Mining the potential prognostic value of synaptosomal-associated protein 25 (SNAP25) in colon cancer based on stromal-immune score

Jinyan Zou [1] , **Darong Duan** [2] , **Changfa Yu** [2] , **Jie Pan** [3] , **Jinwei Xia** [2] , **Zaixing Yang** [2] , **Shasha Cai** [Corresp. 2]

[1] Department of Gastroenterology, Taizhou First People's Hospital, Huangyan Hospital of Wenzhou Medical University, Zhejiang, China, Taizhou, Zhejiang, China

[2] Department of Laboratory Medicine, Taizhou First People's Hospital, Huangyan Hospital of Wenzhou Medical University, Zhejiang, China, Taizhou, Zhejiang, China

[3] Outpatient Department, Taizhou First People's Hospital, Huangyan Hospital of Wenzhou Medical University, Zhejiang, China, Taizhou, Zhejiang, China

Corresponding Author: Shasha Cai
Email address: tyycss@163.com

Background: Colon cancer is one of the deadliest tumors worldwide. Stromal cells and immune cells play important roles in cancer biology and microenvironment across different types of cancer. This study aimed to identify the prognostic value of stromal/immune cell-associated genes for colon cancer in The Cancer Genome Atlas (TCGA) database using bioinformatic technology. **Methods:** The gene expression data and corresponding clinical information of colon cancer were downloaded from TCGA database. Stromal and immune scores were estimated based on the ESTIMATE algorithm. Sanger software was used to identify the differentially expressed genes (DEGs) and prognostic DEGs based on stromal and immune scores. External validation of prognostic biomarkers was conducted in Gene Expression Omnibus (GEO) database. Gene ontology (GO) analysis, pathway enrichment analysis, and gene set enrichment analysis (GSEA) were used for functional analysis. STRING and Cytoscape were used to assess Protein-protein interaction (PPI) network and screen hub genes. Quantitative real-time PCR (qRT-PCR) was used to validate the expression of hub genes in clinical tissues. Synaptosomal-associated protein 25 (SNAP25) was selected for analyzing its correlations with tumor-immune system in TISIDB database. **Results:** Worse overall survivals of colon cancer patients were found in high stromal score group (2963 vs. 1930 days, log-rank test P=0.038) and high immune score group (2894 vs. 2230 days, log-rank test P=0.076). 563 up-regulated and 9 down-regulated genes were identified as stromal-immune score-related DEGs. 70 up-regulated DEGs associated with poor outcomes were identified by COX proportional hazard regression model, and 15 hub genes were selected later. Then, we verified aquaporin 4 (AQP4) and SNAP25 as prognostic biomarkers in GEO database. qRT-PCR results revealed that AQP4 and SNAP25 were significantly elevated in colon cancer tissues compared with adjacent normal tissues

(P=0.003, 0.001). GSEA and TISIDB suggested that SNAP25 involved in cancer-related signaling pathway, immunity and metabolism progresses. **Conclusion:** SNAP25 is a microenvironment-related and immune-related gene that can predict poor outcomes in colon cancer.

1 **Mining the potential prognostic value of synaptosomal-associated protein 25**

2 **(SNAP25) in colon cancer based on stromal-immune score**

3 Jinyan Zou[1], Darong Duan[2], Changfa Yu[2], Jie Pan[3], Jinwei Xia[2], Zaixing Yang[2], Shasha Cai[2]

4 [1]Department of Gastroenterology, Taizhou First People's Hospital, Huangyan Hospital of

5 Wenzhou Medical University, Zhejiang, China

6 [2]Department of Laboratory Medicine, Taizhou First People's Hospital, Huangyan Hospital of

7 Wenzhou Medical University, Zhejiang, China

8 [3]Outpatient Department, Taizhou First People's Hospital, Huangyan Hospital of Wenzhou

9 Medical University, Zhejiang, China.

10 Corresponding Author:

11 Zaixing Yang[2]

12 No 218 Hengjie Road, Huangyan District, Taizhou City, Zhejiang Province, 318020, China

13 E-mail: yangzaixingdiyi@163.com

14 Shasha Cai[2]

15 No 218 Hengjie Road, Huangyan District, Taizhou City, Zhejiang Province, 318020, China

16 E-mail: tyycss@163.com

18 # Abstract

19 **Background:** Colon cancer is one of the deadliest tumors worldwide. Stromal cells and immune

20 cells play important roles in cancer biology and microenvironment across different types of

21 cancer. This study aimed to identify the prognostic value of stromal/immune cell-associated

22 genes for colon cancer in The Cancer Genome Atlas (TCGA) database using bioinformatic

23 technology.

24 **Methods:** The gene expression data and corresponding clinical information of colon cancer were

25 downloaded from TCGA database. Stromal and immune scores were estimated based on the

26 ESTIMATE algorithm. Sanger software was used to identify the differentially expressed genes

27 (DEGs) and prognostic DEGs based on stromal and immune scores. External validation of

28 prognostic biomarkers was conducted in Gene Expression Omnibus (GEO) database. Gene

29 ontology (GO) analysis, pathway enrichment analysis, and gene set enrichment analysis (GSEA)

30 were used for functional analysis. STRING and Cytoscape were used to assess Protein-protein

31 interaction (PPI) network and screen hub genes. Quantitative real-time PCR (qRT-PCR) was

32 used to validate the expression of hub genes in clinical tissues. Synaptosomal-associated protein

33 25 (SNAP25) was selected for analyzing its correlations with tumor-immune system in TISIDB

34 database.

35   **Results:** Worse overall survivals of colon cancer patients were found in high stromal score group

36   (2963 vs. 1930 days, log-rank test P=0.038) and high immune score group (2894 vs. 2230 days,

37   log-rank test P=0.076). 563 up-regulated and 9 down-regulated genes were identified as stromal-

38   immune score-related DEGs. 70 up-regulated DEGs associated with poor outcomes were

39   identified by COX proportional hazard regression model, and 15 hub genes were selected later.

40   Then, we verified aquaporin 4 (AQP4) and SNAP25 as prognostic biomarkers in GEO database.

41   qRT-PCR results revealed that AQP4 and SNAP25 were significantly elevated in colon cancer

42   tissues compared with adjacent normal tissues (P=0.003, 0.001). GSEA and TISIDB suggested

43   that SNAP25 involved in cancer-related signaling pathway, immunity and metabolism

44   progresses.

45   **Conclusion:** SNAP25 is a microenvironment-related and immune-related gene that can predict

46   poor outcomes in colon cancer.

47   **Keywords:** colon cancer, stromal, immune, TCGA, GEO, prognosis

48   Introduction

49   Colon cancer is the third most common malignant tumor and one of the leading causes of

50   cancer-related death worldwide. Since risk factors have been investigated, earlier detection,

51   cancer prevention, surgical techniques, radiotherapy and chemotherapy treatment have been

52   improved, the incidence and mortality of colon cancer has been slowly declined. A model-based

53   estimate showed that 104610 new cases of colon cancer would be diagnosed and 53200 patients

54   would die of colon cancer in United States (Siegel et al. 2020).However, the incidence of

55   colorectal cancer in patients aged <50 years increased by 22% from 2000-2013(Siegel et al.

56   2017). Thus, more attention should be focused on early diagnosis through screening and

57   accurately predicting the survival outcome of patients with colon cancer.

58   Stromal cells and immune cells form the major fraction of colon cancer tissue and are

59   associated with tumor progress, inflammatory and metabolic disorders(Ghesquière et al. 2014;

60   Nilendu et al. 2018). An increasing amount of studies have highlighted the importance of stromal

61   cells and immune cells in cancer biology and microenvironment across different types of

62   cancers(Barros et al. 2018; Garcia-Gomez et al. 2018; Zhan et al. 2017). ESTIMATE is a new

63   method that infers stromal and immune cells in malignant tumors using gene expression

64   signatures(Yoshihara et al. 2013), and has been conducted in acute myeloid leukemia(Yan et al.

