bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Ongoing Replication Stress Response and New Clonal T Cell Development Discriminate Between Liver and Lung Recurrence Sites and Patient Outcomes in Pancreatic Ductal Adenocarcinoma

Short Title: PDAC Replication Stress and TCR Diversity

Jason M. Link[1,2*], Carl Pelz[2], Patrick J. Worth[2,3,4], Sydney Owen[2], Dove Keith[2], Ellen M. Langer[1,2,3], Alison Grossblatt-Wait[2,5], Allison L. Creason[3,6], Julian Egger[6], Hannah Holly[1], Isabel English[1], Kevin MacPherson[1], Motoyuki Tsuda[1], Jeremy Goecks[3,6], Emek Demir[2,3,6], Adel Kardosh[2,3,7], Charles D. Lopez[2,3,7], Brett C. Sheppard[2,3,4], Alex Guimaraes[2,3,8], Brian Brinkerhoff[2,9], Terry K. Morgan[2,3,5,9], Gordon Mills[2,3,10], Jonathan Brody[2,3,4], Rosalie C. Sears[1,2,3*]

AFFILIATIONS

1. Department of Molecular and Medical Genetics, Oregon Health and Science University, Portland, OR
2. Brenden-Colson Center for Pancreatic Care, Oregon Health and Science University, Portland, OR
3. Knight Cancer Institute, Portland, OR
4. Department of Surgery, Oregon Health and Science University, Portland, OR
5. Center for Early Detection Advanced Research, Oregon Health and Science University, Portland, OR
6. Department of Biomedical Engineering, Oregon Health and Science University, Portland, OR
7. Department of Hematology and Oncology, Oregon Health and Science University, Portland, OR
8. Department of Radiology, Oregon Health and Science University, Portland, OR
9. Department of Pathology, Oregon Health and Science University, Portland, OR

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

10. Department of Oncological Sciences, Oregon Health and Science University, Portland, OR

*CORRESPONDING AUTHORS:

Jason M. Link, PhD, Department of Surgery, Division of Surgical Oncology, University of California, Los Angeles, 10833 Le Conte Ave, Suite 54-140 CHS, Los Angeles, CA 90095. e-mail: jmlink@mednet.ucla.edu

Rosalie C. Sears, PhD, Department of Molecular and Medical Genetics, Oregon Health and Science University, 2730 SW Moody Ave, CL4RS, Portland, OR 97201, email: searsr@ohsu.edu

Grant Support: RCS is supported by: U54CA209988, U01CA224012 and R01s CA186241, CA196228; GBM is supported by: 1U01 CA253472 and 5U01 CA217842; J.R.B. is supported by: NIH-NCI R01 CA212600; U01CA224012-03; AACR Grant-15-90-25-BROD, and the Hirshberg Foundation; and foundation support from the Brenden-Colson Center for Pancreatic Care

Disclosures: RCS SAB/Consultant: Novartis, RAPPTA, Larkspur

GBM SAB/Consultant: Amphista, Astex, AstraZeneca, BlueDot, Chrysallis Biotechnology, Ellipses Pharma, ImmunoMET, Infinity, Ionis, Lilly, Medacorp, Nanostring, Nuvectis, PDX Pharmaceuticals, Roche, Signalchem Lifesciences, Tarveda, Turbine, Zentalis Pharmaceuticals;

Stock/Options/Financial: Catena Pharmaceuticals, ImmunoMet, SignalChem, Tarveda, Turbine;

Licensed Technology HRD assay to Myriad Genetics, DSP patents with Nanostring

Data Transparency Statement: Datasets will be made available to other researchers as follows:

RNASeq, GEO; DNASeq, NCBI-SRA; TCRSeq, Adaptive Biotechnologies Immunoseq Tool

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Abstract

Background and Aims: Metastatic pancreatic adenocarcinoma (mPDAC) is lethal, yet a subset of patients who have metastatic disease that spreads only to the lung have better outcomes. We identified unique transcriptomic and immune features that distinguish patients who develop metastases in the liver (liver cohort) versus those with lung-avid but liver-averse mPDAC (lung cohort). **Methods:** We used clinical data from the Oregon Pancreas Tissue Registry to identify PDAC patients with liver and/or lung metastases. Gene expression and genomic alteration data from 290 PDAC tumors were used to identify features unique to patients from the liver and lung cohorts. In parallel, T cell receptor sequencing data from 289 patients were used to identify immune features unique to patients in the lung cohort. **Results:** Lung cohort patients had better survival outcomes than liver cohort patients. Primary tumors from patients in the liver cohort expressed a novel gene signature associated with ongoing replication stress (RS) response predictive of poor patient outcome independent from known subtypes. In contrast, patients with tumors lacking the RS response signature survived longer, especially if their tumors had alterations in DNA damage repair genes. A subset of patients in the lung cohort demonstrated new T cell clonal development in their primary and metastatic tumors leading to diverse peripheral blood TCR repertoires. **Conclusion:** Liver-avid metastatic PDAC is associated with an ongoing RS response, whereas tumors lacking the RS response with ongoing T cell clonal responses may have unique vulnerabilities allowing long-term survival in patients with lung-avid, liver-averse metastatic PDAC.

Keywords: pancreatic cancer, metastatic sites, replication stress, T cell receptor diversity

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Introduction

Patients diagnosed with pancreatic ductal adenocarcinoma (PDAC) face a dismal prognosis with a median overall survival of 1 year and an 11% chance of survival at 5 years; patients who present with metastatic disease (~50%) fare even worse. Of the ~20% of patients eligible for potentially curative resection of the primary tumor, only 30-35% survive >5 years. However, a subset of PDAC patients (~10%) who develop lung metastases survive significantly longer than patients with metastatic spread to other sites[14–17]; in some cases, these patients survive >5 years with untreated, indolent lung metastases[18]. Patients with lung-restricted metastatic PDAC typically have slowly progressing disease, and may benefit from a metastatectomy[19]. In contrast, the common presentation of liver metastases or recurrent disease in the liver is frequently associated with poorer patient outcomes. Although not completely understood, at least two reasons may explain why liver metastases are common and so deadly. First, all venous outflow from the pancreas initially passes through the liver via the portal vein, which likely promotes tumor cell seeding in the liver. Second, the liver itself has been shown to be a nurturing tumor microenvironment for PDAC cells[20], in part, suppressing tumor immunity[21] and allowing PDAC liver metastases to flourish[22].

Another factor contributing to mortality in PDAC patients with liver metastases is the greater likelihood that their tumors will be basal-like[23(p6),24]; a previously described subtype of PDAC associated with poor outcomes. Over the last decade, many studies have categorized PDAC tumors and cell lines into 2 to 6 mutually exclusive subtypes based on gene expression of tumors[5,25–27], and the surrounding stroma containing fibroblasts and infiltrating leukocytes[28,29]. Two consensus subtypes reproducibly emerge from these studies[30]: the basal-like/quasi-mesenchymal/squamoid subtype and the classical/ductal/glandular subtype. These are confirmed by metabolomic[31], proteomic[32] and histologic[33] subtyping, as well as expression of reproducibly identified, hallmark subtype-specific proteins[34]. Single cell analyses have confirmed the existence of the two consensus subtypes (basal-

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

like and classical), and have additionally identified single PDAC tumor cells with a blend of the two subtypes[35,36], suggesting a temporal progression from one subtype to the other, general plasticity between subtypes, or multiple overlapping subtype states. Although patients with classical-subtype tumors survive longer on average than patients with basal-like-subtype tumors, patients with either subtype still experience rapid progression and very poor outcomes. Therefore, other factors, superimposed on subtype, likely determine and drive the fate of patients with PDAC.

Many tumors of the basal-like subtype exhibit gene expression related to replication stress (RS)[37], defined by stalled replication forks during DNA synthesis caused by premature entry into S phase, aberrant DNA replication, transcription/replication collisions, abnormal DNA damage checkpoints or altered DNA damage repair[38]. Failure to resolve RS leads to replication fork collapse, DNA damage, cell cycle arrest and ultimately senescence or cell death. Although aberrantly proliferating cancer cells are unavoidably plagued by RS, some malignant cells evolve response mechanisms to tolerate it, and their ability to survive the pro-mutagenic consequences of replication stress is key to their heterogeneity and aggressiveness. Moreover, RS also leads to cytosolic ssDNA, and consequently a cell intrinsic anti-viral response mediated by the stimulator of interferon genes (STING) that drives inhibition of proliferation and an interferon-mediated immune response meant to incapacitate the cells with cytosolic ssDNA. Thus, the strategies that malignant cells develop to respond to and survive RS (and not just RS *per se*) are critical to selection of drugs that target RS as well as the immune response provoked by RS.

Effective PDAC tumor immunity is rare, and both vaccines and immune checkpoint inhibitors (ICI) designed to generate or reinvigorate anti-tumor T cells have generally failed PDAC patients[39,40] despite success in other solid malignancies. However, exceptional cases exist, demonstrating that adaptive tumor immunity does occur naturally and can be effective in some patients[41]. It remains unclear whether immunotherapy requires the right drug given in tandem to effectively treat PDAC

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

patients and/or whether only a subset of patients will be candidates for therapies that modulate tumor immunity. Although microsatellite instability, mismatch repair deficiency, high tumor mutation burden, and PD-1/PDL-1 are potential biomarkers for immunotherapy efficacy[42–45], there has been little progress identifying biomarkers of the immune response itself that inform treatment decisions for PDAC. Although some precision approaches to monotherapies and immunotherapies can moderately enhance outcomes for a fraction of PDAC patients[44], future success will likely require a more complete understanding of metastatic location, cell intrinsic and microenvironment vulnerabilities, and the tumor immune landscape.

In this study, we generated and interrogated large datasets of tumor DNA and RNA sequencing as well as T cell receptor (TCR) sequencing from matched blood and tumor samples to evaluate the differences between liver-avid versus liver-averse, lung-avid PDAC. We develop a gene-expression signature specific to liver-avid versus liver averse, lung avid disease independent of the basal-like versus classical subtypes, that predicts poor patient outcome; and we identify evidence of ongoing replication stress response associated with this aggressive form of PDAC. Conversely, we show that primary tumors lacking the RS response signature are sensitive to alterations in DNA damage repair genes leading to long-term patient survival. Furthermore, we identify patients with lung-avid, liver-averse disease as a subset of patients with tumors lacking the RS response who exhibit ongoing T cell clonal development in their tumors and resultant high peripheral TCR diversity. Our findings herein identify a gene expression signature associated with RS response and identify tumors lacking this response as an exploitable weakness of PDAC that along with a specific type of functional adaptive tumor immunity can lead to improved outcomes for patients diagnosed with this devastating disease.

Methods

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Tissue Acquisition and Patient Consent

Patient data, blood, and tissues were obtained with informed consent in accordance with the Declaration of Helsinki and were acquired through the Oregon Pancreas Tissue Registry under Oregon Health & Science University IRB protocol #3609.

Clinical Data Collection

Clinical course timepoints, patient demographics, stage, grade, nodal involvement, resection margins, and angiolymphatic invasion were provided as deidentified data by the OHSU cancer registry with quality control data verification in a subset by a surgical pathologist (TM). For all patients in the liver and lung cohorts, we abstracted the site of all lesions proven to be metastatic by biopsy, described as likely metastatic by a radiologist, or that significantly increased in size during progression or decreased in size during treatment as long as a radiologist described the lesion as suspicious for or concerning for metastasis. Clinical imaging was reviewed by AG to validate patient assignments to the liver and lung cohorts. Time to recurrence was calculated from the earliest of either the recurrence date provided by the OHSU cancer registry, or the date of earliest lesion abstracted from CT reports.

