



# Oncogenic Landscape of Somatic Mutations Perturbing Pan-Cancer lncRNA-ceRNA Regulation

Yuanfu Zhang<sup>1†</sup>, Peng Han<sup>2,3†</sup>, Qiuyan Guo<sup>4†</sup>, Yangyang Hao<sup>1</sup>, Yue Qi<sup>1</sup>, Mengyu Xin<sup>1</sup>, Yafang Zhang<sup>5\*</sup>, Binbin Cui<sup>2\*</sup> and Peng Wang<sup>1\*</sup>

<sup>1</sup> College of Bioinformatics Science and Technology, Harbin Medical University, Harbin, China, <sup>2</sup> Department of Colorectal Surgery, Harbin Medical University Cancer Hospital, Harbin, China, <sup>3</sup> Heilongjiang Cancer Research Institute, Harbin, China, <sup>4</sup> Department of Gynecology, The First Affiliated Hospital of Harbin Medical University, Harbin, China, <sup>5</sup> Department of Anatomy, Harbin Medical University, Harbin, China

## OPEN ACCESS

### Edited by:

Kanaga Sabapathy,  
National Cancer Centre Singapore,  
Singapore

### Reviewed by:

Emenike K. Onyido,  
Swansea University Medical School,  
United Kingdom  
Ashish Lal,  
National Institutes of Health (NIH),  
United States

### \*Correspondence:

Peng Wang  
wpgqy@163.com  
Binbin Cui  
13351112888@163.com  
Yafang Zhang  
yafangzhang2008@aliyun.com

<sup>†</sup> These authors share first authorship

### Specialty section:

This article was submitted to  
Molecular Medicine,  
a section of the journal  
Frontiers in Cell and Developmental  
Biology

**Received:** 25 January 2021

**Accepted:** 19 April 2021

**Published:** 17 May 2021

### Citation:

Zhang Y, Han P, Guo Q, Hao Y, Qi Y, Xin M, Zhang Y, Cui B and Wang P (2021) Oncogenic Landscape of Somatic Mutations Perturbing Pan-Cancer lncRNA-ceRNA Regulation. *Front. Cell Dev. Biol.* 9:658346. doi: 10.3389/fcell.2021.658346

Competing endogenous RNAs (ceRNA) are transcripts that communicate with and co-regulate each other by competing for the binding of shared microRNAs (miRNAs). Long non-coding RNAs (lncRNAs) as a type of ceRNA constitute a competitive regulatory network determined by miRNA response elements (MREs). Mutations in lncRNA MREs destabilize their original regulatory pathways. Study of the effects of lncRNA somatic mutations on ceRNA mechanisms can clarify tumor mechanisms and contribute to the development of precision medicine. Here, we used somatic mutation profiles collected from TCGA to characterize the role of lncRNA somatic mutations in the ceRNA regulatory network in 33 cancers. The 31,560 mutation sites identified by TargetScan and miRanda affected the balance of 70,811 ceRNA regulatory pathways. Putative mutations were categorized as high or low based on mutation frequencies. Multivariate multiple regression revealed a significant effect of 162 high-frequency mutations in six cancer types on the expression levels of target mRNAs (ceMs) through the ceRNA mechanism. Low-frequency mutations in multiple cancers perturbing 1624 ceM have been verified by Student's *t*-test, indicating a significant mechanism of changes in the expression level of oncogenic genes. Oncogenic signaling pathway studies involving ceMs indicated functional heterogeneity of multiple cancers. Furthermore, we identified that lncRNA, perturbing ceMs associated with patient survival, have potential as biomarkers. Our collective findings revealed individual differences in somatic mutations perturbing ceM expression and impacting tumor heterogeneity.

**Keywords:** ceRNA, somatic mutation, prognosis, functional analysis, oncogenic pathway

## INTRODUCTION

In recent years, novel, post-transcriptional regulatory mechanisms of competing endogenous RNAs (ceRNAs) have been revealed. RNA molecules, including messenger RNAs (mRNAs), long noncoding RNAs (lncRNAs), circular RNAs, and pseudogene transcripts, can function as competing endogenous RNAs (ceRNAs) to indirectly regulate the expression of relevant target genes by competing with each other for microRNAs (miRNAs) (Salmena et al., 2011). These ceRNAs harbor miRNA response elements (MREs) that bind to miRNA through complementary

sequences and can induce degradation or inhibition of the expression of target genes. In addition, the combination of miRNAs and target genes was a complex network; one miRNA can regulate multiple genes and one gene can be regulated by multiple miRNAs.

lncRNAs were once regarded as byproducts of gene transcription (Quinn and Chang, 2016). However, they are crucial in post-transcriptional regulation through the ceRNA mechanism (Chen et al., 2019) and are dramatic factors that contribute to biological growth and development, aging, diseases, and multiple cancers (Rinn and Chang, 2012; Schmitt and Chang, 2016). For example, MALAT1, which is highly expressed in most cancers, regulates the cell cycle (Tripathi et al., 2013), and PCA3 is an important molecular marker in the early stage of cancer (Lemos et al., 2019). Somatic mutations, which occur in cells other than germ cells and are not inherited, are the substantial cause of most tumors (Jia and Zhao, 2017). Mutation in an miRNA response elements (MRE) of lncRNA can weaken, enhance, or prevent binding to the original miRNA, resulting in an imbalance in the ceRNA regulatory network and altered expression of the relevant target genes (Thomas et al., 2011; Thomson and Dinger, 2016).

The development of sequencing technologies has enabled the identification of somatic mutations associated with tumors (Martincorena and Campbell, 2015). Genetic variations affecting miRNA gene expression have been described (Civelek et al., 2013; Siddle et al., 2014), as has the expression of coding genes whose 3' untranslated regions are targeted by miRNAs (Gamazon et al., 2012; Lu and Clark, 2012). Genetic polymorphisms affecting the regulation of human ceRNAs have been reported (Li M. J. et al., 2017). However, few studies have explored the effects of somatic mutations on ceRNA mechanisms.

Here, we used mutation and RNA-seq profiles from The Cancer Genome Atlas (TCGA) (Tomczak et al., 2015) database to conduct a systematic investigation concerning the effects of lncRNA mutations on the expression of target mRNAs via the ceRNA mechanism in pan-carcinoma. We also studied the impact of significant mutations on oncogenic mechanisms and patient survival.

## MATERIALS AND METHODS

### Data Collection

Information concerning RNA-seq and somatic mutation profiles of 33 cancers were obtained from The Cancer Genome Atlas (Tomczak et al., 2015) (TCGA)<sup>1</sup> database. The GRCh38 v29 version of the human genome annotation data from GENCODE (Harrow et al., 2012)<sup>2</sup>, including the position and sequence information of lncRNAs, was used to annotate somatic mutation profiles. Sequences of miRNA and annotation information were obtained from the miRBase (Kozomara and Griffiths-Jones, 2014)<sup>3</sup> database. Interaction data of miRNA and target genes

(mRNA) that were validated using established experimental methods including the luciferase reporter assay, PCR, and western blotting were collected from miRTarBase (Chou et al., 2018) V8.0<sup>4</sup>. Hallmark (Hanahan and Weinberg, 2011) gene sets were collected from the Molecular Signatures Database (MSigDB Liberzon et al., 2011)<sup>5</sup>.

### Construction of Somatic Mutation-miRNA-lncRNA (ceL)-mRNA (ceM) Unit

Among the numerous somatic mutations in the pan-cancer genome, lncRNA mutations was the focus of this study. We, respectively, define the lncRNA and mRNA involved in the ceRNA regulatory mechanism as ceL and ceM. Sequences approximately 7 nucleotide (nt) upstream and downstream of the lncRNA somatic mutation sites were extracted using the lncRNA annotation from GENCODE, which will be used to construct mutation and control sequences. Considering that the TargetScan (Friedman et al., 2009) software does not recognize short sequences, it was necessary to extract longer upstream nucleotide sequences (14 nt) to offset this impact. TargetScan and miRanda (Betel et al., 2008) (e.g., the miRanda algorithm) are miRNA target gene prediction tools, and therefore were used to predict the miRNA-target relationships of control sequences with strict thresholds of score > 160 and energy < -20 for miRanda and context score < -0.4 for TargetScan. We defined the lncRNA and mRNA involved in the imbalance of ceRNA regulatory mechanism as ceL and ceM, respectively. We selected "mutation-miRNA-lncRNA (ceL)" units with varying binding affinities between the mutation and control sequences, and regarded loss, down, gain, and up as the four conditions of altered lncRNA and miRNA binding affinity (Li M. J. et al., 2017). We further searched for candidate mRNAs (ceMs) controlled by the same miRNA from the miRNA-target gene data of miRTarBase as the last element to construct the somatic mutation-miRNA-ceL-ceM (SMILM) unit. In this context, "putative mutations" are defined as mutations effecting original ceRNA regulation mechanism. This definition has been used in a previous study of genetic associations with ceRNA regulation in the human genome (Li et al., 2020).

### Classification and Definition of Mutations

Somatic mutations do not occur frequently. We defined a site at which at least two samples displayed a mutation as a high-frequency (HF) mutation site. The remaining mutations were defined as low-frequency (LF) mutation sites. The altered binding affinity of lncRNA and miRNA binding was divided into four states (gain, up, loss, and down). Gain, up, loss, and down were scored as +1, +0.5, -1, and -0.5, respectively. The functional score of each mutation was calculated by summing the states of all mutated miRNAs associated with it. For LF mutations, we focused on the mRNAs (ceMs) affected by somatic mutations

<sup>1</sup><https://portal.gdc.cancer.gov/>

<sup>2</sup><https://www.genecodegenes.org/>

<sup>3</sup><http://www.mirbase.org/ftp.shtml>

<sup>4</sup><http://mirtarbase.cuhk.edu.cn/php/index.php>

<sup>5</sup><http://software.broadinstitute.org/gsea/msigdb>

through ceRNA regulatory mechanisms. The possible expression tendency of the mRNA was defined as “up” means that the sum of the mutation scores that regulate this mRNA is greater than zero, “down” means the opposite of “up,” and “none” means that the sum of the mutation scores that regulate this mRNA is equal to zero.

