Integrative spatial omics reveals distinct tumor-promoting multicellular niches and immunosuppressive mechanisms in African American and European American patients with TNBC

Qian Zhu[1,2,3,*,#], Akhila Balasubramanian[4,*], Jaya Ruth Asirvatham[5,*], Badrajee Piyarathna[4], Jaspreet Kaur[6], Nada Mohamed[5], Ling Wu[1], Megha Chatterjee[3], Stacy Wang[1,2], Niloufar Pourfarrokh[5], Uttam Rasaily[3], Yitian Xu[7], Junjun Zheng[7], Deborah Jebakumar[5], Arundathi Rao[5], Shu-hsia Chen[7], Yi Li[1,3,4], Eric Chang[1,3], Xiaoxian Li[8], Ritu Aneja[6], Xiang H.-F. Zhang[1,3,4], Arun Sreekumar[3,4,#]

[*]Contributed equally

[#]Corresponding authors; email: qian.zhu@bcm.edu and arun.sreekumar@bcm.edu

[1]Lester and Sue Smith Breast Center, Baylor College of Medicine, Houston, Texas 77030, USA

[2]Department of Human Molecular Genetics, Baylor College of Medicine, Houston, Texas 77030, USA

[3]Dan L Duncan Comprehensive Cancer Center, Baylor College of Medicine, Houston, Texas 77030, USA

[4]Department of Molecular Cell Biology, Baylor College of Medicine, Houston, Texas 77030, USA

[5]Department of Pathology, Baylor Scott and White Hospital, Temple, Texas 76508, USA

[6]Institute of Biomedical Sciences, Georgia State University, Atlanta, Georgia 30302, USA; current address: School of Health Professionals, The University of Alabama at Birmingham, Alabama 35294, USA

[7]Immune Monitoring Core, Houston Methodist Research Institute, Houston, Texas 77030, USA

[8]Department of Pathology and Laboratory Medicine, Emory University, Atlanta, Georgia 30322, USA

Keywords: triple-negative breast cancer (TNBC), African American, European American, tumor microenvironment.

Abstract

Racial disparities in triple-negative breast cancer (TNBC) outcomes have been reported. However, the biological mechanisms underlying these disparities remain unclear. We integrated imaging mass cytometry and spatial transcriptomics, to characterize the tumor microenvironment (TME) of African American (AA) and European American (EA) patients with TNBC. The TME in AA patients was characterized by interactions between endothelial cells, macrophages, and mesenchymal-like cells, which were associated with poor patient survival. In contrast, the EA TNBC-associated niche is enriched in T-cells and neutrophils suggestive of an exhaustion and suppression of otherwise active T cell responses. Ligand-receptor and pathway analyses of race-associated niches found AA TNBC to be “immune cold” and hence immunotherapy resistant tumors, and EA TNBC as ‘inflamed’ tumors that evolved a distinctive immunosuppressive mechanism. Our study revealed the presence of racially distinct tumor-promoting and immunosuppressive microenvironments in AA and EA patients with TNBC, which may explain the poor clinical outcomes.

Statement of Significance

We have uncovered distinct tumor-promoting and immunosuppressive microenvironments that may contribute to the observed disparities in patient outcomes. Our findings lay the groundwork for developing race-specific therapeutic interventions, potentially improving outcomes for patients with TNBC. Given that chemotherapy and anti-PD1 therapy are currently available treatments for TNBC, our finding of race-specific spatial multicellular niche-associated cell-cell interactions suggest that these treatments may have racially distinct response profiles among TNBC patients. Hence, our novel insights on molecular and cellular interactions in TNBC stratified by race has the potential to inform personalized treatment strategies.

Introduction

A profound racial disparity has been identified between African American (AA) women and their European American (EA) counterparts with respect to the incidence and clinical trajectory of breast cancer (BCa). Specifically, AA women develop BCa at a relatively younger age(1), have a two-fold higher chance of developing Triple Negative (TN) tumors(2), with 40-70 % increased risk of developing stage IV disease(3). Furthermore, cancer-related mortality in AA women was more likely for TNBC than other subtypes, even after adjusting for age, stage, treatment, and socio-economic status(4). Genetic, environmental, and healthcare access/utilization factors may all contribute to this disparity(5). Many studies have examined the genetic component to understand the biological underpinnings of racial disparity(6), while other key factors still need examination and identification.

Studies examining global discovery of gene expression signatures that distinguish AA vs EA TNBC tumors have reported distinct tumor-associated immunologic profiles in AA patients(7) and in patients of African descent(8). The bulk-level RNA sequencing approach used in these studies revealed heterogeneity in TNBC and provided cues on involvement of unique cellular level interactions in the tumor microenvironment (TME), which were not defined. Notably, variations in the composition of the TME between AA and EA BCa patients have been reported(9,10), wherein computational deconvolution of bulk samples has predicted existence of cell types in various combinations in spatial niches. These include Tregs and naïve B cells that have been shown to exhibit positive correlations with African

ancestry(8), while activated mast cells negatively correlated with African ancestry(8). Further, the level of tumor-associated lymphocytes was found to be similar between AA and EA TNBCs(7). These initial findings suggest that the predicted cell types could exist in multiple different combinations, in distinct spatial niches, that could be associated with differences in clinical outcomes in TNBC. Hence, dissecting the spatial granularity in niches within the TME would be of paramount importance to obtain deeper insights into the altered biology associated with AA and EA TNBC.

To address this unmet need, recent studies employing spatial proteomic(11,12) and transcriptomic(13) technologies have provided insights into the molecular, cellular, and spatial phenotypes governing cancer metastasis(14) and tumor recurrence(15). Previous multi-omics experiments have facilitated the creation of machine-learning predictors for treatment response and enhanced our understanding of the biological mechanisms behind molecular cancer phenotypes(16-18). Spatial profiling of proteins in AA and EA patients with BCa revealed racial differences in the proteomes of tumors, although no significant correlation with survival was observed(19). Moreover, spatial transcriptomic analysis of a racially diverse TNBC cohort revealed racial differences in hypoxic tumor content and regions of immune-rich infiltrates(20). However, the composition and spatial heterogeneity of these immune-rich infiltrates remain unclear. Furthermore, whether racial differences exist in the spatial interactions between various components within the TME and their association with disease outcomes remains unknown. Although imaging mass cytometry (IMC)(11), multiplexed ion beam imaging (MIBI)(12), and single-cell RNA sequencing(21) have been used to characterize the TME of TNBC, all patients included in prior studies were of European descent.

We believe that a spatial multi-omics study of a racially diverse TNBC cohort should permit one to understand the biological factors underpinning the large survival gap currently experienced by AA patients with TNBC, while also laying the foundation for developing race-specific therapeutic interventions. To address the caveats of prior studies, for the first time, we have conducted an integrative analysis combining information obtained from spatial single-cell level IMC(22) and re-analysis of spatial transcriptomics data(20) to uncover racially enriched spatial cell-cell interactions that define unique AA and EA TNBC-associated niche. These cell-cell interactions and race-specific niche correlate with patient survival and infer molecular characteristics of EA and AA tumors with a spatial resolution. IMC data measured the expression of

30 proteins at the single cell level, which was then integrated with existing 10X Visium ST data(20), followed by validation of the findings using an independent IMC, 10X ST and Nanostring region of interest (ROI)-based transcriptomics analysis.

Our results for the first time highlight the existence of race-specific multicellular niches and immunosuppressive mechanisms involving tumor cells, immune cells, and endothelial cells in TNBC. Of note, these insights were not discernible from previous bulk-level RNA sequencing studies, as well as from unimodal studies of spatial proteomics or transcriptomics, thus underscoring the importance of spatial multi-omics integration to understand race-specific biological mechanisms.

Results

Overview of IMC

We constructed a racially balanced and clinically matched cohort of 57 surgically resected tissues (majority TNBC, refer **Supplementary Table 1** for clinical information and **Supplementary Figure 1** shows the H&E images) from 26 self-reported African American (AA), and 31 self-reported European American (EA) women. Self-reported race was verified based on a gene expression signature that was found to be upregulated in Western African patients with TNBC(8) (**Supplementary Fig. 2**). All the patients examined have survival outcome information at 10 years. To survey the single-cell spatial interaction landscape, we performed imaging mass cytometry (IMC) on multiple regions of interest (ROIs, see **Supplementary Figure 3**) selected from individual tumors laid out in a tissue microarray (TMA) setting, using a panel of 26 antibodies of immune-regulatory, stromal, and epithelial proteins (**Figure 1 a**).

To begin with, we segmented over 270,000 single cells with spatial coordinates. Using unsupervised clustering with a high number of initializations, we derived 20 high-quality single-cell clusters based on marker protein expression across single-cells (**Figure 1b**). t-SNE plot shows clear demarcation of the single-cell clusters (**Figure 1c**). Differential analysis across the 20 clusters identified uniquely expressed proteins within each cluster. Most clusters contained co-expression of a few markers, after which they were named. For

example, cluster 16 named “16 PanCK VEGF E-Cadherin” contained co-expression of PanCK, VEGF, and E-Cadherin. The 20 clusters covered major epithelial compartment (Cluster 19: PLK1 PD1; Cluster 2: PanCK; Cluster 8: PLK1 PanCK Ki67; Cluster 16: PanCK VEGF E-cadherin), immune compartment (Cluster 12: CD16 CD163 PDL1; Cluster 3: CD11c; Cluster 7: CD152 CD8a; Cluster 10: CD68; Cluster 5: CD11c FoxP3 pHH3; Cluster 15: CD4 CD3), stromal compartment (Cluster 13: CD31 Vimentin AR; Cluster 4: Vimentin PDL1 PLK1), and cells in proliferative, cytotoxic, and hypoxic states (Cluster 6:HIF1a; Cluster 18: Ki67 KIFC1; Cluster 9: GranzymeB CD152 HIF1a, **Supplementary Figure 4a**). **Figure 1d** shows the overlay of expression of E-Cadherin, CD45RA, CD31, CD68, and Vimentin, on the t-SNE, showing distinct clustering patterns of the proteins and their spatial organization in the IMC images (**Figure 1e**).

Spatial cell-cell interactions, rather than cluster abundances, reveal strong race-specific differences predictive of clinical outcome

We next plotted the relative abundance of clusters across all ROIs from all patients belonging to both the races. Interestingly, although most clusters were quite similar in abundances between races, a few were slightly more differentially enriched in a subset of AA versus EA patients. Thus, for example, Cluster 20: CD31 CD45RA (signifying endothelial cells anchored by naïve T cells) was more abundant in subset of AA TNBC, and Cluster 7: CD152 CD8a (signifying exhausted cytotoxic T cells) was more in a subset of EA TNBC. However, as each patient expresses a different combination of markers, the global distribution of cluster abundance was unable to segregate TNBC patients by race for most patients (**Supplemental Figure 4b**).

Given that relative abundances were not sufficient to explain the differences between AA and EA TNBC patients, we hypothesized that the spatial cell-cell interactions might be a stronger indicator of race-specific differences and clinical outcome. Hence, we investigated the spatial interactions between cells of these 20 clusters. We performed spatial proximity enrichment/depletion analysis with Giotto(23) (see Methods) by utilizing all ROIs from all patients. This method constructs a Delauney spatial graph between adjacent cells and computes the number of adjacencies between cells in pairs of clusters in real and randomly shuffled dataset (**Figure 2a**). Differential AA vs EA interactions were selected using a mixed-effects statistical model that accounts for the group-wise repeated observations of patient

ROIs. Using this approach, we identified nine AA-specific, and six EA-specific statistically significant spatial cell-cell interactions (**Figures 2b and 2c**). In AA, five of nine cell-cell interactions were homo-typic interactions suggesting clustering aggregates, and exclusion (**Figure 2b**). In contrast, hetero-typic interactions were observed across all the six EA specific interactions (**Figure 2c**). Despite our earlier findings showing that cluster abundances are not statistically different between races (**Supplemental Figure 4b**), the spatial interactions between the clusters reveal intriguing racially distinct differences. Consistent with this, a closer examination of spatial interactions enriched in AA (**Figure 2b**) revealed many of the top ranked AA-specific interactions are centered around clusters CD31 CD45RA (Cluster 20, denoting endothelial cells anchored by naïve T cells), CD31 Vim (Cluster 13, denoting endothelial cells anchored by Vim^{+} cells), and CD16 CD163 CD68 (Cluster 12, denoting macrophages). Thus for example an interaction denoted by, CD31 Vim -- PLK1 PD1 describes an interaction between Clusters 13 and 19 (**Figure 2b**). On the other hand, interactions in EA patients more frequently involve Granzyme B CD152 HIF1a (Cluster 9, denoting exhausted and cytotoxic immune cells), which interacts with CD68 (Cluster 10, macrophages), CD11c pHH3 (Cluster 3, denoting proliferative dendritic cells), and CD152 CD8a (Cluster 7, denoting cytotoxic and exhausted T cells) (**Figure 2c**). **Figure 2d** shows the spatial representation of the top-ranked AA and EA cell-cell interactions in patient samples. AA-associated interactions are more enriched in AA patients (**Figure 2d**). Similarly, EA-associated interactions are enriched in EA patients (**Figure 2e**). We further verified the endothelial-macrophage interaction observed in AA TNBC using multiplex immunofluorescence staining (**Figure 2f**). Importantly, as expected, using pixel-level CD31-CD163 co-localization analysis, significantly higher co-localization between these two cell types is seen in AA compared to EA TNBC (**Figure 2g**). This observation holds true even after accounting for the number of cells per image (**Figure 2 h**).

To determine the clinical relevance of AA and EA-associated interactions, we examined the association between the degree of AA- and EA-associated interactions within patient ROIs and overall survival (OS) using Kaplan Meier (KM) curves. When the enrichment scores of the top nine AA-associated interactions were combined, a higher interaction summary score was significantly correlated with poor OS in patients with TNBC (**Fig. 3a**). In contrast, no correlation with OS was observed when the combined top six EA-associated interactions were used (**Fig. 3b**). Analysis combining the interaction scores with race revealed a

significant separation of the OS curves for both AA and EA patients, with the separation being more pronounced in AA patients (**Fig. 3c**). Combining EA-associated interaction scores with race as an additional variable provided a significant separation of the OS curves only within the EA group, demonstrating the race-dependent prognostic value of interactions (**Fig. 3d**). These results show that distinct spatial cell-cell interactions shaping the AA and EA tumor-immune microenvironment, with distinct interactions having both race independent and dependent prognostic value.

Tumor microenvironment (TME) architecture of AA tumors is composed of endothelial cells, macrophages, and tumors in a mesenchymal state

Given the limited coverage of protein markers examined using IMC, we expanded the analysis by incorporating spatial transcriptomic data, with the purpose of more extensively characterizing the unique multicellular niches in AA and EA TNBC. We integrated the findings from our IMC data with a publicly available racially balanced TNBC spatial transcriptomic data collection derived from flash frozen TNBC(20). Specifically, the proteins associated with specific AA and EA-associated cell-cell interactions described above were used to derive the corresponding gene queries (AA-Q1 through AA-Q5, and EA-Q1 through EA-Q5), whose expression and spatial localization patterns were interrogated in the TNBC spatial transcriptomics dataset (**Figure 4a**). This has allowed us to validate the AA and EA associated cell-cell interactions identified by IMC in spatial transcriptomics data from TNBC tumors (**Figure 4a). Figure 4b** describes the AA-specific cell-cell interactions (AA-Q1 to AA Q5) that were mapped using this approach. Interestingly, endothelial, and mesenchymal (vimentin)-related cell-cell interactions (AA-Q3, **Figure 4b**), measured by the number of spots exhibiting the co-expression of interaction genes (see Methods), are much more abundant in AA than EA patients (**Figure 4c** black outlines). These spots containing Endothelial-Mesenchymal cell interactions are more spatially clustered in AA than EA samples (see black outlines in **Figure 4c** and spatial clustering score). In contrast, EA-patients displayed a random and uniform distribution for genes in AA-Q3 (**Figure 4c**, EA), indicating that endothelial markers and vimentin (mesenchymal marker) are either not co-expressed or expressed at a lower level in EA patients. Interestingly, as illustrated by a representative AA and EA patient sample, we observe that the top 5 AA-specific cell-cell interactions (AA-Q1 and AA-Q5) are also confined to the same regions in the AA but not in the EA (**Figures 4d**, and **4e**), leading us to believe that components of the 5 AA-specific cell interactions may spatially overlap with each other. To test this, we computed the correlation between the spatial

co-localization profiles of the top 5 queries (AA-Q1 to AA-Q5) in 10 AA samples and 10 EA patient tumors from the TNBC spatial transcriptomics dataset(20) (**Figure 4f**). We observed a higher degree of within-sample cross-query correlation between AA-Q1 to AA-Q5 in AA than EA (**Figure 4f**). Most prominently, AA-Q3 spatially correlates strongly with AA-Q1 through Q2, and Q4 through Q5, but this spatial correlation is entirely missing in EA samples **(Figure 4f)**. Because Q4-Q5 signifies CD163 interactions while Q1-Q2 signifies PECAM1 interactions (or endothelial), Q3 provides a bridge linking all three entities (i.e., Macrophage, Endothelial cells, and Vim+ mesenchymal cells) together **(Figure 4f)**. Overall, these findings suggest that the mesenchymal cell, endothelial cell, and macrophage constitute a multi-member multicellular TME niche that is uniquely recurrent across AA TNBC patients.

