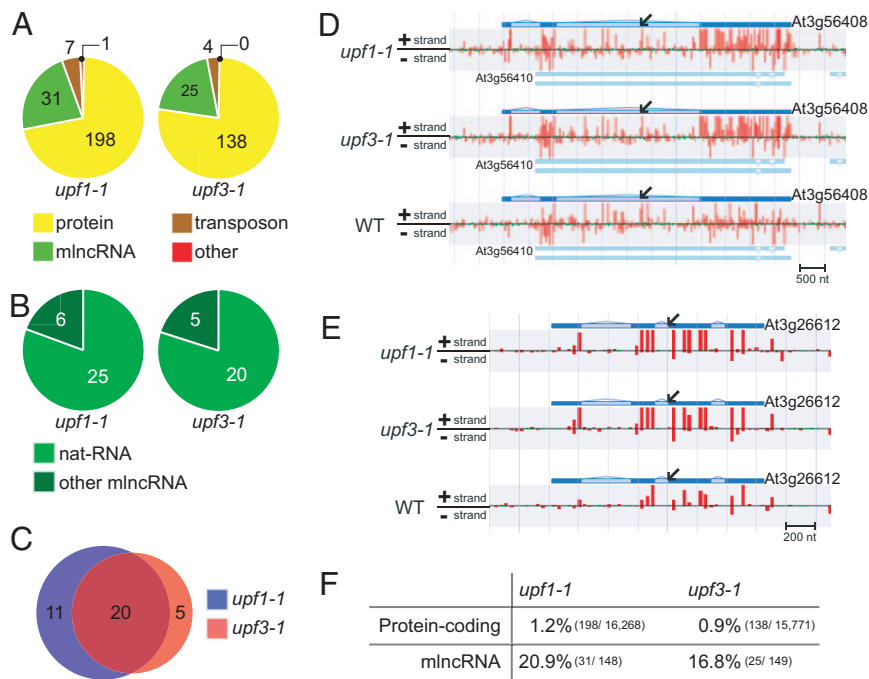
Fig. 2. AGI-annotated transcripts up-regulated in *upf1-1* and *upf3-1*. (*A*) Classification of AGI-annotated transcripts up-regulated ≧1.8-fold in *upf1-1* and *upf3-1*, respectively (*P* initial ≦10^{-8}, FDR α = 0.05). (*B*) Classification of AGI-annotated mlncRNAs up-regulated >1.8-fold (*P* initial ≦10^{-8}, FDR α = 0.05) into 2 types, nat-RNA and other mlncRNA. (*C*) Overlap of AGI-annotated mlncRNAs up-regulated between *upf1-1* and *upf3-1* (*P* initial ≦10^{-8}, FDR α = 0.05). (*D* and *E*) Detection of a nat-RNA (At3g56408) (*D*) and another mlncRNA (At3g26612) (*E*) up-regulated ≧1.8-fold in both *upf1-1* and *upf3-1* (*P* initial ≦10^{-8}, FDR α = 0.05). Black arrows indicate the gene structures of nat-RNA or other mlncRNA in TAIR8. The deep-blue regions are exons and the light-blue regions are introns. The red and green bars indicate the relative signal intensity of probes (red ≧400, green <400). The sense gene, At3g56410, is also shown. The tiling-array expression data are available at http://omicspace.riken.jp/gps/group/psca3. (*F*) Percentages of the protein-coding transcripts and mlncRNAs up-regulated ≧1.8-fold to the expressed protein-coding transcripts and the expressed mlncRNAs (*P* initial ≦10^{-8}, FDR α = 0.05).

same group (Fig. S2). Among the TUs in the same group, 1 TU with the highest intensity was identified as the group-representative TU. We used the predicted gene structure of each group-representative TU for the analyses described below.