## Multivariate Multiple Regression Analyses

Different experimental tools were used for identification of SMILM units mediated by HF mutations (HF-SMILM) and LF mutations (LF-SMILM). In the HF-SMILM unit, multivariate multiple regression models were used to validate whether the expression level normalized by Fragments PerKilobase Million (FPKM) of ceL and ceM conformed to target prediction results (Valente et al., 2014). The fold-change values were used to evaluate the extent of expression changes between two groups of samples. For each SMILM unit, we considered the expression levels of lncRNA ( $E_l$ ) and mRNA ( $E_m$ ) as two independent response variables. As a predictor, the genotype ( $G_t$ ) of an individual was used as the driving variable. Synergistic factors such as the residual expression of miRNAs might also affect the response variables. At the same time, we assumed that the error vector  $\varepsilon = (\varepsilon_1, \varepsilon_2)'$  followed a multivariate Gaussian distribution with an expected value of zero and an unknown covariance matrix. The multivariate multiple regression model constructed for the SMILM unit is:

$$(E_l, E_m) = G_t + M_I r + \varepsilon \quad (1)$$

We used this equation to validate all the SMILM units. We defined  $\eta_l$  and  $\eta_m$  as the regression coefficients of the driving variable  $G_t$ . The influence of  $G_t$  changes on the expression of ceL and ceM was quantified using the regression coefficients  $\eta_l$  and  $\eta_m$ , and the statistical significance of the model was obtained. Since ceL and ceM present a competitive relationship in the ceRNA mechanism, we required that ceL expression changes with genotype and ceM expression changes with genotype followed opposite tendencies ( $\eta_l \times \eta_m < 0, p\text{-value} < 0.05$ ) (Li M. J. et al., 2017).

LF-SMILM unit data were split based on the characterization by ceM, obtaining the somatic mutations, lncRNA, miRNA, and samples corresponding to each ceM. We divided cancer samples into mutated and non-mutated samples according to whether the sample had mutations that affected the expression of specific mRNA (ceM) through the ceRNA mechanism. Student's  $t$ -test was used to compare the ceM expression changes in the two categories of sample. The ceMs with  $p\text{-value} < 0.05$  were retained due to significant changes in their expression affected by putative mutations.

## Construction of ceRNA Regulatory Network

In the SMILM unit validated by multivariate multiple regression models, the mutated lncRNA (ceL), miRNA, and target gene mRNA (ceM) constitute a two-level regulatory relationship. Therefore, we used Cytoscape (Shannon et al.,

2003) to visualize this regulatory relationship in significant SMILM units [mutations-miRNA-lncRNA (ceL)-mRNA (ceM)] (Long et al., 2019).

## Functional Enrichment Analysis Connecting Oncogenic Signaling Pathways

We obtained ceMs whose expression were significantly affected by somatic mutations through the ceRNA mechanism. To assess the role of these ceMs in various cancers, we used the compareCluster function in the R package clusterProfiler to perform functional analyses on multiple pan-cancer gene sets, using threshold  $p\text{valueCutoff} = 0.05$ . Seventeen oncogenic signaling pathways (Malumbres and Barbacid, 2003; Reya and Clevers, 2005; Wade et al., 2013; Moradi-Marjaneh et al., 2018; Ayuk and Abrahamse, 2019; Soleimani et al., 2019, 2020) were collected from articles published between 2008 and 2019. The overlapping signal path was filtered out based on enrichment results.

## Survival Analysis

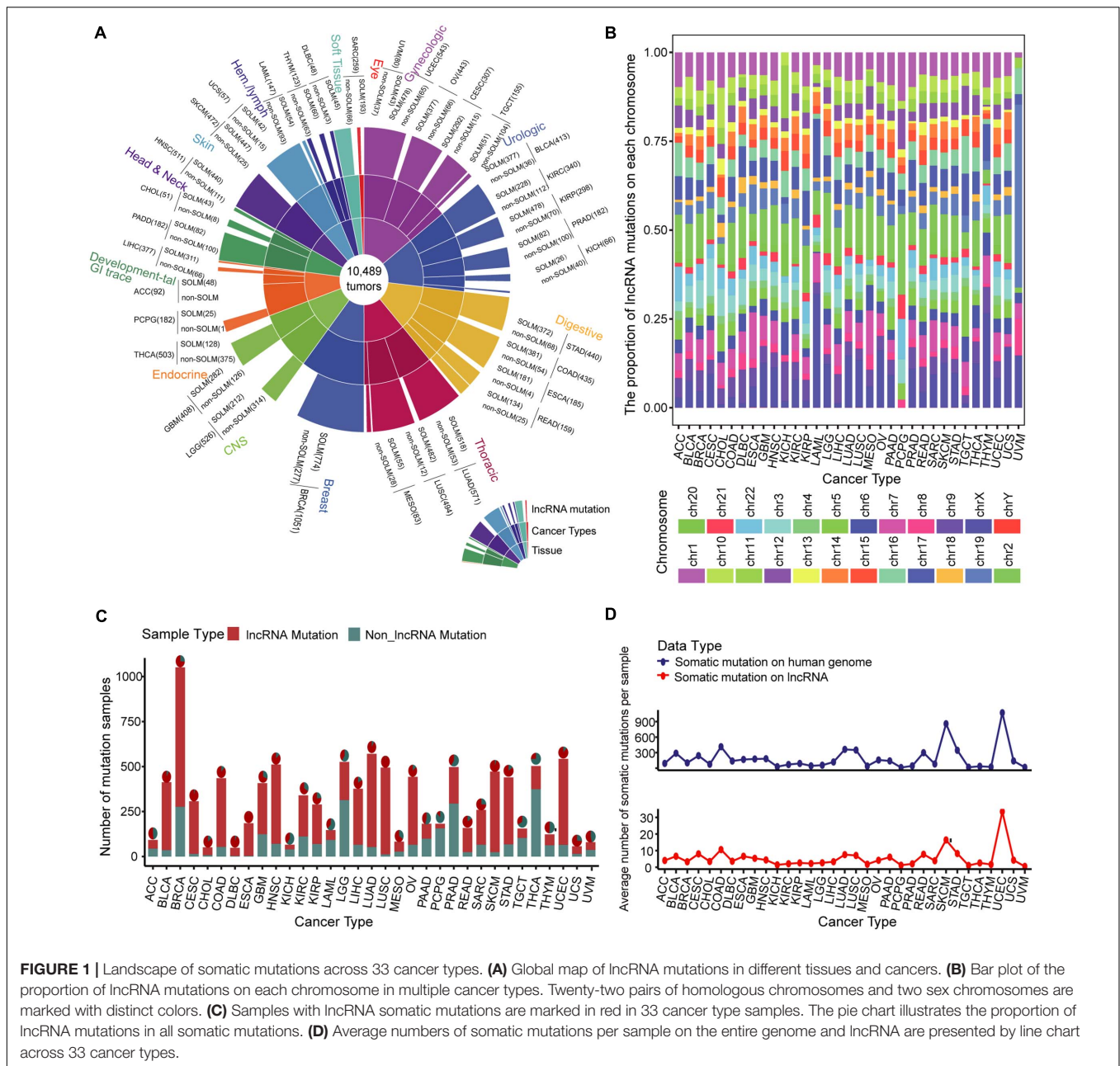
We used the Cox proportional hazards model (Fisher and Lin, 1999) to estimate whether the expression of ceMs regulated by lncRNA mutations according to the ceRNA mechanism was related to patient survival. Hazard ratios (HRs)  $< 1$  and  $p < 0.05$  indicated significant relationships between ceM and reduced risk of death. An HR  $> 1$  indicated the converse. Based on the predicted results, each sample was categorized as one of four types: including “None” means no mutation disrupts the expression of the target gene, “Up-regulated” means the presence of mutations that cause only upregulation of target gene expression, “Down-regulated” means the presence of mutations that cause only downregulation of target gene expression, and “Unknown” means that both mutations resulting in up- and down-regulation of the expression of the target gene are present. The R package for survival was used to create survival curves (Rich et al., 2010). The fitted results were visualized using a ggsurvplot. A  $p\text{-value} < 0.05$  was considered to represent a significant difference in survival.

## RESULTS

### Global Mutation Map Reveals Heterogeneity of Different Tumors

We evaluated samples from the TCGA database collection for which somatic mutation data were available, producing a global map of somatic mutation sample distribution. The map contained 7604 samples with lncRNA mutations in 10,489 samples from 33 cancers (Figure 1A and Supplementary Table 1). We examined the distribution of mutations on chromosomes. The lncRNA mutations in multiple cancer types aggregated differently on chromosomes, especially in kidney renal papillary cell carcinoma (KIPR), acute myeloid leukemia (LAML), pheochromocytoma and paraganglioma (PCPG), thymoma (THYM), and uveal melanoma (UVM),





**FIGURE 1 |** Landscape of somatic mutations across 33 cancer types. **(A)** Global map of lncRNA mutations in different tissues and cancers. **(B)** Bar plot of the proportion of lncRNA mutations on each chromosome in multiple cancer types. Twenty-two pairs of homologous chromosomes and two sex chromosomes are marked with distinct colors. **(C)** Samples with lncRNA somatic mutations are marked in red in 33 cancer type samples. The pie chart illustrates the proportion of lncRNA mutations in all somatic mutations. **(D)** Average numbers of somatic mutations per sample on the entire genome and lncRNA are presented by line chart across 33 cancer types.

compared to those in other tumors (Figure 1B). These findings indicate that the distribution specificity of lncRNA mutations on chromosomes may be the underlying cause of cancer functional heterogeneity. Of all renal cell carcinoma subtypes, the KIPR subtype of kidney cancer has different molecular characteristics and poor survival (Cancer Genome Atlas Research Network, Linehan et al., 2016; Ricketts et al., 2018). Lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) lung cancer subtypes displayed similar mutation distribution profiles on chromosomes, suggesting that cancers in the same tissue site have a similar distribution of mutations on chromosomes. Breast invasive carcinoma (BRCA), kidney chromophobe (KICH), kidney renal clear cell carcinoma (KIRC), and thyroid carcinoma

(THCA) displayed large sample sizes but relatively few lncRNA mutations, suggesting that mutations in lncRNAs have a strong distribution preference among various cancers (Figures 1C,D). lncRNA mutations had low frequencies in the range of 1 to 10%, suggesting that the rates of lncRNA mutations vary among cancer types (Supplementary Figure 1). These findings were consistent with previous studies showing that mutation frequency fluctuates significantly in pan-cancer, and that the mutation rate of some cancers is greatly increased due to missing repair pathways or chromosome integrity checkpoints (Martincorena and Campbell, 2015). We also assessed numbers of mutations per lncRNA in cancer types. A set of lncRNAs with a high mutation frequency was evident for multiple cancers (Supplementary Figure 2). The

lncRNAs XIST, TTN-AS1, STRA6LP, and TSIX had high numbers of mutations in most cancers, which play an important role in the oncogenic mechanism. The lncRNA XIST can regulate X chromosome silent transcription and act as an miRNA sponge upregulating SOD2 to inhibit the development of non-small cell lung cancer (Chen et al., 2016; Liu et al., 2019). TTN-AS1 is an miRNA sponge that regulates cancer development through a ceRNA mechanism (Wang Y. et al., 2020).

