BIOLOGICAL SCIENCES: Biochemistry

Human prefoldin modulates co-transcriptional pre-mRNA splicing

Payán-Bravo L [1,2], Peñate X [1,2] *, Cases I [3], Pareja-Sánchez Y [1], Fontalva S [1,2], Odriozola Y [1,2], Lara E [1], Jimeno-González S [2,5], Suñé C [4], Reyes JC [5], Chávez S [1,2].

[1] Instituto de Biomedicina de Sevilla, Universidad de Sevilla-CSIC-Hospital Universitario V. del Rocío, Seville, Spain.
[2] Departamento de Genética, Facultad de Biología, Universidad de Sevilla, Seville, Spain.
[3] Centro Andaluz de Biología del Desarrollo, CSIC-Universidad Pablo de Olavide, Seville, Spain.
[4] Department of Molecular Biology, Institute of Parasitology and Biomedicine "López Neyra" IPBLN-CSIC, PTS, Granada, Spain.
[5] Andalusian Center of Molecular Biology and Regenerative Medicine-CABIMER, Junta de Andalucia-University of Pablo de Olavide-University of Seville-CSIC, Seville, Spain.
Correspondence: Sebastián Chávez, IBiS, campus HUVR, Avda. Manuel Siurot s/n, Sevilla, 41013, Spain. Phone: +34-955923127: e-mail: schavez@us.es. * Co-corresponding; xenia@us.es.

Abstract

Prefoldin is a heterohexameric complex conserved from archaea to humans that plays a cochaperone role during the cotranslational folding of actin and tubulin monomers. Additional functions of prefoldin in the cell nucleus have been described, including a positive contribution to transcription elongation and chromatin dynamics in yeast. Here we show that prefoldin depletion provoked transcriptional alterations across the genome of human cells. Severe pre-mRNA splicing defects were also detected, particularly under serum stimulation conditions. We found a significant impairment of co-transcriptional splicing during transcription elongation, which explains why the expression of long genes with a high number of introns was affected the most. We detected prefoldin binding to the gene body of transcribed genes and found that its absence caused significant decrease in the levels of Ser2 phosphorylation of the RNA polymerase II carboxy-terminal domain. Moreover, lack of prefoldin reduced the association of the splicing factor U2AF65 with chromatin, an association that is known to be dependent on Ser2 phosphorylation. Altogether the reported results indicate that human prefoldin is able to modulate gene expression by influencing phosphorylation of elongating RNA polymerase II, and thereby regulating co-transcriptional splicing efficiency.

Keywords: Prefoldin, transcription elongation, pre-mRNA splicing

Significance Statement

Prefoldin is a protein complex conserved from archaea to humans. It is known that human prefoldin facilitates folding of cytoskeleton monomers in the cytoplasm, but its role in the cell nucleus is unclear. We found prefoldin bound to chromatin in transcribed genes and detected accumulation of unspliced pre-mRNA precursors when prefoldin was depleted, particularly in serum-activated cells. We also found that lack of prefoldin decreases the levels of a phosphorylated form of RNA polymerase II that is the best-known marker of transcription elongation. This defect, in turn, impairs recruitment of the splicing factor U2AF65, provoking the uncoupling of pre-mRNA splicing from transcription. These results show that prefoldin is, in addition to a cytoplasmic chaperone, a modulator of co-transcriptional pre-mRNA processing.

Introduction

Prefoldin is a protein cochaperone present in all living organisms except eubacteria (1). It is closely related to cytoskeleton assembly as it functions in the co-translational folding of actin and tubulin monomers (2, 3). While this first discovered function is cytoplasmic, prefoldin has also been found to play nuclear roles (4).
In eukaryotes, canonical prefoldin is composed of six different subunits, named PFDN1-6. Subunits PFDN2 and 6, are also components of the Uri/Prefoldin-like complex, involved in the cytoplasmic assembly of several protein complexes(5).
All the six yeast subunits are present in the nucleus and four of them (PFDN1, PFDN2, PFDN5 and PFDN6) are recruited to yeast chromatin in a transcription-dependent manner. Lack of these subunits cause transcription elongation defects, which is reflected in lower levels of long transcripts and genome-wide alterations in the profiles of active RNA polymerase II (RNA pol II) molecules (6). Yeast prefoldin mutants also show synthetic phenotypes when combined with mutations in transcription elongation and chromatin factors, and exhibit defects in cotranscriptional chromatin dynamics, indicating an important contribution of prefoldin to RNA pol II-dependent transcription and other phenomena linked to this process (6).
Human prefoldin subunits PFDN1, PFDN3 and PFDN5 have also been described to be involved in nuclear functions, including removal of HIV integrase (7), inhibition of c-Myc (8) and repression of some genes (9, 10). However, a possible general involvement of human prefoldin in transcriptional and co-transcriptional events has not been investigated. In this work we have explored this possibility and found a significant contribution of prefoldin to the expression of the human genome, particularly at the level of co-transcriptional mRNA splicing.

Materials and methods

<u>Cell culture and transfection</u>
Human HCT116 cells were maintained in McCoy's 5A modified medium supplemented with 10 % foetal bovine serum, 60 mg/l penicillin, and 100 mg/l streptomycin. Additionally, 2 µg/µl G418 were added to resistant cell lines.
For transfection of plasmids and siRNA, Turbofect transfection reagent (Thermo Fisher) and Oligofectamine (Invitrogen) were used, respectively.
Sequences of the siRNAs are as follows: siC (CGUACGCGGAAUACUUCGA); siPFDN2 (CAGCCUAGUGAUCGAUACA); siPFDN5 (AGAGAAGACAGCUGAGGAU).
<u>Generation of PFDN5 null cell line by CRISPR-CAS9</u>
The gRNA TTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCGGTACAGACCAAGTATG was cloned by Gibson assembly (SgiDNA) into the plasmid U6GRNA (Addgene 4148). The resulting plasmid and plasmid hCas9 (Addgene 41815) were cotransfected using Turbofect transfection reagent (Thermo Fisher). PFDN5 null mutant clones were selected by the absence of PFDN5 protein in western blot.

Plasmids and clone selection

PFDN5 ORF (splice variant alpha) was cloned EcoRI/SalI into pCMV-Tag 2A (N-terminal Flag tag, Agilent) or XhoI/EcoRI into pEGFP-C1 (N-terminal GFP tag, Novopro). Expression of Flag-PFDN5 was assessed in the different clones by western blot, whereas in the case of PFDN5-GFP it was by flow cytometry.

Antibodies

The following antibodies were used for immunofluorescence (IF), chromatin immunoprecipitation (ChIP) or western blotting (WB): anti-PFDN5 (Santa Cruz, sc-271150, WB); anti-GAPDH (Santa Cruz, sc-47724, WB); anti-PFDN2 (Bethyl, A304-807A, WB); anti-Rpb1 CTD Ser2 phosphorylated (Abcam, ab5095, ChIP); anti-Rpb3 (Abcam, ab202893, ChIP); anti-Rpb1 8WG16 (Thermo Fisher MA1-10882, ChIP); anti-FLAG (Sigma, F3165, ChIP); anti-Fibrillarin (Abcam, ab4566, IF); anti-U2AF65 (Santa Cruz sc-53942); anti-PRPF19 (Lührmann Laboratory); anti-PRGL1 (Lührmann Laboratory)

RNA extraction and quantitative RT-PCR assays

RNA was extracted with RNAeasy mini kit (Qiagen). 1 µg of DNase treated RNA was retrotranscribed by Super-Script III (Invitrogen) using random hexanucleotides as primers. Relative quantification of specific cDNAs was achieved using SYBR premix (Takara) in a Lyght Cycler (Roche).

RNA library preparation and sequencing

Total RNA was depleted of 18S rRNA and then libraries were prepared. 50 bp paired-end sequencing was performed at the Genomic Unit at CRG (Barcelona, Spain).

RNA-seq analysis

Quality control was done with FSTQC v.0.11.5. Alignment of the reads was performed to Homo_sapiens_GRCh37.74 genome. Differential expression was analysed by Kallisto {Bray, 2016 #654}. Exon ratio was calculated using HISAT2 {Pertea, 2016 #655}. GO analysis was performed using GOFunction (Bioconductor, FDR<0.05). To analyse how the length and number of introns of the genes affects the variation of the siPFDN5 / siControl ($\log_2$(FC) expression differences, genes with the highest expression level were excluded from this analysis (291 genes excluded fout of 11700), since these are generally short (11, 12) and therefore could bias the results. Alternative splicing was performed with DESeq2 (R Studio, FDR<0.05)

FAS Minigene assay

Cells were transfected with 0.5 µg of the reporter plasmid (13). 24 hours later, RNA was extracted AND retrotranscribed with MLV retrotranscriptase (Invitrogen) and a specific Fas primer (AAGCTTGCATCGAATCAGTAG). After 35 PCR cycles, DNA fragments were separated in an agarose gel and the bands quantified using Image Lab 6.0.

Chromatin immunoprecipitation

Human chromatin immunoprecipitation was carried out by fixation (10 minutes, 1 % formaldehyde, 37 ºC) and quenching (5 minutes, 2 % glycine), cells were lysed in lysis buffer (50 mM Tris-HCl, 10 mM EDTA, 1% SDS, pH 8.1 and protease inhibitors). Samples were then sonicated (Bioruptor, Diagenode), centrifuged, and the supernatant incubated with magnetic beads coupled to the specific antibody. Immunoprecipitated

material was washed, treated with proteinase K and the resultant DNA was quantified by qPCR.

Results

Transcriptomic analysis of PFDN5 and PFDN2 deficient human cells

In order to explore the potential role of prefoldin in human gene expression, we conducted experiments to analyse the impact of prefoldin perturbation on the cell transcriptome. We designed siRNAs for *in-vivo* depletion of subunits PFDN2 and PFDN5. PFDN5 is present in the canonical complex, whereas PFDN2 is a component of both the canonical and the Uri/prefoldin-like complex. We reasoned that those changes that are related to the canonical complex would be detected in the two experiments, whereas those related to the Uri/prefoldin-like would be found just in the PFDN2-depletion experiment.

Transfection of HCT116 human colon carcinoma cells with siRNAs that target *PFDN2* and *PFDN5* mRNAs had a 67 % and 55 % efficiency, respectively, in reducing the specific protein after 72 hours of treatment (Sup. Figure S1A).

In order to test the impact of prefoldin depletion on gene expression under different conditions of transcription regulation, we used serum stimulation after starvation, which has long been recognized as a fast and simple system to study global changes in transcription (14). We designed an experiment in which we transfected cells with the siRNAs for 24 hours, then serum-starved them for 48 hours, and finally stimulated by adding serum. We took samples just before and 90 minutes after stimulation and analysed changes in total transcriptome by RNA-seq.

Global changes in RNA-seq signal were apparent in siPFDN5-treated cells both before and after serum stimulation (Figure 1A, left panels). The same was true for siPFDN2-treated cells, although in this case there were less differentially expressed genes (Figure 1A, right panels). The different number of genes affected was also reflected in the different extent of cell death caused by the two siRNAs (Sup. Figure S1B). In both cases, the number of significantly affected genes was higher after serum stimulation (Figure 1A, Sup. Figure S2A), suggesting that prefoldin may play a role in supporting regulated gene expression.

The higher levels of PFDN2 present in human cells, as compared to PFDN5 (15), may explain the lower impact of PFDN2 depletion. Nevertheless, gene ontology (GO) analysis of differentially expressed genes after serum stimulation showed significant enrichment of identical categories in both PFDN2- and PFDN5-dependent genes (Sup. Figure S2B). Before serum stimulation, despite a stronger effect of PFDN5 depletion, there was a significant overlap between the genes affected by PFDN5 and PFDN2 depletions (Sup. Figure S2C). Moreover, most genes affected in both knockdowns had the same direction of change (reduced expression in both or overexpression in both, Sup. Figure S2D), suggesting that most gene expression changes detected in PFDN5- and PFDN2-depleted cells are reflecting shared functions of the two subunits.

Next, we focused on the 198 genes that were more significantly induced ($log_2FC > 1$, FDR < 0.05) by serum in control HCT116 cells (Figure 1B). Induction of 82 and 44 genes were significantly impaired in PFDN5- and PFDN2-depleted cells, respectively (Figure 1B). Comparison of these two subsets of genes resulted in a very significant overlap (Figure 1C).
Considered altogether, these results are compatible with a positive contribution of prefoldin to gene expression, particularly under conditions of gene regulation.

