Prevalence and architecture of posttranscriptionally impaired synonymous mutations in 8,320 genomes across 22 cancer types

Huajing Teng[1,2,†], Wenqing Wei[1,3,†], Qinglan Li[1,3], Meiying Xue[1,3], Xiaohui Shi[4], Xianfeng Li[1,2], Fengbiao Mao [1,*] **and Zhongsheng Sun** [1,*]

[1]Beijing Institutes of Life Science, Chinese Academy of Sciences, Beijing 100101, China, [2]Key laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), Department of Radiation Oncology, Peking University Cancer Hospital & Institute, Beijing 100142, China, [3]University of Chinese Academy of Sciences, Beijing 100049, China and [4]Sino-Danish college, University of Chinese Academy of Sciences, Beijing 100049, China

ABSTRACT

Somatic synonymous mutations are one of the most frequent genetic variants occurring in the coding region of cancer genomes, while their contributions to cancer development remain largely unknown. To assess whether synonymous mutations involved in post-transcriptional regulation contribute to the genetic etiology of cancers, we collected whole exome data from 8,320 patients across 22 cancer types. By employing our developed algorithm, PIVar, we identified a total of 22,948 posttranscriptionally impaired synonymous SNVs (pisSNVs) spanning 2,042 genes. In addition, 35 RNA binding proteins impacted by these identified pisSNVs were significantly enriched. Remarkably, we discovered markedly elevated ratio of somatic pisSNVs across all 22 cancer types, and a high pisSNV ratio was associated with worse patient survival in five cancer types. Intriguing, several well-established cancer genes, including *PTEN*, *RB1* and *PIK3CA*, appeared to contribute to tumorigenesis at both protein function and posttranscriptional regulation levels, whereas some pisSNV-hosted genes, including *UBR4*, *EP400* and *INTS1*, exerted their function during carcinogenesis mainly via posttranscriptional mechanisms. Moreover, we predicted three drugs associated with two pisSNVs, and numerous compounds associated with expression signature of pisSNV-hosted genes. Our study reveals the prevalence and clinical relevance of pisSNVs in cancers, and emphasizes the importance of considering post-transcriptional impaired synonymous mutations in cancer biology.

INTRODUCTION

Somatic synonymous mutations, which do not alter the protein sequences of their host genes (1), are one of the most frequent but rarely investigated genetic changes that occur in the coding regions of human cancer genomes (2). Recent studies have shown that they can act as drivers of cancers by altering RNA splicing, RNA stability and protein translation (3,4), which suggests the existence of uncovered regulatory effects of these 'silent' mutations and highlights the significance to adjust our focus beyond the damaging protein-coding mutations (5). Thus, to provide a comprehensive landscape of cancer genome alterations, it is imperative but still challenging to decipher the role of these 'silent' mutations genome-wide in pathogenesis of cancers.

RNA-binding proteins (RBPs) play versatile roles in posttranscriptional RNA regulation, including splicing, polyadenylation, mRNA stabilization, RNA structure, subcellular localization and transcription (6–9), and their aberrant expression may lead to chaos within the whole regulation network (10). In several cancer types, the abnormal expression of RBPs was associated with patient prognosis (11,12), and numerous mutations occurred in coding regions of RBPs had been implicated in tumorigenesis and progression, such as KHDRBS1 (13) and ELAVL1 (also known as HuR) (14). Interaction between known cancer driver genes and dysregulation of RBPs were also demonstrated. For example, the oncogenic transcription factor c-Myc upregulated the transcription of three heterogeneous nuclear ribonucleoprotein (hnRNP) proteins to control pyruvate kinase mRNA splicing in cancer cells (15). Additionally, in cancer cells, p53 tumor suppressor protein

induced the expression of RNA-binding motif 38 (RBM38), which regulated the stability of targeted mRNAs to promote cell cycle arrest (16). These studies have unveiled the indispensable function of RBPs in carcinogenesis and progression of cancers (17).

It has been demonstrated that RBPs can recognize their RNA substrate via sequence-specific binding motifs (18); therefore, genetic mutations occurred in the binding motifs may disrupt the recognition between RBPs and RNA substrates, resulting in various human diseases (19). For example, a mutation in the 3′ untranslated region of *FMR1* decreased neuronal activity-dependent translation of FMRP by disrupting the binding of HuR, leading to developmental delay in patients (20). However, the effect of synonymous single nucleotide variants (SNVs) on RBP-mediated posttranscriptional regulation in human cancers remains unclear.

To test whether synonymous mutations involved in posttranscriptional regulation contribute to the genetic etiology of cancers, we collected whole exome data from 8,320 patients across 22 cancer types. By employing our developed algorithm, PIVar, we identified a substantial number of posttranscriptionally impaired synonymous SNVs (pisSNVs) and observed the clinical relevance of the somatic pisSNV ratio in 8,320 patients across 22 cancer types. The functional effect of these pisSNVs and their host genes, as well as significantly altered subnetworks containing pisSNV-hosted genes, were further analyzed for their co-occurrence and relative contribution to the etiology of cancers.

MATERIALS AND METHODS

Pipeline for detecting posttranscriptionally impaired SNVs (piSNVs)

To evaluate the impact of mutations on posttranscriptional regulation, we developed a heuristic scoring system, PIVar (https://github.com/WeiWenqing/PIVar), which is inspired by RegulomeDB (21) and centered on the disruption of a protein-RNA interaction via alteration of RNA secondary structure and regulation of gene expression, to identify piSNVs. Firstly, we identified the putative regulatory SNV set as those situated in RBP-binding sites detected by CLIP-seq (crosslinking immunoprecipitation sequencing). Then, functional confidence of specific regulatory SNV was categorized based on their impact on RNA expression, RBP binding, alterations of RNA secondary structure (namely riboSNitch) and miRNA binding (Figure 1A, Supplementary Table S1).

In details, expression quantitative trait loci (eQTLs) and RBP binding sites of 112 CLIP-seq datasets derived from 26 human cell lines or tissues were collected from starBase (22), CLIPdb (23) and extracted sequencing data from Gene Expression Omnibus (GEO) (24). As the RNA-binding domains initially determined the binding specificity and preferences of RBPs on RNA substrate via specific sequence motifs (25), all position weight matrices (PWMs) from the catalog of inferred sequence binding preferences of RBPs (25,26) deposited in the AURA database (27) were used to call motif (match score > 0 and *P*-values < 0.0001) across

human transcriptome. For single-base mutations, we employed the RNAsnp (28) with default parameters to estimate the mutation effects on local RNA secondary structure. For insertions and deletions, we evaluated their effects on RNA secondary structure using the minimal free energy generated by RNAfold (29). In addition, we employed LS-GKM (30) and deltaSVM (31) to predict the impact of SNVs on the binding of specific RBPs by calculating the delta SVM scores. Only the SNVs that produced a change of >5 in the gkm-SVM scores, *P*-values < 0.05 and free energy changes >1 were regarded as SNVs who are likely to affect the binding of RBP. Moreover, we collected response elements for all human microRNAs curated in miRanda (32) and TargetScan v7.0 (33). The mutation effects of specific SNV on microRNA binding were predicted using miR-NASNP v2 (34). Finally, SNVs with functional categories 1–2 listed in Supplementary Table S1 were determined to be piSNVs.

To evaluate the efficiency of our workflow, based on the downloaded RNA-seq and CLIP-seq (103 RBPs) data of HepG2 cell line from the ENCODE (Encyclopedia of DNA Elements) database, we compared the piSNV and pisSNVs we identified with allele-specific RBP-RNA interaction sites predicted by ASPRIN (35), which inferred the interactions by observing the allelic preference of RBPs from CLIP-seq as well as RNA-seq experimental data.

Identification of pisSNVs from somatic mutation data of 22 cancer types and human normal cells

