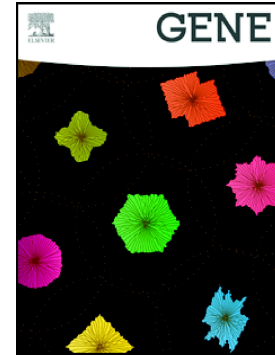Accepted Manuscript

hnRNPDL extensively regulates transcription and alternative splicing

Ruth Zhen Li, Jing Hou, Yaxun Wei, Xiaotian Luo, Yali Ye, Yi Zhang

hnRNPDL extensively regulates transcription and alternative splicing

Ruth Zhen Li[1,2], Jing Hou[2,3], Yaxun Wei[2], Xiaotian Luo[3], Yali Ye[3], Yi Zhang[2,3]

[1]Department of Bioinformatics, School of Basic Medical, Southern Medical University, Guangzhou, Guangdong 510515, China

[2]Center for Genome Analysis, ABLife Inc., Optics Valley International Biomedical Park, Building 18-1, East Lake High-Tech Development Zone, 388 Gaoxin 2nd Road, Wuhan, Hubei 430075, China

[3]Laboratory for Genome Regulation and Human Health, ABLife Inc., Optics Valley International Biomedical Park, Building 18-2, East Lake High-Tech Development Zone, 388 Gaoxin 2nd Road, Wuhan, Hubei 430075, China

Summary

RNA binding proteins (RBPs) are key players of genome regulation. Here we report the transcriptome study of HnRNP D-Like protein, which belongs to the hnRNP family. We used RNA-seq to analyze the global transcript level and alternative splicing on *hnRNPDL* shRNA-treated cells and control. Sh-*hnRNPDL* extensively increased in the expression of genes involved in female pregnancy, cell apoptosis, cell proliferation and cell migration. HnRNPDL regulated alternative splicing of hundreds of genes enriched in transcription regulation and signalling pathways including NOD-like receptor signaling, Notch signaling, and TNF signaling. This study provides the first transcriptome-wide analysis of hnRNPDL regulation of gene expression, which adds to the understanding of critical hnRNPDL functions.

Keywords hnRNPDL, knock down, RNA-seq, gene expression, alternative splicing

Introduction

RNA-binding proteins (RBPs) play key roles in post-transcriptional regulation, which can occur at many different steps in RNA metabolism, including pre-mRNA splicing and polyadenylation, mRNA stability, mRNA localization and translation (Curtis et al., 1995; Wickens et al., 2000; de Moor and Richter, 2001; Johnstone and Lasko, 2001). Many RBPs have been identified as the essential factors during germline and early embryo development. Also, RBPs with essential functions in the development of somatic tissues, including neuron, muscle, hypodermis, and excretory cells, as well as in the timing of development have also been identified (Lee and Schedl, 2006). As for the regulation mechanisms, RBPs have been shown to control the expression of numerous proteins by binding to the respective mRNA species encoding proto-oncogenes, growth factors, cytokines, transcription factors, and other proteins in various cell types (Eberhardt et al., 2007). In particular, RBPs exert regulatory function during cancer metastasis and proliferation via affecting splicing(Venables, 2004; Venables et al., 2009). Thus, it is meaningful to explore the regulatory mechanism of RBPs in cancer tissue or cell lines, which will provide a basis for the cancer treatment.

Heterogeneous nuclear ribonucleoproteins, hnRNPs, are specific groups of RBPs consisting of multiple RNA-binding domains, which include conserved RNP-1 and RNP-2 or hnRNP K homology motifs. Their amino- and carboxyl-terminal domains are commonly divergent (Piñol-Roma and Dreyfuss, 1993; Weighardt et al., 1996b; Krecic and Swanson, 1999). As for their functions, hnRNPs have several key

roles in mRNA metabolism such as pre-mRNA packaging, pre-mRNA processing, mRNA splicing, nuclear-cytoplasmic mRNA shuttling, mRNA decay and translation (Dreyfuss et al., 1993; Weighardt et al., 1996a; Krecic and Swanson, 1999). Over the past two decades, the potential roles of hnRNPs have been appraised in DNA repair, telomere biogenesis, transcription, tumor development and metastasis (Carpenter et al., 2006; Han et al., 2013). Recent studies of the function of hnRNPs in cancer metastasis by regulating alternative splicing have provided insight into the mechanism of this poor prognosis (Xu et al., 2014; Howley et al., 2016; Puppo et al., 2016; Moshiri et al., 2017).

The human DNA- and RNA- binding protein heterogenous nuclear ribonucleoprotein D-like protein (*hnRNPDL*, also known as *JKTBP*) belongs to the family of ubiquitously expressed heterogeneous nuclear ribonucleoproteins (hnRNPs). *JKTBP* has been cloned from human myeloid leukemia by DNA affinity screening, and found to contain two RNA binding domains (Tsuchiya et al., 1998). In fact, there are three isoforms of *JKTBP*, 1, 2 and 1Δ6, but only the dominant form JKTBP1 shuttles between the nucleus and cytoplasm, with JKTBP1 directly interacting with poly (A) $^{+}$ RNA (Kawamura et al., 2002). Furthermore, JKTBP proteins are sequence-specific RNA binding proteins with a binding consensus of ACUAGC, which is different from the most closely related hnRNP D (Kamei and Yamada, 2002). In addition to nuclear export, *JKTBP* has been involved in cap-independent translation by binding to the internal ribosome entry site in the 5′UTR of NF-κB-repressing factor (NRF) and RNA stability by binding to the 3′UTR of NRF mRNA (Omnus et al., 2011). Based on the structural characteristics, *hnRNPDL* has been reported playing important roles in some diseases. For example, a point mutation in exon 6 of *hnRNPDL* causes muscle disorder in the limb-girdle muscular dystrophies (Vieira et al., 2014). Also, the *hnRNPDL* deletion is associated with growth retardation and hypotonia (Hu et al., 2016). In particular, *hnRNPDL* overexpression can lead to abnormal cell proliferation in prostate cancer cells (Wu et al., 2008). Given its binary DNA and RNA binding activity, hnRNPDL is predicted to regulate transcription and alternative splicing, which should contribute to its biological function. However, little information is available in these aspects.

Here, we used unbiased transcriptome analysis to study how hnRNPDL regulates gene transcription and alternative splicing in HeLa cells derived from a cervical cancer patient. We successfully knocked down hnRNPDL by shRNA HeLa. High-throughput RNA sequencing (RNA-seq) were performed for shRNA cells and control cells to define the comprehensive gene expression profiles and identify the genome-wide alternative splicing events regulated by hnRNPDL. The results showed clear changes in transcript profiles and splicing patterns of specific subsets of genes after *hnRNPDL* knockdown in HeLa cells; some of the differentially expressed genes are involved in cell proliferation and migration. Our study provides an important basis and data platform to clarify the role of hnRNPDL in regulation of gene transcription and alternative splicing in tumorigenesis.

Results

Profiling hnRNPDL-regulated gene expression

In order to identify the changes specific to *hnRNPDL*, we examined the response of HeLa cells to the shRNA treatment against *hnRNPDL*, comparing with the empty vector-transduced cells. The efficacy of *hnRNPDL* knockdown, as assessed by RT-PCR, was approximately 50% (Fig. 1A). To comprehensively investigate the hnRNPDL-mediated transcriptional regulation, we constructed cDNA libraries prepared from control and *hnRNPDL* KD cells (two biological replicates). The four RNA-seq

libraries were sequenced on Illumina HiSeq2000 platform to produce 150 nucleotide paired-end reads per sample.

The sequencing data was reviewed to ensure their reliability (Table 1). A total of 84.5M±6.4M raw reads per sample were generated and 80.9M±6.2M clean reads per sample were retained by the removal of adaptors and contaminating sequences. Among these, 77.8M±6.1M were paired-end reads per sample, and an average of 70.3±7.4 million read pairs per sample were aligned to the human genome.

To compare the gene expression patterns across individuals, expression values in units of fragments per kilobase of exon model per million fragments mapped (FPKM) were calculated using an in-house pipeline. RNA-seq yielded 30,012 genes with expression evidence, among which 13851 genes were detected at an expression level of FPKM > 1 in at least one condition (Table S1). Effective depletion of *hnRNPDL* was further confirmed in parallel RNA-seq analysis (Fig. 1B). FPKM values were used to calculate a correlation matrix based on Pearson's correlation coefficient. The diagonal of the heat map showed the Pearson correlation between KD and control cells. The biological replicates were highly correlated (Fig. 1C, Fig. S1A).

