Analysis of prognostic genes in the tumor microenvironment of lung adenocarcinoma

Zhan-yu Xu [Equal first author, 1], Meng-li Zhao [Equal first author, 1], Wen-jie Chen [1], Kun Li [1], Fang-lu Qin [1], Wei-wei Xiang [1], Yu Sun [1], Jiang-bo Wei [1], Li-qiang Yuan [1], Shi-kang Li [Corresp., 1], Sheng-hua Lin [Corresp. 1]

[1] Department of Thoracic and Cardiovascular Surgery, The First Affiliated Hospital of Guangxi Medical University, Nanning, China

Corresponding Authors: Shi-kang Li, Sheng-hua Lin
Email address: 201013064@qq.com, weipeng@stu.gxmu.edu.cn

Background: Prognostic genes in the tumor microenvironment play an important role in immune biological processes and the response of cancer to immunotherapy . Thus, we aimed to assess new biomarkers that are associated with immune/stromal cells in lung adenocarcinomas (LUAD) using the ESTIMATE algorithm, which also significantly affects the prognosis of cancer. **Methods:** The RNA sequencing (RNA-Seq) and clinical data of LUAD were downloaded from the the Cancer Genome Atlas (TCGA). The immune and stromal scores were calculated for each sample using the ESTIMATE algorithm. The LUAD gene chip expression profile data and the clinical data (GSE37745, GSE11969, and GSE50081) were downloaded from Gene Expression Omnibus (GEO) for subsequent validation analysis. Differentially expressed genes were calculated between high and low score groups. Univariate Cox regression analysis was performed on differentially expressed genes (DEGs) between the two groups to obtain initial prognosis genes. These were verified by three independent LUAD cohorts from the GEO database. Multivariate Cox regression was used to identify overall survival-related DEGs. UALCAN and the Human Protein Atlas were used to analyze the mRNA /protein expression levels of the target genes. Immune cell infiltration was evaluated using the Tumor Immune Estimation Resource (TIMER) and CIBERSORT methods, and stromal cell infiltration was assessed using xCell. **Results:** In this study, immune scores and stromal scores are significantly associated with the clinical characteristics of LUAD, including T stage, M stage, pathological stage, and overall survival time. 530 DEGs (18 upregulated and 512 downregulated) were found to coexist in the difference analysis with the immune scores and stromal scores subgroup. Univariate Cox regression analysis showed that 286 of the 530 DEGs were survival-related genes (p < 0.05). Of the 286 genes initially identified, nine prognosis-related genes (CSF2RB, ITK, FLT3, CD79A, CCR4, CCR6, DOK2, AMPD1, and IGJ) were validated from three separate LUAD cohorts. In addition, functional analysis of DEGs also showed that various immunoregulatory molecular pathways, including regulation of

immune response and the chemokine signaling pathways, were involved. Five genes (CCR6, ITK, CCR4, DOK2, and AMPD1) were identified as independent prognostic indicators of LUAD in specific data sets. The relationship between the expression levels of these genes and immune genes was assessed. We found that CCR6 mRNA and protein expression levels of LUAD were greater than in normal tissues. We evaluated the infiltration of immune cells and stromal cells in groups with high and low levels of expression of CCR6 in the TCGA LUAD cohort. In summary, we found a series of prognosis-related genes that were associated with the LUAD tumor microenvironment.

1 **Analysis of prognostic genes in the**

2 **tumor microenvironment of lung adenocarcinoma**

3 Zhan-yu Xu[1#], Meng-li Zhao[1#], Wen-jie Chen[1], Kun Li[1], Fang-lu Qin[1], Wei-wei Xiang[1], Yu

4 Sun[1], Jiang-bo Wei[1], Li-qiang Yuan[1], Shi-kang Li[1*], Sheng-hua Lin[1*]

5 1. Department of Thoracic and Cardiovascular Surgery, The First Affiliated Hospital of Guangxi

6 Medical University, Nanning, Guangxi Zhuang Autonomous Region 530021, P. R. China.

7 *Corresponding author:

8 Shi-kang Li,

9 Department of Thoracic and Cardiovascular Surgery, The First Affiliated Hospital of Guangxi

10 Medical University, Nanning, Guangxi Zhuang Autonomous Region 530021, P. R. China.

11 E-mail address: lsksci10@163.com

12 Sheng-hua Lin,

13 Department of Thoracic and Cardiovascular Surgery, The First Affiliated Hospital of Guangxi

14 Medical University, Nanning, Guangxi Zhuang Autonomous Region 530021, P. R. China.

15 E-mail address: weipeng@stu.gxmu.edu.cn

17 **ABSTRACT**

18 **Background:** Prognostic genes in the tumor microenvironment play an important role in

19 immune biological processes and the response of cancer to immunotherapy. Thus, we aimed to

20 assess new biomarkers that are associated with immune/stromal cells in lung adenocarcinomas

21 (LUAD) using the ESTIMATE algorithm, which also significantly affects the prognosis of

22 cancer.

23 **Methods:** The RNA sequencing (RNA-Seq) and clinical data of LUAD were downloaded from

24 the the Cancer Genome Atlas (TCGA). The immune and stromal scores were calculated for each

25 sample using the ESTIMATE algorithm. The LUAD gene chip expression profile data and the

26 clinical data (GSE37745, GSE11969, and GSE50081) were downloaded from Gene Expression

27 Omnibus (GEO) for subsequent validation analysis. Differentially expressed genes were

28   calculated between high and low score groups. Univariate Cox regression analysis was

29   performed on differentially expressed genes (DEGs) between the two groups to obtain initial

30   prognosis genes. These were verified by three independent LUAD cohorts from the GEO

31   database. Multivariate Cox regression was used to identify overall survival-related DEGs.

32   UALCAN and the Human Protein Atlas were used to analyze the mRNA/protein expression

33   levels of the target genes. Immune cell infiltration was evaluated using the Tumor Immune

34   Estimation Resource (TIMER) and CIBERSORT methods, and stromal cell infiltration was

35   assessed using xCell.

36   **Results:** In this study, immune scores and stromal scores are significantly associated with the

37   clinical characteristics of LUAD, including T stage, M stage, pathological stage, and overall

38   survival time. 530 DEGs (18 upregulated and 512 downregulated) were found to coexist in the

39   difference analysis with the immune scores and stromal scores subgroup. Univariate Cox

40   regression analysis showed that 286 of the 530 DEGs were survival-related genes ($p < 0.05$). Of

41   the 286 genes initially identified, nine prognosis-related genes (CSF2RB, ITK, FLT3, CD79A,

42   CCR4, CCR6, DOK2, AMPD1, and IGJ) were validated from three separate LUAD cohorts. In

43   addition, functional analysis of DEGs also showed that various immunoregulatory molecular

44   pathways, including regulation of immune response and the chemokine signaling pathways, were

45   involved. Five genes (CCR6, ITK, CCR4, DOK2, and AMPD1) were identified as independent

46   prognostic indicators of LUAD in specific data sets. The relationship between the expression

47   levels of these genes and immune genes was assessed. We found that CCR6 mRNA and protein

48   expression levels of LUAD were greater than in normal tissues. We evaluated the infiltration of

49   immune cells and stromal cells in groups with high and low levels of expression of CCR6 in the

50   TCGA LUAD cohort. In summary, we found a series of prognosis-related genes that were

51   associated with the LUAD tumor microenvironment.

52   **Keywords:** Lung adenocarcinoma; Immune and stromal cells; Prognosis; tumor-infiltrating

53

54   **INTRODUCTION**

55   Lung cancer is a malignant disease with the highest morbidity and mortality worldwide in recent

56   years (Bray et al. 2018). Eighty-five percent of lung cancers are non-small cell lung cancer

57   (NSCLC), which can be divided into lung adenocarcinomas (LUAD) and squamous carcinomas

58   (Liu et al. 2017a). There has been an increase in the incidence of LUADs, accounting for 60% of

59   NSCLCs and the most common type of lung cancer (Denisenko et al.). The 5-year survival rate

60   of patients with early-stage NSCLC after surgical treatment is 70%–90%; however,

61   approximately 75% of patients lost their opportunity for surgery by the time of initial diagnosis

62   (Goldstraw et al. 2016). Despite important advances in targeted molecular therapies and

63   immunotherapy for lung cancer, drug resistance still limits success (Ascierto et al. 2019), and

64   new biomarkers are warranted for different populations. This will have important implications

65   for improving patients' early diagnosis rates and for discovering novel targeted therapies.

66   In addition to tumor cells, stromal cells and immune cells as well as tumor-associated normal

67   epithelial cells form malignant solid tumor tissues (Moffitt et al. 2015). Stromal cells contain

68   cancer-related fibroblasts, pericytes, mesenchymal stem cells, blood, and lymphatic endothelial

69   cells (Quail & Joyce 2013). In the tumor microenvironment (TME), stromal cell components and

70   tumor cells interact with other to establish a dynamic connection, which affects the biological

71   function and chemical resistance of tumor cells (Joyce & Pollard 2009). Four types of immune

72   cells are relevant: tumor infiltrating lymphocytes (TILs), tumor-associated macrophages,

73   myeloid-derived suppressor cells, and dendritic cells (Bedognetti et al. 2019). TILs play an

74   important role in immune surveillance and killing of cancerous cells in patients with lung cancer

75   (Ye et al. 2017). NK cell-mediated tumor suppressor function is closely related to tumor

76   progression and is a key factor regulating lung cancer growth and metastasis (Fang et al. 2017).

