The characteristic of epithelial-specific phenotypes and immunosuppressive microenvironment in the context of tumour budding in colorectal cancer

Phimmada Hatthakarnkul [a,b], Natalie Fisher [c], Leonor Schubert Santana [b], Holly Leslie [b], Assya Legrini [b], Aula Ammar [b], Amna Ahmed Mohemmed Malty [b], Kathryn A. F. Pennel [b], Ian Powley [d], Courtney Bull [c], Peter G. Alexander [e], Janos Sztanko [b], Jean A Quinn [b], Jennifer Hay [f], Hannah Morgan [f], Claire Kennedy Dietrich [b], Yoana Doncheva [b], Gerard Lynch [b], Noori Maka [e], Campbell Roxburgh [b], Hester C. van Wyk [e], Xiao Fu [d], Donald C. McMillan [e], Philip Dunne [c,d], Nigel Jamieson [b], Chanitra Thuwajit [g], Joanne Edwards [b]

[a] Biomedical Sciences Program, Faculty of Medicine Siriraj Hospital, Bangkok, Thailand

[b] School of Cancer Sciences, University of Glasgow, Wolfson Wohl Cancer Research Centre, Garscube Estate, Glasgow, United Kingdom.

[c] The Patrick G Johnston Centre for Cancer Research, Queen's University Belfast, Belfast, UK.

[d] CRUK Scotland Institute, Glasgow, United Kingdom.

[e] School of Medicine, University of Glasgow, Glasgow Royal Infirmary, Alexandria Parade, Glasgow, United Kingdom.

[f] Glasgow Tissue Research Facility, University of Glasgow, Queen Elizabeth University Hospital, Glasgow, United Kingdom.

[g] Department of Immunology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand.

Running title: Tumour budding, Tumour microenvironment, Colorectal Cancer

Corresponding Author: Phimmada Hatthakarnkul, School of Cancer Sciences, University of Glasgow, Wolfson Wohl Cancer Research Centre, Glasgow, United Kingdom. phimmada.hatthakarnkul@glasgow.ac.uk. (ORCID https://orcid.org/0000-0002-5274-5480)

Abstract

Background: A small cluster of up to four cells, defined as tumour budding (TB), has been well studied as an independent promising prognostic marker in CRC. However, its underlying mechanism in CRC is unclear.

Methods: Multi-omic approaches from bulk RNA as well as regional GeoMx bulk RNA were used to identify the underlying mechanism of TB and its possible correlation with tumour microenvironment in CRC tissue. The results were validated using immunohistochemistry (IHC) and multiplex immunofluorescence (mIF) staining.

Results: Patients with high TB have a worse outcome and associated with adverse clinical factors. Bulk RNA expression revealed that tumours with high TB are significantly enriched for gene signatures related to epithelial-mesenchymal transition (EMT), TNF-α and TGF-β, in both independent CRC cohorts. Additionally, regional bulk RNA expression from non-tumour (PanCK-) invasive areas showed that tumours with high TB were significantly associated with a decreased immune activity. Multiplex immunofluorescence (mIF) suggested a close proximity of regulatory T cells and pan-macrophages to individual buds.

Discussion: This study confirms the prognostic role of TB in CRC in two independent CRC cohorts and gives an insight into its underlying mechanism and potential interactions with the tumour microenvironment in CRC.

Introduction

Colorectal cancer (CRC) is the third most diagnosed cancer and the second most lethal malignancy worldwide (1). Incidence rates of CRC have been decreasing due to the development of screening programs (2). However, 30% of CRC patients develop synchronous or metachronous metastasis (3). Of these, fewer than 20% with metastatic CRC are alive 5 years after diagnosis (4). Thus, one of the main challenges in cancer research is to develop targeted treatments to prevent metastasis and subsequently increase the survival rate of those with CRC (5).

Tumour budding (TB) is a single cell or small cluster of up to 4 cells found at the invasive front of tumours (6) and has been consistently reported as a prognostic factor associated with tumour recurrence and metastases across solid tumours (7-9). TB has been reported as an independent prognostic marker associated with adverse clinicopathological factors such as lymphatic, venous invasion, and disease recurrence (10-12). Recently, using the ITBCC criteria for TB assessment, the prognostic and predictive values of TB have been reported in stage II CRC from both SACURA (13) and QUASAR (14) trials.

Although the potential prognostic value of TB in CRC has been widely reported, the molecular mechanism relating to budding is poorly understood (9). Recently, we published a literature review detailing the potential underlying mechanisms of TB in CRC as well as a systematic review of how TB may be associated with common CRC mutations (15, 16). Additionally, the role of inflammation in CRC has been repeatedly reported as part of the potential prognostic biomarker in CRC (17-19). The crosstalk between the tumour cells and the immune microenvironment has shown to manipulate the environment for cancer cells to thrive and survive (20, 21). Studies showed that a high number of TB could compromise immune

activity leading to invasion and metastasis in CRC (15, 22). Therefore, in this present study, the underlying mechanism of TB is investigated using bulk and regional spatial transcriptomics as well as the relationship with the tumour microenvironment to understand the biology of TB in CRC.

Methods and Materials

Patient Cohort

Discovery cohort

A cohort of stage I-III CRC patients (n=787) who underwent resection for primary operable CRC in the Glasgow Royal Infirmary (GRI, Glasgow, UK) between 1997 and 2013 was used as the discovery cohort. Patients who died within 30 days of surgery, had emergency operation, or received neoadjuvant therapy were excluded (Figure 1A). Tumour staging was carried out using 5th Edition of the AJCC/UICC-TNM staging system. Clinicopathological data were collected with a minimum of 5 years follow-up years post-resection. The study was approved by the West of Scotland Research Ethics Committee (REC 4: 22/ws/0207) and data are stored in Greater Glasgow and Clyde Safehaven (SH21ON012).

Validation cohort

A publicly available cohort of stage II-III Colon Cancer (CC) patients (n=85), who underwent curative surgery at the Department of Surgery, National Defense Medical College Hospital (Saitama, Japan) between 2002-2003 was used as the validation cohort. The cohort was downloaded from GEO under accession number GSE143985 (23). The data was obtained and processed by NF and CB.

Tumour budding assessment

Following the guidelines from the International Tumor Budding Consensus Conference (ITBCC) 2016 (6), budding scores were assessed in both discovery and validation cohorts. Full CRC sections were stained with Haematoxylin and Eosin (H&E). Ten specific areas (0.785 mm^2) at the invasive tumour front were marked and the number of TBs were determined at 20x magnification. Budding status in the discovery cohort was determined by HW and PH. In validation cohort, TB status was previously assessed, and data was obtained and processed by NF. The densest area of buds represented the budding profile of each CRC case. Low and high TB phenotypes were based on the number of TB (<10 as low and ≥10 as high) found at the invasive area (Figure 1B).

Whole Transcriptomic Assay (TempO-Seq™)

Of 787 patients from discovery cohort, 764 single 4µm thickness tissue sections from CRC resections were used for Templated Oligo-Sequencing (24, 25). (See Supplementary Method 1)

Regional bulk transcriptomic (GeoMx Digital Spatial Profiler; DSP)

Whole CRC tissue sections from the discovery cohort were selected (n=12, 6 low and 6 high TB) and GeoMx RNA Immune Oncology Pathways panel (84 targets) used to identify the gene expression within regions of interest (ROIs) proximal and distal to the invasive front of either tumour-rich (PanCK+) or non-tumour (PanCK-) regions (i.e. stroma-rich region) (Supplementary Figure S1). Additionally, a tissue microarray (TMA), previously made from the invasive budding area with low (n=20) and high (n=23) TB samples were used for the Whole transcriptomic atlas (WTA) GeoMx RNA panels (18,677 targets). Target gene list from both panels are shown in Supplementary Table S1. Regional bulk RNA experiment workflow was illustrated in Supplementary Figure S2. (See Supplementary Method 2)

Immunohistochemistry (IHC)

IHC was performed on a TMA from the invasive high budding area CRC tissue of a discovery cohort (n=180). The weighted histoscore was used to access nuclear cyclinD1 protein expression. (See Supplementary Method 3)

Multiplex immunofluorescence (mIF)

Full CRC sections (n=18) from the discovery cohort were used to perform multiplex immunofluorescence (mIF) with the optimised lymphocyte panel (CD3 CD8 FOXP3 KI67 and PanCK) and myeloid markers (CD68 CD163 CD66b and PanCK). EMT markers were stained in the invasive budding TMA (n=180) (AM) (Supplementary Table S2). The slides were stained using the autostainer (Thermofisher), for full CRC sections and Leica bond for TMA, and scanned by NanoZoomer S60 digital slide scanner (Hamamatsu, USA) with 20x magnification. Cell density was then quantified using Visiopharm (*v2021.02.5.10297*), a digital precision pathology software (See Supplementary Method 4).

