

RESEARCH ARTICLE

A multi-omics study delineates new molecular features and therapeutic targets for esophageal squamous cell carcinoma

Xing Jin¹ | Lei Liu² | Jia Wu¹ | Xiaoxia Jin³ | Guanzhen Yu¹ |
Lijun Jia¹ | Fengying Wang¹ | Minxin Shi² | Haimin Lu² | Jibin Liu² |
Dan Liu¹ | Jing Yang¹ | Hua Li⁴ | Yan Ni⁵ | Qin Luo¹ | Wei Jia⁶ |
Wei Wang² | Wen-Lian Chen¹

¹ Cancer Institute, Longhua Hospital, Shanghai University of Traditional Chinese Medicine, Shanghai, China

² Department of Thoracic Surgery, The Affiliated Tumor Hospital of Nantong University, Nantong, China

³ Department of Pathology, The Affiliated Tumor Hospital of Nantong University, Nantong, China

⁴ Bio-ID Center, School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, China

⁵ The Children's Hospital, National Clinical Research Center for Child Health, Zhejiang University School of Medicine, Hangzhou, China

⁶ Hong Kong Traditional Chinese Medicine Phenome Research Center, School of Chinese Medicine, Hong Kong Baptist University, Kowloon Tong, Hong Kong, China

Correspondence

Wei Jia, Hong Kong Traditional Chinese Medicine Phenome Research Center, School of Chinese Medicine, Hong Kong Baptist University, Kowloon Tong 999077, Hong Kong, China.

Email: weijial@hkbu.edu.hk

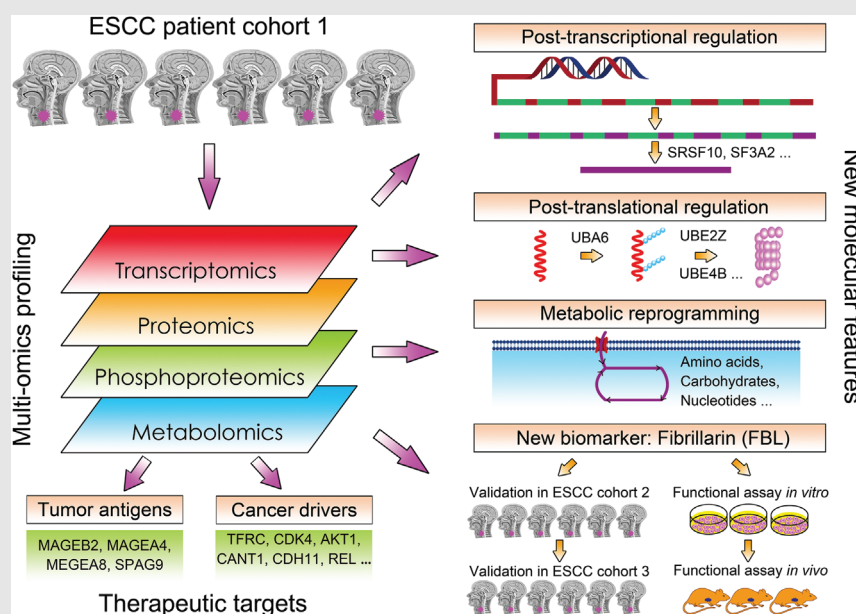
Wei Wang, Department of Thoracic Surgery, The Affiliated Tumor Hospital of Nantong University, Nantong 226361, China.

Email: wangweimz@126.com

Wen-Lian Chen, Cancer Institute, Longhua Hospital, Shanghai University of Traditional Chinese Medicine, Shanghai 200032, China.

Email: chenwl8412@shutcm.edu.cn

Graphical Abstract



1. A new molecular feature of ESCC that involves both active posttranscriptional and posttranslational regulation was unveiled.
2. ESCC-related signaling and metabolic pathways, networks among omics data, and common cancer/testis antigens along with established cancer drivers and kinases were delineated.
3. Proteins with close linkage to ESCC prognosis were discovered, and a new prognostic protein, fibrillarlin (FBL), was further validated, functionally studied, and found to correlate negatively with patient outcomes.

RESEARCH ARTICLE

A multi-omics study delineates new molecular features and therapeutic targets for esophageal squamous cell carcinoma

Xing Jin¹  | Lei Liu²  | Jia Wu¹  | Xiaoxia Jin³  | Guanzhen Yu¹  |
Lijun Jia¹ | Fengying Wang¹  | Minxin Shi²  | Haimin Lu²  | Jibin Liu²  |
Dan Liu¹  | Jing Yang¹  | Hua Li⁴ | Yan Ni⁵  | Qin Luo¹  | Wei Jia⁶  |
Wei Wang² | Wen-Lian Chen¹ 

¹ Cancer Institute, Longhua Hospital, Shanghai University of Traditional Chinese Medicine, Shanghai, China

² Department of Thoracic Surgery, The Affiliated Tumor Hospital of Nantong University, Nantong, China

³ Department of Pathology, The Affiliated Tumor Hospital of Nantong University, Nantong, China

⁴ Bio-ID Center, School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, China

⁵ The Children's Hospital, National Clinical Research Center for Child Health, Zhejiang University School of Medicine, Hangzhou, China

⁶ Hong Kong Traditional Chinese Medicine Phenome Research Center, School of Chinese Medicine, Hong Kong Baptist University, Kowloon Tong, Hong Kong, China

Correspondence

Wei Jia, Hong Kong Traditional Chinese Medicine Phenome Research Center, School of Chinese Medicine, Hong Kong Baptist University, Kowloon Tong 999077, Hong Kong, China.

Email: weijial@hkbu.edu.hk

Wei Wang, Department of Thoracic Surgery, The Affiliated Tumor Hospital of Nantong University, Nantong 226361, China.

Email: wangweimz@126.com

Wen-Lian Chen, Cancer Institute, Longhua Hospital, Shanghai University of Traditional Chinese Medicine, Shanghai 200032, China.

Email: chenwl8412@shutcm.edu.cn

Xing Jin, Lei Liu and Jia Wu contributed equally to this work.

Funding information

National Natural Science Foundation of China, Grant/Award Numbers: 31970708, 81770147, 81802891, 82002953; National Scientific and Technological Major Special

Abstract

Esophageal squamous cell carcinoma (ESCC) is a major histological subtype of esophageal cancer with inferior prognosis. Here, we conducted comprehensive transcriptomic, proteomic, phosphoproteomic, and metabolomic characterization of human, treatment-naïve ESCC and paired normal adjacent tissues (cohort 1, $n = 24$) in an effort to identify new molecular vulnerabilities for ESCC and potential therapeutic targets. Integrative analysis revealed a small group of genes that were related to the active posttranscriptional and posttranslational regulation of ESCC. By using proteomic, phosphoproteomic, and metabolomic data, networks of ESCC-related signaling and metabolic pathways that were closely linked to cancer etiology were unraveled. Notably, integrative analysis of proteomic and phosphoproteomic data pinpointed that certain pathways involved in RNA transcription, processing, and metabolism were stimulated in ESCC. Importantly, proteins with close linkage to ESCC prognosis were identified. By enrolling an ESCC patient cohort 2 ($n = 41$), three top-ranked prognostic proteins X-prolyl aminopeptidase 3 (XPNPEP3), bromodomain PHD finger transcription factor (BPTF), and fibrillarin (FBL) were verified to have increased expression in ESCC. Among these prognostic proteins, only FBL, a well-known nucleolar methyltransferase, was essential for ESCC cell growth in

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Clinical and Translational Medicine* published by John Wiley & Sons Australia, Ltd on behalf of Shanghai Institute of Clinical Bioinformatics

Project of China, Grant/Award Number: 2019ZX09201004-002-013; National Thirteenth Five-Year Science and Technology Major Special Project for New Drug Innovation and Development, Grant/Award Number: 2017ZX09304001; Research fund of Shanghai Municipal Commission of Health, Grant/Award Number: 20174Y0090; Shanghai Rising-Star Program, Grant/Award Number: 18QA1404100; Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning; Shanghai Youth Talent Program; Shanghai Municipal Key Clinical Specialty, Grant/Award Number: shslczdzk03701; Three-Year Plan of Shanghai Municipality for Further Accelerating The Development of Traditional Chinese Medicine, Grant/Award Number: ZY(2018-2020)-CCCX-1016; Shanghai Chenguang Program, Grant/Award Number: 18CG47; Gaofeng Clinical Medicine Grant of Shanghai Municipal Education Commission; Health Commission of Pudong New Area Health and Family Planning Scientific Research Project, Grant/Award Number: PW2019E-1; Xinling Scholar Program of Shanghai University of Traditional Chinese Medicine

vitro and in vivo. Furthermore, a validation study using an ESCC patient cohort 3 ($n = 100$) demonstrated that high FBL expression predicted unfavorable patient survival. Finally, common cancer/testis antigens and established cancer drivers and kinases, all of which could direct therapeutic decisions, were characterized. Collectively, our multi-omics analyses delineated new molecular features associated with ESCC pathobiology involving epigenetic, posttranscriptional, posttranslational, and metabolic characteristics, and unveiled new molecular vulnerabilities with therapeutic potential for ESCC.

KEYWORDS

esophageal squamous cell carcinoma, fibrillarin, molecular feature, multi-omics

1 | INTRODUCTION

Esophageal carcinoma (EC) is the ninth most common cancer and ranks sixth with respect to lethality in the world.¹ In China, the morbidity and mortality of EC are ranked fifth and fourth, respectively, across all cancers.² The major histological type of EC in China is esophageal squamous cell carcinoma (ESCC), accounting for approximately 90% of all EC cases.¹ To decipher the molecular aberrations that drive ESCC tumorigenesis and progression, The Cancer Genome Atlas (TCGA) and other research teams have conducted extensive genomic, epigenomic, and transcriptomic profiling, discovering unique signatures of ESCC that contain frequent genomic amplifications of tumor-promoting genes and those modulating cell cycle and apoptosis genes that have high mutational frequency.^{3–9} Notably, a predominant event in ESCC development is mutation and inactivation of a well-known tumor suppressor, TP53.^{4,9} Another typical hallmark of ESCC is mutation and/or genomic amplification of cell cycle kinases including cyclin D1 (CCND1) and cell division protein kinase 6 (CDK6).^{3,4,6} Additionally, it is

common to observe mutations of genes involved in epigenetic processes and Notch/PI3K/EGFR/Hippo pathways in ESCC.^{3,4,6}

Due to the advances in genomic study, more targeted therapies are being designed for ESCC treatment. Unfortunately, except from HER2-positive ESCC tumors, randomized controlled trials of targeted therapies for other targets, such as EGFR and mesenchymal–epithelial transition pathways, have failed.¹ The results of these trials demonstrate the complexity of ESCC oncogenesis and progression and demonstrate the limitation of genomic profiling alone for identifying effective curative treatments. Hence, a multi-omics approach may provide the necessary information required to unveil more effective molecular targets for ESCC treatment.

In this study, we performed a multi-layer omics study to characterize human, treatment-naïve ESCC tumors and paired normal adjacent tissues (NATs), aiming to delineate the mechanisms of ESCC pathobiology and unveil new therapeutic targets for precise and personalized clinical intervention.

2 | MATERIALS AND METHODS

2.1 | Nano-liquid chromatography-tandem mass spectrometry analysis

2.1.1 | Proteomic analysis using a data-independent approach

Data-dependent analysis (DDA) was performed first to generate a DDA spectral library. Fractionated and reconstituted peptides (~1 μg each fraction) of each pooled sample were resolved using a micro-tip C18 column (75 μm \times 25 cm) packed with ReproSil-Pur C18-AQ, 5 μm resin (Dr. Maisch GmbH, Germany) coupled to a nanoflow HPLC Easy-nLC 1200 system (Thermo Fisher Scientific, cat#LC140) with LC gradient rate at 250 nl/min. Two buffer solutions were used, including buffer A (a mixture of formic acid:H₂O = 1:1000 [vol/vol]) and buffer B (a mixture of formic acid:acetonitrile = 1:1000 [vol/vol]). Pooled sample resolving was achieved with the following separation gradient: 8%–30% buffer B from 0 to 97 min; 30%–100% buffer B from 97 to 100 min; 100% buffer B from 110 to 120 min. Subsequent assays were conducted on a Q-Exactive HF mass spectrometer (Thermo Fisher Scientific, cat#IQLAAEGAAPFALGMBFZ). Positive ion mode was used for detection. The MS1 full scan was set with a range of 300–1800 m/z, and with resolution of 60,000 at m/z 200, AGC target 3e6, and maximum IT 50 ms. A total of 20 MS2 scans were collected after the MS1 scan based on the inclusion list. The MS2 scans were acquired at resolution of 30,000 at m/z 200, AGC target 3e6, maximum IT 120 ms, activation type as HCD, and normalized collision energy at 27 eV.

Subsequently, for data-independent analysis (DIA), 2 μg digested peptides of each case mixed with iRT standard peptides were resolved with the same instrument and buffer solutions used for DDA. Then, the following separation gradient with a minor modification was implemented: 10%–30% buffer B from 0 to 97 min; 30%–100% buffer B from 97 to 100 min; 100% buffer B from 110 to 120 min. The same mass spectrometer was employed for DIA assay with analytic time of 2 h/sample. Each DIA cycle was composed of one full MS1 scan with a range of 350–1650 m/z (scan resolution 120,000 at 200 m/z, AGC target 3e6 and maximum IT 50 ms) and 30 MS2 scans at a DIA mode (scan resolution 30,000 at 200 m/z, AGC target 3e6, Maximum IT auto, activation type as HCD, normalized collision energy at 30 eV with spectral data type set as profile).

2.1.2 | Phosphoproteomic analysis

The iTRAQ-labeled phosphopeptides were enriched and separated with a C18 column (Thermo Scientific Acclaim PepMap100, 100 μm \times 2 cm, NanoViper) packed in a capillary column (Thermo Scientific EASY column, length 10 cm, ID 75 μm , particle size 3 μm , C18-A2) and coupled to a nanoflow HPLC Easy-nLC 1200 system (Thermo Fisher Scientific, cat#LC140) with LC gradient rate at 300 nl/min. Two buffer solutions were used, including buffer A (a mixture of formic acid:H₂O = 1:1000 [vol/vol]) and buffer B (a mixture of formic acid:acetonitrile = 1:1000 [vol/vol]). The same mass spectrometer that was used for proteomic investigation was applied for phosphoproteomic analysis. Positive ion mode was selected for measurement. The MS1 full scan was conducted with a range of 300–1800 m/z, and with resolution of 70,000 at m/z 200, AGC target 1e6, maximum IT 50 ms, and dynamic exclusion time as 60 s. A total of 20 MS2 scans were obtained after the MS1 scan at resolution of 17,500 at 200 m/z, activation type as HCD, isolation window as 2 m/z, normalized collision energy at 30 eV, and underfill at 0.1%.