77   Depicting the immune infiltration of tumor microenvironment or determining the proportion of

78   relevant normal cells in tumor tissues can help in building an accurate tumor prognosis and

79   prediction model. Guo JY et al. demonstrated that chondroitin sulfate proteoglycan serglycin

80   (SRGN), which is secreted by tumor cells and stromal components in the TME, promotes

81   malignant phenotypes through interacting with tumor cell receptor CD44. Increased expression

82 of SRGN is associated with poor prognosis in primary LUAD (Guo et al. 2017). Galectin-3 is a

83 lectin that contributes to TME immunosuppression and regulates diverse functions. Studies

84 suggest that increased Gal-3 expression during cancer progression augments tumor growth,

85 invasiveness, metastatic potential, and immune suppression (Compagno et al. 2014; Farhad et al.

86 2018)

87 The quantification of various cell types allows accurate analysis of dynamically changing

88 immune microenvironments. Experimental methods, such as single-cell sequencing

89 (Kyrochristos et al. 2019), are expensive and cumbersome to operate; therefore, various

90 algorithms have been developed to describe tumor immune environments, including single-

91 sample gene set enrichment analysis (GSEA) (Foroutan et al. 2018), MCPcounter (Petitprez et al.

92 2018), and CIBERSORT (Chen et al. 2018). Yoshihara et al. (Yoshihara et al. 2013) proposed a

93 new algorithm known as ESTIMATE that uses expression data to estimate the number of

94 immune and stromal cells in malignant tumor tissues. This algorithm focuses on the formation of

95 immune and stromal cells from major non-tumor components of tumor samples and identifies the

96 stromal and tumor tissue-specific characteristics associated with immune cell infiltration. Its

97 effectiveness has been shown in glioblastoma (Jia et al. 2018) and breast cancer (Winslow et al.

98 2016). Although accurate immune and stromal scoring of LUAD can be performed, there have

99 been no detailed studies that have analyzed related the expression profiles with ESTIMATE. In

100 this study, the microenvironment-related genes of the LUAD cohort of the TCGA database were

101 studied based on the ESTIMATE score, and the prognostic genes of three different LUAD

102 cohorts from the GEO database were verified. To further elucidate the immunological

103 mechanisms, we assessed the role of the immune microenvironment in the development of

104 LUAD using immune/stromal cell infiltration analysis.

105

106 **Materials and methods**

107 **Database**

108 The level 3 gene expression profile data of lung adenocarcinoma patients (using Illumina

109   HiSeq_RNASeqV2 lung adenocarcinoma RNA expression profiles) were downloaded from the

110   TCGA website (https://portal.gdc.cancer.gov/). The immune and stromal scores were calculated

111   from each sample using the ESTIMATE algorithm. In addition, patient clinical data were

112   downloaded from the TCGA official website (https://portal.gdc.cancer.gov/repository). The

113   LUAD gene chip expression profile data and the clinical data (GSE37745, GSE11969, and

114   GSE50081) of the samples were downloaded from the GEO official website

115   (https://www.ncbi.nlm.nih.gov/geo/) for subsequent validation analysis. To analyze the

116   mRNA/protein expression of CCR6 in primary LUAD and normal tissues, UALCAN

117   (http://ualcan.path.uab.edu) was used, which was based on the TCGA database and Clinical

118   Proteomic Tumor Analysis Consortium. To validate the expression of CCR6, the Human Protein

119   Atlas (http://www.proteinatlas.org/) was used. Tumor Immune Estimation Resource (TIMER)

120   (https://cistrome.shinyapps.io/timer/) was used to assess the correlation between differentially

121   expressed gene (DEG) expression and the immune cell infiltration level. Immune cell infiltration

122   was evaluated using the TIMER and CIBERSORT method, and stromal cell infiltration was

123   determined using xCell (https://xcell.ucsf.edu/).

124

125   **Identification of differentially expressed genes (DEGs)**

126   The limma package in R was used for the analysis of differential expression genes (Ritchie et al.

127   2015), and the cutoff value was set to |log fold change(FC)| > 1, and a p value < 0.05. We plotted

128   a heat map using the cluster analysis results of DEGs in R.

129

130   **GO and KEGG enrichment analysis**

131   To investigate the molecular functions, biological processes, cell components, and signaling

132   pathways involved in DEGs, we used the DAVID database (v6.8, https://david.ncifcrf.gov/) to

133   perform GO analysis and KEGG pathway enrichment. p value < 0.05 and FDR < 0.05 were

134   marked as significant terms.

135

136   **Statistical analysis**

137   The data were assessed for normal distribution, and the central tendency was expressed as the

138   mean ± standard deviation for the data with normal distribution. Means were compared using a

139   Student's t-test. Continuous variables that did not have normal distribution were analyzed using

140   the Wilcoxon rank sum test or Kruskal–Wallis rank sum test. Student's t-tests were used to

141   compare the distribution of scores in the high and low groups, gender, and M stages. The

142   Wilcoxon rank sum test was used to compare the distribution of scores in age, T stages, and N

143   stages. The Kruskal–Wallis rank sum test was used to compare the differences in the distribution

144   of scores of the four different pathological stages. Spearman's rank correlation coefficient was

145   used to measure the correlation between lung cancer stage and overall survival. Determination of

146   survival-related genes was performed using univariate Cox regression analysis (with a p value of

147   <0.05) and the corresponding survival curves were plotted. Multivariate Cox analysis was used

148   to evaluate the contribution of genes as independent prognosis factors for patient survival. All

149   tests were two-tailed, and p values less than 0.05 were considered statistically significant. Al

150   analyses were performed using SPSS for Windows, software version 25.0 (SPSS Inc., Chicago,

151   IL).

152   **RESULTS**

153   **Correlation between immune and stromal scores and clinical characteristics of LUAD**

154   We downloaded the gene expression profiles and clinical data of 492 patients with LUAD from

155   the TCGA database, including 225 males (45.7%), 267 females (54.3%), 157 patients under

156   60 years of age (31.9%), and 335 patients over 60 years of age (68.1%). The ESTIMATE

157   algorithm revealed stromal scores of −1959.31 to 2989.77, and the immune scores of −1355.85

158   to 2905.3. The results of the relationship between the two scores and clinical data (TNM and

159   pathological stage) are shown in Table 1. There were differences in the distribution of stromal

160   scores between the sexes and across M stages. Immune scores also differed with respect to T and

161   pathological stages.

162   We found that the immune and stromal scores of female patients were significantly higher than

163   those of from male patients (p = 0.0069 and 0.0049, respectively) (Figure 1A and 1B). The

164   immune score of T1 was significantly higher than that of T2/T3/T4 stages (p = 0.0002, Figure

165   1C); however, but the distribution of stromal scores was not different between the two groups

166   (p = 0.0999) (Figure 1D). The immune and stromal scores of the M0 stage was higher than those

167   of the M1 stage (p = 0.0098 and 0.0366, respectively) (Figure 1E and 1F). For the different

168   pathological stages of LUAD, the immune scores were different at each stage (p = 0.0276, Figure

169   1G) and gradually decreased from stage I to stage IV (Spearman correlation coefficients:

170   ρ = −0.131, p = 0.004); however, but the difference in stromal scores among the four

171   pathological stages was not statistically significant (p = 0.1525, Figure 1H).

172   Next, we evaluated the correlations between immune (or stromal) scores and overall survival.

173   LUADs (N = 492) were divided into high- and low-score groups, with 246 cases in each group.

174   Stromal scores ranged from −959.31 to 36.85 for the low group and from 39.8 to 2098.77 for the

175   high group. Immune scores ranged from −1355.85 to 952.61 for the low group and 964.93 to

176   2905.3 for the high group. The Kaplan–Meier survival curve analysis for immune scores

177   indicated that the median survival of the low-score group was lower than that of the high-score

178   group (1194 d vs. 1499 d, p = 0.011; log-rank test; Figure 1I). Similarly, for stromal scores, the

179   median survival of the low-score group was lower than that of the high-score group (1235 d vs.

180   1492 d, p = 0.179; log-rank test; Figure 1J), although it was not statistically significant.

181

182   **Comparison of gene expression profiles with LUAD immune scores and stromal scores**

183   We performed differential expression analysis of the high- and low- immune scoring patients and

184   found that 884 genes were downregulated and 55 genes were upregulated (Figure 2A, Table S1).

185   Similarly, differential analysis of the high- and low-stromal scoring groups revealed that 999

186   genes were downregulated and 28 genes were up-regulated (|logFC| > 1, p < 0.05) (Figure 2B,

187   Table S1). In the Venn diagram, 18 genes were found to be upregulated in both groups (Figure

188   2C) and 512 genes were downregulated in both groups (Figure 2D).

189

190 **Correlation of individual DEGs expression with overall survival and Functional**

191 **enrichment analysis of prognostic associated genes**

192 To explore the DEGs associated with overall survival, we performed a univariate Cox regression

193 analysis. The results (Table S2) showed that 286 of the 530 DEGs were survival-related genes

194 ($p < 0.05$). To further determine the relationship between these prognostic DEGs, we used the

195 STRING online tool to obtain a PPI network. This network has 277 nodes and 2143 edges

196 (Figure 3). Functional enrichment clustering of 286 prognostic DEGs was also closely related to

197 the immune pathways. GO analysis results showed that 46 terms were statistically significant,

198 and KEGG enrichment results showed that 22 pathways were significantly enriched ($p < 0.05$,

199 FDR < 0.05). The top GO terms included adaptive immune response, T cell co-stimulation,

200 MHC class II receptor activity, and transmembrane signaling receptor activity (Figure 4A). In

201 addition, most pathways generated from the KEGG pathway enrichment analysis were associated

202 with immune responses, such as cell adhesion molecules, primary immunodeficiency, antigen

203 processing and presentation, and chemokine signaling pathways (Figure 4B).

