



# An Alternative Nested Reading Frame May Participate in the Stress-Dependent Expression of a Plant Gene

Ekaterina V. Sheshukova<sup>1†</sup>, Tatiana V. Komarova<sup>1,2†</sup>, Natalia M. Ershova<sup>1,2</sup>, Anastasia V. Shindyapina<sup>1,2</sup> and Yuri L. Dorokhov<sup>1,2\*</sup>

<sup>1</sup> Department of Genetics and Biotechnology, N.I. Vavilov Institute of General Genetics, Russian Academy of Sciences, Moscow, Russia, <sup>2</sup> A.N. Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University, Moscow, Russia

## OPEN ACCESS

### Edited by:

Ruth Grene,  
Virginia Tech, United States

### Reviewed by:

Elina Welchen,  
National University of the Littoral,  
Argentina  
László Szabados,  
Biological Research Centre (MTA),  
Hungary

### \*Correspondence:

Yuri L. Dorokhov  
dorokhov@genebee.msu.su

<sup>†</sup>These authors have contributed  
equally to this work.

### Specialty section:

This article was submitted to  
Plant Abiotic Stress,  
a section of the journal  
Frontiers in Plant Science

**Received:** 31 August 2017

**Accepted:** 04 December 2017

**Published:** 19 December 2017

### Citation:

Sheshukova EV, Komarova TV,  
Ershova NM, Shindyapina AV and  
Dorokhov YL (2017) An Alternative  
Nested Reading Frame May  
Participate in the Stress-Dependent  
Expression of a Plant Gene.  
Front. Plant Sci. 8:2137.  
doi: 10.3389/fpls.2017.02137

Although plants as sessile organisms are affected by a variety of stressors in the field, the stress factors for the above-ground and underground parts of the plant and their gene expression profiles are not the same. Here, we investigated *NbKPILP*, a gene encoding a new member of the ubiquitous, pathogenesis-related Kunitz peptidase inhibitor (KPI)-like protein family, that we discovered in the genome of *Nicotiana benthamiana* and other representatives of the *Solanaceae* family. The *NbKPILP* gene encodes a protein that has all the structural elements characteristic of KPI but in contrast to the proven *A. thaliana* KPI (AtKPI), it does not inhibit serine peptidases. Unlike roots, *NbKPILP* mRNA and its corresponding protein were not detected in intact leaves, but abiotic and biotic stressors drastically affected *NbKPILP* mRNA accumulation. In search of the causes of suppressed *NbKPILP* mRNA accumulation in leaves, we found that the *NbKPILP* gene is “matryoshka,” containing an alternative nested reading frame (ANRF) encoding a 53-amino acid (aa) polypeptide (53aa-ANRF) which has an amphipathic helix (AH). We confirmed ANRF expression experimentally. A vector containing a GFP-encoding sequence was inserted into the *NbKPILP* gene in frame with 53aa-ANRF, resulting in a 53aa-GFP fused protein that localized in the membrane fraction of cells. Using the 5'-RACE approach, we have shown that the expression of ANRF was not explained by the existence of a cryptic promoter within the *NbKPILP* gene but was controlled by the maternal *NbKPILP* mRNA. We found that insertion of mutations destroying the 53aa-ANRF AH resulted in more than a two-fold increase of the *NbKPILP* mRNA level. The *NbKPILP* gene represents the first example of ANRF functioning as a repressor of a maternal gene in an intact plant. We proposed a model where the stress influencing the translation initiation promotes the accumulation of *NbKPILP* and its mRNA in leaves.

**Keywords:** Kunitz peptidase inhibitor, matryoshka gene, alternative nested reading frame, abiotic and biotic stress, virus, bacteria, gene, expression

## INTRODUCTION

Plants in natural field conditions are constantly exposed to a variety of abiotic and biotic environmental factors (Wang et al., 2003; Chaves and Oliveira, 2004; Nakashima and Yamaguchi-Shinozaki, 2006; Hirel et al., 2007; Bailey-Serres and Voeselek, 2008; Atkinson and Urwin, 2012; Suzuki et al., 2014; Ramegowda and Senthil-Kumar, 2015).

A set of stress factors and, accordingly, gene expression profiles for the underground and above-ground parts of plants differ (Freschet et al., 2015; Islam et al., 2015a,b; Liu et al., 2015; Chmielewska et al., 2016). Wind, rain, and herbivorous insects cause mechanical damages mainly to the above-ground part of the plants (Jaffe and Forbes, 1993; Sampathkumar et al., 2014; Savatin et al., 2014; Lup et al., 2016), which increase opportunities for penetration by bacteria and viruses (Carr et al., 2010; Ryu, 2015). Plant roots suffer primarily from soil bacteria (Alegria Terrazas et al., 2016), root-feeding herbivores (Kergunteuil et al., 2016), and salinity (Hanin et al., 2016). Cold, lack of water or lack of nutrients have a generalized effect on the plant (Suzuki et al., 2014; Ramegowda and Senthil-Kumar, 2015).

The plant, in response to the impact of a stress factor, switches on protective mechanisms, including generalized cellular mobilization and accumulation of mRNAs that direct the synthesis of protective proteins in cells (Downen et al., 2012; Zavaliev et al., 2013; Crisp et al., 2016; Yang et al., 2017; Zhang and Sonnewald, 2017). Termination of the stress-factor impact cancels the synthesis of protective protein mRNAs and is followed by their degradation. Such repeated effects create a kind of “memory” (Crisp et al., 2016), which allows plants to shorten the time for a protective response. The reaction of the plant genome to the stress is carried out in close interaction with the plastid and mitochondrial genomes via anterograde and retrograde signaling (Leister et al., 2017). A main phytohormone, abscisic acid, plays an essential role in modifying the expression levels of stress-responsive genes (Baek et al., 2017; Vishwakarma et al., 2017). Stress-dependent modification of gene expression can be regulated through transcriptional factors (Eulgem and Somssich, 2007; Jang, 2016; Baek et al., 2017; Fang et al., 2017) and post-transcriptional mechanisms including alternative splicing (Shang et al., 2017) and mRNA nucleocytoplasmic trafficking via the nuclear pores (Yang et al., 2017). The translation of mRNA is also a mechanism for responding to stress, especially the stages involving the 5′ terminal oligopyrimidine tract and translation initiation factors eIF4E and eIF2a (Sesma et al., 2017).

It is known that the same genes related to defense provide a different level of mRNA accumulation in leaves and roots after challenge with abiotic factors (De Nardi et al., 2006; Bazakos et al., 2015; Corso et al., 2015; Freschet et al., 2015; Liu et al., 2015; Pucholt et al., 2015; Chmielewska et al., 2016; Fan et al., 2016; Goralski et al., 2016; Gurung et al., 2016; Miao et al., 2016; Nguyen et al., 2016; Kobayashi et al., 2017). Similar leaf-root differences after viral (Sohal et al., 1999; Germundsson et al., 2002; Andika et al., 2016), fungal (Galiana et al., 1997; Okubara and Paulitz, 2005), bacterial (Zdor and Anderson, 1992; de A Gerhardt et al., 1997; Gruau et al., 2015), and insect

attack (Hermsmeier et al., 2001) were noted for various genes such as pathogenesis-related (PR) genes (Fu and Dong, 2013; Parmar et al., 2017) and genes encoding peptidase inhibitors (PIs) (Hörger and van der Hoorn, 2013; Jashni et al., 2015; Grosse-Holz and van der Hoorn, 2016; Misas-Villamil et al., 2016). PIs are classified into superfamilies which are subcategorized further into specific families (Rawlings et al., 2014, 2016). Among them is the Kunitz peptidase inhibitor (KPI) family belonging to the serine PI group (<https://www.ebi.ac.uk/merops>), which includes proteins of ~20 kD with one or two disulfide bonds and a single reactive site (Oliva et al., 2010, 2011; Santamaría et al., 2014; Guo et al., 2015). Functional analysis of the poplar KPI family revealed biochemical diversity and multiplicity in defense reactions (Major and Constabel, 2008). Moreover, a study of white clover has shown that some members of the KPI family are expressed only in roots and not in leaves or other parts of plants (Islam et al., 2015a,b).

Until recently, stress response studies and leaf-root differences in mRNA accumulation studies focused primarily on the mechanisms of transcriptional promoter involvement in gene regulation, activation of transcription factors, and chromatin cytosine methylation in the promoter region of certain genes to prevent the access of transcriptional activators (Komarnytsky and Borisjuk, 2003; Jones and Dangl, 2006; Ryu, 2015; Espinas et al., 2016; Neto et al., 2016; Jiang et al., 2017; Marand et al., 2017). Explanations of the differences in the expression of a gene in leaves and roots are usually based on mechanisms involving transcription factors (Kobayashi et al., 2017), hormones (Fujita et al., 2006), volatile organic compounds (Duran-Flores and Heil, 2016), and mobile RNA transcripts (Saplaoura and Kragler, 2016).

Now it has become clear that the leaf-root differences in gene expression and their reactions to stress cannot be explained by only the regulation of gene transcription. So far, no studies have been conducted on alternative nested reading frames (ANRF) in mRNA and their role in mRNA stability to explain the pattern of certain gene expression levels in leaf and root tissues. Although until recently there was an established opinion that one RNA transcript encodes a single protein, the latest data from proteogenomics pointed to the existence of an exception to this rule, which, in many respects, changes the customary meaning of the term “gene” (Kochetov et al., 2013; Liu and Qian, 2014; Mouilleron et al., 2016). The coding potential of such ANRF was not taken into account in most cases, primarily due to the lack of information on the possibility of their translation to polypeptides in a cell (Andrews and Rothnagel, 2014). Recent discoveries of the cell polypeptides encoded by ANRF created the concept of a gene system which could be described as a “matryoshka” (Ribrioux et al., 2008). Improvement of research techniques led to the detection of double-coding transcripts in mammals, such as the *TRP-1* gene transcript encoding the TYRP1 tumor antigen (Wang et al., 1996), the *caspase 1* (*CASPI* or *ICE*) transcript (Ronsin et al., 1999), the prion protein gene *PRNP* (Vanderperre et al., 2011, 2013), and the gene encoding Ataxin-1 (*ATXN1*) (Bergeron et al., 2013).

In principle, besides the ANRF in the 5′-leader sequence region designated as upstream open reading frames (uORFs),

plant mRNAs contain a plurality of ANRF in the main open reading frame (Hayden and Jorgensen, 2007; Tran et al., 2008; Vaughn et al., 2012). The role of uORF expression in plant stress response was confirmed by ribosomal profiling in normal and stress conditions (Juntawong et al., 2014; Tanaka et al., 2016; Bailey-Serres and Ma, 2017; Schepetilnikov and Ryabova, 2017; Sesma et al., 2017; Xu et al., 2017a,b). Expression of ANRF located in the coding portion of the main gene has only been demonstrated for the maize gene *Zm908* (Dong et al., 2013; Hanada et al., 2013).

Here, we identified and described the properties of a novel *Nicotiana benthamiana* gene encoding a KPI-like protein (NbKPILP). Unlike roots, *NbKPILP* mRNA and its corresponding protein were not detected in intact leaves, but prolonged darkness and bacterial or viral infection stimulated its mRNA accumulation. We found that the *NbKPILP* gene is common for *Solanaceae* plants and belongs to a matryoshka gene family containing an ANRF that encodes a 53-amino acid (aa) polypeptide (53aa-ANRF) which contains an amphipathic helix (AH). Our experimental approaches allowed identification of the 53aa-ANRF that affected the level of *NbKPILP* mRNA accumulation in intact leaves. The existence of a cryptic promoter within the *NbKPILP* gene was excluded using the 5'-RACE approach. We found that insertion of mutations destroying the 53aa-ANRF AH resulted in an increase of *NbKPILP* mRNA accumulation. The *NbKPILP* gene is the first example of an ANRF influencing maternal mRNA accumulation in leaves.

## MATERIALS AND METHODS

### Plant Growth Conditions

*Nicotiana benthamiana* plants were grown in soil in a controlled environment under a 16/8 h day/night cycle.

### Plasmid and Vectors

To create 35S-NbKPILP construct NbKPILP-encoding sequence was obtained by PCR using “NbKPILP(KpnI)d” and “NbKPILP(SalI)r” primers and *N. benthamiana* total cDNA as a template. PCR product was subsequently digested with KpnI and SalI and inserted into pCambia1300-based binary vector containing *Cauliflower mosaic virus* (CaMV) 35S promoter and 35S terminator of transcription (pCambia-35S) via KpnI/SalI sites. To create 35S-AtKPI construct AtKPI-encoding sequence was obtained by PCR using “AtKPI(SacI)d” and “AtKPI(PstI)r” primers and total *A. thaliana* cDNA as a template. PCR product was digested with SacI and PstI and inserted into pCambia-35S SacI/PstI sites.

To obtain the (SS-)NbKPILP-6xHis plasmid, SacI, and HindIII sites were introduced at the 5'- and the 3'-ends of (SS-)NbKPILP-encoding sequence, respectively, through PCR with primers “ss-NbKPILP(SacI)d” and “NbKPILP(HindIII)r.” The (SS-)NbKPILP fragment flanked with SacI and HindIII was cloned into pQE30 (QIAGEN, Holland) plasmid digested with SacI and HindIII, to generate the (SS-)NbKPILP-6xHis construct. To obtain the (SS-)AtKPI-6xHis construct (SS-)AtKPI sequence was amplified using the primer pair “AtKPI(BamHI)d” and “AtKPI(HindIII)r,” PCR product was subsequently digested with

BamHI and HindIII endonucleases. That fragment was cloned into pQE30 (QIAGEN, Holland) plasmid digested with BamHI and HindIII.

For the 35S-NbKPILP(53aa-GFP) and 35S-(SS-)NbKPILP(53aa-GFP) constructs the fragments, containing 1–565 or 70–565 nt of *NbKPILP* ORF, respectively, were amplified using the following pairs of primers: “NbKPILP(KpnI)d”/“53aa\_end(BamHI)r” or “ss-NbKPILP(SacI)d”/“53aa\_end(BamHI)r.” The PCR products were digested with KpnI/BamHI and together with the fragment containing *GFP* flanked with BamHI and PstI were inserted into digested with KpnI and PstI pCambia-35S resulting in 35S-NbKPILP(53aa-GFP) and 35S-(SS-)NbKPILP(53aa-GFP) constructs, respectively.

The 35S-NbKPILP(ACG) construct was created in several steps. First, site-directed mutagenesis using “NbKPILP(KpnI)d”/“53aa(ACG)r” and “53aa(ACG)d”/“NbKPILP(SalI)r” pairs of primers and 35S-NbKPILP plasmid as a template was performed. Second, the final PCR product, containing NbKPILP(ACG) sequence was obtained using overlap PCR approach with “NbKPILP(KpnI)d” and “NbKPILP(SalI)r” pair of primers. The product of overlap PCR was digested with KpnI and SalI and inserted into pCambia-35S via KpnI/SalI sites.

The crTMV-53aa and crTMV-53aa\_mut vectors were created in several steps. First, a PCR fragment containing 53aa-ANRF-encoding sequence (407–568 nt of NbKPILP ORF) was obtained using “53aa(NcoI)d” and “53aa(XhoI)r” pair of primers. For substituted variant (53aa\_mut) the PCR fragment was obtained using “53aa(NcoI)d”/“53aa(AHmut)r” and “53aa(AHmut)d”/“53aa(XhoI)r” pairs of primers for site directed mutagenesis and “53aa(NcoI)d”/“53aa(XhoI)r” pair of primers for overlap PCR. The resulting PCR products, encoding 53aa-ANRF or 53aa\_mut-ANRF, were digested with NcoI and XhoI and inserted into plasmid IC-2 (Komarova et al., 2011), containing 3'-region of crTMV movement protein (MP) gene, 3'-UTR of crTMV and *nos* transcription terminator. The resulting constructs were digested with EcoRI/XhoI and a fragment containing part of *MP*, 53aa-ANRF (or 53aa\_mut-ANRF), crTMV 3'-UTR and *nos* terminator was inserted into TMV-HC vector (Komarova et al., 2011) via EcoRI/XhoI sites resulting in crTMV-53aa or crTMV-53aa\_mut construct, respectively.

To obtain the plasmid encoding NbKPILP(53aa) fused to the DYKDDDDKDYKDDVDDYKDDDDK (3xFLAG) sequence (Ueda et al., 2011), a fragment containing full *NbKPILP* sequence without stop codon was amplified using “NbKPILP(KpnI)d” and “NbKPILP(XbaI)r” primers and subsequently digested with KpnI and XbaI. The 3xFLAG encoding sequence was generated annealing the primers “3xFlag(XbaI)d” and “3xFlag(SalI)r,” resulting in a fragment with overhangs corresponding to XbaI and SalI “sticky” ends. These fragments were inserted into pCambia-35S via KpnI/SalI sites to generate the 35S-NbKPILP(53aa):3xFLAG plasmid. For 35S-NbKPILP(53aa\_mut):3xFLAG construct, fragment containing NbKPILP(53aa\_mut) was obtained by site-directed mutagenesis using primer pairs “NbKPILP(KpnI)d”/“53aa(AHmut)r” and “53aa(AHmut)d”/“NbKPILP(XbaI)r” and 35S-NbKPILP as a template. Then these PCR fragments were used in overlap

PCR with “NbKPILP(KpnI)d”/“NbKPILP(XbaI)r” primers. The resulting product of overlap PCR was digested with KpnI and XbaI and inserted into pCambia-35S together with 3xFLAG-encoding fragment flanked with XbaI and Sall resulting in 35S-NbKPILP(53aa\_mut):3xFLAG.

