Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Cell Rep. Author manuscript; available in PMC 2019 September 19.

Published in final edited form as:

Cell Rep. 2019 September 10; 28(11): 2795–2806.e3. doi:10.1016/j.celrep.2019.08.022.

RNA Binding Protein CELF2 Regulates Signal-Induced Alternative Polyadenylation by Competing with Enhancers of the Polyadenylation Machinery

Rakesh Chatrikhi[1], **Michael J. Mallory**[1], **Matthew R. Gazzara**[1,2], **Laura M. Agosto**[1,3], **Wandi S. Zhu**[4], **Adam J. Litterman**[4], **K. Mark Ansel**[4], **Kristen W. Lynch**[1,3,5,*]

[1]Department of Biochemistry and Biophysics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

[2]Graduate Group in Genomics and Computational Biology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

[3]Graduate Group in Biochemistry and Molecular Biophysics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

[4]Department of Microbiology and Immunology, UC San Francisco, San Francisco, CA 94143, USA

[5]Lead Contact

SUMMARY

The 3′ UTR (UTR) of human mRNAs plays a critical role in controlling protein expression and function. Importantly, 3′ UTRs of human messages are not invariant for each gene but rather are shaped by alternative polyadenylation (APA) in a cell state-dependent manner, including in response to T cell activation. However, the proteins and mechanisms driving APA regulation remain poorly understood. Here we show that the RNA-binding protein CELF2 controls APA of its own message in a signal-dependent manner by competing with core enhancers of the polyadenylation machinery for binding to RNA. We further show that CELF2 binding overlaps with APA enhancers transcriptome-wide, and almost half of 3′ UTRs that undergo T cell signaling-induced APA are regulated in a CELF2-dependent manner. These studies thus reveal CELF2 to be a critical regulator of 3′ UTR identity in T cells and demonstrate an additional mechanism for CELF2 in regulating polyadenylation site choice.

In Brief

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

*Correspondence: klync@pennmedicine.upenn.edu.

AUTHOR CONTRIBUTIONS

R.C. performed the majority of the experimental work, with assistance from M.J.M. W.S.Z., A.J.L., and K.M.A. generated the CRISPR KO clone. M.R.G. carried out the CLIP and DaPars analysis in Figures 5 and 6A. L.M.A. performed the proteomic analysis in Figure 7. R.C. and K.W.L. directed the project and wrote the manuscript.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.celrep.2019.08.022.

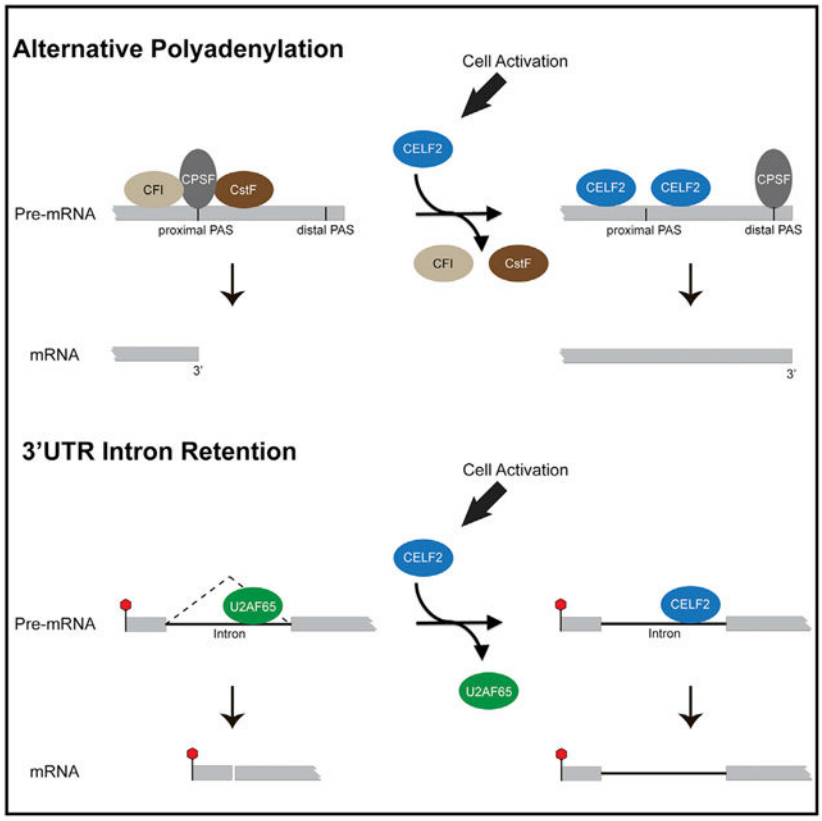
DECLARATION OF INTERESTS

The authors declare no competing interests.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Alternative polyadenylation (APA) is broadly regulated during cellular activation. Chatrikhi et al. demonstrate that the RNA-binding protein CELF2 competes with CFIm25 and CstF64 for binding around polyadenylation sites. Increased expression of CELF2 upon cellular activation alters this competition and is a key driver of activation-induced APA.

Graphical Abstract



INTRODUCTION

Protein expression in cells is controlled not only by regulation of transcription, splicing, and export of coding mRNAs but also through the presence or absence of regulatory sequences located in the 5′ and 3′ UTRs of mRNA. In particular, 3′ UTRs have been shown to be the primary site of microRNA (miRNA) association and function (Bartel, 2009; Mayr, 2017; Tian and Manley, 2017) as well as containing binding sites for proteins that regulate translation, mRNA stability, and subcellular localization of mRNA (Mayr, 2017; Tian and Manley, 2017). The interaction of such regulatory factors with *cis*-acting elements in the 3′ UTR controls the level of protein expression, as well as protein phosphorylation, localization, and activity, by determining where and when a protein is translated in the cell (Basu et al., 2011; Berkovits and Mayr, 2015; Ma and Mayr, 2018). Therefore, the identity of the 3′ UTR plays a critical role in determining the ultimate expression and function of the upstream open reading frame.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Importantly, the 3′ UTR of a given gene is not static but rather can vary in a tissue- and condition-specific manner through at least two mechanisms: alternative splicing and alternative polyadenylation. Alternative splicing within 3′ UTRs, including retention of introns, has been reported in 6%–16% of human protein-coding genes, although this may be an underestimate (Bicknell et al., 2012). Splicing in 3′ UTRs is generally thought to target messages for nonsense-mediated decay and has not been broadly studied (Bicknell et al., 2012; Lejeune and Maquat, 2005). In contrast, much attention has recently been focused on alternative polyadenylation, which occurs over half of human protein-coding genes (Derti et al., 2012; Hoque et al., 2013; Tian et al., 2005; Tian and Manley, 2017) and can alter protein expression or function by altering the presence or absence of binding sites within 3′ UTRs for miRNAs and regulatory proteins (Mayr, 2017; Sandberg et al., 2008; Tian and Manley, 2017).

Polyadenylation is the addition of ~200 adenosine bases to the exposed 3′ hydroxyl group generated by co-transcriptional cleavage of a nascent transcript. Alternative polyadenylation (APA) refers to cases in which cleavage, and subsequent polyadenylation, occurs at different positions on a transcript in a condition or cell type-specific manner (Tian and Manley, 2017). Such APA results in truncation of the open reading frame if cleavage occurs prior to the final coding exons. On the other hand, when both or all of the cleavage events are downstream of the stop codon, APA results in 3′ UTRs of distinct lengths.

Cleavage and polyadenylation is mediated by a large multi-component complex consisting of the core CPSF (cleavage and polyadenylation stimulation factor) subcomplex as well as additional enhancer subcomplexes CFI (cleavage factor I), CFII (cleavage factor II), and CstF (cleavage stimulatory factor) (Shi and Manley, 2015). These four subcomplexes (also known as the cleavage and polyadenylation factors [CPAFs]) associate in a cooperative manner with genomically encoded sequence motifs comprising a hexameric AAUAAA (also known as the polyadenylation sequence [PAS]) and upstream and downstream U/G-rich sequences. The AAUAAA hexamer is the primary binding site for the CPSF complex and directs the location of cleavage, whereas CFI binds to an upstream UGUA and CstF associates with a U/G-rich element downstream of the core hexamer (Shi and Manley, 2015). Importantly, because each of these interactions is relatively weak, the cooperative assembly of CPSF with CFI and CstF helps ensure efficiency of binding and accurate positioning of the overall complex (Tian and Manley, 2017). Consistently, several groups have shown that increased expression or recruitment of CFI and CstF can direct the catalytic activity of CPSF to regions of the transcript that lack a perfect AAUAAA element, thereby enhancing polyadenylation at an otherwise suboptimal (“weak”) site (Brumbaugh et al., 2018; Martin et al., 2012; Takagaki et al., 1996; Tian and Manley, 2017; Zhu et al., 2018). Additional RNA-binding proteins (RBPs) can also promote or hinder the association of CPAFs with particular sites through protein-protein interactions or steric hindrance, respectively (Tian and Manley, 2017). Such regulation of the binding of CPAFs to low- and high-affinity sites, through changes in the expression of individual CPAFs or regulatory RBPs, is thought to drive the majority of APA (Tian and Manley, 2017).

APA has been shown to be broadly regulated in a cell type-and signal-responsive manner, including in response to cellular proliferation. For example, in one study, approximately 5%

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

of genes surveyed undergo APA during antigen-induced proliferative growth of human CD4+ T cells (Sandberg et al., 2008), a required step in a proper immune response. However, only a very few proteins have thus far been identified as essential for regulating the widespread and coordinated APA that has been observed in T cells or other examples of signal-induced APA. We have previously shown that expression of the RBP CELF2 is induced upon stimulation of the Jurkat T cell line and binds broadly to exons, introns, and 3′ UTRs within these cells (Ajith et al., 2016; Mallory et al., 2011). CELF2 is a member of the CUG-BP, Elav-like family (CELF) of binding proteins. It binds preferentially to single-stranded UG and CUG-rich RNA elements and has been widely studied as a regulator of splicing in heart and skeletal muscle, neurons, and T cells (Dasgupta and Ladd, 2012; Gazzara et al., 2017; Ladd et al., 2001). We have previously shown that increased CELF2 protein activity in response to T cell signaling drives alternative splicing of more than 100 genes (Gazzara et al., 2017; Mallory et al., 2015). However, the functional consequence of the pervasive binding of CELF2 to 3′ UTRs has not been explored.

Here we show that CELF2 regulates both 3′ UTR intron retention and APA on its own transcript as well as many others. The regulation of intron retention involves competition between CELF2 and the 65 kDa subunit of the U2-auxiliary factor (U2AF65) for binding to the polypyrimidine tract, consistent with previous studies demonstrating steric hindrance of 3′ splice site use by CELF2 (Dembowski and Grabowski, 2009). In contrast, a widespread role for CELF2 in APA has not previously been described. We show here that CELF2 mediates APA by competing with the CFI and CstF complexes for substrate binding. Increased expression of CELF2 induced upon stimulation of Jurkat T cells results in decreased use of a weak (non-canonical) PAS site in the CELF2 3′ UTR, while mutation of the CFI and CstF binding sites around this weak PAS mimics stimulation and reduces the impact of CELF2 expression. Importantly, transcriptome-wide binding of CELF2 in the 3′ UTR markedly overlaps the binding sites for CFI and CstF. Moreover, we show that knockdown of CELF2 in Jurkat cells abrogates the stimulation-induced APA of many mRNAs, resulting in regulation of mRNA and protein expression. In particular, we demonstrate that CELF2-regulated APA drives CELF2-dependent repression of RBFox2 expression, which we have previously shown to shape splicing programs in many cell types (Gazzara et al., 2017). Together, our data demonstrate a previously unrecognized activity of CELF2 in shaping 3′ UTR identity and suggest that regulation of CELF2 expression may underlie much of the APA that has been previously observed in proliferating cells.

