Enhancers regulate polyadenylation site cleavage and control 3′UTR isoform expression

Buki Kwon[1]*, Mervin M. Fansler[1,2]*, Neil D. Patel[1]*, Shih-Han Lee[1], Jihye Lee[1], Weirui Ma[1], and Christine Mayr[1,2,3]†

[1]Cancer Biology and Genetics Program, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA

[2]Tri-Institutional Training Program in Computational Biology and Medicine, Weill Cornell Graduate College, New York, 10021, NY, USA

*, These authors contributed equally

[3]Lead Contact

†Correspondence: mayrc@mskcc.org

Christine Mayr
Sloan Kettering Institute
1275 York Ave, Box 303
New York, NY 10065
Phone: 646-888-3115

Summary

Enhancers are DNA elements that increase gene expression. mRNA production is determined by transcript production and polyadenylation site (PAS) cleavage activity. We established an assay to measure enhancer-dependent PAS cleavage activity in human cells because PAS cleavage may control alternative 3′UTR isoform expression. We found that enhancers are widespread regulators of cell type-specific differences in 3′UTR isoform usage as they consistently increase cleavage of proximal and weak PAS. Half of tested transcription factors exclusively regulated PAS cleavage without affecting transcript production, whereas co-activators changed both parameters. Deletion of an endogenous enhancer of *PTEN* did not change gene-level mRNA or protein abundance but affected alternative mRNA isoform expression. Enhancer deletion prevented 3′UTR shortening and altered the intrinsic PTEN lipid phosphatase activity without changing the amino acid sequence of PTEN. Taken together, in addition to controlling transcript production, enhancers also regulate PAS cleavage, thus changing 3′UTR isoform usage and protein activity.

Introduction

Comparative genome analyses showed that organismal complexity scales with gene regulation (Levine et al., 2014). It is thought that the increased gene regulatory capacity is largely accomplished by enhancers (Levine et al., 2014). According to the original definition, enhancers are sequences that increase the expression of a reporter gene (Banerji et al., 1981; Gasperini et al., 2020). Currently, it is thought that enhancers mostly affect transcript production (Gasperini et al., 2020), but mRNA processing is essential for the generation of mature mRNAs and this step includes 3′ end cleavage and polyadenylation. Therefore, when disregarding the contribution of mRNA stability, mRNA production of unspliced transcripts is largely determined by the number of produced transcripts and by the cleavage activity of the polyadenylation site (PAS).

Although PAS cleavage activity is difficult to measure directly, its contribution to mRNA expression levels is revealed by single point mutations that occur in PAS or in their surrounding sequence elements. Such mutations were found in the genes encoding α- and β-globin, p53, and prothrombin. These mutations are associated with disease phenotypes, including α- or β-thalassemia, cancer predisposition syndrome, microphthalmia, and thrombophilia despite causing only a 1.5-fold difference in steady-state mRNA levels (Higgs et al., 1983; Orkin et al., 1985; Gehring et al., 2001; Stacey et al., 2011; Johnston et al., 2019).

In addition to *cis*-elements, differential expression of *trans*-acting polyadenylation factors will influence PAS cleavage. As weak elements are often prone to upregulation by *trans*-acting factors, it is likely that weak PAS are the most regulated PAS. This mode of regulation may be especially important for genes that use alternative cleavage and polyadenylation to generate mRNA isoforms with alternative 3′UTRs as the majority of proximal PAS in 3′UTRs are regarded to be intrinsically weak (Legendre and Gautheret, 2003). Here, we used genes that generate alternative 3′UTRs as experimental system to study PAS cleavage activity as a change in cleavage activity of proximal PAS would manifest at the mRNA level as a change in alternative 3′UTR isoform usage.

Approximately half of human genes generate mRNA isoforms with alternative 3′UTRs (Lianoglou et al., 2013). These genes have especially long 3′UTRs that are on average more than four-times longer than genes with constitutive 3′UTRs (Lianoglou et al., 2013). Similar to the higher number of enhancers found in more complex organisms also 3′UTR length correlates with the number of cell types present in an organism which is often used as a measure for

organismal complexity (Chen et al., 2012; Mayr, 2017). The expansion of sequence elements in 3′UTRs may provide more post-transcriptional gene regulatory capacity.

3′UTRs control several aspects of mRNA metabolism including mRNA localization, mRNA stability and translational efficiency (Mayr, 2017, 2019b). Although a change in 3′UTR isoform usage does not affect the protein sequence, alternative 3′UTRs were shown to regulate protein multifunctionality by controlling alternative protein complex assembly (Berkovits and Mayr, 2015; Lee and Mayr, 2019; Mayr, 2019a). 3′UTRs also enable translation of proteins within defined subcellular compartments, which may be another way to influence protein functions (Ma and Mayr, 2018). These data indicate that functional diversity of proteins can be encoded by genomic sequence elements that are transcribed into 3′UTRs, thus providing one explanation of how 3′UTRs can contribute to increased organismal complexity.

As changes in alternative 3′UTR isoform usage can have important functional consequences, it is critical to understand how alternative 3′UTRs are controlled. Through knock-down (KD) experiments of RNA-binding proteins, including polyadenylation and splicing factors, it was shown that these factors often induce global shifts in 3′UTR isoform usage (Jenal et al., 2012; Lackford et al., 2014; Masamha et al., 2014; Gruber et al., 2016; Brumbaugh et al., 2018; Tian and Manley, 2017; Zhu et al., 2018; Chatrikhi et al., 2019). However, genome-wide analyses of 3′UTR isoform usage across cell types and various conditions did not show global shifts, but instead revealed gene- and condition-specific changes in 3′UTR ratios, meaning that the same stimulus in different cell types changed the usage of different PAS (Lianoglou et al., 2013).

In addition to RNA-binding proteins mRNA processing can also be influenced by promoters. For example, promoters may recruit polyadenylation and splicing factors or may influence transcription elongation (Dantonel et al., 1997; Rosonina et al., 2003; Calvo and Manley, 2005; Nagaike et al., 2011), reviewed in (Bentley, 2014; Rambout et al., 2018). Moreover, promoters were reported to influence several aspects of mRNA metabolism, including mRNA decay, mRNA localization, and translation (Trcek et al., 2011; Bregman et al., 2011; Zid and O'Shea, 2014; Slobodin et al., 2017), reviewed in (Haimovich et al., 2013).

As cell type- and condition-specific regulation is often mediated by transcription factors that bind to enhancers (Shlyueva et al., 2014; Gasperini et al., 2020), we established a reporter assay to investigate if promoters and enhancers regulate PAS cleavage activity in mammalian cells. Whereas promoters did not influence PAS cleavage activity in our system, we found that enhancers predominantly increased cleavage activity of proximal PAS. This result was confirmed for endogenous genes as cell type-specific 3′UTR shortening was associated with the

presence of cell type-specific enhancers. Moreover, deletion of an endogenous enhancer revealed that enhancers are required for a switch in alternative 3′UTR isoform usage. For *PTEN*, this mode of regulation is biologically important as the enhancer-mediated 3′UTR isoform change regulates the intrinsic enzymatic activity of PTEN protein in a manner that is independent of overall protein abundance.

Results

Promoters do not regulate PAS cleavage activity

The *PTEN* gene generates alternative 3′UTR isoforms that encode the same protein (Figure 1A) (Lianoglou et al., 2013). A change in cultivation conditions, including media acidification through prolonged culture or the addition of hydrochloric acid, induces 3′UTR shortening of *PTEN* in MCF7 cells (Figures 1B and S1A). We set out to investigate how the switch in 3′UTR isoform expression is regulated and hypothesized that promoters may regulate PAS usage in living cells. To assess promoter-dependent PAS cleavage activity, we established a luciferase reporter system that measures cleavage activity of individual PAS relative to the SV40 virus PAS (Figures 1C, S1B, and Table S1A). As the SV40 PAS is one of the strongest known PAS, it results in cleavage of basically all produced transcripts (Carswell and Alwine, 1989; Schek et al., 1992). Therefore, luciferase activity correlates with transcriptional activity of the promoter when the reporter is terminated by the SV40 PAS, a system that has been widely used to measure transcriptional activity (Smale, 2010). To obtain cleavage activity of a specific test PAS, we generated a similar reporter that instead of the SV40 PAS uses the test PAS for termination. The ratio of luciferase activities obtained from the reporter terminated by the test PAS over the SV40 PAS represents the cleavage activity of the test PAS (Figures 1C and S1B).

We examined promoter-dependent PAS cleavage activity of several promoters derived from viruses or human genes that differ 1000-fold in transcriptional activity (Figure 1D and Tables S1B and S1C). In the context of viral promoters, the proximal PAS (PPAS) of *PTEN* was weak compared to the strong SV40 PAS (Figure 1E). In addition to the SV40 PAS that was used as normalization control, we tested cleavage activity of the strong bovine growth hormone (BGH) PAS that is used in many expression vectors as well as the distal (DPAS) of *PTEN*. We observed a large difference in PAS cleavage activity between the weak PPAS and strong DPAS of *PTEN,* however, different promoters had only a small influence (less than 1.7-fold) on PAS cleavage activity (Figure 1F).

The PTEN enhancer increases cleavage activity of weak PAS

Next, we investigated if enhancers influence PAS cleavage activity and tested the promoter-proximal *PTEN* enhancer (Penh) in the context of the *PTEN* promoter (Pprom; Figures 2A, 2B, and Table S1D). The region of the Penh shows high levels of acetylated H3K27 (H3K27ac) in MCF7 cells and contains a large number of transcription factor binding sites (Figures 2A and S2). Importantly, enhancers were shown to retain their histone marks when transferred to heterologous genes in enhancer reporter assays (Lienert et al., 2011; Hathaway et al., 2012). Although typical enhancers increase transcriptional activity, the Penh modestly reduced transcriptional activity (< 1.6-fold) in the context of the Pprom but increased transcriptional activity in the context of a weak core promoter (Figures 2C and S1C). Addition of the enhancer had little influence on the cleavage activity of strong PAS, such as the BGH PAS or the PTEN DPAS (Figure 2D). However, strikingly, the enhancer increased cleavage activity of the PTEN PPAS by 3.7-fold (Figure 2D).

In this reporter system, relative PAS cleavage activity corresponds to luciferase activity if the different PAS do not influence stability of the reporters. To minimize the elements that affect mRNA stability we used minimal PAS that only contained 100 base pairs of surrounding sequence to ensure proper recognition of PAS (Table S1A). We also measured mRNA stability of the reporters and did not observe enhancer or PAS-dependent differences in stability (Figure S1D). To evaluate if presence of the enhancer changed the transcription start site of the reporter, we performed 5′RACE (rapid amplification of cDNA ends) and observed usage of the annotated transcription start site in the presence or absence of the enhancer (Figure 2E). This indicates that the mature mRNAs produced from the *PTEN* promoter, in the presence or absence of the *PTEN* enhancer, are identical.

Enhancers are known to regulate transcription independently of their orientation and can be located up- or downstream of genes (Shlyueva et al., 2014; Gasperini et al., 2020). We cloned the reverse complement of the enhancer downstream of the PAS and measured PAS cleavage activity. This revealed that the capacity of the enhancer to regulate transcriptional and PAS cleavage activity was largely retained when located downstream and in the opposite orientation (Figures S1E and S1F). Taken together, these data suggest that regulation of PAS cleavage activity is a *bona fide* enhancer activity.

Next, we assessed if the enhancer controls cleavage activity of additional PAS. The cleavage activity of the PPAS from *NUDT21* and *DICER1* increased 3.6-fold in the presence of the enhancer, whereas cleavage activity of non-proximal PAS (NUDT21 DPAS, GAPDH PAS, UBC

PAS) did not change in a coordinated manner (Figure 2F). These observations suggest that in the absence of an enhancer cleavage activity largely depends on the intrinsic strength of a PAS, but intrinsically weak PAS can have high *in vivo* cleavage activity when transcribed from promoters with active enhancers.

