Identification of candidate genes and prognostic value analysis in patients with PDL1-positive and PDL1-negative lung adenocarcinoma

Xiaoguang Qi [1] , **Chunyan Qi** [2] , **Xindan Kang** [1] , **Yi Hu** [Corresp., 1] , **Weidong Han** [3]

[1] Department of Oncology, Chinese PLA General Hospital, Beijing, China

[2] Department of Special Ward, Chinese PLA General Hospital, Beijing, China

[3] Department of Bio-therapeutic, Chinese PLA General Hospital, Beijing, China

Corresponding Author: Yi Hu
Email address: huyi0401@aliyun.com

Background: Increasing bodies of evidence reveal that targeting a programmed cell death protein 1 (PD-1) monoclonal antibody is a promising immunotherapy for lung adenocarcinoma. Although PD receptor ligand 1 (PDL1) expression is widely recognized as the most powerful predictive biomarker for anti-PD-1 therapy, its regulatory mechanisms in lung adenocarcinoma remain unclear. Therefore, we conducted this study to explore differentially expressed genes (DEGs) and elucidate the regulatory mechanism of PDL1 in lung adenocarcinoma. **Methods:** The GSE99995 data set was obtained from the Gene Expression Omnibus (GEO) database. Patients with and without PDL1 expression were divided into PDL1-positive and PDL1-negative groups, respectively. DEGs were screened using R. The Gene Ontology (GO) database and Kyoto Encyclopedia of Genes and Genomes (KEGG) were analyzed using the Database for Annotation, Visualization and Integrated Discovery . Protein–protein interaction (PPI) networks of DEGs was visualized using Cytoscape, and the MNC algorithm was applied to screen hub genes. A survival analysis involving Gene Expression Profiling Interactive Analysis was used to verify the GEO results. Mutation characteristics of the hub genes were further analyzed in a combined study of five datasets in The Cancer Genome Atlas (TCGA) database. **Results:** In total, 869 DEGs were identified, 387 in the PDL1-positive group and 482 in the PDL1-negative group. GO and KEGG analysis results of the PDL1-positive group mainly exhibited enrichment of biological processes and pathways related to cell adhesion and peroxisome proliferators-activated receptors (PPAR) signaling pathway, whereas biological process and pathways associated with cell division and repair were mainly enriched in the PDL1-negative group. The top 10 hub genes were screened during the PPI network analysis. Notably, survival analysis revealed *BRCA1*, mainly involved in cell cycle and DNA damage responses, to be a novel prognostic indicator in lung adenocarcinoma. Moreover, the

prognosis of patients with different forms of lung adenocarcinoma was associated with differences in mutations and pathways in potential hub genes. **Conclusions:** PDL1-positive lung adenocarcinoma and PDL1-negative lung adenocarcinoma might be different subtypes of lung adenocarcinoma. The hub genes might play an important role in PDL1 regulatory pathways. Further studies on hub genes are warranted to reveal new mechanisms underlying the regulation of PDL1 expression. These results are crucial for understanding and applying precision immunotherapy for lung adenocarcinoma.

1 **Identification of candidate genes and prognostic value analysis in patients with PDL1-**

2 **positive and PDL1-negative lung adenocarcinoma**

3 Xiaoguang Qi[1], Chunyan Qi[2], Xindan Kang[1], Yi Hu [1]* and Weidong Han[3]

4 [1] Department of Oncology, Chinese PLA General Hospital, Beijing, China

5 [2] Department of Special Ward, Chinese PLA General Hospital, Beijing, China

6 [3] Department of Bio-therapeutic, Chinese PLA General Hospital, Beijing, China

9 Corresponding Author: Yi Hu,

10 No. 28 Fuxing Road, Haidian District, Beijing, 100853, China.

11 Email address: huyi0401@aliyun.com.

22 **ABSTRACT**

23 **Background:** Increasing bodies of evidence reveal that targeting a programmed cell death

24 protein 1 (PD-1) monoclonal antibody is a promising immunotherapy for lung adenocarcinoma.

25 Although PD receptor ligand 1 (PDL1) expression is widely recognized as the most powerful

26 predictive biomarker for anti-PD-1 therapy, its regulatory mechanisms in lung adenocarcinoma

27 remain unclear. Therefore, we conducted this study to explore differentially expressed genes

28 (DEGs) and elucidate the regulatory mechanism of PDL1 in lung adenocarcinoma. **Methods:**

29 The GSE99995 data set was obtained from the Gene Expression Omnibus (GEO) database.

30 Patients with and without PDL1 expression were divided into PDL1-positive and PDL1-negative

31 groups, respectively. DEGs were screened using R. The Gene Ontology (GO) database and

32 Kyoto Encyclopedia of Genes and Genomes (KEGG) were analyzed using the Database for

33 Annotation, Visualization and Integrated Discovery. Protein–protein interaction (PPI) networks

34 of DEGs was visualized using Cytoscape, and the MNC algorithm was applied to screen hub

35 genes. A survival analysis involving Gene Expression Profiling Interactive Analysis was used to

36 verify the GEO results. Mutation characteristics of the hub genes were further analyzed in a

37 combined study of five datasets in The Cancer Genome Atlas (TCGA) database. **Results:** In total,

38   869 DEGs were identified, 387 in the PDL1-positive group and 482 in the PDL1-negative group.

39   GO and KEGG analysis results of the PDL1-positive group mainly exhibited enrichment of

40   biological processes and pathways related to cell adhesion and peroxisome proliferators-

41   activated receptors (PPAR) signaling pathway, whereas biological process and pathways

42   associated with cell division and repair were mainly enriched in the PDL1-negative group. The

43   top 10 hub genes were screened during the PPI network analysis. Notably, survival analysis

44   revealed *BRCA1*, mainly involved in cell cycle and DNA damage responses, to be a novel

45   prognostic indicator in lung adenocarcinoma. Moreover, the prognosis of patients with different

46   forms of lung adenocarcinoma was associated with differences in mutations and pathways in

47   potential hub genes. **Conclusions:** PDL1-positive lung adenocarcinoma and PDL1-negative lung

48   adenocarcinoma might be different subtypes of lung adenocarcinoma. The hub genes might play

49   an important role in PDL1 regulatory pathways. Further studies on hub genes are warranted to

50   reveal new mechanisms underlying the regulation of PDL1 expression. These results are crucial

51   for understanding and applying precision immunotherapy for lung adenocarcinoma.

52   **INTRODUCTION**

53   Non–small-cell lung cancer (NSCLC) is a leading malignancy threatening human life and health

54   worldwide (Herbst, Morgensztern, & Boshoff,2018; Zhang et al.,2019). Lung adenocarcinoma is

55   the most common type of NSCLC, and it is a more heterogeneous subtype of NSCLC compared

56   with squamous cell carcinoma. Therefore, its individualized treatment and prognosis have

57   attracted considerable attention; understanding its biological characteristics is necessary for

58   achieving such individualized treatment and prognosis. In the past decade, studies have identified

59 tyrosine kinase inhibitors (TKIs) targeting epidermal growth factor receptor (EGFR), anaplastic

60 lymphoma kinase (ALK), and ROS proto-oncogene 1 (ROS1) as potential therapies for lung

61 adenocarcinoma on the basis of genotyping (Sgambato et al., 2018; Singhi et al., 2019).

62 Molecular targeted therapy based on these sensitive targets has considerably enhanced overall

63 survival (OS) in lung adenocarcinoma. Despite the progress in the treatment of lung

64 adenocarcinoma, the mortality rate among patients with advanced lung adenocarcinoma remains

65 high, with the 5-year survival rate being approximately 15% in patients with advanced NSCLC

66 (Blandin Knight et al., 2017). Nevertheless, a breakthrough was achieved in terms of

67 immunotherapy for cancer according to recent reports, with immune checkpoint inhibitors

68 prolonging survival in some patients with various cancer types (Yu & Wang, 2018; Lorigan

69 & Eggermont, 2019; Morse, Hochster, & Benson, 2020). In particular, targeting programmed

70 cell death protein 1 (PD-1) monoclonal antibodies is rapidly becoming a promising therapeutic

71 approach for NSCLC treatment (El-Osta & Jafri, 2019). Patients with >50% PDL1 expression

72 treated using anti-PD-1 therapy showed a longer progression-free survival (PFS) period than did

73 those treated using standard chemotherapy as the first-line treatment for NSCLC. Furthermore, in

74 the past few years, some of these antibodies have been successfully commercialized and

75 approved as first- and second-line treatments for advanced NSCLC (Gridelli et al., 2018; Dafni

76 et al., 2019). In particular, immunotherapy is different from traditional chemotherapy or targeted

77 therapy in that it has durable clinical benefits and fewer side effects in general (Suresh et al.,

78 2018). Hence, immunotherapy is a promising therapy for lung adenocarcinoma; however, its

79 application is limited to a subtype of patients with lung adenocarcinoma. Individualized

80 immunotherapy guided by predictive biomarkers is crucial for improving the prognosis of lung

81 adenocarcinoma.

82 Recent studies have proposed PDL1 expression, tumor mutational burden, and DNA mismatch

83 repair deficiency as biomarkers for anti-PD-1 therapy (Teng et al., 2018; Darvin et al., 2018; El-

84 Osta & Jafri, 2019). In particular, PDL1 expression—recommended by the National

85 Comprehensive Cancer Network guidelines—is widely recognized as the most powerful

86 predictive marker for immunotherapy in lung adenocarcinoma. Moreover, patients with PDL1-

87 positive lung adenocarcinoma are more likely to benefit from immunotherapy than those with

88 PDL1-negative lung adenocarcinoma (Gridelli et al., 2018; Dafni et al., 2019), suggesting that

89 PDL1 expression plays a major role in the pathogenesis of lung adenocarcinoma. However, the

90 mechanism underlying this clinical problem remains unclear. Understanding the regulatory role

91 of PDL1 expression in lung adenocarcinoma is crucial for precision immunotherapy. Currently,

92 PDL1 expression is primarily determined using immunohistochemistry (IHC) assays;

93 nevertheless, IHC assays yield inconsistent results due to variable cutoffs and different

94 antibodies with differing affinities (Hunter, Socinski, & Villaruz, 2018). Hence, such assays

95 cannot comprehensively elucidate the molecular mechanism and enriched pathways underlying

96 the efficacy of immunotherapy. To solve this problem, novel strategies are imperative to explore

97 the intrinsic mechanism of PDL1 expression in the biology of lung adenocarcinoma.

