TITLE

Distinct cell states define the developmental trajectories of mucinous appendiceal neoplasms towards pseudomyxoma metastases

AUTHORS

Carlos Ayala[1]*, Anuja Sathe[2]*, Xiangqi Bai[2], Susan M. Grimes[3], Jeanne Shen[4], George A. Poultsides[1], Byrne Lee[1], Hanlee P. Ji[2,3]

*These authors contributed equally.

INSTITUTIONS

[1] Division of Surgical Oncology, Department of Surgery, Stanford University, Stanford, CA 94305, United States

[2] Division of Oncology, Department of Medicine, Stanford University School of Medicine, Stanford, CA, United States

[3] Stanford Genome Technology Center, Stanford University, Palo Alto, CA, United States

[4] Department of Pathology, Stanford University School of Medicine, Stanford, CA USA

CORRESPONDING AUTHOR

Hanlee P. Ji

Email: genomics_ji@stanford.edu

Mailing address: CCSR 2245, 269 Campus Drive, Stanford, CA-94305, USA

Conflict of interest statement: The authors declare no potential conflicts of interest.

ABSTRACT

Objective: Appendiceal mucinous neoplasms **(AMN)** start in the appendix. Following appendiceal perforation, these tumor metastasize, leading to pseudomyxoma peritonei **(PMP)**. We characterized the distinct cell types and states of AMN and PMP.

Design: Single-cell RNA-seq was conducted on 31 samples including AMNs, PMP metastases and a subset of matched normal appendix or omental tissues. We validated the findings with immunohistochemistry, mass spectrometry on malignant ascites from PMP patients and gene expression data from an independent set of 63 PMPs.

Results: AMNs and PMPs had epithelial tumor cells with goblet features including the elevated expression of mucin genes. Among these tumors, we identified developmental cell states of the epithelium that ranged from progenitor phase to goblet cell differentiation. Metastatic PMP cells had a distinct cell state with gene expression signatures indicative of mTOR and RAS signaling pathways. A series of cancer genes defined this metastatic cell state. Matched PMP metastases from the same patient differed in their expression signatures. The cell state signatures were validated in external datasets containing 63 PMP patients. AMN and PMP tumor microenvironment **(TME)** contained CD4 T follicular helper-like cells and cytotoxic CD8 T cells.

Conclusion: AMN and PMP tumors consisted of progenitor epithelial cells that differentiate into goblet cells. PMP cells had a distinct cell state indicative of metastasis. The TME was infiltrated with T cells, suggesting that immunotherapy is a potential treatment.

SUMMARY BOX

What is already known about this subject?

- Pseudomyxoma peritonei **(PMP)** represents aggressive metastatic dissemination of appendiceal mucinous neoplasms **(AMN)** and has a poor prognosis.
- The cellular and molecular features of both AMN and PMP remain poorly characterized.

What are the new findings?

- AMN and PMP tumor cells demonstrated heterogeneity in their cellular states, that has not been identified previously.
- AMN and PMP tumor cells contain progenitor states that proceed to goblet cell differentiation. PMP tumor cells further differentiate into a metastatic cell state. A single tumor contains these multiple cellular states.
- PMP lesions from the same patient differ in their transcriptional signatures.
- Goblet cell secretory proteins including mucins can be detected in PMP ascites.
- The tumor microenvironment (TME) in AMN and PMP is not immune excluded and contains a rich lymphocyte infiltrate.

How might it impact on clinical practice in the foreseeable future?

- Genes and pathways defining heterogenous AMN and PMP cellular states can be used for targeted therapy.
- AMN and PMP can potentially be treated using immunotherapy.

INTRODUCTION

Appendiceal mucinous neoplasms **(AMN)** are a rare and generally indolent malignancy that originates from the appendix. Given the growth pattern and anatomical origin, AMNs are prone to rupture and invade the abdominal cavity. This pattern of intraabdominal dissemination is referred to as pseudomyxoma peritonei **(PMP)**. Metastatic dissemination of PMP is the cause of significant morbidity [1]. Current treatment for PMP consists of cytoreductive surgery **(CRS)** and heated intraperitoneal chemotherapy **(HIPEC)** using agents such as mitomycin C **(MMC)** as the perfusate [2]. However, only 60% of patients who develop PMP are candidates for CRS and HIPEC [3]. The remainder of patients are treated with systemic chemotherapy, albeit with limited therapeutic effectiveness for most patients. The poor patient outcomes reflect the extent of the metastasis and the specific cellular features of the tumor [4, 5]. There remain fundamental questions about the cellular features and dysregulated biology of tumor cells of AMN and PMP. Determining these features may lead to improved therapeutic strategies.

AMNs are characterized by goblet cell hyperplasia and elevated mucin production beyond normal levels [6]. Normally, goblet cells produce a protective mucus lining for the small and large intestine. The mucus is composed of molecules like Mucin-2 **(MUC2)** [7]. In addition, goblet cells constantly renew and differentiate from the underlying intestinal stem cell population [8]. Based on studies using immunohistochemistry **(IHC)** and gene expression, MUC2 is the predominant mucin that is enriched in AMN and PMP tissue [6]. Some reports have stated that PMPs originate from MUC2-expressing goblet or goblet-like cells [9]. However, the epithelial tumor cells of PMPs do not have the classic histopathologic characteristics of goblet cells [10]. Moreover, little is known about the pathways that distinguish AMNs from metastatic PMPs.

The definitive genomic studies of PMPs were conducted by Levine et al. [5, 11]. They used gene expression analysis on the largest collection of appendiceal cancers to date. They identified genomic signatures that were distinct from colorectal cancer, confirming their unique biology. These gene expression signatures were indicative of oncogenic processes and pathways which have prognostic implications.

This study's goal was characterizing the cellular and molecular features of AMN and PMP. Our analysis was focused on the origins and differentiation of the tumor epithelial cell populations. AMN and PMP cancers have been challenging to study due to their low cellularity and complex composition. In particular, PMPs are complex mixtures of different cell types embedded within a gelatinous matrix. To characterize the specific cell populations from these tumors, we applied single cell RNA sequencing **(scRNA-seq)** on these tumors and determined their gene expression characteristics in a discovery set. This approach enabled us to identify specific epithelial cell subpopulations within these tumors. We discovered specific gene expression markers for the different cell types and validated the presence of some of these marker proteins among samples and ascites. We identified the specific features of the epithelial tumor cell for both types. Subsequently, we validated these findings in an expanded, independent cohort of tumors with gene expression data. This study identified the cell states in tumor epithelial cells that enable progression of AMN to PMP. Finally, the tumor microenvironment **(TME)** of these samples had a diverse immune cell infiltration. This result indicates the potential for immunotherapy treatment among these patients.

METHODS

A complete description is provided in the **Supplemental Methods**. We processed the tumors into single cell suspensions and generated scRNA-seq libraries. Following quality control

measures, integrated clustering analysis was performed to remove batch effects across samples [12-14]. We leveraged a variety of bioinformatics analyses to determine cell identity, differentiation and biologic pathway activity.

Patient and public involvement

Patients or the public were not involved in the design, or conduct, or reporting, and dissemination plans of our research.

RESULTS

Study overview

This study included AMNs without perforation and PMPs **(Table 1)**. Among these tumors, the AMNs were low grade appendiceal mucinous neoplasms **(LAMN)**. The PMPs were low-grade mucinous carcinoma peritonei **(LMCP)** or high-grade mucinous carcinoma peritonei **(HMCP)**. We used multiple approaches to characterize the cellular origins of this disease **(Fig. 1A)**. The first part of the study involved scRNA-seq applied to a discovery cohort of 31 AMNs, PMPs and a subset of matched normal appendix or omental tissue that came from 14 patients. A subset of the PMP samples represented multiple metastatic sites from the same patient. Tumor sites included the omentum, the small bowel mesentery, the liver capsule and the ovaries. In addition, we analyzed several matched normal tissues pairs from the appendiceal or omental tissue.

We validated our findings using immunohistochemistry **(IHC)** for specific cell populations on sections of these tumors. To determine if some of the gene expression markers could be

detected as proteins, we conducted mass spectrometry on ascites samples from three PMP patients that validated a critical feature of the genomic analysis. Using an independent cohort, we confirmed the enrichment of cellular states identified in our discovery cohort using datasets from 63 PMP tumors [15].

Single cell RNA-seq profiling of AMN and PMP tumors

Samples included normal appendix, normal omentum, AMNs or PMPs. Using these tissue samples, the scRNA-seq provided a total of 299,718 single cells **(Supplemental Table 1)**. The next step in the data processing was to minimize potential batch effects on cell clustering. A k-means clustering procedure as part of the Harmony algorithm was used to correct for batch variance in the data [14]. Afterwards, we examined the cell clustering across the cohort of samples. When batch effect occurs, each cluster will have cells from a distinct patient tumor. For this study we observed that each cluster had cells from different patient tumors. This result indicated that using the Harmony algorithm provided adequate removal of batch effects **(Fig. 1B)**. We verified this computationally by calculating a similarity metric called the Adjusted Rand Index **(ARI)** [16]. Comparison of cluster assignments with experimental batch had an ARI of 0.05. The low, near-random value indicated that cluster assignments were independent of experimental batch. Overall, the cell clustering was the result of intrinsic tumor cell properties and not unduly influenced by batch effects.

For each cluster, we assigned cell lineages, also referred to as types, based on the expression of established marker genes, which we list after each cell type **(Fig. 1C - D)**. This marker gene analysis identified epithelial cells (*EPCAM*, *TFF3*, *MUC2*), stromal fibroblasts (*DCN*, *COL1A1*, *LUM*), endothelial cells (*VWF*, *PLVAP*, *PECAM1*), T cells (*CD3D*, *IL7R*, *CD8A*), NK cells

(*NKG7*, *GNLY*), B or plasma cells (*MS4A1*, CD79A), mast cells (*TPSAB1*, *KIT*) and cell lineages associated with macrophage or dendritic cells (*CD68*, *CD14*, *FCGR3A*, *HLA-DRA*). All cell types were identified in varying proportions across samples from normal tissue, AMN and PMP **(Fig. 1E)**. Then, we performed a secondary clustering analysis with batch-correction for each cell lineage – this step improved the characterization of each cell type's transcriptional phenotype **(Supplemental Methods)**.

Delineating the epithelial cell subtypes in AMNs and PMPs

We focused on determining the characteristic of the different epithelial subtypes present in these tumors. From the scRNA-seq data for AMNs, PMPs and normal appendix, we separated the epithelial cells from other cell types and conducted an independent clustering analysis **(Supplemental Fig. 1A)**. We used the samples containing at least 10 epithelial cells **(Supplemental Table 2)**. The threshold for cell counts was required for evaluating cell states and differential expression. We compared the gene expression of each individual cell to a reference single cell atlas of mucosal cell lineages of the human large intestine [17]. This step relied on the SingleR method which can identify cell types in a query dataset using cell-type specific reference transcriptomic datasets **(Supplemental Methods)** [18].

Next, we examined the expression of specific lineage marker genes to provide an additional verification of cell assignment. Among the epithelium there were stem cells (*LGR5*, *ASCL2*, *OLFM4*), goblet cells (*MUC2*, *TFF3*, *SPINK4*, *CLCA1*, *SPDEF*, *FCGBP*) and enterocytes (*CA2*, *SLC26A2*, *FABP1*) **(Fig. 2A)**. These diverse cell types were present throughout all samples. Rare cell types included tuft cells (*POU2F3*, *LRMP*, *TRPM5*) which are chemosensory cells in the epithelial lining of the intestine as well as neuroendocrine cells which were defined by their expression of chromogranin A (*CHGA*).