Table 1 Summary of RNA-seq reads used in analysis

Sample	shHNRNPDL_1st	shHNRNPDL_2nd	shCtrl_1st	shCtrl_2nd	Mean[a]
Raw reads	79394700	80361816	93507914	84592866	84464324±6437771.42
Clean reads	76258020	76745033	89801511	80980963	80946381.75±6272866.89
paired-end reads	73444180	73464018	86459512	77702634	77767586±6130968.75
Total mapped (%[b])	64795719(88.22%)	64887591(88.33%)	80579969(93.2%)	70908159(91.26%)	70292859.5±7430538.26
Uniquely mapped (%[c])	62528563(96.5%)	62661231(96.57%)	77470061(96.14%)	66495398(93.78%)	67288813.25±7032349.38
Splice reads (%[d])	24899723（39.82%）	24517387（39.13%）	29816216（38.49%）	25237326（37.95%）	

a The mean and standard deviation across the 4 samples

b the percentage of paired-end reads that were mapped to the genome

c the percentage of unique reads mapping out of the total mapped reads

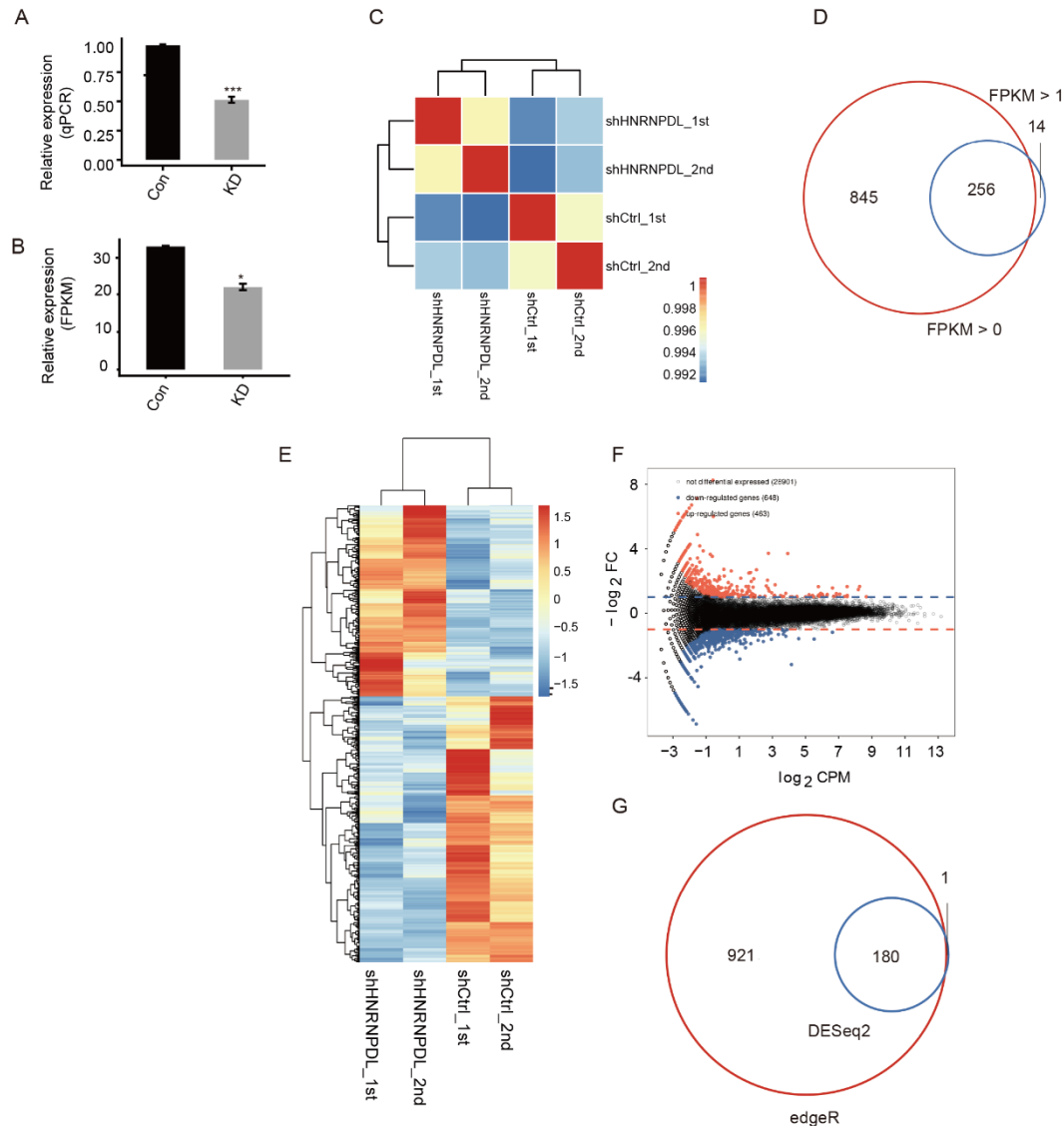d the percentage of uniquely mapped reads that were mapped to splice site

Fig. 1 Differential gene expression in response to *hnRNPDL* knockdown. (A) *HnRNPDL* mRNA expression in HeLa cells after transiently transfected with *hnRNPDL* specific shRNA or control vector, as determined by qRT-PCR. (B) *HnRNPDL* expression values quantified using RNA sequencing data. FPKM values were calculated as explained in Materails and Methods. (C) Heat map showing the hierarchically clustered Pearson correlation matrix resulting from comparing the transcript expression values for control and *hnRNPDL* shRNA-treated samples. (D) Venn diagram showing the overlapped DEGs resulted from all expressed genes and genes with FPKM > 1 in at least one sample. (E) Hierarchical clustering of 1101 DEGs in control and shRNA treated samples. Expression values (FPKM) are log2-transformed and then median-centered by each gene. (F) Presentation of the hnRNPDL regulated genes by MA plot, differentially up-regulated genes (FC≥2, FDR< 0.05) are labeled red, whereas differentially down-regulated (FC≤0.5, FDR< 0.05) are labeled blue. (G) Venn diagram showing the overlapped DEGs resulted from all expressed genes and genes using edgeR and DESeq2 methods.

HnRNPDL extensively regulates HeLa cell transcription

On the base of the high quality RNA-seq data obtained from sh-hnRNPDL and control cells, we then explored genes whose transcription were significantly regulated by hnRNPDL. Differentially expressed genes (DEGs) between the KD and control cells was identified using edgeR (Robinson et al., 2010). When we set the cutoff as fold change ≥2 or ≤0.5 and a 5% false discovery rate (FDR), we surprisingly identified 460 up-regulated and 641 down-regulated genes among all expressed genes, indicating that

hnRNPDL extensively regulates gene transcription. Meanwhile, 97 up-regulated and 173 down-regulated genes were detected among genes with at least one sample expression level of FPKM > 1. Interestingly, DEGs from both sets of expressed genes were highly overlapped (Fig. 1D). Details of the differentially expressed genes can be found in Table S2. Heatmap analysis of the expression patterns of DEGs in RNA-seq samples showed a high consistency of the hnRNPDL-mediated transcription in both data sets (Fig. 1E, Fig. S1B). MA plot was constructed to display the significantly expressed genes that were associated with *hnRNPDL* knockdown (Fig. 1F).

We further used DESeq2 (Anders and Huber, 2010) to perform the DEG analysis on our data, and compared the results between DESeq2 and edgeR. Using a cutoff of FDR < 0.05, we found 88 and 93 up-regulated and down-regulated genes respectively, from the all expressed genes. Apparently, DESeq2 is more stringent than edgeR in recovering DEGs. MA plot of DESeq2 shows that genes with low expression level were not regarded as DEGs compared with edgeR (Fig. S1C). We found that all DEGs except one found by DESeq2 were included by edgeR from the overlapping analysis (Fig. 1G).

HnRNPDL regulates the gene expression involving in multiple cancer-related functions

In order to reveal the potential biological roles of these DEGs, all 1101 DEGs resulted from all expressed genes were subjected to GO and KEGG annotation. The functional analysis of these genes identified 485 genes (227 up-regulated and 258 down-regulated) annotated with GO categories, and 565 genes (263 up-regulated and 302 down-regulated) annotated with categories of KEGG.
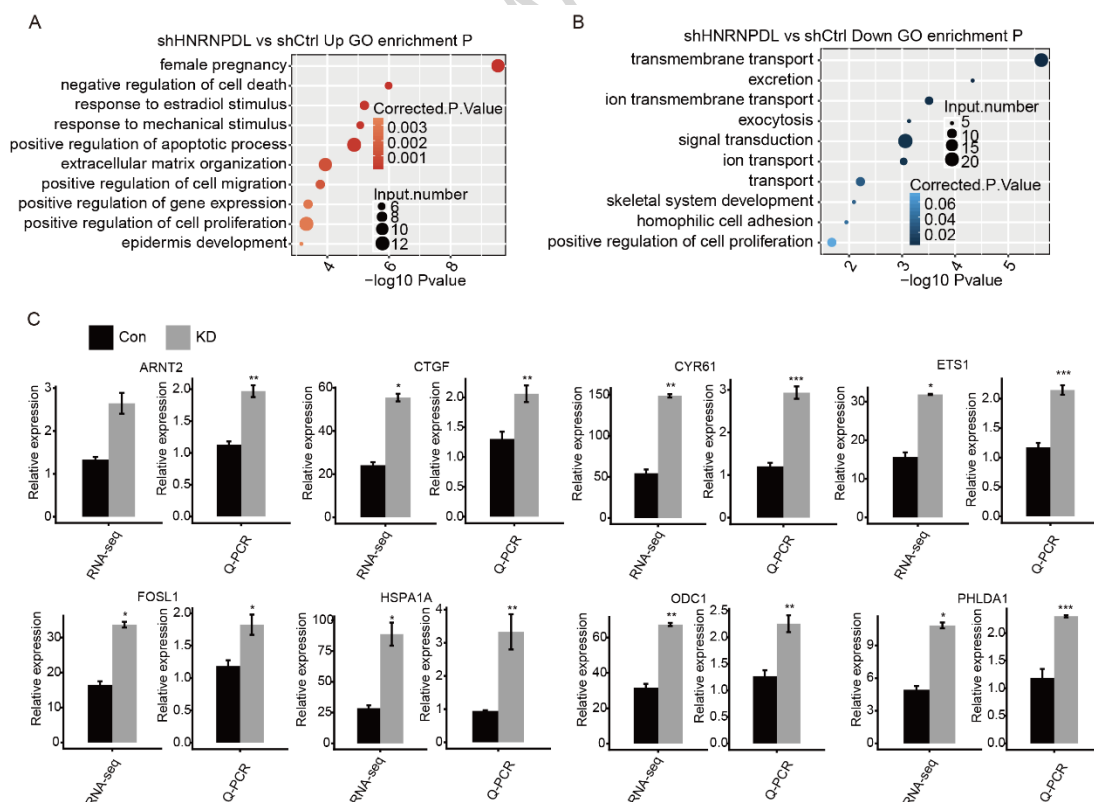

