

Article

Proteome-Wide Analysis Reveals TFEB Targets for Establishment of a Prognostic Signature to Predict Clinical Outcomes of Colorectal Cancer

Zijia Huang ^{1,2,†}, Sheng Zhu ^{3,†}, Ziqin Han ¹, Chen Li ¹, Junze Liang ², Yang Wang ² , Shuixing Zhang ^{1,*} and Jing Zhang ^{1,*} 

¹ Department of Radiology, The First Affiliated Hospital of Jinan University, Jinan University, Guangzhou 510613, China

² MOE Key Laboratory of Tumor Molecular Biology, Key Laboratory of Functional Protein Research of Guangdong Higher Education Institutes, Institute of Life and Health Engineering, College of Life Science and Technology, Jinan University, Guangzhou 510632, China

³ Department of Nuclear Medicine, Affiliated Hospital of Xiangnan University, Xiangnan University, Chenzhou 423000, China

* Correspondence: zsx7515@jnu.edu.cn (S.Z.); zj6410@jnu.edu.cn (J.Z.)

† These authors contributed equally to this work.

Simple Summary: Transcription factor EB (TFEB) is a master modulator of autophagy and lysosomal biogenesis. Dephosphorylation of TFEB at Ser142 and Ser138 determines its nuclear localization and transcriptional activity. The link between TFEB-associated genes and colorectal cancer (CRC) progression and prognosis remains unclear; thus, we performed data-independent acquisition (DIA)-based quantitative proteomics to systematically identify the targets of TFEB. Using stringent statistical criteria, 60 proteins associated with vesicular endocytic trafficking were identified as TFEB targets. Moreover, a prognosis-linked TFEB-related gene signature was developed, showing that patients with higher risk scores had higher epithelial–mesenchymal transition (EMT) scores with worse prognosis. Additionally, a nomogram was constructed by combining clinicopathological parameters and the gene signature to enhance the quantification capacity in risk assessment for individual patients. This research facilitates further mechanistic studies of TFEB, and the TFEB gene signature-based model may provide important information for assisting clinicians to predict CRC patient prognosis.

Abstract: Dephosphorylation of transcription factor EB (TFEB) at Ser142 and Ser138 determines its nuclear localization and transcriptional activity. The link between TFEB-associated genes and colorectal cancer (CRC) progression and prognosis remains unclear. To systematically identify the targets of TFEB, we performed data-independent acquisition (DIA)-based quantitative proteomics to compare global protein changes in wild-type (WT) DLD1 cells and TFEB^{WT}- or TFEB^{S142A/S138A} (activated status)-expressing DLD1 cells. A total of 6048 proteins were identified and quantified in three independent experiments. The differentially expressed proteins in TFEB^{S142A/S138A} versus TFEB^{WT} and TFEB^{WT} versus control groups were compared, and 60 proteins were identified as products of TFEB transcriptional regulation. These proteins were significantly associated with vesicular endocytic trafficking, the HIF-1 signaling pathway, and metabolic processes. Furthermore, we generated a TFEB-associated gene signature using a univariate and LASSO Cox regression model to screen robust prognostic markers. An eight-gene signature (*PLSCR3*, *SERPINA1*, *ATP6V1C2*, *TIMP1*, *SORT1*, *MAP2*, *KDM4B*, and *DDAH2*) was identified. According to the signature, patients were assigned to high-risk and low-risk groups. Higher risk scores meant worse overall survival and higher epithelial–mesenchymal transition (EMT) scores. Additionally, as per the clinicopathological parameters and gene signature, a nomogram was constructed that was utilized to enhance the quantification capacity in risk assessment for individual patients. This research shows that TFEB directly mediates network effects in CRC, and the identified TFEB gene signature-based model may provide important information for the clinical judgment of prognosis.



Citation: Huang, Z.; Zhu, S.; Han, Z.; Li, C.; Liang, J.; Wang, Y.; Zhang, S.; Zhang, J. Proteome-Wide Analysis Reveals TFEB Targets for Establishment of a Prognostic Signature to Predict Clinical Outcomes of Colorectal Cancer. *Cancers* **2023**, *15*, 744. <https://doi.org/10.3390/cancers15030744>

Academic Editor: Mark Molloy

Received: 23 November 2022

Revised: 15 January 2023

Accepted: 20 January 2023

Published: 25 January 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: TFEB; DIA-based proteomics; vesicular endocytic trafficking; colorectal cancer

1. Introduction

Colorectal cancer (CRC), a common malignant tumor in the digestive system, ranks third in terms of incidence and second regarding mortality, according to global cancer statistics [1]. The molecular structure of CRC is extremely heterogeneous and is therefore responsible for the heterogeneous and frequently suboptimal treatment response [2]. Thus, research has focused on molecular subtyping strategies based on single or multi-omics data to categorize patients into subgroups to facilitate risk stratification and disease management [3].

Dysregulation of transcription factors (TFs) leads to pronounced changes in gene expression, which is a common phenomenon of human malignant neoplasias [4,5]. These alterations influence cell proliferation/differentiation, metastasis, and migration, as well as chemotherapeutic resistance, thereby acting as major factors concerning the behavior of tumors [6]. Transcription factor EB (TFEB), which belongs to the MiT-TFE helix TFs, performs a vital function in lysosomal biogenesis, autophagy, metabolism, and the immune response [7]. Numerous reports have indicated that TFEB is upregulated in various types of tumors [7,8], and high levels of TFEB expression in CRC tissues are associated with aggressive clinical characteristics and poor survival of CRC patients [9]. Overexpression of TFEB and upregulation of its transcriptional network appears to be sufficient to drive tumorigenesis in different tissues; therefore, modulation of the activity of TFEB can be considered as a potential therapeutic strategy [10].

The TFEB shuttling between the nucleus and cytoplasm is influenced by its phosphorylation [11]. After being translocated to the nucleus, TFEB activates its target gene's transcription [12]. The nuclear export of TFEB is determined by the phosphorylation status of S142 and S138 near the nuclear export signal (NES). Dephosphorylation of S138/S142 when mTOR is inhibited results in TFEB nuclear retention [13]. The mutation of the aforementioned serines to alanines results in different TFEB types that appear to be localized in the nucleus and are constitutively active [12–14]. It is assumed that many reported targets of MiT members exert critical functions in tumorigenesis [7,15,16]; thus, identification of additional targets may be important for further insights regarding the functions of TFEB and its inhibition in cancer treatment.

In this study, to systematically identify the targets of TFEB, we produced stably wild-type TFEB (TFEB^{WT})- or activated-status TFEB (TFEB^{S142A/S138A})-expressing CRC cells and quantitatively analyzed the proteomes of these cells. Using stringent statistical criteria, 60 proteins were identified as TFEB targets. Moreover, a prognosis-linked gene signature (TFEB related) was developed utilizing the robust biomarkers identified through the combination of different methods. These insights will help further identify TFEB targets and elucidate the functions of TFEB.

2. Materials and Methods

2.1. Cell Lines and Cell Culture

Human DLD1 and 293T cells were cultured as per the respective depositors' recommendations after purchase from the American Type Culture Collection (Manassas, VA, USA).

2.2. Plasmid Construction, Retroviral Infection, and Transfection

TFEB^{WT}-GFP and mutant TFEB^{S142A/S138A}-GFP DNA sequences were amplified by PCR and were cloned into the pEGFP-N1 vectors, followed by subcloning into a pLVX-Puro lentiviral vector. Cells that overexpressed TFEB^{WT} or TFEB^{S142A/S138A} stably were generated as described previously [17]. Briefly, lipofectamine 3000 (Thermo Fisher Scientific, Waltham, MA, USA) was utilized to infect the cells by transfecting 293 T cells with virus skeleton vectors and lentiviral plasmids. CRC cells were infected using culture media

containing viruses and were selected using puromycin. Prior to further experimentation, the selected stably overexpressed wild-type or mutant TFEB cells were maintained for at least two weeks.