The *PTEN* enhancer is required for a switch in 3′UTR isoform expression of *PTEN* at the endogenous locus

To assess if the *PTEN* enhancer regulates PTEN PPAS usage at the endogenous gene locus, we used MCF7 cells that express wild-type *PTEN* (Mendes-Pereira et al., 2009) and deleted the promoter-proximal enhancer using a pair of guide RNAs (Figure 3A). Two clones (delta enhancer 1 (dE1) and dE2) with a heterozygous deletion in the region of the *PTEN* enhancer were used for follow-up experiments (Figures S3A and S3B). Deletion of the enhancer did not affect steady-state *PTEN* mRNA or protein level (Figures 3B and 3C). As the reporter assay suggested that the enhancer is required for increased PPAS usage, we performed a Northern blot of the parental wild-type (WT) and the enhancer deletion clones after cultivating them in normal and acidic media conditions. Under these conditions, WT cells induce the expression of the short *PTEN* 3′UTR isoform at the expense of the long 3′UTR isoform (Figure 3D). In contrast, cells with enhancer deletion did not upregulate the expression of the short *PTEN* 3′UTR isoform. These results indicate that the *PTEN* enhancer is required for *PTEN* 3′UTR shortening as both 3′UTR isoforms have the same mRNA decay rates (Figure S3C).

The switch in 3′UTR isoform expression of *PTEN* increases PTEN protein activity without affecting overall *PTEN* mRNA or protein abundance

Interestingly, the pH-induced 3′UTR ratio change did not affect overall PTEN protein levels (Figure 3E), suggesting that the alternative 3′UTR isoforms are translated with equal efficiency into PTEN protein. As the alternative 3′UTR isoforms do not change the sequence or size of PTEN, PTEN protein encoded by the short or long 3′UTR isoforms cannot be distinguished by western blot. We call the alternative PTEN protein products encoded from the mRNA isoforms with the short or long 3′UTRs PTEN-SU or PTEN-LU, respectively. The relative contribution of PTEN-SU and PTEN-LU to total PTEN protein is indicated by the colored boxes (Figure 3E).

Next, we investigated the biological significance of a *PTEN* 3′UTR ratio change. As PTEN is a phosphatase, we examined if the alternative 3′UTR isoforms influence PTEN activity. PTEN is a major negative regulator of the PI3K pathway and high PTEN activity corresponds to low levels of phosphorylated AKT (pAKT) (Maehama and Dixon, 1998; Li and Sun, 1998). We observed

that PTEN activity correlates with the abundance of PTEN-SU. Under normal cultivation conditions cells generate similar amounts of the short and long 3′UTR isoforms of *PTEN,* corresponding to an intermediate PTEN phosphatase activity (Figure 3F). The pH-induced increase in short 3′UTR isoform expression in WT cells was associated with higher PTEN phosphatase activity, whereas lack of 3′UTR shortening in enhancer deletion cells correlated with lower PTEN activity, indicated by higher levels of pAKT (Figure 3F). To investigate if the difference in PTEN activity is caused by 3′UTR-dependent regulation of intrinsic enzymatic activity, we performed an ELISA and measured PTEN lipid phosphatase activity using PIP3 as substrate. In order to avoid pH-mediated PTEN activity regulation, we obtained MCF7 cells that predominantly express the short *PTEN* 3′UTR isoform using shRNA-mediated KD of the long *PTEN* 3′UTR (LU KD; Figure S3D). Control (Ctrl) KD cells express similar amounts of *PTEN-SU* and *PTEN-LU*. We immunoprecipitated PTEN protein from Ctrl KD and LU KD cells and used equal amounts of total PTEN protein as input for the ELISA (Figure 3G). Intriguingly, we observed higher PTEN lipid phosphatase activity towards PIP3 with PTEN-SU compared with PTEN-LU (Figure 3H). Taken together, these results indicate that PTEN protein translated from the mRNA isoform containing the short 3′UTR has a higher intrinsic activity than PTEN protein translated from the long 3′UTR isoform. As short vs long 3′UTR isoform expression is regulated by the promoter-proximal *PTEN* enhancer, our data indicate that the enhancer regulates protein activity independently of overall protein abundance.

A distal enhancer also regulates PAS cleavage activity

We then asked if PAS cleavage activity is regulated by other enhancers. The *NUDT21* gene encodes an important polyadenylation factor and generates alternative 3′UTRs (Lianoglou et al., 2013). We tested if the enhancer of the *NUDT21* gene regulates PAS cleavage activity. As high levels of H3K27ac are characteristic of enhancers (Shlyueva et al., 2014; Gasperini et al., 2020), we searched for increased H3K27ac levels in the vicinity of the *NUDT21* gene. We detected a ChIP-seq peak with high H3K27ac levels approximately 80 kb downstream of the *NUDT21* gene which is suggestive of a distal enhancer (Figure 4A). We cloned 2 kb of this region that we called the *NUDT21* enhancer (Nenh) and placed it upstream of the Pprom or the *NUDT21* promoter (Nprom) into the luciferase reporters (Figure 4B). In the context of the reporter, the Nenh did not significantly change transcriptional activity of the two promoters (Figure 4C). However, it increased PTEN PPAS cleavage activity between 3.4 and 5.3-fold without affecting cleavage activity of a stronger PAS (Figures 4D and 4E). This suggests that

enhancer-mediated regulation of PAS cleavage activity has the potential to be widespread as two out of two tested enhancers regulated cleavage activity.

Cell type-specific switches in 3′UTR isoform expression that are caused by upregulation of short 3′UTR isoforms are associated with the presence of cell type-specific enhancers

To investigate if enhancers are preferentially associated with cell type-specific changes in 3′UTR isoform expression at endogenous genes, we analyzed cell type-specific changes in gene-level mRNA abundance and 3′UTR isoform usage between hematopoietic stem cells (HSC) and erythrocytes (Ery) (Figure 4F) (Ha et al., 2018; Numata et al., 2020; Cai et al., 2020). The Ery RNA sequencing dataset was chosen as it provided additional information on the presence of cell type-specific enhancers that are associated with specific genes (Cai et al., 2020). We focused on the genes whose gene-level mRNA abundance increased more than two-fold in Ery compared with HSC (Figures 4F and 4G). These genes can be separated into two groups: 1) Both short 3′UTR (*SU*) and long 3′UTR (*LU*) isoforms increase similarly (dSU ≈ dLU, blue) and 2) *LU* expression remains similar and only *SU* increases (dSU/dLU >2, red; Figures 4F-H). We then assigned the previously identified cell type-specific enhancers to their associated genes (Cai et al., 2020). For genes whose mRNA expression was not altered when comparing HSC and Ery, we observed an association of Ery-specific enhancers in 13.8% of cases (no change (NC; Figure 4I). However, when both 3′UTR isoforms were similarly upregulated 23.9% of genes were associated with an Ery-specific enhancer. Upregulation of both 3′UTR isoforms is likely accomplished by increased transcriptional activity. This result was expected as a cell type-specific increase in mRNA abundance is known to be associated with active enhancers (Creyghton et al., 2010). Intriguingly, increased association of cell type-specific enhancers and genes (20.5%) was also observed for genes whose increased mRNA level was solely accomplished by upregulation of their short 3′UTR isoforms as expression of their long 3′UTR isoforms did not change (dSU/dLU >2; Figure 4I).

The simplest explanation for exclusive upregulation of the short 3′UTR isoform is an enhancer-mediated increase in cleavage activity of the proximal PAS, illustrated schematically in Figure 4J (Model 1) which is also supported by our reporter results. We currently cannot rule out that increased mRNA levels that are solely accomplished by upregulation of short 3′UTR isoforms are caused by two consecutive steps of gene regulation that include upregulation of transcript abundance followed by exclusive degradation of long 3′UTR isoforms to previously observed levels (Figure 4J, Model 2). However, in our opinion, this model is unlikely. Taken together, these data confirm that cell type-specific enhancers are associated with genes whose mRNA

levels increase in a cell type-specific manner. Our data support a model wherein cell type-specific enhancers increase transcript production resulting in a similar upregulation of all 3′UTR isoforms (*N* = 86) or increase cleavage activity of the proximal PAS (*N* = 153), thus resulting in cell type-specific 3′UTR shortening.

Transcription factors regulate PAS cleavage activity without affecting transcriptional activity

Next, we set out to identify transcription factors and co-activators that are responsible for the regulation of PAS cleavage activity. We used the *PTEN* enhancer and performed a small-scale shRNA screen. We stably knocked-down individual factors that are known to bind to the Penh in MCF7 cells and that are involved in transcription regulation (Figure S4 and Table S2). We compared the effect of Ctrl KD and transcription factor KD on transcriptional activity and PAS cleavage activity in the context of the Penh-Pprom reporter (Figure 5A). As before, we used the reporter with the SV40 PAS to assess transcriptional activity and the reporter with the PTEN PPAS to determine PAS cleavage activity (Figures S5A and S5B).

In Ctrl KD cells, PTEN PPAS usage was 0.60 (Figure 5B, black dots). As positive control, we knocked-down the polyadenylation factor FIP1L1, which was shown previously to be required for PPAS usage (Lackford et al., 2014). As expected, KD of FIP1L1 decreased PTEN PPAS usage from 0.6 to 0.36 without affecting transcriptional activity (Figure 5B, black dot). Strikingly, KD of 10/21 tested transcription factors had a similar effect (Figure 5B, red dots). These transcription factors included RELA (NF-κB p65), MYC, and FOXA1 and they regulated PAS cleavage activity without having a strong (less than 1.7-fold) effect on transcriptional activity of the reporter (Figure 5C and Tables 1 and S2).

Mutation of MYC binding sites in the enhancer regulates PAS cleavage activity

We tested the effect of RELA or MYC KD on cleavage activity of additional PPAS and again observed decreased cleavage activity in the context of the Penh reporter (Figures 5D and S5C). MYC binding sites (MYC-BS) are E-boxes (Grandori et al., 1996) and the *PTEN* enhancer contains two conserved MYC-BS (Figures 5E and 5F). We next evaluated the influence of MYC-BS mutation on transcriptional and PAS cleavage activity in the context of the reporter. Mutation of the MYC-BS had no influence on transcriptional activity (Figure 5C). It decreased cleavage activity of weak PPAS and did not affect cleavage activity of strong DPAS, thus phenocopying the effect of MYC KD (Figures 5D). These results suggest that transcription factor binding to conserved motifs in the *PTEN* enhancer regulates cleavage activity of proximal PAS.

Co-activators simultaneously regulate transcriptional activity and PAS cleavage activity

As transcription factor binding to enhancers recruits diverse co-activators to promoters, we also tested the effects of co-activator KD on transcriptional and PAS cleavage activity of the reporter (Figure 6A).

In contrast to transcription factors, the majority of tested co-activators (26/35) changed transcriptional activity and PAS cleavage activity at the same time (Figures 6B, purple dots and Table 1). The co-activators included Mediator, general transcription factors, transcription elongation factors, and several histone acetyltransferases (Figures S5A, S5B, and Table S2). KD of the general transcription factor TFIIF or the histone acetyl transferases TIP60 (KAT5) and PCAF (KAT2B) also reduced cleavage activity of additional PPAS (Figure 6C). Moreover, the KD of histone acetyltransferases (PCAF, TIP60, EP300) also decreased PTEN PPAS cleavage activity in the context of the Nenh (Figures 6D and 6E) but had no effect on cleavage activity in reporters that lack the enhancer (Figures S5D and S5E). These results suggest that active enhancers are responsible for increased cleavage activity of weak PAS.

Model of enhancer-mediated control of PTEN protein activity through regulation of 3′UTR isoform usage

In summary, our results obtained from reporter assays, from the genes that undergo cell type-specific 3′UTR shortening, and from the CRISPR-mediated deletion of an endogenous *PTEN* enhancer indicates that enhancers, in addition to regulating transcript production also control PAS cleavage. In the absence of an enhancer or when the enhancer is inactive, the produced transcripts are cleaved and polyadenylated at their 3′ ends based on the intrinsic strength of the PAS (Figure 6F, left panel). In the presence of an active enhancer, PAS cleavage activity of proximal and weak PAS increases, resulting in a change in 3′UTR isoform expression with upregulation of the short 3′UTR isoform (Figure 6F, right panel). For *PTEN*, this mode of regulation did not affect gene-level mRNA or protein levels. However, the relative contribution of PTEN-SU and PTEN-LU to overall PTEN protein has changed with an increased fraction of PTEN-SU. As PTEN-SU has a higher intrinsic phosphatase activity than PTEN-LU, enhancer activation ultimately changes protein activity. We speculate that the change in 3′UTR-dependent protein activity may be due to conformational changes and/or post-translational modifications (see Discussion).

