IWS1 phosphorylation promotes cell proliferation and predicts poor prognosis in *EGFR* mutant lung adenocarcinoma patients, through the cell cycle-regulated *U2AF2* RNA splicing.

Georgios I. Laliotis[1,2,3,12*], Evangelia Chavdoula[1,2], Maria D. Paraskevopoulou[4,13], Abdul D. Kaba[1,2], Alessandro La Ferlita[1,2,5], Satishkumar Singh[6], Vollter Anastas[1,2,7], Salvatore Alaimo[5], Arturo Orlacchio[1,2], Keith A. Nair II[1,2], Vasiliki Taraslia[4,14], Ioannis Vlachos[8,15], Marina Capece[1,2], Artemis Hatzigeorgiou[8], Dario Palmieri[1,2], Christos Tsatsanis[9,10], Lalit Sehgal[6], David P. Carbone[11], Vincenzo Coppola[1,2] and Philip N. Tsichlis[1,2,16*]

[1]The Ohio State University, Department of Cancer Biology and Genetics, Columbus, OH, 43210, USA, [2]The Ohio State University Comprehensive Cancer Center-Arthur G. James Cancer Hospital and Richard J. Solove Research Institute, Columbus, OH, 43210, USA, [3]University of Crete, School of Medicine, Heraklion Crete, 71500, Greece, [4]Molecular Oncology Research Institute, Tufts Medical Center, Boston, MA, 02111, USA, [5]Department of Clinical and Experimental Medicine, Bioinformatics Unit, University of Catania, Catania, 95131, Italy, [6]College of Medicine, Department of Hematology, The Ohio State University, Columbus, OH 43210, [7]Tufts Graduate School of Biomedical Sciences, Program in Genetics, Boston, MA, 02111, USA, [8]DIANA-Lab, Hellenic Pasteur Institute, Athens, 11521, Greece, [9]Department of Clinical Chemistry-Biochemistry, School of Medicine, University of Crete, Heraklion 71110, Crete, Greece, [10]Institute for Molecular Biology and Biotechnology, Foundation for Research and Technology Hellas, Heraklion 70013, Greece, [11]Department of Internal Medicine, Division of Medical Oncology, The Ohio State University Medical Center, Columbus, OH, 43210, USA

Running Title: IWS1 and *U2AF2* RNA splicing in EGFR mutant lung adenocarcinoma.

Present Address :

[12]Sidney Kimmel Comprehensive Cancer Center and Department of Oncology, Johns Hopkins School of Medicine, Baltimore, MD 21205 USA

[13]Takeda Pharmaceuticals U.S.A, Boston, MA, 02139, USA

[14]Center of Basic Research, Biomedical Research Foundation of the Academy of Athens, 11527, Greece

[15]Beth Israel-Deaconess Medical Center, Department Of Pathology, Harvard Medical School, Boston, MA 02215, USA

[15]Lead Contact

*Correspondence should be addressed to: Philip N. Tsichlis (Philip.tsichlis@osumc.edu) and Georgios I. Laliotis (glaliot7@gmail.com)

Abstract

Our previous studies have shown that IWS1 (Interacts with Spt6) is a phosphorylation target of AKT and regulates the alternative RNA splicing of *FGFR2*, linking IWS1 with human Non-Small Cell Lung Cancer. To further address the role of IWS1 in alternative RNA splicing in lung cancer, we performed an RNA-seq study using lung adenocarcinoma cells in which IWS1 was knocked down or replaced by its phosphorylation site mutant. The results identified a novel, exon 2 deficient splice variant of the splicing factor U2 Associated-Factor 2 (*U2AF2*), whose abundance increases, upon the loss of phosphorylated IWS1. This exon encodes part of the U2AF65 Serine-Rich (RS) Domain, which is required for its binding with pre-mRNA Processing factor 19 (Prp19). Here, we show that *U2AF2* exon 2 inclusion depends on phosphorylated IWS1, by promoting Histone H3K36 trimethylation and the assembly of LEDGF/SRSF1 splicing complexes, in a cell-cycle specific manner. Inhibition of the pathway results in the downregulation of cell cycle division associated 5 (*CDCA5*), a phosphorylation target and regulator of ERK, leading to G2/M phase arrest, impaired cell proliferation and tumor growth in mouse xenografts models, an effect more pronounced in *EGFR* mutant cells. Analysis of lung adenocarcinoma samples revealed strong correlations between IWS1 phosphorylation, *U2AF2* RNA splicing, and Sororin/p-ERK levels, especially in *EGFR*, as opposed to *KRAS* mutant patients. More importantly, IWS1 phosphorylation and *U2AF2* RNA splicing pattern are positively correlated with tumor stage, grade, relapse and metastasis, and associated with poor survival in lung adenocarcinoma patients, harboring *EGFR*, but not *KRAS*, mutations. This work highlights the instrumental role of the AKT/p-IWS1 axis to alternative RNA splicing in governing cell cycle progression and tumorigenesis, and proposes this axis as a novel drug target in *EGFR* mutant lung adenocarcinoma, by concomitantly affecting the epigenetic regulation of RNA processing and oncogenic signals.

63 **Introduction**

64 Cells use a plethora of molecular mechanisms to precisely regulate mRNA expression
65 (Glisovic T. et al., 2008[1]). These co- and post-transcriptional processes are orchestrated by a
66 complex RNA-binding protein network. (Jewer M. et al., 2012[2]). One of the main mechanisms
67 of this regulation is alternative RNA splicing. Most human genes harbor introns which are
68 removed during pre-mRNA splicing (Pan Q. et al., 2008[3]), a common mechanism employed
69 by eukaryotic cells to generate multiple transcripts and expand their functional network of the
70 generated protein products. Thus, regulation of alternative RNA splicing affects cellular fate
71 and function and controls the pathobiology of many diseases, including cancer (Paronetto
72 M.P. et al., 2016[4], Zhang X. et al., 2016[5]). Its importance has been demonstrated in different
73 types of cancer, including non-small cell lung carcinoma (NSCLC) (Coomer A.O. et al., 2019[6]).

74 Non-small cell lung carcinoma (NSCLC) is the 2nd most common cancer, with more
75 than 250,000 new cases per year in the US and extremely poor prognosis with <5% 5-year
76 survival (Siegel et al., 2020[7]). NSCLC is further divided into lung adenocarcinoma, squamous
77 cell carcinoma and large cell carcinoma based on histological features. (Yuan M et al, 2019[8]).
78 Notably, up to 69% of NSCLCs harbor mutations in the Epidermal Growth Factor Receptor
79 (*EGFR*) and Kirsten Rat Sarcoma (*KRAS*) genes, affecting major downstream oncogenic
80 signal such as the PI3K/AKT and ERK/MAPK pathway (Sequist et al., 2020[9]).

81 Following our original observations linking the alternative RNA splicing of *FGFR2* with
82 NSCLC (Sanidas et al., 2014[10]), several additional splicing events have been linked to the
83 biology of NSCLC including Bcl-X_L, CD44, Androgen Receptor (AR), HLA-G and PKM, by
84 promoting cell survival, metastasis, chemoresistance, immune surveillance and metabolic
85 advantage, respectively (Li Z. et al., 2016[11], Todaro M. et al., 2014[12], Oltean S. et al., 2014[13],
86 Calabretta S. et al., 2016[14]). Furthermore, another mechanism utilized by lung cancer cells
87 involves the re-programming of alternative RNA splicing through the introduction of de novo
88 mutations which create new splice sites and alternatively spliced isoforms (Vo N.S. et al.,
89 2018[15], Giaj Levra M. et al., 2014[16]). This knowledge has led to an ongoing phase I clinical

90 trial to determine the effects of a modulator of the core splicing function, H3B-8800, in cancer
91 patients (Seiler M., 2018[17]). Thus, deeper comprehension of the molecular regulation of RNA
92 splicing in lung cancer, and other cancer types, is a necessity for proper design of anti-tumor
93 therapies.

94 Mechanistically, alternative RNA splicing is regulated co-transcriptionally and through
95 chromatin modifications in the body of transcribed genes. When the rate of transcription is low,
96 it increases the probability for exons that are not efficiently spliced, to be spliced out of the
97 mature transcript (Listerman et al., 2012[18]). Chromatin modifications are recognized by
98 readers of epigenetic marks, which orchestrate the assembly of molecular complexes that bind
99 to, and functionally regulate RNA-associated enhancers or repressors of splicing (*cis*-acting
100 RNA sequences). This alters the rate of assembly and the composition of spliceosomal
101 complexes (*trans*-acting RNA Binding Proteins-RBPs) (Luco et al., 2010[19], Pradeepa et al.,
102 2012[20], Oltean S. et al., 2014[10]). Cancer cells utilize this machinery in order to manipulate
103 alternative RNA splicing and promote tumorigenesis. Well studied families of RBPs include
104 the Serine/Rich (SR proteins) and heterogeneous nuclear ribonucleoproteins (hNRNPs).
105 (Chen M. et al., 2009[21]). SR proteins bind RNA exonic splicing enhancers (ESE) and intronic
106 splicing enhancers (ISE) and usually promote exon inclusion. (Obeng et al., 2019[22]). On the
107 contrary, hnRNPs bind to exonic splicing silencers (ESS) and intronic splicing silencers (ISS)
108 and in most cases lead to exon skipping. (Martinez-Contreras et al., 2007[23]). Several reports
109 have shown the role of these RBPs in the regulation of alternative RNA splicing and tumor
110 growth in cancer cells, including Serine Rich Specific Factor 1 (SRSF1) and Polypyrimidine
111 Tract Binding Protein (PTBP1) (de Miguel et al., 2014[24], Chen M et al., 2012[25], Jin W et al.,
112 2003[26]).

113 One cellular function, which interfaces with the RNA splicing machinery, is the cell
114 cycle regulation (Dominguez et al., 2016[27]). Progression through the cell cycle depends on
115 periodic changes of RNA metabolism, which can be achieved by multiple mechanisms, one of
116 which is the periodic modulation of RNA splicing (Dominguez et al., 2016[27]). On the other
117 hand, alterations in RNA splicing, are regulated by periodic shifts in the expression and the

118 activity of known cell cycle regulators, such as the Aurora kinases (Moore et al., 2010[28]), the
119 large RS domain-containing protein SON (Ahn et al., 2011[29]), the RNA recognition motif
120 (RRM)-containing protein TgRRM1 (Suvorova et al, 2013[30]) and CLK1 (Dominguez et al.,
121 2016[27]), demonstrating the bidirectional relationship of these cellular functions.

122 We have previously shown that the transcription elongation factor IWS1 and the AKT3
123 kinase play a central role in the regulation of the alternative RNA splicing of *FGFR2*, by
124 promoting the exclusion of exon IIIb. This exclusion event depends on the phosphorylation of
125 IWS1 by AKT3 at Ser720/Thr721. IWS1 binds SPT6, which is attached to the phosphorylated
126 Ser2 of the C-terminal domain (CTD) of RNA polymerase II. SETD2, a Histone H3 trimethyl-
127 transferase, binds to the IWS1/Spt6 complex only when IWS1 is phosphorylated at
128 Ser720/Thr721. SETD2 bound to the phosphorylated and CTD-associated IWS1,
129 trimethylates histone H3 at K36 during transcription. The Histone H3K36me3 marks in the
130 body of actively-transcribed target *FGFR2* gene are recognized by MRG15, which interacts
131 with PTB leading to exclusion of exon IIIb in the mature *FGFR2* mRNA transcript, affecting
132 cell invasion, migration, cell proliferation and tumor growth in mouse xenografts models. More
133 importantly, we had shown robust expression of IWS1 phosphorylation in 21 out of 24 NSCLC
134 samples, which correlates with AKT phosphorylation, introducing the role of IWS1 in lung
135 cancer (Sanidas et al., 2014[10]).

136 Here, to gain insight on the role of IWS1 phosphorylation in the pathogenesis of lung
137 cancer and the epigenetic regulation of alternative RNA splicing genome-wide, we performed
138 RNA-seq studies in lung adenocarcinoma cells NCI-H522, in which IWS1 was knocked down
139 or replaced by its phosphorylation site mutant. The results revealed a novel RNA splice variant
140 of the splicing factor U2 Associated Factor A2 (*U2AF2*), which lacks exon 2 upon loss of
141 phosphorylated IWS1. U2AF65, its protein product, as a splicing factor binds to the
142 Polypyrimidine (Py)-tract/3' splice site and initiates spliceosome assembly, by promoting the
143 interaction between U2 snRNP and the branchpoint (Shen H. et al., 2004[31]). Several reports

144 have illustrated the role of U2AF65 in lung cancer. Specifically, cancer-associated mutations
145 of *U2AF2* in lung cancer patients alter the affinity of its RNA-Binding Domain with Py-tract
146 (Glasser E. et al., 2017[32]). Furthermore, it has been shown that stabilization of U2AF65 by
147 OTUB2 promotes Warburg effect and tumorigenesis in NSCLC, pointing out the importance
148 of this splicing factor in NSCLC. (Li J. et al., 2018[33]). In this report, our goal was to further
149 investigate the molecular mechanisms that control exon inclusion phenotype induced by IWS1
150 phosphorylation, to characterize this novel variant and to highlight its significance in lung
151 carcinogenesis.

152 **Results**

153 **IWS1 expression and phosphorylation regulate alternative mRNA splicing.**

154 We have previously reported that IWS1 phosphorylation at Ser 720/Thr721, primarily
155 by AKT3, resulted in the exclusion of exon IIIb of *FGFR2* in the human NSCLC cell lines NCI-
156 H522 and NCI-H1299 (Sanidas et al., 2014[10]). To explore the molecular mechanisms driving
157 the IWS1 phosphorylation-dependent RNA splicing and gene expression, we performed RNA-
158 Seq to examine the transcriptome of shControl, shIWS1, shIWS1/wild type IWS1 rescue
159 (shIWS1/WT-R) and shIWS1/phosphorylation site IWS1 mutant rescue (shIWS1/MT-R) NCI-
160 H522 cells. To further address the role of IWS1 phosphorylation in the regulation of RNA
161 processing, alongside with these samples, we also engineered shIWS1/IWS1
162 phosphomimetic mutant S720D/T721E rescue (shIWS1/IWS1 DE MT-R) NCI-H522 and NCI-
163 H1299 cells (Sanidas et al., 2014[10]), which were used to validate the putative results of the
164 RNA-seq experiment. Notably, all these rescue clones are engineered to be shIWS1-resistant
165 (Sanidas et al., 2014[10]). First, we confirmed the significant down-regulation of IWS1
166 expression after the transduction of the cells with lentiviral shIWS1 construct and the rescue
167 of IWS1 knock-down with the Flag-tagged wild type, phosphorylation IWS1 site mutant and
168 phosphomimetic IWS1 mutant (**Fig. 1A**). Differential expression analysis of the RNA-seq data,
169 identified 1,621 and 562 differentially expressed genes between shControl and shIWS1 and

170 between shIWS1/WT-R and shIWS1/MT-R cells, respectively (p≤0.01, FDR ≤0.2). 340 genes
171 were detected as differentially expressed in both comparisons, following the same direction in
172 their expression (**Fig. S1A, S1B**). Moreover, 19 out of the FDR-ranked top 100 differentially
173 expressed genes, in shControl versus shIWS1 cells, were also differentially expressed in
174 shIWS1/WT-R versus shIWS1/MT-R, with all of them being regulated similarly in both
175 comparisons (up- or down-regulated) (**Fig. S1C**). Gene Set Enrichment analysis
176 (Subramanian et al., 2005[34]), revealed significant enrichment of genes involved in RNA
177 metabolism and regulation of RNA processing (**Fig. S1D**).

178 Using an unbiased approach, we performed differential exon usage analysis of the
179 RNA-seq data, DEXseq (Anders et al., 2012[35]), to determine the effect of IWS1 expression
180 and phosphorylation on RNA splicing. This analysis identified 1,434 (corresponding to 851
181 genes) and 436 (corresponding to 273 genes) differentially employed exons between
182 shControl versus shIWS1 and shIWS1/WT-R versus shIWS1/MT-R cells, respectively (p
183 ≤0.05). The 1,796 differentially expressed genes and the 692 genes with differential exon
184 usage in shIWS1 versus shControl cells, exhibited an overlap of 165 genes (p≤0.05). Similarly,
185 the 858 differentially expressed genes and the 230 genes with differential exon usage,
186 between shIWS1/MT-R and shIWS/WT-R cells revealed an overlap of 44 genes (**Fig. 1B**).

187 Our earlier studies have shown that the IWS1 expression and phosphorylation promote
188 exon exclusion in the *FGFR2* gene (Sanidas et al., 2014[10]). Therefore, we analysed the exon
189 usage data, and we observed that the most common event associated with the expression
190 and phosphorylation of IWS1, was exon inclusion (**Fig. 1C).** Functional analysis of the IWS1-
191 regulated alternative spliced genes, focused on GO-biological processes, identified significant
192 enrichment of genes involved in RNA processing and RNA splicing, among the top biological

193 processes regulated by IWS1-dependent alternative RNA splicing (**Fig. 1D, 1E**). It is
194 remarkable that the results were similar with the results generated from the functional analysis
195 of the IWS1-dependent differential gene expression. Therefore, these findings imply that the
196 effect of IWS1 on RNA processing may be direct or indirect. The indirect effect may be due to
197 the IWS1 expression and phosphorylation-dependent differential regulation of genes involved
198 in RNA processing.

199 Validation of the RNA-seq, using RT-PCR, showed several examples of alternative
200 RNA splicing of genes which are regulated by IWS1 and IWS1-phosphorylation, via exon
201 inclusion (**Fig. 1, S1G-S1J**). One of these events is the inclusion of exon 2 in the mature
202 mRNA transcript of *U2AF2*, the gene encoding the splicing factor U2AF65. Whereas the
203 predominant *U2AF2* mRNA transcript in shControl and shIWS1/WT-R cells contains exon 2,
204 the predominant transcript in shIWS1 and shIWS1/MT-R cells, is a novel transcript lacking
205 exon 2. To validate these results, we examined the ratio of exons 2 and 3 (E2/E3 *U2AF2*) by
206 RT-PCR in NCI-H522 and NCI-H1299 shControl, shIWS1, shIWS1/WT-R, shIWS1/MT-R and
207 shIWS1/DE MT-R cells. The results revealed a decrease in the E2/E3 ratio in shIWS1 and
208 shIWS1/MT-R cells, relative to shControl cells, which was rescued in shIWS1/DE MT-R (**Fig.
209 1F, 1G**). The decrease in the E2/E3 ratio in shIWS1 and shIWS1/MT-R, relative to shControl
210 cells, and the subsequent rescue of the ratio in shIWS1/DE MT-R cells, was confirmed by
211 quantitative RT-PCR (**Fig. S1E upper panel**). In parallel experiments, qRT-PCR in the same
212 cells revealed increased IIIb/IIIc *FGFR2* ratio in shIWS1 and shIWS1/MT-R cells, relative to
213 shControl cells, which was rescued in shIWS1/DE MT-R, confirming our previous findings on
214 the effect of IWS1 phosphorylation on *FGFR2* alternative RNA splicing (Sanidas et al., 2014[10])
215 (**Fig. S1E lower panel).** Notably, the knockdown of IWS1 or the rescue with the non-
216 phosphorylatable site mutant (shIWS1/MT-R) did not significantly change the total expression
217 of *U2AF2* or the inclusion of *U2AF2* exon 3 in NCI-H522 and NCI-H1299 cells. (**Fig. S1F**).

218 To establish whether IWS1 is directly involved in the molecular complex that regulates
219 *U2AF2* RNA splicing, in NCI-H522 and NCI-H1299 cells transduced with shControl, shIWS1,
220 shIWS1/WT-R and shIWS1/MT-R, we performed chromatin immunoprecipitation (ChIP)

221 assays. The results revealed that both IWS1 WT and S720A/T721A bind equally well to exon
222 2 and exon 3 of *U2AF2* (**Fig. 1H**). Therefore, IWS1 phosphorylation controls *U2AF2* exon 2
223 alternative RNA splicing by regulating events occurring after the binding of IWS1 to chromatin.

224 **IWS1 phosphorylation-dependent mRNA splicing of *U2AF2* is regulated by serum and**
225 **IGF-1 via AKT3.**

226 IWS1 is phosphorylated by AKT3, and to a lesser extent by AKT1, at Ser720/Thr721
227 (Sanidas et al., 2014[10]). Since the phosphorylation and activity of the AKT isoforms can be
228 induced by tyrosine-kinase signals, we questioned whether IGF-1 stimulation of serum starved
229 NCI-H522 and NCI-H1299 cells, promotes *U2AF2* exon 2 inclusion along with the expected
230 AKT activation and IWS1 phosphorylation. The results showed that the E2/E3 *U2AF2* and
231 IIIb/IIIc *FGFR2* ratio, which was used as a control, indeed parallels the activation of AKT and
232 the subsequent phosphorylation of IWS1 (**Fig. 2A, S2A**). To address whether the *U2AF2* exon
233 2 inclusion depends on AKT activation, we treated NCI-H522 and NCI-H1299 cells growing in
234 serum-containing complete media with 5 μM of the pan-AKT inhibitor MK2206, a dose that
235 fully inhibits all AKT isoforms (Sanidas et al., 2014[10]) .The results confirmed that MK2206
236 inhibits both AKT (T308 and S473) and IWS1 phosphorylation (S720) and *U2AF2* exon 2
237 inclusion (**Fig. 2B, S2B upper panel).** Consistently, the same treatment promoted the
238 inclusion of exon IIIb in *FGFR2* mRNA (**Fig. S2B lower panel**). To determine whether it is the
239 AKT3 isoform, which is responsible for the observed effects of AKT inhibition on the *U2AF2*
240 alternative RNA splicing, we transduced NCI-H522 and NCI-H1299 cells with shAKT1,
241 shAKT2 and shAKT3 along with shControl, and we examined its effects on the alternative
242 RNA splicing of IWS1 targets, *U2AF2* exon 2 and *FGFR2* exon IIIb, using RT-PCR and qRT-
243 PCR. Interestingly, the results showed that inhibition of AKT1 and AKT3 phenocopied the
244 treatment with MK2206 (**Fig. 2C, S2C)**. Notably, the effect of AKT1 inhibition was less
245 pronounced, as expected, given that IWS1 is primarily phosphorylated by the AKT3 kinase.
246 These data provide evidence that the physiological regulation of alternative RNA splicing of

247 *U2AF2* by external stimuli depends on the phosphorylation of IWS1 by AKT3, and to a lesser
248 extent, by AKT1.

249 ***U2AF2* Exon 2 inclusion, induced by IWS1 phosphorylation at Ser720/Thr721, depends**
250 **on histone H3K36 trimethylation by SETD2.**

251 We have previously reported that IWS1 phosphorylation by AKT promotes the
252 exclusion of exon IIIb from the mature *FGFR2* mRNA transcript, via a process that depends
253 on Histone H3K36 trimethylation by SETD2, and that the latter is recruited to the CTD of RNA
254 Pol II by phosphorylated IWS1 (Sanidas et al., 2014[10]). To determine whether the *U2AF2* exon
255 2 inclusion phenotype is also dependent on Histone H3K36 trimethylation, we performed ChIP
256 assays in shControl, shIWS1, shIWS1/WT-R and shIWS1/MT-R NCI-H522 and NCI-H1299
257 cells, addressing the abundance of H3K36me3 marks on exons 2 and 3 of *U2AF2*, as well as
258 on the *U2AF2* transcriptional start site (TSS) and the *GAPDH* gene, which were used as
259 controls. The results revealed that the knockdown of IWS1 or the replacement with the
260 phosphorylation-site mutant significantly decreased the abundance of H3K36me3 on *U2AF2*
261 exon 2 and 3 (**Fig. 3A**). Parallel ChIP assays showed the decreased H3K36me3 abundance
262 on *FGFR2* exons IIIb and IIIc in the same cells, which was used as control (**Fig. S3A**). More
263 importantly, inhibition of AKT, using MK2206, phenocopied the H3K36me3 pattern of
264 phosphorylation-deficient IWS1 mutant (**Fig. 3B**). Since IWS1-phosphorylation recruits
265 SETD2 in C-terminal domain of RNA-pol II, we examined the binding of SETD2 on *U2AF2,*
266 *FGFR2* and *GAPDH* gene in the same NCI-H522 and NCI-H1299 cells, transduced using a
267 SETD2 lentiviral construct, with a Hemagglutinin (HA)-tag. The results showed a pattern of
268 SETD2 binding in both IWS1 splicing targets *U2AF2* and *FGFR2*, which parallels the IWS1
269 phosphorylation-dependent pattern of H3K36me3 abundance (**Fig. 3C, Fig. S3B**), suggesting
270 that H3K36 trimethylation is mediated by SETD2.

271 The preceding data suggested that the enzymatically active SETD2 contributes to the
272 IWS1 phosphorylation-dependent regulation of the *U2AF2* alternative RNA splicing. The role

273 of catalytically-active SETD2 was confirmed by SETD2 knockdown and rescue experiments
274 in NCI-H522 and NCI-H1299 cells. Using RT-PCR and qRT-PCR, we observed that the
275 knockdown of SETD2 phenocopies the knockdown of IWS1 on the *U2AF2* and *FGFR2* RNA
276 splicing, an effect rescued by the wild type SETD2, but not by the SETD2 methyltransferase
277 mutant R1625C, which globally silences H3K36me3 (Hacker et. al., 2016[36]) (**Fig. 3D, Fig.**
278 **S3C**). We conclude that the catalytically-active SETD2 is indeed required for the inclusion of
279 exon 2 in the mature *U2AF2* mRNA.

280 Recently, and after the completion of the ChIP experiments, we carried out ChIP-Seq
281 experiments, addressing the binding of IWS1 and SETD2 and the distribution of H3K36me3
282 marks genome-wide in NCI-H522 shIWS1/WT-R and shIWS1/MT-R cells. The unbiased data
283 on the abundance of these markers in the *U2AF2* gene were in general agreement with the
284 ChIP data described above. Specifically, IWS1 was found to bind *U2AF2* E2 independent of
285 its phosphorylation, but SETD2 binding and H3K36me3 abundance on *U2AF2* E2 increased
286 only when IWS1 was phosphorylated (**Fig. 3E**). Furthermore, the same analysis confirms our
287 previous results, by showing the increased H3K36me3 abundance on *FGFR2* exon IIIb only
288 when IWS1 was phosphorylated (**Fig. S3D**) (Sanidas et al., 2014[10]).

289 If the recruitment of SETD2 in C-terminal domain of RNA-pol II depends on the
290 phosphorylation of IWS1, as our data indicate, one would expect that ectopic overexpression
291 of SETD2 would fail to rescue the *U2AF2* alternative RNA splicing phenotype in shIWS1 and
292 shIWS1/MT-R cells. This was confirmed by experiments addressing the rescue of *U2AF2*
293 alternative RNA splicing in NCI-H522 and NCI-H1299 shCon, shIWS1, shIWS1/WT-R and
294 shIWS1/MT-R cells transduced with a SETD2 lentiviral construct, with a Hemagglutinin (HA)-
295 tag (**Fig. S4A**). The failure of SETD2 to rescue the phenotype supports the model of SETD2
296 recruitment by phosphorylated IWS1.

297 However, histone methylation is a dynamic step-by-step process. (Greer et al., 2012[37])
298 and while SETD2 is the only known H3K36 trimethyl-transferase in mammalian cells (Hyun K.

299 et al., 2017[38]), there are several Lysine methyltransferases, which catalyze mono- or di-
300 methylation of Histone H3 at K36 and may influence the SETD2 output. Transfection of the
301 NCI-H522 and NCI-H1299 cells with siRNAs targeting a set of methyltransferases that are
302 known to catalyze Histone H3K36 mono and di-methylation (NSD1, NSD2 and NSD3), or only
303 di-methylation (SMYD2 and ASHL1) (Lucio-Eterovic et al., 2010[339], Q. Qiao et al., 2010[40],
304 Berdasco et al., 2009[41] , Rahman et al., 2011[42], Brown M.A et al., 2006[43], Gregory G.D. et al.,
305 2007[44]), revealed no contribution to the regulation of the alternative RNA splicing of *U2AF2*
306 (**Fig. S4B**).

307 Conversely, Histone methylation is a reversible process (Bannister et al., 2002[45]). The
308 Histone H3K36me3 marks are erased by KDM4A and KDM4C, two members of the KDM4
309 JmjC domain-histone demethylase family (Berry, W.L et al., 2013[46], Cloos, P.A et al., 2007[47],
310 Li, W et al., 2011[48]). Ectopic expression of KDM4A, KDM4B and KDM4C using lentiviral
311 constructs in NCI-H522 and NCI-H1299 cells showed that none of these three altered the
312 IWS1 phosphorylation-dependent alternative RNA splicing pattern of *U2AF2* (**Fig. S4C**).
313 Overall, these data indicate that the reported demethylation of H3K36me3 by members of
314 KDM4/JMJD2 Histone demethylase family does not play a significant role in our model.

315 **The regulation of the alternative RNA splicing of the *U2AF2* exon 2 by IWS1**
316 **phosphorylation, depends on the p52 isoform of the H3K36me3 reader LEDGF.**

317 Based on our preceding studies, the regulation of the *FGFR2* alternative RNA splicing
318 by IWS1 phosphorylation depends on the reading of the Histone H3K36me3 marks by MRG15
319 (Sanidas et al., 2014[10]). To determine whether MRG15 is also the reader of the IWS1-
320 dependent alternative RNA splicing of *U2AF2*, we knocked down MRG15 in both NCI-H522
321 and NCI-H1299 cells. Using RT-PCR and qRT-PCR to monitor the alternative splicing of
322 *U2AF2* in these cells, revealed that it is independent of MRG15 (**Fig. 4A Left panels**). In
323 agreement with this result, the knockdown of the splicing repressor and binding partner of
324 MRG15, PTB, also had no effect on the RNA splicing of *U2AF2* (**Fig. 4A Right panels**), results

325 not surprising since IWS1 phosphorylation mediates exon inclusion signals in our model. In
326 parallel quantitative RT-PCR experiments, the knockdown of MRG15 and PTB in NCI-H522
327 and NCI-H1299, increased the *FGFR2* IIIb/IIIc, as expected (**Fig. S5A**) (Luco et al., 2010[19],
328 Sanidas et al., 2014[10]). To further validate these observations, we knocked down PTBP2 in
329 NCI-H522 and NCI-H1299, a paralog of PTB, that has been reported to rescue PTB functions
330 (Cheung et al., 2009[49]). First, the western blotting in NCI-H522 and NCI-H1299 shPTB cells
331 showed no upregulation of PTBP2 upon loss of PTB (**Fig. 4A**). Furthermore, quantitative RT-
332 PCR results revealed that the regulation of *U2AF2* RNA splicing is independent of PTBP2,
333 with an expected increase of *FGFR2* IIIb/IIIc ratio (**Fig. S5B**).

334 To identify the factor that controls the IWS1 phosphorylation-dependent exon inclusion
335 of *U2AF2* exon 2, we transfected NCI-H522 and NCI-H1299 cells with siRNAs of the known
336 H3K36me3 readers PHF1 (Cai L. et al., 2013[50]), BRPF1 (Vezzoli et al., 2010[51]), MSH6 (Li F.
337 et al., 2014[52]), GLYR-1 (Vermuelen et al., 2010[53]) and LEDGF (Pradeepa et al., 2012[20]), along
338 with a control siRNA. After validating the significant downregulation of the expression by
339 Western blotting, monitoring the effects of these transfections on E2/E3 *U2AF2* ratio using
340 RT-PCR and qRT-PCR, revealed that only the knockdown of LEDGF phenocopied the
341 knockdown of IWS1 on the alternative RNA splicing of *U2AF2*, suggesting that LEDGF is the
342 sole H3K36me3 reader responsible for the observed phenotype (**Fig. 4B, S5C**). To confirm
343 this conclusion and to determine which isoform of LEDGF may be responsible for the
344 phenotype, we used a lentiviral shRNA construct to knockdown LEDGF, targeting the 3' UTR,
345 and we rescued the knockdown by transducing the cells with lentiviral constructs of the
346 alternatively spliced LEDGF isoforms, p75 and p52, with a V5-tag (**Fig. 4C, upper panel**).
347 Monitoring the effects of these transductions by RT-PCR and qRT-PCR, revealed that only
348 the p52 isoform rescues the shLEDGF-induced U2AF2 alternative RNA splicing phenotype.
349 More importantly, the A51P mutant of p52-LEDGF, which cannot bind Histone H3K36me3
350 (Shun et al., 2008[54]), failed to rescue the splicing defect, suggesting that the *U2AF2* exon
351 inclusion depends on the reading of the H3K36me3 marks by p52/LEDGF (**Fig. 4C, S5D**).

352 Notably, the knockdown of LEDGF did not affect the *FGFR2* alternative RNA splicing pattern
353 (**Fig. S5E**), suggesting that although the IWS1 phosphorylation-dependent RNA splicing
354 targets are regulated via H3K36me3, the complexes recruited upon IWS1 phosphorylation,
355 differ.

356 The preceding data provide strong genetic evidence that p52/LEDGF regulates the
357 *U2AF2* alternative RNA splicing by reading the IWS1 phosphorylation-dependent Histone
358 H3K36me3 marks. To validate these observations, chromatin Immunoprecipitation assays for
359 the V5-tagged p52/LEDGF isoform, confirmed that the p52/LEDGF binds *U2AF2* exons 2 and
360 3 only when IWS1 is phosphorylated, observations consistent with SETD2 binding and
361 H3K36me3 abundance (**Fig. 4D, S5F**).

362 If the p52 isoform of LEDGF regulates the alternative RNA splicing of *U2AF2* by
363 reading the Histone H3K36me3 mediated by IWS1-phosphorylation, as suggested by the
364 preceding data, one would expect that overexpression of p52/LEDGF would not rescue the
365 *U2AF2* alternative RNA splicing phenotype in shIWS1 and shIWS1/MT-R cells. This was
366 confirmed by experiments addressing the *U2AF2* RNA splicing in NCI-H522 and NCI-H1299
367 shControl, shIWS1, shIWS1/WT-R and shIWS1/MT-R cells along with a V5-tagged
368 p52/LEDGF construct (**Fig. S5G**). Altogether, these data support the model based on which
369 p52/LEDGF regulates the alternative RNA splicing of *U2AF2* by reading the Histone H3K36
370 trimethylation on the body of *U2AF2* gene, induced by IWS1 phosphorylation at S720/T721,
371 mediated by SETD2.

372 **The p52 isoform of LEDGF regulates the alternative RNA splicing of *U2AF2*, via its**
373 **interaction with the RNA splicing factor SRSF1.**

374 It has been reported that the p52/LEDGF isoform re-localizes in the nucleus, through
375 signals mediated from its unique C-terminal domain, and interacts with the splicing factor
376 SRSF1, affecting its distribution to alternatively spliced genes. (Pradeepa et al., 2012[20]). To

377 investigate a potential role of SRSF1 in *U2AF2* alternative RNA splicing, we knocked down
378 SRSF1 and we showed that its loss phenocopies the knockdown of IWS1. The dependence
379 of the *U2AF2* exon 2 splicing on SRSF1, which was suggested by this result, was confirmed
380 by rescue experiments with wild type SRSF1 (**Fig. 4E, S6A**). Consistent with the LEDGF
381 results, quantitative RT-PCR experiments showed that the knockdown of SRSF1 did not affect
382 the RNA splicing pattern of *FGFR2* (**Figure S6A, right lower panel**). We therefore conclude
383 that SRSF1 regulates the alternative RNA splicing of the *U2AF2* exon 2. More importantly,
384 SRSF1 did not rescue the *U2AF2* RNA splicing phenotype induced by shIWS1 or shIWS1/MT-
385 R (**Fig. S6B**), suggesting that it does not function independently. Instead, we hypothesized
386 that it provides the link between IWS1 phosphorylation-dependent chromatin modifications
387 and p52/LEDGF binding to the RNA splicing machinery. To test this hypothesis, we performed
388 ChIP experiments addressing the binding of SRSF1 to *U2AF2* exons 2 and 3 in shControl,
389 shIWS1, shIWS1/WT-R and shIWS1/MT-R NCI-H522 and NCI-H1299 cells, transduced with
390 a V5-tagged SRSF1 lentiviral construct. The *U2AF2* TSS and GAPDH exon 3 regions were
391 again used as controls. The results confirmed that the binding of SRSF1 to exons 2 and 3
392 parallels the SETD2 and p52/LEDGF binding along with H3K36me3 abundance and occurs
393 only when IWS1 is phosphorylated (**Fig. 4F, S6C**), providing support to the proposed
394 hypothesis.

395 Based on the preceding data, we hypothesize that the binding of SRSF1 to the
396 chromatin-associated p52/LEDGF should bring it into proximity with the nascent pre-mRNA,
397 facilitating their interaction. Analysis of the U2AF2 mRNA sequence, using the web-based
398 pipeline RBP-map (Paz et al., 2014[55]), identified four potential SRSF1 binding sites (2 in
399 *U2AF2* exon 2 and 2 in exon 3) (**Fig. S6E**), providing additional support to this hypothesis.
400 Furthermore, similar analysis of the validated IWS1 splicing targets (SLC12A2, IFT88,
401 STXBP1, C1qTNF6), revealed putative SRSF1 binding sites in both the spliced and adjacent

402 exon (**Fig. S6F, S6G, S6H, S6I**), suggesting a potential universal role of SRSF1 on the IWS1
403 phosphorylation-dependent exon inclusion phenotype. To experimentally address the
404 proposed model, we carried out RNA-IP (RIP) experiments in the same shControl, shIWS1,
405 shIWS1/WT-R and shIWS1/MT-R NCI-H522 and NCI-H1299 cells, focusing on the binding of
406 SRSF1 on the *U2AF2* exon 2, intron 2 and exon 3. The results confirmed that SRSF1 binds
407 primarily to exon 2, but only in the shControl and shIWS/WT-R cells, which parallels its
408 chromatin binding to the H3K36me3-bound p52/LEDGF (**Fig. 4G, S6D**).

409 Based on the data presented in Figure 3 and Figure 4, we propose that IWS1
410 phosphorylation by AKT is required for the recruitment of SETD2 to CTD of RNA Pol II. The
411 latter promotes the trimethylation of Histone H3 at K36 on the body of *U2AF2* and perhaps
412 other target genes. This facilitates the recruitment of the histone H3K36me3-interacting protein
413 p52/LEDGF and its binding partner SRSF1, resulting in exon inclusion of *U2AF2* exon 2 (**Fig.**
414 **4H**).

