Establishment of a ferroptosis-related gene signature for prognosis in lung adenocarcinoma patients

Jingjing Cai [Equal first author, 1] , **Chunyan Li** [Equal first author, 2] , **Hongsheng Li** [3] , **Xiaoxiong Wang** [1] , **Yongchun Zhou** [Corresp. 3]

[1] Molecular Diagnostics Center, The Third Affiliated Hospital of Kunming Medical University/Yunnan Cancer Hospital, Kunming, Yunnan, China

[2] Department of Head and Neck Surgery Section II, The Third Affiliated Hospital of Kunming Medical University/Yunnan Cancer Hospital, Kunming, Yunnan, China

[3] Molecular Diagnostics center, The Third Affiliated Hospital of Kunming Medical University/Yunnan Cancer Hospital, Kunming, Yunnan, China

Corresponding Author: Yongchun Zhou
Email address: chungui7625@163.com

Objective: Lung cancer is the most common malignancy worldwide and exhibits both high morbidity and mortality. In recent years, scientists have made substantial breakthroughs in the early diagnosis and treatment of lung adenocarcinoma (LUAD), however, patient prognosis still shows vast individual differences. In this study, bioinformatics methods were used to identify and analyze ferroptosis-related genes to establish an effective signature for predicting prognosis in LUAD patients.Methods: The gene expression profiles of LUAD patients with complete clinical and follow-up information were downloaded from two public databases, univariate Cox regression and multivariate Cox regression analysis were used to obtain ferroptosis-related genes for constructing the prognos tic risk model, AUC and calibration plot were used to evaluate the predictive accuracy of the FRGS and nomogram. Results: A total of 74 ferroptosis-related differentially expressed genes (DEGs) were identi fied between LUAD and normal tissues from The Cancer Genome Atlas (TCGA) database. A five- gene panel for prediction of LUAD prognosis was established by multivariate regression and was veri fied using the GSE68465 cohort from the Gene Expression Omnibus (GEO) database. Patients were divided into two different risk groups according to the median risk score of the five genes. Based on Kaplan-Meier (KM) analysis, the OS rate of the high-risk group was markedly worse than that of the low-risk group. We also found that risk score was an independent prognostic indicator. The receiver operating characteristic (ROC) curve showed that the proposed model had good prediction ability. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) functional analyses indicated that risk score was prominently enriched in ferroptosis processes. Moreover, at the score of immune-associated gene sets, significant differences were found between the two risk groups.Conclusions: This study demonstrated that ferroptosis-related gene signatures can be used as a potential predictor for the prognosis of LUAD, thus providing a novel strategy for individualized treatment in LUAD patients.

1 Establishment of a ferroptosis-related gene signature for prognosis in lung
2 adenocarcinoma patients

3 Jingjing Cai [1] Chunyan Li [2] Hongsheng Li [1] Xiaoxiong Wang [1] Yongchun Zhou[1*]

4 1. Molecular Diagnostics Center, The Third A☐liated Hospital of Kunming Medical

5 University/Yunnan Cancer Hospital, Kunming, Yunnan, China

6 2. Department of Head and Neck Surgery Section II, The Third A☐liated Hospital of

7 Kunming Medical University/Yunnan Cancer Hospital, Kunming, Yunnan, China

8 Corresponding Author: Yongchun Zhou

9 Email address: chungui7625@163.com

10 (Jing-jing Cai and Chun-yan Li contributed equally to this work, Chun-yan Li is the equal first author)

12 Abstract

13 Objective:

14   Lung cancer is the most common malignancy worldwide and exhibits both high morbidity and

15 mortality. In recent years, scientists have made substantial breakthroughs in the early diagnosis

16 and treatment of lung adenocarcinoma (LUAD), however, patient prognosis still shows vast

17 individual differences. In this study, bioinformatics methods were used to identify and analyze

18 ferroptosis-related genes to establish an effective signature for predicting prognosis in LUAD

19 patients.

20 Methods:

21   The gene expression profiles of LUAD patients with complete clinical and follow-up

22 information were downloaded from two public databases, univariate Cox regression and

23 multivariate Cox regression analysis were used to obtain ferroptosis-related genes for

24 constructing the prognostic risk model, AUC and calibration plot were used to evaluate the

25 predictive accuracy of the FRGS and nomogram.

26 Results:

50   be no significant improvement in patient prognosis, metastasis risk, or recurrence rate, and five-

51   year survival remains at only 15% (Chen, H., et al.,2019). Therefore, it is crucial to find innovative

52   prognostic models to provide better diagnosis and treatment strategies for LUAD patients.

53   Scientists have recently discovered a novel type of programmed cell death that differs from

54   apoptosis and cell necrosis, called ferroptosis, which depends on iron ions and reactive oxygen

55   species (ROS) to induce lipid peroxide accumulation (Latunde-Dada, G.O.,2017; Stockwell, B.R., et

56   al,2017; Conrad, M., et al, 2018). An increasing body of evidence suggests that ferroptosis is

57   involved in the initiation, progression, and suppression of cancer (Fearnhead, H.O., et al, 2017).

58   Thus, induction of ferroptosis may be an emerging target for the treatment of malignant tumors

59   (Liang, C., et al, 2019; Hassannia, B.,et al,2019), and polyunsaturated fatty acid (PUFA) in

60   phospholipids, redox active iron, and lipid peroxidation (LPO) repair defects, may determine the

61   susceptibility of cancer cells to ferroptosis. Current studies have shown that ferroptosis mainly

62   involves two pathways: GSH / GPX4 pathway and FSP1 / CoQ / NADPH pathway (Dixon S J.,

63   et al, 2012). P53 is the most closely related tumor suppressor gene. It can not only induce

64   apoptosis, but also induce ferroptosis. P53 can inhibit the absorption of cystine by systemxc by

65   inhibiting the transcription of SLC7A11 (Wang S J., et al, 2016), resulting in the inhibition of

66   GSH / GPx4 pathway, the reduction of cell antioxidant capacity and the occurrence of ferroptosis.

67   Jiang et al. (Jiang L., et al, 2015) confirmed that SLC7A11 is a new regulatory target of p53 gene.

68   Some genes, such as SLC7A11 (Ji, X., et al,2018), SLC3A2 (Huang, Y., et al,2005), and

69   STYK1 (Lai, Y., et al,2019), are overexpressed in lung cancer cells and mediate the inhibition of

70   ferroptosis. In addition, some lung cancer drugs have been shown to induce ferroptosis. For

71   example, cisplatin is reported to be an inducer of ferroptosis in non-small-cell lung cancer

72   (NSCLC) A549 cells (Guo, J., et al,2018) and cisplatin and erastin (a type I ferroptosis inducer

73   (FIN)) in combination exhibit synergistic effects on anti-lung cancer cell activity. Sorafenib has

74   also been found to induce ferroptosis in the lung cancer cell line NCI-H460 (Lachaier, E., et al,

75   2014). In addition, previous studies have found that GPX4 had high activity in cells with

76   epithelial mesenchymal transition related gene expression. When using first-line chemotherapy

77   drugs to treat melanoma, breast cancer and lung cancer cell lines, the remaining drug-resistant

78   cancer cells were found to have stem cell like characteristics, mesenchymal like gene expression

79   characteristics and GPX4 dependent characteristics (Hangauer M J., 2017). However, there are

80   almost no studies on ferroptosis-related genes and LUAD prognosis. In this research, we

81   explored the prognostic role of ferroptosis-related genes in LUAD.

82   ## Materials & Methods
83   Databases

84   We used the NCI Genomic Data Commons (https://portal.gdc.cancer.gov) for online analysis and

85   visualization of genomic data from The Cancer Genome Atlas (TCGA) to obtain raw data. We

86   downloaded the TCGA-LUAD Htseq_counts.tsv dataset, which contains 528 tumor samples and

87   57 normal samples, and downloaded related phenotype information (e.g., age, sex, and TNM

88   stage) and corresponding survival information (e.g., survival status and time to latest follow-up).

89   We searched the LUAD gene expression dataset from the Gene Expression Omnibus (GEO;

90   http://www.ncbi.nlm.nih.gov/geo/) database and collected a cohort based on the GPL96

91   platform, resulting in 441 cancer patients with complete clinical information.

92   Collection of ferroptosis-related data

93   FerrDb (http://www.zhounan.org/ferrdb) (Zhou, N. and J. Bao,2020) is a manually collected

94   and curated database for the study of markers and regulators of ferroptosis and the association of

95   ferroptosis disease. At present, ferroptosis-related genes (259) have been found and reported in

96   the literature (Supplementary Table S1).

97   Identification of differentially expressed genes (DEGs)

98   The DEGs were obtained using the "limma" software package(Ritchie ME, et al,2015) and

99   those DEGs with log2FC | ≥ 1 and *Padj*< 0.05 were included in subsequent analyses.

