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

[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]Tufts Graduate School of Biomedical Sciences, Program in Genetics, Boston, MA, 02111, USA, [7]DIANA-Lab, Hellenic Pasteur Institute, Athens, 11521, Greece, [8]Department of Clinical Chemistry-Biochemistry, School of Medicine, University of Crete, Heraklion 71110, Crete, Greece, [9]Institute for Molecular Biology and Biotechnology, Foundation for Research and Technology Hellas, Heraklion 70013, Greece, [10]College of Medicine, Department of Hematology, The Ohio State University, Columbus, OH 43210, USA, [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.

[§] These authors contributed equally to this work.

Present Address :

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

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

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

*Corresponding author: Philip N. Tsichlis, Department of Cancer Biology and Genetics. The Ohio State University, and the Ohio State University Comprehensive Cancer Center. Columbus Ohio 43210. Email: Philip.tsichlis@osumc.edu

Abstract

Our previous studies have shown that IWS1 (Interacts with Spt6) is a phosphorylation target of AKT and regulates the alternative RNA splicing of FGFR-2, 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 (SR) 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 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.

Introduction

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

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

Following our original observations linking the alternative RNA splicing of FGFR-2 with NSCLC (Sanidas et al., 2014[10]), several additional splicing events have been linked to the biology of NSCLC including Bcl-X_L, CD44, Androgen Receptor (AR), HLA-G and PKM, by promoting cell survival, metastasis, chemoresistance, immune surveillance and metabolic advantage, respectively (Li Z. et al., 2016[11], Todaro M. et al., 2014[12], Oltean S. et al., 2014[13], Calabretta S. et al., 2016[14]). Furthermore, another mechanism utilized by lung cancer cells involves the re-programming of alternative splicing through the introduction of de novo mutations which create new splice sites and alternatively spliced isoforms (Vo N.S. et al., 2018[15], Giaj Levra M. et al.,

2014[16]). This knowledge has led to an ongoing phase I clinical trial to determine the effects of a modulator of the core splicing function, H3B-8800, in cancer patients (Seiler M., 2018[17]). Thus, deeper comprehension of the molecular regulation of RNA splicing in lung cancer, and other cancer types, is a necessity for proper design of anti-tumor therapies.

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

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

in RNA splicing, are regulated by periodic shifts in the expression and the activity of known cell cycle regulators, such as the Aurora kinases (Moore et al., 2010[28]), the large SR domain-containing protein SON (Ahn et al., 2011[29]), the RNA recognition motif (RRM)-containing protein TgRRM1 (Suvorova et al, 2013[30]) and CLK1 (Dominguez et al., 2016[27]), demonstrating the bidirectional relationship of these cellular functions.

We have previously shown that the transcription elongation factor IWS1 and the Akt3 kinase play a central role in the regulation of the alternative RNA splicing of FGFR-2, by promoting the exclusion of exon 8. This exclusion event depends on the phosphorylation of IWS1 by AKT3 at Ser720/Thr721. IWS1 binds SPT6, which is attached to the phosphorylated Ser2 of the C-terminal domain (CTD) of RNA polymerase II. SETD2, a histone H3 trimethyl-transferase, binds to the IWS1/Spt6 complex only when IWS1 is phosphorylated at Ser720/Thr721. SETD2 bound to the phosphorylated and CTD-associated IWS1, trimethylates histone H3 at K36 during transcription. The histone H3K36me3 marks in the body of actively-transcribed target FGFR-2 gene are recognized by MRG15, which interacts with PTB leading to exclusion of exon 8 in the mature FGFR2 mRNA transcript, affecting cell invasion, migration, cell proliferation and tumor growth in mouse xenografts models. More importantly, we had shown robust expression of IWS1 phosphorylation in 21 out of 24 NSCLC samples, which correlates with AKT phosphorylation, introducing the role of IWS1 in lung cancer (Sanidas et al., 2014[10]).

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

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

Results

IWS1 expression and phosphorylation regulate alternative mRNA splicing.

We have previously reported that IWS1 phosphorylation at Ser 720/Thr721, primarily by AKT3, resulted in the exclusion of exon 8 of FGFR-2 in the human NSCLC cell lines NCI-H522 and NCI-H1299 (Sanidas et al., 2014[10]). To explore the molecular mechanisms driving the IWS1 phosphorylation-dependent RNA splicing and gene expression, we performed RNA-Seq to examine the transcriptome of shControl, shIWS1, shIWS1/wild type IWS1 rescue (shIWS1/WT-R) and shIWS1/phosphorylation site IWS1 mutant rescue (shIWS1/MT-R) NCI-H522 cells. First, we confirmed the significant down-regulation of IWS1 expression after the transduction of the cells with lentiviral shIWS1 construct and the rescue of IWS1 knock-down with the Flag-tagged wild type and phosphorylation IWS1 site mutant (**Fig. 1A**). Differential expression analysis of the RNA-seq data, identified 1,621 and 562 differentially expressed genes between shControl and shIWS1 and between shIWS1/WT-R and shIWS1/MT-R cells, respectively ($p \leq 0.01$, FDR ≤ 0.2). 340 genes were detected as differentially expressed in both comparisons (**Fig. S1A, S1B**). Moreover, 19 out of the FDR-ranked top 100 differentially expressed genes, in shControl versus shIWS1 cells, were also differentially expressed in shIWS1/WT-R versus shIWS1/MT-R (**Fig.**

S1C). Gene Set Enrichment analysis (Subramanian et al., 2005[34]), revealed significant enrichment of genes involved in RNA metabolism and regulation of RNA processing (**Fig. S1D**).

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

Our earlier studies have shown that the IWS1 expression and phosphorylation promote exon exclusion in the *FGFR-2* gene (Sanidas et al., 2014[10]). Therefore, we analysed the exon usage data, and we observed that the most common event associated with the expression and phosphorylation of IWS1, was exon inclusion (**Fig. 1C).** Functional analysis of the IWS1-regulated alternative spliced genes, focused on GO-biological processes, identified significant enrichment of genes involved in RNA processing and RNA splicing, among the top biological processes regulated by IWS1-dependent alternative RNA splicing (**Fig. 1D, 1E**). It is remarkable that the results were similar with the results generated from the functional analysis of the IWS1-dependent differential gene expression. Therefore, these findings imply that the effect of IWS1 on RNA processing may be direct or indirect. The indirect effect may be due to the IWS1 expression and phosphorylation-dependent differential regulation of genes involved in RNA processing.

Validation of the RNA-seq, using RT-PCR, showed several examples of alternative RNA splicing of genes which are regulated by IWS1 and IWS1-phosphorylation, via exon inclusion (**Fig. 1, S1G-S1J**). One of these events is the inclusion of exon 2 in the mature mRNA transcript of *U2AF2*, the gene encoding the splicing factor U2AF65. Whereas the predominant *U2AF2*

mRNA transcript in shControl and shIWS1/WT-R cells contains exon 2, the predominant transcript in shIWS1 and shIWS1/MT-R cells, is a novel transcript lacking exon 2. To validate these results, we examined the ratio of exons 2 and 3 (E2/E3 U2AF2) by RT-PCR in NCI-H522 and NCI-H1299 shControl, shIWS1, shIWS1/WT-R and shIWS1/MT-R cells. The results revealed a decrease in the E2/E3 ratio in shIWS1 and shIWS1/MT-R cells, relative to shControl cells (**Fig. 1F, 1G**). The decrease in the E2/E3 ratio in shIWS1 and shIWS1/MT-R, relative to shControl cells, was confirmed by quantitative RT-PCR (**Fig. S1E**). Notably, the knockdown of IWS1 or the rescue with the non-phosphorylatable site mutant (IWS1/MT-R) did not significantly change the total expression of U2AF2 in NCI-H522 and NCI-H1299 cells. (**Fig. S1F**).

To establish whether IWS1 is directly involved in the molecular complex that regulates U2AF2 RNA splicing, in NCI-H522 and NCI-H1299 cells transduced with shControl, shIWS1, shIWS1/WT-R and shIWS1/MT-R, we performed chromatin immunoprecipitation (ChIP) assays. The results revealed that both IWS1 WT and S720A/T721A bind equally well to exon 2 and exon 3 of *U2AF2* (**Fig. 1H**). Therefore, IWS1 phosphorylation controls U2AF2 exon 2 alternative RNA splicing by regulating events occurring after the binding of IWS1 to chromatin.

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

IWS1 is phosphorylated by AKT3, and to a lesser extent by AKT1, at Ser720/Thr721 (Sanidas et al., 2014[10]). Since the phosphorylation and activity of the AKT isoforms can be induced by tyrosine-kinase signals, we questioned whether IGF-1 stimulation of serum starved NCI-H522 and NCI-H1299 cells, promotes *U2AF2* exon 2 inclusion along with the expected AKT activation and IWS1 phosphorylation. The results showed that the E2/E3 *U2AF2* ratio indeed parallels the activation of AKT and the subsequent phosphorylation of IWS1 (**Fig. 2A, S2A**). To address whether the *U2AF2* exon 2 inclusion depends on AKT activation, we treated NCI-H522 and NCI-H1299 cells growing in serum-containing complete media with 5 μM of the pan-AKT

inhibitor MK2206, a dose that fully inhibits all AKT isoforms (Sanidas et al., 2014[10]) .The results confirmed that MK2206 inhibits both AKT (T308 and S473) and IWS1 phosphorylation (S720) and *U2AF2* exon 2 inclusion (**Fig. 2B, S2B).** To determine whether it is the AKT3 isoform, which is responsible for the observed effects of AKT inhibition on the *U2AF2* alternative RNA splicing, we transfected NCI-H522 and NCI-H1299 cells with siAKT3, along with siControl, and we examined its effects on the alternative splicing of *U2AF2* exon 2, using RT-PCR and qRT-PCR. The results phenocopied the inhibition with MK2206 (**Fig. 2C, S2C)**, providing evidence that the physiological regulation of alternative RNA splicing of *U2AF2* by external stimuli depends on the phosphorylation of IWS1 by AKT3.

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

We have previously reported that IWS1 phosphorylation by AKT promotes the exclusion of exon 8 from the mature FGFR-2 mRNA transcript, via a process that depends on histone H3K36 trimethylation by SETD2, and that the latter is recruited to the CTD of RNA Pol II by phosphorylated IWS1 (Sanidas et al., 2014[10]). To determine whether the U2AF2 exon 2 inclusion phenotype is also dependent on histone H3K36 trimethylation, we performed ChIP assays in shControl, shIWS1, shIWS1/WT-R and shIWS1/MT-R NCI-H522 and NCI-H1299 cells, addressing the abundance of H3K36me3 marks on exons 2 and 3 of *U2AF2*, as well as on the *U2AF2* transcriptional start site (TSS) and the *GAPDH* gene, which were used as controls. The results revealed that the knockdown of IWS1 or the replacement with the phosphorylation-site mutant significantly decreased the abundance of H3K36me3 on *U2AF2* exon 2 and 3 (**Fig. 3A**). More important, inhibition of AKT, using MK2206, phenocopied the H3K36me3 pattern of phosphorylation-deficient IWS1 mutant (**Fig. 3B**). Since IWS1-phosphorylation recruits SETD2 in C-terminal domain of RNA-pol II, we examined the binding of SETD2 on *U2AF2* and *GAPDH* gene in the same NCI-H522 and NCI-H1299 cells, transduced using a SETD2 lentiviral construct,

with a Hemagglutinin (HA)-tag. The results showed a pattern of SETD2 binding, which parallels the IWS1 phosphorylation-dependent pattern of H3K36me3 abundance (**Fig. 3C**), suggesting that H3K36 trimethylation is mediated by SETD2.

The preceding data suggested that the enzymatically active SETD2 contributes to the IWS1 phosphorylation-dependent regulation of the *U2AF2* alternative RNA splicing. The role of catalytically-active SETD2 was confirmed by SETD2 knockdown and rescue experiments in NCI-H522 and NCI-H1299 cells. Using RT-PCR and qRT-PCR, we observed that the knockdown of SETD2 phenocopies the knockdown of IWS1 on the *U2AF2* RNA splicing, an effect rescued by the wild type SETD2, but not by the SETD2 methyltransferase mutant R1625C, which globally silences H3K36me3 (Hacker et. al., 2016[36]) (**Fig. 3D**). We conclude that the catalytically-active SETD2 is indeed required for the inclusion of exon 2 in the mature *U2AF2* mRNA.

Recently, and after the completion of the ChIP experiments, we carried out ChIP-Seq experiments, addressing the binding of IWS1 and SETD2 and the distribution of H3K36me3 marks genome-wide in NCI-H522 shIWS1/WT-R and shIWS1/MT-R cells. The unbiased data on the abundance of these markers in the U2AF2 gene were in general agreement with the ChIP data described above. Specifically, IWS1 was found to bind U2AF2 E2 independent of its phosphorylation, but SETD2 binding and H3K36me3 abundance on U2AF2 E2 increased only when IWS1 was phosphorylated (**Fig. 3E**).

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

(**Fig. S3A**). The failure of SETD2 to rescue the phenotype supports the model of SETD2 recruitment by phosphorylated IWS1.

However, histone methylation is a dynamic step-by-step process. (Greer et al., 2012[37]) and while SETD2 is the only known H3K36 trimethyl-transferase in mammalian cells (Hyun K. et al., 2017[38]), there are several Lysine methyltransferases, which catalyze mono- or di-methylation of histone H3 at K36 and may influence the SETD2 output. Transfection of the NCI-H522 and NCI-H1299 cells with siRNAs targeting a set of methyltransferases that are known to catalyze histone H3K36 mono and di-methylation (NSD1, NSD2 and NSD3), or only di-methylation (SMYD2 and ASHL1) (Lucio-Eterovic et al., 2010[339], Q. Qiao et al., 2010[40], Berdasco et al., 2009[41] , Rahman et al., 2011[42], Brown M.A et al., 2006[43], Gregory G.D. et al., 2007[44]), revealed no contribution to the regulation of the alternative RNA splicing of *U2AF2* (**Fig. S3A**).

Conversely, histone methylation is a reversible process (Bannister et al., 2002[45]). The histone H3K36me3 marks are erased by KDM4A and KDM4C, two members of the KDM4 JmjC domain-histone demethylase family (Berry, W.L et al., 2013[46], Cloos, P.A et al., 2007[47], Li, W et al., 2011[48]). Ectopic expression of KDM4A, KDM4B and KDM4C using lentiviral constructs in NCI-H522 and NCI-H1299 cells showed that none of these three altered the IWS1 phosphorylation-dependent alternative RNA splicing pattern of *U2AF2* (**Fig. S3C**). Overall, these data indicate that the reported demethylation of H3K36me3 by members of KDM4/JMJD2 histone demethylase family does not play a significant role in our model.

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

Based on our preceding studies, the regulation of the *FGFR-2* alternative RNA splicing by IWS1 phosphorylation depends on the reading of the histone H3K36me3 marks by MRG15 (Sanidas et al., 2014[10]). To determine whether MRG15 is also the reader of the IWS1-dependent

alternative RNA splicing of *U2AF2*, we knocked down MRG15 in both NCI-H522 and NCI-H1299 cells. Using RT-PCR and qRT-PCR to monitor the alternative splicing of *U2AF2* in these cells, revealed that it is independent of MRG15 (**Fig. 4A Left panels**). In agreement with this result, the knockdown of the splicing repressor and binding partner of MRG15, PTB, also had no effect on the RNA splicing of *U2AF2* (**Fig. 4A Right panels**), results not surprising since IWS1 phosphorylation mediates exon inclusion signals in our model.

To identify the factor that controls the IWS1 phosphorylation-dependent exon inclusion of *U2AF2* exon 2, we transfected NCI-H522 and NCI-H1299 cells with siRNAs of the known H3K36me3 readers PHF1 (Cai L. et al., 2013[49]), BRPF1 (Vezzoli et al., 2010[50]), MSH6 (Li F. et al., 2014[51]), GLYR-1 (Vermuelen et al., 2010[52]) and LEDGF (Pradeepa et al., 2012[20]), along with a control siRNA. After validating the significant downregulation of the expression by Western blotting, monitoring the effects of these transfections on E2/E3 U2AF2 ratio using RT-PCR and qRT-PCR, revealed that only the knockdown of LEDGF phenocopied the knockdown of IWS1 on the alternative RNA splicing of *U2AF2*, suggesting that LEDGF is the sole H3K36me3 reader responsible for the observed phenotype (**Fig. 4B, S4A**). To confirm this conclusion and to determine which isoform of LEDGF may be responsible for the phenotype, we used a lentiviral shRNA construct to knockdown LEDGF, targeting the 3' UTR, and we rescued the knockdown by transducing the cells with lentiviral constructs of the alternatively spliced LEDGF isoforms, p75 and p52, with a V5-tag (**Fig. 4C, upper panel**). Monitoring the effects of these transductions by RT-PCR and qRT-PCR, revealed that only the p52 isoform rescues the shLEDGF-induced U2AF2 alternative RNA splicing phenotype. More important, the A51P mutant of p52-LEDGF, which cannot bind histone H3K36me3 (Shun et al., 2008[53]), failed to rescue the splicing defect, suggesting that the U2AF2 exon inclusion depends on the reading of the H3K36me3 marks by p52/LEDGF (**Fig. 4C, S4B**).

The preceding data provide strong genetic evidence that p52/LEDGF regulates the U2AF2 alternative RNA splicing by reading the IWS1 phosphorylation-dependent histone H3K36me3 marks. To validate these observations, chromatin Immunoprecipitation assays for the V5-tagged p52/LEDGF isoform, confirmed that the p52/LEDGF binds *U2AF2* exons 2 and 3 only when IWS1 is phosphorylated, observations consistent with SETD2 binding and H3K36me3 abundance (**Fig. 4D, S4C**).

If the p52 isoform of LEDGF regulates the alternative RNA splicing of *U2AF2* by reading the histone H3K36me3 mediated by IWS1-phosphorylation, as suggested by the preceding data, one would expect that overexpression of p52/LEDGF would not rescue the *U2AF2* alternative RNA splicing phenotype in shIWS1 and shIWS1/MT-R cells. This was confirmed by experiments addressing the *U2AF2* RNA splicing in NCI-H522 and NCI-H1299 shControl, shIWS1, shIWS1/WT-R and shIWS1/MT-R cells along with a V5-tagged p52/LEDGF construct (**Fig. S4D**). Altogether, these data support the model based on which p52/LEDGF regulates the alternative RNA splicing of *U2AF2* by reading the histone H3K36 trimethylation on the body of *U2AF2* gene, induced by IWS1 phosphorylation at S720/T721, mediated by SETD2.