415 **The novel U2AF65 splice variant lacking exon 2, does not interact with Prp19.**

416
417 The predominant novel splice variant of the *U2AF2* mRNA that we identified in shIWS1
418 and shIWS1/MT-R cells, lacks exon 2, which encodes part of the N-terminal RS domain of
419 U2AF65 (**Fig 5A**). This domain is responsible for the interaction of U2AF65 with several factors
420 that contribute to mRNA splicing, 3' cleavage and polyadenylation (Millevoi et al., 2006[58], Shen
421 H, et al., 2004[59]). One of these factors is Prp19, a key component of the seven-member
422 ubiquitin ligase complex Prp19C (S.P Chan et al., 2003[60], S.P Chan et al., 2005[61], R. Hogg et
423 al., 2010[62], Chanarat S. et al., 2013[63]).
424 Using co-immunoprecipitation experiments in HEK-293T cells engineered to express
425 the exon-2 containing U2AF65α or the exon 2-excluded U2AF65β variants, with a V5-tag, we
426 confirmed that only the U2AF65α splice variant interact with the endogenous Prp19 (**Fig. 5B**).

427 More importantly, consistent with the previous results, immunoprecipitation of endogenous
428 U2AF65 in NCI-H522 and NCI-H1299 shControl, shIWS1, shIWS1/WT-R and shIWS1-MT-R
429 cells, revealed impaired U2AF65-Prp19 interaction in shIWS1 and shIWS1/MT-R cells, which
430 express primarily U2AF65β, while maintaining interaction with its U2-heterodimer partner,
431 U2AF35, and SF3B1, a U2 snRNP protein (**Fig. 5B, S7A**). These data confirmed that the
432 interaction of U2AF65 with Prp19 depends on the sequence encoded by *U2AF2* exon 2,
433 whose inclusion is regulated by IWS1 phosphorylation.

434 **The splicing of the *U2AF2* mRNA downstream of IWS1 phosphorylation, regulates the**
435 **mRNA splicing and expression of *CDCA5*.**

436 Based on previous reports, U2AF65 binds RNA Pol II, leading to an U2AF65-
437 dependent recruitment of Prp19 to the newly-synthetized pre-mRNA and promoting proper co-
438 transcriptional splicing activation (C.J David et. al., 2011[64]). Furthermore, it had been shown
439 that U2AF65 and Prp19 are necessary for the pre-mRNA splicing and accumulation of Sororin,
440 a component of the cohesin complex, which is encoded by *CDCA5* (Watrin et. al., 2014[65]).
441 Given that aberrant splicing of *U2AF2* in shIWS1 and shIWS1/MT-R cells resulted in the loss
442 of the interaction between U2AF65 and Prp19, we hypothesized that the pre-mRNA splicing
443 of *CDCA5* in these cells will be impaired. We therefore employed qRT-PCR to examine the
444 ratio of spliced and unspliced *CDCA5* mRNA in the same cells. The results confirmed that the
445 mRNA splicing of *CDCA5* was impaired in both the shIWS1 and shIWS1/MT-R cells, with no
446 significant change in the spliced/unspliced ratios of the control *GUSB* mRNA (Watrin et. al.,
447 2014[65]) (**Fig. 5D, S7B, S7C upper panels).** More importantly, the splicing defect was rescued
448 by U2AF65α but not by U2AF65β. (**Fig. 5E, S7D upper panels**).

449 To determine how IWS1 phosphorylation regulates *CDCA5* RNA splicing, we
450 performed RNA Immunoprecipitation (RIP) experiments in shControl, shIWS1, shIWS1/WT-R
451 and shIWS1/MT-R NCI-H522 and NCI-H1299 cells. The results confirmed that both spliced
452 variants of U2AF65 bind equally well *CDCA5* pre-mRNA, as well as the control *GUSB* pre-
453 mRNA, as expected (**Fig. 5D, 5E, S7C, S7D middle panels**) (Shen H. et al., 2004[31]).

454 However, Prp19 binding to the same pre-mRNA regions of *CDCA5* was significantly impaired
455 in shIWS1 and shIWS1/MT-R cells, which predominantly express the U2AF65β isoform (**Fig.**
456 **5D, S7C lower panels**). More importantly, the impaired Prp19 binding to *CDCA5* pre-mRNA
457 was rescued by U2AF65α, but not U2AF65β (**Fig. 5E, S7D lower panels**).

458 Given that only spliced mRNAs are transported out of the nucleus, we also examined
459 the abundance of cytosolic *CDCA5* mRNA in shControl, shIWS1, shIWS1/WT-R, shIWS1/MT-
460 R and shIWS1 rescues with U2AF65α or U2AF65β. To this extent, we fractionated the nuclear
461 and cytosolic compartments of these NCI-H522 and NCI-H1299 cells. Then, using western
462 blotting, we confirmed the validity of the fractionation by the expected expression of Lamin
463 A/C and GAPDH only in the nuclear and cytosolic protein compartment, respectively (**Fig.**
464 **S7E**). The results also confirmed that the mature *CDCA5* mRNA in the cytoplasm was
465 downregulated as expected, and that the RNA transport deficiency is rescued by U2AF65α,
466 but not U2AF65β (**Fig. 5F**).

467 To determine whether the *CDCA5* RNA splicing defect in shIWS1 and shIWS1/MT-R
468 NCI-H522 and NCI-H1299 cells prevents the expression of its protein product, Sororin, we
469 examined the its expression along with the expression of IWS1, p-IWS1, U2AF65 and Prp19,
470 by western blotting. The results confirmed that the expression of Sororin parallels IWS1
471 phosphorylation as it was impaired in shIWS1 and shIWS1/MT-R NCI-H522 and NCI-H1299
472 cells. More importantly, the low expression of Sororin in shIWS1 was rescued by U2AF65α,
473 but not by U2AF65β (**Fig. 5G, S7F upper panels**). Taken together, these data confirm that
474 IWS1 phosphorylation via the alternative RNA splicing of *U2AF2*, is ultimately responsible for
475 the regulation of the mRNA splicing of *CDCA5*.

476 **Sororin and p-ERK form a positive feedback loop, which is activated by IWS1**
477 **phosphorylation and promotes the expression of CDK1 and Cyclin B1.**

478 It has been previously shown that the downregulation of Sororin leads to reduced ERK
479 phosphorylation, in human Colorectal Cancer (CRC) and human Hepatocellular Carcinoma
480 (HCC) (Wang J., et al., 2018[66], A. Zhen et. al., 2019[67]). Since the IWS1 phosphorylation-

481 dependent *U2AF2* alternative RNA splicing regulates the abundance of Sororin, we
482 hypothesized reduced ERK phosphorylation in NCI-H522 and NCI-H1299 shIWS1 and
483 shIWS1/MT-R cells. This was confirmed by experiments showing that ERK phosphorylation
484 at Y202/T204 was reduced in shIWS1 and shIWS1/MT rescue cells, while total ERK was not.
485 More importantly, the reduction of p-ERK levels in shIWS1 cells was rescued by U2AF65α,
486 but not by U2AF65β (**Fig. 5G, S7F lower panels**), confirming the p-IWS1/Sororin/p-ERK axis
487 in lung adenocarcinoma.

488 Sororin is a known target of ERK, which phosphorylates the protein at S79 and S209
489 (Nguyen et. al., 2010[68]). Interestingly, whereas the wild type Sororin and the Sororin
490 phosphomimetic mutant S79E/T209E (Sororin DM-E) rescued the phosphorylation of ERK in
491 shIWS1 cells, the S79A/S209A (Sororin DM-A) mutant did not (**Fig 5G and S7F Lower**
492 **panels**). This suggests the existence of a Sororin-ERK phosphorylation positive feedback loop
493 in lung adenocarcinoma, which is controlled by the AKT-dependent IWS1 phosphorylation,
494 and the *U2AF2* alternative RNA splicing.

495 Based on previous observations, down-regulation of Sororin leads to impaired
496 expression of CDK1 and Cyclin B1 (Wang J., et al., 2018[66], A. Zhen et. al., 2019[67]). To this
497 extent, we hypothesized that IWS1 phosphorylation regulates the expression of CDK1 and
498 Cyclin B1, through the regulation of *U2AF2* RNA splicing and the abundance of Sororin.
499 Consistent with the preceding data, western blotting revealed reduced levels of CDK1 (Y15)
500 phosphorylation, CDK1 and Cyclin B1 levels in shIWS1 and shIWS1/MT-R cells. More
501 importantly, the downregulation of these molecules in shIWS1 cells was rescued by U2AF65α,
502 but not by U2AF65β, wild type Sororin and Sororin DM-E but not by Sororin DM-A (**Fig. 5G,**
503 **S7F lower panels**). These data indicate that the Sororin-ERK phosphorylation feedback loop,
504 downstream of IWS1 phosphorylation, controls the expression of CDK1 and Cyclin B1,
505 through a process depending on *U2AF2* alternative RNA splicing.

506 **IWS1 phosphorylation by AKT3 promotes cellular transformation through *U2AF2***
507 **alternative RNA splicing and Sororin expression.**

508 Based on the proceeding data, IWS1 phosphorylation maintains the establishment of
509 a Sororin/ERK phosphorylation feedback loop in lung adenocarcinoma, through the epigenetic
510 regulation of *U2AF2* alternative RNA splicing. Given that ERK phosphorylation is a well-known
511 signal for cellular transformation (Guo et al., 2020[69], Wee et al., 2017[70]), we then questioned
512 whether the AKT/IWS1 pathway promotes transformation of immortalized Human Bronchial
513 epithelial cells (HBEC-hTERT). To this extent, we overexpressed myristoylated AKT3 (myr-
514 AKT3), the phosphomimetic AKT3 T305/S472DD (AKT3-DD), phosphomimetic IWS1
515 S720D/T721E (IWS1 DE), the two spliced U2AF65 isoforms (U2AF65α, U2AF65β) and
516 several forms of Sororin (Sororin wt, Sororin DM-E, Sororin DM-A) (**Fig. S8A**). Consistent with
517 the ERK phosphorylation pattern, cell transformation assays showed effective transformation
518 of HBEC-hTERT induced by the active AKT3 (myr-AKT3, AKT3-DD) and IWS1 (IWS1 DE).
519 More importantly, the downstream pathway of U2AF65α spliced isoform and active Sororin
520 (Sororin wt, Sororin DM-E) induced cellular transformation, but the exon 2 deficient U2AF65β
521 and inactive Sororin mutant (Sororin DM-A), did not (**Fig. S8B, S8C**). Collectively, these
522 results confirm that IWS1 phosphorylation promotes cellular transformation through the
523 epigenetic regulation of *U2AF2* alternative RNA splicing and Sororin/ERK loop.
524

525 **IWS1 phosphorylation controls the ERK phosphorylation levels, in cells harboring**
526 ***EGFR* and *KRAS* mutations, through *U2AF2* RNA splicing.**

527 *EGFR* and *KRAS* are frequently mutated in human lung adenocarcinoma and promote
528 downstream pro-oncogenic signals through the phosphorylation of ERK (Guo et al., 2020[69],
529 Wee et al., 2017[70]). We therefore addressed the role of IWS1 phosphorylation and *U2AF2*
530 alternative RNA splicing on Sororin and ERK phosphorylation in lung adenocarcinoma cells
531 harboring *KRAS* (A549 and NCI-H460) and *EGFR* (NCI-H1975, PC-9, NCI-H1650) mutations.
532 The results showed that *U2AF2* alternative RNA splicing pattern is independent of the

533 mutational status (**Fig. 5H, upper and lower left panels**). More importantly, the knockdown
534 of IWS1 downregulated the expression of Sororin and ERK phosphorylation (**Fig. 5H, upper**
535 **panel**). These data were striking, because they showed that the effect of Sororin on ERK
536 phosphorylation is dominant over the effect of *KRAS* and *EGFR* mutations. Surprisingly, the
537 reduction of ERK phosphorylation induced by the loss of IWS1 phosphorylation was more
538 robust in *EGFR* as opposed to the *KRAS* mutated cells (**Fig. 5H, lower right panels)**.
539 Altogether, this data suggests that IWS1 phosphorylation controls the *U2AF2* alternative RNA
540 splicing independent of the cellular mutational status and shows robust reduction of the
541 Sororin/ERK phosphorylation axis in lung adenocarcinoma cell lines, harboring *EGFR*
542 mutations.

543 **IWS1 phosphorylation promotes cell proliferation by controlling the Sororin/ERK**
544 **phosphorylation feedback loop, through the novel *U2AF2* RNA splicing pattern.**

545 Given that the Sororin/ERK phosphorylation positive feedback loop, activated by IWS1
546 phosphorylation, promotes the expression of CDK1 and Cyclin B1, we hypothesized that the
547 stimulation of cell proliferation by IWS1, we had observed previously (Sanidas et al., 2014[10]),
548 may depend on the activation of the *U2AF2* RNA splicing-regulated CDCA5/ERK
549 phosphorylation loop. To address this hypothesis, we examined the rate of proliferation of
550 shControl, shIWS1, shIWS1/Sororin WT, shIWS1/Sororin DM-A and shIWS1/Sororin DM-E
551 NCI-H522 and NCI-H1299 cells growing under standard culture conditions. The results
552 showed that cell proliferation was inhibited by shIWS1. More importantly, the inhibition was
553 rescued by wild type Sororin and Sororin DM-E but not by the Sororin DM-A. Significantly, the
554 proliferative effect of the Sororin DM-E was more robust than the proliferative effect of wild
555 type (wt) Sororin **(Fig. 6A).** To further support these data, similar proliferation assays were
556 performed in NCI-H522 and NCI-H1299 shControl, shIWS1, shIWS1/U2AF65α and
557 shIWS1/U2AF65β. The results revealed that the proliferation phenotype in shIWS1 cells was

558 rescued by U2AF65α, but not by U2AF65β, confirming that the regulation of cell proliferation
559 depends on the novel *U2AF2* RNA splicing pattern (**Fig. 6B**).
560 As we pointed out above, the loss of phosphorylated IWS1 results in more robust
561 down-regulation of ERK phosphorylation in *EGFR* as opposed to *KRAS* mutated lung
562 adenocarcinoma cells (**Fig. 5H**), suggesting that the *EGFR* mutated cells would be more
563 sensitive to the loss of IWS1. This was confirmed by experiments showing reduced cell
564 proliferation of A549 (*KRAS* mutant) and NCI-H1975 (*EGFR* mutant) shIWS1 cells (**Fig 6C**).
565 More importantly, in parallel with the data of the reduction of ERK phosphorylation, *EGFR*
566 mutated cell line NCI-H1975, demonstrates the highest reduction in cell proliferation induced
567 by the loss of IWS1 (**Fig. S9A**). Furthermore, the observed inhibition of cell proliferation in
568 A549 and NCI-H1975 shIWS1 cells was rescued by U2AF65α, but not by U2AF65β,
569 suggesting that the *U2AF2* RNA splicing affects cell proliferation in *KRAS* and *EGFR* mutated
570 cells (**Fig. 6D**). Additional support was provided by Western Blot analysis showing that in all
571 cell lines, the cell proliferation marker PCNA was downregulated in shIWS1 cells, an effect
572 rescued by U2AF65α, but not the U2AF65β (**Fig. S9B**). Notably, shIWS1-induced
573 downregulation was more robust in *EGFR* (NCI-H1975, NCI-H1650) as opposed to *KRAS*
574 mutated (A549, NCI-H460) cells (**Fig. S9B**)**.** Altogether, these data suggest that IWS1
575 phosphorylation controls cell proliferation through the regulation Sororin/ERK phosphorylation
576 loop mediated by *U2AF2* RNA splicing.
577 To determine whether IWS1 phosphorylation stimulates cell proliferation by promoting
578 progression through the G2/M phase of the cell cycle, we stained log phase cultures of
579 shControl, shIWS1, shIWS1/U2AF65α and shIWS1/U2AF65β NCI-H522, NCI-H1299, A549
580 and NCI-H1975 cells with Propidium Iodide (PI) and we examined the percentage of cells in
581 different phases of the cell cycle by flow-cytometry. The results confirmed that the knockdown
582 of IWS1 induced prolonged G2/M phase of the cell cycle, an effect rescued by U2AF65α, but
583 not by U2AF65β, confirming the role of IWS1 in cell cycle progression through *U2AF2* RNA
584 splicing (**Fig. 6E**). Consistent with the proliferation data, the G2/M arrest was less pronounced
585 in the *KRAS* mutant cell line, A549 **(Fig. 6E).** Altogether, data presented in figures 5 and 6

586 suggest that the *EGFR* mutated lung adenocarcinoma cell lines are more sensitive than *KRAS*
587 mutant cells upon the loss of IWS1 phosphorylation.

588 **IWS1 phosphorylation and novel *U2AF2* splice variant fluctuate during progression**
589 **through the cell cycle.**

590 The preceding findings suggest that an RNA splicing event, regulated by the AKT-
591 mediated phosphorylation of IWS1, plays a critical role in cell cycle progression. Given that
592 the expression and activity of molecules critically involved in the regulation of the cell cycle,
593 tend to fluctuate as cells transit from one cell cycle phase to the next (Kent et al., 2019[71]), we
594 examined the expression and phosphorylation of IWS1, the *U2AF2* alternative RNA splicing
595 pattern and the expression of SETD2, Sororin and the trimethylation of histone H3K36 in G1,
596 S and G2/M in NCI-H1299 cells. To this end, using carboxyfluorescein succinimidyl ester
597 (CFSE)-like DNA dyes, cells were stained in G1, S and G2/M phase and they were fractionated
598 by FACS sorting (Begum et al, 2013[72]). Western blotting of these fractions confirmed that
599 IWS1, phospho-IWS1, SETD2, Histone H3K36me3 and Sororin are indeed upregulated in S
600 and G2/M. As specific cell cycle phase markers, Cyclin E1 (G1 phase), CDC25A (S phase)
601 and phosphorylated Histone H3 (S10) were used, validating the integrity of the fractionation
602 **(Fig. 6F lower panels)**. More importantly, IWS1 expression and phosphorylation occured
603 more during S phase, whereas Sororin, SETD2 and Histone H3K36me3 levels were more
604 robust during G2/M phase, as previously reported (Watrin et al., 2014[65], Dronamraju et al.,
605 2018[73]) **(Fig. 6F upper panels).** Furthermore, using RNA extracts from the same cells, RT-
606 PCR experiments revealed an *U2AF2* RNA exon inclusion pattern which parallels the
607 expression of IWS1 during the cell cycle **(Fig. 6F).** Moreover, qRT-PCR confirmed that the
608 expression of *IWS1* and *CDCA5* in the RNA level parallels their protein expression, with no
609 change of total *U2AF2* levels (**Fig. 6G**).

610 Chromatin Immuno-Cleavage (ChIC) experiments revealed increased binding of IWS1
611 on *U2AF2* exon 2 and 3 during S and G2/M phase, with significantly more binding during S

612 phase (**Fig. 6H upper panel**). Consistently, the binding of SETD2, the abundance of
613 H3K36me3 and the E2/E3 *U2AF2* ratio parallel the levels and the binding of IWS1 on *U2AF2*
614 gene, although total SETD2 and H3K36me3 are more abundant during G2/M phase,
615 confirming the strong dependence of the pathway on IWS1 (**Fig. 6H, 6I**).

616 Overall, these data provide strong evidence for the model in which IWS1, whose
617 expression increases during S and G2/M, recruits SETD2 to CTD of RNA-Pol II, following
618 phosphorylation by AKT. RNA-Pol II-associated SETD2 trimethylates Histone H3 at K36
619 during transcriptional elongation, resulting in the assembly of p52/LEDGF-SRSF1 complexes
620 which promote the inclusion of exon 2 in the *U2AF2* mRNA. Subsequently, the U2AF65α
621 isoform along with Prp19, facilitate the proper splicing of *CDCA5* pre-mRNA, leading to Sororin
622 accumulation during S and G2/M phase. Finally, Sororin forms a positive feedback loop with
623 ERK phosphorylation, and controls the progression through the cell cycle in lung
624 adenocarcinoma. The events described in this model define a molecular mechanism of a novel
625 splice variant which is epigenetically regulated during the cell cycle, feeding back into its
626 regulation (**Fig. 6J, S9C**).

627 **IWS1 phosphorylation controls tumor growth *in vivo*, by regulating the**
628 ***U2AF2*/Sororin/ERK axis.**

629 Based on our preceding studies, the loss of IWS1 and IWS1 phosphorylation impair
630 tumor growth in a mouse xenograft model (Sanidas et al., 2014[10]). To further validate and
631 expand this observation, we knocked down IWS1 in NCI-H1299, A549 and NCI-H1975 cells
632 and we injected them subcutaneously, along with shControl cells, in immunocompromised
633 NSG mice. The animals were monitored for tumor development and were sacrificed 4 (NCI-
634 H1299 and NCI-H1975) and 6 (A549) weeks later, respectively. The results revealed that the
635 knockdown of IWS1 reduced tumor weight and volume (**Fig. 7A**). To further demonstrate the
636 role of *U2AF2* alternative RNA splicing in tumor growth, we repeated the mouse xenografts
637 models by injecting NCI-H1299 cells transduced with shIWS1 and rescued with the two

638 U2AF65 isoforms, U2AF65α and U2AF65β, along with shControl cells (**Fig. S10A**). The
639 results showed that the reduced tumor volume phenotype in shIWS1-derived tumors was
640 rescued in U2AF65α, but not U2AF65β-derived tumors, confirming the role of the novel *U2AF2*
641 RNA splicing pattern on tumor growth (**Fig. 7B**).

642 Analysis of the percentage reduction of tumor characteristics in the various mouse
643 xenografts cell lines models, showed that the reduction of tumor growth induced by the loss
644 of IWS1 was less pronounced in A549 cells, harboring *KRAS* mutations (**Fig. 7C**). The weak
645 growth reduction of tumours derived from shIWS1 A549 cells, paralleled the weak inhibition of
646 ERK phosphorylation (**Fig 5H**), cell proliferation (**Fig. S9A**) and cell cycle progression (**Fig.**
647 **6E**) induced by the knockdown of IWS1 in these cells.

648 The preceding data confirmed that the knockdown of IWS1 reduces tumor growth. We
649 then proceeded to address the mechanism of the growth inhibition, which had not been
650 adequately addressed in our earlier studies. First, we confirmed the efficiency of the IWS1
651 knockdown, by probing western blots of tumor cell lysates with anti-IWS1 and anti-phospho-
652 IWS1 (S720) antibodies (**Fig. 7D**). Following this, we employed RT-PCR and qRT-PCR on
653 RNA derived from these tumors, to address the usage of exon 2 in the *U2AF2* mRNA. The
654 results confirmed that the knockdown of IWS1 has no effect on the total *U2AF2* mRNA levels
655 (**Fig. S10B**), but promotes the exclusion of exon 2 from *U2AF2* mRNA (**Fig. 7D upper panels,**
656 **S10B**). Probing both tumor lysates and tissue sections with antibodies to regulators and
657 targets of the Sororin/ERK feedback loop, confirmed that the activity of the loop is reduced in
658 tumors derived from shIWS1 cells (**Fig. 7D lower panels, S10C**). Measuring the abundance
659 of the PCNA and Ki-67 proliferation markers by western blotting (PCNA), or
660 Immunohistochemistry (Ki-67), also confirmed the reduced expression of these markers in the
661 shIWS1 tumors (**Fig. 7D lower panels, 7E**).

662 Quantitative analyses of the western blot (p-ERK and PCNA) and IHC data (Ki-67)
663 (**Fig. 7F**), showed that the reduction of all these markers was less pronounced in the *KRAS*
664 mutant A549 tumors, suggesting that *KRAS* mutant cells are more resistant to the loss of
665 IWS1. Altogether, these results are in agreement with the less pronounced effect of shIWS1

666 in the tumor growth in A549-derived tumors (**Fig. 7C**) and with the less pronounced effect of
667 shIWS1 in p-ERK levels (**Fig. 5H**), cell proliferation in culture and degree of G2/M arrest in
668 these cells (**Fig. 6**).

669 **The AKT3/p-IWS1/U2AF2 axis is active in human lung adenocarcinoma and affects the**
670 **clinical outcome of *EGFR* mutant patients.**

671 To determine whether the pathway activated by IWS1 phosphorylation leading to
672 Sororin expression and ERK phosphorylation is active in human lung adenocarcinoma
673 (LUAD), we examined the expression and phosphorylation of several components of this axis,
674 along with the alternative RNA splicing pattern of *U2AF2* in a set of 40 human LUAD samples.
675 For 30 of these tumors, Normal Adjacent Tissue (NAT) was also available and it was tested
676 on parallel with the matching tumor sample (**Fig. 8A**). More importantly, the exon 2-containing
677 *U2AF2* transcript was more abundant in tumor versus normal samples, with no change in total
678 levels of *U2AF2* (**Fig. S11A**). Overall, these results confirmed that the pathway was active in
679 the tumors, but not in the NAT.

680 Human LUAD frequently harbors *KRAS* or *EGFR* mutations and data presented in this
681 report suggest that lung adenocarcinoma cells harboring *EGFR* mutations are more sensitive
682 to the loss of IWS1, as opposed to *KRAS* mutant cells (**Fig. 5, Fig. 6, Fig. 7**). We therefore
683 used monoclonal antibodies, which selectively recognize *KRAS* and *EGFR* mutant forms (Kim
684 et al., 2015[75]), and we identified tumors harboring these mutations (**Fig. 8A**). Comparison of
685 the correlations coefficients of IWS1 and phosphorylated IWS1 with the E2/E3 *U2AF2* and all
686 the components of the downstream Sororin/ERK phosphorylation pathway, revealed strong
687 correlations in the entire cohort with more robust correlations in the *EGFR* than the *KRAS*
688 mutant tumors (**Fig. 8B**). Strikingly, the levels of IWS1 phosphorylation and abundance of
689 exon 2-containing *U2AF2* isoform exhibit strong correlation with more aggressive clinical stage
690 and defines poor survival in the entire patients' and the *EGFR* mutant cohort, but not in the
691 *KRAS* mutated group (**Fig. 8C, 8D**).

692 The preceding data were confirmed by IHC, using sections of a commercially available
693 tissue microarray (TMA) of 50 LUAD with paired NAT. The TMA samples were probed with
694 antibodies against p-IWS1, Sororin, p-ERK, p-CDK1 and EGFR E746-A750 deletion (**Fig.**
695 **S11B, S11C**). The results confirmed that the pathway is more active in the tumors, compared
696 to the NAT samples **(Figure S11D)**. More importantly, IWS1 phosphorylation expression
697 exhibits strong correlations with the components of the Sororin/ERK axis, with more robust
698 correlations in the *EGFR* mutant tumors (**Fig. S11E**). Consistently, the results also showed
699 that the activity of the pathway correlates with clinical stage and histological grade **(Fig. S11F,**
700 **S11G)**. In addition to confirming the western blot data in our set of LUAD patients, the IHC
701 data also demonstrate that the activity of the pathway can be monitored clinically in human
702 lung adenocarcinoma patients by IHC.

703 The data generated from the analysis of our LUAD samples and the TMA were
704 confirmed by data in publicly available databases. Analysis of the data derived from Tumor
705 Cancer Genome Atlas (TCGA), revealed correlations between *IWS1* or *SRSF1*, *U2AF2* E2
706 and other components of the IWS1 phosphorylation pathway (**Fig. 8E**). The results also
707 showed that the *U2AF2* E2/E3 ratio and the expression of *CDCA5* mRNA were significantly
708 upregulated in TCGA LUAD patients with high *IWS1*, compared to the low *IWS1* group (**Fig.**
709 **S11H**). More importantly, *IWS1* expression correlates with more aggressive tumor stage in
710 tumors harboring *EGFR* but not *KRAS* mutations (**Fig. S8I**).

711 Based on recent reports, 55% of lung adenocarcinoma patients relapse and up to 40%
712 of them develop metastatic disease (Popper et al., 2016[76], West et al., 2020[77]). Given that the
713 IWS1 phosphorylation controls oncogenic signals and plays a critical role in patients with lung
714 adenocarcinoma, as suggested in this report, we questioned whether the IWS1-mediated axis
715 is associated to clinical relapse or metastasis in LUAD patients. Analysis of a relapse-related
716 molecular signature database GSE13213 (Tomida et al., 2009[78]), revealed increased
717 expression of the IWS1/Sororin axis in relapsing patients harboring *EGFR*, but not with *KRAS*
718 mutations (**Fig. 8F, S11J**). We therefore conclude that the activation of the IWS1
719 phosphorylation pathway may also contribute to the relapse of lung adenocarcinoma,

720 especially with *EGFR* mutations. Furthermore, analysis of RNA-seq data derived from
721 dissected brain metastatic lesions from lung adenocarcinoma patients (GSE141685), revealed
722 increased expression of the components of IWS1/Sororin axis in the metastatic as opposed
723 to primary lung adenocarcinoma lesion (**Fig. 8G**). Importantly, the *U2AF2* E2/E3 ratio was also
724 elevated in the metastatic tumours, while the IIIb/IIIc ratio of the *FGFR2* mRNA was reduced
725 (**Fig. S11K and S11L**), in agreement with our earlier observations, showing that IWS1
726 phosphorylation promotes *FGFR2* exon IIIb skipping (Sanidas et al, 2014[10]). To further
727 elucidate the role of IWS1 phosphorylation in metastasis in lung adenocarcinoma, analysis of
728 single-cell RNA-Seq (scRNA-seq) data from a recent Patient-Derived LUAD Metastatic model
729 (Laughney et al., 2020[79]), revealed increased expression and robust correlation of IWS1 and
730 CDCA5 in the incipient metastasis providing evidence of the role of IWS1 in the metastatic
731 potential of lung adenocarcinoma (**Fig. 8G**). Moreover, analysis of the RNA-Seq data in the
732 TCGA LUAD dataset, which contains information on cancer-associated mutations, confirmed
733 the link between *IWS1* expression and metastatic disease, but also showed that *IWS1* is again
734 upregulated in tumours derived from patients with metastatic disease, and harboring *EGFR*,
735 but not *KRAS* mutations (**Fig. S11M**).
736 As expected from the preceding data, *IWS1* expression, *U2AF2* exon 2 inclusion and *FGFR2*
737 exon IIIb exclusion, are indicators of poor prognosis in patients with lung adenocarcinomas in
738 the TCGA LUAD dataset, harboring *EGFR*, but not *KRAS* mutations (**Fig. 8H**). In addition, the
739 survival of patients with lung adenocarcinomas in the same dataset, expressing high levels of
740 *IWS1*, was reduced if the tumours also harbored mutated *EGFR* (**Fig. S11N**). Confirmatory to
741 these observations, was the finding that the synergy between *IWS1* expression and *EGFR*
742 mutations, was also observed in the GSE13213/GSE26939 dataset (**Fig. S11O**).
743 Finally, analysis of Whole Exome Sequence (WES) lung adenocarcinoma datasets
744 (Ding et al., 2008[81], Imielinksi et al., 2012[82], Rivzi et al., 2015[83], Jordan et al., 2017[84], Chen et
745 al., 2020[85]) via cBioportal (Cerami et al., 2012[86], Gao et al., 2013[87]) for mutations and copy-
746 number alterations, revealed the occurrence of mutations, fusions and amplification of the

747 AKT/IWS1 pathway described in this report, in patients with Lung Adenocarcinoma (**Figure**
748 **S12A**). Notably, the same analysis revealed the existence of missense and splice site
749 mutations of *U2AF2* on RS domain sequences, along with previously described cancer-
750 associated mutations of the U2AF65 RRM domain, G176E/V, known to affect RNA splicing
751 (Glasser E. et al., 2017[32], Maji et al., 2020[88]) (**Fig. S12B**). Furthermore, consistently with all
752 previous data, *AKT3* and *CDCA5* expression is associated with poor prognosis in patients with
753 lung adenocarcinomas in the TCGA LUAD dataset (**Fig. S12C**).

754 Altogether, these data come in agreement with the *in vitro* (**Fig. 5 and Fig. 6**) and *in*
755 *vivo* (**Fig. 7**) data provided in this report, that indicate that the p-AKT/p-IWS1 axis through the
756 epigenetic regulation of *U2AF2* alternative RNA splicing and oncogenic signals, defines poor
757 survival in lung adenocarcinoma patients harboring *EGFR* mutations.

758 **Discussion**

759 **Regulation of the Sororin/ERK phosphorylation loop from IWS1 phosphorylation**

760 Our studies implicate IWS1 phosphorylation as a regulator of oncogenic signals in lung
761 adenocarcinoma through alternative RNA splicing. We report a novel *U2AF2* splice variant in
762 lung adenocarcinoma which lacks exon 2 upon loss of phosphorylated IWS1. This signaling
763 pathway is initiated by the AKT3-dependent phosphorylation of IWS1, which induces the
764 epigenetic regulation of *U2AF2* alternative RNA splicing, through H3K36me3-mediated
765 signals (**Fig. 3**). The shift in the alternative RNA pattern controls the interaction of U2AF65
766 with Prp19. These two factors regulate the pre-mRNA splicing of *CDCA5* and the expression
767 of its protein product, Sororin, co-transcriptionally (**Fig. 5**). Our data provide insight for the
768 existence of a novel Sororin/ERK phosphorylation feedback loop in lung adenocarcinoma,
769 which is regulated by the alternative RNA splicing of *U2AF2*. Notably, the Sororin-dependent
770 regulation of ERK phosphorylation by IWS1 is dominant over the ERK-activating *EGFR*
771 mutations and has a major impact in the biology of these cells harboring such mutations **(Fig.**

772 **5 and Fig. 6)**. Sororin is a member of the cohesin complex, a seven member protein complex
773 which plays a central role in sister chromatid cohesion and higher chromosomal architecture
774 during meiosis (Ishiguro et al., 2019[89]). This molecular Sororin/ERK phosphorylation switch
775 controls cell proliferation and the progression through the G2/M phase of the cell cycle (**Fig.**
776 **6**). Defects in Sororin expression lead to the activation of spindle assembly checkpoint,
777 delaying the progression through the G2/M phase (Dreier et al., 2011[90]). Given that loss of
778 IWS1 reduced the expression of Sororin, this could explain the G2/M phase delay observed
779 in shIWS1 cells (**Fig. 6**). How the ERK-phosphorylated Sororin promotes the phosphorylation
780 and activation of ERK is currently unknown. Our working hypothesis is that the
781 phosphorylation and activation of ERK is due to signals induced by the interaction of Sororin
782 with its partners in the cohesin complex. If this is the case, the cell may use this mechanism
783 to sense the successful progression from prometaphase to metaphase, in order to activate a
784 molecular switch. This enhances the phosphorylation of Sororin, facilitating progression
785 through the G2/M phase of the cell cycle. In addition, data presented in this report show that
786 the loss of IWS1 leads to impaired expression of CDK1 and Cyclin B1, a phenotype regulated
787 through the Sororin/ERK phosphorylation loop (**Fig. 5**). On the other hand, our data indicate
788 strong correlation of *CDCA5*, *CDK1* and *CCNB1* (Cyclin B1) mRNA levels in LUAD patients,
789 derived from TCGA database (**Fig. 8**). These observations indicate that the induction of Cyclin
790 B1 and CDK1 by Sororin is most likely at the level of transcription. A potential mechanism for
791 this transcriptional regulation, controlled by IWS1, was suggested by earlier studies showing
792 that the Cohesin complex interacts with the Mediator complex. The Mediator-Cohesin
793 complexes are loaded by the NIPBL-Cohesin loading factor to enhancers and core promoters
794 of target genes. Enhancer and core promoter-associated complexes promote loop formation
795 between these segments, and regulate transcription (Kagey et al., 2010[91]). The contribution
796 of this and other mechanisms on the regulation of *CCNB1* and *CDK1* expression is under
797 investigation.

798 **Regulation of RNA processing in multiple layers**

799 Another important conclusion, based on the data presented in this report, is that RNA
800 splicing is a process regulated at multiple levels. IWS1 phosphorylation directly regulates the
801 alternative RNA splicing of *U2AF2*, via a H3K36me3-mediated mechanism (**Fig. 3**) and thus,
802 introducing a new layer of RNA splicing regulation. This shift in alternative RNA splicing pattern
803 affects the binding and formation of the U2AF65/Prp19 interacting complex, required for the
804 efficient splicing of *CDCA5*, which ultimately controls pro-oncogenic ERK phosphorylation
805 signals (**Fig. 5, Fig. 6**). The reason for the multilayered control of RNA splicing by a single
806 RNA splicing regulator could be that this allows a limited number of available pathways to
807 converge, in different combinations, on a large number of RNA splicing events and
808 differentially regulate them. To this extent, our data provide insight on the novel importance of
809 the AKT/IWS1 phosphorylation axis on the fine-tuning of the global epigenetic regulation of
810 RNA splicing cellular program, governing cell cycle regulation and tumor progression. Another
811 important point raised from our data, is that the removal of RS domain of U2AF65,
812 demonstrates a gene-specific indirect effect on intron removal events, through the regulation
813 and recruitment of Prp19 on the target introns. Although U2AF65 is a general splicing factor
814 affecting the branchpoint assembly (Shen H. et al., 2004[31]), earlier observations on the
815 regulation of *CDCA5* pre-mRNA, revealed that components of Prp19 Complex (Prp19 and
816 Cdc5L) along with U2AF65, affected the intron removal of *CDCA5* pre-mRNA, but not the
817 removal of intronic regions in other genes, such as *GUSB* (Watrin et. al., 2014[65]),
818 demonstrating a selective effect of the U2AF65-Prp19 complex on the intronic regions of target
819 genes. Consistently, our RIP data demonstrate the binding of Prp19 in the intronic areas of
820 *CDCA5*, but not *GUSB*, when U2AF65α (exon 2 containing isoform) is expressed (**Fig. 5**),
821 further supporting the selectivity of the latter complex on the target genes. The genome-wide
822 distribution of U2AF65 and Prp19 upon loss of phosphorylated IWS1 and RS domain
823 exclusion, will be addressed in future studies.

824 **IWS1 phosphorylation and *U2AF2* splicing are regulated through the cell cycle, feeding**
825 **back to its regulation, enhancing tumor proliferation.**

826 Given that the cell cycle is an integrated system and that cell cycle regulatory
827 mechanisms tend to also be cell cycle-regulated, we examined and confirmed that IWS1
828 expression and phosphorylation fluctuate through the cell cycle, regulating *U2AF2* RNA
829 splicing. The mechanism of the expression of IWS1 is under investigation. The increased
830 activity of the pathway may also occur due to cell cycle-dependent changes in the activity of
831 AKT. Earlier reports have shown that CDK2-Cyclin A2 complex is activated in S phase and
832 phosphorylates AKT at Ser477/Thr479 (Liu et al., 2014[75]), which further enhances the activity
833 of IWS1 pathway. Furthermore, our results show that the epigenetic complexes assembled
834 upon the phosphorylation of IWS1 on *U2AF2* gene, are regulated in a cell-cycle specific
835 manner. These results come in agreement with previous reports stating that spliceosomal
836 assembly is also regulated through the cell-cycle, implying an extra layer of alternative RNA
837 splicing regulation (Karamysheva et al., 2015[92], Hofmann et al., 2010[93]). More importantly, our
838 results reveal that IWS1 phosphorylation epigenetically controls the dynamics of a novel
839 *U2AF2* splice variant through the cell cycle and regulates the levels of Sororin, which
840 subsequently feeds back to the regulation of the cell cycle and tumor proliferation, through
841 enhancement of ERK oncogenic signals (**Fig. 6**).