65   2019), gastric cancer(Wang et al. 2019), and glioblastoma(Jia et al. 2018).

66   In the current study, we obtained immune and stromal scores of colon cancer based on

67   ESTIMATE. To help elucidate the stromal-immune score-based genes with prognostic value in

68   colon cancer, we obtained gene expression dates from The Cancer Genome Atlas (TCGA), and

69   verified the survival value in a different colon cancer cohort available from the Gene Expression

70   Omnibus (GEO) database. Two hub genes were validated to be prognosis biomarkers and

71   selected for further analysis. We investigated the potential underlying mechanisms of

72   synaptosomal-associated protein 25 (SNAP25) in cancer-related signaling pathway, immunity

73   and metabolism progresses through gene set enrichment analysis (GSEA) and TISIDB database.

74

Materials & Methods

Database and estimation of stromal and immune scores

77   The TCGA level 3 gene expression data and corresponding clinical information for colon

78   cancer patients were obtained from Genomic Date Commons of the National Cancer Institute

79   (http://portal.gdc.cancer.gov). Only patients with gene expression data, follow up information

80   and clinicopathologic information were included in this study. For normalization, the RNA-seq

81   data of all patients was transformed to transcripts per million (TPM) values

82   (https://pubmed.ncbi.nlm.nih.gov/30379987/). The stromal and immune scores for each TCGA

83   sample were conducted by R 3.6.2 using the R package "estimate"

84   (https://pubmed.ncbi.nlm.nih.gov/24113773/).

85   We obtained gene expression profiles and clinical information of 430 colon cancer patients

86   from TCGA database. Among them, 231 (53.7%) cases were male and 199 (46.3%) cases were

87   female. The average age of patients at initial pathological diagnosis was 66.3 years (range: 31-90

88   years). Histologic diagnosis included 369 (85.5%) cases of colon adenocarcinoma and 57

89   (13.3%) cases of colon mucinous adenocarcinoma, 4 (0.9%) cases were not classified. The tumor

90   stage was stage I in 17.4%, stage II in 38.6%, stage III in 29.1, stage IV in 14.4% of cases, and 2

91   cases (0.5%) were of unknown stage. The tumors were located in the left (40%) or right (55.3%)

92   colon according to their anatomic neoplasm subdivision, with 4.7% unknown. Based on

93   ESTIMATE algorithm, we obtained stromal scores (range: -2262.07~1999.52) and immune

94   scores (range: -954.97~3035.59) for all these colon cancer patients.

Correlations between clinicopathologic data and stromal/immune score

96   The correlations between clinicopathologic data and stromal/immune score were analyzed

97   by SPSS 22.0 software (SPSS, Inc., Chicago, IL, USA). Patients with colon cancer were divided

98   into high stromal/immune score (the fourth quartile) and low stromal/immune score groups

99   (quartile 1-3). The stromal/immune score of different clinicopathologic groups was compared by

100   Mann-Whitney U test, and overall survival was estimated by the Kaplan-Meier method and

101   compared by log-rank tests. A value of $P<0.05$ was considered statistically significant.

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

103   DEGs were identified based on immune and stromal scores (high stromal/immune score

104   group vs. low stromal/immune score group) by Sanger_V1.0.8 software (https://shengxin.ren/).

105   Genes with log2 (fold change) >1.5 or <-1 combined with a P value <0.01 were defined as

106   DEGs. The volcano plot of the DEGs was drawn by Sanger_V1.0.8 software, and the venn

107   diagram was drawn on a website Venny 2.1.0

108   (https://bioinfogp.cnb.csic.es/tools/venny/index.html).

109   **Gene ontology (GO) and pathway enrichment analyses**

110   Cellular component (CC), molecular function (MF), biological process (BP) and pathway

111   enrichment analyses were conducted using FunRich 3.1.3

112   (https://pubmed.ncbi.nlm.nih.gov/25921073/). A value of $P<0.05$ was considered as the

113   screening condition.

114   **Survival analysis**

115   A COX proportional hazards model was applied to illuminate prognostic DEGs of colon

116   cancer obtained from TCGA and a P value <0.01 was considered significant. An open source

117   web tool PrognoScan (https://pubmed.ncbi.nlm.nih.gov/19393097/) was conducted to verify the

118   survival outcomes between prognostic DEGs identified and colon cancer patients from GEO

119   database.

120   **Protein-protein interaction (PPI) network construction**

121   The PPI network for DEG-encoded proteins was performed by STRING database

122   (https://pubmed.ncbi.nlm.nih.gov/30476243/) and reconstructed by Cytoscape 3.7.2

123   (https://pubmed.ncbi.nlm.nih.gov/14597658/). The most significant modular analysis was

124   identified by Molecular Complex Detection (MCODE) plugin of Cytoscape, and the plug-in

125   Biological Networks Gene Oncology (BiNGO) of Cytoscape was applied to analysis GO term of

126   hub genes.

127 **Heatmap and clustering analysis**

128   Heatmap and clustering analysis were completed by "heatmap" package.

129 **Quantitative real-time PCR (qRT-PCR)**

130   Total RNA was extracted with Trizol reagent (TaKaRa Bio Inc. Shiga, Japan) from colon

131 cancer and adjacent normal tissues. cDNA was synthesized with PrimeScript™ RT reagent kit

132 (TaKaRa, RR036A). Quantitative real-time PCR (qRT-PCR) was carried out using the TB

133 Green™ Premix Ex Taq™ kit (TaKaRa, RR420A) on ABI step one Real-Time PCR system. The

134 primers were as follows: AQP4, sense strand 5'- GAGCAGGAATCCTCTATC-3', antisense

135 strand 5'- AGTGACATCAGTCCGTTT-3'; SNAP25, sense strand 5'-

136 GTAGTGGACGAACGGGAGC-3', antisense strand 5'- CCTGTCGATCTGGCGATT-3'; GAPDH,

137 sense strand 5'-GTCAACGGATTTGGTCTGTATT-3', antisense strand 5'-

138 AGTCTTCTGGGTGGCAGTGAT-3'.

139   The Institutional Medical Ethics Review Board of Taizhou first people's hospital in

140 Zhejiang Province approval to carry out the study within its facilities (Ethical Application Ref:

141 2019-KY009-03).

142 **Gene set enrichment analysis (GSEA)**

143   According to the expression level of AQP4 (or SNAP25), samples of the complete cohort

144 from TCGA were divided into 2 groups, and implemented using GSEA by Sanger_V1.0.8

145 software. The KEGG gene set biological process database (c2.cp.kegg.v6) were chosen for

146 enrichment analysis. Terms with both P value <0.01 and false discovery rate (FDR) <0.01 were

147 identified.

148 **Mining the immune-related mechanism of SNAP25.**

149   TISIDB (https://pubmed.ncbi.nlm.nih.gov/14597658/) is a friendly web portal integrates

150 988 immune-related anti-tumor genes from 7 databases(Ru et al. 2019). The correlations between

151 immune features and any gene may be explored in 30 TCGA cancer types. In this study, the

152 TISIDB database was used to investigate the associations between expression (or methylation) of

153 SNAP25 and tumor-infiltrating lymphocytes (TILs). However, the relations between TILs and

154 expression (or methylation) of AQP4 in colon cancer were not integrated in TISIDB database. A

155 value of P<0.05 was considered statistically significant.

156

157 **Results**

158   The relationships between stromal/immune score and clinical features

159   The stromal and immune scores were variously distributed between adenocarcinoma and

160   mucinous adenocarcinoma (Figure 1A, C). Both stromal score and immune score of colon

161   adenocarcinoma cases were significantly lower than those of mucinous adenocarcinoma cases

162   (P=0.001, 0.011, respectively). In addition, right colon tumors yielded higher immune scores

163   than those left colon cases (Figure 1D, P<0.001), though no significant differences between left

164   and right colon were found for the stromal scores (Figure 1B, P=0.818).

