Research Article

Identification of Molecular Subtypes and Potential Small-Molecule Drugs for Esophagus Cancer Treatment Based on m⁶A Regulators

Jianjun Li, Hongbo Zhu, Qiao Yang, Hua Xiao, Haibiao Wu, Zhe Fang, Wenjun Li, and Manbo Cai [iD]

Department of Oncology Radiotherapy, The First Affiliated Hospital, Hengyang Medical School, University of South China, No. 69 Chuanshan Road, Hengyang, Hunan, China

Correspondence should be addressed to Manbo Cai; caimanbo@nhfyyy.com

Background. Esophagus cancer (ESCA) is the sixth most frequent cancer in males, with 5-year overall survival of 15%–25%. RNA modifications function critically in cancer progression, and m⁶A regulators are associated with ESCA prognosis. This study further revealed correlations between m⁶A and ESCA development. *Methods*. Univariate Cox regression analysis and consensus clustering were applied to determine molecular subtypes. Functional pathways and gene ontology terms were enriched by gene set enrichment analysis. Protein-protein interaction (PPI) analysis on differentially expressed genes (DEGs) was conducted for hub gene screening. Public drug databases were employed to study the interactions between hub genes and small molecules. *Results*. Three molecular subtypes related to ESCA prognosis were determined. Based on multiple analyses among molecular subtypes, 146 DEGs were screened, and a PPT network of 15 hub genes was visualized. Finally, 8 potential small-molecule drugs (BMS-754807, gefitinib, neratinib, zuclopenthixol, puromycin, sulfasalazine, and imatinib) were identified for treating ESCA. *Conclusions*. This study applied a new approach to analyzing the relation between m⁶A and ESCA prognosis, providing a reference for exploring potential targets and drugs for ESCA treatment.

1. Introduction

Esophagus cancer (ESCA) is the sixth leading cancer with 3.1% incidence and 5.5% mortality worldwide [1]. Males tend to have a higher incidence (4.2%) and mortality (6.8%) than females, mainly due to differences in the smoking and drinking habits between two genders. The major risk factors of ESCA include race, gender, alcohol, tobacco, obesity, gastroesophageal reflux disease (GRED), diet of low fruits and vegetables, and so on [2]. The 5-year overall survival (OS) of the cancer is about 15% to 25%, and poor treatment outcomes are closely associated with late diagnosis and metastasis [3].

In the recent decades, the development of molecular and sequencing technology has deepened the understanding of the genetic causes of ESCA. For example, CCND1, CDK4/ CDK6, and MDM2 genes involved in cell cycle are overexpressed in ESCA patients [4]. High expression of epidermal growth factor receptor (EGFR) is associated with worse prognosis and late clinical stage; therefore EGFR expression could serve as a prognostic biomarker [5, 6]. More importantly, epigenetic factors such as DNA or RNA methylation, histone modifications, and loss of genome imprinting show strong correlation with tumor progression [7–9]. These epigenetic alternations can regulate downstream or upstream gene expression through silencing or activating regulatory factors, resulting in aberrant gene expressions associated with tumor development.

RNA modifications in transcripts are the most commonly detected epigenetic alternation. N⁶-methyladenosine (m⁶A) accounts for a majority of RNA modifications under the control of methyltransferases (writers), binding proteins

(readers), and demethylases (erasers) [10]. Evidence demonstrated that m⁶A modification is involved in tumorigenesis, tumor proliferation, and migration of various types of cancers such as acute myeloid leukemia [11–14], glioblastoma [15, 16], lung cancer [17, 18], hepatocellular carcinoma [19–21], breast cancer [22], and colorectal cancer [23].

Previous studies also discovered a correlation between m⁶A and ESCA. Nagaki et al. proved that knockdown of ALKBH5 can increase m⁶A modification and delay cell cycle progression of esophageal squamous cell carcinoma (ESCC) cells [24]. ALKBH5 has been identified as m⁶A demethylase that facilitates tumor cell proliferation, and a correlation between ALKBH5 and poor prognosis of ESCC patients has been illustrated [24]. Guo et al. observed that high expression of m⁶A reader HNRNPA2B1 was positively associated with ESCA tumor size and lymphatic metastasis [25]. Knockdown of HNRNPA2B1 inhibits tumor cell progression of ESCA cells, indicating HNRNPA2B1 as an oncogenic protein in ESCA development [25]. In addition, HNRNPC and ALKBH5 have been screened and combined as a prognostic signature for predicting ESCA outcomes [26]. These findings provide potential application of m⁶A regulators in clinical treatment.

As m⁶A modification is strongly correlated with tumor proliferation, invasion, and migration, the present study aimed to exploit novel molecular subtypes of ESCA based on m⁶A regulators (writers, readers, and erasers). Furthermore, hub genes associated with ESCA prognosis and potential small-molecule drugs were screened according to molecular subtypes. This study introduced a new strategy of bioinformatics analysis to explore potential drugs for ESCA treatment.

2. Materials and Methods

2.1. Data Information. TCGA-ESCA dataset with the data of gene expression, copy number variation (CNV), single nucleotide variant, methylation, and clinical information was downloaded from The Cancer Genome Atlas (TCGA, https://portal.gdc.cancer.gov/) database. The workflow of defining molecular subtypes of ESCA was shown in Figure 1.

2.2. Genes of m⁶A Writers, Erasers, and Readers. Three types of enzymes (writers, erasers, and readers) related to m⁶A were included. Specifically, m⁶A writers were METTL3, METTL14, WTAP, and KIAA1429. m⁶A erasers were FTO and ALKBH5. m⁶A readers were YTHDC1, YTHDC2, YTHDF1, YTHDF2, YTHDF3, HNRNPA2B1, IGF2BP1, IGF2BP2, and IGF2BP3.

2.3. Data Preprocessing. RNA-seq and methylation data were further processed. For RNA-seq data, samples without clinical follow-up information, overall survival (OS), and survival status were excluded. Genes with transcripts per million (TPM) < 1 in over half of the samples were excluded. Primary solid tumor samples were included. For methylation data, NA value was completed by the KNN function in

impute R package, and beta value was converted to M value. According to cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray, CpGs present in the normal samples were excluded. Unstable methylation sites including CpGs in X and Y chromosomes as well as CpGs in single nucleotide were excluded. Finally, 161 ESCA samples were included (Supplementary Table S1).

2.4. Consensus Clustering. R package of ConsensusClusterPlus (v1.48.0) was used to cluster methylation sites related to ESCA prognosis [27] under the parameters of reps = 100, pItem = 0.8, pFeature = 1, distance = "spearman". D2 algorithm and Euclidean distance were employed for consensus clustering. Cluster numbers k from 2 to 10 were chosen, and the optimal clusters were screened by cumulative distribution function (CDF) curve and consensus CDF.

2.5. Gene Enrichment Analysis and Function Analysis. Single sample gene set enrichment analysis (ssGSEA) in GSVA R package was conducted to calculate the enrichment score of each sample to different functional pathways [28]. WebGestalt (v0.4.3) R package was performed to analyze Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways and gene ontology (GO) enrichment for differentially expressed genes (DEGs).

2.6. Immune Correlation Analysis. We obtained immune checkpoint genes (BTLA, CD200, CD244, LAG3, IDO1, IDO2, PDCD1, CTLA4, PDCD1LG2, TNFRSF8, CD40, TNFSF18, CD86, and CD44) from previous studies to analyze the expression differences of these genes in various molecular subtypes. In addition, we evaluated 28 immune infiltrating cell components by ssGSEA method [29]. To analyze the differences of immune infiltrating cell components in different subtypes, we further evaluated the immune infiltrating score in the sample by using R software package estimate [30], analyzed the differences of immune infiltrating in different subtypes, and evaluated the potential benefits of immunotherapy of different subtypes in imvigor210 [31] by using R software package submap [32].

2.7. Protein-Protein Interaction (PPI) Analysis. STRING (https://string-db.org/) is a database to explore the interaction among known and unknown proteins, including abundant data from current researches, other databases, and data by predicted bioinformatics [33, 34]. The protein interactions of DEGs were analyzed by STRING. PPI result was visualized by Cytoscape (v3.7.2) and further analyzed by cytoHubba to screen hub genes [35–37].

