HAL
open science

The RNA-binding proteins CELF1 and ELAVL1 cooperatively control the alternative splicing of CD44

Géraldine David, David Reboutier, Stéphane Deschamps, Agnès Méreau, William Taylor, Sergi Padilla-Parra, Marc Tramier, Yann Audic, Luc Paillard

To cite this version:

Géraldine David, David Reboutier, Stéphane Deschamps, Agnès Méreau, William Taylor, et al.. The RNA-binding proteins CELF1 and ELAVL1 cooperatively control the alternative splicing of CD44. Biochemical and Biophysical Research Communications, 2022, 626, pp.79-84. 10.1016/j.bbrc.2022.07.073 . hal-03757363

HAL Id: hal-03757363

https://hal.science/hal-03757363v1

Submitted on 14 Dec 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The RNA-binding proteins CELF1 and ELAVL1 cooperatively control the alternative splicing of *CD44*

Géraldine David,[1] David Reboutier,[1] Stéphane Deschamps,[1] Agnès Méreau,[1] William Taylor,[1] Sergi Padilla-Parra,[1,2] Marc Tramier,[1,3] Yann Audic,[1,*] Luc Paillard[1,*]

1 Univ Rennes, CNRS, IGDR (Institut de Génétique et Développement de Rennes)-UMR 6290, F-35000 Rennes, France

2 Present address: Department of Infectious Diseases, King's College London, Faculty of Life Sciences & Medicine, London, UK

3 Univ Rennes, CNRS, Inserm, BIOSIT-UMS 3480, US_S 018, F-35000 Rennes, France

* Corresponding authors: LP (luc.paillard@univ-rennes1.fr) and YA (yann.audic@univ-rennes1.fr).

ACCEPTED MANUSCRIPT / CLEAN COPY

ABSTRACT

CD44 mRNA contains nine consecutive cassette exons, v2 to v10. Upon alternative splicing, several isoforms are produced with different impacts on tumor biology. Here, we demonstrate the involvement of the RNA-binding proteins CELF1 and ELAVL1 in the control of *CD44* splicing. We show by FRET-FLIM that these proteins directly interact in the nucleus. By combining RNAi-mediated depletion and exon array hybridization in HeLa cells, we observe that the exons v7 to v10 of *CD44* are highly sensitive to CELF1 and ELAVL1 depletion. We confirm by RT-PCR that CELF1 and ELAVL1 together stimulate the inclusion of these exons in *CD44* mRNA. Finally, we show in eight different tumor types that high expression of *CELF1* and/or *ELAVL1* is correlated with the inclusion of *CD44* variable exons. These data point to functional interactions between CELF1 and ELAVL1 in the control of *CD44* splicing in human cancers.

KEYWORDS

Gene expression/CD44/RNA-binding protein/Post-transcriptional control/Cancer

ACCEPTED MANUSCRIPT / CLEAN COPY

INTRODUCTION

CD44 encodes a cell surface receptor that is involved in cell-to-matrix adhesion and regulates a number of signaling pathways [1]. While *CD44* expression is a recognized biomarker in several human cancers, understanding of its roles in tumor biology is complicated by the existence of several splice isoforms with different functions. The human *CD44* genes contains 19 exons. Exons 1 to 5 and 15 to 19 are constitutive, whereas exons 6 to14 also called "variable" exons v2-v10 are cassette exons, which can either be skipped or included. Depending on their combinatorial inclusion, 22 different *CD44* isoforms have currently been described [1,2]. The variable exons encode a part of the extracellular domain of CD44 immediately adjacent to the transmembrane domain. Their presence can have an impact on the receptor's capacity to interact with its ligand hyaluronic acid, explaining the different functions of the CD44 isoforms.

Breast epithelial cells essentially express *CD44* isoforms that include the variable exons. During epithelial-mesenchymal transition (EMT), the incorporation of the variable exons decreases. The switch from isoforms including the variable exons (isoforms CD44v) to isoforms devoid of them ("standard" isoform, CD44s) is required for EMT to occur [3]. Similar causal relationships between *CD44* splicing and EMT or tumor invasiveness have been reported in several other cancer types [1]. CD44s is also associated with cancer cell stemness [4].

Positive regulators of variable exons inclusion include SRSF2 in a breast cancer line (exon v6) [5], U2AF2 in melanoma (exons v8-10) [6], and TRA2B in a colon cancer cell line (exon v10) [7]. Negative regulators include PCBP1 in a hepatocarcinoma cell line (exons v3, v5, v6, v8 and v10) [8], HNRNPL and its paralog HNRNPLL in a breast cancer cell line (exon v10) [9] and in colorectal cancer (exon v6) respectively [10], and SRSF9 in a colon cancer cell line (exon v10) [7]. The epithelium-specific splicing factor ESRP1 promotes the formation of CD44v in breast cancer by promoting the inclusion of all variable exons [3,11], whereas ZMAT3 negatively regulates all of them in colorectal carcinoma [12].

ACCEPTED MANUSCRIPT / CLEAN COPY

Here, we show that the RNA-binding proteins CUGBP Elav-like family member 1 (CELF1) and ELAV-like RNA-binding protein 1 (ELAVL1) control *CD44* splicing in most human cancers. In addition to cytoplasmic roles [13–18], both CELF1 and ELAVL1 control alternative splicing [19–21]. We observe that CELF1 and ELAVL1 directly interact in the nucleus and together control the splicing of *CD44* in HeLa cells. Notably, they promote the inclusion of exons v7 to v10. By mining the The Cancer Genome Atlas (TCGA) database, we observe a strong dependency of the expression levels of *CELF1* and *ELAVL1* on *CD44* alternative splicing in 8 different cancer types. We conclude that CELF1 and ELAVL1 control *CD44* alternative splicing in a large variety of tumors.

MATERIALS AND METHODS

Materials

The sequences of the siRNAs used were: *CELF1*: GAGCCAACCUGUUCAUCUA, GCUCUUUAUUGGUAUGAUU, GCUGCAUUAGAAGCUCAGA. *ELAVL1*: GAGGCAAUUACCAGUUUCA, UCUUAAGUUUCGUAAGUUA. Negative control (Ctrl): GUCUAGACGAGUGUGACAU. Luciferase (Luc): CAUUCUAUCCUCUAGAGGAUG.