## Significant Mechanism of lncRNA Mutation Perturbing ceRNA Regulation

We developed a pipeline to assess the effect of somatic mutations perturbing lncRNA-ceRNA regulation in pan-cancer (Figure 2A). To examine the influence of lncRNA mutations on miRNA binding sites according to TargetScan (Friedman et al., 2009) and miRanda (Betel et al., 2008), we used pan-cancer mutation profiles from TCGA. In total, we identified 31,560 putative somatic mutation sites for 33 cancer types in 3,124 putative miRNA target genes (putative lncRNAs). These mutated lncRNAs showed different binding affinities to 2437 miRNAs compared to wild type sequences across 33 cancers (Figure 2B). Considering that the larger numbers of putative mutations in several cancers are due to larger numbers of initial mutations, the proportions of putative mutations compared with the original lncRNA mutations were examined, which reflected the contribution of the mutations-miRNA-ceRNA mechanism in the carcinogenic process across the 33 cancers. PCPG, which had lower number of mutations and average number of mutations per sample, had the highest percentage of putative lncRNA mutations to original somatic mutations on lncRNA (Figure 2C), suggesting that the contribution of the mutations-miRNA-ceRNA mechanism in the oncogenic process is not determined simply by numbers of mutations. Next, for each putative lncRNA (ceL), we found other experimentally verified mRNAs (ceM) targeted by the same miRNA and established a minimal miRNA-ceRNA regulation unit, which we termed the somatic mutation-miRNA-ceL-ceM (SMILM) unit. Taken together, these results reveal significant mechanisms by which mutations perturb gene expression.

The frequency of somatic mutations is lower compared with genetic variations, and an appropriate number of mutation and control samples to explain the relationship between ceL and the corresponding ceM expression in SMILM units is not available (Li M. J. et al., 2017). Therefore, we defined mutations in at least two samples of the same cancer as high-frequency mutations (HF mutations,  $n = 831$ ), and the rest as low-frequency mutations (LF mutations,  $n = 32,823$ ) (Figure 2D). For these two categories of mutations, we separately applied multiple regression and Student's *t*-test to jointly model the contribution of mutations on ceRNA expression variation (see section Materials and Methods).

## Statistical Identification Portrays ceRNA Expression Fluctuation Landscape ceRNA Expression Variation Driven by HF-Mutation

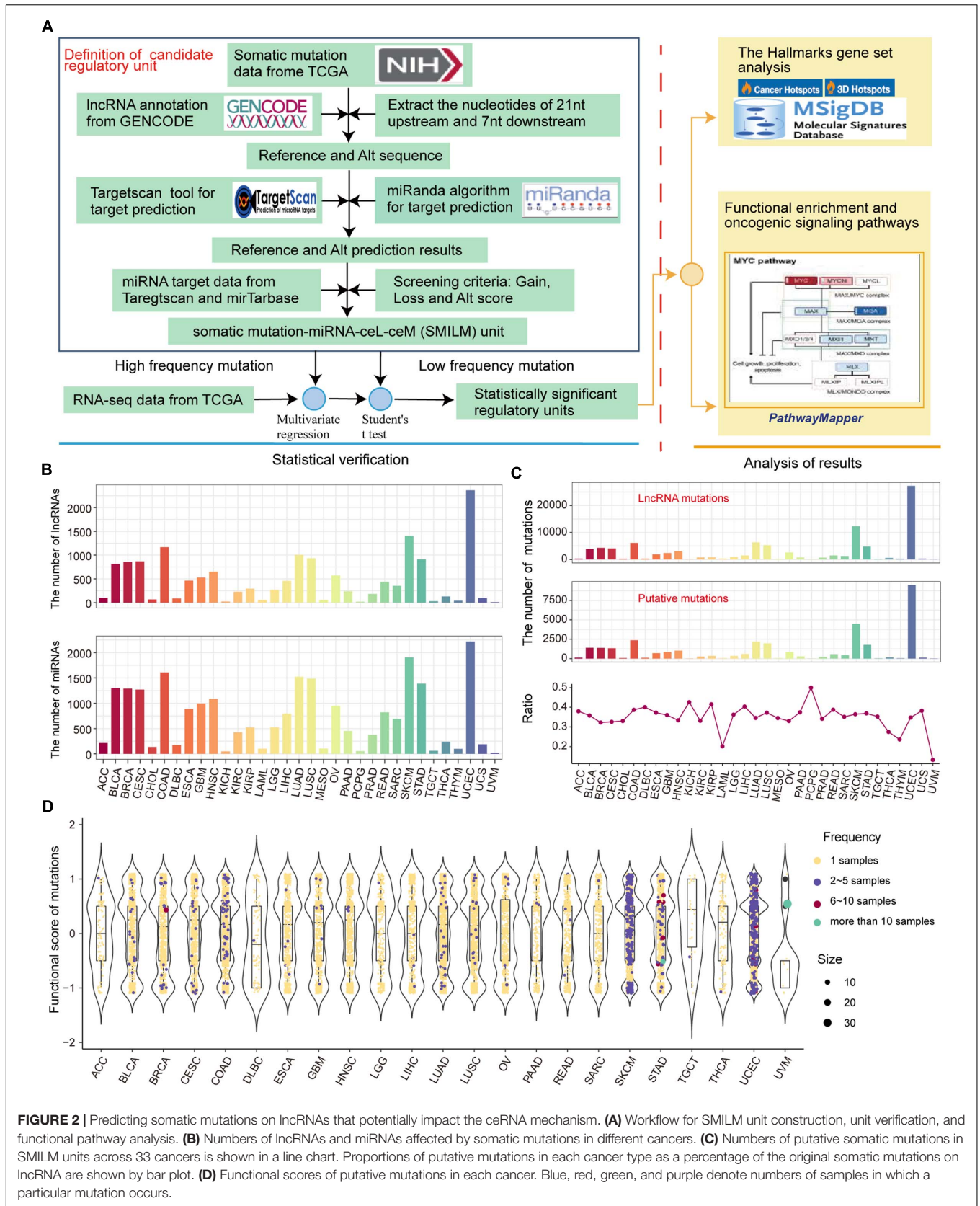
Next, we used a regression model to examine the effect of HF mutations on ceRNA expression levels. We found that only

six cancer types, including colon adenocarcinoma (COAD), head and neck squamous cell carcinoma (HNSC), LUSC, skin cutaneous melanoma (SKCM), STAD, and uterine corpus endometrial carcinoma (UCEC), had putative HF mutations that passed the regression test and identified 293 ceL and ceM genes whose expression levels were significantly correlated with genotypic changes (Figure 3A). Several factors affecting the effectiveness of ceRNAs have been reported, including the expression levels of miRNAs and ceRNAs, as well as binding affinity to miRNA target sites (Ebert and Sharp, 2010; Mukherji et al., 2011; Salmena et al., 2011). Since miRNA expression variation has been considered in the regression model, we focused on ceRNA-centric factors. We further required a consistent direction between the regression coefficient  $\eta_l$  and the changes in the functional prediction score from TargetScan and miRanda. Accordingly, we redefined 162 SMILM units, in which ceL and ceM expression variations displayed opposite and consistent orientations with the target prediction results (Figure 3A). Among 742 putative transition and transversion mutations, 17 mutations were identified to disturb the ceRNA regulation. In addition, three of 59 putative indel mutations were found to disturb the ceRNA regulation. Compared to the original HF mutations, the verified somatic mutations were drastically reduced (Figure 3B). It is likely that ceM expression changes rely not only on a minimal SMILM unit, but also on the interaction of the ceRNA network and other regulatory factors, such as transcription factors (TFs) and DNA methylation.