Using a similar approach as above, we interrogated the pattern of localization for EA-specific interaction queries (EA-Q1 to EA-Q5, **Figure 5a**) in spatial transcriptomics, paying close attention to the patterns exhibited by EA patients. In contrast to our findings with AA-interaction queries, we did not find any spatial clustering pattern for any of the EA-interaction queries, suggesting that EA interactions are uniformly scattered across the entire tumor **(Figure 5b)**. However, the total abundance of the genes constituting the EA clusters was more pronounced in EA compared to AA TNBC (**Figure 5c**). Overall, EA interactions are primarily defined by exhaustive and hypoxic environments (with colocalization of markers GZMB, CTLA4, HIF1A) in EA TNBC.

Niche-specific differential expression analysis reveals additional players associated with multicellular niches and independent mechanisms of immunosuppression

To further identify molecular factors underpinning race-specific cell-cell interaction differences, we extracted spatial transcriptomic spots that are high and low in each interaction of interest and performed niche-specific gene signature analysis (**refer Methods**). To this end, we used spots containing race-specific cell–cell interactions to identify co-expressed genes, termed extended signature genes (ESGs), which were independently associated with AA- and EA-specific interactions. We also performed a differential analysis of ESGs identified per sample between AA and EA patients to identify race-specific ESGs encoding additional niche-associated cell–cell interactions (**Fig. 6a**). As shown in **Figure 6 b**, ESGs associated with AA-specific niche defined by endothelial-macrophage-mesenchymal interactions (AA-Q3) contained additional co-expressing genes that not only marked endothelial cells and

macrophages, but also pointed to Cancer Associated Fibroblast (CAF) shown by COL1A2, DCN, THY1, FN1, CDH11, MXRA5, FSTL1 (**Figure 6b**). The existence of cancer associated fibroblasts (CAFs) was also observed in the remaining 4 AA-specific niche (AA-Q1, Q2, Q4 and Q5, **Supplementary Figures 5a-5e)**. Overall, the cell-type specific expression profiles reveal that the expression of the ESG genes are markedly higher in AA than EA in these respective cell type compartments. Notably, AA-niche ESGs are notably devoid of markers of B-cells, T-cells, neutrophils, and epithelial cells, indicating that these cell types are likely to be present at low frequencies in the AA-niches in AA patients.

In contrast, the ESGs associated with EA-specific niche included genes describing T cells and neutrophils in the background of myeloid cells, for EA-Q2 query (**Figures 6c),** as well as for the five EA-associated niches (EA-Q1 through Q5, **Supplementary Figures 6 a-e**). These genes included S100A8, S100A9, which represent key markers of neutrophils, and CD3D, IL7R, TRBC1, which are markers of T-cells. However, EA-niche ESGs did not contain any markers of CAFs or endothelial cells. Notably, the myeloid signature was more pronounced in the AA-associated niche compared to the EA-associated niche. Nevertheless, expression of ESGs associated with AA-niche and EA-niche were significantly elevated in AA and EA TNBC, respectively (**Figure 6 d, e**). Further, gene set enrichment analysis (GSEA) of the AA associated ESGs highlighted the association between EMT, development of tumor vasculature, and endothelial cells in AA TNBC, and the prevalence of altered immune signaling in EA TNBC for EA-ESGs (**Figures 6 f, g**). In summary, our results lend support to the involvement of endothelial-macrophage interaction, CAFs and tumor vasculature in AA TNBC, and reveal an immune milieu consisting of T cells and neutrophils in a myeloid background, in EA TNBC.

Drivers of cell–cell interactions and T cell states in EA-associated tumor niches

To investigate the drivers of the cell–cell interactions identified in racially distinct tumor niches, we used ESGs associated with AA- and EA-specific niches to characterize ligand-receptor interactions. This analysis revealed different ligand-receptor interactions in AA- and EA-associated niches (**Fig. 7a**). Within the AA niche, most ligand-receptor pairs were related to cell–cell communication, platelet growth factor signaling, EMT, Wnt/Notch signaling, endothelial/extracellular matrix alterations, vascular endothelial growth factor (VEGF) signaling, and integrin-mediated signaling. These ligand-receptor interactions observed in the

AA-associated niche were primarily formed between interacting perivascular-like cells, endothelial cells, CAFs, and macrophages (**Fig. 7b**). In contrast, ligand-receptor interactions in the EA-associated niche were related to immune cells. Analysis of the normalized frequencies of ligand-receptor interactions in the EA-associated niche showed strong interactions between immune cells consisting of neutrophils, T cells, and myeloid cells (**Fig. 7c**). Interactions between inflammatory cytokines (CXCL9, CXCL10, and CXCL11) and CXCR3 were also evident. These ligand-receptor interactions may have downstream consequences on the expression of target genes and activities of signaling pathways such as those depicted in **Figures 7 d-e.**

We next asked whether the T cells identified in the EA-niches were in a functional or dysfunctional state. For this, we examined the ESGs in the EA-niche for the expression of the T-cell exhaustion signature(21).As shown in **Figure 7d,** all the 5 EA-associated niches (EA-Q1 to EA Q5) showed the presence of the 10-gene T-cell exhaustion signature, including TRBC2(24), LAG3(25), HAVCR2(26), and CSF1(27).The number of spots expressing the T-cell exhaustion genes were significantly higher in the EA vs AA tumors (**Figure 7e**). Within each of the 5 EA-niches, the expression of each exhaustion gene is similarly higher in EA **Figures 7 f-k).** Overall, these findings further strengthen the existence of unique TMEs in AA and EA TNBC, with EA TME consisting of signatures of exhausted T cell milieu with neutrophils and myeloid cells.

Independent cohort validations

We next verified the AA and EA-specific niche interactions described in **Figure 4** using a cohort of independent patients that we collected, and a combination of 10X Visium ST and Nanostring GeoMx Digital Spatial Profiling (DSP) system that we profiled. In the first approach, we examined nine independent formalin fixed paraffin embedded (FFPE) TNBC samples (4 AA and 5 EA, see **Supplementary Figures 7 and 8** for H&E images and region of interest markings) using 10X ST platform. We were able to confirm in a race specific manner, the higher expression of AA-ESGs (encoding endothelial-macrophage-EMT interactions) and EA-ESGs (encoding exhausted T cells and neutrophils) in AA and EA tumors, respectively (see **Figures 8a-b** for the top 3 niches of each race). Furthermore, these validated ESGs associated with AA-specific cell-cell interactions clustered strongly within and between themselves (refer to AA patients in **Figures 8c** and **Figure 8d**). In contrast, the validated

ESGs associated with the EA clusters were more diffuse in their distribution (**Figure 8c**, EA patients) while their abundance was highly correlated between the clusters (**Figure 8d**). Nanostring GeoMx DSP analysis of 26 AA and 31 EA TNBC tumors (used for IMC in **Figure 1,** see **Supplementary Figure 3** for region of interest markings) in a tissue-microarray (TMA) setting further confirmed elevated expression of ESGs associated with AA-specific endothelial-macrophage-EMT interactions in AA (**Figure 8e**) and the exhausted T cell interactions in the EA TNBC (**Figure 8f**).

In addition, we also verified the IMC-derived interaction clusters associated with AA and EA TNBC described in **Figure 2 a, b.** For this, we carried out independent IMC analysis of 10 TNBC tissues (5 EA and 5 AA, 9 overlapping with tissues examined in **Figures 8 a, b**, refer **Supplementary Figures 7 and 8** for H&E images). Specifically, as shown in **Supplementary Figure 9a,** strong correlation was observed for the AA and EA- interacting clusters in the discovery and validation data. Furthermore, the validation data confirmed significant interaction between cluster 12 (CD4RA CD68 PLK1) and cluster 15 (AR CD31 Vimentin), reaffirming the endothelial-macrophage-mesenchymal tumor connection. (**Supplementary Figure 9b**). Along similar lines enriched interacting clusters in EA in the validation set revealed the hub cluster (Granzyme B, HIF1a, and myeloperoxidase), which formed significant interactions with exhausted T cells (CD152 CD8a) and macrophage (CD16 CD68 CD163). Importantly, we confirmed the presence of neutrophils in the EA interaction clusters using Myeloperoxidase (MPO) as the marker protein. Overall, IMC validation on an independent cohort confirmed the presence of endothelial-macrophage-mesenchymal niche in AA TNBC and an exhausted T-cell-neutrophil enriched niche in EA TNBC.

Discussion

We present the first integrated multi-omics analysis delineating the distinct TME in AA and EA patients with TNBC. TNBC tissues used in our study were obtained from 26 AA and 32 EA patients who were matched for age, grade, stage, and treatment. Our novel approach integrating IMC-based single-cell proteomics data with publicly available spatial transcriptomic data revealed racially distinct spatially resolved cell–cell interactions, which were verified in an independent cohort using IMC, spatial transcriptomics, and digital spatial profiling.

TNBC is a heterogeneous and racially disparate disease(28,29). AA women are more likely to present with TNBC and have poorer clinical outcomes than EA women(4,30,31). This

increased prevalence of aggressive disease in AA women may contribute to the high TNBC mortality in this population, which is 41% higher in AA patients than in EA patients(32). Although disparities in access to health care may contribute to racial differences in TNBC outcomes, recent studies have revealed biological differences between AA and EA patients with TNBC(8,10,19,33-37). However, most of these studies involved bulk transcriptomics or proteomics analyses, which failed to provide single-cell insights.

To address this gap, we used IMC to identify differences in cell interactions in the TME of AA and EA patients with TNBC. With endothelial cells, macrophages, and mesenchymal-like tumor cells, the TME in AA patients underscores the importance of EMT. In contrast, the TME in EA patients was enriched in immune cell interactions involving T cells, most of which were in an exhausted state. These findings reveal the existence of distinct race-driven immunosuppressive niches, one that is EMT-associated, tumor-promoting and macrophage-driven in AA patients, and another that consists of T cell exhaustion and neutrophils in EA patients. While AA-associated interactions contributed to poor OS in both populations, their effect is worse on AA patients. In contrast, EA-associated interactions were associated with poor OS only in EA but not in AA patients. These findings support the notion that race is important for understanding the significance of TME-associated changes in TNBC outcomes. This notion is consistent with a recent study demonstrating the role of race and ethnicity in immunotherapy response in patients with breast cancer(7).

Although IMC-derived single-cell interactions provide biological and clinical insights, IMC is limited by the number of markers that can be measured and, hence, the number of cell–cell interactions that can be assessed. To address this limitation, we used protein markers of race-specific cell–cell interactions as “queries” and interrogated the spatial transcriptomic data to infer the multicellular niche structure and architecture, composition, associated with EA and AA TNBC tumors. Doing so, it achieves an “expansion” of race-specific cell-cell interactions to glean insights into the additional cell-cell interactions and cell types that are present. The result were the race specific ESGs revealing additional facets of the TME in AA and EA patients that were not discernable through previous global clustering-based approaches, consistent with prior research highlighting the benefits of niche-specific analysis(38).

The AA-associated ESGs additionally revealed the involvement of CAFs in the TME of AA patients. AA-associated ESGs were enriched in EMT, vascular development-related signaling, TGF-β and Wnt/Notch1 signaling pathways, which promote EMT(39). Vascular development was encoded by genes that included VEGF and thrombospondin(40). A metastatic melanoma study identified angiogenesis-related genes as contributors to immunotherapy resistance(41); hence, studies are needed to assess whether the vascular development, CAFs(42), and endothelial cell interactions observed in AA patients confer resistance to anti-PD1 treatments. In EA patients, associated ESGs include genes indicative of exhausted T cells and neutrophils. Interestingly, in ovarian cancer, treatment with chemotherapy instills a spatially exhausted T cell environment(43). In light of this, our findings that show exhausted T cell markers(21) are enriched only in EA patients at baseline suggest that these patients are more likely to be non-responsive to chemotherapy. Together, these findings reveal the suppression of active T cell responses in EA patients with TNBC.

Although TNBC lacks the expression of the three hormonal receptors, recent studies identified subtypes of TNBC. These include basal-like immune-suppressed, immunomodulatory, luminal androgen receptor (LAR), and mesenchymal-like(44). Hammerl et al(45) found that TNBC tissues could be divided into three phenotypes: excluded, ignored, and inflamed. This finding builds upon the work of Gruosso et al(46), which showed the presence of immune-cold, immune-dessert, fully inflamed, stroma-restricted, and margin-restricted TNBC tumors. However, these studies involved single-cell RNA sequencing or relied on the localization and expression of a few markers (e.g., CD8) to describe phenotypes. Moreover, these studies did not assess the TNBC subtypes in racially diverse groups. Our data suggests that EA tumors are most like the inflamed ($P=8.9 \times 10^{-17}$) spatial phenotype, whereas AA tumors tend to belong to the Excluded (P=0.0009) and Ignored phenotype ($P=1.6 \times 10^{-7}$)(45). This further supports the notion that tumors in AA patients are likely to be immunologically cold and, hence, resistant to immunotherapy. Consistent with this, the AA-specific niche (AA-Q1 to AA-Q5) described in our study overlaps with the LAR and mesenchymal-like subtypes described by Jiang et al(44). The presence of androgen receptor in two AA-specific niches (AA-Q3 and AA-Q5) further supports their resemblance with the LAR subtype. Furthermore, the TME in AA patients also resembles the margin-restricted (showing exclusion of T cells) and immune-dessert subtypes described by Grusso et al(46). In

contrast, EA-specific niches (EA-Q1 to EA-Q5) resemble the immunomodulatory subtype described by Jiang et al(44) and the fully inflamed subtype described by Hammeri et al(45).

The implications of the finding that unique spatial niches exist within the TME in AA and EA patients with TNBC are multifold. For the first time, our study revealed architectural differences in the TME, which may influence tumor progression or response to treatment. Our findings provide preliminary evidence to support the development of therapeutic strategies targeting the endothelial-macrophage-EMT axis in AA patients with TNBC. Moreover, the identification of ligand-receptor interactions that encode this axis, including integrin-TGF-β(47), Wnt/Notch1, and VEGF signaling, provides several novel actionable targets(47). For example, clinical-grade Wnt/Notch1 inhibitors affect endothelial cells(48) and EMT(49,50), and could benefit AA patients with TNBC. Furthermore, therapeutic regimens targeting macrophages have been examined in preclinical models of TNBC(51). The identification of a neutrophil-rich TME co-existing with exhausted T cells in EA patients is intriguing and was confirmed using myeloperoxidase as a neutrophil marker in our validation cohort. Recent studies have shown that neutrophils can promote tumor progression(52), and may do so in a macrophage-independent manner(53) by exerting an immunosuppressive function(53,54). Whether such neutrophil-associated immunosuppressive and tumor-promoting features are present in the TME of EA patients with TNBC remains to be determined. These findings, combined with those of previous studies(53,54), suggest that EA patients with TNBC have inflamed tumors that may have evolved distinctive immunosuppressive mechanisms.

Our findings revealing differences in immune landscape between races argues the need to study disparities in TNBC using immunocompetent mouse models. In a recent study examining the TME in different genetically engineered mouse models, two breast cancer prototypes were identified(53). The first prototype involved local and systemic accumulation of neutrophils that overexpressed multiple immunosuppressive molecules(53), similar to those observed in EA patients with TNBC. The second prototype involved local enrichment of macrophages and scarcity of neutrophils(53), similar to what we observed in AA patients. Additional studies comparing the TME in these mouse models with those identified here may reveal new immunocompetent mouse models that could be used to study racial differences in the TME.

One of the limitations of this study is the small IMC panel size used to discover the AA and EA-associated niches. To overcome this, we expanded these niche signatures in an independent 10X spatial transcriptomics analysis. In addition, we validated most of the interactions in an independent cohort using IMC, spatial transcriptomic analysis, and digital spatial profiling. IMC studies for the discovery dataset were conducted on TMA, in which each tumor was represented by a 0.3-mm section. In contrast, the validation dataset included entire tumor sections. This difference could have affected the cluster composition observed for IMC-derived proteins in the discovery and validation phases. Nevertheless, six of the nine AA-specific interactions and all EA-specific interactions were confirmed in the validation dataset.

In conclusion, our results show racial differences in spatial interactions and associated genes in the TME of patients with TNBC. The TME of AA patients is populated by endothelial cells, CAFs, and macrophages, which promote EMT and mesenchymal cell interactions (**Fig. 9a**). These interactions are associated with poor patient survival in both race groups. Furthermore, the tumors of AA patients are immunologically cold and, hence, potentially non-responsive to immunotherapy. In contrast, the TME of EA patients is characterized by exhausted T cells and neutrophils (**Fig. 9b**), which are prognostic only in EA patients. The finding of a neutrophil-rich TME in EA patients is intriguing, and whether neutrophils exert a tumor-promoting function in EA patients remains to be determined. These findings highlight the importance of taking race into consideration when studying alterations in the TME in patients with TNBC.

