

Clonal Heterogeneity in Human Pancreatic Ductal Adenocarcinoma and Its Impact on Tumor Progression

Corresponding Author: Dr Diane Simeone

Version 0:

Decision Letter:

7th Apr 2025

Dear Dr Simeone,

Your Article, "Clonal Heterogeneity in Human Pancreatic Ductal Adenocarcinoma and Its Impact on Tumor Progression" has now been seen by 3 referees. You will see from their comments copied below that while they find your work of considerable potential interest, they have raised quite substantial concerns that must be addressed. In light of these comments, we cannot accept the manuscript for publication, but would be very interested in considering a revised version that addresses these serious concerns.

Reviewer #1 thinks the dataset will likely be of great value for the PDAC community. They have useful guidance for improvement.

Reviewer #2 thinks the work requires substantial re-analysis of the data by integrating more quantitative and statistical approaches. They also suggest to further validate the results in human cell lines. They generally think the level of new biology is limited.

Reviewer #3 is also concerned about the lack of sufficient biology and also requires a more in-depth analysis of the dataset.

We hope you will find the referees' comments useful as you decide how to proceed. If you wish to submit a substantially revised manuscript, please bear in mind that we will be reluctant to approach the referees again in the absence of major revisions.

To guide the scope of the revisions, the editors discuss the referee reports in detail within the team, including with the chief editor, with a view to identifying key priorities that should be addressed in revision and sometimes overruling referee requests that are deemed beyond the scope of the current study. In this case, we ask you to address reviewers' comments in full. We hope that you will find the prioritised set of referee points to be useful when revising your study. Please do not hesitate to get in touch if you would like to discuss these issues further.

If you choose to revise your manuscript taking into account all reviewer and editor comments, please highlight all changes in the manuscript text file. At this stage we will need you to upload a copy of the manuscript in MS Word .docx or similar editable format.

We are committed to providing a fair and constructive peer-review process. Do not hesitate to contact us if there are specific requests from the reviewers that you believe are technically impossible or unlikely to yield a meaningful outcome.

If revising your manuscript:

*1) Include a "Response to referees" document detailing, point-by-point, how you addressed each referee comment. If no action was taken to address a point, you must provide a compelling argument. This response will be sent back to the referees along with the revised manuscript.

*2) If you have not done so already please begin to revise your manuscript so that it conforms to our Article format instructions, available here. Refer also to any guidelines provided in this letter.

*3) Include a revised version of any required Reporting Summary: <https://www.nature.com/documents/nr-reporting-summary.pdf>

It will be available to referees (and, potentially, statisticians) to aid in their evaluation if the manuscript goes back for peer review.

A revised checklist is essential for re-review of the paper.

Please be aware of our [guidelines](https://www.nature.com/nature-research/editorial-policies/image-integrity) on digital image standards.

EXTENDED DATA FIGURES

When re-submitting your manuscript, please ensure that any supplementary figures and tables that are crucial to the manuscript's conclusions are converted into Extended Data figures and tables to increase visibility of these data. Extended Data figures and tables are online-only (present in the online PDF and full-text HTML versions of the paper), peer-reviewed display items that provide essential background to the article but are not included in the main article due to space constraints. A maximum of ten Extended Data display items (figures and tables) is permitted.

You may use the link below to submit your revised manuscript and related files:

Link Redacted

Note: This URL links to your confidential home page and associated information about manuscripts you may have submitted, or that you are reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage.

If you wish to submit a suitably revised manuscript we would hope to receive it within 6 months. If you cannot send it within this time, please let us know. We will be happy to consider your revision so long as nothing similar has been accepted for publication at Nature Genetics or published elsewhere. Should your manuscript be substantially delayed without notifying us in advance and your article is eventually published, the received date would be that of the revised, not the original, version.

Please do not hesitate to contact me if you have any questions or would like to discuss the required revisions further.

Nature Genetics is committed to improving transparency in authorship. As part of our efforts in this direction, we are now requesting that all authors identified as 'corresponding author' on published papers create and link their Open Researcher and Contributor Identifier (ORCID) with their account on the Manuscript Tracking System (MTS), prior to acceptance. ORCID helps the scientific community achieve unambiguous attribution of all scholarly contributions. You can create and link your ORCID from the home page of the MTS by clicking on 'Modify my Springer Nature account'. For more information please visit <http://www.springernature.com/orcid>.

Thank you for the opportunity to review your work.

Sincerely,
Chiara

Chiara Anania, PhD
Associate Editor
Nature Genetics
<https://orcid.org/0000-0003-1549-4157>

Referee expertise:

Referee #1: pancreatic cancer, genomics

Referee #2: pancreatic cancer, genetics, evolution

Referee #3: pancreatic cancer, single-cell and spatial omics

Reviewers' Comments:

Reviewer #1 (Remarks to the Author):

In this highly impressive "atlas" generated by Kalfakakou and colleagues, the authors generate scRNA-seq and spatial transcriptomics on one of the largest series of primary pancreatic cancer samples (n = 62), which include a mixture of endoscopic biopsy samples and resected pancreatic cancers (both treatment naive and neoadjuvant treated). A few of the samples have matched pre and post treatment pairs, while some of the resected samples also have associated local lymph node metastases. The study establishes the remarkable intra-tumoral clonal and transcriptomic heterogeneity of primary pancreatic cancer, and in particular identifies how clonal genetic alterations contribute to both transcriptomic diversity

(including the plasticity of well established lineage states in pancreatic cancer), as well as orchestration of putative tumor microenvironmental interactions. The authors identify chromosome 8q24 amplification (housing Myc oncogene) as a potential driver of classical to basal lineage states in pancreatic cancer, associated with alterations in the immune surveillance program of cancer cells.

In spatial assessment performed on a more limited number of samples, the authors identify a particular "dispersed" clonal pattern that tends to be associated with propensity for nodal metastases. One important translational implication of this study is that uncovering clonal architectural diversity driving pertinent tumor-microenvironmental interactions (e.g., PVR-TIGIT, AXL-GAS6) might dictate therapeutic stratification.

The study is a tour de force and has generated a dataset that is likely to be of great value to the pancreatic cancer research community.

A few questions from this reviewer:

A: The entire premise of clonal heterogeneity and the correlation with transcriptional profiles and lineage states is based on inferring copy number aberrations using inferCNV based on RNA data. Given the absolute pivotal importance of documenting inferCNV findings as "ground truth" for clonal diversity, it is imperative that the data be validated with orthogonal approaches and other rigorous QC.

i: Do the key findings in the paper replicate if an alternative inferred copy number algorithm like CopyKAT is used? It would be important for the authors to validate at least some of their representative data and show that choice of CNV calling algorithm does not impact their conclusions.

ii: Please explicitly confirm that inferCNV or other copy number algorithms are not detecting CNVs (or at least significant recurrent CNVs) in non-neoplastic epithelial cells admixed within the cancer samples.

iii: The availability of matched samples on which spatial analysis is available is a unique strength of the paper but these samples have really not been leveraged adequately. While confirming inferred CNVs (especially Myc amplification) by an orthogonal approach would be technically not possible in the dissociated cells already used in scRNA-seq, the archival serial sections from the ST analyses can be absolutely used to validate key CNV findings in tissue by FISH (also see point B in this regard).

B: I have an analogous concern about the "definitive" vein in which CellphoneDB "inferred" ligand receptor interactions are stated throughout the paper. Please note that while presumptive, these are not functionally validated L-R interactions so I would moderate some of the definitive statements, especially when RNA-based scores are being ascribed to account for the strength of the interaction. The absence of any protein-based validation for presumptive spatially co-localized LRIs driven by clonal alterations like Myc or AXL or PVR amplification in the setting of CNVs is an important omission and one that would significantly reassure the reader that the conclusions are valid if performed in representative examples. I would once again suggest the possibility of leveraging the panel of cases for which ST data is available and using multiplex IF or cyclicIF in conjunction with FISH to demonstrate the link between CNVs, lineage states, and spatially co-localized L-R interactions between cognate cell types (e.g., tumour-T cells or tumour-myeloid cells).

C: The changes in terms of cancer cell and TME interactions in the setting of neoadjuvant chemotherapy are interesting. Are there also changes in the proportion of (or spatial distribution of, when the 11 ST samples are considered) basal/classical/mixed lineage cells following chemotherapy? Essentially what is the phenotype of the post-neoadjuvant therapy "persister" cells in terms of their lineage?

In the neoadjuvant treated ST samples, are there changes in spatial distribution of myCAF and iCAF populations (beyond quantitative alterations shown in Extended Fig 4) that would be pertinent? I may sound like a broken record here but the 11 ST samples are a unique and valuable resource and I would encourage the authors to leverage them as much as possible to "lock down" their conclusions.

D: In the Myc perturbation experiments, are canonical "classical" transcription factors like GATA6 or HNF4a impacted when Myc is knocked down?

E: The dispersed clonal data in Figure 7 is very compelling. Could the authors please explain the apparent paradox that these cells are still predominantly classical and yet harbor Myc amplification? Also while these are by definition localized surgically resected tumors, is there a correlation between presence of dispersed clones in the primary and either timing or patterns of subsequent recurrence?

Reviewer #2 (Remarks to the Author):

In this manuscript, Kalfakakou et al. sought to characterize the evolution of pancreatic adenocarcinoma (PDAC) by integrating genomic, transcriptional, and spatial dynamics. Towards this goal, the authors utilized single cell RNA sequencing of 62 primary tumor sections from 59 patients yielding 443,451 single cells for analyses, in combination with spatial transcriptomics of 13 samples from 11 tumors yielding 53,236 spatial transcriptomic spots. The authors then analyzed these data for inter- and intra-tumoral heterogeneity, clone-specific interactions with cells of the tumor microenvironment (TME), treatment-induced effects, subtype transitions, spatial mapping, and clone dispersal.

The gaps being addressed in this study are defining the extent to which genomic and transcriptional dynamics correlate during PDAC progression, as well as identifying mechanisms driving subtype transitions and evolutionary dynamics across clones, with particular focus on the TME interactions, metastasis, and treatment. The authors hypothesize that distinct clones

within PDACs acquire unique genomics and transcriptional alterations that impact TME interactions, thus driving progression and metastasis.

The authors major claims are: 1) intratumoral clones are heterogeneous with respect to copy number variants (CNVs), transcriptional states, subtype profiles, and TME interactions; 2) the classical to basal transition within each tumor associates with MYC amplification and a depletion of immune related signaling; 3) transcriptional heterogeneity represents a continuum that is impacted by CNV evolution and spatial organization; 4) spatially dispersed clones are enriched for expression of genes related to the epithelial-to-mesenchymal transition and immune suppression; and 5) potential therapeutic interventions for PDAC evolution and metastasis would target TIGIT, PDPR-alpha, the GAS6-AXL axis, and PIGR signaling.

The positive aspects of this study are the large set of PDAC tumor tissues with matched single cell RNA-sequencing and spatial transcriptomics, the six pairs of pre- and post-treatment matched samples, and the inclusion of two primary tumors with one matched lymph node metastasis each. This sample set with genetic and transcriptomic results represents a wealth of data in PDAC, which remains a highly important tumor type to study given its aggressive nature. However, in my view the negative aspects of this study are that there were missed opportunities for statistical analysis and data integration, too much emphasis on a subset of specific cases, and that several of the analyses were performed manually despite algorithms being available. Further, the validation studies do not seem sufficient, as no human cell lines were used and the therapeutic interventions postulated by the authors were not directly tested. Despite the size of the sample set, the novelty of the results is unclear because the authors' major claims stated in the above paragraph have mostly been previously published, as pointed out by the authors in the manuscript or in my specific comments below.

From my perspective, most of my critiques could be addressed with substantial reanalysis of the current data, particularly utilizing quantitative and statistical approaches and replacing manual procedures with published tools. There are relatively few statistical analyses in the current version given the claims put forth by the authors (e.g. very few p-values or FDR values reported given the number of claims made). Although single PDAC tumors can be illustrative, I suggest re-writing the manuscript to include additional analyses across the sample set because much of the current manuscript's text and main figures seem to focus on only a subset of the tumors.

Critiques

Line 110: "the largest cohort studied to date" may need to be clarified. I realize the authors mean the largest cohort of PDAC samples studied using both scRNA-seq and spatial transcriptomics, which may be true. However, for example there are two recent Nature Genetics manuscripts (PMIDs 39227743 and 39294496) that utilized spatial transcriptomics in a similar or higher number of PDACs (n = 13 and 30, respectively). In addition, the latter cohort (PMID 392944960) analyzed a more diverse set of tissues and tumors, including liver metastases. The authors already reference one of these studies (PMID 39227743), but I suggest including the second manuscript as well and perhaps rephrasing or contextualizing the largest cohort statement. A recent study has compiled published data into larger re-analyses of scRNA-seq data consisting of >200 PDAC samples (e.g. PMID 39636224).

Lines 123 - 124 (Figure 1a): Why are >25% of the samples negative for a KRAS mutation? And >20% have no identified genetic driver? These null results impact both biopsies and surgical resections. I assume the lack of driver mutation detection is due to low cellularity in PDAC which does occur often. Yet many of these same tumors indicate >25% epithelial cells in Extended Figure 1k, which does not add up unless many of the epithelial cells are not tumor cells?

Lines 145 - 149 (Extended Figure 1c): There are epithelial cells that seem to form patient specific clusters, but there are also many to overlap across samples in 1c in the central part of the largest blue cluster in 1b.

Lines 171 - 172: Why were the phylogenies inferred manually (and as described in lines 766 - 785)? It is not clear how clusters were identified as distinct or how strict this analysis was, and it seems potentially error-prone because scRNA-seq data presents many challenges for phylogenetic analyses. There are also a number of algorithms to minimize or avoid these documented issues. For example, Liu et al. (PMID 37553438) have recently published a summary of these approaches (see their Introduction) with a novel tool for inferring phylogenies from sc-RNA seq data. The authors should consider using this tool or a similar one to validate their manually derived results. It is also unclear how ploidy or whole genome duplication was handled.

Lines 177 - 219 (Figure 2): Why were these two cases (P323 and P432) selected for these analyses? These two cases remain prominent throughout the manuscript (e.g., Figure 3). Seems like a missed opportunity to extend these analyses to other cases. What is the evidence that "diverse clonal populations contribute to a diversity of transcriptomic signatures"? There is no quantitative analysis in this section that actually supports that statement, instead there are observations on the UMAPs when "cells within a single clone formed distinct clusters, while in others, cells from multiple clones clustered together." For example, it is not clear that there is a relationship between transcriptomic signatures and the number of CNV clones or their genetic divergence from one another (in fact Figure 4 later shows that some cases have two CNV clones and two transcriptomic signatures, whereas others have as many as 8 clones all sharing the same transcriptomic signature).

Lines 246 - 253 (Figure 3b): Were there no KRAS amplifications (PMID 31932696), CDKN2A deletions, or SMAD4 deletions identified? "Alternations" is supposed to be Alterations.

Lines 297 - 306 (Extended Figure 2b): How were low, intermediate, and high defined? On the left side of the heatmap there

are 11 samples (P567.2 through P480.2) that exhibit predominantly “High” interactions with the TME. About one third of these cases (P518, P643, P385, and P480.2) likely have very low cellularity according to the sequencing results. Is it possible that lower cellularity may indicate High TME interactions?

Line 361: This section mentions six cases were analyzed but much of the text and much of Extended Data Figure 3 focuses on one case (P466). This is similar to the critique above that in some sections of the manuscript there is too much emphasis on relatively few cases given the size of the cohort.

Lines 419 - 421: Using the word significantly here requires a sufficiently low p-value or other statistical test.

Lines 426 - 429: This section requires more explanation, it is unclear how this analysis was done.

Lines 464 - 473: Can the authors provide evidence for successful knock down of MYC (also mentioned in lines 823 - 824)? Is it possible to validate this in additional cell lines, preferably human and not just mouse?

Line 489: This is a bit confusing, can the authors provide the correlation value? It appears that BayesPrism annotated the far right as Epithelial cells in 4c and then manually annotated clusters in 4d found these to be CD8+ T cells, for example. Panels 4c and 4d look substantially different, maybe the colors are off?

Line 499-500: Can the authors provide the concordance value? There appear to be several clearly discrepant CNVs in clone 1 across the two datasets, e.g. a deletion in chr. 1 becomes an amplification, an amplification appears in chr. 8q (MYC?), and then an amplification appears in chr. 19. In fact clone 1 in the spatial data resembles clones 2 and 3 as opposed to in the scRNA-seq data.

Line 525: To what extent does dispersal correlate with clone size?

Line 561 (Discussion): The authors need to emphasize what are the unique findings of this study as opposed to confirmatory findings of past studies (including their own). Many of the key findings have been previously published, it is not clear what is novel here. Even a “key finding of this study” (line 594) regarding MYC as potentially a driver of classical to basal transitions has been previously published (PMID 35118421).

Line 634: “in our study” has been written twice.

Line 1117: Extended Figure 4 is missing a title.

Reviewer #3 (Remarks to the Author):

1. Overview

In this manuscript, Kalfakakou and colleagues profiled a cohort of primary human PDAC samples (n = 62) with single-cell RNA-sequencing and spatial transcriptomics. In a limited number of cases, they performed low-passage whole genome sequencing; had paired pre- and post-chemotherapy specimens (n = 6); or had matched lymph node metastases (n = 2). Using inferCNV, they identified intra- and inter-tumoral heterogeneity in clones that were associated with distinct transcriptomic subtypes (e.g., classical, basal, mixed) and interactions with other cell types in the microenvironment. They identified MYC amplification as an important driver of classical-to-basal transition and found that dispersed clones were enriched in the classical-to-basal transition state and associated with greater metastatic proclivity. Overall, the results of this study suggest that phenotypic heterogeneity in PDAC is driven by distinct genetic events, as demonstrated by amplification or deletion of loci harboring key genes associated with malignant cell states.