842 **Interpretation of the same histone modification signal through different epigenetic**
843 **complexes, affecting lung adenocarcinoma development.**

844 Data presented in this report showed that signals originate from AKT phosphorylation
845 may regulate alternative RNA splicing of target genes by different mechanisms, through IWS1
846 phosphorylation. Our previous findings had shown that IWS1 phosphorylation regulates the
847 alternative RNA splicing of *FGFR2* by promoting the exclusion of exon IIIb from the mature

848 transcript (Sanidas et al., 2014[10]). Here we show that, similar to the *FGFR2*, the exon inclusion
849 of *U2AF2* exon 2, is also under the control of the SETD2-dependent Histone H3K36
850 trimethylation in the body of the actively transcribed genes. However, the reader of the Histone
851 H3K36me3 mark is the p52 isoform of LEDGF which interacts with the RNA-binding protein
852 SRSF1, as opposed to the MRG15/PTB complex. Furthermore, we also provide evidence of
853 a more global involvement of SRSF1 in the IWS1 phosphorylation-dependent exon inclusion
854 phenotype (**Fig. S6**). Therefore, although the transduction signal and the regulatory chromatin
855 modification mark is the same for both the exon inclusion and exclusion signals, the effector
856 complexes assembled around H3K36me3, that are responsible for the effects on RNA
857 splicing, differ. Interestingly, both the IWS1 phosphorylation splicing targets affect lung
858 adenocarcinoma development. On the one hand, the *FGFR2* exon exclusion pattern promotes
859 invasion and EMT (Sanidas et al., 2014[10]) and to the other hand, *U2AF2* exon inclusion pattern
860 promotes oncogenic and survival signals, converging towards more aggressive phenotype
861 and clinical profile in lung adenocarcinoma patients, especially with *EGFR* mutations. This
862 isoform specific aggressive clinical phenotype, comes in agreement with recent reports in the
863 literature describing the existence of a cancer-specific splicing addiction network, in which
864 cancer-related isoforms, in our case *U2AF2* and *FGFR2*, manipulate and maintain oncogenic
865 and invasion signals (Bonnal et al., 2020[94], Wang and Aifantis, 2020[95]).

866 **The clinical impact of the p-IWS1/*U2AF2* axis in *EGFR* mutant lung adenocarcinoma.**

867 Throughout this report, we provide evidence that IWS1 phosphorylation, through the
868 epigenetic regulation of *U2AF2* alternative RNA splicing, represents an important oncogenic
869 signal in *EGFR* mutant lung adenocarcinoma. Based on consistent results in cell lines (**Fig.**
870 **6**), animals (**Fig. 7**) and human tumor samples (**Fig. 8**), we showed the major impact of IWS1
871 phosphorylation on proliferation, tumor growth and, eventually, clinical outcome in *EGFR*
872 mutant lung adenocarcinoma patients. Since, the PI3K/AKT axis is triggered by the *EGFR*

873 mutated landscape, and as our data suggest, the activity of the p-IWS1/*U2AF2*/Sororin axis
874 would be increased (**Fig. 8**), further inducing the activation and maintenance of the
875 tumorigenic ERK phosphorylation signals. To this extent, signals derived from the AKT/IWS1
876 phosphorylation pathway facilitates and maintains the oncogenic addiction of these tumors
877 which depends on ERK phosphorylation. Furthermore, our data indicate that IWS1
878 phosphorylation and the RNA splicing pattern of its targets, *U2AF2* and *FGFR2*, selectively
879 correlates not only with tumor grade, stage and patient survival, but also with metastasis and
880 with relapse (**Fig. 8**). Based on these findings, we propose two translational applications for
881 the IWS1 phosphorylation pathway described in this report. First, *EGFR* mutant lung
882 adenocarcinoma patients may benefit from a dual inhibition of EGFR Tyrosine Kinase
883 Inhibitors (TKI) and specific AKT1/3 inhibitors (not currently investigated) or dual AKT/MEK
884 inhibition (MK2206-AZD6244 : NCT01306045), limiting the ERK oncogenic addiction. Based
885 on our recent results, IWS1 phosphorylation also regulates the expression of type I IFNs
886 (Laliotis et al., 2020[96,97]). Furthermore, based on a recent report, the type I IFN signaling is a
887 major determinant of EGFR TKI-sensitivity (Gong et al., 2020[98]). Given that the AKT/p-IWS1
888 axis regulates type I IFN response, the proposed combined EGFR/AKT inhibition could
889 determine the response of *EGFR* mutant LUAD patients through that pathway.

890 Secondly, the IWS1 phosphorylation-dependent cancer splicing addiction can be
891 manipulated by synergistic use of EGFR TKI, inhibitors of the AKT/IWS1 axis and the use of
892 highly isoform-specific antisense oligonucleotides (Splice Switching Oligonucleotides-SSO)
893 and pharmacologic modulators of splicing machinery (Obeng et al., 2019[99]), which are
894 currently under clinical trials (NCT03901469, NCT02711956, NCT02268552, NCT02908685).

895 Collectively, our results suggest that IWS1 phosphorylation by AKT acts as an
896 epigenetic switch that, through H3K36me3-mediated signals, regulates RNA splicing in lung
897 adenocarcinoma maintaining the oncogenic and splicing addiction, and may serve as a novel
898 precision-medicine marker and important drug target in *EGFR* mutant lung adenocarcinoma.

899 **Materials and methods**

900 **Cells and culture conditions, growth factors and inhibitors**

901 Cells, Growth factors, inhibitors, siRNAs and shRNAs, retroviral and lentiviral constructs and
902 experimental protocols are described in detail in the supplemental experimental procedures.
903 Experimental protocols include stimulation with growth factors, transfection with siRNAs,
904 retroviral and lentiviral packaging, and cellular transduction with retroviral and lentiviral
905 constructs.

906 **Cell Proliferation assay, Cell cycle analysis and FACS-sorting of cells in different**
907 **phases of the cell cycle**

908 Cell proliferation of cells growing under normal culture conditions was monitored using the
909 Incucyte S3 Live-Cell Imaging and Analysis System (Essen Biosciences, Ann Arbor, MI).
910 Captured data were analyzed using the native analysis software. The data were normalized
911 to the initial confluence at time point zero. The cell cycle distribution of exponentially growing
912 cells was monitored by FACS analysis of ethanol-fixed, propidium iodide stained, cell cultures.
913 To separate cells in different phases of the cell cycle for further analysis, NCI-H1299 cells
914 were stained with Vybrant™ DyeCycle™ Ruby Stain (Thermo Fisher, Cat. No. V10309) and
915 they were sorted using the BD FACS Aria III cell sorter (BD Biosciences, San Jose, CA). For
916 details on all the above, see Supplemental Experimental Procedures.

917 **Cell transformation assay**

918 Cell transformation assays in immortalized HBEC hTERT were performed using the Cell
919 Transformation Assay Kit-Colorimetric (Abcam Cat No. ab235698), as outlined in
920 Supplemental Experimental Procedures.

921 **RNA-seq analysis**

922 Detailed descriptions of total RNA isolation from NCI-H522 cells, library preparation, RNA
923 sequencing methodology and data analyses are presented in the Supplemental Experimental
924 Procedures

925 **Immunoblotting and Protein Immunoprecipitation**

926 Cells were lysed with RIPA buffer and cell lysates were resolved by electrophoresis in SDS-
927 PAGE and analyzed by immunoblotting. Protein immunoprecipitation was carried out,
928 following established protocols (Mohammed, H. et al, 2016[100]), with modifications. Images
929 were acquired and analyzed, using the Li-Cor Fc Odyssey Imaging System (LI-COR
930 Biosciences, Lincoln, NE). For the lists of antibodies used for immunoprecipitation and western
931 blotting and for other details, see Supplemental Experimental Procedures.

932 **Cloning and site-directed mutagenesis**

933 The cDNA clones used in experiments in this report were purchased from multiple sources.
934 Open reading frames of all the cDNAs were subcloned in lentiviral vectors. Point mutations
935 were introduced into the open reading frames of these clones by site-directed mutagenesis.
936 For details see Supplemental Experimental Procedures.

937

938 **Subcellular Fractionation**

939 Cell pellets were fractionated into nuclear and cytoplasmic fractions, which were used to
940 measure the relative abundance of proteins and RNAs in the nucleus and the cytoplasm. For
941 details, see Supplemental Experimental Procedures.

942 **RT-PCR and qRT-PCR**

943 Total RNA was isolated using the PureLink RNA Kit (Invitrogen, Cat. No 12183018A). Isolated
944 RNA was analyzed by real-time RT-PCR for the expression of the indicated mRNAs. The
945 mRNA levels were normalized based on the levels of *GAPDH* (internal control). The primer
946 sets used are listed in the Supplemental Experimental Procedures.

947 **Chromatin Immunoprecipitation-Sequencing (ChIP-Seq), ChIP and Chromatin Immuno-**
948 **Cleavage (ChIC)**

949 The distribution of chromatin-bound IWS1 and SETD2 and the distribution of histone
950 H3K36me3 genome-wide were addressed by ChIP-Seq experiments. The binding of IWS1
951 and SETD2 and the abundance of histone H3K36me3 on the *U2AF2* and *FGFR2* locus, were
952 addressed by chromatin immunoprecipitation or chromatin Immuno-cleavage (Skene et al.,
953 2018[101]). The latter was employed for experiments with FACS-sorted cells in different phases
954 of the cell cycle. For details, regarding the sorting, the ChIP/ChIC protocols, ChIP-seq library
955 preparation and analysis, see Supplemental Experimental Procedures.

956 **RNA Immunoprecipitation**

957 The binding of RNA binding proteins to regions of the *U2AF2*, *CDCA5* and *GUSB* pre-mRNAs
958 was addressed by RNA Immunoprecipitation. For details, see Supplemental Experimental
959 Procedures.

960 **Tumor xenografts**

961 NOD.*Cg-Prkdc*^{scid} *Il2rg*^{tm1Wjl}/SzJ (NSG) mice were injected subcutaneously with $2x10^6$, $5x10^6$
962 , $1x10^7$ shControl or shIWS1 NCI-H1299, A549 and NCI-H1975 cells, and they were monitored
963 for 4, 6 and 4 weeks, respectively. Tumor metrics, protein and RNA extraction and tumor
964 immunohistochemistry are described in detail in the Supplemental Experimental Procedures.

965 All mouse experiments were reviewed and approved by the Institutional Animal Care and Use
966 Committee (IACUC) of the Ohio State University.

967 **Human Tumor Samples.**

968 Thirty Lung Adenocarcinoma samples with matching normal adjacent tissue were obtained
969 from the Tissue Bank of The Ohio State University. Ten additional lung adenocarcinomas
970 without matching normal tissue had been obtained earlier from the tissue bank of Tufts Medical
971 Center. The latter had been used also in an earlier study on the role of IWS1 in NSCLC
972 (Sanidas et al., 2014[10]). Protein and RNA extraction, and correlations of the abundance of
973 IWS1 and the inclusion of Exon 2 in the *U2AF2* mRNA with tumor stage and patient survival
974 were carried out as described in the Supplemental Experimental Procedures.

975 **Immunohistochemistry**

976 Human lung adenocarcinoma tissue arrays (US Biomax, Cat. No. LC1504) and tumor
977 xenografts from immunodeficient mice, were stained with antibodies. Antibody staining, image
978 acquisition and signal intensity measurements, along with the correlation of IWS1 signal
979 intensity with tumor grade and stage, are described in the Supplemental Experimental
980 Procedures.

981 **TCGA/dataset analysis**

982 TCGA data were downloaded from https://portal.gdc.cancer.gov/ and analysed as described
983 in the Supplemental Experimental Procedures. The microarray and RNA-seq analysis for the
984 publicly available datasets GSE13213, GSE26939 and GSE141685 are outlined in the
985 Supplemental Experimental Procedures.

986 **scRNA-Seq Analysis and t-SNE**

987 The normalized counts from scRNA-Seq data derived from a mouse lung adenocarcinoma
988 metastatic model (GSE123903) were retrieved and analyzed for *IWS1* and *CDCA5*
989 expression, as described before (Laughney et al., 2020[79]). The data were visualized using the
990 Barnes–Hut approximate version of *t*-SNE (Maaten et al., 2008[102])
991 (https://github.com/lvdmaaten/bhtsne)

992 **Data availability**

993 All the source data (uncropped gel images, qPCR, FACS, plates reader and proliferation data),
994 processed counts of RNA-Seq, Junction-Seq splicing platform for visualization of the exon
995 usage results and microscope images derived from this report have been deposited in the
996 Mendeley Dataset in two independent publicly available datasets (Laliotis et al., 2020[103,104]).
997 Raw RNA-Seq and ChIP-Seq data of NCI-H522 cell lines, transduced with lentiviral constructs
998 of shControl, shIWS1, shIWS1/WT-R or shIWS1/MT-R will be deposited in Gene Expression
999 Omnibus (GEO).
1000

1001 **Code availability**

1002 All the code used for the analysis in this report is derived from previously published reports. It
1003 is also explained and cited in the appropriate materials and methods or supplementary
1004 experimental procedures sections.

1005 **Acknowledgments**

1006 The authors wish to thank all the members of the Tsichlis Lab for helpful discussions. We also
1007 thank Drs Carlo M. Croce, Michael Freitas, Wayne Miles, Raphael Pollock, Amanda Toland,
1008 Samir Achaya, Joal Beane and Esmerina Tilli for reviewing the manuscript before the
1009 submission. This work was supported by the NIH grant R01 CA186729 to P.N.T., the NIH
1010 grant R01 CA198117 to P.N.T and V.C, by start-up funds from the OSUCCC to P.N.T,, from

1011 the National Center for Advancing Translational Sciences grant KL2TR002734 to L.S. G.I.L is
1012 supported by the Pelotonia Post-Doctoral fellowship from OSUCCC.

1013 **Author Contributions**

1014 **G.I.L.** Conceptualization, overall experimental design. Performed all the experiments,
1015 analyzed the data, prepared the figures, performed and analyzed the ChIP-seq and wrote the
1016 manuscript. **E.C.** Designed and performed the mouse xenografts, protein and RNA extraction
1017 of mouse tumors, performed the mouse xenografts and human TMA staining studies and
1018 picture acquisition and edited the manuscript **M.D.P.** Bioinformatics analyses of RNA-seq
1019 data. **A.C.** Performed RT-PCR experiments. **A.L.F.** Bioinformatics analyses of Microarray and
1020 TCGA data. **S.S.** Assisted in FACS experiments **V.A.** Analyzed the proliferation experiments.
1021 **K.A.N** Performed RT-PCR experiment and assisted in site-directed mutagenesis **S.A.**
1022 Bioinformatics analyses of scRNA-Seq, Overall Bioinformatics supervision. **A.O.** Extraction of
1023 the tumor samples and edited the manuscript **V.T.** Performed the RNA-seq. **I.V.**
1024 Bioinformatics analyses of RNA-seq data. **M.C.** Extraction of the tumor samples. **A.H.**
1025 Bioinformatics analyses of RNA-seq data. **D.P.** Experimental design and edited the
1026 manuscript. **C.T.** Advised on the design of experiments. **L.S.** Advised on the design of
1027 experiments. **D.P.C.** Advised on the biology of lung cancer and on the design of experiments,
1028 provided cell lines and reagents. **V.C.** Contributed to overall experimental design. **P.N.T.**
1029 Overall experimental design, manuscript writing and editing.

1030 **References**

1031 1. Glisovic, T., Bachorik, J.L., Yong, J. and Dreyfuss, G., 2008. RNA-binding proteins and
1032 post-transcriptional gene regulation. *FEBS letters*, *582*(14), pp.1977-1986.
1033 2. Jewer, M., Findlay, S.D. and Postovit, L.M., 2012. Post-transcriptional regulation in
1034 cancer progression. *Journal of cell communication and signaling*, *6*(4), pp.233-248.

1035 3. Pan, Q., Shai, O., Lee, L.J., Frey, B.J. and Blencowe, B.J., 2008. Deep surveying of
1036 alternative splicing complexity in the human transcriptome by high-throughput
1037 sequencing. *Nature genetics*, *40*(12), p.1413.
1038 4. Paronetto, M.P., Passacantilli, I. and Sette, C., 2016. Alternative splicing and cell
1039 survival: from tissue homeostasis to disease. *Cell Death & Differentiation*, *23*(12),
1040 pp.1919-1929.
1041 5. Zhang, X., Chen, M.H., Wu, X., Kodani, A., Fan, J., Doan, R., Ozawa, M., Ma, J.,
1042 Yoshida, N., Reiter, J.F. and Black, D.L., 2016. Cell-type-specific alternative splicing
1043 governs cell fate in the developing cerebral cortex. *Cell*, *166*(5), pp.1147-1162.
1044 6. Coomer, A.O., Black, F., Greystoke, A., Munkley, J. and Elliott, D.J., 2019. Alternative
1045 splicing in lung cancer. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory*
1046 *Mechanisms*.
1047 7. Siegel, R.L., Miller, K.D. and Jemal, A., 2016. Cancer statistics, 2016. *CA: a cancer*
1048 *journal for clinicians*, *66*(1), pp.7-30.
1049 8. Yuan, M., Huang, L.L., Chen, J.H., Wu, J. and Xu, Q., 2019. The emerging treatment
1050 landscape of targeted therapy in non-small-cell lung cancer. *Signal Transduction and*
1051 *Targeted Therapy*, *4*(1), pp.1-14.
1052 9. Sequist, L.V., Neal, J.W., Jett, J.R. and Ross, M.E., 2016. Personalized, genotype
1053 directed therapy for advanced non-small cell lung cancer. *UpToDate, UpToDate*.
1054 10. Sanidas, I., Polytarchou, C., Hatziapostolou, M., Ezell, S.A., Kottakis, F., Hu, L., Guo,
1055 A., Xie, J., Comb, M.J., Iliopoulos, D. and Tsichlis, P.N., 2014. Phosphoproteomics
1056 screen reveals akt isoform-specific signals linking RNA processing to lung cancer.
1057 *Molecular cell*, *53*(4), pp.577-590.
1058 11. Li, Z., Li, Q., Han, L., Tian, N., Liang, Q., Li, Y., Zhao, X., Du, C. and Tian, Y., 2016.
1059 Pro-apoptotic effects of splice-switching oligonucleotides targeting Bcl-x pre-mRNA in
1060 human glioma cell lines. *Oncology reports*, *35*(2), pp.1013-1019.
1061 12. Todaro, M., Gaggianesi, M., Catalano, V., Benfante, A., Iovino, F., Biffoni, M., Apuzzo,
1062 T., Sperduti, I., Volpe, S., Cocorullo, G. and Gulotta, G., 2014. CD44v6 is a marker of

1063 constitutive and reprogrammed cancer stem cells driving colon cancer metastasis. *Cell*
1064 *stem cell*, *14*(3), pp.342-356.
1065 13. Oltean, S. and Bates, D.O., 2014. Hallmarks of alternative splicing in cancer.
1066 *Oncogene*, *33*(46), pp.5311-5318.
1067 14. Calabretta, S., Bielli, P., Passacantilli, I., Pilozzi, E., Fendrich, V., Capurso, G., Delle
1068 Fave, G. and Sette, C., 2016. Modulation of PKM alternative splicing by PTBP1
1069 promotes gemcitabine resistance in pancreatic cancer cells. *Oncogene*, *35*(16),
1070 pp.2031-2039.
1071 15. Vo, N.S., Reynolds, S.M., Zhao, Y., Climente-González, H., Chai11, S., Wang, F.,
1072 Varghese, R., Huang, M., Liang, W.W., Wyczalkowski, M.A. and Sengupta, S.,
1073 Systematic Analysis of Splice-Site-Creating Mutations in Cancer.
1074 16. Giaj Levra, M., Andre Olaussen, K., Novello, S. and Soria, J.C., 2014. PARP inhibitors:
1075 an interesting pathway also for non-small cell lung cancer?. *Current pharmaceutical*
1076 *design*, *20*(24), pp.3875-3882.
1077 17. Seiler, M., Yoshimi, A., Darman, R., Chan, B., Keaney, G., Thomas, M., Agrawal, A.A.,
1078 Caleb, B., Csibi, A., Sean, E. and Fekkes, P., 2018. H3B-8800, an orally available
1079 small-molecule splicing modulator, induces lethality in spliceosome-mutant cancers.
1080 *Nature medicine*, *24*(4), p.497.
1081 18. Listerman, I., Sapra, A.K. and Neugebauer, K.M., 2006. Cotranscriptional coupling of
1082 splicing factor recruitment and precursor messenger RNA splicing in mammalian cells.
1083 *Nature structural & molecular biology*, *13*(9), p.815.
1084 19. Luco, R.F., Pan, Q., Tominaga, K., Blencowe, B.J., Pereira-Smith, O.M. and Misteli,
1085 T., 2010. Regulation of alternative splicing by histone modifications. *Science*,
1086 *327*(5968), pp.996-1000.
1087 20. Pradeepa, M.M., Sutherland, H.G., Ule, J., Grimes, G.R. and Bickmore, W.A., 2012.
1088 Psip1/Ledgf p52 binds methylated histone H3K36 and splicing factors and contributes
1089 to the regulation of alternative splicing. *PLoS genetics*, *8*(5).

1090 21. Chen, M. and Manley, J.L., 2009. Mechanisms of alternative splicing regulation:
1091 insights from molecular and genomics approaches. *Nature reviews Molecular cell*
1092 *biology*, *10*(11), pp.741-754.
1093 22. Obeng, E.A., Stewart, C. and Abdel-Wahab, O., 2019. Altered RNA Processing in
1094 Cancer Pathogenesis and Therapy. *Cancer discovery*, *9*(11), pp.1493-1510.
1095 23. Martinez-Contreras, R., Cloutier, P., Shkreta, L., Fisette, J.F., Revil, T. and Chabot, B.,
1096 2008. 8 hnRNP proteins and splicing control. *Advances in Experimental Medicine &*
1097 *Biology*, *623*, p.123.
1098 24. de Miguel, F.J., Sharma, R.D., Pajares, M.J., Montuenga, L.M., Rubio, A. and Pio, R.,
1099 2014. Identification of alternative splicing events regulated by the oncogenic factor
1100 SRSF1 in lung cancer. *Cancer research*, *74*(4), pp.1105-1115.
1101 25. Chen, M., David, C.J. and Manley, J.L., 2012. Concentration-dependent control of
1102 pyruvate kinase M mutually exclusive splicing by hnRNP proteins. *Nature structural &*
1103 *molecular biology*, *19*(3), p.346.
1104 26. Jin, W., Bruno, I.G., Xie, T.X., Sanger, L.J. and Cote, G.J., 2003. Polypyrimidine tract-
1105 binding protein down-regulates fibroblast growth factor receptor 1 α-exon inclusion.
1106 *Cancer research*, *63*(19), pp.6154-6157.
1107 27. Dominguez, D., Tsai, Y.H., Weatheritt, R., Wang, Y., Blencowe, B.J. and Wang, Z.,
1108 2016. An extensive program of periodic alternative splicing linked to cell cycle
1109 progression. *Elife*, *5*, p.e10288.
1110 28. Moore, M.J., Wang, Q., Kennedy, C.J. and Silver, P.A., 2010. An alternative splicing
1111 network links cell-cycle control to apoptosis. *Cell*, *142*(4), pp.625-636.
1112 29. Ahn, E.Y., DeKelver, R.C., Lo, M.C., Nguyen, T.A., Matsuura, S., Boyapati, A., Pandit,
1113 S., Fu, X.D. and Zhang, D.E., 2011. SON controls cell-cycle progression by
1114 coordinated regulation of RNA splicing. *Molecular cell*, *42*(2), pp.185-198.
1115 30. Suvorova, E.S., Croken, M., Kratzer, S., Ting, L.M., de Felipe, M.C., Balu, B., Markillie,
1116 M.L., Weiss, L.M., Kim, K. and White, M.W., 2013. Discovery of a splicing regulator
1117 required for cell cycle progression. *PLoS Genet*, *9*(2), p.e1003305.

1118 31. Shen, H. and Green, M.R., 2004. A pathway of sequential arginine-serine-rich domain-
1119 splicing signal interactions during mammalian spliceosome assembly. *Molecular cell*,
1120 *16*(3), pp.363-373.
1121 32. Glasser, E., Agrawal, A.A., Jenkins, J.L. and Kielkopf, C.L., 2017. Cancer-associated
1122 mutations mapped on high-resolution structures of the U2AF2 RNA recognition motifs.
1123 *Biochemistry*, *56*(36), pp.4757-4761.
1124 33. Li, J., Cheng, D., Zhu, M., Yu, H., Pan, Z., Liu, L., Geng, Q., Pan, H., Yan, M. and Yao,
1125 M., 2019. OTUB2 stabilizes U2AF2 to promote the Warburg effect and tumorigenesis
1126 via the AKT/mTOR signaling pathway in non-small cell lung cancer. *Theranostics*, *9*(1),
1127 p.179.
1128 34. Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A.,
1129 Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S. and Mesirov, J.P., 2005. Gene
1130 set enrichment analysis: a knowledge-based approach for interpreting genome-wide
1131 expression profiles. *Proceedings of the National Academy of Sciences*, *102*(43),
1132 pp.15545-15550.
1133 35. Anders, S., Reyes, A. and Huber, W., 2012. Detecting differential usage of exons from
1134 RNA-seq data. *Nature Precedings*, pp.1-1.
1135 36. Hacker, K.E., Fahey, C.C., Shinsky, S.A., Chiang, Y.C.J., DiFiore, J.V., Jha, D.K., Vo,
1136 A.H., Shavit, J.A., Davis, I.J., Strahl, B.D. and Rathmell, W.K., 2016.
1137 Structure/Function analysis of recurrent mutations in SETD2 protein reveals a critical
1138 and conserved role for a SET domain residue in maintaining protein stability and
1139 histone H3 Lys-36 trimethylation. *Journal of Biological Chemistry*, *291*(40), pp.21283-
1140 21295.
1141 37. Greer, E.L. and Shi, Y., 2012. Histone methylation: a dynamic mark in health, disease
1142 and inheritance. *Nature Reviews Genetics*, *13*(5), p.343.
1143 38. Hyun, K., Jeon, J., Park, K. and Kim, J., 2017. Writing, erasing and reading histone
1144 lysine methylation. *Experimental & molecular medicine*, *49*(4), p.e324.

1145 39. Lucio-Eterovic, A.K., Singh, M.M., Gardner, J.E., Veerappan, C.S., Rice, J.C. and
1146 Carpenter, P.B., 2010. Role for the nuclear receptor-binding SET domain protein 1
1147 (NSD1) methyltransferase in coordinating lysine 36 methylation at histone 3 with RNA
1148 polymerase II function. *Proceedings of the National Academy of Sciences*, *107*(39),
1149 pp.16952-16957.
1150 40. Qiao, Q., Li, Y., Chen, Z., Wang, M., Reinberg, D. and Xu, R.M., 2011. The structure
1151 of NSD1 reveals an autoregulatory mechanism underlying histone H3K36 methylation.
1152 *Journal of Biological Chemistry*, *286*(10), pp.8361-8368.
1153 41. Berdasco, M., Ropero, S., Setien, F., Fraga, M.F., Lapunzina, P., Losson, R.,
1154 Alaminos, M., Cheung, N.K., Rahman, N. and Esteller, M., 2009. Epigenetic
1155 inactivation of the Sotos overgrowth syndrome gene histone methyltransferase NSD1
1156 in human neuroblastoma and glioma. *Proceedings of the National Academy of*
1157 *Sciences*, *106*(51), pp.21830-21835.
1158 42. Rahman, S., Sowa, M.E., Ottinger, M., Smith, J.A., Shi, Y., Harper, J.W. and Howley,
1159 P.M., 2011. The Brd4 extra terminal domain confers transcription activation
1160 independent of pTEFb by recruiting multiple proteins, including NSD3. *Molecular and*
1161 *cellular biology*, *31*(13), pp.2641-2652.
1162 43. Brown, M.A., Sims, R.J., Gottlieb, P.D. and Tucker, P.W., 2006. Identification and
1163 characterization of Smyd2: a split SET/MYND domain-containing histone H3 lysine 36-
1164 specific methyltransferase that interacts with the Sin3 histone deacetylase complex.
1165 *Molecular cancer*, *5*(1), p.26.
1166 44. Gregory, G.D., Vakoc, C.R., Rozovskaia, T., Zheng, X., Patel, S., Nakamura, T.,
1167 Canaani, E. and Blobel, G.A., 2007. Mammalian ASH1L is a histone methyltransferase
1168 that occupies the transcribed region of active genes. *Molecular and cellular biology*,
1169 *27*(24), pp.8466-8479.
1170 45. Bannister, A.J., Schneider, R. and Kouzarides, T., 2002. Histone methylation: dynamic
1171 or static?. *Cell*, *109*(7), pp.801-806.

1172 46. Berry, W.L. and Janknecht, R., 2013. KDM4/JMJD2 histone demethylases: epigenetic
1173 regulators in cancer cells. *Cancer research*, *73*(10), pp.2936-2942.
1174 47. Cloos, P.A., Christensen, J., Agger, K., Maiolica, A., Rappsilber, J., Antal, T., Hansen,
1175 K.H. and Helin, K., 2006. The putative oncogene GASC1 demethylates tri-and
1176 dimethylated lysine 9 on histone H3. *Nature*, *442*(7100), p.307.
1177 48. Li, W., Zhao, L., Zang, W., Liu, Z., Chen, L., Liu, T., Xu, D. and Jia, J., 2011. Histone
1178 demethylase JMJD2B is required for tumor cell proliferation and survival and is
1179 overexpressed in gastric cancer. *Biochemical and biophysical research*
1180 *communications*, *416*(3-4), pp.372-378.
1181 49. Cheung, H.C., Hai, T., Zhu, W., Baggerly, K.A., Tsavachidis, S., Krahe, R. and Cote,
1182 G.J., 2009. Splicing factors PTBP1 and PTBP2 promote proliferation and migration of
1183 glioma cell lines. *Brain*, *132*(8), pp.2277-2288.
1184 50. Cai, L., Rothbart, S.B., Lu, R., Xu, B., Chen, W.Y., Tripathy, A., Rockowitz, S., Zheng,
1185 D., Patel, D.J., Allis, C.D. and Strahl, B.D., 2013. An H3K36 methylation-engaging
1186 Tudor motif of polycomb-like proteins mediates PRC2 complex targeting. *Molecular*
1187 *cell*, *49*(3), pp.571-582.
1188 51. Vezzoli, A., Bonadies, N., Allen, M.D., Freund, S.M., Santiveri, C.M., Kvinlaug, B.T.,
1189 Huntly, B.J., Göttgens, B. and Bycroft, M., 2010. Molecular basis of histone H3K36me3
1190 recognition by the PWWP domain of Brpf1. *Nature structural & molecular biology*,
1191 *17*(5), pp.617-619.
1192 52. Li, F., Mao, G., Tong, D., Huang, J., Gu, L., Yang, W. and Li, G.M., 2013. The histone
1193 mark H3K36me3 regulates human DNA mismatch repair through its interaction with
1194 MutSα. *Cell*, *153*(3), pp.590-600.
1195 53. Vermeulen, M., Eberl, H.C., Matarese, F., Marks, H., Denissov, S., Butter, F., Lee,
1196 K.K., Olsen, J.V., Hyman, A.A., Stunnenberg, H.G. and Mann, M., 2010. Quantitative
1197 interaction proteomics and genome-wide profiling of epigenetic histone marks and their
1198 readers. *Cell*, *142*(6), pp.967-980.

1199 54. Shun, M.C., Botbol, Y., Li, X., Di Nunzio, F., Daigle, J.E., Yan, N., Lieberman, J.,
1200 Lavigne, M. and Engelman, A., 2008. Identification and characterization of PWWP
1201 domain residues critical for LEDGF/p75 chromatin binding and human
1202 immunodeficiency virus type 1 infectivity. *Journal of virology*, *82*(23), pp.11555-
1203 11567.Singh, P.K., Plumb, M.R., Ferris, A.L., Iben, J.R., Wu, X., Fadel, H.J., Luke,
1204 B.T., Esnault, C., Poeschla, E.M., Hughes, S.H. and Kvaratskhelia, M., 2015.
1205 LEDGF/p75 interacts with mRNA splicing factors and targets HIV-1 integration to
1206 highly spliced genes. *Genes & development*, *29*(21), pp.2287-2297.
1207 55. Ferris, A.L., Wu, X., Hughes, C.M., Stewart, C., Smith, S.J., Milne, T.A., Wang, G.G.,
1208 Shun, M.C., Allis, C.D., Engelman, A. and Hughes, S.H., 2010. Lens epithelium-
1209 derived growth factor fusion proteins redirect HIV-1 DNA integration. *Proceedings of*
1210 *the National Academy of Sciences*, *107*(7), pp.3135-3140.
1211 56. Paz, I., Kosti, I., Ares Jr, M., Cline, M. and Mandel-Gutfreund, Y., 2014. RBPmap: a
1212 web server for mapping binding sites of RNA-binding proteins. *Nucleic acids research*,
1213 *42*(W1), pp.W361-W367.
1214 57. Anczuków, O., Akerman, M., Cléry, A., Wu, J., Shen, C., Shirole, N.H., Raimer, A.,
1215 Sun, S., Jensen, M.A., Hua, Y. and Allain, F.H.T., 2015. SRSF1-regulated alternative
1216 splicing in breast cancer. *Molecular cell*, *60*(1), pp.105-117.
1217 58. Millevoi, S., Loulergue, C., Dettwiler, S., Karaa, S.Z., Keller, W., Antoniou, M. and
1218 Vagner, S., 2006. An interaction between U2AF 65 and CF Im links the splicing and 3′
1219 end processing machineries. *The EMBO journal*, *25*(20), pp.4854-4864.
1220 59. Shen, H. and Green, M.R., 2004. A pathway of sequential arginine-serine-rich domain-
1221 splicing signal interactions during mammalian spliceosome assembly. *Molecular cell*,
1222 *16*(3), pp.363-373.
1223 60. Chan, S.P., Kao, D.I., Tsai, W.Y. and Cheng, S.C., 2003. The Prp19p-associated
1224 complex in spliceosome activation. *Science*, *302*(5643), pp.279-282.

1225 61. Chan, S.P. and Cheng, S.C., 2005. The Prp19-associated complex is required for
1226 specifying interactions of U5 and U6 with pre-mRNA during spliceosome activation.
1227 *Journal of Biological Chemistry*, *280*(35), pp.31190-31199.
1228 62. Hogg, R., McGrail, J.C. and O'Keefe, R.T., 2010. The function of the NineTeen
1229 Complex (NTC) in regulating spliceosome conformations and fidelity during pre-mRNA
1230 splicing.
1231 63. Chanarat, S. and Sträßer, K., 2013. Splicing and beyond: the many faces of the Prp19
1232 complex. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, *1833*(10),
1233 pp.2126-2134.
1234 64. David, C.J., Boyne, A.R., Millhouse, S.R. and Manley, J.L., 2011. The RNA polymerase
1235 II C-terminal domain promotes splicing activation through recruitment of a U2AF65–
1236 Prp19 complex. *Genes & development*, *25*(9), pp.972-983.
1237 65. Watrin, E., Demidova, M., Watrin, T., Hu, Z. and Prigent, C., 2014. Sororin pre-mRNA
1238 splicing is required for proper sister chromatid cohesion in human cells. *EMBO reports*,
1239 *15*(9), pp.948-955.
1240 66. Wang, J., Xia, C., Pu, M., Dai, B., Yang, X., Shang, R., Yang, Z., Zhang, R., Tao, K.
1241 and Dou, K., 2018. Silencing of CDCA5 inhibits cancer progression and serves as a
1242 prognostic biomarker for hepatocellular carcinoma. *Oncology reports*, *40*(4), pp.1875-
1243 1884.
1244 67. Shen, A., Liu, L., Chen, H., Qi, F., Huang, Y., Lin, J., Sferra, T.J., Sankararaman, S.,
1245 Wei, L., Chu, J. and Chen, Y., 2019. Cell division cycle associated 5 promotes
1246 colorectal cancer progression by activating the ERK signaling pathway. *Oncogenesis*,
1247 *8*(3), p.19.
1248 68. Nguyen, M.H., Koinuma, J., Ueda, K., Ito, T., Tsuchiya, E., Nakamura, Y. and Daigo,
1249 Y., 2010. Phosphorylation and activation of cell division cycle associated 5 by mitogen-
1250 activated protein kinase play a crucial role in human lung carcinogenesis. *Cancer*
1251 *research*, *70*(13), pp.5337-5347

1252 69. Guo, Y.J., Pan, W.W., Liu, S.B., Shen, Z.F., Xu, Y. and Hu, L.L., 2020. ERK/MAPK
1253 signalling pathway and tumorigenesis. *Experimental and Therapeutic Medicine*, *19*(3),
1254 pp.1997-2007.
1255 70. Wee, P. and Wang, Z., 2017. Epidermal growth factor receptor cell proliferation
1256 signaling pathways. *Cancers*, *9*(5), p.52.
1257 71. Kent, L.N. and Leone, G., 2019. The broken cycle: E2F dysfunction in cancer. *Nature*
1258 *Reviews Cancer*, *19*(6), pp.326-338.
1259 72. Begum, J., Day, W., Henderson, C., Purewal, S., Cerveira, J., Summers, H., Rees, P.,
1260 Davies, D. and Filby, A., 2013. A method for evaluating the use of fluorescent dyes to
1261 track proliferation in cell lines by dye dilution. *Cytometry Part A*, *83*(12), pp.1085-1095.
1262 73. Dronamraju, R., Jha, D.K., Eser, U., Adams, A.T., Dominguez, D., Choudhury, R.,
1263 Chiang, Y.C., Rathmell, W.K., Emanuele, M.J., Churchman, L.S. and Strahl, B.D.,
1264 2018. Set2 methyltransferase facilitates cell cycle progression by maintaining
1265 transcriptional fidelity. *Nucleic acids research*, *46*(3), pp.1331-1344.
1266 74. Liu, P., Begley, M., Michowski, W., Inuzuka, H., Ginzberg, M., Gao, D., Tsou, P., Gan,
1267 W., Papa, A., Kim, B.M. and Wan, L., 2014. Cell-cycle-regulated activation of Akt
1268 kinase by phosphorylation at its carboxyl terminus. *Nature*, *508*(7497), pp.541-545.
1269 75. Kim, C.H., Kim, S.H., Park, S.Y., Yoo, J., Kim, S.K. and Kim, H.K., 2015. Identification
1270 of EGFR mutations by immunohistochemistry with EGFR mutation–specific antibodies
1271 in biopsy and resection specimens from pulmonary adenocarcinoma. *Cancer research*
1272 *and treatment: official journal of Korean Cancer Association*, *47*(4), p.653.
1273 76. Popper, H.H., 2016. Progression and metastasis of lung cancer. *Cancer and*
1274 *Metastasis Reviews*, *35*(1), pp.75-91.
1275 77. West, H.J., Shaw, A.T. and Schild, S.E., 2018. Brain metastases in non-small cell lung
1276 cancer.
1277 78. Tomida, S., Takeuchi, T., Shimada, Y., Arima, C., Matsuo, K., Mitsudomi, T., Yatabe,
1278 Y. and Takahashi, T., 2009. Relapse-related molecular signature in lung

1279 adenocarcinomas identifies patients with dismal prognosis. *Journal of clinical*
1280 *oncology*, *27*(17), pp.2793-2799.
1281 79. Laughney, A.M., Hu, J., Campbell, N.R., Bakhoum, S.F., Setty, M., Lavallée, V.P., Xie,
1282 Y., Masilionis, I., Carr, A.J., Kottapalli, S. and Allaj, V., 2020. Regenerative lineages
1283 and immune-mediated pruning in lung cancer metastasis. *Nature Medicine*, *26*(2),
1284 pp.259-269.
1285 80. Cerami, E., Gao, J., Dogrusoz, U., Gross, B.E., Sumer, S.O., Aksoy, B.A., Jacobsen,
1286 A., Byrne, C.J., Heuer, M.L., Larsson, E. and Antipin, Y., 2012. The cBio cancer
1287 genomics portal: an open platform for exploring multidimensional cancer genomics
1288 data.
1289 81. Ding, L., Getz, G., Wheeler, D.A., Mardis, E.R., McLellan, M.D., Cibulskis, K.,
1290 Sougnez, C., Greulich, H., Muzny, D.M., Morgan, M.B. and Fulton, L., 2008. Somatic
1291 mutations affect key pathways in lung adenocarcinoma. *Nature*, *455*(7216), pp.1069-
1292 1075.
1293 82. Imielinski, M., Berger, A.H., Hammerman, P.S., Hernandez, B., Pugh, T.J., Hodis, E.,
1294 Cho, J., Suh, J., Capelletti, M., Sivachenko, A. and Sougnez, C., 2012. Mapping the
1295 hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell*, *150*(6),
1296 pp.1107-1120.
1297 83. Rizvi, N.A., Hellmann, M.D., Snyder, A., Kvistborg, P., Makarov, V., Havel, J.J., Lee,
1298 W., Yuan, J., Wong, P., Ho, T.S. and Miller, M.L., 2015. Mutational landscape
1299 determines sensitivity to PD-1 blockade in non–small cell lung cancer. *Science*,
1300 *348*(6230), pp.124-128.
1301 84. Jordan, E.J., Kim, H.R., Arcila, M.E., Barron, D., Chakravarty, D., Gao, J., Chang, M.T.,
1302 Ni, A., Kundra, R., Jonsson, P. and Jayakumaran, G., 2017. Prospective
1303 comprehensive molecular characterization of lung adenocarcinomas for efficient
1304 patient matching to approved and emerging therapies. *Cancer discovery*, *7*(6), pp.596-
1305 609.

