Multiomic Landscape of Esophageal Squamous Cell Carcinoma: Molecular Profiles and Predictors of Neoadjuvant Chemoradiotherapy Responsiveness and Resistance

Yue Li
Shanghai Chest Hospital Affiliated to Shanghai Jiao Tong University School of Medicine

Jun Liu
Shanghai Chest Hospital Affiliated to Shanghai Jiao Tong University School of Medicine

Wen Yu
Shanghai Chest Hospital Affiliated to Shanghai Jiao Tong University School of Medicine

Xin-Yun Song
Shanghai Chest Hospital Affiliated to Shanghai Jiao Tong University School of Medicine

Xu-Wei Cai
Shanghai Chest Hospital Affiliated to Shanghai Jiao Tong University School of Medicine

Guang-Zhong Wang
CAS Key Laboratory of Computational Biology, Shanghai Institute of Nutrition and Health, University of Chinese Academy of Sciences, Chinese Academy of Sciences

Xiaolong Fu
xlfu1964@hotmail.com
Shanghai Chest Hospital Affiliated to Shanghai Jiao Tong University School of Medicine

Research Article

Keywords: Genetic signature, Multiomics, Treatment response, Response biomarker, Esophageal squamous cell cancer

Posted Date: April 29th, 2024

DOI: https://doi.org/10.21203/rs.3.rs-4295669/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.
Abstract

Background

In neoadjuvant chemoradiotherapy (NCRT), only 1/3 of patients with esophageal squamous cell carcinoma (ESCC) achieve pathologic complete response (pCR). Here, we aimed to depict the biological landscape of ESCC with different responses to neoadjuvant chemoradiotherapy and identify biomarkers to facilitate clinical decision-making.

Method

Tumor specimens before NCRT were obtained for whole exosome sequencing, RNA sequencing, and DIA mass spectrometry. Genomic data were analyzed for significantly mutated genes, copy number alterations, MSI, TMB, and mutational signatures. Transcriptomic and proteomic data were used to examine differentially activated pathways. GSEA and ActivePathways were used for single omics and joint multiomics analyses, respectively. Treatment-resistance biomarkers were identified and confirmed in a separate cohort using IHC.

Result

FBXW7 mutation (Fisher exact test p =0.029) and 9p21.3 cytoband loss(q-value=0.001) are the significant genetic variations in the pCR group. Combined transcriptomic and proteomic analyses revealed that the type I interferon signaling pathways and RIG-I-like receptor signaling pathways were enriched in non-pCR tumors. A biomarker panel of 12 proteins predictive of non-pCR tumors was identified, 10 of which were verified using mIHC in an independent cohort.

Conclusion

We described the multiomic biological characteristics of ESCC with distinct responses to neoadjuvant chemoradiotherapy and proposed a panel of proteins as predictive biomarkers for non-pCR patients.

Introduction

Esophageal cancer ranks seventh in incidence and sixth in mortality rate according to the latest global cancer statistics. China accounts for approximately half of all esophageal cancer cases worldwide, with squamous carcinoma accounting for more than 90% of all cases. More than half of esophageal cancers are locoregional diseases at diagnosis. For resectable locally advanced esophageal cancer, trimodality therapy of neoadjuvant chemoradiotherapy (NCRT) followed by surgery is currently the standard of care. However, significant heterogeneity exists among patients with esophageal cancer. In esophageal squamous cell carcinoma (ESCC), only ~40% of patients achieve pathologic complete response (pCR) after neoadjuvant chemoradiotherapy, and in esophageal adenocarcinoma patients, the ratio is around 20%. Despite complete surgical resection, pCR patients appear to have superior survival outcomes and
longer median time to recurrence than non-pCR patients.\cite{6} In order to improve the treatment outcome in esophageal cancer patients, it is imperative to elucidate the biological basis of NCRT resistance in esophageal cancer, as well as to find biomarkers for treatment-resistant patients to further individualize the treatment strategies, including delayed surgery and organ preservation for pCR patients in ESCC.\cite{7,8}

Previous and ongoing studies have reported different types of biomarkers for esophageal cancer NCRT response prediction, including gene polymorphisms and mRNA, miRNA, and protein markers. However, most previous studies were performed using microarray and immunohistochemistry, and studies taking advantage of next-generation sequencing to systemically discover biomarkers and differentially activated pathways are still lacking.\cite{9}

In this study, we performed multiomics sequencing on pre-treatment biopsied samples of esophageal squamous cell cancer and aimed to comprehensively depict the biological landscape of response and resistance to neoadjuvant chemoradiotherapy. Using whole exome sequencing(WES), transcriptomics, and proteomics data, we elaborated on both the genetic and environmental contributors to NCRT sensitivity and proposed biomarkers of NCRT resistance.

**Materials And Methods**

**Patient enrollment and sample processing**

The recruited patients were histopathologically diagnosed with esophageal squamous cell carcinoma and received neoadjuvant chemoradiotherapy followed by esophagectomy at our institute between 2016-2017.

Endoscopic biopsy samples of the primary tumor and paired peripheral blood samples were obtained from 24 patients before neoadjuvant chemoradiotherapy. The biopsied specimens were washed with PBS, transported in liquid nitrogen, and stored at -80°C until sequencing.

The neoadjuvant chemoradiotherapy regimen consisted of a total radiation dose of 41.4 Gy in 23 fractions, and the chemotherapy regimens were fluorouracil and cisplatin. Surgery was performed within 10 weeks of NCRT, and tumor regression of the surgical specimen was evaluated by pathologists. Pathologic complete response (pCR) was defined as no sign of microscopically viable cancer cells in the primary tumor and lymph nodes encompassed in the radiation field.