Discussion

We established a new assay that allows us to assess promoter and enhancer-dependent PAS cleavage activity in living cells. This assay revealed that promoters that differ 1000-fold in transcriptional activity had little effect on PAS cleavage activity. In contrast, we observed that enhancers regulate PAS cleavage activity substantially. The effect of enhancer activation on strong PAS – that usually are located at the 3′ ends of transcription units – was variable and usually resulted in no change or in a slight downregulation of cleavage activity. However, for six out of six weak PAS that are located in proximal 3′UTRs, we found that the presence of an active enhancer caused a substantial upregulation of cleavage activity (Figures 2D, 2F, 4D, and 4E). These results are supported by a significant association of cell type-specific enhancers with genes that undergo a cell type-specific switch towards 3′UTR shortening (Figures 4F-J) and they were confirmed by the deletion of an endogenous enhancer (Figure 3). In the future, this assay can also be applied to assess the influence of signaling pathways or chromatin states on PAS cleavage activity.

Transcription factors regulate PAS cleavage

Surprisingly, half of the transcription factors that were knocked-down in the context of the reporter only affected PAS cleavage activity without having a strong (less than 1.7-fold) effect on overall transcriptional activity (Figure 5B, Tables 1 and S2). For most transcription factors, we only achieved a partial knock-down, suggesting that we may even underestimate the effects on PAS cleavage activity. Our results indicate that transcription factors play a major role in the regulation of PAS cleavage activity. Although the KD of co-activators and chromatin regulators also regulated PAS cleavage activity, we observed that these factors influenced both transcriptional and PAS cleavage activity simultaneously (Figure 6B, Tables 1 and S2).

In the context of the reporter, we did not observe a promoter- or enhancer-mediated difference in mRNA stability (Figure S1D). The influence of mRNA decay on reporter expression was minimized as the reporter constructs used for measuring transcriptional and cleavage activity only differed in 100 base pairs of sequence surrounding the PAS. Also, the mRNA decay rates of the endogenous alternative 3′UTR isoforms of *PTEN* do not differ in MCF7 cells (Figure S3C). This result confirms previously obtained transcriptome-wide data on alternative 3′UTR isoforms that found a high correlation in mRNA decay rates for the corresponding short and long 3′UTRs (Spies et al., 2013). We present here strong evidence that PAS cleavage activity is a major contributor to the regulation of mRNA isoform abundance. Our study represents the first step in the development of tools that will be able to measure cleavage activity transcriptome-wide

without the need for reporters. Such tools have already been developed to measure transcript production or mRNA stability on a large scale (Mahat et al., 2016; Matsushima et al., 2018).

Enhancer-mediated upregulation of cleavage activity of proximal PAS is widespread at endogenous genes

Enhancers are known to integrate cell type- and condition-specific signals to regulate gene expression (Levine et al., 2014; Gasperini et al., 2020). In addition to regulating gene-level mRNA abundance, our transcriptome-wide analyses reveal that cell type-specific enhancers are also widespread regulators of cell type-specific usage of 3′UTR isoforms as they were significantly associated with genes that underwent 3′UTR shortening (Figure 4I). When taken together with our reporter results this analysis revealed that enhancer-mediated regulation of PAS cleavage is widespread. When analyzing the upregulated mRNAs in erythrocytes compared with hematopoietic stem cells, cell type-specific enhancer-mediated PAS cleavage was observed more frequently as than enhancer-mediated regulation of transcriptional activity (N = 153 vs N = 86; Figure 4I).

We present two possible models for the increased mRNA abundance that is caused by preferential upregulation of the short 3′UTR isoform. The simplest explanation is an increased cleavage activity of the proximal PAS (Figure 4J), a model that is also supported by the results obtained from the reporter assays and the *PTEN* enhancer deletion. Theoretically, it is also possible that an enhancer-mediated increase in transcript production is coupled with exclusive degradation of the long 3′UTR isoforms. This model is unlikely as the expression levels of the long 3′UTR isoforms need to be decreased to the levels observed previously. Our data further point to the possibility that in certain cell types such as HSCs, not all produced transcripts are processed into mature mRNAs (Figure 4J, left panel). These unprocessed transcripts are usually immediately degraded; however, they could represent a reservoir of transcripts that are available for mRNA processing upon activation of signaling pathways, thus achieving fast changes in 3′UTR isoform expression. This represents an exciting direction for future studies.

In addition to the increase in mRNA abundance that is caused by upregulation of short 3′UTR isoforms, there seem to be enhancers that only accomplish a switch in alternative mRNA isoform expression without altering gene level mRNA abundance, as we have observed upon CRISPR-mediated deletion of the *PTEN* enhancer (Figure 3C). The current definition of active enhancers requires mRNA upregulation of the associated gene (Gasperini et al., 2020). With this definition, the *PTEN* enhancer would be missed. In a recent large-scale study that used CRISPRi to repress enhancers only 10% of tested enhancers showed any evidence of

enhancer-mediated regulation of mRNAs levels (Gasperini et al., 2019). It is possible that the number of functional enhancers may increase if mRNA isoform changes would be measured in addition to assessing gene-level abundance. Still, most analysis tools only quantify gene-level mRNA abundance and are not designed to detect alternative mRNA transcript or isoform abundance (Ntranos et al., 2019). This illustrates the need for better tools to assess transcript-level changes and the resulting functional changes in protein activity of highly regulated factors as a large fraction of genes produces alternative mRNA isoforms that only differ in their 3′UTRs (Lianoglou et al., 2013; Lackford et al., 2014; Masamha et al., 2014; Gruber et al., 2016; Brumbaugh et al., 2018; Tian and Manley, 2017).

We currently do not know how transcription factors mediate the regulation of PAS cleavage activity. However, it was previously shown that transcription factors that bind to promoters are able to recruit RNA-binding proteins, including polyadenylation factors (Dantonel et al., 1997; Rosonina et al., 2003; Calvo and Manley, 2005; Nagaike et al., 2011). These factors then bind to the C-terminal domain of RNA polymerase II and may be deposited onto the nascent mRNA in a co-transcriptional manner to regulate mRNA processing (McCracken et al., 1997; Hsin and Manley, 2012; Bentley, 2014; Rambout et al., 2018).

To regulate expression of individual mRNA isoforms, we propose that enhancers recruit specific RNA-binding proteins to promoters that travel with RNA polymerase II and when present result in increased cleavage at weak PAS. Such a mechanism was also proposed for promoter-dependent regulation of post-transcriptional processes in yeast, including the regulation of mRNA stability, cytoplasmic localization, and translation (Harel-Sharvit et al., 2010; Trcek et al., 2011; Bregman et al., 2011; Zid and O'Shea, 2014; Haimovich et al., 2013).

An enhancer controls PTEN protein activity via regulation of alternative 3′UTR isoform expression

For PTEN, we showed that the regulation of PAS cleavage activity did not change overall mRNA or protein levels (Figure 3B-E). However, importantly, differential 3′UTR isoform expression resulted in a difference in PTEN protein activity (Figure 3F). We observed a higher intrinsic enzymatic activity measured as lipid phosphatase activity against PIP3 with PTEN-SU compared with PTEN-LU (Figure 3H). The difference in protein activity is best explained by 3′UTR-dependent effects on protein maturation that either happen co-translationally or right after the protein has been synthesized but when it is still in the proximity of the ribosome (Mayr, 2019a). We showed previously that 3′UTRs recruit proteins to the site of translation that interact with the newly made proteins, thus establishing 3′UTR-dependent protein-protein interactions

(Berkovits and Mayr, 2015; Lee and Mayr, 2019). Alternatively, 3′UTRs can determine translation in specific local environments such as TIS granules which can also affect protein complex assembly (Ma and Mayr, 2018).

Currently, we do not know how the alternative 3′UTRs of *PTEN* regulate the enzymatic activity of PTEN. We envision several possible scenarios: It has been shown that phosphorylation of PTEN in the C-terminus can abrogate PTEN activity as it induces a closed and inactive protein conformation that is no longer able to associate with membranes and is unable to regulate downstream signaling pathways (Worby and Dixon, 2014). It is possible that the long 3′UTR of *PTEN* recruits enzymes that post-translationally modify PTEN protein, thus inducing a conformational change. Moreover, PTEN protein activity is redox-sensitive as the cysteine at the active site can be oxidized (Kwon et al., 2004). The alternative 3′UTRs may promote translation in local environments that differ in redox state or the 3′UTRs may recruit proteins that either increase or decrease the oxidative state of the cysteine in the active site.

Half of genes generate mRNA transcripts with constitutive 3′UTRs (Lianoglou et al., 2013). These genes regulate protein activity through the regulation of mRNA and protein abundance. The majority of genes that generate alternative 3′UTR isoforms encode regulatory factors, including transcription factors, RNA-binding proteins, chromatin regulators, kinases, and ubiquitin enzymes (Lianoglou et al., 2013). These genes have an additional way to regulate protein activity through the regulation of 3′UTR isoform abundance and 3′UTR-dependent activity regulation (Figure 6F). This has the added benefit that different functions of a single protein can be regulated separately. For example, protein abundance of the ubiquitin ligase BIRC3 regulates cell death, but BIRC3-LU has additional functions that include the regulation of B cell migration through 3′UTR-dependent protein complex assembly (Lee and Mayr, 2019). This implies that a change in protein abundance regulates cell death, but a 3′UTR isoform change regulates migration. It is likely that also PTEN-SU and PTEN-LU have different functions. We showed that PTEN-SU has higher enzymatic activity, but PTEN also has functions that do not require its lipid phosphatase activity (Worby and Dixon, 2014), and we speculate that PTEN-LU may accomplish these functions. Nevertheless, these examples illustrate how increased enhancer regulation and increased control of protein functions by alternative 3′UTR isoforms may cooperate with each other to accomplish highly sophisticated gene expression regulation observed in complex organisms.

Acknowledgements

We thank all members of the Mayr lab for helpful discussions and Kristian Helin for critical reading of the manuscript. This work was funded by the NIH Director′s Pioneer Award (DP1-GM123454), Damon Runyon-Rachleff Innovation Award, the Pershing Square Sohn Cancer Research Alliance, and the NCI Cancer Center Support Grant (P30 CA008748).

B.K. designed and performed all experiments regarding endogenous PTEN, M.M.F. performed all computational analyses, N.P. and J.L. cloned the reporter constructs and performed the reporter assays, J.L. evaluated the knock-down efficiency of the shRNAs, S.H.L. performed the western blot for pAKT and immunoprecipitated PTEN for ELISA, W.M. performed the mRNA stability experiment of the reporter. C.M. conceived and supervised the project, designed and analyzed the reporter experiments, and wrote the manuscript with input from all authors.

The authors declare no competing interests.

Figure legends

Figure 1. Promoters do not regulate PAS cleavage activity.

(A) Gene model of *PTEN*. The gene is processed into alternative mRNA isoforms that differ in their 3′UTRs but encode proteins with identical sequence. Transcript cleavage at the proximal PAS (PPAS) generates *PTEN-SU*, whereas transcript cleavage at the distal PAS (DPAS) generates *PTEN-LU*.

(B) Northern blot showing *PTEN* 3′UTR isoform expression in MCF7 cells cultivated in normal (N) or acidified media (Ac; pH = 6.5). The shortest (*PTEN-SU*) and longest (PTEN-LU) 3′UTR isoforms are indicated. Cultivation in acidified media switches 3′UTR isoform ratios and results in 3′UTR shortening of *PTEN*.

(C) Schematic of luciferase reporter constructs used to investigate promoter-dependent PAS cleavage activity. The reporter construct with the SV40 PAS provides transcriptional activity of the promoter. Cleavage activity of a test PAS is obtained by the ratio of the luciferase activities obtained from the test PAS reporter over the SV40 PAS reporter when transcribed from the same promoter. The transcription start site is indicated by the arrow. Rluc, Renilla luciferase. See Fig. S1B for more extensive description.

(D) Transcriptional activities of the indicated reporters in MCF7 cells. Transcriptional activity corresponds to luciferase activity of each reporter when terminated by the SV40 PAS. The promoter sequences are reported in Table S1.

(E) Cleavage activity of the PTEN PPAS is independent of transcriptional activity. PTEN PPAS cleavage activity is similar when the reporter is expressed from four different viral promoters despite a 500-1,000-fold difference transcriptional activity, shown in (D). One-way ANOVA was performed. NS, not significant.

(F) Promoter-dependent PAS cleavage activity of different PAS. Bovine growth hormone (BGH) PAS, PTEN PPAS (proximal PAS of PTEN), PTEN DPAS (distal PAS of PTEN). The PAS sequences are reported in Table S1. One-way ANOVA was performed. **, $P = 3 \times 10^{-9}$; *, $P = 0.04$.