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

It has been reported that the p52/LEDGF isoform re-localizes in the nucleus, through signals mediated from its unique C-terminal domain, and interacts with the splicing factor SRSF1, affecting its distribution to alternatively spliced genes. (Pradeepa et al., 2012[20]). To investigate a potential role of SRSF1 in U2AF2 alternative RNA splicing, we knocked down SRSF1 and we showed that its loss phenocopies the knockdown of IWS1. The dependence of the U2AF2 exon 2 splicing on SRSF1, which was suggested by this result, was confirmed by rescue experiments with wild type SRSF1 (**Fig. 4E, S4E**). We therefore conclude that SRSF1 regulates the alternative

RNA splicing of the U2AF2 exon 2. More importantly, SRSF1 did not rescue the U2AF2 RNA splicing phenotype induced by shIWS1 or shIWS1/MT-R (**Fig. S4F**), suggesting that it does not function independently. Instead, we hypothesized that it provides the link between IWS1 phosphorylation-dependent chromatin modifications and p52/LEDGF binding to the RNA splicing machinery. To test this hypothesis, we performed ChIP experiments addressing the binding of SRSF1 to U2AF2 exons 2 and 3 in shControl, shIWS1, shIWS1/WT-R and shIWS1/MT-R NCI-H522 and NCI-H1299 cells, transduced with a V5-tagged SRSF1 lentiviral construct. The U2AF2 TSS and GAPDH exon 3 regions were again used as controls. The results confirmed that the binding of SRSF1 to exons 2 and 3 parallels the SETD2 and p52/LEDGF binding along with H3K36me3 abundance and occurs only when IWS1 is phosphorylated (**Fig. 4F, S4G**), providing support to the proposed hypothesis.

Based on the preceding data, we hypothesize that the binding of SRSF1 to the chromatin-associated p52/LEDGF should bring it into proximity with the nascent pre-mRNA, facilitating their interaction. Analysis of the U2AF2 mRNA sequence, using the web-based pipeline RBP-map (Paz et al., 2014[55]), identified four potential SRSF1 binding sites (2 in U2AF2 exon 2 and 2 in exon 3) (**Fig. S4I**), providing additional support to this hypothesis. To experimentally address the proposed model, we carried out RNA-IP (RIP) experiments in the same shControl, shIWS1, shIWS1/WT-R and shIWS1/MT-R NCI-H522 and NCI-H1299 cells, focusing on the binding of SRSF1 to the U2AF2 exon 2, intron 2 and exon 3. The results confirmed that SRSF1 binds primarily to exon 2, but only in the shControl and shIWS/WT-R cells, which parallels its chromatin binding to the H3K36me3-bound p52/LEDGF (**Fig. 4G, S4H**).

Based on the data presented in Figure 3 and Figure 4, we propose that IWS1 phosphorylation by AKT is required for the recruitment of SETD2 to CTD of RNA Pol II. The latter promotes the trimethylation of histone H3 at K36 on the body of *U2AF2* and perhaps other target

genes. This facilitates the recruitment of the histone H3K36me3-interacting protein p52/LEDGF and its binding partner SRSF1, resulting in exon inclusion of *U2AF2* exon 2. (**Fig. 4H**)

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

The predominant novel splice variant of the U2AF2 mRNA that we identified in shIWS1 and shIWS1/MT-R cells, lacks exon 2, which encodes part of the N-terminal SR domain of U2AF65 (**Fig 5A**). This domain is responsible for the interaction of U2AF65 with several factors that contribute to mRNA splicing, 3' cleavage and polyadenylation (Millevoi et al., 2006[57], Shen H, et al., 2004[58]). One of these factors is Prp19, a component of the seven-member ubiquitin ligase complex Prp19C (S.P Chan et al., 2003[59], S.P Chan et al., 2005[60], R. Hogg et al., 2010[61], Chanarat S. et al., 2013[62]).

Using co-immunoprecipitation experiments in HEK-293T cells engineered to express the exon-2 containing U2AF65α or the exon 2-excluded U2AF65β variants, with a V5-tag, we confirmed that only the U2AF65α splice variant interact with the endogenous Prp19 (**Fig. 5B**). More important, consistent with the previous results, immunoprecipitation of endogenous U2AF65 in NCI-H522 and NCI-H1299 shControl, shIWS1, shIWS1/WT-R and shIWS1-MT-R cells, revealed impaired U2AF65-Prp19 interaction in shIWS1 and shIWS1/MT-R cells, which express primarily U2AF65β, while maintaining interaction with its U2-heterodimer partner, U2AF35 (**Fig. 5B, S5A**). These data confirmed that the interaction of U2AF65 with Prp19 depends on the sequence encoded by *U2AF2* exon 2, whose inclusion is regulated by IWS1 phosphorylation.

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

Based on previous reports, U2AF65 binds RNA Pol II, leading to an U2AF65-dependent recruitment of Prp19 to the newly-synthetized pre-mRNA and promoting proper co-transcriptional

splicing activation (C.J David et. al., 2011[63]). Furthermore, it had been shown that U2AF65 and Prp19 are necessary for the pre-mRNA splicing and accumulation of Sororin, a component of the cohesin complex, which is encoded by *CDCA5* (Watrin et. al., 2014[64]). Given that aberrant splicing of *U2AF2* in shIWS1 and shIWS1/MT-R cells resulted in the loss of the interaction between U2AF65 and Prp19, we hypothesized that the pre-mRNA splicing of *CDCA5* in these cells will be impaired. We therefore employed qRT-PCR to examine the ratio of spliced and unspliced CDCA5 mRNA in the same cells. The results confirmed that the mRNA splicing of *CDCA5* was impaired in both the shIWS1 and shIWS1/MT-R cells, with no significant change in the spliced/unspliced ratios of the control GUSB mRNA (Watrin et. al., 2014[64]) (**Fig. 5D, S5B, S5C upper panels).** More important, the splicing defect was rescued by U2AF65α but not by U2AF65β. (**Fig. 5E, S5D upper panels**).

To determine how IWS1 phosphorylation regulates CDCA5 RNA splicing, we performed RNA Immunoprecipitation (RIP) experiments in shControl, shIWS1, shIWS1/WT-R and shIWS1/MT-R NCI-H522 and NCI-H1299 cells. The results confirmed that both spliced variants of U2AF65 bind equally well CDCA5 pre-mRNA, as well as the control GUSB pre-mRNA, as expected (**Fig. 5D, 5E, S5C, S5D middle panels**) (Shen H. et al., 2004[31]). However, Prp19 binding to the same pre-mRNA regions of CDCA5 was significantly impaired in shIWS1 and shIWS1/MT-R cells, which predominantly express the U2AF65β isoform (**Fig. 5D, S5C lower panels**). More important, the impaired Prp19 binding to CDCA5 pre-mRNA was rescued by U2AF65α, but not U2AF65β (**Fig. 5E, S5D lower panels**).

Given that only spliced mRNAs are transported out of the nucleus, we also examined the abundance of cytosolic CDCA5 mRNA in shControl, shIWS1, shIWS1/WT-R, shIWS1/MT-R and shIWS1 rescues with U2AF65α or U2AF65β. To this extent, we fractionated the nuclear and cytosolic compartments of these NCI-H522 and NCI-H1299 cells. Then, using western blotting, we confirmed the validity of the fractionation by the expected expression of Lamin A/C and GAPDH only in the nuclear and cytosolic protein compartment, respectively (**Fig. S5E**). The

results also confirmed that the mature CDCA5 mRNA in the cytoplasm was downregulated as expected, and that the RNA transport deficiency is rescued by U2AF65α, but not U2AF65β (**Fig. 5F**).

To determine whether the CDCA5 RNA splicing defect in shIWS1 and shIWS1/MT-R NCI-H522 and NCI-H1299 cells prevents the expression of its protein product, Sororin, we examined the its expression along with the expression of IWS1, p-IWS1, U2AF65 and Prp19, by western blotting. The results confirmed that the expression of Sororin parallels IWS1 phosphorylation as it was impaired in shIWS1 and shIWS1/MT-R NCI-H522 and NCI-H1299 cells. More important, the low expression of Sororin in shIWS1 was rescued by U2AF65α, but not by U2AF65β (**Fig. 5G, S5F upper panels**). Taken together, these data confirm that IWS1 phosphorylation via the alternative RNA splicing of U2AF2, is ultimately responsible for the regulation of the mRNA splicing of CDCA5.

CDCA5 and p-ERK, form a positive feedback loop, which is activated by IWS1 phosphorylation and promotes the expression of CDK1 and Cyclin B1.

It has been previously shown that the downregulation of Sororin leads to reduced ERK phosphorylation, in human Colorectal Cancer (CRC) and human Hepatocellular Carcinoma (HCC) (Wang J., et al., 2018[65], A. Zhen et. al., 2019[66]). Since the IWS1 phosphorylation-dependent *U2AF2* alternative RNA splicing regulates the abundance of Sororin, we hypothesized reduced ERK phosphorylation in NCI-H522 and NCI-H1299 shIWS1 and shIWS1/MT-R cells. This was confirmed by experiments showing that ERK phosphorylation at Y202/T204 was reduced in shIWS1 and shIWS1/MT rescue cells, while total ERK was not. More important, the reduction of p-ERK levels in shIWS1 cells was rescued by U2AF65α, but not by U2AF65β (**Fig. 5G, S5F lower panels**), confirming the p-IWS1/Sororin/p-ERK axis in lung adenocarcinoma.

Sororin is a known target of ERK, which phosphorylates the protein at S79 and S209 (Nguyen et. al., 2010[67]). Interestingly, whereas the wild type Sororin and the Sororin

phosphomimetic mutant S79E/T209E (Sororin DM-E) rescued the phosphorylation of ERK in shIWS1 cells, the S79A/S209A (Sororin DM-A) mutant did not (**Fig 5G and S5F Lower panels**). This suggests a Sororin-ERK phosphorylation positive feedback loop in lung adenocarcinoma, which is controlled by the AKT-dependent IWS1 phosphorylation, and the U2AF2 alternative RNA splicing.

Based on previous observations, down-regulation of Sororin leads to impaired expression of CDK1 and Cyclin B1 (Wang J., et al., 2018[65], A. Zhen et. al., 2019[66]). To this extent, we hypothesized that IWS1 phosphorylation regulates the expression of CDK1 and Cyclin B1, through the regulation of U2AF2 RNA splicing and the abundance of Sororin. Consistent with the preceding data, western blotting revealed reduced levels of CDK1 (Y15) phosphorylation, CDK1 and Cyclin B1 levels in shIWS1 and shIWS1/MT-R cells. More importantly, the downregulation of these molecules in shIWS1 cells was rescued by U2AF65α, but not by U2AF65β, wild type Sororin and Sororin DM-E but not by Sororin DM-A (**Fig. 5G, S5F lower panels**). These data indicate that the Sororin-ERK phosphorylation feedback loop, downstream of IWS1 phosphorylation, controls the expression of CDK1 and Cyclin B1, through a process depending on U2AF2 alternative RNA splicing.

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

EGFR and KRAS are frequently mutated in human lung adenocarcinoma and promote downstream pro-oncogenic signals through the phosphorylation of ERK (Guo et al., 2020[68], Wee et al., 2017[69]). We therefore addressed the role of IWS1 phosphorylation and U2AF2 alternative RNA splicing on Sororin and ERK phosphorylation in lung adenocarcinoma cells harboring K-RAS (A549 and NCI-H460) and EGFR (NCI-H1975, PC-9, NCI-H1650) mutations. The results showed that U2AF2 alternative RNA splicing pattern is independent of the mutational status. (**Fig. 5H, upper and lower left panels**) More important, the knockdown of IWS1 downregulated the

expression of Sororin and ERK phosphorylation (**Fig. 5H, upper panel**). These data were striking, because they showed that the effect of Sororin on ERK phosphorylation is dominant over the effect of K-RAS and EGFR mutations. Surprisingly, the reduction of ERK phosphorylation induced by the loss of IWS1 phosphorylation was more robust in EGFR as opposed to the KRAS mutated cells (**Fig. 5H, lower right panels)**. Altogether, this data suggests that IWS1 phosphorylation controls the U2AF2 alternative RNA splicing independent of the cellular mutational status and shows robust reduction of the Sororin/ERK phosphorylation axis in lung adenocarcinoma cell lines, harboring EGFR mutations.

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

Given that the Sororin/ERK phosphorylation positive feedback loop, activated by IWS1 phosphorylation, promotes the expression of CDK1 and Cyclin B1, we hypothesized that the stimulation of cell proliferation by IWS1, we had observed previously (Sanidas et al., 2014[10]), may depend on the activation of the U2AF2 RNA splicing-regulated CDCA5/ERK phosphorylation loop. To address this hypothesis, we examined the rate of proliferation of shControl, shIWS1, shIWS1/Sororin WT, shIWS1/Sororin DM-A and shIWS1/Sororin DM-E NCI-H522 and NCI-H1299 cells growing under standard culture conditions. The results showed that cell proliferation was inhibited by shIWS1. More importantly, the inhibition was rescued by wild type Sororin and Sororin DM-E but not by the Sororin DM-A. Significantly, the proliferative effect of the Sororin DM-E was more robust than the proliferative effect of wild type (wt) Sororin **(Fig. 6A).** Western blotting revealed reduced expression of the cell proliferation marker PCNA in NCI-H522 and NCI-H1299 shIWS1 cells, which was rescued by U2AF65α, but not U2AF65β, confirming that the regulation of cell proliferation depend on the novel U2AF2 RNA splicing pattern (**Fig. 6B**).

As we pointed out above, the loss of phosphorylated IWS1 results in more robust down-regulation of ERK phosphorylation in EGFR as opposed to K-RAS mutated lung adenocarcinoma

cells (**Fig. 5H**), suggesting that the EGFR mutated cells would be more sensitive to the knockdown of IWS1. This was confirmed by experiments showing reduced cell proliferation of A549 (KRAS mutant) and NCI-H1975 (EGFR mutant) shIWS1 cells. More importantly, in parallel with the data of the reduction of ERK phosphorylation, EGFR mutated cell line NCI-H1975, demonstrates the highest reduction in cell proliferation induced by the loss of IWS1 (**Fig. S6A**). Additional support was provided by experiments showing that the shIWS1-induced downregulation of PCNA was more robust in EGFR (NCI-H1975, NCI-H1650) as opposed to K-RAS mutated (A549, NCI-H460) cells. (**Fig. 6C, S6B**). Moreover, the PCNA downregulation in these cell lines was rescued by U2AF65α, but not the U2AF65β, suggesting that the U2AF2 RNA splicing affects cell proliferation in K-RAS and EGFR-mutated cells (**Fig. 6D, S6B).** Altogether, these data suggest that IWS1 phosphorylation controls cell proliferation through the regulation Sororin/ERK phosphorylation mediated by U2AF2 RNA splicing.

To determine whether IWS1 phosphorylation stimulates cell proliferation by promoting progression through the G2/M phase of the cell cycle, we stained log phase cultures of shControl, shIWS1 NCI-H522, NCI-H1299, A549 and NCI-H1975 cells with Propidium Iodide (PI) and we examined the percentage of cells in different phases of the cell cycle by flow-cytometry. The results confirmed that the knockdown of IWS1 induced prolonged G2/M phase of the cell cycle, an effect less pronounced in KRAS mutant cell line, A549 **(Fig. 6E).** Altogether, data presented in figures 5 and 6 suggest that the EGFR mutated lung adenocarcinoma cell lines are more sensitive than K-RAS mutant cells upon the loss of IWS1 phosphorylation.

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

The preceding findings suggest that an RNA splicing event, regulated by the AKT-mediated phosphorylation of IWS1, plays a critical role in cell cycle progression. Given that the expression and activity of molecules critically involved in the regulation of the cell cycle, tend to

fluctuate as cells transit from one cell cycle phase to the next (Kent et al., 2019[70]), we examined the expression and phosphorylation of IWS1, the U2AF2 alternative RNA splicing pattern and the expression of SETD2, Sororin and the trimethylation of histone H3K36 in G1, S and G2/M in NCI-H1299 cells. To this end, using carboxyfluorescein succinimidyl ester (CFSE)-like DNA dyes, cells were stained in G1, S and G2/M phase and they were fractionated by FACS sorting (Begum et al, 2013[71]). Western blotting of these fractions confirmed that IWS1, phospho-IWS1, SETD2, histone H3K36me3 and Sororin are indeed upregulated in S and G2/M. As specific cell cycle phase markers, Cyclin E1 (G1 phase), CDC25A (S phase) and phosphorylated Histone H3 (S10) were used, validating the integrity of the fractionation **(Fig. 6F lower panels)**. More important, IWS1 expression and phosphorylation occured more during S phase, whereas Sororin, SETD2 and Histone H3K36me3 levels were more robust during G2/M phase, as previously reported (Watrin et al., 2014[64], Dronamraju et al., 2018[72]) **(Fig. 6F upper panels).** Furthermore, using RNA extracts from the same cells, RT-PCR experiments revealed an U2AF2 RNA exon inclusion pattern which parallels the expression of IWS1 during the cell cycle **(Fig. 6F).** Moreover, qRT-PCR confirmed that the expression of IWS1 and CDCA5 in the RNA level parallels their protein expression, with no change of total U2AF2 levels (**Fig. 6G**).

Chromatin Immuno-Cleavage (ChIC) experiments revealed increased binding of IWS1 on U2AF2 exon 2 and 3 during S and G2/M phase, with significantly more binding during S phase (**Fig. 6H upper panel**). Consistently, the binding of SETD2, the abundance of H3K36me3 and the E2/E3 U2AF2 ratio parallel the levels and the binding of IWS1 on U2AF2 gene, although total SETD2 and H3K36me3 are more abundant during G2/M phase (**Fig. 6H, 6I**).

Overall, these data provide strong evidence for the model in which IWS1, whose expression increases during S and G2/M, recruits SETD2 to CTD of RNA-Pol II, following phosphorylation by AKT. RNA-Pol II-associated SETD2 trimethylates histone H3 at K36 during transcriptional elongation, resulting in the assembly of p52/LEDGF-SRSF1 complexes which promote the inclusion of exon 2 in the U2AF2 mRNA. Subsequently, the U2AF65α isoform along

with Prp19, facilitate the proper splicing of CDCA5 pre-mRNA, leading to Sororin accumulation during S and G2/M phase. Finally, Sororin forms a positive feedback loop with ERK phosphorylation, and controls the progression through the cell cycle in lung adenocarcinoma. The events described in this model define a molecular mechanism of a novel splice variant which is epigenetically regulated during the cell cycle, feeding back into its regulation (**Fig. 6J, S6C**).

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

Based on our preceding studies, the loss of IWS1 and IWS1 phosphorylation impair tumor growth in a mouse xenograft model (Sanidas et al., 2014[10]). To further validate and expand this observation, we knocked down IWS1 in NCI-H1299 and A549 cells and we injected them subcutaneously, along with shControl cells, in immunocompromised NSG mice. The animals were monitored for tumor development and were sacrificed 4 and 6 weeks later, respectively. The results revealed that the knockdown of IWS1 reduced tumor weight and volume (**Fig. 7A**). They also showed that the reduction of tumor growth induced by the loss of IWS1 was less pronounced in A549 cells, harboring KRAS mutations (**Fig. 7B**). The weak growth reduction of tumours derived from shIWS1 A549 cells, paralleled the weak inhibition of ERK phosphorylation (**Fig 5H**) and cell proliferation (**Fig S6A**) induced by the knockdown of IWS1 in these cells.