Among the AMN and PMP tumors, the majority of epithelial cell subtype were goblet cells **(Fig. 2B)**. To compare the composition of normal tissue versus AMN or PMP tumors, we conducted a statistical analysis that considered differential cell type abundance, modeling the cell type as a covariate factor **(Supplemental Methods)** [19, 20]. Compared to epithelial cells from normal appendiceal tissue, goblet cells were significantly more abundant in PMPs (adjusted P value 2.49E-50; Wald test) and AMNs (adjusted P value 1.21E-33; Wald test). Both AMN and PMP tumors also had epithelial stem cells. Notably, stem cells were significantly enriched in the AMN samples compared to normal tissue (adjusted P value 2.65E-17; Wald test) or PMP tumors (adjusted P value 1.93E-100; Wald test).

We analyzed the expression of *MUC2*, the predominant extracellular mucin detected in AMN and PMP. Compared to normal appendix cells, the AMN and PMP epithelial cells showed significantly higher *MUC2* expression compared to normal epithelium **(Supplemental Fig. 1B)**. The PMP cells had higher *MUC2* expression compared to the AMN cells (ANOVA FDR P-value < 2.2E-10). We observed a similar pattern of overexpression of a signature comprising genes encoding gel-forming mucins (*MUC2*, *MUC5B*, *MUC5A*, *MUC6*, *MUC19*) [21] **(Supplemental Fig. 1C)** in the PMP cells compared to the AMN cells.

Genomic instability in neoplastic goblet cells from AMNs and PMPs

To confirm that these epithelial cells were neoplastic, we characterized the extent of somatic copy number variations **(CNVs)** using the inferCNV program [22, 23]. The scRNA-seq data across all samples was processed for copy number analysis **(Supplemental Methods)**. Briefly, CNVs were estimated based on the gene expression of a windowed segment and an average expression level over the length of the entire chromosomal arm [24]. This approach identified

CNVs in the range of 10Mb or larger and helped us identify allelic imbalances covering the arms of chromosomes. Either gains or losses are identified. These events are indicators of cancer genomic instability. The somatic CNVs present in goblet cells from AMN and PMP tumor epithelium were compared to normal appendiceal epithelial cells and other non-epithelial cells in the TME.

For all AMN and PMP samples, the goblet epithelial tumor cells had multiple CNVs with some extending over entire chromosome arms **(Fig. 2C)**. Sixteen different chromosome segments had somatic CNVs. Thus, a significant portion of the genome was affected by these changes. In contrast, the normal epithelial cells from the appendix and other cell types had no CNVs or chromosomal imbalances. Large CNVs that included imbalances of entire chromosome arms are indicators of chromosomal instability. Their presence identified these goblet epithelial cells as neoplastic.

A gain of Chr 9q segment was the most common CNV across all of the samples including both AMNs and PMPs. This CNV was present in seven tumors. The aberrant activation of the NOTCH 1 receptor (chromosome 9) has been described in a cohort of patients with low and some high-grade mucinous appendiceal neoplasms [25]. The segments of chromosome arms 8q, 20q and 21q were the next most frequently observed with each being found in at least five tumors. The 8q and 21q gains were unique to the PMPs. We observed CNV gains in chromosome 1q and 20q that have been previously identified in mucinous neoplasms of the appendix [25, 26].

Four PMP patients (P8642, P8639, P8629 and P8605) had multiple metastatic tumors in different anatomic locations that underwent analysis with scRNA-seq. There was evidence of metastatic clonal divergence per an analysis of the tumors matched for each patient. For example, P8639 had three tumor sites including metastasis to the ovary, omentum and the original appendix. All of these tumors had a gain of a chr 21q segment. However, the ovarian metastasis also had a gain of 9q, indicating clonal divergence. P8629 had two metastases: the one to the mesentery had an exclusive gain of 9q compared to the omentum site which had a gain of 21q.

P8605 had three metastatic sites including the ovary, omentum and mesentery. For all three sites, common CNV segment gains included 8q, 16p and 20q. However, the remainder of CNVs tended to be unique to each tumor site. For example, the mesenteric site had unique gains in 1q, 6q and 15q that were not observed among the other sites. This result confirmed that clonal heterogeneity was present among the different tumor sites from the same patient.

Developmental trajectory and cell states among AMN and PMP epithelium

Changes in cell differentiation is a hallmark of cancer development and progression such as what may occur in AMNs as they progress to PMPs. This process requires a cascade of altered gene expression. The quantitative characteristics of these expression changes are referred to as a developmental 'trajectory' using algorithms like Monocle, which was employed in our study [27]. The measurement of this trajectory is referred to as pseudotime. We analyzed the single cell gene expression profiles to determine the pseudotime value and define the cell state. Subsequently, we generated a development order for the cell states and evaluated the specific

features of each state [27]. This analysis focused on the epithelial cells present in the AMN and PMP tumors.

The normal appendiceal epithelium defined the root of the trajectory – we refer to this as State 0 **(Fig. 3A)**. Branching from the root, we identified five other cellular states **(Fig. 3A, B)**. States 1 and 2 had lower pseudotime values; this result was indicative of these cells being in the early steps of development and epithelial differentiation. Both cellular states represented progenitor cells. States 3, 4 and 5 were located more distally on the trajectory with higher values, indicative of later stage of development.

We measured the fraction of epithelial cells for each cellular state across samples **(Fig. 3C, Supplemental Table 3)**. We examined if specific cell states are differentially abundant across normal, AMN or PMP tumors **(Supplemental Methods)**. Our results confirmed that State 3 was most common among epithelial cells from the AMN tumors compared to normal and PMP cells (adjusted P value <= 1.71E-7; Wald test). States 4 and 5 were significantly more common among the PMP cells compared to AMN cells (adjusted P value <= 4.62E-15; Wald test). For an individual tumor's epithelial cells, multiple cell states were present, indicative of intra-tumor heterogeneity. For example, all the tumors in the PMP cohort **(Fig. 3C)** had subpopulations from all four differentiation states in varying proportions.

Genes defining each cell state

To elucidate the cell state properties in greater detail, we performed differential expression and pathway analysis. The root state (i.e., State 0) was associated with normal epithelium and expressed several enterocyte markers (*FABP1*, *GUCA2A*, *CA2*, *ANPEP*, etc.) **(Supplemental**

Table 4). These markers defined the normal epithelial cells found in the appendix. State 1 had had cells with indicators of early differentiation. Cells in this state had high expression of the stem cell marker *OLFM4* as well as the chemokines *CCL20*, *CXCL3*, *CXCL14* and *CXCL1*. State 1 was associated with P8139's AMN tumor cells. Overall, State 1 had features indicative of a progenitor cell state. Cells in this state had a functional feature for secreting chemokines relevant for inflammation.

State 2 cells were present in the remaining AMN tumors and majority of PMP tumors. This state had a gene expression signature associated with progenitor cells **(Supplemental Table 4)**. This progenitor cell state included high expression of *SOX4* that regulates stem cell differentiation [28] and *ELF3* that promotes tumor growth [29]. These transcription factors have been associated with intestinal and goblet cell lineage differentiation, respectively [30, 31]. This state also demonstrated high expression of *SPDEF* and *MUC5AC* which regulate goblet cell phenotype [32]. *SPDEF* is an essential transcriptional factor that is involved in goblet cell differentiation [32]. Overall, these results suggest that the acquisition of goblet cell features is an early developmental event.

State 3 was enriched in AMN cells and State 4 was enriched in PMP cells **(Fig. 3C)**. Compared to the other cell states, States 3 and 4 had significantly higher expression of genes *TFF3*, *MUC5AC*, *MUC2*, *SPINK4*, *WFDC2* and *FCGBP* **(Fig. 3D, Supplemental Table 4)**. These genes are indicative of terminal goblet cell differentiation. Several genes were specific to State 4 including overexpression of *REG4* and tetraspanin genes such as *TM4SF1*, *CD9*, *CD63*, *TSPAN8*, *TSPAN13.* These genes are contributors to proliferation and have been associated with metastasis of cancer cells [33-35].

State 5 was widely represented among the PMPs, which had a higher proportion compared to the AMNs. This state had significant overexpression of genes involved in cancer metastasis and associated with poor prognosis. The differentially expressed genes included *ANXA1*, *FKBP4*, *LCN2*, *VNN1*, *TCN1*, *TACSTD2* (encoding TROP2), *DUOXA2*, *CAECAM5, CAECAM6* and members of the Kallikrein family genes (*KLK6*, *KLK8*, *KLK10* and *KLK11*) [36-47] **(Fig. 3D, Supplemental Table 4)**. State 5's gene expression signature was present in 11 of the PMP tumors among total of 14 (**Supplemental Fig. 2A, B**).

Notably, several genes had higher expression with statistical significance in State 5 compared to the other cell states **(Fig. 3D)**. These genes have not been described for PMP. Citing an example, Annexin A1 (*ANXA1*) is part of the Lipocortin family and a calcium-dependent phospholipid-binding protein. *ANXA1*'s functions include cellular proliferation, apoptosis, progression, and metastasis in cancer [48]. Elevated expression of *KLK8* and *FKBP4* was an exclusive feature of State 5 compared to all other states. *KLK8* has high expression in a variety of tumors including pancreatic and colorectal cancer. In experimental mouse models, KLK8 overexpression had both pro-proliferative and anti-apoptotic effects [49, 50]. The *FKBP4* gene is a member of the immunophilin protein family, which play a role in immunoregulation and basic cellular processes involving protein folding and trafficking. Studies have demonstrated *FKBP4* overexpression across a range of tumor types – its functional role in cancer is less well understood [51]. Regardless, both *KLK8* and *FKBP4* are candidate markers distinguishing the metastatic state of PMPs and may play an important role in their cancer biology.

Epithelial cell states and their defining biological pathways

For each cell state, we conducted a biological pathway analysis with the Gene Set Variation Analysis **(GSVA)** tool. The pathway annotation was based on the Hallmark gene sets – these groups of genes define specific biological pathways and can be assigned to a given cell state [52, 53]. States 1, 4 and 5 had significant upregulation of WNT pathway involved in stem cell maintenance, MYC signaling and cell cycle pathways (ANOVA with Tukey HSD p-value <= 3.6e-04) **(Fig. 3E)**. As noted previously, State 1 was an indicator of progenitor phenotype, State 4 an indicator of goblet cell differentiation in PMPs and State 5 an indicator of metastasis.

States 4 and 5 were also notable for the upregulation of several cancer biology pathways including those involving PI3K signaling, p53 function, KRAS activity and Notch pathways (ANOVA with Tukey HSD p-value <= 7.46e-06) **(Fig. 3E)**. State 5 had significant upregulation of intracellular signaling that contribute to epithelial mesenchymal transition **(EMT)** signaling. These pathways included increased signaling activity of TGF-β, JAK-STAT and angiogenesis signaling (ANOVA with Tukey HSD p-value <= 4.2e-08). In addition, these signaling pathways are associated with cancer development and progression. This result implicates pathways delineating State 5 that support PMP metastasis. These results overlap with what Levine et al. described in their gene expression studies [5, 11].