Figure 2. Enhancers upregulate cleavage activity of proximal and weak PAS.

(A) UCSC genome browser snapshot showing the *PTEN* genomic locus and a zoom into the vicinity of the transcriptional start site (arrow). Also shown are Encode tracks for H3K27ac in

MCF7 and transcription factor (TF) ChIP-seq binding sites. Penh, PTEN enhancer; Pprom, PTEN promoter.

(B) Schematic of reporter constructs to investigate enhancer-dependent PAS cleavage activity. As in Fig. 1C, but instead of comparing the influence of two promoters on PAS cleavage activity, the presence or absence of the Penh in the context of the Pprom is examined.

(C) Transcriptional activity of the Pprom reporter in the presence (Penh-Pprom) or absence (Pprom) of the PTEN enhancer obtained by luciferase activity of the reporters when terminated by the strong SV40 PAS. T-test for independent samples was performed; $P = 0.002$.

(D) Enhancer-dependent PAS cleavage activity of different PAS. The Penh increases PTEN PPAS cleavage activity. T-test for independent samples was performed; **, $P = 1 \times 10^{-8}$; NS, not significant.

(E) The transcription start sites of the Pprom reporters in the presence or absence of Penh was determined by 5′RACE. The Penh is a true enhancer and does not change the transcription start site. The canonical transcription start site (TSS) is used in both reporters and is indicated by the arrow.

(F) As in (D), but additional PAS are shown. T-test for independent samples was performed; **, $P = 1 \times 10^{-6}$; *, $P = 0.001$.

Figure 3. The *PTEN* enhancer is required for 3′UTR shortening of *PTEN* at the endogenous gene locus and regulates PTEN protein activity through a 3′UTR isoform change.

(A) As in Fig. 2A but indicated are the positions of the guide RNAs (gRNA) used for heterozygous deletion of the *PTEN* enhancer in MCF7 cells.

(B) qRT-PCR in wild-type (WT) and enhancer deletion cells (dE1, dE2) shows that steady-state *PTEN* mRNA level are not affected by heterozygous *PTEN* enhancer deletion. *RPL19* mRNA was used for normalization. T-test for independent samples was performed

(C) Western blot of PTEN in WT and enhancer deletion cells was performed in steady-state cultivation conditions. GAPDH was used as loading control.

(D) Northern blot of endogenous *PTEN* from WT and enhancer deletion cells cultivated in normal (N) or acidified (Ac, pH = 6.5) media. A low pH induces 3′UTR shortening of *PTEN* which

is prevented upon lack of the enhancer. The RNA gel is shown as loading control. The colored boxes indicate the contribution of *PTEN-SU* and *PTEN-LU* to total *PTEN* mRNA levels.

(E) Western blot of PTEN in WT and enhancer deletion cells after cultivation in normal or acidified media. GAPDH was used as loading control. The colored boxes indicate the contribution of PTEN-SU and PTEN-LU to total *PTEN* protein levels.

(F) Western blot of pAKT (S473) and total AKT in WT and enhancer deletion cells after cultivation of the cells in normal or acidified media. GAPDH is shown as additional loading control. The colored boxes indicate the contribution of PTEN-SU and PTEN-LU to total *PTEN* protein levels. Increased expression of PTEN-SU correlates with higher PTEN phosphatase activity, corresponding to lower pAKT level. Low amounts of PTEN-SU in enhancer deletion cells correlate with lower PTEN activity and higher pAKT level.

(G) Western blot showing PTEN protein amount that was used as input for PTEN activity assessment by ELISA. To better judge protein levels, increasing amounts of protein were loaded. Ctrl KD cells were obtained by stably expressing a ctrl shRNA, whereas LU KD cells were obtained by stably expressing an shRNA against the long 3′UTR isoform of *PTEN*. The colored boxes indicate the contribution of PTEN-SU and PTEN-LU to total PTEN protein levels.

(H) Enzymatic PTEN lipid phosphatase activity was measured by ELISA with PIP3 as substrate. LU KD cells predominantly generate PTEN-SU which correlates with a higher enzymatic activity. N = 11 biological replicates; T-test for independent samples was performed; **, P = 0.003.

Figure 4. Enhancer-mediated regulation of PAS cleavage is widespread.

(A) UCSC genome browser snapshot showing the *NUDT21* gene locus with Encode ChIP-seq data of H3K27ac. The region that was used as *NUDT21* enhancer (Nenh) is indicated. See Table S1 for details.

(B) Schematic of reporter constructs used to investigate Nenh-dependent PAS cleavage activity in the context of two promoters (Pprom and NUDT21 promoter (Nprom)), shown as in Figure 1C.

(C) Transcriptional activity of the reporters in the presence (Nenh-Pprom or Nenh-Nprom) or absence (Pprom or Nprom) of the Nenh. Transcriptional activity corresponds to luciferase activity of the reporters when terminated by the strong SV40 PAS. T-test for independent samples was performed.

(D) Nenh-dependent PAS cleavage activity in the context of the Pprom. T-test for independent samples was performed; **, *P* = 1 x 10^-8.

(E) Nenh-dependent PAS cleavage activity in the context of the Nprom. T-test for independent samples was performed; **, *P* = 1 x 10^-5.

(F) Changes in gene-level mRNA and 3′UTR isoform expression between Ery and HSC. Each dot represents an mRNA. The difference in short 3′UTR isoform expression is calculated as SU/(SU+LU) in Ery minus HSC. Subgroups of genes are color-coded according to their changes in gene expression and 3′UTR isoform expression. No change (NC; dark grey *N* = 478); gene expression fold change (FC) of Ery/HSC > 2, blue or red. FC in short 3′UTR isoform expression (dSU) is similar to FC in long 3′UTR isoform expression (dLU; dSU ≈ dLU, blue *N* = 360), dSU/dLU >2, red (*N* = 748); other, light grey (*N* = 621).

(G) Changes in gene-level mRNA abundance in the groups from (F). Violin plots denote median, 25th and 75th percentiles.

(H) Changes in 3′UTR isoform expression (dSU, dLU) in the groups from (F). Shown as in (G).

(I) Fraction of genes associated with Ery-specific enhancers (Eryenh+) in the groups from (F). Chi-square test; **, *P*< 0.003. The numbers indicate the number of genes associated with or without Eryenh.

(J) Two potential models to explain upregulation of gene-level mRNA abundance caused by exclusive upregulation of the short 3′UTR isoform in Ery compared to HSC. The lines represent transcripts. Lines without AAA indicate unprocessed transcripts, short lines with AAA indicate short 3′UTR isoforms, whereas longer lines with AAA indicate long 3′UTR isoforms.

Figure 5. PAS usage is regulated by transcription factors that bind to the *PTEN* enhancer.

(A) Schematic of reporter constructs to investigate transcription factor-dependent PAS cleavage activity. As in Figure 1C, but the effect of knock-down (KD) of individual transcription factors known to bind to the Penh on PAS cleavage activity is compared to the effect observed by shRNA controls (Ctrl KD).

(B) Summary of the results of the transcription factor shRNA screen. Shown is the effect of KD of individual transcription factors on transcriptional activity and on PTEN PPAS cleavage activity. Tx, transcription. The grey area on the x-axis denotes no significant change in

transcriptional activity, whereas the grey area on the y-axis denotes no significant change in PAS cleavage activity. See Figure S5 and Table S2 for complete data.

(C) Transcriptional activity upon KD of a subset of transcription factors is shown, as in Fig. 2C. T-test for independent samples was performed.

(D) Transcription factor-dependent cleavage activity of additional PAS is shown after KD of specific transcription factors. T-test for independent samples was performed; *, $P < 0.03$.

(E) UCSC genome browser snapshot showing the vicinity of the transcription start site (arrow) of the *PTEN* gene together with ChIP-seq data for MYC and the sequence conservation track of 100 vertebrates. The positions of the two conserved canonical E-boxes (binding sties for MYC, MYC-BS) are indicated.

(F) Sequence conservation of the two MYC-BS located in the *PTEN* enhancer in different organisms.

Figure 6. PAS usage is regulated by co-activators and histone acetyltransferases.

(A) As in Fig. 5A, but co-activators were knocked-down.

(B) Summary of the results of the co-activator shRNA screen, shown as in Figure 5B.

(C) Co-activator-dependent cleavage activity of additional PAS is shown after KD of specific co-activators and histone acetyl transferases. T-test for independent samples was performed; *, $P < 0.03$.

(D) Cleavage activity of the PTEN PPAS in the context of the Nenh-Pprom reporter comparing the effect of KD of specific histone acetyl transferases with a ctrl KD. T-test for independent samples was performed. **, $P < 0.001$.

(E) As in (D), but cleavage activity of the PTEN PPAS in the context of the Nenh-Nprom reporter was assessed. KD of histone acetyl transferases reduces PTEN PPAS cleavage activity. T-test for independent samples was performed; *, $P < 0.03$.

(F) Model of enhancer-mediated control of PTEN protein activity through regulation of 3′UTR isoform usage (see text).

Table 1. Transcription factors and co-activators that regulate PAS cleavage activity.

Transcription factors that regulate PAS cleavage
FOS
FOXA1
GABPA
IRF1
JUN
JUND
MYC
NFYB
RELA
TCF12
Transcription factors that regulate PAS cleavage and transcription
EGR1
ELF1
RFX5
RXRA
TFAP2A
YY1
Co-activators that regulate PAS cleavage
ING3
SMC3
Co-activators that regulate PAS cleavage and transcription
ACTL6A
BARD1
BRCA1
BRD8
CDC73
CDK9
EP300
EP400
GTF2F1
KAT2B
KAT5
MED1
MORF4L1
NELFB
NELFE
NPM1
SIN3A
SMARCC1
SSRP1
SUPT16H
SUPT4H1
SUPT5H
TAF1
TBP
TCERG1
USP22

Supplementary Figure legends

Figure S1. Enhancer-mediated regulation of PAS cleavage activity.

(A) Northern blot showing *PTEN* 3′UTR isoform expression in MCF7 cells cultivated for 6 days. This results in media acidification (Ac, pH = 6.5). The RNA gel serves as loading control. The shortest (*PTEN-SU*) and longest (*PTEN-LU*) 3′UTR isoforms are indicated. Prolonged culture of MCF7 cells switches 3′UTR isoform ratios and results in 3′UTR shortening of *PTEN*.

(B) Schematic of luciferase reporter assay to measure promoter-dependent cleavage activity of PAS. The SV40 virus late PAS is one of the strongest PAS known and is used as reference PAS. It is the default PAS found in commercial luciferase reporters. As it is assumed that basically all transcripts that are produced from a specific promoter are cleaved and polyadenylated by the SV40 PAS, luciferase activity obtained from reporter A corresponds to transcriptional activity of promoter 1. Such measurements of transcriptional activity are well established (Smale, 2010).

To assess promoter-dependent PAS cleavage activity, we calculate the ratio of luciferase activity obtained from reporter B over A; this corresponds to the cleavage activity of a test PAS in the context of promoter 1. Similarly, we calculate the activity of a test PAS in the context of promoter 2 by calculating the ratio of luciferase activity of reporters D over C.

As only mature mRNA transcripts (that were cleaved and polyadenylated) contribute to luciferase protein production, and thus to luciferase activity, PAS cleavage activity is proportional to luciferase activity. This is true if in a pair of reporters (A and B or C and D) the number of transcripts produced is identical and if their mRNA stability rates are comparable. To minimize the impact of regulatory elements located in the 3′UTR on the stability of the reporters, we used minimal PAS instead of whole 3′UTRs. Minimal PAS consist of endogenous polyadenylation sites (hexamer AAUAAA) together with 100 nucleotides of endogenous surrounding sequence (Table S1). For the most extensively used reporters, promoter-dependent mRNA stability was tested experimentally (Figure S1D).

(C) Transcriptional activity of a weak promoter (TATA) is upregulated by addition of the Penh (Penh-TATA). T-test for independent samples was applied. **, $P = 2 \times 10^{-4}$.

(D) Reporter mRNA transcript stability was assessed after Actinomycin D (ActD) treatment of MCF7 cells expressing the indicated reporters at the indicated time points. The Penh did not influence mRNA stability of the reporters. The sequence context of the SV40 PAS and the PTEN PPAS did not influence the stability of the reporters. One-way ANOVA was performed.

