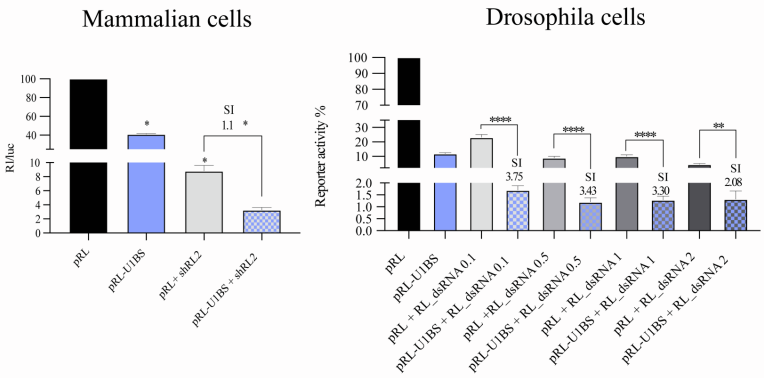
OMTN, Volume *28*

Supplemental information

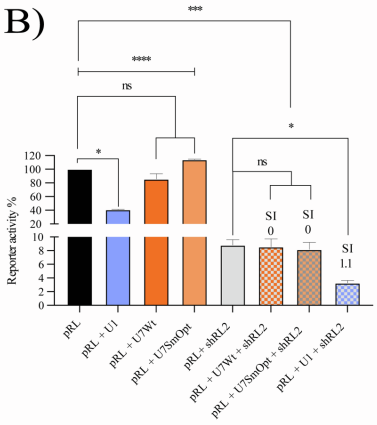
U1A is a positive regulator of the expression of heterologous and cellular genes involved in cell proliferation and migration

Eric Rovira, Beatriz Moreno, Nerea Razquin, Roland Hjerpe, Monika Gonzalez-Lopez, Rosa Barrio, Igor Ruiz de los Mozos, Jernej Ule, Fernando Pastor, Lorea Blazquez, and Puri Fortes

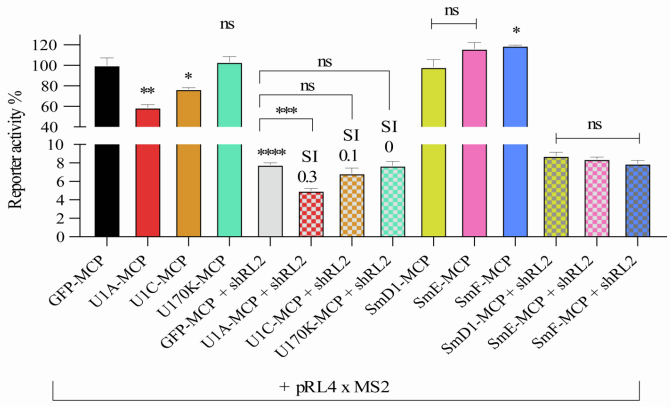
S1A)



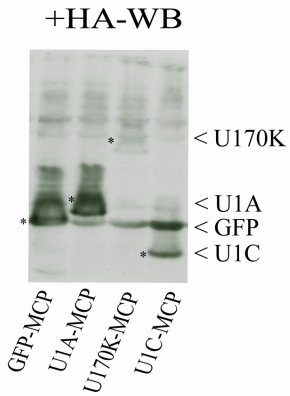
S1B)



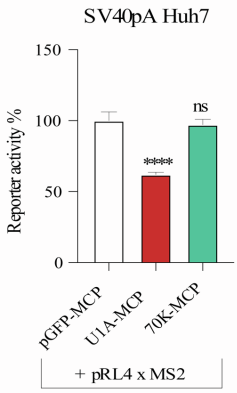
S1C)



S1D)



S1E)



S1F)

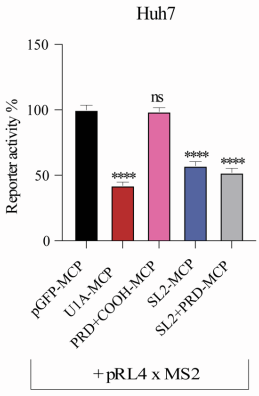
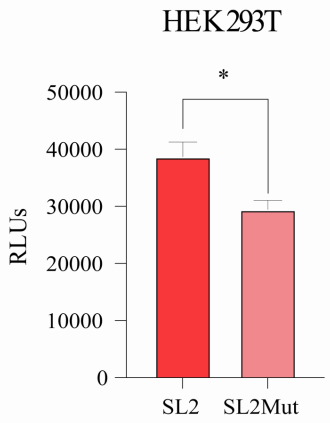
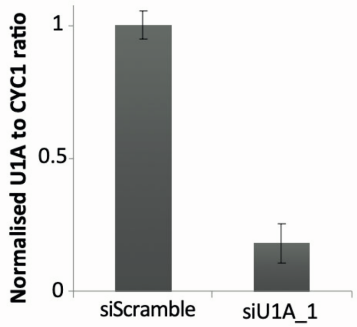