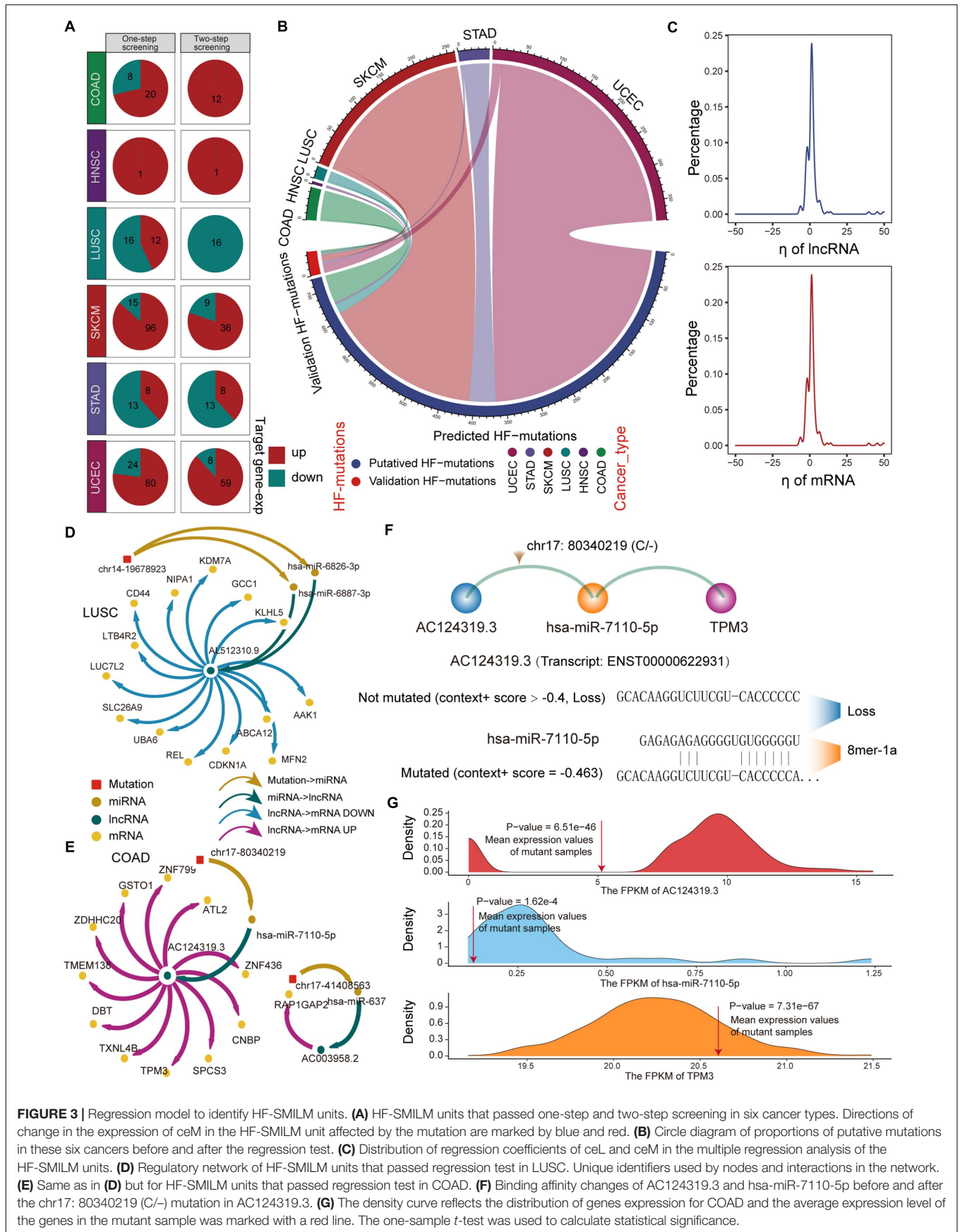
The majority of  $\eta_l$  and  $\eta_m$  were concentrated between  $-10$  and  $10$  in the 162 significant SMILM units, suggesting an important effect of these somatic mutations on ceM expression and ceRNA regulation compared with genetic variation (Li M. J. et al., 2017; Figure 3C). The mutation-ceRNA regulatory relationship was a complex network, where a single mutation affected the affinity to bind multiple miRNAs and thus disturbed the expression of multiple mRNAs (Figures 3D,E and Supplementary Figure 3). For example, an indel (chr17:80340219C) of COAD enhanced the binding affinity of hsa-miR-7110-5p in lncRNA AC124319.3 (ENSG00000280248;  $p$ -value = 3.57E-3,  $\eta_l = -8.03$ ), which competed with TPM3 ( $p$ -value = 6.43E-5,  $\eta_m = 0.79$ ; Figures 3E,F). Further, one-sample *t*-test confirmed that the expression levels of AC124319.3, hsa-miR-7110-5p, and TPM3 in non-mutated samples were statistically different from the average expression of mutant samples (Figure 3G), suggesting that the expression level of miRNA, as a key link with the ceRNA regulatory pathway, is an important factor in identifying the imbalance in the ceRNA mechanism. Taken together, these results suggest that the presence of somatic mutations in lncRNAs could affect the expression of target genes through a ceRNA regulatory pathway.

## Individual Differences in ceM Expression Variation Produced by LF-Mutation

For SMILM units disturbed by LF mutations, we focused on the target ceMs determined by the operability of the experiment and the important role of protein-coding genes in physiological function. To assess the carcinogenic function of LF-SMILM units, we extracted target ceMs involved in cancer hallmarks, which are







**FIGURE 3 |** Regression model to identify HF-SMILM units. **(A)** HF-SMILM units that passed one-step and two-step screening in six cancer types. Directions of change in the expression of ceM in the HF-SMILM unit affected by the mutation are marked by blue and red. **(B)** Circle diagram of proportions of putative mutations in these six cancers before and after the regression test. **(C)** Distribution of regression coefficients of ceL and ceM in the multiple regression analysis of the HF-SMILM units. **(D)** Regulatory network of HF-SMILM units that passed regression test in LUSC. Unique identifiers used by nodes and interactions in the network. **(E)** Same as in **(D)** but for HF-SMILM units that passed regression test in COAD. **(F)** Binding affinity changes of AC124319.3 and hsa-miR-7110-5p before and after the chr17: 80340219 (C/-) mutation in AC124319.3. **(G)** The density curve reflects the distribution of genes expression for COAD and the average expression level of the genes in the mutant sample was marked with a red line. The one-sample *t*-test was used to calculate statistical significance.

important biological processes in cancer development (Agarwal et al., 2015). We divided the expression changes of the target ceM into up, down, and none (Figure 4A, see **Supplementary Methods**). Our data suggest that lncRNA mutations amplify and depress the expression of protein-coding genes (ceMs) through the ceRNA mechanism in multiple cancers to comprehensively impact the carcinogenic process of the hallmarks.

We statistically verified the expression variation of hallmark ceMs affected by LF mutations through the ceRNA mechanism. We found that the expression levels of 1624 ceMs occurring in 32 cancer types were significantly different between the corresponding mutant and control samples and were regulated by 2849 ceL. Cancers with a high number of samples or high lncRNA mutations displayed a small number of identified ceMs (Figure 4B), suggesting that the contribution of mutation-miRNA-ceRNA mechanism is heterogeneous in pan-cancer. Fold change as a measure of change in ceMs expression was found to be clustered between 0.8 and 1.2 (Figure 4C), suggesting that simple statistical metrics mask individualized differences in the expression of mutant interference ceMs. We also found that variation in the expression of target ceM was highly correlated with the change in target prediction score in STAD and Sarcoma (SARC) (Figures 4D,E). This evidence suggests that a ceM could be affected by multiple SMILM units, which have different regulatory effects on the expression of ceM determined by individual differences in putative mutations. Together, these data indicate that individual mutation differences are an important cause of fluctuations in the expression of ceMs.

### ceM Oncogenic Pathways Reveal Pan-Cancer Functional Heterogeneities

CeMs confirmed to be affected by LF mutations in pan-cancer were collected for gene set enrichment analysis (GSEA) (Subramanian et al., 2005) weighing by fold change of ceM expression. We found significantly enriched hallmark gene sets in ceM genes only in COAD and UCEC. Allograft rejection and inflammation response pathways were enriched in ceMs with upregulated expression in both COAD and UCEC (Figures 5A,B), revealing a high similarity in the effects of lncRNA mutations through ceRNA mechanisms in COAD and UCEC. Further, we integrated all ceMs perturbed by HF and LF mutations to analyze the effect of mutation-miRNA-ceRNA mechanism on cellular functions in pan-cancer by KEGG functional enrichment analysis. We found only 21 cancers with significantly enriched functional pathways, primarily involved in energy metabolism, cell metastasis, apoptosis, and functions related to cancer complications (Figure 5C and **Supplementary Figure 4**), indicating that the mutations-miRNA-ceRNA mechanism is widely involved in the development of cancers.

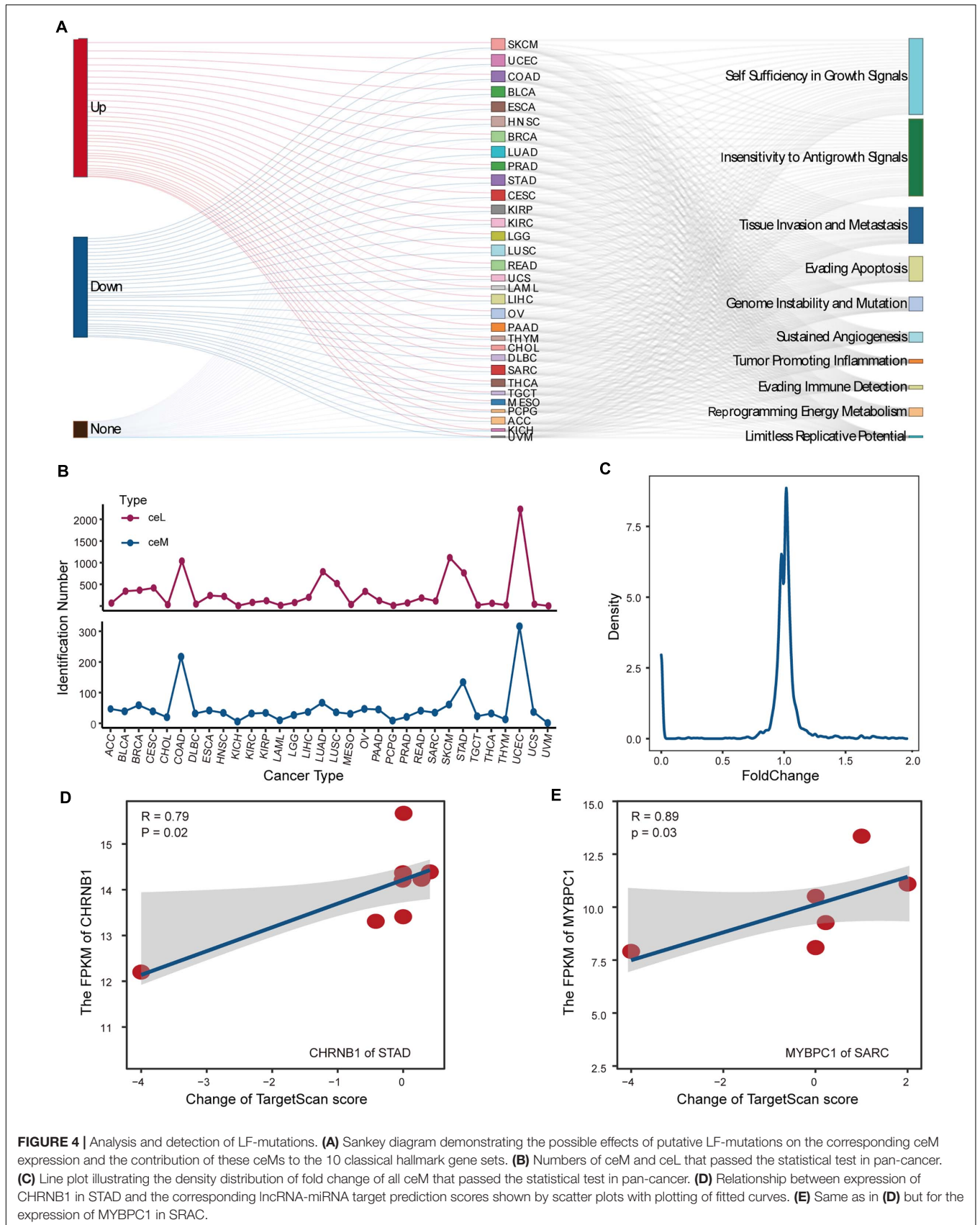
We compiled and reviewed 17 oncogenic signaling pathways verified in articles published between 2008 and 2019 (**Supplementary Table 2**). We only identified eight oncogenic signaling pathways in the KEGG functional enrichment results of eight cancer types (**Supplementary Figure 5**). These include the p53 signaling pathway (Wade et al., 2013), phosphoinositide 3-kinase (PI3K)-AKT signaling pathway, mitogen-activated