2. Major Comments

- Lines 150-152: The authors note that there was no significant differences in cell type composition between treated and untreated samples except a slight increase in the endothelial cell population. This is surprising given that chemotherapy is expected to at least reduce the cancer cell abundance. It would be helpful for the authors comment on this further and to perform this analysis on the paired cohort (n = 6) to see if any changes in cell type abundance are observed in that subset.
- Lines 229-231: If the core biopsies represent a more limited sampling of the tumor compared to the resections, the lack of difference in number of clones would suggest that clones are highly interspersed in space. The authors should comment on this further in the Discussion.
- Lines 239-241: While the authors state their intent to focus on the Moffitt basal and classical signatures, here they mention this small subset of cells that tested negative for both basal and classical signatures. These should be explored further given that several other malignant cell signatures in PDAC have been identified in the literature.
- Lines 267-268: It is surprising that treatment did not affect enriched pathways. How long after treatment were specimens collected? Did the authors do a subset analysis of the 6 paired pre- and post-treatment specimens?
- Lines 306-309: The LRI analysis performed is based on CellPhoneDB was applied to scRNA-seq data and therefore only based on gene expression data. For the pan-low and pan-high interacting tumors: was there any correlation to underlying sample/data quality of the scRNA-seq?
- Lines 325-326: This sentence is an overstatement of what their data can infer regarding LR interactions.
- Lines 467-471: All of the human results were anchored on Moffitt basal and classical but the KMCE mouse PDAC cell lines

do not appear to be enriched in Moffitt basal so the authors have chosen here to expand to other signatures. There appears to be some bias in how certain subtype schemas are being chosen for different aspects of the paper. The authors should be more consistent in the scope of prior malignant signatures applied throughout the study. Moreover, this reviewer would suggest that the authors repeat the KMCE PDAC Myc knockdown experiments in human cell lines that are enriched for Moffitt basal at baseline.

- Lines 491-492: The enrichment of iCAFs over myCAFs in chemotherapy-treated samples has been seen in prior studies such as PMID: 39227743 and should be cited.
- Lines 497-500: A clear advantage of Visium spatial data is the sequence information it provides rather than probe-based quantification. The authors are able to identify some of the same clones in the spatial data that they observed in the scRNA-seq but they stop short of doing LRI on the spatial data. This is a significant missed opportunity.
- Line 530: This is an overstatement. A simple association in two cases without additional validation does not indicate that dispersed clones have a higher propensity to metastasize.
- Lines 541-559: Why wouldn't a fully basal, EMT state be even more aggressive than a classical state with some indication of transition towards a basal state? Does the hybrid nature confer a specific advantage? Is this related to the partial EMT phenotype described previously (PMID: 29920274)? This should be explored and discussed more.
- Lines 586-589: Increased interactions between iCAFs and cancer cells after treatment was previously described and should be cited (PMID: 39227743).
- Lines 594-606: The link between Myc amplification, interaction with TAMs, and metastatic progression in PDAC was previously shown and should be cited in this discussion (PMID: 34551968).

3. Minor Comments

- Throughout the manuscript, the authors refer to individual tumors as examples. This reviewer was unable to find a table or list in the manuscript that includes the treatment status of each tumor, which is important for interpretation of the results.
- While there were 62 primary PDAC tumor samples from 59 patients, only 24 specimens were obtained from surgical resection and the rest were obtained by endoscopic biopsy with limited tissue. This is a limitation that should be discussed given the potential for significant sampling bias.
- Line 210: "coexistence of both classical (C01) and basal (C03, C06)" => should C06 actually be C05?
- Line 216: "Furthermore, we observed the concurrent presence of classical and basal subtypes in many tumors, challenging the traditional binary classification of PDAC." => there are many studies in the literature already demonstrating this, which should be cited as this is not a new finding
- Line 503: Was "clone spC03" confirmed by a pathologist as malignant/invasive for sure? Histologically, it looks like it could represent a preinvasive neoplastic lesion.
- Lines 506-507: The parental clone is classified as "neither" so this statement should be corrected.
- Figure 2c-e: Using the same blue color for C01 (which is mixed) as classical is confusing; similarly using the same tan color for C04 (which is classical) as mixed is confusing.
- Figure 3c: "Columns are clusters and rows are different pathways." This appears to be flipped.
- Figure 4: This seems more suitable as a supplementary figure.
- Figure 6a: Luminal cellular debris was analyzed as "tumor" in the large gland involving spC03. How was this handled analytically?

Version 1:

Decision Letter:

Our ref: NG-A68080R

17th Feb 2026

Dear Dr. Simeone,

Thank you for submitting your revised manuscript "Clonal Heterogeneity in Human Pancreatic Ductal Adenocarcinoma and Its Impact on Tumor Progression" (NG-A68080R). It has now been seen by the original referees and their comments are below. The reviewers find that the paper has improved in revision, and therefore we'll be happy in principle to publish it in Nature Genetics, pending minor revisions to satisfy the referees' final requests and to comply with our editorial and formatting guidelines.

If the current version of your manuscript is in a PDF format, please email us a copy of the file in an editable format (Microsoft Word or LaTeX)-- we can not proceed with PDFs at this stage.

We are now performing detailed checks on your paper and will send you a checklist detailing our editorial and formatting requirements soon. Please do not upload the final materials and make any revisions until you receive this additional information from us.

Thank you again for your interest in Nature Genetics Please do not hesitate to contact me if you have any questions.

Sincerely,

Margot Brandt, PhD
Senior Editor

Reviewer #2 (Remarks to the Author):

In my opinion the authors have thoroughly responded to my critiques as well as those of the other reviewers. The manuscript is substantially improved with new analyses, experimental validation, and methods. The authors were forthright in admitting limitations of their data, and demonstrated a strong effort to address our critiques and revise the manuscript as warranted. My recommendation is to accept the manuscript for publication.

Reviewer #3 (Remarks to the Author):

The authors have adequately addressed my critiques and suggestions.

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Dear Dr Anania,

We thank you and the reviewers for your thoughtful and constructive feedback on our manuscript, *"Clonal Heterogeneity in Human Pancreatic Ductal Adenocarcinoma and Its Impact on Tumor Progression."* We appreciate the thorough evaluation and have carefully and comprehensively addressed each of the reviewers' comments. To ensure that all the concerns raised by the reviewers are thoroughly addressed, we upgraded and streamlined our computational workflows and performed a full-scale re-analysis of our data. We also conducted additional experiments, where requested. Below, we provide a detailed, point-by-point response outlining how we have addressed the reviewers' points through new analyses, additional experiments, and extensive manuscript revisions. We believe these efforts have substantially strengthened our study, and we are grateful for the insights that helped us enhance its impact. We hope the reviewers will now find this revised version suitable for publication in *Nature Genetics*.

Comments from Reviewer #1:

In this highly impressive "atlas" generated by Kalfakakou and colleagues, the authors generate scRNA-seq and spatial transcriptomics on one of the largest series of primary pancreatic cancer samples (n = 62), which include a mixture of endoscopic biopsy samples and resected pancreatic cancers (both treatment naive and neoadjuvant treated). A few of the samples have matched pre and post treatment pairs, while some of the resected samples also have associated local lymph node metastases. The study establishes the remarkable intra-tumoral clonal and transcriptomic heterogeneity of primary pancreatic cancer, and in particular identifies how clonal genetic alterations contribute to both transcriptomic diversity (including the plasticity of well-established lineage states in pancreatic cancer), as well as orchestration of putative tumor microenvironmental interactions. The authors identify chromosome 8q24 amplification (housing Myc oncogene) as a potential driver of classical to basal lineage states in pancreatic cancer, associated with alterations in the immune surveillance program of cancer cells. In spatial assessment performed on a more limited number of samples, the authors identify a particular "dispersed" clonal pattern that tends to be associated with propensity for nodal metastases. One important translational implication of this study is that uncovering clonal architectural diversity driving pertinent tumor-microenvironmental interactions (e.g., PVR-TIGIT, AXL-GAS6) might dictate therapeutic stratification.

The study is a tour de force and has generated a dataset that is likely to be of great value to the pancreatic cancer research community.

We thank the reviewer for their kind words about our study and for recognizing the high level of value to the pancreatic cancer research community. We have carefully addressed each of their comments below.

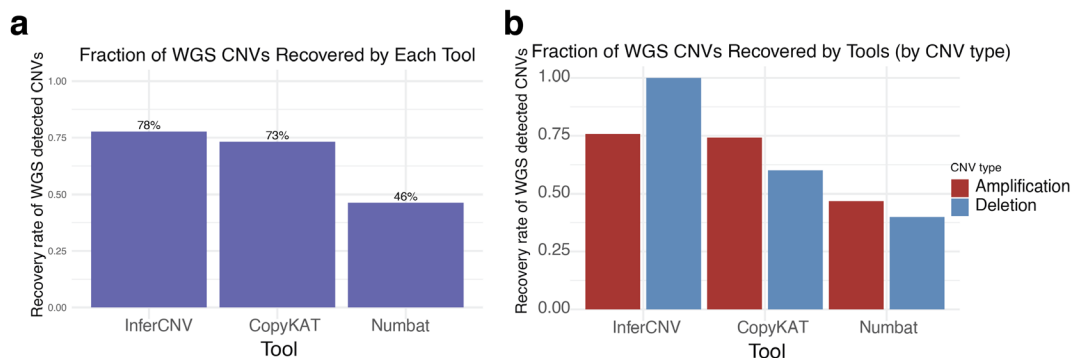
Comments from Reviewer 1:

A) The premise of clonal heterogeneity and the correlation with transcriptional profiles and lineage states is based on inferring copy number aberrations using inferCNV based on RNA data. Given the pivotal importance of documenting inferCNV findings as "ground truth" for clonal diversity, it is imperative that the data be validated with orthogonal approaches and other rigorous QC.

i: Do the key findings in the paper replicate if an alternative inferred copy number algorithm like CopyKAT is used? It would be important for the authors to validate at least some of their representative data and show that choice of CNV calling algorithm does not impact their conclusions.

a) To address this comment, we leveraged the 15 samples for which low-pass whole-genome sequencing (WGS) profiles were available and used as ground truth. We performed a direct, segment-level comparison by matching pseudobulk CNVs inferred from three tools, namely inferCNV, CopyKAT, and Numbat, to WGS-derived CNVs based on genomic interval overlap. Because CNV prediction from single-cell RNA-seq is limited to identifying large-scale chromosomal events, we filtered out WGS CNVs smaller than 500 kb (the smallest CNV detected by any of the three methods) to ensure a fair and comparable evaluation across tools.

This analysis demonstrated an advantage for inferCNV, which retrieved 78% of WGS-called CNVs, compared to 73% for CopyKAT and 46% for Numbat (Rebuttal Fig. 1a). The difference was even more pronounced when deletions and amplifications were considered separately: inferCNV retrieved 100% of WGS-called deletions, while CopyKAT and Numbat retrieved only 58% and 40%, respectively (Rebuttal Fig. 1b). These results are now included in Ext. Fig. 3 of the manuscript.



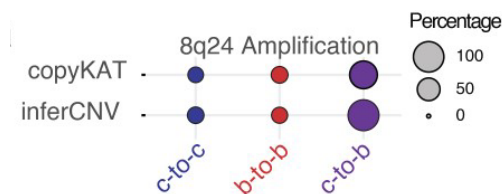
Rebuttal Figure 1: a. Fraction of WGS-detected CNV recovered by the three tools tested. b. Recovery rate of WGS-detected CNVs by tool, per CNV type.

b) To directly address the reviewer's concern and to ensure that our conclusions do not depend on the specific algorithm used, we performed the key biological analyses

using inferCNV and CopyKAT. Given the substantially lower recovery rate of WGS-derived CNVs by Numbat in our benchmarking, we limited the downstream analyses to inferCNV and CopyKA

1. 8q24 amplification in the classical-to-basal transition

We examined the role of 8q24 amplification in the transition from classical to basal phenotypes using both inferCNV and CopyKAT. With CopyKAT, 8q24 remained the most common amplification, being detected in 4 out of 5 classical-to-basal transitions, compared with 5 out of 5 transitions using inferCNV (Rebuttal Fig. 2 – also included in Suppl. Fig. 8).



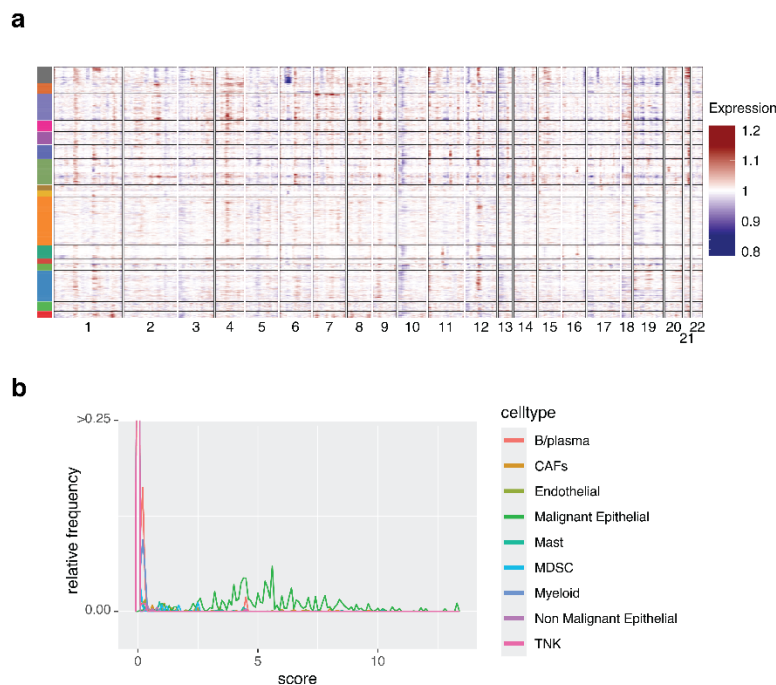
Rebuttal Figure 2: Frequency of 8q24 amplification in classical-to-basal transitions detected by inferCNV and CopyKAT.

2. Concordance of LRI-based translational findings across tools

We compared analyses based on CNVs called by inferCNV and CopyKAT to confirm specific key translational findings, including those uncovering clonal architectural diversity driving relevant tumor–microenvironment interactions with potential therapeutic implications, and found these results were independent of the CNV-calling method. CopyKAT captured the same dominant enrichment and suppression features revealed by inferCNV, reinforcing the robustness of the inferred CNV profiles. The subtle differences in dynamic range -particularly for deletion-associated suppressed LRIs- are consistent with CopyKAT identifying fewer deletions, as we previously showed. Importantly, both tools converged on the major interaction patterns: LR pairs associated with amplifications or deletions consistently exhibited high enrichment or suppression scores, respectively (Rebuttal Fig. 3 – also included in Suppl. Fig. 6). Correlation analysis further confirmed

ii: Please explicitly confirm that inferCNV or other copy number algorithms are not detecting CNVs (or at least significant recurrent CNVs) in non-neoplastic epithelial cells admixed within the cancer samples.

Thank you for this important point. We now refer to these cells as “non-malignant epithelial” rather than “normal,” as we agree that the term “normal” may be misleading in the context of tumor-adjacent tissue, as they may be influenced by fibrotic changes that occur in the pancreas with ductal obstruction. We classified all cells by calculating a CNV score (i.e., the percentage of the genome affected by CNVs). Non-malignant epithelial cells show CNV scores that are comparable to non-epithelial (immune and stromal) cells and distinct from malignant epithelial cells. Moreover, they constitute only ~5% of all the epithelial cells examined, as the study focus was on tumor tissue. We now include a heatmap of non-epithelial cells (Rebuttal Fig. 4a) and a plot showing the distribution of CNV scores across all the cells of all the samples (Rebuttal Fig. 4b), which shows that malignant epithelial cells consistently have a significantly higher score than all other cell types.



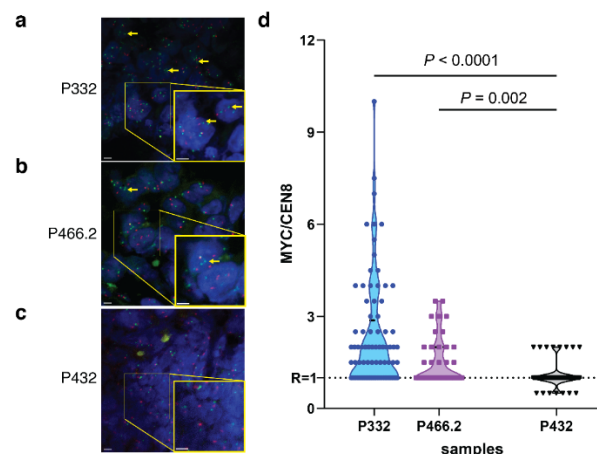
Rebuttal Figure 4: a, Heatmap of inferCNV computed normalized expression of normal epithelial cells. b, Frequency of CNV scores for all the cells of all the samples across cell types.

We do observe infrequent CNV signals in the non-malignant epithelial cells, and some of these signals are recurrent across patients. However, inferCNV does not classify this signal as a true CNV under our thresholding criteria (though we cannot rule out false negatives). Based on the available data, we cannot be certain of their true identity and suspect these may represent small PanIN pre-malignant lesions. It is possible that some

of these cells are indeed genomically normal, while others may represent pre-malignant or genomically altered but non-transforming epithelial states. In light of this ambiguity, and given that this population is small, we have excluded them from our formal analyses. Of note, several studies have shown that somatic CNVs can be detected in pre-malignant tissue, or even in histologically normal epithelial tissues (e.g., Zhou et al., *Nat Genet* 2022; Brunner et al., *Nature* 2019), supporting the idea that low-level CNVs in such cells are not unexpected and not necessarily indicative of malignancy.

iii: The availability of matched samples on which spatial analysis is available is a unique strength of the paper but these samples have really not been leveraged adequately. While confirming inferred CNVs (especially Myc amplification) by an orthogonal approach would be technically not possible in the dissociated cells already used in scRNA-seq, the archival serial sections from the ST analyses can be absolutely used to validate key CNV findings in tissue by FISH (also see point B in this regard).

We thank the reviewer for this excellent suggestion. To validate key CNV findings directly in tissue, we performed FISH for MYC amplification on matched fresh frozen tumor sections from the same samples used in our single cell and spatial transcriptomics analyses. These analyses confirmed the presence of MYC amplification in two samples (P332, P466.2), which were identified as MYC-amplified in our single-cell and spatial analyses, using the chromosome 8 centromere as an internal control. In contrast, no amplification was observed in sample P432 that was not MYC-amplified in our dataset, supporting the specificity of the finding (Rebuttal Fig. 5). The results are now included in the revised manuscript (Ext. Fig. 3).