204

205 **Further identification of prognostic genes in the GEO database**

206 To confirm whether the genes identified in the above steps have prognostic functions in other

207 LUAD cohorts, we used the expression profiles and clinical data of three datasets in the GEO

208 database for verification: GSE37745, GSE11969, and GSE50081. A total of 80 genes in

209 GSE37745, 2703 genes in GSE11969, and 34 genes in GSE50081 were associated with survival.

210 CSF2RB, inducible T cell kinase (ITK), and FLT3, and CD79A were found to be common

211 among the GSE37745, GSE11969, and TCGA datasets (Figure 5A). In addition, CCR4, CCR6,

212 DOK2, AMPD1, and IGJ were found to be common among GSE37745, GSE50081, and TCGA

213 (Figure 5B). Kaplan–Meier survival curves were generated for the nine prognostic genes (Figure

214 5C). The results showed that for all but except CCR4, the overall survival time of the low

215 expression group was significantly lower than that of the high expression group (log-rank

216 $p < 0.05$). This suggests that reduced expression of the eight genes (CSF2RB, ITK, FLT3,

217   CD79A, CCR6, DOK2, AMPD1, and IGJ) and increased expression of one gene (CCR4) can

218   predict poor survival of patients with for LUAD.

219

220   **Immune cell infiltration analysis revealed the correlation between the expression of the**

221   **DEGs and immunocytes**

222   To confirm these findings, we used multivariate Cox regression and found that the five genes,

223   including CCR6, ITK, CCR4, DOK2, and AMPD1, were an independent prognostic indicator for

224   LUAD in their specific data sets (Figures 6A–6D). We investigated immune infiltration using the

225   TIMER deconvolution approach. Interestingly, the expression of the five genes identified

226   positively correlated to the infiltration level of different immune cells, including B cells, CD4$^+$ T

227   cells, CD8$^+$ T cells, macrophage cells, neutrophil cells, and dendritic cells ($p < 0.05$). The

228   expression levels of the five genes were negatively correlated with the tumor purity ($p < 0.05$;

229   Figures 7A–7II).

230

231   **Evaluation of the immune and stromal status between the groups with low and high**

232   **expression of CCR6**

233   Multivariate Cox regression analysis identified that low expression of CCR6 is an independent

234   prognostic factor of poor survival in the TCGA cohort of 484 patients with LUAD (HR 0.29,

235   95％CI 0.12–0.72, $p = 0.007$). According to the results of UALCAN, mRNA and protein

236   expression of CCR6 were all significantly upregulated in primary LUAD tissues as compared

237   with normal samples (all $p < 0.05$; Figures 8A and 8B). Immunohistochemistry staining obtained

238   from The Human Protein Atlas database also demonstrated that CCR6 was upregulated (Figures

239   8C and 8D). According to the median value of CCR6 expression levels, we divided the 484

240   LUAD patients into a low expression group and a high expression group. CIBERSORT and

241   xCell were used to depict the immune cell and stromal cell landscape of CCR6 with high and

242   low expression levels. Figures 8E and 8F show the proportions of immune cells and stromal cells

243   in the 484 LUAD tissues. The differential proportion of infiltration of immune and stromal cells

244 in the group with low CCR6 expression and the group with high CCR6 expression is shown in

245 Figures 8G and 8H. CIBERSORT analysis suggests that high CCR6 expression is associated

246 with recruitment of memory B cells, resting memory CD4 T cells, monocytes, M1 macrophages,

247 resting dendritic cells, and resting mast cells (Figure S1). xCell analysis suggests that high CCR6

248 expression is associated with recruitment of adipocytes, chondrocytes, endothelial cells, and

249 fibroblasts (Figure S1). A comparison of the tumor infiltration levels among tumors with

250 different somatic copy number alterations (SCNAs) for a given gene was determined using

251 TIMER. The results showed that enrichment of the six immune cell types (B cell, CD4$^+$ T cell,

252 CD8$^+$ T cell, macrophage cell, neutrophil cell, and dendritic cell) were significantly

253 downregulated in LUAD with SCNA of CCR6 (Figure 8I).

254

255 **DISCUSSION**

256 Previous studies have shown that the TME can play a key role in promoting tumor progression

257 and increasing mortality (Caetano et al. 2016; Wood et al. 2014). From the perspective of the

258 TME, we identified a series of genes related to the overall survival of patients with LUAD based

259 on data in the TCGA database. By analyzing differences in the overall expression profiles of

260 patients of LUAD cases in high and low scoring groups using the ESTIMATE algorithm, we

261 identified 286 prognostic genes involved in immune response and chemokine signaling.

262 Importantly, we validated these results using three independent LUAD datasets from the GEO

263 database. From the overlap among these datasets, we obtained nine prognostic genes that were

264 associated with the immune microenvironment of LUAD. We found that five (ITK, DOK2,

265 AMPD1, CCR4, and CCR6) out of nine genes were prognostic factors for LUAD in their

266 specific data set.

267 Abnormal expression of protein tyrosine kinases (PTK) is related to tumor invasion and

268 metastasis, tumor neovascularization, and tumor chemotherapy resistance (Knosel et al. 2014).

269 Many drugs have been developed to target PTKs. Although tyrosine kinase inhibitors (TKIs)

270 continue to be involved in targeted therapies for lung cancer, acquired resistance still remains a

271   problem (Ahn et al. 2017). It has been suggested that the mechanism of action of different PTKs

272   needs additional study. IL2 inducible T cell kinase (ITK) is expressed in many immune cells. As

273   a non-receptor tyrosine kinase, it plays an important role in regulating TCRs, chemokine receptor

274   CXCR4, and FCεR-mediated signaling pathways (Sahu & August 2009). The proteins

275   downstream of tyrosine kinase participate in the proliferation and migration of lung cancer

276   (Ghanem et al. 2014), leukemia (Tomoharu et al. 2004), gastric cancer (Liu & Xiao 2014), and

277   other cancer cells via the receptor tyrosine kinase signaling pathway. A study has found that

278   DOK2 is a tumor suppressor gene in LUAD and that knocking out Dok2 in mice accelerates lung

279   tumorigenesis induced by oncogenic EGFR (Berger et al. 2013). We found that low expression

280   levels of ITK, and DOK2 in LUAD were associated with poor prognosis, possibly affecting

281   tumor progression via the transmembrane receptor PTK signaling pathway. However, further

282   studies are required to determine whether TKI can become a new target for lung cancer.

283   The AMPD1 gene encodes adenosine monophosphate deaminase (AMPD), a crucial enzyme in

284   purine nucleotide and energy metabolism, especially in skeletal muscle and cardiac muscle. The

285   role of AMPD1 is to catalyze the conversion of adenosine monophosphate to inosine

286   monophosphate (Smolenski et al. 2014). Xiao Luo et al. identified that AMPD1 was closely

287   associated with the survival of breast cancer patients (Luo et al. 2018). Tianzhou Zha et al. found

288   that the expression of AMPD1 in serum of patients with papillary thyroid carcinoma is closely

289   related to the malignant evolution of PTC and clinical prognosis of patients (Zha et al. 2018).

290   The relationship between AMPD1 and LUAD has not previously been elucidated.

291   Functional enrichment analysis showed that CCR6 was significantly involved in the chemokine

292   signaling pathway. This is consistent with results from previous studies that indicate that in the

293   TME, the communication among cells is dynamically regulated by a complex network of

294   immune factors, including cytokines, chemokines, and numerous growth factors (Shimizu et al.

295   2013). A variety of chemokines promote tumor proliferation, which can play a vital role in tumor

296   progression by helping tumor cells escape immune monitoring (Chang et al. 2018).

297   We found that low expression of CCR6 and high expression of CCR4 were significantly

298   associated with poor prognosis of LUAD, based on a multi-cohort study. Multivariate Cox

299   analysis suggests that CCR6 is an independent prognostic factor in the TCGA LUAD cohort and

300   that CCR4 is an independent prognostic indicator in the two LUAD cohorts, GSE50081 and

301   GSE37745. Wei Liu et al. (Liu et al. 2017b) also found that CCR4 is an independent risk factor

302   for the overall survival of NSCLC, as it functions as a ligand for CCL17 and CCL22. CCR4 can

303   help tumor cells escape the host immune attack by attracting Treg into the TME. Studies have

304   identified that CCR6 can act as a ligand for CCL20 and promote lung metastasis of cancerous

305   adrenal tissue (Raynaud et al. 2010). In a mouse model of lung cancer (Lewis lung carcinoma),

306   expression of CCR6 in tumor cells reduced metastatic potential (Sutherland et al. 2007). Similar

307   to our findings, a study found that increased expression of CCR6 in tumor cells is an independent

308   predictor of a better prognosis in patients with LUAD (Minamiya et al. 2011). A recent study

309   showed that CCL20, in coordination with CCR6, can recruit Treg cells to tumor sites (Zhang et

310   al. 2015). To our knowledge, we have for the first time depicted the landscape of immune cell

311   infiltration and the landscape of CCR6 in stromal cells in LUAD. Our results also suggest that

312   there is a close correlation between CCR6 genomic alterations and immune cell enrichments in

313   LUAD. Therefore, we hypothesized that genetic alteration of CCR6 may play an important role

314   in LUAD. Consistent with both of these previous studies, our analysis reveals an important role

315   of CCR6 in LUAD and its potential value as a biomarker.

316

317   **CONCLUSIONS**

318   In conclusion, using the ESTIMATE algorithm, TIMER, CIBERSORT, and xCell, we analyzed

319   the expression profile data of patients with LUAD from the TCGA database and identified a

320   series of genes related to the TME. Many previously unknown genes were found in addition to

321   the genes already associated with lung cancer or immune microenvironments. Finally, we

322   examined the interaction between CCR6 and TME. This may be important to identify the

323   connecting mechanism between TME and LUAD prognosis.

324

325 **ACKNOWLEDGMENTS**

326 This study was supported by the National Natural Science Foundation of China

327 (NSFC81660488) and the Guangxi Natural Science Foundation under Grant

328 2017GXNSFAA198123.