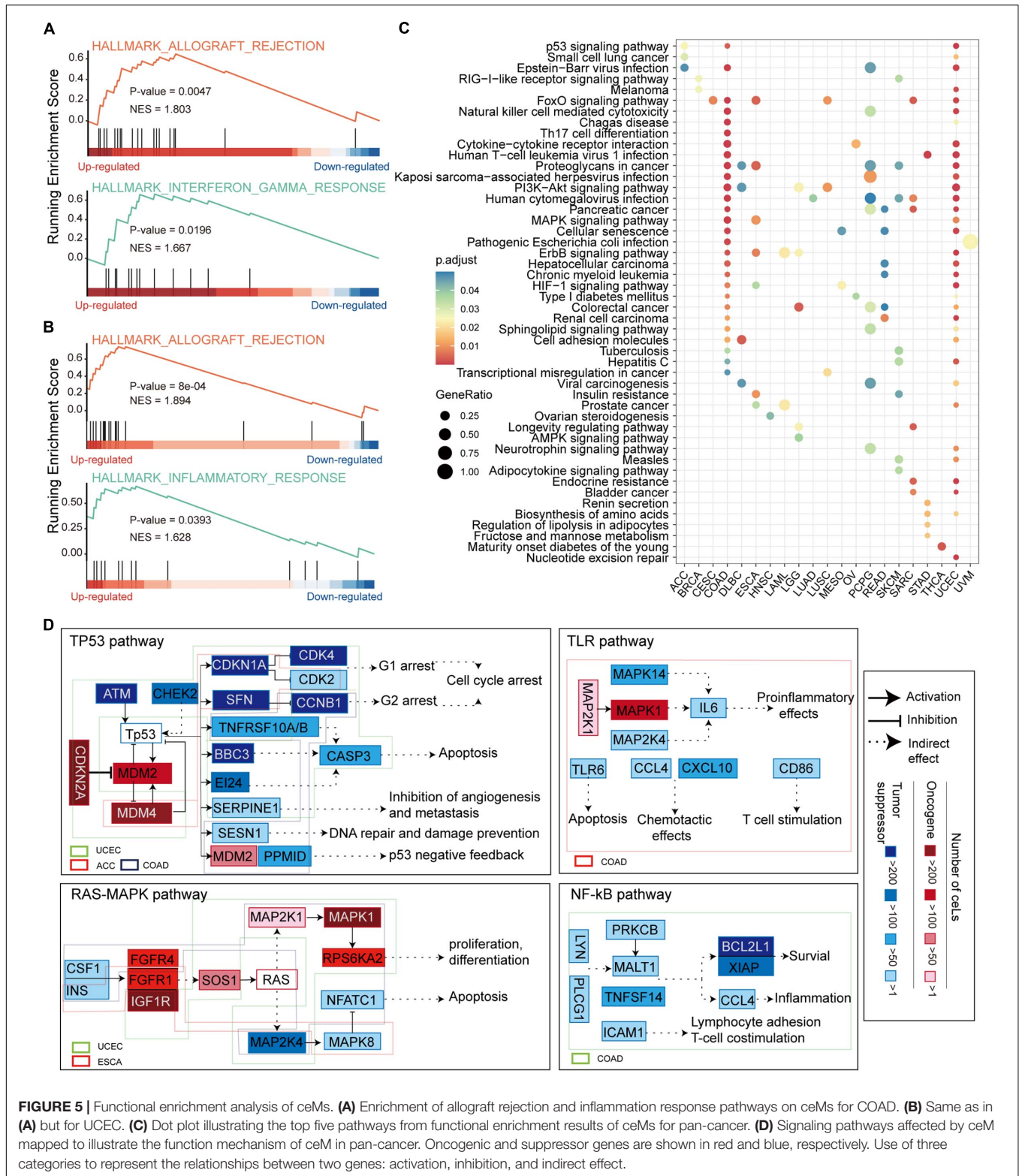
protein kinase (MAPK) signaling pathway (Soleimani et al., 2019), Ras signaling pathway (Malumbres and Barbacid, 2003), Toll-like receptor (TLR) signaling pathway (Moradi-Marjaneh et al., 2018), mammalian target of rapamycin (mTOR) signaling pathway (Ayuk and Abrahamse, 2019), Wnt signaling pathway (Reya and Clevers, 2005) and NF-kappa B (NF-kB) signaling pathway (Soleimani et al., 2020). Regarding these oncogenic signaling pathways, ceMs of UCEC, adrenocortical carcinoma (ACC), and COAD function in the p53 signaling pathway, ceMs of UCEC, and esophageal carcinoma (ESCA) function in the Ras-MAPK and mTOR signaling pathways, and ceMs of brain lower grade glioma (LGG), lymphoid neoplasm diffuse large B-cell lymphoma (DLBC), LUSC, COAD, and UCEC function in the PI3K-AKT signaling pathway (Figure 5D and **Supplementary Figure 6**). These findings suggest that pan-cancer ceMs regulate oncogenic signaling pathways in a flexible manner with certain similarities. The ceMs in COAD were mainly enriched in the p53, TLR, and NF-kB oncogenic signaling pathways, which regulate cell cycle arrest, DNA repair, apoptosis, proinflammatory effects, inflammation, and survival. Furthermore, the ceMs in UCEC regulated cell cycle arrest and apoptosis in the p53 pathway, proliferation and apoptosis in the Ras-MAPK pathway, lipid biosynthesis and autophagy in the mTOR pathway, and angiogenesis and DNA repair in the PI3K-Akt pathway. These results indicate that pan-cancer, where the same oncogenic pathways are regulated through the mutations-miRNA-ceRNA mechanism, is heterogeneous in specific functions. The MDM2 and MDM4 ceMs in COAD and UCEC were found to be important for the stabilization and activation of p53 and could serve as important targets for anti-cancer therapy (Toledo and Wahl, 2007; Wade et al., 2010). The NF- $\kappa$ B signaling pathway is a typical proinflammatory signal transduction pathway and an important target for novel anti-carcinogenic drugs (Lawrence, 2009). mTOR is critical in the pathway and promotes cancer proliferation and metabolism upon overactivation, which is also an important target for cancer therapy (Tian et al., 2019). Taken together, these results suggest that functional variations in pan-cancer that are disturbed by somatic mutations through the ceRNA mechanism may lead to tumor-specific phenotypes.

### Survival Analysis Reveals Biomarker lncRNA in Pan-Cancer

We evaluated the impact of ceMs that participate in the carcinogenic signaling pathway on the survival of cancer patients. Several genes (ceMs) in five cancer types had a significant hazard ratio (HR) value according to Cox proportional hazard model, suggesting a relationship between patient prognosis and lncRNA mutations (**Supplementary Figure 7**). The Kaplan–Meier method was used to plot survival-related ceMs (Figures 6A–C and **Supplementary Figures 8A,B**). We found that samples resulting in the upregulation of EREG in COAD had poor overall survival and that such samples were enriched in stage IV (Figure 6A). It is intriguing to note that according to Kaplan–Meier analysis, there was no significant difference in the survival of BCL2L1 between mutant and control samples (Figure 6A), which may be attributed to the weak effect of







**FIGURE 5 |** Functional enrichment analysis of ceMs. **(A)** Enrichment of allograft rejection and inflammation response pathways on ceMs for COAD. **(B)** Same as in **(A)** but for UCEC. **(C)** Dot plot illustrating the top five pathways from functional enrichment results of ceMs in pan-cancer. **(D)** Signaling pathways affected by ceM mapped to illustrate the function mechanism of ceM in pan-cancer. Oncogenic and suppressor genes are shown in red and blue, respectively. Use of three categories to represent the relationships between two genes: activation, inhibition, and indirect effect.

mutations on ceM expression in several samples. FGFR1, a proven independent prognostic risk factor in patients with resected esophageal squamous cell carcinoma (Wang Y. et al., 2019), was significantly differentially expressed in mutant ESCA

samples ( $p$ -value =  $7.56e-04$ ), and high expression of FGFR1 was associated with poorer patient prognosis (**Figure 6B**). We found that “unknown” samples that perturb MAPK1 expression in UCEC had a better prognosis and the least proportion was

in stage IV (**Figure 6C**). Previous studies have suggested that MAPK1 regulates the metastasis and invasion of cervical cancer through a ceRNA mechanism (Li W. et al., 2017). These results indicate an important contribution of the lncRNA mutation-ceRNA mechanism to the overall survival of cancer patients.

lncRNAs that regulate prognosis-related ceM expression can be critical factors in cancers. High-frequency mutated lncRNAs regulating prognosis-related ceM expression were screened for the identification of cancer-related biomarkers (**Supplementary Table 3, Supplementary Figures 8C,D, and Figures 6D–F**). TTN-AS1 (ENSG00000237298) was identified as a potential biomarker involved in the regulation of both EREG and BCL2L1 (**Figure 6D**). TTN-AS1 has been proven to be associated with the prognosis of COAD and regulate apoptosis and invasion in osteosarcoma (Fu et al., 2019), lung adenocarcinoma (Jia et al., 2019) and colorectal cancer (Wang Y. et al., 2020) via the ceRNA mechanism. The lncRNA GSN-AS1 (ENSG00000235865) regulates BCL2L1 (**Figures 6D,E**), which has been proven to be an important prognostic marker for luminal subtype breast cancer (Yang et al., 2016). lncRNA XIST (ENSG00000229807), TTN-AS1 (ENSG00000237298), and TSIX (ENSG00000270641), which regulate MAPK1 and ERBB3 in UCEC (**Figure 6F**), were shown to be miRNA sponges that control apoptosis via the ceRNA mechanism (Bu et al., 2018; Fu et al., 2019; Li et al., 2020). Taken together, these results suggest that several potential biomarker lncRNAs regulate the expression of protein-coding genes through the ceRNA mechanism to affect patient survival.