The sequences of the PCR primers are: *WNK1* CTCCTCAACAGACAGTGCAG, GAAAGTACCCAGGTGTAGCCA; *PHACTR2* GAAAATTCAAACGGGCACAT, CTTTGAAGCTTTGGGACGAG; *CD44* CAGAAGGAACAGTGGTTTGG, GGGTGGAATGTGTCTTGGTC.

We constructed plasmids encoding fluorescent protein-tagged CELF1 and ELAVL1 by Gateway recombination with the pCS-EGFP-DEST and pCS-Cherry-DEST (Addgene) vectors [22].

Cell manipulations and RNAi

ACCEPTED MANUSCRIPT / CLEAN COPY

In all experiments, HeLa Kyoto cells were transfected with jetPRIME (Polyplus. In double-depletion experiments, we used 5 nM anti-ELAVL1 siRNA (equimolar mix of 2 siRNA) plus 5 nM anti-CELF1 siRNAs (equimolar mix of 3 siRNA). In the other experiments, we used the same concentration of anti-ELAVL1 or anti-CELF1 siRNA plus 5 nM Ctrl siRNA to achieve the same final concentration of total siRNA (10 nM). Control experiments were with 10 nM control siRNAs or 10 nM anti-luciferase siRNAs. After 48 h, we extracted total RNAs with TRIzol (Invitrogen) or NucleoSpin RNA extraction kit (Macherey-Nagel). RNAs were sent to the Plateforme Génomique de Nantes for hybridization on SurePrint G3 human exon microarrays (Agilent), or used as matrices to synthesize cDNA. PCR was done on cDNAs using Cy3-labelled or ^{32}P-labelled primers. The number of PCR cycles was determined empirically for each primer pair to remain within the exponential phase of amplification.

FRET-FLIM

HeLa Kyoto cells were transfected (jetPRIME) with 15 ng of the expression vectors for EGFP-labelled proteins and 150 ng for mCherry-labelled proteins and observed at 37 °C after 24h. We used a fastFLIM system with a plan APO 63x/1.4NA oil immersion objective (Leica). EGFP donor excitation was with a narrow 480/23 nm spectral band selected from a Fianium white-light laser sent to a CSUX1 microscope (Yokogawa) through a dichroic mirror (Di01-T405/488/568/647, Semrock) inside the spinning head. EGFP emissions were filtered on the spinning filter wheel (525/50 nm), then acquired with a time-gated intensified CCD camera. The FLIM calculation to determine the mean fluorescence lifetime on a pixel-by-pixel basis was done online using flimager [23]. Three independent series of transfection were carried out, with ten cells analyzed for each transfection, resulting in a total of 30 nuclei analyzed per condition.

Data analysis

ACCEPTED MANUSCRIPT / CLEAN COPY

The raw microarray data are available in GEO as dataset GSE118981. For TCGA, we retrieved the data from TCGA SpliceSeq [24]. The IDs of alternative splicing events are as follows: inclusion or skipping of exons v2-v10, CD44_ES_15127 and CD44_ES_15130 (these two events differ by 3 nucleotides in exon v5 due to an alternative 3' splice site); exons v3-v10, CD44_ES_15128 and CD44_ES_15131; exons v4-v10, CD44_ES_15129 and CD44_ES_15132; exons v6-v10, CD44_ES_15133; exons v8-v10, CD44_ES_15143. We analyzed all data using in-house R scripts that are available as supplementary material.

RESULTS AND DISCUSSION

CELF1 and ELAVL1 interact in the cell nucleus

We have shown previously that CELF1 and ELAVL1 co-immunoprecipitate [25]. We used fluorescence lifetime imaging-based Förster fluorescence resonance energy transfer (FRET-FLIM) to reveal the subcellular compartment in which the interaction between CELF1 and ELAVL1 occurs. FRET, which reveals close proximity (<10 nm) between the donor (EGFP) and the acceptor (mCherry), is detected by the decrease of the donor lifetime [26]. We expressed EGFP-CELF1 in HeLa cells. The fluorescence lifetime of EGFP-CELF1 fusion was about 2.5 ns (Figure 1A). EGFP lifetime was unchanged in cells co-expressing mCherry-histone 2B (H2B). In contrast, the fluorescence lifetime of EGFP-CELF1 was reduced with co-expressed mCherry-CELF1. FRET between EGFP and mCherry in cells co-expressing EGFP-CELF1 and mCherry-CELF1 was expected as CELF1 homo-oligomerises [27]. Importantly, mCherry-ELAVL1 also reduced the fluorescence lifetime of EGFP-CELF1 to approximately 2.2 ns (Figure 1A). Swapping the fluorophores produced the same results (Figure 1B). These data indicate that CELF1 and ELAVL1 directly interact in the nucleus.

ACCEPTED MANUSCRIPT / CLEAN COPY

135

CELF1 and ELAVL1 control *CD44* splicing

Having demonstrated that CELF1 and ELAVL1 interact in cell nuclei, we hypothesized that they control the splicing pattern of common RNA targets. We depleted CELF1 and/or ELAVL1 in HeLa cells and measured the consequences of these depletions on RNA splicing by exon microarray
140 hybridization. We focused on RNAs bound by both CELF1 and ELAVL1 in CLIP-seq experiments on HeLa cells (Table S1) [28,29]. This analysis, which we performed in triplicate, identified 62 exon probes in 57 genes that were differentially included in mature mRNAs in at least one depletion condition compared to the control, with a false discovery rate of 0.1. Table S2 lists these probes with their normalized exon values (exon probe fluorescence intensity normalized to the expression
145 of the gene) and the splicing indices (SI, log2-ratios of normalized exon values in depleted cells compared to the normalized exon values in control cells).