Specimen Processing

Primary and metastatic PDAC tumor specimens from consented patients at OHSU were processed by the OHSU department of pathology and preserved by standard formalin fixation and paraffin embedding leading to 3μm serial sections stained for hematoxylin & eosin.

Histology Data

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

H&E stained formalin-fixed paraffin-embedded (FFPE) tissue sections from regions corresponding to those extracted for RNA-seq and somatic alteration analyses were independently appraised by two independent pathologists blinded to clinical data for the histologic features shown in Table 1. Overall diagnostic reproducibility kappa statistic was 0.89 [excellent defined as > 0.6]. Consensus diagnoses were used for metric comparisons.

Tempus RNA-seq and Genomic Alteration panel processing

OHSU provided FFPE PDAC specimen blocks along with matched normal blood or tissue to Tempus as part of a contract agreement. OHSU pathologist (TM) and Tempus pathologists marked regions of high tumor content (>20% ratio of tumor to normal nuclei) on H&E stained slides for DNA and RNA extraction. Solid tumor total nucleic acid was extracted from these tumor regions on adjacent FFPE tissue sections using Chemagic 360 sample-specific extraction kits (Perkin Elmer) and digested by proteinase K. RNA was purified from the total nucleic acid by DNase-I digestion. DNA sequencing of 596 genes and whole-transcriptome RNA sequencing were performed as previously described[1]. Briefly, 100 nanograms (ng) of DNA for each tumor sample was mechanically sheared to an average size of 200 base pairs (bp) using a Covaris Ultrasonicator. DNA libraries were prepared using the KAPA Hyper Prep Kit, hybridized to the xT probe set, and amplified with the KAPA HiFi HotStart ReadyMix. One hundred ng of RNA for each tumor sample was heat fragmented in the presence of magnesium to an average size of 200 bp. Library preps were hybridized to the xGEN Exome Research Panel v1.0 (Integrated DNA Technologies) and target recovery was performed using Streptavidin-coated beads, followed by amplification with the KAPA HiFi Library Amplification Kit. The amplified target-captured DNA tumor library was sequenced using 2x126bp PE reads to an average unique on-target depth of 500x (tumor) and 150x (normal) on an Illumina HiSeq 4000. The amplified target-captured RNA tumor library was sequenced using 2x75bp PE reads to an average of 50M

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

reads on an Illumina HiSeq 4000. Samples were further assessed for uniformity with each sample required to have 95% of all targeted bp sequenced to a minimum depth of 300x. Raw fastq files were returned to OHSU.

DNA sequence analysis

DNA Variant detection, visualization, and reporting were performed as previously described[1]. Alignment and mapping was to GRCh37 using Novo align + BWA. CNVs derived from proprietary tumor/normal match analysis using CNAtools. Genomic variants are displayed on waterfall plots using GenVisR R package software[2]. Statistics are Fisher's Exact test to look for significantly changed alteration rates between cohorts.

RNA sequencing analysis

Paired-end fastq sequences were trimmed using trim-galore (ver 0.6.3) and default parameters. Pseudoalignment was performed with kallisto (ver 0.44.0) using genome assembly GRCh38.p5 and gencode (ver 24) annotation; default parameters were used other than the number of threads. The Bioconda package bioconductor-tximport (ver 1.12.1) was used to create gene level counts and abundances (TPMs). Quality checks were assessed with FastQC (ver 0.11.8) and MultiQC (ver 1.7). Quality checks, read trimming, pseudoalignment, and quantitation were performed via a reproducible snakemake pipeline, and all dependencies for these steps were deployed within the anaconda package management system[3,4]. We downloaded the "Pancreatic Adenocarcinoma (TCGA, PanCancer Atlas)" dataset from cbioportal.org.

PurIST analysis

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

PurIST subtype calls and scores were generated using the PurIST method [5] applied to our RNA-Seq data. The PurIST authors provide instructions, R scripts, and gene pairs on GitHub (https://github.com/naimurashid/PurIST).

Development of pORG and pSUB signatures

A two-factor analysis with DESeq2[6] was performed on RNA-Seq counts from 76 primary samples. The two factors modeled were: liver cohort vs lung cohort and basal-like vs classical (from PurIST subtyping). For pORG, the most significant DE genes from the liver vs lung factor (p-value <0.01) were selected, then genes that co-occurred in the top half of the ranked genes from the basal vs classical factor were excluded, resulting in a list of 78 upregulated genes (only genes up in liver cohort were selected). For pSUB, the most significant DE genes from basal-like vs classical factor (p-value <0.01) were selected, then genes that co-occurred in the top half of the ranked genes from the liver vs lung factor were excluded, resulting in a list of 362 upregulated genes, of these, the 100 most significant were selected..

GSEA and GSVA analyses

The GSVA tool[7] was used to calculate relative pORG and pSUB gene set scores across all primaries, all metastases, and all tumors and identify top/bottom quartile cohorts. GSEA[8] was run on clinical cohorts, pORG, pSUB, and PurIST cohorts using the MSigDB database v7.5.1 Hallmark gene set collection[9]. The 7 genes used for the IRDS signature were: STAT1, IFI44, IFIT3, OAS1, IFIT1, G1P2, and MX1

VIPER analyses and Immune cell type enrichment

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

The transcriptional regulon enrichment analysis was performed using VIPER with the TCGA PAAD ARACNe-inferred network[10,11]. Gene expression data was normalized prior to running viper by median centering and scaling. VIPER regulon scores for all primaries were used for cohort comparisons.

Immune Cell Type enrichment analysis was scored using Gene Set Variation Analysis (GSVA) with gene sets defining 16 distinct immune cell types[7,12,13] calculated for all primaries for cohort comparisons. GSVA was performed using log transformed gene expression data with a Gaussian kernel for estimation of the cumulative density function and a normalized enrichment statistic of the KS-like random walk.

Software

R version 3.6.0 was used for GSVA and VIPER. R version 4.1.2 was used with R packages DESeq2, GSVA, msigdbr, gplots, and ggplot. GSEA was run in JAVA using the command line interface.

TCRB Sequencing

Frozen leukocytes and 25 mm thick curls of formalin fixed, paraffin-embedded tumor were submitted to Adaptive Biotechnologies for human TCRB sequencing. Analyses were performed using the Immunoseq tool provided by Adaptive Biotechnologies. Samples with fewer than 100 productive templates were excluded from analyses. The Diversity Metrics Tool was used for richness and evenness metrics, and the Differential Abundance Tool was used to assess overlap between paired tumor and blood samples. The percentage of tumor-exclusive clones was calculated from a list of all rearrangements with ≥10 templates found in both samples combined. For shared, clonal sequences within cohorts, the top 50 rearrangements (by frequency in each sample) were compiled for all samples, and the Immunoseq Sequence Search Tool was used to identify all samples in the cohort

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

that contained any of those CDR3b sequences at any frequency. Only the CDR3b amino sequences found in at least 25% of samples in the cohort were considered shared, clonal sequences.

Statistics

All statistical tests were performed with GraphPad Prism. A Mantel-Cox test was used to compare all Kaplan-Meier survival and recurrence curves. ANOVA was used for plots with multiple comparisons and t-tests were used for plots with single comparisons. For non-Gaussian data (e.g., number of tumor-discrete TCR sequences) we used Kruskal-Wallis tests. Pearson and Spearman correlation coefficients were generated for continuous variables and non-parametric data, respectively

Results

PDAC patients who develop lung and not liver metastases have significantly better outcomes than patients who develop liver metastases

From a de-identified dataset of 1,873 patients diagnosed with and/or treated for PDAC at our institution between 2004 and 2020, we collected a complete set of disease-relevant computed tomography (CT) scans that were available for 234 (Supplemental Dataset 1). From these scans, we identified 35 patients who developed lung metastases but never developed liver metastases (referred to as the “lung cohort”) and 130 patients who developed liver metastases (referred to as the “liver cohort”). Consistent with previous reports, we observed that lung cohort patients fare significantly better by median overall survival (819 days) than liver cohort patients (468 days– Figure 1A). The median survival time for the liver cohort did not noticeably change if patients with both liver and lung metastases were excluded (450 days – Figure 1A). Median survival was also significantly different between patients in the lung and liver cohorts when comparing only patients treated by surgical

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

resection of their primary tumors (876 days vs 549 days, respectively – Supplemental Figure 1A). For patients treated by primary tumor resection, median time to recurrence was significantly longer for patients in the lung cohort than patients in the liver cohort (303 days vs 166 days, respectively – Figure 1B), as was survival after resection (784 vs 498 days, respectively – Figure 1C), which positively correlated with time to recurrence for both cohorts (Supplemental Figure 1B) underscoring the role that recurrence plays in outcomes. Similarly, patients in the lung cohort survived longer after recurrence than patients in the liver cohort, though this difference was not significant (397 days vs 290 days, respectively, P=0.056 – Supplemental Figure 1C), and there was a significant correlation between median survival after recurrence and survival after resection for both cohorts (Supplemental Figure 1D). In contrast, there was no correlation between time to recurrence and survival time after recurrence (Supplemental Figure 1E) suggesting that there are different mechanisms of disease progression during these two clinical time periods.

The survival benefit from the classical subtype is not sufficient to explain the extended survival of lung cohort patients

We generated gene expression data by performing RNA-seq on histologically confirmed tumor-enriched regions from formalin-fixed paraffin embedded primary (N=218) and metastatic (N=72) PDAC tumors from patients at our institution, and then used PurIST[5] to assign conventional subtypes of PDAC (basal-like or classical) to each tumor (Supplemental Dataset 1). Consistent with a previous report[24], we found that tumors from patients in the lung cohort scored significantly more classical than basal-like (Figure 1D). Also consistent with previous reports[25,26], we found an overall median survival benefit for patients with classical tumors compared to those with basal-like tumors (570 vs 394 days, respectively – Figure 1E). Median time to recurrence was also significantly longer for patients with classical-subtype tumors compared to those with basal-like tumors (244 vs 153 days, respectively –

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 1F). However, when we assessed median survival of only patients with classical-subtype tumors in the lung vs liver cohorts, we found that patients in the lung cohort still lived significantly longer than patients in the liver cohort (1681 vs 520 days, respectively – Figure 1E) and also had a longer time to recurrence (303 vs 168 days, respectively – Figure 1F). These results indicate that there is a subtype-independent survival benefit for patients in the lung-cohort relative to the patients in the liver cohort.

Clinical features and genomic alterations associated with metastatic-site avidity

We investigated clinical correlates of the liver and lung cohorts but did not observe differences in gender, race, stage at diagnosis, or tumor grade (Table 1). Patients in the lung cohort were more likely to be treated by resection than patients in the liver cohort (89% vs 64%, respectively – $P<0.001$), which may be due to more advanced disease at presentation for liver-cohort patients, however, the survival advantage in the lung cohort is still evident when only comparing patients treated by resection (Supplemental Fig 1A). A similar proportion of patients were treated with standard-of-care neoadjuvant chemotherapy in both cohorts and neoadjuvant treatment did not appear to influence time to recurrence for either cohort (Supplemental Figure 2).

Two surgical pathologists (TM and BB) performed a comprehensive histologic analysis of H&E stained serial sections adjacent to those extracted for RNA-seq. No significant histologic differences were noted between resection specimens from the two cohorts by grade or stage, including nodal involvement, presence of angiolymphatic invasion, presence of post-resection residual tumor, type of inflammation (acute versus chronic), or the presence/absence of prominent lymphoid aggregates (Table 1). Specimens from both cohorts showed extensive desmoplasia, as expected. We next determined the percentage of malignant cells in our specimens by using mutant allele frequencies determined from amplicon-based high-throughput sequencing of 595 genes known to be altered in

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

cancer. We identified significantly higher tumor content in basal-like vs classical tumors (P=0.016 – Supplemental Figure 3); however, we did not identify a significant difference in the fraction of malignant cells in primary tumors between the liver and lung cohorts, or liver and lung metastases (Supplemental Figure 3).