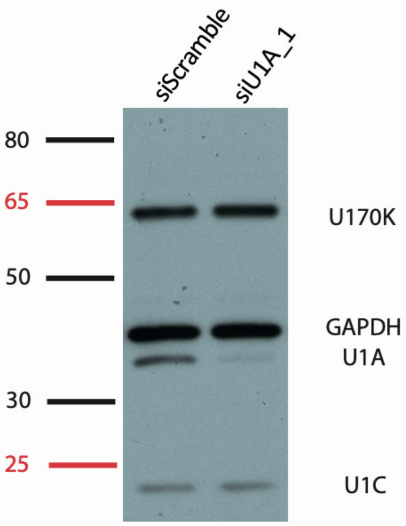
Figure S1. Related to Figure 1: Inhibition of gene expression by RNAi, U1i or U1 snRNP proteins. (A) As in Figure 1A but with an additional shRNA in HeLa cells or increasing doses of dsRNA in Fly cells. (B) As in figure 1B but with an additional shRNA (shRL2). (C) As in figure 1C but with an additional shRNA (shRL2). (D) Expression of the GFP, U1A, U170K or U1C-MCP fusion proteins was verified by Western-blot with an anti-HA antibody. The size of the expected proteins is indicated with an asterisk. (E) As in figure 1D but in Huh7 cells co-transfected with a plasmid expressing a R4LxMS2 plasmid with the SV40 canonical 3'end sequence and plasmids expressing the indicated MCP fused proteins. (F) As in D, but the RL4xMS2 plasmid was co-transfected with plasmids expressing MCP fused to U1A or U1A-truncated fragments. All experiments were performed at least three times in triplicates. Error bars show standard error of the mean. The synergy index (SI) was calculated as described[15]. Either two-tailed Student's t-test or one-way ANOVA were employed to compare two or more independent groups respectively. Non-significant differences are indicated with ns. Significant results are indicated as * p-value < 0.05, ** p-value < 0.01, *** p-value < 0.001 and **** p-value < 0.0001.

S2A)


HEK293T
*
50000
40000
30000
20000
10000
0
RLUs
SL2
SL2Mut

S2B)


Normalised U1A to CYC1 ratio
1
0.5
0
siScramble
siU1A_1


siScramble
siU1A_1
80
65
50
30
25
U170K
GAPDH
U1A
U1C

S2C)

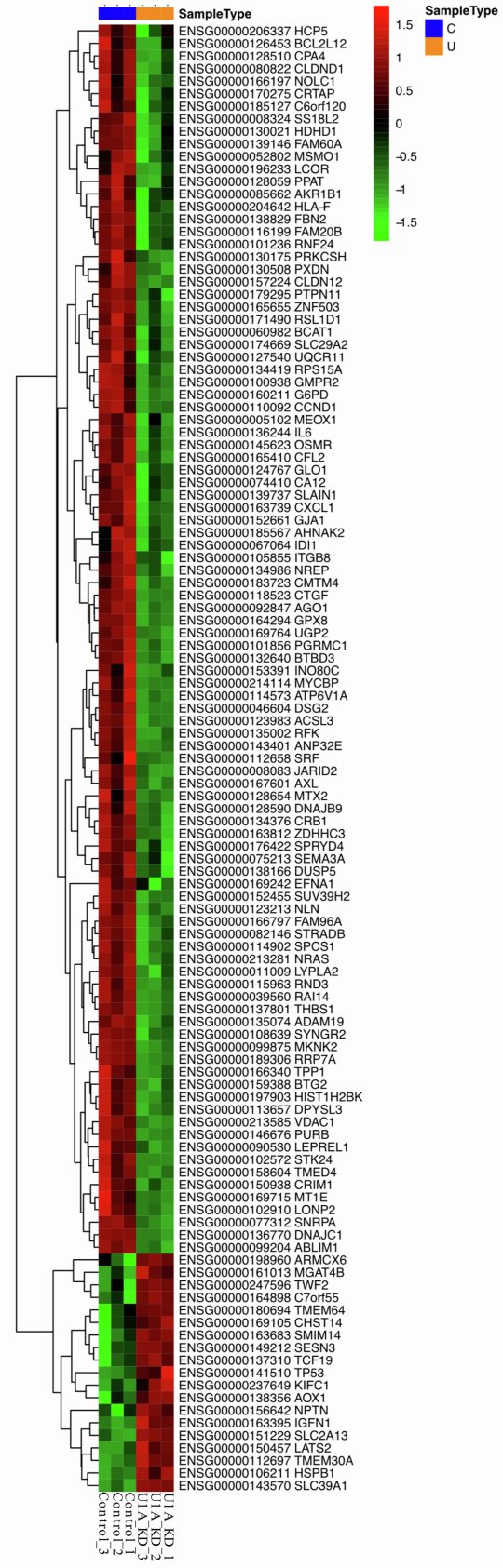
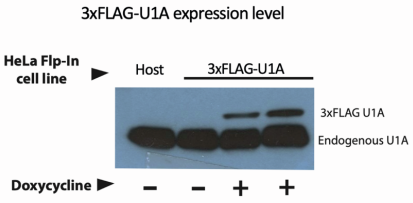
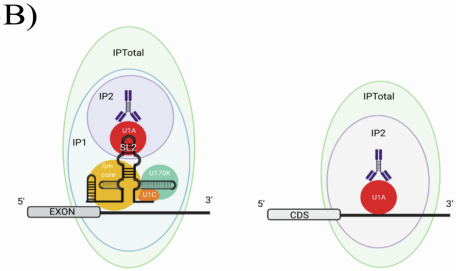