(E) PTEN PPAS cleavage activity was measured in the presence of the Penh1, either located upstream or in the reverse complement (rc) orientation and located downstream of the transcription unit. Presence of the Penh1 regardless of location increases PAS cleavage activity. T-test for independent samples was performed; **, $P = 1 \times 10^{-5}$.

(F) Transcriptional activity of the reporters from (E). T-test for independent samples was performed; **, $P = 1 \times 10^{-5}$. Although the difference in transcriptional activity has a significant *P* value, the fold-change is only approximately 2-fold.

Figure S2. Transcription factors identified by ChIP-seq to bind to the *PTEN* enhancer and promoter. Shown is a UCSC genome browser snap shot for the *PTEN* gene locus of transcription factor ChIP-seq data generated by the Encode project (Consortium, 2012).

Figure S3. CRISPR-mediated deletion of the *PTEN* enhancer in MCF7 cells.

(A) Shown is the gene locus of *PTEN* as in Figure 3A, including the position of the guide RNAs used to delete the *PTEN* enhancer. The obtained deletions in dE1 and dE2 are also shown.

(B) Schematic showing strategy to assess the number of deleted alleles upon CRISPR-mediated *PTEN* enhancer deletion. As MCF7 cells contain more than two alleles for chromosome 10, four PCR amplicons were designed and the DNA content within the deletion was compared to the outside region and normalized to the WT sample. Shown is the fold change in DNA content in the *PTEN* enhancer region in dE1 and dE2 compared with WT cells. One-way ANOVA, $P = 3 \times 10^{-7}$, t-test for independent samples: WT vs dE1: $P = 0.001$, WT vs dE2: $P = 3 \times 10^{-5}$.

(C) Northern blot showing endogenous *PTEN* mRNA expression in MCF7 cells before and after treatment with Actinomycin D (ActD). Right panel shows quantification of the bands and shows that the long and short 3′UTR isoforms of *PTEN* have similar degradation rates.

(D) Northern blot showing *PTEN* mRNA expression using RNA samples from MCF7 cells after stable KD of the long *PTEN* 3′UTR isoform (LU KD) and stable KD using a control shRNA (ctrl KD). Cells with LU KD predominantly express the short 3′UTR isoform of *PTEN*. The RNA gel serves as a loading control.

Figure S4. shRNA-mediated knock-down of transcription factors and co-activators in MCF7 cells. Shown are mRNA levels of transcription factors and co-activators in MCF7 cells, after stable expression of a control shRNA (sh Ctrl) or an shRNA against the indicated factor. *GAPDH* mRNA level serve as loading control.

Figure S5. shRNA screen identifies transcription factors and co-activators that regulate cleavage activity of the PTEN PPAS.

(A) Shown is the influence of KD of individual transcription factors or co-activators on PTEN PPAS activity. KD of FIP1L1 (green; polyadenylation factor) serves as positive control. T-test for independent samples was performed; **, $P < 0.001$; *, $P < 0.02$. See Table S2 for values.

(B) As in (B), but here the influence of KD of individual transcription factors or co-activators on transcriptional activity of the Penh-Pprom reporter is shown. *, indicates a change in transcriptional activity > 2-fold, compared to the average of the ctrls.

(C) Shown is cleavage activity of additional PAS after KD of specific transcription factors (as in Figure 5D). T-test for independent samples was performed; *, $P < 0.01$.

(D) Shown is cleavage activity of the PTEN PPAS in the context of the Pprom reporter comparing the effect of KD of specific histone acetyl transferases with a ctrl KD. T-test for independent samples was performed.

(E) As in (D), but cleavage activity of the PTEN PPAS expressed from the Nprom reporter was examined.

Methods

Luciferase constructs

All luciferase constructs were derived from PIS1 vector (Mayr and Bartel, 2009). It contains the thymidine kinase promoter of Herpes simplex virus (HSV-TK), followed by a renilla luciferase open reading frame, followed by the late SV40 PAS. To obtain reporter constructs with different promoters or PAS, the HSV-TK promoter or the SV40 PAS were exchanged using restriction enzyme digest or Gibson cloning. To obtain reporter constructs with different enhancers, the enhancer sequences (Table S1) were cloned upstream in the sense orientation of the respective promoters. In the case of Penh1, the reverse complement of the sequence was cloned downstream of the PAS. MYC-binding sites (E-boxes with the sequence CACGTG at positions -1149 bp and -1353 bp upstream of the transcription start site of the *PTEN* gene) were mutated to CAAGAA using Quikchange Lightening kit (Agilent).

Luciferase assay

For all experiments, the human breast cancer cell line MCF7 was used which was a gift from the laboratory of Robert Weinberg (Whitehead Institute, Cambridge, USA). Luciferase assays were performed after transfection of constructs with Lipofectamine 2000 in 24-well plates as described previously (Mayr and Bartel, 2009). Each experiment was performed in triplicates. The number of experiments listed in the figures corresponds to biological replicates. In each well, 100 ng of firefly luciferase control reporter plasmid PISO together with 400 ng of renilla luciferase plasmid were transfected. Same molar amounts of plasmid were transfected to account for different construct sizes (400 ng were used for a plasmid of 5000 bp). Firefly and renilla luciferase activities were measured with the Dual-luciferase assay (Promega) 24 hours after transfection. Renilla activity was normalized to firefly activity to control for transfection efficiency. Transcriptional activity of a promoter corresponds to renilla luciferase activity after transfection of the reporter containing the promoter and the SV40 PAS. PAS cleavage activity of a test PAS was obtained by dividing the luciferase activities of the constructs with the test PAS by the SV40 PAS in the context of the same promoter. See legend of Figures S1B for further experimental details.

To assess PAS usage after knock-down of transcription factors, luciferase constructs were transfected into MCF7 cells stably expressing control (ctrl) shRNAs or shRNAs targeting specific

transcription factors or co-activators. PAS usage was calculated as described above. When several shRNAs against a specific factor were available, the results were pooled.

Knock-down of factors using shRNAs

shRNAs were designed using the siRNA selection program from the Whitehead Institute and cloned into pSUPERretropuro. Retroviral particles were obtained as described before (Mayr and Bartel, 2009). Knock-down efficiency was tested by RT-PCR with gene-specific primers and primers for GAPDH.

5' RACE

Renilla luciferase plasmids were transfected as described above and total RNA was extracted after 24 hours using TRIreagent. 5' RACE was performed with the 5' RACE kit (Roche) using gene-specific reverse primers: PTEN-5'RACE-R and R2 (for nested PCR).

Measurement of mRNA stability of the reporters

Luciferase plasmids were transfected as described above. After 24 h, cells were either treated with DMSO or with Actinomycin D (4 µg/ml; Sigma) for the indicated time points. Total RNA was extracted using TRIreagent and was used to generate cDNA using qScript. qRT-PCR was performed using the Roche SYBR Green mastermix on a 7500 HT Fast Real-Time PCR System (Applied Biosystems). Each reaction was performed in triplicate. The experiments were performed at least three times to obtain at least three biological replicates. The primer pairs used to quantify total *PTEN* mRNA (PTEN-stability-F and -R), luciferase mRNA (Rluc-stability-F and -R) and *GAPDH* mRNA (GAPDH-F and GAPDH-stability-R) (for normalization) are listed in the Key Resource Table.

ChIP-seq data sets for MCF7 cells

Binding of MYC to the PTEN promoter was assessed using published ChIP-seq data (GSE33213). The levels of acetylated H3K27 in MCF7 cells were visualized using published ChIP-seq data (GSM946850), generated by the Encode project (Frietze et al., 2012; Consortium, 2012).

Cultivation of MCF7 cells in acidified media

Cultivation of MCF7 cells for 6 days without media change results in acidification of the media (pH = 6.5) and increases the expression of the short 3'UTR isoform of *PTEN*. The 3'UTR ratio

change can be fully recapitulated by cultivating the cells for 24 hours in acidified media (pH = 6.5) using HCl.

Northern blotting

Northern blotting was performed as described previously with modifications (Mayr and Bartel, 2009). Briefly, total RNA was isolated using Trizol reagent (Invitrogen) and 2 µg of polyadenylated mRNA purified from the total RNA with Oligotex (Qiagen) was loaded in each lane. The single- stranded probe was generated by unidirectional PCR reaction as described (Marquardt et al., 2011) with a slight modification. As input for the unidirectional PCR 30 ng of DNA template were used. This consisted of a 700-bp fragment of PTEN coding region amplified from MCF7 cDNA using forward and reverse primers listed in the Key Resource table. The unidirectional PCR reaction (20 µl) was conducted in buffer (100 mM Tris-HCl/pH 8.3, 500 mM KCl, 15 mM $MgCl_2$, 1% Triton-X-100) containing 0.2 mM each of dCTP, dGTP, and dTTP, 0.5 unit of Taq polymerase, the reverse primer from above, and 6 µl of 3000 Ci/mmol [α-^{32}P]dATP (Perkin Elmer). The reaction mixture was initially boiled for 10 min at 95°C and subjected to 35 thermal cycles (95°C for 30 s; 45°C for 30 s; 72°C for 1 min), which was followed by 5-min incubation at 72°C. After the PCR reaction, 5 µl of 0.2 mM EDTA was added to the mixture and boiled for 5 min at 95°C, followed by 2-min chilling on ice. The PCR product was then used for hybridization to probe PTEN transcript.

Deletion of the *PTEN* enhancer using CRISPR-Cas9

The guide RNAs targeting the 5′ and 3′ ends of the *PTEN* enhancer were designed using GuideScan and cloned into pX330 as previously described (Ran et al., 2013; Perez et al., 2017). MCF7 cells were plated on 24-well plates. The next day, the cells were transfected using Lipofectamine 2000 with pmaxGFP and the two pX330 plasmids. After 4 days, GFP-positive cells were sorted into 96-well plates and grown into colonies. Genomic DNA was isolated from each clone and the presence of the *PTEN* alleles was assessed by PCR and sequencing (Figure S3). As MCF7 cells are not diploid at the *PTEN* locus, the presence of *PTEN* enhancer alleles was assessed by qPCR using genomic DNA isolated from MCF7 WT and enhancer deletion cell lines to amplify two regions outside the *PTEN* enhancer (chr10:89,618,718-89,618,843 and chr10:89,620,263-89,620,387) and two regions inside *PTEN* enhancer (chr10:89,621,562-89,621,699 and chr10:89,621,320-89,621,440) in FastStart universal SYBR green master mix (Roche). The qPCR results from the amplicons inside the *PTEN* enhancer were normalized to those from the amplicons outside of the *PTEN* enhancer and were

compared between WT and enhancer deletion cells to determine the extent of loss of the enhancer.

PTEN mRNA levels

PTEN mRNA levels were assessed using qRT-PCR. *PTEN* mRNA was normalized to that of *RPL19* mRNAs.

Western blotting

Western blots were lysed in 2x Laemmli buffer (Alfa Aesar) and performed as described previously (Lee and Mayr, 2019) with the following antibodies: anti-PTEN (1:1000; Santa Cruz Biotechnology), anti-AKT (1:1000; Cell Signaling), anti-phospho AKT (S473) (1:1000; Cell Signaling), and anti-GAPDH (1:500; Santa Cruz Biotechnology). As secondary antibodies anti-mouse IRDye 800 (1:5000; Li-Cor Biosciences) and anti-rabbit IRDye 680 (1:5000; Li-Cor Biosciences) were used.

Immunoprecipitation of PTEN and determination of PTEN lipid phosphatase activity

PTEN was immunoprecipitated as described elsewhere (Papakonstanti et al., 2008) with a minor modification. The MCF7 cells (Ctrl KD, LU KD) were lysed in lysis buffer (50 mM HEPES/pH 7.4, 150 mM NaCl, 1.5 mM $MgCl_2$, 1 mM EGTA, 1 mM DTT, 10% glycerol, 1% Triton X-100, 1X Halt Protease Inhibitor Cocktail (Thermo) and the lysates were immunoprecipitated with anti-PTEN (Santa Cruz Biotechnology). The immune complexes were captured by PureProteome Protein A Magnetic Beads (Millipore) and washed twice with lysis buffer and then with lysis buffer containing 500 mM LiCl. After another two rounds of washing with 10 mM Tris-HCl/pH 8.0, 50 mM NaCl, the resulting immunoprecipitates were subjected to SDS-PAGE, followed by western blotting using anti-PTEN (Cell signaling) for quantification based on GST-PTEN recombinant proteins (Echelon) as a standard. 75-600 fmol of immunoprecipitated PTEN proteins were used to determine PTEN lipid phosphatase activity in vitro by PTEN ELISA kit (Echelon) following the manufacturer's instructions. A dilution series of GST-PTEN recombinant proteins was used as positive control and as standard throughout the ELISA procedure. As LU KD samples usually contained less total PTEN protein than Ctrl KD samples, western blots were used calculate factors to normalize protein amounts. These factors were then applied to normalize the absorbance measured in ELISA for LU KD samples in order to reflect similar input amounts of PTEN protein. Then, the difference in absorbance of ELISA between Ctrl KD and LU KD samples was converted into fold differences in PTEN activity, considering the ratio obtained using GST-PTEN.