CRC spheroids culture

HT-29 CRC cell line was cultured and embedded in 50% Matrigel diluted in DMEM media (Corning, 356231) to form 3D spheroids. The spheroids were grown on coverslips in DMEM media supplemented with 10% fetal bovine serum until an optimal size was achieved. Spheroids were pre-treated with media supplemented with TNF-α (R&D system (aa 80-235), 20ng/ml) for 24 hours followed by media supplemented with both TNF-α (20ng/ml) and TGF-β (R&D system (7666-MB), 10ng/ml) for 24 hours. After that, immunofluorescence (IF) staining was performed. (See Supplementary Method 5)

Statistical analysis

In the discovery cohort, the statistical analysis was performed in IBM SPSS Statistic *(v27.0.)*. Pearson's $\chi 2$ test assessed the relationship between TB status and clinicopathological features. The likelihood ratio test was used when required. Patient survival was assessed by Kaplan-Meier analysis and log rank to test the significance. Univariate and multivariate Cox hazard regression was performed to estimate the hazard ratio (HR) for cancer-specific survival (CSS) and disease-free survival (DFS) to identify any significant prognostic factors in CRC patients. *maxstat (v0.7-25)* and *survminer (v0.4.9)* packages were utilised to determine optimal thresholds for weighted histoscores and counts from IHC and mIF staining.

In the validation cohort, disease free survival analysis was conducted using the *survminer* (*v0.4.9*) and *survival* (*v3.5-3*) R packages. Univariate and multivariate Cox hazard regression was performed to estimate the hazard ratio (HR) for disease-free survival (DFS). This were calculated for budding status with the *MASS* (*v7.3-58.2*) and *finalfi*t (*v1.0.8*) R packages. Correlation analysis was conducted in the *stats* (*v4.2.3*) and visualised within the *corrplot* R package (*v0.92*).

Data Availability

Bulk transcriptomics data of the discovery cohort are deposited with BioProject accession prjna997336. Regional bulk transcriptomic data are available upon request from the corresponding authors.

Results

Clinicopathological parameter of colorectal cancer patient's cohort

In the discovery cohort, after exclusion criteria applied, there were 644 patients included in the analysis. 205 (32%) patients were under 65, 208 (32%) were between 65-74 and 231 (36%) were over 75 years of age, with 353 males and 291 females. 471 (73%) and 173 (27%)

patients were diagnosed with colon and rectal cancer respectively. Cancer-specific survival (CSS) was employed as a primary endpoint for patients' survival in months from the date of surgery until recurrence or cancer-caused mortality. Additionally, disease-free survival (DFS) was employed as a primary endpoint for patients' survival in months without disease after diagnosis or treatment. Clinical characteristics were collected to further investigate the association with patients' survival or other biological markers in CRC (Supplementary Table S3). Of 644 cases, TB could not be identified in 25 cases, therefore 619 patients were included in the final analysis of TB in the discovery cohort as shown in consort diagram (Figure 1A).

The validation cohort was comprised of 85 CC patients with DFS employed as a primary endpoint for patients' survival in months without disease after diagnosis or treatment. The mean age at diagnosis was 66.5 years. Clinical characteristics were collected to further investigate the association with patients' survival or other biological markers in CRC (Supplementary Table S4).

The prognostic role of TB in CRC patients.

When the discovery cohort was dichotomised into low and high TB groups, patients with high TB phenotype (n=180) were associated with poorer CSS (HR=1.934, 95% CI: 1.408-2.658, p<0.001) (Figure 1C). as assessed by Kaplan-Meier curves and log-rank analysis compared to those with low TB (n=439). Life tables demonstrated that 60% of patients with high budding versus 75% of patients with low budding were alive 10 years after initial diagnosis. Additionally, using Pearson's chi-square test, there was a significant positive correlation between high TB and the clinical characteristics of CRC patients such as TNM (TNMIII, p=0.003), T stage (T4, p=0.004), N stage (N2, p=0.002), local recurrence (Yes, p=0.006), distant recurrence (Yes, p=0.028), tumour perforation (positive, p=0.046), venous invasion (positive, p=0.006), peritoneal

involvement (positive, p=0.002), perineural invasion (positive, p=0.012), and a negative association with local inflammatory scores such as Glasgow microenvironment (GMS), a combined score between Klintrup–Mäkinen (KM) and Tumour stromal percentage (TSP) (GMS0, p=0.046). (Figure 1D) (Supplementary Table S5). Interestingly, when stratified based on adverse clinical factors, patients with high TB showed a poorer survival such as those with left-sided colon; p=0.007 and rectal; p=0.001, MMR status (pMMR; p<0.001), later disease staged (TNM III; p=0.010, t3; p=0.010, n1; p=0.018), negative margin involvement; p=0.001, positive venous invasion; p<0.001, distance recurrence; p=0.014, systemic (mGPS2; p=0.037) and local inflammation (GMS2; p=0.011) (Figure 1E-M). In addition, multivariate analysis revealed that age, MMR status, recurrence status, n stage, margin involvement and TB were independent prognostic factors for CRC CSS survival (p=0.006, p=0.032, p<0.001, p<0.001, p<0.001 and p=0.040 respectively) (Table 1). Similarly, patients with high TB (n=431) showed the worst DFS than those with low TB (n=179) (Supplementary Figure S3). Moreover, when subset the cohort into disease type of colon and rectal cancer, tumours with high TB showed the worst CSS and DFS compared to low TB in both diseases suggesting the strong prognostic value of TB in CRC patients (Supplementary Figure S4).

Consistent with the discovery cohort, patients with high TB (n=26) in the validation cohort also had worse DFS outcomes compared to those with low TB (n=59) (HR=11.77, 95% CI;2.54-54.51, p<0.001) (Supplementary Figure S5A). Life tables demonstrated that 65% of patients with high budding versus 96% of patients with low budding had disease free survival 10 years after initial diagnosis. There is a significant positive correlation between high TB and clinical characteristics such as higher TNM stage (stage III, p=0.002) (Supplementary Figure S5B) (Supplementary Table S6). This is consistent with the discovery cohort as higher stage is

significantly associated with high TB. Moreover, when stratified based on adverse clinical factors, patients with high TB were shown to have poorer survival such as those with stage III CRC (p=0.0011) and MSS tumours (p<0.001) (Supplementary Figure S5C-D). In concordance with the discovery cohort, TB was an independent prognostic factor for DFS when entered into Cox regression analysis (Supplementary Table S7).

Bulk RNA transcriptomic analysis

Tissues from CRC tumours were sequenced and transcriptomic data obtained to determine the biological differences between tumours with low and high TB in discovery (n=318) and validation (n=85) cohorts. In the discovery cohort, 9 cases were excluded as budding status was not reported, leaving 309 cases for analysis (Figure 2A). Using the hallmark msigdb database (H), tumours with high TB were enriched for genes related to EMT (padj <0.001, NES=2.46 and padj <0.001, NES=3.06), TGF-β signalling (padj =0.006, NES =1.77 and padj =0.056 NES =1.47), TNF-α signalling via NF-κB (padj = 0.01, NES=1.50 and padj <0.001 NES=1.64) signature for both the discovery and validation cohort respectively (Figure 2B-E). Additionally, a curated gene set (C2), representing canonical pathway and sets corresponding to chemical and genetic perturbation (26), was applied and it was found that tumours with high TB were enriched with genes related to advance gastric cancer (padj <0.001, NES =3.24 and padj<0.001 NES =2.92), invasiveness (padj <0.001, NES = 3.05 and padj<0.001 NES= 3.06), up-regulated of TGF-β/EMT (padj <0.001, NES = 3.00 and padj<0.001 NES= 2.73) signatures for discovery and validation cohort respectively (Figure 2F-I).

CyclinD1 gene expression at the tumour invasive area (PanCK+) and its role in budding formation

To refine our understanding of the molecular mechanism of the budding phenotype, we examined gene expression at the tumour invasive front using the DSP (Figure 3A). Tumour invasive areas (PanCK+) within full tissue sections were selected to compare gene expression between low (n=6) and high (n=6) budding areas at the invasive front (Figure 3B). There was no significant difference in gene expression between the two groups. However, *CCND1*, which encodes for cyclinD1 controlling cell cycle activity, was overexpressed in high budding compared to low budding groups (Figure 3C-D).

To increase sample size and statistical power of our analysis, as well as validate the results at the protein level, cyclinD1 expression was investigated by IHC in the previously made TMA from the invasive front of the high budding areas (n=180) (Figure 3E). The results showed no difference in patient survival when cyclinD1 was assessed within the tumour invasive area (Supplementary Figure S6). However, when cyclinD1 expression was limited to within budding cells only, patients with high cyclinD1 (n=61) are likely to have poorer survival when compared to those with low cyclinD1 (n=88) in budding cells (HR=1.516, 95% CI;0.893-2.573, p=0.123) (Figure 3F), though no significant found. The survival rate after 10 years diagnosis is 51% and 65% for high and low cyclinD1 in budding cells respectively.

From bulk RNA transcriptomic data, the two key gene signatures related to high TB tumours were TGF-β/EMT and TNF-α. These results were observed in both the discovery and validation cohorts (Figure 2B-E). To further investigate the expression of cyclinD1 and the formation of TB *in vitro*, known high TB model HT-29 CRC cell line was used to induce spheroid formation. To investigate if there was any correlation between cyclinD1 expression and the activation of those two signalling pathways within the budding-like region of the spheroids, immunofluorescence (IF) staining of cyclinD1 was performed (Figure 3G). There was a

significant increase in cyclinD1 expression within the budding region of spheroids cultured in supplemented (TNF-α/TGF-β) media compared to control groups (p=0.014) (n=3) (Figure 3H).