2.1.3 | Database searching of proteomic data

First, we used DDA mass spectrometric data to generate a DDA spectral library. DDA data analysis was conducted with MaxQuant software (version 1.5.3.17) and human UniProt database (download in September, 2019) plus iRT peptide sequence (>Biognosys|iRT-Kit|Sequence_fusion LGGNEQVTRYILAGVENSKGTFIIDPGGVIRGTFIIDPAA VIRGAGSSEPVTGLDAKTPVISGGPYEYRVEATFGVDES NAKTPVITGAPYEYRDGLDAASYAPVRADVTPADFSE WSKLLQFGAQGSPFLK). Of note, the DDA analysis was performed with the following parameters: trypsin as the enzyme, 2 for max missed cleavages, carbamidomethyl (C) as the fixed modification, oxidation (M), and acetyl (Protein N-term) as the dynamic modification. Peptides and proteins were identified with a false discovery rate (FDR) <1%. Finally, spectral library was established with the Spectronaut software (Spectronaut Pulsar X_12.0.20491.4, Biognosys) by combining DDA raw files and the results of database searching.

DIA mass spectrometric data were analyzed with Spectronaut software (Spectronaut Pulsar X_12.0.20491.4, Biognosys) and referenced to above established DDA spectral library. The following parameters were used for analysis: “dynamic iRT” was selected for retention time prediction

type, “enabled” was chosen for interference on MS2 level correction, while “disabled” was used for cross-run normalization. FDR of peptide and protein was <1%. The raw proteomic data were deposited to The National Omics Data Encyclopedia (NODE) database (<https://www.biosino.org/node>) at Bio-Med Big Data Center (BMBDC) affiliated with Shanghai Institute of Nutrition and Health (SINH), Chinese Academy of Sciences (CAS), with a project ID of OEP002405.

2.1.4 | Database searching of phosphoproteomic data

The analysis of phosphoproteomic data was carried out with MaxQuant software (version 1.5.5.1) against the human Swiss-Prot database (version: swissprot_human_20422_20190522). The following parameters were used for analysis: “trypsin” as the enzyme, “2” for max missed cleavages, “6 ppm” for main search, “20 ppm” for first search, “20 ppm” for MS/MS tolerance, “carbamidomethyl (C) & iTRAQ8 plex (N-term) & iTRAQ8 plex (K)” as the fixed modifications, “oxidation (M) & acetyl (Protein N-term) & Phospho (STY)” as the variable modifications, “reverse” as the database pattern, and “true” for included contaminants. Peptide, protein, and site were identified with an FDR <1%. The raw phosphoproteomic data were deposited to NODE database at BMBDC affiliated with SINH, CAS, with project ID of OEP002366.

2.2 | Metabolomic analysis

Tissue and cell culture medium metabolites were extracted as previously described.^{10,11} For tissue samples, approximately 20 mg of each tissue sample was weighed and a 250- μ l pre-chilled extraction solvent mixture of chloroform, methanol, and water (vol/vol/vol = 2:5:2) was added. Samples were homogenized for 3 min and then placed in a -20°C freezer for 20 min to precipitate proteins and extract metabolites. For culture media of ESCC cell line KYSE150 (Stem Cell Bank, Chinese Academy of Sciences), 20 μ l medium of each case was collected and the metabolites were extracted with the addition of pre-chilled solvent mixture of chloroform, methanol, and water (vol/vol/vol = 2:5:2).

For above mixtures containing extracted metabolites, after centrifuging at $12,000 \times g$ and 4°C for 10 min, a volume of 150- μ l supernatant was acquired and moved to a clean sample vial. Internal standards were added into the metabolite solution and the mixture was then vacuum-dried at -20°C . The residue was derivatized by use of a two-step procedure and then analyzed based on

those previously described protocols^{12,13} using the Pegasus High-Throughput Gas Chromatography with Time-of-Flight Mass Spectrometer system (Leco Corporation). In brief, a 1- μ l derivatized sample was injected using a mode of splitless under temperature of 270°C for the injector. A flow rate of 1.0 ml/min was used for controlling the carrier gas helium. The temperature of the oven was set at 70°C for 2 min, then raised to 180°C ($10^{\circ}\text{C}/\text{min}$ as the increasing rate), and to 230°C ($6^{\circ}\text{C}/\text{min}$ as the increasing rate), finally to 295°C ($40^{\circ}\text{C}/\text{min}$ as the increasing rate). Oven temperature at 295°C was sustained for 5 min. Notably, the transferline interface and ion source were manipulated at temperatures of 270°C and 220°C , respectively. Mass spectrometer scan was implemented in a range of 50–550 m/z and the data were acquired at a rate of 20 spectra/s. The identities of metabolites were determined by searching the internal library constructed by chemical standards. The raw metabolomic data were deposited to NODE database at BMBDC affiliated with SINH, CAS, with a project ID of OEP002347.

2.3 | Proteome and phosphoproteome data analysis

2.3.1 | Missing value imputation

In the proteomic study, there were 6507 unique proteins to be identified across 48 tissue samples. Proteins that simultaneously possessed 50% missing values in ESCC and NAT tissues were excluded. For the remaining 5511 proteins, missing values were replaced by the smallest non-missing value in the data as previously reported.¹⁴

In the phosphoproteomic study, there were 3215 phosphorylated proteins along with 11,232 phosphosites to be identified. To ensure data reliability in this small sample size ($n = 6$), we extracted 2740 phosphorylated proteins along with 7186 phosphosites with no missing values for further analysis.

2.3.2 | Differential expression analysis

For the proteomic data, the nonparametric and paired two-class Wilcoxon rank-sum test with Bonferroni correction was used. FDR q -values were computed using an R package `qvalue` (v3.10) (<http://github.com/jdstorey/qvalue>). Fold change (FC) values were acquired by dividing the median value of each protein in NAT samples by the median value of corresponding protein in ESCC tumors. Differentially expressed proteins (DEPs) between ESCC tumor and NAT samples were determined by Bonferroni-adjusted $p < .05$, FDR $q < .05$, and FC cutoff as 1.5.

For the phosphoproteomic data, the paired two-class Student's *t*-test with Bonferroni correction was used. FC values were acquired by dividing the mean value of each protein in NAT samples by the mean value of corresponding protein in ESCC tumors. Differentially expressed phosphosites between ESCC tumor and NAT samples were determined by Bonferroni-adjusted $p < .05$ and FC cutoff as 1.5.

2.4 | Metabolomic data analysis

2.4.1 | Missing value imputation

Metabolites that simultaneously possessed 50% missing values in ESCC and NAT tissues were excluded. Subsequently, there were 200 unique metabolites to be identified across 48 tissue samples. Missing values were imputed using the random forest method as previously reported.¹⁵

2.4.2 | Differential expression analysis

We implemented nonparametric and paired two-class Wilcoxon rank-sum test with Bonferroni correction to identify differential metabolites. We then calculated the FDR *q*-value for each metabolite using an R package *qvalue* (v3.10) (<http://github.com/jdstorey/qvalue>). Differential metabolites between ESCC tumor and NAT samples were determined by Bonferroni-adjusted $p < .05$ and FDR $q < .05$.

2.5 | Bioinformatic analysis

2.5.1 | GO and KEGG analyses

Gene ontology (GO) and KEGG analyses were performed using ClueGO,¹⁶ a software package based on Cytoscape, for mRNAs and proteins. In ClueGO, pathway analysis was conducted using two-sided hypergeometric test.

2.5.2 | Gene set enrichment analysis

Gene set enrichment analysis (GSEA) analysis was implemented using an R package *fgsea* for all proteins in this study. FC values (ESCC samples/NAT samples) of all proteins were input for computation. The enriched pathway information was pooled from the GO, Reactome, and KEGG databases.

2.5.3 | Proteomap analysis

We executed an approach, *proteomaps*,¹⁷ to visualize the composition of proteomes with a focus on protein abundances and functions. DEPs between ESCC and NAT samples were recruited, and median values of DEPs of ESCC group and NAT group were used for construction of proteomaps.

2.5.4 | Quantification of pathway activity

All pathway scores of the enrolled samples were inferred by gene set variation analysis (GSVA) method from the GSVA R package.¹⁸ The gene sets used for computation included KEGG, Reactome, and Biocarta. The Benjamini-Hochberg method was used to adjust *p*-values of pathways between ESCC and NAT samples.

2.5.5 | Metabolite set enrichment analysis

Metabolite set enrichment analysis (MSEA) was carried out by using differentially expressed metabolites between ESCC tumors and NATs with an online tool *MetaboAnalyst 4.0*.^{19,20} Before running the analysis, data of metabolites was \log_2 -transformed. During the analysis, the metabolite set library was pathway-associated with metabolite sets (KEGG) (Oct2019).

2.5.6 | Identification of tumor antigens, potential cancer drivers and kinases

Proteomic data were used for identification of tumor antigens, potential cancer drivers and kinases. A cancer/testis (CT) antigen list was downloaded from the CTDatabase (<http://www.cta.lncc.br/modelo.php>). A potential cancer driver list was acquired from a previous study.²¹ A kinase list was obtained from the databases of PhosphoSitePlus and NetworKIN.²² By comparison to these lists, tumor antigens, potential cancer drivers and kinases were identified.

2.5.7 | Drug annotation of potential cancer drivers and key kinases

In this study, two drug databases, DrugBank²³ and PKIDB,²⁴ were used to annotate those identified potential cancer drivers and key kinases with available drugs or inhibitors.

3 | RESULTS

3.1 | Molecular landscape of ESCC by multilayer omics profiling

We enrolled 24 ESCC patients (treatment-naive cases) as cohort 1 and harvested their paired tumor and NAT samples. Each sample underwent RNA-sequencing (RNA-seq), data-independent acquisition (DIA)-based proteomic and nontargeted metabolomic investigations (Figure 1A). Due to the limitation of tissue samples, only three pairs of tumor and NAT tissues were selected for phosphoproteomic assay using mass spectrometry-based isotope tagging for relative and absolute quantification (iTRAQ). Furthermore, esophageal tissues from control mice and ESCC mice induced by the carcinogen 4-nitroquinoline-1-oxide (4-NQO) were collected for RNA-seq and nontargeted metabolomic surveys in order to verify the conservation of ESCC molecular features between distinct species (Figure 1A). Clinical characteristics of ESCC patient cohort 1 are summarized in Figure 1B.

The quality of samples and data was stringently controlled. For patients, their ESCC regions with >80% tumor cells (median [range]: 90% [80%–98%]) were harvested for analysis. The sample quality of RNAs and proteins from all specimens was verified (Figure S1, Table S1). For RNA-seq samples, the sequencing library quality was determined by Agilent 2100 Bioanalyzer, and data quality was evaluated by the Phred quality score (Table S2, Figure S2A). For proteomics analyses, the variability of quality control (QC) samples, data points for each peak, peak capability, internal calibration standards, and distribution of protein false discovery rate (FDR) were analyzed, indicating negligible instrument drift and high quality of the data (Figure S2B–F). For phosphoproteomics analyses, mass error distribution, phosphorylated peptide score distribution, and phosphorylated peptide ratio distribution were assessed (Figure S2G–I), indicating data of high quality. For metabolomics profiling, low variability of QC samples was observed, indicating stability of the measurement system (Figure S2J). For mice, the pathophysiological characteristics of their esophageal tissues were confirmed using hematoxylin and eosin (H&E) staining together with immunohistochemistry (IHC) staining for the esophageal marker keratin 14 and cell proliferation marker Ki-67 (Figure S3A). Furthermore, RNA-seq data quality was confirmed by the Phred score (Figure S3B).

To confirm whether our data were able to accurately capture the molecular features of human ESCC, we first examined the expression of well-established ESCC markers between ESCC tumors and NATs of patients using our multi-omics data. The esophageal carcinogenesis marker keratin 14 (K14, encoded by *KRT14*), growth factor recep-

tors epidermal growth factor receptor (EGFR, encoded by *EGFR*), epidermal growth factor receptor 2 (HER2, encoded by *ERBB2*), nuclear receptor cyclin D1 (encoded by *CCND1*), cell proliferation markers proliferating cell nuclear antigen (PCNA, encoded by *PCNA*), and Ki-67 (encoded by *MKI67*) were selected for analysis. As reported previously, cyclin D1 expression is a common genetic alteration as well as a key driver of ESCC.²⁵ Furthermore, K14, EGFR, cyclin D1, PCNA, and Ki-67 are upregulated both at RNA and protein levels, while HER2 is elevated only at the protein level in ESCC tissues.^{26–31} In agreement with these previous findings, our RNA-seq data demonstrated that *KRT14*, *EGFR*, *CCND1*, *PCNA*, and *MKI67*, but not *ERBB2*, were transcriptionally upregulated in ESCC tissues (Figure S4A). Additionally, our proteomic data revealed that K14, EGFR, HER2, PCNA, and Ki-67 were all increased in ESCC tissues (Figure S4B). However, cyclin D1 was not identified in our proteomic investigation. Therefore, the results of our multi-omics data identified many of the well-known molecular features of ESCC.

Prior to thoroughly analyzing the molecular features of ESCC using our multi-omics data, it was important to ascertain whether our omics data were consistent with established omics datasets. A previous study enrolled 53 ESCC patients, and collected their tumor tissues and matched NAT tissues for gene expression profiling.²⁸ This study revealed that 116 genes were dramatically upregulated, while 43 genes were strikingly downregulated in ESCC tumor tissues. In our RNA-seq data of ESCC patients, among those previously reported, upregulated, 116 genes, 76 of them (65.52%) were remarkably increased in ESCC tumor tissues (Figure S5A). In addition, 32 out of the previously 43 reported downregulated genes (74.41%) were found to be downregulated in the ESCC tumor tissues of our patient cohort (Figure S5B). This result indicated the consistency between our omics data and the published omics data.

