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

Qian Zhu[1,2,3,*,#], Akhila Balasubramanian[4,*], Jaya Ruth Asirvatham[5,*], Megha Chatterjee[3],Badrajee Piyarathna[4], Jaspreet Kaur[6], Nada Mohamed[5], Ling Wu[1], , Stacy Wang[1,2], Niloufar Pourfarrokh[5], Paula Danika Binsol[5], Mahak Bhargava[7], Uttam Rasaily[3], Yitian Xu[8], Junjun Zheng[8], Deborah Jebakumar[5], Arundathi Rao[5], Carolina Gutierrez[1,3], Angela Omilian[9,10], Carl Morrison[9], Gokul M Das[11], Christine Ambrosone[10], Erin H Seeley[12], Shu-hsia Chen[8], Yi Li[1,3,4], Eric Chang[1,3], Xiaoxian Li[13], Elizabeth Baker[14], Ritu Aneja[7], 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]Department of Biology, Georgia State University, Atlanta, Georgia 30302, USA

[7]Department of Nutrition Sciences, School of Health Professions, The University of Alabama at Birmingham, Alabama 35294, USA

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

[9]Department of Pathology and Laboratory Medicine, Roswell Park Comprehensive Cancer Center, Buffalo, NY 14203, USA

[10]Department of Cancer Prevention and Control, Roswell Park Comprehensive Cancer Center, Buffalo, NY 14203, USA

[11]Department of Pharmacology and Therapeutics, Roswell Park Comprehensive Cancer Center, Buffalo, NY 14203, USA

[12]Department of Chemistry, College of Natural Sciences, The University of Texas at Austin, Austin, USA

[13]Department of Pathology and Laboratory Medicine, Emory University, Atlanta, Georgia

30322, USA
[14]Division of Molecular and Translational Biomedicine, Department of Anesthesiology and Perioperative Medicine and Division of Pulmonary, Allergy, and Critical Care Medicine,
The University of Alabama at Birmingham, Alabama 35294, USA

Keywords: triple-negative breast cancer (TNBC), Black American, White American, tumor microenvironment.

Conflict of Interest Disclosure: Jaya Ruth Asirvatham 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.

Abstract

Racial disparities in the clinical outcomes of triple-negative breast cancer (TNBC) have been well-documented, but the underlying biological mechanisms remain poorly understood. To investigate these disparities, we employed a multi-omic approach integrating imaging mass cytometry and spatial transcriptomics to characterize the tumor microenvironment (TME) in self-identified Black American (BA) and White American (WA) TNBC patients. Our analysis revealed that the TME in BA patients is marked by a network of endothelial cells, macrophages, and mesenchymal-like cells, which correlates with reduced patient survival. In contrast, the WA TNBC microenvironment is enriched in T-cells and neutrophils, indicative of T-cell exhaustion and suppressed immune responses. Ligand-receptor and pathway analyses further demonstrated that BA TNBC tumors exhibit a relatively "immune-cold" profile, while WA TNBC tumors display features of an "inflamed" TME, suggesting the evolution of a unique immunosuppressive mechanism. These findings provide insight into racially distinct tumor-promoting and immunosuppressive microenvironments, which may contribute to the observed differences in clinical outcomes among BA and WA TNBC patients.

Statement of Significance

This study identifies distinct tumor microenvironment (TME) profiles in Black and White American TNBC patients, providing new insights into the biological mechanisms driving outcome disparities. Our findings highlight the role of the tumor-endothelial-macrophage niche in these disparities, offering a potential therapeutic target for race-inclusive strategies aimed at improving clinical outcomes. By revealing racial differences in treatment response profiles, this work underscores the necessity for tailored therapies in TNBC. These insights lay the groundwork for the development of inclusive, precision-driven treatment approaches that may help mitigate racial disparities and enhance patient outcomes.

Introduction

A profound racial disparity has been identified between Black American (BA) women and their White American (WA) counterparts with respect to the incidence[1-3] and clinical trajectory[2,3] of breast cancer (BCa). Specifically, BA women develop BCa at a relatively younger age[4], have a two-fold higher chance of developing Triple Negative (TN) tumors[5], with more aggressive disease features (early onset, higher tumor grade, larger tumor size, ER-negative status, distant metastasis)[6]. Additionally, the incidence and mortality of the particularly aggressive triple-negative breast cancer (TNBC) subtype are higher in Black women than in White women[7,8]. Importantly, cancer-related disparities in mortality between BA and WA women was higher for TNBC than other subtypes, even after adjusting for age, stage, treatment, socio-economic status, poverty index, and treatment delay[9,10]. Genetic, environmental, and healthcare access/utilization factors may all contribute to this disparity[11]. Many studies have examined the genetic component to understand the biological underpinnings of racial disparity[12].

Studies examining global discovery of gene expression signatures that distinguish BA vs WA TNBC tumors have reported distinct tumor-associated immunologic profiles in BA patients[13-15] and in patients of African descent[16,17]. 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)[18], which were not defined. Notably, variations in the composition of the TME between BA and WA BCa patients have been reported[19,20], 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[16], while activated mast cells negatively correlated with African ancestry[16]. Further, the level of tumor-associated lymphocytes was found to be similar between BA and WA TNBCs[13]. Controversies regarding the composition of TME and cellular interactions in BA and WA TNBCs could not be resolved due to the limitations of bulk sequencing and the reliance on computational deconvolution. We believe that 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 BA and WA TNBC.

While recent studies employing spatial proteomic[21,22] and transcriptomic[23] technologies have provided insights into the molecular, cellular, and spatial phenotypes governing cancer metastasis[24] and tumor recurrence[25], they have not been extended to the study of BA and WA TNBCs. Previous multi-omics experiments have facilitated the creation of predictors for treatment response and enhanced our understanding of the biological mechanisms behind molecular cancer phenotypes[26-28]. Basic spatial profiling of proteins demonstrated differences between BA and WA BCa patients, despite that these differences were not associated with survival [29]. Moreover, spatial transcriptomics analysis demonstrates that there are racial differences in hypoxic tumor content and regions of immune-rich infiltrates among those with TNBC [30]. 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)[21], multiplexed ion beam imaging (MIBI)[22], and single-cell RNA sequencing[31] have been used to characterize the TME of TNBC, all patients included in these studies were of European descent.

The multifactorial social construct, race, encapsulates both genetic ancestry and the lived environment. Further, we posit that, the experience of racism and being a member of a racialized group differentiates the lived experience of BA and WA women in multiple ways, all of which together result in biological differences. In this study, we examine these biological differences that may be especially pertinent to understanding the poorer outcomes among BA

compared to WA. To date, much of the research examining these biological factors has been done on WA women. An unintended consequence of this is that it can lead to improving treatment options and therapy that are largely developed to address the health of WA. This results in widening disparities in outcomes as we become better at treating BCa in WA, but not BA[32]. Indeed, as BCa treatment and therapy options have increased over the last three decades, resulting in better BCa outcomes overall, there has also been a stark increase in the inequity of BCa outcomes between WA and BA women over this same time period[33]. Understanding the differences in the molecular and cellular interaction in BA women is pertinent to ensuring that racial disparities in outcomes do not widen further. We believe that a spatial multi-omics study of a racially diverse TNBC cohort, as described here, should permit one to understand the biological factors underpinning the large survival gap currently experienced by BA patients with TNBC, while also laying the foundation to develop therapeutic interventions that are inclusive of BA women with BCa. In this spatial multi-omics study, for the first time, we have conducted an integrative analysis combining information obtained from spatial single-cell level IMC[34] and re-analysis of spatial transcriptomics data[30] to uncover racially enriched spatial cell-cell interactions and the larger cell communities that define unique BA and WA TNBC-associated niche. These cell-cell interactions and race-specific niche correlate with patient survival and infer molecular characteristics of WA and BA 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[30], 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 multicellular niches and immunosuppressive mechanisms involving tumor cells, immune cells, and endothelial cells in TNBC that differ between BA and WA. 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

Description of the multi-omic cohort, discovery, and validation datasets

We collected triple-negative breast cancer (TNBC) patient specimens from multiple sources and divided them into discovery and validation cohorts (**Figure 1a**). Imaging mass cytometry (IMC) was performed on a discovery cohort of 57 patients (26 BA and 31 WA) from Baylor Scott and White Hospital (BSW), followed by validation in an independent cohort of 46 patients (15 BA and 31 WA) from Roswell Park Comprehensive Cancer Center (Roswell Park) and an additional 10 patients (5 each BA and WA) from BSW (BSW2). For spatial transcriptomics (ST), 10X Visium ST was conducted on 9 patients (4 BA and 5 WA) from BSW2, and the data were integrated with Visium ST results from 10 patients (6 BA and 4 WA) from a community hospital in Georgia (PI: Dr. Aneja) and 20 patients (10 each BA and WA) from a previously published TNBC ST dataset by Bassiouni et al. (BAS). The BAS dataset was utilized for initial discovery, while the BSW2 and Georgia datasets served for validation of BA and WA specific spatial niche gene signatures. To further support our findings, Nanostring GeoMX DSP was applied to the BSW-Discovery cohort for region-of-interest (ROI) and compartment-based mRNA assessments, providing additional validation for BA and WA-associated niche gene signatures. In total, this data collection effort resulted in spatial immune proteomics data from 113 TNBC patients and spatial gene expression data from 39 TNBC patients, enabling robust identification of spatial niches associated with tumor behavior disparities between BA and WA TNBC patients. **Figure 1b** summarizes the analyses conducted, which include identification of cell-cell interactions and community detection using IMC data, assessment of the prognostic significance of cell communities, spatial localization of cell communities in Visium ST, determination of extended gene signatures (ESG), and ligand-receptor analysis to reveal underlying drivers within the spatial niches.

Overview of IMC Discovery set

We examined a racially balanced and clinically matched cohort of 57 surgically resected TNBC tissue samples, comprising 26 self-reported BA and 31 self-reported WA women (see **Table 1** for clinical details, and **Supplementary Figure 1** for H&E images). Self-reported race provides a more accurate representation of individuals' lived experiences, which may influence the biological factors contributing to cancer-related health disparities. All patients in the cohort had survival outcome data available for

up to 10 years. There was no statistically significant difference in overall survival between BA and WA patients in this cohort (**Table 2**; univariate analysis for race). Additionally, the mutational status of PIK3CA and TP53[8] did not differ significantly between BA and WA patients in our study (**Table 1**).

To investigate the single-cell spatial interaction landscape, we performed IMC on multiple ROIs (see **Supplementary Figure 2**) selected from individual tumors, which were arranged in a tissue microarray (TMA) format (each core with a diameter of 3 mm), using a panel of 26 antibodies targeting immune-regulatory, stromal, and epithelial proteins (overview in **Figure 1c**). ROIs were chosen based on H&E-stained images by a breast pathologist and were subsequently confirmed by a second pathologist. The selected ROIs included areas containing tumor cells intermixed with tumor microenvironment (TME)-related immune and stromal cells. These ROIs were selected from both the tumor center and periphery, categorized as either immune-rich or immune-poor (**Supplementary Figure 3**). An equal number and category of ROIs were chosen from tumors of BA and WA patients (**Table 1**).

We initially segmented over 270,000 single cells based on these ROIs. Through unsupervised clustering with multiple initializations, we identified 16 distinct single-cell types and four multi-cell type clusters based on marker protein expression (**Figure 1d**, clusters 1a to 20a; multi-cell type clusters indicated by asterisk). Differential analysis revealed unique protein expression profiles within each of the 20 clusters, which were named according to co-expressed markers (e.g., Exhausted Cytotoxic T cell, Cluster 7a, characterized by CD152 and CD8a). Six clusters were associated with tumor characteristics (**Figure 1d,** red fonts), while 14 clusters were associated with immune profiles (**Figure 1d**). Immune clusters covered diverse cell types such as exhausted T cells (cluster 7a), helper T-cells (15a), naïve T cells (20a), macrophages (12a, 1a), and dendritic cells (5a and 3a). Tumor compartment covered cells of diverse states, such as proliferative (18a), hypoxic (9a), mesenchymal (4a), and angiogenic (16a) states. The t-SNE plot clearly demarcates the clusters (**Figure 1e**), and **Figure 1f** overlays the expression of E-Cadherin, CD45RA, CD31, CD68, and Vimentin on the t-SNE, highlighting distinct clustering patterns.

Spatial cell-cell interactions, rather than cluster abundances, distinguish BA and WA TNBC.

We analyzed cluster abundances across tumors from patients of both racial groups. While most clusters exhibited similar abundances, several were enriched in specific patient subsets based on race (**Supplementary Figure 4a**). For example, naïve T cells, represented by Cluster 20a (CD31, CD45RA), were more abundant in BA TNBC, whereas exhausted

cytotoxic T cells, identified by Cluster 7a (CD152, CD8a), and were more prevalent in WA TNBC. However, due to the unique marker combinations expressed in each patient's tumor, cluster abundance alone was insufficient to segregate patients based on self-reported race (**Supplementary Figure 4b**).

Given that differences in relative cluster abundances could not fully explain the racial disparities in BA and WA TNBC, we hypothesized that spatial cell-cell interactions might better capture race-associated variations and their clinical implications. To test this, we analyzed cell-cell interactions across 20 clusters using spatial proximity enrichment/depletion analysis with Giotto[35] (see Methods and **Supplementary Figure 5** for details). This analysis involved shuffling cell type labels within each ROI to establish a ROI-specific baseline of interactions, accounting for cellular heterogeneity and abundance differences, and counting interactions against background. Differential interactions between BA and WA patients were identified using a linear mixed model (LMM), which incorporated repeated and hierarchical observations of tumor ROIs from each patient. The resulting landscape of cell-cell interactions, along with specific differential interactions, is shown in **Figures 2a-b**, where positive BA/WA coefficients indicate enriched interactions and negative coefficients indicate depleted interactions in the respective race.

Our analysis revealed six BA- and five WA-specific spatial cell-cell interactions with statistical significance (Padj < 0.05) (**Figures 2b and 2c**). Although cluster abundances did not differ significantly between racial groups (**Supplementary Figure 4b**), spatial interactions highlighted distinctive differences. In BA tumors, key interactions involved naïve T cells—naïve T cells, endo-mesenchymal—tumor cells, and M2 macrophages—mesenchymal cells (**Figure 2b**). In contrast, WA tumors exhibited interactions in a TME, featuring helper T—helper T cells, and myeloid cell—cytotoxic cell, and T cell—cytotoxic cell interactions, with a notable presence of exhaustion marker (CD152) and hypoxia (HIF 1α) (**Figure 2c**). Leave-2-patient-out cross-validation for each race confirmed the robustness of the identified interactions, with consistent coefficients and significance across all subsamples (see 95% CI in **Figures 2b and 2c**). Importantly, when the evaluation was confined solely to the low-grade tumors, highly similar BA/WA coefficients and sets of prioritized spatial cell-cell interactions were obtained, validating the LMM modeling (**Supplementary Figures 6)**.

Figure 2d illustrates the spatial distribution of the top-ranked BA and WA cell-cell interactions across patient samples, with BA tumors showing enrichment in BA-specific interactions. In BA tumors, interactions between M2 macrophages (orange), endothelial cells (black), tumor cells in an epithelial-mesenchymal transition (EMT) state (green), and naïve T cells (blue) are highly clustered. These clustered interactions are not observed in WA tumors. In contrast, WA tumors display increased WA-specific interactions (**Figure 2e**), such as interactions between cytotoxic cells in the hypoxic TME (orange) and exhausted T cells (pink), which are more dispersed compared to the clustered interactions seen in BA tumors. Multiplex immunofluorescence staining in BA TNBC confirmed the endothelial-macrophage interaction (CD31-CD163, **Figure 2f**). Co-localization analysis of CD31 and CD163 (**Figure 2g**) revealed significantly higher co-localization in BA TNBC (n = 27 patients) compared to WA TNBC (n = 44 patients), even after adjusting for cell numbers per image (**Figure 2h**). These findings further validate the spatial cell-cell interactions identified by IMC.

Racially distinct cellular interactions also exist in other IMC cohorts.

To investigate whether similar racially distinct cell-cell interactions are present in other patient cohorts, we performed IMC experiments on TMAs from two additional TNBC patient groups: Roswell Park (see **Table 2** for clinical data, 46 patients) and Baylor Scott and White (BSW2, 10 patients), referred to as validation cohorts. We applied strict quality controls, excluding cores with low cell counts, limited tumor content, or partial cores. Next, we performed cell segmentation, unsupervised single-cell clustering, and spatial proximity analyses to identify BA- and WA-specific cell interactions. Each cohort yielded 20 distinct cell-type clusters, labeled with suffixes "b" (Roswell Park; **Supplementary Fig. 7a**) or "c" (BSW2; **Supplementary Fig. 8a**), respectively.

In Roswell Park, differential BA-specific cell interactions were identified, revealing interactions involved myeloid cells, endothelial cells, proliferating tumor cells, and naïve T cells (**Supplementary Fig. 7b**). Similarly, in BSW2, significant BA interactions featured cells undergoing epithelial-mesenchymal transition (EMT), endothelial cells, macrophages, and naïve T cells (**Supplementary Fig. 8b**), validating findings from the BSW discovery cohort (**Fig. 2b**).

