Both introns and long 3'-UTRs operate as cis-acting elements to trigger nonsense-mediated decay in plants

Sándor Kertész1, Zoltán Kerényi1, Zsuzsanna Mérai1,2, Imre Bartos3, Tamás Pálfy1, Endre Barta1 and Dániel Silhavy1,*

1Agricultural Biotechnology Center, Gödöllő, Hungary, 2Department of Genetics, Eötvös Loránd University, Budapest, Hungary and 3Institute of Physics, Eötvös Loránd University, Budapest, Hungary

Received June 16, 2006; Revised September 1, 2006; Accepted September 22, 2006

ABSTRACT

Nonsense-mediated mRNA decay (NMD) is a eukaryotic quality control mechanism that identifies and eliminates aberrant mRNAs containing a premature termination codon (PTC). Although, key trans-acting NMD factors, UPF1, UPF2 and UPF3 are conserved in yeast and mammals, the cis-acting NMD elements are different. In yeast, short specific sequences or long 3'-untranslated regions (3'-UTRs) render an mRNA subject to NMD, while in mammals’ 3'-UTR located introns trigger NMD. Plants also possess an NMD system, although little is known about how it functions. We have elaborated an agroinfiltration-based transient NMD assay system and defined the cis-acting elements that mediate plant NMD. We show that unusually long 3'-UTRs or the presence of introns in the 3'-UTR can subject mRNAs to NMD. These data suggest that both long 3'-UTR-based and intron-based PTC definition operated in the common ancestors of extant eukaryotes (stem eukaryotes) and support the theory that intron-based NMD facilitated the spreading of introns in stem eukaryotes. We have also identified plant UPF1 and showed that tethering of UPF1 to either the 5' or 3'-UTR of an mRNA results in reduced transcript accumulation. Thus, plant UPF1 might bind to mRNA in a late, irreversible phase of NMD.

INTRODUCTION

Nonsense-mediated mRNA decay (NMD) is a eukaryotic quality control system that identifies and degrades mRNAs containing premature termination codons (PTC), thereby preventing the accumulation of truncated proteins, which might act as dominant-negative mutants. In addition to eliminating aberrant mRNAs, NMD regulates the expression of wild-type genes, as ~5–10% of the yeast, human or Drosophila transcriptome is altered in NMD deficient cells (1–4).

To identify PTC-containing transcripts, the NMD machinery should efficiently discriminate between authentic stop codons and PTCs. At least two (not mutually exclusive) models have been suggested to explain how NMD recognizes PTCs. The first model (often referred to as the pioneer translation or nuclear marking model) suggests that cis-acting NMD elements are present in the coding region but not in the 3'-untranslated region (3'-UTR) of wild-type mRNAs. In the nucleus, NMD trans-acting factors bind to the NMD cis elements and are transported with the mRNAs into the cytoplasm. During the pioneer round of translation, ribosomes displace the NMD trans-acting factors from wild-type mRNAs. However, if translation is stopped at a PTC, the NMD trans-acting factors bound downstream of the PTC are retained on the mRNA. These mRNA-bound NMD factors recruit further NMD components including UPF1, thereby leading to formation of functional NMD complexes and rapid degradation of the PTC-containing mRNA (5–7).

The second model (the faux UTR model) proposes that translation termination on PTC-containing mRNAs is aberrant because their 3'-UTR lacks the factors required for normal termination. This model suggests that aberrant translation termination results in assembly of a functional NMD complex, thus leading to rapid decay of PTC-containing mRNAs (8,9).

UPF3, UPF2 and UPF1 proteins are required for NMD in yeast, worm, Drosophila and mammals (10). Although these core trans-acting NMD factors are conserved, the cis-acting NMD elements are different in mammals and in yeast. In mammalian cells, introns are believed to be the predominant cis-acting NMD elements (11). When either a U12 or a U2 intron is spliced, a multiprotein complex called the exon junction complex (EJC) is deposited on the mRNA 20–25 nt upstream of each exon-exon junction (12,13). As introns

*To whom correspondence should be addressed at H-2101 Gödöllő, P.O. Box 411, Hungary. Tel: +36 28 526 194; Fax: +36 28 526 145; Email: silhavy@abc.hu

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors

© 2006 The Author(s).

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/2.0/uk/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
are rare in the 3'-UTR, the EJC marks the coding regions of mRNAs. UPF3, a component of the mammalian EJC, recruits UPF2 in the perinuclear space. If translation is stopped at a PTC, UPF3–UPF2 complexes located downstream of the PTC can recruit UPF1 and other NMD factors to form functional NMD complex. As translating ribosome remove EJCs located upstream of the stop codon as well as EJCs that are located downstream but in close proximity to a stop codon, mammalian transcripts are targeted by NMD only if a stop codon resides >50–55 nt upstream of an exon–exon junction (14). In yeast, loosely defined sequences called DSE elements can render an mRNA subject to NMD if they are located downstream of a stop codon (15). Hrp1p specifically binds to DSE sequences and might recruit UPF proteins to the mRNAs (16). Both the intron (EJC) and DSE-based PTC definition systems conform to the nuclear marking model. However, a long 3'-UTR can also subject mRNAs to NMD in yeast, *Drosophila* and worm cells (8,17,18). Moreover, mammalian mRNAs with long 3'-UTRs are also targeted by NMD, although less efficiently than mRNAs containing introns in the 3'-UTR (19). The finding that long 5'-UTRs act as NMD cis elements can be explained by the *faux* UTR model. Normal translation termination requires the interaction of terminating ribosome and poly(A)-binding protein (PABP). By preventing this interaction, an unusually long 3'-UTR could lead to aberrant termination and formation of functional NMD complex (8,20).