3′UTR isoform quantification from bulk RNA-seq data

The following GEO datasets were used: Ery (definitive erythrocytes, EryD): GSE112717 and HSC: GSE120705 (Numata et al., 2020; Cai et al., 2020). FASTQ samples were obtained from the Sequence Read Archive (Ery: GSM3081989-GSM3081991; HSC: GSM3408762-GSM3408764) and pseudoaligned for transcript quantification in Salmon v1.3.0 (`salmon quant --gcBias --validateMappings -l A`) to the pre-compiled QAPA v1.3.0 mm10 3' UTR annotation with the full mm10 genome as decoy (Patro et al., 2017; Ha et al., 2018). Isoform usage was quantified using QAPA. Genes used in analysis were filtered based on Num_Events > 1 (more than one annotated isoform), at least 3 TPM for all samples, and a mean proximal isoform usage within 10%-90% for at least one cell type, where proximal isoforms are identified by APA_IDs ending in "_P" (Ha et al., 2018).

Differential gene expression testing between cell types [Ery/HSC] was performed with DESeq2 v1.28.0. Differential 3′UTR isoform usage testing [Ery/HSC] was performed using DEXSeq v1.34.0 (Anders et al., 2012; Love et al., 2014; Soneson et al., 2015). All analyses were performed in R v4.0.2 and Bioconductor v1.13.

Gene subsets were determined using 1) gene-level mRNA expression: two-fold "down", "same'", or two-fold "up", with 5% FDR (non-significant assigned to "same") using DESeq2 results; 2) the difference in short 3′UTR isoform usage (Figure 4F) is calculated as SU/(SU+LU) in Ery minus HSC: -10% "down", "same", or 10% "up", with 5% FDR (non-significant assigned to "same") using DEXSeq adjusted p-values and QAPA PPAU values [Ery – HSC]; and 3) short 3′UTR isoform expression (Figure 4H): two-fold "down", "same", or two-fold "up", using ratio [Ery/HSC] of mean normalized expression values for short 3′UTR isoforms from QAPA results. Normalized expression values were computed by rescaling QAPA TPM values per sample by estimated size factors from DESeq2. dSU is the fold change in short 3′UTR isoform expression, dLU is the fold change in expression of the sum of all 3′UTR isoforms that do not represent the shortest 3′UTR isoform.

Genes associated with Ery-specific enhancers were identified in Supporting Information Dataset_S01 from Cai et al., 2020 (Cai et al., 2020).
Our code will be available at: https://github.com/mfansler/utr-enhancers-pipeline

Statistics

For all pairwise comparisons of PAS usage or transcriptional activity a 2-tailed, 2-sample unequal variance t-test for independent samples was applied. When comparing several samples, a One-way ANOVA was performed. All tests were performed using SPSS.

References

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

Banerji, J., Rusconi, S., and Schaffner, W. (1981). Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. Cell *27*, 299-308.

Bentley, D.L. (2014). Coupling mRNA processing with transcription in time and space. Nature reviews Genetics *15*, 163-175.

Berkovits, B.D., and Mayr, C. (2015). Alternative 3' UTRs act as scaffolds to regulate membrane protein localization. Nature *522*, 363-367.

Bregman, A., Avraham-Kelbert, M., Barkai, O., Duek, L., Guterman, A., and Choder, M. (2011). Promoter elements regulate cytoplasmic mRNA decay. Cell *147*, 1473-1483.

Brumbaugh, J., Di Stefano, B., Wang, X., Borkent, M., Forouzmand, E., Clowers, K.J., Ji, F., Schwarz, B.A., Kalocsay, M., Elledge, S.J.*, et al.* (2018). Nudt21 Controls Cell Fate by Connecting Alternative Polyadenylation to Chromatin Signaling. Cell *172*, 629-631.

Cai, W., Huang, J., Zhu, Q., Li, B.E., Seruggia, D., Zhou, P., Nguyen, M., Fujiwara, Y., Xie, H., Yang, Z.*, et al.* (2020). Enhancer dependence of cell-type-specific gene expression increases with developmental age. Proc Natl Acad Sci U S A *117*, 21450-21458.

Calvo, O., and Manley, J.L. (2005). The transcriptional coactivator PC4/Sub1 has multiple functions in RNA polymerase II transcription. EMBO J *24*, 1009-1020.

Carswell, S., and Alwine, J.C. (1989). Efficiency of utilization of the simian virus 40 late polyadenylation site: effects of upstream sequences. Mol Cell Biol *9*, 4248-4258.

Chatrikhi, R., Mallory, M.J., Gazzara, M.R., Agosto, L.M., Zhu, W.S., Litterman, A.J., Ansel, K.M., and Lynch, K.W. (2019). RNA Binding Protein CELF2 Regulates Signal-Induced Alternative Polyadenylation by Competing with Enhancers of the Polyadenylation Machinery. Cell reports *28*, 2795-2806 e2793.

Chen, C.Y., Chen, S.T., Juan, H.F., and Huang, H.C. (2012). Lengthening of 3'UTR increases with morphological complexity in animal evolution. Bioinformatics *28*, 3178-3181.

Consortium, E.P. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57-74.

Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A.*, et al.* (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc Natl Acad Sci U S A *107*, 21931-21936.

Dantonel, J.C., Murthy, K.G., Manley, J.L., and Tora, L. (1997). Transcription factor TFIID recruits factor CPSF for formation of 3' end of mRNA. Nature *389*, 399-402.

Frietze, S., Wang, R., Yao, L., Tak, Y.G., Ye, Z., Gaddis, M., Witt, H., Farnham, P.J., and Jin, V.X. (2012). Cell type-specific binding patterns reveal that TCF7L2 can be tethered to the genome by association with GATA3. Genome biology *13*, R52.

Gasperini, M., Hill, A.J., McFaline-Figueroa, J.L., Martin, B., Kim, S., Zhang, M.D., Jackson, D., Leith, A., Schreiber, J., Noble, W.S.*, et al.* (2019). A Genome-wide Framework for Mapping Gene Regulation via Cellular Genetic Screens. Cell *176*, 377-390 e319.

Gasperini, M., Tome, J.M., and Shendure, J. (2020). Towards a comprehensive catalogue of validated and target-linked human enhancers. Nature reviews Genetics *21*, 292-310.

Gehring, N.H., Frede, U., Neu-Yilik, G., Hundsdoerfer, P., Vetter, B., Hentze, M.W., and Kulozik, A.E. (2001). Increased efficiency of mRNA 3' end formation: a new genetic mechanism contributing to hereditary thrombophilia. Nat Genet *28*, 389-392.

Grandori, C., Mac, J., Siebelt, F., Ayer, D.E., and Eisenman, R.N. (1996). Myc-Max heterodimers activate a DEAD box gene and interact with multiple E box-related sites in vivo. EMBO J *15*, 4344-4357.

Gruber, A.J., Schmidt, R., Gruber, A.R., Martin, G., Ghosh, S., Belmadani, M., Keller, W., and Zavolan, M. (2016). A comprehensive analysis of 3' end sequencing data sets reveals novel polyadenylation signals and the repressive role of heterogeneous ribonucleoprotein C on cleavage and polyadenylation. Genome Res *26*, 1145-1159.

Ha, K.C.H., Blencowe, B.J., and Morris, Q. (2018). QAPA: a new method for the systematic analysis of alternative polyadenylation from RNA-seq data. Genome biology *19*, 45.

Haimovich, G., Choder, M., Singer, R.H., and Trcek, T. (2013). The fate of the messenger is pre-determined: a new model for regulation of gene expression. Biochim Biophys Acta *1829*, 643-653.

Harel-Sharvit, L., Eldad, N., Haimovich, G., Barkai, O., Duek, L., and Choder, M. (2010). RNA polymerase II subunits link transcription and mRNA decay to translation. Cell *143*, 552-563.

Hathaway, N.A., Bell, O., Hodges, C., Miller, E.L., Neel, D.S., and Crabtree, G.R. (2012). Dynamics and memory of heterochromatin in living cells. Cell *149*, 1447-1460.

Higgs, D.R., Goodbourn, S.E., Lamb, J., Clegg, J.B., Weatherall, D.J., and Proudfoot, N.J. (1983). Alpha-thalassaemia caused by a polyadenylation signal mutation. Nature *306*, 398-400.

Hsin, J.P., and Manley, J.L. (2012). The RNA polymerase II CTD coordinates transcription and RNA processing. Genes Dev *26*, 2119-2137.

Jenal, M., Elkon, R., Loayza-Puch, F., van Haaften, G., Kuhn, U., Menzies, F.M., Oude Vrielink, J.A., Bos, A.J., Drost, J., Rooijers, K.*, et al.* (2012). The poly(A)-binding protein nuclear 1 suppresses alternative cleavage and polyadenylation sites. Cell *149*, 538-553.

Johnston, J.J., Williamson, K.A., Chou, C.M., Sapp, J.C., Ansari, M., Chapman, H.M., Cooper, D.N., Dabir, T., Dudley, J.N., Holt, R.J.*, et al.* (2019). NAA10 polyadenylation signal variants cause syndromic microphthalmia. J Med Genet *56*, 444-452.

Kwon, J., Lee, S.R., Yang, K.S., Ahn, Y., Kim, Y.J., Stadtman, E.R., and Rhee, S.G. (2004). Reversible oxidation and inactivation of the tumor suppressor PTEN in cells stimulated with peptide growth factors. Proc Natl Acad Sci U S A *101*, 16419-16424.

Lackford, B., Yao, C., Charles, G.M., Weng, L., Zheng, X., Choi, E.A., Xie, X., Wan, J., Xing, Y., Freudenberg, J.M.*, et al.* (2014). Fip1 regulates mRNA alternative polyadenylation to promote stem cell self-renewal. EMBO J *33*, 878-889.

Lee, S.H., and Mayr, C. (2019). Gain of Additional BIRC3 Protein Functions through 3'-UTR-Mediated Protein Complex Formation. Mol Cell *74*, 701-712 e709.

Legendre, M., and Gautheret, D. (2003). Sequence determinants in human polyadenylation site selection. BMC genomics *4*, 7.

Levine, M., Cattoglio, C., and Tjian, R. (2014). Looping back to leap forward: transcription enters a new era. Cell *157*, 13-25.

Li, D.M., and Sun, H. (1998). PTEN/MMAC1/TEP1 suppresses the tumorigenicity and induces G1 cell cycle arrest in human glioblastoma cells. Proc Natl Acad Sci U S A *95*, 15406-15411.

Lianoglou, S., Garg, V., Yang, J.L., Leslie, C.S., and Mayr, C. (2013). Ubiquitously transcribed genes use alternative polyadenylation to achieve tissue-specific expression. Genes Dev *27*, 2380-2396.

Lienert, F., Wirbelauer, C., Som, I., Dean, A., Mohn, F., and Schubeler, D. (2011). Identification of genetic elements that autonomously determine DNA methylation states. Nat Genet *43*, 1091-1097.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome biology *15*, 550.

Ma, W., and Mayr, C. (2018). A Membraneless Organelle Associated with the Endoplasmic Reticulum Enables 3'UTR-Mediated Protein-Protein Interactions. Cell *175*, 1492-1506 e1419.

Maehama, T., and Dixon, J.E. (1998). The tumor suppressor, PTEN/MMAC1, dephosphorylates the lipid second messenger, phosphatidylinositol 3,4,5-trisphosphate. J Biol Chem *273*, 13375-13378.

Mahat, D.B., Kwak, H., Booth, G.T., Jonkers, I.H., Danko, C.G., Patel, R.K., Waters, C.T., Munson, K., Core, L.J., and Lis, J.T. (2016). Base-pair-resolution genome-wide mapping of active RNA polymerases using precision nuclear run-on (PRO-seq). Nature protocols *11*, 1455-1476.