Rebuttal Figure 5: DNA FISH of MYC (8q24.2). a-c, Representative fluorescence images and d, quantifications for MYC copy numbers in MYC-amplified samples P332, P466.2, and negative control P432. a-c, MYC (green) and control (red) FISH probes targeting MYC and the centromere of chromosome 8; yellow arrows point to nuclei showing amplified MYC. Scale bar, 5 μm. Inserts, magnified view of regions shown in the square Field of View (FOV). d, Dotted and violin plot of per-nuclei copy

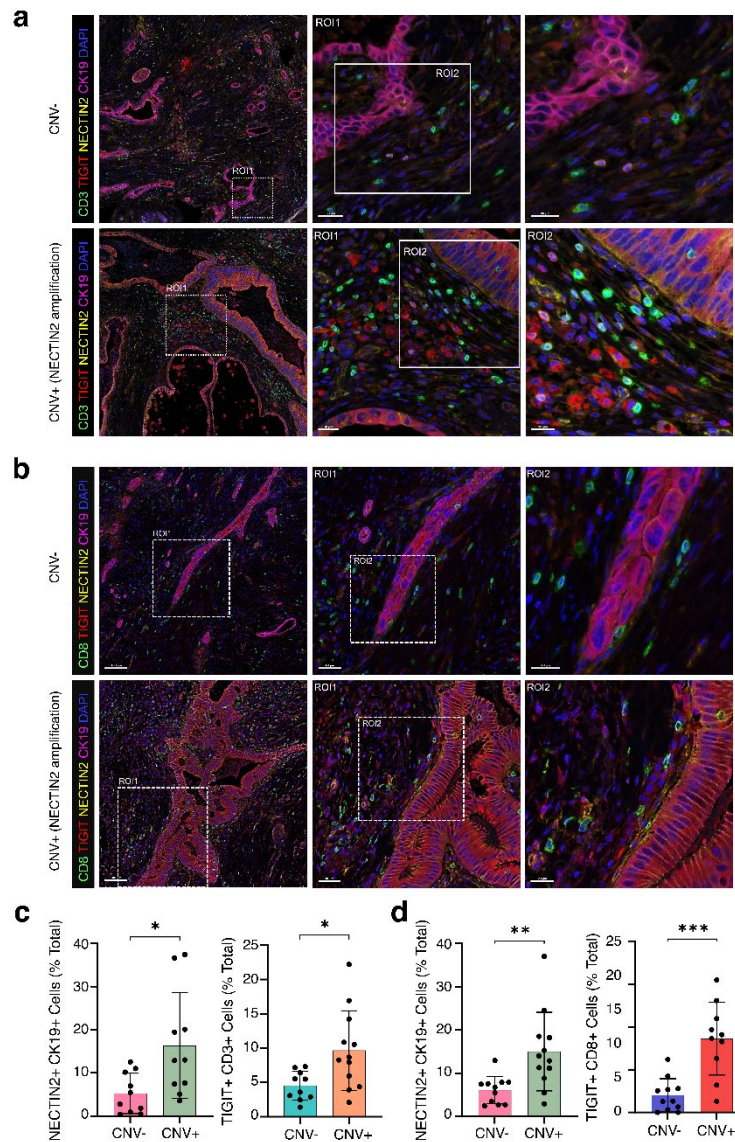
number ratio for MYC in each sample. Statistics calculated by one-way ANOVA followed by multiple t-test (two-sided).

B): I have an analogous concern about the "definitive" vein in which CellphoneDB "inferred" ligand receptor interactions are stated throughout the paper. Please note that while presumptive, these are not functionally validated L-R interactions so I would moderate some of the definitive statements, especially when RNA-based scores are being ascribed to account for the strength of the interaction. The absence of any protein-based validation for presumptive spatially co-localized LRIs driven by clonal alterations like Myc or AXL or PVR amplification in the setting of CNVs is an important omission and one that would significantly reassure the reader that the conclusions are valid if performed in representative examples. I would suggest the possibility of leveraging the panel of cases for which ST data is available and using multiplex IF or cyclicIF in conjunction with FISH to demonstrate the link between CNVs, lineage states, and spatially co-localized L-R interactions between cognate cell types (e.g., tumour-T cells or tumour-myeloid cells).

We acknowledge the need for cautious interpretation of LRI predictions and have revised the manuscript to describe LRIs using terms such as “suggested” or “putative.” To experimentally assess a key interaction, NECTIN2–TIGIT, we attempted FISH and multiplex IF on archival ST sections. Because no validated commercial probes for the NECTIN2 locus exist, we designed custom FISH probes; however, these probes failed to generate interpretable signal, and the FISH experiment could not be used for validation.

To evaluate the predicted CNV–LRI relationship, we therefore focused on multiplex IF. Guided by our single-cell CNV profiles, we selected two tumors: one carrying the 19p13 amplification encompassing NECTIN2 (P466.2) and one without (P432). As expected from the scRNA-seq–derived CNV calls, the amplified tumor showed markedly higher NECTIN2 expression and increased TIGIT signal. Importantly, mIF revealed clear spatial co-localization of NECTIN2⁺ epithelial cells with TIGIT⁺ CD3⁺ T cells and even more strongly with TIGIT⁺ CD8⁺ T cells, both of which were significantly enriched in the CNV⁺ tumor. These orthogonal results support CNV-associated enhancement of NECTIN2–

TIGIT signaling in the tumor harboring the predicted NECTIN2 amplification (Rebuttal Fig. 6, also included in Ext. Fig. 5).



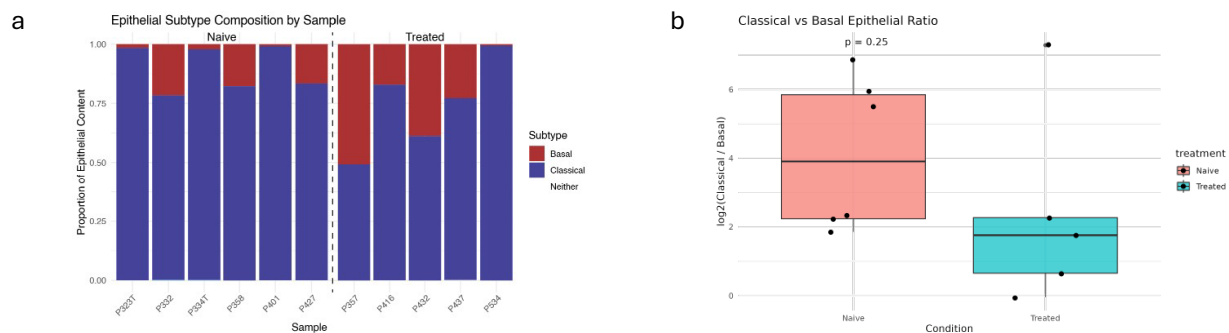
Rebuttal Figure 6: a-b, Representative multiplex immunofluorescence images of pancreatic ductal adenocarcinoma (PDAC) tissue sections with (CNV⁺; P466.2) or without (CNV⁻; P432) a NECTIN2 copy-number variation. a, Co-localization of NECTIN2⁺ CDK19⁺ cells with TIGIT⁺ CD3⁺ cells. b, Co-localization of NECTIN2⁺ CDK19⁺ cells with TIGIT⁺ CD8⁺ cells. Sections were stained for CK19 (magenta), NECTIN-2 (yellow), TIGIT (red), and CD3 or CD8 (green), with nuclei counterstained using DAPI (blue). Scale bars: 100 μ m (overview), 30 μ m (ROI1), and 20 μ m (ROI2). c, Quantification of the percentage of CK19⁺ cells co-expressing NECTIN2 (left) and CD3⁺ cells co-expressing TIGIT (right) in PDAC tissue sections shown in a. Data are presented as mean $\pm$ SD (* p < 0.05, two-tailed t-test). d, Quantification of the percentage of CK19⁺ cells co-expressing NECTIN2 (left) and CD8⁺ cells co-expressing TIGIT (right) in PDAC tissue sections shown in b. Data are presented as mean $\pm$ SD (* p < 0.05, two-tailed t-test).

C) The changes in terms of cancer cell and TME interactions in the setting of neoadjuvant chemotherapy are interesting. Are there also changes in the proportion of (or spatial distribution of, when the 11 ST samples are considered) basal/classical/mixed lineage cells following chemotherapy? Essentially what is the phenotype of the post-neoadjuvant therapy "persister" cells in terms of their lineage?

In the neoadjuvant treated ST samples, are there changes in spatial distribution of myCAF and iCAF populations (beyond quantitative alterations shown in Extended Fig 4) that would be pertinent? I may sound like a broken record here but the 11 ST samples are a unique and valuable resource and I would encourage the authors to leverage them as much as possible to "lock down" their conclusions.

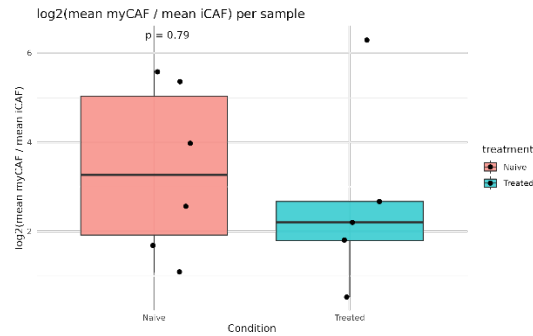
To further explore the reviewer's question about whether neoadjuvant chemotherapy affects lineage composition and the spatial organization of fibroblast subsets in our spatial transcriptomics cohort, we performed additional analyses comparing treated and untreated samples. Specifically, we assessed changes in the proportions and spatial distribution of basal, classical, and mixed lineage clones, as well as myCAF and iCAF populations, following chemotherapy.

Subtype analysis revealed a modest increase in the basal subtype in the treated spatial transcriptomics samples; however, the sample size is limited and does not provide sufficient statistical power to draw definitive conclusions (Rebuttal Fig. 7a,b).



Rebuttal Figure 7: a, Subtype composition per sample and condition (naïve vs treated) in the spatial samples. b, Ratio of spots presenting classical/basal subtype in naïve and treated spatial samples.

The analysis of the distribution of iCAFs and myCAFs in the tissue didn't reveal any significant differences either, as shown in Rebuttal Fig. 8.

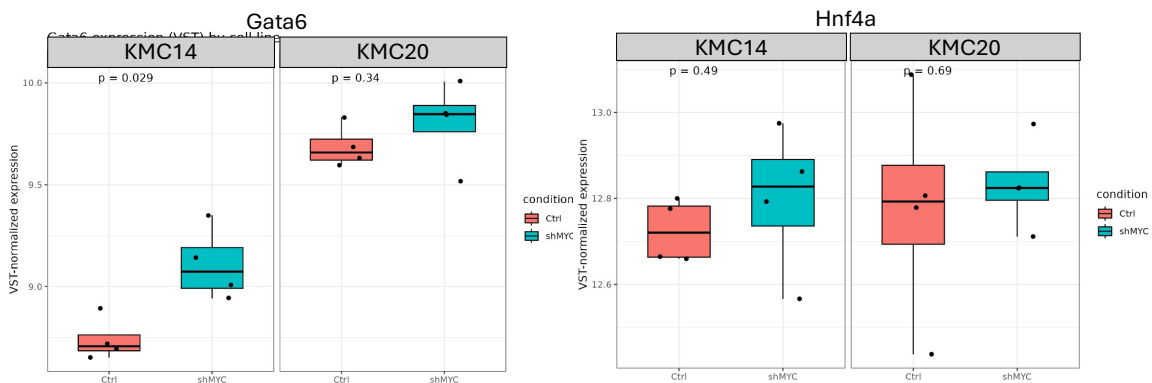


Rebuttal Figure 8: Ratio of myCAF/iCAF in spots in naïve and treated spatial samples.

D) In the Myc perturbation experiments, are canonical "classical" transcription factors like *GATA6* or *HNF4a* impacted when Myc is knocked down?

In our murine cell line MYC perturbation experiments, *Gata6* expression increased in one of the two mouse cell line clones, while *Hnf4a* expression remained unchanged (Rebuttal Fig. 9). Because gene-level responses can vary between experimental models, our original analysis focused on subtype signature changes as a more integrative measure of pathway-level alterations.

As part of the revision, we have now performed equivalent MYC knockdown experiments in human PDAC cell lines using RT-PCR. In this human context, we observed an increase in classical subtype markers (*EPCAM*, *CDH1*, *MUC1*, and *FOXA2*) together with a decrease in basal markers (*TP63*, *KRT5*, *S100A2*, and *VIM*) (Rebuttal Fig. 10, also presented in Suppl. Fig. 7) after MYC KD.



Rebuttal Figure 9: *Gata6* and *Hnf4a* expression remain stable after successful Myc knockdown in two murine cell line clones.

E) The dispersed clonal data in Figure 7 is very compelling. Could the authors please explain the apparent paradox that these cells are still predominantly classical and yet harbor Myc amplification? Also while these are by definition localized surgically resected tumors, is there a correlation between presence of dispersed clones in the primary and either timing or patterns of subsequent recurrence?

The observation that dispersed clones remain predominantly classical despite harboring MYC amplification reflects our finding that MYC may be required but not sufficient for the classical-to-basal transition and we believe that these cells are probably in a transition state. Several lines of evidence from our data indicate that these cells occupy a transitional or intermediate state. First, dispersed clones show upregulation of basal-associated markers (LY6D, KRT6A, KRT15, FGFBP1) and EMT-related genes (FN1, S100A4, S100A5) despite their overall classical identity (Fig. 8i; Suppl. Fig. 8i). Second, these clones exhibit depletion of immune-related pathways (Fig. 8j,l,m), consistent with our earlier observation that immune evasion accompanies the transition toward a basal phenotype. Third, enrichment of EMT and hypoxia pathways in dispersed clones (Fig. 8j,k) supports enhanced invasive and adaptive capacity, features characteristic of basal-like states. This hypothesis is discussed in the manuscript and aligns with our broader conclusion that MYC amplification, along with other signaling alterations such as suppression of immune signaling pathways, may contribute to lineage plasticity in PDAC.

We also evaluated whether dispersed clones in the primary tumor were associated with recurrence patterns. One patient was not followed after surgery due to transfer of care. Among the remaining ten patients, three experienced documented recurrences: two with dispersed clones (at 5 and 9 months) and one without dispersed clones (at 17 months). Several additional deaths occurred without a recorded recurrence, all in patients whose tumors contained dispersed clones (deaths at 16, 18, 25, and 29 months). Thus, across the cohort, tumors with dispersed clones consistently showed early disease progression, whether captured as a recurrence or as death without documented recurrence. In contrast, tumors without dispersed clones showed either later recurrence (17 months) or longer-term survival (alive at 28, 43, and 47 months). Although the cohort size is too small for formal statistical testing, the observed pattern is consistent with the molecular differences we identified between dispersed and non-dispersed clones.

Comments from Reviewer #2:

In this manuscript, Kalfakakou et al. sought to characterize the evolution of pancreatic adenocarcinoma (PDAC) by integrating genomic, transcriptional, and spatial dynamics. Towards this goal, the authors utilized single cell RNA sequencing of 62 primary tumor sections from 59 patients yielding 443,451 single cells for analyses, in combination with spatial transcriptomics of 13 samples from 11 tumors yielding 53,236 spatial

transcriptomic spots. The authors then analyzed these data for inter- and intra-tumoral heterogeneity, clone-specific interactions with cells of the tumor microenvironment (TME), treatment-induced effects, subtype transitions, spatial mapping, and clone dispersal.

The gaps being addressed in this study are defining the extent to which genomic and transcriptional dynamics correlate during PDAC progression, as well as identifying mechanisms driving subtype transitions and evolutionary dynamics across clones, with particular focus on the TME interactions, metastasis, and treatment. The authors hypothesize that distinct clones within PDACs acquire unique genomics and transcriptional alterations that impact TME interactions, thus driving progression and metastasis.

The authors major claims are: 1) intratumoral clones are heterogeneous with respect to copy number variants (CNVs), transcriptional states, subtype profiles, and TME interactions; 2) the classical to basal transition within each tumor associates with MYC amplification and a depletion of immune related signaling; 3) transcriptional heterogeneity represents a continuum that is impacted by CNV evolution and spatial organization; 4) spatially dispersed clones are enriched for expression of genes related to the epithelial-to-mesenchymal transition and immune suppression; and 5) potential therapeutic interventions for PDAC evolution and metastasis would target TIGIT, PDFR-alpha, the GAS6-AXL axis, and PIGR signaling.

The positive aspects of this study are the large set of PDAC tumor tissues with matched single cell RNA-sequencing and spatial transcriptomics, the six pairs of pre- and post-treatment matched samples, and the inclusion of two primary tumors with one matched lymph node metastasis each. This sample set with genetic and transcriptomic results represents a wealth of data in PDAC, which remains a highly important tumor type to study given its aggressive nature. However, in my view the negative aspects of this study are that there were missed opportunities for statistical analysis and data integration, too much emphasis on a subset of specific cases, and that several of the analyses were performed manually despite algorithms being available. Further, the validation studies do not seem sufficient, as no human cell lines were used and the therapeutic interventions postulated by the authors were not directly tested. Despite the size of the sample set, the novelty of the results is unclear because the authors' major claims stated in the above paragraph have mostly been previously published, as pointed out by the authors in the manuscript or in my specific comments below.

From my perspective, most of my critiques could be addressed with substantial reanalysis of the current data, particularly utilizing quantitative and statistical approaches and replacing manual procedures with published tools. There are relatively few statistical analyses in the current version given the claims put forth by the authors (e.g. very few p-

values or FDR values reported given the number of claims made). Although single PDAC tumors can be illustrative, I suggest re-writing the manuscript to include additional analyses across the sample set because much of the current manuscript's text and main figures seem to focus on only a subset of the tumors.

We thank the reviewer for their thoughtful and constructive feedback, as well as for recognizing the value of our dataset and the integration of single-cell RNA sequencing and spatial transcriptomics across a large PDAC cohort. We appreciate their insightful comments and have addressed each point in detail below. In response, we reanalyzed all data de novo using fully automated methods, added statistical evaluations that were previously missing, and incorporated new analyses and experiments where needed to strengthen the manuscript. We believe these revisions directly address the reviewer's concerns and substantially improve the clarity and rigor of the study.