2.8. Databases of Small-Molecule Drugs. Databases of L1000 fireworks display (L1000FWD, https://maayanlab.cloud/L1000FWD/) [38], Drug-Gene Interaction database (DGIdb, https://dgidb.org/) [39, 40], and The Connectivity
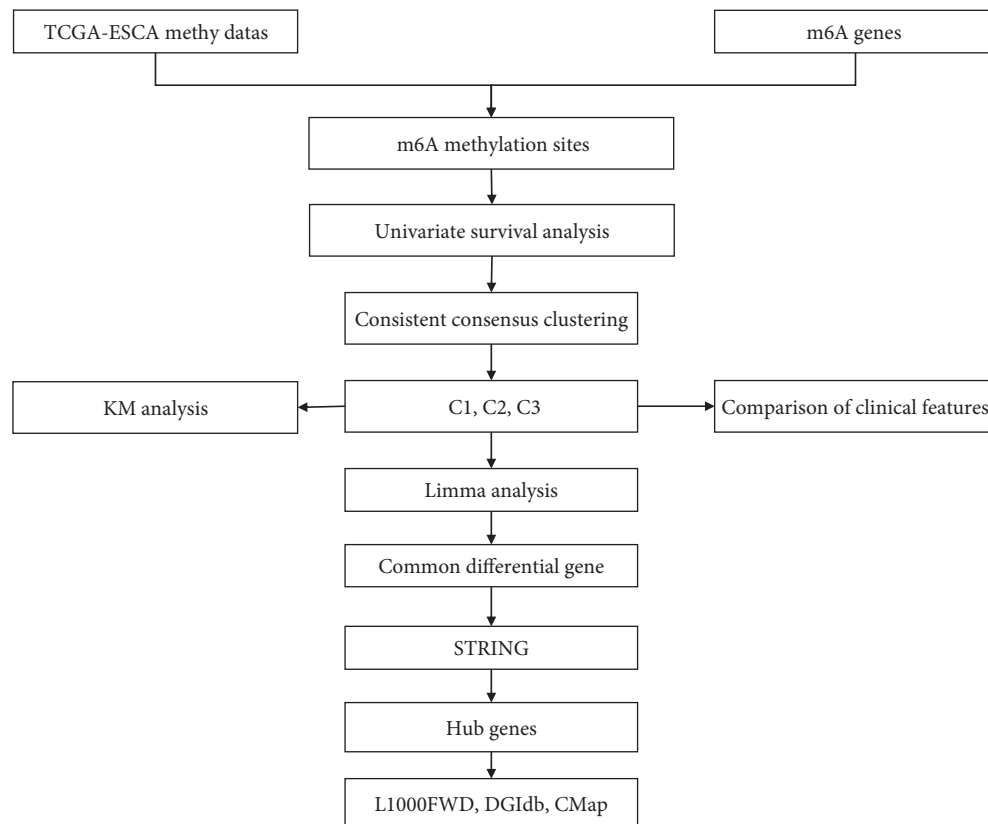
FIGURE 1: The workflow of developing m⁶A-related molecular subtypes and screening potential small-molecular drugs for treating ESCA.

Map (CMap, https://clue.io/) [41, 42] were included for screening small molecules interacting with hub genes. L1000FWD includes over 16000 small molecules and gene expression data from tumor cell lines of 1000 drugs. DGIdb stores over 10000 drugs and 15000 interactions between drugs and genes. CMap contains over 19000 small molecules corresponding to 25200 biological entities. The function of small molecules associated with hub genes came from the National Library of Medicine (PubChem, https://pubchem.ncbi.nlm.nih.gov/#query=).

3. Results

3.1. Consensus Clustering of Methylation Sites on 15 m⁶A-Related Genes. To develop a molecular subtyping system based on m⁶A methylation sites, association between m⁶A methylation sites and ESCA prognosis was analyzed. Coxph function in *R* package survival was used to perform univariate Cox regression analysis between 221 methylation sites and OS, survival status. 9 methylation sites associated with prognosis were screened ($p < 0.05$, Supplementary Table S2). Then 161 ESCA samples were clustered based on the 9 methylation sites with consensus clustering in ConsensusClusterPlus *R* package. As shown in Figure 2, when cluster number $k = 3$, CDF did not show great change; meanwhile, the relative change in area under CDF curve was the maximum, suggesting that $k = 3$ was the optimal. Therefore, under $k = 3$, 161 ESCA samples were clustered into three subtypes of C1, C2, and C3.

Survival analysis manifested significant OS difference in the three subtypes, with the most favorable prognosis detected in C3 subtype ($p = 0.018$, Figure 3). However, no difference of OS was observed between C1 and C2 subtypes. Gene mutation analysis showed that the top mutated gene was TP53 and that C3 subtype had the least mutations of the top 20 mutated genes among three subtypes (Supplementary Figure S1). Such results suggested a relation between gene mutations and prognosis.

3.2. The Relation between the Three Subtypes and Clinical Features. We next assessed the relation between subtypes and clinical features, including *T* stage, *N* stage, *M* stage, stages I to IV, grade, and risk factors of gender and smoking history. Although only the distribution of *G* stage (*G*1, *G*2, and *G*3) showed a close relation with the three subtypes ($p < 0.05$), there was a modest tendency indicating that C3 subtype had the lower proportion of stages with more invasive features than C1 and C2 subtypes (Figures 4(a)–4(e)). For risk factors, males accounted for a significantly higher proportion than female due to a higher smoking in males, and the female proportion was the highest in C3 subtype (Figure 4(f)). In addition, the number of nonsmokers was more in C3 subtype (tobacco = 1) than C1 and C2 subtypes; however, no significant difference was detected among tobacco groups (Figure 4(g)). We compared the three molecular subtypes with the previously reported three molecular subtypes (CIN, GS, and MSI) [43]. We observed
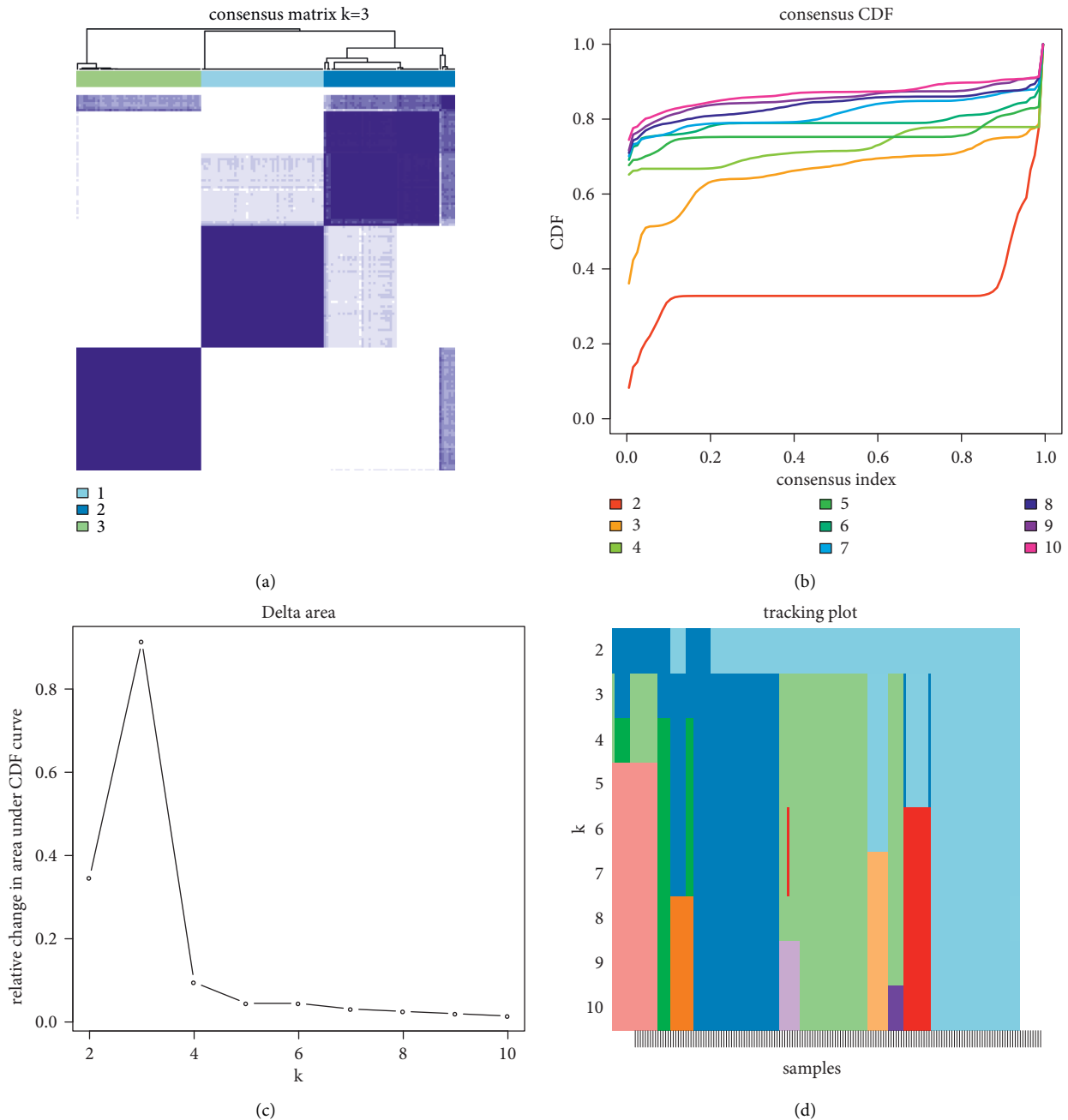
(a)

(b)

(c)

(d)

FIGURE 2: Consensus clustering of 161 ESCA samples based on m^{6}A methylation sites. (a) A consensus matrix when $k = 3$ where k represents cluster number. (b) Consensus CDF when $k = 2$ to 10. (c) The relative change in area under CDF curve when $k = 2$ to 10. (d) Tracking plot of samples when $k = 2$ to 10.

that MSI subtypes are mainly related to C2 (Figure 4(h)). For example, the distribution of C1, C2, and C3 in CIN and GS subtypes is similar, suggesting that the new three molecular subtypes can be used as a supplement to the previously reported molecular subtypes.