Fig.2 Gene ontology analysis and q-PCR validation of the up-regulated and down-regulated gene. (A-B) The top 10 representative GO Biological Process terms of *hnRNPDL*-regulated genes. (C) Validation of DEGs by identical q-PCR assay.

Based on the cutoff criterion, the upregulated DEGs were enriched in 49 GO terms, and the downregulated DEGs in 23 GO terms (see Table S3 for details). In the biological process terms of GO analysis, the up-regulated genes in the *hnRNPDL*-KD cells were mainly associated with cancer-related pathways including negative regulation of cell death, positive regulation of apoptotic process, extracellular matrix organization, positive regulation of cell migration, positive regulation of gene expression, positive regulation of cell proliferation (Fig. 2A). The down-regulated genes were mostly enriched in the transmembrane transport, the excretion and the signal transduction (Fig. 2B). When the adjusted *p*-value for KEGG pathways were set as <0.05, a number of cancer-related pathways including Estrogen, Jak-STAT and TGFβ signaling pathways were enriched (Fig. S1D-E; for details see Table S3).

The similar analysis was performed for edgeR DEGs resulted from the FPKM>1 genes, four top 10 biological terms were overlapped with those from the all expressed genes, which included positive regulation of apoptotic process, extracellular matrix organization, and positive regulation of cell proliferation (for details see Table S3). Down-regulated genes from those of FPKM>1 only enriched in signal transduction functional term, which was also the 5th-enriched term from those of all expressed genes (Fig. S1F-G). GO analysis for DESeq2-resulted DEGs, showing a good consistency with results from edgeR DEGs (Fig. S1H-I).

In order to verify the RNA-seq results, we conducted q-PCR analysis. CDNA was prepared from the control and *hnRNPDL* KD cells. Eight upregulated DEGs subjected to q-PCR analysis including *CTGF, HSPALA, CYR61, FOSL1, PHLDA1, ETS1, ARNT2 and ODC1*. The selected genes were enriched in the apoptosis, cell proliferation or other pathway of the GO analysis and the FPKM of these genes was higher than 1 in at least one sample. The results of this experiment are presented in Fig. 2C. We observed a statistically significant increase in *CTGF, HSPALA, CYR61, FOSL1, PHLDA1, ETS1, ARNT2 and ODC1*, in agreement with the RNA-seq analysis.

Identification of hnRNPDL-dependent alternative splicing events

One key aim of this study was to gain insight into the role of hnRNPDL on alternative splicing (AS) regulation. Therefore, we further used transcriptome sequencing to explore the hnRNPDL-dependent AS events in HeLa cells. We obtained 67.3M±7.0M uniquely mapped reads from *hnRNPDL*-KD and control HeLa cells, in which 37.95%~39.82% were junction reads (Table 1). Overall, when compared to the reference genome annotation, we detected 70.5% of annotated exons (258,972 out of 367,321 annotated exon). 168,599 annotated and 220,259 novel splice junctions were detected using TopHat2 (Kim et al., 2013).
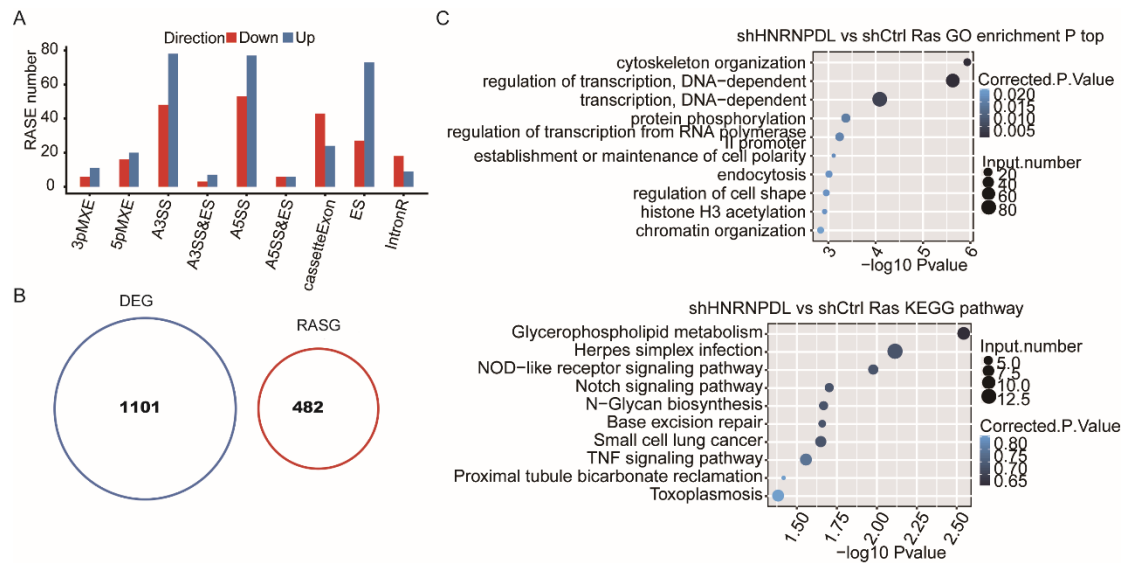
Fig.3 Identification and functional analysis of hnRNPDL-regulated splicing events. (A) Classification of different AS types regulated by hnRNPDL protein. (B) The overlap analysis between hnRNPDL-regulated differentially expressed genes (DEG) and alternative splicing gene (RASG) (C) the top 10 GO biological process analysis and KEGG functional pathway of alternative splicing gene

In order to investigate the global changes in AS occurring, AS events were analyzed from the RNA-seq dataset using ABLas software tool that we have previously developed (Xia et al., 2017). We detected 21,406 known alternative splicing events (ASEs) which have been annotated in the model genes, and we also detected 44,821 novel ASEs, excluding intron retention (Table S4).

When we used the two repeats as replicates and applied a stringent cutoff of p-value ≤0.05 and changed AS ratio ≥ 0.15 (See Material and Method) to identify high-confidence regulated alternative splicing events (RASEs), 554 RASEs were resulted. Complete RASEs can be found in Table S5 including 27 known intron-retention (IR) RASEs and 527 non-IR (NIR) RASEs. The NIR RASEs included 126 alternative 3′splice site (A3SS), 130 alternative 5′splice site (A5SS), 100 exon skipping (ES) and 67 Cassette Exon. The other event types included mutually exclusive 5′UTR (5pMXE, 36), mutually exclusive 3′UTR (3pMXE, 17), mutually exclusive exon (MXE, 29), alternative 5′splice site & exon skipping (A5SS & ES, 12) and alternative 3′splice site & exon skipping (A3SS & ES, 10) (Fig. 3A). These data suggested that hnRNPDL globally regulates alternative splicing events in HeLa cells.

We also combined two repeats into one sample pair, and ran ABLas and rMATS software (Shen et al., 2014) using cutoffs of FDR<=0.05 and differential ratio of 0.15. A total of 1455 and 704 exon skipping RASEs, which also included Cassette Exon, were identified using rMATS and ABLas, respectively. Among these RASEs, 282 were overlapped (Figure S2A). The other types of RASEs were not overlapped well. When we compared ES-type RASEs recovered from the analysis of the two experimental repeats as replicates using rMATS, the result was very similar to that from the combined one sample pair (Figure S2B). However, ABLas results from two different analysis yielded more different results, probably we applied t-test for replicate analysis (Figure S2C, Table S5). Different RASE analysis softwares have been developed using different criteria, it is conceivable to see difference in the output.

We decided to continue the study using ABLas results from treating samples as replicates for the following reasons. First, rMATS analysis and ABLas analysis using one sample pair were too sensitive in detecting RASEs. Second, ABLas classified more types of alternative splicing events. Third, we have validated the analyzed results using qPCR a number of independent projects, and found that ABLas NIR RASE results were quite reliable (Jin et al., 2017; Xia et al., 2017).

To rule out the possibility that the increase in AS events could be attributed to simply an upregulation in transcription, we analyzed whether the expression levels of the genes found to undergo hnRNPDL-regulated alternative splicing were also under transcriptional control. We found that genes under significant transcriptional control did not overlap with those under splicing control (Fig. 3B).