RESULTS

CELF2 Binds Its Own 3′ UTR and Induces Intron Retention and APA

Our previous transcriptome-wide mapping of CELF2 binding to RNA in the Jurkat T cells revealed the surprising discovery that CELF2 binding is enriched in 3′ UTRs (Ajith et al., 2016) (Figure 1A). Although CELF2 has been described primarily as a regulator of splicing (Dasgupta and Ladd, 2012; Ladd et al., 2001), the extensive 3′ UTR binding suggests further undetermined roles of CELF2 in RNA biogenesis. To begin to investigate the functional impact of CELF2 3′ UTR binding, we looked for 3′ UTRs that exhibited strong binding and potential regulation in the Jurkat T cells.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Interestingly, one of the 3′ UTRs in which we observe extensive binding of CELF2 protein by crosslinking immunoprecipitation sequencing (CLIP-seq) is the 3′ UTR of the CELF2 mRNA itself (Figure 1B). Moreover, we have previously shown that both the retention of an intron and use of competing APA sites (PAS2 versus PAS3) is altered in the CELF2 3′ UTR upon stimulation of Jurkat cells with the phorbol ester PMA (Mallory et al., 2015) (Figure 1B). Importantly, PMA stimulation of Jurkat cells also results in a significant increase in CELF2 protein expression through an increase in both transcription and mRNA stability (Mallory et al., 2011, 2015). Analysis of protein expression, 3′ UTR intron retention (IR), and APA across a time course of stimulation revealed that the predominant spike in CELF2 protein occurs in the first 8–12 h of PMA stimulation, whereas significant changes in IR or APA of the CELF2 3′ UTR are not observed until 12–16 hours after PMA treatment (Figures 1C and 1D). The fact that increased CELF2 protein precedes a change in IR or APA suggests a role of CELF2 protein in autoregulation of its own 3′ UTR. To more directly test the causality of CELF2 protein expression in 3′ UTR regulation, we used a doxycycline-inducible promoter driving a CELF2 cDNA expression construct. Strikingly, we find that doxycycline induction of CELF2 protein is sufficient to induce both IR and APA in the endogenous CELF2 3′ UTR, strongly suggesting that CELF2 protein directly regulates IR and APA in the CELF2 3′ UTR (Figures 1E and S1).

CELF2 Is Necessary and Sufficient for IR by Inhibiting U2AF65 Binding

We first sought to further investigate the regulation of IR by CELF2, as this is consistent with known activities of CELF2. We generated a CELF2-knockout (KO) Jurkat cell line by mutation of exon 5 by CRISPR (see STAR Methods) to introduce a stop codon, resulting in cells that lack any detec] CELF2 protein but retain near normal mRNA expression (Figure 2A). In this cell line we observe no increased CELF2 3′ UTR IR upon treatment with PMA, demonstrating that CELF2 protein is necessary for PMA-induced retention of the 3′ UTR intron (Figure 2A). We observe a similar loss of 3′ UTR IR in Jurkat cells that are depleted of CELF2 by short hairpin RNA (shRNA) (Figure S2), although in these cells there is less CELF2 mRNA as well as less protein, thus hindering robust conclusions. To determine, conversely, if CELF2 is sufficient for IR in the absence of other T cell factors, we used HeLa cells that naturally lack any detectable CELF2 protein (Figure 2B). Consistent with a causative role for CELF2 in IR, we detect no IR in RNA from a CELF2 3′ UTR intron reporter construct transfected into HeLa cells either on its own or co-transfected with cDNA encoding GFP, while cotransfection with CELF2 cDNA induces almost complete IR (Figure 2B). Together, these data demonstrate that CELF2 protein is necessary and sufficient to induce retention of its own 3′ UTR intron.

Previous studies have shown that CELF2 represses 3′ splice site use by competing with U2AF65 (Dembowski and Grabowski, 2009). Close examination of the CLIP-seq data reveals that CELF2 CLIP peaks exactly overlap the 3′ polypyrimidine track of the 3′ UTR intron (Figure 2C), which is the sequence to which U2AF65 binds to promote splicing (Agrawal et al., 2016). To ask whether CELF2 competes with U2AF65 for binding to the 3′ UTR intron, we carried out a UV crosslinking experiment using a radiolabeled RNA corresponding to 120 nt around the 3′ splice site of the 3′ UTR intron and nuclear extract from unstimulated and PMA-stimulated Jurkat cells (Figure 2D). Consistent with the CLIP

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

data, we find that both U2AF65 and CELF2 crosslink to the 3′ UTR intron in extract from unstimulated cells. By comparison, in extracts from stimulated cells, CELF2 crosslinking increases while U2AF65 crosslinking decreases (Figure 2D). We conclude that the reduction in U2AF65 crosslinking is a direct result of increased CELF2 expression, as we observe a similar loss of U2AF65 signal when we add recombinant CELF2 protein to extract from unstimulated cells (Figure 2E). Although in many instances competition between CELF2 and U2AF65 lead to exon skipping, repression of the 3′ splice site of a terminal intron is predicted to result in IR, as no alternative 3′ splice site is available (see Discussion). Interestingly, in our previously published RNA sequencing (RNA-seq) data from wild-type and CELF2-depleted JSL1 cells (Gazzara et al., 2017), we find several instances of CELF2-dependent IR events in 3′ UTRs (Table S1; Figure S2). Although IR is not a focus of the present study, these data further highlight that the previously described role of CELF2 in 3′ splice site regulation can result in IR as well as exon skipping.

CELF2 Is Necessary and Sufficient for the Switch to PAS3 by Inhibiting Enhancer Elements Required for PAS2

In contrast to its well-documented activity as a splicing regulator, CELF2 has not previously been shown to control cleavage and polyadenylation. To determine if, in addition to IR, CELF2 is necessary and sufficient to directly regulate APA of its 3′ UTR, we again used the KO, knockdown, and HeLa systems described above (Figure 2). As we observed for IR, we find that the PMA-induced switch from PAS2 to PAS3 is abolished in the CELF2-KO and CELF2-knockdown Jurkat cells (Figure 3A; Figure S3). Moreover, PAS3 is only used in HeLa cells upon co-transfection with CELF2 (Figure 3B). Importantly, the use of PAS3 is not dependent on the IR activity of CELF2, as the CELF2-dependent use of PAS3 is equivalent in 3′ UTR PAS reporters that contain or lack the intron (Figure 3B, +Intron versus−Intron). Therefore, we conclude that CELF2 directly activates an APA switch from PAS2 to PAS3.

To begin to understand how CELF2 may regulate APA, we analyzed the sequence features around PAS2 and PAS3. PAS3 has a canonical AAUAAA hexamer, whereas PAS2 has a divergent AUUAAA motif (Figure 3C). The AUUAAA motif is known to recruit the CPSF complex less efficiently than the canonical sequence (Sheets et al., 1990). However, we also note the presence of two UGUA enhancer elements 40–50 nt upstream of the PAS2 hexamer (Figure 3C, box), as well as an extended U/G-rich enhancer element downstream (Figure 3C, underline). The upstream enhancer (UE) UGUA is the binding site for CFIm25, the RNA-binding subunit of the CFI complex (Yang et al., 2010), while the downstream enhancer (DE) is bound by CstF64, the RNA-binding subunit of CstF. Both CFI and CstF facilitate recruitment of the CPSF complex through protein-protein interactions (Shi and Manley, 2015). Notably, both the UE and DE sequences closely mirror the U/G-rich binding preference of CELF2. We also observe that the PAS2-associated UGUA motifs and U/G-rich enhancer fall directly within CELF2 CLIP peaks (Figure 3C). Therefore, we hypothesized that CELF2 may compete with CFIm25 or CstF64 for binding upstream or downstream of PAS2, respectively, thus decreasing the efficiency of PAS2 recognition to favor PAS3.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

To directly test if CELF2 inhibits the binding of CFIm25 and/or CstF64 around PAS2, we carried out UV crosslinking using recombinant proteins and an RNA corresponding to PAS2 and approximately 85 nt of upstream and downstream sequence (see STAR Methods). Consistent with our identification of binding sites for CELF2, CFim25 and CstF64 within this region of the CELF2 3′ UTR, we readily observe crosslinking of all three individual proteins to the RNA (Figures 4A and 4B). Moreover, when CstF64 and CFIm25 are co-incubated with RNA, crosslinking of both of these proteins is observed (Figure S4). By contrast, addition of CELF2 results in reduced binding of CFIm25 and CstF64, in a dose-dependent manner, with a concomitant increase in binding of CELF2. This competition in binding is observed whether CFIm25 and CstF64 are alone or in combination with one another (Figures 4A and 4B; Figure S4A). As a control, no competition in observed between CFIm25/CstF64 and hnRNP L or hnRNP K (Figures 4A and 4B; Figure S4B) despite the fact that these proteins have been observed to bind to 3′ UTRs and C-rich elements, respectively (Figure S4). We also note that CELF2 competitively inhibits the binding of CFIm25 and CstF64 even when these proteins are present in molar excess, suggesting that CELF2 binds to this RNA with higher affinity than either CFIm25 or CstF64.

To provide further evidence of competition between CELF2 and CFIm25/CstF64 in cells, we also carried out mutagenesis of the PAS reporter construct used in Figure 3B. Specifically we engineered mutations that disrupt binding of CFIm25, CstF64, and CELF2 around PAS2 and tested the impact of these mutations on use of PAS2 and PAS3 (Figure 4C). Consistent with the prediction that use of PAS2 requires cooperative enhancing activity from CFI and CstF, mutation of either the CFIm25 (mUE) or CstF64 (mDE) binding site markedly reduced PAS2 use in favor of PAS3 in the absence of CELF2 (Figure 4C). In both cases, CELF2 expression still resulted in a further increase in PAS3, presumably because of competition of CELF2 with the non-mutated element. However, mutation of both the UE and DE (mU/D) surrounding PAS2 is sufficient to flip APA fully to PAS3 and abrogate any impact of CELF2 addition (Figure 4C). A similar, though less striking, result is also observed in constructs in which mutations are made that disrupt CFIm25 and CstF64 binding but maintain CELF2 binding (mU/D-C2; Figure S4). We can also rule out the alternative possibility that CELF2 also directly enhances PAS3, given that we observe a similar CELF2-induced reduction in PAS2 when PAS3 is replaced by the heterologous, unregulated PAS site from the reporter vector (Figure 4C, swPAS3; see Figure S4 for quantification). Moreover, improving the strength of PAS2 by inserting a consensus AAUAAA motif also increased the relative use of PAS2 over PAS3 in both the absence and presence of CELF2 (Figure 4C). Thus, from both the biochemical and cellular data, we conclude that the primary mechanism by which CELF2 causes a switch from PAS2 to PAS3 is by inhibiting the use of PAS2 through competing with factors such as CFIm25 and CstF64 for binding to the required enhancer elements. This model is further consistent with the lack of CELF2 binding downstream of PAS3 and the fact that in the absence of CELF2 protein, PAS2 is used constitutively.