For the WA group, Roswell Park showed WA interactions among exhausted T cells, Tregs, memory T cells, and M2 macrophages within a hypoxic TME (**Supplementary Fig. 7c**). BSW2 demonstrated WA-specific interactions involving exhausted T cells, helper T cells, dendritic cells, and M2 macrophages **(Supplementary Fig. 8c**). These are consistent with the exhausted T cell phenotype in the BSW discovery cohort (**Fig. 2c**). Importantly, despite differences in cluster definitions across the three IMC datasets due to unsupervised clustering, a consistent set of cell types emerged across each race within the respective cohorts. Therefore, we next sought to determine if these recurrent interactions formed higher-order cellular communities and if these communities had prognostic implications.

Cellular interactions analyzed in multiple IMC cohorts reveal existence of recurrent cell community structure that further associate with clinical outcomes in each cohort.

To investigate whether cell-cell interactions extend beyond pairwise relationships to form organized cellular communities within the TME, we expanded the list of differential interactions (with Padj = 0.20) for each cohort and performed community detection. This involved identifying connected components followed by pruning (see Methods). As the number of interactions increased, distinct communities began to emerge, showing a high degree of consistency in the BA- and WA-specific communities across the three independent IMC datasets (BSW discovery, BSW2, and Roswell Park validation). A total of two and four communities respectively, were identified for the BA and WA groups (**Figures 3 a, b** and **Supplementary Figure 9**). The dominant community in BA (termed BA-Community-1) was characterized by recurring interactions involving “M2 Macrophages-Endothelial Cells-Mesenchymal Cells-Naïve T Cells” (**Figure 3a**). In contrast, the top community in WA (termed WA-Community-1) was composed of “Exhausted Cytotoxic T Cells-Helper T Cells-Macrophages-Hypoxia” (**Figure 3b**).

We subsequently assessed the prognostic value of the BA- and WA-specific communities by examining their prevalence within the ROIs and correlating these with overall survival (OS) in patients. As previously reported, race did not significantly influence OS between the two patient groups in both the BSW Discovery and Roswell Park (**Tables 3 and 4,** Race). To evaluate the collective impact of the BA and WA communities, we computed a composite interaction score for the communities by first aggregating the individual cell-cell interaction scores within each BA and WA community for each patient sample, then summed across all identified BA or WA communities to form a *BA* or *WA-community-all* score per patient. Patients were stratified into high- and low-score groups based on the average *community-all* score across patients. The association of these scores with OS was examined using Kaplan-

Meier (KM) curves in both the BSW-Discovery and Roswell-Validation IMC cohorts. Notably, after adjusting for age, race, and stage, higher BA-community scores were linked to worse OS in both the BSW-Discovery and Roswell-Validation cohorts (**Figure 3c**). In contrast, the WA-community score was only prognostic in the Roswell Park-Validation cohort (**Figure 3d**). We next assessed the prognostic value of BA- and WA-associated communities in each race group. BA-associated communities were a significant predictor of OS for BA patients, across both the above-mentioned cohorts (**Supplementary Figures 10 a, c**). WA-communities were a significant predictor of OS for the WA patients in both the cohorts (**Supplementary Figures 10 b, d**). Specifically, the prognostic value of the WA-communities were more pronounced for WA vs BA TNBC patients (Supplementary Figures 10 a, c). Furthermore, combining the BA and WA-interaction scores was no longer prognostic, highlighting the unique prognostic attributes of BA and WA-associated interaction scores (**Supplementary Figures 11 a, b**).

Consistent with these findings, in the multivariate analysis, the BA-associated community remained significantly associated with OS even after adjusting for age, race, and stage in both datasets (**Tables 3 and 4**). In the BSW-Discovery cohort, the BA-associated community also maintained significance when adjusted for BMI (**Table 3**, P = 0.021). Conversely, the WA-associated community was significant in the multivariate analysis only in the Roswell Park-Validation cohort (**Table 4**). However, none of the single-cell clusters were prognostic in BA or WA patients in the BSW Discovery data (**Supplementary Table 1**). Therefore, we conclude that single-cell abundances do not influence race-specific survival outcomes. In contrast, **Figures 3** in our manuscript highlights the association between communities, made of interacting cell clusters, and overall survival in both a race-dependent and independent manner.

Taken together, our analysis demonstrates that the BA-associated community is a robust prognostic marker for overall survival across multiple patient cohorts, independent of clinical factors such as age, race, and stage. This highlights the potential of cell-cell interactions within the BA-community as a key determinant of tumor progression. In contrast, the WA-associated community's prognostic value appears to be cohort-specific, warranting further investigation into its role in specific patient populations. These findings underscore the importance of cellular interactions in shaping the TME and suggest that targeting community-specific interactions may offer novel therapeutic avenues for improving patient outcomes.

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

To extend our analysis beyond the limited coverage of protein markers examined with IMC, we integrated spatial transcriptomic data to gain a more comprehensive understanding of the distinct multicellular niches in BA and WA TNBC. We combined our IMC results with a publicly available, racially balanced spatial transcriptomic dataset from 20 TNBC patients (n = 10 each for BA and WA) derived from flash-frozen tumor samples[30]. Genes encoding proteins associated with the top BA and WA communities identified through IMC were interrogated within this spatial transcriptomic dataset to assess their spatial localization across the 20 TNBC tumors (**Figures 4a, b**). To start, protein markers identified by IMC, such as CD31 and CD45RA, were converted to their corresponding genes, *PECAM1* and *PTPRC*, in the spatial transcriptomic dataset. Then, the genes belonging to each BA-Community 1 or WA-Community 1 (**Figure 4b**) were spatially plotted to reveal their spatial localization patterns in tumors.

Notably, the localized regions demonstrating the co-expression of BA-Community-1 genes (characterizing the Endo-Mesenchymal-Mac-Naïve T niche) are significantly more spatially clustered in BA patients compared to WA patients (**Figure 4c** black outlines, and spatial clustering score **Figure 4d**). Strong spatial clustering observed across 10 BA TNBC patients indicates that this is a recurrent feature in BA tumors. Conversely, WA patients exhibited a random and uniform distribution for these genes (**Figure 4c,** WA), suggesting that endothelial markers and vimentin (mesenchymal marker) are either not co-expressed or expressed at diminished levels in WA patients. Taken together, these findings imply that the mesenchymal cell, endothelial cell, and macrophage form a multi-member multicellular TME niche that is distinctly recurrent across BA TNBC.

In contrast to our observations with the BA query, the localized spots for the WA-Community-1 query genes are uniformly distributed throughout the entire tumors (**Supplementary Figure 12**) rather than confined to any region, suggesting an infiltration of exhausted T cells throughout the tumors. Instead, WA tumors were characterized by a higher presence of WA-Community-1 spots, irrespective of their spatial configuration (**Supplementary Figure 12**). In summary, WA interactions are chiefly characterized by exhaustive immune and hypoxic

environments (illustrated by the diffused co-localization of markers *GZMB, CTLA4, HIF1A*) in WA TNBC.

Niche-specific differential expression analysis reveals additional players associated with BA and WA-tumor associated multicellular niches

To further delineate molecular factors underlying race-associated communities, we extracted ST spots that exhibit high and low expression for each community and performed niche-specific gene signature analysis (**refer Methods**). In this context, spots characterized by elevated expression of BA- or WA-community genes were compared with spots exhibiting low expression within the respective ST samples, to identify co-expressed genes called **extended signature genes** (ESGs). A differential analysis of ESGs between BA and WA tumor further identified race-associated ESGs that could suggest additional cell types in the niches (**Figure 5a**). With these ESGs, we also observed a higher number of spots localized to each respective race-associated community (**Figure 5d**).

As illustrated in **Figure 5b**, ESGs associated with BA-Community-1 query implicated endothelial cells and macrophages, and additionally Cancer Associated Fibroblasts (CAFs). The cell-type specific expression profiles reveal that the expression of the ESG genes are markedly higher in BA than WA tumors within these specific cell type compartments (**Figure 5b**). Notably, CAFs are marked by the higher expression of *SPARC, COL4A1, TAGLN, CALD1, FBLN1, RARRES2, CCDC80, SFRP2,* and *VCAN* in BA patients. Intriguingly, BA-niche ESGs lack markers of B-cells, T-cells, neutrophils, and epithelial cells. Gene set enrichment analysis (GSEA) of the BA-associated ESGs revealed a strong association with processes such as extracellular matrix organization (P=2.3E-28), epithelial-mesenchymal transition (EMT, P= 6.6E-28), tumor vasculature development (P=1.2E-46), and endothelial cell activity (P=1E-18) in BA TNBC (**Figure 5e**).

In contrast, the ESGs associated with the WA-Community 1 (**Figure 5c**) included genes that characterize T cells and additionally neutrophils. GSEA of the WA-associated ESGs highlighted alterations in immune signaling pathways within WA TNBC (**Figure 5e**). These ESGs included *S100A8* and *S100A9*, which are well-established markers for neutrophils, as well as *CD3E, IL2RG*, and *LTB,* which are indicative of T cell populations. Importantly, the WA-niche ESGs did not include markers for CAFs or endothelial cells (**Figure 5c**). Notably,

neutrophils were not evaluated in the initial BSW-discovery IMC panel, as it was devoid of any neutrophil marker; however, it was identified by this ST-derived ESG analysis. Therefore, to corroborate these results, in formulating the IMC antibody panel for the BSW2 and Roswell validation cohorts, we incorporated the neutrophil marker Myeloperoxidase (MPO) to measure its expression. Our results showed that the expression of *MPO* strongly correlated with WA-Community-1. Specifically, *MPO* expression was highly enriched in cluster 9c of the BSW2 dataset and cluster 1b of the Roswell Park IMC dataset (**Supplementary Figures 13 a, b**). The interaction between neutrophils and exhausted T cells is more pronounced in WA vs BA TNBC (**Supplementary Figures 13 c, d**). These findings confirm the presence of neutrophils within the top WA community and confirms the approach of using ESGs to identify additional niche constituents.

Drivers in BA and WA tumor niches, and the predominance of exhausted T cells in WA niches

To investigate the drivers within the racially distinct spatial environments, we used ESGs, as described earlier, to extract and reveal a unique set of ligand-receptor pairs within the BA and WA niches (**Figure 6a**). In the BA niche, most ligand-receptor interactions were linked to processes such as cell-cell communication, platelet-derived growth factor (*PDGF*) signaling, epithelial-mesenchymal transition (EMT), Wnt/Notch signaling, endothelial/extracellular matrix remodeling, vascular endothelial growth factor (*VEGF*) signaling, and integrin-mediated signaling. These interactions primarily involved perivascular-like cells, endothelial cells, CAFs, and macrophages (**Figure 6b**). In contrast, the ligand-receptor interactions in the WA niche were primarily associated with immune cells, as supported by the normalized frequencies of the interactions between neutrophils, T cells, and myeloid cells (**Figure 6c**). Notably, inflammatory cytokines such as *CXCL9, CXCL10,* and *CXCL11* showed prominent interactions with the *CXCR3* receptor. These ligand-receptor interactions may have downstream consequences on the expression of target genes and activities of signaling pathways**.**

To further explore the functional state of T cells within the WA niches, we examined the expression of a T-cell exhaustion signature in the ESGs associated with the WA niche. As shown in **Figure 6d**, the WA-associated niche (WA-Community-1) exhibited the presence of an 11-gene T-cell exhaustion signature, including *TRBC2*[36]*, LAG3*[37]*, HAVCR2*[38]*, and CSF1*[39]**.**

The number of spots expressing these exhaustion genes was significantly higher in WA tumors compared to BA tumors (**Figure 6e**). Collectively, these results underscore the unique TME features in BA and WA TNBC, with the WA TME exhibiting a signature characterized by exhausted T cells, along with neutrophils and myeloid cells, contributing to a distinct immune landscape.

Validation of BA and WA-ESG signatures in independent clinical cohorts

We then validated the BA and WA-associated ESGs described in **Figure 4** by employing a combination of 10X Visium ST and Nanostring GeoMx DSP systems. As previously outlined in **Figure 1,** tumor microarrays (TMAs) were created and Visium ST profiling was conducted on two independent cohorts: BSW2 (9 samples, 5 BA and 4 WA, **Supplementary Figures 14 and 15** for H&E images) and Georgia-Validation (10 samples, 6 BA and 4 WA, **Supplementary Figures 16 and 17** for H&E images). Using both cohorts, we confirmed the significantly higher expression of BA-ESGs, which encode endothelial-macrophage-EMT interactions, in BA tumors, and WA-ESGs, which encode exhausted T cells and neutrophils, in WA tumors (**Figure 7a** for the BSW2 cohort, and **Figure 7b** for the Georgia cohort, highlighting the top BA-Community-1 and WA-Community-1). Additionally, through spatial localization, we observed that BA-community genes spatially clustered strongly within and between themselves in BA tumors, whereas no such clustering was seen in WA tumors in both validation cohorts (**Figure 7c**).

Nanostring GeoMx evaluation of 26 BA and 31 WA TNBC tumors, employed for imaging mass cytometry in **Figure 1** (refer to **Supplementary Figure 2** for indicated regions of interest), further substantiated the increased expression of ESGs linked with the BA niche in BA tumors and the WA niche in WA TNBC (**Figure 7d**). Importantly, the immune compartment (CD45-positive) demonstrated a notable increase in ESG expression (Figure 7d), while the epithelial compartment (PanCK-positive) did not exhibit such an increase (**Supplementary Figure 18**), highlighting the compartment-specific characteristics of the BA and WA cellular communities. Collectively, these findings offer insights into the distinct spatial ecosystems underlying WA-specific immunosuppression.

Discussion

In this study, we present the first integrated multi-omics analysis examining the distinct tumor microenvironments (TMEs) of BA and WA patients with TNBC. Using advanced spatial technologies such as IMC and spatial transcriptomics, we identified race-associated spatially resolved cell–cell interactions, community structures, and niche signatures (i.e., ESGs) in BA and WA TNBC tumors. Importantly, these findings were reproducible across three independent IMC datasets comprising over 100 patients. This reproducibility reinforces the robustness and generalizability of our observations, providing confidence that the distinct TMEs we describe are not limited to a single cohort or dataset. Additionally, we validated our findings using two independent cohorts, BSW2 and Georgia, with the 10X Visium and Nanostring GeoMx Digital Spatial Profiling (DSP) technologies.

Our study builds upon recent studies that have highlighted differences in TNBC between BA and WA patients[16,20,29,40-44], often relying on bulk transcriptomics and proteomics analyses. These studies have uncovered distinct molecular features in TNBC tumors from these two racial groups, but they often failed to provide insights at the single-cell level. The bulk approach cannot resolve the complexities of the TME, nor can it account for the spatial relationships between various cell types. Our study addresses this gap by using single-cell proteomics to map the cellular architecture and interactions within the TME of both BA and WA TNBC, revealing how these spatially organized cell communities may drive tumor progression and influence clinical outcomes.

Our spatial analysis identified distinct racial differences in TNBC tumors. Although cluster abundances were similar between BA and WA patients, spatial distributions and interactions differed significantly. BA tumors showed tightly clustered, immune-suppressive communities involving naïve T cells, dendritic cells, and M2 macrophages within hypoxic regions, promoting immune tolerance, EMT, and potentially metastasis[45]. Conversely, WA tumors displayed communities pointing to interactions of exhausted T cells and myeloid cells, indicative of chronic immune suppression. This type of TME may hinder effective anti-tumor immunity and contribute to poor treatment response. Overall, our study emphasizes that the organization and composition of cellular communities in the TME are key drivers of racial disparities in TNBC and may serve as targets for tailored therapeutic interventions.

While IMC-derived single-cell interactions provide valuable biological insights, its capacity is limited by the number of measurable markers, restricting the range of cell–cell interactions that can be examined. To overcome this constraint, we utilized protein markers indicative of race-associated communities as "queries" and integrated them with spatial transcriptomic data to reconstruct the multicellular niche structure, architecture, and composition in BA and WA TNBC tumors. This strategy enabled us to extend the scope of race-associated cell-cell interactions and gain deeper insights into the additional cell types and their interactions within the TME. Through this approach, we identified race-associated ESGs, which uncovered previously unrecognized aspects of the TME in both BA and WA tumors—insights that were not apparent through prior global clustering methods. These findings align with earlier research emphasizing the advantages of niche-specific analysis for understanding complex tumor ecosystems[46].

The BA-associated ESGs emphasized the significant role of cancer-associated fibroblasts (CAFs) and were enriched in pathways promoting epithelial-mesenchymal transition, vascular development, TGF-β, Wnt/Notch1 signaling, and extracellular matrix (ECM) organization. The ESGs further revealed that BA tumors had niches rich in macrophages, accompanied by suppressed CD8+ T cell function, which aligns with immune suppression driven by TGF-β through ECM remodeling[47], and a reduction in metabolites such as arginine, along with the secretion of ornithine by tumor-associated macrophages[47] that suppress T cell functions.
Additionally, vascular development genes (*VEGF* and *thrombospondin*[48]) suggest potential resistance mechanisms to anti-PD1 treatments, warranting further investigation. Notably, a metastatic melanoma study identified angiogenesis-related genes as contributors to immunotherapy resistance[49,50]. Conversely, WA-associated ESGs showed enrichment of exhausted T cell and neutrophil markers, paralleling chemotherapy-induced immune suppression in ovarian cancer. In ovarian cancer, treatment with chemotherapy instills a spatially exhausted T cell environment[51]. This indicates WA patients may be less responsive to chemotherapy, highlighting distinct immunosuppressive mechanisms in WA TNBC.