98 Because of advances in next-generation sequencing technology, high-throughput sequencing

99 results on cancers have been increasingly released on the Gene Expression Omnibus (GEO), a

100 public database; this can thus enable understanding the molecular mechanism of PDL1

101   expression and pathogenesis of lung adenocarcinoma. Using next-generation sequencing

102   technology, previous studies have successfully identified key biomarkers in lung

103   adenocarcinoma, which has proved to be essential in understanding the molecular mechanism of

104   tumors (Tang et al., 2018; Li et al., 2019). Because PDL1 expression might affect certain key

105   pathways in immunotherapy for lung adenocarcinoma, exploring the differences between

106   patients with PDL1-positive and those with PDL1-negative lung adenocarcinoma would provide

107   insights into the regulatory mechanisms of immunotherapy. Hence, it is reasonable to explore the

108   differences in gene expression profiles and enriched pathways between the two groups of

109   patients. Using the GEO database, we studied a gene expression dataset (GEO accession number:

110   GSE99995) and comparatively analyzed sequencing data of gene expression between patients

111   with PDL1-positive and PDL1-negative lung adenocarcinoma using the Agilent oligonucleotide

112   microarray system. The aim of the study was to explore the differences in gene expression

113   profiles and enriched pathways between the two groups of patients and identify the potential

114   biomarkers predicting patient prognosis in order to elucidate the regulatory mechanism of PDL1

115   in immunotherapy, which could be crucial in guiding precision immunotherapy and improve

116   prognosis in lung adenocarcinoma.

117   **MATERIAL AND METHODS**

118   **Gene expression profile data**

119   The GEO database is a public functional genomics data repository

120   (https://www.ncbi.nlm.nih.gov/geo/). We obtained gene expression profile data from the public

121   GEO database based on the keywords "PDL1 expression," "lung adenocarcinoma," and "homo

122   sapiens." GSE99995 was retrieved. To prevent interferon expression levels having effects on

123   PDL1 gene expression profiles, only patients with lung adenocarcinoma with low interferon

124   expression levels were included. Patients with and without PDL1 expression were divided into

125   PDL1-positive (3 people, average age: 52.33 ± 4.04 years, tumor stage IIIA) and PDL1-negative

126   (3 people, average age: 51.33 ± 3.06 years, tumor stage IIIA) groups, respectively.

127   **DEGs analysis and mapping**

128   A series of matrix files and platforms were downloaded and processed using R. Screening of

129   differentially expressed genes (DEGs) between the two groups as well as heat mapping and

130   volcano mapping were performed using Limma. A log FoldChange of >2 and an adjusted P

131   value of <0.05 indicated the presence of DEGs.

132   **Gene Ontology and Kyoto Encyclopedia of Genes and Genomes pathway enrichment**

133   **analysis**

134   We used Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG)

135   pathway enrichment analyses to understand gene functional annotation and functional

136   enrichment, respectively. We used the Database for Annotation, Visualization and Integrated

137   Discovery (DAVID) to perform GO and KEGG annotation of DEGs. DAVID is an online

138   database (https://david.ncifcrf.gov/) for estimating functional domains and biological

139   implications. Fisher's exact test was employed for analyses of pathways, diseases, and functions.

140   A P value of <0.05 is recommended because it denotes the significance of GO terms and KEGG

141   pathway enrichment in genes. The top 10 GO terms and KEGG pathway enrichment results were

142   mapped using Hmisc and ggplot2 in R.

143   **Gene set enrichment analysis pathway enrichment analysis and validation**

144   Gene Set Enrichment Analysis (GSEA) is a computational method—based on the analysis of all

145   genes—that determines whether an a priori defined set of genes shows statistically significant

146   differences between two biological states. In the KEGG pathway analysis, the threshold value

147   was set according to the expression level, which could only be analyzed in genes with

148   significantly different expression levels. To avoid limitations in our results, GSEA was used

149   again to consider the effect of all DEGs, not limited to those with significantly different

150   expression levels. Therefore, GSEA was used to comprehensively analyze the differences in

151   gene pathway enrichment results between the two groups.

152   **Protein–protein interaction network and analysis of hub genes**

153   STRING is an online database designed to evaluate and predict protein–protein interactions

154   (PPIs) (https://string-db.org/cgi/input.pl). First, STRING was used in this study to analyze the

155   PPI network of different genes in the two patient groups. Isolated nodes were removed, and the

156   results of the interaction network were downloaded and then imported into Cytoscape (version

157   3.7.2) for subsequent analysis. Second, the PPI network was constructed and visualized using

158   Cytoscape (version 3.7.2) and cytoHubba. The MNC algorithm was employed to screen and

159   identify the hub genes that might be key candidate genes with crucial regulatory functions.

160   **Clinical characteristics and survival analysis related to hub genes**

161   Gene Expression Profiling Interactive Analysis (GEPIA) is an online cancer data mining site that

162   is based on the TCGA and GTEx databases and uses a standard processing pipeline

163   (http://gepia.cancer-pku.cn/detail.php). In this study, the online survival analysis tool of GEPIA

164   was used to verify our previous results obtained using the TCGA and GTEx databases.

165   Furthermore, differences in the expression of hub genes in lung adenocarcinoma and adjacent

166   normal tissues were analyzed. Moreover, the relationships among differences in expression of

167   hub genes, pathological staging, and prognosis in lung adenocarcinoma were analyzed.

168   Furthermore, P values of <0.05 were considered statistically significant.

169   **Mutation characteristics and pathway analyses of hub genes**

170   We investigated the mutation characteristics and possible functional mechanisms of the

171   differential expression of hub genes in lung adenocarcinoma. Therefore, data sets containing

172   gene expression profiles of patients with lung adenocarcinoma were included in the TCGA

173   database (e.g., Broad, Cell 2012; MSKCC, Science [2015]; TCGA, Firehose Legacy; TCGA,

174   Nature [2014]; and TCGA, PanCancer Atlas). A combined study of five data sets including 1598

175   patients and 1600 samples were included in the analysis.

176   GSCALite is a web-based analysis platform that analyzes an entire set of genes in cancers

177   (http://bioinfo.life.hust.edu.cn/web/GSCALite/), including alterations in DNA or RNA of cancer-

178 related genes, the activity of 10 cancer-related pathways, and miRNA regulatory network for

179 genes. To understand the cancer-related pathways of hub genes, we distinguished hub gene

180 expression between pathway activation and inhibition groups according to pathway scores in

181 GSCALite.

182 **RESULTS**

183 **Identification and mapping of DEGs**

184 Our results revealed that the PDL1-positive and PDL1-negative groups comprised 34,729 genes.

185 According to criteria of a log FoldChange of >2, and an adjusted P value of <0.05, a total of 869

186 DEGs were identified, with 387 and 482 DEGs in the PDL1-positive and PDL1-negative groups,

187 respectively. The corresponding heat map and volcano map are shown in Figure 1.

188 **GO Terms and KEGG pathway enrichment analysis of DEGs**

189 The analysis of DEGs by using GO Terms was divided into three parts, namely biological

190 processes, cellular components, and molecular functions. The GO analysis of the PDL1-positive

191 group mainly revealed the enrichment of pathways involved in cell adhesion (BP, GO:0007155)

192 plasma membrane (CC, GO:0005886), and cadmium ion binding (MF, GO:0046870). The GO

193 analysis of the PDL1-negative group mainly revealed the enrichment of pathways involved in

194 cell division (BP, GO:0051301), nucleoplasm (CC, GO:0005654), and DNA binding

195 (MF,GO:0003677). The detailed results are shown in Table 1. The KEGG pathway enrichment

196 results in the PDL1-positive group revealed that the enriched pathways were mainly those

197   involved in cell adhesion molecules, PPAR signaling, and ECM-receptor interaction. However,

198   the KEGG pathway enrichment results of the PDL1-negative group revealed that the enriched

199   pathways were mainly those involved in DNA replication, cell cycle, and mismatch repair (Table

200   2). The detailed results are shown in Table 2. The top 10 GO terms and KEGG enriched

201   pathways were mapped using the Hmisc and ggplot2 packages. The results are shown in Figure 2.

202   **GSEA enrichment analysis and validation**

203   The results were validated through GSEA analysis after all the DEGs were considered. The

204   GSEA results also provided the KEGG results. However, a new pathway enrichment was

205   observed in the PDL1-positive group, namely the T-cell receptor pathway, which was not

206   observed in KEGG results of DEGs (Fig. 3).

207   **Construction of PPI network and analysis of hub genes**

208   The STRING database consists of known and predicted PPIs. The interactions include direct

209   (physical) and indirect (functional) associations. To explore the expression relationships between

210   the DEGs, a PPI network was constructed in our study. Our results showed that the PPI network

211   consisted of 570 nodes and 2937 edges (Fig. 4A). The average node degree was 10.3 and the

212   average local clustering coefficient was 0.404 (PPI enrichment p value: <1.0e-16). In addition,

213   the MNC algorithm was used to screen and identify hub genes that might be key genes with

214   crucial regulatory functions. The top 10 hub genes, BUB1B, CDC45, BUB1, TTK, BRCA1,

215   TOP2A, NDC80, RFC4, MCM2, and DTL, were identified in our study (Fig. 4B). The results

216   were ranked using the MNC method, as shown in Table 3.

217   Analysis of the mechanism of hub genes in the TCGA database

218   To prevent the limitations caused by using a single GEO database, our results were verified using

219   the TCGA and GTEx databases. First, comparative analysis of hub gene expression in lung

220   adenocarcinoma and adjacent normal tissue was performed in our study, and our results showed

221   that the expression levels of BUB1B, CDC45, BUB1, TTK, TOP2A, NDC80, MCM2, and DTL

222   were significantly higher in lung adenocarcinoma tissues than in adjacent normal lung tissues (P

223   < 0.01). However, although the expression levels of BRCA1 and RFC4 were higher in lung

224   adenocarcinoma tissues than in adjacent normal tissues, the differences were not significant (Fig.

225   5).

226   In addition, the relationship between hub genes and pathological staging were analyzed. The

227   results showed that BUB1B (F value = 5.22, p = 0.00148), CDC45 (F value = 2.86, P = 0.0364),

228   BUB1 (F value = 5.22, P = 0.00149), TTK (F value = 5.06, P = 0.00185), BRCA1 (F value = 4.9,

229   P = 0.0023), TOP2A (F value = 2.88, P = 0.0354), NDC80(F value = 3.58, P = 0.014), MCM2 (F

230   value = 2.78, P = 0.0407), and DTL (F value = 5.97, P = 0.000535) were positively correlated

231   with pathological staging, and the differences were statistically significant. However, regarding

232   the correlation between RFC4 and pathological staging, although RFC4 was positively correlated

233   with pathological staging, the expression of RFC4 did not differ significantly between the two

234   groups (F value = 1.82, P = 0.142; Fig. 6).