Prefoldin deficiency causes stronger expression defects in long genes with a high number of introns

Yeast prefoldin has an impact on transcription elongation (6). If this were the case for human prefoldin, we would expect that longer genes would have more problems to be expressed properly in prefoldin knockdowns. To test this possibility, we divided the genes into quintiles according to their length and plotted the differential expression before and after serum stimulation. Before serum stimulation, in cells treated with siPFDN5 or siPFDN2, there was no sign of negative correlation between gene length and differential expression (Figure 2A, left panel and Sup. Figure S3A, left panels). If any, we detected a slight positive effect in very long genes in siPFDN5-treated cells that was not observed in siPFDN2-treated cells (Figure 2A and Sup. Figure S3A, left panels). In contrast, after stimulation we detected a consistent negative effect of gene length over gene expression in both siRNA treatments (Figure 2A and Sup. Figure S3A, right panels).
In the human genome, gene length is proportional to the number of introns that a gene contains (16). Therefore, our next question was whether the number of introns correlated to the change in expression caused by prefoldin knockdowns. Genes were divided into quintiles depending on their number of introns. When differential expression of each group of genes was plotted, a negative relationship between expression and number of introns was apparent only after serum stimulation (Figure 2B, Sup. Figure S3B). These results are very similar to the previous ones, indicating an effect of the structural properties of genes (length, number of introns) in the expression under prefoldin deficiency conditions, pointing towards an effect of prefoldin in transcription elongation.
The effect of gene length could be a consequence of the effect of the number of introns, or viceversa. In a third possibility, both features could be independently contributing to the observed effect. To distinguish between these possibilities, we decided to plot the induction of the genes against their length. Since the two correlations were observed under serum stimulating conditions, we focused on those genes that were induced by serum in the control cells. These genes were further divided in three groups depending on their number of introns (Figure 2C). The negative effect of gene length in the capacity of a gene to undergo induction after serum addition was evident even in control cells (Figure 2C, left panel). In these same control cells, the number of introns did not seem to lower induction further. Strikingly, in the prefoldin knockdowns, there was a clear negative effect of the number of introns in gene induction (Figure 2C, Sup. Figure

S3C). Taken altogether, these results point towards a possible influence of prefoldin in transcription elongation or, at least, in those events of the gene expression process, like pre-mRNA splicing, that take place co-transcriptionally.

Splicing efficiency after serum stimulation is reduced in prefoldin knockdown cells

To further investigate the possible role of prefoldin in splicing, we looked for defects in splicing in the RNA-seq data. Since polyA selection was not used for RNA-seq library preparation, we expected to find a certain, though low, level of unspliced mRNA precursors in the data set. In order to test this, we divided the reads from the sequencing reaction in two groups that we call "exonic reads" (those present in the mature mRNA) and "gene reads" (all the reads contained within genes, including those corresponding to unspliced precursors). For each gene, we divided the number of exonic reads by that of gene reads. Thus, we defined the exon ratio of each gene in each condition. The exon ratio is used here as a proxy for the maturation state of the products of a given gene in a given sample, with a ratio of 1 meaning that all the RNAs detected are matured mRNAs, and a ratio lower than 0 meaning that unspliced precursors were detected. As expected, most of the genes in our data presented a ratio close to 1, but other values of the ratio were clearly detectable (Sup. Figure S4). In control cells, the exon ratios increased after serum stimulation (Figure 3A and Sup. Figure S4). This increase was especially frequent in highly expressed genes (Figure 3A). However, in prefoldin siRNA treated cells, the exon ratio decreased after serum stimulation and, again, this decrease was particularly frequent in highly expressed genes (Figure 3A).
In order to exclude that these results were influenced by the stressfull conditions of siRNA transfection, we validated them in PFDN5 KO cells obtained by CRISPR-CAS9 (Figure 3B). Levels of pre-mRNA for two introns of two different genes were measured by RT-PCR in control and PFDN5 KO cells. Higher levels of pre-mRNA in serum-stimulated cells were detected for all four introns in the mutant cells (Figure 3C).
To further test the implication of prefoldin in splicing, we measured alternative splicing in a transient transfection assay (minigene assay). PFDN5 KO cells showed a mild increase in exon skipping compare to control cells, and this increase was corrected by expressing an ectopic copy of PFDN5 (Sup. Figure 5A).
Exon skipping is one the possible types of alternative splicing events. We analysed different types of alternative splicing events in the PFDN5 knockdown cells RNAseq data, by using a bioinformatic approach (17). We detected significant alterations in the different types analysed, in two directions: significant decreases of some events that were detected in control cells, and new alternative splicing events that were not detected in control cells (Sup. Figure 5B). These results suggest that prefoldin depletion affected pre-mRNA splicing in a very general manner, without introducing any specific bias in exon definition.

Cotranscriptional splicing is delayed in prefoldin knockdown cells

Hitherto, we have shown evidence of human prefoldin being connected to pre-mRNA splicing and the expression of long genes. Since splicing can occur co-transcriptionally (18), the simplest explanation for this double connection of prefoldin would be that it is involved in co-transcriptional splicing. To test whether the prefodin-depleted cells were in fact defective in co-transcriptional splicing, we decided to do a DRB-mediated transcription inhibition and release experiment in siPFDN5-treated cells. DRB is an inhibitor of CDK9 that blocks transcription reversibly (19). So, when DRB is washed out, polymerases start synchronously and the first round of transcription can be followed by RT-PCR. In order to study both transcription and cotranscriptional splicing, primer pairs were designed to amplify the cDNA derived from different kinds of unspliced precursors of the gene *CTNNBL1*, a long gene previously used as a model in this type of experiments (20, 21). We measured transcription elongation rate by the appearance of unspliced precursors with primer pairs spanning exon-intron junctions (Figure 4A). We could not detect any significant difference in the times required to detect exon 5-intron 5 or exon 6-intron 6 precursors. As expected, no signal of the intron15-exon 16 precursor was detected at the same times (Figure 4A). These results ruled out a significant alteration of RNA pol II elongation rate by prefoldin depletion.
Using the same RT-PCR samples, we measured co-transcriptional splicing by the appearance of precursors containing either introns 5 or 6 but lacking the preceding one, already spliced (Figure 4B). While in control cells this kind of precursors were detectable after 50 minutes of DRB wash, in siPFDN5-treated cells we could not detect the same precursors before the end of the experiment (Figure 4B). These results indicate that prefoldin is required for normal coupling between transcription and splicing.

Prefoldin is present in the body of transcribed genes and contributes to the recruitment of the U2AF65 splicing factor to elongating RNA pol II

We then explored whether the functional involvement of prefoldin in co-transcriptional splicing correlated with its physical presence in the transcribed gene. Prefoldin has been found in the nucleus in different organisms (6, 22–25). In yeast it has also been showed to locate in the body of transcribed genes (6). To test this in human cells by ChIP, a Flag-PFDN5 stable cell line was constructed (Sup. Figure S6). We found a reproducible PFDN5 signal over the gene bodies of the genes tested, that was absent further downstream (Figure 5A).
One possible mechanism on how prefoldin might influence co-transcriptional splicing is by favouring the action of splicing factors. In order to search for such a potential functional interaction between prefoldin and pre-mRNA splicing factors we made use of TheCellMap (https://thecellmap.org/). This database contains genetic interaction data for thousands of yeast mutants. Considering that yeast prefoldin has been also previously connected to transcription elongation (6), we reasoned that yeast genetic interactions might be informative for those functions of prefolding that has been conserved during the evolution of the eukaryotic kingdom. When we systematically looked for correlations between the six prefoldin subunits and splicing factors, we obtained the results shown in Supplementary figure S7A. Most of the correlations found

pointed to a functional interaction between prefoldin and the PRP19 complex (PRP19C). Indeed, PRP19C has been reported to be recruited to transcribed chromatin through interaction with the splicing factor U2AF65 in a CTD phosphorylation-dependent manner (26). Therefore, we decided to test the recruitment of U2AF65 to transcribed genes in PFDN5 KO cells (Figure 5B). In the absence of prefoldin, U2AF65 levels in the two genes tested were reduced compared to control cells (Figure 5B). No reduction was found for elongating RNA pol II levels (Figure 5C). Moreover, total protein levels of U2AF65 and several components of PRP19C remained constant (Sup. Figure S8C-E).

In yeast cells, lack of PFDN1 causes a defect in the phosphorylation levels of the Ser2 residue of the RNA pol II largest subunit CTD (6). CTD phosphorylation in Ser2 (CTD-Ser2P) is known to favour the recruitment of some splicing factors (27) and, among them, U2AF65 (26). Thus, we reasoned if the Ser2 phosphorylation defect were conserved in humans, reduced phosphorylation levels could result in impairment of U2AF65 recruitment, which in turn would originate cotranscriptional splicing problems, and the kind of splicing defects that we observed in our RNA-seq data. Indeed, when we measured the level of CTD-Ser2P by chromatin immunoprecipitation in human PFDN5 KO cells, we observed a significant reduction of the signal (Figure 5D). In contrast, total levels of RNA pol II in chromatin did not change (Figure 5C). CDK9 is the mayor kinase phosphorylating Ser2 of the CTD (28). Lack of prefoldin did not produce significant changes in the cellular levels of CDK9 (Supplementary Figure S8A-B). Taken altogether, these results indicate that prefoldin contributes to preserve RNA pol II phosphorylation, splicing factor recruitment and co-transcriptional splicing in human cells.

Discussion

Prefoldin is well known for its cytoplasmic role in the cotranslational folding of tubulin and actin monomers (29, 30). Prefoldin has also been shown to play additional functions in the cell nucleus (4, 31), including a positive role in transcription elongation and cotranscriptional chromatin dynamics across the yeast genome (6). The results described in this work extend this view as they show a general contribution of prefoldin to human gene expression. Several results support the connection between human prefoldin and transcription elongation: a) prefoldin depletion preferentially impairs the expression of long genes; b) gene induction after serum stimulation in prefoldin-deficient cells is reduced in longer genes; c) PFDN5 is found associated to the bodies of transcribed genes (Figure 5); d) lack of prefoldin reduces RNA pol II CTD-Ser2P (Figure 5), a specific marker of transcription elongation (32).

The detailed analysis of the transcriptome of prefoldin-depleted human cells detected significant alterations of pre-mRNA splicing, lower levels of expression in genes with a high number of introns, and impaired induction of those serum-activated genes that contain a higher number of introns. Moreover, gene length and number of introns synergistically impaired gene induction by serum in prefoldin-depleted cells. These data points towards a positive influence of prefoldin in co-transcriptional pre-mRNA

splicing. Our experiments analysing DRB-synchronized populations of RNA pol II molecules confirmed that prefoldin-depleted cells were defective in co-transcriptional splicing.

Our results suggest that prefoldin influences co-transcriptional pre-mRNA splicing by favouring the recruitment of the splicing machinery during transcription elongation. Pre-mRNA splicing is a sequential process that consists of spliceosome assembly, pre-catalytic activation and catalysis (33). It is known that pre-mRNA splicing is coupled to transcription elongation (34, 35). We found that genetic interactions point to a closer correlation of prefoldin subunits with the PRP19 complex, a splicing factor involved in the pre-catalytic activation step (36), than with the snRNP components of the spliceosome (Supplementary Figure S7). Similar level of correlation was also found between several prefoldin subunits and the splicing factor PRP2 (Supplementary Figure S7), which also acts during remodelling of the spliceosomal B^{act} complex to the catalytically activated B* complex, just before step one of splicing (37). The action of PRP19C is mediated by U2AF65, which helps recruit this complex to the phosphorylated RNA pol II CTD domain (26). This mechanism explains coupling of transcription elongation and pre-mRNA splicing activation (26). We found that lack of prefoldin provokes a decrease in RNA pol II CTD-Ser2P and in the levels of U2AF65 present in the body of transcribed genes. We, therefore, propose a scenario where the positive contribution of prefoldin to Ser2 phosphorylation is necessary for the recruitment of PRP19C/U2AF65, which in turn allows efficient co-transcriptional splicing (26). We cannot exclude that prefoldin, by favouring the recruitment of U2AF65 to elongating RNA pol II, is also contributing to transcription elongation itself. In fact, Mud2, the yeast U2AF65 homolog, favours both transcription elongation and pre-mRNA splicing (38).

Another, non-mutually exclusive possibility is that canonical prefoldin helps the assembly of the spliceosome. The co-chaperone function of prefoldin could be necessary for this. Other large complexes are assisted by the Uri/prefoldin-like complex for their assembly (5, 39, 40). Moreover, the prefoldin-like component RUVBL1/RUVBL2 regulates assembly of the U5 small nuclear ribonucleoprotein (39), and plant canonical prefoldin is involved in maintaining the levels of the LSM2-8 complex through Hsp90 (41). However, this potential role of human prefoldin cannot explain all the results presented here, since an assembly failure would affect spliceosome activity in general, not only co-transcriptional splicing.

Prefoldin is a stable heterohexameric complex (15). However, there is also evidence that some of the individual subunits have a function by themselves (42). Human PFDN5, for instance, is able to inhibit the transactivation domain of c-Myc by recruiting histone deacetylases (9) and promoting c-myc degradation (43). This interaction of c-Myc and PFDN5 is subunit-specific and does not involve other prefoldin subunits (9). As we detected clear parallelisms between PFDN2 and PFDN5 in the analyses that we performed, we conclude that the effects on transcription elongation and pre-mRNA splicing that we describe herein likely reflects the function of the canonical prefoldin complex.