Expression of EMT-related genes in TB in CRC

Due to the small sample size and small panel of genes (n=84) used to determine the underlying mechanism of TB, we further investigated the wider panel of RNA GeoMx WTA genes (n=18,815) in CRC TMA previously made from the invasive front of low (n=20) and high (n=23) budding regions (Figure 4A). PanCK staining was used to mask tumour (PanCK+) invasive areas for further analysis (Figure 4B). Interestingly, pairwise GSEA revealed of the invasive high TB tumour areas compared to low TB revealed an upregulation of EMT-related genes (padj <0.001, NES=1.76) and down-regulation of genes regarding the activation of KRAS (padj <0.001, NES=2.19) (Figure 4C-D).

The consistent finding from both the bulk and regional bulk RNA expression data was that tumours with high TB are highly enriched with an EMT signature. Therefore, a further study was performed to investigate if budding cells undergo EMT. mIF staining of two EMT markers (E-cadherin and β-catenin) was performed in the previously made CRC TMA from an invasive high budding area (n=180) from the discovery cohort. Interestingly, within the budding population, only a small proportion of TBs were found to have low E-cadherin and high β-catenin expression which represents the stimulation of EMT. (Supplementary Figure S7).

Immune-related genes expressed in PanCK- non-epithelial cells reveal the possible role of immunosuppression in TB.

In addition to the tumour region, regional bulk transcriptomic RNA from the whole CRC sections (n=12) can also determine gene expression from non-tumour (PanCK-) cells representing the immune-microenvironment. Thus, gene expression in PanCK- invasive region

was compared between low (n=6) and high (n=6) TB (Figure 5A). The results showed that *CXCL9* and *CXCL10* were significantly expressed in the invasive PanCK- area with low TB compared to high TB (Figure 5B-C) and suggests that expression of these chemokines may suppress the budding phenotype. To understand if samples with high TB are entirely devoid of immune activation or if a continuum/gradient exist, further investigation was carried out to identify gene expression between invasive proximal (closer to the budding area) and distal (distant further away from the budding area) PanCK- area (Figure 5D). Immune-related genes, such as *NKG7, CMKLR1, IFNAR1, CD47, IL6, ITGB8, FAS, BCL2, TNF, CXCR6, CD3E, and VSIR,* were significantly expressed in the distal compared to the proximal area (Figure 5E-F).

To further understand the role of specific immune phenotypes surrounding TB at the invasive margin, tissue sections from the discovery cohort (n=18) were utilised to perform mIF staining for specific phenotypes (Supplementary Figure S8).

In tumours with high TB, the percentage of regulatory T cells (CD3+FOXP3+) is significantly higher in the proximal compared to distal areas (p=0.012). There was no significant difference in the percentage of cytotoxic T cells (CD3+CD8+), although a smaller number of T cells was found at the proximal compared to distal areas. (Figure 5G). These results suggest an immune activity further away from the high budding region, which was not observed at the bud-rich invasive front.

The number of pan-macrophages (CD68+) is significantly higher in the proximal when compared to the distal area from the high budding area (n=9) (p=0.004). Additionally, there is a significantly lower number of CD68+CD163+ cells in the proximal compared to the distal area (p=0.004) in high (n=9) budding groups (Figure 5H). The results revealed that a higher number of CD66b+ cells, a neutrophil marker, were found at the proximal compared to the distal from

the high budding area (n=9) (p=0.09) (Figure 5I). The results demonstrated no significant difference when comparing the percentage of immune phenotypes at the proximal area surrounded with budding cells between tumours exhibiting either low or high budding phenotype. (Supplementary Figure S9).

Spatial relationship between budding cluster and its surrounded immune cells.

Further analysis of the mIF was performed to identify individual budding clusters and their relationship with the surrounding immune microenvironment. Radius analysis was performed to determine the density of immune phenotypes within an area of 50 microns from each budding clusters (Figure 5J). The averaged counts showed that there is a higher number of cytotoxic and regulatory T cells in tumours with low TB compared to high TB phenotype. Additionally, a higher number of macrophages (CD68+ and CD68+CD163+) were found in low TB phenotype whereas a higher number of CD66b+ cells within 50 µm from budding clusters in tumours with the high number of TB was reported (Figure 5K).

Furthermore, nearest distance analysis was performed to identify if the distance from budding clusters and specific phenotypes could reveal additional crosstalk or cell-cell interaction (Figure 5L). Interestingly, the budding clusters were observed to be closer to the regulatory T cells (CD3+CD8+) than cytotoxic T cells (CD3+FOXP3+), with an average distance of 41.31 um and 48.33 um respectively (Figure 5M). The findings showed that CD68+ pan-macrophages are significantly closer to budding clusters than CD68+CD163+ pro-inflammatory macrophages (27.49 vs 52.73 µm, $p<0.0001$) and CD66b+ (27.49 vs 42.97 µm, $p<0.001$) cells. Moreover, the budding clusters are significantly closer to neutrophils (CD66b+) than pro-inflammatory macrophages (CD68+CD163+) (42.97 vs 52.73 µm, p=0.028) (Figure 5N). It can be concluded

that budding clusters are significantly closer in proximity to pan-macrophages, neutrophil, and pro-inflammatory macrophages (Supplementary Table S8).

To validate the findings from mIF analysis, gene expression from non-tumour (PanCK-) regions was profiled using the GeoMx® WTA panel on cores from the invasive budding area of colorectal tumours (Figure 5O). GSEA revealed a significant decrease in gene signatures associated with immune activity in the invasive PanCK- area of high TB samples such as immune response activation (padj < 0.001, NES =-2.60), regulation of lymphocyte activation (padj <0.001, NES = -2.41) and B cell-mediated immunity (padj <0.001, NES = -2.71) compared to low TB invasive PanCK- area (Figure 5P). Gene expression deconvolution was performed to quantify cell populations based on the gene expression data. Results showed that there is a significantly low proportion of natural killer cells (NK) (p=0.01) and plasma cells (p=0.02) (Figure 5Q) at the PanCK- invasive area of high TB compared to low TB. This confirms our previous finding that, when tumours exhibit a low TB phenotype, there is an increase in inflammation whereas immunosuppressive activity is likely to be found within an invasive area of tumours with high TB.

Discussion

Colorectal cancer (CRC) is one of the most diagnosed cancers worldwide and within the UK (1). The development of screening programs has improved patient survival. However, some patients still have a poor prognosis due to metastatic disease (27). TB has recently been reported as a promising independent prognostic marker in CRC (6, 9, 28). This has resulted in an international consensus conference (ITBCC) on budding assessment, and it has been strongly recommended that TB status should be included in clinical reports in CRC (6, 29-31). Recently, TB has been added to the Royal Society Guideline. In both the discovery and validation CRC

patient cohorts, high TB was shown to have a strong prognostic value as well as a significant correlation with adverse clinical factors. Among those factors, the high TNM stage was found significantly associated with high TB phenotype in both cohorts. Additionally, TB was found to be an independent prognostic factor when entering multivariate analysis in both cohorts. These results confirm that TB is a promising prognostic biomarker in CRC patients as reported by other peer-reviewed studies (32).

Results from GSEA bulk transcriptomic data showed a significant enrichment for gene signatures related to TNF-α, TGF-β and EMT in tumours with high TB. However, gene expression from whole tissue may not fully describe the underlying mechanism of budding cells. This led to the use of regional bulk RNA transcriptomic GeoMx to determine the gene expression, in different compartments of tumour (PanCK+) and non-tumour (PanCK-) within the invasive budding-rich regions. *CCND1* gene, which encodes cyclinD1, was found to be highly upregulated in the invasive edge of tumours with high TB. The prognostic value of cyclinD1 in CRC patients has consistently been reported (33). However, no study has shown its correlation with TB in CRC. Our finding showed that patients with high expression of cyclinD1 within TB cells have poorer survival compared to those with low cyclinD1, although this was not significant due to the small sample size.

The synergistic role of TNF-α and TGF-β has been used as a model to study tumour progression and metastasis (34, 35). Using laser microdissection within the invasive budding area, Li and colleagues suggested an integration between TNF-α and TGF-β pathways in the development of invasive CRC tumours (36). To understand the formation of TB *in vitro*, 3D spheroids of known high TB CRC line HT-29 were used (37). Results showed an increase in cyclinD1 expression within the budding-like regions in spheroids cultured in media

supplemented with TNF-α and TGF-β. Studies have suggested that TB cells may not undergo proliferative activity, thus are able to avoid detection from the immune system which allows them to invade other parts of the body (38, 39). In agreement with human CRC, micrometastatses cells in human-like mouse model were shown to have decreased stemness and cell proliferation (40). This increased as the metastasis grew, suggesting a pre-metastasis stage of TB developing the niche for further invasion. Despite this study not demonstrating a significant association between cyclinD1 and high TB phenotype due to the underpower of the sample size, the strong trend is noteworthy, and merits further investigate as a potential target regulating budding formation. When produced, buds may undergo an early stage of the cell cycle, hence less proliferative, to maintain their ability to survive undetected before extravasation.