We used the same gene panel DNA sequencing data to compare genes with alterations between basal-like and classical cohorts, and between liver and lung cohorts. The top 10 alterations are shown for each cohort (Supplemental Figure 2 and Supplemental Dataset 2). Among primary tumors, we found a significant increase in the frequency of alterations for ABL1, MKI67, and SMARCA4 in basal-like tumors, and for ATM, CUX1, GNAS, and TGFBR2 in classical tumors (Supplemental Table 1A). In liver-cohort primary tumors, alterations in TP53 were significantly more common, whereas PMS2 and PTEN alterations were more common in lung-cohort primary tumors (Supplemental Figure 4 and Supplemental Table 1A). Among metastases, we found 10 genes more likely to be altered in basal-like versus classical tumors (Supplemental Figure 4 and Supplemental Table 1B), but none that were different between the liver-cohort and lung-cohort metastases. Consistent with previous reports[46], we found similar genes altered between paired tumors and metastases from the same patient (Supplemental Table 2).

Gene-expression signatures in primary tumors specific to metastatic organotropism (pORG) or previously defined subtypes (pSUB) independently predict overall survival

Since outcome is influenced by recurrence site within the same subtype (Figure 1E and F), we sought to identify gene-expression pathways in primary tumors that may predict liver-avidity versus lung avidity/liver aversity of primary tumor cells without being influenced by the higher percentage of basal-like tumors in the liver cohort (Figure 1D). We ran a two-factor analysis with DESeq2[6] to identify differentially expressed (DE) genes in primary tumors from the liver cohort vs lung cohort and from

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

basal-like vs classical primary tumors (Supplemental Dataset 3). In order to focus on the biology of metastatic organotropism independent from previously published subtypes[25,28,30], we excluded the top DE genes for established subtypes from the DE metastatic organotropism genes to obtain a **p**rimary **org**anotropism signature discriminating between liver-avid vs lung-avid, liver-averse tumors that we termed pORG (78 upregulated genes – Supplemental Dataset 3). We also performed this process for the DE **p**rimary tumor conventional **sub**type basal-like vs classical genes, subtracting the top DE organotropism genes to obtain a gene signature termed pSUB (100 upregulated genes - Supplemental Dataset 3).

We used the GSVA tool[7] to run a Gene Set Variation Analysis of our primary tumor samples for both the pORG and pSUB gene sets. This analysis yields an activity score for each tumor relative to all other tumors (Supplemental Dataset 1). As expected, primary tumors from the liver cohort had significantly higher pORG scores than primary tumors from the lung cohort, but the pORG signature score did not significantly separate primary basal-like tumors from classical tumors (Figure 2A). Conversely, while pSUB significantly separated basal-like from classical tumors, it did not significantly separate the liver and lung cohort primary tumors (Figure 2A). We also produced pORG and pSUB GSVA scores for metastatic samples. These pORG scores did not distinguish liver-cohort from lung-cohort metastases (Figure 2B, left panel), indicating that this signature - and the functions it represents - are likely specific to primary tumors. Similar to other subtyping methods, the pSUB score did distinguish basal-like from classical metastases (Figure 2B, right panel) and also distinguished metastases from the liver and lung cohorts, underscoring previous reports that subtype biology carries over into metastases and that lung metastases are more likely to be classical (Figure 1D). GSVA scores for all specimens were not significantly different between all primary and all metastatic tumors for either pORG or pSUB (Supplemental Figure 5A).

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

We found that overall survival significantly correlated with primary tumor pORG and pSUB GSVA scores independently (Supplemental Figure 5B). Relatedly, we found significant differences in overall survival between the top and bottom quartiles of all primary tumor specimens classified by pORG or pSUB GSVA scores for both our dataset (Figure 2C, top panels) and the pancreatic adenocarcinoma patient dataset (PAAD) reported by The Cancer Genome Atlas (Figure 2C, bottom panels). Thus, even though the pORG gene set is devoid of genes enriched in basal-like compared to classical tumors, and vice versa, each independently predicts outcome in our dataset and in the independent TCGA PAAD dataset, suggesting that each of these non-overlapping signatures can identify unique, clinically relevant biology.

Distinct cancer hallmark pathways are enriched by pORG and pSUB gene sets

We used the cancer Hallmarks GSEA database to investigate pathways enriched in primary tumors with high vs low pORG scores and between the liver vs lung cohorts. We expected that specimens with high versus low pORG scores may more robustly identify GSEA pathways enriched in liver-avid primary tumors than the comparison between the liver and lung cohorts, as the latter comparison is influenced by subtype, and would be influenced by the inherent limitations of clinical-cohort assignments. Consistent with this, both high versus low pORG and liver-cohort versus lung-cohort comparisons showed similar pathway enrichments, but the liver vs lung comparison generally yielded lower normalized enrichment scores (Figure 3A and Supplemental Dataset 4). Both comparisons showed enrichment for pathways related to oncogene-mediated replication stress; specifically: G2M Checkpoint, E2F Targets, Mitotic Spindle, MYC Targets V1, and Interferon Alpha Response, as well as pathways related to cell metabolism and mitogenesis. The only hallmark pathway significantly enriched in low pORG or lung cohort was myogenesis. Specimens from both comparisons clustered by their most significantly enriched pathways (NES >1.7, nom p-value = <0.05) as expected (Figure

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

3B). We performed similar analyses of hallmark gene sets enriched in primary tumors comparing specimens defined as basal-like versus classical by PurIST, specimens with high versus low PurIST scores, and specimens with high versus low pSUB scores. All three of these comparisons revealed enrichment for pathways: Epithelial to Mesenchymal Transition (EMT), Apical Junction, hypoxia and angiogenesis, plausibly capturing the relationship between hypoxic stress and EMT[47] (Figure 3C left panel and Supplemental Dataset 4). Additionally, the high pSUB score tumors showed enrichment of glycolysis and mTORC signaling. Inversely, we found pathways for Bile Acid Metabolism and Pancreas Beta Cells enriched in both classical subtype and low pSUB score primary tumors (Figure 3C and Supplemental Dataset 4). The significantly, positively enriched hallmark pathways (NES >1.7, nom p-value = <0.05) in these comparisons segregated tumor groupings as expected (Figure 3D). Together, these data indicate that separating metastatic organotropism and conventional molecular subtype gene expression by using our pORG and pSUB signatures can identify unique biologic pathway enrichments, with a high pORG score correlating with replication stress response pathways and a high pSUB score correlating with EMT and hypoxia pathways. Similar hallmark pathways were enriched in high pORG and high pSUB samples from the independent TCGA PAAD dataset (Supplemental Figure 6).

Low pORG tumors are less tolerant to defects in DNA damage repair

A recent report by Dreyer et al. suggested that treatment-agent efficacy may depend on both replication stress (RS) and DNA Damage Response (DDR) gene alteration status – dividing patients into 4 categories based on the presence or absence of those two factors[37]. We considered that overall survival might be affected by these factors apart from treatment. Since our data indicate that liver-avid, high pORG score primary tumors are enriched for pathways indicative of ongoing RS, we divided patients into 4 similar categories by evenly splitting primary tumors by high and low pORG

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

score and also by the presence or absence of a known DDR gene alteration[48]. Although patients with high pORG scoring tumors fared poorly regardless of DDR gene status, patients with low pORG scores survived significantly longer (P=0.012) if their primary tumors had altered DDR genes (Figure 3E). This result suggests that the effects of DDR gene alterations on tumor cell survival may be ameliorated in tumor cells with high pORG scores that have upregulated pathways responding to ongoing replication stress; consistent with activation of cell cycle checkpoint pathways to survive DNA damage and replication stress as well as MYC activity upregulating DNA damage response pathways in tumors with high pORG scores[49,50]. Additionally, liver-cohort primary tumors with DDR gene alterations had some of the highest pORG scores compared to those without DDR gene alterations (Figure 3F); suggesting that the presence of DDR gene alterations themselves may promote molecular mechanisms supporting tumor-cell responses to replication stress and DNA damage in order to avoid mitotic catastrophe.

High pORG and liver-cohort tumors are enriched in transcription networks related to DNA damage repair, cell cycle promotion, and immune cell depletion

We performed transcriptional regulon enrichment analysis using VIPER and the TCGA pancreatic PAAD ARACNe-inferred network[10,11] to identify regulatory protein activity scores between cohort comparisons (Supplemental Dataset 5). Consistent with GSEA results from the Hallmarks of Cancer gene sets (Figure 3), we observed significantly increased activity of proteins involved in DNA replication (ORC2 and MCM2-7 proteins – Figure 4A), and the response to replication stress (CHEK1, ATR, RAD51, BRCA1/2, BARD1, and PALB2 – Figure 4B) in high pORG vs low pORG and liver cohort versus lung cohort samples, while parallel proteins involved in double-strand break response like ATM and ARID1A were significantly decreased in activity (Figure 4B). Concurrently, positive cell cycle regulators (CyclinD, CDK2,4,6 and E2F1,2 – Figure 4C) were significantly

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

increased in high versus low pORG and liver versus lung samples while the CDK inhibitor CDKN1B (P27) was significantly decreased (Figure 4C). We also found increased VIPER regulon activity scores in high vs low pORG for IFN alpha/beta receptor subunits 1 and 2 (mean DE 1.75 and 0.89, respectively – P<0.0001 for both), consistent with enrichment of IFNa response in high pORG samples by GSEA (Figure 3A and B). Together, these results support ongoing replication stress response mechanisms in high pORG/liver cohort PDAC.

The enrichment of gene expression related to Type 1 IFN (IFN) signaling in high pORG tumors does not match the well-known transient effects of IFN (inhibition of infected/tumor cell proliferation[51] and upregulation of leukocyte effector functions[52]). However, the downstream activity of IFN is pleiotropic and changes dramatically over time; in fact, chronic IFN signaling in cancer induces an IFN-related DNA Damage Resistance Signature (IRDS) of gene expression that is associated with tumor cell resistance to DNA damage[53–55] and escape from tumor immunity[56]. We used a set of IRDS genes reported to distinguish outcomes in cancer[53,55] and found a significant enrichment in high pORG versus low pORG tumors (Figure 4D, left panel). A similar trend, though not significant, was seen in the liver-cohort versus lung-cohort primary tumors (Figure 4D, right panel). Two genes in the IRDS gene set matched VIPER regulons (STAT1 and MX1) and as expected, these were both significantly enriched in high pORG and liver-cohort tumors over low pORG and lung-cohort tumors, respectively (Figure 4E).

Chronic IFN signaling can also limit tumor immunity by favoring infiltration of suppressive leukocytes and inactivating adaptive immune cells[57]. To investigate this in our tumor cohorts we calculated the mean expression of sets of genes related to common lymphocyte subsets[12,58,59]. We found a significant enrichment of B cells and subsets of T cells (notably, CD4+ follicular helper and effector memory cells) in low pORG versus high pORG tumors (Figure 4F), and a similar pattern for lung-cohort versus liver-cohort tumors, though the latter comparisons were not significant. Taken

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

together, these results suggest that tumors without ongoing RS and chronic IFN signaling are enriched for lymphocyte subtypes.