It is known that coupling to transcription is not essential for splicing to occur (44). This can explain why we did not detect a possible contribution of prefoldin to pre-mRNA splicing under serum starvation conditions. We found that, in control cells, serum stimulation provoked increased efficiency of splicing, reflected in higher exon/intron ratios in total RNA reads. The transcriptional spur caused by serum stimulation might explain this phenomenon through co-transcriptional splicing. If that were the case, a positive contribution of prefoldin to splicing would be only expected under serum conditions, as we found.

Human prefoldin is involved in some cellular regulations like the epithelial-mesenchymal transition (EMT). PFDN1 has been described to mediate EMT in lung cancer cell lines and metastasis in mouse models (10). We have recently confirmed that overexpression of prefoldin subunits associates with the risk of mortality and metastasis in non-small cell lung cancer (45). The current model suggests that prefoldin overexpression would alter the expression pattern of EMT regulatory genes (10). The results of this study may contribute to elucidate how prefoldin can play such a regulatory function.

Acknowledgements

We thank all the people of the IBiS gene expression lab for helpful discussion, Victoria Begley for English editing, and Cindy Will and Reinhard Lührmann for anti PRPC19 antibodies.

Funding

This work has been supported by grants from the Spanish Ministry of Economy and Competitiveness [BFU2016-77728-C3-1-P to S.C. and BFU2017-85420-R to J.C.R.], the Andalusian Government [P12-BIO1938MO, BIO271 and US-1256285 to S.C. and BIO321 to J.C.R.] and European Union funds (FEDER) and by FPI grant from Junta de Andalucía to L.P.-B.

Conflict of interest

The authors declare not to have any conflict of interest.

References

1. R. Arranz, J. Martín-Benito, J. M. Valpuesta, “Structure and function of the cochaperone Prefoldin” in *Prefoldins: The New Chaperones*, (2018), pp. 119–131.
2. S. Chávez, P. Puerto-Camacho, “Prefoldins” in *ELS*, (John Wiley & Sons, Ltd, 2016), pp. 1–8.
3. V. F. Lundin, M. R. Leroux, P. C. Stirling, Quality control of cytoskeletal proteins and human disease. *Trends Biochem Sci* **35**, 288–297 (2010).
4. G. Millán-Zambrano, S. Chávez, Nuclear functions of prefoldin. *Open biol* **4**, 630–642 (2014).

5. S. Boulon, E. Bertrand, B. Pradet-Balade, HSP90 and the R2TP co-chaperone complex: Building multi-protein machineries essential for cell growth and gene expression. *RNA Biol* **9**, 148–154 (2012).
6. G. Millán-Zambrano, *et al.*, The prefoldin complex regulates chromatin dynamics during transcription elongation. *PLoS Genet* **9**, e1003776 (2013).
7. A. Mousnier, *et al.*, von Hippel Lindau binding protein 1-mediated degradation of integrase affects HIV-1 gene expression at a postintegration step. *PNAS* **104**, 13615–20 (2007).
8. K. Mori, *et al.*, MM-1, a Novel c-Myc-associating Protein That Represses Transcriptional Activity of c-Myc. *J Biol Chem* **273**, 29794–29800 (1998).
9. A. Satou, T. Taira, S. M. M. Iguchi-Ariga, H. Ariga, A novel transrepression pathway of c-Myc: Recruitment of a transcriptional corepressor complex to c-Myc by MM-1, a c-Myc-binding protein. *J Biol Chem* **276**, 46562–46567 (2001).
10. D. Wang, *et al.*, Prefoldin 1 promotes EMT and lung cancer progression by suppressing cyclin A expression. *Oncogene* **36**, 885–898 (2017).
11. C. I. Castillo-Davis, S. L. Mekhedov, D. L. Hartl, E. V. Koonin, F. A. Kondrashov, Selection for short introns in highly expressed genes. *Nat Genet* **31**, 415–8 (2002).
12. F. Chiaromonte, W. Miller, E. E. Bouhassira, Gene length and proximity to neighbors affect genome-wide expression levels. *Genome Res* **13**, 2602–2608 (2003).
13. S. Becerra, M. Montes, C. Hernández-Munain, C. Suñé, Prp40 pre-mRNA processing factor 40 homolog B (PRPF40B) associates with SF1 and U2AF 65 and modulates alternative pre-mRNA splicing in vivo. *RNA* **21**, 438–457 (2015).
14. V. R. Iyer, *et al.*, The transcriptional program in the response of human fibroblasts to serum. *Science* **283**, 83–7 (1999).
15. M. Miyazawa, *et al.*, Prefoldin subunits are protected from ubiquitin-proteasome system-mediated degradation by forming complex with other constituent subunits. *Int J Biol Chem* **286**, 19191–19203 (2011).
16. R. Regulapati, A. Bhasi, C. K. Singh, P. Senapathy, Origination of the Split Structure of Spliceosomal Genes from Random Genetic Sequences. *PLoS ONE* **3**, e3456 (2008).
17. G. P. Alamancos, A. Pagès, J. L. Trincado, N. Bellora, E. Eyras, Leveraging transcript quantification for fast computation of alternative splicing profiles. *RNA* **21**, 1521–1531 (2015).
18. T. Nojima, *et al.*, RNA Polymerase II Phosphorylated on CTD Serine 5 Interacts with the Spliceosome during Co-transcriptional Splicing. *Mol Cell* **72**, 369–379.e4 (2018).
19. R. M. Sheridan, N. Fong, A. D'Alessandro, D. L. Bentley, Widespread backtracking by RNA Pol II is a major effector of gene activation, 5' pause release, termination, and transcription elongation rate. *Mol Cell* **73**, 107–118.e4 (2019).
20. J. Singh, R. A. Padgett, Rates of in situ transcription and splicing in large human genes. *Nat Struct Mol Biol* **16**, 1128–1133 (2009).
21. S. Jimeno-González, *et al.*, Defective histone supply causes changes in RNA polymerase II elongation rate and cotranscriptional pre-mRNA splicing. *PNAS* **112**, 14840–14845 (2015).
22. Y. Hagio, *et al.*, Distinct localizations and repression activities of MM-1 isoforms toward c-Myc. *J Cell Biochem* **97**, 145–155 (2006).
23. A. Abe, *et al.*, Prefoldin plays a role as a clearance factor in preventing

proteasome inhibitor-induced protein aggregation. *J Biol Chem* **288**, 27764–76 (2013).
24. A. Locascio, M. A. Blázquez, D. Alabadí, Dynamic Regulation of Cortical Microtubule Organization through Prefoldin-DELLA Interaction. *Curr Biol* **23**, 804–809 (2013).
25. H. Tsuchiya, T. Iseda, O. Hino, Identification of a Novel Protein (VBP-1) Binding to the von Hippel-Lindau (VHL) Tumor Suppressor Gene Product. *Cancer Res* **56** (1996).
26. C. J. David, A. R. Boyne, S. R. Millhouse, J. L. Manley, The RNA polymerase II C-terminal domain promotes splicing activation through recruitment of a U2AF65-Prp19 complex. *Genes Dev* **25**, 972–983 (2011).
27. B. Gu, D. Eick, O. Bensaude, CTD serine-2 plays a critical role in splicing and termination factor recruitment to RNA polymerase II in vivo. *Nucleic Acids Res* **41**, 1591–1603 (2013).
28. C. W. Bacon, I. D'Orso, CDK9: a signaling hub for transcriptional control. *Transcription* **10**, 57–75 (2019).
29. R. Arranz, J. Martín-Benito, J. M. Valpuesta, "Structure and function of the cochaperone prefoldin" in *Prefoldins: The New Chaperones*, (2018), pp. 119–131.
30. D. Gestaut, *et al.*, The Chaperonin TRiC/CCT Associates with Prefoldin through a Conserved Electrostatic Interface Essential for Cellular Proteostasis. *Cell* **177**, 751–765.e15 (2019).
31. L. Payán-Bravo, X. Peñate, S. Chávez, "Functional Contributions of Prefoldin to Gene Expression" in *Prefoldins: The New Chaperones*, (2018), pp. 1–10.
32. S. Egloff, S. Murphy, Cracking the RNA polymerase II CTD code. *Trends in Genet* **24**, 280–288 (2008).
33. C. Yan, R. Wan, Y. Shi, Molecular Mechanisms of pre-mRNA Splicing through Structural Biology of the Spliceosome. *Cold Spring Harb Perspect Biol* **11**, a032409 (2019).
34. M. Montes, S. Becerra, M. Sánchez-Álvarez, C. Suñé, Functional coupling of transcription and splicing. *Gene* **501**, 104–117 (2012).
35. L. Herzel, D. S. M. Ottoz, T. Alpert, K. M. Neugebauer, Splicing and transcription touch base: co-transcriptional spliceosome assembly and function. *Nat Rev Mol Cell Biol* **18**, 637–650 (2017).
36. S.-P. Chan, The Prp19p-Associated Complex in Spliceosome Activation. *Science* **302**, 279–282 (2003).
37. P. Bao, C. Höbartner, K. Hartmuth, R. Lührmann, Yeast Prp2 liberates the 5′ splice site and the branch site adenosine for catalysis of pre-mRNA splicing. *RNA* **23**, 1770–1779 (2017).
38. R. Minocha, *et al.*, Mud2 functions in transcription by recruiting the Prp19 and TREX complexes to transcribed genes. *Nucleic Acids Res* **46**, 9749–9763 (2018).
39. P. Cloutier, *et al.*, R2TP/Prefoldin-like component RUVBL1/RUVBL2 directly interacts with ZNHIT2 to regulate assembly of U5 small nuclear ribonucleoprotein. *Nat Commun* **8**, 15615 (2017).
40. N. Izumi, *et al.*, AAA+ Proteins RUVBL1 and RUVBL2 Coordinate PIKK Activity and Function in Nonsense-Mediated mRNA Decay. *Sci Signaling* **3**, ra27-ra27 (2010).
41. D. Esteve-Bruna, *et al.*, Prefoldins contribute to maintaining the levels of the spliceosome LSM2–8 complex through Hsp90 in Arabidopsis. *Nucleic Acids Res* **48**, 6280–6293 (2020).

42. A. F. Amorim, D. Pinto, L. Kuras, L. Fernandes, Absence of Gim proteins, but not GimC complex, alters stress-induced transcription. *BBA-Gene Regul Mech* **1860**, 773–781 (2017).
43. Y. Kimura, *et al.*, MM-1 facilitates degradation of c-Myc by recruiting proteasome and a novel ubiquitin E3 ligase. *Int J Oncol* **31**, 829–836 (2007).
44. J. Han, J. Xiong, D. Wang, X.-D. Fu, Pre-mRNA splicing: where and when in the nucleus. *Trends Cell Biol* **21**, 336–343 (2011).
45. X. Peñate, *et al.*, Overexpression of canonical prefoldin associates with the risk of mortality and metastasis in non-small cell lung cancer. *Cancers* **12** (2020).

Figure legends

Figure 1: Depletion of PFDN5 and PFDN2 alter gene expression in human cells. HCT166 cells were treated with the different siRNAs for 24 hours, then serum-starved for 48 hours, then gene expression was induced by adding serum, and samples were taken just before and 90 minutes after stimulation. **A)** MA plots of RNA-seq experiments. The logarithm of the fold change (log_2FC (siControl versus siPFDN)) is shown against the logarithm of the level of gene expression (defined for each gene as the mean of the counts per million of all the samples). The graphs on the left show the transcriptomic changes of the samples transfected with the siPFDN2 with respect to the control, and those on the right, those of the samples transfected with the siPFDN5. Unaffected genes are represented in black, while in red are all those genes showing $|log_2FC| > 0$, FRD < 0.05. The upper panels represent the data of the samples taken before serum stimulation, while the lower panels represent those taken after serum stimulation. **B)** Representation of the number of genes induced by serum in control cells (log_2FC (t=90 versus t=0 minutes of serum treatment) > 1, FDR < 0.05), and those whose stimulation was significantly affected by siPFDN2 or siPFDN5 (FDR>0.05). **C)** Overlap between those genes whose induction by serum was impaired by PFDN2 and PFDN5 depletion. The p value after a hypergeometric test is also shown.

Figure 2. Prefoldin deficiency impairs the induction of long genes with a high number of introns. A) Expression differences between siPFDN5 and siControl cells (log_2 FC) with respect to the gene length, before and after serum stimulation. The genes were separated into quintiles. The * represents p<0.005 in a Student´s t test comparing each quintile to quintile 1. **B)** The same analysis of A with respect to the number of introns that each gene contains, before and after serum stimulation. The genes were separated into quintiles. The * represents p<0.005 in a Student´s t test comparing each quintile to quintile 1. **C)** Serum regulated genes ($|log_2FC| > 0$ and FDR < 0.05, N=1035) were divided into three groups according to the tercile of the number of introns. Serum-dependent FC (Log_2) is represented against the gene length (Log_{10}), in cells transfected with the siControl (left panel) or siPFDN5 (right panel).