## DISCUSSION

In this study, we integrated mutation data from 33 cancer types with RNA-seq profiles from TCGA to explore the association between somatic mutations in lncRNA and the regulatory mechanism of ceRNA. Using multivariate multiple regression and statistical analyses, we identified 162 significant HF-SMILM units and many ceMs perturbed by LF mutations from pan-cancers. The mutations-miRNA-ceRNA mechanism appeared to be dynamic, with individual differences in the regulation of ceRNA expression due to mutation specificity. In addition, we characterized the function of ceMs in pan-cancer through oncogenic signaling pathway studies and survival analysis, identifying biomarker lncRNAs that regulate the expression of ceM associated with patient survival. These findings provide a new perspective to explain the role of lncRNA mutations in post-transcriptional gene regulation.

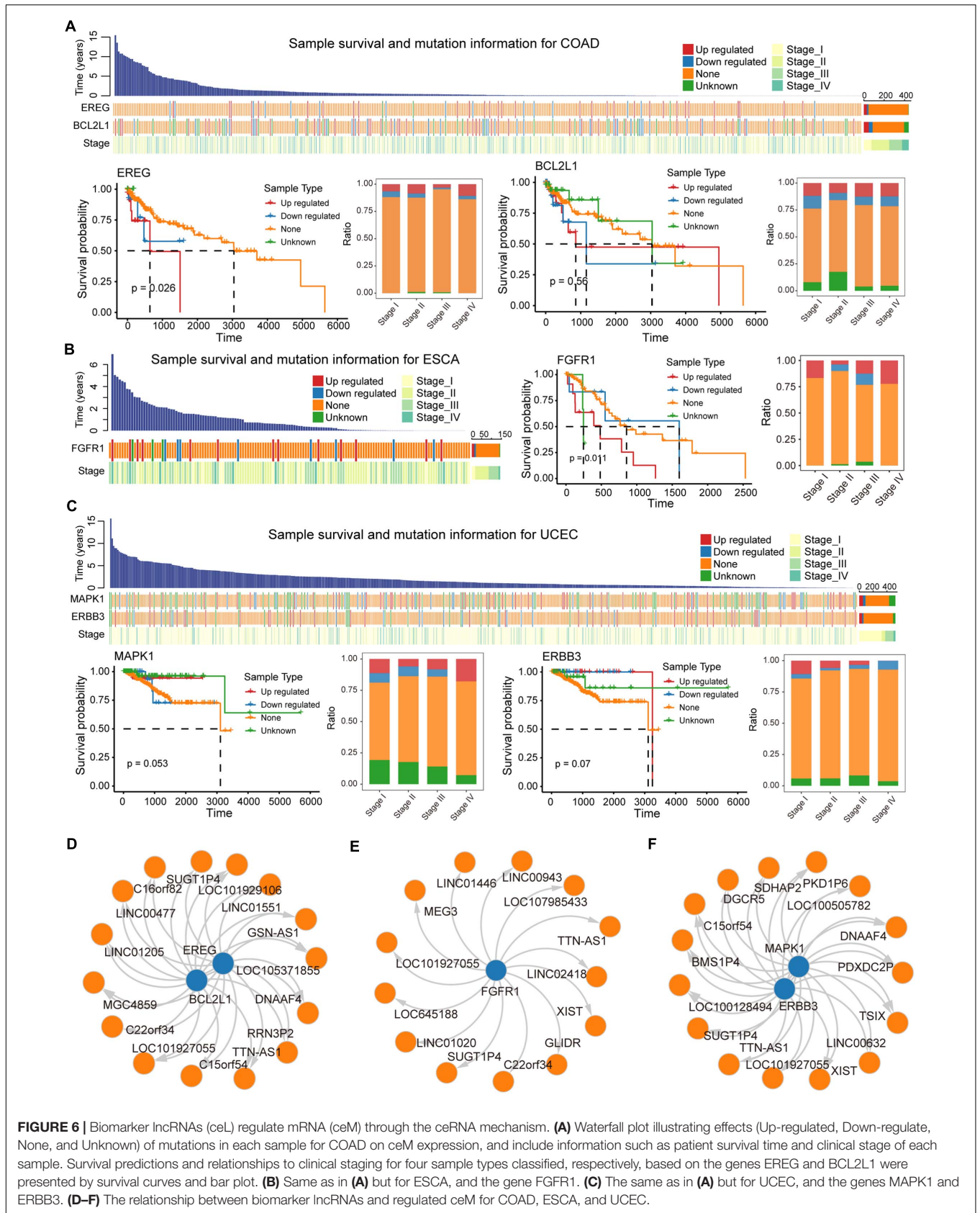
Although we used both TargetScan and miRanda tools to predict potential SMILM units with rigorous screening of scores and energetics, it is also possible that our lncRNA mutation-miRNA detection missed target sites not predicted by either tool. Alternatively, we considered the union or intersection of multiple miRNA-target prediction algorithms, including PITA (Kertesz et al., 2007) and RNAhybrid (Kruger and Rehmsmeier, 2006); however, unions might introduce many false positives and intersections might introduce false negatives. Our current standards provide a reasonable and reliable reference for further

functional research, and experimental methods such as CLIP-seq and RIP-seq can be used to overcome these shortcomings. Another potential limitation of our study is that we only considered one competing unit for detection in a complex ceRNA regulatory network. Changes in the expression of a node gene in the ceRNA regulatory network perturbs the balance of the entire network (Levine et al., 2007). The cascade effect caused by miRNA redistribution and ceRNA competition from a global perspective requires building a more complex algorithm to more accurately describe the complete response of the entire network.

It is certainly a significant idea that we consider lncRNA somatic mutations in terms of the ceRNA regulatory mechanism. The effects of genetic variation on ceRNA regulation have been revealed by previous studies (Cheng et al., 2015; Ghanbari et al., 2015; Li M. J. et al., 2017), and a large database of correlations has emerged (Li et al., 2014). However, few studies have addressed somatic mutations (Wang P. et al., 2020). Previous studies have focused on the perturbation of ceRNA mechanisms by genetic variations (Gamazon et al., 2012; Ghanbari et al., 2015; Li M. J. et al., 2017; Wang P. et al., 2020). For example, one of our prior studies determined the effect of somatic mutations of lncRNA on ceRNA mechanisms in pan-cancers (Wang P. et al., 2020). Li et al. have explored the genetic associations with ceRNA regulation in the human genome. These studies focused on methodology development and dataset construction on how to connect genomic variations and ceRNA expression. Thus, further studies evaluating the effects of mutations on ceRNA expression and downstream function are needed. Our research aimed to provide new insights into the oncogenic mechanism from the perspective of somatic mutations perturbing the ceRNA mechanism. Further, we expanded the scope of our analysis to study the effect of mutations on perturbing biological networks, functions, and clinical phenotypes. We believe that our analysis will be helpful for dissecting disease pathology caused by personalized mutations and further contribute to precision medicine.

Despite the limitations of our study, our findings reveal one of the underlying causes of changes in the physiological functions in cancer, which will help advance the development of precision medicine. Using mutation and transcriptome data from multiple cancers, we found many cancer-specific SMILM units and identified ceMs affected by ceLs. These results complement recent studies on the mechanism by which lncRNA mutations perturb ceRNA (Bhattacharya and Cui, 2016; Wang P. et al., 2020). By verifying the HF-mutation SMILM unit, we discovered the mutation-mediated ceRNA expression fluctuation mechanism. We also found individual differences in changes in the expression of ceM. The diverse distribution of HF and LF mutations in different cancers and genes was consistent with the genetic heterogeneity of different cancers. Thus, we performed a specific investigation on specific cancer types and genes based on the background of tumor heterogeneity. This strategy has been previously used to characterize individual disease pathologies caused by cancer-specific or gene-specific mutations (Lawrence et al., 2013; Zack et al., 2013; ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium, 2020). We believe that our exhaustive analysis will be helpful for dissecting disease pathology





**FIGURE 6 |** Biomarker lncRNAs (ceL) regulate mRNA (ceM) through the ceRNA mechanism. **(A)** Waterfall plot illustrating effects (Up-regulated, Down-regulate, None, and Unknown) of mutations in each sample for COAD on ceM expression, and include information such as patient survival time and clinical stage of each sample. Survival predictions and relationships to clinical staging for four sample types classified, respectively, based on the genes EREG and BCL2L1 were presented by survival curves and bar plot. **(B)** Same as in **(A)** but for ESCA, and the gene FGFR1. **(C)** The same as in **(A)** but for UCEC, and the genes MAPK1 and ERBB3. **(D-F)** The relationship between biomarker lncRNAs and regulated ceM for COAD, ESCA, and UCEC.

caused by personalized mutations and contribute to precision medicine. Importantly, we performed functional enrichment and oncogenic signature pathway studies, as well as survival analysis of ceMs from multiple cancers. We identified that pan-cancer has functional heterogeneity in mutation-ceRNA mechanisms, primarily enriched in cell proliferation and apoptosis, DNA repair, and immune regulation. Furthermore, FLT1 of LUSC, ITGB1 of DLBC, MDM4, and CDKN1A of ACC, MAPK1, and ERBB3 of UCEC, FGFR1 of ESCA, and EREG and BCL2L1 of COAD have been strongly associated with patient survival in their respective cancers. Furthermore, the biomarker lncRNAs, which regulate the above genes and are mutated with HF, contribute to the development of clinical research.