100   Constructing and validating a prognostic ferroptosis-related gene signature

101   The ferroptosis-related genes associated with prognosis in LUAD patients were identified

102   using univariate Cox regression analysis. Genes with an adjusted *P*-value of < 0.05 were

103   included for further analysis. The STRING online database (STRING; http://string-db.org)

104   (v11.0) (Szklarczyk, D., et al, 2019) was used to analyze the interactions among genes with

105   significant differences in univariate analysis, and a protein-protein interaction (PPI) network was

106   constructed. Using Cytoscape (v3.7.1) (Smoot, M.E., et al, 2011) to further visualize the DEGs, the

107   top ten genes were screened by the degree method using the cytoHubba plug-in. Multivariate

108   Cox regression analysis was used to obtain the gene panel for constructing the prognostic risk

109   model, which was determined as risk score = $\sum$(Coefi $\times$ Expi). According to the median

110   valuescalculated using the R software packages "survminer" and "survival", the tumor samples

111   were divided into high-risk or low-risk groups. Kaplan-Meier (KM) survival was further plotted

112   using the "survival" package (Ranstam J, Cook JA, 2017). A time-dependent receiver operating

113   characteristic (ROC) curve drawn with the R package "survivalROC" was used to test the

114   accuracy of the model prediction (Szklarczyk, D., et al, 2019)

115   Nomogram generation

116   The R package "RMS" was used to draw a compound nomogram based on risk scores and

117   clinicopathological characteristics, and consistency between the predicted and actual results was

118   evaluated using a calibration curve.

119   Biological pathway of ferroptosis-related prognostic genes

120   Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment

121   analyses (Ashburner M., et al, 2000)of DEGs in the two risk groups were performed to determine

122   their main biological functions. The "clusterProfiler" package with R language was used to

123   generate corresponding files.Single sample gene set enrichment analysis (ssGSEA) and the "gsva"

124   package (Hänzelmann S., et al, 2013) in R were used for scoring each sample in the two risk

125   groups based on 29 immune-related genes (Supplementary Table S2) sets to explore the

126   correlations among different risk groups and immune status.

127   Statistical analysis

128   Univariate Cox regression, multivariate Cox regression, wilcoxon testand log rank tests

129   were used in this study. Statistical analysis and related figures were generated using R software

130   (v3.5.3) (Diboun, I., et al, 2006). In this study, Padj< 0.05 was regarded as statistically

131   significant.

132   Results

133   Patient characteristics and DEGs

134   Our study flow chart is shown in Fig. 1. In total, there were 528 LUAD patients in the

135   TCGA dataset and 441 LUAD patients in the GSE68465 dataset. Their clinical characteristics

136   are shown in Table 1.

137   We obtained 5 438 DEGs in the TCGA database according to | log2FC | ≥ 1 and Padj< 0.05.

138   We used Venn online analysis to identify and visualize overlapping DEGs in the two databases

139   and downloaded the Venn diagram (Fig. 2a). The heat map and volcano map (Fig. 2b, c) showed

140   48 up-regulated genes and 26 down-regulated genes.The list of DEGs is shown in Supplement 3.

141   PPI analysis of DEGs

142   Using univariate Cox regression analysis, we identified 17 genes related to prognosis in

143   LUAD patients (Fig. 3a), the correlations of which are shown in Fig. 3b. We then imported the

144   above DEGs into the STRING database, and a PPI network was obtained after the lowest

145   confidence was set to 0.015 and genes without interactions were removed (Fig. 3c).

146   Subsequently, we used Cytoscape and cytoHubba to sketch the top 10 genes. SLC7A11,

147   SLC7A11, and GDF15 were determined to be the top hub genes (Fig. 3d).

148   Construction and validation of a prognostic model in TCGA and GEO cohorts

149   We obtained a five-gene panel as a prognostic signature using the multivariate Cox

150   regression model. The forest plot of the five-gene panel is shown in Fig. 4a. The five screened

151   genes were Arachidonate 15-Lipoxygenase (ALOX15), DNA Damage Inducible Transcript 4

152   (DDIT4), Hepatocyte Nuclear Factor 4 Alpha (HNF4A), Interleukin 33 (IL33), and Growth

153   Differentiation Factor 15 (GDF15). Among them, DDIT4 and HNF4A were highly expressed in

154   tumor tissues, whereas ALOX15, IL33, and GDF15 were lowly expressed (Table 2). Risk score

155   = (-0.061165052 × expression ALOX15) + (0.169594473 × expression DDIT4) + (0.072356312

156   × expression HNF4A) + (-0.131023567 × expression IL33) + (-0.107882858 × expression

157   GDF15). Therefore, the risk scores of LUAD patients in the two databases were calculated

158   according to the Cox regression model (composed of five genes).

159   Patients were dichotomized into low-risk score (n = 251) and high-risk score groups (n =

160   250) based on the median cut-off value (Fig. 5a). Compared with the low-risk group, the high-

161   risk group showed a higher risk of death (Fig. 5c). Based on three-dimensional (3D) principal

162   component analysis (PCA), the prognosis model clearly distinguished the LUAD tumor samples

163   of the two risk groups (Fig. 5e). Furthermore, KM curve analysis showed that the OS of the high-

164   risk group was significantly worse than that of the low-risk group (Fig. 5g, $P < 0.001$). The

165   clinical correlation heat map of the two risk groups is shown in Fig. 5i. A time-dependent ROC

166   curve was used to test the predictive performance of the OS risk score composed of the five-gene

167   panel. Results showed that the AUCs of one-year, two-years, and three-years were 0.704, 0.668,

168   and 0.693, respectively (Fig. 5k).

169   We also observed similar survival curves, survival statuses, risk scores, patient distributions,

170   and clinical prognostic characteristics in the validated GSE68465 dataset (Fig. 5b, d, f, h, j).

171   When we performed 1-, 2-, and 3-year ROC curve analyses to assess the predictive capacity of

172   the prognostic five-gene signature, we found the AUCs for 1-, 2-, and 3-year OS predictions

173   were 0.653, 0.644, and 0.617, respectively(Fig. 5l).

174   Nomogram establishment

175   Based on the two cohorts, we constructed a nomogram using clinicopathological

176   characteristics (age, sex, TNM stage, grade) and risk scores, respectively, and calculated the total

177   score of each patient to predict the one-year, two-year, and three-year survival rates of the

178   LUAD patients (Fig. 6a). Further calibration curves showed that the third year OS predicted by

179   the nomogram was in good agreement with actual OS (Fig. 6b).

180   Assessment of independent prognostic value of risk score

181   To evaluate the relationship between risk score and prognosis of the five-gene model, we

182   used risk score as an index and clinicopathological characteristics of LUAD patients for

183   univariate and multivariate Cox regression analyses. Available variables of the TCGA cohort

184   included age, sex, TMN stage, and risk score. The GSE68465 cohort included age, sex, TN stage,

185   and risk score. In the TCGA cohort, the risk score was determined as an independent prognostic

186   factor for OS (HR = 1.809, 95% CI = 1.292-2.535, $P < 0.001$; HR = 1.496, 95% CI = 1.056–

187   2.121, P = 0.023) (Fig. 7a). The GSE68465 cohort results were consistent (HR = 1.625, 95% CI

188   = 1.254–2.105, P< 0.001; HR = 1.425, 95% CI = 1.098–1.850, P = 0.008) (Fig. 7b).

189   Functional enrichment analysis in TCGA and GEO cohorts

190   We used the R language clusterProfiler package to analyze enrichment in biological

191   functions (GO) and KEGG pathways for the 1 285 DEGs in the high- and low-risk groups in the

192   TCGA cohort. In total, 40 important functional annotations (P< 0.05, Fig. 8a) were enriched in

193   molecular functions, among which five were iron related (P< 0.05, Fig. 8a), including a variety

194   of enzymes, gluconosyltransferase and oxidoreductase activity. Unfortunately, the 16

195   significantly enriched pathogenic KEGG pathways (P< 0.05, Fig. 8c) were not associated with

196   ferroptosis. In the GSE68465 validation cohort, multiple functional iron-related molecules were

197   enriched (P < 0.05, Fig. 8b). In addition, one iron-related pathway, namely ferroptosis, was

198   enriched in the KEGG validation cohort (P< 0.05, Fig. 8d).

199   We used ssGSEA to further score the samples from different risk groups in the TCGA and

200   GSE68465 cohorts. Differences in different immune cells, functions, and pathways were

201   detected between the two groups (Fig. 9a-b). The high-risk group in the TCGA cohort showed

202   lower scores in immune-related cells, such as mast cells, neutrophils, dendritic cells (DCs), and T

203   helper cells, with only natural killer (NK) cells showing higher scores (all adjusted P < 0.05, Fig.

204   9a).Similar results were obtained for the GSE68465 cohort (all adjusted P< 0.05, Fig. 9b).For

205   pathways, the high-risk group in the two cohorts showed lower scores for type II and type I IFN

206   responses (all adjusted P< 0.05, Fig. 9b).

207   Discussion

208   While the incidence rate of lung cancer is no longer the top priority, the mortality rate is still

209   first among cancer deaths, accounting for more than 180 000 cases [20].At present, LUAD is still

210   the most common histological subtype of lung cancer, with high mortality among Asians,

211   females, and non-smoking patients (Chen, J., et al, 2020). Although LUAD treatment has made

212   great progress in recent years, most patients with LUAD still exhibit poor prognosis and a low

213   five-year survival rate due to tumor recurrence and metastasis.Therefore, a reliable prognostic

214   biomarker is crucial for evaluating and predicting prognosis in LUAD patients. Recently,

215   bioinformatics analyses have become an important screening tool in cancer research(Huang,

216   C.C.,et al,2019).

217   Most previous studies have focused on the relationship between ferroptosis and

218   tumorigenesis, development, proliferation, and invasion, with only one study by Liang et al.