Although TNBC lacks the expression of the three hormonal receptors, recent studies have identified additional subtypes of TNBC. These include basal-like immune-suppressed, immunomodulatory, luminal androgen receptor (LAR), and mesenchymal-like subtypes[52]. In addition, Hammerl et al[53] found that TNBC tissues could be divided into three phenotypes:

excluded, ignored, and inflamed. This finding builds upon the work of Gruosso et al[54], which showed the presence of immune-cold, immune-desert, fully inflamed, stroma-restricted, and margin-restricted TNBC tumors. Notably, these studies involved single-cell RNA sequencing or relied on the localization and expression of a few markers (e.g., CD8) to describe phenotypes. Additionally, these studies did not assess the TNBC subtypes in racially diverse groups.

Our data, as shown in **Supplementary Figure 19**, suggests that WA tumors resemble the inflamed spatial phenotype ($P=8.9 \times 10^{-17}$) whereas BA tumors tend to align more closely with the Excluded ($P = 0.0009$) and Ignored phenotypes ($P=1.6 \times 10^{-7}$)[53]. This further supports the notion that tumors in BA patients are likely to be immunologically cold, and therefore resistant to immunotherapy. Consistent with this, the BA-Community-1 identified in our study overlaps with the LAR and mesenchymal-like subtypes described by Jiang et al[52]. The presence of androgen receptor in BA-Community-1 further supports its resemblance to the LAR subtype. Additionally, the TME in BA tumors also exhibits features of the margin-restricted (showing exclusion of T cells) and immune-desert subtypes described by Gruosso et al[54]. In contrast, WA-associated community-1 resembles the immunomodulatory subtype described by Jiang et al. and the fully inflamed subtype described by Hammerl et al[53]. Our results also highlight the importance of spatial datasets in understanding the function of immune cell types, as the cell type abundances or deconvoluted cell proportions from bulk RNAseq studies may underestimate the complexity of tumor microenvironments in BA and WA patients. Previous findings regarding which race has more T cells and Treg cells and what are their implications would likely need to be reassessed with the spatial technologies. The new spatial immunophenotypes described by Gruosso et al and Hammerl et al are a step towards this direction. Our work further expanded this knowledge by including additional cell-cell interactions and community network as features of race-specific TNBC tumors.

The implications of our finding that unique spatial niches exist within the TME of tumors from BA and WA patients are significant and multifaceted. For the first time, our study revealed architectural differences within the TME, which may influence both tumor progression and response to treatment. These findings provide preliminary evidence supporting the potential development of therapeutic strategies targeting the endothelial-macrophage-EMT axis, which may be especially beneficial for BA patients with TNBC. The identification of ligand-receptor

interactions that regulate this axis—such as integrin-TGF-β[55], Wnt/Notch1, and VEGF signaling—highlights several novel, actionable targets[55]. For example, clinical-grade Wnt/Notch1 inhibitors, which affect endothelial cells[56] and EMT[57,58], could be a promising avenue for BA patients with TNBC. Additionally, therapeutic regimens targeting macrophages have shown promise in preclinical models of TNBC[59]. The identification of a neutrophil-rich TME co-existing with exhausted T cells in WA tumors is particularly intriguing and was confirmed using myeloperoxidase as a neutrophil marker in our validation cohort. Recent studies suggest that neutrophils can promote tumor progression[60] and may do so independently of macrophages[61], exerting an immunosuppressive function[61,62]. These findings, in conjunction with previous studies [61,62], suggest that WA patients with TNBC have inflamed tumors that have evolved distinctive immunosuppressive mechanisms, which may provide further insight into resistance to current therapies.

Building on the above discussion of the TME biology associated with racial differences, our findings suggest that race-specific cell-cell interactions have important implications for survival outcomes in TNBC. While BA-tumor associated communities contributed to poor OS in both BA and WA populations, the impact of WA-tumor associated communities on OS was more pronounced in WA patients, and not in BA patients. Specifically, WA-Community-1 was associated with poor OS in WA patients in the Roswell cohort, further supporting the notion that racial differences in TME interactions may have differential prognostic significance. Importantly, these associations remained significant after adjusting for key clinical factors such as age, race, and stage, highlighting the robustness of these findings.

These observations underscore the potential of race-associated cell-cell interactions as strong prognostic markers of survival in TNBC. The ability to identify specific cellular communities and their interactions within the TME offers new avenues for understanding how these interactions influence disease progression in a race-dependent and independent manner. Our results are in line with recent research that demonstrates the role of race and ethnicity in shaping immunotherapy responses in breast cancer patients[13], further validating the significance of these findings for the development of personalized therapies.

It is crucial to recognize that while our study focuses on the biological and clinical aspects of race in TNBC, race itself is a social construct deeply intertwined with structural inequities. These inequities, including limited access to healthcare, lack of culturally appropriate care,

poor built environments, and higher levels of chronic stress, disproportionately affect minority populations and contribute to poorer health outcomes. Moreover, health inequities extend beyond race and are prevalent across various groups experiencing social stratification, which compounds the challenges faced by marginalized populations. While advancements in medical treatment and prevention have reduced overall mortality rates, structural inequalities have led to a widening gap in health outcomes for disadvantaged groups[32]. Our research suggests that the unique TME observed in BA patients' tumors, characterized by endothelial-macrophage-EMT axis alterations, could be amenable to targeted therapies that are currently under investigation. By conducting racially inclusive research, we can better understand how different social determinants shape the biology of cancer, thus paving the way for more equitable therapeutic strategies. Ultimately, addressing these inequities through comprehensive research can reduce the disparities in breast cancer outcomes, ensuring that health improvements do not leave behind underserved populations.

In conclusion, our study provides compelling evidence of racial differences in the spatial interactions and associated gene signatures within the TME of TNBC patients, as illustrated in **Figure 8**. The characterization of the TME of BA patients shares significant similarities with the previously described "Tumor Microenvironment of Metastasis (TMEM) doorway" model[63-66]. In the BA TME (**Figure 8a**), perivascular macrophages, mesenchymal cancer cells, and endothelial cells form a tightly connected network that promotes endothelial cell junction dissociation, thereby creating vascular openings that enable cancer cells to intravasate. These unique spatial interactions in BA tumors are strongly correlated with poor survival outcomes in both racial groups, underscoring their clinical relevance. In contrast, the WA TME (**Figure 8b**) is characterized by distinct niches enriched in exhausted T cells and neutrophils. This finding raises important questions regarding the potential tumor-promoting roles of neutrophils, which require further investigation. Collectively, these findings emphasize the importance of incorporating racially diverse samples in TME research to fully appreciate the complex biological and clinical implications of racial differences in TNBC. Understanding these differences could pave the way for more personalized and effective therapeutic strategies that address the unique needs of patients from diverse racial backgrounds.

Methods:

TNBC patient tissue microarrays: De-identified breast cancer patient samples with at least 10-year follow-up and self-reported race, were obtained from Baylor Scott and White hospital (BSWH), Temple, Texas in the form of formalin fixed paraffin embedded (FFPE) tissue. Tissue cores from 57 tumors (26 BA tumors and 31 WA tumors) spread across 8 tissue microarrays (TMAs) were used for this study. In addition to the self-reported race, clinical information such as body mass index (BMI), receptor status, tumor grade, chemotherapy status, presence of metastases, time to clinical follow up, PIK3Ca and TP53 mutation status, and information on Diabetes Miletus (DM), as well as detailed histopathology of the tumors were available to us (refer **Table 1** for clinical information and **Supplementary Figure 1** for H&E images). 2 BA and 2 WA patients received neoadjuvant chemotherapy, the remaining patients were treatment naïve at the time of sample collection. All tumors were collected at surgery. Following this, in the adjuvant setting, all the patients were treated at the same hospital using the standard of care regimens. None of them received immune check point therapy. Two patients were enrolled in a Southwest Oncology Group (SWOG) clinical trial. These samples obtained from Baylor Scott and White hospital were used for generating BSW-Discovery data and Nanostring Geo Mx spatial transcriptomics validation data.
Additionally, for validation studies using IMC, five FFPE TNBC tissues each from BA and WA were obtained from Baylor Scott and White and the data from these samples were termed BSW2 validation. Also, 46 FFPE TNBC tissues (15 BA and 31 WA) from Roswell Park Comprehensive Cancer Center were used for IMC-based validation.

For 10X Spatial Transcriptomics, we examined five WA and four BA TNBC samples described above from Baylor Scott and White (BSW2 validation) hospital. In addition, 10X spatial transcriptomics data for four WA and six BA samples from the Georgia cohort was obtained in collaboration with Dr. Aneja from University of Alabama at Birmingham. All the human studies were performed under IRB (Institutional Review Board) protocols 020-393 and 130559 (Baylor Scott and White Hospital), H-28445 (Baylor College of Medicine), and 300009407 (University of Alabama at Birmingham).

Region of Interest (ROI) selection and segmentation: For each patient core (in the BSW Discovery cohort, 3 mm diameter) 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 performed by Dr. Asirvatham, breast pathologist and independently verified by Dr. Danika (resident pathologist, **Supplementary Figure 2**). A total of 98 ROIs (each 600um-by-600um) 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). **Table 1** describes the numbers of ROIs belonging to each of these four groups across all the specimens analyzed. Importantly, the number of TCIR, TCIP, TPIP and TPIR analyzed were comparable across tumors from both BA and WA patients (**Table 1**). Furthermore, H&E-based stratification of tumor-associated ROIs into immune rich and immune poor regions was confirmed by quantification of CD45 + immune cells (**Supplementary Figure 3**). Notably, IMC and Nanostring GeoMx analysis were conducted on adjacent TMA sections containing identically marked ROIs.

The following guidelines were used to select the ROIs. TMA cores were at least 3 mm in diameter. Tumor periphery was defined as within 1mm of tumor perimeter when identifiable. Tumor center was defined as greater than 1mm from the tumor perimeter. Stromal tumor infiltrating lymphocytes were determined using the International TILs Working Group 2014 guidelines[67] on H&E slides, following completion of available tutorials at tilsinbreastcancer.org. Less than 10% was considered immune poor and greater than 50% was considered immune rich. A single pathologist (JRA) selected all ROIs at a single sitting. These ROI were verified by a second reviewer (PBD) with 100% agreement.

Imaging mass cytometry (Mass CyTOF), BSW-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. In terms of distribution of ROIs among patients in this BSW-Discovery cohort, a total of 98 ROIs were examined that included 47 ROIs from 26 BA patients, and 51 ROIs from 31 WA patients. This included one BA patient with 3 ROIs, 19 BA patients with 2 ROIs and 6 BA patients with 1 ROI. For WA, we had 20 patients with 2 ROIs and 11 patients with 1 ROI.

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[21,34,68]. 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 2**. 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), BSW-2 and Roswell Park Validation data sets: Here 5 WA and 5 BA TNBC FFPE tumors (whole sections) were analyzed from BSW-2 validation cohort and 15 BA and 31 WA TNBC FFPE tumors (1 mm core size) were analyzed from Roswell Park validation cohort (Summarized clinical data in **Table 2**). Regions of Interest (ROIs) were selected by Drs. Asirvatham and Gutierrez respectively, for the two cohorts. In the BSW-2 validation cohort, ROI size was 700um-by-700um. For the Roswell Park validation cohort, ROI size was 600um-by-600um and each ROI covered the entire 1mm core. The IMC analyses were 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 [69]. A total of 27 and 32 metal-tagged antibodies respectively, were used to examine the BSW-2 and Roswell Park validation samples (refer **Supplementary Tables 3 and 4**) 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 9 WA and 10 BA TNBC FFPE tumors from two independent cohorts were analyzed (4 BA and 5 WA from BSW2 and 6 BA and 4 WA from Georgia cohort, See **Supplementary Figures 14-17** for H&E images and ROI markings). The 10X spatial transcriptomics analyses were conducted by the Immune Monitoring Core at the Houston Methodist Research Institute for the BSW2 cohort, and the Advanced Genomics Core at University of Michigan for the Georgia cohort.

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): NanoString platform was used to profile the patient samples in the BSW-Discovery cohort, the same patient cohort used in BSW-Discovery IMC 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. **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[70]. 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 (and confirmed by an independent pathologist Dr. Binsol) 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 BA and 31 WA 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 BA TNBC. The samples used were the BSW-Discovery patient cohort, the same used to generate BSW-Discovery IMC data and Nanostring GeoMx DSP data. 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[71]. 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 was quantified per image. This was repeated for all the 103 immunofluorescent images of WA and BA 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[68], CellProfiler[72], and CellPose[73]. We started by applying a Gaussian Blur in ImageJ[74] 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[75] 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 in IMC: 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[35]. 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[22] in the Giotto paper[35]. Next, we conducted group-wise comparison to extract racial group-associated cell-cell interactions (CCIs). For this analysis, race groups were self-reported Black American (BA) or White American (WA) TNBC patients. We used a mixed linear statistical model to find racial group associated 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[76]. 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[76]. 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 BA or WA, and patient is the patient identifier of ROI. Multiple ROIs are associated with each patient (i.e. making patient the random effect), whereas race is the main (or fixed) effect. The alternative null model has the design function "score ~ 1 + (1|patient)" which excludes the race associated effect. ANOVA testing was performed between these two models to select only the interaction pairs whose scores are most affected by race.

***Cell Community detection in IMC*:** Racially distinct cell-cell interactions that pass FDR cut off 0.20, with a positive BA-coefficient (for BA-specific CCI) or positive WA-coefficient (for WA-specific CCI) in the linear model, were selected to form cell communities. For the selected race-specific CCIs, X, consisting of interactions a—b, where a, b are cell type clusters, we form a graph G where nodes are the clusters, and the edges connect two clusters if there is a CCI between them. We next obtain connected components of G, which will result in an initial set of cell communities. Each community will undergo pruning steps: If a community consists of just one node with self-interaction, it is removed. If a community consists of weak "bridging" edges between sub-communities, these bridging edges are removed, and the resultant sub-communities will become communities. Such bridging edges are usually apparent from the topology of the community graph, since there is usually one bridging edge connecting two sub-communities (if sub-community structure exists within the community).

***Community matching between IMC datasets*:** Cell community matching is possible between the 3 IMC datasets utilized in this study, since all datasets use same or similar protein antibody panel with >90% of antibodies shared across datasets. Thus, to match cell communities that are found independently from each IMC dataset, each community is first reduced to a set of member cell clusters and further reduced to these clusters' differentially expressed protein markers. We next performed matching at the protein marker levels. The two communities (one in each IMC dataset) with the highest number of protein marker matches are declared a matching community. The top matched community are termed BA-Community1 and WA-Community1 respectively for each race. In the special case where a community in dataset 1 looks to be a union of two communities in dataset 2 (which would produce an equally likely match to each of two communities), we will go to dataset 3 to decide whether or not to split the community in dataset 1 or join the two smaller communities in dataset 2 in matching with dataset 1. BA- and WA-Community1 represents the top matching community across the 3 IMC datasets. Weaker scoring matches are termed Community2, 3, etc. Dataset-specific communities are ranked last.

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[30] and in our own TNBC cohorts (used for validation) to generate normalized gene expression matrices per ST sample with Giotto[35].

Targeted analysis of spatial transcriptomic datasets

Query co-localization in ST dataset: Each community was converted from protein names to gene names. The list of gene names in the community forms a race-specific-query, such as BA-community 1. 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 BA and WA groups.

Spatial Clustering: 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[77] 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 (i.e., BA-Community-1):

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

This silhouette coefficient[78] 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[79] 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 then compared between BA and WA patient groups.

Derivation of Extended Signature Genes of BA- and WA-associated communities: To derive extended signature genes (ESGs) of BA- and WA- race-associated communities, for each community query (say BA-Community-1), we stratified the Visium spots per ST sample into interact-high and interact-low groups based on QCS score of that community query per spot. The community-high group consists of Visium spots in the top 10% of spots with the highest QCS. The community-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 community-high and community-low spots within the ST sample, forming Extended Signature Genes of the query. Recurrent BA-associated ESGs were next derived by requiring ESGs to be present in at least 50% of BA patient samples and the percentage of BA patients with ESGs as signature must be higher than the percentage in WA patients.

Cell-type specific expression of extended signature genes of BA- and WA-associated niches: We integrated single-cell RNA sequencing data of breast cancer atlas[80] with spatial transcriptomics to derive cell-type specific gene expression profiles in BA- and WA-interaction-targeted niches. The cell-type specific expression of gene g in cell type c, in a niche N that is defined by WA or BA-Community, 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 community-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[80], and Wigerblad et al[81], and Alvarez-Breckenridge et al[82], which provided the neutrophil subsets.