With the rapid development of high-throughput technologies, an increasing number of large biological data sets can be obtained at the whole-transcriptome level. This makes it difficult to dissect the individual pathologies behind the fast-growing datasets. Although many novel biomarkers have been identified by *in vivo* or *in vitro* experimental methods, identifying new disease-biomarker associations based on traditional, one-by-one experimental studies are expensive, complex, and time-consuming. To overcome these problems, a bioinformatics strategy has been used in previous studies to dissect gene regulation and revealed valuable results (Du et al., 2013; Li et al., 2015; Wang P. et al., 2015). We believe that our analyses will provide novel insights into mutations affecting lncRNA-associated regulatory mechanisms at the transcriptional level. Both the method and predictions could serve as helpful references for future experimental and functional dissections of lncRNAs.

## CONCLUSION

Our study provides a global landscape of the effects of lncRNA somatic mutations on the ceRNA mechanism in pan-cancer. Our findings extend existing knowledge on the relevance of lncRNA mutations in functions related to cancer via the ceRNA mechanism. The integration of mutation and RNA

expression data from tumor samples enhances the interpretation of the identified SMILMs, helping to improve the reliability of the predictions; thus, this approach may provide more precise theoretical guidance for experimental studies and clinical applications.

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

## AUTHOR CONTRIBUTIONS

PW, BC, and YaZ conceived and designed the experiments. YuZ, PH, QG, and YH analyzed data. YQ and MX collected the data. YuZ, PH, and QG validated the method and data. YuZ and PH wrote this manuscript. All authors read and approved the final manuscript.

## FUNDING

This work was supported by the National Natural Science Foundation of China (32070622, 81702766, 82072640, and 81902646), Heilongjiang Provincial Natural Science Foundation (LH2020H046), Postdoctoral Science Foundation of China (2020M670922), and Postdoctoral Foundation of Heilongjiang Province (LBH-Z19077).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcell.2021.658346/full#supplementary-material>

## REFERENCES

- Agarwal, V., Bell, G. W., Nam, J. W., and Bartel, D. P. (2015). Predicting effective microRNA target sites in mammalian mRNAs. *eLife* 4:e05005. doi: 10.7554/eLife.05005.028
- Ayuk, S. M., and Abrahamse, H. (2019). mTOR signaling pathway in cancer targets photodynamic therapy *in vitro*. *Cells* 8:431. doi: 10.3390/cells8050431
- 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. doi: 10.1093/nar/gkm995
- Bhattacharya, A., and Cui, Y. (2016). Somamir 2.0: a database of cancer somatic mutations altering microRNA-ceRNA interactions. *Nucleic Acids Res.* 44, D1005–D1010. doi: 10.1093/nar/gkv1220
- Bu, Y., Zheng, D., Wang, L., and Liu, J. (2018). lncRNA TSIX promotes osteoblast apoptosis in particle-induced osteolysis by down-regulating miR-30a-5p. *Connect. Tissue Res.* 59, 534–541. doi: 10.1080/03008207.2017.1413362
- Cancer Genome Atlas Research Network, Linehan, W. M., Spellman, P. T., Ricketts, C. J., Creighton, C. J., Fei, S. S., et al. (2016). Comprehensive molecular characterization of papillary renal-cell carcinoma. *N. Engl. J. Med.* 374, 135–145. doi: 10.1056/NEJMoa1505917
- Chen, C. K., Blanco, M., Jackson, C., Aznauryan, E., Ollikainen, N., Surka, C., et al. (2016). Xist recruits the X chromosome to the nuclear lamina to enable chromosome-wide silencing. *Science* 354, 468–472. doi: 10.1126/science.aae0047
- Chen, Y., Lin, Y., Bai, Y., Cheng, D., Bi, Z., and Long Noncoding, A. (2019). RNA (lncRNA)-associated competing endogenous RNA (ceRNA) network identifies eight lncRNA biomarkers in patients with osteoarthritis of the knee. *Med. Sci. Monit.* 25, 2058–2065. doi: 10.12659/MSM.915555
- Cheng, D. L., Xiang, Y. Y., Ji, L. J., and Lu, X. J. (2015). Competing endogenous RNA interplay in cancer: mechanism, methodology, and perspectives. *Tumour Biol.* 36, 479–488. doi: 10.1007/s13277-015-3093-z
- Chou, C. H., Shrestha, S., Yang, C. D., Chang, N. W., Lin, Y. L., Liao, K. W., et al. (2018). miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions. *Nucleic Acids Res.* 46, D296–D302. doi: 10.1093/nar/gkx1067
- Civelek, M., Hagopian, R., Pan, C., Che, N., Yang, W. P., Kayne, P. S., et al. (2013). Genetic regulation of human adipose microRNA expression and its