2.3. Immunoblotting

Immunoblotting assays were performed as previously described [18]. SDS-PAGE was performed on the extracted proteins, which were then transferred onto PVDF membranes (BIO-RAD, Hercules, CA, USA). The gel bands formed after exposure to appropriate antibodies (primary and secondary) were visualized using the ECL reagent (BIO-RAD) and imaged by Tanon 5200-Multi (Tanon Science & Technology, Shanghai, China). The primary antibodies (against TFEB, GFP, LAMP2, NDRG1, SMPD1, and actin) and the HRP-conjugated secondary antibodies (goat anti-rabbit antibody) were supplied by Proteintech (Wuhan, China). Original uncropped images for immunoblotting are presented in Figure S1.

2.4. Cell Proliferation Assays

For the proliferation assay, CRC cells stably expressing TFEB^{WT} or TFEB^{S142A/S138A} were seeded in 96-well microplates at a density of 1.5×10^3 cells per well and were cultured for four days. One day after seeding, cells were stained with CCK-8 (Beyotime, Jiangsu, China). The automated microplate spectrophotometer (BioTek Instruments, Winooski, VT, USA) was employed to read the plate after 1.5 h at the 450 nm wavelength. Absorbance was normalized against the absorbance on the first day and was calculated. Each experiment was performed in triplicate.

2.5. Confocal Examination

DLD1 cells stably expressing TFEB^{WT}-GFP or TFEB^{S142A/S138A}-GFP were plated on confocal dishes (NEST, Wuxi, China) and were incubated for 24 h. Prior to imaging, DAPI (Beyotime) was used as a stain to label the nucleus. TFEB^{WT} treated with Torin (TargetMol, Shanghai, China) was used as positive control. A confocal laser scanning microscope (LSM 880 with AiryScan, Carl Zeiss, Oberkochen, Germany) was utilized to visualize and record the GFP fluorescence.

2.6. Pinocytosis Assays

Pinocytosis was examined using a TRITC-dextran (70 kDa; Ruixi Biological Technology Co., Ltd., Xi'an, China) uptake assay. The incubation of the stably expressing TFEB^{WT}-GFP or TFEB^{S142A/S138A}-GFP DLD1 cells with 1 mg/mL TRITC-dextran was performed for 2 h at 37 °C, followed by thorough washing and processing for immunofluorescence assays as described above. The quantification of TRITC-dextran uptake was conducted through the "Analyze Particles" feature in Fiji software (ImageJ v1.4). Cells that absorbed TRITC-dextran were quantified with the aid of a flow cytometer (BD Biosciences, CA, USA) at 594 nm.

2.7. Mass Spectrometry (MS) Analysis

To assess global protein alteration, data-independent acquisition (DIA)-based proteomics was performed as per our prior study [19]. Briefly, cells were subjected to cell lysis buffer with subsequent trypsin digestion and desalination. Following the addition of iRT-Standard (Biognosys, Cambridge, MA, USA) to each sample, the Orbitrap Fusion Lumos mass spectrometer (Thermo Fisher Scientific) was employed for producing the DIA-MS raw data. The raw data were searched against the UniProt human protein database (<http://www.uniprot.org>, accessed on 20 January 2021). Among all mentioned alterations, the variable ones were considered to be the acetylation of lysine and the protein N-terminal as well as oxidation (M), whereas a fixed modification was considered to be carbamidomethylation (C). The peptide and protein identification false discovery rate cutoff values were set at 1%, and their expression levels were measured with the aid of the Spectronaut software v14 (Omicsson, Shanghai, China).

2.8. Analysis of Differentially Expressed TFEB Targets

The differentially expressed proteins (DEPs) in the TFEB^{S142A/S138A} and TFEB^{WT} groups (TFEB^{S142A/S138A}/TFEB^{WT}) and TFEB^{WT} and control groups (TFEB^{WT}/control) were obtained using the R software v4.0.5 “limma” package [20], and the criteria used to define DEPs were as follows: fold change (FC) > 1.2 and $p < 0.05$. TFEB targets were identified as FC (TFEB^{S142A/S138A}/TFEB^{WT}) > FC (TFEB^{WT}/control) > 1.2 and $p < 0.05$.

2.9. Gene Ontology Analysis

The enriched pathways of TFEB targets were identified through the R package “clusterProfiler” [21,22] utilizing GO (including CC: cellular component, BP: biological pathway, and MF: molecular function) and KEGG functional enrichment analyses. Utilizing the entire human genome as a reference, the functional annotation chart options set $p < 0.05$ for determining the significant biological pathways.

2.10. Publicly Available mRNA Data

The relevant clinical information and RNA-seq data on colon cancer (COAD) from The Cancer Genome Atlas (TCGA) database were acquired by using the online tool Sangerbox (<http://sangerbox.com/>, accessed on 6 April 2022) [23], and were included in the training group. Subsequently, an expression matrix file (GSE17536) via the Gene Expression Omnibus (GEO) (<http://www.ncbi.nlm.nih.gov/geo/>, accessed on 20 May 2022) was utilized for external verification.

2.11. Signature Generation

The “Survival” R package was employed for univariate Cox regression analysis to screen out the differentially expressed TFEB targets linked to prognosis, whereas further operations such as the Cox regression model of the least absolute shrinkage and selection operator (LASSO) were fit using the “Glmnet” R package to further identify the most robust candidates. The gene expression levels were normalized and calculated by the corresponding LASSO Cox coefficients, which were then used to establish the TFEB-related risk score (TRS) as follows:

$$\text{TRS} = \sum_i \text{Coefficient (mRNA}_i) \times \text{Expression (mRNA}_i)$$

2.12. Classification, Prediction, and Validation in GEO and TCGA

As per the formula of the signature, the patients were classified into two risk groups (high-risk and low-risk groups) depending on the TRS values. The threshold values and area under the curve were obtained and analyzed using the receiver operating characteristic (ROC), a time-dependent function. The aforementioned values were calculated for 1-, 3-, and 5-year overall survival (OS) and relapse-free survival to validate the performance of the signature using the “survivalROC” R package. The predictive significance of the risk score formula concerning prognosis was assessed using log-rank tests and Kaplan–Meier (K-M) survival curve analyses. Time-dependent concordance index (C-index) and time-dependent receiver operating characteristic (tROC) analyses were used to compare the predictive capacity of survival among different variables using the R packages “survConcordance” and “survivalROC”.

2.13. Gene Set Enrichment Analysis (GSEA)

GSEA Software V3.0 (<http://software.broadinstitute.org/gsea/index.jsp>, accessed on 18 August 2022) [24] and h.all.v7.4.symbols.gmt gene set (<http://software.broadinstitute.org/gsea/index.jsp>, accessed on 18 August 2022) [24] were used to investigate the associated pathways in different risk subgroups. Based on the gene expression profile in TCGA and phenotypic grouping, the minimum gene set was set as 5, and the maximum gene was set as 5000. After a thousand resamples, the functional annotation chart options set $p < 0.05$ for determining the significant hallmark gene sets.

2.14. Calculation of Epithelial–Mesenchymal Transition (EMT) Scores

The EMT scores were calculated as previously reported by Milena P. Mak et al. [25], using the pan-cancer EMT signature (Table S2 [25]). For each CRC patient, the EMT score was calculated as the mean expression of the mesenchymal markers minus the mean expression of the epithelial markers.

3. Results

3.1. Constitutive Nuclear Import of TFEB Promotes CRC Proliferation

Dephosphorylation of the two highly evolutionarily conserved serines S138 and S142, which are situated near a NES (Figure 1A,B), is critical for its nuclear retention and activation. To examine the role of activated TFEB in CRC, a series of constructs was generated in which the GFP tag was fused to TFEB^{WT} or TFEB^{S138A/S142A}. Substantial expression of the TFEB-GFP fusion protein was observed in TFEB^{WT}- and TFEB^{S138A/S142A}-transfected cells through Western blotting using anti-TFEB and anti-GFP antibodies (Figure 1C). Confocal microscopy showed that mutations in S138 and S142 caused almost complete nuclear retention of TFEB (Figure 1D,E), as compared to TFEB^{WT}, similar to the mTOR inhibitor Torin. The CCK-8 assay showed that CRC cells overexpressing both WT and mutant TFEB displayed a pro-proliferation phenotype, as compared to the control, while TFEB^{S138A/S142A} displayed a higher proliferation ability than TFEB^{WT} cells (Figure 1F), suggesting an oncogenic role of activated TFEB in CRC.

