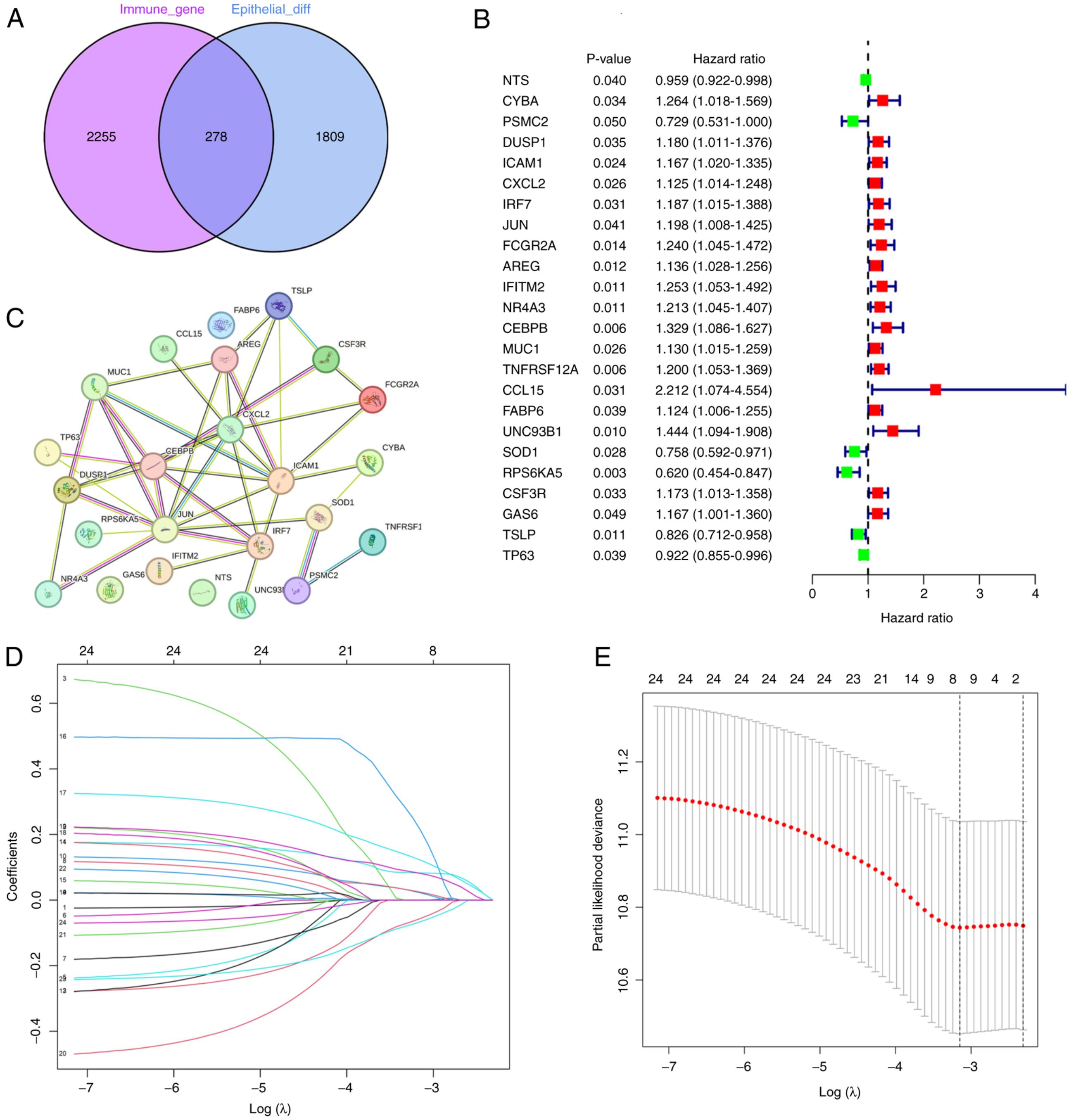

Figure S1. Filter for intersection gene and LASSO-Cox gene analyses. (A) Intersection of epithelial immune genes. (B) Gene set expression of genes associated with prognosis. (C) Protein-protein interaction network among 24 cox genes. (D) LASSO regression analysis coefficient distribution diagram and (E) LASSO regression analysis parameter graph. LASSO, least absolute shrinkage and selection operator.