Peripheral blood T cell receptor repertoires are more diverse in patients with lung-tropic, liver-averse disease

We performed sequencing of gene segments comprising the complementarity determining region 3 (CDR3) of T cell receptor (TCRB) genes from 289 blood samples, 175 primary tumors (141 overlapping with the RNA-seq dataset), and 43 metastatic tumors (33 overlapping with the RNA-seq dataset). Within this dataset, we analyzed blood samples from 77 patients in the liver cohort and 16 patients in the lung cohort, of which 60 and 16 were matched with tumor samples from the same patient, respectively. Consistent with the gene-expression data showing enrichment for some T cell subsets in low pORG tumors (Figure 4F), we found a higher density of TCR templates in samples from low pORG tumors (Figure 5A, p= .0.012) but no difference between basal-like and classical subtypes. We did not find a higher density of TCR templates in lung-cohort tumors relative to liver-cohort tumors (Figure 5A), although this is consistent with the lack of a clear enrichment of T cells in lung-cohort over liver-cohort tumors (Figure 4F). As expected, the density of TCR templates per ng of tumor tissue significantly influenced the number of productive rearrangements we analyzed (i.e., the observed richness of the population, Supplemental Figure 7A), and we found significantly more productive rearrangements in low pORG tumors compared to high pORG tumors (Supplemental Figure 7B). These data suggest that mechanisms represented by a high pORG signature may restrict the density of T cells in the tumor and reduce the richness of analyzed TCR repertoires.

T cell repertoire diversity is a function of both richness (the abundance of TCRs) and evenness (the distribution of the frequencies of clones within a sample – a very clonal repertoire would have low evenness). We developed an Evenness Index that is the combination of three metrics unrelated to

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

richness and unlikely to be affected by differences in the number of TCR productive rearrangements analyzed (Pielou Evenness, Simpsons Evenness, and Clone Distribution Slope)[60]. We did not find a difference in TCR repertoire evenness between any of the tumor cohorts indicating that the overall shapes of the liver-cohort and lung-cohort tumor repertoires are not different (Supplemental Figure 7C and Figure 5B). In contrast, we found significantly more TCR repertoire evenness (i.e., less clonality) within peripheral blood samples from patients in the lung cohort (Figure 5C). We considered that although evenness/clonality was similar between liver-cohort and lung-cohort tumors, that the nature of those clonal responses may differ, and that difference may manifest in the peripheral blood TCR. To investigate this, we used data from 214 matched pairs of tumor and blood samples (91% collected on the same day) and looked for unique TCRB clones in tumor samples that were absent in matched blood samples (i.e., tumor-distinct clones), suggestive of new clonal development. We found that lung-cohort tumors harbored significantly more of these tumor-distinct clones (Figure 5D). This result is consistent with ongoing, new T cell clonal development in lung-cohort tumors that is similar in evenness/clonality to other tumors at any one timepoint, but as new clones emigrate to the periphery, TCR diversity increases in blood. Consistent with this conclusion, we found a positive correlation between the Evenness Index in blood samples and the percentage of tumor-distinct clones for all 214 patients with matched blood and tumor samples (Supplemental Figure 7D). Additionally, we found that the Evenness Index in tumor samples negatively correlated with the percentage of tumor-distinct clones (Supplemental Figure 7E), underscoring that new clonal development contributes to overall clonality of the tumor T cell repertoire.

Our data suggest that T cell responses in liver-cohort tumors are not directed at new antigens over time, but instead may reflect chronic reactions to common PDAC-associated (or other) antigens present since tumor initiation. To investigate these responses, we selected published CDR3b sequences that have been experimentally confirmed as part of TCRs specific for KRAS G12/13

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

alterations[61], a tumor-initiating mutation present in ~90% of PDAC tumors. We identified 4 of these CDR3b sequences in liver cohort blood and tumor samples, and 3 of those 4 CDR3b sequences in lung-cohort blood and tumor samples (Supplemental Dataset 6). Consistent with repeated expansion of the same clones over time, we found these putative mutant KRAS-specific CDR3b sequences at significantly higher frequencies in the analyzed repertoires of liver-cohort tumor samples compared to the lung-cohort tumor samples (Figure 6A), suggesting that T cell responses to tumor-initiating neoepitopes are more likely to expand in liver-cohort tumors. These sequences existed at similarly low frequencies in liver-cohort and lung-cohort blood samples (Figure 6A), consistent with limited expansion of these clones in the periphery. The presence of these putative mutant KRAS-specific CDR3b sequences in tumors did not correlate with better outcomes; rather the median overall survival for patients with KRAS-specific CDR3β sequences in tumors was shorter for both liver-cohort (391 days) and lung-cohort (565 days) patients compared to the median survival for all patients in these cohorts (468 and 819 days, respectively – Figure 1A).

Lung-cohort tumors are a subset of low-pORG tumors that may have capitalized on the higher density of tumor infiltrating T cells (Supplemental Fig 7E) to generate ongoing, multi-clonal responses in tumors. We considered that the number of public TCRs – i.e., identical TCRs specific for the same critically immunogenic epitopes – could represent common responses within each cohort. To investigate this, we identified the 50 most frequent CDR3β amino acid sequences from the TCR repertoire of each patient and compiled them in to four separate lists: liver-cohort blood, liver-cohort tumor, lung-cohort blood and lung-cohort tumor (Supplemental Dataset 6), and then looked for these clonal sequences in the entire TCR repertoires of all samples in the same cohort. We then identified any of these clonal CDR3b sequences that were shared by at least 25% of the samples in the same cohort. Thus, we identified CDR3b sequences that were *shared* by a large fraction of samples in the cohort and were dominantly *clonal* in at least one sample. None of the shared, clonal sequences

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

matched those reported to be mutant KRAS specific[61]. We found similar numbers of these shared, clonal CDR3β sequences in liver-cohort tumors (N=21) as in lung-cohort tumors (N=24); in contrast, we found many more shared, clonal CDR3β sequences among blood samples from patients in the liver cohort (N=174) than from patients in the lung cohort (N=41, Figure 6B). This result is consistent with repeated responses to a limited set of common tumor-initiating neoepitopes in liver-cohort tumors over time, increasing the chance that tumors from different patients would generate identical TCRs. These shared, clonal sequences may accumulate in the periphery even though they are not found in many liver-cohort tumors at any one point in time. In contrast to responses to putative mutant KRAS-specific tumor-initiating epitopes, the shared, clonal responses in lung-cohort tumors, existed at higher frequencies than in liver-cohort tumors (Figure 6C). These results suggest that liver-cohort tumors may promote clonal responses against persistent neoepitopes (Figure 6A), whereas lung-cohort tumors may be more likely to make clonal responses against new (though uncharacterized) tumor antigens (Figure 6C) as they respond to new antigens over time (Figure 5D). As expected, some of the shared, clonal CDR3b sequences have been previously identified in TCRs known to recognize common infectious agents in humans[62] (Figure 6B and C), suggesting that poor expansion of T cell clones may be a characteristic of liver-cohort patients regardless of relevance to tumor immunity. We considered that the shared, clonal sequences we identified in lung-cohort tumors may also be present in some liver-cohort tumors (and vice versa) as public TCRs are unlikely to be perfectly cohort-specific. In fact, we found the shared, clonal sequences from each tumor cohort in the other tumor cohort; but we were more likely to find shared, clonal liver-cohort tumor CDR3b sequences in lung-cohort tumors compared to the inverse (Figure 6D). This underscores that these public/shared, clonal sequences are being generated by both cohorts, but more commonly present (Figure 6D) and expanded (Figure 6C) in lung-cohort tumors.

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

PDAC tumors that promote new, clonal T cell responses over time may be more likely to generate T cell-mediated tumor immunity that suppresses growth of liver metastases. In lung-cohort patients, this seemingly more effective tumor immunity exists in patients concurrently with metastatic disease (even at resection, patients with future recurrent disease must have occult metastases), however, we expected that metastases would harbor less clonal outgrowth as progressive disease escapes tumor immunity. In fact, we found this to be true for metastases in general, as they had significantly greater evenness in their TCR repertoires than did primary tumors (Figure 6E), consistent with our data above showing that greater TCR evenness in tumors correlates with fewer tumor-distinct clones (Supplemental Figure 8E). In stark contrast though, the three metastases we were able to analyze from our lung cohort harbored more tumor-distinct clones than any other metastases in our study and lower evenness (Figure 6E, F). Together these data support the concept that lung-cohort patients are uniquely able to generate ongoing, clonal T cell responses to critically immunogenic antigens in both their primary and metastatic tumors, likely contributing to greater peripheral TCR diversity and possibly suppression of liver metastases.

Discussion

Previous efforts to dissect PDAC tumors into subtypes used unbiased approaches with gene expression, proteomics, or metabolomics data to describe mutually exclusive subsets and defining signatures[25,26,28,29,31,33]. We approached the problem differently by classifying tumors based on a well-known clinical outcome associated with metastatic organotropism; specifically, longer-term survival among patients with lung-avid, liver-averse disease and poor outcomes with liver-avid disease[14–16,63]. We generated two new large datasets obtained from only patients with proven adenocarcinoma in the pancreas or Ampulla of Vater tumors that were histologically pancreaticobiliary: 1) a 290 specimen RNA-seq dataset with genomic alterations from tumor-enriched regions from primary and metastatic

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

tumors and 2) TCRB sequencing of 289 blood and 175 matched tumor samples, mostly overlapping with the RNA-seq dataset. As previously reported[24], tumors from patients in the lung cohort were unlikely to be categorized as the basal-like subtype; whereas classical subtype tumors were common in both cohorts. Uniquely, we found that patients with classical-subtype primary tumors fared significantly worse if their disease was liver-avid rather than lung-avid and liver-averse. This finding led us to investigate whether a unique gene signature exists for liver-avid tumors apart from gene-expression related to previously described subtypes.

We extracted a gene-expression signature for liver-avid primary tumors (pORG) that was devoid of genes overexpressed in basal-like primary tumors, addressing the overlap between these two subtypes and demonstrated that the pORG signature can independently predict patient outcome. High pORG and liver-avid primary tumors were enriched for pathways indicating ongoing replication stress and the response to replication stress by engagement of DNA repair processes, along with an increase in the activity of cell cycle drivers and mitogenic signaling (Figure 3 and 4, Supplemental Figure 6). Liver-avid tumors with somatic alterations in DDR genes had some of the highest pORG signature scores (Figure 3F), suggesting that PDAC tumor cells can avoid the detrimental effects of ongoing DNA damage by adopting strong RS response mechanisms. Conversely, patients live longer if their PDAC tumors have low pORG scores (Figure 2C), particularly if they harbor a DDR gene mutation (Figure 3E) and likely fail to productively respond to replication stress caused by DDR defects. A high pORG signature and ongoing response to replication stress was also associated with chronic interferon signaling, which unlike an acute type 1 interferon response, can induce an IFN-related DNA Damage Resistance Signature (IRDS)[53–55] and reduced tumor immunity[56].

Diverse TCRB repertoires associate with positive outcomes in PDAC patients[41,64]. Consistent with these reports, we identified high TCR evenness/diversity in peripheral blood from lung-cohort patients relative to patients in the liver cohort and most other patients with PDAC (Figure 5C and

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplemental Figure 7C) suggesting that there is something unique about tumor immunity in patients with lung-avid, liver-averse disease. We also found that lung-cohort patients had more T cell clones found in their tumors that were absent in paired blood samples compared to liver-cohort patients (Figure 5D); suggesting that there are new clonal responses in lung-cohort tumors that have not yet emigrated to blood, consistent with greater diversity building in the periphery over time (Supplemental Figure 7D). There are several possible explanations for enhanced T cell clonal development in lung-cohort tumors, including: a predilection for diversity in the pre-immune repertoire, greater adjuvanticity of the tumor (including the microenvironment) that promotes responses to many antigens, or a greater number of immunogenic neoepitopes or other tumor associated antigens generated in the tumor over time. A thorough comparison of immune competence in lung-cohort patients along with the immune contexture and neoepitope burden of their tumors may help parse these etiologies. It is also possible that only patients who stochastically develop a diverse T cell repertoire directed toward the right antigens end up with liver-averse disease.

