nature research

Peer Review Information

Journal: Nature Genetics
Manuscript Title: A single-cell and spatially resolved atlas of human breast cancers
Corresponding author name(s): Dr Alexander Swarbrick

Reviewer Comments & Decisions:

Decision Letter, initial version:

11th Nov 2020

Dear Alex,

Firstly, I'd like to thank you and your co-authors for your patience during the review process.

Your Article, "An integrated multi-omic cellular atlas of human breast cancers" has now been seen by 3 referees. You will see from their comments below that while they find your work of interest, some important points are raised. We are interested in the possibility of publishing your study in Nature Genetics, but would like to consider your response to these concerns in the form of a revised manuscript before we make a final decision on publication.

Broadly speaking, all three reviewers appear supportive of the timeliness of the paper, and the dataset analysed. However, they all raise issues regarding technical details that we believe should be addressed.

Referee #1 asks for more methodological details and further characterisation of the results. They think that the Gene Modules are more interesting and novel than the scSubtype signature analysis. Their comments on the spatial analysis suggests major improvements are required.
Referee #2 questions the robustness of the immune cell data, and asks for significantly more detail on this. They also request more statistical testing, and also to compare the results presented here to other published data.
Referee #3 appreciates the value of the dataset but asks interesting questions regarding HER2 heterogeneity.

We note that all three referees' reports had some concordant themes: for example, more comparisons to external data (PAM50, METABRIC) and more technical/statistical detail, and further validation of

findings (e.g. using IHC/FISH). We think that reviewer #2's comment on the immune cell enrichment is important and should be fully addressed. Finally, we also note that the reviewers made relatively few comments about the spatial transcriptomics data, suggesting that this part of the manuscript could be improved and better-integrated with the remainder of the analysis.

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

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

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

When revising your manuscript:

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

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

*3) Include a revised version of any required Reporting Summary:
https://www.nature.com/documents/nr-reporting-summary.pdf
It will be available to referees (and, potentially, statisticians) to aid in their evaluation if the manuscript goes back for peer review.
A revised checklist is essential for re-review of the paper.

Please be aware of our guidelines on digital image standards.

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

[REDACTED]

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

We hope to receive your revised manuscript within four to eight weeks. If you cannot send it within

this time, please let us know.

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

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

We look forward to seeing the revised manuscript and thank you for the opportunity to review your work.

Sincerely,

Michael Fletcher, PhD
Associate Editor, Nature Genetics

ORCID: 0000-0003-1589-7087

Referee expertise:

Referee #1: single-cell and spatial methods

Referee #2: breast cancer, immunology

Referee #3: breast cancer, single-cell and spatial methods

Reviewers' Comments:

Reviewer #1:
Remarks to the Author:
The authors have collected a large database and are to be commended for an amazing job. The manuscript not only presents a large dataset, but also provides a comprehensive classification of the subtypes of microenvironment and the modules of expression in cancer cells. Finally, the authors mount an extensive analysis which integrates a previously published large dataset (METABRIC) to generate interesting insights into breast cancer 'ecotypes'. This is an impressive and important work. We do have the following concerns:

It is not clear how the cells in Figure 1G are selected. For example for tumor 3838, figure 1D indicates there are malignant cells present, however this tumor does not appear to have any epithelial cells in 1C. Furthermore, in 1G this tumor is absent from the analysis, suggesting again that there are no

cancer cells. The parts of Figure 1 should be coherent.

The method of Figure 1G is generally not clear even after reading the methods section. In the Tirosh et al method there are cells that are required to be used as a reference. Here this is either different or not described. It seems that what the authors must be doing is using the cells that cluster across patients as a reference (i.e., non-cancer; Fig. 1F) but this is not described at all.

scSubtype is presented as a primary method in the paper whereas their gene modules are much more novel and useful. In Figure 2, the authors mostly analyze the cancer cell transcriptomes using a pseudo-bulk RNAseq method scSubtype to assign the subtypes of individual cells. This method has a circular-logic element to it as the scSubtype gene signatures are based on a "bulk" signature of cancer cells of the tumor which are later shown to be a mixture of the different subtypes. It would instead be more useful to focus directly on the integrative clustering method that truly utilizes the power of the single cell data and expand further on their findings from this analysis. That is, it would be more useful to move the scSubtype method to the supplementary and move the analysis of the gene modules from the supplementary fig S5 to the main figure. Further, the authors should consider better characterizing the gene modules and comparing to subtype signatures (either from scSubtype or bulk PAM50 analysis) to provide a better understanding of the biological relevance of these gene modules. Finally, it would be informative to correlate the presence of these gene modules with patient outcomes using TCGA or METABRIC data. Lastly, in Fig 2E, all 26 samples have not been shown and it would important to include all samples.

In the analysis of lymphocytes, the authors claim to "provide insights into the immunotherapeutic strategies most appropriate for each subtype of disease". The evidence is that across sub-populations receptor-ligand pairs (immune checkpoint molecules) are differentially expressed. However, this is difficult to appraise since a P-value is not provided. Specifically, the authors claim "significantly higher expression" with no statistical support. What should be provided is a comparison between the average expression levels of the gene/pair in each tumor across the subtypes. Moreover, it would be appropriate to substantiate any claim with an independent method such as immunofluorescence, or average expression from published datasets.

Similarly, the authors write that "We observed a substantially reduced proportion of LAM 1:FABP5 cells in the HER2+ tumors". However, no P-value is provided although in the caption it is mentioned that the t-test is significant without mentioning specifics.

In Figure 5, the authors present a spatial transcriptomics analysis. Overall the results are not dramatic as no novel claims are made and the analysis is riddled with unresolved dependencies with other manuscripts: 1. Stereoscope is not a published method and therefore we do not know if it has been through peer-review. It's difficult to assess its validity. For example, what do the numbers in Figure 5B signify. 2. In Figure 5, the authors use an independent dataset (cited as Lundeberg et al); it seems problematic that if this were published the other dataset may not be available. Moreover, there are aspects of the results which are in the supplement that can be better used for the main figure: there are subplots in Figure S10D that would be better shown in the main Figure 5. Finally, the authors state "We earlier predicted that gene modules were associated with unique microenvironments." There is no mention regarding this prediction earlier in the section where the gene modules are identified neither a justification to support this prediction. It is confusing the colors between Fig. 5A and B are different yet correspond to the same regions in the slide.

The ecotype analysis shown in Figure 6 is very interesting, however additional validation could be provided. As the prediction of the ecotype heavily hinges on CIBERSORT, it might be useful to use an independent method of deconvolution to support their findings.

Reviewer #2:
Remarks to the Author:

A major contribution in Wu et al is development of a molecular signature to classify single cell transcriptional profiles of breast cancer into different molecular subtypes, as well as to evaluate infiltrating immune and parenchymal cell clusters, and their associated correlates. For subtyping, the initial strategy employed was to assume PAM50 is the gold standard and then devise scSubtype to maximize similarity to PAM50 subtype assignments. This is a reasonable approach. A summary of the overlap between genes in scSubtype and PAM50 would aid in interpretation of scSubtype.

scSubtype is applied to individual cells by computing the average gene expression of each subtype signature and assigning subtype to the highest score. There is no information provided about the robustness of the assignments for an individual cell. The authors should comment on this, particularly for samples classified as 'discordant' with bulk PAM50 scores.

scSubtype signatures are "validated" by calculating the degree of epithelial cell differentiation and proliferation, both of which are associated with the molecular intrinsic subtypes. This seems a bit circular because the signatures were optimized against the PAM50 intrinsic subtypes. At a minimum, there should be an assessment of the overlap of genes across the various signatures.

True validation of scSubtype signatures in an external single cell dataset would strengthen the claims about its robustness.

With that immune cells occupy (at most) between 10-20% of the cellularity in a breast tumor (based on FACs studies), robustness of the immune cell data is questionable given that scRNA Seq was not performed following any sort of enrichment for CD45 cells by either positive or negative selection. Data should be provided on proportions of cellular compartments, and significance of read depths for lower percentage populations so that interpretation can be appropriately gauged.

CIBERSORT deconvolution: Fig S11 shows correlations between deconvoluted pseudobulk and actual proportions that from 0-1. However, the main text claims that there is good correlation between these measures. Some additional analyses to support the claim is needed, or else the authors need to roll back their interpretation of the findings.

Application of scSubtype to deconvolute METABRIC is an interesting use case for the signatures and provides some additional insights into classification of breast tumors. Curtis et al (PMID: 22522925) describe 10 molecular subtypes of breast cancer. The authors should comment on the relationship of their ecosystem signatures to the previously published subtypes.

Missing details and/or stats in several places, for example:
- Fig 2D (lines 197-199): No indication of whether the differences across subtypes are significant
- Supp Fig 5: No stats to support the comparisons of GM expression across scSubtypes

• The authors claim that tumors in E3 were associated with poor overall survival, consistent with known poor prognosis of Basal-like and highly proliferative tumors (line 515-517), however this does not appear to be supported by data.

Other comments
Introduction mentions data indicating CD8 infiltrations correlate with outcome but fail to mention impact on myeloid subsets also in correlating with outcome, or the significant body of literature identifying myeloid subsets as significant regulators of T cell functionality
'parenchyma cells' in introduction should be defined for the reader. Based on text later in results, this is an odd definition as infiltrating immune cells would also be considered as "parenchymal". Terminology here is not standard and confusing. Do the uthors instead mean to refer to mesenchymal lineage cells?

CD68 was used to identify a myeloid cluster in Fig 1. CD68 only identifies a subset of myeloid cells, predominantly monocytes and macrophages, not DCs, eosinophils, basophils, neutrophils or mast cells. This should be clarified for the reader. In Fig 3, immune cells are subsetted with deeper resolution, but again, myeloid subsetting is for monocytes, macrophages, and DCs only. Were eosinophils, basophils, mast cells not identified? Were they absent? Clarification on this topic is needed. Is their absence a result of bulk scRNAseq without any CD456 enrichment where read depth was limited by paucity of these subsets? Should be discussed. How are subsets identified concordant (or not) with published FACs studies of breast cancer?

The authors identified 6 macrophage clusters. Do any of these identify the macrophage-signature that predicts clinical outcome (in Metabric) reported by Pollard (PMID 30930117) via RNAseq from FACs sorted macrophages?

B cell cluster is identified in Fig 1, but not included in Fig 3. Why? What B cell state is represented in the cluster (Fig 1)? Any indication of plasma cells? B cell clusters are evaluated in Fig S6 but more details on why absent in Fig 3 are needed.

Spatial transcriptomic data in Fig 5 – CAF-T cell neighborhoods were identified. What about myeloid-T cell considering association of LAM2 and T cells in (Fig 3, S10). Are these suppressive myeloid subsets by gene expression, e.g., hi IL-10, ROS, iNOS)?

Methods for scSubtype are reasonable and leverage existing tools in Seurat package, however the description of how the methods were implemented is a bit confusing and should be written for readers not familiar with Seurat package functions.

In general, the figure legends are not very informative. Additional details about what is plotted would aid in interpretation.

Suggest moving Supp Fig 5B to main to provide overview summary of molecular features in each gene module.

Reviewer #3:

Remarks to the Author:
Wu et al. have conducted single cell profiling of 26 primary breast tumours using RNA sequencing and CITE-sequencing. They map major cell types and lineage relationships among cancer cells and the TME. They explore inter and intra-tumoural heterogeneity with respect to the intrinsic subtypes of breast cancer and formulate a single cell subtype classifier. Using these data they also deconvolve bulk RNAseq data and classify tumours into 'ecotypes' based on their cell composition. In addition, they use the Visium 10x platform to conduct analysis of RNA expression with spatial resolution in a small subset of tumours.

This is a well presented and timely report that does a reasonable job of viewing established breast cancer subtypes through a single cell lens. I judge the work to be of moderate novelty and impact; many of the observations are not as original as claimed.

Points to be addressed:
The novelty and significance of some of the findings are frequently over stated. There have been several single cell analyses of breast tumours. What amounts to 'comprehensive' is also highly subjective. The language should be moderated to acknowledge efforts that have gone before.

Established cell signatures were used to annotate cell clusters. A limitation of cell signatures is that they are typically prototypes and do not account for the variability in expression profiles that may arise in a dynamic environment like a tumour. How well did cell signatures match clusters? How did the authors deal with poorly concordant cluster signatures?

The single cell subtyping classifier is a nice addition. However, need it be so complex? After all the key criteria are whether a cell expresses hormone receptors, proliferation markers, HER2 and basal cytokeratins. I would like to see how a much simpler classifier performs in these data, for example one using as few as three genes (see Haibe-Kains et al JNCI 2012).

Several tumours are shown to contain HER2 cells that do not belong to the HER2 subtype. One HER2 tumour (4066) is identified as having almost half of its cells as non-HER2. This could be of major clinical significance in the use of targeted therapy. Can these cells be identified in situ using IHC and/or FISH? If not, how should they be interpreted?

In this report, as in many others, HER2 positive breast cancer is treated as a single subtype. But this does not reflect the clinical reality. Hormone receptor expression stratifies HER2 positive disease into two groups that differ in every meaningful way - prognosis, treatment, patterns of relapse. It would be interesting to know how this distinction impacted the cellular ecosystem of HER2 positive tumours here.

I typically use the term 'parenchyma' or 'parenchymal' to refer to the body of a tissue, and not specifically referring to stromal/endothelial cells as the authors seem to. I suggest reconsidering this.

With respect to 'ecotypes': how useful were the true single cell data in disease stratification? If the authors use prototypical gene signatures (of the TME and/or epithelial cells), deconvolve and call subtypes again, is there a difference in outcome prediction? How well do the subtypes compare to ecotypes and PAM50?

Pseudotime analyses are used to interpret stromal and endothelial cell profiles. Analyses of this sort

should be conducted cautiously because they will always return some sort of answer. What leads the authors to believe that these cell types do in fact differentiate in this manner? Other forms of differentiation, from mesenchymal stem cells for example, are possible.

Author Rebuttal to Initial comments

Wu et. al.
Detailed responses and revisions

SUMMARY OF REVISIONS:

We thank the reviewers and the editors for their valuable time and effort committed towards improving our manuscript. We have been able to address the comments and questions raised with new analysis or clarifications of our results or methods. We believe that these changes have significantly improved the manuscript. We have used tracked changes in the manuscript and describe them below in our detailed point-by-point response.

RESPONSE TO EDITORIAL COMMENTS AND OTHER MAJOR CHANGES:

Referee #1 asks for more methodological details and further characterisation of the results. They think that the Gene Modules are more interesting and novel than the scSubtype signature analysis.

Response: We appreciate the referees' comments, which have allowed us to improve the manuscript. We have added substantial methodological details to the results and methods sections (see response to referee's comments 1.1 & 1.2 below). We agree that the gene module results are interesting and have moved data into the main figure to bolster that analysis. However, we also think the scSubtype work is an important complement to it and a link to established clinically-relevant methods of bulk tumour subtyping. Referees 2 & 3 also acknowledge the importance of scSubtype (see 2.1, 3.3) and so we respectfully suggest that it remains in the main body of the manuscript.

Their comments on the spatial analysis suggests major improvements are required.

Response: We have addressed all of the referee's comments with clarifications and revisions of the spatial analysis. We note that the method used for deconvolution is now published, and that our findings hold-up regardless of whether we use the additional Her2+ dataset (see 1.6).

Referee #2 questions the robustness of the immune cell data, and asks for significantly more detail on this. They also request more statistical testing, and also to compare the results presented here to other published data.

Response: We have provided substantial additional details on the immune analysis, both in the manuscript and in additional data for the referees. We have conducted additional statistical analyses, which largely support our conclusions (see 2.1, 2.2). Finally as suggested by the referee, we have conducted new analysis of myeloid-T cell interactions in the single cell and spatial datasets, which reveals new insights when integrated with the literature suggested by the referee (2.12, 2.13).

Referee #3 appreciates the value of the dataset but asks interesting questions regarding HER2 heterogeneity.

Response: Indeed the referee makes important points about heterogeneity of HER2+ disease. We have clarified our nomenclature of HER2+ disease and conducted further analysis of the association of HER2+ disease subsets with ecotypes, as requested (see 3.4, 3.5).

We note that all three referees' reports had some concordant themes: for example, more comparisons to external data (PAM50, METABRIC) and more technical/statistical detail, and further validation of findings (e.g. using IHC/FISH).

Response: We appreciate this feedback and have included substantial comparisons to external data when appropriate (see 1.3c, 1.4, 2.1, 2.4, 2.7). We have provided more detailed statistical analyses, which mostly support our claims (see 1.4, 1.5, 2.1, 2.2, 2.8, 2.9, 2.12, 3.2). In cases where they do not support our conclusions, we have altered them accordingly. We have not conducted additional experimental IHC or FISH experiments due to time constraints and limitations in validating highly multiplexed data using low-plex methodology. However, we have provided supporting evidence from analysis of published datasets.

We think that reviewer #2's comment on the immune cell enrichment is important and should be fully addressed.

Response: We agree and have provided an extensive response to this question. In short, the single cell technology used in this manuscript preferentially captures immune cells. In fact, the majority (50.3%) of filtered cells in the dataset are immune. For this reason, we consider this dataset to be well suited to immune analysis. We also provide further metrics to demonstrate the quality of the immune cell data generated (see 2.5, 2.12).

Finally, we also note that the reviewers made relatively few comments about the spatial transcriptomics data, suggesting that this part of the manuscript could be improved and better-integrated with the remainder of the analysis.

Response: We appreciate this request and have improved our integration of the single cell and spatial transcriptomics. We include additional text linking the rationale for spatial analysis back to the single cell data (see 1.6). We have also conducted new analyses of macrophage function that uses the single cell- and spatial-transcriptomics datasets to identify evidence for T cell suppression by our novel macrophage subsets (see 2.13).

<u>Additional reanalysis of ecotypes</u>

In response to questions from all 3 reviewers on ecotyping (See 1.7, 2.6, 2.7 below), we conducted extensive additional ecotype analysis, including comparison to results obtained from an independent method of deconvolution, DWLS. These show that the ecotyping results are robust and not an artefact of the Cibertsortx method nor overly influenced by cell types with lower deconvolution accuracy. Furthermore, we provide new analysis of the stratification of Her2+ tumours by ecotype and the relationship of

ecotypes with the genomic IntCluster classifier. These analyses have improved confidence in ecotyping and support the conclusion that ecotypes are not simply surrogates of previously-identified transcriptomic or genomic subtypes.

POINT BY POINT RESPONSES TO REVIEWERS:

Reviewer #1:

The authors have collected a large database and are to be commended for an amazing job. The manuscript not only presents a large dataset, but also provides a comprehensive classification of the subtypes of microenvironment and the modules of expression in cancer cells. Finally, the authors mount an extensive analysis which integrates a previously published large dataset (METABRIC) to generate interesting insights into breast cancer 'ecotypes'. This is an impressive and important work. We do have the following concerns:

1.1 It is not clear how the cells in Figure 1G are selected. For example for tumor 3838, figure 1D indicates there are malignant cells present, however this tumor does not appear to have any epithelial cells in 1C. Furthermore, in 1G this tumor is absent from the analysis, suggesting again that there are no cancer cells. The parts of Figure 1 should be coherent.

> **Response**: We'd like to thank the reviewer for pointing out this lack of coherency. This has arisen due to the way in which we selected samples for detailed epithelial cell analysis. As we carried out 'within-sample' analysis for all steps related to the epithelial cell analyses (i.e., inferCNV, integrative clustering to identify gene-modules, and scSubtype) we removed all samples with less than 200 epithelial cells from our analysis (samples CID3838, CID4398, CID3946 and CID4040). As Fig 1A-C shows all the cells in our final, filtered dataset, the epithelial cells from these 4 samples are not included in this figure. Our intention with figures 1D-F was to highlight the areas of overlap between cells across all of our samples (1D) and with respect to clinical subtype (1E) and whether the cells had been identified as neoplastic using inferCNV (1F).
>
> **Revision**: To improve coherency, we have re-plotted these figures with the epithelial cells from these 4 samples removed. It follows that this is also why these samples (including CID3838) were not included in the figure 1G related to neoplastic cells identified using inferCNV. We have also clarified this point in the methods, page 31 and lines 4-5.
> We also noticed that sample CID4398 was mistakenly included in the inferCNV heatmap of neoplastic cells, shown in fig 1G. We have updated figure 1G and removed sample CID4398.

1.2 The method of Figure 1G is generally not clear even after reading the methods section. In the Tirosh et al method there are cells that are required to be used as a reference. Here this is either different or not described. It seems that what the authors must be doing is using the cells that cluster across patients as a reference (i.e., non-cancer; Fig. 1F) but this is not described at all.

> **Response**: We thank the reviewer for correctly pointing out that this important detail of our implementation of the inferCNV was left out of the methods. To clarify, the reference cells were all patient-matched stromal cells (except for

CAFs, as there is a possibility that some of these may be cancer cells with a mesenchymal phenotype). All epithelial cells were used as the observations.

Revision: We have added this to the text at page 30 lines 14-15.

1.3a scSubtype is presented as a primary method in the paper whereas their gene modules are much more novel and useful. In Figure 2, the authors mostly analyze the cancer cell transcriptomes using a pseudo-bulk RNAseq method scSubtype to assign the subtypes of individual cells. This method has a circular-logic element to it as the scSubtype gene signatures are based on a "bulk" signature of cancer cells of the tumor which are later shown to be a mixture of the different subtypes. It would instead be more useful to focus directly on the integrative clustering method that truly utilizes the power of the single cell data and expand further on their findings from this analysis. That is, it would be more useful to move the scSubtype method to the supplementary and move the analysis of the gene modules from the supplementary fig S5 to the main figure.

Response: We thank the reviewer for their interest and suggestions regarding the breast cancer gene modules. While we agree with the comments regarding a "circular-logic element" of scSubtype, we would like to reiterate that this was circular by design. The intrinsic subtypes are a very well accepted classifier of breast cancer biology and a clinically-actionable predictor of risk. Our objective was to apply the 'bulk' pam50 classifier at single cell resolution. In this way we link the observed neoplastic heterogeneity to well-accepted breast cancer phenotypes with clinically-relevant associations. To achieve this we made a necessary assumption that the bulk subtype of a tumour will match the most frequent neoplastic cell subtype. This work is complemented by the gene module work which takes a data-driven approach to finding the most frequent ontologies associated with intra-tumoral heterogeneity.