Figure S2. Distribution of risk scores in the training cohort, the test cohort and the whole cohort, respectively. Heat map of 4 epithelial immune-related genes expressions for the (A) training, (B) test and (C) whole group. Risk curve for the risk scores in the (D) training group, (E) test group and (F) whole group. Scatterplot showing the survival status of each patient in the (G) training, (H) test and (I) whole group. AREG, amphiregulin; MUC1, mucin-1; FABP6, fatty acid-binding protein 6; TSLP, thymic stromal lymphopoietin.

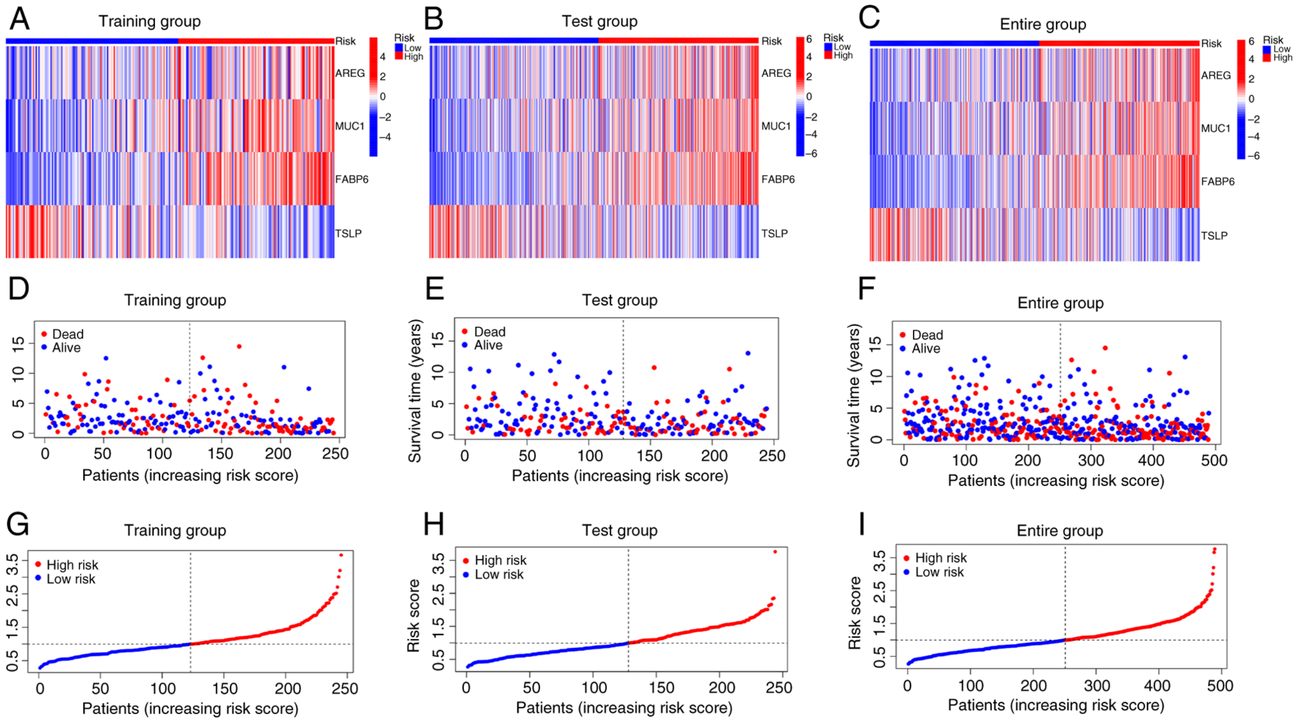

Figure S3. Survival prediction and differential expression of model genes in patients with high or low risk. Kaplan-Meier curve of (A) AREG, (B) FABP6, (C) TSLP and (D) MUC1. Differential expression of (E) AREG, (F) FABP6, (G) TSLP and (H) MUC1. AREG, amphiregulin; MUC1, mucin-1; FABP6, fatty acid-binding protein 6; TSLP, thymic stromal lymphopoietin.