235   Finally, the relationship between the hub genes and prognosis was analyzed. The results showed

236   that patients with lung adenocarcinoma with low expression levels of BUB1B (Log rank P =

237   3.8e−05, P (HR) = 5.1e−05), CDC45 (log rank P = 0.0032, p (HR) = 0.0035), BUB1 (log rank P

238   = 0.0024, p (HR) = 0.0026), TTK [log rank P = 0.00029, P (HR) = 0.00034], BRCA1 [log rank

239   P = 0.0035, P (HR) = 0.0038], TOP2A [log rank P = 0.011, P (HR) = 0.012], NDC80 [log rank P

240   = 0.0027, P (HR) = 0.003], MCM2 [log rank P = 0.02, P (HR) = 0.021], and DTL [log rank P =

241   0.0016, P (HR) = 0.0018] had significantly higher overall survival than those with high gene

242   expression levels (P < 0.05). However, no significant correlation was observed between RFC4

243   gene expression level and overall survival (log rank P = 0.059; Fig. 7).

244   **Analysis of the mechanism of hub genes in the TCGA database**

245   Furthermore, analysis of the lung adenocarcinoma was performed using the database TCGA. The

246   results showed that the hub genes had different mutation frequencies, ranging from 1% to 5%.

247   For example, BUB1B (4%), CDC45 (1.9%), and DTL (5%) (Fig. 8A). Moreover, the potential

248   hub genes exhibited different mutation forms in lung adenocarcinoma, which may facilitate

249   further exploration of the function of these hub genes in lung adenocarcinoma. For example,

250   BUB1B mainly exhibited deep deletion, whereas *BRCA1* mainly exhibited amplification (Fig.

251   8B). However, these different mutant forms of hub genes were involved in different signaling

252   pathways influencing the development of lung adenocarcinoma; for example, *BRCA1* was

253   mainly involved in the cell cycle and DNA damage response pathways (Fig. 9A, B).

254   **DISCUSSION**

255   Lung adenocarcinoma is a highly heterogeneous type of cancer, and its individualized treatment

256   has attracted considerable attention. An increasing number of studies have recently indicated that

257   immunotherapy, especially anti-PD-1 therapy, is a promising strategy for treating lung

258   adenocarcinoma. Some anti-PD-1 antibodies have been successfully commercialized and

259   approved as first- and second-line immunotherapy options for advanced NSCLC in the past few

260   years. Nevertheless, the overall response rate for such immunotherapy is only approximately

261   20% (Greillier, Tomasini, & Barlesi, 2018), and the efficacy of such immunotherapy is affected

262   by PDL1 expression. In particular, PDL1 expression has been proposed as a predictive

263   biomarker, suggesting that it plays a major role in immune regulation in lung adenocarcinoma.

264   Therefore, understanding the effects of PDL1 expression on the biological behavior and the

265   efficacy of immunotherapy in lung adenocarcinoma is crucial. However, the use of IHC alone to

266   assess PDL1 expression cannot provide a complete explanation for the molecular mechanism and

267   enriched pathways underlying the efficacy of immunotherapy. With the development of next-

268   generation sequencing technology and bioinformatics, a novel strategic solution to this problem

269   may be identified (Morganti et al., 2019). Furthermore, previous studies have used bioinformatic

270   analysis to explore the core genes of lung adenocarcinoma and their malignant transformation

271   mechanisms (Yuan et al., 2017; Yeh et al., 2019). However, no research has been conducted on

272   PDL1-positive and PDL1-negative patients. Understanding DEGs and differences in the

273   biological processes and enrichment pathways between PDL1-positive and PDL1-negative

274   patients is crucial. Hence, by using gene bioinformatics analysis, we explored the differences in

275   gene expression profiles and enrichment pathways between the two groups of patients and

276 identified the potential key biomarkers that could be used for predicting disease prognosis in

277 patients. The elucidation of the regulatory mechanism of PDL1 is crucial to improve precision

278 immunotherapy for lung adenocarcinoma.

279 Our study demonstrated that even in patients with lung adenocarcinoma, differences in gene

280 expression were observed between those in the PDL1-positive and PDL1-negative group, thus

281 suggesting that these two types of patients should be treated differently.

282 The important findings in our study in the GO Terms and KEGG pathway analysis were the

283 differences between the PDL1-positive and PDL1-negative groups of patients with lung

284 adenocarcinoma. Our results reveal that the PDL1-positive group mainly demonstrated

285 enrichment of adhesion-related biological processes and the PPAR signaling pathway. Adhesion-

286 related biological processes have been previously reported to be associated with invasion and

287 metastasis in patients with lung adenocarcinoma (Stevens et al., 2017; Alonso-Nocelo et al.,

288 2018), resulting in relatively high levels of organ and tissue metastases and poor prognosis. Real-

289 world studies have revealed similar results (Wang et al., 2015; Okita et al., 2017). Previous

290 studies have reported that PPAR, an anti-inflammatory molecule with a potent, could promote

291 tumor proliferation, angiogenesis, inflammation, and metastasis in lung adenocarcinoma (Reka et

292 al., 2011; Ammu et al., 2019), suggesting that it is a vital and potential therapeutic target for lung

293 adenocarcinoma. Moreover, Lv and Wang (2015) discovered that the PPAR signaling pathway

294 played a major role in malignant transformation of cells among nonsmoking patients with lung

295 adenocarcinoma. These results are consistent with our findings. Although several studies have

296   suggested the relationship between PPAR and lung adenocarcinoma, the role of PPAR in

297   immunotherapy for lung adenocarcinoma has yet to be reported. Moreover, recent studies have

298   demonstrated that PPAR could enhance indoleamine 2,3-dioxgenase-1 (IDO) activity and

299   promote the generation of regulatory T cells in melanoma; hence, the use of PPAR inhibitors

300   could enhance cancer immunotherapy (Poupot et al., 2014). On the basis of our results along

301   with the results of these studies, we hypothesized that the PPAR pathway might play a crucial

302   role in immunotherapy for lung adenocarcinoma. A comprehensive analysis of our results

303   indicated that although PDL1 expression was associated with the immune response pathway, it

304   might have potential to contribute to invasion and metastasis in patients with PDL1-positive lung

305   adenocarcinoma.

306   Conversely, the results of the PDL1-negative group mainly revealed enrichment of division-

307   related biological processes and repair systems pathways, which indicated that these pathways

308   might be able to repair defects in apoptosis or recombination to reduce the production of

309   neoantigens. Some clinical studies have shown that PDL1-negative patients tend to have lower

310   tumor mutation burdens or fewer neoantigens than PDL1-positive patients (Xia et al., 2017;Chan

311   et al., 2019), which might be related to its involvement in various repair systems. Further

312   research must be conducted to confirm our conclusions.

313   In addition, the GSEA results were similar to those for the KEGG pathway analysis. More

314   importantly, a notable and novel finding was reported in the PDL1-positive group, namely that

315   related to the T-cell receptor pathway, which has been proved to be crucial for achieving a

316   favorable immunotherapeutic response in NSCLC (Van De & Borst, 2015). This finding

317   suggests that PDL1 might be involved in the regulation of the T-cell pathway; therefore,

318   understanding the effects of PDL1 expression on tumor microenvironmental immune cells and

319   mediation of T-cell pathways is necessary.

320   In our study, 10 hub genes screened in the GEO database were verified and further analyzed

321   using the TCGA and GTEx databases. Our study further analyzed the prognostic value and

322   possible mechanism of hub genes in lung adenocarcinoma considering their importance.

323   Several hub genes, such as BUB1B, CDC45, BUB1, TTK, TOP2A, MCM2, NDC80, and DTL,

324   have been reported to be associated with poor prognosis in patients with lung adenocarcinoma

325   (Hayama et al., 2006; He et al., 2019; Liu et al., 2017; Perez-Peña et al., 2017; Song et al., 2018;

326   Sun et al., 2020). Our findings corroborate these results. Another major finding of our study is

327   that BRCA1 is a predictor in patients with lung adenocarcinoma.

328   BRCA1 encodes a nuclear phosphoprotein that helps maintain genomic stability. It is involved in

329   biological processes such as cell cycle regulation, replication, and mitotic spindle assembly.

330   Previous studies have suggested that BRCA1 overexpression regulates drug response in

331   chemotherapy, is related to the efficacy of EGFR-TKIs, and is prevalent in patients with NSCLC

332   with early disease onset (Reguart et al., 2008; Sun et al., 2018). Recent studies have analyzed the

333   potential prognostic role of BRCA1 in early-stage NSCLC, suggesting that BRCA1 is a predictor

334   of survival in only stage III NSCLC (Hu et al., 2019). Although several studies have

335   demonstrated the role of BRCA1 in NSCLC, the prognostic value and possible mechanism of

336 BRCA1 in lung adenocarcinoma remain unclear. Our study results suggest that BRCA1

337 overexpression was associated with poor prognosis in lung adenocarcinoma. In particular, our

338 results indicate that BRCA1 was mainly involved in cell cycle and DNA damage responses in

339 lung adenocarcinoma, which might produce new antigens and enhance immune responses.

340 BRCA1 mutation was also reported to be associated with tumor neoantigen production, immune

341 cell invasion, and PDL1 expression in ovarian cancer (Strickland et al., 2016). Overall, these

342 findings suggest that BRCA1 plays a key role in the immunomodulatory pathway of lung

343 adenocarcinoma.

344 RFC4, also named activator 1, is a protein complex consisting of five subunits measuring 140, 40,

345 38, 37, and 36 kD. Telomere C-strand (Lagging Strand) Synthesis and E2F-mediated regulation

346 of DNA replication are some of its related pathways. A previously conduced weighted gene co-

347 expression network analysis revealed that it was a prognosis-related biomarker in lung

348 adenocarcinoma (Yi et al., 2020). However, our GEPIA survival analysis showed that RFC4

349 overexpression was positively correlated with pathological staging but was not associated with

350 poor prognosis in patients with lung adenocarcinoma. Possible reasons for these inconsistent

351 findings are differences in sample selection strategies and algorithms. Accordingly, the

352 prognostic value of RFC4 for lung adenocarcinoma warrants further study.

