Identification of Ferroptosis-Related Hub Genes and Their Association with Immune Infiltration in Chronic Obstructive Pulmonary Disease by Bioinformatics Analysis

Yi-Can Yang¹, Meng-Yu Zhang¹, Jian-Yu Liu¹, Yuan-Yuan Jiang², Xiu-Li Ji³, Yi-Qing Qu²

¹Department of Pulmonary and Critical Care Medicine, Qilu Hospital, Cheleoo College of Medicine, Shandong University; Shandong Key Laboratory of Infectious Respiratory Diseases, Jinan, People’s Republic of China; ²Department of Pulmonary and Critical Care Medicine, Qilu Hospital of Shandong University; Shandong Key Laboratory of Infectious Respiratory Diseases, Jinan, People’s Republic of China; ³Department of Pulmonary Disease, Jinan Traditional Chinese Medicine Hospital, Jinan, People’s Republic of China

Correspondence: Yi-Qing Qu, Department of Pulmonary and Critical Care Medicine, Qilu Hospital, Shandong University, Shandong Key Laboratory of Infectious Respiratory Diseases, Wenhuaxi Road 107#, Jinan, 250012, People’s Republic of China, Tel +86 531 8216 9335, Fax +86 531 8296 7544, Email quyiqing@sdu.edu.cn

Purpose: Ferroptosis and immune infiltration are involved in the pathogenesis of chronic obstructive pulmonary disease (COPD). We aim to identify ferroptosis-related hub genes and analyze their association with immune infiltration in COPD through bioinformatics methods.

Materials and Methods: The mRNA microarray data of GSE38974 were downloaded from Gene Expression Omnibus to obtain differentially expressed genes (DEGs). The DEGs were intersected with ferroptosis-related genes (FRGs) from FerrDb to obtain differentially expressed FRGs. GO and KEGG enrichment and protein–protein interaction (PPI) analyses of differentially expressed FRGs were conducted in R software and STRING database. The key module and hub genes were screened by Cytoscape software. MiRNAs, transcription factors and signal molecules were predicted in miRNet and NetworkAnalyst. The disease correlation in the Comparative Toxicomics Database (CTD) and the receiver operating characteristic (ROC) curves of hub genes were analyzed. Immune infiltration was evaluated by CIBERSORT algorithm. Spearman correlation analyses were conducted between hub genes and differentially infiltrated immune cells.

Results: Fifteen differentially expressed FRGs were identified, which were enriched in some terms involving airway inflammatory responses and structural remodeling. Five hub genes were selected including HIF1A, IL6, PTGS2, CDKN1A and ATM. Inference scores in CTD indicated their association with COPD. Two miRNAs, five transcription factors and one signal molecule were predicted. The combination of hub genes could be a fine diagnostic indicator of COPD (AUC: 0.981, CI: 0.940-1.000). Immune infiltration evaluation showed that monocytes and M0 macrophages were upregulated in COPD lung tissues, while CD8 T cells, activated NK cells, M2 macrophages, resting dendritic cells and resting mast cells were downregulated. The hub genes were significantly associated with differentially infiltrated immune cells.

Conclusion: We identified five ferroptosis-related hub genes (HIF1A, IL6, PTGS2, CDKN1A and ATM) in COPD, and found that they may influence the pathogenesis of COPD by regulating ferroptosis and thus affecting infiltrating immune cells.

Keywords: COPD, ferroptosis, immune infiltration, bioinformatics

Introduction

Chronic obstructive pulmonary disease (COPD) has an increasing rate of incidence in recent years and causes three million deaths annually, which brings about a heavy economic burden.¹ Currently, there are no effective target drugs applied to clinical practice so it is urgent to mine promising drug targets. Airway inflammation is an important feature and contributes to the pathogenesis and progression of COPD.² Ferroptosis refers to the programmed cell death induced...
by lipid peroxidation via iron-dependent pathway with unique morphological and biological features. Usually under environmental stresses or intra/inter-cellular signaling, many metabolic products such as reactive oxygen species (ROS) and phospholipid containing polyunsaturated fatty acid chain(s) (PUFA-PL) can trigger phospholipid peroxidation. Previous study proved that ferroptosis was involved in the pathogenesis of COPD. Stimulated by cigarette smoke, bronchial epithelium produced reactive oxygen species, which induced lipid peroxidation, membrane damage and even ferroptosis. Glutathione peroxidase 4 (GPX4) – a vital antioxidant regulator – is also impaired during ferroptosis. Cigarette smoke extract altered ferroptosis-related genes expression in bronchoalveolar epithelial cells. Hypermethylation of the nuclear factor erythroid 2-related factor 2 (Nrf2) promoter could inhibit Nrf2/GPX4 axis, thus affecting ferroptosis in COPD. Otherwise, many studies suggest that various immune cells play vital roles in chronic airway diseases such as COPD. Innate immune cells, which were enhanced in small airways, modulated airway inflammation and remodeling. Previous study reported that CD8+ T cells enhanced ferroptosis-specific lipid peroxidation in tumor cells. The proliferation of B cells and antibody production was influenced by iron ion regulating the expression of Cyclin E1. Macrophages recognized oxidized phospholipids on the cell surface to clear ferroptosis cells via toll-like receptor 2 (TLR2).

The aim of this research is to identify ferroptosis-related hub genes and their association with immune cells infiltration in COPD lung tissues compared with normal ones. Additionally, we intend to construct interactive networks of hub genes with miRNAs, transcription factors and signal molecules and evaluate the diagnostic values of hub genes.

**Materials and Methods**

**Data Acquisition**

The mRNA expression microarray data of GSE38974, including 23 patients with COPD and 9 normal controls, were extracted from the Gene Expression Omnibus (GEO) datasets. The platform was GPL4133 Agilent-014850 Whole Human Genome Microarray 4x44K G4112F (Feature Number version). Lung tissues from 9 smokers with no evidence of obstructive lung disease and 23 smokers with COPD were examined for mRNA expression. All the clinical information including age, gender, sample source, smoking history, GOLD stage and FVC group was publicly accessible in GEO database (https://www.ncbi.nlm.nih.gov/geo/geo2r/?acc=GSE38974&platform=GPL4133).

The 259 ferroptosis-related genes (FRGs) were downloaded from the FerrDb database.