The study was reviewed and approved by the Institutional Review Board of Shanghai Chest Hospital(registration number: KS2040). Waiver of written informed consent for the biopsy samples was granted by the above IRB.

**DNA extraction and whole exome sequencing(WES)**
Genomic DNA was extracted from tumor tissues and paired peripheral blood samples using a TIANamp Genomic DNA Kit (TIANGEN). Samples without blood controls were excluded from the study. DNA was quantified using a Nanodrop spectrophotometer (Thermo Fisher Scientific, Inc., Wilmington, DE, USA), and the integrity was examined by 1% agarose electrophoresis. Exosomes were captured using Agilent SureSelect Human All Exon v6 library (Agilent Technologies, USA) following the manufacturer's protocol. The DNA libraries were amplified and sequenced on an Illumina sequencing platform (NovaSeq 6000, Illumina, Inc., San Diego, CA, USA), and 150 bp paired-end reads were generated.

RNA extraction and RNA sequencing

RNA extraction was performed using TRlzol reagent (Invitrogen), and the integrity was assessed using an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). Samples showing significant degradation during nucleic acid electrophoresis were removed. Libraries were constructed using the VAHTS Universal V6 RNA-seq Library Prep Kit according to the manufacturer's instructions and sequenced on the Illumina Novaseq 6000 platform.

DIA proteomics

Approximately 100 mg of frozen tumor samples from each patient were used for protein extraction using phenol extraction buffer. The total protein concentrations were measured, and the same quantity of protein from each sample was digested and desalted. Samples with an insufficient quantity of protein were removed. A Q-Exactive-HF mass spectrometer (Thermo Fisher Scientific) and Easy nanoLC-1200 (Thermo Fisher Scientific) were used for both the shotgun proteomics and DIA experiments. A survey scan from m/z 350 to 1250 m/z at 35,000 resolution were adopted. Spectronaut Pulsar™ 15.3.210906.50606 (Biognosys, Switzerland) search and library construction were performed using default settings. The DIA data were analyzed with Spectronaut, and the results were filtered based on a value cutoff of 0.01 (equivalent to FDR < 1%).

WES data analysis

The raw data from paired tumor/blood samples were demultiplexed and converted into the FASTQ format. Raw reads were pre-processed using fastp(Version:0.20.0) to obtain high-quality reads. Using BWA, clean reads were aligned to the GRCh37 reference human genome (version:0.7.17). Based on the matched tumor and blood samples, MuTect(Version:2.0) software was used to identify somatic SNV/INDEL sites, and ANNOVAR software was used for annotation. With default parameters, MuSiC (version 0.4.1-1) further uses somatic variants (including SNVs and INDELs) to identify significantly mutated genes (SMGs). SMGs have a noticeably higher mutation rate than the background.
By using the Non-negative Matrix Factorization (NMF) method and aligning it with the known mutational signatures from the COSMIC database, the mutational signature of somatic SNV was retrieved. The purity and ploidy of the tumor samples were determined using a probabilistic model in Sequenza (Version: 2.1.2). Genomic Identification of Significant Targets in Cancer (GISTIC, version 2.0.23) was used to identify significant focal CNV (copy number variation) events and estimate log2-transformed, gene-level, or segment-level CNV ratios. The GISTIC calls included -2 (deletion), -1 (loss), 0 (diploid), 1 (gain), and 2 (gain) (amplification).

The proportion of MSI sites in each tumor sample was determined using MSIsensor software. The number of bases covered by the exome capture kit (34.3944 Mb) was divided by the total number of non-silent somatic SNV and indels inside the exonic regions to determine TMB.

For neoantigen analysis, individual human leukocyte antigen (HLA) typing was performed using Optitype from the WES data. For each non-synonymous SNV or non-frameshift indel, we built an amino acid FASTA sequence (mutant) with 23 amino acids, containing the mutated amino acid at position 12 and its corresponding wild-type peptides (normal). For frameshift mutations, the FASTA sequence was built from 12 amino acids upstream of the mutation towards the end of the transcript using the CCDS database from NCBI. IEDB netMHCpan 4.1 was used to predict the HLA class I binding strength of each peptide. Mutant peptides with binding strength (IC50) < 500 nM and corresponding normal peptides with IC50 > 500 nM were predicted as neoantigens. Neoantigens with fragments per kilobase per million (FPKM) > 1 were defined as expressed.

**RNA-seq data analysis**

Fastp was used to first process raw reads in fastq format, removing any low-quality reads to provide clean reads. Using HISAT2, the clean reads were mapped to the reference genome. To evaluate the distribution of reads on the gene body, RseqQC was used. Because of the distribution bias on the 3’ end of mRNAs, read counts within the 3’UTR region of each gene were obtained by HTSeq-count to ensure the accuracy of mRNA expression quantification.

**Gene expression profile analysis**

For RNA-seq data, FPKM was calculated to normalize gene expression. Differential expression analysis of transcriptomic data was carried out using DESeq2 (R package v1.38.3) and edgeR (R package v3.40.2). For proteomics data, limma (R package v3.54.1) and t-test were used for differential expression analysis.