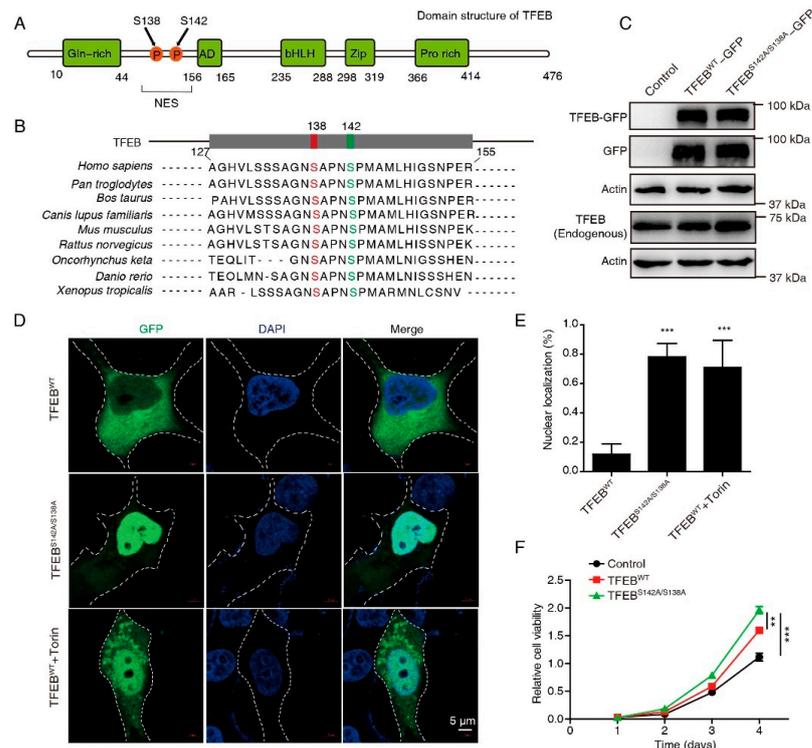


Figure 1. Mutations in S138 and S142 caused almost complete nuclear retention of TFEB. (A) Schematic of the human TFEB protein domains. S138 and S142 are localized in the proximity of a NES. (B) S138 and S142 of TFEB are evolutionarily conserved in the indicated species. Alignment of the sequences near TFEB S138 and S142 is shown. (C) Western blots of endogenous and exogenous TFEB in CRC cells transfected with wild-type TFEB-GFP or TFEB-GFP mutant (S142A/S138A). Uncropped immunoblots are provided in the Figure S1. (D,E) Localization of TFEB^{WT} and TFEB^{S142A/S138A} in the cytosol and in the nucleus was detected by confocal microscopy. Representative images (D) and statistical results (E) are shown. Torin1 (250 nm), a positive control. Cell outlines in white dotted lines. DAPI was used to label the nucleus. Scale bars, 5 μm. Mean ± SD, n = 10 cells per condition, unpaired *t*-test. (F) Cell viability of indicated cells were analyzed using a CCK8 assay. Mean ± SEM, n = 3, unpaired *t*-test. ** *p* < 0.01, *** *p* < 0.001.

3.2. DIA-Based Proteomics Profiling Reveals Potential TFEB Targets

To systematically identify downstream proteins regulated by TFEB, the proteome of DLD1-TFEB^{WT}, -TFEB^{S142A/S138A}, and control cell lines were analyzed through DIA-based proteomics (Figure 2A). Proteins from these lines were profiled in triplicate, and 6048 proteins (Table S1) were identified and quantified at 1% FDR. Ridge plots were employed to show the expression distribution of these proteins in different groups (Figure 2B), and no significant difference was observed among each set of triplicates; however, the expression distribution differed between the three groups ($p = 9.7 \times 10^{-4}$). To examine the difference between the proteins measured in the three groups and to validate the reproducibility of the triplicate measurements in each experiment, a PCA was performed. This revealed that the samples could be divided into TFEB^{WT}, TFEB^{S142A/S138A}, and control groups, thus confirming the similarity of each set of triplicate measurements (Figure 2C). We analyzed DEPs in TFEB^{S142A/S138A} versus TFEB^{WT} and TFEB^{WT} versus control groups through volcano plots and obtained a list of 480 upregulated and 964 downregulated proteins in the TFEB^{S142A/S138A}/TFEB^{WT} groups and 982 and 117 in the TFEB^{WT}/control groups (>1.2 -fold change, $p < 0.05$; Figure 2D, Table S2). To better identify and recognize targets transcriptionally regulated by TFEB translocation to the nucleus without being affected by TFEB overexpression, we selected proteins that were upregulated in the TFEB^{WT}/control groups and remained upregulated when the proteome of TFEB^{S142A/S138A} cells was compared with that of TFEB^{WT} cells (Figure 2E). After filtering, 60 proteins showed gradually increased expression levels, in the order of control, TFEB^{WT}, and TFEB^{S142A/S138A} (Figure 2F, Table S3). Subsequently, three randomly selected proteins (LAMP2, NDRG1, SMPD1) from the 60 proteins were verified by immunoblotting and qRT-PCR. The results showing the quantitative trend of the three DEPs in MS were consistent with the immunoblotting and qRT-PCR results (Figure S2A,B). The high accuracy of our MS result provided a basis for the subsequent analysis of the 60 TFEB-regulated proteins.

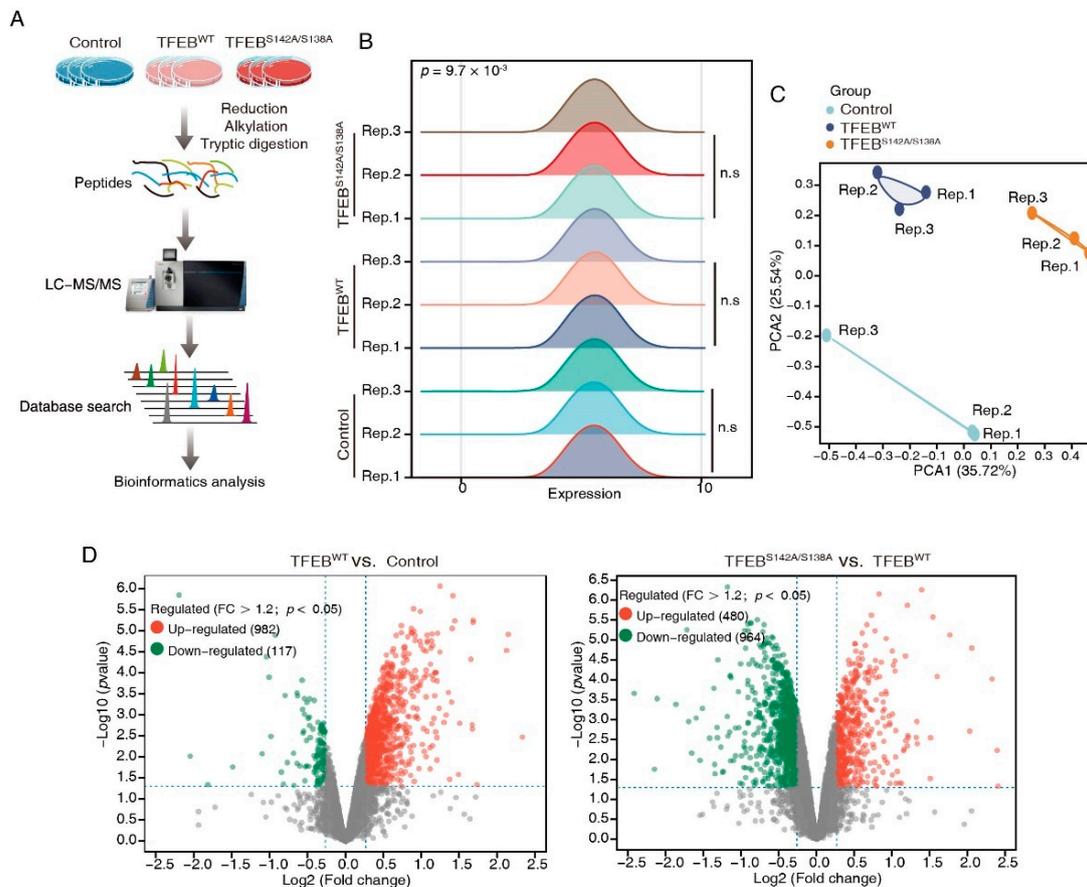


Figure 2. Cont.