3.3. The Enrichment of Metabolism Pathways in the Three Subtypes. Compared with normal cells, tumor cells are more active in acquiring energy through metabolism pathways to promote cell proliferation and migration.

Therefore, we speculated that the activity of tumor cells in metabolism pathways could indicate the condition of cancer patients' prognosis. To examine whether there was an association between subtypes and metabolism pathways, ssGSEA was conducted to calculate the enrichment score of each sample. Eight major metabolism pathways, including nitrogen metabolism, nicotinate and nicotinamide metabolism, histidine metabolism, glyoxylate and dicarboxylate metabolism, glycerophospholipid metabolism, glycerolipid metabolism, drug metabolism cytochrome p450, and glutathione metabolism, were analyzed.
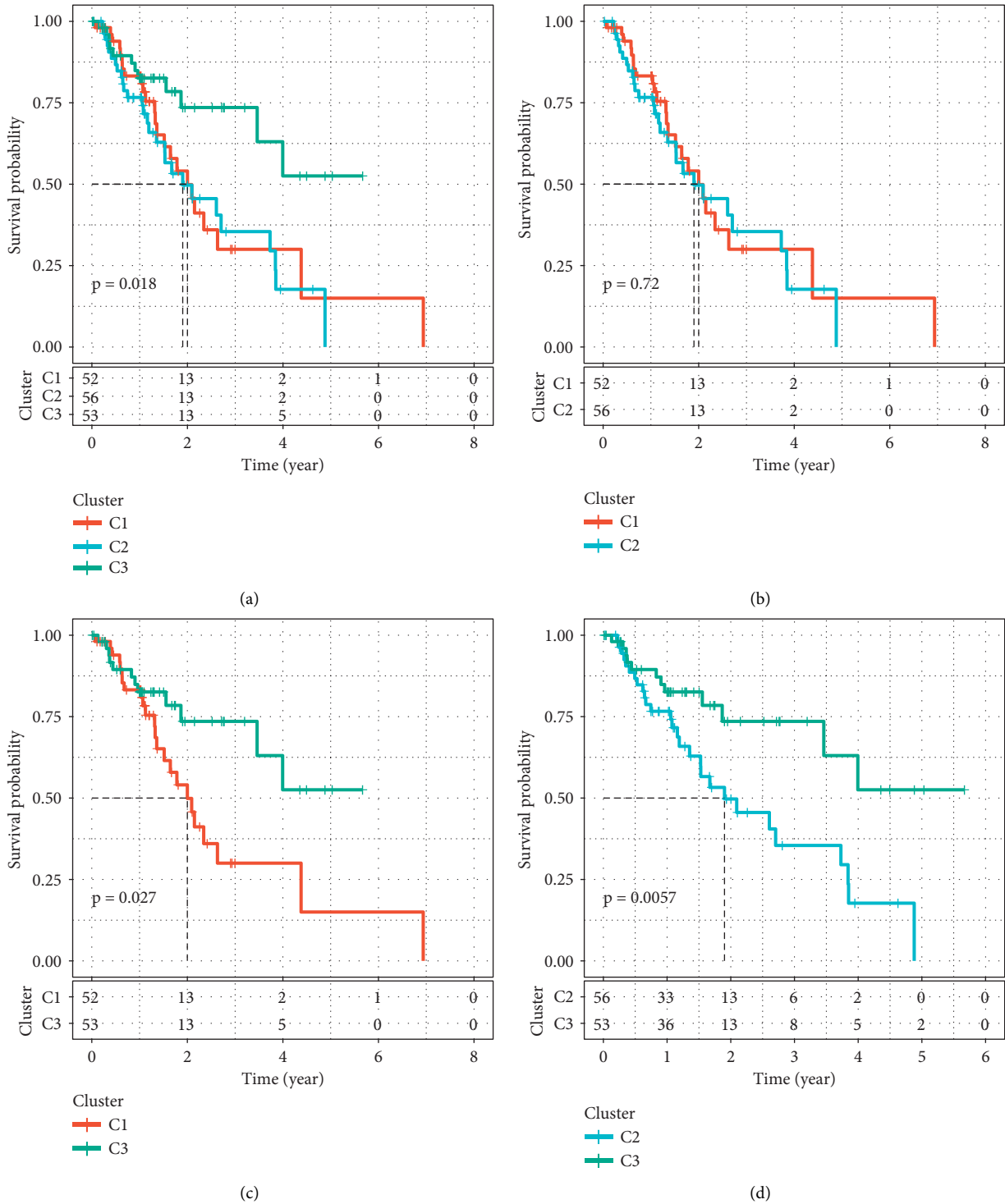
(a)

(b)

(c)

(d)

FIGURE 3: Kaplan-Meier survival curve of C1, C2, and C3 subtypes. (a) Survival analysis among three subtypes ($p = 0.018$). (b) Survival analysis between C1 and C2 subtypes ($p = 0.72$). (c) Survival analysis between C1 and C3 subtypes ($p = 0.027$). (d) Survival analysis between C1 and C3 subtypes ($p = 0.0057$). Log-rank test was performed.

The result exhibited that C3 subtype had the lowest enrichment score in these pathways, suggesting that C3 subtype with favorable prognosis was relatively inactive in metabolism pathways (Figure 5).

3.4. Immune Correlation of Different Molecular Subtypes. Immunotherapy is a promising clinical treatment method. In order to characterize the potential benefits of immunotherapy of different molecular subtypes, we first compared
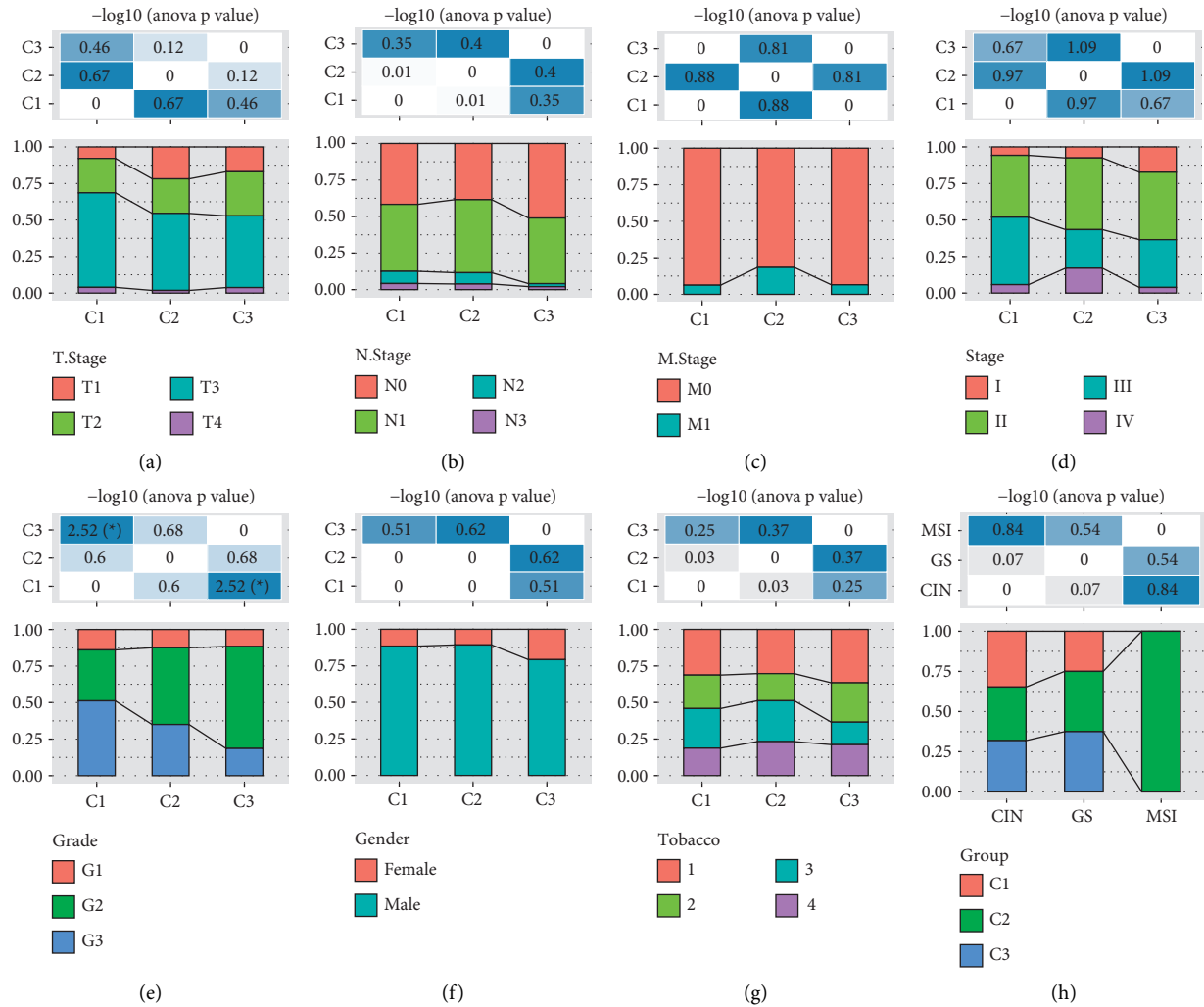
FIGURE 4: The relation between three subtypes and clinical features, including T stage (a), N stage (b), M stage (c), stages I to IV (d), grade (e), gender (f), tobacco (g), and TCGA molecular subtypes (h). ANOVA was performed. $^*p < 0.05$.

the differences of immune infiltration in the immune microenvironment of the three molecular subtypes. It can be observed that C1 subtype has higher matrix components and higher tumor purity (Figure 6(a)). We also observed the differences of multiple immune cell infiltration in patients with three molecular subtypes (Figure 6(b)). For example, C3 subtype has higher effector memory CD8 T cell and activated B cell, and C1 subtype has the highest regulatory T cell. These results show that the three molecular subtypes have different immune microenvironment characteristics. Further, we analyzed the expression differences of immune checkpoint genes in the three molecular subtypes and observed that 10 (71%) immune checkpoint genes had significant expression differences (Figure 6(c)), of which CD40 had the most significant expression difference. In addition, we also observed that C1 subtype was significantly correlated with CR/PR (Figure 6(d)). This suggests that C1 subtype may benefit from immunotherapy of PD-L1.

