nature portfolio

# Peer Review File

## Convergent alterations in the tumor microenvironment of MYC-driven human and murine prostate cancer

Reviewers' Comments:

Reviewer #1:
Remarks to the Author:
Summary:
The authors describe changes in the issue microenvironment in prostate cancer by performing single-cell RNA-sequencing (scRNA-seq) and pathology of the comparative biology between human prostate cancer using a genetically engineered mouse model (GEMM) of prostate cancer. They conclude that MYC signaling was the top up-regulated pathway in human cancers. They found that many non-malignant cell states in the tumor microenvironment (TME), including epithelial, immune, and fibroblast cell compartments, were conserved across individuals. Then, using a GEMM of prostate epithelial cell-specific MYC activation in two mouse strains, they found that MYC oncogene-expressing neoplastic cells, directly and indirectly, reprogrammed the TME during carcinogenesis. They described that the progression from a precursor-enriched to invasive-cancer-enriched state was accompanied by a cell-intrinsic switch from pro- immunogenic to immunosuppressive transcriptional programs with coinciding enrichment of immunosuppressive myeloid and Treg cells in the immune microenvironment. They conclude that MYC signaling reshapes TME of prostate cancer.
Overall this is a very clearly presented, very well organized, highly detailed descriptive study and the data and observations are worthy of an eventual report. However, as presented the findings do not contain any mechanistic or causal experimentation. Further, the findings as described are largely not novel, since they conclude what multiple other investigators have described that oncogenes like MYC causally regulate the TME and this plays a causal role in tumorigenesis and has implications for therapeutic responses. Pointedly, the effectors and molecules they described have largely if not entirely already been described, as regulated by MYC and many other reports do perform the mechanistic experiments, to demonstrate a causal role for tumor progression, and/or therapeutic response. Hence, this study as written is incomplete. A revised study with mechanistic and causal experimentation and a better effort to integrate what is common and uniquely found by these investigators with the existing literature would be an excellent report. Other issues that should be addressed are detailed below
Major Issues:
1. The authors findings are entirely descriptive, albeit well presented and interesting, but there is no direct causal demonstration of a role of any specific molecular or cellular effector. However, in the literature, numerous investigators have already shown that oncogenes including MYC play a direct and indirect role in shaping the TME and that this can play a role in tumorigenesis, immune evasion and immune sensitivity. This has been seen across many different kinds of cancer, different tumor models, human cancers. For this study to make a significant contribution to the field, they would have to consider these findings and explain how these findings uncover any mechanistically or clinically relevant finding.
2. A major strength of the study is the use of comparative human and mouse analysis. However, a potential weakness is that they presume there is a simple MYC gene signature. There are multiple "signatures" in the literature, and generally it appears that what MYC regulates is contextual. By performing an analysis where they assume that MYC has a simple single defined signature they create a circumstance where their findings could be in a sense tautological. They choose one way to describe how MYC regulates a single set of genes and then say they see this, and therefore MYC is the driver. However, the common set of genes that have been described are more likely reflective of genes that are generally associated with cancer, rather than the exclusive genes regulated only by MYC. Hence, concluding that MYC is always involved may be true, but I do not think they should assume they have demonstrated this by simply starting with one MYC signature.
3. The authors say that they have described evaluated relevant times points to interrogate the progression of changes that occur with respect to the TME. However, human prostate cancer appears to occur over decades and a challenge of the GEMM models is that they occur over a much shorter time frame. Generally, mouse models would appear to best reflect the time course and changes of relatively highly aggressive cancers. The authors should consider this.
4. The authors are somewhat selective about literature they consider. There are numerous publications

that have interrogated how the MYC oncogene may regulate the TME and it would be useful for them to consider to what extent what they found has been already described versus unique and to focus on some integration of these previous reports with their own findings, then focus on what they have found conceptually and mechanistically is novel.

Reviewer #2:
Remarks to the Author:
Graham et al perform a single cell study on primary human prostate tumors and mouse prostates from 2 different Myc overexpressors. The manuscript is expertly written and the data a solid addition to the field. Clearly much thought and work went into the project. I only have minor suggestions. Please note that as a key person responsible for nomenclature, my comments
are meant to help the entire field get on the same page even though it might seem like nitpicking. Hopefully these suggestions won't take too much time as they are mostly computational.
1. The labeling of the club cell cluster as 'intermediate/atrophic' cells is problematic. Your own recent paper labeled PIA as club (like) cells. I think it's confusing for the field to rename a cell 'type' according to a histopathological phenotype. Please stick with club or club-like. I also think it's wrong to continue to conflating the PIA phenotype with actual cancer insinuating it is a precursor when your own work shows that it is tumor adjacent. Surely if it were a precursor it would be possible to detect CNV in club (like) cells. Is it not quite possible that the club-like cells within the same gland coexist with tumor cells without being clones (and are maybe even induced to that state by tumor cells)? We have enough evidence to show that the luminal to club plasticity is a normal response to androgen depletion and not restricted to cancer. I'm actually quite surprised there isn't any evidence yet that club cells are precursors or harbor mutations (it would make sense), but we should be careful about insinuating they are precursors until there is quality evidence (trajectory analysis doesn't count!). This should be explicitly stated in the manuscript so there is no confusion.

2. This is bound to be a highly valuable dataset for the field so I would ask for due diligence on cell type annotation. Please identify club (not 'atrophic'), hillock, ionocyte, and neuroendocrine cells in Figure 2 as well as periprostatic/subglandular (APOD+) and interstitial fibroblasts (C7+) (Figure 7D). The full transcriptomic identification of these cell types can be found here: PMID: 34173975. Of note, we are now labeling the subglandular fibroblasts as 'periprostatic' because they can be distinguished from subglandular fibroblasts from other organs. Remember the ultimate goal for the field of science is to create nomenclature that can distinguish these cells in multi-organ datasets. This is why we can't just say 'luminal', we have to say 'prostate luminal' or 'urethral luminal' when multiple tissues are represented in a dataset. Additionally, I think it's important to show the wild type and Myc annotated datasets side-by-side (using the split.by function) and not overlapped by color even though it takes up more space in the figure. People need to see whether the induced clusters (Ly6d basal, psca luminal and timp fibs) are present in the wild type and therefore represent an expansion of a preexisting cell type vs. a change in cell state. Even if it needs to be placed in a supplementary figure, it has to be visible for the reader somewhere. On that note, it is necessary to show each of these markers used to annotate a myc-induced cell state in situ in the wild type lower urinary tract. It is known that Psca labels the entire urothelium. Ly6d labels urethral urothelium. I can see Timp1 expression in normal fibroblasts from our own data set, but where are they located in the wild type lower urinary tract? Is your Timp1+ fib an expansion of a preexisting, specially localized fibroblast or an induction of subglandular or interstitial fibroblasts? Timp1 staining is warranted in wild-type and myc lower urinary tracts to answer this for the reader. If the answer is that you can't identify a preexisting Timp1+ fibroblast, then that is still informative. Of note, what you are calling 'interstitial' fibs in mouse prostate is what we called 'prostate' fibs and what you are calling 'subglandular' is what we identify as 'ductal' based on staining of the entire lower urinary tract. Clearly, the field still needs to come to a consensus on nomenclature, but I feel the term subglandular is confusing for a fibroblast that isn't surrounding prostate glands, but rather the proximal ducts. It's your call, but since you cited our work on this, I thought I'd point it out.

3. It is stated that Ly6d is a urothelial marker, but the urethra wasn't included in the dissection so they can't be urothelial cells. Here is where things become semantic and confusing and I feel like we have t o be careful not to be too prostate-centric. First, I would be highly surprised if there weren't at least a f ew Krt4+/Psca+ urethral luminal cells in wild type prostate as they project into the proximal prostate d ucts and are also found in distal tips. Please check closer and annotate if so. Second, we have found th at differentiating between a preexisting urethral luminal cell and a preexisting
prostate luminal cell that has undergone lineage plasticity to resemble a club (human) or urethral luminal (mouse) cell due to low androgen is quite difficult. Whether it's Myc o/e, Pten ko or castration/ADT/5ARI, this induced urethral lineage identity looks almost identical to preexisting urethral cells. It is therefore quite important to identify these cells in the wildtype data for comparison to the Myc-induced cell profile as others have done. It is also important to note that these Krt4+/Psca+ urethral urothelial cells are NOT a preexisting pool of proximal progenitors as has been definitively shown in multiple lineage tracing studies at this point. I do applaud your efforts to differentiate between this preexisting cell type and the induced cell state based on Psca expression, but as a single marker this is not enough and could be accomplished computationally by comparing to the preexisting cell type. If they can't be identified, you could use our dataset of urethra as a reference.

I hope these concerns are construed as constructive and don't delay too much.

Reviewer #3:
Remarks to the Author:
Most efforts in PCa research are focussed on understanding mCRPC and NEPC, but basic research in the aetiology of prostate cancer is required to better understand how benign lesions progress to adenocarcinoma and locally invasive disease. The importance of the microenvironment and paracrine signalling in supporting PCa progression is well known, but mechanistically the specific contributions remain poorly defined. In this article Graham and colleagues performed scSeq in a cohort of localised PCa patients. Their analysis revealed that although significant heterogeneity in clustering of luminal cells can be observed inter-patients, the heterogeneity of the TME between patients is less striking.

Using the Hi-MYC mouse in 2 strains they performed clustering analysis and identified two epithelial clusters enriched for MYC. These clusters MYC1 and MYC2 differ in their enrichment between lobes. This analysis also led to the hypothesis that these MYC1/2 cells are responsible for the formation of the basal and luminal Ly6d clusters which are not enriched in MYC expression.

They then perform an elegant analysis of potential ligand-receptor enrichment between the identified clusters and basal Ly6d cells. This revealed a potential key role of interferon signalling in the paracrine interactions between MYC cells and Ly6d cluster. Their results also provides a compelling model for the switch to an immune cold phenotype in localised PCa.

My main concern relates to the exclusive use of the Hi-MYC model. For example, validating some of their key findings using the HoxB13 MYC mouse, if tissue sections are readily available, could also demonstrate that the Ly6d basal cell cluster is not Hi-MYC mouse specific. Considering that Bieberich and DeMarzo who characterised the HoxB13-MYC mouse are co-author on this manuscript it might be easy to source relevant tissues for analysis. The probasin promoter used to generate the Hi-MYC mouse can be leaky as reported by Bieberich and colleagues (PMC5006678):
“Another potential advantage of the Hoxb13 regulatory elements for driving prostate expression of MYC and loss of Pten is that we have not observed any leakiness of expression of Hoxb13-based reporter genes, MYC overexpression or Pten loss in any of these animals, whereas leakiness of the rat probasin promoter into stromal elements has been observed (58)”

Minor:
As published in an addendum to the original Hi-MYC article from 2003, the MYC protein expressed in this transgene contains additional C-terminal amino acids. So it is possible that some of the observed effects are related to this. Probably a simple mention in the discussion that the MYC expressed in this model may not behave exactly like WT Myc would be sufficient.

REVIEWER COMMENTS

**Reviewer #1,** expertise in MYC signaling and the TME (Remarks to the Author):

Summary:
The authors describe changes in the issue microenvironment in prostate cancer by performing single-cell RNA-sequencing (scRNA-seq) and pathology of the comparative biology between human prostate cancer using a genetically engineered mouse model (GEMM) of prostate cancer. They conclude that MYC signaling was the top up-regulated pathway in human cancers. They found that many non-malignant cell states in the tumor microenvironment (TME), including epithelial, immune, and fibroblast cell compartments, were conserved across individuals. Then, using a GEMM of prostate epithelial cell-specific MYC activation in two mouse strains, they found that MYC oncogene-expressing neoplastic cells, directly and indirectly, reprogrammed the TME during carcinogenesis. They described that the progression from a precursor-enriched to invasive-cancer-enriched state was accompanied by a cell-intrinsic switch from pro-immunogenic to immunosuppressive transcriptional programs with coinciding enrichment of immunosuppressive myeloid and Treg cells in the immune microenvironment. They conclude that MYC signaling reshapes TME of prostate cancer.
Overall this is a very clearly presented, very well organized, highly detailed descriptive study and the data and observations are worthy of an eventual report. However, as presented the findings do not contain any mechanistic or causal experimentation. Further, the findings as described are largely not novel, since they conclude what multiple other investigators have described that oncogenes like MYC causally regulate the TME and this plays a causal role in tumorigenesis and has implications for therapeutic responses. Pointedly, the effectors and molecules they described have largely if not entirely already been described, as regulated by MYC and many other reports do perform the mechanistic experiments, to demonstrate a causal role for tumor progression, and/or therapeutic response. Hence, this study as written is incomplete. A revised study with mechanistic and causal experimentation and a better effort to integrate what is common and uniquely found by these investigators with the existing literature would be an excellent report. Other issues that should be addressed are detailed below

Major Issues:
1. The authors findings are entirely descriptive, albeit well presented and interesting, but there is no direct causal demonstration of a role of any specific molecular or cellular effector. However, in the literature, numerous investigators have already shown that oncogenes including MYC play a direct and indirect role in shaping the TME and that this can play a role in tumorigenesis, immune evasion and immune sensitivity. This has been seen across many different kinds of cancer, different tumor models, human cancers. For this study to make a significant contribution to the field, they would have to consider these findings and explain how these findings uncover any mechanistically or clinically relevant finding.

*We agree with the reviewer that the role of MYC in shaping the TME of human cancers has been richly studied in multiple cancer types. In our revised manuscript, we have taken care to*

*better contextualize and highlight the aspects of novelty in our manuscript with respect to this rich prior literature.*

*One key aspect of novelty of our study is the finding that MYC-induced changes to the prostate TME significantly differ from the precursor to invasive carcinoma stages, and to our knowledge, this is the first scRNA-seq study to report a MYC-driven cell intrinsic pro-inflammatory to immunosuppressive switch in both prostate cancer patient tissue samples and MYC-driven models of prostate cancer. As noted by the reviewer, various studies in other tumor types, such as liver, breast, and bladder, have reported that cancer cell activation of MYC can indirectly modulate cytokine expression to be immunosuppressive (PMID: 36966168, 31965581, 34251762, 31186238, 32716585). MYC has also been found to directly modulate expression of checkpoint molecules to facilitate immune evasion (PMID: 36525476, 26966191, 36897988, 34321275; see more detailed description below in major issue #4).*

*However, our scRNA-seq prostate cancer studies of the murine MYC-driven model of prostate cancer (Hi-Myc) show that initially, MYC activation is not immunosuppressive but pro-inflammatory in prostate luminal cells. Persistent MYC activation during tumor progression drives a cell-intrinsic switch from a pro-inflammatory precursor state to immunosuppressive programs associated with the downregulation of tumor suppressor pathways during progression to carcinoma.*

*Prompted by the reviewer's comments, we further interrogated the downstream MYC targets in our human prostatectomy scRNA-seq data to examine which MYC targets reflect this inflammatory switch. In this expanded analysis of several MYC gene signatures, we again saw that many of the MYC targets were upregulated in both the cancer and the atrophy/intermediate clusters. Notably, this atrophy cluster includes luminal cells of proliferative inflammatory atrophy (PIA), a previously described prostate cancer precursor reported by co-author Angelo De Marzo (PMID: 10595928) and known to show the earliest evidence of irregular MYC upregulation in luminal cells (PMID: 18567993). Consistent with MYC activation in precursor and carcinoma, MYC targets were upregulated in both the cancer and atrophy clusters. However, there was a distinct difference in the expression of genes known to be downregulated by MYC. In particular, we observed downregulated MYC targets in the cancer cluster but not in the atrophy cluster (Supplementary Figure S8A&B). Furthermore, these downregulated MYC targets significantly overlapped with tumor suppressor/gatekeeper pathways P53 and apoptosis and the pro-inflammatory TNFA signaling via NFKB pathway (Supplementary Figure S8C&D). Likewise, examining the mouse data, we found that the MYC downregulated genes in the luminal MYC cells of invasive carcinoma (Hi-Myc 10 months) compared to the precursor cells (Hi-Myc 6 months) also had significant overlap with apoptosis, P53, and TNFA signaling via NFKB (Supplementary Figure S21A&B).*

*Additionally, we observed that the switch from cell-intrinsic pro-inflammatory to immunosuppressive switch occurred over a period of latency between the 6mo precursor enriched state to the 10mo invasive cancer enriched state, and was accompanied by downregulation of key tumor suppressor and gatekeeper pathways, including PTEN, p53, and DNA repair, all of which are highly relevant for human prostate carcinogenesis. These data suggest that initial MYC activation leads to pro-neoplastic pathway activation and the engagement of tumor suppressive and pro-inflammatory pathways. The progression to invasive cancer involves a period of latency, with downregulation of the tumor suppressor and*

*inflammatory pathways. These observations led to the hypothesis that the forced loss of these tumor suppressor pathways could greatly accelerate the switch from pro-inflammatory to immunosuppressive pathways. To test this hypothesis directly, in the revised manuscript, we showed that conditional, forced disruption of the PTEN tumor suppressor, combined with conditional MYC activation in the mouse prostate (BMPC mice), greatly accelerated the pro-inflammatory to immunosuppressive switch compared to MYC activation alone (Supplemental Figure S21I-L, [updated manuscript lines 423-456]). This switch was also associated with the known acceleration of progressing from precursor to invasive and metastatic cancer in the BMPC animals.*

*Collectively, these additional findings emphasize that MYC targets are differentially regulated during the malignant transformation of prostate luminal cells from precursor to carcinoma and that the switch from immunogenic to immunosuppressive programs is likely driven by progressive downregulation of tumor suppressive pathways on top of MYC activation. This is a critically important observation, as a previously published scRNA-seq data of Hi-Myc mice examining transcriptional changes in luminal cells was performed in 3-month-old mice (PMID: 35562350). At 3 months, this is well before MYC-activated luminal cells undergo progression from neoplastic precursors to fully invasive adenocarcinoma. Consequently, these critical cell-intrinsic changes in luminal cells during carcinogenesis were missed.*

*Based on the reviewer's feedback, we have updated the manuscript to include this additional analysis in the context of previously published findings, highlighting known MYC-induced immunosuppressive changes reported in various tumor types and emphasizing the novelty of our findings in prostate cancer [updated manuscript lines 122 - 129, 236 - 257, 358 - 374, 423 - 456, 503 - 507, 524 - 544].*

*Regarding clinical relevance, in the discussion section, we highlight the potential utility of TREM2 antibodies, as there are ongoing clinical trials, and now also highlighted the ongoing pre-clinical and clinical trials targeting MYC in this revised version.*

2. A major strength of the study is the use of comparative human and mouse analysis. However, a potential weakness is that they presume there is a simple MYC gene signature. There are multiple “signatures” in the literature, and generally it appears that what MYC regulates is contextual. By performing an analysis where they assume that MYC has a simple single defined signature they create a circumstance where their findings could be in a sense tautological. They choose one way to describe how MYC regulates a single set of genes and then say they see this, and therefore MYC is the driver. However, the common set of genes that have been described are more likely reflective of genes that are generally associated with cancer, rather than the exclusive genes regulated only by MYC. Hence, concluding that MYC is always involved may be true, but I do not think they should assume they have demonstrated this by simply starting with one MYC signature.

