



ARTICLE

<https://doi.org/10.1038/s41467-019-13993-7>

OPEN

U1 snRNP regulates cancer cell migration and invasion in vitro

Jung-Min Oh¹ , Christopher C. Venters¹, Chao Di¹, Anna Maria Pinto¹, Lili Wan¹, Ihab Younis¹, Zhiqiang Cai¹, Chie Arai¹, Byung Ran So¹, Jingqi Duan¹ & Gideon Dreyfuss¹ ^{*}

Stimulated cells and cancer cells have widespread shortening of mRNA 3'-untranslated regions (3'UTRs) and switches to shorter mRNA isoforms due to usage of more proximal polyadenylation signals (PASs) in introns and last exons. U1 snRNP (U1), vertebrates' most abundant non-coding (spliceosomal) small nuclear RNA, silences proximal PASs and its inhibition with antisense morpholino oligonucleotides (U1 AMO) triggers widespread premature transcription termination and mRNA shortening. Here we show that low U1 AMO doses increase cancer cells' migration and invasion in vitro by up to 500%, whereas U1 over-expression has the opposite effect. In addition to 3'UTR length, numerous transcriptome changes that could contribute to this phenotype are observed, including alternative splicing, and mRNA expression levels of proto-oncogenes and tumor suppressors. These findings reveal an unexpected role for U1 homeostasis (available U1 relative to transcription) in oncogenic and activated cell states, and suggest U1 as a potential target for their modulation.

¹ Howard Hughes Medical Institute, Department of Biochemistry and Biophysics, University of Pennsylvania School of Medicine, Philadelphia, PA 19104-6148, USA. *email: gdreyfuss@hhmi.upenn.edu

Widespread shortening of messenger RNA (mRNA) 3'-untranslated regions (3'UTRs) and switches to short mRNA isoforms is a common feature and contributing factor to cell stimulation, seen in immune cells and neurons, and oncogenicity^{1–6}. These shortening events occur due to a shift in usage of more upstream polyadenylation signals (PASs) in the last exon and in introns. In quiescent cells, these PASs are generally silenced by U1 snRNP (U1), vertebrates' most abundant non-coding small nuclear RNP, which is necessary for production of full-length RNA polymerase II transcripts from protein-coding genes and long non-coding RNAs⁷. This U1 activity, called telescripting, requires RNA base-pairing of U1 snRNA 5'-end, which is also required for U1's role in 5' splice site (5'ss) recognition. U1 antisense morpholino oligonucleotides (U1 AMO), which inhibits U1:pre-mRNA base-pairing, triggers widespread premature cleavage and polyadenylation (PCPA), as well as inhibits splicing^{8,9}. Transfection of a high U1 AMO dose that masks all, or nearly all, of U1 snRNA 5'-end causes drastic PCPA from cryptic PASs frequently found in the 5'-side of the intron in pre-mRNAs of thousands of genes^{8,9}, especially in long introns of large genes (>39 kb). In contrast, small genes (<6.8 kb) are generally PCPA-resistant and many of them are upregulated in this environment. Importantly, these small genes are enriched in functions related to cell survival and acute cell stress response⁷. The drastic PCPA from high U1 AMO in many genes obscures other effects that are more readily detected with low U1 AMO doses. These changes include 3'UTR shortening (shifts to usage of more proximal PASs in tandem PASs) and the production of shorter mRNA isoforms from PAS usage in introns⁹. This revealed that low U1 AMO recapitulates known shifts to shorter mRNA isoforms that occurs in stimulated neuronal cells. We have shown previously that the stimulation-induced rapid transcription upregulation creates transient telescripting deficit because U1 levels cannot rise in step with the transcription surge from immediate early genes⁹. U1 synthesis is a slower process involving nuclear export of pre-snRNAs, SMN complex-mediated snRNP assembly in the cytoplasm, and re-import to the nucleus¹⁰. Consequently, some pre-mRNAs transcribed in this time window (~2–6 h post stimulation) are processed to shorter mRNA isoforms due to PCPA in introns. For example, pre-mRNA processing of *homer-1*, which encodes a synaptogenesis scaffold protein, shifted from full-length mRNA to a shorter mRNA isoform that encodes antagonistic activity due to PCPA in an intron¹¹. Importantly, low U1 AMO recapitulated the same isoform switching⁹.

Here, we investigated if low U1 AMO and U1 over-expression could also modulate cell phenotype. Our studies uncover a role for U1 in regulating cancer cells' phenotype, and identify widespread and diverse transcriptome changes resulting from modulating U1 availability. These changes are consistent with U1's central role in splicing and telescripting, and include many with known functions in cancer.

Results

U1 level changes modulate hallmark phenotypes of cancer cells in vitro. We used standard in vitro assays to determine if moderate U1 inhibition has an effect on proliferation, migration, or invasion of cancer cells, which serve as quantitative measures of oncogenic phenotype. Various low U1 AMO doses (2.5–250 pmole) or control, non-targeting AMO (cAMO) were transfected into HeLa cells, a cervical carcinoma cell line (Fig. 1). These U1 AMO doses masked ~15–30% of U1 snRNA 5'-ends, making it inaccessible for base-pairing corresponding to the U1 AMO dose^{8,9}. As shown in Fig. 1a, 62.5 pmole U1 AMO moderately increased (38%) cell proliferation after 48–72 h. Higher U1 AMO

doses (≥ 250 pmole) were toxic and reduced overall cell numbers. Remarkably, low U1 AMO dose-dependently enhanced cell migration and invasion by up to ~500% in 24 h with peak activity at 62.5 pmole (Fig. 1b–e). The increased migration and invasion reflect true enhancements that could not be accounted for by the comparatively small increase in cell number (Fig. 1a). U2 AMO, which interferes with U2 snRNP's function in splicing¹², did not enhance any of these phenotypes over the entire dose range, indicating that the specificity of the U1 AMO effects (Supplementary Fig. 1).

To further examine the effects of U1 on cell phenotype, we over-expressed U1 (U1 OE) from a plasmid carrying U1 snRNA gene with its native promoter and termination elements. This achieved 20% and 40% U1 level increases compared to empty plasmid by transfecting 1 μ g and 1.5 μ g of this plasmid, respectively (Supplementary Fig. 2). The U1 OE significantly and dose-dependently attenuated cell migration (25–50%) and cell invasion (25–65%) after 24 h (Fig. 2). U1 OE dose-dependently increased the amount of U1 snRNP, determined by the amount of U1 snRNA in anti-Sm immunoprecipitations⁹.

