

AWARD NUMBER: W81XWH-20-1-0577

TITLE: The Influence of Adipogenic Progenitors and Duffy-Null Phenotype on the Normal Breast and Breast Cancer Biology of Women of African Descent

PRINCIPAL INVESTIGATOR: Harikrishna Nakshatri

CONTRACTING ORGANIZATION: Indiana University, Indianapolis, IN

REPORT DATE: July 2022

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Development Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE			<i>Form Approved</i> OMB No. 0704-0188		
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE July 2022		2. REPORT TYPE Annual		3. DATES COVERED 01Jul2021-30Jun2022	
4. TITLE AND SUBTITLE: The Influence of Adipogenic Progenitors and Duffy-Null Phenotype on the Normal Breast and Breast Cancer Biology of Women of African Descent			5a. CONTRACT NUMBER		
			5b. GRANT NUMBER W81XWH-20-1-0577		
			5c. PROGRAM ELEMENT NUMBER		
6. AUTHOR(S) Harikrishna Nakshatri E-Mail: hnakshat@iupui.edu			5d. PROJECT NUMBER		
			5e. TASK NUMBER		
			5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Indiana University School of Medicine, Indianapolis, IN 46202, USA			8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Development Command Fort Detrick, Maryland 21702-5012			10. SPONSOR/MONITOR'S ACRONYM(S)		
			11. SPONSOR/MONITOR'S REPORT NUMBER(S)		
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for Public Release; Distribution Unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT: Breast cancers in women of African ancestry tend to be highly aggressive and metastatic compared to breast cancers in women of European ancestry. Among women of African descent, breast cancers in those who are duffy-heterozygous or duffy-null genotype are even more aggressive. In the year-1 report, as part of aim 1, we showed increased number of a stromal cell type called PZP (PROCR+/ZEB1+PDGFRalpha+) cells in healthy breast tissues of women of African ancestry compared to European or Latina ancestry. PZP cells display fibroadipogenic and mesenchymal stem-like properties. In this report we show that PZP cells can be differentiated into adipocytes or osteocytes under appropriate growth conditions. In co-culture conditions, PZP:epithelial cell communication resulted in luminal epithelial cells acquiring basal/stem cell characteristics and increased expression of IL-6 suggesting the impact of this communication on the microenvironment and breast epithelial hierarchy. Consistent with this possibility, the level of phospho-STAT3, which is a downstream target of IL-6, was higher in the normal and cancerous breast tissues of women of African ancestry compared to women of European ancestry. PZP cells transformed with HRas ^{G12V} ± SV40-T/t antigens generated metaplastic carcinoma in NSG mice suggesting that these cells could be the cell-of-origin of metaplastic breast cancers. Incidence of metaplastic carcinoma is higher in women of African ancestry compared to women of European ancestry. Collectively, these results identify a stromal cell component that could influence the biology of breast cancer in AA women. When PZP cells are implanted along with transformed breast epithelial cells, they localize like stroma in a tumor. These observations now allow us to determine whether their presence decrease sensitivity of tumors to chemotherapy. For aim 2, we generated transformed cell lines from a donor who is duffy heterozygous and two from duffy-null. We observed elevated levels of phosphorylated STAT3 and NF-kappaB in transformed cells from duffy-mutant carriers compared to duffy-wild type donors. Thus, while breast tissues from women of African ancestry in general have higher phospho-STAT3, its levels further increase in women who are additionally duffy-heterozygous or duffy-homozygous. Thus, drugs targeting IL6 and/or STAT3 may be effective on tumors of women of African descent.					
15. SUBJECT TERMS None listed.					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER (include area code)
Unclassified	Unclassified	Unclassified	Unclassified	74	USAMRDC

TABLE OF CONTENTS

	<u>Page</u>
1. Introduction	4
2. Keywords	4
3. Accomplishments	4
4. Impact	11
5. Changes/Problems	12
6. Products	12
7. Participants & Other Collaborating Organizations	13
8. Special Reporting Requirements	N/A
9. Appendices	Submitted manuscript