*We agree with the reviewer that the role of MYC is context-dependent. Given that multiple signatures have been reported downstream of MYC activity in various contexts, we have expanded our analysis beyond the three MYC signatures initially included in the analysis of our prostatectomy dataset (Hallmark MYC targets V1, Hallmark MYC targets V2, and Chi Dang targets) seen in Figure 2 and Supplementary Figure S7 and the publicly available TCGA primary*

*prostate cancer RNAseq dataset in Supplementary Figure S6. Our revised manuscript now includes 20 MYC signatures/gene sets, which include up and down-regulated MYC targets from previously published studies. These findings are included in Supplementary Figure S8. As described above in response to Major Issue #1, this expanded MYC gene signature analysis was instrumental in revealing additional insights regarding the MYC-driven cell-intrinsic switch from inflammatory to immunosuppressive during prostate cancer carcinogenesis. Changes to the manuscript's text reflecting these additional studies and findings can be found in the results section [updated manuscript lines 236 - 257, 423 - 456].*

*Furthermore, the use of the genetically engineered mouse models where MYC activation alone showed strong parallels and revealed longitudinal dynamics of the activation and consequences of MYC activation to human prostate cancer provided important mechanistic evidence that the observed alterations in MYC signature were driven by MYC activation, and not just a generic consequence of carcinogenesis.*

3. The authors say that they have described evaluated relevant times points to interrogate the progression of changes that occur with respect to the TME. However, human prostate cancer appears to occur over decades and a challenge of the GEMM models is that they occur over a much shorter time frame. Generally, mouse models would appear to best reflect the time course and changes of relatively highly aggressive cancers. The authors should consider this.

*As with most mouse models of cancer, the timeline is often accelerated compared to the naturally occurring progression in humans. In our study, we strategically selected prostate cancer progression by stage (normal, precursor, and primary disease), which we refer to as time points. We agree this is an important consideration and we have now revised the text to clarify this in the description of the mouse model [updated manuscript lines 343 - 346]. In the Hi-Myc model, while the timeline of prostate cancer progression is compressed compared to humans, this model does not progress beyond organ-confined disease, emulating a non-aggressive phenotype. Prompted by the reviewer's comments, we performed additional pathway analysis in human and mouse data, revealing that in addition to the down-regulation of gatekeeper pathways such as apoptosis and P53, PTEN pathway downregulation and mTOR activation coincided with the cell-intrinsic switch of MYC target genes from precursor to invasive carcinoma (Supplemental Figures 8C-F and 21A-D). Notably, PTEN, a well-established negative regulator PI3K/AKT/mTOR signaling is lost by biallelic inactivation in ~20% of primary prostate cancer and half of aggressive castration-resistant prostate cancer (PMID: 29460925). Collectively, these data suggested that the loss of a negative regulator of PI3K/AKT/mTOR signaling would accelerate the MYC-driven alterations in the prostate TME and promote a disease phenotype more representative of aggressive prostate cancer. To that end, we employed the BMPC prostate cancer mouse model developed by co-authors De Marzo and Biebrich (PMID: 26554830). In this model, MYC activation and PTEN loss are driven by the prostate luminal-specific HoxB13 regulatory locus, with extensive PIN-like lesions at 3 months, progressing to metastatic disease at 6 months. We generated scRNA-seq libraries from the prostate (3 months) and metastatic lesion (6 months) of BMPC mice and integrated these data with our WT and Hi-Myc datasets (Supplemental Figure 21E-H). The TME of the 3-month BMPC mouse prostate resembled the TME of the Hi-Myc 6-month prostate, including Ly6d-*

*expressing basal cells, Trem2 macrophages, and Timp1 fibroblasts (Supplemental Figure S21I-K). Likewise, the cell-intrinsic pro-inflammatory to immunosuppressive switch observed in Hi-Myc (6 months vs 10 months) occurred in a much-accelerated timeline in the BMPC mice (3 months vs 6 months) (Supplemental Figure 21L), representing a more aggressive phenotype.*

4. The authors are somewhat selective about literature they consider. There are numerous publications that have interrogated how the MYC oncogene may regulate the TME and it would be useful for them to consider to what extent what they found has been already described versus unique and to focus on some integration of these previous reports with their own findings, then focus on what they have found conceptually and mechanistically is novel.

*We thank the reviewer for these constructive comments. We have now developed a more complete contextualization of the prior literature and improved the focus on the novel aspects of the current manuscript with respect to this prior literature. Previously published studies in various tumor models have shown that cell-intrinsic MYC activation in cancer cells can suppress immune cell infiltration. For example, in a liver cancer model, MYC activation suppressed MHCI-mediated antigen presentation in tumor cells, while inactivating MYC would drive macrophage and T cell recruitment to the TME (PMID: 36525476). In leukemia models, MYC promoted immune evasion in cancer by upregulating CD47 and PD-L1 (PMID: 26966191) and a glycoimmune checkpoint ligand (PMID: 36897988). In a breast cancer model, MYC was reported to suppress the cGAS-STING pathway, resulting in low immune infiltration in the TME (PMID: 34321275). As one important novel aspect of our current manuscript, our scRNA-seq analysis of the murine MYC-driven model of prostate cancer (Hi-Myc) shows that initially, MYC activation is not immunosuppressive but rather pro-inflammatory in prostate luminal cells with significant infiltration of macrophages and activated T cells to the TME (Supplemental Figures S15). How MYC ultimately drives an immunosuppressive TME in prostate cancer is not well understood. Here, we show that persistent MYC activation during tumor progression, accompanied by a downregulation of TNFA signaling and tumor suppressor pathways such as apoptosis, P53, and PTEN (or mTOR activation) resulted in a cell-intrinsic switch from a pro-inflammatory precursor state to immunosuppressive carcinoma, as detailed above in our responses to Major Issues 1-3.*

**Reviewer #2,** expertise in scRNA-seq for prostate cancer (Remarks to the Author):

Graham et al perform a single cell study on primary human prostate tumors and mouse prostates from 2 different Myc overexpressors. The manuscript is expertly written and the data a solid addition to the field. Clearly much thought and work went into the project. I only have minor suggestions. Please note that as a key person responsible for nomenclature, my comments are meant to help the entire field get on the same page even though it
might seem like nitpicking. Hopefully these suggestions won't take too much time as they are mostly computational.

*We thank the reviewer for highlighting the strengths of this manuscript and its importance to the field. We agree wholeheartedly that there is great value in developing consensus in the field*

*with regards to a common nomenclature around identified cell types/states for organs and systems. We thank the reviewers and others for their leadership and contributions in advancing t his type of effort for the field. Thanks to the reviewer's constructive feedback, where we saw clear agreement with gene expression signatures for cell types/states that aligned with previously defined nomenclature, we have now revised the nomenclature to reflect that. However, in a few instances, our data have shown that there may be more nuance needed in applying the appropriate nomenclature and the suggested terms would not be fully consistent with the states that we observed. These determinations were assessed with great depth and rigor to make sure we follow the spirit of this reviewer's constructive criticisms. In those instances, we have indicated why we think the nomenclature we previously used is more appropriate, both based on data that we present in this revised manuscript, as well as prior data from two and a half decades of studying prostate cancer cell types/states using molecular pathologic, genetic, and epigenetic studies in our collaborative team. Additionally, we thank the reviewer for disclosing his identity, and hope the reviewer will be open to working together with us and others in the field to develop a highly informed and nuanced consensus on the few cell types/states that could not easily be conformed to terminologies suggested by this reviewer.*

1. The labeling of the club cell cluster as 'intermediate/atrophic' cells is problematic. Your own recent paper labeled PIA as club (like) cells. I think it's confusing for the field to rename a cell 'type' according to a histopathological phenotype. Please stick with club or club-like. I also think it's wrong to continue to conflating the PIA phenotype with actual cancer insinuating it is a precursor when your own work shows that it is tumor adjacent. Surely if it were a precursor it would be possible to detect CNV in club (like) cells. Is it not quite possible that the club-like cells within the same gland coexist with tumor cells without being clones (and are maybe even induced to that state by tumor cells)? We have enough evidence to show that the luminal to club plasticity is a normal response to androgen depletion and not restricted to cancer. I'm actually quite surprised there isn't any evidence yet that club cells are precursors or harbor mutations (it would make sense), but we should be careful about insinuating they are precursors until there is quality evidence (trajectory analysis doesn't count!). This should be explicitly stated in the manuscript so there is no confusion.

*We agree with the reviewer that the goal of developing consistent nomenclature for cell types in t he prostate is worthwhile to pursue, and recognize significant efforts to generate single-cell atlases and share this critical resource with the field. However, we are concerned that using the label of club or club-like to describe broadly the cluster of cells that we classified as atrophy/intermediate has multiple major issues. Furthermore, we would like to clarify multiple of the comments made by the reviewer with respect to PIA and its relationship with cancer, and appreciate the opportunity presented by the reviewer to make sure our discussion of these cell types and the PIA lesions reflect what we have uncovered about these lesions over the last 20+ years of research in studying them across our team using morphological and molecular genetic and epigenetic studies since the initial description of PIA lesions led by the De marzo lab in 1999 (PMID: 10595928, 11068311, 11550208, 11795433, 12707036, 12937133, 14607218, 14755684, 15880529, 17384581, 18302221, 19574101,*

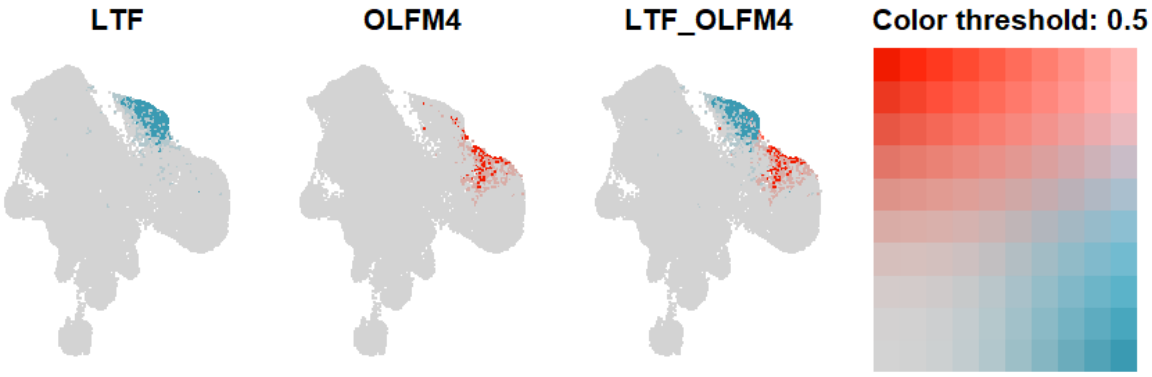
*22212087, 29422309, 30900284, 34341114, 37347938, 37550827). We have organized our response to this reviewer comment in two major themes to address these issues:*

a. *Clarifications about PIA and its relationship to PIN and invasive cancer*

*First, we emphasize that our intention was not to imply that atrophy always leads to cancer or is a cancer-exclusive process, but rather that the inflammatory conditions that give rise to a subtype of atrophy called proliferative inflammatory atrophy (PIA) can form a hotbed of proliferative epithelial cells that can give rise to neoplastic cells (PMID: 29089606, 34341114, 10595928, 26813971). While PIA does not always lead to cancer, extensive literature exists supporting the role of PIA as a precursor of prostate cancer, given that PIA often merges with the prostatic intraepithelial neoplasia, an established prostate cancer precursor, and even with invasive cancer lesions (PMID: 19507201, 10595928, 11068311). Furthermore, evidence of cancer-associated molecular alterations has been reported in PIA, including p53 alterations (PMID: 12210488) and gain of chromosome 8 centromeric DNA (PMID: 11337374). ERG rearrangement and activation, which occurs in roughly half of all prostate cancer cases (PMID: 25915839), has been reported in PIA (PMID: 34341114). A small fraction of PIA cells can harbor epigenetic DNA hypermethylation at GSTP1 and other genes leading to their silencing, similar to the majority of PIN and cancer lesions (PMID: 37347938, 30082453). Nonethless, the majority of PIA cells/lesions do not harbor genetic or epigenetic alterations seen in human cancers, but rather show a phenotype consistent with stress response and injury/regeneration – with evidence to support this at the histomorphological, and molecular genetic/epigenetic and gene/protein expression levels. The luminal epithelial cells in these PIA lesions appear atrophic at a cellular scale (instead of having a tall columnar cell appearance characteristic of totally benign prostate luminal epithelial cells, they have a stunted/atrophic appearance), and can express genes suggesting a state that can be considered intermediate between luminal and basal epithelial cell types (e.g. mix of typically basal and luminal cytokeratins) and often have strong induction of stress response genes (e.g. GSTP1, GSTA1, PTGS2, LTF). Collectively, this data supports the idea that the inflammatory conditions driving PIA can form the hotbed from which PIN and/or invasive cancer lesions can arise. We have updated our text to emphasize these points [updated manuscript lines 82 - 97].*

*Additionally, prostatic atrophy is not a homogenous population of epithelial cells but rather is comprised of many different variations with distinct and nuanced biology. Although rarely discussed, there are two major classes of prostate atrophy (Prostate Cancer: Biology, Genetics, and the New Therapeutics, 2nd Edition, Chapter 2). The first could be considered hormonal or diffuse atrophy, which is induced after androgen deprivation or , androgen receptor blockade. This, or a variation of it, was studied recently by Joseph et al. (PMID: 34928497). However, the second type, called focal atrophy, which includes PIA, is not known to be induced by androgen blockade but can occur at any age and is associated with inflammation. The consensus by leading experts in the field has recognized that focal prostatic atrophy is a broad categorization of various subtypes (PMID: 17001160). How gene expression modules in luminal intermediate cells in focal atrophy and luminal cells in hormonal atrophy compare has not been determined. While potentially sharing some transcriptional similarities, these atrophy subtypes are at least somewhat unique and more detailed studies to describe these differences are underway in our team.*