1) There are missed opportunities for statistical analysis and data integration, too much emphasis on a subset of specific cases, and that several of the analyses were performed manually despite algorithms being available. Most of my critiques could be addressed with substantial reanalysis of the current data, particularly utilizing quantitative and statistical approaches and replacing manual procedures with published tools. There are relatively few statistical analyses in the current version given the claims put forth by the authors (e.g. very few p-values or FDR values reported given the number of claims made). Although single PDAC tumors can be illustrative, I suggest re-writing the manuscript to include additional analyses across the sample set because much of the current manuscript's text and main figures seem to focus on only a subset of the tumors.

We thank the reviewer for this comment. In our original submission, we focused on illustrative examples primarily to explain our analytical approach. We would like to note that several global analyses were already included. For example, in Figure 4, we summarized the number of epithelial clones across samples, identified recurrently altered loci across clones, and performed pathway analysis at the cohort level; we also quantified CNV differences between classical and basal clones, shown in Suppl. Figure 4. Similarly, in Figure 5, we analyzed clone–TME interaction heterogeneity across the full dataset and systematically assessed how CNVs in ligand/receptor genes shape these interactions across samples.

That said, we agree that a more comprehensive assessment would strengthen our manuscript, and in particular that statistical evaluation was missing. In response to Reviewer #2's suggestions, we have incorporated quantitative and statistical evaluations throughout, we provide additional global insights, and we present individual examples more selectively to explain our approach. In the next comments we present a new pathway-level assessment to quantitatively evaluate clonal–transcriptomic relationships,

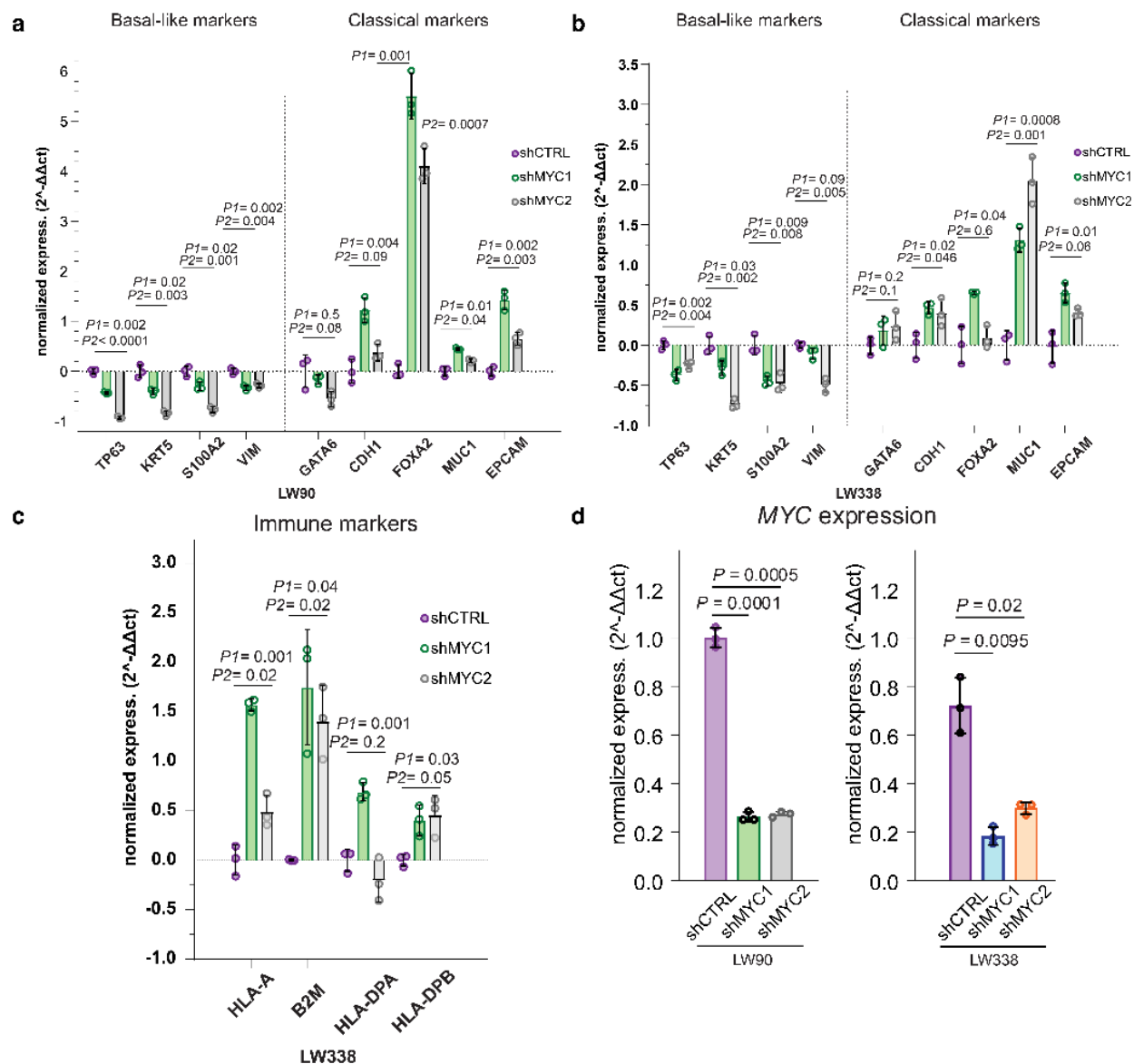
an expansion of our global analyses of the matched samples, and validation of our LRI findings in the spatial cohort. We have also incorporated statistical tests across analyses and FDR-adjusted values to support the robustness of our results. Most importantly, we replaced manual approaches with established computational and statistical methods. All the additional analyses and statistical methods are presented in detail in the following comments.

2) Validation studies do not seem sufficient, as no human cell lines were used and the therapeutic interventions postulated by the authors were not directly tested. Despite the size of the sample set, the novelty of the results is unclear because the authors' major claims stated in the above paragraph have mostly been previously published.

In response to the reviewer's concern, we have now performed additional experiments using primary human PDAC cell lines. In a basal-like human PDAC line harboring MYC amplification, shRNA-mediated MYC knockdown induced a transcriptional shift toward a classical state, consistent with our model that MYC amplification drives the transition from classical to basal phenotypes (Rebuttal Fig. 10 – also presented in Suppl. Figure 7 of the manuscript).

We also assessed immune-related genes in one of the two MYC-amplified lines (LW338). MYC knockdown led to upregulation of antigen-presentation markers (HLA-A, B2M, HLA-DPA, HLA-DPB), indicating that MYC activity contributes to the immune-depleted state characteristic of basal-like clones (Rebuttal Fig. 10c). We did not extend immune-gene analysis to the second primary human PDAC line (LW90) because key immune markers were expressed at extremely low baseline levels in the parental cells, making it uninformative for assessing MYC-dependent regulation.

These findings, together with our murine PDAC data and human scRNA-seq analyses, provide functional validation of the proposed MYC-dependent lineage switch in both mouse and human systems.



Rebuttal Figure 10: a, b, Quantification of basal-like markers (TP63, KRT5, S100A2, and VIM) and classical markers (CDH1, FOXA2, MUC1, and EPCAM) in LW90 and LW338 cells stably expressing either a scrambled shRNA control (purple), MYC-targeting shRNA1 (green), or shRNA2 (gray). c, Quantification of immune related markers (HLA-A, B2M, HLA-DPA, HLA-DPB) in LW338 cells stably expressing either a scrambled shRNA control (purple), MYC-targeting shRNA1 (green), or shRNA2 (gray). d, Quantification of MYC expression by RT-PCR in LW90 cells and LW338 cells stably expressing either a scrambled shRNA control (shCTRL), shRNA1, or shRNA2 targeting the coding region of MYC. Data represent three independent biological replicates and are presented as mean $\pm$ SD. Statistical significance was assessed using one-way ANOVA followed by two-sided multiple t-tests

Previous work that has linked MYC to basal-like/squamous PDAC (Hayashi et al, PMID 35118421) used multiregion bulk sequencing and copy-number analysis with FACETS to show that MYC amplification is frequently subclonal, associated with squamous features, and that in four out of five MYC-amplified cases the regions with squamous morphology fall within a lineage derived from the MYC-amplified population. Their phylogenetic

inferences were based on anatomically distinct tissue regions (different areas of the primary tumor and multiple matched metastases), with each region treated as a single evolutionary unit. In PDAC organoids, MYC overexpression could induce squamous marker expression in a subset of models with chromatin-modifier mutations but did not consistently alter morphology. These findings support a role for MYC in promoting squamous/basal programs, but they do not resolve clone-level transcriptional states within individual tumors or test whether MYC loss can reverse basal-like phenotypes in PDAC.

Our study resolves these lineage relationships at clone-level resolution within single primary tumors by integrating scRNA-seq-derived CNV profiles with transcriptional subtypes. This approach reveals that basal-like epithelial clones can arise from classical ancestors within the same tumor, in association with focal amplification of the MYC locus. This evolutionary transition cannot be detected from region-level bulk profiling. We further show that these MYC-amplified basal clones preferentially expand in immune-depleted microenvironments. Finally, using both murine and primary human PDAC models with endogenous MYC amplification, we demonstrate that MYC knockdown shifts gene expression toward a classical program and reactivates immune pathways, providing direct loss-of-function evidence that MYC activity is required to maintain the basal state. Together, these findings extend prior region-level observations to a mechanistic, clone-resolved model of MYC-driven classical-to-basal transitions in PDAC.

3) Line 110: “the largest cohort studied to date” may need to be clarified. I realize the authors mean the largest cohort of PDAC samples studied using both scRNA-seq and spatial transcriptomics, which may be true. However, there are two recent Nature Genetics manuscripts (PMIDs 39227743 and 39294496) that utilized spatial transcriptomics in a similar or higher number of PDACs (n = 13 and 30, respectively). In addition, the latter cohort (PMID 392944960) analyzed a more diverse set of tissues and tumors, including liver metastases. The authors already reference one of these studies (PMID 39227743), but I suggest including the second manuscript as well and perhaps rephrasing or contextualizing the largest cohort statement. A recent study has compiled published data into larger re-analyses of scRNA-seq data consisting of >200 PDAC samples (e.g. PMID 39636224).

We thank the reviewer for this clarification. We have updated the text to cite both recent Nature Genetics studies (PMIDs 39227743 and 39294496) and to rephrase our statement accordingly. Our cohort is distinct in representing one of the largest prospectively collected and deeply characterized sets of primary PDAC tumors profiled with both single-cell RNA-seq and spatial transcriptomics from the same samples, with matched pre- and post-treatment pairs and detailed clinical annotation. We have also added reference to

the recent meta-analysis (PMID 39636224), which aggregates over 200 publicly available PDAC datasets, to appropriately contextualize our work within the current literature.

4) Lines 123 - 124 (Figure 1a): Why are >25% of the samples negative for a KRAS mutation? And >20% have no identified genetic driver? These null results impact both biopsies and surgical resections. I assume the lack of driver mutation detection is due to low cellularity in PDAC which does occur often. Yet many of these same tumors indicate >25% epithelial cells in Extended Figure 1k, which does not add up unless many of the epithelial cells are not tumor cells?

We appreciate the reviewer's attention to this detail. The absence of driver mutations in some samples is in part due to unavailable sequencing data, that is now denoted as "Data not available" in Figure 1a (now clarified in the revised legend). For these cases, no residual tumor material was available to perform sequencing. When excluding such NA cases, 85% of sequenced tumors in our cohort harbored a KRAS mutation, and 93% had at least one driver mutation in KRAS, TP53, SMAD4, or CDKN2A. These frequencies are consistent with large US and international studies, which report KRAS mutations in 82–95% of PDACs and mutations in at least one of these four genes in over 90% of cases (PMID 35948708, 27526321, 32025013).

5) Lines 145 - 149 (Extended Figure 1c): There are epithelial cells that seem to form patient specific clusters, but there are also many to overlap across samples in 1c in the central part of the largest blue cluster in 1b.

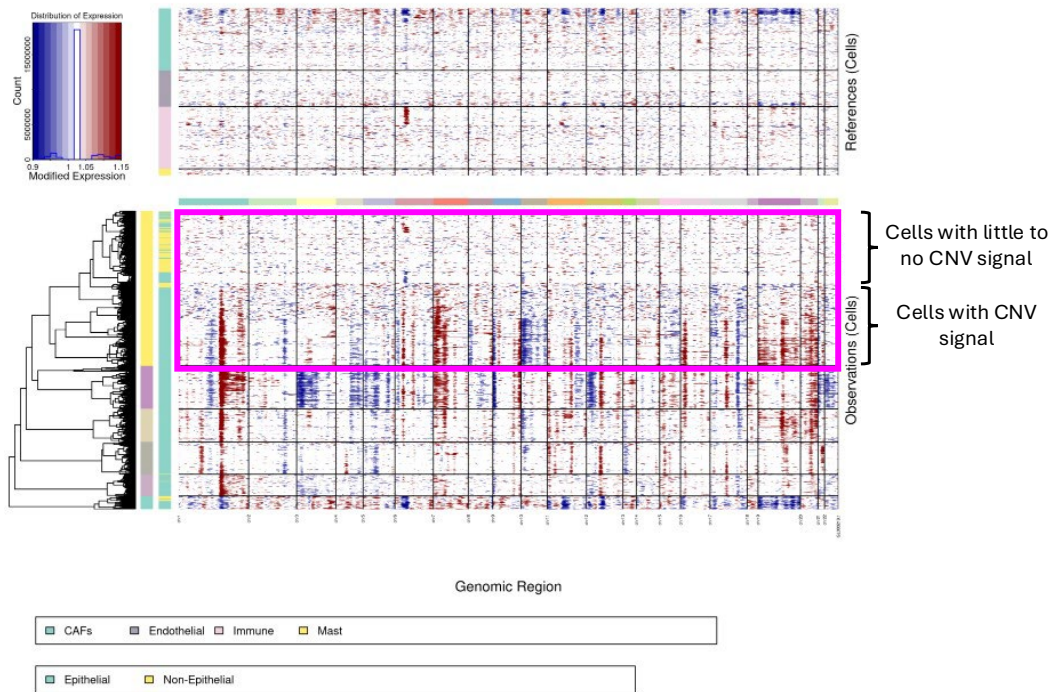
To clarify the extent of patient specificity of transcriptional clusters within the epithelial compartment, we quantified the patient composition of each transcriptional cluster using normalized Shannon entropy. After performing Leiden clustering on the epithelial cells, we selected a resolution that yielded at least as many clusters as samples (64 clusters for 62 samples), ensuring sufficient granularity to detect patient-specific structure. Using an entropy threshold of 0.3 to define patient specificity, 76.6% of clusters (49/64) were classified as patient-specific. Among the 15 clusters that did not meet this criterion, five contained >30% non-malignant epithelial cells, for which patient specificity is not expected; excluding these increases the proportion of patient-specific clusters to 83%. We have updated Extended Fig. 1 (panels d–f) to include these analyses and added corresponding descriptions in the Results ("The landscape of the PDAC TME at the single-cell and spatial levels") and Methods ("Quantification of sample specificity of clusters using Shannon entropy") sections.

6) Lines 171 - 172: *Why were the phylogenies inferred manually (and as described in lines 766 - 785)? It is not clear how clusters were identified as distinct or how strict this analysis was, and it seems potentially error-prone because scRNA-seq data presents many challenges for phylogenetic analyses. There are also a number of algorithms to minimize or avoid these documented issues. For example, Liu et al. (PMID 37553438) have recently published a summary of these approaches (see their Introduction) with a novel tool for inferring phylogenies from sc-RNA seq data. The authors should consider using this tool or a similar one to validate their manually derived results. It is also unclear how ploidy or whole genome duplication was handled.*

We recognize the reviewer's concerns regarding the potential limitations of manual cluster selection and phylogeny inference from scRNA-seq data. To clarify, our initial approach followed the strategy used in one of the early studies applying inferCNV for clonal inference, which also manually constructed phylogenetic trees (PMID 35948708). However, we agree that manual procedures may introduce bias. To address this, we fully automated our pipeline, as detailed below, to ensure reproducibility and to minimize subjectivity in tree construction.

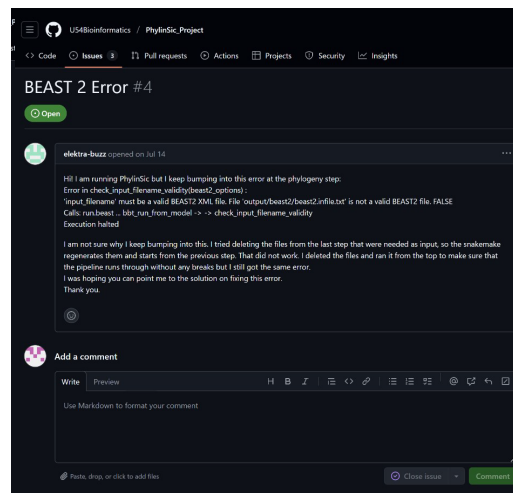
Regarding clone (cell cluster) identification, the workflow is fully automated all but one samples. Across the 50 tumors, the automated inferCNV-based pipeline identified 285 clones. Only one sample required a small, targeted refinement step. In that case, the automated inferCNV thresholds merged two biologically distinct groups: cells with clear CNV signal and a population with little to no CNV signal (see Rebuttal Fig. 11). Because this represented an obvious under-clustering event, we "manually" separated these two populations, by forcing inferCNV to call one more cluster, while leaving all other clone

boundaries unchanged. The description of this process was indeed missing from the methods and is now detailed in a new Methods section “Clone Identification”.



Rebuttal Figure 11: InferCNV output. Clusters of cells with the same CNV profile (clones) are identified in an unsupervised way.

Regarding the phylogenetic tree construction, we first attempted to run the tool PhylinSic recommended by the reviewer (Liu et al., PMID 37553438). However, we encountered a reproducible execution error that prevented the tool from running to completion, and despite contacting the authors, we were unable to obtain a resolution (Reb. Fig. 12).