329

330 **CONFLICTS OF INTEREST**

331 The authors declare that there are no conflicts of interest.

332

333 **REFERENCES**

334 Ahn M-J, Sun J-M, Lee S-H, Ahn JS, and Park K. 2017. EGFR TKI combination with
335   immunotherapy in non-small cell lung cancer. *Expert opinion on drug safety* 16:465-469.
336   10.1080/14740338.2017.1300656

337 Ascierto PA, Bifulco C, Buonaguro L, Emens LA, Ferris RL, Fox BA, Delgoffe GM, Galon J,
338   Gridelli C, Merlano M, Nathan P, Odunsi K, Okada H, Paulos CM, Pignata S, Schalper
339   KA, Spranger S, Tortora G, Zarour H, Butterfield LH, and Puzanov I. 2019. Perspectives
340   in immunotherapy: meeting report from the "Immunotherapy Bridge 2018" (28-29
341   November, 2018, Naples, Italy). *J Immunother Cancer* 7:332. 10.1186/s40425-019-0798-
342   3

343 Bedognetti D, Ceccarelli M, Galluzzi L, Lu R, Palucka K, Samayoa J, Spranger S, Warren S, Wong
344   KK, Ziv E, Chowell D, Coussens LM, De Carvalho DD, DeNardo DG, Galon J, Kaufman
345   HL, Kirchhoff T, Lotze MT, Luke JJ, Minn AJ, Politi K, Shultz LD, Simon R, Thorsson
346   V, Weidhaas JB, Ascierto ML, Ascierto PA, Barnes JM, Barsan V, Bommareddy PK, Bot
347   A, Church SE, Ciliberto G, De Maria A, Draganov D, Ho WS, McGee HM, Monette A,
348   Murphy JF, Nistico P, Park W, Patel M, Quigley M, Radvanyi L, Raftopoulos H, Rudqvist
349   NP, Snyder A, Sweis RF, Valpione S, Zappasodi R, Butterfield LH, Disis ML, Fox BA,
350   Cesano A, and Marincola FM. 2019. Toward a comprehensive view of cancer immune
351   responsiveness: a synopsis from the SITC workshop. *J Immunother Cancer* 7:131.
352   10.1186/s40425-019-0602-4

353 Berger AH, Ming C, Alessandro M, Janas JA, Masaru N, Bronson RT, Taylor BS, Marc L, Linda
354   VA, and Katerina P. 2013. DOK2 Inhibits EGFR-Mutated Lung Adenocarcinoma. *Plos*
355   *One* 8:e79526.

356 Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, and Jemal A. 2018. Global cancer
357   statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36
358   cancers in 185 countries. *CA: A Cancer Journal for Clinicians*.

359 Caetano MS, Zhang H, Cumpian AM, Gong L, Unver N, Ostrin EJ, Daliri S, Chang SH, Ochoa
360   CE, Hanash S, Behrens C, Wistuba, II, Sternberg C, Kadara H, Ferreira CG, Watowich SS,
361   and Moghaddam SJ. 2016. IL6 Blockade Reprograms the Lung Tumor Microenvironment

362   to Limit the Development and Progression of K-ras-Mutant Lung Cancer. *Cancer Res*
363   76:3189-3199. 10.1158/0008-5472.Can-15-2840

364 Chang WA, Yen MC, Hung JY, Yang CJ, Jian SF, Yeh IJ, Liu KT, Hsu YL, and Kuo PL. 2018.
365   Investigation of the role of tumor necrosis factor-like weak inducer of apoptosis in non-
366   small cell lung cancer. *Oncology Reports* 39:573-581.

367 Chen B, Khodadoust MS, Liu CL, Newman AM, and Alizadeh AA. 2018. Profiling Tumor
368   Infiltrating Immune Cells with CIBERSORT. *Methods in molecular biology (Clifton, NJ)*
369   1711:243-259. 10.1007/978-1-4939-7493-1_12

370 Compagno D, Gentilini LD, Jaworski FM, Perez IG, Contrufo G, and Laderach DJJG. 2014.
371   Glycans and galectins in prostate cancer biology, angiogenesis and metastasis.   24:899-
372   906.

373 Denisenko TV, Budkevich IN, Zhivotovsky BJCD, and Disease. Cell death-based treatment of
374   lung adenocarcinoma.   9:117.

375 Fang F, Xiao W, and Tian Z. 2017. NK cell-based immunotherapy for cancer. *Semin Immunol*
376   31:37-54. 10.1016/j.smim.2017.07.009

377 Farhad M, Rolig AS, and Oncoimmunology WLRJ. 2018. The role of Galectin-3 in modulating
378   tumor growth and immunosuppression within the tumor microenvironment.   7:-.

379 Foroutan M, Bhuva DD, Lyu R, Horan K, Cursons J, and Davis MJ. 2018. Single sample scoring
380   of molecular phenotypes. *BMC Bioinformatics* 19:404. 10.1186/s12859-018-2435-4

381 Ghanem T, Bracken J, Kasem A, Wen GJ, and Mokbel K. 2014. mRNA expression of DOK1-6 in
382   human breast cancer. *World Journal of Clinical Oncology* 5:156-163.

383 Goldstraw P, Chansky K, Crowley J, Rami-Porta R, Asamura H, Eberhardt WE, Nicholson AG,
384   Groome P, Mitchell A, and Bolejack V. 2016. The IASLC Lung Cancer Staging Project:
385   Proposals for Revision of the TNM Stage Groupings in the Forthcoming (Eighth) Edition
386   of the TNM Classification for Lung Cancer. *J Thorac Oncol* 11:39-51.
387   10.1016/j.jtho.2015.09.009

388 Guo JY, Hsu HS, Tyan SW, Li FY, Shew JY, Lee WH, and Chen JY. 2017. Serglycin in tumor
389   microenvironment promotes non-small cell lung cancer aggressiveness in a CD44-
390   dependent manner. *Oncogene* 36:2457-2471. 10.1038/onc.2016.404

391 Jia D, Li S, Li D, Xue H, Yang D, and Liu Y. 2018. Mining TCGA database for genes of prognostic
392   value in glioblastoma microenvironment. *Aging (Albany NY)* 10:592-605.
393   10.18632/aging.101415

394 Joyce JA, and Pollard JW. 2009. Microenvironmental regulation of metastasis. *Nat Rev Cancer*
395   9:239-252. 10.1038/nrc2618

396 Knosel T, Kampmann E, Kirchner T, and Altendorf-Hofmann A. 2014. [Tyrosine kinases in soft
397   tissue tumors]. *Pathologe* 35 Suppl 2:198-201. 10.1007/s00292-014-1958-3

398 Kyrochristos ID, Ziogas DE, Goussia A, Glantzounis GK, and Roukos DH. 2019. Bulk and Single-
399   Cell Next-Generation Sequencing: Individualizing Treatment for Colorectal Cancer.
400   *Cancers (Basel)* 11. 10.3390/cancers11111809

401 Liu G, Pei F, Yang F, Li L, Amit A, Liu S, Buchan J, and William CJIJoMS. 2017a. Role of
402   Autophagy and Apoptosis in Non-Small-Cell Lung Cancer.   18:367-.

Liu HS, and Xiao HS. 2014. MicroRNAs as potential biomarkers for gastric cancer. *World Journal of Gastroenterology* 20:12007-12017.

Liu W, Wei X, Li L, Wu X, Yan J, Yang H, and Song F. 2017b. CCR4 mediated chemotaxis of regulatory T cells suppress the activation of T cells and NK cells via TGF-β pathway in human non-small cell lung cancer. *Biochemical & Biophysical Research Communications* 488:196.

Luo X, Yu H, Song Y, and Physiology TSJJoC. 2018. Integration of metabolomic and transcriptomic data reveals metabolic pathway alteration in breast cancer and impact of related signature on survival. 234:13021-13031.

Minamiya Y, Saito H, Takahashi N, Ito M, Toda H, Ono T, Konno H, Motoyama S, and Ogawa JIJTB. 2011. Expression of the chemokine receptor CCR6 correlates with a favorable prognosis in patients with adenocarcinoma of the lung. 32:197-202.

Moffitt RA, Marayati R, Flate EL, Volmar KE, Loeza SG, Hoadley KA, Rashid NU, Williams LA, Eaton SC, Chung AH, Smyla JK, Anderson JM, Kim HJ, Bentrem DJ, Talamonti MS, Iacobuzio-Donahue CA, Hollingsworth MA, and Yeh JJ. 2015. Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma. *Nat Genet* 47:1168-1178. 10.1038/ng.3398

Petitprez F, Vano YA, Becht E, Giraldo NA, de Reyniès A, Sautès-Fridman C, and Fridman WH. 2018. Transcriptomic analysis of the tumor microenvironment to guide prognosis and immunotherapies. *Cancer immunology, immunotherapy : CII* 67:981-988. 10.1007/s00262-017-2058-z

Quail DF, and Joyce JA. 2013. Microenvironmental regulation of tumor progression and metastasis. *Nat Med* 19:1423-1437. 10.1038/nm.3394

Raynaud CM, Olaf M, Philippe D, Frédéric C, Ken André O, Vincent DM, Fabrice A, Laure S, and Jean-Charles S. 2010. Expression of chemokine receptor CCR6 as a molecular determinant of adrenal metastatic relapse in patients with primary lung cancer. *Clinical Lung Cancer* 11:187-191.

Ritchie ME, Belinda P, Di W, Yifang H, Law CW, Wei S, and Smyth GK. 2015. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research* 43:e47.