*In terms of focal atrophy and PIA, there are ongoing efforts in our research group to leverage scRNA-seq analysis to define these various atrophic/intermediate subtypes, as there appear to be transcriptionally unique subgroups. For example, OLFM4 and LTF, as part of broader gene expression modules, have distinct regional expression within the atrophy/intermediate cluster (see UMAPs below). We believe these transcriptionally distinct subgroups reflect distinct atrophy/intermediate cell states. While the extensive characterization of atrophy/intermediate cell subtypes is beyond the scope of this cancer-focused study, we plan on sharing these research findings in a future manuscript that includes a comprehensive analysis of all prostate zones (peripheral, central, and transition), with a thorough characterization of luminal, basal, and atrophy/intermediate cell types*




b. *Club-like is not synonymous with atrophy/intermediate cells*

*In their landmark study in 2018 published in Cell Reports, they used FACS and single cell and bulk RNA sequencing to develop a single-cell atlas of normal prostate (PMID: 30566875). In that study, they identified the presence of cells with very high transcriptional similarity to club cells in the lung, with a key defining feature that they expressed the secretoglobin genes SCGB1A1 and SCGB3A1, which are well known markers of lung club cells. While these cells were identified in cells isolated from combined central/transition zone prostate tissues, they were nearly absent in the peripheral zone in that study (PMID: 30566875, Figure 5B). Consistent with these previous findings, we also find that in our peripheral zone tissues, SC GB1A1 expression is very rare and SCGB1A1 or SCGB3A1 expressing cells were not identified in our peripheral zone atrophy/intermediate cell cluster (Supplemental Figure S1E). Thus, we feel it would be too presumptive to broadly equate the atrophy/intermediate cells in our peripheral zone tissues to club cells, for which the most well-defined markers are the SCGB1A1 and similar secretoglobin genes, which are absent in our atrophy/intermediate clusters.*

*While we sampled all zones of the prostate in our study, our single-cell analyses were performed exclusively in the peripheral zone, except for our inferred copy number variation analysis, where patient-matched tissues collected from the central and transition zones that were histologically evaluated as cancer-free were used as normal controls (Figure 2C, Supplemental Figures S3-5). To make sure we weren't simply missing detection of the secretoglobin genes due to some technical issue, we now went back to our scRNA-seq data from the central and transition zone tissues, and found that SCGB1A1 and SCGB3A1 genes are*

*indeed robustly expressed in the intermediate cell cluster in the transition zone, and largely absent in the peripheral zone (Supplementary Figure S1E), again consistent with the previous findings. Histological assessment of our peripheral zone tissues showed extensive focal atrophy consistent with PIA, suggesting that a substantial proportion of our epithelial cells include atrophy/intermediate cells from PIA.*

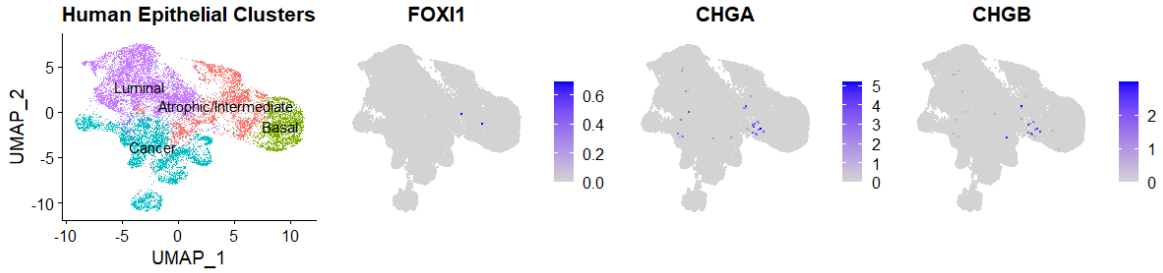
*Interestingly, although the atrophy/intermediate cells characterized in our study lack expression of the secretoglobin genes, they do have multiple genes that are also expressed in the club-like cells that were present in the transition zone and previously described. These include the OLFM4 and PIGR genes. Based on the expression of these genes in PIA lesions, in a recent study led by Franklin Huang and Karen Sfanos, in which Angelo De Marzo was also a co-author, it was suggested that the luminal cells in PIA lesions are club-like cells (PMID: 37550827). In another study, Franklin Huang's group termed these cells "cancer club cells" (PMID: 35013146), presumably due to their presence in peripheral zone tissues containing cancer lesions. However, in both of those studies, it was shown that secretoglobin expression was very rare to absent in these peripheral zone cells. In situ analysis of the LTF and PIGR expression revealed that these were indeed positive in atrophy/intermediate cells of PIA lesions, and that SCGB1A1 expression was very rare in these cells (PMID: 37550827). In our own data, we have seen that LTF (and associated genes such as PIGR) and OLFM4 (and associated genes) are expressed in distinct subsets of atrophy/intermediate cells. Based on these detailed analyses, triggered in large part because of the comments from this reviewer that led us to challenge our initial labels, we have now arrived at the conclusion that the atrophy/intermediate cells are a broad categorization composed of multiple subsets of transcriptionally distinct cell states that are recognizable due to the large number of epithelial cells evaluated in our study. Secretoglobin expressing club-like cells are possibly a rare subset of these atrophy/intermediate cells. In an ongoing study, we are thoroughly characterizing the various atrophy/intermediate cell states/subtypes and plan to develop a report describing these findings. In the meanwhile, we feel it is most appropriate to retain the umbrella term atrophy/intermediate cells to describe this population of cells.*

2. This is bound to be a highly valuable dataset for the field so I would ask for due diligence on cell type annotation. Please identify club (not 'atrophic'), hillock, ionocyte, and neuroendocrine cells in Figure 2 as well as periprostatic/subglandular (APOD+) and interstitial fibroblasts (C7+) (Figure 7D). The full transcriptomic identification of these cell types can be found here: PMID: 34173975. Of note, we are now labeling the subglandular fibroblasts as 'periprostatic' because they can be distinguished from subglandular fibroblasts from other organs. Remember the ultimate goal for the field of science is to create nomenclature that can distinguish these cells in multi-organ datasets. This is why we can't just say 'luminal', we have to say 'prostate luminal' or 'urethral luminal' when multiple tissues are represented in a dataset. Additionally, I think it's important to show the wild type and Myc annotated datasets side-by-side (using the split.by function) and not overlapped by color even though it takes up more space in the figure. People need to see whether the induced clusters (Ly6d basal, psca luminal and timp fibs) are present in the wild type and therefore represent an expansion of a preexisting cell type vs. a change in cell state. Even if it needs to be placed in a supplementary figure, it has to be visible for the reader somewhere. On that note, it is necessary to show each of these markers used to annotate a

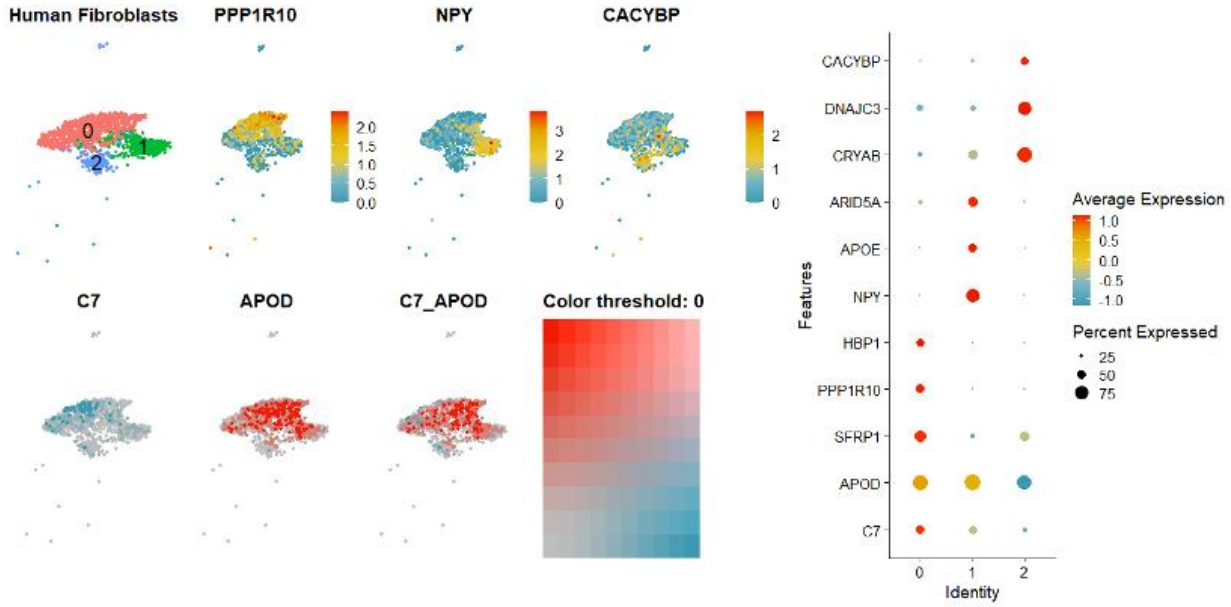
myc-induced cell state in situ in the wild type lower urinary tract. It is known that Psca labels the entire urothelium. Ly6d labels urethral urothelium. I can see Timp1 expression in normal fibroblasts from our own data set, but where are they located in the wild type lower urinary tract? Is your Timp1+ fib an expansion of a preexisting, specially localized fibroblast or an induction of subglandular or interstitial fibroblasts? Timp1 staining is warranted in wild-type and myc lower urinary tracts to answer this for the reader. If the answer is that you can't identify a preexisting Timp1+ fibroblast, then that is still informative. Of note, what you are calling 'interstitial' fibs in mouse prostate is what we called 'prostate' fibs and what you are calling 'subglandular' is what we identify as 'ductal' based on staining of the entire lower urinary tract. Clearly, the field still needs to come to a consensus on nomenclature, but I feel the term subglandular is confusing for a fibroblast that isn't surrounding prostate glands, but rather the proximal ducts. It's your call, but since you cited our work on this, I thought I'd point it out.

*In regards to the ionocytes, authors (Karthaus et al. 2020) who originally identified the Foxi1-expressing ionocyte-like luminal epithelial cluster in the mouse prostate noted that they did not identify a distinct human ionocyte cluster in their scRNA-seq data (PMID: 32355025). Consistent with this published finding, we observed only 2 luminal cells that expressed FOXI1 and consequently did not update Figure 2. We included the Foxi1 expressing luminal cluster in Figure 3 of mouse data and have updated the name of this cluster from Foxi1 to "Ionocytes".*

*To address the issue of neuroendocrine cells specifically, we would like to highlight that Karthaus et al. did not observe a distinct neuroendocrine cluster (PMID: 32355025). They attributed this to the rarity of the neuroendocrine cells. Likewise, Joseph et al. also did not observe a distinct neuroendocrine cluster, as neuroendocrine cells were also scattered in UMAP space (PMID: 30566875). In our data, neuroendocrine cells (CHGA+ CHGB+) appear rare and do not form a distinct cluster on our UMAP. This likely is due to being an extremely rare population in our dataset (21 cells). As a result, we did not update Figure 2 with a neuroendocrine cluster.*



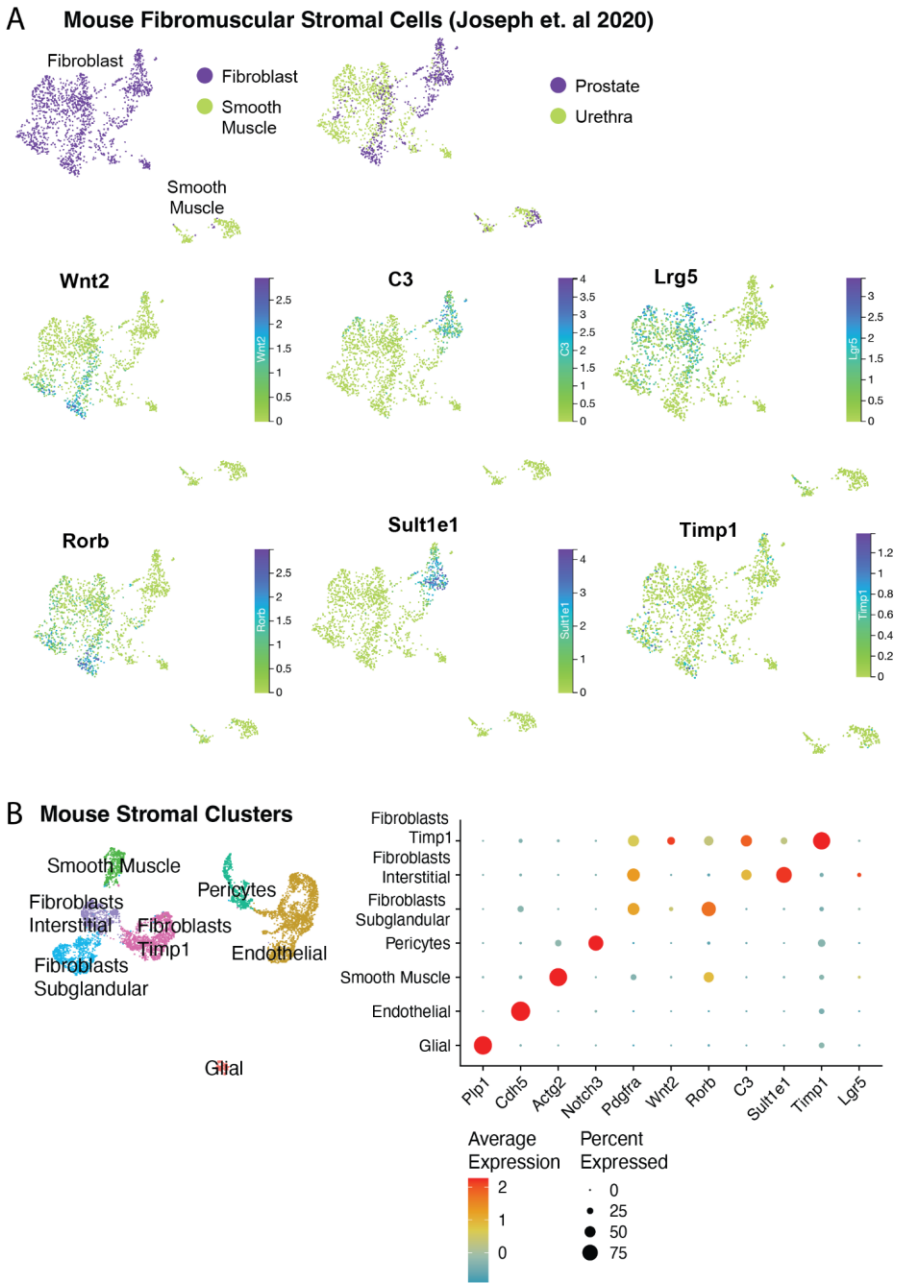

*In regards to our human prostate fibroblast data, when we performed clustering analysis on the subset of fibroblast cells, we were able to identify three distinct subclusters (PPP1R10, NPY, and CACYBP) but were unable to derive distinct C7 or APOD subclusters, as these genes appeared to be co-expressed in our human fibroblast dataset.*




*As recommended by the reviewer, we have updated supplemental figures for the wild type and Hi-Myc annotated datasets to show side-by-side UMAPs of epithelial (Figure S12C), immune (Figure S17B), and stromal cells (Figure S18C). We would also like to highlight that the cell proportion plots (see Figures 4C, 6C, and 7B) show the relative abundance of cell types for each sample/library by genotype and account for potential skewing from differences in library sizes that may not be readily apparent by viewing split UMAPs. Previously published studies by João D. Barros-Silva et al. immunostained for LY6D in mouse prostates, and showed a rare subset of luminal and basal cells expressed LY6D and this population was significantly expanded following ADT (PMID: 30566873). We originally proposed that Ly6d-expressing luminal and basal cells reflect a cell state change, as we also observe rare Ly6d-expressing cells in the WT mouse prostate. Consequently, it is possible that these Ly6d-expressing epithelial cells could represent an expansion of a pre-existing population. We have updated the text in our manuscript to raise this possibility and have also included an updated analysis describing a urothelial cluster and Hi-Myc-specific epithelial subtypes, which we refer to as Reactive Luminal and Reactive Basal (see our response to Issue #3 below, [updated manuscript lines 292 - 294, 955 - 965]).*

*To answer the question if Timp1-expressing fibroblasts are located in the WT lower urinary tract, we were able to examine the reviewer's publicly available dataset of mouse stromal cells from the prostate and lower urinary tract on the CZ CellxGene platform (https://cellxgene.cziscience.com/e/a810e511-c18b-4b2a-8fdf-98a6a0d433a7.cxg/ ). The data suggests that there does not appear to be a distinct Timp1-expressing fibroblast population in the prostate or lower urinary tract (Figure Below, Panel A). Consistent with this data, in our mouse prostate dataset, Timp1 Fibroblasts are absent in WT mouse prostates (Figures 7B, S20D). Interestingly, individual cells in the Timp1 Fibroblast cluster show mutually exclusive expression of Rorb or Sult1e1, suggesting that the Timp1 fibroblast cells were induced or derived from subglandular or interstitial fibroblasts (Supplememtal Figures S18D and S20B). Additionally, the Timp1 fibroblast cluster shows expression for Wnt2 and C3 (Figure Below, Panel B).*

*Based on gene expression, it does appear that Wnt2-expressing "ductal" cells are the equivalent of our Rorb-expressing "subglandular" fibroblasts, and the C3-expressing "prostate" fibroblasts are the equivalent of the "interstitial" fibroblasts (panel A). For consistency, we used the same nomenclature as our previously published single-cell atlas of the mouse prostate (PMID: 36373171). We selected "subglandular" to describe a fibroblast population surrounding the prostatic glandular epithelium, composed of both acini and ducts, consistent with Michael Ittmann's description of the human and mouse prostate (PMID: 29038334). Alternatively, the term "sub-acinar" may be used since we observed Rorb-expressing fibroblasts surrounding the prostatic acini in the anterior and dorsal lobes (PMID: 36373171, Figure 6A). However, Joseph et al. describe Wnt2-expressing fibroblasts surrounding glandular ducts and the acini of the anterior lobe (PMID: 34173975). As the reviewer suggested, subglandular may be interpreted to mean surrounding the prostate organ. However, describing this Rorb/Wnt2 fibroblast population as either "ductal" or "sub-acinar" may also be limiting as this fibroblast population is found surrounding both the prostatic glandular ducts and acini.*




3. It is stated that Ly6d is a urothelial marker, but the urethra wasn't included in the dissection so they can't be urothelial cells. Here is where things become semantic and confusing and I feel like we have to be careful not to be too prostate-centric. First, I would be highly surprised if there weren't at least a few Krt4+/Psca+ urethral luminal cells in wild type prostate as they project into the proximal prostate ducts and are also found in distal tips. Please check closer and annotate if so. Second, we have found that differentiating between a preexisting urethral luminal cell and a preexisting prostate luminal cell that has undergone lineage plasticity to resemble a club (human) or urethral luminal (mouse) cell due to low androgen is quite difficult. Whether it's Myc o/e, Pten ko or castration/ADT/5ARI, this induced urethral lineage identity

looks almost identical to preexisting urethral cells. It is therefore quite important to identify these cells in the wildtype data for comparison to the Myc-induced cell profile as others have done. It is also important to note that these Krt4+/Psca+ urethral urothelial cells are NOT a preexisting pool of proximal progenitors as has been definitively shown in multiple lineage tracing studies at this point. I do applaud your efforts to differentiate between this preexisting cell type and the induced cell state based on Psca expression, but as a single marker this is not enough and could be accomplished computationally by comparing to the preexisting cell type. If they can't be identified, you could use our dataset of urethra as a reference.