165   The potential association of overall survival and stromal/immune score was explored by

166   classifying 430 colon cancer patients into high and low score groups based on their stromal or

167   immune scores. As shown in Figure 1E, F, the median overall survival of patients with a low

168   stromal score was longer than those in high score group (2963 vs. 1930 days, log-rank test

169   P=0.038); consistently, the median overall survival of patients with a low immune score was

170   longer than those in high score group (2894 vs. 2230 days, log-rank test P=0.076), although there

171   was no statistically significant difference. However, patients with both a high stromal score and a

172   high immune score were found to have significantly worse survival than those with low scores

173   (1891 vs. 2974 days, log-rank test P=0.039) (Figure 1G).

174   Identification of DEGs by stromal and immune scores in colon cancer

175   After standardization of the RNA-Seq data for all 430 colon cancer patients obtained from

176   TCGA database, we identified 4881 and 1512 DEGs based on stromal and immune scores,

177   respectively. As shown in the volcano plots of DEGs for stromal/immune score (Figure 2A, B),

178   4791 genes were up-regulated and 90 genes were down-regulated for the comparison based on

179   stromal score. Similarly, 1113 genes were up-regulated and 399 genes were down-regulated

180   based on high immune score group vs. low immune score group. Through Venn diagram (Figure

181   2C, D) analysis, 563 shared up-regulated DEGs and 9 shared down-regulated DEGs from

182   stromal score and immune score groups were identified and selected for subsequent analysis.

183   Functional and pathway enrichment analyses

184   Go and pathway enrichment analyses of the above 572 genes were performed (Figure 2E, F,

185   G, H). For CC, DEGs were mainly associated with plasma membrane, extracellular and

186   extracellular space. With regard to MF, genes were mainly clustered in receptor activity, cell

187   adhesion molecule activity and B cell receptor activity. DEGs in the BP category primarily

188   enriched in immune response, cell communication and signal transduction. The pathway

189   enrichment analysis showed genes were mainly enriched in epithelial-to-mesenchymal transition,

190   peptide ligand-binding receptors and GPCR ligand binding.

191   **Identification of prognostic DEGs in colon caner**

192   The COX proportional hazard regression model was constructed to identify potential

193   prognostic DEGs in colon cancer. Among the 563 shared up-regulated DEGs and 9 shared down-

194   regulated DEGs, 70 up-regulated DEGs associated with poor outcomes were shown in Figure 3.

195   **PPI network construction and hub gene analysis among prognostic DEGs**

196   To further explore the interplay among the 70 identified prognostic DEGs, we conducted a

197   PPI network containing 30 nodes and 53 edges based on STRING tool and Cytoscape software

198   (Figure 4A). Module analysis using MCODE was constructed, and 15 hub genes were selected

199   (Figure 4A). CC, MF, BP analyses of the total 15 hub genes were performed using BiNGO

200   (Figure 4B). The 15 hub genes were mainly associated with plasma membrane, transporter

201   activity, secretion and channel activity.

202   **Heatmap and clustering analysis of 15 hub genes**

203   The expression level of 15 hub genes in "dead" and "alive" groups was shown in Figure 4C.

204   **Verifying the survival outcomes of hub genes in the GEO database**

205   The survival outcomes of 15 hub genes were identified on the PrognoScan web tool, which

206   provided overall survival of GSE12945, GSE17536 and GSE17537 datasets for colorectal

207   cancer. Then, AQP4 and SNAP25 were verified (Figure 5) to be significantly associated with

208   overall survival according to both the log-rank test and COX proportional hazards regression

209   analysis (all P <0.05).

210   **The expression of two hub genes in 20 colon cancer and adjacent normal tissue sample set**

211   **using qRT-PCR.**

212   The expression of AQP4 and SNAP25 was validated in 20 pairs of clinical tissues using

213   qRT-PCR. Interestingly, the mRNA relative expression levels of both AQP4 and SNAP25 were

214   significantly elevated in colon cancer tissues compared with adjacent normal tissues (P=0.003,

215   0.001) (Figure 6).

216   **GSEA using TCGA database**

217   To further investigate the underlying mechanism of AQP4 and SNAP25 in colon cancer,

218   KEGG pathway enrichment analysis was performed by GSEA. For AQP4, "vascular smooth

219   muscle contraction" (NES=2.17, P<0.001, FDR=0.006) gene set was prominently enriched. For

220 SNAP25, 57 gene sets were enriched, including 11 gene sets which were cancer-related

221 processes (Figure 7). Besides, high expression of SNAP25 might also be involved in "B cell

222 receptor signaling pathway" (NES=2.06, P=0.002, FDR=0.002), "cell adhesion molecules cams"

223 (NES=2.05, P=0.008, FDR=0.002), "chemokine signaling pathway" (NES=2.07, P=0.002,

224 FDR=0.002), "complement and coagulation cascades" (NES=2.08, P<0.001, FDR=0.002), "T

225 cell receptor signaling pathway" (NES=2.06, P<0.001, FDR=0.002), "adipocytokine signaling

226 pathway" (NES=1.96, P<0.001, FDR=0.005), "aldosterone regulated sodium reabsorption"

227 (NES=2.02, P<0.001, FDR=0.003), "glycosaminoglycan biosynthesis heparan sulfate"

228 (NES=2.00, P<0.001, FDR=0.004), and "insulin signaling pathway" (NES=1.94, P=0.002,

229 FDR=0.007). This suggested that immunity and metabolism may be as well involved in the

230 underlying mechanism of SNAP25 in colon cancer.

231 **Regulation of immune molecules by SNAP25**

232 The spearman's correlations between lymphocytes and expression, methylation of SNAP25

233 were performed in TISIDB database (Figure 8). The associations between the expression of

234 SNAP25 and immune-related signatures of TILs types were shown in Figure 8A, and the greatest

235 correlations including natural killer cell (NK; r=0.493, P<2.2e-16), macrophage (r=0.45, P<2.2e-

236 16), mast cell (r=0.448, P<2.2e-16), and natural killer T cell (NKT; r=0.447, P<2.2e-16) were

237 shown in Figure 8B. The correlations between methylation of SNAP25 and lymphocytes were

238 shown in Figure 8C, and Figure 8D displayed the remarkable negative correlations including

239 plasmacytoid dendritic cell (pDC; r=-0.419, P=2.96e-14), type 1 T helper cell (Th1; r=-0.406,

240 P=4.06e-13), T follicular helper cell (Tfh; r=-0.391, P=3.78e-12), and NKT (r=-0.391, P=3.61e-

241 12). Therefore, the potential underlying mechanism of SNAP25 in colon cancer may be involved

242 in the regulation of the above TILs.

243 # Discussion

244 Solid tumor tissues comprise not only malignant cells but also tumor microenvironment,

245 including immune cells, stromal cells, epithelial cells, fibroblasts, vascular cells and signaling

246 molecules. Accumulating evidence clarifies that tumor microenvironment plays a crucial role in

247 tumor growth, progression, metastasis, prognosis, and treatment(Petitprez et al. 2018; Wu & Dai

248 2017). In the current study, we focused on stromal and immune scores, which reflect the

249 microenvironment of tumor and hence contribute to survival prediction in colon cancer.

250 Meanwhile, our results were in accordance with previous specific insights(Koi & Carethers

251 2017; Zhang et al. 2018), and might provide extra data in the mining of interaction between

252 tumor and environment in colon cancer.