Sahu N, and August A. 2009. ITK inhibitors in inflammation and immune-mediated disorders. *Curr Top Med Chem* 9:690-703. 10.2174/156802609789044443

Shimizu K, Okita R, and Nakata M. 2013. Clinical significance of the tumor microenvironment in non-small cell lung cancer. *Ann Transl Med* 1:20. 10.3978/j.issn.2305-5839.2013.06.01

Smolenski RT, Rybakowska I, Turyn J, Romaszko P, Zabielska M, Taegtmeyer A, Słomińska EM, Kaletha KK, and Barton PJ. 2014. AMP deaminase 1 gene polymorphism and heart disease-a genetic association that highlights new treatment. *Cardiovasc Drugs Ther* 28:183-189. 10.1007/s10557-013-6506-5

Sutherland A, ., J-F M, Maho A, ., and Parmentier M, . 2007. Expression of the chemokine receptor CCR6 in the Lewis lung carcinoma (LLC) cell line reduces its metastatic potential in vivo. *Cancer Gene Therapy* 14:847-857.

444 Tomoharu Y, Masaki S, Atsushi I, Asuka I, Yasuhiro E, Mitsujiro O, Kazuho H, Hisaaki S,
445   Katsuko S, and Kohichiro T. 2004. Role of Dok-1 and Dok-2 in myeloid homeostasis and
446   suppression of leukemia. *Journal of Experimental Medicine* 200:1681-1687.
447 Winslow S, Lindquist KE, Edsjo A, and Larsson C. 2016. The expression pattern of matrix-
448   producing tumor stroma is of prognostic importance in breast cancer. *BMC Cancer* 16:841.
449   10.1186/s12885-016-2864-2
450 Wood SL, Pernemalm M, Crosbie PA, and Whetton AD. 2014. The role of the tumor-
451   microenvironment in lung cancer-metastasis and its relationship to potential therapeutic

452   targets ☆. *Cancer Treatment Reviews* 40:558-566.

453 Ye L, Zhang F, Li H, Yang L, Lv T, Gu W, and Song Y. 2017. Circulating Tumor Cells Were
454   Associated with the Number of T Lymphocyte Subsets and NK Cells in Peripheral Blood
455   in Advanced Non-Small-Cell Lung Cancer. *Dis Markers* 2017:5727815.
456   10.1155/2017/5727815
457 Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torresgarcia W, Treviño V,
458   Shen H, Laird PW, and Levine DA. 2013. Inferring tumour purity and stromal and immune
459   cell admixture from expression data. *Nature Communications* 4:2612.
460 Zha T, Wu HJE, and Medicine T. 2018. Expression of serum AMPD1 in thyroid carcinoma and
461   its clinical significance.
462 Zhang CY, Qi Y, Li XN, Yang Y, Liu DL, Zhao J, Zhu DY, Wu K, Zhou XD, Zhao SJB, and
463   Pharmacotherapy. 2015. The role of CCL20/CCR6 axis in recruiting Treg cells to tumor
464   sites of NSCLC patients.   69:242-248.
465

466 **FIGURE LEGENDS**

467 **Figure 1. Relationships among immune and stromal scores, clinical characteristics, and**

468 **overall survival.**

469 (A-H) Distribution of stromal and immune scores based on gender, T stage, M stage, and

470 pathological stage. (I, J) Kaplan–Meier survival analysis of low and high score groups based on

471 immune and stromal scores. (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$)

472

473 **Figure 2. Comparison of gene expression profiles with immune scores and stromal scores in**

474 **LUAD.**

475 (A, B) Heatmaps were drawn for the DEGs obtained from the respective high and low group

476 comparisons of the immune scores and stromal scores. (C, D) Venn diagrams drawn from the

477 DEGs that are up-regulated or down-regulated in immune scores and stromal scores.

478

Figure 3. PPI network of 286 prognostic DEGs.

480

Figure 4. GO terms and KEGG analysis of 286 DEGs.

(A) GO analysis of 286 DEGs. (B) KEGG enrichments of 286 DEGs. GO – gene ontology;

KEGG – Kyoto Encyclopedia of Genes and Genomes; MF – molecular function; CC – cellular

component; BP – biological process.

485

Figure 5. Correlation between the expression of nine prognostic genes and overall survival

in GEO. (A) The Venn diagram shows prognostic genes identified based on GSE11969,

GSE37745, and TCGA cross-validation. (B) The Venn diagram shows prognostic genes

identified based on GSE50081, GSE37745, and TCGA cross-validation. (C-K) Kaplan–Meier

survival curves were generated for the nine prognostic genes (AMP11, CCR4, CCR6, CD79A,

CSF2RB, DOK2, FLT3, IGJ, and ITK) extracted from the comparison of groups of high (red

line) and low (blue line) gene expression.

Figure 6. The multivariate Cox regression analysis of 9 prognostic genes.

(A–D) The HR and p value from the multivariate Cox regression of the selected prognostic genes

in TCGA, GSE11969, GSE37745, and GSE50081 (* $p < 0.05$, ** $p < 0.01$).

496

Figure 7. Correlation analysis between the expression of the DEGs (CCR6, ITK, CCR4,

DOK2, AMPD1) and the infiltration levels of B cells, CD8$^+$ T cells, CD4$^+$ T cells,

macrophages, neutrophils, and dendritic cells.

(A–G) The TIMER database showed that CCR6 expression was negatively correlated with tumor

purity and was positively correlated with the infiltration of six types of immune cells in lung

adenocarcinoma.

(H-N) The TIMER database showed that ITK expression was negatively correlated with the

tumor purity and was positively correlated with the infiltration of six types of immune cells in

505   lung adenocarcinoma.

506   (O-U) The TIMER database showed that CCR4 expression was negatively correlated with the

507   tumor purity and was positively correlated with the infiltration of six types of immune cells in

508   lung adenocarcinoma.

509   (V-BB) The TIMER database showed that AMPD1 expression was negatively correlated with

510   the tumor purity and was positively correlated with the infiltration of six types of immune cells

511   in lung adenocarcinoma.

512   (CC-II) The TIMER database showed that DOK2 expression was negatively correlated with the

513   tumor purity and was positively correlated with the infiltration of six types of immune cells in

514   lung adenocarcinoma.

515

516   **Figure 8. mRNA and protein expression and immune and stromal status of CCR6 in**

517   **LUAD.**

518   (A–B) mRNA and protein expression of CCR6 in primary LUAD tissues as compared with

519   normal samples (UALCAN). (C–D) CCR6 protein was not expressed in normal lung tissues,

520   whereas high expression was observed in LUAD tissues (The Human Protein Atlas). (E–F) Bar

521   charts of 22 immune cell proportions (CIBERSORT) and 14 stromal cell proportions (xCell)

522   between the groups with low and high expression of CCR6 in LUAD. (G–H) Differential

523   expression of different types of immune and stromal cells between the groups with low and high

524   expression of CCR6 in LUAD. (I) The association of SCNAs of CCR6 with immune infiltration

525   in LUAD (TIMER).

526

527   **TABLE LEGENDS**

528   Table 1. The relationship between the two scores and clinical characteristics.

529

530   **SUPPORTING INFORMATION**

531   Table S1. List of DEGs in the immune score group and the stromal score group.

532   Table S2. 286 survival-related genes were analyzed using univariate Cox regression.

533   Figure S1. Comparative expression of immune cells and stromal cells between the groups with

534   low and high expression of CCR6 in LUAD.

Figure 1

Relationships among immune and stromal scores, clinical characteristics, and overall survival.

Distribution of stromal and immune scores based on gender (A), T stage (B), M stage (C) ,

and pathological stage (D) . Kaplan–Meier survival analysis of low and high score groups

based on stromal (E) and immune scores (F). (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$)