Identification of differentially expressed genes among the three subtypes and functional analysis were done.

As no difference of OS was found between C1 and C2 subtypes, and C3 subtype had the optimal prognosis, we also analyzed the DEGs between C1 and C3 and between C2 and C3 subtypes. Between C1 and C3 subtypes, 193 DEGs (132 upregulated genes and 61 downregulated genes) were identified under conditions of $p < 0.05$ and |fold change (FC)| > 1.5 using Limma R package (Figure 7(a)). Then 193 DEGs were further assessed with GO function analysis and KEGG pathways using WebGestalt R package. GO analysis showed that 432 terms of biological process, 27 terms of cellular component, and 41 terms of molecular function were annotated with significant differences between C1 and C3 subtypes ($p < 0.05$). The top 10 enriched terms of biological process, cellular component, and molecular function were displayed (Figures 7(b)–7(d)). However, no KEGG pathways with significant difference between C1 and C3 subtypes were found. Moreover, between C2 and C3 subtypes, we identified 1673 DEGs incorporating 685 upregulated and 988 downregulated genes (Supplementary Figure S2) and annotated 35 KEGG pathways, 1181 terms of biological process, 132 terms of cellular component, and 153
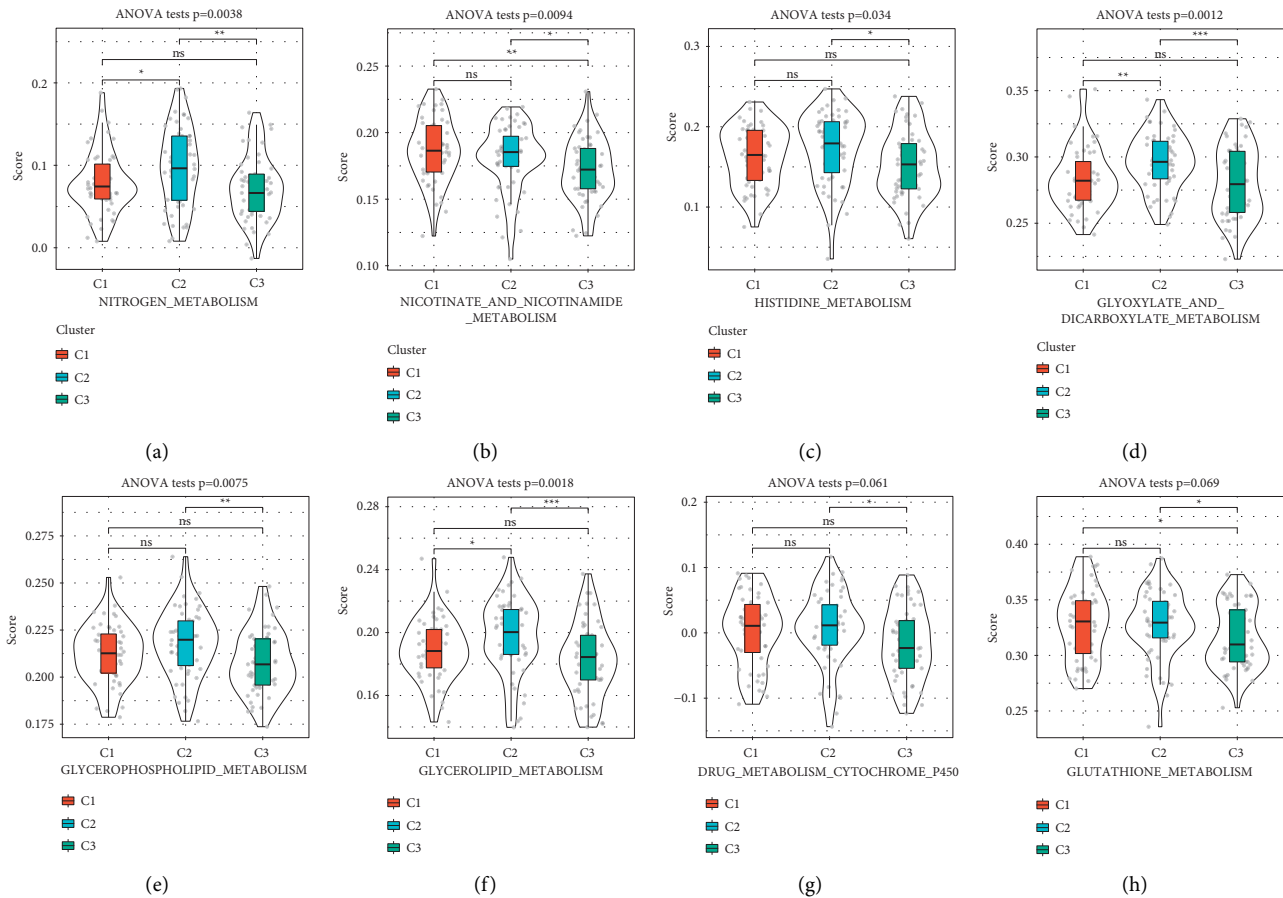
FIGURE 5: The relation between the three subtypes and metabolism pathways, including nitrogen metabolism (a), nicotinate and nicotinamide metabolism (b), histidine metabolism (c), glyoxylate and dicarboxylate metabolism (d), glycerophospholipid metabolism (e), glycerolipid metabolism (f), drug metabolism cytochrome p450 (g), and glutathione metabolism (h). ANOVA was performed. ns, no significance. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

terms of molecular function. The top 10 enriched terms were shown in Supplementary Figure S3. Among these terms, epidermal cell differentiation, striated muscle cell differentiation, skin development, epidermis development, and epithelial cell differentiation were all annotated in the top 10 biological processes between C1 and C3 and between C2 and C3 (Figure 7 and Supplementary Figure S3).

3.5. Construction of PPI Networks and Hub Gene Identification.

Next, mutually upregulated and downregulated DEGs between C1 and C3 and between C2 and C3 subtypes were examined. 146 mutual DEGs including 97 upregulated and 49 downregulated ones were identified for constructing PPI networks (Figure 8(a)). The bioinformatics tools in STRING database were applied to assess the interactions among 146 proteins of DEGs. Subsequently, Cytoscape was applied to visualize the PPI network and cytoHubba was performed to screen hub genes from the network (Figure 8(b)). Finally, the following top 15 hub genes were identified: OCLN, TFF1, BMP4, KRT18, CLDN3, CLDN4, KRT8, TFAP2A, PPARG, AGR2, GATA4, EPCAM, SNAI2, EGFR, and TMPRSS2. We further evaluated the expression differences of these 15 genes in cancer and adjacent tumors. We observed that GATA4, AGR2, and PPARG were significantly underexpressed in tumor samples (Supplementary Figure S4A). We further evaluated the methylation level of CpG sites in the promoter region of these 15 genes in each sample. It can be observed that there is a higher methylation level in cancer samples as a whole, in particular, GATA4 and TFAP2A (Supplementary Figure S4B). We used ssGSEA to evaluate the enrichment scores of six important immune pathways and further analyzed the correlation between these 15 genes and these immune pathways. It was observed that there was a higher correlation between these genes and weak correlation with immune pathways, among which EPCAM was the most correlated with immune pathways (Supplementary Figure S4(c)).