It was further revealed that the genes regulated by hnRNPDL-mediated alternative splicing were highly enriched for the cytoskeleton organization, regulation of transcription and protein phosphorylation (GO biological process terms, Fig. 3C). Enriched KEGG pathways (p-value greater than 0.05) included those involved in NOD-like receptor signaling pathway, Notch signaling pathway, and TNF signaling pathway (Fig. 3C). Genes involved in each GO term or KEGG pathway were provided in Table S6.

To validate hnRNPDL-regulated alternative splicing events identified from the RNA-seq data, the 23 RASEs classified into IR, ES, Cassette Exon, A5SS and A3SS were selected for q-PCR validation. Most of these genes in RASEs belonged to the enriched functional terms or KEGG pathways. For IR RASEs, we selected those with intronic average base >10 in at least one sample. For NIR RASEs, we selected those events with one class of junction reads >100 in at least one sample. The p-value of these events differed greatly, most were between 0.01-0.045 (Table S7). PCR primer pairs were listed in Table S8, and were designed to amplify the long splicing isoforms and the short splicing isoforms in the same reaction. Out of 23 tested events, 16 alternative splicing events validated by q-PCR were in agreement with the RNA-seq results. The 16 validated splicing events were located in the following genes, *CCAR1*, *IMMP1L*, *TNFRSF12A*, *RHOT1*, *MINK1*, *PUM2*, *NGLY1*, *NDEL1*, *CEP63*, *NR2C2*, *EPB41*, *SF1*, *ZNF200*, *B4GALT2*, *C11orf73*, and *DNPH1* (Fig. 4 and Fig.S3).
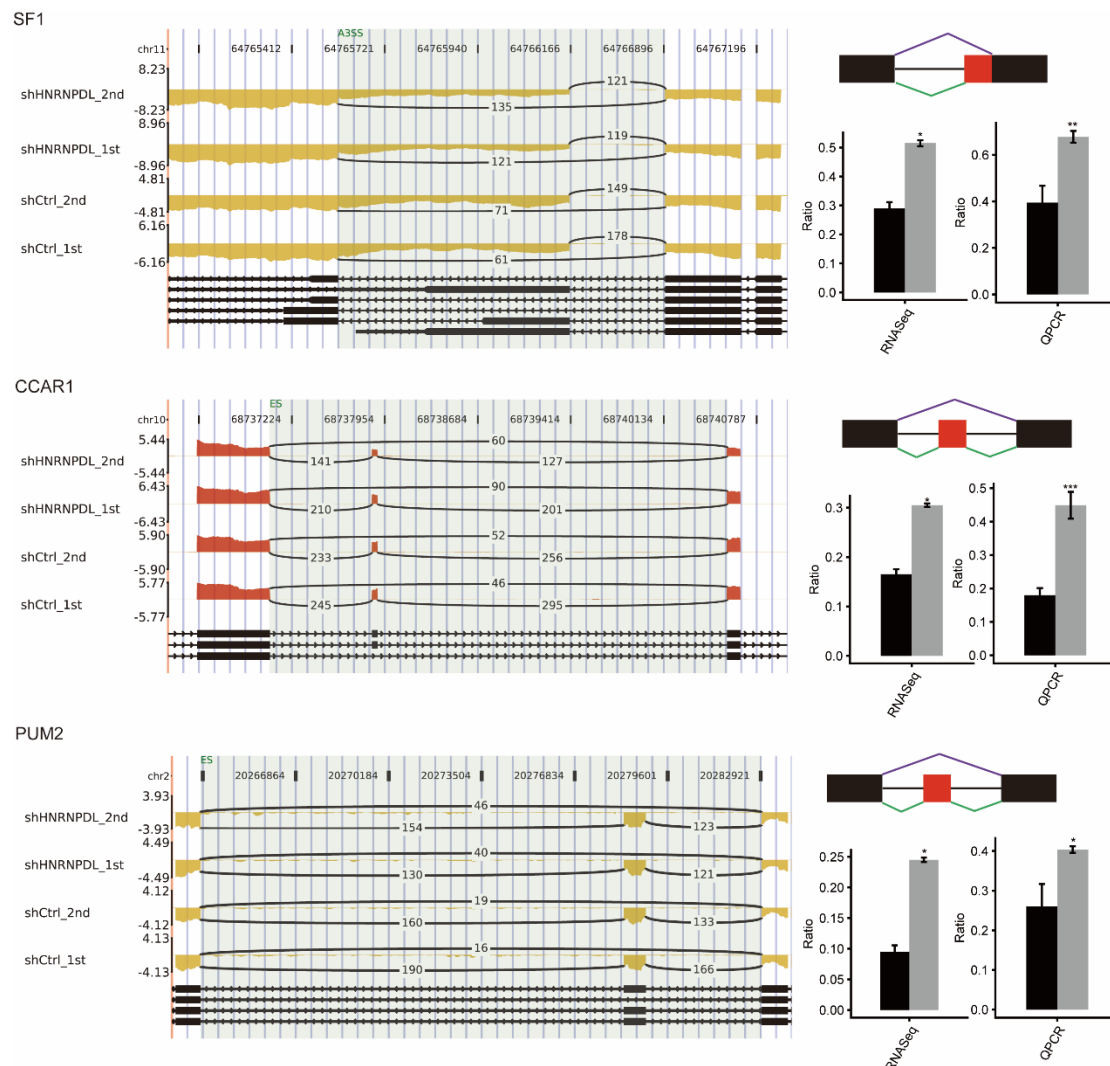
Fig.4 Validation of hnRNPDL-affected AS events. (Left panel) IGV-sashimi plots showing alternative splicing changes that occurred in HeLa cells depleted of hnRNPDL. The transcripts for the gene are shown below. (Right panel, top) the schematic diagrams depict the structures of alternative splicing events, AS1 (shown in purple) and AS2 (shown in green); exon sequences are denoted by boxes, and intron sequences by the horizontal line. (Right panel, bottom) RNA-seq quantification and RT-PCR validation of alternative splicing events. The altered ratio of AS events in RNA-seq were calculated using the formula: AS1 junction reads/AS1junction reads +AS2 junction reads; while the altered ratio of AS events in q-PCR were calculated using the formula: AS1 transcripts level/ AS2 transcripts level.

A recent report shows that hnRNPDL controls the expression of itself and hnRNPD by alternative splicing of cassette exons in their 3'UTRs (Kemmerer et al., 2018). We examined whether knock-down of hnRNPDL expression could influence the expression and alternative splicing of *hnRNPD*. RNA-seq results showed that expression of *hnRNPD* was slightly increased upon the reduction of hnRNPDL expression (Fig. S4A), consistent with the report (Kemmerer et al., 2018). HnRNPDL regulation of the alternative splicing of its own and hnRNPD was not evident in this study (Fig. S4B).

Further validation of hnRNPDL-regulated gene expression and alternative splicing

In order to further validate the confidence of the DEGs and alternative splicing events (ASEs) detected in this study, we performed additional set of experiments starting from shRNA treatment. We showed that both the protein level and mRNA level of hnRNPDL after shRNA treatment were down-regulated (Fig. 5A). We then performed q-PCR analysis of DEGs and ASEs after reverse transcription, using the total RNA prepared from the newly prepared cells. Among the 8 DEGs validated in Fig. 2, expression of *ARNT2* did not yield sufficient PCR signal, *FOSL1* showed no significant change, and all the increased expression of all other six genes upon HnRNPDL silencing were confirmed in the new experiments (Fig. 5B). Among the 16 ASEs validated in Figure 4 and Figure S2, we found that alternative splicing of pre-mRNAs from *CCAR1, TNFRSF12A, RHOT1, MINK1, NGLY1, NR2C2, EPB41, B4GALT2, C11orf73* were the same in both experiments (Fig. 5C). Alternative splicing of *SF1* pre-mRNA showed an inverse regulation by HnRNPDL in the new expriment, and that of *IMMP1L*, *PUM2*, *NDEL1*, *CEP63*, *ZNF200* did not show significant change.
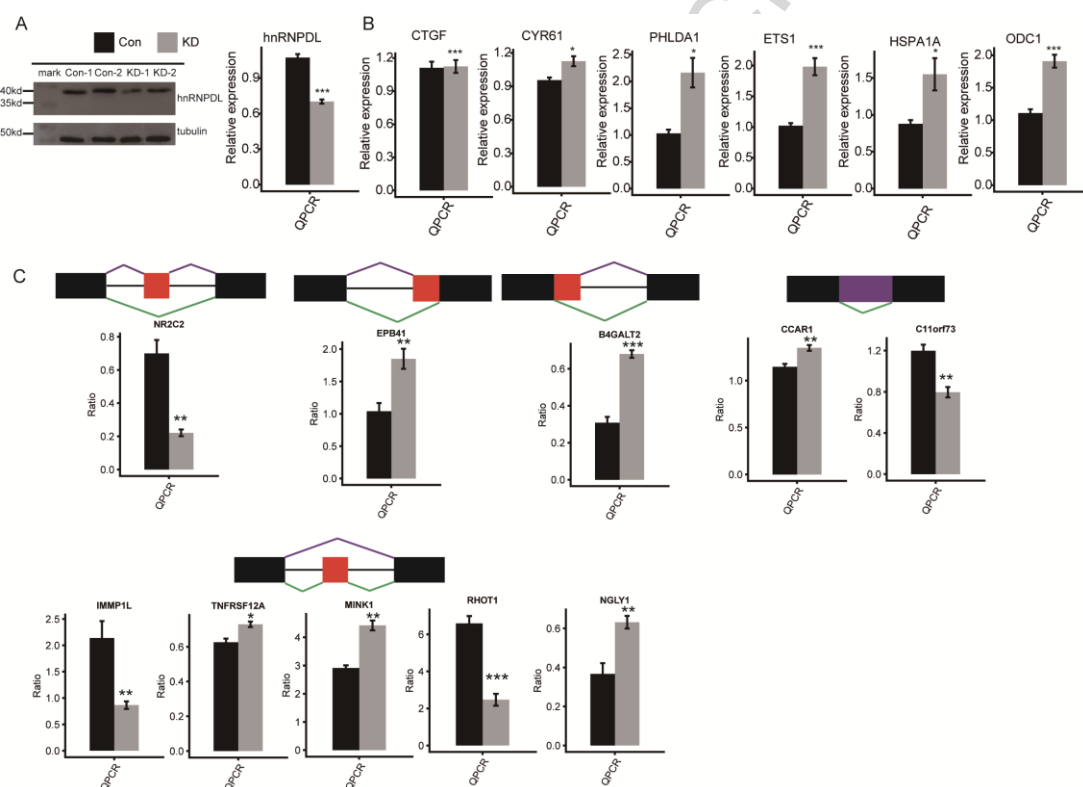
Fig.5 A new set of experiments validating HnRNPDL-regulated gene expression and alternative splicing. (A) Western blot analysis showed hnRNPDL protein expression decreased after shRNA treatment. (B) Q-PCR validation of hnRNPDL-regulated gene expression with DEGs shown in Fig. 2 being tested. The genes showing the same change pattern after hnRNPDL treatment were shown. (C) Q-PCR validation of hnRNPDL-regulated alternative splicing events with those shown in Fig. 4 and Fig. S2 being tested. The ASEs showing the same change pattern after hnRNPDL treatment were shown.