Our results suggest that patients who develop liver metastases (a common outcome in PDAC) – relative to those with lung-avid, liver-averse disease – are more likely to generate clonal responses against persistent neoepitopes present since tumor initiation, at the expense of generating clonal responses to critically immunogenic tumor-associated antigens as they arise over time. This is consistent with a liver-avid primary tumor environment that favors weakly clonal responses of T cells prone to exhaustion over robust naïve T cell expansion against newly arising antigens[61]. Tumors are regularly refreshed with new T cell clones from the blood[65], which in liver-cohort tumors may lead to a repeated cycle of T cell tumor exit, tumor entry, and weak clonal responses to a small pool of tumor initiating antigens. Our data supports that the distinction between high frequency T cell clones in liver-cohort and lung-cohort tumors (Figures 6A and C, respectively) is that the former represents clonal responses to a small pool of tumor initiating antigens that are decreasingly effective and the latter

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

represents T cell expansion against new immunogenic antigens on which many lung-cohort patients have converged. In addition to this, we found that lung-cohort patients were uniquely able to generate tumor-distinct clones in both primary and metastatic tumors (Figure 6F), indicating that ongoing T cell clonal responses at several tumor sites may contribute to systemic resistance to PDAC progression.

In this report we identified a gene-expression signature representing ongoing replication stress response (i.e., high DNA replication and persistent mitogenesis and metabolism, along with concurrent DNA damage repair processes) that is associated with liver-avid, primary PDAC tumors. Tumors with low enrichment for this signature were less aggressive, particularly if DDR-participating genes were altered, suggesting that low pORG tumors may be more sensitive to therapeutics that interrupt DDR pathways. In concert with the low RS response of low pORG tumors, we found a high density of T cells and TCR sequences (Figure 4F and 5A) – one factor that may contribute to better outcomes for patients with low pORG tumors. A small subset of patients with low pORG tumors may direct those dense T cell infiltrates toward ongoing, clonal responses to critically immunogenic antigens leading to PDAC-relevant TCR diversity in the periphery and liver-averse disease. Our results point to two avenues for therapeutic intervention that may function additively: 1) exploitation of tumors lacking RS response mechanisms by promoting DNA damage (e.g., platinum chemotherapeutics) and interrupting DDR pathways (e.g., PARP inhibitors), and 2) immunomodulatory approaches (e.g., vaccines and immune checkpoint inhibition) that promote ongoing T cell responses to new tumor-associated antigens over time.

Acknowledgments

We thank Jen Jen Yeh and laboratory for advice in performing PurIST analyses. We thank members of the Brenden-Colson Center for Pancreatic Care, the Sears Lab, and the Brody Lab for helpful discussions. We thank the Oregon Pancreas Tissue Registry, the OHSU Research Data Warehouse, and the Knight Biomedical Library for providing de-identified patient metadata and specimens, and the Knight Biostatistics Shared Resource for helpful discussion. The Knight shared resources are

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

supported by Knight Cancer Institute P30 CA69533. These studies were supported by NIH-NCI U54CA209988, U01CA224012 and R01s CA186241, CA196228 to RCS; NIH-NCI 1U01 CA253472 and 5U01 CA217842 to GBM; and NIH-NCI R01 CA212600, U01CA224012-03, AACR Grant-15-90-25-BROD, and the Hirshberg Foundation to JBM; and foundation support from the Brenden-Colson Center for Pancreatic Care.

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

References

1. Beaubier N, Tell R, Lau D, et al. Clinical validation of the tempus xT next-generation targeted oncology sequencing assay. Oncotarget 2019;10:2384–2396.

2. Skidmore ZL, Wagner AH, Lesurf R, et al. GenVisR: Genomic Visualizations in R. Bioinforma Oxf Engl 2016;32:3012–3014.

3. Köster J, Rahmann S. Snakemake—a scalable bioinformatics workflow engine. Bioinformatics 2012;28:2520–2522.

4. Bray NL, Pimentel H, Melsted P, et al. Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol 2016;34:525–527.

5. Rashid NU, Peng XL, Jin C, et al. Purity Independent Subtyping of Tumors (PurIST), A Clinically Robust, Single-sample Classifier for Tumor Subtyping in Pancreatic Cancer. Clin Cancer Res 2020;26:82–92.

6. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15:550.

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

8. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005;102:15545–15550.

9. Liberzon A, Birger C, Thorvaldsdóttir H, et al. The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst 2015;1:417–425.

10. Alvarez MJ, Shen Y, Giorgi FM, et al. Functional characterization of somatic mutations in cancer using network-based inference of protein activity. Nat Genet 2016;48:838–847.

11. Lachmann A, Giorgi FM, Lopez G, et al. ARACNe-AP: gene network reverse engineering through adaptive partitioning inference of mutual information. Bioinformatics 2016;32:2233–2235.

12. Bareche Y, Buisseret L, Gruosso T, et al. Unraveling Triple-Negative Breast Cancer Tumor Microenvironment Heterogeneity: Towards an Optimized Treatment Approach. JNCI J Natl Cancer Inst 2019;112:708–719.

13. Tamborero D, Rubio-Perez C, Muiños F, et al. A Pan-cancer Landscape of Interactions between Solid Tumors and Infiltrating Immune Cell Populations. Clin Cancer Res 2018;24:3717–3728.

14. Downs-Canner S, Zenati M, Boone BA, et al. The Indolent Nature of Pulmonary Metastases from Ductal Adenocarcinoma of the Pancreas. J Surg Oncol 2015;112:80–85.

15. He C, Huang X, Zhang Y, et al. The impact of different metastatic patterns on survival in patients with pancreatic cancer. Pancreatology 2021;21:556–563.

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

16. Liang X, Cheng Y, Zhou W, et al. Systematic Pan-Cancer Population-Based Analysis Reveals the Incidence and Prognosis of Lung Metastases at Diagnosis. J Oncol 2021;2021:9999968.

17. Wangjam T, Zhang Z, Zhou XC, et al. Resected pancreatic ductal adenocarcinomas with recurrence limited in lung have a significantly better prognosis than those with other recurrence patterns. Oncotarget 2015;6:36903–36910.

18. Link JM, Liudahl SM, Betts CB, et al. Tumor-Infiltrating Leukocyte Phenotypes Distinguish Outcomes in Related Patients With Pancreatic Adenocarcinoma. JCO Precis Oncol 2021;5:PO.20.00287.

19. Arnaoutakis GJ, Rangachari D, Laheru DA, et al. Pulmonary Resection for Isolated Pancreatic Adenocarcinoma Metastasis: An Analysis of Outcomes and Survival. J Gastrointest Surg Off J Soc Surg Aliment Tract 2011;15:1611–1617.

20. Lee JW, Stone ML, Porrett PM, et al. Hepatocytes direct the formation of a pro-metastatic niche in the liver. Nature 2019;567:249–252.

21. Yu J, Green MD, Li S, et al. Liver metastasis restrains immunotherapy efficacy via macrophage-mediated T cell elimination. Nat Med 2021;27:152–164.

22. Houg DS, Bijlsma MF. The hepatic pre-metastatic niche in pancreatic ductal adenocarcinoma. Mol Cancer 2018;17:95.

23. O'Kane GM, Grünwald BT, Jang G-H, et al. GATA6 Expression Distinguishes Classical and Basal-like Subtypes in Advanced Pancreatic Cancer. Clin Cancer Res 2020;26:4901–4910.

24. Dreyer SB, Upstill-Goddard R, Legrini A, et al. Genomic and molecular analyses identify molecular subtypes of pancreatic cancer recurrence. Gastroenterology 2021. Available at: https://www.sciencedirect.com/science/article/pii/S0016508521035228 [Accessed September 18, 2021].

25. Collisson EA, Sadanandam A, Olson P, et al. Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy. Nat Med 2011;17:500–503.

26. Bailey P, Chang DK, Nones K, et al. Genomic analyses identify molecular subtypes of pancreatic cancer. Nature 2016;531:47–52.

27. Chan-Seng-Yue M, Kim JC, Wilson GW, et al. Transcription phenotypes of pancreatic cancer are driven by genomic events during tumor evolution. Nat Genet 2020:1–10.

28. Moffitt RA, Marayati R, Flate EL, et al. Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma. Nat Genet 2015;advance online publication. Available at: http://www.nature.com/ng/journal/vaop/ncurrent/full/ng.3398.html [Accessed September 15, 2015].

29. Puleo F, Nicolle R, Blum Y, et al. Stratification of Pancreatic Ductal Adenocarcinomas Based on Tumor and Microenvironment Features. Gastroenterology 2018;155:1999-2013.e3.

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

30. Raphael BJ, Hruban RH, Aguirre AJ, et al. Integrated Genomic Characterization of Pancreatic Ductal Adenocarcinoma. Cancer Cell 2017;32:185-203.e13.

31. Daemen A, Peterson D, Sahu N, et al. Metabolite profiling stratifies pancreatic ductal adenocarcinomas into subtypes with distinct sensitivities to metabolic inhibitors. Proc Natl Acad Sci U S A 2015.

32. Law HC-H, Lagundžin D, Clement EJ, et al. The Proteomic Landscape of Pancreatic Ductal Adenocarcinoma Liver Metastases Identifies Molecular Subtypes and Associations with Clinical Response. Clin Cancer Res 2020;26:1065–1076.

33. Knudsen ES, Vail P, Balaji U, et al. Stratification of Pancreatic Ductal Adenocarcinoma: Combinatorial Genetic, Stromal, and Immunologic Markers. Clin Cancer Res 2017;23:4429–4440.

34. Noll EM, Eisen C, Stenzinger A, et al. CYP3A5 mediates basal and acquired therapy resistance in different subtypes of pancreatic ductal adenocarcinoma. Nat Med 2016;22:278–287.

35. Bernard V, Semaan A, Huang J, et al. Single-Cell Transcriptomics of Pancreatic Cancer Precursors Demonstrates Epithelial and Microenvironmental Heterogeneity as an Early Event in Neoplastic Progression. Clin Cancer Res 2019;25:2194–2205.

36. Laise P, Turunen M, Maurer HC, et al. Pancreatic Ductal Adenocarcinoma Comprises Coexisting Regulatory States with both Common and Distinct Dependencies. bioRxiv 2020:2020.10.27.357269.

37. Dreyer SB, Upstill-Goddard R, Paulus-Hock V, et al. Targeting DNA Damage Response and Replication Stress in Pancreatic Cancer. Gastroenterology 2021;160:362-377.e13.

38. Labrie M, Brugge JS, Mills GB, et al. Therapy resistance: opportunities created by adaptive responses to targeted therapies in cancer. Nat Rev Cancer 2022:1–17.

39. Zheng L, Ding D, Edil BH, et al. Vaccine-Induced Intratumoral Lymphoid Aggregates Correlate with Survival Following Treatment with a Neoadjuvant and Adjuvant Vaccine in Patients with Resectable Pancreatic Adenocarcinoma. Clin Cancer Res 2021;27:1278–1286.

40. Morrison AH, Byrne KT, Vonderheide RH. Immunotherapy and Prevention of Pancreatic Cancer. Trends Cancer 2018;4:418–428.