A full list of the oligonucleotides used for cloning is presented in Table S1.

## Transcription Start Site Determination Using a Step-Out Rapid Amplification of cDNA 5'-End (5'-RACE) Approach

The 5'-RACE of *NbKPILP* cDNA was performed using the Mint RACE cDNA amplification set (Evrogen, Russia) according to manufacturer's instructions. The following gene-specific primers were used: “pr1;” “pr2” (Table S1).

## Agroinjection Experiments

The *Agrobacterium tumefaciens* strain GV3101 was transformed with individual binary constructs and grown at in LB medium supplemented with 50 mg/l rifampicin, 25 mg/l gentamycin, and 50 mg/l kanamycin. The *Agrobacterium* from an overnight culture were resuspended in 10 mM MES buffer (pH 5.5) supplemented with 10 mM MgSO<sub>4</sub> and adjusted to a final OD<sub>600</sub> of 0.1. Agroinjection was performed using nearly fully expanded *N. benthamiana* leaves attached to the intact plant. A bacterial suspension was infiltrated into the leaf tissue using a 2-ml syringe, after which the plants were grown under greenhouse conditions at a 16/8 h light/dark photoperiod unless otherwise specified.

## *N. tabacum* Inoculation with TMV

Lower leaves of *N. tabacum* cv. Samsun plants were mechanically inoculated with TMV virions (100 mg/ml) in 50 mM sodium phosphate buffer, pH 7.0, in the presence of Celite, as described previously (Dorokhov et al., 1981).

## GFP Visualization

GFP fluorescence in the inoculated leaves was monitored by illumination with a handheld UV source,  $\lambda = 366$  nm. To visualize 53aa-GFP subcellular localization, imaging was performed 72 h after agroinfiltration with 35S-*NbKPILP*(53aa-GFP) in the presence of a vector encoding silencing suppressor P19 from *Tomato bushy stunt virus* using an LSM510 confocal laser scanning microscope (Zeiss, Germany). Excitation wavelength was 488 nm and detection window was 493–550 nm.

## Western Blot Analysis

For Western blot analysis, the proteins from agroinjected leaves were divided into S17, P17, P1, and CW crude fractions according to Deom et al. (1990) with modifications. Briefly, frozen plant material was ground to a powder in liquid nitrogen followed by addition of three volumes of ice-cold-modified GB buffer (100 mM Tris, pH 8.0, 0.4 M sucrose, 10 mM KCL, 5 mM MgCl<sub>2</sub>, 10 mM  $\beta$ -mercaptoethanol). The obtained slurry was filtered through a double-layered Miracloth (Millipore/Merck). The material retained on the filter was collected and washed (30–60 min incubation followed by centrifugation at 1,000  $\times g$ ) 5–8 times with GB-buffer supplemented with 0.1% Triton X-100

(the final wash was performed without Triton X-100) to obtain the CW-enriched fraction. The filtrate was centrifuged at 1,000  $\times g$  for 10 min to obtain the P1 (pellet enriched with nuclei) fraction. Supernatant was further divided into S17 (supernatant enriched with soluble proteins) and P17 (pellet enriched with membrane non-soluble proteins) fractions after centrifugation at 17,000  $\times g$  for 30 min. Pellets from CW, P1 and P17 fractions were resuspended in one volume of 1xPBS.

Total protein was extracted from plant tissues using TriReagent (MRC) according to the manufacturer's instructions.

Aliquotes from all fractions were analyzed through SDS-polyacrylamide gel electrophoresis and blotted onto polyvinylidene difluoride membranes (GE Healthcare). For GFP or NbKPILP detection, the membranes were probed with corresponding antibodies: goat anti-GFP antibodies conjugated with horseradish peroxidase (Rockland Immunochemicals) or mouse polyclonal antibodies against recombinant NbKPILPx6His. Anti-mouse antibodies conjugated with horseradish peroxidase (Rockland Immunochemicals) were used as secondary antibodies. The bands were visualized using the chemiluminescence ECL kit (GE Healthcare).

## Q-PCR Analysis of Transcript Concentrations

Total RNA was extracted from plant tissues using TriReagent (MRC) according to the manufacturer's instructions. The RNA concentration was determined using a Nanodrop ND-1000 spectrophotometer (Isogen Life Sciences). All RNA samples had a 260:280 absorbance ratio between 1.9 and 2.1. The synthesis of the first strand, followed by real-time qPCR, was performed as described in Dorokhov et al. (2012). Briefly, 0.1 mg of random hexamers and 0.1 mg of oligo-dT primer were added to 2 mg of total RNA to obtain cDNA through reverse transcription using Superscript II reverse-transcriptase (Invitrogen), according to the manufacturer's protocol. Real-time quantitative PCR was carried out using the iCycler iQ real-time PCR detection system (Bio-Rad). Target genes were detected using sequence-specific primers (Table S2) and Eva Green master mix (Syntol) according to the manufacturer's instructions. Each sample was run in triplicate, and a non-template control was added to each run. A minimum of five biological replicates were performed.

## Statistics

Student's *t*-test was performed using Excel (Microsoft, Redmond, WA). *P* < 0.05 were considered significant.

## RESULTS

### Identification and Analysis of the *N. benthamiana* Gene Encoding KPI-Like Protein (NbKPILP)

In a study of the transcriptome of *N. benthamiana* leaves infected with *Tobacco mosaic virus* (TMV), we detected an increased accumulation of mRNA (EMBL ID FN687760) encoding KPI-like protein (NbKPILP) (NCBI accession number D4IHB9), which according to the MEROPS classification (<http://merops>.

sanger.ac.uk; Rawlings et al., 2014, 2016) belongs to the I3 family (clan IC), of which the soybean Kunitz trypsin inhibitor (*Glycine max*) is a typical member. To determine if there are any introns present within the *NbKPILP* ORF, the *N. benthamiana* genomic DNA and primers designed for the cDNA amplification were used to amplify the corresponding fragment. The amplified fragment length was 606-bp for both genomic DNA and cDNA. The genomic fragment sequence perfectly matched the cDNA-derived fragment (data not shown), which indicated the absence of an intron in the ORF-encoding part of chromosomal *NbKPILP* gene. Thus, like other plant KPIs (Jofuku and Goldberg, 1989; Ishikawa et al., 1994; Saarikoski et al., 1996; Gruden et al., 1997; Ashida et al., 2000; Tian et al., 2007; Huang et al., 2010; Bhattacharjee et al., 2014; Bunyatang et al., 2016), the *NbKPILP* gene does not contain introns, as shown by the chromosomal DNA sequence that we identified (data not shown) and a sequence search in the Sol Genomics Network (<http://solgenomics.net/>) and the *N. benthamiana* Genome and Transcriptome database (<http://bentgenome.qut.edu.au/>). The predicted amino acid sequence has high homology to KPILP of other *Solanaceae* species, including not only the earlier described *N. glutinosa* biotic cell death-associated protein (NgCDM1; Suh et al., 2003) but also KPILP sequences from tobacco, potato, tomato, and pepper (Figure 1). *NbKPILP* contains a predicted 24-aa signal sequence and has homology to members of the I3 family such as soybean KTI which possesses inhibitor activity against serine peptidase and is encoded by the *KTI3* gene (GmKTI3) (29.17% identity; Jofuku and Goldberg, 1989), *N. tabacum* KPII (Huang et al., 2010), and the *A. thaliana* KPI (EMBL ID Q8RXD5) (33.3% identity; Li et al., 2008). Figure 1 also shows alignment of *NbKPILP* with other sequences from the *N. benthamiana* genome: genes encoding miraculin-like protein 1, NbMLP1 (39.47% identity), NbMLP2 (35.75% identity), and NbMLP3, which is completely identical to *NbKPILP* (100.00% identity; Goodwin et al., 2012). Using the MEGA7 software (<http://www.megasoftware.net/>), a phylogenetic tree including KPIs and KPILPs with highest sequence identity to *NbKPILP*, was constructed (Figure S1). Inspection of the phylogenetic tree revealed that *NbKPILP* belongs to the same branch as other KPILPs from *Solanaceae* plants.

The *NbKPILP* likely undergoes glycosylation similar to *Synsepalum dulcificum* miraculin (MIRA) (Takahashi et al., 1990) which shows the presence of N-linked glycans at the Asn-71 and Asn-215 positions (the numbering includes signal sequences) (Figure 1). *NbKPILP* has putative N-glycosylation sites in Asn-60, Asn-86, and Asn-136 (Figure 1) according to a prediction by the NetNGlyc programme (<http://www.cbs.dtu.dk/services/NetNGlyc/>).

The *NbKPILP* has main sequence signatures of the KPI family (Major and Constabel, 2008; Huang et al., 2010; Guo et al., 2015) such as (i) the Kunitz motif, (ii) four cysteine residues arranged into two intra-chain disulfide bridges, and (iii) the reactive loop. However, there are significant differences in comparison with GmKPI3, AtKPI, or NbMLP1, for example (Figure 1). *NbKPILP* contains a lysine residue instead of glycine in the Kunitz motif and substitutions in the P1 reactive site that determines substrate specificity (Ser instead of Arg or Lys and Ala instead of Ile/Val;

Goodwin et al., 2012; Guo et al., 2015). The reactive loop of *NbKPILP*, according to a prediction by PeptideCutter (<http://www.expasy.org/tools/peptidecutter/>), does not contain a P1 reactive site residue that interacts with trypsin, such as Arg63-Ile64 in the Kunitz-type soybean trypsin inhibitor (De Meester et al., 1998). Thus, *NbKPILP* is unlikely to have trypsin inhibitory activity. To test this prediction, we obtained a recombinant *E. coli* strain producing *NbKPILP* containing 6xHis tag and lacking the signal sequence [(SS-)*NbKPILP*-6xHis]. (SS-)AtKPI-6xHis was used as a positive control with experimentally proven KPI activity *in vitro* (Li et al., 2008). Two substrates were used to evaluate the inhibitory activity of proteins: (i) p-nitroaniline, a synthetic substrate of trypsin which is characterized by high specificity of trypsin/substrate interaction and (ii) azocasein, a natural substrate of trypsin. (SS-)AtKPI-6xHis demonstrated a 90% decrease in protease activity with the synthetic substrate, and a 55% reduction with the natural substrate whereas (SS-) *NbKPILP*-6xHis did not show any trypsin inhibition on either substrate *in vitro* (data not shown).

### ***NbKPILP* mRNA Accumulation in *N. benthamiana* Leaves Is Increased after Incubation in Darkness and Foreign Protein Overproduction**

Biologically, the expression of KPI and KPILP in an intact plant varies widely in roots and leaves (Huang et al., 2010; Goodwin et al., 2012). We examined adult plants of *N. benthamiana* (Figure 2A) and did not identify *NbKPILP* in leaves using Western blot analysis (Figure 2B). Conversely, the protein analysis of the roots detected a set of protein bands, among which there was a major double band corresponding to proteins of 30 and 29 kDa, which significantly exceeds the predicted weight of a mature *NbKPILP* (19 kDa). Subcellular root fractionation showed the presence of *NbKPILP* mainly in the membrane fraction (data not shown), which is consistent with the presence of a signal sequence in the protein (Figure 1). It can be assumed that the significant differences in the content of *NbKPILP* in leaves and roots may occur due to differences in the level of accumulation of *NbKPILP* mRNA. We explored the *NbKPILP* mRNA levels in roots and mature leaves of healthy intact plants and showed that *NbKPILP* mRNA content in leaves was negligible, while in the roots it was almost three orders higher (Figure 2). The *N. benthamiana* Gene Expression Atlas programme (<http://sefapps02.qut.edu.au/atlas/tREX6.php>) confirmed our observation and also showed an increased expression of the *NbKPILP* gene in flowers (Figure S2A).

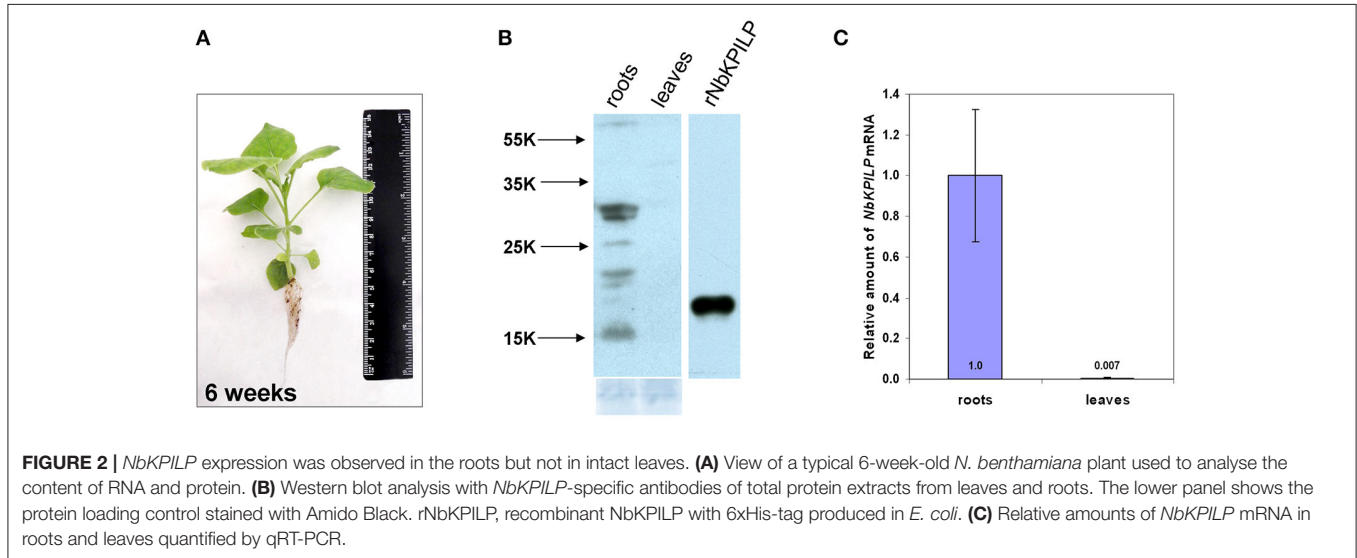
We analyzed the *NbKPILP* mRNA accumulation in the leaves and roots of *N. benthamiana* plants of different ages, starting from 2-week-old seedlings to 20-week-old flowering plants (Figure S2B). We observed that the *NbKPILP* mRNA content in leaves was low and independent of plant age, while the *NbKPILP* mRNA content in roots increased with plant age and was highest at the flowering stage (Figure S2C).

Thus, the intact *N. benthamiana* roots and leaves are radically different in *NbKPILP* content due to differences in the level of *NbKPILP* mRNA accumulation.