1306 85. Chen, J., Yang, H., Teo, A.S.M., Amer, L.B., Sherbaf, F.G., Tan, C.Q., Alvarez, J.J.S.,
1307 Lu, B., Lim, J.Q., Takano, A. and Nahar, R., 2020. Genomic landscape of lung
1308 adenocarcinoma in East Asians. *Nature Genetics*, *52*(2), pp.177-186.
1309 86. Cerami, E., Gao, J., Dogrusoz, U., Gross, B.E., Sumer, S.O., Aksoy, B.A., Jacobsen,
1310 A., Byrne, C.J., Heuer, M.L., Larsson, E. and Antipin, Y., 2012. The cBio cancer
1311 genomics portal: an open platform for exploring multidimensional cancer genomics
1312 data.
1313 87. Gao, J., Aksoy, B.A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S.O., Sun, Y.,
1314 Jacobsen, A., Sinha, R., Larsson, E. and Cerami, E., 2013. Integrative analysis of
1315 complex cancer genomics and clinical profiles using the cBioPortal. *Science signaling*,
1316 *6*(269), pp.pl1-pl1.
1317 88. Maji, D., Glasser, E., Henderson, S., Galardi, J., Pulvino, M.J., Jenkins, J.L. and
1318 Kielkopf, C.L., 2020. Representative cancer-associated U2AF2 mutations alter RNA
1319 interactions and splicing. *Journal of Biological Chemistry*, pp.jbc-RA120.
1320 89. Ishiguro, K.I., 2019. The cohesin complex in mammalian meiosis. *Genes to Cells*,
1321 *24*(1), pp.6-30.
1322 90. Dreier, M.R., Bekier, M.E. and Taylor, W.R., 2011. Regulation of sororin by Cdk1-
1323 mediated phosphorylation. *Journal of cell science*, *124*(17), pp.2976-2987.
1324 91. Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L.,
1325 Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S. and Taatjes, D.J., 2010.
1326 Mediator and cohesin connect gene expression and chromatin architecture. *Nature*,
1327 *467*(7314), pp.430-435.
1328 92. Karamysheva, Z., Díaz-Martínez, L.A., Warrington, R. and Yu, H., 2015. Graded
1329 requirement for the spliceosome in cell cycle progression. *Cell Cycle*, *14*(12), pp.1873-
1330 1883.
1331 93. Hofmann, J.C., Husedzinovic, A. and Gruss, O.J., 2010. The function of spliceosome
1332 components in open mitosis. *Nucleus*, *1*(6), pp.447-459.

1333 94. Bonnal, S.C., López-Oreja, I. and Valcárcel, J., 2020. Roles and mechanisms of
1334 alternative splicing in cancer—implications for care. *Nature Reviews Clinical Oncology*,
1335 pp.1-18.
1336 95. Wang, E. and Aifantis, I., 2020. RNA Splicing and Cancer. *Trends in Cancer*.
1337 96. Georgios I. Laliotis, Adam D. Kenney, Evangelia Chavdoula, Arturo Orlacchio, Abdul
1338 K. Kaba, Alessandro La Ferlita, Vollter Anastas, Christos Tsatsanis, Joal D. Beane,
1339 Lalit Sehgal, Vincenzo Coppola, Jacob S. Yount, Philip N. Tsichlis, 2020, 'IWS1
1340 phosphorylation by AKT kinase controls the nucleocytoplasmic export of type I IFNs
1341 and the sensitivity of lung adenocarcinoma cells to oncolytic viral infection, through
1342 U2AF2 RNA splicing', bioRxiv 2020.12.26.424461; doi:
1343 https://doi.org/10.1101/2020.12.26.424461
1344 97. Laliotis, G.I., Kenney, A.D., Orlacchio, A., Coppola, V., Yount, J.S. and Tsichlis, P.N.,
1345 2020. Abstract LB-183: Sensitivity of cancer cells to oncolytic viruses is defined by
1346 IWS1 phosphorylation dependent epigenetic regulation of U2AF2 splicing and
1347 nucleocytoplasmic export of type I IFN transcripts.
1348 98. Gong, K., Guo, G., Panchani, N., Bender, M.E., Gerber, D.E., Minna, J.D., Fattah, F.,
1349 Gao, B., Peyton, M., Kernstine, K. and Mukherjee, B., 2020. EGFR inhibition triggers
1350 an adaptive response by co-opting antiviral signaling pathways in lung cancer. *Nature*
1351 *Cancer*, *1*(4), pp.394-409.
1352 99. Obeng, E.A., Stewart, C. and Abdel-Wahab, O., 2019. Altered RNA Processing in
1353 Cancer Pathogenesis and Therapy. *Cancer discovery*, *9*(11), pp.1493-1510.
1354 100. Mohammed, H., Taylor, C., Brown, G.D., Papachristou, E.K., Carroll, J.S. and
1355 D'santos, C.S., 2016. Rapid immunoprecipitation mass spectrometry of endogenous
1356 proteins (RIME) for analysis of chromatin complexes. *Nature protocols*, *11*(2), p.316.
1357 101. Skene, P.J., Henikoff, J.G. and Henikoff, S., 2018. Targeted in situ genome-
1358 wide profiling with high efficiency for low cell numbers. *Nature protocols*, *13*(5), p.1006.
1359 102. Maaten, L.V.D. and Hinton, G., 2008. Visualizing data using t-SNE. *Journal of*
1360 *machine learning research*, *9*(Nov), pp.2579-2605.

1361 103. Laliotis, Georgios I.; Chavdoula, Evangelia; Paraskevopoulou, Maria D.; Kaba,
1362 Abdul; La Ferlita, Alessandro; Singh, Satish; Anastas, Vollter; Nair II, Keith A.;
1363 Orlacchio, Arturo; Taraslia, Vasiliki; Vlachos, Ioannis; Capece, Marina; Hatzigeorgiou,
1364 Artemis; Palmieri, Dario; Tsatsanis, Christos; Alaimo, Salvatore; Sehgal, Lalit;
1365 Carbone, David P.; Coppola, Vincenzo; Tsichlis, Philip N. (2020), "IWS1
1366 phosphorylation promotes cell proliferation and predicts poor prognosis in EGFR
1367 mutant lung adenocarcinoma patients, through the cell cycle-regulated U2AF2 RNA
1368 splicing.", Mendeley Data, V4, **DOI:** 10.17632/y82yx8gfkj.4
1369 104. Laliotis, Georgios; Chavdoula, Evangelia; Tsichlis, Philip N. (2020), "IWS1
1370 phosphorylation promotes cell proliferation and predicts poor prognosis in EGFR
1371 mutant lung adenocarcinoma patients, through the cell cycle-regulated U2AF2 RNA
1372 splicing_images_depository", Mendeley Data, V1, doi:10.17632/ts44bfg45k.1
1373

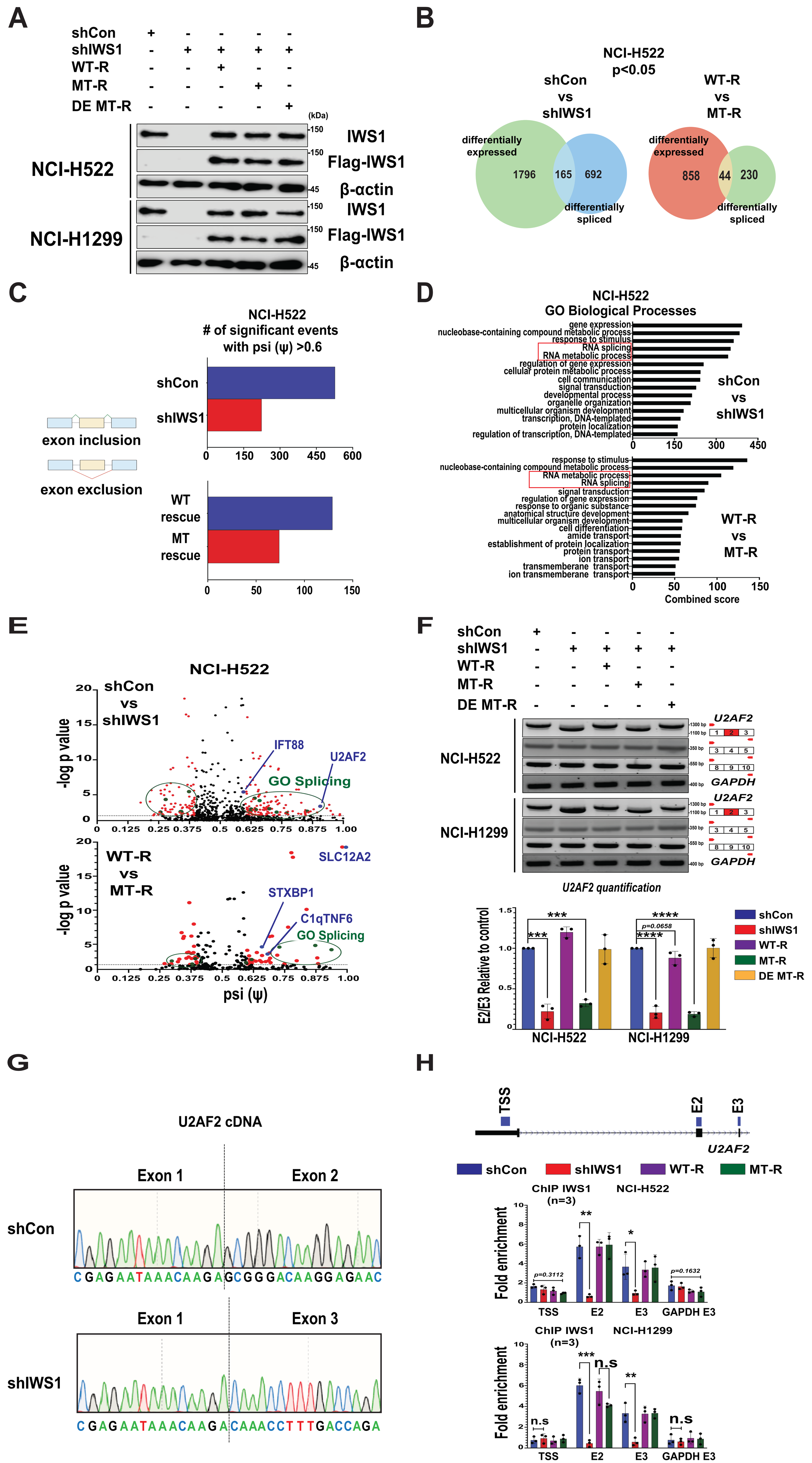

A
shCon
shIWS1
WT-R
MT-R
DE MT-R
(kDa)
NCI-H522
NCI-H1299
IWS1
Flag-IWS1
β-αctin
B
NCI-H522
p<0.05
shCon vs shIWS1
WT-R vs MT-R
differentially expressed
differentially spliced
1796
165
692
858
44
230
C
NCI-H522
# of significant events with psi (ψ) >0.6
shCon
shIWS1
exon inclusion
exon exclusion
WT rescue
MT rescue
D
NCI-H522
GO Biological Processes
shCon vs shIWS1
WT-R vs MT-R
Combined score
E
NCI-H522
shCon vs shIWS1
WT-R vs MT-R
-log p value
psi (ψ)
IFT88
U2AF2
GO Splicing
SLC12A2
STXBP1
C1qTNF6
F
U2AF2
GAPDH
U2AF2 quantification
E2/E3 Relative to control
shCon
shIWS1
WT-R
MT-R
DE MT-R
NCI-H522
NCI-H1299
G
U2AF2 cDNA
Exon 1
Exon 2
Exon 3
shCon
shIWS1
CGAGAATAAACAAGAGCGGGACAAGGAGAAC
CGAGAATAAACAAGACAAACCTTTGACCAGA
H
TSS
E2
E3
U2AF2
shCon
shIWS1
WT-R
MT-R
ChIP IWS1 (n=3)
NCI-H522
NCI-H1299
Fold enrichment
p=0.3112
p=0.1632
n.s
GAPDH E3

Figure 1. IWS1 expression and/or phosphorylation regulate alternative mRNA splicing

A. Western blots of lysates of NCI-H522 and NCI-H1299 cells, transduced with the indicated constructs, were probed with anti-IWS1, anti-Flag-tag and anti-β-actin (control) antibodies.

B. Overlaps between differentially-expressed and differentially spliced genes in shControl and shIWS1 or shIWS1/WT-R and shIWS1/MT-R NCI-H522 cells ($q<0.05$) are illustrated, using Venn diagrams.

C. Bar graphs show the number of alternative splicing events with exon inclusion with a percentage spliced in (psi/ψ) >0.6 in shControl versus shIWS1 (upper panel) and shIWS1/WT-R versus shIWS1/MT-R (lower panel) NCI-H522 cells. The comparisons were limited to alternative splicing events with a p value < 0.05. This shows that IWS1 phosphorylation promotes exon inclusion (see diagram on the left) in multiple alternatively spliced genes.

D. GO analysis, based on differences in the abundance of alternative splicing events in shControl versus shIWS1 (upper panel) and shIWS1/WT-R versus shIWS1/MT-R (lower panel) NCI-H522 cells. Comparisons were limited to alternative splicing events whose abundance changes significantly with the expression or phosphorylation of IWS1($p<0.05$). Red boxes highlight gene sets involved in the regulation of RNA processing.

E. Volcano plots of the -log p value vs the exon inclusion levels of all the exon inclusion and exon exclusion alternative splicing events, detected by DEXseq in the comparisons between shControl and shIWS1 (upper panel), or shIWS1/WT-R and shIWS1/MT-R (lower panel) NCI-H522 cells. The statistically significant events ($p<0.05$) with a percentage spliced in (psi/ψ) level of >0.6 or <0.4 are shown in red. Statistically significant events in genes in the GO functions RNA splicing or RNA metabolic processes are shown in green. Alternatively-spliced IWS1 targets validated in this report are shown in blue.

(*U2AF2, IFT88, SLC12A2, STXBP1, C1qTNF6)*. The validation of these through RT-PCR is shown in Figure S1.

F. *IWS1 phosphorylation regulates U2AF2 alternative mRNA splicing.* (Upper panel) RT-PCR of *U2AF2*, using oligonucleotide primers that map in exons 1 and 3, exons 3 and 5 or exons 8 and 10 (*U2AF2* expression control). RNA transcripts containing or lacking exon 2 are distinguished based on the size of the amplified cDNA derived from these transcripts. RT-PCR was carried out using RNA derived from shControl, shIWS1 shIWS1/WT-R, shIWS1/MT-R and shIWS1/DE MT-R NCI-H522 and NCI-H1299 cells. *hGAPDH* was used as the loading control. (Lower panel) The relative abundance of the amplified cDNA bands in the upper panel was determined, and the ratio between exons E2 and E3 was calculated relative to the ratio in shControl cells, which was given the arbitrary value of 1.

G. Sequencing chromatograms showing the junction between exons 1 and 2 and exons 1 and 3 in the two alternatively spliced *U2AF2* RNA transcripts.

H. *IWS1 binds exons 2 and 3, but not the transcription start site (TSS) of the U2AF2 gene.* (Upper) UCSC browser snapshot showing exons 1, 2 and 3 of the human *U2AF2* gene. The map position of the PCR primer sets used in the ChIP experiments in this figure is indicated by blue marks. (Lower) ChIP assays addressing the binding of IWS1 on the *U2AF2* and *GAPDH* genes in shControl, shIWS1 shIWS1/WT-R and shIWS1/MT-R NCI-H522 and NCI-H1299 cells. Bars show the mean fold enrichment (anti-IWS1 IP, vs IgG control IP) in IWS1 binding, in shIWS1 relative to shControl cells or in shIWS1/MT-R relative to shIWS1/WT-R cells ±SD. Data were normalized relative to the input (2%)

All assays were done in triplicate, on three biological replicates. n.s : non-significant *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001. (one-side unpaired t-test)

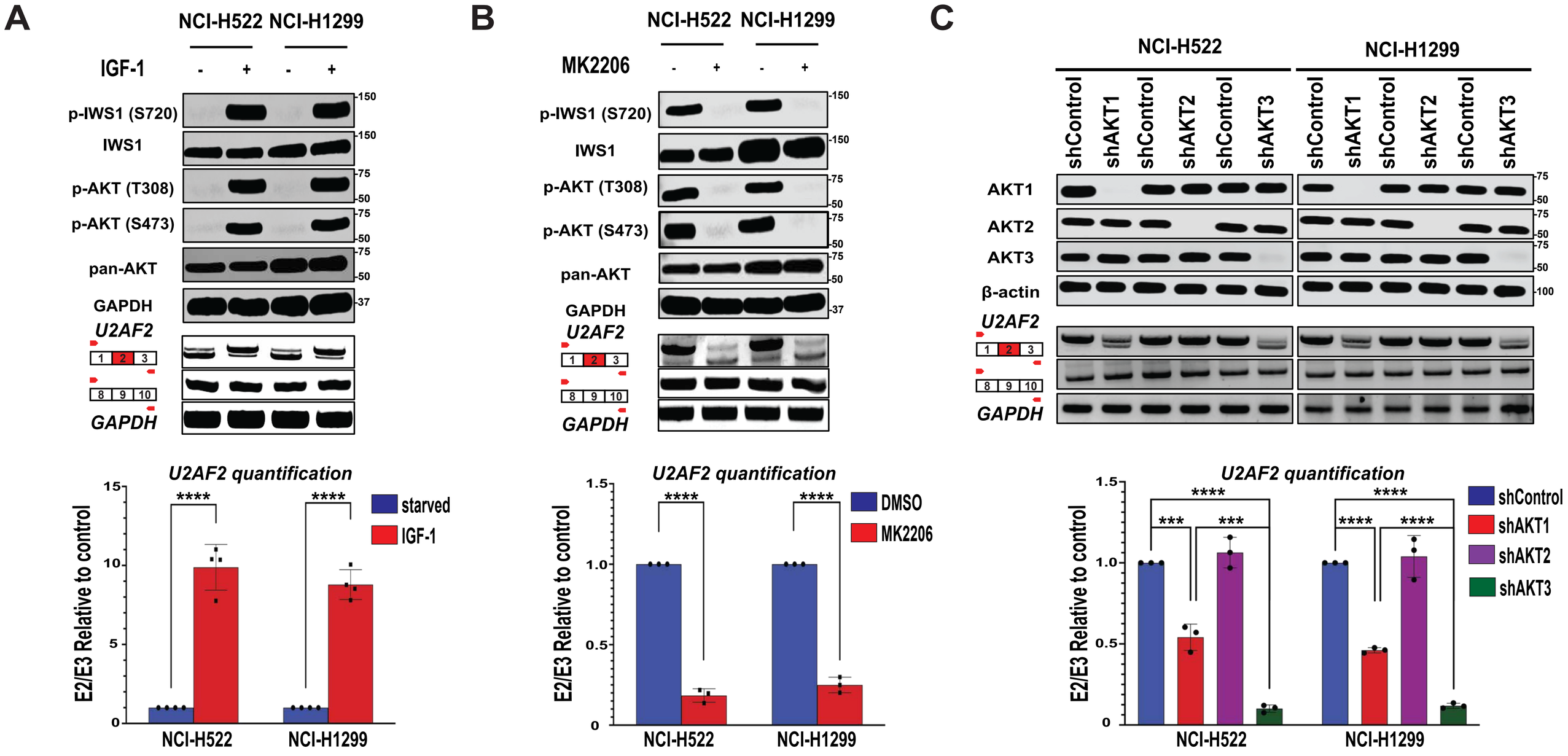

A
NCI-H522 NCI-H1299
IGF-1 - + - +
p-IWS1 (S720)
IWS1
p-AKT (T308)
p-AKT (S473)
pan-AKT
GAPDH
U2AF2
GAPDH
U2AF2 quantification
E2/E3 Relative to control
starved
IGF-1
NCI-H522
NCI-H1299
B
NCI-H522 NCI-H1299
MK2206 - + - +
p-IWS1 (S720)
IWS1
p-AKT (T308)
p-AKT (S473)
pan-AKT
GAPDH
U2AF2
GAPDH
U2AF2 quantification
E2/E3 Relative to control
DMSO
MK2206
NCI-H522
NCI-H1299
C
NCI-H522
NCI-H1299
shControl
shAKT1
shAKT2
shAKT3
AKT1
AKT2
AKT3
β-actin
U2AF2
GAPDH
U2AF2 quantification
E2/E3 Relative to control
shControl
shAKT1
shAKT2
shAKT3
NCI-H522
NCI-H1299

Figure 2. IWS1 phosphorylation-dependent mRNA splicing of *U2AF2* is regulated by serum and IGF-1-induced signals transduced by AKT3.

A. *IGF-1 promotes the IWS1-dependent U2AF2 exon 2 (E2) inclusion.* (Upper panel) Following serum starvation for 24 hours, NCI-H522 and NCI-H1299 cells were stimulated with IGF-1 or harvested. Four hours later, the stimulated cells were lysed and all the lysates were probed with the indicated antibodies. (Middle panel) RT-PCR with mRNA derived from the cells shown in the upper panel and oligonucleotide primers mapping to exons 1 and 3 of the *U2AF2* gene, show that IGF1-induced signals promote *U2AF2* exon 2 inclusion. (Lower panel) The *U2AF2* gene *E2/E3* ratio was calculated following quantification of the RT-PCR products in the middle panel. The bars show this ratio (mean± SD) in IGF1-stimulated NCI-H522 and NCI-H1299 cells relative to the ratio in untreated serum-starved cells.

B. *U2AF2 exon 2 (E2) inclusion depends on AKT kinase.* (Upper panel) NCI-H522 and NCI-H1299 cell lysates, harvested following a 4 hr treatment with MK2206 (5μM) or DMSO, were probed with the indicated antibodies. (Middle panel) RT-PCR reactions, using mRNA derived from the cells in the upper panel and oligonucleotide primers mapping in *U2AF2* exons 1 and 3 show that inhibiting AKT inhibits the inclusion of exon 2 in mature *U2AF2* mRNA transcripts in both cell lines. (Lower panel) The *U2AF2* mRNA *E2/E3* ratio was calculated following quantification of the RT-PCR products in the middle panel. The bars show this ratio (mean± SD) in MK2206-treated (5μM) NCI-H522 and NCI-H1299 cells relative to the ratio in untreated cells.

C. *U2AF2 exon 2 (E2) inclusion depends on AKT3***.** (Upper panel) NCI-H522 and NCI-H1299 cell lysates, transduced with shControl, shAKT1, shAKT2 and shAKT3 were probed with the indicated antibodies. (Middle panel) RT-PCR reactions, using mRNA derived from the cells in the upper panel and oligonucleotide primers mapping in *U2AF2* exons 1 and 3 show that knocking down AKT3, and to a lesser extent AKT1, inhibits the inclusion of exon 2 in mature

U2AF2 mRNA transcripts in both cell lines. (Lower panel) The *U2AF2* mRNA *E2/E3* ratio was calculated following quantification of the RT-PCR products in the middle panel. The bars show this ratio (mean± SD) in shAKT1, shAKT2 and shAKT3 NCI-H522 and NCI-H1299 cells relative to the ratio in shControl cells.

All experiments in this figure were done in triplicate, on three biological replicates. n.s: non-significant *$p<0.05$, **$p<0.01$, ***$p<0.001$, ****$p<0.0001$ (one-side unpaired t-test).

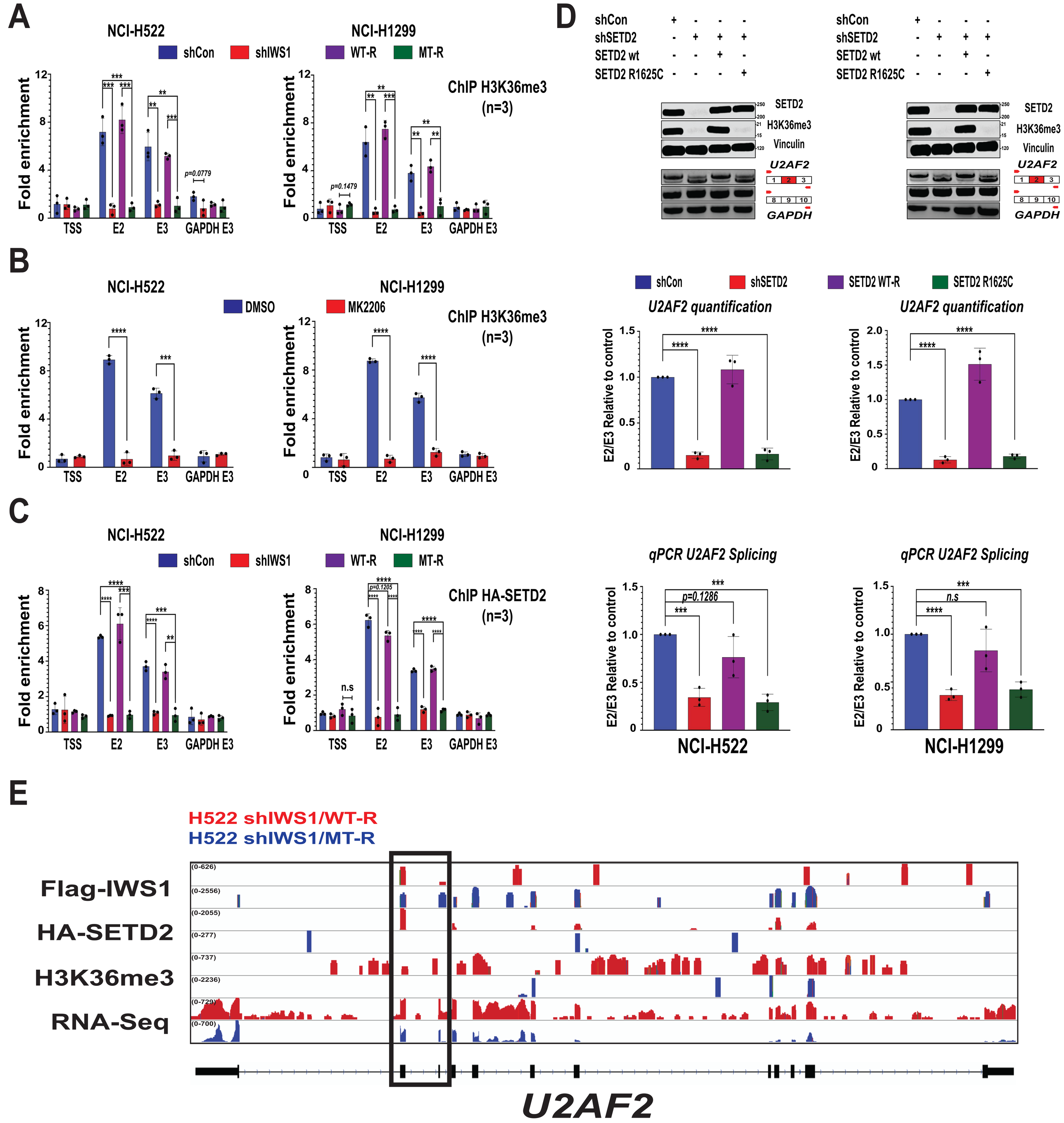

A
NCI-H522
NCI-H1299
shCon
shIWS1
WT-R
MT-R
ChIP H3K36me3
(n=3)
Fold enrichment
TSS
E2
E3
GAPDH E3
p=0.0779
p=0.1479
B
DMSO
MK2206
C
ChIP HA-SETD2
(n=3)
p=0.1205
n.s
D
shCon
shSETD2
SETD2 wt
SETD2 R1625C
SETD2
H3K36me3
Vinculin
U2AF2
GAPDH
SETD2 WT-R
U2AF2 quantification
E2/E3 Relative to control
qPCR U2AF2 Splicing
p=0.1286
E
H522 shIWS1/WT-R
H522 shIWS1/MT-R
Flag-IWS1
HA-SETD2
H3K36me3
RNA-Seq
U2AF2

Figure 3. *U2AF2* Exon 2 inclusion, induced by IWS1 phosphorylation at Ser720/Thr721, depends on H3K36 trimethylation by SETD2.

A. ChIP assays showing the abundance of H3K36me3 on the *U2AF2* and *GAPDH* genes in shControl, shIWS1 shIWS1/WT-R and shIWS1/MT-R NCI-H522 (left) and NCI-H1299 cells (right). Bars show the mean fold enrichment in H3K36me3 (anti-H3K36me3 IP, vs IgG control IP) in the indicated regions of the *U2AF2* gene, in shControl, shIWS1, shIWS1/WT-R and shIWS1/MT-R cells ±SD. Data were normalized relative to the input (2%).

B. *AKT inhibition interferes with the H3K36me3 trimethylation in the U2AF2 gene*. ChIP assays showing the abundance of H3K36me3 on the *U2AF2* and *GAPDH* genes in NCI-H522 (left) and NCI-H1299 cells (right), before and after the treatment with MK2206 (5μM) or DMSO. The bars show the mean fold enrichment of H3K36me3 (anti-H3K36me3 IP, vs IgG control IP) in the indicated regions of the *U2AF2* gene, in shControl, shIWS1 shIWS1/WT-R and shIWS1/MT-R cells ±SD. Data were normalized relative to the input (2%).

C. ChIP assays showing the binding of HA-SETD2 to the indicated regions of the *U2AF2* and *GAPDH* genes, in shControl, shIWS1 shIWS1/WT-R and shIWS1/MT-R NCI-H522 (left) and NCI-H1299 cells (right), transduced with a lentiviral HA-SETD2 construct. The bars show the mean fold enrichment in SETD2 binding (anti-HA IP, vs IgG control IP) in the indicated regions of the *U2AF2* gene, in shControl, shIWS1, shIWS1/WT-R and shIWS1/MT-R cells ±SD. Data were normalized relative to the input (2%). The expression of HA-SETD2 is presented in Figure S4A.

D. *The IWS1-regulated alternative RNA splicing of U2AF2, depends on enzymatically active SETD2*. (Upper panel) Lysates of NCI-H522 (left) and NCI-H1299 (right) cells transduced with shControl or shSETD2 lentiviral constructs and shSETD2 cells rescued with wild type

SETD2, or the catalytically inactive SETD2 R1625C mutant. were probed with the indicated antibodies. RT-PCR, using RNA derived from these cells and oligonucleotide primers mapping in *U2AF2* exons 1 and 3, revealed that the inclusion of exon 2 in the mature U2AF2 mRNA transcripts depends on the expression of enzymatically active SETD2. (Middle panel) The RT-PCR products in the experiment in the upper panel were quantified. Bars show the E2/E3 ratio relative to the shControl, which was given the value of 1. (Lower panel) The ratio of the *U2AF2* exons 2 and 3 was determined by quantitative RT-PCR, using the mRNA isolated from the cells in the upper panel. Again, bars show the E2/E3 ratio relative to the shControl.

Error bars indicate SD. All experiments in this figure were done on three biological replicates, in triplicate. n.s: non-significant *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001 (one-side unpaired t-test).

E. Snapshots of the integrative genomic viewer showing the distribution of Flag-IWS1, HA-SETD2, and histone H3K36me3 marks, within the *U2AF2* gene, obtained from a ChIP-Seq analysis, along with the distribution of RNA reads obtained from the RNA-seq analysis of shIWS1/WT-R and shIWS1/MT-R NCI-H522 cells. Scale represents reads per million (RPM). Snapshots of peaks detected in both biological replicates are shown. The black box outlines *U2AF2* exons 2 and 3 and the adjacent regions.

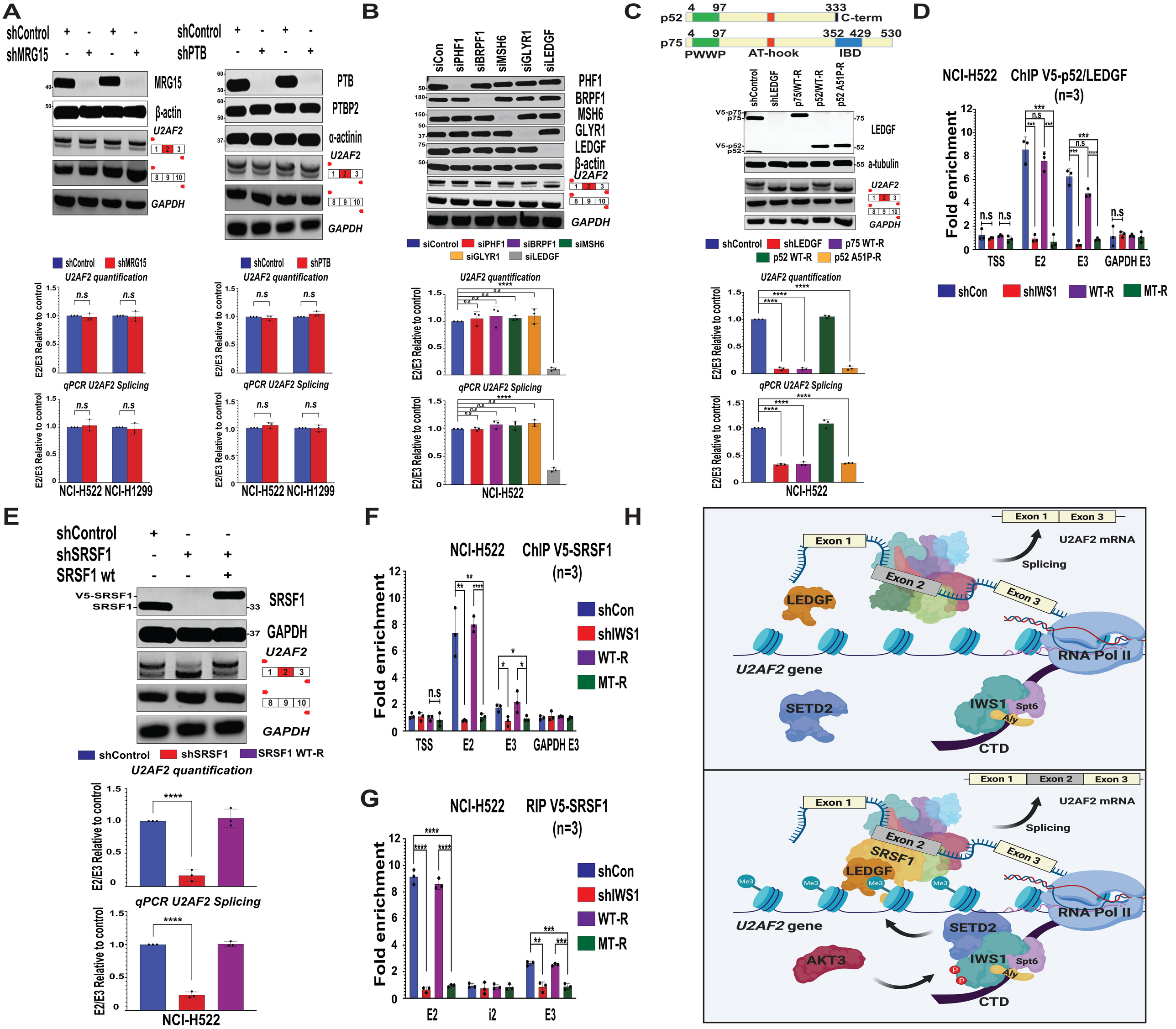

A
shControl
shMRG15
MRG15
β-actin
U2AF2
GAPDH
shPTB
PTB
PTBP2
α-actinin
shControl
shMRG15
shPTB
U2AF2 quantification
qPCR U2AF2 Splicing
E2/E3 Relative to control
n.s
NCI-H522 NCI-H1299
B
siCon
siPHF1
siBRPF1
siMSH6
siGLYR1
siLEDGF
PHF1
BRPF1
MSH6
GLYR1
LEDGF
β-actin
U2AF2
GAPDH
siControl
siPHF1
siBRPF1
siMSH6
siGLYR1
siLEDGF
U2AF2 quantification
qPCR U2AF2 Splicing
NCI-H522
C
p52
p75
C-term
PWWP
AT-hook
IBD
shControl
shLEDGF
p75/WT-R
p52/WT-R
p52 A51P-R
LEDGF
a-tubulin
U2AF2
GAPDH
shControl
shLEDGF
p75 WT-R
p52 WT-R
p52 A51P-R
U2AF2 quantification
qPCR U2AF2 Splicing
NCI-H522
D
NCI-H522 ChIP V5-p52/LEDGF
(n=3)
Fold enrichment
TSS
E2
E3
GAPDH E3
shCon
shIWS1
WT-R
MT-R
E
shControl
shSRSF1
SRSF1 wt
SRSF1
GAPDH
U2AF2
shControl
shSRSF1
SRSF1 WT-R
U2AF2 quantification
qPCR U2AF2 Splicing
NCI-H522
F
NCI-H522 ChIP V5-SRSF1
(n=3)
Fold enrichment
TSS
E2
E3
GAPDH E3
shCon
shIWS1
WT-R
MT-R
G
NCI-H522 RIP V5-SRSF1
(n=3)
Fold enrichment
E2
i2
E3
shCon
shIWS1
WT-R
MT-R
H
Exon 1
Exon 2
Exon 3
U2AF2 mRNA
Splicing
LEDGF
SRSF1
U2AF2 gene
RNA Pol II
SETD2
IWS1
Spt6
Aly
CTD
AKT3
Me3
P

Figure 4. The regulation of the alternative splicing of the *U2AF2* exon 2 by IWS1 phosphorylation, depends on the p52 isoform of the H3K36me3 reader LEDGF and its splicing partner SRSF1.