Alternative splicing (AS) of mRNA allows for the expression of multiple RNA isoforms and contributes to the complexity of the proteome.^{32,33} Hence, we examined gene expression including their respective isoforms from the RNA-seq data. We carried out unsupervised hierarchical clustering and principal component analysis (PCA) using the multilayer omics data to ascertain human ESCC molecular features. When compared to NATs, ESCC tissues showed distinct signatures at gene, gene isoform, protein, phosphoprotein, and metabolite levels (Figures 1C and S6). As the phosphoproteomic investigation was only executed in three pairs of samples, phosphoproteomic data were excluded in the subsequent multivariate analysis. Next, integrated clustering of RNA-seq, proteomic and metabolomic data was performed using the iCluster algorithm,³⁴ and the results clearly discriminated ESCC

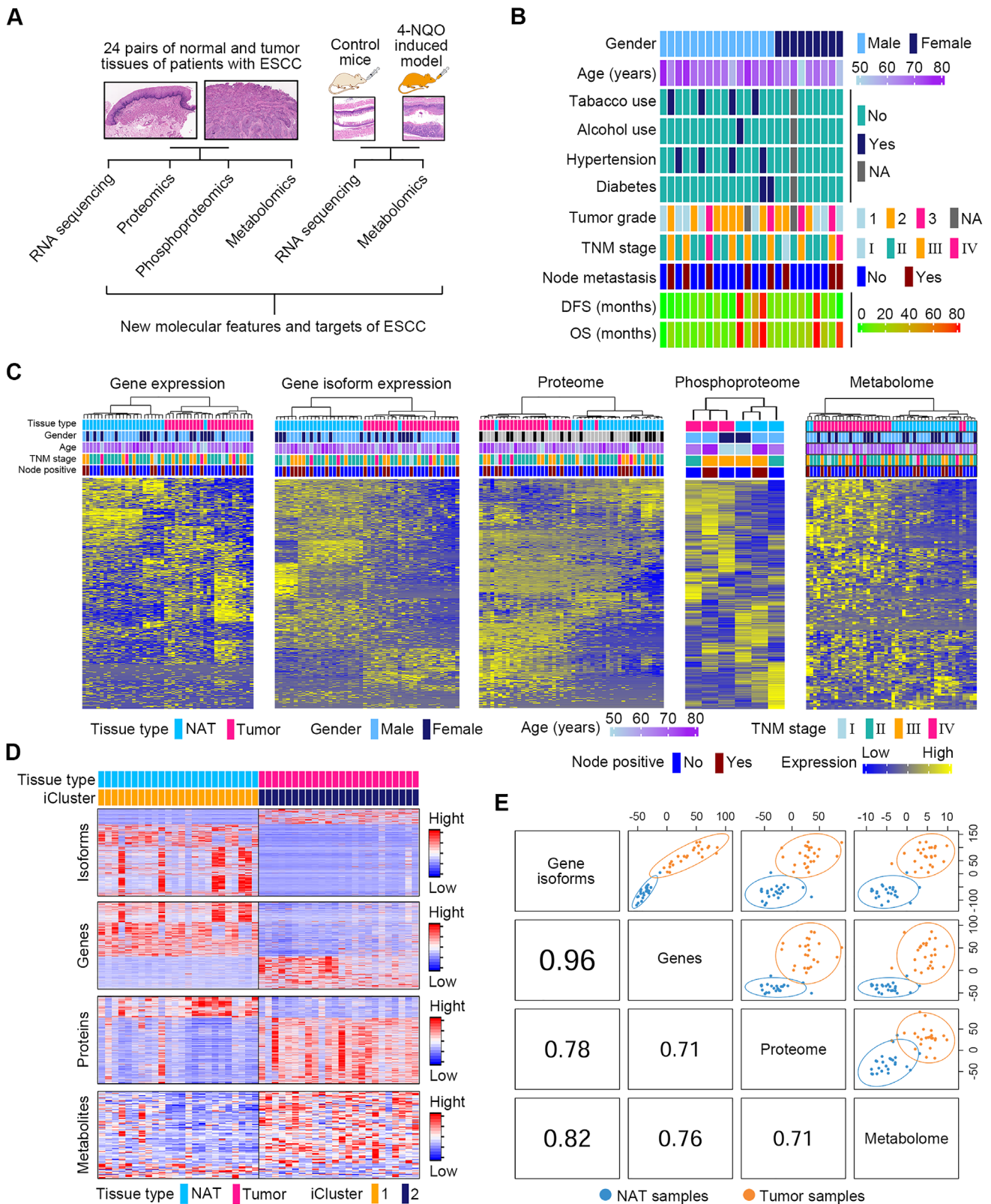


FIGURE 1 Multilayer omics profiling of ESCC and NAT tissues. (A) Strategy of multilayer omics investigations using paired ESCC and NAT samples from patient cohort 1 ($n = 24$) along with esophageal tissues from control mice and carcinogen-induced ESCC mice. (B) Heatmap showing clinical parameters of ESCC patients for multilayer omics investigations. (C) Unsupervised hierarchical clustering of multi-omics data of ESCC patients. Clinical parameters from 48 tissue samples profiled with each omics data are also depicted. (D) Integrated clustering of four molecular layers of data showed that tissue samples of patients with ESCC fell into two groups by iCluster that were virtually identical to histological classes ESCC and NAT. (E) Correlation analysis between any two omics profiles using DIABLO algorithm

from NAT samples (Figure 1D). Finally, we executed the DIABLO algorithm³⁵ to extract molecular profiles from the omics data for all patient samples, which allowed computation of the correlation between any two molecular profiles. The result revealed strong correlations between any two omics profiles (Figure 1E), demonstrating a consistent difference between ESCC and NAT samples at distinct molecular layers.

Subsequently, the RNA-seq and metabolomic data of mice were used to execute unsupervised hierarchical clustering and PCA analysis. As compared to esophageal tissues from control mice, esophageal tissues from ESCC mice exhibited distinct molecular features at gene, gene isoform, and metabolite levels (Figure S7), demonstrating that the alteration of molecular features in ESCC was conserved between different species.

3.2 | Discordance between transcriptome and proteome revealed active posttranscriptional and posttranslational regulation

To ascertain the global association between transcriptome and proteome, we computed gene-wise (inter-sample) correlation of 6174 mRNA–protein pairs across all patient samples. We observed a striking discordance between mRNAs and proteins as showed by a median gene-wise correlation value of 0.07 and a low percentage (15.69%) of mRNA–protein pairs with significant positive Spearman correlations (Figure S8A). Discordance between mRNAs and protein expression levels indicated active posttranscriptional regulation in these tissue samples.

The next step involved determination of the mRNA–protein relationship differences between the ESCC and NAT samples of patients. First, we executed gene-wise correlation analysis. Although a remarkable discrepancy between mRNAs and protein expression levels was observed in both groups of samples as shown by the low median correlation of 0.06 for ESCC tissues and 0 for NATs, the tumor tissues exhibited a right shift of distribution of correlation values relative to NATs (Kolmogorov–Smirnov test $p < 2.2 \times 10^{-16}$) (Figure 2A), indicating more positive correlations in tumor tissues. Next, we calculated sample-wise (intra-sample) correlations between mRNAs and proteins. The median correlation values for ESCC and NAT tissues were 0.20 and 0.15, respectively. Notably, ESCC tissues displayed an overt right shift of distribution of correlation values as compared to NATs (Kolmogorov–Smirnov test $p = 2.34 \times 10^{-5}$) (Figure 2B), demonstrating more positive correlations in tumor tissues. Furthermore, analysis of the FC of mRNA and protein expression between ESCC tumors and NATs showed that most of

proteins were upregulated in tumor tissues not only when corresponding mRNAs were increased, but also when corresponding mRNAs were undisturbed or reduced (Figure 2C). However, most of proteins were downregulated in NAT tissues regardless of any expression patterns of corresponding mRNAs (Figures 2C and S8B). Additionally, more upregulated mRNAs were observed in ESCC tumors relative to NATs (Figure 2C). Together, these results indicated that mRNA transcription and translation were enhanced, and posttranscriptional regulation was hyperactive in ESCC tissues. Moreover, to verify the active posttranscriptional control in ESCC, we performed validation assay using three randomly selected genes with negative correlation between their mRNAs and proteins, including *SMNDC1*, *MTHFD2*, and *PNO1*. RT-qPCR and Western blotting measurements showed that although mRNAs of these genes were not altered in ESCC tissues as relative to NATs, their protein products *SMNDC1*, *MTHFD2*, and *PNO1* were markedly elevated (Figure S8C,D), thus underscoring the high activity of posttranscriptional regulation in ESCC.

We hypothesized that those genes with increased expression at both mRNA and protein levels (namely with high transcriptional and translational activities) in ESCC tumors acted as core upstream signals that led to the observed difference in posttranscriptional modulation between ESCC and NAT tissues. To validate our hypothesis, we enrolled genes ($n = 480$, accounting for 7.77% [480/6174]) with increased mRNA abundance in tumor tissues and with positive correlation between mRNAs and corresponding proteins across all tissues for analysis. GO biological process analysis revealed that activities of RNA processing, RNA slicing, and gene expression were remarkably enhanced in ESCC tumors as evidenced by the top 10 enriched pathways (Figures 2D and S8E). Of importance, among these 480 genes, those producing proteins with extremely high levels in ESCC tissues, included serine and arginine rich splicing factor 10 (SRSF10), splicing factor 3A subunit 2 (SF3A2), cleavage stimulation factor subunit 2 (CSTF2), and replication timing regulatory factor 1 (RIF1) (Figure 2E). In our proteomic data for ESCC tumors, SRSF10, SF3A2, and CSTF2 were positively correlated to proliferation markers Ki-67 and/or PCNA (Figure 2F). In addition, SRSF10 and SF3A2 were negatively associated with differentiation marker S100A14 (Figure 2F). These results indicated that SRSF10, SF3A2, and CSTF2 may actively participate in ESCC malignancy.

We then hypothesized that the protein activities described above were involved in upstream signaling that would affect posttranscriptional control via altering the proteome. Hence, we performed GSEA using all proteins and selected gene sets involved in posttranscrip-

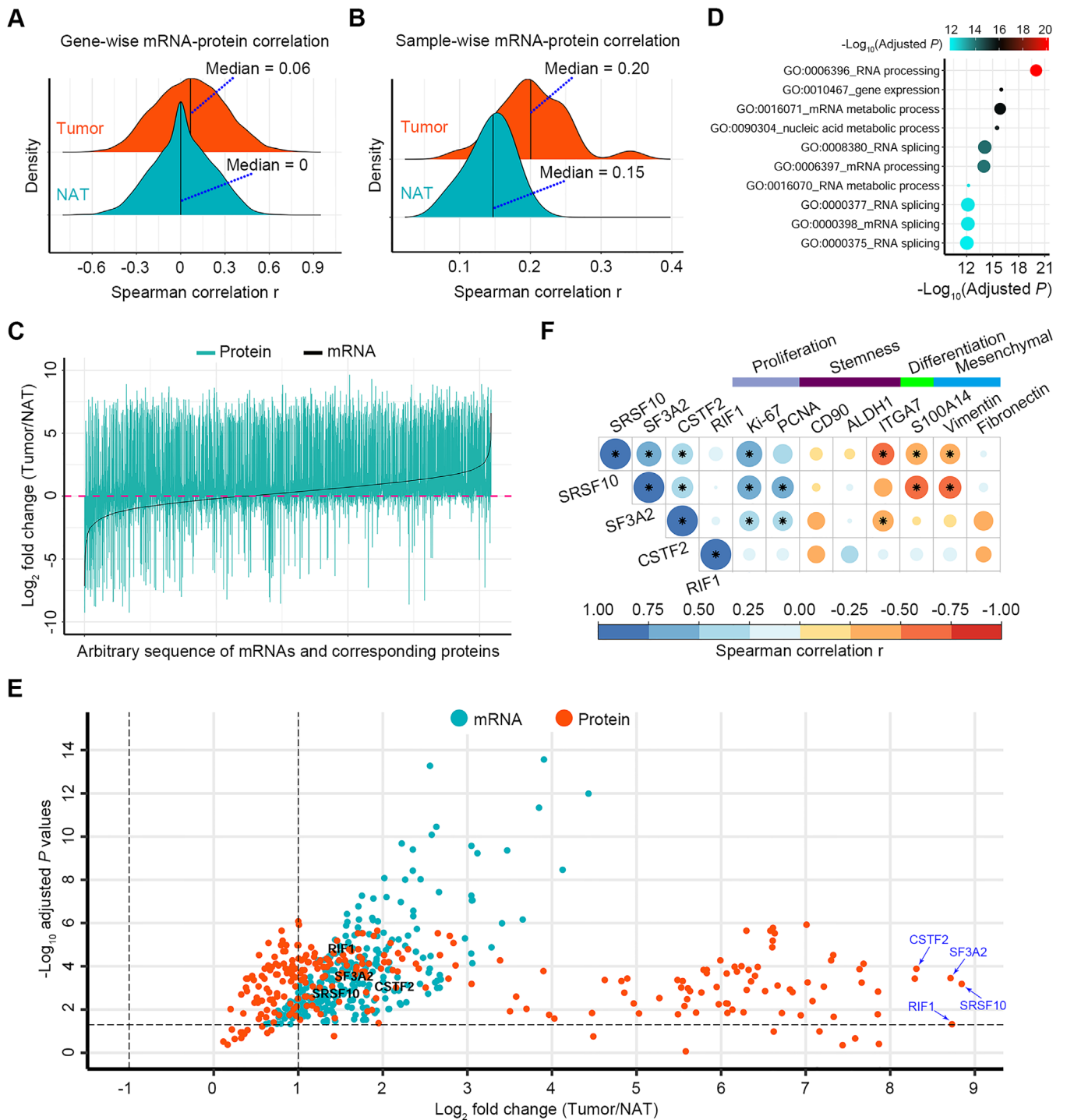


FIGURE 2 Correlations between transcriptomic and proteomic data of ESCC patients. (A and B) Density ridgeline plots showing gene-wise correlations (A) and sample-wise correlations (B) in tumors and NATs, respectively. (C) Multiple line plots displaying the fold change (FC) values of mRNAs and corresponding proteins between ESCC tumors and NATs. mRNAs were aligned by their FC values with an increasing order. (D) GO biological process analysis of 480 genes with high transcriptional and translational activities in ESCC tumors. The top 10 enriched pathways were selected for presentation. Each node size represents the percentage of measured genes in a specific pathway. (E) Volcano plot displaying the FC values (log_2 transformed) and p -values ($-\text{log}_{10}$ transformed) of 480 genes with high transcriptional and translational activities in ESCC tumors. p -Values of RNAs were derived from edgeR analysis by comparing ESCC and NAT samples, while p -values of proteins were obtained from the nonparametric and paired two-class Wilcoxon rank-sum test with Bonferroni correction by comparing ESCC and NAT samples. (F) Spearman correlation analysis between four proteins with extremely high expression in ESCC tumors and key neoplastic markers. Star marks in heatmap cells represent the correlations between proteins of interest and neoplastic markers with $p < .05$. The areas of circles show the absolute value of corresponding correlation coefficients

tional and posttranslational level control ($p < .05$). As expected, tumors were significantly more enriched than NATs in transcription and translation-related functions, including RNA processing/transcription, protein translation, protein modification, and proteolysis (Figure 3A). Then, we extracted proteins key for posttranscriptional and posttranslational regulation, including members of RNA transcription, the nonsense-mediated mRNA decay (NMD) pathway that regulates the abundance of a large number of cellular RNAs,³⁶ eukaryotic initiation factor (eIF) complex key for protein translation, and protein ubiquitin proteasome system (UPS) key for protein degradation. Among them, most of the proteins with differential expression between ESCC and NAT tissues showed increased upregulation in tumors (Figure 3B–E), consistent with the expedited rates of RNA transcription, RNA decay, protein translation, and proteolysis in ESCC tissues. Among those DEPs between ESCC and NAT samples involved in pathways listed in Figure 3A, five of them with extremely high expression in ESCC are highlighted in Figure S9, including SRSF10, U6 snRNA-associated Sm-like protein LSM6, SF3A2, mitochondrial ribosomal protein L21 (MRPL21), and ubiquitin conjugating enzyme E2 A (UBE2A).