I hope these concerns are construed as constructive and don't delay too much.

*The reviewer is describing an interesting and previously described benign histological feature called "urothelial metaplasia" (PMID: 33070957). In the normal prostate, the urothelial epithelium can extend into periurethral ducts of the peripheral prostate and may even be the result of a hyperplastic periurethral duct. As the reviewer notes, differentiating between the urethral urothelial cells from induced lineage plastic prostate epithelial cells can be challenging, as they share several marker genes in common, including Krt4 and Psca. To identify which of these cells in the Luminal Psca Ly6d cluster were urothelial cells found in WT mouse prostates and urethra, we subsetted the Luminal cluster expressing both Psca and Ly6d, integrated samples by strain, and performed dimensionality reduction and clustering analysis. We identified three distinct clusters with unique gene expression profiles and cell proportions in WT and Hi-Myc prostates (Figure 4A-C, Supplemental Figure S11A-B). One cluster was identified as luminal prostate cells, referred to as "Luminal" as they expressed several prostate luminal-specific genes, including Agr2, Pbsn, and Tgm4 (Supplemental Figure S11B). Leveraging the publicly available dataset of WT mouse epithelial cells from the prostate and lower urinary tract on the CZ CellxGene platform (https://cellxgene.cziscience.com/e/20634fa3-f3cf-44b5-8bc3-b825610bfe8c.cxg/) from Joseph et al. (PMID: 32497356), we confirmed that these differentially expressed genes were consistent with a prostate luminal identity (Supplemental Figure S11C-D). The other two clusters were found to both express Psca and Krt4. One cluster, referred to as the "Urothelial" cluster, had comparable cell abundance between WT and Hi-Myc prostates with no significant fold change difference (Figure 4C, supplemental Figure S12C). Moreover, examining the Joseph et al. dataset confirmed that the differentially expressed genes, such as Upk1a, Gsdmc2, and Ppp1r1b, in the "Urothelial" cluster were consistent with a urethra urothelial identity (supplemental Figure S11B&E). The other cluster, referred to as "Reactive Luminal" was observed exclusively in the Hi-Myc mouse prostate and makes up most of the cells in the subsetted Psca+/Ly6d+ Luminal cluster (Figure 4C, Supplemental Figures S11A&B, S12C). Notably, this cluster expressed several genes, including Tpm2, Ly6c1, and Calcb, that were not observed in the "Urothelial" cluster in our dataset, nor the urethra and prostate epithelial clusters from the publicly available dataset provided by Joseph et al (PMID: 32497356) (Supplemental Figure S11B&F). Notably, the "Urothelial" and "Reactive Luminal" subclusters were distinct when we annotated both populations in the complete dimensionality reduction plot including all epithelial cells (Figure 4A, Supplemental Figure S11G, S12B&C ). Collectively, these additional analyses support the idea of a unique Psca+/Ly6d+ Luminal population observed in the Hi-Myc mouse prostate TME consistent with an induced prostatic*

*epithelial cell type distinct from the urethra urothelial cells. We have included these additional analyses in Figure 4, Supplemental Figures S11, S12 and have updated the manuscript text to incorporate these additional findings [updated manuscript lines 955 - 965].*

**Reviewer #3,** expertise in prostate cancer GEMM (Remarks to the Author):

Most efforts in PCa research are focussed on understanding mCRPC and NEPC, but basic research in the aetiology of prostate cancer is required to better understand how benign lesions progress to adenocarcinoma and locally invasive disease. The importance of the microenvironment and paracrine signalling in supporting PCa progression is well known, but mechanistically the specific contributions remain poorly defined. In this article Graham and colleagues performed scSeq in a cohort of localised PCa patients. Their analysis revealed that although significant heterogeneity in clustering of luminal cells can be observed inter-patients, the heterogeneity of the TME between patients is less striking.

Using the Hi-MYC mouse in 2 strains they performed clustering analysis and identified two epithelial clusters enriched for MYC. These clusters MYC1 and MYC2 differ in their enrichment between lobes. This analysis also led to the hypothesis that these MYC1/2 cells are responsible for the formation of the basal and luminal Ly6d clusters which are not enriched in MYC expression.

They then perform an elegant analysis of potential ligand-receptor enrichment between the identified clusters and basal Ly6d cells. This revealed a potential key role of interferon signalling in the paracrine interactions between MYC cells and Ly6d cluster. Their results also provides a compelling model for the switch to an immune cold phenotype in localised PCa.

My main concern relates to the exclusive use of the Hi-MYC model. For example, validating some of their key findings using the HoxB13 MYC mouse, if tissue sections are readily available, could also demonstrate that the Ly6d basal cell cluster is not Hi-MYC mouse specific. Considering that Bieberich and DeMarzo who characterised the HoxB13-MYC mouse are co-author on this manuscript it might be easy to source relevant tissues for analysis. The probasin promoter used to generate the Hi-MYC mouse can be leaky as reported by Bieberich and colleagues (PMC5006678):
“Another potential advantage of the Hoxb13 regulatory elements for driving prostate expression of MYC and loss of Pten is that we have not observed any leakiness of expression of Hoxb13-based reporter genes, MYC overexpression or Pten loss in any of these animals, whereas leakiness of the rat probasin promoter into stromal elements has been observed (58)”

*We are appreciative that the reviewer recognized the novel contributions of our study in the context of prostate TME and deemed our research approach and findings compelling. Although we use two different mouse strains, we recognize that the MYC transgene expression construct in the Hi-Myc model has limitations that may impact our observations and ultimately, our conclusions. We agree that our manuscript would benefit from additional experiments to validate*

*our original findings described in the Hi-Myc models. To that end, we include additional scRNA-seq data and analysis of BMPC mice, a Hoxb13-driven mouse model of prostate cancer with MYC activation and PTEN loss (PMID: 26554830). As pointed out by the reviewer, the Hoxb13 regulatory elements perform better as a prostate luminal-specific driver than the rat probasin promoter employed in the Hi-Myc model. In the BMPC model, PIN-like lesions are observed at 3 months, and evidence of metastatic progression is observed at 6 months. In the revised manuscript, we generated scRNA-seq libraries from the prostate (3 months) and metastatic lymph node lesion (6 months) of BMPC mice and integrated these data with our WT and Hi-Myc datasets (Supplemental Figures S21E-H). Notably, the TME of the BMPC mouse prostate was comparable to the TME of the Hi-Myc prostate. Specifically, Ly6d-expressing Reactive Basal cells, Trem2 macrophages, and Timp1 fibroblasts were enriched in the tissues of BMPC mice compared to WT prostates (Supplemental Figures S21I-K). Furthermore, the cell-intrinsic pro-inflammatory to immunosuppressive switch observed in Hi-Myc (6 months vs 10 months) also occurred in the BMPC mice (3 months vs 6 months) (Supplemental Figure S21L). Altogether, these additional studies in the BMPC mice validate and recapitulate the key MYC-driven alterations to the TME we initially observed in Hi-Myc mice and human prostates.*

Minor:
As published in an addendum to the original Hi-MYC article from 2003, the MYC protein expressed in this transgene contains additional C-terminal amino acids. So it is possible that some of the observed effects are related to this. Probably a simple mention in the discussion that the MYC expressed in this model may not behave exactly like WT Myc would be sufficient.

*As recommended by the reviewer, we reference the addendum to the originally published Hi-Myc study regarding the additional amino acids of the expressed human MYC transgene [updated manuscript lines 482 - 486]. While Hi-Myc is a commonly used mouse model in the prostate cancer field, this potential limitation in the Hi-Myc mouse model and the associated MycCaP cell line is not widely known or discussed. We agree that it is appropriate to raise this issue, and have now done so in the revised manuscript.*

Reviewers' Comments:

Reviewer #1:
Remarks to the Author:
The authors have addressed many of my comments. There remain a few concerns that reduce significance od the paper.
1. It would be important to compare the relative abundance of each cluster type, such as Luminal MYC1, MYC2, at different time points in different stages of progression.
2. The TP53 pathway was significant in enrichment analyses but the authors hypothesized that PTEN loss could potentially drive MYC activation. They should examine if the alterations co-occur in the accelerating immunosuppressive tumor microenvironment (TME)? PTEN loss demonstrates heterogeneity within the samples.
3. The changes in the MYC programs throughout the progression in the mouse models (BMPC, HiMYC and Wt 6mo vs 10months) should be described.
4. The results identify EGR2, WT1 and other TFs with a high correlation with TGFB receptors. Hence it is not clear why the authors chose EGR4 as the transcription factor of interest in their study, given the high correlation of several other factors with TGFB markers?
5. There were incorrect references to the figures, supplementary figures 8, 9,
6. There were typos such as "Epithelial" in the title of Figure 4 should be addressed.

Reviewer #2:
Remarks to the Author:
see my comments in the attachment

**Reviewer #2,** expertise in scRNA-seq for prostate cancer (Remarks to the Author):

Graham et al perform a single cell study on primary human prostate tumors and mouse prostates from 2 different Myc overexpressors. The manuscript is expertly written and the data a solid addition to the field. Clearly much thought and work went into the project. I only have minor suggestions. Please note that as a key person responsible for nomenclature, my comments are meant to help the entire field get on the same page even though it
might seem like nitpicking. Hopefully these suggestions won't take too much time as they are mostly computational.

*We thank the reviewer for highlighting the strengths of this manuscript and its importance to the field. We agree wholeheartedly that there is great value in developing consensus in the field with regards to a common nomenclature around identified cell types/states for organs and systems. We thank the reviewers for their leadership and contributions in advancing this type of effort for the field. Thanks to the reviewer's constructive feedback,
where we saw clear agreement with gene expression signatures for cell types/states that aligned with previously defined nomenclature, we have now revised the nomenclature to reflect that. However, in a few instances, our data have shown that there may be more nuance needed in applying the appropriate nomenclature and the suggested terms would not be fully consistent with the states that we observed. These determinations were assessed with great depth and rigor to make sure we follow the spirit of this reviewer's constructive criticisms. In those instances, we have indicated why we think the nomenclature we previously used is more appropriate, both based on data that we present in this revised manuscript, as well as prior data from two and a half decades of studying prostate cancer cell types/states using molecular pathologic, genetic, and epigenetic studies in our collaborative team. Additionally, we thank the reviewer for disclosing his identity, and hope the reviewer will be open to working together with us and others in the field to develop a highly informed and nuanced consensus on the few cell types/states that could not easily be conformed to terminologies suggested by this reviewer.*

1. The labeling of the club cell cluster as 'intermediate/atrophic' cells is problematic. Your own recent paper labeled PIA as club (like) cells. I think it's confusing for the field to rename a cell 'type' according to a histopathological phenotype. Please stick with club or club-like. I also think it's wrong to continue to conflating the PIA phenotype with actual cancer insinuating it is a precursor when your own work shows that it is tumor adjacent. Surely if it were a precursor it would be possible to detect CNV in club (like) cells. Is it not quite possible that the club-like cells within the same gland coexist with tumor cells without being clones (and are maybe even induced to that state by tumor cells)? We have enough evidence to show that the luminal to club plasticity is a normal response to androgen depletion and not restricted to cancer. I'm actually quite surprised there isn't any evidence yet that club cells are precursors or harbor mutations (it would make sense), but we should be careful about insinuating they are precursors until there is quality evidence (trajectory analysis doesn't count!). This should be explicitly stated in the manuscript so there is no confusion.

*We agree with the reviewer that the goal of developing consistent nomenclature for cell types in the prostate is worthwhile to pursue, and recognize the significant efforts to generate single-cell atlases and share this critical resource with the field. However, we are concerned that using the label of club or club-like to describe broadly the cluster of cells that we classified as atrophy/intermediate has multiple major issues. Furthermore, we would like to clarify multiple of the comments made by the reviewer with respect to PIA and its relationship with cancer, and appreciate the opportunity presented by the reviewer to make sure our discussion of these cell types and the PIA lesions reflect what we have uncovered about these lesions over the last 20+ years of research in studying them across our team using morphological and molecular genetic and epigenetic studies since the initial description of PIA lesions led by the De marzo lab in 1999 (PMID: 10595928, 11068311, 11550208, 11795433, 12707036, 12937133, 14607218, 14755684, 15880529, 17384581, 18302221, 19574101, 22212087, 29422309, 30900284, 34341114, 37347938, 37550827). We have organized our response to this reviewer comment in two major themes to address these issues:*

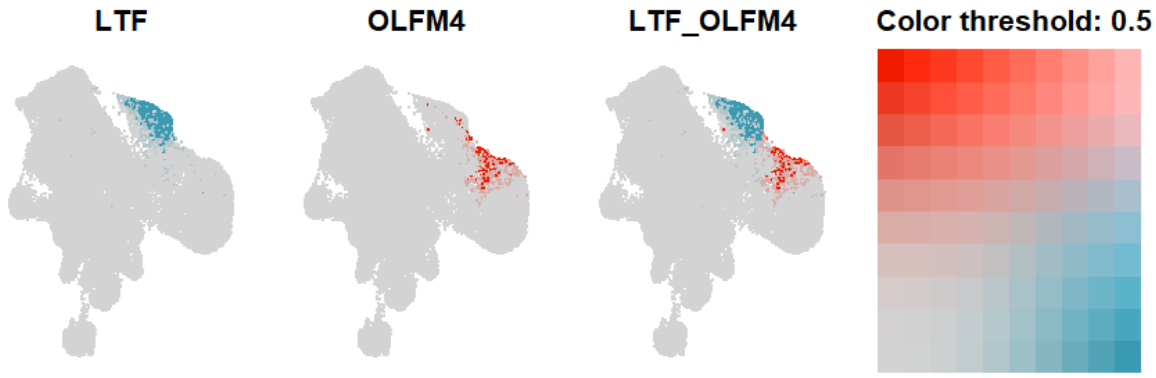
a. *<u>Clarifications about PIA and its relationship to PIN and invasive cancer</u>*