We chose two RNAs that are differently controlled by CELF1 and ELAVL1 to validate these results by RT-PCR (Figure 2). *PHACTR2* contains a cassette exon with a higher SI following CELF1 depletion (1.07) than following ELAVL1 depletion (0.28, Table S2). Accordingly, in RT-
150 PCR, depleting CELF1 strongly increased the inclusion of the cassette exon whereas depleting ELAVL1 had virtually no effect (Figure 2A). Conversely, *WNK1* is more strongly controlled by ELAVL1 than by CELF1 in microarray experiments (SI = 1.05 and SI = 0.43, respectively, Table S2), and in RT-PCR (Figure 2B). Moreover, for both *PHACTR2* and *WNK1*, the splicing patterns in non-transfected cells are very similar to those in cells transfected with two different control siRNA,
155 that used in microarray experiments (Ctrl) or an anti-luciferase siRNA. Therefore, the control siRNA has no effect on the splicing patterns of these transcripts. In addition, the modifications of the splicing patterns were the same when using the siRNAs against either *CELF1* or *ELAVL1* individually or as mixed pools, suggesting that the changes observed in splicing patterns were not due to siRNA off-target effects. Hence, the experiments shown in Figure 2 validate the exon

ACCEPTED MANUSCRIPT / CLEAN COPY

microarray results.

For 58 of the 62 differential probes, the splicing indices have the same sign after depletion of CELF1 as after depletion of ELAVL1 (Figure 3A). This suggests that CELF1 and ELAVL1 control the splicing patterns of their presumptive common targets in the same direction ($p = 2*10^{-10}$, chi-squared test). Of these differential probes, 2 are localized within exons of *CD44*, and we focused on this gene. Microarray hybridization demonstrated that depleting CELF1 and/or ELAVL1 significantly reduced the inclusion of variable exons v8 and v10. The inclusion of v7 and v9 was also reduced albeit statistically not significantly. The inclusion of the flanking exons v6 and 15 was the same in all conditions (Figure 3B, Table S2). Interestingly, CLIP-seq experiments [28] have identified several CELF1-binding clusters on the *CD44* pre-mRNA (Figure 3B), suggesting that CELF1 can directly control *CD44* splicing. We performed RT-PCR using primers targeting exons v6 and 15. Of the 16 isoforms which can theoretically be produced by the combinatorial inclusion or skipping of four exons, we detected seven, whose identities were confirmed by Sanger sequencing of the amplimeres (Figure 3C). We quantified the percentages of inclusion of variable exons v7 to v10 from the cumulative percentages of all isoforms containing the exons of interest. Depleting CELF1 or ELAVL1 reduced the inclusion of all of these exons. Furthermore, depleting both proteins reduced their incorporation into mature mRNAs even more (Figure 3C). Hence, CELF1 and ELAVL1 act together to stimulate the inclusion of the variable exons v7 to v10 of *CD44*.

The inclusion of *CD44* internal exons is higher in cancer samples with high *CELF1* and *ELAVL1* transcript levels

The above data show that, in HeLa cells, CELF1 and ELAVL1 stimulate the usage of the *CD44* variable exons v7 to v10. We next wanted to test the relevance of these findings to human cancers using The Cancer Genome Atlas (TCGA) database. We analyzed all splicing events that

ACCEPTED MANUSCRIPT / CLEAN COPY

involve *CD44* variable exons v2 to v10 using TCGA SpliceSeq [24]. The alternative splicing events that are annotated in TCGA SpliceSeq are the inclusion or skipping of exons v2 to v10, v3 to v10, v4 to v10, v6 to v10 and v8 to v10 between exons 5 and 15. The inclusion of exons v7 to v10 between exons v6 and 15 (as in Figure 3) is not annotated in TCGA SpliceSeq. We detected alternative splicing of *CD44* between exons 5 and 15 for 18 TCGA cohorts of 27. For each of these 5 alternative splicing events in the 18 cohorts, we calculated the median percentage of inclusion of the internal exons in half the samples that express *CELF1* or *ELAVL1* the most and in half the samples that express *CELF1* or *ELAVL1* the least. We did the same for samples that simultaneously express *CELF1* and *ELAVL1* the least or the most. Figure 4A, lower panel, shows the difference between the groups of high and low expression, and the associated *p*-values (Wilcoxon tests). Positive values indicate cases where variable exons of *CD44* are more included in patients with high levels of *CELF1* or *ELAVL1*. When considering only statistically significant differences ($p < 10^{-2}$), the inclusion of the variable exons is systematically higher in patients with high levels of *CELF1* (19/19) or both *CELF1* and *ELAVL1* (12/12), but not of *ELAVL1* (8/18). At a first glance, this suggests that ELAVL1 is not related to *CD44* splicing in human tumors. However, the density plot in Figure 4A, upper panel, indicates that the difference between high and low levels of expression is greater when both *CELF1* and *ELAVL1* levels are used to construct the high/low groups than when only *CELF1* is ($p = 1.8*10^{-4}$, Wilcoxon test). Figure 4B shows the cohorts with significantly higher inclusion of *CD44* variable exons in samples with high expression of *CELF1*, *ELAVL1*, or both. As an example, Figure 4C shows that the low *CELF1* expression group of breast cancer (BRCA) samples have less inclusion of the variable exons v8-v10 (left panel). The same is true for the low *ELAVL1* expression group (middle panel). The effect size is larger when *CELF1* and *ELAVL1* expression are used together to define the high/low groups (right panel). Together, these data show that CELF1 and ELAVL1 contribute to the control of *CD44* alternative splicing in human tumors, by promoting together the inclusion of *CD44* variable exons.

ACCEPTED MANUSCRIPT / CLEAN COPY

Conclusion

The RNA-binding proteins ELAVL1 and CELF1 have documented roles in epithelial-mesenchymal transition (EMT). ELAVL1 binds to and stabilizes the *SNAI* (Snail) and *ZEB1* mRNAs. This supports translation of the encoded proteins, which is crucial for EMT [30–32]. ELAVL1 is also involved in sustaining TGFbeta signaling during EMT [33]. CELF1 activates the translation of core genes that drive EMT [14]. Consequently, both ELAVL1 and CELF1 are considered positive regulators of EMT. Here, we find that they promote the inclusion of *CD44* variable exons in several cancers. Since the switch to CD44 isoforms devoid of the variable exons is mandatory for EMT [1–4], the control exerted by CELF1 and ELAVL1 on *CD44* splicing suggests that they negatively regulate EMT. How is it possible to reconcile previous data with ours? Splicing takes place in the nucleus while the control of mRNA stability and translation are cytoplasmic events. It could be that CELF1 and ELAVL1 inhibit EMT onset in the nucleus by promoting a splicing pattern of *CD44* that maintains epithelial identity, whereas in the cytoplasm they activate pro-EMT factors. Interestingly, a translocation of ELAVL1 from the nucleus to the cytoplasm is observed during adrenaline-mediated induction of EMT in pancreatic cancer cells [33]. It will be important in future studies to systematically document the subcellular localisation of ELAVL1 and CELF1 before, during and after EMT, and the repertoires of controlled genes during these phases.