Similar experiments on other cancer cells, including human lung adenocarcinoma (A549) and breast adenocarcinomas (MCF-7 and MDA-MB-231), demonstrated the generality of the U1-related effects on phenotype. U1 AMO enhanced migration of A549, MCF-7, and MDA-MB-231 by 58–72% and increased invasion of A549 and MDA-MB-231 by 53–64% (Supplementary Table 1). Conversely, U1 OE attenuated the migration and invasion of these cell lines by ~50% compared to the control levels (Supplementary Table 1).

U1 level changes cause numerous and diverse transcriptome changes. We used high-throughput RNA sequencing (RNA-seq) to determine transcriptome effects resulting from U1 level modulation. To enhance detection of nascent RNAs, as U1 functions co-transcriptionally⁷, we metabolically labeled RNAs with 4-thiouridine (4-shU) for 2 h (at 6–8 h post-transfection with AMOs and 22–24 h of U1 OE), and sequenced the thiol-selected RNAs. This protocol also readily detects near steady-state mRNAs. Reads were mapped to the human genome (UCSC, hg38) and filtered for unique alignments. The statistics of the RNA-seq datasets, normalized for sequencing depth as reads per million (RPM), are shown in Supplementary Table 2. Over 9100 genes expressed to RPKM ≥ 1 in the control samples for U1 AMO and U1 OE, and were included in further analysis. This revealed numerous and diverse transcriptome changes caused by U1 AMO, including in mRNA expression, 3'UTR length, and alternative splicing affecting thousands of genes. The number of genes affected by each type of change are listed in Supplementary Table 3. Frequently, more than one type of change was detected in transcripts of the same gene (18–47%; Supplementary Fig. 3). Generally, the number of events for each type of change increased with U1 AMO dose and there was extensive overlap (31–97%) between events detected at the two doses tested (Supplementary Table 3). To maintain a high confidence in the transcriptome changes that were analyzed, we did all further analyses and discussion on events that were detected in both doses of U1 AMO or U1 OE, respectively.

Confirming earlier observations from datasets with lower resolution and sequencing depth⁹, low U1 AMO elicited widespread 3'UTR shortening readily detected in genome browser views, for example, *TOP2A*, *TFRC*, and *CDC25A* (Fig. 3). A common approach to identify changes in locations of 3'-poly(A)s is to use oligo(dT)-primed RNA-seq¹³. However, it does not provide adequate information on overall transcriptome changes due to inherent 3' bias. An alternative method, DaPars, uses a regression model to deduce alternative PAS usage among multiple

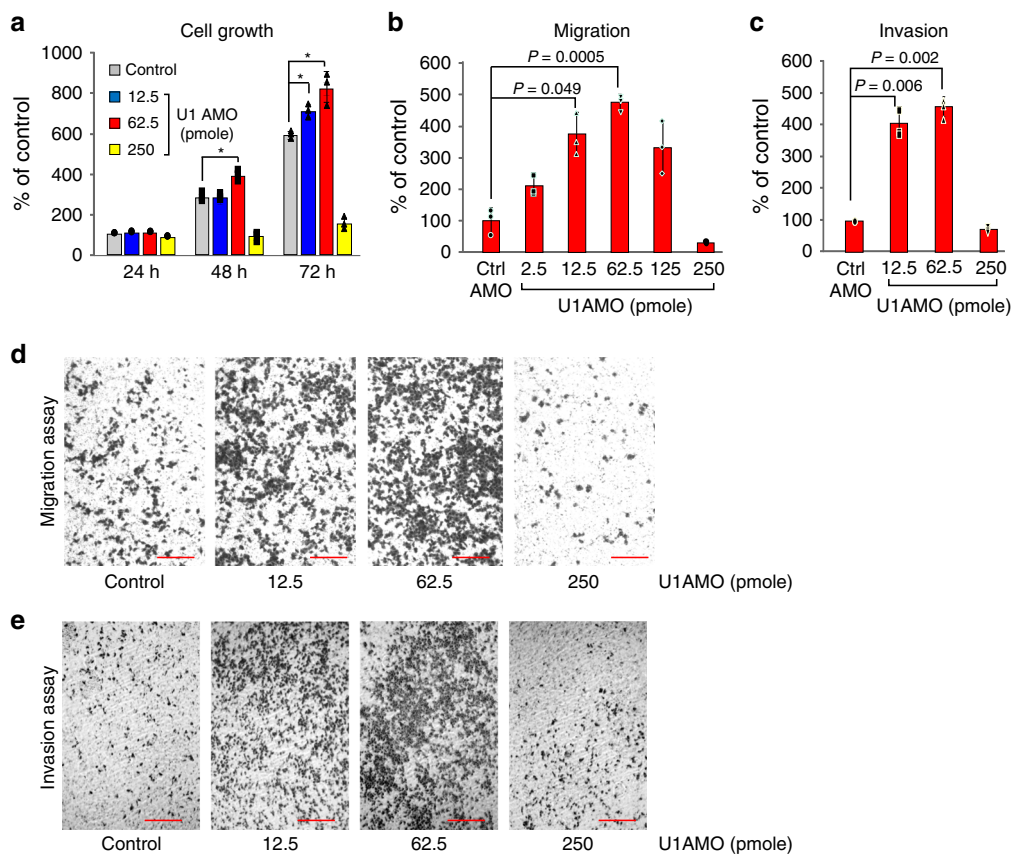


Fig. 1 Low U1 AMO dose-dependently enhances migration and invasion of HeLa cells in vitro. **a** HeLa cells were transfected with the indicated amounts of U1 or control AMOs. Cell proliferation was determined from cell numbers using the Cell Titer-Glo Luminescent Cell Viability Assay at the times indicated. Data are represented as mean \pm standard deviation (SD) ($n = 3$, independent cell culture). P -value was calculated with two-tailed Student's t -test. The asterisk indicates $P < 0.05$. **b, c** Migrating and invading cells were measured 24 h post-transfection. 1.5×10^5 and 2.5×10^5 cells were cultured for the migration and invasion assays, respectively. Cells migrated and invaded from an upper chamber to a bottom chamber over 24 h and were counted using the cell staining solution Cytoselect 24-well cell migration assay kit (Cell Biolabs). The number of migrated and invaded cells were compared with each control. Data are represented as mean \pm SD ($n = 3$, independent cell culture). P -value was calculated with two-tailed Student's t -test. **d, e** Migrating and invading cells were stained 24 h after transfection and visualized by phase-contrast microscopy (10x magnification). The number of migrated cells were 111, 344, 472, and 26 in control, 2.5, 62.5, and 250 U1 AMO (pmole), respectively. The number of invaded cells were 124, 472, 532, and 88 in control, 2.5, 62.5, and 250 U1 AMO (pmole), respectively. Scale bar = 250 μm . Source data are provided as a Source Data file.