41. Balachandran VP, Łuksza M, Zhao JN, et al. Identification of unique neoantigen qualities in long-term survivors of pancreatic cancer. Nature 2017;551:512–516.

42. Le DT, Durham JN, Smith KN, et al. Mismatch-repair deficiency predicts response of solid tumors to PD-1 blockade. Science 2017;357:409–413.

43. Lawlor RT, Mattiolo P, Mafficini A, et al. Tumor Mutational Burden as a Potential Biomarker for Immunotherapy in Pancreatic Cancer: Systematic Review and Still-Open Questions. Cancers 2021;13:3119.

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

44. Pishvaian MJ, Blais EM, Brody JR, et al. Overall survival in patients with pancreatic cancer receiving matched therapies following molecular profiling: a retrospective analysis of the Know Your Tumor registry trial. Lancet Oncol 2020;21:508–518.

45. Li H, Merwe PA van der, Sivakumar S. Biomarkers of response to PD-1 pathway blockade. Br J Cancer 2022:1–13.

46. Makohon-Moore AP, Zhang M, Reiter JG, et al. Limited heterogeneity of known driver gene mutations among the metastases of individual patients with pancreatic cancer. Nat Genet 2017;49:358–366.

47. Hapke RY, Haake SM. Hypoxia-induced epithelial to mesenchymal transition in cancer. Cancer Lett 2020;487:10–20.

48. Heeke AL, Pishvaian MJ, Lynce F, et al. Prevalence of Homologous Recombination–Related Gene Mutations Across Multiple Cancer Types. JCO Precis Oncol 2018;2. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6139373/ [Accessed October 31, 2020].

49. Herold S, Herkert B, Eilers M. Facilitating replication under stress: an oncogenic function of MYC? Nat Rev Cancer 2009;9:441–444.

50. Carey JasonPW, Karakas C, Bui T, et al. Synthetic Lethality of PARP Inhibitors in Combination with MYC Blockade Is Independent of BRCA Status in Triple-Negative Breast Cancer. Cancer Res 2018;78:742–757.

51. Musella M, Manic G, De Maria R, et al. Type-I-interferons in infection and cancer: Unanticipated dynamics with therapeutic implications. Oncoimmunology 2017;6:e1314424.

52. McNab F, Mayer-Barber K, Sher A, et al. Type I interferons in infectious disease. Nat Rev Immunol 2015;15:87–103.

53. Weichselbaum RR, Ishwaran H, Yoon T, et al. An interferon-related gene signature for DNA damage resistance is a predictive marker for chemotherapy and radiation for breast cancer. Proc Natl Acad Sci U S A 2008;105:18490–18495.

54. Khodarev NN, Beckett M, Labay E, et al. STAT1 is overexpressed in tumors selected for radioresistance and confers protection from radiation in transduced sensitive cells. Proc Natl Acad Sci U S A 2004;101:1714–1719.

55. Duarte CW, Willey CD, Zhi D, et al. Expression Signature of IFN/STAT1 Signaling Genes Predicts Poor Survival Outcome in Glioblastoma Multiforme in a Subtype-Specific Manner. PLoS ONE 2012;7:e29653.

56. Musella M, Galassi C, Manduca N, et al. The Yin and Yang of Type I IFNs in Cancer Promotion and Immune Activation. Biology 2021;10:856.

57. Boukhaled GM, Harding S, Brooks DG. Opposing Roles of Type I Interferons in Cancer Immunity. Annu Rev Pathol Mech Dis 2021;16:167–198.

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

58. Bindea G, Mlecnik B, Tosolini M, et al. Spatiotemporal Dynamics of Intratumoral Immune Cells Reveal the Immune Landscape in Human Cancer. Immunity 2013;39:782–795.

59. Chtanova T, Newton R, Liu SM, et al. Identification of T cell-restricted genes, and signatures for different T cell responses, using a comprehensive collection of microarray datasets. J Immunol Baltim Md 1950 2005;175:7837–7847.

60. DeWolf S, Grinshpun B, Savage T, et al. Quantifying size and diversity of the human T cell alloresponse. JCI Insight 3:e121256.

61. Lowery FJ, Krishna S, Yossef R, et al. Molecular signatures of antitumor neoantigen-reactive T cells from metastatic human cancers. Science 2022;375:877–884.

62. Bagaev DV, Vroomans RMA, Samir J, et al. VDJdb in 2019: database extension, new analysis infrastructure and a T-cell receptor motif compendium. Nucleic Acids Res 2020;48:D1057–D1062.

63. Kruger SF, Lohneis A, Abendroth A, et al. Prognosis and tumor biology of pancreatic cancer patients with isolated lung metastases: translational results from the German multicenter AIO-YMO-PAK-0515 study. ESMO Open 2022;7. Available at: https://www.esmoopen.com/article/S2059-7029(22)00008-4/fulltext [Accessed February 6, 2022].

64. Hopkins AC, Yarchoan M, Durham JN, et al. T cell receptor repertoire features associated with survival in immunotherapy-treated pancreatic ductal adenocarcinoma. JCI Insight 3. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6124515/ [Accessed September 21, 2020].

65. Wu TD, Madireddi S, Almeida PE de, et al. Peripheral T cell expansion predicts tumour infiltration and clinical response. Nature 2020;579:274–278.

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure Legends

Figure 1: Outcomes for patients with liver metastases with or without lung metastases (liver cohort) and lung metastases without liver metastases (lung cohort): A) Overall survival, B) Days between resection and recurrence, and C) Survival since resection. D) PusIST subtyping scores for primary (black) and metastatic (green) tumor specimens from patients in the liver and lung cohorts. Blue bars represent means. Two-tailed Students t-test P value is reported. E) Overall survival and F) Days to recurrence for patients categorized by PurIST subtype and liver/lung cohorts. For charts with survival endpoints, patients not deceased were censored at last known vital status. Patients who died <30 days after resection were omitted. A Mantel-Cox test was used to compare all Kaplan-Meier survival and recurrence curves.

Figure 2: GSVA scores for the pORG (left panels) and pSUB (right panels) gene sets were calculated for the indicated cohorts from A) primary tumors and B) metastatic tumors. Blue bars represent means. One-way ANOVA was used to calculate indicated P values. C) Patient survival based on top and bottom quartiles of their primary tumors by pORG (left panels) and pSUB (right panels) GSVA scores from the OHSU dataset (top panels) and TCGA PAAD dataset (bottom panels). Patients who died within 30 days after resection are not shown. A Mantel-Cox test was used to compare Kaplan-Meier survival curves.

Figure 3: A) Normalized enrichment scores (NES) and are shown for hallmark GSEA pathways if any of the comparisons listed on each plot reached a NES >1.7 and nominal P<0.05. B) Heatmaps of GSVA scores (high=red, low=green) for primary tumors showing pathway (NES >1.7 and nominal P<0.05) and specimen clustering from the cohorts indicated on each plot. C) and D) Displayed as in A and B for the indicated cohort comparisons. E) Overall survival for patients with high or low pORG

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

GSVA scores (top 50% or bottom 50%) stratified by known presence or absence of a somatic alteration in a DDR-related gene. A Mantel-Cox test was used to compare Kaplan-Meier survival curves. F) pORG GSVA scores for primary tumors from patients in the liver cohort categorized by a known alteration (DDR defect) or no variant detected (DDR intact) in a DDR-related gene. Two-tailed Students t-test P value is reported.

Figure 4: Mean differential VIPER regulon activity scores with significant ($P<0.05$ for all shown; one-way ANOVA) enrichment in high pORG vs low pORG (red) and liver-cohort vs lung-cohort (blue) primary tumors for regulons related to A) DNA replication B) DNA damage repair C) Cell cycle. D) GSEA plots and normalized enrichment scores (NES) for the IRDS gene set for the indicated comparisons. E) Mean differential VIPER regulon activity scores with significant ($P<0.05$ for all shown; one-way ANOVA) enrichment for 2 IRDS-related regulons in high pORG vs low pORG (red) and liver-cohort vs lung-cohort (blue) primary tumors. F) Enrichment of gene sets defining leukocyte subtypes in low pORG vs high pORG (pink) and lung-cohort vs liver-cohort (light blue) primary tumors ($P<0.05$ for pORG comparisons; one-way ANOVA).

Figure 5: A) The density of productive TCRB templates (per ng of tumor tissue) in tumors from the indicated cohorts. B) The Evenness Index of the TCRB repertoires from tumors in each indicated cohort. C) The Evenness Index of the TCRB repertoires from blood samples in each indicated cohort. D) The percentage of unique tumor TCRB CDR3β sequences with ≥10 templates detected in tumor samples, but none detected in paired (i.e., from the same patient) blood samples. For all, blue bars represent means. One-way ANOVA was used to calculate indicated P values.

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 6: A) The frequency of any putative mutant KRAS-specific CDR3β sequences within the TCR repertoire for each sample containing at least one of those sequences (each dot represents one sample). B) The number of shared, clonal sequences found for each cohort; red fractions of bars represent the number of CDR3β sequences with reported specificity for common human pathogens. C) The average frequency of each shared, clonal CDR3β sequence in all samples containing that sequence (each dot represents one shared, clonal CDR3β sequence); red dots represent CDR3β sequences with reported specificity for a common human pathogen. D) The percentage of samples containing at least one of the alternate cohort's shared, clonal sequences (each dot represents one CDR3β sequence). E) The Evenness Index score for all primary and metastatic samples; purple dots represent samples from liver-cohort patients and blue dots represent samples from lung-cohort patients. F) The percentage of tumor-distinct clones in primary and metastatic tumors. For all, blue bars represent means. One-way ANOVA was used to calculate indicated P values for A, C, and F. Two-tailed Student's t-test was used to calculate P values for D and E.

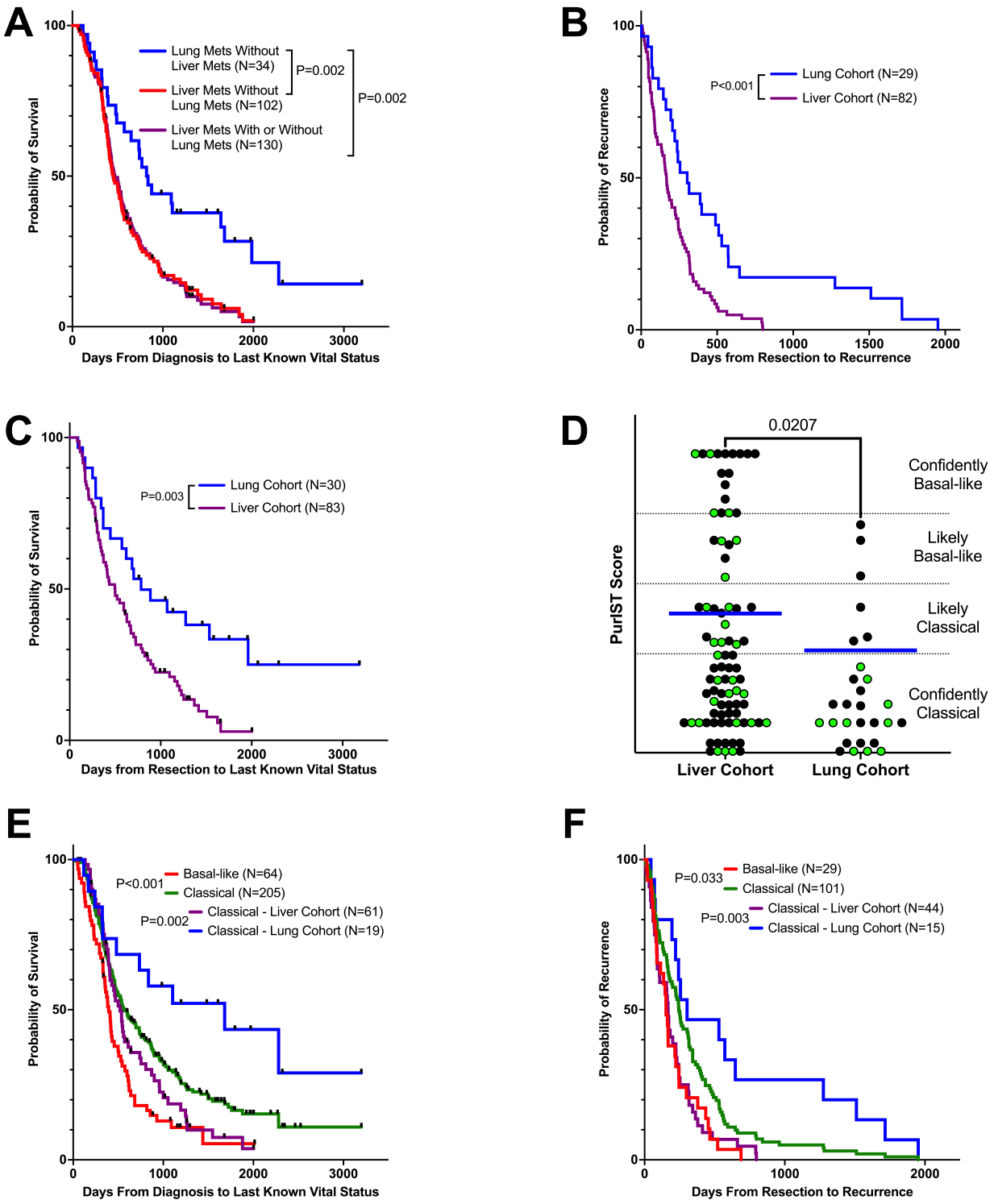
bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