The events of NMD-mediated mRNA decay downstream of functional NMD complex formation are not well understood. In yeast and human cells, NMD complex formation results in decapping and deadenylation of PTC-containing mRNA, while in *Drosophila*, it leads to cleavage of aberrant mRNAs close to the PTC (21). Decapped and/or deadenylated mRNAs are degraded in specific cellular compartments called P-bodies by normal mRNA decay pathways involving XRN1 and the exosome (22). In animals, SMG-1, SMG-5, SMG-6 and SMG-7 are also required for NMD. SMG-1 phosphorylates UPF1, while SMG-5, SMG-6 and SMG-7 are involved in dephosphorylation of UPF1 (23).

NMD also operates in higher plants, although very little is known about mechanism or function of plant NMD systems (24,25). A computer search identified one putative ortholog for UPF1, UPF2 and UPF3 in *Arabidopsis* genome (10). Indeed, alternative splicing products containing PTCs accumulate to high levels in an *Arabidopsis* line carrying a UPF3 mutation (26). Moreover, it has been reported that the null mutant of putative UPF1 is lethal (27,28), while the missense mutant *lbat* shows strong phenotype (27,29). The cis-acting NMD elements have not yet been identified in plants. The findings that PTC-containing mRNAs derived from intronless genes are efficiently targeted by plant NMD (30–33) suggest that long 3'-UTRs or DSE-like elements can act as cis factors in plants. The role of plant introns in PTC definition is still under debate. In rice waxy gene, splicing of an intron upstream of the PTC affected the efficiency of NMD. These data were interpreted to indicate that splicing and NMD were coupled in plants, as in mammals (34). However, in mammals only introns positioned downstream of the PTC could act as NMD cis elements. Moreover, other reports suggested that plant introns present in the 3'-UTR did not subject mRNAs to NMD (35,36).

We have elaborated an agroinfiltration-based transient NMD test system to define the cis-acting elements of plant NMD and to identify plant UPF1. We show that both long 3'-UTRs and introns located in the 3'-UTR act as NMD cis factors. As both long 3'-UTR-based and intron-based PTC identification systems operate in plants as well as in animals, it is likely that these PTC definition systems already existed in the common ancestor of stem eukaryotes. We have also shown that tethering of UPF1 to either the 5'- or 3'-UTR causes a dramatic reduction of target mRNA levels, suggesting that in plants, UPF1 binds to the mRNA in a late, irreversible phase of NMD.

**MATERIALS AND METHODS**

**Plasmid constructs**

For the agroinfiltration assays, genes were cloned into the BIN61S binary vector between the 35S promoter and terminator sequences. Binary vectors contain all cis sequences required for *Agrobacterium*-mediated plant transformation or transient expression (37). P14 was previously described (38). Brief descriptions of cloning of unpublished constructs are given here, whereas primer sequences are available as Supplementary Figure 1.