SampleType
C
U
1.5
1
0.5
0
-0.5
-1
-1.5
ENSG00000206337 HCP5
ENSG00000126453 BCL2L12
ENSG00000128510 CPA4
ENSG00000080822 CLDND1
ENSG00000166197 NOLC1
ENSG00000170275 CRTAP
ENSG00000185127 C6orf120
ENSG00000008324 SS18L2
ENSG00000130021 HDHD1
ENSG00000139146 FAM60A
ENSG00000052802 MSMO1
ENSG00000196233 LCOR
ENSG00000128059 PPAT
ENSG00000085662 AKR1B1
ENSG00000204642 HLA-F
ENSG00000138829 FBN2
ENSG00000116199 FAM20B
ENSG00000101236 RNF24
ENSG00000130175 PRKCSH
ENSG00000130508 PXDN
ENSG00000157224 CLDN12
ENSG00000179295 PTPN11
ENSG00000165655 ZNF503
ENSG00000171490 RSL1D1
ENSG00000060982 BCAT1
ENSG00000174669 SLC29A2
ENSG00000127540 UQCR11
ENSG00000134419 RPS15A
ENSG00000100938 GMPR2
ENSG00000160211 G6PD
ENSG00000110092 CCND1
ENSG00000005102 MEOX1
ENSG00000136244 IL6
ENSG00000145623 OSMR
ENSG00000165410 CFL2
ENSG00000124767 GLO1
ENSG00000074410 CA12
ENSG00000139737 SLAIN1
ENSG00000163739 CXCL1
ENSG00000152661 GJA1
ENSG00000185567 AHNAK2
ENSG00000067064 IDI1
ENSG00000105855 ITGB8
ENSG00000134986 NREP
ENSG00000183723 CMTM4
ENSG00000118523 CTGF
ENSG00000092847 AGO1
ENSG00000164294 GPX8
ENSG00000169764 UGP2
ENSG00000101856 PGRMC1
ENSG00000132640 BTBD3
ENSG00000153391 INO80C
ENSG00000214114 MYCBP
ENSG00000114573 ATP6V1A
ENSG00000046604 DSG2
ENSG00000123983 ACSL3
ENSG00000135002 RFK
ENSG00000143401 ANP32E
ENSG00000112658 SRF
ENSG00000008083 JARID2
ENSG00000167601 AXL
ENSG00000128654 MTX2
ENSG00000128590 DNAJB9
ENSG00000134376 CRB1
ENSG00000163812 ZDHHC3
ENSG00000176422 SPRYD4
ENSG00000075213 SEMA3A
ENSG00000138166 DUSP5
ENSG00000169242 EFNA1
ENSG00000152455 SUV39H2
ENSG00000123213 NLN
ENSG00000166797 FAM96A
ENSG00000082146 STRADB
ENSG00000114902 SPCS1
ENSG00000213281 NRAS
ENSG00000011009 LYPLA2
ENSG00000115963 RND3
ENSG00000039560 RAI14
ENSG00000137801 THBS1
ENSG00000135074 ADAM19
ENSG00000108639 SYNGR2
ENSG00000099875 MKNK2
ENSG00000189306 RRP7A
ENSG00000166340 TPP1
ENSG00000159388 BTG2
ENSG00000197903 HIST1H2BK
ENSG00000113657 DPYSL3
ENSG00000213585 VDAC1
ENSG00000146676 PURB
ENSG00000090530 LEPREL1
ENSG00000102572 STK24
ENSG00000158604 TMED4
ENSG00000150938 CRIM1
ENSG00000169715 MT1E
ENSG00000102910 LONP2
ENSG00000077312 SNRPA
ENSG00000136770 DNAJC1
ENSG00000099204 ABLIM1
ENSG00000198960 ARMCX6
ENSG00000161013 MGAT4B
ENSG00000247596 TWF2
ENSG00000164898 C7orf55
ENSG00000180694 TMEM64
ENSG00000169105 CHST14
ENSG00000163683 SMIM14
ENSG00000149212 SESN3
ENSG00000137310 TCF19
ENSG00000141510 TP53
ENSG00000237649 KIFC1
ENSG00000138356 AOX1
ENSG00000156642 NPTN
ENSG00000163395 IGFN1
ENSG00000151229 SLC2A13
ENSG00000150457 LATS2
ENSG00000112697 TMEM30A
ENSG00000106211 HSPB1
ENSG00000143570 SLC39A1
U1A_KD_1
U1A_KD_2
U1A_KD_3
Control_1
Control_2
Control_3

Figure S2. Related to Figure 2: Free U1A binds target genes and is a positive regulator of gene expression. (A) As in Figure 2A but in HEK293T cells. (B) U1A mRNA or U1 snRNP specific proteins were measured by qRT-PCR (top) or Western Blot (bottom) in HeLa cells transfected with control or U1A-targeting siRNAs. Note that that depletion of U1A does not affect the levels of U170K or U1C. (C) Heatmap showing the results of the Log2 FC expression changes from the Volcano plot in Figure 2B. The colour code gradient is indicated to the right.

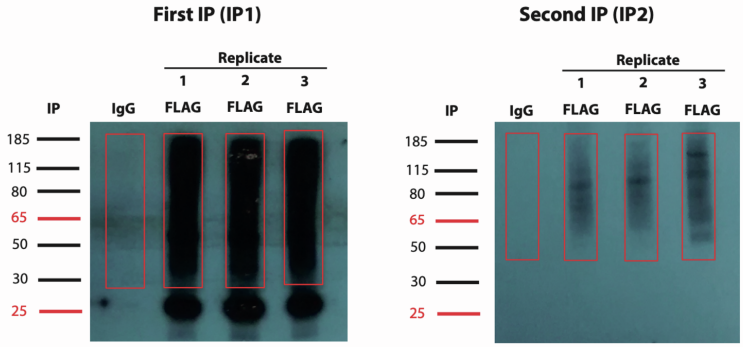
S3A)



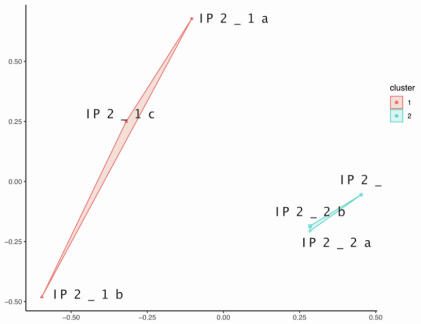
S3B)



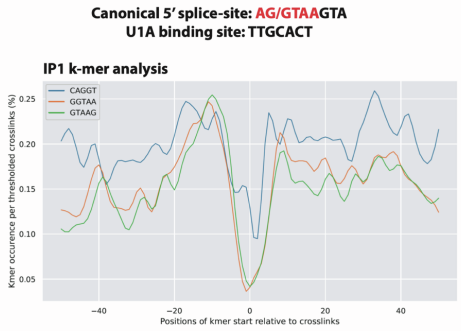
S3C)



S3D)



S3E)



S3F)