253   Then, we screened out 572 microenvironment-related DEGs, and found they were mainly

254 enriched in plasma membrane (CC), receptor activity (MF), immune response (BP), and

255 epithelial-to-mesenchymal transition (pathway). Afterwards, AQP4, ASTN1, ATP2B3, CADM3,

256 CD22, CD37, CD79B, CLVS2, CXCR5, FCRLA, GABRG2, MS4A1, NRXN1, SNAP25, and

257 SYT4 were identified as prognostic hub genes, and two of them were verified to be prognosis

258 biomarkers in GEO database. qRT-PCR results revealed that AQP4 and SNAP25 were

259 significantly elevated in colon cancer tissues compared with adjacent normal tissues (P=0.003,

260 0.001). Next, we investigated the underlying mechanism of AQP4 and SNAP25 in colon cancer

261 by GSEA, and found that the high expression of SNAP25 might be involved in cancer-related

262 signaling pathway, immunity and metabolism processes. Further researches in TISIDB database

263 indicated greatest positive correlations between SNAP25 expression and TILs (NK, NKT,

264 macrophage, mast), and greatest negative correlations between SNAP25 methylation and TILs

265 (pDC, Th1, Tfh, NKT). TILs are associated with prognosis for the survival of various tumors in

266 many previous studies. TILs are also reported as a predictive biomarker in colon cancer(Zhao et

267 al. 2019). However, some studies found no significant association between TILs and overall

268 survival. It remains controversial on the prognostic value of TILs in colon cancer may be due to

269 different TIL responses or subsets, different patient populations, different study designs, and the

270 heterogeneity of immune infiltrate. Further studies should be designed to provide more insight

271 into the study of TILs and survival.

272   SNAP25 belongs to the soluble N-ethylmaleimide-sensitive factor attachment protein

273 receptor (SNARE) complex, which is essential for neurotransmitter release, synaptic, secretory

274 vesicle exocytosis, minimal fusion machinery, cell-to-cell signaling, and the regulation of ion

275 channels(Baker & Hughson 2016; Wang et al. 2017; Yoon & Munson 2018). As reported,

276 SNAP25 is potentially important for normal vesicle fusion and lysosomal trafficking(Manca et

277 al. 2014; Mu et al. 2018). Kyoko Kobayashi et al(Kobayashi et al. 2016) found that SNAP25

278 protein was expressed in 46% (77/168) of diffuse large B-cell lymphoma (DLBCL) patients and

279 associated with CD5 expression (P=0.018). Huang CJ et al(Huang et al. 2017) displayed the

280 expression and functional significance of SNAP25 in medulloblastoma.

281   Although the clinical significances and functions of SNAP25 for colon cancer have not

282   been previously reported, it may serve as prognosis biomarker according to this study. To our

283   greatest interest, SNAP25 is the most highly interconnected nodes (Figure 4A), and involves in

284   many cancer-related processes consists of adherens junction, calcium signaling pathway,

285   cytokine receptor interaction, ECM receptor interaction, hedgehog signaling pathway, JAK-

286   STAT signaling pathway, MAPK signaling pathway, TGF beta signaling pathway, Toll-like

287   receptor signaling pathway, VEGF signaling pathway, and WNT signaling pathway (Figure 7).

288   Besides, SNAP25 is also involved in immunity and metabolism processes such as B cell receptor

289   signaling pathway, cell adhesion molecules cams, chemokine signaling pathway, complement

290   and coagulation cascades, T cell receptor signaling pathway, adipocytokine signaling pathway,

291   aldosterone regulated sodium reabsorption, glycosaminoglycan biosynthesis heparan sulfate and

292   insulin signaling pathway. Those may bring novel insights into the potential underlying

293   mechanism of SNAP25 in colon cancer. In addition, this paper reveals the significant

294   correlations between SNAP25 and lymphocytes (NK, macrophage, mast cell, NKT), which

295   indicates the potential association of tumor microenvironment and SNAP25. Recently, a lot of

296   attention has been paid to tumor microenvironment and immune evasion for future diagnoses and

297   treatments of malignant tumors. Immunotherapies have been revolutionizing tumor treatment,

298   although immunological response in different patients is heterogeneous. NK cells participate in

299   tumor immunosurveillance, and are one of the most promising therapies for various types of

300   cancer. However, NK cell populations may shape with altered reactivity in malignant

301   tumors(Hofer & Koehl 2017). In colorectal cancer, a high level of mast cells was confirmed with

302   poor survival, and the density of innate immune cells (macrophages, mast cells, neutrophils, and

303   immature dendritic cells) increased with tumor stage(Koi & Carethers 2017). Therefore,

304   SNAP25 is extremely closely associated with various types of TILs, and has the potential to

305   serve as a prognosis biomarker and an immunotherapeutic target for colon cancer.

306   However, our study presents a number of limitations. One major limitation is that the

307   present research was mainly based on previous data from TCGA and GEO, therefore, future

308   investigations in vivo and vitro are needed to investigate the effect of SNAP25 in colon cancer.

309   A second issue is that since the integrated bioinformatics analysis was focus on immune

310   microenvironment for colon cancer, and the DEGs were identified based on immune and stromal

311   scores, there was no validation for these DEGs based on control samples. Finally, we are lacking

312 of the specimens of colon cancer and adjacent normal tissues. So, we only validate the

313 expressions of AQP4 and SNAP25 in 20 pairs of tissues by qRT-PCR.

Conclusions

315 In summary, SNAP25 is a microenvironment-related and immune-related gene that can

316 predict poor outcomes in colon cancer. Bioinformatic analysis suggests that SNAP25 is involved

317 in cancer-related signaling pathway, immunity and metabolism processes, which may provide a

318 new target for investigating the underling mechanism of colon cancer.

319

Acknowledgements

321 This work was supported by one grant from Health and Family planning Commission of

322 Zhejiang Province (2020KY1045) and one grant from Taizhou Science and Technology Bureau

323 in Zhejiang Province (1901ky62). And the authors would like to thank Qingqing Xia (Central

324 laboratory of Taizhou First People's Hospital) for her technical assistance in qRT-PCR.

325

References

327 Baker RW, and Hughson FM. 2016. Chaperoning SNARE assembly and disassembly. *Nat Rev Mol Cell Biol*
328   17:465-479.
329 Barros MR, Jr., de Melo CML, Barros MLCMGR, de Cássia Pereira de Lima R, de Freitas AC, and Venuti A. 2018.
330   Activities of stromal and immune cells in HPV-related cancers. *Journal of experimental & clinical cancer*
331   *research : CR* 37:137-137.
332 Garcia-Gomez A, Rodríguez-Ubreva J, and Ballestar E. 2018. Epigenetic interplay between immune, stromal and
333   cancer cells in the tumor microenvironment. *Clinical immunology (Orlando, Fla)* 196:64-71.
334 Ghesquière B, Wong BW, Kuchnio A, and Carmeliet P. 2014. Metabolism of stromal and immune cells in health
335   and disease. *Nature* 511:167-176.
336 Hofer E, and Koehl U. 2017. Natural Killer Cell-Based Cancer Immunotherapies: From Immune Evasion to
337   Promising Targeted Cellular Therapies. *Frontiers in immunology* 8:745-745.
338 Huang CJ, Lee CL, Liu CY, Huang SH, Hou JW, Chen YH, Chien CC, Ho CM, Lo WC, and Hung KL. 2017.
339   Detection of lower levels of SNAP25 using multiple microarray systems and its functional significance in
340   medulloblastoma. *Int J Mol Med* 39:1195-1205.
341 Jia D, Li S, Li D, Xue H, Yang D, and Liu Y. 2018. Mining TCGA database for genes of prognostic value in
342   glioblastoma microenvironment. *Aging* 10:592-605.
343 Kobayashi K, Yamaguchi M, Miyazaki K, Imai H, Yokoe K, Ono R, Nosaka T, and Katayama N. 2016. Expressions
344   of SH3BP5, LMO3, and SNAP25 in diffuse large B-cell lymphoma cells and their association with clinical
345   features. *Cancer Med* 5:1802-1809.
346 Koi M, and Carethers JM. 2017. The colorectal cancer immune microenvironment and approach to immunotherapies.
347   *Future oncology (London, England)* 13:1633-1647.
348 Manca P, Mameli O, Caria MA, Torrejon-Escribano B, and Blasi J. 2014. Distribution of SNAP25, VAMP1 and
349   VAMP2 in mature and developing deep cerebellar nuclei after estrogen administration. *Neuroscience*
350   266:102-115.
351 Mu Y, Yan X, Li D, Zhao D, Wang L, Wang X, Gao D, Yang J, Zhang H, Li Y, Sun Y, Wei Y, Zhang Z, Chang X,
352   Yao Z, Tian S, Zhang K, Terada LS, Ma Z, and Liu Z. 2018. NUPR1 maintains autolysosomal efflux by
353   activating SNAP25 transcription in cancer cells. *Autophagy* 14:654-670.
354 Nilendu P, Sarode SC, Jahagirdar D, Tandon I, Patil S, Sarode GS, Pal JK, and Sharma NK. 2018. Mutual
355   concessions and compromises between stromal cells and cancer cells: driving tumor development and drug
356   resistance. *Cellular oncology (Dordrecht)* 41:353-367.