The regulatory genes underlying epithelial cell states

Next, we identified the regulatory genes that defined the transcriptional characteristics of these cell states **(Fig. 3F)**. We identified transcription factors that regulate the expression of downstream target genes. For this analysis we relied on the SCENIC algorithm. It identifies modules of interconnected genes containing the regulatory transcription factor and its target genes. These networked modules are called regulons [54].

For example, States 3 and 4 were associated with goblet cell differentiation - both states had significantly high activity of *SPDEF* (ANOVA with a Tukey HSD p-value = 2.68e-11). As we noted before, this transcription factor gene is a regulator of lineage commitment and differentiation into goblet cells [55]. The identification of this gene validated the results from the trajectory analysis, which showed that the tumor epithelial cells had a goblet cell identity.

There were other specific regulons observed in State 5, which was associated with PMP metastasis **(Fig. 3F)**. We identified regulatory genes including *TFAP2A*, *TEAD4* and *CD59*. These genes have been shown to promote EMT in lung and colon cancer, among other malignancies [56-58]. Overall, the regulon analysis identified a series of regulatory genes that contribute to metastatic PMPs.

Heterogeneity among multiple site PMP metastases and their cell states

Using the results of the trajectory analysis, we examined the cellular heterogeneity for the subset of patients with multiple metastatic samples. The tumors originated from different abdominal sites. Three PMP patients (P8605, P8629, P8639) had tissues from different metastatic sites in the abdomen and an adequate number of cells (10 or greater) for this analysis. As noted previously, CNV alterations indicated that there was clonal divergence among these samples. Therefore, it was of interest to see if the cell states also had heterogeneity.

For the samples with adequate cell numbers, we determined the differences in cell states among these metastatic sites per patient, performed differential expression analysis for intra-sample variation and identified the specific genes that defined different metastatic sites.

When we compared the different tumor sites from a given patient, each metastasis had varying proportion of different cell state **(Fig. 3C)**. For example, P8605's ovarian metastasis had a higher proportion of cells in State 2 compared to the mesenteric site. This correlated with the observation that a 14q CNV gain was identified only in the ovarian metastasis (**Fig. 2C**). The inverse was observed for State 5 where the mesenteric site had a higher proportion of these cells versus the ovary metastasis. The mesenteric metastasis had unique gains of 1q, 6q and 15q.

We examined the gene expression differences among the metastatic sites. As noted previously, P8605 had three metastases located in the mesentery, the omentum and the ovary. Each site had distinct gene expression signatures. For the metastasis in the mesentery, we observed significantly increased expression of metastasis-related genes *ANXA1* and *KLK8* compared to the omental and ovarian implants (adjusted p-value < 0.05) **(Supplemental Fig. 3A)**. The omental metastasis had a higher expression of the goblet cell differentiation markers *FCGBP* and *TFF3*. The ovarian site had increased expression of the progenitor state marker *SOX4*. The three lesions also differed in their expression of *GNAS*, a frequently mutated gene in PMP that regulates mucin production [59].

Patient P8629 had two metastatic sites involving the omentum and mesentery. The omental metastasis had a higher expression (adjusted p-value < 0.05) of the goblet cell differentiation

markers *TFF3*, *FCGBP*, *MUC2* and *SPINK4* **(Supplemental Fig. 3B)**. Conversely, the mesentery lesion had enriched expression of the progenitor markers *ELF3* and *GNAS*.

Patient P8639 had two metastatic sites from the omentum and ovary. In addition, we had scRNA-seq data from this patient's primary appendiceal cancer. The metastatic sites had differential expression of the following genes: *TCN1*, the transcobalamin gene, has been associated with poor prognosis and metastasis in gastric and colon cancer [38, 60]; *WFDC2* is a marker for goblet cells; *GNAS* frequently has mutations in appendiceal cancers **(Supplemental Fig. 3C)**. *TCN* had notably higher expression than omental or ovary metastasis. *WFDC2* expression was higher in the omental metastasis compared to the primary site and other metastasis. Altogether, this result confirms that for each of these patients, comparing the PMP metastasis sites showed that they had different properties of stemness, goblet cell identity and metastasis differentiation.

Protein markers present in PMP tumors

One of our major findings was that the tumor epithelial cells have a cell state defined by a goblet cell phenotype. To determine if there were specific markers of goblet cell and metastatic differentiation cell states, we used a combination of immunohistochemistry **(IHC)** and mass spectrometry on a subset of the PMP tumors (P8629, P8639 and P8605) and ascites (P5853, P8605 and P8629) respectively. For the IHC study, we used antibody markers (MUC2 and AGR2) that defined goblet calls. We also examined the protein expression of ANXA1. As we described, this gene was enriched in State 5 that defined PMP metastasis.

For the three PMPs (P8629, P8639 and P8605), MUC2 and AGR2 demonstrated intense staining among the epithelial tumor cells **(Fig. 4A)**. Likewise, ANXA1 was detected in all PMPs albeit with less intense staining than MUC2. Altogether, our results show that all markers were expressed among the tumor epithelial cells from the metastasis in the mesentery (P8629), the ovary (P8639) and the mesentery (P8605).

Malignant ascites samples were available from three PMP patients (P5853, P8605 and P8629). We determined if the secretory products of goblet cells were secreted into the ascites using mass-spectrometry **(Supplemental Methods)**. Overall, we identified 299 proteins in P5853, 300 in P8605 and 276 in P8629 **(Supplemental Table 5)**. The results were reported using a base-10 logarithm of the p-value for each protein [61]. This statistical value represents the likelihood of the peptide spectrum match being random. The mass spectrometry results revealed a number of enriched proteins based on a significance threshold in which the protein p-value is two-fold lower than the negative control decoy proteins [61].

Across all samples, we detected significant elevated expression of the mucin proteins MUC2, MUC5AC and MUC5B as well as goblet cell markers AGR2 and FCGBP **(Fig. 4B)**. MUC2 had the most significant log p-value. All three patients' ascites had detectable ANXA1 levels. From the trajectory analysis, we had identified *ANXA1* as a marker of State 5, which is associated with metastasis **(Fig. 3D, Supplemental Table 4)**. P5853's ascites had goblet cell secretory products TFF3 and CLCA1 [8]. In summary, these protein analysis results support a goblet cell identity for the PMP tumors. More importantly, our results demonstrate a direct role for these cells in the secretion of their intracellular products into malignant ascites.

Evidence of PMP cell states in independent cohorts

We validated our scRNA-seq findings using independent gene expression data (Affymetrix microarrays) from 63 PMPs [5]. We performed gene set variation analysis **(GSVA)** [52] on this expression data. For this analysis, we used the expression signatures representing the different cell states **(Fig. 5A)**. For each cell state, we identified the differentially expressed genes with statistical significance (adjusted p-value <= 0.05, log fold change > 0.25) **(Supplemental Table 4)**. Next, we aggregated the states that were overlapped and were closely related. States 1 and 2 represented the progenitor state. States 3 and 4 represented the goblet cell state. We examined State 5, associated with PMP metastasis and State 0 (i.e., root) separately. The GSVA analysis provided a quantitative enrichment score, which indicates whether the PMPs expressed the gene expression signature of any given cell state.

The majority of samples had positive GSVA enrichment scores for the progenitor state (61.9%), the goblet cell differentiation state (58.7%) and the metastatic state (57.1%) **(Fig. 5B)**. The root state was detected in 49.2% samples. Hence, cell states from our single-cell discovery cohort were detected in independent PMP cohorts. The low cellularity of these tumors makes it challenging to detect tumor epithelium specific features from conventional bulk sequencing assays. This result indicates that having identified granular tumor epithelial cell states using scRNA-seq analysis, we can identify them in bulk datasets using their gene expression signatures. Furthermore, the goblet state enrichment scores were significantly correlated with metastasis **(Fig. 5C)** and progenitor **(Fig. 5D)** scores across all samples. This result confirmed the co-existence of heterogenous tumor epithelial cell states within a PMP tumor, which we identified using trajectory analysis in the discovery cohort.

Infiltrating immune cells in the AMN and PMP microenvironment

We characterized the lymphocytes in the TME of the AMNs and PMPs. Our major focus was determining if these tumors contain infiltrating lymphocytes. Initially, we annotated lymphocytes by comparing individual cell transcriptomes to a reference pan-cancer atlas of tumor infiltrating immune cells and verified assignments based on the expression of marker genes **(Supplemental Methods)** [62].

We identified B and plasma cells (*MS4A1*, *CD19*, *etc.*) **(Fig. 6A)**. In addition, we detected T cell subsets such as naïve-like cells (*CCR7*, *SELL*), regulatory T cells **(Tregs)** (*FOXP3*, *IL2RA*, *etc.*) and follicular helper-like cells **(TFh-like)** (*CXCL13*, *TNFRSF18*, *CTLA4, etc.*). We identified cytotoxic CD8 T cells (*CD8A*, *CD8B*, *CCL5*, *etc.*) and a subset of exhausted CD8 T cells (*CD8A*, *CD8B*, *CXCL13*, *PDCD1*, *CTLA4*, *etc.*). We also detected NK cells (*KLRD1*, *KLRF1*, *GNLY*, *etc.*) [62, 63].

Cell subsets were detected across normal, AMN and PMP samples (**Fig. 6B**). This indicates that AMN and PMP tumors contain a rich lymphocytic infiltrate that includes B cells, TFh-like CD4 T cells, cytotoxic CD8 T cells and NK cells. While the absolute number of cells varied by sample **(Supplemental Table 6)**, all tumors contained TFh-like CD4 T cells, cytotoxic CD8 T and NK cells. Very few cells mapped to the exhausted CD8 phenotype. The TFh-like cells have been linked to anti-tumor immune response via the generation of tertiary lymphoid structures [64, 65]. Both cytotoxic CD8 and NK cells had effector cytokine expression (*GZMA*, *GZMB*, *GZMK*, *IFNG*, *PRF1*) **(Fig. 6A)**. These effectors molecules are required to orchestrate an anti-tumor immune response [63]. Overall, these results suggest that the AMN and PMPs TME is not “cold” and contains infiltrating lymphocytes with potential anti-tumor activity.

DISCUSSION

The cellular characteristics of AMNs and their progression to PMP have been a challenge to study due to the rarity of these tumors, their low cellularity and their complex gelatinous composition. The epithelial cell population is often diluted among the other cell types and mucinous components. To overcome the challenges of determining their cellular features, we used scRNA-seq on a set of AMNs and PMPs. Our results identified and characterized the cellular features of low abundance tumor epithelial cells present in gelatinous PMP tumors.

An important aspect of our study was determining the developmental origin and trajectory of goblets cells [66]. The discovery of MUC2 expression implicated an enteric rather than ovarian origin for PMP tumors [6, 67]. That being said, prior studies have not firmly established whether the detection of MUC2 was a result of acquired expression of non-gastrointestinal tumor cells and/or cells with an intestinal lineage origin. Our analysis conclusively established a goblet cell identity of intestinal origin. Moreover, we observed tumor cells with goblet cell features in peritoneal metastases (e.g., omentum or peritoneum). These tissues normally lack intestinal epithelial cells and thus confirmation of their presence implicated the intestinal origin of these cells. The genomic and proteomic markers of goblet cell identity may be useful in PMP diagnosis.