To investigate the differentially activated cellular functional pathways in different response groups across transcriptomics and proteomics, R package clusterProfiler was used. We also performed a multiomics joint analysis using ActivePathways with the RNA and protein expression of all genes as input data. Adjusted p-values were calculated using the Benjamini-Hochberg method.
Public dataset acquisition

The clinical information of 185 samples from the TCGA ESCA cohort was obtained from the TCGA data portal, and 36 out of 185 ESCC samples with recorded radiotherapy responses were chosen for biomarker validation analysis.

ESCC cell culture and Treatment

The human ESCC cell line TE-1 (RRID:CVCL_1759) and KYSE510 (RRID:CVCL_1354) were purchased from the Cell Bank/Stem Cell Bank, Chinese Academy of Sciences. Both cell lines have been authenticated using STR profiling within the last three years. All cellular experiments were performed with mycoplasma-free cells.

TE-1 and KYSE510 cells were cultured in RPMI1640 medium (Cytiva) with 10% FBS (Gibco, Life Technologies Inc.) at 37°C in 5% CO2 /95% air.

Cell Radiation treatment: When growing up >70-80% confluence, ESCC cells were irradiated using the Small Animal Radiation Research Platform (SARRP)(Xstrahl Inc.) at a dose rate of 2.4 Gy/min.

Interferon treatment: TE1 was subjected to a 16-day continuous low-dose IFNβ (Chamot Biotechnology Co., Ltd) stimulation regimen(5 IU/ml), with media replacement every other day, and the addition of fresh IFNβ.

Apoptosis will be detected through staining of the cells with Annexin V and propidium iodide(PI) solution(LIANKE Bio.), followed by flow cytometry analysis.

FBXW7 Gene knockdown and IFN level measurement

FBXW7 gene knockdown was accomplished by employing siRNA with the following sequence: Sense (5’-3’): CGGGUGAAUUUAUUCGAAATT; Antisense (5’-3’): UUUCGAAUAUAUCACCCGTT. Human ESCC cell line KYSE510 was used for the gene knockdown. KYSE510 cells were seeded to reach a confluence of approximately 70% before transfection. The transfection reagent used was GP-transfect-Mate (GenePharma), with the siRNA working concentration of 100 nmol/L. In the gene knockdown group, transfection was executed with a mixture of transfect-mate (5 μL in 1 mL) and siRNA (100nM), while the control group received transfect-mate exclusively. The transfection efficiency was evaluated using qRT-PCR.

Intracellular type I IFN levels were quantified using IFNα and IFNβ ELISA kits(XY-biotechnology). IFN level assessment was conducted 48 hours after the siRNA transfection. Treated cells were subjected to freeze-thaw cycles to induce cell lysis, and the supernatant was collected for measurement of IFNα and IFNβ levels.
**Immunofluorescence**

Immunofluorescence was performed for STAT1, EIF2AK2, MX1, BST2, TRIM21, SAMHD1, IFI44L, GBP1, PARP14, ISG15, HLA-B, and IFIT3, and the tumor cells were stained with P63. The TSAPLus fluorescent multiplex immunohistochemistry kit (Wuhan Servicebio Technology Co., Ltd.) was used for antigen staining. The antibodies used are listed in table S1. The stained sections were scanned using Pannoramic DESK (3DHISTECH Ltd.).

**Statistical analysis**

Statistics were performed using R (version 4.0.4). Unless otherwise specified, the continuous variables and categorical variables are compared with student t-test and Chi-squared test; Benjamini and Hochberg FDR (False Discovery Rate) procedure was employed to adjust p-values for multiple comparisons; tests were two sided, and a p-value or an adjusted p-value of less than 0.05 considered significant. The correlation analysis was performed by calculating Spearman or Pearson correlation coefficient.

The joint analysis of transcriptomics and proteomics data was performed using ActivePathways in R. The significance of the enriched pathways across multiomics data was evaluated by hypergeometric P-value, which was corrected as $Q_{\text{pathway}}$ using the Holm Bonferroni method of family-wise error rate (FWER). $^{11}$ Pathways with $Q_{\text{pathway}} < 0.05$ were reported as significant.

Significantly mutated genes (SMG) in two response groups were analyzed using Fisher exact test. The TMB difference were analyzed using Mann-Whitney U test. In CNV analysis, the adjusted p-value using the Benjamini and Hochberg FDR procedure was calculated by GISTIC2.0 package and reported as q-value.

Kaplan-Meier plotter (kmplot.com/analysis/) was used to evaluate the correlation of the biomarker panel expression level and the survival outcomes of ESCC patients, which utilizes gene expression data sources including TCGA, EGA, and GEO etc. $^{12}$

**Results**

**Patient information and Quality control of samples**

Among the 24 patients, 11 achieved pCR (54%). The baseline clinical characteristics including sex, age, smoking history, tumor location, tumor clinical T stage and N status were similar between two response groups (table S2).

WES was performed on 18 tumor samples, as six lacked blood control. Seven RNA samples showed significant degradation on electrophoresis, and one sample contained insufficient protein for
spectrometry. These samples were not included in the experiment. In total, RNA sequencing and DIA spectrometry were performed on 17 and 23 samples, respectively (table S2).

**Interferon signaling and RIG-I-like receptor signaling pathways were enriched in non-pCR tumors in the joint analysis of transcriptomics and proteomics**

RNA-seq was performed in 17 of the 24 patients. Genes differentially expressed between pCR and non-pCR patients were analyzed using EdgeR and DESeq2 in R. The 85 genes at the intersection of the two results were identified as differential genes (p < 0.005, \( \log_2 FC > 0.58 \) or \(<-0.58 \)) (figure S1, table S3). GSEA analysis highlighted that the most significantly enriched pathways in the non-pCR group were the type I interferon signaling pathway and the defense response to viruses (adjusted p-value = 0.000). Conversely, the pCR patients showed significant enrichment in keratinization-related pathways (adjusted p-value < 0.001) (figure 1A and table S4). Furthermore, in non-pCR patients, the MDA-5 signaling pathway, an upstream regulator of interferon signaling, was enriched, as revealed by Gene Ontology (GO) enrichment analysis of the differential genes (adjusted p-value = 0.051).