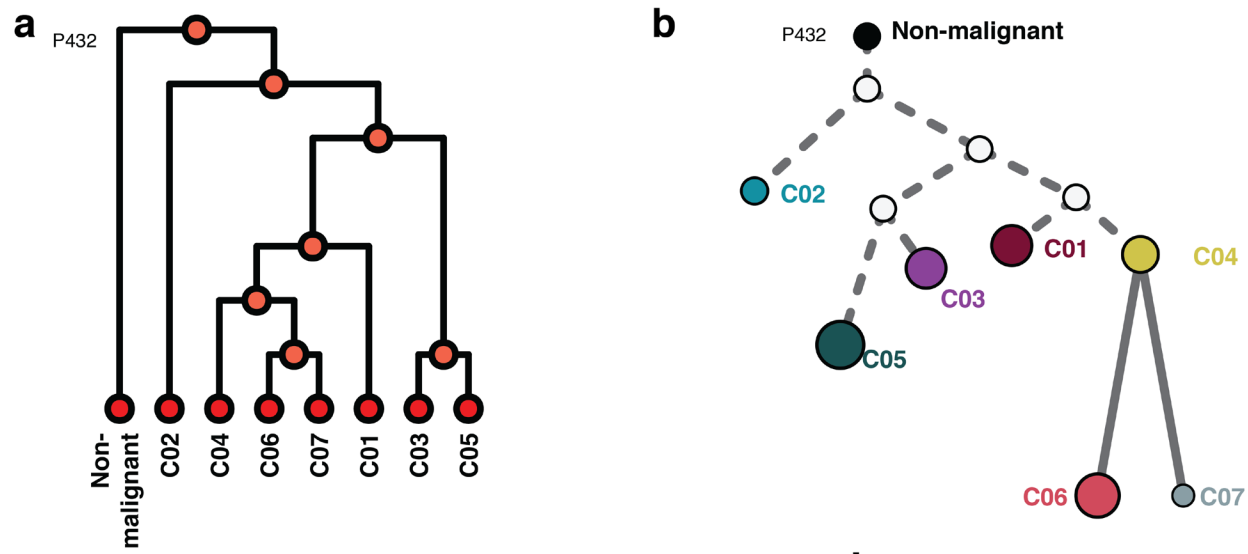
Rebuttal Figure 12: Screenshot of the issue opened in PhylinSic's GitHub page.

Because we still aimed to implement an automated phylogeny-inference method, we used a maximum parsimony approach, which is well-suited for CNV data and has been previously applied for reconstructing clonal relationships from copy-number profiles (PMID 35948708, 27526321).

We note that although PhylinSic addresses several challenges of scRNA-seq-based SNV phylogenies, our analysis relies on CNVs inferred by inferCNV. CNV-based trees follow distinct evolutionary dynamics from SNV-based trees: a substantial proportion of CNV phylogenies exhibit punctuated evolution (~57% of diploid tumors) rather than gradual accumulation of alterations (PMID 27526321, 32025013). Maximum parsimony is therefore a natural fit for CNV-based evolution and provides an appropriate automated, topology-based tree reconstruction.

However, maximum parsimony does not specify ancestor–descendant relationships. Therefore, after computing the parsimony tree, we applied a small, rule-based refinement step that preserves the topology. Specifically, for each pair of directly connected clones, we examined whether the CNV profile of one clone was a subset of the other. When this was the case, and when (a) the reverse was not true, and (b) the subset clone carried a smaller overall CNV burden, we designated the subset clone as the ancestor. In all other cases, the two connected clones were interpreted as sharing a common ancestor (either observed in the data or unobserved but implied by parsimony). This step does not modify the tree structure produced by maximum parsimony; it only resolves ancestor–descendant directionality using an objective criterion derived from CNV patterns.

Rebuttal Fig. 13 provides an example in which the final refined tree is identical to the original parsimony output. Full methodological details are included in the updated Methods section (“Phylogenetic tree construction”), and we have updated the relevant figures (Fig. 2, Suppl. Fig. 1, Suppl. Fig. 2).



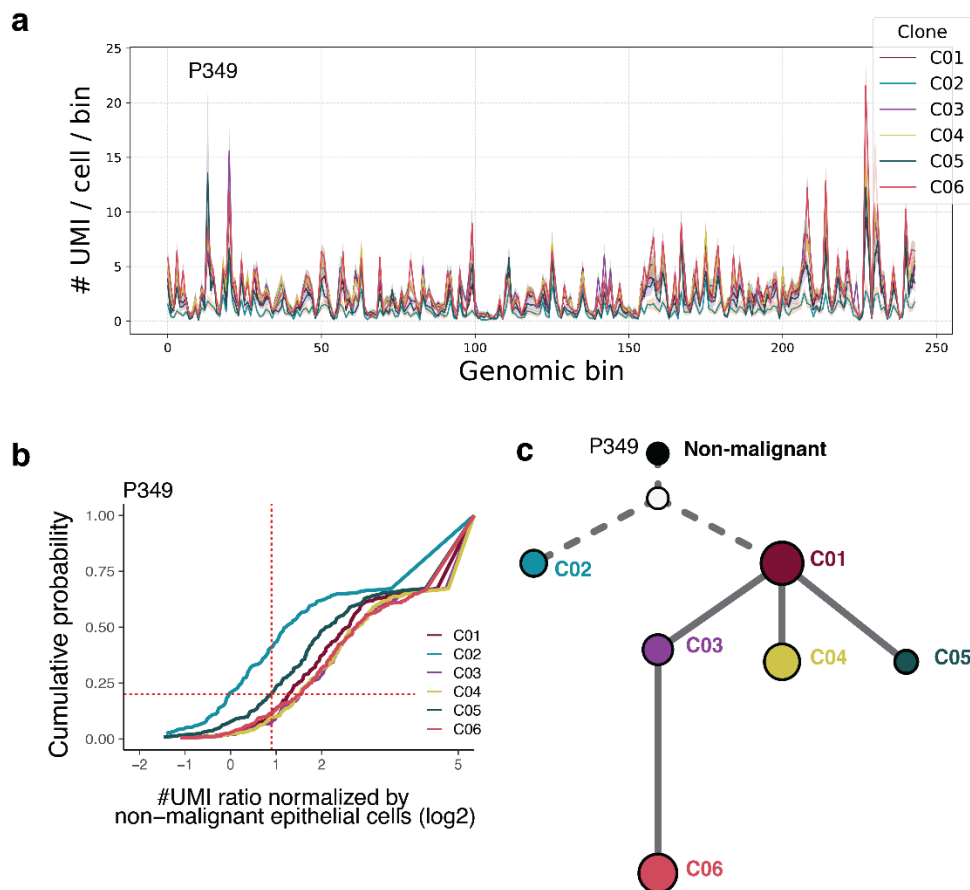
Rebuttal Figure 13: a. Output of maximum parsimony for sample P432. b. Tree constructed for P432 after refining for ancestral relationships.

Measuring ploidy and WGD is a challenge for scRNA-seq-based phylogenetic analyses, and we note that there is currently no consensus approach in the literature for reliable detection of WGD events from scRNA-seq data. To address the reviewer's comment, we performed an exploratory assessment of WGD by comparing the total UMI counts of malignant clones to those of non-malignant epithelial cells within the same sample. Relative UMI shifts were then used only as supportive evidence to further validate our phylogenetic tree construction, under the premise that if a clone has undergone WGD, its descendants should also reflect WGD. We classified a clone as WGD-positive if over 80% of its genomic positions exhibited a $\log_2$ UMI ratio (malignant epithelial cells vs. non-malignant epithelial cells) exceeding 0.9, consistent with the ~ 2 -fold shift expected for whole-genome doubling.

Across all tumors, we identified WGD in 32.6% of samples, which aligns with reported frequencies in the literature (PMID 30013179, 37400777). In one sample, shown in Rebuttal Fig. 14, an entire branch of the phylogeny, comprising clones C01, C03, C04, C05, and C06, displayed elevated UMI levels indicative of a shared WGD event, whereas clone C02 did not. This demonstrates that WGD occurred once along that lineage and was inherited by all downstream clones. In all other tumors containing a WGD-positive clone, that clone was always observed at a leaf of the phylogenetic tree, meaning it had no detectable descendants in our dataset. This pattern is fully consistent with our phylogenetic framework: once a clone acquires WGD, all its descendants should retain this signal, and the placement of WGD-positive clones either across an entire branch or exclusively at leaf nodes supports the internal consistency of our inferred trees.

Given the limitations of using UMI counts as a proxy for ploidy in scRNA-seq, we have restricted this analysis exclusively to validating tree construction and have not drawn further conclusions from it. We believe this cautious approach accounts for potential WGD while avoiding overinterpretation beyond the scope of our study.

Full methodological details for WGD analysis are included in the Methods section “Inferring whole genome duplication in tumor clones”.



Rebuttal Figure 14: Ploidy analysis for sample P349, shows Whole Genome Duplication for clones C01, C03, C04, C05, C06, but not for C02. a, Averaged Unique Molecular Identifier (UMI) counts across tumor clones. b, Cumulative distribution of relative ploidy inference across tumor clones. c, Phylogenetic tree.

7) Lines 177 - 219 (Figure 2): Why were these two cases (P323 and P432) selected for these analyses? These two cases remain prominent throughout the manuscript (e.g., Figure 3). Seems like a missed opportunity to extend these analyses to other cases. What is the evidence that “diverse clonal populations contribute to a diversity of transcriptomic signatures”? There is no quantitative analysis in this section that actually supports that statement, instead there are observations on the UMAPs when “cells within a single clone formed distinct clusters, while in others, cells from multiple clones clustered together.” For

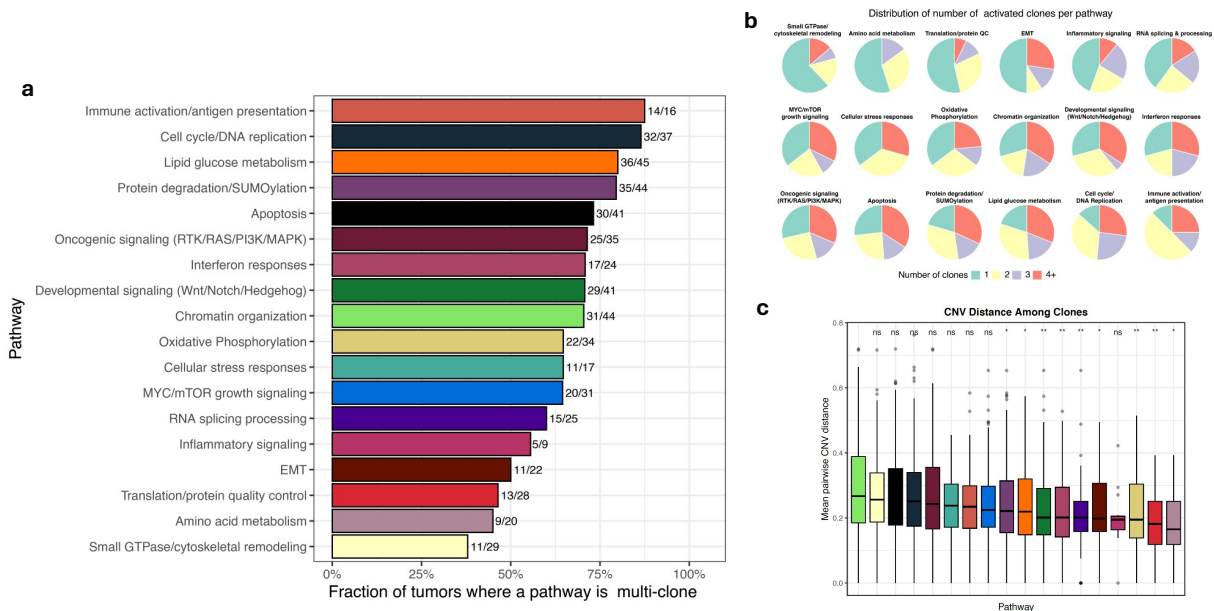
example, it is not clear that there is a relationship between transcriptomic signatures and the number of CNV clones or their genetic divergence from one another (in fact Figure 4 later shows that some cases have two CNV clones and two transcriptomic signatures, whereas others have as many as 8 clones all sharing the same transcriptomic signature).

We appreciate the reviewer's suggestion to clarify case selection and to provide more quantitative support for the relationship between clonal and transcriptomic diversity. Cases P323 and P432 were selected as illustrative examples to demonstrate our per-sample analytical workflow; however, these analyses have been performed across the entire cohort, as shown in Suppl. Fig. 2. For instance, Figure 4 compiles cohort-level analyses, including the number of epithelial clones per sample, recurrently altered loci across clones, and pathway enrichment results. We also report CNV differences between classical and basal clones, in Suppl. Figure 4. Likewise, Figure 5 assesses clone-TME interaction heterogeneity across the entire cohort and systematically evaluates how CNVs in ligand and receptor genes modulate these interactions across samples. Our classical-to-basal transition analysis (Fig. 7, Suppl. Fig. 7) was also performed across the entire cohort, but only five samples exhibited a transition; downstream analyses were therefore restricted to those samples. We do agree that some additional analyses would strengthen our manuscript, and in particular that statistical evaluation was missing. In response, we incorporated statistical tests across analyses; we performed a new pathway-level assessment to quantitatively evaluate clonal-transcriptomic relationships, which is described in this response; and we expanded our global analyses of the matched samples, as detailed in our response to comment 10. In addition, following a comment from Reviewer 3, we now include a validation of our LRI findings in the spatial cohort as well (Ext. Fig. 4).

To directly address the reviewer's request for global, quantitative evidence linking clonal diversity to transcriptomic diversity, we expanded our analysis to test whether individual clones exhibit distinct pathway activation patterns across the cohort. Specifically, we assessed, for each pathway and for each clone, whether that clone shows significantly elevated pathway activity relative to all other clones in the dataset. This framework allowed us to quantify, in a principled and cohort-wide manner, the extent to which different clones activate different transcriptional programs.

Using this approach, we found substantial heterogeneity in pathway activation across tumors and across clones. Several pathways, such as protein degradation/SUMOylation, and chromatin organization were activated in the majority of tumors (45/50, 44/50, and 44/50, respectively), whereas inflammatory signaling pathways were activated in only 9/50 tumors. Within tumors, immune activation and cell cycle-related pathways were the most frequently multiclonal, with 87.5% and 86.5% of tumors showing activation of more than one clone for these pathways, respectively. In contrast, small GTPase-related pathways were multiclonal in only 38% of tumors with pathway activation (Rebuttal Fig.

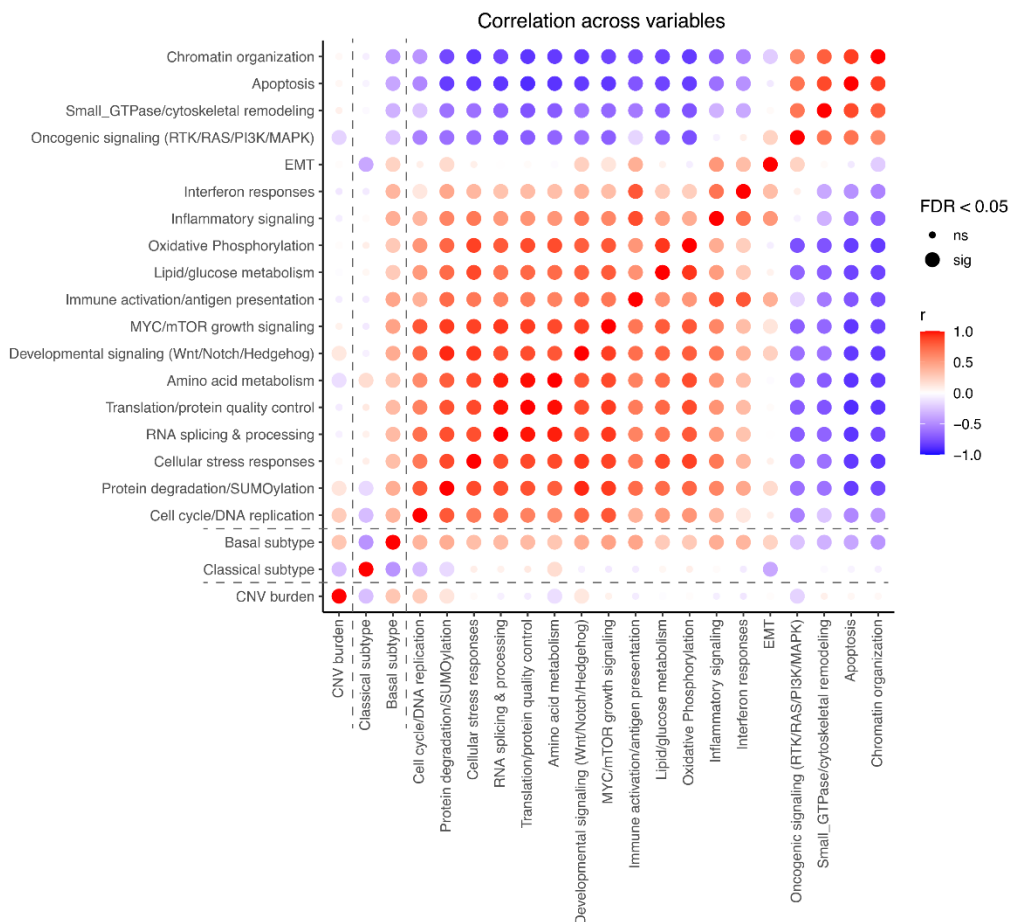
15a,b – also included in Fig. 4 and Suppl. Fig. 4). We further examined whether CNV divergence among clones explains pathway divergence (Rebuttal Fig. 15c – also included in Suppl. Fig. 4). Although a few comparisons reached statistical significance, effect sizes were small, and no consistent trend was observed across pathways. Overall, these results show that pathway activity may differ across clones in some tumors, although these differences are not systematically explained by CNV divergence, and many pathways are shared across clones. We have refined the text to clarify that transcriptional programs do not show a consistent relationship with CNV divergence.



Rebuttal Figure 15: a, Fraction of tumors where a pathway is multiclonal. b, Distribution of number of activated clones per sample per pathway. c, CNV distance among clones that are activated for the same pathway within a sample.

Finally, we wanted to further examine how pathway activity relates to transcriptomic subtype and CNV burden, so we computed pairwise correlations across all variables (Rebuttal Fig. 16 – also included in Fig. 4). Our goal was to determine whether specific pathway programs co-vary with clonal genomic burden or align with established PDAC subtypes. This analysis revealed that CNV burden was positively correlated with the basal subtype and inversely correlated with the classical subtype, consistent with higher genomic instability in basal tumors. Pathways related to cell cycle, translation, and protein degradation formed a cluster positively associated with the basal subtype. In contrast, pathways involving chromatin organization, apoptosis, and small GTPase/cytoskeletal remodeling showed inverse correlations with both this pathway cluster and the basal subtype, while exhibiting little association with the classical subtype. Oncogenic signaling was inversely correlated with CNV burden, in line with the early acquisition of KRAS

mutations. EMT displayed a distinct pattern, showing strong positive correlation with the basal subtype and negative correlation with the classical subtype.