We identified the trajectory of goblet cell development and differentiation in both AMN and PMP. This process included progenitor cell states with further differentiation into a metastatic cell state. Notably, PMPs had a specific cell state associated with metastasis and dysregulated pathways such as EMT activity, an indicator of their high metastatic potential. Using an

independent gene expression data from a prior study of PMPs, we validated the signatures defining these states [5]. Importantly, we identified new genes that were associated with this metastatic cell state including *ANXA1*, *KLK8* and *FKBP4*. These genes have been implicated as contributors to cancer biology and metastasis among other tumor types but not in appendiceal carcinomas. Further studies of these genes may provide greater insight into the biology of PMP metastasis.

Goblet secretory cell fate in the gastrointestinal tract has been proposed to be specified by WNT signaling following NOTCH inhibition[68]. Surprisingly, our trajectory analysis revealed NOTCH pathway activation in States 4 and 5. Interestingly, conditional NOTCH activation of post-mitotic epithelium in the small intestine of mice resulted in goblet cell fate [69]. Furthermore, a chromosome 9 germline mutation resulting in *NOTCH1* activation has been reported in patients with AMN. Our single-cell analysis of copy number changes identified amplification in chromosome 9 among most of the patient tumors. Lastly, mutations in KRAS, one of the two most commonly mutated genes in AMN, resulted in HES1 (NOTCH target) activation and goblet hyperplasia in a mouse model of colorectal cancer [26, 70]. Overall, these findings suggest the role of NOTCH signaling in goblet cell differentiation and the development of AMN. In addition, NOTCH activation in appendiceal mucinous tumors represents a novel target as ongoing clinical trials continue to study the efficacy of NOTCH inhibitors in patient with solid tumors.

Gastrointestinal goblet cells generate a protective mucosal barrier composed of glycoproteins and mucin [8]. In addition to the goblet cell gene expression features, we detected secretory products of goblet cells in ascites from PMP patients. The mass spectrometric analysis of patient derived ascites identified MUC2, AGR2 and MUC5AC in all the tested patients. This is consistent with prior IHC and western blot assays detecting MUC2 and MUC5AC in PMP tissue

and ascites [71, 72]. Furthermore, we detected a heterogenous composition of mucins in our scRNA-seq analysis for patients in the AMN and PMP cohorts, with *MUC2* and *MUC5AC* detected consistently in all patients. This study identified goblet cell specific secretory products in patient derived ascites that can be traced to the transcriptome of tumor cells. Altogether, these results show that the goblet tumor cells contribute to the mucinous phenotype of this cancer.

Current mainstem therapy for PMP patients consists of cytoreductive surgery and heated intra-peritoneal chemotherapy (HIPEC) using mitomycin C as the perfusate. Systemic chemotherapy is used in suboptimal surgical candidates or as palliation [73]. This study identified signaling pathways that point towards potential targets including: tetraspanin and kallikrein family genes, MYC, PI3K-AKT-mTOR and RAS signaling pathways. In support, genomic studies have shown, at lower frequency, mutations in *SMAD4*, *ATM*, *PIK3CA*, *AKT* and *JAK3* in AMN and PMP tissues [74]. Furthermore, we also identified considerable heterogeneity in the abundance of tumor cell states within the same tumor and at different metastatic sites in the same PMP patient. The observed heterogeneity may be a source of resistance to both existing chemotherapy regimens and may guide future tailored targeted therapy to prevent resistance in patients.

Within the AMN and PMP's immune TME, both tumor types contained both CD4 and CD8 sub-populations linked to anti-tumor immune responses [64, 65]. This result is supported by a recent study that identified immune enriched TME in a subset of PMP patients using spatial gene expression analysis [75]. Prior experiments have shown that PMP tumors respond to immune checkpoint blockade or cancer neoantigen vaccination based on a common mutation in *GNAS* [76, 77]. Despite the presence of these immune cells, their anti-tumor activity could be

hindered by the presence of gelatinous mucin in PMP [78]. Therefore, increasing their cytotoxicity via immunotherapy in conjunction with reducing the mucin barrier via mucolytics or surgical debulking is a potential way of improving anti-tumor response from these cells [79]. This could be particularly advantageous for PMP that contains widespread disseminated lesions and for those patients that show limited response to current debulking and HIPEC treatment regimens.

DISCLOSURE OF POTENTIAL CONFLICTS OF INTEREST

None to disclose.

DATA AVAILABILITY

Sequencing data will be released under dbGAP identifier phs001818. Single cell data objects will be available on https://dna-discovery.stanford.edu/research/datasets/.

AUTHORS' CONTRIBUTIONS

CA and AS were involved in conception and design of the study, development of methodology, acquisition of data, analysis and interpretation of data and writing of the manuscript. XB, SG, and JS were involved in the analysis and interpretation of data. GAP and BL were involved in design of the study and clinical translational component. HPJ oversaw the conception and design of the study, data analysis, interpretation of data and writing of the manuscript.

ACKNOWLEDGEMENTS

We are grateful to all patients who participated in the study. This work was supported by the Don and Ruth Seiler Fund. Other support came from US National Institutes of Health grants R01HG006137 (HPJ) and U01CA217875 (HPJ, AS). HPJ also received support from the Clayville Foundation.

Table 1. AMN and PMP samples

Tumor type	Patient ID	Tissue sites for single-cell analysis	No. of samples for single-cell analysis	Mass spectrometry on ascitic fluid
AMN	P8139	Appendix tumor	1	No
AMN	P8142	Normal appendix	2	No
AMN	P8153	Appendix tumor and normal	2	No
AMN	P8333	Appendix tumor and normal	2	No
AMN	P8617	Appendix tumor and normal	2	No
PMP	P8336	Peritoneum tumor	1	No
PMP	P5853	Peritoneum tumor	1	Yes
PMP	P8587	Omentum tumor	1	No
PMP	P8605	Appendix mesentery, omentum and ovary tumors	3	Yes
PMP	P8606	Appendix and small bowel mesentery tumors, normal omentum	3	No
PMP	P8620	Appendix and omentum tumors, normal omentum	3	No
PMP	P8629	Appendix mesentery and small bowel mesentery tumors	2	Yes
PMP	P8639	Appendix, omentum, ovary and appendix mesentery tumors	4	No
PMP	P8642	Appendix mesentery, omentum, liver capsule tumors, normal omentum	4	No

AMN: Appendiceal mucinous neoplasm, PMP: pseudomyxoma peritonei

FIGURE LEGENDS

Figure 1. (A) Schematic representation of study design. (B-C) UMAP representation of dimensionally reduced data following batch correction and graph-based clustering of all datasets annotated by (B) samples and (C) cell lineages. (D) Dot plot depicting average expression levels of specific lineage-based marker genes together with the percentage of cells expressing the marker. (E) Proportion of cell lineages detected from each sample annotated by normal, AMN and PMP designation.

Figure 2. (A) Heatmap depicting expression of lineage marker genes following reference-based assignment of tumor and normal epithelial cells. (B) Proportion of epithelial cell lineages detected from each sample annotated by normal, AMN and PMP designation. (C) Heatmap representation of inferred single-cell CNV signatures of goblet tumor epithelial cells per sample compared to reference cells from normal epithelium, immune and stromal cells. CNV legend: state 1 = complete loss, state 2 = loss of one copy, state 3 = neutral, state 4 = addition of one copy, state 5 = addition of two copies, state 6 = addition of more than 2 copies.

Figure 3. (A-B) Trajectory plots of epithelial cells colored by their (A) pseudotime values and additionally annotated by state, (B) condition. (C) Proportion of trajectory states within epithelial cells detected from each sample annotated by normal, AMN and PMP designation. (D) Violin plots indicating expression of selected genes in each trajectory branch. * indicates cell state with significant expression (adjusted p-value < 0.05) for that gene. (E) Heatmap depicting average GSVA enrichment score of selected Hallmark pathways per trajectory branch (ANOVA Tukey HSD p-value < 0.05). (F) Heatmap depicting average Area Under the Curve value of respective

regulon per trajectory branch (ANOVA Tukey HSD p-value < 0.05). Numbers in parentheses indicate the number of genes in the examined gene regulatory network.

Figure 4. (A) Representative images of immunohistochemical staining of respective tumor samples for MUC2, AGR2, ANXA1 or representative isotype control. Scale bar = 50 μm. (B) Selected proteins detected in mass spectrometry of PMP ascites samples with their detection p-value.

Figure 5. (A) Schematic representation of signature enrichment analysis. (B) GSVA scores for respective tumor epithelial cell states in validation bulk dataset samples. (C-D) Scatter plots showing Spearman correlation of respective tumor epithelial cell state signatures across all samples from the validation bulk datasets.

Figure 6. (A) Heatmap depicting average expression levels of respective genes in lymphocyte subsets. (B) UMAP representation of dimensionally reduced data following batch correction and graph-based clustering of all datasets annotated by lymphocyte subsets across normal, AMN and PMP samples.

SUPPLEMENTAL DATA

Supplemental Methods. Format: PDF

Supplemental Figure S1-S3. Format: PDF

Supplemental Tables S1 – S6. Format: XLSX

REFERENCES

1. Choudry, H.A. and R.K. Pai, *Management of Mucinous Appendiceal Tumors.* Ann Surg Oncol, 2018. **25**(8): p. 2135-2144.
2. Valle, S.J., et al., *Hyperthermic Intraperitoneal Chemotherapy (HIPEC) Methodology, Drugs and Bidirectional Chemotherapy.* Indian J Surg Oncol, 2016. **7**(2): p. 152-9.
3. Bartlett, D.J., et al., *Mucinous appendiceal neoplasms: classification, imaging, and HIPEC.* Abdom Radiol (NY), 2019. **44**(5): p. 1686-1702.
4. Isella, C., et al., *Improved Outcome Prediction for Appendiceal Pseudomyxoma Peritonei by Integration of Cancer Cell and Stromal Transcriptional Profiles.* Cancers (Basel), 2020. **12**(6).
5. Levine, E.A., et al., *Prognostic Molecular Subtypes of Low-Grade Cancer of the Appendix.* J Am Coll Surg, 2016. **222**(4): p. 493-503.
6. O'Connell, J.T., et al., *Pseudomyxoma peritonei is a disease of MUC2-expressing goblet cells.* Am J Pathol, 2002. **161**(2): p. 551-64.
7. Birchenough, G.M., et al., *New developments in goblet cell mucus secretion and function.* Mucosal Immunol, 2015. **8**(4): p. 712-9.
8. Kim, Y.S. and S.B. Ho, *Intestinal goblet cells and mucins in health and disease: recent insights and progress.* Curr Gastroenterol Rep, 2010. **12**(5): p. 319-30.
9. Choudry, H.A., et al., *Mucin as a therapeutic target in pseudomyxoma peritonei.* J Surg Oncol, 2012. **106**(7): p. 911-7.
10. Carr, N.J., et al., *The histopathological classification, diagnosis and differential diagnosis of mucinous appendiceal neoplasms, appendiceal adenocarcinomas and pseudomyxoma peritonei.* Histopathology, 2017. **71**(6): p. 847-858.

11. Levine, E.A., et al., *Gene expression profiling of peritoneal metastases from appendiceal and colon cancer demonstrates unique biologic signatures and predicts patient outcomes.* J Am Coll Surg, 2012. **214**(4): p. 599-606; discussion 606-7.

12. Butler, A., et al., *Integrating single-cell transcriptomic data across different conditions, technologies, and species.* Nat Biotechnol, 2018. **36**(5): p. 411-420.

13. Hafemeister, C. and R. Satija, *Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression.* Genome Biol, 2019. **20**(1): p. 296.