***Ligand-receptor pair analysis*:** We download all human ligand-receptor pairs from CellPhoneDB database[83]. We next computed enrichment of ligand-receptor pairs within ESGs of BA- and WA-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 BA-Community-1 and WA-Community-1. Ligand-receptor pair results were next summarized into

a cell-type interaction network as shown in **Figure 6a** 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[53] supplementary material. To compare these existing phenotype signatures with our WA-niche associated 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[84] web server to compute gene-set enrichment statistics against the gene-set database[85] 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 BA- and WA-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-associated niche genes in TNBC BA and WA patients (see **Supplementary Figures 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 BA and WA samples, and t-test statistics (testing WA>BA or BA>WA 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 BA patients or higher in WA patients in two independent validation cohorts (19 TNBC, 10 BA and 9 WA, see **Supplementary Figures 14-17**). Specifically, for each community query BA-Community-1 and WA-Community-1, we extracted community-high and -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 high and low spots that were defined and compared this log-fold change in BA samples and WA 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. In BSW-Discovery and Roswell Park cohorts, survival analysis was done independently per cohort based on IMC-derived community profiles observed in patient ROIs. Specifically for each patient, we derive a *per-community-score* as follows: we compute the sum of interaction scores of the constituent cell-cell interactions of the community, deriving a community score. Then we sum the community scores of all BA-identified communities or WA-identified communities, deriving a *community-all* score per BA-communities or WA-communities. Patients are next stratified based on whether the *community-all* score of each patient is higher or lower than the average across all patients. Kaplan Meier curves were next created using the survival library in R. We further examined the association between BA- and WA-associated communities with overall survival in a univariate and multivariate setting using Cox Proportional Hazard Analysis (**Table 2**). In the latter, we controlled for age, race and stage.

Cox Proportional Hazard Analysis: We are interested in associating the BA-community and WA-community scores with patients' survival in the BSW-Discovery and Roswell Park Validation cohorts. Each cohort has documented patient attributes such as age, stage, race, BMI, and mutation status, in addition to 10 year survival status. We built both univariate models to examine the association of BA-/WA-Community scores to survival, and multivariate models to examine these associations while accounting for covariates such as age, stage, and race. In the univariate setting, the association was performed by the coxph function in R by setting "Surv(time,status)~BAscore" where BAscore is the BA-Community score for

individual patient. In the multivariate setting, the association was made by setting “Surv(time, status)~BAscore+bmi+age+stage+race”. This examines the association of BAscore while accounting for other covariates. Both the aforementioned simple model and the race-stratified Cox model were attempted (e.g. “Surv(time, status)~BAscore+bmi+age+stage+strata(race)”) and similar coefficients and P values were obtained.

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 WA and BA groups.

Contributions:

AB, MC, JK, WL, MB, 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, JRA and PDB marked the ROIs for IMC, 10X ST and Nanostring GeoMx DSP analysis for the BSW discovery, BSW2 validation and Georgia validation, CG marked the ROIs for IMC for the Roswell Park validation, NM, NP compiled the clinical data for the TNBC samples, JRA, S-hC, XL, AO, CM, GMD, CA provided reagents and resources, YL, EC, EB 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 research was also supported by NCI grant P30CA016056 (Cancer Center Support Grant Development Funds) to the Biorepository and Laboratory Services Shared Resource at the Roswell Park Comprehensive Cancer Center, and P30CA013148 to the O'Neal Comprehensive Cancer Center at University of Alabama at Birmingham. The authors acknowledge Callee Arnold, Histo-technologist at BSWH for technical expertise. Editorial assistance was provided by Christos Evangelou, PhD.

References

1 Kohler, B. A. *et al.* Annual Report to the Nation on the Status of Cancer, 1975-2011, Featuring Incidence of Breast Cancer Subtypes by Race/Ethnicity, Poverty, and State. *J Natl Cancer Inst* **107**, djv048 (2015). https://doi.org/10.1093/jnci/djv048

2 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)* **10** (2018). https://doi.org/10.3390/cancers10120514

3 Dietze, E. C., Sistrunk, C., Miranda-Carboni, G., O'Regan, R. & Seewaldt, V. L. Triple-negative breast cancer in African-American women: disparities versus biology. *Nat Rev Cancer* **15**, 248-254 (2015). https://doi.org/10.1038/nrc3896

4 Singh, M. *et al.* Distinct breast cancer subtypes in women with early-onset disease across races. *Am J Cancer Res* **4**, 337-352 (2014).

5 Amirikia, K. C., Mills, P., Bush, J. & Newman, L. A. Higher population-based incidence rates of triple-negative breast cancer among young African-American women : Implications for breast cancer screening recommendations. *Cancer* **117**, 2747-2753 (2011). https://doi.org/10.1002/cncr.25862

6 DeSantis, C. E. *et al.* Breast cancer statistics, 2019. *CA: a cancer journal for clinicians* **69**, 438-451 (2019).

7 Kanaan, Y. M. *et al.* Metabolic profile of triple-negative breast cancer in African-American women reveals potential biomarkers of aggressive disease. *Cancer Genomics Proteomics* **11**, 279-294 (2014).

8 Prakash, O. *et al.* Racial Disparities in Triple Negative Breast Cancer: A Review of the Role of Biologic and Non-biologic Factors. *Front Public Health* **8**, 576964 (2020). https://doi.org/10.3389/fpubh.2020.576964

9 Parise, C. A. & Caggiano, V. Disparities in race/ethnicity and socioeconomic status: risk of mortality of breast cancer patients in the California Cancer Registry, 2000-2010. *BMC Cancer* **13**, 449 (2013). https://doi.org/10.1186/1471-2407-13-449

10 Tannenbaum, S. L., Koru-Sengul, T., Miao, F. & Byrne, M. M. Disparities in survival after female breast cancer diagnosis: a population-based study. *Cancer Causes Control* **24**, 1705-1715 (2013). https://doi.org/10.1007/s10552-013-0246-5

11 Ademuyiwa, F. O. *et al.* Impact of body mass index on clinical outcomes in triple-negative breast cancer. *Cancer* **117**, 4132-4140 (2011). https://doi.org/10.1002/cncr.26019

12 Keenan, T. *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* **33**, 3621-3627 (2015). https://doi.org/10.1200/JCO.2015.62.2126

13 Newman, L. A. *et al.* Tumor-Associated Lymphocytes and Breast Cancer Survival in Black and White Women. *JAMA Surgery* (2024). https://doi.org/10.1001/jamasurg.2023.8024

14 O'Meara, T. *et al.* Immune microenvironment of triple-negative breast cancer in African-American and Caucasian women. *Breast Cancer Res Treat* **175**, 247-259 (2019). https://doi.org/10.1007/s10549-019-05156-5

15 Yao, S. *et al.* Breast Tumor Microenvironment in Black Women: A Distinct Signature of CD8+ T-Cell Exhaustion. *J Natl Cancer Inst* **113**, 1036-1043 (2021). https://doi.org/10.1093/jnci/djaa215

16 Martini, R. *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* **12**, 2530-2551 (2022). https://doi.org/10.1158/2159-8290.CD-22-0138

17 Pitt, J. J. *et al.* Characterization of Nigerian breast cancer reveals prevalent homologous recombination deficiency and aggressive molecular features. *Nat Commun* **9**, 4181 (2018). https://doi.org/10.1038/s41467-018-06616-0

18 Thorsson, V. *et al.* The Immune Landscape of Cancer. *Immunity* **48**, 812-830 e814 (2018). https://doi.org/10.1016/j.immuni.2018.03.023
19 Charan, M. *et al.* Molecular and Cellular Factors Associated with Racial Disparity in Breast Cancer. *Int J Mol Sci* **21** (2020). https://doi.org/10.3390/ijms21165936
20 Martin, D. N. *et al.* Differences in the tumor microenvironment between African-American and European-American breast cancer patients. *PLoS One* **4**, e4531 (2009). https://doi.org/10.1371/journal.pone.0004531
21 Ali, H. R. *et al.* Imaging mass cytometry and multiplatform genomics define the phenogenomic landscape of breast cancer. *Nat Cancer* **1**, 163-175 (2020). https://doi.org/10.1038/s43018-020-0026-6
22 Keren, L. *et al.* A Structured Tumor-Immune Microenvironment in Triple Negative Breast Cancer Revealed by Multiplexed Ion Beam Imaging. *Cell* **174**, 1373-1387 e1319 (2018). https://doi.org/10.1016/j.cell.2018.08.039
23 Stahl, P. L. *et al.* Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* **353**, 78-82 (2016). https://doi.org/10.1126/science.aaf2403
24 Risom, T. *et al.* Transition to invasive breast cancer is associated with progressive changes in the structure and composition of tumor stroma. *Cell* **185**, 299-310 e218 (2022). https://doi.org/10.1016/j.cell.2021.12.023
25 Uttam, S. *et al.* Spatial domain analysis predicts risk of colorectal cancer recurrence and infers associated tumor microenvironment networks. *Nat Commun* **11**, 3515 (2020). https://doi.org/10.1038/s41467-020-17083-x
26 Sammut, S.-J. *et al.* Multi-omic machine learning predictor of breast cancer therapy response. *Nature* **601**, 623-629 (2022). https://doi.org/10.1038/s41586-021-04278-5
27 Krassowski, M., Das, V., Sahu, S. K. & Misra, B. B. State of the Field in Multi-Omics Research: From Computational Needs to Data Mining and Sharing. *Front Genet* **11**, 610798 (2020). https://doi.org/10.3389/fgene.2020.610798
28 Zhang, B. *et al.* Proteogenomic characterization of human colon and rectal cancer. *Nature* **513**, 382-387 (2014). https://doi.org/10.1038/nature13438
29 Omilian, A. R. *et al.* Multiplexed digital spatial profiling of invasive breast tumors from Black and White women. *Mol Oncol* **16**, 54-68 (2022). https://doi.org/10.1002/1878-0261.13017
30 Bassiouni, R. *et al.* Spatial Transcriptomic Analysis of a Diverse Patient Cohort Reveals a Conserved Architecture in Triple-Negative Breast Cancer. *Cancer Res* **83**, 34-48 (2023). https://doi.org/10.1158/0008-5472.CAN-22-2682
31 Tietscher, S. *et al.* A comprehensive single-cell map of T cell exhaustion-associated immune environments in human breast cancer. *Nat Commun* **14**, 98 (2023). https://doi.org/10.1038/s41467-022-35238-w
32 Phelan, J. C. & Link, B. G. Controlling disease and creating disparities: a fundamental cause perspective. *J Gerontol B Psychol Sci Soc Sci* **60 Spec No 2**, 27-33 (2005). https://doi.org/10.1093/geronb/60.special_issue_2.s27
33 Giaquinto, A. N. *et al.* Breast Cancer Statistics, 2022. *CA Cancer J Clin* **72**, 524-541 (2022). https://doi.org/10.3322/caac.21754
34 Giesen, C. *et al.* Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nat Methods* **11**, 417-422 (2014). https://doi.org/10.1038/nmeth.2869
35 Dries, R. *et al.* Giotto: a toolbox for integrative analysis and visualization of spatial expression data. *Genome Biol* **22**, 78 (2021). https://doi.org/10.1186/s13059-021-02286-2
36 Ferrari, M. *et al.* Structure-guided engineering of immunotherapies targeting TRBC1 and TRBC2 in T cell malignancies. *Nat Commun* **15**, 1583 (2024). https://doi.org/10.1038/s41467-024-45854-3

37 Grebinoski, S. *et al.* Autoreactive CD8+ T cells are restrained by an exhaustion-like program that is maintained by LAG3. *Nature Immunology* **23**, 868-877 (2022). https://doi.org/10.1038/s41590-022-01210-5
38 Anderson, A. C., Joller, N. & Kuchroo, V. K. Lag-3, Tim-3, and TIGIT: Co-inhibitory Receptors with Specialized Functions in Immune Regulation. *Immunity* **44**, 989-1004 (2016). https://doi.org/10.1016/j.immuni.2016.05.001
39 Kersten, K. *et al.* Spatiotemporal co-dependency between macrophages and exhausted CD8(+) T cells in cancer. *Cancer Cell* **40**, 624-638.e629 (2022). https://doi.org/10.1016/j.ccell.2022.05.004
40 Lindner, R. *et al.* Molecular phenotypes in triple negative breast cancer from African American patients suggest targets for therapy. *PLoS One* **8**, e71915 (2013). https://doi.org/10.1371/journal.pone.0071915
41 Huo, D. *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* **25**, 4835-4846 (2016). https://doi.org/10.1093/hmg/ddw305
42 Haddad, S. A. *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* **37**, 870-877 (2016). https://doi.org/10.1093/carcin/bgw067
43 Newman, L. A. & Kaljee, L. M. Health Disparities and Triple-Negative Breast Cancer in African American Women: A Review. *JAMA Surg* **152**, 485-493 (2017). https://doi.org/10.1001/jamasurg.2017.0005
44 Piyarathna, D. W. B. *et al.* ERR1- and PGC1α-associated mitochondrial alterations correlate with pan-cancer disparity in African Americans. *The Journal of Clinical Investigation* **129**, 2351-2356 (2019). https://doi.org/10.1172/JCI127579
45 Jinesh, G. G. & Brohl, A. S. Classical epithelial-mesenchymal transition (EMT) and alternative cell death process-driven blebbishield metastatic-witch (BMW) pathways to cancer metastasis. *Signal Transduct Target Ther* **7**, 296 (2022). https://doi.org/10.1038/s41392-022-01132-6
46 Mason, K. *et al.* Niche-DE: niche-differential gene expression analysis in spatial transcriptomics data identifies context-dependent cell-cell interactions. *Genome Biol* **25**, 14 (2024). https://doi.org/10.1186/s13059-023-03159-6
47 Tharp, K. M. *et al.* Tumor-associated macrophages restrict CD8(+) T cell function through collagen deposition and metabolic reprogramming of the breast cancer microenvironment. *Nat Cancer* (2024). https://doi.org/10.1038/s43018-024-00775-4
48 Zhang, X. *et al.* Thrombospondin-1 modulates vascular endothelial growth factor activity at the receptor level. *Faseb j* **23**, 3368-3376 (2009). https://doi.org/10.1096/fj.09-131649
49 Hugo, W. *et al.* Genomic and Transcriptomic Features of Response to Anti-PD-1 Therapy in Metastatic Melanoma. *Cell* **165**, 35-44 (2016). https://doi.org/10.1016/j.cell.2016.02.065
50 Costa, A. *et al.* Fibroblast Heterogeneity and Immunosuppressive Environment in Human Breast Cancer. *Cancer Cell* **33**, 463-479 e410 (2018). https://doi.org/10.1016/j.ccell.2018.01.011
51 Launonen, I.-M. *et al.* Chemotherapy induces myeloid-driven spatial T-cell exhaustion in ovarian cancer. *bioRxiv*, 2024.2003.2019.585657 (2024). https://doi.org/10.1101/2024.03.19.585657
52 Jiang, Y. Z. *et al.* Genomic and Transcriptomic Landscape of Triple-Negative Breast Cancers: Subtypes and Treatment Strategies. *Cancer Cell* **35**, 428-440.e425 (2019). https://doi.org/10.1016/j.ccell.2019.02.001
53 Hammerl, D. *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* **12**, 5668 (2021). https://doi.org/10.1038/s41467-021-25962-0