**Identification of Differentially Expressed FRGs**

The GEO2R is a web tool internally stalled in GEO database specialized for analyzing differentially expressed genes between experimental group and control group. The GEO2R analysis between COPD group and control group was performed on the GEO datasets and the result of differential expression analysis was downloaded for further analysis. The cut-off criteria for differential gene expression were the absolute value of log fold change (FC) >1 and P value <0.05. The gene list of differential expression analysis and FRGs were intersected to obtain the differentially expressed FRGs. The expressions of differentially expressed FRGs were plotted using the R package Complex Heatmap. The expression differences of differentially expressed FRGs between COPD group and normal group in GSE38974 were compared using Kruskal–Wallis test and Dunn’s test. The correlation analysis of differentially expressed FRGs was performed using Spearman’s correlation.

**Gene Ontology (GO) Terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathways Analyses**

The GO and KEGG enrichment analyses were conducted using the R package clusterProfiler. The screening criteria for significant terms were adjusted P values less than 0.05 and q values less than 0.2. Combined with logFC values, the enrichment analyses were performed by calculating the Z-scores using the R package GOplot.
STRING Database and Cytoscape Software
The STRING database is an online tool for analyzing protein–protein interaction. The PPI analysis was carried out using the STRING database.\textsuperscript{21} Cytoscape is a computer software that graphically displays, analyzes and edits the network, which contains multiple plugins. The results were processed and visualized in Cytoscape software (version 3.9.0). The key module was screened by the Molecular Complex Detection (MCODE) plugin. The differentially expressed FRGs were ranked by degrees and the top five genes were considered to be hub genes by cytoHubba plugin.

The Comparative Toxicomics Database
The Comparative Toxicomics Database (CTD) provides integrated information on complex interactions among chemical exposures, genes, proteins and diseases.\textsuperscript{22} In this study, we used it to estimate the inference scores of hub genes in several respiratory tract diseases.

miRNet, NetworkAnalyst and Encyclopedia of RNA Interactomes (ENCORI)
There are multiple online tools to analyze the interaction between non-coding RNAs and genes and predict target molecules. The miRNet is a useful online tool centering around miRNAs and their interacting molecules.\textsuperscript{23} The NetworkAnalyst is an integrated and powerful database for gene expression analysis and construction of interacting networks.\textsuperscript{24} In this article, they were utilized to explore and visualize the networks between hub genes and miRNAs, transcription factors and signal molecules. The ENCORI focuses on predicting RNA interaction.\textsuperscript{25} Here, it was used to predict the upstream molecules lncRNAs targeting screened miRNAs. The screening criterion was set as strict stringency (the number of Ago CLIP-seq experiments is no less than five) and the top three lncRNAs were selected.

The Receiver Operating Characteristic (ROC) Curves of Hub Genes
The logistic regression model of hub genes was constructed using glm function in R software, and the ROC curves were plotted using R package pROC. The ROC curve of each hub gene can help us determine whether its expression has diagnostic value to some extent. The true positive rate (TPR) or sensitivity refers to the number of true positive samples detected divided by the number of all true positive samples. The false-positive rate (FPR) refers to the number of false-positive samples detected divided by the number of all true negative samples. Specificity refers to the number of true negative samples detected divided by the number of all true negative samples. The abscissa represents 1 – specificity and the ordinate represents sensitivity. The area under the curve (AUC) is used to determine the prediction accuracy. The AUC is usually between 0.5 and 1.0. The ROC curve has low/moderate/high accuracy when the AUC is 0.5–0.7/0.7–0.9/ more than 0.9, respectively. The Youden’s index (sensitivity plus specificity minus one) is used to assess the authenticity of the model. As it gets closer to 1.0, the model is much more authentic.

CIBERSORT Algorithm
CIBERSORT is an R/web tool for deconvolution of expression matrices of human immune cell subtypes based on the principle of linear support vector regression.\textsuperscript{26} By the way of the CIBERSORT algorithm, we analyzed the proportions of 22 types of immune cells infiltration in patients with COPD and normal controls. The infiltration differences in patients with COPD and normal controls were compared using Kruskal–Wallis test and Dunn’s test. The Spearman correlation analysis was carried out to show the correlation within differentially infiltrated immune cells and the association between hub genes and differentially infiltrated immune cells.

Statistical Methods
All the statistical calculations were conducted in R software (version 3.6.3). The corresponding R packages were described as above. The statistical significance was marked as follows: ns, p≥0.05; \*p< 0.05; \**p < 0.01; \***p < 0.001. A p.adjust was the corrected p value obtained by the p value correction method; a q-value was an adjusted p-value, taking into account the false discovery rate (FDR). As the p value/p.adjust/q-value is less than 0.0001, scientific notation (exponent, E) is used. For instance, 0.0000267 is written as 2.67E-05.
Results
Identification of Differentially Expressed FRGs
The design of this research was shown in the flow chart (Figure 1). Principal component analysis (PCA) was conducted to show that there was a good degree of clustering between the two groups (Figure 2A). After intersection, 102 genes were obtained (Figure 2B). Under the condition of absolute values of logFC > 1 and P values <0.05, 15 differentially expressed FRGs were discovered including 11 upregulated genes and 4 downregulated genes (Table 1). The 15 differentially expressed FRGs between COPD and normal groups were presented in volcano plot and heatmap (Figure 2C and D). The volcano plot showed the distribution of gene expression between COPD and normal groups. Genes with an adjusted P-value <0.05 and absolute fold-change value > 1 were considered as differentially expressed genes. Each point represented one gene. Red dots indicated significantly upregulated genes and blue dots indicated significantly downregulated genes. The top three upregulated genes included IL6, ATM and TNFAIP3 and the top two downregulated genes included IL33 and TGFBR1. Furthermore, the expression differences of 15 differentially expressed FRGs between the COPD and normal groups in GSE38974 were shown in box plots (Figure 3A). As shown in Figure 3A, 15 ferroptosis-related genes were all significantly differentially expressed between COPD group and normal samples, which was consistent with Figure 2C and D. To explore the expression correlation of these ferroptosis-related genes, correlation analysis was performed. The Spearman correlation analyses of 15 differentially expressed FRGs were shown in heatmap (Figure 3B). As shown in Figure 3B, the expressions of most up-regulating genes were significantly

![Figure 1 Flow chart.](https://doi.org/10.2147/COPD.S348569)
positively correlated with each other, so were the down-regulating genes. Similarly, the expressions of the most up-regulated genes had a significantly negative correlation with those of the four down-regulated genes.