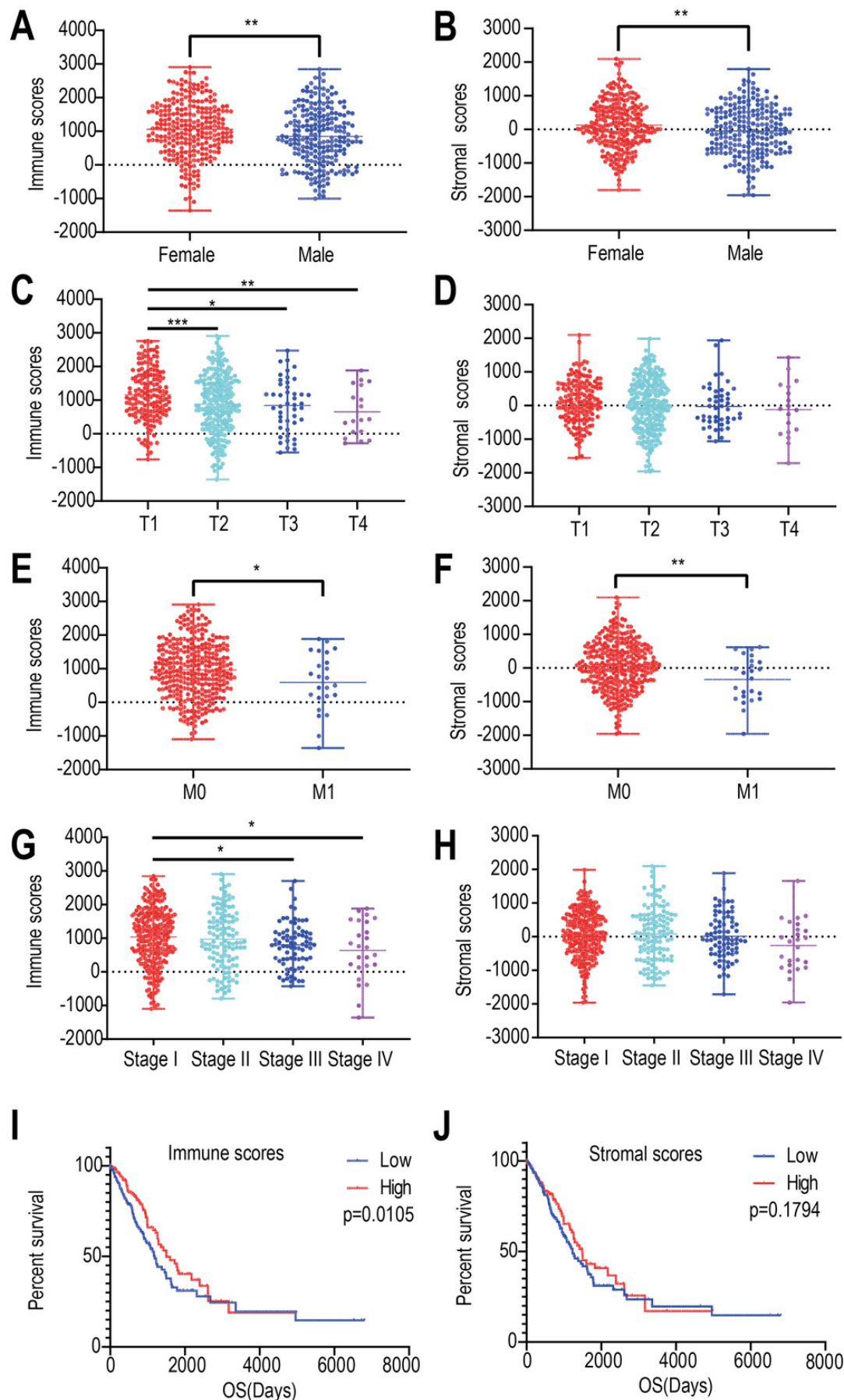



Figure 2

Comparison of gene expression profiles with immune scores and stromal scores in LUAD.

Heatmaps were drawn for the DEGs obtained from the respective high and low group comparisons of the immune scores (A) and stromal scores (B). Venn diagrams drawn from the DEGs that are up- or down-regulated in immune scores (C) and stromal scores (D).

A **Immune scores**

Low   High

B **Stromal scores**

Low   High

C Up-regulated DEGs in stromal score group

(15.4%)   18
(27.7%)   37
(56.9%)

Up-regulated DEGs in immune score group

D Down-regulated DEGs in stromal score group

487
(35.5%)   512
(37.3%)   372
(27.1%)

Down-regulated DEGs in immunescore group

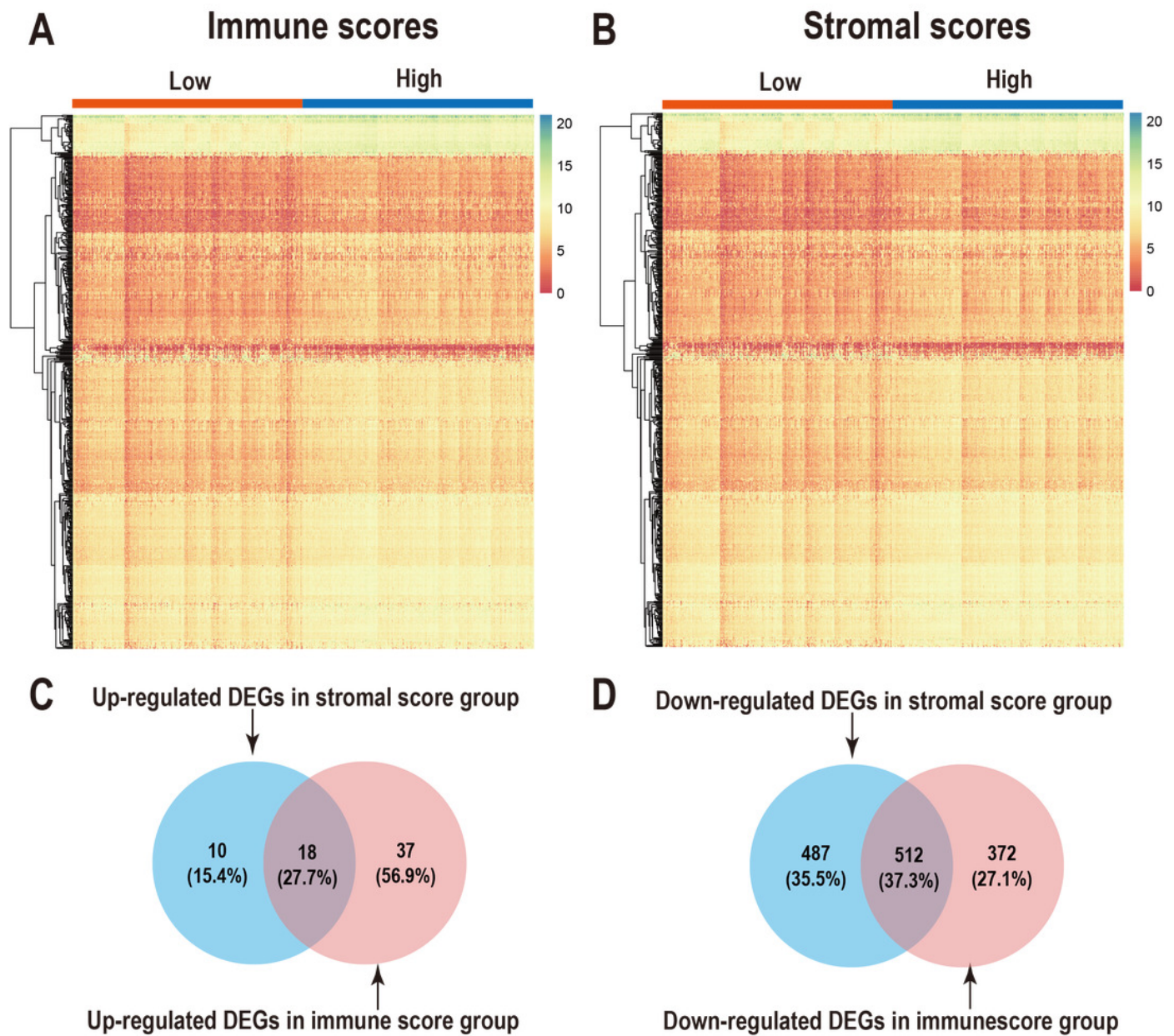



Figure 3

PPI network of 286 prognostic DEGs.

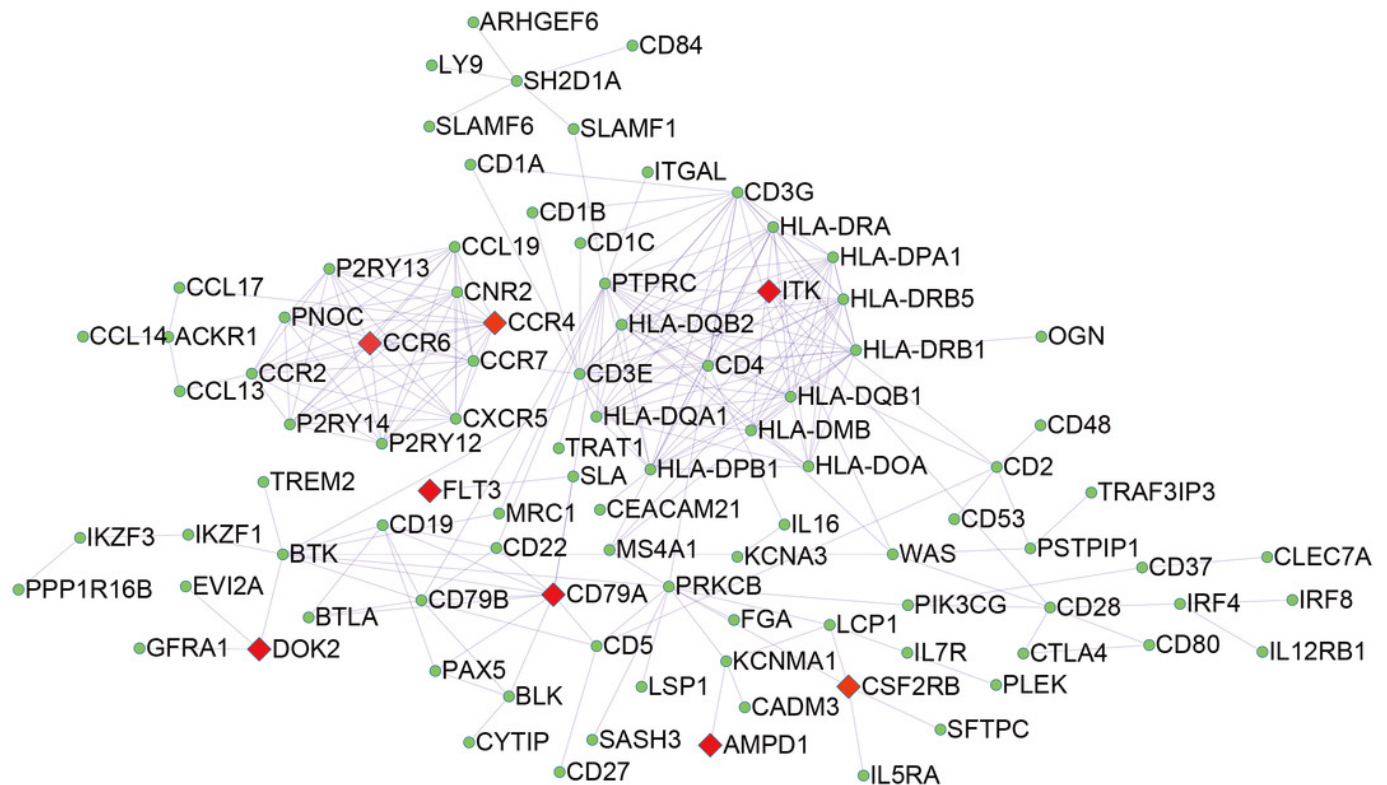



Figure 4

GO term and KEGG analysis of 286 DEGs.