Marquardt, S., Hazelbaker, D.Z., and Buratowski, S. (2011). Distinct RNA degradation pathways and 3' extensions of yeast non-coding RNA species. Transcription *2*, 145-154.

Masamha, C.P., Xia, Z., Yang, J., Albrecht, T.R., Li, M., Shyu, A.B., Li, W., and Wagner, E.J. (2014). CFIm25 links alternative polyadenylation to glioblastoma tumour suppression. Nature *510*, 412-416.

Matsushima, W., Herzog, V.A., Neumann, T., Gapp, K., Zuber, J., Ameres, S.L., and Miska, E.A. (2018). SLAM-ITseq: sequencing cell type-specific transcriptomes without cell sorting. Development *145*.

Mayr, C. (2017). Regulation by 3'-Untranslated Regions. Annual review of genetics *51*, 171-194.

Mayr, C. (2019a). 3' UTRs Regulate Protein Functions by Providing a Nurturing Niche during Protein Synthesis. Cold Spring Harbor symposia on quantitative biology *84*, 95-104.

Mayr, C. (2019b). What Are 3' UTRs Doing? Cold Spring Harb Perspect Biol *11*.

Mayr, C., and Bartel, D.P. (2009). Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. Cell *138*, 673-684.

McCracken, S., Fong, N., Yankulov, K., Ballantyne, S., Pan, G., Greenblatt, J., Patterson, S.D., Wickens, M., and Bentley, D.L. (1997). The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. Nature *385*, 357-361.

Mendes-Pereira, A.M., Martin, S.A., Brough, R., McCarthy, A., Taylor, J.R., Kim, J.S., Waldman, T., Lord, C.J., and Ashworth, A. (2009). Synthetic lethal targeting of PTEN mutant cells with PARP inhibitors. EMBO molecular medicine *1*, 315-322.

Nagaike, T., Logan, C., Hotta, I., Rozenblatt-Rosen, O., Meyerson, M., and Manley, J.L. (2011). Transcriptional activators enhance polyadenylation of mRNA precursors. Mol Cell *41*, 409-418.

Ntranos, V., Yi, L., Melsted, P., and Pachter, L. (2019). A discriminative learning approach to differential expression analysis for single-cell RNA-seq. Nat Methods *16*, 163-166.

Numata, A., Kwok, H.S., Zhou, Q.L., Li, J., Tirado-Magallanes, R., Espinosa Angarcia, V., Hannah, R.L., Park, J., Wang, C.Q., Krishnan, V.*, et al.* (2020). Lysine acetyltransferase Tip60 is required for hematopoietic stem cell maintenance. Blood.

Orkin, S.H., Cheng, T.C., Antonarakis, S.E., and Kazazian, H.H., Jr. (1985). Thalassemia due to a mutation in the cleavage-polyadenylation signal of the human beta-globin gene. EMBO J *4*, 453-456.

Papakonstanti, E.A., Zwaenepoel, O., Bilancio, A., Burns, E., Nock, G.E., Houseman, B., Shokat, K., Ridley, A.J., and Vanhaesebroeck, B. (2008). Distinct roles of class IA PI3K isoforms in primary and immortalised macrophages. Journal of cell science *121*, 4124-4133.

Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods *14*, 417-419.

Perez, A.R., Pritykin, Y., Vidigal, J.A., Chhangawala, S., Zamparo, L., Leslie, C.S., and Ventura, A. (2017). GuideScan software for improved single and paired CRISPR guide RNA design. Nat Biotechnol *35*, 347-349.

Rambout, X., Dequiedt, F., and Maquat, L.E. (2018). Beyond Transcription: Roles of Transcription Factors in Pre-mRNA Splicing. Chemical reviews *118*, 4339-4364.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nature protocols *8*, 2281-2308.

Rosonina, E., Bakowski, M.A., McCracken, S., and Blencowe, B.J. (2003). Transcriptional activators control splicing and 3'-end cleavage levels. J Biol Chem *278*, 43034-43040.

Schek, N., Cooke, C., and Alwine, J.C. (1992). Definition of the upstream efficiency element of the simian virus 40 late polyadenylation signal by using in vitro analyses. Mol Cell Biol *12*, 5386-5393.

Shlyueva, D., Stampfel, G., and Stark, A. (2014). Transcriptional enhancers: from properties to genome-wide predictions. Nature reviews Genetics *15*, 272-286.

Slobodin, B., Han, R., Calderone, V., Vrielink, J., Loayza-Puch, F., Elkon, R., and Agami, R. (2017). Transcription Impacts the Efficiency of mRNA Translation via Co-transcriptional N6-adenosine Methylation. Cell *169*, 326-337 e312.

Smale, S.T. (2010). Luciferase assay. Cold Spring Harbor protocols *2010*, pdb prot5421.

Soneson, C., Love, M.I., and Robinson, M.D. (2015). Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000Research *4*, 1521.

Spies, N., Burge, C.B., and Bartel, D.P. (2013). 3' UTR-isoform choice has limited influence on the stability and translational efficiency of most mRNAs in mouse fibroblasts. Genome Res *23*, 2078-2090.

Stacey, S.N., Sulem, P., Jonasdottir, A., Masson, G., Gudmundsson, J., Gudbjartsson, D.F., Magnusson, O.T., Gudjonsson, S.A., Sigurgeirsson, B., Thorisdottir, K.*, et al.* (2011). A germline variant in the TP53 polyadenylation signal confers cancer susceptibility. Nat Genet *43*, 1098-1103.

Tian, B., and Manley, J.L. (2017). Alternative polyadenylation of mRNA precursors. Nature reviews Molecular cell biology *18*, 18-30.

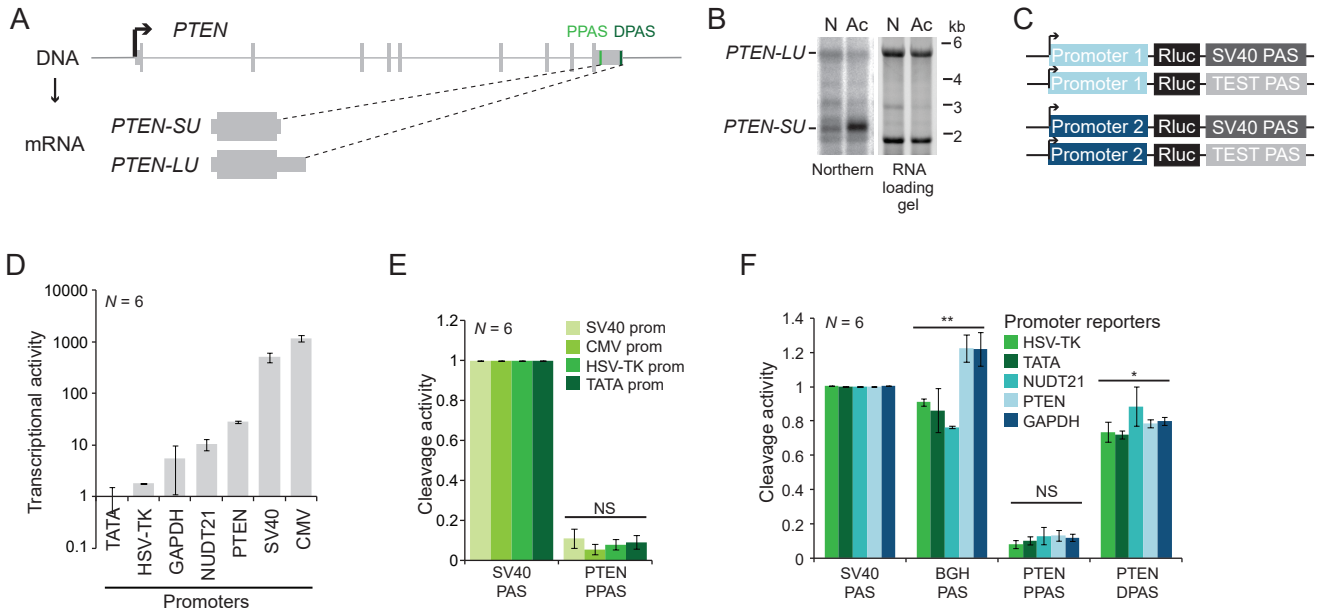
Trcek, T., Larson, D.R., Moldon, A., Query, C.C., and Singer, R.H. (2011). Single-molecule mRNA decay measurements reveal promoter- regulated mRNA stability in yeast. Cell *147*, 1484-1497.

Worby, C.A., and Dixon, J.E. (2014). Pten. Annu Rev Biochem *83*, 641-669.

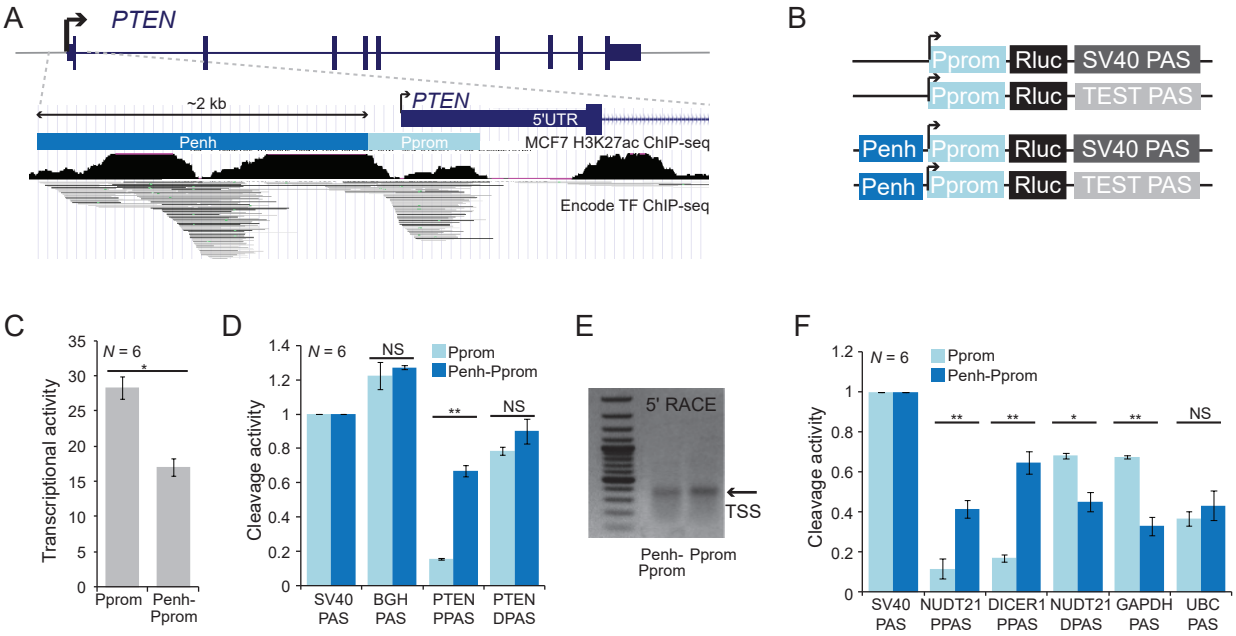
Zhu, Y., Wang, X., Forouzmand, E., Jeong, J., Qiao, F., Sowd, G.A., Engelman, A.N., Xie, X., Hertel, K.J., and Shi, Y. (2018). Molecular Mechanisms for CFIm-Mediated Regulation of mRNA Alternative Polyadenylation. Mol Cell *69*, 62-74 e64.

Zid, B.M., and O'Shea, E.K. (2014). Promoter sequences direct cytoplasmic localization and translation of mRNAs during starvation in yeast. Nature *514*, 117-121.