The GO and KEGG Enrichment Analyses of Differentially Expressed FRGs

To analyze the potential biological functions of these differentially expressed ferroptosis-related genes, we carried out GO and KEGG enrichment analyses by way of R software. In total, 739 biological processes (BPs), 21 cellular components (CCs), 21 molecular functions (MFs) and 26 KEGG pathways were enriched. The GO term results exhibited that differentially expressed FRGs were mainly involved in regulation of smooth muscle cell proliferation, membrane microdomain, membrane raft, caveola, cytokine receptor binding, cytokine activity, and transforming growth factor beta receptor binding. The KEGG analysis indicated that differentially expressed FRGs participated in cell senescence pathway, FoxO signaling pathway and HIF1 signaling pathway (Figure 4A and B, Table 2). After combining with expression levels (logFC), the Z-scores showed that all the significant terms could be positively regulated by differentially expressed FRGs (Z-scores >0) (Figure 4C and D). These findings implied that 15 differentially expressed FRGs may participate in inflammatory responses and airway remodeling in COPD pathogenesis.
Construction of Protein–Protein Interaction (PPI) Network and Identification of Key Module and Hub Genes

To determine the interactive relationship among differentially expressed FRGs, the protein–protein interaction analysis was conducted. The interaction of 15 candidate genes was analyzed in STRING database, and the results were visualized in Cytoscape software. The results showed that these differentially expressed FRGs interacted with each other (Figure 5A) and displayed the interaction number of each gene (Figure 5B, Supplementary Table 1). In total, there existed 15 nodes and 134 edges. The MCODE plugin analysis showed that there existed one key module containing 11 nodes and 50 edges including GDF15, IL6, ATF3, PTGS2, TGFBR1, HIF1A, CDKN1A, ATM, HMOX1, TNFAIP3 and MYB (Figure 5C). The cytoHubba plugin analysis identified five hub genes, including HIF1A, IL6, PTGS2, CDKN1A and ATM (Figure 5D). The Venn diagram indicated the overlap of predicted hub genes (Figure 5E). The detailed information of five hub genes can be seen in Supplementary Table 2.

Evaluation of Correlation Between Hub Genes and Respiratory Tract Diseases

In order to estimate the theoretical association between predicted hub genes and chemical/environmental exposures, the five hub genes were analyzed in the Comparative Toxicomics Database and four respiratory tract diseases were chosen including COPD, chronic bronchitis, pulmonary emphysema and non-small cell lung cancer (NSCLC). The average inference scores of five hub genes in COPD (46.85) were higher than those in chronic bronchitis (35.72) and pulmonary emphysema (17.17) but lower than those in NSCLC (55.56) (Figure 6). The findings implied that five predicted hub genes might participate in multiple pathophysiological processes in respiratory diseases.

Construction of the Networks Between Hub Genes with miRNAs, Transcription Factors and Signal Molecules

The hub genes could probably play a role by acting as transcription factors and vital signal molecules or interacting with intracellular non-coding RNAs. In order to predict upstream or downstream molecules of five hub genes and speculate on the mechanism of action of each hub gene, interactive network analysis was conducted. The five hub genes were uploaded to the miRNet online database to analyze the interaction with miRNAs and transcription factors in human lung tissues. In total, 44 miRNAs were predicted and two miRNAs, hsa-let-7b-5p and hsa-miR-1-3p, both targeting five hub genes, were selected for further exploration (Figure 7A, Supplementary Table 3). The transcription factor-hub gene regulatory network consisted of 217 interactions between 164 transcription factors and five hub genes. Five transcription

| Gene Symbol | LogFC | Change | P value |
|-------------|-------|--------|---------|
| IL6         | 2.347 | Up     | 0.001156|
| ATM         | 1.675 | Up     | 3.76E-10|
| TNFAIP3     | 1.617 | Up     | 2.76E-05|
| SLC7A5      | 1.458 | Up     | 0.000313|
| HIF1A       | 1.36  | Up     | 6.89E-06|
| CDKN1A      | 1.273 | Up     | 0.000218|
| PTGS2       | 1.262 | Up     | 0.006918|
| HMOX1       | 1.22  | Up     | 0.002393|
| ATF3        | 1.193 | Up     | 0.008531|
| GDF15       | 1.077 | Up     | 0.001905|
| PML         | 1.062 | Up     | 1.85E-06|
| ENPP2       | −1.047| Down   | 0.002512|
| MYB         | −1.095| Down   | 0.000536|
| TGFBR1      | −1.227| Down   | 5.97E-07|
| IL33        | −1.456| Down   | 3.84E-06|
Figure 3 Expression and correlation of differentially expressed FRGs. (A) Box plot of 15 differentially expressed FRGs in COPD group compared with normal group. The significance markers are shown as: *, P<0.05; **, P<0.01; ***, P<0.001. (B) Heatmap of correlation of 15 differentially expressed FRGs.
factors including EGR1, NFKB1, RELA, SP1 and STAT3, which had the highest connectivity with hub genes, were selected (Figure 7B, Supplementary Table 3). The ENCORI database was used to screen upstream lncRNAs of the two miRNAs. Six lncRNAs were predicted: NUTM2A-AS19, XIST and NEAT1 (targeting hsa-let-7b-5p); RMRP, MALAT1 and AL162431.2 (targeting hsa-miR-1-3p) (Figure 7C, Supplementary Table 3). Next, five hub genes were uploaded into the NetworkAnalyst database to analyze the interaction with signal molecules. TP53 was prominent due to interacting with three hub genes in a network of signal molecules (Figure 7D, Supplementary Table 3).

