

Light regulates widespread plant alternative polyadenylation through the chloroplast

M. Guillermina Kubaczka¹, Micaela A. Godoy Herz¹, Wei-Chun Chen², Dinghai Zheng²,
Ezequiel Petrillo¹, Bin Tian^{2,3} and Alberto R. Kornblihtt^{1*}

¹Universidad de Buenos Aires (UBA), Facultad de Ciencias Exactas y Naturales,
Departamento de Fisiología, Biología Molecular y Celular and CONICET-UBA, Instituto
de Fisiología, Biología Molecular y Neurociencias (IFIBYNE), 1428 Buenos Aires,
Argentina.

²Department of Microbiology, Biochemistry and Molecular Genetics, Rutgers New Jersey
Medical School, Newark, NJ 07103, USA

³Gene Expression and Regulation Program, and Center for Systems and Computational
Biology, The Wistar Institute, Philadelphia, PA 19104, USA.

*Corresponding author (ark@fbmc.fcen.uba.ar)

Abstract

Transcription of eukaryotic protein-coding genes generates immature mRNAs that are subjected to a series of processing events, including capping, splicing, cleavage and polyadenylation (CPA) and chemical modifications of bases. Alternative polyadenylation (APA) greatly contributes to mRNA diversity in the cell. By determining the length of the 3' untranslated region, APA generates transcripts with different regulatory elements, such as miRNA and RBP binding sites, which can influence mRNA stability, turnover and translation. In the model plant *Arabidopsis thaliana*, APA is involved in the control of seed dormancy and flowering. In view of the physiological importance of APA in plants, we decided to investigate the effects of light/dark conditions and compare the underlying mechanisms to those elucidated for alternative splicing (AS). We found that light controls APA in approximately 30% of *Arabidopsis* genes. Similar to AS, the effect of light on APA requires functional chloroplasts, is not affected in mutants of the phytochrome and cryptochrome photoreceptor pathways and is observed in roots only when the communication with the photosynthetic tissues is not interrupted. Furthermore, mitochondrial activity is necessary for the effect of light in roots but not in shoots. However, unlike AS, coupling with transcriptional elongation does not seem to be involved since light-dependent APA regulation is neither abolished in mutants of the TFIIS transcript elongation factor nor universally affected by chromatin relaxation caused by the histone deacetylase inhibition. Instead, regulation seems to be linked to light-elicited changes in the abundance of constitutive CPA factors, also mediated by the chloroplast.

Introduction

Transcription of eukaryotic protein-coding genes generates immature mRNAs that are subjected to a series of co-transcriptional processing events, including capping, splicing, cleavage and polyadenylation (CPA) and chemical modifications of bases. CPA is responsible for cleavage of the nascent RNA and addition of a poly(A) tail to cleaved RNA. Most of the core CPA factors found in yeast and mammalian cells, about 20 in total, have homologous counterparts in the plant *Arabidopsis thaliana*, in particular within the complexes CPSF (cleavage and polyadenylation specificity factor), CstF (cleavage stimulation factor) and CFII (cleavage factor II). Equivalent subunits of mammalian CFI

(cleavage factor I), however, have not been identified in plants (1). In all studied organisms, core CPA factors act together with the poly(A) polymerase and other RNA-binding proteins to yield poly(A)-tailed mature mRNAs. A key step in this process is the recognition of RNA motifs that define the polyA site (PAS). The AAUAAA hexamer is the most prominent PAS motif in metazoans, which is recognized by CPSF. However, this sequence has only been reported in approximately 10% of the *Arabidopsis* transcripts, which opens the yet-unsolved question of how the PAS is specified in plants (1).

Similar to alternative splicing (AS), alternative polyadenylation (APA) greatly contributes to mRNA diversity in the cell (2). By affecting the length and sequence of the 3' untranslated region (3'UTR), APA generates transcripts with different architectures of target sites for microRNAs, which affects mRNA turnover and translation. When coupled with AS, APA can also alter C-terminal segments of the polypeptides encoded by a single gene. For example, this mechanism is involved in the control of seed dormancy in plants. The *Arabidopsis* *DOG1* (delay of germination 1) gene gives rise to two mRNA isoforms that are generated by the use of proximal and distal (with respect to the promoter) PASs, respectively (3). The short and long mRNAs encode polypeptides of 30 and 32 kDa respectively, differing in a non-conserved C-terminal stretch. However, only the protein encoded by the short mRNA is functionally active in promoting dormancy. Mutations in CPA factors that lead to reduced use of the proximal PAS cause weakened seed dormancy (3). APA of non-coding RNAs also has fundamental biological roles in plants. The best studied example is the role of the long non-coding RNA *COOLAIR* in the control of flowering in *Arabidopsis*. *COOLAIR* is an antisense transcript of the flowering locus C gene (*FLC*) that displays two APA sites. When the distal site is used, a transcription-permissive histone mark is deployed along the *FLC* gene which allows for *FLC* expression and repression of flowering. Instead, if the proximal site is used, the short isoform of *COOLAIR* base-pairs with complementary *FLC* sequences, forming a so-called R-loop, which promotes the deployment of a transcription-repressive histone mark, which subsequently inhibits *FLC* expression and promotes flowering (4, 5).

We have previously shown that light regulates *Arabidopsis* AS through the chloroplast. We demonstrated that the photosynthetic electronic transport chain initiates a chloroplast retrograde signaling that regulates nuclear splicing (6). Later we found that light promotes RNAPII elongation while in darkness elongation is lower (7). These

changes in transcriptional elongation are causative of the observed changes in AS, indicating that the chloroplast control responds to the kinetic coupling between transcription and RNA processing found in mammalian cells (for reviews see references 8 and 9), and providing unique evidence that coupling is important for a whole organism to respond to environmental signals.

In view of the physiological importance of APA in plants, we decided to investigate the effects of light/ dark conditions and the underlying mechanisms and to compare them to the ones elucidated for AS. We report here that light deeply affects *Arabidopsis* 3'UTR APA in 28.7% (3,467) of 12,070 genes assessed by genome-wide analyses. Three gene groups are clearly defined. Those in which light promotes preferred usage of proximal PASs, those in which distal PASs are preferred and those in which light has not effect on APA at all. Similar to AS, the effect of light on APA requires functional chloroplasts, is not affected in mutants of the phytochrome and cryptochrome photoreceptor pathways, is observed in roots when the communication with the photosynthetic tissues is not interrupted. Furthermore, like in AS, mitochondrial activity is necessary for the APA effect of light in roots but not in shoots (10). However, unlike AS, coupling with transcriptional elongation does not seem to be involved in changes in APA triggered by light. Alternatively, regulation seems to be linked to light-elicited changes in the abundance of constitutive CPA factors, also mediated by the chloroplast.

Results

Arabidopsis genes display substantial 3'UTR APA isoform expression

In order to investigate the role of light/ dark exposure in *Arabidopsis* APA, we applied a light regime similar to the one used in our previous AS studies (6, 7, 10). Briefly, seedlings were grown for two weeks in constant white light to minimize interference from the circadian clock and then transferred to light or dark conditions for different time periods (Fig. 1A). To obtain APA regulation information genome-wide, we subjected total RNA to RNA sequencing by using the 3' region extraction and deep sequencing (3'READS) method (11, detailed in Materials and Methods). We generated ~82.5 million PAS-supporting reads from the two sample groups (Suppl. Table 1). After clustering of adjacent cleavage sites that resulted from heterogenous cleavage for each PAS (Suppl. Fig.

1A, see Materials and Methods for detail), we identified 160,881 PASs in the *Arabidopsis* genome. Consistent with previous studies (12, 13), plant PASs identified by 3'READS are surrounded with A/T-rich sequences, with a prominent T-rich peak around -10 nt and a modest A-rich peak around -20 nt (cleavage site is set to position 0, [Suppl. Fig. 1B](#)). By analyzing hexamers in neighboring regions of the PAS, we found substantially enriched TA-rich motifs surrounding *Arabidopsis* PASs ([Suppl. Fig. 1C](#)), which is similar to PASs in *S. cerevisiae* but distinct from those in metazoans (14-17).

Overall, 59% of all PASs were mapped to the last exon of mRNA genes, mostly in 3'UTRs. In addition, 10% were in introns, 6% in coding exons and 22% in intergenic regions ([Fig. 1B, left](#)). Notably, the vast majority of the PAS-supporting reads (96%) were mapped to the last exon ([Fig. 1B, right](#)), indicating that transcripts using last exon PASs are much more abundant than those using PASs in other regions.