357 Petitprez F, Vano YA, Becht E, Giraldo NA, de Reyniès A, Sautès-Fridman C, and Fridman WH. 2018.
358   Transcriptomic analysis of the tumor microenvironment to guide prognosis and immunotherapies. *Cancer*
359   *Immunology, Immunotherapy* 67:981-988.
360 Ru B, Wong CN, Tong Y, Zhong JY, Zhong SSW, Wu WC, Chu KC, Wong CY, Lau CY, Chen I, Chan NW, and
361   Zhang JW. 2019. TISIDB: an integrated repository portal for tumor–immune system interactions.
362   *Bioinformatics* 35:4200-4202.
363 Siegel RL, Miller KD, Fedewa SA, Ahnen DJ, Meester RGS, Barzi A, and Jemal A. 2017. Colorectal cancer
364   statistics, 2017. *CA Cancer J Clin* 67:177-193.
365 Siegel RL, Miller KD, and Jemal A. 2020. Cancer statistics, 2020. *CA Cancer J Clin* 70:7-30.
366 Wang H, Wu X, and Chen Y. 2019. Stromal-Immune Score-Based Gene Signature: A Prognosis Stratification Tool
367   in Gastric Cancer. *Frontiers in Oncology* 9.
368 Wang T, Li L, and Hong W. 2017. SNARE proteins in membrane trafficking. *Traffic* 18:767-775.
369 Wu T, and Dai Y. 2017. Tumor microenvironment and therapeutic response. *Cancer letters* 387:61-68.
370 Yan H, Qu J, Cao W, Liu Y, Zheng G, Zhang E, and Cai Z. 2019. Identification of prognostic genes in the acute
371   myeloid leukemia immune microenvironment based on TCGA data analysis. *Cancer Immunol Immunother*
372   68:1971-1978.
373 Yoon TY, and Munson M. 2018. SNARE complex assembly and disassembly. *Curr Biol* 28:R397-R401.
374 Yoshihara K, Shahmoradgoli M, Martinez E, Vegesna R, Kim H, Torres-Garcia W, Trevino V, Shen H, Laird PW,
375   Levine DA, Carter SL, Getz G, Stemke-Hale K, Mills GB, and Verhaak RG. 2013. Inferring tumour purity
376   and stromal and immune cell admixture from expression data. Nat Commun 4:2612.
377 Zhan H-X, Zhou B, Cheng Y-G, Xu J-W, Wang L, Zhang G-Y, and Hu S-Y. 2017. Crosstalk between stromal cells
378   and cancer cells in pancreatic cancer: New insights into stromal biology. *Cancer letters* 392:83-93.
379 Zhang M, Pang H-J, Zhao W, Li Y-F, Yan L-X, Dong Z-Y, and He X-F. 2018. VISTA expression associated with
380   CD8 confers a favorable immune microenvironment and better overall survival in hepatocellular carcinoma.
381   *BMC Cancer* 18.
382 Zhao Y, Ge X, He J, Cheng Y, Wang Z, Wang J, and Sun L. 2019. The prognostic value of tumor-infiltrating
383   lymphocytes in colorectal cancer differs by anatomical subsite: a systematic review and meta-analysis.
384   *World J Surg Oncol* 17:85.
385

Figure 1

Figure 1. Stromal and immune scores are associated with the clinical features of colon cancer and their overall survival.

(A) Distribution of stromal scores of colon cancer between different histologic diagnosis. (B) The correlation between stromal scores and anatomic neoplasm subdivision. (C) Distribution of immune scores of colon cancer between different histologic diagnosis. (D) The correlation between immune scores and anatomic neoplasm subdivision. (E) Kaplan-Meier survival curve for patients with low vs. high stromal scores. (F) Kaplan-Meier survival curve for patients with low vs. high immune scores. (G) Kaplan-Meier survival curve for patients with low vs. high both stromal and immune scores.
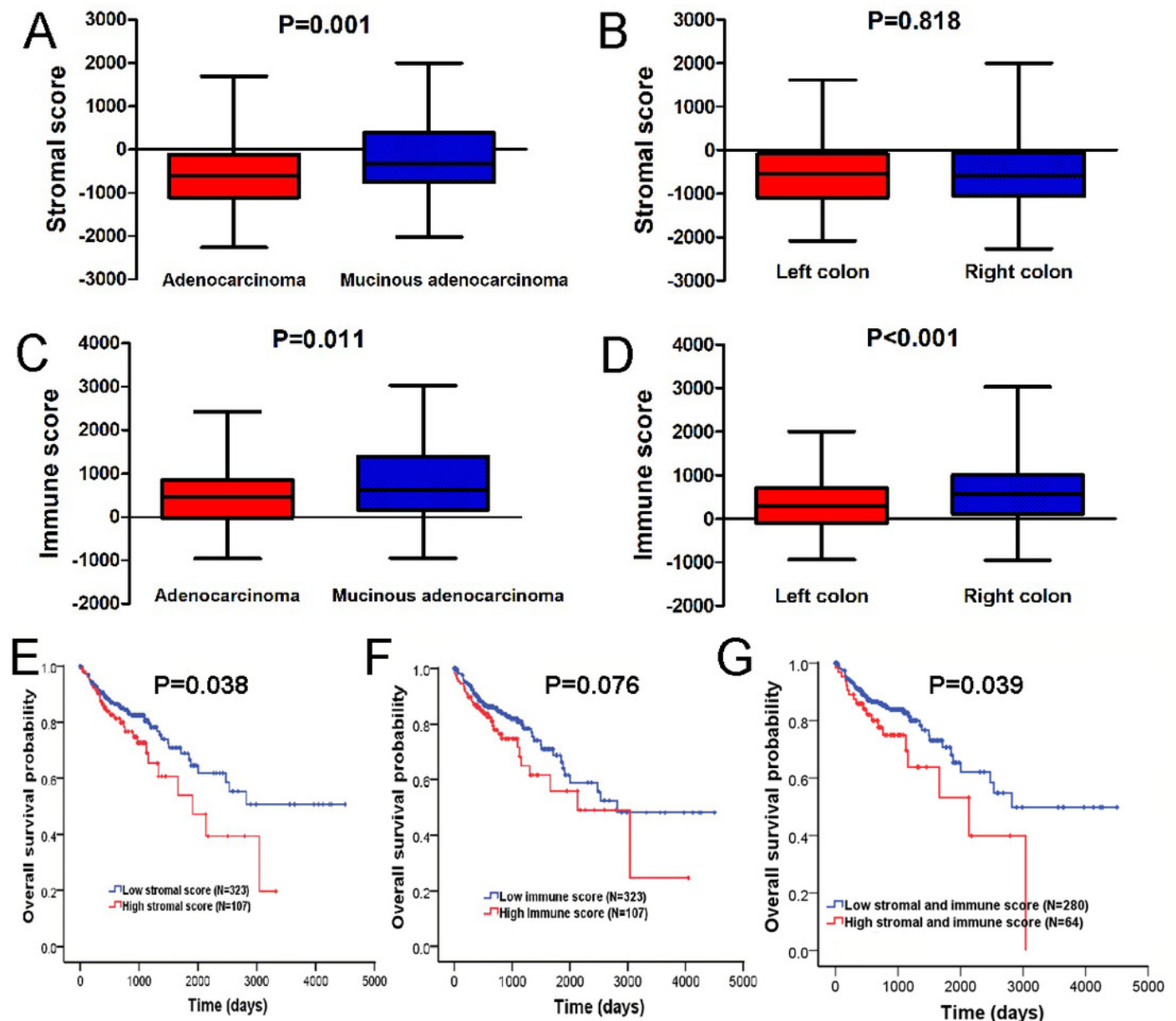
Figure 2

Figure 2. Expression profiles and biological functions of DEGs based on stromal and immune scores.

