Supporting Information

Mallory et al. 10.1073/pnas.1423695112



Fig. S1. RT-PCR analysis of the expression of endogenous and cDNA-derived CELF2 mRNA. Low-cycle RT-PCR of CELF2 mRNA derived from the endogenous gene (*Top*, eCELF2) or stably integrated cDNA (*Bottom*, rCELF2).



Fig. S2. CELF2 protein expression is unaffected by depletion of Dicer in JSL1 T cells. (A) Western blot for indicated proteins using lysates from JSL1 cells depleted of Dicer protein expression by treatment with a morpholino oligonucleotides. FADD serves as a positive control for a protein known to be under control of miRNAs/Dicer in T cells. hnRNP L serves as a loading control. (B) Quantification of two to three replicates of blots as shown in A.

Modeling CELF2 mRNA level

A simple model for the increase and eventual stabilization of CELF2 mRNA level can be derived as follows:

r(t) is the relative level of CELF2 mRNA level at time t after PMA stimulation.

p is the rate of transcription of new CELF2 mRNAs.

g is the rate of degradation of CELF2 mRNAs.

r(t) is related to p and g like:

S N N C

 $dr(t) / dt = p - g^*r(t)$

The solution of this model is $r(t) = p/g + C^*e^{-gt}$. In order for r(0) = 1, set C = (1 - p/g) such that:

$$r(t) = p/g + (1 - p/g) * e^{-gt}$$

We observe a 4.5-fold increase in p and a 2.5-fold decrease in g from unstimulated to fully stimulated JSL1. This suggests a 4.5*2.5 = 11.25-fold increase in the steady state level of CELF2 mRNA.



Fig. S3. Modeling of CELF2 mRNA expression based on observed regulation of transcription and stability demonstrates that the observed increases in transcription and stability are sufficient to account for the ~12-fold increase in mRNA expression that has been experimentally measured.

Key

PNAS PNAS

- --- = reported polyA cleavage site (Ensembl 8/8/14)
- ---_ = proposed polyA recognition hexamer site
- **nn** = experimentally validated polyA cleavage site (this study)

aaa = intron

AAA = exon

/ = exon-intron boundary

ISTOP 80