(A) GO analysis of 286 DEGs . (B) KEGG enrichments of 286 DEGs. GO – gene ontology; KEGG – Kyoto Encyclopedia of Genes and Genomes; MF – molecular function; CC – cellular component; BP – biological process.

ID	Description
GO:0006955	immune response
GO:0002250	adaptive immune response
GO:0031295	T cell costimulation
GO:0002504	antigen processing and presentation of peptide or polysaccharide antigen via MHC class
GO:0050776	regulation of immune response
GO:0009897	external side of plasma membrane
GO:0005887	integral component of plasma membrane
GO:0005886	plasma membrane
GO:0042613	MHC class II protein complex
GO:0016021	integral component of membrane
GO:0071556	integral component of lumenal side of endoplasmic reticulum membrane
GO:0032395	MHC class II receptor activity
GO:0004888	transmembrane signaling receptor activity
GO:0004872	receptor activity
GO:0042605	peptide antigen binding
GO:0030246	carbohydrate binding

Figure 5

Correlation between expression of nine prognosis-related genes and overall survival in GEO.

(A) The Venn diagram shows prognosis-related genes based on GSE11969, GSE37745, and TCGA cross-validation. (B) The Venn diagram shows prognosis-related genes based on GSE50081, GSE37745, and TCGA cross-validation. (C) Kaplan–Meier survival curves were generated for nine prognosis genes (AMP11, CCR4, CCR6, CD79A, CSF2RB, DOK2, FLT3, IGJ, and ITK) extracted from the comparison of groups of high (red line) and low (blue line) gene expression.

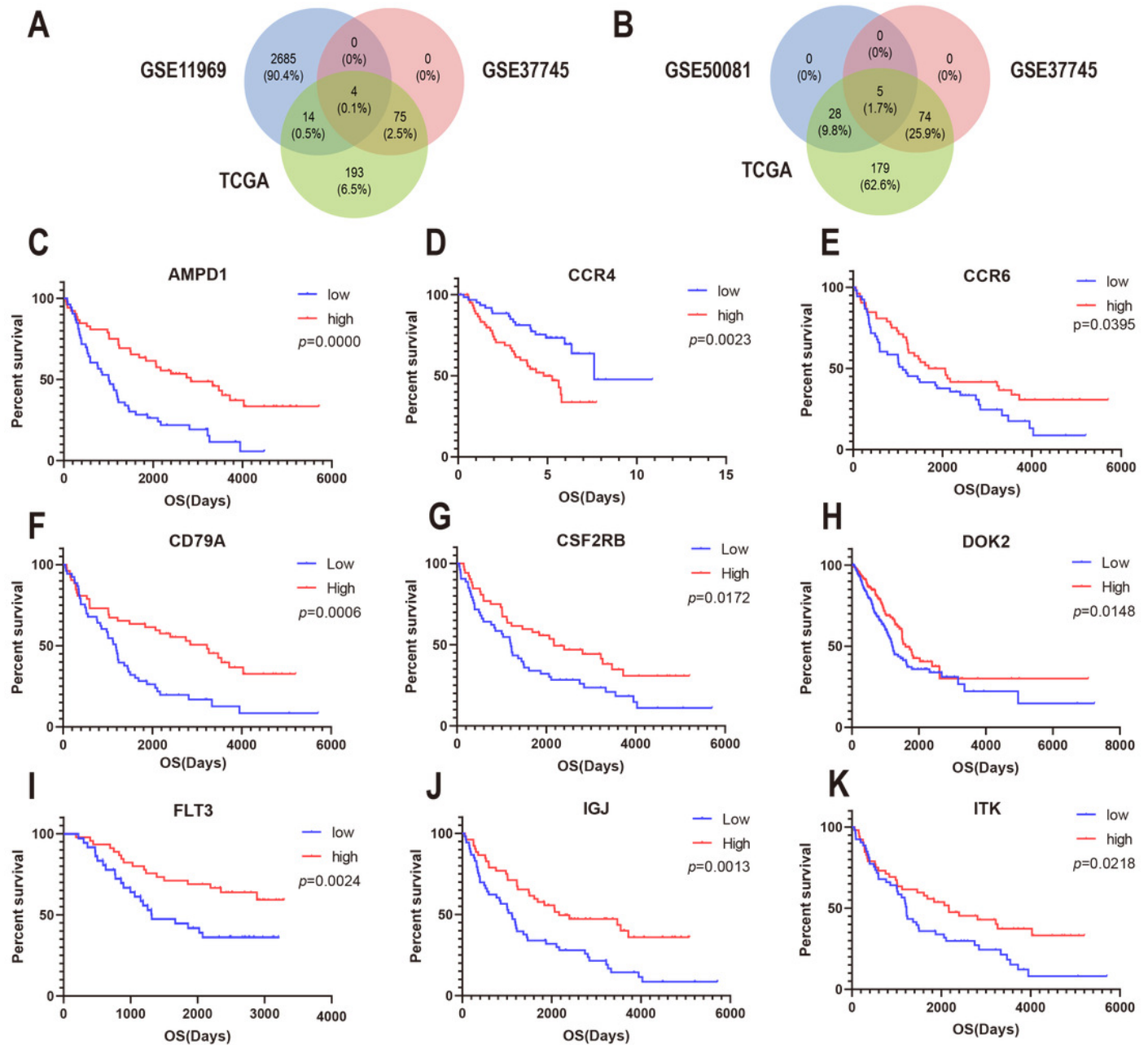



Figure 6

The multivariate Cox regression analysis of 9 prognostic genes.

The HR and P-value from the multivariate Cox regression of selected prognostic genes in TCGA (A), GSE11969 (B), GSE37745 (C), GSE50081 (D). (* p < 0.05, ** p < 0.01)

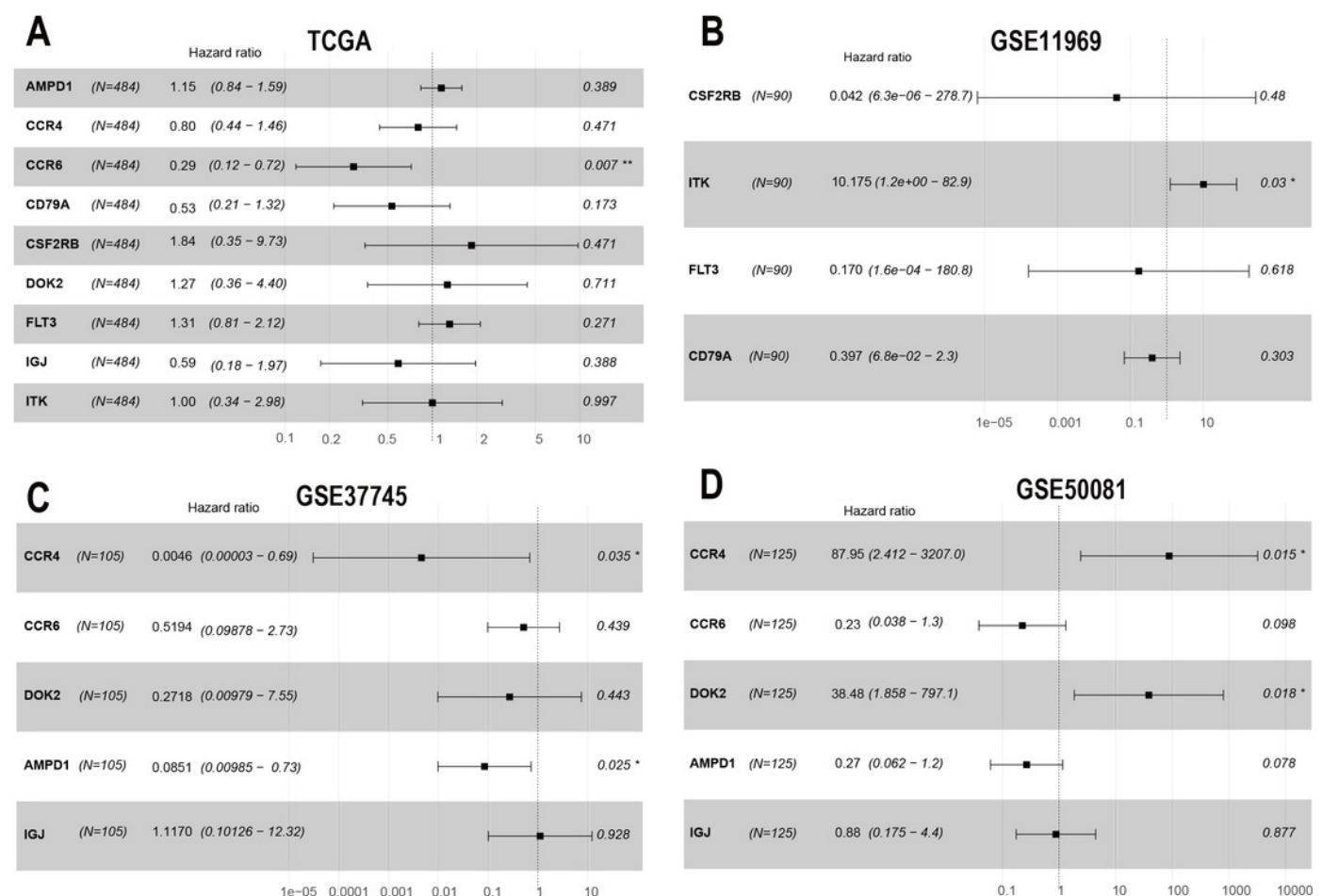



Figure 7

Correlation analysis between the expressions of mainly identified DEGs (CCR6, ITK, CCR4, DOK2, AMPD1) and infiltration levels of B cell, CD8+ T cell, CD4+ T cell, macrophage, neutrophil, and dendritic cell.