The preceding data confirmed that the knockdown of IWS1 reduces tumour growth. Following confirmation, we proceeded to address the mechanism of the growth inhibition, which had not been adequately addressed in our earlier studies. First, we confirmed the efficiency of the IWS1 knockdown, by probing western blots of tumour cell lysates with anti-IWS1 and anti-phospho-IWS1 (S720) antibodies (**Fig. 7C**). Following this, we employed RT-PCR and qRT-PCR on RNA derived from these tumours, to address the usage of exon 2 in the *U2AF2* mRNA. The results confirmed that the knockdown of IWS1 has no effect on the total *U2AF2* mRNA levels (**Fig. S7A**), but promotes the exclusion of exon 2 from *U2AF2* mRNA (**Fig. 7C, upper panels,**

7D). Probing both tumour lysates and tissue sections with antibodies to regulators and targets of the Sororin/ERK feedback loop, confirmed that the activity of the loop is reduced in tumours derived from shIWS1 cells (**Fig. 7C, lower panels, S7B**). Measuring the abundance of the PCNA and Ki-67 proliferation markers by western blotting (PCNA), or Immunohistochemistry (Ki-67), also confirmed the reduced expression of these markers in the shIWS1 tumours (**Fig. 7C, lower panels, 7E**).

Quantitative analyses of the western blot (p-ERK and PCNA) and IHC data (Ki-67) (**Fig. 7F**), showed that the reduction of all these markers was less pronounced in the K-RAS mutant A549 tumors suggesting that KRAS mutant cells are more resistant to the loss of IWS1. Altogether, these results are in agreement with the less pronounced effect of shIWS1 in the tumor growth in A549-derived tumors (**Fig. 7B**) and with the less pronounced effect of shIWS1 in p-ERK levels (**Fig. 5H**), cell proliferation in culture and degree of G2/M arrest in these cells (**Fig. 6**).

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

To determine whether the pathway activated by IWS1 phosphorylation leading to Sororin expression and ERK phosphorylation is active in human lung adenocarcinoma (LUAD), we examined the expression and phosphorylation of several components of this axis, along with the alternative RNA splicing pattern of U2AF2 in a set of 40 human LUAD samples. For 30 of these tumors, Normal Adjacent Tissue (NAT) was also available and it was tested on parallel with the matching tumor sample. (**Fig. 8A**). More importantly, the exon 2-containing U2AF2 transcript was more abundant in tumor versus normal samples, with no change in total levels of U2AF2. (**Fig. S8A**). Overall, these results confirmed that the pathway was active in the tumors, but not in the NAT.

Human LUAD frequently harbors K-RAS or EGFR mutations and data presented in this report suggest that lung adenocarcinoma cells harboring EGFR mutations are more sensitive to

the loss of IWS1, as opposed to KRAS mutant cells (**Fig. 5 and Fig. 6**). We therefore used monoclonal antibodies, which selectively recognize K-RAS and EGFR mutant forms (Kim et al., 2015[74]), and we identified tumors harboring these mutations (**Fig. 8A**). Comparison of the correlations coefficients of IWS1 and phosphorylated IWS1 with the E2/E3 U2AF2 and all the components of the downstream Sororin/ERK phosphorylation pathway, revealed strong correlations in the entire cohort with more robust correlations in the EGFR than the KRAS mutant tumors (**Fig. 8B**). Strikingly, the levels of IWS1 phosphorylation and abundance of exon 2-containing U2AF2 isoform exhibit strong correlation with more aggressive clinical stage and defines poor survival in the entire patients' and the EGFR mutant cohort, but not in the KRAS mutated group (**Fig. 8C, 8D**).

The preceding data were confirmed by IHC, using sections of a commercially available tissue microarray (TMA) of 50 LUAD with paired NAT. The TMA samples were probed with antibodies against p-IWS1, Sororin, p-ERK, p-CDK1 and EGFR E746-A750 deletion (**Fig. S8B, S8C**). The results confirmed that the pathway is more active in the tumors, compared to the NAT samples **(Figure S8D).** More importantly, IWS1 phosphorylation expression exhibits strong correlations with the components of the Sororin/ERK axis, with more robust correlations in the EGFR mutant tumors (**Fig. S8E**). Consistently, the results also showed that the activity of the pathway correlates with clinical stage and histological grade **(Fig. S8F, S8G).** In addition to confirming the western blot data in our set of LUAD patients, the IHC data also demonstrate that the activity of the pathway can be monitored clinically in human tumors by IHC.

The data generated from the analysis of our LUAD samples and the TMA were confirmed by data in publicly available databases. Analysis of the data derived from Tumor Cancer Genome Atlas (TCGA), revealed correlations between IWS1 or SRSF1, U2AF2 E2 and other components of the IWS1 phosphorylation pathway (**Fig. 8E**). The results also showed that the U2AF2 E2/E3 ratio and the expression of CDCA5 mRNA were significantly upregulated in TCGA LUAD patients with high IWS1, compared to the low IWS1 group (**Fig. S8H**). More importantly, IWS1 expression

correlates with more aggressive tumor stage in tumors harboring EGFR but not KRAS mutations (**Fig. S8I**).

Based on recent reports, 55% of lung adenocarcinoma patients relapse and up to 40% of them develop metastatic disease. (Popper et al., 2016[75], West et al., 2020[76]). Given that the IWS1 phosphorylation controls oncogenic signals and plays a critical role in patients with lung adenocarcinoma, as suggested in this report, we questioned whether the IWS1-mediated axis is associated to clinical relapse or metastasis in LUAD patients. Analysis of a relapse-related molecular signature database GSE13213 (Tomida et al., 2009[77]), revealed increased expression of the IWS1/Sororin axis in relapsing patients harboring EGFR, but not with KRAS mutations (**Fig. 8F, S8J**). We therefore conclude that the activation of the IWS1 phosphorylation pathway may also contribute to the relapse of lung adenocarcinoma, especially with EGFR mutations. Furthermore, analysis of RNA-seq data derived from dissected brain metastatic lesions from lung adenocarcinoma patients (GSE141685), revealed increased expression of the components of IWS1/Sororin axis in the metastatic as opposed to primary lung adenocarcinoma lesions. Importantly, the *U2AF2* mRNA E2/E3 ratio was also elevated in the metastatic tumours, while the E8/E9 ratio of the *FGFR-2* mRNA was reduced (**Fig. S8K and S8L**), in agreement with our earlier observations, showing that IWS1 phosphorylation promotes *FGFR-2* exon 8 skipping (Sanidas et al, 2014[10]). Analysis of the RNA-Seq data in the TCGA LUAD dataset, which contains information on cancer-associated mutations, confirmed the link between IWS1 expression and metastatic disease, but also showed that IWS1 is again upregulated in tumours derived from patients with metastatic disease, and harbouring EGFR, but not K-RAS mutations (**Fig. S8M**).

As expected from the preceding data, IWS1 expression, *U2AF2* exon 2 inclusion and *FGFR-2* exon 8 exclusion, are indicators of poor prognosis in patients with lung adenocarcinomas in the TCGA LUAD dataset, harbouring EGFR, but not K-RAS mutations (**Fig. 8H**). In addition, the survival of patients with lung adenocarcinomas in the same dataset, expressing high levels of IWS1, was reduced if the tumours also harboured a mutated EGFR (**Fig. S8N**). Confirmatory to

these observations, was the finding that the synergy between IWS1 expression and EGFR mutations, was also observed in the GSE13213/GSE26939 dataset (**Fig. S8O**).

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

Discussion

Regulation of the Sororin/ERK phosphorylation loop from IWS1 phosphorylation

Our studies implicate IWS1 phosphorylation as a regulator of oncogenic signals in lung adenocarcinoma through alternative RNA splicing. We report a novel U2AF2 splice variant in lung adenocarcinoma which lacks exon 2 upon loss of phosphorylated IWS1. This signaling pathway is initiated by the AKT3-dependent phosphorylation of IWS1, which induces the epigenetic regulation of U2AF2 alternative RNA splicing, through H3K36me3-mediated signals (**Fig. 3**). The shift in the alternative RNA pattern controls the interaction of U2AF65 with Prp19. These two factors regulate the pre-mRNA splicing of CDCA5 and the expression of its protein product, Sororin, co-transcriptionally (**Fig. 5**). Our data provide insight for the existence of a novel Sororin/ERK phosphorylation feedback loop in lung adenocarcinoma, which is regulated by the alternative RNA splicing of U2AF2. Notably, the Sororin-dependent regulation of ERK phosphorylation by IWS1 is dominant over the ERK-activating EGFR mutations and has a major impact in the biology of these cells harboring such mutations **(Fig. 5 and Fig. 6)**. Sororin is a member of the cohesin complex, a seven member protein complex which plays a central role in sister chromatid cohesion and higher chromosomal architecture during meiosis (Ishiguro et al., 2019[78]). This molecular Sororin/ERK phosphorylation switch controls cell proliferation and the progression through the G2/M phase of the cell cycle (**Fig. 6**). Defects in Sororin expression lead

to the activation of spindle assembly checkpoint, delaying the progression through the G2/M phase (Dreier et al., 2011[79]). Given that loss of IWS1 reduced the expression of Sororin, this could explain the G2/M phase delay observed in shIWS1 cells (**Fig. 6**) How the ERK-phosphorylated Sororin promotes the phosphorylation and activation of ERK is currently unknown. Our working hypothesis is that the phosphorylation and activation of ERK is due to signals induced by the interaction of Sororin with its partners in the cohesin complex. If this is the case, the cell may use this mechanism to sense the successful progression from prometaphase to metaphase, in order to activate a molecular switch. This enhances the phosphorylation of Sororin, facilitating progression through the G2/M phase of the cell cycle. In addition, data presented in this report show that the loss of IWS1 leads to impaired expression of CDK1 and Cyclin B1, a phenotype regulated through the Sororin/ERK phosphorylation loop (**Fig. 5**). On the other hand, our data indicate strong correlation of CDCA5, CDK1 and Cyclin B1 mRNA levels in LUAD patients, derived from TCGA database (**Fig. 8**) These observations indicate that the induction of cyclin B1 and CDK1 by Sororin is most likely at the level of transcription. A potential mechanism for this transcriptional regulation, controlled by IWS1, was suggested by earlier studies showing that the Cohesin complex interacts with the Mediator complex. The Mediator-Cohesin complexes are loaded by the NIBPL Cohesin loading factor to enhancers and core promoters of target genes. Enhancer and core promoter-associated complexes promote loop formation between these segments, and regulate transcription (Kagey et al., 2010[80]). The contribution of this and other mechanisms on the regulation of CCNB1 and CDK1 expression is under investigation.

Regulation of RNA processing in multiple layers

Another important conclusion, based on the data presented in this report, is that RNA splicing is a process regulated at multiple levels. IWS1 phosphorylation directly regulates the alternative RNA splicing of U2AF2, via a H3K36me3-mediated mechanism (**Fig. 3**) and thus, introducing a new layer of RNA splicing regulation. This shift in alternative RNA splicing pattern

affects the binding and formation of the U2AF65/Prp19 interacting complex, required for the efficient splicing of CDCA5, which ultimately controls pro-oncogenic ERK phosphorylation signals (**Fig. 5, Fig. 6**). The reason for the multilayered control of RNA splicing by a single RNA splicing regulator could be that this allows a limited number of available pathways to converge, in different combinations, on a large number of RNA splicing events and differentially regulate them. To this extent, our data provide insight on the novel importance of the AKT/IWS1 phosphorylation axis on the fine-tuning of the global epigenetic regulation of RNA splicing cellular program, governing cell cycle regulation and tumor progression.

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

Given that the cell cycle is an integrated system and that cell cycle regulatory mechanisms tend to also be cell cycle-regulated, we examined and confirmed that IWS1 expression and phosphorylation fluctuate through the cell cycle, regulating U2AF2 RNA splicing. The mechanism of the expression of IWS1 is under investigation. The increased activity of the pathway may also occur due to cell cycle-dependent changes in the activity of AKT. Earlier reports have shown that CDK2-Cyclin A2 complex is activated in S phase and phosphorylates AKT at Ser477/Thr479 (Liu et al., 2014[73]), which further enhances the activity of IWS1 pathway. Furthermore, our results show that the epigenetic complexes assembled upon the phosphorylation of IWS1 on U2AF2 gene, are regulated in a cell-cycle specific manner. These results come in agreement with previous reports stating that spliceosomal assembly is also regulated through the cell-cycle, implying an extra layer of alternative RNA splicing regulation. (Karamysheva et al., 2015[81], Hofmann et al., 2010[82]). More importantly, our results reveal that IWS1 phosphorylation epigenetically controls the dynamics of a novel U2AF2 splice variant through the cell cycle and regulates the levels of Sororin, which subsequently feeds back to the regulation of the cell cycle and tumor proliferation, through enhancement of ERK oncogenic signals (**Fig. 6**).

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

Data presented in this report showed that signals originate from AKT phosphorylation may regulate alternative RNA splicing of target genes by different mechanisms, through IWS1 phosphorylation. Our previous findings had shown that IWS1 phosphorylation regulates the alternative RNA splicing of FGFR-2 by promoting the exclusion of exon 8 from the mature transcript. (Sanidas et al., 2014[10]) Here we show that, similar to the FGFR-2, the exon inclusion of U2AF2 exon 2, is also under the control of the SETD2-dependent histone H3K36 trimethylation in the body of the actively transcribed genes. However, the reader of the histone H3K36me3 mark is the p52 isoform of LEDGF which interacts with the RNA-binding protein SRSF1, as opposed to the MRG15/PTB complex. Therefore, although the transduction signal and the regulatory chromatin modification mark is the same for both the exon inclusion and exclusion signals, the effector complexes assembled around H3K36me3, that are responsible for the effects on RNA splicing differ. Interestingly, both the IWS1 phosphorylation splicing targets affect lung adenocarcinoma development. On the one hand, the FGFR-2 exon exclusion pattern promotes invasion and EMT (Sanidas et al., 2014[10]) and to the other hand, U2AF2 exon inclusion pattern promotes oncogenic and survival signals, converging towards more aggressive phenotype and clinical profile in lung adenocarcinoma patients, especially with EGFR mutations. This isoform specific aggressive clinical phenotype, comes in agreement with recent reports in the literature describing the existence of a cancer-specific splicing addiction network, in which cancer-related isoforms, in our case U2AF2 and FGFR-2, manipulate and maintain oncogenic and invasion signals (Bonnal et al., 2020[83], Wang and Aifantis, 2020[84]).

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

Throughout this report, we provide evidence that IWS1 phosphorylation, through the epigenetic regulation of U2AF2 alternative RNA splicing, represents an important oncogenic signal in EGFR mutant lung adenocarcinoma. Based on consistent results in cell lines (**Fig. 6**), animals (**Fig. 7**) and human tumor samples (**Fig. 8**), we showed the major impact of IWS1 phosphorylation on proliferation, tumor growth and, eventually, clinical outcome in EGFR mutant lung adenocarcinoma patients. Since, the PI3K/AKT axis is triggered by the EGFR mutated landscape, and as our data suggest, the activity of the IWS1 phosphorylation/U2AF2/Sororin axis would be increased (**Fig. 8**), further inducing the activation and maintenance of the tumorigenic ERK phosphorylation signals. To this extent, signals derived from the AKT/IWS1 phosphorylation pathway facilitates and maintains the oncogenic addiction of these tumors which depends on ERK phosphorylation. Furthermore, our data indicate that IWS1 phosphorylation and the RNA splicing pattern of its targets, U2AF2 and FGFR-2, selectively correlates not only with tumor grade, stage and patient survival, but also with metastasis and with relapse (**Fig. 8**). Based on these findings, we propose two translational applications for the IWS1 phosphorylation pathway described in this report. First, EGFR mutant lung adenocarcinoma patients may benefit from a dual inhibition of EGFR and specific AKT1/3 inhibitors (not currently investigated) or dual AKT/MEK inhibition (MK2206-AZD6244 : NCT01306045), limiting the ERK oncogenic addiction. Secondly, the IWS1 phosphorylation-dependent cancer splicing addiction can be manipulated by synergistic use of EGFR Tyrosine Kinase Inhibitors, inhibitors of the AKT/IWS1 axis and the use of highly isoform-specific antisense oligonucleotides and pharmacologic modulators of splicing machinery (Obeng et al., 2019[85]), which are currently under clinical trials (NCT03901469, NCT02711956, NCT02268552, NCT02908685).

Collectively, our results suggest that IWS1 phosphorylation by AKT acts as an epigenetic switch that, through H3K36me3-mediated signals, regulates RNA splicing in lung

adenocarcinoma maintaining the oncogenic and splicing addiction, and may serve as a novel precision-medicine marker and important drug target in EGFR mutant lung adenocarcinoma.

Materials and methods

Cells and culture conditions, growth factors and inhibitors

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

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

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

RNA-seq analysis

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

Immunoblotting and Protein Immunoprecipitation

Cells were lysed with RIPA buffer and cell lysates were resolved by electrophoresis in SDS-PAGE and analyzed by immunoblotting. Protein immunoprecipitation was carried out, following established protocols (Mohammed, H. et al, 2016[86]), with modifications. Images were acquired and analyzed, using the Li-Cor Fc Odyssey Imaging System (LI-COR Biosciences, Lincoln, NE). For the lists of antibodies used for immunoprecipitation and western blotting and for other details, see Supplemental Experimental Procedures.

Cloning and site-directed mutagenesis

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

Subcellular Fractionation

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

RT-PCR and qRT-PCR

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

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

The distribution of chromatin-bound IWS1 and SETD2 and the distribution of histone H3K36me3 genome-wide were addressed by ChIP-Seq experiments. The binding of IWS1 and SETD2 and the abundance of histone H3K36me3 on the *U2AF2* locus, were addressed by chromatin immunoprecipitation or chromatin Immuno-cleavage (Skene et al., 2018[87]). The latter was employed for experiments with FACS-sorted cells in different phases of the cell cycle. For details, regarding the sorting, the ChIP/ChIC protocols, ChIP-seq library preparation and analysis, see Supplemental Experimental Procedures.

RNA Immunoprecipitation

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

Tumor xenografts.

NOD.*Cg-Prkdc*^{scid} *Il2rg*^{tm1Wjl}/SzJ (NSG) mice were injected subcutaneously with $2x10^6$ or $5x10^6$ shControl or shIWS1 NCI-H1299 and A549 cells, and they were monitored for 4 and 6 weeks, respectively. Tumor metrics, protein and RNA extraction and tumor immunohistochemistry are

described in detail in the Supplemental Experimental Procedures. All mouse experiments were reviewed and approved by the Institutional Animal Care and Use Committee (IACUC) of the Ohio State University.