A correlation between the invasive PanCK+ high budding area and the EMT signature was observed in WTA GeoMx panel. Similarly, using microdissected RNA sequencing, De Smedt and colleagues reported the positive EMT gene signature in the invasive budding region compared to its main tumour core suggesting the possibility that EMT activation regulates budding formation in CRC (41). However, in our study mIF of known EMT markers was found only within a few populations of budding cells suggesting perhaps an involvement of non-EMT signalling that drives TB. This may also explain the partial EMT activity within TB, thus the selected markers could be crucial to understand the full activation of this pathway (42, 43). Recently, investigated in a mouse CRC model, micrometastases cells were reported to maintain expression of EPCAM and E-cadherin, suggesting they retain an epithelial phenotype and that EMT may not play an important role within these micro budding cells (40), validating our findings. In addition, a study by Fisher et al. suggested that the enrichment of the signature may also come from non-epithelial cells within the TME which also express genes within the EMT

signature, therefore mimicking an enrichment for EMT activity (44). As we observed an enrichment for EMT within bulk transcriptional analysis which was not validated at the epithelial level, in combination with the recent literature on epithelial cells within micrometastases retaining an epithelial phenotype (44), our finding fit with the literature.

The analysis of regional bulk RNA GeoMx was also used to investigate gene expression within a non-tumour (PanCK-) area. Our study has reported significantly higher expression of *CXCL9* and *CXCL10* genes in low TB tumours. This is in agreement with previous reports, which observed favorable prognostic value of these chemokines in CRC (45, 46). Moreover, gene expression results from DSP showed that immune-related activity was decreased in proximal compared to distal of an invasive area when the tumour exhibited high TB. The single cell mIF staining demonstrated a decreased number of cytotoxic T cells (CD3+CD8+) while an increase in immunosuppressive regulatory (CD3+FOXP3+) T cells was observed at the proximal compared to the distal invasive area of high TB tumours. Additionally, when a bigger sample size was studied, a significant expression of NK and plasma cells was reported at the proximal invasive PanCK- of low compared to high TB cases. These results suggest that TB could have an immunosuppressive role which alters their microenvironment leading to disease metastasis in CRC.

The spatial crosstalk between budding cells and their surrounding immune phenotypes was demonstrated and that CD3+FOXP3+ regulatory T cells were closer in distance to budding cells when compared to the CD3+CD8+ cytotoxic cells. This suggests that TB could perhaps recruit regulatory cells which then influence the stimulation of pro-tumorigenic activity. Interestingly, CD68+ pan-macrophages, a member of the other two cell populations within the panel, were found to be closest in distance to TB. Recent study reported the engulfment of

tumour budding by CD68+ pan-macrophages, confirming our finding as well as suggesting the defensive role of the innate immune response against budding cells (47). Additionally, the regulation of PD-L1/PD-1 by tumour-associated macrophages has previously been reported (48). This mechanism is worth exploring as the finding by Lin et al. has recently reported that, within the budding area, there is an increased immune cell expression of PD-L1 leading to immune evasion of budding cells (49).

Our study has the limitations. Firstly, cyclinD1 expression within TB has been accessed in high TB cases (n=180). Though there is a trend in its prognostic within TB, due to the small sample, there is a need to further validate this result in a bigger cohort. Additionally, as a study by Haddad et al. has recently reported that pseudobudding can occur due to gland disruption and inflammation, which may misrepresent TB in CRC (38). They suggested that using PanCK staining may not represent true TB and that the pseudobudding phenotype should be taken into consideration in TB assessment. Future work should exploit single-cell technology to capture the true biology of budding cells.

In conclusion, our study utilised a series of multi-omics approaches to include bulk and regional bulk RNA transcriptomic data validated by conventional IHC and multiplex IF staining to understand the underlying mechanism of TB, both at an epithelial-level and within the TME. Our findings point towards a novel epithelial-specific phenotype, differing to conventional EMT, which has not been previous explored as well the presence of an immunosuppressive microenvironment surrounding TB in CRC (Supplementary Figure S10). This study paves the way for further characterisation of the epithelial budding phenotype to identify a therapeutic target for TB in combination with the exploitation of the immune microenvironment.

Acknowledgement

The authors wish to acknowledge the Glasgow Tissue Research Facility and NHS Greater Glasgow and Clyde biorepository for helping in preparing the tissue for this study. The authors gratefully acknowledge for Mahidol University and National Research Council of Thailand (2022) for funding and financial support (PH) (N41B650195). The authors would also like to acknowledge Dr. Suyanee Thongchot for her accommodation and Siriraj Research Fund, Faculty of Medicine Siriraj Hospital for the financial support (R016533013). The authors also grateful for Bowel Cancer UK for the funding (AM) (22PT00058). The authors would also like to acknowledge the Academic unit of Surgery at the Glasgow Royal Infirmary. The authors wish to acknowledge Dr Catherine Winchester for help reviewing the manuscript prior to submission. The authors also wish to acknowledge CRUK funding which enable this research (Scottish Cancer Centre CTRQQR-2021/100006) (JE) as well as Chief Scientific Office (CSO) Scotland funding (EPD/22/13) (KP, JE).

Competing interests

Professor D.C. McMillan is on the editorial board for BJC.

Authors' contributions

P.Hatthakarnkul: Data curation, Software, Formal analysis, Funding acquisition, Validation, Investigation, Visualization, Methodology, Writing-original draft, Project administration and Writing-review and editing.
N.Fisher: Resources, Data curation, Software, Formal analysis, Investigation, Methodology, Writing-review and editing
L.Schubert Santana: Software, Formal analysis, Investigation, Methodology, Writing-review and editing
H.Leslie: Investigation, Methodology
A.Legrini: Software, Methodology
A.Ammar: Investigation, Methodology
A.Matly: Formal analysis, Methodology
K.Pennel: Methodology, Writing-review and editing
I.Powley: Software, Methodology
C.Bull: Formal analysis
P.Alexander: Methodology
J.Sztanko: Formal analysis,
J.Quinn: Supervision, Writing-review and editing
J.Hay: Resources, Methodology
H.Morgan: Resources, Methodology
C.Dietrich: Methodology
Y.Doncheva: Methodology
G.Lynch: Writing-review and editing
N.Maka: Resources
C.Roxburgh: Resources
H.van Wyk: Resources, Supervision, Methodology
X.Fu: Supervision, Methodology, Writing-review and editing
D.McMillan: Resources, Writing-review and editing
P.Dunne: Resources
N.Jamieson: Resources
C.Thuwajit: Resources, Supervision, Writing-review and editing
J.Edwards: Conceptualization, Resources, Supervision, Funding acquisition, Writing-review and editing

References

1. Rawla P, Sunkara T, Barsouk A. Epidemiology of colorectal cancer: incidence, mortality, survival, and risk factors. Prz Gastroenterol. 2019;14(2):89-103.
2. Navarro M, Nicolas A, Ferrandez A, Lanas A. Colorectal cancer population screening programs worldwide in 2016: An update. World J Gastroenterol. 2017;23(20):3632-42.
3. Vayrynen V, Wirta EV, Seppala T, Sihvo E, Mecklin JP, Vasala K, et al. Incidence and management of patients with colorectal cancer and synchronous and metachronous colorectal metastases: a population-based study. BJS Open. 2020;4(4):685-92.
4. Biller LH, Schrag D. Diagnosis and Treatment of Metastatic Colorectal Cancer: A Review. JAMA. 2021;325(7):669-85.
5. Xie YH, Chen YX, Fang JY. Comprehensive review of targeted therapy for colorectal cancer. Signal Transduct Target Ther. 2020;5(1):22.
6. Lugli A, Kirsch R, Ajioka Y, Bosman F, Cathomas G, Dawson H, et al. Recommendations for reporting tumor budding in colorectal cancer based on the International Tumor Budding Consensus Conference (ITBCC) 2016. Mod Pathol. 2017;30(9):1299-311.
7. Zlobec I, Hädrich M, Dawson H, Koelzer VH, Borner M, Mallaev M, et al. Intratumoural budding (ITB) in preoperative biopsies predicts the presence of lymph node and distant metastases in colon and rectal cancer patients. British Journal of Cancer. 2013;110(4):1008-13.
8. Koelzer VH, Zlobec I, Lugli A. Tumor budding in colorectal cancer--ready for diagnostic practice? Hum Pathol. 2016;47(1):4-19.
9. Lugli A, Zlobec I, Berger MD, Kirsch R, Nagtegaal ID. Tumour budding in solid cancers. Nat Rev Clin Oncol. 2021;18(2):101-15.
10. Salhia B, Trippel M, Pfaltz K, Cihoric N, Grogg A, Ladrach C, et al. High tumor budding stratifies breast cancer with metastatic properties. Breast Cancer Res Treat. 2015;150(2):363-71.
11. Pai RK, Cheng YW, Jakubowski MA, Shadrach BL, Plesec TP, Pai RK. Colorectal carcinomas with submucosal invasion (pT1): analysis of histopathological and molecular factors predicting lymph node metastasis. Mod Pathol. 2017;30(1):113-22.
12. Slik K, Blom S, Turkki R, Välimäki K, Kurki S, Mustonen H, et al. Combined epithelial marker analysis of tumour budding in stage II colorectal cancer. The Journal of Pathology: Clinical Research. 2019;5(1):63-78.
13. Ueno H, Ishiguro M, Nakatani E, Ishikawa T, Uetake H, Matsuda C, et al. Prospective Multicenter Study on the Prognostic and Predictive Impact of Tumor Budding in Stage II Colon Cancer: Results From the SACURA Trial. J Clin Oncol. 2019;37(22):1886-94.
14. Mitrovic B, Handley K, Assarzadegan N, Chang HL, Dawson HAE, Grin A, et al. Prognostic and Predictive Value of Tumor Budding in Colorectal Cancer. Clin Colorectal Cancer. 2021;20(3):256-64.
15. Hatthakarnkul P, Quinn JA, Ammar A, Lynch G, Van Wyk H, McMillan DC, et al. Molecular mechanisms of tumour budding and its association with microenvironment in colorectal cancer. Clin Sci (Lond). 2022;136(8):521-35.