3.6. Screening of Small Molecules Related to Hub Genes.

The 15 hub genes were screened from DEGs between C1 and C3, C2 and C3 were considered to be closely related to ESCA prognosis, and this also suggested that these genes could be the targets for ESCA treatment. Therefore, we introduced three databases of L1000FWD, DGIdb, and CMap with abundant data of the interactions between small-molecule
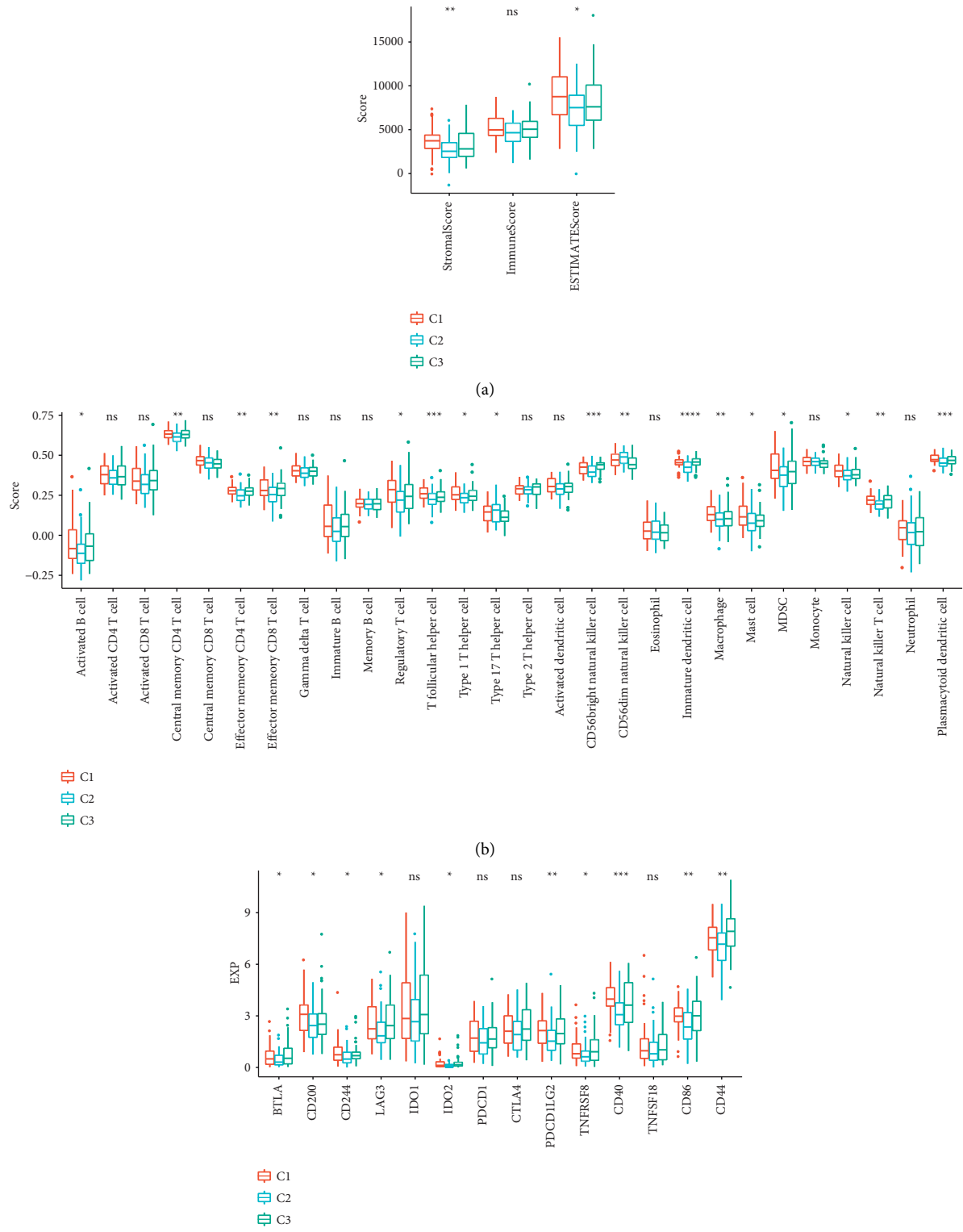
(a)
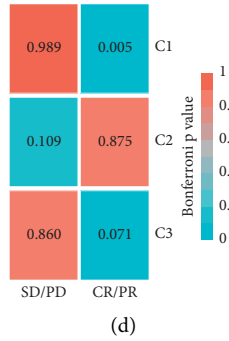

(b)

(c)

Figure 6: Continued.

(d)

FIGURE 6: Immune correlation of three molecular subtypes. (a) Different distribution of immune infiltration of different molecular subtypes. (b) The distribution of immune cell infiltration in different molecular subtypes was different. (c) The expression and distribution of immune checkpoint genes of different molecular subtypes were different. (d) Correlation of immunotherapeutic response of PD-L1 with different molecular subtypes. Analysis of variance was used to test the difference between multiple groups of samples.
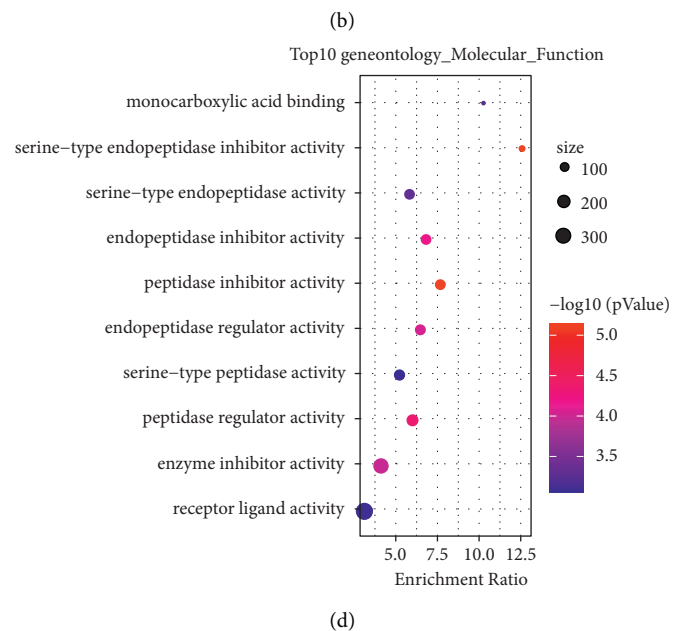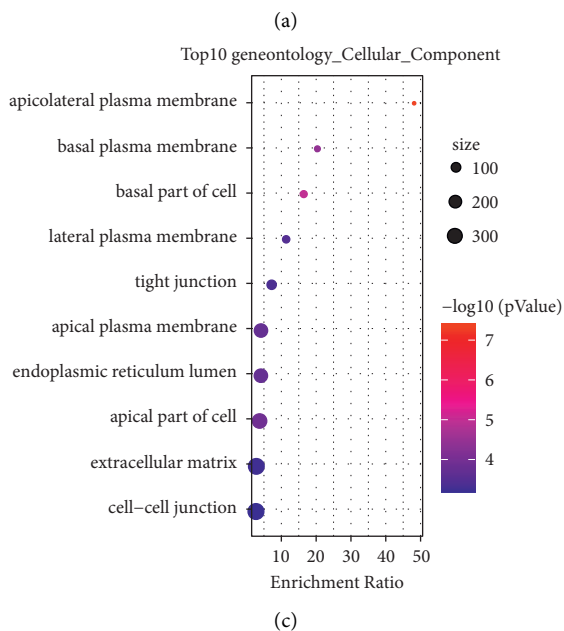

(a)

(b)

(c)