Using transcript abundance of 5% as the cutoff to call an isoform ([Suppl. Fig. 2A](#)), we found that 81.6% of plant mRNA genes display APA in our samples. On average, an APA gene expresses 2.9 APA isoforms ([Suppl. Fig. 2B](#)). Of all APA genes, 91.9% have APA sites exclusively in the last exon and 4.7% have APA sites in both the last exon and upstream introns/exons ([Fig. 1C](#)). Therefore, most plant genes display 3'UTR APA through alternative usage of PASs in the last exon.

For genes that do not display APA, the median value of their 3'UTR size is 151 nt ([Fig. 1D](#)). By contrast, for genes that expression 3'UTR APA isoforms, the median values for the shortest 3'UTR and the longest 3'UTR are 125 nt and 214 nt, respectively ([Fig. 1D](#)). The alternative 3'UTR size has a median value of 78 nt across genes ([Suppl. Fig. 2C](#)). Therefore, 3'UTR APA in plants can potentially alter a substantial portion (30-60%) of the 3'UTR sequence, impacting post-transcriptional control of gene expression. We also found that the longest 3'UTR isoform is typically expressed at a higher level than the shortest 3'UTR isoform ([Suppl. Fig. 2D](#)). In addition, the PASs of the shortest and longest isoforms are surrounded with distinct mRNA motifs ([Suppl. Fig. 2E](#)), with the AATAAA motif highly enriched for the last PAS (longest 3'UTR isoform) and TGT motif for the first PAS (shortest 3'UTR isoform).

Light/dark switch elicits widespread APA isoform changes

147 We next examined APA isoform changes in light vs. dark conditions. Because most genes
148 display 3'UTR APA, we focused on the top two most expressed 3'UTR APA isoforms for
149 each gene. For simplicity, they are named proximal PAS (pPAS) isoform and distal PAS
150 (dPAS) isoform, respectively (Fig. 1E). We calculated their relative expression (RE, dPAS
151 vs. pPAS) in each sample group and then compared RE between the groups (light vs.
152 dark), yielding the value RED (relative expression difference).

153 Using p-value cutoff of 0.05 (DEXSeq), we identified 3,467 genes that showed
154 significant differences in APA isoform expression levels between light and dark
155 conditions (Fig. 1F). Strikingly, genes showing 3'UTR lengthening, i.e., upregulation of
156 dPAS isoform relative to pPAS isoform, outnumbered those showing 3'UTR shortening by
157 3.3-fold (Fig. 1F). By using Gene Ontology analysis, we found that genes showing 3'UTR
158 lengthening tend to have functions in various metabolic pathways, such as ATP, cellular
159 amide, hexose, etc. (Suppl. Table 2).

160 Representative genes showing 3'UTR lengthening, 3'UTR shortening or no change
161 are shown in Suppl. Fig. 3. *HTA9* and *RKH* are two paradigmatic examples of 3'UTR
162 lengthening and 3'UTR shortening genes, respectively (Fig. 2A). Using PCR primers
163 targeting alternative 3'UTR sequences of their isoforms and upstream common regions
164 (Suppl. Table 3), we validated our 3'READS data by RT-qPCR.

165 Our 3'READS data also detected widespread gene expression changes in
166 light/dark conditions (Suppl. Fig. 4A). We found that there is a significant bias for
167 upregulated genes having 3'UTR shortening and downregulated genes having 3'UTR
168 lengthening ($P = 0.002$, Fisher's exact test, Suppl. Fig. 4B). This result indicates that 3'UTR
169 APA is relevant to gene expression changes.

170 **The chloroplast is the sensor for the light effect on APA**

171 Next we wanted to address whether signaling through photoreceptors was involved in
172 APA regulation in response to light. We exposed seedlings of *Arabidopsis* mutants for the
173 red/far red-light photoreceptor phytochrome A (phyA-201, 18) and for the blue light
174 photoreceptor cryptochrome cry1 (cry1-1, 19) and cry2 (fha-1, 20) to our light regime
175 protocol (Fig. 2B-D), using the wild type Landsberg erecta (Ler) background as control.
176 APA was assessed by RT-qPCR in the *HTA9* and *RKH* genes, as examples of 3'UTR
177 lengthening and shortening by light respectively (Figs. 2B and 2C), and the *FAD6* gene
178 (Fig. 2D) as a gene with no change in APA. Both types of mutants behaved similarly to

179 wild type seedlings for all three genes, which refutes the notion that photoreceptors play a
180 role in on APA.

181 Because retrograde signals from the chloroplast have been shown to modulate
182 nuclear gene expression (6, 10, 21), we reasoned that chloroplast functions might be
183 involved in light-elicited APA regulation. To this end, we used the herbicide DCMU [3-
184 (3,4-dichlorophenyl)-1,1-dimethylurea] (22) that blocks the photosynthetic electron transport
185 from photosystem II to the plastoquinones. Interestingly, DCMU inhibited the effect of
186 light on APA in both *HTA9* (Fig. 2E) and *RKH* (Fig. 2F) genes but had no effect on *FAD6*
187 (Fig. 2G), suggesting that chloroplast function is necessary for modulating APA in
188 response to light.

190 **The light effect on APA is sensed by the photosynthetic tissues**

191 To obtain further evidence of the involvement of the chloroplast in APA we performed
192 dissection experiments. Because roots have no chloroplasts, we reasoned that if seedlings
193 were cut to separate roots from green photosynthetic tissues (shoots) (Figs. 3A and 3B),
194 APA differences upon light/dark treatments should only be observed in the green tissues.
195 Similarly to what happens with AS (6), in the case of the *HTA9* gene the effect was
196 observed both in dissected leaves and roots (Fig. 3C, left) when dissection was performed
197 6 hr after light/dark treatment. This led us to perform a dissection experiment in which
198 shoots were separated from roots before the treatment (Fig. 3C, right). In these conditions
199 dissected shoots retain the same light response for *HTA9* APA as undissected seedlings,
200 but light has no effect on APA the in dissected roots. In the case of *RKH* APA, the light
201 effect was only observed in shoots independently of whether dissection was performed
202 after or before treatment (Fig. 3D). The negative control *FAD6* was unresponsive in
203 neither shoot nor roots in both dissection protocols (Fig. 3E). These results strongly
204 reinforce the evidence that light modulates APA through the chloroplast. The *HTA9*
205 dissection experiments suggest that a signal generated by light in the green tissues can
206 move to the roots to modulate APA in a similar way as in the green parts of the plant.

208 **Sugars and mitochondrial activity modulate the light effect on *HTA9* APA in the roots.**

209 In view if the evidence emerging from DCMU inhibition and dissection experiments, we
210 decided to further investigate the signaling from shoots to roots. Sugars are the main

photosynthates in terrestrial plants. These are generated in the green tissues and either metabolized in their cells or loaded into the phloem to feed non-photosynthetic tissues. In a previous study (10), it was shown that sucrose, the most important phloem-mobile sugar of *Arabidopsis* (23), is responsible of mediating AS changes in the roots triggered by light exposure of the shoots. More interestingly, externally applied sucrose solutions mimicked the effect of light on AS patterns in the root but had no major effect in the shoots. In agreement with AS results, incubation with 100 mM sucrose of seedlings subjected to the light/dark regime that were later dissected (Fig. 4A, POST) did not alter the light effect on *HTA9* APA in the shoots, but abolished it in the roots, in a way that mimics the effect of light in roots kept in the dark (note the height of the black bar that corresponds to roots treated with sucrose in Fig. 4A). In order to rule out any osmotic effect of sucrose we used equal concentrations of sorbitol as negative control. When dissection was performed before light/dark treatments (Fig. 4 B, PRE) incubation with sucrose had neither effect on the change in *HTA9* APA observed in the shoots nor on the lack of it observed in the roots.

Sugars have dual roles in plants, serving both as metabolic fuel and as signaling molecules (24). The signaling pathway in which sugars activate target of rapamycin (TOR) kinase and subsequently gene expression has been involved in the regulation of AS by sugars in *Arabidopsis* roots (10). Xiong et al. (25) revealed that glucose activation of TOR kinase in *Arabidopsis* meristems depended on glycolysis-mitochondria-mediated energy. This is consistent with findings that the respiratory chain/oxidative phosphorylation uncoupler 2,4-dinitrophenol (DNP) suppresses mammalian target of rapamycin (mTOR) activation in brain (26). Since DNP was also shown to abolish changes in root AS induced by light exposure of the green tissues (10), we investigated the effects of DNP on APA. In order to be able to assess DNP effects in the roots, experiments were carried out in the POST protocol, i.e., where excision was performed after the light/dark or drug treatments. Fig. 4C shows that DNP does not suppress the light effect on APA in the shoots, but completely abolishes the conspicuous light effect in the roots, in a way that imitates darkness. Control experiments with the APA unresponsive *FAD6* gene are shown in Suppl. Fig. 5.