Methods:

TNBC patient tissue microarrays: De-identified breast cancer patient samples with at least 10-year follow-up were obtained from Baylor Scott and White hospital, Temple, Texas in the form of formalin fixed paraffin embedded (FFPE) tissue. The study was performed under IRB (Institutional Review Board) protocols 020-393 and 130559 of Baylor Scott and White Hospital, and H-28445 of Baylor College of Medicine. Tissue cores from 57 tumors (26 AA tumors and 31 EA tumors) spread across 8 tissue microarrays (TMAs) were used for this study. In addition to the race, clinical information such as receptor status, tumor grade, chemotherapy status, presence of metastases, time to clinical follow up, as well as detailed histopathology of the tumors were available to us (refer **Supplementary Table 1** for clinical information and **Supplementary Figure 1** for H&E images).

Assessment of ancestry of the self-reported TNBC patients: To categorize the TNBC samples according to the delineated African Ancestry, we utilized a gene signature derived from a publication where gene signatures associated African American ancestry were described(8). This signature contained 613 genes. We used the Nanostring GeoMx Digital Spatial Profiling data conducted on the 57 TNBC tissues in this study. A total of 279 genes overlapped with the 613 genes described in the above-mentioned publication(8). A PCA plot was used to examine the separation of TNBC samples in our data set into AA and EA groups based on the 279 gene signature.

Region of Interest (ROI) selection and segmentation: For each patient core analyzed using imaging mass cytometry, 10X spatial transcriptomics or Nanostring Geo Mx Digital

Spatial Profiling (DSP), ROIs were selected based on two parameters- tumor location (center vs periphery) and amount of immune infiltration observed in the H&E section (immune rich vs immune poor). This was done by the breast pathologist Dr. Asirvatham. A total of 98 ROIs were selected across the 8 TMAs (and 56 TNBC tissues), each belonging to one of four categories- Tumor Center Immune Rich (TCIR), Tumor Center Immune Poor (TCIP), Tumor Periphery Immune Rich (TPIR), and Tumor Periphery Immune Poor (TPIP). **Supplementary Table 1** describes the numbers of ROIs belonging to each of these four groups across all the specimens analyzed.

Imaging mass cytometry (Mass CyTOF), Discovery data set: For the discovery data, a panel of 26 antibodies were selected based on the markers expressed in common immune cell populations- including T cells, macrophages, dendritic cells, and natural killer (NK) cells- as well as those expressed on tumor cells- including markers for hypoxia, angiogenesis, proliferation, epithelial mesenchymal transition (EMT), and immunosuppression.

Mass CyTOF antibody conjugation and staining were done in collaboration with the Aneja lab at Georgia State University. Lanthanide metal-labeled and Iridium intercalator antibodies were purchased from Fluidigm. Unconjugated antibodies were conjugated using the Max Par X8 labeling kits from Fluidigm. The concentrations of the conjugated antibodies were assessed using the NanoDrop system and the final concentration was adjusted to 500 µg/ml. The conjugated antibodies were stored in an antibody stabilizer at 4 degrees Celsius. Descriptions of antibodies and isotope tags are described in **Supplementary Table 2.**

Staining of the tissues was performed as previously described(11,22,55). Briefly, slides were de-paraffinized in xylene and rehydrated in alcohol. Antigen retrieval was performed with pre-heated Tris-EDTA buffer (pH 9) at 95 degrees Celsius in a de-cloaking chamber (Biocare Medical). The slides were cooled and blocked with 3% BSA (in PBS) for 1 hour. Slides were then incubated with metal-tagged antibodies (1:50) overnight at 4 degrees Celsius. Counterstaining of the nuclei was performed with Iridium intercalator (1:200 dilution).

Image acquisition by Hyperion and data analysis- Image acquisition was done in collaboration with the Flow Cytometry core at Baylor College of Medicine. Tissue analysis was performed using a Helios time of flight mass cytometer coupled to the Hyperion Imaging System (Fluidigm). Prior to acquisition, the imaging system was auto tuned using a 3-element tuning slide (Fluidigm) according to manufacturer's instructions. ROIs for imaging were

selected as described in **Supplementary Figure 3**. Following the flushing of the ablation chamber with Helium, the tissue sections were ablated in a spot-by-spot fashion by a UV laser spot at 200Hz frequency and 1µm resolution. The results were stored in Fluidigm's MCD format and exported as 16-bit OME tiff format for downstream quantification.

Imaging mass cytometry (Mass CyTOF), Validation data set: Here 5 EA and 4 AA TNBC FFPE tumors were analyzed (See Supplementary Figure 7 and 8 for H&E images and ROI markings). This analysis was done by the Immune Monitoring Core at the Houston Methodist Research Institute. Sample preparation commenced with the staining of tissues using pathologist-verified, metal-tagged antibodies, which were optimized for the CyTOF® imaging system (56). A total of 27 metal-tagged antibodies were used in this analysis (refer **Supplementary Table 3**) These antibodies allowed comprehensive analysis of immune, stromal, and tumor cell heterogeneity, in addition to various cell subsets and functional phenotypes within the tumor microenvironment (TME). All the antibodies were prepared according to the manufacturer's protocols provided by Standard BioTools, measured for absorbance, and stored in Candor PBS Antibody Stabilization solution (Candor Bioscience) at 4°C. FFPE tissue sections were subjected to baking, dewaxing in xylene, rehydration through graded alcohols, and heat-induced epitope retrieval in an EZ-Retriever System (BioGenex) at 95°C using a Tris-Tween20 buffer at pH 9 for 20 minutes. After blocking with 3% BSA in TBS, the sections were incubated overnight with an antibody master mix, followed by washing, and staining for nuclear identification using Cell-ID Intercalator (Standard BioTools). Subsequently, the slides were washed, air-dried, and stored for ablation. The sections were then ablated using the Hyperion system (Standard BioTools) for data acquisition. Data acquisition followed the method described earlier for the Discovery data set.

10X Visium Spatial Transcriptomic Profiling: Here 5 EA and 4 AA TNBC FFPE tumors were analyzed (See Supplementary Figure 7 and 8 for H&E images and ROI markings). This analysis was done by the Immune Monitoring Core at the Houston Methodist Research Institute.

cDNA libraries were prepared following the guidelines outlined in the Visium CytAssist Spatial Gene Expression for FFPE User Guide. FFPE tissue sections of 5 µm thickness were mounted on Superfrost™ Plus Microscope Slides (Fisherbrand™) and subjected to H&E staining after deparaffinization. Following imaging, the cover slips were removed from the

sections, and the sections were processed for hematoxylin de-staining and de-crosslinking. The glass slide bearing the tissue section underwent processing using the Visium CytAssist instrument (10x Genomics) to facilitate the transfer of analytes to a Visium CytAssist Spatial Gene Expression slide, which features a 0.42 cm^2 capture area with 4,992 spatial barcodes. Subsequent to this, the probe extension and library construction were carried out according to the standard Visium for FFPE workflow, following the manufacturer's protocol, but outside the instrument. The libraries were then sequenced using paired-end dual-indexing (28 cycles for Read 1, 10 cycles for i7 index, 10 cycles for i5 index, and 50 cycles for Read 2) on an Illumina NovaSeq X platform, achieving an average of 30,000 reads per spot. The resulting FASTQ files, together with the H&E images, were processed using Space Ranger version 2.1.0 (10x Genomics), using the GRCh38-2020-A reference genome.

Gene expression analysis using GeoMX DSP system (NanoString): This was used to profile 57 FFPE TNBC tissues (26 AA and 31 EA) which were also used to generate the "Discovery data". **Supplementary Figures 1 and 2** describe the H&E images, selected ROIs and the immunofluorescence staining with PanCK and CD45 for each tissue. The latter was used as the morphology markers to segment the tissues into epithelial and immune cell compartments as described below. **Supplementary Table 1** describes the various categories of ROIs selected by the breast pathologist Dr. Asirvatham for the Nanostring analysis.

We queried the transcriptome in these samples using the GeoMX Human Whole Transcriptome Atlas (WTA), which measures ~18,000 protein-coding genes. GeoMX DSP analysis was performed as previously described(57). In brief, the 8 TMAs were stained with mRNA hybridization probes attached to UV-photo cleavable indexing oligonucleotides. The slides were stained with two morphology markers- Pan Cytokeratin (PanCK, to identify tumor cells), and CD45 (to identify immune cells) in addition to a nuclear stain (DAPI). Specific regions of interest (ROIs) were selected by the breast pathologist Dr. Asirvatham under the guidance of the morphology markers. Each ROI was thus divided into three segments based on the positive/negative staining of the morphology markers- Tumor segment (PanCK +ve CD45 -ve), Immune segment (PanCK -ve CD45 +ve), and Stroma segment (PanCK -ve CD45-ve). 293 segments (98 tumor segments, 98 immune segments, and 97 stromal segments) were generated in this fashion, from 57 tumors (26 AA and 31 EA tumors, respectively). Since the stromal segment was defined using negative selection, it contained a mixed population of cells, and hence was not considered for downstream analysis. The slides

were exposed to UV light, thereby releasing the indexing oligonucleotides. The oligonucleotides from each ROI were then collected into specific wells of a microplate and counted using the nCounter system.

Multiplex Immunofluorescence Analysis: This was used to validate the co-localization of endothelial cells and macrophages in AA TNBC. A total of 27 AA and 44 EA TNBC samples (used to generate “Discovery phase” IMC data and “Validation phase” Nanostring GeoMx DSP data) were used. These tissues were distributed across 8 TMAs. Briefly, FFPE patient TMA slides were subjected to baking at 60°C for two hours and washed with xylene to remove excess paraffin. Subsequently, the slides underwent rehydration by incubating in a series of ethanol solutions at various concentrations (100%, 95%, 70%, 50%, 30%, 0%). After rinsing with PBS, the slides were immersed in 1x Target Retrieval Solution, pH 9 (Dako, S2367) at 115°C for 15 minutes in a pressure cooker for antigen retrieval. Following antigen retrieval, the tissue sections were permeabilized and blocked using 10% normal donkey serum in PBS-GT (2% fish gelatin, 0.1% Triton-X100 in PBS) for 1 hour at room temperature. Primary antibodies (goat anti-CD31, R&D, AF3628; rabbit anti-CD163, Abcam, ab182422) were diluted to 1:100 in PBS-GT and applied to the tissue sections, which were then incubated in a humid chamber at 4°C overnight. The following day, the slides were washed with PBS three times and incubated with secondary antibodies (Alexa Fluor 488-conjugated Donkey anti-Rabbit-IgG, Jackson ImmunoResearch, 711-545-152; Alexa Fluor 555-conjugated Donkey anti-Goat IgG, Thermo Scientific, A-21432) diluted to 1:500 in PBS-GT for 2 hours at room temperature.

After two washes with PBS, the slides were stained with Hoechst (20 µg/mL) in PBS for 5 minutes at room temperature, followed by two additional PBS washes. ProlongTM Gold antifade mountant was applied, and a cover slip was carefully placed. The slides were left to cure in the dark at room temperature overnight. Finally, images were acquired using a Zeiss LSM 780 confocal microscope, and the fluorescent images were processed using Zen software (Zeiss).

Immunofluorescence quantification and co-localization: Multi-channel TIFF images were first separated into individual color channels: Red – representing CD31, Green – CD163, and Blue – Hoechst. We conducted pixel-level co-localization between CD31 and CD163 channel images in ImageJ(58). First, individual channel was smoothed using Gaussian Blur at default

setting. Then signal intensities per channel were auto-thresholded to preserve the upper 5-10% of the image histogram. Signal intensities were segmented by using the Find Maxima function in ImageJ, and locations of segmented pixels of CD31 and CD163 intensities were recorded into Region of Interest (ROI) Manager. Next, we computed co-localization frequencies between CD31 and CD163 segmented pixels by the following procedure. The number of pairs of pixels (where one from CD31 pixel list and one from CD163) that are separated by <40-pixel Euclidean distance apart were quantified per image. This was repeated for all of the 103 immunofluorescent images of EA and AA TNBC tumors that were acquired.

Imaging Mass Cytometry Analysis: Our IMC data, stored in MCD files, were processed using the imctools to generate multichannel TIFF images for protein markers. For cell segmentation, we used tools such as HistoCAT(55), CellProfiler(59), and CellPose(60). We started by applying a Gaussian Blur in ImageJ(61) to smooth the Ir191 and Ir193 DNA channel images, enhancing signal continuity within cell nuclei. Deep learning tool CellPose was applied to perform cell segmentation on DNA channels. For images with unsatisfactory segmentation, ilastik(62) was used for semi-supervised pixel classification; a few cells were manually segmented to train the classifier for segmenting the entire image. HistoCAT uses cell masks to compile a cell-by-protein matrix, summarizing pixel intensities for each cell and protein marker. Cell position was determined by the centroid position of cell masks obtained from DNA channel segmentation.

Clustering and differential protein analysis: We applied log-transformation and z-scoring, first across all cells and then across all protein markers. Next, we used K-means clustering with a high number of random starts (nstart=100,000) for reliable centroid generation. This was used because K-means promoted greater tolerance for noise and variation in data processing, was less sensitive and more robust to outlier expression, and avoided the generation of ROI (region of interest)-specific clusters. We determined 20 K-means clustering showing robust cluster-specific expression. We next determined differentially expressed proteins across clusters by performing all one-vs-one comparisons. Statistical significance was reached if P<=0.05. We deemed a protein to be differentially expressed for a cluster C_k if the protein is significant in 17 out of 19 one-vs-one comparisons between C_k and C_n where

$n = 1 \ldots N$ excluding k. After that, each cluster was named by the list of differential markers expressed in the cluster.

Spatial interaction analysis: We constructed a Delauney spatial graph for single cells in each ROI and used it with cluster labels for spatial proximity analysis in the Giotto pipeline(23). Giotto's Cell Proximity Enrichment tool, coupled with 1000 simulations, identified cell-type interaction enrichment or depletion compared to random simulations with shuffled labels. This analysis generated a z-score for each cell-type pair, indicating enrichment (positive z-score) or depletion (negative z-score) of interactions per ROI. Giotto is a validated pipeline for inferring spatial interactions from Multiplexed Ion Beam Imaging (MIBI), IMC and compatible data. We have previously illustrated an example of analyzing a MIBI TNBC dataset(12) in the Giotto paper(23). Next, we conducted group-wise comparison to extract racial group-specific cell-cell interactions (CCIs). For this analysis, race groups were self-reported African American (AA) or European American (EA) TNBC patients. We used a mixed linear statistical model to find racial group specific CCIs. The linear mixed model accounts for variations derived from patients while allowing each patient to have repeated observations in the form of multiple ROIs per patient(63). This approach yields more accurate P-values than standard tests that falsely assume independence between ROIs. We used the lme4 R package for this analysis(63). In this linear model, we write the design function as "score ~ 1 + race + (1|patient)", where score is the interaction score of a particular cell-type pair in ROI, race is either AA or EA, and patient is the patient identifier of ROI. Multiple ROIs are associated with each patient (i.e. making patient the random effect), where as race is the main (or fixed) effect. The alternative null model has the design function "score ~ 1 + (1|patient)" which excludes the race effect. ANOVA testing was performed between these two models to select only the interaction pairs whose scores are most affected by race.

Pre-processing of 10X Visium Spatial Transcriptomics data: We have used Space Ranger to align the reads to human hg38. Following this, we loaded the resulting hdf5 file into Giotto for preliminary analyses. This includes dimensionality reduction, UMAP (Uniform Manifold Approximation and Projection), and KNN-based Leiden clustering. We preprocessed each spatial transcriptomics sample in Bassiouni et al(20) and in our own TNBC cohort (used for validation, refer **Supplementary Figures 7 and 8**) to generate normalized gene expression matrices per ST sample with Giotto(23).

Targeted analysis of spatial transcriptomic datasets

Query co-localization of cell-cell interaction pairs: Each cell-cell interaction (CCI) was converted from protein names to gene names. The list of gene names in CCI forms a race-specific-query, such as AA-Q1 (one of AA-Q1 through AA-Q5). For a given query, we computed a query co-localization score (QCS) per spot x per ST sample A by summing the scaled and log-normalized expression of query genes in the spot:

$$QCS_{x,A} = \frac{1}{|Q|}\sum_{q\in Q} Expr_{x,A}(q)$$

. $QCS_{x,A}$ was deemed significant if it exceeded (P<0.05) the summed score of the randomly shuffled case, whereby the expression of each query gene was randomly distributed among all the in-tissue Visium spots in the sample, and thus destroys the dependence among query genes. A per-sample QCS score is next quantified as $QCS_A = |QCS_{x,A} > QCS_{shuffled,A}|$ which is equal to the total number of Visium spots with significant QCS scores. This QCS_A was compared between AA and EA racial groups.