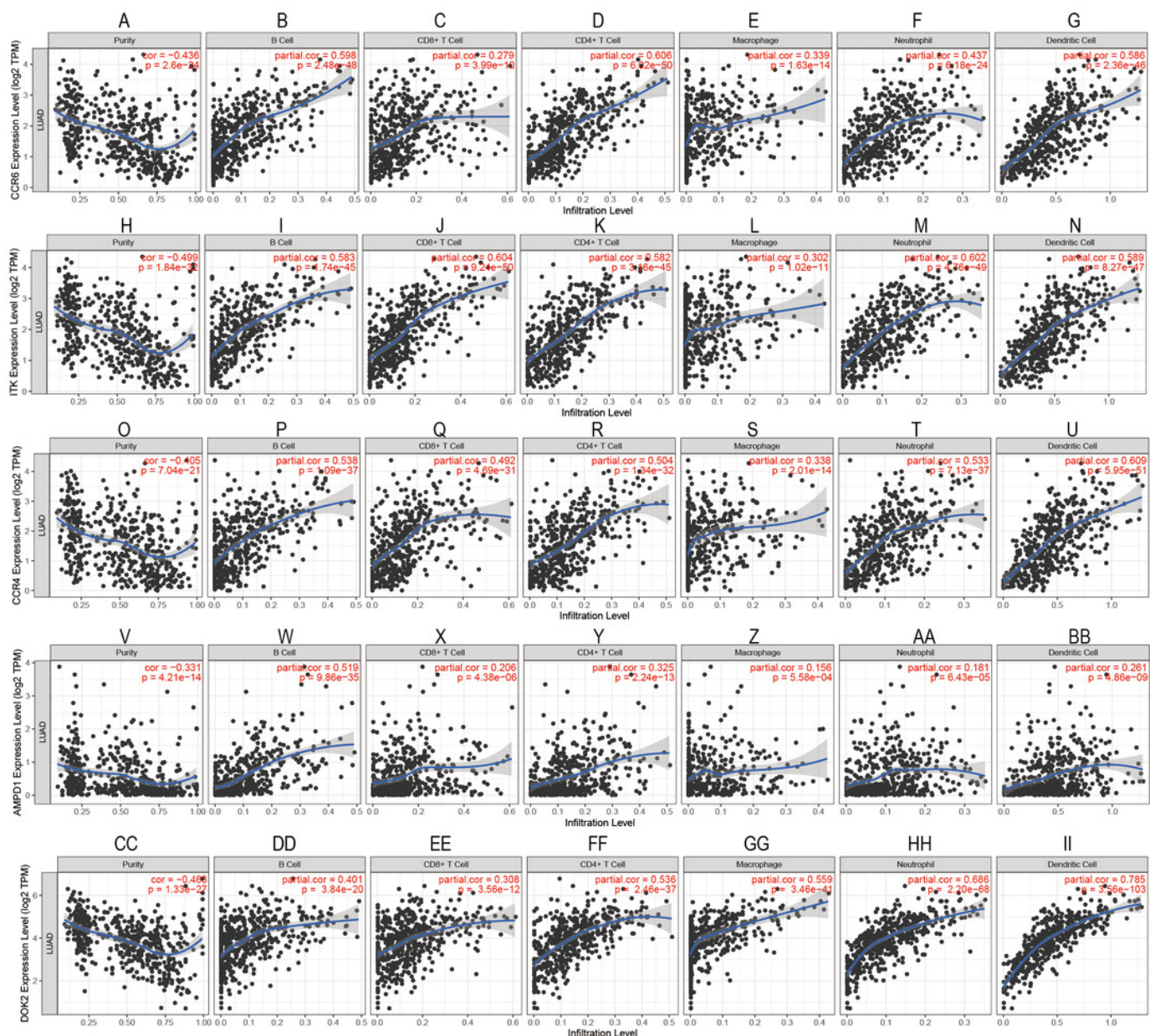



Figure 8

mRNA and protein expressions, immune and stromal status of CCR6 in LUAD.

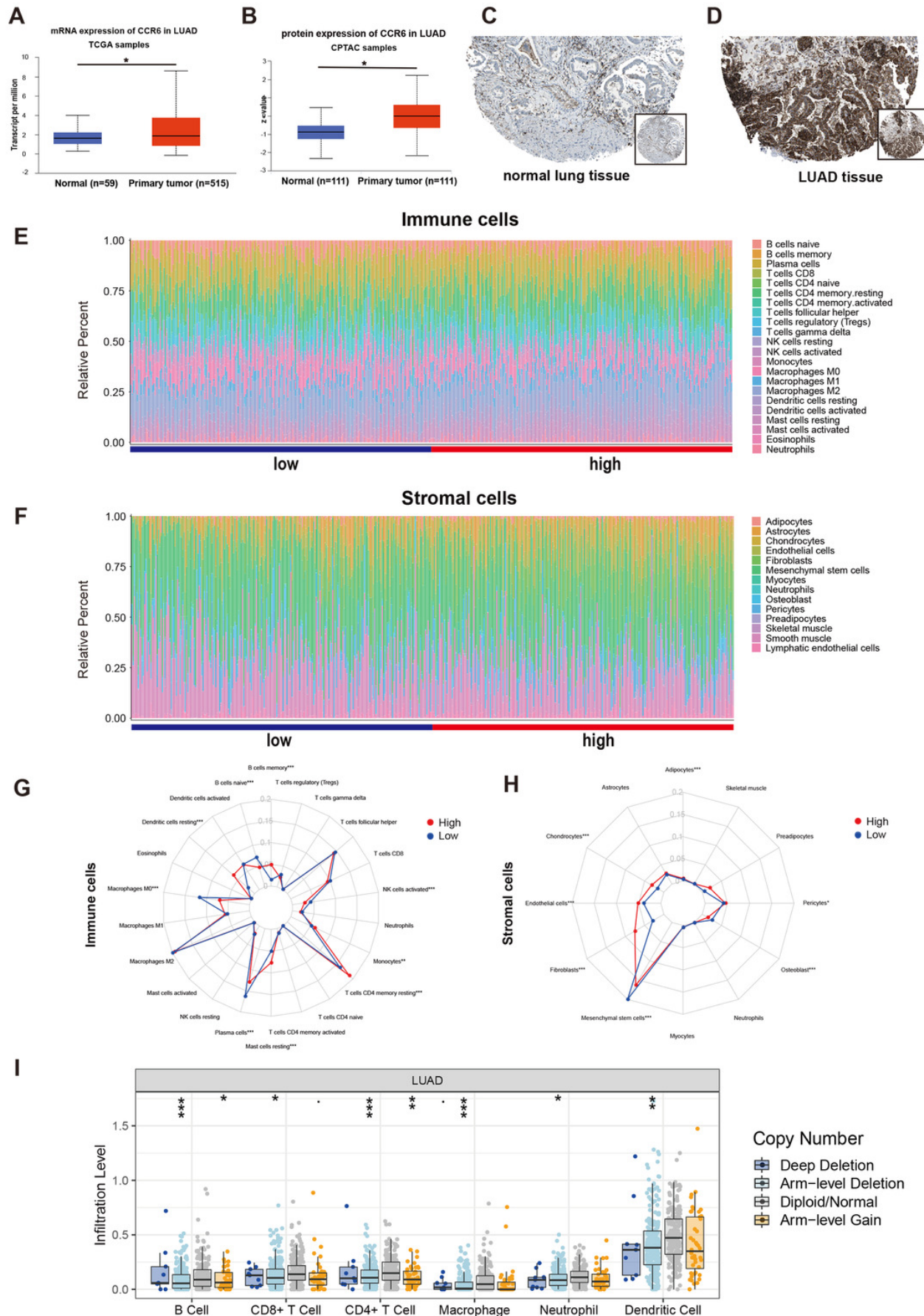
(A) mRNA and protein expressions of CCR6 in primary LUAD tissues compared to normal samples (UALCAN). (B) CCR6 protein were not expressed in normal lung tissues, whereas its high expressions were observed in LUAD tissues (The Human Protein Atlas). (C) Bar charts of 22 immune cell proportions (CIBERSORT) and 14 stromal cell proportions (xCell) between low and high expression of CCR6 in LUAD. (D) Differential expression of different types of immune cells and stromal cells between low and high expression of CCR6 in LUAD . (E) The association of SCNAs of CCR6 with immune infiltration in LUAD (TIMER).

The relationship between the two scores and clinical characteristics.

Table 1. The relationship between the two scores and clinical characteristics.

Variable	Patients	Stromal scores	P value	Immune scores	P value
Scores			<0.000**		<0.000**
Low	246	−539.40 ± 27.69		490.60 ± 45.44	
High	246	635 ± 26.75		1437 ± 41.31	
Age, years			0.223		0.100
≤60	157	235.10		231.1	
>60	335	251.84		253.72	
Sex			0.007**		0.005**
Male	225	−48.34 ± 48.91		849.60 ± 55.64	
Female	267	128.80 ± 43.52		1060 ± 49.79	
T stage			0.101		0.001**
T1	167	260.99		277	
T2/T3/T4	325	238.85		230.40	
N stage			0.768		0.155
N0	318	242.34		247.45	
N1/N2/N3	163	238.39		228.42	
M stage			0.010**		0.037*
M0	324	54.67 ± 40.45		961.70 ± 46.01	
M1	24	−343.60 ± 137.40		592.90 ± 177.2	
Stage			0.137		0.023*
I	262	249.35		256.45	
II	118	246.36		242.92	
III	79	232.61		210.16	
IV	25	183.72	□	196.48	□

Tips: p < 0.05*; p < 0.01**