*First, we emphasize that our intention was not to imply that atrophy always leads to cancer or is a cancer-exclusive process, but rather that the inflammatory conditions that give rise to a subtype of atrophy called proliferative inflammatory atrophy (PIA) can form a hotbed of proliferative epithelial cells that can give rise to neoplastic cells (PMID: 29089606, 34341114, 10595928, 26813971). While PIA does not always lead to cancer, extensive literature exists supporting the role of PIA as a precursor of prostate cancer, given that PIA often merges with the prostatic intraepithelial neoplasia, an established prostate cancer precursor, and even with invasive cancer lesions (PMID: 19507201, 10595928, 11068311). Furthermore, evidence of cancer-associated molecular alterations has been reported in PIA, including p53 alterations (PMID: 12210488) and gain of chromosome 8 centromeric DNA (PMID: 11337374). ERG rearrangement and activation, which occurs in roughly half of all prostate cancer cases (PMID: 25915839), has been reported in PIA (PMID: 34341114). A small fraction of PIA cells can harbor epigenetic DNA hypermethylation at GSTP1 and other genes leading to their silencing, similar to the majority of PIN and cancer lesions (PMID: 37347938, 30082453). Nonethless, the majority of PIA cells/lesions do not harbor genetic or epigenetic alterations seen in human cancers, but rather show a phenotype consistent with stress response and injury/regeneration – with evidence to support this at the histomorphological, and molecular genetic/epigenetic and gene/protein expression levels. The luminal epithelial cells in these PIA lesions appear atrophic at a cellular scale (instead of having a tall columnar cell appearance characteristic of totally benign prostate luminal epithelial cells, they have a stunted/atrophic appearance), and can express genes suggesting a state that can be considered intermediate between luminal and basal epithelial cell types (e.g. mix of typically basal and luminal cytokeratins) and often have strong induction of stress response genes (e.g. GSTP1, GSTA1, PTGS2, LTF). Collectively, this data supports the idea that the inflammatory conditions driving PIA can form the hotbed from which PIN and/or invasive cancer lesions can arise. We have updated our text to emphasize these points [updated manuscript lines 82 - 97].*

*Additionally, prostatic atrophy is not a homogenous population of epithelial cells but rather is comprised of many different variations with distinct and nuanced biology. Although*

*rarely discussed, there are two major classes of prostate atrophy (Prostate Cancer: Biology, Genetics, and the New Therapeutics, 2nd Edition, Chapter 2). The first could be considered hormonal or diffuse atrophy, which is induced after androgen deprivation or , androgen receptor blockade. This, or a variation of it, was studied recently by Joseph et al. (PMID: 34928497). However, the second type, called focal atrophy, which includes PIA, is not known to be induced by androgen blockade but can occur at any age and is associated with inflammation. The consensus by leading experts in the field has recognized that focal prostatic atrophy is a broad categorization of various subtypes (PMID: 17001160). How gene expression modules in luminal intermediate cells in focal atrophy and luminal cells in hormonal atrophy compare has not been determined. While potentially sharing some transcriptional similarities, these atrophy subtypes are at least somewhat unique and more detailed studies to describe these differences are underway in our team.*

*In terms of focal atrophy and PIA, there are ongoing efforts in our research group to leverage scRNA-seq analysis to define these various atrophic/intermediate subtypes, as there appear to be transcriptionally unique subgroups. For example, OLFM4 and LTF, as part of broader gene expression modules, have distinct regional expression within the atrophy/intermediate cluster (see UMAPs below). We believe these transcriptionally distinct subgroups reflect distinct atrophy/intermediate cell states. While the extensive characterization of atrophy/intermediate cell subtypes is beyond the scope of this cancer-focused study, we plan on sharing these research findings in a future manuscript that includes a comprehensive analysis of all prostate zones (peripheral, central, and transition), with a thorough characterization of luminal, basal, and atrophy/intermediate cell types*



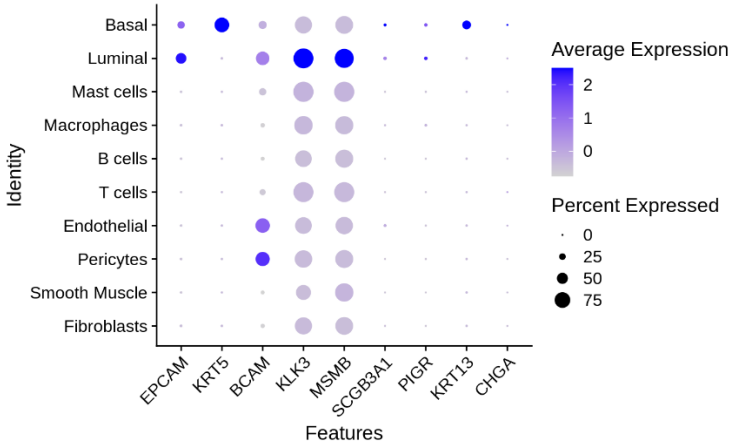

b. *<u>Club-like is not synonymous with atrophy/intermediate cells</u>*

*In their landmark study in 2018 published in Cell Reports, they used FACS and single cell and bulk RNA sequencing to develop a single-cell atlas of normal prostate (PMID: 30566875). In that study, they identified the presence of cells with very high transcriptional similarity to club cells in the lung, with a key defining feature that they expressed the secretoglobin genes SCGB1A1 and SCGB3A1, which are well known markers of lung club cells. While these cells were identified in cells isolated from combined central/transition zone prostate tissues, they were nearly absent in the peripheral zone in that study (PMID: 30566875, Figure 5B). Consistent with these previous findings, we also find that in our peripheral zone tissues, SC GB1A1 expression is very rare and SCGB1A1 or SCGB3A1 expressing cells*

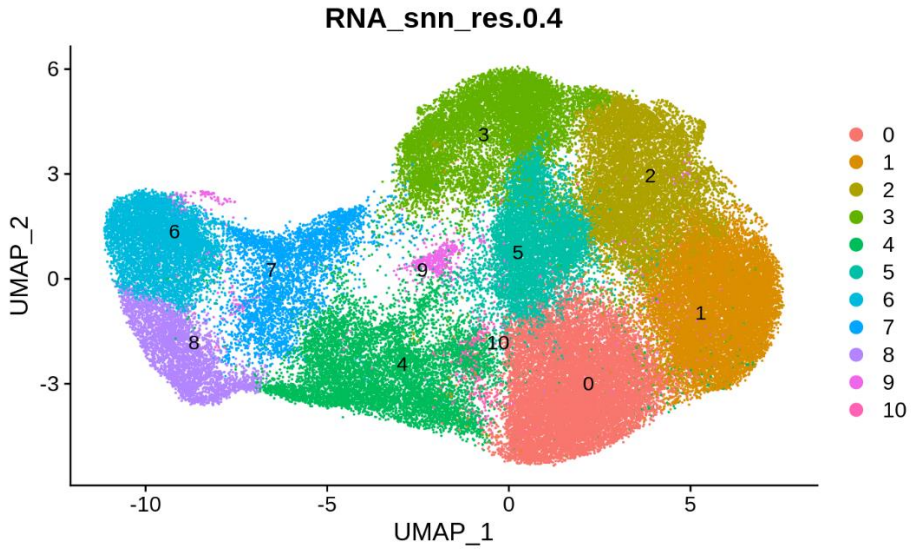
*were not identified in our peripheral zone atrophy/intermediate cell cluster (Supplemental Figure S1E). Thus, we feel it would be too presumptive to broadly equate the atrophy/intermediate cells in our peripheral zone tissues to club cells, for which the most well-defined markers are the SCGB1A1 and similar secretoglobin genes, which are absent in our atrophy/intermediate clusters.*

*It is simply not true that there is limited secretoglobin expression in your data. Moreover, expression of a single gene does not signify cell type. I have a few suggestions based on my analysis of your Seurat object:*

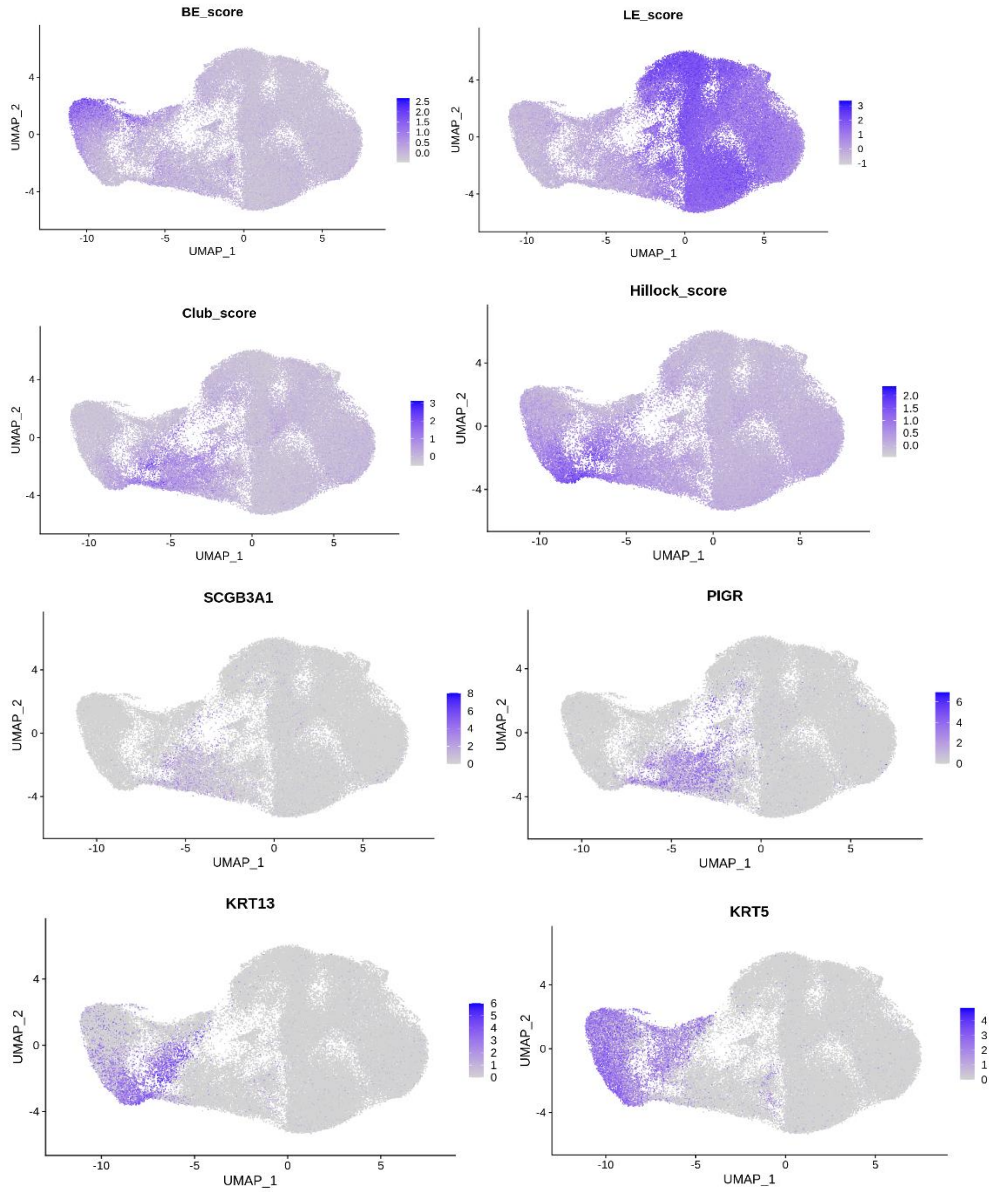
*The data need ambient RNA correction. There is background luminal epithelial gene expression across all clusters signifying incomplete washing before barcoding. I would highly suggest using CellBender. I'm guessing this is why your clustering looks a little wonky.*

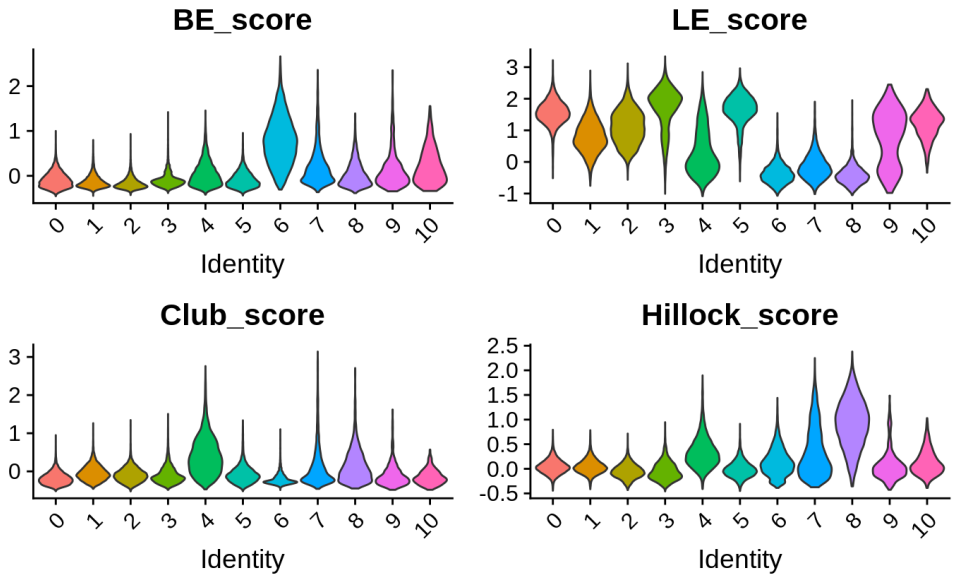



*I would also suggest reclustering after subsetting. The following image is from your prostatectomy data where we subset your epithelia and then reclustered. (Of note, we use log transformation instead of SCTransform if you want to try).*



*I ran a quick correlation analysis using gene scores from basal, luminal, club and hillock reference data on your data and the results are quite similar to what the Sykes lab published (Figure 2, Nature Comms 2023, 14:663). The OLFM4+ cluster in your image above is hillock and the LTF+ are club. I have no doubt that there exists a spectrum of states for club and hillock in cancer as we and others have observed, but you cannot state that you won't annotate them because you think they should not exist in PZ specimens, especially since others have already published that they do (with in situ validation) (PMID: 36750562, 37550827). I know you are partial to the PIA terminology, but it really isn't appropriate for single cell data unless you have some spatial correlation. In fact, I would argue that PIA is simply a spectrum of luminal to club intermediate states as Franklin Huang suggested (NOT an intermediate between luminal and basal). Your own data suggest so.*






*While we sampled all zones of the prostate in our study, our single-cell analyses were performed exclusively in the peripheral zone, except for our inferred copy number variation analysis, where patient-matched tissues collected from the central and transition zones that were histologically evaluated as cancer-free were used as normal controls (Figure 2C, Supplemental Figures S3-5). To make sure we weren't simply missing detection of the secretoglobin genes due to some technical issue, we now went back to our scRNA-seq data from the central and transition zone tissues, and found that SCGB1A1 and SCGB3A1 genes are indeed robustly expressed in the intermediate cell cluster in the transition zone, and largely absent in the peripheral zone (Supplementary Figure S1E), again consistent with the previous fin dings. Histological assessment of our peripheral zone tissues showed extensive focal atrophy consistent with PIA, suggesting that a substantial proportion of our epithelial cells include atrophy/intermediate cells from PIA.*

*It is a stretch of the data to claim that every cell in the cluster you define as 'intermediate' belongs to the PIA phenotype. Unless you have spatial evidence linking genotype to phenotype like we (PMID: 34928497) and others (PMID: 36750562, 37550827) provide, then you cannot simply say it's nuanced. The burden of proof is now on you to show that PIA is not a spectrum of club cell states. Your own data suggest that it is. The transcriptome of the intermediate cluster is nearly identical to club cells from normal prostatic urethra. In fact, intermediate cells are not between luminal and basal or luminal and hillock, but only between luminal and club as we definitively show in PMID: 34928497. No one in the current single cell era only uses single genes to define a cell type. I'm not sure why you are specifically focusing on SCGB1A1 as the definitive marker of club cells as there are many more. Labeling an entire cluster of cells along the spectrum of urothelial club and prostate luminal as a single phenotype (PIA) that includes a heterogeneous mixture is inappropriately simplifying the data to match the desired annotation of your phenotype. This must be corrected.*

*Interestingly, although the atrophy/intermediate cells characterized in our study lack expression of the secretoglobin genes, they do have multiple genes that are also expressed in*

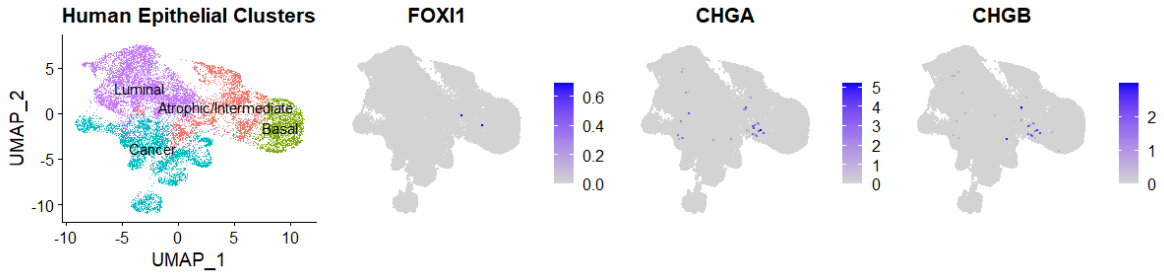
*the club-like cells that were present in the transition zone and previously described. These includ e the OLFM4 and PIGR genes. Based on the expression of these genes in PIA lesions, in a recent study led by Franklin Huang and Karen Sfanos, in which Angelo De Marzo was also a co-author, it was suggested that the luminal cells in PIA lesions are club-like cells (PMID: 37550827). In another study, Franklin Huang's group termed these cells "cancer club cells" (PMID: 35013146), presumably due to their presence in peripheral zone tissues containing cancer lesions. However, in both of those studies, it was shown that secretoglobin expression was very rare to absent in these peripheral zone cells. In situ analysis of the LTF and PIGR expression revealed that these were indeed positive in atrophy/intermediate cells of PIA lesions, and that SCGB1A1 expression was very rare in these cells (PMID: 37550827). In our own data, we have seen that LTF (and associated genes such as PIGR) and OLFM4 (and associated genes) are expressed in distinct subsets of atrophy/intermediate cells. Based on these detailed analyses, triggered in large part because of the comments from this reviewer that led us to challenge our initial labels, we have now arrived at the conclusion that the atrophy/intermediate cells are a broad categorization composed of multiple subsets of transcriptionally distinct cell states that are recognizable due to the large number of epithelial cells evaluated in our study. Secretoglobin expressing club-like cells are possibly a rare subset of these atrophy/intermediate cells. In an ongoing study, we are thoroughly characterizing the various atrophy/intermediate cell states/subtypes and plan to develop a report describing these findings. In the meanwhile, we feel it is most appropriate to retain the umbrella term atrophy/intermediate cells to describe this population of cells.*

