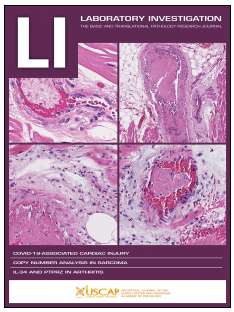
Journal Pre-proof



Expression of IDO1 in tumor microenvironment significantly predicts the risk of recurrence/distant metastasis for patients with esophageal squamous cell carcinoma

Cen Qiu, Qingyuan Lin, Saiyan Ji, Chenjie Han, Qingyuan Yang

PII: S0023-6837(23)00206-4

DOI: https://doi.org/10.1016/j.labinv.2023.100263

Reference: LABINV 100263

To appear in: *Laboratory Investigation*

Received Date: 21 June 2023

Revised Date: 6 October 2023

Accepted Date: 9 October 2023

Please cite this article as: Qiu C, Lin Q, Ji S, Han C, Yang Q, Expression of IDO1 in tumor microenvironment significantly predicts the risk of recurrence/distant metastasis for patients with esophageal squamous cell carcinoma, *Laboratory Investigation* (2023), doi: https://doi.org/10.1016/j.labinv.2023.100263.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2023 United States & Canadian Academy of Pathology. Published by ELSEVIER INC. All rights reserved.

Expression of IDO1 in tumor microenvironment significantly predicts the risk of recurrence/distant metastasis for patients with esophageal squamous cell carcinoma

Cen, Qiu*; Qingyuan, Lin*; Saiyan, Ji; Chenjie, Han; Qingyuan, Yang

Department of Pathology, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China, 200011

*These authors contributed equally to the study.

Corresponding author: Qingyuan, Yang, No. 639, Manufacturing Bureau Road, Shanghai, China, 200011. Email: shengzhou-2005@163.com. Tel: 86-021-53315212.

Competing interests: All the listed authors declare that no competing interests exist.

Abstract

In this study, we aimed to explore immune markers predicting locoregional recurrence/distant metastasis (R/M) for patients with esophageal squamous cell carcinoma (ESCC) post-surgical intervention by using a novel high-throughput spatial tool to quantify multiple immune proteins expressed in ESCC and lymphocytes in tumor microenvironment (TME-L). First, formalin-fixed paraffin-embedded tissues from surgical patients with ESCC (n=94) were constructed on a microarray，which were then divided into discovery (n=36) and validation cohorts (n=58). Using a newly developed GeoMx digital spatial profiling tool, 31 immune proteins in paired ESCC and TME-L, morphologically segmented by PANCK and CD45, respectively, from the discovery cohort were quantified, releasing 2232 variables. Next, the correlation

matrix was analyzed using the Corrplot package in R Studio, resulting in six closely correlated clusters. The Least Absolute Shrinkage and Selection Operator regression scoring model predictive of R/M risk with superior specificity was successfully established based on three hierarchically clustered immune proteins: ARG1 in ESCC/PANCK+, STING and IDO1 in TME-L/CD45+. Moreover, the expression of IDO1 in TME-L, rather than in ESCC, significantly predicted the R/M risk score with an area under curve of 0.9598. Besides, its correlation with R/M status was further validated by dual immunohistochemistry staining of IDO1 and CD45 in discovery and validation cohorts. Above all, our findings not only provide a more accurate scoring approach based on quantitative immune proteins for the prediction of R/M risk, but also propose that IDO1 in TME-L potentially plays a driving role in mediating R/M in ESCC.

Keywords: esophageal squamous cell carcinoma; digital spatial profiling; PanCK; CD45; recurrence/distant metastasis prediction; tumor microenvironment

Introduction

Esophageal squamous cell carcinoma (ESCC) is the predominant pathological type of esophageal cancer (EC) in China (1). Despite advances in diagnosis, prognosis, and improvements in neoadjuvant and adjuvant treatments, the overall 5-year survival rate of patients with ESCC is only 40%, with locoregional recurrence and/or distant metastasis (R/M) post-surgical intervention as the primary causes (1). The pathological TNM stage and treatment regimen have been proposed as risk-based

strategies for relapse prediction (2). To our knowledge, the equation developed by Wang et al. is the only scoring model for predicting postoperative recurrence in patients with ESCC (3). However, its predictive accuracy is not superior to that of TNM stage. Additionally, it is of limited clinical importance in guiding early interventions to prevent relapse. Thus, a more accurate R/M predictive model based on quantifiable targeted molecules is urgently required.

Multiple immune proteins have been suggested as prognostic markers for patients with ESCC in different clinical scenarios. A C-index based on the semiquantitative levels of four immune markers in different regions of ESCCs , including CD8 and FOXP3 in tumor infiltrating lymphocytes (TIL), CD33 in myeloid-derived suppressor cells (MDSC), and PD-L1 in ESCC, has shown superior precision in predicting 5-year overall survival (OS) compared to the TNM staging system (4), suggesting that the spatial distribution of immune proteins, known as the immune context (5), has different clinical significance. Unfortunately, no predictive R/M models have been established based on immune proteins identified using quantitative high-throughput spatial analysis.

GeoMx digital spatial profiling (DSP) is a newly developed methodology that combines high-throughput sequencing, photochemistry, digital light projection, and microfluidic technologies to advance semiquantitative immunohistochemistry (IHC) and RNA in situ hybridization techniques. It has been commonly used to assess the spatial distribution of proteins and RNA in formalin-fixed paraffin-embedded tissues

(FFPE) (6). For instance, Kulasinghe et al. has recently employed this tool to study the differentially expressed immune and PI3K/AKT signaling proteins between patients with and without relapse, and uncover that NF1 in tumor compartment as well as ERα in stromal compartment were significantly indicative of therapy response in triple-negative breast cancer (TNBC) (7).

Thus, in the present study, we used this platform to quantitatively detect a panel of 31 immune proteins, including 13 immune cell markers, 8 immune activation markers, and 10 immunotherapy targets, in ESCC and lymphocytes in tumor microenvironment (TME-L), which were morphologically segmented by positive staining of PANCK (PANCK+) and CD45 (CD45+), respectively. Next, the correlation matrix of all variables revealed six closely correlated clusters, which were subsequently used to construct the R/M risk scoring model using a Least Absolute Shrinkage and Selection Operator (LASSO) regression method. Interestingly, a three-protein-based R/M risk scoring model with better specificity and positive predictive value (PPV) than the UICC8 staging system was successfully established. Moreover, the IDO1 expression in TME-L was not only linearly related to above R/M risk score, but also specifically predicted it with an area under curve (AUC) of 0.9598. Besides, results from dual IHC staining of CD45 and IDO1 in discovery and validation cohorts confirmed its close correlation with R/M status. Above all, findings in our study not only provide a more accurate scoring approach based on quantitative immune proteins for the prediction of R/M risk, but also propose that IDO1 in TME-L potentially plays a driving role in

mediating R/M in ESCC.

Materials and Methods

Study population and ethics

94 patients with ESCC, who underwent radical surgery without neoadjuvant therapy at Huashan Hospital of Fudan University between September 2012 and November 2017 were included in the study. The annual follow-up visits of patients' survival and disease progression including R/M were conducted by clinicians. Of note, the R/M here is defined as patients who were diagnosed with locoregional recurrence and/or distant metastasis after surgical intervention. The median follow-up duration was 58 months (range, 2-72 months). This study was approved by the Institutional Research Ethics Committee of Huashan Hospital (protocol no. KY2022-614). All the procedures were performed in accordance with the 1975 Declaration of Helsinki.