In summary, our newly added experiments conformed that the HnRNPDL-regulated gene expression and alternative splicing are confident.

Discussion

RNA-sequencing offers several key advantages: At first, unlike the hybridization-based approaches, RNA-seq is not limited to detecting transcripts that correspond to the existing genomic sequence. For example, 454-based RNA-seq has been used to sequence the transcriptome of the Glanville fritillary butterfly (Vera et al., 2008). RNA-seq can also reveal sequence variations (for example, SNPs) in the transcribed regions (Cloonan et al., 2008; Morin et al., 2008). RNA-seq does not have an upper limit for quantification, which correlates with the number of sequences obtained. Consequently, it has a large dynamic range of expression levels over which transcripts can be detected: a greater than 9,000-fold range was estimated in a study that analyzed 16 million mapped reads in *Saccharomyces cerevisiae* (Nagalakshmi et al., 2008), and a range spanning five orders of magnitude was estimated for 40 million mouse sequence reads (Mortazavi et al., 2008). RNA-seq also has shown to be highly accurate for quantifying expression levels, as determined by using quantitative PCR (qPCR) and spike-in RNA controls of known concentration (Nagalakshmi et al., 2008). The results of RNA-seq also show high levels of reproducibility, for both technical and biological replicates (Cloonan et al., 2008; Nagalakshmi et al., 2008).

In this work, we have performed the transcriptome analysis of *hnRNPDL* knock-down and control cells.

Prior studies point to hnRNPDL affecting mRNA nuclear export (Kawamura et al., 2002) and translation (Omnus et al., 2011). Using RNA-seq, we have identified a large number of up-regulated and down-regulated genes, revealing that the inactivation of *hnRNPDL* has widespread effects.

Our functional analysis indicates that these changes in gene expression may impact female pregnancy，cell death, cell proliferation, cell migration and multiple signaling pathways. Although hnRNPDL is known to bind to single-stranded DNA or double-stranded DNA, its transcriptional regulation has not been identified to date. The transcriptional functions of hnRNPDL might be affected by interaction with other proteins (Tsuchiya et al., 1998). The effect of hnRNPs affecting transcription through protein-protein interaction has been described for hnRNP K (LH et al., 1998) and hnRNP U (MK and VM, 1999).

In this study, we used HeLa cell which was originated from a cervical cancer patient. Interestingly, we found that hnRNPDL represses the expression of female pregnancy genes encoding pregnancy-specific glycoprotein (PSGs), *ETS1* (v-ets avian erythroblastosis virus E26 oncogene homolog 1), *FOSL1* (FOS-like antigen 1), *ENDOU* (endonuclease, poly U-specific), and *FOSB* (FBJ murine osteosarcoma viral oncogene homolog B) (Table S2). PSGs were high in the syncytiotrophoblast starting as early as about the time of implantation (Grudzinskas et al., 1977). Other studies have reported that PSGs were expressed in fetal liver, salivary gland, testis, myeloid cells and intestine (Horne et al., 1977; Borjigin et al., 1990; Shupert and Chan, 1993). However, the expression level in these tissues appears to be very low (Salahshor et al., 2005). In addition, *PSG9*, a member of PSG family, was upregulated in cancer cells and it appears to be useful as a biomarker for the early detection of colorectal cancer (Salahshor et al., 2005). *ETS1* is a proto-oncogene expressing in various cancer cells and associated with tumor progression(Trojanowska, 2000; Tien et al., 2004). *ETS1* is involved in villus cytotrophoblast differentiation (Kessler et al., 2015), and is expressed in normal human placenta (Takai et al., 2006). In addition, *ETS1* mRNA levels increased with the uterine cervical cancer disease stage (Fujimoto et al., 2002). We show that expression of *ETS1* (FDR = 2.02E-42) is negatively regulated by hnRNPDL.

Similarly, *FOSL1* (also known as *FRA-1*) and *FOSB*, which belong to the Fos family of AP-1 transcription factors, were elevated in *hnRNPDL* KD cell lines. *FOSL* is emerging as a key regulator in tumor cell EMT/MET plasticity (Dhillon and Tulchinsky, 2015).

In this study, the control HeLa cells exhibited the low expression level of PSGs (FPKM<1, Table S2); however, the expression of PSGs were upregulated after *hnRNPDL* knock-down, indicating that these genes could play additional roles in cancer cells. Further, the *hnRNPDL* expressions were strong in the mouse normal neuronal cells, oocytes and spermatocytes (Akagi et al., 2000). We proposed that PSGs might be associated with cervical cancer, in addition to their association with pregnancy. And the transcriptional activity of the PSG family may be modulated by hnRNPDL protein.

We have also shown that upon *hnRNPDL* knock-down, genes in a number of tumorigenesis functions including cell death, cellular proliferation and cellular migration are upregulated. These include *CTGF* (connective tissue growth factor) and *CYR61* (cysteine rich-61), which are the members of the connective tissue factor CCN (CYR61, CTGF, NVO) family. The CCN proteins function as ligands of the integrin to regulate cell proliferation, apoptosis, cell migration/adhesion, and angiogenesis(Yeger and Perbal, 2007). TAZ interacting with TEAD family member functionally activates *CTGF* and *CYR61* transcription (Lai et al., 2011). KEGG analysis also revealed a number of tumorigenesis-related signaling pathways. HnRNPDL-repressed genes are enriched in estrogen signaling pathway, which include *ESR1, ADCY2, CREB5, HSPA6, HSPA1B, HSPA1A, MMP9, GABBR2*. In addition to ESR1 discussed above, ADCY2, CREB5, MM9 are all known to be involved in tumorigenesis (van Kempen and Coussens, 2002; Belotti et al., 2003; Yu et al., 2011; Qi and Ding, 2014). Heat shock proteins play important roles in cancer, particularly in estrogen-related cancers, and has been widely used in cancer diagnosis, prognosis, prediction (Veldscholte et al., 1992; Ciocca and Calderwood, 2005). *IL7R, SOCS1, GH1, IL23A, IL24, SPRY4* in Jak-STAT pathways. Interestingly, most genes in this pathway are involved in immune response. Taken together, our results suggest that hnRNPDL might function as a tumor suppressor by repressing the expression of a number of oncogenes. The mechanism of this transcriptional repression requires further investigation.

Notably, the vast majority of AS events occurred in genes involved in regulation of transcription. Examples of these genes include *SF1* (Splicing Factor 1), *CCAR1* (Cell Division Cycle and Apoptosis Regulator 1) and *NR2C2* (Nuclear Receptor Subfamily 2 Group C Member 2).

We have observed *hnRNPDL*-dependent alternative splicing of *SF1* resulted from the use of a cryptic donor site in intron 12 and generated the long isoform in *hnRNPDL*-knockdown cells. SF1 was initially isolated as a factor necessary for pre-spliceosome formation *in vitro* (Krämer, 1992). Previous studies reported that the role of the unique C-termini of different *SF1* isoforms was elusive, sequences within the divergent proline-rich portion. The proline-rich region of C-termini have been showed to interact with formin-binding proteins, which raised the possibility that the function of different *SF1* isoforms was regulated by interaction with these proteins (Krämer et al., 1998).