The ROC Curves of Hub Genes
To determine the diagnostic value in discriminating COPD patients from normal controls, the ROC curves of each hub gene were plotted using R software. The logistic regression model of hub genes was constructed based on glm function. The formula was “-88.166 + 3.0089*HIF1A + −2.8988*IL6 + 2.8957*PTGS2 + 3.2435*CDKN1A + 7.3934*ATM”. As
shown in Figure 8, the expression levels of HIF1A (AUC: 0.923, CI: 0.804-1.00) and ATM (AUC: 0.976, CI: 0.926-1.000) had high predictive accuracy (Figure 8A and E). The expression levels of IL6 (AUC: 0.826, CI: 0.608-1.000) and CDKN1A (AUC: 0.860, CI: 0.653-1.000) had moderate predictive accuracy (Figure 8B and D). The expression level of PTGS2 had low predictive accuracy (AUC: 0.681, CI: 0.471-0.892) (Figure 8C). The AUC of combination of five hub genes was 0.981 (CI: 0.940-1.000) (Figure 8F). When the cut-off threshold was 1.398, the sensitivity, specificity and Youden index were 0.957, 1.000 and 0.957, respectively. These results indicated that this model had high accuracy and authenticity to distinguish COPD group from normal group.

The Immune Cells Infiltration Characteristics in Patients with COPD and Normal Controls

The infiltrating status of various immune cells in lung tissues had obvious differences (Figure 9A). Monocytes and macrophages accounted for the majority of all infiltrating cells, especially in COPD lung tissues. The infiltration differences in both groups are shown in Figure 9B. Seven types of immune cells, including CD8 T cells, activated NK cells, monocytes, M0 macrophages, M2 macrophages, resting dendritic cells and resting mast cells, had differential infiltration in patients with COPD compared with normal controls. The adjusted P-values of seven kinds of immune cells were 0.002, 0.001, 0.025, <0.001, 0.002, 0.008 and 0.020, respectively. Monocytes and M0 macrophages were upregulated in COPD lung tissues, while CD8 T cells, activated NK cells, M2 macrophages, resting dendritic cells and resting mast cells were downregulated. Figure 10A reveals the correlations between differentially infiltrated immune cells. Monocytes had positive correlations with CD8 T cells, M2 macrophages and resting dendritic cells (r = −0.39, −0.67 and −0.46, respectively). M0 macrophages had inverse correlations with CD8 T cells and activated NK cells (r = −0.38 and −0.58, respectively). However, CD8 T cells had positive correlations with M2 macrophages, resting dendritic cells and resting mast cells (r = 0.54, 0.40 and 0.56, respectively). Resting mast cells were positively associated with M2 macrophages and resting dendritic cells (r = 0.36 and 0.61, respectively). The correlations between the expression of hub genes and differentially infiltrated immune cells are displayed in Figure 10B. Positive associations were observed between monocytes and IL6, monocytes and PTGS2, monocytes and CDKN1A (r = 0.53, 0.42 and 0.62, respectively). M0 macrophages were also positively associated with HIF1A and ATM (r = 0.50 and 0.52, respectively). However, CD8

| Ontology | ID       | Description                                      | GeneRatio | p.adjust | q.value |
|----------|----------|--------------------------------------------------|-----------|----------|---------|
| BP       | GO: 0033002 | Muscle cell proliferation                       | 7/15      | 5.12E-07 | 1.95E-07|
| BP       | GO: 0048660 | Regulation of smooth muscle cell proliferation | 6/15      | 1.42E-06 | 5.41E-07|
| BP       | GO: 0048659 | Smooth muscle cell proliferation               | 6/15      | 1.42E-06 | 5.41E-07|
| BP       | GO: 0045765 | Regulation of angiogenesis                      | 7/15      | 3.07E-06 | 1.17E-06|
| BP       | GO: 0070482 | Response to oxygen levels                       | 7/15      | 3.07E-06 | 1.17E-06|
| CC       | GO: 0045121 | Membrane raft                                   | 3/15      | 0.016    | 0.010   |
| CC       | GO: 0098857 | Membrane microdomain                           | 3/15      | 0.016    | 0.010   |
| CC       | GO: 0005901 | Caveola                                          | 2/15      | 0.016    | 0.010   |
| CC       | GO: 0098589 | Membrane region                                 | 3/15      | 0.016    | 0.010   |
| CC       | GO: 1902554 | Serine/threonine protein kinase complex         | 2/15      | 0.016    | 0.010   |
| MF       | GO: 0005125 | Cytokine activity                               | 3/15      | 0.033    | 0.017   |
| MF       | GO: 0005160 | Transforming growth factor beta receptor binding| 2/15      | 0.033    | 0.017   |
| MF       | GO: 0005126 | Cytokine receptor binding                       | 3/15      | 0.033    | 0.017   |
| MF       | GO: 0031625 | Ubiquitin protein ligase binding                | 3/15      | 0.033    | 0.017   |
| MF       | GO: 0044389 | Ubiquitin-like protein ligase binding           | 3/15      | 0.033    | 0.017   |
| KEGG     | hsa04066   | HIF-1 signaling pathway                         | 4/14      | 0.003    | 0.002   |
| KEGG     | hsa04068   | FoxO signaling pathway                          | 4/14      | 0.003    | 0.002   |
| KEGG     | hsa04218   | Cellular senescence                             | 4/14      | 0.004    | 0.003   |
| KEGG     | hsa05202   | Transcriptional misregulation in cancer         | 4/14      | 0.006    | 0.004   |
| KEGG     | hsa05167   | Kaposi sarcoma-associated herpesvirus infection| 4/14      | 0.006    | 0.004   |
Figure 5 PPI network, key module and hub genes of differentially expressed FRGs. (A) The PPI among 15 differentially expressed FRGs. (B) The interaction number of each differentially expressed FRG. (C) Key module of the PPI network screened by MCODE plugin. (D) Hub genes screened by cytoHubba plugin. (E) The overlap of predicted hub genes.
T cells were strongly negatively associated with HIF1A, IL6, PTGS2 and CDKN1A ($r=-0.68$, $-0.83$, $-0.72$ and $-0.79$, respectively). The remaining several types of immune cells also had weakly to moderately negative correlations with the expression of most of the hub genes as displayed in the heatmap.