Of all the predicted non-AGI TUs, 77 and 59 TUs were up-regulated ≧1.8-fold in *upf1-1* and *upf3-1*, respectively (*P* initial ≦10^{-8}, FDR α = 0.05) (Table S5 and Table S6). Of them, at least 46.8% and 47.5% were supported by cDNA clones. Of the increased non-AGI TUs, 51 and 42 in *upf1-1* and *upf3-1*, respectively, were the TUs for putative nat-RNAs, and 26 and 17, respectively, were the TUs for putative intergenic ncRNAs (incRNA) (Fig. 3*A*). Here, we identified TUs overlapping >1 nt with antisense AGI code genes as nat-RNAs and the rest as incRNAs. Of the non-AGI TUs that were increased, 39 were found in both *upf1-1* and *upf3-1* (Fig. 3*B*). For example, G2116, a putative nat-RNA that was derived from the antisense strand of the gene for PHYTOALEXIN-DEFICIENT 4 (PAD4; At3g52430), was up-regulated in both mutants (Fig. 3*C*), and Group 2870 (G2870), a putative incRNA, was also up-regulated in both mutants (Fig. 3*D*). These results suggest that more mlncRNAs were regulated by the NMD system than estimated from the analysis of AGI-annotated genes shown in Fig. 2.

Validation of the Tiling-Array Results. To validate the tiling-array experiments, quantitative reverse transcription-PCR (RT-PCR) of the mlncRNAs with ≧1.8-fold increase in *upf1-1* and *upf3-1* (*P* initial ≦10^{-8}) (Table S3 and Table S4) was performed. 5′ rapid amplification of cDNA ends (RACE)-based RT-PCR was used to detect strand-specific bands of nat-RNAs (Fig. 4*A*), because sense and antisense transcripts should arise from both strands at the same locus. The sample RNAs were dephosphorylated, decapped, ligated with 5′ adapter, and reverse-

transcribed followed by quantitative PCR by using a 5′ forward primer annealed to the adapter and gene-specific reverse primer (Fig. 4*A*). However, the common quantitative RT-PCR was used to analyze the other mlncRNAs and incRNAs. Among 30 mlncRNAs up-regulated in the array analysis, accumulation of at least 28 mlncRNAs including 17 nat-RNAs was increased in *upf1-1* and *upf3-1* compared with those of the wild type (Fig. 4*B*). Similar results were obtained by using plants containing another mutant allele of *UPF3*, *upf3-2* (Fig. S3).

The 5′RACE-based RT-PCR described above can detect only the capped transcripts but not uncapped transcripts. To examine whether uncapped mlncRNAs were also targeted by NMD, we directly ligated the 5′ adapter to the sample RNAs followed by quantitative RT-PCR. If uncapped mlncRNAs were targeted, this experiment could detect higher signals in the *upf* mutants than in the wild type. However, no mlncRNA-specific amplified bands were detected in either the wild type, *upf1-1* or *upf3-1* (Fig. S4). This result indicated that the capped mlncRNAs, but not uncapped mlncRNAs, were the major substrates of NMD.

To compare the sequences of the up-regulated mlncRNAs among the wild type, *upf1-1* and *upf3-1*, we cloned and sequenced the PCR-amplified products of 5 AGI-annotated mlncRNAs. The sequences were highly homologous (≧99%), and the 5′-end-closest termination codons, which are putative triggers of NMD, were conserved among the wild type, *upf1-1* and *upf3-1*. The results described above suggested that the native mlncRNAs were recognized as aberrant transcripts and were suppressed by the NMD system.