(A) Volcano plot showing up-regulated DEGs in red and down-regulated DEGs in blue for the comparison based on high and low stromal score groups. (B) Volcano plot showing up-regulated DEGs in red and down-regulated DEGs in green for the comparison based on high and low immune score groups. (C, D) Venn diagrams showing 563 shared up-regulated DEGs (C) and 9 shared down-regulated DEGs (D) from stromal score and immune score groups. (E, F, G, H) Top six Go terms and pathways enriched by DEGs.

Figure 3

Figure 3. Forest plot of hazard ratios (HR) for 70 prognostic DEGs in colon cancer. HR and 95% confidence intervals (CI) were obtained by the COX proportional hazards model.

HR and 95% confidence intervals (CI) were obtained by the COX proportional hazards model.

Prognostic gene	HR (95%CI)	P value
PDE4DIP1	146.84(4.42-4884.01)	0.0053
CTB-60B18.6	100.35(3.12-3229.54)	0.0093
AC113194.1	26.00(2.61-258.61)	0.0054
SLITRK2	5.66(2.06-15.52)	0.0007
CLVS2	5.64(2.49-12.81)	<0.0001
AL137026.1	7.16(1.78-28.72)	0.0055
CACNG5	3.88(1.57-9.59)	0.0034
CXCR5	3.38(1.78-6.41)	0.0002
ANKRD34C	4.15(2.01-8.56)	0.0001
AL161781.2	2.83(1.44-5.59)	0.0027
H3-2	3.07(1.39-6.76)	0.0055
AQP4	2.82(1.52-5.22)	0.0010
LINC02295	2.87(1.35-6.09)	0.0061
CHST9	2.79(1.29-6.03)	0.0093
AC023794.5	2.20(1.28-3.78)	0.0045
NRXN1	1.99(1.18-3.34)	0.0096
CUX2	2.21(1.49-3.30)	0.0001
GABRG2	2.26(1.53-3.33)	<0.0001
PANTR1	1.93(1.47-2.54)	<0.0001
HCN4	1.84(1.37-2.47)	0.0001
NTRK1	1.72(1.22-2.43)	0.0020
AP000894.2	2.00(1.27-3.12)	0.0025
ASTN1	2.04(1.40-2.96)	0.0002
SIGLEC22P	1.54(1.11-2.14)	0.0091
PHF21B	1.70(1.17-2.46)	0.0055
PLEKHG4B	1.79(1.29-2.48)	0.0004
SYT4	1.30(1.16-1.45)	<0.0001
HMP19	1.32(1.08-1.62)	0.0072
IGF1	1.31(1.08-1.58)	0.0065
AC245041.2	1.34(1.17-1.53)	<0.0001
JPH4	1.42(1.21-1.66)	<0.0001
ATP2B3	1.41(1.14-1.75)	0.0015

Prognostic gene	HR (95%CI)	P value
IGLON5	1.12(1.07-1.18)	<0.0001
KIRREL2	1.16(1.09-1.24)	<0.0001
SEZ6	1.13(1.07-1.20)	<0.0001
TMEM59L	1.12(1.06-1.18)	<0.0001
CHST8	1.13(1.07-1.19)	<0.0001
SBK2	1.11(1.06-1.17)	<0.0001
PCSK2	1.17(1.09-1.26)	<0.0001
AC245041.1	1.12(1.06-1.19)	0.0001
ACTBL2	1.26(1.12-1.43)	0.0002
ADCYAP1	1.17(1.08-1.28)	0.0003
MUC16	1.21(1.09-1.34)	0.0003
AC083967.1	1.26(1.11-1.43)	0.0004
MEOX2	1.06(1.01-1.09)	0.0005
CIDEA	1.11(1.04-1.17)	0.0006
CLEC17A	1.23(1.09-1.39)	0.0007
SNAP25	1.10(1.04-1.16)	0.0007
HAND1	1.02(1.01-1.04)	0.0009
LY6K	1.11(1.04-1.18)	0.0010
CD79B	1.03(1.01-1.05)	0.0011
FAM129C	1.21(1.08-1.37)	0.0013
FCRL1	1.16(1.06-1.27)	0.0014
CD37	1.03(1.01-1.05)	0.0014
FCRL3	1.21(1.08-1.37)	0.0016
CD22	1.06(1.02-1.09)	0.0020
LINC01857	1.11(1.04-1.20)	0.0029
TUSC5	1.06(1.02-1.10)	0.0032
PLA2G2D	1.03(1.01-1.05)	0.0034
ANXA8	1.09(1.03-1.15)	0.0035
IGHV3OR16-7	1.16(1.05-1.28)	0.0041
HAS1	1.07(1.02-1.13)	0.0046
EBF3	1.22(1.06-1.41)	0.0050
FCRLA	1.11(1.03-1.19)	0.0058
MS4A1	1.03(1.01-1.06)	0.0067
GPR87	1.05(1.01-1.10)	0.0075
U62631.1	1.23(1.06-1.43)	0.0079
CADM3	1.12(1.03-1.21)	0.0088
TAGLN3	1.15(1.03-1.29)	0.0098
FCRL2	1.24(1.11-1.39)	0.0001

Figure 4

Figure 4. PPI network, GO analysis and expression of hub gene.