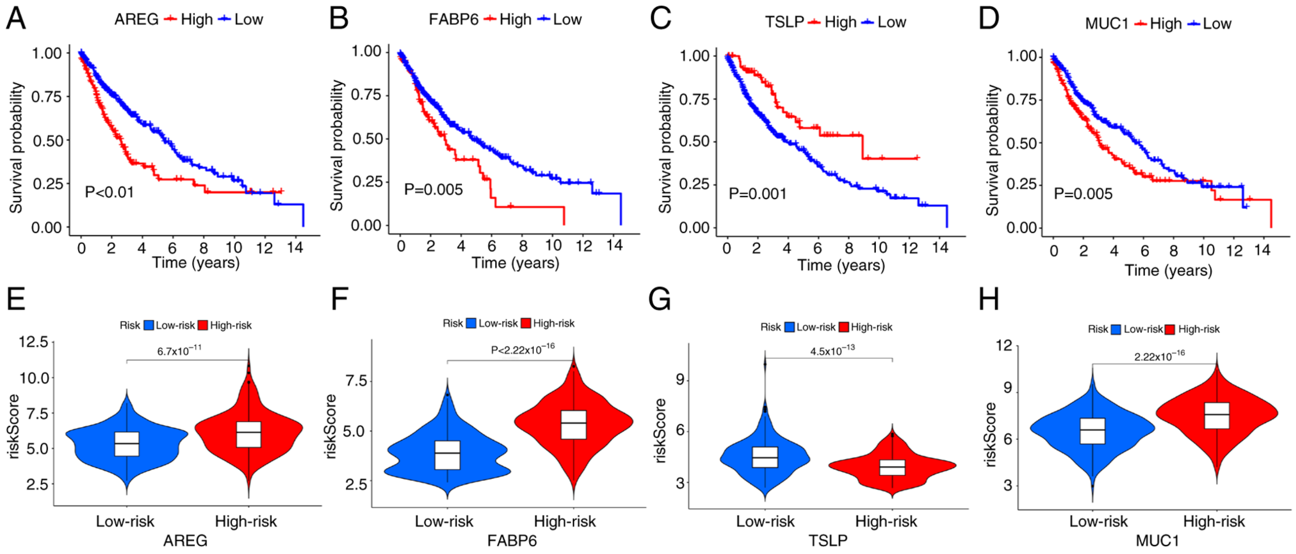

Figure S4. Model and clinical relationship. (A) Univariate and (B) multivariate cox analysis. (C) Differences in clinical factors between the high and low risk groups. (D) Heat map of clinical traits in high and low risk groups. $^{*}P<0.05$; $^{**}P<0.01$. T, tumor stage; N, lymph node stage; M, metastasis stage.