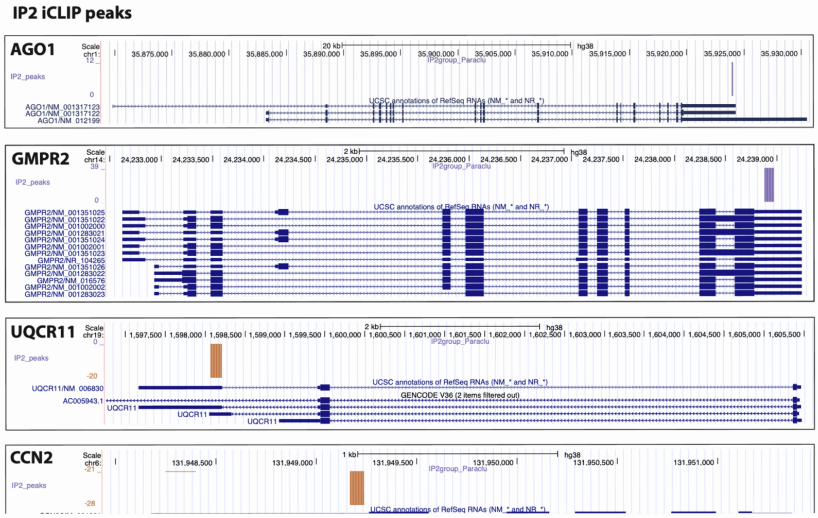


Figure S3. Related to Figure 2: Free U1A binds target genes and is a positive regulator of gene expression. (A) Endogenous and 3xFLAG-U1A expression level by Western-Blot with an anti-U1A antibody in host or 3xFLAG-U1A HeLa Flp-In stable cells. The expression of 3XFLAG-U1A is inducible by doxycycline. (B) Schematic representation of the U1 snRNP factors that may coimmunoprecipitate after one (IP1) or two (IP2) rounds of immunoprecipitation with FLAG antibody in 3XFLAG-U1A Hela Flp-In cell line. (C) Radioactive gel visualization of protein-RNA complexes separated by SDS-PAGE and transferred into a nitrocellulose membrane after one (IP1) or two (IP2) rounds of immunoprecipitation with IgG or FLAG antibody in 3XFLAG-U1A Hela Flp-In cell line. (D) PCA plot obtained with the iCLIP peak scores for each gene identified in 3 replicates of IP1 and IP2 iCLIP libraries. (E) Graph with the k-mer sequence enriched in IP1 iCLIP data, which shows U1 snRNP consensus sequence. (F) Visualization of U1A binding regions in UCSC obtained from IP2 iCLIP peaks.

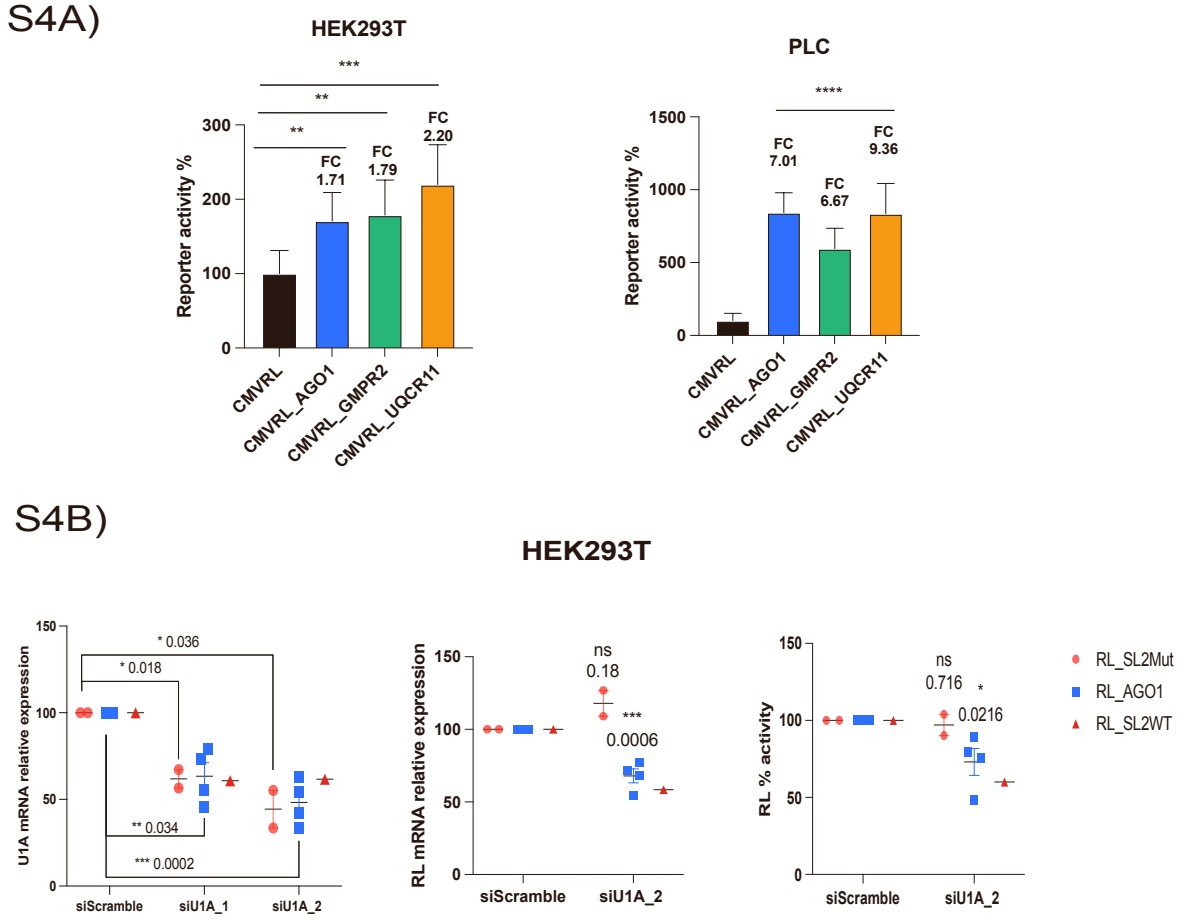


Figure S4. Related to Figure 3: U1 binding sequences increase reporter expression in 293T and PLC cells. (A) As in Figure 3C but in HEK293T and PLC cells (B) As in Figure 3D but in HEK293T clones transfected with an additional U1A-targeting siRNA (siU1A_2). The left graph represents U1A mRNA expression level obtained by qPCR in each HEK293T clone transfected with scramble or U1A siRNAs. Each dot represents the mean of three replicates performed on the same clone. The activity of RL was normalized to that of the control for each clone to estimate the relative RL activity