(d)
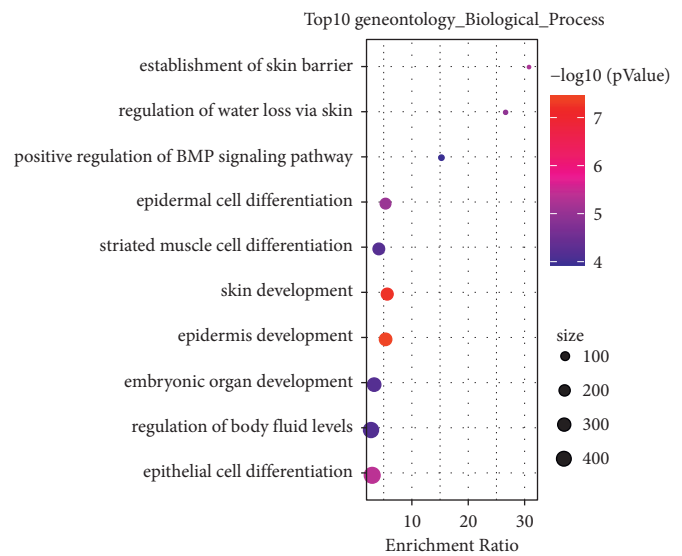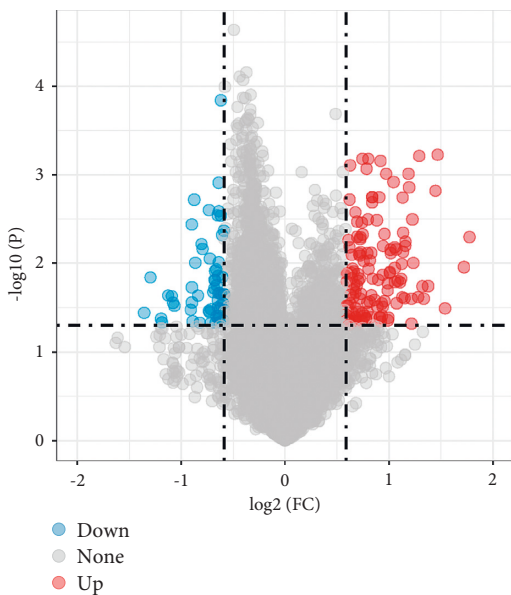
FIGURE 7: The volcano plot of DEGs between C1 and C3 subtypes. Blue represents downregulated genes and red represents upregulated genes. FC, fold change (a). The top 10 annotated terms of biological process (b), cellular component (c), and molecular function (d) between C1 and C3 subtypes. Dot size represents the gene numbers. The annotated terms were displayed in vertical axis and the enrichment ratio of each term was displayed in horizontal axis.
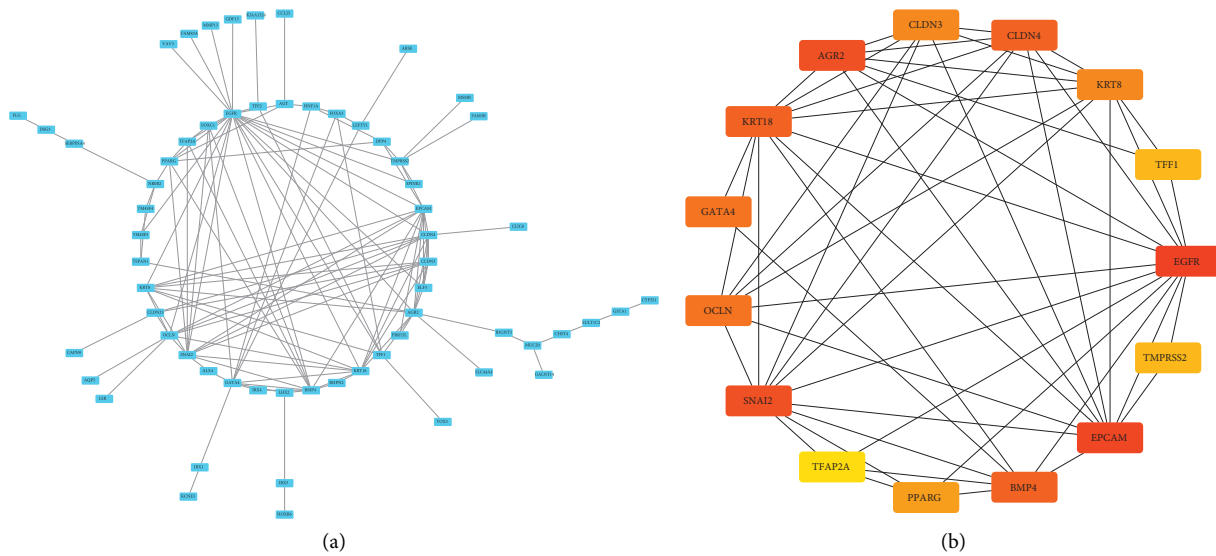
(a)

(b)

FIGURE 8: PPI analysis of 146 upregulated and downregulated genes. (a) A PPI network of 146 DEGs constructed by STRING. (b) A PPI network of 15 hub genes screened by Cytoscape and cytoHubba. Black line represents the interactions between proteins.

drugs and genes. If one drug is negatively associated with expression of one gene related to ESCA, the drug could be considered as a potential drug for ESCA treatment. Within three databases, we screened a total of 598 small molecules having interactions with hub genes, including 96 from L1000FWD, 439 from DGIdb, and 63 from CMap. By overlapping these small molecules in three databases, we observed 3 small molecules (BMS-754807, gefitinib, and neratinib) were overlapped between L1000FWD and DGIdb, 3 small molecules (zuclopenthixol, puromycin, and naringenin) were overlapped between L1000FWD and CMap, 2 small molecules (sulfasalazine and imatinib) were overlapped between DGIdb and CMap (Figure 9). Among these 8 small molecules, BMS-754807, gefitinib, neratinib, and imatinib have antitumor activity. Zuclopenthixol, as a Dopamine receptor antagonist, is a drug for treating schizophrenia. Puromycin is an aminoglycoside antibiotic, and sulfasalazine is a nonsteroid anti-inflammatory drug. These drugs may specifically target hub genes and take function in suppressing tumor cell proliferation and invasion, although further experiment and analysis are needed for illustrating their function and mechanism in antitumor activity.

4. Discussion

A number of epigenetic studies on ESCA have revealed the significance of epigenetic regulation on ESCA development; however, the role of m^{6}A modification on ESCA has not been systematically studied. Only several studies have found that some m^{6}A regulators, such as ALKBH5, HNRNPA2B1, and HNRNPC, have strong relation with ESCA prognosis [24–26]. Inspired from the previous researches, we focused on analyzing a total of 15 m^{6}A regulators and identified three new molecular subtypes associated with clinical features and ESCA prognosis. Furthermore, we constructed a PPI network based on DEGs screened from the three subtypes and
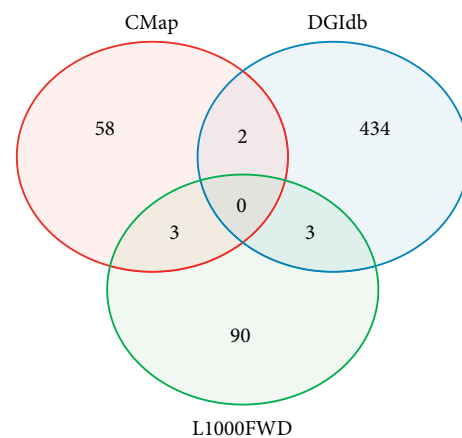

FIGURE 9: Screening of small molecules related to hub genes from L1000FWD, DGIdb, and CMap databases.

determined 15 prognosis-related hub genes from the PPI network.

Some of the 15 hub genes have been reported to be associated with tumor progression of ESCA. For example, TFF1 encodes a mucosa protector factor, and it is silenced in the early stage of ESCA development resulting from high methylation of TFF1 promoter [44]. BMP4 and EPCAM are involved in inducing epithelial-mesenchymal transition (EMT) and promoting tumor cell migration of ESCA [45–47]. Low expression of CLDN4 is indicative of a poor prognosis of ESCC [48]. High expression of TFAP2A is correlated with favorable OS of ESCC patients [49]. EGFR is highly expressed in ESCA and some other cancer types; moreover, it is seen as a promising target for inhibiting tumor aggression [50]. Although some hub genes have not been found to be correlated with ESCA development, their relations with other cancer types have been previously demonstrated.

Using small-molecule databases, 8 potential drugs closely interacting with the 15 hub genes were identified. These drugs negatively associated with expression of the hub genes can be considered as potential drugs for treating ESCA. Among the 8 drugs, gefitinib, neratinib, and imatinib have been commercialized for clinical treatment of specific cancers. Gefitinib is an EGFR tyrosine kinase inhibitor that can hinder tumor cell proliferation and angiogenesis and has been commercially applied in treating non-small-cell lung cancer [51]. Clinical trials of gefitinib in advanced ESCA patients demonstrated a partial response and stable disease rate of between 24% and 39%, showing a relatively positive effect [52–54]. Neratinib is a tyrosine kinase inhibitor targeting HER1, HER2, and HER4 and can effectively improve disease-free survival of HER2-positive breast cancer patients given with chemotherapy and trastuzumab [55, 56]. Imatinib, a tyrosine kinase inhibitor targeting Bcr-Abl tyrosine kinase, could suppress disease progression and extend overall survival of chronic myeloid leukemia and gastrointestinal stromal tumors [57, 58].

BMS-754807 has not been used to treat cancers; however, evidence suggested a promising application of it in clinical practice. BMS-754807 is an inhibitor of targeting insulin-like growth factor-1 receptor/insulin receptor (IGF-1R/IR) signaling pathway, which has been proven to be effective in suppressing tumor cell proliferation of xenograft tumor models of several cancer types [59–61]. Study found that sulfasalazine could enhance cisplatin-induced cytotoxic effects on advanced gastric cancer and bladder cancer [62, 63]. The remaining two drugs zuclopenthixol and puromycin have not been reported to be related to cancer therapy, but they still may have the potential to target hub genes related to ESCA prognosis, according to our analysis.

This study did not differentiate two molecular types of esophagus cancer (squamous cell carcinoma and adenocarcinoma), which may affect the results of molecular subtypes to some extent. In addition, further study on these hub genes and small molecules are needed to demonstrate their functions in clinical practice. Importantly, this study applied a new approach to analyzing the relation between m^{6}A and ESCA prognosis and provided a valuable reference to explore potential targets and drugs for ESCA treatment.

5. Conclusions

In conclusion, this study determined three molecular subtypes of ESCA based on m^{6}A regulators and identified 8 potential small-molecule drugs closely interacting with hub genes through integrative analysis. The new molecular subtypes were effective in classifying ESCA patients into low-risk and high-risk groups. The 15 hub genes screened from DEGs among three subtypes can be potential targets for treating ESCA. The 8 small-molecule drugs closely interacting with the hub genes may be promising drugs for ESCA patients.