Regulation of the mlncRNAs by NMD Depends on Their Translation. The NMD system is a translation-dependent pathway (9). To obtain further experimental support, we treated wild-type seed-
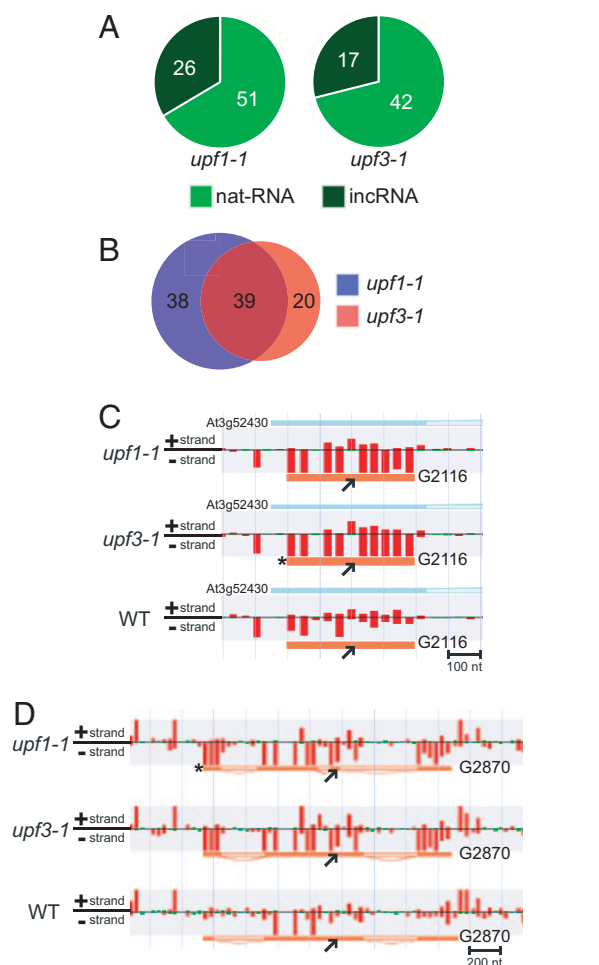
Fig. 3. Predicted non-AGI TUs up-regulated in *upf1-1* and *upf3-1*. (*A*) Classification of predicted mlncRNAs up-regulated ≧1.8-fold in *upf1-1* and *upf3-1* into 2 types, nat-RNA and incRNA (*P* initial ≦10⁻⁸, FDR α = 0.05). (*B*) Overlap of non-AGI TUs up-regulated ≧1.8-fold (*P* initial ≦10⁻⁸, FDR α = 0.05) in *upf1-1* and *upf3-1*. (*C* and *D*) Detection of a putative nat-RNA (G2116) (*C*) and a putative incRNA (G2870) (*D*) up-regulated >1.8-fold in both *upf1-1* and *upf3-1* (*P* initial ≦10⁻⁸, FDR α = 0.05). Black arrows indicate the predicted gene structures. The deep-orange regions are putative exons and the light-orange regions are putative introns. The red and green bars indicate the relative signal intensity of probes (Red ≧400, Green <400). Asterisks show the group representatives in G2116 and G2870, respectively. The sense gene, At3g52430 (PAD4), is also shown as a blue bar.

lings with cycloheximide (CHX), which inhibits translation and thus temporarily suppresses the NMD system (17, 21, 23), and examined the expression of some mlncRNAs. Of the 9 detected mlncRNAs, accumulation of 7 mlncRNAs in the CHX-treated (CHX+) plants was up-regulated compared with those in the mock-treated (CHX-) plants (Fig. 5). These results, and the data in Fig. 4, indicate that the NMD system suppresses aberrant mlncRNAs.

Discussion

In this study, we showed that many mlncRNAs could serve as NMD substrates (Fig. 1) and that not only protein-coding transcripts but also many mlncRNAs, including nat-RNAs, were up-regulated in *upf1-1* and *upf3-1* (Figs. 2 and 3). The location of the 5′-end-closest termination codons of the up-regulated AGI-annotated mlncRNAs is in agreement with the previously proposed consensus of the aberrant RNA molecules targeted by the NMD system (21–23). This finding suggests that the NMD
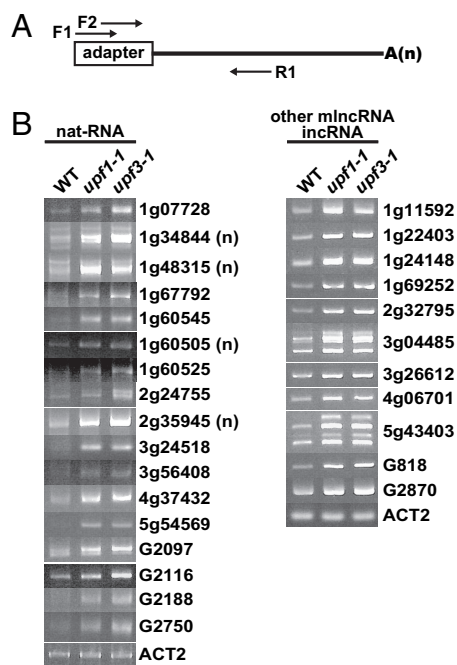