16. Hatthakarnkul P, Quinn JA, Matly AAM, Ammar A, van Wyk HC, McMillan DC, et al. Systematic review of tumour budding and association with common mutations in patients with colorectal cancer. Crit Rev Oncol Hematol. 2021;167:103490.
17. McMillan DC. The systemic inflammation-based Glasgow Prognostic Score: a decade of experience in patients with cancer. Cancer Treat Rev. 2013;39(5):534-40.
18. Lu X, Guo W, Xu W, Zhang X, Shi Z, Zheng L, et al. Prognostic value of the Glasgow prognostic score in colorectal cancer: a meta-analysis of 9,839 patients. Cancer Manag Res. 2019;11:229-49.
19. Proctor MJ, Morrison DS, Talwar D, Balmer SM, O'Reilly DS, Foulis AK, et al. An inflammation-based prognostic score (mGPS) predicts cancer survival independent of tumour site: a Glasgow Inflammation Outcome Study. Br J Cancer. 2011;104(4):726-34.
20. Gao LF, Zhong Y, Long T, Wang X, Zhu JX, Wang XY, et al. Tumor bud-derived CCL5 recruits fibroblasts and promotes colorectal cancer progression via CCR5-SLC25A24 signaling. J Exp Clin Cancer Res. 2022;41(1):81.
21. Nearchou IP, Lillard K, Gavriel CG, Ueno H, Harrison DJ, Caie PD. Automated Analysis of Lymphocytic Infiltration, Tumor Budding, and Their Spatial Relationship Improves Prognostic Accuracy in Colorectal Cancer. Cancer Immunol Res. 2019;7(4):609-20.
22. van Wyk HC, Park JH, Edwards J, Horgan PG, McMillan DC, Going JJ. The relationship between tumour budding, the tumour microenvironment and survival in patients with primary operable colorectal cancer. Br J Cancer. 2016;115(2):156-63.
23. Shinto E, Yoshida Y, Kajiwara Y, Okamoto K, Mochizuki S, Yamadera M, et al. Clinical Significance of a Gene Signature Generated from Tumor Budding Grade in Colon Cancer. Ann Surg Oncol. 2020;27(10):4044-54.
24. Yeakley JM, Shepard PJ, Goyena DE, VanSteenhouse HC, McComb JD, Seligmann BE. A trichostatin A expression signature identified by TempO-Seq targeted whole transcriptome profiling. PloS one. 2017;12(5):e0178302.
25. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15-21.
26. Matthews L, Gopinath G, Gillespie M, Caudy M, Croft D, de Bono B, et al. Reactome knowledgebase of human biological pathways and processes. Nucleic Acids Res. 2009;37(Database issue):D619-22.
27. Zeineddine FA, Zeineddine MA, Yousef A, Gu Y, Chowdhury S, Dasari A, et al. Survival improvement for patients with metastatic colorectal cancer over twenty years. NPJ Precis Oncol. 2023;7(1):16.
28. Sirin AH, Sokmen S, Unlu SM, Ellidokuz H, Sarioglu S. The prognostic value of tumor budding in patients who had surgery for rectal cancer with and without neoadjuvant therapy. Tech Coloproctol. 2019;23(4):333-42.
29. Haddad TS, Lugli A, Aherne S, Barresi V, Terris B, Bokhorst JM, et al. Improving tumor budding reporting in colorectal cancer: a Delphi consensus study. Virchows Arch. 2021;479(3):459-69.
30. Graham RP, Vierkant RA, Tillmans LS, Wang AH, Laird PW, Weisenberger DJ, et al. Tumor Budding in Colorectal Carcinoma: Confirmation of Prognostic Significance and Histologic Cutoff in a Population-based Cohort. Am J Surg Pathol. 2015;39(10):1340-6.

31. Zlobec I, Berger MD, Lugli A. Tumour budding and its clinical implications in gastrointestinal cancers. Br J Cancer. 2020;123(5):700-8.
32. Rogers AC, Winter DC, Heeney A, Gibbons D, Lugli A, Puppa G, et al. Systematic review and meta-analysis of the impact of tumour budding in colorectal cancer. Br J Cancer. 2016;115(7):831-40.
33. Li Y, Wei J, Xu C, Zhao Z, You T. Prognostic significance of cyclin D1 expression in colorectal cancer: a meta-analysis of observational studies. PLoS One. 2014;9(4):e94508.
34. Liu ZW, Zhang YM, Zhang LY, Zhou T, Li YY, Zhou GC, et al. Duality of Interactions Between TGF-beta and TNF-alpha During Tumor Formation. Front Immunol. 2021;12:810286.
35. Hahn S, Nam MO, Noh JH, Lee DH, Han HW, Kim DH, et al. Organoid-based epithelial to mesenchymal transition (OEMT) model: from an intestinal fibrosis perspective. Sci Rep. 2017;7(1):2435.
36. Li H, Zhong A, Li S, Meng X, Wang X, Xu F, et al. The integrated pathway of TGFbeta/Snail with TNFalpha/NFkappaB may facilitate the tumor-stroma interaction in the EMT process and colorectal cancer prognosis. Sci Rep. 2017;7(1):4915.
37. Georges LMC, De Wever O, Galvan JA, Dawson H, Lugli A, Demetter P, et al. Cell Line Derived Xenograft Mouse Models Are a Suitable in vivo Model for Studying Tumor Budding in Colorectal Cancer. Front Med (Lausanne). 2019;6:139.
38. Haddad TS, van den Dobbelsteen L, Ozturk SK, Geene R, Nijman IJ, Verrijp K, et al. Pseudobudding: ruptured glands do not represent true tumor buds. J Pathol. 2023;261(1):19-27.
39. Dawson H, Koelzer VH, Karamitopoulou E, Economou M, Hammer C, Muller DE, et al. The apoptotic and proliferation rate of tumour budding cells in colorectal cancer outlines a heterogeneous population of cells with various impacts on clinical outcome. Histopathology. 2014;64(4):577-84.
40. Canellas-Socias A, Cortina C, Hernando-Momblona X, Palomo-Ponce S, Mulholland EJ, Turon G, et al. Metastatic recurrence in colorectal cancer arises from residual EMP1(+) cells. Nature. 2022;611(7936):603-13.
41. De Smedt L, Palmans S, Andel D, Govaere O, Boeckx B, Smeets D, et al. Expression profiling of budding cells in colorectal cancer reveals an EMT-like phenotype and molecular subtype switching. Br J Cancer. 2017;116(1):58-65.
42. Grigore AD, Jolly MK, Jia D, Farach-Carson MC, Levine H. Tumor Budding: The Name is EMT. Partial EMT. J Clin Med. 2016;5(5).
43. Yang J, Antin P, Berx G, Blanpain C, Brabletz T, Bronner M, et al. Guidelines and definitions for research on epithelial-mesenchymal transition. Nat Rev Mol Cell Biol. 2020;21(6):341-52.
44. Fisher NC, Byrne RM, Leslie H, Wood C, Legrini A, Cameron AJ, et al. Biological Misinterpretation of Transcriptional Signatures in Tumor Samples Can Unknowingly Undermine Mechanistic Understanding and Faithful Alignment with Preclinical Data. Clin Cancer Res. 2022;28(18):4056-69.
45. Chen J, Ye X, Pitmon E, Lu M, Wan J, Jellison ER, et al. IL-17 inhibits CXCL9/10-mediated recruitment of CD8(+) cytotoxic T cells and regulatory T cells to colorectal tumors. J Immunother Cancer. 2019;7(1):324.

46. Zhang J, Tao J, Gao RN, Wei ZY, He YS, Ren CY, et al. Cytotoxic T-Cell Trafficking Chemokine Profiles Correlate With Defined Mucosal Microbial Communities in Colorectal Cancer. Front Immunol. 2021;12:715559.
47. Koelzer VH, Canonica K, Dawson H, Sokol L, Karamitopoulou-Diamantis E, Lugli A, et al. Phenotyping of tumor-associated macrophages in colorectal cancer: Impact on single cell invasion (tumor budding) and clinicopathological outcome. Oncoimmunology. 2016;5(4):e1106677.
48. Pu Y, Ji Q. Tumor-Associated Macrophages Regulate PD-1/PD-L1 Immunosuppression. Front Immunol. 2022;13:874589.
49. Lin JR, Wang S, Coy S, Chen YA, Yapp C, Tyler M, et al. Multiplexed 3D atlas of state transitions and immune interaction in colorectal cancer. Cell. 2023;186(2):363-81 e19.