219   (Liang, J.Y., et al,2020) exploring ferroptosis-related genes and survival rates in patients with

220   hepatocellular carcinoma (HCC). In this study, we used two independent databases and

221   constructed a ferroptosis-related gene panel to predict OS in patients with LUAD. Firstly, we

222   performed univariate Cox regression analysis on 74 ferroptosis-related DEGs in LUAD patients

223   and found that 17 genes were excellent predictors of prognosis, including seven down-regulated

224   and 10 up-regulated genes. Five ferroptosis-related genes (i.e., ALOX15, ddit4, HNF4A, IL33,

225   and GDF15) were further screened by multivariate regression analysis. We then divided patients

226   into high- and low-risk groups according to their median risk score. KM curve analysis showed

227   that the high-risk group was associated with poorer survival compared with the low-risk group.

228   An ROC curve was used to evaluate the prognostic reliability of this signature. The nomogram

229   results indicated that the risk-score model may be an effective method to predict survival status

230   in LUAD patients over different years. We applied GO and KEGG pathway enrichment analysis

231   and found that risk score was significantly correlated with the biological functions and pathways

232   of ferroptosis. SsGSEA further showed that there were significant differences in the immune

233   status of the DEGs between the high-risk group and low-risk group.

234   ALOX15 encodes a member of the lipoxygenase family of proteins,which can regulate

235   inflammation and immune responses. Recent studies have shown that ALOX15 is not only

236   involved in apoptosis, but also in autophagy and ferroptosis (Dixon, S.J., et al, 2012). ALOX15 is

237   involved in ferroptosis through multiple pathways (Stoyanovsky, D.A., et al, 2019), including

238   regulating the activation of Ras selective lethal small molecular 3 (RSL3) (Probst, L., et al,2017),

239   reducing the activation of glutathione (GSH), and forming a complex with

240   phosphatidylethanolamine binding protein 1(PEBP1) (Wenzel, S.E., et al,2017). Mounting

241   evidence indicates that ALOX15 is down-regulated in many human cancers, including colorectal

242   (Shureiqi, I., et al, 2000), prostate (Tang, S., et al, 2002), breast (Jiang, W.G., et al, 2003) and lung

243   cancers(Gonzalez, A.L., et al,2004). Gonzalez et al. (Gonzalez, A.L., et al, 2004) also reported that

244   the expression of 15-LOX-2 is higher in better differentiated NSCLC and is negatively correlated

245   with tumor grade and tumor cell proliferation. However, whether ALOX15 participates in the

246   occurrence of NSCLC by targeting ferroptosis remains unclear. Previous studies have indicated

247   that the use of 12/15-LOX inhibitors or the silencing of ALOX15 expression can prevent cancer

248   cells (including Calu-1 human NSCLC) with RAS expression from cell death in erastin- and

249   RSL3-induced ferroptosis (Shintoku, R., et al, 2017).

250   DDIT4 is a stress-response protein whose main function is to inhibit mTOR under stressful

251   conditions.DDIT4 is considered an oncogene (Smith, E.R.,X.X. Xu;2009), and its overexpression

252   is significantly associated with poorer prognosis in tumor patients(Tirado-Hurtado, I., et al,2018).

253   Jin et al. (Jin, H.O., et al, 2019) reported that constitutive overexpression of DDIT4 can lead to

254   HSP27 and HSP70 induction and AKT activation. This mechanism is related to lung cancer cell

255   survival and IR resistance, indicating that DDIT4 may be a therapeutic target for lung cancer.

256   GDF15 is a member of the transforming growth factor-beta superfamily, and its association

257   with cancer can depend on cell state and tumor environment. Recent study suggests that GDF15

258   is lowly expressed in NSCLC, and its down-regulation is associated with poor prognosis in such

259   patients (Lu, X., et al, 2018).GDF15 is also suggested to inhibit the growth and bone metastasis of

260   LUAD A549 cells by targeting the TGF-β/Smad signaling pathway (Duan, L., et al,

261   2019).However, GDF15 is up-regulated in some malignant tumors, such as gastric cancer in

262   humans. For example, Chen et al. (Chen, L., et al, 2020) recently reported on the role of GDF15 in

263   gastric cancer cell (MGC803) ferroptosis.GDF15 also plays an important role in erastin-induced

264   ferroptosis by affecting the function of system Xc and regulating the expression of SLC7A11

265   (Chen, L., et al, 2020).However, the role of GDF15 in LUAD ferroptosis has not yet been reported

266   and needs to be further elucidated.

267   HNF4A is a nuclear transcription factor, which is highly expressed in most cancers and is

268   significantly associated with poor prognosis. Moreover, HNF4A is reported to play a role in

269   ferroptosis (Dai, C., et al, 2020). Research has indicated that HNF4A expression is up-regulated

270   in HCC(Dai, C., et al,2020), which can increase the synthesis of GSH and inhibit ferroptosis by

271   up regulating the expression of STMN1 (a ferroptosis down-regulated factor) (Xu, L., et al,

272   2001).In lung cancer, activation of GSH biosynthesis-related genes can also lead to the inhibition

273   of ferroptosis (Zhang, X., et al,2019).However, whether and how dysregulated expression of

274   HNF4A regulates GSH production in LUAD remain to be explored.

275   Martin-Sanchez, D., et al. (Martin-Sanchez, D., et al,2017) showed that IL-33 release is related

276   to ferroptosis, and ferroptosis in acute kidney injury may regulate inflammation by activating IL-

277   33.However, the relationship between ferroptosis in cancer and IL-33 is not clear.Previous

278   studies have reported that CD8 + T cells induce ferroptosis of tumor cells *in vivo*(Wang, W., et al,

279   2019; Tang, R., et al,2020) Our study also explored the relationship between risk score and

280   immune activity, but the mechanism between ferroptosis-related genes of LUAD and tumor

281   immunity needs to be further clarified.

282   The five genes included in our model play important roles in the occurrence and

283   development of LUAD, and most are related to ferroptosis in malignant tumors. Therefore, this

284   model could be a useful prognostic indicator of LUAD. However, our research has some

285   limitations that most bioinformatics analysis studies share. Firstly, all our data are from public

286   databases, so it will be necessary to verify the prognostic value of the model in clinical samples.

287   Secondly, this study failed to explore the underlying molecular mechanism of ferroptosis-related

288   genes in the occurrence and development of LUAD.

289   Conclusions

290   We developed a prognosis signature of five ferroptosis-related genes (ALOX15, DDIT4,

291   HNF4A, IL33, GDF15), which showed good reliability. Analysis of two independent databases

292   demonstrated that this model was independently related to OS in LUAD patients, and thus may

293   be a good predictor of LUAD prognosis.

294   Abbreviations

295   LUAD Lung adenocarcinoma;

296   UCSC The University of California, Santa Cruz

297   TCGA The Cancer Genome Atlas;

298   GEO Gene Expression Omnibus

299   DEGs Differentially expressed genes

300   OS Overall survival

301   ROC Receiver operating characteristic

302   GO Gene Ontology

303   KEGG Kyoto Encyclopedia of Genes and Genomes

304   SsGSEA Single-sample gene set enrichmentanalysis

305   3D PCA Principal component analysis

306   HR Hazard ratio

307   CI Confidence interval

308   DCs Dendritic Cells

309   K–M   Kaplan–Meier

310   ALOX15   The five screened genes were Arachidonate 15-Lipoxygenase

311   DDIT4   DNA Damage Inducible Transcript 4

312   HNF4A   Hepatocyte Nuclear Factor 4Alpha

313   IL33   Interleukin 33

314   GDF15   Growth Differentiation Factor

315   References

316   Siegel, R.L., K.D. Miller, and A. Jemal, Cancer statistics, 2020. CA Cancer J Clin, 2020. 70(1):

317   p. 7-30.

318   Testa, U., G. Castelli, and E. Pelosi, Lung Cancers: Molecular Characterization, Clonal

319   Heterogeneity and Evolution, and Cancer Stem Cells. Cancers (Basel), 2018. 10(8).

320   Eberhardt, W.E. and M. Stuschke, Multimodal treatment of non-small-cell lung cancer. Lancet,

321   2015. 386(9998): p. 1018-20.

322   Chen H, Carrot-Zhang J, Zhao Y, Hu H, Freeman SS, Yu S, Ha G, Taylor AM, Berger AC,

323   Westlake L, Zheng Y, Zhang J, Ramachandran A, Zheng Q, Pan Y, Zheng D, Zheng S,

324   Cheng C, Kuang M, Zhou X, Zhang Y, Li H, Ye T, Ma Y, Gao Z, Tao X, Han H, Shang J,

325   Yu Y, Bao D, Huang Y, Li X, Zhang Y, Xiang J, Sun Y, Li Y, Cherniack AD, Campbell

326   JD, Shi L, Meyerson M., Genomic and immune profiling of pre-invasive lung

327   adenocarcinoma. Nat Commun, 2019. 10(1): p. 5472.

328   Latunde-Dada, G.O., Ferroptosis: Role of lipid peroxidation, iron and ferritinophagy. Biochim

329   Biophys Acta Gen Subj, 2017. 1861(8): p. 1893-1900.

330   Stockwell BR, Friedmann Angeli JP, Bayir H, Bush AI, Conrad M, Dixon SJ, Fulda S, Gascón S,

331   Hatzios SK, Kagan VE, Noel K, Jiang X, Linkermann A, Murphy ME, Overholtzer M,

332   Oyagi A, Pagnussat GC, Park J, Ran Q, Rosenfeld CS, Salnikow K, Tang D, Torti FM,

333   Torti SV, Toyokuni S, Woerpel KA, Zhang DD., Ferroptosis: A Regulated Cell Death

334   Nexus Linking Metabolism, Redox Biology, and Disease. Cell, 2017. 171(2): p. 273-285.

335 Conrad M, Kagan VE, Bayir H, Pagnussat GC, Head B, Traber MG, Stockwell BR., Regulation

336   of lipid peroxidation and ferroptosis in diverse species. Genes Dev, 2018. 32(9-10): p.