Figure 3. Prefoldin deficiency reduces splicing efficiency under serum stimulation conditions. A) The change of the exon ratio (exonic reads / total reads) after serum stimulation is represented. The genes were divided into three groups according to their

expression level (high, medium and low levels). The expression level of each gene was defined as the mean number of reads, per million of all the samples, after normalizing by gene length. **B)** Western blot showing the absence of PFDN5 protein in the PFDN5 KO cell line. **C)** Control and PFDN5 KO cells were treated as in Figure 1 and the level of pre-mRNA in two different regions of the *SRRM2* and *FASN* genes was determined by amplifying an intron-exon junction. The graphs represent the average and the standard deviation obtained from three different biological replicates. Pre-mRNA data was normalized first to mature mRNA level, and then to time zero.

Figure 4. PFDN5 depletion impairs cotranscriptional splicing. A) Pre-mRNA levels at different loci throughout the *CTNNBL1* gene in cells transfected with control siRNA (red line) or PFDN5 siRNA (blue line), and treated with DRB for 3 hours. Samples were taken every 10 minutes after washing DRB out. Pre-mRNA data was normalized first to the mature mRNA, and next to the DRB untreated sample. A scheme of the *CTNNBL1* gene is also shown, and the different amplicons used are depicted above each graph. **B)** Events of cotranscriptional splicing (measured as the presence of pre-mRNA without the indicated intron) at different points throughout the *CTNNBL1* gene in cells treated as in A). The different amplicons used are depicted above each graph.

Figure 5. Lack of prefoldin decreases U2AF65 recruitment to transcribed chromatin and Ser2 phosphorylation of elongating RNA pol II. A) The levels of Flag-PFDN5 were measured by ChIP in different regions of the *CTNNBL1* (upper panels) and *CD44* (lower panels) gene using anti-Flag antibody in control (light pink) and Flag-PFDN5 cells (dark pink). An intergenic region in chromosome 5 was used as the non-transcribed control. **B)** The levels of U2AF65 were measured by ChIP in different regions of the *CTNNBL1*(upper panels) and *CD44* (lower panels) gene, using anti-U2AF65 antibody in control (red bars) and PFDN5 KO cells (blue bars). **C)** The levels of RNA pol II were measured by ChIP in different regions of the *CTNNBL1* (upper panels) and *CD44* (lower panels) gene, using the 8WG16 antibody in control (red bars) and PFDN5 KO cells (blue bars). **D)** The levels of Ser2-phosphorylated RNA pol II were measured in different regions of the *CTNNBL1* (upper panels) and *CD44* (lower panels) gene using an anti-Ser2P-CTD specific antibody, in control (red bars) and PFDN5 KO cells (blue bars). RNA pol II CTD-Ser2P values were normalized to total RNA pol II levels. Average values and standard deviation are represented in all graphs. Four biological replicates were performed. * P value <0.05; ** <0.005; *** <0.0005. A total number of 12 technical replicates were considered for Student´s t test.

Supplementary figure 1. Efficiency of the siRNA-mediated depletion of PFDN2 and PFDN5. A) Lefts panels: *PFDN2* and *PFDN5* mRNA levels in HCT116 cells, 24, 48 and 72 hours after being transfected with siPFDN2 or siPFDN5, respectively, or with a siRNA control (siControl). Rights panels: Protein levels analysed by Western blotting, using antibodies against PFDN2 or PFDN5 and GAPDH as a loading control, after transfecting the cells with siPFDN2, siPFDN5, respectively, or with siControl. The

results were quantified with *Image Lab* software. **B)** Cell viability assay on the same cell line transfected with siPFDN2 or siPFDN5 and the siControl separately. The cells were counted every 24 hours for 3 days, and the percentage of living cells relative to the control was represented. **A, B)** Averaged values ± the standard deviation obtained from three different experiments are represented. * Student´s t test with P value <0.05 and ** Student´s t test with P value <0.005.

Supplementary figure 2. Differential expression of PFDN2- and PFDN5-deficient cells. A) Representation of the number of genes differentially expressed in the cells transfected with siPFDN2 or with siPFDN5 at each time, before and after serum stimulation. **B)** The genes affected by siPFDN2 or siPFDN5 after serum stimulation were functionally grouped into Gene Ontology categories. Those categories that were commonly affected by both siRNAs are represented in the table. **C)** Representation of the number of genes differentially expressed in the cells transfected with siPFDN2 or with siPFDN5 before serum stimulation, and the genes that coincide. *P* value of the hypergeometric test is shown. n indicates total number of genes included in the analysis. **D)** Paralelism of the expression changes produced by siPFDN2 and siPFDN5. Changes in mRNA levels of genes affected by both siPFDN2 and siPFDN5 before serum stimulation are shown.

Supplementary figure 3. PFDN2 deficiency impairs the induction of long genes with a high number of introns. A) Expression differences between siPFDN2 and siControl cells ($\log_2$ FC) with respect to the gene length, before and after serum stimulation. The genes were separated into quintiles. The * represents p<0.005 in a Student´s t test comparing each quintile to quintile 1. **B)** The same analysis of A with respect to the number of introns that each gene contains, before and after serum stimulation. The genes were separated into quintiles. The * represents p<0.005 in a Student´s t test comparing each quintile to quintile 1. **C)** Serum regulated genes ($|\log_2 FC| > 0$ and FDR < 0.05, N=1035) were divided into three groups according to the tercile of the number of introns. Serum-dependent FC (Log_2) is represented against the gene length (Log_{10}), in cells transfected with the siControl (left panel) or siPFDN2 (right panel).

Supplementary figure 4. Exonic ratio distribution before and after serum stimulation. The exonic ratio (vertical axis) is represented against the level of gene expression (horizontal axis) of the cells transfected with the siControl (red), siPFDN2 (green) and siFDN5 (blue) before and after serum stimulation.

Supplementary figure 5. PFDN5 deficiency has a mild general impact on alternative splicing. A) Lack of PFDN5 causes accumulation of intron reads after serum stimulation in a *FASN* minigene assay. **A** schematic drawing of the alternative isoforms of the *Fas* minigen.is shown on the left. On the right, bar histogram of short isoform / long isoform ratios of the *Fas* minigen, in control cells, PFDN5 KO cells, and PFDN5 KO cells transfected with GFP-PFDN5. The graph represents the average

values ± the standard deviation obtained from five different experiments. **B)** PFDN5 depletion causes general alteration in alternative splicing. SUPPA computational tool (Alamancos et al, 2015) was used for exploring the impact of siPFDN5 on alternative splicing. Bars histogram representing the percentage of annotated alternative splicing events that occur significantly less frequently in PFDN5-deficient cells is shown in the left panel. On the right, bars histogram representing the frequency of alternative splicing events that appear in PFDN5-deficient cells (at least in 10 % of the reads) and are not present in control cells. SE: skipping exon; MX: mutually exclusive exons; A5: alternative 5' splice-site; A3: alternative 3' splice-site; RI: retained intron; AF: alternative first exon; AL: alternative last exon.

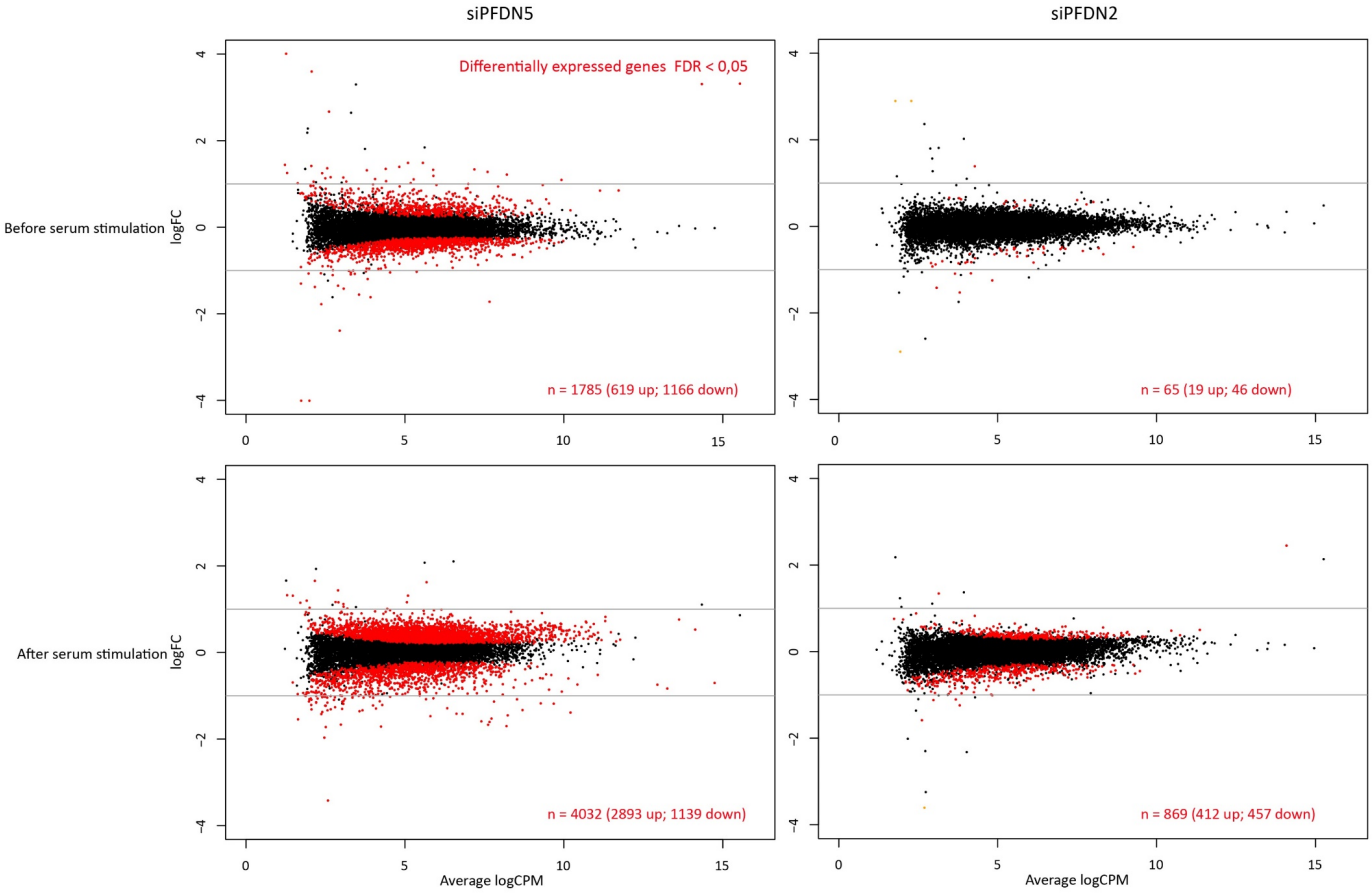
Supplementary figure 6. HCT116 cell line expressing Flag-PFDN5. Expression of Flag-PFDN5 in HCT116 stable cell line. A western blot is shown, after using antibodies against PFDN5 and GAPDH as a loading control.

Supplementary figure 7. Correlations between prefoldin and splicing mutants in the Cell Map database of genetic interactions. A) Genetic correlations between prefoldin and splicing mutants were explored in the Cell Map database (thecellmap.org). Pairs of prefoldin/pre-mRNA splicing mutants showing Pearson correlations coefficients higher than 0.13 are shown in the table. **B)** The two factors that showed general correlations with prefoldin mutants, PRP19C and PRP2, participate in the activation step of pre-mRNA splicing. Spliceosome assembly and splicing reaction. snRNPs subunits U1, U2, U4, U5 and U6 are represented; as well as the splicing factors Prp2 and Prp19C.

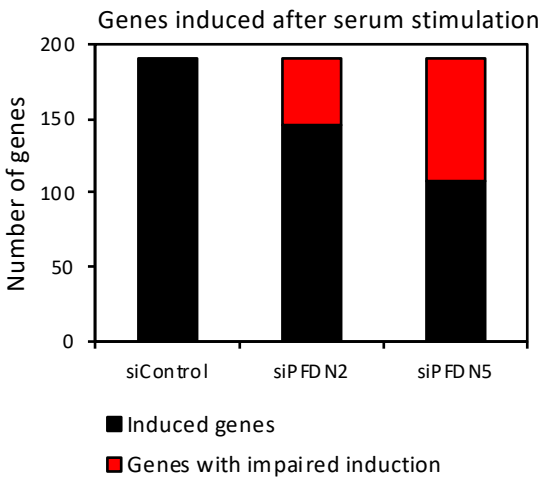
Supplementary figure 8. Total levels of CDK9 and different splicing factors remain constant in the absence of PFDN5. A) CDK9 mRNA levels of cells treated as in Figure 1A, measured by RT-PCR. **B)** CDK9 protein levels were determined by western blot in control and PFDN5 KO cells. GAPDH was used as a loading control. **C-E)** Protein levels determined by western blotting, using antibodies against PRPF19 (C), PRGL1 (D) and U2AF65 (E) splicing factors and GAPDH as a loading control. The results were quantified with *Image Lab* software. **A-E)** Averaged values ± the standard deviation obtained from three different experiments are represented.