tandem 3'UTR PASs from standard RNA-seq¹⁴. Frequently, 3'UTRs have multiple tandem PASs, resulting in complex mixtures of mRNAs with various 3'UTR lengths that make it difficult to resolve by DaPars. We reasoned that a shift in usage to proximal 3'UTR PASs would decrease the overall amount of transcription in the 3'UTR, while a shift to more distal PASs will increase it. As changes in 3'UTR amount could result from mRNA expression level changes alone, we calculated the ratio of RNA-seq reads in the 3'UTR to the reads in the portion of the CDS in the same last exon (heretofore LECDS). This normalizes 3'UTR signal to mRNA expression. A decrease or increase in LECDS in a sample compared to control suggests 3'UTR shortening or lengthening, respectively.

LECDS-identified 3'UTR shortening in 1919 genes in both 12.5 and 62.5 U1 AMO ($p \leq 0.01$), and only a small fraction (130 genes) had more reads in 3'UTRs (Fig. 3a). In contrast, U1 OE caused a 3' increase in reads (longer 3'UTRs) in 3210 mRNAs, while it shortened 3'UTRs in only 344 genes (Fig. 3b). The large overlap of the lengthened (~84%) and shortened genes (~76%) in both U1 OE (1 μg and 1.5 μg) and U1 AMO samples (U1 AMO 12.5 and 62.5), respectively, showed a strong validation of the RNA-seq data from separate biological experiments. Although LECDS is a much simpler calculation and lacks the ability to

define the location of the PAS used, it nevertheless called 3'UTR shortening in 79% of the mRNAs identified as such by DaPars (Supplementary Fig. 4). Further confirmation of 3'UTR shortening and lengthening is shown for select examples in Fig. 3c, d, and 3'RACE shows that they result from APA (Supplementary Fig. 5). Notably, the fraction of genes that had 3'UTR length changes corresponded to the U1 AMO or OE dose, and U1 OE caused transcription in some genes to extend farther downstream from canonical gene ends. Thus, U1 has a role in regulation of both proximal and distal PASs.

As expected for U1's role in splicing, U1 AMO caused alternative splicing (AS) changes in ~700 genes, including A5'SS and A3'SS, cassette exon (CE) and intron retention (IR) events (Supplementary Table 3 and Fig. 4)¹⁵. Examples of AS and confirmation of select cases by RT-PCR are shown in Fig. 4b and Supplementary Fig. 6. Many of the observed AS changes could contribute to the oncogenic phenotype. For example, ataxia-telangiectasia, a cancer predisposition disorder caused by mutations in the ATM gene¹⁶, has two introns retained (intron 1 and 33) and is downregulated in U1 AMO compared to cAMO (Figs. 4b and 5). A similar number of splicing changes is observed with knockdowns of other splicing factors^{15,17,18}.

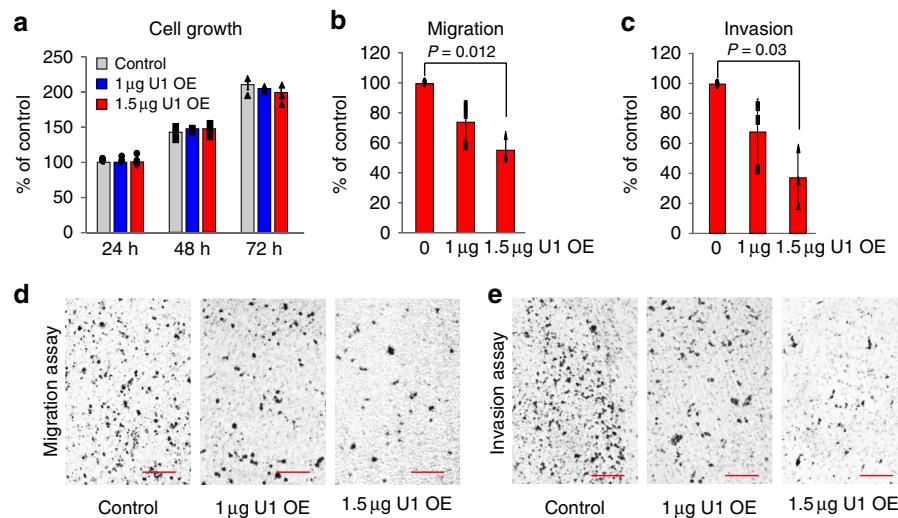


Fig. 2 U1 over-expression decreases cell migration and invasion of HeLa cells in vitro. HeLa cells were transfected with control empty vector or U1 expression plasmids. **a** Cell growth determined in 24 h intervals. The data are represented as mean \pm SD ($n = 3$, independent cell culture). **b, c** Cell migration and invasion were measured 24 h after transfection. The bars indicate mean \pm SD ($n = 3$, independent cell culture). P -value was calculated with two-tailed Student's t -test. **d, e** Cell migration and invasion were determined 24 h post-transfection and imaged by phase-contrast microscopy (10 \times magnification). Scale bar = 250 μ m. Source data are provided as a Source Data file.

Gene ontology analysis (GO term) revealed expression level changes induced by U1 level modulation in multiple oncogenes, tumor suppressors and cell cycle related genes. For example, *ATR*, *RBI*, *ATM*, and several *CDC* (Cell division cycle) genes were downregulated in U1 AMO, but upregulated in U1 OE. On the other hand, oncogene and cancer related genes, such as *MYC*, *FOS*, *JUN*, and growth promoting genes were upregulated in U1 AMO, but downregulated in U1 OE (Fig. 5 and Supplementary Fig. 7). Downregulation could be explained by PCPA in an intron, which is frequently difficult to detect without exosome inhibition as they are rapidly eliminated^{19–21}. Alternatively, downregulation could be caused by inhibition of transcription initiation, as a secondary effect to downregulation or mis-splicing of transcription factors. Upregulation could be caused by PCPA in long genes, which could decrease competition for transcription and splicing factors for non-PCPAed genes, as well as in response to cell stress resulting from PCPA in other genes^{7,22}. Nevertheless, concomitant upregulation of oncogenes and downregulation of tumor suppressor genes likely contribute to the phenotypic changes in the wake of U1 AMO. Among genes affected by U1 AMO are splicing factors, which have been linked to myelodysplastic syndromes, chronic lymphocytic leukemia and other cancers (Supplementary Fig. 8)^{17,23–25}.