14. Korsunsky, I., et al., *Fast, sensitive and accurate integration of single-cell data with Harmony.* Nat Methods, 2019. **16**(12): p. 1289-1296.

15. Newman, A.M., et al., *Determining cell type abundance and expression from bulk tissues with digital cytometry.* Nat Biotechnol, 2019. **37**(7): p. 773-782.

16. Hubert, L. and P. Arabie, *Comparing partitions.* Journal of Classification, 1985. **2**(1): p. 193-218.

17. Elmentaite, R., et al., *Cells of the human intestinal tract mapped across space and time.* Nature, 2021. **597**(7875): p. 250-255.

18. Aran, D., et al., *Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage.* Nat Immunol, 2019. **20**(2): p. 163-172.

19. Ramachandran, P., et al., *Resolving the fibrotic niche of human liver cirrhosis at single-cell level.* Nature, 2019. **575**(7783): p. 512-518.

20. Haber, A.L., et al., *A single-cell survey of the small intestinal epithelium.* Nature, 2017. **551**(7680): p. 333-339.

21. Nguyen, B., et al., *The genomic landscape of carcinomas with mucinous differentiation.* Sci Rep, 2021. **11**(1): p. 9478.

22. Patel, A.P., et al., *Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma.* Science, 2014. **344**(6190): p. 1396-401.

23. Tirosh, I., et al., *Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq.* Science, 2016. **352**(6282): p. 189-96.

24. Tickle, T., et al. *inferCNV of the Trinity CTAT Project*. 2019 [cited 2020; Available from: https://github.com/broadinstitute/inferCNV.

25. LaFramboise, W.A., et al., *Discrimination of low- and high-grade appendiceal mucinous neoplasms by targeted sequencing of cancer-related variants.* Mod Pathol, 2019. **32**(8): p. 1197-1209.

26. Alakus, H., et al., *Genome-wide mutational landscape of mucinous carcinomatosis peritonei of appendiceal origin.* Genome Med, 2014. **6**(5): p. 43.

27. Qiu, X., et al., *Reversed graph embedding resolves complex single-cell trajectories.* Nat Methods, 2017. **14**(10): p. 979-982.

28. Moreno, C.S., *SOX4: The unappreciated oncogene.* Semin Cancer Biol, 2020. **67**(Pt 1): p. 57-64.

29. Wang, H., et al., *Overexpression of ELF3 facilitates cell growth and metastasis through PI3K/Akt and ERK signaling pathways in non-small cell lung cancer.* Int J Biochem Cell Biol, 2018. **94**: p. 98-106.

30. Gracz, A.D., et al., *Sox4 Promotes Atoh1-Independent Intestinal Secretory Differentiation Toward Tuft and Enteroendocrine Fates.* Gastroenterology, 2018. **155**(5): p. 1508-1523 e10.

31. Ng, A.Y., et al., *Inactivation of the transcription factor Elf3 in mice results in dysmorphogenesis and altered differentiation of intestinal epithelium.* Gastroenterology, 2002. **122**(5): p. 1455-66.

32. Noah, T.K., et al., *SAM pointed domain ETS factor (SPDEF) regulates terminal differentiation and maturation of intestinal goblet cells.* Exp Cell Res, 2010. **316**(3): p. 452-65.

33. Heo, K. and S. Lee, *TSPAN8 as a Novel Emerging Therapeutic Target in Cancer for Monoclonal Antibody Therapy.* Biomolecules, 2020. **10**(3).

34. Fu, F., et al., *Role of Transmembrane 4 L Six Family 1 in the Development and Progression of Cancer.* Front Mol Biosci, 2020. **7**: p. 202.

35. Zhang, J., et al., *The Clinical Significance and Mechanisms of REG4 in Human Cancers.* Front Oncol, 2020. **10**: p. 559230.

36. Chai, C.Y., et al., *VNN1 overexpression is associated with poor response to preoperative chemoradiotherapy and adverse prognosis in patients with rectal cancers.* Am J Transl Res, 2016. **8**(10): p. 4455-4463.

37. Liang, Z. and X. Li, *Identification of ANXA1 as a potential prognostic biomarker and correlating with immune infiltrates in colorectal cancer.* Autoimmunity, 2021. **54**(2): p. 76-87.

38. Liu, G.J., et al., *High expression of TCN1 is a negative prognostic biomarker and can predict neoadjuvant chemosensitivity of colon cancer.* Sci Rep, 2020. **10**(1): p. 11951.

39. Liu, X., et al., *Elevated expression of KLK8 predicts poor prognosis in colorectal cancer.* Biomed Pharmacother, 2017. **88**: p. 595-602.

40. Mange, A., et al., *FKBP4 connects mTORC2 and PI3K to activate the PDK1/Akt-dependent cell proliferation signaling in breast cancer.* Theranostics, 2019. **9**(23): p. 7003-7015.

41. Meng, W., et al., *FKBP4 Accelerates Malignant Progression of Non-Small-Cell Lung Cancer by Activating the Akt/mTOR Signaling Pathway.* Anal Cell Pathol (Amst), 2020. **2020**: p. 6021602.

42. Nugteren, S. and J.N. Samsom, *Secretory Leukocyte Protease Inhibitor (SLPI) in mucosal tissues: Protects against inflammation, but promotes cancer.* Cytokine Growth Factor Rev, 2021. **59**: p. 22-35.

43. Santiago-Sanchez, G.S., et al., *Biological Functions and Therapeutic Potential of Lipocalin 2 in Cancer.* Int J Mol Sci, 2020. **21**(12).

44. Shvartsur, A. and B. Bonavida, *Trop2 and its overexpression in cancers: regulation and clinical/therapeutic implications.* Genes Cancer, 2015. **6**(3-4): p. 84-105.

45. Wei, Z., et al., *Targeting secretory leukocyte protease inhibitor (SLPI) inhibits colorectal cancer cell growth, migration and invasion via downregulation of AKT.* PeerJ, 2020. **8**: p. e9400.

46. Zhu, X., et al., *Clinicopathological Analysis and Prognostic Assessment of Transcobalamin I (TCN1) in Patients with Colorectal Tumors.* Med Sci Monit, 2020. **26**: p. e923828.

47. Gebauer, F., et al., *Carcinoembryonic antigen-related cell adhesion molecules (CEACAM) 1, 5 and 6 as biomarkers in pancreatic cancer.* PLoS One, 2014. **9**(11): p. e113023.

48. Araujo, T.G., et al., *Annexin A1 as a Regulator of Immune Response in Cancer.* Cells, 2021. **10**(9).

49. Hua, Q., et al., *Upregulation of KLK8 Predicts Poor Prognosis in Pancreatic Cancer.* Front Oncol, 2021. **11**: p. 624837.

50. Hua, Q., et al., *KLK8 promotes the proliferation and metastasis of colorectal cancer via the activation of EMT associated with PAR1.* Cell Death Dis, 2021. **12**(10): p. 860.

51. Xiong, H., et al., *FKBP4 is a malignant indicator in luminal A subtype of breast cancer.* J Cancer, 2020. **11**(7): p. 1727-1736.

52. Hänzelmann, S., R. Castelo, and J. Guinney, *GSVA: gene set variation analysis for microarray and RNA-seq data.* BMC Bioinformatics, 2013. **14**: p. 7.

53. Liberzon, A., et al., *The Molecular Signatures Database (MSigDB) hallmark gene set collection.* Cell Syst, 2015. **1**(6): p. 417-425.

54. Aibar, S., et al., *SCENIC: single-cell regulatory network inference and clustering.* Nat Methods, 2017. **14**(11): p. 1083-1086.

55. Gregorieff, A., et al., *The ets-domain transcription factor Spdef promotes maturation of goblet and paneth cells in the intestinal epithelium.* Gastroenterology, 2009. **137**(4): p. 1333-45 e1-3.

56. Xiong, Y., et al., *TFAP2A potentiates lung adenocarcinoma metastasis by a novel miR-16 family/TFAP2A/PSG9/TGF-beta signaling pathway.* Cell Death Dis, 2021. **12**(4): p. 352.

57. Liu, Y., et al., *Increased TEAD4 expression and nuclear localization in colorectal cancer promote epithelial-mesenchymal transition and metastasis in a YAP-independent manner.* Oncogene, 2016. **35**(21): p. 2789-800.

58. Geller, A. and J. Yan, *The Role of Membrane Bound Complement Regulatory Proteins in Tumor Development and Cancer Immunotherapy.* Front Immunol, 2019. **10**: p. 1074.

59. Lin, Y.L., R. Ma, and Y. Li, *The biological basis and function of GNAS mutation in pseudomyxoma peritonei: a review.* J Cancer Res Clin Oncol, 2020. **146**(9): p. 2179-2188.

60. Zhu, X., et al., *Clinicopathological Analysis and Prognostic Assessment of TCN1 in Patients with Gastric Cancer.* Surg Innov, 2021: p. 15533506211045318.

61. Bern, M., Y.J. Kil, and C. Becker, *Byonic: Advanced Peptide and Protein Identification Software.* Current Protocols in Bioinformatics, 2012. **40**(1): p. 13.20.1-13.20.14.

62. Nieto, P., et al., *A single-cell tumor immune atlas for precision oncology.* Genome Res, 2021. **31**(10): p. 1913-1926.

63. van der Leun, A.M., D.S. Thommen, and T.N. Schumacher, *CD8(+) T cell states in human cancer: insights from single-cell analysis.* Nat Rev Cancer, 2020. **20**(4): p. 218-232.

64. Zheng, L., et al., *Pan-cancer single-cell landscape of tumor-infiltrating T cells.* Science, 2021. **374**(6574): p. abe6474.

65. Zhang, Y., et al., *Single-cell analyses reveal key immune cell subsets associated with response to PD-L1 blockade in triple-negative breast cancer.* Cancer Cell, 2021. **39**(12): p. 1578-1593 e8.

66. Valasek, M.A. and R.K. Pai, *An Update on the Diagnosis, Grading, and Staging of Appendiceal Mucinous Neoplasms.* Adv Anat Pathol, 2018. **25**(1): p. 38-60.

67. O'Connell, J.T., C.M. Hacker, and S.H. Barsky, *MUC2 is a molecular marker for pseudomyxoma peritonei.* Mod Pathol, 2002. **15**(9): p. 958-72.

68. Qi, Z. and Y.G. Chen, *Regulation of intestinal stem cell fate specification.* Sci China Life Sci, 2015. **58**(6): p. 570-8.

69. Zecchini, V., et al., *Notch signaling regulates the differentiation of post-mitotic intestinal epithelial cells.* Genes Dev, 2005. **19**(14): p. 1686-91.

70. Feng, Y., et al., *Mutant KRAS promotes hyperplasia and alters differentiation in the colon epithelium but does not expand the presumptive stem cell pool.* Gastroenterology, 2011. **141**(3): p. 1003-1013 e1-10.

71. Akashi, S., et al., *Mucinous Balls Tangled With Mesothelial Cells and MUC2-Positive Cancer Cells in the Ascites of Pseudomyxoma Peritonei.* Diagn Cytopathol, 2016. **44**(7): p. 628-31.

72. Mall, A.S., et al., *MUC2, MUC5AC and MUC5B in the mucus of a patient with pseudomyxoma peritonei: biochemical and immunohistochemical study.* Pathol Int, 2007. **57**(8): p. 537-47.

