



OPEN

Identification of an mRNA isoform switch for HNRNPA1 in breast cancers

Murat Erdem^{1,5}, İbrahim Ozgul^{1,5}, Didem Naz Dioken¹, Irmak Gurcuoglu¹, Sezen Guntekin Ergun^{2,4}, Rengul Cetin-Atalay², Tolga Can^{2,3} & Ayse Elif Erson-Bensan^{1,2✉}

Roles of HNRNPA1 are beginning to emerge in cancers; however, mechanisms causing deregulation of HNRNPA1 function remain elusive. Here, we describe an isoform switch between the 3'-UTR isoforms of *HNRNPA1* in breast cancers. We show that the dominantly expressed isoform in mammary tissue has a short half-life. In breast cancers, this isoform is downregulated in favor of a stable isoform. The stable isoform is expressed more in breast cancers, and more HNRNPA1 protein is synthesized from this isoform. High HNRNPA1 protein levels correlate with poor survival in patients. In support of this, silencing of HNRNPA1 causes a reversal in neoplastic phenotypes, including proliferation, clonogenic potential, migration, and invasion. In addition, silencing of HNRNPA1 results in the downregulation of microRNAs that map to intragenic regions. Among these miRNAs, miR-21 is known for its transcriptional upregulation in breast and numerous other cancers. Altogether, the cancer-specific isoform switch we describe here for HNRNPA1 emphasizes the need to study gene expression at the isoform level in cancers to identify novel cases of oncogene activation.

Advances in RNA-sequencing (RNA-seq) methods revealed more than 90% of human genes to produce tissue-specific alternative mRNA isoforms, which add substantial complexity to the transcriptome in higher vertebrates^{1–3}. Potential mechanisms that cause and coordinate this isoform diversity are; alternative use of transcription start sites, alternative splicing, and alternative polyadenylation^{4,5}. The resulting mRNA isoforms may or may not share the same 5'UTRs (untranslated regions), coding sequences (CDSs), or 3'UTRs^{6,7}. As a result, differences in the CDSs and UTRs may alter mRNA stability, cellular localization, and translation rate, leading to changes in protein functions or levels³. Hence, the type of isoforms and their relative expression levels become important variables for gene expression regulation in normal and disease states. As we begin to appreciate the depth and extent of isoform diversity in the transcriptome, new findings point out changes in relative ratios of mRNA isoforms, referred to as isoform switches. Accumulating evidence highlights the role of isoform switches with biological impact in normal tissues and diseases, including cancer^{8–10}. For example, a shorter isoform of *BCL*, *BCL-XS*, activates apoptosis and is a tumor suppressor, whereas the longer isoform, *BCL-XL*, is an oncogene blocking apoptosis¹¹. The ratio of these two oppositely functioning mRNA isoforms is altered in cancers in favor of the longer isoform, enhancing survival¹². Similarly, expression of a *BRAF* (B-Raf proto-oncogene) isoform that lacks the RAS binding domain promotes resistance to *BRAF* inhibition in a group of melanoma patients^{4,13}.

High throughput efforts provide further evidence that pathologic shift of isoform ratios may disrupt protein-protein interactions in different cancers¹⁴. Hence, the discovery of cancer-specific isoform switches holds promising potential as diagnostic biomarkers and therapy targets^{9,15}.

This study describes a cancer-specific isoform switch for a versatile RNA binding protein (RBP), HNRNPA1. We took a combinatorial in silico and in vitro approach to identify and verify the isoform switch between 3'-end isoforms of *HNRNPA1*. We show that *HNRNPA1* isoforms with different 3'UTRs have different half-lives and the cancer-specific isoform switch contributes to increased HNRNPA1 protein abundance in breast cancer cells. We also provide new insight into HNRNPA1 function in indirectly modulating the expression of intragenic microRNAs such as miR-21, a well-known oncomiR. These results emphasize the importance of gene expression analysis at the isoform level in cancer cells, revealing unknown oncogene activation cases with biological impact.

¹Department of Biological Sciences, Middle East Technical University (METU), Dumlupinar Blv No: 1 Üniversiteler Mah., Cankaya, Ankara 06800, Turkey. ²Cancer Systems Biology Laboratory, CanSyL, Graduate School of Informatics, Middle East Technical University, 06800 Ankara, Turkey. ³Department of Computer Engineering, Middle East Technical University (METU), Dumlupinar Blv No: 1 Üniversiteler Mah, Ankara 06800, Turkey. ⁴Present address: Department of Medical Biology, Hacettepe University, Ankara, Turkey. ⁵These authors contributed equally: Murat Erdem and İbrahim Ozgul. ✉email: erson@metu.edu.tr

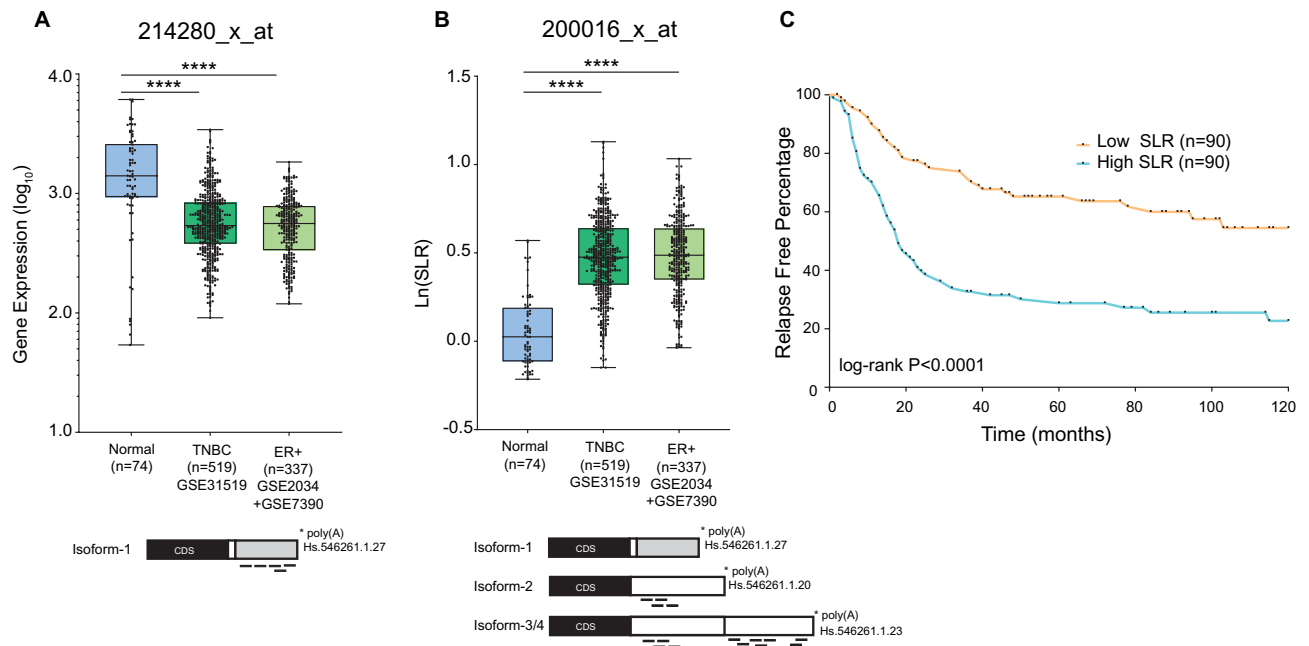


Figure 1. Isoform level expression of *HNRNPA1*. **(A)** *HNRNPA1* Isoform-1 expression in breast cancer patients compared with normal breast tissue (**** $p < 0.0001$, unpaired t-test) in GSE31519, GSE2034, and GSE7390 datasets. Positions of Affymetrix distal probes (214280_x_at) recognizing only Isoform-1 are shown. **(B)** SLR values (Isoform 2 + 3 + 4/Isoform 3 + 4) of *HNRNPA1* in breast cancer patients compared with normal breast tissue (**** $p < 0.0001$, unpaired t-test) in GSE31519, GSE2034, and GSE7390 datasets. Positions of Affymetrix distal probes (200016_x_at) recognizing Isoform-2,3,4 are shown. **(C)** High expression (of Isoform-2, 3, and 4) correlates with poor relapse-free survival in TNBC patients. GSE31519 dataset was grouped according to the top 25% (High) and bottom 25% (Low) expressing patients. Expression values are based on proximal probes of 200016_x_at, excluding levels of Isoform-1. Hazard ratio (95% CI) is 2.709 (log-rank $p < 0.0001$).

Results

Isoform level analysis. In a targeted screen for *HNRNPA1* expression in breast cancers, we re-analyzed GSE31519, GSE2034, GSE7390 datasets using APADetect, an algorithm to detect isoform level expression differences based on differential poly(A) site usage^{16,17}. We analyzed data sets for probe-level differences based on the positions of poly(A) sites. Probes proximal to poly(A) sites generally recognize all isoforms, whereas distal probes recognize longer isoforms. Ratios of proximal to distal probe sets were calculated for normal and cancer samples. Significant changes in the signal intensities were reported as SLR ((Short + Long)/Long ratio). In breast cancer patients ($n = 856$), independent from tumor type, we observed significant downregulation ($p < 0.0001$) of an *HNRNPA1* isoform that ends with a distal poly(A) site on the gene locus (Hs.546261.1.27) (Fig. 1A). The downregulated isoform (hereon called Isoform-1) has a different 3'UTR than other isoforms of *HNRNPA1* due to the inclusion of two non-coding terminal exons (exon 12, 13). Only the distal probes of 214280_x_at (Affymetrix probe set ID) recognize Isoform-1 specifically.

On the contrary, the SLR of other isoforms was high in patients compared to normal breast tissue (Fig. 1B). The Isoform-2, 3, 4 were co-detected (but could not be distinguished) and quantified by the distal probes of the 200016_x_at probe set as these isoforms have the same terminal exon (exon 11). These isoforms either have a shorter 3'UTR (Isoform-2) or a longer 3'UTR (Isoform-3, 4). Overall, breast cancer patients have an increased isoform ratio (Short + Long isoforms)/Long isoforms) for *HNRNPA1*. The ratio shift was significant; however, we could not determine individual expression levels because both proximal and distal probes recognize multiple isoforms. Interestingly, increased expression of isoforms detected by proximal probes of 200016_x_at correlates with patient relapse times in the GSE31519 cohort (Fig. 1C). These data suggested that isoforms are differentially expressed in patients.