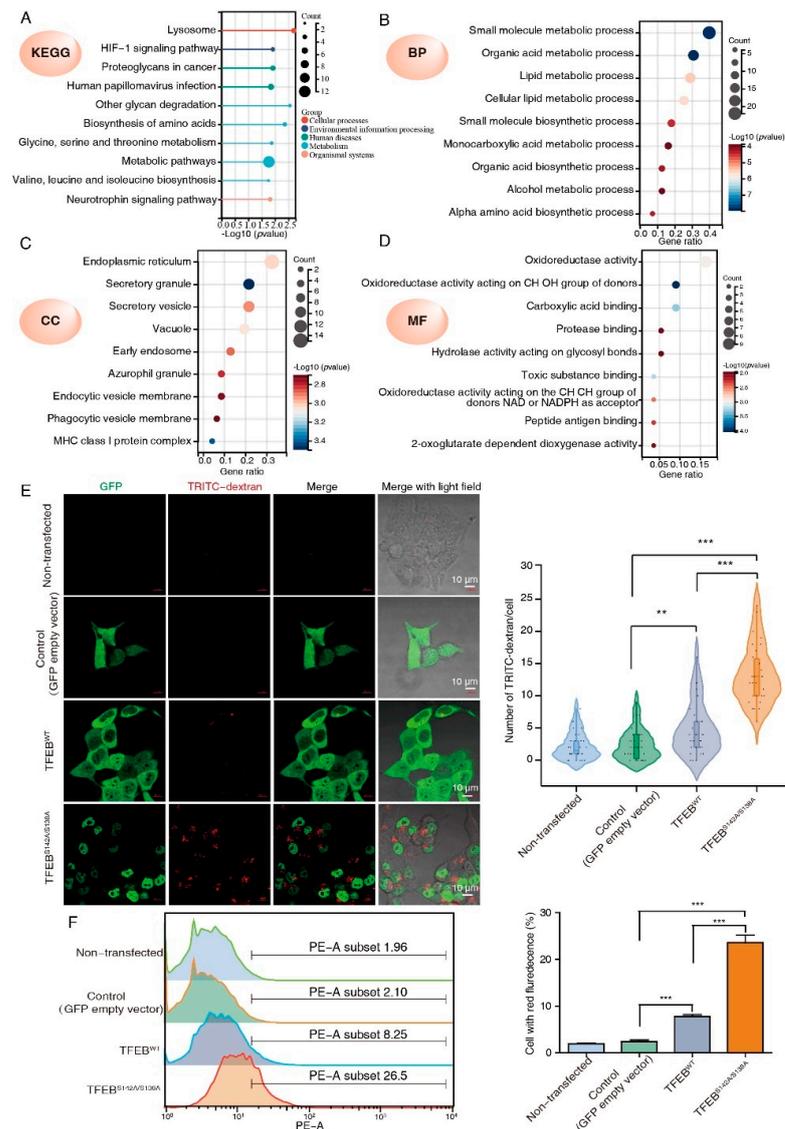


Figure 3. GO classification and pathway analysis of the TFEB transcriptional regulated proteins. (A) KEGG analysis of the 60 TFEB transcriptional regulated proteins. (B–D) The same proteins were subjected to the GO analysis (BP, biological process; CC, cellular component; MF, molecular function). (E,F) DLD1 cells stably expressing TFEB^{WT} or TFEB^{S142A/S138A} were incubated with TRITC-dextran (0.2 mg/mL) in a growth medium at 37 °C for 2 h. The representative confocal images show internalized TRITC-dextran (red dots) in indicated cells. Scale bars, 10 μm. Quantification is indicated on the right. Mean ± SD, $n = 30$ cells per condition, unpaired t -test. (E) Pinocytosis quantification using TRITC-dextran in indicated cells by flow cytometry. (F) Mean ± SEM, $n = 3$, unpaired t -test. ** $p < 0.001$, *** $p < 0.001$.

3.4. Establishment of a TFEB Target Signature for Prognosis

The 60 TFEB targets were analyzed through univariate Cox regression, and promising candidates with a cutoff value of the log-rank test of $p < 0.05$ were utilized for robust prognostic marker detection by inclusion in the LASSO Cox regression model. Overfitting was eliminated by tenfold cross-validation, with an optimal λ value of 0.007 (Figure 4A). A set of eight genes (*PLSCR3*, *SERPINA1*, *ATP6V1C2*, *TIMP1*, *SORT1*, *MAP2*, *KDM4B*, and *DDAH2*) as well as their individual LASSO coefficients were generated (Figure 4B). Their corresponding hazard ratios and LASSO coefficient distributions are listed in Figure 4C,D. As a result, the TRS formula: $TRS = \sum_i \text{Coefficient}(\text{mRNA}_i) \times \text{Expression}(\text{mRNA}_i)$ was generated.