Unlike AS, light control of *HTA9* APA is not affected by transcriptional elongation.

We used two experimental approaches to evaluate if the effect of light on *HTA9* APA was linked to changes in transcriptional elongation according to the kinetic coupling mechanism. To inhibit transcriptional elongation rate, we assessed an *Arabidopsis* mutant of the TFIIS transcription elongation factor. TFIIS is a factor required for RNAPII processivity that stimulates RNAPII to reassume elongation after pausing (27, 28). The *tfiis* mutant was achieved by replacing the key amino acids D290 and E291 of the acidic loop responsible for TFIIS stimulatory activity by alanines (28) as described in yeast (29), giving rise to a dominant negative phenotype showing a range of developmental defects, such as defective growth and serrated leaves. We have previously reported that in the *tfiis* mutant, transcript elongation is reduced and the change in *AtRS31* AS induced by light is abolished, mimicking the effects of darkness (7). While this is confirmed as a positive control in Fig. 5A, the effect of light on *HTA9* APA is not affected in the *tfiis* mutant (Fig. 5C). The negative control *FAD6* behaves similarly (Fig. 5E). On the activation side, we explored the use of a histone deacetylase inhibitor (HDAC) trichostatin A (TSA). By promoting higher histone acetylation and chromatin opening, HDACs have been proved to promote elongation in animal (30, 31) and plant cells (7), and in the latter, to mimic the chloroplast-mediated light effect on *AtRS31* AS. Again, while this is confirmed as positive control in Fig. 5B, the effect of light on *HTA9* APA is not affected by TSA treatment (Fig. 5D) and the negative control *FAD6* behaves similarly (Fig. 5F).

Changes in CPA factor mRNA abundance

In view of the fact that the light effect on APA does not seem to follow the kinetic coupling mechanism, we decided to investigate if the changes in APA in the light could be explained by changes in the expression of CPA factors measured at the mRNA level. We assessed three subunits of CPSF and two of CstF in shoots and roots obtained from excision performed after the light/dark treatment. All three CPSF subunits (100, 160 and 30) exhibited reduction in their steady state mRNA levels in the dark both in the shoots and in the roots (Fig. 6A). The effect is smaller in the roots (note smaller values for the white bars) suggesting that a signal generated in the green tissues reaches the roots but is not as effective as the signal in the shoots. In the case of CstF, while no changes in mRNA levels were observed for CstF64 neither in the shoots nor in the roots, CstF77 mRNA abundance drops in the dark in the shoots but is not affected in the roots (Fig. 6B). To

investigate if, as shown for APA in Figs. 2E-G, the chloroplast is the sensor for the upregulation of CPSF100, 160 and 30 and CstF77 by light, we treated whole seedlings with the electron transport chain inhibitor DCMU. Consistently, the DCMU treatment abolished the light/dark effect on mRNA levels for these subunits (Figs. 6C and 6D).

Discussion

We report here an unforeseen, potent regulation of APA by light/dark conditions in plants. It should be noted that the effects here described occur in an intact living organism (*Arabidopsis* seedlings) and under a fundamental physiological external signal (light). Although this investigation was initially inspired by our previous work on the effect of light on plant AS (6, 7, 10), our findings indicate that, despite some similarities, APA regulation has fundamental differences with respect to that of AS. Light/dark conditions affect 3'UTR APA of a substantial number of *Arabidopsis* genes (approximately 30%), which highlights its global impact. In the regulated genes light promotes the usage of both dPAS (75% of genes) and pPAS (the remaining 25%). Because of the predominance of dPAS usage and the robustness of its response, we decided to use the APA event of a gene encoding a histone H2A variant (*HTA9*) as a model to investigate the mechanisms involved. An ideal event to study would have been APA in the long non-coding RNA COOLAIR (AT5G01675) that is key in the control of flowering (4, 5), but unfortunately, COOLAIR is not expressed in seedlings (32), the paradigm system of our previous work on splicing and the present work on polyadenylation.

Similar to the regulation of AS, the light effect on APA is not mediated by phytochrome nor cryptochrome photoreceptors but by the chloroplast. The correct functioning of the photosynthetic electron transport chain is necessary since its inhibition by DCMU abolishes the effect. This is consistent with the fact that the light effect is observed both in the green tissues and the roots as long as their connection is not interrupted by dissection. Since the roots have no functional chloroplasts, the absence of effect in isolated roots indicates that a signaling molecule must travel through the phloem from shoots to roots. We identify this molecule as the main photosynthate sucrose because it fully mimics the effect of light in the roots when seedlings are kept in the dark. Most interestingly, we found that root mitochondrial activity is necessary in the roots since the respiratory chain/oxidative phosphorylation uncoupler 2,4-dinitrophenol (DNP)

abolishes the light effect fully mimicking the effect of darkness in the roots when seedlings are kept in the light. The opposite effects of sucrose and DNP allowed us to hypothesize that, like in AS regulation (10), the APA mechanism could involve sugar activation of TOR kinase, previously shown to depend on mitochondria-generated energy (25). Notably, 3'UTR size regulation through APA is increasingly associated with cell metabolism in mammalian cells (33). Whether plants and metazoans share similar mechanisms in energy-mediated APA regulation is to be examined in the future.

Two experiments of different nature indicate that, unlike AS, the effect on APA is not linked to the control of RNAPII elongation. The light effect on APA is neither abolished in a mutant of the transcription elongation factor TFIIS nor in seedling treated with the histone deacetylase inhibitor TSA, previously shown to cause higher transcript elongation due to histone acetylation and chromatin relaxation. Important controls show that both the TFIIS mutant and TSA treatment suppress the light effect on AS, the former mimicking darkness and the latter mimicking light as published (7). In search for an alternative mechanism we wondered if light would be affecting the expression of cleavage/polyadenylation factors. We found indeed that light upregulates mRNA levels for CPSF100, 160 and 30 and CstF77 and that this increase is suppressed by treatment with DCMU. In particular, *Arabidopsis* CPSF30 has been reported to interact with other polyadenylation factors like FIPS5, CPSF100, and CstF77 and can be considered as a central hub in the protein-protein interaction network of plant polyadenylation complex subunits (34-36). Indeed, *Arabidopsis* PAS selection, is determined by the presence or absence of this factor (37). The precise mechanism by which CPSF30 controls *Arabidopsis* APA remains elusive. However, the current model for canonical plant PA signaling involves three discreet cis-elements collaborating to the effective 3' end formation of mRNAs: (i) the far upstream element (FUE), consisting of an extended U and G-rich region situated more than 50 nucleotides upstream from the PAS; (ii) the near-upstream element (NUE), a 6 to 10 nucleotide-A-rich region situated 10 to 30 nucleotides upstream from the PAS; and (iii) the cleavage element (CE), a U-rich region centered around the PAS. Although most functioning aspects of these elements have not yet been characterized, CPSF30 was shown to play a role in the functioning of the NUE because PAS that are only used in the wild type, but not the CPSF30-deficient mutant, possess the

characteristic A-rich NUE signature, while PAS used only in the CPSF30-deficient mutant lack this signature (37). In mammalian models, there is accumulated evidence that a variety of core CPA factors regulate APA (38-43). Moreover, a model was proposed in which the choice of PASs is dependent on the strengths of the cis-elements present in the PAS and the relative usage dependent on the competition between PASs for the available CPA factors (44). Upon loss or diminishment of core CPA factor(s) the relative strengths of all PASs decrease. However, any factor that would increase the window of opportunity for the CPA factors to recognize pPAS would lead to a shift toward the pPAS, such as RNAPII pausing, slowing RNAPII elongation, or increasing the distance between pPAS and dPAS. Even though there is still no evidence of this model in plants, it is an interesting putative scenario to see the light as a regulator of CPA factors abundance and PAS selection. Interestingly, we found that the distance between two APA sites is important for both types of APA isoform regulation, with 3'UTR lengthening and those with 3'UTR shortening (Suppl. Fig. 6).

As 3'UTRs are very rich in regulatory elements, the physiological consequences of the changes in APA decisions in response to light could be wide and affect many pathways. In particular, the gene with the event chosen as output, *HTA9*, encodes a histone H2A variant (H2A.Z). This variant has been associated with environmental responses to temperature and stress (45, 46). On the other hand, it was shown that H2A.Z-containing nucleosomes wrap DNA more tightly than canonical H2A nucleosomes, which may affect RNAPII elongation and, in turn, AS (7). In any case, further work will be necessary to better understand the biological roles of the novel link between plant APA and light reported here.