We downloaded the somatic mutation data of 8,320 patients across 22 cancer types, including skin cutaneous melanoma (SKCM), lung squamous cell carcinoma (LUSC), lung adenocarcinoma (LUAD), esophageal carcinoma (ESCA), head and neck squamous cell carcinoma (HNSC), bladder urothelial carcinoma (BLCA), cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC), stomach adenocarcinoma (STAD), rectum adenocarcinoma (READ), glioblastoma multiforme (GBM), liver hepatocellular carcinoma (LIHC), uterine corpus endometrioid carcinoma (UCEC), colon adenocarcinoma (COAD), kidney renal clear cell carcinoma (KIRC), pancreatic adenocarcinoma (PAAD), sarcoma (SARC), kidney renal papillary cell carcinoma (KIRP), ovarian serous cystadenocarcinoma (OV), breast invasive carcinoma (BRCA), brain lower grade glioma (LGG), prostate adenocarcinoma (PRAD), thyroid carcinoma (THCA), from the Broad Institute/The Cancer Genome Atlas (TCGA) data portal (http://gdac. broadinstitute.org/). Taking samples from the DSMNC database (Database of Somatic Mutations in Normal Cells) (36) as a control, we got 0.77 million somatic SNVs occurring in over 579 human normal cells from the DSMNC database. The piSNVs and pisSNVs that could destroy the binding between RNA and RBPs were identified from all these somatic mutation data using PIVar, and the distribution of pisSNV ratio (ratio of pisSNV to synonymous SNV) between control and each cancer type was compared using Wilcoxon rank-sum test. For a given sample, the ratio of posttranscriptionally impaired non-synonymous SNV (pinsSNV) to non-synonymous SNV was also calculated, and then the distribution of pinsSNV ratio within each can-
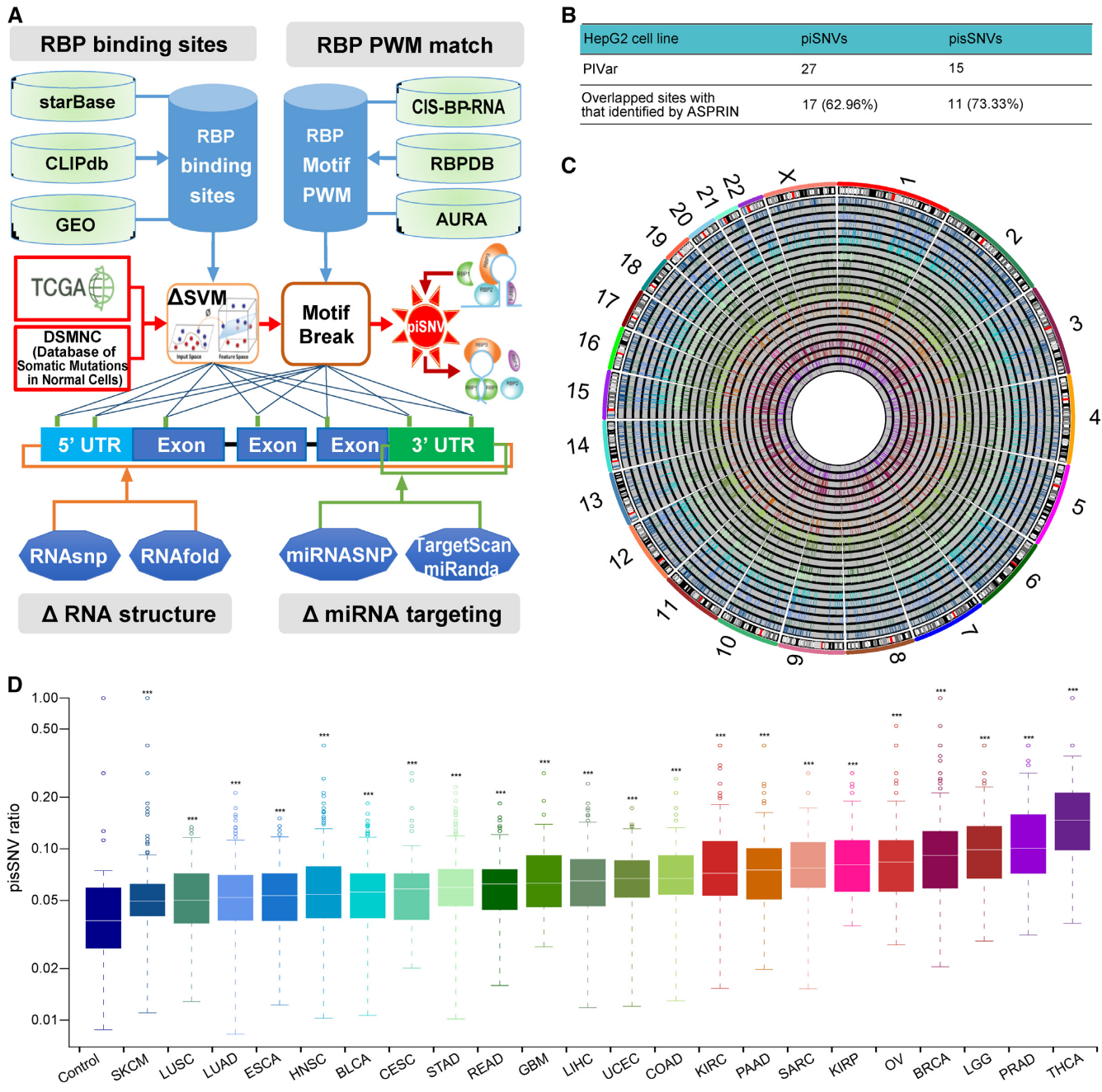
Figure. 1 Posttranscriptional impaired synonymous SNVs (pisSNVs) identified in TCGA pan-cancer. (**A**) Workflow for identifying posttranscriptionally impaired SNVs. (**B**) Evaluation of the impact of piSNVs identified by PIVar on posttranscriptional regulation through allele-specific binding activity (inferred by ASPRIN (35)) of 103 RBPs based on the CLIP-seq and RNA-seq data of the HepG2 cell line. (**C**) Genome-wide distribution of pisSNVs identified in 22 TCGA cancer types. The circle adjacent to the karyotypes and the innermost circle show lines representing the distribution of pisSNVs identified in SKCM and THCA, respectively. Other circles from outermost to innermost are arranged according to the order of cancer types listed in (**D**) (from left to right). (**D**) Elevated ratio of somatic pisSNVs in TCGA pan-cancer compared with that of control from the DSMNC database (*** $P<0.001$). For a given sample, the ratio of pisSNV to synonymous SNV was calculated, and then the distribution of pisSNV ratio within each cancer type was compared with that in control samples using Wilcoxon rank-sum test.

cer type was compared with that in pisSNV ratio using Wilcoxon rank-sum test.

Clinical survival analysis

We downloaded corresponding clinical data of each cancer type from cBioPortal (http://www.cbioportal.org/). Then, the function 'surv_cutpoint' from R package 'survminer' was applied to determine the optimal pisSNV ratio based on disease-free survival (DFS) information of the patients. DFS in patients with high or low pisSNV ratio was compared using the Kaplan–Meier method and the log-rank test. Multivariate Cox proportional regression analysis accounting for age, gender, TNM stage and race of patient was also performed to analyze the association between pisSNV ratio and clinical outcome.

RBP motif enrichment of pisSNV loci and their impact on RBP binding

Many RBPs interact with mRNAs via a limited set of modular RNA-binding domains (37), which initially determine the specificity and preferences of RNA binding with specific sequence motifs (25). Therefore, 247 PWMs from the catalog of the inferred RNA binding motif deposited in the AURA database (27) were used to call motif matches in the transcribed regions of the genome. After the integrative analysis of the RBP motif and CLIP-seq-derived RBP binding for all identified pisSNVs, Fisher exact tests were used to explore the RBPs impacted by pisSNVs (FDR < 0.05 and OR > 1). To further evaluate the impact of pisSNVs on RBP binding, the crystal structures of the RBP-RNA complex (PDB ID: 5EN1 and 2LEC) were downloaded from the Protein Data Bank (https://www.rcsb.org/) and visualized by PyMOL (https://pymol.org).

Chemiluminescent electrophoresis mobility shift assays

To assess the effect of RNA mutation on binding of RBP, electrophoresis mobility shift assays were performed using LightShift Chemiluminescent RNA EMSA kit (Catalog # 20158; Thermo Scientific, Rockford, USA). Two purified RBPs were used in the assay. Of which, PCBP3 was purchased from OriGene Technologies (Catalog # TP329176; Rockville, USA), and PTBP1 was kindly provided by Dr Yuanchao Xue (Institute of Biophysics, Chinese Academy of Sciences, Beijing, China). Then, 200 ng RBP was pre-incubated with 100 μg ml^{-1} tRNA in 1× RNA EMSA binding buffer for 10 minutes at room temperature. After that, 80 fmol synthesized 3′-biotin-labeled wild-type or point-mutated RNA oligos (Supplementary Table S2) were respectively added to the mixture (15 μl final volume) and incubated for 20 minutes at room temperature. Then, 3.75 μl 5× loading buffer was added to the 15 μl RBP-RNA mixture and immediately loaded into the pre-run TBE polyacrylamide gel, and ran at 100 V for 45–60 min in cooled 0.5× TBE buffer. Samples were then transferred to positively charged nylon membrane (Thermo Scientific, Rockford, USA), and crosslinked with UV-light crosslinking instrument equipped with 254 nm bulbs. The subsequent blocking, washing and detection were performed according to the manufacturer's instructions.

Differential expression analysis

We downloaded the RNA-seq read count data of 22 cancer types and corresponding normal tissues from the Broad Institute. The R package 'DESeq2' was used to assay the expression of 2,042 pisSNV-hosted genes and 35 enriched RBPs identified in the previous step, and only genes with fold-change >2 and FDR <0.05 were considered to be differentially expressed genes.

Identification of co-existed protein damaged mutations in pisSNV-hosted genes

All genetic mutations occurred in 2,042 pisSNV-hosted genes across 22 cancer types were curated, and their functional consequences on protein were predicted using Varcards (38). Protein damage mutations were then defined by loss-of-function mutations or missense mutations with a damaging score ≥0.8 predicted using ReVe (39). After that, gene-length corrected occurrence frequency of these genes with protein damaging mutations and posttranscriptional impaired synonymous mutations in all 8,320 cancer patients were analyzed.

Network analysis

We used the HotNet2 workflow (40) to mine significantly mutated subnetworks according to the literature-based Human Protein Reference Database (HPRD) network (http://www.hprd.org/), and the sample frequency of each pisSNV-hosted gene was taken as the network heat score to identify significant subnetworks with default parameters.

KEGG enrichment analysis

To explore the function of pisSNV-hosted gene within each subnetworks, we analyzed KEGG enrichment using the R package 'clusterProfiler' (41), with a Bonferroni correction test, and identified significant pathways with FDR values <0.05.