ACKNOWLEDGEMENTS

The results shown here are based in part on data generated by the TCGA Research Network (http://cancergenome.nih.gov/). Microarray hybridizations were carried out on the Plateforme Génomique de Nantes.

This research was funded by a grant from the Ligue contre le Cancer (comités 35, 22, 29) to

ACCEPTED MANUSCRIPT / CLEAN COPY

LP. GD was supported by a joint doctoral fellowship from the Ligue contre le Cancer and by the Région Bretagne (ARED).

REFERENCES

[1] M.H. Mesrati, S.E. Syafruddin, M.A. Mohtar, A. Syahir, CD44: A Multifunctional Mediator of Cancer Progression, Biomolecules. 11 (2021). https://doi.org/10.3390/BIOM11121850.

[2] L. Prochazka, R. Tesarik, J. Turanek, Regulation of alternative splicing of CD44 in cancer, Cell. Signal. 26 (2014) 2234–2239. https://doi.org/10.1016/J.CELLSIG.2014.07.011.

[3] R.L. Brown, L.M. Reinke, M.S. Damerow, D. Perez, L.A. Chodosh, J. Yang, C. Cheng, CD44 splice isoform switching in human and mouse epithelium is essential for epithelial-mesenchymal transition and breast cancer progression, J. Clin. Invest. 121 (2011) 1064–1074. https://doi.org/10.1172/JCI44540.

[4] H. Zhang, R.L. Brown, Y. Wei, P. Zhao, S. Liu, X. Liu, Y. Deng, X. Hu, J. Zhang, X.D. Gao, Y. Kang, A.M. Mercurio, H.L. Goel, C. Cheng, CD44 splice isoform switching determines breast cancer stem cell state, Genes Dev. 33 (2019) 166–179. https://doi.org/10.1101/gad.319889.118.

[5] T.J. Loh, H. Moon, S. Cho, D.W. Jung, S.E. Hong, D.H. Kim, M.R. Green, X. Zheng, J. Zhou, H. Shen, SC35 promotes splicing of the C5-V6-C6 isoform of CD44 pre-mRNA, Oncol. Rep. 31 (2014) 273–279. https://doi.org/10.3892/OR.2013.2812.

[6] P. Zhang, S. Feng, G. Liu, H. Wang, A. Fu, H. Zhu, Q. Ren, B. Wang, X. Xu, H. Bai, C. Dong, CD82 suppresses CD44 alternative splicing-dependent melanoma metastasis by mediating U2AF2 ubiquitination and degradation, Oncogene. 35 (2016) 5056–5069. https://doi.org/10.1038/ONC.2016.67.

[7] J. Oh, Y. Liu, N. Choi, J. Ha, D. Pradella, C. Ghigna, X. Zheng, H. Shen, Opposite Roles of Tra2β and SRSF9 in the v10 Exon Splicing of CD44, Cancers (Basel). 12 (2020) 1–12. https://doi.org/10.3390/CANCERS12113195.

[8] T. Zhang, X.H. Huang, L. Dong, D. Hu, C. Ge, Y.Q. Zhan, W.X. Xu, M. Yu, W. Li, X. Wang, L. Tang, C.Y. Li, X.M. Yang, PCBP-1 regulates alternative splicing of the CD44 gene and inhibits invasion in human hepatoma cell line HepG2 cells, Mol. Cancer. 9 (2010). https://doi.org/10.1186/1476-4598-9-72.

[9] T.J. Loh, S. Cho, H. Moon, H.N. Jang, D.R. Williams, D.W. Jung, I.C. Kim, C. Ghigna, G. Biamonti, X. Zheng, H. Shen, hnRNP L inhibits CD44 V10 exon splicing through interacting with its upstream intron, Biochim. Biophys. Acta. 1849 (2015) 743–750. https://doi.org/10.1016/J.BBAGRM.2015.01.004.

ACCEPTED MANUSCRIPT / CLEAN COPY

[10] K. Sakuma, E. Sasaki, K. Kimura, K. Komori, Y. Shimizu, Y. Yatabe, M. Aoki, HNRNPLL, a newly identified colorectal cancer metastasis suppressor, modulates alternative splicing of CD44 during epithelial-mesenchymal transition, Gut. 67 (2018) 1103–1111. https://doi.org/10.1136/GUTJNL-2016-312927.

[11] T. Yae, K. Tsuchihashi, T. Ishimoto, T. Motohara, M. Yoshikawa, G.J. Yoshida, T. Wada, T. Masuko, K. Mogushi, H. Tanaka, T. Osawa, Y. Kanki, T. Minami, H. Aburatani, M. Ohmura, A. Kubo, M. Suematsu, K. Takahashi, H. Saya, O. Nagano, Alternative splicing of CD44 mRNA by ESRP1 enhances lung colonization of metastatic cancer cell, Nat. Commun. 3 (2012). https://doi.org/10.1038/NCOMMS1892.

[12] B.R. Muys, D.G. Anastasakis, D. Claypool, L. Pongor, X.L. Li, I. Grammatikakis, M. Liu, X. Wang, K. V. Prasanth, M.I. Aladjem, A. Lal, M. Hafner, The p53-induced RNA-binding protein ZMAT3 is a splicing regulator that inhibits the splicing of oncogenic CD44 variants in colorectal carcinoma, Genes Dev. 35 (2021) 102–116. https://doi.org/10.1101/GAD.342634.120.