Fig. 4. Quantitative RT-PCR analysis of AGI and non-AGI mlncRNAs. (*A*) Illustration of 5′RACE-based RT-PCR used for nat-RNA detection. The sample RNA was subjected to dephosphorylation and then decapping reaction. The RNA adapter was ligated to the 5′ end of the RNA. After the RT reaction by using a specific oligo(dT) primer, first PCR was performed by using F1 and R1 primer set. When specific signals were not detectable in the first PCR, additional second (nested) PCR was performed by using the first PCR product as a template and F2 and R1 primer set. (*B*) Detection of selected nat-RNAs, incRNAs, and other mlncRNAs with ≧1.8-fold increase (*P* initial ≦10⁻⁸). *ACT2* mRNA was used as an internal control. (n) indicates the result of the second (nested) PCR.

system genome-wide suppresses aberrant mlncRNAs including nat-RNAs. Identification of the NMD-targeted mlncRNAs may help identify the roles of these mlncRNAs.

Some short ORFs that could encode short peptides have been reported to function in biologically important processes (28, 29). In fact, the short ORFs of many mlncRNAs, which could be recognized by NMD, could produce short peptides. It is unknown whether these putative peptides could be translated and whether they have an important function in biological processes. However, their overaccumulation might have an unfavorable effect on plant cells. Thus, NMD may coordinate the amount of these short peptides by suppressing the short peptide-encoding mlncRNAs before the translation.

The NMD system is considered to be a mechanism involved in the degradation of aberrant mRNAs that contain a premature termination codon (PTC) resulting from unexpected errors such as genomic mutations, transcriptional errors, and missplicing (9). However, the percentage of up-regulated mlncRNAs to all expressed mlncRNAs was much higher than the percentage of up-regulated mRNAs to all expressed mRNAs in *upf1-1* and *upf3-1* (Fig. 2*F* and Fig. S3). This result suggests that one of the most important roles of NMD is the suppression of the mlncRNAs that are recognized as aberrant transcripts.

Many transcripts including hundreds of intergenic ncRNAs are overaccumulated in RNAi knockdown lines of core subunits of the exosome in *Arabidopsis* (3). The transcripts up-regulated in the *upf* mutants may overlap with exosome substrates identified previously, because, in the NMD pathway, the RNAs recognized as the aberrant transcripts by the UPF complex should be degraded from the 3′ end by deadenylation and
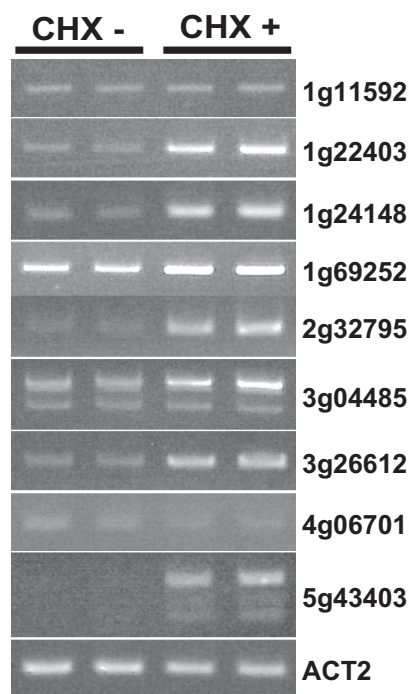
Fig. 5. Quantitative RT-PCR analysis of some mlncRNAs in the cycloheximide (CHX)-treated plants. Two independent samples from CHX-untreated (CHX-) and CHX-treated (CHX+) plants, respectively, were loaded. *ACT2* mRNA was used as an internal control.

subsequent 3′→5′ exonuclease activity, which is probably included in the exosome (14–16). However, only subtle overlaps (6 AGI-annotated transcripts) were found between the transcripts up-regulated in the *upf* mutants and exosome substrates (Table S1 and Table S2). In addition, the previously identified population of exosome substrates does not include any nat-RNAs. These differences are probably because of the difference of growth condition, age of plants used, and statistical analysis method.