To create binary PHA clone, PHA fragment was PCR amplified with PHA1-F/PHA2-R primers from common bean, cloned first to pBluescript KS, and then it was moved into BIN61S. PCR-based *in vitro* mutagenesis (PHA1mut-F/PHA2mut-R primers) of PHA was carried out with Quick-Change Mutagenesis Kit (Stratagene) to create PHA-m. PHA-s, which contains the ORF of PHA-m from ATG to PTC, was PCR amplified from PHA-m (PHA1-F/PHA2mut-R primers), and then cloned into BIN61S. UPF1 was RT–PCR amplified (Stratagene, ProSTAR® ULTRA HF RT–PCR System) from *Arabidopsis thaliana* leaf RNA with UPF1-F/UPF1-R primers, cloned into pBluescript KS (KSUFP1), and then UPF1 was moved into BIN61S. UPF1DN was obtained by introducing point mutations into KSUPF1 with R863C-F/R863C-R primers, and then the mutant fragment was cloned into BIN61S. To create GFP binary construct, GFP fragment lacking the endoplasmic retention signal was PCR amplified from 35SGFP construct (39) with mGFP4-F/mGFP4-R primers and cloned to BIN61S, abc, bc, c, b and a fragments were PCR amplified from PHA-m with PHAB1-F/PHAX1-R, PHAB2-F/PHAX1-R, PHAB3-F/PHAX1-R, PHAB2-F/PHAX2-R, PHAB1-F/PHAX3-R primers, respectively. These PCR products were cloned just downstream of the stop codon of binary GFP vector. P-G700, P-G500 and P-G300 PCR fragments were generated with GUS-F/GUS700-R, GUS-F/GUS500-R, GUS-F/GUS300-R primers, and then these fragments were cloned into PHA-s. ST-LS1 (referred to as Ls in this manuscript) intron was PCR amplified with Ls-F and Ls-R primers from GUS-intron construct (40), it was cloned into pBluescript KS (KSLs), and then it was moved downstream of GFP-c. To create P-99Ls, Ls fragment was cloned from KSLs into PHA-s-78 binary vector, which contained PHA-s ORF + 78 nt downstream of PTC. To obtain P-28Ls, PHA-s was cloned to KSLs and then PHA-sLs fragment was moved into BIN61S. To obtain P-28 and P-99 constructs, RT–PCR...
products were generated with pFF-R and pKS KpnI primers from RNA samples isolated from P-28Ls and P-99Ls infiltrated leaves, and then these fragments were cloned into BIN61S. \(5^\prime\)-UTR fragment was PCR amplified with LN-F and LN-R primers from \(5^\prime\)-UTR region of \(\lambda\)N construct (41) and cloned into BIN61S. To obtain \(\lambda\)N-UPF1, UPF1 from KSUPF1 was cloned into \(\lambda\)N binary clone. BoxB tethering target sequence, which contains five direct repeats of boxB, was PCR amplified from 5boxB plasmid (41) with 5BBF/5BB-R primers, and then it was cloned into the \(3^\prime\)-UTR region of Gulf-c and PHA-s-78 binary constructs (GFP3’boxB and P-3’boxB). To create GFP5’boxB and P-5’boxB \(3^\prime\)-UTR tethering target plasmids, boxB was cloned into the \(5^\prime\)-UTR region of GFP and PHA-s binary vectors.

**Agroinfiltration assays and GFP imaging**

*Agrobacterium tumefaciens* infiltration (agroinfiltration) is a plant transient assay system. Infiltration of *Nicotiana tabacum* or *Nicotiana benthamiana* leaves with culture of an *Agrobacterium* containing a binary vector leads to transient expression of the gene that is cloned into the expression cassette of binary vector. To co-express different genes, agrobacterium cultures each carrying one binary vector should be mixed before infiltration. As agroinfiltration elicits RNA silencing, co-infiltration of a silencing suppressor is required for efficient agroinfiltration-mediated expression. PHA or PHA-derived constructs were agroinfiltrated into *N. tabacum* leaves, while for better GFP visualization GFP and GFP-derived constructs were expressed in *N. benthamiana* leaves. Agroinfiltration assays and GFP detections were carried out as described (42). For co-infiltration, respective *A. tumefaciens* cultures were mixed before infiltration. Unless it is indicated in the text or in the figure legends, the final OD of each culture was 0.2 (OD\(_{600} = 0.2\)) in the mixture.

**Cycloheximide treatment, RT–PCR assay and RNA gel blot analysis**

Agroinfiltrated leaves were collected at 3 d.p.i., cut into small pieces and treated for 3 h with 10 \(\mu\)g/ml cycloheximide (Sigma) as described (43). RNA extraction and RNA gel blot analysis was carried out as described (44). PCR fragments labelled with random priming method were used for northern analyses. Phosphorimage measurements were used to quantify expression of mRNAs. RT–PCR analysis was carried out with QIAGEN OneStep RT–PCR Kit to test splicing efficiency. pFF-R and mGFP4for primers were used for testing the splicing of GFP-cLs. For P-28Ls and P-99Ls samples pFF-R and pKS KpnI primers were used, which amplifies the mRNA of test constructs as well as the mRNA of PHA internal control. PCR fragments were directly sequenced (GFP-cLs) or cloned and sequenced (P-28Ls and P-99Ls).

**Western blot analysis**

Proteins were separated in an 8\% SDS–polyacrylamide gel, then transferred on to a Hybond-C Extra filter (Amersham Pharmacia Biotech). N-terminal region of UPF1 (1–123 amino acids) was expressed as a glutathione \(S\)-transferase (GST) fusion protein in *Escherichia coli* (GST-UPF1-N), then rabbit polyclonal antibodies (\(\alpha\)-UPF1) raised against GST-UPF1-N were used for western analysis. Transiently expressed *Arabidopsis* UPF1 or UPF1DN proteins were easily detected with \(\alpha\)-UPF1, while endogenous *N. tabacum* or *N. benthamiana* UPF1 could not be detected with this antibody.