353 The present study has several limitations. First, only a single GSE was searched and analyzed in

354 the GEO database. To prevent bias resulting from the analysis of a single data set, we adopted an

355 integrated bioinformatic analysis approach. Second, the main objective of this study was to

356 explore the influence of PDL1 expression on related gene expression profiles in patients with

357 lung adenocarcinoma. Hence, this study was divided into expression in PDL1-positive and

358 PDL1-negative groups; however, no further distinction was made in the PDL1-positive group

359 between patients with high (50%) and low (<1%) levels of PDL1 expression. Different PDL1

360 expression levels might have caused differences in the expression profiles of genes, which

361 warrants further confirmation. Third, some key genes have been studied in various tumors,

362 including non–small-cell lung cancer. However, this study identified core genes that might be

363 related to or regulated by the expression of PDL1 in lung adenocarcinoma, which will assist in

364 the further discovery of mechanisms for designing novel immunotherapeutic options.

365 This study revealed the DEGs and different biological pathways between PDL1-positive and

366 PDL1-negative patients with lung adenocarcinoma, speculating that these two types of patients

367 might have different subtypes of lung adenocarcinoma. The identification and verification of hub

368 genes through integrated bioinformatic analysis revealed that they related to immune response

369 pathways and prognosis in patients with lung adenocarcinoma. In particular, they had different

370 mutations and were involved in different pathways in lung adenocarcinoma.

371 **CONCLUSIONS**

372 PDL1-positive lung adenocarcinomaand PDL1-negative lung adenocarcinoma might be different

373 subtypes of lung adenocarcinoma. Potential hub genes might be involved in PDL1 regulatory pathways,

374 and further research is warranted to reveal new mechanisms underlying the regulation of PDL1 expression.

375 This could be of great significance for precision immunotherapy for lung adenocarcinoma.

REFERENCES

376   1.Herbst R S , Morgensztern D, Boshoff C. 2018.The biology and management of non-small cell lung cancer.

377   Nature 553: 446-454 DOI: 10.1038/nature25183.

379   2.Zhang L, Peng R, Sun Y, Wang J, Chong X, Zhang Z. 2019. Identification of key genes in non-small cell lung

380   cancer by bioinformatics analysis. Peer J 7:e8215 DOI: 10.7717/peerj.8215.

381   3.Singhi E K, Horn L, Sequist L V, Heymach J, Langer C J. 2019. Advanced Non–Small Cell Lung Cancer:

382   Sequencing Agents in the EGFR-Mutated/ALK-Rearranged Populations. American Society of Clinical Oncology

383   Educational Book 39: e187-e197 DOI: 10.1200/EDBK_237821.

384   4.Sgambato A, Casaluce F, Maione P, Gridelli C. 2018. Targeted therapies in non-small cell lung cancer: a focus on

385   ALK/ROS1 tyrosine kinase inhibitors. Expert review of anticancer therapy 18: 71-80

386   DOI: 10.1080/14737140.2018.1412260.

387   5.Blandin Knight S, Crosbie PA, Balata H, Chudziak J, Hussell T, Dive C. 2017. Progress and prospects of early

388   detection in lung cancer. Open biology 7: 170070 DOI: 10.1098/rsob.170070.

389   6.Yu X, Wang X. 2018. Tumor immunity landscape in non-small cell lung cancer. Peer J 6:e4546

390   DOI: 10.7717/peerj.4546.

391   7.Lorigan P, Eggermont AMM. 2019. Anti-PD1 treatment of advanced melanoma: development of criteria for a safe

392   stop. annals of oncology 30:1038-1040 DOI: 10.1093/annonc/mdz182.

393   8.Morse MA, Hochster H, Benson A. 2020. Perspectives on Treatment of Metastatic Colorectal Cancer with

394   Immune Checkpoint Inhibitor Therapy. Oncologist 25:33-45 DOI: 10.1634/theoncologist.2019-0176.

395   9. El-Osta H, Jafri S. 2019. Predictors for clinical benefit of immune checkpoint inhibitors in advanced non-small-

396   cell lung cancer: a meta-analysis. Immunotherapy 11:189-199   DOI: 10.2217/imt-2018-0086.

397   10.Dafni U, Tsourti Z, Vervita K, Peters S. 2019. Immune checkpoint inhibitors, alone or in combination with

398   chemotherapy, as first-line treatment for advanced non-small cell lung cancer. A systematic review and network

399   meta-analysis. Lung Cancer 134:127-140 DOI: 10.1016/j.lungcan.2019.05.029.

400   11.Gridelli C, Ascierto PA., Grossi F, Baldini E, Favaretto A, Garassino MC, Morabito A,Migliorino M R, Rossi A,

401   de Marinis F. 2018. Second-line treatment of advanced non-small cell lung cancer non-oncogene addicted: New

402   treatment algorithm in the era of novel immunotherapy. Current clinical pharmacology 13: 76-84

403   DOI: 10.2174/1574884713666180711160008.

404   12.Suresh K, Naidoo J, Lin CT, Danoff S. 2018. Immune checkpoint immunotherapy for non-small cell lung cancer:

405   benefits and pulmonary toxicities. Chest 154: 1416-1423 DOI: 10.1016/j.chest.2018.08.1048.

406   13.Teng, F, Meng, X., Kong, L, Yu J, 2018. Progress and challenges of predictive biomarkers of anti PD-1/PD-L1

407   immunotherapy: a systematic review. Cancer letters 414:166-173 DOI: 10.1016/j.canlet.2017.11.014.

408   14.Darvin, P, Toor, SM, Nair, VS, Elkord E. 2018. Immune checkpoint inhibitors: recent progress and potential

409   biomarkers. Experimental & molecular medicine 50:1-11 DOI: 10.1038/s12276-018-0191-1.

410   15.Hunter KA, Socinski MA, Villaruz LC. 2018. PD-L1 testing in guiding patient selection for PD-1/PD-L1

411   inhibitor therapy in lung cancer. Molecular diagnosis & therapy 22:1-10 DOI: 10.1007/s40291-017-0308-6.

412   16.Li Z, Sang M, Tian Z, Liu Z, Lv J, Zhang F, Shan B. 2019. Identification of key biomarkers and potential

413   molecular mechanisms in lung cancer by bioinformatics analysis. Oncology letters 18: 4429-4440

414   DOI: 10.3892/ol.2019.10796.

415   17.Tang Q, Zhang H, Kong M, Mao X, Cao X. 2018. Hub genes and key pathways of non-small lung cancer

416   identified using bioinformatics. Oncology letters 16: 2344-2354 DOI: 10.3892/ol.2018.8882.

417   18.Greillier L, Tomasini P, Barlesi F. 2018. The clinical utility of tumor mutational burden in non-small cell lung

418   cancer. Translational lung cancer research. 7:639-646 DOI: 10.21037/tlcr.2018.10.08.

419   19.Morganti S, Tarantino P, Ferraro E, D'Amico P, Duso BA, Curigliano G. 2019. Next Generation Sequencing

420   (NGS): A Revolutionary Technology in Pharmacogenomics and Personalized Medicine in Cancer. Advances in

421   Experimental Medicine and Biology 1168: 9-30 DOI: 10.1007/978-3-030-24100-1_2.

422   20.Yuan J, Zhang N, Yin L, Zhu H, Zhang L, Zhou L, Yang M. 2017. Clinical implications of the autophagy core

423   gene variations in advanced lung adenocarcinoma treated with gefitinib. Scientific reports 7: 1-10

424   DOI: 10.1038/s41598-017-18165-5.

425   21.Yeh SJ, Chang CA, Li CW, Wang LH, Chen BS. 2019. Comparing progression molecular mechanisms between

426   lung adenocarcinoma and lung squamous cell carcinoma based on genetic and epigenetic networks: big data

427   mining and genome-wide systems identification. Oncotarget 10:3760-3806.DOI: 10.18632/oncotarget.26940.

428   22.Stevens LE, Cheung WK, Adua SJ, Arnal-Estapé A, Zhao M, Liu Z, Brewer K, Herbst RS, Nguyen DX. 2017.

429   Extracellular matrix receptor expression in subtypes of lung adenocarcinoma potentiates outgrowth of

430   micrometastases. Cancer research 77: 1905-1917 DOI: 10.1158/0008-5472.CAN-16-1978.

431   23.Alonso-Nocelo M, Raimondo TM, Vining KH, López-López R, de la Fuente M, Mooney DJ. 2018. Matrix

432   stiffness and tumor-associated macrophages modulate epithelial to mesenchymal transition of human

433   adenocarcinoma cells. Biofabrication 10: 035004 DOI: 10.1088/1758-5090/aaafbc.

434   24.Okita R, Maeda A, Shimizu K, Nojima Y, Saisho S, Nakata M. 2017. PD-L1 overexpression is partially regulated

435   by EGFR/HER2 signaling and associated with poor prognosis in patients with non-small-cell lung cancer. Cancer

436   Immunology, Immunotherapy 66: 865-876 DOI: 10.1007/s00262-017-1986-y.

437   25.Wang A, Wang HY, Liu Y, Zhao MC, Zhang HJ, Lu ZY, Fang YC, Chen XF,Liu GT. 2015. The prognostic

438   value of PD-L1 expression for non-small cell lung cancer patients: a meta-analysis. European Journal of Surgical

439   Oncology 41: 450-456 DOI: 10.1016/j.ejso.2015.01.020.

440   26.Ammu VRK, Garikapati KK, Krishnamurthy PT, Chintamaneni PK, Pindiprolu SKS.   2019. Possible role of

441   PPAR-γ and COX-2 receptor modulators in the treatment of Non-Small Cell lung carcinoma. Medical hypotheses

442   124: 98-100 DOI: 10.1016/j.mehy.2019.02.024.

443   27.Reka AK, Goswami MT, Krishnapuram R, Standiford TJ, Keshamouni VG. 2011. Molecular cross-regulation

444   between PPAR-γ and other signaling pathways: implications for lung cancer therapy. Lung cancer 72:154-159

445   DOI: 10.1016/j.lungcan.2011.01.019.

446   28.Lv, M., & Wang L. 2015. Comprehensive analysis of genes, pathways, and TFs in nonsmoking Taiwan females

447   with lung cancer. Experimental lung research 41:74-83 DOI: 10.3109/01902148.2014.971472.