We have found that *hnRNPDL* depletion significantly alters the inclusion of known variable exons of *CCAR1, TNFRSF12A, MINK1, PUM2, NGLY1, and NR2C2* (details in Fig. 4 and Fig. S3). HnRNPL was also reported to regulate the alternative splicing of the RNA-binding protein PUM2 and CCAR1.

Inclusion of the variable exon regulates overall protein expression or alters the domain structure of the protein (Ganesh et al., 2014). Previous studies suggest that CCAR1 can serve as a key intracellular transducer of either proliferation or apoptosis signaling pathways in response to different signals (Rishi et al., 2003; Kim et al., 2008). In addition, CCAR1 function as a coactivator interacting with β-catenin to activate Wnt signaling (Ou et al., 2009).

The sequence organization, and DNA- and RNA-binding properties of hnPNPDL suggest that it could have dual functions in mRNA synthesis at the transcriptional and post-transcriptional levels. In this study, we have successfully applied RNA-seq technology to study the role of hnRNPDL in a human cervical cancer cell line, demonstrating hnPNPDL regulation of both transcription and alternative splicing. And these two layers of regulation seems to be uncoupled from each other. We show hnPNPDL negatively regulates the transcription of genes involved in tumorigenesis-related functions including cell death, cell proliferation, cell migration, and Jak-STAT signaling pathway. Interestingly, hnPNPDL also represses the expression of genes enriched in female pregnancy and estrogen signaling pathway, and most genes in these functions are also associated with tumorigenesis. It also regulates the alternative splicing of genes enriched NOD-like receptor, Notch and TNF signaling pathways, as well as transcription factors. The findings presented provide several molecular mechanisms for the role of hnRNPDL in physiological and pathological states such as cancers.

Acknowledgements

Material and methods

Cell culture and transfections

HeLa cells were cultured under standard conditions with Dulbecco's modified Eagle's medium (DMEM) with 10% fetal bovine serum (FBS), 100 μg/mL streptomycin, and 100 U/mL penicillin. shRNA transfection of HeLa cells was performed using lipofectAMINE 2000 (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's protocol. Transfected cells were harvested after 48 h for q-PCR analysis.

The efficient hnRNPDL-shRNA sequence is CGGATACTTCTGAAGAACAAA.

Assesment of knockdown by shRNA

Candidate control genes for assessment of effects of *hnRNPDL* shRNA were *GAPDH* (glyceraldehyde-3-phosphate dehydrogenase). CDNA synthesis was done by standard procedures and real time PCR was performed on the Bio-Rad S1000 with Bestar SYBR Green RT-PCR Master Mix (DBI Bioscience, Shanghai, China). Transcript levels for the genes analyzed were measured in comparison with the housekeeping gene *GAPDH* as an internal reference standard, using the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001).

RNA extraction and sequencing

The HeLa cells were ground into fine powder before RNA extraction. Total RNA was extracted by the hot phenol method. The RNA was further purified with two phenol-chloroform treatments and then treated with RQ1 DNase (Promega, Madison, WI, USA) to remove DNA. The quality and quantity of the purified RNA were redetermined by measuring the absorbance at 260 nm/280 nm (A260/A280) using Smartspec Plus (BioRad, USA). The integrity of RNA was further verified by 1.5% agarose gel electrophoresis.

For each sample, 10μg of the total RNA was used for RNA-seq library preparation. Polyadenylated mRNAs were purified and concentrated with oligo(dT)-conjugated magnetic beads (Invitrogen, Carlsbad, CA, USA) before directional RNA-seq library preparation. The purified mRNAs were then iron fragmented at 95°C followed by end repair and 5' adaptor ligation. Then, reverse transcription was performed with RT primer harboring 3' adaptor sequence and randomized hexamer. The cDNAs were purified, amplified, and stored at -80°C until they were used for sequencing.

For high-throughput sequencing, the libraries were prepared following the manufacturer's instructions. Illumina Nextseq 500 system was used to collect data from 151-bp pair-end sequencing (ABlife Inc., Wuhan, China).

RNA-Seq Raw Data Clean and Alignment

Raw reads containing more than 2-N bases were first discarded. Then adaptors and low quality bases were trimmed from raw sequencing reads using FASTX-Toolkit (Version 0.0.13). The short reads less than 16nt were also dropped. After that, clean reads were aligned to the GRch38 genome by TopHat2 (Kim et al., 2013) with 4 mismatches. Uniquely mapped reads were used to calculate reads number and FPKM value (FPKM represents fragments per kilobase and per million) for each gene.

Differentially Expressed Genes (DEG) analysis

FPKM (paired-end fragments per kilobase of exon per million fragments mapped), was used to evaluate the expression level of genes. To screen out the differentially expressed genes (DEGs), we applied the software edgeR [28], which was specifically used to analyze the differential expression of genes using raw RNA-Seq reads. To determine whether a gene was differentially expressed, we analyzed the results based on the fold change (fold change ≥2 or ≤0.5) and false discovery rate (FDR<0.05).

To predict the gene function and calculate the functional category distribution frequency, Gene Ontology (GO) analyses and enriched KEGG pathway were identified using KOBAS 2.0 server (Xie et al., 2011). Hypergeometric test and Benjamini-Hochberg FDR controlling procedure were used to define the enrichment of each pathway (corrected *p*-value<0.05).

Alternative splicing analysis

The alternative splicing events (ASEs) and regulated alternative splicing events (RASEs) between the samples were defined and quantified by using the ABLas pipeline as described previously (Xia et al., 2017). In brief, detection of seven types of ASEs was based on the splice junction reads. The eight types of ASE included Cassette exon (CassetteExon), Exon skipping (ES), Mutual exclusive

exon skipping (MXE), A5SS, A3SS, the MXE combined with alternative 5' promoter (5pMXE), and with alternative polyadenylation site (3pMXE). Intron retention was calculated based on the average base in the intronic and the adjacent exonic region.

After detecting the ASEs in each RNA-seq sample, fisher's exact test was chosen to calculate the significant p-value, with the alternative reads and model reads of samples as input data, respectively. The changed ratio of alternatively spliced reads and constitutively spliced reads between compared samples was defined as the RASE ratio. The *p*-value < 0.05 and RASE ratio > 0.2 were set as the threshold for RASEs detection.

Real time qPCR validation of DEGs and AS events

In this study, to elucidate the validity of the RNA-seq data, quantitative real-time PCR (qRT-PCR) was performed for some selected DEGs, and normalized with the reference gene GAPDH. The same RNA samples for RNA-seq were used for qRT-PCR. The PCR conditions are consisted of denaturing at 95˚C for 10 min, 40 cycles of denaturing at 95˚C for 15 s, annealing and extension at 60˚C for 1 min. PCR amplifications were performed in triplicate for each sample.

Primers for qPCR analysis were listed in Table S8.

Western blot analysis

HeLa cells line after 48-h transfection were collected in each group, lysed on ice for 30 min in RIPA buffer containing 50 mM Tris-HCl (pH 7.4), 150 mM NaCl, 1.0% deoxycholate, 1% Triton X-100, 1 mM EDTA and 0.1% SDS. The samples were centrifuged (12,000 rpm, 5 min), and supernatants were stored separately for backup. The protein was stored and frozen in refrigerator at 4°C after the protein concentration of every sample using BCA protein assay kit. Fifteen µg protein of each group were boiled for 10 min with protein denaturants separated by 10% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and transferred into a PVDF membrane (Millipore).The membranes were blocked in 5% skim milk diluted in phosphate-buffered saline
Containing 0.1% Tween-20 (PBST) one hour at room time. Then the membranes were incubated at 4°C overnight with the following primary antibodies:rabbit anti-human hnRNPDL (1 : 1000, A10721, ABclonal); rabbit anti-human Tublin (1:2000,Abclonal) was used as a loading control. After washing with PBST three times (5 min each time), the membranes were subsequently incubated with a goat anti-rabbit IgG (1 : 10000, as014, Abclonal) was marked by horseradish peroxidase (HRP) for 45min, and then washed by PBST three times (5min each time). The visualization of protein bands was realized by enhanced chemiluminescence (ECL) (JP001B250, CLINX).

References

Akagi, T., Kamei, D., Tsuchiya, N., Nishina, Y., Horiguchi, H., Matsui, M., Kamma, H. and Yamada, M., 2000. Molecular characterization of a mouse heterogeneous nuclear ribonucleoprotein D-like protein JKTBP and its tissue-specific expression. Gene 245, 267-73.

Anders, S. and Huber, W., 2010. Differential expression analysis for sequence count data. Genome Biol 11, R106.

Belotti, D., Paganoni, P., Manenti, L., Garofalo, A., Marchini, S., Taraboletti, G. and Giavazzi, R., 2003.

Matrix metalloproteinases (MMP9 and MMP2) induce the release of vascular endothelial growth factor (VEGF) by ovarian carcinoma cells: implications for ascites formation. Cancer Res 63, 5224-9.

Borjigin, J., Tease, L.A., Barnes, W. and Chan, W.Y., 1990. Expression of the pregnancy-specific beta 1-glycoprotein genes in human testis. Biochemical & Biophysical Research Communications 166, 622-9.