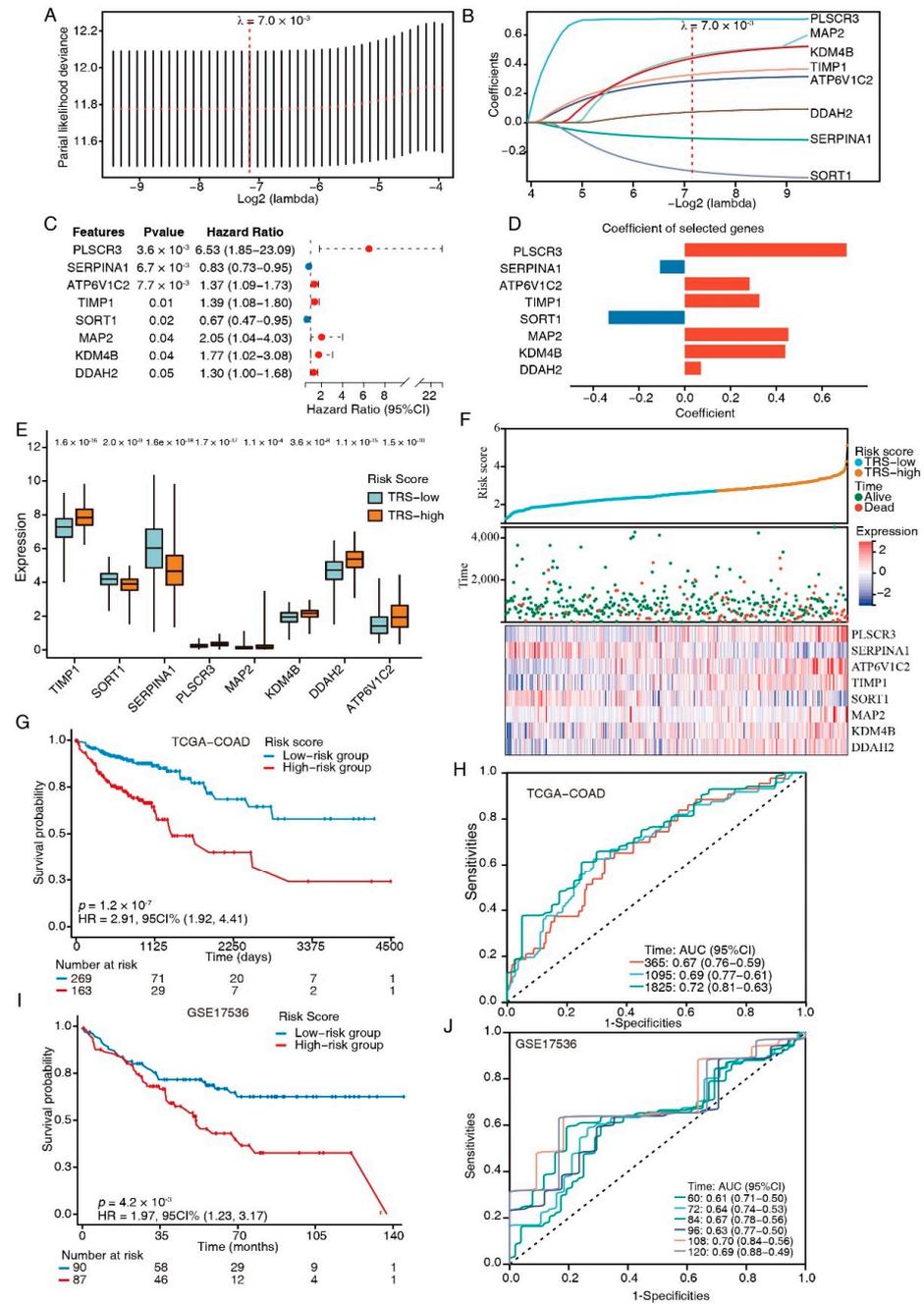


Figure 4. TFEB-related gene signature establishment. (A) Selection of the optimal LASSO parameter lambda; vertical lines indicate the optimal values. (B) LASSO coefficient profiles of the 8 candidate genes. (C) Forest plots of the 8-gene signature. (D) Coefficients of the eight genes. (E) Boxplots showing the 8-gene signature expression levels between high- and low-risk-score groups. (F) The distributions of the risk score and survival status of CRC patients, and the heatmap showing the eight genes' expression profiles. (G) K–M survival curves for patients at high and low risk in the TCGA-COAD cohort. (H) ROC plot and AUC scores for predicted OS by the TRS in the TCGA-COAD cohort. (I) Kaplan–Meier survival curves of OS between high- and low-risk patients in the GSE17536 dataset. (J) ROC plot and AUC scores for predicted OS by the TRS in the GSE17536 dataset.

We next divided the CRC patients from the training set into high- and low-risk groups according to their TRS. The expression levels of risk-type genes (*PLSCR3*, *ATP6V1C2*, *TIMP1*, *MAP2*, *KDM4B*, and *DDAH2*) were elevated in the high-risk category in contrast with the low-risk category, whereas the expression levels of protective-type genes

(*SERPINA1* and *SORT1*) had the opposite trend (Figure 4E). Ranked by TRS, the CRC patients were examined regarding the link between survival times and survival status (Figure 4F). The eight genes' expression patterns were represented by a heatmap, which demonstrated the predictive ability of the eight-gene signature in the prognostic values in patients as well as their potential influence on the onset and progression of malignancies.

A higher TRS was associated with a poorer prognosis for patients than lower scores, according to a K-M analysis ($HR = 2.91$, $p = 1.2 \times 10^{-7}$; Figure 4G). The prognostic model's predictive ability was tested using the ROC curves; the AUCs of these curves were estimated at around 0.7, indicating good predictive performance (Figure 4H). The same model was utilized to verify the eight-gene signature in the testing dataset (GSE17536, $n = 177$). The K-M survival curves depicted increased survival rates for the low-risk category (Figure 4I), which was congruent with the training group data. The ROC curves depicted diverse AUC values ranging from 0.61 to 0.7 (Figure 4J). Collectively, these results suggested that the models based on the eight-gene signature had high specificity and sensitivity.

3.5. Tumor-Biology-Associated Potential Functions of the TFEB Signature

To gain further insights into the features of the TFEB signature, the link between the cancer hallmark gene sets and TRS was examined (Figure 5A). The group with high risk showed enrichment of the EMT, angiogenesis, myogenesis, hedgehog signaling, apical junction, notch signaling, and coagulation as per the GSEA data, whereas protein secretion was reduced. Among these enriched pathways, EMT produced the highest enrichment score; therefore, EMT scores for the CRC patients were subsequently calculated. Compared to normal tissues, cancer tissues had increased EMT scores (Figure 5B), and the high-risk group had higher EMT scores (Figure 5C). Furthermore, the EMT score positively correlated with the TRS (Pearson's $R = 0.41$, $p = 4.3 \times 10^{-19}$, Figure 5D). These results indicated that EMT made pronounced contributions to the prognosis of COAD patients with a higher TRS.

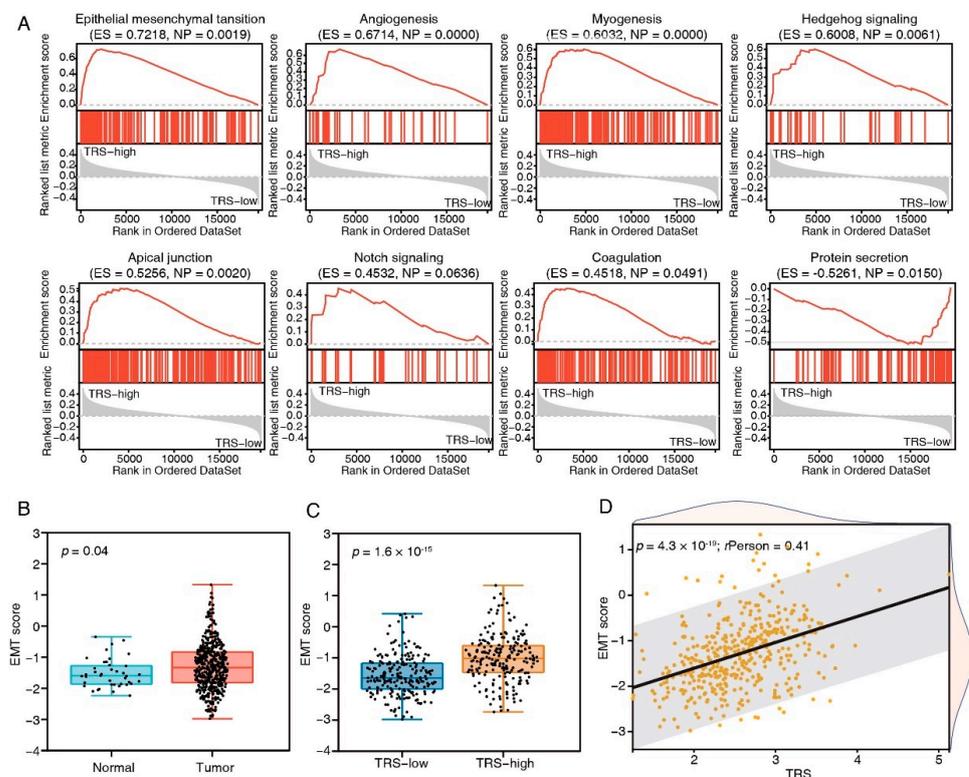


Figure 5. Biological pathways of TFEB score subtypes. (A) GSEA enrichment plots showing the 8 most enriched cancer hallmark pathways ($p < 0.05$). (B) The comparison of EMT scores in tumor and normal tissues in the TCGA-COAD cohort. (C) The comparison of EMT scores in high- and low-risk groups in the TCGA-COAD cohort. (D) The correlation analysis between EMT scores and risk scores.