It is known that pre-mRNA alternative splicing (AS) is a key posttranscriptional process.³⁷ To provide more evidence for the active posttranscriptional regulation in ESCC, we analyzed seven common AS events in our RNA-seq data, including alternate acceptor site (AA), alternate donor site (AD), alternate promoter (AP), alternate terminator (AT), exon skip (ES), mutually exclusive exons (ME), and retained intron (RI) using the SpliceSeq tool.³⁸ We calculated percent-splice-in (PSI) values of RNA transcripts, which reveal how efficiently these sequences are spliced into transcripts for a specific AS event.³⁹ There were 3883 RNA transcripts with statistically differential PSI values between ESCC tumors and NATs, and more AA, AD, AT, ES, and RI events were observed in tumors (Figure 3F). This was consistent with previous studies that showed increased AS events in breast cancer and colorectal cancer tissues relative to NAT tissues.^{40,41}

3.3 | Integrative illustration of molecular pathways and metabolic signatures

Due to the high discordance between mRNAs and proteins, we used proteomic data together with phosphoproteomic and metabolomic data to delineate the multilayer molecular alterations of human ESCC. First, we analyzed pathways dysregulated in ESCC using proteomic data. There were 2890 DEPs in ESCC tissues (adjusted $p < .05$,

FDR $q < .05$, FC cutoff as 1.5) (Figure S10A). We then constructed proteomaps¹⁷ to cluster the DEPs according to their KEGG pathway annotations and observed a distinct difference between ESCC tumors and NATs (Figure 4A). ESCC tumors were dominated by higher levels of spliceosome, histone, and ribosome-related proteins along with lower levels of cytoskeleton proteins. Subsequently, we performed GSVA¹⁸ using those DEPs to quantify pathway activation. A total of 157 pathways were found to be significantly perturbed in ESCC tumors (adjusted $p < .05$) (Figures 4B and S11). Predominantly activated pathways in ESCC included those related to RNA transcription/processing/metabolism, DNA synthesis and repair, protein synthesis, proteolysis, and cell cycle. Conversely, pathways related to cell junctions and interactions were strikingly attenuated in ESCC tumors. Additionally, ectopic stimulation of oncogenic signaling pathways, including MAPK, Notch, Wnt, and mTOR, was found in ESCC tumors. We then extended pathway analysis to the phosphoproteomic data. There were 517 differentially expressed phosphosites in ESCC tumors (adjusted $p < .05$, FC cutoff as 1.5) (Figure S10B), indicating increased activity of posttranslational modification. GSVA analysis using the corresponding proteins of these phosphosites revealed the activation of pathways involved in RNA transcription, processing, and metabolism in human ESCC tumors (adjusted $p < .05$) (Figure 4C).

Subsequently, we analyzed the metabolic signatures of human ESCC using metabolomic data. There were 56.50% metabolites (113/200) with differential expression in ESCC tumors (adjusted $p < .05$, FDR $q < .05$). Among these differentially expressed metabolites (DEMs), 83.19% (94/113) of them were upregulated (Figure 4D), indicating an active metabolic feature of ESCC tumors. We then performed MSEA^{19,20} using those 113 DEMs and observed a total of 17 metabolic pathways with remarkable perturbation in ESCC tumors (adjusted $p < .05$, FDR $q < .05$) (Figure 4E). Of note, 58.82% (10/17) of these pathways were involved in amino acid metabolism, indicating that amino acid metabolism was predominantly disturbed in human ESCC. In line with this finding, measurement of spent culture media of a human ESCC cell line KYSE150 manifested that ESCC cells readily imported and consumed many extracellular amino acids, including cystine, arginine, proline, and so on (Figure S12).

Both proteomic and phosphoproteomic data pinpointed an increased activity in pathways related to RNA transcription, processing, and metabolism, indicating that these pathways were crucial for ESCC pathobiology. These pathways are thoroughly analyzed in the next section. We then performed an integrated analysis for the most

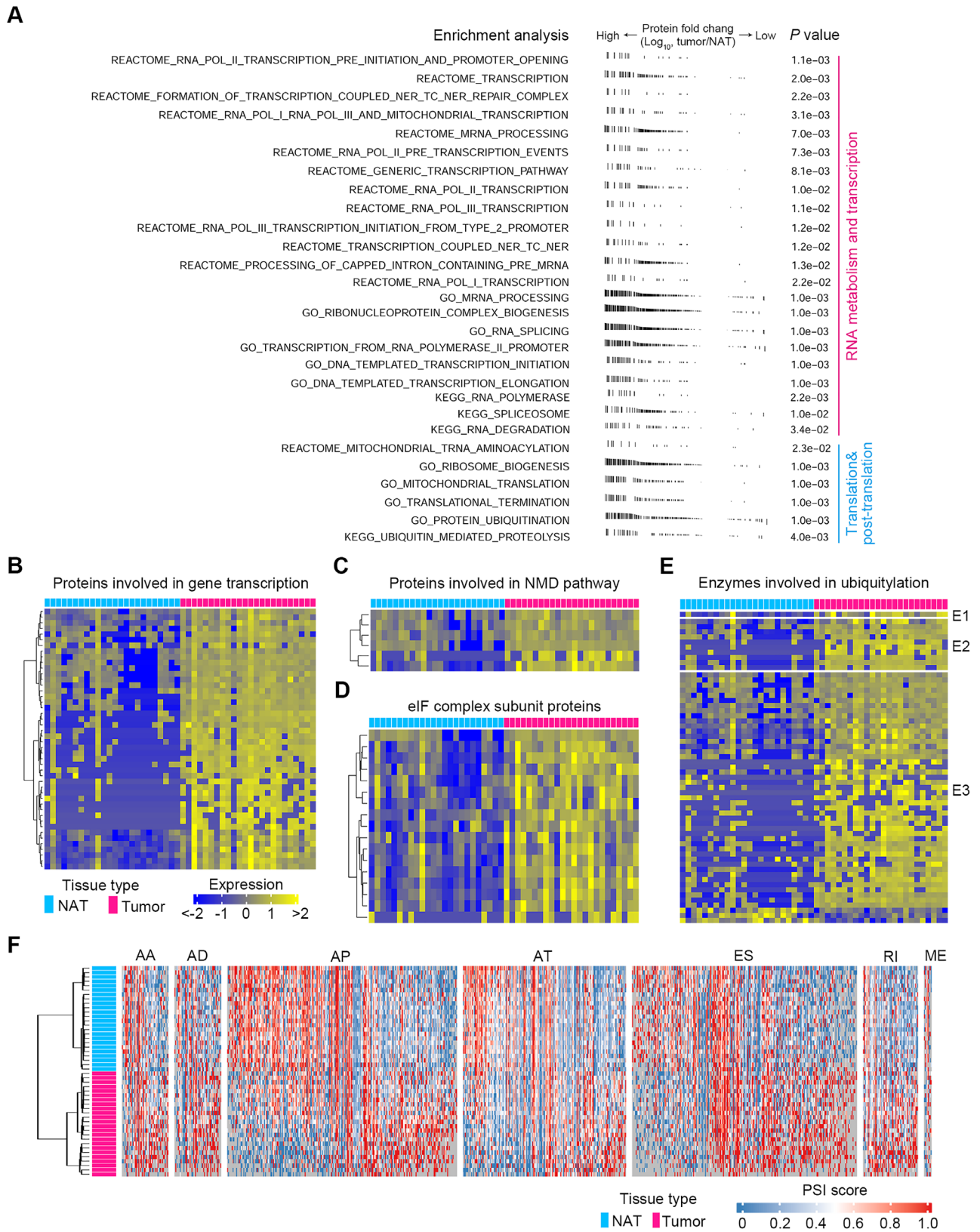


FIGURE 3 Mechanistic investigations of transcriptional, posttranscriptional, translational, and posttranslational modulations in ESCC. (A) GSEA analysis of all proteins showing active gene sets involved in transcriptional, posttranscriptional, translational, and posttranslational level control in human ESCC tissues. One vertical line in the figure represents one protein. (B–E) Differentially expressed proteins between ESCC and NAT samples involved in gene transcription (B), NMD pathway (C), eIF complex subunit (D), and ubiquitylation (E). (F) Heatmap displaying RNA transcripts with statistically differential PSI values between ESCC and NAT samples. Seven alternative splicing events were enrolled for analysis

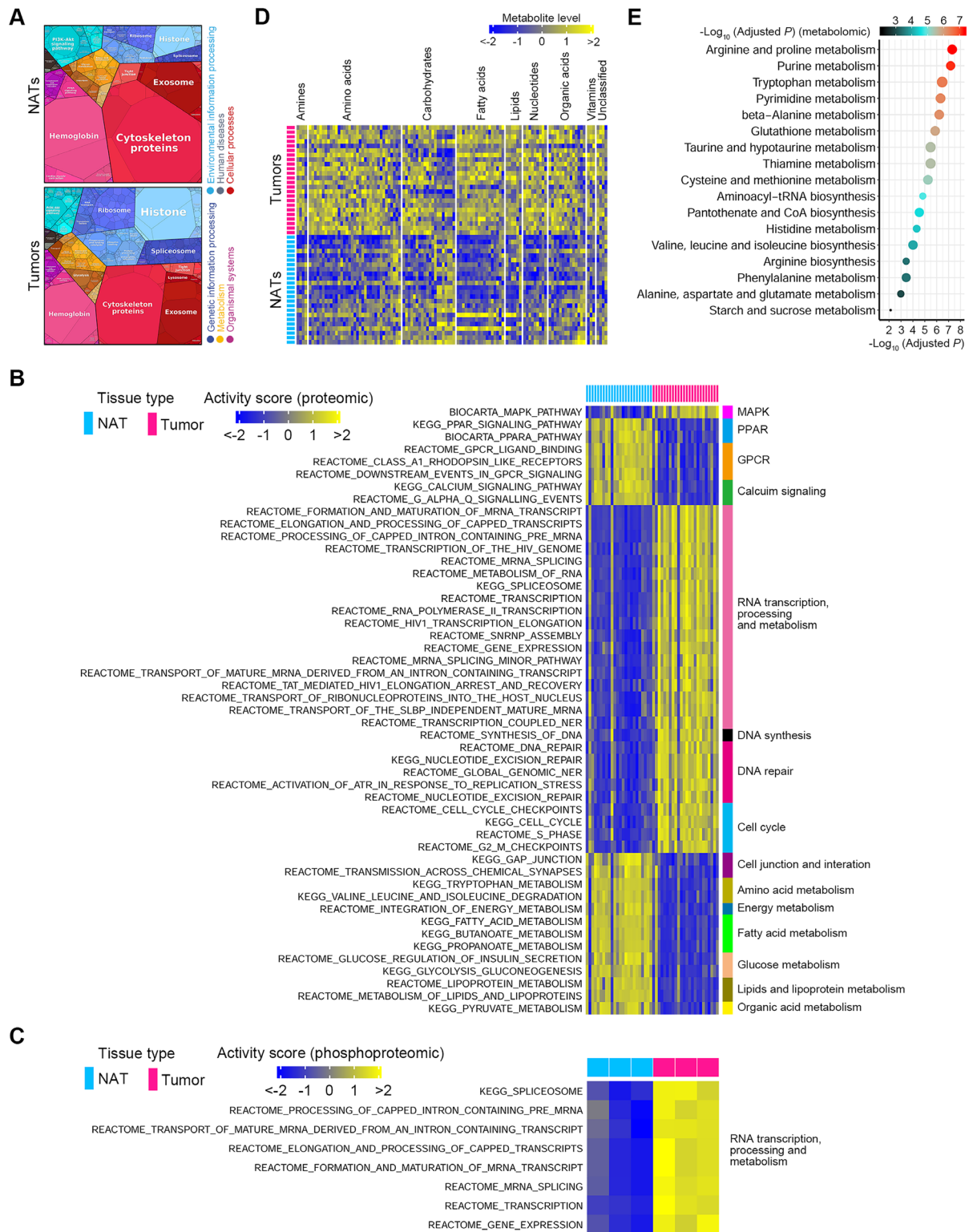


FIGURE 4 Multi-omics characterization of molecular features of ESCC. (A) Differential functional categories between NAT and ESCC specimens as illustrated by Proteomaps using proteomic data. Each polygon corresponds to a single KEGG pathway, and the size was correlated with the ratio between the two groups of samples. The maps show high dissimilarity between NAT and ESCC tissues. (B) Activity scores of the top 50 significantly perturbed pathways according to protein levels between ESCC and NAT tissues. Up- and downregulated pathways are shown as blue and yellow, respectively. (C) Activity scores of significantly disturbed pathways according to protein phosphorylation levels between ESCC and NAT tissues. Up- and downregulated pathways are indicated as blue and yellow, respectively. (D) Differentially expressed metabolites between ESCC tumors and NATs as exhibited by the heatmap. Notably, most of these metabolites were upregulated in ESCC tumors. (E) Significantly perturbed metabolic pathways in human ESCC. The node size represents the statistic q -values of metabolic pathways derived from MSEA analysis

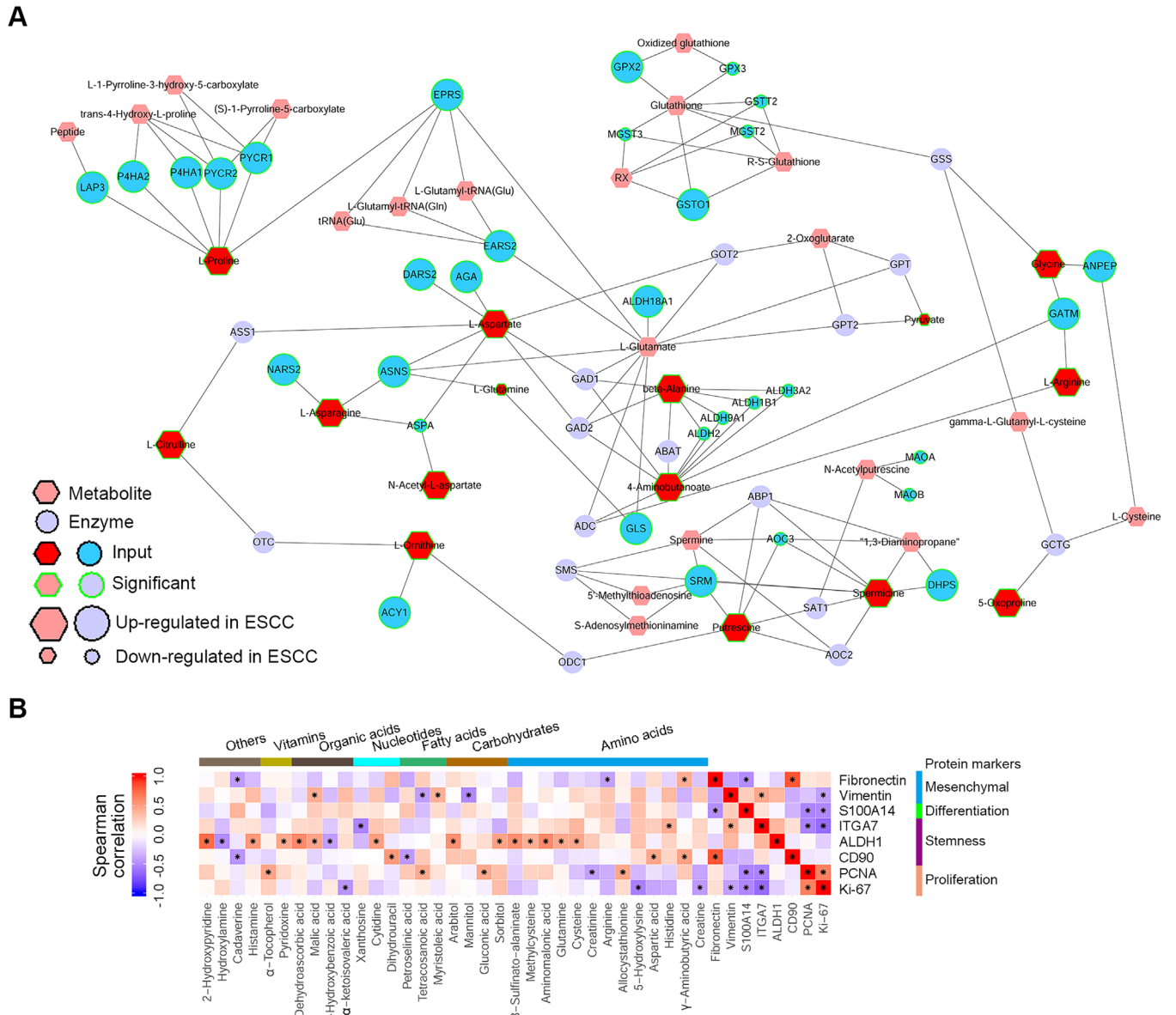


FIGURE 5 Integrative analysis of molecular signatures of ESCC using proteomic and metabolomic data. (A) Integrative analysis of significantly altered metabolic enzymes and metabolites for the most perturbed metabolic pathway in ESCC, arginine and proline metabolism. (B) Spearman correlation analysis between DEMs and DEPs representing key cellular phenotypes of ESCC. Metabolites with significant correlation to at least one phenotypic protein marker are selected for display. Star marks in heatmap cells represent correlations between metabolites and proteins with $p < .05$

perturbed metabolic pathway in ESCC, arginine and proline metabolism, using metabolomic and proteomic data. The network highlighted that ESCC tumors selectively expressed several metabolic enzymes to expedite amino acid production (Figure 5A).