[13] M. Cibois, C. Gautier-Courteille, A. Vallee, L. Paillard, A strategy to analyze the phenotypic consequences of inhibiting the association of an RNA-binding protein with a specific RNA, RNA. 16 (2010) 10–15. https://doi.org/10.1261/rna.1742610.

[14] A. Chaudhury, S. Cheema, J.M. Fachini, N. Kongchan, G. Lu, L.M. Simon, T. Wang, S. Mao, D.G. Rosen, M.M. Ittmann, S.G. Hilsenbeck, C.A. Shaw, J.R. Neilson, CELF1 is a central node in post-transcriptional regulatory programmes underlying EMT, Nat. Commun. 7 (2016) 13362. https://doi.org/10.1038/ncomms13362.

[15] I.A. Vlasova, N.M. Tahoe, D. Fan, O. Larsson, B. Rattenbacher, J.R. Sternjohn, J. Vasdewani, G. Karypis, C.S. Reilly, P.B. Bitterman, P.R. Bohjanen, Conserved GU-rich elements mediate mRNA decay by binding to CUG-binding protein 1, Mol. Cell. 29 (2008) 263–270. https://doi.org/10.1016/j.molcel.2007.11.024.

[16] A. Masuda, H.S. Andersen, T.K. Doktor, T. Okamoto, M. Ito, B.S. Andresen, K. Ohno, CUGBP1 and MBNL1 preferentially bind to 3’ UTRs and facilitate mRNA decay, Sci. Rep. 2 (2012) 209. https://doi.org/10.1038/srep00209.

[17] N. Mukherjee, D.L. Corcoran, J.D. Nusbaum, D.W. Reid, S. Georgiev, M. Hafner, M. Ascano, T. Tuschl, U. Ohler, J.D. Keene, Integrative regulatory mapping indicates that the RNA-binding protein HuR couples pre-mRNA processing and mRNA stability, Mol. Cell. 43 (2011) 327–339. https://doi.org/10.1016/j.molcel.2011.06.007.

[18] C.-Y.A. Chen, N. Xu, A.-B. Shyu, Highly selective actions of HuR in antagonizing AU-rich element-mediated mRNA destabilization, Mol. Cell. Biol. 22 (2002) 7268–7278.

[19] A. Kalsotra, X. Xiao, A.J. Ward, J.C. Castle, J.M. Johnson, C.B. Burge, T.A. Cooper, A postnatal switch of CELF and MBNL proteins reprograms alternative splicing in the developing heart, Proc. Natl. Acad. Sci. U. S. A. 105 (2008) 20333–20338. https://doi.org/10.1073/pnas.0809045105.

ACCEPTED MANUSCRIPT / CLEAN COPY

[20] A. V. Philips, L.T. Timchenko, T.A. Cooper, Disruption of splicing regulated by a CUG-binding protein in myotonic dystrophy, Science. 280 (1998) 737–741.

[21] S.-H. Chang, O. Elemento, J. Zhang, Z.W. Zhuang, M. Simons, T. Hla, ELAVL1 regulates alternative splicing of eIF4E transporter to promote postnatal angiogenesis, Proc. Natl. Acad. Sci. U. S. A. 111 (2014) 18309–18314. https://doi.org/10.1073/pnas.1412172111.

[22] J.A. Villefranc, J. Amigo, N.D. Lawson, Gateway compatible vectors for analysis of gene function in the zebrafish, Dev. Dyn. 236 (2007) 3077–3087. https://doi.org/10.1002/DVDY.21354.

[23] A. Leray, S. Padilla-Parra, J. Roul, L. Heliot, M. Tramier, Spatio-Temporal Quantification of FRET in living cells by fast time-domain FLIM: a comparative study of non-fitting methods [corrected], PLoS One. 8 (2013) e69335. https://doi.org/10.1371/journal.pone.0069335.

[24] M. Ryan, W.C. Wong, R. Brown, R. Akbani, X. Su, B. Broom, J. Melott, J. Weinstein, TCGASpliceSeq a compendium of alternative mRNA splicing in cancer, Nucleic Acids Res. 44 (2016) D1018-1022. https://doi.org/10.1093/nar/gkv1288.

[25] O. Le Tonqueze, B. Gschloessl, A. Namanda-Vanderbeken, V. Legagneux, L. Paillard, Y. Audic, Chromosome wide analysis of CUGBP1 binding sites identifies the tetraspanin CD9 mRNA as a target for CUGBP1-mediated down-regulation, Biochem. Biophys. Res. Commun. 394 (2010) 884–889. https://doi.org/10.1016/j.bbrc.2010.03.020.

[26] S. Padilla-Parra, M. Tramier, FRET microscopy in the living cell: different approaches, strengths and weaknesses, BioEssays News Rev. Mol. Cell. Dev. Biol. 34 (2012) 369–376. https://doi.org/10.1002/bies.201100086.

[27] B. Cosson, C. Gautier-Courteille, D. Maniey, O. Ait-Ahmed, M. Lesimple, H.B. Osborne, L. Paillard, Oligomerization of EDEN-BP is required for specific mRNA deadenylation and binding, Biol. Cell. 98 (2006) 653–665. https://doi.org/10.1042/BC20060054.

[28] O. Le Tonqueze, B. Gschloessl, V. Legagneux, L. Paillard, Y. Audic, Identification of CELF1 RNA targets by CLIP-seq in human HeLa cells, Genomics Data. 8 (2016) 97–103. https://doi.org/10.1016/j.gdata.2016.04.009.

[29] P.J. Uren, S.C. Burns, J. Ruan, K.K. Singh, A.D. Smith, L.O.F. Penalva, Genomic analyses of the RNA-binding protein Hu antigen R (HuR) identify a complex network of target genes and novel characteristics of its binding sites, J. Biol. Chem. 286 (2011) 37063–37066. https://doi.org/10.1074/jbc.C111.266882.

[30] R. Dong, P. Chen, K. Polireddy, X. Wu, T. Wang, R. Ramesh, D.A. Dixon, L. Xu, J. Aube, Q. Chen, An RNA-Binding Protein, Hu-antigen R, in Pancreatic Cancer Epithelial to Mesenchymal Transition, Metastasis, and Cancer Stem Cells, Mol. Cancer Ther. 19 (2020) 2267–2277. https://doi.org/10.1158/1535-7163.MCT-19-0822.