Revision: We respectfully suggest that scSubtype remains in the main figure however, we do agree that Figure S5B would be more appropriately placed in the main figure, and have updated this in the revised manuscript.

1.3b Further, the authors should consider better characterizing the gene modules and comparing to subtype signatures (either from scSubtype or bulk PAM50 analysis) to provide a better understanding of the biological relevance of these gene modules.

Response: We thank the reviewer for the interest in our gene-module based approach. With regards to comparing the gene-modules to scSubtype signatures, this analysis is specifically shown in figure S5F of the manuscript, where the scSubtype score enrichments, for the cells assigned to each gene-module are plotted.

With regards to further characterization of the biological relevance of the gene-modules, we curated a set of 78 gene-signatures related to breast cancer biology and compared them to each of the gene-modules. Briefly, we calculated a score (using AUCell) for each gene-signature (including the 7 gene-modules), in each neoplastic cell. We then generated a pair-wise correlation matrix using these per-cell enrichment scores and extracted those with highest correlations (top-10) to each gene-module (Figure R1).

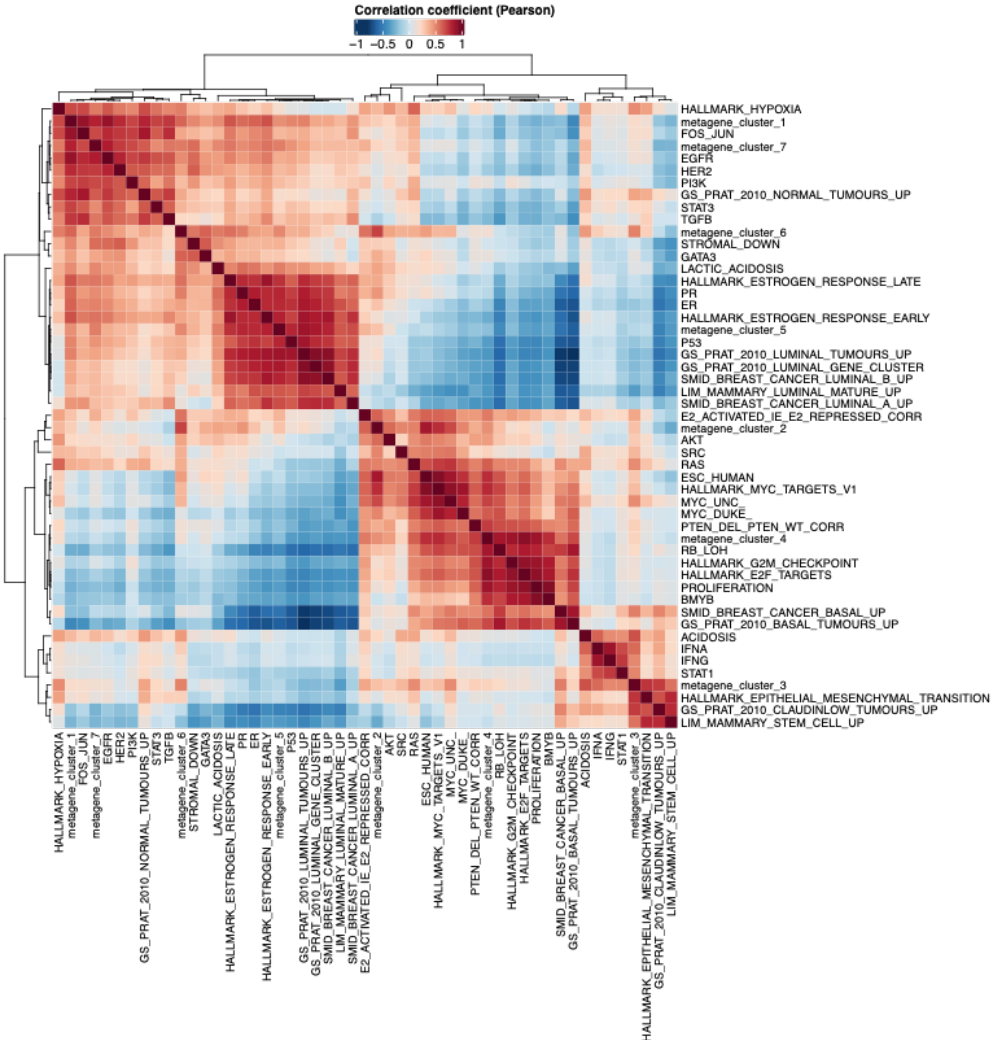

Figure R1: Pearson correlation matrix of AUCell scores from all neoplastic cells. The top-10 correlated gene-sets to each of the 7 gene-modules is shown.

A number of gene-set clusters are apparent, broadly related to individual gene-modules. For instance: GM1 is related to FOS_JUN, EGFR and PI3K signalling and luminal tumours; GM3 is related to EMT, Claudin-low breast cancers and interferon signalling; GM4 is correlated with proliferation and basal-like breast cancer tumours; GM5 is a strongly related to luminal breast cancers; and GM6 forms a small cluster related to GATA3 signalling.

This analysis reinforces our findings that a number of our gene-modules align with breast cancer subtypes, at both bulk tumour and cellular level (scSubtype) and suggests other disease pathways that may be of relevance for understanding the role that these gene-modules play in driving tumour heterogeneity.

Revisions: Again we thank the reviewer for interest in this analysis. In the interests of space, we have not included the additional gene signature analysis in the revised manuscript. However, we would be willing to incorporate if the reviewer and editor consider it an important addition.

1.3c Finally, it would be informative to correlate the presence of these gene modules with patient outcomes using TCGA or METABRIC data.

Response: As requested by the reviewer, we have investigated the prognostic association of the seven breast cancer gene modules using Kaplan Meier analysis in the METABRIC[1] and TCGA[2] cohorts, as well as a third independent cohort by Harrell *et al.*[3]. This revealed two survival trends that were consistently identified across multiple cohorts. As seen in the survival curves below (Figure R2), we found that patients with a high gene module 4 (GM4) score had worse overall survival (red line). As GM4 was enriched for proliferation and cell cycle related pathways, this result is expected from the known prognostic association of proliferation markers and signatures[4]. Furthermore, patients with a high GM5 score had better overall survival (red line). As GM5 was enriched for ER gene pathways, this result is consistent with the better prognostic outcomes of luminal breast cancers. No other gene modules revealed significant survival trends that were detected across two or more cohorts (data not shown).

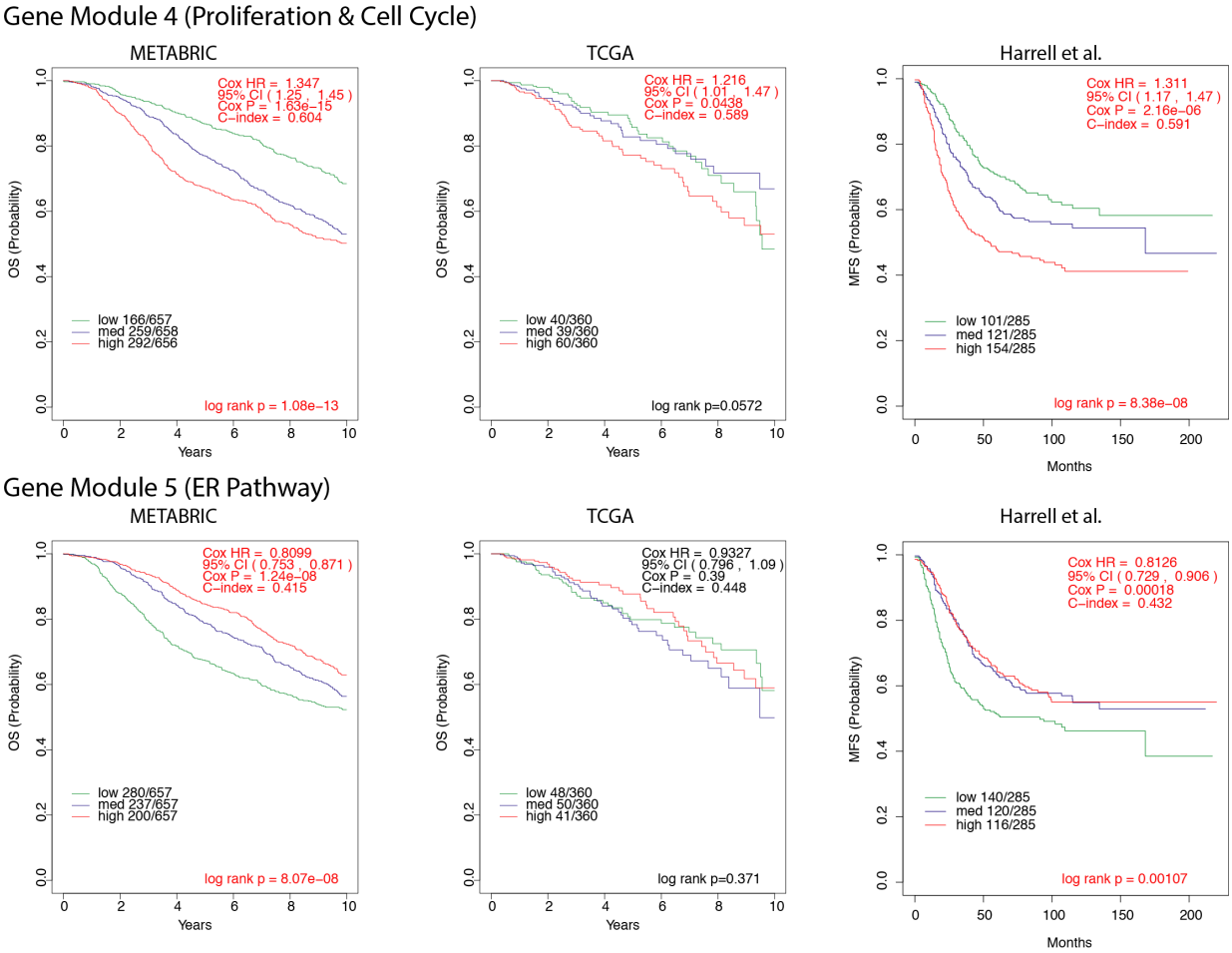

Figure R2. Kaplan Meier curves showing the prognostic association between gene modules 4 (proliferation and cell cycle) and 5 (ER pathway) across three independent datasets from the METABRIC, TCGA and Harrell *et al.* studies. Univariate Kaplan–Meier analysis with Cox–Mantel log-rank test was performed using the survival package in R. Patients were stratified into 2 or 3 equal sized groups based on the rank order of the gene expression module. Overall Survival or Relapse Free Survival were used as the end points.

While we agree that correlating these gene modules with patient outcomes is interesting, we believe that directly examining the prognostic associations of single-cell breast cancer signatures in bulk tumor cohorts can be confounded by neoplastic heterogeneity and by non-malignant cell types that have overlapping transcriptional profiles. To this regard, we were not surprised to find no novel survival trends in other gene modules. For instance, GM3 was enriched for pathways related to immune regulation (e.g. interferon response and antigen presentation) and EMT, which can be confounded in bulk studies by infiltrating immune cells and mesenchymal cells, respectively.

Revisions: As the association of GM4 and GM5 with prognosis are as expected, we have decided to not include these results in our manuscript in the interest of space. However, we are open to incorporating it at the discretion of the reviewer and editor.

1.3d Lastly, in Fig 2E, all 26 samples have not been shown and it would important to include all samples.

Response: It is correct that not all 26 samples were included in fig 2 (we assume the reviewer is referring to figure panels 2B, F and G) as only 21 samples had sufficient neoplastic cells to be included in the integrative clustering analysis. For the identification of gene-modules we removed any samples with less than 50 neoplastic cells. This is why sample CID44041 is not included in figure 2F and 2G. For scSubtype, we applied a slightly more stringent cell filtering threshold of at least 150 neoplastic cells, which is why sample CID3586 was additionally removed from scSubtype analysis and is not included in figure 2B.

Revisions: We have updated the text and methods to improve the clarity of these extra neoplastic cell based, sample filtering steps (see: page 8, lines 2-3; page 10, line 12; page 32, line 12; and page 34, line 14)

1.4 In the analysis of lymphocytes, the authors claim to "provide insights into the immunotherapeutic strategies most appropriate for each subtype of disease". The evidence is that across sub-populations receptor-ligand pairs (immune checkpoint molecules) are differentially expressed. However, this is difficult to appraise since a P-value is not provided. Specifically, the authors claim "significantly higher expression" with no statistical support. What should be provided is a comparison between the average expression levels of the gene/pair in each tumor across the subtypes. Moreover, it would be appropriate to substantiate any claim with an independent method such as immunofluorescence, or average expression from published datasets.

Response: We apologise to the reviewer for the missing statistics for ligand-receptor comparisons across subtypes in our manuscript. We acknowledge that statistical support is critical to support these claims. To address this, we have included the p-values from our differential gene expression analysis across subtypes for each gene using the MAST method (Figure R3; Figure R4). This reveals that the majority of changes reported are significant.

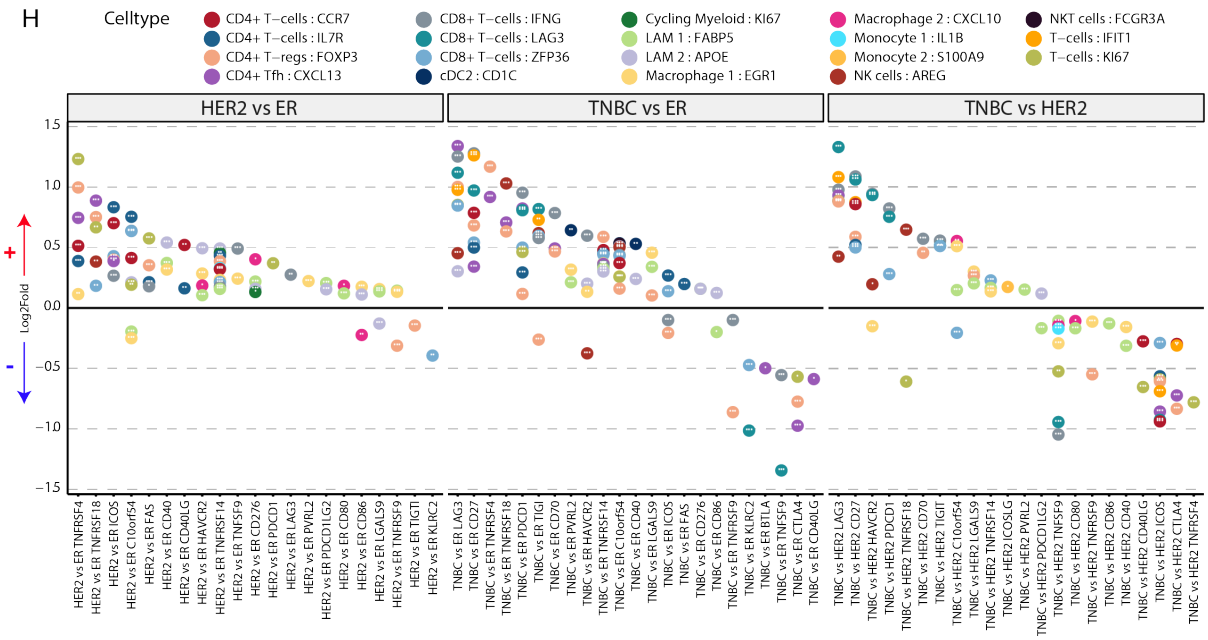

Figure R3. Log2fold differential gene expression change of surface immunomodulatory receptor genes between BrCa subtypes, stratified by immune cell subset. Only statistically-significant differences using a pair-wise comparison are shown (bonferroni corrected values using MAST). P-values denoted by asterisks: *p < 0.05, p < 0.01, *p < 0.001 and ****p < 0.0001.

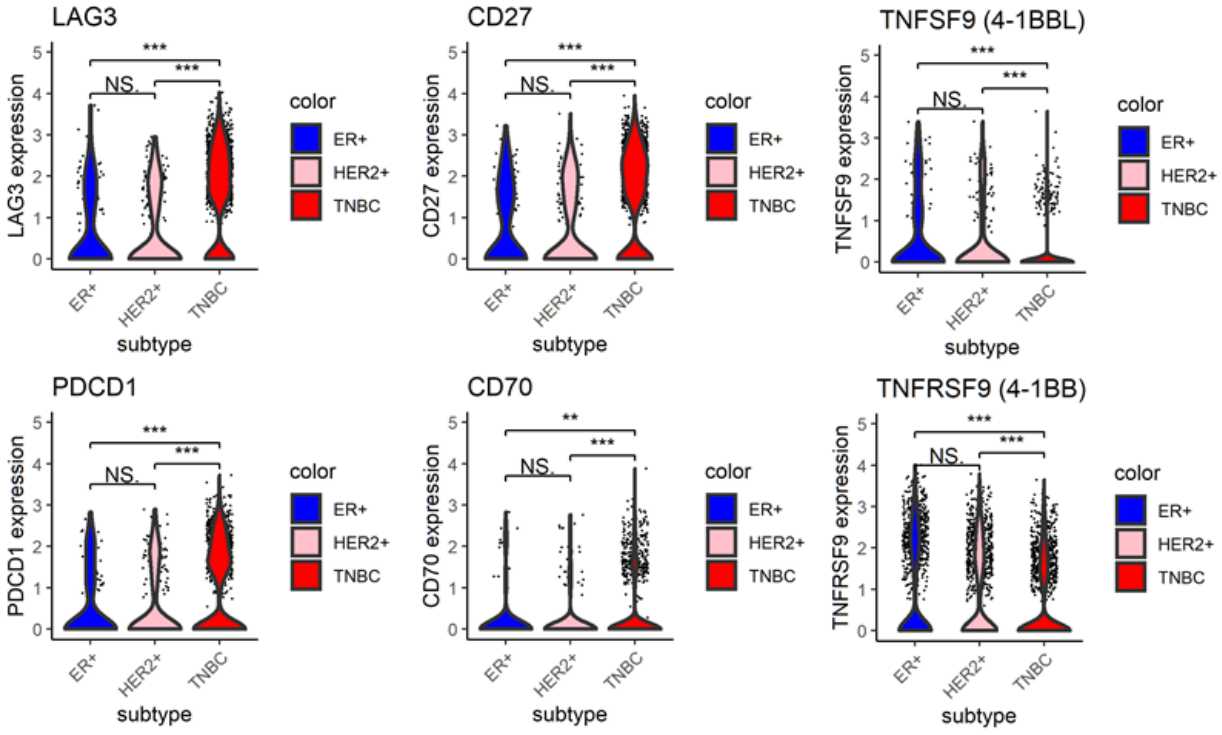

Figure R4. Violin plots showing the log-normalised expression of immunoregulatory genes LAG3, CD27, PDCD1, CD70, for the T-cell : LAG3 cluster across clinical BrCa subtypes analysed by scRNA-Seq. Statistical significance was computed using paired student t-test. P-values denoted by asterisks: *p < 0.05, p < 0.01, *p < 0.001 and ****p < 0.0001.

Furthermore, we agree with the referee that support from an independent dataset would substantiate these claims. We have investigated the expression of these immuno-modulatory genes in published datasets such as METABRIC and TCGA (Figure R5). While there are obvious caveats of exploring single-cell expression in bulk tumor data, we do observe a consistent enrichment of markers in Basal-like breast cancers which were up-regulated in CD8+ T-cells from TNBC tumors in our scRNA-Seq cohort. As seen in the expression plots below, this includes *CD70*, *PDCD1* (PD-1), LAG3 and *CD27*. In the case of PD-1, our findings are also consistent with our recent immunofluorescence study [5], which shows that TNBC have a higher infiltration of PD-1+ T-cells compared to luminal BrCa.

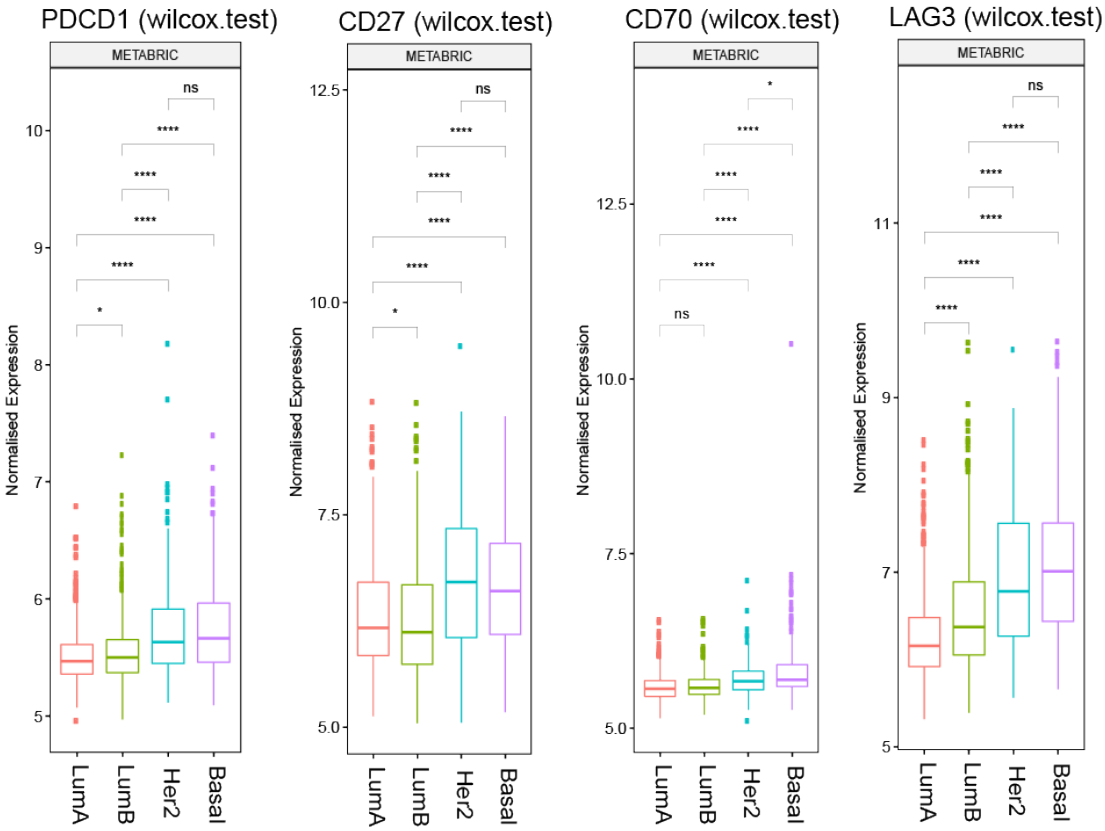

Figure R5. Normalised expression of immunoregulatory genes *PDCD1, CD27* and *CD70* in the METABRIC and TCGA cohorts. Tumors are grouped by PAM50 subtype classification. Statistical significance was computed using a wilcoxon test. P-values denoted by asterisks: *p < 0.05, p < 0.01, *p < 0.001 and ****p < 0.0001.