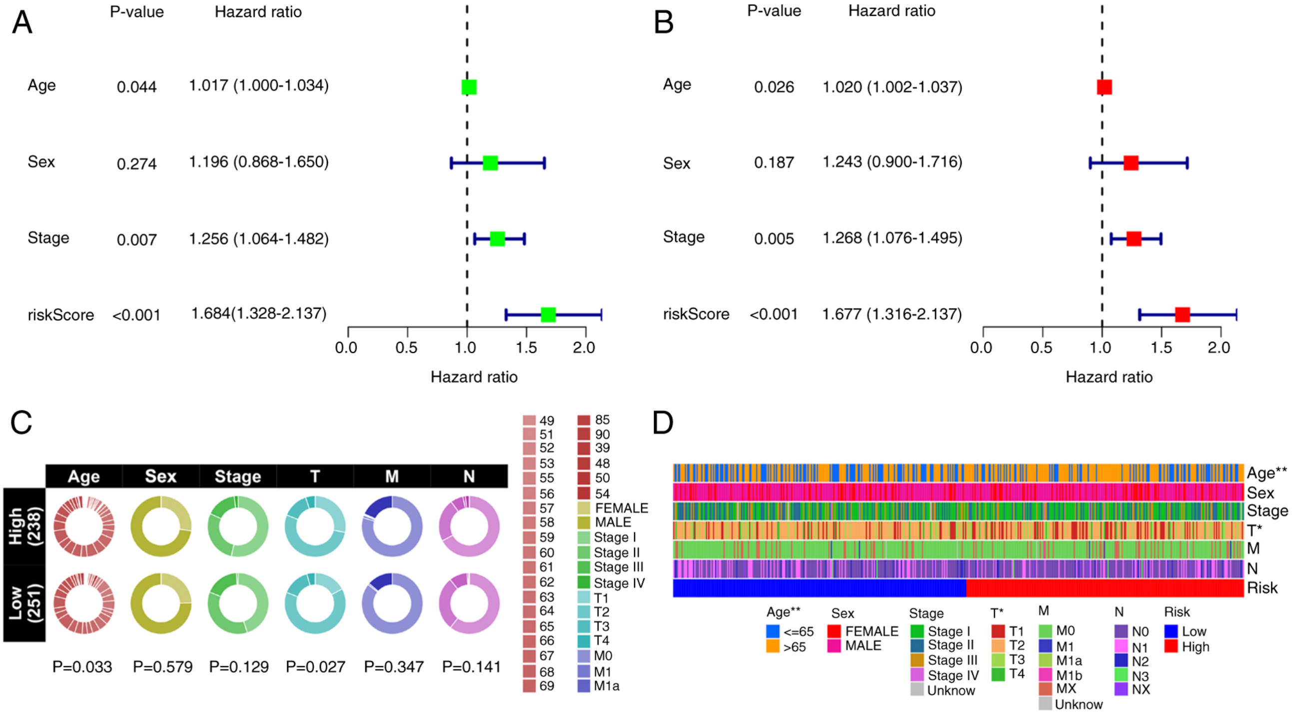

Figure S5. Clinical traits as prospective estimates of prognosis and external database validation. Correlation analysis of risk scorewithage (A); Correlation analysis of risk sore with sex (B) and (C) stage (C); The Kaplan-Meier analysis curves for the patients stratified by risk score and age (D). Kaplan-Meier analysis curves for the patients stratified by risk score and sex (E); The Kaplan-Meier analysis curves for the patients stratified by risk score and stage (F); (G) T stage and (H) N stage. (I) Gene Expression Omnibus external validation. T, tumor; N, lymph node.