Rebuttal Figure 16: Correlation across pathways, subtypes, and CNV burden of clones.

Full methodological details for this new pathway analysis are included in Methods sections “Pathway analysis” and “Quantification of pathway activity at the clonal level”.

8) Lines 246 - 253 (Figure 3b): Were there no *KRAS* amplifications (PMID 31932696), *CDKN2A* deletions, or *SMAD4* deletions identified? “Alterations” is supposed to be Alterations.

We thank the reviewer for this observation. The *CDKN2A* and *SMAD4* loci (9p21 and 18q21, respectively) were indeed frequently altered in our clones (9p21: 55% altered, 46% deleted; 18q21: 62.5% altered, 53% deleted), while the *KRAS* locus was amplified in 16.5% of clones; since our figure displays only the loci that are altered in more than 60% of clones, only the *SMAD4* locus was highlighted in the oncoprint. To improve clarity, we have updated the figure and its legend to indicate that only the most common alterations are shown, and we also report the presence and frequency of all the events in

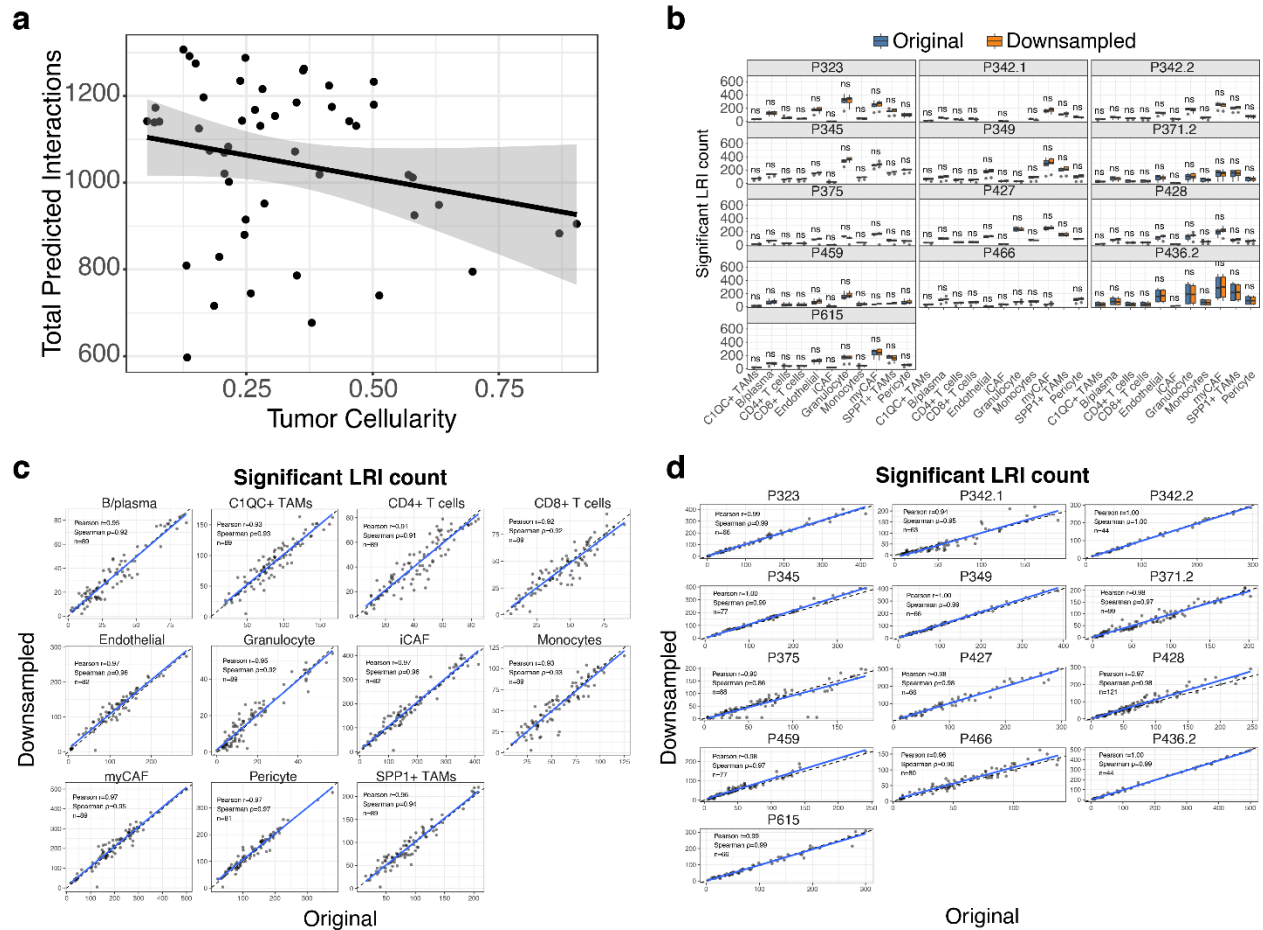
Suppl. Table 3. The typographical error “alternations” has also been corrected to “alterations.”

9) Lines 297 - 306 (Extended Figure 2b): How were low, intermediate, and high defined? On the left side of the heatmap there are 11 samples (P567.2 through P480.2) that exhibit predominantly “High” interactions with the TME. About one third of these cases (P518, P643, P385, and P480.2) likely have very low cellularity according to the sequencing results. Is it possible that lower cellularity may indicate High TME interactions?

We have now clarified specifically how interaction levels were defined- “Low”, “Intermediate”, and “High” groups were based on total predicted ligand-receptor interactions (LRIs) per sample as computed by CellPhoneDB, with samples into tertiles: the top 33% were labeled “High”, the bottom 33% “Low”, and the remaining third “Intermediate.” The description of this categorization has been added to the Methods section (section “Ligand-receptor interaction analysis”).

To address the reviewer’s observation that samples exhibiting “High” TME interaction levels tend to have lower tumor cellularity, we correlated the total number of unique interactions per tumor sample, to each tumor’s cellularity (Rebuttal Fig. 17a). This initial correlation wasn’t significant, but to further assess whether a high or low cellularity tumor could introduce technical bias to our interaction analysis (i.e., whether lower cellularity samples might spuriously appear to have more LRIs), we performed a downsampling analysis restricted to the top 25% most cellular samples ($n = 13$, cellularity 49–91%). In this analysis, malignant epithelial cells were downsampled to match a cellularity equivalent to the median of the bottom 25% cellular samples (median=16%), after which we re-ran the sample-level LRI analysis. As shown in the resulting boxplots and correlation plots (Rebuttal Fig. 17b-c), LRI counts remained similar between the original and downsampled datasets, and no systematic differences were observed within individual samples (paired Wilcoxon rank-sum test FDR-corrected p-values non-significant, Pearson $r > 0.9$, FDR-corrected p-values $\lll 0.01$ for all cell types and samples).

These results demonstrate that tumor cellularity does not confound the detection or quantification of significant LRIs in our analyses.

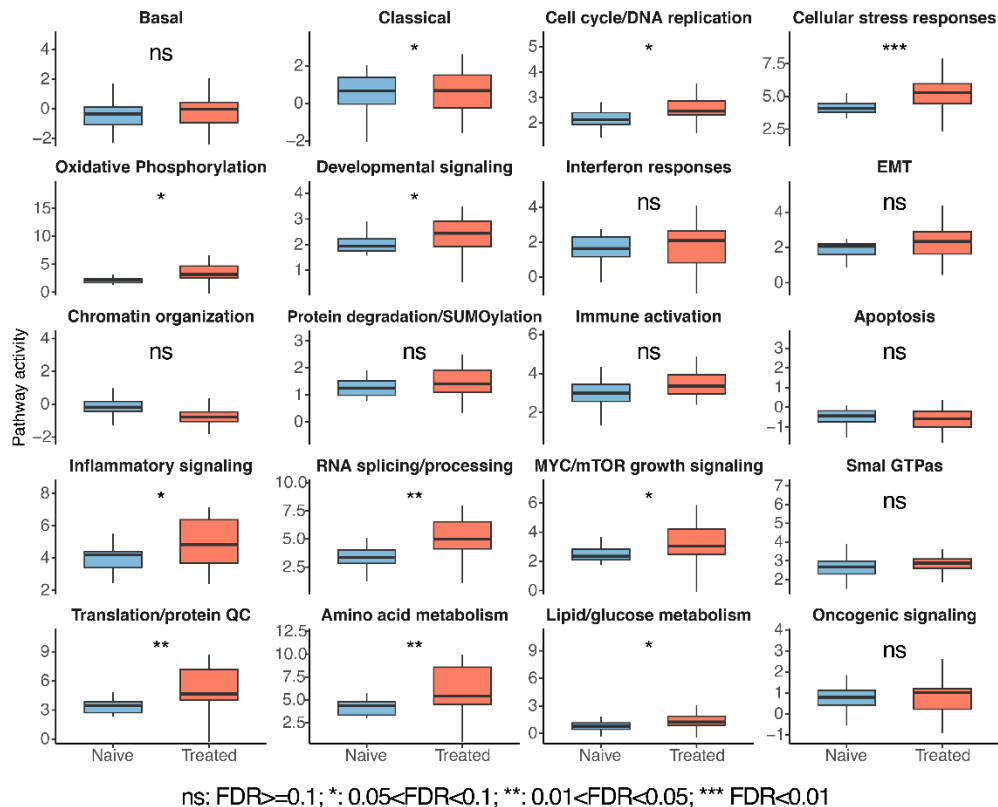


Rebuttal Figure 17: a, Correlation of predicted number of interactions to tumor cellularity per sample. **b**, Comparison of significant LRI counts between tumor clones and the TME in the original high-cellularity samples and in the same samples after downsampling malignant cells to achieve 16% cellularity. **c**, Correlation of significant LRI counts between tumor clones and individual TME cell types across samples; each dot represents one clone. **d**, Correlation of significant LRI counts between tumor clones and all TME cell types within each sample; each dot represents one clone.

10) Line 361: This section mentions six cases were analyzed but much of the text and much of Extended Data Figure 3 focuses on one case (P466). This is similar to the critique above that in some sections of the manuscript there is too much emphasis on relatively few cases given the size of the cohort.

We agree that the original presentation placed disproportionate emphasis on the matched case P466, which we initially highlighted because it clearly illustrated the analytical framework. In response, we removed panels that were unique to this sample and added a new Extended Figure 7 that summarizes all six matched pre- and post-treatment cases. We also incorporated two additional panels describing cell type composition across matched samples and the number of clones exhibiting different patterns of elimination or expansion following therapy.

In addition, we have expanded our analysis beyond ligand–receptor interactions to specifically assess pathway activity dynamics across treatment. We compared pathway activity in pre- versus post-treatment samples using a linear mixed-effect model (pathway activity $\sim$ Treatment + (1|sample)), adjusting p-values across pathways using Benjamini–Hochberg FDR. Although limited by sample size ($n = 6$), several biosynthetic and metabolic programs, including cell-cycle, oxidative phosphorylation, MYC/mTOR signaling, and RNA processing, were modestly increased post-treatment (Rebuttal Fig. 18 – also included in Ext. Fig. 6).



Rebuttal Figure 18: Pathway activity in matched samples collected before and after treatment with chemotherapy.

11) Lines 419 - 421: Using the word significantly here requires a sufficiently low p-value or other statistical test.

Due to the limited number of classical-to-basal transitions in our cohort ($n = 5$), formal statistical testing (e.g., Fisher's exact test) was not feasible. To assess whether the 8q24 (MYC locus) amplification was disproportionately associated with this transition, we performed a permutation-based enrichment test. This analysis yielded an empirical $p = 0.0002$, while no other transition type showed significant enrichment for 8q24 amplification. The text has been revised accordingly.

12) Lines 426 - 429: This section requires more explanation, it is unclear how this analysis was done.

Upon revisiting the analysis, we determined that the number of tumors carrying classical-to-basal transition CNVs with available survival data was too small to support a statistically robust comparison. Because the results were therefore underpowered and potentially misleading, we have removed the survival analysis from the revised manuscript.

13) Lines 464 - 473: Can the authors provide evidence for successful knock down of MYC (also mentioned in lines 823 - 824)? Is it possible to validate this in additional cell lines, preferably human and not just mouse?

We have now included a plot demonstrating downregulation of MYC expression in KMCE cells (Suppl. Figure 8c). In addition, we performed new experiments by knocking down MYC in human PDAC cell lines with high baseline MYC expression. These experiments showed effective MYC suppression in knockdown experiments and downregulation of basal markers and upregulation of classical markers in both knockdown and inhibitor experiments, showing congruence of our finding across multiple species (see Rebuttal Fig. 10 and Suppl. Fig. 8e-g).

14) Line 489: This is a bit confusing, can the authors provide the correlation value? It appears that BayesPrism annotated the far right as Epithelial cells in 4c and then manually annotated clusters in 4d found these to be CD8+ T cells, for example. Panels 4c and 4d look substantially different, maybe the colors are off?

We have updated the figure and legend to ensure consistent labeling and clearer interpretation. The color mismatch between panels c and d of Ext. Fig. 2 (previously Ext. Fig. 4) was unintentional and has now been corrected. We performed a chi-square test with permutation (100,000 iterations) to assess the association between the automatic (BayesPrism-based) and manual cell type annotations. The analysis showed a strong and consistent association (χ^2 statistic = 76,154, Cramér's V = 0.53, p = 9×10^{-6}). The statistical test is now added to the methods (section "Spatial transcriptomics library preparation, sequencing, and data processing"), and the p-value is added in the results

15) Line 499-500: Can the authors provide the concordance value? There appear to be discrepant CNVs in clone 1 across the two datasets, e.g. a deletion in chr. 1 becomes an

amplification, an amplification appears in chr. 8q (MYC?), and then an amplification appears in chr. 19. In fact clone 1 in the spatial data resembles clones 2 and 3 as opposed to the scRNA-seq data.

We thank the reviewer for this observation. To clarify, clone labels (e.g., C01, C02, etc. in the scRNA-seq data and spC01, spC02, etc. in the spatial data) are dataset-specific and do not correspond one-to-one across modalities. We have now revised the spatial clone labels in the manuscript to spA, spB, etc., to avoid confusion. Our intent was to highlight that the overall CNV landscapes captured by single-cell and spatial profiling of the same samples were consistent, rather than to imply direct clone-level correspondence.

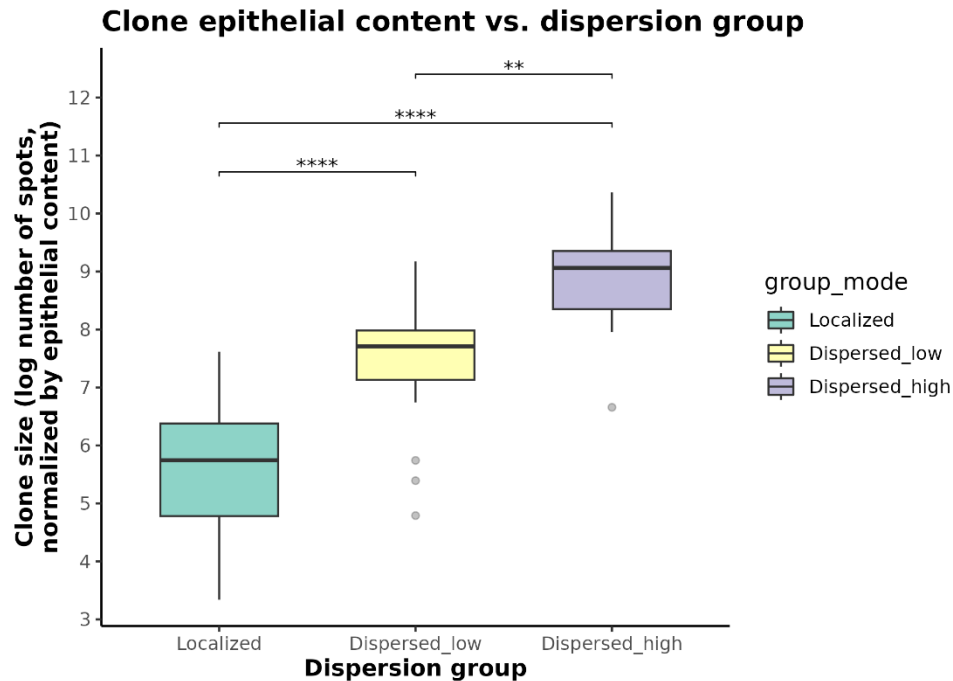
We did not expect to identify identical clones across modalities, as also noted in the manuscript, given inherent selection biases such as differences in sampling regions and processing methods. At the sample level, CNV patterns were largely preserved between modalities: on average, 79% (per sample min: 65%; max: 94%) of CNVs identified in single-cell data were also detected in the matched spatial dataset, and 73% (per sample min: 51%; max: 87.5%) of CNVs identified in the spatial data were found in single-cell data. The higher overlap observed when comparing spatial data against single-cell data may reflect the greater clonal resolution of single-cell data, which can capture additional subclonal diversity not always represented in spot-based spatial measurements. In contrast, the spatial modality likely reflects the dominant clones within each tissue section, resulting in a higher fraction of shared CNVs when comparisons are made at the sample level. To quantify the extent of similarity despite these differences, we computed the Jaccard similarity index between CNV profiles of all clone pairs within each matched sample and defined clones with a similarity >0.4 as concordant. On average, across all samples, approximately 57% of spatial clones and 44% of single-cell clones exhibited high similarity to a clone in the other dataset, with some samples reaching up to 100% and 83%, respectively.

We now present this data in the manuscript (Results section “Subpopulations within the tumor epithelial compartment reveal a dynamic process of clonal evolution”, third to last paragraph).

Line 525: To what extent does dispersal correlate with clone size?

We examined whether clone size (number of spots per clone normalized by BayesPrism-predicted epithelial content) is associated with spatial dispersal. Using our already defined dispersion groups (Localized clones, Lowly dispersed clones, Highly dispersed clones), we compared clone sizes across categories with pairwise Wilcoxon tests. Clone size increased significantly with higher levels of dispersal (all pairwise FDR adjusted p -values

< 0.01; Rebuttal Fig. 19 – also presented in Suppl. Fig. 8), indicating that larger clones tend to occupy a broader spatial area within the tissue.