For immunoregulatory genes 4-1BBL (*TNFSF9*) and (*TNFRSF9*/CD137) found to be enriched in *LAG3*/c8 cells from ER+ and HER2+ patients, these were not differentially expressed in METABRIC and TCGA. This may be a limitation of estimating single-cell expression in bulk tumors, nonetheless, claims were removed from manuscript.

Revisions: P-values for single cell expression claims have now been included in revised manuscript. Specifically, the following modifications were made to address reviewers comments:

- We have added statistical significance values for the expression of dysfunctional and cytotoxic gene signature across all T-cell / ILC clusters. Previous Figure S6C is now relocated to Fig S7B.
- We have included a figure to demonstrate that dysfunctional T-cell gene expression scores are statistically different across cancer subtypes for cells of CD8+ T : LAG3 and CD8+ T : IFNG clusters (Fig S7C).
- We have included a figure to visualise Log2fold expression of all statistically significant receptor-ligand genes described in Figure 3I, stratified by celltype, across breast cancer subtypes (Fig S7D)
- Violin plots with statistical analysis for each gene claimed in text; PD1, CD27, CD70, LAG (Fig S7E).
- The expression of markers PDCD1 (PD-1), CD27, LAG3 and CD70 in the METABRIC cohort across breast cancer subtypes to support the single-cell differential gene expression findings (Fig S7F). This is further discussed in the text page 13 lines 8-11.
- We have also discussed a recent immunofluorescence study by Millar et al 2020[5] in the text page 13 lines 11-12, which supports claims of increased PD1 expression in tissue resident T-cells in TNBC in comparison to HER2+/ER+.

1.5 Similarly, the authors write that "We observed a substantially reduced proportion of LAM 1:FABP5 cells in the HER2+ tumors". However, no P-value is provided although in the caption it is mentioned that the t-test is significant without mentioning specifics.

Response: We apologise to the reviewer for this oversight in our manuscript. Descriptions of the p-values in the figure legends were mistakenly removed through editing and have now been re-incorporated into the figure legend of Figure 3 of the revised manuscript.

1.6 In Figure 5, the authors present a spatial transcriptomics analysis. Overall the results are not dramatic as no novel claims are made and the analysis is riddled with unresolved dependencies with other manuscripts: 1. Stereoscope is not a published method and therefore we do not know if it has been through peer-review. It's difficult to assess its validity. For example, what do the numbers in Figure 5B signify. 2. In Figure 5, the authors use an independent dataset (cited as Lundeberg et al); it seems problematic that if this were published the other dataset may not be available. Moreover, there are aspects of the results which are in the supplement that can be better used for the main figure: there are subplots in Figure S10D that would be better shown in the main Figure 5. Finally, the authors state "We earlier predicted that gene modules were associated with unique microenvironments." There is no mention regarding this prediction earlier in the section where the gene modules are identified neither a justification to support this prediction. It is confusing the colors between Fig. 5A and B are different yet correspond to the same regions in the slide.

Response: We acknowledge the referee's concerns regarding the use of unpublished methods and datasets in our study. In regards to stereoscope method we applied for spatial deconvolution, this method has now been published [6].

While we acknowledge that the Lundeberg *et al.* dataset used in our study is unpublished, we would like to highlight that our findings of stromal-immune co-localisations hold true without the inclusion of these additional HER2+ spatial transcriptomics datasets. This is evident in the heatmap in Figure 5H, where Pearson Correlations between stromal-immune cell types are summarised across the four TNBC and two ER+ spatial datasets exclusive to this study, and the seven HER2+ spatial datasets from the Lundeberg *et al.* study. For example, this shows that negative correlations between iCAFs and myCAFs, and positive correlations between iCAFs and CD4+ and CD8+ T-cells, hold true across a majority of cases from both studies. The use of this independent dataset provides important additional support for our biological findings across the Her2 subtype of breast cancer. A preprint describing this data is available online[7] and this manuscript is currently under revision. As such we expect the timing of its publication to coincide with ours. Dr Lundeberg makes all published spatial transcriptomics datasets available online (https://www.spatialresearch.org/resources-published-datasets/).

Revisions:

- We have clarified the data in Fig 5B where scaled deconvolution abundances, as estimated using stereoscope, are shown. This has been updated in the figure legend for Figure 5 of the revised manuscript.
- We thank the referee for the suggestion that Figure S10D could be better shown in the main figure. We agree with this, and this has now been updated in Figure 5C-D of the revised manuscript.
- We also apologise that our statement "We earlier predicted that gene modules were associated with unique microenvironments" was not clear. This is now clarified in the text p19 lines 4-6 of the revised manuscript.
- As requested, the colours related to pathology regions are now consistent between Fig. 5A and Fig. 5B of the revised manuscript.

1.7 The ecotype analysis shown in Figure 6 is very interesting, however additional validation could be provided. As the prediction of the ecotype heavily hinges on CIBERSORT, it might be useful to use an independent method of deconvolution to support their findings.

Response: We thank the reviewer for this suggestion and agree that a second method of deconvolution could be used to support our conclusions. To address this we have used an independent method (DWLS[8]) to provide additional validation to our CIBERSORTx based ecotype. Similar to CIBERSORTx, DWLS is also specifically designed to utilise single-cell data and performed well in a recent benchmark[9]. However, DWLS uses a different method, dampened weighted least squares, so is not redundant to CIBERSORT, which uses support vector regression. Comparison between the deconvolution results showed no significant difference (wilcox p-value >0.05) in correlation coefficients, between actual and predicted cell-type fractions, using pseudo-bulk samples (Figure R6).

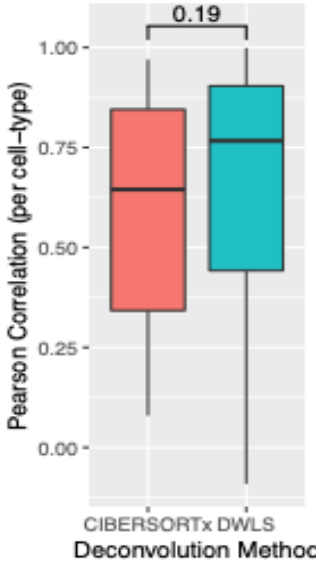

Figure R6: Boxplots comparing the Pearson correlations between the actual cell-fractions captured by scRNA-Seq and those predicted by CIBERSORTx and DWLS, using pseudo-bulk expression profiles. Wilcox p-value.

This indicated that both methods were performing comparably across all cell-types.

We next compared the performance of DWLS when applied to the METABRIC cohort. Similar to fig 6B, we assessed the performance of the DWLS method when predicting the scSubtype and cycling cell abundances across the PAM50 tumour subtypes. This resulted in similar performance to CIBERSORTx, where: (i) each scSubtype cell-type was significantly enriched in tumours with associated bulk PAM50 subtype classifications; and (ii) cycling cells were enriched in the more proliferative tumours, namely basal-like, Her2E and LumB. An exception to this was seen for the LumA_SC cell-type, where only a modest difference (p=0.034) was seen between LumA and B tumours and no difference (p=0.86) between LumA and Her2E tumours (Figure R7). This indicates that CIBERSORTx performs slightly better (see Figure 6B in the manuscript) in the deconvolution of scSubtype defined cell-types, in particular LumA tumours.

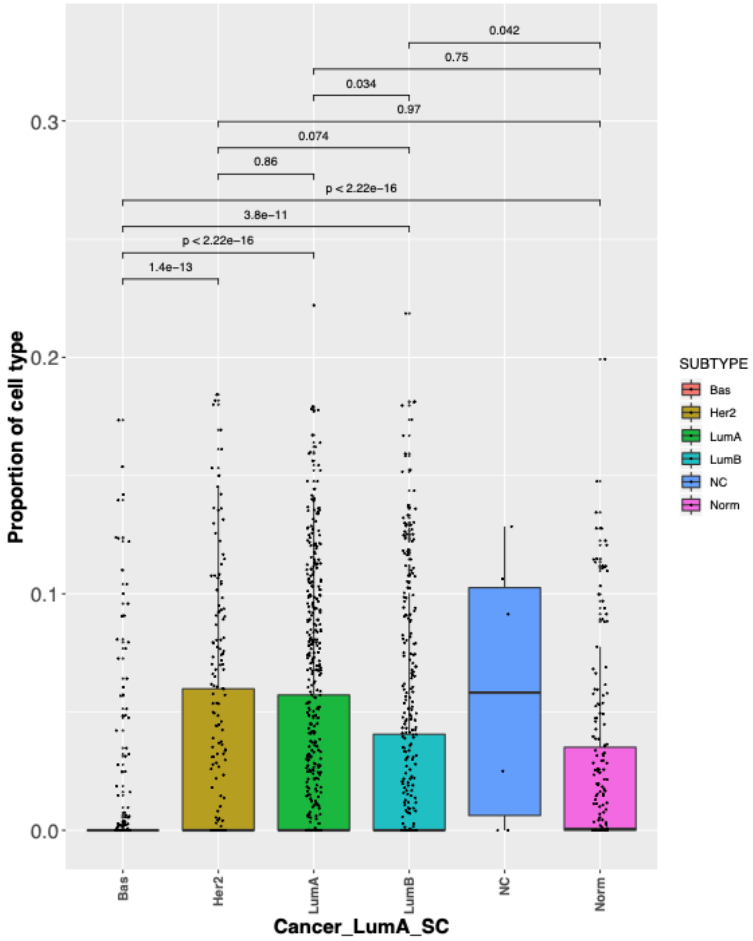

Figure R7: Boxplot comparing the DWLS predicted cell-fractions of LumA_SC cells in each METABRIC patient tumor, stratified by PAM50 subtypes. p-values calculated using Wilcox test.

Next we approached the analysis by comparing the concordance between nine tumour ecotypes generated from CIBERSORTx versus DWLS calculated cell-fractions. Our hypothesis being that if the deconvolution approaches were performing similarly they should show pairs of ecotypes with substantial overlap of tumours. The Table R1 below, listing the nine ecotype pairs with the highest average overlap, shows that this was broadly the case. Importantly, all nine of the CIBERSORTx ecotypes are part of this top-ranked list.

cibersortx ecotype	dwls ecotype	sample overlap	sample overlap (%)
E1	E9	113	43.1
E2	E3	221	79.5
E3	E6	157	82.3
E4	E2	185	75.7
E5	E8	64	34.3
E6	E1	58	28.1
E7	E4	145	70.6
E8	E7	104	45.8
9	E7	63	32.8

Table R1: Tumour ecotype overlaps between using CIBERSORTx and DWLS.

Finally, we used these overlapping tumours to define nine “common” ecotypes consisting of 1110 tumours whose ecotype calls were concordant between the two methods. To test if these common ecotypes showed similar properties to the original CIBERSORTx ecotypes we compared their molecular subtype (Figure R8) and cell-type compositions (Figure R9), and their ability to stratify prognostic groups of patients (Figure R9).

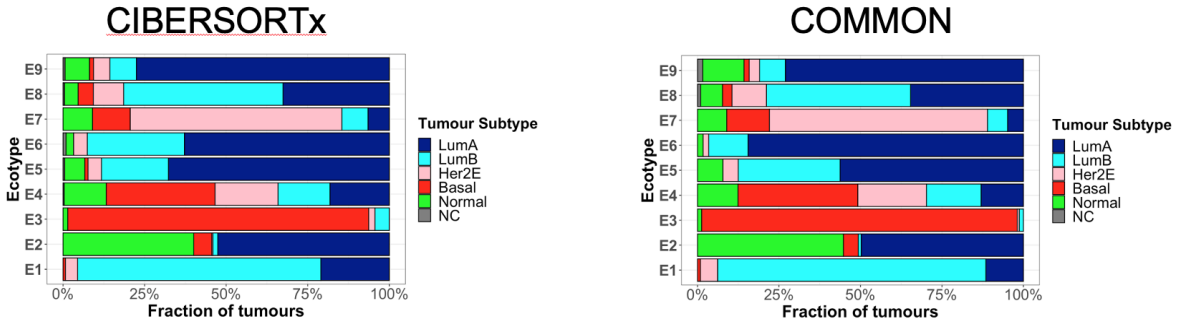

Figure R8: Relative proportion of the PAM50 molecular subtypes of the ecotypes generated from CIBERSORTx (LHS) and the common tumors in each ecotype, when combining CIBERSORTx and DWLS results (RHS).

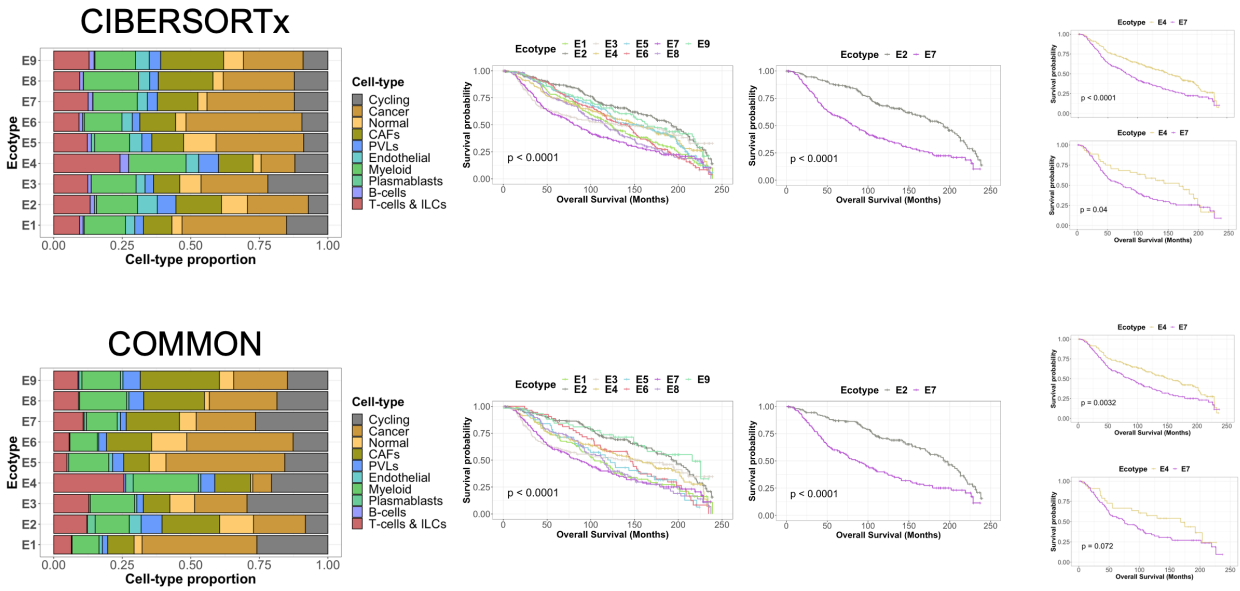

Figure R9: Comparison between ecotypes generated from CIBERSORTx (top) and the common tumors in each ecotype, when combining CIBERSORTx and DWLS results (bottom). Relative average proportion of the major cell-types enriched in the ecotypes (LHS). KM plots showing overall survival analysis of the ecotypes (RHS).

This comparison shows that the common ecotypes look very similar to the ecotypes defined from using CIBERSORTx alone. For example: ecotype E3 predominantly consists of basal-like tumours, the largest proportion of cycling cells, and relates to poor prognosis. E7 predominantly consists of Her2E tumours and relates to poor prognosis; E4 consists of a mixture of all subtypes and is enriched for immune cells; and E2 shows good prognosis, predominantly luminal A and normal-like subtypes, and lowest cycling cell content. Additionally, stratification of Her2E patients in the immune high (E4) and low (E7) ecotypes

shows similar trends (although no longer significant, p=0.07) towards improved survival for those in E4.

In summary, we have shown that the concordance when comparing ecotypes derived from the independent methods is high and maintains clinical and cellular composition features. Thus indicating that the ecotypes are not a unique feature of CIBERSORTx.

Revisions:

- We have added panel S11B to show a comparison between CIBERSORTx and DWLS when deconvoluting the pseudo-bulk datasets
- We have added panel S11J to show the consistency of survival results when using ecotypes formed from the overlap of CIBERSORTx and DWLS.
- We have added Table S11 showing the tumour overlap between the ecotypes when using CIBERSORTx and DWLS.
- The manuscript text has been updated (page 22, lines 3-4 and page 23, lines 10-14) to reflect these figure and table additions.
- The methods have been updated to describe the DWLS approach (page 40, line 20; page 41, lines 10-16; page 41, lines 22-26; page 42, lines 1-7).
- Figure S11 legend text has been updated accordingly.

Reviewer #2:

2.1 A major contribution in Wu et al is development of a molecular signature to classify single cell transcriptional profiles of breast cancer into different molecular subtypes, as well as to evaluate infiltrating immune and parenchymal cell clusters, and their associated correlates. For subtyping, the initial strategy employed was to assume PAM50 is the gold standard and then devise scSubtype to maximize similarity to PAM50 subtype assignments. This is a reasonable approach. A summary of the overlap between genes in scSubtype and PAM50 would aid in interpretation of scSubtype.

Response: We agree with the reviewer that the overlap of scSubtype and PAM50 gene lists would aid the interpretation of the methods. As seen in below (Figure R10), this revealed only 4 genes in common between the two gene lists. This included two genes in common with Basal_SC (*ACTR3B* and *KRT14*) and two genes in common with the Her2E_SC (*ERBB2* and *GRB7*) signatures (Table S4).

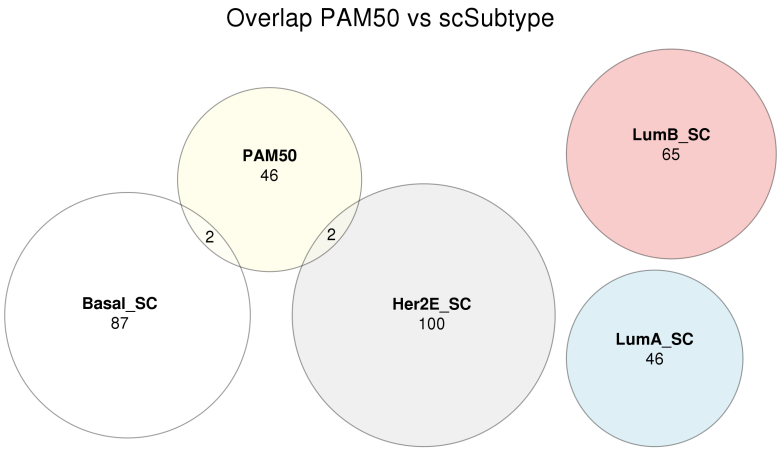

Figure R10. Overlap of scSubtype and PAM50 gene lists.

Revision: We have incorporated these overlaps into the results section of the text on page 8 lines 14-16.

2.2 scSubtype is applied to individual cells by computing the average gene expression of each subtype signature and assigning subtype to the highest score. There is no information provided about the robustness of the assignments for an individual cell. The authors should comment on this, particularly for samples classified as 'discordant' with bulk PAM50 scores.

Response: It is technically challenging to calculate the robustness of each cell without a gold standard external guide of some type, which does not exist for a single-cell intrinsic subtype classifier. It is worth noting the well-accepted methods such as Pam50 (Prosigna), OncotypeDX or Mammaprint also do not implement a robustness measure or any threshold when assigning subtype to a sample, yet are is widely accepted as a clinical tests. We believe that the robustness of scSubtype will only improve upon greater sample numbers applied in future scRNA-Seq studies.

Revisions: To ensure this is clearly communicated, we have modified or removed parts of our text in the revised manuscript describing the "accuracy", "discordance", and "classifier" nature of our scSubtype results, as seen in the abstract (page 4) and the text (page 8, line 21; page 24, line 20; page 25, lines 8-9). We have also incorporated a sentence into the discussion page 25 lines 14-16 acknowledging that the performance of scSubtype will improve upon greater sample numbers in future scRNA-Seq studies.

2.3 scSubtype signatures are "validated" by calculating the degree of epithelial cell differentiation and proliferation, both of which are associated with the molecular intrinsic subtypes. This seems a bit circular because the signatures were optimized against the PAM50 intrinsic subtypes. At a minimum, there should be an assessment of the overlap of genes across the various signatures.

Response: The reviewer raises an important point regarding the validation of intrinsic breast cancer subtypes with the differentiation scores (Dscore) and proliferation scores. However, it is important to consider that PAM50, Dscore and proliferation scores were independently developed in 3 different publications, and using distinct methods.

Dscore was derived from models of normal breast epithelial cell differentiation states rather than from breast cancers[4]. Dscore reflects the differentiation axis from mammary stem cells to luminal progenitors and mature luminal cells, which were all established using gene expression data from each of these FACS purified subpopulations[10,11]. The application of Dscore to breast cancers subsequently revealed important correlations with clinical subtypes, for instance, higher Dscores were found in luminal cancers, whereas lower Dscores are found in basal-like breast cancers[4]. Therefore our validation approach is not circular in nature. We observed consistent Dscore trends with our scSubtype classifier, which revealed high Dscores in LumA_SC and LumB_SC subtypes compared to Basal_SC and HER2_SC. This provides important correlations with established breast cancer biology. While we agree that assessing the overlap would be informative, this is not possible as the Dscore is computed using a centroid based predictor with information from ~20 thousand genes.