[31] R. Dong, J.G. Lu, Q. Wang, X.L. He, Y.K. Chu, Q.J. Ma, Stabilization of Snail by HuR in the process of hydrogen peroxide induced cell migration, Biochem. Biophys. Res. Commun. 356

ACCEPTED MANUSCRIPT / CLEAN COPY

(2007) 318–321. https://doi.org/10.1016/J.BBRC.2007.02.145.

[32] J. Sun, X. Gu, N. Wu, P. Zhang, Y. Liu, S. Jiang, Human antigen R enhances the epithelial-mesenchymal transition via regulation of ZEB-1 in the human airway epithelium, Respir. Res. 19 (2018). https://doi.org/10.1186/S12931-018-0805-0.

[33] J. Pu, X. Zhang, H. Luo, L. Xu, X. Lu, J. Lu, Adrenaline promotes epithelial-to-mesenchymal transition via HuR-TGFβ regulatory axis in pancreatic cancer cells and the implication in cancer prognosis, Biochem. Biophys. Res. Commun. 493 (2017) 1273–1279. https://doi.org/10.1016/J.BBRC.2017.09.146.

ACCEPTED MANUSCRIPT / CLEAN COPY

FIGURE LEGENDS

Figure 1. CELF1 and ELAVL1 directly interact in the nucleus.

A, Representative images of HeLa cells transfected by a plasmid driving the expression of EGFP-CELF1 (left), or co-transfected by the same plasmid and a plasmid that drives the expression of mCherry-histone 2B, CELF1, or ELAVL1. The EGFP lifetime channel is shown. **B**, Box plot of the distribution of EGFP fluorescence lifetimes in all measured pixels in 30 nuclei. The difference between the control (no mCherry plasmid) and the test (CELF1-mCherry plasmid) conditions is highly significant ($p < 2.2*10^{-16}$, Kruskal-Wallis test). **C**, Same as **B** but EGFP is fused to ELAVL1.

Figure 2. RT-PCR validation of microarray results.

A, In *PHACTR2*, a cassette exon is more repressed by CELF1 than by ELAVL1 in microarray experiments. *Middle panel,* RT-PCR with primers located in the flanking exons (arrows in *upper panel*) with RNA extracted from cells previously treated with the indicated siRNAs: NT, not treated; Ctrl, same control siRNA as in microarray experiments; luc, luciferase; C1, C2, C3, *CELF1* siRNA 1 to 3; C123, mixture of 3 *CELF1* siRNAs*;* E1, E2, *ELAVL1* siRNA 1 and 2; E12, mixture of 2 *ELAVL1* siRNAs; CE, mixture of 3 *CELF1* and 2 *ELAVL1* siRNAs. *Lower panels,* percentages of cassette exon-containing isoform (percent spliced in) and mean values from 3 independent experiments. *p*, Dunnett's test p-value comparing the NT with the other conditions (ns, not significant, $p > 0.05$). **B**, Same as **A,** but for *WNK1*, which contains a cassette exon more strongly repressed by ELAVL1 than by CELF1.

Figure 3. CELF1 and ELAVL1 control together the inclusion of exons v7 to v10 of *CD44* mRNA in HeLa cells.

A, Number of exons that are either repressed (positive splicing index after depletion) or stimulated by CELF1 or ELAVL1. **B**, Genomic regions of *CD44* between exons v6 and 15. The splicing

ACCEPTED MANUSCRIPT / CLEAN COPY

indices in cells treated by siRNAs against *CELF1* (upper), *ELAVL1* (middle) or both (lower) are indicated as green or red bars. The data for CLIPseq of CELF1 in HeLa cells [28] are shown in blue. **C**, Upper left, schematic showing the position of the RT-PCR primers. Upper right, one representative RT-PCR experiment from HeLa cells transfected with the indicated siRNA. Ctrl, same control siRNA as in microarray experiments; luc, luciferase; CELF1, mixture of 3 *CELF1* siRNAs; ELAVL1, mixture of 2 *ELAVL1* siRNAs; C+E, mixture of 3 *CELF1* and 2 *ELAVL1* siRNAs. Bottom, quantification of 3 independent experiments. The percentages of inclusion of exon v7, v8 and v9, and v10, were calculated from bands 1, 3, and 6, bands 1 and 4, and bands 1, 2, and 5, respectively. Asterisks indicate statistical significance following Dunnett's test comparing the Ctrl siRNA and the other conditions (exact *p*-values, v7 0.025, 0.014 and $1.2*10^{-3}$; v8 and v9, 0.01, $5.5*10^{-3}$ and $5.9*10^{-4}$; v10, $9.2*10^{-3}$, $4.0*10^{-3}$ and $4.2*10^{-4}$).

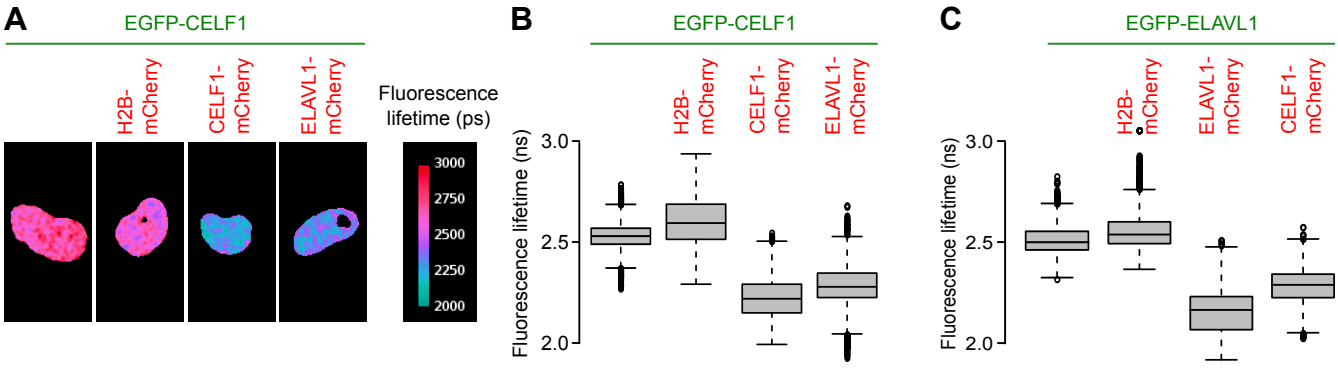
Figure 4. The inclusion of *CD44* variable exons is elevated in cancer patients with high levels of *CELF1* and *ELAVL1*.