Tissue microarray construction

FFPE tissue slides of 94 ESCC patients receiving surgical interventions were stained with haematoxylin and eosin (H&E), and pathologists circled the tumor regions to construct the tissue microarray (TMA) with one core per patient. Then, the TMA block with a core diameter of 2.0 mm was constructed by Outdo Biotech (Shanghai) and arranged in a 12-column, 8-row format. Subsequently, the TMA block was continuously sliced for downstream assays.

GeoMx digital spatial profiling

A total of 42 markers tested by the GeoMx digital spatial profiling (DSP) tool with

nCounter system (Table S1) (8), categorized as background controls (n=6), housekeeping controls (n=3), morphological markers (n=2), immune cell markers (n=13), immune activation status markers (n=8), and immune oncology (IO) drug targeted markers (n=10), were performed by FynnBio (Shandong). Briefly, the TMA slide was co-stained with oligo-conjugated antibodies (Nanostring, Seatle, US) targeting the above markers and the nuclei were counterstained with SYTO13 (NanoString, Seatle, US). Subsequently, the slide was scanned to select regions of interest (ROI) in the paired ESCC/PANCK+ and TME-L/CD45+ compartments from the same patient. 36 samples were randomly selected from the overall study cohort to from the discovery cohort. The protein oligos in the ROIs were then UV-photocleaved and capillary-aspirated into a 96-well collection plate, followed by sealing with Fluoromount G mounting medium (Northern Biotech, Waltham, US) and hybridization of the probe with the Hyb Code pack master mix (NanoString, Seatle, US). The samples were then automatically prepared by the nCounter Prep Station and counted using the nCounter Analyzer. In particular, the original nCount data from all ROIs were quality checked and normalized to housekeeping proteins according to the commercially established pipeline (NanoString, Seatle, US) prior to downstream analysis.

Bioinformatics analysis of DSP data

The correlation matrix analysis of immune proteins (n=31) in the PANCK+ and CD45+ compartments of 36 samples was performed in R Studio using the Corrplot package

(version 0.92) with specific arguments (sig.level=0.05, insig='blank', order='hclust'). Next, hierarchically clustered immune proteins were used to establish the LASSO regression model in R Studio using the glmnet package (version 4.1-4) to predict R/M risk in patients with ESCC after surgery. The risk score was calculated following the formula: risk score= $\sum_1^n Coef(i) * \exp(i)$, with Coef (i) meaning the coefficient of each variable and exp (i) representing the expression level of corresponding variable.

Single and dual immunohistochemistry staining and interpretation

The TMA block was sliced into 3 μm-thick slides and automatically stained with the indicated antibodies on a Benchmark Ultra instrument (Roche, USA), accompanied with their corresponding negative and positive controls according to the manufacture's protocol. Finally, for single-marker staining, the antibodies were stained brown with horseradish peroxidase (HRP)-mediated 3,3-diaminobenzidine (DAB). For dual-marker staining, CD45 was stained brown with HRP-mediated DAB, and IDO1 or STING was stained pink with alkaline phosphatase-mediated Fast Red conversion. The slides were then scanned as digital images using a NanoZoomer S60 scanner (Hamamatsu Photonics, Japan) with an NDP Scanner. Representative results were obtained by exporting selected regions from the NDP.view2 software (version 2.8.24). The antibodies used for IHC staining are summarized in Table S2.

To interpret the staining signals of STING and cGAS, the percentage of positive signals (cytoplasmic brown) and their intensity in ESCC/PANCK+ and TME-L/CD45+ cells, respectively, was independently and blindly evaluated by two pathologists. In

accordance with previous scoring methods (9), the IHC score was calculated by multiplying the staining intensity score with the percentage score to obtain a value between 0 and 12. Regarding to the interpretation of dual staining signals, the percentage of cells with pink staining was evaluated in ESCC/PANCK+ and TME-L/CD45+ cells. STING was considered low or high with signals of <20% or ⩾ 20%, whereas IDO1 was defined as low or high with signals of <1% or ⩾1%.

Statistics

Continuous data were analyzed and graphically illustrated as mean ± SD using GraphPad Prism 8.0 (GraphPad, La Jolla, CA, USA), while categorical data were analyzed with Pearson's chi-square test using SPSS statistics (version 25.0). Kaplan-Meier curves were analyzed using the logrank test (Mantel-Cox). Independent OS prognostic markers were analyzed using the univariate and multivariate Cox regression tests according to the rule of ten events per variable (10). Statistical significance was considered when $p<0.05$, which was further classified into the following four groups: $p<0.05$, $p<0.01$, $p<0.001$, and $p<0.0001$.

Results

Demographic data of patients in the discovery and validation cohorts

A total of 94 patients with ESCC without neoadjuvant therapy were included in this study, whose FFPE blocks were used for TMA construction with one core per patient. To establish the R/M risk-scoring model, the overall cohort was randomly divided into a discovery cohort with 36 patients and a validation cohort with 58 patients. A detailed

workflow chart of this study is shown in Figure 1.

The discovery cohort included 25 men (69.4%) and 11 women (30.6%) with a median age of 67 years at diagnosis. Up to 38.9% of the tumors (14/36) were pathologically diagnosed as having at least one metastatic lymph node lesion (LNM^{yes}). Similarly, 38.9% of the patients were clinically diagnosed with UICC8 stage III/IV (11). During the postoperative follow-up period, 30.6% (11/36) of patients were diagnosed with R/M, which was clinically confirmed according to the Chinese Society of Clinical Oncology (CSCO) guidelines (12). Similarly, the demographics of patients in the validation cohort were not significantly different from those in the discovery cohort. The detailed clinical characteristics of the two cohorts are summarized in Table S3. No patient received immunotherapy during postoperative treatment.

Differential expression of immune proteins in ESCC and TME-L

In this study, we used a newly developed GeoMx DSP tool with nCounter system (6) that simultaneously profiled multiple proteins in a quantitative manner in FFPE tissues to investigate the distribution of immune proteins (n=31) in ESCC (PANCK+) and TME-L (CD45+) in the discovery cohort (Table S1) (8). Representative paired ROIs in PANCK+ and CD45+ compartments from the same patient are shown in Figure 2A.

After normalization, 2232 variables (Table S4) were used for downstream analysis. Interestingly, most immune proteins were highly expressed in TME-L/CD45+ cells, compared to ESCC/PANCK+ cells (Figure 2B), with exception for four IO drug targeting markers (TIM3, ARG1, IDO1, and LAG3) and one immune activation marker

(CD127). The tumor cell-intrinsic expression of cGAS and STING, which are known to modulate adaptive immune response (13), has been recently uncovered to remodel the TME by inducing infiltration of $CD8^+$ T cells and expression of type I interferon in radiation treated ESCC (14), implicating that cGAS and STING play a critical role in ESCC. Therefore, to confirm the reliability of the DSP tool for distinguishing the differential expression of immune proteins in ESCC and TME-L, we semi-quantified the cGAS and STING in the validation cohort using single IHC staining. As shown in Figure 2C and D, STING expression was significantly higher in TME-L/CD45+ tissues than that in ESCC/PANCK+ tissues. Consistently, as reported in previous studies (15), STING was significantly correlated with cGAS in TME-L cells (r=0.3698, p=0.0288), although cGAS was not differentially expressed in ESCC and TME-L cells. In summary, our high-throughput profiling using the DSP tool reliably provided initial evidence that several immune proteins are highly expressed in the TME-L of ESCC tissues.