U1 level changes alters expression of cancer genes. To explore the potential role of U1 dependent 3'UTR length changes on the oncogenic phenotype we observed, we interrogated our gene set against cancer gene databases [Sanger²⁶ and UCSF (waldman.ucsf.edu/GENES/completechroms.html)]. These archives include oncogenes induced through various mechanisms, including mutation, chromosomal translocation, or loss of miRNA repression, and the analysis revealed that a large number of oncogenes incurred 3'UTR length changes. The oncogenes affected at each U1 level, 204 in total, are listed in Supplementary Table 4 and include cell cycle regulation (*CDC25A*, *CCNB1*), apoptosis (*BCL6*, *BRCA1*), cell migration (*FGFR1*, *FYN*), extracellular matrix remodeling (*TIMP2*), signaling (*EGFR*), transcription (*WNT5A*), metastasis and tumor progression (*EWSR1*, *APC*, *BRAF*).

There are numerous examples of cancers resulting from oncogene upregulation due to loss of miRNA repression, either

because the relevant miRNA is downregulated or its target in the 3'UTR has been removed^{27–30}. U1 level changes recapitulated the same miRNA 3'UTR target elimination or restoration in many genes (Fig. 3c, d and Supplementary Fig. 5). For *CDC25A*, an essential phosphatase for the G1-S transition, 3'UTR shortening eliminates several miRNA-binding sites, including let-7, miR-15, and miR-21 (Fig. 3c). An increase in *CDC25A* protein due to alleviation of miRNA-mediated repression exacerbates hepatic cyst formation and colon cancer^{31–33}. 3'UTR shortening in RAS oncogene family member, *RAB10*, which removes target sites for miRs-103/107, is found in numerous cancer cell lines and results in a dramatic increase of this protein⁴. As shown in Fig. 3d, 3'UTR of *RAB10* is already shortened in HeLa cells, and U1 OE reverses this shortening and restores the corresponding miRNA-binding sites. Similarly, U1 OE lengthened 3'UTR of *KRAS* to include a let-7 binding site (Fig. 3d). This region contains a polymorphism that has been shown to impair let-7 binding, and is prognostic for breast cancer aggressiveness³⁴. U1 AMO also did not change the associated miRNA expression (Supplementary Fig. 9). In addition to loss or gain of miRNA-binding sites, 3'UTR length or amount changes can affect other mRNA regulating elements, such as AU-rich elements (AREs), with roles in cancer phenotype. For example, U1 OE-induced 3'UTR lengthening in *c-Fos*, a gene controlling cell proliferation, differentiation, survival and tumorigenesis^{35,36}, restores an ARE and strongly decreases mRNA level (>2-fold, Fig. 5).

Discussion

The numerous transcriptome changes resulting from moderate perturbation of U1 homeostasis, namely, available U1 relative to the amount of transcription, fall into clear categories that can be explained by U1's known functions in telescripting and splicing. Many of the changes are known to be oncogenic on their own, in humans and mice. For example, U1 AMO recapitulates activation of proto-oncogenes, including 3'UTR shortening (e.g., *EGFR*) or upregulation (e.g., *MYC*), and downregulation or ORF-disruptive splicing changes in tumor suppressors (e.g., *ATM*). While it is possible that a change in one or a few of the affected genes is particularly important, our study strongly suggests that the cumulative effect of multiple transcriptome changes that affect a