Potential clinical drug analysis

We downloaded somatic mutation data for 1,001 cancer cell lines and natural log IC50 of the dose–response curve for all screened cell line/drug combinations from the Genomics of Drug Sensitivity in Cancer (http://www.cancerrxgene.org/). Next, we used analysis of variance (ANOVA) to identify the associations between posttranscriptionally impaired synonymous mutations and drug responses according to previous publication (42). Cohen's *d* was used to quantify the effect size, and the resulting *P* values were corrected by FDR. To further explore the therapeutic effects of pisSNV-hosted genes, the gene expression profiles of each identified pisSNV-hosted gene in each cancer type were compared with drug response signatures listed in the Connectivity Map (CMAP) build 02 (Broad Institute) (43).

RESULTS

Pipeline for detecting posttranscriptionally impaired SNVs (piSNVs)

To investigate the potential impact of genomic mutations on posttranscriptional regulation, we developed PIVar according to the functional confidence of variants based on multi-omic experimental data (Figure 1A). As a pilot study, we first analyzed the mutation data of HepG2 cell line from the ENCODE database using PIVar, and identified 27 piSNVs and 15 pisSNVs in the cell line. A recently developed computational method, ASPRIN (35), could infer RBP–RNA interactions by observing the allelic preference of RBPs from CLIP-seq as well as RNA-seq experimental data, which provided us a method to evaluate the efficiency of our workflow. We used it to analyze allele-specific binding of 103 RBPs based on the CLIP-seq and RNA-seq data from the same cell line, and identified 987 allele-specific RBP–RNA interaction sites in the exon regions. Seventeen (62.96%) piSNVs and 11 (73.33%) pisSNVs obtained through our pipeline were overlapped with the allele-specific RBP–RNA interaction sites identified by ASPRIN (Figure 1B; Supplementary Table S3), which suggests that PIVar was more stringent for identifying the impact of genetic mutations on posttranscriptional regulation network.

Elevated ratio of somatic pisSNVs across cancer types

Inspired by previous studies in which genetic mutations can disrupt the RBP recognition of RNA substrates (20,44) and many RBPs play important roles in tumorigenesis (11–16,35), we then employed PIVar (Figure 1A) to analyze the somatic mutation spectrum of 22 cancer types to explore the correlation between mutations and binding of RBPs. In total, we identified 98,260 nonredundant piSNVs across 22 cancer types that could destroy the binding between mRNA and the corresponding RBP. Synonymous mutations can function as driver mutations in human cancers by disrupting RNA splicing or RBP binding instead of altering the sequence of encoded proteins directly (4); thus, we focused on the previously neglected 'silent' mutations and observed a total of 22,948 synonymous piSNVs (pisSNVs) across 22 cancer types (Figure 1C; Supplementary Table S4). Taking samples from the DSMNC database as a control, we observed a significantly higher ratio of pisSNVs in TCGA pan-cancer compared with that of the control (Figure 1D). Also, we observed synonymous mutations had a higher mutation load on posttranscriptional regulation comparing to non-synonymous mutations in 15 cancer types (Supplementary Figure S1). The significant distinction between cancer and normal control samples unveiled the prevalence of posttranscriptionally impaired synonymous mutations in cancer genomes, implying their contribution to cancer etiology.

Clinical relevance of the pisSNV ratio across cancer types

To further identify the clinical relevance of the elevated pisSNV ratio in cancers, we stratified the patients in each cancer type according to the ratio of pisSNV to synonymous SNVs identified in a given sample. We found that DFS of patients with seven cancer types, namely, BLCA,

CESE, LUAD, OV, SKCM, STAD and THCA, had a significant correlation with the pisSNV ratio (Figure 2; Supplementary Figure S2), and the patients with a high pisSNV ratio in each cancer type had a worse survival situation than those with a low pisSNV ratio. Additionally, the association between the increased ratio of pisSNV and the DFS of cancer patients remained statistically significant in five cancer types, namely, BLCA, LUAD, OV, SKCM and STAD, even after adjusting for age, gender and TNM stage of patient in the multivariate Cox regression analysis. These results further revealed the contribution of pisSNVs to survival of cancer patients and provided clinical evidence that the pisSNV ratio could serve as a potential prognostic biomarker for several types of cancers.

Functional effect of pisSNVs on RBP binding

After obtaining the association between clinical outcome and pisSNV ratio, we wanted to determine the functional effect of these pisSNVs on posttranscriptional regulation, especially on RBP binding. It has been demonstrated that genetic mutations on RNA substrates can disrupt the RBP recognition toward them (20). To provide experimental evidence that pisSNVs disrupt the binding of RBPs, we performed EMSA on three randomly selected pisSNVs who predicted to alter the binding of PCBP3 or PTBP1 (Figure 3A, B; Supplementary Figure S3). As shown in Figure 3A, the binding of PCBP3 to unmutated RNA probe of *DAB2,* a tumor suppressor by dictating the TGF-β responses of tumor cells (45), was stronger, while the probe with the pisSNV occurring in *DAB2* showed visible differences in their binding to PCBP3. Similar results were also found in pisSNV of *ZFHX3* or *USP9X* on binding of PTBP1 (Figure 3B; Supplementary Figure S3).

In addition, our analyses of the crystal structure of the RBP-RNA complex supported the effect of some identified pisSNVs on RBP bindings. For example, a pisSNV that occurs in the gene *ADCY7* might influence the binding with hnRNP A2/B1, a key player in the posttranscriptional regulation of the maturation, transport, and metabolism of long noncoding RNAs/mRNA (46) and an essential regulator in the development and progression of breast cancer (47). Specifically, the co-crystal structure of the hnRNPA2B1-RNA complex (PDB ID: 5EN1) revealed that hnRNPA2B1 could bind to 5′-AGGACUG-3′ RNA oligonucleotide (46) (Figure 3C). By screening the identified pisSNVs, we found that *ADCY7*, whose expression significantly correlated with the overall survival of acute myeloid leukemia patients (48), had a mutation from G to A at the second position of the above motif. Another RBP affected by a pisSNV is SRSF2, the RBP that plays an important role in the regulation of alternative splicing events (49), and it was demonstrated that SRSF2 could bind to the 5′-UGGAGU-3′ RNA oligonucleotide (PDB ID: 2LEC) (50) (Figure 3D). We inferred that two pisSNVs in UGGAGU sequence motif residing in *DAB2* and *ZFHX3* could destroy the binding between SRSF2 and their corresponding RNAs.