Human Tumor Samples.

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

Immunohistochemistry

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

TCGA/dataset analysis

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

Data availability

All the source data derived from this report have been deposited in the Mendeley Dataset (Laliotis et al., 2020[88]). Raw RNA-Seq and ChIP-Seq data of NCI-H522 cell lines, transduced with lentiviral constructs of shControl, shIWS1, shIWS1/WT-R or shIWS1/MT-R will be deposited in Gene Expression Omnibus (GEO).

Acknowledgments

The authors wish to thank all the members of the Tsichlis Lab for helpful discussions. We also thank Dr Samir Achaya, Dr Joal Beane and Dr Esmerina Tilli for reviewing the manuscript before the submission. This work was supported by the NIH grant R01 CA186729 to P.N.T., the NIH grant R01 CA198117 to P.N.T and V.C, by start-up funds from the OSUCCC to P.N.T,, from the National Center for Advancing Translational Sciences grant KL2TR002734 to L.S. G.I.L is supported by the Pelotonia Post-Doctoral fellowship from OSUCCC.

References

1. Glisovic, T., Bachorik, J.L., Yong, J. and Dreyfuss, G., 2008. RNA-binding proteins and post-transcriptional gene regulation. *FEBS letters*, *582*(14), pp.1977-1986.
2. Jewer, M., Findlay, S.D. and Postovit, L.M., 2012. Post-transcriptional regulation in cancer progression. *Journal of cell communication and signaling*, *6*(4), pp.233-248.
3. Pan, Q., Shai, O., Lee, L.J., Frey, B.J. and Blencowe, B.J., 2008. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nature genetics*, *40*(12), p.1413.
4. Paronetto, M.P., Passacantilli, I. and Sette, C., 2016. Alternative splicing and cell survival: from tissue homeostasis to disease. *Cell Death & Differentiation*, *23*(12), pp.1919-1929.

5. Zhang, X., Chen, M.H., Wu, X., Kodani, A., Fan, J., Doan, R., Ozawa, M., Ma, J., Yoshida, N., Reiter, J.F. and Black, D.L., 2016. Cell-type-specific alternative splicing governs cell fate in the developing cerebral cortex. *Cell*, *166*(5), pp.1147-1162.
6. Coomer, A.O., Black, F., Greystoke, A., Munkley, J. and Elliott, D.J., 2019. Alternative splicing in lung cancer. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*.
7. Siegel, R.L., Miller, K.D. and Jemal, A., 2016. Cancer statistics, 2016. *CA: a cancer journal for clinicians*, *66*(1), pp.7-30.
8. Yuan, M., Huang, L.L., Chen, J.H., Wu, J. and Xu, Q., 2019. The emerging treatment landscape of targeted therapy in non-small-cell lung cancer. *Signal Transduction and Targeted Therapy*, *4*(1), pp.1-14.
9. Sequist, L.V., Neal, J.W., Jett, J.R. and Ross, M.E., 2016. Personalized, genotype directed therapy for advanced non-small cell lung cancer. *UpToDate, UpToDate*.
10. Sanidas, I., Polytarchou, C., Hatziapostolou, M., Ezell, S.A., Kottakis, F., Hu, L., Guo, A., Xie, J., Comb, M.J., Iliopoulos, D. and Tsichlis, P.N., 2014. Phosphoproteomics screen reveals akt isoform-specific signals linking RNA processing to lung cancer. *Molecular cell*, *53*(4), pp.577-590.
11. Li, Z., Li, Q., Han, L., Tian, N., Liang, Q., Li, Y., Zhao, X., Du, C. and Tian, Y., 2016. Pro-apoptotic effects of splice-switching oligonucleotides targeting Bcl-x pre-mRNA in human glioma cell lines. *Oncology reports*, *35*(2), pp.1013-1019.
12. Todaro, M., Gaggianesi, M., Catalano, V., Benfante, A., Iovino, F., Biffoni, M., Apuzzo, T., Sperduti, I., Volpe, S., Cocorullo, G. and Gulotta, G., 2014. CD44v6 is a marker of constitutive and reprogrammed cancer stem cells driving colon cancer metastasis. *Cell stem cell*, *14*(3), pp.342-356.
13. Oltean, S. and Bates, D.O., 2014. Hallmarks of alternative splicing in cancer. *Oncogene*, *33*(46), pp.5311-5318.

14. Calabretta, S., Bielli, P., Passacantilli, I., Pilozzi, E., Fendrich, V., Capurso, G., Delle Fave, G. and Sette, C., 2016. Modulation of PKM alternative splicing by PTBP1 promotes gemcitabine resistance in pancreatic cancer cells. *Oncogene*, *35*(16), pp.2031-2039.
15. Vo, N.S., Reynolds, S.M., Zhao, Y., Climente-González, H., Chai11, S., Wang, F., Varghese, R., Huang, M., Liang, W.W., Wyczalkowski, M.A. and Sengupta, S., Systematic Analysis of Splice-Site-Creating Mutations in Cancer.
16. Giaj Levra, M., Andre Olaussen, K., Novello, S. and Soria, J.C., 2014. PARP inhibitors: an interesting pathway also for non-small cell lung cancer?. *Current pharmaceutical design*, *20*(24), pp.3875-3882.
17. Seiler, M., Yoshimi, A., Darman, R., Chan, B., Keaney, G., Thomas, M., Agrawal, A.A., Caleb, B., Csibi, A., Sean, E. and Fekkes, P., 2018. H3B-8800, an orally available small-molecule splicing modulator, induces lethality in spliceosome-mutant cancers. *Nature medicine*, *24*(4), p.497.
18. Listerman, I., Sapra, A.K. and Neugebauer, K.M., 2006. Cotranscriptional coupling of splicing factor recruitment and precursor messenger RNA splicing in mammalian cells. *Nature structural & molecular biology*, *13*(9), p.815.
19. Luco, R.F., Pan, Q., Tominaga, K., Blencowe, B.J., Pereira-Smith, O.M. and Misteli, T., 2010. Regulation of alternative splicing by histone modifications. *Science*, *327*(5968), pp.996-1000.
20. Pradeepa, M.M., Sutherland, H.G., Ule, J., Grimes, G.R. and Bickmore, W.A., 2012. Psip1/Ledgf p52 binds methylated histone H3K36 and splicing factors and contributes to the regulation of alternative splicing. *PLoS genetics*, *8*(5).
21. Chen, M. and Manley, J.L., 2009. Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nature reviews Molecular cell biology*, *10*(11), pp.741-754.

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

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

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

46. Berry, W.L. and Janknecht, R., 2013. KDM4/JMJD2 histone demethylases: epigenetic regulators in cancer cells. *Cancer research*, *73*(10), pp.2936-2942.

47. Cloos, P.A., Christensen, J., Agger, K., Maiolica, A., Rappsilber, J., Antal, T., Hansen, K.H. and Helin, K., 2006. The putative oncogene GASC1 demethylates tri-and dimethylated lysine 9 on histone H3. *Nature*, *442*(7100), p.307.

48. Li, W., Zhao, L., Zang, W., Liu, Z., Chen, L., Liu, T., Xu, D. and Jia, J., 2011. Histone demethylase JMJD2B is required for tumor cell proliferation and survival and is overexpressed in gastric cancer. *Biochemical and biophysical research communications*, *416*(3-4), pp.372-378.

49. Cai, L., Rothbart, S.B., Lu, R., Xu, B., Chen, W.Y., Tripathy, A., Rockowitz, S., Zheng, D., Patel, D.J., Allis, C.D. and Strahl, B.D., 2013. An H3K36 methylation-engaging Tudor motif of polycomb-like proteins mediates PRC2 complex targeting. *Molecular cell*, *49*(3), pp.571-582.

50. Vezzoli, A., Bonadies, N., Allen, M.D., Freund, S.M., Santiveri, C.M., Kvinlaug, B.T., Huntly, B.J., Göttgens, B. and Bycroft, M., 2010. Molecular basis of histone H3K36me3 recognition by the PWWP domain of Brpf1. *Nature structural & molecular biology*, *17*(5), pp.617-619.

51. Li, F., Mao, G., Tong, D., Huang, J., Gu, L., Yang, W. and Li, G.M., 2013. The histone mark H3K36me3 regulates human DNA mismatch repair through its interaction with MutSα. *Cell*, *153*(3), pp.590-600.

52. Vermeulen, M., Eberl, H.C., Matarese, F., Marks, H., Denissov, S., Butter, F., Lee, K.K., Olsen, J.V., Hyman, A.A., Stunnenberg, H.G. and Mann, M., 2010. Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell*, *142*(6), pp.967-980.

53. Shun, M.C., Botbol, Y., Li, X., Di Nunzio, F., Daigle, J.E., Yan, N., Lieberman, J., Lavigne, M. and Engelman, A., 2008. Identification and characterization of PWWP domain residues

critical for LEDGF/p75 chromatin binding and human immunodeficiency virus type 1 infectivity. *Journal of virology*, *82*(23), pp.11555-11567.Singh, P.K., Plumb, M.R., Ferris, A.L., Iben, J.R., Wu, X., Fadel, H.J., Luke, B.T., Esnault, C., Poeschla, E.M., Hughes, S.H. and Kvaratskhelia, M., 2015. LEDGF/p75 interacts with mRNA splicing factors and targets HIV-1 integration to highly spliced genes. *Genes & development*, *29*(21), pp.2287-2297.

54. Ferris, A.L., Wu, X., Hughes, C.M., Stewart, C., Smith, S.J., Milne, T.A., Wang, G.G., Shun, M.C., Allis, C.D., Engelman, A. and Hughes, S.H., 2010. Lens epithelium-derived growth factor fusion proteins redirect HIV-1 DNA integration. *Proceedings of the National Academy of Sciences*, *107*(7), pp.3135-3140.
55. Paz, I., Kosti, I., Ares Jr, M., Cline, M. and Mandel-Gutfreund, Y., 2014. RBPmap: a web server for mapping binding sites of RNA-binding proteins. *Nucleic acids research*, *42*(W1), pp.W361-W367.
56. Anczuków, O., Akerman, M., Cléry, A., Wu, J., Shen, C., Shirole, N.H., Raimer, A., Sun, S., Jensen, M.A., Hua, Y. and Allain, F.H.T., 2015. SRSF1-regulated alternative splicing in breast cancer. *Molecular cell*, *60*(1), pp.105-117.
57. Millevoi, S., Loulergue, C., Dettwiler, S., Karaa, S.Z., Keller, W., Antoniou, M. and Vagner, S., 2006. An interaction between U2AF 65 and CF Im links the splicing and 3′ end processing machineries. *The EMBO journal*, *25*(20), pp.4854-4864.
58. Shen, H. and Green, M.R., 2004. A pathway of sequential arginine-serine-rich domain-splicing signal interactions during mammalian spliceosome assembly. *Molecular cell*, *16*(3), pp.363-373.
59. Chan, S.P., Kao, D.I., Tsai, W.Y. and Cheng, S.C., 2003. The Prp19p-associated complex in spliceosome activation. *Science*, *302*(5643), pp.279-282.
60. Chan, S.P. and Cheng, S.C., 2005. The Prp19-associated complex is required for specifying interactions of U5 and U6 with pre-mRNA during spliceosome activation. *Journal of Biological Chemistry*, *280*(35), pp.31190-31199.

61. Hogg, R., McGrail, J.C. and O'Keefe, R.T., 2010. The function of the NineTeen Complex (NTC) in regulating spliceosome conformations and fidelity during pre-mRNA splicing.
62. Chanarat, S. and Sträßer, K., 2013. Splicing and beyond: the many faces of the Prp19 complex. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, *1833*(10), pp.2126-2134.
63. David, C.J., Boyne, A.R., Millhouse, S.R. and Manley, J.L., 2011. The RNA polymerase II C-terminal domain promotes splicing activation through recruitment of a U2AF65–Prp19 complex. *Genes & development*, *25*(9), pp.972-983.
64. Watrin, E., Demidova, M., Watrin, T., Hu, Z. and Prigent, C., 2014. Sororin pre-mRNA splicing is required for proper sister chromatid cohesion in human cells. *EMBO reports*, *15*(9), pp.948-955.
65. Wang, J., Xia, C., Pu, M., Dai, B., Yang, X., Shang, R., Yang, Z., Zhang, R., Tao, K. and Dou, K., 2018. Silencing of CDCA5 inhibits cancer progression and serves as a prognostic biomarker for hepatocellular carcinoma. *Oncology reports*, *40*(4), pp.1875-1884.
66. Shen, A., Liu, L., Chen, H., Qi, F., Huang, Y., Lin, J., Sferra, T.J., Sankararaman, S., Wei, L., Chu, J. and Chen, Y., 2019. Cell division cycle associated 5 promotes colorectal cancer progression by activating the ERK signaling pathway. *Oncogenesis*, *8*(3), p.19.
67. Nguyen, M.H., Koinuma, J., Ueda, K., Ito, T., Tsuchiya, E., Nakamura, Y. and Daigo, Y., 2010. Phosphorylation and activation of cell division cycle associated 5 by mitogen-activated protein kinase play a crucial role in human lung carcinogenesis. *Cancer research*, *70*(13), pp.5337-5347
68. Guo, Y.J., Pan, W.W., Liu, S.B., Shen, Z.F., Xu, Y. and Hu, L.L., 2020. ERK/MAPK signalling pathway and tumorigenesis. *Experimental and Therapeutic Medicine*, *19*(3), pp.1997-2007.
69. Wee, P. and Wang, Z., 2017. Epidermal growth factor receptor cell proliferation signaling pathways. *Cancers*, *9*(5), p.52.

70. Kent, L.N. and Leone, G., 2019. The broken cycle: E2F dysfunction in cancer. *Nature Reviews Cancer*, *19*(6), pp.326-338.

71. Begum, J., Day, W., Henderson, C., Purewal, S., Cerveira, J., Summers, H., Rees, P., Davies, D. and Filby, A., 2013. A method for evaluating the use of fluorescent dyes to track proliferation in cell lines by dye dilution. *Cytometry Part A*, *83*(12), pp.1085-1095.

72. Dronamraju, R., Jha, D.K., Eser, U., Adams, A.T., Dominguez, D., Choudhury, R., Chiang, Y.C., Rathmell, W.K., Emanuele, M.J., Churchman, L.S. and Strahl, B.D., 2018. Set2 methyltransferase facilitates cell cycle progression by maintaining transcriptional fidelity. *Nucleic acids research*, *46*(3), pp.1331-1344.

73. Liu, P., Begley, M., Michowski, W., Inuzuka, H., Ginzberg, M., Gao, D., Tsou, P., Gan, W., Papa, A., Kim, B.M. and Wan, L., 2014. Cell-cycle-regulated activation of Akt kinase by phosphorylation at its carboxyl terminus. *Nature*, *508*(7497), pp.541-545.

74. Kim, C.H., Kim, S.H., Park, S.Y., Yoo, J., Kim, S.K. and Kim, H.K., 2015. Identification of EGFR mutations by immunohistochemistry with EGFR mutation–specific antibodies in biopsy and resection specimens from pulmonary adenocarcinoma. *Cancer research and treatment: official journal of Korean Cancer Association*, *47*(4), p.653.

75. Popper, H.H., 2016. Progression and metastasis of lung cancer. *Cancer and Metastasis Reviews*, *35*(1), pp.75-91.

76. West, H.J., Shaw, A.T. and Schild, S.E., 2018. Brain metastases in non-small cell lung cancer.

77. Tomida, S., Takeuchi, T., Shimada, Y., Arima, C., Matsuo, K., Mitsudomi, T., Yatabe, Y. and Takahashi, T., 2009. Relapse-related molecular signature in lung adenocarcinomas identifies patients with dismal prognosis. *Journal of clinical oncology*, *27*(17), pp.2793-2799.

78. Ishiguro, K.I., 2019. The cohesin complex in mammalian meiosis. *Genes to Cells*, *24*(1), pp.6-30.

79. Dreier, M.R., Bekier, M.E. and Taylor, W.R., 2011. Regulation of sororin by Cdk1-mediated phosphorylation. *Journal of cell science*, *124*(17), pp.2976-2987.

80. Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S. and Taatjes, D.J., 2010. Mediator and cohesin connect gene expression and chromatin architecture. *Nature*, *467*(7314), pp.430-435.

81. Karamysheva, Z., Díaz-Martínez, L.A., Warrington, R. and Yu, H., 2015. Graded requirement for the spliceosome in cell cycle progression. *Cell Cycle*, *14*(12), pp.1873-1883.

82. Hofmann, J.C., Husedzinovic, A. and Gruss, O.J., 2010. The function of spliceosome components in open mitosis. *Nucleus*, *1*(6), pp.447-459.

83. Bonnal, S.C., López-Oreja, I. and Valcárcel, J., 2020. Roles and mechanisms of alternative splicing in cancer—implications for care. *Nature Reviews Clinical Oncology*, pp.1-18.

84. Wang, E. and Aifantis, I., 2020. RNA Splicing and Cancer. *Trends in Cancer*.

85. Obeng, E.A., Stewart, C. and Abdel-Wahab, O., 2019. Altered RNA Processing in Cancer Pathogenesis and Therapy. *Cancer discovery*, *9*(11), pp.1493-1510.

86. Mohammed, H., Taylor, C., Brown, G.D., Papachristou, E.K., Carroll, J.S. and D'santos, C.S., 2016. Rapid immunoprecipitation mass spectrometry of endogenous proteins (RIME) for analysis of chromatin complexes. *Nature protocols*, *11*(2), p.316.

87. Skene, P.J., Henikoff, J.G. and Henikoff, S., 2018. Targeted in situ genome-wide profiling with high efficiency for low cell numbers. *Nature protocols*, *13*(5), p.1006.