364 **Materials and methods**

365

366 **Plant material, growth conditions and drug treatments.** The *Arabidopsis* Columbia
 367 ecotype (Col 0) and Landsberg erecta (Ler) were used as wild type, according to the
 368 mutants assessed. Seeds were stratified for three days in the dark at 4°C and then
 369 germinated on Murashige and Skoog 0.5x (MS) medium containing 1% (w/v) agar.
 370 Standard treatment protocol: *A. thaliana* seedlings were grown in Petri dishes with MS
 371 medium at a constant temperature of 22°C under constant white light provided by
 372 fluorescent tubes with an intensity of irradiance between 70 and 100 $\mu\text{mol}/\text{m}^2\text{sec}$ (20
 373 seeds per dish) for a period of 2 weeks and then transferred for 48hr to darkness. After
 374 this period, seedlings were kept in the dark or transferred to the light for 6 hours. At the
 375 end of this light/ dark treatment, seedlings were harvested in liquid nitrogen. For all
 376 pharmacological treatments, drugs were added 1 hour prior to the light/ dark treatment.
 377 For the DCMU [3-(3,4-dichlophenyl)-1,1-dimethylurea; Sigma] treatment seedlings were
 378 transferred to 6-wells plates and incubated in 20 μM DCMU. For the sucrose treatment,
 379 plants on agar plates were covered with 10 mL of 100 mM sucrose or sorbitol, used as
 380 osmotic control. Vacuum was applied for five minutes to facilitate drug uptake. For the
 381 DNP (2,4-dinitrophenol; Sigma) treatment, the drug was added up to 20 μM , using
 382 ethanol was as vehicle control.

383

384 **3'READS and PAS identification.** Total RNA extraction of seedlings was carried out by
 385 using the RNeasy Plant Mini Kit (Qiagen) following manufacturer's instructions. The 3'
 386 region extraction and deep sequencing (3'READS) method were described previously (11)
 387 Libraries were sequenced on an Illumina HiSeq machine (2x150 bases) at Admera Health
 388 (New Jersey, USA). 3'READS data were analyzed as previously described (17, 47, 48).
 389 Briefly, after adapter sequence removal, reads were mapped to the *Arabidopsis thaliana*
 390 reference genome (TAIR10) by using the Bowtie2 program (49). Sequences with a
 391 mapping quality (MAPQ) score ≥ 10 and ≥ 2 nongenomic Ts at the 5' end after alignment
 392 were considered as PAS-supporting (PASS) reads and were used for subsequence
 393 analysis. Identified cleavage sites within 24 nt from one another were clustered into PAS
 394 clusters (50). 3'READS data statistics are shown in [Suppl. Table 1](#). Gene annotation was
 395 based on the TAIR10 and the Ensembl databases.

396

397 **Analysis of APA isoforms.** Analysis of 3'UTR APA was based on the top two most
 398 expressed APA isoforms of a gene, which were named proximal and distal PAS isoforms.
 399 To eliminate spurious APA isoforms, we further required that the number of PASS reads
 400 for the minor APA isoform (the second most expressed) to be above 5% of all isoforms
 401 combined. For 3'UTR APA analysis, only the APA sites in the last exon were used. The
 402 relative expression (RE) of two PAS isoforms, e.g., pPAS and dPAS, was calculated by
 403 $\log_2(\text{RPM})$ of dPAS vs. pPAS, where RPM was reads per million PASS reads. Relative
 404 expression difference (RED) of two isoforms in two comparing samples was based on the
 405 difference in RE for the two isoforms between the two samples. DEXSeq was used to
 406 derive statistically significant APA events ($\text{FDR} < 0.05$) (51).

407

408 **Gene expression analysis.** The DESeq method (R Bioconductor) (52) was used to analyze
 409 gene expression changes. Significantly regulated genes are those with adjusted $p < 0.05$
 410 and fold change > 1.2 . PAS reads of each gene were summed to represent gene
 411 expression.

412

413 **PAS motif analysis.** The PROBE program was used to examine sequence motifs around
 414 the PASs (53). The genomic region surrounding each PAS was divided into four
 415 subregions: -60 to -31 nt, -30 to -1 nt, +1 to +30 nt, and +31 to +60 nt. The observed
 416 frequency of each k-mer in a subregion was enumerated and compared with the expected
 417 frequency based on randomized sequences of the region. Randomization was carried out
 418 by using the first-order Markov chain model (53). The enrichment score (Z-score) was
 419 calculated based on the difference between the observed and expected frequencies. The
 420 Fisher's exact test was used to determine significance.

421

422 **Gene Ontology analysis.** The GOstats hypergeometric test (R Bioconductor) was used to
 423 test for significant association of genes with gene ontology (GO) terms. GO annotation for
 424 *Arabidopsis thaliana* was obtained from org.At.tair.db (R Bioconductor). GO terms
 425 associated with more than 1,000 genes were considered too generic and were discarded. To
 426 remove redundant GO terms, each reported GO term was required to have at least 25% of
 427 the genes that were not associated with another term with a more significant p -value.

428

429 **APA analysis by RT-qPCR.** Seedlings were grown following specifications given in Plant
430 material, growth conditions and drug treatments. Samples were harvested and total RNA
431 was purified using TRIzol (Invitrogen). 500 ng of RNA were further used to synthesize
432 cDNA with MMLV-RT enzyme (Invitrogen) and oligo-dT as primer following the
433 manufacturer's instructions. Synthesized cDNAs were amplified with 1.5 U of Taq DNA
434 polymerase (Invitrogen) and SYBR Green (Roche) using the Eppendorf Mastercycler
435 Realplex. Primer sequences for qRT-PCR are available in [Suppl. Table 3](#). RT-qPCR
436 experiments were quantified with $n \geq 3$, where n = about 25-30 *Arabidopsis* seedlings
437 growing in one Petri dish. Changes considered significant show differences with a p value
438 < 0.05 (two-tailed Student's t test).

439

440 **Acknowledgments**

441 We thank E. Martín, V. Roselló and V. Buggiano for technical assistance. This work was
442 supported by grants from the Agencia Nacional de Promoción Científica y Tecnológica of
443 Argentina (PICT-2019-862), the Universidad de Buenos Aires (UBACYT
444 20020170100046BA) and the Lounsbery Foundation (USA) to ARK and NIH grants
445 GM084089 and GM129069 to BT. M.A.G-H, E.P. and A.R.K. are career investigators and
446 M.G.K. received a fellowship from the Consejo Nacional de Investigaciones Científicas y
447 Técnicas of Argentina (CONICET).

448

449 **References**

- 450 1. J. Yang, Y. Cao, L. Ma. Co-transcriptional RNA processing in plants: exploring from the
451 perspective of polyadenylation. *Int J. Mol. Sci.* **22**, 3300 (2021).
- 452 2. B. Tian, J. L. Manley. Alternative polyadenylation of mRNA precursors. *Nat. Rev. Mol.*
453 *Cell Biol.* **18**, 18-30 (2017).
- 454 3. M. Cyrek *et al.* Seed dormancy in *Arabidopsis* is controlled by alternative
455 polyadenylation of DOG1. *Plant Physiol.* **170**, 947-955 (2016).
- 456 4. F. Liu, S. Marquardt, C. Lister, S. Swiezewski, C. Dean. Targeted 3' processing of
457 antisense transcripts triggers *Arabidopsis* FLC chromatin silencing. *Science* **327**, 94-97
458 (2010).