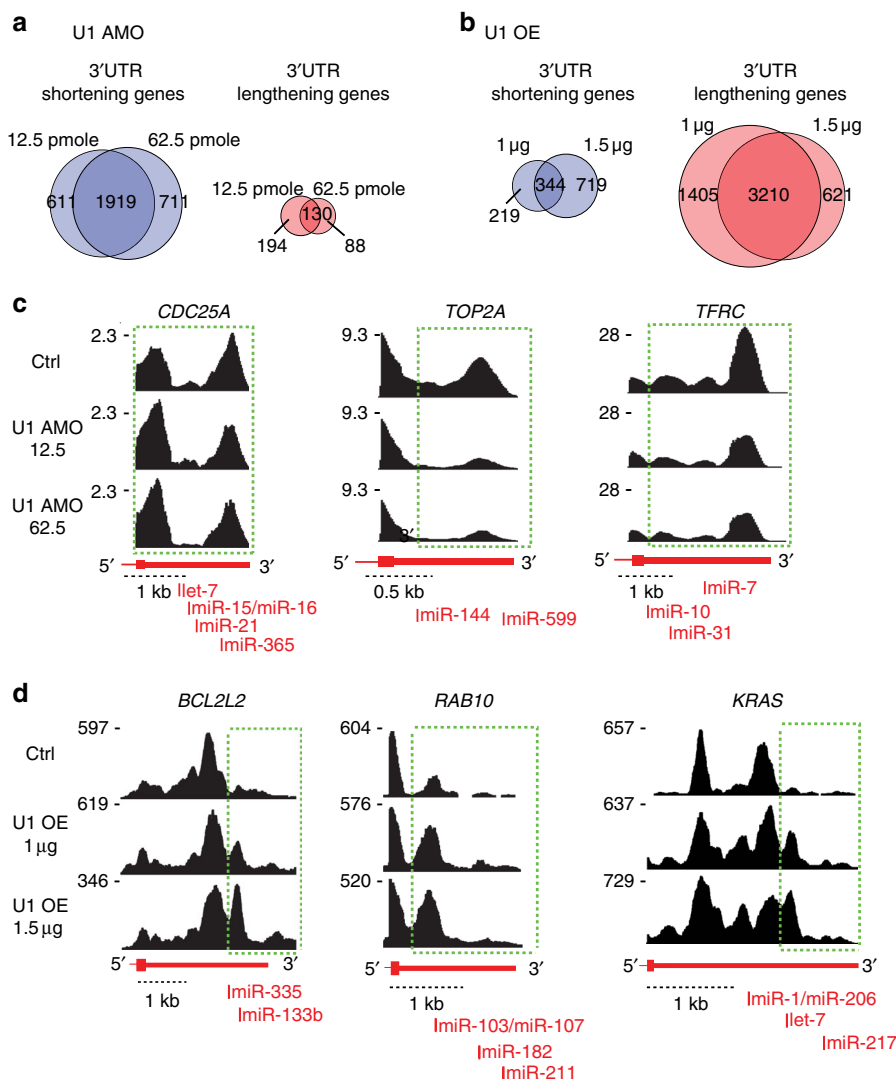


Fig. 3 Examples of genes with 3'UTR shortening by U1 AMO and genes with 3'UTR lengthening from U1 over-expression. Venn diagrams of LECDS-identified genes affected by 3'UTR shortening versus 3'UTR lengthening in U1 AMO 12.5 and 62.5 pmole (**a**) and U1 OE 1 and 1.5 µg (**b**) samples. **c** Cells transfected with control or U1 AMO (12.5 and 62.5 pmole, 8 h) were labeled with 4-thiouridine to select and sequence nascent transcripts. RNA-seq maps of representative genes from two biological experiments are shown. Green dotted boxes indicate the affected regions where shortening in the 3'UTR occurs. **d** Cells transfected with control or U1 expression plasmids (1 µg and 1.5 µg, 24 h) were labeled with 4-thiouridine to select nascent transcripts. Green dotted boxes indicate the regions where there was a switch towards longer 3'UTRs. miRNA target sites that are involved in cancer and expressed in HeLa are shown.

fraction of transcripts in many relevant genes drives the phenotype changes. Our findings demonstrate that U1 homeostasis plays an important role in maintaining normal gene expression balance, and prevent the transcriptome from defaulting to a state in which multiple oncogenic drivers are activated and counter-acting tumor-suppressors are diminished. In that sense, U1's activity is tumor suppressor-like.

Several other factors that can regulate 3'UTR length have been described, particularly components of the cleavage and polyadenylation machinery (CPA)^{13,37–41}. Knockdown of CPSF6/CFIm68, CPSF5/CFIm25, and PABPN1 cause widespread 3'UTR shortening^{40,42,43}. Furthermore, CPSF5 and CPSF6 are down-regulated in cancers and their knockdowns in cell lines enhances oncogenicity^{14,40,44,45}. It is possible that these CPA factors (CPAFs) work in concert with U1, as recent evidence for a complex of U1 with CPAFs suggests²². 3' UTR shortening analysis of the RNA-seq of CPSF5 knockdown in HeLa cells by LECDS-identified 1210 genes (63% overlap with U1 AMO) that

are also shortened in U1 AMO (Supplementary Fig. 10), suggesting potentially extensive commonality of targets. A recent study described that U1 over-expression in PC-12 rat neuronal cell line upregulated cancer related genes, including MYC and FOS⁴⁶. While these results appear to be inconsistent with our observations, they may be explained by the different cell types used in these studies and by different U1 levels relative to transcription.

Mutations in several splicing factors, particularly U2-associated factors and SR proteins, have been shown to drive many cancers^{23,24}. However, the potential role of U1 in cancer has not been tested previously. It is intriguing that a major U1 gene cluster in humans at chr1p36.13 flanks a breakpoint linked to many cancers⁴⁷. The potential effect of these genomic abnormalities on U1 levels remains to be determined, as well as U1's potential role in tumors. Mounting evidence indicate that U1 AMO and U1 over-expression mimic biological processes^{7,9,20,48–50}. Thus, these experimental tools for dose-dependent modulation of

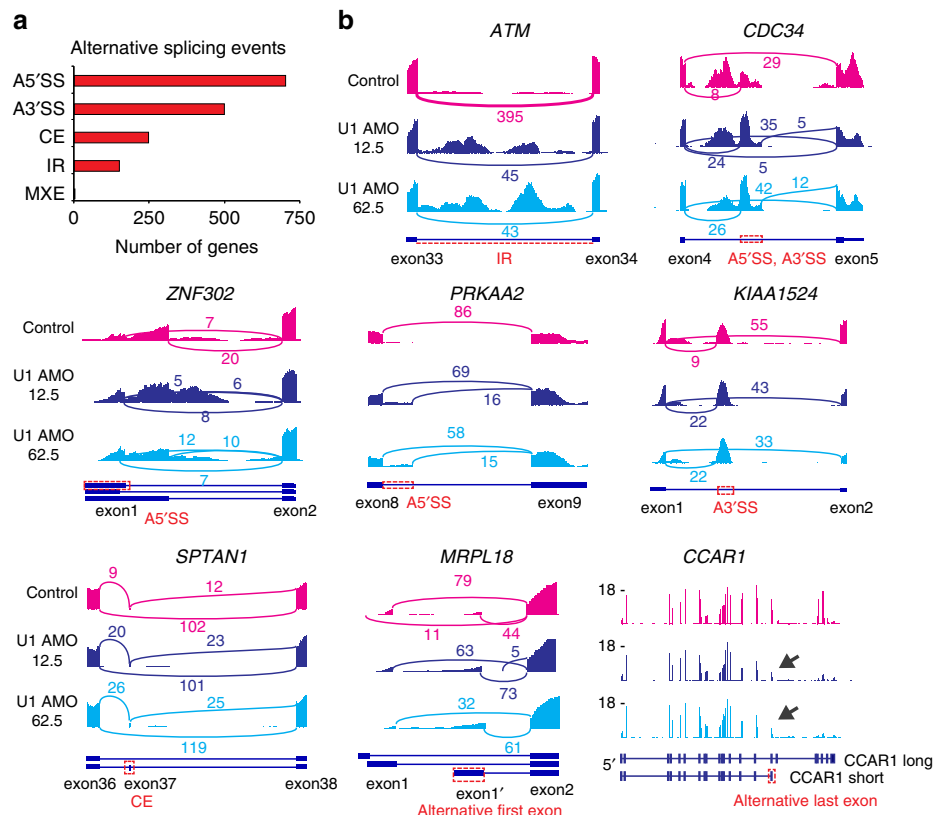


Fig. 4 Low U1 AMO causes widespread alternative splicing changes in HeLa cells. **a** Summary of alternative splicing events as identified with JUM analysis. Each alternative splicing event happened in both low U1 AMO 12.5 and 62.5 pmole in HeLa cells. (A5'SS: alternative 5' splice site; A3'SS: alternative 3' splice site; CE: cassette exon; IR: intron retention; MXE: mutually exclusive exons) **b** Sashimi plots and genome browser view of each alternative splicing change in the low U1 AMO 12.5 and 62.5 pmole.