In view of the importance of metabolites in cellular phenotype determination,^{42,43} we conducted correlation analysis between DEMs and DEPs representing key cellular phenotypes of ESCC (Figure 5B). Overtly, 13 metabolites showed positive linkage to aldehyde dehydro-

genase 1 (ALDH1), a well-established stemness marker of ESCC.^{44,45} These metabolites included five amino acids (cysteine, glutamine, aminomalonic acid, methylcysteine, and 3-sulfinato-alaninate), two carbohydrates (sorbitol and arabinol), one nucleotide (cytidine), two organic acids (malic acid and dehydroascorbic acid), one vitamin (pyridoxine), and two unclassified metabolites (histamine and 2-hydroxypyridine). These results suggested that altered metabolites might be involved in impacting ESCC stemness.

3.4 | Stimulated pathways of RNA transcription, processing, and metabolism unraveled by proteomic and phosphoproteomic profiling

As mentioned above, both proteomic and phosphoproteomic data pinpointed those pathways involved in RNA transcription, processing, and metabolism (Figure 4B,C), thus implicating the importance of these pathways for ESCC malignancy. Consequently, the features of these pathways were thoroughly dissected. The percentage values of remarkably upregulated proteins in ESCC tumors accounting for totally enriched and differential proteins in each pathway were 97.75% (87/89), 100.00% (99/99), 100.00% (67/67), 100.00% (77/77), 100.00% (52/52), 100.00% (62/62), and 100.00% (63/63) for spliceosome (KEGG), processing of capped intron-containing pre-mRNA (REACTOME), formation and maturation of mRNA transcript (REACTOME), mRNA splicing (REACTOME), transport of mature mRNA derived from an intron-containing transcript (REACTOME), elongation and processing of capped transcripts (REACTOME), and transcription (REACTOME), respectively (Figures 6 and S13). While the percentage values of significantly enhanced phosphosites in ESCC tumors accounting for totally enriched and differential phosphosites in each pathway were 87.50% (14/16), 96.15% (25/26), 100.00% (15/15), 95.83% (23/24), 95.00% (19/20), 93.75% (15/16), and 100.00% (17/17), respectively (Figures 6 and S13). Collectively, these findings showed that these molecular pathways were aberrantly stimulated in ESCC tumors.

3.5 | Identification of protein markers with prognostic potential

Of interest, for the DEPs in ESCC tissues of patient cohort 1, we sought to identify those proteins closely associated with patient survival and potentially involved in ESCC progression. In this small cohort ($n = 24$), higher TNM stage and lymph node metastasis, two well-known risk factors for ESCC, were associated with increased hazard of disease relapse and death with a borderline significance (Figure S14A), indicating the robustness of the prognostic data. Univariate Cox model was fitted to assess the association between each protein and patient survival. In order to enhance the recognition of potential prognostic proteins in this small cohort, we selected a borderline p -value of .08 together with a 95% confidence interval of hazard ratios (HRs) to perform statistical analysis as previously reported,⁴⁶ and a total of 118 proteins with prognostic potential were identified. Among these proteins, 66 of them displayed positive associations with hazard of disease

relapse or death (Figure 7A), while the remaining proteins exhibited negative linkage to hazard of disease relapse or death (Figure S14B). KEGG pathway analysis revealed that these 118 prognostic proteins were mainly enriched in pathways of phagosome, complement and coagulation cascades, p53 signaling pathway, and fructose and mannose metabolism ($p < .05$) (Figure 7B).

For the 66 prognostic proteins positively linked to HRs, their changes in ESCC tissues of patients at protein and corresponding mRNA levels were analyzed. Twenty-three of these proteins showed consistent expression direction with their corresponding mRNAs, indicating that these proteins were modulated at the transcriptional level (Figures 7C and S14C). While the remaining prognostic proteins exhibited inconsistent expression direction with their corresponding mRNAs, indicating that these proteins were regulated at the posttranscriptional level (Figures 7C and S14C). Of note, 14 genes, including 3-hydroxyacyl-CoA dehydratase 2 (*HACD2*), RNA binding motif protein 3 (*RBM3*), mitochondrial ribosomal protein L14 (*MRPL14*), PEST proteolytic signal containing nuclear protein (*PCNP*), X-prolyl aminopeptidase 3 (*XPNPEP3*), TBC1 domain family member 5 (*TBCID5*), bromodomain PHD finger transcription factor (*BPTF*), centrosomal protein 170B (*CEP170B*), protein tyrosine phosphatase non-receptor type 13 (*PTPN13*), N-alpha-acetyltransferase 50, NatE catalytic subunit (*NAA50*), small nuclear ribonucleoprotein polypeptide F (*SNRPF*), nucleoporin 43 (*NUP43*), fibrillarin (*FBL*), and coiled-coil-helix-coiled-coil-helix domain containing 2 (*CHCHD2*), were significantly upregulated at both protein and mRNA levels (Figure 7C). The mRNA transcription of these genes was assessed in ESCC mice induced by 4-NQO. The result showed that three of them, including *RBM3*, *BPTF*, and *FBL*, were significantly increased in esophageal tissues of ESCC mice as relative to that of control mice (Figure 7D), demonstrating that upregulation of these three genes in ESCC was conserved in distinct species. Next, we selected five of the most upregulated proteins in ESCC tumors, including *HACD2*, *RBM3*, *MRPL14*, *PCNP*, and *XPNPEP3*, together with *BPTF* and *FBL* for further investigation.

3.6 | Validation and functional assays highlighting FBL as a new unfavorable prognostic biomarker

Subsequently, we enrolled an ESCC patient cohort 2 ($n = 41$, Table S3) and performed Western blot assays to validate the expression of the above seven proteins with prognostic potential in ESCC tissues. The results showed that *FBL*, *XPNPEP3*, and *BPTF* were remarkably increased in ESCC tumors as relative to paired NATs (Figure 8A).

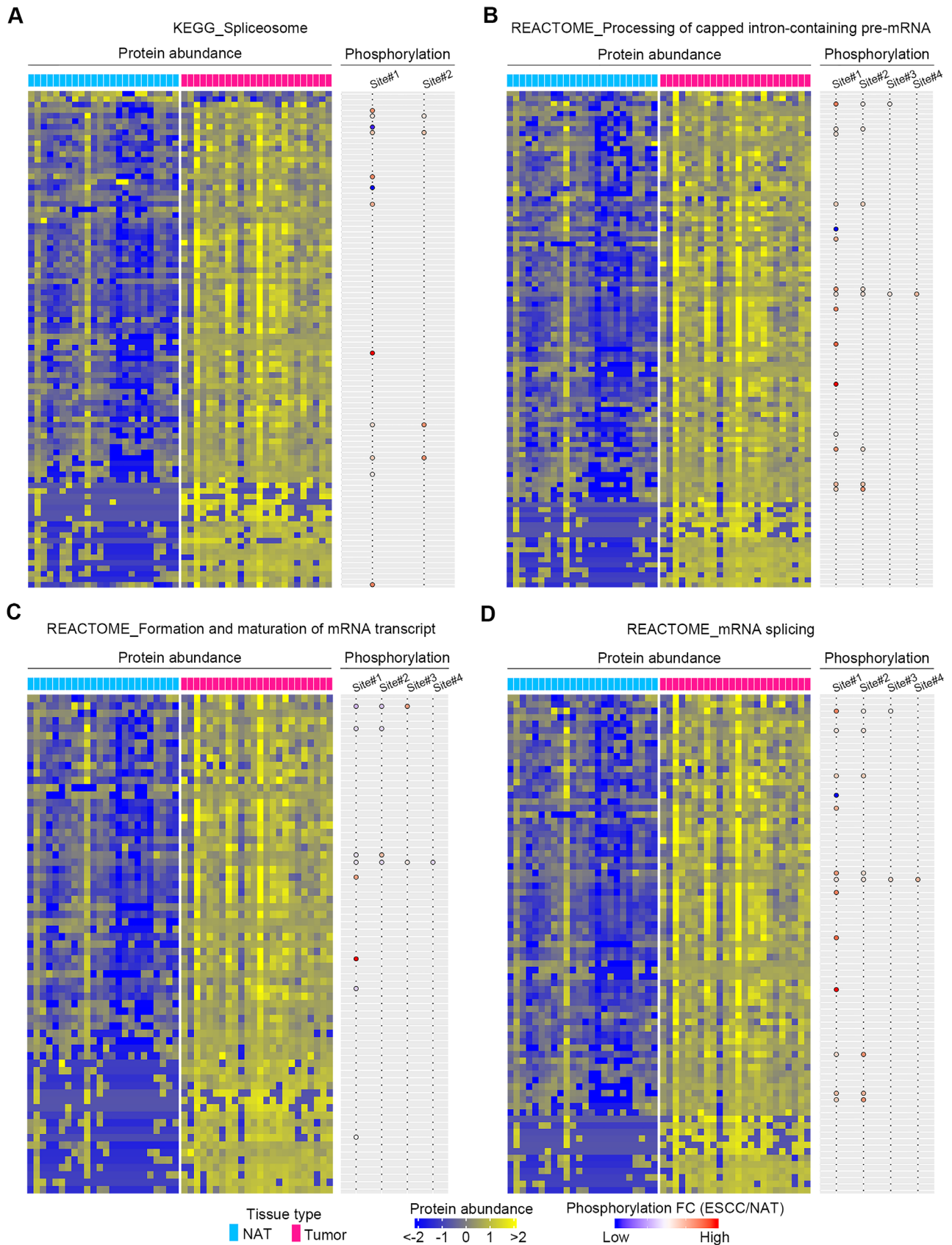
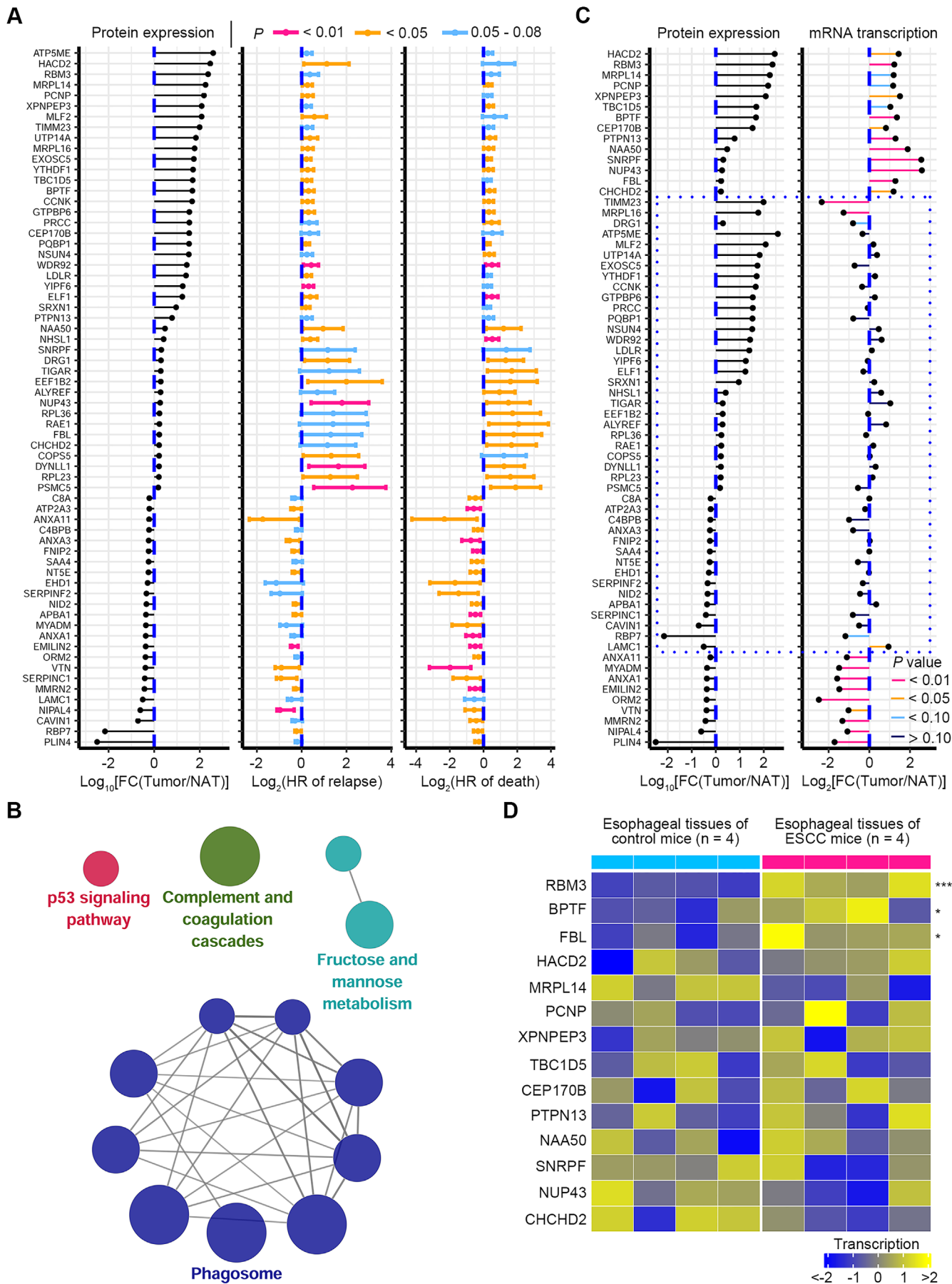


FIGURE 6 Essential pathways of ESCC identified by the integrative analysis of proteomic and phosphoproteomic data. (A–D) Heatmaps showing DEPs between ESCC tumors and NATs in the enriched pathways, including spliceosome (KEGG), processing of capped intron-containing pre-mRNA (REACTOME), formation and maturation of mRNA transcript (REACTOME), and mRNA splicing (REACTOME). Balloon plots on the right reveal differential phosphosites of proteins in the matched heatmaps on the left



By contrast, PCNP was not altered, while RBM3, MRPL14, and HACD2 were not detected in human ESCC tumors (Figure S15A). These results manifested the inconsistency between some mRNAs and their protein products, and further highlighted the active posttranscriptional modulation in ESCC. Furthermore, a previously reported gene expression dataset GSE23400,²⁸ which contained 53 pairs of ESCC tumor tissues and matched NAT tissues of patients, was enrolled for re-analysis. The result verified that *FBL* transcription was dramatically elevated in ESCC tumors (Figure S15B).