88. Laliotis, Georgios I.; Chavdoula, Evangelia; Paraskevopoulou, Maria D.; Kaba, Abdul; La Ferlita, Alessandro; Anastas, Vollter; Orlacchio, Arturo; Taraslia, Vasiliki; Vlachos, Ioannis; Hatzigeorgiou, Artemis; Capece, Marina; Palmieri, Dario; Sehgal, Lalit; Carbone, David P.; Coppola, Vincenzo; Tsichlis, Philip N.; Alaimo, Salvatore; Tsatsanis, Christos

(2020), “IWS1 phosphorylation promotes cell proliferation and predicts poor prognosis in EGFR mutant lung adenocarcinoma patients, through the cell cycle-regulated U2AF2 RNA splicing.”, Mendeley Data, v2

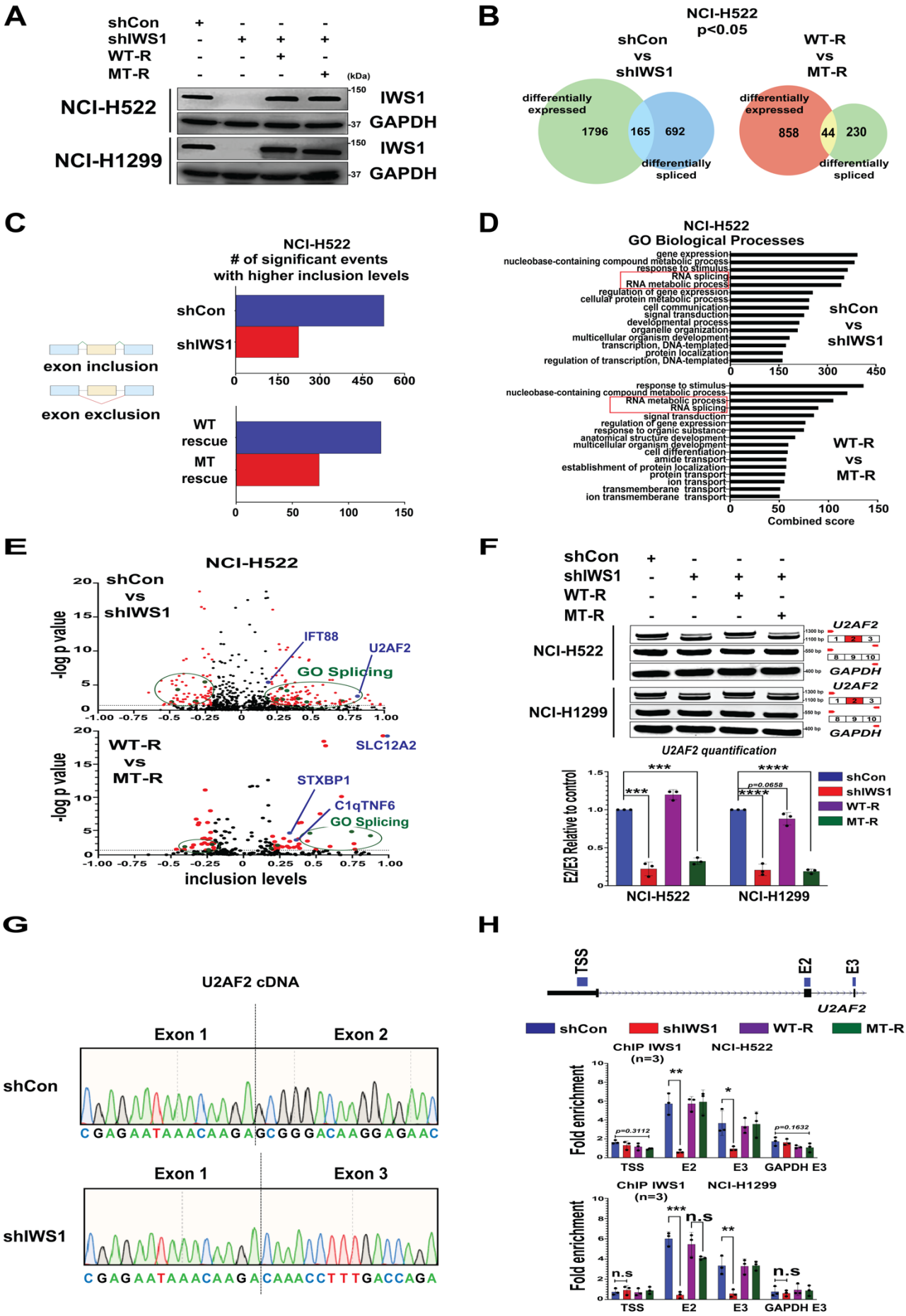

A
shCon
shIWS1
WT-R
MT-R
(kDa)
NCI-H522
NCI-H1299
IWS1
GAPDH
B
NCI-H522
p<0.05
shCon vs shIWS1
WT-R vs MT-R
differentially expressed
differentially spliced
1796
165
692
858
44
230
C
NCI-H522
# of significant events with higher inclusion levels
exon inclusion
exon exclusion
shCon
shIWS1
WT rescue
MT rescue
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
inclusion levels
IFT88
U2AF2
GO Splicing
SLC12A2
STXBP1
C1qTNF6
F
U2AF2
GAPDH
U2AF2 quantification
E2/E3 Relative to control
p=0.0658
shCon
shIWS1
WT-R
MT-R
G
U2AF2 cDNA
Exon 1
Exon 2
Exon 3
shCon
shIWS1
CGAGAATAAACAAGAGCGGGACAAGGAGAAC
CGAGAATAAACAAGACAAACCTTTGACCAGA
H
TSS
E2
E3
U2AF2
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 and anti-GAPDH (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 ($p<0.05$) are illustrated, using Venn diagrams. **C**. Bar graphs show the number of alternative splicing events with exon inclusion (inclusion level >20%) 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 an inclusion level of >0.2 or <-0.2 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, C1qTNF6, STXBP1)*. 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 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 and shIWS1/MT-R NCI-H522 and NCI-H1299 cells. GAPDH 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)

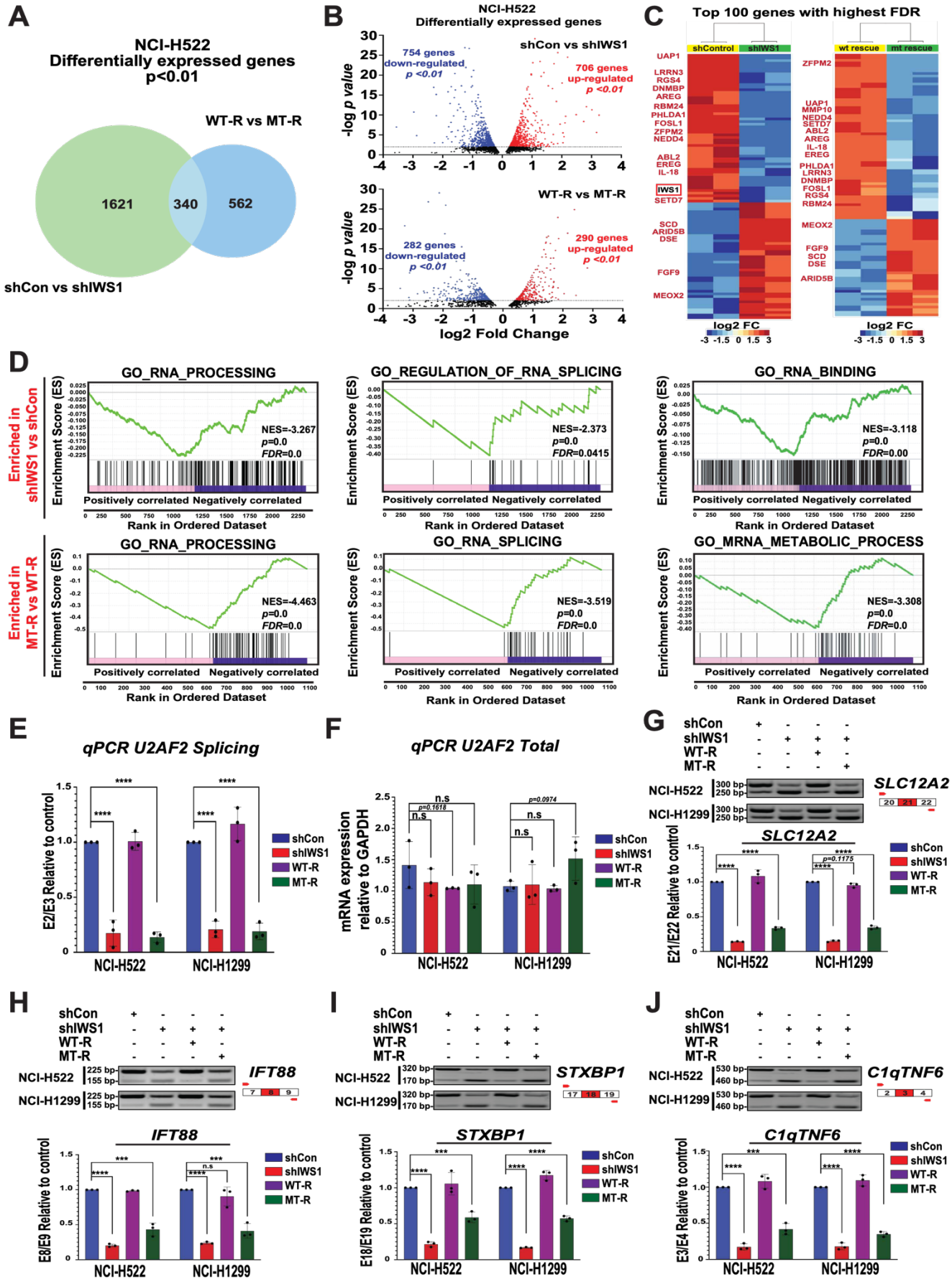

A
NCI-H522
Differentially expressed genes
p<0.01
WT-R vs MT-R
1621
340
562
shCon vs shIWS1
B
NCI-H522
Differentially expressed genes
shCon vs shIWS1
754 genes
down-regulated
p <0.01
706 genes
up-regulated
p <0.01
WT-R vs MT-R
282 genes
down-regulated
p <0.01
290 genes
up-regulated
p <0.01
-log p value
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
log2 FC
D
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
Enriched in shIWS1 vs shCon
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
Enriched in MT-R vs WT-R
Enrichment Score (ES)
Positively correlated
Negatively correlated
Rank in Ordered Dataset
E
qPCR U2AF2 Splicing
E2/E3 Relative to control
NCI-H522
NCI-H1299
shCon
shIWS1
WT-R
MT-R
F
qPCR U2AF2 Total
mRNA expression relative to GAPDH
n.s
p=0.1618
p=0.0974
G
SLC12A2
E21/E22 Relative to control
p=0.1175
H
IFT88
E8/E9 Relative to control
I
STXBP1
E18/E19 Relative to control
J
C1qTNF6
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 ($p<0.01$) 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 (highest FDR) in shIWS1 relative to shControl and in shIWS1/MT-R relative to shIWS1/WT-R NCI-H522 cells. Genes present in both sets 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.* Quantitative RT-PCR showing the *E2/E3* ratio in the *U2AF2* mRNA transcripts in shControl, shIWS1, shIWS1/WT-R and shIWS1/MT-R NCI-H522 and NCI-H1299 cells. Bars show the E2/E3 ratio (mean ± SD) in all the cell lines, normalized relative to the shControl. **F.** Quantitative RT-PCR, showing the abundance of all the *U2AF2* mRNA transcripts 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*. **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)

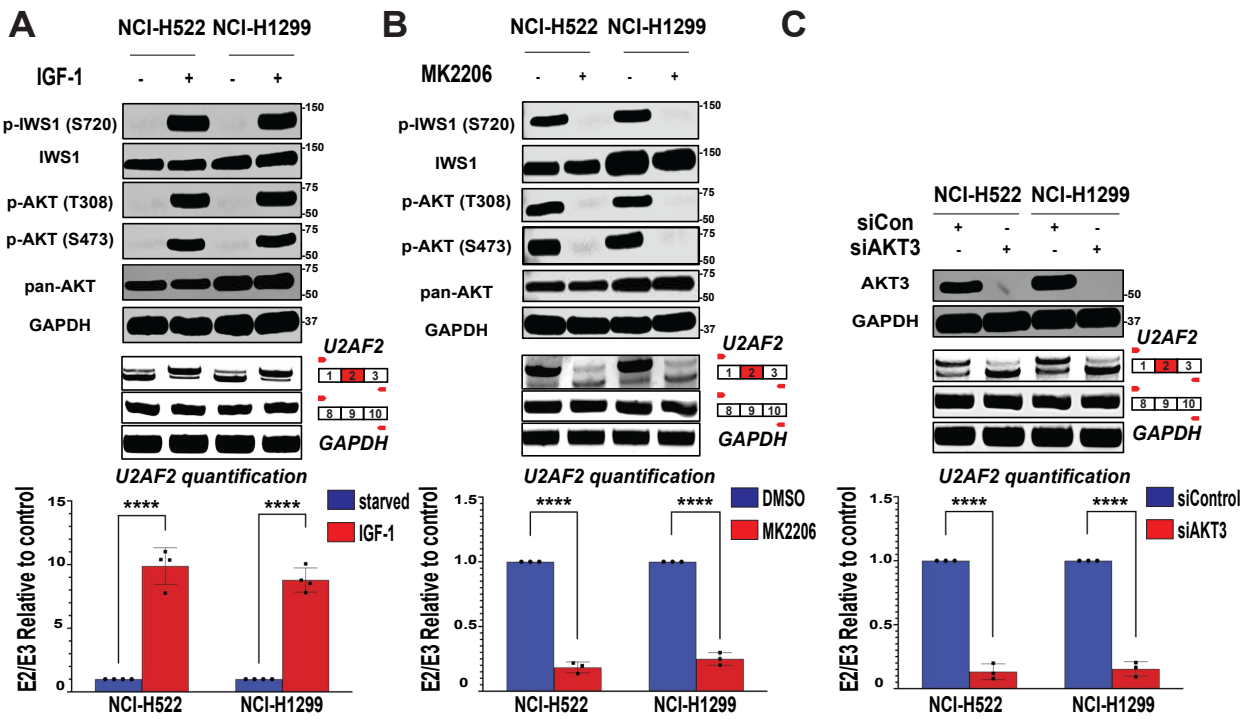


Figure 2. IWS1 phosphorylation-dependent mRNA splicing of U2AF2 is regulated by serum and IGF-1-induced signals transduced by AKT3

A. *IGF-1 promotes the IWS1-dependent U2AF2 exon 2 (E2) inclusion.* (Upper panel) Following serum starvation for 24 hours, NCI-H522 and NC-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 IGF-1 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.* (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 DMSO-treated cells. **C.** *U2AF2 exon 2 (E2) inclusion depends on AKT3***.** (Upper panel) NCI-H522 and NCI-H1299 cell lysates, harvested 48 hr after transfection with siAkt3 or siControl, 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 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 *siAKT3*-transfected NCI-H522 and NCI-H1299 cells relative to the ratio in mock-transfected 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).

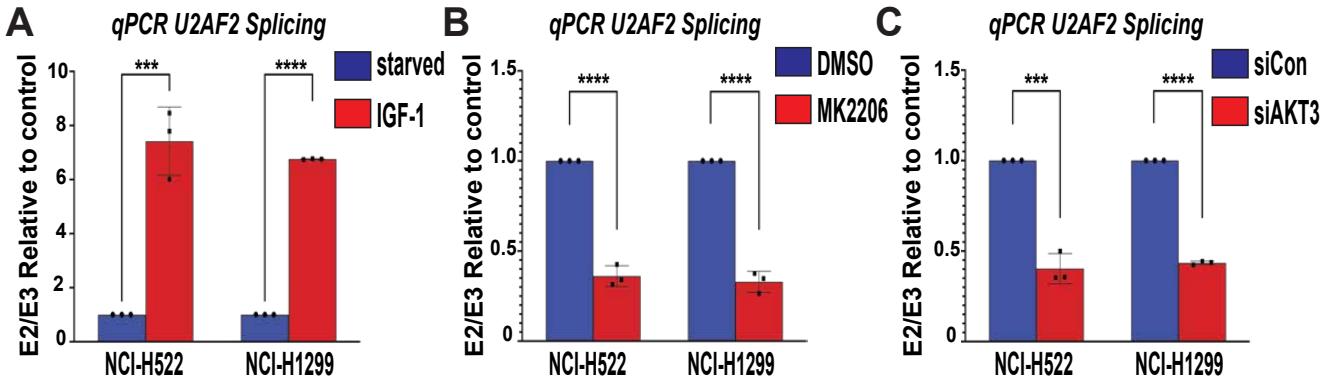


Figure S2 (relative to Figure 2). IWS1 phosphorylation-dependent mRNA splicing of U2AF2 is regulated by serum and IGF-1-induced signals transduced by AKT3

A. B. and C. The *U2AF2* mRNA E2/E3 ratio in the cells in figure 2 was also measured by quantitative RT-PCR. Bars show the E2/E3 ratio in IGF-1, MK2206 and siAKT3-treated 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)

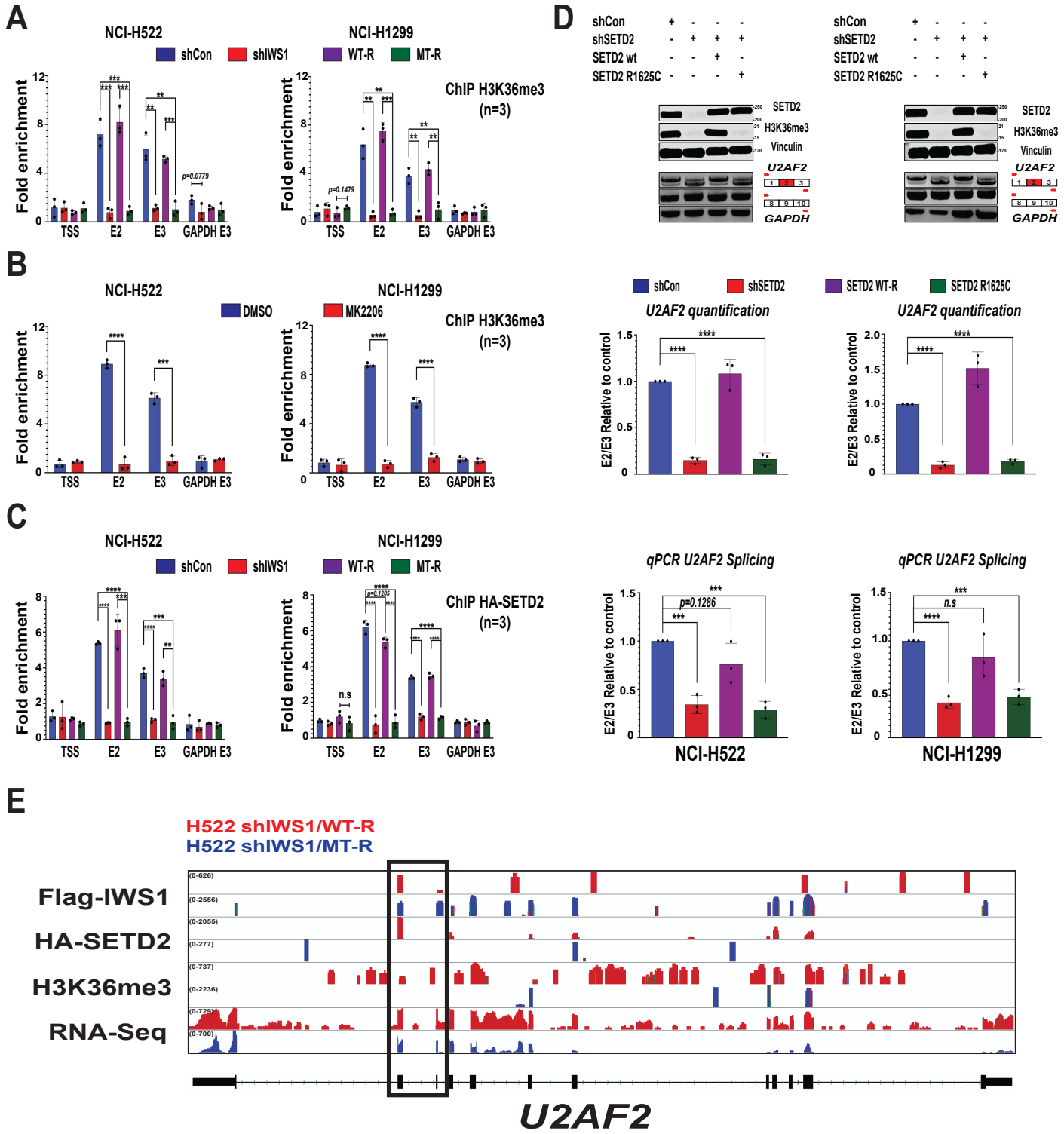

A
NCI-H522
NCI-H1299
shCon
shIWS1
WT-R
MT-R
ChIP H3K36me3
(n=3)
Fold enrichment
TSS
E2
E3
GAPDH E3
p=0.0779
p=0.1479
B
DMSO
MK2206
U2AF2 quantification
E2/E3 Relative to control
shSETD2
SETD2 WT-R
SETD2 R1625C
C
ChIP HA-SETD2
(n=3)
p=0.1206
n.s
qPCR U2AF2 Splicing
p=0.1286
D
SETD2
H3K36me3
Vinculin
U2AF2
GAPDH
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 S3A. **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.