		* * * * *	
GmKPI3	-----MKSTIFFLF-----LFCFAF-TTSYLP SAIADFLVDNEG NPLENGGTYI LSD		46
CaKPIILP	-----M-----KISGILLLS SLL-SFTL FQITKSEPVVD TNNEQVM TGSYYI LPA		45
SlKPIILP	-----MKI LLLS CFL-YFTL FQITKSEPVVD TNNEQVRPGYTYI LPA		42
NbKPIILP	-----MK-----IISRILL SCLL-FLT L FQ-VKAE PVLD TNKQEIRPGYTYI LPA		45
NbMPLP3	-----MK-----IISRILL SCLL-FLT L FQ-VKAE PVLD TNKQEIRPGYTYI LPA		45
NtKPIILP	-----MK-----IISRILL SCLL-FLT L FQ-VKAE PVLD TNKQEV RPYTYI LPA		45
NgKPIILP	-----MK-----IISRILL SCLL-FLT L FQ-VKTE PVLD TNKQEIRPGYTYI LPA		45
AtKPI	MTKTTKTMNPK-FYLVL-----ALTAVLASNAYGAVVDIDGNAMF-HESYVYLPV		48
NbMPLP2	-----MVPF--LLVALSTSSFFLVKAQDVPEPVLDVSGNPFVRTGVNYF LPA		45
NtKPI1	-----MNTLLLLLSLVIPIALCVFNPSRFLAGSSPSPVLDINGDKVKVGLN YFVLPV		53
NbMPLP1	-----MKTQQLF L P F--LIFTXSFNSFLSSAAEAPPVVDIAGK KLR TGIDYI L P V		50
MIRA	--MKELTMLSLSFFFVS---ALLAAAANPLLSAADSAPNPVLDIDGKELRTGTNYIYIVPV		55
	#		
GmKPI3	IT-AF-GGIRAAPT---GNERCPL-TVVQSRNELDKGIGTI I SSP--YRIRFIAEGHPLS		98
CaKPIILP	TT-GNGGGLTLAQGA---NGS C P L-NVYQAQNSQSGVLP LRLLMVNSSSGLVIDENEAIN		100
SlKPIILP	ASAGSGGGLTLAKGE---NGS C P L-DVFQAQNSQSLGIP LKFLMVNSSAGLVIDENEDIN		98
NbKPIILP	TT-ANGGGLTLAKGE---NGS C P L-DVFQAQNVQSTGLP LKFLMVNSSAGLVIDENEDIN		100
NbMPLP3	TT-ANGGGLTLAKGE---NGS C P L-DVFQAQNVQSTGLP LKFLMVNSSAGLVIDENEDIN		100
NtKPIILP	TTAANGGGLTLAKGE---NGS C P L-DVFQAQNVQSTGLP LKFLMVNSSAGLVIDENEDIN		101
NgKPIILP	TTGANGGGLTLAKGE---NGS C P L-DVFQAQNSQSGVLP LKFLMVNSSAGLVIDENEDIN		101
AtKPI	IR-GRGGGLTLAGRG---GQPCPY-DIVQESS EVD EGI PVKFSNWRLK VAF-VPE SQNLN		102
NbMPLP2	GR-GNGGGLQVASIRN---RTNPL-VVSHADESSIGGYLQFSPVNPENI-IRISTDLN		99
NtKPI1	IR-GRGGGLLPSNVKQ---NNTCPR-DI I QNSDEVQEG L P V V F A P F N T K K G V - V R L S I D L N		108
NbMPLP1	VR-GRGGGLTLDSTG---NES C P L D A V V Q E Q K E I N N G F P L T F T P V N P K K G V - I R E S T D L N		105
MIRA	LR-DHGGGLTVSATT P N G T F V C P P - R V V Q T R K E V D H D R P L A F F P E N P K E D V - V R V S T D L N		112
	#		
GmKPI3	LKFDSFAVIMLCVGIPT EWSVVED-----LPEGPAVKIGENK DAMD G W F R L E R V S D D E		151
CaKPIILP	IKFAAPKYESI CK-KSTVWKI E E G-----LVTGGIKGGLVSGT SCLFTIQKYED--		150
SlKPIILP	IKFAAQRYVSI CN-VSTVWKI E D G-----IVTTGGIKGGS E N G T S T S L F T I Q K Y E D --		148
NbKPIILP	IKFAAPRYVSI CN-KSTVWKI E D G-----FVSTGGIKGGS E N G T A T S L F T I Q K Y E D --		150
NbMPLP3	IKFAAPRYVSI CN-KSTVWKI E D G-----FVSTGGIKGGS E N G T A T S L F T I Q K Y E D --		150
NtKPIILP	IKFAAPRYVSI CN-KSTVWKI -D G-----FVTGGIKGGS E N G T A T S L F T I Q K Y E D --		150
NgKPIILP	IKFAAPRYVSI CN-KSTVWKI E D G-----FVTTGGIKGGS E N G T A T S L F T I Q K Y E D --		151
AtKPI	IETDVGAT I C I ---QSTYWRVGEF D H E R K Q Y F V V A G P K P E G F G Q D S L K S F F K I E K S G E --		157
NbMPLP2	VKFT-SIHISD---SSTVWRINTE-IIPQRYLVTVGGVEGNPGRETLSNWF SIDRYED--		152
NtKPI1	VRFV-TPTICA---RETIWKLGT Y D D K L K Q Y F I V T G G V E G N P G P Q T L S S W F K I E K L G T --		162
NbMPLP1	IKFS-AASICV---QTTLWLKDDFDETTGKYFITIGNENGNPGRETI SNWFKIEKFER--		159
MIRA	INFS-AFMPCRWT-SSTVWRLLDKYDESTGQYFVTIGGVKGNPGPETISSWFKIEEFCGS--		169
	^		
GmKPI3	FNNYKLVFCPQQAED-----DKCGDIGISIDHDDGTRRLVVSKNKPLVVQFQKLDKESL		205
CaKPIILP	--AYALQYWRPKGRTFVCPKLS CGYIGIS PVA-NGSRRLAVN-RTACKIVFKKA----		201
SlKPIILP	--AYALQYCPRATGCSFICPRL LCGYIGI L T A E - N G S R H L A V N - R P V F K I V F R K A ----		199
NbKPIILP	--VYALQYCPRATGCSFICPRL LCGYIGI S P A A - N G S R R L A V N - R P V F K I V F K K V ----		201
NbMPLP3	--VYALQYCPRATGCSFICPRL LCGYIGI S P A A - N G S R R L A V N - R P V F K I V F K K V ----		201
NtKPIILP	--AYALQYCPRATGCSFICPRL LCGYIGI A P A A - N G S R R L A V N - R P V F K I V F K K A ----		201
NgKPIILP	--AYALQYCPRAAGCSFICPRL LCGYIGI A P A A - N G S R R L A V N - R P I F K I E P K K A S D T E V		207
AtKPI	--DAYKFVFCPRT CDS---GNPKCSDVGI F I D E - L G V R R L A L S - D K P F L V M F K K A N V T E V		210
NbMPLP2	--AYKLVYCPGV CET---CRPFCGDIGI L V E G - S K R V L F L R F - D K P L K V T F H K I ----		199
NtKPI1	--DYKFVFCPSVCKI---CKVICDKVGIYTK--DGVRFLALS-DTPLRVMFKKT F ----		209
NbMPLP1	--DYKLVYCPV CNF---CKVICDKIGI F I Q - D G T R R L A L S - D V P F K V M F K K A ----		205
MIRA	--GFYKLVFCPTVCGS---CKVKCGDVG I Y I D Q - K G R R R L A L S - D K P F A F E F N K T Y V F --		220
GmKPI3	AKKNHGLSRSE-----	216	
CaKPIILP	-----	201	
SlKPIILP	-----	199	
NbKPIILP	-----	201	
NbMPLP3	-----	201	
NtKPIILP	-----	201	
NgKPIILP	NFSQFDIYNRCGFSNTYMNKSY	229	
AtKPI	SSKTM-----	215	
NbMPLP2	-----	199	
NtKPI1	-----	209	
NbMPLP1	-----	205	
MIRA	-----	220	

**FIGURE 1 |** Amino acid sequence alignment of NbKPIILP with predicted amino acid sequences of other homologous proteins. The alignment was performed using the CLUSTAL programme (<http://www.ebi.ac.uk/Tools/msa/clustalo>). KPI-like proteins, CaKPIILP (*Capsicum annuum* KPIILP, sequence ID Ntab-BX\_AWOK-SS1956), *N. benthamiana* miraculin-like proteins (NbMPLP1, 2, and 3) (Goodwin et al., 2012) and NtKPIILP (*N. tabacum* KPIILP, sequence ID Ntab-BX\_AWOK-SS1956) were retrieved from the Sol Genomics Network (<http://solgenomics.net/>). SlKPIILP, *Solanum lycopersicum* KPI-like protein (accession number K4BJT7) and NgKPIILP, *N. glutinosa* biotic cell death-associated protein (accession number Q850R9) were retrieved from the NCBI database. GmKPI3, *Glycine max* KPI encoded by the *KT13* gene (NCBI accession number P01070) (De Meester et al., 1998), *N. tabacum* KPI1 (NCBI accession number B8Y888) (Huang et al., 2010), AtKPI, *Arabidopsis* (Continued)

**FIGURE 1** | *thaliana* KPI (NCBI accession number Q8RXD5) (Li et al., 2008), *Richadella dulcifica/Synsepalum dulcificum* miraculin (MIRA) (NCBI accession number P13087). Signal sequences are underlined. The known structural features of GmKPI are annotated as follows: asterisks (\*) above the alignment denote the Kunitz motif, # and ^ signs denote pairs of Cys residues forming disulfide bonds, the reactive loop sequence is highlighted in gray, the P1 reactive site residues responsible for trypsin inhibition activity according to De Meester et al. (1998), Li et al. (2008) and Huang et al. (2010) and the PeptideCutter ([http://web.expasy.org/peptide\\_cutter/](http://web.expasy.org/peptide_cutter/)) prediction are bolded and underlined. The NetNGlyc 1.0 server (<http://www.cbs.dtu.dk/services/NetNGlyc/>) predicted N-glycosylated sites are marked with bolded red text.



We assumed that like other KPIs (Huang et al., 2010) and KPILP (Suh et al., 2003; Goodwin et al., 2012; Islam et al., 2015a), *NbKPILP* mRNA accumulation is stimulated by the action of abiotic and biotic factors. Therefore, we investigated the effect of abiotic factors such as prolonged darkness on the accumulation of *NbKPILP* mRNA in leaves. For this, we incubated the plants in darkness for different periods and analyzed *NbKPILP* mRNA content in leaves. **Figure 3** shows that after 48 h of darkness *NbKPILP* mRNA content of the leaves significantly increased and continued to increase up to 72 h, reaching almost a 50-fold increase after a 96-h darkness period. However, when the plants were transferred back to a normal light/darkness photoperiod, *NbKPILP* mRNA content returned to the original level. Therefore, prolonged exposure of the plants to darkness induces *NbKPILP* mRNA accumulation in leaves.

We expected that biotic factors such as TMV infection, which is associated with impaired functioning of the chloroplast and the appearance of mosaic symptoms, should lead to the stimulation of *NbKPILP* mRNA accumulation. To test this hypothesis, we used tobacco (*Nicotiana tabacum* L. cv. Samsun) systemically infected with TMV, whose upper leaves showed mosaic symptoms, i.e., alternation of dark green spots (islands, as described by Atkinson and Matthews, 1970) characterized with a low content of TMV and light, yellow-green leaf areas enriched with TMV particles (Atkinson and Matthews, 1970; Moore et al., 2001; **Figure 4A**). Our analysis of both regions of mosaic leaves revealed a huge amount of TMV coat protein accumulated in yellow-green islands (data not shown) and a significant increase in *NbKPILP* mRNA content in yellow-green areas compared with

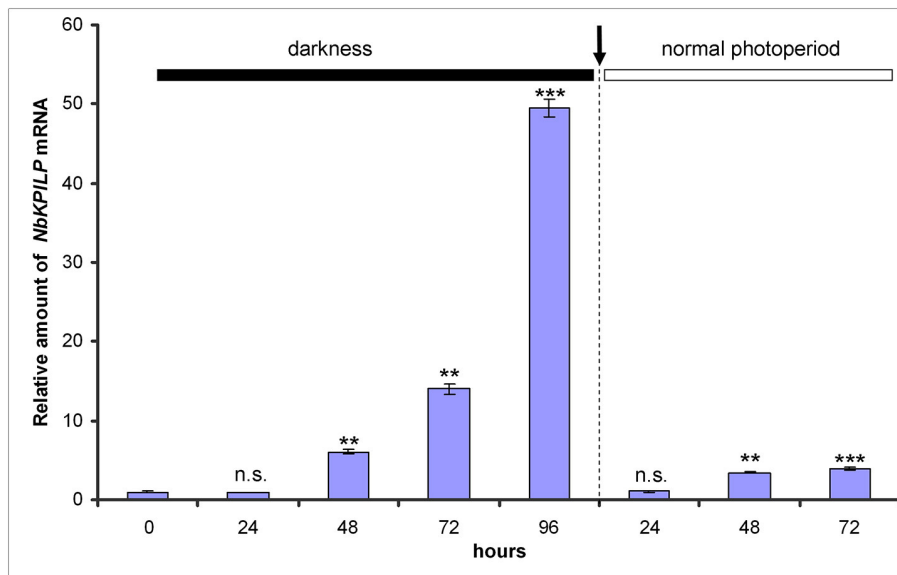
dark green islands and especially compared with leaves of the uninfected plants (**Figure 4B**).

Synthesis of a foreign protein in *N. benthamiana* leaves may also lead to the stimulation of *NbKPILP* mRNA accumulation. We chose GFP as a foreign protein, and used two types of crTMV-based vectors (**Figure 5A**) with different productive capacities for GFP synthesis in plants. The control vector, crTMV-GFP, provided moderate production of GFP (**Figure 5B**) and its modified variant, crTMV(i)-GFP, containing multiple introns in the crTMV cDNA sequence (Komarova et al., 2012), mediated excessive GFP production in leaves 4 days after agroinjection (**Figure 5B**). In parallel with the increase in accumulation of GFP mRNA (**Figure 5C**), *NbKPILP* mRNA content also increased (**Figure 5D**).

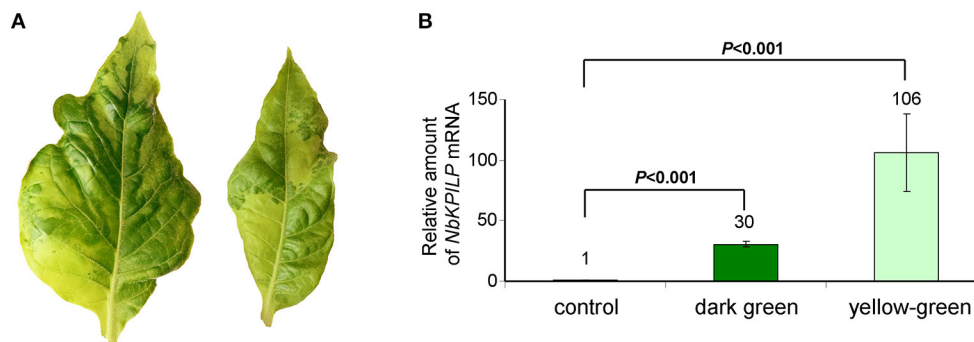
We concluded that factors of abiotic and biotic stress stimulate *NbKPILP* mRNA accumulation in the leaves.

### ***NbKPILP* Is a Matryoshka Gene Containing a Translatable, Alternative Nested Reading Frame Encoding a 53-aa Polypeptide (53aa-ANRF)**

Biologically, the amount of *NbKPILP* in leaves can be determined by both the transcriptional activity of the gene and the stability of its mRNA in the cytoplasm. Recently, the role of upstream open reading frames (uORFs) in 5'-UTRs in mRNA translation and stability was shown (Tanaka et al., 2016; Bailey-Serres and Ma, 2017; Xu et al., 2017b). To identify putative uORFs, we identified the transcription start site and the sequence of the



**FIGURE 3** | Darkness stimulates the accumulation of *NbKPILP* mRNA in the leaves. Plants were kept in darkness for 24, 48, 72, or 96 h, and the leaf samples were taken for analysis of *NbKPILP* mRNA by qRT-PCR. The data shown represent five independent experiments. The difference in *NbKPILP* mRNA content between control plants and plants after different periods of incubation in darkness (24, 48, 72, 96, and 24 h after return to a normal photoperiod) is shown. \*\*\* $P < 0.001$ , n.s. = not significant (Student's *t*-test).



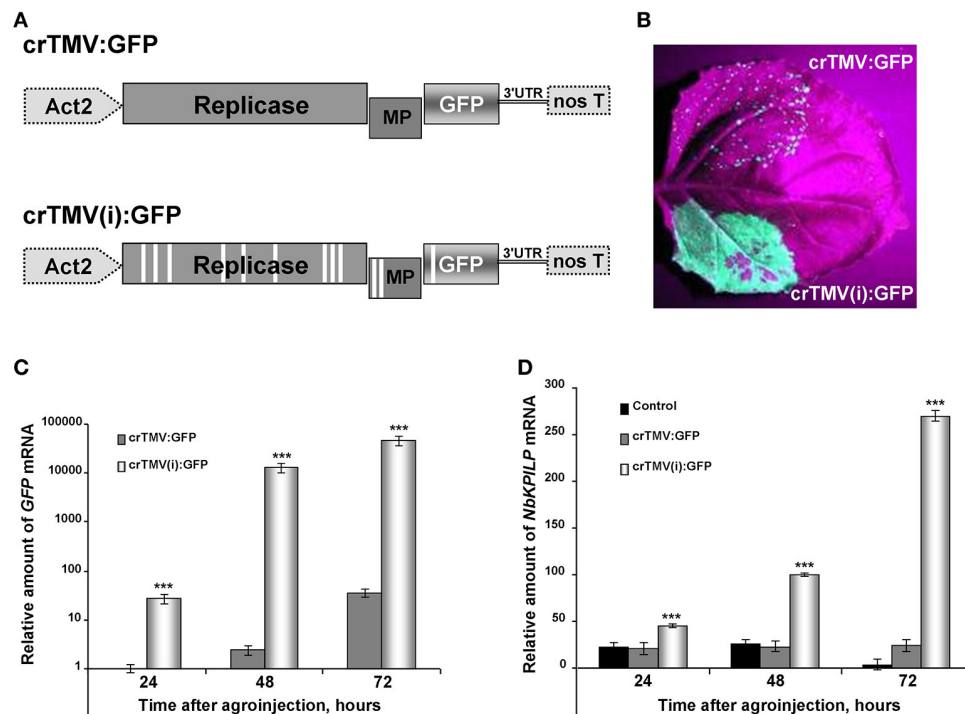
**FIGURE 4** | The yellow-green areas of TMV-infected leaves with mosaic symptoms are characterized by a high content of *NbKPILP* mRNA. **(A)** Mosaic symptoms of *N. tabacum* cv. Samsun leaves systemically infected with TMV. **(B)** Relative amounts of *NbKPILP* mRNA in dark-green and yellow-green areas of leaves with mosaic symptoms quantified by qRT-PCR. The data represents five independent experiments, and the standard error bars are indicated. The statistical significance of the difference in *NbKPILP* mRNA content between mosaic areas and control non-infected leaves is shown (Student's *t*-test).

*NbKPILP* mRNA 5'-UTR using the 5'-Rapid Amplification of Complementary Ends (5'-RACE) approach. As *NbKPILP* mRNA content in the intact leaf is very low (Figure 2), we used two approaches to increase it: (a) incubation of the plants in darkness for 4 days according to the scheme shown in Figure 3 and (b) agroinjection of *N. benthamiana* leaves with a plasmid directing *NbKPILP* mRNA synthesis.