Abbreviations

CDF:	Cumulative distribution function
CMap:	The connectivity map
CNV:	Copy number variation
DEGs:	Differentially expressed genes
DGIdb:	Drug-gene interaction database
EGFR:	Epidermal growth factor receptor
ESCA:	Esophagus cancer
ESCC:	Esophageal squamous cell carcinoma
FC:	Fold change
GO:	Gene ontology
GRED:	Gastroesophageal reflux disease
IGF-1R/IR:	Insulin-like growth factor-1 receptor/insulin receptor
KEGG:	Kyoto encyclopedia of genes and genomes
L1000FWD:	L1000 fireworks display
m^{6}A:	N^6-methyladenosine
OS:	Overall survival
PPI:	Protein-protein interaction
ssGSEA:	Single sample gene set enrichment analysis
TCGA:	The cancer genome atlas
TPM:	Transcripts per million.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

Supplementary Materials

Supplementary Figure S1. The mutation patterns of top 20 mutated genes in C1 (a), C2 (b), and C3 (c) subtypes. *Supplementary Figure S2.* The volcano plot of DEGs between C2 and C3 subtypes. Blue represents downregulated genes and red represents upregulated genes. FC, fold change. *Supplementary Figure S3.* The top 10 annotated terms of biological processes (a), cellular component (b), molecular function (c), and KEGG pathways (D) between C2 and C3 subtypes. Dot size represents the gene numbers. The annotated terms were displayed in vertical axis and the enrichment ratio of each term was displayed in horizontal axis. *Supplementary Figure S4.* a: Differential expression of 15 genes in cancer and adjacent tissues; b: Heatmap of methylation distribution of promoter region methylation sites of 15 genes in various types of samples; c: Correlation heatmap of 15 genes and 6 immune related pathways. *Supplementary Table S1.* The clinical information of 161

ESCA samples. *Supplementary Table S2*. 9 m^{6}A methylation sites screened by univariate Cox regression analysis. (*Supplementary Materials*)

References

[1] H. Sung, J. Feerlay, R. L. Siegel et al., "Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: a Cancer Journal for Clinicians*, vol. 71, no. 3, pp. 209–241, 2021.

[2] M. J. D. Arnal, Á. Ferrández Arenas, and Á. L. Arbeloa, "Esophageal cancer: risk factors, screening and endoscopic treatment in Western and Eastern countries," *World Journal of Gastroenterology*, vol. 21, no. 26, pp. 7933–7943, 2015.

[3] A. Pennathur, M. K. Gibson, B. A. Jobe, and J. D. Luketich, "Oesophageal carcinoma," *The Lancet*, vol. 381, no. 9864, pp. 400–412, 2013.

[4] Y. Song, L. Li, Y. Ou et al., "Identification of genomic alterations in oesophageal squamous cell cancer," *Nature*, vol. 509, no. 7498, pp. 91–95, 2014.

[5] W. Zhang, H. Zhu, X. Liu et al., "Epidermal growth factor receptor is a prognosis predictor in patients with esophageal squamous cell carcinoma," *The Annals of Thoracic Surgery*, vol. 98, no. 2, pp. 513–519, 2014.

[6] Z. Gao, X. Meng, D. Mu, X. Sun, and J. Yu, "Prognostic significance of epidermal growth factor receptor in locally advanced esophageal squamous cell carcinoma for patients receiving chemoradiotherapy," *Oncology Letters*, vol. 7, no. 4, pp. 1118–1122, 2014.

[7] T. D. Ahrens, M. Werner, and S. Lassmann, "Epigenetics in esophageal cancers," *Cell and Tissue Research*, vol. 356, no. 3, pp. 643–655, 2014.

[8] A. Nebbioso, F. P. Tambaro, C. Dell'Aversana, and L. Altucci, "Cancer epigenetics: moving forward," *PLoS Genetics*, vol. 14, no. 6, Article ID e1007362, 2018.

[9] Q. Lan, P. Y. Liu, J. Haase, J. L. Bell, S. Hüttelmaier, and T. Liu, "The critical role of RNA m6A methylation in cancer," *Cancer Research*, vol. 79, no. 7, pp. 1285–1292, 2019.

[10] X.-Y. Chen, J. Zhang, and J.-S. Zhu, "The role of m6A RNA methylation in human cancer," *Molecular Cancer*, vol. 18, no. 1, p. 103, 2019.

[11] Z. Li, H. Weng, R. Su et al., "FTO plays an oncogenic role in acute myeloid leukemia as a N 6 -methyladenosine RNA demethylase," *Cancer Cell*, vol. 31, no. 1, pp. 127–141, 2017.

[12] L. P. Vu, B. F. Pickering, Y. Cheng et al., "The N6-methyladenosine (m6A)-forming enzyme METTL3 controls myeloid differentiation of normal hematopoietic and leukemia cells," *Nature Medicine*, vol. 23, no. 11, pp. 1369–1376, 2017.

[13] I. Barbieri, K. Tzelepis, L. Pandolfini et al., "Promoter-bound METTL3 maintains myeloid leukaemia by m6A-dependent translation control," *Nature*, vol. 552, no. 7683, pp. 126–131, 2017.

[14] H. Weng, H. Huang, H. Wu et al., "METTL14 inhibits hematopoietic stem/progenitor differentiation and promotes leukemogenesis via mRNA m6A modification," *Cell Stem Cell*, vol. 22, no. 2, pp. 191–205, 2018.

[15] S. Zhang, B. S. Zhao, A. Zhou et al., "m6A demethylase ALKBH5 maintains tumorigenicity of glioblastoma stem-like cells by sustaining FOXM1 expression and cell proliferation program," *Cancer Cell*, vol. 31, no. 4, pp. 591–606, 2017.

[16] Q. Cui, H. Shi, P. Ye et al., "m6A RNA methylation regulates the self-renewal and tumorigenesis of glioblastoma stem cells," *Cell Reports*, vol. 18, no. 11, pp. 2622–2634, 2017.

[17] S. Lin, J. Choe, P. Du, R. Triboulet, and R. I. Gregory, "The m 6 A methyltransferase METTL3 promotes translation in human cancer cells," *Molecular Cell*, vol. 62, no. 3, pp. 335–345, 2016.

[18] Y. Du, G. Hou, H. Zhang et al., "Sumoylation of the m6A-RNA methyltransferase METTL3 modulates its function," *Nucleic Acids Research*, vol. 46, no. 10, pp. 5195–5208, 2018.

[19] M. Chen, L. Wei, C.-T. Law et al., "RNA N6-methyladenosine methyltransferase-like 3 promotes liver cancer progression through YTHDF2-dependent posttranscriptional silencing of SOCS2," *Hepatology*, vol. 67, no. 6, pp. 2254–2270, 2018.

[20] J. Z. Ma, F. Yang, C. C. Zhou et al., "METTL14 suppresses the metastatic potential of hepatocellular carcinoma by modulating N 6 -methyladenosine-dependent primary MicroRNA processing," *Hepatology*, vol. 65, no. 2, pp. 529–543, 2017.

[21] S. Müller, M. Glaß, A. K. Singh et al., "IGF2BP1 promotes SRF-dependent transcription in cancer in a m6A- and miRNA-dependent manner," *Nucleic Acids Research*, vol. 47, no. 1, pp. 375–390, 2019.

[22] C. Zhang, D. Samanta, H. Lu et al., "Hypoxia induces the breast cancer stem cell phenotype by HIF-dependent and ALKBH5-mediated m6A-demethylation of NANOG mRNA," *Proceedings of the National Academy of Sciences*, vol. 113, no. 14, pp. E2047–E2056, 2016.

[23] A. Tanabe, K. Tanikawa, M. Tsunetomi et al., "RNA helicase YTHDC2 promotes cancer metastasis via the enhancement of the efficiency by which HIF-1α mRNA is translated," *Cancer Letters*, vol. 376, no. 1, pp. 34–42, 2016.

[24] Y. Nagaki, S. Motoyama, T. Yamaguchi et al., "m6A demethylase ALKBH5 promotes proliferation of esophageal squamous cell carcinoma associated with poor prognosis," *Genes to Cells*, vol. 25, no. 8, pp. 547–561, 2020.

[25] H. Guo, B. Wang, K. Xu et al., "m6A reader HNRNPA2B1 promotes esophageal cancer progression via up-regulation of ACLY and ACC1," *Frontiers in Oncology*, vol. 10, Article ID 553045, 2020.

[26] L.-C. Xu, J.-X. Pan, and H.-D. Pan, "Construction and validation of an m6A RNA methylation regulators-based prognostic signature for esophageal cancer," *Cancer Management and Research*, vol. 12, pp. 5385–5394, 2020.

[27] M. D. Wilkerson and D. N. Hayes, "Consensus cluster plus: a class discovery tool with confidence assessments and item tracking," *Bioinformatics*, vol. 26, no. 12, pp. 1572-1573, 2010.

[28] S. Hänzelmann, R. Castelo, and J. Guinney, "GSVA: gene set variation analysis for microarray and RNA-seq data," *BMC Bioinformatics*, vol. 14, no. 1, p. 7, 2013.

[29] P. Charoentong, F. Finotello, M. Angelova et al., "Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade," *Cell Reports*, vol. 18, no. 1, pp. 248–262, 2017.