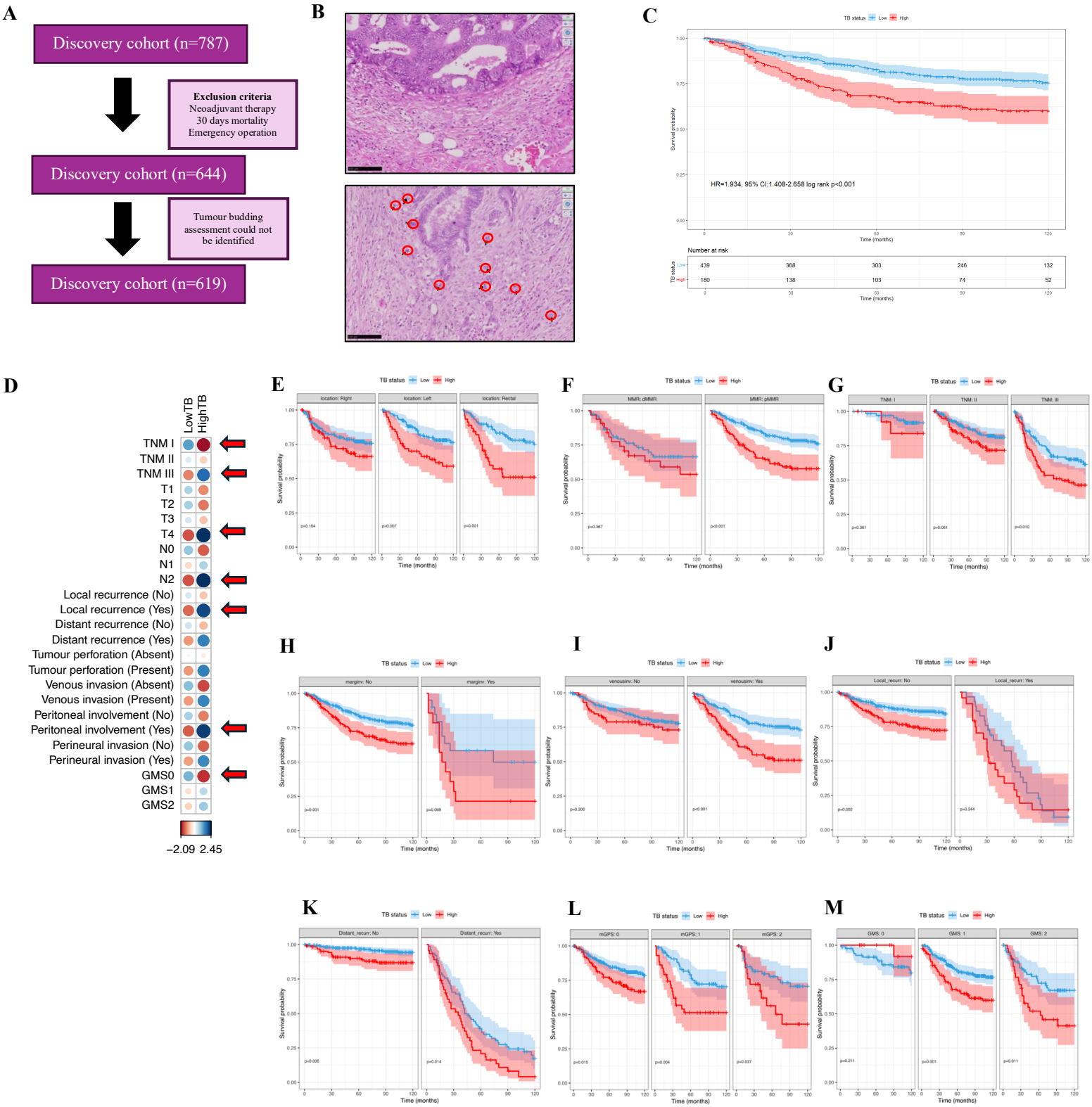
A
Discovery cohort (n=787)
Exclusion criteria
Neoadjuvant therapy
30 days mortality
Emergency operation
Discovery cohort (n=644)
Tumour budding assessment could not be identified
Discovery cohort (n=619)
B
C
TB status Low High
HR=1.934, 95% CI;1.408-2.658 log rank p<0.001
Survival probability
Time (months)
Number at risk
Low 439 368 303 246 132
High 180 138 103 74 52
D
LowTB
HighTB
TNM I
TNM II
TNM III
T1
T2
T3
T4
N0
N1
N2
Local recurrence (No)
Local recurrence (Yes)
Distant recurrence (No)
Distant recurrence (Yes)
Tumour perforation (Absent)
Tumour perforation (Present)
Venous invasion (Absent)
Venous invasion (Present)
Peritoneal involvement (No)
Peritoneal involvement (Yes)
Perineural invasion (No)
Perineural invasion (Yes)
GMS0
GMS1
GMS2
−2.09 2.45
E
location: Right
location: Left
location: Rectal
F
MMR: dMMR
MMR: pMMR
G
TNM: I
TNM: II
TNM: III
H
marginv: No
marginv: Yes
I
venousinv: No
venousinv: Yes
J
Local_recurr: No
Local_recurr: Yes
K
Distant_recurr: No
Distant_recurr: Yes
L
mGPS: 0
mGPS: 1
mGPS: 2
M
GMS: 0
GMS: 1
GMS: 2

Figure 1 *(A) Simplified consort diagram summarizing the number of patients in discovery cohort originally (n=787), after applied exclusion criteria (n=644) and after budding assessment (n=619). (B) Representative image of tumours with low (top) and high (bottom) budding phenotype in CRC H&E stained sections. (C) Kaplan-Meier survival analysis based on tumour budding phenotype for cancer specific survival (CSS) in discovery cohort. (D) The correlation plot for the Pearson's chi square test showed the residuals for TB phenotypes against clinical factors; positive residuals are in blue, suggesting a positive association and negative residuals are red, suggesting a negative association. The bigger size of the circle, the more significant association was found, red arrow indicates the statistically significant association. (E-M) Kaplan-Meier survival analysis based on tumour budding phenotype stratified by clinical factors for cancer specific survival (CSS) in discovery cohort.*