To gain an overall insight into the impact of these pisSNVs on RBP binding, we performed motif and binding site enrichment analysis. The integrative analysis of the RBP
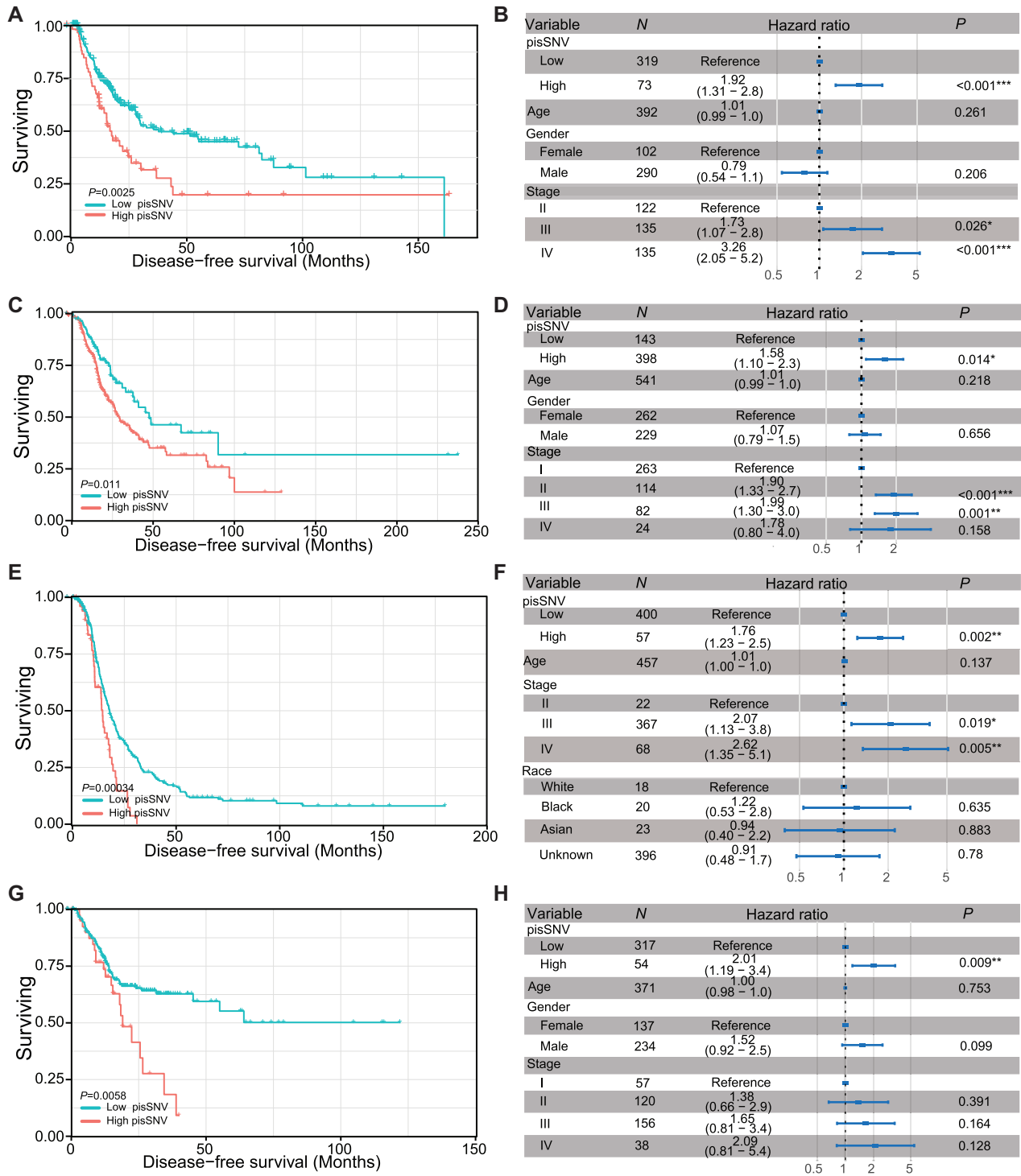
Figure 2. Disease-free survival related to the pisSNV ratio across cancer types. Subjects from the BLCA (**A**, **B**), LUAD (**C**, **D**), OV (**E**, **F**) and STAD (**G**, **H**) cohort were stratified according to the pisSNV ratio. Multivariate Cox regression analysis of the pisSNV ratio after taking into account age, gender and TNM stage of patients.
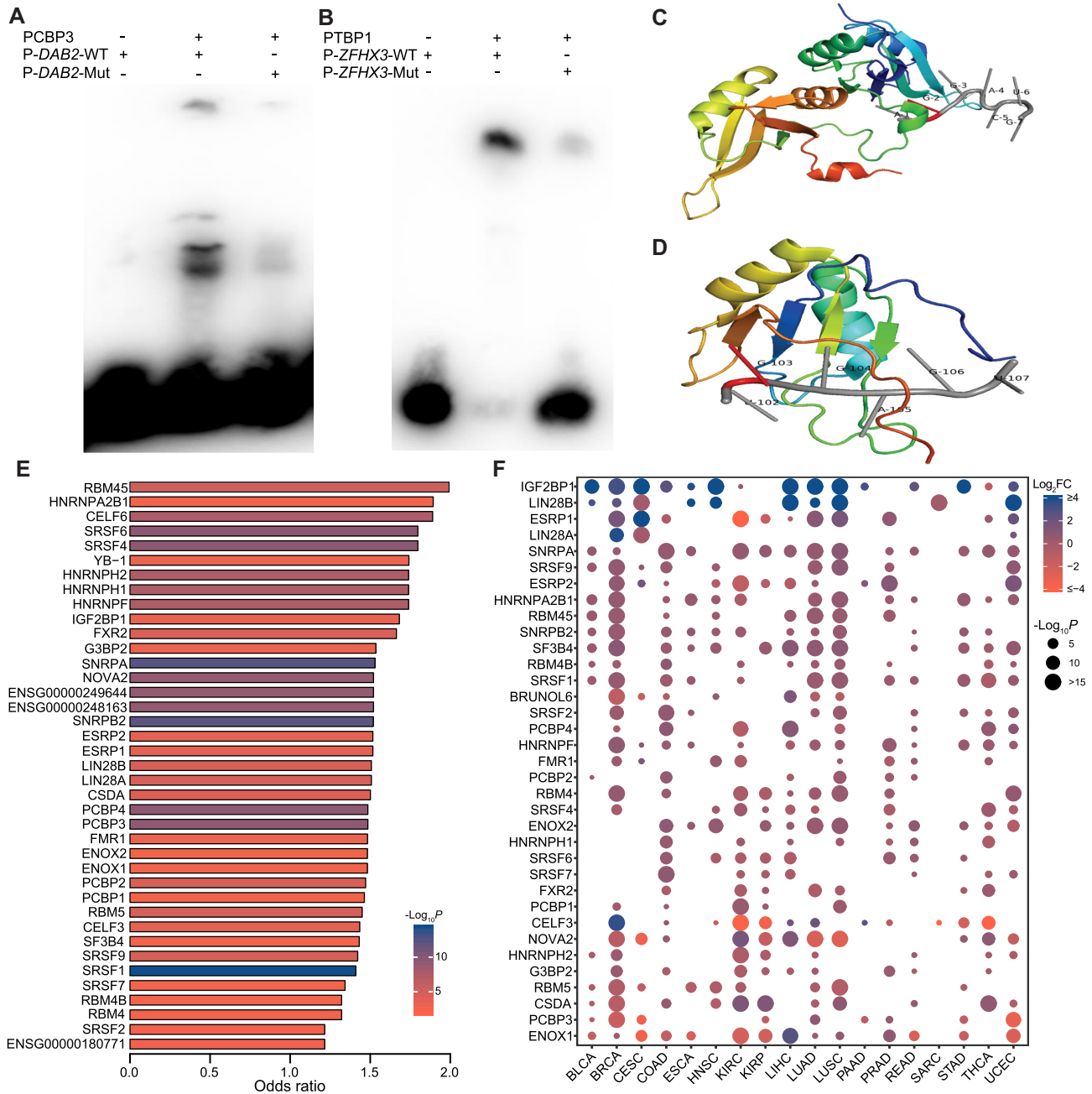
Figure 3. Functional effects of pisSNVs on RBP binding. Electrophoretic mobility shift assays (EMSA) results show the impact of pisSNVs on the recognition of PCBP3 (**A**) and PTBP1 (**B**) to their RNA targets. The co-crystal structure of the hnRNPA2B1 (**C**; PDB ID: 5EN1) or SRSF2 (**D**; PDB ID: 2LEC)-RNA complex displayed the binding locations of certain RBPs on specific RNA motifs. (**E**) RBP motif enrichment and CLIP-seq-derived RBP binding site analysis of pisSNVs identified 35 significantly enriched RBPs. (**F**) Expression of RBPs in different cancer types.

motif and CLIP-seq-derived RBP binding for all identified pisSNVs revealed 35 significantly enriched RBPs (Figure 3E; Supplementary Table S5). As previous studies demonstrated that abnormal expression of RBPs contributed to cancer development (11,12), we screened the differential expression values of these RBPs between tumor and corresponding normal tissues in different cancer types (Figure 3F). The expression data suggest that dysregulation of these RBPs occurs ubiquitously in cancers. For example,

IGF2BP1, a posttranscriptional regulator required for tumor cell proliferation, invasion, and chemoresistance (51), was differentially upregulated in 14 cancer types. ENOX1, a critical mediator of endothelial cell radiosensitization and a potential cancer therapy target (52), was differentially downregulated in 10 cancer types. Collectively, these data provide evidence to support the hypothesis that the identified pisSNVs contribute to cancer pathogenesis by affecting the binding of some cancer-associated RBPs.
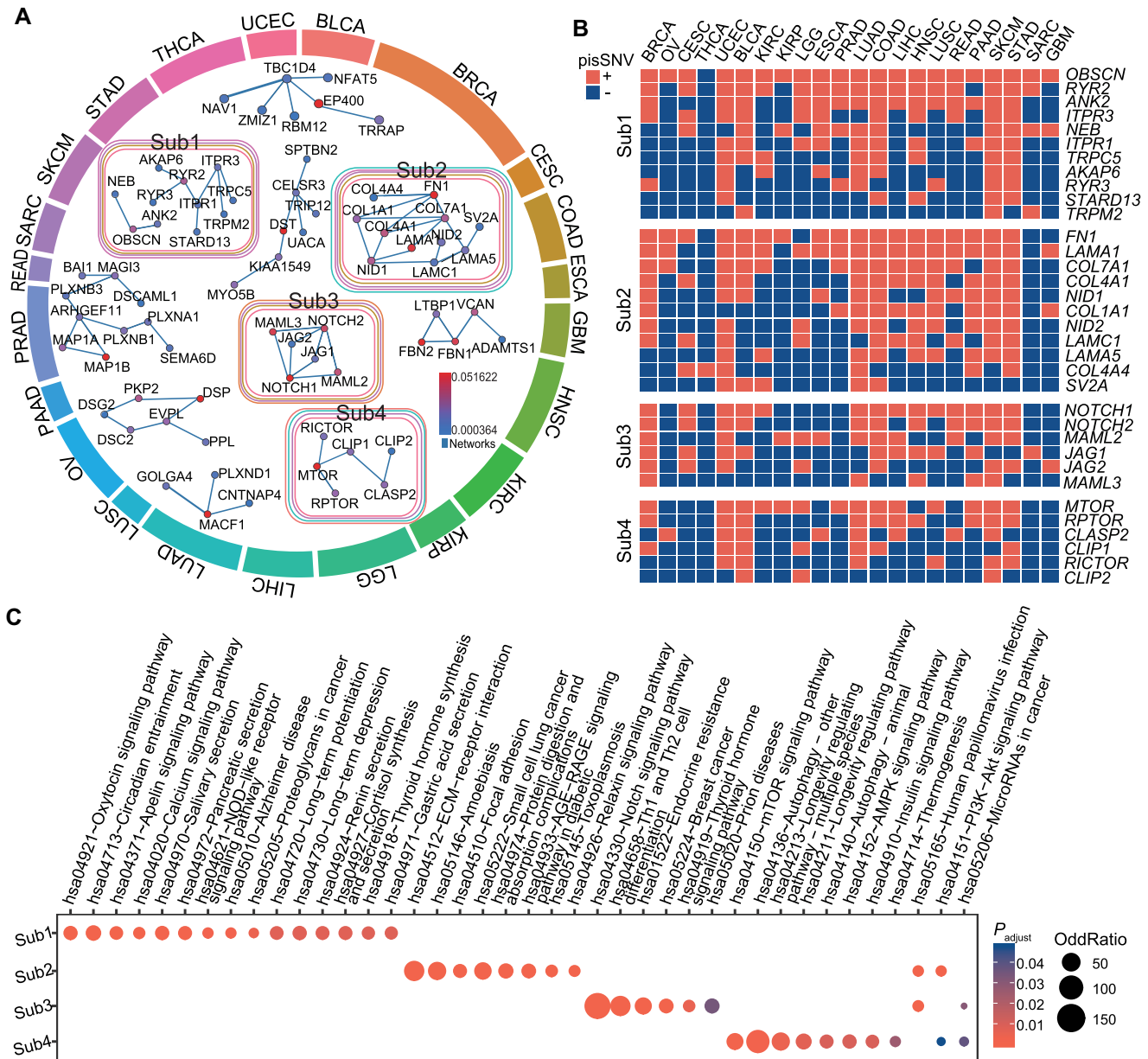
Figure 4. The co-occurrence and expression of pisSNV-hosted genes in TCGA pan-cancer. The co-occurrence (**A**) and expression (**B**) of the top 5% (102) gene-length corrected frequently occurred pisSNV-hosted genes in different cancer types. (**C**) The scatter plot shows the gene-length corrected occurrence frequency of genes with posttranscriptional impaired synonymous mutations and protein damaging mutations.