448   29.Poupot M, Boissard F, Betous D, Bardouillet L, Fruchon S, L'Faqihi-Olive F, Mekaouche M, Ingoure S, Sicard

449   H, Dubreuilh G, Fournié JJ. 2014. The PPARα pathway in Vγ9Vδ2 T cell anergy. Cellular and Molecular Biology

450   Letters 19: 649-658 DOI: 10.2478/s11658-014-0218-0.

451   30.Xia H, Shen J, Hu F, Chen S, Huang H, Xu Y, Ma H. 2017. PD-L1 over-expression is associated with a poor

452   prognosis in Asian non-small cell lung cancer patients. Clinica Chimica Acta 469: 191-194

453   DOI: 10.1016/j.cca.2017.02.005.

454   31.Chan TA, Yarchoan M, Jaffee E, Swanton C, Quezada SA, Stenzinger A, Peters S. 2019.Development

455   of tumor mutation burden as an immunotherapy biomarker: utility for the oncology clinic. Annals of Oncology 30:

456   44-56 DOI: 10.1093/annonc/mdy495.

457   32.Van De Ven K, Borst J. 2015. Targeting the T-cell co-stimulatory CD27/CD70 pathway in cancer

458   immunotherapy: rationale and potential. Immunotherapy 7: 655-667 DOI: 10.2217/imt.15.32.

459   33.Song YJ, Tan J, Gao XH, Wang LX. 2018. Integrated analysis reveals key genes with prognostic value in lung

460   adenocarcinoma. Cancer management and research 10: 6097 DOI: 10.2147/CMAR.S168636.

461   34.He W, Fu L, Yan Q, Zhou Q, Yuan K, Chen L, Han Y. 2019. Gene set enrichment analysis and meta-analysis

462   identified 12 key genes regulating and controlling the prognosis of lung adenocarcinoma. Oncology letters 17: 5608-

463   5618 DOI: 10.3892/ol.2019.10236.

464   35.Hayama S, Daigo Y, Kato T, Ishikawa N, Yamabuki T, Miyamoto M, Tsuchiya E, Kondo S, Nakamura Y. 2006.

465   Activation of CDCA1-KNTC2, members of centromere protein complex, involved in pulmonary carcinogenesis.

466   Cancer Research 66:10339-10348 DOI: 10.1158/0008-5472.CAN-06-2137.

467   36.Liu YZ, Wang BS, Jiang YY, Cao J, Hao JJ, Zhang Y,Xu X, Cai Y, Wang MR. 2017. MCMs expression in lung

468   cancer: implication of prognostic significance. Journal of Cancer 8: 3641 DOI: 10.7150/jca.20777.

469   37.Perez-Peña J, Corrales-Sánchez V, Amir E, Pandiella A, Ocana A. 2017. Ubiquitin-conjugating enzyme E2T

470   (UBE2T) and denticleless protein homolog (DTL) are linked to poor outcome in breast and lung cancers. Scientific

471   reports 7: 1-8 DOI: 10.1038/s41598-017-17836-7.

472   38.Sun ZY, Wang W, Gao H, Quan-Fang Chen QF.2020. Potential therapeutic targets of the nuclear division cycle

473   80 (NDC80) complexes genes in lung adenocarcinoma J. Cancer 11: 2921-2934 doi:10.7150/jca.41834.

474   39.Reguart N, Cardona AF, Carrasco E, Gomez P, Taron M, Rosell R. 2008. BRCA1: A new genomic marker for

475   non–small-cell lung cancer. Clinical lung cancer 9: 331-339 DOI: 10.3816/CLC.2008.n.048.

476   40.Sun X, Cui F, Yin H, Wu D, Wang N, Yuan M, Fei Y, Wang Q. 2018. Association between EGFR mutation and

477   expression of BRCA1 and RAP80 in non-small cell lung cancer. Oncology letters 16:2201-2206

478   DOI: 10.3892/ol.2018.8938.

479   41.Hu X, Yang D, Li Y, Li L, Wang Y, Chen P, Xu S, Pu X, Zhu W, Deng P, Ye J, Zhang H, LIzaso A, Liu H, Mao

480   X, Huang H, Chu Q, Hu C. 2019. Prevalence and clinical significance of pathogenic germline BRCA1/2 mutations

481   in Chinese non-small cell lung cancer patients. Cancer Biology & Medicine 16:556 DOI: 10.20892/j.issn.2095-

482   3941.2018.0506.

483   42.Strickland KC, Howitt BE, Shukla SA, Rodig S, Ritterhouse LL, Liu JF, Matulonis UA, Konstantinopoulos PA.

484   2016. Association and prognostic significance of BRCA1/2-mutation status with neoantigen load, number of tumor-

485   infiltrating lymphocytes and expression of PD-1/PD-L1 in high grade serous ovarian cancer. Oncotarget 7: 13587

486   DOI: 10.18632/oncotarget.7277.

487   43.Yi M, Li T, Qin S, Yu S, Chu   Q, Li A, Wu K. 2020. Identifying Tumorigenesis and Prognosis-Related Genes

488   of Lung Adenocarcinoma: Based on Weighted Gene Coexpression Network Analysis. BioMed Research

489   International 4169691 DOI: 10.1155/2020/4169691.

Figure 1

Heat map and volcanic plot of DEGs comparing patients with PDL1-positive and PDL1-negative lung adenocarcinoma.

(A) Heat map of DEGs. (B) Volcanic plot of DEGs. Red represents high expression levels in the PDL1-positive group, whereas blue represents high expression levels in the PDL1-negative group. Gray represents no difference in expression levels between the two groups.

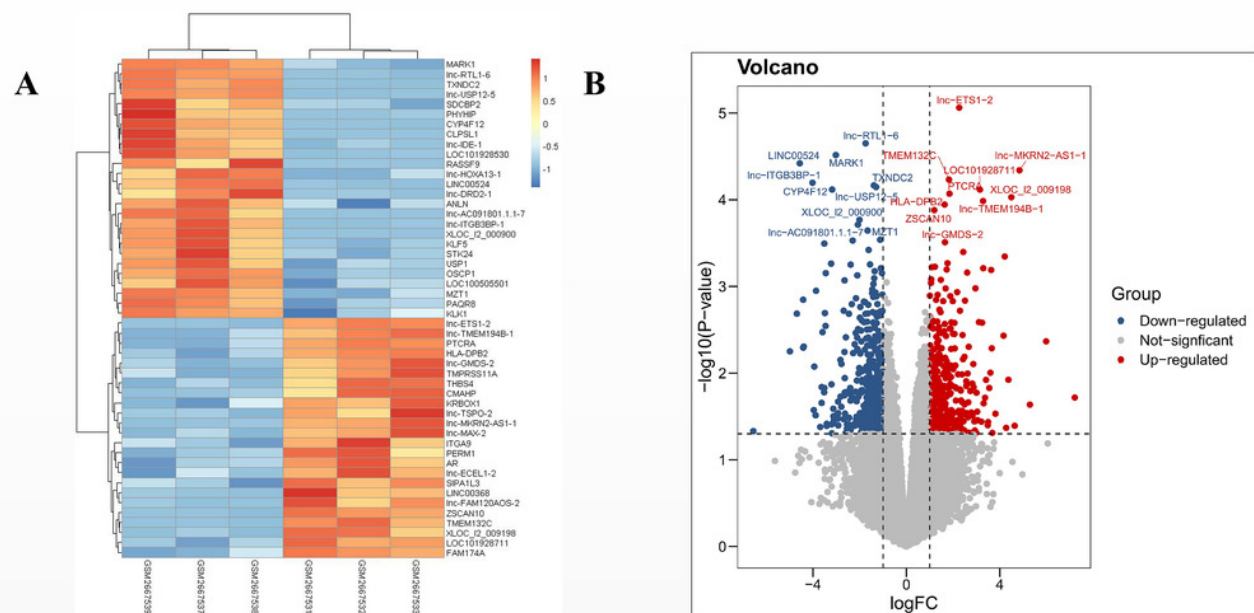



Figure 2

Bubble mapping of the top 10 GO Terms and KEGG pathway enrichment analysis data of DEGs

(A) GO analysis of DEGs in biological process. (B) GO analysis of DEGs in cellular components. (C) GO analysis of DEGs in terms of molecular function. (D) KEGG enrichment analysis of DEGs. A high gene ratio represents a high level of enrichment. The size of the dot indicates the number of target genes in the pathway and the color of the dot reflects the p value range. (E) GO Chord plot of DEGs. (F) KEGG Chord plot of DEGs.

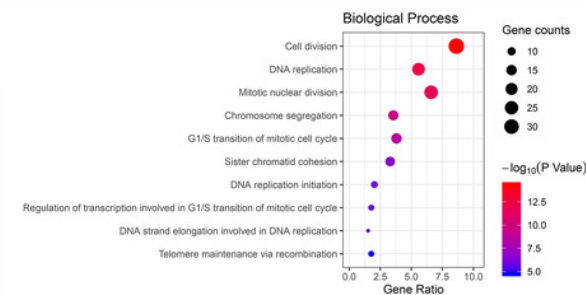
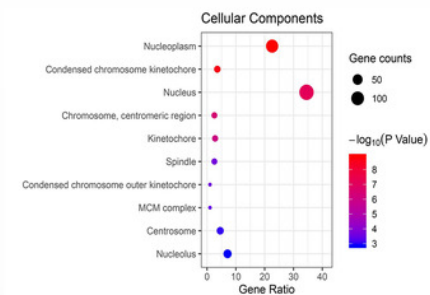
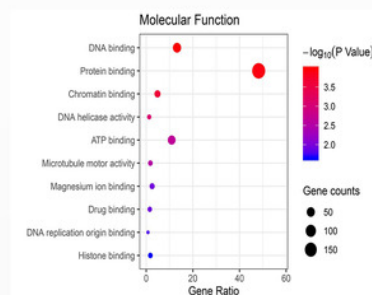
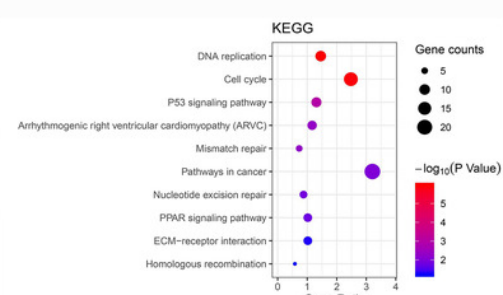
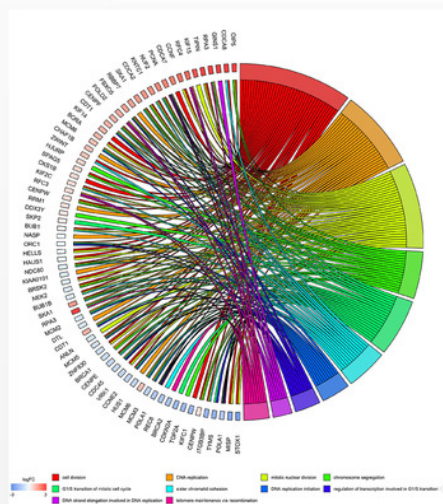
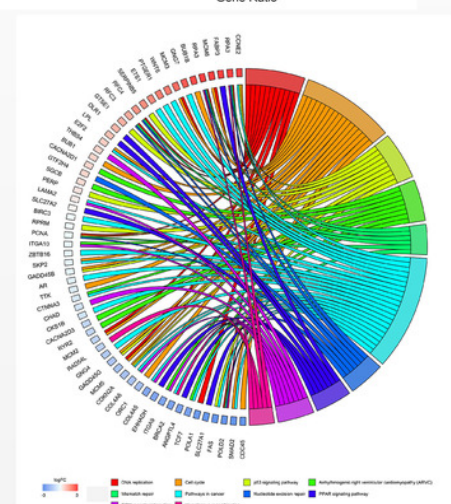



Figure 3

GSEA enrichment analysis and validation.