A. *The H3K36me3 reader MRG15 and its binding partner PTB, play no role in the regulation of U2AF2 alternative RNA splicing.* (Upper panel) Western blots of lysates of NCI-H522 and NCI-H1299 cells transduced with shControl and shMRG15 (left) or shControl and shPTB (right) were probed with the indicated antibodies. RT-PCR, using RNA isolated from these cells and *U2AF2* exon 1 and exon 3 or exon 8 and 10 oligonucleotide primers shows that the knockdown of neither MRG15 nor PTB affects the alternative splicing of *U2AF2*. (Middle and lower panel) The RT-PCR results were confirmed by quantitative RT-PCR. Bar graphs show the E2/E3 ratio in shMRG15 and shPTB NCI-H522 and NCI-H1299 cells, relative to the shControl ± SD.

B. *The alternative splicing of the U2AF2 exon 2 is regulated by the H3K36me3 reader LEDGF.* (Upper Panel) The listed H3K36me3 readers were efficiently knocked down by siRNA transfection of NCI-H522 cells, as determined by probing western blots of transfected cell lysates with the indicated antibodies. RT-PCR, using RNA isolated from these cells and *U2AF2* exon 1 and exon 3 or exon 8 and 10 oligonucleotide primers shows that only the knockdown of LEDGF affects the alternative splicing of *U2AF2*. (Middle and lower panel) The RT-PCR results were confirmed by quantitative RT-PCR. Bar graphs show the E2/E3 ratio in NCI-H522 cells, after the knockdown of the indicated readers, relative to the siControl ± SD.

C. *The alternative RNA splicing of the U2AF2 exon 2 is regulated by the p52, but not the p75 isoform of LEDGF.*(Upper panel) Schematic representation of the p52 and p75 isoforms of LEDGF showing the position of the PWWP domain (green), the AT hook-like domain (red), the unique to p52 8 aa long C-terminal domain (black) and the unique to p75

integrase binding domain (IBD) (blue) (Ferris et al., 2010[55]). (Middle panel) Western blots of lysates of shControl, shLEDGF, shLEDGF/LEDGF-p75-R,shLEDGF/LEDGF-p52-WT-R,shLEDGF/LEDGF-p52A51P-R NCI-H522 cells, were probed with the indicated antibodies. RT-PCR, using RNA isolated from these cells and *U2AF2* exons 1 and 3 or exons 8 and 10 oligonucleotide primers shows that only the wild type p52 isoform of LEDGF rescues the alternative splicing of *U2AF2* in shLEDGF-transduced cells. The p75 isoform and the p52 mutant A51P, which does not bind H3K36me3, failed to rescue. (Lower panels) The RT-PCR results were confirmed by quantitative RT-PCR. Bar graphs show the E2/E3 ratio in shLEDGF-transduced NCI-H522 cells, before and after the rescue with the indicated constructs, relative to the shControl ± SD.

D. *The recruitment of p52/LEDGF to the U2AF2 gene depends on IWS1 phosphorylation.* ChIP assays showing the binding of p52/LEDGF to the *U2AF2* and *GAPDH* genes in shControl, shIWS1 shIWS1/WT-R and shIWS1/MT-R NCI-H522 cells, transduced with a lentiviral V5-p52/LEDGF construct. The bars show the mean fold enrichment in p52/LEDGF binding (anti-V5 IP, vs IgG control IP) to the indicated regions of the *U2AF2*gene, in shControl, shIWS1 shIWS1/WT-R and shIWS1/MT-R cells ±SD. Data were normalized relative to the input (2%). The expression of V5-p52/LEDGF is shown in figure S5G.

E. *The alternative RNA splicing of the U2AF2 exon 2 depends on SRSF1.* (Upper panel) Western blots of lysates of shControl, shSRSF1 and shSRSF1/SRSF1 WT-R NCI-H522 cells, were probed with the indicated antibodies. RT-PCR, using RNA isolated from these cells and *U2AF2* exons 1 and 3 or exons 8 and 10 oligonucleotide primers, shows that the knockdown of SRSF1 promotes a shift in the RNA splicing pattern of *U2AF2*, which favors the exclusion of exon 2 from the mature transcripts, and that wild type SRSF1 rescues the shift. (Lower panels) The RT-PCR results were confirmed by quantitative RT-PCR. Bar

graphs show the E2/E3 ratio in shSRSF1-transduced NCI-H522 cells, before and after rescue with wild type SRSF1, relative to the shControl ±SD.

F. *The recruitment of SRSF1 to the U2AF2 gene depends on IWS1 phosphorylation.* ChIP assays showing the binding of SRSF1 to the *U2AF2* and *GAPDH* genes in shControl, shIWS1 shIWS1/WT-R and shIWS1/MT-R NCI-H522 cells, transduced with a lentiviral V5-SRSF1 construct. The bars show the mean fold enrichment in SRSF1 binding (anti-V5 IP, vs IgG control IP) to the indicated regions of the *U2AF2* gene, in shControl, shIWS1 shIWS1/WT-R and shIWS1/MT-R cells ±SD. Data were normalized relative to the input (2%). The V5-SRSF1 expression is shown in Figure S6B.

G. *The recruitment of SRSF1 to the U2AF2 RNA depends on IWS1 phosphorylation.* RIP assays showing the binding of SRSF1 to the *U2AF2* RNA in the NCI-H522 cells in Fig. 4F. The bars show the mean fold enrichment in SRSF1 binding in the indicated regions of the *U2AF2* RNA (anti-V5 IP, vs IgG control IP) ±SD. Data were normalized relative to the input (2%). All assays in this figure were done in triplicate, on three biological replicates. n.s : non-significant *$p<0.05$, **$p<0.01$, ***$p<0.001$, ****$p<0.0001$. (one-side unpaired t-test)

H. *Model of the regulation of U2AF2 alternative RNA splicing by AKT3-mediated IWS1 phosphorylation at Ser720/Thr721.* The phosphorylation of IWS1 recruits SETD2 to the Spt6/IWS1/Aly complex in the CTD of RNA Pol II. SetD2 in the complex, trimethylates histone H3 at K36 during transcriptional elongation. p52/LEDGF and its RNA-binding partner SRSF1, bind histone H3K36me3. Binding of SRSF1 to exon 2 promotes inclusion of this exon in the mature *U2AF2* transcript.

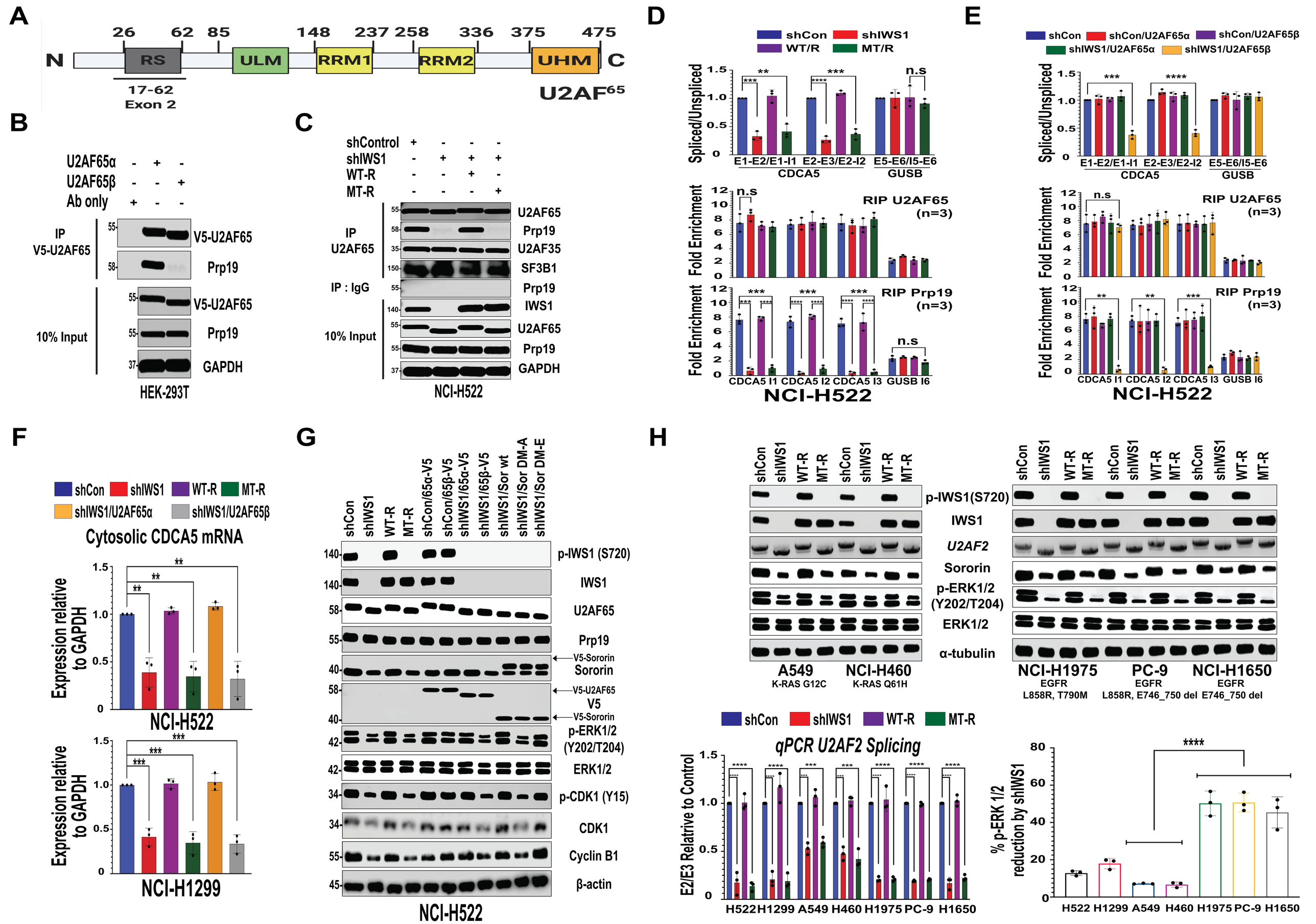

A
N
C
26
62
85
148
237
258
336
375
475
RS
ULM
RRM1
RRM2
UHM
17-62
Exon 2
U2AF65
B
U2AF65α
U2AF65β
Ab only
IP
V5-U2AF65
V5-U2AF65
Prp19
10% Input
V5-U2AF65
Prp19
GAPDH
HEK-293T
C
shControl
shIWS1
WT-R
MT-R
IP
U2AF65
U2AF65
Prp19
U2AF35
SF3B1
IP : IgG
Prp19
IWS1
10% Input
U2AF65
Prp19
GAPDH
NCI-H522
D
shCon
shIWS1
WT/R
MT/R
Spliced/Unspliced
E1-E2/E1-I1
E2-E3/E2-I2
E5-E6/I5-E6
CDCA5
GUSB
n.s
RIP U2AF65
(n=3)
Fold Enrichment
RIP Prp19
(n=3)
CDCA5 I1
CDCA5 I2
CDCA5 I3
GUSB I6
NCI-H522
E
shCon
shCon/U2AF65α
shCon/U2AF65β
shIWS1/U2AF65α
shIWS1/U2AF65β
Spliced/Unspliced
E1-E2/E1-I1
E2-E3/E2-I2
E5-E6/I5-E6
CDCA5
GUSB
n.s
RIP U2AF65
(n=3)
Fold Enrichment
RIP Prp19
(n=3)
CDCA5 I1
CDCA5 I2
CDCA5 I3
GUSB I6
NCI-H522
F
shCon
shIWS1
WT-R
MT-R
shIWS1/U2AF65α
shIWS1/U2AF65β
Cytosolic CDCA5 mRNA
Expression relative to GAPDH
NCI-H522
NCI-H1299
G
shCon
shIWS1
WT-R
MT-R
shCon/65α-V5
shCon/65β-V5
shIWS1/65α-V5
shIWS1/65β-V5
shIWS1/Sor wt
shIWS1/Sor DM-A
shIWS1/Sor DM-E
p-IWS1 (S720)
IWS1
U2AF65
Prp19
V5-Sororin
Sororin
V5-U2AF65
V5
V5-Sororin
p-ERK1/2
(Y202/T204)
ERK1/2
p-CDK1 (Y15)
CDK1
Cyclin B1
β-actin
NCI-H522
H
shCon
shIWS1
WT-R
MT-R
p-IWS1(S720)
IWS1
U2AF2
Sororin
p-ERK1/2
(Y202/T204)
ERK1/2
α-tubulin
A549
K-RAS G12C
NCI-H460
K-RAS Q61H
NCI-H1975
EGFR
L858R, T790M
PC-9
EGFR
L858R, E746_750 del
NCI-H1650
EGFR
E746_750 del
shCon
shIWS1
WT-R
MT-R
qPCR U2AF2 Splicing
E2/E3 Relatrive to Control
H522
H1299
A549
H460
H1975
PC-9
H1650
% p-ERK 1/2
reduction by shIWS1

Figure 5. IWS1 phosphorylation controls the CDCA5/ERK phosphorylation feedback loop, through U2AF2 alternative RNA splicing.

A. Domain organization of the U2AF65 protein, showing the Serine/Arginine Rich (SR) domain, the U2AF-ligand motif (ULM), two RNA recognition motifs (RRMs) and the U2AF Homology Motif (UHM). Exon 2 encodes the N-terminal RS domain. U2AF65 and U2AF35 interact through the U2AF65 UHM domain. The numbers on top identify the amino acids at the boundaries of the indicated motifs.

B. *The novel U2AF65β spliced variant, lacking exon 2, does not interact with Prp19.* HEK-293T cells were transduced with lentiviral constructs V5-U2AF65α, or V5-U2AF65β. Anti V5-U2AF65 immunoprecipitates from these cells were probed with anti-Prp19 or anti-V5-tag antibodies. (Lower panel). Western blots of the input lysates, probed with the indicated antibodies are also shown.

C. *The U2AF65-Prp19 interaction depends on IWS1-phosphorylation.* (Upper panel) Anti-U2AF65 immunoprecipitated from lysates of shControl, shIWS1 shIWS1/WT-R and shIWS1/MT-R NCI-H522 cells were probed with anti-U2AF65, anti-Prp19, and anti-U2AF35 antibodies. Anti-IgG mouse isotype control immunoprecipitates were probed with the anti-Prp19 antibody. Western blots of the input lysates, probed with the indicated antibodies are also shown.

D. *IWS1-phosphorylation controls CDCA5 RNA splicing.* (Upper panel). The ratio of spliced to unspliced *CDCA5* RNA transcripts was examined by quantitative RT-PCR, using total RNA isolated from shControl, shIWS1 shIWS1/WT-R and shIWS1/MT-R NCI-H522 cells and the listed oligonucleotide primers (E1-E2/E1-I1 and E2-E3/E2-I2). The same was done with the transcripts of a control gene (*GUSB*) (Oligonucleotide primers E5-E6/I5-E6). (Middle and lower panels) RIP assays in the same cells show that whereas the binding of Prp19 to the *CDCA5* RNA depends on the phosphorylation of IWS1, the binding of

U2AF65 does not. The bars show the mean fold enrichment in U2AF65 and Prp19 binding in the indicated regions of the *CDCA5* pre-mRNA (anti-U2AF65 or anti-Prp19 IP, vs IgG control IP) ± SD. Data were normalized relative to the input (2%). The primer location within *CDCA5* and *GUSB* can be found in Figure S7B.

E. *U2AF2 alternative RNA splicing, downstream of IWS1 phosphorylation controls CDCA5 splicing.* (Upper panel). The ratio of spliced to unspliced CDCA5 RNA transcripts was examined in shControl NCI-H522 cells that were either not rescued, or rescued with U2AF65α, or U2AF65β, and in shIWS1 NCI-H522 cells rescued with the same U2AF65α, or U2AF65β-encoding constructs. Quantitative RT-PCR was carried out and presented, as in Fig 5D. The GUSB gene was used again as the control. (Middle and lower panels) RIP assays in the same cells show that whereas U2AF65α rescues the binding of Prp19 to the CDCA5 RNA in shIWS1 cells, U2AF65β does not. Moreover, U2AF65α and U2AF65β, bound equally to the *CDCA5* RNA. The presentation of the data was the same as in Fig 5D.

F. *The abundance of mature CDCA5 mRNA in the cytoplasm is under the control of the IWS1-phosphorylation-dependent expression of U2AF65α.* The abundance of the cytosolic *CDCA5* mRNA was measured by quantitative RT-PCR in the cytosolic fraction of shIWS1, shIWS1/WT-R, shIWS1/MT-R, shIWS1/U2AF65α-R and shIWS1/U2AF5β-R NCI-H522 (upper panel) and NCI-H1299 (lower panel) cells. Bars show the mean cytosolic *CDCA5* mRNA expression normalized to *GAPDH* ± SD in NCI-H522 and NCI-H1299 shIWS1, shIWS1/WT-R, shIWS1/MT-R, shIWS1/U2AF65α-R and shIWS1/U2AF5β-R cells compared to shControl. The accuracy of the fractionation was validated with the experiment in Figure S7E.

G. *CDCA5 and p-ERK, form a positive feedback loop, which is activated by IWS1 phosphorylation and U2AF2 alternative RNA splicing and promotes the expression of CDK1 and Cyclin B1.* Lysates of shControl, shIWS1, shIWS1/WT-R and shIWS1/MT-R

NCI-H522 cells, along with shControl/U2AF65α, shControl/U2AF65β, shIWS1/U2AF65α-R, shIWS1/U2AF5β-R and shIWS1/CDCA5-WT-R, shIWS1/CDCA5-DM-A-R and shIWS1/CDCA5-DM-D-R NCI-H522 cells, were probed with the indicated antibodies. (DM-A is the S79/S209 CDCA5 mutant and DM-E is the S79/S209EE CDCA5 mutant).

H. *IWS1 phosphorylation controls the phosphorylation of ERK in cells harboring EGFR and KRAS mutations.* Lysates of shControl, shIWS1, shIWS1/WT-R and shIWS1/MT-R *KRAS* (upper left panels) and *EGFR* mutant cell lines were probed with the indicated antibodies. ERK phosphorylation was reduced in all shIWS1 and shIWS1/MT-R cells, including the ones with *KRAS* and *EGFR* mutations. RT-PCR, using RNA isolated from these cells and *U2AF2* exon 1 and exon 3 oligonucleotide primers, confirmed the exclusion of exon 2 from the mature *U2AF2* mRNA, in all the shIWS1-transduced cells. (Lower left panel) The RT-PCR results were confirmed by quantitative RT-PCR. Bar graphs show the E2/E3 ratio in all the shIWS1, shIWS1/WT-R, and shIWS1/MT-R cells, relative to the shControl ± SD. (Lower right panel) Quantification of the reduction of ERK-phosphorylation (Y202/T204), induced by shIWS1 in the indicated cell lines. Bars show the percent reduction, normalized to tubulin. The comparison between the non-*EGFR* mutant (NCI-H522, NCI-H1299, A549 and NCI-H460) and *EGFR* mutant cells (NCI-H1975, PC-9, NCI-H1650) was performed with one-way ANOVA statistical test.

All assays in this figure were done in triplicate, on three biological replicates. n.s : non-significant *$p<0.05$, **$p<0.01$, ***$p<0.001$, ****$p<0.0001$ (one-side unpaired t-test).

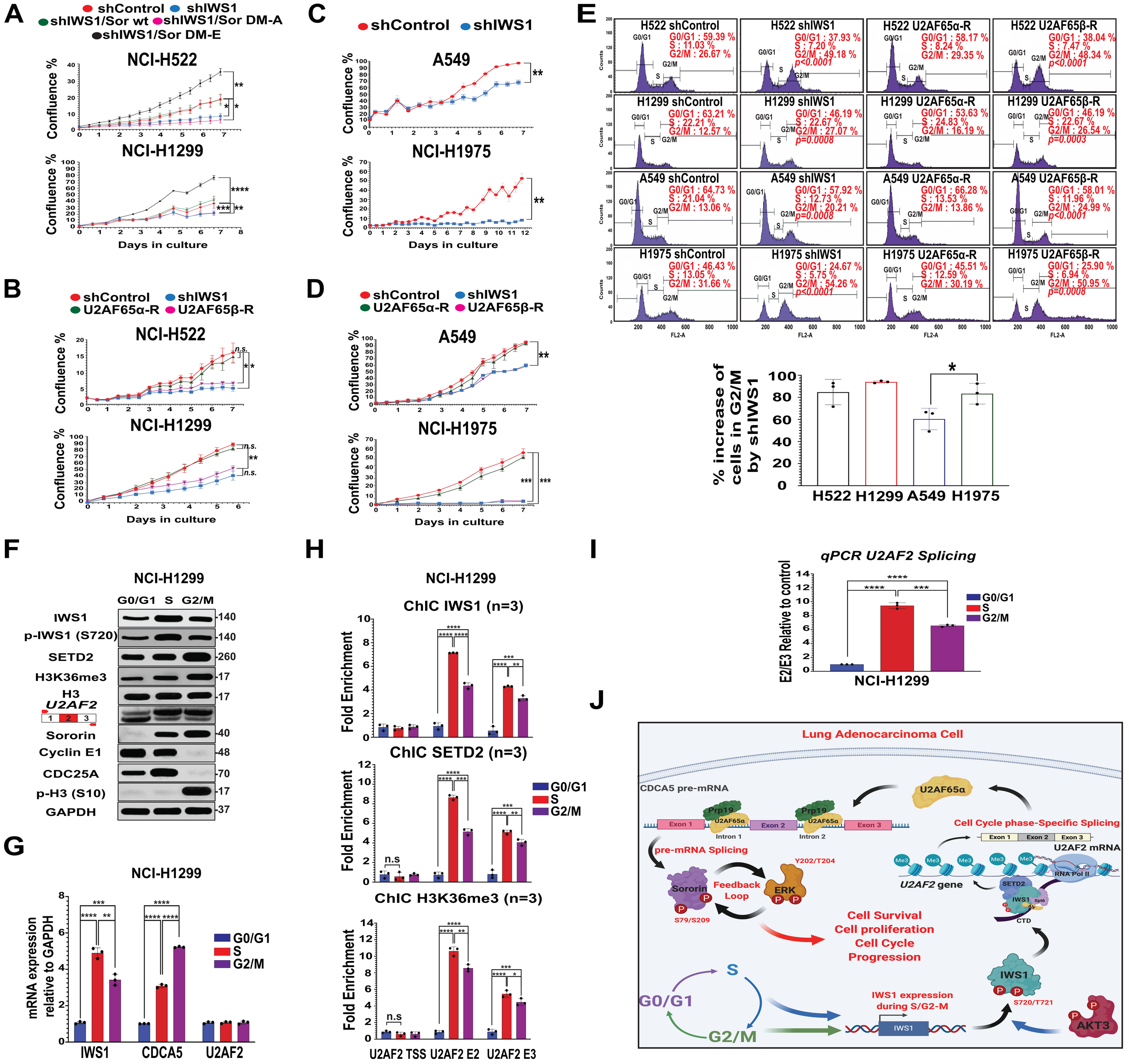

A
shControl
shIWS1
shIWS1/Sor wt
shIWS1/Sor DM-A
shIWS1/Sor DM-E
NCI-H522
NCI-H1299
Confluence %
Days in culture
B
shControl
shIWS1
U2AF65α-R
U2AF65β-R
NCI-H522
NCI-H1299
Confluence %
Days in culture
C
shControl
shIWS1
A549
NCI-H1975
Confluence %
Days in culture
D
shControl
shIWS1
U2AF65α-R
U2AF65β-R
A549
NCI-H1975
Confluence %
Days in culture
E
H522 shControl
G0/G1 : 59.39 %
S : 11.03 %
G2/M : 26.67 %
H522 shIWS1
G0/G1 : 37.93 %
S : 7.20 %
G2/M : 49.18 %
p<0.0001
H522 U2AF65α-R
G0/G1 : 58.17 %
S : 8.24 %
G2/M : 29.35 %
H522 U2AF65β-R
G0/G1 : 38.04 %
S : 7.47 %
G2/M : 48.34 %
p<0.0001
H1299 shControl
G0/G1 : 63.21 %
S : 22.21 %
G2/M : 12.57 %
H1299 shIWS1
G0/G1 : 46.19 %
S : 22.67 %
G2/M : 27.07 %
p=0.0008
H1299 U2AF65α-R
G0/G1 : 53.63 %
S : 24.83 %
G2/M : 16.19 %
H1299 U2AF65β-R
G0/G1 : 46.19 %
S : 22.67 %
G2/M : 26.54 %
p=0.0003
A549 shControl
G0/G1 : 64.73 %
S : 21.04 %
G2/M : 13.06 %
A549 shIWS1
G0/G1 : 57.92 %
S : 12.73 %
G2/M : 20.21 %
p=0.0008
A549 U2AF65α-R
G0/G1 : 66.28 %
S : 13.53 %
G2/M : 13.86 %
A549 U2AF65β-R
G0/G1 : 58.01 %
S : 11.96 %
G2/M : 24.99 %
p<0.0001
H1975 shControl
G0/G1 : 46.43 %
S : 13.05 %
G2/M : 31.66 %
H1975 shIWS1
G0/G1 : 24.67 %
S : 5.75 %
G2/M : 54.26 %
p<0.0001
H1975 U2AF65α-R
G0/G1 : 45.51 %
S : 12.59 %
G2/M : 30.19 %
H1975 U2AF65β-R
G0/G1 : 25.90 %
S : 6.94 %
G2/M : 50.95 %
p=0.0008
Counts
FL2-A
% increase of cells in G2/M by shIWS1
H522 H1299 A549 H1975
F
NCI-H1299
G0/G1 S G2/M
IWS1
p-IWS1 (S720)
SETD2
H3K36me3
H3
U2AF2
Sororin
Cyclin E1
CDC25A
p-H3 (S10)
GAPDH
G
NCI-H1299
mRNA expression relative to GAPDH
G0/G1
S
G2/M
IWS1
CDCA5
U2AF2
H
NCI-H1299
ChIC IWS1 (n=3)
ChIC SETD2 (n=3)
ChIC H3K36me3 (n=3)
Fold Enrichment
n.s
U2AF2 TSS
U2AF2 E2
U2AF2 E3
I
qPCR U2AF2 Splicing
E2/E3 Relative to control
NCI-H1299
J
Lung Adenocarcinoma Cell
CDCA5 pre-mRNA
U2AF65α
Prp19
Exon 1
Exon 2
Exon 3
Intron 1
Intron 2
pre-mRNA Splicing
Cell Cycle phase-Specific Splicing
U2AF2 mRNA
U2AF2 gene
RNA Pol II
SETD2
IWS1
CTD
Sororin
S79/S209
Feedback Loop
ERK
Y202/T204
Cell Survival
Cell proliferation
Cell Cycle Progression
S
G0/G1
G2/M
IWS1 expression during S/G2-M
IWS1
S720/T721
AKT3

Figure 6. IWS1 phosphorylation promotes cell proliferation by controlling the Sororin/ERK phosphorylation feedback loop, through *U2AF2* RNA splicing.

A. Growth curves of the indicated NCI-H522 (upper) and NCI-H1299 (lower) cells in media supplemented with 10% FBS. Cell proliferation was measured every 6 hours using an Incucyte live cell imager. Images of live cells in three independent cultures for each cell type were taken and analyzed using the native Incucyte confluence masking software. Results were expressed as confluence percentages ± SD. For simplicity, 12 and not 6 hour time points are shown. P values were calculated for the endpoint measurements, using the one-side unpaired t-test. *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001

B. *U2AF2 alternative RNA splicing affects cell proliferation.* Growth curves of the indicated NCI-H522 (upper) and NCI-H1299 (lower) cells in media supplemented with 10% FBS. The acquisition, representation and statistical analysis were performed similar to Fig. 6A.

C. *IWS1 phosphorylation promotes cell proliferation more robustly in the EGFR mutant lung adenocarcinoma cell line NCI-H1975, than in the KRAS mutant cell line A549.* Growth curves of shControl and shIWS1 A549 (*KRAS* mutant) and NCI-H1975 (*EGFR* mutant) cells in media supplemented with 10% FBS. Cell proliferation was measured as in A.

D. Growth curves of the indicated A549 (upper) and NCI-H1975 (lower) cells in media supplemented with 10% FBS. The acquisition, representation and statistical analysis were performed similar to Fig. 6A.

E. *IWS1 phosphorylation regulates progression through the G2/M phase of the cell cycle.* Cell cycle profiles of the indicated propidium iodide (PI)-stained cell lines, transduced with shControl, shIWS1 (Left upper panels) or rescued by U2AF65α oρ U2AF65β lentiviral constructs (Right upper panels). The cells were plated in equal numbers and they were harvested from semi-confluent cultures 48 hours later. Figure shows one representative, out of three biological replicates for each cell line. Numbers in red show the mean

percentage of cells in different phases of the cell cycle, measured based on the three biological replicates. The percent of cells in G2/M in shIWS1 cells was compared with the percent of cells in G2/M in shControl cells and the p value of the difference was calculated with the one-side unpaired t-test. (Lower panel) Bars show the shIWS1-induced percent change of the percentage of cells in G2/M in the indicated cell lines. The error bars show the SD of the percent change in 3 biological replicates. The percent change is smallest in the *KRAS* mutant cell line A549. Supplemental Table S4 lists the data for all the replicates and all the conditions. The statistical analysis between the NCI-H1975 and A549 was performed using one-sided unpaired t-test.

F. *IWS1 phosphorylation and U2AF2 alternative RNA splicing fluctuate during progression through the cell cycle.* Protein extracts derived from Exponentially growing NCI-H1299 cell cultures were separated into fractions enriched for cells in G0/G1, S and G2/M. Lysates of these cell fractions were probed with the indicated antibodies. RT-PCR, using RNA from the same cell fractions, and *U2AF2* exons 1 and 3 oligonucleotide primers, shows that *U2AF2* exon 2 inclusion in the mature *U2AF2* mRNA increases in cells in S and G2/M.

G. The abundance of the *IWS1*, *CDCA5* and *U2AF2* mRNAs in the G0/G1, S and G2/M phases of the cell cycle in NCI-H1299 cells was measured by quantitative RT-PCR and is presented relative to the abundance of GAPDH ± SD.

H. *IWS1 phosphorylation regulates the U2AF2 alternative RNA splicing in a cell-cycle specific manner*. ChIC assays showing the binding of IWS1 and SETD2 (upper and middle panels respectively) and the abundance of H3K36me3 (lower panel) in the *U2AF2* gene, in NCI-H1299 cells in the G0/G1, S and G2/M phases of the cell cycle. The bars show the mean fold enrichment in IWS1 and SETD2 binding, and in H3K36me3 abundance, in the indicated regions of the *U2AF2* gene ±SD.

I. Quantitative RT-PCR showing the *E2/E3 U2AF2* ratio in NCI-H1299 cells. Bars show the mean E2/E3 *U2AF2* ratio in NCI-H1299 cellular fractions enriched for cells in S and G2/M phase relative to G0/G1 enriched fractions. In 6G to 6I, the error bars show the SD of at least 3 biological replicates. *$p<0.05$, **$p<0.01$, ***$p<0.001$, ****$p<0.0001$. (one-side unpaired t-test)

J. *Model of the regulation of the Sororin/ERK loop and cell cycle progression via the U2AF2 alternative RNA splicing, through the cell cycle.* IWS1 expression occurs specifically during the S and G2/M phase of the cell cycle. Following this cell cycle specific expression, IWS1 is phosphorylated by Akt3 at S720/T721 and orchestrates the assembly of epigenetic complexes on *U2AF2* gene, which translate the SETD2-mediated H3K36me3 signal into shifts of the alternative RNA splicing pattern of U2AF2, in a cell cycle specific manner. Subsequently, the U2AF65α isoform along with Prp19, facilitate the proper splicing of *CDCA5* pre-mRNA, leading to accumulation of Sororin during S and G2/M phase. Finally, Sororin forms a positive feedback loop with ERK phosphorylation. Activation of this loop plays an important role in the maintenance of ERK phosphorylation, and in the progression through the G2/M phase of the cell cycle in lung adenocarcinoma.

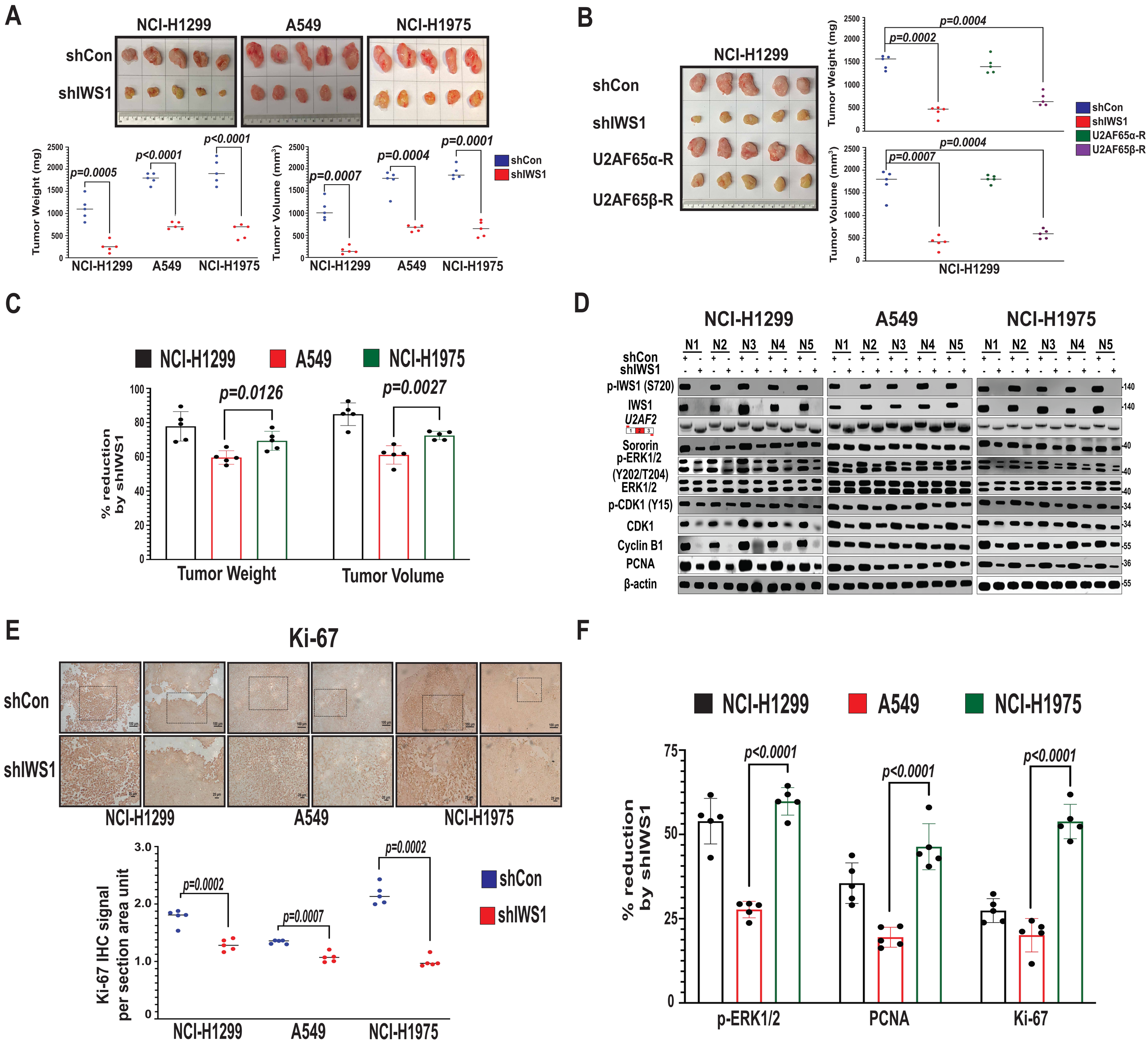

A
NCI-H1299
A549
NCI-H1975
shCon
shIWS1
Tumor Weight (mg)
Tumor Volume (mm³)
p=0.0005
p<0.0001
p<0.0001
p=0.0007
p=0.0004
p=0.0001
shCon
shIWS1
B
NCI-H1299
shCon
shIWS1
U2AF65α-R
U2AF65β-R
Tumor Weight (mg)
Tumor Volume (mm³)
p=0.0004
p=0.0002
p=0.0004
p=0.0007
shCon
shIWS1
U2AF65α-R
U2AF65β-R
C
NCI-H1299
A549
NCI-H1975
% reduction by shIWS1
p=0.0126
p=0.0027
Tumor Weight
Tumor Volume
D
NCI-H1299
A549
NCI-H1975
N1 N2 N3 N4 N5
shCon
shIWS1
p-IWS1 (S720)
IWS1
U2AF2
Sororin
p-ERK1/2 (Y202/T204)
ERK1/2
p-CDK1 (Y15)
CDK1
Cyclin B1
PCNA
β-actin
E
Ki-67
shCon
shIWS1
NCI-H1299
A549
NCI-H1975
Ki-67 IHC signal per section area unit
p=0.0002
p=0.0007
p=0.0002
shCon
shIWS1
F
NCI-H1299
A549
NCI-H1975
% reduction by shIWS1
p<0.0001
p<0.0001
p<0.0001
p-ERK1/2
PCNA
Ki-67

Figure 7. IWS1 phosphorylation controls tumor growth *in vivo*, by regulating the U2AF2/Sororin/ERK axis.