To obtain cDNA suitable for the identification of the 5'-end of *NbKPILP* mRNA, we used the template-switching principle to add an adapter sequence to the 5'-end of the cDNA (Matz et al., 1999). Then, we performed PCR with sequence-specific reverse primers, pr1 or pr2, complementary to the *NbKPILP* region of mRNA, and a forward primer corresponding to the adapter

sequence (Figure 6A). We obtained only one PCR product for each reverse primer (pr1 or pr2; Figure 6B). Sequencing of these products allowed us to identify the transcription start site and 5'-UTR (Figure 6C). The mRNA transcription start site is located 23 nts upstream of *NbKPILP* ORF and contains a uORF with the sequence ATGCATTAA (Figure 6C). Although it is known that such start-stop codons in the 5'-UTR are able to influence the stability of mRNA (Tanaka et al., 2016), we are sure that this is not the case, since the uORF is located directly at the 5'-end. In fact, in agroinjection experiments, comparison of constructs with or without this uORF did not reveal differences in the accumulation of *NbKPILP* mRNA in the leaves (data not shown).





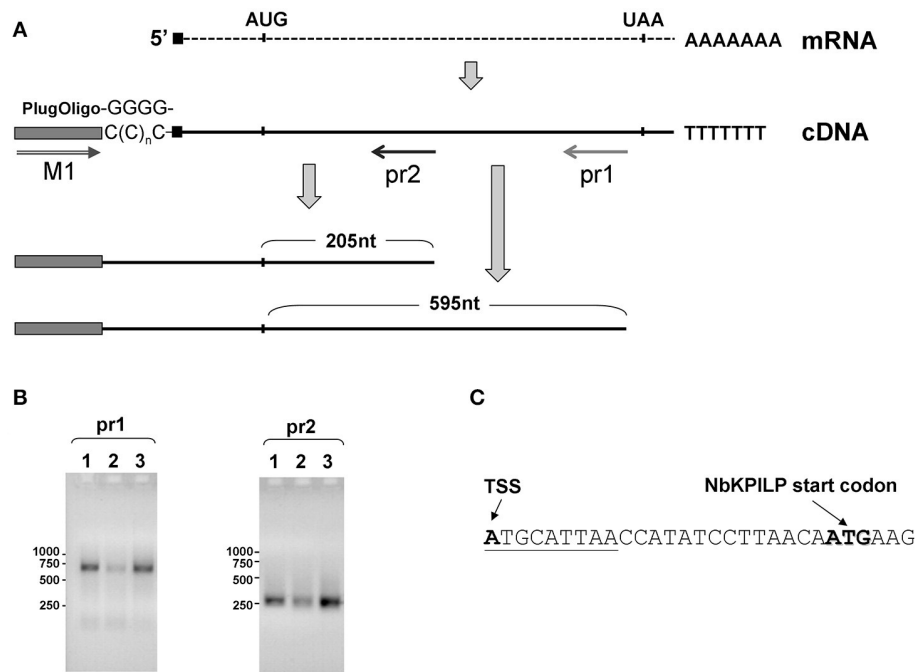
**FIGURE 5** | Overproduction of GFP as a foreign protein in the leaves stimulates the accumulation of *NbKPILP* mRNA. **(A)** Schematic representation of the crTMV-based vector crTMV:GFP and its intron-containing variant crTMV(i):GFP. Act2, *Arabidopsis thaliana act2* gene promoter; genes encoding replicase, movement protein (MP) and GFP, the 3'-untranslated region are shown; nos T stands for terminator of transcription. Introns are denoted by light rectangles. **(B)** UV visualization of GFP in leaves infiltrated with crTMV:GFP (upper part of the leaf) and crTMV(i):GFP (lower part of the leaf) on the 4th day after infiltration. **(C)** Relative amounts of GFP mRNA in *N. benthamiana* leaves agroinjected with crTMV-based vectors quantified by qRT-PCR. GFP mRNA accumulation mediated by crTMV:GFP vector was set as 1. The statistical significance of the difference between GFP mRNA levels mediated by two crTMV-based vectors is shown. \*\*\* $P < 0.001$  (Student's *t*-test). **(D)** Relative amounts of *NbKPILP* mRNA in *N. benthamiana* leaves agroinjected with crTMV-based vectors quantified by qRT-PCR. *NbKPILP* mRNA level of the intact leaf was set as 1. Control, leaf agroinfiltrated with empty vector pBin19. The statistical significance of the difference between *NbKPILP* mRNA levels mediated by two crTMV-based vectors is shown. \*\*\* $P < 0.001$  (Student's *t*-test).

However, looking closer at the mechanism of *NbKPILP* mRNA translation, we see that the first AUG codon (**Figure 7A**) is in an unfavorable context and the ribosomal 43S Pre-Initiation Complex (PIC) is likely to skip this codon via the leaky scanning mechanism (Kozak, 1986). For dicots, aaA(A/C)aAUGGcu is the optimal context for the start AUG codon (Joshi et al., 1997; Gupta et al., 2016). The next start AUG codon located 429 nts downstream from the 5' end of the *NbKPILP* mRNA is in a favorable context (agAaaAUGGga; **Figure 7A**). The second start AUG codon opens an alternative nested reading frame (ANRF) encoding the 53-aa polypeptide (53aa-ANRF) whose presence can also be predicted for other *KPILP* of *Solanaceae* plants (**Figures S3, S4**) but which is absent from true KPIs, such as *AtKPI* in *A. thaliana* (**Table 1, Table S4**).

According to a prediction by the Phyre2 programme, the 53aa-ANRF contains a transmembrane amphipathic helix (AH) (**Table 1, Figure S5**) indicating that the hypothetical 53-aa polypeptide is targeted to the cell membrane. To identify whether the 53-aa polypeptide is synthesized, we used an experimental approach in which we fused the 53aa-ANRF encoding sequence with the *GFP* gene, following the general design and context of

the *NbKPILP* gene. This technique would allow us to detect 53aa-GFP using a Western blot assay in different cell fractions. We created two vectors encoding the N-terminal part of *NbKPILP* up to the end of the 53aa-ANRF fused to GFP with a signal sequence [35S-*NbKPILP*(53aa-GFP)] (**Figure 7B**) and without a signal sequence [35S-(SS-)*NbKPILP*(53aa-GFP)] (**Figure 7C**). Then, *N. benthamiana* leaves were agroinjected with these constructs and GFP synthesis was examined by fluorescence microscopy which revealed a small amount of fluorescent single cells in leaves (**Figure S6**). To make sure that the cells synthesized GFP fused with 53-aa polypeptide, we performed a Western blot analysis of subcellular fractions of leaves agroinjected with the abovementioned binary vectors. **Figures 7D,E** show that in accordance with the predicted amphipathic properties, GFP fused with the 53-aa polypeptide was detected mainly in the cell wall fraction. Its synthesis can be seen in the leaves agroinjected with constructs encoding *NbKPILP*(53aa-GFP) with (**Figure 7D**) and without a signal sequence (**Figure 7E**).

The results of 5'-RACE confirmed that no additional shorter variants of *NbKPILP* cDNA detected in the cell (**Figure 6B**),



**FIGURE 6 |** Transcription start site and *NbKPILP* mRNA 5'-UTR identification. **(A)** Schematic representation of the 5'-RACE procedure: cDNA preparation using addition of non-templated dCMP residues to the 3'-end of full-length cDNA by reverse transcriptase (Schmidt and Mueller, 1999) followed by the cDNA second strand synthesis using "PlugOligo" primer and further amplification using "M1" and sequence-specific primers. The position of the primers is marked with arrows. AUG and UAA correspond to start and stop codons of the *NbKPILP* ORF. Sequence complementary to PlugOligo primer designated with gray box; double line with an arrowhead stands for the forward M1 primer; pr1 and pr2, positions of the reverse primers for PCR. **(B)** Agarose gel electrophoresis of PCR products obtained using the adapter primer and pr1 or pr2. Lanes correspond to the following samples: 1—leaves from the plant incubated in darkness for 96 h, 2—leaves from the plant incubated at a normal photoperiod for 24 h after 96-h darkness, 3—leaves 3 days after agroinjection with 35S-*NbKPILP*. **(C)** The nucleotide sequence of the *NbKPILP* mRNA 5'-UTR. The putative uORF is underlined. TSS, transcription start site.

therefore the expression of ANRF could not be explained by the existence of a cryptic promoter within the *NbKPILP* gene but is likely to be controlled by the maternal *NbKPILP* mRNA.

We concluded that *NbKPILP* mRNA is capable of directing the synthesis of the 53-aa polypeptide *in vivo*.

### The Role of 53aa-ANRF in the Accumulation of *NbKPILP* mRNA in *N. benthamiana* Leaves

To assess whether ANRF expression affects the level of mRNA accumulation in the leaves, we used two approaches. First, we assumed that the *AtKPI* gene, which does not contain ANRF according to predictions (Table 1), is probably expressed more actively in *N. benthamiana* leaves than *NbKPILP*. We compared the level of mRNA accumulation in agroinjected leaves expressing *NbKPILP* or *AtKPI* from 35S-based constructs (Figure 8A). Figure 8B shows that the level of *AtKPI* mRNA in agroinjected *N. benthamiana* leaves is almost two orders higher than *NbKPILP* mRNA.

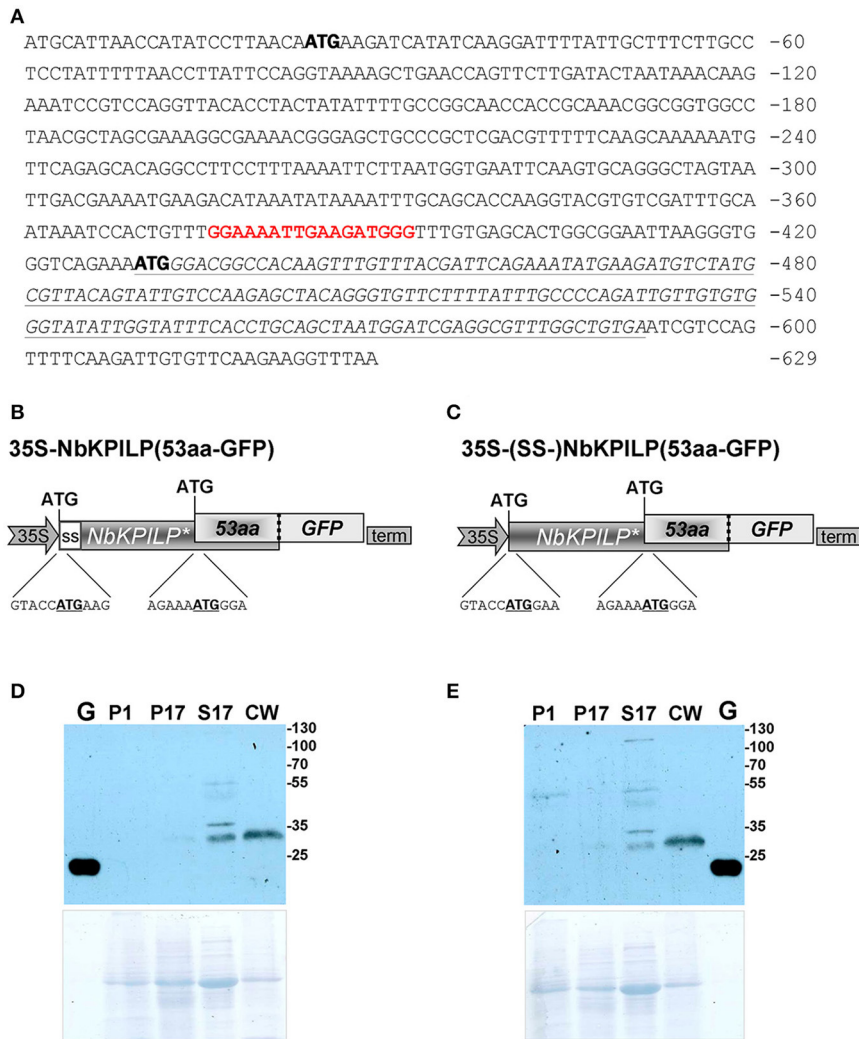
Then, we hypothesized that the optimal context of the 53aa-ANRF start codon makes it more preferable for scanning PIC, resulting in suppression of recognition and translational initiation from the *NbKPILP* AUG codon. To test this hypothesis, we inserted a nucleotide substitution into the ANRF AUG

start codon, resulting in an ACG codon (Figure 9A) and thus lowering the probability of 53aa-ANRF translation. Figure 9B shows that leaf agroinjection with 35S-*NbKPILP*(ACG) resulted in a 10-fold increase in mRNA content in comparison to 35S-*NbKPILP*(ATG). Moreover, the results of a Western blot analysis of leaf proteins were consistent with the mRNA analysis (Figure 9). Antibodies to *NbKPILP*-6xHis have not revealed *NbKPILP*-specific proteins in intact leaves. Only after leaf agroinjection with 35S-*NbKPILP* we detected a weak signal of the 30-kDa protein. For the mutant 35S-*NbKPILP*(ACG), however, we observed very intense protein bands, including three major and two minor protein bands, the presence of which is likely due to the manifestation of various predicted *NbKPILP* glycoforms (Figure 1).

We concluded that the accumulation of *NbKPILP* and its mRNA depends on the expression of 53aa-ANRF.

### Effect of the 53aa-ANRF Amphipathic Helix (AH) on *NbKPILP* Expression

It can be assumed that the inhibitory effect of ANRF on the accumulation of mRNA in the intact leaf is related to the amphipathic nature of the 53-aa polypeptide (Figure S5). It is known that the synthesis of a protein containing AH leads to endoplasmic reticulum membrane modification and



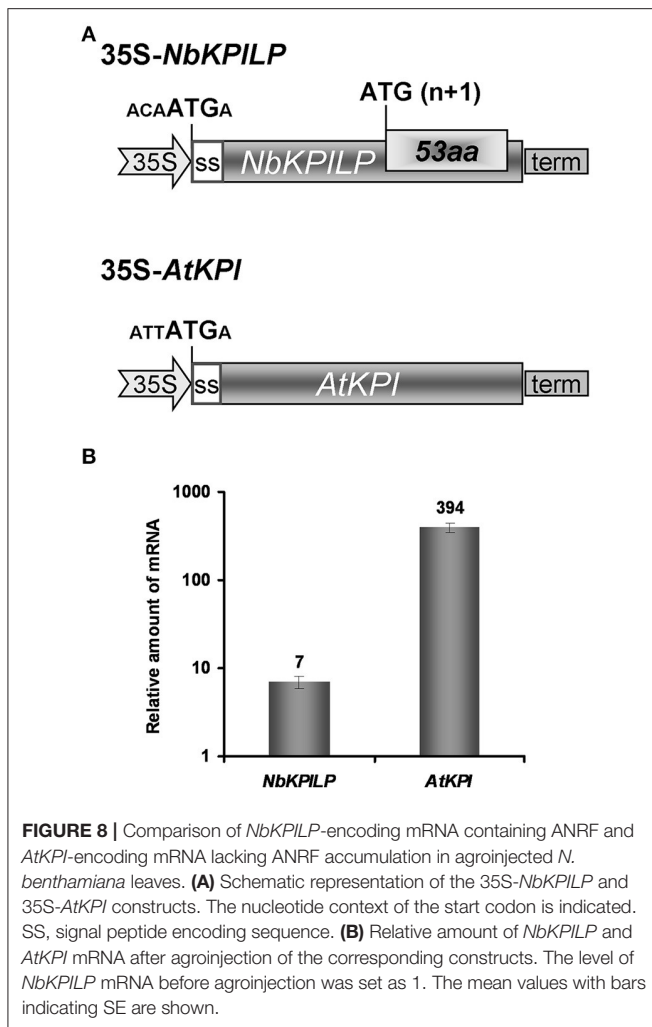
**FIGURE 7** | *NbKPILP* mRNA contains an ANRF translated into short 53-aa polypeptide (53aa-ANRF) in *N. benthamiana* leaves. **(A)** The sequence of the *NbKPILP* cDNA with the designated potential start codons (bold). Polypurine block sequence is highlighted in bolded red; 53aa-ANRF is italicized and underlined. **(B,C)** Schematic representation of the 35S-*NbKPILP*(53aa-GFP) **(B)** and 35S-(SS-)-*NbKPILP*(53aa-GFP) **(C)** vectors comprising the following elements: *CaMV* 35S promoter, *NbKPILP* with or without 24-aa signal sequence (SS-) and 53aa-ANRF fused with GFP. **(D,E)** Western blot analysis using anti-GFP antibodies of *N. benthamiana* leaf proteins at 3 days after agroinjection with the 35S-*NbKPILP*(53aa-GFP) **(D)** or 35S-(SS-)-*NbKPILP*(53aa-GFP) **(E)** vectors. Cell homogenate was fractionated into P1, P17, S17, and CW fractions. G, the total soluble protein from leaves 3 days after agroinjection with 35S-GFP vector. The lower panels show the protein loading control stained with Amido Black. All agroinfiltrations were performed in the presence of a vector encoding silencing suppressor P19 from *Tomato bushy stunt virus*.