To begin confirming the in silico patient data, we first validated the 3'-ends of isoforms by 3'-RACE, cloning, and sequencing (Supplementary Figs. S1–S3). Next, we tested breast cancer cell lines ($n = 18$) and a panel of breast cancer patient cDNAs ($n = 25$) by RT-qPCR (Supplementary Fig. S4). We detected an increased ratio of isoforms compared to Isoform-1 in 40% of breast cancer cell lines and approximately 90% of patient samples. However, the culprit of microarray data and RT-qPCR was the use of probes and primers recognizing more than one isoform. To delineate isoform-specific expression, we turned to RNA-seq data of Genotype-Tissue Expression (GTEx) and TCGA datasets. We compared the expression of isoforms in GTEx normal tissue samples to TCGA tumor samples using UCSC Xena, which allows comparison of the two datasets^{18–20}.

Surprisingly, Isoform-1, which is low in breast cancer patients (Fig. 1A), is the dominant isoform in normal mammary tissue (ENST00000547566.1) (Fig. 2). Isoform-2 (ENS ENST00000330752.12) has a short 3'UTR, and it is overexpressed in Luminal A, Luminal B, HER2 enriched, basal-like, and normal-like breast

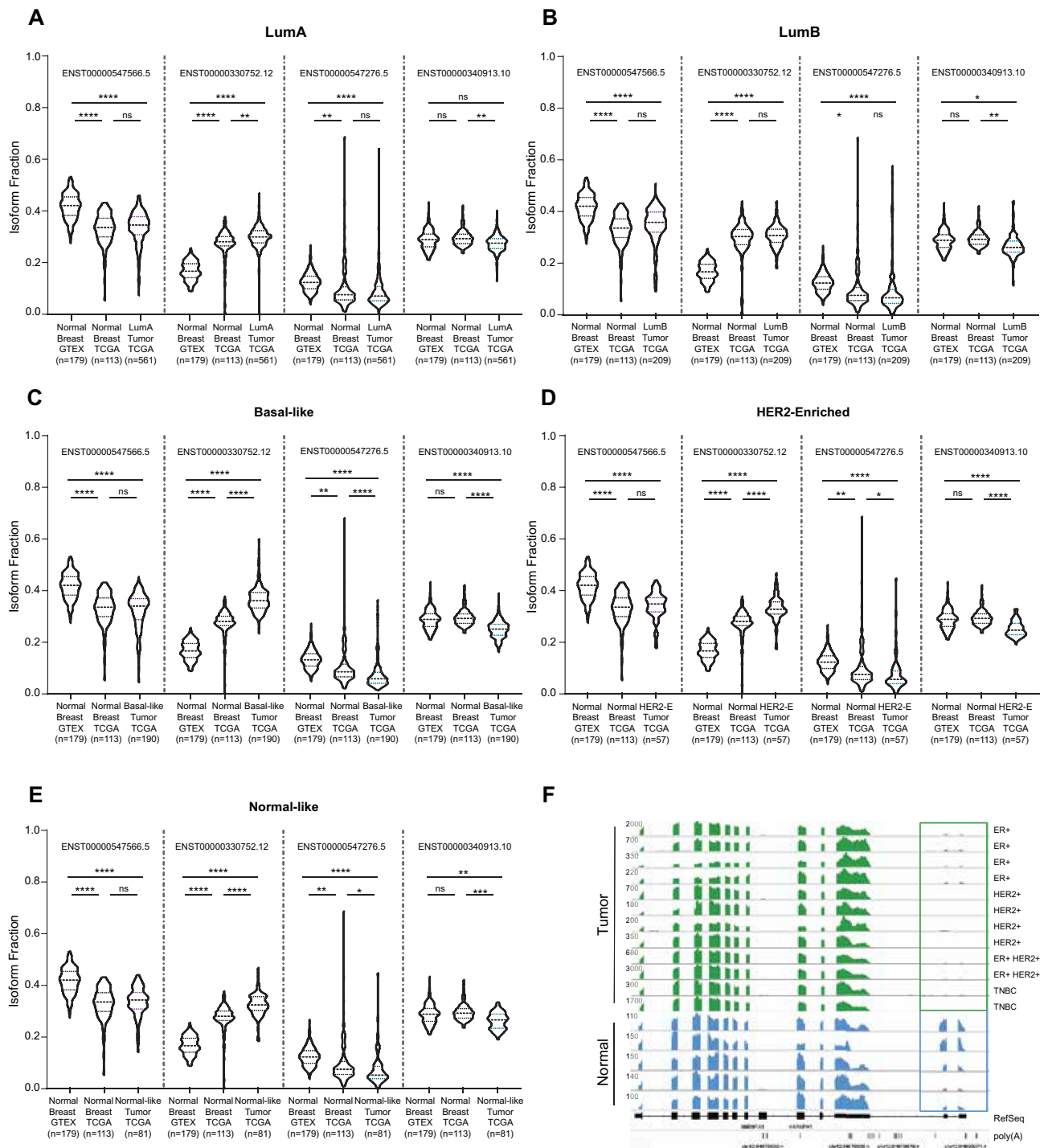


Figure 2. Isoform fractions of *HNRNPA1* isoforms (Isoform-1, 2, 3, 4) in normal breast (GTEx), normal adjacent tissue (TCGA), and in Luminal A (A), Luminal B (B), Basal-like (C), HER2-enriched (D), Normal-like breast cancers (E) in the TCGA dataset (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$, one-way ANOVA, Tukey's HSD), (F) Representative single-cell RNA-seq data for normal cell types and in breast cancer cells (GSE75688, GSE113197). The boxed area shows the reads from the exons unique to Isoform-1.

cancers compared to GTEx normal mammary tissue. Expression of Isoform-3 (ENST00000547276.5) is further decreased in all breast cancer subtypes, despite not being very abundant in normal breast tissue. Isoform-4 (ENST00000340913.10) is also downregulated in all subtypes compared to adjacent normal or GTEx normal tissue. Isoform-4 also has a long 3'UTR identical to Isoform-3 but has an exon-8 insertion (156 bp) which does not alter the reading frame.

Notably, the expression levels of some isoforms in the matched peritumor tissue of breast cancer patients were in an intermediate state compared with the normal tissues in the GTEx dataset. This pattern suggests that the

adjacent normal tissue may be distinct from healthy tissue and the tumor, as was also reported in other studies²¹. The isoform switch was also evident in single-cell RNA-sequencing data of normal mammary cells compared with breast cancer cells (GSE113197, GSE75688) (Fig. 2F).

With these results, the reason behind SLR change in breast cancers (Fig. 1B) became apparent. Higher SLR was due to the upregulation of Isoform-2 and downregulation of other isoforms.

Isoform switch and HNRNPA1 protein levels. At this point, we wanted to investigate the functional consequences of the isoform switch. However, we were surprised to find out that Isoform-1 appears as a non-coding transcript (NR_135167). The coding sequence of HNRNPA1 ends in exon 10, and all isoforms share the same stop site despite having different terminal exons (Supplementary Fig. S5). Hence to find out whether Isoform-1 is coding for a peptide, we first calculated the coding potential using the Coding Potential Calculator 2 (CPC2) algorithm²² and saw that it was similar to other isoforms (Fig. 3A). To verify the coding potential experimentally, we performed ribosomal affinity purification (TRAP) of translated mRNAs, performed RT-qPCR, and normalized the polysome-bound transcript levels to no-TRAP control cells. These results showed that all isoforms were associated with the immunoprecipitated polysomes. *XIST* non-coding RNA was used as a negative control (Fig. 3B).

Next, we cloned the coding sequence of *HNRNPA1* along with the different 3'UTRs of the isoforms. Isoform-1 has a unique 3'UTR (Iso-1-3'UTR). Other isoforms share the same terminal exons, but Isoform-2 has a shorter 3'UTR (S-3'UTR) compared to the long 3'UTRs (L-3'UTR) of Isoform-3 and Isoform-4. Therefore, we tested whether these three types of 3'UTRs had different effects on protein levels. We transfected MCF7 and MDA-MB-231 cells with the HA-tagged HNRNPA1 protein expression constructs. Protein expression was detected by western blotting. Of interest, the level of HNRNPA1 protein encoded by the construct with Iso-1-3'UTR was markedly lower than the other isoforms (Fig. 3C). Because transfection efficiency could be a reason for this observation, we cloned the three different 3'UTRs downstream of a reporter gene and transiently transfected cells for a dual luciferase assay where transfection efficiencies were normalized. Here too, the luciferase reporters for the S-3'UTR and L-3'UTR had significantly higher activities than Iso-1-3'UTR in both cells (Fig. 3D). Results from the 3'UTR-reporter system, along with forced expression of HA-tagged proteins, suggested that expression of Isoform-1 correlated with lower protein levels. Since we did not see a difference in ribosome association of isoforms (Fig. 3A), we tested whether mRNA half-lives could affect HNRNPA1 protein abundance. We tested mRNA levels of *HNRNPA1* isoforms following actinomycin D treatment for 12 h to prevent new transcription. RT-qPCR results showed that Isoform-1 had a short half-life, comparable to *MYC* mRNA, well-known for its short half-life²³ (Fig. 4A). In contrast, other isoforms were still stable after 12 h when we finalized the experiment in MCF7 and MDA-MB-231 cells. Similar decay rates were determined in MCF10A cells (non-tumorigenic mammary epithelial) (Supplementary Fig. S6).

We also treated cells with actinomycin D and cycloheximide, an inhibitor of ribosomal elongation²⁴. Interestingly, cycloheximide treatment for only 3 h had a dramatic recovery effect only for Isoform 1 (> 3.5 fold in MCF7, > 7.5 fold in MDA-MB-231) (Fig. 4B, Supplementary Fig. S7). This quick recovery suggests that Isoform-1 is co-translationally degraded, as cycloheximide is also known to inhibit mRNA decay²⁵.