**Discussion**

Accumulating evidence indicates that ferroptosis participates in the pathogenesis of COPD. Previous review summarized that ferroptosis can affect inflammation through immunogenicity and ferroptosis inhibitors may benefit certain diseases through their anti-inflammatory effects. However, more research is required to better our understanding of ferroptosis in pathogenesis of COPD. In our study, we obtained 15 differentially expressed FRGs in patients with COPD compared with normal controls through bioinformatics analysis. Several hub genes were reported in the previous study. For instance, HIF1A, as a switch gene, was upregulated in COPD cases using network-based analysis implemented by SWIM software. CDKN1A played important functions in the development and progression of COPD. In our study, the enrichment analyses of 15 differentially expressed FRGs were conducted to explore their potential functions. The results indicated that they were associated with airway inflammatory response and remodeling. For example, cell senescence occurs in many pathological processes in COPD, which is consistent with previous reports. Cell senescence impedes iron-mediated cell death pathways by impairing ferritinophagy, a lysosomal process that promotes ferritin degradation.

Next, we constructed a PPI network of 15 differentially expressed FRGs and first identified five ferroptosis-related hub genes, including HIF1A, IL6, PTGS2, CDKN1A and ATM. To further explore the correlation between hub genes and diseases, we analyzed the inference scores for four respiratory tract diseases in CTD and found that five hub genes were closely correlated with COPD and other respiratory tract diseases. These findings reminded us that it was vital to clarify the mechanism of action of these genes in COPD pathogenesis.

Bioinformatics methods provide us with a convenient way to predict crosstalk networks and screen potential biomarkers in COPD. A large number of miRNAs and IncRNAs were reported to be involved in COPD initiation and development. MALAT1/miR-146a/COX2 (namely PTGS2) axis affected the lung function of patients with COPD. Some non-coding RNA targets including miR-195, miR-181c and TUG1 are viable for alleviating COPD in vivo. However, to our knowledge, previous articles reporting the correlation between non-coding RNAs and ferroptosis mainly focused on multiple cancers. The present study constructed the networks of hub genes with miRNAs and transcription factors in miRNet database and identified two key miRNAs, namely, hsa-let-7b-5p and hsa-miR-1-3p. The hsa-let-7b-5p participated in endothelial mitochondrial dynamics and acted as a biomarker for diagnosing Parkinsonian Syndromes. The hsa-miR-1-3p inhibited lung adenocarcinoma cell tumorigenesis and improved gefitinib resistance.
Additionally, the upstream lncRNAs of two miRNAs were predicted using ENCORI database and we found six lncRNAs with the most experimental evidence. Among them, NUTM2A-AS19 and AL162431.2 are newly reported in this study. XIST and MALAT1 played an important role in mitochondrial dysfunction, cell senescence and epigenetic alterations in COPD pathogenesis under the condition of tobacco smoke exposure. NEAT1 promoted activation of inflammasomes in macrophages. RMRP promoted the progression of NSCLC via competing with miR-1-3p. Thus, we speculate that the following axes may regulate ferroptosis in EGFR mutant lung cancer cell. 

Figure 7 The interaction network of hub genes in miRNet and network Analyst. (A) The network of hub genes with miRNAs. (B) The network of hub genes with transcription factors. The fuchsia nodes represent hub genes and the green nodes represent transcription factors. The five transcription factors that connect with at least four hub genes are labelled. (C) The network of hub genes with signal molecules. The signal molecules that connect with at least two hub genes are labelled. (D) The predicted lncRNA-miRNA-hub gene regulatory network. Yellow diamonds represent lncRNAs, green ellipses represent miRNAs, and blue rectangles represent hub genes.
COPD pathogenesis including NUTM2A-AS19 or XIST or NEAT1/hsa-let-7b-5p/hub gene axes and RMRP or MALAT1 or AL162431.2/hsa-miR-1-3p/hub gene axes. Moreover, the identification of five important transcription factors including EGR1, NFKB1, RELA, SP1 and STAT3 would be the groundwork for molecular mechanisms of ferroptosis in COPD pathogenesis. EGR1 was indispensable for MUC5AC expression induced by cigarette smoke in human bronchial epithelial cells. Genetic knockdown of RELA (NFKB subunit) diminished IL6 production in HBE cells. SP1 was crucial for anti-inflammatory molecule IL10 secretion in the phototherapy effect in HBE cells. STAT3 was a vital molecule in regulating the expression of inflammatory cytokines in COPD murine model. Notably, the signal molecule network revealed that TP53 connected with IL6, CDKN1A and ATM, suggesting that TP53 may be a potential driver of COPD towards lung cancer. This indicated that ferroptosis may also participate in COPD-related carcinogenesis.

Previous studies concentrated on the construction of a ferroptosis-related gene model for prognosis in cancer. For example, researchers screened ten ferroptosis-related genes, which served as potential prognostic biomarkers. To testify the diagnostic values of hub genes, we conducted ROC analyses and discovered that each of them varied in predictive accuracy, while combination of five genes could serve as a fine model to distinguish patients with COPD from normal controls (AUC: 0.981, CI: 0.940-1.000).