The R/M risk scoring model was established based on three closely correlated immune proteins in ESCC

The findings from the above correlation analysis between STING and cGAS prompted us to explore the mutual correlation among all immune proteins profiled in PANCK+ and CD45+ compartments. As shown in Figure 3A, proteins in the CD45+ compartment, including ARG1, LAG3, CTLA4, PDL2 (PD-L2), FN1 (fibronectin 1), PD1 (PD-1), CD137, CD56, GITR, B7H3 (CD276) and CD25, were significantly

positively correlated (Table S5, cluster 1). Similarly, CD40, beta-2-microglobulin (B2M), CD4, CD3 and CD8, expressed in the CD45+ compartment, were also significantly positively correlated with each other (Table S5, cluster 2), indicating a specific subpopulation of T cells in the TME of ESCC. Alternatively, the proteins expressed in the PANCK+ compartment were also closely correlated. For example, GZMB, ICOS, CD27, FN1, CD56, CD137, LAG3, PDL2, CD20, PD1, B2M (Table S5, cluster 3), B2M, STING, VISTA, and OX40L (Table S5, cluster 4). Furthermore, some proteins expressed in ESCC/PANK+ cells were significantly associated with proteins expressed in the TME-L/CD45+ cells. For example, CD68 and CD11c expressed in TME-L positively correlated with PDL1, B7H3, CD127, CD4, CD40, CD68, CD3, and CD8 expression in ESCC (Table S5, cluster 5), and CD80 and ARG1 expressed in ESCC correlated with IDO1, PDL1, ICOS, CD127, CD80, GZMB, OX40L, STING, and TIM3 expressed in TME-L cells (Table S5, cluster 6). As suggested by the relationship between STING and cGAS, we speculated that these six closely correlated groups may play vital roles in the response of patients to adjuvant therapies after radical surgery, and warrant further mechanistic studies.

Notably, there were also some negatively correlated immune proteins in CD45+ and PANCK+ compartments. For example, CD80 in ESCC, PD1 in TME-L cells, LAG3 in ESCC, and the T cell markers CD3, CD8, and CD4 indicated the presence of a negative cross-talk between ESCC and TME-L mediated by these immune proteins. However, the IO drug-targeting marker VISTA in TME-L was negatively correlated

with CD20 and CD11c, both of which were also expressed in TME-L, implicating a negative regulation of VISTA in the ESCC signals presented by dendritic cells (CD11c+) to B lymphocytes (CD20+) in the TME of ESCC (16).

Recurrence, including R/M after primary tumor surgical resection, is a well-known cause of poor prognosis in patients with ESCC (3) and cannot be specifically predicted for early prevention (3). Therefore, unlike prognostic models (4), we attempted to establish a LASSO model that predicted R/M risk based on correlated clusters identified above. Finally, according to the tuning parameter lamda,we successfully established the best LASSO model composed of three immune proteins: ARG1 in PANCK+ (ARG1_PANCK), STING, and IDO1 in CD45+ (STING_CD45, IDO1_CD45), with an intercept of 1.18 (Figure 3B). And the risk score formula is $4.37x10^{-4}$ x (IDO1_CD45) -$9.13x10^{-6}$ x (STING_CD45)-$7.44x10^{-5}$ x (ARG1_PANCK).

R/M risk scores specifically predicted the incidence of R/M for patients with ESCC after surgical resection

The R/M risk score from above LASSO model was further evaluated for the predictive accuracy using the ROC analysis. As shown in Figure 4A, the area under the curve (AUC) was 0.7455 when 1.493 was used as the best cut-off value of R/M risk score (p=0.0204), and the sensitivity and specificity were 63.6% and 96.0%, respectively (Table S6). Furthermore, we calculated its negative predictive value (NPV) and positive predictive value (PPV), which are 85.7% and 87.5%, respectively (Table S6). According to previous studies (4), LNM and UICC8 tumor stages, which are

independent prognostic factors in most solid tumors, are closely correlated with tumor recurrence and, to some extent, are used as predictive factors for R/M status in the clinical setting. Therefore, to evaluate the superiority of our R/M risk scoring model, we further compared its predictive validity with the status of the LNM and UICC8 stages in the discovery cohort. As shown in Table S6, all three predictors had similar NPV. However, although the sensitivity of our R/M risk score was slightly lower than that of the UICC8 stage and LNM, it showed better PPV and specificity, suggesting that it can more specifically predict patients with ESCC to truly undergo R/M, as well as not to undergo after surgical interventions.

As indicated above, relapse always leads to a poor prognosis. Therefore, we investigated whether the R/M risk score was an independent prognostic marker in the discovery cohort. Interestingly, of all the immune proteins profiled using the DSP tool, high levels of CD44 in ESCC (CD44_PANCK) independently predicted poorer OS with a hazard ratio (HR) of 5.654 (Figure 4B, p=0.003), which was also supported by previous studies using a semiquantitative IHC staining method (17). Furthermore, R/M risk score was also an independent prognostic marker in ESCC (Figure 4B, p=0.002, HR=5.479). The 5-year OS rate of patients with high R/M risk scores (>1.493, R/M risk score^{high}) was significantly lower than that of patients with low R/M risk scores (<1.493, R/M risk score^{low}) (Figure 4C, 12% vs 70%, p=0.0118, HR=4.993). Notably, IDO1, ARG1 and STING proteins, which were included in the R/M risk scoring model, expressed in ESCC/PANCK+ and TME-L/CD45+ cells, were

not prognostic markers for patients with ESCC (Figure 4B). Together, our immune proteins-based scoring model is superior to LNM and TNM stages in predicting R/M status for patients with ESCC after radical surgery.

IDO1 expressed in TME-L accurately predicted the R/M risk for ESCC patients

To simplify the clinical translation of our model and alleviate the economic burden for patients, we further analyzed the correlation of the R/M risk score with the expression of three immune proteins, namely IDO1_CD45, STING_CD45, and ARG1_PANCK, which were used to construct the LASSO model. Surprisingly, we found that IDO1_CD45 was linearly correlated with R/M risk score (Figure 5A, R^2=0.5556, p<0.0001), whereas STING_CD45 and ARG1_PANCK were not significantly correlated (Figure 5B and C, p>0.05). Furthermore, the ROC curve constructed to evaluate the predictive validity of IDO1_CD45 for R/M risk score demonstrated an AUC of up to 0.9598 (Figure 5D, p<0.0001) when taking 708.9 as its best cut-off value. Specifically, IDO1 expressed in PANCK+ (IDO1_PANCK) could not predict R/M risk score (Figure 5E), although its expression in TC has previously been reported to be a prognostic marker for patients with ESCC (18).