1 Introduction: African American (AA) women suffer higher mortality from triple negative breast cancer (TNBC) than Caucasian women. By contrast, breast cancer in Hispanic and Native American women is less prevalent and these women have a better outcome. Whether worst outcome in AA women is due to an increased incidence of TNBC or unique biological factors that promote aggressive biology is an important but unresolved challenge in cancer disparity research. Recent studies have also demonstrated that duffy heterozygous/null phenotype, which is most commonly inherited in people of sub-Saharan ancestry, is associated with aggressive breast cancer biology in African American women, which provided first hint to the impact of genetic ancestry influencing the biology of breast cancer. We previously reported that the normal breasts in AA women are enriched for cells that express ZEB1, an epithelial to mesenchymal transition associated transcription factor. In in vitro studies, these cells also expressed PROCRA and PDGFRalpha (PDGFR α). Therefore, we labelled these cells as PZP cells. One aim of the proposal is focused on further characterizing these cells. In our preliminary studies, we had observed that breast epithelial cells from duffy-heterozygous and duffy-null carriers have higher levels of activated cMET signaling. Since cMET pathway is associated with chemotherapy resistance and metastasis, the second aim is focused on testing the hypothesis that duffy-heterozygous and duffy-null phenotype confers aggressive metastatic property to breast cancer through cMET pathway.

2 Keywords: Breast cancer, PROCRA, ZEB1, PDGFRalpha, Duffy, drug resistance, metastasis.

3 Accomplishments:

Specific Aim 1. To investigate the intrinsic and extrinsic effects of PZP cells enriched in AA women on tumorigenesis	Cell lines and cohorts	Timeline Months	Current status
Major Task 1. Trans-differentiating properties of PZP cells of the normal breast:			
Subtask 1 (6.1.2): Characterize PZP cells for trans-differentiation upon treatment with various ligands of PDGFR.	KTB40 and KTB42. Note that we have four other similar cell lines in stock but only two will be used (KTB32, KTB53, KTB55 and KTB59)	1-5	Completed
Subtask 2 (6.1.3): Determine how cancer cell-induced factors modulate trans-differentiation	Cancer cell lines MCF-7, T47-D, SK-BR-3, BT-474, HCC1937, MDA-MB-468, SUM149PT, HCC70, HCC1187, DU4475, BT-549, HS578T, MDA-MB-231 and MDA-MB-436. PZP lines KTB40 and KTB42	2-12	Work in progress
<i>Milestone Achieved: demonstrated that PZP cells upon trans-differentiation into fibroblasts, adipocytes or osteoblasts alter tumor progression</i>			
Major Task 2 (6.1.4): Determine whether cancer cell-derived factors cause PDGFRα isoform switching or receptor dimerization to enhance differentiation to fibroblasts			
Subtask 3: Determine the effects of conditioned media from cell lines on PDGFR α isoforms in PZP cells.	Conditioned media from cancer cell lines listed in subtask 2	13-15	Early stages
<i>Milestones achieved: Conditioned media from select cell lines block the generation of decoy PDGFRα receptor.</i>			
Major Task 3 (6.1.5). The effects of trans-differentiated PZP cells on invasive and drug sensitivity of various breast cancer cell lines:			
Subtask 4: Generate GFP labeled PZP cells and trans-differentiate cells into fibroblasts, adipocytes, or osteoblasts	Parental and transdifferentiated KTB40 and KTB42	10-14	GFP+ cells have been generated
Subtask 5: Perform co-culture experiments to determine the influence of various trans-differentiated PZP cells on cancer cell invasion	Cancer cell lines described in subtask 2 plus oncogene transformed breast epithelial cell lines from two each of Caucasian, Hispanic and African American women, parental	14-20	Identified IL-6 as the factor induced in PZP cells in response to epithelial cell interaction