A. *IWS1 phosphorylation at Ser720/Thr721 promotes tumor growth in xenograft lung adenocarcinoma model in NSG mice*. NSG mice were injected subcutaneously with shControl or shIWS1 NCI H1299, A549 or NCI-H1975 cells. (N=5 mice/group). (Left panels) Images of the induced tumors, harvested at 4 weeks (NCI-H1299, NCI-H1975 cells) and 6 weeks (A549 cells) from the time of inoculation (Right panels) Scatter plots showing the tumor weight (up) and volume (down) of the harvested tumors. The horizontal lines indicate mean tumor weight or volume. Statistical analyses were done using the paired t-test.

B. *Inclusion of U2AF2 E2 promotes tumor growth in xenograft lung adenocarcinoma model in NSG mice.* NSG mice were injected subcutaneously with shControl, shIWS1, shIWS1/U2AF65α or shIWS1/U2AF65β NCI-H1299 cells. (N=5 mice/group). (Left panels) Images of the induced tumors, harvested at 4 weeks from the time of inoculation (Right panels) Scatter plots showing the tumor weight (up) and volume (down) of the harvested tumors. The horizontal lines indicate mean tumor weight or volume. Statistical analyses were done using the paired t-test.

C. From the weight and volume of the tumors induced by shIWS1-transduced cells and shControl-transduced cells in Figure 7A, we calculated the shIWS1-induced percent reduction of both tumor weight and volume. Bars show the mean percent reduction. The error bars show the SD of the percent reduction in each group. Statistical analyses were done using the one-sided unpaired t-test.

D. *IWS1 phosphorylation controls the Sororin/ERK phosphorylation axis in vivo.* Cell lysates derived from NCI-H1299, A549 and NCI-H1975 (shControl and shIWS1) mouse xenografts, were probed with the indicated antibodies. RT-PCR, using RNA isolated from

these xenografts and *U2AF2* exon 1 and exon 3 oligonucleotide primers shows that the knockdown of IWS1 in the xenografts results in *U2AF2* exon 2 exclusion, as expected.

E. (Upper panels) Ki-67 staining of tumor xenografts of shControl and shIWS1-transduced NCI-H1299 (left), A549 (middle) and NCI-H1975 (right) cells. Formalin-fixed, paraffin-embedded tumor samples were stained with a Ki-67 antibody. Secondary antibody was HRP-labelled. The boxes delineate the area of higher magnification shown in the image below. Scale bar in the right corner of each image. (Lower panel) Scatter plots showing the Ki-67 IHC signal relative to the section area in shControl and shIWS1 NCI-H1299, A549 and NCI-H1975 tumors. The horizontal line shows the mean Ki-67 signal in the indicated groups of xenografts. Statistical analyses were performed, using the paired t-test.

F. *KRAS mutant tumors are more resistant to the loss of IWS1 in vivo.* Percentage of reduction of the described marker induced by shIWS1 in the tumors derived from the NCI-H1299, A549 and NCI-H1975 cell lines. Bars show the mean reduction of the markers, derived from the Western blots data in Figure 7D (p-ERK and PCNA) and IHC data in Figure 7E (Ki-67). The error bars are SD of the values in each mice group. The statistics were performed with one-sided unpaired t-test.

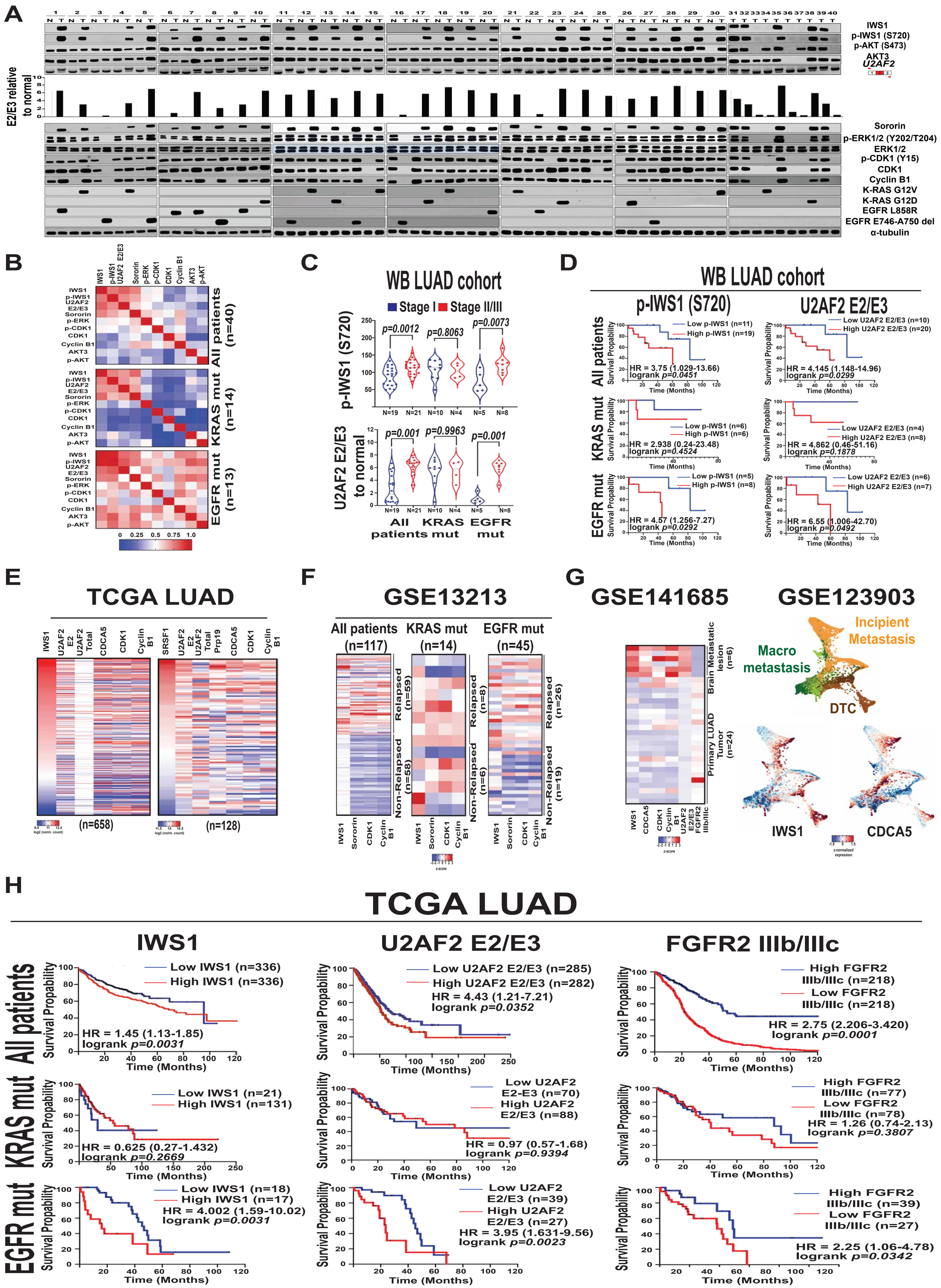

A
IWS1
p-IWS1 (S720)
p-AKT (S473)
AKT3
U2AF2
E2/E3 relative to normal
Sororin
p-ERK1/2 (Y202/T204)
ERK1/2
p-CDK1 (Y15)
CDK1
Cyclin B1
K-RAS G12V
K-RAS G12D
EGFR L858R
EGFR E746-A750 del
α-tubulin
B
All patients (n=40)
KRAS mut (n=14)
EGFR mut (n=13)
C
WB LUAD cohort
Stage I
Stage II/III
p-IWS1 (S720)
U2AF2 E2/E3 to normal
p=0.0012
p=0.8063
p=0.0073
p=0.001
p=0.9963
p=0.001
All patients
KRAS mut
EGFR mut
D
WB LUAD cohort
p-IWS1 (S720)
U2AF2 E2/E3
All patients
KRAS mut
EGFR mut
Low p-IWS1 (n=11)
High p-IWS1 (n=19)
HR = 3.75 (1.029-13.66)
logrank p=0.0451
Low U2AF2 E2/E3 (n=10)
High U2AF2 E2/E3 (n=20)
HR = 4.145 (1.148-14.96)
logrank p=0.0299
Low p-IWS1 (n=6)
High p-IWS1 (n=6)
HR = 2.938 (0.24-23.48)
logrank p=0.4524
Low U2AF2 E2/E3 (n=4)
High U2AF2 E2/E3 (n=8)
HR = 4.862 (0.46-51.16)
logrank p=0.1878
Low p-IWS1 (n=5)
High p-IWS1 (n=8)
HR = 4.57 (1.256-7.27)
logrank p=0.0292
Low U2AF2 E2/E3 (n=6)
High U2AF2 E2/E3 (n=7)
HR = 6.55 (1.006-42.70)
logrank p=0.0492
Survival Propability
Time (Months)
E
TCGA LUAD
(n=658)
(n=128)
F
GSE13213
All patients (n=117)
KRAS mut (n=14)
EGFR mut (n=45)
Relapsed
Non-Relapsed
G
GSE141685
GSE123903
Brain Metastatic lesion (n=6)
Primary LUAD Tumor (n=24)
Incipient Metastasis
Macro metastasis
DTC
IWS1
CDCA5
H
TCGA LUAD
IWS1
U2AF2 E2/E3
FGFR2 IIIb/IIIc
All patients
KRAS mut
EGFR mut
Low IWS1 (n=336)
High IWS1 (n=336)
HR = 1.45 (1.13-1.85)
logrank p=0.0031
Low U2AF2 E2/E3 (n=285)
High U2AF2 E2/E3 (n=282)
HR = 4.43 (1.21-7.21)
logrank p=0.0352
High FGFR2 IIIb/IIIc (n=218)
Low FGFR2 IIIb/IIIc (n=218)
HR = 2.75 (2.206-3.420)
logrank p=0.0001
Low IWS1 (n=21)
High IWS1 (n=131)
HR = 0.625 (0.27-1.432)
logrank p=0.2669
Low U2AF2 E2-E3 (n=70)
High U2AF2 E2/E3 (n=88)
HR = 0.97 (0.57-1.68)
logrank p=0.9394
High FGFR2 IIIb/IIIc (n=77)
Low FGFR2 IIIb/IIIc (n=78)
HR = 1.26 (0.74-2.13)
logrank p=0.3807
Low IWS1 (n=18)
High IWS1 (n=17)
HR = 4.002 (1.59-10.02)
logrank p=0.0031
Low U2AF2 E2/E3 (n=39)
High U2AF2 E2/E3 (n=27)
HR = 3.95 (1.631-9.56)
logrank p=0.0023
High FGFR2 IIIb/IIIc (n=39)
Low FGFR2 IIIb/IIIc (n=27)
HR = 2.25 (1.06-4.78)
logrank p=0.0342
Survival Propability
Time (Months)

Figure 8. The p-IWS1/U2AF2 pathway is active in human lung adenocarcinomas and impacts tumor grade, stage, metastatic potential and treatment relapse in patients with *EGFR* mutant, but not *KRAS* mutant tumors.

A. Lysates from 30 LUAD samples, paired with NAT and 10 unpaired LUAD samples were probed with the indicated antibodies. (Fifth line from the top). RT-PCR, using RNA isolated from these tumors, and *U2AF2* exon 1 and exon 3 oligonucleotide primers, shows that phosphor-IWS1 correlates with *U2AF2* exon 2 inclusion, as expected. (Sixth line) *U2AF2* E2/E3 exon ratio in the tumors in the upper panel, relative to the average of the 30 normal lung samples.

B. *IWS1 and phosphor-IWS1 correlate with U2AF2 exon 2 inclusion and the activity of the Sororin/ERK axis, most robustly in EGFR mutant lung adenocarcinomas.* Heatmaps of the correlation coefficients between the indicated components of the IWS1 phosphorylation pathway, in the entire cohort (left), in *KRAS* mutant (middle) and in *EGFR* mutant tumors (right). The correlation coefficients were calculated using simple linear regression. The values and the statistical confidence of all the comparisons can be found in Supplementary Table S5.

C. Violin plots showing the abundance of IWS1 phosphorylation (left) and the *U2AF2* E2/E3 ratio (right) in stage I and Stage II/III tumors. Data shown for all tumors in A and B, and selectively for *EGFR* or *KRAS* mutant tumors. The horizontal black lines indicate mean values for p-IWS1 levels and *U2AF2* E2/E3 ratios. Statistical analyses were performed using the one-sided unpaired t-test.

D. Kaplan-Meier Curves showing the impact of the abundance of phosphor-IWS1 (left) and *U2AF2* E2/E3 ratio (right) on patient survival in the LUAD cohort in Fig. 8A. Separate curves are shown again for all the patients, or for the patients with *KRAS* or *EGFR* mutant

tumors. The effect of phosphor-IWS1 levels and *U2AF2* exon 2 inclusion on patient survival is again significant in *EGFR* mutant, but not in *KRAS* mutant tumors. Statistical analyses were performed using the log rank test and Cox's proportional hazards model.

E. Heat Maps showing the correlation of IWS1 or SRSF1 with components of the IWS1/U2AF2 pathway in the TCGA LUAD database. The RNA-seq expression values from each individual patient were expressed as log2 normalized counts and they were used to calculate the correlation coefficients.

F. *The IWS1 phosphorylation axis correlated with relapse in lung adenocarcinoma patients harboring EGFR mutations.* Heatmaps showing the expression of IWS1, CDCA5, CDK1 and CCNB1 in relapsed and in non-relapsed tumors. Microarray analyses of lung adenocarcinomas in the entire cohort (left), *KRAS* mutant (middle) or *EGFR* mutant (right) cohort in the GSE13213 dataset. For quantification of gene expression we used the z-scores of the microarray signals.

G. *IWS1 expression and the U2AF2 exon 2 inclusion and FGFR2 exon IIIb exclusion pathways correlate strongly with metastasis in lung adenocarcinomas.* (Left panel) Heat Maps showing the expression of IWS1 and its downstream targets in the IWS1/U2AF2 pathway in resected brain metastatic lesions and primary lung adenocarcinomas. RNA-seq data from the GSE141685 dataset and primary lung adenocarcinomas derived from TCGA LUAD database. The TCGA LUAD samples were picked randomly, independent of their clinical stage or mutational status. For quantification of gene expression we used the z-scores of the normalized RNA-Seq reads. The primary tumors from the TCGA LUAD database are listed in Table S6. (Right Upper panel) Force-directed layout/t-SNE of all metastatic tumor cells isolated from patient-derived xenograft mice (Laughney et al., 2020[79]) colored by source, derived from scRNA-Seq data from the GSE123903. (Right lower panel) Force-directed layout (as in the upper panel) of all xenograft tumor cells colored by *z*-normalized *IWS1* and *CDCA5* expression.

H. *The IWS1 phosphorylation axis defines poor survival in lung adenocarcinoma patients harboring EGFR mutations.* Kaplan-Meier Curves showing the impact of IWS1 expression (left), *U2AF2* E2/E3 ratio (middle) and *FGFR2* IIIb/IIIc ratio (right) on patient survival in patients in the TCGA LUAD database. Separate curves are shown for all the patients, or for the patients with K*RAS* or *EGFR* mutant tumors. The effect of IWS1 levels and *U2AF2* exon 2 inclusion in patient survival is significant in *EGFR* mutant, but not in *KRAS* mutant tumors. Statistical analyses were performed using the log rank test and Cox's proportional hazards model.

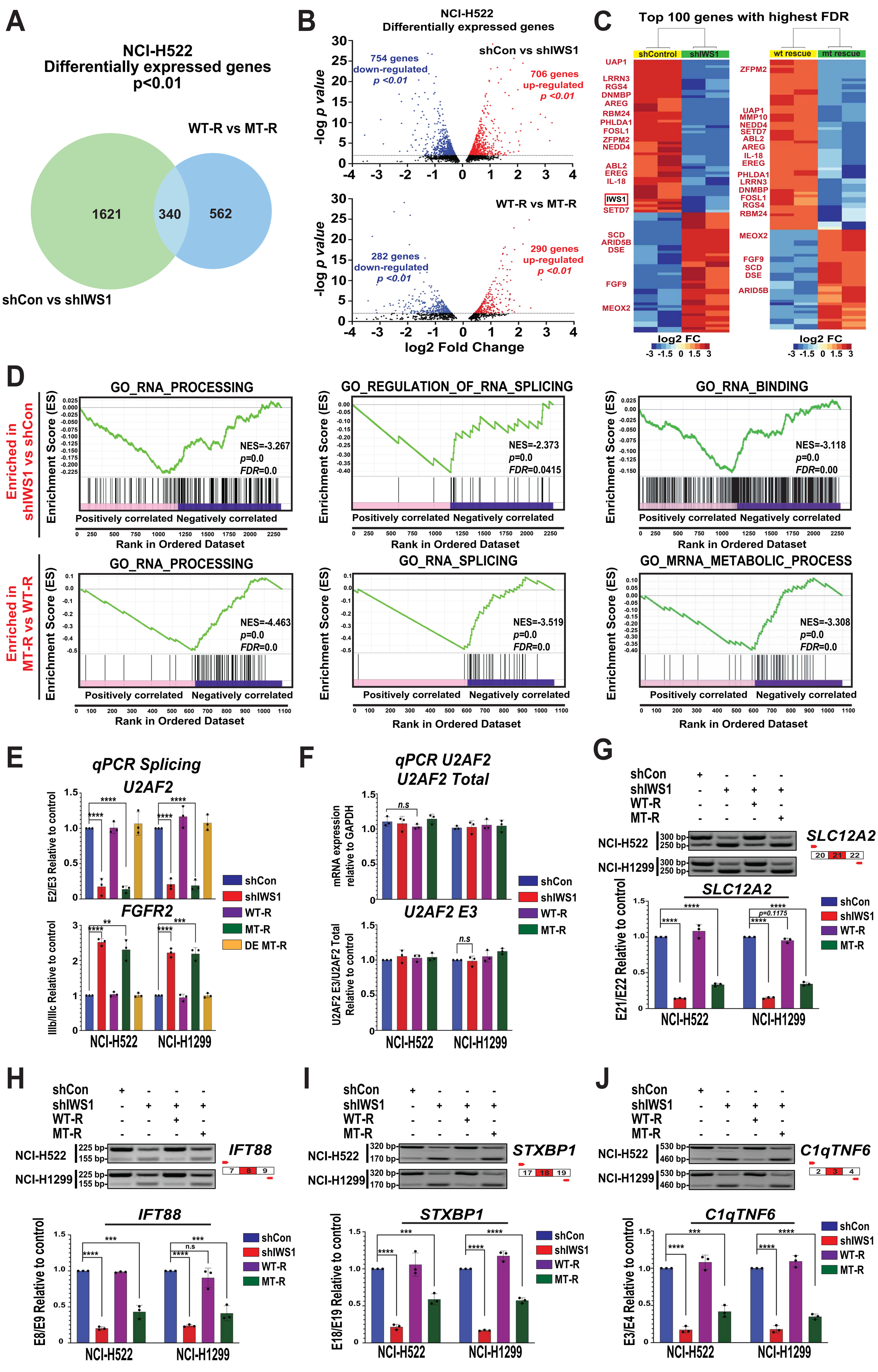

A
NCI-H522
Differentially expressed genes
p<0.01
WT-R vs MT-R
1621
340
562
shCon vs shIWS1
B
NCI-H522
Differentially expressed genes
shCon vs shIWS1
754 genes
down-regulated
p <0.01
706 genes
up-regulated
p <0.01
WT-R vs MT-R
282 genes
down-regulated
p <0.01
290 genes
up-regulated
p <0.01
-log p value
log2 Fold Change
C
Top 100 genes with highest FDR
shControl
shIWS1
wt rescue
mt rescue
UAP1
LRRN3
RGS4
DNMBP
AREG
RBM24
PHLDA1
FOSL1
ZFPM2
NEDD4
ABL2
EREG
IL-18
IWS1
SETD7
SCD
ARID5B
DSE
FGF9
MEOX2
MMP10
SETD7
log2 FC
-3 -1.5 0 1.5 3
D
Enriched in shIWS1 vs shCon
Enriched in MT-R vs WT-R
GO_RNA_PROCESSING
NES=-3.267
p=0.0
FDR=0.0
GO_REGULATION_OF_RNA_SPLICING
NES=-2.373
p=0.0
FDR=0.0415
GO_RNA_BINDING
NES=-3.118
p=0.0
FDR=0.00
GO_RNA_PROCESSING
NES=-4.463
p=0.0
FDR=0.0
GO_RNA_SPLICING
NES=-3.519
p=0.0
FDR=0.0
GO_MRNA_METABOLIC_PROCESS
NES=-3.308
p=0.0
FDR=0.0
Enrichment Score (ES)
Positively correlated
Negatively correlated
Rank in Ordered Dataset
E
qPCR Splicing
U2AF2
FGFR2
E2/E3 Relative to control
IIIb/IIIc Relative to control
NCI-H522
NCI-H1299
shCon
shIWS1
WT-R
MT-R
DE MT-R
F
qPCR U2AF2
U2AF2 Total
mRNA expression relative to GAPDH
U2AF2 E3
U2AF2 E3/U2AF2 Total Relative to control
n.s
G
shCon
shIWS1
WT-R
MT-R
NCI-H522
NCI-H1299
300 bp
250 bp
SLC12A2
20 21 22
E21/E22 Relative to control
p=0.1175
H
IFT88
225 bp
155 bp
7 8 9
E8/E9 Relative to control
I
STXBP1
320 bp
170 bp
17 18 19
E18/E19 Relative to control
J
C1qTNF6
530 bp
460 bp
2 3 4
E3/E4 Relative to control

Figure S1 (relative to Figure 1). IWS1 expression and/or phosphorylation regulate alternative mRNA splicing

A. Venn diagram shows the genes that are differentially expressed in shControl versus shIWS1 or shIWS1/WT-R versus shIWS1/MT-R NCI-H522 cells (q<0.01) and the overlap between the two sets.

B. Volcano plots of all the differentially expressed genes in shControl versus shIWS1(upper panel) and shIWS1/WT-R versus shIWS1/MT-R NCI-H522 cells (lower panel). Genes upregulated or downregulated with a p value <0.01 are shown in red and blue, respectively.

C. Heatmaps of the top 100 differentially-expressed genes (lowest FDR) in shIWS1 relative to shControl and in shIWS1/MT-R relative to shIWS1/WT-R NCI-H522 cells. Genes present in both sets and regulated to the same direction are shown in red. The abundance of IWS1 was significantly lower in shIWS1 cells, relative to the shControl cells, as expected (red box).

D. Gene Set Enrichment Analysis (GSEA) of genes differentially expressed in shControl versus shIWS1 (upper panels) and in shIWS1/WT-R versus shIWS1/MT-R (lower panels) NCI-H522 cells demonstrates significant enrichment in genes involved in RNA processing. (NES–normalized enrichment score, FDR–false discovery ratio)

E. *IWS1 expression and phosphorylation regulate the alternative splicing of the U2AF2 exon 2 and FGFR2 exon IIIb.* Quantitative RT-PCR showing the *E2/E3* and *IIIb/IIIc* ratio in the *U2AF2* (upper panel) and *FGFR2* (lower panel) mRNA transcripts in shControl, shIWS1, shIWS1/WT-R, shIWS1/MT-R and shIWS1/DE MT-R NCI-H522 and NCI-H1299 cells. Bars show the E2/E3 or IIIb/IIIc ratio (mean ± SD) in all the cell lines, normalized relative to the shControl.

F. Quantitative RT-PCR, showing the abundance of the total *U2AF2* mRNA transcripts (upper panel) or the U2AF2 E3/U2AF2 total ratio in shControl, shIWS1, shIWS1/WT-R and shIWS1/MT-R NCI-H522 and NCI-H1299 cells. Bars show the total *U2AF2* mRNA (mean ± SD) relative to *GAPDH* and the U2AF2 E3/U2AF2 Total (mean ± SD) relative to shControl.

G. **H. I** and **J**. IWS1 expression and phosphorylation regulate the alternative RNA splicing of *SLC12A2* (exon 21), *IFT88* (exon 8), *STXBP1* (exon 18) and *C1qTNF6* (exon 3). In all cases, IWS1 promotes exon inclusion. Upper panels show the electrophoresed products of the corresponding RT-PCR reactions. These products were quantified and the ratios of the alternatively spliced to non-alternatively spliced adjacent exons were normalized relative to the ratio in the shControl cells. Bars show the relative exon ratios (mean ± SD) in shControl, shIWS1, shIWS1/WT-R and shIWS1/MT-R NCI-H522 and NCI-H1299 cells. Assays were done in triplicate, on three biological replicates. n.s : non-significant *$p<0.05$, **$p<0.01$, ***$p<0.001$, ****$p<0.0001$ (one-side unpaired t-test).

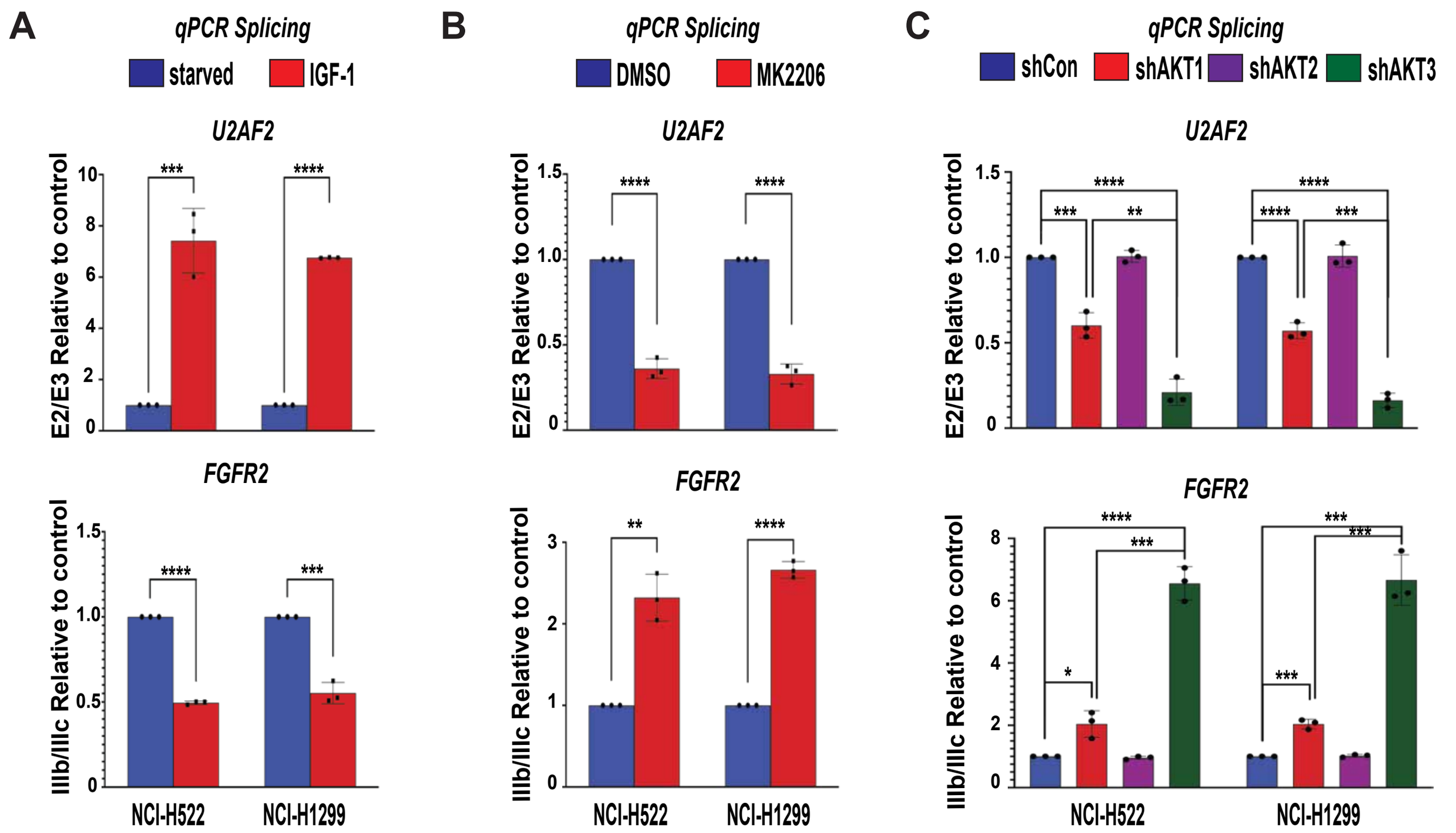

A
qPCR Splicing
starved
IGF-1
U2AF2
E2/E3 Relative to control
FGFR2
IIIb/IIIc Relative to control
NCI-H522
NCI-H1299
B
qPCR Splicing
DMSO
MK2206
U2AF2
E2/E3 Relative to control
FGFR2
IIIb/IIIc Relative to control
NCI-H522
NCI-H1299
C
qPCR Splicing
shCon
shAKT1
shAKT2
shAKT3
U2AF2
E2/E3 Relative to control
FGFR2
IIIb/IIIc Relative to control
NCI-H522
NCI-H1299

Figure S2 (relative to Figure 2).IWS1 phosphorylation-dependent mRNA splicing of U2AF2 is regulated by serum and IGF-1-induced signals transduced by AKT3

A. B. and C. (Upper panels) The *U2AF2* mRNA E2/E3 ratio in the cells in Figure 2 was measured by quantitative RT-PCR. Bars show the E2/E3 ratio in IGF-1, MK2206 and shAKT1, shAKT2 and shAKT3 cells relative to the control cells. (Lower panels) The *FGFR2* mRNA IIIb/IIIc ratio in the same cells was also measured by quantitative PCR. Bars show the IIIb/IIIc ratio in IGF-1, MK2206 and shAKT1, shAKT2 and shAKT3 cells relative to the control cells. Error bars indicate SD. All experiments in this figure were done on three biological replicates, in triplicate. n.s: non-significant *$p<0.05$, **$p<0.01$, ***$p<0.001$, ****$p<0.0001$ (one-side unpaired t-test).

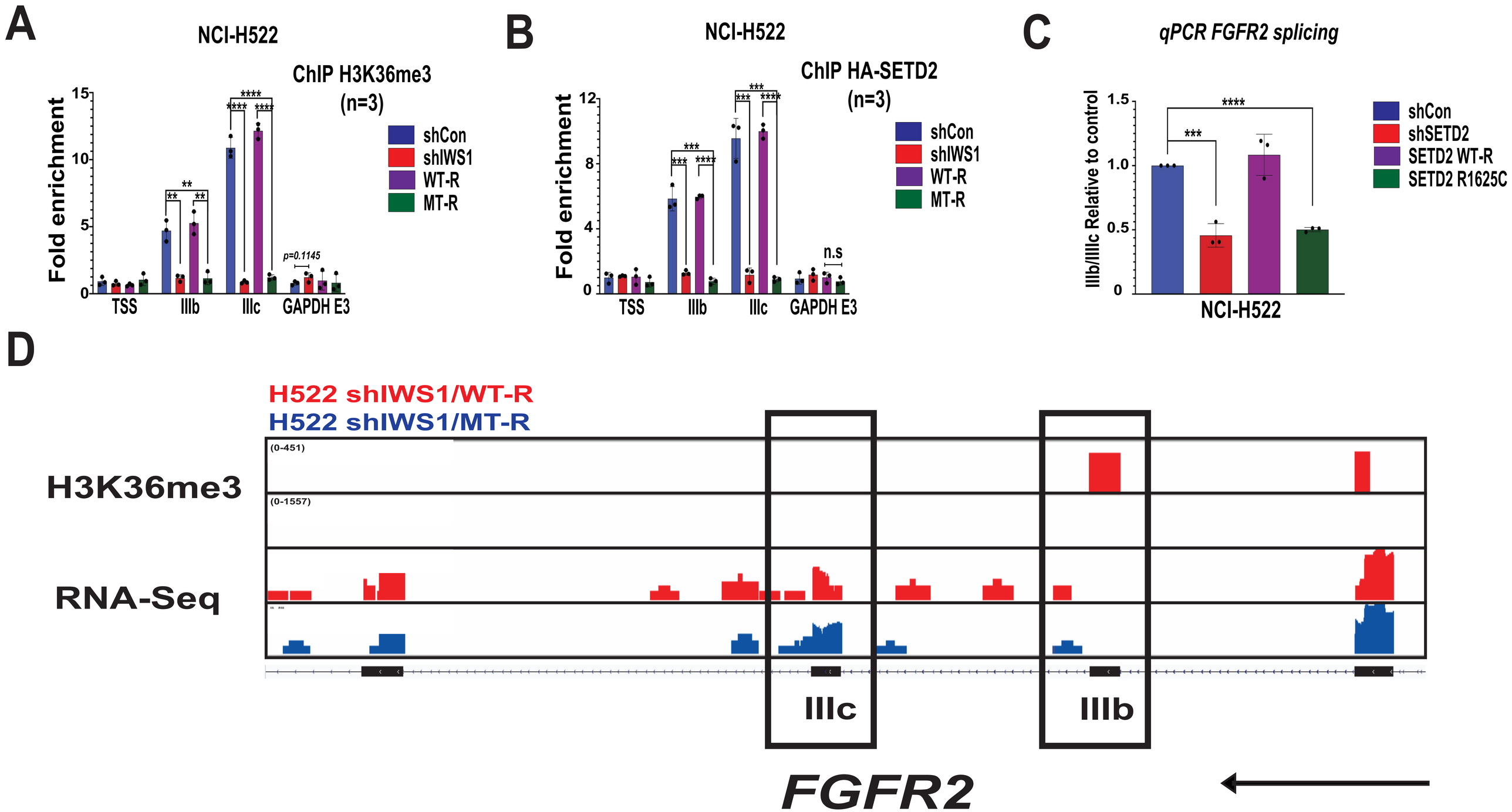

A
NCI-H522
ChIP H3K36me3
(n=3)
Fold enrichment
shCon
shIWS1
WT-R
MT-R
TSS
IIIb
IIIc
GAPDH E3
p=0.1145
B
NCI-H522
ChIP HA-SETD2
(n=3)
Fold enrichment
n.s
C
qPCR FGFR2 splicing
IIIb/IIIIc Relative to control
shSETD2
SETD2 WT-R
SETD2 R1625C
NCI-H522
D
H522 shIWS1/WT-R
H522 shIWS1/MT-R
H3K36me3
RNA-Seq
(0-451)
(0-1557)
IIIc
IIIb
FGFR2

Figure S3 (relative to Figure 3). *FGFR2* Exon IIIb exclusion, induced by IWS1 phosphorylation at Ser720/Thr721, depends on H3K36 trimethylation by SETD2.

A. ChIP assays showing the abundance of H3K36me3 on the *FGFR2* and *GAPDH* genes in shControl, shIWS1 shIWS1/WT-R and shIWS1/MT-R NCI-H522 cells. Bars show the mean fold enrichment in H3K36me3 (anti-H3K36me3 IP, vs IgG control IP) in the indicated regions of the *FGFR2* gene, in shControl, shIWS1, shIWS1/WT-R and shIWS1/MT-R cells ±SD. Data were normalized relative to the input (2%).

B. ChIP assays showing the binding of HA-SETD2 to the indicated regions of the *FGFR2* and *GAPDH* genes, in shControl, shIWS1 shIWS1/WT-R and shIWS1/MT-R NCI-H522, transduced with a lentiviral HA-SETD2 construct. The bars show the mean fold enrichment in SETD2 binding (anti-HA IP, vs IgG control IP) in the indicated regions of the *FGFR2* gene, in shControl, shIWS1, shIWS1/WT-R and shIWS1/MT-R cells ±SD. Data were normalized relative to the input (2%)

C. *The IWS1-regulated alternative RNA splicing of FGFR2, depends on enzymatically active SETD2*. The ratio of the *FGFR2* exons IIIb and IIIc was determined by quantitative RT-PCR, using the mRNA isolated from the cells from cells in Figure 3D. Again, bars show the *FGFR2* IIIb/IIIc ratio relative to the shControl. Error bars indicate SD. All experiments in this figure were done on three biological replicates, in triplicate. n.s: non-significant *$p<0.05$, **$p<0.01$, ***$p<0.001$, ****$p<0.0001$ (one-side unpaired t-test).

D. Snapshots of the integrative genomic viewer showing the distribution of histone H3K36me3 marks, within the *FGFR2* gene, obtained from a ChIP-Seq analysis, along with the distribution of RNA reads obtained from the RNA-seq analysis of shIWS1/WT-R and shIWS1/MT-R NCI-H522 cells. Scale represents reads per million (RPM). Snapshots of peaks detected in both biological replicates are shown. The black box outlines *FGFR2* exons IIIb and IIIc and the adjacent regions. Arrow shows the direction of transcription.