U1 should facilitate studies on cancer and stimulated cells behavior, and suggest U1 as a potential clinical target for a wide range of disorders.

Methods

Cell culture and treatment. Cells (HeLa, A549, MCF-7 and MDA-MB-231 obtained from ATCC) were maintained in Dulbecco's Modified Eagle Medium supplemented with 10% fetal bovine serum (FBS), 10 units ml⁻¹ penicillin and 10 µg ml⁻¹ streptomycin at 37°C, and 5% CO₂. All cell lines were tested for mycoplasma contamination and were not authenticated. U1, U2 antisense, and control morpholino oligonucleotide sequences were 5'-GGTATCTCCCCTGCCAGGTAAGTAT-3', 5'-CCTCTTACCTCAGTTACAATTTATA-3' and 5'-TGATAAGAACAGATACTACTTGA-3', respectively⁸. Oligonucleotides were transfected into 1 × 10⁶ cells using a Neon transfection system (Invitrogen) according to the manufacturer's instruction to achieve the desired final doses indicated in the text and figures.

Proliferation assay. The Cell Titer-Glo Luminescent Cell Viability Assay (Promega) was used according to the manufacturer's instructions to measure cell proliferation. The cells were transfected with AMOs and seeded in triplicate in 96-well plates at a density of 1 × 10⁴ cells per well. Cells were incubated in media containing 1% FBS for 3 days and proliferation was measured every 24 h.

Migration and invasion assays. For migration studies, a standard assay was used to determine the number of cells that traversed a porous polycarbonate membrane in response to a chemo-attractant (higher serum concentration)^{51,52} using the Cytoselect 24-well cell migration assay (Cell Biolabs). Invasion was measured using BD BioCoat Matrigel invasion chambers (BD Bioscience). In both assays, cells were transfected with AMOs, and 1.5 × 10⁵ and 2.5 × 10⁵ cells per well were seeded in an upper chamber in serum free media for the migration and invasion assay, respectively. The lower chamber was filled with media containing 10% FBS. After 24 h, cells passing through polycarbonate membrane were stained and counted according to the manufacturer's instructions.

Metabolic RNA labeling, isolation, and RNA-seq. 4-thiouridine (250 µM) was added to cells between 6–8 h after U1 AMOs transfection. Total RNA was extracted with Trizol (Invitrogen) and poly(A) RNA purified on Oligotex beads (Qiagen). Free thiols on poly(A) mRNA were reacted with 0.2 mg ml⁻¹ biotin-HPDP for 2 h to label RNA that incorporated 4-thiouridine. RNA was then purified on M-280 streptavidin Dynabeads (Invitrogen), cDNA was synthesized using Ovation RNA-Seq System V2 (NuGEN) and libraries for Illumina sequencing were constructed using Encore NGS Library System (NuGEN) according to the manufacturer's instructions.

Mapping RNA-seq reads. RNA-seq reads were aligned to reference genome UCSC/hg38 using STAR⁵³ version 2.6.0b with default settings. Reads per exon were grouped, from which RPKM (Read Per Kilobase of exon model per Million mapped reads)⁵⁴ values were calculated. Only genes with RPKM ≥ 1 were included in further analyses.

Differential gene expression, alternative splicing, and miRNA analyses. The Gfold algorithm⁵⁵ was used to calculate differential expression using the default parameters for detecting reliable expression change on all exons of the longest full-length isoform excluding the 3'UTR. A 1.5-fold change was set with a *P*-value < 0.01 to identify genes with significant expression change. The Junction Usage Model (JUM v2.0.2)¹⁵ analysis was performed to identify differentially spliced isoforms in experimental samples compared to controls and quantify their expression levels by computing the ΔPSI (difference of Percent Spliced Isoform). Significant differential alternative splicing events were detected by performing a X² likelihood-ratio test followed by Benjamini–Hochberg (BH) multiple testing with an adjusted *P*-value ≤ 0.01. For stringent inclusion, we also restricted ΔPSI ≥ 10%. TargetScan miRNA regulatory sites were downloaded from UCSC Genome browser (<http://genome.ucsc.edu/>). The coordinate of each predicted miRNA site was compared to those genes for which the 3'UTR showed changes in length. For miRNA microarray, HeLa cells were transfected with control or 62.5 pmole U1 AMO for 8 h. Total RNA was extracted using miRNeasy kit (Qiagen) according to manufacturer recommendations. GeneChip miRNA 3.0 array (Affymetrix) was used to determine the miRNA expression and experiments and analysis were performed by Molecular Profiling Facility at University of Pennsylvania.