Subsequently, dual IHC staining of markers in the validation cohort was used to confirm the above findings in the discovery cohort, representative data of which are shown in Figure 5F. As expected, results from the dual IHC staining showed that the expression of IDO1 in TME-L, rather than in ESCC, was significantly correlated with R/M status, while the STING protein, expressed in ESCC and TME-L, was not

associated with R/M status (Table 1). Besides, the expression of IDO1_CD45, shown as nCount data, was also significantly different in patients with or without R/M status (Figure 5G, p=0.0273). Taken together, our findings suggest that IDO1 in TME-L cells, either quantitatively detected by DSP or semi-quantitatively tested by dual IHC staining, is highly associated with R/M in ESCC.

Discussion

With advances in technologies involving spatial evaluation of tumor features, an increasing number of studies have focused on identifying transcriptional immune characteristics by combining single cell transcriptomic sequences, RNA sequences, and high-plex DSP transcriptomic sequences in different molecular stratified subgroups (19, 20). However, to the best of our knowledge, no study has reported the landscape of immune proteins in ESCC and TME-L using high-plex spatial technology. In this study, we used a newly developed DSP tool to spatially quantify a panel of 31 immune proteins, including factors participating in immune cell activation, immune checkpoints (IC) that influence the crosstalk between TC and immune cells, and immune cell markers, in ESCC and TME-L. Interestingly, our findings revealed that some ICs, such as TIM3, LAG3, and IDO1, were not differentially expressed in ESCC and TME-L. IDO1 is an intracellular enzyme catalyzing tryptophan degradation into kynurenine, and its expression has been shown to increase in the stroma of multiple cancers, inducing negative feedback for cancer evasion from immunosurveillance (21). Upregulation of LAG3, also known as CD223, is required to control overt

activation and prevent the onset of autoimmunity, which has been recognized as the third inhibitory receptor for targeted cancer immunotherapy beyond CTLA4 and PD1-PDL1 (22). Taken together, we speculate that TIM3, LAG3 and IDO1 might not participate in tumorigenesis by suppressing immune responses in ESCC.

In this study, CD40 in the CD45+ compartment was closely correlated with B2M, which participates in antigen presentation, in the CD3+/CD4+/CD8+ T lymphocyte subpopulation, suggesting that it may interact with its ligand CD40L to transform non-activated antigen-presenting cells into cells that are potent inducers of T cell immunity, thus priming tumor immunity in the TME (23). Furthermore, the expression of ICs including CTLA4, PD-L2, PD-1, CD137, B7H3, ARG1, CD137, and LAG3 in the TME-L was highly correlated with CD56, a marker of natural killer (NK) cells, and may represent a specific immune feature of exhausted NK cells in the TME (24), which may possibly mediate the immune evasion of ESCC during treatment (25). Above all, it is reasonable to speculate that these six newly identified, closely correlated clusters of immune proteins may provide a rationale for pharmaceutical companies to develop more precise immunotherapies for patients with ESCCs in the future.

The recurrence rate for patients with ESCC who undergo radical resection surgery without neoadjuvant therapies is as high as 45.8%, the distant metastasis rate of which is approximately 23.3% (26), which poses a high risk of poor survival for these patients. In this study, based on the hierarchical clusters above, we successfully established an R/M predictive scoring model with superior specificity and PPV, which

were better than those of LNM and UICC8 stages (2, 3). In particular, IDO1 in TME-L was linearly correlated with the R/M risk score and specifically predicted the risk score with an AUC of 0.9598, demonstrating that quantitative examination of IDO1 in TME-L/CD45+ is sufficient to predict the R/M risk for patients with ESCC after radical surgery. Furthermore, this finding suggests that IDO1 in TME-L potentially plays a driving role in mediating R/M in ESCC. To detect IDO1 in TME-L more easily in a clinical setting and alleviate the economic burden for patients, we developed a dual IHC staining approach and the results showed that IDO1 in CD45+ cells was significantly correlated with the R/M status in discovery and validation cohorts. Notably, the independent prediction of IDO1 in CD45+ cells for R/M status is less superior to the R/M risk score, with sensitivity and specificity of 50% and 85%, respectively. Besides, it is not an independent prognostic marker in ESCC, whereas the R/M risk score independently predicts the poor survival for ESCC patients. Taken together, we propose that the highly correlated factors IDO1_CD45, STING_CD45 and ARG1_PANCK which are used to construct the R/M risk predictive model possibly interact with each other, with IDO1_CD45 as the driver, to co-regulate the R/M process in ESCC, thus leading to poor survivals. However, further mechanism studies will facilitate our understanding of their exact roles in regulating R/M. Above all, our findings not only provide a more accurate scoring approach based on quantitative immune proteins for the prediction of R/M risk, but also propose that IDO1 in TME-L potentially plays a driving role in mediating R/M in ESCC.

Journal Pre-proof

However, our study had some limitations. First, the sample size of the discovery cohort was relatively small, and the predictive performance of the LASSO model must be further verified in larger cohorts using the same quantitative strategy. Second, we did not intensively study the clinical significance of the six closely correlated immune protein groups. More mechanistic and clinical studies are needed to broaden their applications in predicting OS and immunotherapeutic responses in ESCC. Third, the development of a quantitative evaluation of the results from dual IHC staining using a more economical strategy will facilitate the clinical translation of our findings in the future. In summary, our findings not only provide a more accurate approach based on quantitative expression of immune proteins for the prediction of R/M incidence, but also shed a novel insight into spatially distributed IDO1 in mediating R/M in ESCC.

Acknowledgements

We deeply appreciate the contribution of Professor Ning Wu from the Department of Cardiothoracic Surgery, Huashan Hospital of Fudan University, for clinical consultation and sample collection. We also sincerely thank the executive director Dr. Tianhai, Ji in our department for moral support and encouragement all the time.

Declarations

Ethics approval and consent to participate

This retrospective pilot study was approved by the Institutional Research Ethics Committee of Huashan Hospital (protocol no. KY2022-614).

Consent for publication

Not applicable in this study.

Authors' contributions

QY conceived and planned the study. CQ evaluated all IHC results. QL performed all IHC assays. SJ and CH collected all archived FFPE samples and performed follow-ups. QY, CQ, and QL obtained ethical approval and analyzed all data. QY organized and wrote the manuscript. All authors provided critical feedback and helped shape the research, analysis, and the manuscript.

Fundings

This study was financially supported by Fundamental Research Program Funding of Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine (JYZZ195). And it was also supported by The Cross Disciplinary Research Project (JYJC202106) and the Biobank Project (YBKB201911), both of which are from Shanghai Ninth People's hospital, Shanghai Jiao Tong University School of Medicine.

Data availability

All datasets are presented in the main manuscript or in additional supporting files whenever possible.

References

1. He Y, Liang D, Du L, Guo T, Liu Y, Sun X, et al. Clinical characteristics and survival of 5283 esophageal cancer patients: A multicenter study from eighteen hospitals across six regions in China. Cancer communications. 2020;40(10):531-44.

2. Zhou S, Liu S, Zhang L, Guo S, Shen J, Li Q, et al. Recurrence Risk Based on

Pathologic Stage After Neoadjuvant Chemoradiotherapy in Esophageal Squamous Cell Carcinoma: Implications for Risk-Based Postoperative Surveillance Strategies. Annals of surgical oncology. 2018;25(12):3639-46.

3. Wang W, Chen Y, Zhang X, Xi K, Feng X, Zhang L. A Model Shows Utility in Predicting Postoperative Recurrence and Distant Metastasis in Curatively Resected Esophageal Squamous Cell Cancer. Cancer control : journal of the Moffitt Cancer Center. 2019;26(1):1073274819852965.