- 459 5. C. Xu, *et al.* R-loop resolution promotes co-transcriptional chromatin silencing. *Nat.*
460 *Commun.* **12**, 1790 (2021).
- 461 6. E. Petrillo *et al.* A chloroplast retrograde signal regulates nuclear alternative splicing.
462 *Science* **344**, 427-430 (2014).
- 463 7. M. A. Godoy Herz *et al.* Light regulates plant alternative splicing through the control of
464 transcriptional elongation. *Mol. Cell* **73**, 1066-1074 (2019).
- 465 8. A. R. Kornblihtt *et al.* Alternative splicing: a pivotal step between eukaryotic
466 transcription and translation. *Nat. Rev. Mol. Cell Biol.* **14**, 153-165 (2013).
- 467 9. T. Saldi, M. A. Cortazar, R. M. Sheridan, D. L. Bentley. Coupling of RNA polymerase II
468 transcription elongation with pre-mRNA splicing. *J. Mol. Biol.* **428**, 69-81 (2016).
- 469 10. S. Riegler *et al.* Light regulates alternative splicing outcomes via the TOR kinase
470 pathway. *Cell Rep.* **36**, 109676 (2021).
- 471 11. D. Zheng, X. Liu, B. Tian. 3'READS+, a sensitive and accurate method for 3' end
472 sequencing of polyadenylated RNA. *RNA* **22**, 1631-1639 (2016).
- 473 12. A. Sherstnev *et al.* Direct sequencing of *Arabidopsis thaliana* RNA reveals patterns of
474 cleavage and polyadenylation. *Nat. Struct. Mol. Biol.* **8**, 845-852 (2012).
- 475 13. S. Zhu *et al.* PlantAPAdb: A Comprehensive Database for Alternative Polyadenylation
476 Sites in Plants. *Plant Physiol.* **182**, 228-242 (2020).
- 477 14. X. Liu *et al.* Comparative analysis of alternative polyadenylation in *S. cerevisiae* and *S.*
478 *pombe*. *Genome Res.* **27**, 1685-1695 (2017).
- 479 15. X. Liu *et al.* Transcription elongation rate has a tissue-specific impact on alternative
480 cleavage and polyadenylation in *Drosophila melanogaster*. *RNA* **23**, 1807-1816 (2017).
- 481 16. C. H. Jan, R. C. Friedman, J. G. Ruby, D. P. Bartel. Formation, regulation and evolution
482 of *Caenorhabditis elegans* 3'UTRs. *Nature* **469**, 97-101 (2011).
- 483 17. M. Hoque *et al.* Analysis of alternative cleavage and polyadenylation by 3' region
484 extraction and deep sequencing. *Nat. Methods.* **10**, 133-139 (2013).
- 485 18. A. Nagatani, J.W. Reed, J. Chory. Isolation and initial characterization of *Arabidopsis*
486 mutants that are deficient in phytochrome A. *Plant Physiol.* **102**, 269-277 (1993).
- 487 19. M. Koornneef, E. Rolf, C.J.P. Spruit. Genetic control of light-inhibited hypocotyl
488 elongation in *Arabidopsis thaliana* (L.) Heynh. *Z. Pflanzenphysiol.* **100**, 147-160 (1980).
- 489 20. M. Koornneef, C.J. Hanhart, J.H. van der Veen. A genetic and physiological analysis of
490 late flowering mutants in *Arabidopsis thaliana*. *Mol. Genet. Genom.* **229**, 57-66 (1991).

- 491 21. M. Ruckle, L. Burgoon, L. Lawrence, C. Sinkler, R. Larkin. Plastids are major regulators
492 of light signaling in Arabidopsis. *Plant Physiol.* **159**, 366-390 (2012).
- 493 22. A. Khandelwal, T. Elvitigala, B. Ghosh, R. S. Quatrano. Arabidopsis transcriptome
494 reveals control circuits regulating redox homeostasis and the role of an AP2 transcription
495 factor. *Plant Physiol.* **148**, 2050-2058 (2008).
- 496 23. K. Wippel, N. Sauer. Arabidopsis SUC1 loads the phloem in suc2 mutants when
497 expressed from the SUC2 promoter. *J. Exp. Bot.* **63**, 669-679 (2012).
- 498 24. J. Wind, S. Smeekeens, J. Hanson. Sucrose: metabolite and signaling molecule.
499 *Phytochem.* **71**, 1610-1614 (2010).
- 500 25. Y. Xiong *et al.* Glucose-TOR signaling reprograms the transcriptome and activates
501 meristems. *Nature* **496**, 181-186 (2013).
- 502 26. D. Liu *et al.* The mitochondrial uncoupler DNP triggers brain cell mTOR signaling
503 network reprogramming and CREB pathway up-regulation. *J. Neurochem.* **134**, 677-692
504 (2015).
- 505 27. R. N. Fish, C. M. Kane. Promoting elongation with transcript cleavage stimulatory
506 factors. *Biochim. Biophys. Acta* **1577**, 287-307 (2002).
- 507 28. J. Dolata *et al.* NTR1 is required for transcription elongation checkpoints at alternative
508 exons in Arabidopsis. *EMBO J.* **34**, 544-58 . (2015).
- 509 29. S. Sigurdsson, A. B. Dirac-Svejstrup, J. Q Svejstrup. Evidence that transcript cleavage is
510 essential for RNA polymerase II transcription and cell viability. *Mol. Cell* **38**, 202-210
511 (2010).
- 512 30. G. Dujardin *et al.* How slow RNA polymerase II elongation favors alternative exon
513 skipping. *Mol. Cell* **54**, 683-690 (2014).
- 514 31. L. E. Marasco *et al.* Counteracting chromatin effects of a splicing-correcting antisense
515 oligonucleotide improves its therapeutic efficacy in spinal muscular atrophy. *Cell* **185**,
516 2057-2070 (2022).
- 517 32. A. V. Klepikova, A. S. Kasianov, E. S Gerasimov, M. D. Logacheva, A. A. Penin. A high
518 resolution map of the Arabidopsis thaliana developmental transcriptome based on RNA-
519 seq profiling. *Plant J.* **88**, 1058-1070 (2016).
- 520 33. Q. Zhang, B. Tian. The emerging theme of 3'UTR mRNA isoform regulation in
521 reprogramming of cell metabolism. *Biochem. Soc. Trans.* **28**, 1111-1119 (2023).

522 34. K. P. Forbes, B. Addepalli, A. G. Hunt. An Arabidopsis Fip1 homolog interacts with
523 RNA and provides conceptual links with a number of other polyadenylation factor
524 subunits. *J. Biol. Chem.* **281**, 176–186 (2006).

525 35. R. Q. Xu *et al.* The 73 kDa subunit of the cleavage and polyadenylation specificity
526 factor (CPSF) complex affects reproductive development in Arabidopsis. *Plant Mol. Biol.*
527 **61**, 799–815 (2006).

528 36. S. A. Bell, A. G. Hunt. The Arabidopsis ortholog of the 77 kDa subunit of the cleavage
529 stimulatory factor (AtCSTF-77) involved in mRNA polyadenylation is an RNA-binding
530 protein. *FEBS Lett.* **584**, 1449–1454 (2010).

531 37. P. E. Thomas *et al.* Genome-wide control of polyadenylation site choice by CPSF30 in
532 Arabidopsis. *Plant Cell* **24**, 4376–4388 (2012).

533 38. G. Martin, A. R. Gruber, W. Keller, M. Zavolan. M. Genome-wide analysis of pre-
534 mRNA 3' end processing reveals a decisive role of human cleavage factor I in the
535 regulation of 3' UTR length. *Cell Rep.* **1**, 753–763 (2012).

536 39. T. Kubo, T. Wada, Y. Yamaguchi, A. Shimizu, H. Handa. Knock-down of 25 kDa
537 subunit of cleavage factor Im in hela cells alters alternative polyadenylation within 3'-
538 UTRs. *Nucleic Acids Res.* **34**, 6264–6271 (2006).

539 40. C. P. Masamha *et al.* CFIm25 links alternative polyadenylation to glioblastoma tumour
540 suppression. *Nature* **510**, 412–416 (2014).

541 41. M. Jenal *et al.* The poly(A)-binding protein nuclear 1 suppresses alternative cleavage
542 and polyadenylation sites. *Cell* **149**, 538–553 (2012).

543 42. C. Yao *et al.* Transcriptome-wide analyses of CSTF64-RNA interactions in global
544 regulation of mRNA alternative polyadenylation. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 18773–
545 18778 (2012).

546 43. W. Li *et al.* Systematic profiling of poly(A)+ transcripts modulated by core 3' end
547 processing and splicing factors reveals regulatory rules of alternative cleavage and
548 polyadenylation. *PLoS Genet.* **11**: e1005166 (2015).

549 44. P. Tang, Y. Zhou. Alternative polyadenylation regulation: insights from sequential
550 polyadenylation. *Transcription* **13**, 89-95 (2022).

551 45. S. V. Kumar, P. A. Wigge. H2A.Z-containing nucleosomes mediate the thermosensory
552 response in Arabidopsis. *Cell* **140**, 136-47 (2010).

553 46. W. Sura *et al.* Dual Role of the Histone Variant H2A.Z in Transcriptional regulation of

554 stress-response genes. *Plant Cell*. 29, 791-807 (2017).

555 47. D. Zheng, B. Tian. Polyadenylation site-based analysis of transcript expression by
556 3'READS. *Methods Mol. Biol.* **1648**, 65-77 (2017).

557 48. D. Zheng *et al.* Cellular stress alters 3'UTR landscape through alternative
558 polyadenylation and isoform-specific degradation. *Nat. Commun.* **9**, article number
559 2268 (2018).