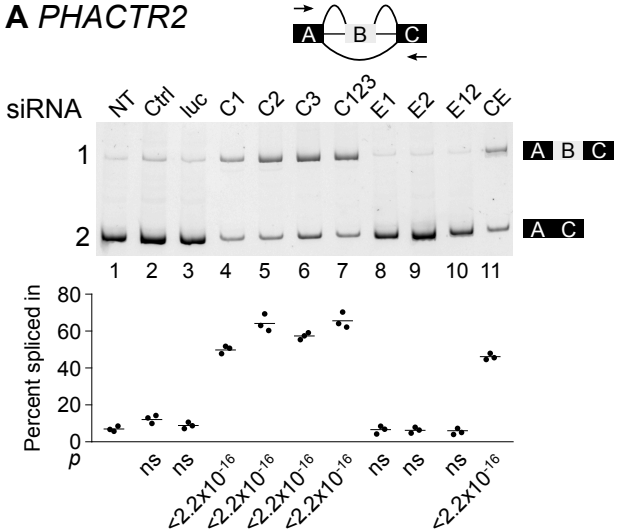
A, *Lower panel*, Volcano plot showing the statistical significance (*p*-value of a Wilcoxon test, -log10 scale) of the difference between the percentage of exon inclusion (PSI, percent spliced in) in high expression groups compared to the corresponding low expression groups. Each dot is one alternative splicing event (group of exons) pertaining to *CD44* variable exons in one TCGA cohort. *Upper panel*, density plot showing the distribution of the differences of exon inclusion for events with *p*-values below 10^{-2}. **B**, Table summarizing the TCGA cohorts in which the indicated variable exons are significantly ($p < 10^{-2}$) more included in patients with high levels of the indicated transcripts (color code as in **A**). **C**, *Left panel*, we separated the patients in the TCGA breast cancer (BRCA) panel (n = 1079) into two groups of the same size based on the level of *CELF1* mRNA. For each of the two groups, we show the distribution of the percent of inclusion of variable exons v8-v10. The *p*-value of a Wilcoxon test is shown. *Middle panel*, same as left, but grouped based on *ELAVL1* levels. *Right panel*, same as left, but grouped based on both *CELF1* and *ELAVL1*

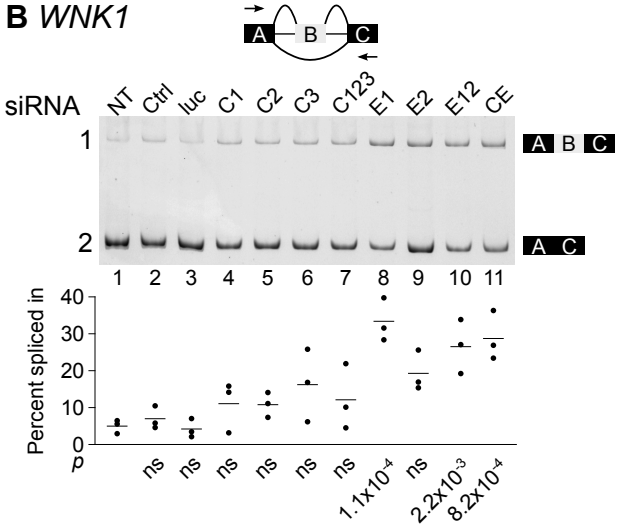
ACCEPTED MANUSCRIPT / CLEAN COPY

expression. High group, n = 255; Low group, n = 254.

ACCEPTED MANUSCRIPT / CLEAN COPY


A
EGFP-CELF1
H2B-mCherry
CELF1-mCherry
ELAVL1-mCherry
Fluorescence lifetime (ps)
3000
2750
2500
2250
2000
B
EGFP-CELF1
H2B-mCherry
CELF1-mCherry
ELAVL1-mCherry
Fluorescence lifetime (ns)
3.0
2.5
2.0
C
EGFP-ELAVL1
H2B-mCherry
ELAVL1-mCherry
CELF1-mCherry
Fluorescence lifetime (ns)
3.0
2.5
2.0

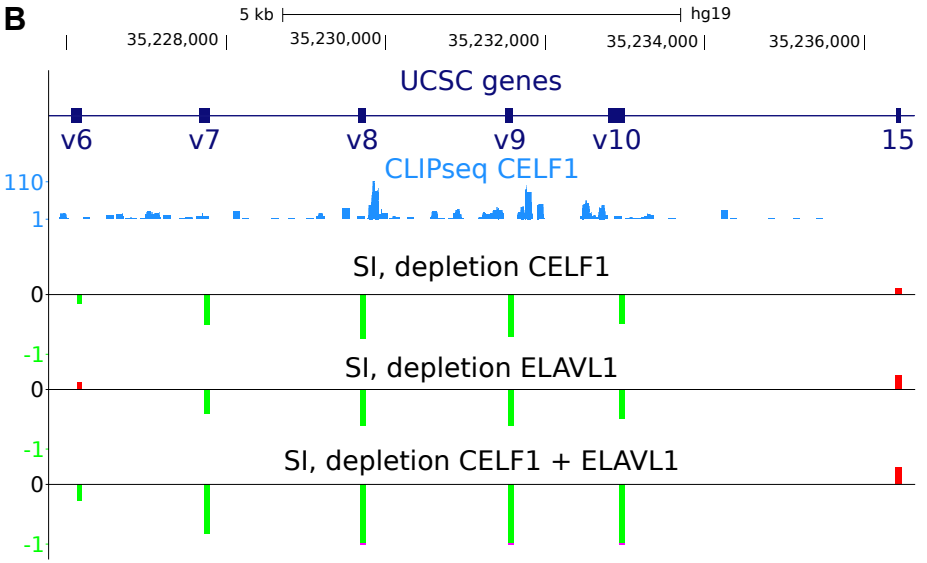

A PHACTR2
A
B
C
siRNA
NT
Ctrl
luc
C1
C2
C3
C123
E1
E2
E12
CE
1
2
A B C
A C
1 2 3 4 5 6 7 8 9 10 11
Percent spliced in
80
60
40
20
0
p
ns
ns
<2.2x10-16
<2.2x10-16
<2.2x10-16
<2.2x10-16
ns
ns
ns
<2.2x10-16