4. Duan J, Xie Y, Qu L, Wang L, Zhou S, Wang Y, et al. A nomogram-based immunoprofile predicts overall survival for previously untreated patients with esophageal squamous cell carcinoma after esophagectomy. Journal for immunotherapy of cancer. 2018;6(1):100.

5. Galon J, Costes A, Sanchez-Cabo F, Kirilovsky A, Mlecnik B, Lagorce-Pages C, et al. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. Science. 2006;313(5795):1960-4.

6. Merritt CR, Ong GT, Church SE, Barker K, Danaher P, Geiss G, et al. Multiplex digital spatial profiling of proteins and RNA in fixed tissue. Nature biotechnology. 2020;38(5):586-99.

7. Kulasinghe A, Monkman J, Shah ET, Matigian N, Adams MN, O'Byrne K. Spatial Profiling Identifies Prognostic Features of Response to Adjuvant Therapy in Triple Negative Breast Cancer (TNBC). Frontiers in oncology. 2021;11:798296.

8. Monkman J, Taheri T, Ebrahimi Warkiani M, O'Leary C, Ladwa R, Richard D, et al.

High-Plex and High-Throughput Digital Spatial Profiling of Non-Small-Cell Lung Cancer (NSCLC). Cancers. 2020;12(12).

9. Maynard A, McCoach CE, Rotow JK, Harris L, Haderk F, Kerr DL, et al. Therapy-Induced Evolution of Human Lung Cancer Revealed by Single-Cell RNA Sequencing. Cell. 2020;182(5):1232-51 e22.

10. van Smeden M, Moons KG, de Groot JA, Collins GS, Altman DG, Eijkemans MJ, et al. Sample size for binary logistic prediction models: Beyond events per variable criteria. Statistical methods in medical research. 2019;28(8):2455-74.

11. Bertero L, Massa F, Metovic J, Zanetti R, Castellano I, Ricardi U, et al. Eighth Edition of the UICC Classification of Malignant Tumours: an overview of the changes in the pathological TNM classification criteria-What has changed and why? Virchows Archiv : an international journal of pathology. 2018;472(4):519-31.

12. Yongtao WCH. Interpretation of the 2020 Guidelines for the Dagnosis and Treatment of Esophageal Cancer by Chinese Society of Clinical Oncology. J Cancer Control Treat. 2020;33(4):285-90.

13. Haase S, Banerjee K, Mujeeb AA, Hartlage CS, Nunez FM, Nunez FJ, et al. H3.3-G34 mutations impair DNA repair and promote cGAS/STING-mediated immune responses in pediatric high-grade glioma models. The Journal of clinical investigation. 2022;132(22).

14. Nakajima S, Mimura K, Kaneta A, Saito K, Katagata M, Okayama H, et al. Radiation-Induced Remodeling of the Tumor Microenvironment Through Tumor

Cell-Intrinsic Expression of cGAS-STING in Esophageal Squamous Cell Carcinoma. International journal of radiation oncology, biology, physics. 2023;115(4):957-71.

15. Hopfner KP, Hornung V. Molecular mechanisms and cellular functions of cGAS-STING signalling. Nature reviews Molecular cell biology. 2020;21(9):501-21.

16. Graham Martinez C, Barella Y, Kus Ozturk S, Ansems M, Gorris MAJ, van Vliet S, et al. The immune microenvironment landscape shows treatment-specific differences in rectal cancer patients. Frontiers in immunology. 2022;13:1011498.

17. Okamoto K, Ninomiya I, Ohbatake Y, Hirose A, Tsukada T, Nakanuma S, et al. Expression status of CD44 and CD133 as a prognostic marker in esophageal squamous cell carcinoma treated with neoadjuvant chemotherapy followed by radical esophagectomy. Oncology reports. 2016;36(6):3333-42.

18. Zhou S, Yang H, Zhang J, Wang J, Liang Z, Liu S, et al. Changes in Indoleamine 2,3-Dioxygenase 1 Expression and CD8+ Tumor-Infiltrating Lymphocytes after Neoadjuvant Chemoradiation Therapy and Prognostic Significance in Esophageal Squamous Cell Carcinoma. International journal of radiation oncology, biology, physics. 2020;108(1):286-94.

19. Carter JM, Polley MC, Leon-Ferre RA, Sinnwell J, Thompson KJ, Wang X, et al. Characteristics and Spatially Defined Immune (micro)landscapes of Early-stage PD-L1-positive Triple-negative Breast Cancer. Clinical cancer research : an official journal of the American Association for Cancer Research. 2021;27(20):5628-37.

20. Wu SZ, Al-Eryani G, Roden DL, Junankar S, Harvey K, Andersson A, et al. A

single-cell and spatially resolved atlas of human breast cancers. Nature genetics. 2021;53(9):1334-47.

21. Charehjoo A, Majidpoor J, Mortezaee K. Indoleamine 2,3-dioxygenase 1 in circumventing checkpoint inhibitor responses: Updated. International immunopharmacology. 2023;118:110032.

22. Andrews LP, Marciscano AE, Drake CG, Vignali DA. LAG3 (CD223) as a cancer immunotherapy target. Immunological reviews. 2017;276(1):80-96.

23. Toes RE, Schoenberger SP, van der Voort EI, Offringa R, Melief CJ. CD40-CD40Ligand interactions and their role in cytotoxic T lymphocyte priming and anti-tumor immunity. Seminars in immunology. 1998;10(6):443-8.

24. Kurachi M. CD8(+) T cell exhaustion. Seminars in immunopathology. 2019;41(3):327-37.

25. Khan M, Arooj S, Wang H. NK Cell-Based Immune Checkpoint Inhibition. Frontiers in immunology. 2020;11:167.

26. Liu S, Wen J, Yang H, Li Q, Chen Y, Zhu C, et al. Recurrence patterns after neoadjuvant chemoradiotherapy compared with surgery alone in oesophageal squamous cell carcinoma: results from the multicenter phase III trial NEOCRTEC5010. European journal of cancer. 2020;138:113-21.

Journal Pre-proof

Table 1. Association of R/M status with clinicopathological factors and indicated immune proteins in ESCC tissues from discovery and validation cohorts.