Of the AGI-annotated transcripts up-regulated in *upf1-1* and *upf3-1*, 8 transcripts were derived from transposon genes (Fig. 2A, Table S1 and Table S2). In addition, 4 up-regulated nat-RNAs were derived from the antisense strands of nonexpressed sense genes annotated as transposons or pseudogenes (Table S1 and Table S2). Taken together, of the 97 non-AGI TUs up-regulated (Fig. 3A), 21 (22%) TUs include short segment(s) of the repeat sequence(s) originating from transposable elements (*E* value <0.01) (Table S5 and Table S6). The transposon-associated ncRNAs have been reported to be suppressed by DNA methylation at the transcriptional level (2, 6). The relationship between DNA methylation and the suppression of mlncRNA by the NMD system is as yet unknown.

Nuclear cap binding protein 80 (CBP80) promotes the interaction of UPF1 with UPF2 in mammals (30). The *Arabidopsis* CBP80 homolog, ABA HYPERSENSITIVE1 (ABH1), is involved in premRNA splicing and processing of miRNA precursor (7, 8). Inactivation of ABH1 that results in decreased levels of mature miRNAs is accompanied by apparent stabilization of the precursors. However, our tiling-array analysis of *upf* mutants did not reveal any remarkable accumulation of miRNA precursors. Therefore, the NMD system is probably involved in the suppression of the aberrant mlncRNAs, not in the processing of miRNA.

Materials and Methods

Plant Materials and RNA Extraction. The Col-0 ecotype of *Arabidopsis* was used in this study. The mutants *upf1-1* (point mutation), *upf3-1* (SALK_025175), and *upf3-2* (SALK_097931) were as described previously (14–16). Plants were grown in plastic dishes (30 plants per plastic dish) containing GM agar (0.85%) medium supplemented with 1% sucrose <16-h-light/8-h-dark (40–80 μmol of photons m^{-2} sec^{-1}) essentially as described in ref. 4. Total RNA was extracted from 15-day-old seedlings by using Isogen reagent (NIPPON GENE), precipitated with 1/3 volume of 8M LiCl, and resolved in RNase-free DEPC water.

Whole-Genome-Tiling Array and Analysis. The GeneChip *Arabidopsis* tiling-array set (1.0F Array and 1.0R Array, Affymetrix) was used (2, 4). Eight micrograms per array of the total RNA extracted from 15-day-old seedlings was used for probe synthesis. Probe synthesis, array hybridization, and computational analyses of RNA expression were performed as described in ref. 4. Three independent biological replicates were performed for each strand array. The ARTADE-based method (*P* initial $\leq 10^{-8}$) was used to detect the expressed AGI-annotated genes and predict the non-AGI TUs from the expression data (4, 27). We identified the AGI transcripts and the non-AGI TUs predominantly up-regulated in both *upf* mutants by Mann–Whitney *U* test (FDR α = 0.05) as described in ref. 4.

The *Arabidopsis* genome annotation used in this analysis was based on the TAIR8 genome version (ftp://ftp.arabidopsis.org/home/tair/Genes/TAIR8_genome_release/TAIR8_functional_descriptions) as of May 5, 2008. Of the AGI-annotated genes annotated as ''other RNA,'' the genes annotated as ''Potential natural antisense gene'' were identified as nat-RNAs, and the rest except transacting siRNA precursors were as other mlncRNAs in this study. When 1 AGI-annotated gene had some structural variations, 1 variant with the youngest variant number was selected for use in the analyses of array data.

CHX Treatment. Fifteen-day-old *Arabidopsis* seedlings were vacuum-infiltrated with 10 μg/ml CHX (Nacalai Tesque) in the buffer (0.046 g/L Murashige and Skoog Plant Salt Mixture, 0.3 g/L sucrose, pH 5.8). The seedlings were incubated on the bench at room temperature for 3 h, followed by total RNA extraction.