DIA mass spectrometry was performed in 23 of the 24 patients, identifying 4080 proteins. Analysis using PCA and unsupervised hierarchical clustering with valid proteins revealed distinct protein expression patterns in patients with distinct treatment responses (figure 1S). Combined limma and t-test identified 385 differentially expressed proteins (p < 0.05, \( \log_2 FC > 0.58 \) or \(<-0.58 \)) between pCR and non-pCR patients (figure 1S). Gene ontology enrichment analysis was performed based on the differential proteins, and the top 30 enriched pathways were all derived from the non-pCR group (figure 1B). The top enriched pathways included RNA catabolic-related and neutrophil activation pathways.

In accordance with the transcriptomic enrichment results, the type I interferon signaling pathway was also among the significant pathways enriched in the non-pCR group (adjusted p-value = 0.003). In addition, the dsRNA response pathway and Wnt signaling pathway enriched in proteomics corresponded with the MDA-5 signaling pathway and Wnt protein secretion pathway in transcriptomics, respectively (figure 1C).

To identify the consistently enriched pathways across different gene expression omics while also detecting pathways not apparent in any single-omic dataset, we integrated the transcriptomics and the proteomics data using ActivePathways package in R.11 In genes elevated in the non-pCR group, the joint analysis of ActivePathways highlighted 315 genes that were significantly enriched in 165 pathways (figure 1D). Approximately 20% of the enriched pathways were supported by data from both RNA and protein levels, and 24% of the pathways were identified by multiomics joint analysis. Notably, the type I interferon signaling and interferon regulatory pathways were frequently observed in these two categories. Additionally, our analysis highlighted the significant enrichment of the upstream pathway of interferon, the "RIG-I signaling pathway," in the non-pCR group (Q_{pathway} = 0.015).

Both RIG-I and MDA5 are members of retinoic acid-inducible gene I (RIG-I)-like receptors (RLRs). RLRs are RNA sensors localized in the cytosol and could induce type I interferon production in the detection of
unusual nucleic acids.\textsuperscript{13} Our data support that the activation of both the RLR pathway and its downstream interferon signaling pathway contributes to neoadjuvant chemoradiotherapy treatment resistance in ESCC. It's well established that the JAK-STAT pathway is the predominant downstream signal transduction pathway after interferon signal activation.\textsuperscript{14} In our differential expression gene analysis, STAT1 was significantly up-regulated in non-pCR group in both transcriptomics data (p-value=0.001, FC= 0.43) and proteomics data (p-value=0.03, FC= 0.48). Our findings are corroborated by previous research, which revealed that RIG-I-induced STAT1 activation leads to chemotherapy and radiotherapy resistance in breast cancer.\textsuperscript{15}

**Genetic characteristics of patients receiving NCRT with different response**

A total of 18 biopsied samples were subjected to whole exome sequencing with the corresponding blood samples as controls, among which 9 samples (50\%) were pCR after NCRT.

First, we evaluated the general genetic differences between the pCR and non-pCR groups (figure 2). Tumor mutation burden (TMB), described as mutations (SNVs or indels) per megabase, was not significantly different between groups (pCR median 10.32, SD3.63, IQR4.84; non-pCR median 9.36, SD4.75, IQR3.05; Mann-Whitney U test, p=0.840). The differences in other genetic variations (table S5), including total somatic CNA length (t-test p=0.608 for copy number gain, p=0.392 for copy number loss), somatic variation (SV) counts (t-test p=0.358), microsatellite instability (MSI) (t-test p=0.729), and tumor purity (t-test p=0.889), were not statistically significant between the two response groups. Gene mutation signatures were investigated via nucleotide mutation signature and the Catalog of Somatic Mutations in Cancer (COSMIC) database v3.\textsuperscript{16} the results revealed that the most common nucleotide mutations were C>T and T>C among all samples; there was no nucleotide mutational difference between the two groups (figure S2). The COSMIC signatures SBS25 and SBS9 were detected in all patients, and SBS9 was more frequently detected in the pCR group (figure S2). We further analyzed the differences in neoantigen load between the groups. Neither neoantigen load nor expressed neoantigen load was associated with the pathological response (figure S3).