	Discovery cohort		Chi-square test p-value	Validation cohort		Chi-square test p-value
	R/M status^{no}	R/M status^{yes}		R/M status^{no}	R/M status^{yes}	
N (%)	**25 (69.4%)**	**11 (30.6%)**		**41 (70.7%)**	**17 (29.3%)**	
Sex						
Male	17 (68.0%)	8 (72.7%))		35 (85.4%)	14 (82.4%)	
Female	8 (32.0%)	3 (27.3%)	1.000	6 (14.6%)	3 (17.6%)	1.000
Lesion site						
Upper	5 (20.0%)	5 (45.5%)		9 (22.0%)	0 (0.0%)	
Middle	16 (64.0%)	4 (36.4%)		20 (48.8%)	12 (70.6%)	
Lower	4 (16.0%)	2 (18.2%)	0.240	12 (29.3%)	5 (29.4%)	0.092
Diameter						
≤2cm	10 (40.0%)	3 (37.5%)		11 (26.8%)	2 (11.8%)	
>2cm	15 (60.0%)	5 (62.5%)	1.000	30 (73.2%)	15 (88.2%)	0.307
Differentiation						
Well/Moderate	17 (68.0%)	9 (81.8%)		31 (75.6%)	14 (82.4%)	
Poor	8 (32.0%)	2 (18.2%)	0.688	10 (24.4%)	3 (17.6%)	0.736
LNM_status						
No	18 (72.0%)	4 (36.4%)		26 (63.4%)	7 (41.2%)	
Yes	7 (28.0%)	7 (63.6%)	0.067	15 (36.6%)	10 (58.8%)	0.120
UICC8_stage						
I/II	19 (76.0%)	3(27.3%)		29 (70.7%)	8 (47.1%)	
III/IV	6 (24.0%)	8(72.7%)	0.010	12 (29.3%)	9 (52.9%)	0.088
STING_PANCK						
Low	18 (72.0%)	7 (63.6%)		22 (53.7%)	9 (56.3%)	
High	7 (28.0%)	4 (36.4%)	0.703	19 (46.3%)	7 (43.8%)	0.860
STING_CD45						
Low	13 (52.0%)	6 (54.5%)		11 (26.8%)	6 (37.5%)	
High	12 (48.0%)	5 (45.5%)	0.888	30 (73.2%)	10 (62.5%)	0.523
IDO1_PANCK						
Low	16 (69.6%)	6 (60.0%)		34 (85.0%)	13 (86.7%)	
High	7 (30.4%)	4 (40.0%)	0.696	6 (15.0%)	2 (13.3%)	1.000
IDO1_CD45						
Low	13 (59.1%)	1 (10.0%)		32 (80%)	7 (46.7%)	

Journal Pre-proof

Journal Pre-proof

High	9 (40.9%)	9 (90.0%)	0.019	8 (20.0%)	8 (53.3%)	0.022
PD-L1						
Negative	22 (88.0%)	9 (81.8%)		35 (94.6%)	15 (93.8%)	
Positive	3 (12.0%)	2 (18.2%)	0.631	2 (5.4%)	1 (6.3%)	1.000

Journal Pre-proof

Figures

Figure 1. Workflow chart of the study.

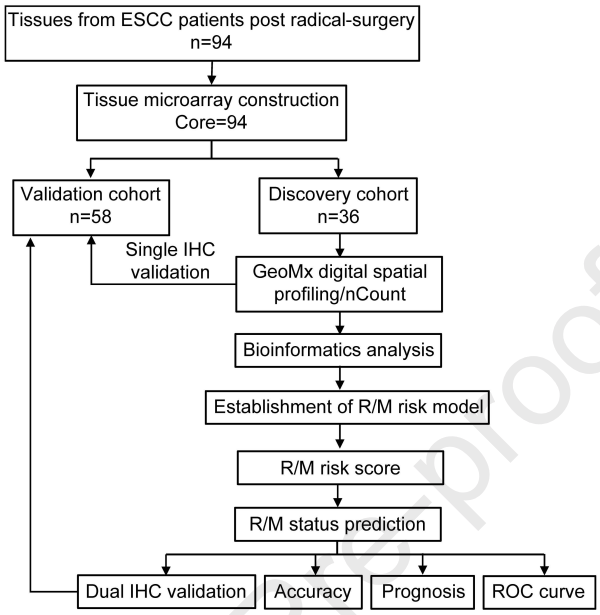
Figure 2. Expression of immune proteins is significantly higher in CD45+ compartments than that in PANCK+ compartments of ESCC tissues. A. Representative ROIs in paired PANCK+ compartment (green) and CD45+ compartment (pink) from the same specimen selected for the GeoMx DSP assay. The nuclei were stained with SYTO13 (blue). **B.** Differential expression of immune proteins in paired PANCK+ compartment and CD45+ compartment of ESCC tissues in the discovery cohort. Data are shown as mean ± SD and statistically analyzed using two-sided Student's t-test. **C.** The bar plot showing the differential expression of immune proteins indicated as IHC scores in the validation cohort, which are shown as mean ± SD and statistically analyzed using two-sided Student's t-test. Statistical significance was shown as $p<0.05$ (*). ns, non-significant. **D.** Representative single IHC staining of indicated proteins in the validation cohort. Scale bars, 100 μm and 250 μm. H&E, hematoxylin and eosin stain.

Figure 3. A R/M risk scoring model is successfully established based on three immune proteins using the LASSO regression analysis. A. The correlation matrix showing the mutual relationships among immune proteins expressed in PANCK+ compartment (green) and CD45+ compartment (pink) of ESCC tissues in the discovery cohort. Corr_r, correlation coefficient. **B.** The tuning parameter (lamda) selected to establish the best R/M risk predictive scoring model.

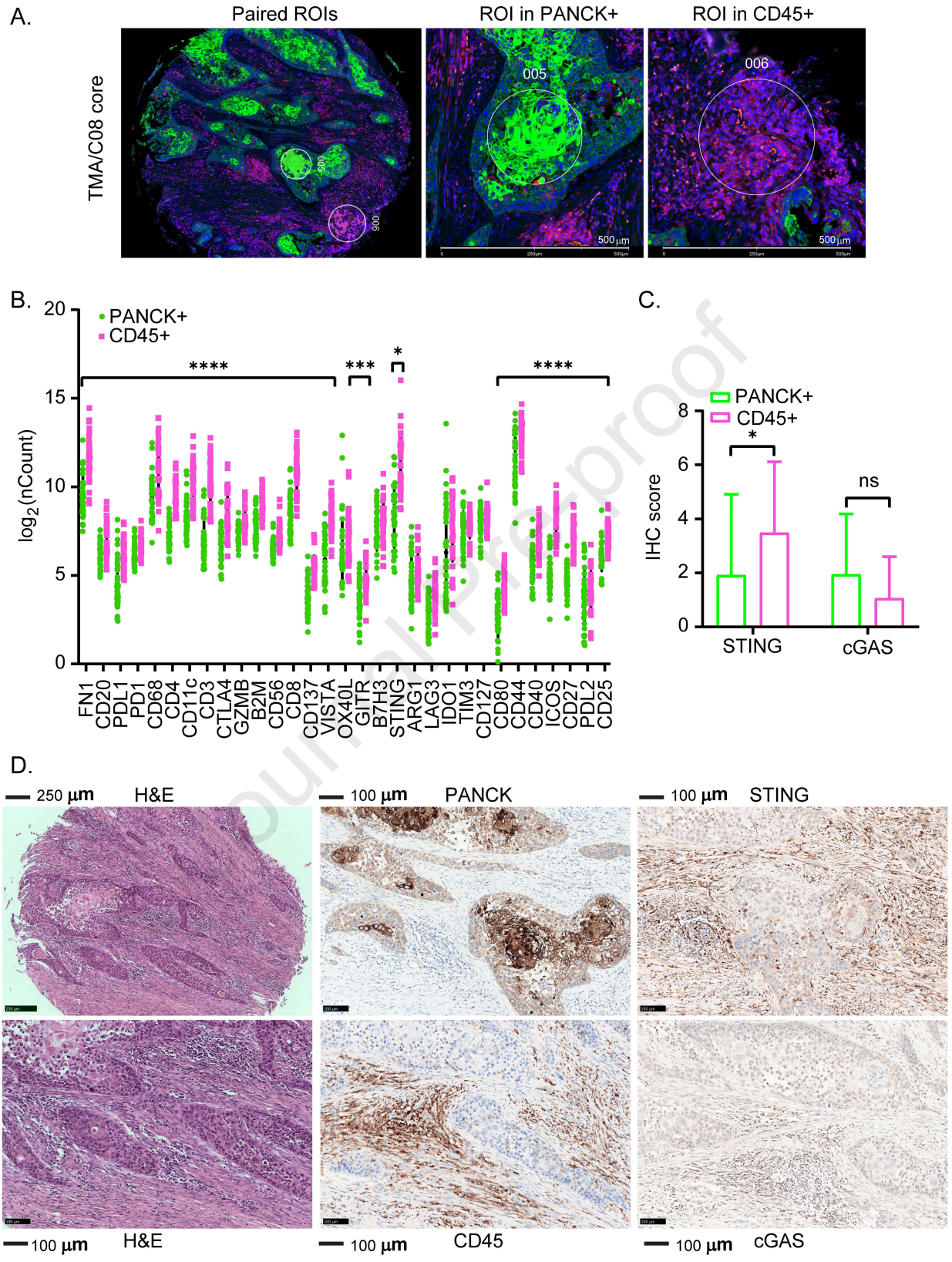
Figure 4. The R/M risk score specifically predicts the R/M status for ESCC patients in the discovery cohort. A. The ROC curves of R/M risk score predictive of R/M status with an AUC of 0.7455. **B.** The forest plot showing the univariate and multivariate Cox regression survival analysis used to identify the independent prognostic markers in the discovery cohort. **C.** The Kaplan-Meier survival curves showing the differential survival ratio between patients with Risk score^{low} (<1.493) and Risk score^{high} (>1.493). CI, confidence interval, HR, Hazard ratio.