		Liver Cohort		Lung Cohort	
Clinical Characteristics		N=130		N=35	
Gender	Female	59	45%	14	40%
	Male	71	55%	21	60%
Race	White	119	92%	32	91%
	Asian	5	3%	1	3%
	Unknown	6	5%	2	6%
Primary Tumor Site	Pancreas (Adenocarcinoma Only)	127	98%	35	100%
	Ampulla of Vater (Pancreaticobiliary Type Only)	3	2%	0	0%
Stage at Diagnosis	Stage 1a	1	1%	1	3%
	Stage 1b	9	7%	4	11%
	Stage 2a	15	12%	6	17%
	Stage 2b	51	39%	17	49%
	Stage 3	19	15%	3	9%
	Stage 4	33	25%	4	11%
	No Data	2	2%	0	0%
Grade	1 - Well Differentiated	1	1%	2	6%
	2 - Moderately Differentiated	49	38%	8	23%
	3 - Poorly Differentiated	27	21%	12	34%
	Not Determined	53	41%	13	37%
Treated by resection		N=83	64% (of all)	N=31	89% (of all)
Resection Details	Neoadjuvant Treatment	17	20%	11	35%
	No Residual Tumor	73	88%	28	90%
	Microscopic Residual Disease Present	10	12%	3	10%
	Angiolymphatic Invasion	63	76%	18	58%
	Tumor Involved in Regional Lymph Nodes	62	75%	20	65%
Histology Analysis Performed		N=67		N=24	
Histology	Acute Inflammation	34	51%	9	38%
	Chronic Inflammation	39	58%	12	50%
	Plasmacytoid Inflammation	21	31%	12	50%
	Lymphoid Aggregates/TLSs	21	31%	10	42%
	Perineural Invasion	19	28%	9	38%
	Desmoplasia	64	96%	21	88%

Table 1: Patient demographics, disease characteristics, and tumor-specimen histology parameters for patients categorized into liver and lung cohorts. Percentages for resection details are only from primary tumor resections and percentages for histology are only from tumors with histology analyzed. Percentage of patients treated by resection was the only comparison significantly different between liver and lung cohorts ($P<0.001$, two-tailed Fisher's exact test).

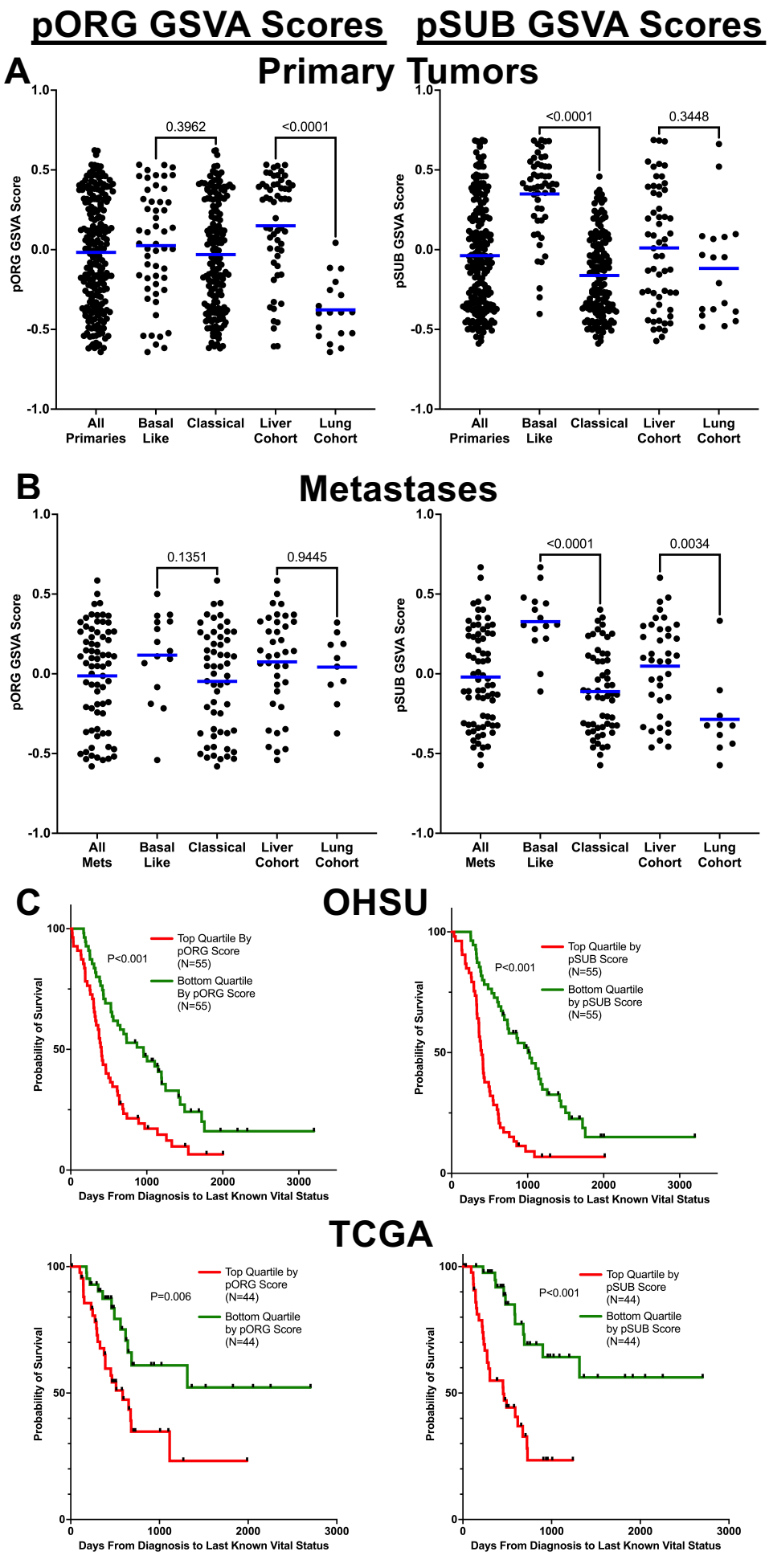
bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 1



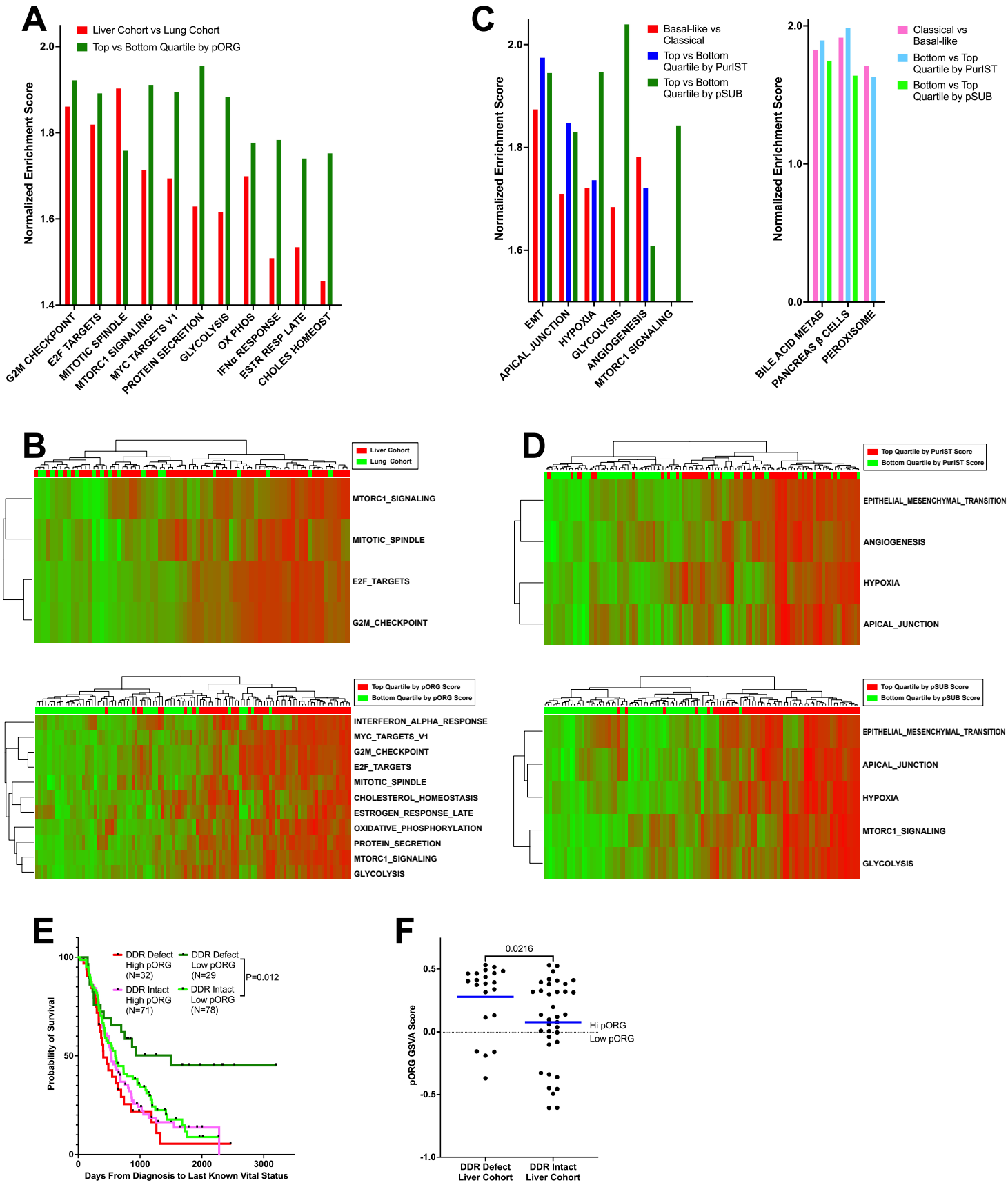
bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 2