2. This is bound to be a highly valuable dataset for the field so I would ask for due diligence on cell type annotation. Please identify club (not 'atrophic'), hillock, ionocyte, and neuroendocrine cells in Figure 2 as well as periprostatic/subglandular (APOD+) and interstitial fibroblasts (C7+) (Figure 7D). The full transcriptomic identification of these cell types can be found here: PMID: 34173975. Of note, we are now labeling the subglandular fibroblasts as 'periprostatic' because they can be distinguished from subglandular fibroblasts from other organs. Remember the ultimate goal for the field of science is to create nomenclature that can distinguish these cells in multi-organ datasets. This is why we can't just say 'luminal', we have to say 'prostate luminal' or 'urethral luminal' when multiple tissues are represented in a dataset. Additionally, I think it's important to show the wild type and Myc annotated datasets side-by-side (using the split.by function) and not overlapped by color even though it takes up more space in the figure. People need to see whether the induced clusters (Ly6d basal, psca luminal and timp fibs) are present in the wild type and therefore represent an expansion of a preexisting cell type vs. a change in cell state. Even if it needs to be placed in a supplementary figure, it has to be visible for the reader somewhere. On that note, it is necessary to show each of these markers used to annotate a myc-induced cell state in situ in the wild type lower urinary tract. It is known that Psca labels the entire urothelium. Ly6d labels urethral urothelium. I can see Timp1 expression in normal fibroblasts from our own data set, but where are they located in the wild type lower urinary tract? Is your Timp1+ fib an expansion of a preexisting, specially localized fibroblast or an induction of subglandular or interstitial fibroblasts? Timp1 staining is warranted in wild-type and myc lower urinary tracts to answer this for the reader. If the answer is that you can't identify a preexisting Timp1+ fibroblast, then that is still informative. Of note, what you are calling 'interstitial' fibs in

mouse prostate is what we called 'prostate' fibs and what you are calling 'subglandular' is what we identify as 'ductal' based on staining of the entire lower urinary tract. Clearly, the field still needs to come to a consensus on nomenclature, but I feel the term subglandular is confusing for a fibroblast that isn't surrounding prostate glands, but rather the proximal ducts. It's your call, but since you cited our work on this, I thought I'd point it out.

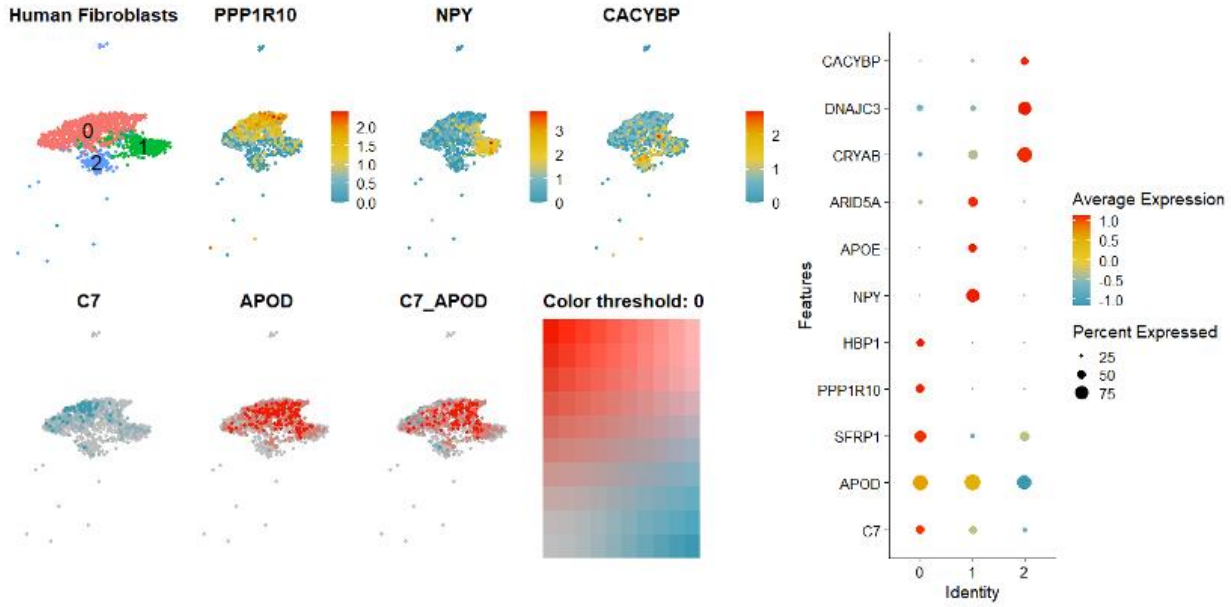
*In regards to the ionocytes, authors (Karthaus et al. 2020) who originally identified the Foxi1-expressing ionocyte-like luminal epithelial cluster in the mouse prostate noted that they did not identify a distinct human ionocyte cluster in their scRNA-seq data (PMID: 32355025). Consistent with this published finding, we observed only 2 luminal cells that expressed FOXI1 and consequently did not update Figure 2. We included the Foxi1 expressing luminal cluster in Figure 3 of mouse data and have updated the name of this cluster from Foxi1 to "Ionocytes".*

*To address the issue of neuroendocrine cells specifically, we would like to highlight that Karthaus et al. did not observe a distinct neuroendocrine cluster (PMID: 32355025). They attributed this to the rarity of the neuroendocrine cells. Likewise, Joseph et al. also did not observe a distinct neuroendocrine cluster, as neuroendocrine cells were also scattered in UMAP space (PMID: 30566875). In our data, neuroendocrine cells (CHGA+ CHGB+) appear rare and do not form a distinct cluster on our UMAP. This likely is due to being an extremely rare population in our dataset (21 cells). As a result, we did not update Figure 2 with a neuroendocrine cluster.*




*In regards to our human prostate fibroblast data, when we performed clustering analysis on the subset of fibroblast cells, we were able to identify three distinct subclusters (PPP1R10, NPY, and CACYBP) but were unable to derive distinct C7 or APOD subclusters, as these genes appeared to be co-expressed in our human fibroblast dataset.*

*This is quite interesting and seems to mimic your mouse data that 2 different fibroblasts take on a single CAF signature. However, I would also argue that the fibroblast numbers are suspiciously low perhaps due to the digestion method. Maybe more data will shed light on this, but your identification of a single fibroblast cluster mimics the 2023 Sykes paper as well.*




*As recommended by the reviewer, we have updated supplemental figures for the wild type and Hi-Myc annotated datasets to show side-by-side UMAPs of epithelial (Figure S12C), immune (Figure S17B), and stromal cells (Figure S18C). We would also like to highlight that the cell proportion plots (see Figures 4C, 6C, and 7B) show the relative abundance of cell types for each sample/library by genotype and account for potential skewing from differences in library sizes that may not be readily apparent by viewing split UMAPs. Previously published studies by João D. Barros-Silva et al. immunostained for LY6D in mouse prostates, and showed a rare subset of luminal and basal cells expressed LY6D and this population was significantly expanded following ADT (PMID: 30566873). We originally proposed that Ly6d-expressing luminal and basal cells reflect a cell state change, as we also observe rare Ly6d-expressing cells in the WT mouse prostate. Consequently, it is possible that these Ly6d-expressing epithelial cells could represent an expansion of a pre-existing population. We have updated the text in our manuscript to raise this possibility and have also included an updated analysis describing a urothelial cluster and Hi-Myc-specific epithelial subtypes, which we refer to as Reactive Luminal and Reactive Basal (see our response to Issue #3 below, [updated manuscript lines 292 - 294, 955 - 965]).*

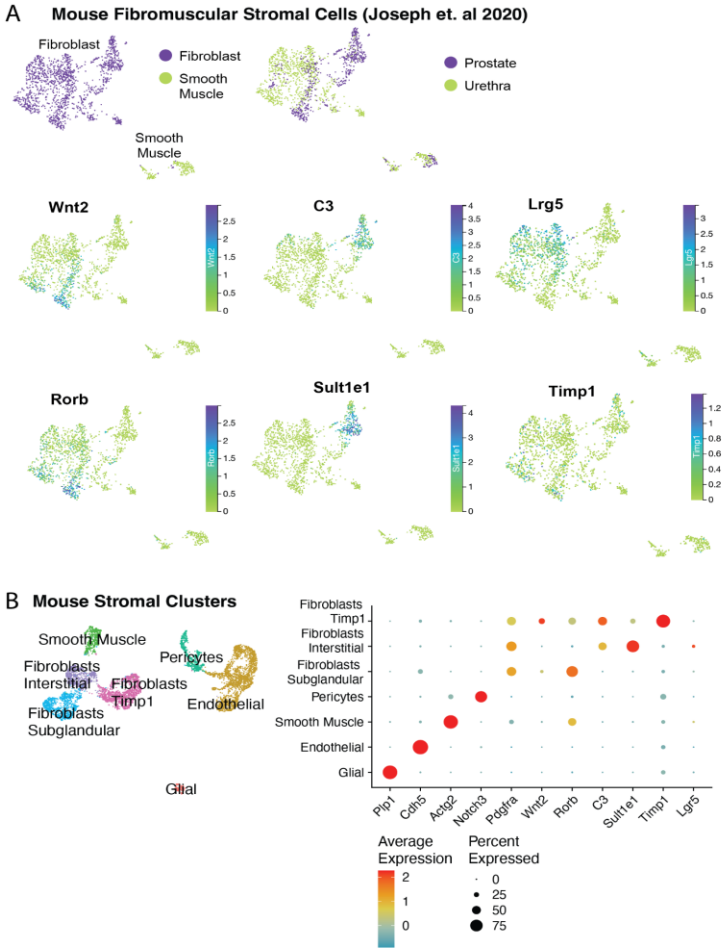
*I really like the annotation of these as 'reactive' as it signifies conversion of a preexisting prostate cell type even if it now looks more like urothelium than prostate.*

*To answer the question if Timp1-expressing fibroblasts are located in the WT lower urinary tract, we were able to examine the reviewer's publicly available dataset of mouse stromal cells from the prostate and lower urinary tract on the CZ CellxGene platform (https://cellxgene.cziscience.com/e/a810e511-c18b-4b2a-8fdf-98a6a0d433a7.cxg/ ). The data suggests that there does not appear to be a distinct Timp1-expressing fibroblast population in the prostate or lower urinary tract (Figure Below, Panel A). Consistent with this data, in our mouse prostate dataset, Timp1 Fibroblasts are absent in WT mouse prostates (Figures 7B, S20D). Interestingly, individual cells in the Timp1 Fibroblast cluster show mutually exclusive expression of Rorb or Sult1e1, suggesting that the Timp1 fibroblast cells were induced or*

*derived from subglandular or interstitial fibroblasts (Supplememtal Figures S18D and S20B). Additionally, the Timp1 fibroblast cluster shows expression for Wnt2 and C3 (Figure Below, Panel B).*

*Based on gene expression, it does appear that Wnt2-expressing "ductal" cells are the equivalent of our Rorb-expressing "subglandular" fibroblasts, and the C3-expressing "prostate" fibroblasts are the equivalent of the "interstitial" fibroblasts (panel A). For consistency, we used the same nomenclature as our previously published single-cell atlas of the mouse prostate (PMID: 36373171). We selected "subglandular" to describe a fibroblast population surrounding the prostatic glandular epithelium, composed of both acini and ducts, consistent with Michael Ittmann's description of the human and mouse prostate (PMID: 29038334). Alternatively, the term "sub-acinar" may be used since we observed Rorb-expressing fibroblasts surrounding the prostatic acini in the anterior and dorsal lobes (PMID: 36373171, Figure 6A). However, Joseph et al. describe Wnt2-expressing fibroblasts surrounding glandular ducts and the acini of the anterior lobe (PMID: 34173975). As the reviewer suggested, subglandular may be interpreted to mean surrounding the prostate organ. However, describing this Rorb/Wnt2 fibroblast population as either "ductal" or "sub-acinar" may also be limiting as this fibroblast population is found surrounding both the prostatic glandular ducts and acini.*

*I think we can agree that the fibroblast annotations are muddled because of the differences between lobes, but at some point we all have to start calling each cell type the same thing.*



3. It is stated that Ly6d is a urothelial marker, but the urethra wasn't included in the dissection so they can't be urothelial cells. Here is where things become semantic and confusing and I feel l ike we have to be careful not to be too prostate-centric. First, I would be highly surprised if there weren't at least a few Krt4+/Psca+ urethral luminal cells in wild type prostate as they project into the proximal prostate ducts and are also found in distal tips. Please check closer and annotate if so. Second, we have found that differentiating between a preexisting urethral luminal cell and a preexisting prostate luminal cell that has undergone lineage plasticity to resemble a club (human) or urethral luminal (mouse) cell due to low androgen is quite difficult. Whether it's Myc o/e, Pten ko or castration/ADT/5ARI, this induced urethral lineage identity looks almost identical to preexisting urethral cells. It is therefore quite important to identify these cells in the wildtype data for comparison to the Myc-induced cell profile as others have done. It is also important to note that these Krt4+/Psca+ urethral urothelial cells are NOT a preexisting pool of proximal progenitors as has been definitively shown in multiple lineage tracing studies at this point. I do applaud your efforts to differentiate between this preexisting cell type and the induced cell state based on Psca expression, but as a single marker this is not enough and could be accomplished computationally by comparing to the preexisting cell type. If they can't be identified, you could use our dataset of urethra as a reference.

I hope these concerns are construed as constructive and don't delay too much.

*The reviewer is describing an interesting and previously described benign histological feature called "urothelial metaplasia" (PMID: 33070957). In the normal prostate, the urothelial epithelium can extend into periurethral ducts of the peripheral prostate and may even be the result of a hyperplastic periurethral duct. As the reviewer notes, differentiating between the urethral urothelial cells from induced lineage plastic prostate epithelial cells can be challenging, as they share several marker genes in common, including Krt4 and Psca. To identify which of these cells in the Luminal Psca Ly6d cluster were urothelial cells found in WT mouse prostates and urethra, we subsetted the Luminal cluster expressing both Psca and Ly6d, integrated samples by strain, and performed dimensionality reduction and clustering analysis. We identified three distinct clusters with unique gene expression profiles and cell proportions in WT and Hi-Myc prostates (Figure 4A-C, Supplemental Figure S11A-B). One cluster was identified as luminal prostate cells, referred to as "Luminal" as they expressed several prostate luminal-specific genes, including Agr2, Pbsn, and Tgm4 (Supplemental Figure S11B). Leveraging the publicly available dataset of WT mouse epithelial cells from the prostate and lower urinary tract on the CZ CellxGene platform (https://cellxgene.cziscience.com/e/20634fa3-f3cf-44b5-8bc3-b825610bfe8c.cxg/) from Joseph et al. (PMID: 32497356), we confirmed that these differentially expressed genes were consistent with a prostate luminal identity (Supplemental Figure S11C-D). The other two clusters were found to both express Psca and Krt4. One cluster, referred to as the "Urothelial" cluster, had comparable cell abundance between WT and Hi-Myc prostates with no significant fold change difference (Figure 4C, supplemental Figure S12C). Moreover, examining the Joseph et al. dataset confirmed that the differentially expressed genes, such as Upk1a, Gsdmc2, and Ppp1r1b, in the "Urothelial" cluster were consistent with a urethra urothelial identity (supplemental Figure S11B&E). The other cluster, referred to as "Reactive*

*Luminal" was observed exclusively in the Hi-Myc mouse prostate and makes up most of the cells in the subsetted Psca+/Ly6d+ Luminal cluster (Figure 4C, Supplemental Figures S11A&B, S12C). Notably, this cluster expressed several genes, including Tpm2, Ly6c1, and Calcb, that were not observed in the "Urothelial" cluster in our dataset, nor the urethra and prostate epithelial clusters from the publicly available dataset provided by Joseph et al (PMID: 32497356) (Supplemental Figure S11B&F). Notably, the "Urothelial" and "Reactive Luminal" subclusters were distinct when we annotated both populations in the complete dimensionality reduction plot including all epithelial cells (Figure 4A, Supplemental Figure S11G, S12B&C ). Collectively, these additional analyses support the idea of a unique Psca+/Ly6d+ Luminal population observed in the Hi-Myc mouse prostate TME consistent with an induced prostatic epithelial cell type distinct from the urethra urothelial cells. We have included these additional analyses in Figure 4, Supplemental Figures S11, S12 and have updated the manuscript text to incorporate these additional findings [updated manuscript lines 955 - 965].*

Truly well done on these revisions.