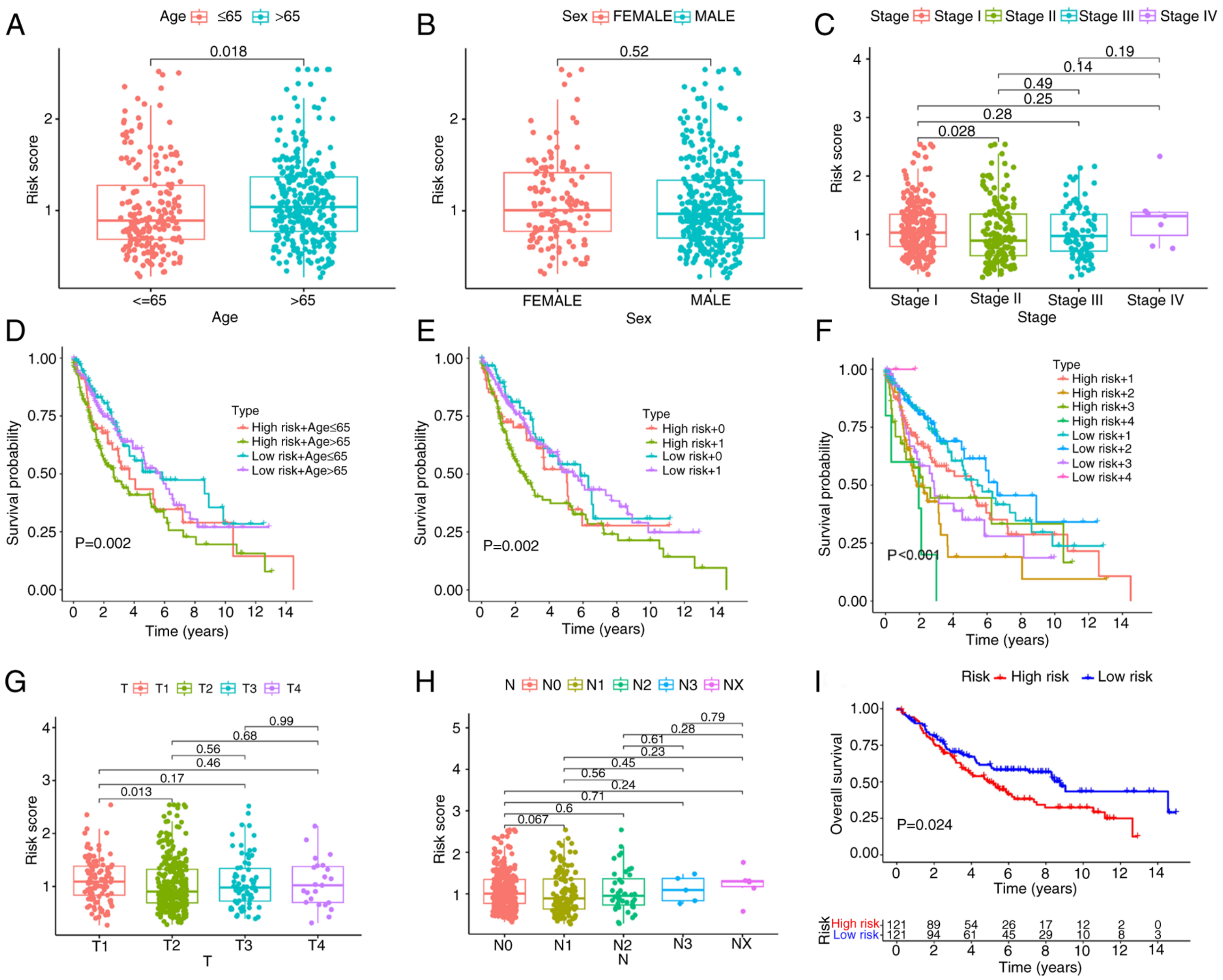

Figure S6. Immune checkpoint and IMvigor210 cohort external validation. (A) Kaplan-Meier curves for patients with high-risk and low-risk scores in the IMvigor210 cohort. (B) Differences in CR/PR and SD/PD among different risk groups in IMvigor210 cohort (C) IMvigor210 cohort prediction of immunotherapy outcomes. (D) PVR gene expression was different in different risk groups. (E) Correlation between risk score and PVR gene expression (F) TNFRSF14 gene expression was different in different risk groups. (G) Correlation between riskscore and PVR gene expression. AUC, area under the curve; CI, confidence interval; CR, complete response; PR, partial response; PD, progressive disease; SD, stable disease; PVR, poliovirus replication cell adhesion molecule; TNFRSF14, tumor necrosis factor receptor superfamily Member 14.