3.6. Combination of the TRS and Clinicopathological Characteristics Improves Survival Prediction

To compare the TRS with the crucial clinical factors, we analyzed the TRS differences between subgroups with distinct clinicopathological features. In the TCGA dataset, T3–4, stage III–IV, and N1–2 patients had an increased TRS (Figure 6A); similarly, in the testing dataset GSE17536, the TRS was increased in stage 3–4 and grade 3 (Figure S3). To enhance the quantification capacity in risk assessment, we constructed a nomogram using the TRS and six clinical features including gender; age; clinical N, M, and T; and stage (Figure 6B). The actual 3-year survival times were congruent with the anticipated survival times, as per the calibration plots (Figure 6C). The K–M survival analyses indicated significant OS in the risky group (Figure 6D). The prediction capacity of the nomogram was assessed through tROC analyses, with an average AUC above 0.8, which was markedly better than pathological TNM staging. (Figure 6E). Furthermore, the nomogram also showed a robust performance in the prediction of survival rates in the testing dataset (Figure S4A,B). These results suggested that the TRS-based nomogram had a stable and powerful performance in the prediction of survival rates.

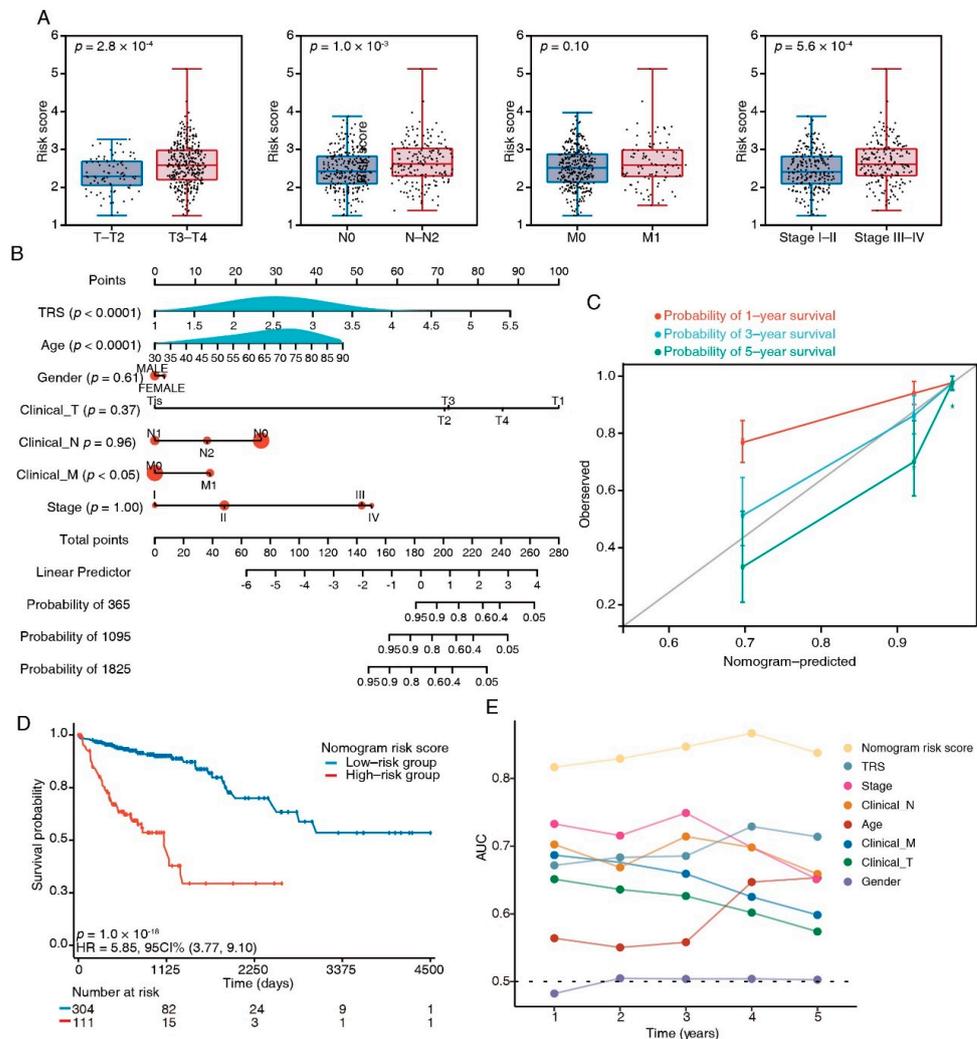


Figure 6. Nomograms showing the results of prognostic models using the TFEB signature and TNM staging system to predict OS of patients with CRC. (A) Distribution of TRS in clinicopathological variables, e.g., tumor (T), lymph node status (N), metastasis (M), and clinical stage. (B) Nomogram based on risk score and clinical factors. (C) Calibration plots of the nomogram for predicting the probability of 1-, 3-, and 5-year survival, ● represents the mean and * represents the media. (D) Kaplan–Meier curves of OS between high-risk and low-risk patients in the TCGA-COAD cohort. (E) tROC analysis to assess the accuracy of the nomogram.

4. Discussion

Aberrant regulation of MiT/TEF transcription factors plays a critical role in tumorigenesis [7–10,12]; however, the network directly mediated by TFEB in CRC remains obscure. In the current study, we produced TFEB^{S138A/S142A}, a continuously activated form, to systematically investigate direct targets of TFEB by performing in-depth DIA-MS-based quantitative proteomics. A total of 60 TFEB-regulated DEPs were identified, and a univariate and LASSO Cox regression model was performed to select robust prognostic biomarkers for the creation of a TFEB-related gene signature. Additionally, the accuracy and prediction capability were enhanced through the construction of a nomogram model according to the clinicopathological features and TFEB-related gene signatures.

As a master transcriptional regulator of cell catabolism, TFEB is translocated from the cytoplasm to the nucleus to control gene expression that is relevant to lysosomal and autophagosomal biogenesis [12,26]. Our results consistently showed that TFEB carried Ser-to-Ala mutations of both Ser138 and Ser142 and showed constitutive nuclear localization, which significantly promotes CRC growth. The activation of TFEB influences various downstream cellular processes that are vital for phagocyte functions, such as lysosome functioning and autophagy, endocytosis, phagocytosis, endoplasmic reticulum restoration, mitochondrial homeostasis, lipid and glucose metabolism, and the production of cytokine and chemokine [27]. In congruence with these results, we found that these proteins were localized in the endoplasmic reticulum and structures associated with secretion and endocytosis, and they were involved in the regulation of lysosome processes, metabolic processes, HIF-1 signaling, oxidoreductase activity, and several binding functions, indicating that the functions of TFEB are not limited to autophagy and lysosome biogenesis.

Vesicular endocytic trafficking is essential in regulating tumor metastasis and growth. As per the type of cargo, internalization route, and scission mechanism, vesicular endocytic trafficking occurs in three general types in the cell: pinocytosis, phagocytosis, and endocytosis [28]. TFEB induces cellular endocytosis and phagocytosis to mediate lysosomal biogenesis, autophagy flux, and MTORC1 signaling [27,29]. In this research, the cellular components of TFEB targets were linked to vesicle membranes during phagocytosis and endocytosis, and an unidentified TFEB function of enhancing the pinocytosis pathway was detected (Figure 3E,F), indicating that TFEB is involved in regulating all three general modes of vesicular endocytic trafficking. These findings may help explain why active TFEB can promote CRC cell proliferation (Figure 1E) and pronounced TRS group enrichment in the carcinogenic activation pathways (Figure 5).

Based on the identified TFEB-associated genes, we established a prognostic CRC signature with a combination of eight genes (*PLSCR3*, *SERPINA1*, *ATP6V1C2*, *TIMP1*, *SORT1*, *MAP2*, *KDM4B*, and *DDAH2*) and an equation for calculating a TRS (Figure 4). The poorer prognosis observed in the group with higher TRS values through survival analysis was confirmed using independent cohorts. Additionally, GSEA analysis revealed considerably increased enrichment levels concerning the carcinogenic activation pathways in the group with higher TRS values, i.e., EMT, angiogenesis, myogenesis, hedgehog signaling, apical junction, notch signaling, and coagulation. These findings provided a more thorough and credible reason for the worse prognosis in patients belonging to the high-TRS subgroup. Additionally, a nomogram was generated containing the TRS and other clinicopathological parameters. In comparison with various characteristics, a stable and powerful performance was shown by the nomogram regarding the prediction of survival rates during follow-up at different times (Figure 6). The resulting data indicated the high potential of the TRS for CRC patient prognosis, and the capability of the nomogram as a reliable method for diagnosing CRC patients in clinical settings.