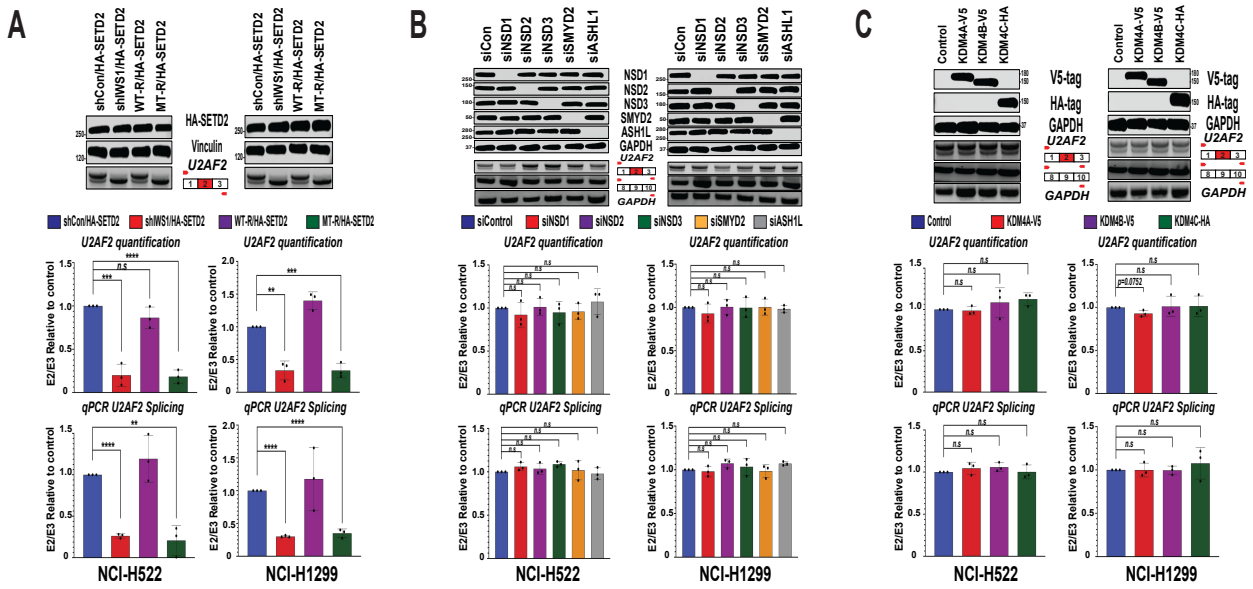


Figure S3 (relative to Figure 3). *U2AF2* Exon 2 inclusion, induced by IWS1 phosphorylation at Ser720/Thr721, depends on H3K36 trimethylation by SETD2.

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 methyltransferases 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) .

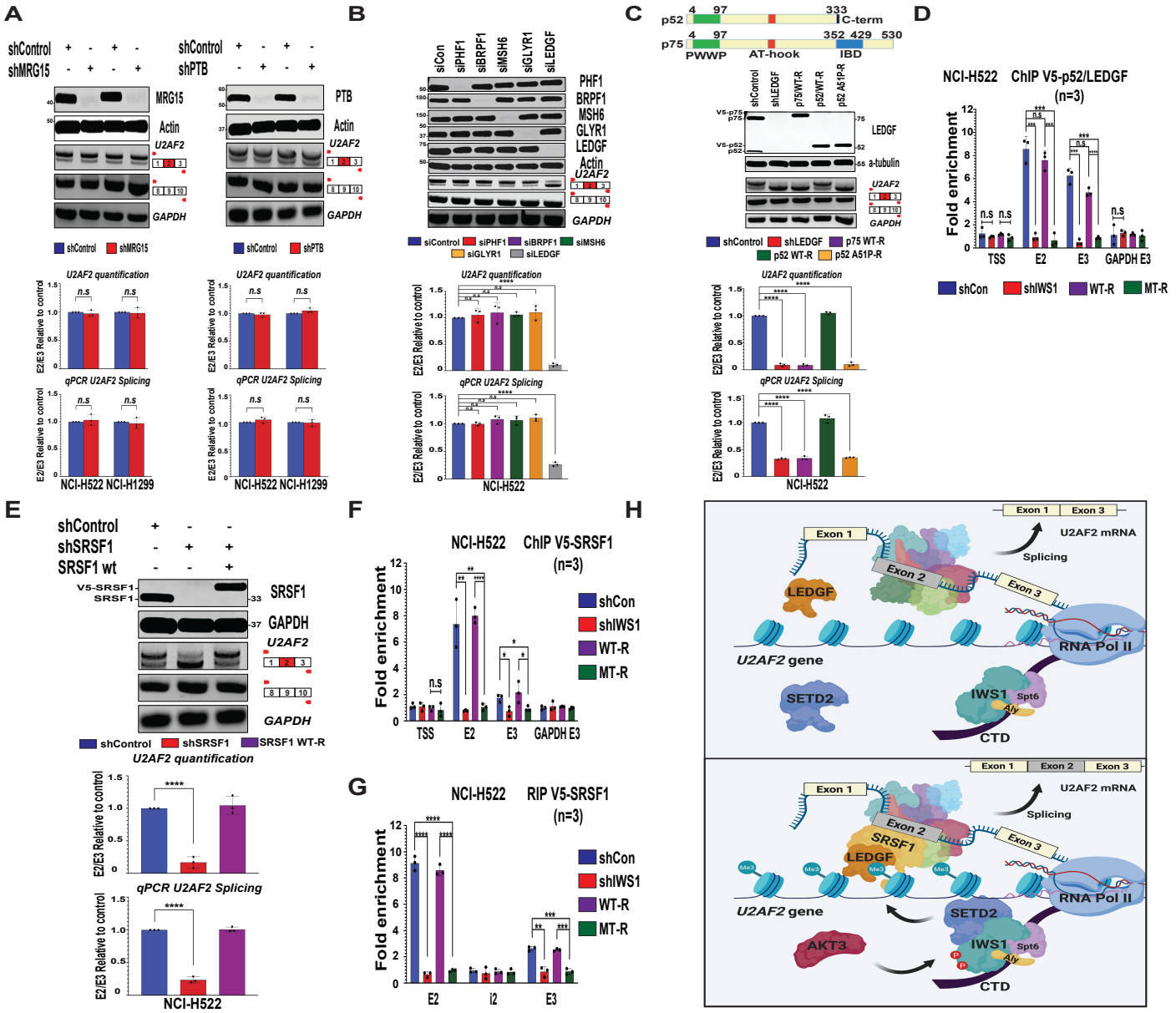


Figure 4. The regulation of the alternative splicing of the *U2AF2* exon 2 by IWS1 and IWS1 phosphorylation, depends on the p52 isoform of the H3K36me3 reader LEDGF and its splicing partner SRSF1.

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[54]). (Middle panel) Western blots of lysates of shControl, shLEDGF, shLEDGF/LEDGF-p75-R,shLEDGF/LEDGF-p52WT-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 S4D. **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 S4F. **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.

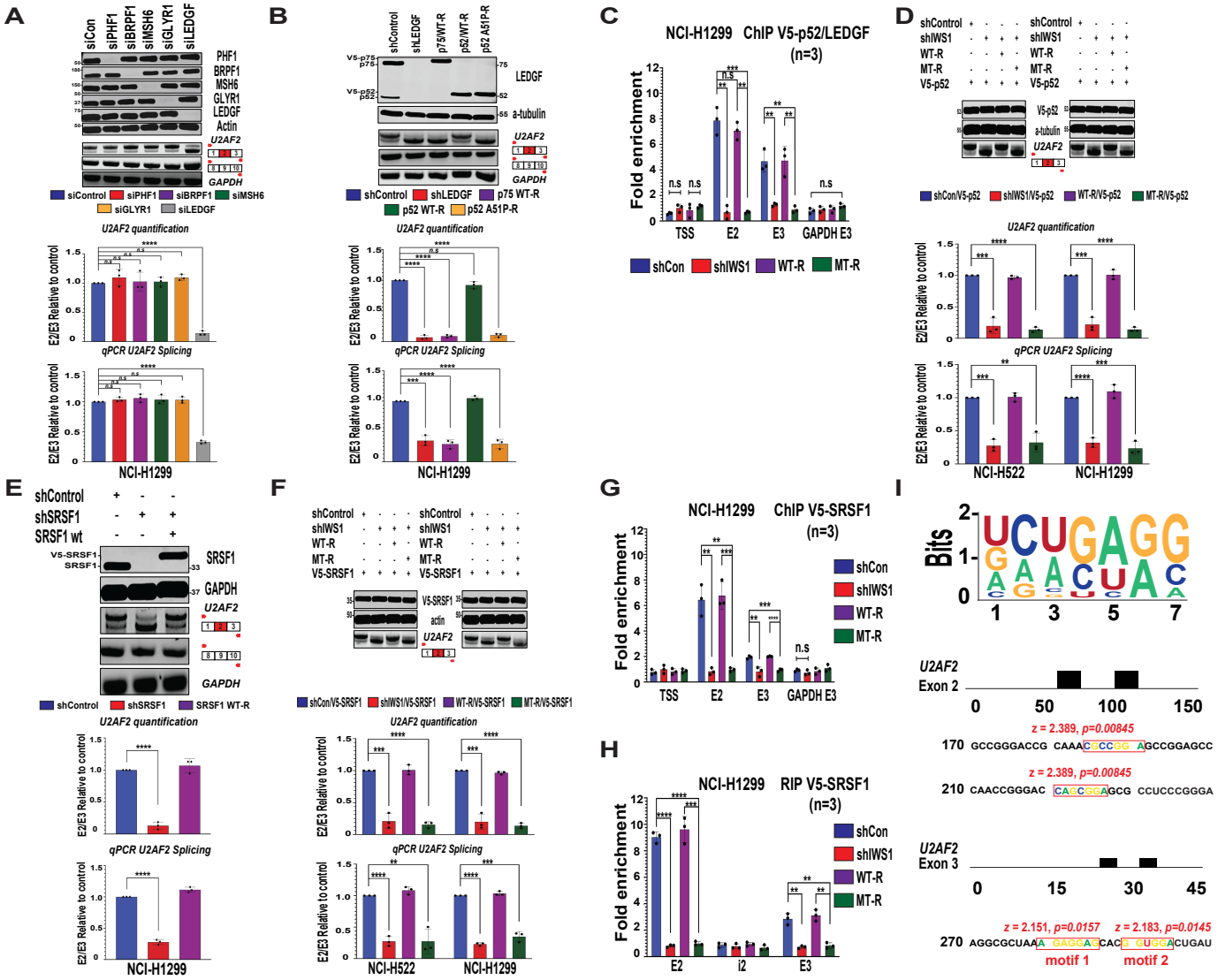


Figure S4 (relative to Figure 4). The regulation of the alternative splicing of the *U2AF2* exon 2 by *IWS1* and *IWS1* phosphorylation, depends on the p52 isoform of the H3K36me3 reader *LEDGF* and its splicing partner *SRSF1*.

A. The knockdown of the indicated histone H3K36me3 readers in NCI-H1299 cells confirmed that only the knockdown of LEDGF results the exclusion of exon 2 from the mature *U2AF2* mRNA transcript. This is a repeat of the experiment in figure 4B, which was done in NCI-H522 cells. The results confirmed that LEDGF controls the mRNA splicing of *U2AF2* in multiple cell lines. **B.** 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. The results confirmed that the p52 isoform of wild type LEDGF controls the mRNA splicing of *U2AF2* in multiple cell lines. **C.** 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. **D.** *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. **E.** *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. **F.** *V5-SRSF1-WT 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. **G.** The experiments in NCI-H522 cells, in figure 4F, was repeated in NCI-H1299 cells with identical results. **H.** The experiments in NCI-H522 cells, in figure 4G, was repeated in NCI-H1299 cells with identical results. **I.** (Upper panel) The 7-nt SRSF1 binding motif, involved in the regulation of alternative RNA splicing by SRSF1 (Anczuków et al., 2015[56]) (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)

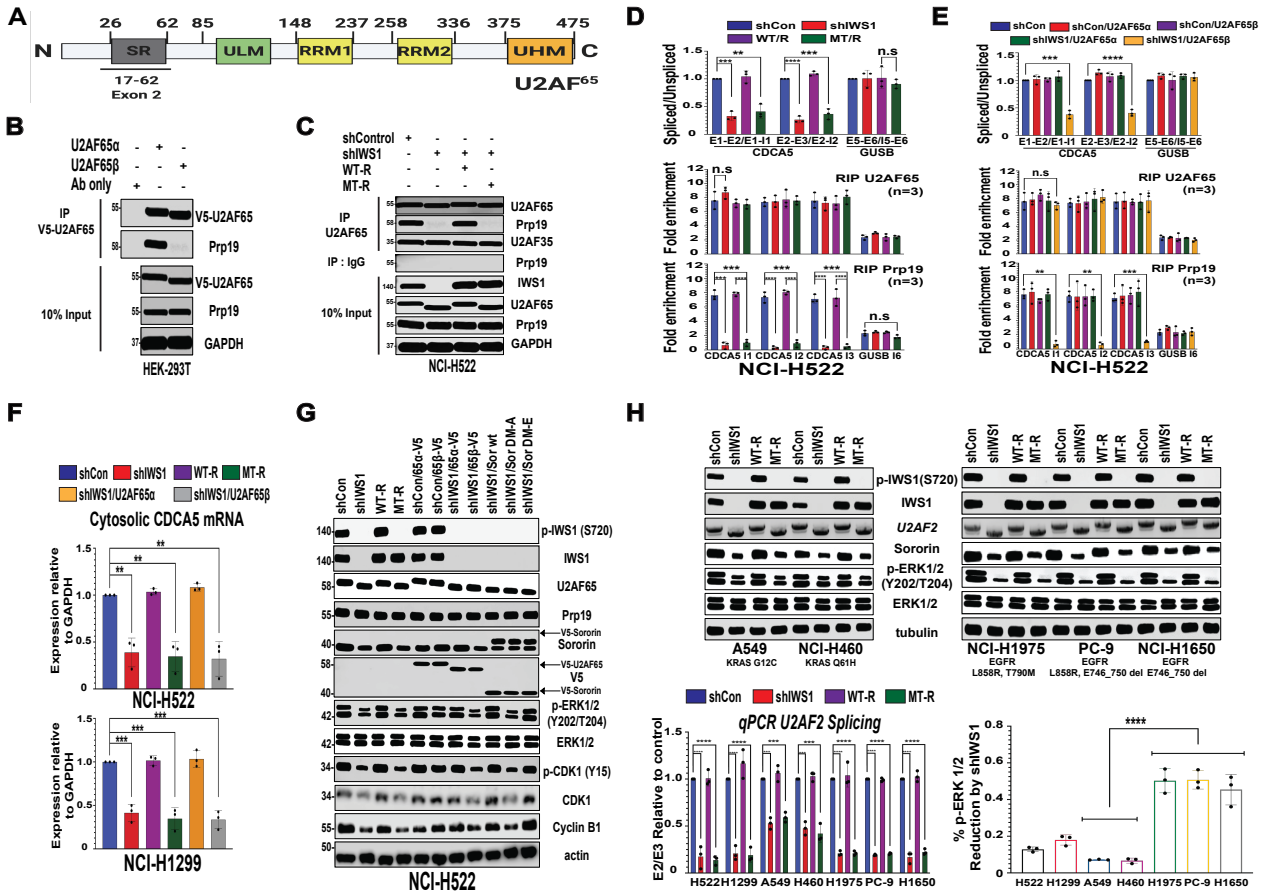


Figure 5. IWS1 phosphorylation controls the CDCA5/ERK phosphorylation feedback loop, through U2AF2 alternative RNA splicing.

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 SR 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, doesn't 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 S5B. **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 S5E.

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 K-RAS mutations.* Lysates of shControl, shIWS1, shIWS1/WT-R and shIWS1/MT-R K-RAS (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 K-RAS 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 KRAS mutant (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).