54 Gruosso, T. *et al.* Spatially distinct tumor immune microenvironments stratify triple-negative breast cancers. *J Clin Invest* **129**, 1785-1800 (2019). https://doi.org/10.1172/JCI96313
55 Bagati, A. *et al.* Integrin alphavbeta6-TGFbeta-SOX4 Pathway Drives Immune Evasion in Triple-Negative Breast Cancer. *Cancer Cell* **39**, 54-67 e59 (2021). https://doi.org/10.1016/j.ccell.2020.12.001
56 Akil, A. *et al.* Notch Signaling in Vascular Endothelial Cells, Angiogenesis, and Tumor Progression: An Update and Prospective. *Front Cell Dev Biol* **9**, 642352 (2021). https://doi.org/10.3389/fcell.2021.642352
57 Wang, Z., Li, Y., Kong, D. & Sarkar, F. H. The role of Notch signaling pathway in epithelial-mesenchymal transition (EMT) during development and tumor aggressiveness. *Curr Drug Targets* **11**, 745-751 (2010). https://doi.org/10.2174/138945010791170860
58 Yuan, X. *et al.* Notch signaling and EMT in non-small cell lung cancer: biological significance and therapeutic application. *J Hematol Oncol* **7**, 87 (2014). https://doi.org/10.1186/s13045-014-0087-z
59 Singh, S. *et al.* Chemotherapy Coupled to Macrophage Inhibition Induces T-cell and B-cell Infiltration and Durable Regression in Triple-Negative Breast Cancer. *Cancer Res* **82**, 2281-2297 (2022). https://doi.org/10.1158/0008-5472.CAN-21-3714
60 Zheng, C. *et al.* Neutrophils in triple-negative breast cancer: an underestimated player with increasingly recognized importance. *Breast Cancer Research* **25**, 88 (2023). https://doi.org/10.1186/s13058-023-01676-7
61 Kim, I. S. *et al.* Immuno-subtyping of breast cancer reveals distinct myeloid cell profiles and immunotherapy resistance mechanisms. *Nat Cell Biol* **21**, 1113-1126 (2019). https://doi.org/10.1038/s41556-019-0373-7
62 Welte, T. *et al.* Oncogenic mTOR signalling recruits myeloid-derived suppressor cells to promote tumour initiation. *Nat Cell Biol* **18**, 632-644 (2016). https://doi.org/10.1038/ncb3355
63 Harney, A. S. *et al.* Real-Time Imaging Reveals Local, Transient Vascular Permeability, and Tumor Cell Intravasation Stimulated by TIE2hi Macrophage-Derived VEGFA. *Cancer Discov* **5**, 932-943 (2015). https://doi.org/10.1158/2159-8290.CD-15-0012
64 Harney, A. S. *et al.* The Selective Tie2 Inhibitor Rebastinib Blocks Recruitment and Function of Tie2(Hi) Macrophages in Breast Cancer and Pancreatic Neuroendocrine Tumors. *Mol Cancer Ther* **16**, 2486-2501 (2017). https://doi.org/10.1158/1535-7163.MCT-17-0241
65 Pignatelli, J. *et al.* Invasive breast carcinoma cells from patients exhibit MenaINV- and macrophage-dependent transendothelial migration. *Sci Signal* **7**, ra112 (2014). https://doi.org/10.1126/scisignal.2005329
66 Roussos, E. T. *et al.* Mena invasive (Mena(INV)) and Mena11a isoforms play distinct roles in breast cancer cell cohesion and association with TMEM. *Clin Exp Metastasis* **28**, 515-527 (2011). https://doi.org/10.1007/s10585-011-9388-6
67 Salgado, R. *et al.* The evaluation of tumor-infiltrating lymphocytes (TILs) in breast cancer: recommendations by an International TILs Working Group 2014. *Ann Oncol* **26**, 259-271 (2015). https://doi.org/10.1093/annonc/mdu450
68 Schapiro, D. *et al.* histoCAT: analysis of cell phenotypes and interactions in multiplex image cytometry data. *Nat Methods* **14**, 873-876 (2017). https://doi.org/10.1038/nmeth.4391
69 Xu, Y. *et al.* Patients with Lung Cancer of Different Racial Backgrounds Harbor Distinct Immune Cell Profiles. *Cancer Res Commun* **2**, 884-893 (2022). https://doi.org/10.1158/2767-9764.CRC-22-0057
70 Merritt, C. R. *et al.* Multiplex digital spatial profiling of proteins and RNA in fixed tissue. *Nat Biotechnol* **38**, 586-599 (2020). https://doi.org/10.1038/s41587-020-0472-9
71 Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* **9**, 671-675 (2012). https://doi.org/10.1038/nmeth.2089

72 Carpenter, A. E. *et al*. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol* **7**, R100 (2006). https://doi.org/10.1186/gb-2006-7-10-r100
73 Stringer, C., Wang, T., Michaelos, M. & Pachitariu, M. Cellpose: a generalist algorithm for cellular segmentation. *Nat Methods* **18**, 100-106 (2021). https://doi.org/10.1038/s41592-020-01018-x
74 Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676-682 (2012). https://doi.org/10.1038/nmeth.2019
75 Berg, S. *et al.* ilastik: interactive machine learning for (bio)image analysis. *Nat Methods* **16**, 1226-1232 (2019). https://doi.org/10.1038/s41592-019-0582-9
76 Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* **67**, 1 - 48 (2015). https://doi.org/10.18637/jss.v067.i01
77 Zhu, Q., Shah, S., Dries, R., Cai, L. & Yuan, G. C. Identification of spatially associated subpopulations by combining scRNAseq and sequential fluorescence in situ hybridization data. *Nat Biotechnol* (2018). https://doi.org/10.1038/nbt.4260
78 Rousseeuw, P. J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics* **20**, 53-65 (1987).
79 Obayashi, T. & Kinoshita, K. COXPRESdb: a database to compare gene coexpression in seven model animals. *Nucleic Acids Res* **39**, D1016-1022 (2011). https://doi.org/10.1093/nar/gkq1147
80 Wu, S. Z. *et al.* A single-cell and spatially resolved atlas of human breast cancers. *Nat Genet* **53**, 1334-1347 (2021). https://doi.org/10.1038/s41588-021-00911-1
81 Wigerblad, G. *et al.* Single-Cell Analysis Reveals the Range of Transcriptional States of Circulating Human Neutrophils. *J Immunol* **209**, 772-782 (2022). https://doi.org/10.4049/jimmunol.2200154
82 Alvarez-Breckenridge, C. *et al.* Microenvironmental Landscape of Human Melanoma Brain Metastases in Response to Immune Checkpoint Inhibition. *Cancer Immunol Res* **10**, 996-1012 (2022). https://doi.org/10.1158/2326-6066.CIR-21-0870
83 Efremova, M., Vento-Tormo, M., Teichmann, S. A. & Vento-Tormo, R. CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nat Protoc* **15**, 1484-1506 (2020). https://doi.org/10.1038/s41596-020-0292-x
84 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* **35**, W193-200 (2007). https://doi.org/10.1093/nar/gkm226
85 Merico, D., Isserlin, R., Stueker, O., Emili, A. & Bader, G. D. Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. *PLoS One* **5**, e13984 (2010). https://doi.org/10.1371/journal.pone.0013984

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

Variable	**Black American (BA, n=26, Value ± SD)**	**White American (WA, n=31, Value ± SD)**	**P value BA vs WA**
Age	58.2 ± 10.89	60.06 ± 13.00	0.4
Regions of Interest (ROI)	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			0.06
Grade 1	0	1	
Grade 2	2	7	
Grade 3	24	23	
Tumor Stage			0.18
Stage 1	9	14	
Stage 2	12	16	
Stage 3	3	0	
Stage 4	1	1	
Unknown	2	0	
BMI	35.14 ± 9.09	29.16 ± 6.09	0.07
Mets detected at surgical resection of primary tumor			
Yes	1	1	
No	25	30	
Unknown	0	0	
Neoadjuvant Chemotherapy			
Yes	2	2	
No	24	29	
TP53 Mutation H			
Yes	1	4	
No	25	27	
PIK3CA Mutation H			
Yes	2	3	
No	14	11	
Unknown	10	17	
PIK3CA Mutation E			
Yes	4	6	
No	12	8	
Unknown	10	17	
Median Survival (months)	131.5±61.5	126±44.6	0.51

Table 2. Summarized clinical information of patients in the Tissue Microarray (TMA) from Roswell Park Validation cohort, examined using imaging mass cytometry (IMC).

Variable	**Black American (n=15, Value ± SD)**	**White American (n=31, Value ± SD)**	**P value**
Tumor Grade			0.16
Grade 1	0	0	
Grade 2	0	2	
Grade 3	13	29	
Unknown	2	0	
Tumor Stage			0.10
Stage 1	3	9	
Stage 2	6	18	
Stage 3	5	4	
Stage 4	1	0	
Unknown	0	0	
AGE	53.66 ± 14.54	56.83 ± 15.19	0.5
Median Survival (months)	79.00 ± 38.09	98.00 ± 34.13	0.7

Table 3: Table showing results of Cox Proportional Hazard Analysis for the Baylor Scott and White Discovery Data

	Hazard Ratio (HR)	**P-value**
Univariate model		
Race	1.55	0.5
Stage	1.76	0.14
Grade	0.82	0.35
Age	0.98	0.59
Body Mass Index (BMI)	1.02	0.6
Diabetes Mellitus Status	0.86	0.85
PIK3CA Mutation Status	0.75	0.53
RAS Mutation Status	0.40	0.2
TP53 Mutation Status	0.40	0.18
BA-communities-all	3.45	0.06
WA-communities-all	1.18	0.40
Multivariate model		
Adjusting for age:		
BA-communities-all	3.38	0.065
WA-communities-all	1.15	0.417
Adjusting for race:		
BA-communities-all	3.73	0.066
WA-communities-all	1.41	0.315
Adjusting for stage:		
BA-communities-all	4.66	0.037
WA-communities-all	1.02	0.49
Adjusting for age, race, and stage:		
BA-communities-all	5.35	0.037
WA-communities-all	1.21	0.397
Adjusting for age, race, BMI, stage:		
BA-communities-all	7.69	0.021
WA-communities-all	1.27	0.368

Table 4: Table showing results of Cox Proportional Hazard Analysis for the Roswell Park Validation Data

Cox-proportional Hazard Ratios (for Roswell Park Cohort)		
	Hazard Ratio	**P-value**
Univariate model		
stage	1.75	0.125
age	1.03	0.061
race	1.61	0.264
BA-communities-all	2.15	0.113
WA-communities-all	4.16	0.036
Multivariate model, adjusting for stage, age, race		
BA-communities-all	3.75	0.032
WA-communities-all	6.11	0.025

Supplementary Table 1. Association of cluster abundance with patient race and TNBC prognosis

Cluster ID		Abundance (BA vs WA) P-val	HR	HR (P-val)	BA HR	BA HR (P-val)	WA HR	WA HR (P-val)
1	CD16 CD163 PDL1	0.84	0.978	0.67	0.93	0.537	1.09	0.36
2	PanCK	**0.003**	0.996	0.931	0.88	0.67	1.009	0.83
3	CD11c pHH3	0.428	0.969	0.614	0.98	0.758	0.6	0.57
4	Vim AR PDL1 PLK1	0.058	1.008	0.9	0.78	0.36	1.05	0.37
5	CD11c FOXP3 pHH3 Ecad	0.137	1.007	0.85	1.02	0.55	0.81	0.46
6	HIF1a	0.973	1.03	0.487	1.004	0.949	1.06	0.34
7	CD152 CD8a	0.094	0.983	0.779	1.02	0.88	0.98	0.82
8	PLK1 PanCK Ki67	0.629	0.828	0.376	0.87	0.54	0.78	0.52
9	GzmB CD152 HIF1a	0.221	**1.1**	0.169	1.09	0.18	0.91	0.76
10	CD68	0.53	0.926	0.426	0.93	0.589	0.93	0.61
11	CD3 CD45 CD4 CD8a CD45RO	0.21	0.93	0.642	0.919	0.697	0.94	0.8
12	CD16 CD163 CD68	0.35	1.002	0.923	0.997	0.95	1.02	0.74
13	CD31 Vim AR	**0.0083**	1.007	0.956	0.986	0.96	1.059	0.74
14	CD45 CD45RA CD45RO	0.059	0.971	0.629	0.95	0.65	1.17	0.44
15	CD4 CD3	0.879	1.06	0.337	1.1	0.429	1.06	0.42
16	PanCK VEGF Ecad	0.388	**1.02**	**0.114**	0.996	0.9	**1.09**	**0.004**
17	KIFC1 GzmB Ki67	0.295	**1.05**	**0.08**	1.02	0.34	**1.11**	**0.1**
18	Ki67 KIFC1	0.728	1.01	0.782	1.03	0.45	0.94	0.63
19	PLK1 PD1	**0.0088**	0.937	0.607	1.15	0.67	0.92	0.64
20	CD31 CD45RA	0.148	1.02	0.693	1.06	0.31	0.93	0.71

Supplementary Table 2: List of antibodies and the metal tags used for Imaging Mass Cytometry (IMC) analysis of BSW-Discovery samples examined in Tissue Microarray format.

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 BSW2 validation samples

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

Supplementary Table 4: List of antibodies and the metal tags used for Imaging Mass Cytometry (IMC) analysis of Roswell Park validation samples

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
160Gd	CD14
161Dy	CD152/CTLA4
162Dy	NOS2
163Dy	VEGF
164Dy	MPO
165Ho	HIF1a
166Er	CD45RA
167Er	Granzyme B
168Er	Ki-67
169Tm	Arginase-1
170Er	CD3
173Yb	CD45RO
175Lu	KIFC1
176Yb	pHH3
Pr141	CD8a
Nd142	CD20

Figure Legends

Figure 1: Imaging Mass Cytometry (IMC) Profiling of Self-Reported Black American (BA) and White American (WA) Triple Negative Breast Cancer (TNBC). a) Schematic representation of the clinical cohorts analyzed, and the corresponding types of analyses performed for each group. For the BSW2 validation cohort sample numbers 10/9 describe samples examined using imaging mass cytometry and 10X spatial transcriptomics, respectively. b) Overview of the analytical workflow employed throughout the study. c) Diagram illustrating the imaging mass cytometry-based immune cell profiling process applied to BA and WA TNBC tumors, utilizing a tissue microarray (TMA) format. d) Unsupervised clustering of individual cells, segmented from IMC data of BA and WA TNBC tumors, identified 20 distinct clusters. The differential protein markers and their corresponding nomenclature (columns) across these 20 clusters (rows) are shown. e) Representative t-SNE plots illustrating the distribution of the 20 clusters as described in panel (d) within the expression space. f) Overlay of expression patterns for E-Cadherin, CD45RA, CD31, CD68, and Vimentin onto the t-SNE plots.

Figure 2: Spatial Proximity Analysis Uncovers Distinct Cell-Cell Interactions in Black American (BA) and White American (WA) Triple Negative Breast Cancer (TNBC) Using Imaging Mass Cytometry (IMC). a) Landscape of cell-cell interactions in BA (left) and WA (right) TNBC, based on pairwise interactions between the 20 cell clusters defined in **Figure 1d.** Race-specific interactions were identified using a linear mixed model, where the interaction score is modeled as score ~ race + patient, with race as a fixed effect and patient as a random effect. Coefficients for the BA (left) and WA (right) terms are shown, with positive coefficients indicating enriched interactions and negative coefficients indicating depleted interactions. Significant BA-specific interactions are highlighted with red boxes, while significant WA-specific interactions are highlighted with blue boxes. b) Top six spatially resolved cell-cell interactions in BA TNBC (blue bars) compared to WA TNBC (red bars), with significance determined by -log10 P values (green bars). BA TNBC tumors are characterized by Endothelial-Macrophage-Mesenchymal (Endo-Mac-Vim) interactions. Error bars represent the 95% confidence interval (CI) derived from 100 leave-2-patient-out subsampling. c) Same as panel (b), but showing the top five spatially resolved cell-cell interactions in WA TNBC (red bars) compared to BA TNBC (blue bars), with significance computed by -log10 P values (green bars). WA TNBC tumors are marked by immune exhaustion-related interactions. Error

bars represent the 95% CI derived from 100 leave-2-patient-out subsampling. d) Spatial illustration of key BA-associated cell-cell interactions identified by IMC, showing co-localization of Endothelial (black) and Macrophage (orange) clusters in BA patients. e) Spatial illustration of key WA-associated cell-cell interactions identified by IMC, highlighting the prominent interactions between Cytotoxic cells (orange) and Exhausted T cells (pink) in WA tumors, suggesting immune exhaustion. f) Multiplex immunofluorescence validation of Endothelial (CD31)-Macrophage (CD163) interactions in BA TNBC, but not in WA TNBC. Representative images from three BA and three WA TNBC patients are shown. Magnified images (from the large white box) of areas marked by small white boxes are included as insets. g, h) Pixel-level co-localization quantification of Endothelial-Macrophage interaction, shown in panel (f), across the entire tumor tissue (g) and per cell within the tumor (h) in BA and WA TNBC. P values for panels (g) and (h) were calculated using the Mann-Whitney test.

Figure 3: Identification of Recurrent Cellular Communities and Their Association with Clinical Outcomes in BA and WA TNBC Cohorts. a) **BA-Community-1**: In the BA cohort, community detection of cellular interactions, derived from expanding differential interactions (FDR = 0.20) and identifying connected components, revealed a dominant community (BA-Community-1) characterized by recurring interactions between “M2 Macrophages-Endothelial Cells-Mesenchymal Cells-Naïve T Cells.” This community was consistently observed across the BSW discovery, BSW2, and Roswell Park datasets and was associated with clinical outcomes in BA TNBC. b) **WA-Community-1**: Using the approach described in a), in the WA cohort, the dominant community (WA-Community-1) comprised “Exhausted Cytotoxic T Cells-Helper T Cells-Macrophages-Hypoxia.” This community was consistently found across the BSW discovery, BSW2, and Roswell Park datasets. For additional detected communities in BA and WA, see **Supplementary Figure 9**. c) Kaplan Meier (KM) plot showing significantly poor overall survival among TNBC patients when patients were stratified by the sum of scores of all BA communities (1&2) after adjusting for age, race and stage (**see Figure 2b**), in the BSW-discovery and Roswell Park validation cohort. d) KM plot showing the stratification of TNBC patients based on the sum of scores of all WA communities (1-4) after adjusting for age, race and stage (**see Figure 2c**), in the BSW-discovery and Roswell Park validation cohort. Stratification into the high and low groups was performed using average. All P values computed using one-tailed log-rank test.

Figure 4: BA-Associated Community Exhibits Unique Spatial Clustering in TNBC Spatial Transcriptomics Data. a) Overview of the approach used to identify the spatial localization of BA- and WA-associated communities in TNBC tissues. BA or WA-associated communities, defined by IMC-derived protein markers of interacting cell types, were converted to gene names and interrogated in a racially balanced TNBC spatial transcriptomics (ST) dataset to assess spatial patterns of co-localization for the query genes. b) The top BA-associated community (BA-community 1) was interrogated in the ST dataset[30]. Protein symbols (blue) and corresponding gene symbols (black) are shown. c) Representative example of spatial co-localization for BA-community 1 in 10 BA and 10 WA TNBC samples from the ST dataset. BA-community 1 genes show strong co-localization (marked by black outlines) in BA TNBC, but not in WA TNBC. The scale bar represents the strength of co-localization, defined as the averaged log-transformed normalized read count. d) Quantification of spatial clustering of co-localized spots for BA-community 1 in BA (blue) and WA (red) TNBC tumors. Each bar represents the average of 10 BA and 10 WA samples shown in panel (c). Spatial clustering of BA-community 1 was significantly higher in BA TNBC compared to WA TNBC, with statistical significance determined by T-test.