<u>TGA</u> TCCTAACCCCAGAGGCTCCCTGCTCTCATTTTAGCTTTCTTAGG/gtaagtcccacgagccagcctgtctcaacagggaag
ccctacccagtttgccata <mark>attaaa</mark> acttgggctactttgtttctattgag ta attcctcctc <mark>ca</mark> gttgtgccactgta 320
ttctcctttgttttgcgatttgaatctccttttacctttttttt
catatacacaaataaatggcatcttgagaatttaaaaggcaaacaaa
actggtgccccgagagcactgcctggagaattcacccctccttgcccggcgccccgccccggagggggggg
agaggttaacttggtggcctaggagaggggagaagccaggagaagcacttactccaaccacgtctttccactacatcat 640
cttgttttactaagtaggattttattttagggttcaaagaaacatgacatttattggtcaaaatattacactggttgtt 720
attttgttattgttttattttagtttttagaaaggattaatgtacaaaaagatttagagatctgtttcttaaagctacag 800
ggtttaaaaaaataaaatgagtgaaaatacttgatgtttcttgaaagataaatttaataataaata
acataaataaaaaagaaagccacaggcctgtaaattttatttgtaaaaaagcatttctatttttatacaaaatttttact 960
gcctcagaaataatggaagttatttattgcctattgttaacttattgggtacctaaagagcagttctctgtctagctag
accttctgaagtagtctctagaggaattgttctaggaatgggctttttttcaccgttgaaccctagacctgaagttcatg
ctttatcctctcgttgtttgtatctgtctgattattttgtttg
ctttccccgtatgtcacatttgttgaaactggaccttctccccctgtccctcaccccaaaacacccggaatccatcc
ttcgctctctccagatggattctcttggggtttcattgtgctgtggataaggagtgtaagaaatgcaaattatgtgaatg 1360
gctcggagactccctaatgacctaagatttgcattagtttttctcctgcacccttaaaagtgattttgttgccgctgcat 1440
agattctgtgtaactttttactcttccttgttttttccctag/ACATCTTCATGCCCGTTAGTTCATCGTTTGCCTAGCAT
GTCCCTGTGGCGTCTCAAAAAAAGTTTCATCGTCCCGTCATTGTTTCTGATGTCTTTCTGACCTCACATCATATTTGGT PAS-1a
TCTCCTACTGACCTTTGATCTAGTTTGACCTTTG <mark>AAATTT</mark> GCATGTGACCTCATCTAGCTATGA <mark>AT</mark> TCTGGGAAGTCAAT 1680
GTGAAAAACATTGCTGCATTCATGCAAGACTGAAATTTATTATTAGACAAATTCATTATAGAAAAAACCTGTGGCAAAAA 1760
GTAACAATAGCAATATGTGTCCAGGGACACAGAATGTTGGTTTCTAACAGACTACTTCCAAAAACAGTTTGAGAAAAA 1920
TTTGGGGCAAATGCTACCTATTTGTGTCACCTTTTGCTGAACTCACAGTTAGACAATCCATGGTTTAATGCACATGAAAT 2160
TACCTATATTTTATACTGTTTCAATGTACAGGAGAAAGGTTACTGTAAACTGTGTTATGTTGGTGCTTCTGTGAATGAG PAS-1b 2240
TTGTGGTTTCATCATGAGTCTTAATGTTCTTTGTTGAT <mark>AAGACA</mark> AGTTTAGAATTGGTTTACT <mark>TA</mark> ATA <mark>CA</mark> AAAAAAAAA 2320
AAAGAATTTCAAAAAAAAAAAGTTGTTTGCTTAAAAAAAA
2100