CELF2 Binding Overlaps Extensively with APA Enhancer Activities Transcriptome-wide

We next wanted to test the possibility that CELF2 also drives a broader program of activation-induced APA. As discussed in the introduction, widespread changes in APA have been observed upon activation of T cells and other transitions between differentiation and

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

proliferation (Brumbaugh et al., 2018; Ji and Tian, 2009; Sandberg et al., 2008), although in most cases the mechanism(s) driving such signal-induced APA have not been fully determined. The fact that CELF2 expression increases upon activation of Jurkat T cells (Figure 1C), and CELF2 regulates activation-induced APA of its own transcript, suggests that increased expression of CELF2 may be one of the drivers of activation-induced APA transcriptome-wide. Consistently, when we overlap the location of CELF2 binding in 3′ UTRs with PAS sequences used in lymphoid tissues (spleen, lymph node, and thymus from APASdb; You et al., 2015), we find that more than 10% of these PAS sequences have a flanking CELF2 binding site within 100 nt (Figure 5A). By comparison, hnRNP L, which also binds extensively in 3′ UTRs in T cells (Shankarling et al., 2014), is found within 100 nt of fewer than 2% of PAS sequences (Figure 5A). Mapping of the CELF2 and hnRNP L CLIP peaks around the PAS sequences identified in various lymphoid tissues further highlights the tendency of CELF2 to bind around PAS sites, with particular enrichment immediately upstream of the hexamer (Figure 5B). To determine if this profile of CELF2 binding might overlap with CFI and CstF, we compared our CELF2 CLIP data with CLIP-seq data on the CFI and CstF complexes in the K562 lymphoblast cell line (www.encodeproject.org). We find that 6% and 10% of CELF2 binding sites directly overlap with CstF and CFI binding sites, respectively (Figure 5C). Again, in contrast, only 2.5% and 1% of hnRNP L peaks overlap with CstF and CFI sites, respectively (Figure 5D). The overlap of CELF2 particularly with CFI binding is also apparent in a density map of CLIP peaks around PAS sites (Figure 5E).

CELF2 Regulates a Broad Program of Activation-Induced APA in Jurkat Cells

To directly assess the functional role of CELF2 on APA, especially of that triggered in a signal-responsive manner, we analyzed our pre-existing RNA-seq data from Jurkat cells (Gazzara et al., 2017) with the DaPars algorithm (Xia et al., 2014). Comparison of the transcriptome of wild-type JSL1 cells preand post-stimulation identified 218 genes that undergo signal-responsive APA with a change of at least 20% between alternative sites (Table S2; Figure 6A). Remarkably, in almost half of these genes (96 of 218) we also observe APA changes upon depletion of CELF2 as detected by DaPars analysis and validated by 3′ RACE (Table S2; Figures 6A–6E; Figure S5). In many of these cases, depletion of CELF2 in stimulated cells clearly induces a polyadenylation pattern similar to unstimulated wild-type cells (Figures 6B–6E). Moreover, consistent with the meta-gene analysis in Figure 5, for many of the CELF2-dependent APA events we observe binding of CELF2 proximal to the PAS sites and overlapping binding of the CFIm and CstF complexes (Figures 6B–6E). These data confirm that CELF2 is a major driver of activation-induced APA, at least in the Jurkat T cells, and suggests that CELF2 generally regulates APA in a manner analogous its role as a repressor of splicing, namely, by competing with the binding of critical components of the processing machinery (Figure 6F; see Discussion).

The regulation of APA is predicted to affect protein expression through control of mRNA stability and/or translation. Indeed, we note that several of the genes for which we observe significant APA in Jurkat cells upon stimulation and depletion of CELF2 also exhibit large changes in mRNA abundance (Figure 7A; Table S3A), although we observe no specific correlation between 3′ UTR shortening or lengthening and transcript abundance (Table

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

S3A). We also analyzed recent mass spectrometry data from our Jurkat cells (L.M.A., B.A. Garcia, and K.W.L., unpublished data) to determine if protein expression is altered for any of the genes in which we observe CELF2 and/or stimulation-induced APA. Again, although we observe no general trend between APA and protein abundance, we do detect several instances of altered protein abundance (Table S3). For example, for SLC2A3 and LRCH4, we observe that the change in APA correlates with significant (~11- and 5-fold, respectively) changes in protein expression (Figure 7B; Table S3), while only SLC2A3 also has a significant change in mRNA abundance (Figure 7A). A more modest change in protein (~2-fold) is also observed for five additional genes for which we observe CELF2-dependent APA (Table S3B; Figure 7B, note log scale). The sensitivity of this protein abundance data was confirmed by the lack of change in the abundance of hnRNP L, and marked increase in CD69 upon stimulation of Jurkat cells, consistent with previous studies (Shankarling et al., 2014; Topp et al., 2008). Thus, these data suggest that at least a subset of CELF2-dependent APA events may regulate translation to shape the proteome in activated cells.

We were particularly intrigued by the identification of RBFOX2 mRNA as a target of CELF2-dependent APA, as we have previously reported that CELF2 represses the expression of RBFOX2 protein to control splicing of a broad program of RBFOX2-target genes; however, the mechanism by which CELF2 regulates RBFOX2 expression was unknown (Gazzara et al., 2017). RBFOX2 has two alternative sites of polyadenylation, the more distal of which is only used upon depletion of CELF2 in Jurkat cells (Figure 7C). To ask if CELF2-mediated repression of this distal site controls the expression of RBFOX2 protein, we transfected cells with an antisense morpholino oligo (AMO) that is complementary to this distal PAS site. As predicted, this AMO inhibits use of the distal PAS even in cells depleted of CELF2 (Figure 7C, RACE). Remarkably, inhibition of the distal PAS completely blocked the increase in RBFOX2 protein that is typically induced upon depletion of CELF2 (Figure 7C, western blot). Therefore, we conclude that CELF2-induced regulation of APA has a direct impact on RBFOX2 expression, and thus the subsequent gene expression profile that we have previously shown is a result of the functional antagonism between these two proteins (Gazzara et al., 2017).

DISCUSSION

APA is broadly observed in human transcripts and has been shown to be highly regulated in response to cell stimulation (Derti et al., 2012; Hoque et al., 2013; Sandberg et al., 2008; Tian et al., 2005; Tian and Manley, 2017). Here we identify a previously unrecognized role for the RBP CELF2 in regulating APA and provide evidence that CELF2 controls an extensive program of APA transcriptome-wide. Specifically, we show that CELF2 binding overlaps and competes with polyadenylation enhancer activities located upstream and downstream of the primary PAS site (Figure 6F). At a transcriptome-wide level, this competition may be most generally with CFI bound to UGUA sites upstream of the PAS sequence, given the marked peak of CELF2 binding ~100 nt before the PAS exactly overlapping the peak of CFI binding (Figures 5B and 5C). However, in the case of regulation of its own APA, CELF2 activity is also sensitive to the presence of the enhancer element downstream of PAS2 that binds CstF (Figure 4). Notably, at a transcriptome-wide level, CELF2 binding is enriched downstream of the PAS compared with the control protein

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

hnRNP L (Figure 5B). Moreover, some CstF binding is observed upstream of the PAS in addition to the dominant binding peak downstream (Figures 5C and 5E). Previously, Shi and colleagues showed that CFI promotes recruitment of CstF as well as CPSF, concluding that the CPAFs assemble cooperatively around PAS sites (Zhu et al., 2018). Taken together, these data suggest that CELF2 functions at multiple points to disrupt the efficient assembly of an active CPAF complex, including obscuring the binding site of CFI as well as CstF. Which interactions are most sensitive to CELF2 competition and most disruptive to PAS site use is likely highly dependent on the precise sequences and context of each gene.

The activity of CELF2 in regulating APA through competition with core enhancer activities is highly reminiscent of at least one mechanism by which splicing is regulated by CELF2, namely, competition with the splicing machinery at the 3′ splice site (Figure 6F). Dembowski and Grabowski (2009) showed that CELF2 causes exon skipping by binding across the branchpoint sequence of the 3′ splice site. Such interactions block association of U2AF65 and the U2 snRNP with the intron (Dembowski and Grabowski, 2009; Dujardin et al., 2010), thereby favoring use of the downstream competing 3′ splice site. Consistently, we have observed that ~25% of exons in Jurkat cells that are skipped in a CELF2-dependent manner exhibit binding of CELF2 within the 3′ splice site region (Ajith et al., 2016). Here we find that binding of CELF2 to the 3′ splice site of a terminal intron similarly blocks U2AF65 binding but in this case causes IR. IR is indeed the expected consequence of U2AF65 repression of a terminal 3′ splice site, as there is no alternative 3′ splice site downstream to allow assembly of the catalytic spliceosome complex. Importantly, we show here that although IR and APA can both occur in the same 3′ UTR, and both involve inhibition of core machinery, these regulatory processes are mechanistically uncoupled. Specifically, we find no difference in CELF2-induced APA in the presence or absence of the intron (Figure 3B), and our intron reporter lacks both PAS2 and PAS3 (Figure 2B). Therefore, we conclude that the CELF2 shapes 3′ UTR identity by at least two independent processes: APA and alternative splicing.

The identification of CELF2 as a regulator of polyadenylation provides an important new function to this protein. Interestingly, previous studies have determined that the close homolog CELF1 also exhibits enrichment of binding to 3′ UTRs (Wang et al., 2015). However, in these studies CELF1 was reported to function primary in cytoplasmic mRNA stability and decay (Masuda et al., 2012; Wang et al., 2015). Consistently, at least in Jurkat cells, CELF1 is dominantly expressed in the cytoplasm, while CELF2 is largely nuclear (Mallory et al., 2011). Moreover, in our knockdown and KO models of CELF2 we see no change in CELF1 expression (Martinez et al., 2015), yet we see a functional change in APA and splicing, indicating that CELF1 cannot functionally substitute for CELF2 in these processes. Therefore, we conclude that although CELF1 and CELF2 both bind extensively to 3′ UTRs, the functional consequence of these binding events is largely distinct.

Importantly, although we demonstrate here that CELF2 regulates many instances of activation-induced APA, we assume we have underestimated the scope of CELF2 mediated regulation of APA, as DaPars analysis has been shown to miss some APA events, especially among lowly expressed genes or those that contain more than two PAS sites (Xia et al., 2014). Our focus in this study has been on the identification and mechanism of CELF2 as an

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

APA regulatory factor; however, moving forward, it will be exciting to extend our studies here to identify the full program of CELF2 mediated APA in T cells as well as other tissues in which CELF2 is highly expressed such as muscle and neurons.