**TABLE 1** | *KPILP* matryoshka genes with embedded ANRFs.

#	Plant and protein name	Leaf expression	EMBL ID	Identities*, %	Length of the corresponding protein (aa)	ANRF-encoded peptide contains AH**
1	<i>Nicotiana benthamiana</i> KPILP (NbKPILP)	No	FN687760	100	201	Yes
2	<i>Nicotiana glutinosa</i> KPILP (NgKPILP)	No	AF208022	94.5	229	Yes
3	<i>Solanum tuberosum</i> KPILP (StKPILP)	NA	XP_006353918	85.2	198	Yes
4	<i>Solanum lycopersicum</i> KPILP (SIKPILP)	NA	XM_004235444	84.8	199	Yes
5	<i>Trifolium repens</i> KPI4 (TrKPI4)	No	KF022201.1	28.0	210	Yes
6	<i>Hevea brasiliensis</i> KPI (HbKPI)	NA	EM_PL:KM979450	36.5	196	Yes
7	<i>Arabidopsis thaliana</i> KPI (AtKPI)	NA	At1g73260	33.3	205	No

\*Compared to *NbKPILP*.

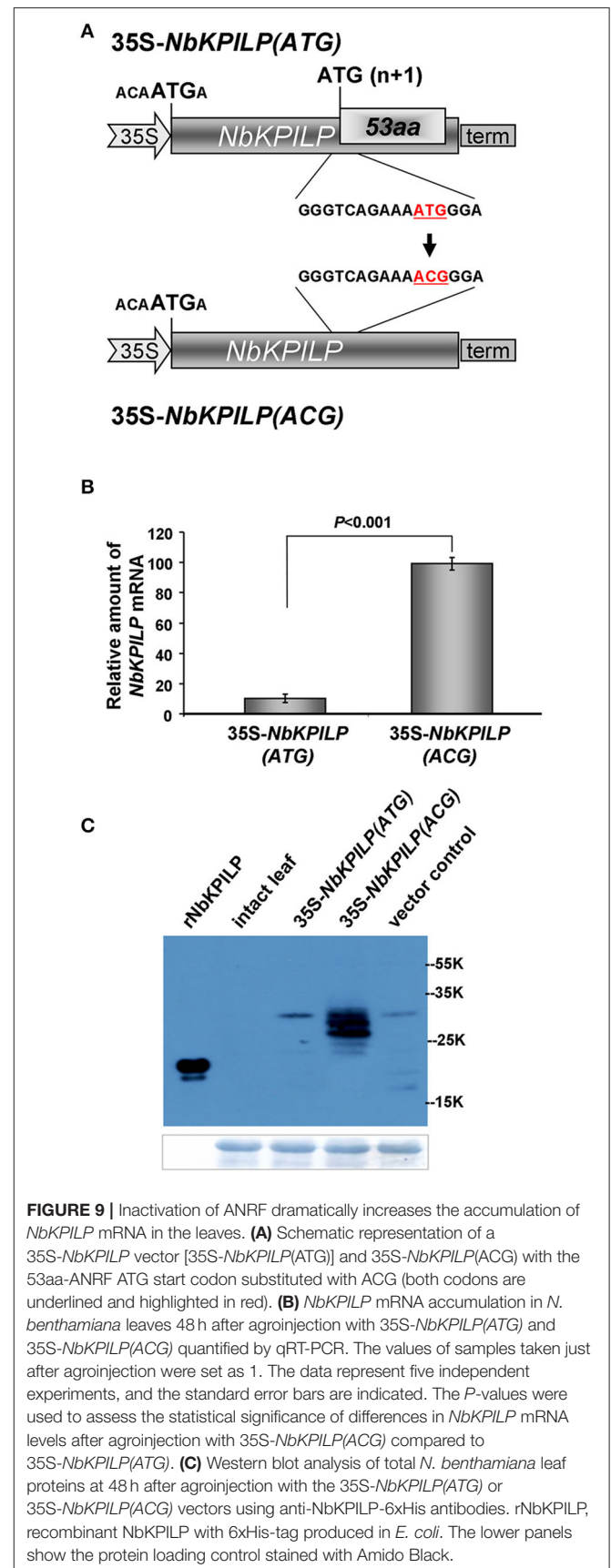
\*\*Predicted by <http://www.cbs.dtu.dk/services/TMHMM-2.0/>.

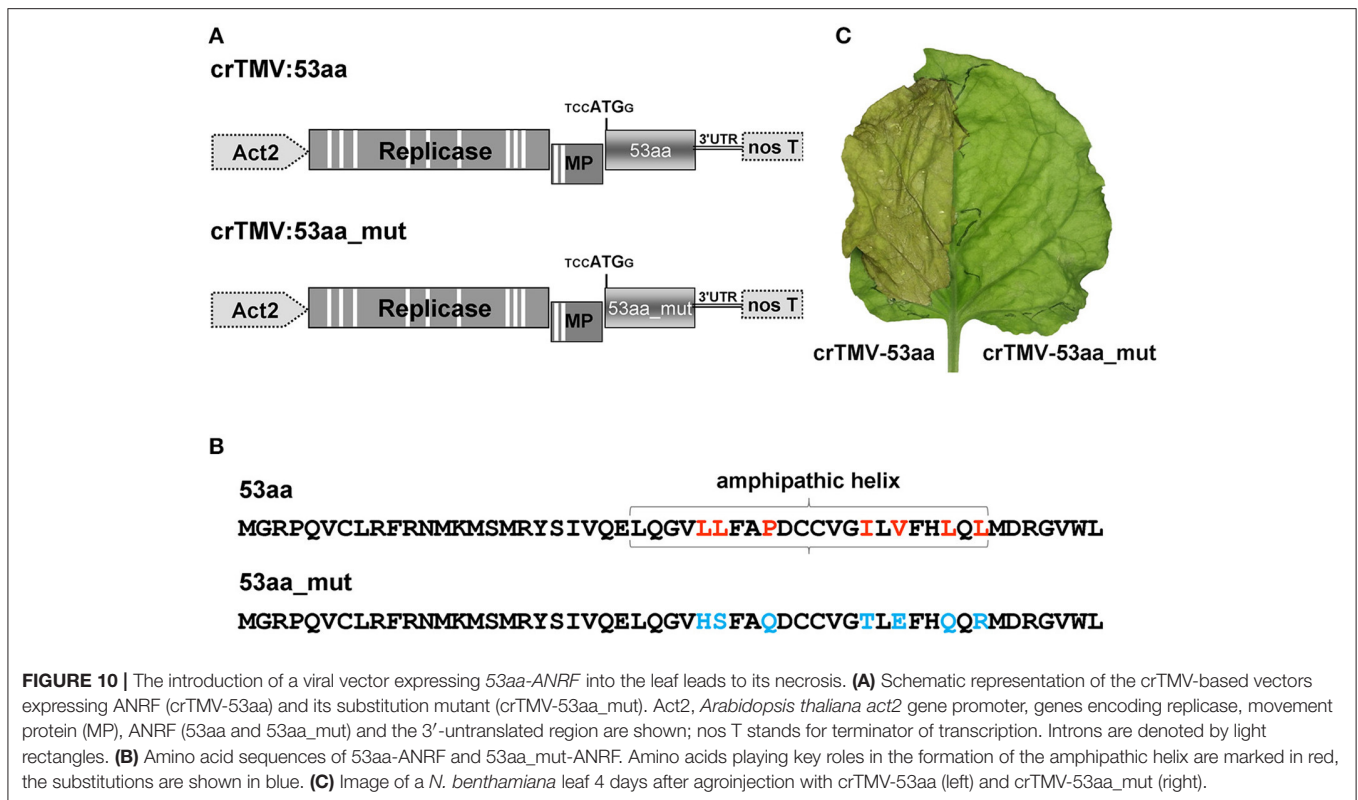


to cell death (Hashimoto et al., 2015). To test the potential ability of 53aa-ANRF to cause cellular death, we created two viral vectors based on the crTMV genome (**Figure 10A**). In the first vector, crTMV-53aa, the coat protein gene was replaced by a sequence encoding 53aa-ANRF. The second vector, crTMV-53aa\_mut, encoded a mutant of 53aa-ANRF (53aa\_mut-ANRF) containing amino acid substitutions (**Figure 10B**) that led to the destruction of the AH (**Figure S7**). **Figure 10C** shows that the agroinjection of a half-leaf by the crTMV-53aa vector led to necrosis. However, crTMV-53aa\_mut, encoding the 53-aa polypeptide without AH, did not induce cell death.

We concluded that 53aa-ANRF is potentially capable of influencing the membrane structure and potentially playing a role in cell death.

In the next step, we investigated the effect of AH-containing 53aa-ANRF membrane protein synthesis on maternal *NbKPILP* mRNA accumulation in the leaves. To perform agroinjection experiments, we created a 35S-based vector 35S-*NbKPILP*(53aa):3xFLAG (**Figure 11A**) which has the natural organization of *NbKPILP* mRNA and allows one to distinguish





between the synthesis of endogenous *NbKPILP* mRNA and synthesis of mRNA directed by the 35S-based expression vector. Of note, the injection of *Agrobacterium* leads to the accumulation of endogenous *NbKPILP* mRNA, and, hence, it is necessary to distinguish the experimental mRNA from the endogenous *NbKPILP* mRNA. For this purpose, we added a 3xFLAG-tag encoding sequence (Ueda et al., 2011) to the 3'-terminal end of *NbKPILP* to obtain an *NbKPILP*:3xFLAG fusion protein.

To understand the effect of AH on the accumulation of *NbKPILP* mRNA, we created a 35S-*NbKPILP*(53aa\_mut):3xFLAG, that encoded a 53aa\_mut-ANRF with amino acid substitutions that cause destruction of AH, but the nucleotide sequence was mutated in a way that preserved the natural amino acid sequence of *NbKPILP*. To study the relative amount of *NbKPILP*:3xFLAG mRNA expression, we performed qRT-PCR using FLAG-specific primers (Table S2). Figure 11B shows that removal of the AH from the 53aa-ANRF resulted in more than a two-fold increase in the *NbKPILP*:3xFLAG mRNA level.

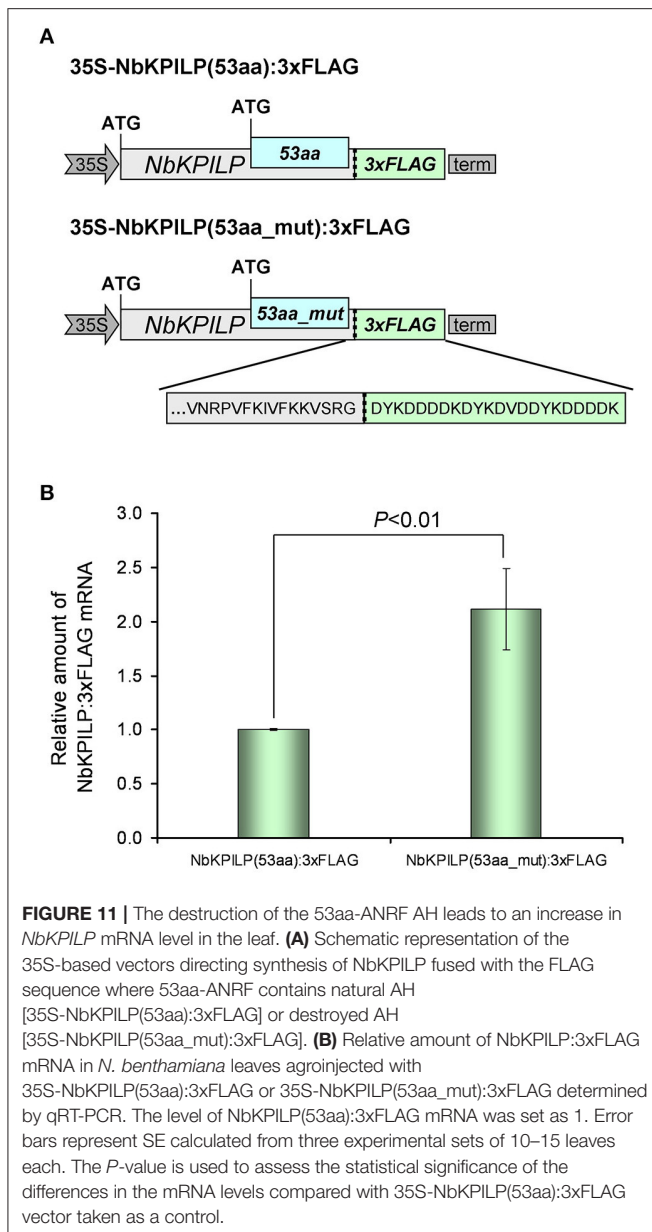
We concluded that the amphipathic nature of the 53-aa polypeptide is responsible for inhibition of *NbKPILP* mRNA accumulation in leaves of intact plants.

## DISCUSSION

Plant KPIs are normally represented by multigene families ranging from several members to several tens of members, as

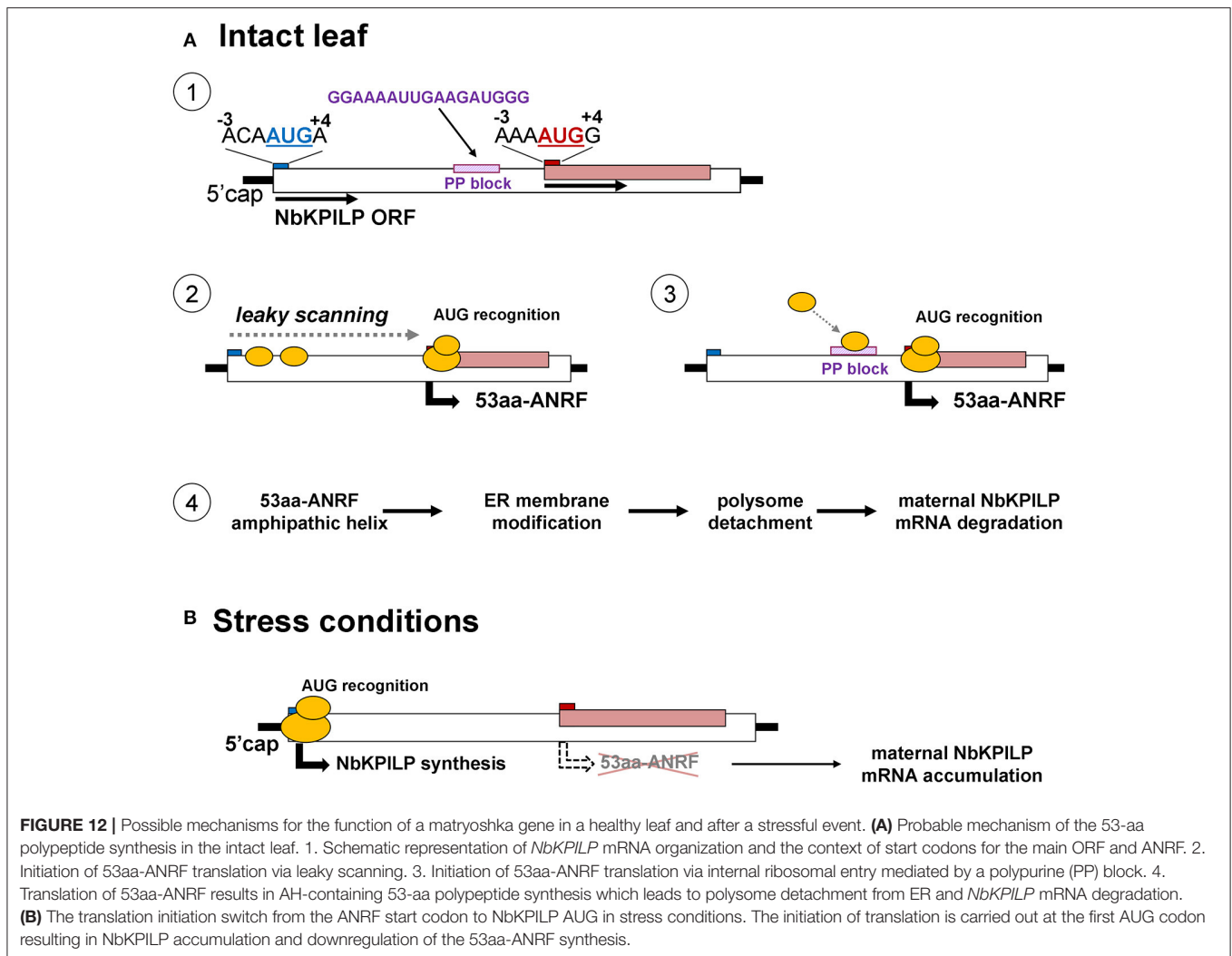
in the case of the poplar (*Populus trichocarpa*), where a wide range of within-family identity was found from low (38%) to high homology (96%) (Major and Constabel, 2008; Philippe et al., 2009). Some members of this family show their activity in organs only under stressful conditions (Rustgi et al., 2017). The upstream regions of the *NbMLP1*, *NbMLP2*, and *NbMLP3* genes were identified, and promoter elements responsible for defense responses were studied to explain the activation of these genes during stress (Goodwin et al., 2012). Here, our studies of the *N. benthamiana* genome revealed a matryoshka gene organization and another way of regulating gene activity in which ANRF functions as a repressor of a maternal gene in the plant. The *NbKPILP* gene is not unique and its expression pattern indicates no detectable mRNA accumulation in leaves of intact plants, which is also characteristic of other KPIs such as *NtKPI1* (Huang et al., 2010), *NgKPILP* (Suh et al., 2003), *TrKPI4* (Islam et al., 2015a), and *NbMLP1*, *NbMLP2*, and *NbMLP3* (Goodwin et al., 2012). Our analysis revealed that the *NbKPILP* gene belongs to a six-member group, which includes *NgKPILP*, *StKPILP*, *SlKPILP*, *TrKPI4*, and *HbKPI* (Table 1, Table S4). All these genes are characterized by (i) the presence of an ANRF whose start codon is in a more favorable context than the start codon of the main ORF, (ii) the presence of an 18- to 25-nucleotide polypurine block ~30 nts upstream of ANRF, and (iii) the 30- to 53-amino acid long membrane protein encoded by the ANRF.