(A, B, C, D) GSEA enrichment analysis in the PDL1-positive group. (E, F, G, H) GSEA enrichment analysis in the PDL1-negative group .

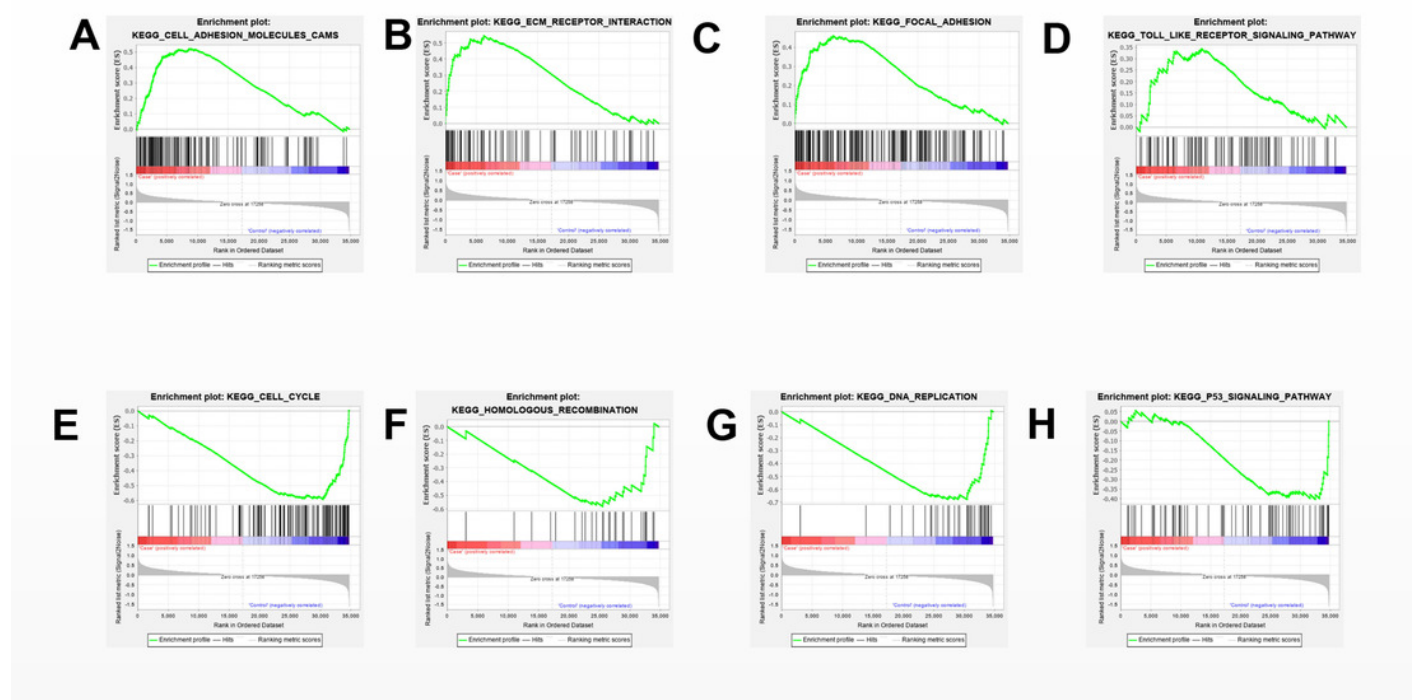



Figure 4

Construction of PPI networks and analysis of hub genes.

(A) The PPI network was constructed using Cytoscape. (B) The top 10 hub genes were screened using the MNC algorithm. The color of the hub gene holds importance according to the MNC algorithm.

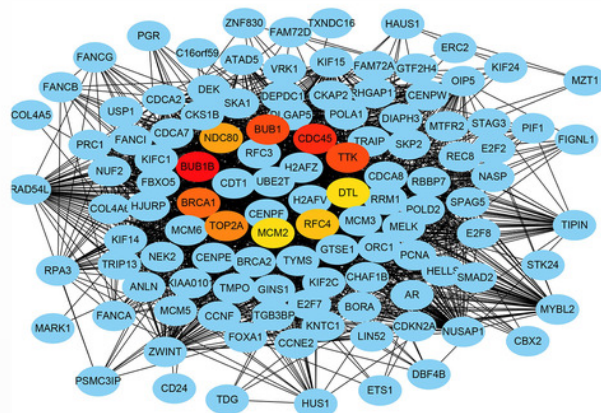
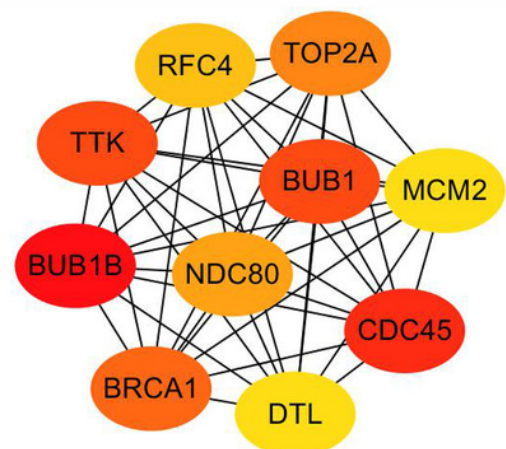



Figure 5

Analysis of hub gene expression in lung adenocarcinoma.

The red and gray boxes represent lung adenocarcinoma and adjacent normal tissue, respectively. (A) BUB1B; (B) CDC45; (C) BUB1, (D) TTK; (E) BRCA1; (F) TOP2A;(G) NDC80; (H) RFC4; (I) MCM2; and (J) DTL.

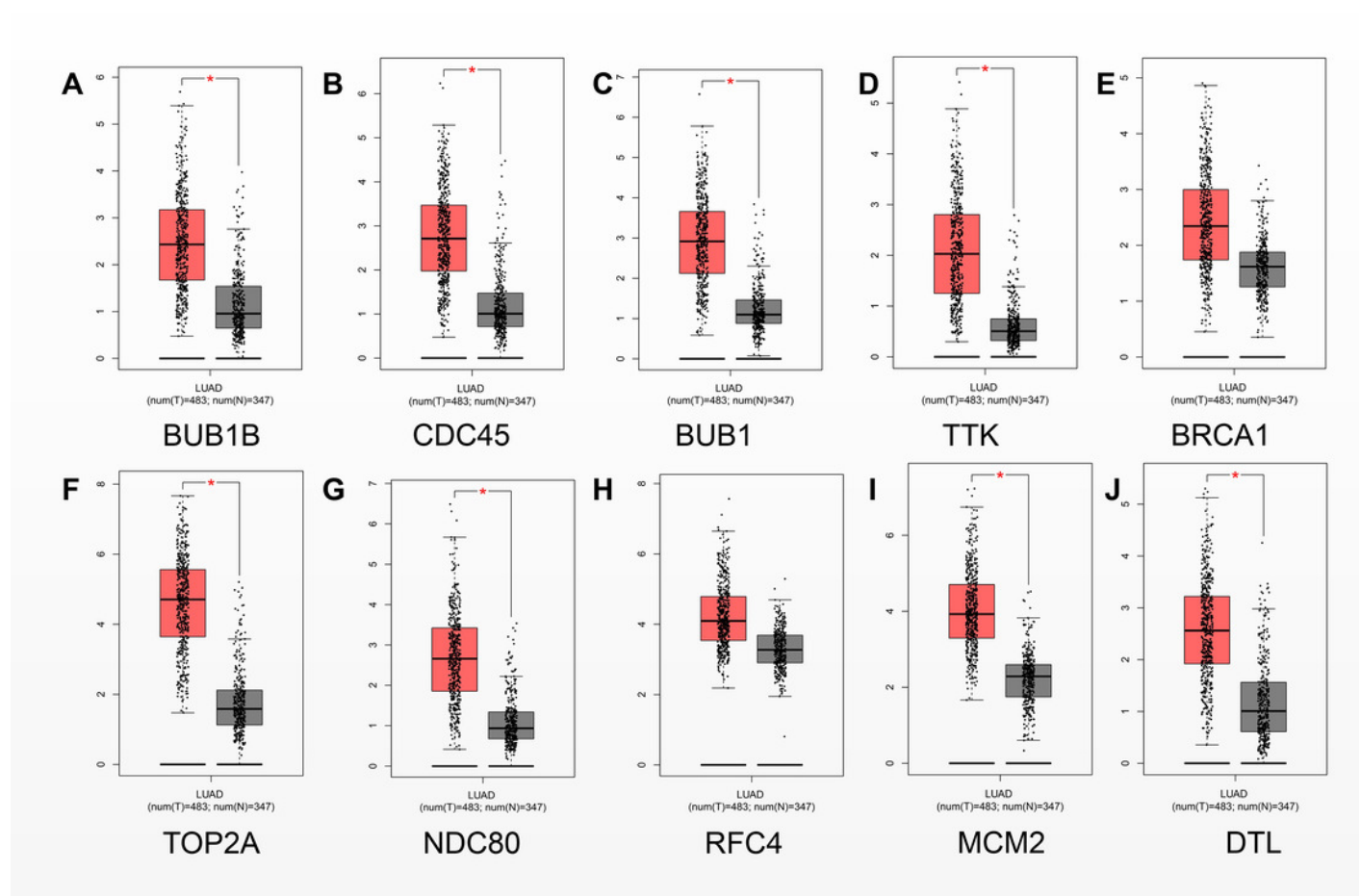



Figure 6

Analysis of the relationship between hub genes and pathological staging in lung adenocarcinoma.