	and transdifferentiated KTB40 and KTB42		
Subtask 6: Perform drug sensitivity studies of co-cultured cells	Cell lines described in subtask 5	20-24	Ongoing but focus may shift to STAT3 inhibitors due to current data
<i>Milestone achieved: PZP cells, depending on type of trans-differentiation altered sensitivity of cancer cells to chemotherapy and influenced their invasive properties</i>			
Major Task 4 (6.1.6). The effects of trans-differentiated PZP cells on growth and metastatic properties of breast cancer cells <i>in vivo</i>:			
Subtask 7: Determine the effects of trans-differentiated PZP cells on growth and metastasis of cancer cells	Cell line: One transformed cell line derived from epithelial cells of Caucasian and another from African American women, MCF-7 and MDA-MB-468 cell line. Parental and transdifferentiated KTB40 or KTB42 Animal: NSG mice. Cohort size: 336 (4 tumor lines with 7 types of PZP cells, 12 animals per group)	20-28	Partially done. PZP cells line tumor cells <i>in vivo</i> and recapitulate human breast tumor stroma <i>in vivo</i>
<i>Milestone(s) Achieved: Demonstrated the role of PZP cells in growth and metastasis of tumor cells in vivo.</i>			
Major Task 5 (6.1.7): Determine whether the PDGFRα inhibitor nilotinib can reduce PZP cells activity and increase chemosensitivity			
Subtask 8: Implant tumor cells with PZP cells and determine sensitivity of tumors to nilotinib \pm paclitaxel.	Cell lines: A transformed cell line derived from African American women plus one PZP cell line. Mice: NSG Cohort: 96 (48 animals without PZP cells and 48 with PZP cells. Four treatment groups, 12 per group. Control, nilotinib, paclitaxel and both drugs.	22-32	To be initiated next month
<i>Milestones Achieved: Nilotinib increases sensitivity of tumors to chemotherapeutic drug paclitaxel</i>			
Major Task 6 (6.1.8): Intrinsic tumorigenic properties of PZP cells:			
Subtask 9: Transform parental PZP or PZP cells trans-differentiated into epithelial cells and determine their tumorigenic properties	Cell lines: Three PZP cell lines transformed using two different oncogenes- total six cell lines. Mice: NSG. Cohort: Six cell lines, 12 animals per cell line. Total 72 animals.	20-28	Already complete. Determined that PZP cells are cell of origin of metaplastic breast cancers
<i>Milestone achieved: Epithelial trans-differentiated PZP cells generate tumors with distinct characteristics</i>			
Specific Aim 2. To investigate the role of hyperactive c-MET signaling in breast tumorigenesis under the duffy-null background			
Major Task 7 (6.2.2). <i>In vitro</i> characterization of immortalized and transformed duffy-wild type, duffy-heterozygous and duffy-null breast epithelial cell lines:			
Subtask 10: Generate immortalized and transformed cell lines with duffy wild type, duffy-heterozygous and duffy-null background	Cell lines: Two each of duffy wild type, duffy-heterozygous and duffy-null immortalized cell lines; each transformed with two sets of oncogenes. 18 cell	1-12	Generated immortalized and transformed variants of one wild type and one duffy-heterozygous breast epithelial cells.

	lines (six immortalized and 12 transformed)		
Subtask 11: Determine CCL2 and CXCL8 mediated signaling in all cell types by proteomics and RNA-seq	Cell lines described in subtask 10	12-24	Did not observe much of an effect of CCL2 and IL-8 on signaling but basal NF-kB and STAT3 signaling are higher under duffy-null and/or heterozygous background
<i>Milestone achieved: Distinct CCL2 and CXCL8 signaling in duffy-heterozygous and duffy-null epithelial cells compared to duffy wild type cells.</i>			
Major Task 8 (6.2.3): Stem cells properties of duffy-null/heterozygous cells			
Subtask 12: Determine tumor-initiating capacity of transformed duffy-wild type, duffy-heterozygous and duffy-null cell lines	Cell lines: two each of transformed cell lines in duffy-wild type, duffy-heterozygous and duffy-transformed background. Mice: NSG Cohort: 3 cell dilutions, 5 per dilution, 2 cell lines per category and three categories, Total 90 animals	24-28	Although Duffy-null cells were difficult to immortalize, transformed cells are highly aggressive and formed tumors in NSG mice. Serial dilution studies are underway.
<i>Milestones achieved: Transformed cell lines under duffy-null background have higher tumor-initiating capacity.</i>			
Major Task 9: The influence of duffy phenotype on drug sensitivity			
Subtask 13: Determine the sensitivity of immortalized and transformed cells to doxorubicin, paclitaxel, and cisplatin	Cell lines: two each of immortalized and transformed cell lines under duffy-wild types, duffy-heterozygous and duffy-null background	12-17	Did not show difference in sensitivity to paclitaxel at higher doses but studies at lower doses are underway.
<i>Milestones achieved: Transformed cells under duffy-null and duffy-heterozygous background are resistant to chemotherapy.</i>			
Major Task 10 (6.2.5). The ability of c-MET inhibitors to sensitize transformed duffy-null cells to chemotherapy:			
Subtask 14: Determine sensitivity of duffy wild type, duffy heterozygous, and duffy null immortalized and transformed cells to crizotinib or carbozantinib with and without chemotherapeutic drugs <i>in vitro</i>	Cell lines: same as described in subtask 13	13-20	Yet to be initiated
Major Task 11 (6.2.6). In vivo effects of duffy-null/heterozygous phenotype on tumorigenicity, metastasis and drug sensitivity			
Subtask 15: Measure tumor growth rate and lung metastasis of transformed duffy-wild type, duffy-heterozygous and duffy-null cell derived tumors	Cell lines: Transformed cell lines from each category but transformed by two different sets of oncogenes Mice: NSG Cohort: 12 transformed cell lines, 12 animals per cell line and 144 animals	20-26	Yet to be initiated
Subtask 16: Measure the effects of Crizotinib with and without chemotherapy on tumor growth and metastasis	Cell lines: One each of transformed cell line under duffy-wild type, duffy-heterozygous and duffy-null background. Mice: NSG mice. Cohort: 72 animals per cell line, each cell line in 6 groups (control, crizotinib, chemo-1, chemo-2,	27-34	Yet to be initiated.