Figure 5. Significantly altered subnetworks and pathways containing pisSNV-hosted genes. (**A**) HotNet2 consensus subnetworks are arranged according to the enriched cancer types. Colored outlines surrounding each network indicate the cancer types that are enriched, with the color corresponding to the cancer types of the outmost circle. Protein interactions between a subnetwork are derived from the HPRD networks. The co-occurrence (**B**) and enriched pathways (**C**) of the pisSNV-hosted genes from different subnetworks.

The co-occurrence of pisSNV-hosted genes in TCGA pan-cancer

To explore the influence of pisSNVs at the gene level, we summarized the somatic pisSNVs obtained in each cancer type and obtained 2,042 nonredundant pisSNV-hosted genes. A large proportion of pisSNV-hosted genes occurred in a cancer-specific manner, and the top three gene-length corrected frequently occurred pisSNV-hosted genes were *LRP1B*, *LAMA1* and *HERC2*, which occurred in at least 17 cancer types (Figure 4A). In total of 8,320 patients, the percentage of patients with pisSNVs of *LRP1B* and *LAMA1* was 2.16 and 0.81, respectively. *LRP1B* was thought to

function as a tumor suppressor (53), and microRNA-mediated inactivation of LRP1B could increase the growth and invasive capacity of cancer cells (54). Altered protein expression of LAMA1, posttranscriptionally regulated by miR-202, contributed to cell proliferation and migration of ESCC (55). Our data provide evidence that *LRP1B* and *LAMA1* may contribute to tumorigenesis through post-transcriptional regulation at another level. We also downloaded RNA sequencing data from TCGA to evaluate the differential expression values of these pisSNV-hosted genes between tumor and corresponding normal tissues in different cancer types. From the bubble chart of the
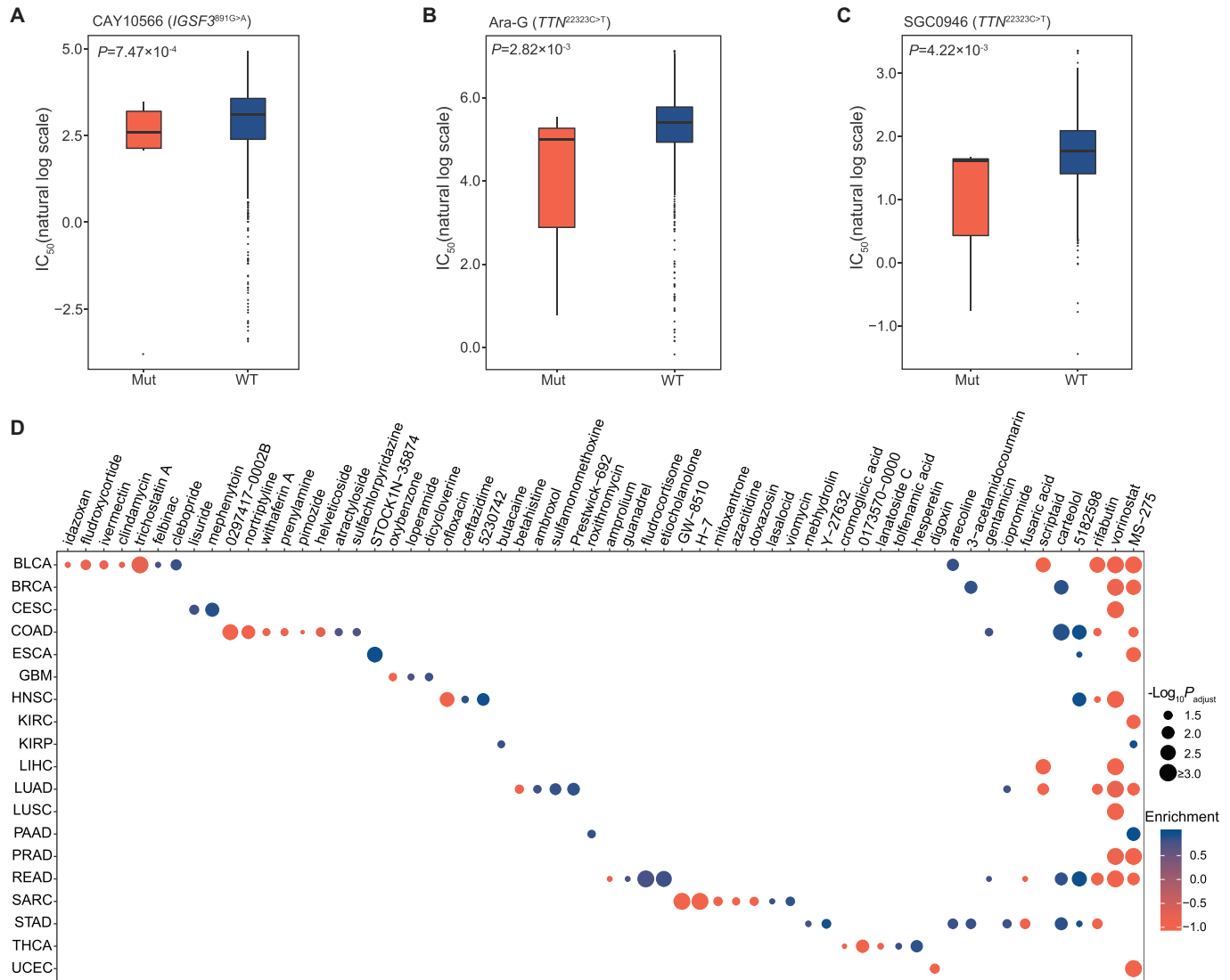
Figure 6. Drugs identified based on pisSNV and expression of pisSNV-hosted genes in different cancer types (**A**–**C**) Significant pharmacogenomic interactions for mutant (Mut) versus wild-type (WT) cell lines identified by ANOVA. (**D**) Drugs identified using CMAP in different cancer types.

top 5% frequently occurred pisSNV-hosted genes (Figure 4B), we found that the majority of them were differentially expressed in each cancer type. For example, *IQGAP3* and *CENPE* were upregulated in 15 cancer types, while *ITGA8* and MYH11 was downregulated in 16 and 15 cancer types, respectively, which, from the perspective of expression, reveals the functional relevance of these pisSNV-hosted genes in cancer tumorigenesis. When we summarized the gene-length corrected frequency of occurrence of 2,042 genes based on the status of posttranscriptional impact and/or function damage (Figure 4C), we found that 470 of 8,320 cancer patients carried protein-damaging mutations in *LRP1B*, the highest frequently occurred pisSNV-hosted gene. Thus, we inferred that *LRP1B* contributed to tumorigenesis at both the protein function and posttranscriptional regulation levels. This was supported by another evidence that some identified pisSNV-hosted genes are well-established cancer genes with a high deleterious mutation rate, such as *PTEN*, *RB1*, *COL11A1*, *PIK3CA*, *CSMD3*

and *DNAH5*. Interestingly, we found that some pisSNV-hosted gene, such as *UBR4*, *EP400* and *INTS1* (Supplementary Table S6), function during carcinogenesis mainly via posttranscriptional manner. The identification of these pisSNV-hosted genes highly expands our understanding of cancer biology.