337   602-619.

338 Fearnhead, H.O., P. Vandenabeele, and T. Vanden Berghe, How do we fit ferroptosis in the

339   family of regulated cell death? Cell Death Differ, 2017. 24(12): p. 1991-1998.

340 Liang C, Zhang X, Yang M, Dong X., Recent Progress in Ferroptosis Inducers for Cancer

341   Therapy. Adv Mater, 2019. 31(51): p. e1904197.

342 Hassannia, B., P. Vandenabeele, and T. Vanden Berghe, Targeting Ferroptosis to Iron Out

343   Cancer. Cancer Cell, 2019. 35(6): p. 830-849.

344 Dixon SJ, Lemberg KM, Lamprecht MR, Skouta R, Zaitsev EM, Gleason CE, Patel DN, Bauer

345   AJ, Cantley AM, Yang WS, Morrison B 3rd, Stockwell BR., Ferroptosis: an iron-

346   dependent form of nonapoptotic cell death[J].Cell,2012,149(5):1060-1072.

347 Wang SJ, Li D, Ou Y, Jiang L, Chen Y, Zhao Y, Gu W. , Acetylation is crucial for p53-mediated

348   ferroptosis and tumor suppression[J].Cell Rep,2016,17(2):366-373.

349 Jiang L, Kon N, Li T, Wang SJ, Su T, Hibshoosh H, Baer R, Gu W., Ferroptosis as a p53-

350   mediated activity during tumour suppression[J].Nature,2015,520 (7545):57-62.

351 Ji X, Qian J, Rahman SMJ, Siska PJ, Zou Y, Harris BK, Hoeksema MD, Trenary IA, Heidi C,

352   Eisenberg R, Rathmell JC, Young JD, Massion PP. , xCT (SLC7A11)-mediated

353   metabolic reprogramming promotes non-small cell lung cancer progression. Oncogene,

354   2018. 37(36): p. 5007-5019.

355   Huang Y, Dai Z, Barbacioru C, Sadée W., Cystine-glutamate transporter SLC7A11 in cancer

356   chemosensitivity and chemoresistance. Cancer Res, 2005. 65(16): p. 7446-54.

357   Lai Y, Zhang Z, Li J, Li W, Huang Z, Zhang C, Li X, Zhao J. , STYK1/NOK correlates with

358   ferroptosis in non-small cell lung carcinoma. Biochem Biophys Res Commun, 2019.

359   519(4): p. 659-666.

360   Guo J, Xu B, Han Q, Zhou H, Xia Y, Gong C, Dai X, Li Z, Wu G., Ferroptosis: A Novel Anti-

361   tumor Action for Cisplatin. Cancer Res Treat, 2018. 50(2): p. 445-460.

362   Lachaier E, Louandre C, Godin C, Saidak Z, Baert M, Diouf M, Chauffert B, Galmiche A. ,

363   Sorafenib induces ferroptosis in human cancer cell lines originating from different solid

364   tumors. Anticancer Res, 2014. 34(11): p. 6417-22.

365   Hangauer MJ, Viswanathan VS, Ryan MJ, Bole D, Eaton JK, Matov A, Galeas J, Dhruv HD,

366   Berens ME, Schreiber SL, McCormick F, McManus MT., Drug-tolerant persister cancer

367   cells are vulnerable to GPX4 inhibition[J].Nature,2017,551(7679):247-250.

368   Zhou, N. and J. Bao, FerrDb: a manually curated resource for regulators and markers of

369   ferroptosis and ferroptosis-disease associations. Database (Oxford), 2020. 2020.

370   Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK., limma powers differential

371   expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res.

372   2015 Apr 20;43(7):e47.

373   Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva

374   NT, Morris JH, Bork P, Jensen LJ, Mering CV. , STRING v11: protein-protein

375   association networks with increased coverage, supporting functional discovery in

376   genome-wide experimental datasets. Nucleic Acids Res, 2019. 47(D1): p. D607-d613.

377   Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T. , Cytoscape 2.8: new features for data

378   integration and network visualization. Bioinformatics, 2011. 27(3): p. 431-2.

379   Ranstam J, Cook JA. Kaplan-Meier curve. Br J Surg. 2017 Mar;104(4):442.

380   Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K,

381   Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese

382   JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. , Gene ontology: tool for the

383   unification of biology. The Gene Ontology Consortium. Nat Genet. 2000 May;25(1):25-9.

384   Hänzelmann S, Castelo R, Guinney J., GSVA: gene set variation analysis for microarray and

385   RNA-seq data. BMC Bioinformatics. 2013 Jan 16;14:7.

386   Diboun I, Wernisch L, Orengo CA, Koltzenburg M., Microarray analysis after RNA

387   amplification can detect pronounced differences in gene expression using limma. BMC

388   Genomics, 2006. 7: p. 252.

389   https://www.iarc.fr/faq/latest-global-cancer-data-2020-qa/.

390   Chen J, Yang H, Teo ASM, Amer LB, Sherbaf FG, Tan CQ, Alvarez JJS, Lu B, Lim JQ, Takano

391   A, Nahar R, Lee YY, Phua CZJ, Chua KP, Suteja L, Chen PJ, Chang MM, Koh TPT,

392   Ong BH, Anantham D, Hsu AAL, Gogna A, Too CW, Aung ZW, Lee YF, Wang L, Lim

393   TKH, Wilm A, Choi PS, Ng PY, Toh CK, Lim WT, Ma S, Lim B, Liu J, Tam WL,

394   Skanderup AJ, Yeong JPS, Tan EH, Creasy CL, Tan DSW, Hillmer AM, Zhai W.

395   Genomic landscape of lung adenocarcinoma in East Asians. Nat Genet, 2020. 52(2): p.

396   177-186.

397   Huang, C.C., M. Du, and L. Wang, Bioinformatics Analysis for Circulating Cell-Free DNA in

398   Cancer. Cancers (Basel), 2019. 11(6).

399 Liang JY, Wang DS, Lin HC, Chen XX, Yang H, Zheng Y, Li YH. A Novel Ferroptosis-related

400   Gene Signature for Overall Survival Prediction in Patients with Hepatocellular

401   Carcinoma. Int J Biol Sci, 2020. 16(13): p. 2430-2441.

402 Dixon SJ, Lemberg KM, Lamprecht MR, Skouta R, Zaitsev EM, Gleason CE, Patel DN, Bauer

403   AJ, Cantley AM, Yang WS, Morrison B 3rd, Stockwell BR. Ferroptosis: an iron-

404   dependent form of nonapoptotic cell death. Cell, 2012. 149(5): p. 1060-72.

405 Stoyanovsky DA, Tyurina YY, Shrivastava I, Bahar I, Tyurin VA, Protchenko O, Jadhav S,

406   Bolevich SB, Kozlov AV, Vladimirov YA, Shvedova AA, Philpott CC, Bayir H, Kagan

407   VE. Iron catalysis of lipid peroxidation in ferroptosis: Regulated enzymatic or random

408   free radical reaction? Free Radic Biol Med, 2019. 133: p. 153-161.

409 Probst L, Dächert J, Schenk B, Fulda S. Lipoxygenase inhibitors protect acute lymphoblastic

410   leukemia cells from ferroptotic cell death. Biochem Pharmacol, 2017. 140: p. 41-52.

411 Wenzel SE, Tyurina YY, Zhao J, St Croix CM, Dar HH, Mao G, Tyurin VA, Anthonymuthu TS,

412   Kapralov AA, Amoscato AA, Mikulska-Ruminska K, Shrivastava IH, Kenny EM, Yang

413   Q, Rosenbaum JC, Sparvero LJ, Emlet DR, Wen X, Minami Y, Qu F, Watkins SC,

414   Holman TR, VanDemark AP, Kellum JA, Bahar I, Bayır H, Kagan VE. PEBP1 Wardens

415   Ferroptosis by Enabling Lipoxygenase Generation of Lipid Death Signals. Cell, 2017.

416   171(3): p. 628-641.e26.

417 Shureiqi I, Chen D, Lee JJ, Yang P, Newman RA, Brenner DE, Lotan R, Fischer SM, Lippman

418   SM. 15-LOX-1: a novel molecular target of nonsteroidal anti-inflammatory drug-induced

419   apoptosis in colorectal cancer cells. J Natl Cancer Inst, 2000. 92(14): p. 1136-42.

420 Tang S, Bhatia B, Maldonado CJ, Yang P, Newman RA, Liu J, Chandra D, Traag J, Klein RD,

421   Fischer SM, Chopra D, Shen J, Zhau HE, Chung LW, Tang DG.Evidence that

422  arachidonate 15-lipoxygenase 2 is a negative cell cycle regulator in normal prostate

423  epithelial cells. J Biol Chem, 2002. 277(18): p. 16189-201.

424 Jiang, W.G., A. Douglas-Jones, and R.E. Mansel, Levels of expression of lipoxygenases and

425  cyclooxygenase-2 in human breast cancer. Prostaglandins Leukot Essent Fatty Acids,

426  2003. 69(4): p. 275-81.

427 Gonzalez AL, Roberts RL, Massion PP, Olson SJ, Shyr Y, Shappell SB. 15-Lipoxygenase-2

428  expression in benign and neoplastic lung: an immunohistochemical study and correlation

429  with tumor grade and proliferation. Hum Pathol, 2004. 35(7): p. 840-9.