Fig. S4. (Continued)

2480 GATGTTTTTTTTTTAAAGAAAAAAGTGAAAATATATAGTGCCAAATTCCAAAGGTACTTCCTAGAGCT CAGT 2560 GTGTTTCTTGTGAGAAGTAATTTGATAACATGGGTATTTTATTATGTGTTTTGTATAAATCCCTAATATTTAAA ΔΑΑΑΑ 2640 AACAAAACAAAAAAAGGTTACAAAGTTTGTTAACTTGCTATCCTGTGGTCTTGTTGCCTGAAATTGTTATTGTT TGTTAT 2720 TTCTCTCTGATGTTTTTGTAAGACATTGTATAAGTGCCCATGTCCCACTTTTTTAACCACTCCGCACATCAGTGCTGTG 2800 AAGGCAACCTCACCATGTATTTTCTTCATAATCTATGGAAACCTCTAAGGTGAGAAAGTTTTGAACTTTTAACCC ттст 2880 2960 TTTTTCCCTGCATCTATCCTCTAAGTTGTTTCGGTTTGACTACTTTGTTCTTTGGTTAAGATCCAAAAGAAAACAGAAAA 3040 CAATTCCACGAGGCCAATCTAAAGGGAAAAAATCCTACACTACTTTTACTACTTTTGATTATTTCTCATTTTTGGGAAAA 3120 GAATTCCTAATGTGCTACTAGAATTCCTTCTTCAGTTTTAACGAGTAATTGGATAAACCCTGAGGGAAAACGGAGGTAGA 3200 TTCAGCACCTAACAATCCTGTATGCTTTTGAGATCACGTTTAGTGCTATGTCCTAGTCTAGAATATTTTCATATACCTTG 3280 CAGTAAAACGACTTTGTGGCAGGACAGTCTCTTGAGGGGTTTTGTTTCTGTTTCCTAAATACTCCTAAATAATATTTCTA 3360 ATCAGCCATTATGCTGGGGCATCTCTGATCCCAGTAGGTACCTCTGAATATACCAGGTGTCTGGAGTTAGAAGCCCATAG 3440 CCCTTTCCCAGCCTTTTTGGTTTTTTTTTTTTTTTGAACACATTTCATCTAAGTAAAGCTCAGTTCTTTATCACAATTTACTGA 3520 CCAAATACCTAGCACCAGTTCCTGCTGCCACTTTTTAAAGTGCCATATGACTTTCTACGAACAGGTACCTTGCTGTCTTG 3600 ACAAATCCTAATGTCACGCCTACAGCCCCAACACAAGCTCCAGTCTTCCTCTCGGCATGCCCTGGAAGCTTCT TCCAGT 3680 CTTCCTCTTCGGCATGCCCTGGAAGCTTCTTGGCCTCAGCTCCCCTTCCCCGCTCAGCACCCTGTTAGGATCAGTGTGTG 3760 CTTCT 3840 TATAGCACATGCACTTCCTTACAATGACATGATTTGTATTATCCTCACATGTGTTTACTACTGCTGGGGGCCTTCC TCAT 3920 CCTCTGAGGGCTATTTTGTACTTTCTGCAGCAATCAGCTTAATAACAACACTTATTGCACCTGTCTCTCTGAGAACAC 4000 GGTGTGTCTCGACACGTACCACGTACGTGGAAACACAAGAGCCCACCACTTGAATTTCTAAGACCATTTCATTCTGAAAC 4080 4160 CCACCTGCATGCATCTCCCCCAGAGGAAACACTGAGGGTAGGGGACAGGAGGCTCAGGACGCGCCCTCTGAATC GAGTGT 4240 TTCTTCTTCACAAGTCACCAAGAGAGGAGAAGTGGGGGGAAAGTCCTTTTTGCCCTTCTCCAAAAAATAACCTTCCACAGA 4320 4400 CAGAAGAAAACACACACAGGGAAACCGCTTTTTTTAATCAATTGTAGAGAATAGTCATTTTTAATCTAAATTAGAGAATT 4480 4560 4640 4720 TCATTTTGTTTAGAATTACAAAATAGTTTTTAAATATTGTCTGAGAAAAGCCAAAGTTAATGCAACCTAGTGGA ACTGT 4800 AAGACCATTTGAGTATTGTTTGTTTTATTGATGCATTTGGATTTTGTTGTTGATGGAATTTGAGCCAAAAAAA ΑΑΤΑΟ 4880 GCAGGCTTTCCTATTTCTACAACTGATTGTACTTATGCATTTTGTACCAGTGGAACTTTTTATACTGGAGATTAAAAAAA 4960 AAATGGAAATTTTTGTGGCTTGCTCTGGTGGGCCCCTGACAATGACTGATTTCAAGTTTGATTTCGGGTTGATT GATTGA 5040 TTGATTGATAGAAAGAAAGTTGCTTTTCTTTTGAGAATTAAAAACTTTGGCTTGATTTCTTTTTCCCTTTGCTTATATC 5120 TAGCATTAGAATTTTGTCTTAAAATAACAGCGGTAAGTTTCACTTTTTATTCTGTATTGTGCAGTTACACAATAAGGTAA 5200 TTAGATTTAGAAGTACTCAGTCACTTTAAGTGGATAAATGTATTAGTTAAAACTTTAGGGTTTGCTTTTTGC GTTTAG 5280 ATCAAAGTTTTTTCTGATTCTTCTGTCCTCATTGTGAACATAACCGTGTAGTTGAAACAGTCAAACTTATTTTTGTAATG

Fig. S4. (Continued)