	crizotinib plus chemo-1, crizotinib plus chemo-2.) Three cell lines. 216 animals		
<i>Milestone achieved: Duffy-heterozygous or duffy-null compared to duffy wild type transformed epithelial cells are more sensitive to crizotinib or carbozantinib ± chemotherapy both in vivo and in vitro</i>			
Subtask 17: Write Manuscript based on results of aims 1 and 2.		12-16 and 33-36	

Specific aims:

Aim 1: To investigate the intrinsic and extrinsic effects of PZP cells enriched in AA women on tumorigenesis.

Outcome: The majority of data presented in first year report and part of the submitted manuscript are not included in the report. The manuscript titled “Influence of genetic ancestry on breast stromal cells provides biologic basis for increased incidence of metaplastic breast cancer in women of African descent” has been submitted for peer-review (currently to Nature Communications) and the pre-print is available through <https://doi.org/10.1101/2022.07.14.500115>. bioRxiv

Results not described in the report but part of last year’s report include:

- 1) Generation of PZP cell lines from breast tissues of healthy women of African Ancestry (AA)
- 2) PZP cells express lobular fibroblast markers
- 3) PZP cells are enriched in the normal breasts of women of African ancestry compared to women of European and Latina ancestry
- 4) Modeling the effects of PZP cells on breast tumorigenesis
- 5) PZP-epithelial cell interaction alters the expression of specific genes including transgelin and IL6.
- 6) Potential role of PZP cells in immune cell modulation in the microenvironment.
- 7) Transformed PZP cells generate metaplastic carcinoma.

Trans-differentiating properties of PZP cells

In order to determine to what extent PZP cells show similarity to adipogenic progenitors described in the mouse mammary gland¹ and fibroadipogenic progenitors (FAPs) described in the skeletal muscle² with respect to trans-differentiation into adipogenic and osteogenic lineages, PZP cells were subjected to adipogenic and osteogenic differentiation growth conditions. Indeed, PZP cells differentiated into adipocytes (Fig. 1A) and the differentiated cells expressed adipocyte differentiation marker PPAR γ (Fig. 1B). Adipogenic trans-differentiation capabilities of PZP cells was further confirmed through RUNX1 overexpression (Fig. 1C). PZP cells cultured in osteogenic media for 21 days showed osteogenic differentiation as evident from mineralization of matrix with Ca²⁺ and positive alizarin red staining (Fig. 1D).