Rebuttal Figure 19: Comparison of number of cells (clone sizes) in localized, lowly dispersed, and highly dispersed clones.

Line 561 (Discussion): The authors need to emphasize what are the unique findings of this study as opposed to confirmatory findings of past studies (including their own). Many of the key findings have been previously published, it is not clear what is novel here. Even a “key finding of this study” (line 594) regarding MYC as potentially a driver of classical to basal transitions has been previously published (PMID 35118421).

The reviewer raises an important point regarding novelty. To clarify our study’s distinct contributions, we note the following:

- Largest study to date examining clonal evolution within primary PDAC:** There have been previous studies characterizing PDAC heterogeneity using either single-cell RNA sequencing or spatial transcriptomics, but these have generally involved smaller cohorts or lacked integrated clonal reconstruction. In our study, we analyze 62 primary tumors, representing the largest prospectively collected cohort profiled with both modalities to date, and reconstruct clone-level evolutionary lineages across these samples. This cohort size provides the statistical power and diversity needed to identify recurrent patterns of clonal

diversification, subtype transitions, and tumor–microenvironment interactions that are difficult to discern in smaller datasets.

- **Intra-tumor phylogenetic resolution of the classical-to-basal transition:** Prior work (e.g., PMID: 35118421) linked MYC amplification to squamous/basal phenotypes across different primary or metastatic sites. What had not been shown is how this relationship unfolds *within the same tumor*. In our study, we reconstructed clonal lineages based on CNV profiles and then examined the transcriptional identity of each clone along these lineages. Across all cases with a classical-to-basal transition, the basal clone consistently emerged from a classical ancestor that had newly gained MYC. This provides, to our knowledge, the first intra-tumor evidence that MYC amplification marks the evolutionary step at which a basal-like lineage appears.
- **Immune depletion during the classical-to-basal transition:** Across all transition pairs, we observe that MYC gain is accompanied by a coordinated loss of immune-response programs. This pattern is not present in classical clones that retain 8q24 amplification without transitioning, suggesting that immune depletion is not a passive consequence of MYC gain but may represent an additional requirement for the emergence of a basal-like lineage. This provides a mechanistic framework that links genomic change (MYC amplification) with transcriptional rewiring (loss of immune programs) during lineage divergence within individual tumors.
- **Data-driven and cross-species validation:** Our functional data supports a cell-intrinsic role for MYC in maintaining the basal state. In both murine and human MYC-amplified PDAC models, MYC knockdown reverses basal-like features and restores classical and immune-response programs. Together, these results provide direct genetic and mechanistic evidence that MYC amplification is not only associated with, but can also drive and sustain, the basal lineage within PDAC.
- **CNVs define subtypes in PDAC:** Beyond MYC, we show that CNVs at loci encoding classical or basal marker genes distinguish classical from basal clones. Classical clones preferentially carry amplifications of classical-marker loci and deletions of basal loci, while basal clones show the opposite pattern. These CNVs act as subtype-defining genomic features that help explain the transcriptional differences between PDAC subtypes.
- **Inclusion of pre- and post-treatment matched samples provides rare insight into treatment-driven clonal dynamics:** Our cohort includes six pairs of matched tumor samples collected before and after chemotherapy, offering an opportunity to observe how clonal compositions and tumor–microenvironment interactions change in response to treatment. Such matched datasets are uncommon in PDAC,

and this design allowed us to identify patterns of clonal persistence, elimination, and microenvironmental remodeling associated with therapeutic exposure.

- **Clone specific interactions with the TME are defined by CNVs:** In addition to identifying clone-specific differences in ligand–receptor interactions, our analysis reveals a direct link between tumor CNVs and the strength of corresponding microenvironmental signaling axes. Amplifications consistently enhanced ligand–receptor engagement, whereas deletions suppressed it, establishing a genetically encoded mechanism governing how malignant clones communicate with the TME. We strengthened this conclusion by adding spatial transcriptomic and multiplex immunofluorescence validation, both of which independently recapitulate CNV-associated modulation of interaction strength (Rebuttal Fig. 6 & 23).
- **Spatially dispersed clones facilitate invasion to lymph nodes:** Spatial transcriptomics with clone-level CNV mapping associates highly dispersed clones (defined by local Moran's I) are the ones that metastasize to regional lymph nodes. These clones display early upregulation of basal marker genes and EMT program, depletion of immune pathways, and MYC amplification, consistent with a transitional, invasion-competent state. This provides spatially resolved evidence that clonal distribution patterns relate to metastatic behavior.

We now highlight these contributions explicitly in the revised Discussion.

Line 634: “in our study” has been written twice. This has been corrected.

Line 1117: Extended Figure 4 is missing a title. A title has now been added in Ext. Fig. 2 (previously Ext. Fig. 4).

Comments from Reviewer #3:

1. Overview

In this manuscript, Kalfakakou and colleagues profiled a cohort of primary human PDAC samples ($n = 62$) with single-cell RNA-sequencing and spatial transcriptomics. In a limited number of cases, they performed low-passage whole genome sequencing; had paired pre- and post-chemotherapy specimens ($n = 6$); or had matched lymph node metastases ($n = 2$). Using inferCNV, they identified intra- and inter-tumoral heterogeneity in clones that were associated with distinct transcriptomic subtypes (e.g., classical, basal, mixed) and interactions with other cell types in the microenvironment. They identified MYC amplification as an important driver of classical-to-basal transition and found that

dispersed clones were enriched in the classical-to-basal transition state and associated with greater metastatic proclivity. Overall, the results of this study suggest that phenotypic heterogeneity in PDAC is driven by distinct genetic events, as demonstrated by amplification or deletion of loci harboring key genes associated with malignant cell states.

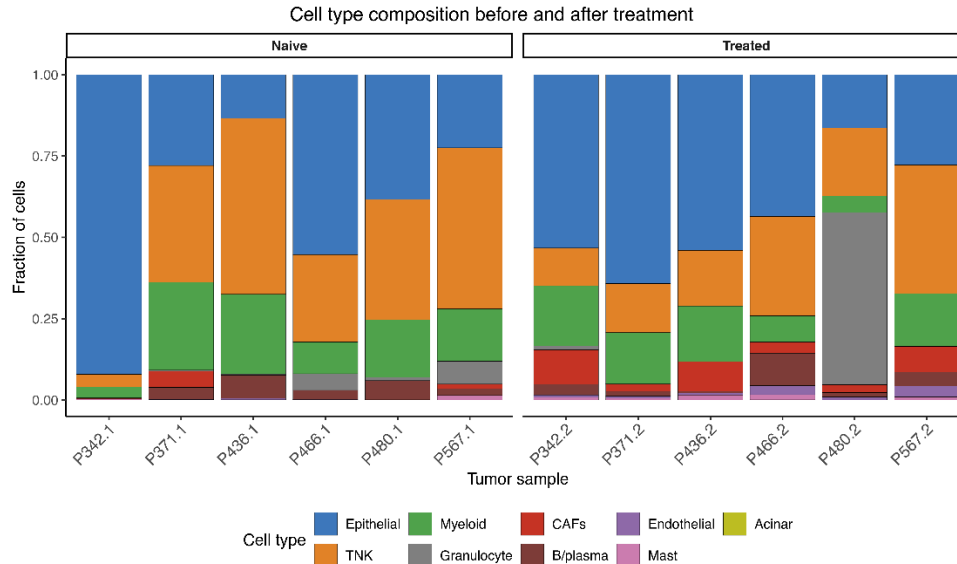
We appreciate the reviewer's summary of our study and recognition of its main findings. We have carefully considered the points raised and address each specific comment in detail below.

2. Major comments

a) Lines 150-152: The authors note that there was no significant differences in cell type composition between treated and untreated samples except a slight increase in the endothelial cell population. This is surprising given that chemotherapy is expected to at least reduce the cancer cell abundance. It would be helpful for the authors comment on this and perform this analysis on the paired cohort (n = 6) to see if any changes in cell type abundance are observed in that subset.

We appreciate the reviewer's insightful comment. As noted in the manuscript, our initial differential composition analysis across the full cohort (n = 62) revealed no significant changes in major cell types between treated and untreated tumors, apart from a modest increase in endothelial cells (Extended Fig. 1k). We believe this pattern may reflect the biological and clinical heterogeneity of the cohort, as well as limited statistical power when comparing across individuals with diverse treatment histories and tumor characteristics.

To more directly assess treatment-associated changes, we repeated the analysis within the subset of six patients with matched pre- and post-treatment samples using *propeller* in a paired design (Rebuttal Fig. 20 – also included in Ext. Fig. 6). Consistent with our cohort-level findings, few differences were detected. After FDR correction, only Endothelial and CAF populations showed modest but significant increases in treated samples (*FDR* = 0.008, 0.01, respectively).



Rebuttal Figure 20: Cell type composition in matched samples collected before and after treatment with chemotherapy.

b) Lines 229-231: *If the core biopsies represent a more limited sampling of the tumor compared to the resections, the lack of difference in number of clones would suggest that clones are highly interspersed in space. The authors should comment on this further in the Discussion.*

We agree this is an important observation. In our analysis, the number of detected clones did not differ significantly between core biopsies and surgical resections (Suppl. Fig. 4b), despite the expected differences in sampled tissue volume. While our data do not allow us to formally assess spatial intermixing, the similar number of clones detected in core biopsies and resections may indeed reflect a degree of spatial interspersion, at least within the sampled regions. We now note this possibility in the revised Discussion (second paragraph), while also acknowledging that sampling bias and technical resolution may influence these observations.

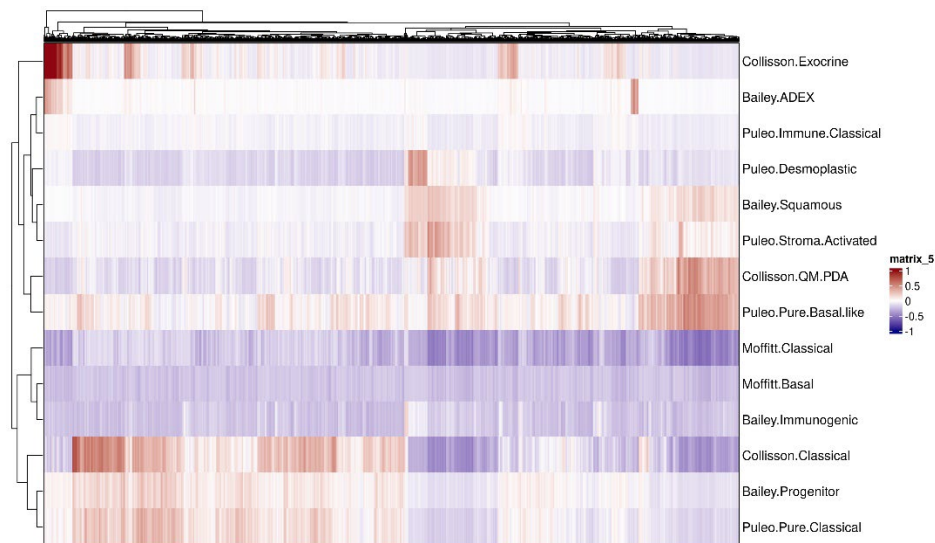
c) Lines 239-241: *While the authors state their intent to focus on the Moffitt basal and classical signatures, here they mention this small subset of cells that tested negative for both basal and classical signatures. These should be explored further given that several other malignant cell signatures in PDAC have been identified in the literature.*

We agree that further characterization of the subset of epithelial cells that did not score as either basal or classical according to the Moffitt signatures could provide additional insight into malignant cell heterogeneity. As shown in Suppl. Fig. 4b, when applied to all

the malignant epithelial cells of the cohort, the Moffitt Classical signature correlates strongly with the Puleo Pure Classical, Bailey Progenitor, and Collisson Classical signatures, while the Moffitt Basal signature correlates with the Puleo Pure Basal, Puleo QM-PDA, and Bailey Squamous signatures. Other proposed PDAC subtypes were not readily observed in our data.

In response to this comment, we performed a focused analysis of the epithelial cells that did not classify as either classical or basal. We examined whether these cells are enriched for alternative malignant programs described in prior PDAC studies and found that this subset indeed contains very rare populations expressing other published subtype signatures (Rebuttal Fig. 21 – also presented in Suppl. Fig. 4), including Collisson Exocrine, Bailey ADEX, Puleo Immune Classical, Puleo Desmoplastic, Bailey Squamous, Puleo Stroma Activated, Collisson QM-PDA, Puleo Pure Basal like, Moffitt Classical, Moffitt Basal, Bailey Immunogenic, Collisson Classical, Bailey Progenitor, and Puleo Pure Classical.

Notably, these “neither” cells represent only ~6% of all epithelial cells, limiting the strength of biological inference. We therefore retain the classification of this group as “neither” basal nor classical within the Moffitt framework, while noting that their expression profiles suggest the presence of rare malignant programs. These findings are now described in the revised Results (Section “Characterizing the clonal intra- and intertumoral heterogeneity in PDAC”, third paragraph).



Rebuttal Figure 21: Subtype signature enrichment in cells with negative enrichment for both the Moffitt Classical & Basal signatures.

d) Lines 267-268: It is surprising that treatment did not affect enriched pathways. How long after treatment were specimens collected? Did the authors do a subset analysis of the 6 paired pre- and post-treatment specimens?

While exact timing of collection after treatment with chemotherapy varied between patients depending on clinical scheduling, the typical interval between treatment completion and surgery was 4-8 weeks, with an average of 4.7 weeks, as documented in clinical records.

To address how chemotherapy induces changes in the transcriptional profile of PDAC tumors, we performed pathway-level analyses within the six paired pre- and post-treatment samples using a linear mixed-model framework (*pathway activity ~ Treatment + (1|sample)*). Although the small number of pairs limits statistical power, several biosynthetic and metabolic programs, including cell-cycle, oxidative phosphorylation, MYC/mTOR signaling, and RNA processing, were modestly increased post-treatment, indicating that the transcriptional effects of therapy are subtle but detectable within matched samples (Rebuttal Fig. 18). A variation of this plot is now also included in Ext. Figure 6.

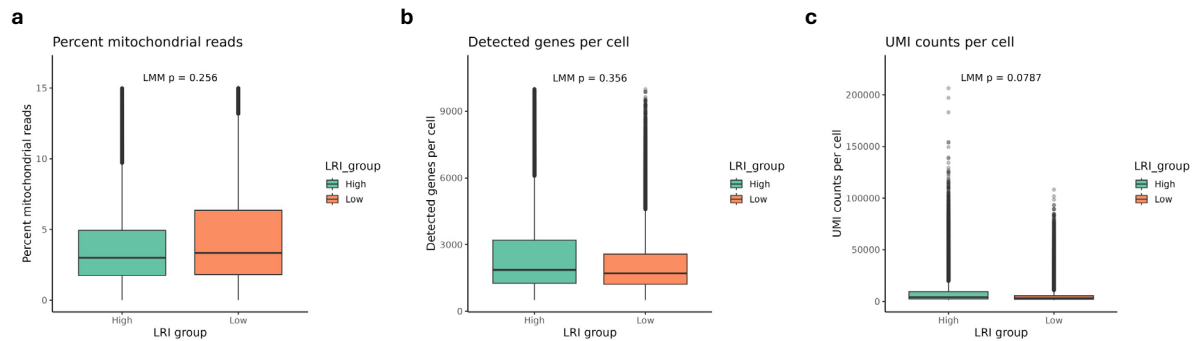
e) Lines 306-309: The LRI analysis performed is based on CellPhoneDB was applied to scRNA-seq data and therefore only based on gene expression data. For the pan-low and pan-high interacting tumors: was there any correlation to underlying sample/data quality of the scRNA-seq?

To evaluate whether the observed differences in total LRI counts across tumors could be influenced by scRNA-seq data quality, we examined standard QC metrics, including the number of detected genes per cell, total UMI counts, and the percentage of mitochondrial reads.

Importantly, as part of our standard preprocessing pipeline, all samples were subjected to identical QC thresholds: only cells with ≥ 1200 UMIs, ≥ 500 detected genes, and $\leq 15\%$ mitochondrial reads were retained. These criteria ensure that all samples entering the analysis have comparable baseline data quality.

To further assess whether any residual differences in data quality might explain variation in LRI counts, we classified tumors into high or low interaction groups (based on $\geq 10/13$ cell types tested) and fit linear mixed-effects models for each QC metric, with sample included as a random effect to control for sample-level variability.

Across all three QC metrics we found no significant differences between high- and low-interacting tumors after accounting for sample effects (Rebuttal Fig. 22). These results indicate that the differences in total LRI counts are not attributable to differences in scRNA-seq data quality between the groups.



Rebuttal Figure 22: Comparison of percent of mitochondrial reads (a), detected genes per cell (b), and UMI counts per cell (c), in tumors with high vs low interaction count with the TME. p-values are calculated with linear mixed-effect models.

We also examined the effect of tumor content on the number of interactions, as suggested by reviewer #2. First, we correlated the total number of unique interactions per tumor sample, to each tumor's cellularity, which wasn't significant (Rebuttal Fig. 17a). To further assess whether a high or low cellularity tumor could introduce technical bias to our interaction analysis (i.e., whether lower cellularity samples might spuriously appear to have more LRIs), we rerun CellPhoneDB for the top 25% most cellular samples ($n = 13$, cellularity 49–91%), after downsampling to match a cellularity equivalent to the median of the bottom 25% cellular samples (median=16%). This analysis showed that LRI counts remained similar between the original and downsampled datasets, and no systematic differences were observed within individual samples (paired Wilcoxon rank-sum test FDR-corrected p-values non-significant, Pearson $r > 0.9$, FDR-corrected p-values $<<<0.01$ for all cell types and samples) (Rebuttal Fig. 17b-c).

f) Lines 325-326: This sentence is an overstatement of what their data can infer regarding LR interactions.

We agree and the sentence has now been revised to: “Our analysis suggests that CNV alterations may influence LRIs within the TME, with potential implications for immune evasion, stromal remodeling, and therapeutic strategies.” (Section “Clonal diversity shapes a heterogeneous landscape of interactions with the PDAC TME”, 5th paragraph)

g) Lines 467-471: All of the human results were anchored on Moffitt basal and classical but the KMCE mouse PDAC cell lines do not appear to be enriched in Moffitt basal so the authors have chosen here to expand to other signatures. There appears to be some bias

in how certain subtype schemas are being chosen for different aspects of the paper. The authors should be more consistent in the scope of prior malignant signatures applied throughout the study. Moreover, this reviewer would suggest that the authors repeat the KMCE PDAC Myc knockdown experiments in human cell lines that are enriched for Moffitt basal at baseline.