Significantly altered subnetworks containing pisSNV-hosted genes

To further assess the interactive relationship among the 2,042 pisSNV-hosted genes, we employed HotNet2 (40) to mine the subnetwork in the HPRD network according to the occurrence frequency of each gene, and identified 4 significantly altered subnetworks (Figure 5). Overall, the altered subnetworks included several well-known cancer pathways, such as the phosphoinositide 3-kinase (PI3K)-AKT, mTOR and NOTCH signaling pathways. Subnetwork 1 consisted of *ITPR1*, *ITPR3*, *RYR2* (2.25% of 8,320

samples), *RYR3* and *TRPM2*, which are involved in thyroid hormone synthesis, pancreatic secretion, and apelin signaling. This subnetwork also contained *OBSCN* (2.21% of all samples), which co-occurred as pisSNV-hosted genes in >20 cancer types. Subnetwork 2 was related to five cancer types, UCEC, LUAD, COAD, SKCM and STAD, and the genes included were *COL1A1*, *COL4A1*, *COL4A4*, *FN1*, *LAMA5*, *LAMC1* and *LAMA1*, which are enriched in PI3K-Akt signaling pathway, suggesting the shared etiology of these cancer types at post-transcriptional level. Subnetwork 3 contained multiple members of the NOTCH signaling pathway in four cancer types, BRCA, UCEC, COAD and SKCM, including *NOTCH1*, *NOTCH2*, *MAML2*, *JAG1*, *JAG2* and *MAML3*. The last most posttranscriptionally impaired pan-cancer subnetwork consisted of *MTOR*, *RPTOR*, *CLASP2*, *CLIP1*, *RICTOR* and *CLIP2*, which are enriched in the mTOR signaling pathway, AMPK signaling pathway, insulin signaling pathway and PI3K-Akt signaling pathway. Posttranscriptionally impaired synonymous mutations of *MTOR*, a reported oncogene in multiple cancers (56), co-occurred in 14 cancer types. Interestingly, we also identified some newly cancer-relevant genes within these subnetworks. Our findings indicated that these pisSNV-hosted genes group together with well-established cancer genes, contributing to the carcinogenesis of cancers at the posttranscriptional level.

Therapeutic implications of pisSNVs and pisSNV-hosted genes

To investigate the therapeutic implications of these identified pisSNVs, we analyzed pairwise interactions between mutation data of pisSNVs and dose–response curve from the 1,001 cancer cell lines, and identified two drug-response associated pisSNVs (Figure 6A–C). As shown in Figure 6A, high sensitivity (low IC_{50} value) for CAY10566, a steroyl-CoA desaturase 1 (SCD1) inhibitor which inhibits the conversion from saturated to monounsaturated, long-chain fatty acyl-CoAs (57), was significantly associated with *IGSF3* c.891 G>A mutation. In addition, Ara-G and SGC0946 had higher response to *TTN* c.22323 C>T mutated cell lines in comparison with wild-type cell lines (Figure 6B). To further explore the potential therapeutic effects of the identified cancer-associated pisSNV-hosted genes, we employed the CMAP workflow (43) to identify clinical drugs based on the expression signature of these pisSNV-hosted genes. In total, 60 compounds were identified, and the majority of them were identified to have an effect on a specific type of cancer (Figure 6D). Two inhibitors of histone deacetylase, namely, MS-275 and vorinostat, were identified as highly ranked compounds with antagonistic effects on gene expression signatures associated with 9 cancer types (Figure 6D). In addition, compound 5182598 and carteolol were found to be able to induce the biological state encoded in the signature associated with five and four cancer types, respectively (Figure 6D). Collectively, the identification of above drugs and compounds might be potentially beneficial for treating patients with specific cancer types.

DISCUSSION

Recent studies have shown that synonymous mutations could function as cancer drivers by altering RNA splicing, RNA stability and protein translation (3,4,58–60). However, it is still challenging to decipher the role of these mutations genome-wide in the carcinogenesis of cancers. Numerous studies have found that RBPs play an important role in posttranscriptional regulation (6–8,35), and in the cancer genome, abnormal expression of RBPs has a significant effect on cancer phenotype (10–12). Therefore, we speculated that the synonymous mutations present in the cancer genome may disrupt the binding of RBPs by altering the secondary structure of RNA, thereby affecting the translation, transportation or degradation of the RNA and ultimately resulting in the carcinogenesis and progression of cancer. To verify this, we presented a new approach, PIVar, according to the functional confidence of variants based on multi-omic experimental data, for identifying pisSNVs from pan-cancer genome data. When we used the allele-specific binding preference of 103 RBPs derived from the CLIP-seq and RNA-seq data of the HepG2 cell line, we observed that 73.33% of the pisSNVs identified in the cell line could be validated. Also, our EMSA results on three randomly selected pisSNVs provide experimental evidence that the identified pisSNVs could disrupt the binding of their corresponding RBPs. Thus, we provided an efficient and reliable tool for genome-wide deciphering of the role of these synonymous mutations in the development of cancers at the posttranscriptional regulation level.

By employing PIVar, we identified approximately twenty-three thousands nonredundant pisSNVs spanning 2,042 genes across 22 cancer types. Compared with that of control samples from the DSMNC database, an elevated ratio of somatic pisSNVs across cancer types was observed, and a high ratio of pisSNVs could worsen the patients' survival of five cancer types, including BLCA, LUAD, OV, SKCM and STAD, even when adjusted by multivariate Cox regression. The results provided clinical evidence of the ability of the pisSNV ratio to serve as a potential prognostic biomarker for these five kinds of cancers. It should be noted that some identified pisSNV-hosted genes were well-established cancer-relevant genes with a high deleterious mutation rate, including *PTEN*, *RB1*, *COL11A1*, *PIK3CA*, *CSMD3*, *DNAH5* and *LRP1B*. For example, *LRP1B*, a member of low density lipoprotein receptor-related protein family, had the highest gene-length corrected frequency of occurrence as a pisSNV-hosted gene, while 470 of 8,320 cancer patients was also identified to carry protein damaging mutations of *LRP1B*. Down-regulation of *LRP1B* was reported to promote the growth and migration of colon cancer cells (61), and its deletion was associated with acquired chemotherapy resistance to liposomal doxorubicin in high-grade serous ovarian cancers (62). Thus, the pisSNV-hosted genes tend to contribute to tumorigenesis at both the protein functional and posttranscriptional regulation levels. Interestingly, we found that some pisSNV-hosted gene, such as *UBR4*, *EP400* and *INTS1* (Supplementary Table S6), might exert their function during carcinogenesis mainly via posttranscriptional mechanisms. The identification of these pisSNV-hosted genes that were not previously well charac-

terized based on missense mutation or expression, highly expands our understanding of cancer biology. It should be mentioned that only impaired structure and posttranscriptional regulation of pisSNV-hosted genes were considered into the current discussion, while some genes may work as transcription factors or epigenetic regulators and exert their function at transcriptional level, thus comprehensive functions of these identified pisSNV-hosted gene and their functional validation is still needed in the future study.

Clinical implication analysis unveiled two drug-response associated pisSNVs and three pisSNV-associated compounds including CAY10566, Ara-G and SGC0946. Of the compounds responding to *TTN* c.22323 C>T synonymous mutation, Ara-G is a metabolite of nelarabine, which could be used for chemotherapy in T-cell acute lymphoblastic leukemia (63). In addition, SGC0946 can work as an inhibitor of histone lysine methyltransferase for H3K79 and selectively kill mixed lineage leukemia cells (64). Drug susceptibility predictions based on the expression signature of identified pisSNV-hosted genes revealed two highly ranked compounds, that is, MS-275 and vorinostat, with effects on 9 cancer types. MS-275, also known as entinostat and SNDX-275, is a benzamide histone deacetylase inhibitor and was undergoing clinical trials for the treatment of various cancers (65). A recent study suggested that vorinostat (also known as suberanilohydroxamic acid, SAHA) possesses some activity against recurrent glioblastoma multiforme, resulting in a median overall survival of 5.7 months (66). Moreover, combination of vorinostat and carboplatin as well as paclitaxel in the treatment of advanced non-small-cell lung carcinoma (NSCLC) showed improved response rates and increased median progression-free survival and overall survival (67). The identification of these compounds might be beneficial for patients with specific cancer types.

In summary, for the first time, our study reveals the prevalence and clinical relevance of pisSNVs in cancers, and provides valuable resource for future post-transcriptional regulation researches, which may facilitate prognosis of specific cancer types and development of new therapeutic strategies for cancers.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

Conflict of interest statement. None declared.