These results indicated that the isoform switch results in differential expression of isoforms with different mRNA stabilities, affecting protein levels. However, we also tested whether 3'UTRs may regulate the localization of HNRNPA1 protein, as was suggested for a few interesting cases^{26,27}. In this case, the nuclear localization of HNRNPA1 was independent of 3'UTR sequences of isoforms (Supplementary Fig. S8).

Overall, these results showed that Isoform-1 is a rapidly degraded mRNA, possibly better regulating the protein level of HNRNPA1. This unstable isoform is low in breast cancers, whereas Isoform-2 is upregulated. Because this switch would indicate upregulation of HNRNPA1 protein, we were curious to investigate HNRNPA1 protein levels in patient samples. Hence, we took advantage of a quantitative liquid chromatography/mass spectrometry-based proteome analysis dataset, which used protein extracts from breast tumors and adjacent non-cancerous tissues²⁸. In this dataset, HNRNPA1 protein was significantly high in 52 tumors compared to normal tissues and 13 basal-like tumors compared with normal tissue (Fig. 5A). High HNRNPA1 protein levels in these patients correlated with decreased survival, strengthening the significance of the oncogenic role of HNRNPA1 (Fig. 5B). Moreover, in an independent dataset of Clinical Proteomic Tumor Analysis Consortium (CPTAC), HNRNPA1 protein was also overexpressed in luminal, HER2+, and TNBC tumors (Fig. 5C). Of note, post-translational modifications and protein-protein interactions are likely to introduce additional layers of regulation to HNRNPA1 activity in cells.

HNRNPA1 silenced models and intragenic miRNAs. Next, to begin addressing the biological relevance of HNRNPA1 overexpression in breast cancers, we generated stable shRNA constructs to target *HNRNPA1* expression in MCF7 and MDA-MB-231 cells (Fig. 5D) and tested these models for changes in their neoplastic phenotypes. HNRNPA1 is a versatile RNA-binding protein involved in many aspects of RNA biology, so we found a significant reversal of neoplastic phenotypes in both silencing models. We observed loss of clonogenicity (Fig. 5E), decreased proliferation (Fig. 5F), decreased motility (Fig. 5G), decreased migration and invasion capability (Fig. 5H) upon sustained silencing of HNRNPA1.

Next, to shed light on the possible effects of HNRNPA1 activity in breast cancers, we turned to microRNAs (miRNAs) as a less explored aspect of HNRNPA1 function. HNRNPA1 has been implicated in promoting or hindering the processing steps of pri-miR-18a and pri-let-7a-1 by direct binding to loop regions^{29,30}. Because HNRNPA1 has fundamental roles in RNA biology, we wanted to test whether other miRNAs would be affected by HNRNPA1 silencing. We used *HNRNPA1* silenced MCF7 and MDA-MB-231 cells (Fig. 6A,D) to screen approximately 800 miRNAs using the NanoString technology (nCounter Human miRNA assay). We detected

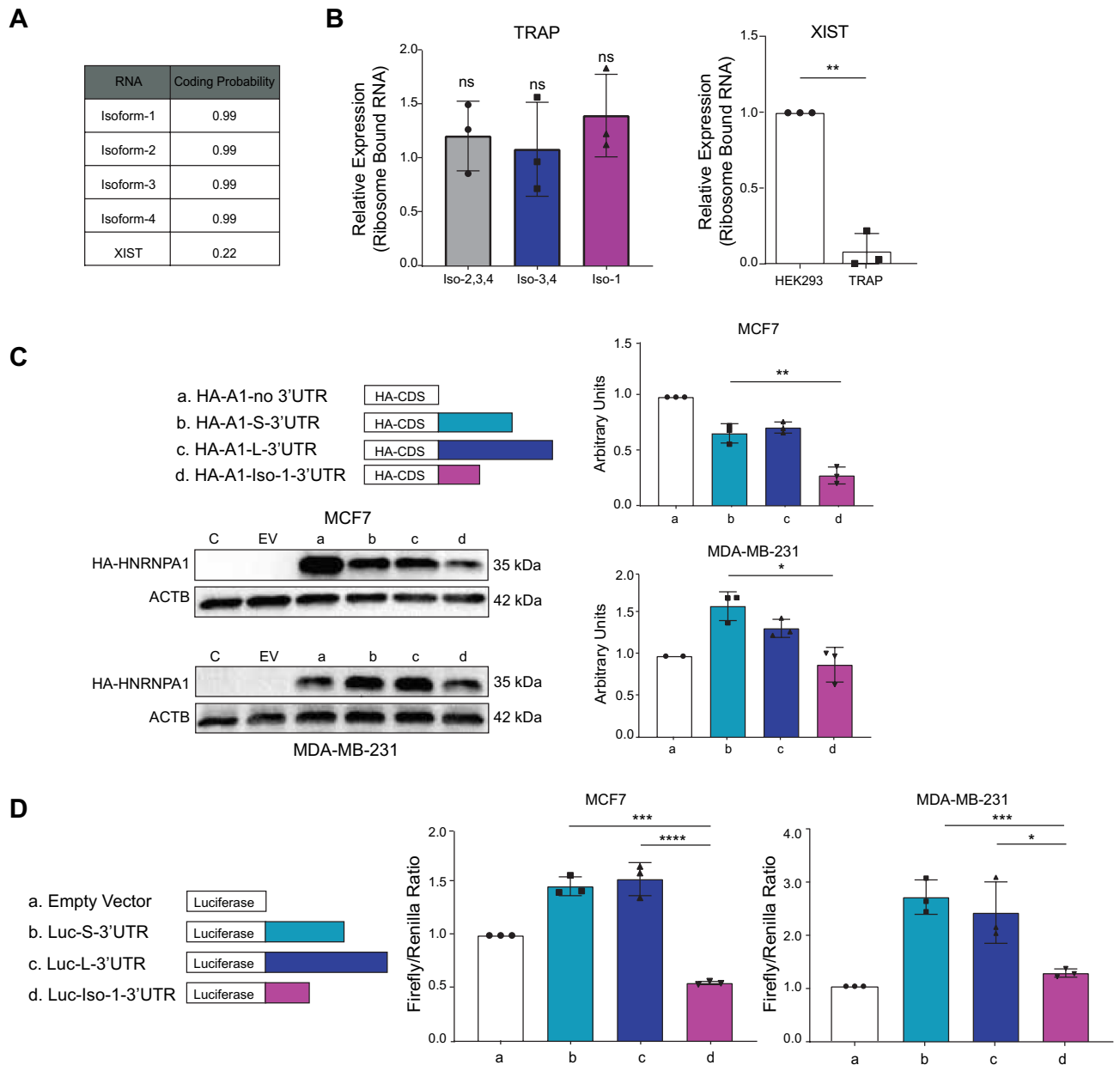


Figure 3. Translation of *HNRNPA1* isoforms. **(A)** Coding potentials of *HNRNPA1* isoforms determined by CPC2 are shown. *XIST* is a non-coding RNA. **(B)** TRAP in EGFP-L10A expressing HEK293 cells. Levels of TRAP RNAs were normalized to their expression values in untransfected HEK293 cells to eliminate the bias of expression levels (ns: not significant, $n = 3$ independent RNA-IPs, one-way ANOVA, Tukey's HSD). *XIST*, a non-coding RNA, was used as a negative control (** $p < 0.01$; $n = 3$, Student's t-test). **(C)** Cells were transiently transfected with indicated vectors, and lysates were collected. HA antibody was used to detect *HNRNPA1* levels. Same blots were hybridized with ACTB antibody. The image is representative of 3 independent experiments. Graphs show densitometric quantification of bands (* $p < 0.05$, ** $p < 0.01$, one-way ANOVA, Tukey's HSD), uncropped images are presented in Supplementary Fig. S13. **(D)** Different 3'UTRs were cloned downstream of the luciferase gene in the pMIR vector. Cells were transiently transfected, and *Firefly/Renilla* luciferase read-outs from the constructs were normalized to that of empty pMIR (* $p < 0.05$, ** $p < 0.01$, **** $p < 0.0001$; $n = 3$ independent transfections, one-way ANOVA, Tukey's HSD).

the downregulation of mature miRNAs in *HNRNPA1* silenced cells. We identified common or cell line-specific miRNAs downregulated upon *HNRNPA1* silencing (Supplementary Fig. S9). A total of 43 miRNAs (70% of significantly changed miRNAs) in MCF7-sh and 32 miRNAs (94% of significantly changed miRNAs) in MDA-MB-231-sh cells were downregulated compared to NT controls. Interestingly, these miRNAs were enriched for their predicted and confirmed mRNA targets in cancer-related pathways and signaling cascades (Fig. 6C,F) which may partly explain the phenotypic changes we detected in *HNRNPA1* silenced cancer cells.