73. Lin, Y.L., et al., *Consensuses and controversies on pseudomyxoma peritonei: a review of the published consensus statements and guidelines.* Orphanet J Rare Dis, 2021. **16**(1): p. 85.

74. Stein, A., et al., *Molecular and Genetic Markers in Appendiceal Mucinous Tumors: A Systematic Review.* Ann Surg Oncol, 2020. **27**(1): p. 85-97.

75. Su, J., et al., *Prognostic Molecular Classification of Appendiceal Mucinous Neoplasms Treated with Cytoreductive Surgery and Hyperthermic Intraperitoneal Chemotherapy.* Ann Surg Oncol, 2020. **27**(5): p. 1439-1447.

76. Forsythe, S.D., et al., *Organoid Platform in Preclinical Investigation of Personalized Immunotherapy Efficacy in Appendiceal Cancer: Feasibility Study.* Clin Cancer Res, 2021. **27**(18): p. 5141-5150.

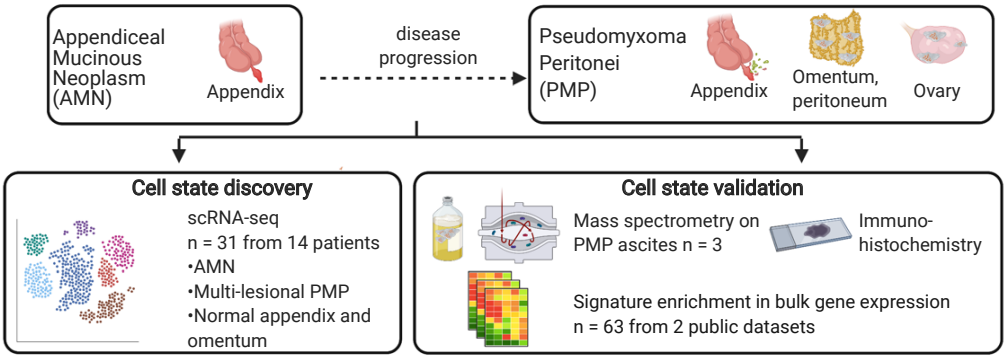
77. Flatmark, K., et al., *Peptide vaccine targeting mutated GNAS: a potential novel treatment for pseudomyxoma peritonei.* J Immunother Cancer, 2021. **9**(10).

78. Sommariva, A., et al., *Novel Perspectives in Pseudomyxoma Peritonei Treatment.* Cancers (Basel), 2021. **13**(23).

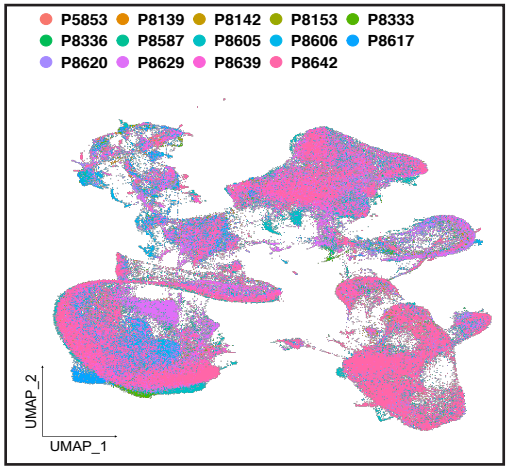
79. Robert, C., *A decade of immune-checkpoint inhibitors in cancer therapy.* Nat Commun, 2020. **11**(1): p. 3801.

Figure 1

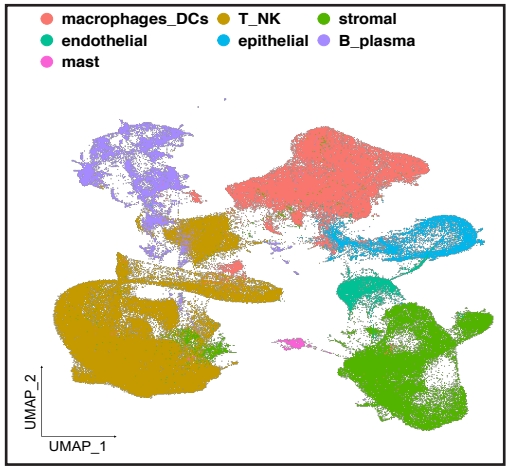
A



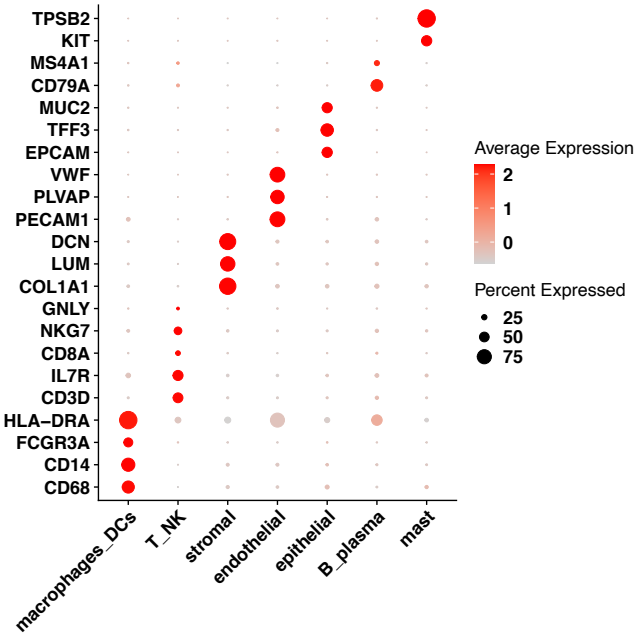
B



C



D



E

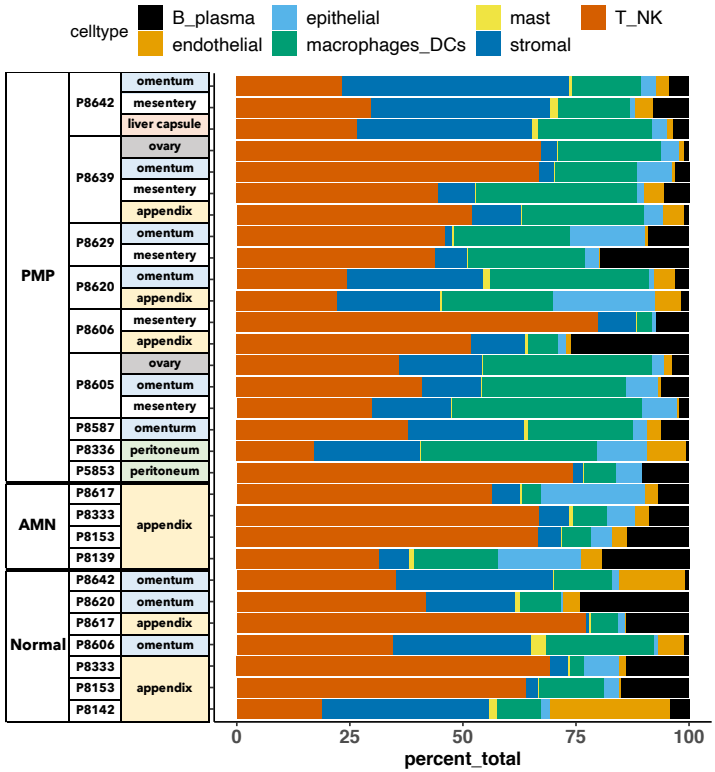
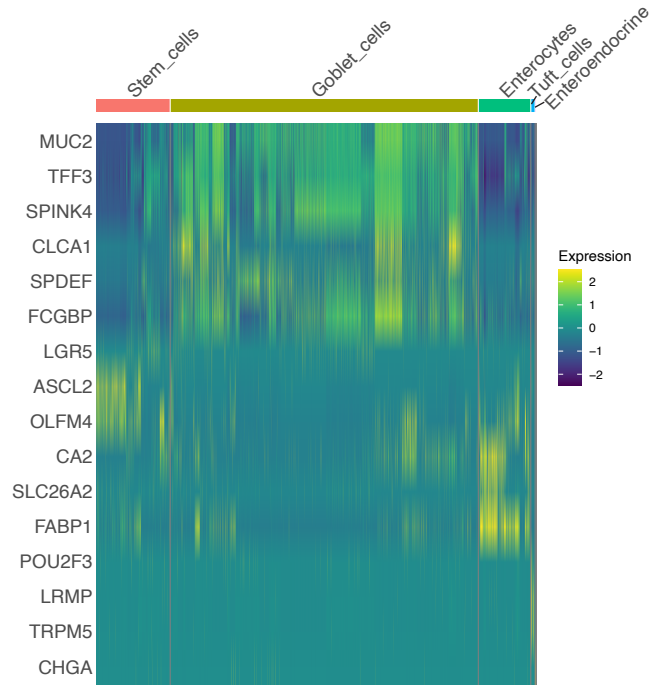
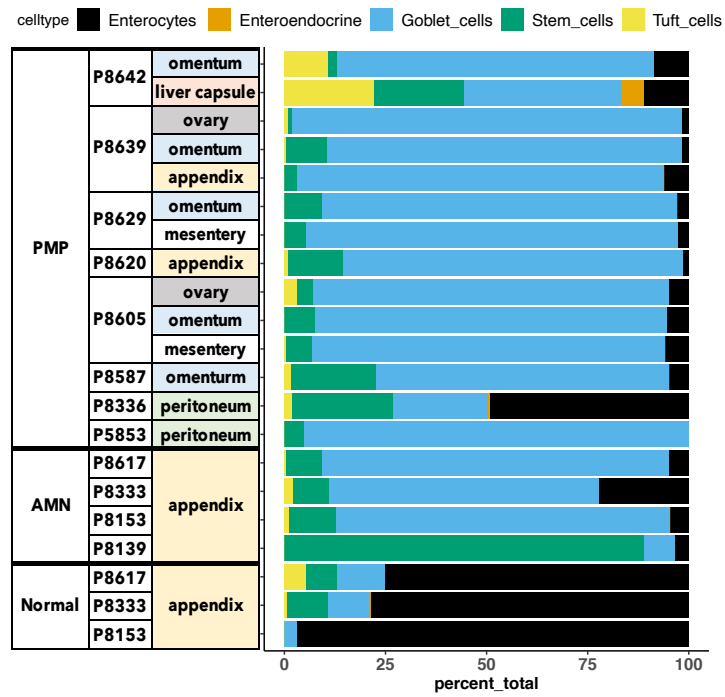


Figure 2

A


Stem_cells
Goblet_cells
Enterocytes
Tuft_cells
Enteroendocrine
MUC2
TFF3
SPINK4
CLCA1
SPDEF
FCGBP
LGR5
ASCL2
OLFM4
CA2
SLC26A2
FABP1
POU2F3
LRMP
TRPM5
CHGA
Expression
2
1
0
−1
−2

B


celltype
Enterocytes
Enteroendocrine
Goblet_cells
Stem_cells
Tuft_cells
PMP
P8642
omentum
liver capsule
P8639
ovary
omentum
appendix
P8629
omentum
mesentery
P8620
appendix
P8605
ovary
omentum
mesentery
P8587
omenturm
P8336
peritoneum
P5853
peritoneum
AMN
P8617
P8333
P8153
P8139
appendix
Normal
P8617
P8333
P8153
appendix
0
25
50
75
100
percent_total