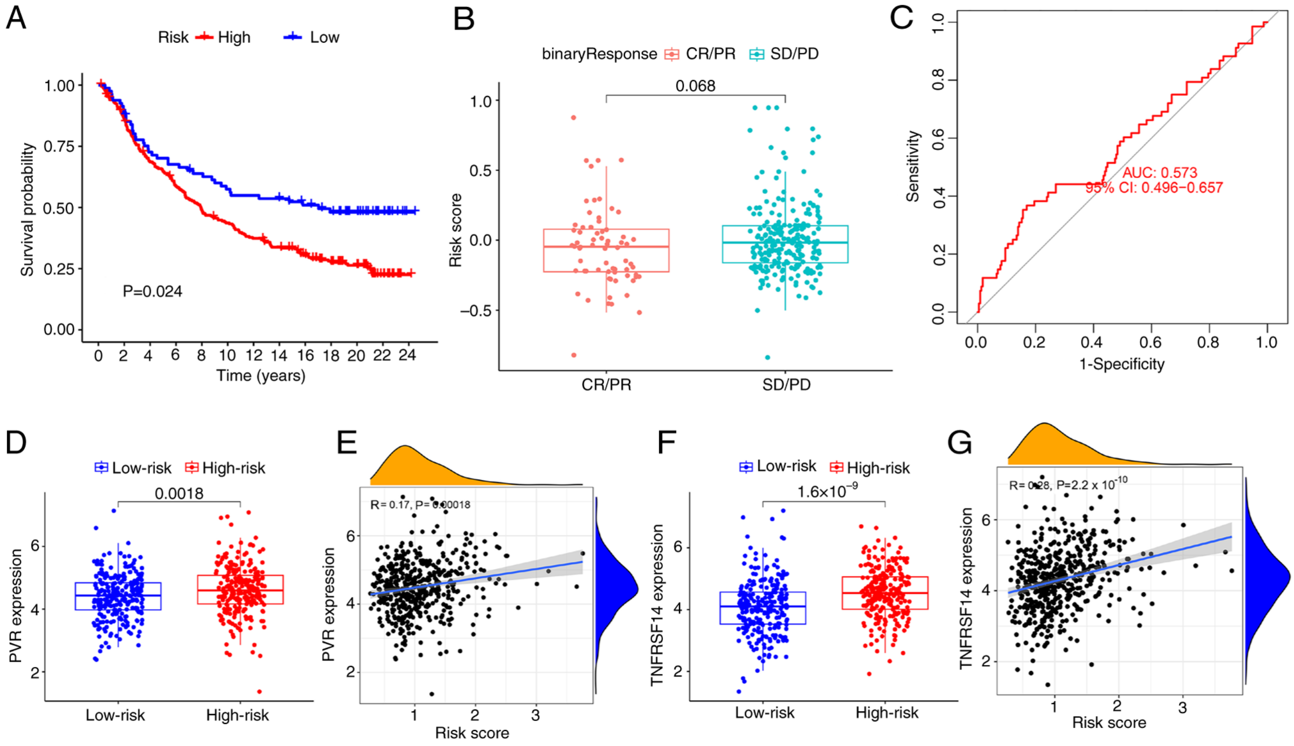

Figure S7. Other immune checkpoints differences. (A) CD40; (B) CD44; (C) CD70; (D) CD80; (E) CD86; (F) CD200; (G) CD244; (H) CD274; (I) CTLA4; (J) ICOS; (K) ICOSLG; and (L) IDOL1. CTLA4, cytotoxic T-lymphocyte-associated protein 4; ICOS, Inducible. T Cell Costimulator; ICOSLG, Inducible T Cell Costimulator Ligand; IDOL1, Indoleamine 2,3-Dioxygenase 1.