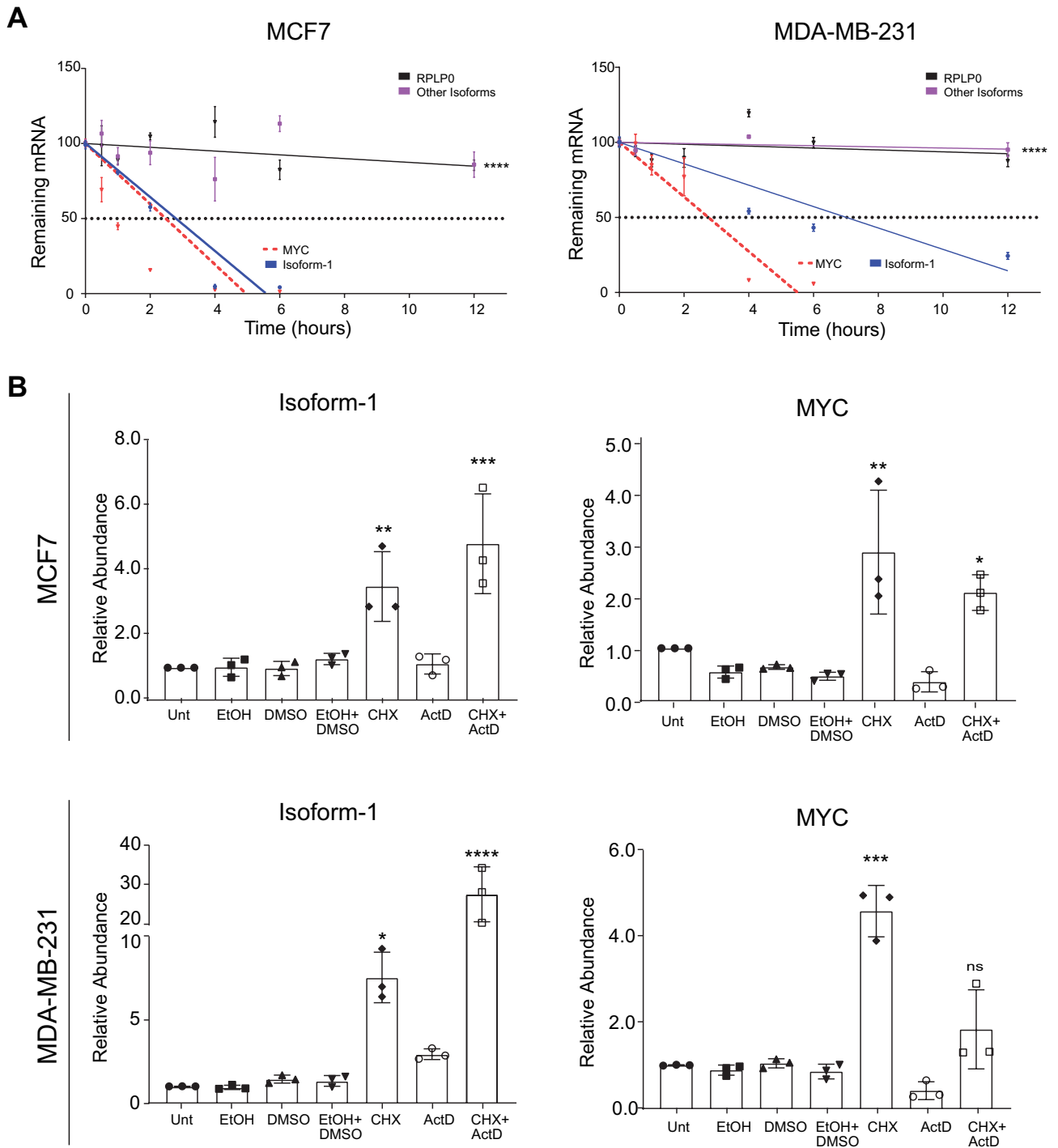


Figure 4. Stabilities of isoforms. **(A)** Remaining *HNRNPA1* transcripts in actinomycin D treated cells were detected by RT-qPCR. *RPLP0* as a stable mRNA and *MYC* as an unstable mRNA were used as controls (n = 3 independent treatments, ****p < 0.0001; n = 3, Student's t-test). **(B)** Cells were treated with actinomycin D and/or cycloheximide (CHX, 100 µg/mL) for 3 h to prevent transcription and translation. EtOH (Ethanol) and DMSO are carrier controls. Cells were collected and RNA was isolated for RT-qPCR (*p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001, ns: not significant, n = 3 independent experiments, student's t-test).

Notably, most downregulated miRNAs in both models were intragenic and mapped to introns of host genes (Fig. 6B,E). We reasoned that decreased expression of host genes might explain the downregulation of these miRNAs. Indeed, low miRNA read counts correlated with downregulated mature miR-27b-3p and pri-miR-27b-3p levels along with its host gene *C9ORF3* on 9q22.32 in MDA-MB-231 cells upon *HNRNPA1* silencing (Fig. 7A). These results suggested that the downregulation of miR-27b was due to decreased transcription of the host gene.

Next, miR-21 caught our attention as one of the most abundantly expressed and studied miRNAs in breast and other cancers^{31,32}. RT-qPCR verified low read counts for miR-21. Pri-miR21 levels were also low in MCF7 cells upon HNRNPA1 silencing (Fig. 7B). MiR-21 gene resides within the intron 11 of *VMP1* (Vacuole Membrane Protein-1) on 17q23.2. Hence, we tested whether *VMP1* was also downregulated in HNRNPA1 silenced cells to explain the mechanism behind decreased pri-miR-21 levels. However, there was only a minimal decrease in *VMP1* mRNA, which was unlikely to explain low levels of pri-miR-21 in HNRNPA1 silenced cells (Fig. 7B). However, *VMP1* mRNA is not the only source for miR-21 biogenesis; additional miR-21 promoters and primary transcripts have been characterized from within the terminal intronic regions of *VMP1*³³. To test whether the activity of this promoter region³⁴ was different in HNRNPA1 silenced cells, we cloned the well-defined promoter region for miR-21, a 433 bp region between –3770 to –3337 relative to the hairpin, into the pGL3-Basic promoter vector, driving *Firefly* luciferase expression. Transfection efficiency was monitored with pRL-TK driving the expression of the *Renilla* luciferase. Control cells (NT-sh) and HNRNPA1 silenced cells were transiently transfected with both vectors. We observed that the luciferase enzyme activity from the pGL3-miR-21 promoter was approximately 70% lower in HNRNPA1 silenced cells than in the control cells (NT) (Fig. 7C). These findings collectively show that miR-21 and pri-miR-21 levels were downregulated mainly due to the decreased activity of the miR-21 promoter in HNRNPA1 silenced cells.

To test the functionality of miR-21 downregulation on potential targets, we chose to generate a miR-21 sensor rather than testing known mRNA targets because HNRNPA1 loss is likely to alter levels/functions of many other coding and non-coding genes. Hence, we cloned two complementary binding sites for miR-21 downstream of *Firefly* luciferase CDS. We transfected control and HNRNPA1 silenced cells with this sensor. As a result, we detected higher luciferase activity from the miR-21 sensor in HNRNPA1 silenced cells due to less miR-21 binding to the 3'UTR of the luciferase mRNA (Fig. 7D). A mutant construct lacking the seed sequences of miR-21 had similar luciferase activities in control and HNRNPA1 silenced cells, showing the specificity of the sensor (Supplementary Fig. S10). miR-21 is upregulated in breast cancers, and this upregulation impacts the overall survival of ER+ breast cancers (Supplementary Fig. 10, Fig. 7E). While the effect of HNRNPA1 on miR-21 transcription is possibly indirect, an HNRNPA1 guided network may hold the potential to decipher transcriptional deregulation of miRNAs implicated in cancers.

Finally, we also took an independent approach and targeted the HNRNPA1 gene locus with CRISPR/Cas9. We confirmed decreased expression levels of miRNAs in HNRNPA1 deleted cells (Supplementary Fig. S11); however, the cells were not viable for continued culturing, showing HNRNPA1 dependency of cells. Indeed, most breast cancer cell lines have low “gene effect scores” indicating a high likelihood that HNRNPA1 is an essential gene in depletion assays (Supplementary Fig. S12). Our data and dependency scores collectively suggest that the HNRNPA1 function is critical and that cells cannot rescue its loss. Notably, because disease mutations have been reported for HNRNPA1 along with mutations in HNRNPA2B1³⁵, we still looked into expression patterns of the two transcripts. We found no significant correlation in more than a thousand breast cancer patient samples (Supplementary Fig. S12).

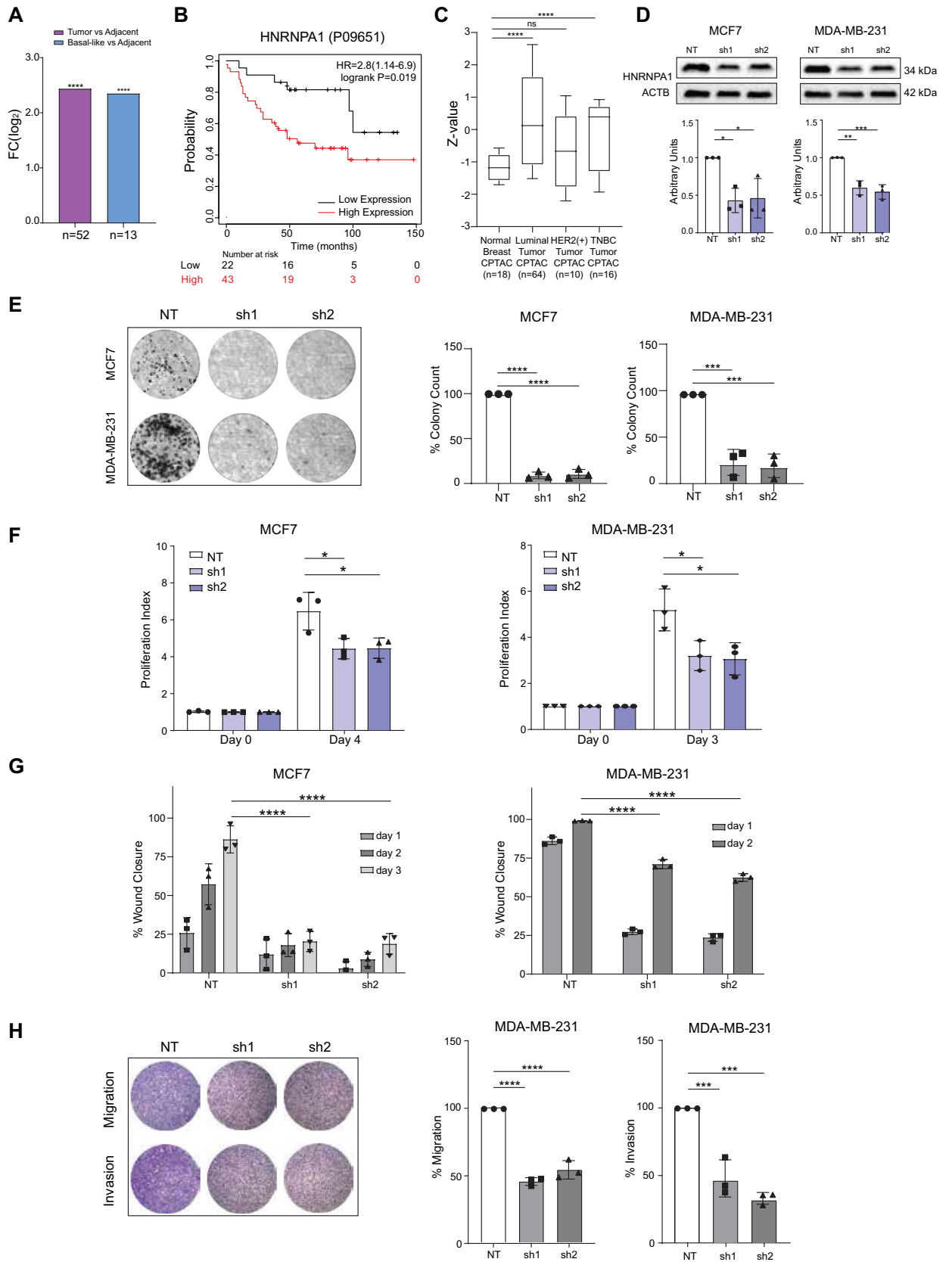
Overall, we report an isoform switch for HNRNPA1 and provide insight into the oncogenic roles of HNRNPA1 as a versatile RNA-binding protein whose expression is critical for the neoplastic phenotypes of breast cancer cells. Our findings, specifically on miR-21, may help understand how oncogenic miRNAs are frequently elevated in cancers.