PAS-2 5360 TATGTTATTGTGTGATGCAGTTTTTTGCTTCTGTCTCCAATATTAAACCATTTTCCTAATACTTGTTTCTCTCT CTGCGT 5440 GTTGTATTGTTGGTAGTCATTATATGTTGGTGATACATCTGCACACCTCACTGTTTCACGTATCTGTTTTTTC TATG 5520 TTGTGTAAAAAGATACAGTCGATTCCACTTAAGTGAATATCTGATTTGGGGAAGAGGAAACTGCACAGCAGCAC TTTG 5600 ATTTATGTACAACCGCCCACTTGAACTCCTGCTTCCAAGATTTATACACTTTTTTCCTACAGTTCACCTCAGTAACTGCC 5680 TTTGGGGTTGGATTTGCCAAGGTGTTTTACGGTCTTTTGACCACACGGTTAATTTCCCCTTCTCCCCTCCC CATGA 5760 GTCTATCCACGCAGTTCTTAACATACATTCCAAACTGCTGCGGGGTTTCCTCCACACCCACGAGGCCAAACCCGCATC 5840 ATTACAGTCCAACTCTTCTTAGACACTAATCACTTTCTAAAAGAAGTGAGTTCTCCAGGGAAGAAAAATCAACTTAGCCA 5920 TATTTA 6000 TTTCTACAAGCTAATTGAATCCAGGAGCAGCTTTAATTATTAACACTAACGGAAGAGAAAAGAAGTATTTCCAAGGGCTC 6080 TACTGT 6160 CCTTTTCCAAACAAAAGCTAATAACGCCATACGCATCCACACACCTCCTCCTGGATGAACCTAAGTCTCGTCC ACTGT 6240 CACCCCAAGGCCAGTTATCAAAAACTGTTCCTTCTGCCCTCAAAGACTGAAGCCGCAGGCCCTGTTCTGCCTCTGCTC 6320 AGGAATCTGATTGCTCTTAAAGTGCTCTTACAAGATTCCGTCGATGTTTGCTCCCTCTGTCTTATCTCCTTC CAGTAT 6400 TTTCAGACGCACAAGCTCTGACTCAGGCCACCCAGCACCTGGCAGCCTTCCTCTGCGTGCCAGTGAAATCTCTCCCAGTA 6480 GGTGCTCAGGCTCTTCACCCTCTCGTTGCAGCAGCCCCTCTGGCCCGAGATTGTATTTTCTAGTCTGCTCAAGAACAGAG 6560 GGGCCTGGCGAGGTGGAGGAAGGGAGGCTCCCTCCCGGGGCTGGCCGCCTCAGTGATGGTGATGCCACCAGC CAGCC 6640 TGCACGGATGGAGCTCACTTCCTACCATCACTTCTGCTTTCAGTCTCCTCATAGCCTTCGGCAGTCTCCTCTGAAAACAC 6720 ACTGCCTACGTTAGTGGGAAAGGATGGCTAAGGGTGTGATTTCTTTTTATTAGGGAGGTGGGAGGGGGGATGTTGGGAGGC 6800 ATGGGGGGTGATTAAATTATCATTTCCAAGGTGCAAGTGTTTATTTTTCAAATCTATAAGTACATAGTTCCACC ATTTTG 6880 GCACAAGACAAAATACTCATGCTAAGCAAAGGAAGAGAAAGACAAAGCCAGTTTTTGTTGCTTTCTAAAGCAACAAATA 6960 CATA 7040 GGGAT 7120 7200 GGCTCGGACTCTTAGCTGGCAGGCTAGGACACACAGTGAACATTAATGTACTGTGAATCGTTCCTGATAAGTGAATAAAA 7280 CATTGTGACCAAACAATCAGCTTATTCACTTATCAGGAATCGACTGTTCTGCTAATTTGCTGTTGTTGTTTT гстст 7360 GTTGTAGTTCACTATTTTTGCTGTGTTCTGTTTTCTCCTCTCATGCTTGGGCAAATCACTGGGAATATTATCTT CATGTG 7440 ACCATGAAACGTTTCTATTGAGTGAAAATGATATCTTAACAAAATCTATGCACTTGCTATCAGGAACACAATA GGATG 7520 'AS-3a AAGTTA TGTCTTATATATTGAACTATATAGTACTCGATTTCTTA<mark>AATAAA</mark>GCTTAA<mark>GA</mark>AAG<mark>GA</mark>CTCTGTTGTGTGTG 7600 ATGTGATTATTTTCAAAGCAGTGTTTCTAAGGGGTATTTTTGTTCTTTTAAGTCTTGAGTGATACAGGATATTTTTCATT \S-3b AAAATTATATCACTGGCTTTATGACTTAATATTCAATAAATTGACTTT

Fig. S4. Human CELF2 3'UTR sequence and annotation.