S5)

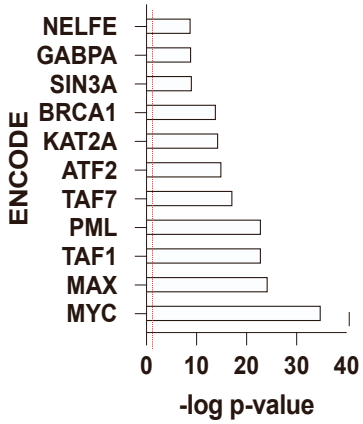


Figure S5. Related to Figure 4: U1A mRNA levels associate with cancer prognosis and cell growth. (A) Bar plot representing the enrichment analyses for ENCODE transcription factors in the list of genes that overlap in Figure 4F (U1A-bound genes from IP2 iCLIP data and genes whose expression positively associates (R>0.5) with U1A mRNA levels in TCGA LIHC data).

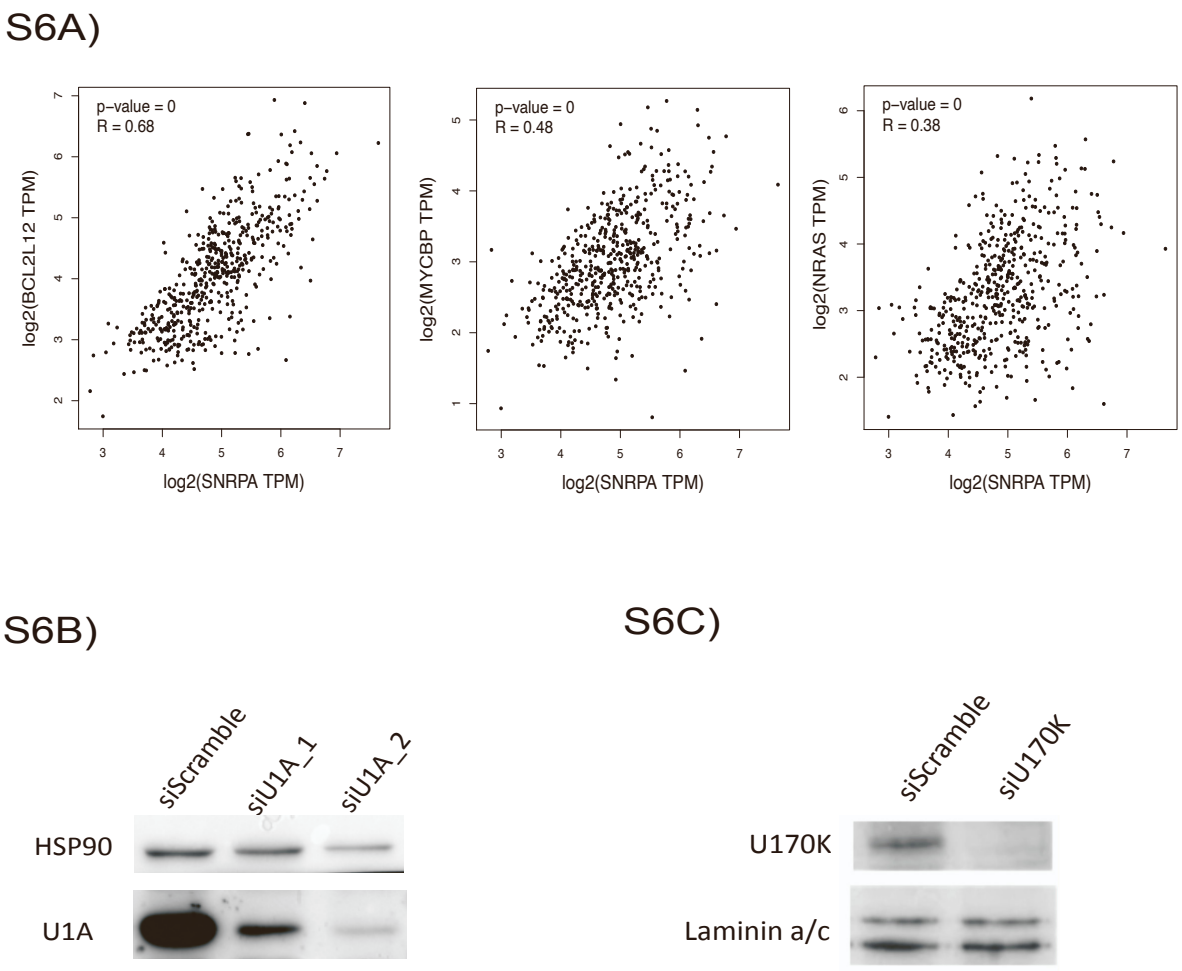
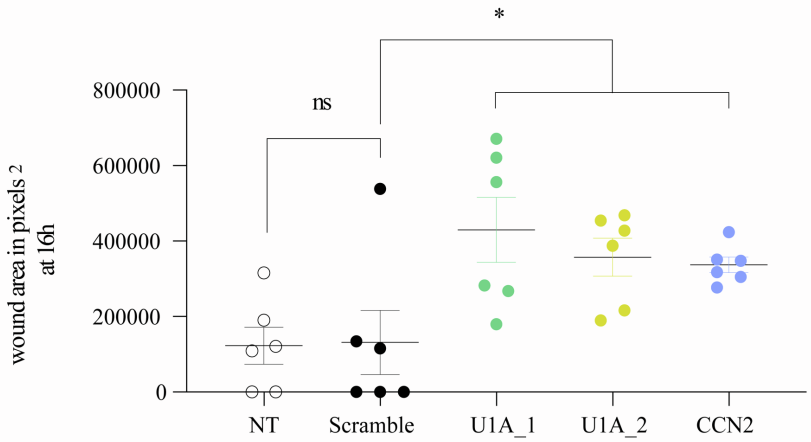