Discussion

Mechanisms leading to alternative processing of mRNAs are gaining more attention as we begin to appreciate the complexities of cancer transcriptomes^{36,37}. Accordingly, widespread expression of alternatively spliced or polyadenylated isoforms has been described in cancers^{9,14,38,39}. As part of this complexity, cancer-specific isoform switches change the ratio of mRNA isoforms that may differ in their CDSs or 3'UTR sequences, consequently modulating protein functions in cancer cells. Hence, an increased appreciation of isoform switches may help the discovery of overlooked cancer-related genes and provide new avenues for diagnostic and therapeutic applications.

HNRNPA1 is a versatile protein involved in diverse aspects of RNA biology, including mRNA trafficking, telomere maintenance, regulation of mRNA stability, and splicing by antagonizing or enhancing other splicing proteins. HNRNPA1 can also bind to AU-rich elements and UAGGGA(U)-motifs in the 3'UTRs, and possesses RNA chaperone activity, promoting RNA–RNA interactions. In addition, HNRNPA1 has been implicated in transcriptional activation by binding to and destabilizing G-quadruplex structures within promoters^{40–42}. Hence deregulation of HNRNPA1 abundance may have diverse and indirect consequences.

Our work here demonstrates an isoform switch for HNRNPA1 in breast cancers. HNRNPA1 has four similar mRNA isoforms in normal breast tissue as described in the GTEx database. All isoforms mainly differ at their 3'UTRs. Our integrated *in silico* approach combining isoform level analysis of microarrays, RNA-seq, and single-cell RNA-seq data allowed the discovery and confirmation of the isoform switch. The microarray data clearly showed downregulation of Isoform-1 because distal probes only recognize this isoform. Interestingly, the ratio of other isoforms was high, and the increased expression of these isoform(s) correlated with patient survival. However, it was unclear which isoform was increased because the probes recognized more than one isoform in the microarray data. To this end, the use of GTEx and TCGA datasets revealed isoform-specific expression patterns in breast cancers. Isoform-1, the dominant transcript in mammary tissue, was downregulated in all PAM50 groups compared to GTEx normal tissues. Other minor isoforms (Isoform-3 and 4) were also lower in tumors than adjacent normal or GTEx normal tissue. In contrast, Isoform-2 was the only isoform that was upregulated in breast cancers. This pattern suggested proximal polyadenylation to favor Isoform-2, which has the most proximal poly(A) site, over Isoform-1 and other isoforms with distal poly(A) sites. We wanted to understand the consequence of this switch, and we found that the dominant isoform (Isoform-1) in breast tissue has a unique



◀**Figure 5.** HNRNPA1 protein levels and function in breast cancers. (A) Upregulation of HNRNPA1 protein in breast tumors compared with adjacent tissue pairs (n = 52) and in basal-like breast cancers (n = 13) compared with adjacent tissue. Fold change values were taken from Tang et al., study²⁸ (****p < 0.0001, Wald test), (B) Kaplan–Meier plots were generated from the KM Plotter database (<http://kmplot.com/analysis/>). Plots represent the percentage of overall survival in breast cancer patients in the Tang et al., study. (HNRNPA1 protein high: red and low: black) (p = 0.019, HR = 2.8, and FDR = 50%). (C) HNRNPA1 protein expression in luminal, HER2+ and TNBC groups is from the CPTAC data from UALCAN (<http://ualcan.path.uab.edu/analysis.html>) (****p < 0.0001). Log2 Spectral count ratio values from CPTAC were first normalized within each sample profile, and then normalized across samples. Z-values represent standard deviations from the median. (D) Two independent clones (sh1, sh2) are shown with decreased HNRNPA1 protein levels compared to non-targeting (NT) shRNA transfected controls. The same blots were hybridized with ACTB antibody to test sample loading. The image is representative of 3 independent experiments. Graphs show densitometric quantification of bands, uncropped images are presented in Supplementary Fig. S13 (*p < 0.05, **p < 0.01, ***p < 0.001; n = 3, one-way ANOVA, Tukey's HSD). (E) Colony formation of NTsh and HNRNPA1sh cells after 7 or 14 days in MCF7 and MDA-MB-231 cells, respectively. Colonies were counted and analyzed using CountPHICS software (***p < 0.001, ****p < 0.0001; n = 3 independent experiments, one-way ANOVA, Tukey's HSD). (F) Effect of HNRNPA1 silencing on proliferation rates detected by MTT (*p < 0.05; n = 3 independent experiments, one-way ANOVA, Tukey's HSD). (G) Effect of HNRNPA1 silencing on wound healing property of cells. Cell layers in each well were scratched by a pipette tip. Closure of wound in each well was examined at days 0, 1, 2, and 3. (***p < 0.001 and ****p < 0.0001; n = 3, one-way ANOVA, Tukey's HSD). H. Effect of HNRNPA1 knockdown on cell migration and invasion in MDA-MB-231 cells. Cells were allowed to pass through the transwell for 15 h or Matrigel-coated membranes for 18 h (***p < 0.001 and ****p < 0.0001; n = 3 independent experiments, one-way ANOVA, Tukey's HSD).

3'UTR sequence and is quite unstable, possibly controlling the levels of HNRNPA1 protein. On the other hand, Isoform-2 is stable and translated more than Isoform-1, as was shown by forced expression and reporter assays. These results suggested that this switch may lead to overexpression of HNRNPA1 protein. Indeed independent proteome datasets revealed overexpression of HNRNPA1 protein in tumors and a correlation with poor survival. In support of an oncogenic role, depletion of HNRNPA1 had a significant effect on neoplastic phenotypes in RNAi silenced cell models. The decreased neoplastic phenotypes in vitro were substantial in the RNAi models. Of note, our CRISPR/Cas9 knockout models did not survive. This observation is in agreement with cell dependency scores listed in DepMap and canSAR datasets. Hence these data suggested HNRNPA1 function is critical and possibly not recovered by other members of the HNRNPs.

Given all the diverse roles of HNRNPA1, we sought to provide additional insight into the HNRNPA1 function in breast cancers. A high throughput miRNA expression assay showed a global downregulation of mature miRNA levels. Further analyses showed that the majority of these miRNAs were located within host genes. Among these, we showed miR-21 promoter activity was decreased, and pri-miR-21 levels were downregulated. While the effect on global downregulation of miRNAs is possibly an indirect consequence of HNRNPA1 loss, it will be essential to delineate the HNRNPA1 downstream players responsible for the transcription of pri-miRNAs listed here. Considering these results and the known roles of HNRNPA1 in RNA metabolism⁴¹, upregulation of HNRNPA1 through the isoform switch may significantly affect different transcriptome components. Of note, while Isoform-2 is upregulated, other isoforms still contribute to HNRNPA1 protein synthesis. The switch enhances protein overexpression but hinders detection of overexpression at the transcript level.

Overall, our data emphasize that focusing on isoform level changes is essential to decoding the cancer transcriptome in higher resolution. This perspective may allow the identification of new oncogene activation cases where overall mRNA levels may not change significantly or common driver mutations do not exist at the genome level. In addition, isoform-specific expression data could also be critical to study isoform-specific post-translational modifications of proteins. Therefore, looking for isoform switches in cancer transcriptomes is a promising strategy to discover new cancer genes with biological impact. The isoform switch we describe for HNRNPA1 has implications in breast cancer and possibly other malignancies.

Methods

Isoform level analysis. CEL files of GSE31519, GSE2034, GSE7390 datasets (and normal breast tissue arrays listed in Supplementary Table S4) were analyzed by APADetect for isoform level quantification, as was described^{16,17}. Briefly, data sets were analyzed for probe-level differences based on the positions of poly(A) sites. Ratios of proximal to distal probe sets were calculated in normal and cancer samples. Significant changes in the ratio of proximal/distal probe sets, separated by poly(A) sites, were reported as ln(SLR) ((Short + Long)/Long ratio).

Cell lines. MCF7, MDA-MB-231, and MCF10A cell lines (ATCC HTB-22, HTB-26, and CRL-10317) were grown as suggested by the manufacturer. Cells were checked regularly for mycoplasma contamination by PCR.