Figure 1

A


siPFDN5
siPFDN2
Differentially expressed genes FDR < 0,05
Before serum stimulation
After serum stimulation
logFC
n = 1785 (619 up; 1166 down)
n = 65 (19 up; 46 down)
n = 4032 (2893 up; 1139 down)
n = 869 (412 up; 457 down)
Average logCPM

B


Genes induced after serum stimulation
Number of genes
siControl
siPFDN2
siPFDN5
Induced genes
Genes with impaired induction

C

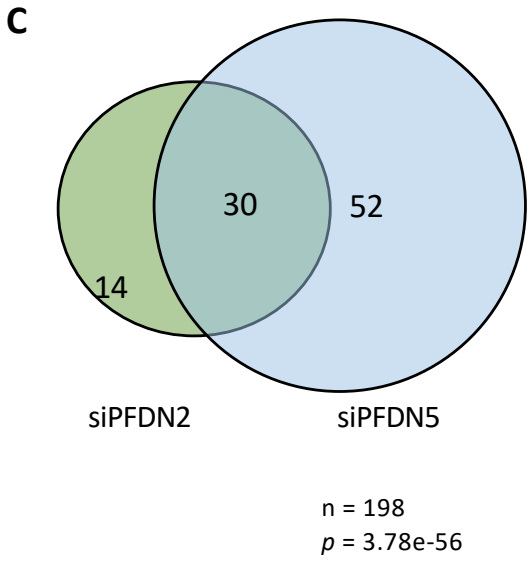

30
52
14
siPFDN2
siPFDN5
n = 198
p = 3.78e-56

Figure 2

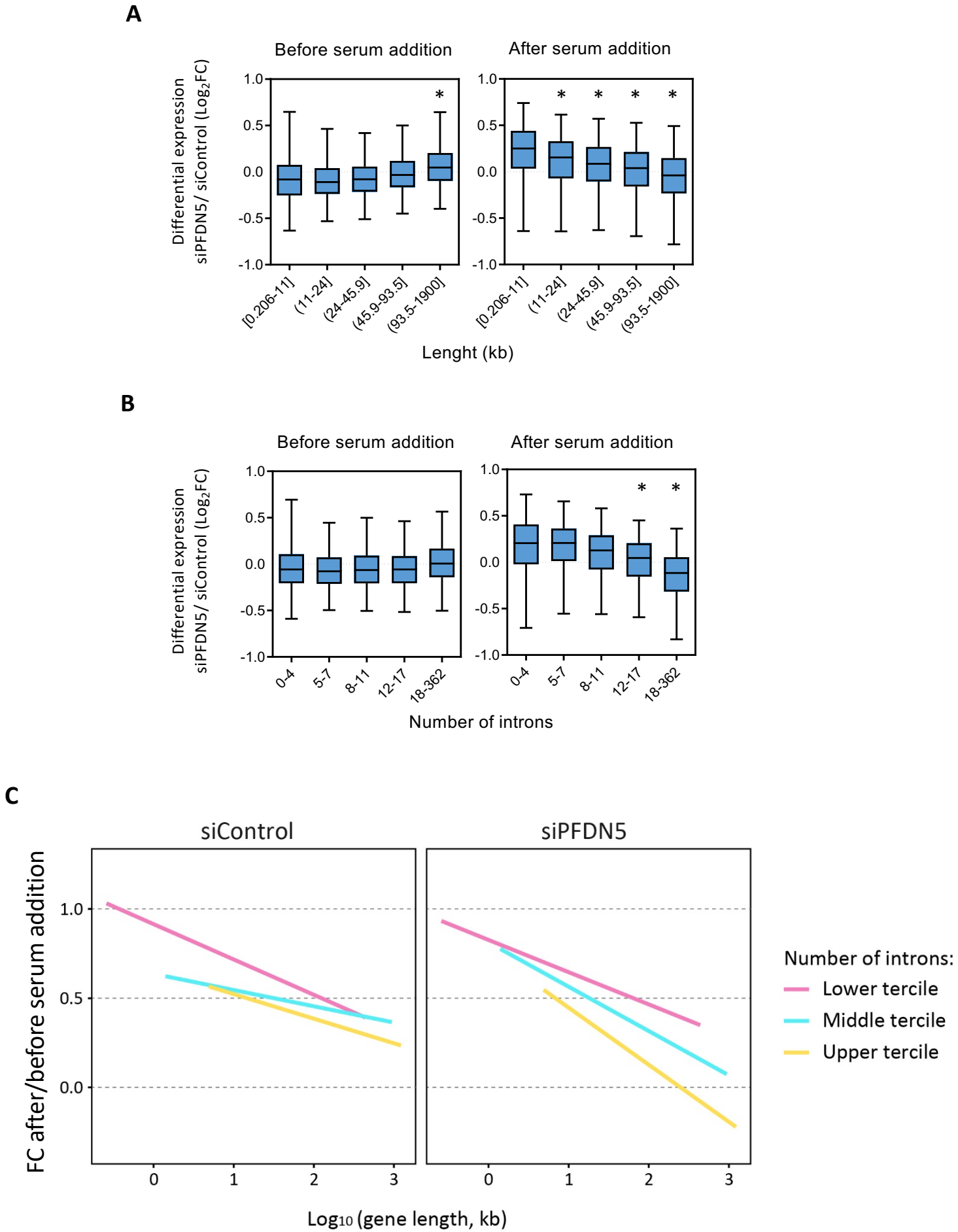
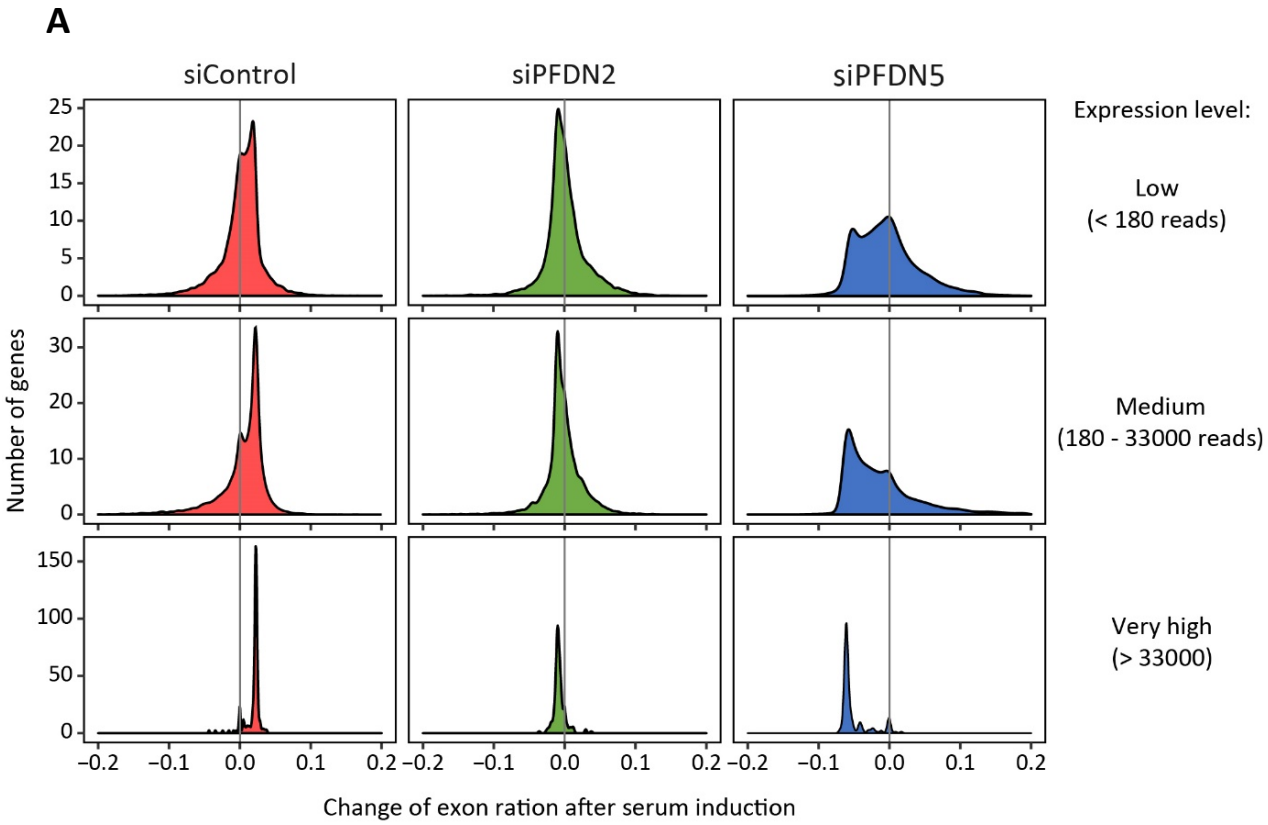
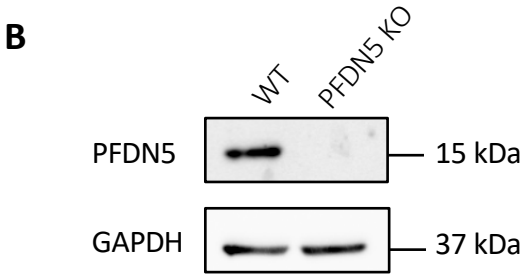

A
Before serum addition
After serum addition
Differential expression
siPFDN5/ siControl (Log2FC)
1.0
0.5
0.0
-0.5
-1.0
*
[0.206-11]
(11-24]
(24-45.9]
(45.9-93.5]
(93.5-1900]
Lenght (kb)
B
Before serum addition
After serum addition
Differential expression
siPFDN5/ siControl (Log2FC)
0-4
5-7
8-11
12-17
18-362
Number of introns
C
siControl
siPFDN5
FC after/before serum addition
1.0
0.5
0.0
0
1
2
3
Log10 (gene length, kb)
Number of introns:
Lower tercile
Middle tercile
Upper tercile

Figure 3


A
siControl
siPFDN2
siPFDN5
Expression level:
Low
(< 180 reads)
Medium
(180 - 33000 reads)
Very high
(> 33000)
Number of genes
Change of exon ration after serum induction


B
WT
PFDN5 KO
PFDN5
15 kDa
GAPDH
37 kDa

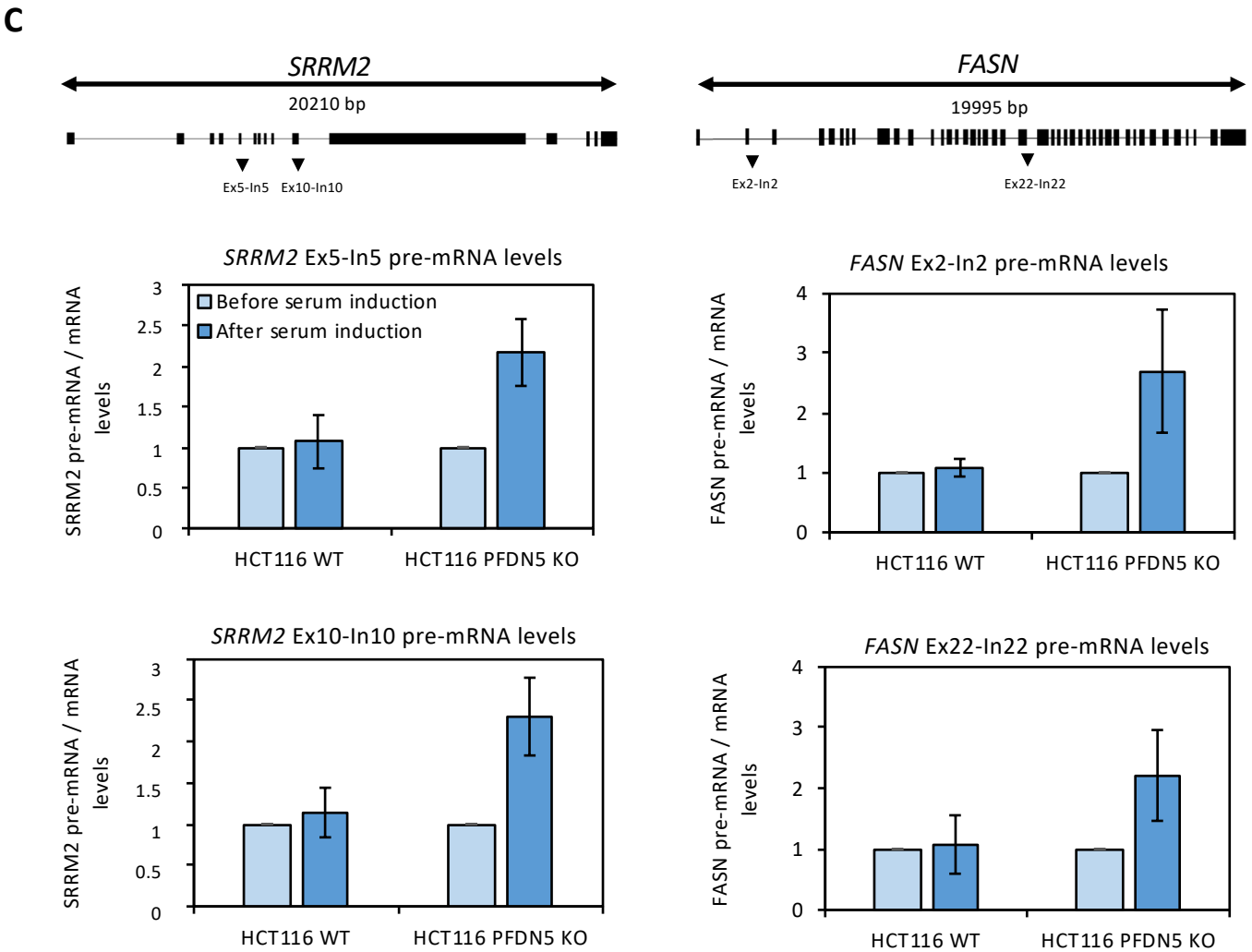

C
SRRM2
20210 bp
Ex5-In5
Ex10-In10
FASN
19995 bp
Ex2-In2
Ex22-In22
SRRM2 Ex5-In5 pre-mRNA levels
Before serum induction
After serum induction
SRRM2 pre-mRNA / mRNA levels
HCT116 WT
HCT116 PFDN5 KO
FASN Ex2-In2 pre-mRNA levels
FASN pre-mRNA / mRNA levels
SRRM2 Ex10-In10 pre-mRNA levels
FASN Ex22-In22 pre-mRNA levels