C

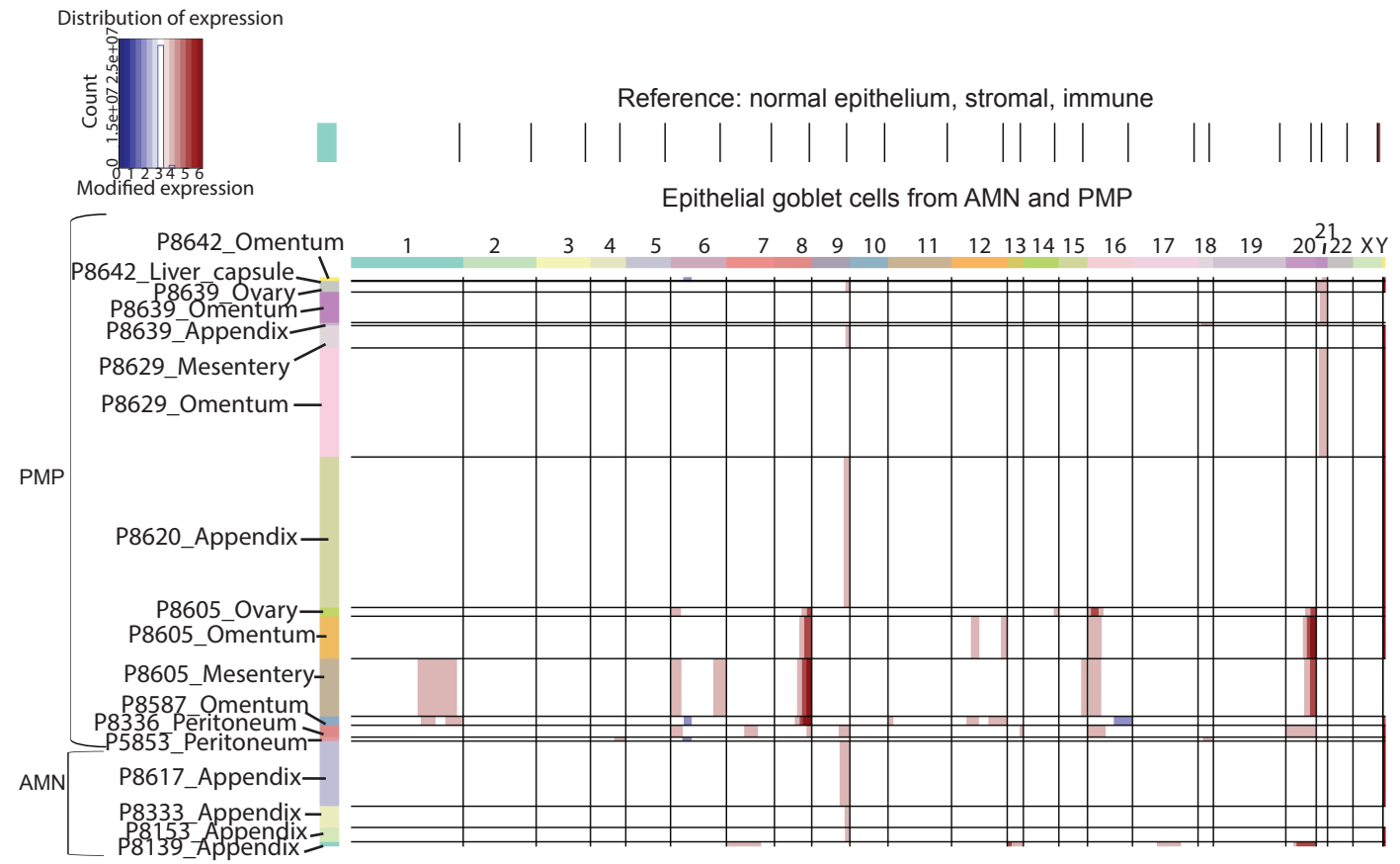

Distribution of expression
Count
0
1.5e+07
2.5e+07
0 1 2 3 4 5 6
Modified expression
Reference: normal epithelium, stromal, immune
Epithelial goblet cells from AMN and PMP
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 XY
P8642_Omentum
P8642_Liver_capsule
P8639_Ovary
P8639_Omentum
P8639_Appendix
P8629_Mesentery
P8629_Omentum
PMP
P8620_Appendix
P8605_Ovary
P8605_Omentum
P8605_Mesentery
P8587_Omentum
P8336_Peritoneum
P5853_Peritoneum
AMN
P8617_Appendix
P8333_Appendix
P8153_Appendix
P8139_Appendix

Figure 3

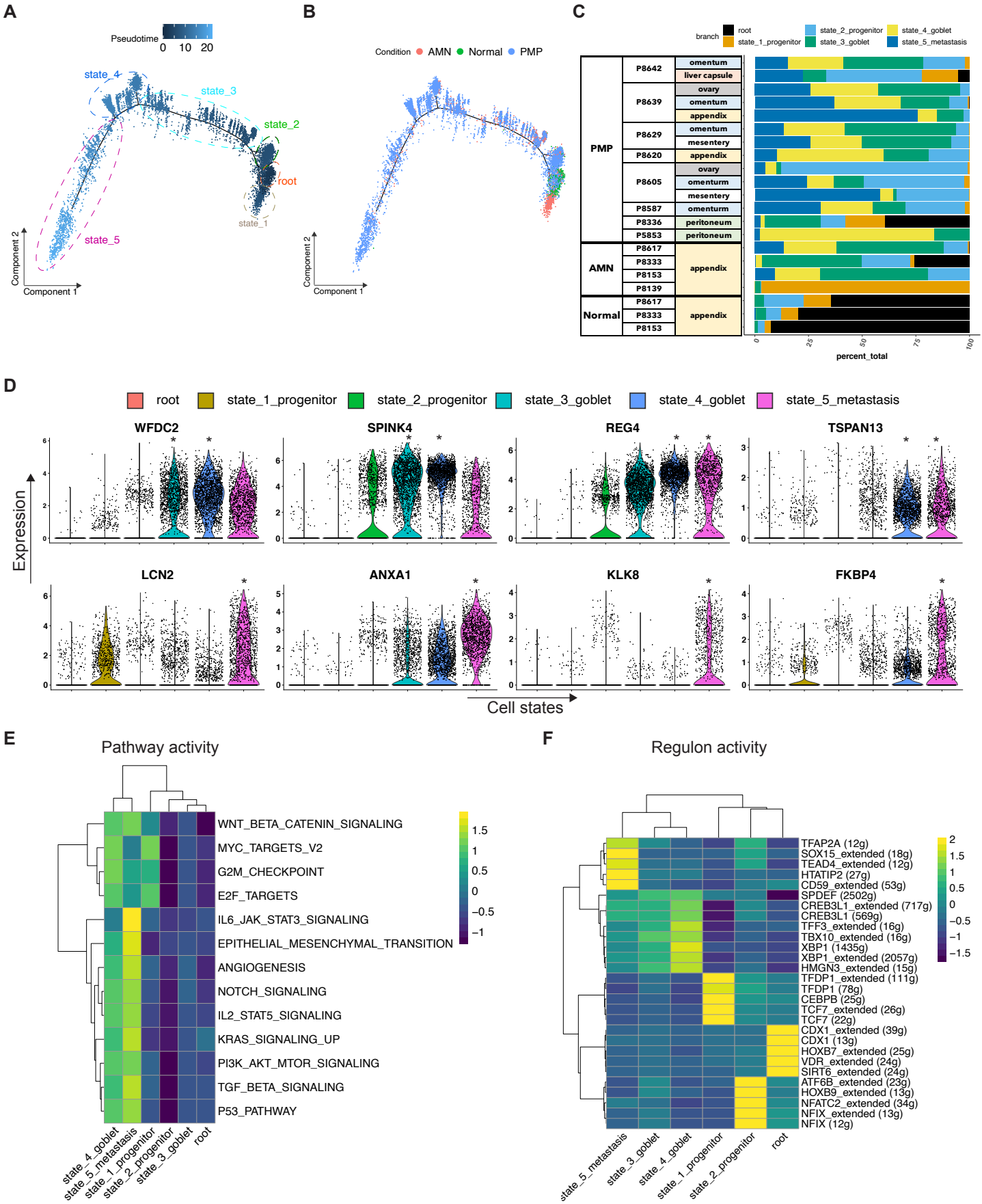
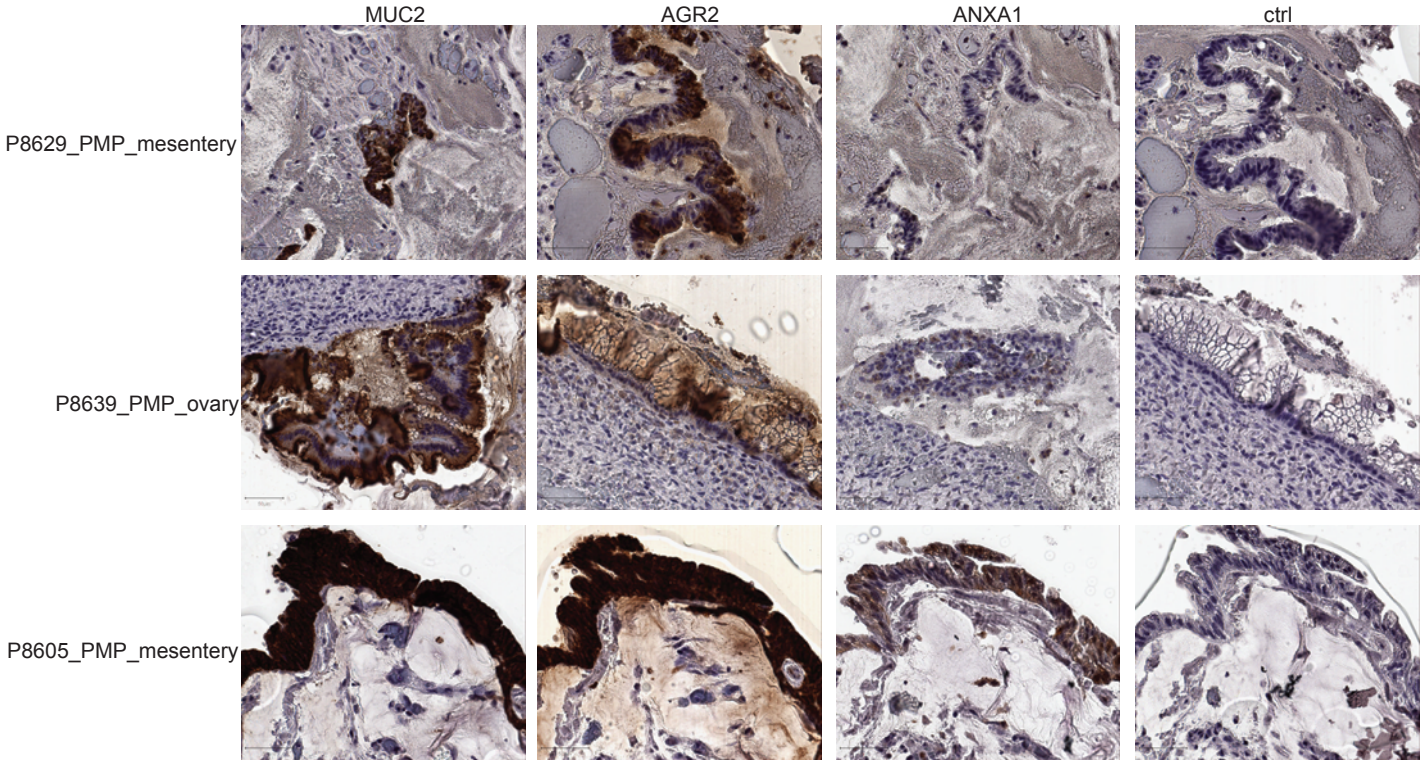