We then explored the impact of *FBL*, *BPTF*, and *XPNPEP3* on the malignancy of ESCC cells. Two human ESCC cell lines KYSE150 and Eca109 were employed to delete *FBL*, *BPTF*, and *XPNPEP3* individually by using two distinct guide RNAs for each gene. In vitro studies revealed that *FBL* abrogation dramatically downregulated the expression of a cell proliferation marker PCNA and remarkably restrained ESCC cell growth (Figure 8B). Moreover, *FBL* ablation in ESCC cells elicited G1 phase arrest and restrained the expression of cyclin D1 key for G1/S transition, whereas its abrogation did not influence cell apoptosis (Figure S15C–E). In addition, *FBL* deletion in KYSE150 cells remarkably repressed the activity of PI3K/AKT signaling, as demonstrated by the downregulation of phosphorylated AKT at Thr308 and Ser473, respectively (Figure S15F). The role of PI3K/AKT signaling in promoting G1/S transition has been previously well established.⁴⁷ Indeed, suppression of PI3K/AKT signaling by using a PI3K inhibitor LY294002 caused increased G1 phase arrest in KYSE150 cells (Figure S15G,H). In vivo studies revealed that *FBL* deletion markedly impaired the neoplastic growth of KYSE150 xenograft tumors (Figure 8C,D). However, ablation of *BPTF* or *XPNPEP3* did not influence ESCC cell propagation (Figure S15I,J). Together, these results demonstrated that *FBL*, but not *BPTF* and *XPNPEP3*, was essential for ESCC cell growth via activation of PI3K/AKT signaling and promotion of G1/S transition.

Finally, we enrolled an ESCC patient cohort 3 ($n = 100$, Table S4) to conduct an IHC staining assay study of *FBL*, and results confirmed its upregulated expression in ESCC tumors relative to NATs, as well as its usefulness as a prognostic biomarker (Figure 8E). Of importance, high *FBL* expression predicted inferior overall survival and relapse-free survival of patients with ESCC (Figure 8F). Notably, analysis of the TCGA RNA-seq data showed that high transcription of *FBL* was not associated with dismal overall survival of patients with ESCC ($n = 78$, Figure S15K), indicating no relevance between *FBL* mRNA and ESCC patient prognosis. Collectively, these findings demonstrated that high expression of *FBL* at protein level in tumor tissues was indicative of poor prognosis of ESCC patients.

3.7 | Recapitulation of key molecular events participating in ESCC development

As mentioned above, *FBL* was crucial for ESCC cell growth in vitro and in vivo. *FBL* is a nucleolar methyltransferase that mainly functions in site-specific methylation of rRNA and histone H2A, thus promoting ribosome assembly and early embryonic development.^{48,49} Hence, its increased presence in tumors indicated that epigenetic modulation and mRNA translation in ribosomes may be involved in ESCC development. Second, as an important posttranscriptional process, AS was active in ESCC tumors (Figure 3F). Moreover, pathways related to RNA transcription, processing, and metabolism were stimulated in ESCC tumors (Figure 6). Additionally, UPS, a key pathway for proteolysis, was stimulated in ESCC tumors (Figures 3A,E and S9). Together, these results indicated that posttranscriptional and posttranslational regulation participated in ESCC development. Finally, a series of metabolites and metabolic pathways were upregulated in ESCC tumors (Figure 4D,E), implying that activated metabolism was required for sustaining the malignancy of ESCC. In summary, we inferred that molecular events in epigenetic,

FIGURE 7 Discovery of protein markers with prognostic value. (A) Total of 66 DEPs with positive correlation to HRs of disease relapse and death were identified. For these prognostic proteins, their 95% confidential intervals of HRs of relapse or death are illustrated. HRs >1 correspond to an increased risk of death/relapse compared with the lower abundance of proteins, while HRs <1 correspond to a reduced risk of death/relapse compared with the lower abundance of proteins. (B) KEGG pathway analysis showing that 118 proteins with prognostic potential are enriched in pathways ($p < .05$), including phagosome, complement and coagulation cascades, p53 signaling pathway, and fructose and mannose metabolism. Each node represents a pathway, with node size reflecting pathway enrichment significance. (C) Expression direction (tumor/NAT) between 66 proteins with prognostic potential and their corresponding mRNAs. Twenty-three of these proteins that are out of the dashed quadrilateral reveal consistent expression direction with their corresponding mRNAs, whereas the remaining proteins that are highlighted by the dashed quadrilateral display inconsistent expression direction with their corresponding mRNAs. (D) For the 14 genes with increased expression at both mRNA and protein levels in human ESCC tumors shown in (C), their mRNA transcription was analyzed in esophageal tissues of ESCC and control mice. Three of these genes are significantly upregulated in esophageal tissues of ESCC mice. * $p < .05$, *** $p < .001$, two-tailed Student's *t*-test

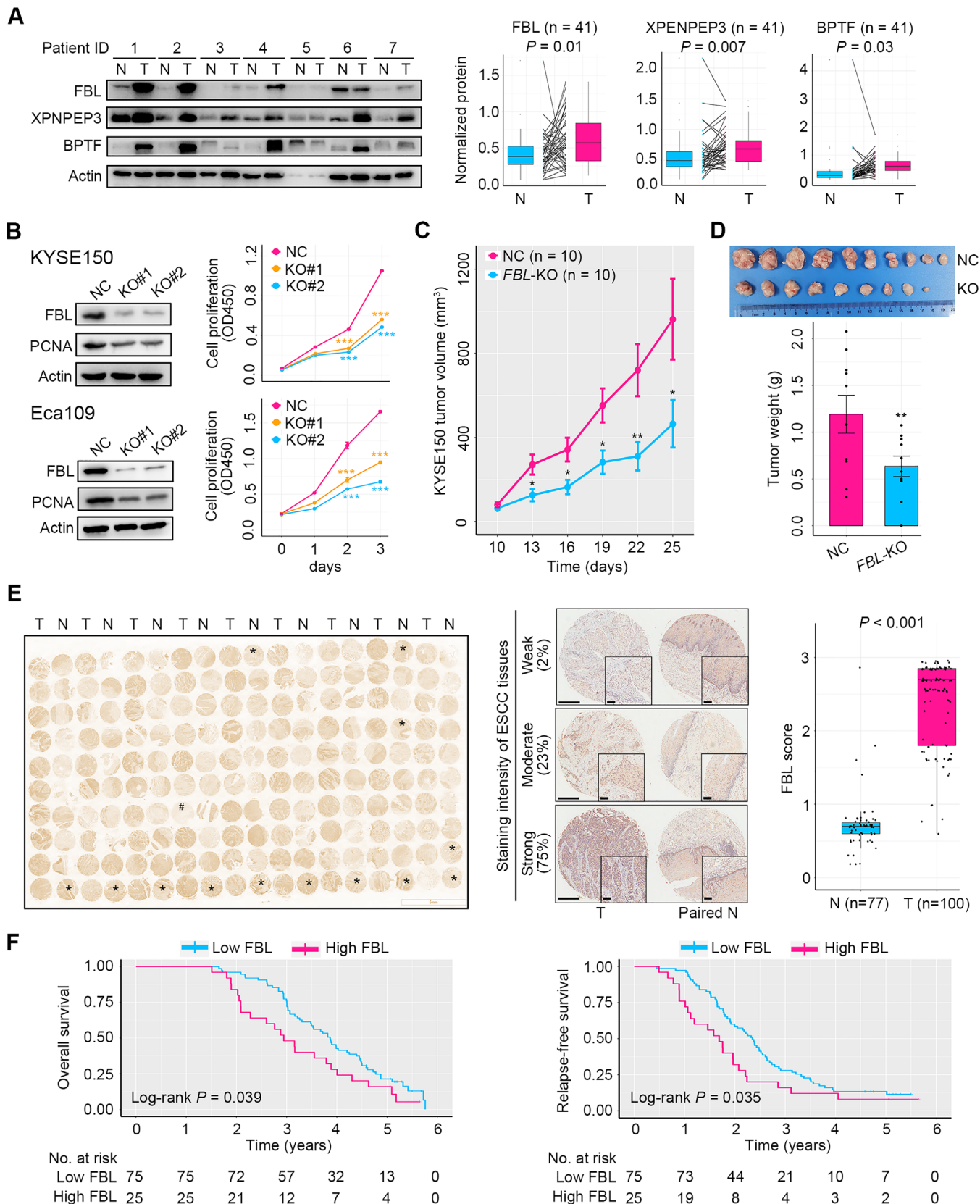


FIGURE 8 Validation and functional assays of prognostic protein markers. (A) Validation of the expression of FBL, XPENPEP3, and BPTF between ESCC tumors and NATs using a patient cohort 2 ($n = 41$). Western blot and quantitative results displayed an increase in expression of these proteins in ESCC tumors (T) relative to NATs (N). Representative Western blot images are shown. p -Values were computed using the Wilcoxon rank-sum test. (B) Knockdown of *FBL* in ESCC cells KYSE150 and Eca109 and the consequential impact on PCNA expression and cell growth in vitro. (C) *FBL* ablation curtailed the growth of KYSE150 xenograft tumors ($n = 10$ for each group).

posttranscriptional, posttranslational, and metabolic layers cooperated closely to promote ESCC development and progression (Figure 9).

3.8 | Tumor antigens and drug annotation of potential cancer drivers and key kinases

Identification of new tumor antigens including CT antigens would afford more opportunities for vaccine development for cancer immunotherapy.⁵⁰ By using the proteomic data, four known CT antigens, MAGE family member B2 (MAGEB2), MAGE family member A4 (MAGEA4), MAGE family member A8 (MAGEA8), and sperm-associated antigen 9 (SPAG9), were found to be significantly increased in ESCC tumors with FC range from 1.84 to 223.50 as compared with NATs (adjusted $p < .05$, FDR $q < .05$) (Figure 10A).

In addition, identification of altered potential cancer drivers and kinases would improve our understanding of cancer biology and give rise to new therapeutic targets.^{21,51,52} We used our proteomic data to conduct the investigation. By comparison of our data to a list of potential cancer drivers described previously,²¹ we identified 32 potential drivers with more than four-fold increase in ESCC tissues (adjusted $p < .05$, FDR $q < .05$) (Figure 10B). Furthermore, by data comparison to a kinase list from PhosphoSitePlus and NetworKIN,²² we found 29 known kinases with more than a four-fold increase in ESCC tissues (adjusted $p < .05$, FDR $q < .05$) (Figure 10C). Subsequently, two drug databases, DrugBank²³ and PKIDB,²⁴ were used for drug annotation for these potential cancer drivers and kinases elevated in ESCC. The results revealed that 21.88% (7/32) of cancer drivers and 86.21% (25/29) of kinases possessed targeted inhibitors (Figure 10B,C). These inhibitors could be tested as new therapeutics for ESCC.

4 | DISCUSSION

Previous omics studies of ESCC mainly focused on elucidating the genomic aberrations of this malignancy.³⁻⁹

Notably, by conducting a comprehensive genomic analysis of 158 ESCC cases, Song et al. identified several new mutated genes as novel oncogenes of ESCC and also found a series of new gene mutations that were potentially involved in the activity regulation of histones and several essential signaling pathways.⁵³ However, molecular perturbations in ESCC at posttranscriptional and posttranslational levels were not distinguished in these studies.

The current study provides a comprehensive characterization of the molecular systems of ESCC at the transcriptomic, proteomic, phosphoproteomic, and metabolomic levels. We demonstrated that expression patterns of genes, gene isoforms, proteins, phosphosites, and metabolites were all overtly altered in ESCC tumors relative to NATs. The conservation of modified expression patterns of genes, gene isoforms, and metabolites was verified in carcinogen-induced ESCC mice. Of importance, integrative analysis of transcriptomic and proteomic data revealed a remarkable discrepancy between mRNAs and the corresponding proteins in ESCC tumors, hence identifying high activities of ESCC tumors in posttranscriptional and posttranslational processes, including RNA transcription, RNA AS, RNA decay, protein translation, and proteolysis. It is reported that posttranscriptional processes, such as RNA AS, RNA stability, and RNA decay, play a vital role in tumorigenesis and tumor progression.^{37,54} For example, increased mRNA stability of *SEMA4D* regulated by HuR promotes cell proliferation and migration of ESCC cells.⁵⁵ In the current study, we found that a panel of 480 genes with high expression both at RNA and protein levels, together with a series of differential proteins discovered by proteomic profiling, were potentially involved in modulating posttranscriptional and posttranslational processes of ESCC. Notably, among these proteins, those with extremely high expression in ESCC tumors included SRSF10, SF3A2, CSTF2, RIF1, LSM6, MRPL21, and UBE2A. SRSF10, a well-known splicing factor, mediates AS of interleukin 1 receptor accessory protein (IL1RAP) and promotes tumorigenesis in the cervix.⁵⁶ SF3A2 and CSTF2 are RNA-binding proteins.²¹ SF3A2 stimulates inclusion of exon 14 of the histone methyltransferase enhancer of zeste 2 polycomb repressive complex 2 subunit (EZH2) and has pro-proliferative activity in renal cancer.⁵⁷ CSTF2 induces 3'-UTR shortening of Rac family small GTPase 1 (RAC1) to exacerbate

(D) Images and weight of KYSE150 xenograft tumors with or without *FBL* deletion ($n = 10$ for each group). (E) Left panel, IHC staining image showing FBL expression between ESCC (T) and NAT (N) tissues from patient cohort 3 ($n = 100$). Samples in the microtissue array that are not orderly arranged are marked with stars for ESCC tissues or hashes for NATs. Middle panel, representative images of ESCC tissues with weak, moderate, or strong FBL staining. The percentage of tumor tissues of each staining intensity is revealed. Paired normal tissues are displayed. Right panel, quantification of IHC staining. p -Value was obtained from the Wilcoxon rank-sum test. (F) Overall survival and relapse-free survival curves of ESCC patients of cohort 3 ($n = 100$) stratified by low and high FBL expressions from IHC staining. Error bars represent mean \pm SEM. * $p < .05$, ** $p < .01$, *** $p < .001$, two-tailed Student's t -test