Although immune infiltration in malignancies keeps attracting the attention of researchers, very few reports explored the immune infiltration in COPD. The spatially confined eosinophil-rich type 2 microenvironments were identified in COPD. The proportion of T cells decreased in the lungs of current smokers and patients with COPD, whereas the proportion of macrophages increased. In our study, we uncovered the immune infiltration status in patients with COPD compared with normal controls. Monocytes were the majority of immune cells in both groups and increased prominently in COPD group. Monocytes, as an essential part of innate immune system, influence human diseases both by direct effects and by differentiating into macrophages. The cytokine response of monocytes to bacteria was compromised in
Figure 9 Immune infiltration of COPD lung tissues compared with normal tissues. (A) Stack bar chart of proportions of the immune cells infiltration. (B) Box plot of proportions of the immune cells infiltration. The significance markers are shown as: ns, P>0.05; *, P<0.05; **, P<0.01; ***, P<0.001.
smoking-induced COPD and thus impaired immune response. Macrophages are plastic in response to various tissue microenvironment and external stimuli. We found that the proportion of M0 macrophages increased markedly, which could serve as reserves ready for polarization stimuli. M1 macrophages primarily take part in pro-inflammatory responses, however, they were not observed to increase remarkably in this study. M2 macrophages, which primarily participate in anti-inflammatory responses, decreased dramatically in COPD group, suggesting that their functions may be undermined in COPD pathogenesis. CD8 T cells were observed to decrease drastically, which was inconsistent with previous researches. The number of IFN-γ-producing CD8+ and CD4+ lymphocytes increased in the lungs of patients with COPD. The frequencies of CD8+ T cell subsets increased observably in patients with COPD compared with normal controls and non-smokers. It may be partial due to the difference between statistics-based bioinformatics methods and flow cytometry assays. Activated NK cells decreased in COPD lung tissues. However, evidence from other studies revealed that the proportions of NK cells increased in BAL fluid of patients with COPD. Another two researches claimed that the number of NK cells in the lung parenchyma of patients with COPD was at the same level as that in the peripheral blood, and bronchoalveolar lavage fluid in healthy smokers. It seems that NK cells in lung tissues may have different effects compared with those from blood and BALF. Moreover, we found resting dendritic cells and resting mast cells, which were the minority of infiltrating immune cells, decreased in patients with COPD. Mast cells can interact with multiple immune cells and structural cells and thereby facilitate inflammatory responses, airway remodeling and angiogenesis. The functions of them might be antagonized by monocytes and M0 macrophages in some degree in COPD lung tissues. The proportions of monocytes and M0 macrophages were positively associated with most of hub genes, whereas CD8 T cells, activated NK cells, M2 macrophages, resting dendritic cells and resting mast cells were negatively associated most of the hub genes. It can be inferred that the functions of monocyte and M0 macrophages may be promoted by these hub genes. The other types of infiltrating cells, for instance, CD8 T cells, were likely to be inhibited by these hub genes. HIF1A was validated to drive ferroptosis in clear cell carcinomas and ATM was essential for promoting ferroptosis. IL6 and PTGS2 were confirmed as the downstream markers of ferroptosis. CDKN1A was required to suppress ferroptosis. We speculate that these high-expressed hub genes in COPD group may get involved in the regulation of ferroptosis in structural cells of pulmonary parenchyma and thus affect the infiltrating immune cells residing in pulmonary interstitium or recruited from peripheral blood, which could lead to differentially histopathological changes in the lungs of patients with COPD. Taken together, the immune cells infiltration contributes to the pathogenesis of COPD in a sophisticated manner and more research is in urgent need to elucidate the situation.
Our study had obvious limitations. The number of cases included in our study was relatively small. Due to the lack of detailed clinical information, correlations between hub genes and clinical characteristics cannot be explored. Another apparent deficiency is that we did not perform basic experiments to validate the expression of hub genes and their correlation with immune cells. For now, our study can provide a theoretical basis for further explorations of ferroptosis-related phenotypes in COPD research.

**Conclusion**
We identified five ferroptosis-related hub genes (HIF1A, IL6, PTGS2, CDKN1A and ATM) in COPD, a combination of which had diagnostic value. Two miRNAs, five transcription factors and one signal molecule were predicted to target these hub genes, and the lncRNA-miRNA-hub gene regulatory network was constructed. Ferroptosis-related hub genes were significantly associated with immune infiltration in the lung tissues of patients with COPD.

**Abbreviations**
AUC, area under the curve; BP, biological process; CC, cellular component; CI, confidence interval; COPD, chronic obstructive pulmonary disease; CTD, Comparative Toxicomics Database; ENCORI, Encyclopedia of RNA Interactomes; FC, fold change; FRG, ferroptosis-related gene; FVC, forced vital capacity; GEO, Gene Expression Omnibus; GO, Gene Ontology; GOLD, Global Initiative for Chronic Obstructive Lung Disease; GPX4, glutathione peroxidase 4; KEGG, Kyoto Encyclopedia of Genes and Genomes; MCODE, Molecular Complex Detection; MF, molecular function; Nrf2, nuclear factor erythroid 2-related factor 2; NSCLC, non-small cell lung cancer; PCA, principal component analysis; PPI, protein–protein interaction; PUFA-PL, phospholipid containing polyunsaturated fatty acid chain(s); ROC, receiver operating characteristic; ROS, reactive oxygen species; TLR2, toll-like receptor 2.

**Ethics Statement**
This study was reviewed by Medical Ethics Committee of Qilu Hospital of Shandong University and exempted from ethical approval due to the usage of human data from the open and public Gene Expression Omnibus database.

**Funding**
This study was supported by the National Natural Science Foundation of China (grant No. 81800039) and the Jinan Clinical Research Center for Prevention and Control Project of Major Respiratory Diseases (grant No. 201912011).

**Disclosure**
The authors report no conflicts of interest in this work.

**References**
1. Rabe KF, Watz H. Chronic obstructive pulmonary disease. *Lancet*. 2017;389(10082):1931–1940. doi:10.1016/S0140-6736(17)31222-9
2. Brightling C, Greening N. Airway inflammation in COPD: progress to precision medicine. *Eur Respir J*. 2019;54(2):1900651. doi:10.1183/13993003.00651-2019
3. Sun Y, Chen P, Zhai B, et al. The emerging role of ferroptosis in inflammation. *Biomed Pharmacother*. 2020;127:110108. doi:10.1016/j.biopha.2020.110108
4. Jiang X, Stockwell BR, Conrad M. Ferroptosis: mechanisms, biology and role in disease. *Nat Rev Mol Cell Biol*. 2021;22(4):266–282. doi:10.1038/s41580-020-00324-8
5. Yoshida M, Minagawa S, Araya J, et al. Involvement of cigarette smoke-induced epithelial cell ferroptosis in COPD pathogenesis. *Nat Commun*. 2019;10(1):3145. doi:10.1038/s41467-019-10991-7
6. Lian N, Zhang Q, Chen J, Chen M, Huang J, Lin Q. The role of ferroptosis in bronchoalveolar epithelial cell injury induced by cigarette smoke extract. *Front Physiol*. 2021;12:751206. doi:10.3389/fphys.2021.751206
7. Su LJ, Zhang JH, Gomez H, et al. Reactive oxygen species-induced lipid peroxidation in apoptosis, autophagy, and ferroptosis. *Oxid Med Cell Longev*. 2019;2019:5080843. doi:10.1155/2019/5080843
8. Yan B, Ai Y, Sun Q, et al. Membrane damage during ferroptosis is caused by oxidation of phospholipids catalyzed by the oxidoreductases POR and CYB5R1. *Mol Cell*. 2021;81(2):355–369. doi:10.1016/j.molcel.2020.11.024
9. Selit TM, Proneth B, Conrad M. Role of GPX4 in ferroptosis and its pharmacological implication. *Free Radic Biol Med*. 2019;133:144–152. doi:10.1016/j.freeradbiomed.2018.09.014
10. Zhang Z, Fu C, Liu J, et al. Hypermethylation of the Nrf2 promoter induces ferroptosis by inhibiting the Nrf2-GPX4 axis in COPD. *Int J Chron Obstruct Pulmon Dis*. 2021;16:3347–3362. doi:10.2147/COPD.S340113