STAR★METHODS

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Kristen W. Lynch (klync@pennmedicine.upenn.edu). Reporter constructs generated in this study are all freely available upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

To generate CELF2 knockout clones, we used CRISPR editing that targeted CELF2 in Jurkat cells. Briefly, a custom designed crRNA (AUUUUCUGUCUUCCACAGCU) was annealed to tracrRNA (Dharmacon). This annealed gRNA complex (80 μM) was then mixed 1:1 by volume with 40 μM *S. pyogenes* Cas9-NLS (University of California Berkeley QB3 Macrolab) to a final concentration of 20 μM Cas9 ribonucleotide complex (RNP). This complexed gRNA:Cas9 RNP was nucleofected into Jurkat cells with the SE. Cell Line 96-well Nucleofector Kit using a 4-D Nucleofector following the manufacturer's recommendations (Lonza). Edited cells were cultured for 2–3 days before single-cell sorted (FACS Aria) into 96-well plates to generate single-cell clones. Clones were screened by sequencing of genomic DNA and by western blot. CELF2 depletion (knock-down) or overexpression was done using previously described Jurkat cells stably expressing doxycycline inducible FLAG-tagged CELF2 cDNA or shRNA targeting CELF2 (Mallory et al., 2015). All Jurkat cells were cultured in in RPMI (Corning) supplemented with 5% heat-inactivated fetal bovine serum (FBS) (GIBCO) as described previously (Lynch and Weiss, 2000). HeLa cells were cultured in DMEM (Corning) with 10% heat-inactivated FBS. For stimulations, Jurkat T cells were cultured with 20 ng/mL PMA (Sigma-Aldrich) for 48 h (RT–PCR and 30RACE), or the indicated time points for western blot analyses. Jurkat cells are male, HeLa cells are female.

METHOD DETAILS

Transfections—HeLa cells were transfected with indicated plasmids containing CELF2 3′UTR constructs (4 μg total DNA) by Lipofectamine 2000 (Invitrogen) transfection reagent according to the manufacturer's instructions. Cells were harvested 24 hours post-transfection for RT-PCR, 3′RACE and western blot analyses. To alter RBFOX2 APA Jurkat cells were transfected with 10 nmol of an AMO blocking distal polyA site of RBFOX2 3′UTR (5′ GAAGCACTGTTTTTAAATAAAAGAGAGAAACACCA 3′; GeneTools) by electroporation as previously described (Lynch and Weiss, 2000; Rothrock et al., 2003). Cells were incubated with AMO for 16 h after transfection and then treated with doxycycline to induce depletion of CELF2 for 48 hours.

RT-PCR—RNA from Jurkat and HeLa cells were isolated using RNABee (Tel-Test, Inc.) according to the manufacturer's instructions. Low-cycle RT-PCR was carried out and analyzed as previously described (Lynch and Weiss, 2000; Rothrock et al., 2003) using

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

sequence-specific primers for individual genes. Primer sequences used for RT-PCR experiments are provided in Table S4.

3′RACE—RACE-Ready cDNA were produced and transcript 3′ ends were identified by SMARTer RACE cDNA Amplification Kit (Clontech) according to the manufacturer's instructions. RACE-Ready cDNA were analyzed on agarose gels. Gene specific forward primers used for amplification of RACE-Ready cDNA are listed in Table S4.

Western blots—Western blots were carried out as previously described (Melton et al., 2007). Briefly, 10 μg of total protein lysates were loaded into 10% 37.5:1 bis-acrylamide SDS-PAGE gels. Antibodies used for western blot analyses are listed in the Key Resources Table.

UV crosslinking and immunoprecipitation—*In vitro* UV crosslinking assays were performed as previously described (Ajith et al., 2016). Briefly, CELF2 3′UTR constructs were *in vitro* transcribed from a T7 promoter and body-labeled with ^{32}P α-UTP. Nuclear extracts from unstimulated or stimulated Jurkat cells or purified recombinant proteins were incubated with 3′UTR RNAs, UV cross-linked, RNase-digested, and resolved on an SDSPAGE. For UV crosslinking followed by immunoprecipitation, UV cross-linking reactions after RNase digestion were incubated with corresponding primary antibodies overnight at 4°C in RIPA buffer followed by incubation with protein-G Sepharose beads (GE) and elution. 200 ng of purified recombinant CELF2 protein (Ajith et al., 2016) or hnRNPL protein (Chiou et al., 2013) was mixed with nuclear extract from unstimulated Jurkat cells for UV crosslinking followed by immunoprecipitation in Figure 2E. Amounts of purified recombinant CELF2, CFIm25, CstF64, hnRNPL and hnRNPK proteins used for *in vitro* UV crosslinking experiments are as indicated in the Figures 4 and S4. Sources of the recombinant proteins are listed in Key Reources Table. Antibodies used for immunoprecipitation are listed in Key Resources Table.

Mutagenesis—Specific protein binding sequences in CELF2 3′UTR constructs (Figures 3D and 3E) were mutated using QuickChange Lightening Site-Directed Mutagenesis kit (Agilent Technologies) according to manufacturer's instructions. Sequence specific primers used for mutagenesis are listed in Table S4.

LC-MS/MS analysis—Protein sample processing for mass spectrometry (MS) was carried out as described previously (L.M.A., B.A. Garcia, and K.W.L., unpublished data). In brief, unstimulated and PMA stimulated JSL1 cells were harvested after 48 hours of treatment and lysed in a sequential, two-step manner in order to fractionate cells into cytoplasmic versus nuclear fractions. Proteins were digested with trypsin and analyzed using a Dionex UltiMate 3000 (Thermo Scientific, San Jose, CA, USA) liquid chromatography (LC) system coupled online with a Q Exactive HF-X instrument (Thermo Scientific, San Jose, CA, USA). LC was configured with a 75 μm ID × 20 cm Reprosil-Pur C_{18}-AQ (3 μm; Dr. Maisch GmbH, Germany) reverse phase capillary column packed in-house. Data was analyzed by the Sequest HT search engine in Proteome Discoverer v2.3 (Thermo Scientific), using a full human proteome database with canonical protein sequences (SwissProt, release 2018_05)

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

and a 1%false discovery rate (FDR). Cytoplasmic and nuclear samples were considered as fractions of the same cellular condition.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analysis for the quantification of differential alternative polyadenylation in Table S2 was done according to the DaPars pipeline using the standard default settings (Xia et al., 2014). For the proteomic data (Table S3), significant changes were determined by performing a two-tailed homoscedastic t test between unstimulated and stimulated conditions; our significance cutoff was p value < 0.05 (or 4.32 when–log2 transformed). For the RT-PCR analysis of intron retention in Figures 2 and S2, we provide standard deviation as calculated from three independent replicate experiments.

DATA AND CODE AVAILABILITY

The unique proteomic data used for this study is available in the PRIDE database (PDX012556).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

This work was funded by an NIH grant R35 GM118048 to K.W.L. L.M.A. was supported by an NSF pre-doctoral fellowship. We would like to thank Yong sheng Shi and members of the Lynch lab for helpful discussion and comments on the manuscript and Yoseph Barash and Ben Garcia for guidance with data analysis.

REFERENCES

Agrawal AA, Salsi E, Chatrikhi R, Henderson S, Jenkins JL, Green MR, Ermolenko DN, and Kielkopf CL (2016). An extended U2AF(65)-RNA-binding domain recognizes the 3′ splice site signal. Nat. Commun 7, 10950. [PubMed: 26952537]

Ajith S, Gazzara MR, Cole BS, Shankarling G, Martinez NM, Mallory MJ, and Lynch KW (2016). Position-dependent activity of CELF2 in the regulation of splicing and implications for signal-responsive regulation in T cells. RNA Biol 13, 569–581. [PubMed: 27096301]

Bartel DP (2009). MicroRNAs: target recognition and regulatory functions. Cell 136, 215–233. [PubMed: 19167326]

Basu SK, Malik R, Huggins CJ, Lee S, Sebastian T, Sakchaisri K, Quiñones OA, Alvord WG, and Johnson PF (2011). 3′UTR elements inhibit Ras-induced C/EBP0β post-translational activation and senescence in tumour cells. EMBO J 30, 3714–3728. [PubMed: 21804532]

Berkovits BD, and Mayr C (2015). Alternative 3′ UTRs act as scaffolds to regulate membrane protein localization. Nature 522, 363–367. [PubMed: 25896326]

Bicknell AA, Cenik C, Chua HN, Roth FP, and Moore MJ (2012). Introns in UTRs: why we should stop ignoring them. BioEssays 34, 1025–1034. [PubMed: 23108796]

Brumbaugh J, Di Stefano B, Wang X, Borkent M, Forouzmand E, Clowers KJ, Ji F, Schwarz BA, Kalocsay M, Elledge SJ, et al. (2018). Nudt21 controls cell fate by connecting alternative polyadenylation to chromatin signaling. Cell 172, 106–120.e21. [PubMed: 29249356]

Chiou NT, Shankarling G, and Lynch KW (2013). hnRNP L and hnRNP A1 induce extended U1 snRNA interactions with an exon to repress spliceosome assembly. Mol. Cell 49, 972–982. [PubMed: 23394998]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Dasgupta T, and Ladd AN (2012). The importance of CELF control: molecular and biological roles of the CUG-BP, Elav-like family of RNA-binding proteins. Wiley Interdiscip. Rev. RNA 3, 104–121. [PubMed: 22180311]

Dembowski JA, and Grabowski PJ (2009). The CUGBP2 splicing factor regulates an ensemble of branchpoints from perimeter binding sites with implications for autoregulation. PLoS Genet 5, e1000595. [PubMed: 19680430]

Derti A, Garrett-Engele P, Macisaac KD, Stevens RC, Sriram S, Chen R, Rohl CA, Johnson JM, and Babak T (2012). A quantitative atlas of polyadenylation in five mammals. Genome Res 22, 1173–1183. [PubMed: 22454233]

Dujardin G, Buratti E, Charlet-Berguerand N, Martins de Araujo M, Mbopda A, Le Jossic-Corcos C, Pagani F, Ferec C, and Corcos L (2010). CELF proteins regulate CFTR pre-mRNA splicing: essential role of the divergent domain of ETR-3. Nucleic Acids Res 38, 7273–7285. [PubMed: 20631008]

Gazzara MR, Mallory MJ, Roytenberg R, Lindberg JP, Jha A, Lynch KW, and Barash Y (2017). Ancient antagonism between CELF and RBFOX families tunes mRNA splicing outcomes. Genome Res 27, 1360–1370. [PubMed: 28512194]

Hoque M, Ji Z, Zheng D, Luo W, Li W, You B, Park JY, Yehia G, and Tian B (2013). Analysis of alternative cleavage and polyadenylation by 3′ region extraction and deep sequencing. Nat. Methods 10, 133–139. [PubMed: 23241633]

Ji Z, and Tian B (2009). Reprogramming of 3′ untranslated regions of mRNAs by alternative polyadenylation in generation of pluripotent stem cells from different cell types. PLoS ONE 4, e8419. [PubMed: 20037631]

Ladd AN, Charlet N, and Cooper TA (2001). The CELF family of RNA binding proteins is implicated in cell-specific and developmentally regulated alternative splicing. Mol. Cell. Biol 21, 1285–1296. [PubMed: 11158314]

Lejeune F, and Maquat LE (2005). Mechanistic links between nonsense-mediated mRNA decay and pre-mRNA splicing in mammalian cells. Curr. Opin. Cell Biol 17, 309–315. [PubMed: 15901502]