430 Shintoku R, Takigawa Y, Yamada K, Kubota C, Yoshimoto Y, Takeuchi T, Koshiishi I, Torii S,

431  Lipoxygenase-mediated generation of lipid peroxides enhances ferroptosis induced by

432  erastin and RSL3. Cancer Sci, 2017. 108(11): p. 2187-2194.

433 Smith, E.R. and X.X. Xu, REDD1, a new Ras oncogenic effector. Cell Cycle, 2009. 8(5): p. 675-

434  6.

435 Tirado-Hurtado, I., W. Fajardo, and J.A. Pinto, DNA Damage Inducible Transcript 4 Gene: The

436  Switch of the Metabolism as Potential Target in Cancer. Front Oncol, 2018. 8: p. 106.

437 Jin HO, Hong SE, Kim JY, Kim MR, Chang YH, Hong YJ, Lee JK, Park IC, Induction of

438  HSP27 and HSP70 by constitutive overexpression of Redd1 confers resistance of lung

439  cancer cells to ionizing radiation. Oncol Rep, 2019. 41(5): p. 3119-3126.

440 Lu X, He X, Su J, Wang J, Liu X, Xu K, De W, Zhang E, Guo R, Shi YE, EZH2-Mediated

441  Epigenetic Suppression of GDF15 Predicts a Poor Prognosis and Regulates Cell

442  Proliferation in Non-Small-Cell Lung Cancer. Mol Ther Nucleic Acids, 2018. 12: p. 309-

443  318.

444   Duan L, Pang HL, Chen WJ, Shen WW, Cao PP, Wang SM, Liu LL, Zhang HL., The role of

445   GDF15 in bone metastasis of lung adenocarcinoma cells. Oncol Rep, 2019. 41(4): p.

446   2379-2388.

447   Chen L, Qiao L, Bian Y, Sun X. , GDF15 knockdown promotes erastin-induced ferroptosis by

448   decreasing SLC7A11 expression. Biochem Biophys Res Commun, 2020. 526(2): p. 293-

449   299.

450   Dai C, Chen X, Li J, Comish P, Kang R, Tang D. Transcription factors in ferroptotic cell death.

451   Cancer Gene Ther. 2020 Sep;27(9):645-656.

452   Xu L, Hui L, Wang S, Gong J, Jin Y, Wang Y, Ji Y, Wu X, Han Z, Hu G. Expression profiling

453   suggested a regulatory role of liver-enriched transcription factors in human hepatocellular

454   carcinoma. Cancer Res. 2001 Apr 1;61(7):3176-81.

455   Zhang X, Du L, Qiao Y, Zhang X, Zheng W, Wu Q, Chen Y, Zhu G, Liu Y, Bian Z, Guo S,

456   Yang Y, Ma L, Yu Y, Pan Q, Sun F, Wang J. Ferroptosis is governed by differential

457   regulation of transcription in liver cancer. Redox Biol. 2019 Jun;24:101211.

458   Alvarez SW, Sviderskiy VO, Terzi EM, Papagiannakopoulos T, Moreira AL, Adams S, Sabatini

459   DM, Birsoy K, Possemato R. NFS1 undergoes positive selection in lung tumours and

460   protects cells from ferroptosis. Nature. 2017 Nov 30;551(7682):639-643.

461   Martin-Sanchez D, Ruiz-Andres O, Poveda J, Carrasco S, Cannata-Ortiz P, Sanchez-Niño MD,

462   Ruiz Ortega M, Egido J, Linkermann A, Ortiz A, Sanz AB. Ferroptosis, but Not

463   Necroptosis, Is Important in Nephrotoxic Folic Acid-Induced AKI. J Am Soc Nephrol.

464   2017 Jan;28(1):218-229.

465   Wang W, Green M, Choi JE, Gijón M, Kennedy PD, Johnson JK, Liao P, Lang X, Kryczek I,

466   Sell A, Xia H, Zhou J, Li G, Li J, Li W, Wei S, Vatan L, Zhang H, Szeliga W, Gu W, Liu

467   R, Lawrence TS, Lamb C, Tanno Y, Cieslik M, Stone E, Georgiou G, Chan TA,

468   Chinnaiyan A, Zou W. CD8+ T cells regulate tumour ferroptosis during cancer

469   immunotherapy. Nature. 2019 May;569(7755):270-274.

470   Tang R, Xu J, Zhang B, Liu J, Liang C, Hua J, Meng Q, Yu X, Shi S. Ferroptosis, necroptosis,

471   and pyroptosis in anticancer immunity. J Hematol Oncol. 2020 Aug 10;13(1):110.

Figure 1

Figure 1. Flow diagram of the bioinformatic analysis in this study
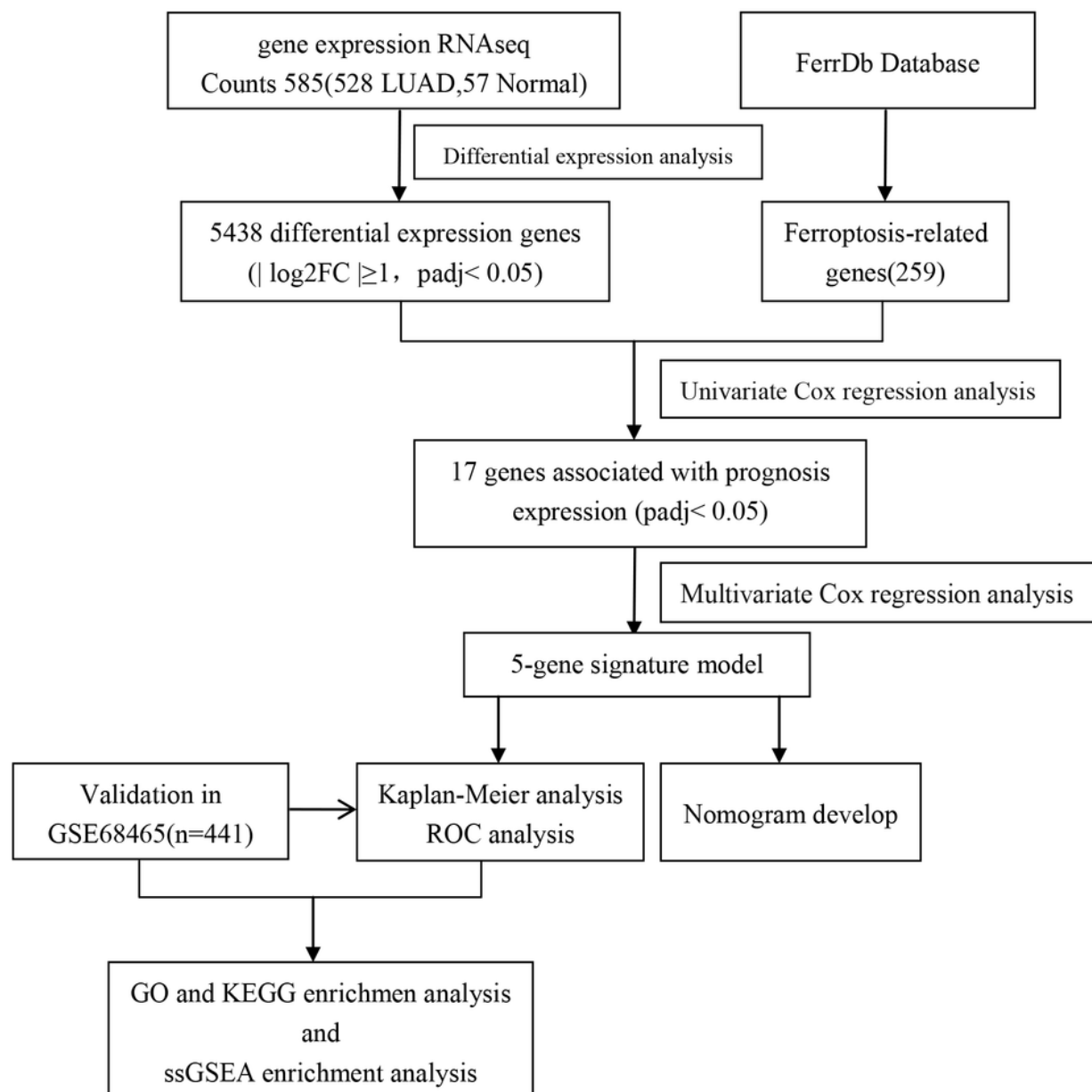
Figure 2