**RESULTS**

**Agroinfiltration-based transient NMD test system**

Transient assays have been successfully used to study NMD in mammals and *Drosophila*. Therefore, as a first step towards characterization of plant NMD, we wanted to establish an efficient transient NMD assay system. It has been shown that the wild-type bean phytohemagglutinin (PHA) mRNA is stable, while the mutant mRNA containing a PTC (PHA-m) is efficiently targeted by NMD (30,31). In transgenic tobacco cells, the half-lives of PHA-m transcripts were reduced 3-fold relative to the PHA transcripts; consequently PHA-m mRNAs accumulated to \(\sim\)5 times lower levels than PHA transcripts (31). We have adapted this well-characterized NMD reporter gene system to establish an agroinfiltration-based transient plant NMD assay (also see Materials and Methods). PHA and PHA-m, which contains a PTC at codon 79, were cloned into binary vectors (Figure 1), and then *N. tabacum* leaves were infiltrated with *Agrobacterium* cultures carrying PHA or PHA-m binary plasmid. As an internal control, the ORF of PHA-m from ATG to the PTC (PHA-s) was also cloned into a binary vector (Figure 1), and then the *Agrobacterium*-containing PHA-s was co-infiltrated with either PHA or PHA-m. Moreover, as agroinfiltration leads to strong transgene-induced RNA silencing, P14 silencing suppressor (Figure 1) was also co-infiltrated (38). Thus, leaves were co-infiltrated with PHA + PHA-s + P14 or with PHA-m + PHA-s + P14 cultures. At 3 days post infiltration (d.p.i.) RNA gel blot assays were performed to monitor mRNA accumulation. PHA and PHA-m transcript levels were normalized to the levels of the PHA-s common internal control; therefore PHA and PHA-m mRNA levels could be compared. The finding that PHA-m mRNA levels were \(\sim\)10 times lower than PHA mRNA levels suggested that PHA-m transcripts were targeted by NMD (Figure 2A, compare lanes 3–4 to 1–2 and Supplementary Figure 2). As NMD is a translation-dependent RNA degradation system, treatment with the translation inhibitor cycloheximide stabilizes PTC-containing mRNAs. To confirm that NMD is responsible for the reduced PHA-m mRNAs levels, PHA and PHA-m infiltrated leaves were treated with cycloheximide. As expected, cycloheximide treatment resulted in strongly increased PHA-m mRNA levels (3,75\%\), but did not enhance PHA mRNA accumulation (Figure 2B). These data support the conclusion that in agroinfiltrated leaves PHA-m mRNA levels are downregulated by NMD.

In worms, the NMD and RNA silencing pathways overlap (45). Thus, because UPF1 may be involved in RNA silencing in plants (28), it was possible that the P14 silencing suppressor modified the NMD response. To exclude this possibility, we repeated the experiments in the absence of P14 but at a lower temperature (15\°C) where RNA silencing is inactive (46). We found that at 15\°C, in the absence of P14, PHA-m mRNAs accumulated to much lower levels
than PHA transcripts (Figure 2C). These data suggest that, in contrast to RNA silencing, NMD is active even at low temperature and indicate that P14 does not affect NMD. Consistently, PHA transcripts were ~10 times more abundant than PHA-m mRNAs when these genes were expressed at the normal growth temperature from a specific binary vector (NOSas) that barely triggers silencing (S. Kertesz, Z. Merai and D. Silhavy, unpublished data).

As these data suggest that P14 has no effect on NMD, further agroinfiltration-based NMD experiments were carried out at 22°C in the presence of the P14 silencing suppressor.

Identification of Arabidopsis UPF1

UPF1 is the key trans-acting factor responsible for NMD in yeast as well as in animals. Based on sequence similarities, it has been suggested that At5g47010 encodes UPF1 (UPF1) in Arabidopsis (10,27). Moreover, as many PTC-containing transcripts are upregulated in a missense mutant or in T-DNA insertion mutants of At5g47010, it is likely that Arabidopsis UPF1 is a functional ortholog of yeast UPF1 (28,29). However, direct evidence proving that Arabidopsis UPF1 is involved in NMD have not yet been reported. To test whether plant UPF1 plays a role in NMD, it was amplified by RT–PCR, cloned into a binary vector, and then the effect of overexpression of UPF1 on NMD was studied in agroinfiltration assays. As co-infiltration of UPF1 did not result in increased PHA-m transcript levels, we concluded that overexpression of Arabidopsis UPF1 did not affect NMD in tobacco leaves (data not shown).

It has also been reported that changing of a conserved arginine to cysteine in the RNA helicase domain of either yeast (R779C) or human (R844C) UPF1 leads to a dominant-negative protein (47,48). We hypothesized that if At5g47010 encodes the functional Arabidopsis UPF1, a similar change of the corresponding arginine (R863C) will also lead to a dominant-negative protein (UPF1DN, see Figure 1), and that expression of UPF1DN will interfere with NMD. To test these hypotheses, PHA-m control and PHA-m NMD reporter constructs were co-infiltrated with UPF1DN. Importantly, co-expression of UPF1DN selectively increased the levels of PTC-containing PHA-m transcripts, providing evidence that UPF1DN interfered with NMD (Figure 2A, compare lanes 7–8 to 3–4 and Supplementary Figure 2). Western-blot assays confirmed that UPF1 and UPF1DN proteins accumulated to comparable levels (data not shown); thus, the different effects of these proteins on NMD are not due to differences in abundance. These data strongly suggest that UPF1 plays a role in plant NMD, and that UPF1DN is a dominant-negative mutant. In further experiments, UPF1DN co-infiltration assays were used to identify NMD-targeted mRNAs.