Lynch KW, and Weiss A (2000). A model system for activation-induced alternative splicing of CD45 pre-mRNA in T cells implicates protein kinase C and Ras. Mol. Cell. Biol 20, 70–80. [PubMed: 10594010]

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.e19. [PubMed: 30449617]

Mallory MJ, Jackson J, Weber B, Chi A, Heyd F, and Lynch KW (2011). Signal- and development-dependent alternative splicing of LEF1 in T cells is controlled by CELF2. Mol. Cell. Biol 31, 2184–2195. [PubMed: 21444716]

Mallory MJ, Allon SJ, Qiu J, Gazzara MR, Tapescu I, Martinez NM, Fu XD, and Lynch KW (2015). Induced transcription and stability of CELF2 mRNA drives widespread alternative splicing during T-cell signaling. Proc. Natl. Acad. Sci. U S A 112, E2139–E2148. [PubMed: 25870297]

Martin G, Gruber AR, Keller W, and Zavolan M (2012). Genome-wide analysis of pre-mRNA 3′ end processing reveals a decisive role of human cleavage factor I in the regulation of 3′ UTR length. Cell Rep 1, 753–763. [PubMed: 22813749]

Martinez NM, Agosto L, Qiu J, Mallory MJ, Gazzara MR, Barash Y, Fu XD, and Lynch KW (2015). Widespread JNK-dependent alternative splicing induces a positive feedback loop through CELF2-mediated regulation of MKK7 during T-cell activation. Genes Dev 29, 2054–2066. [PubMed: 26443849]

Masuda A, Andersen HS, Doktor TK, Okamoto T, Ito M, Andresen BS, and Ohno K (2012). CUGBP1 and MBNL1 preferentially bind to 3′ UTRs and facilitate mRNA decay. Sci. Rep 2, 209. [PubMed: 22355723]

Mayr C (2017). Regulation by 3′-untranslated regions. Annu. Rev. Genet 51, 171–194. [PubMed: 28853924]

Melton AA, Jackson J, Wang J, and Lynch KW (2007). Combinatorial control of signal-induced exon repression by hnRNP L and PSF. Mol. Cell. Biol 27, 6972–6984. [PubMed: 17664280]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Rothrock C, Cannon B, Hahm B, and Lynch KW (2003). A conserved signal-responsive sequence mediates activation-induced alternative splicing of CD45. Mol. Cell 12, 1317–1324. [PubMed: 14636588]

Sandberg R, Neilson JR, Sarma A, Sharp PA, and Burge CB (2008). Proliferating cells express mRNAs with shortened 3′ untranslated regions and fewer microRNA target sites. Science 320, 1643–1647. [PubMed: 18566288]

Shankarling G, Cole BS, Mallory MJ, and Lynch KW (2014). Transcriptome-wide RNA interaction profiling reveals physical and functional targets of hnRNP L in human T cells. Mol. Cell. Biol 34, 71–83. [PubMed: 24164894]

Sheets MD, Ogg SC, and Wickens MP (1990). Point mutations in AAUAAA and the poly (A) addition site: effects on the accuracy and efficiency of cleavage and polyadenylation in vitro. Nucleic Acids Res 18, 5799–5805. [PubMed: 2170946]

Shi Y, and Manley JL (2015). The end of the message: multiple protein-RNA interactions define the mRNA polyadenylation site. Genes Dev 29, 889–897. [PubMed: 25934501]

Takagaki Y, Seipelt RL, Peterson ML, and Manley JL (1996). The polyadenylation factor CstF-64 regulates alternative processing of IgM heavy chain pre-mRNA during B cell differentiation. Cell 87, 941–952. [PubMed: 8945520]

Thompson MG, Muñoz-Moreno R, Bhat P, Roytenberg R, Lindberg J, Gazzara MR, Mallory MJ, Zhang K, García-Sastre A, Fontoura BMA, and Lynch KW (2018). Co-regulatory activity of hnRNP K and NS1-BP in influenza and human mRNA splicing. Nat. Commun 9, 2407. [PubMed: 29921878]

Tian B, and Manley JL (2017). Alternative polyadenylation of mRNA precursors. Nat. Rev 18, 18–30.

Tian B, Hu J, Zhang H, and Lutz CS (2005). A large-scale analysis of mRNA polyadenylation of human and mouse genes. Nucleic Acids Res 33, 201–212. [PubMed: 15647503]

Topp JD, Jackson J, Melton AA, and Lynch KW (2008). A cell-based screen for splicing regulators identifies hnRNP LL as a distinct signal-induced repressor of CD45 variable exon 4. RNA 14, 2038–2049. [PubMed: 18719244]

Vaquero-Garcia J, Barrera A, Gazzara MR, González-Vallinas J, Lahens NF, Hogenesch JB, Lynch KW, and Barash Y (2016). A new view of transcriptome complexity and regulation through the lens of local splicing variations. Elife 5, e11752. [PubMed: 26829591]

Wang ET, Ward AJ, Cherone JM, Giudice J, Wang TT, Treacy DJ, Lambert NJ, Freese P, Saxena T, Cooper TA, and Burge CB (2015). Antagonistic regulation of mRNA expression and splicing by CELF and MBNL proteins. Genome Res 25, 858–871. [PubMed: 25883322]

Xia Z, Donehower LA, Cooper TA, Neilson JR, Wheeler DA, Wagner EJ, and Li W (2014). Dynamic analyses of alternative polyadenylation from RNA-seq reveal a 3′-UTR landscape across seven tumour types. Nat. Commun 5, 5274. [PubMed: 25409906]

Yang Q, Gilmartin GM, and Doublié S (2010). Structural basis of UGUA recognition by the Nudix protein CFI(m)25 and implications for a regulatory role in mRNA 3′ processing. Proc. Natl. Acad. Sci. USA 107, 10062–10067. [PubMed: 20479262]

You L, Wu J, Feng Y, Fu Y, Guo Y, Long L, Zhang H, Luan Y, Tian P, Chen L, et al. (2015). APASdb: a database describing alternative poly(A) sites and selection of heterogeneous cleavage sites downstream of poly(A) signals. Nucleic Acids Res 43, D59–D67. [PubMed: 25378337]

Zhu Y, Wang X, Forouzmand E, Jeong J, Qiao F, Sowd GA, Engelman AN, Xie X, Hertel KJ, and Shi Y (2018). Molecular mechanisms for CFIm-mediated regulation of mRNA alternative polyadenylation. Mol. Cell 69, 62–74.e4. [PubMed: 29276085]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Highlights

- The RNA-binding protein CELF2 competes for binding with 3′ end processing factors
- CELF2 regulates a broad program of cellular alternative polyadenylation (APA)
- Increased CELF2 protein upon cell stimulation drives stimulation-induced APA
- CELF2 regulates RBFOX2 expression by limiting use of a distal polyadenylation site

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

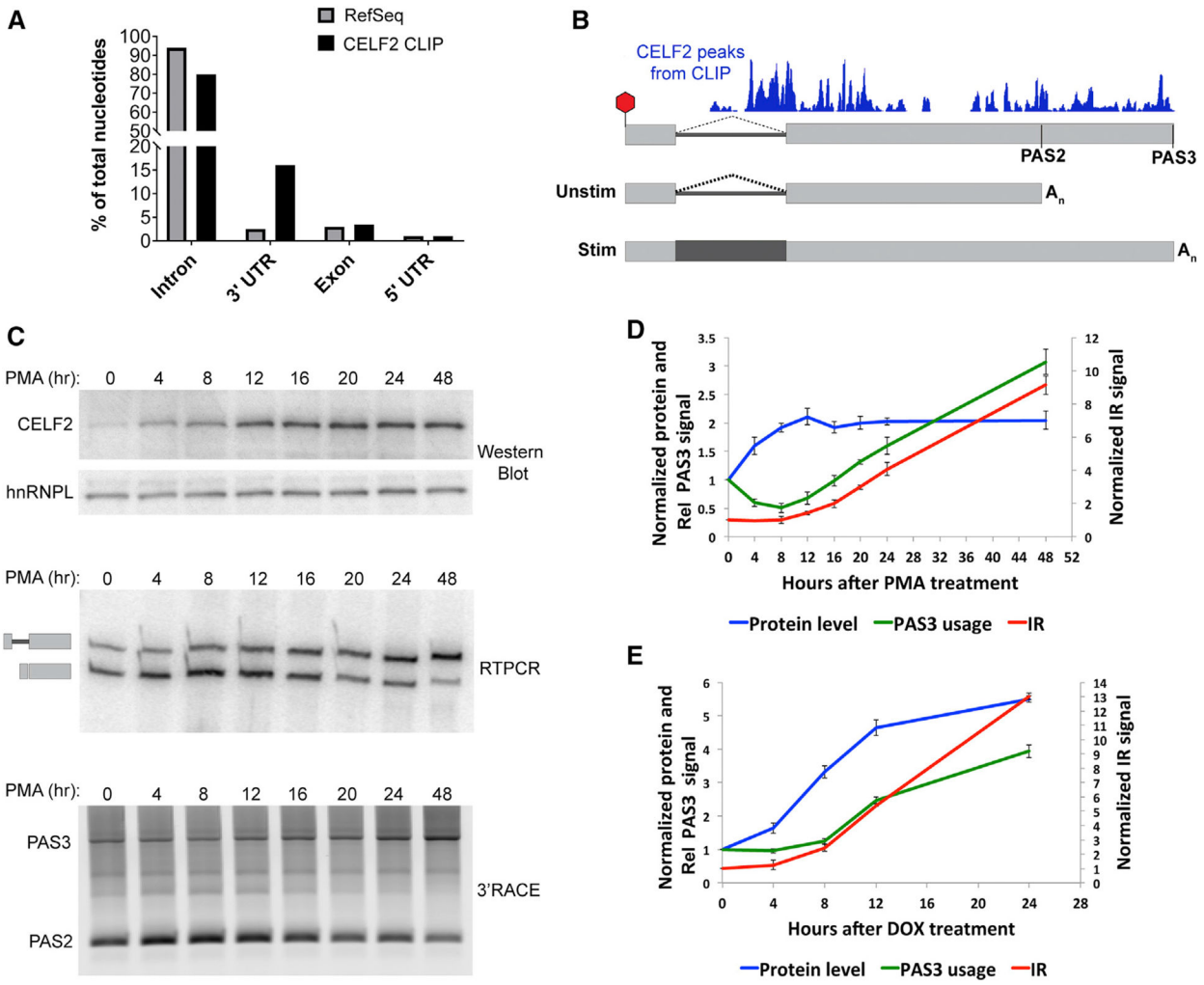


Figure 1. CELF2 Binds Its Own 3′ UTR and Induces Intron Retention and Alternative Polyadenylation

(A) Transcriptome distribution of CELF2 CLIP-seq peaks (black; from Ajith et al., 2016) compared with total Ref-Seq transcriptome (gray).

(B) Schematic of CELF2 CLIP peaks mapped along the 3′ UTR of CELF2 mRNA. Light gray boxes are exons, thin and thick dark gray line and/or box are the spliced and retained intron. The relative location of two alternative PAS sites (PAS2 and PAS3) is indicated. No CLIP peaks for CELF2 are detected downstream of PAS3.