Spatial Clustering and Cross-Query Correlation: In addition to QCS, we also quantified a Spatial Clustering score, and conducted cross-query correlation analysis. In the former, to compute Spatial Clustering, we adopted the silhouette coefficient metric previously described in Zhu et al(64) that measures the spatial coherence of gene expression pattern. Here, the pattern was represented by spots showing significant QCS (given the label 1) against the background of remaining spots (given the label 0). Spatial Clustering (δ), computed by silhouette coefficients, was next assessed individually for each query (AA-Q1 through AA-Q5):

$$\delta = 1/|L_1|\sum_{s_i\in L_1}(m_i - n_i)/\max(m_i, n_i)$$

This silhouette coefficient(65) assesses the spatial distance associated with two sets, L_1 (spots given label 1) and L_0 (spots given label 0). For a given spot s_i in set L_1, m_i is defined as the average distance between s_i and any spot in L_0, and n_i is defined as the average distance between s_i and any spot in L_1. For distance, we used rank-normalized, exponentially transformed distance which prioritizes local physical distance between two spots. The distance between spots s_i and s_j is defined as $r(s_i, s_j) = 1 - q^{rank_d(s_i,s_j)-1}$ where $rank_d(s_i, s_j)$

is the mutual rank(66) of s_i and s_j in vectors of Euclidean distances $\{Euc(s_i,*)\}$ and $\{Euc(s_j,*)\}$. q is a rank-weighting constant, set at 0.95.

Spatial clustering was compared between AA and EA patient groups. For the analysis of cross-query correlation, for each query, we made a vector of QCS scores for all spots per query, $Q_1 = [QCS_1, QCS_2, QCS_3, \dots] for\ spots\ 1,2,3$. We next computed pairwise Pearson correlation coefficients between vectors Q1, Q2, … Q5. The purpose of this analysis was to assess the spatial correlations between the 5 AA-specific queries, as it revealed unique differences between AA and EA patient samples.

Derivation of Extended Signature Genes of AA- and EA-associated cell-cell interaction niches: To derive extended signature genes (ESGs) of AA- and EA- race-associated and interaction-targeted niches, for each interaction query AA-Qi, we stratified the Visium spots per ST sample into interact-high and interact-low groups based on QCS score of that query per spot. The interaction-high group consists of Visium spots in the top 10% of spots with the highest QCS. The interaction-low group consists of spots in the bottom 10% with lowest QCS. These high- and low-groups were defined per ST sample. Per-sample differential gene expression analysis was next performed between the interact-high and interact-low spots within the ST sample, forming Extended Signature Genes of the query. Recurrent AA-associated ESGs were next derived by requiring ESGs to be present in at least 50% of AA patient samples and the percentage of AA patients with ESGs as signature must be higher than the percentage in EA patients.

Cell-type specific expression of extended signature genes of AA- and EA-associated niches: We integrated single-cell RNA sequencing data of breast cancer atlas(67) with spatial transcriptomics to derive cell-type specific gene expression profiles in AA- and EA-interaction-targeted niches. The cell-type specific expression of gene g in cell type c, in a niche N that is defined by EA or AA-query, shortened as $Expr_{g,c,N}$, was derived from:

$$Expr_{g,c,N} = AvgExpr_{g,c} \times V_{g,N}$$

where $AvgExpr_{g,c}$ is the average expression of g in cell type c in the scRNAseq data, and $V_{g,N}$ is the ST average-subtracted expression of g in spot s in niche N in 10X Visium sample.

Niche is made up of spots in the interaction-high group defined earlier. Briefly, $V_{g,N}$ is defined as:

$$V_{g,N} = \sum_{s \in N} VExpr_{g,s} /|N| - \sum_{s \in S} VExpr_{g,s} /|S|$$

Here, $VExpr_{g,s}$ is the Visium expression of g in spot s. $V_{g,N}$ is the result of subtracting the background expression (summed over all spots S in the ST sample) from the niche-expression (summed over only spots in niche N). For cell types, we iterated over all Cancer Associated Fibroblasts (CAF), endothelial, Peri-Vascular Like (PVL) cell types, as well as all subsets of immune cells in Wu et al(67), and Wigerblad et al(68), and Alvarez-Breckenridge et al(69), which provided the neutrophil subsets.

***Ligand-receptor pair analysis*:** We download all human ligand-receptor pairs from CellPhoneDB database(70). We next computed enrichment of ligand-receptor pairs within ESGs of AA- and EA-associated cell niches by identifying pairs where both the ligand and receptor genes were present within the ESGs. This information was mapped to different niches AA-Q1 through Q5 for AA-samples, and EA-Q1 through Q5 for EA-samples. Ligand-receptor pair results were next summarized into a cell-type interaction network as shown in **Figure 7a** by the following procedure. The cell-interaction score CI between two cell types c_1 and c_2 is given by:

$$CI(c_1, c_2) = \sum_{LR} \frac{I(LR, c_1, c_2)}{N(LR)}$$

Where LR is a ligand-receptor pair – we iterated over all ligand-receptors that are enriched among the ESGs. $I(LR, c_1, c_2)$ is an indicator function that is 1 if LR is expressed in c_1 and c_2, and 0 otherwise. $N(LR)$ is the total number of cell type pairs in which ligand-receptor is expressed, defined as:

$$N(LR) = \sum_{c_1} \sum_{c_2} I(LR, c_1, c_2)$$

Overlap analysis with existing TNBC subtype gene signatures: Inflamed, Excluded, and Ignored signatures of TNBC were downloaded from Hammerl et al(45) supplementary material. To compare these existing phenotype signatures with our EA-niche specific ESG

signatures, we computed the number of overlapping genes and its statistical significance using the hypergeometric distribution test.

Gene-Set Enrichment Analysis of Niche-Genes: We used the g:Profiler(71) web server to compute gene-set enrichment statistics against the gene-set database(72) compiled by the Bader lab. This database has compiled all current GO Biological Process, and all pathway gene sets and continuously updated. We downloaded Human_GOBP_AllPathways_no_GO_iea.gmt and uploaded it directly to g:Profiler for enrichment analysis of AA- and EA-associated ESGs. Default enrichment settings were used for the analysis.

NanoString GeoMx DSP Validation Analysis: We interrogated the CD45+ immune and PanCK epithelial compartment gene expression matrices to compare the expression level of race-specific niche genes in TNBC AA and EA patients (see **Supplementary Figure 1 and 2**). For each patient and each ROI, a summary score was computed for each niche-associated ESG gene set based on summing log-normalized gene expression for all genes in each niche-gene set. The score was compared between AA and EA samples, and t-test statistics (testing EA>AA or AA>EA depending on the niche) were computed.

10X Visium Spatial Transcriptomics Validation Analysis: We validated the niche specific ESGs by asking whether these ESGs are higher in AA patients or higher in EA patients in our validation cohort (9 TNBC, 4 AA and 5 EA, see **Supplementary Figures 7 and 8**). Specifically, for each cell-cell interaction query AA-Q1…Q5 and EA-Q1…Q5, we extracted interact-high and interact-low spots per ST sample in our validation cohort. Next, for the corresponding niche specific ESGs, for each gene g we computed

$$\log_2\left(\frac{avgExpr_{g,high}}{avgExpr_{g,low}}\right)$$

using the interact-high and -low spots that were defined and compared this log-fold change in AA samples and EA samples in validation cohort.

Kaplan Meier (KM) Survival Analysis: Survival analysis was performed using the survival R package and visualized using the ggplot2 package autoplot function. Patients were stratified into interaction-high and interaction-low groups based on whether the interaction score of the

patient is higher or lower than the average score. All 9 AA-specific interactions and 6 EA-specific interactions (as shown in **Figures 2b and 2c**) were tested. Summed interaction score (of 9 or 6 interactions) was also tested.

Visualization: Spatial gene expression profiles were plotted using scatter function in Python matplotlib library. Correlation maps were plotted with seaborn library, heatmap function, and plasma colormap. Gene expression heatmaps were plotted with seaborn library clustermap function with colormap set to Spectral. Protein co-localization figures (IMC) were generated for defined cell clusters using the scatter function. Lastly, we used lineplot function (seaborn package) to compare gene expression between EA and AA groups.

Contributions:

AB, JK, WL, MC, UR, YX: collected experimental data. AB wrote the manuscript. BP, JZ: pre-processed IMC, 10X ST and Nanostrong GeoMx DSP data, DJ, AR and JRA created the TMAs, JRA selected the patient samples and marked the ROIs for IMC, 10X ST and Nanostring GeoMx DSP analysis, NM, NP compiled the clinical data for the TNBC samples, JRA, S-hC, XL provided reagents and resources, YL, EC and XHFZ critically reviewed the data and provided scientific inputs, QZ developed and performed all the novel integrative analysis, conceptualized the study, wrote the manuscript, and ASK conceptualized, directed, and funded the study, and wrote the manuscript. All authors discussed the results and commented on the manuscript.

Acknowledgements: This study was partly funded by RO1CA227904 (ASK and XZHF) and 1 P20 CA284971-01 (ASK), and a Cancer Prevention Research Institute of Texas (CPRIT) grant RR220035 (QZ). Dr. Zhu is a CPRIT scholar in Cancer Research. Dr. Zhang is a McNair scholar. Dr. Sreekumar is supported by Charles C Bell Jr-endowed professorship. The authors acknowledge Dr. Liang Zhang and Kathy Ton from the NanoString TAP team for performing the GeoMX Human Whole Transcriptome Atlas assay and assisting with the initial data analysis. IMC data collection was supported by the Cytometry and Cell Sorting Core at Baylor College of Medicine with funding from the CPRIT Core Facility Support Award (CPRIT-RP180672), the NIH (CA125123 and RR024574) and the assistance of Joel M. Sederstrom.

Confocal microscopy was supported by the Integrated Microscopy Core at BCM with funding from NIH (DK56338, and CA125123), CPRIT (RP150578), the Dan L. Duncan Comprehensive Cancer Center, and the John S. Dunn Gulf Coast Consortium for Chemical Genomics.The authors acknowledge Callee Arnold, histo-technologist at BSWH for technical expertise. JRA served as an advisor for Roche. Xiaoxian Li has served as an advisor for Astra Zeneca, Roche, Eli Lilly, and Onviv, and Champions Oncology has funded research in Xiaoxian Li's lab. Dr. Sreekumar reports grants from the Agilent Foundation, non-financial support from Sri Sathya Sai Institute for Higher Learning, India, and personal fees from Karkinos Health Care Pvt. Ltd., India, outside the submitted work. Editorial assistance was provided by Christos Evangelou, PhD.

References

1. Singh M, Ding Y, Zhang LY, Song D, Gong Y, Adams S, *et al.* Distinct breast cancer subtypes in women with early-onset disease across races. Am J Cancer Res **2014**;4(4):337-52.
2. Amirikia KC, Mills P, Bush J, Newman LA. Higher population-based incidence rates of triple-negative breast cancer among young African-American women : Implications for breast cancer screening recommendations. Cancer **2011**;117(12):2747-53 doi 10.1002/cncr.25862.
3. Chen L, Li CI. Racial disparities in breast cancer diagnosis and treatment by hormone receptor and HER2 status. Cancer Epidemiol Biomarkers Prev **2015**;24(11):1666-72 doi 10.1158/1055-9965.EPI-15-0293.
4. Akinyemiju T, Moore JX, Altekruse SF. Breast cancer survival in African-American women by hormone receptor subtypes. Breast Cancer Res Treat **2015**;153(1):211-8 doi 10.1007/s10549-015-3528-7.
5. Ademuyiwa FO, Groman A, O'Connor T, Ambrosone C, Watroba N, Edge SB. Impact of body mass index on clinical outcomes in triple-negative breast cancer. Cancer **2011**;117(18):4132-40 doi 10.1002/cncr.26019.
6. Keenan T, Moy B, Mroz EA, Ross K, Niemierko A, Rocco JW, *et al.* Comparison of the Genomic Landscape Between Primary Breast Cancer in African American Versus White Women and the Association of Racial Differences With Tumor Recurrence. J Clin Oncol **2015**;33(31):3621-7 doi 10.1200/JCO.2015.62.2126.
7. Newman LA, Chen Y, Martini R, Demaria S, Formenti S, Elemento O, Davis MB. Tumor-Associated Lymphocytes and Breast Cancer Survival in Black and White Women. JAMA Surgery **2024** doi 10.1001/jamasurg.2023.8024.
8. Martini R, Delpe P, Chu TR, Arora K, Lord B, Verma A, *et al.* African Ancestry-Associated Gene Expression Profiles in Triple-Negative Breast Cancer Underlie Altered Tumor Biology and Clinical Outcome in Women of African Descent. Cancer Discov **2022**;12(11):2530-51 doi 10.1158/2159-8290.CD-22-0138.
9. Charan M, Verma AK, Hussain S, Misri S, Mishra S, Majumder S, *et al.* Molecular and Cellular Factors Associated with Racial Disparity in Breast Cancer. Int J Mol Sci **2020**;21(16) doi 10.3390/ijms21165936.
10. Martin DN, Boersma BJ, Yi M, Reimers M, Howe TM, Yfantis HG, *et al.* Differences in the tumor microenvironment between African-American and European-American breast cancer patients. PLoS One **2009**;4(2):e4531 doi 10.1371/journal.pone.0004531.
11. Ali HR, Jackson HW, Zanotelli VRT, Danenberg E, Fischer JR, Bardwell H, *et al.* Imaging mass cytometry and multiplatform genomics define the phenogenomic landscape of breast cancer. Nat Cancer **2020**;1(2):163-75 doi 10.1038/s43018-020-0026-6.
12. Keren L, Bosse M, Marquez D, Angoshtari R, Jain S, Varma S, *et al.* A Structured Tumor-Immune Microenvironment in Triple Negative Breast Cancer Revealed by Multiplexed Ion Beam Imaging. Cell **2018**;174(6):1373-87 e19 doi 10.1016/j.cell.2018.08.039.
13. Stahl PL, Salmen F, Vickovic S, Lundmark A, Navarro JF, Magnusson J, *et al.* Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. Science **2016**;353(6294):78-82 doi 10.1126/science.aaf2403.
14. Risom T, Glass DR, Averbukh I, Liu CC, Baranski A, Kagel A, *et al.* Transition to invasive breast cancer is associated with progressive changes in the structure and composition of tumor stroma. Cell **2022**;185(2):299-310 e18 doi 10.1016/j.cell.2021.12.023.
15. Uttam S, Stern AM, Sevinsky CJ, Furman S, Pullara F, Spagnolo D, *et al.* Spatial domain analysis predicts risk of colorectal cancer recurrence and infers associated tumor microenvironment networks. Nat Commun **2020**;11(1):3515 doi 10.1038/s41467-020-17083-x.

16. Sammut S-J, Crispin-Ortuzar M, Chin S-F, Provenzano E, Bardwell HA, Ma W, *et al.* Multi-omic machine learning predictor of breast cancer therapy response. Nature **2022**;601(7894):623-9 doi 10.1038/s41586-021-04278-5.
17. Krassowski M, Das V, Sahu SK, Misra BB. State of the Field in Multi-Omics Research: From Computational Needs to Data Mining and Sharing. Front Genet **2020**;11:610798 doi 10.3389/fgene.2020.610798.
18. Zhang B, Wang J, Wang X, Zhu J, Liu Q, Shi Z, *et al.* Proteogenomic characterization of human colon and rectal cancer. Nature **2014**;513(7518):382-7 doi 10.1038/nature13438.
19. Omilian AR, Sheng H, Hong CC, Bandera EV, Khoury T, Ambrosone CB, Yao S. Multiplexed digital spatial profiling of invasive breast tumors from Black and White women. Mol Oncol **2022**;16(1):54-68 doi 10.1002/1878-0261.13017.
20. Bassiouni R, Idowu MO, Gibbs LD, Robila V, Grizzard PJ, Webb MG, *et al.* Spatial Transcriptomic Analysis of a Diverse Patient Cohort Reveals a Conserved Architecture in Triple-Negative Breast Cancer. Cancer Res **2023**;83(1):34-48 doi 10.1158/0008-5472.CAN-22-2682.
21. Tietscher S, Wagner J, Anzeneder T, Langwieder C, Rees M, Sobottka B, *et al.* A comprehensive single-cell map of T cell exhaustion-associated immune environments in human breast cancer. Nat Commun **2023**;14(1):98 doi 10.1038/s41467-022-35238-w.
22. Giesen C, Wang HA, Schapiro D, Zivanovic N, Jacobs A, Hattendorf B, *et al.* Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. Nat Methods **2014**;11(4):417-22 doi 10.1038/nmeth.2869.
23. Dries R, Zhu Q, Dong R, Eng CL, Li H, Liu K, *et al.* Giotto: a toolbox for integrative analysis and visualization of spatial expression data. Genome Biol **2021**;22(1):78 doi 10.1186/s13059-021-02286-2.
24. Ferrari M, Righi M, Baldan V, Wawrzyniecka P, Bulek A, Kinna A, *et al.* Structure-guided engineering of immunotherapies targeting TRBC1 and TRBC2 in T cell malignancies. Nat Commun **2024**;15(1):1583 doi 10.1038/s41467-024-45854-3.
25. Grebinoski S, Zhang Q, Cillo AR, Manne S, Xiao H, Brunazzi EA, *et al.* Autoreactive CD8+ T cells are restrained by an exhaustion-like program that is maintained by LAG3. Nature Immunology **2022**;23(6):868-77 doi 10.1038/s41590-022-01210-5.
26. Anderson AC, Joller N, Kuchroo VK. Lag-3, Tim-3, and TIGIT: Co-inhibitory Receptors with Specialized Functions in Immune Regulation. Immunity **2016**;44(5):989-1004 doi 10.1016/j.immuni.2016.05.001.
27. Kersten K, Hu KH, Combes AJ, Samad B, Harwin T, Ray A, *et al.* Spatiotemporal co-dependency between macrophages and exhausted CD8(+) T cells in cancer. Cancer Cell **2022**;40(6):624-38.e9 doi 10.1016/j.ccell.2022.05.004.
28. Siddharth S, Sharma D. Racial Disparity and Triple-Negative Breast Cancer in African-American Women: A Multifaceted Affair between Obesity, Biology, and Socioeconomic Determinants. Cancers (Basel) **2018**;10(12) doi 10.3390/cancers10120514.
29. Prakash O, Hossain F, Danos D, Lassak A, Scribner R, Miele L. Racial Disparities in Triple Negative Breast Cancer: A Review of the Role of Biologic and Non-biologic Factors. Front Public Health **2020**;8:576964 doi 10.3389/fpubh.2020.576964.
30. Carey LA, Perou CM, Livasy CA, Dressler LG, Cowan D, Conway K, *et al.* Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. JAMA **2006**;295(21):2492-502 doi 10.1001/jama.295.21.2492.
31. Dietze EC, Sistrunk C, Miranda-Carboni G, O'Regan R, Seewaldt VL. Triple-negative breast cancer in African-American women: disparities versus biology. Nat Rev Cancer **2015**;15(4):248-54 doi 10.1038/nrc3896.