In principle, the matryoshka gene design can manifest itself on two levels. In the first level, the chromosome level,



the “embedded” gene is located in introns and exons of a larger chromosomal gene. The variety of mechanisms of gene expression is determined not only by the number of protein coding genes but also by the existence of alternative transcription initiation sites and alternative splicing (Mercer and Mattick, 2013; Mercer et al., 2015). Interestingly, unlike in plants, animal genes that encode KPI contain introns (He et al., 2003; Chakrabarty et al., 2006; Isaeva et al., 2012; Chen et al., 2013) and with the help of alternative splicing, can control KPI expression (Sandbrink et al., 1994; Itoh et al., 1999; Kirchhofer et al., 2003; Wood et al., 2014). In plants, there is no alternative splicing for KPIs, so the function of ANRF during mRNA translation can be a mechanism involving stress-regulated gene expression. To explain the low content of *NbKPILP* and its

mRNA in intact leaves, we propose mechanisms that allow the synthesis of the 53-aa polypeptide which modifies ER-membranes in cells of intact leaves (Figure 12). In general, eukaryotic translational initiation is a multistep process that includes PIC assembly, its attachment to the mRNA, mRNA circularization, scanning, start codon selection, and 60S subunit joining (Haimov et al., 2015; Sesma et al., 2017). The first probable mechanism of the 53-aa polypeptide synthesis is based on ribosomal leaky scanning (Figure 12A). Usually, translation is initiated at the first 5′-proximal AUG codon in an acceptable context. The extent of “leakiness” depends on the nucleotide context surrounding AUG. The first AUG codon of *NbKPILP* mRNA is not in the optimal context for dicots (AANAUGGC) (Figure 12A-1; Kim et al., 2014; Gupta et al., 2016), therefore, the formation of the initiating complex is more preferable at the second AUG initiating codon, which is in the favorable context (Figure 12A-2). PIC scanning for over 400 nucleotides is likely to create the risk of ribosome stalling and mRNA degradation (Kawaguchi and Bailey-Serres, 2005; Hsu et al., 2016). Ribosome shunting, which occurs when the ribosome bypasses or shunts parts of the 5′-UTR on its way to the AUG start codon, is unlikely here (Miras et al., 2017; Terenin et al., 2017) since there is a PP block 39 nts upstream of the ANRF AUG codon which can perform the role of an internal ribosome entry site (Dorokhov et al., 2002; Figure 12A-3). In either case, both mechanisms could lead to the synthesis of the 53-aa polypeptide, whose AH is potentially capable of ER-membrane modification, i.e., polyribosome detachment leading to mRNA degradation (Hashimoto et al., 2015; Figure 12A-4). Under stress conditions, the synthesis of the 53-aa polypeptide and the degradation of *NbKPILP* mRNA is downregulated, since the initiation of translation is carried out at the first AUG codon, even in the unfavorable context (Figure 12B). As our experiments showed, stress factors affected the functioning of chloroplasts. It can be assumed that the change in the mode of translation of *NbKPILP* mRNA under stress is associated with the synthesis or destruction of a certain chloroplast factor, as shown in the study of the uORF participation in the repression of the transcription factor AtHB1 (Ribone et al., 2017). The question arises whether the matryoshka gene organization of the *NbKPILP* gene and the stress-regulating mechanism of its mRNA accumulation in leaves are widespread. Although the current assembly of *N. benthamiana* genome is incomplete, we found ~300,000 sORF in 41,354 transcripts of *N. benthamiana* according to the fifth version of transcriptome (Nakasugi et al., 2014), using the following criteria of a matryoshka gene organization: (a) mRNA contains nested short ORF (translated products between 10 and 100 amino acids), (b) sORF is translatable according to the sORFfinder prediction (Hanada et al., 2010), and (c) AUG codon of the sORF is in favorable Kozak’s context (consensus aMNauggS). It turned out that 1,409 ANRFs fall on 1,332 unique transcripts, in which 95% contain single ANRF, 4.4%–2 ANRFs and 0.6%–3 or 4 ANRFs. However, the proportion of ANRF among transcripts with different and identical levels of expression in roots and leaves was practically the same. Thus, the ANRF mechanism is likely to be used mainly for specific genes and tasks.



## AUTHOR CONTRIBUTIONS

YD and TK: conceived of the hypotheses and designed the experiments; ES: performed most of the experiments; TK, AS, and NE: performed some experiments; YD, ES, TK, and AS: evaluated the data and drafted the outline of the manuscript; YD, ES, and TK revised and finalized the manuscript; all the authors read and approved the manuscript.

## ACKNOWLEDGMENTS

We gratefully thank Dr. Y. E. Dunaevsky for help in the trypsin inhibitory activity measurements. The work of ES, TK, NE, and YD on the study of *NbKPILP* expression was performed at the N.I. Vavilov Institute of General Genetics of RAS and funded by the Russian Science Foundation (project No. 16-14-00002). The work of ES on obtaining the bacterial producer of *NbKPILP* and studying the trypsin inhibition activity of recombinant *NbKPILP* *in vitro* was supported by the Russian Foundation for Basic Research (project No. 16-34-00062). The

bioinformatic analysis and study of the *NbKPILP* expression pattern were performed within the framework of a project supported by the President of Russian Federation grant (MD-5697.2016.4).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2017.02137/full#supplementary-material>

**Figure S1 |** Molecular phylogenetic analysis of KPIs and KPILPs with highest sequence identity to *NbKPILP* (see **Table S3**) calculated from the amino acid sequences by the maximum likelihood method and based on the JTT matrix-based model (Jones et al., 1992). The evolutionary analysis was conducted in MEGA7.

**Figure S2 |** The *NbKPILP* mRNA accumulation in *N. benthamiana*. **(A)** The *NbKPILP* mRNA accumulation in *N. benthamiana* organs according to the *N. benthamiana* Gene Expression Atlas programme (<http://sefapps02.qut.edu.au/atlas/tREX6.php>). **(B)** Pictures of different ages of *N. benthamiana* plants isolated from soil. **(C)** The *NbKPILP* mRNA accumulation in *N. benthamiana* leaves and roots. The plants of different ages ranging from 2-week-old seedlings to

20-week-old flowering plants shown in (B), were used to analyse *NbKPILP* mRNA by qRT-PCR. The statistical significance of the difference in *NbKPILP* mRNA content between plants of different ages (5, 8, and 20 weeks) and seedlings (2 weeks) is shown. \*\* $P < 0.01$ ; \*\*\* $P < 0.001$  (Student's *t*-test).

**Figure S3** | Nucleotide sequence alignment of *N. benthamiana* Kunitz peptidase inhibitor-like (*NbKPIL*) mRNA (EMBL ID FN687760) with other homologous *Solanaceae* mRNAs: SmKPIL, *Solanum melongena* KPIL (Sol genomics sequence ID Sme2.5\_02047.1\_g00009.1) NtKPIL (*Nicotiana tabacum* KPIL, Sol genomics sequence ID Ntab-BX\_AWOK-SS1956), CaKPIL (*Capsicum annuum* KPIL, sequence ID Ntab-BX\_AWOK-SS1956), StKPIL, *Solanum tuberosum* KPIL (EMBL ID XP\_006353918), SIKPIL, *Solanum lycopersicum* KPIL (EMBL ID XM\_004235444) and NgKPIL, *Nicotiana glutinosa* biotic cell death-associated protein (EMBL ID AF208022). The polypurine block and ANRF are marked by red and yellow, respectively. Asterisks mark the nucleotides which are the same for all sequences. Hyphens represent sequence gaps. The alignment was executed by the Clustal Omega programme (<http://www.ebi.ac.uk/Tools/msa/clustalo>).

**Figure S4** | Amino acid sequence alignment of polypeptides encoded by ANRFs.

**Figure S5** | 53aa-ANRF contains an amphipathic helix as predicted by the Phyre2 programme (<http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index>).

**Figure S6** | 53aa-GFP localizes predominantly in the CW. GFP imaging by confocal microscopy of *N. benthamiana* leaf at 3 days after agroinjection with the 35S-*NbKPILP(53aa-GFP)* in the presence of a vector encoding silencing suppressor P19 from *Tomato bushy stunt virus*. Images are projections of several confocal sections. Upper panel represents a confocal image superimposed on a bright field image of the same cell.

**Figure S7** | Amino acid substitutions of 53aa-ANRF (A) resulted in AH destruction (B) as predicted by the Phyre2 programme.

**Table S1** | Oligonucleotides used for cloning and 5'-RACE.

**Table S2** | Oligonucleotides used for qPCR.

**Table S3** | Kunitz protease inhibitors and Kunitz protease inhibitor-like proteins.

**Table S4** | *KPILP* matryoshka genes with embedded ANRFs.

## REFERENCES

- Alegria Terrazas, R., Giles, C., Paterson, E., Robertson-Albertyn, S., Cesco, S., Mimmo, T., et al. (2016). Plant-microbiota interactions as a driver of the mineral turnover in the rhizosphere. *Adv. Appl. Microbiol.* 95, 1–67. doi: 10.1016/bs.aambs.2016.03.001
- Andika, I. B., Kondo, H., and Sun, L. (2016). Interplays between soil-borne plant viruses and rna silencing-mediated antiviral defense in roots. *Front. Microbiol.* 7:1458. doi: 10.3389/fmicb.2016.01458
- Andrews, S. J., and Rothnagel, J. A. (2014). Emerging evidence for functional peptides encoded by short open reading frames. *Nat. Rev. Genet.* 15, 193–204. doi: 10.1038/nrg3520
- Ashida, Y., Matsushima, A., Tsuru, Y., Hirota, T., and Hirata, T. (2000). Isolation and sequencing of a cDNA clone encoding a 20-kDa protein with trypsin inhibitory activity. *Biosci. Biotechnol. Biochem.* 64, 1305–1309. doi: 10.1271/bbb.64.1305
- Atkinson, N. J., and Urwin, P. E. (2012). The interaction of plant biotic and abiotic stresses: from genes to the field. *J. Exp. Bot.* 63, 3523–3543. doi: 10.1093/jxb/ers100
- Atkinson, P. H., and Matthews, R. E. (1970). On the origin of dark green tissue in tobacco leaves infected with tobacco mosaic virus. *Virology* 40, 344–356. doi: 10.1016/0042-6822(70)90411-3
- Baek, W., Lim, C. W., and Lee, S. C. (2017). Functional analysis of the pepper protein phosphatase, CaAIPP1, and its interacting partner CaAIRF1: modulation of ABA signaling and the drought stress response. *Plant Cell Environ.* 40, 2359–2368. doi: 10.1111/pce.13039
- Bailey-Serres, J., and Ma, W. (2017). Plant biology: an immunity boost combats crop disease. *Nature* 545, 420–421. doi: 10.1038/nature22497
- Bailey-Serres, J., and Voisenek, L. A. (2008). Flooding stress: acclimations and genetic diversity. *Annu. Rev. Plant Biol.* 59, 313–339. doi: 10.1146/annurev.arplant.59.032607.092752
- Bazakos, C., Manioudaki, M. E., Sarpoglou, E., Spano, T., and Kalaitzis, P. (2015). 454 Pyrosequencing of Olive (*Olea europaea* L.) transcriptome in response to salinity. *PLoS ONE* 10:e0143000. doi: 10.1371/journal.pone.0143000
- Bergeron, D., Lapointe, C., Bissonnette, C., Tremblay, G., Motard, J., and Roucou, X. (2013). An out-of-frame overlapping reading frame in the ataxin-1 coding sequence encodes a novel ataxin-1 interacting protein. *J. Biol. Chem.* 288, 21824–21835. doi: 10.1074/jbc.M113.472654
- Bhattacharjee, N., Banerjee, S., and Dutta, S. K. (2014). Cloning, expression and mutational studies of a trypsin inhibitor that retains activity even after cyanogen bromide digestion. *Protein Expr. Purif.* 96, 26–31. doi: 10.1016/j.pep.2014.01.013
- Bunyatang, O., Chirapongsatongkul, N., Bangrak, P., Henry, R., and Churngchow, N. (2016). Molecular cloning and characterization of a novel bi-functional  $\alpha$ -amylase/subtilisin inhibitor from *Hevea brasiliensis*. *Plant Physiol. Biochem.* 101, 76–87. doi: 10.1016/j.plaphy.2016.01.014
- Carr, J. P., Lewsey, M. G., and Palukaitis, P. (2010). Signaling in induced resistance. *Adv. Virus Res.* 76, 57–121. doi: 10.1016/S0065-3527(10)76003-6
- Chakrabarty, A., Green, J. A., and Roberts, R. M. (2006). Origin and evolution of the TKDP gene family. *Gene* 373, 35–43. doi: 10.1016/j.gene.2005.12.033
- Chaves, M. M., and Oliveira, M. M. (2004). Mechanisms underlying plant resilience to water deficits: prospects for water-saving agriculture. *J. Exp. Bot.* 55, 2365–2384. doi: 10.1093/jxb/erh269
- Chen, Z., Luo, F., Feng, J., Yang, W., Zeng, D., Zhao, R., et al. (2013). Genomic and structural characterization of Kunitz-type peptide LmKTT-1a highlights diversity and evolution of scorpion potassium channel toxins. *PLoS ONE* 8:e60201. doi: 10.1371/journal.pone.0060201
- Chmielewska, K., Rodziewicz, P., Swarczewicz, B., Sawikowska, A., Krajewski, P., Marczak, E., et al. (2016). Analysis of drought-induced proteomic and metabolomic changes in barley (*Hordeum vulgare* L.) leaves and roots unravels some aspects of biochemical mechanisms involved in drought tolerance. *Front. Plant Sci.* 7:1108. doi: 10.3389/fpls.2016.01108
- Corso, M., Vannozzi, A., Maza, E., Vitulo, N., Meggio, F., Pitacco, A., et al. (2015). Comprehensive transcript profiling of two grapevine rootstock genotypes comprehending in drought susceptibility links the phenylpropanoid pathway to enhanced tolerance. *J. Exp. Bot.* 66, 5739–5752. doi: 10.1093/jxb/erv274
- Crisp, P. A., Ganguly, D., Eichten, S. R., Borevitz, J. O., and Pogson, B. J. (2016). Reconsidering plant memory: intersections between stress recovery, RNA turnover, and epigenetics. *Sci. Adv.* 2:e1501340. doi: 10.1126/sciadv.1501340
- de A Gerhardt, L. B., Sachetto-Martins, G., Contarini, M. G., Sandroni, M., de P Ferreira, R., de Lima, V. M., et al. (1997). *Arabidopsis thaliana* class IV chitinase is early induced during the interaction with *Xanthomonas campestris*. *FEBS Lett.* 419, 69–75. doi: 10.1016/S0014-5793(97)01332-X
- De Meester, P., Brick, P., Lloyd, L. F., Blow, D. M., and Onesti, S. (1998). Structure of the Kunitz-type soybean trypsin inhibitor (STI): implication for the interactions between members of the STI family and tissue-plasminogen activator. *Acta Crystallogr. D Biol. Crystallogr.* 54, 589–597. doi: 10.1107/S0907444997015849
- De Nardi, B., Dreos, R., Del Terra, L., Martellosi, C., Asquini, E., Tornincasa, P., et al. (2006). Differential responses of *Coffea arabica* L. leaves and roots to chemically induced systemic acquired resistance. *Genome* 49, 1594–1605. doi: 10.1139/g06-125
- Deom, C. M., Schubert, K. R., Wolf, S., Holt, C. A., Lucas, W. J., and Beachy, R. N. (1990). Molecular characterization and biological function of the movement protein of tobacco mosaic virus in transgenic plants. *Proc. Natl. Acad. Sci. U.S.A.* 87, 3284–3288.
- Dong, X., Wang, D., Liu, P., Li, C., Zhao, Q., Zhu, D., et al. (2013). Zm908p11, encoded by a short open reading frame (sORF) gene, functions in pollen tube growth as a profilin ligand in maize. *J. Exp. Bot.* 64, 2359–2372. doi: 10.1093/jxb/ert093
- Dorokhov, Y. L., Komarova, T. V., Petrunia, I. V., Frolova, O. Y., Pozdyshev, D. V., and Gleba, Y. Y. (2012). Airborne signals from a wounded leaf facilitate