11. Bu T, Wang LF, Yin YQ. How do innate immune cells contribute to airway remodeling in COPD progression?. *Int J Chron Obstruct Pulmon Dis*. 2020;15:107–116. doi:10.2147/COPD.S235054

12. Wang W, Green M, Choi JE, et al. CDS+ T cells regulate tumour ferroptosis during cancer immunotherapy. *Nature*. 2019;569(7755):270–274. doi:10.1038/s41586-019-1170-y

13. Jiang Y, Li C, Wu Q, et al. Iron-dependent histone 3 lysine 9 demethylation controls B cell proliferation and humoral immune responses. *Nat Commun*. 2019;10(1):2935. doi:10.1038/s41467-019-11002-5

14. Luo X, Gong HB, Gao HY, et al. Oxynated phosphatidylserine/ethanolamine navigates phagocytosis of ferroptotic cells by interacting with TLR2. *Cell Death Differ*. 2021;28(6):1971–1989. doi:10.1038/s41418-020-00719-2

15. Ezzie ME, Crawford M, Cho JH, et al. Gene expression networks in COPD: microRNA and mRNA regulation. *Thorax*. 2012;67(2):122–131. doi:10.1136/thoraxjnl-2011-200089

16. Barrett T, Willhite SE, Ledoux P, et al. NCBI GEO: archive for functional genomics data sets–update. *Nucleic Acids Res*. 2013;41(D1):D991–D995. doi:10.1093/nar/gks1193

17. Zhou N, Bao J. FerrDb: a manually curated resource for regulators and markers of ferroptosis and ferroptosis-disease associations. *Database*. 2020;2020:baa021. doi:10.1093/database/baa021

18. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*. 2016;32(18):2847–2849. doi:10.1093/bioinformatics/btw313

19. Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology. The gene ontology consortium. *Nat Genet*. 2000;25(1):21–29. doi:10.1038/75556

20. Kanelitha M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000;28(1):27–30. doi:10.1093/nar/28.1.27

21. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012;16(5):284–287. doi:10.1089/omi.2011.0118

22. Walter W, Sánchez-Cabo F, Ricote M. GOplot: an R package for visually combining expression data with functional analysis. *Bioinformatics*. 2015;31(17):2912–2914. doi:10.1093/bioinformatics/btv300

23. Szklarczyk D, Gable AL, Lyon D, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res*. 2019;47(D1):D607–D613. doi:10.1093/nar/gky1131

24. Davis AP, Grondin CJ, Johnson RJ, et al. Comparative toxicogenomics database (CTD): update 2021. *Nucleic Acids Res*. 2021;49(D1):D1138–D1143. doi:10.1093/nar/gkaa891

25. Chang L, Zhou G, Soufan O, Xia J. miRNA Net 2.0: network-based visual analytics for microRNA functional analysis and systems biology. *Nucleic Acids Res*. 2020;48(W1):W244–W251. doi:10.1093/nar/gszaa467

26. Zhou G, Soufan O, Ewalad J, Hancock REW, Basu N, Xia J. NetworkAnalyst 3.0: a visual analytics platform for comprehensive gene expression profiling and meta-analysis. *Nucleic Acids Res*. 2019;47(W1):W234–W241. doi:10.1093/nar/gkz240

27. Li JH, Liu S, Zhou H, Qu LH, Yang JH. starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res*. 2014;42(D1):D92–D97. doi:10.1093/nar/gkt1248

28. Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. 2015;12(5):453–457. doi:10.1038/nmeth.3337

29. Paci P, Fiscon G, Conte F, et al. Integrated transcriptomic correlation network analysis identifies COPD molecular determinants. *Sci Rep*. 2020;10(1):3361. doi:10.1038/s41598-020-60228-7

30. Yang D, Yan Y, Hu F, Wang T. Death Differ. 2013;20(1):1495. doi:10.1038/s41467-019-09482-6

31. Bu T, Wang LF, Yin YQ. How do innate immune cells contribute to airway remodeling in COPD progression?. *Int J Chron Obstruct Pulmon Dis*. 2020;15:107–116. doi:10.2147/COPD.S235054

32. Sun L, Xu A, Li M, et al. Effect of methylation status of lncRNA-MALAT1 and microRNA-146a on pulmonary function and expression level of COX2 in patients with chronic obstructive pulmonary disease. *Front Cell Dev Biol*. 2021;9:667624. doi:10.3389/fcell.2021.667624

33. Mei D, Tan WSD, Tay Y, Mukhopadhyay A, Wong WSF. Therapeutic RNA strategies for chronic obstructive pulmonary disease. *Trends Pharmacol Sci*. 2020;41(7):475–486. doi:10.1016/j.tips.2020.04.007

34. Hu Q, Li J, Nitta K, et al. FGFR1 is essential for N-acetyl-seryl-aspartyl-lysyl-proline regulation of mitochondrial dynamics by upregulating microRNA let-7b-5p. *J Cell Biochem*. 2020;111(3):2242–2240. doi:10.1002/jcb.2017.12.089

35. Stashko P, Hejl AM, Heegaard NHH, et al. The biomarker potential of cell-free microRNA from cerebrospinal fluid in parkinsonian syndromes. *Mov Disord*. 2019;34(2):246–254.