(A) PPI network contained 30 nodes and 53 edges based on STRING tool and Cytoscape software was constructed. The hub genes were selected by MCODE in yellow with 9 nodes and 22 edges in the left module, and 6 nodes and 12 edges in the right module. (B) The GO analysis of hub genes was performed using BiNGO. The color of the node represented the corrected P-value of ontologies, and the size of the node represented the numbers of involved genes in the ontologies. P<0.01 was considered statistically significant. (C) The expression of 15 hub genes in TCGA cohort by z-score, with red represents higher expression and blue represents lower expression.
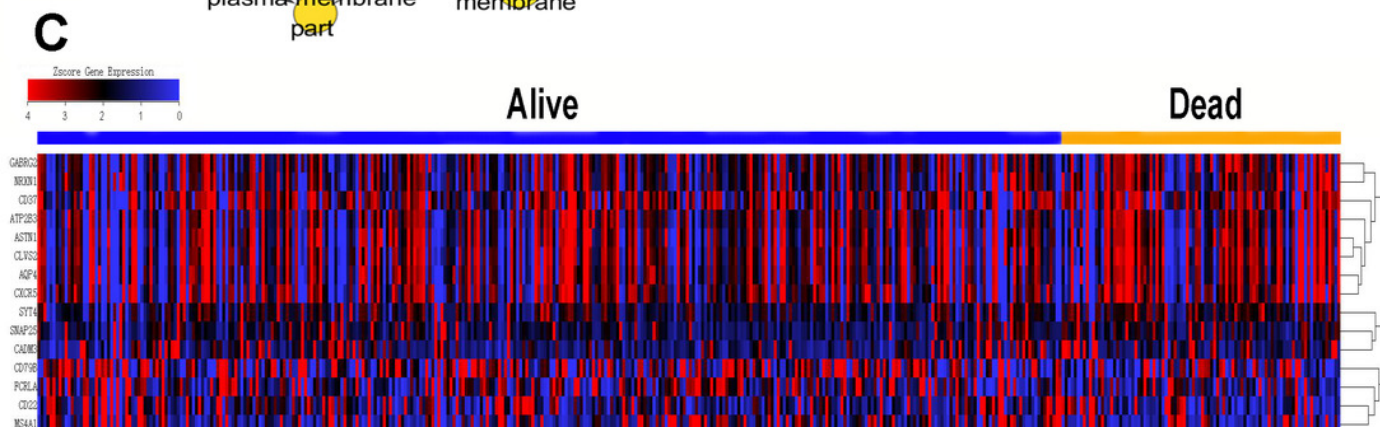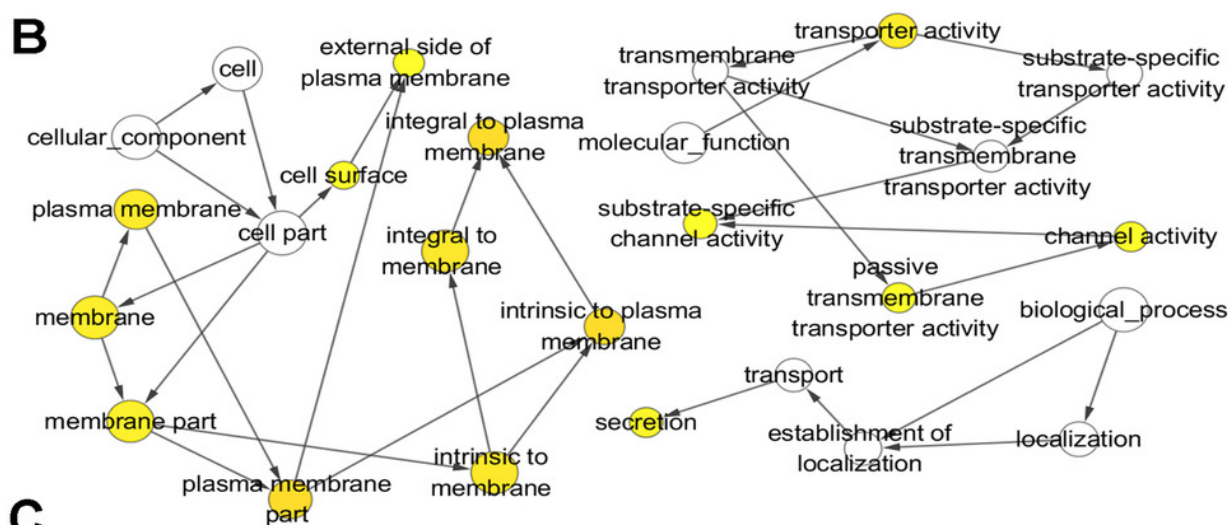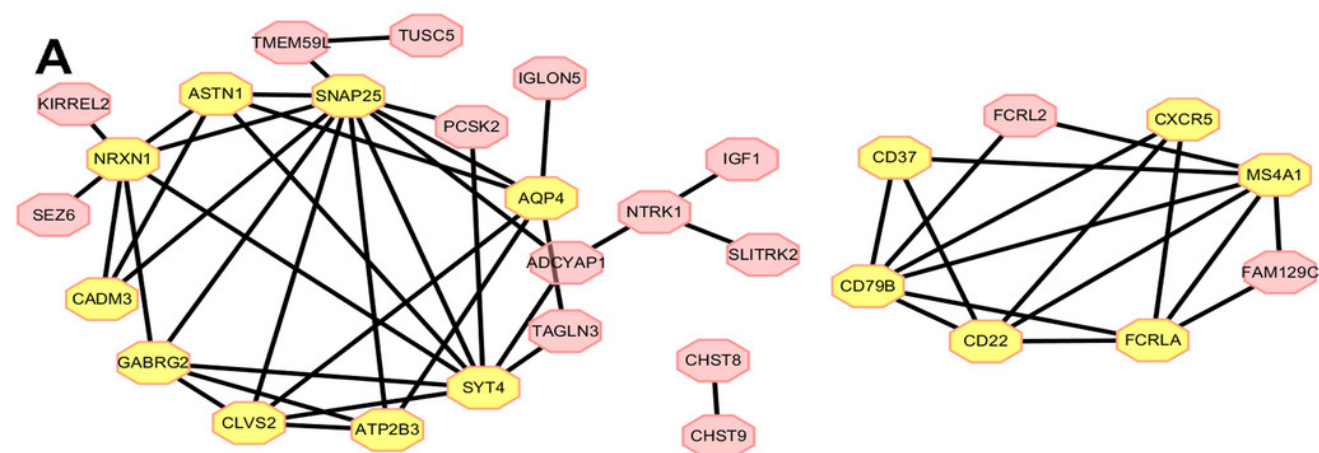
Figure 5

Figure 5. verifying the survival outcomes of hub genes in the GEO database.

Kaplan-Meier survival curves with the log-rank test and COX proportional hazards regression analysis were generated for prognostic genes. The dotted blue/ red lines indicated 95% confidence intervals of overall survival probability. P<0.05 was considered statistically significant.
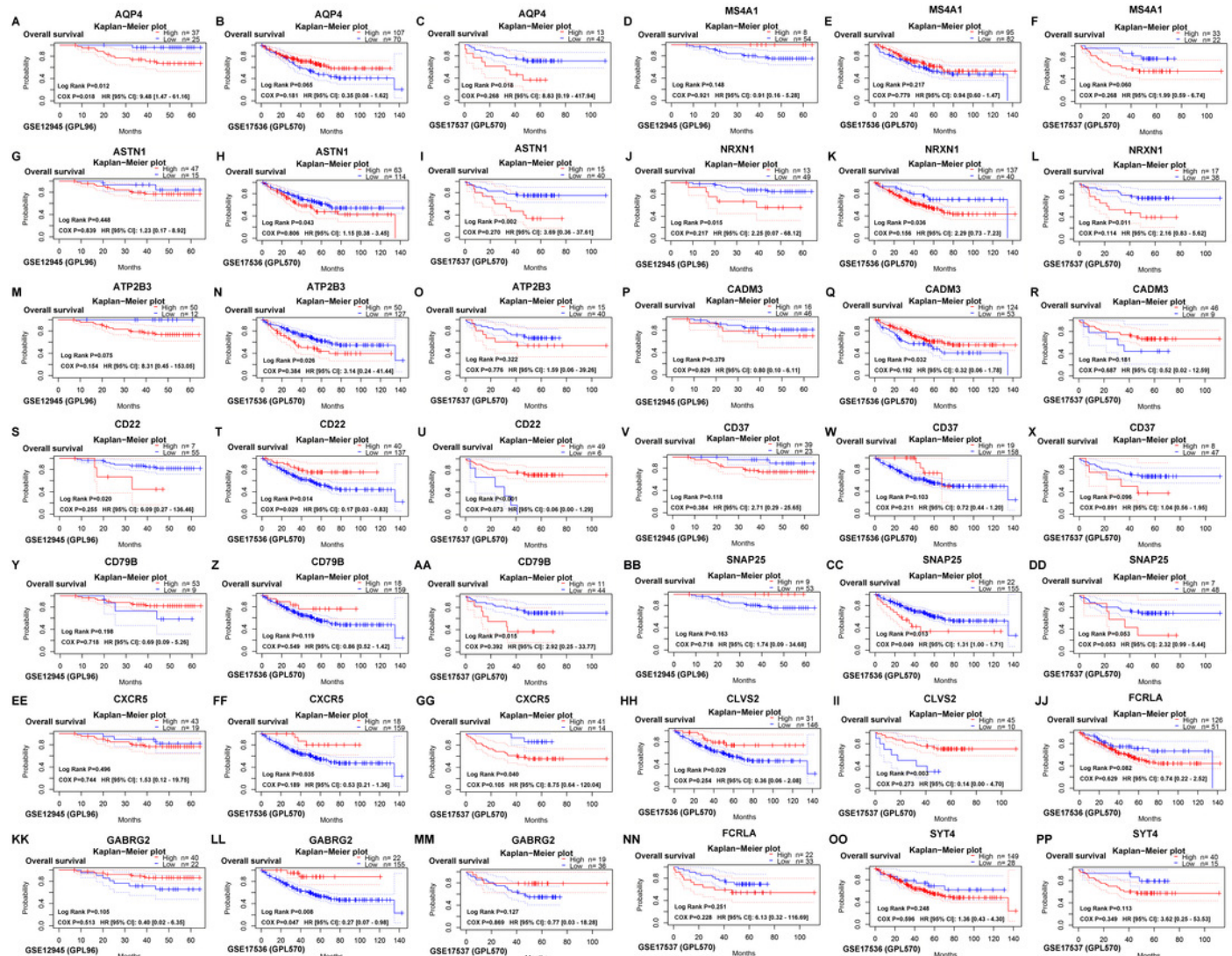
Figure 6