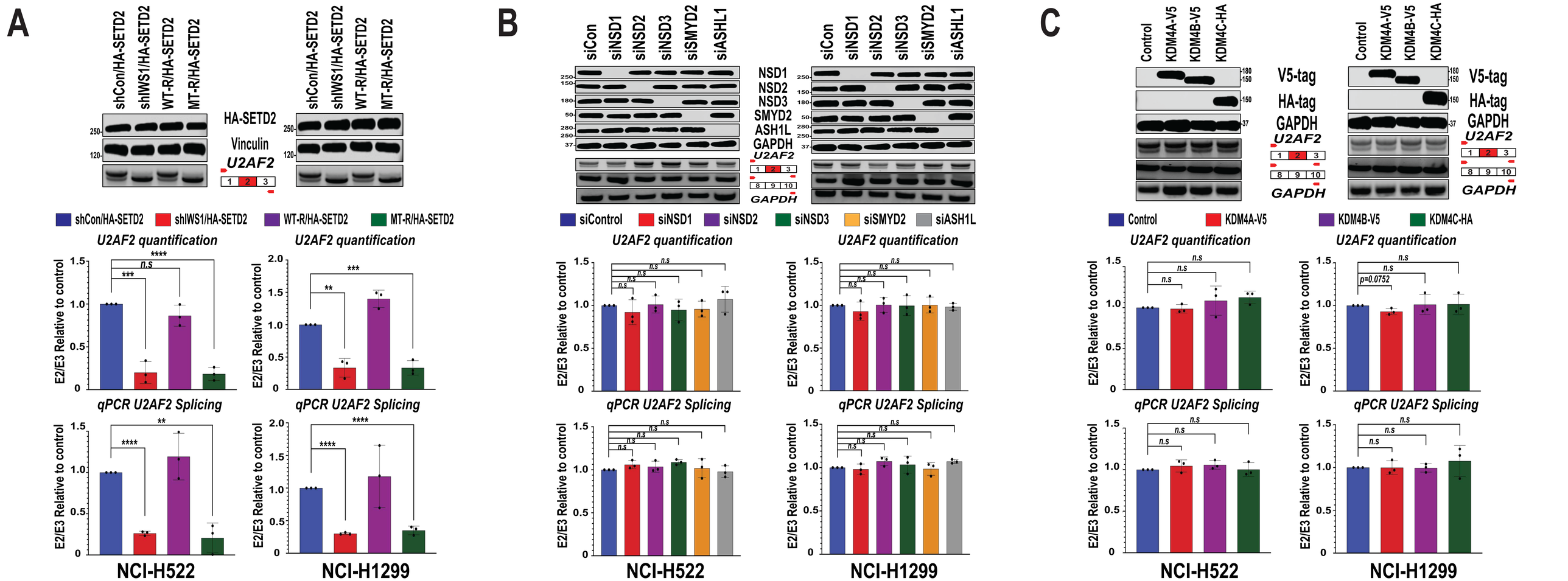

A
shCon/HA-SETD2
shIWS1/HA-SETD2
WT-R/HA-SETD2
MT-R/HA-SETD2
HA-SETD2
Vinculin
U2AF2
250
120
1 2 3
shCon/HA-SETD2 shIWS1/HA-SETD2 WT-R/HA-SETD2 MT-R/HA-SETD2
U2AF2 quantification
E2/E3 Relative to control
qPCR U2AF2 Splicing
NCI-H522
NCI-H1299
B
siCon siNSD1 siNSD2 siNSD3 siSMYD2 siASHL1
NSD1
NSD2
NSD3
SMYD2
ASH1L
GAPDH
U2AF2
8 9 10
siControl siNSD1 siNSD2 siNSD3 siSMYD2 siASH1L
C
Control KDM4A-V5 KDM4B-V5 KDM4C-HA
V5-tag
HA-tag
GAPDH
U2AF2
Control KDM4A-V5 KDM4B-V5 KDM4C-HA
n.s
p=0.0752

Figure S4 (relative to Figure 3). *U2AF2* Exon 2 inclusion, induced by IWS1 phosphorylation at Ser720/Thr721, depends on H3K36 trimethylation by SETD2.

A. *SETD2 does not rescue the shIWS1 and shIWS1/MT-R U2AF2 alternative splicing phenotype.* (Upper panels-First two lines) Lysates of shControl, shIWS1 shIWS1/WT-R and shIWS1/MT-R NCI-H522 and NCI-H1299 cells, transduced with a lentiviral construct of wild type HA-SETD2, were probed with the indicated antibodies. (Upper panels-third line) RT-PCR reactions, using mRNA derived from the same cells, and oligonucleotide primers mapping in *U2AF2* exons 1 and 3 show that SETD2 does not rescue the shIWS1-induced exclusion of exon 2 from the *U2AF2* mRNA. (Middle panel) *U2AF2 E2/E3* ratio in HA-SETD2-transduced shIWS1, shIWS1/WT-R and shIWS1/MT-R cells, relative to HA-SETD2-transduced shControl cells. Data based on quantification of the RT-PCR products above. (Lower panel) The RT-PCR results were confirmed by quantitative RT-PCR. Bar graphs show again the E2/E3 ratio in HA-SETD2-transduced shIWS1, shIWS1/WT-R and shIWS1/MT-R cells, relative to HA-SETD2-transduced shControl cells ± SD.

B. (Upper panels. first 6 lines) The listed H3K36 methyltransferase were efficiently knocked down in NCI-H522 and NCI-H1299 cells, by siRNA transfection, as determined by probing western blots of transfected cell lysates with the indicated antibodies. (Upper panels, lines 7 and 9) RT-PCR, using RNA isolated from these cells, and *U2AF2* exon 1 and exon 3 or exon 8 and 10 oligonucleotide primers shows that none of these methyltransferases control the splicing pattern of the *U2AF2* mRNA. (Middle panels) *U2AF2 E2/E3* ratio in the siRNA-transfected cells, relative to the siControl cells. Data based on quantification of the RT-PCR products above. (Lower panels) The RT-PCR results were confirmed by quantitative RT-PCR. Bar graphs show the *E2/E3* ratio in NCI-H522 and NCI-H1299 cells, after the knockdown of the indicated methyltransferases, relative to the siControl cells ± SD.

C. (Upper Panels, first 3 lines) The listed H3K36 demethylases were overexpressed in NCI-H522 and NCI-H1299 cells, using viral constructs (pLX304 constructs for KDM4A and KDM4B and a pBabe-puro construct for KDM4C). Expression was determined by probing western blots of transduced cell lysates with the indicated antibodies. (Upper panels 4th and 5th lines) RT-PCR, using RNA isolated from these cells, and *U2AF2* exon 1 and exon 3 or exon 8 and 10 oligonucleotide primers shows that none of these demethylases control the splicing pattern of the *U2AF2* mRNA. (Middle panels) *U2AF2 E2/E3* ratio in the transduced cells, relative to the vector-transduced (pLX304 EV plus pBabe-puro EV) cells. Data based on quantification of the RT-PCR data above. (Lower panels) The RT-PCR results were confirmed by quantitative RT-PCR. Bar graphs show the *E2/E3* ratio in NCI-H522 and NCI-H1299 cells, after transduction with the indicated demethylases, relative to the Control cells ± SD. All assays in this figure were done in triplicate, on three biological replicates. n.s : non-significant *$p<0.05$, **$p<0.01$, ***$p<0.001$, ****$p<0.0001$. (one-side unpaired t-test)

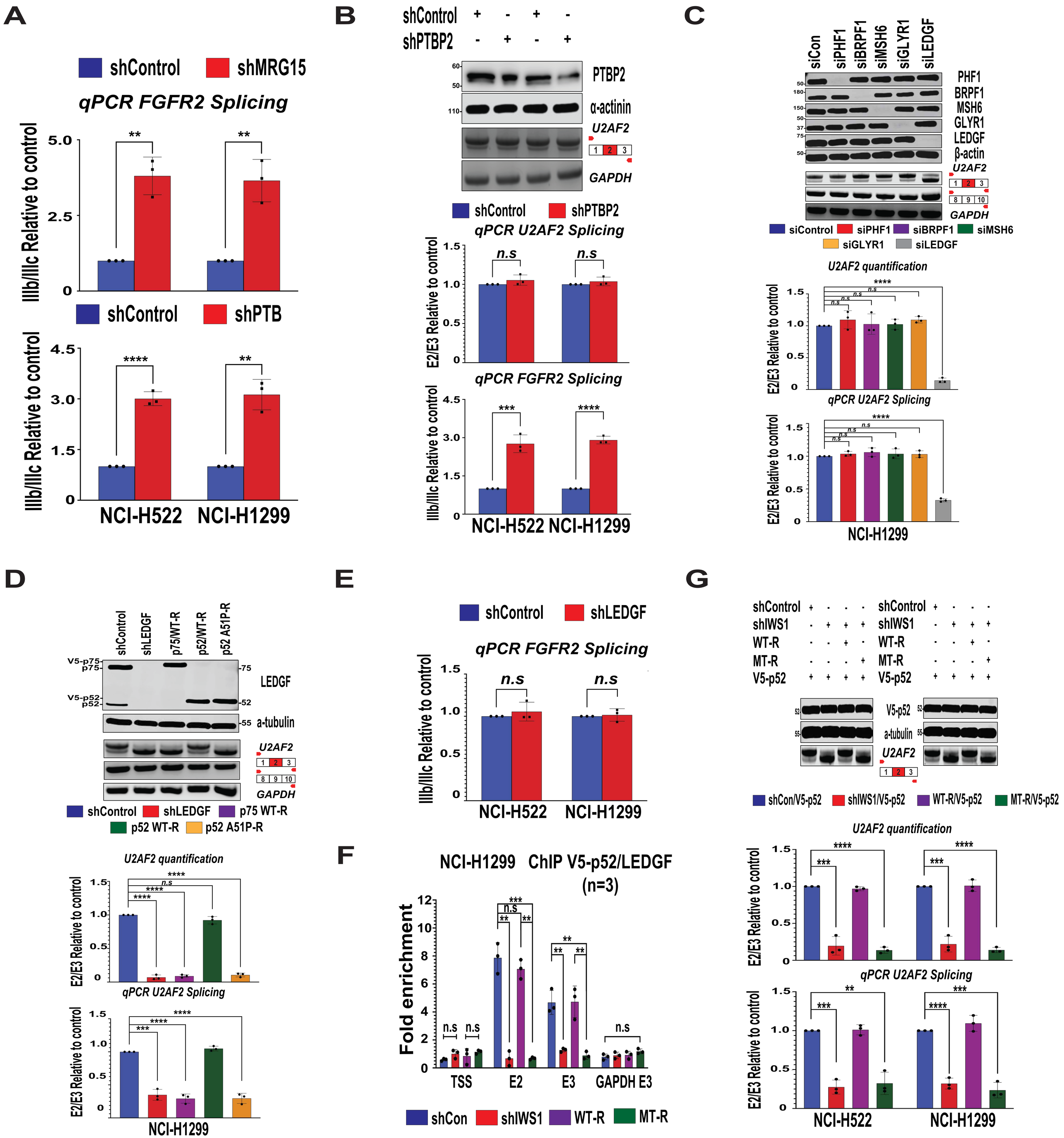

A
shControl
shMRG15
qPCR FGFR2 Splicing
IIIb/IIIc Relative to control
NCI-H522
NCI-H1299
shPTB
B
shPTBP2
PTBP2
α-actinin
U2AF2
GAPDH
qPCR U2AF2 Splicing
E2/E3 Relative to control
n.s
qPCR FGFR2 Splicing
C
siCon
siPHF1
siBRPF1
siMSH6
siGLYR1
siLEDGF
PHF1
BRPF1
MSH6
GLYR1
LEDGF
β-actin
siControl
U2AF2 quantification
D
shLEDGF
p75/WT-R
p52/WT-R
p52 A51P-R
LEDGF
a-tubulin
p75 WT-R
p52 WT-R
p52 A51P-R
E
F
NCI-H1299 ChIP V5-p52/LEDGF
(n=3)
Fold enrichment
TSS
E2
E3
GAPDH E3
shCon
shIWS1
WT-R
MT-R
G
V5-p52
shCon/V5-p52
shIWS1/V5-p52
WT-R/V5-p52
MT-R/V5-p52

Figure S5 (relative to Figure 4). The regulation of the alternative splicing of the *U2AF2* exon 2 by *IWS1* phosphorylation, depends on the p52 isoform of the H3K36me3 reader *LEDGF* and its splicing partner *SRSF1*.

- **A.** *The H3K36me3 reader MRG15 and its binding partner PTB, control the FGFR2 alternative RNA splicing.* The ratio of the *FGFR2* exons IIIb and IIIc was determined by quantitative RT-PCR, using the mRNA isolated from the NCI-H522 (upper panel) and NCI-H1299 (lower panel) cells in Figure 4A. Again, bars show the *FGFR2* IIIb/IIIc ratio relative to the shControl. Error bars indicate SD.
- **B.** *The paralog of PTB, PTBP2 plays no role in the regulation of U2AF2 alternative RNA splicing.* (Upper panel) Western blots of lysates of NCI-H522 and NCI-H1299 cells transduced with shControl and shPTBP2 were probed with the indicated antibodies. (Middle and lower panel) Quantitative RT-PCR, using RNA isolated from these cells shows that the knockdown of PTBP2 does not affect the alternative splicing of *U2AF2,* but affects the *FGFR2* splicing pattern. Bar graphs show the *U2AF2* E2/E3 and *FGFR2* IIIb/IIIc ratio in shPTBP2 NCI-H522 and NCI-H1299 cells, relative to the shControl ± SD.
- **C.** The knockdown of the indicated histone H3K36me3 readers in NCI-H1299 cells confirmed that only the knockdown of LEDGF results in the exclusion of exon 2 from the mature *U2AF2* mRNA transcript. This is a repeat of the experiment in figure 4C, which was done in NCI-H522 cells. The results confirmed that LEDGF controls the mRNA splicing of *U2AF2* in multiple cell lines.
- **D.** The expression of LEDGF-p75-WT, LEDGF-p52-WT and LEDGF-p52A51P (a LEDGF mutant, which has lost the ability to bind H3K36me3) in shLEDGF-transduced NCI-H1299 cells, confirmed that only LEDGF-p52-WT can rescue the shLEDGF-induced exclusion of exon 2 from the mature *U2AF2* mRNA transcript. This is a repeat of the experiment in figure 4C, which was done in NCI-H522 cells.

E. *The H3K36me3 reader LEDGF, plays no role in the regulation of FGFR2 alternative RNA splicing.* The ratio of the *FGFR2* exons IIIb and IIIc was determined by quantitative RT-PCR, using the mRNA isolated from the NCI-H522 and NCI-H1299 cells in Figure 4C. Bars show the *FGFR2* IIIb/IIIc ratio in shLEDGF relative to the shControl. Error bars indicate SD.

F. ChIP assays addressing the binding of V5-LEDGF-p52 in TSS, exon 2 and exon 3 of U2AF2, in shControl, shIWS1, shIWS1/WT-R and shIWS1/MT-R NCI-H1299 cells, confirmed the binding of this LEDGF isoform in both, exon 2 and exon 3. This is a repeat of the experiment in figure 4D, which was done in NCI-H522 cells. The results confirmed that the p52 isoform of wild type LEDGF controls the mRNA splicing of *U2AF2,* by binding the U2AF2 exons 2 and 3, in multiple cell lines.

G. *V5-LEDGF-p52 does not rescue the shIWS1 and shIWS1/MT-R U2AF2 alternative splicing phenotype.* (Upper panel-First two lines) shControl, shIWS1 shIWS1/WT-R and shIWS1/MT-R NCI-H522 (left panel) and NCI-H1299 (right panel) cells were transduced with a lentiviral construct of V5-p52/LEDGF. The expression of V5-p52-LEDGF was confirmed by probing lysates of these cells with an anti-V5 or an anti-α-tubulin (loading control) antibody. (Upper panel-Third line) RT-PCR, using RNA isolated from these cells and *U2AF2* exons 1 and 3 oligonucleotide primers shows that V5-p52-LEDGF does not rescue the shIWS1 and shIWS1/MT-R *U2AF2* alternative splicing phenotype. (Middle panel) Bars show the mean *E2/E3 U2AF2* RNA ratio in NCI-H522 and NCI-H1299 shIWS1, shIWS1/WT rescue and shIWS1/MT rescue cells, in the experiment in the top panel, normalized to the shControl cells ± SD (one-side unpaired t-test) (Lower panel) The RT-PCR results were confirmed by quantitative RT-PCR. All assays in this figure were done in triplicate, on three biological replicates. n.s : non-significant *$p<0.05$, **$p<0.01$, ***$p<0.001$, ****$p<0.0001$ (one-side unpaired t-test).

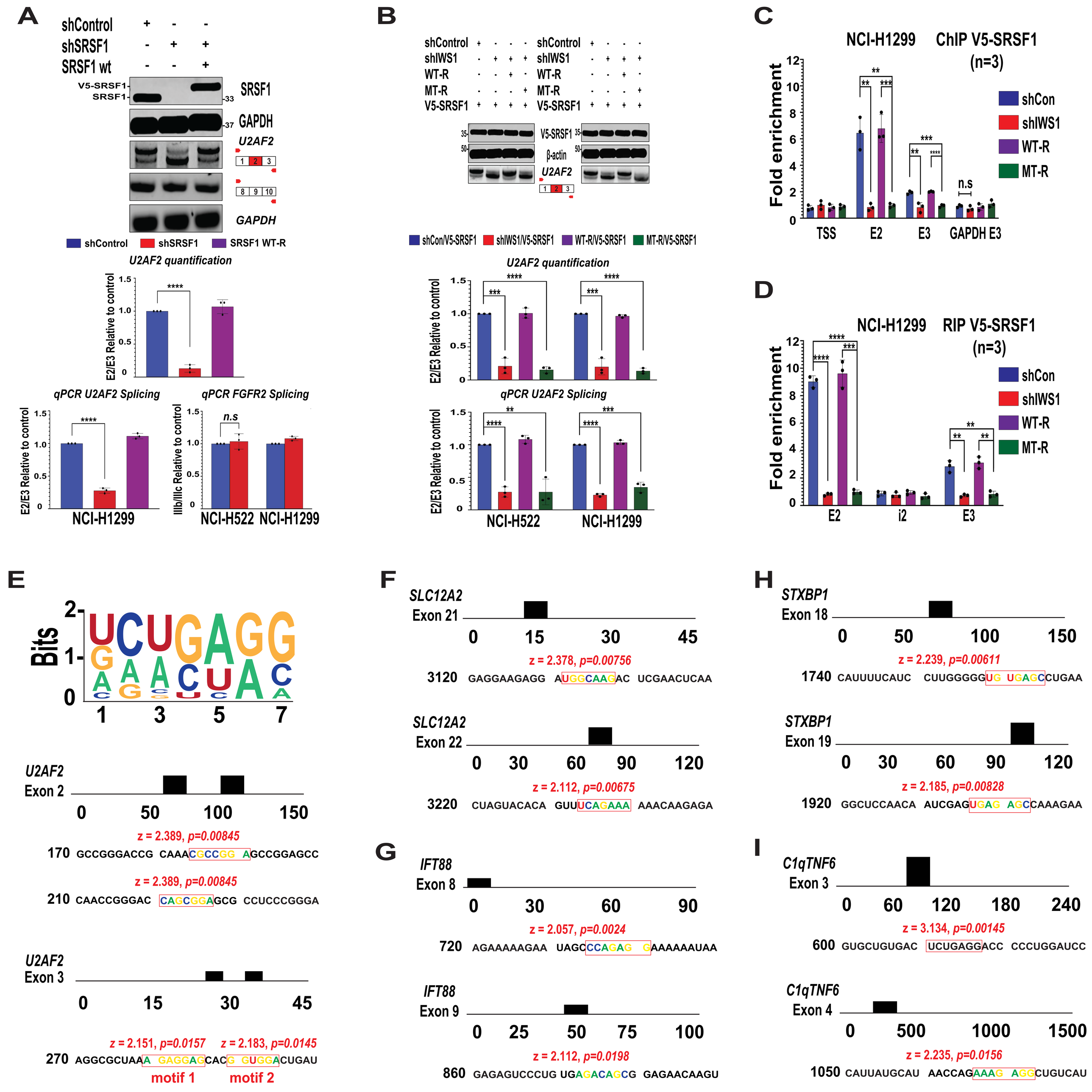

A
shControl
shSRSF1
SRSF1 wt
V5-SRSF1
SRSF1
SRSF1
GAPDH
U2AF2
GAPDH
shControl
shSRSF1
SRSF1 WT-R
U2AF2 quantification
E2/E3 Relative to control
qPCR U2AF2 Splicing
qPCR FGFR2 Splicing
IIIb/IIIc Relative to control
n.s
NCI-H1299
NCI-H522
B
shControl
shIWS1
WT-R
MT-R
V5-SRSF1
β-actin
U2AF2
shCon/V5-SRSF1
shIWS1/V5-SRSF1
WT-R/V5-SRSF1
MT-R/V5-SRSF1
U2AF2 quantification
qPCR U2AF2 Splicing
C
NCI-H1299 ChIP V5-SRSF1
(n=3)
Fold enrichment
shCon
shIWS1
WT-R
MT-R
TSS
E2
E3
GAPDH E3
n.s
D
NCI-H1299 RIP V5-SRSF1
(n=3)
i2
E
Bits
U2AF2
Exon 2
z = 2.389, p=0.00845
170 GCCGGGACCG CAAACGCCGG AGCCGGAGCC
z = 2.389, p=0.00845
210 CAACCGGGAC CAGCGGAGCG CCUCCCGGGA
U2AF2
Exon 3
z = 2.151, p=0.0157
z = 2.183, p=0.0145
270 AGGCGCUAAA GAGGAGCACG GUGGACUGAU
motif 1
motif 2
F
SLC12A2
Exon 21
z = 2.378, p=0.00756
3120 GAGGAAGAGG AUGGCAAGAC UCGAACUCAA
SLC12A2
Exon 22
z = 2.112, p=0.00675
3220 CUAGUACACA GUUUCAGAAA AAACAAGAGA
G
IFT88
Exon 8
z = 2.057, p=0.0024
720 AGAAAAAGAA UAGCCCAGAG GAAAAAAUAA
IFT88
Exon 9
z = 2.112, p=0.0198
860 GAGAGUCCCUG UGAGACAGCG GAGAACAAGU
H
STXBP1
Exon 18
z = 2.239, p=0.00611
1740 CAUUUUCAUC CUUGGGGGUG UGAGCCUGAA
STXBP1
Exon 19
z = 2.185, p=0.00828
1920 GGCUCCAACA AUCGAGUGAG AGCCAAAGAA
I
C1qTNF6
Exon 3
z = 3.134, p=0.00145
600 GUGCUGUGAC UCUGAGGACC CCCUGGAUCC
C1qTNF6
Exon 4
z = 2.235, p=0.0156
1050 CAUUAUGCAU AACCAGAAAG AGGCUGUCAU

Figure S6 (relative to Figure 4). The regulation of the alternative splicing of the *U2AF2* exon 2 by *IWS1* phosphorylation, depends on the p52 isoform of the H3K36me3 reader *LEDGF* and its splicing partner *SRSF1*.

A. *The alternative RNA splicing of the U2AF2 exon 2 depends on SRSF1.* (Upper panel) The experiment in figure 4E, which was carried out in NCI-H522 cells, was repeated in NCI-H1299 cells, with identical results. (Lower panel) As a confirmation of the independence of *FGFR2* RNA splicing to SRSF1, quantitative RT-PCR from NCI-H522 cells in Figure 4E and NCI-H1299 cells above, was used to determine the *FGFR2* IIIb/IIIc ratio. Again, bars show the *FGFR2* IIIb/IIIc ratio in shSRSF1 relative to the shControl. Error bars indicate SD.

B. *SRSF1 does not rescue the shIWS1 and shIWS1/MT-R U2AF2 alternative splicing phenotype.* (Upper panel-First two lines) shControl, shIWS1 shIWS1/WT-R and shIWS1/MT-R NCI-H522 (left panel) and NCI-H1299 (right panel) cells were transduced with a lentiviral construct of V5-SRSF1. The expression of V5-SRSF1 was confirmed by probing lysates of these cells with an anti-V5 or an anti-β-actin (loading control) antibody. (Upper panel-Third line) RT-PCR, using RNA isolated from these cells and *U2AF2* exons 1 and 3 oligonucleotide primers shows that V5-SRSF1 does not rescue the shIWS1 and shIWS1/MT-R *U2AF2* alternative splicing phenotype. (Middle panel) Bars show the mean *E2/E3 U2AF2* RNA ratio in NCI-H522 and NCI-H1299 shIWS1, shIWS1/WT-R and shIWS1/MT rescue cells, in the experiment in the top panel, normalized to the shControl cells ± SD (one-side unpaired t-test) (Lower panel) The RT-PCR results were confirmed by quantitative RT-PCR.

C. The experiments in NCI-H522 cells, in figure 4F, were repeated in NCI-H1299 cells with identical results.

D. The experiments in NCI-H522 cells, in figure 4G, were repeated in NCI-H1299 cells with identical results. All assays in this figure were done in triplicate, on three biological replicates. n.s : non-significant *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001. (one-side unpaired t-test).

E. (Upper panel) The 7-nt SRSF1 binding motif, involved in the regulation of alternative RNA splicing by SRSF1 (Anczuków et al., 2015[52]) (Middle and Lower panels) Boxes show the location of the SRSF1 binding motifs in *U2AF2* exons 2 and 3 and their nucleotide sequence. The height of the boxes is proportional to the -log(*p-value*) and it is presented relative to the height of the box mapping between nucleotides 184 and 192. The binding scores (z-scores) and p values are calculated using the RBPmap pipeline (http://rbpmap.technion.ac.il/index.html)

F., G., H., and I. The validated splicing targets of IWS1 phosphorylation in Fig. S1 (*SLC12A2, IFT88, STXBP1 and C1qTNF6*) were subjected to SRSF1 binding sites analysis as above.

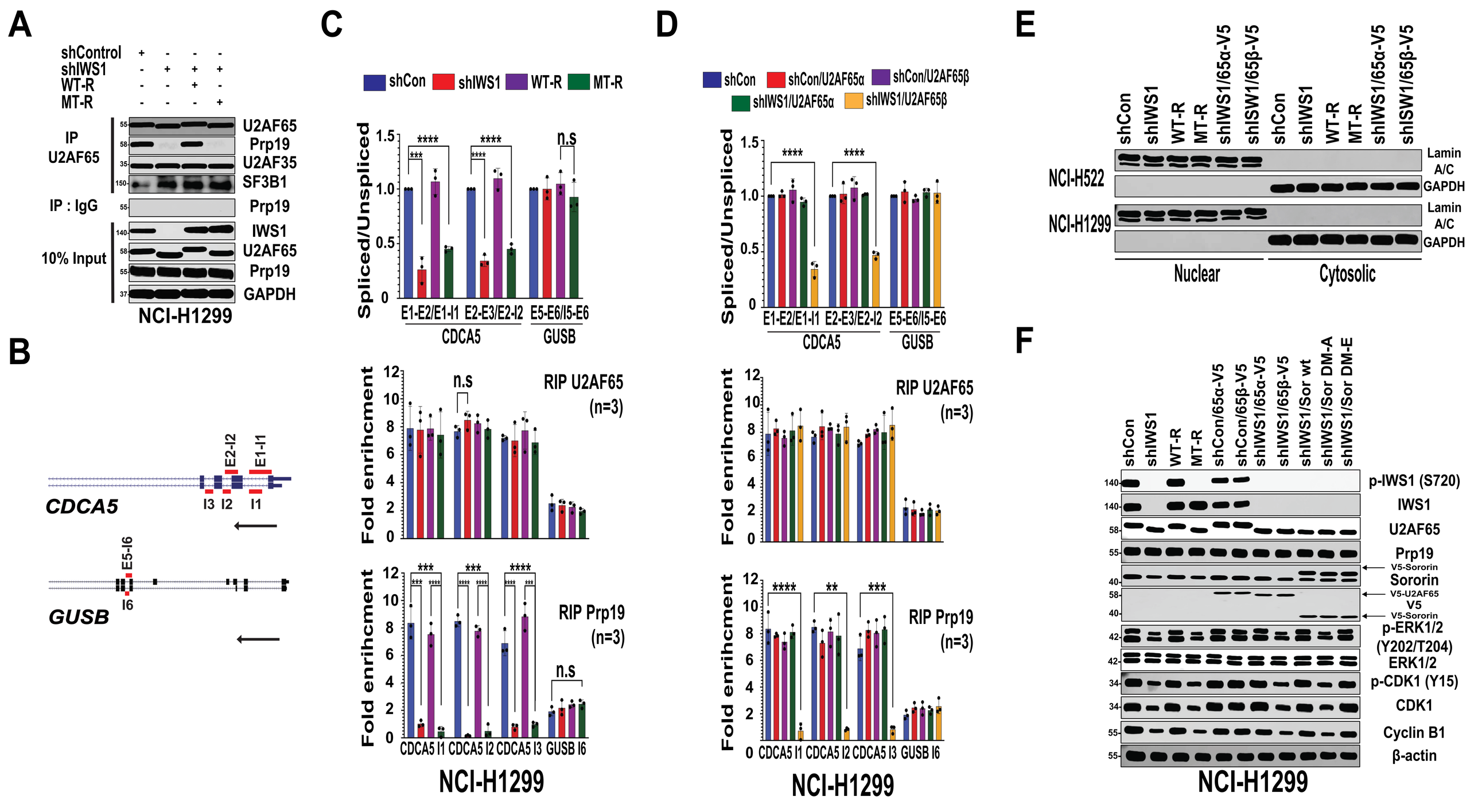

A
shControl
shIWS1
WT-R
MT-R
IP U2AF65
U2AF65
Prp19
U2AF35
SF3B1
IP : IgG
Prp19
10% Input
IWS1
U2AF65
Prp19
GAPDH
NCI-H1299
B
CDCA5
E2-I2
E1-I1
I3 I2 I1
GUSB
E5-I6
I6
C
shCon
shIWS1
WT-R
MT-R
Spliced/Unspliced
E1-E2/E1-I1 E2-E3/E2-I2 E5-E6/I5-E6
CDCA5
GUSB
n.s
RIP U2AF65
(n=3)
Fold enrichment
RIP Prp19
(n=3)
CDCA5 I1 CDCA5 I2 CDCA5 I3 GUSB I6
NCI-H1299
D
shCon
shCon/U2AF65α
shCon/U2AF65β
shIWS1/U2AF65α
shIWS1/U2AF65β
Spliced/Unspliced
E1-E2/E1-I1 E2-E3/E2-I2 E5-E6/I5-E6
CDCA5
GUSB
RIP U2AF65
(n=3)
Fold enrichment
RIP Prp19
(n=3)
CDCA5 I1 CDCA5 I2 CDCA5 I3 GUSB I6
NCI-H1299
E
shCon
shIWS1
WT-R
MT-R
shIWS1/65α-V5
shSW1/65β-V5
NCI-H522
NCI-H1299
Lamin A/C
GAPDH
Nuclear
Cytosolic
F
shCon
shIWS1
WT-R
MT-R
shCon/65α-V5
shCon/65β-V5
shIWS1/65α-V5
shIWS1/65β-V5
shIWS1/Sor wt
shIWS1/Sor DM-A
shIWS1/Sor DM-E
p-IWS1 (S720)
IWS1
U2AF65
Prp19
V5-Sororin
Sororin
V5-U2AF65
V5
V5-Sororin
p-ERK1/2
(Y202/T204)
ERK1/2
p-CDK1 (Y15)
CDK1
Cyclin B1
β-actin
NCI-H1299

Figure S7 (relative to Figure 5). IWS1 phosphorylation controls the CDCA5/ERK phosphorylation feedback loop, through U2AF2 alternative RNA splicing.

A. The experiment in NCI-H522 cells, in figure 5C, was repeated in NCI-H1299 cells with identical results.

B. UCSC browser snapshot showing the exon position of the relevant portions of the human *CDCA5* and *GUSB* genes. The map position of the sequences amplified by PCR is indicated with red lines. The arrows indicate the direction of transcription.

C. The quantitative RT-PCR and RIP experiments in NCI-H522 cells, in figure 5D, were repeated in NCI-H1299 cells with identical results.

D. The quantitative RT-PCR and RIP experiments in NCI-H522 cells, in figure 5E, were repeated in NCI-H1299 cells with identical results.

E. The cytosolic and nuclear compartment of the indicated NCI-H522 (upper two lines) and NCI-H1299 cells (lower two lines) were probed with the indicated antibodies, to validate the fractionation.

F. The experiment in NCI-H522 cells, in figure 5G, was repeated in NCI-H1299 cells with identical results. All assays in this figure were done in triplicate, on three biological replicates. n.s : non-significant *$p<0.05$, **$p<0.01$, ***$p<0.001$, ****$p<0.0001$. (one-side unpaired t-test).

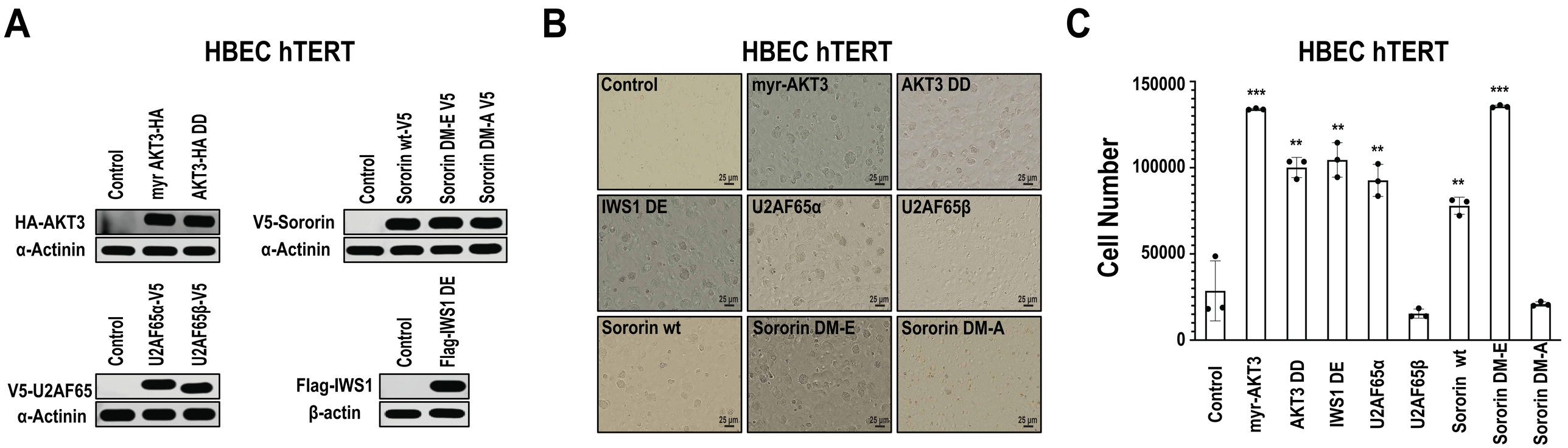

A
HBEC hTERT
Control
myr AKT3-HA
AKT3-HA DD
HA-AKT3
α-Actinin
Control
Sororin wt-V5
Sororin DM-E V5
Sororin DM-A V5
V5-Sororin
α-Actinin
Control
U2AF65α-V5
U2AF65β-V5
V5-U2AF65
α-Actinin
Control
Flag-IWS1 DE
Flag-IWS1
β-actin
B
HBEC hTERT
Control
myr-AKT3
AKT3 DD
IWS1 DE
U2AF65α
U2AF65β
Sororin wt
Sororin DM-E
Sororin DM-A
25 µm
C
HBEC hTERT
Cell Number
150000
100000
50000
0
***
**
Control
myr-AKT3
AKT3 DD
IWS1 DE
U2AF65α
U2AF65β
Sororin wt
Sororin DM-E
Sororin DM-A

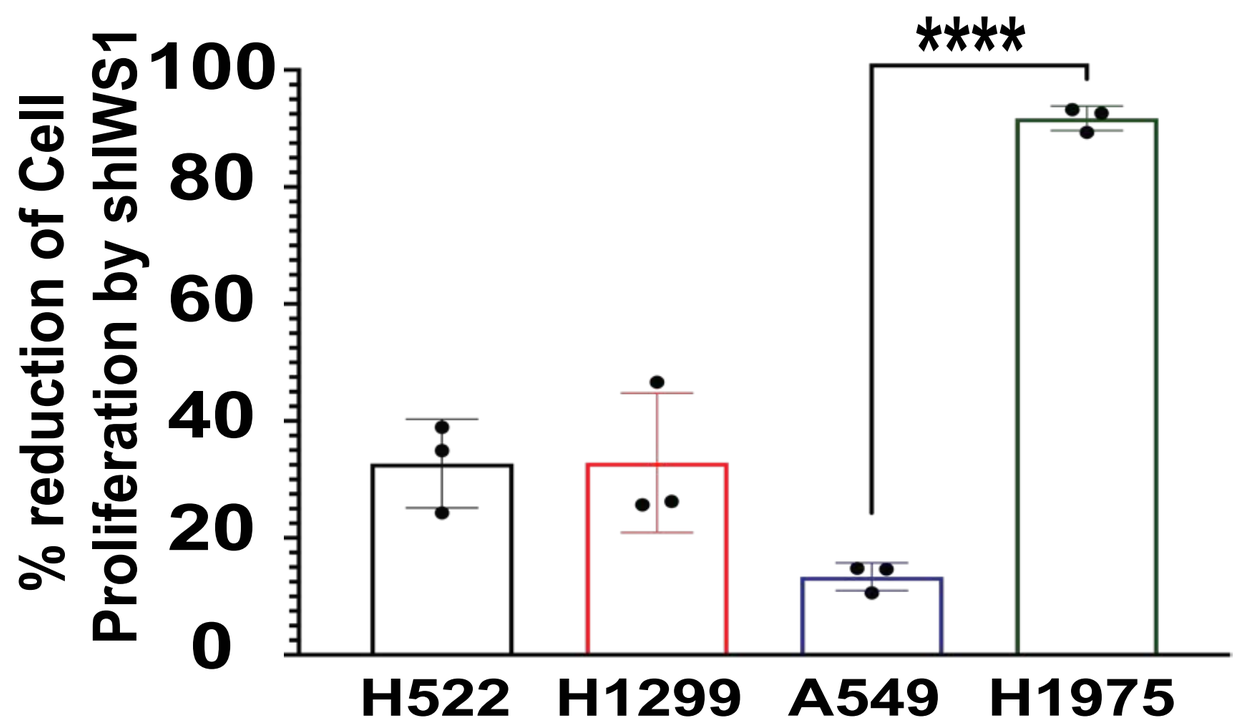
Figure S8 (relative to Figure 5). IWS1 phosphorylation by AKT3 promotes cellular transformation through *U2AF2* alternative RNA splicing and Sororin expression.

A. Western blots of lysates of HBEC hTERT cells, transduced with the indicated constructs, were probed with the indicated antibodies.