**Gene mutations**

Next, we performed a significantly mutated gene (SMG) analysis of the two response groups. TP53 was the most mutated gene with a mutation frequency of 72\% across all samples (figure 3A). FBXW7 was the only gene with significantly different mutation frequencies between the groups (Fisher exact test p-value=0.029), with five out of nine pCR patients carrying non-synonymous mutations, in contrast to no mutation in the non-pCR group. Mutation types included missense and frameshift mutations in the F-box, COG2319, and WD40 domains (figure 3B).
In a previous study, Gstalder et al. demonstrated that inactivation of FBXW7 in a melanoma model was associated with impaired tumor-intrinsic expression of the dsRNA sensors MDA5 and RIG-I, and thus decreased induction of type I interferon expression, which leads to anticancer therapy resistance. Given the observed differences in RLR and type I interferon signaling pathway expression levels between pCR and non-pCR patients, we sought to investigate if these expression variations were linked to FBXW7 mutation status. To do so, we conducted a differential gene expression and enrichment analysis, comparing samples with FBXW7 mutations to those with FBXW7 wild-type status using transcriptomic data. The results of GSEA revealed a notable reduction in the expression of the type I interferon signaling pathway in FBXW7 mutated samples (figure 3C and 3D). Meanwhile, GO enrichment analysis showed that the interferon upstream RLR pathways, including the MDA-5 signaling pathway (adjusted p-value = 0.002) and the RIG-I signaling pathway (adjusted p-value = 0.018), were both downregulated in FBXW7 mutated samples by GO enrichment analysis. RNA expression data revealed positive correlations between FBXW7 expression and key RLR pathway genes (Spearman correlation coefficient IFIH1 r = 0.602, DDX58 (RIG-I) r = 0.400, TBK1 r = 0.694), as well as type I interferon-induced JAK-STAT pathway key gene expression (Spearman correlation coefficient STAT1 r = 0.449, JAK1 r = 0.588, TYK2 r = 0.306) (figure 3E) in 17 samples that underwent RNA-seq. Moreover, we analyzed the TCGA ESCC dataset (esophageal carcinoma, TCGA, Firehose Legacy) on cBioPortal, and verified the co-expression tendency of FBXW7 with key genes in the RLR and JAK-STAT pathways (measured by Spearman correlation coefficient, table S6).

In addition to its regulatory effect on RLR and interferon signaling, FBXW7 was also reported to contribute to DNA repair via nonhomologous end-joining and recovery of the cell cycle by promoting P53 degradation after DNA-damaging therapies. Inactivation of FBXW7 enhanced tumor radiosensitivity and chemosensitivity. The GO enrichment analysis in FBXW7 mutated samples in our cohort exhibited a significantly lowered expression in “DNA damage response” pathway (adjusted p-value = 0.002).

Collectively, our data suggest that inactivation of FBXW7 in ESCC is sensitive to neoadjuvant chemoradiotherapy. It might exert an effect on the treatment response through the regulation of interferon-related pathways and DNA damage response pathways.

**Copy number variations**

Copy number variations (CNV) were investigated in all samples using GISTIC (figure S4). In the non-pCR group, the most significant copy number variations were 20q13.33 amplification (q-value = 0.000) and 4q35.2 deletion (q-value = 0.044). In the pCR group, 8q24.3 (q-value = 0.000) was the top amplified cytoband, encompassing genes that encode a series of microRNAs and oncogenes such as PARP10. The most significant cytoband loss in the pCR group was 9p21.3 (q-value = 0.001), which is a 917kb long chromosome band containing CDKN2A/B, MTAP, and a type I IFN gene cluster (figure 3F, 3G). Loss of chromosome 9p21.3 has been reported as the most common homozygous deletion in human cancers. While 9p21.3 deletion almost always involves CDKN2A, co-deletion of the IFN gene cluster only occurs in...
certain patients, which disrupts the tumor microenvironment and changes the tumor biological behaviors in these patients.\textsuperscript{20}

In our pCR cohort, GISTIC results showed that five out of nine pCR patients carry 9p21.3 copy number variations, including CDKN2A deletion and co-deletion of IFN genes. We presume that the chromosome 9p21.3 IFN cluster deletion may partially contribute to the relatively lower expression of the type I IFN signaling pathway in the pCR group at both the transcriptomic and proteomic levels. In the ESCC TCGA cohort (esophageal carcinoma, TCGA, Firehose Legacy) from cBioPortal, 51 out of 96 patients carry 9p21.3 deletion, among which 13 patients had co-deletion of IFN genes. Initially, we found no significant difference in the expression level of the interferon signaling pathway between patients with and without non-IFN cluster 9p21.3 copy number deletion. However, when comparing patients with and without 9p21.3 deletions that included the IFN cluster, we observed a significant increase in the expression of RIG-I, IFN signaling, and JAK-STAT pathways in the copy number deletion group. (figure S4). 36 patients in the TCGA ESCA cohort had recorded response evaluation of radiotherapy, among which three patients were found to carry 9p21.3 IFN cluster deletion, and all of their radiotherapy responses were complete remission (CR).

The findings in the TCGA cohort reaffirmed our assumption that chromosome 9p21.3 copy number deletion with IFN-cluster is a sensitive genetic variance that enhances ESCC neoadjuvant chemoradiotherapy treatment response, and it might exert its pro-therapeutic effect by downregulating interferon signaling in the tumor cells.

**Assessment of the IFN effect on ESCC cell line radiosensitivity**

Previous studies have reported that constitutive elevation of type I IFN was seen in some cancers.\textsuperscript{21} And the chronic exposure to low dose type I IFN may promote cell resistance to DNA damage through elevated expression of a subset of IFN induced genes including STAT1, STAT2 and IRF9.\textsuperscript{22} To investigate whether chronic exposure to type I IFN promotes radiotherapy resistance of esophageal squamous cell cancer, we treated the ESCC cell line TE1 with low-dose IFN\(\beta\) (5 IU/ml) continuously for 16 days. The treated cells and untreated control received irradiation of 0, 2, 4, 8Gy, and apoptotic percentages were measured 48h after radiation.