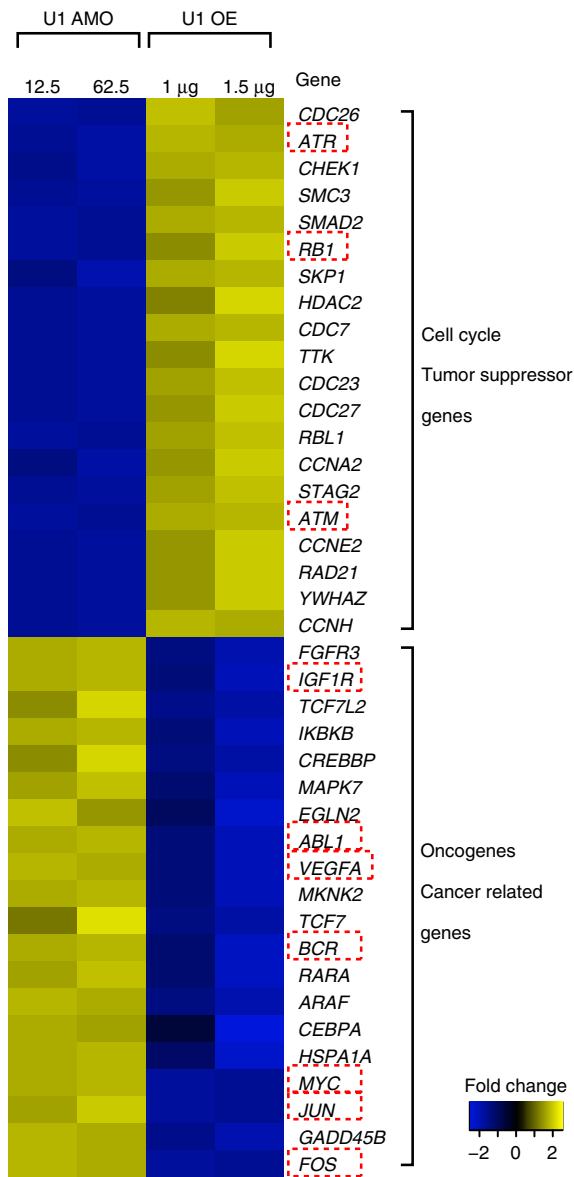


Fig. 5 U1 AMO upregulated oncogenes and downregulates tumor suppressor genes and U1 over-expression antagonizes the effect in HeLa cells. Heat map showing genes related to cell cycle and tumor suppressor function are downregulated in U1 AMO and upregulated in U1 OE. Oncogene and cancer related genes are upregulated in U1 AMO and downregulated in U1 OE (1.5-fold change, P -value < 0.01 calculated by GFold⁵⁵). Source data are provided as a Source Data file.

LECDs calculations. We used both the 3'UTR and coding sequence of the terminal exon to determine the change in reads in 3'UTR of the longest (full-length) isoform. As 3'UTR reads can change with expression level changes, such as transcription up- or downregulation, 3'UTR signals were compared to that from the last exon's coding sequence (LECDs). If the 3'UTR reads are significantly decreased or increased relative to those from the LECDs, the 3'UTR is called as shortening or lengthening, respectively. The significance of this read change was detected using a Fisher's Exact test followed by Benjamini-Hochberg (BH) multiple testing with an adjusted P -value ≤ 0.01 . This provides an accurate read-out of the net change in 3'UTR expression in a gene's transcript and can be used based on total RNA-seq without specialized poly(A) mapping.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

All sequencing data described are available on GEO under the accession number GSE140543. The source data underlying Figs. 1a–c, 2a–c, 5 and Supplementary Figs. 1a, b, 2, 5, 6, 7b, d are provided as a Source Data file. All data are available from the corresponding author upon reasonable request.

Code availability

Code for the analyses described in this study is available from the corresponding author upon request.

Received: 26 July 2019; Accepted: 11 December 2019; Published online: 07 January 2020

References

- Flavell, S. W. et al. Genome-wide analysis of MEF2 transcriptional program reveals synaptic target genes and neuronal activity-dependent polyadenylation site selection. *Neuron* **60**, 1022–1038 (2008).
- 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).
- Ji, Z. & Tian, B. Reprogramming of 3' untranslated regions of mRNAs by alternative polyadenylation in generation of pluripotent stem cells from different cell types. *PLoS ONE* **4**, e8419 (2009).
- Mayr, C. & Bartel, D. P. Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell* **138**, 673–684 (2009).
- Singh, I. et al. Widespread intronic polyadenylation diversifies immune cell transcriptomes. *Nat. Commun.* **9**, 1716 (2018).
- Lee, S. H. et al. Widespread intronic polyadenylation inactivates tumour suppressor genes in leukaemia. *Nature* **561**, 127–131 (2018).
- Oh, J. M. et al. U1 snRNP telescripting regulates a size-function-stratified human genome. *Nat. Struct. Mol. Biol.* **24**, 993–999 (2017).
- Kaida, D. et al. U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. *Nature* **468**, 664–648 (2010).
- Berg, M. G. et al. U1 snRNP determines mRNA length and regulates isoform expression. *Cell* **150**, 53–64 (2012).
- Matera, A. G. & Wang, Z. A day in the life of the spliceosome. *Nat. Rev. Mol. Cell Biol.* **15**, 108–121 (2014).
- Niibori, Y., Hayashi, F., Hirai, K., Matsui, M. & Inokuchi, K. Alternative poly(A) site-selection regulates the production of alternatively spliced vesl-1/homer1 isoforms that encode postsynaptic scaffolding proteins. *Neurosci. Res* **57**, 399–410 (2007).
- Matter, N. & Konig, H. Targeted 'knockdown' of spliceosome function in mammalian cells. *Nucleic Acids Res* **33**, e41 (2005).
- Li, W. et al. Systematic profiling of poly(A)+ transcripts modulated by core 3' end processing and splicing factors reveals regulatory roles of alternative cleavage and polyadenylation. *PLoS Genet.* **11**, e1005166 (2015).
- Masamha, C. P. et al. CFIm25 links alternative polyadenylation to glioblastoma tumour suppression. *Nature* **510**, 412–416 (2014).
- Wang, Q. & Rio, D. C. JUM is a computational method for comprehensive annotation-free analysis of alternative pre-mRNA splicing patterns. *Proc. Natl Acad. Sci. USA* **115**, E8181–E8190 (2018).
- Gilad, S. et al. Genotype-phenotype relationships in ataxia-telangiectasia and variants. *Am. J. Hum. Genet.* **62**, 551–561 (1998).
- Alsafadi, S. et al. Cancer-associated SF3B1 mutations affect alternative splicing by promoting alternative branchpoint usage. *Nat. Commun.* **7**, 10615 (2016).
- Seiler, M. et al. Somatic mutational landscape of splicing factor genes and their functional consequences across 33 cancer types. *Cell Rep.* **23**, 282–296 e4 (2018).
- Ntini, E. et al. Polyadenylation site-induced decay of upstream transcripts enforces promoter directionality. *Nat. Struct. Mol. Biol.* **20**, 923–928 (2013).
- Almada, A. E., Wu, X., Kriz, A. J., Burge, C. B. & Sharp, P. A. Promoter directionality is controlled by U1 snRNP and polyadenylation signals. *Nature* **499**, 360–363 (2013).
- Ogami, K. et al. An Mtr4/ZFC3H1 complex facilitates turnover of unstable nuclear RNAs to prevent their cytoplasmic transport and global translational repression. *Genes Dev.* **31**, 1257–1271 (2017).
- So, B. R. et al. A complex of U1 snRNP with cleavage and polyadenylation factors controls telescripting, regulating mRNA transcription in human cells. *Mol. Cell* **76**, 590–599 e4 (2019).
- Yoshida, K. et al. Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature* **478**, 64–69 (2011).
- Ilagan, J. O. et al. U2AF1 mutations alter splice site recognition in hematological malignancies. *Genome Res* **25**, 14–26 (2015).