560 49. B. Langmead, S. L. Salzberg. Fast gapped-read alignment with Bowtie 2. *Nat. Methods*
561 **9**, 357-359 (2012).

562 50. B. Tian, J. Hu, H. Zhang, C. S. Lutz. A large-scale analysis of mRNA polyadenylation
563 of human and mouse genes. *Nucleic. Acids Res.* **33** 201-212 (2005).

564 51. S. Anders, A. Reyes, W. Huber. Detecting differential usage of exons from RNA-seq
565 data. *Genome Res.* **22**, 2008-2017 (2012).

566 52. S. Anders, W. Huber. Differential expression analysis for sequence count data. *Genome*
567 *Biol.* **11**, R106 (2010).

568 53. J. Hu, C. S. Lutz, J. Wilusz, B. Tian. Bioinformatic identification of candidate cis-
569 regulatory elements involved in human mRNA polyadenylation. *RNA* **10**, 1485-1493.
570 (2005).

571

572 Figure Legends

573

574 **Figure 1. Light/dark conditions elicit widespread APA changes in *Arabidopsis thaliana*.**

575 **A.** Protocol scheme of the light/dark regime used in this study. Total RNA from seedlings
576 was subject to 3'READS for gene expression and APA analyses. **B.** Distribution of
577 identified PASs (left) and PAS reads (right) in different regions of the plant genome. **C.**
578 Different types of APA genes identified in this study. **D.** Top, schematic of 3'UTR APA
579 isoform; bottom, 3'UTR sizes for mRNAs of genes without 3'UTR APA (single 3'UTR) and
580 mRNAs with the longest or shortest 3'UTRs of genes with 3'UTR APA. The median value
581 for each group is indicated. **E.** Diagram showing 3'UTR APA analysis. The two most
582 abundant APA isoforms per gene were selected for comparison, which are named
583 proximal PAS (pPAS) and distal PAS (dPAS) isoforms, respectively. The distance between
584 the two PASs is considered alternative 3'UTR (aUTR). **F.** Scatterplot showing genes with
585 pPAS and dPAS isoform abundance differences between light- and dark-treated
586 seedlings. Results represent analyses of 12,873 genes in three biological replicates. Genes
587 with significantly (FDR < 0.05, DEXseq analysis) higher or lower abundance of pPAS
588 isoforms in light vs. dark conditions are shown in blue and red respectively.

589

590 **Figure 2. The light effect on APA is sensed by the chloroplast (genetic and biochemical**

591 **evidence).** **A.** 3'READS data for two *Arabidopsis* genes with opposite APA changes. Top: a
592 representative event [*HTA9* gene (AT1G52740), reads in red] with higher usage of dPAS in
593 the light. Bottom: a representative event [*RKH* gene (AT5G15270), reads in blue] with
594 higher usage of pPAS in the light. For each APA event two pairs of primers were designed
595 to validate the APA changes using RT-qPCR: amplicon dPAS (dark orange) only exists if
596 the dPAS is used (long isoform); amplicon cod is common to all isoforms in the upstream
597 coding region. Changes in APA are quantified as ratios of dPAS/cod amplicons relative
598 mRNA expression levels for every gene. **B-D.** Light is not sensed by photoreceptors. APA
599 response to light/dark in different *Arabidopsis* phytochrome and cryptochrome mutant
600 genotypes in a Landsberg erecta background (wt, Ler). Three selected genes are shown:
601 *HTA9*, whose APA events increases its dPAS usage in the light (A), *RKH*, whose APA
602 event diminishes its dPAS usage in the light (B) and *FAD6*, whose APA event is not
603 affected by the light/dark conditions, and serves as a negative control. **E-G.** Effect of the

604 photosynthetic electron transfer chain inhibitor DCMU on the light/dark effect on APA
 605 events of the *HTA9* (E), *RKH* (F) and *FAD6* (G) genes. Seedlings were grown in constant
 606 light, transferred to darkness for 48 hr. and then treated with 20 μ M DCMU during a 6-hr.
 607 light/dark further incubation. RT-qPCR experiments were quantified with $n \geq 3$, where n
 608 = ~ 25 -30 *Arabidopsis* seedlings growing in one Petri dish. White and black bars represent
 609 light and dark treatments respectively. Changes considered significant show differences
 610 with a p value < 0.05 (two-tailed Student's t test). *** = $p < 0.001$; ** = $p < 0.01$; * = $p < 0.05$;
 611 NS (not significant) = $p > 0.05$.

612

613 **Figure 3. The light effect on APA is sensed by the photosynthetic tissues (anatomical**
 614 **evidence). A-B.** Schemes for dissections of *Arabidopsis* seedlings performed after (post-) or
 615 before (pre-) the light/dark treatment. **C-E.** APA isoform analysis in green tissue (shoots)
 616 and in the roots of post- (left) and pre- dissection (right) light/dark treatments of the
 617 *HTA9* (C), *RKH* (D) and *FAD6* (E) genes. Bar colors and RT-qPCR conditions were as in
 618 Figure 2.

619

620 **Figure 4. Sugars and mitochondrial activity modulate the light effect on APA in the**
 621 **roots. HTA9 gene APA isoform analysis in green tissues (shoots) and roots of Arabidopsis**
 622 **seedlings dissected post- (A and C) and pre- (B) light/dark treatments. A and B.**
 623 Incubations were performed with 100 mM sucrose or sorbitol (negative osmotic control)
 624 as indicated. C. Post-treatment excision protocol. Incubations were performed with 20 μ M
 625 of the mitochondrial uncoupler dinitrophenol (DNP) or vehicle (ctrl.) as indicated. Bar
 626 colors and RT-qPCR conditions were as in Figure 2.

627

628 **Figure 5. AS and APA respond differently to factors affecting transcript elongation.**
 629 Effects of genetic disruption (*tftis* mutant) of the transcription elongation factor TFIIS (A,
 630 C and E) and of treatment with the histone deacetylase inhibitor TSA (B, D and F) on AS
 631 of the *Arabidopsis RS31* gene (A and B) and on APA of the *HTA9* (C and D) and *FAD6* (E
 632 and F) genes. Bar colors and RT-qPCR conditions were as in Figure 2.

633

634 **Figure 6. Light upregulates mRNA levels of cleavage/polyadenylation factors through**
 635 **the chloroplast. A and B.** RT-qPCR quantification of mRNA levels encoding subunits of

636 CPSF (A) and CstF (B) in *Arabidopsis* shoots and roots obtained in a post-light/dark
 637 treatment excision experiment. C and D. Effect of the photosynthetic electron transfer
 638 chain inhibitor DCMU on the light/dark effect on cleavage/polyadenylation factor
 639 mRNA levels in whole seedlings for CPSF (C) and CstF (D). Factor mRNA levels were
 640 relativized to mRNA levels of protein phosphatase 2A (PP2A). Bar colors and RT-qPCR
 641 conditions were as in Figure 2.