- viral spreading and induce antibacterial resistance in neighboring plants. *PLoS Pathog.* 8:e1002640. doi: 10.1371/journal.ppat.1002640
- Dorokhov, Y. L., Miroshnichenko, N. A., Alexandrova, N. M., and Atabekov, J. G. (1981). Development of systemic TMV infection in upper noninoculated tobacco leaves after differential temperature treatment. *Virology* 108, 507–509. doi: 10.1016/0042-6822(81)90458-X
- Dorokhov, Y. L., Skulachev, M. V., Ivanov, P. A., Zvereva, S. D., Tjulkina, L. G., Merits, A., et al. (2002). Polypurine (A)-rich sequences promote cross-kingdom conservation of internal ribosome entry. *Proc. Natl. Acad. Sci. U.S.A.* 99, 5301–5306. doi: 10.1073/pnas.082107599
- Dowen, R. H., Pelizzola, M., Schmitz, R. J., Lister, R., Dowen, J. M., Nery, J. R., et al. (2012). Widespread dynamic DNA methylation in response to biotic stress. *Proc. Natl. Acad. Sci. U.S.A.* 109, E2183–2191. doi: 10.1073/pnas.1209329109
- Duran-Flores, D., and Heil, M. (2016). Sources of specificity in plant damaged-self recognition. *Curr. Opin. Plant Biol.* 32, 77–87. doi: 10.1016/j.pbi.2016.06.019
- Espinas, N. A., Saze, H., and Saijo, Y. (2016). Epigenetic control of defense signaling and priming in plants. *Front. Plant Sci.* 7:1201. doi: 10.3389/fpls.2016.01201
- Eulgem, T., and Somssich, I. E. (2007). Networks of WRKY transcription factors in defense signaling. *Curr. Opin. Plant Biol.* 10, 366–371. doi: 10.1016/j.pbi.2007.04.020
- Fan, W., Zhao, M., Li, S., Bai, X., Li, J., Meng, H., et al. (2016). Contrasting transcriptional responses of PYR1/PYL/RCAR ABA receptors to ABA or dehydration stress between maize seedling leaves and roots. *BMC Plant Biol.* 16:99. doi: 10.1186/s12870-016-0764-x
- Fang, C., Zhang, P., Jian, X., Chen, W., Lin, H., Li, Y., et al. (2017). Overexpression of Lsi1 in cold-sensitive rice mediates transcriptional regulatory networks and enhances resistance to chilling stress. *Plant Sci. Int. J. Exp. Plant Biol.* 262, 115–126. doi: 10.1016/j.plantsci.2017.06.002
- Freschet, G. T., Swart, E. M., and Cornelissen, J. H., (2015). Integrated plant phenotypic responses to contrasting above- and below-ground resources: key roles of specific leaf area and root mass fraction. *New Phytol.* 206, 1247–1260. doi: 10.1111/nph.13352
- Fu, Z. Q., and Dong, X. (2013). Systemic acquired resistance: turning local infection into global defense. *Annu. Rev. Plant Biol.* 64, 839–863. doi: 10.1146/annurev-arplant-042811-105606
- Fujita, M., Fujita, Y., Noutoshi, Y., Takahashi, F., Narusaka, Y., Yamaguchi-Shinozaki, K., et al. (2006). Crosstalk between abiotic and biotic stress responses: a current view from the points of convergence in the stress signaling networks. *Curr. Opin. Plant Biol.* 9, 436–442. doi: 10.1016/j.pbi.2006.05.014
- Galiana, E., Bonnet, P., Conrod, S., Keller, H., Panabières, F., Ponchet, M., et al. (1997). RNase activity prevents the growth of a fungal pathogen in tobacco leaves and increases upon induction of systemic acquired resistance with elicitor. *Plant Physiol.* 115, 1557–1567. doi: 10.1104/pp.115.4.1557
- Germundsson, A., Sandgren, M., Barker, H., Savenkov, E. L., and Valkonen, J. P., (2002). Initial infection of roots and leaves reveals different resistance phenotypes associated with coat protein gene-mediated resistance to Potato mop-top virus. *J. Gen. Virol.* 83, 1201–1209. doi: 10.1099/0022-1317-83-5-1201
- Goodwin, P. H., Xie, W., and Valliani, M. (2012). Three genes of miraculin-like proteins from *Nicotiana benthamiana* with dissimilar putative structures show highly similar patterns of induction following bacterial and fungal infections. *Eur. J. Plant Pathol.* 134, 795–810. doi: 10.1007/s10658-012-0056-8
- Goralski, M., Sobieszczyńska, P., Obrepalska-Stepłowska, A., Swiercz, A., Zmienko, A., and Figlerowicz, M. (2016). A gene expression microarray for *Nicotiana benthamiana* based on *de novo* transcriptome sequence assembly. *Plant Methods* 12:28. doi: 10.1186/s13007-016-0128-4
- Grosse-Holz, F. M., and van der Hoorn, R. A., (2016). Juggling jobs: roles and mechanisms of multifunctional protease inhibitors in plants. *New Phytol.* 210, 794–807. doi: 10.1111/nph.13839
- Gruau, C., Trotel-Aziz, P., Villaume, S., Rabenoelina, F., Clément, C., Baillieux, F., et al. (2015). *Pseudomonas Fluorescens* PTA-CT2 triggers local and systemic immune response against botrytis cinerea in grapevine. *Mol. Plant-Microbe Interact.* 28, 1117–1129. doi: 10.1094/MPMI-04-15-0092-R
- Gruden, K., Strukelj, B., Ravnikar, M., Poljsak-Prijatelj, M., Mavric, I., Brzin, J., et al. (1997). Potato cysteine proteinase inhibitor gene family: molecular cloning, characterisation and immunocytochemical localisation studies. *Plant Mol. Biol.* 34, 317–323. doi: 10.1023/A:1005853026333
- Guo, J., Erskine, P. T., Coker, A. R., Wood, S. P., Cooper, J. B., Mareš, M., et al. (2015). Structure of a Kunitz-type potato cathepsin D inhibitor. *J. Struct. Biol.* 192, 554–560. doi: 10.1016/j.jsb.2015.10.020
- Gupta, P., Rangan, L., Ramesh, T. V., and Gupta, M. (2016). Comparative analysis of contextual bias around the translation initiation sites in plant genomes. *J. Theor. Biol.* 404, 303–311. doi: 10.1016/j.jtbi.2016.06.015
- Gurung, B., Bhardwaj, P. K., and Talukdar, N. C. (2016). Subtractive transcriptome analysis of leaf and rhizome reveals differentially expressed transcripts in *Panax sokpayensis*. *Funct. Integr. Genomics.* 16, 619–639. doi: 10.1007/s10142-016-0517-9
- Haimov, O., Sinvani, H., and Dikstein, R. (2015). Cap-dependent, scanning-free translation initiation mechanisms. *Biochim. Biophys. Acta* 1849, 1313–1318. doi: 10.1016/j.bbagr.2015.09.006
- Hanada, K., Akiyama, K., Sakurai, T., Toyoda, T., Shinozaki, K., and Shiu, S. H. (2010). sORF finder: a program package to identify small open reading frames with high coding potential. *Bioinforma. Oxf. Engl.* 26, 399–400. doi: 10.1093/bioinformatics/btp688
- Hanada, K., Higuchi-Takeuchi, M., Okamoto, M., Yoshizumi, T., Shimizu, M., Nakaminami, K., et al. (2013). Small open reading frames associated with morphogenesis are hidden in plant genomes. *Proc. Natl. Acad. Sci. U.S.A.* 110, 2395–2400. doi: 10.1073/pnas.1213958110
- Hanin, M., Ebel, C., Ngom, M., Laplace, L., and Masmoudi, K. (2016). New insights on plant salt tolerance mechanisms and their potential use for breeding. *Front. Plant Sci.* 7:1787. doi: 10.3389/fpls.2016.01787
- Hashimoto, M., Komatsu, K., Iwai, R., Keima, T., Maejima, K., Shiraishi, T., et al. (2015). Cell death triggered by a putative amphipathic helix of radish mosaic virus helicase protein is tightly correlated with host membrane modification. *Mol. Plant-Microbe Interact.* 28, 675–688. doi: 10.1094/MPMI-01-15-0004-R
- Hayden, C. A., and Jorgensen, R. A. (2007). Identification of novel conserved peptide uORF homology groups in *Arabidopsis* and rice reveals ancient *Eukaryotic* origin of select groups and preferential association with transcription factor-encoding genes. *BMC Biol.* 5:32. doi: 10.1186/1741-7007-5-32
- He, N., Yakiyama, M., Fujii, H., Banno, Y., and Yamamoto, K. (2003). Genomic structure and expression analysis of the gene encoding a silkworm basic Kunitz-type chymotrypsin inhibitor. *Biochim. Biophys. Acta* 1628, 71–77. doi: 10.1016/S0167-4781(03)00118-0
- Hermesmeier, D., Schittko, U., and Baldwin, I. T. (2001). Molecular interactions between the specialist herbivore *manduca sexta* (*Lepidoptera, Sphingidae*) and its natural host *Nicotiana attenuata*. I. Large-scale changes in the accumulation of growth- and defense-related plant mRNAs. *Plant Physiol.* 125, 683–700. doi: 10.1104/pp.125.2.683
- Hirel, B., Le Gouis, J., Ney, B., and Gallais, A. (2007). The challenge of improving nitrogen use efficiency in crop plants: towards a more central role for genetic variability and quantitative genetics within integrated approaches. *J. Exp. Bot.* 58, 2369–2387. doi: 10.1093/jxb/erm097
- Hörger, A. C., and van der Hoorn, R. A. L. (2013). The structural basis of specific protease-inhibitor interactions at the plant-pathogen interface. *Curr. Opin. Struct. Biol.* 23, 842–850. doi: 10.1016/j.sbi.2013.07.013
- Hsu, P. Y., Calviello, L., Wu, H.-Y. L., Li, F.-W., Rothfels, C. J., Ohler, U., et al. (2016). Super-resolution ribosome profiling reveals unannotated translation events in *Arabidopsis*. *Proc. Natl. Acad. Sci.* 113, E7126–E7135. doi: 10.1073/pnas.1614788113
- Huang, H., Qi, S. D., Qi, F., Wu, C. A., Yang, G. D., and Zheng, C. C., (2010). NtKTI1, a Kunitz trypsin inhibitor with antifungal activity from *Nicotiana tabacum*, plays an important role in tobacco's defense response. *FEBS J.* 277, 4076–4088. doi: 10.1111/j.1742-4658.2010.07803.x
- Isaeva, M. P., Chausova, V. E., Zelepuga, E. A., Guzev, K. V., Tabakmakher, V. M., Monastyrnaya, M. M., et al. (2012). A new multigene superfamily of Kunitz-type protease inhibitors from sea anemone *Heteractis crispata*. *Peptides* 34, 88–97. doi: 10.1016/j.peptides.2011.09.022
- Ishikawa, A., Ohta, S., Matsuoka, K., Hattori, T., and Nakamura, K. (1994). A family of potato genes that encode Kunitz-type proteinase inhibitors: structural comparisons and differential expression. *Plant Cell Physiol.* 35, 303–312.
- Islam, A., Leung, S., Burgess, E. P., Laing, W. A., Richardson, K. A., Hofmann, R. W., et al. (2015a). Knock-down of transcript abundance of a family of Kunitz proteinase inhibitor genes in white clover (*Trifolium repens*) reveals

- a redundancy and diversity of gene function. *New Phytol.* 208, 1188–1201. doi: 10.1111/nph.13543
- Islam, A., Mercer, C. F., Leung, S., Dijkwel, P. P., and McManus, M. T. (2015b). Transcription of biotic stress associated genes in white clover (*Trifolium repens* L.) differs in response to cyst and root-knot *Nematode* infection. *PLoS ONE* 10:e0137981. doi: 10.1371/journal.pone.0137981
- Itoh, H., Kataoka, H., Hamasuna, R., Kitamura, N., and Koono, M. (1999). Hepatocyte growth factor activator inhibitor type 2 lacking the first Kunitz-type serine proteinase inhibitor domain is a predominant product in mouse but not in human. *Biochem. Biophys. Res. Commun.* 255, 740–748. doi: 10.1006/bbrc.1999.0268
- Jaffe, M. J., and Forbes, S. (1993). Thigmomorphogenesis: the effect of mechanical perturbation on plants. *Plant Growth Regul.* 12, 313–324. doi: 10.1007/BF00027213
- Jang, J.-C. (2016). Arginine-rich motif-tandem CCH zinc finger proteins in plant stress responses and post-transcriptional regulation of gene expression. *Plant Sci. Int. J. Exp. Plant Biol.* 252, 118–124. doi: 10.1016/j.plantsci.2016.06.014
- Jashni, M. K., Mehrabi, R., Collemare, J., Mesarich, C. H., and de Wit, P. J. G. M. (2015). The battle in the apoplast: further insights into the roles of proteases and their inhibitors in plant-pathogen interactions. *Front. Plant Sci.* 6:584. doi: 10.3389/fpls.2015.00584
- Jiang, J., Ma, S., Ye, N., Jiang, M., Cao, J., and Zhang, J. (2017). WRKY transcription factors in plant responses to stresses. *J. Integr. Plant Biol.* 59, 86–101. doi: 10.1111/jipb.12513
- Jofuku, K. D., and Goldberg, R. B. (1989). Kunitz trypsin inhibitor genes are differentially expressed during the soybean life cycle and in transformed tobacco plants. *Plant Cell* 1, 1079–1093. doi: 10.1105/tpc.1.1.1079
- Jones, D. T., Taylor, W. R., and Thornton, J. M. (1992). The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* 8, 275–282. doi: 10.1093/bioinformatics/8.3.275
- Jones, J. D., and Dangl, J. L. (2006). The plant immune system. *Nature* 444, 323–329. doi: 10.1038/nature05286
- Joshi, C. P., Zhou, H., Huang, X., and Chiang, V. L. (1997). Context sequences of translation initiation codon in plants. *Plant Mol. Biol.* 35, 993–1001. doi: 10.1023/A:1005816823636
- Juntawong, P., Girke, T., Bazin, J., and Bailey-Serres, J. (2014). Translational dynamics revealed by genome-wide profiling of ribosome footprints in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 111, E203–212. doi: 10.1073/pnas.1317811111
- Kawaguchi, R., and Bailey-Serres, J. (2005). mRNA sequence features that contribute to translational regulation in *Arabidopsis*. *Nucleic Acids Res.* 33, 955–965. doi: 10.1093/nar/gki240
- Kergunteuil, A., Bakhtiari, M., Formenti, L., Xiao, Z., Defossez, E., and Rasmann, S. (2016). Biological control beneath the feet: a review of crop protection against insect root herbivores. *Insects* 7:70. doi: 10.3390/insects7040070
- Kim, Y., Lee, G., Jeon, E., Sohn, E., Ju, Lee, Y., Kang, H., et al. (2014). The immediate upstream region of the 5'-UTR from the AUG start codon has a pronounced effect on the translational efficiency in *Arabidopsis thaliana*. *Nucleic Acids Res.* 42, 485–498. doi: 10.1093/nar/gkt864
- Kirchhofer, D., Peek, M., Li, W., Stamos, J., Eigenbrot, C., Kadkhodayan, S., et al. (2003). Tissue expression, protease specificity, and Kunitz domain functions of hepatocyte growth factor activator inhibitor-1B (HAI-1B), a new splice variant of HAI-1. *J. Biol. Chem.* 278, 36341–36349. doi: 10.1074/jbc.M304643200
- Kobayashi, A., Miura, S., and Kozaki, A. (2017). INDETERMINATE DOMAIN PROTEIN binding sequences in the 5'-untranslated region and promoter of the SCARECROW gene play crucial and distinct roles in regulating SCARECROW expression in roots and leaves. *Plant Mol. Biol.* 94, 1–13. doi: 10.1007/s11103-016-0578-0
- Kochetov, A. V., Prayaga, P. D., Volkova, O. A., and Sankaramakrishnan, R. (2013). Hidden coding potential of *Eukaryotic* genomes: nonAUG started ORFs. *J. Biomol. Struct. Dyn.* 31, 103–114. doi: 10.1080/07391102.2012.691367
- Komarnytsky, S., and Borisjuk, N. (2003). Functional analysis of promoter elements in plants. *Genet. Eng. (N.Y.)* 25, 113–141. doi: 10.1007/978-1-4615-0073-6\_6
- Komarova, T. V., Kosorukov, V. S., Frolova, O. Y., Petrunia, I. V., Skrypnik, K. A., Gleba, Y. Y., et al. (2011). Plant-made trastuzumab (*Herceptin*) inhibits HER2/Neu+ cell proliferation and retards tumor growth. *PLoS ONE* 6:e17541. doi: 10.1371/journal.pone.0017541
- Komarova, T. V., Sheval, E. V., Pozdyshev, D. V., Kolesnikova, V. S., and Dorokhov, Y. L. (2012). Rapid and massive green fluorescent protein production leads to formation of protein Y-bodies in plant cells. *Biochem. (Mosc.)* 77, 603–608. doi: 10.1134/S0006297912060065
- Kozak, M. (1986). Point mutations define a sequence flanking the AUG initiator codon that modulates translation by *Eukaryotic* ribosomes. *Cell* 44, 283–292. doi: 10.1016/0092-8674(86)90762-2
- Leister, D., Wang, L., and Kleine, T. (2017). Organellar gene expression and acclimation of plants to environmental stress. *Front. Plant Sci.* 8:387. doi: 10.3389/fpls.2017.00387
- Li, J., Brader, G., and Palva, E. T. (2008). Kunitz trypsin inhibitor: an antagonist of cell death triggered by phytopathogens and fumonisin B1 in *Arabidopsis*. *Mol. Plant* 1, 482–495. doi: 10.1093/mp/ssp013
- Liu, B., and Qian, S. B. (2014). Translational reprogramming in cellular stress response. *Wiley Interdiscip. Rev. RNA* 5, 301–315. doi: 10.1002/wrna.1212
- Liu, H., Sultan, M. A., Liu, X. L., Zhang, J., Yu, F., and Zhao, H. X. (2015). Physiological and comparative proteomic analysis reveals different drought responses in roots and leaves of drought-tolerant wild wheat (*Triticum boeoticum*). *PLoS ONE* 10:e0121852. doi: 10.1145/2818302
- Lup, S. D., Tian, X., Xu, J., and Pérez-Pérez, J. M. (2016). Wound signaling of regenerative cell reprogramming. *Plant Sci. Int. J. Exp. Plant Biol.* 250, 178–187. doi: 10.1016/j.plantsci.2016.06.012
- Major, I. T., and Constabel, C. P. (2008). Functional analysis of the Kunitz trypsin inhibitor family in poplar reveals biochemical diversity and multiplicity in defense against herbivores. *Plant Physiol.* 146, 888–903. doi: 10.1104/pp.107.106229
- Marand, A. P., Zhang, T., Zhu, B., and Jiang, J. (2017). Towards genome-wide prediction and characterization of enhancers in plants. *Biochim. Biophys. Acta* 1860, 131–139. doi: 10.1016/j.bbagr.2016.06.006
- Matz, M., Shagin, D., Bogdanova, E., Britanova, O., Lukyanov, S., Diatchenko, L., et al. (1999). Amplification of cDNA ends based on template-switching effect and step-out PCR. *Nucleic Acids Res.* 27, 1558–1560. doi: 10.1093/nar/27.6.1558
- Mercer, T. R., and Mattick, J. S. (2013). Understanding the regulatory and transcriptional complexity of the genome through structure. *Genome Res.* 23, 1081–1088. doi: 10.1101/gr.156612.113
- Mercer, T. R., Clark, M. B., Andersen, S. B., Brunck, M. E., Haerty, W., Crawford, J., et al. (2015). Genome-wide discovery of human splicing branchpoints. *Genome Res.* 25, 290–303. doi: 10.1101/gr.182899.114
- Miao, L., Zhang, L., Raboanahary, N., Lu, G., Zhang, X., Xiang, J., et al. (2016). Transcriptome analysis of stem and globally comparison with other tissues in *Brassica napus*. *Front. Plant Sci.* 7:1403. doi: 10.3389/fpls.2016.01403
- Miras, M., Miller, W. A., Truniger, V., and Aranda, M. A. (2017). Non-canonical translation in plant RNA viruses. *Front. Plant Sci.* 8:494. doi: 10.3389/fpls.2017.00494
- Misas-Villamil, J. C., van der Hoorn, R. A., and Doehlemann, G. (2016). Papain-like cysteine proteases as hubs in plant immunity. *New Phytol.* 212, 902–907. doi: 10.1111/nph.14117
- Moore, C. J., Sutherland, P. W., Forster, R. L., Gardner, R. C., and MacDiarmid, R. M. (2001). Dark green islands in plant virus infection are the result of posttranscriptional gene silencing. *Mol. Plant Microbe Interact.* 14, 939–946. doi: 10.1094/MPMI.2001.14.8.939
- Moulleron, H., Delcourt, V., and Roucou, X. (2016). Death of a dogma: *Eukaryotic* mRNAs can code for more than one protein. *Nucleic Acids Res.* 44, 14–23. doi: 10.1093/nar/gkv1218
- Nakashima, K., and Yamaguchi-Shinozaki, K. (2006). Regulons involved in osmotic stress-responsive and cold stress-responsive gene expression in plants. *Physiol. Plant.* 126, 62–71. doi: 10.1111/j.1399-3054.2005.00592.x
- Nakasugi, K., Crowhurst, R., Bally, J., and Waterhouse, P. (2014). Combining transcriptome assemblies from multiple de novo assemblers in the allo-tetraploid plant *Nicotiana benthamiana*. *PLoS ONE* 9:e91776. doi: 10.1371/journal.pone.0091776
- Neto, J. R. C. F., da Silva, M. D., Pandolfi, V., Crovella, S., Benko-Iseppon, A. M., and Kido, E. A. (2016). Epigenetic signals on plant adaptation: a biotic stress perspective. *Curr. Protein Pept. Sci.* 18, 352–367. doi: 10.2174/1389203717666160724194653