Reviewer #3:
Remarks to the Author:
The authors have addressed my initial criticisms and have provided new data using the HoxB13 mouse model.

**Reviewer 1**

The authors have addressed many of my comments. There remain a few concerns that reduce significance [of] the paper.

We thank Reviewer 1 for their feedback and are pleased to know that our revised manuscript addressed many of the initial concerns raised by the reviewer. Below, we address the additional follow-up concerns in our updated manuscript, with point-by-point responses.

1. It would be important to compare the relative abundance of each cluster type, such as Luminal MYC1, MYC2, at different time points in different stages of progression.

We agree that it can be helpful and important to assess the relative cell abundance for the key cell types enriched in the TME of MYC-driven prostate cancer across stages (WT to precursor to invasive to metastasis). Key cell types include the Reactive Basal cluster, Trem2 Macrophages, Timp1 Fibroblasts, and Luminal MYC clusters. To perform this analysis, we have grouped the Luminal MYC 1 and 2 clusters for technical reasons. The Luminal MYC 2 cluster is a ventral-enriched subpopulation and is not present for 10-month Hi-MYC mice, as only the dorsal and lateral lobes were dissected and sequenced for this time-point. As a comparator, we also include the cell populations of normal basal, luminal, and fibroblast clusters, as well as the Trem2-negative macrophage cells (see Supplemental Figure S21L) and update the manuscript text [Lines 453-456].

2. The TP53 pathway was significant in enrichment analyses but the authors hypothesized that PTEN loss could potentially drive MYC activation. They should examine if the alterations co-occur in the accelerating immunosuppressive tumor microenvironment (TME)? PTEN loss demonstrates heterogeneity within the samples.

We thank the reviewer for the opportunity to further clarify the roles of TP53 and PTEN pathways in the context of MYC-driven prostate cancer. In human cancers, genetic mutation or loss of TP53 or PTEN is associated with aggressive prostate cancer and the development of metastatic disease. This aggressiveness is well known to be further compounded with the co-occurring loss of TP53 and PTEN in the same tumors (PMID: 29460925). We explored the co-occurrence of TP53 pathway downregulation and PTEN/mTORC pathway dysregulation in the Hi-Myc and BMPC models (see Supplemental Figure S22B-C and updated manuscript text [Lines 456-459]). In the invasive carcinoma-enriched Hi-Myc (10 months) and metastatic BMPC (6 months), we observed suppression of P53 signaling compared to precursor-enriched Hi-Myc (6 months) and BMPC (3 months) models. In the Hi-Myc model, we observed upregulated mTOR signaling at the carcinoma-enriched time point compared to the precursor-enriched time point. In contrast, mTOR signaling was elevated in the BMPC mice at the precursor-enriched and metastatic timepoints. In the BMPC model, MYC is activated, and PTEN is lost. Consequently, the loss of PTEN is expected to upregulate mTOR signaling.

3. The changes in the MYC programs throughout the progression in the mouse models (BMPC, HiMYC and Wt 6mo vs 10months) should be described.

As recommended by the reviewer, we have performed gene signature analyses for several MYC programs for each stage of progression in the WT, Hi-Myc, and BMPC mice. Interestingly, we observed that MYC programs were especially upregulated in metastatic prostate cancer cells (see Supplemental Figure S22A and updated manuscript text [Lines 456-459).

4. The results identify EGR2, WT1 and other TFs with a high correlation with TGFB receptors. Hence it is not clear why the authors chose EGR4 as the transcription factor of interest in their study, given the high correlation of several other factors with TGFB markers?

We agree with the reviewer that transcription factors other than EGR4 highly correlate with TGFB receptors, including EGR2 and WT1. We highlight EGR4 because it is known to directly regulate transcription of *Timp1* and the fibrosis-associated collagens *Col1a1, Col1a2, Col3a1, and Col5a2.* This was not the case with the other transcription factors, which have been reported to be upstream for only a subset of these genes. However, it is quite possible that multiple transcription factors may be involved in the upregulation of *Timp1* and fibrosis-associated collagens in fibroblasts, and not exclusively EGR4. We have updated the manuscript text to state this possibility [Lines 412-416, 516].

5. There were incorrect references to the figures, supplementary figures 8, 9,
As noted by the reviewer, there were incorrect references to the supplemental figures. Upon closer inspection, we discovered that in the PDF supplemental document, Figure 10 was mistakenly placed as Figure 1. Consequently, all figures S1-S10 were unintentionally shifted by one. We have corrected this error in an updated version.

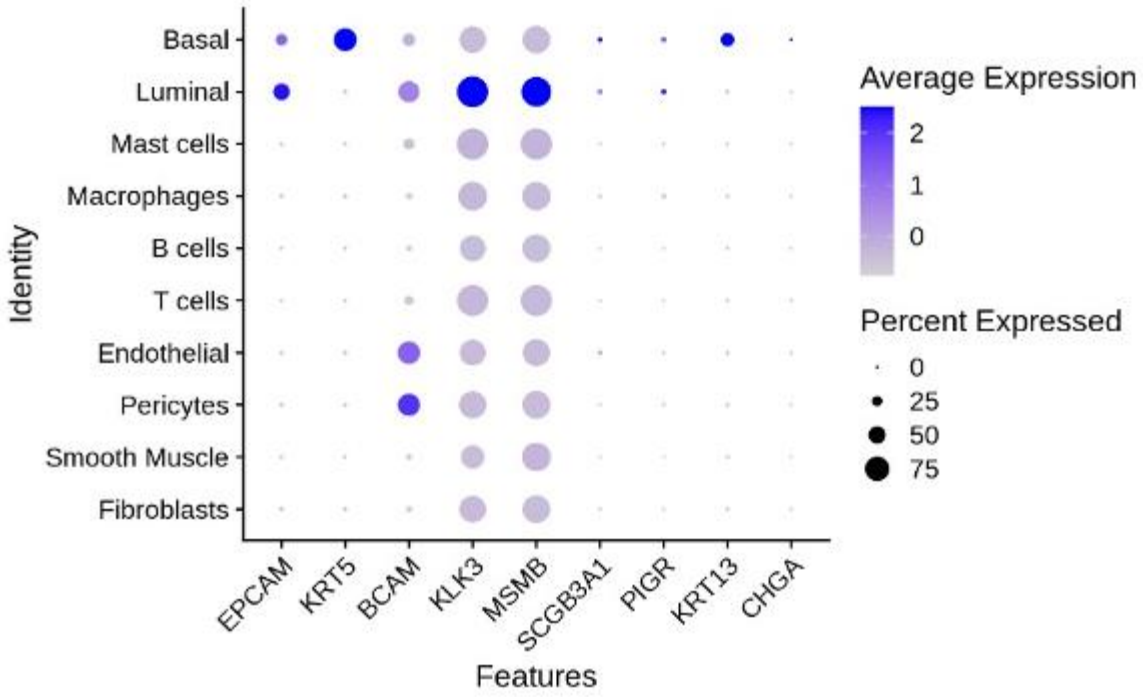
6. There were typos such as "Epithelial" in the title of Figure 4 should be addressed.
We thank the reviewer for indicating typos in the manuscript. We have double-checked Figure 4 and the associated figure legend and updated it.

**Reviewer 2**

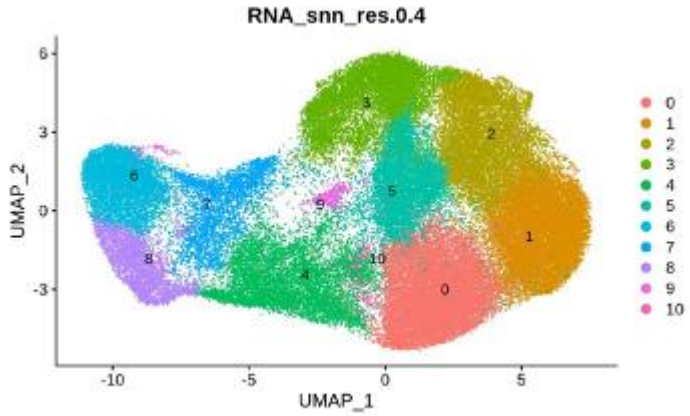
We thank Reviewer 2 for recognizing our efforts and complimenting the work of our revised manuscript. For brevity, we have included the comments from Reviewer 2 where a response or additional analysis was required.

It is simply not true that there is limited secretoglobin expression in your data. Moreover, expression of a single gene does not signify cell type. I have a few suggestions based on my analysis of your Seurat object:

The data need ambient RNA correction. There is background luminal epithelial gene expression across all clusters signifying incomplete washing before barcoding. I would highly suggest using CellBender. I'm guessing this is why your clustering looks a little wonky.

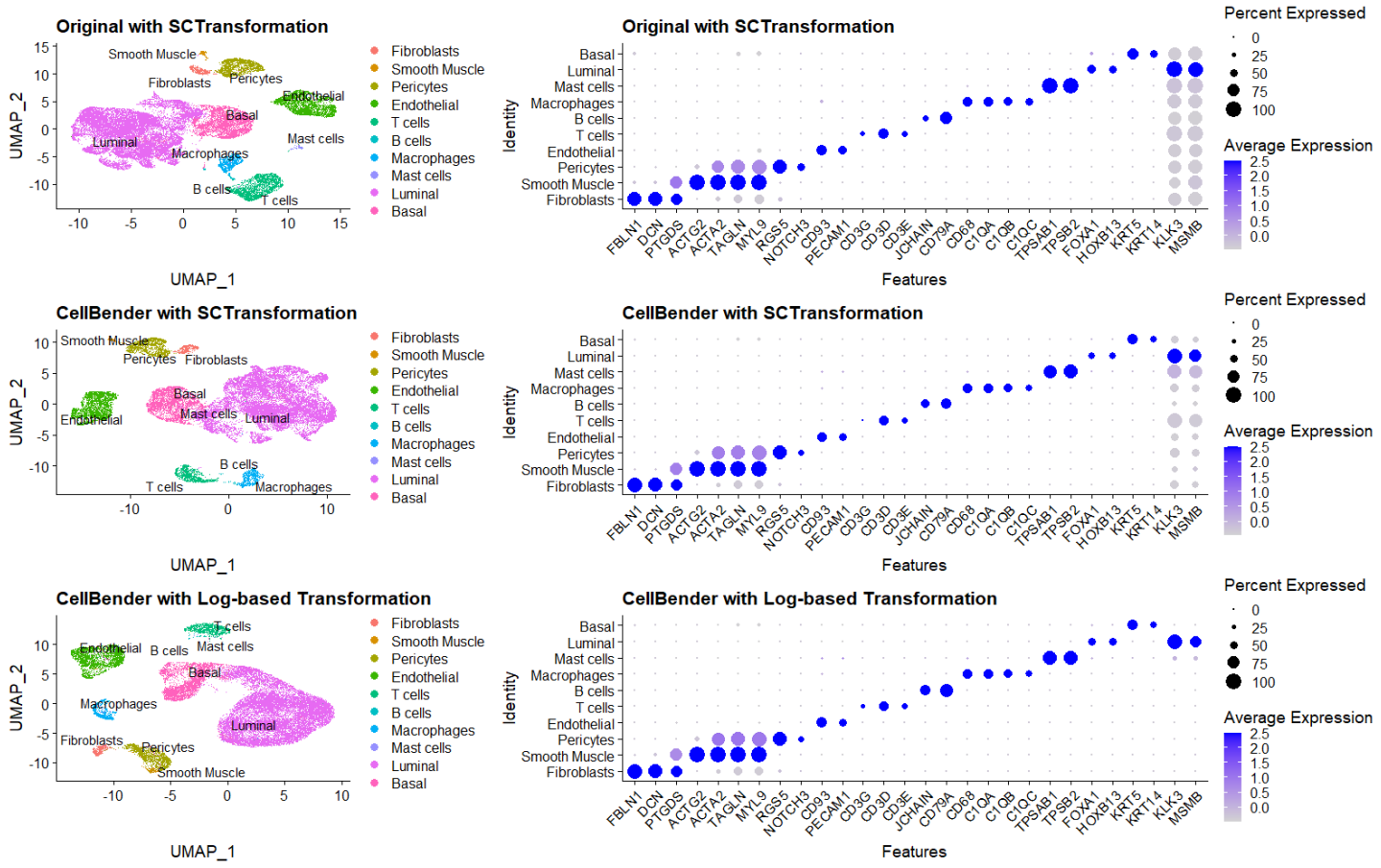



I would also suggest reclustering after subsetting. The following image is from your prostatectomy data where we subset your epithelia and then reclustered. (Of note, we use log transformation instead of SCTransform if you want to try).

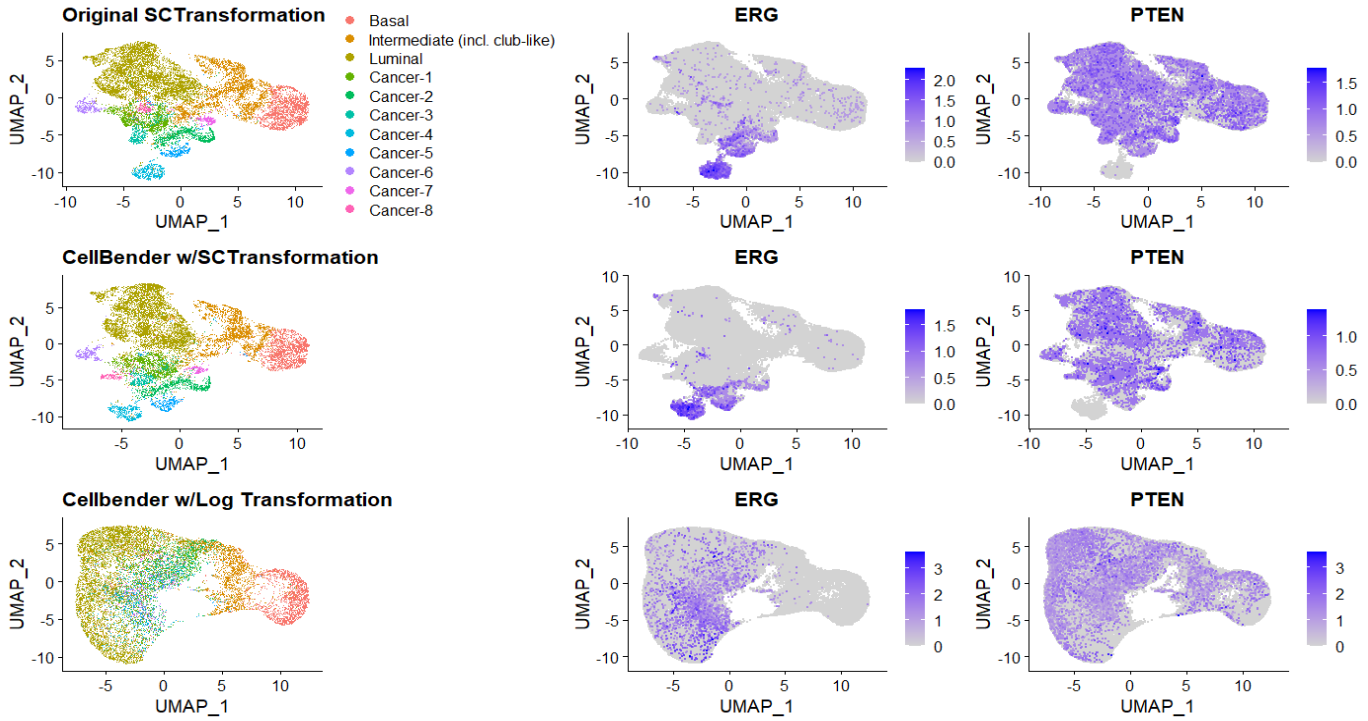



In earlier iterations of our analysis, we previously applied log-based transformation. However, we ultimately elected to normalize with SCTransformation, as variance stabilization using regularized negative binomial regression has been previously shown to outperform log-based normalization methods in variable gene selection, dimensionality reduction, and differential gene expression analysis (PMID: 31870423, 35042561). With respect to CellBender, we agree that this can be a useful package to correct for ambient noise in scRNA-seq data. However, we have also found that while CellBender can be very useful when such ambient noise can interfere with cell type clustering and identification, it can also lead to over-correction and obscuring of cell types, especially those occurring with low frequency. In the case of the current study, the reviewer's observation of background expression of prostate luminal epithelial genes *KLK3* and *MSMB* across the cell clusters is duly noted. Nonetheless, the average expression (z-score scaled) in the dot plot provided by the reviewer confirms that there are significant differences in the relative expression of *KLK3* and *MSMB* in the prostate luminal cells (Z-score > 2) and any non-luminal cell types (Z-score < 0). Consequently, we found that the impact of the ambient expression of genes such as *KLK3* and *MSMB* on cell type-specific clustering was minimal and did not impact our downstream analyses. To illustrate this, we applied CellBender with