(A) BUB1B; (B) CDC45; (C) BUB1; (D) TTK; (E) BRCA1; (F) TOP2A; (G) NDC80; (H) RFC4; (I) MCM2; and (J) DTL.

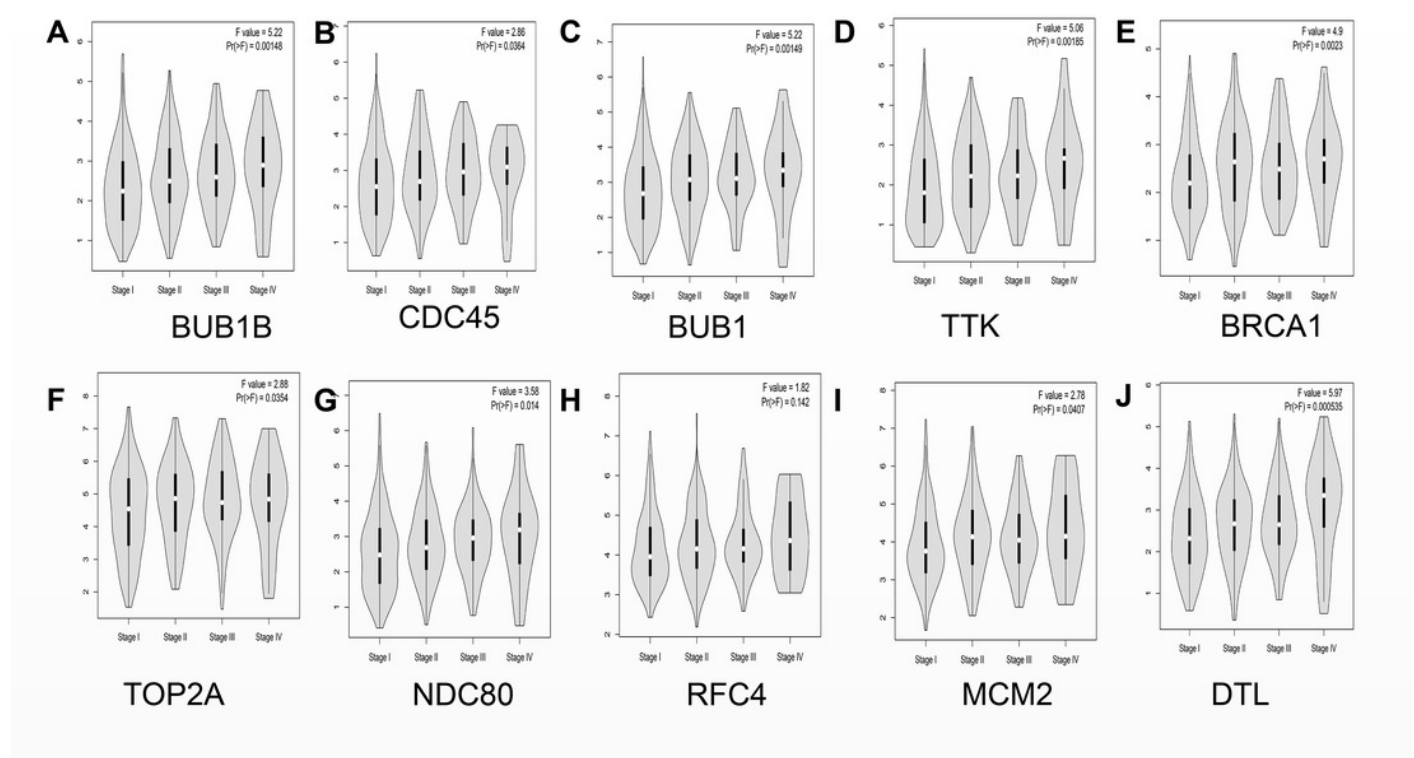



Figure 7

Survival analysis for patients with lung adenocarcinoma in relation to the expression of hub genes.

(A) BUB1B; (B) CDC45; (C) BUB1; (D) TTK; (E) BRCA1; (F) TOP2A; (G) NDC80; (H) RFC4; (I) MCM2; and (J) DTL. Red and blue represent high and low expression levels of hub genes, respectively.

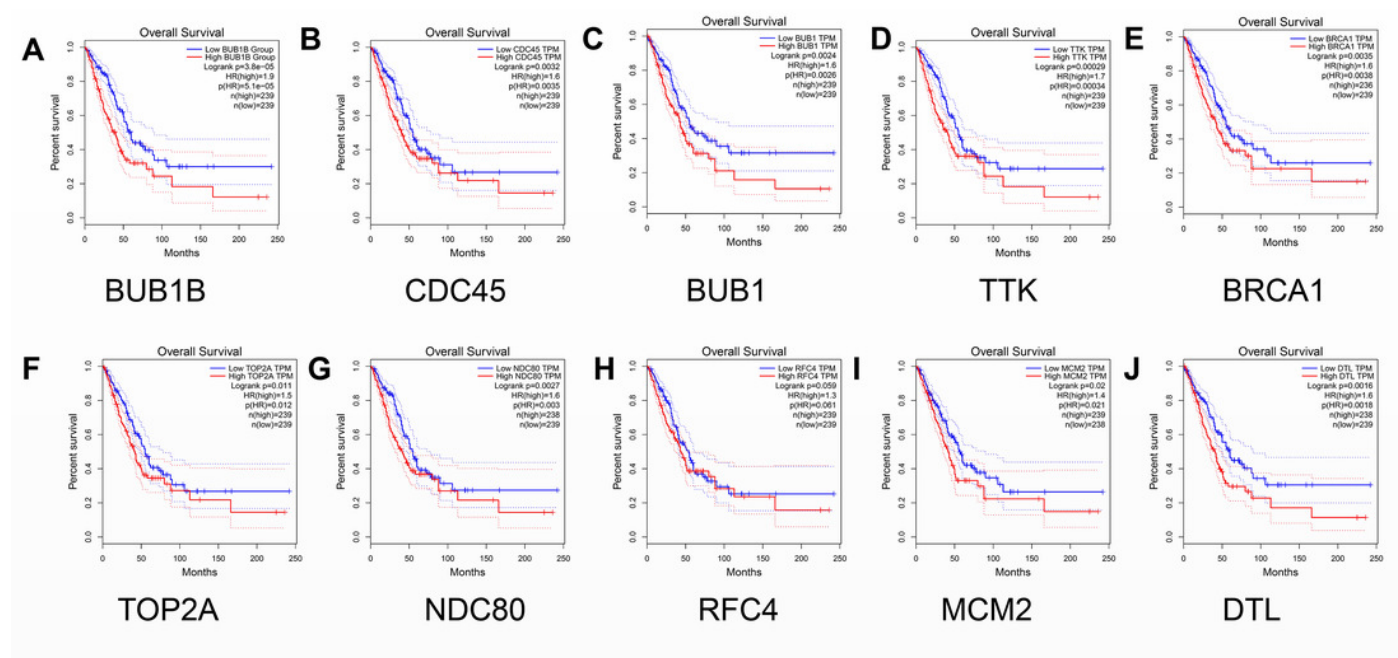



Figure 8

Analysis of mutation characteristics of hub genes.

(A) Matrix heatmap shows genomic alterations of hub genes in five lung data sets (Broad, Cell [2012]; MSKCC, Science [2015]; TCGA, Firehose Legacy; TCGA, Nature [2014]; and TCGA, PanCancer Atlas). (B) The alteration frequencies of hub genes across five studies on lung adenocarcinoma.