Fig. 1

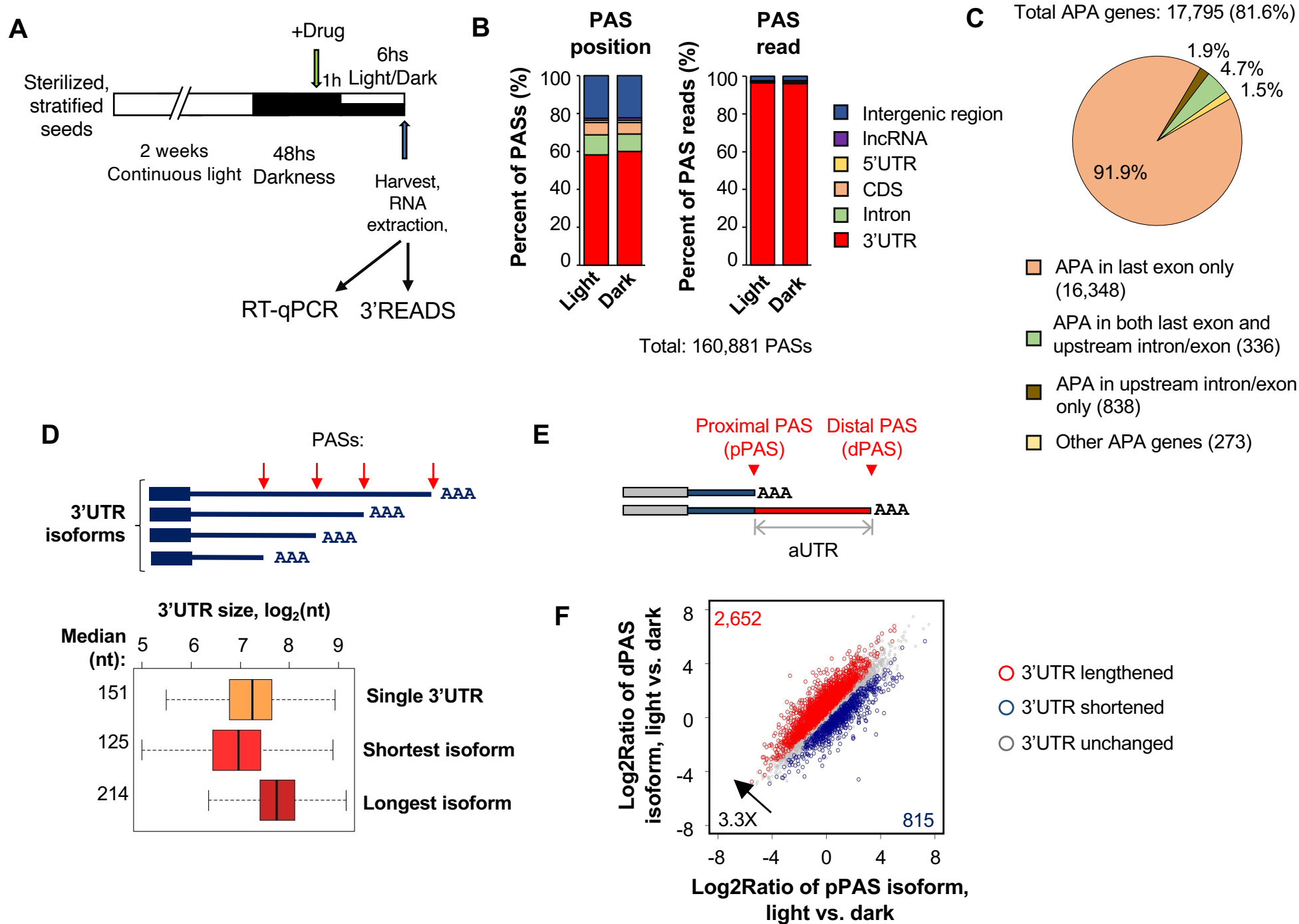


Fig. 2

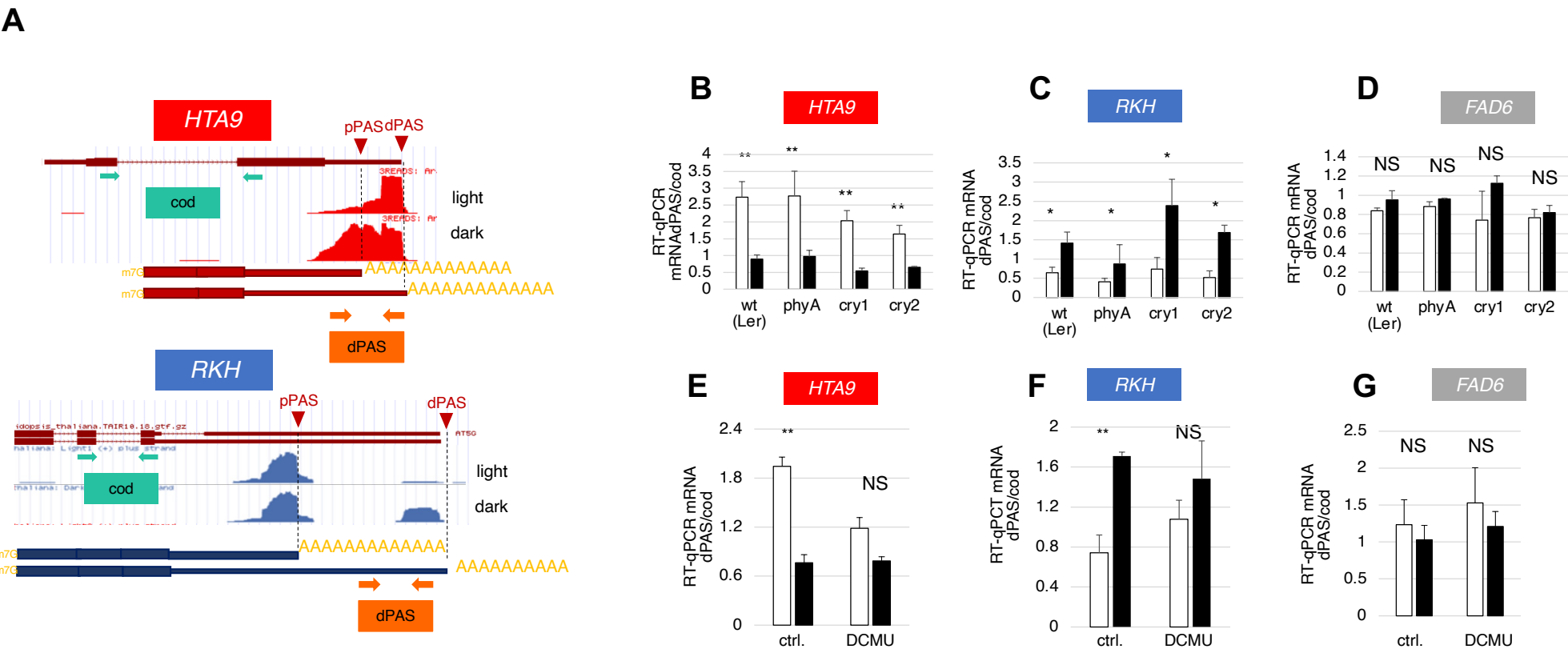


Fig. 3

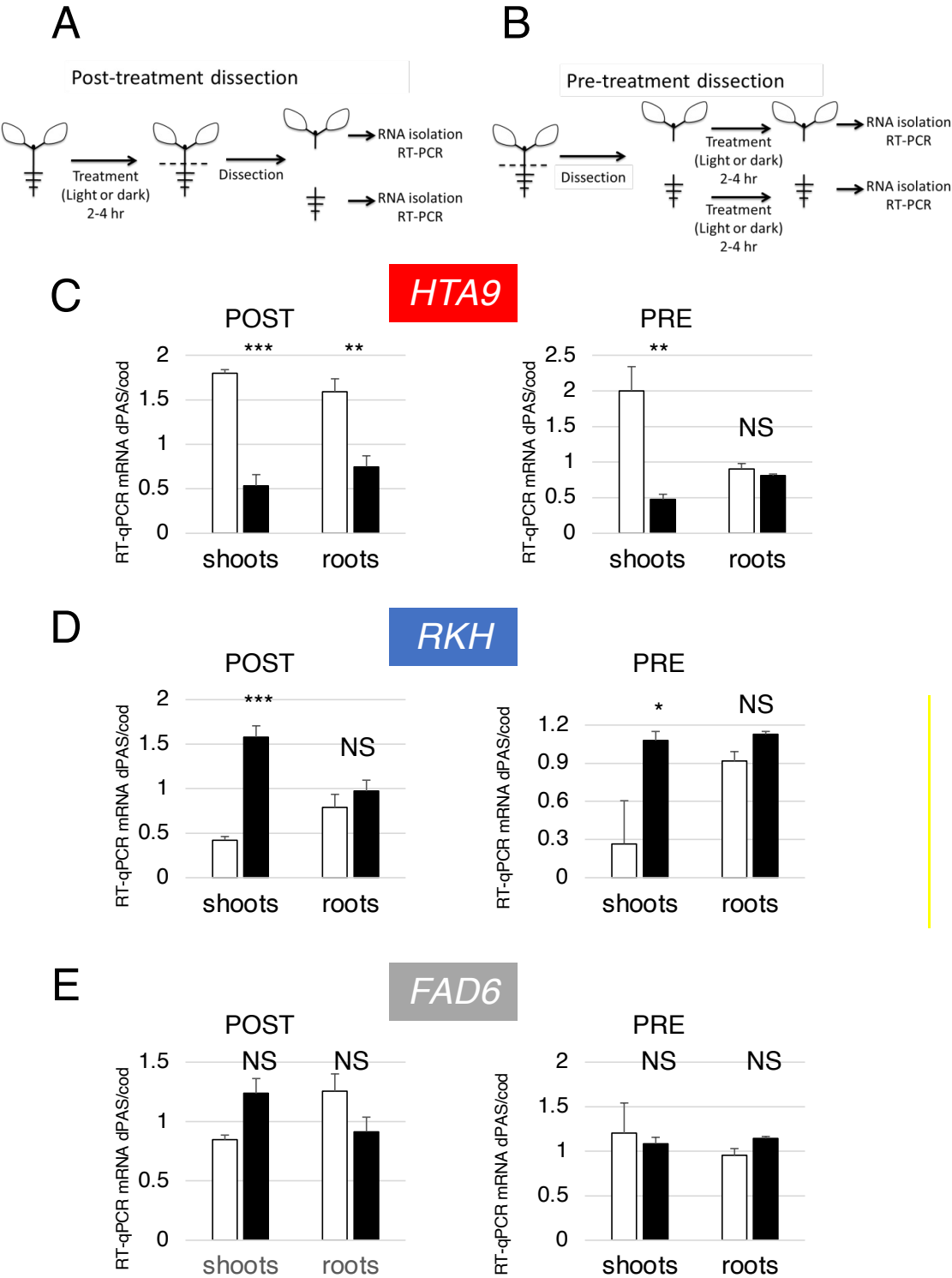