Fig. 55. Additional analysis of CELF2 3'UTR. (A) 3' RACE using forward primers indicated on the schematic (*Upper*) to identify sites of cleavage/PAS along the entire 3'UTR of CELF2. Products derived from confirmed genuine polyadenylation sites are noted with a "PAS" label. Note that none of these primers are position to detect PAS2. Products that correspond to annealing of the reverse 3' RACE primer to genomically encoded stretches of adenosines are noted by a "gPA." The prominent band observed with primer 4 could either be PAS1 or an extensive gPA; however, because we have no evidence for use of PAS1 in the Northern blot (Fig. 4C), we assign this as a gPA. Multiple faint bands detected with primers 5 and 6 most likely correspond to gPA, as no potential PAS is within 1 kb. (*B*) RT-PCR showing an increase in the retention of intron 13–14 in response to stimulation with PMA. Primers I-5' and I-3' interrogate the presence of RNA contain the 5' and 3' end of the intron retained with the adjoining exon. Primers Ex13–E14 interrogate the junction of exon 13 to exon 14 with the intron spliced out. The length of the intron precludes amplification from exon 13 to exon 14 across the retained intron. Actin is a loading control.



Fig. 56. Splicing of CELF2 3'UTR does not target message for NMD. (*A*) Western blot of Upf1 depletion by transfection with an antisense morpholino oligonucleotide (AMO). hnRNP L is used as a loading control. (*B*) Primers used to detect intron 13 retained (IR) and intron 13 spliced (IS) forms of CELF2. Upstream primer is specific for exon 6 to avoid potential ambiguity because of exon 6 skipping, although this does not appear to induce NMD (see *E*). (*C*, *Left*) Quantification of CELF2 messages with the intron retained (blue) and spliced (red) versions of the 3'UTR measured 60 h after Upf1 knock-down as in *A*. (*Right*) Parallel analysis of the NMD-resistant (blue) and NMD-sensitive (red) versions of Tra2B mRNA, used as a positive control for inhibition of NMD. (*D*, *Left*) Quantification of CELF2 messages with the intron retained (blue) and spliced (red) versions of the 3'UTR measured at indicated times after treatment with cycloheximide. (*Right*) Parallel analysis of the NMD-resistant (blue) and Spliced (red) versions of Tra2B mRNA, used as a positive control for inhibition of NMD. (*D*, *Left*) NMD. (*E*) RT-PCR with primers to exon 5 and 7 show no difference in exon 6 skipped product upon Upf1 KD, indicating that this splicing event does not trigger significant NMD in Jurkat cells.



Fig. 57. CELF2 mRNA export to cytoplasm is independent of whether 3'UTR is spliced or retained. (A) Schematic of primers used to detect the intron retained (IR) and intron spliced (IS) versions of the CELF2 3'UTR. (B) RT-PCR analysis of indicated RNAs in nuclear and cytoplasmic pools. Purity of nucleo-cytoplasm was assessed by MALAT1, a strongly retained nuclear RNA (*Upper*). Strong retention of incompletely spliced versions of actin was also observed (*Lower*). In contrast, both the IR and IS versions of the CELF2 mRNA are observed in the cytoplasm with similar efficiencies, indicative that retention of intron 13–14 does not induce nuclear retention. Actin mRNA is used as a control of another mRNA efficiently exported from the nucleus. Percent cytoplasmic message was calculated at (cyto)/(ctyo+nuclear) from replicate experiments. The variation in the data is shown in parentheses.



Fig. S8. Isolation of CD4/CD8 DP and DN thymocytes. Representative FACS plots of gating to isolate DP and DN thymocytes. Single-cell thymocytes suspensions were stained with anti-CD8-APC, anti-CD4-FITC, and a cocktail of antibodies to deplete nonthymocytes that were labeled with PE. Cells were first gated for PE⁻ (*Upper Left*), then by forward scatter/side scatter for live cells (*Lower Left*), and finally DN and DP cells were collected based on the gates shown (*Bottom Right*). Typical yields were $\sim 2 \times 10^6$ DP and 5×10^5 DN cells.