RT-qPCRs and 3'RACE. MIQE guidelines were followed for RT-qPCRs⁴³. RNA isolation, cDNA synthesis and RT-qPCR were performed as described^{16,17} (Supplementary Table S1). RPLP0 was used as reference gene for RT-qPCRs. Breast cancer cell lines and Breast Cancer cDNA array IV (Origene, BCRT104) were used as described^{16,17}. Rapid amplification of cDNA ends (RACE)-specific cDNA synthesis was performed using the 3' RACE Kit (Roche). Cells were treated with actinomycin D (Tocris Bioscience) (2 µg/mL for MCF7, 10 µg/mL for

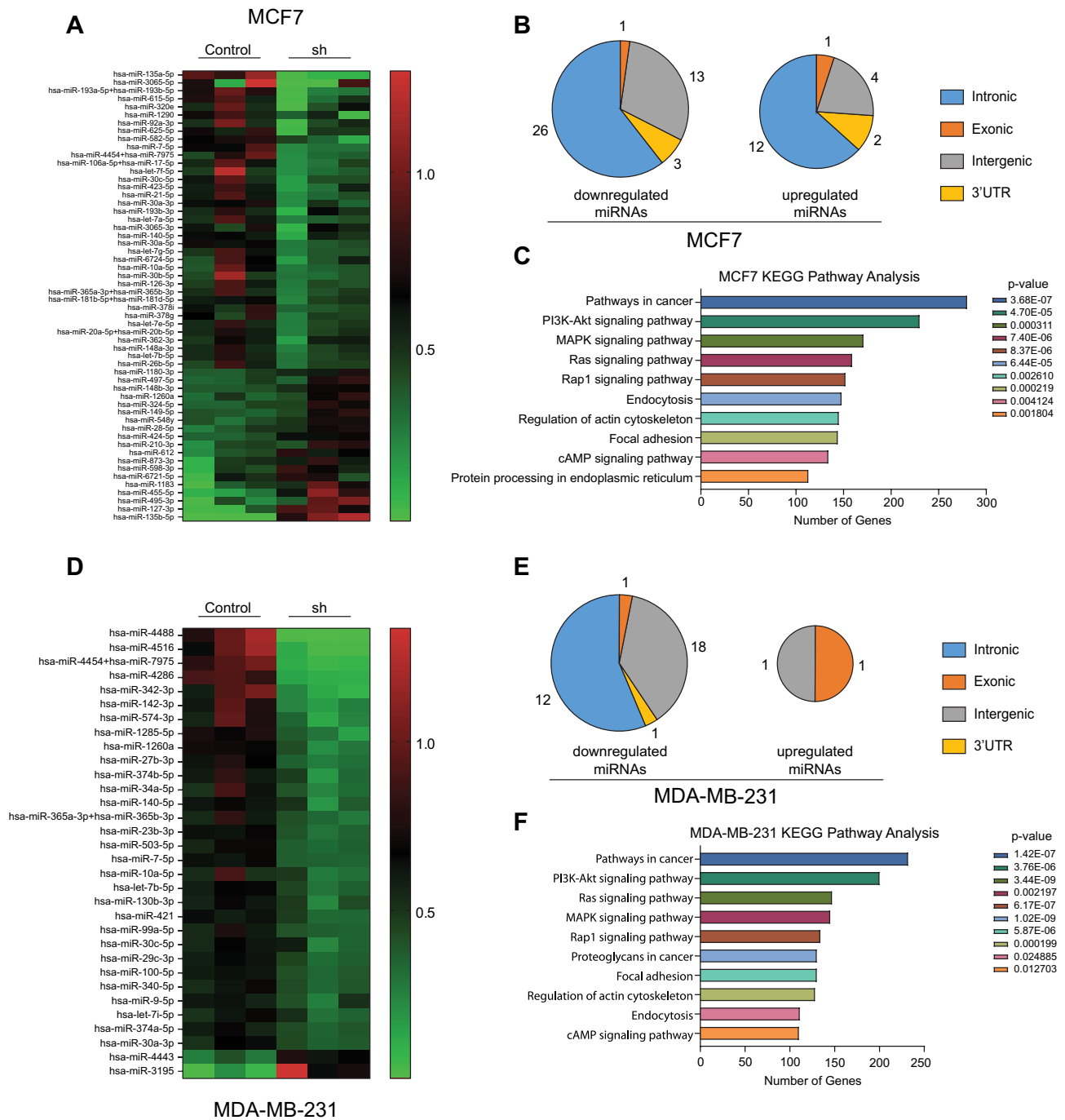


Figure 6. Effect of *HNRNPA1* silencing on miRNAs. **(A)** RNA isolated from *HNRNPA1* silenced, and control cells were detected by NanoString miRNA panel. Heatmap shows miRNA expression fold changes (<0.6 and >1.5) in *HNRNPA1*sh (A1-sh) MCF7 cells. **(B)** Pie charts visualize the proportion of differentially expressed miRNAs grouped according to their genomic features in *HNRNPA1*sh (A1-sh) MCF7 cells (blue color for the intronic miRNAs, orange for exonic miRNAs, gray for intergenic, and yellow for miRNAs located in 3'UTRs of host genes). **(C)** Biological pathways affected by miRNAs whose expression levels were changed upon long-term *HNRNPA1* silencing in MCF7 were determined by DIANA-mirPath. **(D)** Heatmap shows miRNA expression fold changes (<0.6 and >1.5) in A1-sh MDA-MB-231 cells. **(E)** Pie charts visualize the proportion of differentially expressed miRNAs grouped according to their genomic features in *HNRNPA1*sh (A1-sh) MDA-MB-231 cells. **(F)** Biological pathways affected by *HNRNPA1* regulated miRNAs in MDA-MB-231 cells.

MDA-MB-231) to determine decay rates. For miRNA quantification, cDNAs were synthesized from total RNA with TaqMan MicroRNA Reverse Transcription Kit (Applied Biosystems, 4427975). TaqMan Universal Master Mix II (Applied Biosystems, 4440040) was used with hsa-miR-21 (000397), hsa-miR-27b (000409), and control RNU43 (001095).

HNRNPA1 expression and silencing experiments. HNRNPA1 isoforms with Hemagglutinin (HA) tag sequence were PCR amplified (Supplementary Table S2) using Phusion High-Fidelity DNA Polymerase (NEB) and cloned into pcDNA 3.1 (-) (Thermo Fisher). *HNRNPA1* short-hairpin⁴⁴ and non-targeting (NT) shRNA oligos were cloned into pSUPER retro.neo-GFP (OligoEngine) (a gift from Dr. Uygur Tazebay). Two monoclonal cell lines were picked and expanded. Protein levels were detected by western blotting using anti-HA (Abcam ab91110), anti-HNRNPA1 (Abcam ab177152), and anti- β -actin (Santa-Cruz sc-47778). Bands were visualized in the ChemiDoc MP Imaging System (Bio-Rad). Blots were minimally cut prior to hybridization with HNRNPA1, were stripped off, and re-hybridized with ACTB. Therefore, the whole-length blots are not provided. The raw images for the blots are given in Supplementary Fig. S13.

Luciferase assays. 3'UTRs of isoforms (211, 350 bp, and 702 bp) were cloned into pMIR-Report (pMIR) (Ambion) (Supplementary Table S3). Cells were co-transfected with pMIR (*Firefly* luciferase) and phRL-TK (*Renilla* luciferase) using TurboFect (Thermo Fisher). Twenty-four hours after transfection, cells were collected, and dual luciferase activities were measured with Dual-Luciferase[®] Reporter Assay (Promega). For promoter activity assay, miR-21 canonical promoter³⁴ was cloned into pGL3-Basic and was co-transfected with phRL-TK. For the miR21 sensor, two complementary binding sites were cloned downstream of the luciferase gene into pMIR. Mutant sensors lack the seed sequences of miR21.

Translating ribosomal affinity purification (TRAP). EGFP-L10A HEK293 cell line was generated by stable transfection with the pEGFP-C1/RPL10A construct. TRAP was designed and performed as described⁴⁵.

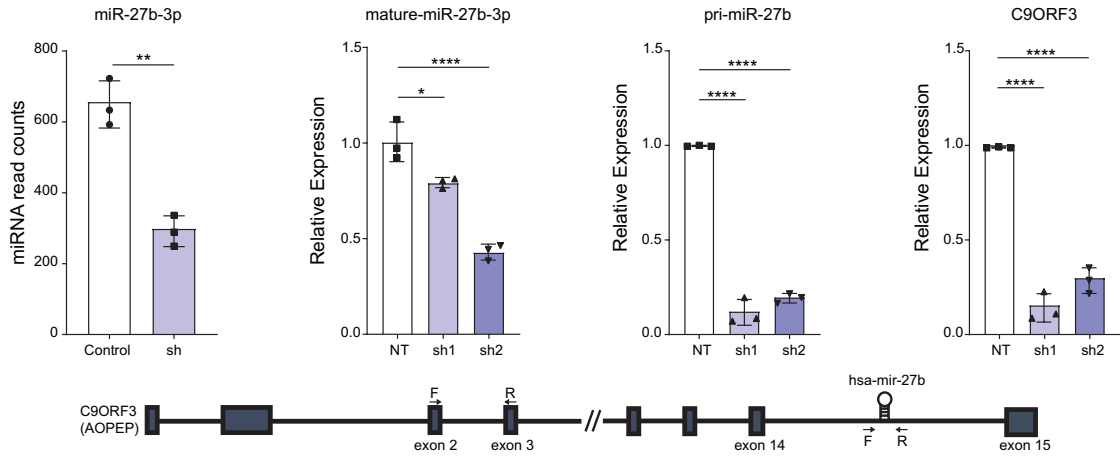
Datasets. Data were derived from public domain resources. RSEM TPM data in the Genotype-Tissue Expression Project (GTEx) (<https://gtexportal.org>)⁷ and The Cancer Genome Atlas (TCGA), Genomic Data Commons Data Portal (GDC Data Portal) (<https://portal.gdc.cancer.gov>) were retrieved from TCGA TARGET GTEx study of UCSC Xena, Xena Toil RNA-Seq Recompute Compendium (<https://toil.xenahubs.net>) (Jan.16, 2021). TCGA TARGET GTEx study in USCS Xena contains re-analyzed data by the same RNA-Seq pipeline for TCGA and GTEx samples. Thus, the batch effect caused by different computational analyses is eliminated for comparison of tumor vs. normal expression¹⁹. The clinical data for TCGA-BRCA samples containing PAM50 status were downloaded from TCGA by the TCGAAbiolinks R package version 2.20.0⁴⁶. Isoform fraction values were calculated for the four transcripts of *HNRNPA1*. Isoform fractions were calculated by dividing the individual isoform expressions (TPM) by the total expression of all isoforms, as described previously⁹. HNRNPA1 protein levels in tumors were retrieved from Tang et al.²⁸. Protein expression for HNRNPA1 in breast cancer subtypes was determined using the CPTAC data from the UALCAN (<http://ualcan.path.uab.edu/analysis.html>) database⁴⁷. For single-cell RNA-seq data, representative samples were selected from GSE75688 and GSE113197 (Supplementary Table S4). RNA-Seq data analyses were performed through the Cancer Genomics Cloud (CGC), powered by Seven Bridges⁴⁸. Cell dependency scores were retrieved from DepMap portal (<https://depmap.org/portal/>) using canSAR (<https://cansarblack.icr.ac.uk/>)^{49–53}. Microarray datasets are also listed in Supplementary Table S4.