Fig. 4

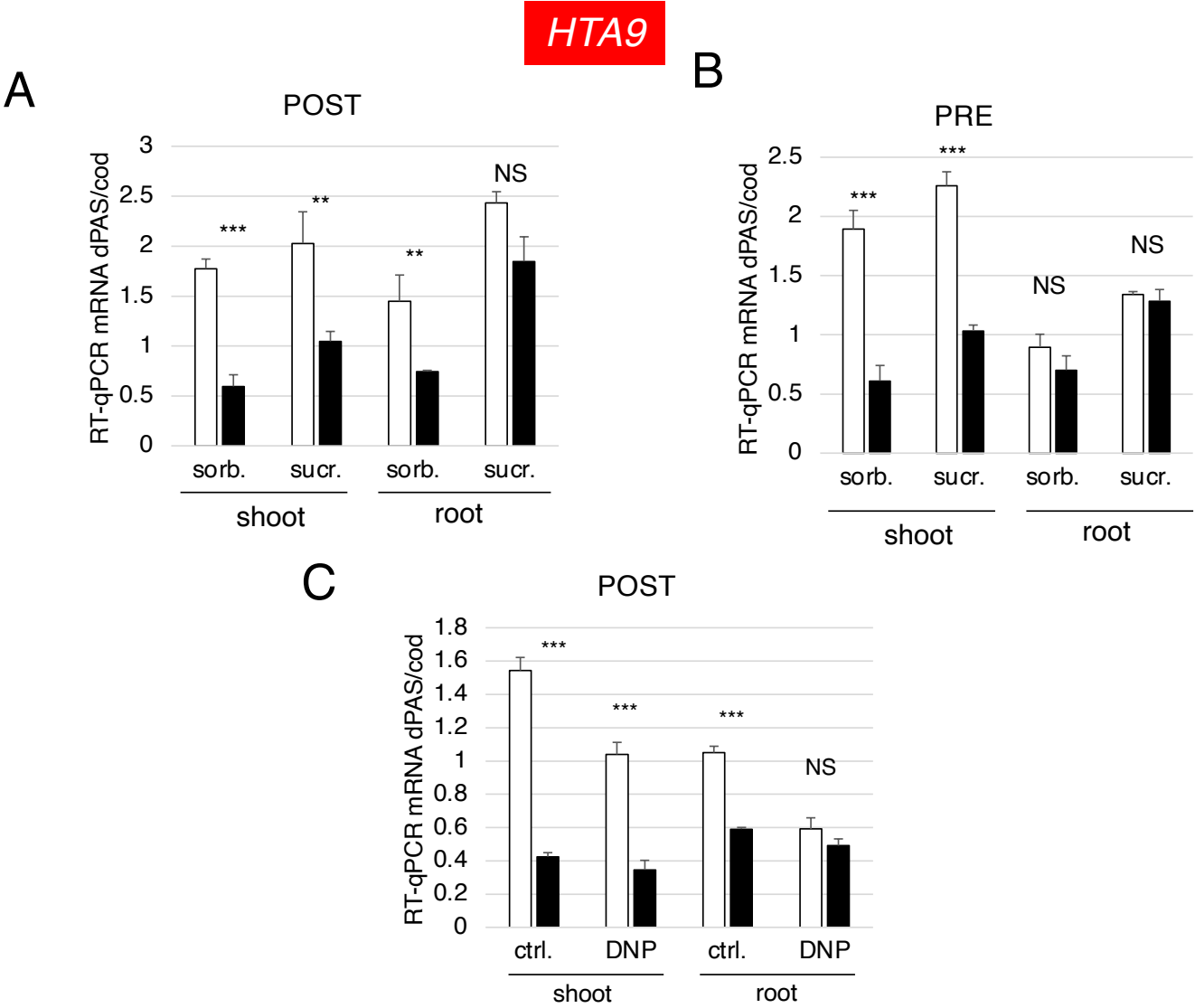


Fig. 5

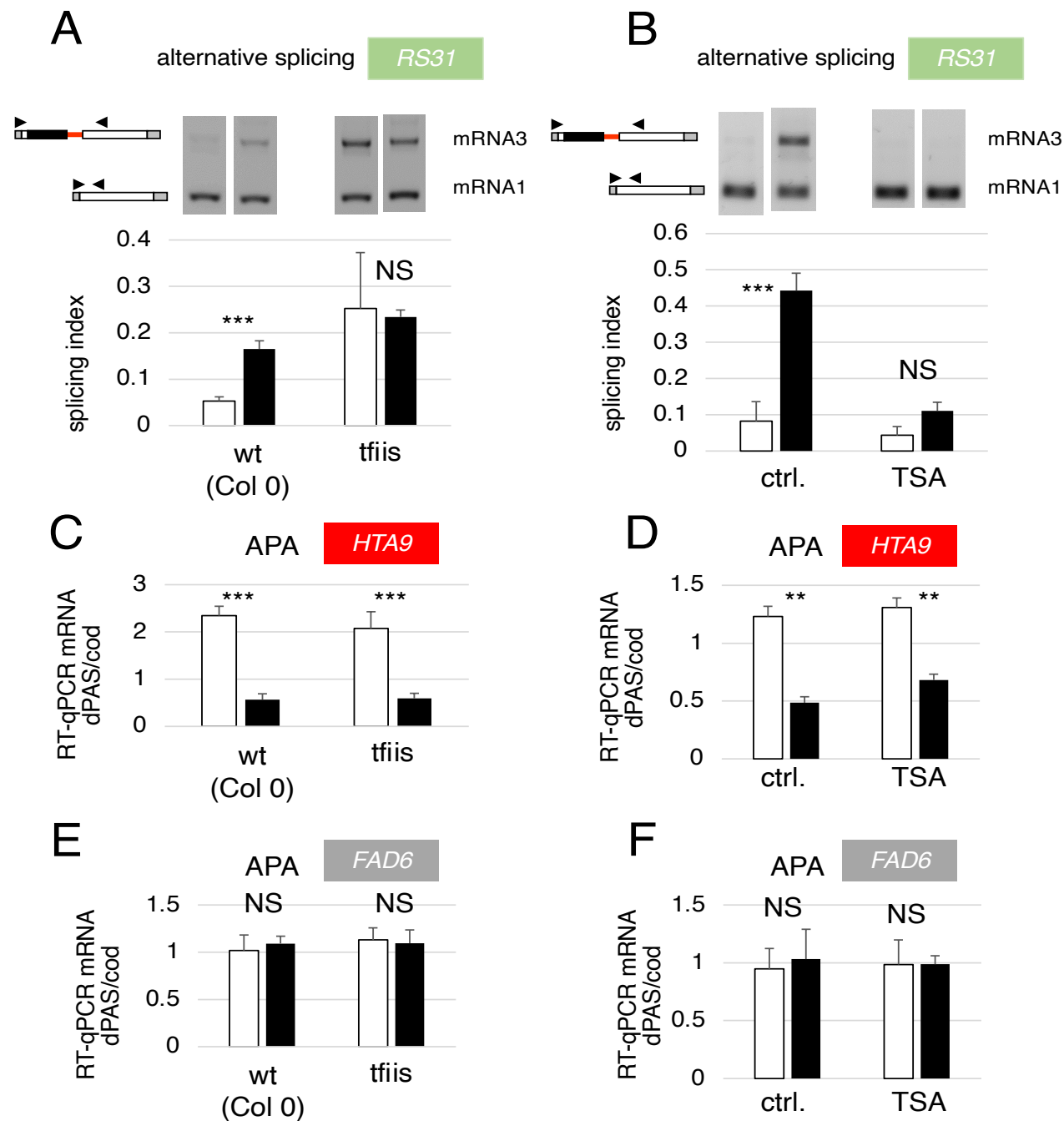


Fig. 6

mRNA abundance