Carpenter, B., MacKay, C., Alnabulsi, A., MacKay, M., Telfer, C., Melvin, W.T. and Murray, G.I., 2006. The roles of heterogeneous nuclear ribonucleoproteins in tumour development and progression. Biochimica et Biophysica Acta (BBA)-Reviews on Cancer 1765, 85-100.

Ciocca, D.R. and Calderwood, S.K., 2005. Heat shock proteins in cancer: diagnostic, prognostic, predictive, and treatment implications. Cell Stress Chaperones 10, 86-103.

Cloonan, N., Forrest, A.R., Kolle, G., Gardiner, B.B., Faulkner, G.J., Brown, M.K., Taylor, D.F., Steptoe, A.L., Wani, S., Bethel, G., Robertson, A.J., Perkins, A.C., Bruce, S.J., Lee, C.C., Ranade, S.S., Peckham, H.E., Manning, J.M., McKernan, K.J. and Grimmond, S.M., 2008. Stem cell transcriptome profiling via massive-scale mRNA sequencing. Nat Methods 5, 613-9.

Curtis, D., Lehmann, R. and Zamore, P.D., 1995. Translational regulation in development. Cell 81, 171.

de Moor, C.H. and Richter, J.D., 2001. Translational control in vertebrate development, International Review of Cytology. Academic Press, pp. 567-608.

Dhillon, A.S. and Tulchinsky, E., 2015. FRA-1 as a driver of tumour heterogeneity: a nexus between oncogenes and embryonic signalling pathways in cancer. Oncogene 34, 4421-8.

Dreyfuss, G., Matunis, M.J., Pinolroma, S. and Burd, C.G., 1993. hnRNP Proteins and the Biogenesis of mRNA. Annual Review of Biochemistry 62, 289.

Eberhardt, W., Doller, A., Akool el, S. and Pfeilschifter, J., 2007. Modulation of mRNA stability as a novel therapeutic approach. Pharmacol Ther 114, 56-73.

Fujimoto, J., Aoki, I., Toyoki, H., Khatun, S. and Tamaya, T., 2002. Clinical implications of expression of ETS-1 related to angiogenesis in uterine cervical cancers. Annals of Oncology 13, 1598-1604.

Ganesh, S., Cole, B.S., Mallory, M.J. and Lynch, K.W., 2014. Transcriptome-wide RNA interaction profiling reveals physical and functional targets of hnRNP L in human T-cells. Molecular & Cellular Biology 34, 71.

Grudzinskas, J.G., Lenton, E.A., Gordon, Y.B., Kelso, I.M., Jeffrey, D., Sobowale, O. and Chard, T., 1977. Circulating levels of pregnancy-specific beta1-glycoprotein in early pregnancy. British Journal of Obstetrics & Gynaecology 84, 740.

Han, N., Li, W. and Zhang, M., 2013. The function of the RNA-binding protein hnRNP in cancer metastasis. Journal of Cancer Research & Therapeutics 9 Suppl, S129.

Horne, C.H., Towler, C.M. and Milne, G.D., 1977. Detection of pregnancy specific beta1-glycoprotein in formalin-fixed tissues. Journal of Clinical Pathology 30, 19.

Howley, B.V., Hussey, G.S., Link, L.A. and Howe, P.H., 2016. Translational regulation of inhibin betaA by TGFbeta via the RNA-binding protein hnRNP E1 enhances the invasiveness of epithelial-to-mesenchymal transitioned cells. Oncogene 35, 1725-35.

Hu, X., Chen, X., Wu, B., Soler, I.M., Chen, S. and Shen, Y., 2016. Further defining the critical genes for the 4q21 microdeletion disorder. American Journal of Medical Genetics Part A 173, 120.

Jin, L., Li, G., Yu, D., Huang, W., Cheng, C., Liao, S., Wu, Q. and Zhang, Y., 2017. Transcriptome analysis reveals the complexity of alternative splicing regulation in the fungus Verticillium dahliae. BMC Genomics 18.

Johnstone, O. and Lasko, P., 2001. Translational regulation and RNA localization in Drosophila oocytes and embryos. Annu Rev Genet 35, 365-406.

Kamei, D. and Yamada, M., 2002. Interactions of heterogeneous nuclear ribonucloprotein D-like protein JKTBP and its domains with high-affinity binding sites. Gene 298, 49-57.

Kawamura, H., Tomozoe, Y., Akagi, T., Kamei, D., Ochiai, M. and Yamada, M., 2002. Identification of the nucleocytoplasmic shuttling sequence of heterogeneous nuclear ribonucleoprotein D-like protein JKTBP and its interaction with mRNA. Journal of Biological Chemistry 277, 2732-9.

Kemmerer, K., Fischer, S. and Weigand, J.E., 2018. Auto- and cross-regulation of the hnRNPs D and DL. RNA 24, 324-331.

Kessler, C.A., Stanek, J.W., Stringer, K.F. and Handwerger, S., 2015. ETS1 induces human trophoblast differentiation. Endocrinology 156, 1851-9.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R. and Salzberg, S.L., 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol 14, R36.

Kim, J.H., Yang, C.K., Heo, K., Roeder, R.G., An, W. and Stallcup, M.R., 2008. CCAR1, a key regulator of mediator complex recruitment to nuclear receptor transcription complexes. Molecular Cell 31, 510.

Krämer, A., 1992. Purification of splicing factor SF1, a heat-stable protein that functions in the assembly of a presplicing complex. Molecular & Cellular Biology 12, 4545-4552.

Krämer, A., Quentin, M. and Mulhauser, F., 1998. Diverse modes of alternative splicing of human splicing factor SF1 deduced from the exon-intron structure of the gene. Gene 211, 29-37.

Krecic, A.M. and Swanson, M.S., 1999. hnRNP complexes: composition, structure, and function. Current Opinion in Cell Biology 11, 363-371.

Lai, D., Ho, K.C., Hao, Y. and Yang, X., 2011. Taxol resistance in breast cancer cells is mediated by the hippo pathway component TAZ and its downstream transcriptional targets Cyr61 and CTGF. Cancer Res 71, 2728-38.

Lee, M.H. and Schedl, T., 2006. RNA-binding proteins. WormBook, 1-13.

LH, M., CJ, C., BJ, S., WH, T. and SC, L., 1998. Identification of heterogeneous nuclear ribonucleoprotein K (hnRNP K) as a repressor of C/EBPbeta-mediated gene activation. The Journal of biological chemistry 273, 10784-91.

Livak, K.J. and Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25, 402-8.

MK, K. and VM, N., 1999. hnRNP U inhibits carboxy-terminal domain phosphorylation by TFIIH and represses RNA polymerase II elongation. Molecular and Cellular Biology 19, 6833-44.

Morin, R., Bainbridge, M., Fejes, A., Hirst, M., Krzywinski, M., Pugh, T., McDonald, H., Varhol, R., Jones, S. and Marra, M., 2008. Profiling the HeLa S3 transcriptome using randomly primed cDNA and massively parallel short-read sequencing. Biotechniques 45, 81-94.

Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L. and Wold, B., 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 5, 621-8.

Moshiri, A., Puppo, M., Rossi, M., Gherzi, R. and Briata, P., 2017. Resveratrol limits epithelial to mesenchymal transition through modulation of KHSRP/hnRNPA1-dependent alternative splicing in mammary gland cells. Biochim Biophys Acta 1860, 291-298.

Nagalakshmi, U., Wang, Z., Waern, K., Shou, C., Raha, D., Gerstein, M. and Snyder, M., 2008. The transcriptional landscape of the yeast genome defined by RNA sequencing. Science 320,

1344-9.

Omnus, D.J., Mehrtens, S., Ritter, B., Resch, K., Yamada, M., Frank, R., Nourbakhsh, M. and Reboll, M.R., 2011. JKTBP1 is involved in stabilization and IRES-dependent translation of NRF mRNAs by binding to 5' and 3' untranslated regions. Journal of Molecular Biology 407, 492-504.

Ou, C.Y., Kim, J.H., Yang, C.K. and Stallcup, M.R., 2009. Requirement of cell cycle and apoptosis regulator 1 for target gene activation by Wnt and beta-catenin and for anchorage-independent growth of human colon carcinoma cells. Journal of Biological Chemistry 284, 20629-37.

Piñol-Roma, S. and Dreyfuss, G., 1993. Cell cycle-regulated phosphorylation of the pre-mRNA-binding (heterogeneous nuclear ribonucleoprotein) C proteins. Molecular & Cellular Biology 13, 5762-70.

Puppo, M., Bucci, G., Rossi, M., Giovarelli, M., Bordo, D., Moshiri, A., Gorlero, F., Gherzi, R. and Briata, P., 2016. miRNA-Mediated KHSRP Silencing Rewires Distinct Post-transcriptional Programs during TGF-beta-Induced Epithelial-to-Mesenchymal Transition. Cell Rep 16, 967-78.

Qi, L. and Ding, Y., 2014. Involvement of the CREB5 regulatory network in colorectal cancer metastasis. Yi Chuan 36, 679-84.