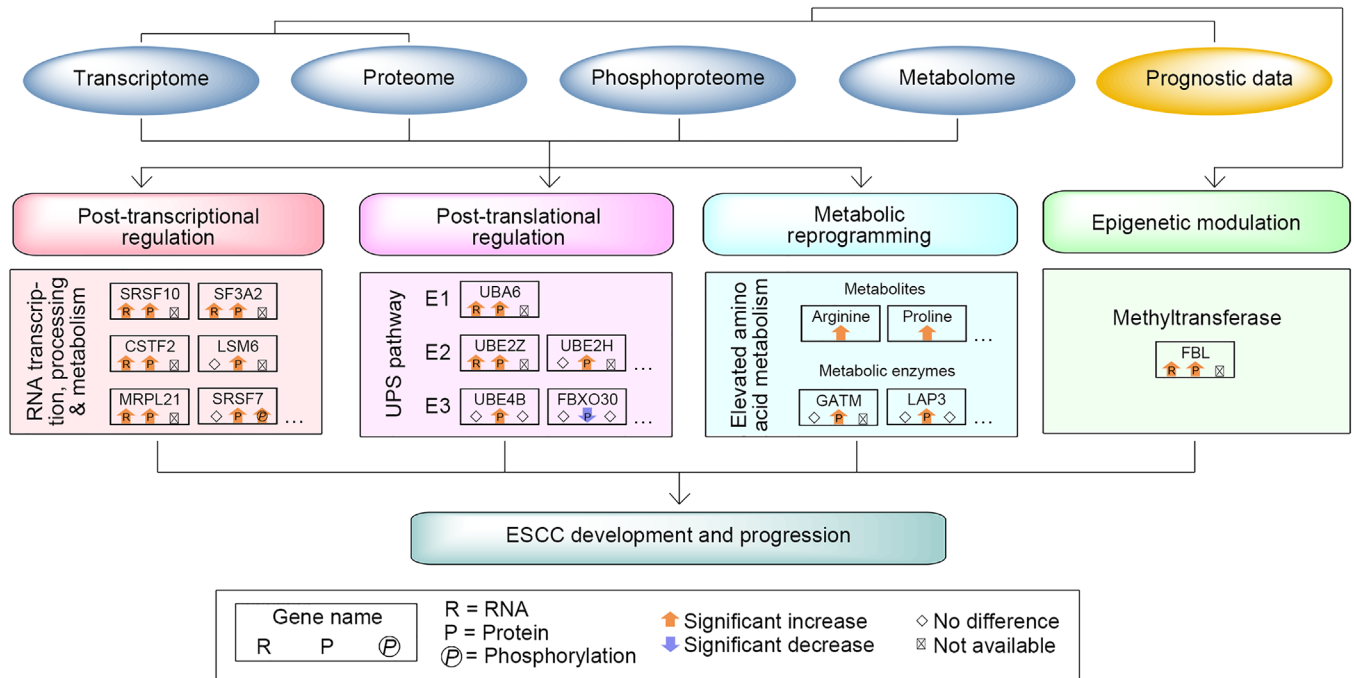


FIGURE 9 A model depicting molecular events during ESCC development identified by the multilayer study. Altered molecular events of epigenetic modulation, postranscriptional and postranslational regulation, and metabolic reprogramming, which were potentially involved in ESCC development and progression, are depicted

cellular malignancy in urothelial carcinoma.⁵⁸ RIF1 modulates replication timing regulation and activates expression of human telomerase reverse transcriptase to expedite epithelial ovarian cancer growth.⁵⁹ LSM6 and MRPL21 are RNA-binding proteins.²¹ LSM6 participates in reducing E-cadherin expression, thus promoting cell migration in breast cancer.⁶⁰ MRPL21 is highly expressed in several cancers and could be used as a biomarker for cancer prediction.⁶¹ UBE2A, an E2 ubiquitin-conjugating enzyme, involved in DNA damage repair by catalyzing the ubiquitination of different target proteins, promotes cell cycle progression and tumorigenesis.⁶² Due to the well-established tumorigenic roles of these proteins, it is reasonable to hypothesize that active postranscriptional and postranslational regulation is potential oncogenic driver of ESCC.

To the best of our knowledge, the present study is the first to integrate proteomic, phosphoproteomic, and metabolomic data to thoroughly portray the perturbations of signaling and metabolic pathways in ESCC. Analyses of proteomic and phosphoproteomic data revealed that pathways involved in RNA metabolism, transcription, and translation were enhanced at both protein and phosphorylation levels, thus underscoring postranscriptional processes as possible etiological determinants of ESCC. There is increasing, consistent evidence to corroborate the importance of this biological process in ESCC development and progression.^{63,64} Future investigations are necessary to

complete our understanding of the biological mechanisms that determine ESCC malignancy. Additionally, our proteomic and metabolomic data revealed an enhanced amino acid metabolism in ESCC tumors, indicating an addiction of ESCC cells to amino acids. Indeed, previous studies have found that a series of amino acids were remarkably upregulated in ESCC tumors.^{65,66} Intriguingly, we observed that several metabolites were positively linked to an ESCC stemness marker ALDH1, indicating a close relationship between metabolism and ESCC stemness. Further work is required to ascertain the biological functions and translational potential of these abnormal metabolic pathways and metabolites in ESCC.

An essential finding of this study is the identification of crucial proteins with prognostic potential for ESCC. For 66 of them, which are negatively linked to patient survival, their prognostic value is worthy of a validation in future independent cohorts. It should be particularly noted that this study yields a newly identified ESCC prognostic marker, FBL, a nucleolar methyltransferase that mainly functions in site-specific methylation of rRNA and histone H2A, ribosome assembly, and early embryonic development.^{48,49,67,68} FBL is involved in rDNA synthesis during the interphase of the cell cycle, and is required for normal nuclear morphology and cancer cell growth.^{49,68} To the best of our knowledge, there are no published studies reporting the role of FBL in ESCC. Here, we found that FBL is highly expressed in ESCC tissues, negatively associ-

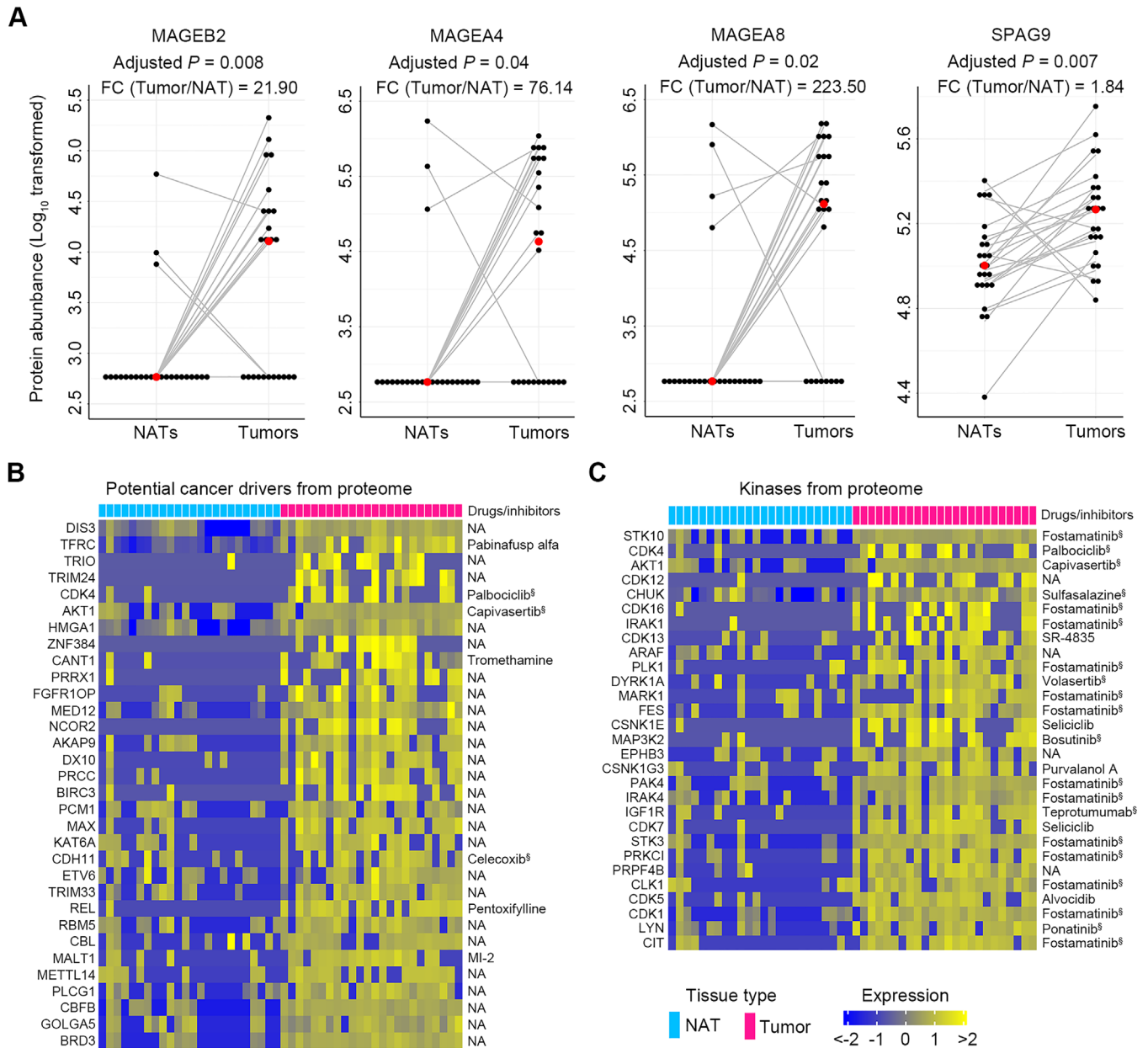


FIGURE 10 Identification of tumor antigens and drug annotation of potential cancer drivers and key kinases in ESCC. (A) Four tumor antigens with significant increases in ESCC tumors were identified by comparison of our proteomic data to a cancer/testis antigen list downloaded from the CTDdatabase. Red dots represent the median values of each group. p -Values were computed using the Wilcoxon rank-sum test with Bonferroni correction. (B) A total of 32 potential cancer drivers with more than four-fold increase in ESCC tissues were identified by comparison of our proteomic data to a list of potential cancer drivers provided by a previous study. (C) A total of 29 kinases with more than four-fold increase in ESCC tissues were discovered by comparison of our proteomic data to a kinase list from databases, PhosphoSitePlus and NetworKIN. Identified cancer drivers and key kinases were annotated with targeted drugs or inhibitors by searching the databases, DrugBank and PKIDB. [§]Drugs or inhibitors approved by Food and Drug Administration for clinical trial. NA, not available

ated with patient prognosis, and vital for ESCC cell growth via stimulation of PI3K/AKT signaling and promotion of cell cycle progression, indicating that FBL is a potential therapeutic target against ESCC. The underlying molecular mechanism of how FBL regulates PI3K/AKT signaling requires further investigation.

Finally, therapeutic indications from this study should be evaluated. First, molecular pathways involved in post-transcriptional and posttranslational regulation are potential new therapeutic targets for ESCC treatment. For example, seven of these pathway proteins identified by GO and GSEA analyses, including SRSF10, SF3A2, CSTF2,

RIF1, LSM6, MRPL21, and UBE2A, are extremely highly expressed in ESCC and have well-established tumor-promoting roles, thus supporting them as latent targets for curative treatment of ESCC. In addition, UPS, a key pathway involved in posttranslational regulation, is an appealing target for cancer therapy.⁶⁹ Our proteomic data reveals that many UPS enzymes are dramatically upregulated in ESCC tumors, implying that inhibitors of UPS may be useful against ESCC. In line with this reasoning, a previous study reported that a well-known proteasome inhibitor, bortezomib was able to induce cell cycle arrest and cell apoptosis, thus potentiating cytotoxicity of radiation therapy for ESCC.⁷⁰ Second, those prognosis-associated pathways enriched by KEGG analysis could be exploited as therapeutic targets for ESCC. For instance, one of the inhibited pathways, the p53 signaling pathway, drives the oncogenesis of ESCC as previously reported.²⁵ Third, due to four CT antigens with an overt increase in ESCC tumors, it may be possible to develop vaccines to specifically eliminate ESCC cells. Finally, based on the drug annotation for potential cancer drivers and key kinases identified by our proteomic data, we can test the efficacy of these inhibitors in preclinical ESCC models as well as clinical trials in the future and develop new attractive therapeutics for the treatment of ESCC.

5 | CONCLUSIONS

By using a multi-omics approach, we deciphered new molecular events involved in ESCC development and progression, including aberrant methyltransferase expression, hyperactive posttranscriptional and posttranslational regulation, and reshaped metabolism. These findings have deepened our understanding of ESCC pathobiology and provided new prognostic biomarkers for risk stratification of ESCC patients. Furthermore, these findings unveiled new therapeutic targets and strategies for ESCC.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (31970708, 81770147, 81802891, 82002953), National Scientific and Technological Major Special Project of China (2019ZX09201004-002-013), National Thirteenth Five-Year Science and Technology Major Special Project for New Drug Innovation and Development (2017ZX09304001), Research fund of Shanghai Municipal Commission of Health (20174Y0090), Shanghai Rising-Star Program (18QA1404100), Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning, Shanghai Youth Talent Program, Shanghai Municipal Key Clinical

Specialty (shslczdzk03701), Three-Year Plan of Shanghai Municipality for Further Accelerating The Development of Traditional Chinese Medicine (ZY(2018-2020)-CCCX-1016), Shanghai Chenguang Program (18CG47), Gaofeng Clinical Medicine Grant of Shanghai Municipal Education Commission, Health Commission of Pudong New Area Health and Family Planning Scientific Research Project (PW2019E-1), and Xinling Scholar Program of Shanghai University of Traditional Chinese Medicine.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

ETHICS STATEMENT

All participants provided informed written consent in accordance with the regulation of the Institutional Review Board of the Affiliated Tumor Hospital of Nantong University (Approval number 2019-022) in agreement with the Declaration of Helsinki. Studies of carcinogen-induced ESCC mice were approved and conducted by Wuhan Servicebio Technology Co., Ltd (Approval number 2018015). Subcutaneous tumor xenograft experiments were approved by the Institutional Animal Care and Use Committee of Longhua Hospital (Approval number LHERAW-19038).

AUTHORS CONTRIBUTIONS

Conceptualization: Wen-Lian Chen, Wei Wang, and Wei Jia. Multilayer omics study, data mining, and bioinformatics analysis: Wen-Lian Chen, Hua Li, and Yan Ni. Pathology, H&E, and IHC staining assays: Xiaoxia Jin and Guanzhen Yu. Molecular biology and in vitro and in vivo studies: Xin Jing, Jia Wu, Dan Liu, Jing Yang, and Qin Luo. Patient enrollment, tissue samples, and clinical data: Lei Liu, Jia Wu, Fengying Wang, Minxin Shi, Haimin Liu, Jibin Liu, and Wei Wang. Manuscript writing: Wen-Lian Chen. Writing - review and editing: Lijun Jia, Hua Li, and Wei Jia.

DATA AVAILABILITY STATEMENT












Our raw data have been deposited to the National Omics Data Encyclopedia (NODE) database at Bio-Med Big Data Center affiliated with Shanghai Institute of Nutrition and Health, Chinese Academy of Sciences, with project IDs of OEP002359 for RNA-seq dataset, OEP002405 for proteomic dataset, OEP002366 for phosphoproteomic dataset, and OEP002347 for metabolomic dataset. The web link for NODE database is <https://www.biosino.org/node>.