REFERENCES

1. Hunt,R.C., Simhadri,V.L., Iandoli,M., Sauna,Z.E. and Kimchi-Sarfaty,C. (2014) Exposing synonymous mutations. *Trends Genet.*, **30**, 308–321.
2. Greenman,C., Stephens,P., Smith,R., Dalgliesh,G.L., Hunter,C., Bignell,G., Davies,H., Teague,J., Butler,A., Stevens,C. *et al.* (2007) Patterns of somatic mutation in human cancer genomes. *Nature*, **446**, 153–158.
3. Soussi,T., Taschner,P.E. and Samuels,Y. (2017) Synonymous somatic variants in human cancer are not infamous: a plea for full disclosure in databases and publications. *Hum. Mutat.*, **38**, 339–342.
4. Supek,F., Minana,B., Valcarcel,J., Gabaldon,T. and Lehner,B. (2014) Synonymous mutations frequently act as driver mutations in human cancers. *Cell*, **156**, 1324–1335.
5. Xiang,Y., Ye,Y., Lou,Y., Yang,Y., Cai,C., Zhang,Z., Mills,T., Chen,N.Y., Kim,Y., Muge Ozguc,F. *et al.* (2018) Comprehensive characterization of alternative polyadenylation in human cancer. *J. Natl. Cancer Inst.*, **110**, 379–389.
6. Pereira,B., Billaud,M. and Almeida,R. (2017) RNA-binding proteins in cancer: old players and new actors. *Trends Cancer*, **3**, 506–528.
7. Lunde,B.M., Moore,C. and Varani,G. (2007) RNA-binding proteins: modular design for efficient function. *Nat. Rev. Mol. Cell Biol.*, **8**, 479–490.
8. Glisovic,T., Bachorik,J.L., Yong,J. and Dreyfuss,G. (2008) RNA-binding proteins and post-transcriptional gene regulation. *FEBS Lett.*, **582**, 1977–1986.
9. Ouyang,H., Zhang,K., Fox-Walsh,K., Yang,Y., Zhang,C., Huang,J., Li,H., Zhou,Y. and Fu,X.D. (2017) The RNA binding protein EWS is broadly involved in the regulation of pri-miRNA processing in mammalian cells. *Nucleic Acids Res.*, **45**, 12481–12495.
10. Wang,Z.L., Li,B., Luo,Y.X., Lin,Q., Liu,S.R., Zhang,X.Q., Zhou,H., Yang,J.H. and Qu,L.H. (2018) Comprehensive genomic characterization of RNA-Binding proteins across human cancers. *Cell Rep.*, **22**, 286–298.
11. Hopkins,T.G., Mura,M., Al-Ashtal,H.A., Lahr,R.M., Abd-Latip,N., Sweeney,K., Lu,H., Weir,J., El-Bahrawy,M., Steel,J.H. *et al.* (2016) The RNA-binding protein LARP1 is a post-transcriptional regulator of survival and tumorigenesis in ovarian cancer. *Nucleic Acids Res.*, **44**, 1227–1246.
12. Wurth,L., Papasaikas,P., Olmeda,D., Bley,N., Calvo,G.T., Guerrero,S., Cerezo-Wallis,D., Martinez-Useros,J., Garcia-Fernandez,M., Huttelmaier,S. *et al.* (2016) UNR/CSDE1 drives a post-transcriptional program to promote melanoma invasion and metastasis. *Cancer Cell*, **30**, 694–707.
13. Frisone,P., Pradella,D., Di Matteo,A., Belloni,E., Ghigna,C. and Paronetto,M.P. (2015) SAM68: signal transduction and RNA metabolism in human cancer. *Biomed. Res. Int.*, **2015**, 528954.
14. Wang,J., Guo,Y., Chu,H., Guan,Y., Bi,J. and Wang,B. (2013) Multiple functions of the RNA-binding protein HuR in cancer progression, treatment responses and prognosis. *Int. J. Mol. Sci.*, **14**, 10015–10041.
15. David,C.J., Chen,M., Assanah,M., Canoll,P. and Manley,J.L. (2010) HnRNP proteins controlled by c-Myc deregulate pyruvate kinase mRNA splicing in cancer. *Nature*, **463**, 364–368.
16. Zhang,J., Xu,E., Ren,C., Yan,W., Zhang,M., Chen,M., Cardiff,R.D., Imai,D.M., Wisner,E. and Chen,X. (2014) Mice deficient in Rbm38, a target of the p53 family, are susceptible to accelerated aging and spontaneous tumors. *Proc. Natl. Acad. Sci. U. S. A.*, **111**, 18637–18642.
17. Wurth,L. (2012) Versatility of RNA-binding proteins in cancer. *Comp. Funct. Genomics*, **2012**, 178525.
18. Feng,H., Bao,S., Rahman,M.A., Weyn-Vanhentenryck,S.M., Khan,A., Wong,J., Shah,A., Flynn,E.D., Krainer,A.R. and Zhang,C. (2019) Modeling RNA-binding protein specificity in vivo by precisely registering protein-RNA crosslink sites. *Mol. Cell*, **74**, 1189–1204.
19. Singh,B., Trincado,J.L., Tatlow,P.J., Piccolo,S.R. and Eyras,E. (2018) Genome sequencing and RNA-motif analysis reveal novel damaging noncoding mutations in human tumors. *Mol. Cancer Res.*, **16**, 1112–1124.
20. Suhl,J.A., Muddashetty,R.S., Anderson,B.R., Ifrim,M.F., Visootsak,J., Bassell,G.J. and Warren,S.T. (2015) A 3′ untranslated region variant in FMR1 eliminates neuronal activity-dependent translation of FMRP by disrupting binding of the RNA-binding protein HuR. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, E6553–E6561.
21. Boyle,A.P., Hong,E.L., Hariharan,M., Cheng,Y., Schaub,M.A., Kasowski,M., Karczewski,K.J., Park,J., Hitz,B.C., Weng,S. *et al.* (2012) Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res.*, **22**, 1790–1797.
22. Smith,A.K., Kilaru,V., Kocak,M., Almli,L.M., Mercer,K.B., Ressler,K.J., Tylavsky,F.A. and Conneely,K.N. (2014) Methylation quantitative trait loci (meQTLs) are consistently detected across ancestry, developmental stage, and tissue type. *BMC Genomics*, **15**, 145.

23. Yang,Y.C., Di,C., Hu,B., Zhou,M., Liu,Y., Song,N., Li,Y., Umetsu,J. and Lu,Z.J. (2015) CLIPdb: a CLIP-seq database for protein-RNA interactions. *BMC Genomics*, **16**, 51.

24. Barrett,T., Wilhite,S.E., Ledoux,P., Evangelista,C., Kim,I.F., Tomashevsky,M., Marshall,K.A., Phillippy,K.H., Sherman,P.M., Holko,M. *et al.* (2013) NCBI GEO: archive for functional genomics data sets–update. *Nucleic Acids Res.*, **41**, D991–D995.

25. Ray,D., Kazan,H., Cook,K.B., Weirauch,M.T., Najafabadi,H.S., Li,X., Gueroussov,S., Albu,M., Zheng,H., Yang,A. *et al.* (2013) A compendium of RNA-binding motifs for decoding gene regulation. *Nature*, **499**, 172–177.

26. Cook,K.B., Kazan,H., Zuberi,K., Morris,Q. and Hughes,T.R. (2011) RBPDB: a database of RNA-binding specificities. *Nucleic Acids Res.*, **39**, D301–D308.

27. Dassi,E., Malossini,A., Re,A., Mazza,T., Tebaldi,T., Caputi,L. and Quattrone,A. (2012) AURA: Atlas of UTR Regulatory Activity. *Bioinformatics*, **28**, 142–144.

28. Sabarinathan,R., Tafer,H., Seemann,S.E., Hofacker,I.L., Stadler,P.F. and Gorodkin,J. (2013) The RNAsnp web server: predicting SNP effects on local RNA secondary structure. *Nucleic Acids Res.*, **41**, W475–W479.

29. Gruber,A.R., Lorenz,R., Bernhart,S.H., Neubock,R. and Hofacker,I.L. (2008) The Vienna RNA websuite. *Nucleic Acids Res.*, **36**, W70–W74.

30. Lee,D. (2016) LS-GKM: a new gkm-SVM for large-scale datasets. *Bioinformatics*, **32**, 2196–2198.

31. Lee,D., Gorkin,D.U., Baker,M., Strober,B.J., Asoni,A.L., McCallion,A.S. and Beer,M.A. (2015) A method to predict the impact of regulatory variants from DNA sequence. *Nat. Genet.*, **47**, 955–961.

32. Betel,D., Wilson,M., Gabow,A., Marks,D.S. and Sander,C. (2008) The microRNA.org resource: targets and expression. *Nucleic Acids Res.*, **36**, D149–D153.

33. Agarwal,V., Bell,G.W., Nam,J.W. and Bartel,D.P. (2015) Predicting effective microRNA target sites in mammalian mRNAs. *eLife*, **4**, e05005.

34. Gong,J., Liu,C., Liu,W., Wu,Y., Ma,Z., Chen,H. and Guo,A.Y. (2015) An update of miRNASNP database for better SNP selection by GWAS data, miRNA expression and online tools. *Database*, **2015**, bav029.

35. Bahrami-Samani,E. and Xing,Y. (2019) Discovery of Allele-Specific Protein-RNA interactions in human transcriptomes. *Am. J. Hum. Genet.*, **104**, 492–502.

36. Miao,X., Li,X., Wang,L., Zheng,C. and Cai,J. (2019) DSMNC: a database of somatic mutations in normal cells. *Nucleic Acids Res.*, **47**, D971–D975.

37. Lunde,B.M., Moore,C. and Varani,G. (2007) RNA-binding proteins: modular design for efficient function. *Nat. Rev. Mol. Cell Biol.*, **8**, 479–490.

38. Li,J., Shi,L., Zhang,K., Zhang,Y., Hu,S., Zhao,T., Teng,H., Li,X., Jiang,Y., Ji,L. *et al.* (2018) VarCards: an integrated genetic and clinical database for coding variants in the human genome. *Nucleic Acids Res.*, **46**, D1039–D1048.

39. Li,J., Zhao,T., Zhang,Y., Zhang,K., Shi,L., Chen,Y., Wang,X. and Sun,Z. (2018) Performance evaluation of pathogenicity-computation methods for missense variants. *Nucleic Acids Res.*, **46**, 7793–7804.