25. Kim, E. et al. SRSF2 mutations contribute to myelodysplasia by mutant-specific effects on exon recognition. *Cancer Cell* **27**, 617–630 (2015).
26. Futreal, P. A. et al. A census of human cancer genes. *Nat. Rev. Cancer* **4**, 177–183 (2004).
27. Nicoloso, M. S., Spizzo, R., Shimizu, M., Rossi, S. & Calin, G. A. MicroRNAs—the micro steering wheel of tumour metastases. *Nat. Rev. Cancer* **9**, 293–302 (2009).
28. Nana-Sinkam, S. P. & Croce, C. M. Non-coding RNAs in cancer initiation and progression and as novel biomarkers. *Mol. Oncol.* **5**, 483–491 (2011).
29. Ebert, M. S. & Sharp, P. A. Roles for MicroRNAs in conferring robustness to biological processes. *Cell* **149**, 515–524 (2012).
30. Mendell, J. T. & Olson, E. N. MicroRNAs in stress signaling and human disease. *Cell* **148**, 1172–1187 (2012).
31. Johnson, C. D. et al. The let-7 microRNA represses cell proliferation pathways in human cells. *Cancer Res.* **67**, 7713–7722 (2007).
32. Lee, S. O. et al. MicroRNA15a modulates expression of the cell-cycle regulator Cdc25A and affects hepatic cystogenesis in a rat model of polycystic kidney disease. *J. Clin. Invest.* **118**, 3714–3724 (2008).
33. Wang, P. et al. microRNA-21 negatively regulates Cdc25A and cell cycle progression in colon cancer cells. *Cancer Res.* **69**, 8157–8165 (2009).
34. Paranjape, T. et al. A 3'-untranslated region KRAS variant and triple-negative breast cancer: a case-control and genetic analysis. *Lancet Oncol.* **12**, 377–386 (2011).
35. Milde-Langosch, K. The Fos family of transcription factors and their role in tumorigenesis. *Eur. J. Cancer* **41**, 2449–2461 (2005).
36. Silvestre, D. C., Gil, G. A., Tomasini, N., Bussolino, D. F. & Caputto, B. L. Growth of peripheral and central nervous system tumors is supported by cytoplasmic c-Fos in humans and mice. *PLoS ONE* **5**, e9544 (2010).
37. Takagaki, Y., Seipelt, R. L., Peterson, M. L. & Manley, J. L. The polyadenylation factor CstF-64 regulates alternative processing of IgM heavy chain pre-mRNA during B cell differentiation. *Cell* **87**, 941–952 (1996).
38. Di Giarmartino, D. C., Nishida, K. & Manley, J. L. Mechanisms and consequences of alternative polyadenylation. *Mol. Cell* **43**, 853–866 (2011).
39. de Klerk, E. et al. Poly(A) binding protein nuclear 1 levels affect alternative polyadenylation. *Nucl. Acids Res.* **40**, 9089–9101 (2012).
40. Martin, G., Gruber, A. R., Keller, W. & Zavolan, M. Genome-wide analysis of pre-mRNA 3' end processing reveals a decisive role of human cleavage factor I in the regulation of 3' UTR length. *Cell Rep.* **1**, 753–763 (2012).
41. Yao, C. et al. Transcriptome-wide analyses of CstF64-RNA interactions in global regulation of mRNA alternative polyadenylation. *Proc. Natl Acad. Sci. USA* **109**, 18773–18778 (2012).
42. Jenal, M. et al. The poly(a)-binding protein nuclear 1 suppresses alternative cleavage and polyadenylation sites. *Cell* **149**, 538–553 (2012).
43. Gruber, A. R., Martin, G., Keller, W. & Zavolan, M. Means to an end: mechanisms of alternative polyadenylation of messenger RNA precursors. *Wiley Interdiscip. Rev. RNA* **5**, 183–196 (2014).
44. Yu, K. et al. A precisely regulated gene expression cassette potently modulates metastasis and survival in multiple solid cancers. *PLoS Genet.* **4**, e1000129 (2008).
45. Park, H. J. et al. 3' UTR shortening represses tumor-suppressor genes in trans by disrupting ceRNA crosstalk. *Nat. Genet.* **50**, 783–789 (2018).
46. Cheng, Z. et al. Gene expression profiling reveals U1 snRNA regulates cancer gene expression. *Oncotarget* **8**, 112867–112874 (2017).
47. Henrich, K. O., Schwab, M. & Westermann, F. 1p36 tumor suppression—a matter of dosage? *Cancer Res.* **72**, 6079–6088 (2012).
48. Chiu, A. C. et al. Transcriptional pause sites delineate stable nucleosome-associated premature polyadenylation suppressed by U1 snRNP. *Mol. Cell* **69**, 648–663 e7 (2018).
49. Vorlova, S. et al. Induction of antagonistic soluble decoy receptor tyrosine kinases by intronic polyA activation. *Mol. Cell* **43**, 927–939 (2011).
50. Langemeier, J. et al. A complex immunodeficiency is based on U1 snRNP-mediated poly(A) site suppression. *EMBO J.* **31**, 4035–4044 (2012).
51. Tsujii, M. et al. Cyclooxygenase regulates angiogenesis induced by colon cancer cells. *Cell* **93**, 705–716 (1998).
52. Kramer, N. et al. In vitro cell migration and invasion assays. *Mutat. Res.* **752**, 10–24 (2012).
53. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
54. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* **5**, 621–628 (2008).
55. Feng, J. et al. GFOLD: a generalized fold change for ranking differentially expressed genes from RNA-seq data. *Bioinformatics* **28**, 2782–2788 (2012).

Acknowledgements

We thank members of our laboratory for helpful discussions and comments on the manuscript. This work was supported by the US National Institutes of Health (R01GM112923 to G.D.). G.D. is an Investigator of the Howard Hughes Medical Institute.

Author contributions

J.M.O., A.M.P., L.W., I.Y. and G.D. conceived and designed the project. J.M.O., A.M.P., L.W., I.Y., C.A., Z.C., B.R.S. and J.D. performed the experiments. C.C.V. and C.D. performed the informatics analyses. J.M.O., C.C.V., C.D. and G.D. wrote the manuscript with input from all the authors. G.D. was responsible for the project's performance.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41467-019-13993-7>.

Correspondence and requests for materials should be addressed to G.D.

Peer review information *Nature Communications* thanks Donald Rio, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permission information is available at <http://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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020