(C) Representative detection of CELF2 protein (by western blot, top), CELF2 mRNA 3′ UTR intron retention (by RT-PCR, middle), and switch from PAS2 to PAS3 (by 3′ RACE, bottom) following stimulation of Jurkat cells with PMA for the indicated time points. HnRNPL is used as a loading control for the western blot, as this has previously been shown not to change in response to PMA (Shankarling et al., 2014).

(D) Quantification of the data in (C) along with replicates (n = 3). PAS3 quantification is relative to PAS2 (i.e., PAS3/[PAS3 + PAS2]).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(E) Quantification of assays such as in (C) but following induction of CELF2 protein with a doxycycline-driven CELF2 cDNA construct (n = 3). Error bars in both (D) and (E) represent SD. Representative gels for (E) are shown in Figure S1.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

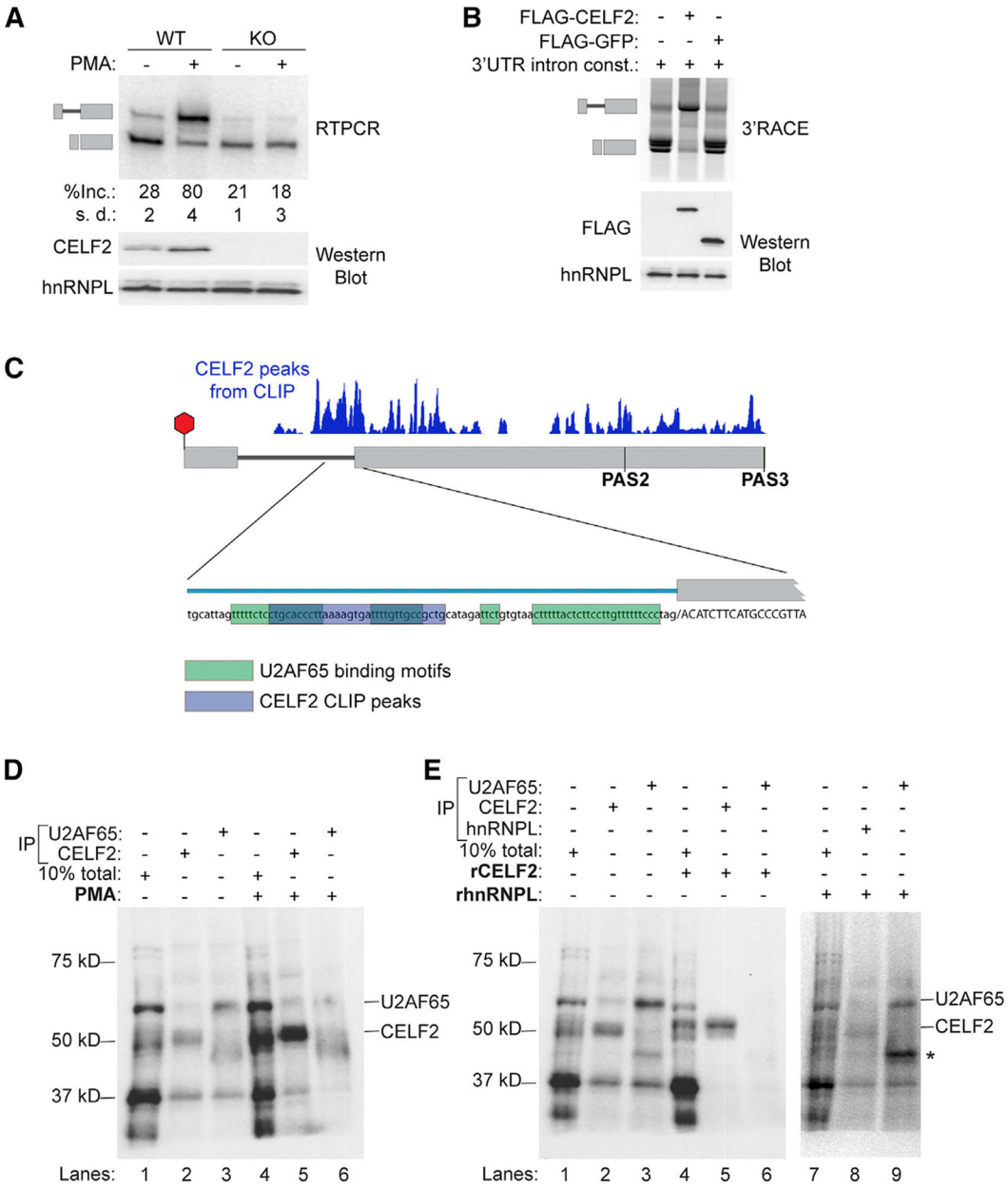


Figure 2. CELF2 Is Necessary and Sufficient for Intron Retention by Inhibiting U2AF65 Binding

(A) Analysis of CELF2 3′ UTR intron retention by RTPCR in parental wild-type Jurkat cells versus cells that express no detectable CELF2 protein because of CRISPR-mediated frameshift (KO). Western blot confirms lack of CELF2 protein with hnRNP L as loading control. See also Figure S2.

(B) Analysis of intron retention in CELF2 3 ′UTR intron reporter construct that contains the first 2.5 kb of the CELF2 3′ UTR (inclusive of the whole intron and flanking sequences), upon transfection of HeLa cells with the reporter alone or together with CELF2 or GFP cDNA as indicated.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(C) Schematic of the CELF2 3′ UTR highlighting the sequences in the 3′ splice site region of the intron and location of U2AF65 and CELF2 binding sites, as well as the results of CLIP-seq.

(D) *In vitro* UV crosslinking of nuclear extract from unstimulated (−PMA) or stimulated (+PMA) Jurkat cells with *in vitro* transcribed, radiolabeled RNA corresponding to 120 nt encompassing the intron sequences highlighted in (C). "+IP" lanes are those in which immunoprecipitation with the indicated antibody was done after crosslinking and RNase treatment. Precipitated product is then run on the SDS-PAGE gel.

(E) Same as (D), but all reactions are done with nuclear extract from unstimulated Jurkat cells, in the absence (−) or presence (+) of 200 ng purified recombinant CELF2 or recombinant hnRNPL. Asterisk indicates a protein that binds non-specifically to Protein-G Sepharose beads.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

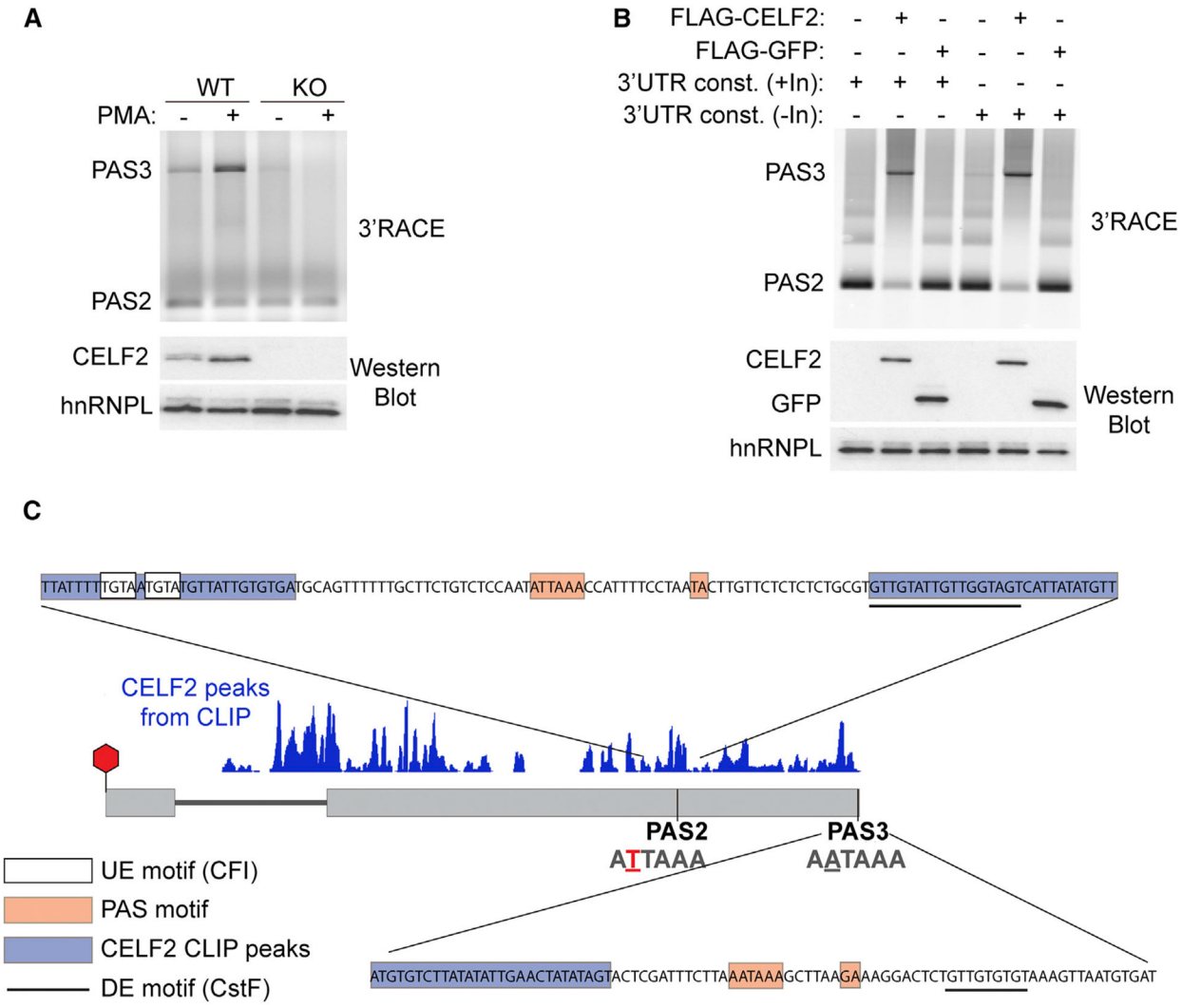


Figure 3. CELF2 Is Necessary and Sufficient for the Switch to PAS3 and Binds to Sequences Overlapping Enhancer Elements Required for PAS2

(A) Analysis of CELF2 polyadenylation site use by 3′ RACE in parental wild-type Jurkat cells versus cells that express no detectable CELF2 protein because of CRISPR-mediated frameshift (KO). Western blot confirms lack of CELF2 protein with hnRNP L as loading control, as also shown in Figure 2A. See also Figure S3.

(B) Analysis of polyadenylation in CELF2 3′ UTR PAS reporter construct that contains either the entire CELF2 3′ UTR (+Intron), or just sequences after the intron (−Intron), when transfected into HeLa cells either alone or together with CELF2 or GFP cDNA as indicated.