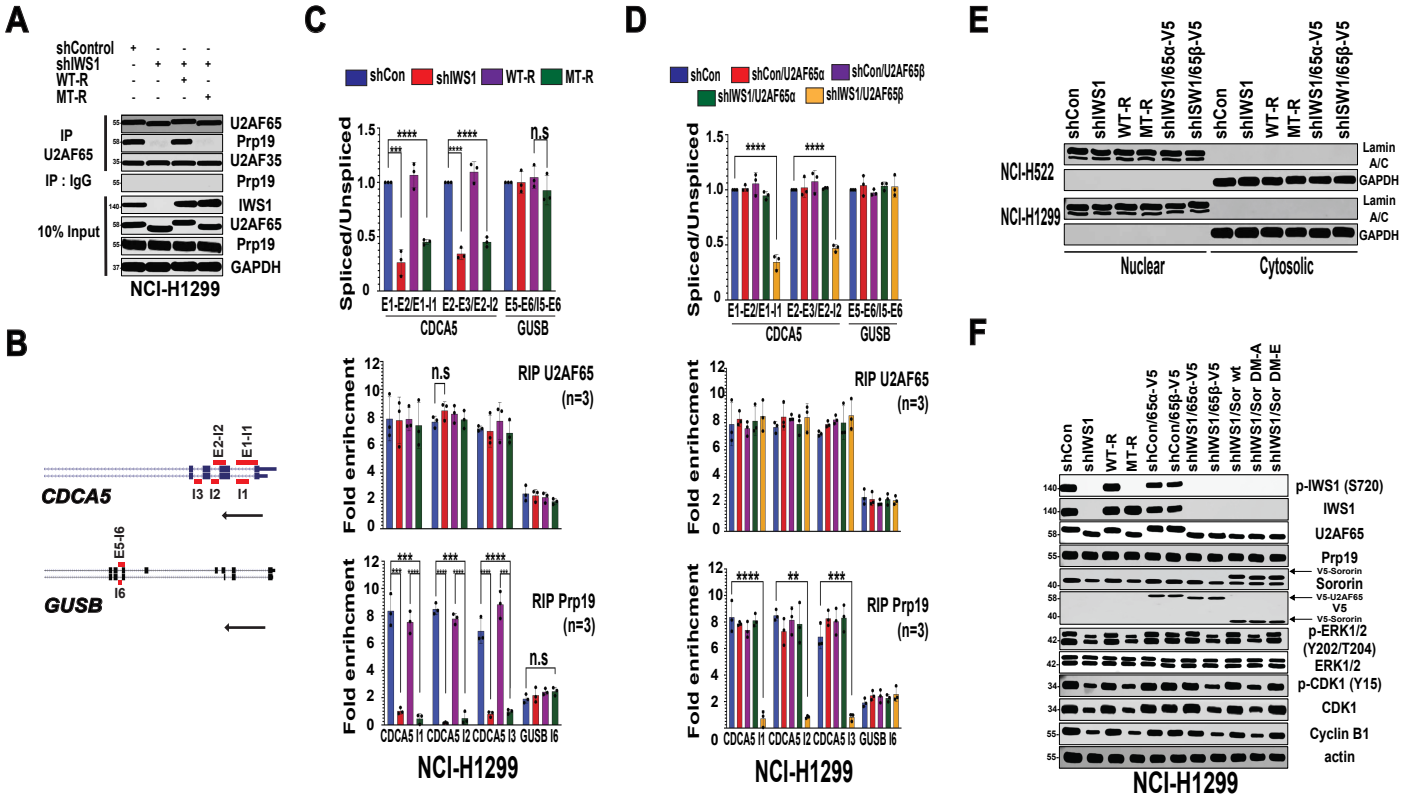


Figure S5 (relative to Figure 5). IWS1 phosphorylation controls the CDCA5/ERK phosphorylation feedback loop, through U2AF2 alternative RNA splicing.

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.

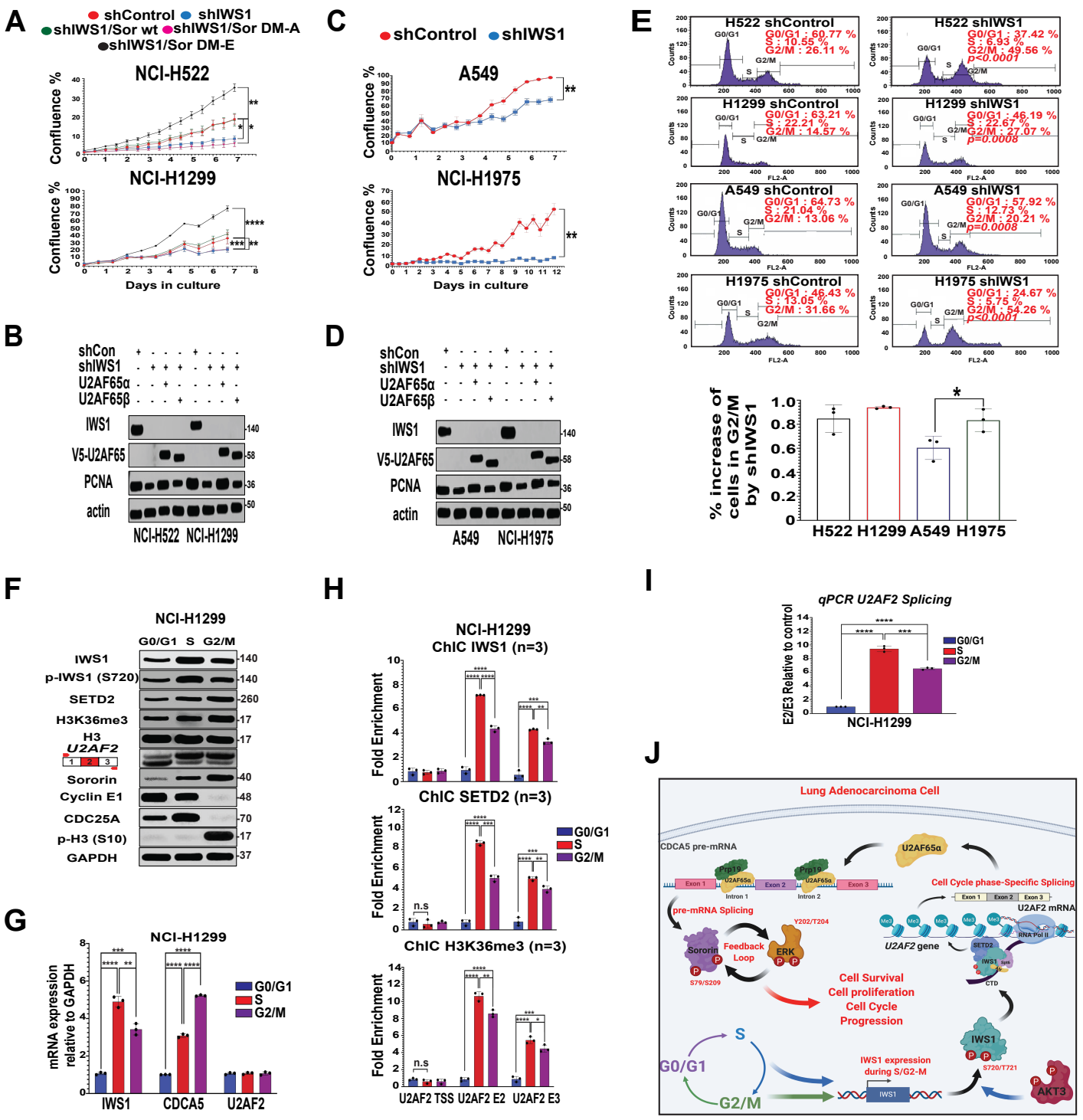

A
shControl
shIWS1
shIWS1/Sor wt
shIWS1/Sor DM-A
shIWS1/Sor DM-E
NCI-H522
NCI-H1299
Confluence %
Days in culture
B
shCon
shIWS1
U2AF65α
U2AF65β
IWS1
V5-U2AF65
PCNA
actin
NCI-H522
NCI-H1299
C
shControl
shIWS1
A549
NCI-H1975
Confluence %
Days in culture
D
shCon
shIWS1
U2AF65α
U2AF65β
IWS1
V5-U2AF65
PCNA
actin
A549
NCI-H1975
E
H522 shControl
G0/G1 : 60.77 %
S : 10.55 %
G2/M : 26.11 %
H522 shIWS1
G0/G1 : 37.42 %
S : 6.93 %
G2/M : 49.56 %
p<0.0001
H1299 shControl
G0/G1 : 63.21 %
S : 22.21 %
G2/M : 14.57 %
H1299 shIWS1
G0/G1 : 46.19 %
S : 22.67 %
G2/M : 27.07 %
p=0.0008
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
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
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
G0/G1
S
G2/M
NCI-H1299
J
Lung Adenocarcinoma Cell
CDCA5 pre-mRNA
U2AF65α
Prp19
Exon 1
Exon 2
Exon 3
Intron 1
Intron 2
pre-mRNA Splicing
Sororin
S79/S209
Feedback Loop
ERK
Y202/T204
Cell Cycle phase-Specific Splicing
U2AF2 mRNA
U2AF2 gene
RNA Pol II
SETD2
IWS1
CTD
Cell Survival
Cell proliferation
Cell Cycle Progression
IWS1
S720/T721
AKT3
S
G0/G1
G2/M
IWS1 expression during S/G2-M

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.* Lysates derived from the indicated NCI-H522 and NCI-H1299 cells were probed with the indicated antibodies. In agreement with the data in A, PCNA was downregulated by shIWS1 in both cell lines, and its downregulation was rescued by U2AF65α, but not U2AF65β. **C.** *IWS1 phosphorylation promotes cell proliferation more robustly in the EGFR mutant lung adenocarcinoma cell line NCI-H1975, than in the K-RAS 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.** Cell lysates derived from shControl and shIWS1 A549 and NCI-H1975 cells, before and after rescue with U2AF65α, or U2AF65β, were probed with the indicated antibodies. PCNA was again downregulated by shIWS1 in both cell lines, and its downregulation was rescued by U2AF65α, but not U2AF65β. **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 (left upper panels) or shIWS1 (Right upper panels) lentiviral constructs. 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 K-RAS 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.

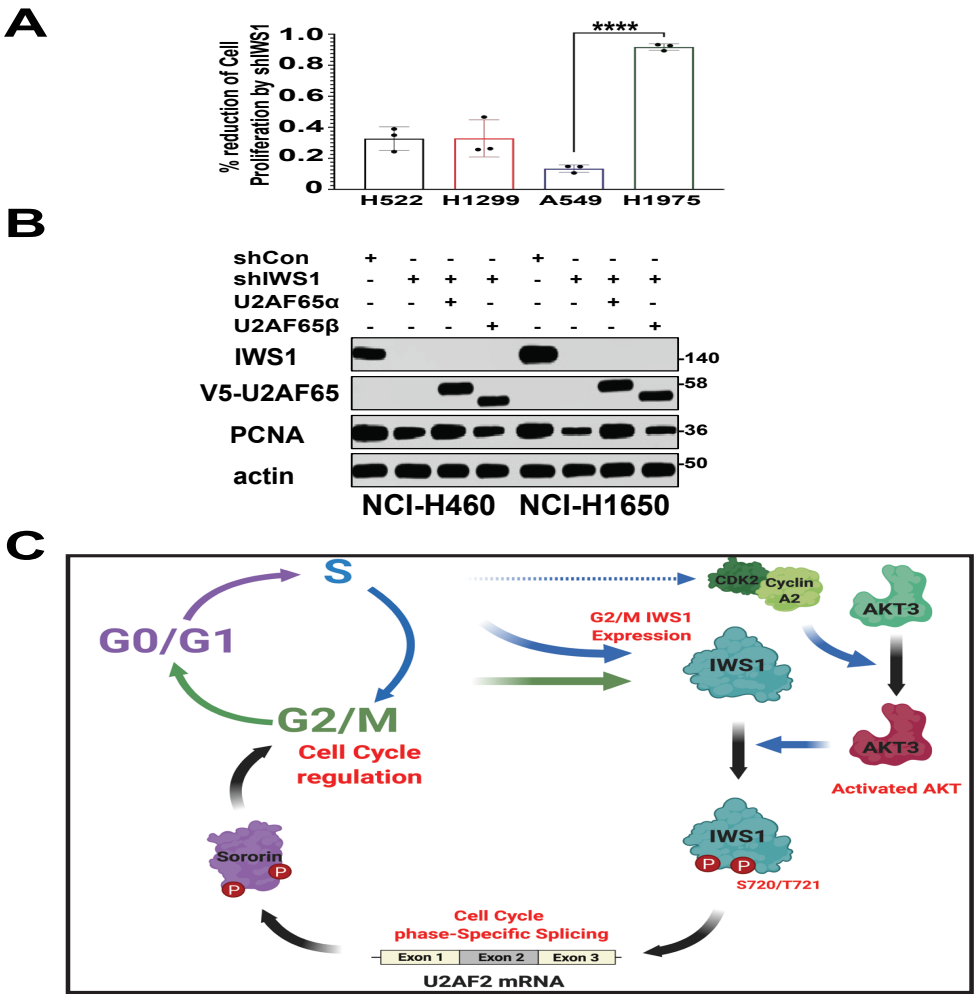


Figure S6 (relative to Figure 6). IWS1 phosphorylation promotes cell proliferation by controlling the Sororin/ERK phosphorylation feedback loop, through U2AF2 RNA splicing.

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.** Cell lysates derived from the indicated NCI-H460 and NCI-H1650 cells, were probed with the indicated antibodies, in a repeat of the experiments in figures 6B and 6D, in two additional cell lines. **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[73]) 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.

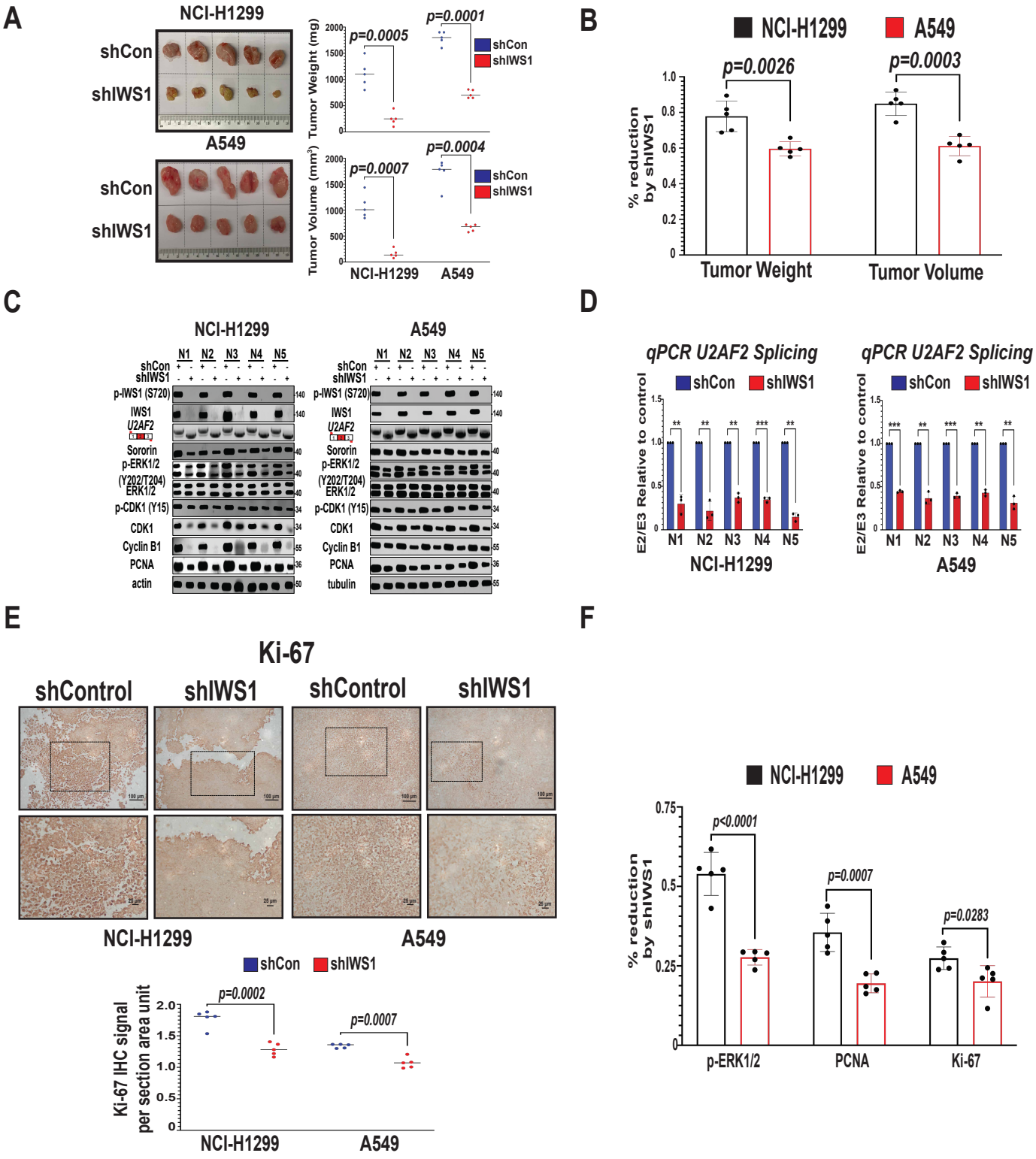
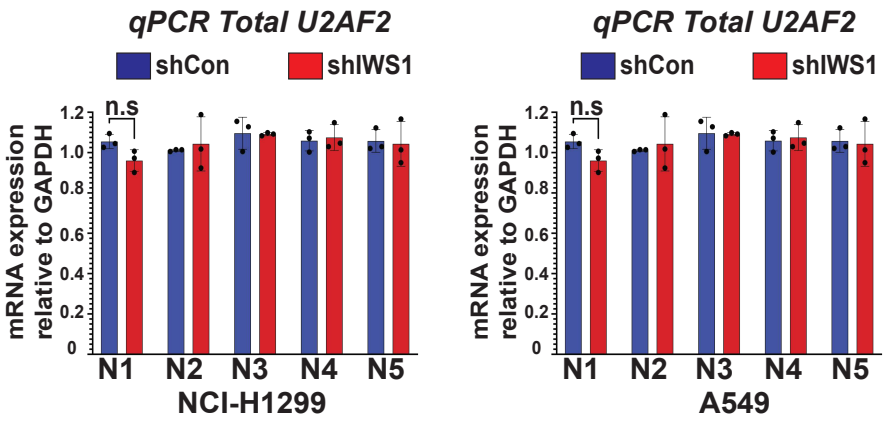

A
NCI-H1299
shCon
shIWS1
A549
shCon
shIWS1
Tumor Weight (mg)
p=0.0005
p=0.0001
shCon
shIWS1
Tumor Volume (mm³)
p=0.0007
p=0.0004
NCI-H1299
A549
B
NCI-H1299
A549
p=0.0026
p=0.0003
% reduction by shIWS1
Tumor Weight
Tumor Volume
C
NCI-H1299
A549
N1 N2 N3 N4 N5
shCon
shIWS1
p-IWS1 (S720)
IWS1
U2AF2
Sororin
p-ERK1/2 (Y202/T204)
ERK1/2
p-CDK1 (Y15)
CDK1
Cyclin B1
PCNA
actin
tubulin
D
qPCR U2AF2 Splicing
shCon
shIWS1
E2/E3 Relative to control
N1 N2 N3 N4 N5
NCI-H1299
A549
E
Ki-67
shControl
shIWS1
shControl
shIWS1
NCI-H1299
A549
shCon
shIWS1
Ki-67 IHC signal per section area unit
p=0.0002
p=0.0007
NCI-H1299
A549
F
NCI-H1299
A549
% reduction by shIWS1
p<0.0001
p=0.0007
p=0.0283
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 or A549 cells. (N=5 mice/group). (Left panels) Images of the induced tumors, harvested at 4 weeks (NCI-H1299 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.** From the weight and volume of the tumors induced by shIWS1-transduced cells and shControl-transduced cells, 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. **C.** *IWS1 phosphorylation controls the Sororin/ERK phosphorylation axis in vivo.* Cell lysates derived from NCI-H1299 and A549 (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. **D.** *IWS1 phosphorylation controls U2AF2 alternative RNA splicing in vivo.* Quantitative RT-PCR showing the *E2/E3 U2AF2* ratio using RNA derived from tumors in Fig 7C. Bars show the mean E2/E3 U2AF2 in tumors derived from NCI-H1299 (left) or A549 (right) shIWS1 cells relative to those from shControl. The error bars are SD of at least 3 technical replicates. **$p<0.001$, ***$p<0.001$. (paired t-test) **E.** (Upper panels) Ki-67 staining of tumor xenografts of shControl and shIWS1-transduced NCI-H1299 (left) and A549 (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 and A549 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 and A549 cell lines. Bars show the mean reduction of the markers, derived from the Western blots data in Figure 7C (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


qPCR Total U2AF2
shCon
shIWS1
n.s
mRNA expression relative to GAPDH
1.2
1.0
0.8
0.6
0.4
0.2
0
N1
N2
N3
N4
N5
NCI-H1299
qPCR Total U2AF2
shCon
shIWS1
n.s
mRNA expression relative to GAPDH
N1
N2
N3
N4
N5
A549