Long 3'-UTR can act as NMD cis element in plant

We next wanted to identify the cis-acting elements of the plant NMD system. As PHA-m is a strong NMD target, whereas PHA or PHA-s mRNAs are not targeted by NMD, it is likely that sequences between the PTC of PHA-m and its original stop codon (referred to as abc region,
Inactive. Infiltrated plants were grown at low temperature where RNA silencing is carried out as described for (A), except that P14 was not co-infiltrated.

Levels of Cyc treated samples are shown relative to the mRNA levels of the treated control samples (N. tabacum co-infiltrated with buffer (C0)). (Figure 3A). Importantly, co-infiltration of UPF1DN (Figure 3A, compare lanes 7–8 to 3–4, and see Supplementary Figure 3) or cycloheximide treatment (Figure 3B, compare lanes 3–4 to 1–2) led to selective increase of GFP-abc mRNA levels. These data indicate that GFP-abc transcripts are targeted by NMD. Thus, we concluded that the abc region contained all the necessary information for PTC definition.

Although it has been shown that in plants PTC-containing mRNAs transcribed from intronless genes are targeted by NMD (30–33), it is not known if DSE-like short, specific destabilizing sequences or unusually long 3’-UTRs subject these mRNAs to NMD. To distinguish between these two possibilities we tried to map the DSE-like destabilizing sequences in the abc region. The abc region was divided into three sequences ~equal in size (~200 nt), and then each region was cloned downstream of the stop codon of the GFP binary construct yielding GFP-a, GFP-b and GFP-c, respectively (Figure 1). To test if any of these constructs is targeted by NMD, GFP-a, GFP-b, GFP-c test and GFP control constructs were agroinfiltrated alone or were co-infiltrated with UPF1DN. We have found that GFP-a, GFP-b and GFP-c transcripts accumulated to comparable levels and that these mRNAs were less abundant than the control GFP transcript. It is likely that GFP-a, GFP-b and GFP-c mRNAs were moderately targeted by NMD because co-infiltration of UPF1DN resulted in slightly increased transcript levels (Figure 3C and Supplementary Figure 4). These data suggest that the abc region does not contain DSE-like destabilizing sequences, thus it is likely that the extended 3’-UTRs trigger NMD in plants. Moreover, as GFP-abc was much more efficiently targeted by NMD than GFP-c (Figure 3D), GFP-a or GFP-b (data not shown), it is likely that the length of 3’-UTR defined the efficacy of NMD. Indeed, we found that GFP-bc mRNAs accumulated to higher levels than GFP-abc but were less abundant than GFP-b or GFP-c transcripts (Figure 3D and Supplementary Figure 5).

To further confirm that long 3’-UTRs and not DSE-like sequences act as NMD cis elements in plants, 700, 500 and 300 nt long bacterial sequences derived from GUS gene were cloned downstream of PHA-s (P-G700, P-G500, P-G300, see Figure 1). It is unlikely that bacterial sequences contain plant DSE-like sequences, thus the finding that bacterial sequences destabilized plant transcripts further supported the model that long 3’-UTRs triggered NMD in plants (Figure 3E). Consistently, the longest GUS segment caused the strongest reduction of PHA-s mRNA levels (Figure 3E, compare lanes 7–8 to 3–4).

**Introns located in the 3’-UTR could act as NMD cis elements in plant**

Introns are efficient NMD cis elements in mammals. To clarify if introns located in the 3’-UTR could act as NMD
cis elements in plants, we cloned the well-characterized potato ST-LS1 intron (referred to as Ls) (40) downstream of the GFP-c reporter gene (GFP-c-Ls, Figure 1). The effect of Ls intron on mRNA accumulation was studied in agroinfiltration assays by comparing the expression levels of the GFPc-Ls test and the GFP-c control constructs. The green fluorescence was much weaker in GFPc-Ls infiltrated plants than in GFP-c infiltrated samples (Figure 4A Supplementary Figure 6A). Consistently, GFPc-Ls mRNAs accumulated to low levels, while GFP-c mRNAs were abundant. RT–PCR assays confirmed that Ls was efficiently spliced from GFPc-Ls (Supplementary Figure 6C). Co-infiltration of UPF1DN with GFPc-Ls resulted in stronger fluorescence and in dramatically increased GFPc-Ls mRNA levels (11×). Therefore, we concluded that the 3′-UTR located Ls intron could trigger NMD-mediated degradation of GFPc-Ls mRNA (Figure 4A and Supplementary Figure 6B). Consistent with these data, cycloheximide treatment led to increased GFPc-Ls mRNA levels (Figure 4B).