In comparison to the control group, TE1 cells treated with 5 IU/ml IFN\(\beta\) for 16 days exhibited a significantly reduced number of apoptotic cells following exposure to 2 Gy (t-test p=0.031), 4 Gy (t-test p=0.018), and 8 Gy (t-test p=0.012) radiation (figure 4A-B), indicating increased radioresistance after low-dose IFN\(\beta\) treatment. The results supported the hypothesis that the primary radioresistance observed in NCRT treatment may derive from the constitutive activation of type I IFN signaling in ESCC.
Knockdown of FBXW7 alters intracellular IFN levels

To evaluate the effect of FBXW7 mutation on the ESCC type I IFN expression levels, we used siRNA to knockdown FBXW7 in KYSE510 cell line. After transfection for 48 hours, the expression of FBXW7 was downregulated to 40%. The intracellular levels of both IFNα and IFNβ were significantly decreased in the FBXW7-siRNA group compared with the control group (IFNα p=0.000, IFNβ p=0.010) (figure 4C). This result supports our hypothesis that the loss-of-function mutation of FBXW7 could lower the constitutive expression level of type I IFN in ESCC, and may further alter the radiosensitivity through this effect.

Interferon stimulated genes (ISGs) as resistance biomarkers of ESCC patients treated with neoadjuvant chemoradiotherapy

After neoadjuvant chemoradiotherapy, patients with pathological complete remission have distinct gene expression profiles at the protein level compared with non-pCR patients. We aimed to define a protein biomarker panel that can identify NCRT-resistant tumors. Protein-protein interaction (PPI) analysis was performed using 385 differentially expressed proteins in the two response groups. Next, the genes in the PPI network were ranked based on their topological score calculated using the cytoHubba MCC method. 45 hub genes were identified at a cutoff score of 50000. The hub genes were further selected using transcriptomic data, and the final protein biomarkers were defined as the intersection of PPI hub genes and transcriptomic differential expression genes (DESeq2, p-value<0.05) (table S7). The biomarker panel comprised 12 proteins: STAT1, EIF2AK2, MX1, BST2, TRIM21, SAMHD1, IFI44L, GBP1, PARP14, ISG15, HLA-B, and IFIT3. All proteins had relatively elevated expression in the non-pCR group; thus, the panel could identify ESCC patients with potential NCRT resistance. The corresponding genes of these proteins were all members of the interferon-stimulated gene family, and active interactions were observed within their protein-protein interaction (PPI) network (figure 5A).

The 12-protein biomarker panel was tested in the 36-patient subset with a recorded radiotherapy response from TCGA ESCA cohort (table S8). GSEA demonstrated significant enrichment of the biomarkers in non-complete response (radiographic progressive disease, stable disease) patients (NES -1.852, adjusted p-value 0.002) (figure 5B), as expected. Stratified by high and low levels of the biomarker panel expression, Kaplan-Meier plotter (kmplot.com/analysis/) was used to assess the survival outcomes of ESCC patients from databases including TCGA, EGA, and GEO etc. Patients with higher biomarker expression level had significantly impaired overall survival (median OS, 18.9 months) compared with the low-expression subset (median OS, 42.1 months) (HR:2.29, 95% CI 1.00-5.23; log-rank p = 0.044) (figure 5C).

To further test the discriminative potential of the 12 NCRT-resistant biomarkers, we performed immunofluorescence staining of the biomarkers using FFPE slides from an independent cohort of NCRT-
treated ESCC patients. Each biomarker was stained on pre-treatment (NCRT) biopsied primary tumor samples from 5 pCR patients and 5 non-pCR patients. Biomarker positive score (BPS) was defined as the percentage of tumor cells with positive biomarker staining at any intensity (the number of positive tumor cells divided by the total number of tumor cells). The BPS values of the 12 biomarkers are shown in figure 5D (p-values for t-test are listed in table S9). Ten out of the 12 biomarkers had significantly elevated expression in the tumor cells of the non-pCR patients using two-sample t-test, including IFIT3 (p=0.000), BST2 (p=0.000), SAMHD1 (p=0.000), HLA-B (p=0.000), EIF2AK2 (p=0.001), GBP1 (p=0.002), IFI44L (p=0.003), MX1 (p=0.006), STAT1 (p=0.026), ISG15 (p=0.038) (figure 5E and figure S5).

Discussion

Trimodality therapy with neoadjuvant chemoradiotherapy followed by surgery is currently the standard treatment for locally advanced esophageal cancer. However, approximately 60% of esophageal squamous cell cancers and ~80% of esophageal adenocarcinomas cannot achieve a complete pathological response due to intrinsic or acquired tumor chemoradiation resistance. Currently, the mechanism of resistance to NCRT in esophageal cancer has not been comprehensively investigated using a systemic multiomics approach. Moreover, treatment response biomarkers evidenced by high-throughput techniques are still lacking.9

In our study, we aimed to unveil the biological basis for therapeutic resistance by jointly analyzing genomics, transcriptomics, and proteomics data from pre-treatment biopsy samples of 24 ESCC patients. Our findings highlighted the interferon signaling pathway and its upstream RIG-like receptor pathway as key players in poor responders. At the genetic level, we identified FBXW7 loss-of-function mutations and chromosome 9p21.3 copy number deletion as sensitive genetic variations to NCRT. We linked these variations to the interferon and RLR pathway expressions, validated in the TCGA ESCC dataset. Based on proteomic and transcriptomic insights, we proposed a 12-protein biomarker panel (STAT1, EIF2AK2, MX1, BST2, TRIM21, SAMHD1, IFI44L, GBP1, PARP14, ISG15, HLA-B, and IFIT3) from the interferon stimulated genes (ISGs) family to identify treatment-resistant patients. Preliminary validation in the TCGA ESCC subset and confirmation in an independent ESCC cohort supported the potential of these biomarkers.