Six of the eight genes served as risky types in our study, and most of them have been studied in many cancers. For example, involved in both tumor progression and metastasis, *ATP6V1C2* functions in the biological process of transferring hydrogen ions [30]. *TIMP1*, termed tissue inhibitor matrix metalloproteinase 1, has been widely studied in various cancers. In colorectal cancer, *TIMP1* promotes metastasis through the FAK-PI3K/AKT and

MAPK pathways [31]. *MAP2*, a crucial regulator of microtubule dynamics, exerts dual effects on cancer progression [32]. Our research revealed that *MAP2* correlated with a poor prognosis of CRC and served as a risk biomarker. *KDM4B* promotes the progression of colorectal cancer and glucose metabolism by promoting TRAF6-mediated AKT activation [33]. *DDAH2* was reported to be involved in the invasion of lung adenocarcinoma by promoting angiogenesis [34]. The expression of *SERPINA1* and *SORT1* has previously been shown to be increased in several types of cancers [35,36]; however, the two proteins served as two protective biomarkers in the current study. Notably, *PLSCR3*, with the highest risk coefficient in our model, has rarely been investigated in tumor research. Therefore, more research on CRC is needed to fully understand the molecular processes related to the TFEB of the novel gene signature.

5. Conclusions

Taken together, we performed a proteome-wide analysis of potential TFEB targets, and our insights facilitate further mechanistic studies of TFEB. Furthermore, a novel TFEB-regulated gene signature was established to identify CRC patients with increased risk. By incorporating this with clinicopathological characteristics, we assembled a nomogram to quantitatively assess individual patients' risk scores. The TFEB gene signature-based model may prove to be efficient in assisting the clinician to predict patient prognosis, thereby leading to the effective management of CRC on an individual basis.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/cancers15030744/s1>, Figure S1: Uncropped Western Blots for Figures 1 and S2A, Figure S2: Verification of randomly selected DEPs, Figure S3: Distribution of TRS in stage and grade, Figure S4: The external validation of the nomogram, Table S1: The quantifiable protein list in TFEB^{WT}, TFEB^{S142A/S138A}, and control datasets, Table S2: The differentially expressed proteins in TFEB^{S142A/S138A} vs. TFEB^{WT} and TFEB^{WT} vs. control datasets, Table S3: The protein list of TFEB targets (fold change: TFEB^{S142A/S138A} vs. TFEB^{WT} > TFEB^{WT} vs. control > 1.2; $p < 0.05$).

Author Contributions: Conceptualization, J.Z., Y.W. and S.Z. (Shuixing Zhang); methodology, Z.H. (Zijia Huang) and J.Z.; software, J.Z.; validation, Z.H. (Zijia Huang), S.Z. (Sheng Zhu) and Z.H. (Ziqin Han); formal analysis, J.Z.; investigation, S.Z. (Sheng Zhu) and C.L.; resources, J.Z.; data curation, Z.H. (Zijia Huang) and J.Z.; writing—original draft preparation, J.Z.; writing—review and editing, Y.W. and S.Z. (Shuixing Zhang); visualization, J.Z., Z.H. (Zijia Huang) and J.L.; supervision, S.Z. (Shuixing Zhang); project administration, S.Z. (Shuixing Zhang); funding acquisition, S.Z. (Shuixing Zhang), J.Z. and S.Z. (Sheng Zhu). All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (grant number 82227802), the Basic Research Funds for Central universities of Jinan University (grant number 21622110) and Clinical Medical Technology Innovation Guidance Project of Hunan Province (grant number 22020SK52202).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data are available from the corresponding authors upon request.