Figure 5: Distinct Gene Signatures Define Unique Cell Types in BA and WA Tumor Niches. a) Overview of the approach used to identify niche-associated gene signatures. For BA- and WA-associated communities (BA-community 1 and WA-community 1), extended signature genes (ESGs) were derived based on co-expression within query (i.e., BA-community 1 or WA-community 1 genes) co-localized spots. ESGs were computed per sample and compared between BA and WA, with genes significantly elevated in BA (i.e., BA > WA for BA-niche) or in WA (i.e., WA > BA for WA-niche) selected for further analysis. Each query generated a distinct BA-niche and WA-niche, composing of ST spots uniquely expressing BA-community 1 or WA-community 1 genes. b) Representative heat map showing cell-type-specific expression of ESGs associated with the BA niche defined by expansion of BA-community 1. These ESGs correspond to various cell types, including Cancer-Associated Fibroblasts (CAF), endothelial cells, and myeloid cells, while the BA-niche is poor in T-cells and neutrophils. c) Same as in (b), but for the heat map showing cell-type-specific expression of ESGs associated with the WA niche defined by expansion of WA-community 1. ESGs in this niche predominantly correspond to exhausted T-cells and neutrophils, while the WA-niche is poor in CAFs and endothelial cells. d) Average quantification of the number of co-localized

Visium spots in BA (blue) and WA (red) TNBC corresponding to BA- and WA-associated niches. Statistical significance was computed using a T-test. e, f) Gene set enrichment analysis (GSEA) of ESGs associated with the BA-niche (e) and WA-niche (f), with -log10 P-values shown.

Figure 6: Ligand-Receptor Interactions and ESG Analysis in BA and WA Niches Reveal Distinct Mechanisms of Tumor Progression and Immune Suppression. a) Ligand-receptor interactions derived from ESGs associated with the BA niche (BA-community 1) reveal the activation of multiple tumor-promoting pathways, including those involving cell surface proteins, adhesion molecules, extracellular matrix components, endothelial cells, and key signaling pathways such as platelet-derived growth factor receptor (PDGFR), Transforming Growth Factor beta (TGF-ß), WNT, Notch, Thrombospondin, and Placental Growth Factor (PGF). In contrast, ligand-receptor interactions in the WA niche (WA-community 1) are dominated by cytokine activities and pathways related to T-cell exhaustion. b) Normalized frequencies of ligand-receptor interactions in the BA-niche, analyzed using ESGs from BA-community 1, are presented in a matrix format with ligands (rows) and receptors (columns). c) Normalized frequencies of ligand-receptor interactions in the WA-niche, analyzed using ESGs from WA-community 1, are similarly presented. d) Average expression of genes associated with T-cell exhaustion[31] in WA-community 1 in WA and BA TNBC tissues. The expression of T-cell exhaustion-related genes was significantly higher in WA TNBC compared to BA TNBC, with significance computed using a T-test. e) Quantification of the average number of spots containing genes associated with T-cell exhaustion in 10 BA and 10 WA TNBC spatial transcriptomics samples[30]. The number of co-localized spots was significantly higher in WA TNBC compared to BA TNBC, with significance determined using a Kolmogorov-Smirnov test.

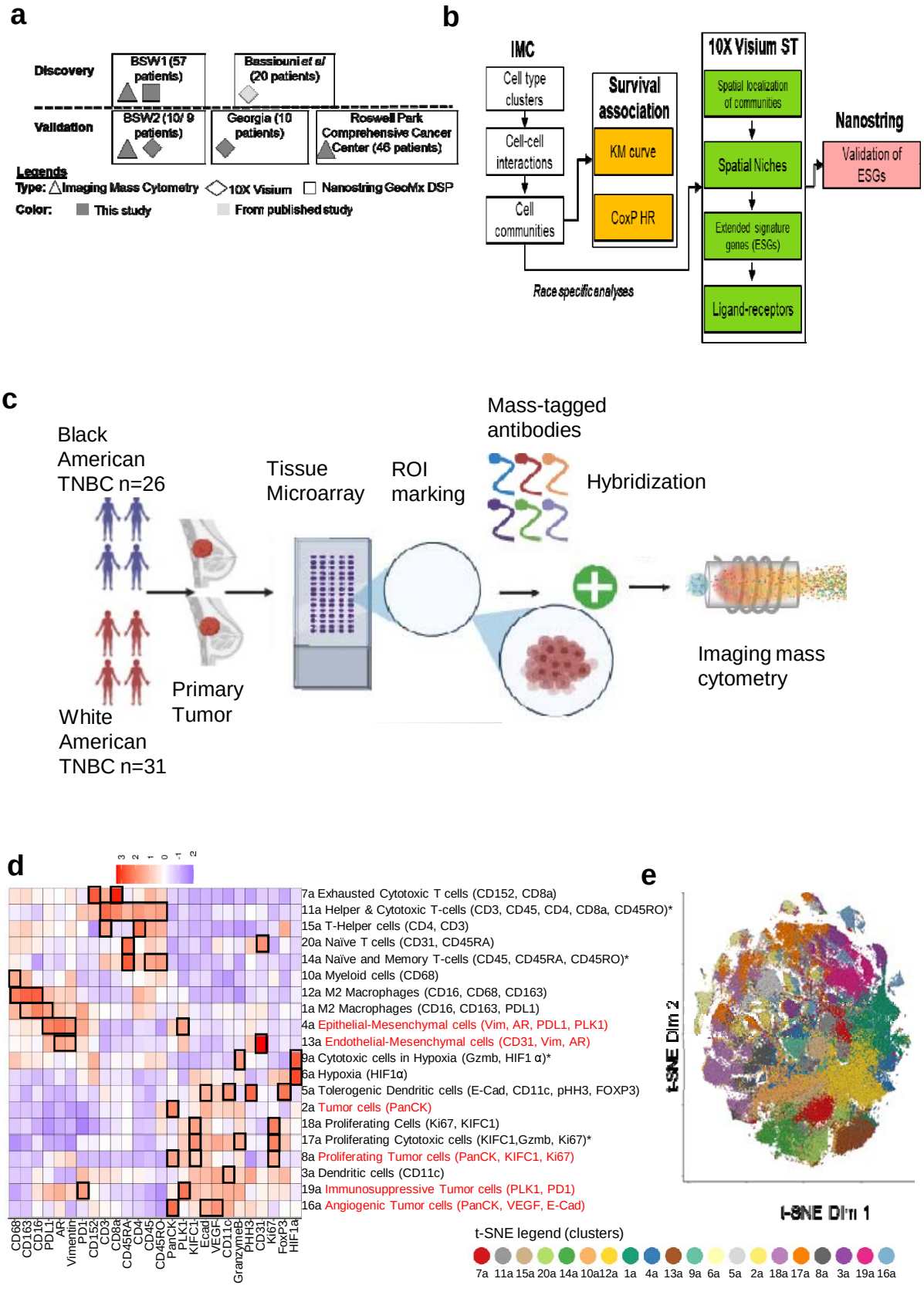
Figure 7: Independent Validation of BA and WA Niches Using Spatial Transcriptomics and GeoMX DSP System (NanoString). a) **Spatial Transcriptomics Validation in BSW2 Cohort**: Top panel: Spatial transcriptomics analysis of 4 independent FFPE BA tumors and 5 WA tumors from the BSW2 validation cohort verifies the elevated expression of ESGs associated with BA-community 1 in BA TNBC compared to WA. Bottom panel: The same analysis verifies the elevated expression of ESGs associated with WA-community 1 in WA TNBC compared to BA. b) **Spatial Transcriptomics Validation in Georgia Cohort**: Top

panel: Spatial transcriptomics analysis of 6 FFPE BA tumors and 4 WA tumors from the Georgia validation cohort confirms the elevated expression of ESGs associated with BA-community 1 in BA compared to WA TNBC. Bottom panel: The same analysis verifies the elevated expression of ESGs associated with WA-community 1 in WA compared to BA TNBC. c) **Representative Co-localization of ESGs Associated with BA-Niche**: Spatial co-localization of ESGs associated with BA-niche is observed in BA TNBC (top), but not in WA TNBC (bottom). d) **GeoMX Validation of BA and WA Niche Enrichment**: Analysis of Nanostring GeoMx data from 26 BA and 31 WA TNBC tumors used for IMC in **Figure 1** shows significantly higher enrichment of BA-community 1 (Endo-Mac-EMT) in the immune compartments (CD45 positive) of BA compared to WA TNBC. Similar analysis shows significantly higher enrichment of WA-community 1 (exhausted T cells) in the immune compartments of WA compared to BA TNBC. Significance was computed using T-tests.

Figure 8: Distinct Tumor Microenvironments (TME) in BA and WA TNBC. a) **BA Tumor Microenvironment (TME)**: The TME in BA TNBC is defined by a dynamic and highly integrated interaction between endothelial cells, macrophages, and tumor cells undergoing epithelial-mesenchymal transition (EMT). In this model, tumor cells express vimentin, a key marker of EMT, which enhances their motility and ability to invade surrounding tissues. A critical feature of the BA TME is the presence of a Tumor Microenvironment of Metastasis (TMEM), where the interplay between endothelial cells and macrophages creates specialized "doorways" that facilitate the intravasation of tumor cells into the bloodstream. These TMEM "doorways" are regions where the endothelial cells undergo modifications that increase permeability, enabling tumor cells to enter the vasculature more easily. The macrophages in this context are key players, secreting factors that promote endothelial cell destabilization and fostering the creation of these metastatic niches. The resulting tumor-promoting environment is immune-suppressive, limiting the ability of host immune cells to effectively target and eliminate the tumor. This BA-specific TME, combined with the TMEM mechanisms, supports tumor progression and the establishment of metastatic lesions. b) **WA Tumor Microenvironment (TME)**: In contrast, the WA TME is marked by a prominent immune exhaustion phenotype. Here, exhausted T-cells and neutrophils are prevalent, set against a backdrop of myeloid cells that drive immune suppression. This immune-suppressive environment is in part maintained by the interactions between exhausted T-cells and the cytokine milieu, which promotes im-

mune dysfunction. While the WA TME does not exhibit the same TMEM "doorways" as seen in BA TNBC, the immune-exhaustive nature of this microenvironment still creates challenges for tumor-targeting immune responses. In this context, the accumulation of myeloid cells, particularly macrophages, perpetuates a cycle of immune suppression, preventing effective anti-tumor immunity and allowing the tumor to persist and potentially metastasize.

Figure 1


a
Discovery
BSW1 (57 patients)
Bassiouni et al (20 patients)
Validation
BSW2 (10/ 9 patients)
Georgia (10 patients)
Roswell Park Comprehensive Cancer Center (46 patients)
Legends
Type: Imaging Mass Cytometry
10X Visium
Nanostring GeoMx DSP
Color: This study
From published study
b
IMC
Cell type clusters
Cell-cell interactions
Cell communities
Survival association
KM curve
CoxP HR
Race specific analyses
10X Visium ST
Spatial localization of communities
Spatial Niches
Extended signature genes (ESGs)
Ligand-receptors
Nanostring
Validation of ESGs
c
Black American TNBC n=26
White American TNBC n=31
Primary Tumor
Tissue Microarray
ROI marking
Mass-tagged antibodies
Hybridization
Imaging mass cytometry
d
7a Exhausted Cytotoxic T cells (CD152, CD8a)
11a Helper & Cytotoxic T-cells (CD3, CD45, CD4, CD8a, CD45RO)*
15a T-Helper cells (CD4, CD3)
20a Naïve T cells (CD31, CD45RA)
14a Naïve and Memory T-cells (CD45, CD45RA, CD45RO)*
10a Myeloid cells (CD68)
12a M2 Macrophages (CD16, CD68, CD163)
1a M2 Macrophages (CD16, CD163, PDL1)
4a Epithelial-Mesenchymal cells (Vim, AR, PDL1, PLK1)
13a Endothelial-Mesenchymal cells (CD31, Vim, AR)
9a Cytotoxic cells in Hypoxia (Gzmb, HIF1 α)*
6a Hypoxia (HIF1α)
5a Tolerogenic Dendritic cells (E-Cad, CD11c, pHH3, FOXP3)
2a Tumor cells (PanCK)
18a Proliferating Cells (Ki67, KIFC1)
17a Proliferating Cytotoxic cells (KIFC1,Gzmb, Ki67)*
8a Proliferating Tumor cells (PanCK, KIFC1, Ki67)
3a Dendritic cells (CD11c)
19a Immunosuppressive Tumor cells (PLK1, PD1)
16a Angiogenic Tumor cells (PanCK, VEGF, E-Cad)
CD68 CD163 CD16 PDL1 AR Vimentin PD1 CD152 CD3 CD8a CD45RA CD4 CD45 CD45RO PanCK PLK1 KIFC1 Ecad VEGF CD11c GranzymeB PHH3 CD31 Ki67 FoxP3 HIF1a
e
t-SNE Dim 2
t-SNE Dim 1
t-SNE legend (clusters)
7a 11a 15a 20a 14a 10a 12a 1a 4a 13a 9a 6a 5a 2a 18a 17a 8a 3a 19a 16a
f
E-Cadherin
CD45RA
CD31
CD68
Vimentin
t-SNE Dim 2

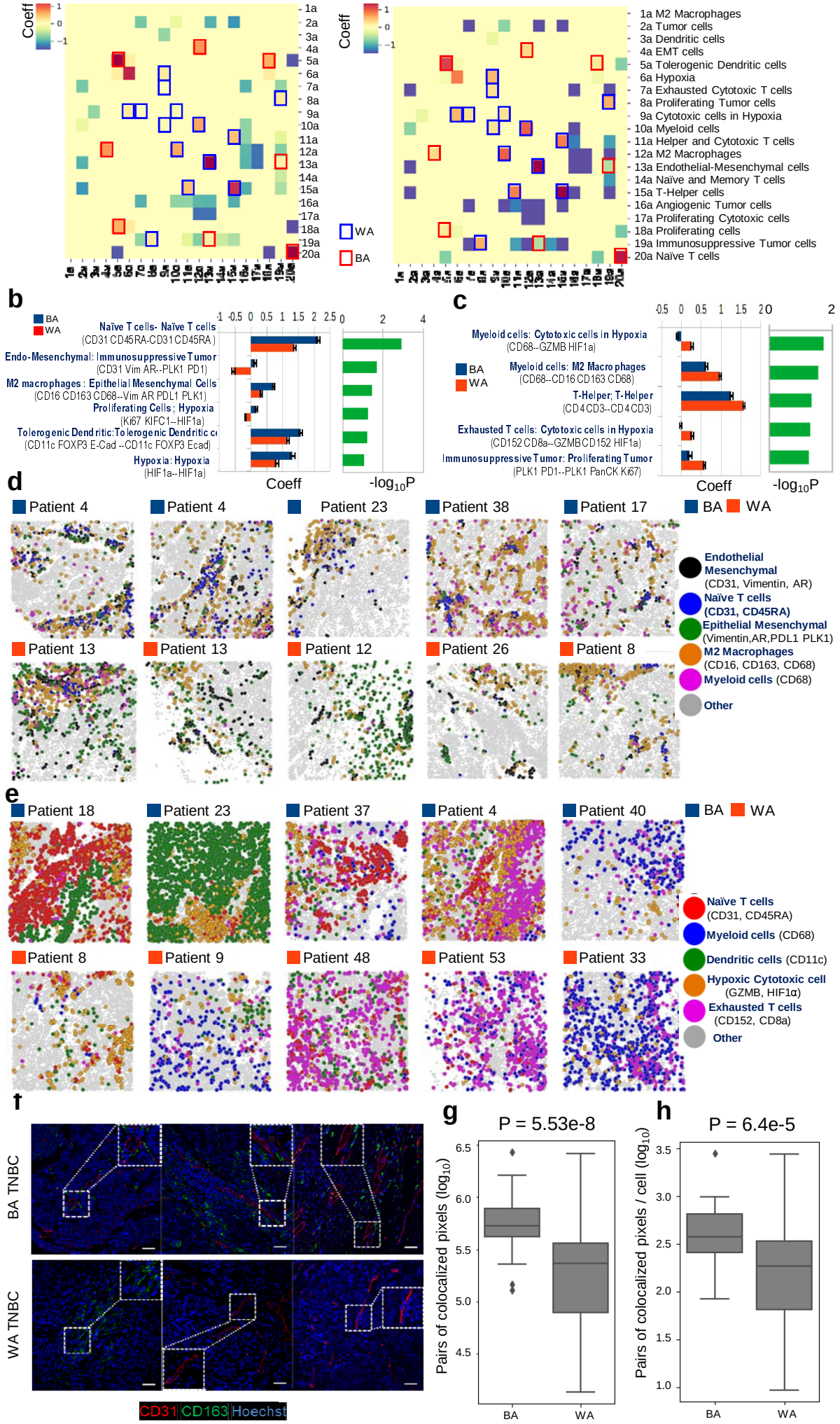
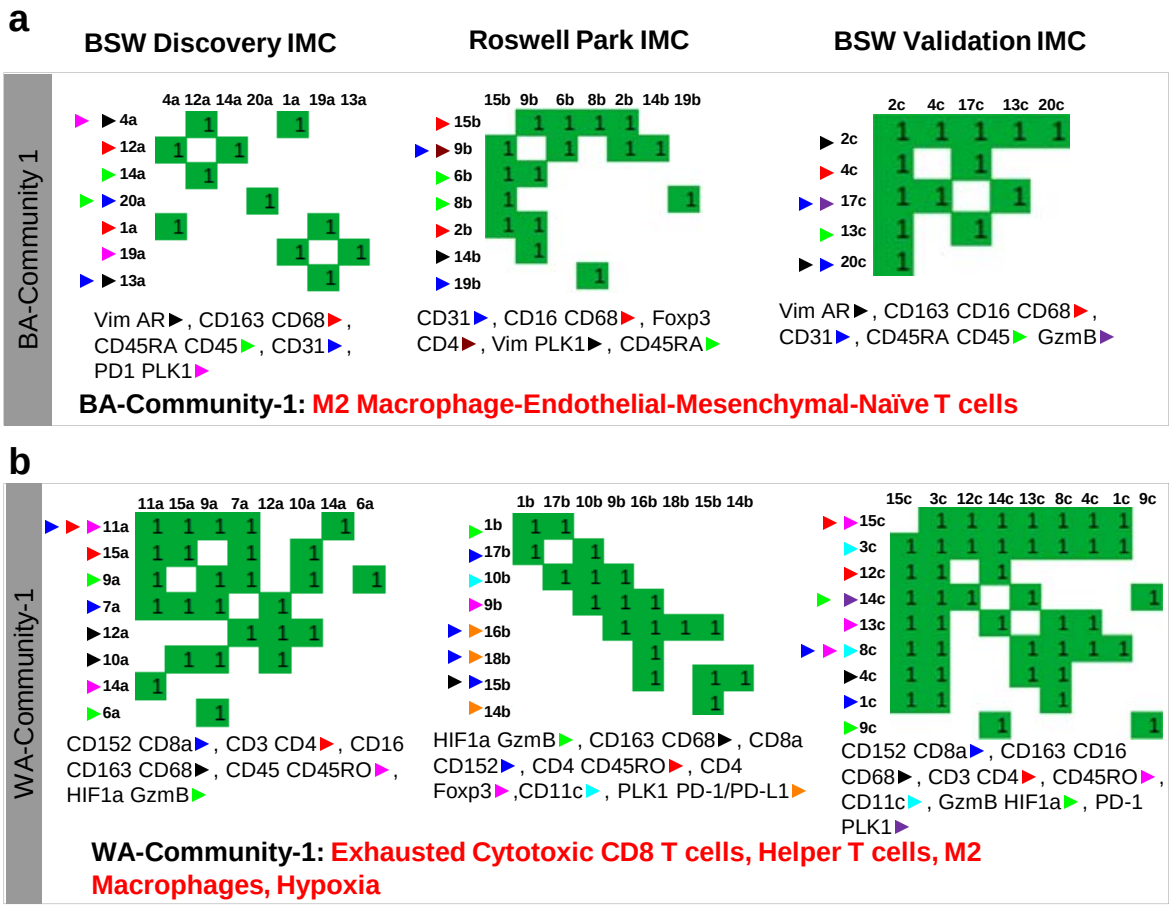