B WNK1
A
B
C
siRNA
NT
Ctrl
luc
C1
C2
C3
C123
E1
E2
E12
CE
1
2
A B C
A C
1 2 3 4 5 6 7 8 9 10 11
Percent spliced in
40
30
20
10
0
p
ns
ns
ns
ns
ns
ns
1.1x10-4
ns
2.2x10-3
8.2x10-4

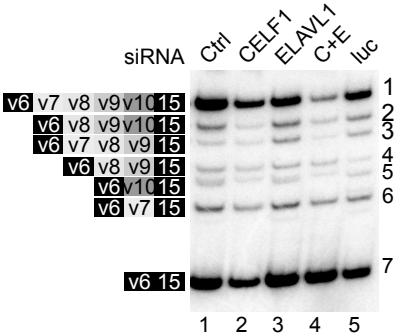
A

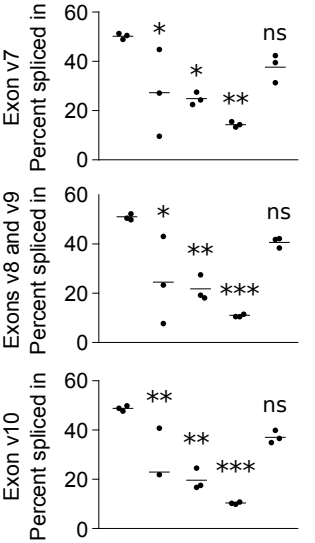
	ELAVL1-repressed	ELAVL1-stimulated
CELF1-repressed	42	3
CELF1-stimulated	1	16

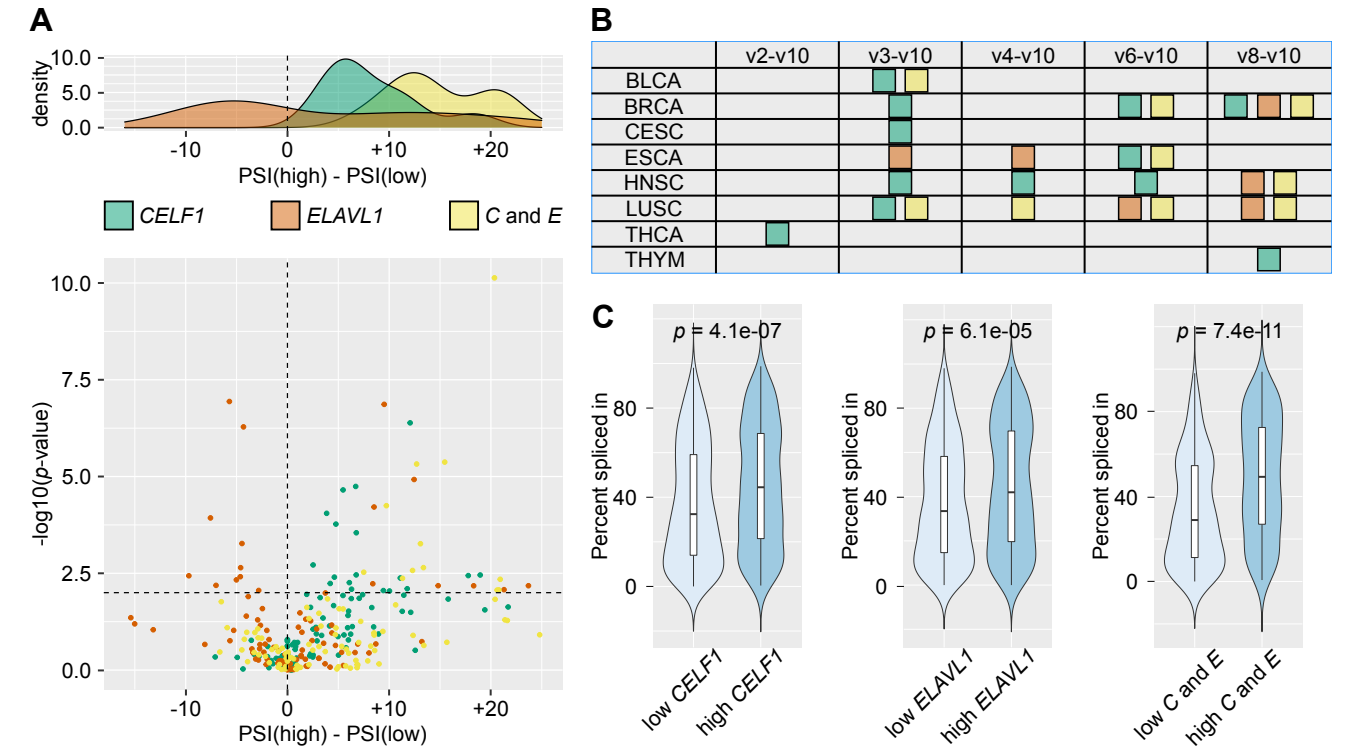
B



C






A
density
10.0
5.0
0.0
-10
0
+10
+20
PSI(high) - PSI(low)
CELF1
ELAVL1
C and E
-log10(p-value)
10.0
7.5
5.0
2.5
0.0
-10
0
+10
+20
PSI(high) - PSI(low)
B
v2-v10
v3-v10
v4-v10
v6-v10
v8-v10
BLCA
BRCA
CESC
ESCA
HNSC
LUSC
THCA
THYM
C
p = 4.1e-07
p = 6.1e-05
p = 7.4e-11
Percent spliced in
80
40
0
low CELF1
high CELF1
low ELAVL1
high ELAVL1
low C and E
high C and E