ORCID

Xing Jin  <https://orcid.org/0000-0003-2942-1469>

Lei Liu  <https://orcid.org/0000-0003-0256-0944>

Jia Wu  <https://orcid.org/0000-0002-2697-9160>

Xiaoxia Jin  <https://orcid.org/0000-0002-2388-5281>
 Guanzhen Yu  <https://orcid.org/0000-0001-6193-8148>
 Fengying Wang  <https://orcid.org/0000-0002-2833-2379>
 Minxin Shi  <https://orcid.org/0000-0001-6521-1261>
 Haimin Lu  <https://orcid.org/0000-0003-0877-9824>
 Jibin Liu  <https://orcid.org/0000-0002-1603-6468>
 Dan Liu  <https://orcid.org/0000-0002-7069-9940>
 Jing Yang  <https://orcid.org/0000-0001-9764-5136>
 Yan Ni  <https://orcid.org/0000-0003-1779-7266>
 Qin Luo  <https://orcid.org/0000-0003-0255-8430>
 Wei Jia  <https://orcid.org/0000-0002-3739-8994>
 Wen-Lian Chen  <https://orcid.org/0000-0002-1359-2549>

REFERENCES

- Lagergren J, Smyth E, Cunningham D, Lagergren P. Oesophageal cancer. *Lancet*. 2017;390(10110):2383-2396.
- Chen W, Zheng R, Zhang S, et al. Cancer incidence and mortality in China in 2013: an analysis based on urbanization level. *Chin J Cancer Res*. 2017;29(1):1-10.
- Cancer Genome Atlas Research N, Analysis Working Group, Asan University, et al. Integrated genomic characterization of oesophageal carcinoma. *Nature*. 2017;541(7636):169-175.
- Gao YB, Chen ZL, Li JG, et al. Genetic landscape of esophageal squamous cell carcinoma. *Nat Genet*. 2014;46(10):1097-1102.
- Zhang L, Zhou Y, Cheng C, et al. Genomic analyses reveal mutational signatures and frequently altered genes in esophageal squamous cell carcinoma. *Am J Hum Genet*. 2015;96(4):597-611.
- Sawada G, Niida A, Uchi R, et al. Genomic landscape of esophageal squamous cell carcinoma in a Japanese population. *Gastroenterology*. 2016;150(5):1171-1182.
- Deng J, Chen H, Zhou D, et al. Comparative genomic analysis of esophageal squamous cell carcinoma between Asian and Caucasian patient populations. *Nat Commun*. 2017;8(1):1533.
- Chang J, Tan W, Ling Z, et al. Genomic analysis of oesophageal squamous-cell carcinoma identifies alcohol drinking-related mutation signature and genomic alterations. *Nat Commun*. 2017;8:15290.
- Chen XX, Zhong Q, Liu Y, et al. Genomic comparison of esophageal squamous cell carcinoma and its precursor lesions by multi-region whole-exome sequencing. *Nat Commun*. 2017;8(1):524.
- Chen WL, Jin X, Wang M, et al. GLUT5-mediated fructose utilization drives lung cancer growth by stimulating fatty acid synthesis and AMPK/mTORC1 signaling. *JCI Insight*. 2020;5(3):e131596.
- Qiu Y, Cai G, Zhou B, et al. A distinct metabolic signature of human colorectal cancer with prognostic potential. *Clin Cancer Res*. 2014;20(8):2136-2146.
- Chen WL, Wang JH, Zhao AH, et al. A distinct glucose metabolism signature of acute myeloid leukemia with prognostic value. *Blood*. 2014;124(10):1645-1654.
- Wang JH, Chen WL, Li JM, et al. Prognostic significance of 2-hydroxyglutarate levels in acute myeloid leukemia in China. *Proc Natl Acad Sci U S A*. 2013;110(42):17017-17022.
- Gatto L, Lilley KS. MSnbase-an R/Bioconductor package for isobaric tagged mass spectrometry data visualization, processing and quantitation. *Bioinformatics*. 2012;28(2):288-289.
- Wei R, Wang J, Su M, et al. Missing value imputation approach for mass spectrometry-based metabolomics data. *Sci Rep*. 2018;8(1):663.
- Bindea G, Mlecnik B, Hackl H, et al. ClueGO: a Cytoscape plugin to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics*. 2009;25(8):1091-1093.
- Liebermeister W, Noor E, Flamholz A, et al. Visual account of protein investment in cellular functions. *Proc Natl Acad Sci U S A*. 2014;111(23):8488-8493.
- Hanzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics*. 2013;14:7.
- Xia J, Wishart DS. MSEA: a web-based tool to identify biologically meaningful patterns in quantitative metabolomic data. *Nucleic Acids Res*. 2010;38(Web Server issue):W71-W77.
- Chong J, Soufan O, Li C, et al. MetaboAnalyst 4.0: towards more transparent and integrative metabolomics analysis. *Nucleic Acids Res*. 2018;46(W1):W486-W494.
- Sebestyen E, Singh B, Minana B, et al. Large-scale analysis of genome and transcriptome alterations in multiple tumors unveils novel cancer-relevant splicing networks. *Genome Res*. 2016;26(6):732-744.
- Wiredja DD, Koyuturk M, Chance MR. The KSEA app: a web-based tool for kinase activity inference from quantitative phosphoproteomics. *Bioinformatics*. 2017;33(21):3489-3491.
- Law V, Knox C, Djoumbou Y, et al. DrugBank 4.0: shedding new light on drug metabolism. *Nucleic Acids Res*. 2014;42(Database issue):D1091-D1097.
- Carles F, Bourg S, Meyer C, Bonnet P. PKIDB: a curated, annotated and updated database of protein kinase inhibitors in clinical trials. *Molecules*. 2018;23(4):908.
- Opitz OG, Harada H, Suliman Y, et al. A mouse model of human oral-esophageal cancer. *J Clin Invest*. 2002;110(6):761-769.
- Tang XH, Knudsen B, Bemis D, Tickoo S, Gudas LJ. Oral cavity and esophageal carcinogenesis modeled in carcinogen-treated mice. *Clin Cancer Res*. 2004;10(1 Pt 1):301-313.
- Lam AK. Molecular biology of esophageal squamous cell carcinoma. *Crit Rev Oncol Hematol*. 2000;33(2):71-90.
- Su H, Hu N, Yang HH, et al. Global gene expression profiling and validation in esophageal squamous cell carcinoma and its association with clinical phenotypes. *Clin Cancer Res*. 2011;17(9):2955-2966.
- Tan C, Qian X, Guan Z, et al. Potential biomarkers for esophageal cancer. *Springerplus*. 2016;5:467.
- Okuno Y, Nishimura Y, Kashi I, Ono K, Hiraoka M. Prognostic values of proliferating cell nuclear antigen (PCNA) and Ki-67 for radiotherapy of oesophageal squamous cell carcinomas. *Br J Cancer*. 1999;80(3-4):387-395.
- Kimos MC, Wang S, Borkowski A, et al. Esophagin and proliferating cell nuclear antigen (PCNA) are biomarkers of human esophageal neoplastic progression. *Int J Cancer*. 2004;111(3):415-417.
- He C, Zhou F, Zuo Z, Cheng H, Zhou R. A global view of cancer-specific transcript variants by subtractive transcriptome-wide analysis. *PLoS One*. 2009;4(3):e4732.

33. El Marabti E, Younis I. The cancer spliceome: reprogramming of alternative splicing in cancer. *Front Mol Biosci.* 2018;5:80.
34. Shen R, Olshen AB, Ladanyi M. Integrative clustering of multiple genomic data types using a joint latent variable model with application to breast and lung cancer subtype analysis. *Bioinformatics.* 2009;25(22):2906-2912.
35. Singh A, Shannon CP, Gautier B, et al. DIABLO: an integrative approach for identifying key molecular drivers from multi-omics assays. *Bioinformatics.* 2019;35(17):3055-3062.
36. Hug N, Longman D, Caceres JF. Mechanism and regulation of the nonsense-mediated decay pathway. *Nucleic Acids Res.* 2016;44(4):1483-1495.
37. Jewer M, Findlay SD, Postovit LM. Post-transcriptional regulation in cancer progression: microenvironmental control of alternative splicing and translation. *J Cell Commun Signal.* 2012;6(4):233-248.
38. Ryan MC, Cleland J, Kim R, Wong WC, Weinstein JN. Splice-Seq: a resource for analysis and visualization of RNA-Seq data on alternative splicing and its functional impacts. *Bioinformatics.* 2012;28(18):2385-2387.
39. Schafer S, Miao K, Benson CC, et al. Alternative splicing signatures in RNA-seq data: percent spliced in (PSI). *Curr Protoc Hum Genet.* 2015;87:11.16.1-11.16.14.
40. Liu J, Li H, Shen S, et al. Alternative splicing events implicated in carcinogenesis and prognosis of colorectal cancer. *J Cancer.* 2018;9(10):1754-1764.
41. Eswaran J, Horvath A, Godbole S, et al. RNA sequencing of cancer reveals novel splicing alterations. *Sci Rep.* 2013;3:1689.
42. Liu JY, Wellen KE. Advances into understanding metabolites as signaling molecules in cancer progression. *Curr Opin Cell Biol.* 2020;63:144-153.
43. Li X, Egervari G, Wang Y, Berger SL, Lu Z. Regulation of chromatin and gene expression by metabolic enzymes and metabolites. *Nat Rev Mol Cell Biol.* 2018;19(9):563-578.
44. Wang D, Plukker JTM, Coppes RP. Cancer stem cells with increased metastatic potential as a therapeutic target for esophageal cancer. *Semin Cancer Biol.* 2017;44:60-66.
45. Islam F, Gopalan V, Wahab R, Smith RA, Lam AK. Cancer stem cells in oesophageal squamous cell carcinoma: identification, prognostic and treatment perspectives. *Crit Rev Oncol Hematol.* 2015;96(1):9-19.
46. Hackshaw A, Kirkwood A. Interpreting and reporting clinical trials with results of borderline significance. *BMJ.* 2011;343:d3340.
47. Duronio RJ, Xiong Y. Signaling pathways that control cell proliferation. *Cold Spring Harb Perspect Biol.* 2013;5(3):a008904.
48. Shubina MY, Musinova YR, Sheval EV. Nucleolar methyltransferase fibrillarin: evolution of structure and functions. *Biochemistry (Mosc).* 2016;81(9):941-950.
49. Amin MA, Matsunaga S, Ma N, et al. Fibrillarin, a nucleolar protein, is required for normal nuclear morphology and cellular growth in HeLa cells. *Biochem Biophys Res Commun.* 2007;360(2):320-326.
50. Dou Y, Kawaler EA, Cui Zhou D, et al. Proteogenomic characterization of endometrial carcinoma. *Cell.* 2020;180(4):729-748.
51. Clark DJ, Dhanasekaran SM, Petralia F, et al. Integrated proteogenomic characterization of clear cell renal cell carcinoma. *Cell.* 2019;179(4):964-983.
52. Archer TC, Ehrenberger T, Mundt F, et al. Proteomics, post-translational modifications, and integrative analyses reveal molecular heterogeneity within medulloblastoma subgroups. *Cancer Cell.* 2018;34(3):396-410.
53. Song Y, Li L, Ou Y, et al. Identification of genomic alterations in oesophageal squamous cell cancer. *Nature.* 2014;509(7498):91-95.
54. Audic Y, Hartley RS. Post-transcriptional regulation in cancer. *Biol Cell.* 2004;96(7):479-498.
55. Wang Y, Zhao H, Zhi W. SEMA4D under the posttranscriptional regulation of HuR and miR-4319 boosts cancer progression in esophageal squamous cell carcinoma. *Cancer Biol Ther.* 2020;21(2):122-129.
56. Liu F, Dai M, Xu Q, et al. SRSF10-mediated ILIRAP alternative splicing regulates cervical cancer oncogenesis via miLIRAP-NF-kappaB-CD47 axis. *Oncogene.* 2018;37(18):2394-2409.
57. Chen K, Xiao H, Zeng J, et al. Alternative splicing of EZH2 pre-mRNA by SF3B3 contributes to the tumorigenic potential of renal cancer. *Clin Cancer Res.* 2017;23(13):3428-3441.
58. Chen X, Zhang JX, Luo JH, et al. CSTF2-induced shortening of the RAC1 3'UTR promotes the pathogenesis of urothelial carcinoma of the bladder. *Cancer Res.* 2018;78(20):5848-5862.
59. Liu YB, Mei Y, Long J, et al. RIF1 promotes human epithelial ovarian cancer growth and progression via activating human telomerase reverse transcriptase expression. *J Exp Clin Cancer Res.* 2018;37(1):182.
60. Wang L, Wrobel JA, Xie L, et al. Novel RNA-affinity proteogenomics dissects tumor heterogeneity for revealing personalized markers in precision prognosis of cancer. *Cell Chem Biol.* 2018;25(5):619-633.
61. Kim HJ, Maiti P, Barrientos A. Mitochondrial ribosomes in cancer. *Semin Cancer Biol.* 2017;47:67-81.
62. Hosseini SM, Okoye I, Chaleshtari MG, et al. E2 ubiquitin-conjugating enzymes in cancer: implications for immunotherapeutic interventions. *Clin Chim Acta.* 2019;498:126-134.
63. Lin C, Wang Y, Wang Y, et al. Transcriptional and posttranscriptional regulation of HOXA13 by lncRNA HOTTIP facilitates tumorigenesis and metastasis in esophageal squamous carcinoma cells. *Oncogene.* 2017;36(38):5392-5406.
64. Wu D, He X, Wang W, et al. Long noncoding RNA SNHG12 induces proliferation, migration, epithelial-mesenchymal transition and stemness of esophageal squamous cell carcinoma cells via post-transcriptional regulation of BMI1 and CTNNB1. *Mol Oncol.* 2020;14(9):2332-2351.
65. Sun C, Li T, Song X, et al. Spatially resolved metabolomics to discover tumor-associated metabolic alterations. *Proc Natl Acad Sci U S A.* 2019;116(1):52-57.
66. Tokunaga M, Kami K, Ozawa S, et al. Metabolome analysis of esophageal cancer tissues using capillary electrophoresis-time-of-flight mass spectrometry. *Int J Oncol.* 2018;52(6):1947-1958.
67. Tessarz P, Santos-Rosa H, Robson SC, et al. Glutamine methylation in histone H2A is an RNA-polymerase-I-dedicated modification. *Nature.* 2014;505(7484):564-568.
68. Iyer-Bierhoff A, Krogh N, Tessarz P, et al. SIRT7-dependent deacetylation of fibrillarin controls histone H2A methylation and rRNA synthesis during the cell cycle. *Cell Rep.* 2018;25(11):2946-2954.e5.
69. Wertz IE, Wang X. From discovery to bedside: targeting the ubiquitin system. *Cell Chem Biol.* 2019;26(2):156-177.

70. Lioni M, Noma K, Snyder A, et al. Bortezomib induces apoptosis in esophageal squamous cell carcinoma cells through activation of the p38 mitogen-activated protein kinase pathway. *Mol Cancer Ther.* 2008;7(9):2866-2875.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Jin X, Liu L, Wu J, et al. A multi-omics study delineates new molecular features and therapeutic targets for esophageal squamous cell carcinoma. *Clin Transl Med.* 2021;11:e538. <https://doi.org/10.1002/ctm2.538>