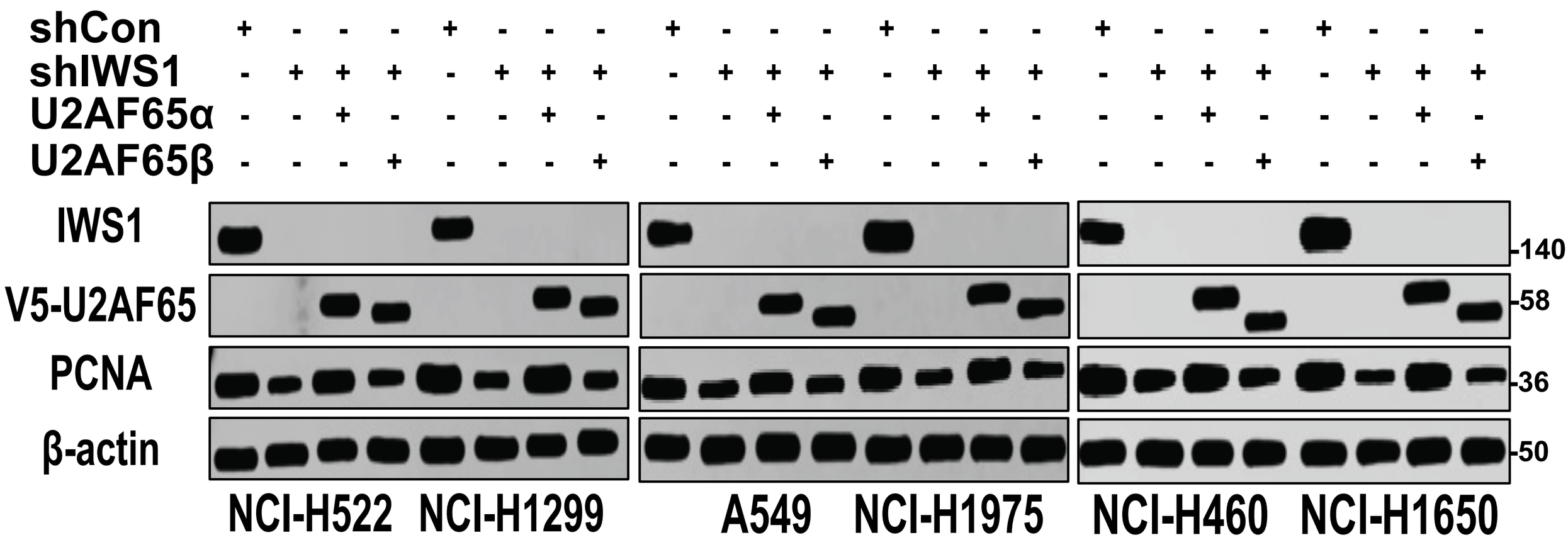
B. *The AKT/IWS1 axis promotes cellular transformation*. HBEC hTERT cells in Figure S8A were subjected to cell transformation assay for 7 days. At that time point the cells were imaged at Incucyte live cell imager. Scale bar in the right corner of each image.

C. After imaging, the cells were incubated for 4 hours in WST solution as outlined in Cell Transformation Assay Kit (Abcam) and the number of transformed cells was measured as the absorbance in 450nm. All assays in this figure were done in triplicate, on three biological replicates. n.s : non-significant *$p<0.05$, **$p<0.01$, ***$p<0.001$, ****$p<0.0001$. (one-side unpaired t-test).

A


% reduction of Cell Proliferation by shIWS1
100
80
60
40
20
0
****
H522
H1299
A549
H1975

B


shCon
shIWS1
U2AF65α
U2AF65β
IWS1
V5-U2AF65
PCNA
β-actin
-140
-58
-36
-50
NCI-H522
NCI-H1299
A549
NCI-H1975
NCI-H460
NCI-H1650

C

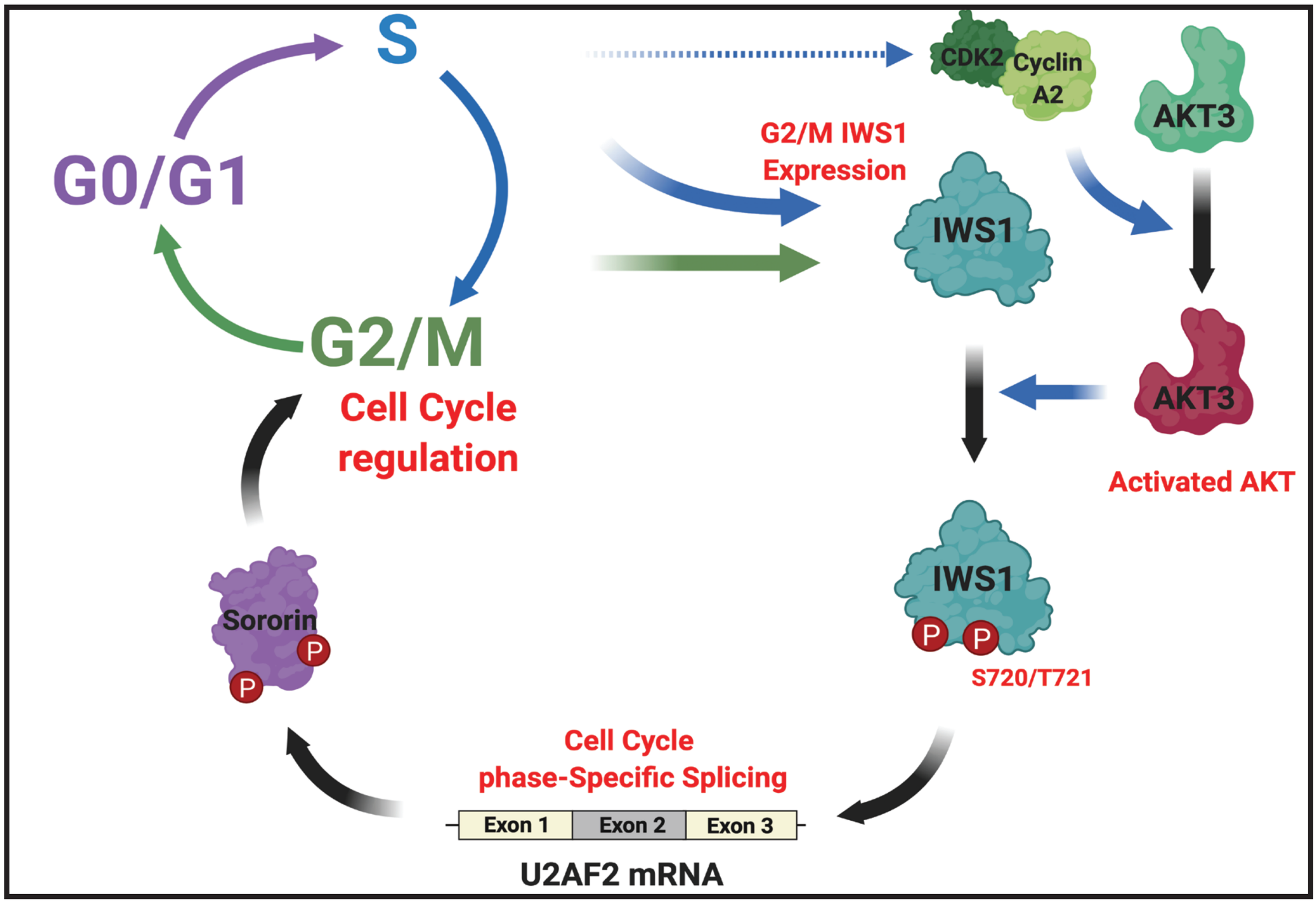

S
G0/G1
G2/M
Cell Cycle regulation
CDK2
Cyclin A2
AKT3
G2/M IWS1 Expression
IWS1
AKT3
Activated AKT
IWS1
P
P
S720/T721
Sororin
P
P
Cell Cycle phase-Specific Splicing
Exon 1
Exon 2
Exon 3
U2AF2 mRNA

Figure S9 (relative to Figure 6). IWS1 phosphorylation promotes cell proliferation by controlling the Sororin/ERK phosphorylation feedback loop, through *U2AF2* RNA splicing.

A. Percent reduction of cell proliferation, induced by shIWS1 in the indicated cell lines. Bars show the mean reduction of cell proliferation, derived from the end-point data in Figure 6A and 6C. The error bars show the SD of at least 3 biological replicates. The statistics were performed using one-sided unpaired t-test (NCI-H522 vs A549 : p = 0.0225, NCI-H1299 vs A549 : 0.0493, A549 vs NCI-H1975 : p = 0.0001, NCI-H522 vs NCI-H1975 : p = 0.0174, NCI-H1299 vs NCI-H1975 : p = 0.0011).

B. Lysates derived from the indicated cells were probed with the indicated antibodies. In agreement with the data in Fig. 6A, 6C PCNA was downregulated by shIWS1 in both cell lines, and its downregulation was rescued by U2AF65α, but not U2AF65β.

C. *Model of the regulation of U2AF2 alternative RNA splicing, downstream of IWS1 phosphorylation, through the cell cycle.* IWS1 is expressed during S and G2/M (this report) and published data indicate that AKT is also activated as the cells enter S phase, via phosphorylation by CDK2/Cyclin A2 (Liu et al, 2014[74]). This induces a cell cycle specific alternative RNA splicing, resulting in inclusion of exon 2 in *U2AF2* mRNA. Subsequently, this isoform regulates the splicing and expression of Sororin, which in turn regulates the progression through the G2/M phase of the cell cycle.

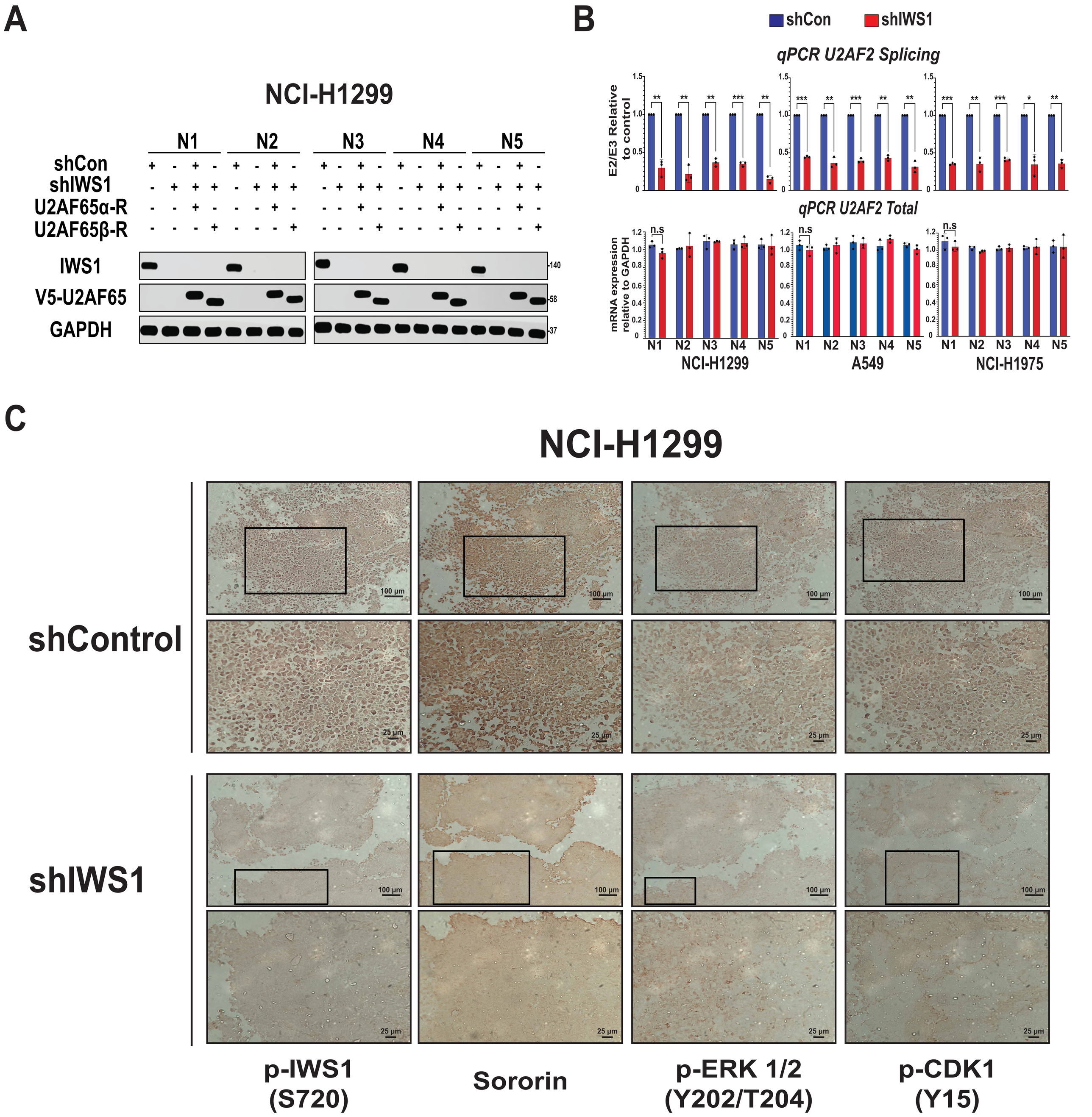

A
NCI-H1299
N1
N2
N3
N4
N5
shCon
shIWS1
U2AF65α-R
U2AF65β-R
IWS1
V5-U2AF65
GAPDH
140
58
37
B
shCon
shIWS1
qPCR U2AF2 Splicing
E2/E3 Relative to control
qPCR U2AF2 Total
mRNA expression relative to GAPDH
n.s
NCI-H1299
A549
NCI-H1975
C
NCI-H1299
shControl
shIWS1
100 µm
25 µm
p-IWS1 (S720)
Sororin
p-ERK 1/2 (Y202/T204)
p-CDK1 (Y15)

Figure S10 (Relative to Figure 7). IWS1 phosphorylation controls tumor growth *in vivo*, by regulating the U2AF2/Sororin/ERK axis.

A. Cell lysates derived from NCI-H1299 shControl, shIWS1, shIWS1/U2AF65α and shIWS1/U2AF65β mouse xenografts, were probed with the indicated antibodies.

B. *IWS1 phosphorylation controls U2AF2 alternative RNA splicing in vivo.* Quantitative RT-PCR showing the *E2/E3 U2AF2* ratio (upper panel) or total *U2AF2* in lysates derived from tumors in Fig 7C. Bars show *U2AF2* E2/E3 ratio or the mean *U2AF2* levels, relative to GAPDH, in tumors derived from NCI-H1299 (left), A549 (middle) and NCI-H1975 (right) shControl and shIWS1 cells. The error bars are SD of at least 3 technical replicates. **p<0.001, ***p<0.001. (paired t-test), n.s : non-significant (paired t-test)

C. *IWS1 controls the Sororin/ERK phosphorylation axis in vivo.* (Upper panels) Phospho-IWS1 (Ser720), Sororin phosphor-ERK (Y202/T204) and phospho-CDK1 (Y15) staining, of NCI-H1299 shControl (upper) and shIWS1 (lower) tumor xenografts. Formalin-fixed, paraffin-embedded tumor samples from the experiment in Figure 7 were stained with the indicated antibodies. Secondary antibody was HRP-labelled. The staining of individual tumors with different antibodies was done, using sequential tumor sections, so that we could determine whether the expression and/or phosphorylation of proteins of interest spatially overlapped. Boxes delineate the area of higher magnification, shown in the image below. Scale in the right lower corner of each image.

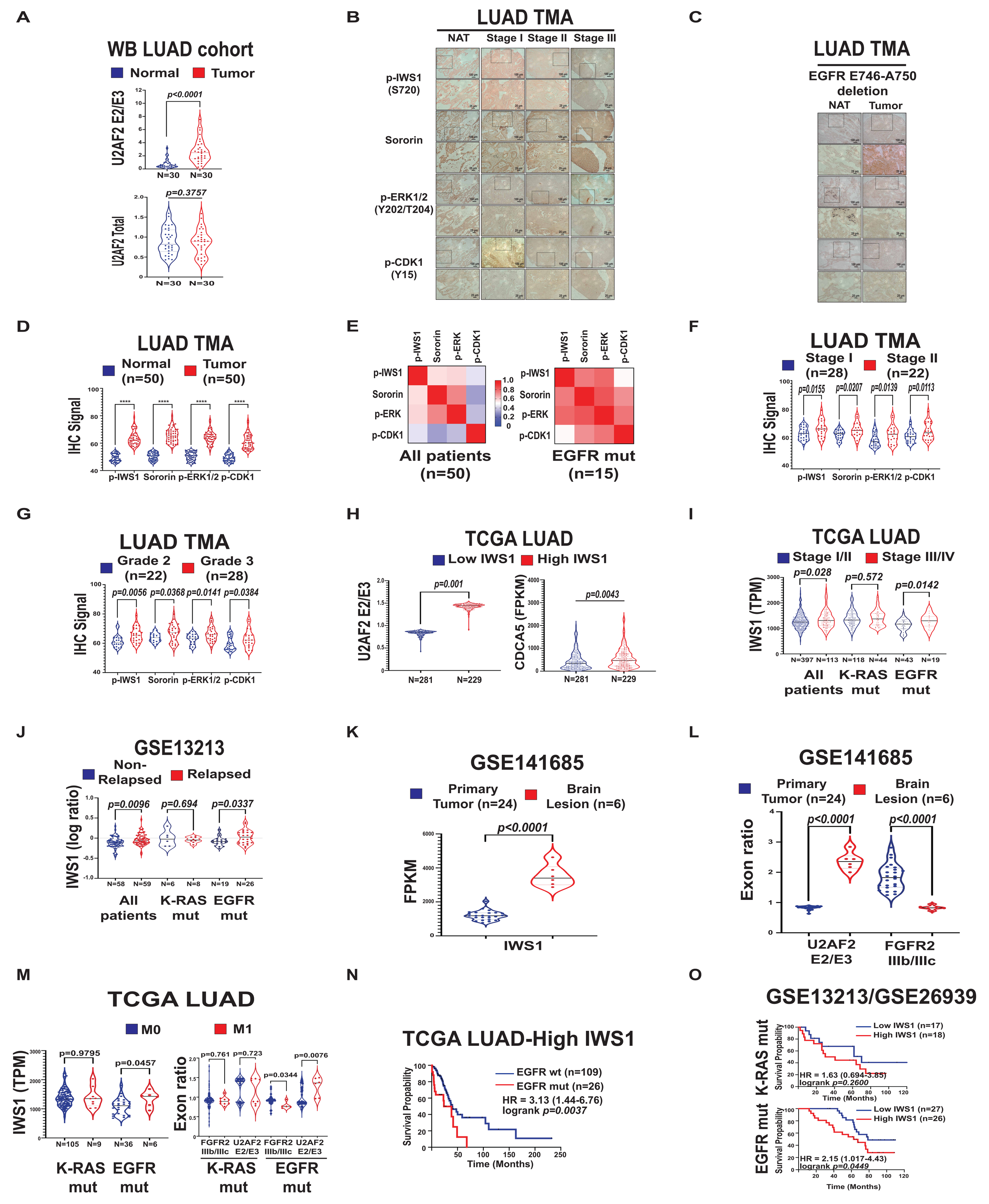

A
WB LUAD cohort
Normal
Tumor
U2AF2 E2/E3
p<0.0001
N=30
N=30
U2AF2 Total
p=0.3757
N=30
N=30
B
LUAD TMA
NAT
Stage I
Stage II
Stage III
p-IWS1 (S720)
Sororin
p-ERK1/2 (Y202/T204)
p-CDK1 (Y15)
C
LUAD TMA
EGFR E746-A750 deletion
NAT
Tumor
D
LUAD TMA
Normal (n=50)
Tumor (n=50)
IHC Signal
p-IWS1
Sororin
p-ERK1/2
p-CDK1
E
p-IWS1
Sororin
p-ERK
p-CDK1
All patients (n=50)
EGFR mut (n=15)
F
LUAD TMA
Stage I (n=28)
Stage II (n=22)
IHC Signal
p=0.0155
p=0.0207
p=0.0139
p=0.0113
p-IWS1
Sororin
p-ERK1/2
p-CDK1
G
LUAD TMA
Grade 2 (n=22)
Grade 3 (n=28)
IHC Signal
p=0.0056
p=0.0368
p=0.0141
p=0.0384
p-IWS1
Sororin
p-ERK1/2
p-CDK1
H
TCGA LUAD
Low IWS1
High IWS1
U2AF2 E2/E3
p=0.001
N=281
N=229
CDCA5 (FPKM)
p=0.0043
N=281
N=229
I
TCGA LUAD
Stage I/II
Stage III/IV
IWS1 (TPM)
p=0.028
p=0.572
p=0.0142
N=397
N=113
N=118
N=44
N=43
N=19
All patients
K-RAS mut
EGFR mut
J
GSE13213
Non-Relapsed
Relapsed
IWS1 (log ratio)
p=0.0096
p=0.694
p=0.0337
N=58
N=59
N=6
N=8
N=19
N=26
All patients
K-RAS mut
EGFR mut
K
GSE141685
Primary Tumor (n=24)
Brain Lesion (n=6)
FPKM
p<0.0001
IWS1
L
GSE141685
Primary Tumor (n=24)
Brain Lesion (n=6)
Exon ratio
p<0.0001
p<0.0001
U2AF2 E2/E3
FGFR2 IIIb/IIIc
M
TCGA LUAD
M0
M1
IWS1 (TPM)
p=0.9795
p=0.0457
N=105
N=9
N=36
N=6
K-RAS mut
EGFR mut
Exon ratio
p=0.761
p=0.723
p=0.0344
p=0.0076
FGFR2 IIIb/IIIc
U2AF2 E2/E3
FGFR2 IIIb/IIIc
U2AF2 E2/E3
K-RAS mut
EGFR mut
N
TCGA LUAD-High IWS1
Survival Propability
EGFR wt (n=109)
EGFR mut (n=26)
HR = 3.13 (1.44-6.76)
logrank p=0.0037
Time (Months)
O
GSE13213/GSE26939
K-RAS mut
Low IWS1 (n=17)
High IWS1 (n=18)
HR = 1.63 (0.694-3.85)
logrank p=0.2600
EGFR mut
Low IWS1 (n=27)
High IWS1 (n=26)
HR = 2.15 (1.017-4.43)
logrank p=0.0449
Survival Propability
Time (Months)

Figure S11 (relative to Figure 8). The effect of the Akt3/p-IWS1/U2AF2 axis in lung adenocarcinoma

A. *The overall abundance of the U2AF2 mRNA in lung adenocarcinomas and the adjacent normal tissue is similar, but the E2/E3 ratio is higher in the tumors.* Quantitative RT-PCR showing the *U2AF2 E2/E3* ratio (top) and the total *U2AF2* mRNA (bottom) in lysates derived from the 30 LUAD samples and the paired normal adjacent tissue in Fig 8A. The black lines in the violin plots show the mean *U2AF2 E2/E3* ratio relative to the matching normal samples, or *U2AF2* levels, relative to *GAPDH*. Statistical analyses were done using the paired t-test.

B. *IWS1 phosphorylation correlates with U2AF2 alternative RNA splicing and the Sororin/ERK axis in lung adenocarcinoma patients***.** Sequential sections of a commercially available tissue microarray (TMA) of 50 LUAD with paired NAT, were probed with the indicated antibodies. Figure shows the staining of a representative NAT sample and representative samples of Stage I, II and III lung adenocarcinomas (one of each). Boxes delineate the area of higher magnification, shown in the image below. Scale in the right lower corner of each image.

C. The TMA sections used before were stained with the specific monoclonal antibody against EGFR E746-A750 deletion. Figure shows three representative lung adenocarcinomas with EGFR deletion (E746-A750 deletion) and the corresponding NATs. Boxes delineate the area of higher magnification, shown in the image below. Scale in the right lower corner of each image.

D. *The overall activity of the p-IWS1/CDCA5 is higher in human lung adenocarcinomas than in the adjacent normal tissue.* Violin plots showing the abundance of IWS1 phosphorylation, along with the abundance of Sororin, phospho-ERK and phospho-CDK1, in the LUAD and NAT samples in the tissue microarray in B and C. The black line indicates

the mean p-IWS1, Sororin, p-ERK and p-CDK1 levels. Statistical analyses were done using the paired t-test.

E. *IWS1 phosphorylation exhibits a more robust correlation with Sororin, phospho-ERK and phospho-CDK1 in lung adenocarcinomas with EGFR mutations.* Heatmaps showing the correlation coefficient between the indicated components of the IWS1 phosphorylation pathway in the entire TMA cohort (left) and in the *EGFR* mutant cohort (right). The correlations were calculated using simple linear regression. The statistics and the p values for all the comparisons, can be found in Supplementary Table S5.

F. *The activity of the p-IWS1/CDCA5 pathway correlates with tumor stage in human lung adenocarcinomas.* Violin plots comparing the phosphorylation of IWS1 and the abundance of Sororin, phospho-ERK and phospho-CDK1 in stage I and Stage II/III lung adenocarcinomas in the group of 50 tumors sampled in the commercial TMA described above. Statistical analyses were performed using the one-sided unpaired t-test.

G. *The activity of the p-IWS1/CDCA5 pathway correlates with tumor grade in human lung adenocarcinomas.* Violin plots comparing the phosphorylation of IWS1 and the abundance of Sororin, phospho-ERK and phospho-CDK1 in histological grade 2 and grade 3 lung adenocarcinomas in the group of 50 tumors in the commercial TMA described above. Statistical analyses were performed using the one-sided unpaired t-test.

H. The expression of IWS1 correlates with Exon 2 inclusion in the *U2AF2* mRNA transcripts (upper panel) and with the expression of *CDCA5* (lower panel) in the TCGA LUAD dataset. The black line indicates the mean expression of *U2AF2* E2/E3 ratio and *CDCA5*. Statistical analyses were performed using the one-sided unpaired t-test.

I. Violin plots showing the abundance of IWS1 in stage I/II and in Stage III/IV tumors of patients in the TCGA LUAD database. Data shown for all tumors, and selectively for *EGFR* or *KRAS* mutant tumors. The horizontal lines indicate mean values for IWS1 levels. Statistical analyses were performed using the one-sided unpaired t-test.

J. Violin plots showing the abundance of IWS1 in relapsed and non-relapsed tumors of patients in the GSE13213 dataset. Data shown for all tumors, and selectively for *EGFR* or *KRAS* mutant tumors. The horizontal lines indicate mean values for IWS1. Statistical analyses were performed using the one-sided unpaired t-test.

K. (Left panel) Violin plots showing the expression of IWS1 in transcripts per million (TPM) in *KRAS* and *EGFR* mutant, primary and metastatic lung adenocarcinomas in the TCGA LUAD database. IWS1 expression is higher in tumors from patients with metastatic disease (M1) only if the tumors harbour *EGFR* mutations. (Right panel) Exon E2/E3 ratio in the *U2AF2* mRNA and exon IIIb/IIIc ratio in the *FGFR2* mRNA in *KRAS* and *EGFR* mutant lung adenocarcinomas. E2/E3 ratio in the *U2AF2* mRNA is higher and the E8/E9 ratio in the *FGFR2* mRNA is lower in tumors from patients with metastatic disease (M1), only if the tumors harbour *EGFR* mutations. The black lines indicate the mean IWS1 expression, and the mean *U2FA2* E2/E3 and *FGFR2* IIIb/IIIc ratios. Statistical analyses were performed using the one-sided unpaired t-test.

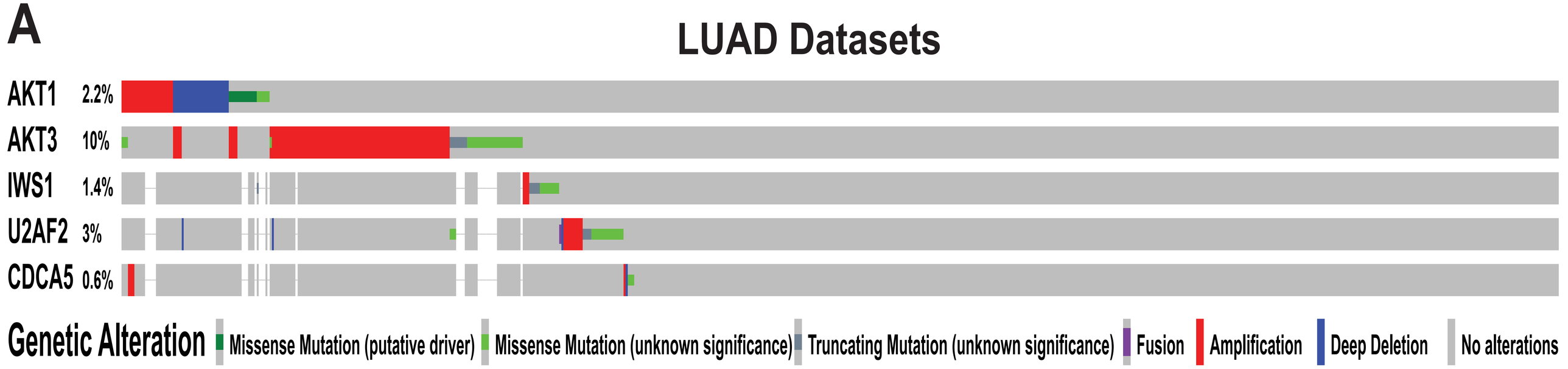
L. The expression of *IWS1* (upper panel) and the *E2/E3* ratio in mature *U2AF2* mRNA transcripts (lower panel-left) are higher in brain metastases of lung adenocarcinomas than in primary tumors, while the *IIIb/IIIc* ratio in mature *FGFR2* transcripts (lower panel-right) is lower in brain metastases relative to primary tumors. The black lines indicate the mean expression of IWS1 and the mean *U2AF2 E2/E3* and *FGFR2 IIIb/IIIc* ratio. Violin plots were based on the same data shown in Figure 8G. Statistical analyses were performed using the one-sided unpaired t-test.

M. The *E2/E3* ratio in mature *U2AF2* mRNA transcripts are higher in the same patient group as Fig S11L. The black lines indicate the mean *U2AF2 E2/E3* and *FGFR2 IIIb/IIIc* ratio. Violin plots were based on the same data shown in Figure 8G. Statistical analyses were performed using the one-sided unpaired t-test.

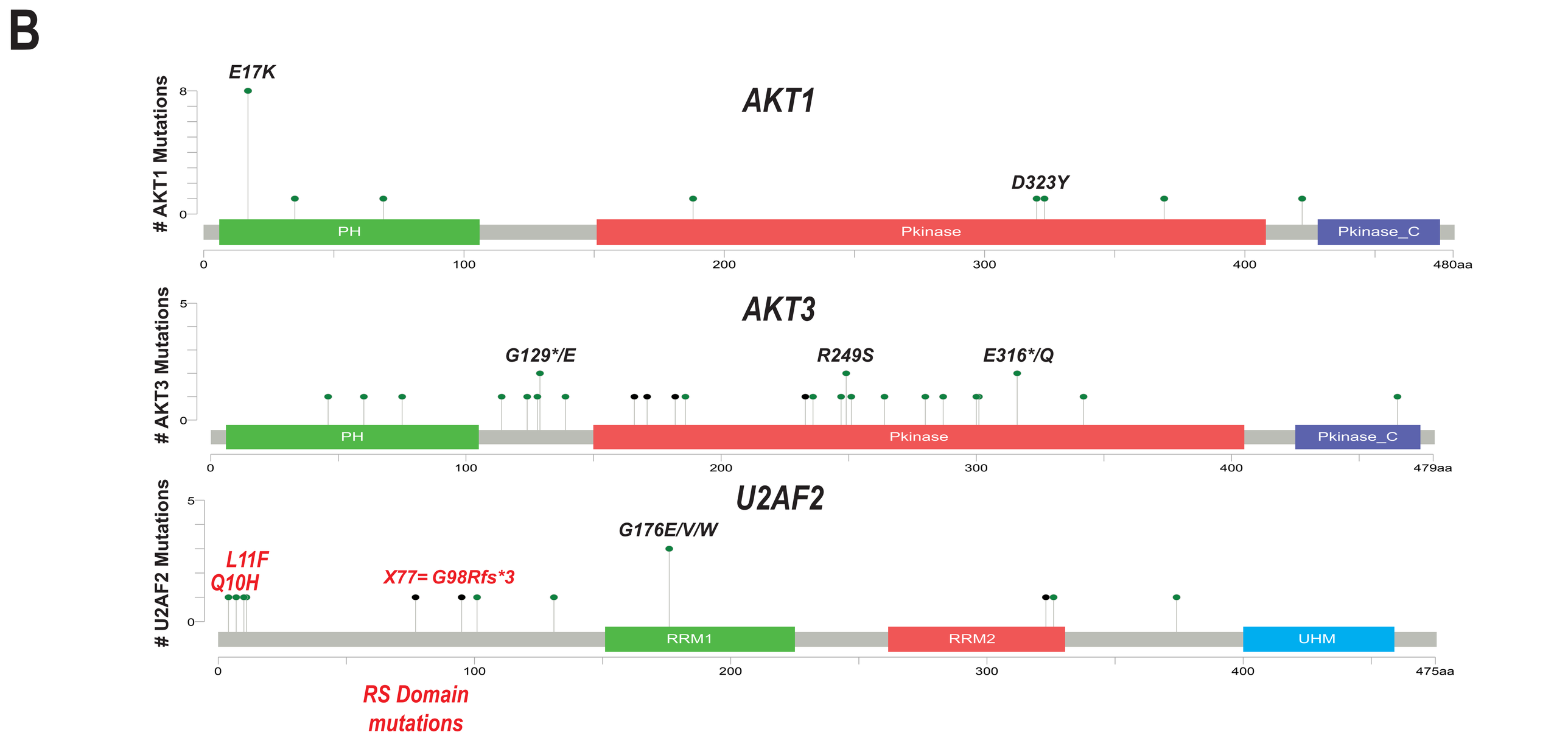
N. *EGFR mutations predict poor survival in patients expressing high levels of IWS1.* Kaplan-Meier curves showing the impact of *EGFR* mutations on survival, in patients with lung adenocarcinomas expressing high levels of IWS1. Statistical analyses were performed using the log rank test and Cox's proportional hazards model.

O. *The abundance of IWS1 has a stronger negative impact on the survival of patients with lung adenocarcinomas harbouring EGFR, than those with lung adenocarcinomas harbouring KRAS mutations.* Kaplan-Meier curves showing the impact of the abundance of IWS1 on the survival of patients with lung adenocarcinomas harbouring *KRAS* mutations (upper panel) or *EGFR* mutations (lower panel) in the GSE13213/GSE26969 dataset. Statistical analyses were performed using the log rank test and Cox's proportional hazards model.

A


LUAD Datasets
AKT1 2.2%
AKT3 10%
IWS1 1.4%
U2AF2 3%
CDCA5 0.6%
Genetic Alteration
Missense Mutation (putative driver)
Missense Mutation (unknown significance)
Truncating Mutation (unknown significance)
Fusion
Amplification
Deep Deletion
No alterations

B


AKT1
AKT1 Mutations
E17K
D323Y
PH
Pkinase
Pkinase_C
0
100
200
300
400
480aa
AKT3
AKT3 Mutations
G129*/E
R249S
E316*/Q
PH
Pkinase
Pkinase_C
479aa
U2AF2
U2AF2 Mutations
G176E/V/W
L11F
Q10H
X77= G98Rfs*3
RRM1
RRM2
UHM
475aa
RS Domain mutations

C

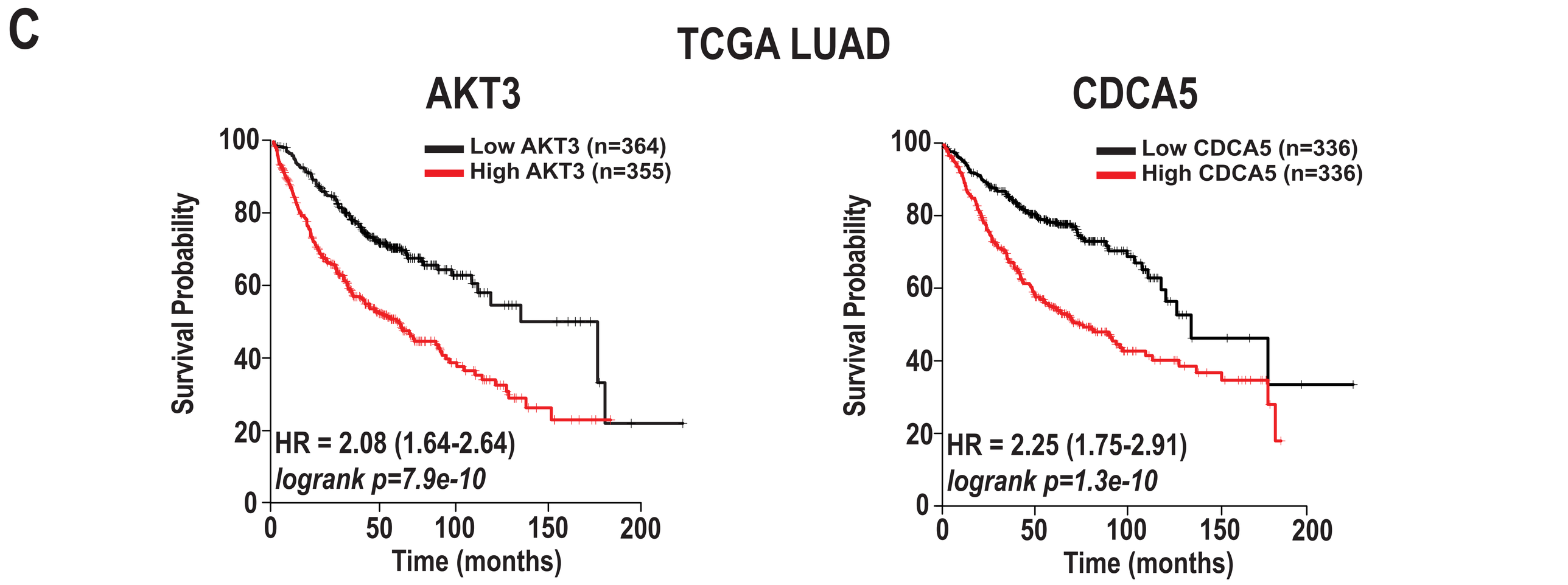

TCGA LUAD
AKT3
Low AKT3 (n=364)
High AKT3 (n=355)
Survival Probability
HR = 2.08 (1.64-2.64)
logrank p=7.9e-10
Time (months)
CDCA5
Low CDCA5 (n=336)
High CDCA5 (n=336)
Survival Probability
HR = 2.25 (1.75-2.91)
logrank p=1.3e-10
Time (months)

Figure S12 (relative to Figure 8). The effect of the Akt3/p-IWS1/U2AF2 axis in lung adenocarcinoma

A. *The AKT/IWS1/U2AF2 pathway is genetically altered in patients with Lung Adenocarcinoma.* Alterations of the AKT/IWS1 pathway were visualized by cBioPortal for Cancer Genomics, using all the available LUAD datasets (TSP- Ding et al., 2008[81], Broad-Imielinksi et al., 2012[82], MSKCC-Rivzi et al., 2015[83], Jordan et al., 2017[84], OncoSG-Chen et al., 2020[85], TCGA Firehose Legacy). Mutation, deletion, amplification, fusion and other alterations are shown in different colors.

B. The distribution of the mutations of the above patients in *AKT1*, *AKT3* and *U2AF2* genes. The most common mutations in these patients are shown in black. Regarding *U2AF2,* the missense and splice mutations of the RS domain can be found in red. The complete list of the mutations in this pathway can be found in Supplementary Table S7.

C. Kaplan-Meier Curves showing the impact of *AKT3* expression (left) and *CDCA5* (right) on patient survival in patients in the TCGA LUAD database. Statistical analyses were performed using the log rank test and Cox's proportional hazards model.