Figure 6 The relative mRNA expression levels of AQP4 (A) and SNAP25 (B) in colon cancer tissues and adjacent normal tissues were confirmed by qRT-PCR.
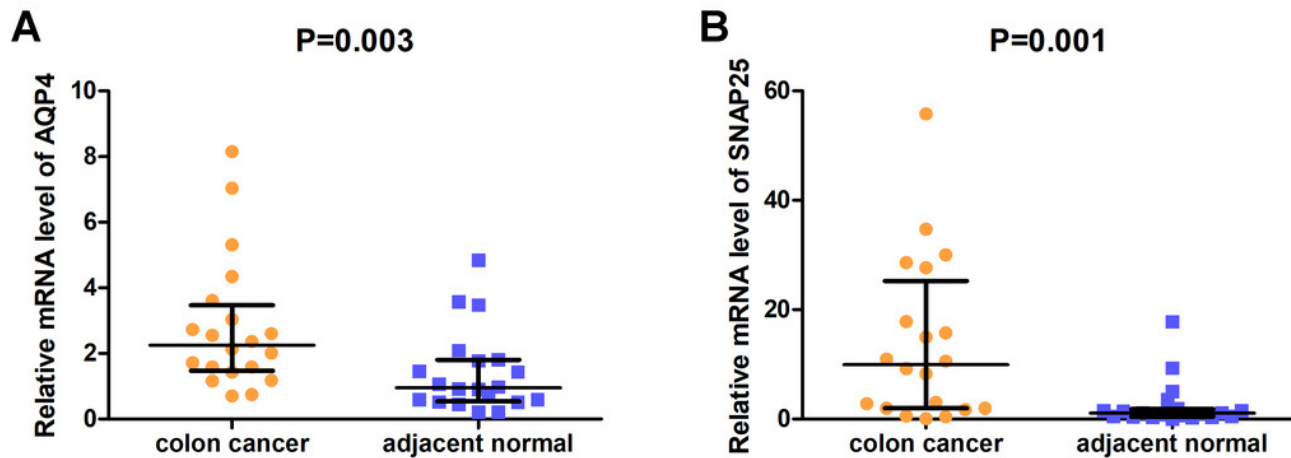
Figure 7

Figure 7. GSEA enrichment plots showed that eleven gene sets related to tumor signaling pathways were enriched in the high SNAP25 expression group.
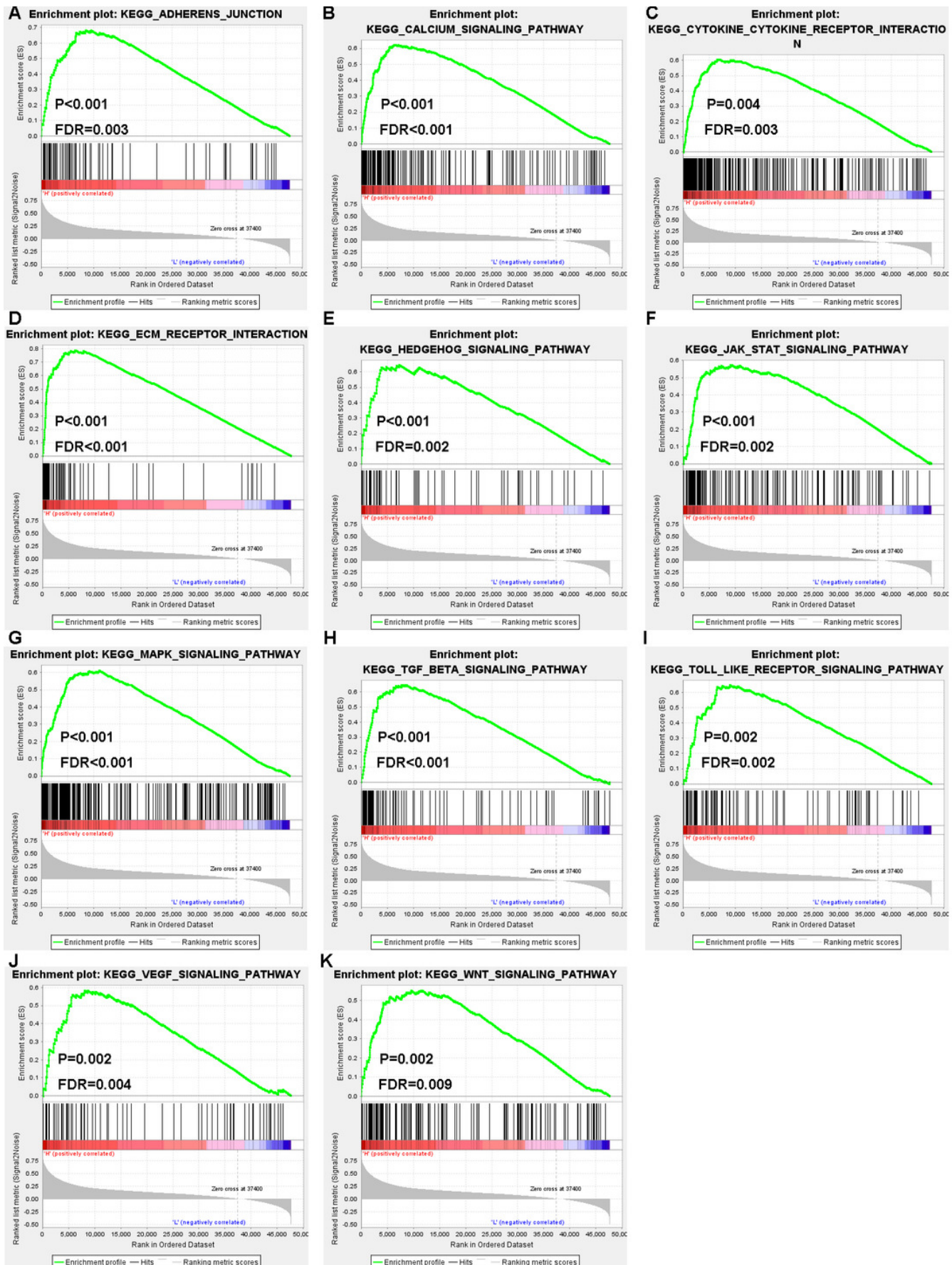
Figure 8

Figure 8. Spearman's correlations between SNAP25 and lymphocytes (TISIDB).

(A) Relations between the SNAP25 expression and abundance of TILs across human cancers.
(B) Top 4 greatest positive correlations between SNAP25 expression and TILs. (C) Relations
between the SNAP25 methylation and abundance of TILs across human cancers. (D) Top 4
greatest negative correlations between SNAP25 methylation and TILs.