Rishi, A.K., Zhang, L., Boyanapalli, M., Wali, A., Mohammad, R.M., Yu, Y., Fontana, J.A., Hatfield, J.S., Dawson, M.I. and Majumdar, A.P.N., 2003. Identification and Characterization of a Cell Cycle and Apoptosis Regulatory Protein-1 as a Novel Mediator of Apoptosis Signaling by Retinoid CD437. Journal of Biological Chemistry 278, 33422-35.

Robinson, M.D., McCarthy, D.J. and Smyth, G.K., 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139.

Salahshor, S., Goncalves, J., Chetty, R., Gallinger, S. and Woodgett, J.R., 2005. Differential gene expression profile reveals deregulation of pregnancy specific beta1 glycoprotein 9 early during colorectal carcinogenesis. BMC Cancer 5, 66.

Shen, S., Park, J.W., Lu, Z.X., Lin, L., Henry, M.D., Wu, Y.N., Zhou, Q. and Xing, Y., 2014. rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. Proc Natl Acad Sci U S A 111, E5593-601.

Shupert, W.L. and Chan, W.Y., 1993. Pregnancy specific beta 1-glycoprotein in human intestine. Molecular & Cellular Biochemistry 120, 159-70.

Takai, N., Ueda, T., Narahara, H. and Miyakawa, I., 2006. Expression of c-Ets1 protein in normal human placenta. Gynecologic & Obstetric Investigation 61, 15-20.

Tien, H., Maria, T. and Watson, D.K., 2004. Ets proteins in biological control and cancer. Journal of Cellular Biochemistry 91, 896-903.

Trojanowska, M., 2000. Ets factors and regulation of the extracellular matrix. Oncogene 19, 6464.

Tsuchiya, N., Kamei, D., Takano, A., Matsui, T. and Yamada, M., 1998. Cloning and characterization of a cDNA encoding a novel heterogeneous nuclear ribonucleoprotein-like protein and its expression in myeloid leukemia cells. Journal of Biochemistry 123, 499-507.

van Kempen, L.C. and Coussens, L.M., 2002. MMP9 potentiates pulmonary metastasis formation. Cancer Cell 2, 251-2.

Veldscholte, J., Berrevoets, C.A., Brinkmann, A.O., Grootegoed, J.A. and Mulder, E., 1992. Anti-androgens and the mutated androgen receptor of LNCaP cells: differential effects on binding affinity, heat-shock protein interaction, and transcription activation. Biochemistry 31, 2393-9.

Venables, J.P., 2004. Aberrant and Alternative Splicing in Cancer. Cancer Research 64, 7647-7654.

Venables, J.P., Klinck, R., Koh, C.S., Gervaisbird, J., Bramard, A., Inkel, L., Durand, M., Couture, S., Froehlich, U. and Lapointe, E., 2009. Cancer-associated regulation of alternative splicing. Nature Structural & Molecular Biology 16, 670.

Vera, J.C., Wheat, C.W., Fescemyer, H.W., Frilander, M.J., Crawford, D.L., Hanski, I. and Marden, J.H., 2008. Rapid transcriptome characterization for a nonmodel organism using 454 pyrosequencing. Mol Ecol 17, 1636-47.

Vieira, N.M., Naslavsky, M.S., Licinio, L., Kok, F., Schlesinger, D., Vainzof, M., Sanchez, N., Kitajima, J.P., Gal, L. and Cavaçana, N., 2014. A defect in the RNA-processing protein HNRPDL causes limb-girdle muscular dystrophy 1G (LGMD1G). Human molecular genetics 23, 4103.

Weighardt, F., Biamonti, G. and Riva, S., 1996a. The roles of heterogeneous nuclear ribonucleoproteins (hnRNP) in RNA metabolism. Bioessays 18, 747-756.

Weighardt, F., Biamonti, G. and Riva, S., 1996b. The roles of heterogeneous nuclear ribonucleoproteins (hnRNP) in RNA metabolism. Bioessays 18, 747–756.

Wickens, M., Goodwin, E.B., Kimble, J., Strickland, S. and Hentze, M.W., 2000. Translational Control of Developmental Decisions. Sonenberg N.translational Control of Gene Expression.cold 39.

Wu, Y.Y., Li, H., Lv, X.Y., Wei, Q., Li, X., Liu, X.Y., Zhou, Q. and Wei, Y.Q., 2008. Overexpression of JKTBP1 induces androgen‐independent LNCaP cell proliferation through activation of epidermal growth factor‐receptor (EGF‐R). Cell Biochemistry & Function 26, 467-477.

Xia, H., Chen, D., Wu, Q., Wu, G., Zhou, Y., Zhang, Y. and Zhang, L., 2017. CELF1 preferentially binds to exon-intron boundary and regulates alternative splicing in HeLa cells. Biochim Biophys Acta 1860, 911-921.

Xie, C., Mao, X., Huang, J., Ding, Y., Wu, J., Dong, S., Kong, L., Gao, G., Li, C.Y. and Wei, L., 2011. KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. Nucleic Acids Res 39, W316-22.

Xu, Y., Gao, X.D., Lee, J.H., Huang, H., Tan, H., Ahn, J., Reinke, L.M., Peter, M.E., Feng, Y., Gius, D., Siziopikou, K.P., Peng, J., Xiao, X. and Cheng, C., 2014. Cell type-restricted activity of hnRNPM promotes breast cancer metastasis via regulating alternative splicing. Genes Dev 28, 1191-203.

Yeger, H. and Perbal, B., 2007. The CCN family of genes: a perspective on CCN biology and therapeutic potential. J Cell Commun Signal 1, 159-64.

Yu, S.J., Yu, J.K., Ge, W.T., Hu, H.G., Yuan, Y. and Zheng, S., 2011. SPARCL1, Shp2, MSH2, E-cadherin, p53, ADCY-2 and MAPK are prognosis-related in colorectal cancer. World J Gastroenterol 17, 2028-36.

Highlights

1. Transcriptome-wide study of the targets of hnRNPDL regulation in HeLa cells
2. hnRNPDL targets gene expression in cell apoptosis, proliferation and migration.
3. HnRNPDL regulates alternative splicing of genes in several signalling pathways.

Abbreviations list

RBP RNA Binding Protein

hnRNP Heterogeneous nuclear ribonucleoproteins

hnRNP K Heterogeneous Nuclear Ribonucleoprotein K

hnRNPDL Heterogeneous Nuclear Ribonucleoprotein D Like

JKTBP JKT41-Binding Protein

hnRNP D Heterogeneous Nuclear Ribonucleoprotein D

NRF NF-κB Repressing Factor

NF-κB Nuclear Factor Kappa B

RNA-Seq RNASequencing

qRT-PCR Quantitative Reverse Transcription Polymerase Chain Reaction

FPKM fragments per kilo base of exon model per million fragments mapped

KD knockdown

FC fold change

FDR false discovery rate

DEG differentially expressed gene

GO Gene ontology

KEGG Kyoto Encyclopedia of Genes and Genomes

CTGF Connective Tissue Growth Factor

HSPA1A Heat Shock Protein Family A (Hsp70) Member 1A

CYR61 cysteine rich-61

FOSL1 Fos Proto-Oncogene Like 1

PHLDA1 Pleckstrin Homology Like Domain Family A Member 1

ETS1 V-Ets Avian Erythroblastosis Virus E26 Oncogene homolog 1

ARNT2 Aryl Hydrocarbon Receptor Nuclear Translocator 2

ODC1 Ornithine Decarboxylase 1

AS alternative splicing

ASE alternative splicing event

RASE regulated alternative splicing event

A3SS alternative 3′splice site

A5SS alternative 5′splice site

ES exon skipping

5pMXE mutually exclusive 5′UTR

3pMXE mutually exclusive 3′UTR

MXE mutually exclusive exon

IR intron retention

A5SS & ES alternative 5′splice site & exon skipping

A3SS & ES alternative 3′splice site & exon skipping

CCAR1 Cell Division Cycle and Apoptosis Regulator 1

IMMP1L Inner Mitochondrial Membrane Peptidase Subunit 1

TNFRSF12A TNF Receptor Superfamily Member 12A

RHOT1 Ras Homolog Family Member T1

MINK1 Misshapen Like Kinase 1

PUM2 Pumilio RNA Binding Family Member 2

NGLY1 N-Glycanase 1

NDEL1 NudE Neurodevelopment Protein 1 Like 1

CEP63 Centrosomal Protein 63

NR2C2 Nuclear Receptor Subfamily 2 Group C Member 2

EPB41 Erythrocyte Membrane Protein Band 4.1

SF1 Splicing Factor 1

ZNF200 Zinc Finger Protein 200

B4GALT2 Beta-1, 4-Galactosyltransferase 2

C11orf73 chromosome 11 open reading frame 73

DNPH1 2'-Deoxynucleoside 5'-Phosphate N-Hydrolase 1

hnRNP U Heterogeneous Nuclear Ribonucleoprotein U

PSG pregnancy specific glycoprotein

ENDOU endonuclease, poly U-specific

FOSB FBJ murine osteosarcoma viral oncogene homolog B

EMT epithelial-mesenchymal transition

MET mesenchymal-epithelial transition

NVO nephroblastoma overexpressed protein

TAZ Tafazzin

NOD nucleotide-binding oligomerization domain

TNF tumor necrosis factor

EDTA Ethylenediaminetetraacetic acid

SDS sodium dodecyl sulfate

BCA bicinchoninic acid