For proliferation scores[12], these were 11 genes related to the cell cycle that were initially used to compare against clinical assays, such as the Ki67 IHC proliferation index, for their prognostic abilities to predict patient outcome. While it is true that there is a strong overlap with PAM50 (9 of 11 genes shared), none of these 11 genes overlap with our scSubtype signatures (Figure R11). As we observed higher proliferation scores in cells classified as Basal_SC and Her2_SC compared to LumA_SC and LumB_SC, this independent correlation fits with the higher proliferative nature of Basal and Her2 breast cancers[12].

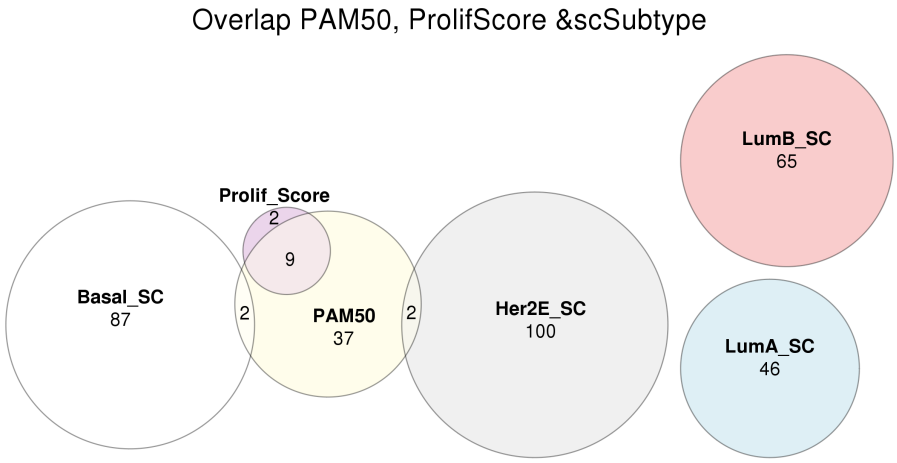

Figure R11. Overlap of scSubtype, PAM50 and proliferation score gene lists.

Taken overall, scSubtype provides novel signatures for classifying intrinsic molecular subtypes on the single cell level, which importantly fits with known breast cancer biology. We agree that reporting this overlap between scSubtype signatures and Dscores/Proliferation scores in our manuscript would help clarify this point to the reader.

Revisions: We have included these overlaps into the respective sections of the text page 8, lines 14-16 and page 33, lines 22-25 of the revised manuscript. We also acknowledge that the use of the term “validate” may not be appropriate here, which we have clarified in the text on page 10 line 1 of the revised manuscript.

2.4 True validation of scSubtype signatures in an external single cell dataset would strengthen the claims about its robustness.

Response: We agree with the reviewer that validation of scSubtype signatures using an external dataset would strengthen its claims. In our study, we tried to best control for this by splitting our scRNA-Seq tumor cohort into training and test sets in our scSubtype methodology. As for most classifiers, the robustness of scSubtype signatures will likely improve when using a larger number of samples across training and test sets in future scRNA-Seq studies. We do acknowledge that there are externally published scRNA-Seq datasets of primary human breast cancers[10,13] however, we do not believe they are suitable for this purpose. Most of these studies include datasets that are experimentally enriched for immune cells rather than tumor cells [14,15]. Most importantly, these datasets do not have matching bulk RNA-Seq and single cell RNAseq on the same tumors, like we have for our study, which is importantly required for a true validation against the gold standard clinical assay (which is based upon bulk mRNA expression).

Revision: We have included some discussion on these aspects in the text of the revised manuscript on page 25 lines 14-16.

2.5 With that immune cells occupy (at most) between 10-20% of the cellularity in a breast tumor (based on FACs studies), robustness of the immune cell data is questionable given that scRNA Seq was not performed following any sort of enrichment for CD45 cells by either positive or negative selection. Data should be provided on proportions of cellular compartments, and significance of read depths for lower percentage populations so that interpretation can be appropriately gauged.

Response: While it is true that we did not perform any enrichment or selection for immune cells prior to scRNA-Seq, it is widely accepted that droplet-based single-cell methods, including the 10X Chromium platform used in our study, are biased due to cell size, flow rates and droplet sizes. This has been explored in a large benchmarking study by Slyper et al. [16] which showed an enrichment of leukocytes in single-cell analysis of cancer tissues compared to matching single nuclei analysis (Slyper et al., 2020 Fig. 6). The natural selection of immune cells using this methodology can also be seen as beneficial for studying the immune milieu of cancers and tissue types with lower immune infiltration, as no extensive experimental enrichments are required.

In response to the requested data for the proportions of each cellular compartment, this is shown for each of the major cell types and states in Figure 1C of the manuscript. On average, 50.3% of cells from each tumor were immune cells (B-cells, Plasmablasts, T-cells and Myeloid cells), which ranged between 39.8%, 52.5% and 69.2% for ER+, TNBC and HER2+ tumors, respectively. If we have misinterpreted the reviewer request here, we are happy to provide other data upon clarification.

The read depths in each cellular compartment can be observed in Figure S2 C-D of the original manuscript, which shows the number of unique molecular identifiers (UMIs) and genes detected across each of the major cell types and states. In general, lymphocytes such as T- and B-cells show a lower average number of UMIs and genes per cell compared to more transcriptionally active cells such as cancer cells and proliferating 'Cycling' cells. These trends across cell types metrics are consistent with those previously reported in other cancer types (Lambrechts et al. 2018, Fig.1D[17] and Slyper et al., 2020 Extended Data Fig. 2 [16]).

Despite this, we sequenced our single cell libraries extensively and our T- and B-cells clusters have read depths that are slightly higher than those reported by previous studies [16] with an average of ~2,000 and ~1,000 UMIs and genes per cell (Fig. S2C-D). In addition, we have applied sophisticated and stringent filtering methods to ensure the highest possible quality of our final dataset, including the Emptydrops method, which distinguishes real cell droplets from empty droplets that hold similar expression profiles to the ambient 'noise' RNA profile. We hope these clarifications of our methodology, metrics and data processing steps satisfy the reviewers concerns about the quality and robustness of our immune cells.

2.6 CIBERSORT deconvolution: Fig S11 shows correlations between deconvoluted pseudobulk and actual proportions that from 0-1. However, the main text claims that there is good correlation between these measures. Some additional analyses to support the claim is needed, or else the authors need to roll back their interpretation of the findings.

Response: We thank the reviewer for the opportunity to clarify this point. Our reference to a "good" correlation was referring to the average of correlation coefficients (0.64) for all 46 cell-types (S11A, inset). However, we acknowledge that deconvolution methods have limitations and that not all cell-types perform equally, as shown by the individual results when deconvoluting the pseudo-bulk samples. To investigate this further we calculated the statistical significance of these correlations for each cell-type and summarised according to their major cell-type (Figure R12).

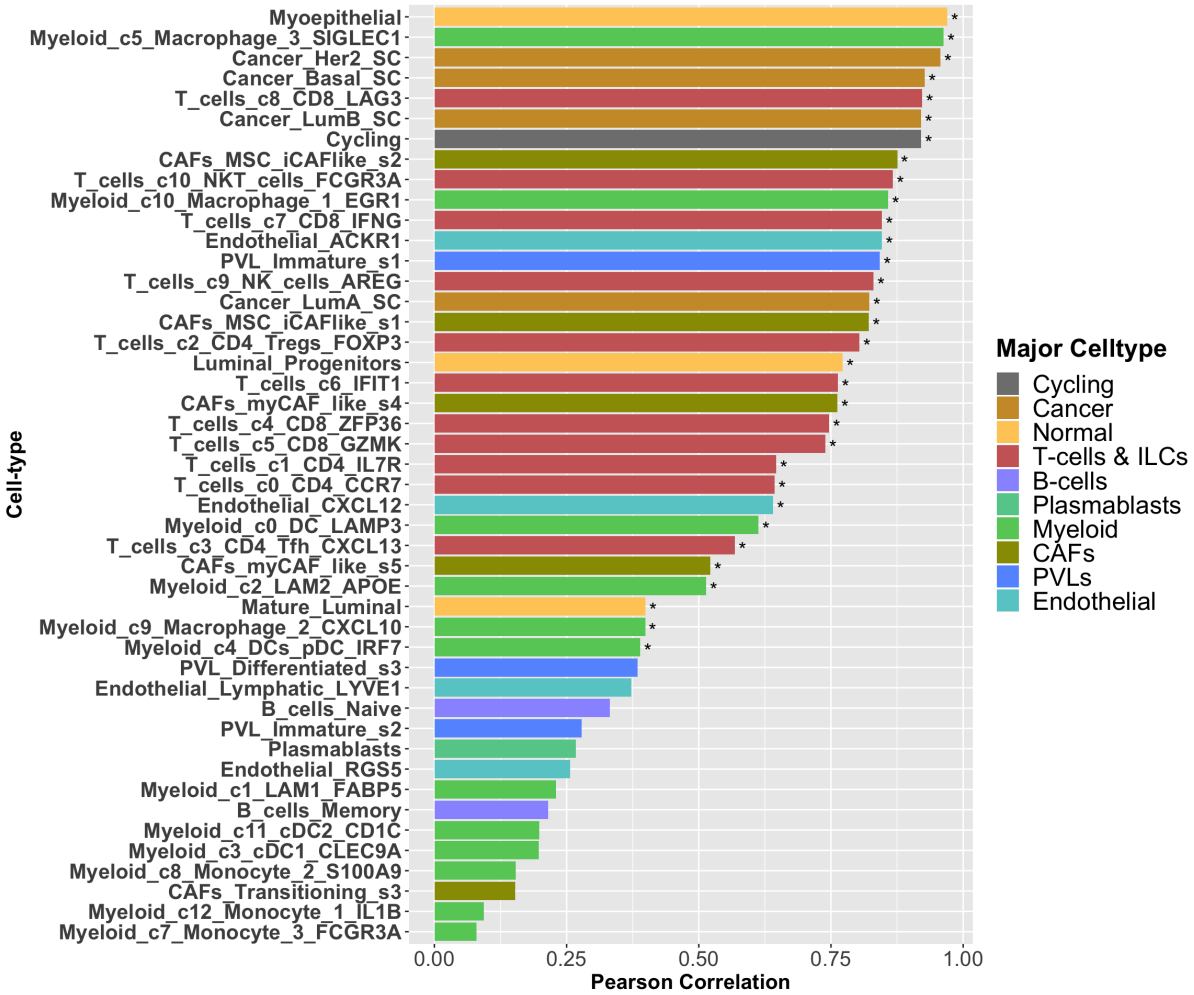

Figure R12: Barplot of the Pearson correlation between the actual cell-fractions captured by scRNA-Seq and the CIBERSORTx predicted fractions from pseudo-bulk expression profiles. * denotes a significant correlation $p<0.05$.

Clearly, a majority of cell-types consistently perform well (32/46 have p-value<0.05). These include: all epithelial cells (cancer and normal); all T-cell &

ILC subsets; cycling; and most of the CAF cell-types. In contrast, B-cells & plasmablasts, as well as a number of myeloid (*c1_LAM1_FABP5, c11_cDC2_CD1C, c3_cDC1_CLEC9A, c8_Monocyte_2_S100A9, c12_Monocyte_1_IL1B, c7_Monocyte_3_FCGR3A*), Endothelial (*Endothelial_Lymphatic_LYVE1*, *Endothelial_RGS5*), PVL (*Differentiated_s3*, *Immature_s2*), and one CAF (*CAFs_Transitioning_s3*) cell-type, didn't reach significance in correlation.

One explanation why some cell-types perform less well could be that they are rarer cell-types with low abundance. Indeed, on average the cell-types with significant correlation between actual and predicted cell-type abundances have significantly higher actual abundance in the pseudo-bulk samples ($p<0.001$, Wilcox-test) (Figure R13). This indicates that the rarity of a cell-type may be having some impact on the deconvolution performance.

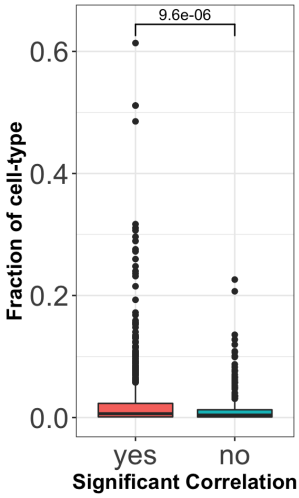

Figure R13: Boxplot comparing the actual cell-abundances, captured via scRNA-Seq, of those cell types with and without a significant correlation with predicted cell-type abundance using CIBERSORTx. Wilcox test p-value.

Although rare cell-types can be a challenge for deconvolution methods, our analysis in response to REF 1.7 shows that our ecotypes are robust across two deconvolution methods. Importantly, the alternative deconvolution approach that we used (DWLS) was specifically developed to help improve the identification of rarer cell-types, indicating that a combined approach may be beneficial. However, even when comparing the performance of both methods we can see that some cell-types remain challenging to predict accurately by either method, as seen by the 8 cell-types in the "significant_in_none" grouping (Figure R14).

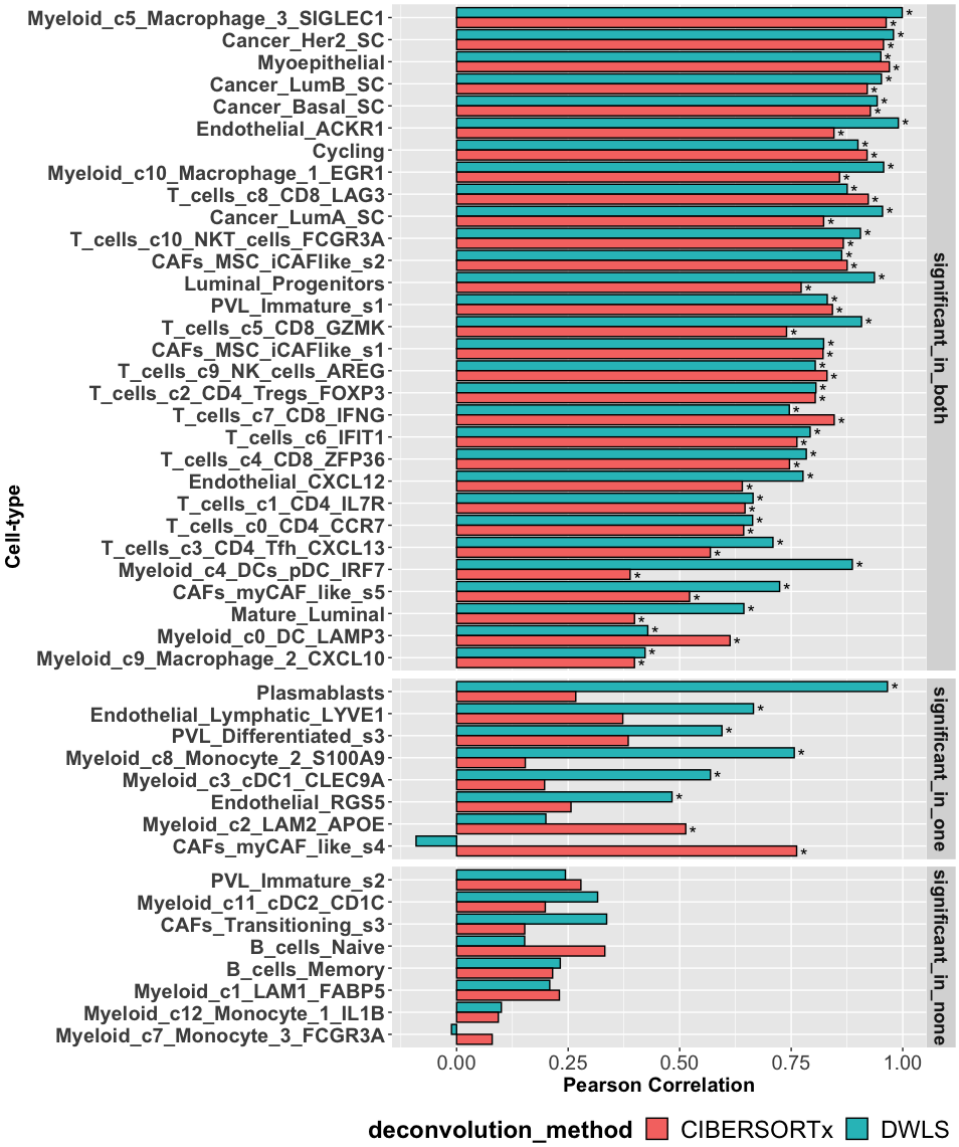

Figure R14: Barplot comparing the Pearson correlation, for each of the cell-types in the subset level of the BrCa cell taxonomy, between the actual cell-fractions captured by scRNA-Seq and the CIBERSORTx and DWLS predicted fractions from pseudo-bulk expression profiles. * denotes a significant correlation p<0.05 between actual and predicted cell-type abundance.

Ultimately, it's important to estimate the impact that these poorer performing cell-types might have on the tumour ecotypes identified from METABRIC. Because deconvolution methods perform better when the most complete set of representative cell-types are used, we were hesitant to repeat the cell-type deconvolution analysis with these poorly performing cell-types removed, as it may have an unintended impact on the performance of the other cell-types. Instead, we repeated the consensus clustering to re-generate the tumour ecotypes using only the 32 cell-types that were significantly correlated (p<0.05). The hypothesis being that if the ecotypes were similar, after removal of the poorer performing cell-types, we could increase our confidence in the robustness of the ecotypes and that they are not unduly dependent on or influenced by the poorer performing cell-types. This is broadly what we observed. When using only the 32 significantly correlated cell-types we identified similar numbers of tumour ecotypes. To allow a direct comparison between the

ecotypes generated with all cell-types we evaluated the concordance of the nine ecotypes by calculating the tumour overlaps between those generated using all cell-types and only the significant ones. Our hypothesis being that if the cell-types with lower correlation scores weren't substantially affecting the ecotypes they should show pairs of ecotypes with substantial tumour overlaps. The table below (Table R2), listing the nine ecotype pairs with the highest average overlap, shows that this was the case.

ecotype (all cell-types)	ecotype (significant cell-types)	sample overlap	sample overlap (%)
E1	E1	234	81
E2	E3	237	86
E3	E2	182	90
E4	E4	216	78
E5	E7	148	69
E6	E5	195	86
E7	E6	154	81
E8	E8	133	59
E9	E8	71	37

Table R2: Tumour ecotype overlaps between using CIBERSORTx on all or significantly correlated cell-types.

These overlaps show that most of the "all cell-types" tumour ecotypes have a greater than 75% overlap of samples with the "significant cell-types" tumour ecotypes, showing their robustness. We then used these overlapping tumours to define nine "common" ecotypes consisting of 1570 tumours whose ecotype calls were concordant between the two methods, a clustered heatmap of the cell abundances for these ecotypes is shown below (Figure R15).

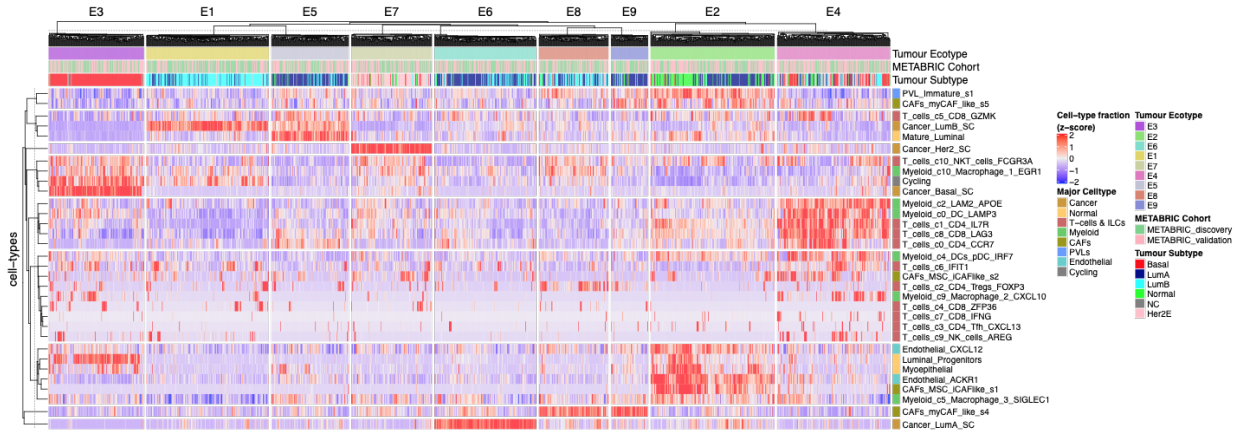

Figure R15: Heatmap of ecotypes formed from the common METABRIC tumors (columns) identified from combining ecotypes generated using CIBERSORTx with all, or the 32 significantly correlated cell-types (rows), when using CIBERSORTx on pseudo-bulk samples.

To test if these common ecotypes showed similar properties to the ecotypes using all cell-types we compared their molecular subtype (Figure R16) and cell-type compositions (Figure R17).

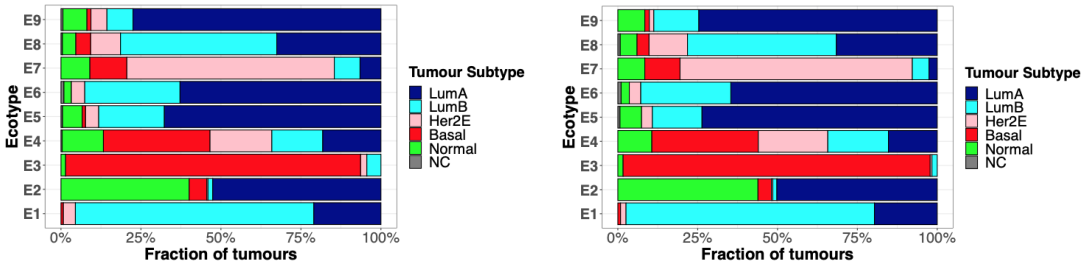

Figure R16: Relative proportion of the PAM50 molecular subtypes of the ecotypes generated from all cell-types (LHS) and the common tumors in each ecotype, when combining CIBERSORTx consensus clustering results from using all or the 32 significant cell-types (RHS).