(C) Schematic of the CELF2 3′ UTR highlighting the sequences surrounding the PAS2 and PAS3 sites, highlighting the PAS hexamer sequence (orange), location of the upstream enhancer (UE; box) and downstream enhancer (DE; underline) sequences, and location of CELF2 binding sites (blue).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

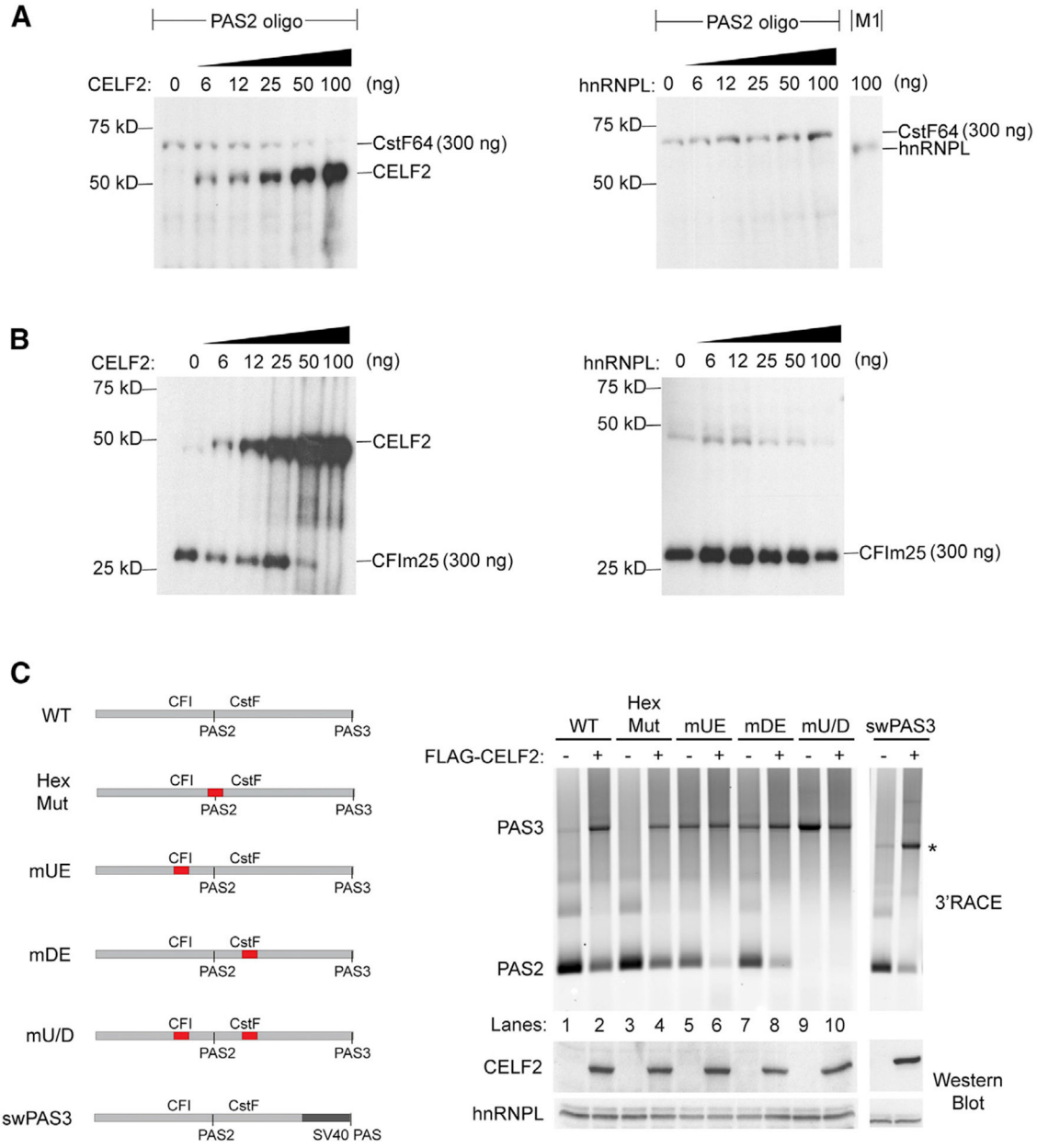


Figure 4. CELF2 Competes with CFIm25 and CstF64 to Regulate APA of Its Own 3′ UTR

(A and B) *In vitro* UV crosslinking analysis of indicated amount of recombinant CstF64 (A) or CFIm25 (B) protein bound to radiolabeled PAS2 oligonucleotide in the absence or presence of increasing amounts of recombinant CELF2 (left) or hnRNPL (right). hnRNPL protein crosslinked to M1 oligonucleotide (Thompson et al., 2018) is shown as a positive control for hnRNPL binding to RNA. The lane with hnRNP L alone was run on a different gel. (C) Left, schematic of reporter constructs and right, 3′ RACE analysis of mutant constructs upon transfection in HeLa cells with (+) or without (−) cotransfected CELF2 cDNA. Asterisk marks the product resulting from use of SV40 PAS. Red boxes indicate location of mutations that disrupt binding of CstF64, CFIm25, and CELF2, as indicated in the text. swPAS3 lacks all sequences around the native PAS3 (shown as a dark gray box),

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

such that the vector-encoded non-PMA responsive SV40 polyA is the only other PAS present.

See also Figure S4.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

A

Total lymphoid PAS	15064	
CELF2 flanking	1791	11.9%
hnRNPL flanking	236	1.6%

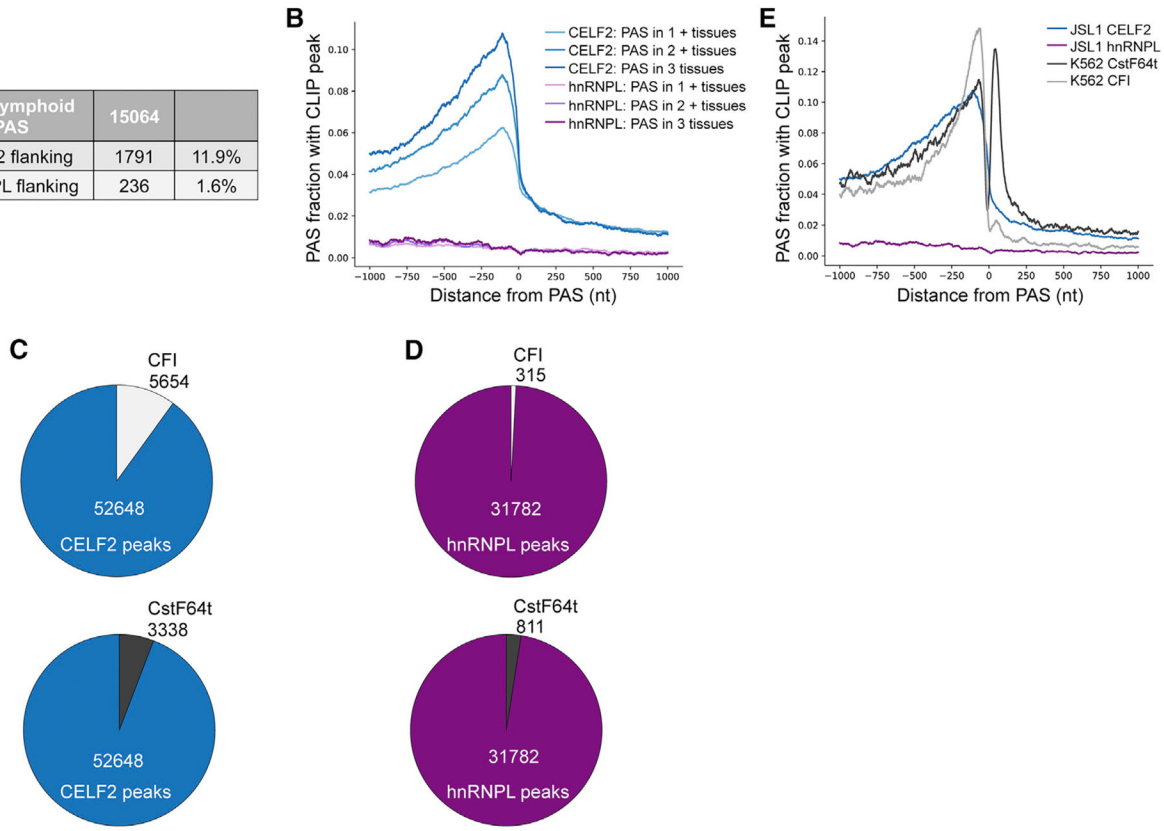


Figure 5. CELF2 Binding Overlaps Extensively with APA Enhancer Activities Transcriptome-wide

(A) Number of PAS sites identified in at least two lymphoid tissues (total), from APASdb (You et al., 2015), that overlap with CELF2 or hnRNP L CLIP-peaks from Jurkat cells (Ajith et al., 2016; Shankarling et al., 2014).

(B) Mapping of CLIP-peak density across a 1 kb interval before and after PAS sites identified in lymphoid tissues. Color intensity indicates confidence of PAS site, on the basis of identification in one (light), two (medium), or three (dark) of the lymphoid tissues.

(C) Number of CFI (top) or CstF (bottom) CLIP-peaks from K562 cells (www.encodeproject.org) that overlap with CELF2 CLIP-peaks from Jurkat cells (Ajith et al., 2016; Shankarling et al., 2014). Total number of CELF2 peaks and total number of overlapped peaks are given.

(D) Same as (C) but with hnRNP L CLIP-peaks from Jurkat cells (Ajith et al., 2016; Shankarling et al., 2014).

(E) Same as in (C) but mapping of the CFI and CstF Encode data as well as the CELF2 and hnRNP L data around the highest confidence PAS sites.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