Figure 4

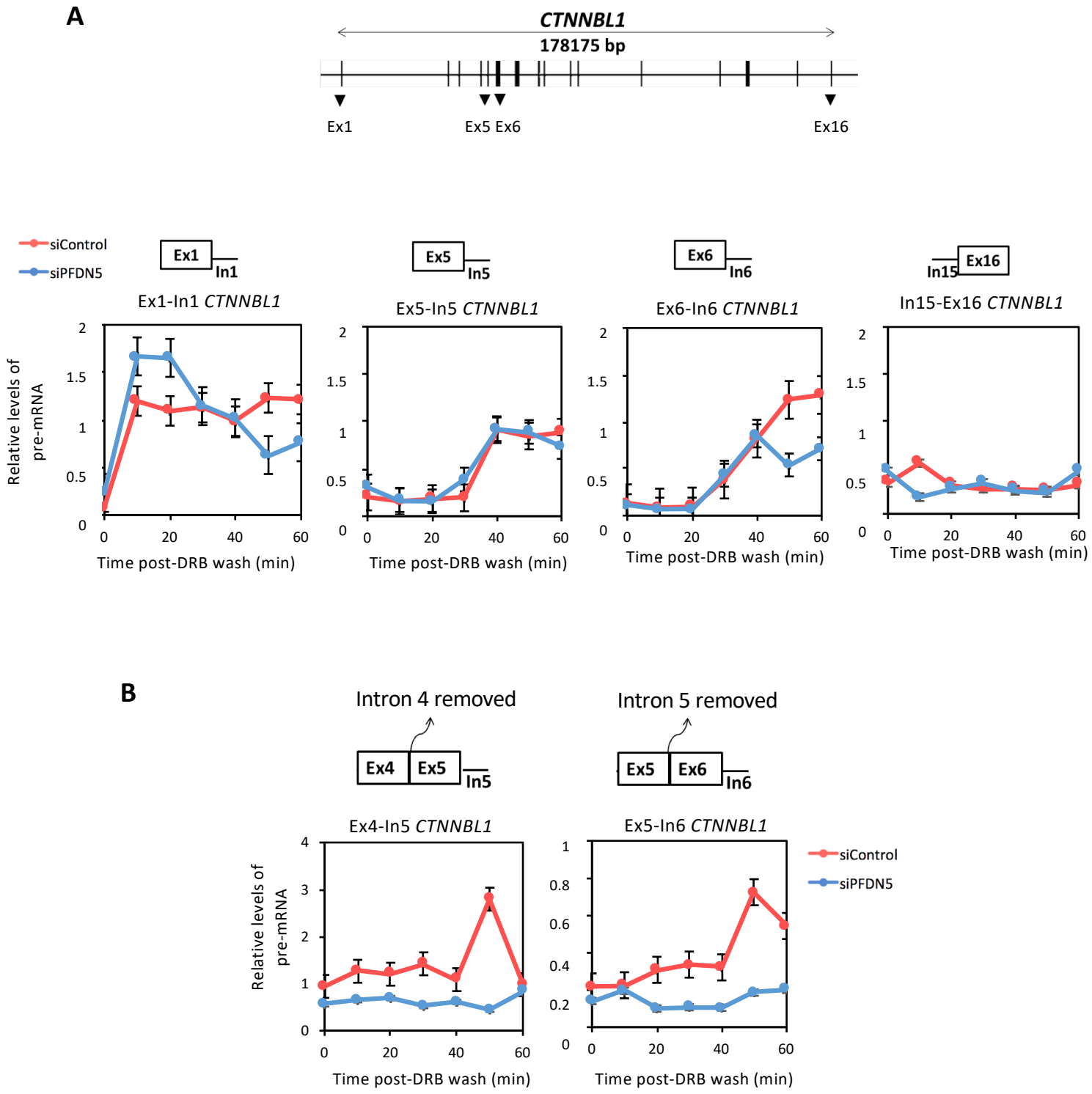
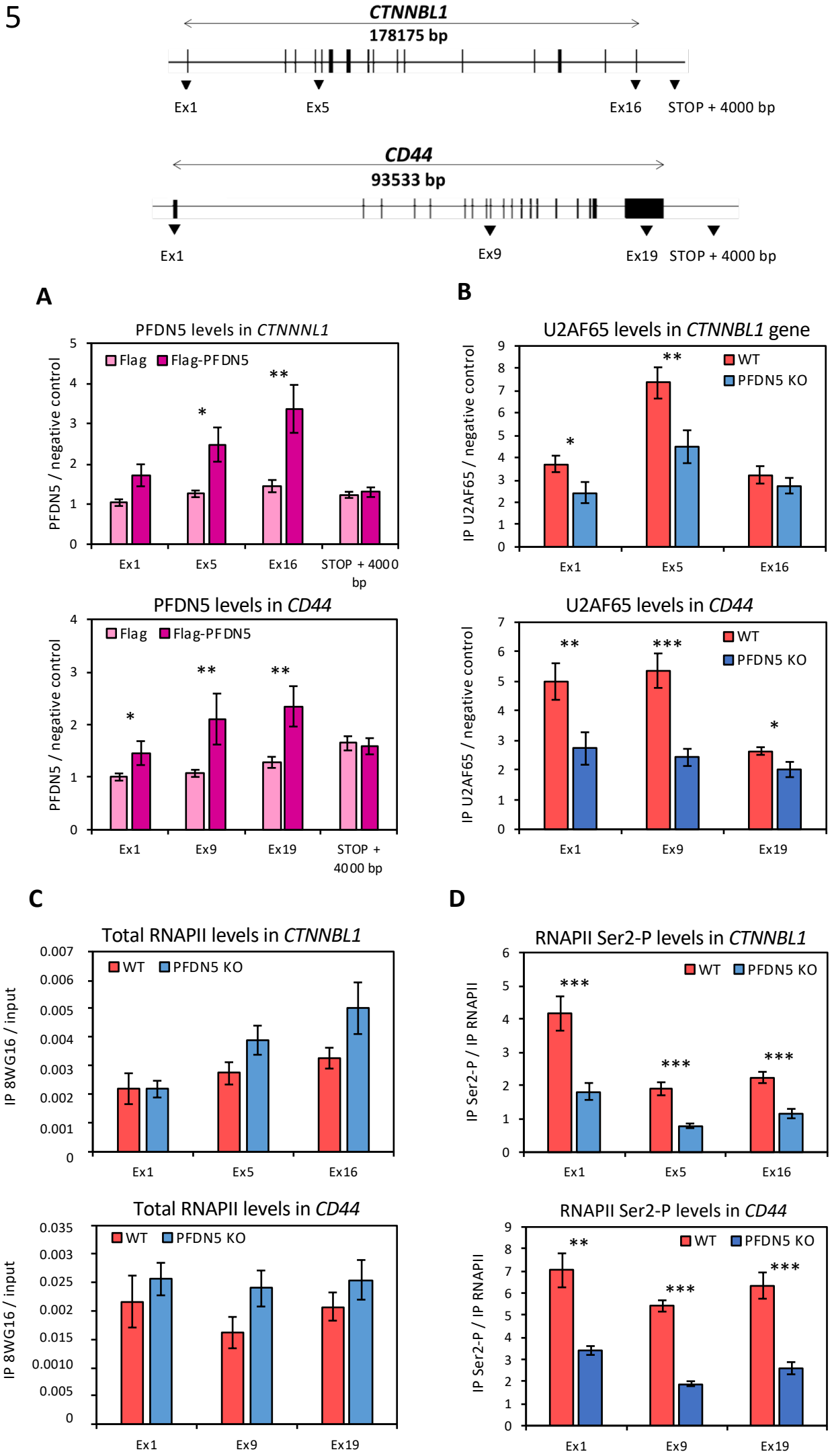
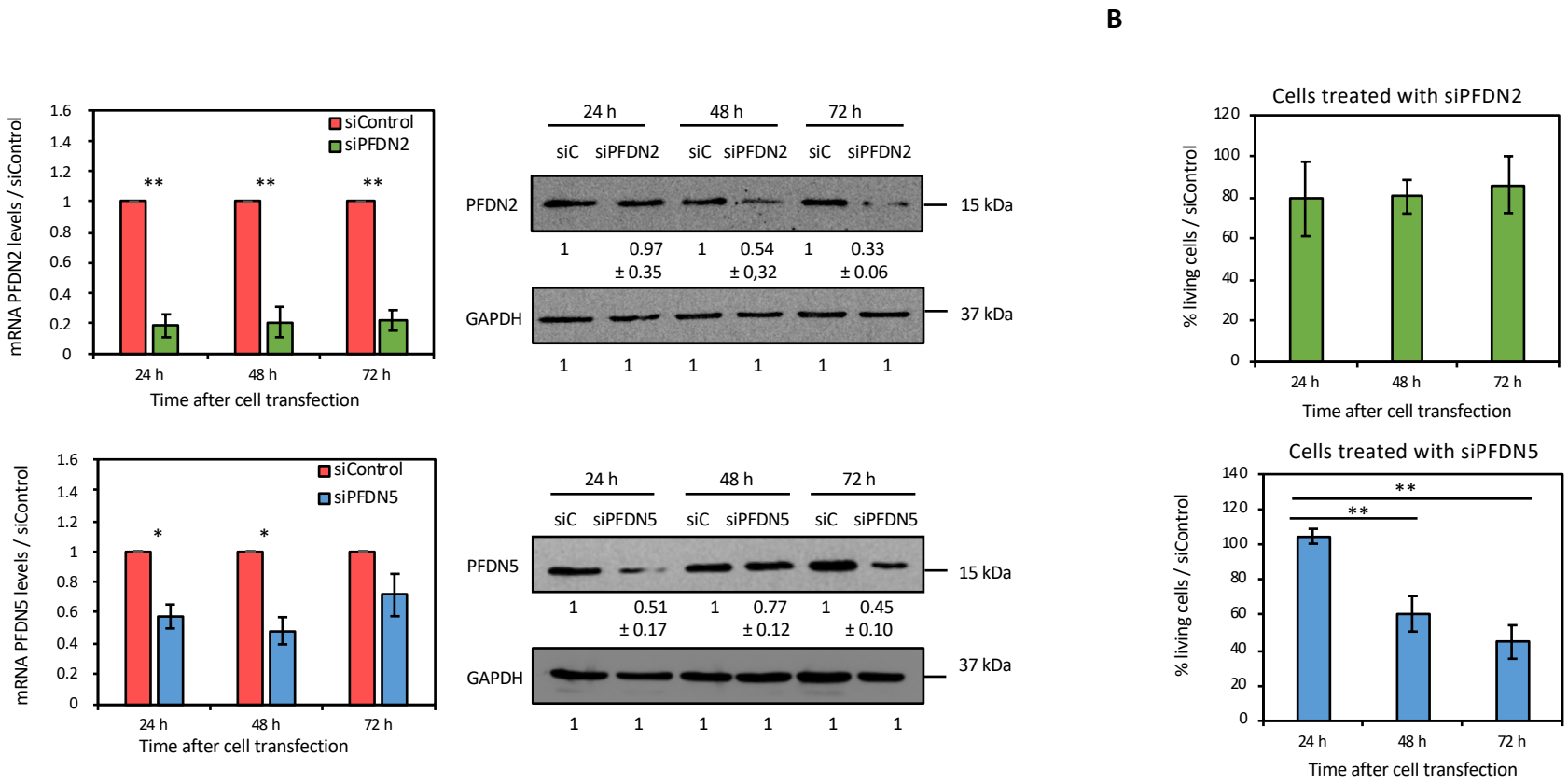