To exclude that the effect of Ls intron is gene specific, we have also cloned the Ls intron 28 and 99 nt downstream of the GFPc-Ls test and the GFP-c control constructs. The green fluorescence was much weaker in GFPc-Ls infiltrated plants than in GFP-c infiltrated samples (Figure 4A Supplementary Figure 6A). Consistently, GFPc-Ls mRNAs accumulated to low levels, while GFP-c mRNAs were abundant. RT–PCR assays confirmed that Ls was efficiently spliced from GFPc-Ls (Supplementary Figure 6C). Co-infiltration of UPF1DN with GFPc-Ls resulted in stronger fluorescence and in dramatically increased GFPc-Ls mRNA levels (11×). Therefore, we concluded that the 3′-UTR located Ls intron could trigger NMD-mediated degradation of GFPc-Ls mRNA (Figure 4A and Supplementary Figure 6B). Consistent with these data, cycloheximide treatment led to increased GFPc-Ls mRNA levels (Figure 4B).

To exclude that the effect of Ls intron is gene specific, we have also cloned the Ls intron 28 and 99 nt downstream of the stop codon of the PHA-s binary construct yielding P-28Ls and P-99Ls, respectively (Figure 1), and then leaves were infiltrated with either of these constructs. In these experiments PHA was used as an internal control. RT–PCR confirmed that the Ls intron was efficiently spliced to both P-28Ls and P-99Ls mRNAs (Supplementary Figure 6D). We found that P-99Ls mRNAs accumulated to low levels in the infiltrated leaves, while P-28Ls mRNAs were abundant (Figure 4C). Interestingly, co-infiltration of UPF1DN resulted in increased accumulation of P-99Ls mRNAs (4.3×) but only slightly enhanced the levels of P-28Ls mRNAs (1.66×) (Figure 4C). Consistently, cycloheximide treatment resulted in increased P-99Ls mRNA levels...
Figure 4. Introns located in the 3′-UTR could act as NMD cis elements. (A) Accumulation of the GFP-c and GFP-cLs mRNAs in leaves of N. benthamiana infiltrated with GFP-c + P14 (lanes 1 and 2), GFP-cLs+P14 (lanes 3 and 4), GFP-c + P14 + UPF1DN (lanes 5 and 6) or GFP-cLs + P14 + UPF1DN (lanes 7 and 8) cultures. GFP-c and GFP-cLs mRNA levels were normalized to the corresponding P14 levels, and then mean GFP-c mRNA level was taken as 1. Photos show the GFP activity in the infiltrated leaves. (B) Effect of cycloheximide (Cyc) on the accumulation of the GFP-c and GFP-cLs transcripts. (C) The destabilizing effect of the 3′-UTR located Ls intron is position-dependent. N. tabacum leaves were infiltrated with P-28Ls + PHA + P14 or with P-99Ls + PHA + P14 + UPF1 (lanes 1–2 and 5–6, respectively). These mixtures were also co-infiltrated with UPF1DN (lanes 3–4 and 7–8). To show that the different 3′-UTR length of P-28Ls and P-99Ls transcripts is responsible for the different accumulation, identical but intronless constructs (P-28 and P-99) were infiltrated and analysed in parallel. Blot was hybridized with a PHA-s probe. Upper bands show PHA control transcript. Lower bands show P14 control transcript. Samples and photos were taken at 3 d.p.i.

Tethering of UPF1 to either the 3′- or 5′-UTR of a mRNA leads to low target transcript levels

In mammals, tethering of UPF1 to the 3′-UTR of a wild-type mRNA leads to rapid degradation of the targeted mRNA (49). In contrast, tethering of UPF1 to the 5′-UTR or to the open reading frame of an intronless mRNA enhances its translation, because UPF1 promotes mRNA polyosome association (50). Interestingly, UPF2, UPF3 and certain EJC proteins act similarly, tethering of these proteins to the 3′-UTR triggers rapid mRNA degradation, while their tethering to the ORF enhances translation of target mRNAs (50, 51). Since introns act as NMD cis elements in both mammals and plants, we have postulated that in addition to NMD, plant UPF1 also stimulates translation. To test these hypotheses we constructed a binary vector, which expresses UPF1 as a λN fusion protein (λN-UPF1, Figure 1). λN specifically binds a short RNA sequence called the boxB (52, 53). To create a 3′-UTR target construct, five copies of boxB as direct repeats were inserted downstream of GFP reporter gene (GFP5′boxB, see Figure 1), and then GFP5′boxB was infiltrated alone or was co-infiltrated with λN-UPF1. As controls, GFP3′boxB was also co-infiltrated with λN or UPF1 binary constructs. While co-infiltration of λN or UPF1 affected the target mRNA levels only slightly, GFP3′boxB transcript levels were strongly reduced in λN-UPF1 co-infiltrated samples (Figure 5A, compare lanes 7–8 to 1–2 and Supplementary Figure 7). Similarly, co-infiltration of λN-UPF1 also reduced the transcript levels of P-3′boxB (see Figure 1), another 3′-UTR tethering target construct that contained five boxB repeats downstream of PHA-s reporter gene (Figure 5C). These data indicate that tethering of plant UPF1 to the 3′-UTR of a wild-type mRNA leads to downregulation of the targeted mRNA.