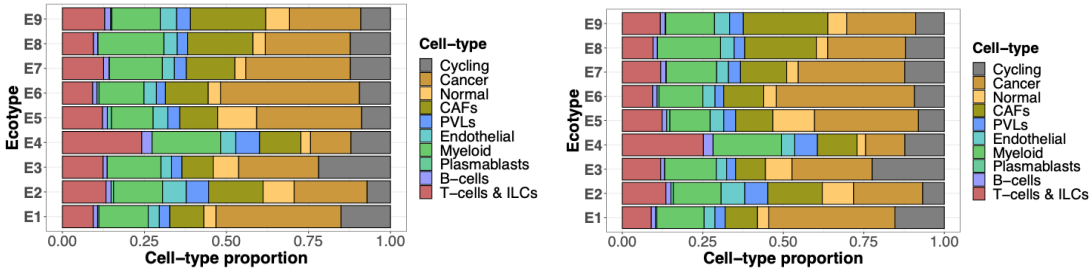

Figure R17: Relative average proportion of the major cell-types enriched in the ecotypes from all cell-types (LHS) and the common tumors in each ecotype, when combining CIBERSORTx consensus clustering results from using all or the 32 significant cell-types (RHS).

Finally, we wanted to look at the effect on survival of removing poorly correlated cell-types. As expected from the substantial sample overlaps between ecotypes derived from all cell-types and just the significantly correlated ones, similar overall survival was seen for the “significant cell-types” ecotypes. In particular, patients in E3 (mostly basal-like tumours) and E7 (mostly Her2E tumours) again had the worst prognosis (see figure 6F for result using all cell-types) and Figure R18. When stratifying the Her2E patients in the immune depleted (E7) and immune rich (E4) ecotypes we again see increased overall survival in the patients with significantly increased immune cell abundance (log-rank p-value=0.034) (Figure R18).

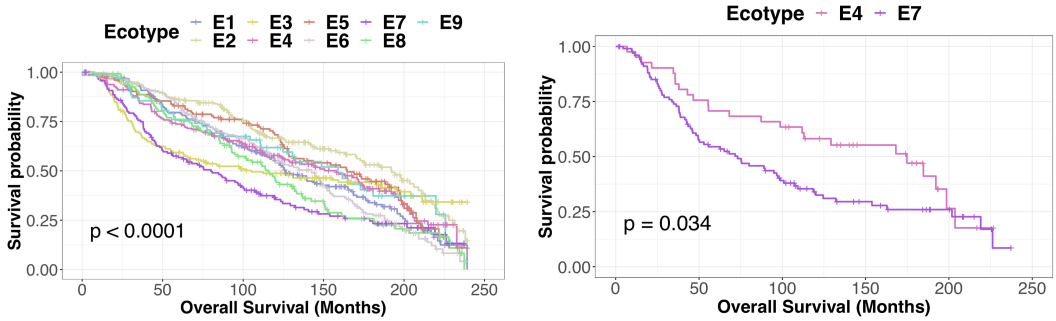

Figure R18: Kaplan-Meier (KM) plot of all patients with common tumors in each of the ecotypes (LHS) and Her2E patients with tumors in ecotypes E4 and E7 (RHS).

Taken together these new analyses indicate that the ecotyping approach is robust and not unduly affected by the presence of poorer performing cell-types. However, we acknowledge that care should be taken when interpreting results specifically related to these cell-types. In particular, a more conservative approach, with those cell-types removed, should also be used for corroboration.

Revisions:

- We have updated figure S11A to indicate the cell-types that have significant correlations between actual and CIBERSORTx predicted cell abundances.
- We have removed S11B, related to per-sample correlations, as we focus specifically on the accuracy of the cell-type correlations which we think are more informative for interpreting our findings.
- We have added panels S11C-H to reflect the ecotyping results when using only the significantly correlated cell-types.
- We have added Table S10 showing the tumour overlap between the ecotypes when using all cell-types or only the significantly correlated cell-types.
- The manuscript text has been updated (page 21, line 26; page 22, lines 1-4; page 23, lines 10-14) to reflect these figures and table additions.
- To be conservative in our interpretation of the cellular composition of the ecotypes, we have removed specific reference in the text to any cell-types that have insignificant deconvolution performance (i.e., Memory B-cells (page 22, line 21) and S1009 monocytes (page 23, line 5).

2.7 Application of scSubtype to deconvolute METABRIC is an interesting use case for the signatures and provides some additional insights into classification of breast tumors. Curtis et al (PMID: 22522925) describe 10 molecular subtypes of breast cancer. The authors should comment on the relationship of their ecosystem signatures to the previously published subtypes.

Response: To assess the relationship between our tumour ecotypes and the int-clusters a similar approach to assessing the PAM50 subtype relationship, shown in figure 6D, was used. There is not a unique association between ecotypes and the int-clusters, however some interesting relationships can be seen (Figure R19).

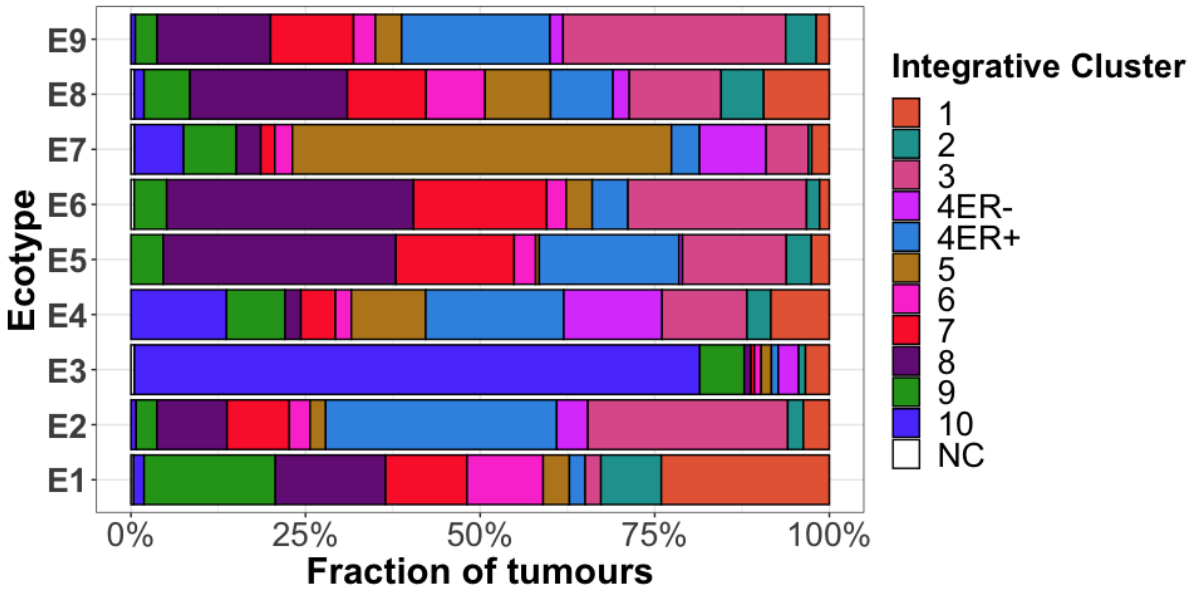

Figure R19: Relative proportion of the METABRIC integrative cluster annotations of the tumors in each ecotype (ecotypes generated using CIBERSORTx across all cell-types).

E3, which is enriched for basal-like tumours and cycling cells, has a high proportion of cancers from **int-cluster 10**, also associated with basal-like tumours and similarly poor 5-year survival; **E7** has a high proportion of cancers from **int-cluster 5**, which is defined by ERBB2 amplification and enrichment of the HER2E tumours. This is the worst prognosis group in both the METABRIC and our ecotype analysis; **E2** is the best prognosis ecotype, consisting of mostly normal-like and LumA tumours. This associates with both intCluster 3 (mostly LumA) and the ER+ patients in int-cluster 4 (mostly normal-like and LumA). This has some similarity to the METABRIC paper, where the best prognosis is for int-cluster 3 patients, however, our approach was able to additionally group these two good prognosis genomic clusters together.

Finally, ecotype **E4** doesn't have a clear association with either a specific int-cluster or PAM50 subtype, suggesting that our cell deconvolution based approach has identified a mixed subtype tumour group that is not easily resolved by bulk genomic studies, reflected by the role of the stromal and immune cells in defining this ecotype. In conclusion, these results suggest that ecotypes are not a simple surrogate for molecular subtypes and genomic subtypes.

Revisions: We agree that this is an important point and will be of wider interest to the community. We have added a new figure panel to show the association between ecotypes and METABRIC int-clusters (figure S11K) and supporting text to the manuscript (page 23, lines 16-26).

Missing details and/or stats in several places, for example:
2.8 • Fig 2D (lines 197-199): No indication of whether the differences across subtypes are significant

Response: We thank the reviewer for noting this lack of statistics. To resolve this we have used the Wilcox test to estimate the significance of the differences

between the proliferation and differentiation scores, shown in figure 2D, for each of the 4 scSubtypes. This showed that all pairwise comparisons were significant (<0.001).
Revisions: We have clarified this in the legend of figure 2.

2.9 • Supp Fig 5: No stats to support the comparisons of GM expression across scSubtypes

Response: We thank the reviewer for noting this lack of supporting statistics. To resolve this we have estimated the significance of the differences between the scSubtype scores of cells assigned to each of the gene-modules (fig S5F). We have approached this in 2 ways: (i) A Kruskal-Wallis test was performed to calculate the significance between any of the four scSubtype score groups in each of the gene-module groups; (ii) Wilcox tests were used to identify which scSubtype had significantly increased scSubtype scores in the cells assigned to each gene-module. Four (one-sided) Wilcox tests were performed within each gene-module grouping, to compare the scSubtype scores of each scSubtype to the rest of the scSubtype scores. P-values were adjusted for multiple testing.

Revisions: We have updated figure S5F, and the associated figure legend, to reflect these statistical tests.

2.10 • The authors claim that tumors in E3 were associated with poor overall survival, consistent with known poor prognosis of Basal-like and highly proliferative tumors (line 515-517), however this does not appear to be supported by data.

Response: We thank the reviewer for the opportunity to clarify this point. Our intention is to show that ecotype E3, which predominantly consists of basal-like tumours and is enriched for Basal_SC and cycling cells, has a poor prognosis as expected from other studies. Although this is not the ecotype with the worst 20-year overall survival, these patients do have a particularly poor 5-year prognosis, similar to METABRIC int-cluster 10.

Revision: To clarify and support this we have added a specific reference to figure 6F (page 23, line 3) and updated the text (page 23, line 3) to clarify that this is specifically related to poor 5-year survival.

Other comments
2.11 Introduction mentions data indicating CD8 infiltrations correlate with outcome but fail to mention impact on myeloid subsets also in correlating with outcome, or the significant body of literature identifying myeloid subsets as significant regulators of T cell functionality
'parenchyma cells' in introduction should be defined for the reader. Based on text later in results, this is an odd definition as infiltrating immune cells would also be considered as "parenchymal". Terminology here is not standard and confusing. Do the uthors instead mean to refer to mesenchymal lineage cells?

Response: We acknowledge the reviewers comment on our citation of myeloid literature in our introduction. We agree that the role of myeloid cells in the clinical

outcome of breast cancer and regulation of anti-tumour immunity is established and should be included. We also agree with the reviewers comment that our use of 'parenchymal cells' may be confusing, and does not fit with standard terminology. We are indeed referring to cells of the mesenchymal lineage, including fibroblasts, perivascular cells and endothelial cells, which we agree would be more appropriate terminology.

Revisions: We have now incorporated several key myeloid related studies in our introduction of the text page 5 lines 21-23 in our revised manuscript. We have also replaced all instances of 'parenchymal' with 'mesenchymal' in the text and figures of our revised manuscript.

2.12 CD68 was used to identify a myeloid cluster in Fig 1. CD68 only identifies a subset of myeloid cells, predominantly monocytes and macrophages, not DCs, eosinophils, basophils, neutrophils or mast cells. This should be clarified for the reader. In Fig 3, immune cells are subsetted with deeper resolution, but again, myeloid subsetting is for monocytes, macrophages, and DCs only. Were eosinophils, basophils, mast cells not identified? Were they absent? Clarification on this topic is needed. Is their absence a result of bulk scRNAseq without any CD45 enrichment where read depth was limited by paucity of these subsets? Should be discussed. How are subsets identified concordant (or not) with published FACs studies of breast cancer?

Response: Regarding our myeloid cell classifications, the expression of *CD68* in Figure 1B is used for visualisation purposes only, and that a single marker is not used for the cell classification. We rather applied two levels of cell type classification using gene signatures and a cluster grouping approach to ensure that all cells were subsequently examined through reclustering analysis at higher resolutions. The first was applying the Garnett method[18] using published cell type signatures. Here, myeloid related signatures including gene sets for monocytes, macrophages and DCs were extracted from the XCell database[19] and included for Garnett classification. Secondly, we utilised Garnett's 'cluster extension' approach where all neighbouring cells that are clustered together are classified as the majority call within that cluster, including 'unknown' cells. This conservative approach ensures that every unclassified 'unknown' cell that shares similar transcriptional profiles are annotated and further examined by reclustering.

In response to the missing cell types mentioned by the referee (including eosinophils, basophils and mast cells), We acknowledge their presence in breast cancers, previously described by Ruffell *et al.* [20]. However, these cell types were not detected in our study. This is likely due to (i) relative rarity of these cells and (ii) loss due to tissue dissociation or capture bias. In a study by Azizi, Pe'er and colleagues (2018)[14], no eosinophils or basophils are described, while only very small numbers of mast cells are identified. These findings are consistent with previous IHC studies that report no tumor infiltrating eosinophils in breast cancers [21]. Furthermore, mast cells have been associated with low tumor grade and estrogen positivity[21]. Considering that less than half of our patient tumor cohort are estrogen positive and predominantly consists of high grade tumors, it

is not unexpected that mast cells were not detected without prior immune enrichment. Similarly, previous FACS studies suggest have shown that basophils make up less than 5% of CD45+ cells in treatment naive breast cancers[20], suggesting targeted enrichment strategies are required to study them by scRNA-Seq.

Revisions: We acknowledge these caveats of our immune dataset, which we have discussed in the text of the revised manuscript (page 13, line 25; page 14, line 1; page 24, lines 12-14).

2.13 The authors identified 6 macrophage clusters. Do any of these identify the macrophage-signature that predicts clinical outcome (in Metabric) reported by Pollard (PMID 30930117) via RNAseq from FACs sorted macrophages?

Response: We thank the reviewer for the suggested literature regarding tumor-associated macrophages (TAM). The TAM RNA signature reported by Pollard[22] correlates to some degree with all macrophages, with Macrophage clusters MAC : CXCL10, MAC : SIGLEC1 and MAAC : EGR1 associating most highly. The distinct high expression of CCL8 found within MAC : CXCL10 cluster strongly suggests the most likely candidate for the CCL8 TAM phenotype identified in Pollard[22].

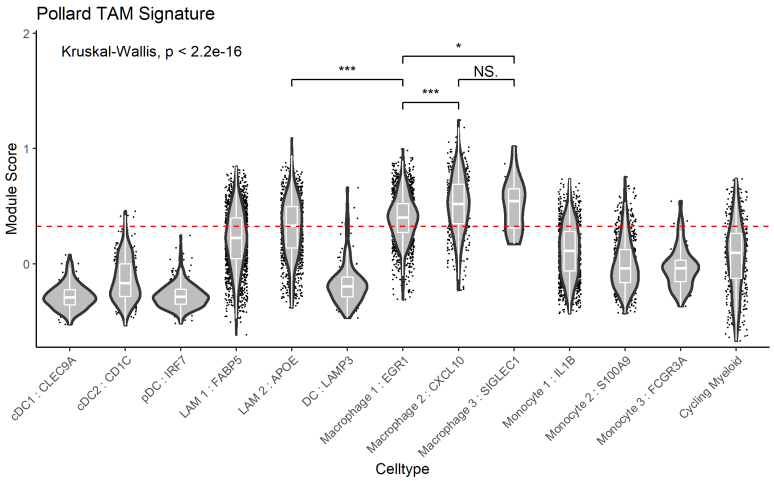

Figure R20. Enrichment of the TAM gene signature from Cassetta et al. (2019) across myeloid clusters, as estimated using ModuleScore function in seurat v3. A paired t-test was performed for cell types of interest. P-values denoted by asterisks: *p < 0.05, p < 0.01, *p < 0.001 and ****p < 0.0001. Dashed red line marks median expression value of signature score.

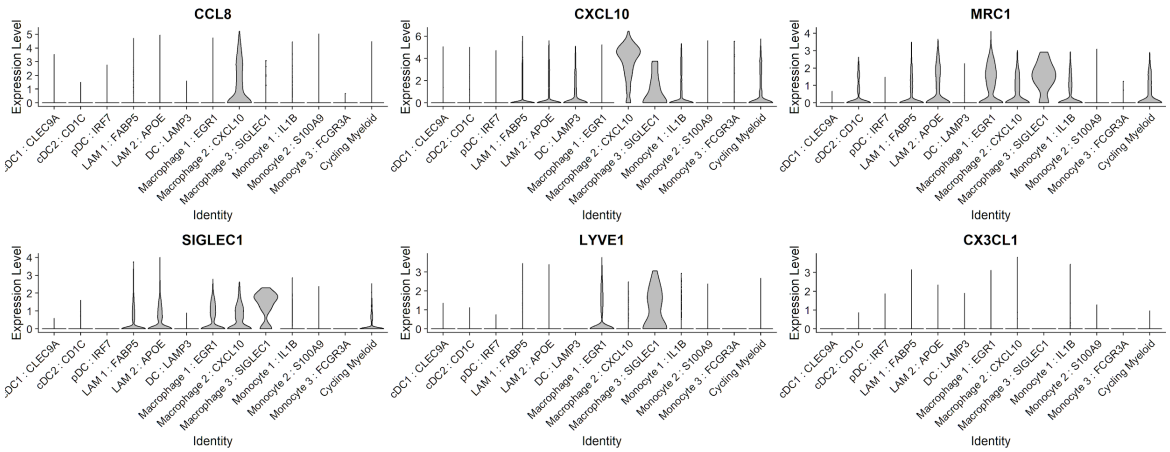

Figure R21. Expression of selected TAM markers from Cassetta *et al.* (2019) across myeloid clusters.

Revisions: We have included the Pollard Tumour-associated macrophage (TAM) signature[22] and scRNA log-normalised expression value of CCL8 found in our myeloid dataset as a supplementary figure (Fig S7I).

2.14 B cell cluster is identified in Fig 1, but not included in Fig 3. Why? What B cell state is represented in the cluster (Fig 1)? Any indication of plasma cells? B cell clusters are evaluated in Fig S6 but more details on why absent in Fig 3 are needed.

Response: Compared to the T/ILC and myeloid data, the B cell data did not cluster to high resolution. This is likely an artefact of variable Ig gene expression confounding clustering and differential gene expression analysis, as noted in the original submission. This is a common issue with analysis of B cells using this technology. Nonetheless,we identified 2 major sub-clusters of memory and naive B cells. In the interest of space, the B cell data is reported in supplementary figure S6.
Plasma-like cells were detected in multiple cases. As mature plasma cells are thought to primarily reside in the bone marrow, we have conservatively named all cells of this lineage 'plasmablasts' as seen in Fig 1 and Fig S6. Similar confounders, however at a greater intensity, afflict the clustering and DEG analysis of plasma cells due to their even greater proportion of immunoglobulin genes.

2.15 Spatial transcriptomic data in Fig 5 – CAF-T cell neighborhoods were identified. What about myeloid-T cell considering association of LAM2 and T cells in (Fig 3, S10). Are these suppressive myeloid subsets by gene expression, e.g., hi IL-10, ROS, iNOS)?

Response: To address whether myeloid subsets are suppressive, we first looked at the expression of the genes noted by the reviewer. The expression of IL-10 was mainly detected in the c10_Macrophage_1_EGR1 cluster, while both LAM clusters (c1_LAM1_FABP5 and c2_LAM2_APOE) showed very sparse expression. iNOS could not be detected by scRNA-Seq, while both markers were too sparsely detected across both single cell and spatial transcriptomics datasets, likely due to gene drop-out (data not shown).

To account for gene dropout, we instead scored signatures of immune suppression, separated by known ligand mediated versus secreted genes. We used IL10, TGFB1, IL4, IL13, IL37, IL33, IL35, IL12A, EBI3, IL27, IL23A for module scoring of secretable immune molecules and VSIR, CD276, CD274, PDCD1LG2, LGALS9, TNFRSF14, PVR, FASLG, TNFSF10, VTCN1, B7H4 along with ectoenzymes ENTPD1, NT5E, for surface bound molecule mediated suppressions. Signature score of inhibitory surface markers revealed a distinctly high expression of Macrophage 2 : CXCL10 cluster in comparison to all other macrophages. Expression of secreted immune suppressor genes was found to be not significant within the Macrophage clusters (Mac 1, Mac 2, Mac 3) however significant when compared across to LAM1/2 macrophages.