Coeff
1a M2 Macrophages
2a Tumor cells
3a Dendritic cells
4a EMT cells
5a Tolerogenic Dendritic cells
6a Hypoxia
7a Exhausted Cytotoxic T cells
8a Proliferating Tumor cells
9a Cytotoxic cells in Hypoxia
10a Myeloid cells
11a Helper and Cytotoxic T cells
12a M2 Macrophages
13a Endothelial-Mesenchymal cells
14a Naïve and Memory T cells
15a T-Helper cells
16a Angiogenic Tumor cells
17a Proliferating Cytotoxic cells
18a Proliferating cells
19a Immunosuppressive Tumor cells
20a Naïve T cells
WA
BA
b
Naïve T cells- Naïve T cells
(CD31 CD45RA-CD31 CD45RA)
Endo-Mesenchymal: Immunosuppressive Tumor
(CD31 Vim AR--PLK1 PD1)
M2 macrophages : Epithelial Mesenchymal Cells
(CD16 CD163 CD68--Vim AR PDL1 PLK1)
Proliferating Cells ; Hypoxia
(Ki67 KIFC1--HIF1a)
Tolerogenic Dendritic:Tolerogenic Dendritic c
(CD11c FOXP3 E-Cad --CD11c FOXP3 Ecad)
Hypoxia: Hypoxia
(HIF1a--HIF1a)
Coeff
-log10P
c
Myeloid cells: Cytotoxic cells in Hypoxia
(CD68--GZMB HIF1a)
Myeloid cells: M2 Macrophages
(CD68--CD16 CD163 CD68)
T-Helper; T-Helper
(CD4 CD3--CD4 CD3)
Exhausted T cells: Cytotoxic cells in Hypoxia
(CD152 CD8a--GZMB CD152 HIF1a)
Immunosuppressive Tumor: Proliferating Tumor
(PLK1 PD1--PLK1 PanCK Ki67)
Coeff
-log10P
d
Patient 4
Patient 23
Patient 38
Patient 17
Patient 13
Patient 12
Patient 26
Patient 8
BA
WA
Endothelial Mesenchymal (CD31, Vimentin, AR)
Naïve T cells (CD31, CD45RA)
Epithelial Mesenchymal (Vimentin,AR,PDL1 PLK1)
M2 Macrophages (CD16, CD163, CD68)
Myeloid cells (CD68)
Other
e
Patient 18
Patient 23
Patient 37
Patient 4
Patient 40
Patient 8
Patient 9
Patient 48
Patient 53
Patient 33
Naïve T cells (CD31, CD45RA)
Myeloid cells (CD68)
Dendritic cells (CD11c)
Hypoxic Cytotoxic cell (GZMB, HIF1α)
Exhausted T cells (CD152, CD8a)
Other
f
BA TNBC
WA TNBC
CD31 CD163 Hoechst
g
P = 5.53e-8
Pairs of colocalized pixels (log10)
BA
WA
h
P = 6.4e-5
Pairs of colocalized pixels / cell (log10)
BA
WA

Figure 2

Figure 3.


a
BSW Discovery IMC
Roswell Park IMC
BSW Validation IMC
BA-Community 1
4a 12a 14a 20a 1a 19a 13a
Vim AR, CD163 CD68, CD45RA CD45, CD31, PD1 PLK1
15b 9b 6b 8b 2b 14b 19b
CD31, CD16 CD68, Foxp3 CD4, Vim PLK1, CD45RA
2c 4c 17c 13c 20c
Vim AR, CD163 CD16 CD68, CD31, CD45RA CD45 GzmB
BA-Community-1: M2 Macrophage-Endothelial-Mesenchymal-Naïve T cells
b
WA-Community-1
11a 15a 9a 7a 12a 10a 14a 6a
CD152 CD8a, CD3 CD4, CD16 CD163 CD68, CD45 CD45RO, HIF1a GzmB
1b 17b 10b 9b 16b 18b 15b 14b
HIF1a GzmB, CD163 CD68, CD8a CD152, CD4 CD45RO, CD4 Foxp3,CD11c, PLK1 PD-1/PD-L1
15c 3c 12c 14c 13c 8c 4c 1c 9c
CD152 CD8a, CD163 CD16 CD68, CD3 CD4, CD45RO, CD11c, GzmB HIF1a, PD-1 PLK1
WA-Community-1: Exhausted Cytotoxic CD8 T cells, Helper T cells, M2 Macrophages, Hypoxia

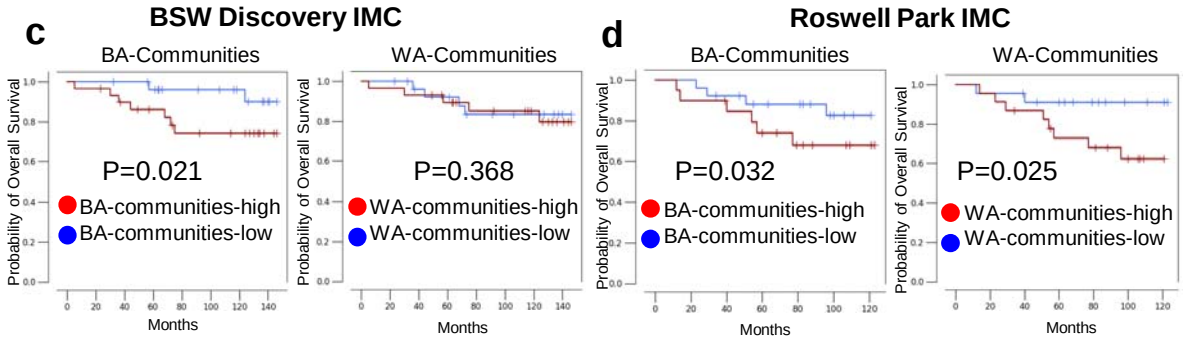

c
BSW Discovery IMC
BA-Communities
WA-Communities
Probability of Overall Survival
P=0.021
BA-communities-high
BA-communities-low
P=0.368
WA-communities-high
WA-communities-low
Months
d
Roswell Park IMC
BA-Communities
WA-Communities
P=0.032
BA-communities-high
BA-communities-low
P=0.025
WA-communities-high
WA-communities-low
Months

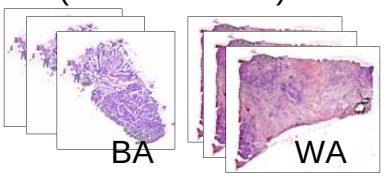
Figure 4

a

Query
IMC-derived BA or WA Community

Proteins → Genes

TNBC spatial transcriptomics
(Bassiouni *et al*)



- Spatial-Localization
- Extended Gene Signature (ESG, Fig. 5)

b **BA Query**:

BA-Community1: M2 Macrophage-Endothelial-Mesenchymal-Naïve T cells

Protein: CD31 CD45RA Vim AR CD16 CD68 CD163

Gene: *PECAM1 PTPRC VIM AR FCGR3A FCGR3B CD68 CD163*

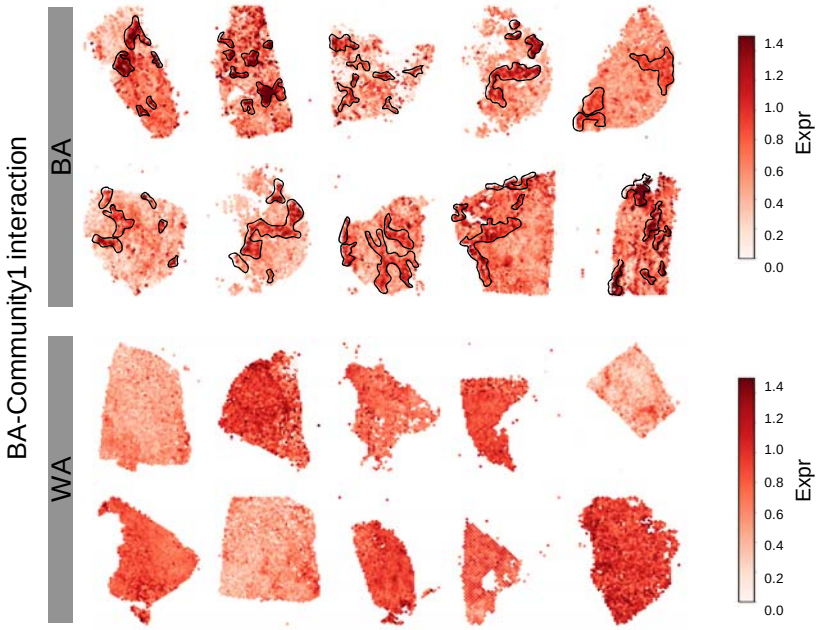
WA Query:

WA-Community1: Exhausted Cytotoxic CD8 T cells, Helper T cells, M2 Macrophages, Hypoxia

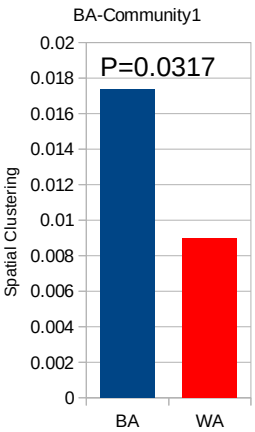
Protein: CD152 CD8a CD4 CD163 CD68 CD45RO HIF1a GzmB PLK1 PD-1 CD3

Gene: *CTLA4 CD8A CD4 CD163 CD68 PTPRC HIF1A GZMB PLK1 PDCD1 CD3E CD3D CD3G*

c **Spatial Localization (Niche)**



d



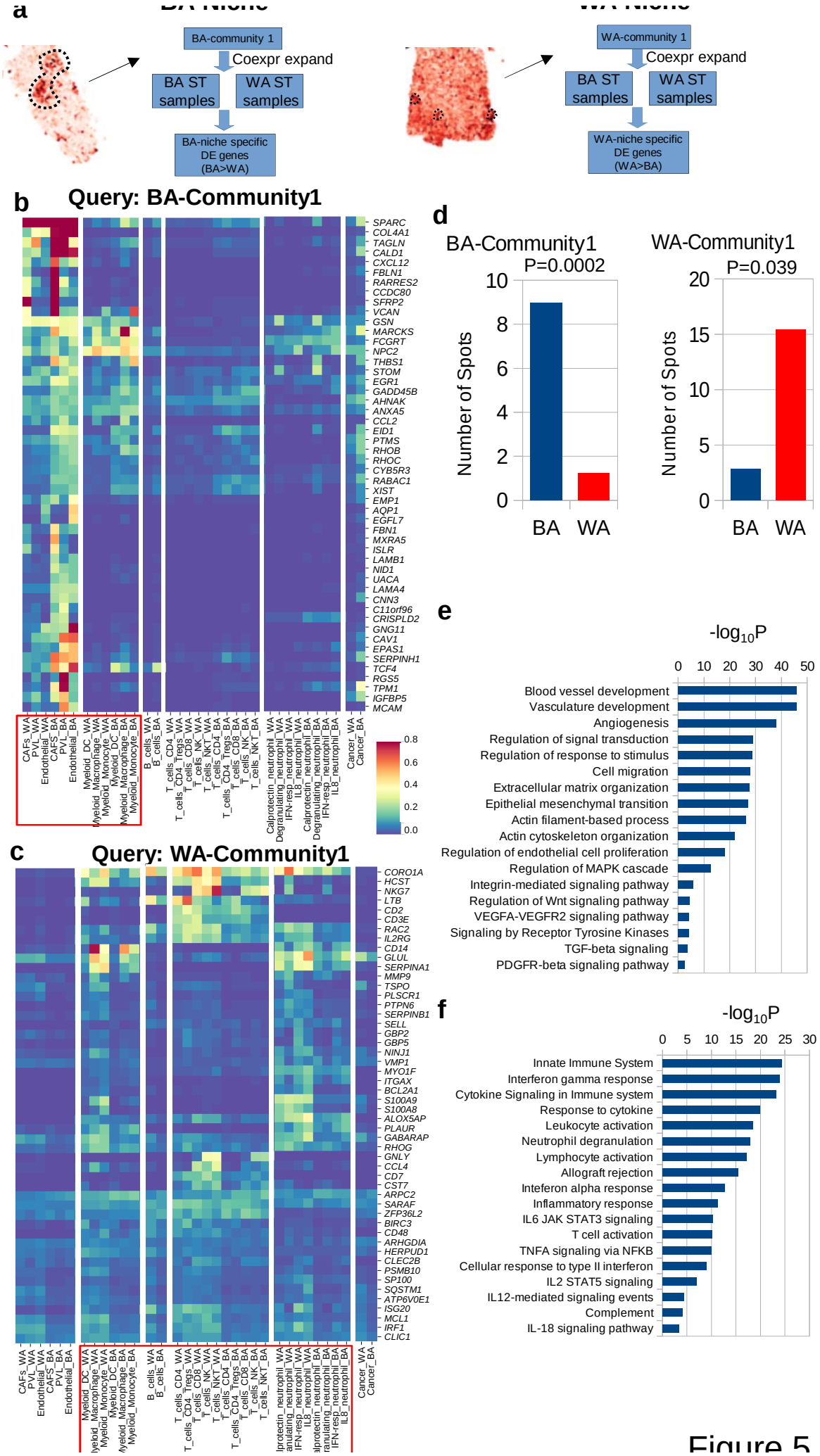
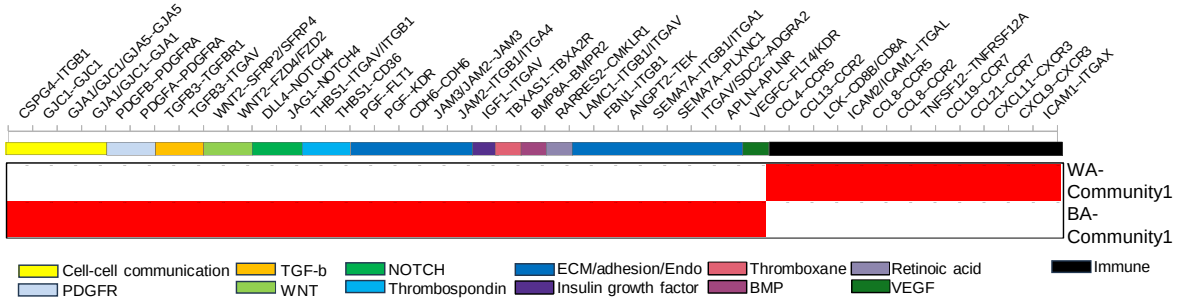