We next tested the effect of UPF1 tethering to the 5′-UTR of a reporter mRNA. To create a 5′-UTR tethering target construct, five boxB sequences as direct repeats were cloned into the 5′-UTR region of GFP reporter gene (GFP5′boxB, see Figure 1). GFP5′boxB was infiltrated alone or was co-infiltrated with λN-UPF1, with λN or with UPF1, respectively. Unfortunately, we failed to detect GFP activity in the GFP5′boxB infiltrated leaves (perhaps because inserted sequences contain ATGs), therefore we could not analyse the effect of 5′-UTR tethering of UPF1 on translation efficiency. However, we could study the effect of 5′-UTR tethering of UPF1 on target mRNA accumulation. Surprisingly, co-infiltration of λN-UPF1 severely reduced the levels of GFP5′boxB transcripts, while co-infiltration of λN or UPF1 did not modify the accumulation of GFP5′boxB mRNAs (Figure 5B). Tethering of UPF1 to the 5′-UTR region of...
the PHA-s reporter construct (P-5'-boxB, see Figure 1) also led to dramatically reduced target mRNA levels (Figure 5D).

It is unlikely that tethering of UPF1 to an mRNA would lead to transcriptional silencing of the corresponding gene; therefore we suggest that tethering of UPF1 to either the 3'- or 5'-UTR causes reduced transcript accumulation because it triggers rapid degradation of targeted mRNA. As NMD is a translation-coupled decay system, the finding that 5'-UTR tethering of UPF1 also causes low transcript levels is an unexpected result. It is possible that in plants, unlike in mammals, UPF1 joins the NMD complex in a late, irreversible phase of complex formation, thus artificial tethering of UPF1 to any part of an mRNA results in rapid decay of that transcript.

DISCUSSION

**Agrobacterium-based transient plant NMD assay**

Transient assays have been efficiently used to analyse NMD in mammalian and Drosophila cells. As transient NMD assays often require the expression of several different genes, and because agroinfiltration is the best method to co-express many different genes in plants (54), we have established an agroinfiltration-based transient plant NMD test system. We have shown that PTC-containing mRNAs accumulate to low levels relative to wild-type controls and that the levels of PTC-containing mRNAs are selectively increased by cycloheximide treatment or by co-expressing UPF1DN, a dominant-negative mutant of plant UPF1 (Figures 2 and 3). These data indicate that in agroinfiltrated leaves PTC-containing mRNAs were targeted by NMD. We think that this versatile transient NMD test system, in combination with UPF1DN co-expression and λN-boxB tethering assay systems can become useful tools to characterize plant NMD.

**Cis elements of plant NMD**

Taking the advantage of the transient NMD test system, we have identified the cis elements of plant NMD. We found that in plants stop codons are identified as PTC if the 3'-UTR is unusually long or if an intron is located in the 3'-UTR. Previous reports have shown that intronless mRNAs can be degraded by NMD in plants (30–33). Consistently we have found that intronless mRNAs are targeted by NMD if a PTC is introduced into the coding region or if stuffer sequences are cloned into their 3'-UTRs (Figures 2 and 3). As we failed to identify DSE-like destabilizing sequences and because insertion of either bacterial or plant sequences into the 3'-UTR triggered NMD, we concluded that long 3'-UTR subjected intronless mRNAs to NMD in plants. Moreover, we have shown that this effect was size-dependent, mRNAs with longer 3'-UTR were more effectively targeted by NMD than transcript with shorter 3'-UTR (Figure 3D and E). Long 3'-UTR could also trigger NMD in yeast, Drosophila, worm and mammalian cells (8,17–19,55). The faux UTR model of yeast NMD suggests that long 3'-UTR (and perhaps other unusual 3'-UTRs) causes aberrant translation termination and that aberrant termination leads to the formation of a functional NMD complex (9). As PTC-containing yeast mRNAs can be protected from NMD if PABP is tethered downstream of the stop codon, it is likely that long 3'-UTR causes aberrant termination by inhibiting the interaction of terminating ribosome with PABP (8). Our findings that long 3'-UTR triggers NMD in...
plants and that this transcript destabilizing effect depends on the size of the 3'-UTR can be explained if the faux UTR model is valid for plant NMD. Therefore, we suggest that aberrant translation termination also leads to NMD in plants and that increasing the distance between terminating ribosome and PABP results in aberrant termination.

Introns could also act as NMD cis elements in plants. The finding that incorporation of Ls intron into the 3'-UTR of either the GFP or PHA-s transcripts subjects these mRNAs to NMD (Figure 4) supports this conclusion. Moreover, the effect of plant introns on mRNA stability is position-dependent. mRNAs carrying Ls intron 99 nt downstream of the stop codon were targeted by NMD, while transcripts carrying the same intron 28 nt downstream of the stop codon did not trigger NMD (Figure 4C and D). In mammals, introns located less than 50–55 nt downstream of the stop codon also fail to trigger NMD. As 3'-UTR located introns trigger NMD in a similar position-dependent manner in both plants and mammals, we suggest that in plants, like in mammals, EJC connects splicing and NMD. Indeed, the putative orthologs of most EJC components were identified in plants (56,57). However, in plants, unlike in mammals, the 3'-UTR tethering of EJC proteins does not result in the decay of targeted mRNAs (Supplementary Figure 8). Therefore further experiments are required to prove that plant EJCs are directly involved in NMD.