Acknowledgments: The authors thank Xing-Feng Yin, Jinan University and Zheng-Hua Sun, Jinan University for their technical assistance with mass spectrometry. The authors thank Li-Ye Zhong, Jinan University for her technical assistance with stable cell line construction.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J. Clin.* **2021**, *71*, 209–249. [[CrossRef](#)]
2. Chen, B.; Scurrah, C.R.; McKinley, E.T.; Simmons, A.J.; Ramirez-Solano, M.A.; Zhu, X.; Markham, N.O.; Heiser, C.N.; Vega, P.N.; Rolong, A.; et al. Differential pre-malignant programs and microenvironment chart distinct paths to malignancy in human colorectal polyps. *Cell* **2021**, *184*, 6262–6280. [[CrossRef](#)]
3. Salvucci, M.; Crawford, N.; Stott, K.; Bullman, S.; Longley, D.B.; Prehn, J.H.M. Patients with mesenchymal tumours and high Fusobacteriales prevalence have worse prognosis in colorectal cancer (CRC). *Gut* **2022**, *71*, 1600–1612. [[CrossRef](#)]
4. Bhagwat, A.S.; Vakoc, C.R. Targeting Transcription Factors in Cancer. *Trends Cancer* **2015**, *1*, 53–65. [[CrossRef](#)]
5. Lee, T.I.; Young, R.A. Transcriptional regulation and its misregulation in disease. *Cell* **2013**, *152*, 1237–1251. [[CrossRef](#)]
6. Laissue, P. The forkhead-box family of transcription factors: Key molecular players in colorectal cancer pathogenesis. *Mol. Cancer* **2019**, *18*, 5. [[CrossRef](#)]
7. Napolitano, G.; Ballabio, A. TFEB at a glance. *J. Cell Sci.* **2016**, *129*, 2475–2481. [[CrossRef](#)]
8. Perera, R.M.; Stoykova, S.; Nicolay, B.N.; Ross, K.N.; Fitamant, J.; Boukhali, M.; Lengrand, J.; Deshpande, V.; Selig, M.K.; Ferrone, C.R.; et al. Transcriptional control of autophagy-lysosome function drives pancreatic cancer metabolism. *Nature* **2015**, *524*, 361–365. [[CrossRef](#)]
9. Liang, J.; Jia, X.; Wang, K.; Zhao, N. High expression of TFEB is associated with aggressive clinical features in colorectal cancer. *OncoTargets Ther.* **2018**, *11*, 8089–8098. [[CrossRef](#)]
10. Li, Y.; Hodge, J.; Liu, Q.; Wang, J.; Wang, Y.; Evans, T.D.; Altomare, D.; Yao, Y.; Murphy, E.A.; Razani, B.; et al. TFEB is a master regulator of tumor-associated macrophages in breast cancer. *J. Immunother. Cancer* **2020**, *8*, e000543. [[CrossRef](#)]
11. Puertollano, R.; Ferguson, S.M.; Brugarolas, J.; Ballabio, A. The complex relationship between TFEB transcription factor phosphorylation and subcellular localization. *EMBO J.* **2018**, *37*, e98804. [[CrossRef](#)] [[PubMed](#)]
12. Settembre, C.; Di Malta, C.; Polito, V.A.; Garcia Arencibia, M.; Vetrini, F.; Erdin, S.; Erdin, S.U.; Huynh, T.; Medina, D.; Colella, P.; et al. TFEB links autophagy to lysosomal biogenesis. *Science* **2011**, *332*, 1429–1433. [[CrossRef](#)]
13. Napolitano, G.; Esposito, A.; Choi, H.; Matarese, M.; Benedetti, V.; Di Malta, C.; Monfregola, J.; Medina, D.L.; Lippincott-Schwartz, J.; Ballabio, A. mTOR-dependent phosphorylation controls TFEB nuclear export. *Nat. Commun.* **2018**, *9*, 3312. [[CrossRef](#)] [[PubMed](#)]
14. Li, Y.; Xu, M.; Ding, X.; Yan, C.; Song, Z.; Chen, L.; Huang, X.; Wang, X.; Jian, Y.; Tang, G.; et al. Protein kinase C controls lysosome biogenesis independently of mTORC1. *Nat. Cell Biol.* **2016**, *18*, 1065–1077. [[CrossRef](#)] [[PubMed](#)]
15. Yang, M.; Tang, B.; Wang, S.; Tang, L.; Wen, D.; Vlodaysky, I.; Yang, S.M. Non-enzymatic heparanase enhances gastric tumor proliferation via TFEB-dependent autophagy. *Oncogenesis* **2022**, *11*, 49. [[CrossRef](#)]
16. Li, C.; Wang, L. TFEB-dependent autophagy is involved in scavenger receptor OLR1/LOX-1-mediated tumor progression. *Autophagy* **2022**, *18*, 462–464. [[CrossRef](#)]
17. Wang, Y.; Zhang, J.; Zheng, C.C.; Huang, Z.J.; Zhang, W.X.; Long, Y.L.; Gao, G.B.; Sun, Y.; Xu, W.W.; Li, B.; et al. C20orf24 promotes colorectal cancer progression by recruiting Rin1 to activate Rab5-mediated mitogen-activated protein kinase/extracellular signal-regulated kinase signalling. *Clin. Transl. Med.* **2022**, *12*, e796. [[CrossRef](#)]
18. Zhang, J.; Zhou, Y.; Li, N.; Liu, W.T.; Liang, J.Z.; Sun, Y.; Zhang, W.X.; Fang, R.D.; Huang, S.L.; Sun, Z.H.; et al. Curcumin Overcomes TRAIL Resistance of Non-Small Cell Lung Cancer by Targeting NRH:Quinone Oxidoreductase 2 (NQO2). *Adv. Sci.* **2020**, *7*, 2002306. [[CrossRef](#)]
19. Sun, Y.; Yang, Y.M.; Hu, Y.Y.; Ouyang, L.; Sun, Z.H.; Yin, X.F.; Li, N.; He, Q.Y.; Wang, Y. Inhibition of nuclear deacetylase Sirtuin-1 induces mitochondrial acetylation and calcium overload leading to cell death. *Redox Biol.* **2022**, *53*, 102334. [[CrossRef](#)]
20. Ritchie, M.E.; Phipson, B.; Wu, D.; Hu, Y.; Law, C.W.; Shi, W.; Smyth, G.K. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **2015**, *43*, e47. [[CrossRef](#)]
21. Yu, G.; Wang, L.G.; Han, Y.; He, Q.Y. clusterProfiler: An R package for comparing biological themes among gene clusters. *OMICS* **2012**, *16*, 284–287. [[CrossRef](#)] [[PubMed](#)]
22. Wu, T.; Hu, E.; Xu, S.; Chen, M.; Guo, P.; Dai, Z.; Feng, T.; Zhou, L.; Tang, W.; Zhan, L.; et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation* **2021**, *2*, 100141. [[CrossRef](#)] [[PubMed](#)]
23. Shen, W.; Song, Z.; Zhong, X.; Huang, M.; Shen, D.; Gao, P.; Qian, X.; Wang, M.; He, X.; Wang, T.; et al. Sangerbox: A comprehensive, interaction-friendly clinical bioinformatics analysis platform. *iMeta* **2022**, *1*, e36. [[CrossRef](#)]
24. Subramanian, A.; Tamayo, P.; Mootha, V.K.; Mukherjee, S.; Ebert, B.L.; Gillette, M.A.; Paulovich, A.; Pomeroy, S.L.; Golub, T.R.; Lander, E.S.; et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 15545–15550. [[CrossRef](#)] [[PubMed](#)]
25. Mak, M.P.; Tong, P.; Diao, L.; Cardnell, R.J.; Gibbons, D.L.; William, W.N.; Skoulidis, F.; Parra, E.R.; Rodriguez-Canales, J.; Wistuba, I.I.; et al. A Patient-Derived, Pan-Cancer EMT Signature Identifies Global Molecular Alterations and Immune Target Enrichment Following Epithelial-to-Mesenchymal Transition. *Clin. Cancer Res.* **2016**, *22*, 609–620. [[CrossRef](#)]
26. Saftig, P.; Haas, A. Turn up the lysosome. *Nat. Cell Biol.* **2016**, *18*, 1025–1027. [[CrossRef](#)] [[PubMed](#)]
27. Irazoqui, J.E. Key Roles of Mit Transcription Factors in Innate Immunity and Inflammation. *Trends Immunol.* **2020**, *41*, 157–171. [[CrossRef](#)] [[PubMed](#)]
28. Khan, I.; Steeg, P.S. Endocytosis: A pivotal pathway for regulating metastasis. *Br. J. Cancer* **2021**, *124*, 66–75. [[CrossRef](#)]

29. Nnah, I.C.; Wang, B.; Saqcena, C.; Weber, G.F.; Bonder, E.M.; Bagley, D.; De Cegli, R.; Napolitano, G.; Medina, D.L.; Ballabio, A.; et al. TFEB-driven endocytosis coordinates MTORC1 signaling and autophagy. *Autophagy* **2019**, *15*, 151–164. [[CrossRef](#)]
30. Li, G.; Huang, J.; Chen, S.; He, Y.; Wang, Z.; Peng, J. High Expression of ATP6V1C2 Predicts Unfavorable Overall Survival in Patients With Colon Adenocarcinoma. *Front. Genet.* **2022**, *13*, 930876. [[CrossRef](#)]
31. Song, G.; Xu, S.; Zhang, H.; Wang, Y.; Xiao, C.; Jiang, T.; Wu, L.; Zhang, T.; Sun, X.; Zhong, L.; et al. TIMP1 is a prognostic marker for the progression and metastasis of colon cancer through FAK-PI3K/AKT and MAPK pathway. *J. Exp. Clin. Cancer Res.* **2016**, *35*, 148. [[CrossRef](#)] [[PubMed](#)]
32. Yi, R.; Feng, J.; Yang, S.; Huang, X.; Liao, Y.; Hu, Z.; Luo, M. miR-484/MAP2/c-Myc-positive regulatory loop in glioma promotes tumor-initiating properties through ERK1/2 signaling. *J. Mol. Histol.* **2018**, *49*, 209–218. [[CrossRef](#)] [[PubMed](#)]
33. Li, H.; Lan, J.; Wang, G.; Guo, K.; Han, C.; Li, X.; Hu, J.; Cao, Z.; Luo, X. KDM4B facilitates colorectal cancer growth and glucose metabolism by stimulating TRAF6-mediated AKT activation. *J. Exp. Clin. Cancer Res.* **2020**, *39*, 12. [[CrossRef](#)] [[PubMed](#)]
34. Shiozawa, T.; Iyama, S.; Toshima, S.; Sakata, A.; Usui, S.; Minami, Y.; Sato, Y.; Hizawa, N.; Noguchi, M. Dimethylarginine dimethylaminohydrolase 2 promotes tumor angiogenesis in lung adenocarcinoma. *Virchows Arch.* **2016**, *468*, 179–190. [[CrossRef](#)] [[PubMed](#)]
35. Berger, K.; Rhost, S.; Rafnsdottir, S.; Hughes, E.; Magnusson, Y.; Ekholm, M.; Stal, O.; Ryden, L.; Landberg, G. Tumor co-expression of progranulin and sortilin as a prognostic biomarker in breast cancer. *BMC Cancer* **2021**, *21*, 185. [[CrossRef](#)] [[PubMed](#)]
36. Wu, C.C.; Lu, Y.T.; Yeh, T.S.; Chan, Y.H.; Dash, S.; Yu, J.S. Identification of Fucosylated SERPINA1 as a Novel Plasma Marker for Pancreatic Cancer Using Lectin Affinity Capture Coupled with iTRAQ-Based Quantitative Glycoproteomics. *Int. J. Mol. Sci.* **2021**, *22*, 6079. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.