A
CTNNBL1
178175 bp
Ex1
Ex5 Ex6
Ex16
siControl
siPFDN5
Ex1
In1
Ex1-In1 CTNNBL1
Ex5
In5
Ex5-In5 CTNNBL1
Ex6
In6
Ex6-In6 CTNNBL1
In15
Ex16
In15-Ex16 CTNNBL1
Relative levels of pre-mRNA
Time post-DRB wash (min)
B
Intron 4 removed
Ex4
Ex5
In5
Ex4-In5 CTNNBL1
Intron 5 removed
Ex5
Ex6
In6
Ex5-In6 CTNNBL1
siControl
siPFDN5
Relative levels of pre-mRNA
Time post-DRB wash (min)

Figure 5

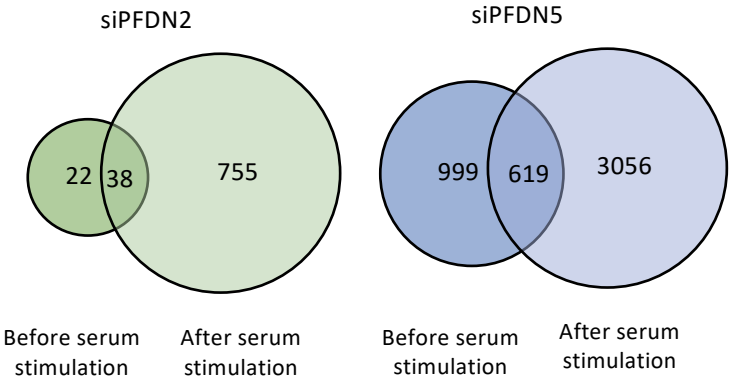

CTNNBL1
178175 bp
Ex1
Ex5
Ex16
STOP + 4000 bp
CD44
93533 bp
Ex1
Ex9
Ex19
STOP + 4000 bp
A
PFDN5 levels in CTNNNL1
PFDN5 / negative control
Flag
Flag-PFDN5
Ex1
Ex5
Ex16
STOP + 4000 bp
PFDN5 levels in CD44
PFDN5 / negative control
Flag
Flag-PFDN5
Ex1
Ex9
Ex19
STOP + 4000 bp
B
U2AF65 levels in CTNNBL1 gene
IP U2AF65 / negative control
WT
PFDN5 KO
Ex1
Ex5
Ex16
U2AF65 levels in CD44
IP U2AF65 / negative control
WT
PFDN5 KO
Ex1
Ex9
Ex19
C
Total RNAPII levels in CTNNBL1
IP 8WG16 / input
WT
PFDN5 KO
Ex1
Ex5
Ex16
Total RNAPII levels in CD44
IP 8WG16 / input
WT
PFDN5 KO
Ex1
Ex9
Ex19
D
RNAPII Ser2-P levels in CTNNBL1
IP Ser2-P / IP RNAPII
WT
PFDN5 KO
Ex1
Ex5
Ex16
RNAPII Ser2-P levels in CD44
IP Ser2-P / IP RNAPII
WT
PFDN5 KO
Ex1
Ex9
Ex19

Supplementary figure 1. Efficiency of the siRNAs-mediated depletion of PFDN2 and PFDN5.



Supplementary figure 2. Differential expression of a group of genes affected in siPFDN2 and siPFDN5.

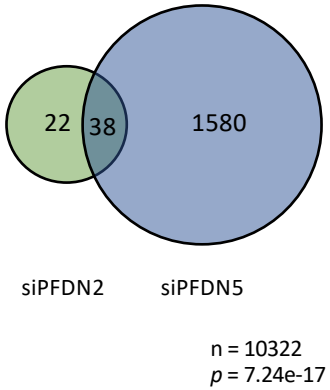
A)



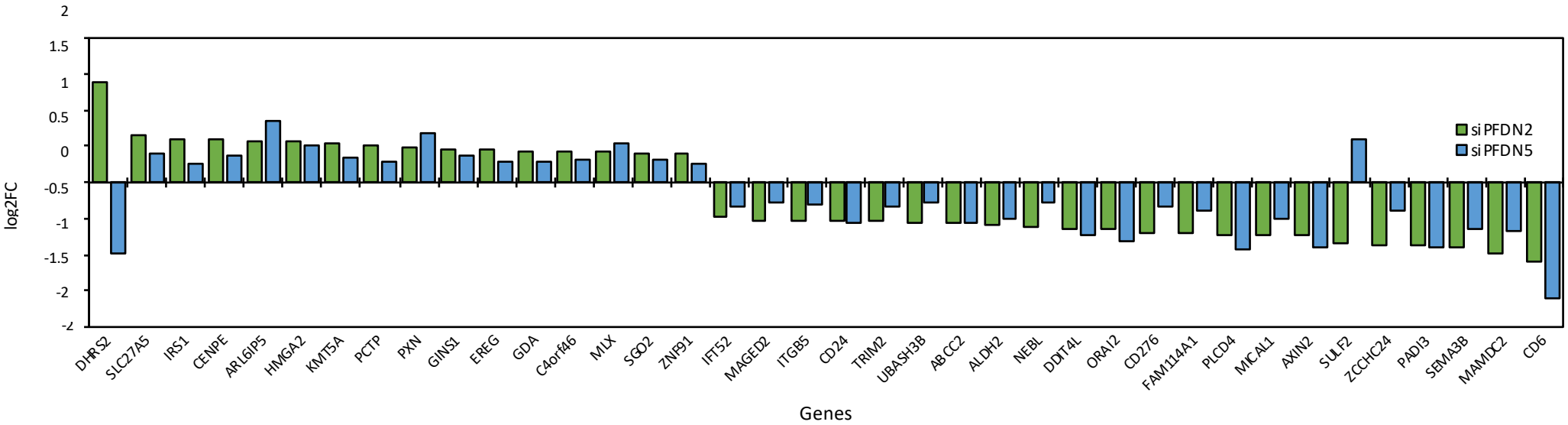
B)

GOs whose expression level changes in deficient cells in PFDN2 and PFDN5 after serum stimulation		
GO	***p* siPFDN2**	***p* siPFDN5**
cell adhesion	2.78E-14	6.30E-09
nervous system development	7.14E-12	0
animal organ morphogenesis	3.17E-10	1.09E-07
regulation of cell differentiation	1.26E-08	8.76E-07
vasculature development	2.33E-08	6.24E-07
negative regulation of developmental process	4.56E-06	1.89E-05
synapse organization	2.30E-05	2.87E-05
cytoskeleton organization	2.78E-05	2.32E-07
heart development	4.70E-05	1.94E-06

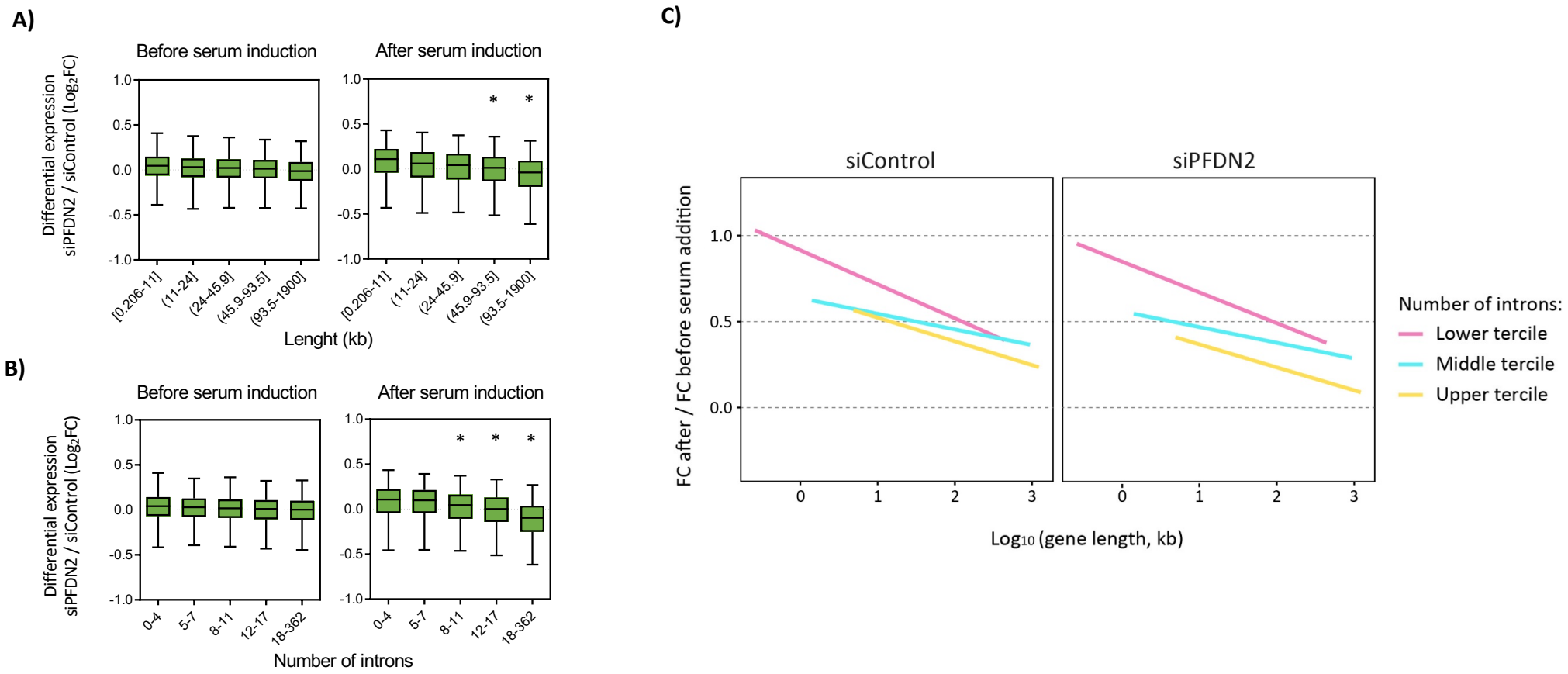
C)



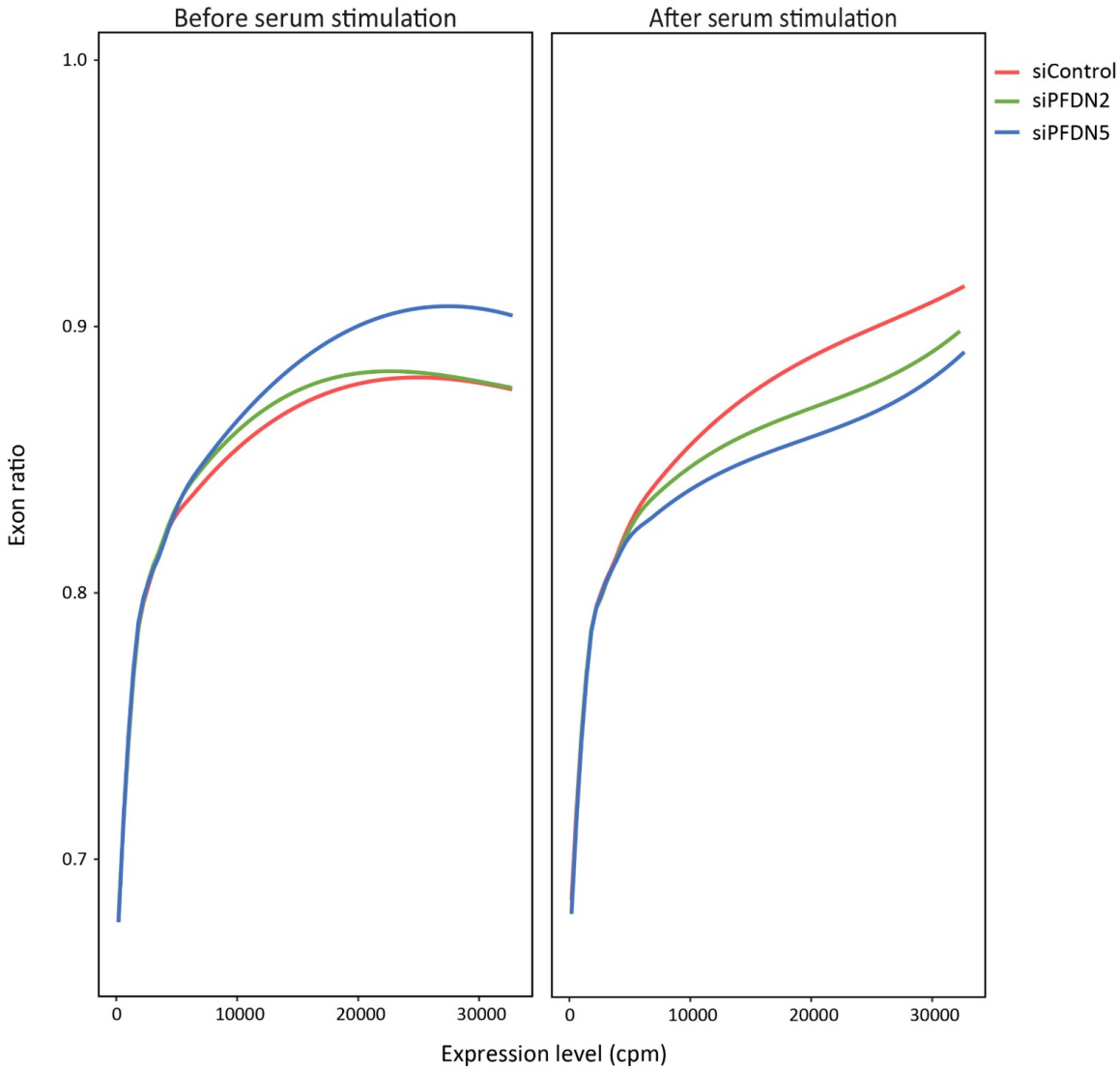
D)