Quantitative RT-PCR Analysis. The total RNA was subjected to DNase I (Takara) treatment before the RT reaction. cDNA was synthesized from 4 μg of the total RNA in 40 μL of mixture by using a Primescript 1st strand cDNA synthesis kit (Takara) and oligo(dT) primer. PCR was performed in 20 μL of mixture with 0.5 μL of the RT product, Ex Taq polymerase (Takara) and the respective gene-specific primer sets. The cycling parameters were 94 °C for 2 min and 30 sec, 25 or 30 cycles of 20 sec at 94 °C, 30 sec at 55 °C, 30 sec at 72 °C, and a final elongation step at 72 °C for 2 min and 30 sec. Primer sets used are listed in Table S7. PCR products were subjected to electrophoresis on 2% agarose gel followed by ethidium bromide staining for visualization.

Quantitative 5′RACE-Based RT-PCR analysis. The total RNA was subjected to DNase I (Takara) treatment before the series of treatments. GeneRacer Kit (Invitrogen) was used according to the modified protocol described below. The total RNA (5 μg) was treated with calf intestinal phosphatase (CIP) followed by treatment of tobacco acid pyrophosphatase (TAP) to remove the 5′ cap structure. Three micrograms of the resulting RNA were subjected to ligation with GeneRacer RNA oligo adapter to 5′ end of the RNA by using T4 RNA ligase. The ligated RNA was reverse-transcribed to synthesize cDNA by using SuperScript III RT and the GeneRacer Oligo dT primer in 20 μL of mixture. First PCR was performed in 20 μL of mixture with 0.25 μL of RT product, Ex Taq polymerase, GeneRacer 5′ forward primer, and the gene-specific reverse primer. The cycling parameters of first PCR were 94 °C for 2 min and 30 sec, 40 cycles of 20 sec at 94 °C, 30 sec at 60 °C, 30 sec at 72 °C, and a final elongation step at 72 °C for 2 min and 30 sec. Second (nested) PCR was performed in a 20 μL of mixture with 1 μL of the first PCR product, Ex Taq polymerase, GeneRacer 5′ nested forward primer, and the gene-specific reverse primer. The cycling parameters of nested PCR were 94 °C for 2 min and 30 sec, 20 cycles of 20 sec at 94 °C, 30 sec at 60 °C, 30 sec at 72 °C, and a final elongation step at 72 °C for 2 min and 30 sec. Visualization was as described above. The forward and reverse primers used are listed in Table S8.