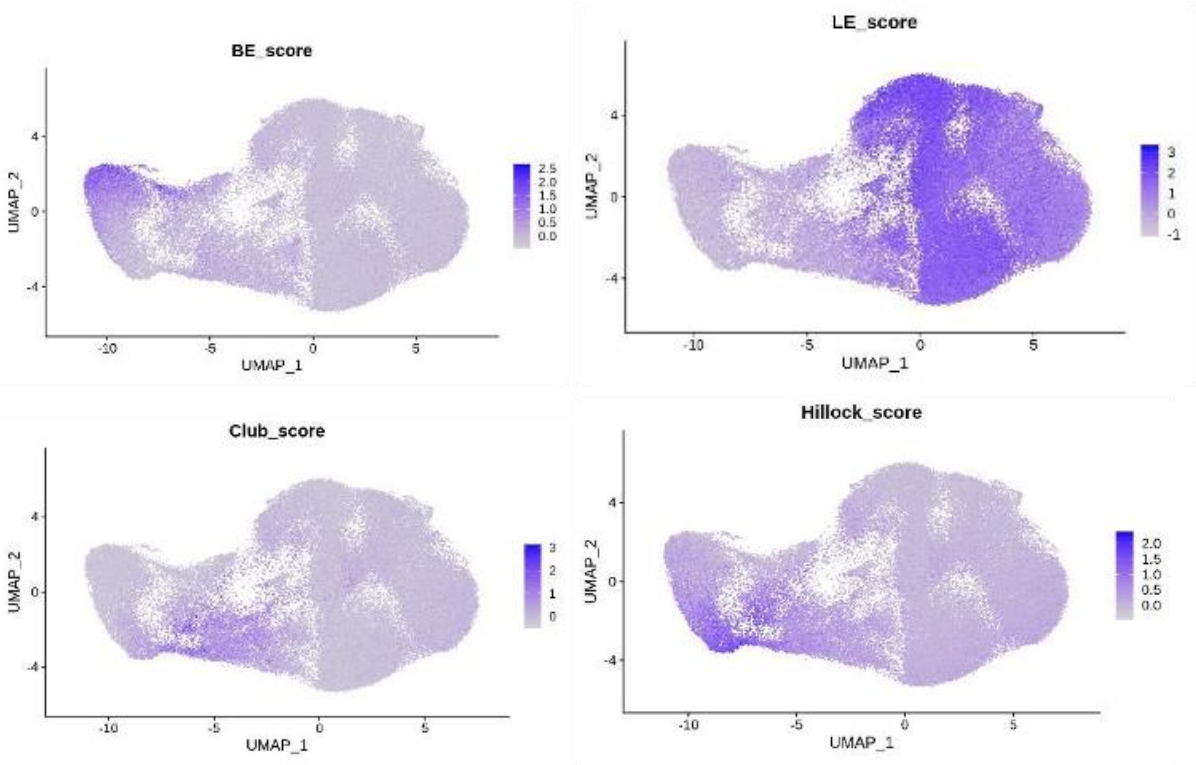
SCTransformation (2nd row) and CellBender with log-based transformation (applying log-based transformation with integration, 3rd row) to our dataset. As shown in our UMAP and dot plot of our original analysis provided below, the marker gene-specific expression (such as prostate luminal specific genes *HOXB13* and *FOXA1*) is cell type-specific, and the cell type-specific clusters are distinct and partitioned in UMAP space (first row). However, when we applied CellBender and transferred the cell type annotations from our original seurat object to the seurat object with CellBender and SCTransformation, we lost distinct B cell and Mast cell clusters (second row), which are certainly known to be present in human prostate. If we applied CellBender with log-based transformation and integration, in addition to losing B cell and mast cell clusters, we also lost partitioning of pericyte and smooth muscle clusters in UMAP space (third row). Based on these analyses, we conclude that our original strategy of data processing and normalization is preferable to retain distinct clustering of rarer cell types.

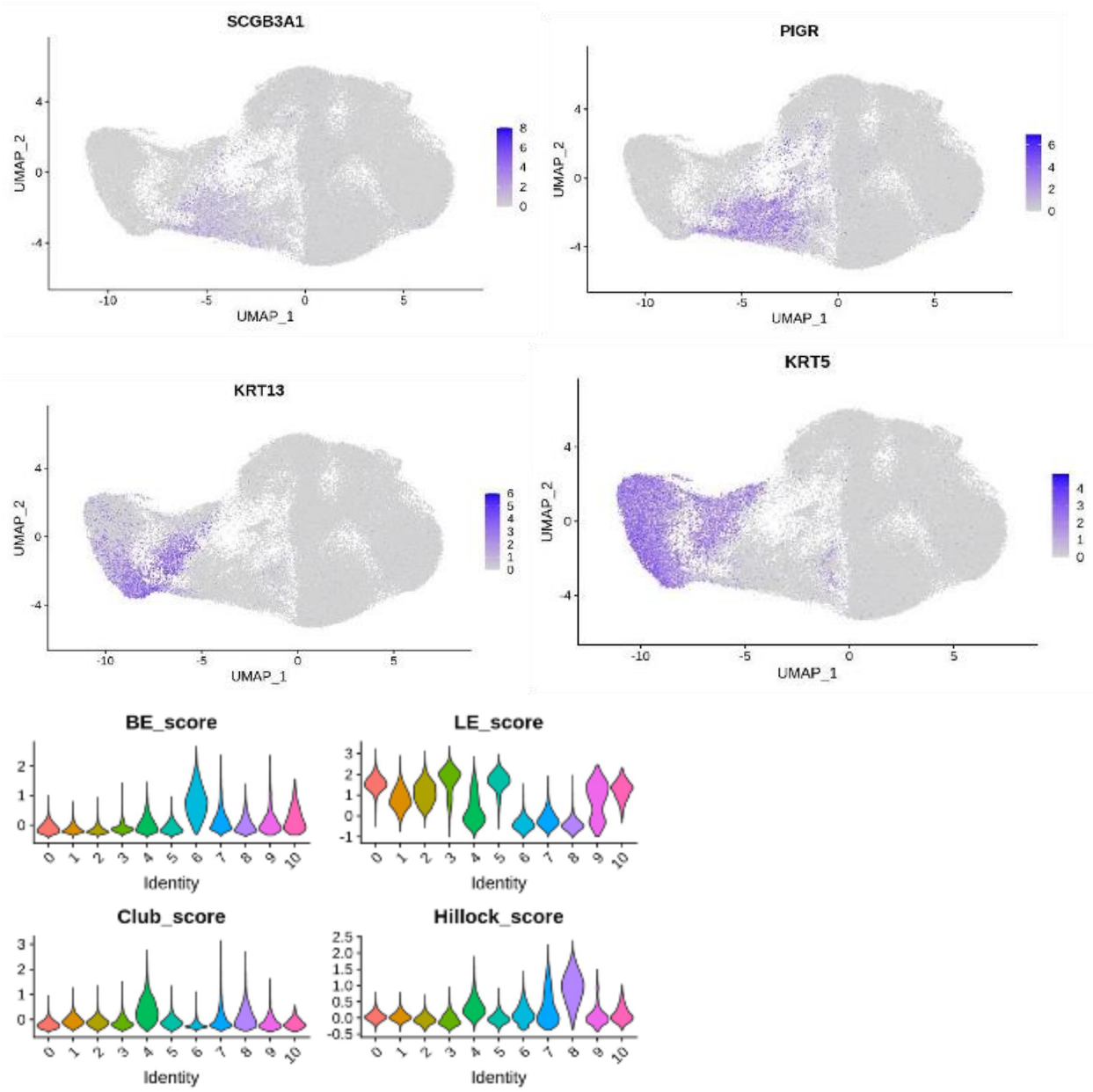



When we subset the epithelial cells, renormalize, and recluster (shown in Figure 2 of our manuscript and described in the methods section), we are able to detect heterogeneous cancer clusters. Inter and intra-tumor heterogeneity is a well-known feature of prostate cancer. In our study, we applied several approaches, including computational (inferCNV) and in situ methods (IHC and CISH staining), to validate our dimensionality reduction and clustering analysis strategy. As suggested by the reviewer, we investigated the use of CellBender, and CellBender with log-based transformation and integration in our subsetted analysis of epithelial cell clusters. A side-by-side comparison of CellBender with SCTransformation vs. SCTransformation (original version) of the epithelial subset shows very similar clustering in UMAP space, with the original ERG-positive/PTEN-negative cluster (Cancer-4) and ERG-positive clusters (Cancer-2, Cancer-5) preserved. However, CellBender with log-based transformation and integration shows a loss of the distinct cancer heterogeneity we originally captured and validated using orthogonal approaches (e.g., ERG gain and PTEN loss). For this cancer-focused study, we have decided to preserve our original version of normalization, dimensionality reduction, and clustering of epithelial cells.




I ran a quick correlation analysis using gene scores from basal, luminal, club and hillock reference data on your data and the results are quite similar to what the Sykes lab published (Figure 2, Nature Comms 2023, 14:663). The OLFM4+ cluster in your image above is hillock and the LTF+ are club. I have no doubt that there exists a spectrum of states for club and hillock in cancer as we and others have observed, but you cannot state that you won't annotate them because you think they should not exist in PZ specimens, especially since others have already published that they do (with in situ validation) (PMID: 36750562, 37550827). I know you are partial to the PIA terminology, but it really isn't appropriate for single cell data unless you have some spatial correlation. In fact, I would argue that PIA is simply a spectrum of luminal to club intermediate states as Franklin Huang suggested (NOT an intermediate between luminal and basal). Your own data suggest so.






It is a stretch of the data to claim that every cell in the cluster you define as ‘intermediate’ belongs to the PIA phenotype. Unless you have spatial evidence linking genotype to phenotype like we (PMID: 34928497) and others (PMID: 36750562, 37550827) provide, then you cannot simply say it’s nuanced. The burden of proof is now on you to show that PIA is not a spectrum of club cell states. Your own data suggest that it is. The transcriptome of the intermediate cluster is nearly identical to club cells from normal prostatic urethra. In fact, intermediate cells are not between luminal and basal or luminal and hillock, but only between luminal and club as we definitively show in PMID: 34928497. No one in the current single cell era only uses single genes to define a cell type. I’m not sure why you are specifically focusing on SCGB1A1 as the definitive marker of club cells as there are many more. Labeling an entire cluster of cells along the spectrum of urothelial club and prostate luminal as a single phenotype (PIA) that includes a heterogeneous mixture is inappropriately simplifying the data to match the desired annotation of your phenotype. This must be corrected.

We agree with the reviewer that a single gene is not always sufficient to define a cell type or state. The choice of marker genes used to describe club-like cells in prostate tissues has been variable across spatial-based studies. For example, the genes LTF, PIGR, OLFM4, SCGB1A1, and SCGB3A1 are the marker genes

described in PMID: 36750562 (Strand), while MMP3, LCN2, and WFD2 are described in PMID: 36750562 (Sykes). In situ staining in PMID: 37550827 (by Sfanos and De Marzo) shows MMP7, CP, PIGR, LTF, and relatively infrequent expression of SCGB1A1. The differences in how various groups describe club-like cells and the selection of marker genes used for single-cell, spatial transcriptomics, and in situ staining to annotate these cells showcase the ongoing ambiguity in the field. We agree with the reviewer that this may reflect a spectrum of intermediate cell states. As for our original annotation “Atrophic/Intermediate”, it was not our intention to claim that all cells in this cluster belong to PIA; hence the umbrella term “atrophy” was used in our original cluster annotation, which includes multiple atrophic subtypes beyond PIA. However, we agree with the reviewer that broader spatial transcriptomic data would be needed to firmly conclude that these cells represent the atrophy cells in PIA or other types of atrophy. We are carrying out a detailed study of PIA and other atrophy lesions using spatial transcriptomic and other integrated modalities and will defer a detailed characterization of subsets of PIA and the cell types present therein for that study – such an analysis is certainly out of the scope of the current paper that is focused on MYC-driven carcinogenesis. We also acknowledge that some of the intermediate cells identified in our study show club-like gene expression as the reviewer has helped to point out in his re-analysis of our data. In order to signify this, and to incorporate the nomenclature used in prior prostate scRNA-seq reports, we have re-labeled the cluster to “Intermediate including club-like” in all of our human data figures and have updated the text in the manuscript to reflect this updated cluster labeling.

Reviewers' Comments:

Reviewer #1:
Remarks to the Author:
Nature communications: NCOMMS-23-46528B
1.The authors need to compare the relative abundance of each cluster type, such as Luminal MYC1, MYC2, at different time points in different stages of progression.

The authors have addressed the comment by including cell abundance plots for different stages, and the signatures have been scored accordingly.

2.The TP53 pathway was significant in enrichment analyse. But the authors hypothesized that PTEN loss could potentially drive MYC activation? They should examine if the alterations co-occur in the accelerating immunosuppressive tumor microenvironment (TME)? PTEN loss demonstrates heterogeneity within the samples.

As suggested, the authors have included a figure displaying the gene signature scores for the P53 and PTEN/mTORC pathways and analyzed their co-occurrence.

3. The changes in the MYC programs throughout the progression in the mouse models (BMPC, HiMYC and Wt 6mo vs 10months) should be described.

The authors have addressed this concern in the revised manuscript by incorporating MYC program scores for various stages.

4.The results identify EGR2, WT1 and other TFs with a high correlation with TGFB receptors. Hence it is not clear why the authors chose EGR4 as the transcription factor of interest in their study? Especially given they observed a high correlation of several other factors with TGFB markers?

The authors' clarification regarding the selection of EGR4 in the manuscript is appreciated. The rationale behind highlighting EGR4, particularly its direct regulation of Timp1 and fibrosis-associated collagens, is well-considered. It's understood that the upregulation of these genes may involve the contribution of multiple transcription factors, and the updated manuscript text effectively addresses this aspect.

5.Incorrect references and typos were corrected

Reviewer #2:
Remarks to the Author:
Truly well done on the paper. I appreciate the willingness to try my suggestions. This will be an excellent addition to the field.

## OVERVIEW

*We thank the reviewers for their constructive feedback and thoughtful recommendations throughout the various iterations of the manuscript. As evidenced by the reviewer's comments (shown below), we are pleased that we have adequately addressed all the concerns and issues raised by reviewers in this most recent version of the manuscript. We also thank Reviewer #2 for complementing the study and highlighting its usefulness to the field. Because the reviewers raised no additional concerns or issues, point-by-point responses were not included below.*

## REVIEWERS' COMMENTS

Reviewer #1 (Remarks to the Author):

Nature communications: NCOMMS-23-46528B
1.The authors need to compare the relative abundance of each cluster type, such as Luminal MYC1, MYC2, at different time points in different stages of progression.

The authors have addressed the comment by including cell abundance plots for different stages, and the signatures have been scored accordingly.

2.The TP53 pathway was significant in enrichment analyse. But the authors hypothesized that PTEN loss could potentially drive MYC activation? They should examine if the alterations co-occur in the accelerating immunosuppressive tumor microenvironment (TME)? PTEN loss demonstrates heterogeneity within the samples.

As suggested, the authors have included a figure displaying the gene signature scores for the P53 and PTEN/mTORC pathways and analyzed their co-occurrence.

3. The changes in the MYC programs throughout the progression in the mouse models (BMPC, HiMYC and Wt 6mo vs 10months) should be described.

The authors have addressed this concern in the revised manuscript by incorporating MYC program scores for various stages.

4.The results identify EGR2, WT1 and other TFs with a high correlation with TGFB receptors. Hence it is not clear why the authors chose EGR4 as the transcription factor of interest in their study? Especially given they observed a high correlation of several other factors with TGFB markers?

The authors' clarification regarding the selection of EGR4 in the manuscript is appreciated. The rationale behind highlighting EGR4, particularly its direct regulation of Timp1 and fibrosis-associated collagens, is well-considered. It's understood that the upregulation of these genes may involve the contribution of multiple transcription factors, and the updated manuscript text effectively addresses this aspect.

5.Incorrect references and typos were corrected

Reviewer #2 (Remarks to the Author):

Truly well done on the paper. I appreciate the willingness to try my suggestions. This will be an excellent addition to the field.