[30] K. Yoshihara, M. Shahmoradgoli, E. Martínez et al., "Inferring tumour purity and stromal and immune cell admixture from expression data," *Nature Communications*, vol. 4, no. 1, p. 2612, 2013.

[31] A. V. Balar, M. D. Galsky, J. E. Rosenberg et al., "Atezolizumab as first-line treatment in cisplatin-ineligible patients with locally advanced and metastatic urothelial carcinoma: a single-arm, multicentre, phase 2 trial," *Lancet (London, England)*, vol. 389, no. 10064, pp. 67–76, 2017.

[32] Y. Hoshida, J.-P. Brunet, P. Tamayo, T. R. Golub, and J. P. Mesirov, "Subclass mapping: identifying common subtypes in independent disease data sets," *PLoS One*, vol. 2, no. 11, Article ID e1195, 2007.

[33] B. Snel, G. Lehmann, P. Bork, and M. A. Huynen, "STRING: a web-server to retrieve and display the repeatedly occurring

neighbourhood of a gene," *Nucleic Acids Research*, vol. 28, no. 18, pp. 3442–3444, 2000.

[34] D. Szklarczyk, A. L. Gable, D. Lyon et al., "STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets," *Nucleic Acids Research*, vol. 47, no. D1, pp. D607–D613, 2019.

[35] P. Shannon, A. Markiel, O. Ozier et al., "Cytoscape: a software environment for integrated models of biomolecular interaction networks," *Genome Research*, vol. 13, no. 11, pp. 2498–2504, 2003.

[36] D. Otasek, J. H. Morris, J. Bouças, A. R. Pico, and B. Demchak, "Cytoscape Automation: empowering workflow-based network analysis," *Genome Biology*, vol. 20, no. 1, p. 185, 2019.

[37] C. H. Chin, S.-H. Chen, H.-H. Wu, C.-W. Ho, M.-T. Ko, and C.-Y. Lin, "cytoHubba: identifying hub objects and subnetworks from complex interactome," *BMC Systems Biology*, vol. 8, no. Suppl 4, p. S11, 2014.

[38] Z. Wang, A. Lachmann, A. B. Keenan, and A. Ma'ayan, "L1000FWD: fireworks visualization of drug-induced transcriptomic signatures," *Bioinformatics*, vol. 34, no. 12, pp. 2150–2152, 2018.

[39] M. Griffith, O. L. Griffith, A. C. Coffman et al., "DGIdb: mining the druggable genome," *Nature Methods*, vol. 10, no. 12, pp. 1209-1210, 2013.

[40] S. L. Freshour, S. Kiwala, K. C. Cotto et al., "Integration of the drug-gene interaction database (DGIdb 4.0) with open crowdsource efforts," *Nucleic Acids Research*, vol. 49, no. D1, pp. D1144–d1151, 2021.

[41] J. Lamb, E. D. Crawford, D. Peck et al., "The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease," *Science*, vol. 313, no. 5795, pp. 1929–1935, 2006.

[42] A. Subramanian, R. Narayan, S. M. Corsello et al., "A next generation connectivity Map: L1000 platform and the first 1,000,000 profiles," *Cell*, vol. 171, no. 6, p. 1437, 2017.

[43] Cancer Genome Atlas Research Network, "Integrated genomic characterization of oesophageal carcinoma," *Nature*, vol. 541, no. 7636, pp. 169–175, 2017.

[44] I. M. Gonzaga, S. C. Soares Lima, M. C. Nicolau et al., "TFF1 hypermethylation and decreased expression in esophageal squamous cell carcinoma and histologically normal tumor surrounding esophageal cells," *Clinical Epigenetics*, vol. 9, no. 1, p. 130, 2017.

[45] C. Kestens, P. D. Siersema, G. J. A. Offerhaus, and J. W. P. M. van Baal, "BMP4 signaling is able to induce an epithelial-mesenchymal transition-like phenotype in barrett's esophagus and esophageal adenocarcinoma through induction of SNAIL2," *PLoS One*, vol. 11, no. 5, Article ID e0155754, 2016.

[46] J. Zhang, A. Luo, F. Huang, T. Gong, and Z. Liu, "SERPINE2 promotes esophageal squamous cell carcinoma metastasis by activating BMP4," *Cancer Letters*, vol. 469, pp. 390–398, 2020.

[47] C. Driemel, H. Kremling, S. Schumacher et al., "Context-dependent adaption of EpCAM expression in early systemic esophageal cancer," *Oncogene*, vol. 33, no. 41, pp. 4904–4915, 2014.

[48] C. O. Sung, S. Y. Han, and S.-H. Kim, "Low expression of claudin-4 is associated with poor prognosis in esophageal squamous cell carcinoma," *Annals of Surgical Oncology*, vol. 18, no. 1, pp. 273–281, 2011.

[49] Y. Zhang, Y. Xu, Z. Li et al., "Identification of the key transcription factors in esophageal squamous cell carcinoma," *Journal of Thoracic Disease*, vol. 10, no. 1, pp. 148–161, 2018.

[50] S. Sigismund, D. Avanzato, and L. Lanzetti, "Emerging functions of the EGFR in cancer," *Molecular oncology*, vol. 12, no. 1, pp. 3–20, 2018.

[51] R. S. Herbst, M. Fukuoka, and J. Baselga, "Gefitinib - a novel targeted approach to treating cancer," *Nature Reviews Cancer*, vol. 4, no. 12, pp. 956–965, 2004.

[52] D. J. Adelstein, C. P. Rodriguez, L. A. Rybicki, D. I. Ives, and T. W. Rice, "A phase II trial of gefitinib for recurrent or metastatic cancer of the esophagus or gastroesophageal junction," *Investigational New Drugs*, vol. 30, no. 4, pp. 1684–1689, 2012.

[53] D. R. Ferry, M. Anderson, K. Beddard et al., "A phase II study of gefitinib monotherapy in advanced esophageal adenocarcinoma: evidence of gene expression, cellular, and clinical response," *Clinical Cancer Research*, vol. 13, no. 19, pp. 5869–5875, 2007.

[54] Y. Xu, Z. Xie, Y. Shi et al., "Gefitinib single drug in treatment of advanced esophageal cancer," *Journal of Cancer Research and Therapeutics*, vol. 12, no. Supplement, pp. C295–c297, 2016.

[55] A. Chan, S. Delaloge, F. A. Holmes et al., "Neratinib after trastuzumab-based adjuvant therapy in patients with HER2-positive breast cancer (ExteNET): a multicentre, randomised, double-blind, placebo-controlled, phase 3 trial," *The Lancet Oncology*, vol. 17, no. 3, pp. 367–377, 2016.

[56] M. Martin, F. A. Holmes, B. Ejlertsen et al., "Neratinib after trastuzumab-based adjuvant therapy in HER2-positive breast cancer (ExteNET): 5-year analysis of a randomised, double-blind, placebo-controlled, phase 3 trial," *The Lancet Oncology*, vol. 18, no. 12, pp. 1688–1700, 2017.

[57] A. Hochhaus, R. A. Larson, F. Guilhot et al., "Long-term outcomes of imatinib treatment for chronic myeloid leukemia," *New England Journal of Medicine*, vol. 376, no. 10, pp. 917–927, 2017.

[58] J. A. Fernández, D. Ferreras, J. J. Ruiz-Manzanera et al., "Characteristics and prognosis of jejunoileal gastrointestinal stromal tumours (GISTs) in the era of imatinib: a comparative study with gastric GISTs," *Clinical and Translational Oncology: Official Publication of the Federation of Spanish Oncology Societies and of the National Cancer Institute of Mexico*, vol. 23, no. 7, pp. 1368–1376, 2021.

[59] X. Hou, F. Huang, L. F. Macedo et al., "Dual IGF-1R/InsR inhibitor BMS-754807 synergizes with hormonal agents in treatment of estrogen-dependent breast cancer," *Cancer Research*, vol. 71, no. 24, pp. 7597–7607, 2011.

[60] N. Awasthi, C. Zhang, W. Ruan, M. A. Schwarz, and R. E. Schwarz, "BMS-754807, a small-molecule inhibitor of insulin-like growth factor-1 receptor/insulin receptor, enhances gemcitabine response in pancreatic cancer," *Molecular Cancer Therapeutics*, vol. 11, no. 12, pp. 2644–2653, 2012.

[61] S. E. Franks, R. A. Jones, R. Briah, P. Murray, and R. A. Moorehead, "BMS-754807 is cytotoxic to non-small cell lung cancer cells and enhances the effects of platinum chemotherapeutics in the human lung cancer cell line A549," *BMC Research Notes*, vol. 9, no. 1, p. 134, 2016.

[62] K. Shitara, T. Doi, O. Nagano et al., "Phase 1 study of sulfasalazine and cisplatin for patients with CD44v-positive gastric cancer refractory to cisplatin (EPOC1407)," *Gastric Cancer*, vol. 20, no. 6, pp. 1004–1009, 2017.

[63] K. Ogihara, E. Kikuchi, S. Okazaki et al., "Sulfasalazine could modulate the CD 44v9- xCT system and enhance cisplatin-induced cytotoxic effects in metastatic bladder cancer," *Cancer Science*, vol. 110, no. 4, pp. 1431–1441, 2019.