a
BA-community 1
Coexpr expand
BA ST samples
WA ST samples
BA-niche specific DE genes (BA>WA)
WA-community 1
Coexpr expand
BA ST samples
WA ST samples
WA-niche specific DE genes (WA>BA)
b
Query: BA-Community1
SPARC
COL4A1
TAGLN
CALD1
CXCL12
FBLN1
RARRES2
CCDC80
SFRP2
VCAN
GSN
MARCKS
FCGRT
NPC2
THBS1
STOM
EGR1
GADD45B
AHNAK
ANXA5
CCL2
EID1
PTMS
RHOB
RHOC
CYB5R3
RABAC1
XIST
EMP1
AQP1
EGFL7
FBN1
MXRA5
ISLR
LAMB1
NID1
UACA
LAMA4
CNN3
C11orf96
CRISPLD2
GNG11
CAV1
EPAS1
SERPINH1
TCF4
RGS5
TPM1
IGFBP5
MCAM
CAFs_WA
PVL_WA
Endothelial_WA
CAFs_BA
PVL_BA
Endothelial_BA
Myeloid_DC_WA
Myeloid_Macrophage_WA
Myeloid_Monocyte_WA
Myeloid_DC_BA
Myeloid_Macrophage_BA
Myeloid_Monocyte_BA
B_cells_WA
B_cells_BA
T_cells_CD4_WA
T_cells_CD4_Tregs_WA
T_cells_CD8_WA
T_cells_NK_WA
T_cells_NKT_WA
T_cells_CD4_BA
T_cells_CD4_Tregs_BA
T_cells_CD8_BA
T_cells_NK_BA
T_cells_NKT_BA
Calprotectin_neutrophil_WA
Degranulating_neutrophil_WA
IFN-resp_neutrophil_WA
IL8_neutrophil_WA
Calprotectin_neutrophil_BA
Degranulating_neutrophil_BA
IFN-resp_neutrophil_BA
IL8_neutrophil_BA
Cancer_WA
Cancer_BA
0.8
0.6
0.4
0.2
0.0
c
Query: WA-Community1
CORO1A
HCST
NKG7
LTB
CD2
CD3E
RAC2
IL2RG
CD14
GLUL
SERPINA1
MMP9
TSPO
PLSCR1
PTPN6
SERPINB1
SELL
GBP2
GBP5
NINJ1
VMP1
MYO1F
ITGAX
BCL2A1
S100A9
S100A8
ALOX5AP
PLAUR
GABARAP
RHOG
GNLY
CCL4
CD7
CST7
ARPC2
SARAF
ZFP36L2
BIRC3
CD48
ARHGDIA
HERPUD1
CLEC2B
PSMB10
SP100
SQSTM1
ATP6V0E1
ISG20
MCL1
IRF1
CLIC1
d
BA-Community1
P=0.0002
Number of Spots
10
8
6
4
2
0
BA
WA
WA-Community1
P=0.039
Number of Spots
20
15
10
5
0
BA
WA
e
-log10P
0
10
20
30
40
50
Blood vessel development
Vasculature development
Angiogenesis
Regulation of signal transduction
Regulation of response to stimulus
Cell migration
Extracellular matrix organization
Epithelial mesenchymal transition
Actin filament-based process
Actin cytoskeleton organization
Regulation of endothelial cell proliferation
Regulation of MAPK cascade
Integrin-mediated signaling pathway
Regulation of Wnt signaling pathway
VEGFA-VEGFR2 signaling pathway
Signaling by Receptor Tyrosine Kinases
TGF-beta signaling
PDGFR-beta signaling pathway
f
-log10P
0
5
10
15
20
25
30
Innate Immune System
Interferon gamma response
Cytokine Signaling in Immune system
Response to cytokine
Leukocyte activation
Neutrophil degranulation
Lymphocyte activation
Allograft rejection
Inteferon alpha response
Inflammatory response
IL6 JAK STAT3 signaling
T cell activation
TNFA signaling via NFKB
Cellular response to type II interferon
IL2 STAT5 signaling
IL12-mediated signaling events
Complement
IL-18 signaling pathway

Figure 5

a



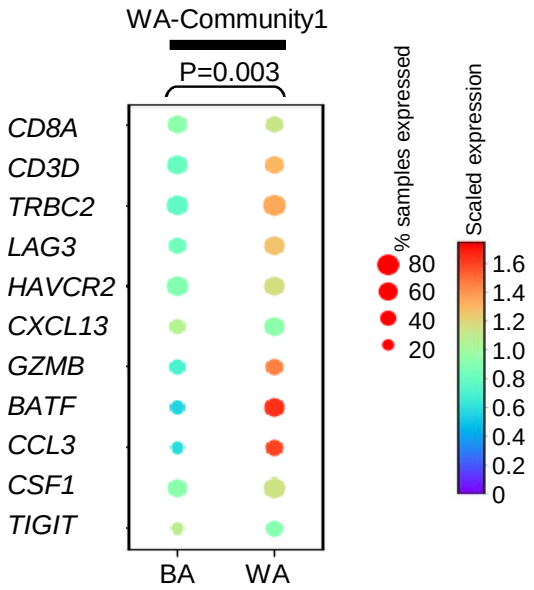
b

Ligands \ Receptors	CAF	Endo	PVL	Mac	Mono	Myo-Epi	Cancer	DC
CAF	3.83	3.91	1.81	1.26	0.25	0.26	0.11	0
Endo	2	7.22	0.89	0.36	0.25	0.33	0.11	0.25
PVL	2.13	5.05	3.4	1.09	0.25	0.51	0.33	0
Cancer	0.23	0.59	0.22	0.07	0	0.05	0.11	0
DC	0.25	0.75	0.75	0.25	0.25	0.25	0	0.25
Mono	0.12	0.48	0.22	0.07	0	0.05	0	0
Myo-Epi	0.78	0.15	0.38	0.07	0	0.05	0	0
Mac	0.33	0.33	0.33	0.17	0	0	0	0

c

Ligands	CAF	Cancer	Mac	Mono	T	DC	Neutro	B
CAF	0	0	0.1	0.1	0.1	0.1	0.1	0
DC	0	0	0.08	0	0.31	0.23	0.15	0.31
PVL	0	0	0	0	0.33	0.33	0	0.33
Endo	0	0	0.8	0.88	0.97	0.47	0.47	0.17
Mac	0.5	0.5	2	2.5	1.25	1	0.25	0.08
T	0	0	0.5	0	2.06	0.06	0.06	0.06
Mono	0	0	0.5	0.5	0.31	0.06	0.31	0.15
Neutro	0	0	0.17	0.25	0.23	0.06	0.23	0.06

d



e

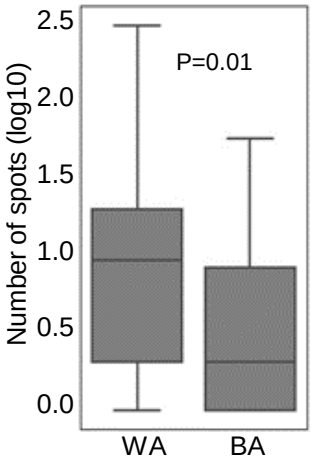


Figure 6

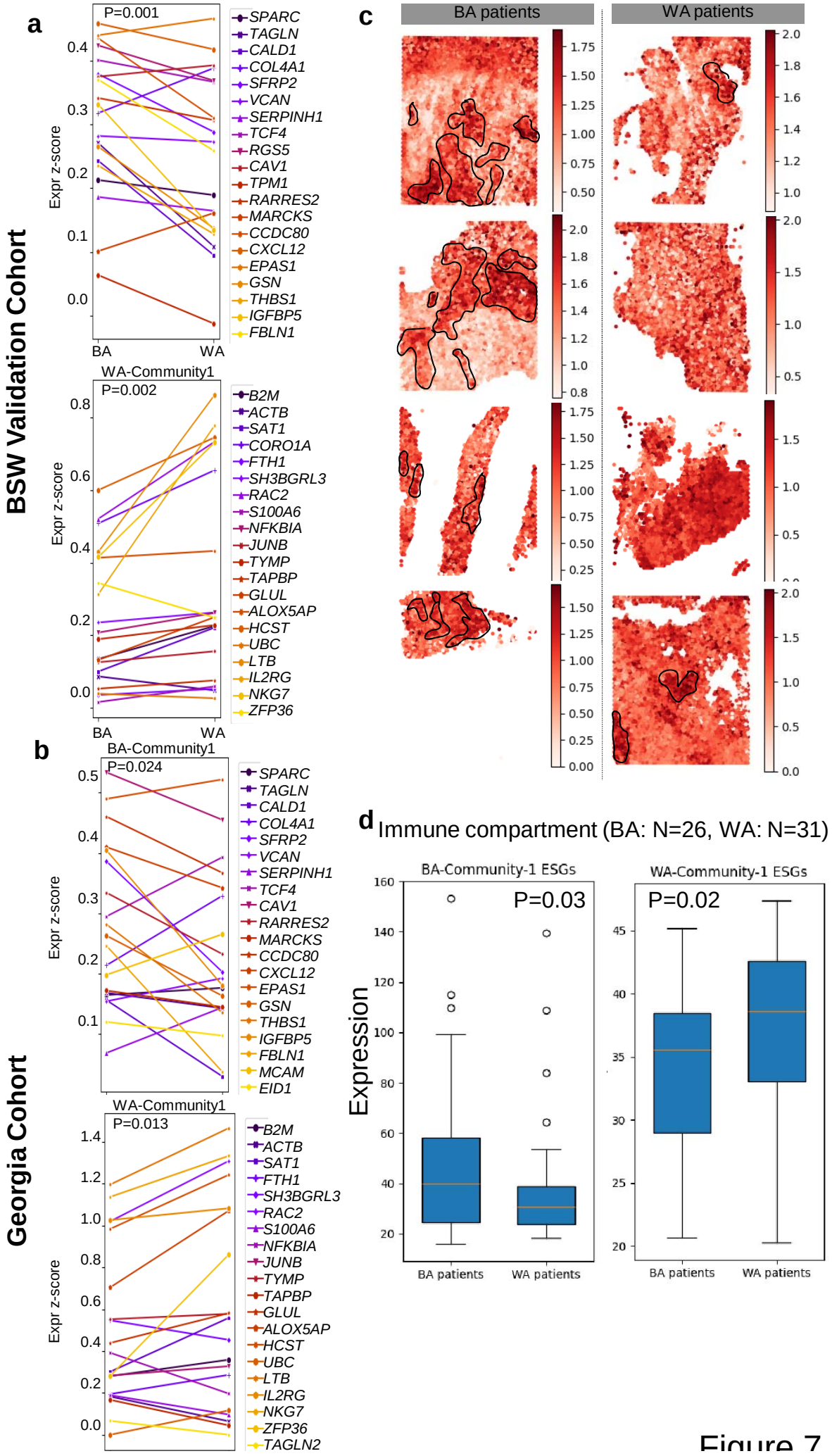
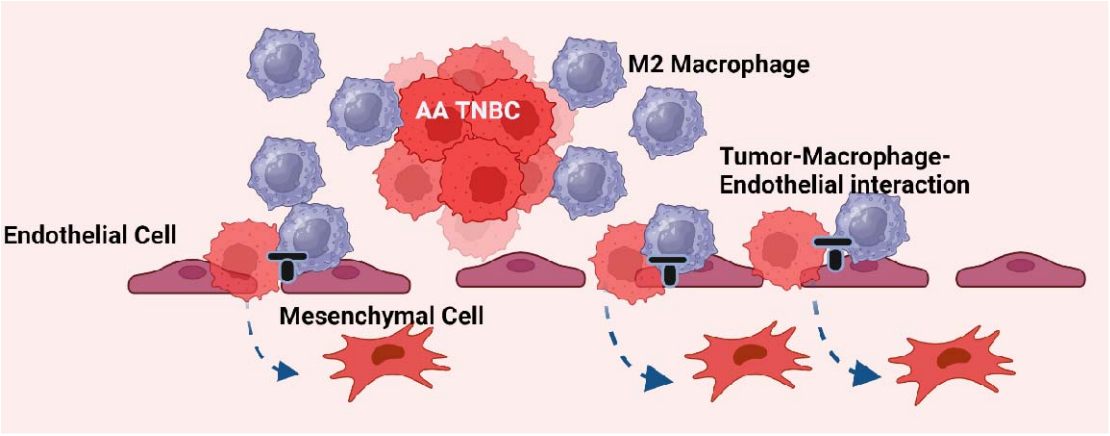

a
BSW Validation Cohort
P=0.001
Expr z-score
BA
WA
SPARC
TAGLN
CALD1
COL4A1
SFRP2
VCAN
SERPINH1
TCF4
RGS5
CAV1
TPM1
RARRES2
MARCKS
CCDC80
CXCL12
EPAS1
GSN
THBS1
IGFBP5
FBLN1
WA-Community1
P=0.002
B2M
ACTB
SAT1
CORO1A
FTH1
SH3BGRL3
RAC2
S100A6
NFKBIA
JUNB
TYMP
TAPBP
GLUL
ALOX5AP
HCST
UBC
LTB
IL2RG
NKG7
ZFP36
b
Georgia Cohort
BA-Community1
P=0.024
SPARC
TAGLN
CALD1
COL4A1
SFRP2
VCAN
SERPINH1
TCF4
CAV1
RARRES2
MARCKS
CCDC80
CXCL12
EPAS1
GSN
THBS1
IGFBP5
FBLN1
MCAM
EID1
WA-Community1
P=0.013
B2M
ACTB
SAT1
FTH1
SH3BGRL3
RAC2
S100A6
NFKBIA
JUNB
TYMP
TAPBP
GLUL
ALOX5AP
HCST
UBC
LTB
IL2RG
NKG7
ZFP36
TAGLN2
c
BA patients
WA patients
d
Immune compartment (BA: N=26, WA: N=31)
BA-Community-1 ESGs
P=0.03
WA-Community-1 ESGs
P=0.02
Expression
BA patients
WA patients

Figure 7

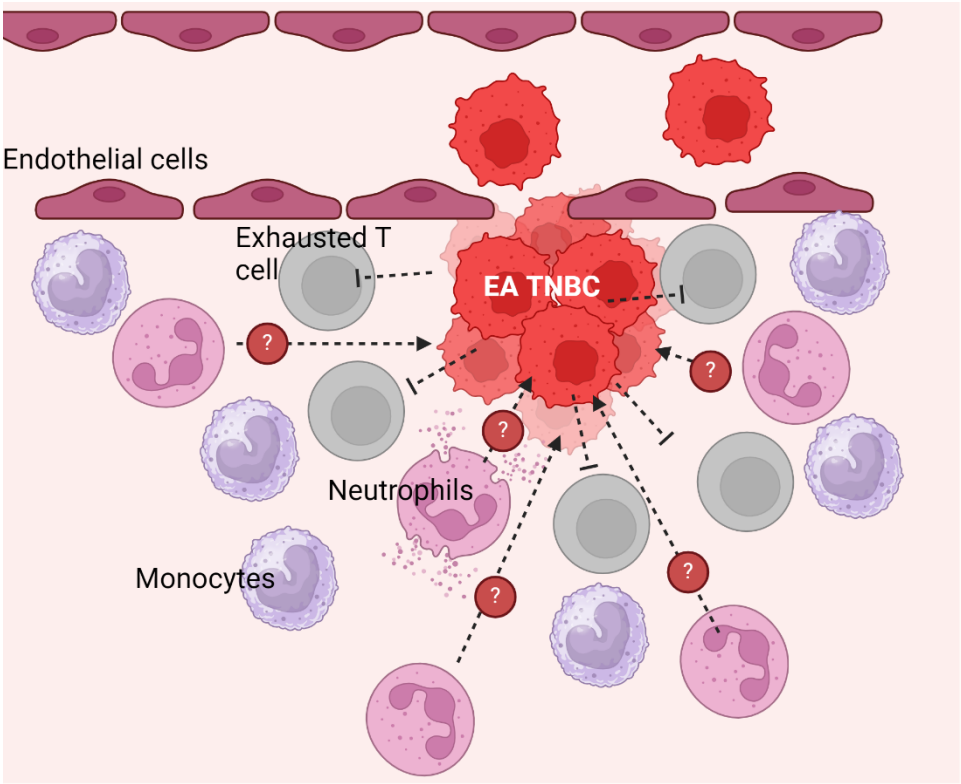
Figure 8

a

BA TNBC Tumor Microenvironment



TMEM doorway

b

WA TNBC Tumor Microenvironment