1. Yamada K, *et al.* (2003) Empirical analysis of transcriptional activity in the *Arabidopsis* genome. *Science* 302:842–846.
2. Zhang X, *et al.* (2006) Genome-wide high resolution mapping and functional analysis of DNA methylation in *Arabidopsis. Cell* 126:1189–1201.
3. Chekanova JA, *et al.* (2007) Genome-wide high resolution mapping of exosome substrates reveals hidden features in the *Arabidopsis* transcriptome. *Cell* 131:1340–1353.
4. Matsui A, *et al.* (2008) *Arabidopsis* transcriptome analysis under drought, cold, high-salinity and ABA treatment conditions using a tiling array. *Plant Cell Physiol* 49:1135–1149.
5. Seki M, *et al.* (2002) Functional annotation of a full-length *Arabidopsis* cDNA collection. *Science* 296:141–145.
6. Lister R, *et al.* (2008) Highly integrated single-base resolution maps of the epigenome in *Arabidopsis. Cell* 133:523–536.
7. Gregory BD, *et al.* (2008) A link between RNA Metabolism and silencing affecting *Arabidopsis* development. *Dev Cell* 14:854–866.
8. Laubinger S, *et al.* (2008) Dual roles of the nuclear cap-binding complex and SERRATE in pre-mRNA splicing and microRNA processing in *Arabidopsis thaliana. Proc Natl Acad Sci USA* 105:8795–8800.
9. Maquat LE (2004) Nonsense-mediated mRNA decay: Splicing, translation and mRNP dynamics. *Nat Rev Mol Cell Biol* 5:89–99.
10. Behm-Ansmant I, *et al.* (2007) mRNA quality control: An ancient machinery recognizes and degrades mRNAs with nonsense codons. *FEBS Lett* 581:2845–2853.
11. Conti E, Izaurralde E (2005) Nonsense-mediated mRNA decay: Molecular insights and mechanistic variations across species. *Curr Opin Cell Biol* 17:316–325.
12. Culbertson MR, Leeds PF (2003) Looking at mRNA decay pathways through the window of molecular evolution. *Curr Opin Genet Dev* 13:207–214.
13. Rehwinkel J, Raes J, Izaurralde E (2006) Nonsense-mediated mRNA decay: Target genes and functional diversification of effectors. *Trends Biochem Sci* 31:639–646.
14. Mitchell P, Tollervey D (2003) An NMD pathway in yeast involving accelerated deadenylation and exosome-mediated 3′→5′ degradation. *Mol Cell* 11:1405–1413.
15. Lejeune F, Li X, Maquat LE (2003) Nonsense-mediated mRNA decay in mammalian cells involves decapping, deadenylating, and exonucleolytic activities. *Mol Cell* 12:675–687.
16. He F, *et al.* (2003) Genome-wide analysis of mRNAs regulated by the nonsense-mediated and 5′ to 3′ mRNA decay pathways in yeast. *Mol Cell* 12:1439–1452.
17. Hori K, Watanabe Y (2005) UPF3 suppresses aberrant spliced mRNA in *Arabidopsis. Plant J* 43:530–540.
18. Yoine M, *et al.* (2006) The *lba1* mutation of UPF1 RNA helicase involved in nonsense-mediated mRNA decay cause pleiotropic phenotypic changes and altered sugar signaling in *Arabidopsis. Plant J* 47:49–62.
19. Arciga-Reyes L, Wootton L, Kieffer M, Davies B (2006) UPF1 is required for nonsense-mediated mRNA decay (NMD) and RNAi in *Arabidopsis. Plant J* 47:480–489.
20. Yoine M, Nishii T, Nakamura K (2006) *Arabidopsis* UPF1 RNA helicase for nonsense-mediated mRNA decay is involved in seed size control and is essential for growth. *Plant Cell Physiol* 47:572–580.
21. Kertesz S, *et al.* (2006) Both introns and long 3′-UTRs operate as *cis*-acting elements to trigger nonsense-mediated decay in plants. *Nucleic Acids Res* 34:6147–6157.
22. Schwartz AM, *et al.* (2006) Stability of plant mRNAs depends on the length of the 3′-untranslated region. *Biochemistry* (*Moscow*) 71:1377–1384.
23. Hori K, Watanabe Y (2007) Context analysis of termination codons in mRNA that are recognized by plant NMD. *Plant Cell Physiol* 48:1072–1078.
24. Kerenyi Z, *et al.* (2008) Inter-kingdom conservation of mechanism of nonsense-mediated mRNA decay. *EMBO J* 27:1585–1595.
25. Mendell JT, Dietz HC (2001) When the message goes awry: Disease-producing mutations that influence mRNA content and performance. *Cell* 107:411–414.
26. Mendell JT, *et al.* (2004) Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. *Nat Genet* 36:1073–1078.
27. Toyoda T, Shinozaki K (2005) Tiling array-driven elucidation of transcriptional structures based on maximum-likelihood and Markov models. *Plant J* 43:611–621.
28. Galindo MI, Pueyo JI, Fouix S, Bishop SA, Couso JP (2007) Peptides encoded by short ORFs control development and define a new eukaryotic gene family. *PLoS Biol* 5:e106.
29. Kondo T, *et al.* (2007) Small peptide regulators of actin-based cell morphogenesis encoded by a polycistronic mRNA. *Nat Cell Biol* 9:660–665.
30. Hosoda K, Kim YK, Lejeune F, Maquat LE (2005) CBP80 promotes interaction of Upf1 with Upf2 during nonsense-mediated mRNA decay in mammalian cells. *Nat Struct Mol Biol* 12:893–901.