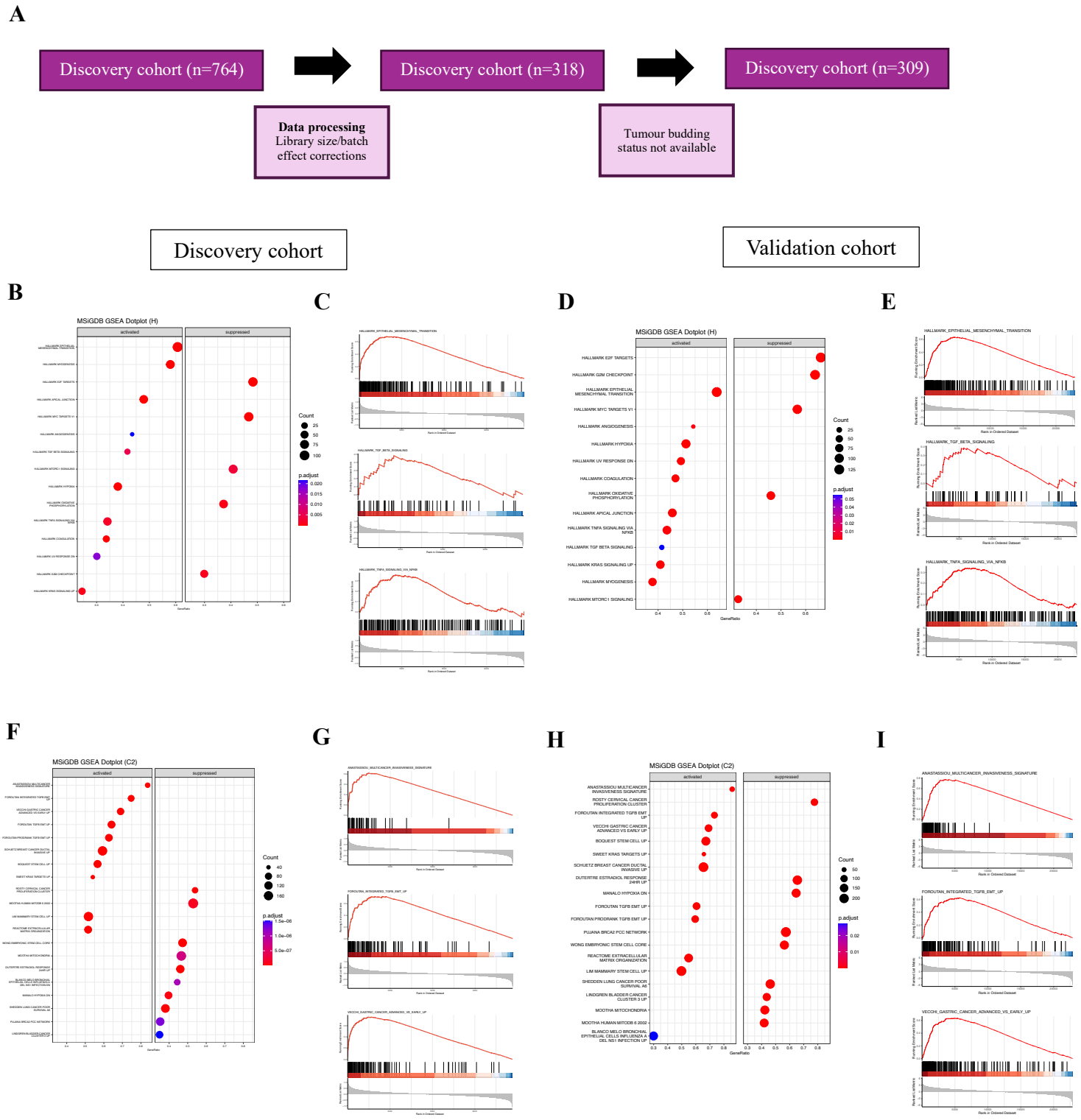

Figure 2 *(A) Simplified consort diagram summarizing the number of patients in discovery cohort used in transcriptomic analysis. Gene Set Enrichment Analysis (GSEA) in discovery and validation cohorts between low (n=209,59) and high TB (n=100,26) respectively. Dot plot based GSEA analysis shows top 10 gene sets significantly up and down-regulated in (B, D) hallmark and (F, H) curated gene set pathways. (C,E,G and I) Enrichment plot shows the differences of signalling pathways involved between tumours with low and high TB.*

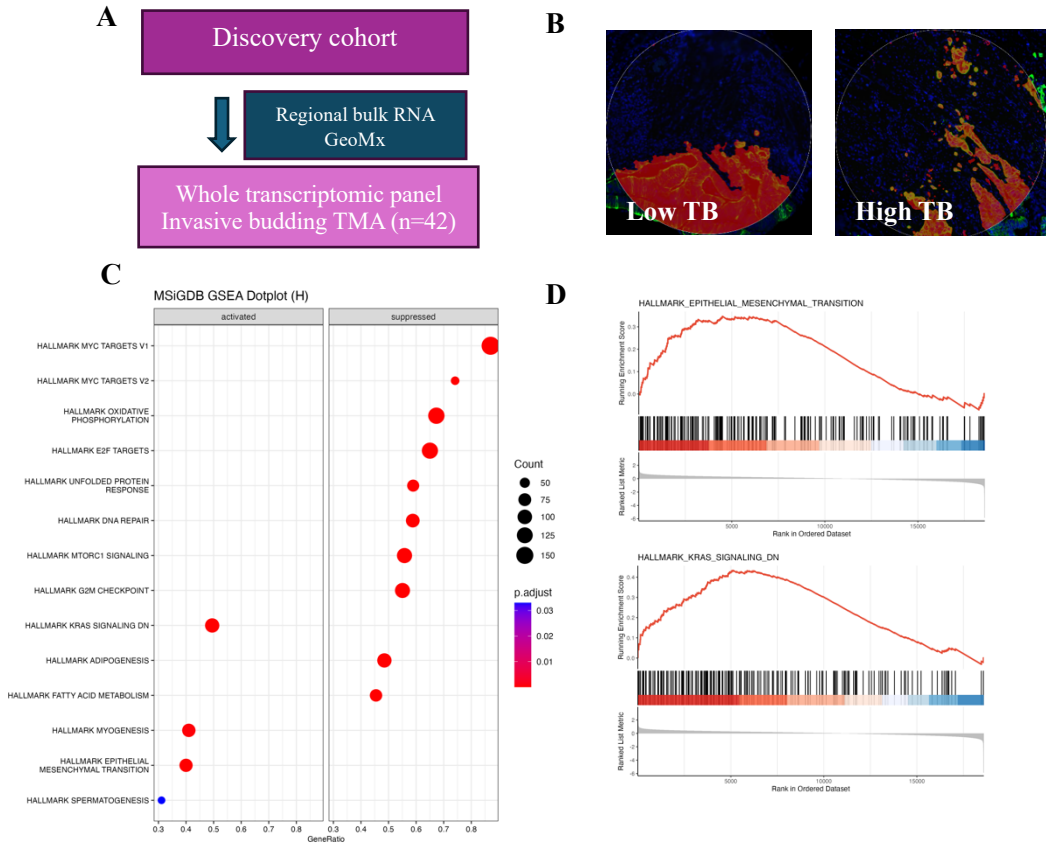

Figure 4 *(A) Simplified consort diagram summarizing the number of patients in discovery cohort used in regional bulk RNA GeoMx (B) The representative images of TMA from invasive area marked epithelial tissue (red) of tumours with low (n=20) and high (n=23) TB for further analysis. (C) Dot plot based GSEA analysis shows top 10 gene sets significantly up and down-regulated in hallmark gene set pathways. (D) Enrichment plot shows the differences of signalling pathways involved between tumours with low and high TB.*

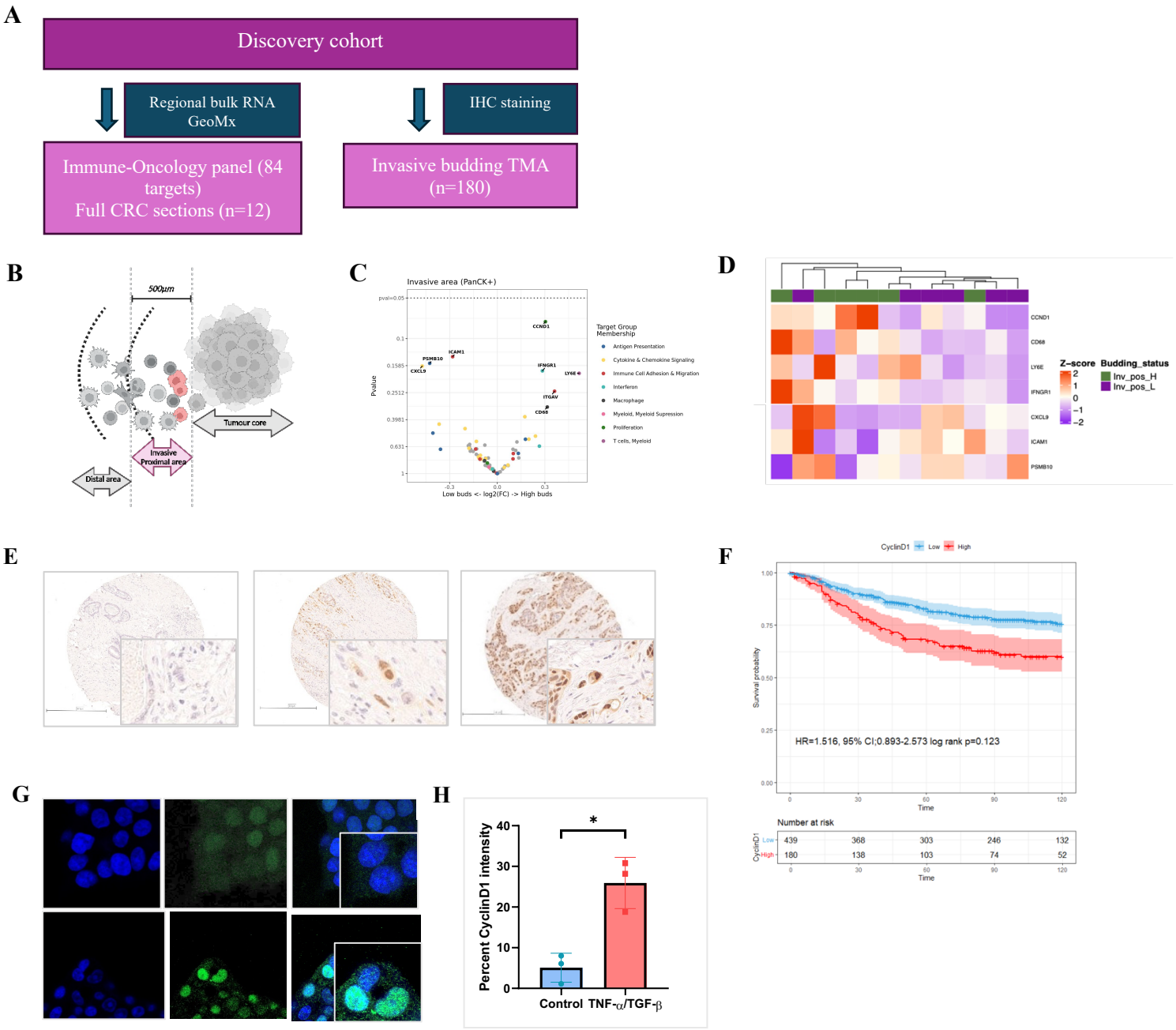

Figure 3 *(A) Simplified consort diagram summarizing the number of patients in discovery cohort used in regional bulk RNA GeoMx and IHC staining (B) Graphic shows the region of tumour invasive area used to determine the differential genes expression between tumours with low and high TB. (C) Volcano plot of gene expression in the invasive tumour area between tumours with high and low budding profile (n=12; low 6 and high 6). (D) Heatmap illustrating the expression of the genes in the two different groups. (E) IHC staining of CyclinD1 protein expression (low, medium and high) in TMA of budding cells (n=190) and (F) Kaplan-Meier survival analysis based on cyclinD1 expression for cancer specific survival (CSS) in budding cells. (G) CyclinD1 immunofluorescence staining (green color) in the budding region of the HT-29 spheroids cultured in normal (top) and supplemented with TNF-α / TGF-β media (bottom) (n=3). (H) Bar plots showed the percent intensity of cyclinD1 compared between control and supplemented (TNF-α / TGF-β) groups (n=3).*