Phenotype assays. Colonies were grown for 7 or 14 fourteen days for MCF7 and MDA-MB-231, respectively. Colonies were photographed and counted via the ImageJ⁵⁴. Wounded cell layers were monitored for up to 48 h for wound healing. Images were captured (Olympus Corp.), and wound widths were measured (ImageJ). Colonies were counted and analyzed using CountPHICS (<http://www.fuw.edu.pl/~bbrzozow/FizMed/countPHICS.html>)⁵⁵. Images of wounds are given in Supplementary Fig. S13. MTT was performed as described⁵⁶. Migration and invasion assays were performed using a transwell system with an 8- μ m pore size (Corning). Cells were allowed to migrate for 15 h or invade Matrigel-coated membranes for 18 h. Cells on the apical surface were fixed, stained, and counted as described⁵⁶. All assays were repeated three independent times with at least three technical replicates.

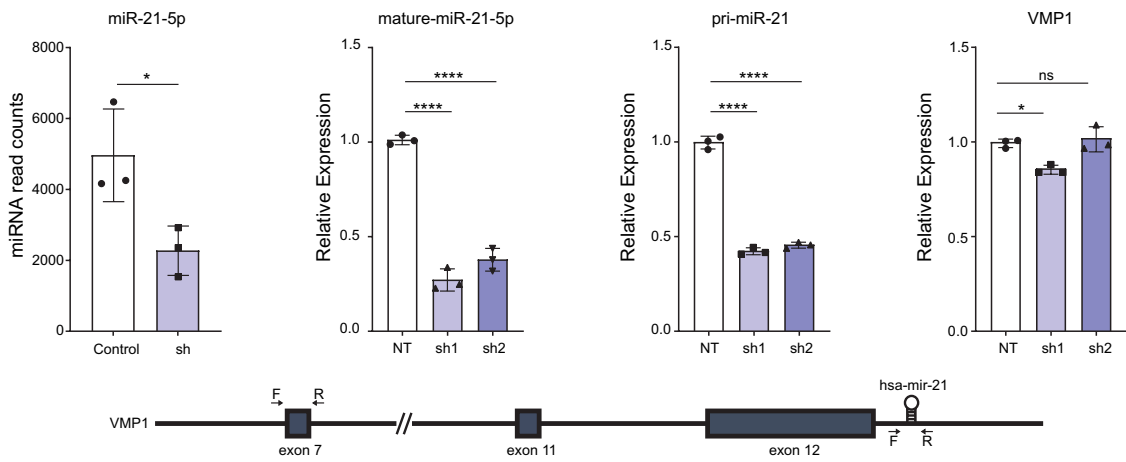
NanoString nCounter miRNA assay. NanoString nCounter Human miRNA V3 was performed according to manufacturer's instructions (NanoString Technologies) at CanSyL/M.E.T.U. Significance was calculated using an unpaired t-test for the three technical replicates. Significant expression changes were listed based on fold changes (<0.6 and >1.5) and p-values ($p < 0.05$). Heatmaps were drawn using GraphPad Prism 8.0.2. Biological pathways affected by miRNAs were determined by using DIANA TOOLS mirPath v.3⁵⁷.

Survival analysis. Expression of isoforms, determined by the proximal probes of 200016_x_at, was used to group patients in the GSE31519 dataset. Patients were grouped according to top 25% (High, $n = 90$) and bottom 25% (Low, $n = 90$) expressers. The survival graph for HNRNPA1 protein was from the cohort described in Tang et al.²⁸. Hazard ratio (HR) with 95% confidence intervals and log-rank p-value were calculated using Kaplan-Meier Plotter⁵⁸.

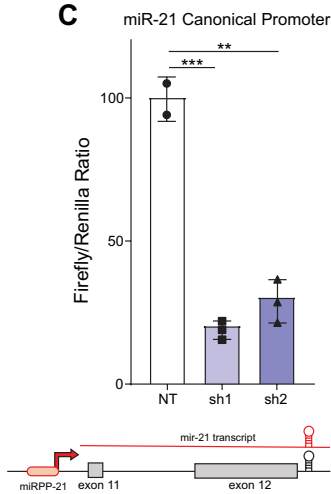
A



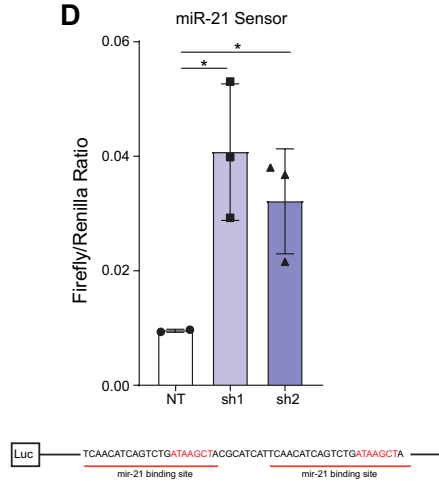
B



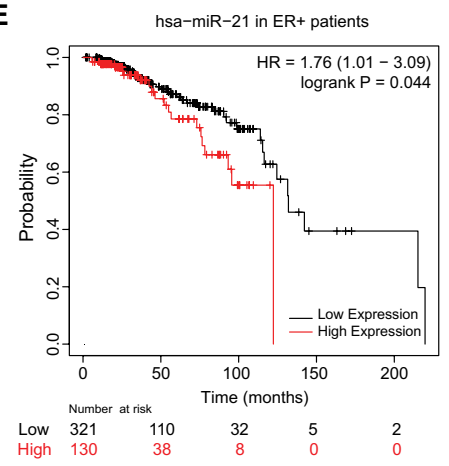
C



D



E



◀Figure 7. MiRNA levels in HNRNPA1 silenced cells. **(A)** miR-27b-3p count reads were decreased in HNRNPA1 silenced MDA-MB-231 cells on the NanoString miRNA panel. Mature miRNA levels in HNRNPA1 silenced MDA-MB-231 cells were detected by TaqMan miRNA PCR ($n=2$ biological replicates), pri-miR-27b and host gene *C9ORF3* expression levels were quantified by RT-qPCR ($*p < 0.05$, $**p < 0.01$, $****p < 0.0001$; $n=3$ biological replicates, one-way ANOVA, Tukey's HSD). Gene structure of host gene, miR-27b, and primer positions are shown below the graphs. **(B)** miR-21 count reads were decreased in HNRNPA1 silenced MCF7 cells. Mature miRNA levels in HNRNPA1 silenced MCF7 cells were detected by TaqMan miRNA PCR, pri-miR-21, and host gene (*VMP1*) expression levels were quantified by RT-qPCR. ($*p < 0.05$, $**p < 0.01$, $****p < 0.0001$; $n=3$, one-way ANOVA, Tukey's HSD). Gene structure of host gene, miR-21 and primer positions are shown below the graphs. **(C)** Control and HNRNPA1 silenced MCF7 cells were transfected with miR-21 promoter-reporter and *Renilla* luciferase vector. Normalized luciferase activity is shown in HNRNPA1 silenced cells compared to controls (miRPP-21 is the canonical mir-21 promoter) ($**p < 0.01$ and $***p < 0.001$ $n=3$, one-way ANOVA, Tukey's HSD). **(D)** miR21 sensor activity is low in control MCF7 cells compared to HNRNPA1 silenced cells ($*p < 0.05$, $n=3$, one-way ANOVA, Tukey's HSD). **(E)** High miR-21 expression correlates with poor overall survival in ER+ breast cancer patients (TCGA) (from KM-plotter database) (logrank $p=0.044$, HR=1.76, FDR>50%).

Data availability

All data generated or analyzed during this study are included in this submitted article and Supplementary Information.