40. Leiserson,M.D., Vandin,F., Wu,H.T., Dobson,J.R., Eldridge,J.V., Thomas,J.L., Papoutsaki,A., Kim,Y., Niu,B., McLellan,M. *et al.* (2015) Pan-cancer network analysis identifies combinations of rare somatic mutations across pathways and protein complexes. *Nat. Genet.*, **47**, 106–114.

41. Yu,G., Wang,L.G., Han,Y. and He,Q.Y. (2012) clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*, **16**, 284–287.

42. Zhao,J., Cheng,F. and Zhao,Z. (2017) Tissue-specific signaling networks rewired by major somatic mutations in human cancer revealed by proteome-wide discovery. *Cancer Res.*, **77**, 2810–2821.

43. Lamb,J., Crawford,E.D., Peck,D., Modell,J.W., Blat,I.C., Wrobel,M.J., Lerner,J., Brunet,J.P., Subramanian,A., Ross,K.N. *et al.* (2006) The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science*, **313**, 1929–1935.

44. Yang,E.W., Bahn,J.H., Hsiao,E.Y., Tan,B.X., Sun,Y.W., Fu,T., Zhou,B., Van Nostrand,E.L., Pratt,G.A., Freese,P. *et al.* (2019) Allele-specific binding of RNA-binding proteins reveals functional genetic variants in the RNA. *Nat.Commun.*, **10**, 1338.

45. Hannigan,A., Smith,P., Kalna,G., Lo Nigro,C., Orange,C., O'Brien,D.I., Shah,R., Syed,N., Spender,L.C., Herrera,B. *et al.* (2010) Epigenetic downregulation of human disabled homolog 2 switches TGF-beta from a tumor suppressor to a tumor promoter. *J. Clin. Invest.*, **120**, 2842–2857.

46. Wu,B., Su,S., Patil,D.P., Liu,H., Gan,J., Jaffrey,S.R. and Ma,J. (2018) Molecular basis for the specific and multivariant recognitions of RNA substrates by human hnRNP A2/B1. *Nat. Commun.*, **9**, 420.

47. Alarcon,C.R., Goodarzi,H., Lee,H., Liu,X., Tavazoie,S. and Tavazoie,S.F. (2015) HNRNPA2B1 is a mediator of m(6)A-Dependent Nuclear RNA processing events. *Cell*, **162**, 1299–1308.

48. Li,C., Xie,J., Lu,Z., Chen,C., Yin,Y., Zhan,R., Fang,Y., Hu,X. and Zhang,C.C. (2015) ADCY7 supports development of acute myeloid leukemia. *Biochem. Biophys. Res. Commun.*, **465**, 47–52.

49. Zhang,J., Lieu,Y.K., Ali,A.M., Penson,A., Reggio,K.S., Rabadan,R., Raza,A., Mukherjee,S. and Manley,J.L. (2015) Disease-associated mutation in SRSF2 misregulates splicing by altering RNA-binding affinities. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, E4726–E4734.

50. Daubner,G.M., Clery,A., Jayne,S., Stevenin,J. and Allain,F.H. (2012) A syn-anti conformational difference allows SRSF2 to recognize guanines and cytosines equally well. *EMBO J.*, **31**, 162–174.

51. Huang,X., Zhang,H., Guo,X., Zhu,Z., Cai,H. and Kong,X. (2018) Insulin-like growth factor 2 mRNA-binding protein 1 (IGF2BP1) in cancer. *J. Hematol. Oncol.*, **11**, 88.

52. Venkateswaran,A., Sekhar,K.R., Levic,D.S., Melville,D.B., Clark,T.A., Rybski,W.M., Walsh,A.J., Skala,M.C., Crooks,P.A., Knapik,E.W. *et al.* (2014) The NADH oxidase ENOX1, a critical mediator of endothelial cell radiosensitization, is crucial for vascular development. *Cancer Res.*, **74**, 38–43.

53. Liu,C.X., Li,Y., Obermoeller-McCormick,L.M., Schwartz,A.L. and Bu,G. (2001) The putative tumor suppressor LRP1B, a novel member of the low density lipoprotein (LDL) receptor family, exhibits both overlapping and distinct properties with the LDL receptor-related protein. *J. Biol. Chem.*, **276**, 28889–28896.

54. Prazeres,H., Torres,J., Rodrigues,F., Pinto,M., Pastoriza,M.C., Gomes,D., Cameselle-Teijeiro,J., Vidal,A., Martins,T.C., Sobrinho-Simoes,M. *et al.* (2011) Chromosomal, epigenetic and microRNA-mediated inactivation of LRP1B, a modulator of the extracellular environment of thyroid cancer cells. *Oncogene*, **30**, 1302–1317.

55. Meng,X., Chen,X., Lu,P., Ma,W., Yue,D., Song,L. and Fan,Q. (2016) MicroRNA-202 inhibits tumor progression by targeting LAMA1 in esophageal squamous cell carcinoma. *Biochem. Biophys. Res. Commun.*, **473**, 821–827.

56. Bailey,M.H., Tokheim,C., Porta-Pardo,E., Sengupta,S., Bertrand,D., Weerasinghe,A., Colaprico,A., Wendl,M.C., Kim,J., Reardon,B. *et al.* (2018) Comprehensive characterization of cancer driver genes and mutations. *Cell*, **173**, 371–385.

57. Tesfay,L., Paul,B.T., Konstorum,A., Deng,Z., Cox,A.O., Lee,J., Furdui,C.M., Hegde,P., Torti,F.M. and Torti,S.V. (2019) Stearoyl-coa desaturase 1 protects ovarian cancer cells from ferroptotic cell death. *Cancer Res.*, **79**, 5355–5366.

58. Li,X., Shi,L., Wang,Y., Zhong,J., Zhao,X., Teng,H., Shi,X., Yang,H., Ruan,S., Li,M. *et al.* (2019) OncoBase: a platform for decoding regulatory somatic mutations in human cancers. *Nucleic Acids Res.*, **47**, D1044–D1055.

59. Mao,F., Xiao,L., Li,X., Liang,J., Teng,H., Cai,W. and Sun,Z.S. (2016) RBP-Var: a database of functional variants involved in regulation mediated by RNA-binding proteins. *Nucleic Acids Res.*, **44**, D154–D163.

60. Mao,F., Liu,Q., Zhao,X., Yang,H., Guo,S., Xiao,L., Li,X., Teng,H., Sun,Z. and Dou,Y. (2017) EpiDenovo: a platform for linking regulatory de novo mutations to developmental epigenetics and diseases. *Nucleic Acids Res.*, **46**, D92–D99.

61. Wang,Z., Sun,P., Gao,C., Chen,J., Li,J., Chen,Z., Xu,M., Shao,J., Zhang,Y. and Xie,J. (2017) Down-regulation of LRP1B in colon cancer promoted the growth and migration of cancer cells. *Exp. Cell Res.*, **357**, 1–8.

62. Cowin,P.A., George,J., Fereday,S., Loehrer,E., Van Loo,P., Cullinane,C., Etemadmoghadam,D., Ftouni,S., Galletta,L., Anglesio,M.S. *et al.* (2012) LRP1B deletion in high-grade serous

ovarian cancers is associated with acquired chemotherapy resistance to liposomal doxorubicin. *Cancer Res.*, **72**, 4060–4073.

63. Homminga,I., Zwaan,C.M., Manz,C.Y., Parker,C., Bantia,S., Smits,W.K., Higginbotham,F., Pieters,R. and Meijerink,J.P. (2011) In vitro efficacy of forodesine and nelarabine (ara-G) in pediatric leukemia. *Blood*, **118**, 2184–2190.

64. Yu,W., Chory,E.J., Wernimont,A.K., Tempel,W., Scopton,A., Federation,A., Marineau,J.J., Qi,J., Barsyte-Lovejoy,D., Yi,J. *et al.* (2012) Catalytic site remodelling of the DOT1L methyltransferase by selective inhibitors. *Nat. Commun.*, **3**, 1288.

65. Kato,Y., Yoshimura,K., Shin,T., Verheul,H., Hammers,H., Sanni,T.B., Salumbides,B.C., Van Erp,K., Schulick,R. and Pili,R. (2007) Synergistic in vivo antitumor effect of the histone deacetylase inhibitor MS-275 in combination with interleukin 2 in a murine model of renal cell carcinoma. *Clin. Cancer Res.*, **13**, 4538–4546.

66. Galanis,E., Jaeckle,K.A., Maurer,M.J., Reid,J.M., Ames,M.M., Hardwick,J.S., Reilly,J.F., Loboda,A., Nebozhyn,M., Fantin,V.R. *et al.* (2009) Phase II trial of vorinostat in recurrent glioblastoma multiforme: a north central cancer treatment group study. *J. Clin. Oncol.*, **27**, 2052–2058.

67. Ramalingam,S.S., Maitland,M.L., Frankel,P., Argiris,A.E., Koczywas,M., Gitlitz,B., Thomas,S., Espinoza-Delgado,I., Vokes,E.E., Gandara,D.R. *et al.* (2010) Carboplatin and Paclitaxel in combination with either vorinostat or placebo for first-line therapy of advanced non-small-cell lung cancer. *J. Clin. Oncol.*, **28**, 56–62.