A
Pseudotime
0 10 20
state_4
state_3
state_2
root
state_1
state_5
Component 2
Component 1
B
Condition AMN Normal PMP
Component 2
Component 1
C
branch
root
state_1_progenitor
state_2_progenitor
state_3_goblet
state_4_goblet
state_5_metastasis
PMP
AMN
Normal
P8642
omentum
liver capsule
P8639
ovary
omentum
appendix
P8629
omentum
mesentery
P8620
appendix
P8605
ovary
omenturm
mesentery
P8587
omenturm
P8336
peritoneum
P5853
peritoneum
P8617
P8333
P8153
P8139
appendix
P8617
P8333
P8153
appendix
0
25
50
75
100
percent_total
D
root
state_1_progenitor
state_2_progenitor
state_3_goblet
state_4_goblet
state_5_metastasis
WFDC2
SPINK4
REG4
TSPAN13
LCN2
ANXA1
KLK8
FKBP4
Expression
Cell states
E
Pathway activity
WNT_BETA_CATENIN_SIGNALING
MYC_TARGETS_V2
G2M_CHECKPOINT
E2F_TARGETS
IL6_JAK_STAT3_SIGNALING
EPITHELIAL_MESENCHYMAL_TRANSITION
ANGIOGENESIS
NOTCH_SIGNALING
IL2_STAT5_SIGNALING
KRAS_SIGNALING_UP
PI3K_AKT_MTOR_SIGNALING
TGF_BETA_SIGNALING
P53_PATHWAY
1.5
1
0.5
0
−0.5
−1
state_4_goblet
state_5_metastasis
state_1_progenitor
state_2_progenitor
state_3_goblet
root
F
Regulon activity
TFAP2A (12g)
SOX15_extended (18g)
TEAD4_extended (12g)
HTATIP2 (27g)
CD59_extended (53g)
SPDEF (2502g)
CREB3L1_extended (717g)
CREB3L1 (569g)
TFF3_extended (16g)
TBX10_extended (16g)
XBP1 (1435g)
XBP1_extended (2057g)
HMGN3_extended (15g)
TFDP1_extended (111g)
TFDP1 (78g)
CEBPB (25g)
TCF7_extended (26g)
TCF7 (22g)
CDX1_extended (39g)
CDX1 (13g)
HOXB7_extended (25g)
VDR_extended (24g)
SIRT6_extended (24g)
ATF6B_extended (23g)
HOXB9_extended (13g)
NFATC2_extended (34g)
NFIX_extended (13g)
NFIX (12g)
2
1.5
1
0.5
0
−0.5
−1
−1.5
state_5_metastasis
state_3_goblet
state_4_goblet
state_1_progenitor
state_2_progenitor
root

Figure 4

A


MUC2
AGR2
ANXA1
ctrl
P8629_PMP_mesentery
P8639_PMP_ovary
P8605_PMP_mesentery

B

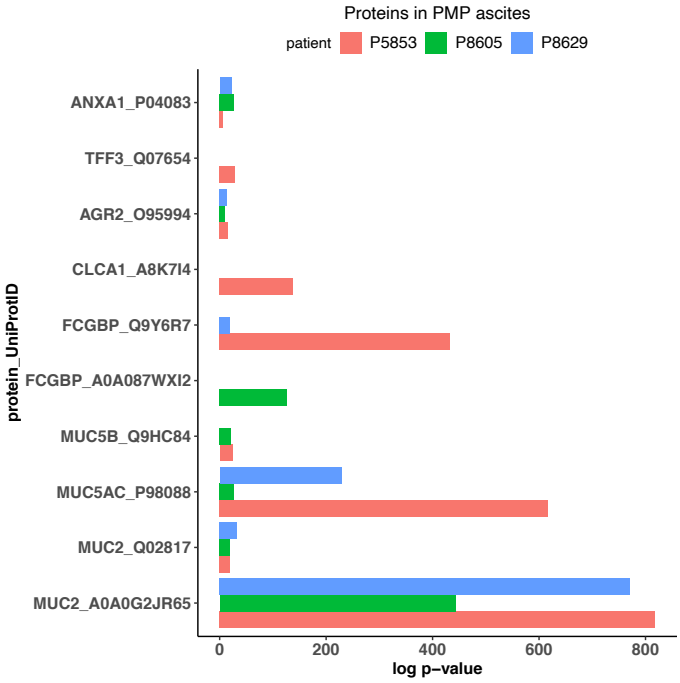
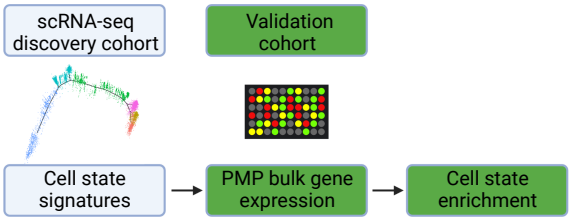
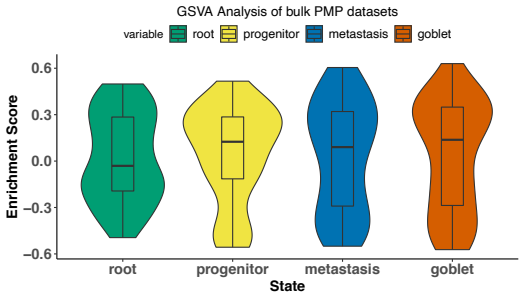

Proteins in PMP ascites
patient
P5853
P8605
P8629
ANXA1_P04083
TFF3_Q07654
AGR2_O95994
CLCA1_A8K7I4
FCGBP_Q9Y6R7
FCGBP_A0A087WXI2
MUC5B_Q9HC84
MUC5AC_P98088
MUC2_Q02817
MUC2_A0A0G2JR65
protein_UniProtID
0
200
400
600
800
log p−value

Figure 5

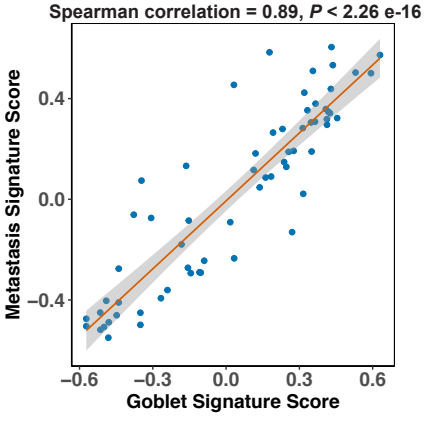
A



B



C



D

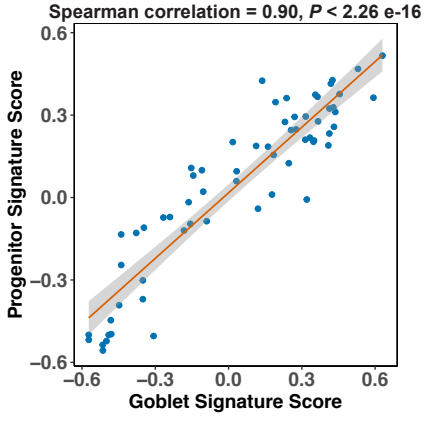
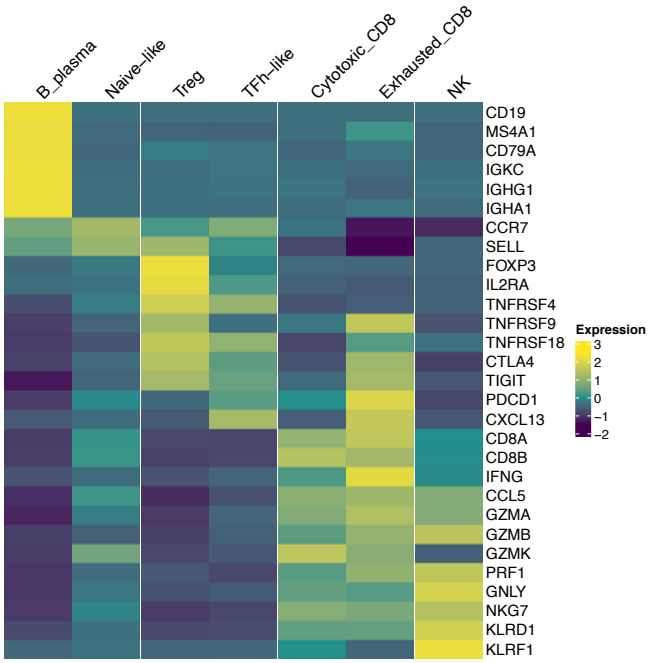
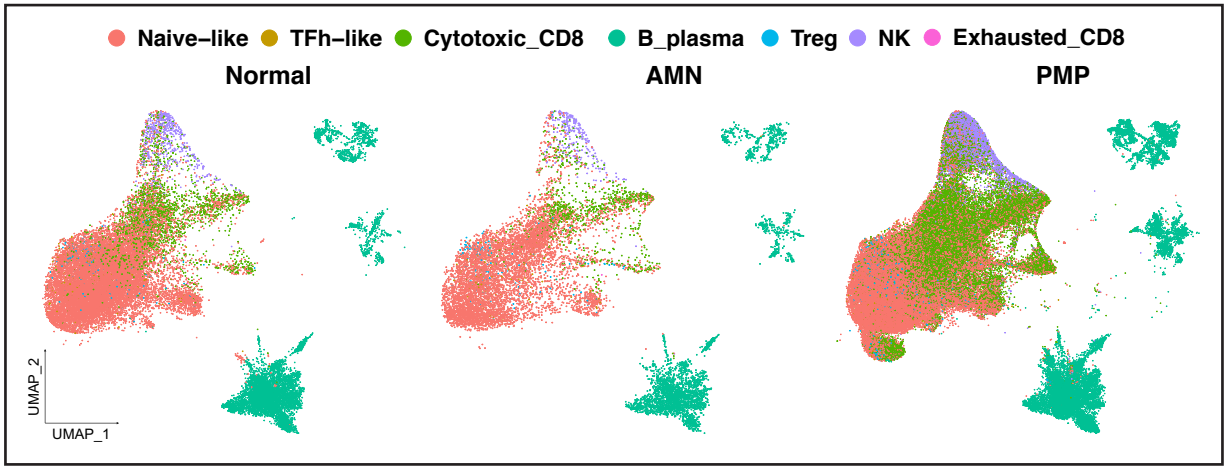


Figure 6

A


B_plasma
Naive−like
Treg
TFh−like
Cytotoxic_CD8
Exhausted_CD8
NK
CD19
MS4A1
CD79A
IGKC
IGHG1
IGHA1
CCR7
SELL
FOXP3
IL2RA
TNFRSF4
TNFRSF9
TNFRSF18
CTLA4
TIGIT
PDCD1
CXCL13
CD8A
CD8B
IFNG
CCL5
GZMA
GZMB
GZMK
PRF1
GNLY
NKG7
KLRD1
KLRF1
Expression
3
2
1
0
−1
−2

B


Naive−like
TFh−like
Cytotoxic_CD8
B_plasma
Treg
NK
Exhausted_CD8
Normal
AMN
PMP
UMAP_2
UMAP_1