32. Barrow MA, Martin ME, Coffey A, Andrews PL, Jones GS, Reaves DK, *et al.* A functional role for the cancer disparity-linked genes, CRYbetaB2 and CRYbetaB2P1, in the promotion of breast cancer. Breast Cancer Res **2019**;21(1):105 doi 10.1186/s13058-019-1191-3.
33. Lindner R, Sullivan C, Offor O, Lezon-Geyda K, Halligan K, Fischbach N, *et al.* Molecular phenotypes in triple negative breast cancer from African American patients suggest targets for therapy. PLoS One **2013**;8(11):e71915 doi 10.1371/journal.pone.0071915.
34. Huo D, Feng Y, Haddad S, Zheng Y, Yao S, Han YJ, *et al.* Genome-wide association studies in women of African ancestry identified 3q26.21 as a novel susceptibility locus for oestrogen receptor negative breast cancer. Hum Mol Genet **2016**;25(21):4835-46 doi 10.1093/hmg/ddw305.
35. Haddad SA, Ruiz-Narvaez EA, Haiman CA, Sucheston-Campbell LE, Bensen JT, Zhu Q, *et al.* An exome-wide analysis of low frequency and rare variants in relation to risk of breast cancer in African American Women: the AMBER Consortium. Carcinogenesis **2016**;37(9):870-7 doi 10.1093/carcin/bgw067.
36. Newman LA, Kaljee LM. Health Disparities and Triple-Negative Breast Cancer in African American Women: A Review. JAMA Surg **2017**;152(5):485-93 doi 10.1001/jamasurg.2017.0005.
37. Piyarathna DWB, Balasubramanian A, Arnold JM, Lloyd SM, Karanam B, Castro P, *et al.* ERR1- and PGC1α-associated mitochondrial alterations correlate with pan-cancer disparity in African Americans. The Journal of Clinical Investigation **2019**;129(6):2351-6 doi 10.1172/JCI127579.
38. Mason K, Sathe A, Hess PR, Rong J, Wu CY, Furth E, *et al.* Niche-DE: niche-differential gene expression analysis in spatial transcriptomics data identifies context-dependent cell-cell interactions. Genome Biol **2024**;25(1):14 doi 10.1186/s13059-023-03159-6.
39. Zhang J, Tian XJ, Xing J. Signal Transduction Pathways of EMT Induced by TGF-β, SHH, and WNT and Their Crosstalks. J Clin Med **2016**;5(4) doi 10.3390/jcm5040041.
40. Zhang X, Kazerounian S, Duquette M, Perruzzi C, Nagy JA, Dvorak HF, *et al.* Thrombospondin-1 modulates vascular endothelial growth factor activity at the receptor level. Faseb j **2009**;23(10):3368-76 doi 10.1096/fj.09-131649.
41. Hugo W, Zaretsky JM, Sun L, Song C, Moreno BH, Hu-Lieskovan S, *et al.* Genomic and Transcriptomic Features of Response to Anti-PD-1 Therapy in Metastatic Melanoma. Cell **2016**;165(1):35-44 doi 10.1016/j.cell.2016.02.065.
42. Costa A, Kieffer Y, Scholer-Dahirel A, Pelon F, Bourachot B, Cardon M, *et al.* Fibroblast Heterogeneity and Immunosuppressive Environment in Human Breast Cancer. Cancer Cell **2018**;33(3):463-79 e10 doi 10.1016/j.ccell.2018.01.011.
43. Launonen I-M, Erkan EP, Niemiec I, Junquera A, Hincapié-Otero M, Afenteva D, *et al.* Chemotherapy induces myeloid-driven spatial T-cell exhaustion in ovarian cancer. bioRxiv **2024**:2024.03.19.585657 doi 10.1101/2024.03.19.585657.
44. Jiang YZ, Ma D, Suo C, Shi J, Xue M, Hu X, *et al.* Genomic and Transcriptomic Landscape of Triple-Negative Breast Cancers: Subtypes and Treatment Strategies. Cancer Cell **2019**;35(3):428-40.e5 doi 10.1016/j.ccell.2019.02.001.
45. Hammerl D, Martens JWM, Timmermans M, Smid M, Trapman-Jansen AM, Foekens R, *et al.* Spatial immunophenotypes predict response to anti-PD1 treatment and capture distinct paths of T cell evasion in triple negative breast cancer. Nat Commun **2021**;12(1):5668 doi 10.1038/s41467-021-25962-0.
46. Gruosso T, Gigoux M, Manem VSK, Bertos N, Zuo D, Perlitch I, *et al.* Spatially distinct tumor immune microenvironments stratify triple-negative breast cancers. J Clin Invest **2019**;129(4):1785-800 doi 10.1172/JCI96313.

47. Bagati A, Kumar S, Jiang P, Pyrdol J, Zou AE, Godicelj A, *et al.* Integrin alphavbeta6-TGFbeta-SOX4 Pathway Drives Immune Evasion in Triple-Negative Breast Cancer. Cancer Cell **2021**;39(1):54-67 e9 doi 10.1016/j.ccell.2020.12.001.
48. Akil A, Gutiérrez-García AK, Guenter R, Rose JB, Beck AW, Chen H, Ren B. Notch Signaling in Vascular Endothelial Cells, Angiogenesis, and Tumor Progression: An Update and Prospective. Front Cell Dev Biol **2021**;9:642352 doi 10.3389/fcell.2021.642352.
49. Wang Z, Li Y, Kong D, Sarkar FH. The role of Notch signaling pathway in epithelial-mesenchymal transition (EMT) during development and tumor aggressiveness. Curr Drug Targets **2010**;11(6):745-51 doi 10.2174/138945010791170860.
50. Yuan X, Wu H, Han N, Xu H, Chu Q, Yu S, *et al.* Notch signaling and EMT in non-small cell lung cancer: biological significance and therapeutic application. J Hematol Oncol **2014**;7:87 doi 10.1186/s13045-014-0087-z.
51. Singh S, Lee N, Pedroza DA, Bado IL, Hamor C, Zhang L, *et al.* Chemotherapy Coupled to Macrophage Inhibition Induces T-cell and B-cell Infiltration and Durable Regression in Triple-Negative Breast Cancer. Cancer Res **2022**;82(12):2281-97 doi 10.1158/0008-5472.CAN-21-3714.
52. Zheng C, Xu X, Wu M, Xue L, Zhu J, Xia H, *et al.* Neutrophils in triple-negative breast cancer: an underestimated player with increasingly recognized importance. Breast Cancer Research **2023**;25(1):88 doi 10.1186/s13058-023-01676-7.
53. Kim IS, Gao Y, Welte T, Wang H, Liu J, Janghorban M, *et al.* Immuno-subtyping of breast cancer reveals distinct myeloid cell profiles and immunotherapy resistance mechanisms. Nat Cell Biol **2019**;21(9):1113-26 doi 10.1038/s41556-019-0373-7.
54. Welte T, Kim IS, Tian L, Gao X, Wang H, Li J, *et al.* Oncogenic mTOR signalling recruits myeloid-derived suppressor cells to promote tumour initiation. Nat Cell Biol **2016**;18(6):632-44 doi 10.1038/ncb3355.
55. Schapiro D, Jackson HW, Raghuraman S, Fischer JR, Zanotelli VRT, Schulz D, *et al.* histoCAT: analysis of cell phenotypes and interactions in multiplex image cytometry data. Nat Methods **2017**;14(9):873-6 doi 10.1038/nmeth.4391.
56. Xu Y, Zhang L, Thaiparambil J, Mai S, Perera DN, Zhang J, *et al.* Patients with Lung Cancer of Different Racial Backgrounds Harbor Distinct Immune Cell Profiles. Cancer Res Commun **2022**;2(8):884-93 doi 10.1158/2767-9764.CRC-22-0057.
57. Merritt CR, Ong GT, Church SE, Barker K, Danaher P, Geiss G, *et al.* Multiplex digital spatial profiling of proteins and RNA in fixed tissue. Nat Biotechnol **2020**;38(5):586-99 doi 10.1038/s41587-020-0472-9.
58. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nat Methods **2012**;9(7):671-5 doi 10.1038/nmeth.2089.
59. Carpenter AE, Jones TR, Lamprecht MR, Clarke C, Kang IH, Friman O, *et al.* CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biol **2006**;7(10):R100 doi 10.1186/gb-2006-7-10-r100.
60. Stringer C, Wang T, Michaelos M, Pachitariu M. Cellpose: a generalist algorithm for cellular segmentation. Nat Methods **2021**;18(1):100-6 doi 10.1038/s41592-020-01018-x.
61. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, *et al.* Fiji: an open-source platform for biological-image analysis. Nat Methods **2012**;9(7):676-82 doi 10.1038/nmeth.2019.
62. Berg S, Kutra D, Kroeger T, Straehle CN, Kausler BX, Haubold C, *et al.* ilastik: interactive machine learning for (bio)image analysis. Nat Methods **2019**;16(12):1226-32 doi 10.1038/s41592-019-0582-9.
63. Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software **2015**;67(1):1 - 48 doi 10.18637/jss.v067.i01.

64. Zhu Q, Shah S, Dries R, Cai L, Yuan GC. Identification of spatially associated subpopulations by combining scRNAseq and sequential fluorescence in situ hybridization data. Nat Biotechnol **2018** doi 10.1038/nbt.4260.
65. Rousseeuw PJ. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics **1987**;20:53-65.
66. Obayashi T, Kinoshita K. COXPRESdb: a database to compare gene coexpression in seven model animals. Nucleic Acids Res **2011**;39(Database issue):D1016-22 doi 10.1093/nar/gkq1147.
67. Wu SZ, Al-Eryani G, Roden DL, Junankar S, Harvey K, Andersson A, *et al.* A single-cell and spatially resolved atlas of human breast cancers. Nat Genet **2021**;53(9):1334-47 doi 10.1038/s41588-021-00911-1.
68. Wigerblad G, Cao Q, Brooks S, Naz F, Gadkari M, Jiang K, *et al.* Single-Cell Analysis Reveals the Range of Transcriptional States of Circulating Human Neutrophils. J Immunol **2022**;209(4):772-82 doi 10.4049/jimmunol.2200154.
69. Alvarez-Breckenridge C, Markson SC, Stocking JH, Nayyar N, Lastrapes M, Strickland MR, *et al.* Microenvironmental Landscape of Human Melanoma Brain Metastases in Response to Immune Checkpoint Inhibition. Cancer Immunol Res **2022**;10(8):996-1012 doi 10.1158/2326-6066.CIR-21-0870.
70. Efremova M, Vento-Tormo M, Teichmann SA, Vento-Tormo R. CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. Nat Protoc **2020**;15(4):1484-506 doi 10.1038/s41596-020-0292-x.
71. Reimand J, Kull M, Peterson H, Hansen J, Vilo J. g:Profiler--a web-based toolset for functional profiling of gene lists from large-scale experiments. Nucleic Acids Res **2007**;35(Web Server issue):W193-200 doi 10.1093/nar/gkm226.
72. Merico D, Isserlin R, Stueker O, Emili A, Bader GD. Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. PLoS One **2010**;5(11):e13984 doi 10.1371/journal.pone.0013984.

Supplementary Table 1. Summarized clinical information of patients in the Tissue Microarray (TMA) examined using imaging mass cytometry (IMC) and Nanostring Geo Mx Digital Spatial Profiling (refer Supplementary Figure 1 for H&E images).

Variable	**African American (n=26)**	**European American (n=31)**
AGE	58 ±11.09	60.06 ± 13.00
Receptor Expression		
ER		
negative	26	31
positive	0	0
Unknown	0	0
PR		
negative	26	31
positive	0	0
Unknown	0	0
HER2/NEU		
negative	25	31
positive	1	0
Unknown	0	0
Regions of Interest (ROI) Selected based on Immune Infiltration Profile (using H&E, multiple/patient)	47	51
Tumor Center Immune Poor	10	14
Tumor Center Immune Rich	17	15
Tumor Periphery Immune Poor	1	0
Tumor Periphery Immune Rich	19	22
Tumor Grade		
Grade 1	0	1
Grade 2	1	7
Grade 3	24	23
Unknown	1	0

Tumor Stage		
Stage 1	8	10
Stage 1A	1	4
Stage 2A	6	11
Stage 2B	5	5
Stage 3A	4	0
Stage 3B	0	0
Stage 4	1	1
Unknown	1	0
BMI	35.28 ± 9.27	29.16 ± 6.09
Mets		
Yes	3	1
No	22	30
Unknown	1	0

Supplementary Table 2: List of antibodies and the metal tags used for Imaging Mass Cytometry (IMC) analysis of discovery samples examined in Tissue Microarray format (refer Supplementary Figure 2 for H&E images).

Metal	**Antibody**
143Nd	Vimentin
144Nd	PLK1
145Nd	AR
146Nd	CD16
147Sm	CD163
148Nd	Pan Cy-tokeratin
150Nd	PD-L1
151Eu	CD31
152Sm	CD45
154Sm	CD11c
155Gd	FOXP3
156Gd	CD4
158Gd	E-Cadherin

159Tb	CD68
161Dy	CD152/CTLA4
162Dy	CD8a
163Dy	VEGF
164Dy	HIF1a
165Ho	PD-1
166Er	CD45RA
167Er	Granzyme B
168Er	Ki-67
170Er	CD3
173Yb	CD45RO
175Lu	KIFC1
176Yb	pHH3

Supplementary Table 3: List of antibodies and the metal tags used for Imaging Mass Cytometry (IMC) analysis of validation samples (refer Supplementary Figures 7 and 8 for H&E images).

Metal Tag	Antibody
143Nd	Vimentin
144Nd	PLK1
145Nd	AR
146Nd	CD16
147Sm	CD163
148Nd	Pan Cytokeratin
149Sm	CD31
150Nd	PD-L1
151Eu	PD-1
152Sm	CD45
154Sm	CD11c
155Gd	FOXP3
156Gd	CD4
158Gd	E-Cadherin
159Tb	CD68
161Dy	CD152/CTLA4
162Dy	CD8a
163Dy	VEGF
164Dy	MPO
165Ho	HIF1a
166Er	CD45RA

167Er	Granzyme B
168Er	Ki-67
170Er	CD3
173Yb	CD45RO
175Lu	KIFC1
176Yb	pHH3

Figure Legends

Figure 1: Imaging Mass Cytometry (IMC) analysis of African American (AA) and European American (EA) triple negative breast cancer (TNBC). a) Illustration of the workflow used to conduct IMC analysis of AA and EA TNBC tumors in a tissue microarray (TMA) format. b) Unsupervised clustering of single cells segmented from IMC data obtained from AA and EA TNBC tumors resulted in 20 clusters. c) Representative t-SNE plot showing the distribution of 20 clusters described in (b). d) Overlay of expressions of E-Cadherin, CD45RA, CD31, CD68, and Vimentin, on the t-SNE. e) Spatial distribution of proteins described in panel (d) in TNBC tumors.