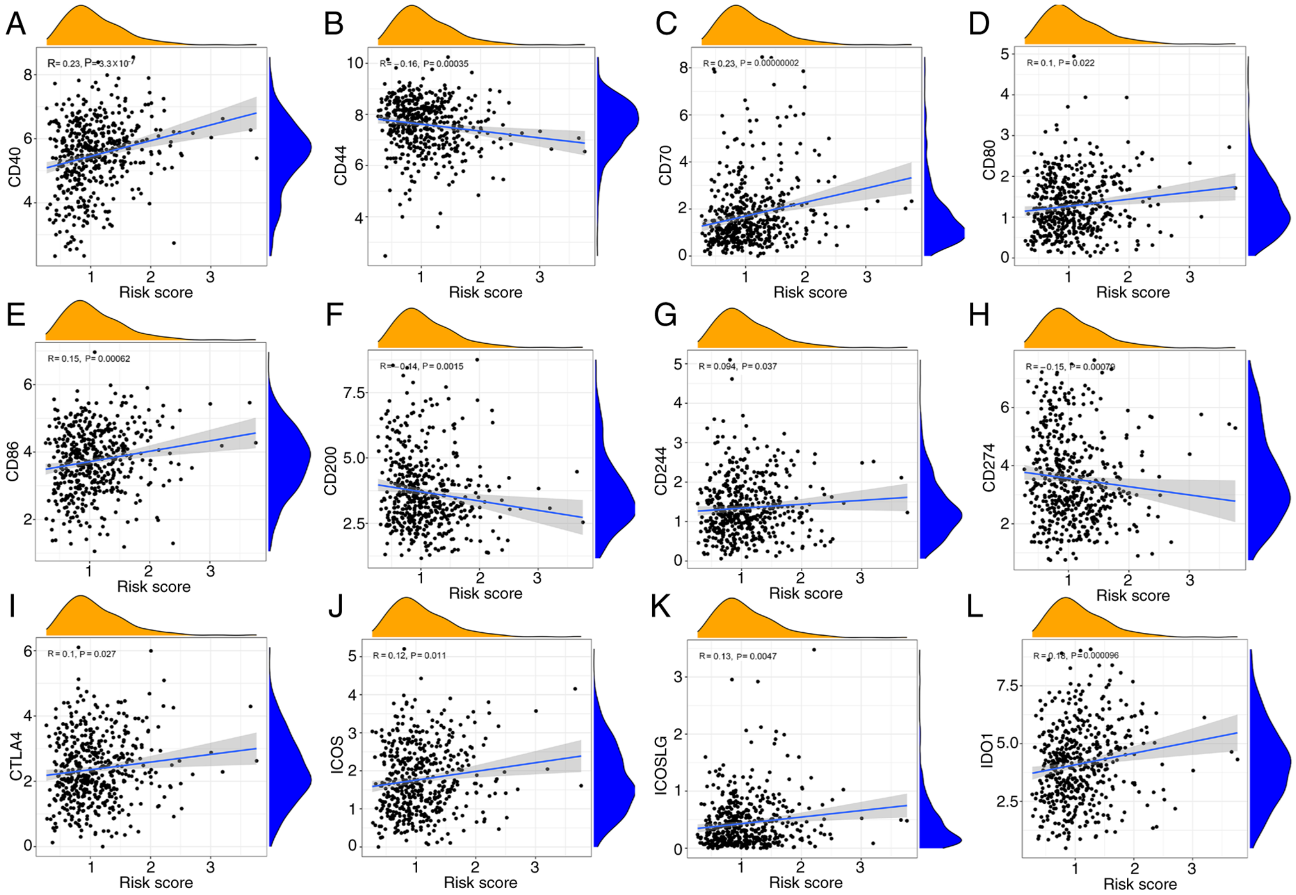

Figure S8. *In vitro* experimental validation of the risk model. (A) Immunohistochemistry images of the normal and tumor status of genes from the Human Protein Atlas database. (B) Reverse transcription-quantitative PCR results. $^{**}P<0.05$; $^{***}P<0.01$. AREG, amphiregulin; MUC1, mucin-1; FABP6, fatty acid-binding protein 6; TSLP, thymic stromal lymphopoietin.