Figure 2. The differentially ferroptosis-related genes of LUAD in the TCGA database and FerrDb database.
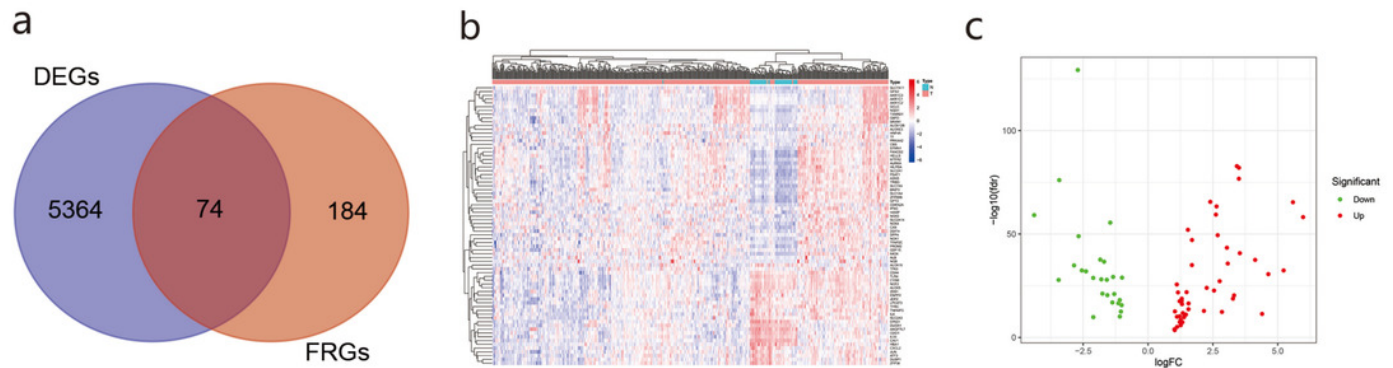
Figure 3

Figure 2. The differentially ferroptosis-related genes of LUAD in the TCGA database and FerrDb database.
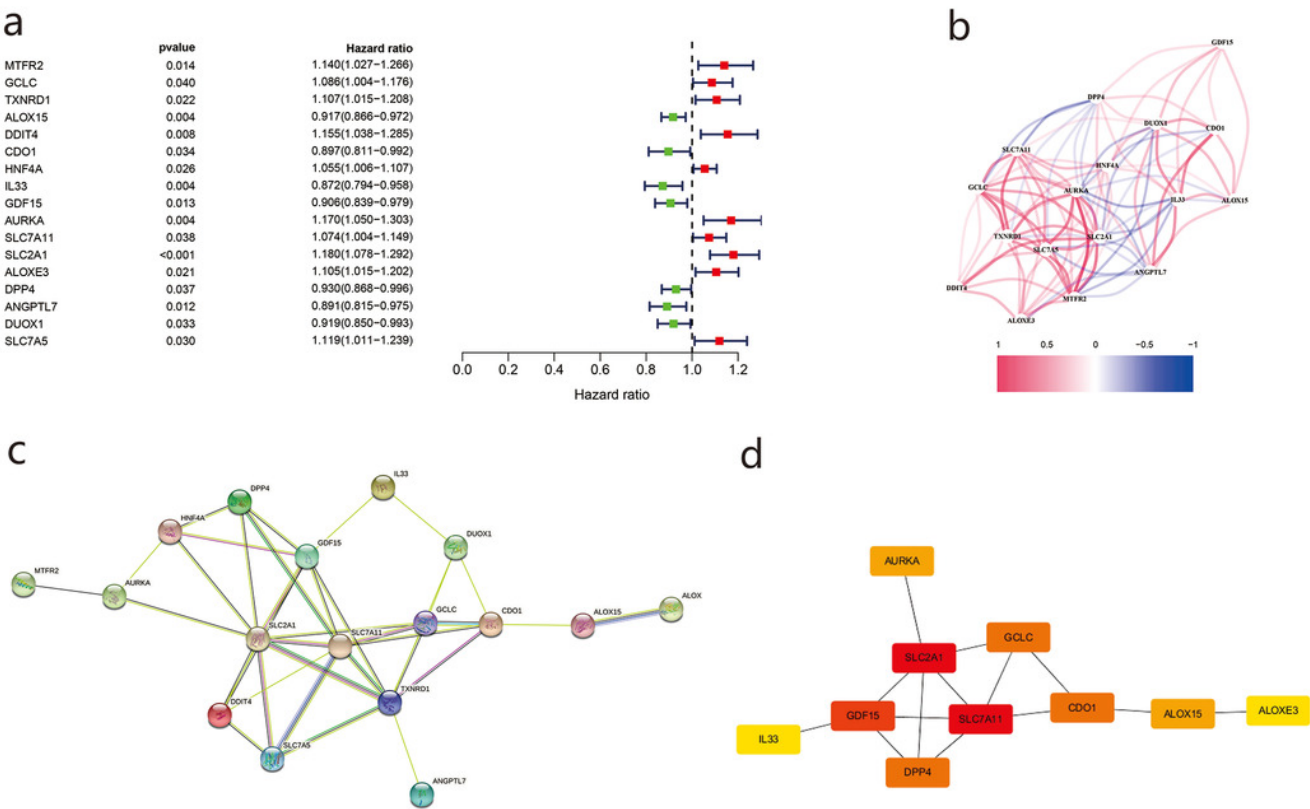
Figure 4

Fig.4 Forest plot of multivariate Cox proportional hazards regression analysis of overall survival for 5-gene signature model in the TCGA cohort.
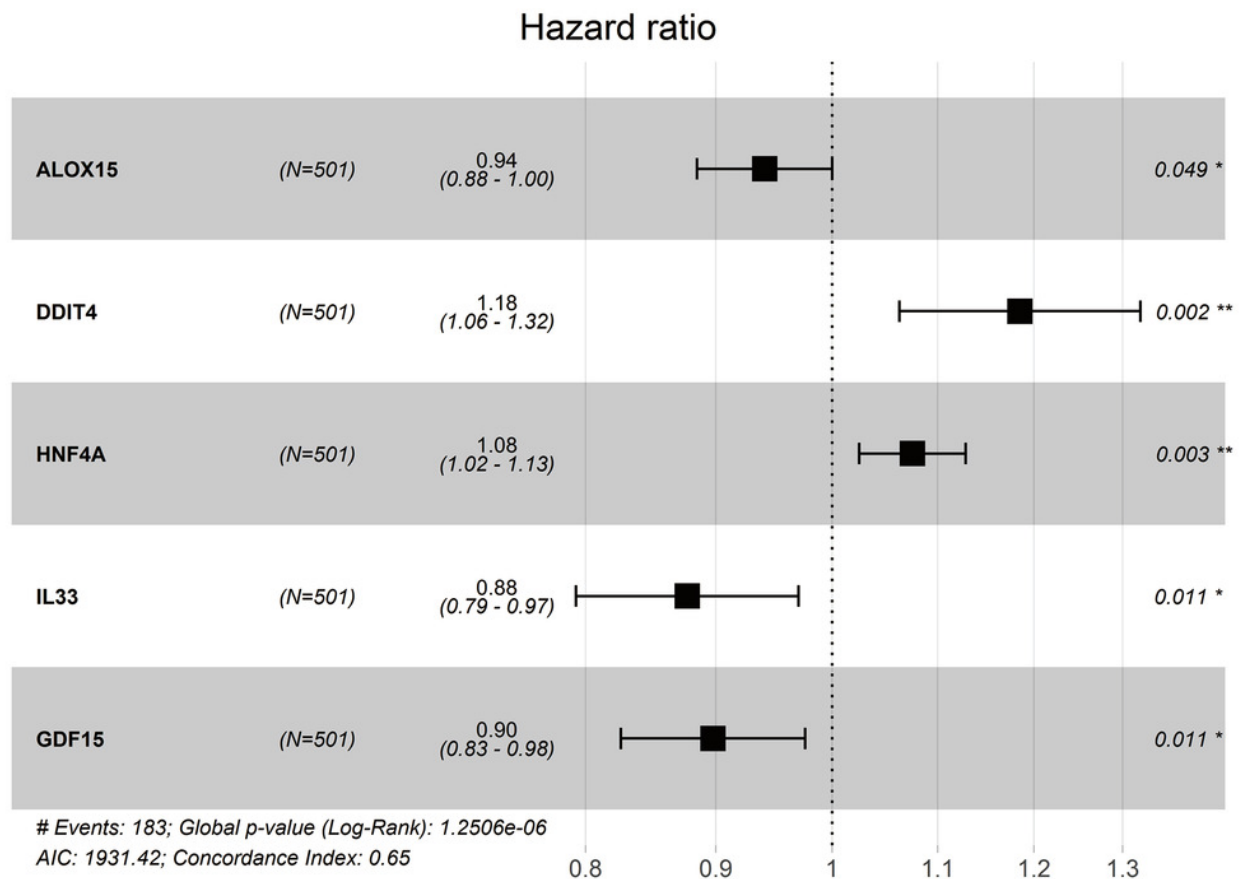
Figure 5