36. Li T, Wang X, Jing L, Li Y. MiR-1-3p inhibits lung adenocarcinoma cell tumorigenesis via targeting protein regulator of cytokinesis 1. *Front Oncol*. 2019;9:120. doi:10.3389/fonc.2019.00120

37. Jiao D, Chen J, Li Y, et al. miR-1-3p and miR-206 sensitizes HGF-induced gefitinib-resistant human lung cancer cells through inhibition of c-met signalling and EMT. *J Cell Mol Med*. 2018;22(7):3526–3536. doi:10.1111/jcmm.13629

38. Devadoss D, Long C, Langley RJ, et al. Long noncoding transcriptome in chronic obstructive pulmonary disease. *Am J Respir Cell Mol Biol*. 2019;61(6):678–688. doi:10.1165/rcmb.2019-0184TR

39. Zhang P, Cao L, Zhou R, Yang X, Wu M. The IncRNA neat1 promotes activation of inflammasomes in macrophages. *Nat Commun*. 2019;10(1):1495. doi:10.1038/s41467-019-09482-6

40. Wang Y, Luo X, Liu Y, Han G, Sun D. Long noncoding RNA RMRP promotes proliferation and invasion via targeting miR-1-3p in non-small-cell lung cancer. *J Cell Biochem*. 2019;120(9):15170–15181. doi:10.1002/jcb.28779

41. Wang SB, Zhang C, Xu XG, et al. Early growth response factor 1 is essential for cigarette smoke-induced MUC5AC expression in human bronchial epithelial cells. *Biochem Biophys Res Commun*. 2017;490(2):147–154. doi:10.1016/j.bbrc.2016.10.014

42. Wu YF, Li ZY, Dong LL, et al. Inactivation of MTOR promotes autophagy-mediated epithelial injury in particulate matter-induced airway inflammation. *Autophagy*. 2020;16(3):435–450. doi:10.1080/15548627.2019.1628536
43. Brito A, Santos T, Herculano K, et al. The MAPKInase signaling and the stimulatory protein-1 (Sp1) transcription factor are involved in the phototherapy effect on cytokines secretion from human bronchial epithelial cells stimulated with cigarette smoke extract. *Inflammation*. 2021;44(4):1643–1661. doi:10.1007/s10753-021-01448-5

44. Kim SH, Hong JH, Yang WK, et al. Herbal combinational medication of Glycyrrhiza glabra, Agastache rugosa containing glycyrrhizic acid, tiliain inhibits neutrophilic lung inflammation by affecting CXCL2, interleukin-17/STAT3 signal pathways in a murine model of COPD. *Naturis*. 2020;12(4):926. doi:10.3390/n12040926

45. Ren Z, Hu M, Wang Z, et al. Ferroptosis-related genes in lung adenocarcinoma: prognostic signature and immune, drug resistance, mutation analysis. *Front Genet*. 2021;12:672904. doi:10.3389/fgene.2021.672904

46. Jogdand P, Siddhuraj P, Mori M, et al. Eosinophils, basophils and type 2 immune microenvironments in COPD-affected lung tissue. *Eur Respir J*. 2020;55(5):1900110. doi:10.1183/13993003.00110-2019

47. Cruz T, López-Giraldo A, Noell G, et al. Multi-level immune response network in mild-moderate chronic obstructive pulmonary disease (COPD). *Respir Res*. 2019;20(1):152. doi:10.1186/s12931-019-1105-z

48. Kapellos TS, Bonaguro L, Gemünd I, et al. Human monocyte subsets and phenotypes in major chronic inflammatory diseases. *Front Immunol*. 2019;10. doi:10.3389/fimmu.2019.02035

49. Knobloch J, Panek S, Yanik SD, et al. The monocyte-dependent immune response to bacteria is suppressed in smoking-induced COPD. *J Mol Med*. 2019;97(6):817–828. doi:10.1007/s00109-019-01778-w

50. Mark NM, Kargl J, Busch SE, et al. Chronic obstructive pulmonary disease alters immune cell composition and immune checkpoint inhibitor efficacy in non-small lung cancer. *Am J Respir Crit Care Med*. 2018;197(3):325–336. doi:10.1164/rcrm.201704-0795OC

51. Zhuang H, Li N, Chen S, et al. Correlation between level of autophagy and frequency of CD8+ T cells in patients with chronic obstructive pulmonary disease. *J Int Med Res*. 2020;48(9):300060520952638. doi:10.1177/0300060520952638

52. Eriksson Ström J, Pourazar J, Linder R, et al. Cytotoxic lymphocytes in COPD airways: increased NK cells associated with disease, iNKT and NKT-like cells with current smoking. *Respir Res*. 2018;19(1):244. doi:10.1186/s12931-018-0940-7

53. Freeman CM, Stolberg VR, Crudgington S, et al. Human CD56+ cytotoxic lung lymphocytes kill autologous lung cells in chronic obstructive pulmonary disease. *Plos One*. 2014;9(7):e103840. doi:10.1371/journal.pone.0103840

54. Hodge G, Mukaro V, Holmes M, Reynolds PN, Hodge S. Enhanced cytotoxic function of natural killer and natural killer T-like cells associated with decreased CD94 (Kp43) in the chronic obstructive pulmonary disease airway. *Respirology*. 2013;18(2):369–376. doi:10.1111/j.1440-1843.2012.02287.x

55. Freeman CM, Curtis JL. Lung dendritic cells: shaping immune responses throughout chronic obstructive pulmonary disease progression. *Am J Respir Cell Mol Biol*. 2017;56(2):152–159. doi:10.1165/rcmb.2016-0272TR

56. Givi ME, Redegeld FA, Folkerts G, Mortaz E. Dendritic cells in pathogenesis of COPD. *Curr Pharm Des*. 2012;18(16):2329–2335. doi:10.2174/138161212800166068

57. Mortaz E, Folkegels G, Redegeld F. Mast cells and COPD. *Palm Pharmacol Ther*. 2011;24(4):367–372. doi:10.1016/j.pupt.2011.03.007

58. Soltani A, Ewe YP, Lim ZS, et al. Mast cells in COPD airways: relationship to bronchodilator responsiveness and angiogenesis. *Eur Respir J*. 2012;39(6):1361–1367. doi:10.1183/09031936.00084411

59. Zou Y, Palte MJ, Delk AA, et al. A GPX4-dependent cancer cell state underlies the clear-cell morphology and confers sensitivity to ferroptosis. *Nat Commun*. 2019;10(1):1617. doi:10.1038/s41467-019-09277-9

60. Chen PH, Wu J, Ding CC, et al. Kinome screen of ferroptosis reveals a novel role of ATM in regulating iron metabolism. *Cell Death Differ*. 2020;27(3):1008–1022. doi:10.1038/s41418-019-0393-7

61. Linkermann A, Skouta R, Himmerkus N, et al. Synchronized renal tubular cell death involves ferroptosis. *Proc Natl Acad Sci U S A*. 2014;111(47):16836–16841. doi:10.1073/pnas.1415518111