Constitutive activation of IFN pathway and expression of interferon-stimulated genes (ISGs) in cancer are related to genotoxic agents’ treatment resistance.25 In a previous study, a series of STAT1-centered ISGs that were found to be highly expressed in radioresistant tumor nu61, and were designated as IFN-related DNA damage signature (IRDS). IRDS were proved to be associated with radioresistance and chemoresistance in different cell lines, including melanoma, breast cancer, lung cancer, and colon cancer.26,27 In clinical settings, a classifier incorporating seven ISGs, namely STAT1, IFI44, IFIT3, OAS1, IFIT1, G1P2, and MX1, successfully categorized patients’ IRDS status. The classifier was validated across multiple independent clinical breast cancer datasets, demonstrating its reliability in predicting recurrence risk following radiotherapy and chemotherapy. Additionally, elevated expression levels of certain other ISGs, including DDX60, STAT1, OAS1, IFI6, and IFI27, were found to be associated with unfavorable radiotherapy outcomes in cancer patients.28 Our proposed NCRT treatment biomarkers demonstrated a
significant alignment with those documented in the literature, affirming the reliability of our biomarker panel.

The mechanisms by which cancer cells constitutively overexpress ISGs and how elevated ISGs promote radioresistance and chemoresistance are poorly understood. Extensive literature supports that STAT1 is the core gene responsible for therapeutic resistance and oncogenic phenotypes.\textsuperscript{25} Beolens et al. found that exosome RNAs (exoRNA) originating from cancer stroma could induce ISGs in cancer cells through RIG-I-dependent activation of STAT1, which would then amplify the NOTCH3 signaling to expand the therapy-resistant cells with stemness phenotype.\textsuperscript{15} Specific ISGs downstream of STAT1 were also reported to promote oncogenicity or genotoxic stress protection. For example, ISG15 was involved in reinforcing the stem-like phenotype of pancreatic ductal adenocarcinoma cancer stem cells.\textsuperscript{29} IFIT1 and IFIT3 were found to promote tumor growth and metastasis in oral squamous cell carcinoma.\textsuperscript{30} PARP14 was found to promote cell-cycle progression through cyclin D1 mRNA stabilization, and its overexpression is related to treatment resistance and poor prognosis in multiple cancers.\textsuperscript{31,32}

FBXW7, known as a tumor suppressor gene, is a key component of E3 ligase that targets the degradation of a series of proto-oncogenes, such as cyclin E and MYC. Its mutations are commonly observed in different cancer types. In our study, the mutation frequency of FBXW7 was 56% in the pCR group and 0% in the non-pCR group, and the mutation types included missense and frameshift mutations. These mutations included missense and frameshift variants, which either disrupted FBXW7 dimerization or its ability to bind to substrates. FBXW7 loss-of-function mutations have been reported to increase radiation sensitivity. Yang et al. demonstrated FBXW7 knockout led to increased radiotherapy efficacy both in vitro and in vivo as a result of impaired double-strand break repair.\textsuperscript{33} In another research, depletion of FBXW7 was found to sensitize cancer cells to radiation through stabilization of p53 and induction of cell cycle arrest and apoptosis.\textsuperscript{19} Besides its established role in DNA damage repair, our results also suggest FBXW7 mutations confer chemoradiation sensitivity through impaired RLR and IFN signaling pathway. While early research seldom discusses FBXW7's role in immunity, a few recent studies have revealed its importance in regulating the IFN signaling pathway. Song et al.\textsuperscript{34} first described the role of FBXW7 in antiviral immunity as a stabilizer of RIG-I, and FBXW7 knockout led to decreased RIG-I protein levels and impaired type I IFN signaling. In a murine melanoma model, a lack of FBXW7 was also found to be associated with diminished expression of both RIG-I and MDA-5, as well as impaired induction of IFN signaling.

In esophageal cancer, chromosome 9p21.3 deletion is the most frequent homozygous deletion among all CNAs, and it is thought to be one of the early driving events in precancer lesions.\textsuperscript{35} Co-deletion of the IFN cluster in the 9p21.3 deletion has been associated with the disruption of type I IFN signaling, thus causing tumor immune evasion and immunotherapy resistance.\textsuperscript{20,36} Our results show that deletion of 9p21.3 sensitizes esophageal squamous cell cancer to chemoradiation, potentially through impaired IFN pathway activation.
Our study had some limitations. Genetic variations related to NCRT resistance are highly heterogeneous, and the gene mutations and CNVs proposed in this study are insufficient to explain all scenarios. Furthermore, multiomics sequencing was not performed in all the recruited samples owing to the quality control of the frozen specimens. Only 13 patients had complete sequencing data from all three omics levels.

In summary, our multiomics analysis provided valuable insights into the distinctive biological profiles of ESCC patients responding differently to neoadjuvant chemoradiotherapy. Through the integration of genomic, transcriptomic, and proteomic data, we uncovered the consistent activation of the IFN and RLR signaling pathways as prominent characteristics in NCRT-resistant patients. Additionally, we introduced a promising ISG-based biomarker panel with the potential to predict treatment resistance in these patients.