Figure 5. Construction and predictive accuracy in different risk models with TCGA and GSE68465 dataset.
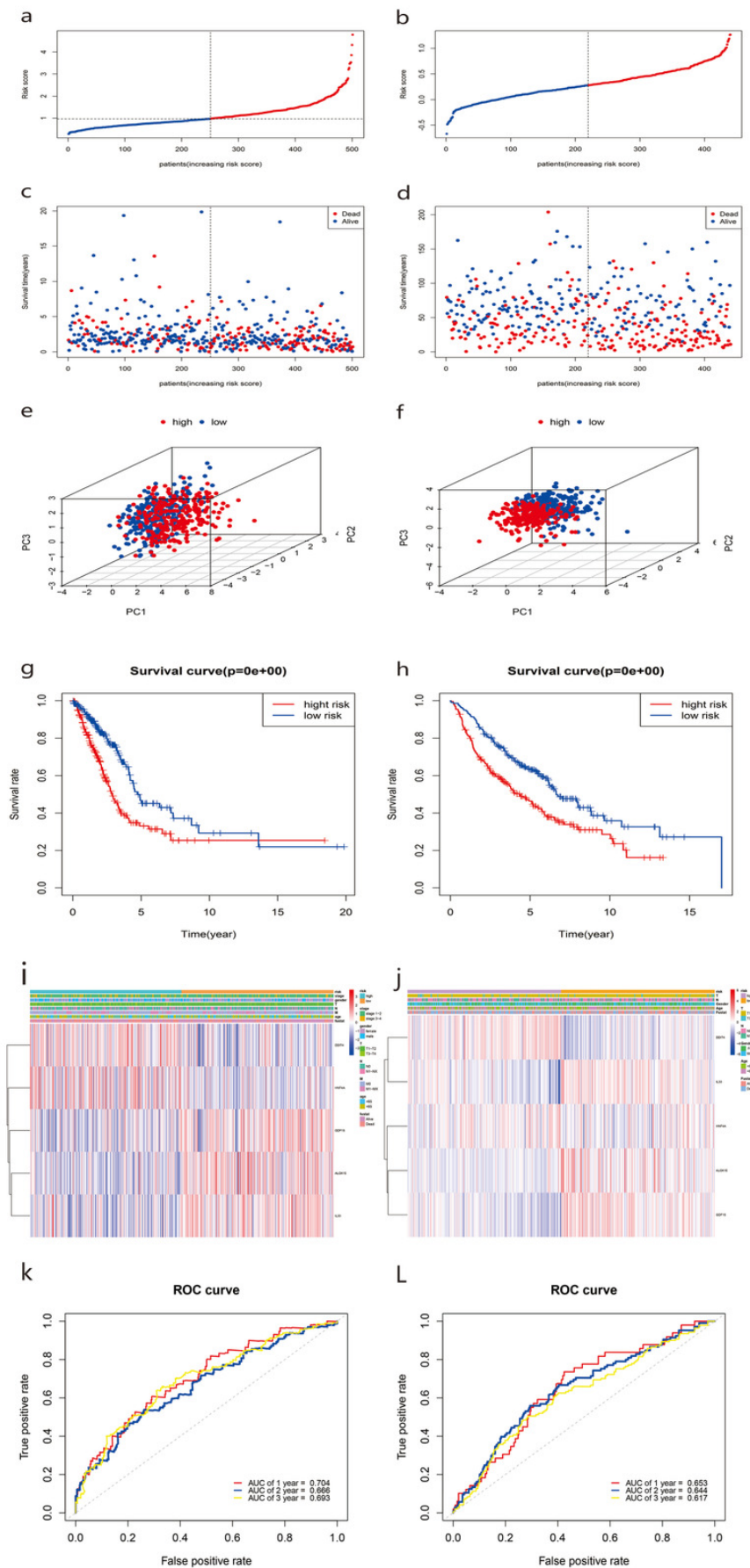
Figure 6

Figure 6. Construction of survival prediction nomogram and calibration plot of the nomogram.
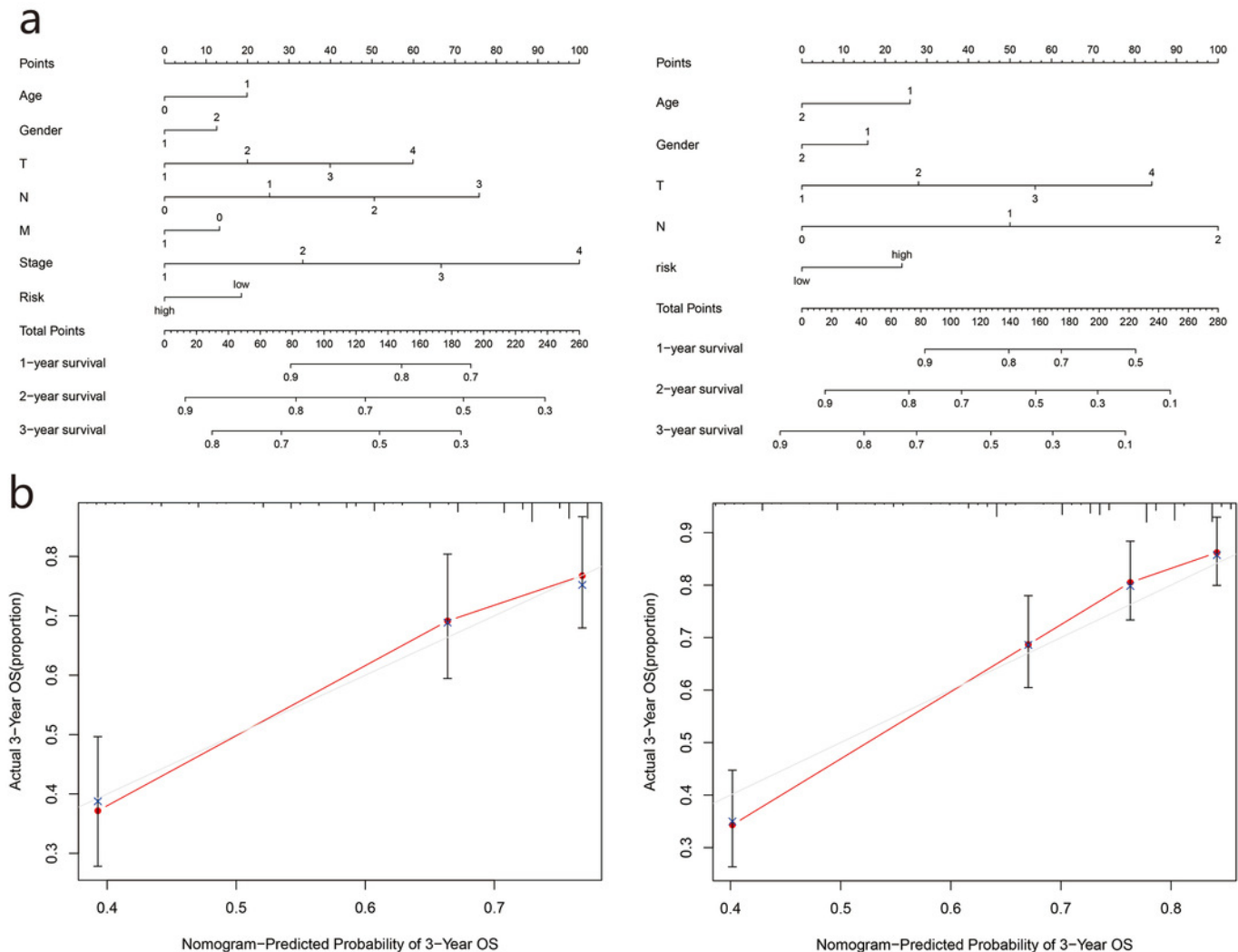
Figure 7

Figure 7. Forest plot of univariate and multivariate Cox regression analysis in TCGA set (a) and GSE68465 validation set(b).
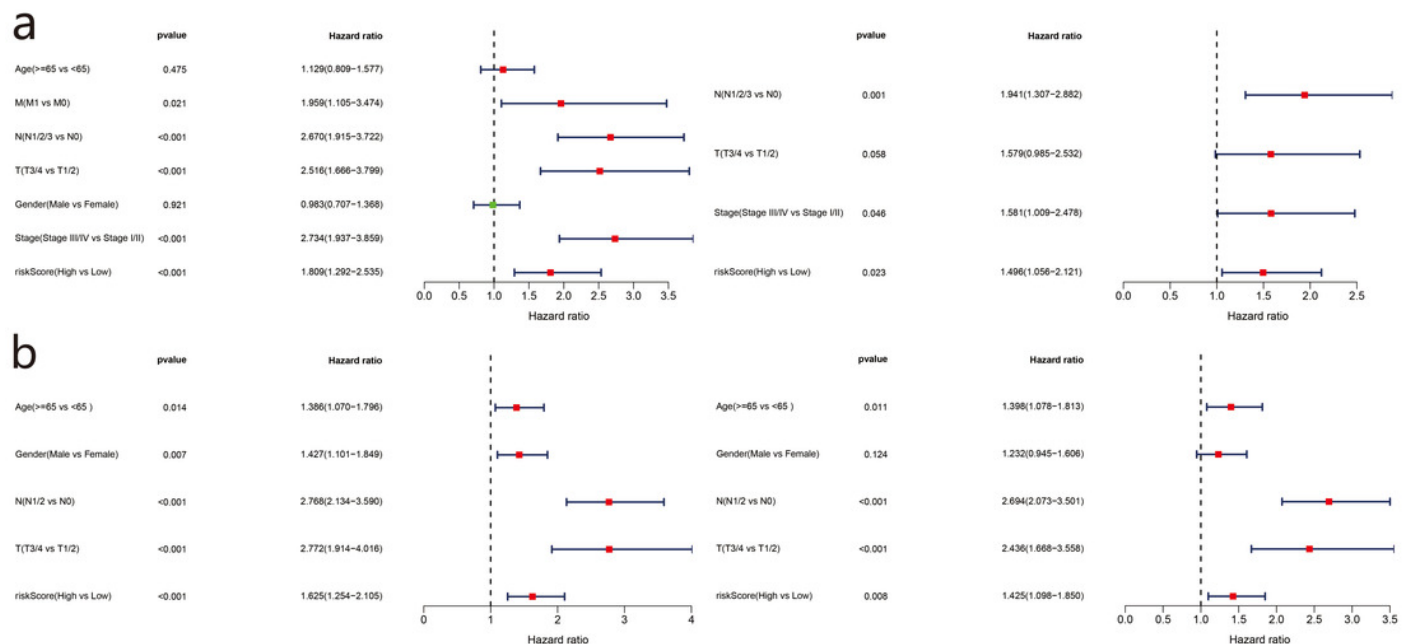
Figure 8