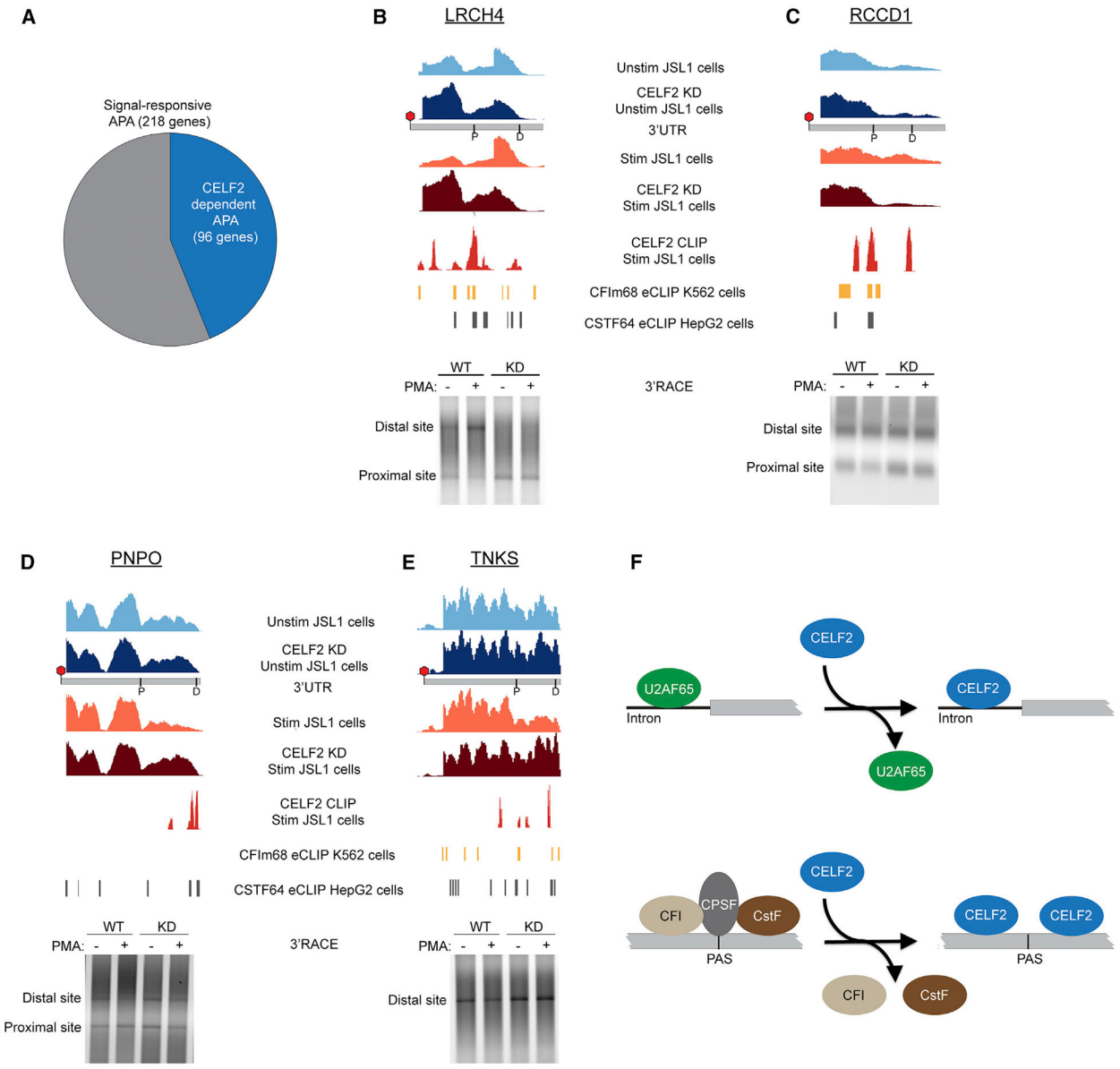


Figure 6. CELF2 Regulates a Broad Program of Activation-Induced APA in Jurkat Cells

(A) Number of CELF2-regulated APA events that overlap with total activation-induced APA events, as detected by DaPars analysis (Xia et al., 2014) of RNA-seq data from Jurkat cells. (B–E) 3′ RACE analysis of APA of the genes LRCH4 (B), RCCD1 (C), PNPO (D), and TNKS (E) in wild-type (WT) or CELF2-deficient (KD) cells before (−) or after (+) stimulation with PMA. For each panel, the gene name, 3′ UTR schematic, and RNA-seq tracks are shown from the indicated cell conditions. Proximal (P) and distal (D) PAS sites are indicated. CELF2 CLIP peaks from stimulated Jurkat cells (bottom bright red tracks), CFIm68 CLIP peaks from K562 cells (bottom yellow tracks), and CstF64 CLIP peaks from HepG2 cells (bottom gray tracks) are also shown. See Figure S5 for more examples.

(F) Model showing CELF2 regulation of 3′ UTR IR and APA by competition with core processing machinery. Top: CELF2 promotes 3′ UTR IR by inhibiting U2AF65 binding to 3′ splice site. Bottom: CELF2 regulates APA by competing with core enhancer factors CFI and CstF upstream and downstream of PAS, respectively.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

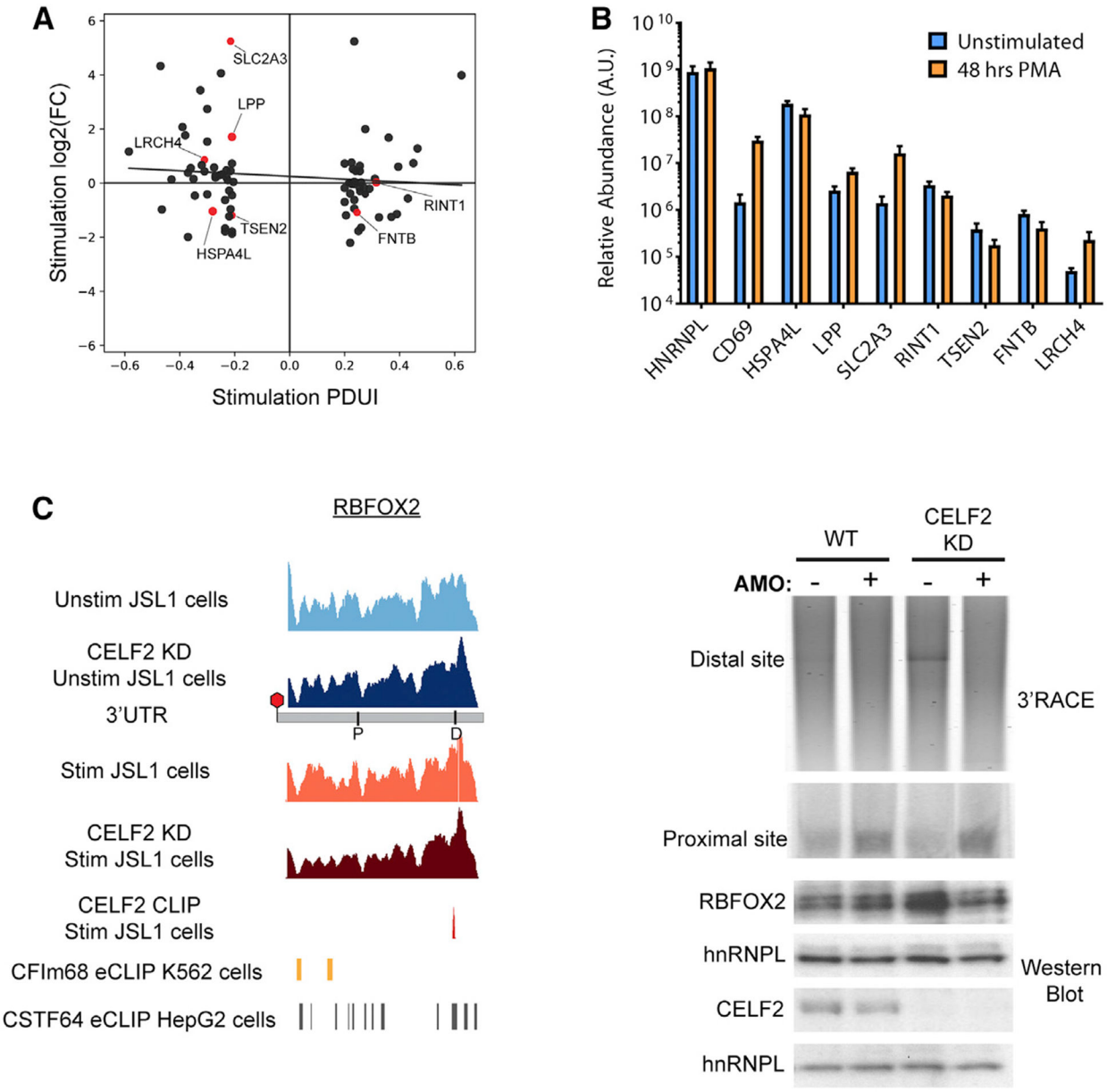


Figure 7. Functional Impacts of CELF2-Mediated APA

(A) A scatterplot of stimulation-induced change in steady-state transcript level (log_2[FC]) versus APA shifts (percentage distal polyA site usage index [PDUI]) for CELF2 target genes. Genes that also exhibit changes in their encoded protein (see B) are highlighted in red and labeled.

(B) Abundance of proteins encoded by genes with CELF2 and/or activation-dependent APA (plus proteins known to change [CD69] or not [hnRNP L]) upon stimulation. Relative abundance is calculated from shotgun proteomics of Jurkat cells before and after stimulation, as described in the STAR Methods. Error bars represent SE from at least two independent experiments.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(C) 3′ RACE analysis of APA of RBFOX2 in wild-type (WT) or CELF2-deficient (KD) cells without (−) or with (+) transfection of AMO targeting RBFOX2 distal site. 3′ UTR schematic and RNA-seq tracks are shown from the indicated cell conditions. Proximal(P) and distal (D) PAS sites are indicated. CELF2 CLIP peaks from stimulated Jurkat cells (bottom bright red tracks), CFIm68 CLIP peaks from K562 cells (bottom yellow tracks), and CstF64 CLIP peaks from HepG2 cells (bottom gray tracks) are also shown. Western blots show that inhibition of the distal PAS completely blocked the increase in RBFOX2 protein that is typically induced upon depletion of CELF2. hnRNPL is used as a loading control. AMO, antisense morpholino oligo.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
CELF2	University of Florida ICBR	HL1889; RRID:AB_1078584
U2AF65	Sigma	U4758; RRID:AB_262122
hnRNPL	Abcam	Ab6106; RRID:AB_305294
FLAG-tag	Cell Signaling Technology	2368S; RRID:AB_10707327
RBFOX2	Bethyl	(A300–864A); RRID: AB_609476
Chemicals, Peptides, and Recombinant Proteins		
RPMI 1640	Corning	10-040-CV
DMEM	Corning	10-013-CV
FBS	GIBCO	16000-044
Lipofectamine 2000 reagent	Invitrogen	52887
Protein G Sepharose beads	GE Healthcare	17-0618-01
Recombinant CFIm25	Abcam	ab104669
Recombinant CstF64	Origene	TP304450
Recombinant CELF2	Dr. Kristen W. Lynch	Ajith et. al., 2016
Recombinant hnRNPL	Dr. Kristen W. Lynch	Chiou et. al., 2013
Recombinant hnRNPK	Dr. Kristen W. Lynch	Thompson et. al., 2018
S. pyogenes Cas9-NLS	UC Berkeley QB3 Macrolab	N/A
Critical Commercial Assays		
SMARTer RACE cDNA Amplification Kit	Clontech	634858
QuickChange Lightening Site-Directed Mutagenesis kit	Agilent Technologies	210518–5
Nucleofector Kit using a 4-D Nucleofector	Lonza	V4XC-2012
Deposited Data		
Proteomic data	Dr. Benjamin A. Garcia and Dr. Kristen W. Lynch	PRIDE database: (PDX012556)
Experimental Models: Cell Lines		
Human: JSL1 Jurkat T cells	Dr. Kristen W. Lynch	Lynch and Weiss, 2000
Human: HeLa cells	ATCC	N/A
Human: CELF2 KO Jurkat T cells	This study	N/A
Oligonucleotides		
Primers for RTPCR and 3'RACE	This study	See Table S4
Primers for 3′UTR mutagenesis	This study	See Table S4
Primers for UV crosslinking	This study	See Table S4
Recombinant DNA		
Plasmid: CELF2 3′UTR intron reporter	This study	N/A
Plasmid: CELF2 3′UTR PAS (+Intron) reporter	This study	N/A
Plasmid: CELF2 3′UTR PAS (−Intron) reporter	This study	N/A
Plasmid: mUE-C2	This study	N/A
Plasmid: mDE-C2	This study	N/A

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Plasmid: mU/D-C2	This study	N/A
Plasmid: mUE	This study	N/A
Plasmid: mDE	This study	N/A
Plasmid: mU/D	This study	N/A
Plasmid: sw-PAS3	This study	N/A
Software and Algorithms		
DaPARs analysis	Xia et. al., 2014	https://code.google.com/p/dapars/
MAJIQ algorithm	Vaquero-Garcia et al., 2016	https://majiq.biociphers.org