Figure 2: Spatial proximity analysis reveals unique spatial cell-cell interactions in African American (AA) and European American (EA) TNBC determined using IMC. a) Unique interactions in AA (left) and EA (right) TNBC formed between pairs of 20 spatially resolved clusters. AA tumors have stronger spatially resolved cell-cell interactions with a predominance of homo-typic interactions as shown for clusters 5, 6, 13, 15 and 20. EA tumors have weaker spatially resolved cell-cell interactions with prevalence of hetero-typic interactions as shown for clusters 5 and 20, and clusters 18 and 20. b) Top nine spatially resolved cell-cell interactions in AA (blue bars) compared to EA (red bars) TNBC, with associated significance computed using -log10 P value (green bars). AA TNBC tumors are marked by Endothelial-Macrophage-Vimentin (Endo-Mac-Vim) interactions. c) Same as in (b) but showing top six spatially resolved cell-cell interactions in EA (red bars) compared to AA (blue bars) TNBC with associated significance computed using -log10 P value (green bars). EA tumors are marked by interactions that suggest immune exhaustion. d) Spatial illustration of key AA-associated cell-cell interactions identified using IMC. e) Spatial illustration of key EA-associated cell-cell interactions identified using IMC. f) Multiplex immunofluorescence validation of Endo (CD31)-Mac (CD163) interaction in AA TNBC but not in EA TNBC. Representative images from 3 AA and EA TNBC patients are shown. Magnified image (refer large white box) of areas marked (small white box) are included as inset in each image. g, h) pixel-level co-localization quantification of spatial clustering of Endo-Mac interaction shown in panel (f) across the entire tumor tissue and per cell within the tumor, in AA and EA TNBC.

Figure 3. AA-associated cell-cell interactions correlate with patient survival. a) Kaplan Meier (KM) plot showing significantly poor overall survival among TNBC patients when patients

were stratified by the sum of top nine AA associated cell-cell interactions (see Figure 2b). b) KM plot showing the stratification of TNBC patients based on the sum of top six EA associated cell-cell interactions (see Figure 2c). EA-interactions correlate minimally with overall survival. c) Stratification of TNBC patients based on the combination of the nine AA associated cell-cell interactions (see Figure 2b) and self-reported race information. A larger survival gap is observed for AA TNBC compared to EA TNBC. (d) Same as c, but for EA-specific cell-cell interactions. Combining the top six EA associated cell-cell interactions with race allows us to correlate EA-specific interactions to EA TNBC patient survival. All P values computed using one-tailed log-rank test.

Figure 4: AA associated cell-cell interactions show unique spatial clustering in the TNBC spatial transcriptomics data. a) Overall approach used to identify spatial localization of AA and EA-specific cell-cell interactions in TNBC tissues. Each query is an AA or EA-specific interaction, specified first by IMC-derived protein markers of interacting cell types, then converted to gene names. The query is interrogated in the TNBC ST collection20 to check spatial pattern of co-localization of query genes. (b) Top five AA-specific interactions (AA-Q1 to AA-Q5) interrogated in the ST data set. The protein symbols and corresponding gene symbols are in blue and black fonts, respectively. (c) Representative example of spatial co-localization for AA-Q3 in 10 AA and 10 EA samples obtained from the TNBC ST data set. Importantly, genes belonging to AA-Q3 show strong co-localization (marked by black outlines) only in AA but not in EA TNBC. Scale bar represents the strength of co-localization, defined as averaged log-transformed normalized read count. (d) Same as in (c) but for top five AA-specific interactions (AA-Q1 to AA-Q5) in a representative AA and EA TNBC. AA-Q1 to AA-Q5 show strongly similar spatial patterns in AA but not in EA TNBC. (e) Quantification of spatial clustering of query-co-localized spots for AA-Q1 to AA-Q5 in AA (Blue) and EA (Red) TNBC tumors. Each bar is an average of 10 samples. For all the five interactions, spatial clustering was significantly higher in the AA TNBC compared to EA TNBC. Significance computed using T-statistics. (f) Correlation plots for AA (top 2 rows) and EA TNBC samples (bottom 2 rows) examining the correlation of spatial patterns between the top five AA-enriched queries (AA-Q1 to AA-Q5). Each plot corresponds to a sample and contains 5 columns and 5 rows representing the top five AA-enriched interactions (AA-Q1 to AA-Q5). The color of each cell in the plot denotes Pearson correlation. A higher cross-query correlation is indicative of a multicellular niche enriched in AA TNBC, made of endothelial (AA-Q1, Q3), macrophage (AA-

Q4, Q5), and vimentin (AA-Q3, Q5) populations. Such strong cross-query correlation is not observed in EA TNBC.

Figure 5: EA associated cell-cell interactions show diffused distribution in the published TNBC spatial transcriptomics data. (a) Top five EA-specific interactions (EA-Q1 to EA-Q5) interrogated in the ST data set. The protein symbols and corresponding gene symbols are in blue and black fonts, respectively. These top five EA-specific interactions constitute the EA-specific niche. These interactions are mostly defined by GZMB, CTLA4 and HIF1A. (b) Representative example of diffused distribution of proteins in EA-Q3 in 10 AA and 10 EA samples obtained from the TNBC ST data set. Genes belonging to EA-Q3 show higher abundance in EA compared to AA TNBC. The scale bar represents the average log-transformed normalized expression of genes in EA-Q3. (c) Bar plot showing the difference in abundance for the genes associated with EA-specific interactions in AA (blue bar) and EA (red bar) TNBC (10 each). To obtain the abundance measure, the number of spots with EA query co-localization in each ST sample was quantified, and fold-over-AA enrichment was computed. Significance computed using T-statistics.

Figure 6: Distinct gene signatures define unique cell types in AA and EA niches.

(a) Overall approach used to identify niche-associated gene signatures. For AA and EA-specific interactions (AA-Q1 to Q5, and EA-Q1 to Q5), independently, a set of extended signature genes (ESG) was obtained based on co-expression within the query-co-localized spots. ESGs were next computed to select those that were significantly elevated in AA (i.e., AA>EA for AA-niche) or in EA (i.e., EA>AA for EA-niche). Each query thus generates one niche, producing 5 AA-niches and 5 EA-niches. (b) Representative heat map showing cell-type specific expression for ESGs associated with AA niche defined by expansion of AA-Q3. These ESGs in combination encode various cell types as described below the heat map. AA-Q3 niche is enriched in Cancer-Associated Fibroblasts (CAF), endothelial cells and cells of myeloid lineage. AA-Q3 is poor in T-cells and neutrophils. (c) same as in (b) but for heat map showing cell-type specific expression for ESGs associated with EA niche defined by expansion of EA-Q3. These ESGs in combination encode for exhausted T cells and neutrophils and are poor in Cancer-Associated Fibroblasts (CAF) and endothelial cells. (d) Average quantification of number of co-localized spots in AA (blue) and EA (red) TNBC corresponding to ESGs associated with AA-specific niches (AA-Q1 to AA-Q5). (e) Same as in (d) but for EA-specific niches (EA-Q1 to EA-Q5). Significance for panels (d) and (e) computed

using T-statistics. (f-g) Gene set enrichment analysis of ESGs associated with AA-niche (f) and those associated with EA-niche (g). -log10Pvalue is shown.

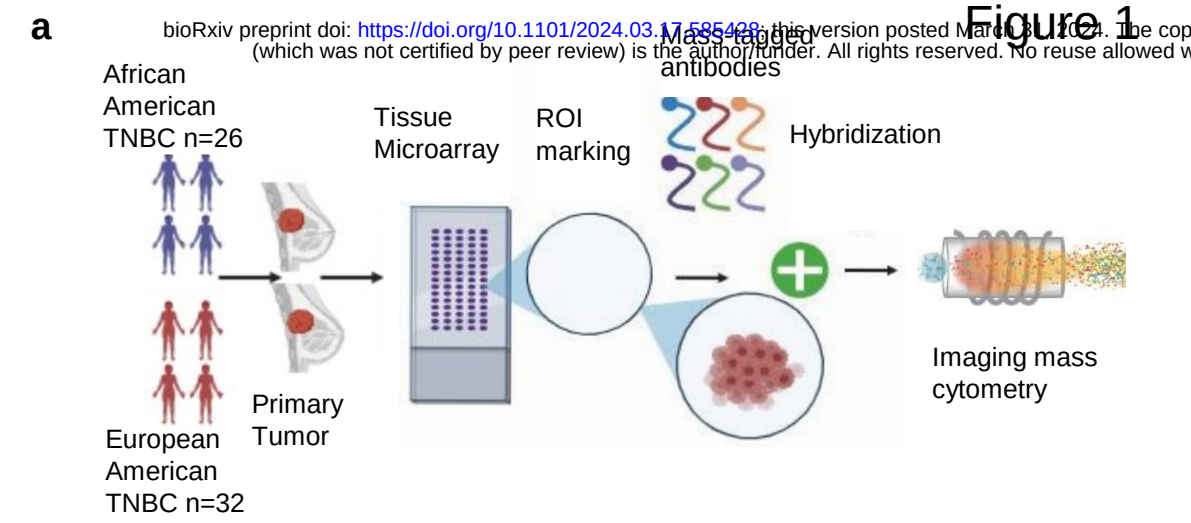
Figure 7: Ligand-receptor interactions and analysis of ESGs associated with AA and EA niche reveal distinct mechanisms for promoting tumor progression and instilling immune suppression. (a) AA niche-specific ligand-receptor interactions (AA-Q1 to AA-Q5) derived from the ESGs reveal activation of multiple tumor-promoting pathways that include cell surface proteins, adhesion molecules, extracellular matrix components, endothelial cells, and multiple signaling pathways including platelet derived growth factor (PDGFR), Transforming Growth Factor beta (TGF-ß), WNT, Notch, Thrombospondin, and Placental Growth Factor (PGF). In contrast, EA-niche specific ligand-receptor interactions (EA-Q1 to EA-Q5) highlight cytokine activities and T-cell exhaustion. (b) Normalized frequencies of ligand (rows) and receptor (columns) interactions in AA-niche, analyzed using ESGs of AA-Q3. (c) Normalized frequencies of ligand-receptor interactions in EA-niche, analyzed using ESGs of EA-Q3. (d) Prevalence of known genes associated with T-cell exhaustion21 within each of the EA-associated niches (EA-Q1 to EA-Q5). Green and white boxes indicate presence and absence of the gene in each of the niches. (e) Quantification of average number of spots containing the gene associated with T-cell exhaustion in AA (10) and EA (10) TNBC ST samples. Significance computed using T-statistics. (f)-(j) Average expression of individual genes associated with T-cell exhaustion in each of the 5 niches. Panels d–h represent expression in EA-niches 1-5. Significance computed using T-statistics. Importantly, across all the 5 niches, genes associated with T-cell exhaustion or immune suppression were significantly more pronounced in EA compared to AA TNBC.

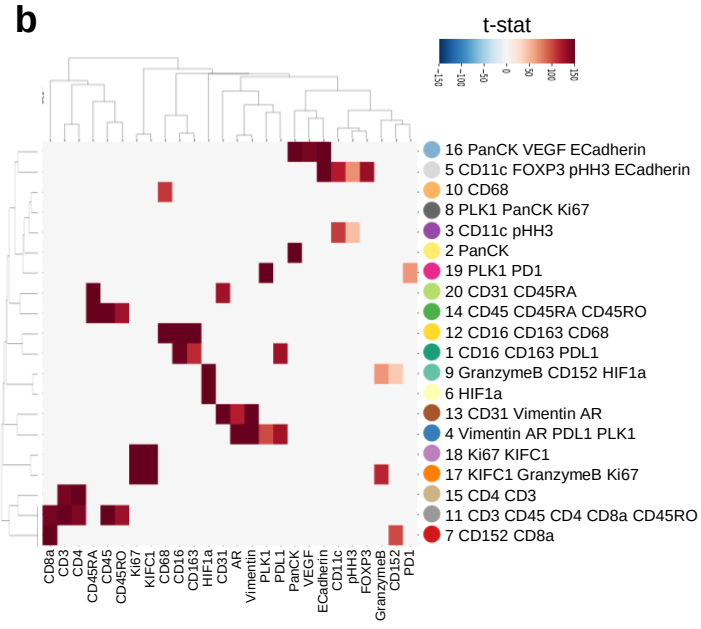
Figure 8: Independent validation of AA and EA niches using spatial transcriptomics and GeoMX DSP system (NanoString). (a) Spatial transcriptomics analysis conducted on 4 independent formalin fixed paraffin embedded (FFPE) AA tumors and 5 EA tumors verifies elevated expression of ESGs associated with AA-Q3 to AA-Q5 in AA compared to EA TNBC. (b) Same as in (a) but verifying the elevated expression of ESGs associated with EA-Q3 to EA-Q5 in EA compared to AA TNBC. (c) Representative spatial co-localization of ESGs associated with AA-Q3 niche in AA (AA patients) but not in EA (EA patients) TNBC. The correlation plots on the right show higher inter-niche correlation within AA TNBC for AA-associated niche (AA-Q1 to AA-Q5, upper panel), which is not seen in the EA TNBC samples

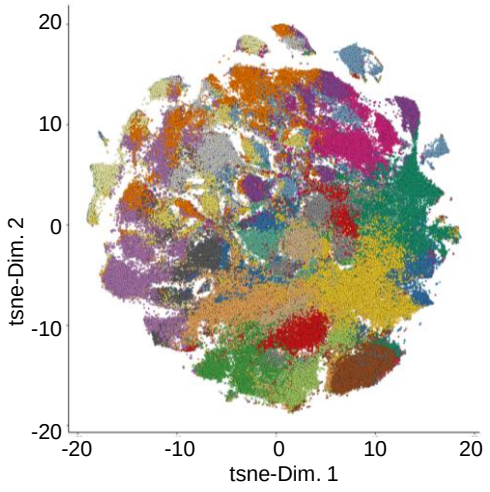
(lower panel). (d) Analysis of Nanostring GeoMx data from 26 AA and 32 EA TNBC tumors used for IMC in Figure 1, shows significant enrichment (computed by difference in Z-score) of cell-cell interactions representing Endo-Mac-EMT in either or both the tumor cell (red bar) or immune cell (blue bar) compartment in AA compared to EA (upper panel). The lower panel shows the significant enrichment (computed by difference in Z-score) of cell-cell interactions representing Immune Exhaustion and Immune Suppression in either or both the tumor cell (red bar) or immune cell (blue bar) compartment in EA compared to AA (upper panel). Significance computed using T-statistics.

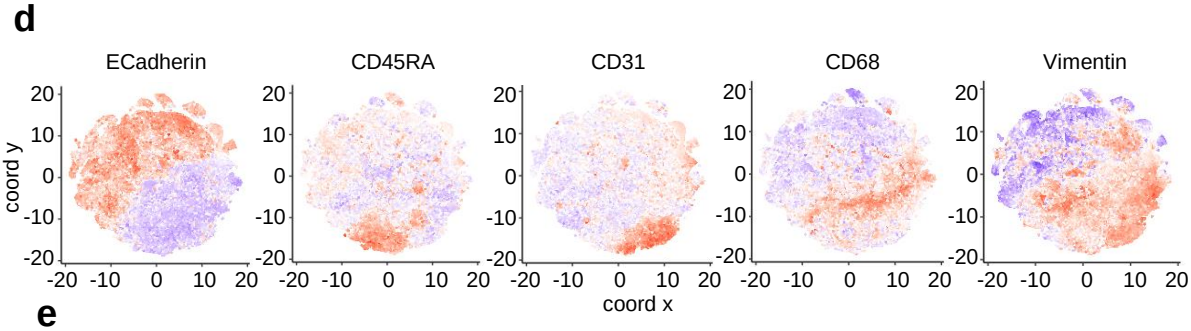
Figure 9: AA and EA TNBC have distinct Tumor Microenvironments (TME). (a) Cartoon representation of AA TME consists of close interaction between endothelial cells-macrophages and tumor cells expressing vimentin undergoing epithelial mesenchymal transition. These niches promote tumor progression and suppress host-immune surveillance. (b) Cartoon representation of EA TME highlighting the prevalence of immune exhaustion and immune suppressive niches. These are encoded by exhausted T-cells and neutrophils in the background of myeloid cells.