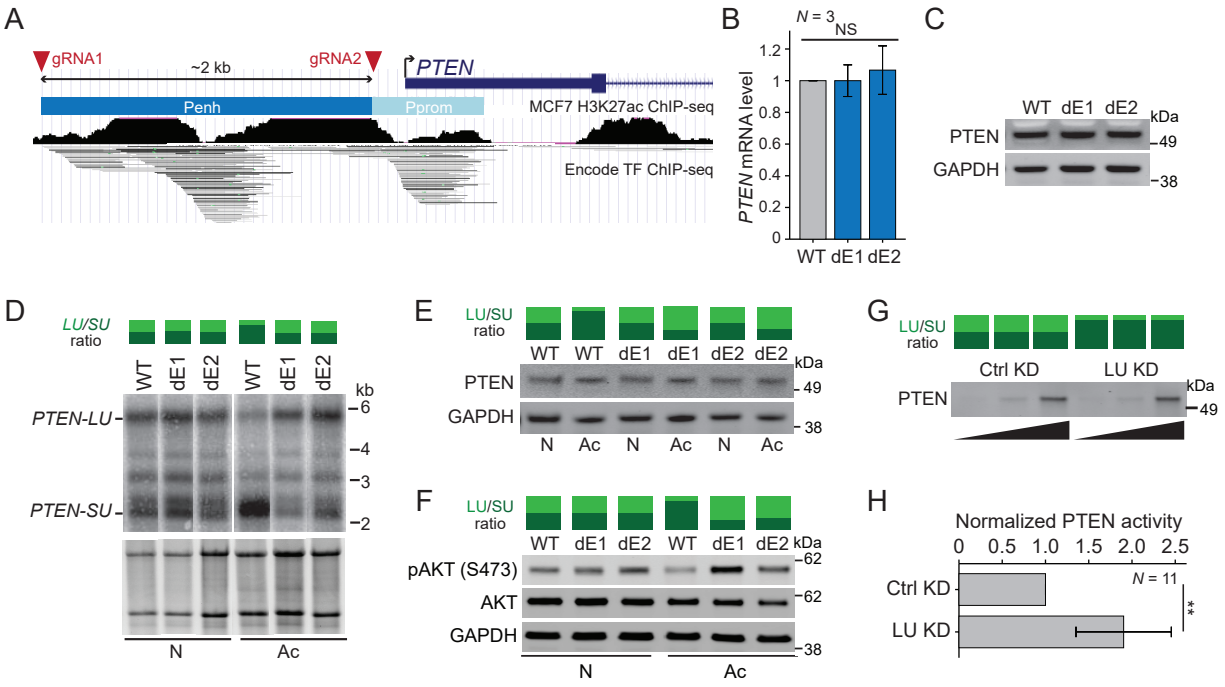
Kwon, Figure 1



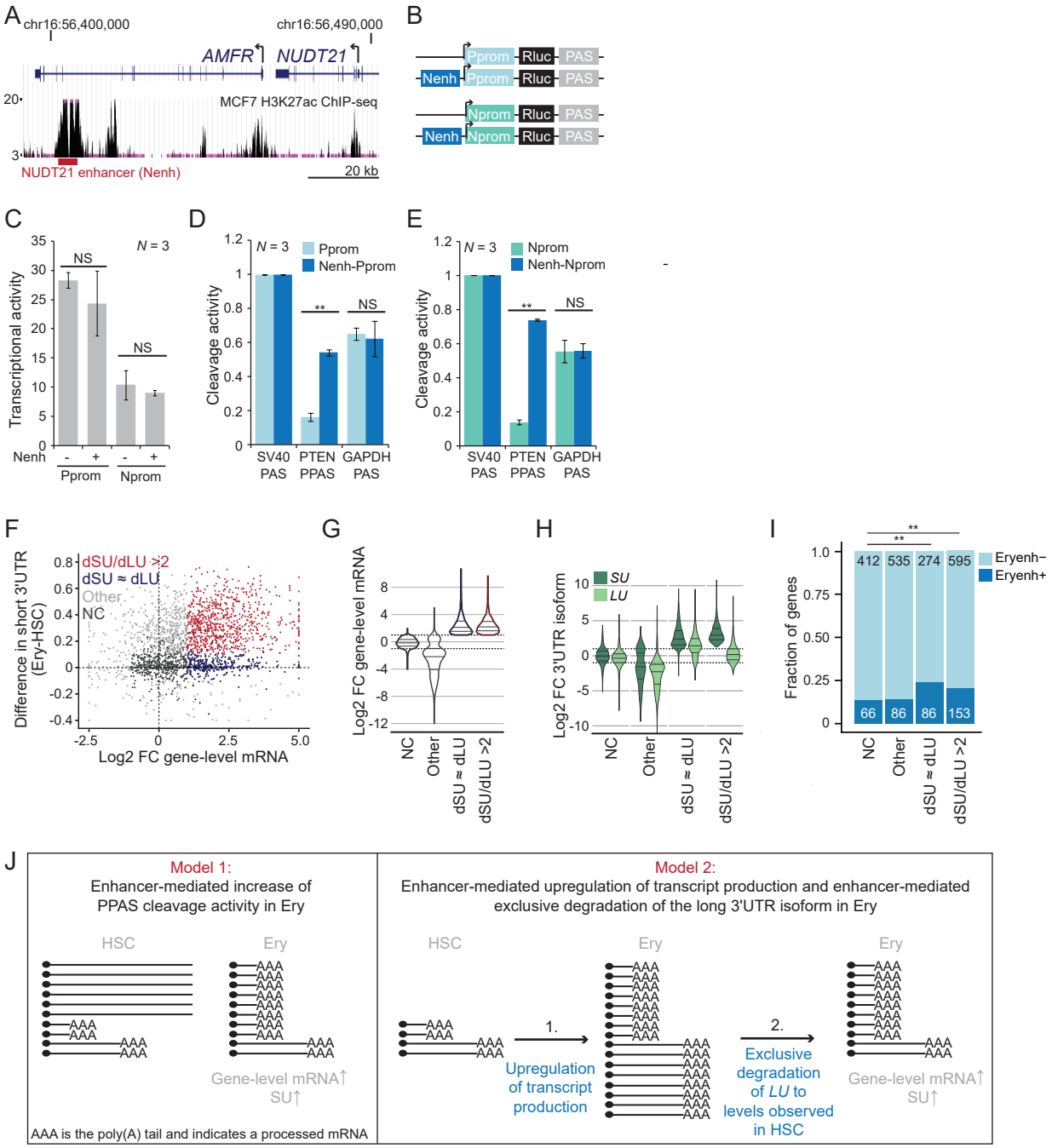
Kwon, Figure 2



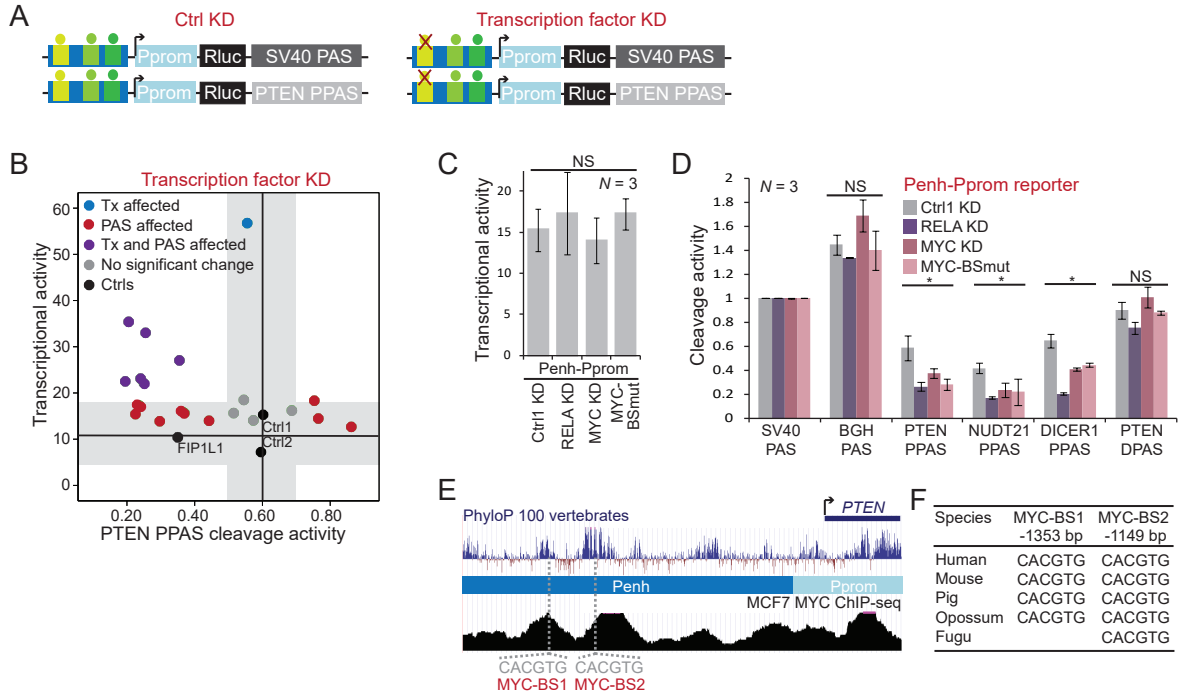
Kwon, Figure 3

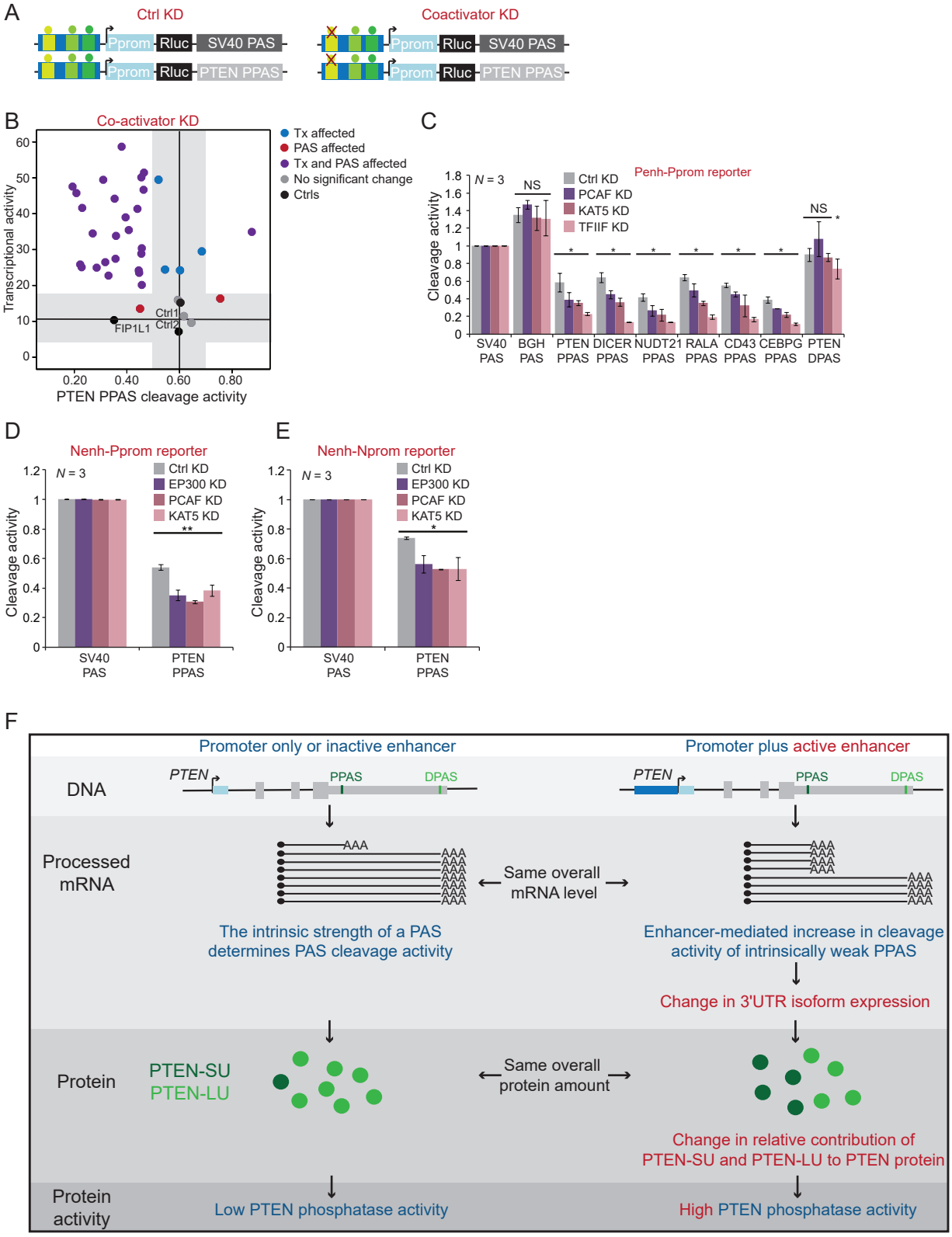


Kwon, Figure 4



Kwon, Figure 5



Species	MYC-BS1 -1353 bp	MYC-BS2 -1149 bp
Human	CACGTG	CACGTG
Mouse	CACGTG	CACGTG
Pig	CACGTG	CACGTG
Opossum	CACGTG	CACGTG
Fugu		CACGTG

Kwon, Figure 6