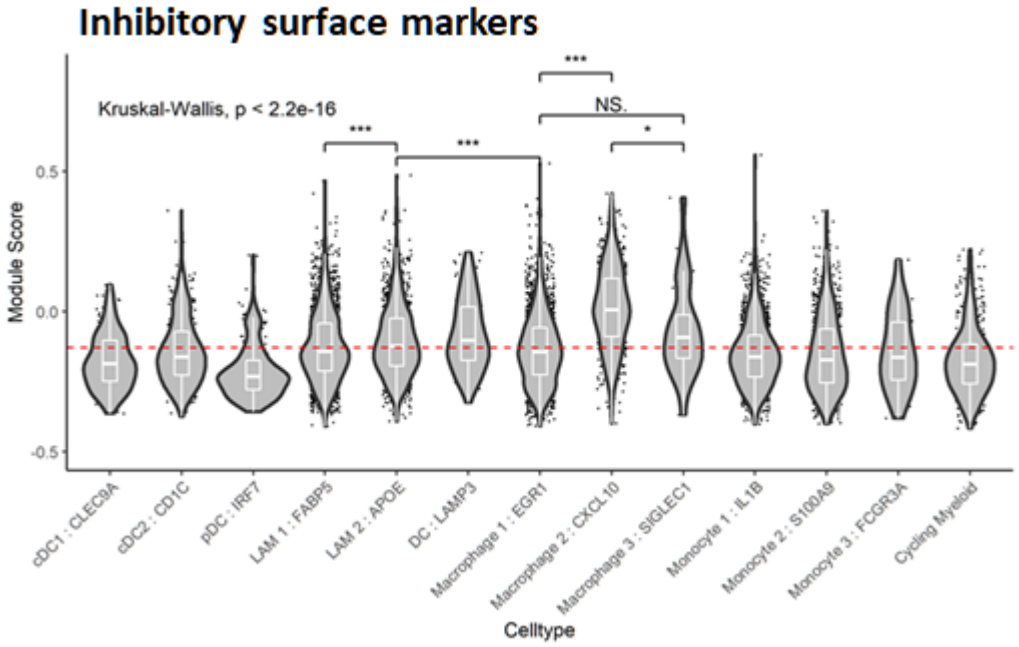

Figure R22. Gene signature score of surface markers previously associated with suppression across myeloid clusters, as estimated using ModuleScore function in seurat v3. A paired t-test was performed for cell types of interest. P-values denoted by asterisks: *p < 0.05, p < 0.01, *p < 0.001 and ****p < 0.0001. Dashed red line marks median expression value of signature score.

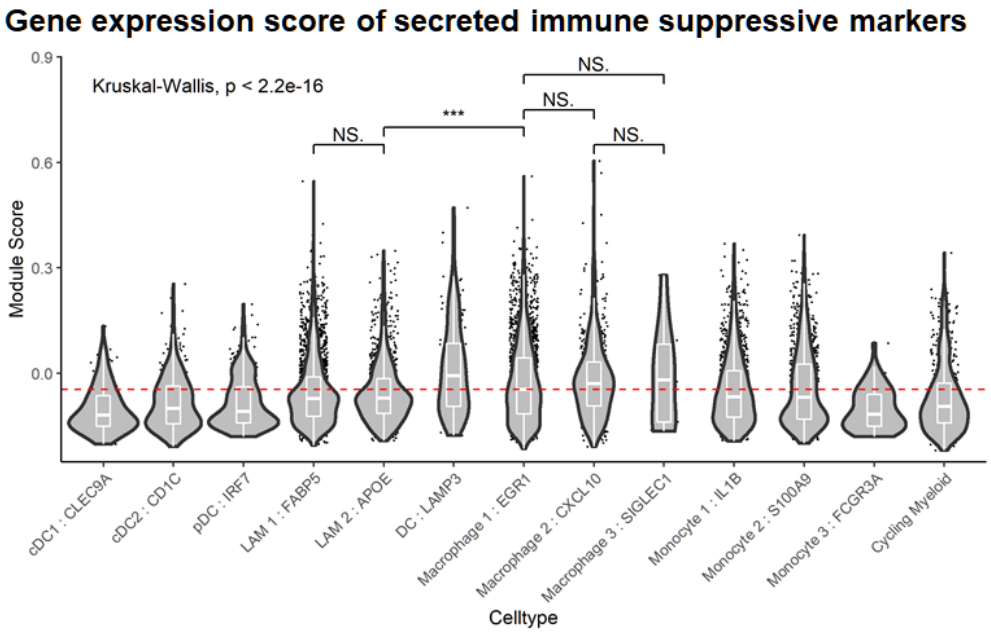

Figure R23. Gene signature score of secreted immune suppressive markers previously associated with suppression across myeloid clusters, as estimated using ModuleScore function in seurat v3. A paired t-test was performed for cell types of interest. P-values denoted by asterisks: *p < 0.05, p < 0.01, *p < 0.001 and ****p < 0.0001. Dashed red line marks median expression value of signature score.

In light of this result, we examined our spatial transcriptomics data to reveal associations between T-cells and the Mac:CXCL10/c9 cluster in addition to the reported interaction of LAM2 cells with T cells (Fig 6F). Mac:*CXCL10*/c9 cells were mostly enriched in spots annotated as 'invasive cancer + stroma + lymphocytes' (Fig S10E). As seen in the plots of cell abundance over tissue spots (Figure R24) and the revised heatmap from Figure 5H of Pearson correlations in the revisions below, the Mac:CXCL10/c9 cluster showed consistent positive correlations with CD4+ and CD8+ T-cells across a majority of breast tumors, suggesting they are spatially localised to similar tumor regions. Interestingly, Mac:CXCL10/c9 cells also express the highest levels of CCL8, corresponding to the repressive TAM phenotype reported by Pollard et al. discussed in rebuttal point 2.13. Considering the expression of PD-L1 and PD-L2 in Mac:CXCL10/c9 cells, as identified earlier by scRNA-Seq and CITE-Seq (Figure 3H), spatial proximity with T-cells suggests these immune niches have some functional relevance in regulating anti-tumor immunity.

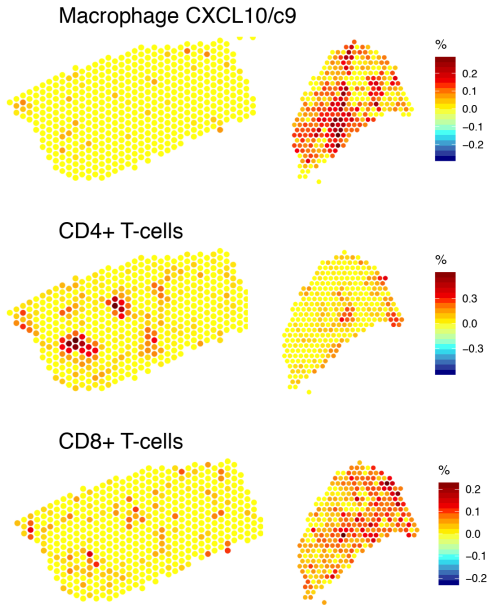

Figure R24. Scaled deconvolution values for Macrophage CXCL10/c9, CD4+ and /CD8+ T-cells overlaid onto tissue spots as in Figure 5A, as estimated using Stereoscope. Representative tissue mostly enriched for DCIS (left) and invasive cancer (right) from TNBC case CID44971 is shown.

Revisions: We have included these results related to the spatial associations between the macrophage CXCL10/c9 cluster and CD4/CD8 T-cells in Figure 5H (Figure R25). We have also placed a plot of the cell abundances (as above in Figure R24) over tissue spots in Figure 5K. These findings are now discussed in the text page 21 lines 9-15 in the revised manuscript.

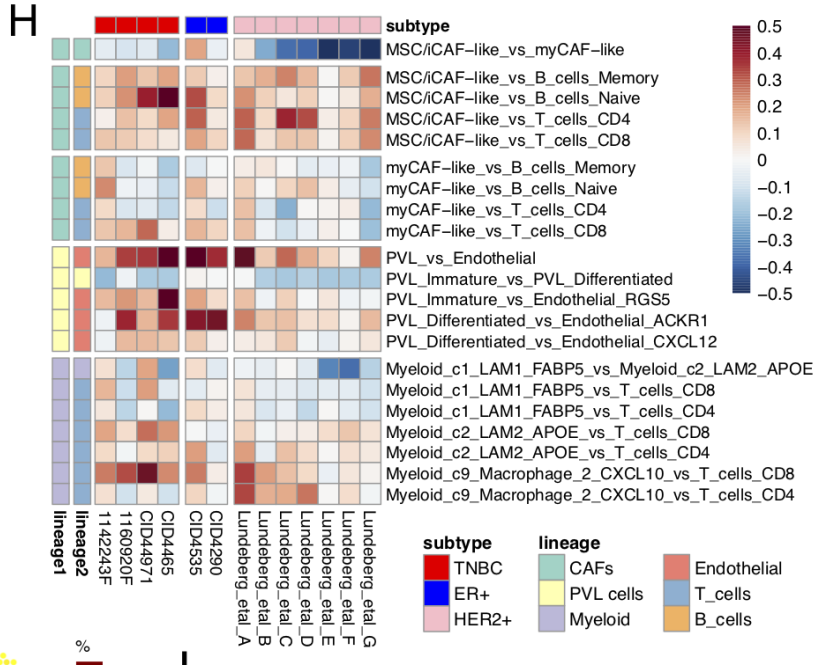

Figure R25. Extract from figure 5H of the revised manuscript below, where we have included the associations between Myeloid_c9_Macrophage_2_CXCL10 and CD4/CD8 T-cells across spatial datasets. Pearson correlations between deconvolution abundances, as estimated using stereoscope, as shown.

2.16 Methods for scSubtype are reasonable and leverage existing tools in Seurat package, however the description of how the methods were implemented is a bit confusing and should be written for readers not familiar with Seurat package functions.

> **Revisions:** We have updated this methods sections with greater detail explaining the seurat functions applied to scSubtype in the text (page 32, line 23-25; page 33, lines 1-5) of the revised manuscript.

2.17 In general, the figure legends are not very informative. Additional details about what is plotted would aid in interpretation.

> **Revision:** We have provided greater detail to all our figure legends of the revised manuscript.

2.18 Suggest moving Supp Fig 5B to main to provide overview summary of molecular features in each gene module.

> **Revisions**: We have moved S5B to figure 2E. Re-numbered and re-arranged other panels in figure 2 to fit and amended accordingly in the text and figure 2 and S5 legends.

Reviewer #3:

Remarks to the Author:

Wu et al. have conducted single cell profiling of 26 primary breast tumours using RNA sequencing and CITE-sequencing. They map major cell types and lineage relationships among cancer cells and the TME. They explore inter and intra-tumoural heterogeneity with respect to the intrinsic subtypes of breast cancer and formulate a single cell subtype classifier. Using these data they also deconvolve bulk RNAseq data and classify tumours into 'ecotypes' based on their cell composition. In addition, they use the Visium 10x platform to conduct analysis of RNA expression with spatial resolution in a small subset of tumours.

This is a well presented and timely report that does a reasonable job of viewing established breast cancer subtypes through a single cell lens. I judge the work to be of moderate novelty and impact; many of the observations are not as original as claimed.

Points to be addressed:

3.1 The novelty and significance of some of the findings are frequently over stated. There have been several single cell analyses of breast tumours. What amounts to 'comprehensive' is also highly subjective. The language should be moderated to acknowledge efforts that have gone before.

> **Response:** We acknowledge that our use of the term 'comprehensive' is subjective. It was used to convey the strengths of our study in providing an in depth and integrated analysis of the neoplastic, immune and stromal compartments of breast tumors, whereas previous studies have mostly focused on profiling these cell types in isolation. With regards to the previous breast cancer single-cell based publications, we have tried to include reference to studies that we are aware of in the introduction and at other relevant points.

However, if the reviewer has a specific study in mind that we may have missed we would welcome the opportunity to include this.

Revisions: We have moderated our language to avoid the use of 'comprehensive' in the text of the revised manuscript (page 4, line 4; page 6, line 19; page 7, line 3; page 24, line 4; page 24, line 26).

3.2 Established cell signatures were used to annotate cell clusters. A limitation of cell signatures is that they are typically prototypes and do not account for the variability in expression profiles that may arise in a dynamic environment like a tumour. How well did cell signatures match clusters? How did the authors deal with poorly concordant cluster signatures?

Response: The referee raises a valid concern regarding the use of established cell signatures to annotate our single-cell clusters, which raises an interesting point regarding the distinction between cell types and cell states. While we acknowledge that these prototypical signatures are derived from other tissues and technologies, we would like to reiterate that these were only applied to classify our major cell lineages. We used a carefully curated list of established signatures from databases and the primary literature to ensure that all cell types in our scRNA-Seq dataset were covered. We subsequently examined each major cell type in great detail to annotate our finer cell subsets using unbiased reclustering methods, independent analysis, and correlations with extensive literature and expertise in breast cancer, immune and stromal biology.

For the annotation of major cell lineages, we also applied a conservative approach to deal with cell types whose expression profiles may vary in the dynamic tumour microenvironment. All cells that first pass our quality control parameters and cell filtering methods (Emptydrops) are run through the cell annotation method Garnett, which classifies poorly concordant cells as 'unknown'. Here we used Garnett's cluster extension approach, which reclassifies these 'unknown' cells in each cluster to the majority cell type call within that cluster. For example, 'unknown' cells within a cluster that predominantly contains myeloid cells may have expression profiles that slightly deviate from prototypical signatures however, these unknown cells are reclassified as myeloid cells and are further examined in subsequent analyses. In this example, more detailed myeloid annotations are subsequently defined by examining the expression of established monocyte, macrophage and DC markers, cell surface molecules by CITE-Seq (Figure 3H) and enrichment of more refined signatures from the literature (Figure 3F). In another example of the mesenchymal cells, prototypical fibroblast signatures were first applied to classify cells with the major CAF lineage annotation. We then reanalysed CAFs using an independent pseudotime method to define all the dynamic CAF states present in breast tumors (Figure 4C), and further correlated them to established CAF signatures from the literature (Figure S9E-F).

As requested by the reviewer, we have also scored these prototypical signatures in our single cell dataset to show how well they matched clusters (figure below). This was performed using the AUCell method, which is independent of our cell classification method Garnett. Here we see a specific enrichment of each

signature in their respective major cell annotation groups, for example, Myeloid signatures enriched in the Myeloid cluster, suggesting that they are highly suitable for cell type classification in breast cancers. This highlights that these prototypical signatures broadly reflect major lineage cell types rather than their dynamic cell states in a tumour.

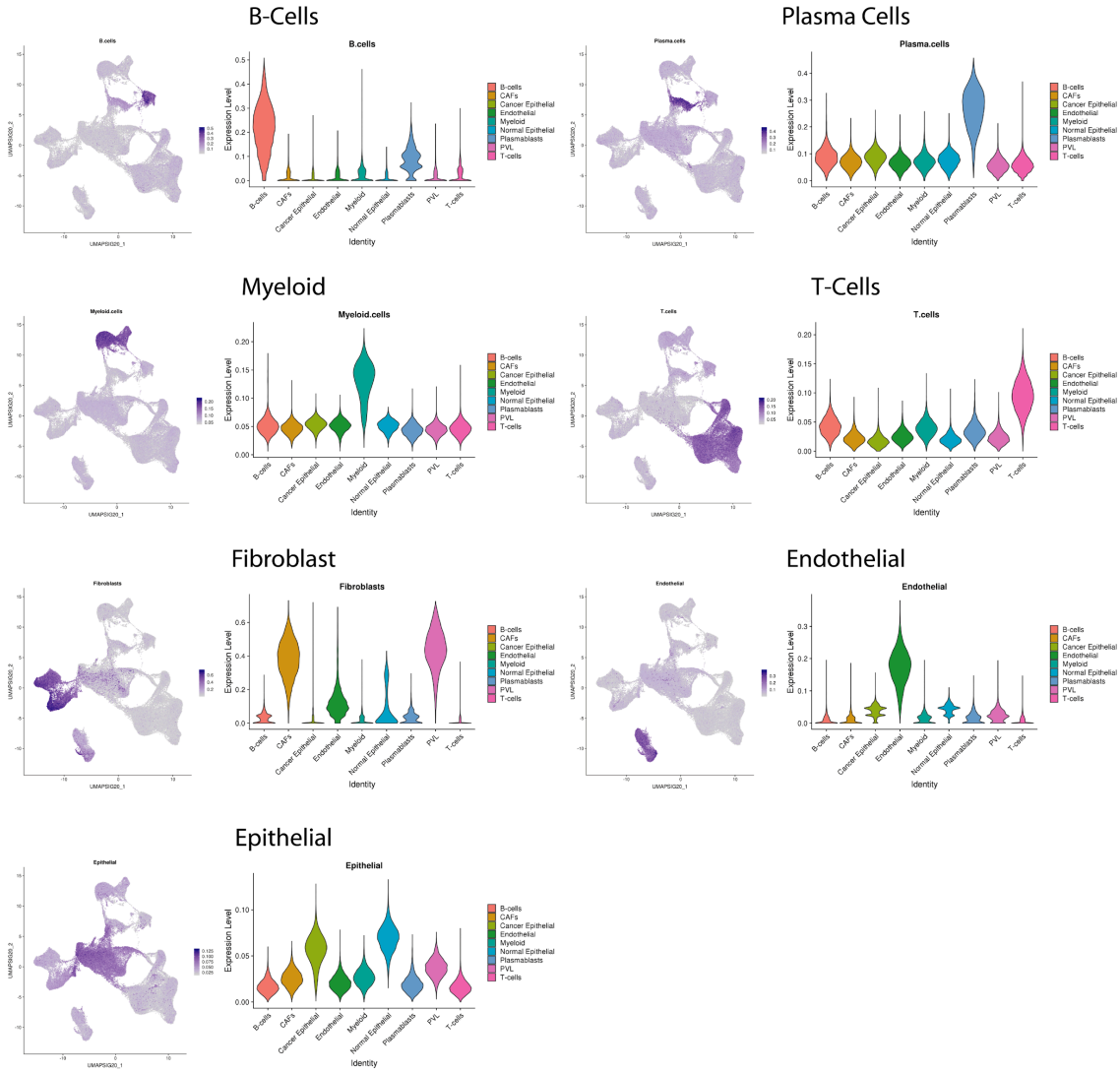

Figure R26. Signature scores for the gene lists used for the classification of major cell lineages using Garnett. Gene lists were extracted from the XCell database and the Lim *et al.* study. Scores were computed using AUCell, which is an independent method to Garnett. For each cell type, signature scores are shown as featureplots in UMAP space (left), as in Figure 1A, and as violin plots (right).

3.3 The single cell subtyping classifier is a nice addition. However, need it be so complex? After all the key criteria are whether a cell expresses hormone receptors, proliferation markers, HER2 and basal cytokeratins. I would like to see how a much simpler classifier performs in these data, for example one using as few as three genes (see Haibe-Kains et al JNCI 2012).

Response: We thank the reviewer for their comment regarding the complexity of our single cell subtyping method. While studies by Haibe-Kains et al. (2012) have suggested that a three gene model can be used for an adequate classification of the clinical subtypes of breast cancers in bulk analysis, these studies did not fully explore its clinical relevance. Further discussion on this specific matter can be found in a detailed rebuttal[23], which suggested that larger gene panels provide significantly greater prognostic information. For instance, larger gene panels such as that used in PAM50 can provide significantly greater prognostic information, predict response to chemotherapy, and provide a greater understanding of breast cancer diversity compared to the three gene model. The latter point is particularly important as our study aimed to understand the phenotypic diversity of breast cancers at single cell resolution. Furthermore, classification using fewer genes can not maintain the same concordance for subtype classification, particularly for Luminal B and Her2E tumors[23].

We would also like to reiterate that our method is not overly complex in nature. For cell classification, the average expression of the four subtypes signatures are first computed, then every neoplastic cell is classified by the highest of the four signature scores. This is simpler than some previous classifiers.

Lastly, due to the inherent sparsity of current scRNA-Seq technologies, models applying a small number of gene parameters to the single-cell space, such as the three gene model, may be technically challenging. When we assessed the PAM50 genes in single cell space, we observed that for each gene, an average of 80% of all the neoplastic cells had zero values (Table R3). One important example was that ~37.8% of all luminal cells, and ~68.7% of all neoplastic cells, had 0 reads for the estrogen receptor gene *ESR1*. The sparsity of PAM50 was our main rationale for the need to derive new signatures for intrinsic molecular subtyping that are compatible with the unique features of single-cell data.

Gene	Percentage of neoplastic cells with 0 counts (%)
SLC39A6	39.02
BAG1	43.82
MLPH	44.71
MYC	47.47
GPR160	60.97
ERBB2	61.53
CXXC5	62.99
FOXA1	63.16
BLVRA	64.45
ESR1	68.66
GRB7	68.73
MAPT	68.78
PHGDH	71.91
UBE2T	76.34
MDM2	79.16
SFRP1	81.47
ACTR3B	82.00
PTTG1	82.01
TYMS	83.00
TMEM45B	83.17
CENPF	83.29
NAT1	83.92
UBE2C	83.98
CCNB1	84.25

Gene	Percentage of neoplastic cells with 0 counts (%)
BIRC5	84.72
EGFR	85.17
BCL2	86.75
CDC20	87.42
CDH3	87.70
FGFR4	87.86
MKI67	88.37
PGR	88.96
RRM2	89.44
KRT17	89.94
KRT5	89.95
KRT14	91.08
MYBL2	91.50
ANLN	91.97
CDC6	93.20
MELK	93.27
MIA	93.95
KIF2C	94.32
CEP55	94.36
FOXC1	94.68
MMP11	95.84
CCNE1	96.10
EXO1	96.46

Table R3. Sparsity of PAM50 genes in scRNA-Seq data. For each PAM50 gene, the percentage of neoplastic cells with zero reads are shown.

3.4 Several tumours are shown to contain HER2 cells that do not belong to the HER2 subtype. One HER2 tumour (4066) is identified as having almost half of its cells as non-HER2. This could be of major clinical significance in the use of targeted therapy. Can these cells be identified in situ using IHC and/or FISH? If not, how should they be interpreted?

Response: The reviewer is correct in stating that “HER2 cells" could be found in several tumors that do not belong to the HER2 subtype (Fig 2B) however, we would like to clarify that our findings are referring to the presence of HER2-Enriched (HER2E) cells by intrinsic subtype (scSubtype) in clinically HER2-negative (cHER2-) tumors. These are two distinct classifications based on molecular intrinsic subtype (HER2E) or clinical pathology guided by Her2 IHC/amplification status (cHER2). The HER2E subtype is known to have significant clinical heterogeneity, and is composed of both cHER2+ or cHER2– tumors[24], and also those that are ER+ and ER–. The presence of HER2E cells in cHER2- tumors indicates that these cells have transcriptional similarities to the HER2-Enriched subtype tumors; this does not necessarily mean these cells are cHER2+ as we know that 20-30% of HER2E tumors are not cHER2+.