Figure 1









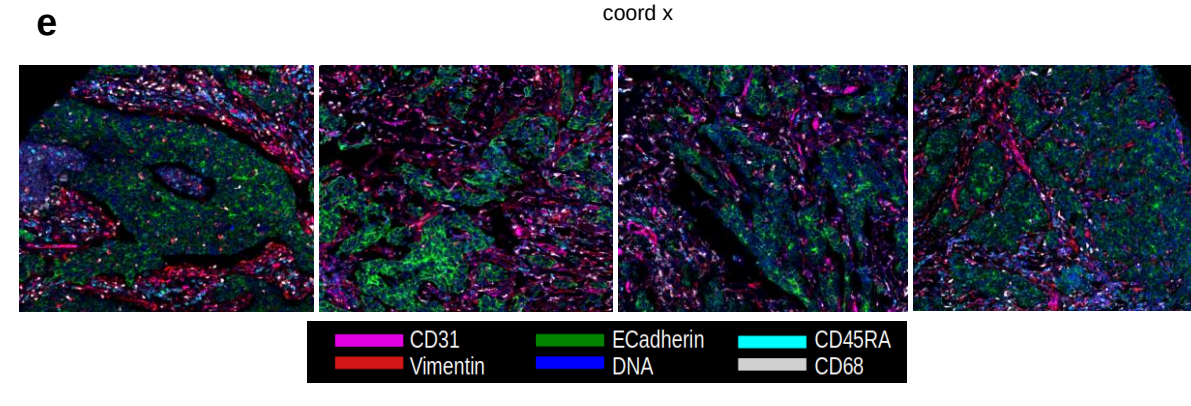


Figure 2

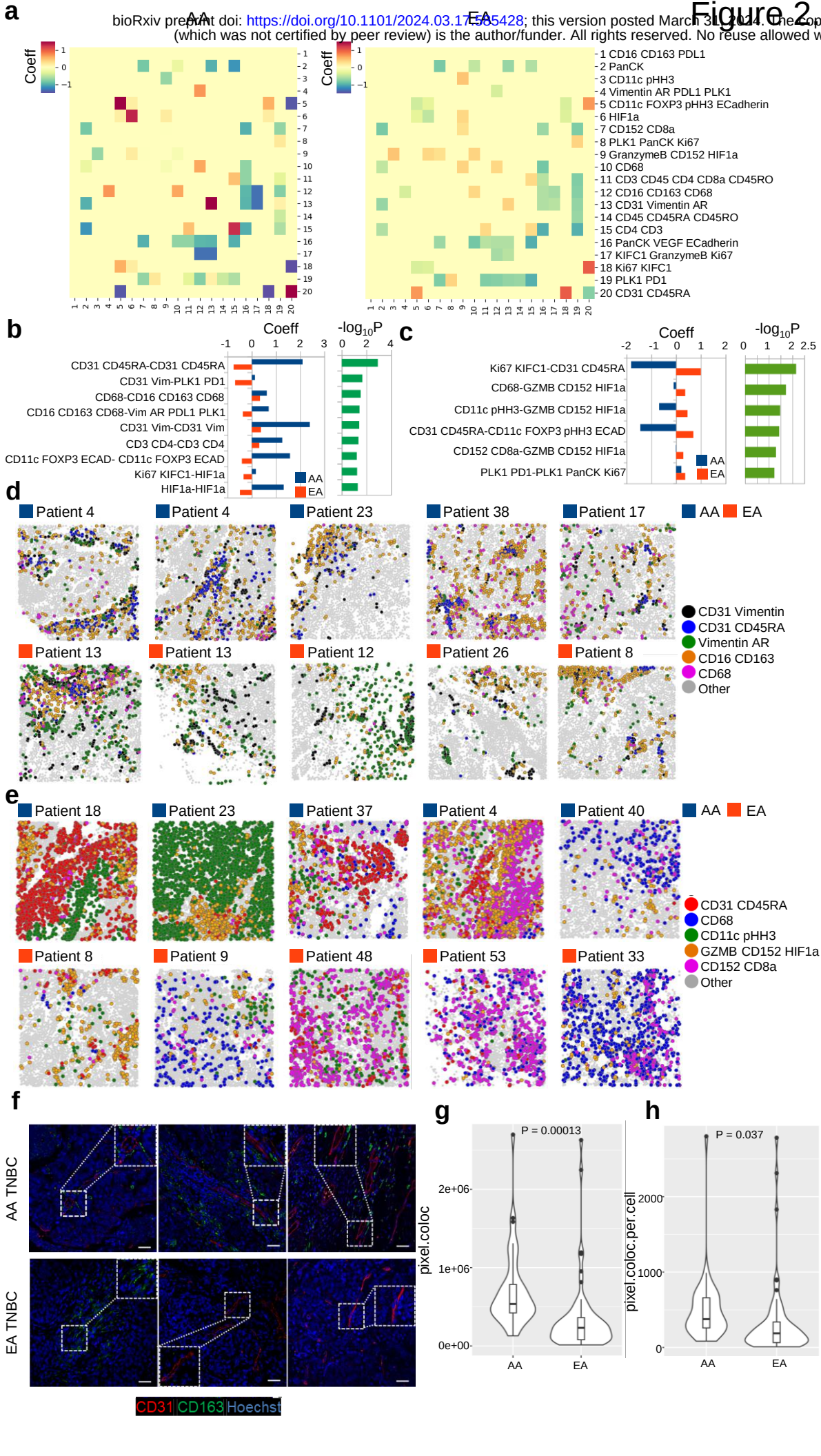
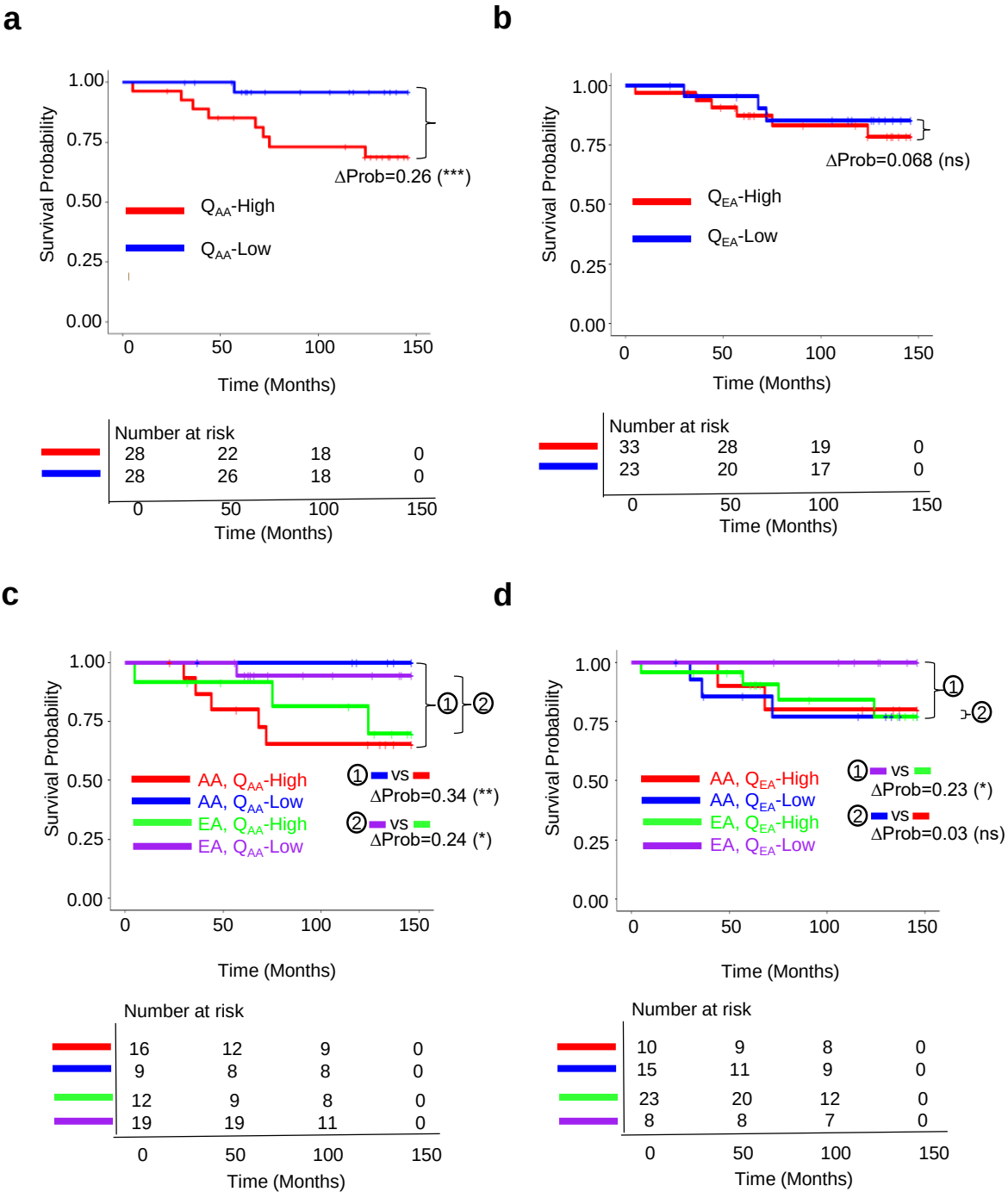
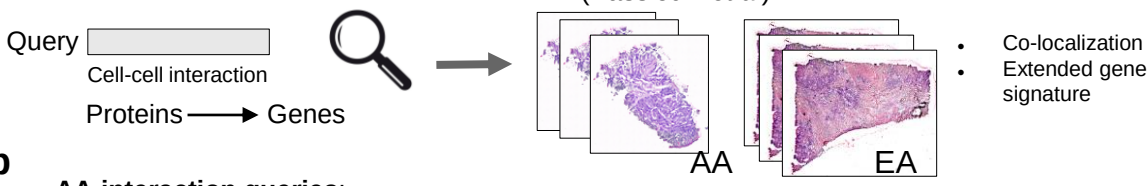


Figure 3



a

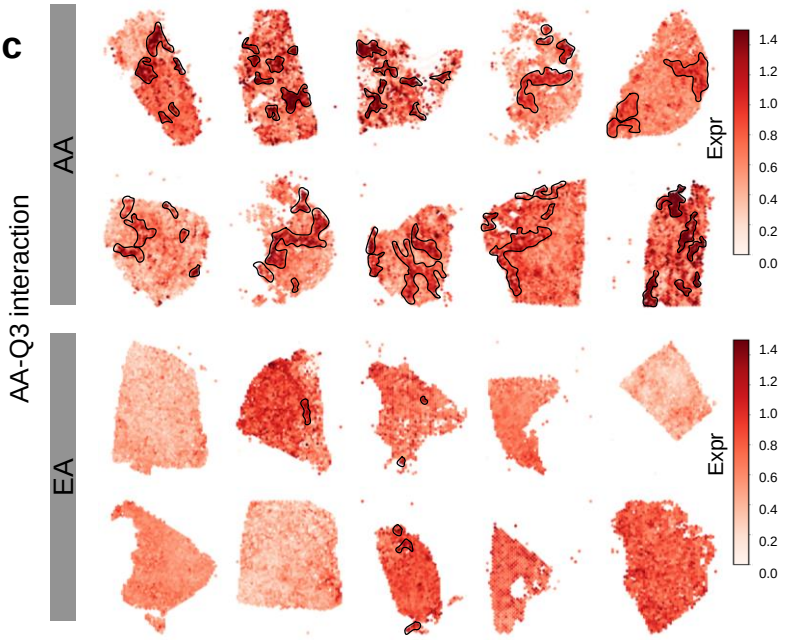


b

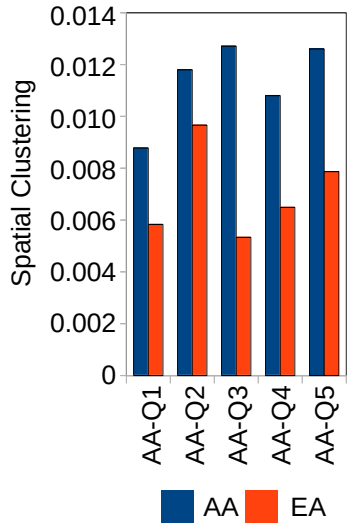
AA interaction queries:

AA-Q1: *PECAM1 PTPRC* (CD31 CD45RA)
AA-Q2: *PECAM1 PTPRC CCR7 SELL IL7R IL2RG* (CD31 CD45RA[#])
AA-Q3: *PECAM1 VIM AR PLK1 PDCD1* (CD31 Vim AR PDL1 PLK1)
AA-Q4: *CD68 FCGR3A FCGR3B CD163* (CD68 CD16 CD163)
AA-Q5: *CD68 CD163 FCGR3A FCGR3B VIM AR PLK1* (CD68 CD163 CD16 Vim AR PLK1)