- consequences for metabolic traits. *Hum. Mol. Genet.* 22, 3023–3037. doi: 10.1093/hmg/ddt159
- Du, Z., Fei, T., Verhaak, R. G., Su, Z., Zhang, Y., Brown, M., et al. (2013). Integrative genomic analyses reveal clinically relevant long noncoding RNAs in human cancer. *Nat. Struct. Mol. Biol.* 20, 908–913. doi: 10.1038/nsmb.2591
- Ebert, M. S., and Sharp, P. A. (2010). Emerging roles for natural microRNA sponges. *Curr. Biol.* 20, R858–R861. doi: 10.1016/j.cub.2010.08.052
- Fisher, L. D., and Lin, D. Y. (1999). Time-dependent covariates in the Cox proportional-hazards regression model. *Annu. Rev. Public Health* 20, 145–157. doi: 10.1146/annurev.publhealth.20.1.145
- Friedman, R. C., Farh, K. K., Burge, C. B., and Bartel, D. P. (2009). Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.* 19, 92–105. doi: 10.1101/gr.082701.108
- Fu, D., Lu, C., Qu, X., Li, P., Chen, K., Shan, L., et al. (2019). LncRNA TTN-AS1 regulates osteosarcoma cell apoptosis and drug resistance via the miR-134-5p/MBTD1 axis. *Aging (Albany NY)* 11, 8374–8385. doi: 10.18632/aging.102325
- Gamazon, E. R., Ziliak, D., Im, H. K., LaCroix, B., Park, D. S., Cox, N. J., et al. (2012). Genetic architecture of microRNA expression: implications for the transcriptome and complex traits. *Am. J. Hum. Genet.* 90, 1046–1063. doi: 10.1016/j.ajhg.2012.04.023
- Ghanbari, M., Franco, O. H., de Looper, H. W., Hofman, A., Erkeland, S. J., and Dehghan, A. (2015). Genetic variations in microRNA-binding sites affect microRNA-mediated regulation of several genes associated with cardio-metabolic phenotypes. *Circ. Cardiovasc. Genet.* 8, 473–486. doi: 10.1161/CIRCGENETICS.114.000968
- Hanahan, D., and Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *Cell* 144, 646–674. doi: 10.1016/j.cell.2011.02.013
- Harrow, J., Frankish, A., Gonzalez, J. M., Tapanari, E., Diekhans, M., Kokocinski, F., et al. (2012). GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res.* 22, 1760–1774. doi: 10.1101/gr.135350.111
- ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium (2020). Pan-cancer analysis of whole genomes. *Nature* 578, 82–93. doi: 10.1038/s41586-020-1969-6
- Jia, P., and Zhao, Z. (2017). Impacts of somatic mutations on gene expression: an association perspective. *Brief Bioinform.* 18, 413–425.
- Jia, Y., Duan, Y., Liu, T., Wang, X., Lv, W., Wang, M., et al. (2019). LncRNA TTN-AS1 promotes migration, invasion, and epithelial mesenchymal transition of lung adenocarcinoma via sponging miR-142-5p to regulate CDK5. *Cell Death Dis.* 10:573. doi: 10.1038/s41419-019-1811-y
- Kertesz, M., Iovino, N., Unnerstall, U., Gaul, U., and Segal, E. (2007). The role of site accessibility in microRNA target recognition. *Nat. Genet.* 39, 1278–1284. doi: 10.1038/ng2135
- Kozomara, A., and Griffiths-Jones, S. (2014). miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res.* 42, D68–D73. doi: 10.1093/nar/gkt1181
- Kruger, J., and Rehmsmeier, M. (2006). RNAhybrid: microRNA target prediction easy, fast and flexible. *Nucleic Acids Res.* 34, W451–W454. doi: 10.1093/nar/gkl243
- Lawrence, M. S., Stojanov, P., Polak, P., Kryukov, G. V., Cibulskis, K., Sivachenko, A., et al. (2013). Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* 499, 214–218. doi: 10.1038/nature12213
- Lawrence, T. (2009). The nuclear factor NF- $\kappa$ B pathway in inflammation. *Cold Spring Harb. Perspect. Biol.* 1:a001651. doi: 10.1101/cshperspect.a001651
- Lemos, A. E. G., Matos, A. D. R., Ferreira, L. B., and Gimba, E. R. P. (2019). The long non-coding RNA PCA3: an update of its functions and clinical applications as a biomarker in prostate cancer. *Oncotarget* 10, 6589–6603. doi: 10.18632/oncotarget.27284
- Levine, E., Zhang, Z., Kuhlman, T., and Hwa, T. (2007). Quantitative characteristics of gene regulation by small RNA. *PLoS Biol.* 5:e229. doi: 10.1371/journal.pbio.0050229
- Li, J. H., Liu, S., Zhou, H., Qu, L. H., and Yang, J. H. (2014). starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res.* 42, D92–D97. doi: 10.1093/nar/gkt1248
- Li, L., Lv, G., Wang, B., and Kuang, L. (2020). XIST/miR-376c-5p/OPN axis modulates the influence of proinflammatory M1 macrophages on osteoarthritis chondrocyte apoptosis. *J. Cell Physiol.* 235, 281–293. doi: 10.1002/jcp.28968
- Li, M. J., Zhang, J., Liang, Q., Xuan, C., Wu, J., Jiang, P., et al. (2017). Exploring genetic associations with ceRNA regulation in the human genome. *Nucleic Acids Res.* 45, 5653–5665. doi: 10.1093/nar/gkx331
- Li, W., Liang, J., Zhang, Z., Lou, H., Zhao, L., Xu, Y., et al. (2017). MicroRNA-329-3p targets MAPK1 to suppress cell proliferation, migration and invasion in cervical cancer. *Oncol. Rep.* 37, 2743–2750. doi: 10.3892/or.2017.5555
- Li, Y., Shao, T., Jiang, C., Bai, J., Wang, Z., Zhang, J., et al. (2015). Construction and analysis of dynamic transcription factor regulatory networks in the progression of glioma. *Sci. Rep.* 5:15953. doi: 10.1038/srep15953
- Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdottir, H., Tamayo, P., and Mesirov, J. P. (2011). Molecular signatures database (MSigDB) 3.0. *Bioinformatics* 27, 1739–1740. doi: 10.1093/bioinformatics/btr260
- Liu, J., Yao, L., Zhang, M., Jiang, J., Yang, M., and Wang, Y. (2019). Downregulation of LncRNA-XIST inhibited development of non-small cell lung cancer by activating miR-335/SOD2/ROS signal pathway mediated pyroptotic cell death. *Aging (Albany NY)* 11, 7830–7846. doi: 10.18632/aging.102291
- Long, J., Xiong, J., Bai, Y., Mao, J., Lin, J., Xu, W., et al. (2019). Construction and investigation of a lncRNA-associated ceRNA regulatory network in cholangiocarcinoma. *Front. Oncol.* 9:649. doi: 10.3389/fonc.2019.00649
- Lu, J., and Clark, A. G. (2012). Impact of microRNA regulation on variation in human gene expression. *Genome Res.* 22, 1243–1254. doi: 10.1101/gr.132514.111
- Malumbres, M., and Barbacid, M. (2003). RAS oncogenes: the first 30 years. *Nat. Rev. Cancer* 3, 459–465. doi: 10.1038/nrc1097
- Martincorena, I., and Campbell, P. J. (2015). Somatic mutation in cancer and normal cells. *Science* 349, 1483–1489. doi: 10.1126/science.aab4082
- Moradi-Marjaneh, R., Hassanian, S. M., Fiuji, H., Soleimanpour, S., Ferns, G. A., Avan, A., et al. (2018). Toll like receptor signaling pathway as a potential therapeutic target in colorectal cancer. *J. Cell Physiol.* 233, 5613–5622. doi: 10.1002/jcp.26273
- Mukherji, S., Ebert, M. S., Zheng, G. X., Tsang, J. S., Sharp, P. A., and van Oudenaarden, A. (2011). MicroRNAs can generate thresholds in target gene expression. *Nat. Genet.* 43, 854–859. doi: 10.1038/ng.905
- Quinn, J. J., and Chang, H. Y. (2016). Unique features of long non-coding RNA biogenesis and function. *Nat. Rev. Genet.* 17, 47–62. doi: 10.1038/nrg.2015.10
- Reya, T., and Clevers, H. (2005). Wnt signalling in stem cells and cancer. *Nature* 434, 843–850. doi: 10.1038/nature03319
- Rich, J. T., Neely, J. G., Paniello, R. C., Voelker, C. C., Nussenbaum, B., and Wang, E. W. (2010). A practical guide to understanding Kaplan-Meier curves. *Otolaryngol. Head Neck Surg.* 143, 331–336. doi: 10.1016/j.otohns.2010.05.007
- Ricketts, C. J., De Cubas, A. A., Fan, H., Smith, C. C., Lang, M., Reznik, E., et al. (2018). The cancer genome atlas comprehensive molecular characterization of renal cell carcinoma. *Cell Rep.* 23, 313–326.e5. doi: 10.1016/j.celrep.2018.03.075
- Rinn, J. L., and Chang, H. Y. (2012). Genome regulation by long noncoding RNAs. *Annu. Rev. Biochem.* 81, 145–166. doi: 10.1146/annurev-biochem-051410-092902
- Salmena, L., Poliseno, L., Tay, Y., Kats, L., and Pandolfi, P. P. (2011). A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell* 146, 353–358. doi: 10.1016/j.cell.2011.07.014
- Schmitt, A. M., and Chang, H. Y. (2016). Long noncoding RNAs in cancer pathways. *Cancer Cell* 29, 452–463. doi: 10.1016/j.ccell.2016.03.010
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303
- Siddle, K. J., Deschamps, M., Tailleux, L., Nedelec, Y., Pothlichet, J., Lugo-Villarino, G., et al. (2014). A genomic portrait of the genetic architecture and regulatory impact of microRNA expression in response to infection. *Genome Res.* 24, 850–859. doi: 10.1101/gr.161471.113
- Soleimani, A., Rahmani, F., Ferns, G. A., Ryzhikov, M., Avan, A., and Hassanian, S. M. (2020). Role of the NF- $\kappa$ B signaling pathway in the pathogenesis of colorectal cancer. *Gene* 726:144132. doi: 10.1016/j.gene.2019.144132
- Soleimani, A., Rahmani, F., Saeedi, N., Ghaffarian, R., Khazaei, M., Ferns, G. A., et al. (2019). The potential role of regulatory microRNAs of RAS/MAPK



- signaling pathway in the pathogenesis of colorectal cancer. *J. Cell Biochem.* 120, 19245–19253. doi: 10.1002/jcb.29268
- Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* 102, 15545–15550. doi: 10.1073/pnas.0506580102
- Thomas, L. F., Saito, T., and Saetrom, P. (2011). Inferring causative variants in microRNA target sites. *Nucleic Acids Res.* 39:e109. doi: 10.1093/nar/gkr414
- Thomson, D. W., and Dinger, M. E. (2016). Endogenous microRNA sponges: evidence and controversy. *Nat. Rev. Genet.* 17, 272–283. doi: 10.1038/nrg.2016.20
- Tian, T., Li, X., and Zhang, J. (2019). mTOR signaling in cancer and mTOR inhibitors in solid tumor targeting therapy. *Int. J. Mol. Sci.* 20:755. doi: 10.3390/ijms20030755
- Toledo, F., and Wahl, G. M. (2007). MDM2 and MDM4: p53 regulators as targets in anticancer therapy. *Int. J. Biochem. Cell Biol.* 39, 1476–1482. doi: 10.1016/j.biocel.2007.03.022
- Tomczak, K., Czerwinska, P., and Wiznerowicz, M. (2015). The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Contemp. Oncol. (Pozn)* 19, A68–A77. doi: 10.5114/wo.2014.47136
- Tripathi, V., Shen, Z., Chakraborty, A., Giri, S., Freier, S. M., Wu, X., et al. (2013). Long noncoding RNA MALAT1 controls cell cycle progression by regulating the expression of oncogenic transcription factor B-MYB. *PLoS. Genet.* 9:e1003368. doi: 10.1371/journal.pgen.1003368
- Valente, G., Castellanos, A. L., Vanacore, G., and Formisano, E. (2014). Multivariate linear regression of high-dimensional fMRI data with multiple target variables. *Hum. Brain Mapp.* 35, 2163–2177. doi: 10.1002/hbm.22318
- Wade, M., Li, Y. C., and Wahl, G. M. (2013). MDM2, MDMX and p53 in oncogenesis and cancer therapy. *Nat. Rev. Cancer* 13, 83–96. doi: 10.1038/nrc3430
- Wade, M., Wang, Y. V., and Wahl, G. M. (2010). The p53 orchestra: Mdm2 and Mdmx set the tone. *Trends Cell Biol.* 20, 299–309. doi: 10.1016/j.tcb.2010.01.009
- Wang, P., Li, X., Gao, Y., Guo, Q., Ning, S., Zhang, Y., et al. (2020). LnCeVar: a comprehensive database of genomic variations that disturb ceRNA network regulation. *Nucleic Acids Res.* 48, D111–D117. doi: 10.1093/nar/gkz887
- Wang, P., Ning, S., Zhang, Y., Li, R., Ye, J., Zhao, Z., et al. (2015). Identification of lncRNA-associated competing triplets reveals global patterns and prognostic markers for cancer. *Nucleic Acids Res.* 43, 3478–3489. doi: 10.1093/nar/gkv233
- Wang, Y., Jiang, F., Xiong, Y., Cheng, X., Qiu, Z., and Song, R. (2020). LncRNA TTN-AS1 sponges miR-376a-3p to promote colorectal cancer progression via upregulating KLF15. *Life Sci.* 244:116936. doi: 10.1016/j.lfs.2019.116936
- Wang, Y., Wu, Y., Li, J., Lai, Y., Zhou, K., and Che, G. (2019). Prognostic and clinicopathological significance of FGFR1 gene amplification in resected esophageal squamous cell carcinoma: a meta-analysis. *Ann. Transl. Med.* 7:669. doi: 10.21037/atm.2019.10.69
- Yang, F., Lv, S. X., Lv, L., Liu, Y. H., Dong, S. Y., Yao, Z. H., et al. (2016). Identification of lncRNA FAM83H-AS1 as a novel prognostic marker in luminal subtype breast cancer. *Onco Targets Ther.* 9, 7039–7045. doi: 10.2147/OTT.S110055
- Zack, T. I., Schumacher, S. E., Carter, S. L., Cherniack, A. D., Saksena, G., Tabak, B., et al. (2013). Pan-cancer patterns of somatic copy number alteration. *Nat. Genet.* 45, 1134–1140. doi: 10.1038/ng.2760

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Zhang, Han, Guo, Hao, Qi, Xin, Zhang, Cui and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.