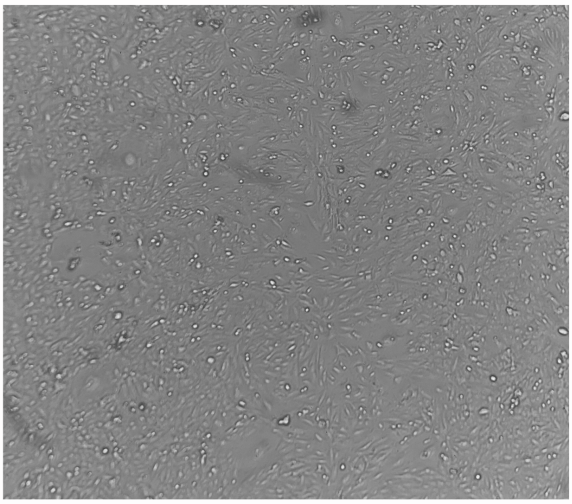
Figure S6. Related to figure 5: U1A upregulates several oncogenes and promotes cell migration. (A) Association between U1A and BCL2L12, NRAS or MYCBP mRNA levels in LIHC using TCGA data. All analyses have been performed using GEPIA2. (B) U1A protein levels obtained by Western Blot in JHH6 cells after inhibition with two different U1A targeting siRNAs. HSP90 was used as a loading control. (C) U170K protein levels obtained by Western Blot in JHH6 cells after transfection with scramble or U170K targeting siRNAs. Lamina/b was used as a loading control.

S7A)

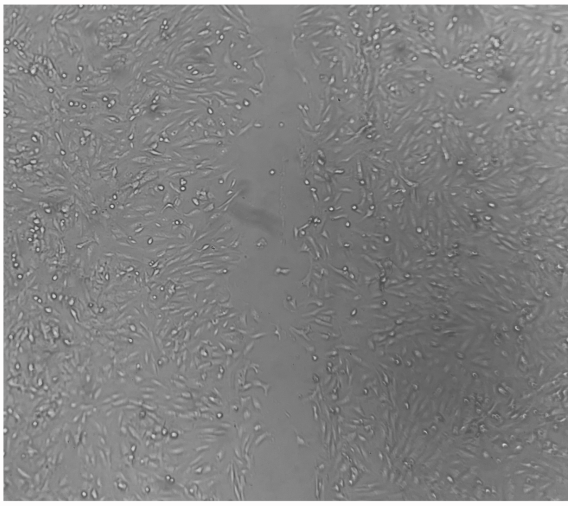


S7B)

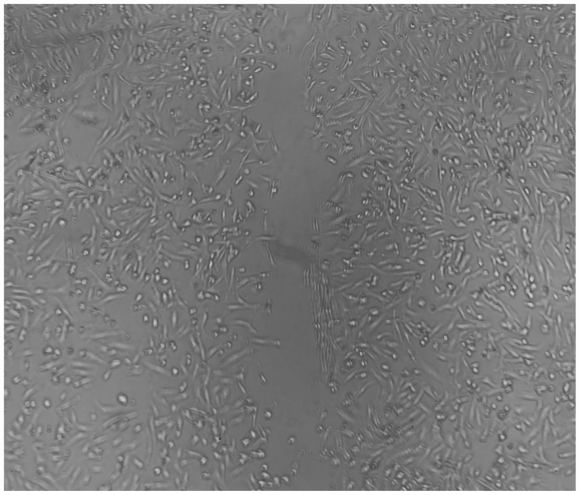
SiScramble



SiU1A_1



SiU1A_2



SiCCN2

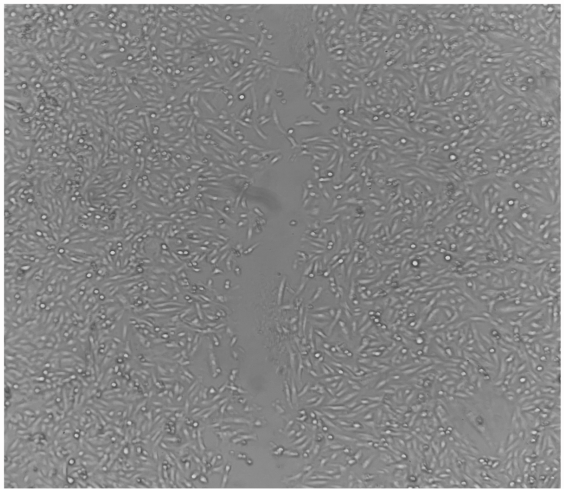


Figure S7. Related to Figure 5: U1A knockdown affects cell migration. JHH6 cells were transfected with two independent siRNAs targeting U1A, an siRNA targeting CCN2 and a scramble siRNA. 24h after transfection the wound healing assay was set on M12 well plates. (A) Quantification with FIJI image software analysis of the cell-depleted area in squared pixels of images taken16 hours after the wound was performed. Each dot plotted represents a replicate. Error bars show standard error of the mean. One-way ANOVA was employed to compare the means of all siRNA-transfected cells. Two tailed Student's t test was employed to compare the not transfected (NT) and Scramble means. Ns indicates not significant, * indicates p-value <0.05. (B) Representative images taken under bright field microscope 16 hours after the wound was performed.