Table S1.	Primers	used for	indicated	experiments
-----------	---------	----------	-----------	-------------

Experiment	Forward	Reverse		
RT-PCR				
rCELF2	TCTGCTCGACAGCACGCAGTG	ATTGTGCAGTGCATTCTGGGC		
eCELF2	GGATTACAAGGATGACGACG	ATTGTGCAGTGCATTCTGGGC		
Actin	CCCTGGCATTGCCGACAGG	GCCGATCCACACGGAGTACT		
APP	CGTGGAGCTCCTTCCCGTGAATGG	CCCACCATGAGTCCAATGATTGCACC		
CTTN	CGCCGTTGGCTTTGAGTATCAAGGC	CTTATCCATCCGATCCTTCTGCACC		
FUS	GGCAGCGGTGGCTATGGACAGC	CGGAGTCATGACGTGATCCTTGG		
MRPL42	CCAAAATTCGGGCGTACATGTGCAG	GTGTGTTCATATGGAATGTCCACAGAAGGG		
NAB2	GGAGAACAGAGTCACCCTGAAATCC	GAGGTCCTGGGTCTGAAGACACC		
OPA1	GTTCTCCGGAAGAAACGGCGTTTAG	TGATGAATGCCTTTGTCATCTTTCTGC		
PPP1R12A	CCAAGCACCACATCAACAACAAG	CTGTGTTGATCTTCTAGATTGTCTTGC		
SRPK2	F1:GTCGTCCTCTTCAGAAAGGCCGGAG	TTGCAGTAGTCCGCAGGGTCCTCTTGC		
	F2:CCGGAAAGTGCTGGCCATTCAGGC			
Tra2B	GCCTCCTTAAGGAAGGTGCAAGAGG	R1:GTCAAATGACGACTTCCGCATTTTCC		
		R2:AGCGAGACCGTGACCGGGTATAATGC		
CELF2-E5/7	GCAGATAGTGAAAAGTCCAACGCTGTGG	CCCAGGTGGCAGTGTTGAGCTGCTGC		
CELF2-E6	TGCATCAGTCTCAGACCATGGAG			
CELF2-E12	CACAGGAATTTGGAGACCAGG			
CELF2-I13 (IR)		AAGTTGGCAATGTGGTCCTCCTCTGC		
CELF2-E14 (IS)		GCTAGGCAAACGATGAACTAACGGGC		
MALAT1	AGATTTCCCAAGCAGACAGC	ACCGCACAGCTCGGGCGAG		
qPCR				
eCELF2	CCAGGGTAGGGCTGATAAGG	TGAGTGATCCAAAGCTCCGT		
Actin	CCCTGGCATTGCCGACAGG	GCCGATCCACACGGAGTACT		
3' RACE				
1F (PASi)	AGGCCTCTCGAGCCTAACCCCAGAGGCTCCCTGC			
2F	AGGCCTCTCGAGCATAAATACATAAATAAAAAAAAAGAAAG			
3F	AGGCCTCTCGAGGATTTTGTTGCCGCTGCATAGATTCTGTG			
4F	AGGCCTCTCGAGGTTGGTGCTTCTGTGAATTAAGTTGTGG			
5F	AGGCCTCTCGAGCCCTGCATCTATCCTCTAAGTTGTTTCGG			
6F	AGGCCTCTCGAGGGTTATAGTTGGCTTGTGCTACTCTGG			
7F	AGGCCTCTCGAGCGATTCCACTTAAGTGAATATCTGATTTGG			
8F	AGGCCTCTCGAGGGAGGTAAAGTTTCTCACACTCAAGTCG			
9F	AGGCCTCTCGAGGGCTAAGGGTGTGATTTCTTTTTATTAGG			
PAS2	GGCTTGATTTCTTTTTTCCCTTTGCTTATATCTAGC			
PAS3	GCACTTGCTATCAGGAAACAATACTGG			

Dataset S1. Complete RASL-Seq data PMA and CELF2 knock-down

Dataset S1

PNAS PNAS

Dataset S2. RASL-Seq data for ~200 PMA-responsive genes

Dataset S2

Dataset S3. RASL-Seq data for thymocytes

Dataset S3