Our data that introns can act as NMD cis elements in plants are apparently conflicting with previous studies. It has been suggested that introns do not trigger NMD in barley, because mRNAs containing introns downstream of the PTC and the control transcripts accumulated to comparable levels (35). However, the transcript which was used as a control in that study also carried a PTC, therefore the control transcript could be also targeted by NMD. Indeed, both the intron containing and control mRNAs accumulated to 3-fold lower levels than the corresponding wild-type mRNA (35). Rose has shown that incorporation of the Arabidopsis ubiquitin intron 80 nt downstream of the stop codon of the GUS reporter gene did not result in reduced reporter mRNA levels (36). In mammals, different introns trigger NMD with different efficiency (58). It is possible that the Ls intron triggers NMD more effectively than the ubiquitin intron. Alternatively, in plants, the intron should be more distant than 80 nt downstream of the stop codon to trigger NMD.

Identification of NMD cis elements allows the recognition of Arabidopsis genes, which could be regulated by NMD. Structural targets could be transcripts with an intron in the 3'-UTR, mRNAs with long 3'-UTR, or mRNAs containing an upstream ORF (uORF) in their 5'-UTR region. In silico analysis shows, that ~30% of plant mRNAs contain uORF. We have identified ~1000 genes whose 3'-UTRs are longer than 500 nt and found that 3.6% of all Arabidopsis genes contain intron in their 3'-UTRs (list of genes are available as Additional Materials, http://www.abc.hu/RNA/). Therefore, it is likely that NMD also plays a role in regulation of many wild-type genes in plants.

Evolution of PTC definition
As the putative UPF1, 2 and 3 orthologs can be identified in each eukaryotic lineages and because the NMD system has not been found in prokaryotes, it is likely that NMD evolved in the stem eukaryotes (56,59). However, the evolution of NMD system is not well understood. As long 3'-UTR triggers NMD in yeast, Drosophila, worm and human cells (19, 20, 21), it has been suggested that this PTC definition system is ancient (19), perhaps it was already present in the stem eukaryotes. Our result that long 3'-UTRs also trigger NMD in plants supports this hypothesis. The evolution of intron-based PTC definition is more debated. Based on the findings that introns are not NMD cis elements in yeast or Drosophila, it has been suggested that the intron-based PTC definition evolved late, only when alternative splicing had become dominant in certain animal lineages (11). An alternative model of NMD evolution has proposed that the intron-based PTC definition has already operated in the stem eukaryotes. This model suggests that splicing and NMD were coupled by the EJC in these ancient organisms (56,59). As the EJC components (56) and many introns (60,61) are highly conserved in eukaryotes, this model hypothesizes that in stem eukaryotes EJC-based NMD could efficiently eliminate PTC-containing mRNAs of spliced transcripts, thereby providing primary positive selection for the intron containing alleles. Consequently, NMD facilitated the rapid spreading of ancient introns in stem eukaryotes (56,59). The finding that 3'-UTR located introns trigger NMD in a position-dependent manner in plants is consistent with this model. This model also predicts that introns should be ~evenly distributed within the coding regions (56,59). Indeed, we and others (59,62) have shown that plant introns are distributed relatively equally along the coding regions except the very 5' and 3' regions (Additional Materials).

Taken together, our findings suggest that both long 3'-UTR and intron-based PTC definition systems could operate in stem eukaryotes. In lineages, in which intron loss dominated (yeast, Drosophila etc.), splicing and NMD could be evolutionarily uncoupled, thus long 3'-UTR became the dominant NMD cis element. In contrast, in the extremely intron-dense lineages, where alternative splicing is very widely used (mammals), introns could become the more efficient and dominating NMD cis elements. Although, plants are intron-dense organisms, intronless genes are also frequently found in the plant genome, thus NMD machinery has evolved under dual constrains, it should efficiently identify PTC-containing mRNAs derived from either intronless or intron containing genes. Therefore, both long 3'-UTR and intron-based PTC recognition machinery should work efficiently in plants.

SUPPLEMENTARY DATA
Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS
The authors are grateful to N. Gehring and M. W. Hentze for kindly providing 5boxB and lRN plasmids. The authors thank D. Baulcombe for 3SGFP construct. The authors are grateful to I. Janosi for helping with database analysis and to G. Giczey for useful corrections. The authors thank Edina Kapuszta for technical assistance and G. Takacs for helping
with figure preparations. This research was supported by grants from the Hungarian Scientific Research Fund (OTKA) (T15 042787, K60102) and from the National Office for Research and Technology (4/064/2004, ‘Búzakalász’). D.S. was supported by the Bolyai János Scholarship. S.K. and Zs.M. are graduate students of Eötvös Loránd University ‘Classical and molecular genetics’ PhD program. Funding to pay the Open Access publication charges for this article was provided by the Hungarian Scientific Research Fund (OTKA) (T15 042787, K60102).

Conflict of interest statement. None declared.

REFERENCES