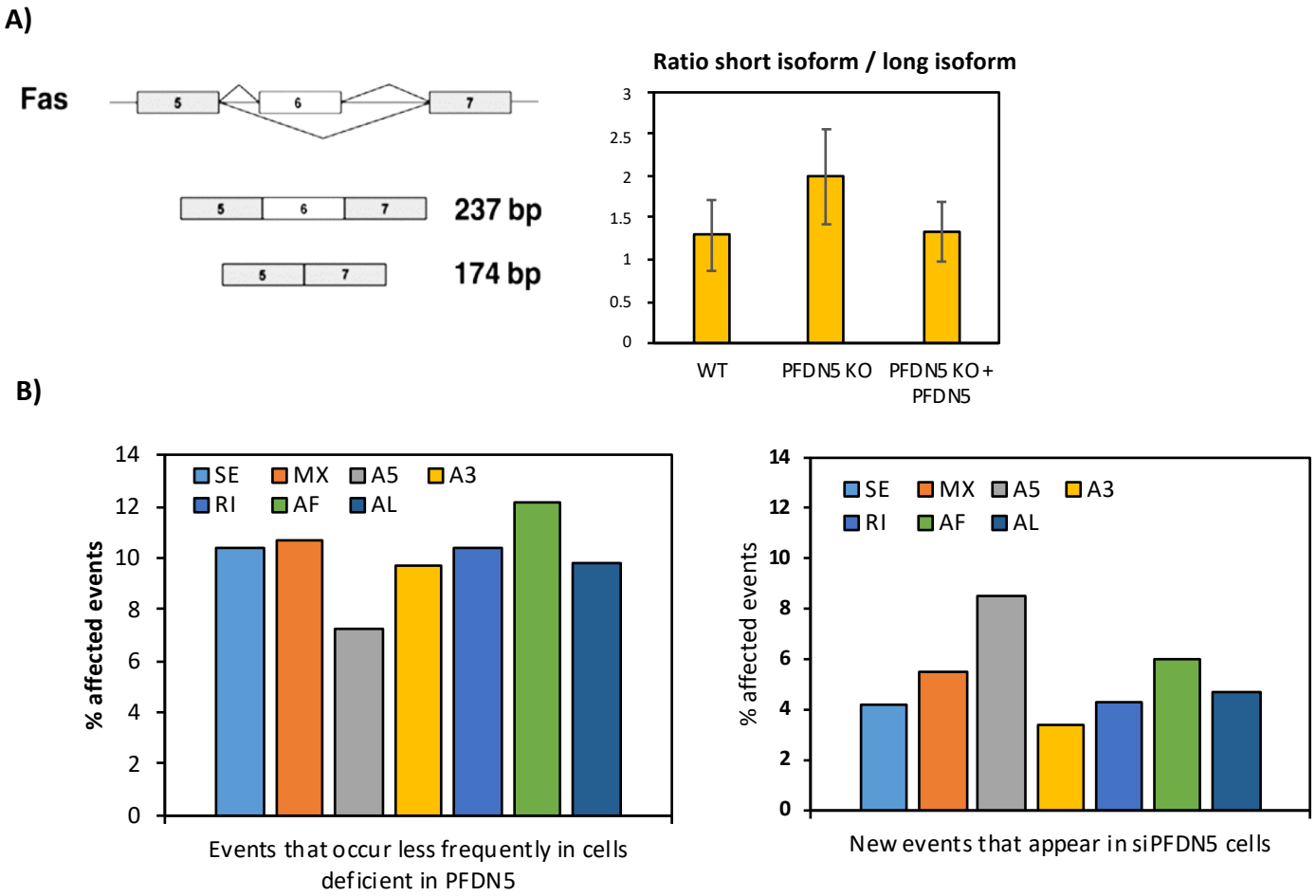
Supplementary figure 3. PFDN2 impairs the induction of long genes with a high number of introns.



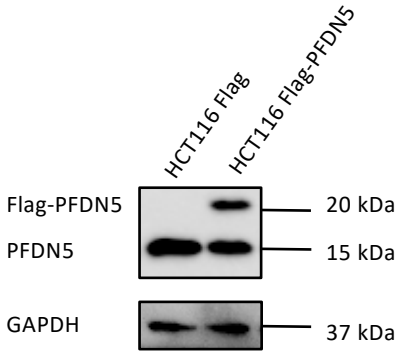
Supplementary figure 4. Exonic ratio distribution before and after serum induction.



Supplementary figure 5. PFDN5 deficiency has a mild general impact on alternative splicing.



Supplementary figure 6. HCT116 cell line expressing Flag-PFDN5.



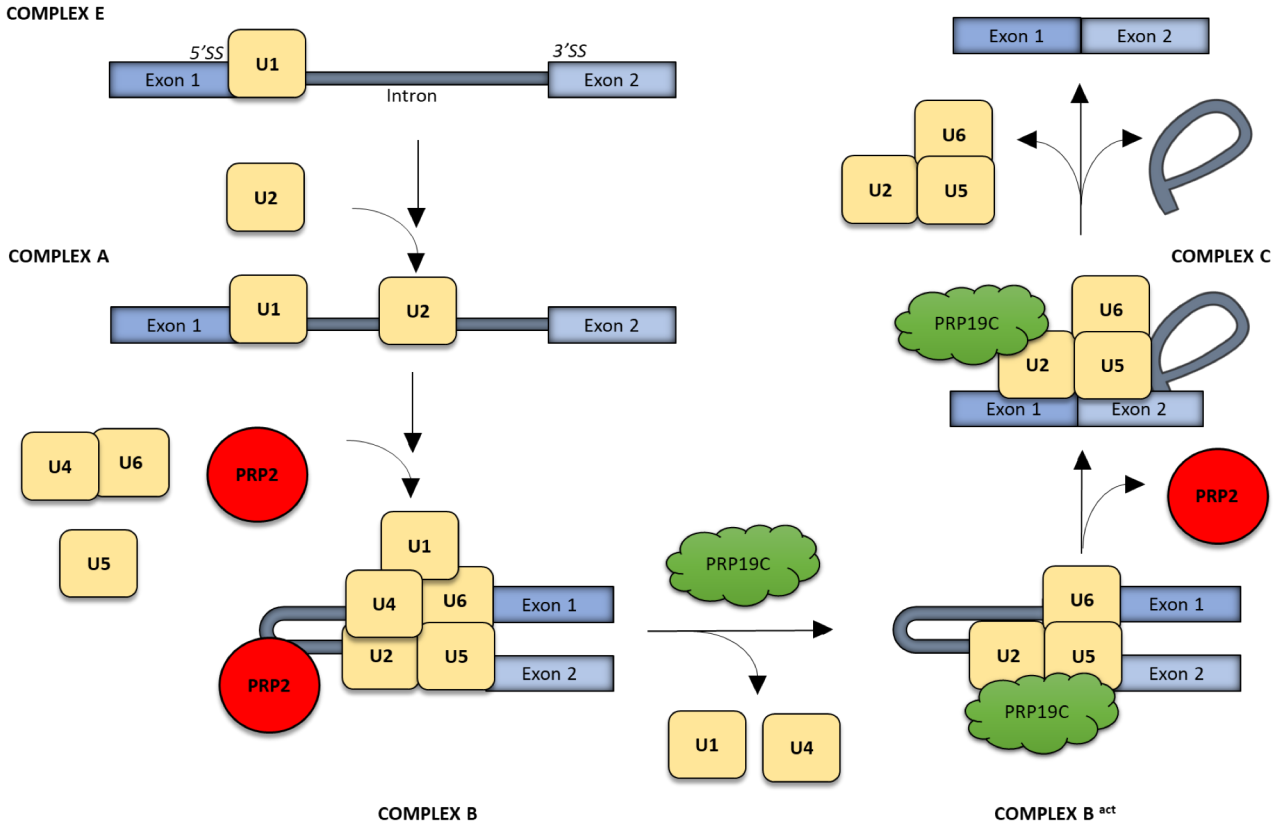
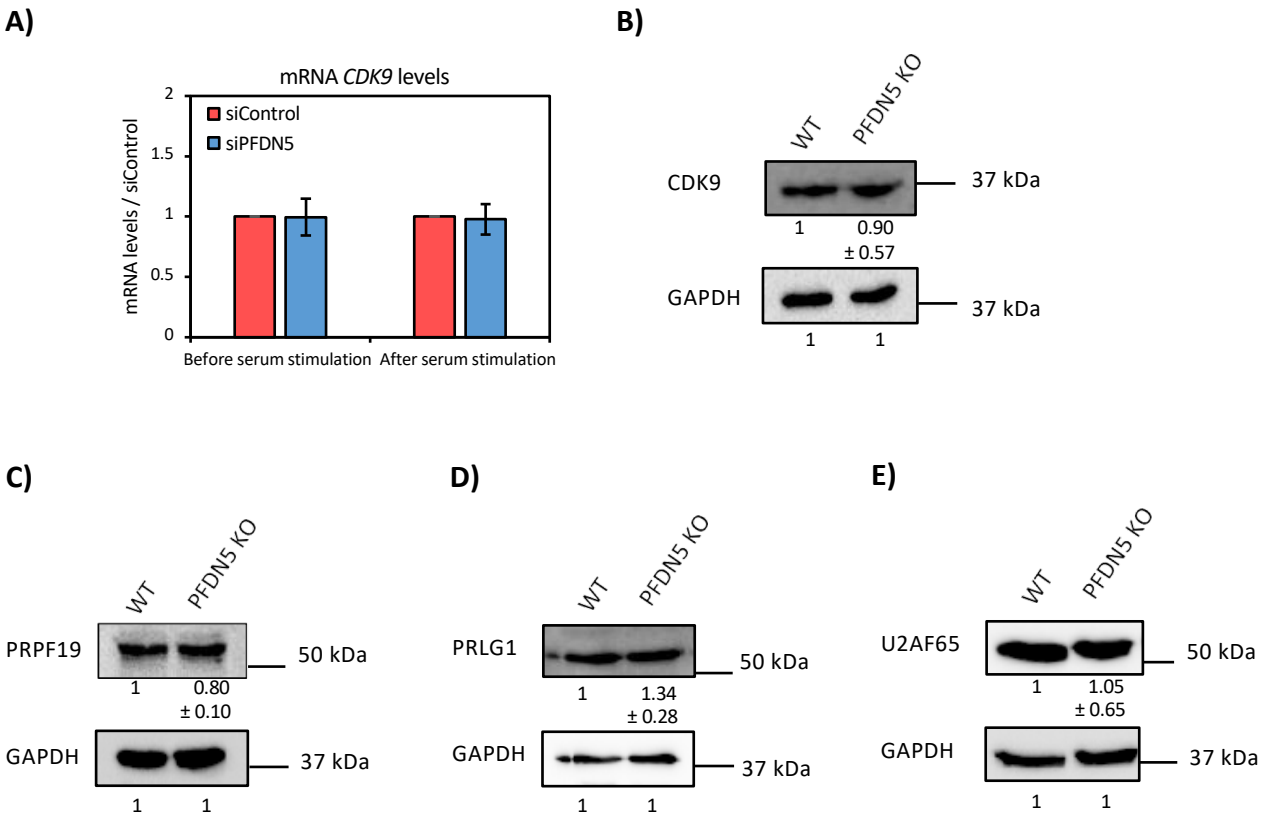
Supplementary figure 7. Correlations between prefoldin and splicing mutants.

A)

	Non-snRNP					snRNP			
	PRP19C								
	cef1-13	*clf1-1*	*prp46-5001*	*syf1-5001*	*prp2-1*	*lsm6Δ*	*lsm8-5003*	*smx3-5001*	*hsh49-5002*
pfdn1Δ	X						X	X	
pfdn2Δ	X		X		X				
pfdn3Δ	X								
pfdn4Δ	X	X		X	X	X			
pfdn5Δ	X								
pfdn6Δ	X	X			X				X

Table. Pearson correlation coefficient ≥ 0.13
Source: https://thecellmap.org/

B)



Supplementary figure 8. Total levels of CDK9 and different splicing factors remain constant in the absence of PFDN5.