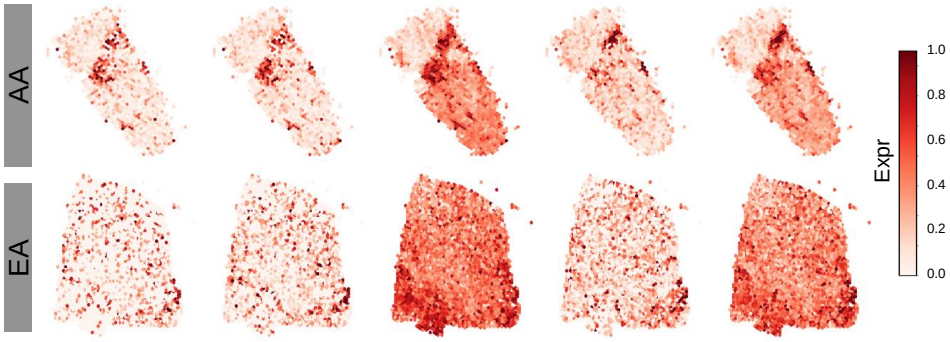
c



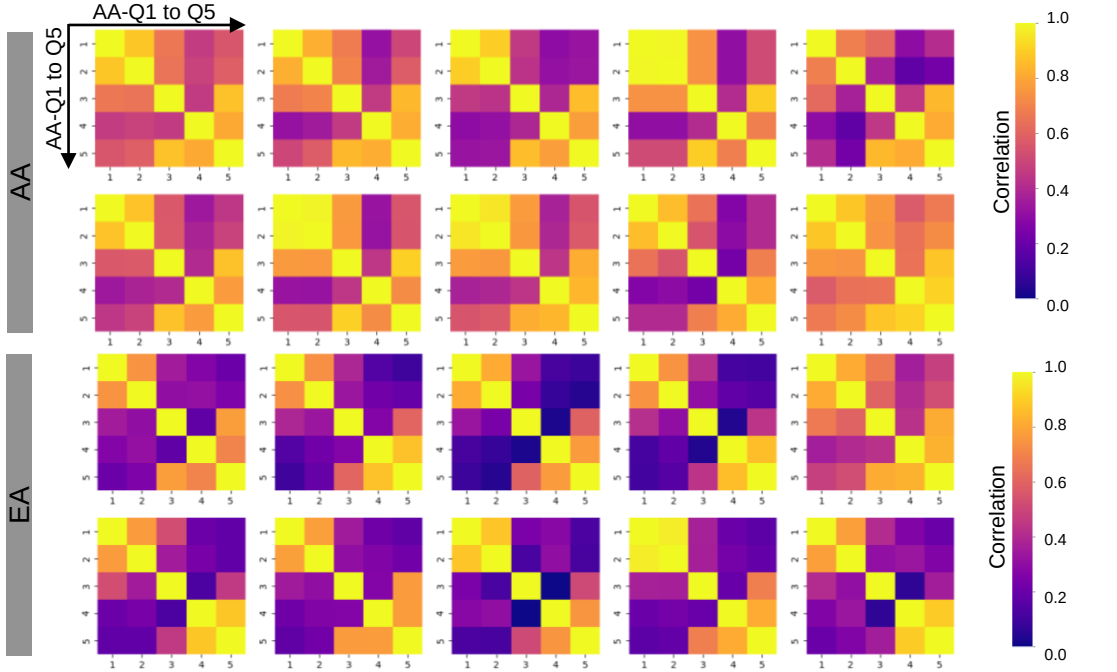
d



e



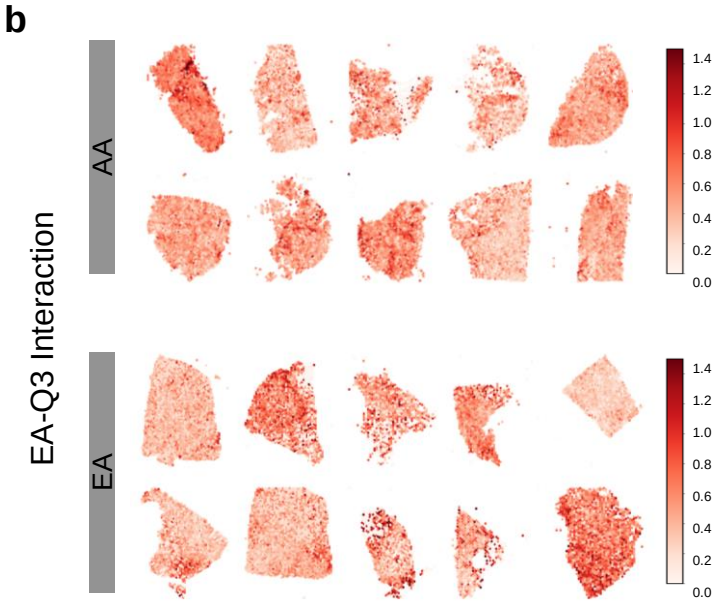
f

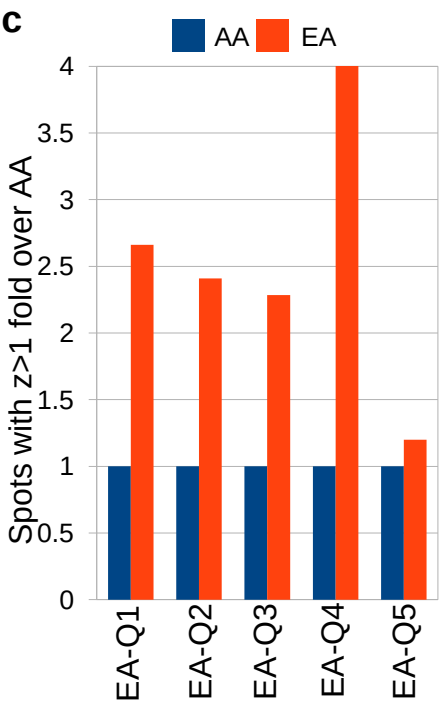


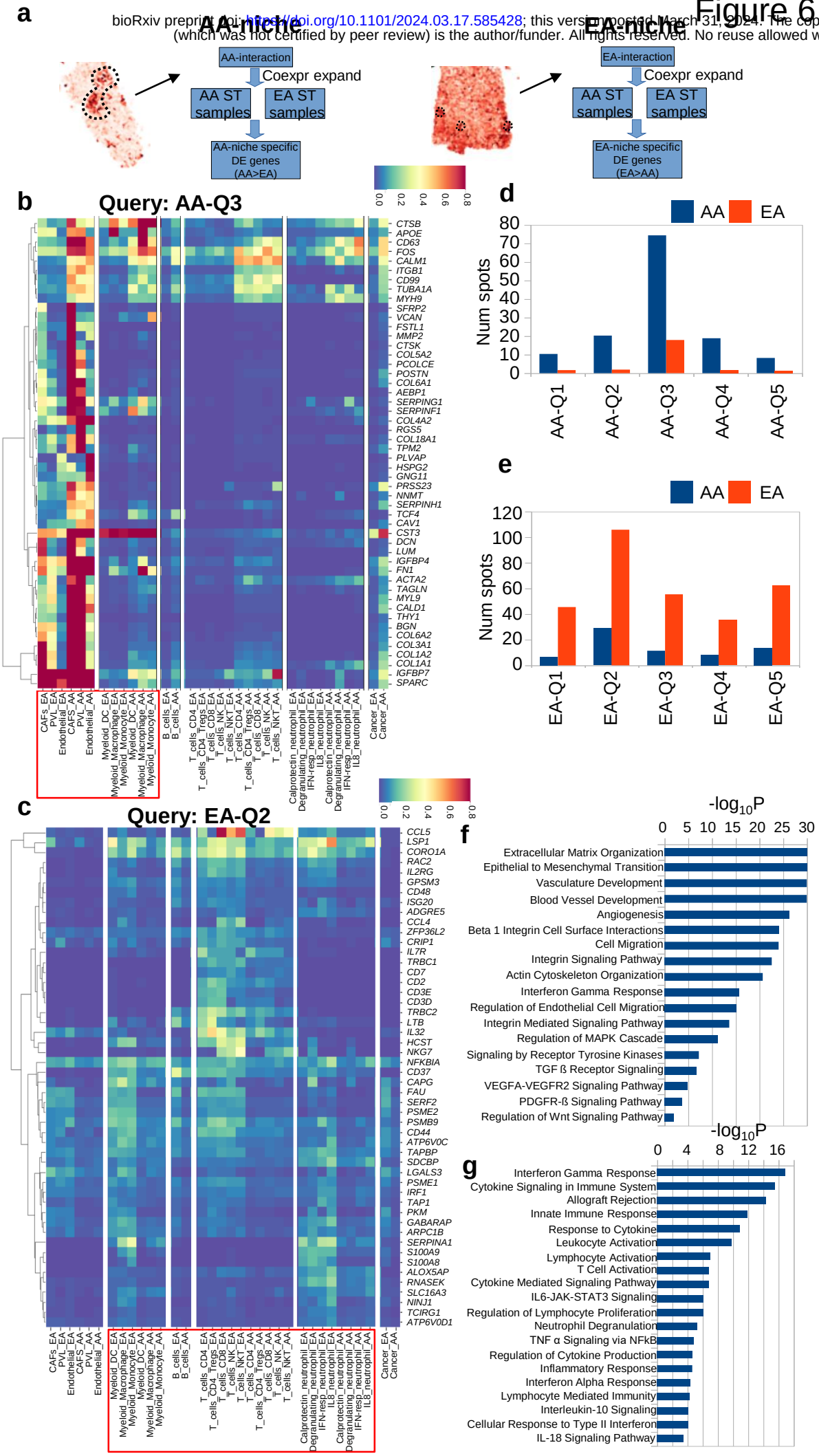
a

EA Interaction queries:

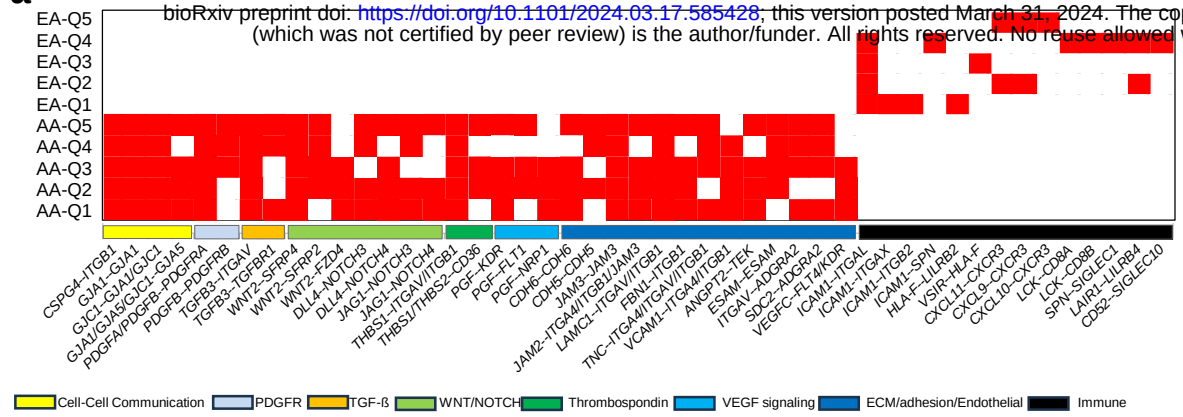
EA-Q1: *MKI67 KIFC1 PECAM1 PTPRC* (Ki67 KIFC1 CD31 CD45RA)
EA-Q2: *CD68 GZMB CTLA4 HIF1A* (CD68 GZMB CD152 HIF1a)
EA-Q3: *ITGAX GZMB CTLA4 HIF1A* (CD11C GZMB CD152 HIF1a)
EA-Q4: *PECAM1 PTPRC ITGAX FOXP3 CDH1* (CD31 CD45RA CD11c FOXP3 ECAD)
EA-Q5: *CTLA4 GZMB CD8A HIF1A* (CD152 GZMB CD8a HIF1a)







a



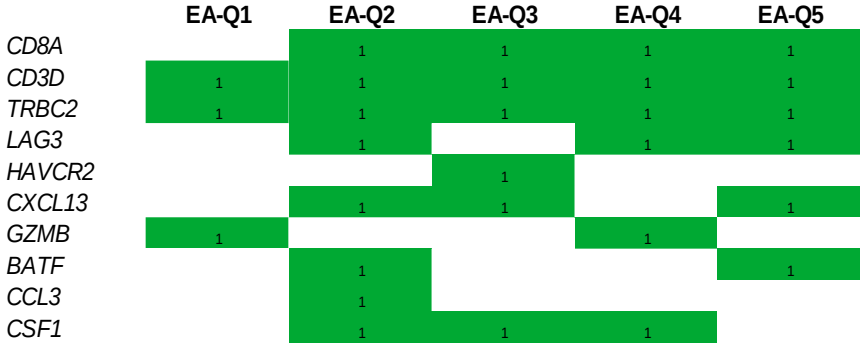
b

	PVL	Endo	CAF	Mac	Mono	Myoepith	Myeloid	Cancer
Perivascular Like (PVL)	4.86	6.58	3.52	3.02	0.5	0	0.75	0.5
Endothelial (Endo)	4.77	9	2.19	0.77	0	0	0	0.25
Cancer Associated Fibroblast (CAF)	2.41	4	5.58	2.66	0	1	0.5	0
Macrophage (Mac)	0.36	2	0.86	0.61	0	0	0	0
Monocyte (Mono)	2	0.75	0.5	0.75	0.5	0	0	0
Myo-epithelial (Myoepith)	0.25	0.25	1.25	0.75	0	0.5	0	0
Myeloid	0	0	0	0	0	0	0	0
Cancer	0	0.75	0.25	0.25	0	0	0	0.25

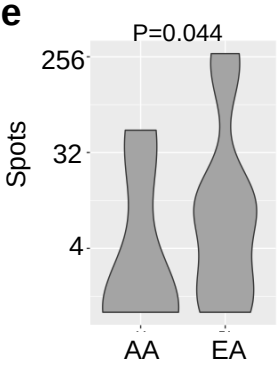
c

	Neutro	T	Mac	Mono	DC	Endo	B
Neutrophil (Neutro)	0.5	0.33	0.33	1	0.16	0.33	0
T-cells (T)	0.33	2	0	0.33	0	0	0
Macrophage (Mac)	0	1.5	1	0	1.5	0	0
Monocyte (Mono)	0.16	0	1.33	1	0.16	0	0
Dendritic Cells (DC)	0	0	0	0	0	0	0
Endothelial (Endo)	0.16	0	0.33	1	0.16	0	0
B-cells (B)	0.16	0	0	0.16	0	0	0

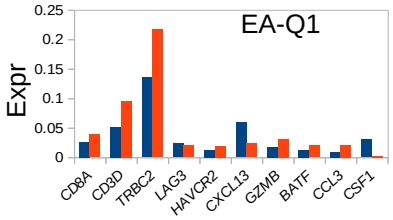
d



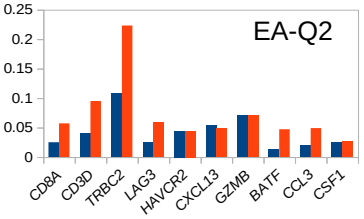
e



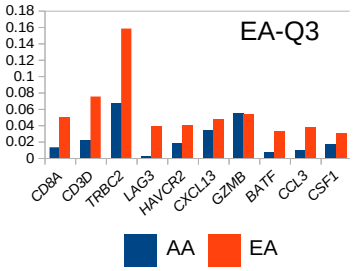
f



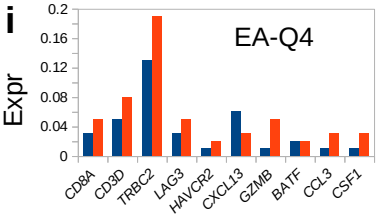
g



h



i



j

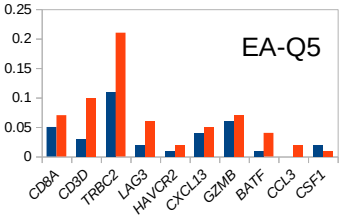


Figure 7

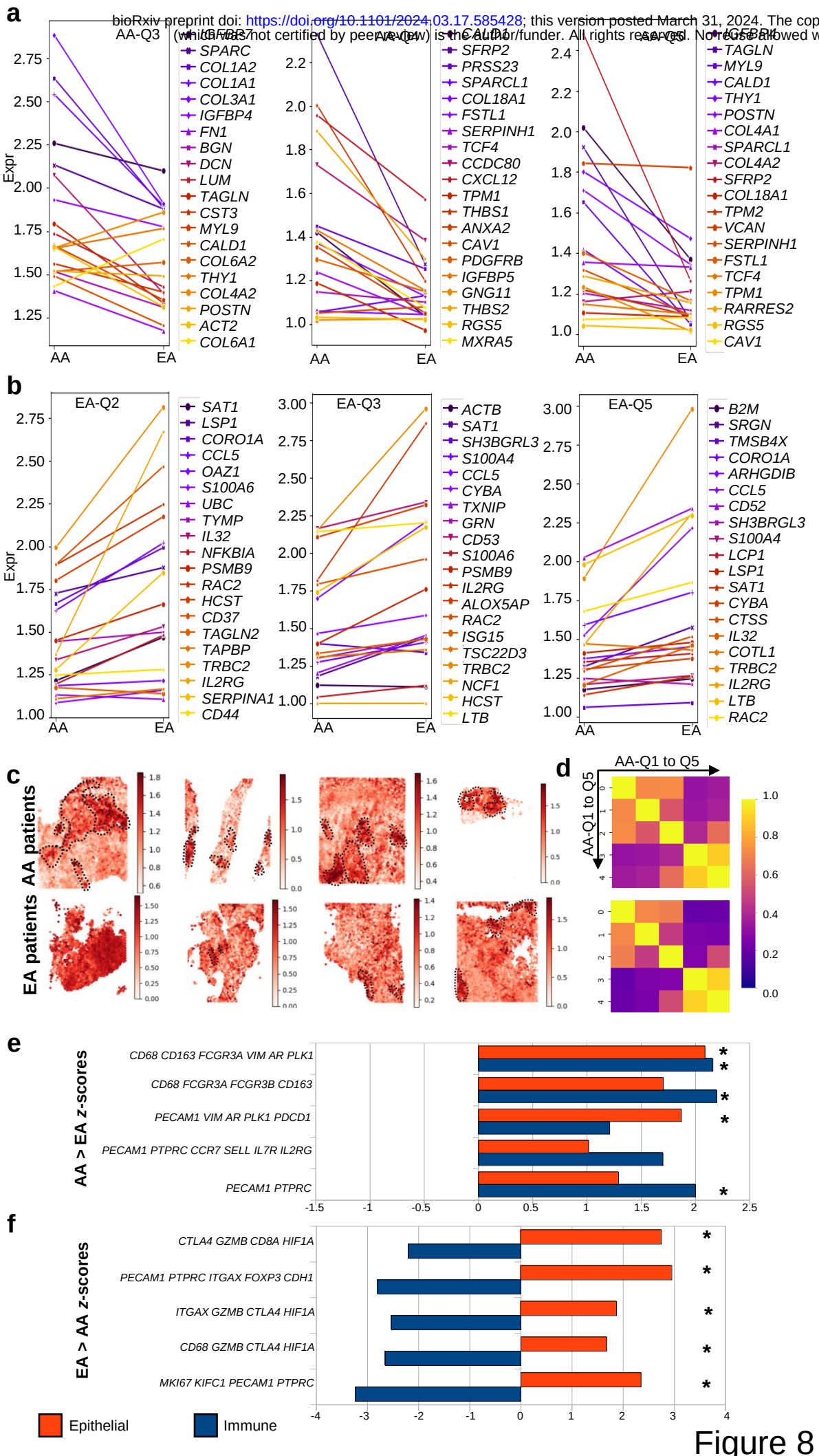
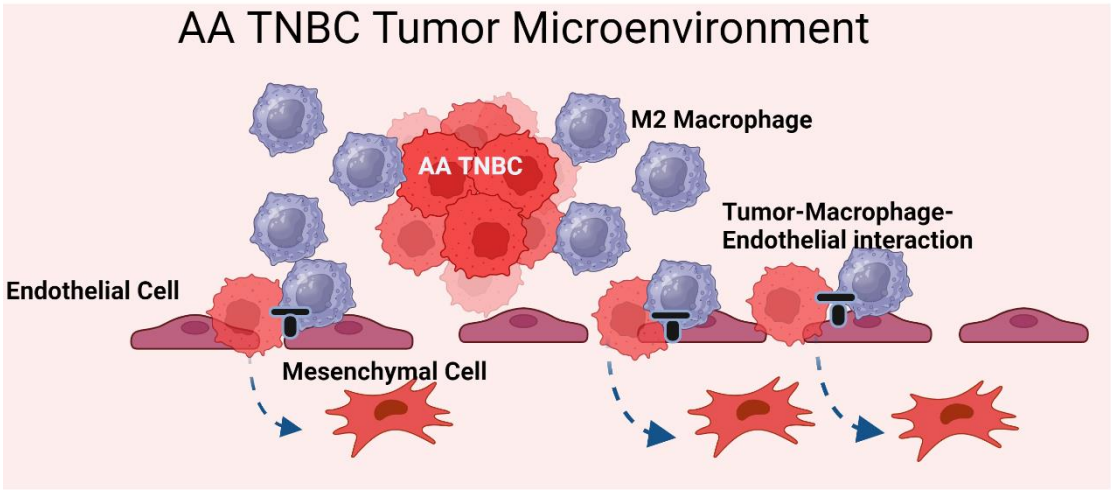


Figure 8

a



Interaction

b