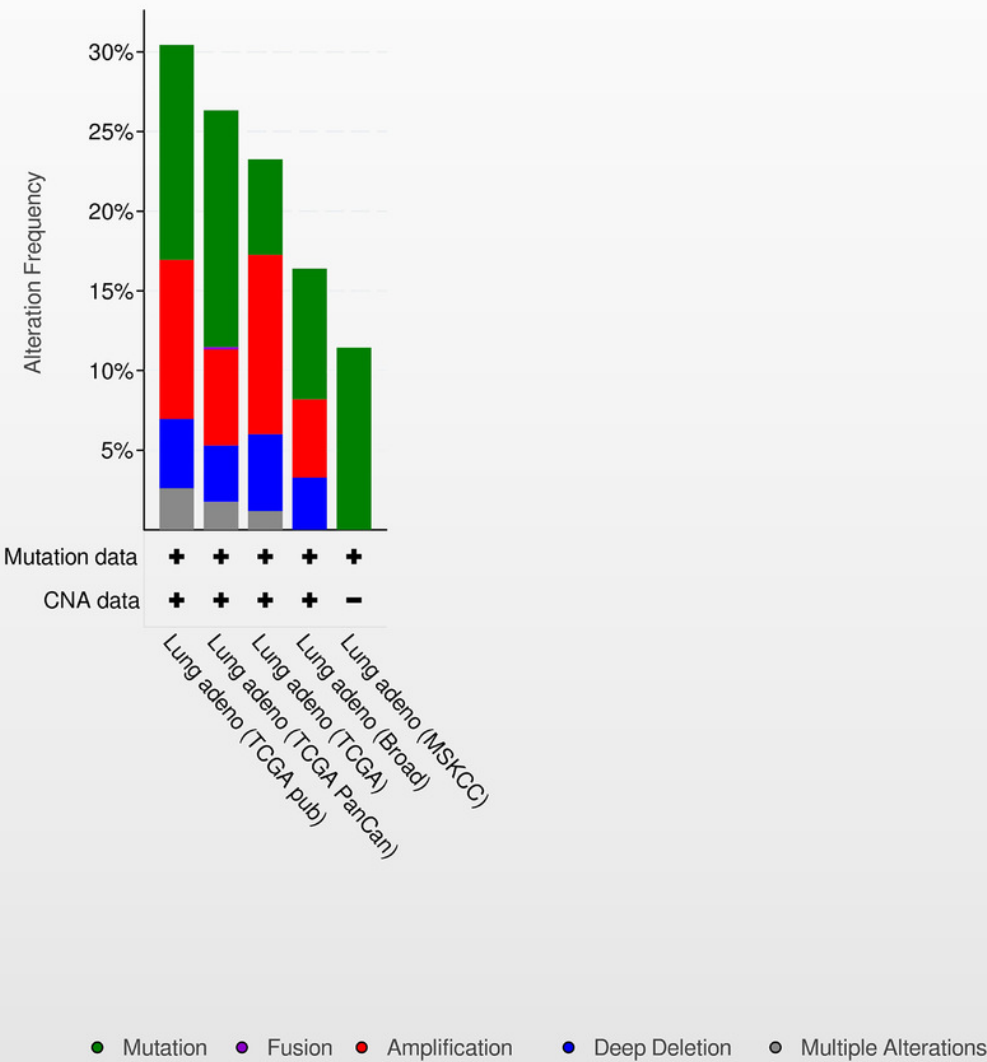
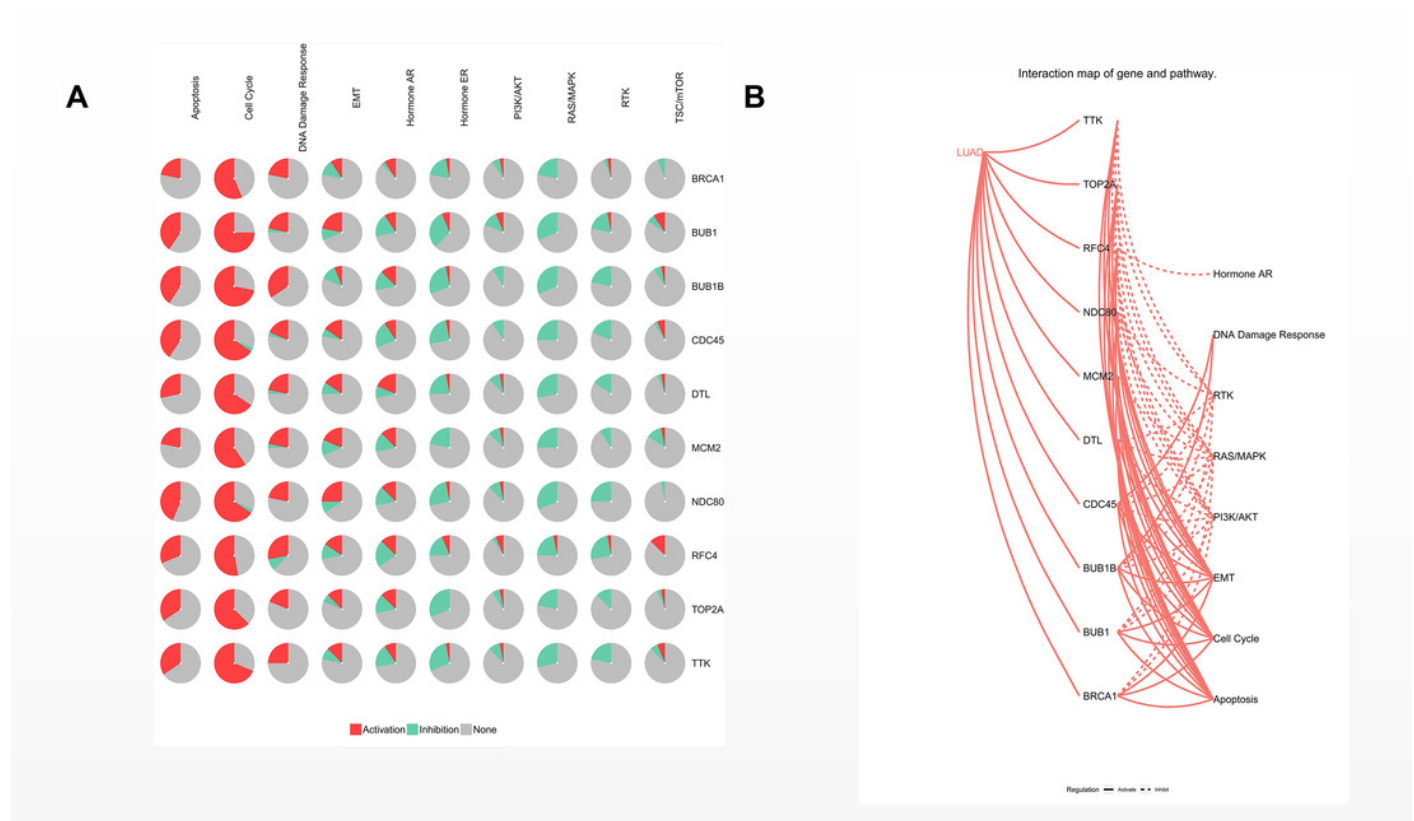



Figure 9

Analysis of pathways of Hub genes.

(A) Pie chart of hub genes involved in pathways. (B) Interaction map of hub genes and pathways in lung adenocarcinoma.

Top 10 GO Terms

1   Table 1 Top 10 GO Terms

PDL1 positive group

Category	Term	Count	%	P-Value	FDR
GOTERM_BP_FAT	cell adhesion	21	8.713692946	2.88E-04	0.469256078
GOTERM_BP_FAT	biological adhesion	21	8.713692946	2.94E-04	0.478025436
GOTERM_BP_FAT	cell-cell adhesion	12	4.979253112	5.17E-04	0.841085346
GOTERM_BP_FAT	regulation of cell motion	10	4.149377593	5.50E-04	0.894179746
GOTERM_BP_FAT	MAPKKK cascade	9	3.734439834	0.00173479	2.79437728
GOTERM_CC_FAT	plasma membrane	81	33.60995851	1.28E-08	1.64E-05
GOTERM_CC_FAT	intrinsic to plasma membrane	33	13.69294606	2.53E-05	0.032468756
GOTERM_CC_FAT	intrinsic to membrane	93	38.58921162	5.00E-05	0.064061563
GOTERM_CC_FAT	integral to plasma membrane	31	12.86307054	1.03E-04	0.132137329
GOTERM_CC_FAT	membrane raft	8	3.319502075	0.001913071	2.423429637
GOTERM_MF_FAT	cadmium ion binding	3	1.244813278	0.006756465	8.971494974
GOTERM_MF_FAT	sugar binding	8	3.319502075	0.012297355	15.76520126
GOTERM_MF_FAT	carbohydrate binding	11	4.564315353	0.01503554	18.94591653
GOTERM_MF_FAT	calcium ion binding	20	8.298755187	0.024264811	28.86459093
GOTERM_MF_FAT	lipid binding	12	4.979253112	0.02867722	33.1973471

PDL1 negative group

Category	Term	Count	%	P-Value	FDR
GOTERM_BP_DIRECT	cell division	34	8.629441624	4.29E-15	7.22E-12
GOTERM_BP_DIRECT	DNA replication	22	5.583756345	4.68E-13	7.79E-10

GOTERM_BP_DIRECT	mitotic nuclear division	26	6.598984772	2.39E-12	3.98E-09
GOTERM_BP_DIRECT	chromosome segregation	14	3.553299492	1.75E-10	0.000000291
GOTERM_BP_DIRECT	G1/S transition of mitotic cell cycle	15	3.807106599	3.32E-09	0.00000554
GOTERM_CC_DIRECT	nucleoplasm	89	22.58883249	1.26E-09	0.00000169
GOTERM_CC_DIRECT	condensed chromosome kinetochore	14	3.553299492	2.29E-09	0.00000306
GOTERM_CC_DIRECT	nucleus	136	34.5177665	6.77E-08	0.0000906
GOTERM_CC_DIRECT	chromosome, centromeric region	10	2.538071066	0.00000043	0.000576
GOTERM_CC_DIRECT	kinetochore	11	2.791878173	0.00000104	0.001390211
GOTERM_MF_DIRECT	DNA binding	52	13.19796954	0.000106	0.150583673
GOTERM_MF_DIRECT	protein binding	190	48.22335025	0.000124	0.175586276
GOTERM_MF_DIRECT	chromatin binding	19	4.822335025	0.000252	0.356042613
GOTERM_MF_DIRECT	DNA helicase activity	5	1.269035533	0.000804	1.132060211
GOTERM_MF_DIRECT	ATP binding	43	10.91370558	0.002267608	3.161395589

Top 10 KEGG pathway enrichment results

1   Table 2 Top 10 KEGG pathway enrichment results

PDL1 positive group

Category	Term	Count	%	P-Value	Genes	FDR
KEGG_PATHWAY	Arrhythmogenic right ventricular cardiomyopathy (ARVC)	7	2.904564315	0.001300189	LAMA2, ITGA9, CACNA2D1, RYR2, ITGA10, CACNA2D3, CTNNA3	1.414211029
KEGG_PATHWAY	Cell adhesion molecules (CAMs)	7	2.904564315	0.019094239	NCAM2, ITGA9, SELP, CDH15, CD22, CLDN22, HLA-DQA1	19.02707303
KEGG_PATHWAY	PPAR signaling pathway	5	2.074688797	0.024587412	LPL, SLC27A1, OLR1, FABP3, ANGPTL4	23.85529626
KEGG_PATHWAY	Complement and coagulation cascades	5	2.074688797	0.024587412	KNG1, CD55, CR2, F3, CFD	23.85529626
KEGG_PATHWAY	ECM-receptor interaction	5	2.074688797	0.045850772	LAMA2, ITGA9, ITGA10, CHAD, THBS4	40.17953232

PDL1 negative group

Category	Term	Count	%	P-Value	Genes	FDR
KEGG_PATHWAY	DNA replication	10	2.538071066	4.44E-09	RFC3, RFC4, POLD2, PCNA, POLA1, MCM2, MCM3, MCM5, CM6, RPA3	0.00000536
KEGG_PATHWAY	Cell cycle	15	3.807106599	1.25E-08	E2F2, SKP2, TTK, SMAD2, MCM2, MCM3, MCM5, MCM6, CCNE2, CDC45, CDKN2A, PCNA, BUB1, BUB1B, ORC1	0.0000151
KEGG_PATHWAY	Mismatch repair	5	1.269035533	0.000497	RFC3, RFC4, POLD2,	0.598123244

				PCNA, RPA3	
KEGG_PATHWAY	Nucleotide excision repair	6	1.52284264	0.001008028	RFC3, RFC4, POLD2, PCNA, GTF2H4, RPA3	1.210410161
KEGG_PATHWAY	p53 signaling pathway	6	1.52284264	0.004883612	CCNE2, CDKN2A, SERPINB5, RPRM, PERP, GTSE1	5.740021178

Top 10 hub genes ranked using the MNC method

1   Table 3 Top 10 hub genes ranked using the MNC method

Rank	Name	Score
1	BUB1B	81
2	CDC45	79
3	BUB1	78
3	TTK	78
5	BRCA1	75
6	TOP2A	73
7	NDC80	72
8	RFC4	71
9	MCM2	69
9	DTL	69