B

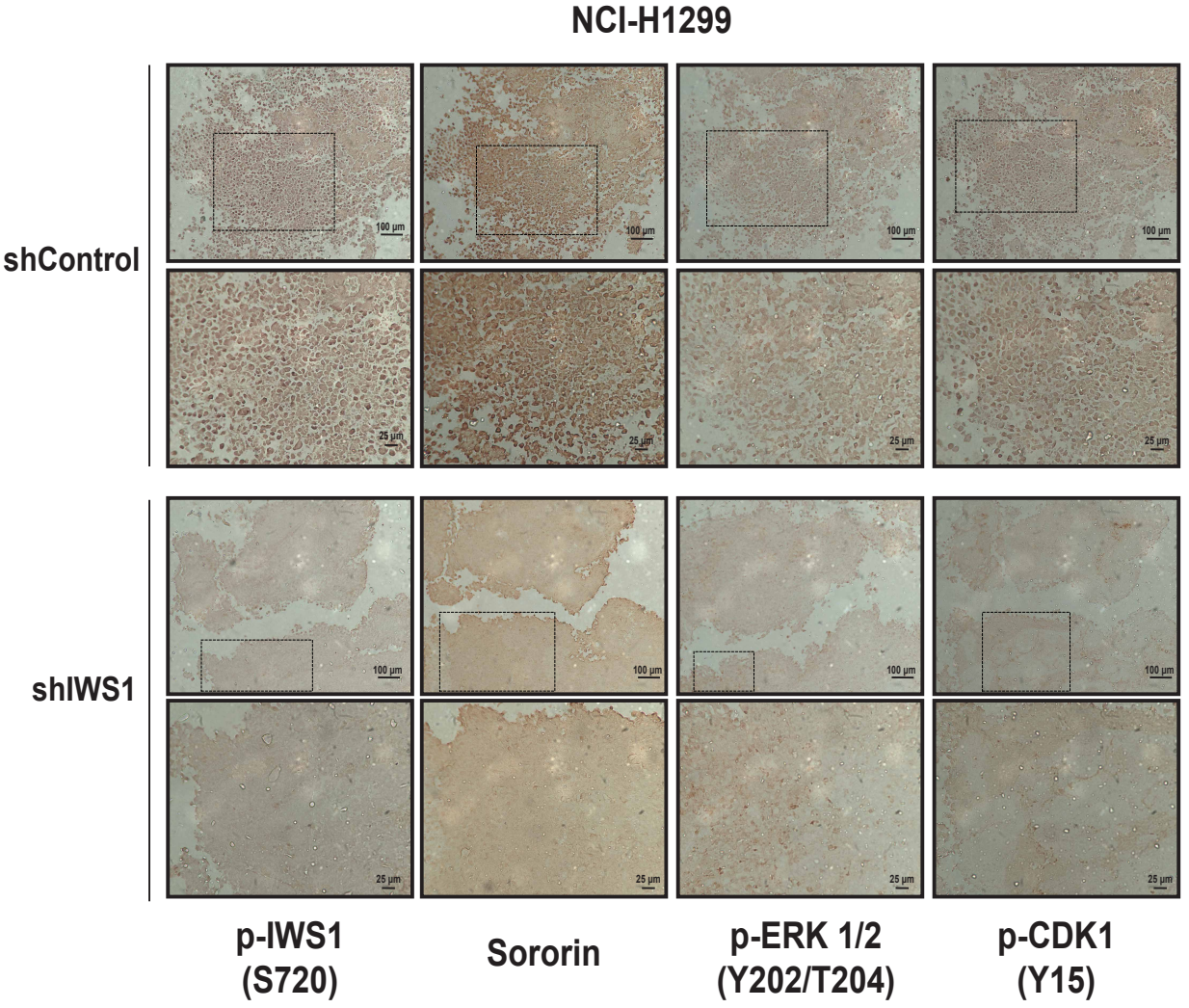

NCI-H1299
shControl
shIWS1
100 µm
25 µm
p-IWS1 (S720)
Sororin
p-ERK 1/2 (Y202/T204)
p-CDK1 (Y15)

Figure S7 (Relative to Figure 7). IWS1 phosphorylation controls tumor growth *in vivo*, by regulating the U2AF2/Sororin/ERK axis.

A. Quantitative RT-PCR showing the total U2AF2 in lysates derived from tumors in Fig 7C. Bars show the mean U2AF2 levels, relative to GAPDH, in tumors derived from NCI-H1299 (left) or A549 (right) shControl and shIWS1 cells. The error bars are SD of at least 3 technical replicates. n.s : non-significant (paired t-test) **B.** *IWS1 controls the Sororin/ERK phosphorylation axis in vivo.* (Upper panels) Phospho-IWS1 (Ser720), Sororin, phospho-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.

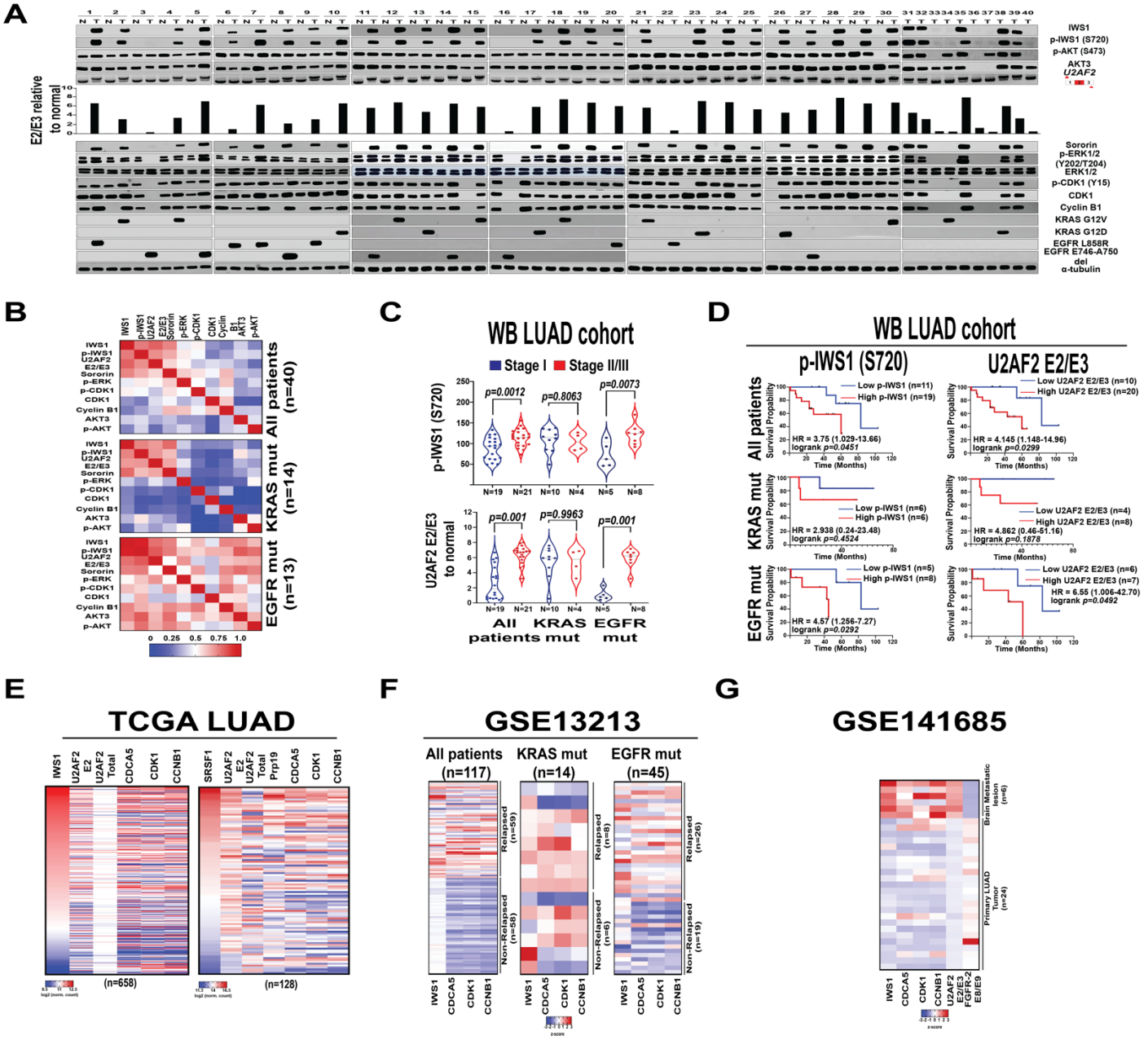

A
E2/E3 relative to normal
IWS1
p-IWS1 (S720)
p-AKT (S473)
AKT3
U2AF2
Sororin
p-ERK1/2 (Y202/T204)
ERK1/2
p-CDK1 (Y15)
CDK1
Cyclin B1
KRAS G12V
KRAS G12D
EGFR L858R
EGFR E746-A750 del
α-tubulin
B
All patients (n=40)
KRAS mut (n=14)
EGFR mut (n=13)
C
WB LUAD cohort
Stage I
Stage II/III
p-IWS1 (S720)
U2AF2 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
Survival Propability
Time (Months)
HR = 3.75 (1.029-13.66)
logrank p=0.0451
HR = 4.145 (1.148-14.96)
logrank p=0.0299
HR = 2.938 (0.24-23.48)
logrank p=0.4524
HR = 4.862 (0.46-51.16)
logrank p=0.1878
HR = 4.57 (1.256-7.27)
logrank p=0.0292
HR = 6.55 (1.006-42.70)
logrank p=0.0492
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
Brain Metastatic lesion
Primary LUAD Tumor
H
TCGA LUAD
IWS1
U2AF2 E2/E3
FGFR-2 E8/E9
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 IWS1 (n=21)
High IWS1 (n=131)
HR = 0.625 (0.27-1.432)
logrank p=0.2669
Low IWS1 (n=18)
High IWS1 (n=17)
HR = 4.002 (1.59-10.02)
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
Low U2AF2 E2-E3 (n=70)
High U2AF2 E2/E3 (n=88)
HR = 0.97 (0.57-1.68)
logrank p=0.9394
Low U2AF2 E2/E3 (n=39)
High U2AF2 E2/E3 (n=27)
HR = 3.95 (1.631-9.56)
logrank p=0.0023
High FGFR-2 E8/E9 (n=218)
Low FGFR-2 E8/E9 (n=218)
HR = 2.75 (2.206-3.420)
logrank p=0.0001
High FGFR-2 E8/E9 (n=77)
Low FGFR-2 E8/E9 (n=78)
HR = 1.26 (0.74-2.13)
logrank p=0.3807
High FGFR-2 E8/E9 (n=39)
Low FGFR-2 E8/E9 (n=27)
HR = 2.25 (1.06-4.78)
logrank p=0.0342

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 *K-RAS* mutant tumors.

A. Lysates form 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 *K-RAS* 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 *K-RAS* 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 8A. Separate curves are shown again for all the patients, or for the patients with *K-RAS* 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 *K-RAS* 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/CDCA5/ERK/CDK1/CCNB1 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), *K-RAS* 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 8 exclusion pathways correlate strongly with metastasis in lung adenocarcinomas.* 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. **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 *FGFR-2* E8/E9 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 *K-RAS* mutant tumors. Statistical analyses were performed using the log rank test and Cox's proportional hazards model.

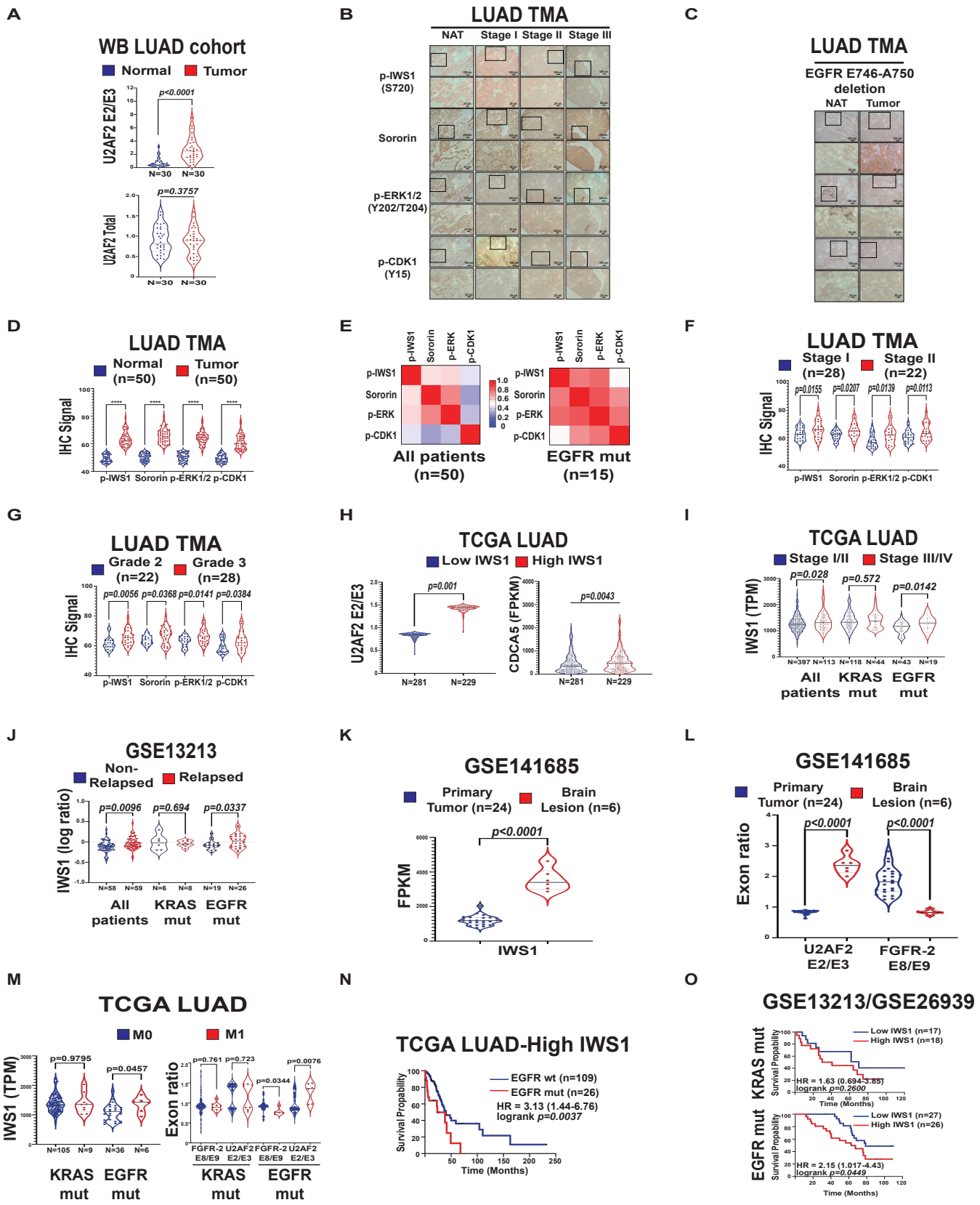

A
WB LUAD cohort
Normal
Tumor
U2AF2 E2/E3
p<0.0001
N=30
N=30
U2AF2 Total
p=0.3757
N=30
N=30
B
LUAD TMA
NAT
Stage I
Stage II
Stage III
p-IWS1 (S720)
Sororin
p-ERK1/2 (Y202/T204)
p-CDK1 (Y15)
C
LUAD TMA
EGFR E746-A750 deletion
NAT
Tumor
D
LUAD TMA
Normal (n=50)
Tumor (n=50)
IHC Signal
p-IWS1
Sororin
p-ERK1/2
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
KRAS 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
KRAS 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
FGFR-2 E8/E9
M
TCGA LUAD
M0
M1
IWS1 (TPM)
p=0.9795
p=0.0457
N=105
N=9
N=36
N=6
KRAS mut
EGFR mut
Exon ratio
p=0.761
p=0.723
p=0.0344
p=0.0076
FGFR-2 E8/E9
U2AF2 E2/E3
FGFR-2 E8/E9
U2AF2 E2/E3
KRAS 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
KRAS mut
Low IWS1 (n=17)
High IWS1 (n=18)
Survival Propability
HR = 1.63 (0.694-3.85)
logrank p=0.2600
Time (Months)
EGFR mut
Low IWS1 (n=27)
High IWS1 (n=26)
Survival Propability
HR = 2.15 (1.017-4.43)
logrank p=0.0449
Time (Months)

Figure S8 (relative to Figure 8). The AKT/p-IWS1/U2AF2 axis is active in human lung adenocarcinoma and affects the clinical outcome of EGFR mutant patients.

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***.** 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 *K-RAS* 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 *K-RAS* mutant tumors. The horizontal lines indicate mean values for IWS1. Statistical analyses were performed using the one-sided unpaired t-test. **K.** The expression of *IWS1* is higher in brain metastases of lung adenocarcinomas than in primary tumors, while the *E8/E9* 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 *FGFR-2 E8/E9* ratio. Violin plots were based on the same data shown in Figure 8G. Statistical analyses were performed using the one-sided unpaired t-test. **L.** The *E2/E3* ratio in mature *U2AF2* mRNA transcripts are higher in the same patient group as Fig S8L. The black lines indicate the mean *U2AF2 E2/E3* and *FGFR-2 E8/E9* 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.** (Left panel) Violin plots showing the expression of IWS1 in transcripts per million (TPM) in *K-RAS* 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 harbor EGFR mutations. (Right panel) Exon E2/E3 ratio in the *U2AF2* mRNA and exon E8/E9 ratio in the *FGFR-2* mRNA in *K-RAS* 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 harbor *EGFR* mutations. The black lines indicate the mean IWS1 expression, and the mean *U2FA2* E2/E3 and *FGFR-2* E8/E9 ratios. 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 harboring EGFR, than those with lung adenocarcinomas harboring K-RAS mutations.* Kaplan-Meier curves showing the impact of the abundance of IWS1 on the survival of patients with lung adenocarcinomas harboring K-RAS 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.