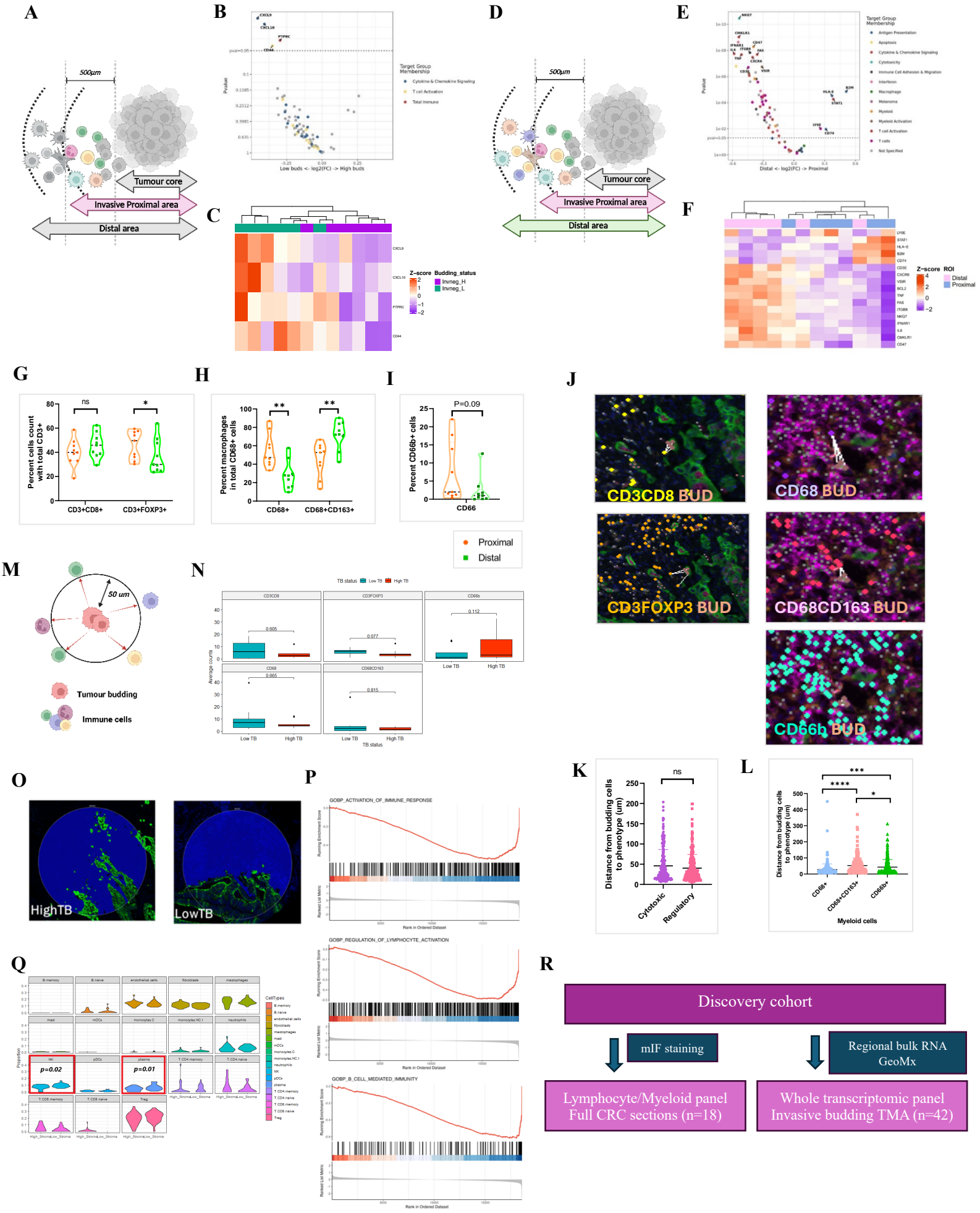
A
500µm
Tumour core
Invasive Proximal area
Distal area
B
Target Group Membership
Cytokine & Chemokine Signaling
T cell Activation
Total Immune
Low buds <- log2(FC) -> High buds
C
Z-score
Budding_status
Invneg_H
Invneg_L
D
500µm
Tumour core
Invasive Proximal area
Distal area
E
Target Group Membership
Distal <- log2(FC) -> Proximal
F
Z-score
ROI
Distal
Proximal
G
Percent cells count with total CD3+
CD3+CD8+
CD3+FOXP3+
H
Percent macrophages in total CD68+ cells
CD68+
CD68+CD163+
I
Percent CD66b+ cells
P=0.09
CD66
Proximal
Distal
J
CD3CD8 BUD
CD68 BUD
CD3FOXP3 BUD
CD68CD163 BUD
CD66b BUD
M
50 um
Tumour budding
Immune cells
N
TB status
Low TB
High TB
CD3CD8
CD3FOXP3
CD66b
CD68
CD68CD163
Average counts
K
Distance from budding cells to phenotype (um)
Cytotoxic
Regulatory
L
Distance from budding cells to phenotype (um)
CD68+
CD68+CD163+
CD66b+
Myeloid cells
O
HighTB
LowTB
P
GOBP_ACTIVATION_OF_IMMUNE_RESPONSE
GOBP_REGULATION_OF_LYMPHOCYTE_ACTIVATION
GOBP_B_CELL_MEDIATED_IMMUNITY
Q
p=0.02
p=0.01
R
Discovery cohort
mIF staining
Regional bulk RNA GeoMx
Lymphocyte/Myeloid panel
Full CRC sections (n=18)
Whole transcriptomic panel
Invasive budding TMA (n=42)

Figure 5 *(A) Graphic shows the region of invasive non-tumour area used to determine the differential genes expression compared between tumours with low (n=6) and high TB (n=6). (B) Volcano plot revealed genes significantly expressed compared between two groups. (C) Heatmap illustrated the expression of the genes in two different groups (persian green; lowTB and purple; highTB). (D) Graphic shows the region of interested; distal and proximal stromal (PanCK-) from invasive area, used to determine the differential genes expression in tumours with high TB (n=6). (E) Volcano plot revealed genes significantly expressed compared between two distinct areas. (F) Heatmap illustrated the expression of the genes in two different areas (pink; distal and cyan; proximal). (G) Boxplot shows a percent count of cytotoxic (CD3+CD8+) and T regulatory (CD3+FOXP3+) cells compared between invasive (orange) and distant stromal (green) area in tumours with high TB. Boxplot shows a percent count of (H) CD68+ and CD68+CD163+ cells and (I) CD66b+ cells compared between invasive (orange) and distant stromal (green) area in tumours with high TB. (J) Representative image of the distance measured from individual budding cluster to immune cells. (K-L) Dotplots shows the distance from budding clusters to each immune phenotype (um). (M) Illustration of radius analysis from the budding clusters within the range of 50um. (N) Boxplot shows the average counts of each immune phenotype within 50um from TB cells. (O) The representative images of TMA from invasive area marked non-epithelial tissue (blue) of low (n=20) and high (n=20) TB from GeoMx (P) Enrichment plot shows the differences of genes signature associated with immune response between non-tumour area between low and high TB. (Q) Cellular deconvolution for the proportion of immune cells compared between low and high TB. (R) Simplified consort diagram summarizing the number of patients in discovery cohort used in mIF staining panels and regional bulk RNA GeoMx.*

Table 1 *Univariate and Multivariate analysis for cancer specific survival (CSS) in discovery cohort*

Clinicopathological characteristics	Univariable analysis		Multivariate analysis	
	Hazard ratio (95% CI)	***P***	**Hazard ratio (95% CI)**	***P***
Sex	1.286 (0.94-1.76)	0.117		
Age	1.29 (1.07-1.56)	0.008	1.49 (1.12-1.99)	0.006
MMR status	0.65 (0.49-0.85)	0.002	0.58 (0.36-0.95)	0.032
Recurrence status	4.29 (3.64-5.07)	<0.001	4.72 (3.59-6.19)	<0.001
TNM stage	2.64 (2.02-3.46)	<0.001		
T stage	1.89 (1.49-2.39)	<0.001		
N stage	2.19 (1.80-2.67)	<0.001	1.87 (1.33-2.63)	<0.001
Margin involvement	4.22 (2.69-6.63)	<0.001	6.67 (3.16-14.07)	<0.001
Perineural invasion	3.63 (2.35-5.60)	<0.001		
Peritoneal involvement	2.10 (1.52-2.89)	<0.001		
Tumour perforation	2.79 (1.23-6.31)	0.014		
Venous invasion	1.60 (1.17-2.19)	0.003		
TB	1.93 (1.40-2.65)	<0.001	1.64 (1.02-2.62)	0.040

MMR = Mismatch repair, TNM = Tumour Node Metastasis, TB = Tumour budding