**Declarations**

**ADDITIONAL INFORMATION:**

**Funding Statement:** Shanghai Chest Hospital Project of Collaborative Innovation (YJXT20190202Z) program.

**Corresponding author:** Prof. Dr. Xiaolong Fu, Department of Radiation Oncology, Shanghai Chest Hospital, Shanghai Jiao Tong University. 241 West Huaihai Rd., Shanghai, China. E-mail: xlfu1964@hotmail.com; Dr. Guang-Zhong Wang, CAS Key Laboratory of Computational Biology, Shanghai Institute of Nutrition and Health, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai 200031, China. E-mail: guangzhong.wang@picb.ac.cn

**Conflict of interest disclosure statement:** The authors declare no potential conflicts of interest.

**Data Access Statement:** Research data supporting this publication will be provided upon request.

We thank the Shanghai Leading Talent Program for financial support.

**Data availability**

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

**References**


Figures
Figure 1

Differential gene expression and pathway enrichment in transcriptomics level and proteomics level.

A, RNA-seq data was used in GSEA pathway enrichment analysis (enrichment in the pCR group, enrichment score > 0; enrichment in the non-pCR group, enrichment score < 0). Top 30 enriched pathways are presented.
B, GO enrichment analysis of the differential proteins (DIA mass spectrometry data was used for analysis). Top 30 differential pathways were enriched in the non-pCR group.

C, Selected corresponding pathways enriched in transcriptomics and proteomics.

D, ActivePathways enrichment map showing pathways enriched in transcriptomics, proteomics, and omics joint analysis. Nodes represent pathways annotated by Gene Ontology. Magenta coloring represents evidence from transcriptomics; orange coloring represents evidence from proteomics; pink coloring represents evidence from joint omics.
Figure 2

Genetic landscape

The plot displays the overall genetic landscape of pCR tumors compared to non-pCR tumors. Tumor mutation burden (TMB) includes single nucleotide variations (SNVs) and insertion-deletion mutations (indel). TMB was similar between response groups (Mann-Whitney U test, $p=0.840$). Gene
mutation signatures analysis included single nucleotide substitution signatures (second line) and COSMIC signatures (third line). Total somatic I length (t-test p = 0.608 for copy number gain, p = 0.392 for copy number loss), somatic variation (SV) counts (t-test p = 0.358), microsatellite instability (MSI) (t-test p = 0.729), and tumor purity (t-test p = 0.889), were not statistically significant between the two response groups.
Gene mutations and copy number variation analysis

A, Gene mutation plot. The top bar of the plot shows the mutation burden of each sample, parted by synonymous mutation (Syn) and non-synonymous mutation (NonSyn). The major part of the plot displays the top mutated genes in each sample. The left column shows the mutation frequency of each gene.

B, Types of FBXW7 gene mutations detected in WES in our pCR cohort. The y-axis represents the mutated sample number.

C, GSEA enrichment analysis using transcriptomic data shows type I interferon signaling pathway is enriched in FBXW7 wild-type samples and has decreased expression in FBXW7 mutated samples (which are also pCR samples).

D, The gene expression heatmap of the type I interferon signaling pathway in FBXW7 mutated versus wild-type samples.

E, FBXW7 expression level is positively correlated with key gene expression levels from (RIG-I)-like receptor (RLR) pathway and JAK-STAT pathway.

F, 9p21.3 cytoband copy number difference between pCR and non-pCR group.

G, Schematic view of cytoband 9p21.3 and the genes mapped on this region. The IFN cluster contains an array of type I interferon genes.
Figure 4

In vitro assessment of IFN effects on radiosensitivity and FBXW7 KD effect on IFN level

A-B, Esophageal squamous cell cancer cell line TE1 was treated with low-dose (5 IU/ml) IFNβ continuously for 16 days and received radiotherapy. Cell apoptotic rates were measured 48h after radiation. Cells
treated with IFNβ shows increased radiation resistance in 2,4,8Gy radiation doses. Comparisons were made using t-test: 0Gy(p=0.073), 2 Gy(p=0.031), 4 Gy(p=0.018), and 8 Gy(p=0.012). Error bars show SEM.

ns: Not significant, * p<0.05

C,FBXW7 siRNA knockdown lowers KYSE510 intracellular type I IFN levels. Intracellular IFNa and IFNβ levels were measured by ELISA assay 48h after transfection. Control group receives transfer-mate only. Error bars show SEM.
Figure 5

Type I IFN signaling pathway genes as biomarkers for NCRT treatment resistance

A. Protein-protein interaction network of biomarker proteins plotted by STRING. Edges represent protein-protein associations. Known interactions: blue edge - from curated databases; magenta edge-...

B, GSEA analysis of the biomarker panel in the TCGA ESCC cohort shows significant enrichment of the biomarkers in non-CR patients.

C, Survival analysis of ESCC patients using Kaplan-Meier plotter (kmplot.com/analysis/) shows a significantly prolonged median OS of 42.1 months in patients with low expression of the biomarkers, compared with the median OS of 18.9 months in patients with high expression of the biomarkers.

D, Biomarker positive scores (BPS) for each gene. BPS is defined as the percentage of tumor cells with positive staining at any intensity. Comparisons are made using t-test. IFIT3(p=0.000), BST2(p=0.000), SAMHD1(p=0.000), HLA-B(p=0.000), EIF2AK2(p=0.001), GBP1(p=0.002), IFI44L(p=0.003), MX1(p=0.006), STAT1(p=0.026), ISG15(p=0.038), TRIM21(p=0.073), PARP14(P=0.0306).

*p<0.05, **p<0.01, ***p<0.005

E, The positive and negative staining of representative samples for each biomarker gene. Blue-DAPI; Pink-P63(tumor cell staining); Green-biomarker.

**Supplementary Files**

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementtables.V2.pdf
- suppfigures.pdf