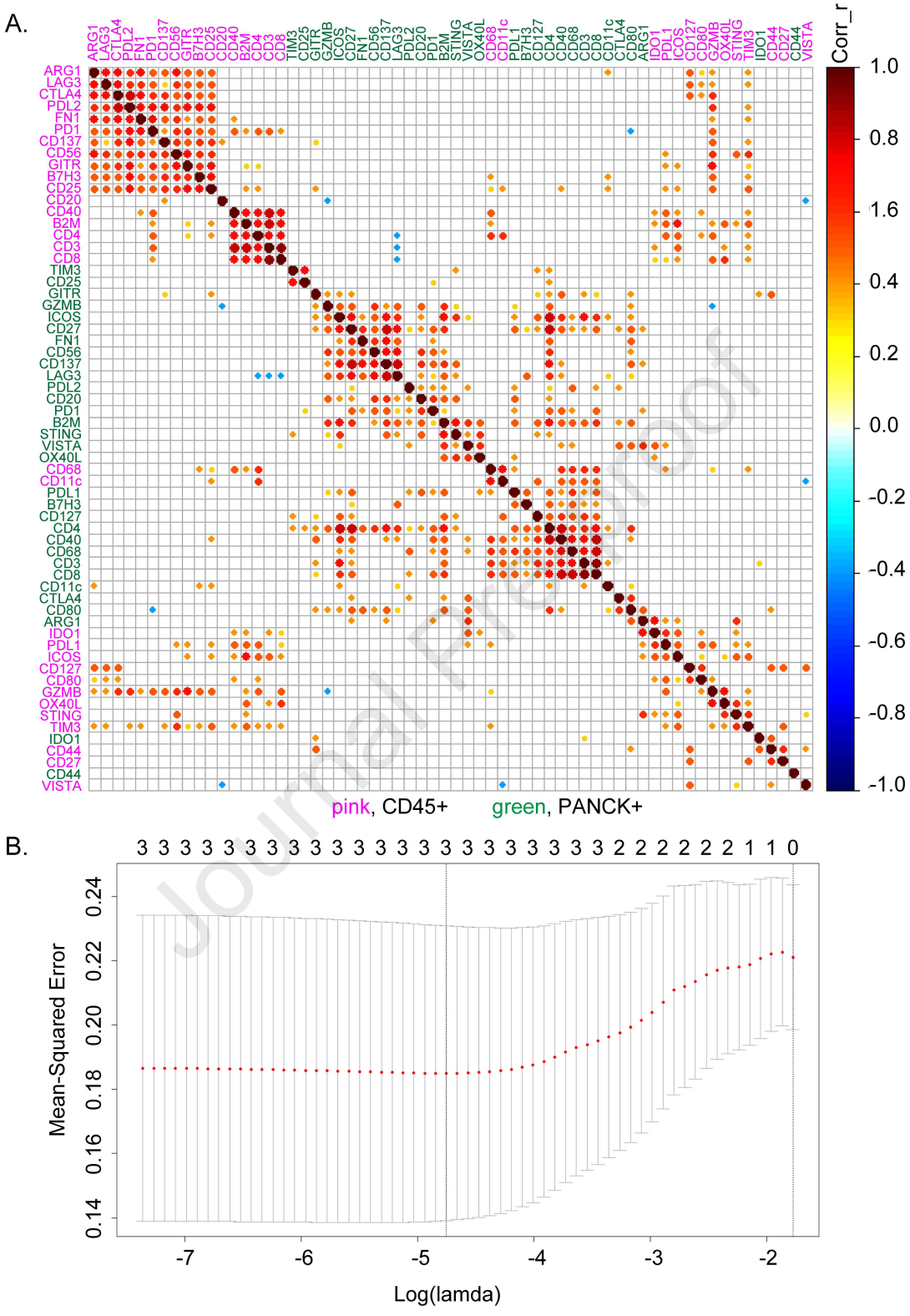
Figure 5. The expression of IDO1 in CD45+ compartment is highly correlated with R/M status in ESCC. A-C. The linear correlation analysis of R/M risk score with protein levels, shown as nCount data, of IDO1 in CD45+ compartment (IDO1_CD45, A) and STING in CD45+ compartment (STING_CD45, B), ARG1 in the PANCK+ compartment (ARG1_PANCK, C) in the discovery cohort. **D-E.** The ROC curves of IDO1_CD45 (D) and IDO1_PANCK (E) expression, shown as nCount data, predictive of R/M risk score in the discovery cohort. **F.** Representative images of dual IHC staining of indicated proteins in ESCC tissues from patients with (R/M status^{yes}) or without (R/M status^{no}) R/M in the validation cohort. Scale bar, 100 μm. **G.** The bar plot showing the differential expression of IDO1_CD45 and STING_CD45, indicated as nCount data, in ESCC patients with or without R/M.