Similarly, cHER2+ tumours can contain cells with transcriptomes more closely related to the LumA, LumB or Basal intrinsic subtypes than to HER2E. The particular case that the reviewer has highlighted (CID4066) is a cHER2+/ER+ breast cancer (70% 3+ ER, 100% HER2-ISH AMPLIFIED 3+; Table S1). Our findings that this tumor contains a mixture of non HER2E cells from the LumA and LumB intrinsic subtypes (Fig 2B) highlights the power of single cell and scSubtype for identifying intra-tumor heterogeneity, and it is quite possible this tumor might showed a less than robust response to trastuzumab due to this mixed molecular nature. This intriguing finding cannot be corroborated here and will be the focus of future studies. We hope that our explanation of these distinct

HER2 definitions have clarified the reviewer's request for the use of IHC and/or FISH to identify the "non-HER2" cells in this cHER2+ tumor.

Revision: We have clarified the definition of molecular and clinical HER2 subtypes in the introduction of the text page 5, lines 10-11.

3.5 In this report, as in many others, HER2 positive breast cancer is treated as a single subtype. But this does not reflect the clinical reality. Hormone receptor expression stratifies HER2 positive disease into two groups that differ in every meaningful way - prognosis, treatment, patterns of relapse. It would be interesting to know how this distinction impacted the cellular ecosystem of HER2 positive tumours here.

Response: We thank the reviewer for raising these important points regarding the challenges in defining and treating HER2+ breast cancers. We acknowledge that these tumours can have significant heterogeneity in both ERBB2 and ER signalling. Due to low numbers of HER2+ tumours in this study (2 x HER2+/ER+, 3 x HER2+/ER-) we weren't confident in the power to robustly analyse specific cellular differences between these tumour types. We look forward to exploring this in detail as more cellular datasets of breast cancer are generated by the community.

However, we are able to address this question in our deconvolution analysis. Specifically, we extracted all 247 tumours classified as HER2+ from the METABRIC cohort and stratified them according to their ER status. This resulted in 139 HER2+/ER- and 108 HER2+/ER+ tumours. First we carried out unsupervised hierarchical clustering on the CIBERSORTx deconvolution results to identify patterns of difference in cell-type composition between the 2 (Figure R27). This clearly highlights the separation between luminal cell-types, and Her2_SC and Basal_SC cell-types, which are significantly enriched in the HER2+/ER+ and HER2+/ER- tumours, respectively. In addition, a number of T-cells & ILC cell-types are significantly enriched in the HER2+/ER- tumours (CD4+ IL7R, NKT, CD8+ LAG3, CD4 Tfh).

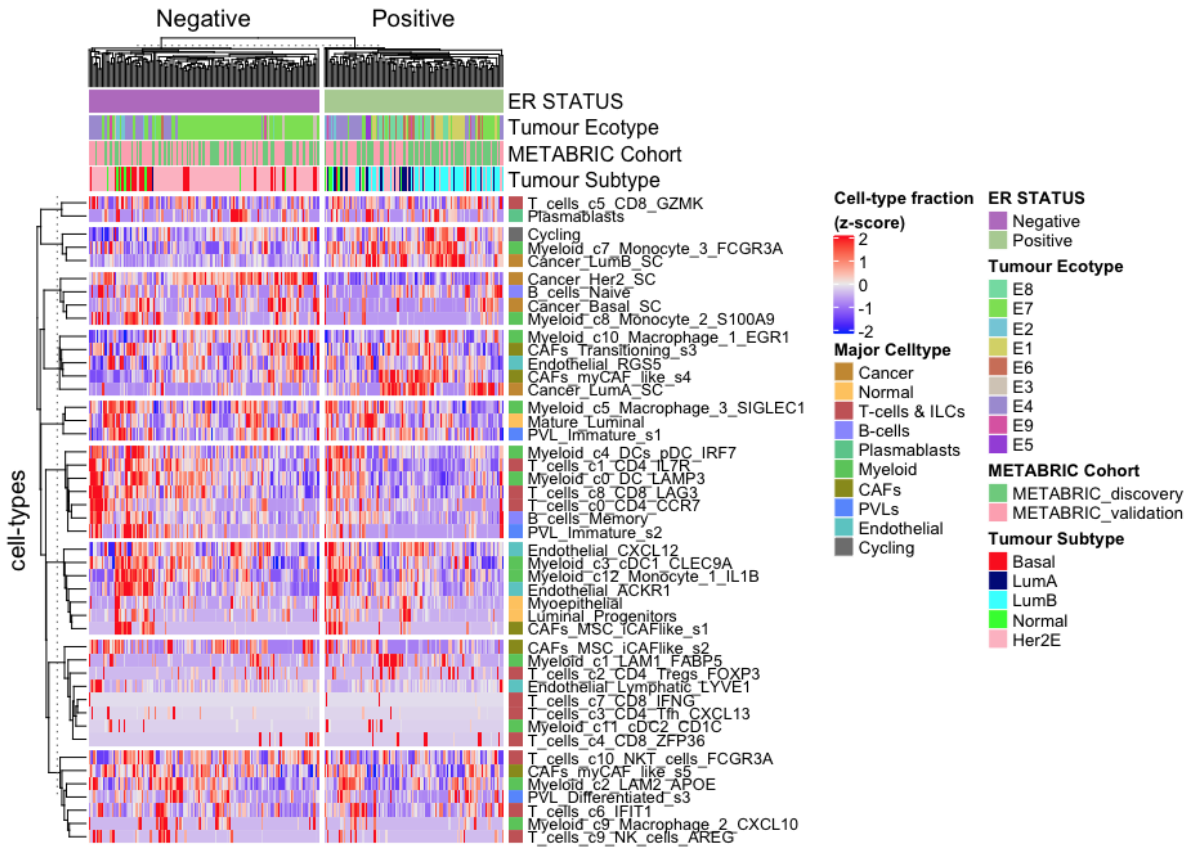

Figure R27: Heatmap of scaled CIBERSORTx cell-type abundances in all HER2+ tumours in METABRIC, stratified on ER status.

In addition, as we specifically note a survival advantage of Her2E patients in ecotype E4 compared to E7, we looked into the HER2 and ER status of these tumours (Figure R28).

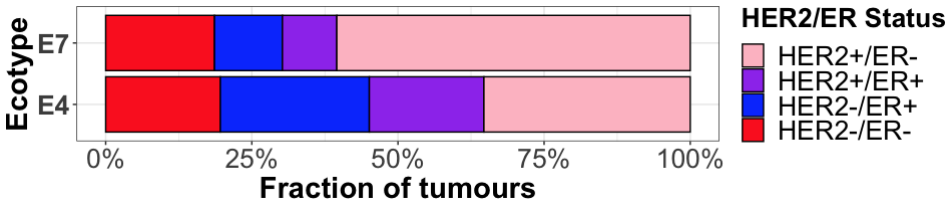

Figure R28: Proportion of Her2E METABRIC tumours assigned to ecotypes E4 and E7, stratified by both their HER2 and ER status.

Clearly there is an enrichment for HER2+/ER- tumours in E7 (which we showed to have worse prognosis) and HER2+/ER+ in E4 (which we showed to have better prognosis). This suggests that ER status may be impacting and confounding the survival analysis and that we can not with reasonable certainty state whether this survival advantage is related to the predicted cellular composition, the ER status, or a combination of the 2.

Revisions: Again we thank the reviewer for raising these important points regarding the heterogeneity of HER2+ breast cancers. In the interests of space, we have not included the analysis of the HER2+/ER+ vs HER2+/ER- cellular compositions in the manuscript, but would be willing to incorporate if the reviewer and editor consider it an important addition. In light of ER status

possibly confounding the survival analysis of the Her2E patients assigned to ecotypes E4 and E7 we have decided to take a conservative approach and remove these findings from the manuscript. In summary, we have removed Figure 6I and related text (page 23, line 8).

3.6 I typically use the term 'parenchyma' or 'parenchymal' to refer to the body of a tissue, and not specifically referring to stromal/endothelial cells as the authors seem to. I suggest reconsidering this.

Response: We thank the reviewer for their suggestion. We agree that our use of 'parenchymal cells' may be confusing. We believe 'mesenchymal' better fits with the standard terminology of stromal/endothelial cells.

Revisions: We have replaced the use of ''parenchymal' with 'mesenchymal' in all instances of the text and figures of our revised manuscript.

3.7 With respect to 'ecotypes': how useful were the true single cell data in disease stratification? If the authors use prototypical gene signatures (of the TME and/or epithelial cells), deconvolve and call subtypes again, is there a difference in outcome prediction? How well do the subtypes compare to ecotypes and PAM50?

Response: We agree with the reviewer that this is an interesting question. However, the high resolution to which we have defined our taxonomy of breast cancer cells and the novelty of our dataset means that many of our cell-types do not have suitable prototypic gene signatures available from breast cancers. In addition, we have used cell-type deconvolution methods (CIBERSORTx and DWLS) that are specifically designed to work with annotated single-cell data. Unfortunately, a consequence of this is that we are not able to directly compare the performance of specific prototypic signatures using this approach.

3.8 Pseudotime analyses are used to interpret stromal and endothelial cell profiles. Analyses of this sort should be conducted cautiously because they will always return some sort of answer. What leads the authors to believe that these cell types do in fact differentiate in this manner? Other forms of differentiation, from mesenchymal stem cells for example, are possible.

Response: We agree with the reviewer that caution should always be taken from single-cell pseudotime analyses. We would like to reiterate that analysis using conventional clustering methods including seurat v3 revealed very similar structures for each of our mesenchymal cell types (indicated by the UMAPs with pseudotime states overlaid in Fig. S10C of original manuscript). In addition, our interpretation of these results are drawn from a strong body of literature surrounding the markers expressed by each mesenchymal state, as well as several *in vitro* studies in other tissue or disease settings. We do acknowledge that this body of work should be better highlighted and clarified in the text.

Our conclusions that these stromal cells differentiate in this manner are drawn from the expression of established mesenchymal stem/stromal cell (MSC) markers in early pseudotime states. For example, early CAF state s1 expressed canonical MSC markers including *ALDH1A1*[25,26], *KLF4*[27], *CD34*[25,28] and *LEPR*[29].

These markers are expressed by a variety of normal tissue and bone marrow stem cells and have been shown to play important roles in self renewal and regulating cell differentiation. The expression of these markers all decrease as CAFs transition towards later states s5, suggesting that the earlier states resemble more of a progenitor like state. In contrast, the expression of established markers of differentiated myofibroblasts, including *ACTA2* (aSMA), are increased in later CAF states s5[30]. Differentiated myofibroblasts are further known to synthesise extracellular matrix components and remodelling enzymes[31]. To this regard, we find collagens (including *COL1A1*, *COL1A2* and *COL12A1*) and ECM remodelling proteases (including *MMP11* and *MMP14*) to be highly expressed by later CAF states s5. As we have reported in our study, these CAF states show similarities to CAF subsets described in pancreas cancer[32,33]. These studies have show that iCAFs can be induced towards myCAF-like phenotypes *in vitro* upon distinct tumor mono-, co- and trans-well culture conditions[32,33], thus suggesting that the CAF states described in our study likely differentiate in a similar manner. These studies also show the plasticity of CAF differentiation, suggesting that the CAF states described here may also be interconvertible.

Similarly for perivascular cells, early state s1 also expressed canonical MSC markers[34] including *CD44*[35], high *PDGFRB* (CD140b)[36] and *THY1* (CD90)[37]. Later perivascular state s3 had higher expression of several cytoskeletal components such as *MYH11*, *MYLK*, *TAGLN*, *ACTA2*, *CNN1*, *ACT2G,* which have been similarly described by in-vitro studies generating smooth muscle cells from MSCs[38,39].

In contrast, states of endothelial cells (ECs) did not share markers of MSCs but rather showed similarities to distinct physiological stalk and tip-like ECs. For instance, EC state s1 were enriched for canonical markers associated with stalk cells such as *ACKR1*, *SELE* and *SELP*[40], while states s2 and s3 were enriched for canonical tip cell markers such as *DLL4*[41,42]. Angiogenesis is known to be a dynamic process between stalk and tip cells, which is regulated by hypoxia, VEGF and Notch signalling[43]. Upon stimulation, stalk cells transition towards tip cells which display migratory behavior, whereas stalk cells proliferate to support sprout elongation and vessel formation. These states are thought to be dynamic and interchangeable[44], suggesting that the states described in our study can likely differentiate bi-directionally.

Lastly, we acknowledge the reviewers' statement that differentiation from other more primitive progenitor cells are possible, such as from mesenchymal stem cells. Although this is a very interesting question, mesenchymal stem cells were not detected in our dataset, likely from their rare abundance or are poorly captured using the current methodology.

Revisions: We have incorporated several studies further supporting the directionality of CAF, PVL and endothelial differentiation in our revised manuscript. Furthermore, we have included caveats of this analysis, stating that mesenchymal states may not necessarily differentiate in this manner, are likely interconvertible, and may arise from other progenitor cells like mesenchymal

stem cells. These revisions can be found in the text in page 18 lines 4-6 and page 27 lines 3-9 of the revised manuscript.

References

1. Curtis, C. *et al.* The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* **486**, 346-52 (2012).
2. Cancer Genome Atlas, N. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61-70 (2012).
3. Harrell, J.C. *et al.* Genomic analysis identifies unique signatures predictive of brain, lung, and liver relapse. *Breast Cancer Res Treat* **132**, 523-35 (2012).
4. Prat, A. *et al.* Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res* **12**, R68 (2010).
5. Millar, E. *et al.* TILs Immunophenotype in Breast Cancer Predicts Local Failure and Overall Survival: Analysis in a Large Radiotherapy Trial with Long-Term Follow-Up. *Cancers (Basel)* **12**(2020).
6. Andersson, A. *et al.* Single-cell and spatial transcriptomics enables probabilistic inference of cell type topography. *Communications Biology* **3**, 565 (2020).
7. Andersson, A. *et al.* Spatial Deconvolution of HER2-positive Breast Tumors Reveals Novel Intercellular Relationships. *bioRxiv*, 2020.07.14.200600 (2020).
8. Tsoucas, D. *et al.* Accurate estimation of cell-type composition from gene expression data. *Nat Commun* **10**, 2975 (2019).
9. Avila Cobos, F., Alquicira-Hernandez, J., Powell, J.E., Mestdagh, P. & De Preter, K. Benchmarking of cell type deconvolution pipelines for transcriptomics data. *Nat Commun* **11**, 5650 (2020).
10. Raouf, A. *et al.* Transcriptome analysis of the normal human mammary cell commitment and differentiation process. *Cell Stem Cell* **3**, 109-18 (2008).
11. Lim, E. *et al.* Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. *Nat Med* **15**, 907-13 (2009).
12. Nielsen, T.O. *et al.* A comparison of PAM50 intrinsic subtyping with immunohistochemistry and clinical prognostic factors in tamoxifen-treated estrogen receptor-positive breast cancer. *Clin Cancer Res* **16**, 5222-32 (2010).
13. Qian, J. *et al.* A pan-cancer blueprint of the heterogeneous tumor microenvironment revealed by single-cell profiling. *Cell Res* **30**, 745-762 (2020).
14. Azizi, E. *et al.* Single-Cell Map of Diverse Immune Phenotypes in the Breast Tumor Microenvironment. *Cell* **174**, 1293-1308 e36 (2018).
15. Savas, P. *et al.* Single-cell profiling of breast cancer T cells reveals a tissue-resident memory subset associated with improved prognosis. *Nat Med* **24**, 986-993 (2018).
16. Slyper, M. *et al.* A single-cell and single-nucleus RNA-Seq toolbox for fresh and frozen human tumors. *Nature Medicine* **26**, 792-802 (2020).
17. Lambrechts, D. *et al.* Phenotype molding of stromal cells in the lung tumor microenvironment. *Nat Med* **24**, 1277-1289 (2018).
18. Pliner, H.A., Shendure, J. & Trapnell, C. Supervised classification enables rapid annotation of cell atlases. *Nature Methods* **16**, 983-986 (2019).
19. Aran, D., Hu, Z. & Butte, A.J. xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol* **18**, 220 (2017).
20. Ruffell, B. *et al.* Leukocyte composition of human breast cancer. *Proc Natl Acad Sci U S A* **109**, 2796-801 (2012).
21. Amini, R.M. *et al.* Mast cells and eosinophils in invasive breast carcinoma. *BMC Cancer* **7**, 165 (2007).

22. Cassetta, L. *et al.* Human Tumor-Associated Macrophage and Monocyte Transcriptional Landscapes Reveal Cancer-Specific Reprogramming, Biomarkers, and Therapeutic Targets. *Cancer Cell* **35**, 588-602 e10 (2019).
23. Prat, A., Parker, J.S., Fan, C. & Perou, C.M. PAM50 assay and the three-gene model for identifying the major and clinically relevant molecular subtypes of breast cancer. *Breast Cancer Res Treat* **135**, 301-6 (2012).
24. Prat, A. *et al.* Molecular features and survival outcomes of the intrinsic subtypes within HER2-positive breast cancer. *J Natl Cancer Inst* **106**(2014).
25. Gentry, T. *et al.* Simultaneous isolation of human BM hematopoietic, endothelial and mesenchymal progenitor cells by flow sorting based on aldehyde dehydrogenase activity: implications for cell therapy. *Cytotherapy* **9**, 259-74 (2007).
26. Tomita, H., Tanaka, K., Tanaka, T. & Hara, A. Aldehyde dehydrogenase 1A1 in stem cells and cancer. *Oncotarget* **7**, 11018-32 (2016).
27. Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861-72 (2007).
28. Sidney, L.E., Branch, M.J., Dunphy, S.E., Dua, H.S. & Hopkinson, A. Concise review: evidence for CD34 as a common marker for diverse progenitors. *Stem Cells* **32**, 1380-9 (2014).
29. Zhou, B.O., Yue, R., Murphy, M.M., Peyer, J.G. & Morrison, S.J. Leptin-receptor-expressing mesenchymal stromal cells represent the main source of bone formed by adult bone marrow. *Cell Stem Cell* **15**, 154-68 (2014).
30. Hinz, B. *et al.* The myofibroblast: one function, multiple origins. *Am J Pathol* **170**, 1807-16 (2007).
31. Shinde, A.V., Humeres, C. & Frangogiannis, N.G. The role of alpha-smooth muscle actin in fibroblast-mediated matrix contraction and remodeling. *Biochim Biophys Acta Mol Basis Dis* **1863**, 298-309 (2017).
32. Biffi, G. *et al.* IL1-Induced JAK/STAT Signaling Is Antagonized by TGFbeta to Shape CAF Heterogeneity in Pancreatic Ductal Adenocarcinoma. *Cancer Discov* (2018).
33. Ohlund, D. *et al.* Distinct populations of inflammatory fibroblasts and myofibroblasts in pancreatic cancer. *J Exp Med* **214**, 579-596 (2017).
34. Lv, F.J., Tuan, R.S., Cheung, K.M. & Leung, V.Y. Concise review: the surface markers and identity of human mesenchymal stem cells. *Stem Cells* **32**, 1408-19 (2014).
35. Zhu, H. *et al.* The role of the hyaluronan receptor CD44 in mesenchymal stem cell migration in the extracellular matrix. *Stem Cells* **24**, 928-35 (2006).
36. Buhring, H.J. *et al.* Novel markers for the prospective isolation of human MSC. *Ann N Y Acad Sci* **1106**, 262-71 (2007).
37. Jones, E.A. *et al.* Isolation and characterization of bone marrow multipotential mesenchymal progenitor cells. *Arthritis Rheum* **46**, 3349-60 (2002).
38. Talele, N.P., Fradette, J., Davies, J.E., Kapus, A. & Hinz, B. Expression of alpha-Smooth Muscle Actin Determines the Fate of Mesenchymal Stromal Cells. *Stem Cell Reports* **4**, 1016-30 (2015).
39. Brun, J. *et al.* Smooth Muscle-Like Cells Generated from Human Mesenchymal Stromal Cells Display Marker Gene Expression and Electrophysiological Competence Comparable to Bladder Smooth Muscle Cells. *PLoS One* **10**, e0145153 (2015).
40. Thiriot, A. *et al.* Differential DARC/ACKR1 expression distinguishes venular from non-venular endothelial cells in murine tissues. *BMC Biology* **15**, 45 (2017).

41. Mailhos, C. *et al.* Delta4, an endothelial specific notch ligand expressed at sites of physiological and tumor angiogenesis. *Differentiation* **69**, 135-44 (2001).
42. Ubezio, B. *et al.* Synchronization of endothelial Dll4-Notch dynamics switch blood vessels from branching to expansion. *Elife* **5**(2016).
43. Blanco, R. & Gerhardt, H. VEGF and Notch in tip and stalk cell selection. *Cold Spring Harb Perspect Med* **3**, a006569 (2013).
44. Jakobsson, L. *et al.* Endothelial cells dynamically compete for the tip cell position during angiogenic sprouting. *Nat Cell Biol* **12**, 943-53 (2010).

Decision Letter, first revision:

22nd Feb 2021

Dear Alex,

Our thanks to you and your co-authors for your patience during the prolonged review of your revised manuscript.

Your Article, "An integrated multi-omic cellular atlas of human breast cancers" has now been seen by 2 referees; as I've previously informed you, the final reviewer was unable to submit a report due to extenuating circumstances and we concluded that there was sufficient support to make a decision at this point.

You will see from their comments below that while they find your work has substantially improved in revision and both referees are now supportive of publication, there are still a small number of important points remaining. We would like to consider your response to these concerns in the form of a revised manuscript before we make a final decision on publication.

In brief, reviewer #1 suggests including some of the additional figures presented in your rebuttal letter in the revised article; and ensuring that the spatial transcriptomic data from the Andersson et al. manuscript is available at publication.

Reviewer #3 has two more substantial comments: the first, suggesting that the question of HER2 clinical/intrinsic subtype discordance could be more thoroughly investigated using your dataset; and the deconvolution could be repeated using "conventional" methods.

In our editorial view, we think that all of these points should be addressed before publication; we think that the clear guidance provided means that any further proposed work is not unduly onerous. We also note that reviewer #3 does not outright request additional in-depth analysis of HER2 status discordance (saying, "in any case", that further explanation is "warranted"), but we agree with their suggestion that clear conclusions could likely be drawn from your rich dataset.

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

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

We are committed to providing a fair and constructive peer-review process. Do not hesitate to contact

us if there are specific requests from the reviewers that you believe are technically impossible or unlikely to yield a meaningful outcome.