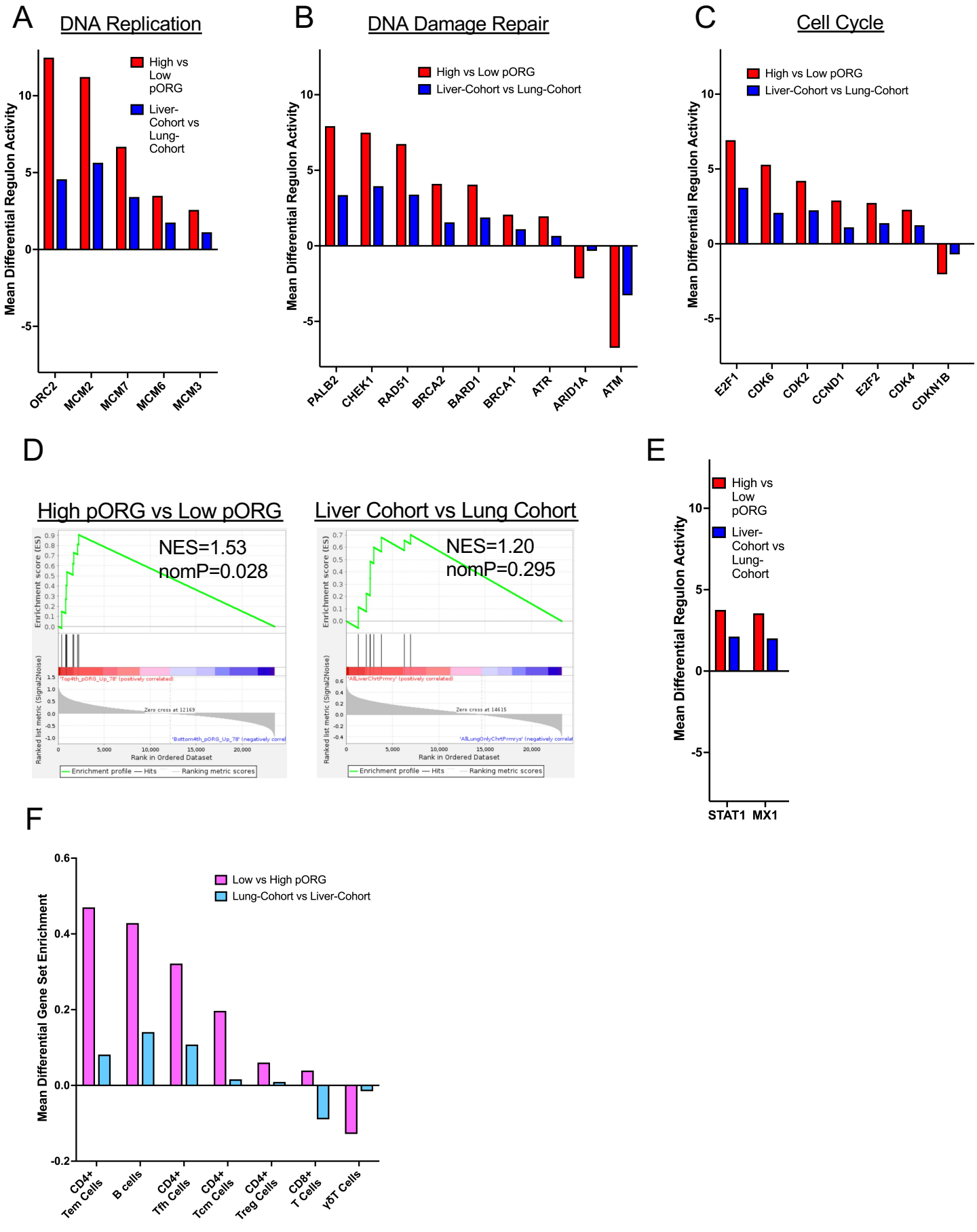
bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 3



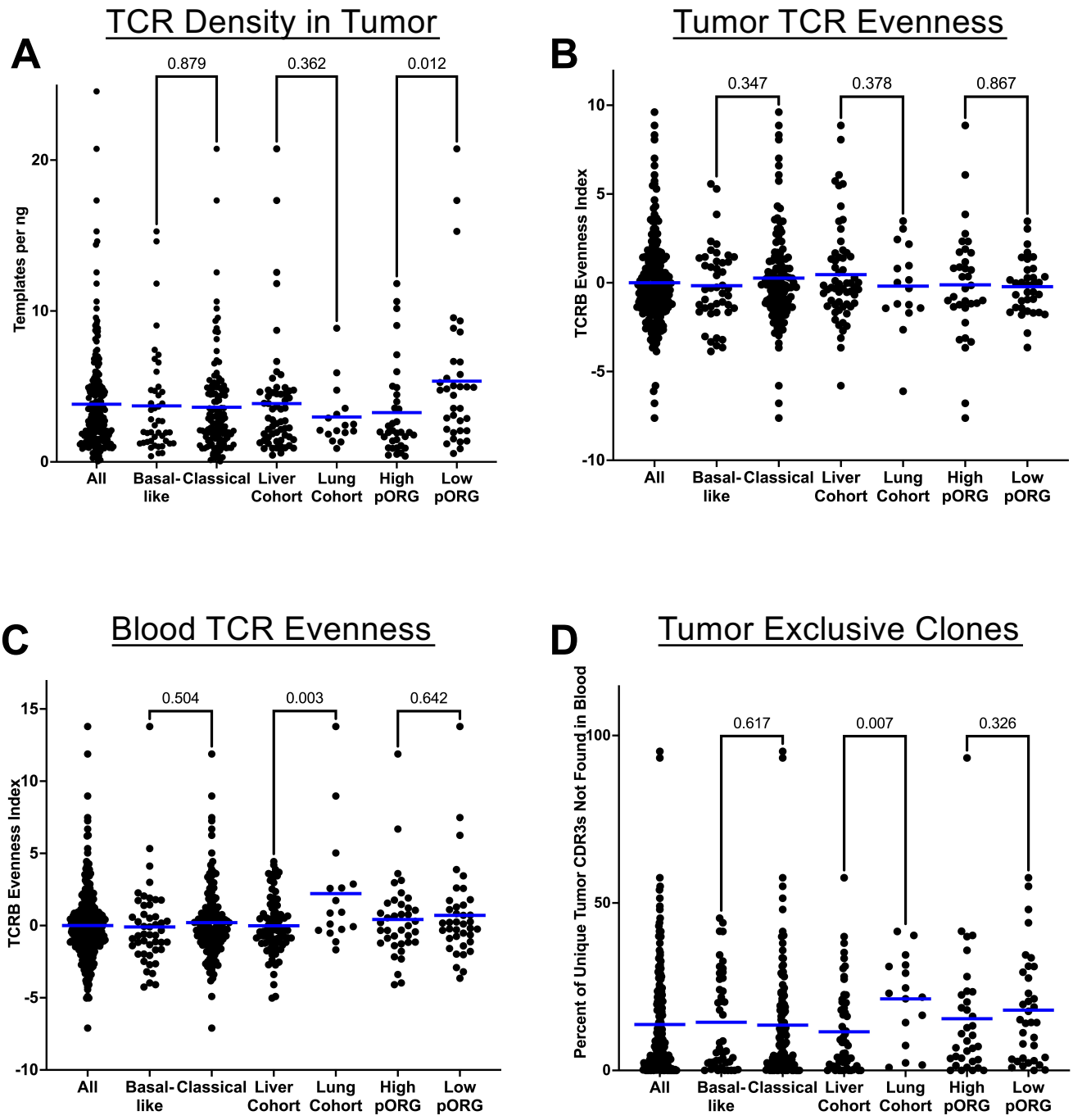
bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 4



bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 5



bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490552; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

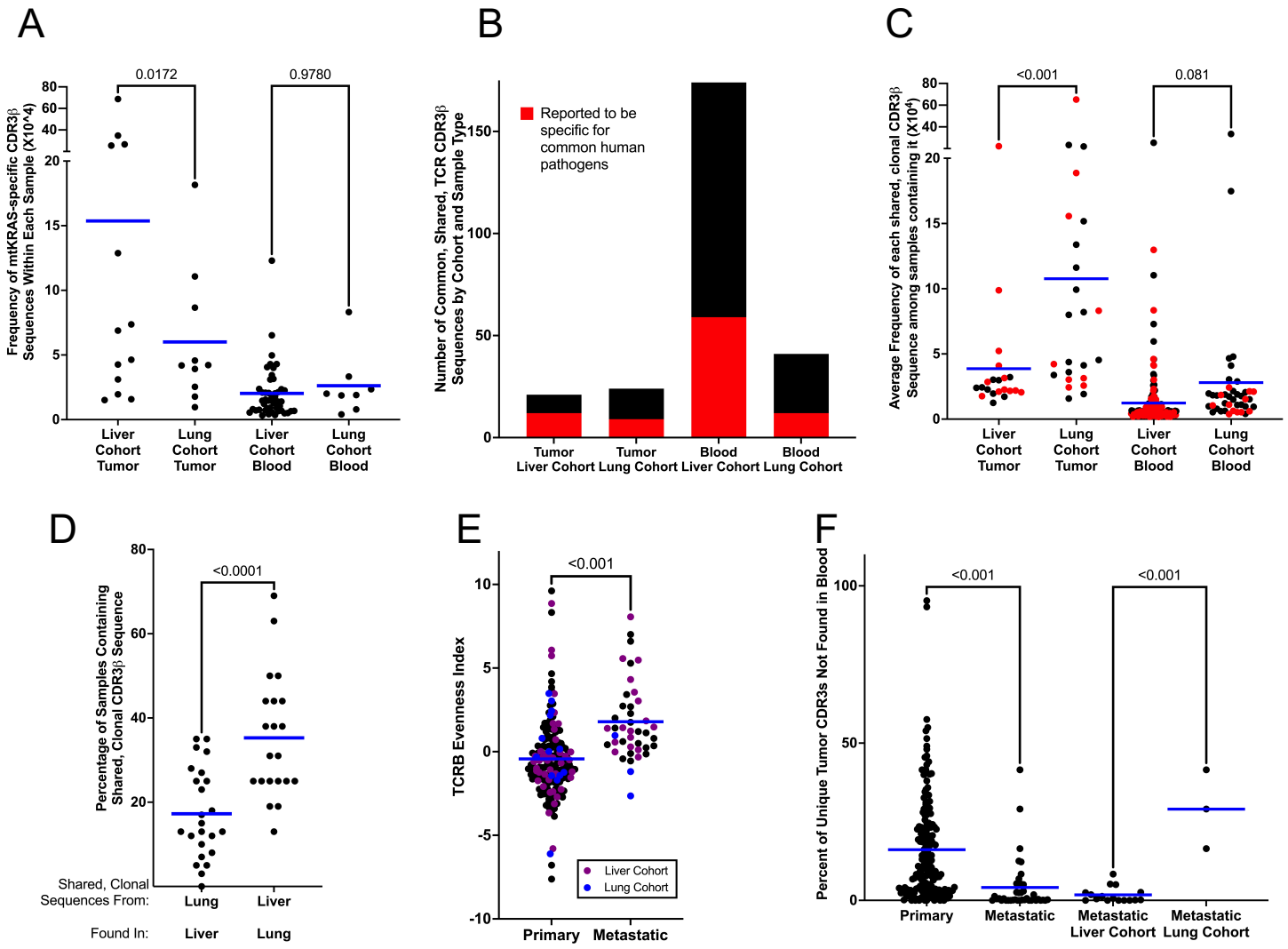



Figure 6