Journal Pre-proof



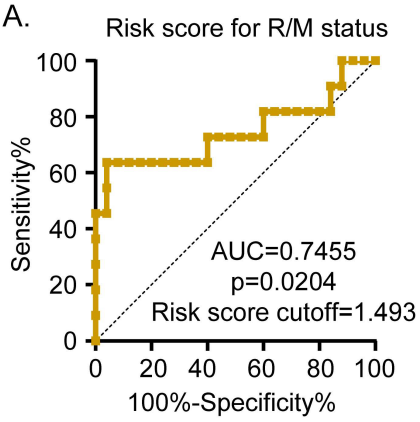
Journal Pre-proof


A.
Paired ROIs
ROI in PANCK+
ROI in CD45+
TMA/C08 core
005
006
500 µm
B.
PANCK+
CD45+
****
***
*
****
log2(nCount)
FN1
CD20
PDL1
PD1
CD68
CD4
CD11c
CD3
CTLA4
GZMB
B2M
CD56
CD8
CD137
VISTA
OX40L
GITR
B7H3
STING
ARG1
LAG3
IDO1
TIM3
CD127
CD80
CD44
CD40
ICOS
CD27
PDL2
CD25
C.
PANCK+
CD45+
IHC score
*
ns
STING
cGAS
D.
250 µm H&E
100 µm PANCK
100 µm STING
100 µm H&E
100 µm CD45
100 µm cGAS

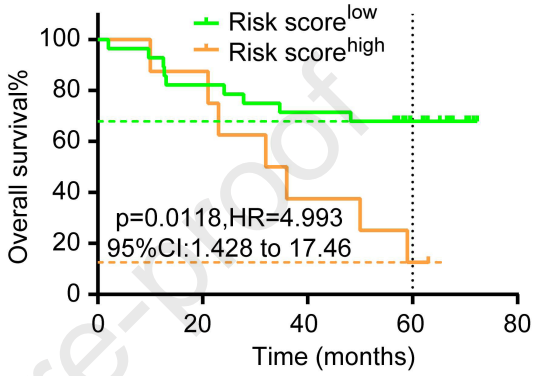


Journal Pre-proof

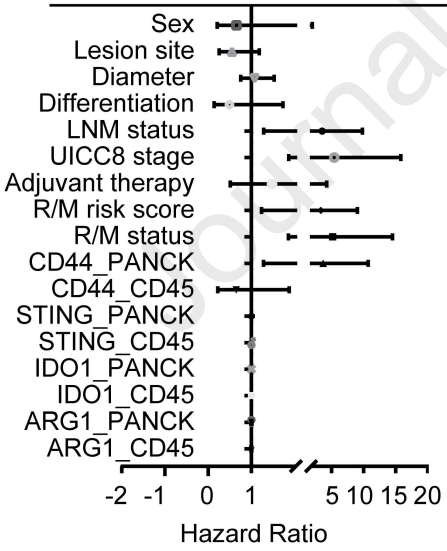
A.



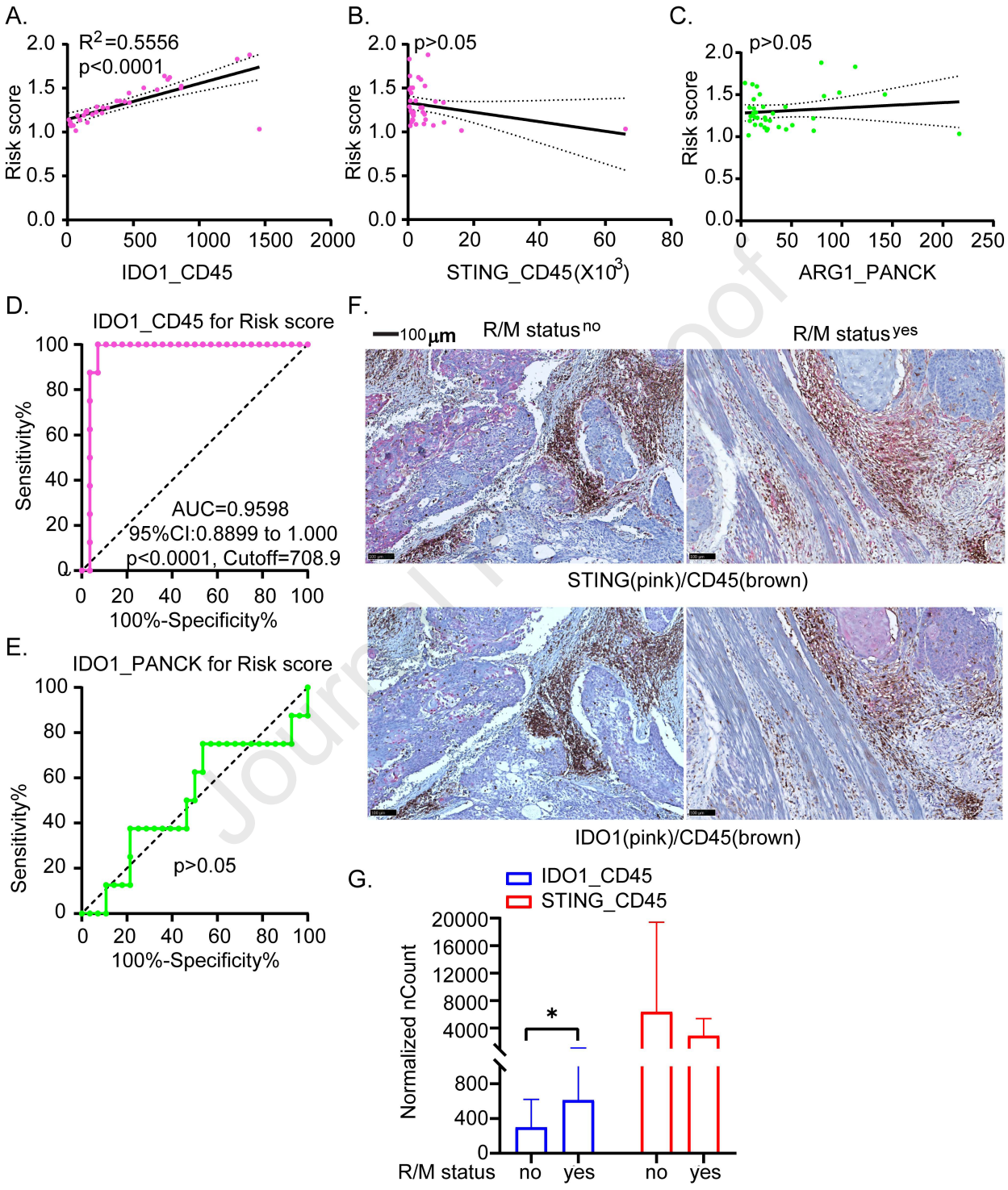
C.



B.



	Univariate Cox analysis p-value	Multivariate Cox analysis p-value
Sex	0.459	
Lesion site	0.126	
Diameter	0.698	
Differentiation	0.267	
LNM status	0.015	
UICC8 stage	0.002	
Adjuvant therapy	0.470	
R/M risk score	0.018	0.002
R/M status	0.002	
CD44_PANCK	0.016	0.003
CD44_CD45	0.423	
STING_PANCK	0.369	
STING_CD45	0.584	
IDO1_PANCK	0.585	
IDO1_CD45	0.440	
ARG1_PANCK	0.716	
ARG1_CD45	0.924	


A.
R2=0.5556
p<0.0001
Risk score
IDO1_CD45
B.
p>0.05
Risk score
STING_CD45(X10^3)
C.
p>0.05
Risk score
ARG1_PANCK
D.
IDO1_CD45 for Risk score
Sensitivity%
100%-Specificity%
AUC=0.9598
95%CI:0.8899 to 1.000
p<0.0001, Cutoff=708.9
E.
IDO1_PANCK for Risk score
Sensitivity%
100%-Specificity%
p>0.05
F.
100μm
R/M status no
R/M status yes
STING(pink)/CD45(brown)
IDO1(pink)/CD45(brown)
G.
IDO1_CD45
STING_CD45
Normalized nCount
R/M status
no
yes
no
yes
*