When revising your manuscript:

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

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

*3) Include a revised version of any required Reporting Summary:
https://www.nature.com/documents/nr-reporting-summary.pdf
It will be available to referees (and, potentially, statisticians) to aid in their evaluation if the manuscript goes back for peer review.
A revised checklist is essential for re-review of the paper.

Please be aware of our guidelines on digital image standards.

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

[REDACTED]

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

We hope to receive your revised manuscript within four to eight weeks. If you cannot send it within this time, please let us know.

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

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

We look forward to seeing the revised manuscript and thank you for the opportunity to review your work.

Sincerely,

Michael Fletcher, PhD
Associate Editor, Nature Genetics

ORCID: 0000-0003-1589-7087

Reviewers' Comments:

Reviewer #1:
Remarks to the Author:
The authors have capably addressed all of our concerns related to statistical considerations, sample representation in the figures, as well as performing additional validations.

The authors provide important additional analyses in the rebuttal but some of these are not included in the actual manuscript. For example, Figure R1 is not included in the revision for "in the interests of space". We recommend that this figure in particular be included in the publication though as it does add to an understanding of the dataset.

With respect to the public release of ST datasets discussed on page 13 of the rebuttal, the authors expect simultaneous publication of the two manuscripts (ref. 7, Andersson et al.). We suggest ensuring that the data is available upon publication, since ST is a central aspect of this study.

Reviewer #3:
Remarks to the Author:
The authors have adequately addressed most of my concerns. My congratulations to them on this substantial advance in our understanding of cellular heterogeneity in breast cancer.

Two points are outstanding:

3.4 I understand that the authors are distinguishing between the HER2 intrinsic subtype and clinical HER2 positivity. The significance of the HER2 intrinsic subtype, where discordance with the clinical assay is found, remains unclear. The authors explain that these differences arise owing to 'transcriptional similarities'. This is rather ambiguous, and may leave some readers scratching their heads. The single cell data that the authors have generated gives them an opportunity to determine whether intratumoural cellular heterogeneity may explain discordance between clinical and intrinsic subtype designations. In any case, a more extensive explanation is warranted as the significance of HER2E in HER2 negative tumours remains unclear.

3.7 I expressed my point poorly. My question is: if the authors were to use a conventional deconvolution method (based on a reference prototypical gene signature matrix) such as xCell to deconvolve bulk data, and then call ecotypes based on these scores, what would the resulting tumour

subtypes (ecotypes) look like? How would they compare to the tumour subtypes called based on the authors' scRNAseq data? And would their prognostic impact be lower, comparable or greater?

Author Rebuttal, first revision:

Wu et al.
Detailed responses and revisions

POINT BY POINT RESPONSE TO REVIEWERS:

Reviewer #1:
Remarks to the Author:
The authors have capably addressed all of our concerns related to statistical considerations, sample representation in the figures, as well as performing additional validations.

Response: We thank the reviewer for their supportive comments.

1.1. The authors provide important additional analyses in the rebuttal but some of these are not included in the actual manuscript. For example, Figure R1 is not included in the revision for "in the interests of space". We recommend that this figure in particular be included in the publication though as it does add to an understanding of the dataset.

Response: We agree with the reviewer and have included this figure as supplementary figure S5B in the paper.

Revision: We have added this figure as fig S5B and re-numbered other panels in this figure accordingly. We have edited the revised manuscript on page 10, lines 24-26 and also the legend text of supplementary figure 5.

1.2. With respect to the public release of ST datasets discussed on page 13 of the rebuttal, the authors expect simultaneous publication of the two manuscripts (ref. 7, Andersson et al.). We suggest ensuring that the data is available upon publication, since ST is a central aspect of this study.

Response: We agree with the reviewer that all spatial transcriptomics data should be made available upon publication. All relevant ST data used in our study from the Andersson et al. manuscript has since been deposited into the Zenodo public data repository (DOI: 10.5281/zenodo.3957257).

Revision: We have added these details into the Data Availability section of our revised manuscript on page 42 line 25, page 43 lines 1-2.

Reviewer #3:

Remarks to the Author:

The authors have adequately addressed most of my concerns. My congratulations to them on this substantial advance in our understanding of cellular heterogeneity in breast cancer.

Response: We thank the reviewer for their supportive comments.

Two points are outstanding:

3.4 I understand that the authors are distinguishing between the HER2 intrinsic subtype and clinical HER2 positivity. The significance of the HER2 intrinsic subtype, where discordance with the clinical assay is found, remains unclear. The authors explain that these differences arise owing to 'transcriptional similarities'. This is rather ambiguous, and may leave some readers scratching their heads. The single cell data that the authors have generated gives them an opportunity to determine whether intratumoural cellular heterogeneity may explain discordance between clinical and intrinsic subtype designations. In any case, a more extensive explanation is warranted as the significance of HER2E in HER2 negative tumours remains unclear.

Response: We agree with the reviewer that the relationship between HER2 clinical and intrinsic subtypes is an interesting question, and this topic has been extensively addressed using bulk tumor profiling (see Prat *et al.*[1] and Schettini *et al.*[2]). We also agree that insights into intratumoural cellular heterogeneity using scRNAseq may help to resolve the differences seen when comparing IHC status versus multi-gene gene expression subtype. However, the present study was designed to comprehensively map the cellular landscape of breast cancer across many cell types and all clinical subtypes, so this is not an appropriate cohort in which to answer this specific question concerning a minority clinical subtype. Our study includes 3 clinically Her2+/ER- and 1 Her2+/ER+ case. Similarly, when we select non-HER2+ cases in which there are at least 100 HER2E_SC cells, we identify 1 ER+ case and 3 TNBC cases. Therefore we respectfully suggest that the

reviewer's question would be best addressed by a focused study of neoplastic heterogeneity across a substantial number of HER2+ cases. To perform detailed comparative analyses of single cell agreement or discordances here would not be robust and likely subject to overfitting.

3.7 I expressed my point poorly. My question is: if the authors were to use a conventional deconvolution method (based on a reference prototypical gene signature matrix) such as xCell to deconvolve bulk data, and then call ecotypes based on these scores, what would the resulting tumour subtypes (ecotypes) look like? How would they compare to the tumour subtypes called based on the authors' scRNAseq data? And would their prognostic impact be lower, comparable or greater?

Response: We thank the reviewer for clarifying their question and their interest in our tumour ecotyping approach. As requested we have used xCell and its associated suite of 64 cell-type signatures to estimate the enrichment of cell-type abundance in the METABRIC cohorts (Figure R1). To assess the performance of the approach, we first identified the most enriched cell-types across all ~2000 breast cancers (Figure R1). Although epithelial cells were highly enriched, as expected, the most highly enriched cell-type was MPP (Multi-Potent Progenitors). This is surprising as it is not a cell-type that, to our knowledge, is known to be present in breast cancers. Additionally, Chondrocytes and Keratinocytes were also within the top 10 cell types enriched. The predicted high abundance of these cell types immediately highlight the limitations of biological interpretation, in a breast cancer specific context, based on the cell-type signatures derived from diverse tissue and disease contexts in the xCell dataset.

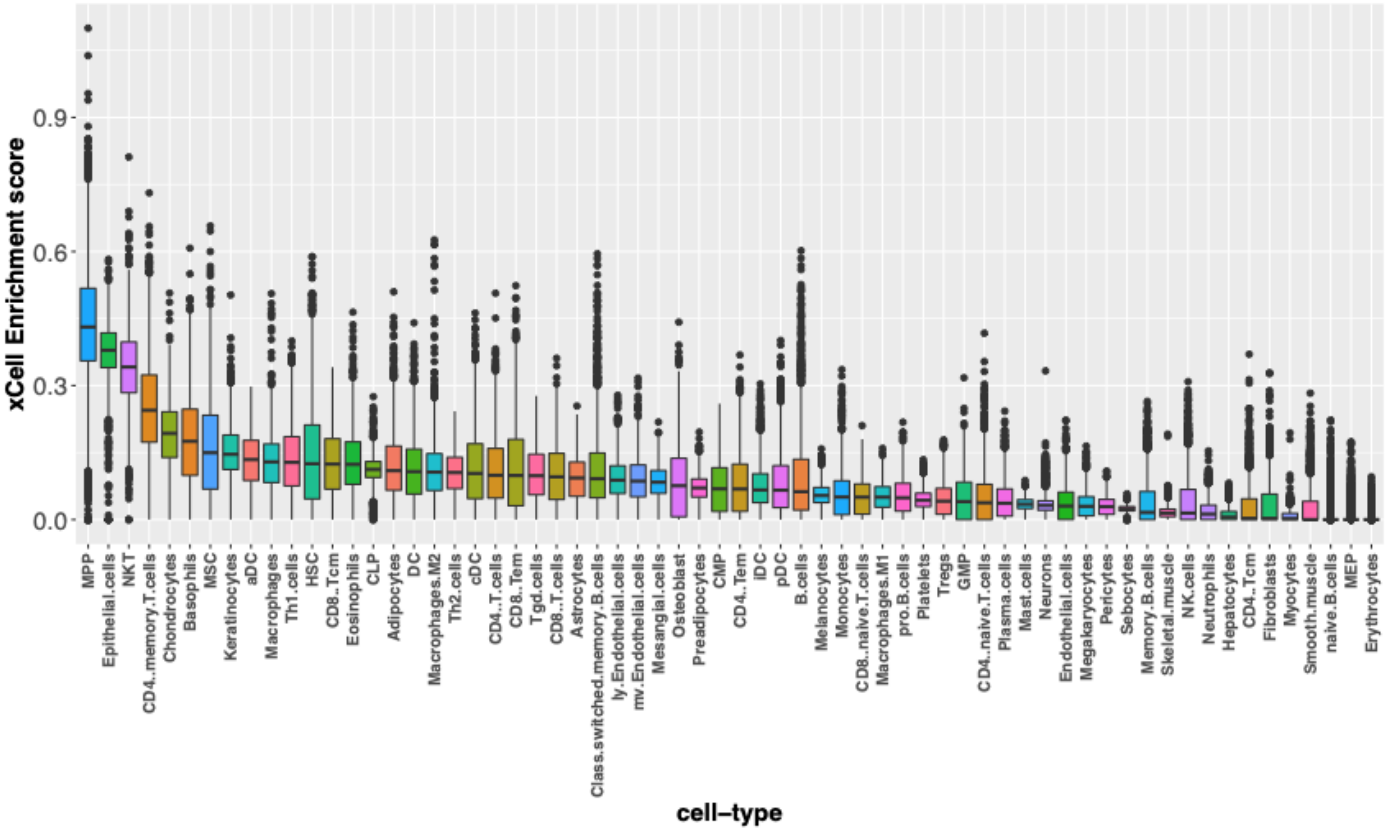

Figure R1: *Boxplots showing the distributions of xCell identified cell-types, across all METABRIC samples, ordered by the median enrichment score.*

Regardless, we proceeded to generate tumour ecotypes using our consensus clustering approach. This resulted in three ecotype clusters. To allow a direct comparison between the xCell-derived ecotypes and those using our detailed breast cancer cell taxonomy, we set the number of ecotype clusters to be 9 (Figure R2).

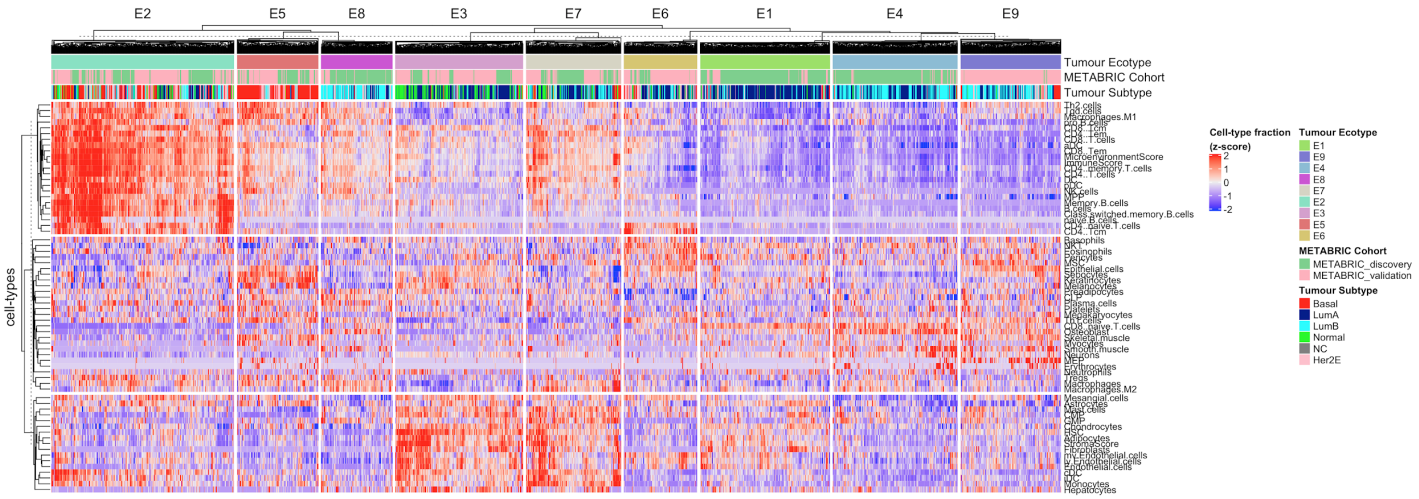

Figure R2: *Heatmap showing the pattern of cell-type enrichment across the 9 ecotypes generated with xCell on the METABRIC patient cohort.*

There were some similarities between the tumour ecotypes generated from the two approaches, for example: (i) a poor-prognosis ecotype, consisting predominantly of basal-like tumours with a low immune-score (E5, in figures R2 and R3). This most closely corresponds (62% overlap) to ecotype E3 in figure 6 of our manuscript; and (ii) a good prognosis ecotype, consisting mostly of LumA and normal-like tumours (E3, in figures R2 and R3). This most closely corresponds (54% overlap) to ecotype E2 in figure 6 of our manuscript.

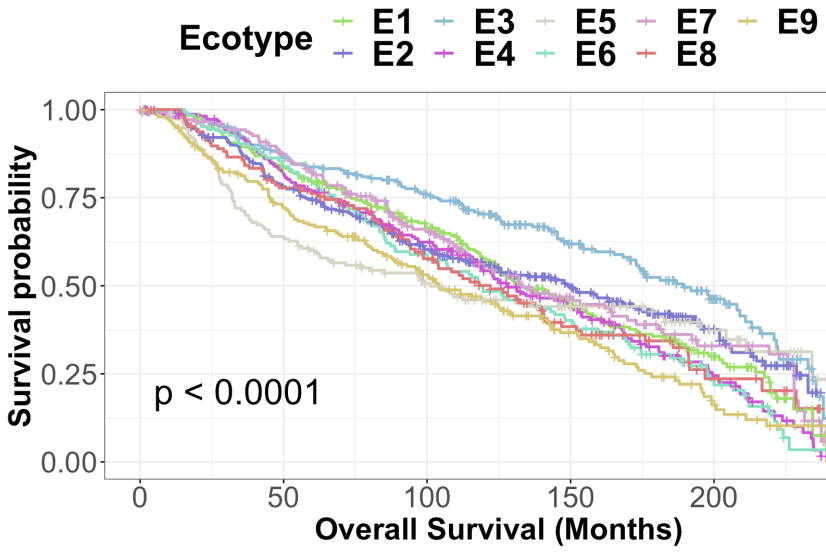

Figure R3: *Survival analysis of xCell derived ecotypes from the METABRIC cohort.*

However, significant limitations of the xCell method were also apparent:

- It was not able to identify the poor prognosis Her2E group of patients that was identified as ecotype E7 when using our single-cell derived cell-types.
- The ecotyping results using prototypical signatures are biologically difficult to interpret. For example, the xCell E5 ecotype, is defined by cell-types including sebucytes and keratinocytes (figure R2), both specialised cells found in the skin but not the breast. In contrast, our corresponding brca-signature derived ecotype (E3), is defined by the presence of: Basal_SC, Luminal-progenitor and cycling cells, which are all hallmarks of basal-like breast cancers, indicating that the method is performing well.
- Finally, the ecotype clustering appears to be heavily influenced by batch effects associated with each cohort. For example, xCell derived ecotypes E4 and E9 almost exclusively consist of tumours from the METABRIC discovery and validation cohorts, respectively (Figure R2). We consider this to be a significant limitation that will markedly limit the ability to carry out robust, integrated, meta-analyses of tumour ecotypes across datasets. Importantly, our method using CIBERSORTx and our BrCa single-cell data does not appear to significantly suffer from this phenomenon (see figure 6 of our manuscript).

For these reasons we conclude that our approach, driven by our novel disease-specific cellular signatures and using deconvolution methods specifically developed for working with single-cell datasets, are more robust than those that utilise a broader set of gene-signatures such as xCell.

References

1. Prat, A. *et al.* Molecular features and survival outcomes of the intrinsic subtypes within HER2-positive breast cancer. *J Natl Cancer Inst* **106**(2014).
2. Schettini, F. *et al.* HER2-enriched subtype and pathological complete response in HER2-positive breast cancer: A systematic review and meta-analysis. *Cancer Treat Rev* **84**, 101965 (2020).

Decision Letter, second revision:

Our ref: NG-A55835R1

13th Apr 2021

Dear Alex,

Thank you for submitting your revised manuscript "An integrated multi-omic cellular atlas of human breast cancers" (NG-A55835R1). It has now been seen by the original Reviewer #3 and their comment is below. They find that the paper has improved in revision, and therefore we'll be happy in principle to publish it in Nature Genetics, pending minor revisions to satisfy the referees' final requests and to comply with our editorial and formatting guidelines.

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

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

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

Sincerely,

Michael Fletcher, PhD
Associate Editor, Nature Genetics

ORCID: 0000-0003-1589-7087

Reviewer #3 (Remarks to the Author):

The authors have adequately addressed all of my remaining questions.

Final Decision Letter:

In reply please quote: NG-A55835R2 Swarbrick

8th Jul 2021

Dear Alex,

I am delighted to say that your manuscript "A single-cell and spatially resolved atlas of human breast cancers" has been accepted for publication in an upcoming issue of Nature Genetics.

Prior to setting your manuscript, we may make minor changes to enhance the lucidity of the text and with reference to our house style. We therefore ask that you examine the proofs most carefully to ensure that we have not inadvertently altered the sense of your text in any way.

Once your manuscript is typeset and you have completed the appropriate grant of rights, you will receive a link to your electronic proof via email with a request to make any corrections within 48 hours. If, when you receive your proof, you cannot meet this deadline, please inform us at rjsproduction@springernature.com immediately.

Your paper will be published online after we receive your corrections and will appear in print in the next available issue. You can find out your date of online publication by contacting the Nature Press Office (press@nature.com) after sending your e-proof corrections. Now is the time to inform your Public Relations or Press Office about your paper, as they might be interested in promoting its publication. This will allow them time to prepare an accurate and satisfactory press release. Include your manuscript tracking number (NG-A55835R2) and the name of the journal, which they will need when they contact our Press Office.

Before your paper is published online, we shall be distributing a press release to news organizations worldwide, which may very well include details of your work. We are happy for your institution or funding agency to prepare its own press release, but it must mention the embargo date and Nature Genetics. Our Press Office may contact you closer to the time of publication, but if you or your Press Office have any enquiries in the meantime, please contact press@nature.com.

Acceptance is conditional on the data in the manuscript not being published elsewhere, or announced in the print or electronic media, until the embargo/publication date. These restrictions are not intended to deter you from presenting your data at academic meetings and conferences, but any enquiries from the media about papers not yet scheduled for publication should be referred to us.

Please note that Nature Genetics is a Transformative Journal (TJ). Authors may publish their research with us through the traditional subscription access route or make their paper immediately open access through payment of an article-processing charge (APC). Authors will not be required to make a final decision about access to their article until it has been accepted. Find out more about Transformative Journals

<B>Authors may need to take specific actions to achieve compliance with funder and institutional open access mandates. For submissions from January 2021, if your research is supported by a funder that requires immediate open access (e.g. according to Plan S principles) then you should select the gold OA route, and we will direct you to the compliant route where possible. For authors selecting the subscription publication route our standard licensing terms will need to be accepted, including our self-archiving policies. Those standard licensing terms will supersede any other terms that the author or any third party may assert apply to any version of the

manuscript.

Please note that Nature Research offers an immediate open access option only for papers that were first submitted after 1 January, 2021.

You will not receive your proofs until the publishing agreement has been received through our system.

If you have any questions about our publishing options, costs, Open Access requirements, or our legal forms, please contact ASJournals@springernature.com

If you have posted a preprint on any preprint server, please ensure that the preprint details are updated with a publication reference, including the DOI and a URL to the published version of the article on the journal website.

To assist our authors in disseminating their research to the broader community, our SharedIt initiative provides you with a unique shareable link that will allow anyone (with or without a subscription) to read the published article. Recipients of the link with a subscription will also be able to download and print the PDF.

As soon as your article is published, you will receive an automated email with your shareable link.

You can now use a single sign-on for all your accounts, view the status of all your manuscript submissions and reviews, access usage statistics for your published articles and download a record of your refereeing activity for the Nature journals.

An online order form for reprints of your paper is available at https://www.nature.com/reprints/author-reprints.html. Please let your coauthors and your institutions' public affairs office know that they are also welcome to order reprints by this method.

If you have not already done so, we invite you to upload the step-by-step protocols used in this manuscript to the Protocols Exchange, part of our on-line web resource, natureprotocols.com. If you complete the upload by the time you receive your manuscript proofs, we can insert links in your article that lead directly to the protocol details. Your protocol will be made freely available upon publication of your paper. By participating in natureprotocols.com, you are enabling researchers to more readily reproduce or adapt the methodology you use. Natureprotocols.com is fully searchable, providing your protocols and paper with increased utility and visibility. Please submit your protocol to https://protocolexchange.researchsquare.com/. After entering your nature.com username and password you will need to enter your manuscript number (NG-A55835R2). Further information can be found at https://www.nature.com/nprot/.

Sincerely,

Michael Fletcher, PhD
Associate Editor, Nature Genetics

ORCID: 0000-0003-1589-7087