- Nguyen, V. N., Vo, K. T., Park, H., Jeon, J.-S., and Jung, K. H. (2016). A systematic view of the MLO family in rice suggests their novel roles in morphological development, diurnal responses, the light-signaling pathway, and various stress responses. *Front. Plant Sci.* 7:1413. doi: 10.3389/fpls.2016.01413
- Okubara, P. A., and Paulitz, T. C. (2005). Root defense responses to fungal pathogens: a molecular perspective. *Plant Soil* 274, 215–226. doi: 10.1007/s11104-004-7328-9
- Oliva, M. L. V., Ferreira, R., da, S., Ferreira, J. G., de Paula, C. A. A., Salas, C. E., and Sampaio, M. U. (2011). Structural and functional properties of kunitz proteinase inhibitors from *Leguminosae*: a mini review. *Curr. Protein Pept. Sci.* 12, 348–357. doi: 10.2174/138920311796391061
- Oliva, M. L. V., Silva, M. C. C., Sallai, R. C., Brito, M. V., and Sampaio, M. U. (2010). A novel subclassification for Kunitz proteinase inhibitors from *Leguminosae* seeds. *Biochimie* 92, 1667–1673. doi: 10.1016/j.biochi.2010.03.021
- Parmar, N., Singh, K. H., Sharma, D., Singh, L., Kumar, P., Nanjundan, J., et al. (2017). Genetic engineering strategies for biotic and abiotic stress tolerance and quality enhancement in horticultural crops: a comprehensive review. *3 Biotech* 7:239. doi: 10.1007/s13205-017-0870-y
- Philippe, R. N., Ralph, S. G., K ulheim, C., Jancsik, S. I., and Bohlmann, J. (2009). Poplar defense against insects: genome analysis, full-length cDNA cloning, and transcriptome and protein analysis of the poplar Kunitz-type protease inhibitor family. *New Phytol.* 184, 865–884. doi: 10.1111/j.1469-8137.2009.03028.x
- Pucholt, P., Sj odin, P., Weih, M., R nnerberg-W stljung, A. C., and Berlin, S. (2015). Genome-wide transcriptional and physiological responses to drought stress in leaves and roots of two willow genotypes. *BMC Plant Biol.* 15:244. doi: 10.1186/s12870-015-0630-2
- Ramegowda, V., and Senthil-Kumar, M. (2015). The interactive effects of simultaneous biotic and abiotic stresses on plants: mechanistic understanding from drought and pathogen combination. *J. Plant Physiol.* 176, 47–54. doi: 10.1016/j.jplph.2014.11.008
- Rawlings, N. D., Barrett, A. J., and Finn, R. (2016). Twenty years of the MEROPS database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.* 44, D343–350. doi: 10.1093/nar/gkv1118
- Rawlings, N. D., Waller, M., Barrett, A. J., and Bateman, A. (2014). MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.* 42, D503–D509. doi: 10.1093/nar/gkt953
- Ribone, P. A., Capella, M., Arce, A. L., and Chan, R. L. (2017). A uORF represses the transcription factor AtHBI in aerial tissues to avoid a deleterious phenotype. *Plant Physiol.* 175, 1238–1253. doi: 10.1104/pp.17.01060
- Ribrioux, S., Br unger, A., Baumgarten, B., Seuwen, K., and John, M. R. (2008). Bioinformatics prediction of overlapping frameshifted translation products in mammalian transcripts. *BMC Genomics* 9:122. doi: 10.1186/1471-2164-9-122
- Ronsin, C., Chung-Scott, V., Poullion, I., Aknouche, N., Gaudin, C., and Triebel, F. (1999). A non-AUG-defined alternative open reading frame of the intestinal carboxyl esterase mRNA generates an epitope recognized by renal cell carcinoma-reactive tumor-infiltrating lymphocytes *in situ*. *J. Immunol. Baltim. Md* 163, 483–490.
- Rustgi, S., Boex-Fontvieille, E., Reinbothe, C., von Wettstein, D., and Reinbothe, S. (2017). Serpin1 and WSCP differentially regulate the activity of the cysteine protease RD21 during plant development in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* 114, 2212–2217. doi: 10.1073/pnas.1621496114
- Ryu, C.-M. (2015). Against friend and foe: type 6 effectors in plant-associated bacteria. *J. Microbiol. Seoul Korea* 53, 201–208. doi: 10.1007/s12275-015-5055-y
- Saarikoski, P., Clapham, D., and von Arnold, S. (1996). A wound-inducible gene from *Salix viminalis* coding for a trypsin inhibitor. *Plant Mol. Biol.* 31, 465–478. doi: 10.1007/BF00042221
- Sampathkumar, A., Yan, A., Krupinski, P., and Meyerowitz, E. M. (2014). Physical forces regulate plant development and morphogenesis. *Curr. Biol.* 24, R475–483. doi: 10.1016/j.cub.2014.03.014
- Sandbrink, R., Masters, C. L., and Beyreuther, K. (1994). APP gene family: unique age-associated changes in splicing of Alzheimer's beta A4-amyloid protein precursor. *Neurobiol. Dis.* 1, 13–24.
- Santamar a, M. E., Diaz-Mendoza, M., Diaz, I., and Martinez, M. (2014). Plant protein peptidase inhibitors: an evolutionary overview based on comparative genomics. *BMC Genomics* 15:812. doi: 10.1186/1471-2164-15-812
- Saplaoura, E., and Kragler, F. (2016). Mobile transcripts and intercellular communication in plants. *Enzymes* 40, 1–29. doi: 10.1016/bs.enz.2016.07.001
- Savatin, D. V., Gramegna, G., Modesti, V., and Cervone, F. (2014). Wounding in the plant tissue: the defense of a dangerous passage. *Front. Plant Sci.* 5:470. doi: 10.3389/fpls.2014.00470
- Schepetilnikov, M., and Ryabova, L. A. (2017). Auxin signaling in regulation of plant translation reinitiation. *Front. Plant Sci.* 8:1014. doi: 10.3389/fpls.2017.01014
- Schmidt, W. M., and Mueller, M. W. (1999). CapSelect: a highly sensitive method for 5' CAP-dependent enrichment of full-length cDNA in PCR-mediated analysis of mRNAs. *Nucleic Acids Res.* 27:e31.
- Sesma, A., Castresana, C., and Castellano, M. M. (2017). Regulation of translation by TOR, eIF4E and eIF2  in plants: current knowledge, challenges and future perspectives. *Front. Plant Sci.* 8:644. doi: 10.3389/fpls.2017.00644
- Shang, X., Cao, Y., and Ma, L. (2017). Alternative splicing in plant genes: a means of regulating the environmental fitness of plants. *Int. J. Mol. Sci.* 18:432. doi: 10.3390/ijms18020432
- Sohal, A. K., Love, A. J., Cecchini, E., Covey, S. N., Jenkins, G. I., and Milner, J. J. (1999). Cauliflower mosaic virus infection stimulates lipid transfer protein gene expression in *Arabidopsis*. *J. Exp. Bot.* 50, 1727–1733. doi: 10.1093/jxb/50.341.1727
- Suh, M. C., Oh, S.-K., Kim, Y.-C., Pai, H.-S., and Choi, D. (2003). Expression of a novel tobacco gene, NgCDM1, is preferentially associated with pathogen-induced cell death. *Physiol. Mol. Plant Pathol.* 62, 227–235. doi: 10.1016/S0885-5765(03)00049-3
- Suzuki, N., Rivero, R. M., Shulaev, V., Blumwald, E., and Mittler, R. (2014). Abiotic and biotic stress combinations. *New Phytol.* 203, 32–43. doi: 10.1111/nph.12797
- Takahashi, N., Hitotsuya, H., Hanzawa, H., Arata, Y., and Kurihara, Y. (1990). Structural study of asparagine-linked oligosaccharide moiety of taste-modifying protein, miraculin. *J. Biol. Chem.* 265, 7793–7798.
- Tanaka, M., Sotta, N., Yamazumi, Y., Yamashita, Y., Miwa, K., Murota, K., et al. (2016). The minimum open reading frame, AUG-stop, induces boron-dependent ribosome stalling and mRNA degradation. *Plant Cell* 28, 2830–2849. doi: 10.1105/tpc.16.00481
- Terenin, I. M., Smirnova, V. V., Andreev, D. E., Dmitriev, S. E., and Shatsky, I. N. (2017). A researcher's guide to the galaxy of IRESs. *Cell. Mol. Life Sci. CMLS* 74, 1431–1455. doi: 10.1007/s00018-016-2409-5
- Tian, W. M., Peng, S. Q., Wang, X. C., Shi, M. J., Chen, Y. Y., and Hu, Z. H. (2007). Vegetative storage protein in *Litchi chinensis*, a subtropical evergreen fruit tree, possesses trypsin inhibitor activity. *Ann. Bot.* 100, 1199–1208. doi: 10.1093/aob/mcm216
- Tran, M. K., Schultz, C. J., and Baumann, U. (2008). Conserved upstream open reading frames in higher plants. *BMC Genomics* 9:361. doi: 10.1186/1471-2164-9-361
- Ueda, M., Manabe, Y., and Mukai, M. (2011). The high performance of 3XFLAG for target purification of a bioactive metabolite: a tag combined with a highly effective linker structure. *Bioorg. Med. Chem. Lett.* 21, 1359–1362. doi: 10.1016/j.bmcl.2011.01.038
- Vanderperre, B., Lucier, J. F., Bissonnette, C., Motard, J., Tremblay, G., Vanderperre, S., et al. (2013). Direct detection of alternative open reading frames translation products in human significantly expands the proteome. *PLoS ONE* 8:e70698. doi: 10.1371/journal.pone.0070698
- Vanderperre, B., Staskevicius, A. B., Tremblay, G., McCoy, M., O'Neill, M. A., Cashman, N. R., et al. (2011). An overlapping reading frame in the PRNP gene encodes a novel polypeptide distinct from the prion protein. *FASEB J.* 25, 2373–2386. doi: 10.1096/fj.10-173815
- Vaughn, J. N., Ellingson, S. R., Mignone, F., and Arnim, A., von (2012). Known and novel post-transcriptional regulatory sequences are conserved across plant families. *RNA* 18, 368–384. doi: 10.1261/rna.031179.111
- Vishwakarma, K., Upadhyay, N., Kumar, N., Yadav, G., Singh, J., Mishra, R. K., et al. (2017). Abscisic acid signaling and abiotic stress tolerance in plants: a review on current knowledge and future prospects. *Front. Plant Sci.* 8:161. doi: 10.3389/fpls.2017.00161
- Wang, R. F., Parkhurst, M. R., Kawakami, Y., Robbins, P. F., and Rosenberg, S. A. (1996). Utilization of an alternative open reading frame of a normal gene in generating a novel human cancer antigen. *J. Exp. Med.* 183, 1131–1140. doi: 10.1084/jem.183.3.1131
- Wang, W., Vinocur, B., and Altman, A. (2003). Plant responses to drought, salinity and extreme temperatures: towards genetic engineering for stress tolerance. *Planta* 218, 1–14. doi: 10.1007/s00425-003-1105-5

- Wood, J. P., Ellery, P. E. R., Maroney, S. A., and Mast, A. E. (2014). Biology of tissue factor pathway inhibitor. *Blood* 123, 2934–2943. doi: 10.1182/blood-2013-11-512764
- Xu, G., Greene, G. H., Yoo, H., Liu, L., Marqués, J., Motley, J., et al. (2017a). Global translational reprogramming is a fundamental layer of immune regulation in plants. *Nature* 545, 487–490. doi: 10.1038/nature22371
- Xu, G., Yuan, M., Ai, C., Liu, L., Zhuang, E., Karapetyan, S., et al. (2017b). uORF-mediated translation allows engineered plant disease resistance without fitness costs. *Nature* 545, 491–494. doi: 10.1038/nature22372
- Yang, Y., Wang, W., Chu, Z., Zhu, J. K., and Zhang, H. (2017). Roles of nuclear pores and nucleo-cytoplasmic trafficking in plant stress responses. *Front. Plant Sci.* 8:574. doi: 10.3389/fpls.2017.00574
- Zavaliev, R., Levy, A., Gera, A., and Epel, B. L. (2013). Subcellular dynamics and role of Arabidopsis  $\beta$ -1,3-glucanases in cell-to-cell movement of tobamoviruses. *Mol. Plant-Microbe Interact.* 26, 1016–1030. doi: 10.1094/MPMI-03-13-0062-R
- Zdor, R. E., and Anderson, A. J. (1992). Influence of root colonizing bacteria on the defense responses of bean. *Plant Soil* 140, 99–107. doi: 10.1007/BF00012811
- Zhang, H., and Sonnewald, U. (2017). Differences and commonalities of plant responses to single and combined stresses. *Plant J. Cell Mol. Biol.* 90, 839–855. doi: 10.1111/tpj.13557

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Sheshukova, Komarova, Ershova, Shindyapina and Dorokhov. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.