We appreciate the reviewer's comment and fully agree on the importance of consistent subtype frameworks. Throughout the study, we primarily relied on the Moffitt classification for human data, as we have demonstrated that it is broadly consistent with other PDAC subtype schemas. For the mouse experiments, we included additional subtype signatures due to the well-recognized lack of a one-to-one correspondence between mouse and human gene expression profiles, which is an issue not specific to the Moffitt classification but inherent to cross-species comparisons.

It is worth noting that the Moffitt Classical signature is also the most enriched in the KMCE mouse dataset, which supports its relevance in this model (Fig. 7e). In response to the reviewer's suggestion, we additionally performed *MYC* knockdown experiments in two basal-like human PDAC cell lines harboring *MYC* amplification and enriched for the basal signature. This was also proposed by Reviewer 2 (comment 2). The results showed that *MYC* knockdown in basal-like human PDAC cells led to downregulation of basal markers and upregulation of classical markers, indicating a shift toward a classical transcriptional state - consistent with our model in which *MYC* amplification drives the transition from classical to basal phenotypes (Rebuttal Fig. 10). These results are presented in Suppl. Fig. 7f-h and in the revised Results (Section "MYC is required to maintain basal identity in mouse and human PDAC models", last paragraph).

h) Lines 491-492: The enrichment of iCAFs over myCAFs in chemotherapy-treated samples has been seen in prior studies such as PMID: 39227743 and should be cited.

Thank you for pointing this out. We have now cited PMID: 39227743 in the revised manuscript to acknowledge prior evidence of iCAF enrichment following chemotherapy.

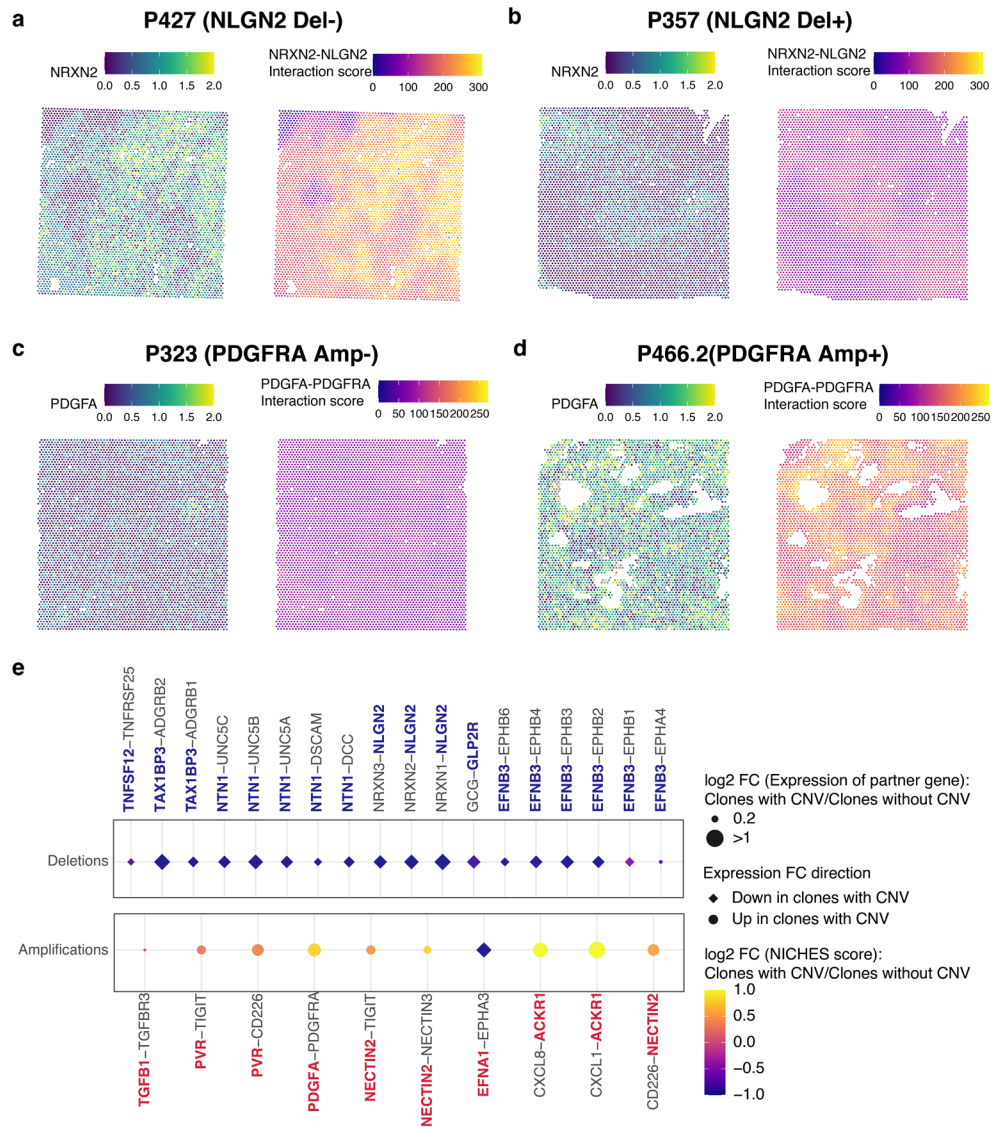
i) Lines 497-500: A clear advantage of Visium spatial data is the sequence information it provides rather than probe-based quantification. The authors are able to identify some of the same clones in the spatial data that they observed in the scRNA-seq but they stop short of doing LRI on the spatial data. This is a significant missed opportunity.

We have now extended our analysis to investigate LRIs directly in the spatial transcriptomics data, thereby providing spatial support for the CNV-associated communication patterns identified in the single-cell data.

First, we performed multiplex immunofluorescence (mIF) for a representative ligand–receptor pair (NECTIN2-TIGIT), in tumors with and without CNVs. The proportion of NECTIN2⁺CK19⁺ epithelial cells and TIGIT⁺CD8⁺ & TIGIT⁺CD3⁺ immune cells was significantly higher in the tumor with the NECTIN2 amplification (Rebuttal Fig. 5,6), confirming increased co-localization of the two proteins in the presence of CNVs.

Next, we examined the spatial expression patterns of selected ligands and receptors in the Visium data to visualize their distribution within the tissue (Rebuttal Fig. 23a-d). To explore potential spatial co-occurrence, we use NICHES to compute interaction scores using the CellPhoneDB database as input and visualized the resulting spatial interaction maps (Rebuttal Fig. 23a-d).

To formally quantify CNV effects on ligand–receptor interactions, we computed $\log_2$ fold-changes in both NICHES interaction scores and partner-gene expression clones with or without a CNV, pooled across all spatial samples. The resulting summary plot (Fig. 23e) shows that clones harboring amplifications exhibit increased interaction strength for ligand–receptor pairs in which the amplified gene participates (highlighted in red), whereas clones carrying deletions exhibit reduced interaction strength for pairs involving the deleted gene (highlighted in blue). In this plot, dot color represents the $\log_2$ fold-change in NICHES interaction score (clones with versus clones without a CNV), dot size reflects the absolute fold-change of partner-gene expression change in the microenvironment surrounding each clone, and dot shape indicates whether the partner-gene expression is increased or decreased. Importantly, partner-gene expression refers to the ligand or receptor not affected by the CNV, reflecting changes in the composition or activation state of neighboring, non-epithelial cells rather than expression of the CNV-affected gene itself. Together, these new analyses integrate the spatial data into our study, supporting the conclusion that CNVs in epithelial clones systematically influence ligand–receptor signaling intensity within the tumor microenvironment. The Rebuttal Fig. 23 is now included in the manuscript as Extended Fig. 4.



Rebuttal Figure 23: a-d, Representative spatial expression and spatial interaction (computed with NICHES) maps for selected ligand-receptor (L-R) pairs affected by copy-number alterations (a-b, NLGN2 deletions suppress NRXN2-NLGN2 interaction, c-d, PDGFRA amplifications increase PDGFA-PDGFA interaction). For each tumor sample, the left panel shows the spatial distribution of the partner gene expression (i.e., the ligand or receptor not affected by the CNV), and the right panel shows the corresponding spatial NICHES interaction score for the L-R pair. Expression and interaction scores are visualized on matched spatial grids. e, Summary of CNV-associated changes in interaction scores across all spatial clones and samples for high-confidence deletion- and amplification-associated L-R pairs. Each point summarizes the fold-change in (i) interaction strength (color; NICHES log₂ fold-change between clones with and without a CNV) and (ii) partner-gene expression (size; mean expression of the partner gene in the neighborhood of clones with and without a CNV). Shapes indicate the direction of partner-gene expression change (circle, up; diamond, down). Ligands or receptors encoded within the CNV region are labeled in color (deletions: blue, amplifications: red).

j) *Line 530: This is an overstatement. A simple association in two cases without additional validation does not indicate that dispersed clones have a higher propensity to metastasize.*

The language has been revised to avoid overstatement and now reads: “Given the observed association between dispersed clones and lymph node metastases in these two cases, and considering the established link between the basal subtype and increased EMT activity and metastasis, we sought to investigate the subtypes of clones in our spatial dataset.”

k) *Lines 541-559: Why wouldn't a fully basal, EMT state be even more aggressive than a classical state with some indication of transition towards a basal state? Does the hybrid nature confer a specific advantage? Is this related to the partial EMT phenotype described previously (PMID: 29920274)? This should be explored and discussed more.*

We thank the reviewer for this insightful comment. We agree that a fully basal/mesenchymal EMT state is not necessarily the most metastasis-competent state. Foundational work from the Weinberg laboratory has established that EMT is best understood as a dynamic continuum, and that cells in intermediate or transitional EMT states can display enhanced metastatic potential due to retained epithelial programs combined with newly acquired mesenchymal traits (PMID: 19487818, 21436443, 27878502, 28187288, 30459476). Fully mesenchymal states often exhibit reduced proliferative capacity and impaired colonization ability, whereas intermediate states tend to be more plastic, stress-resistant, and tumor-initiating.

In PDAC specifically, Aiello et al. (PMID: 29920274) demonstrated that tumor cells frequently adopt a partial EMT phenotype, maintaining epithelial transcripts while activating mesenchymal and motility programs through alternative EMT trajectories. This intermediate state supports dissemination without requiring full mesenchymal conversion.

We appreciate the opportunity to clarify our interpretation. We do not conclude that hybrid or transitioning states are inherently more aggressive than fully basal/EMT states. Instead, we propose that the dispersed clones we observe may represent an intermediate or early phase of lineage plasticity. In our spatial dataset, fully basal clones were relatively rare, limiting our ability to systematically assess differences in metastatic behavior between transitional and fully basal states. Consequently, we refrain from making claims about their relative aggressiveness. Rather, we note that dispersed clones retain classical/epithelial identity while exhibiting elevated basal markers. This pattern is consistent with a transitional or partial EMT-like state, as described in the works above. While we cannot definitively assign a partial EMT identity from transcriptomic data alone,

the elevated basal markers, the enrichment of the EMT signature, and the observed enhanced dissemination capacity align with the concept that hybrid EMT states confer specific metastatic advantages, rather than a complete EMT program. We now discuss this more explicitly in the revised manuscript.

l) *Lines 586-589: Increased interactions between iCAFs and cancer cells after treatment was previously described and should be cited (PMID: 39227743).*

As noted in our response to an earlier comment (Lines 491–492), we have now cited PMID: 39227743 in the revised manuscript to acknowledge prior work describing iCAF enrichment following chemotherapy.

m) *Lines 594-606: The link between Myc amplification, interaction with TAMs, and metastatic progression in PDAC was previously shown and should be cited in this discussion (PMID: 34551968).*

Thank you for this suggestion. We had already cited PMID: 34551968 in the context of MYC-driven metastatic behavior, and we now also explicitly highlight its findings regarding MYC-dependent engagement of tumor-associated macrophages in the revised Discussion.

Minor Comments

- *Throughout the manuscript, the authors refer to individual tumors as examples. This reviewer was unable to find a table or list in the manuscript that includes the treatment status of each tumor, which is important for interpretation of the results.*

We agree and have now included a supplementary table listing treatment status, sample source (biopsy or resection, stage, and other relevant metadata for all tumors (Suppl. Table 2).

- *While there were 62 primary PDAC tumor samples from 59 patients, only 24 specimens were obtained from surgical resection and the rest were obtained by endoscopic biopsy with limited tissue. This is a limitation that should be discussed given the potential for significant sampling bias.*

To assess potential sampling bias, we compared biopsy and resection samples and found that although resections expectedly yield a higher number of recovered tumor cells, there were no systematic differences in number of detected clones (Suppl. Figure 4a,b). The

majority of our samples (38 of 62 specimens) were indeed obtained via biopsy, reflecting the reality that many patients with pancreatic cancer do not undergo surgical resection and that endoscopic biopsies represent the most common and clinically relevant route for tissue acquisition. We therefore believe that including biopsy-derived samples is essential for capturing the biological diversity present across patient presentations and for ensuring that analytical approaches developed here generalize to the types of specimens most frequently obtained in clinical practice. We have now included a brief discussion of this potential limitation in the revised manuscript (Discussion, second paragraph).

- *Line 210: “coexistence of both classical (C01) and basal (C03, C06)” => should C06 actually be C05?*

Thank you for catching this typo which has been corrected in the revised text.

- *Line 216: “Furthermore, we observed the concurrent presence of classical and basal subtypes in many tumors, challenging the traditional binary classification of PDAC.” => there are many studies in the literature already demonstrating this, which should be cited as this is not a new finding.*

We agree and have now revised this sentence, and also cited relevant studies (e.g., PMID: 36781852, 40269162, 31273297) in the revised sentence to acknowledge prior observations of coexisting subtypes in PDAC .

- *Line 503: Was “clone spC03” confirmed by a pathologist as malignant/invasive for sure? Histologically, it looks like it could represent a preinvasive neoplastic lesion.*

This is a fair observation, and we agree that, based on H&E morphology alone, spC03 (now renamed spC) could plausibly represent either an invasive clone or a high-grade preinvasive lesion such as PanIN-3. We had previously consulted with a pathologist, who noted this ambiguity in the histological image. However, spC harbors multiple, complex CNVs consistent with late-arising malignant clones observed elsewhere in the tumor. Given its genomic profile and its similarity to other invasive clones in the same sample, we interpret spC as likely representing a malignant population. We have now clarified this point in the revised manuscript (Section “Subpopulations within the tumor epithelial compartment reveal a dynamic process of clonal evolution”, second to last paragraph), noting both the histological ambiguity and the supporting genomic evidence.

- *Lines 506-507: The parental clone is classified as “neither” so this statement should be corrected.*

We have revised the sentence to accurately reflect the classification of the parental clone in the scRNA-seq data.

- *Figure 2c-e: Using the same blue color for C01 (which is mixed) as classical is confusing; similarly using the same tan color for C04 (which is classical) as mixed is confusing.*

We’ve updated the color scheme for the subtypes throughout the manuscript to ensure clearer and more consistent representation of clone identities and subtypes.

- *Figure 3c: “Columns are clusters and rows are different pathways.” This appears to be flipped.*

Thank you – we’ve corrected the figure legend accordingly.

- *Figure 4: This seems more suitable as a supplementary figure.*

We appreciate the suggestion, but we believe Figure 6 (previously Fig. 4) presents a core aspect of the manuscript—namely, the relationship between clonal phylogenies and transcriptional subtypes across the cohort. It synthesizes key findings and helps convey the heterogeneity and complexity of clonal evolution in PDAC. For this reason, we have kept it in the main figures but have clarified the layout and labeling to improve clarity. We are open to further adjustments per reviewer’s input on our thought process regarding this point.

- *Figure 6a: Luminal cellular debris was analyzed as “tumor” in the large gland involving spC03. How was this handled analytically?*

We thank the reviewer for catching this. In our analysis pipeline, luminal cellular debris had been excluded prior to downstream annotation and clone assignment. However, the version of the figure included in the initial submission did not reflect this filtering step. We have now corrected Figure 3a (Previously Fig. 6a) to accurately display the curated data.

Dear Dr. Brandt,

We thank you and the reviewers for your thoughtful and constructive feedback on our revised manuscript, "*Clonal Heterogeneity in Human Pancreatic Ductal Adenocarcinoma and Its Impact on Tumor Progression*" and for the positive assessment of our work. We are grateful that the reviewers find that the manuscript has been substantially improved and are pleased that it is considered acceptable in principle for publication in Nature Genetics, pending minor revisions. We also consent to the publication of the reviewer comments, author rebuttal letters and editorial decision letters as a Supplementary item.

Below we provide a brief point-by-point response to the final referee comments.

Reviewer #2

Comment: In my opinion the authors have thoroughly responded to my critiques as well as those of the other reviewers. The manuscript is substantially improved with new analyses, experimental validation, and methods. The authors were forthright in admitting limitations of their data, and demonstrated a strong effort to address our critiques and revise the manuscript as warranted. My recommendation is to accept the manuscript for publication.

Response: We thank Reviewer #2 for the careful assessment of our manuscript throughout the review process, for the constructive critiques that helped improve the study, and for the positive final evaluation. We appreciate the reviewer's recognition of the additional analyses, validation experiments, and clarifications of the study's limitations.

Reviewer #3

Comment: The authors have adequately addressed my critiques and suggestions.

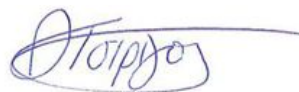
Response: We thank Reviewer #3 for the thoughtful evaluation of our manuscript and for the constructive suggestions, which helped strengthen the revised study. We appreciate the reviewer's positive assessment that the concerns have been adequately addressed.

We address the remaining editorial and formatting requirements in the author guidance file and provide all requested files in editable format.

Sincerely,



Diane Simeone, MD
Director, Moores Cancer Center
Chugai Pharmaceutical Chair in Cancer
Professor of Surgery
UC San Diego Health



Aristotelis Tsirigos, PhD
Professor of Medicine
Professor of Pathology
Co-director, Precision Medicine Division
NYU Grossman School of Medicine