Table S1: Genes deregulated after U1A depletion

Table S2: U1A targets identified by iCLIP (IP2)

Table S3: Nucleic acid sequences used in this study:

pCMV-RL plasmid constructs:

Name	Sequence (5’ -> 3’)
pCMV-RL-pL-WT-U1BS polylinker	GCTAGCtcgcgAAGCTTgtcgacTCTAGAtgcaattcgtggCAGGTAAGTATaattcgGAATTCagatctCTCGAGatatcGCTAGC
pCMV-RL-pL-Mut-U1BS polylinker	GCTAGCtcgcgAAGCTTgtcgacTCTAGAtgcaattcgtggCAtGgAAcTATaattcgGAATTCagatctCTCGAGatatcGCTAGC
pCMV-RL-pL-4xMS2	TCTAGATACGGTACTTATTGCCAAGAAAGCACGAGCATCAGCCGTGCCTCCAGGTCGAATCTTCAAACGACGACGATCACGCGTCGCTCCAGTATTCCAGGGTTCATCAGATCCGGTAACCTACGGTACTTATTGCCAAGAAAGCACGAGCATCAGCCGTGCCTCCAGGTCGAATCTTCAAACGACGACGATCACGCGTCGCTCCAGTATTCCAGGGTTCATCAGATCCGAATTC
MALAT1 3’end sequence	TCTAGACAGTAGGGTCATGAAGGTTTTTCTTTTCCTGAGAAAACAACACGTATTGTTTTCTCAGGTTTTGCTTTTTGGCCTTTTTCTAGCTTAAAAAAAAAAAAAGCAAAAGATGCTGGTGGTTGGCACTCCTGGTTTCCAGGACGGGGTTCAAATCCCTGCGGCGTCTTTGCTTTGACTACTAATCTGTCTTCAGGACTCTTTCTGTATTTCTCCTTTTCTCTGCAGGTGCTAGTTCTTGGAGTTTTGGGGAGGTGGGAGGTAACAGCACAATATCTTTGAACTATATACATCCTTGATGTATAATTTGTCAGGAGCTTGACTTGATTGTATATTCATATTTACACGAGAACCTAATATAACTGCCTTGTCTTTTTCAGGTAATAGCCTGCAGCTGGTGTTTTGAGAAGCCCTACTGCTGAAAACTTAACAATTTTGTGTAATAAAAATGGAGAAGCTCTAAAGGATCC
SL2 WT sequence	Fw: CTAGTTCGACATCGTCGCGATTGCACTCCCGCGACGGTCGACAC Rv: AATTGTGTCGACCGTCGCGGGAGTGCAATCGCGACGATGTCGAA
SL2 Mut sequence	Fw: CTAGTTCGACATCGTCGCGAGGATCCTCCCGCGACGGTCGACAC Rv: AATTGTGTCGACCGTCGCGGGAGTGCAATCGCGACGATGTCGAA
3´UTR of UQCR11	TCTAGAGGTGCTAGCTCCTCCCACCTCAGCACCTGCTGCATCTGGAGCAGCCCAAGCCTCAGGATGGACAAGAGGAAACCCACAGCTCAGCTTCAGGCTTCTTATGTTTCTGAAAACAGCTTGGATATTTTAATGCACGTTGCATTAAAGAATTC
3’UTR of GMPR2	TCTAGATAGACCTGAGCAGTTCTACCCTCCCAAGGCACCAGTACTCTACCATGGGGCATCCCAAGTGGGGTCCTCACCCATCCCAGCTACTGCAGCTCTGTATTACTTTGTCATTTCCTGTTGTCTCACTCCTGAGGGCTCCTGCAGTAACTCTGTACTTCTCTATCTGCACACACAAAATGCCCAAGGCACTCACTGGGGAGGAAGCAAGGAAGCAAACAGTCTGAGAAAATGATGCAAGAAAATCAAATGGGAATCTGGGGACCCAACACAACATCCTGAAGATTATTAAAAGGAAAAGATGCTGATTGGTACATAAATCTTTTACATGGCCTTGGaggcctGGAGGCAGGCTTTTAGAATCATGTTTTGTTAATCCGCTTCACTAAATTGGACCTTCACATATCTAAAAAGCTCTGAAGTGTTTGTATATTTGAAATACCTCAATAAAGAATTC

3’UTR of AGO1	TCTAGAAATTACTTTCCTGTGCACACAGTCCAGCCTAATTGGTATGTGATGTTGC ACTTAGCAGCCATGTGGTGGGCATGTGTGACTACTCTGGTTTTCACTTTAGTTTC TAAACTTTTTATCCCTCTCAAGTCCAGCATGGATGGGGAAATGTCTCTGGtaCCC CACAGCTGTGTACTTGTTTGCATTTGTTTCCCTTTGAGATTTGTGTTTGTGTCCTG CTTTGAGCTGTACCTTGTCCAGTCCATTGTGAAATTATCCCAGCAGCTGTAATGT ACAGTTCCTTCTGAAGCAAGCAACATCAGCAGCAGCAGCAGCAGCACAAT TCTGTGTTTTATAAAGAATTC

Plasmids expressing U snRNAs:

Name	Sequence (5’ -> 3’)
U7aU1MUT-SmWT	AATAGTTCCATGGCAAATTTGTCTAGTAGGCTTTCTGGCTTTTTACC
3’U7aU1MUT-SmU1	AATAGTTCCATGGCAAATTTTTGGAGTAGGCTTTCTGGCTTTTTACC

Tethering oligos:

Name	Sequence (5’ -> 3’)
U1A-tether	Fw: GGATCCACCGGTCGCCACCATGGCAGTTCCCGAGACCCG Rv: TGCAGAATCGATCTACTTCTTGGCAAAGGAGA
U1C-tether	Fw: GGATCCACCGGTCGCCACCATGCCCAAG TTTTATTGTGA Rv: TGCAGAATCGATTTATCTGTCTGGTCGAGTCA
U170K-tether	Fw:GGATCCACCGGTCGCCACCATGACCCAGTTCCTGCCGCC Rv: TGCAGAATCGATTCACTCCGGCGCAGCCTCCA
AgeINotI-SmB	F: ATCCACCGGTTGGCGGCCGCATGACGGTGGGCAAGAGCAG R: CAGAAGATCTATCGATCTAGGGCCTTGGTGGGCGCA
XmaINotI-SmD1	F: ATCCCCCGGGTGGCGGCCGCATGAAGCTCGTGAGATTTTT R: CAGAAGATCTATCGATTTATCGCCTAGGACCCCCTC
XmaINotI-SmD2	F: ATCCCCCGGGTGGCGGCCGCATGAGCCTCCTCAACAAGCC R: CAGAAGATCTATCGATCTACTTGCCGGCGATGAGCG
XmaINotI-SmD3	F: ATCCCCCGGGTGGCGGCCGCATGTCTATTGGTGTGCCGAT R: CAGAAGATCTATCGATTTATCTTCGCTTTTGAAAGA
XmaINotI-SmE	F: ATCCCCCGGG TGGCGGCCGCATGGCGTACCGTGGCCAGGG R: CAGAAGATCTATCGATC TAGTTGGAGACACTTTGTA
XmaINotI-SmF	F: ATCCCCCGGGTGGCGGCCGCATGAGTTTACCCCTCAATCC R: CAGAAGATCTATCGATCTATTCTCTCATTTCC CCAT
XmaINotI-SmG	F: ATCCCCCGGGTGGCGGCCGCATGAGCAAAGCTCACCCTCC R: CAGAAGATCTATCGATTTATACTCGTTCCAAGGCTT
U1A-101-ClaI	R: CAGAATCGATTTA GAAGGTGCCTTTCATCTTGGC
U1A-115-ClaI	R: CAGAATCGATTTAGCTCTTGG GCTTCCTCTTCT
U1A-102-AgeI	F: ATCCACCGGTCGCCACCATGGTGGAGCGGGACCGCAAG
PRD-Wt	F: CCGGAGCGGGACCGCAAGCGGGAGAAGAGGAAGCCCAAGAGCTAAA R: TCGATTTAGCTCTTGGGCTTCCTCTTCTCCCGCTTGCGGTCCCGCT

In-fusion primers:

Name	Sequence (5’ -> 3’)
SBSfiT7F	GTGTCGTGAAAACTACCCCAAGCTGGCCTTGAGGCCTAATACGACTCACTATAGG
NotCMVRLSfipASBR	CTCTAGAGAATTGATCCCCAAGCTTGGCCTGACAGGCCGTCTGCTCGAAGCGGCCGC

Primers to generate dsRNA for *Drosophila melanogaster*:

Name	Sequence (5’ -> 3’)
T7RLF	TAATACGACTCACTATAGGGAGAGATAACTGGTCCGCAGTGGT
T7RLR	TAATACGACTCACTATAGGGAGACAACATGGTTTCCACGAAGA

RNA interference (RNAi):

Name	Sequence (5’ -> 3’)
siU1A_1	CCUUUAAGACUUACCUCAA
siU1A_2	GAGGUUUGGUUUUUCACAA
siU170K	CCUUUAAGACUUACCUCAA
siScramble	*SIC001 - Merck*
shRL1	S: GATCCCCGAAAGTTTATGATCCAGAATTCAAGAGATTCTGGATCATAAACTTTCTTTTTGGAAA AS: AGCTTTTCCAAAAAGAAAGTTTATGATCCAGAATCTCTTGAATTCTGGATCATAAACTTTCGGG
shRL2	S: GATCCCCGCATCAAGATAAGATCAAATTCAAGAGATTTGATCTTATCTTGATGCTTTTTGGAAA AS: AGCTTTTCCAAAAAGCATCAAGATAAGATCAAATCTCTTGAATTTGATCTTATCTTGATGCGGG

qRT-PCR primers:

Name	Sequence (5’ -> 3’)
U1A_Fw	ACCCGCCCTAACCACACTAT
U1A_Rv	GGCTCCGTGATACCAGGATA
AGO1_Fw	AGATCAAAGTCTGGGCCATC
AGO1_Rv	GCAGCCCTGAGTAGGTGTTC
GMPR2_Fw	CAGCAAACTCTTGCCACTGA
GMPR2_Rv	CAGCAAACTCTTGCCACTGA

UQCR11_Fw	GGTACCCAGTCCAGGATCAG
UQCR11_Rv	TACCGGGAGCTGGTCAAG
RL_Fw	GCAAATCAGGCAAATCTGGT
RL_Rv	CCATTCATCCCATGATTCAA
U170K_Fw	TCCGGAGAATGGGTATTTGA
U170K_Rv	GGACAAACTCAAGTGGCCAA
BCL2L12_Fw	AGACCGCAAGTTGAGTGGAGGA
BCL2L12_Rv	AGCCTCACCACGCCTAAGGAAG
CCN2_Fw	ACCTGTGGGATGGGCATCTC
CCN2_Rv	CGGATGCACTTTTTGCCCTTC
MYCBP_Fw	AGAGCTGCTTCGCCTAGAACTG
MYCBP_Rv	ATTCAGCACGCTTCTCCTCCTG
NRAS_Fw	GAAACCTCAGCCAAGACCAGAC
NRAS_Rv	GGCAATCCCATACAACCCTGAG
RRAS_Fw	GATCTGCCTTGTTCCCGACCAA
RRAS_Rv	CTGCTGGTGTTCGCCATTAACG
FN1_Fw	CTGCGAGAGTAAACCTGAA
FN1_Rv	TCCCAGATCATGGAGTCTTT
CDH2_Fw	ATGTGCATGAAGGACAGCCT
CDH2_Rv	CATGCCATCTTCATCCACCT
RPLP0_Fw	CCATTCATCCCATGATTCAA
RPLP0_Rv	ACTCAGGATTTCAATGGTGCC