Figure 8. Functional enrichment analysis of two risk groups by Gene Ontology (GO) and Kyoto Encyclopedia of genes and genomes (KEGG)
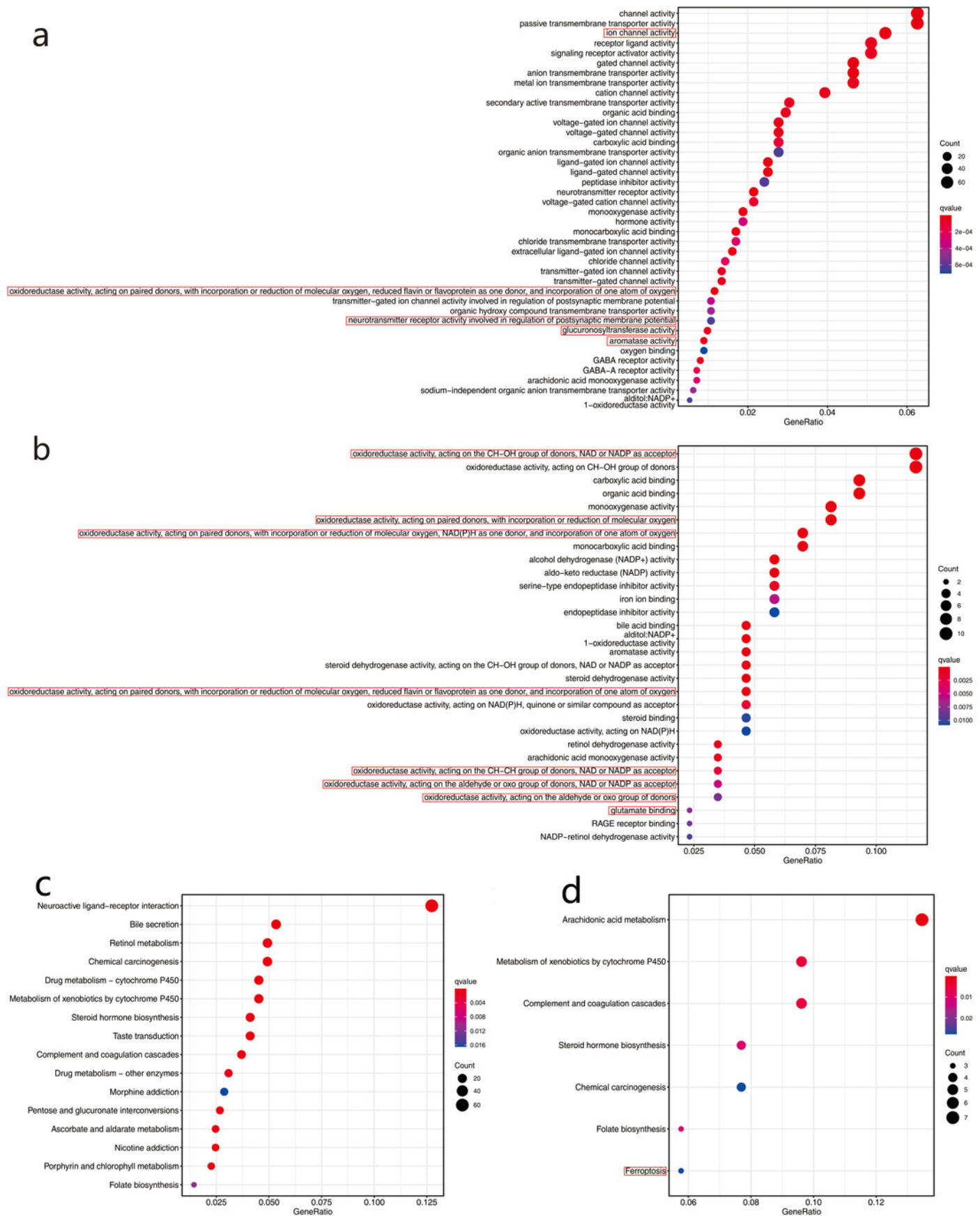
Figure 9

Figure 9. Comparison of the ssGSEA scores between two risk groups
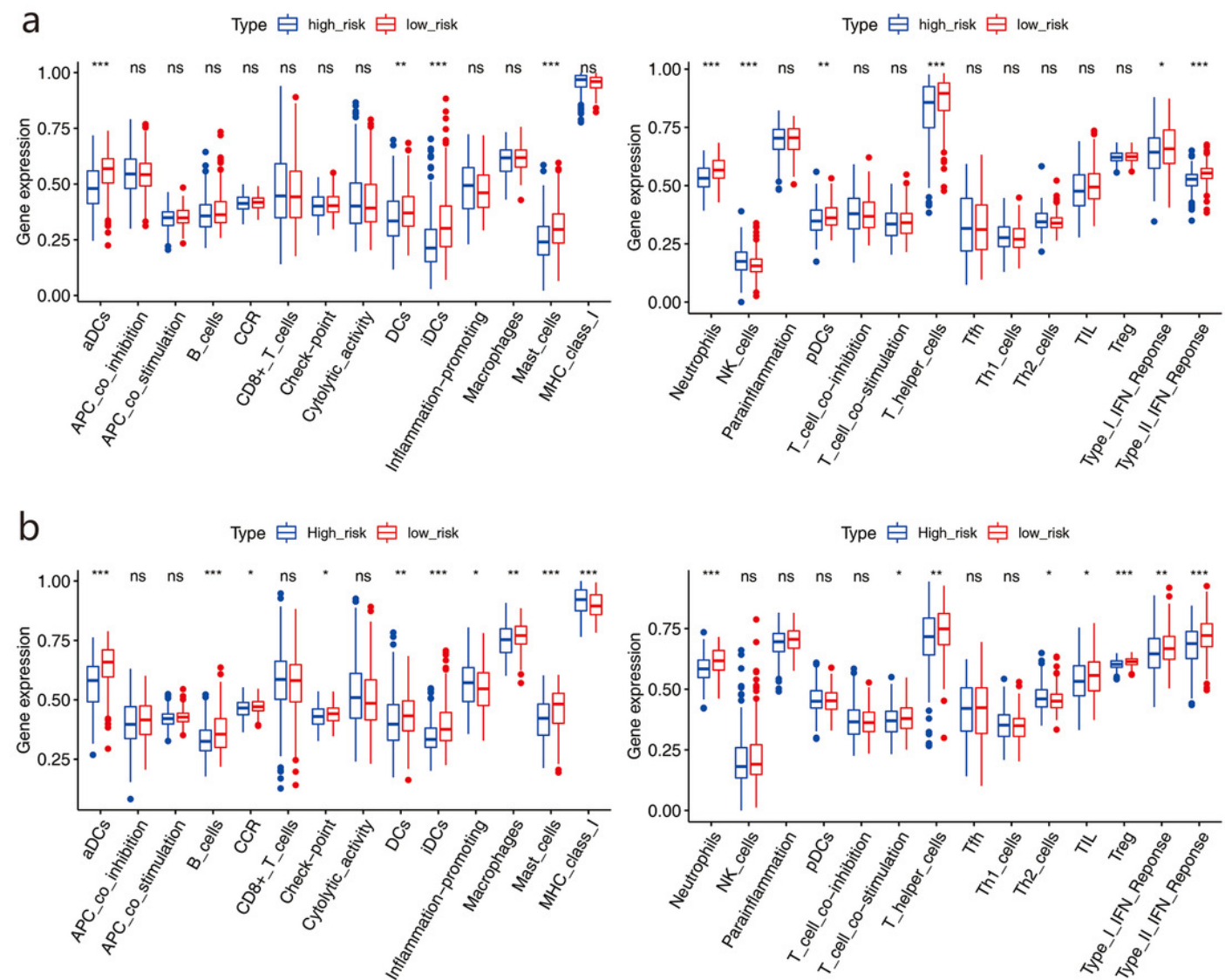
Table.1 Clinicopathological characteristics of the LUAD patients

1   Table.1 Clinicopathological characteristics of the LUAD patients

Characteristics	TCGA (n = 501)		GSE68465(n = 441)	
	Number of cases	%	Number of cases	%
Age (years)				
≥65	270	53.9	229	51.9
< 65	231	46.1	212	48.1
Gender				
Male	232	46.3	222	50.3
Female	269	53.7	219	49.7
T				
T1-2	436	87	401	90.9
T3-4	65	13	40	9.1
N				
N0	329	65.7	299	59.7
N1-Nx	172	34.3	142	40.3
M				
M0	336	67.1	NA	NA
M1-Mx	165	32.9	NA	NA
Stage				
Stage 1-2	394	78.6	NA	NA
Stage 3-4	107	21.4	NA	NA

Table2. The prognostic value of five ferroptosis-related genes

2 Table2.   The

Gene	Coef	HR	HR.95L	HR.95H	P vaule
ALOX15	-0.061165052	0.940667967	0.88	1.00	0.049434651
DDIT4	0.169594473	1.184824275	1.06	1.32	0.002251552
HNF4A	0.072356312	1.075038325	1.02	1.13	0.003200809
IL33	-0.131023567	0.877197101	0.79	0.97	0.010695027
GDF15	-0.107882858	0.897732752	0.83	0.98	0.011022652

3 prognostic value of five ferroptosis-related genes

4 Notes: Coef is the risk coefficient of each gene, if the value of coef > 0, it is regarded as the risk factor of

5 prognosis, otherwise it is regarded as the protective factor of prognosis. HR, hazard ratio.