Received: 15 July 2021; Accepted: 6 December 2021

Published online: 27 December 2021

References

- Pan, Q., Shai, O., Lee, L. J., Frey, B. J. & Blencowe, B. J. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.* **40**, 1413–1415 (2008).
- Barbosa-Morais, N. L. *et al.* The evolutionary landscape of alternative splicing in vertebrate species. *Science* **338**, 1587–1593 (2012).
- Wang, E. T. *et al.* Alternative isoform regulation in human tissue transcriptomes. *Nature* **456**, 470–476 (2008).
- Oltean, S. & Bates, D. O. Hallmarks of alternative splicing in cancer. *Oncogene* **33**, 5311–5318 (2014).
- Sandberg, R., Neilson, J. R., Sarma, A., Sharp, P. A. & Burge, C. B. Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. *Science* **320**, 1643–1647 (2008).
- Dunham, I. *et al.* An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
- Melé, M. *et al.* The human transcriptome across tissues and individuals. *Science* **348**, 660–665 (2015).
- Trapnell, C. *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511–515 (2010).
- Vitting-Seerup, K. & Sandelin, A. The landscape of isoform switches in human cancers. *Mol. Cancer Res.* **15**, 1206–1220 (2017).
- Sebestyén, E., Zawisza, M. & Eyras, E. Detection of recurrent alternative splicing switches in tumor samples reveals novel signatures of cancer. *Nucleic Acids Res.* **43**, 1345–1356 (2015).
- Cloutier, P. *et al.* Antagonistic effects of the SRp30c protein and cryptic 5' splice sites on the alternative splicing of the apoptotic regulator Bcl-x. *J. Biol. Chem.* **283**, 21315–21324 (2008).
- Stevens, M. & Oltean, S. Modulation of the apoptosis gene Bcl-x function through alternative splicing. *Front. Genet.* **10**, 804 (2019).
- Poulikakos, P. I. *et al.* RAF inhibitor resistance is mediated by dimerization of aberrantly spliced BRAF(V600E). *Nature* **480**, 387–390 (2011).
- Kahraman, A., Karakulak, T., Szklarczyk, D. & von Mering, C. Pathogenic impact of transcript isoform switching in 1209 cancer samples covering 27 cancer types using an isoform-specific interaction network. *Sci. Rep.* **10**, 14453 (2020).
- Zhao, W., Hoadley, K. A., Parker, J. S. & Perou, C. M. Identification of mRNA isoform switching in breast cancer. *BMC Genomics* **17**, 1–12 (2016).
- Akman, B. H., Can, T. & Erson-Bensan, A. E. Estrogen-induced upregulation and 3'-UTR shortening of CDC6. *Nucleic Acids Res.* **40**, 10679–10688 (2012).
- Akman, H. B., Oyken, M., Tuncer, T., Can, T. & Erson-Bensan, A. E. 3'UTR shortening and EGF signaling: Implications for breast cancer. *Hum. Mol. Genet.* **24**, 6910–6920 (2015).
- Carithers, L. J. & Moore, H. M. The genotype-tissue expression (GTEx) project. *Biopreserv. Biobank.* **13**, 307–308 (2015).
- Vivian, J. *et al.* Toil enables reproducible, open source, big biomedical data analyses. *Nat. Biotechnol.* **35**, 314–316 (2017).
- Goldman, M. J. *et al.* Visualizing and interpreting cancer genomics data via the Xena platform. *Nat. Biotechnol.* **38**, 675–678 (2020).
- Aran, D. *et al.* Comprehensive analysis of normal adjacent to tumor transcriptomes. *Nat. Commun.* **8**, 1–14 (2017).
- Kang, Y. J. *et al.* CPC2: A fast and accurate coding potential calculator based on sequence intrinsic features. *Nucleic Acids Res.* **45**, W12–W16 (2017).
- Herrick, D. J. & Ross, J. The half-life of c-myc mRNA in growing and serum-stimulated cells: Influence of the coding and 3' untranslated regions and role of ribosome translocation. *Mol. Cell. Biol.* **14**, 2119–2128 (1994).
- Schneider-Poetsch, T. *et al.* Inhibition of eukaryotic translation elongation by cycloheximide and lactimidomycin. *Nat. Chem. Biol.* **6**, 209–217 (2010).
- Beelman, C. A. & Parker, R. Differential effects of translational inhibition in cis and in trans on the decay of the unstable yeast MFA2 mRNA. *J. Biol. Chem.* **269**, 9687 (1994).
- Berkovits, B. D. & Mayr, C. Alternative 3' UTRs act as scaffolds to regulate membrane protein localization. *Nature* **522**, 363–367 (2015).
- Mayr, C. Regulation by 3'-untranslated regions. *Annu. Rev. Genet.* **51**, 171–194 (2017).
- Tang, W. *et al.* Integrated proteotranscriptomics of breast cancer reveals globally increased protein-mRNA concordance associated with subtypes and survival. *Genome Med.* **10**, 94 (2018).
- Michlewski, G. & Cáceres, J. F. Antagonistic role of hnRNP A1 and KSRP in the regulation of let-7a biogenesis. *Nat. Struct. Mol. Biol.* **17**, 1011–1018 (2010).
- Michlewski, G., Guil, S. & Cáceres, J. F. Stimulation of pri-miR-18a processing by hnRNP A1. *Adv. Exp. Med. Biol.* **700**, 28–35 (2010).
- Yan, L. X. *et al.* MicroRNA miR-21 overexpression in human breast cancer is associated with advanced clinical stage, lymph node metastasis and patient poor prognosis. *RNA* **14**, 2348–2360 (2008).

32. Wang, H. *et al.* MicroRNA-21 promotes breast cancer proliferation and metastasis by targeting LZTFL1. *BMC Cancer* **19**, 5951 (2019).
33. Ribas, J. *et al.* A novel source for miR-21 expression through the alternative polyadenylation of VMP1 gene transcripts. *Nucleic Acids Res.* **40**, 6821–6833 (2012).
34. Fujita, S. *et al.* miR-21 gene expression triggered by AP-1 is sustained through a double-negative feedback mechanism. *J. Mol. Biol.* **378**, 492–504 (2008).
35. Kim, H. J. *et al.* Mutations in prion-like domains in hnRNP A2B1 and hnRNP A1 cause multisystem proteinopathy and ALS. *Nature* **495**, 467–473 (2013).
36. Obeng, E. A., Stewart, C. & Abdel-Wahab, O. Altered RNA processing in cancer pathogenesis and therapy. *Cancer Discov.* **9**, 1493–1510 (2019).
37. Erson-Bensan, A. E. RNA-biology ruling cancer progression? Focus on 3'UTRs and splicing. *Cancer Metast. Rev.* **39**, 887–901 (2020).
38. Lee, S. H. *et al.* Widespread intronic polyadenylation inactivates tumour suppressor genes in leukaemia. *Nature* **561**, 127–131 (2018).
39. Begik, O., Oyken, M., Alican, T. C., Can, T. & Erson-bensan, A. E. Alternative polyadenylation patterns for novel gene discovery and classification in cancer alternative polyadenylation patterns for novel gene discovery and classification in cancer. *Neoplasia* **19**, 574–582 (2017).
40. Clarke, J. P., Thibault, P. A., Salapa, H. E. & Levin, M. C. A Comprehensive analysis of the role of hnRNP A1 function and dysfunction in the pathogenesis of neurodegenerative disease. *Front. Mol. Biosci.* **8**, 217 (2021).
41. Gebauer, F., Schwarzl, T., Valcárcel, J. & Hentze, M. W. RNA-binding proteins in human genetic disease. *Nat. Rev. Genet.* **22**, 185–198 (2021).
42. Roy, R., Huang, Y., Seckl, M. J. & Pardo, O. E. Emerging roles of hnRNP A1 in modulating malignant transformation. *Wiley Interdiscipl. Rev. RNA* **8**, e1431 (2017).
43. Bustin, S. A. *et al.* The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem.* **55**, 611–622 (2009).
44. Guil, S., Long, J. C. & Cáceres, J. F. hnRNP A1 relocalization to the stress granules reflects a role in the stress response. *Mol. Cell Biol.* **26**, 5744–5758 (2006).
45. Heiman, M., Kulicke, R., Fenster, R. J., Greengard, P. & Heintz, N. Cell type-specific mRNA purification by translating ribosome affinity purification (TRAP). *Nat. Protoc.* **9**, 1282–1291 (2014).
46. Antonia, C. *et al.* TCGAAbiolinks: An R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res.* **44**, e71 (2016).
47. Chandrasekar, D. S. *et al.* UALCAN: A portal for facilitating tumor subgroup gene expression and survival analyses. *Neoplasia* **19**, 649–658 (2017).
48. Lau, J. W. *et al.* The cancer genomics cloud: Collaborative, reproducible, and democratized: A new paradigm in large-scale computational research. *Can. Res.* **77**, e3–e6 (2017).
49. Coker, E. A. *et al.* canSAR: Update to the cancer translational research and drug discovery knowledgebase. *Nucleic Acids Res.* **49**, D1074–D1082 (2021).
50. Meyers, R. M. *et al.* Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nat. Genet.* **49**, 1779–1784 (2017).
51. Dempster, J. M. *et al.* Extracting biological insights from the project achilles genome-scale CRISPR screens in cancer cell lines. *BioRxiv* **20**, 720243 (2019).
52. Dempster, J. M. *et al.* Chronos: A CRISPR cell population dynamics model. *BioRxiv* **25**, 432728 (2021).
53. Ghandi, M. *et al.* Next-generation characterization of the cancer cell line encyclopedia. *Nature* **569**, 503–508 (2019).
54. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671 (2012).
55. Brzozowska, B. *et al.* Freeware tool for analysing numbers and sizes of cell colonies. *Radiat. Environ. Biophys.* **58**, 109–117 (2019).
56. Akhavantabasi, S., Sapmaz, A., Tuna, S. & Erson-Bensan, A. E. MiR-125b targets ARID3B in breast cancer cells. *Cell Struct. Funct.* **37**, 27–38 (2012).
57. Vlachos, I. S. *et al.* DIANA-miRPath v3.0: Deciphering microRNA function with experimental support. *Nucleic Acids Res.* **43**, W460–W466 (2015).
58. Nagy, Á., Munkácsy, G. & Györfy, B. Pancancer survival analysis of cancer hallmark genes. *Sci. Rep.* **11**, 787 (2021).

Acknowledgements

We thank Dr. Deniz Cansen Kahraman from CanSyL for the GFP-Sorting experiments and Esra Cicek for sharing actinomycin-treated MCF10A cDNAs. The results shown here are in whole or part based upon data generated by the TCGA Research Network (<https://www.cancer.gov/tcga>). The Seven Bridges Cancer Genomics Cloud has been funded in whole or in part with Federal Funds from the National Cancer Institute, National Institutes of Health, Contract No. HHSN261201400008C and ID/IQ Agreement No. 17X146 under Contract No. HHSN261201500003I.

Author contributions

M.E. and I.O. helped conceptualize the work and conducted the experiments, D.N.D. curated and analyzed patient data, I.G. conducted TRAP experiments, S.G.E. and R.C.A. helped conduct the experiments, T.C. developed the APADetect algorithm and helped in silico analysis, A.E.E.B. conceptualized the idea, oversaw the project, analyzed the data, wrote and reviewed the manuscript. All authors reviewed the manuscript.

Funding

Turkish Association for Cancer Research and Control Agency (Antalya Branch) funded AEED, TÜBİTAK 1002 (218Z109) funded ME, and METU research grant (TEZ-D-108-2021-10779) for graduate students funded IO.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-04007-y>.

Correspondence and requests for materials should be addressed to A.E.E.-B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021