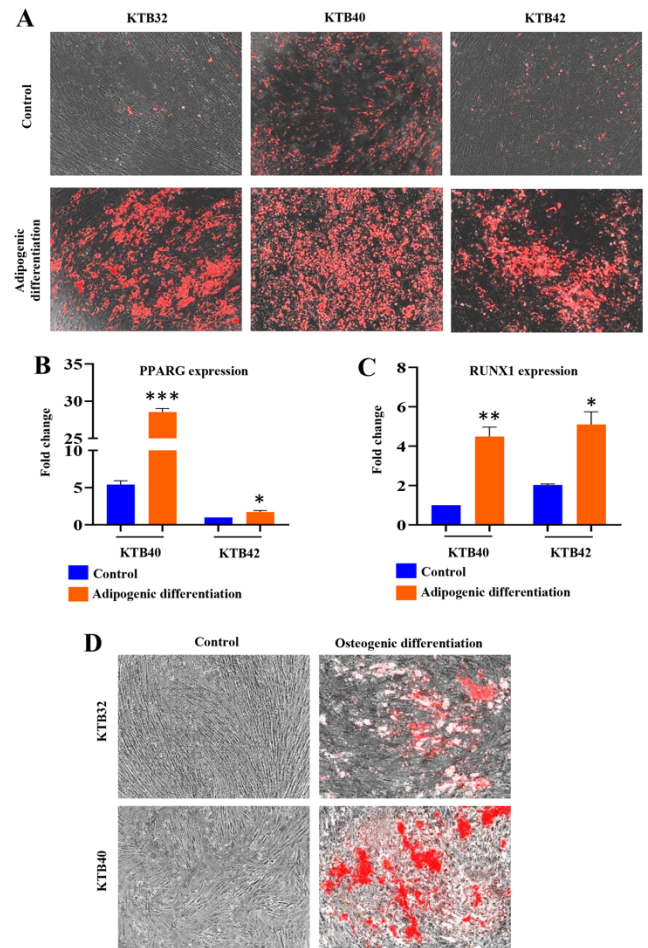


Figure 1: Trans-differentiating properties of PZP cells. (A) PZP cells undergo adipogenic differentiation under appropriate growth condition. Neural lipids stain red upon Oil Red-O staining. (B) Adipogenic differentiated PZP cells show elevated level of adipocytes differentiation marker PPAR γ . (C) Adipogenic differentiation of PZP cells was confirmed through RUNX1 overexpression. (D) PZP cells undergo osteogenic differentiation under appropriate growth condition. Mineralization of matrix with Ca²⁺ was visualized by alizarin red staining. *p<0.05, **p<0.01, ***p<0.001 by ANOVA.

PZP cells cultured in osteogenic media for 21 days showed osteogenic differentiation as evident from mineralization of matrix with Ca²⁺ and positive alizarin red staining (Fig. 1D).

Thus, PZP cells enriched in the breasts of women of AA could correspond to multipotent cells that can transdifferentiate into different cell types based on environmental cues.

The effects of breast cancer: PZP cell interaction on trans-differentiation of epithelial luminal progenitor cells

In order to further investigate the intercellular communication between PZP and epithelial cells, we generated stable tdTomato-labeled KTB40 and KTB42 cell lines using pCDH-EF1-Luc2-P2A-tdTomato lentivirus. To determine whether co-culturing alter the phenotype of epithelial cells, tdTomato-labeled PZP (KTB40, KTB42), epithelial (KTB34, KTB39), and co-cultured PZP and epithelial cells were analyzed by flow cytometry using CD49f and EpCAM, which can differentiate luminal mature (CD49f-/EpCAM+), luminal progenitor (CD49f+/EpCAM+) and basal/stem (CD49f+/EpCAM-) cells. Co-cultured epithelial cells displayed an increase in CD49f⁺/EpCAM^{-/low} and reduced CD49f^{-/low}/EpCAM^{+/high} cell populations in co-culture condition (Fig. 2A-C). tdTomato-Red⁺ PZP cells on their own were CD49f-/EpCAM-, a small fraction of these cells displayed a unique CD49f⁺/EpCAM⁺ phenotype under co-culture condition (Fig. 2D and E). These results suggest that similar to mouse mammary fibroadipogenic cells¹, PZP cells can potentially acquire epithelial characteristics under specific conditions. Our attempts of characterize CD49f⁺/EpCAM⁺ trans-differentiated PZP cells, collected via flow cytometry, were not successful because of dominant growth of few contaminating PZP cells in the sorted population of cells.

Normal breast and/or breast tumors of AA women have higher levels of phospho-STAT3.

Since PZP-epithelial cell interaction resulted in elevated IL-6 expression *in vitro* and PZP cells are present at higher levels in the normal breasts of women of African ancestry (AA), it is expected that signals downstream of IL-6 should be higher in the breast tissues of AA women compared to European ancestry (EA) women. Phosphorylation of STAT3 at residue S727 is one the major downstream signal upon binding of IL-6 to its receptor and S727 phosphorylated STAT3 promotes mesenchymal to epithelial transition of early disseminated cancer cells³. We used phospho-STAT3 as a surrogate marker to determine IL-6 activity in the normal breast tissues and breast tumors of AA and EA women. Although cells in the KTB-normal contained lower levels of phospho-STAT3 as we could estimate on positivity, positivity was still higher with healthy breast tissues of AA women compared to EA women (Fig. 3A-C). A modestly higher levels of phospho-STAT3 were also observed in tumors of AA compared to EA women (Fig. 3C). Consistent with elevated levels of PZP cells in normal

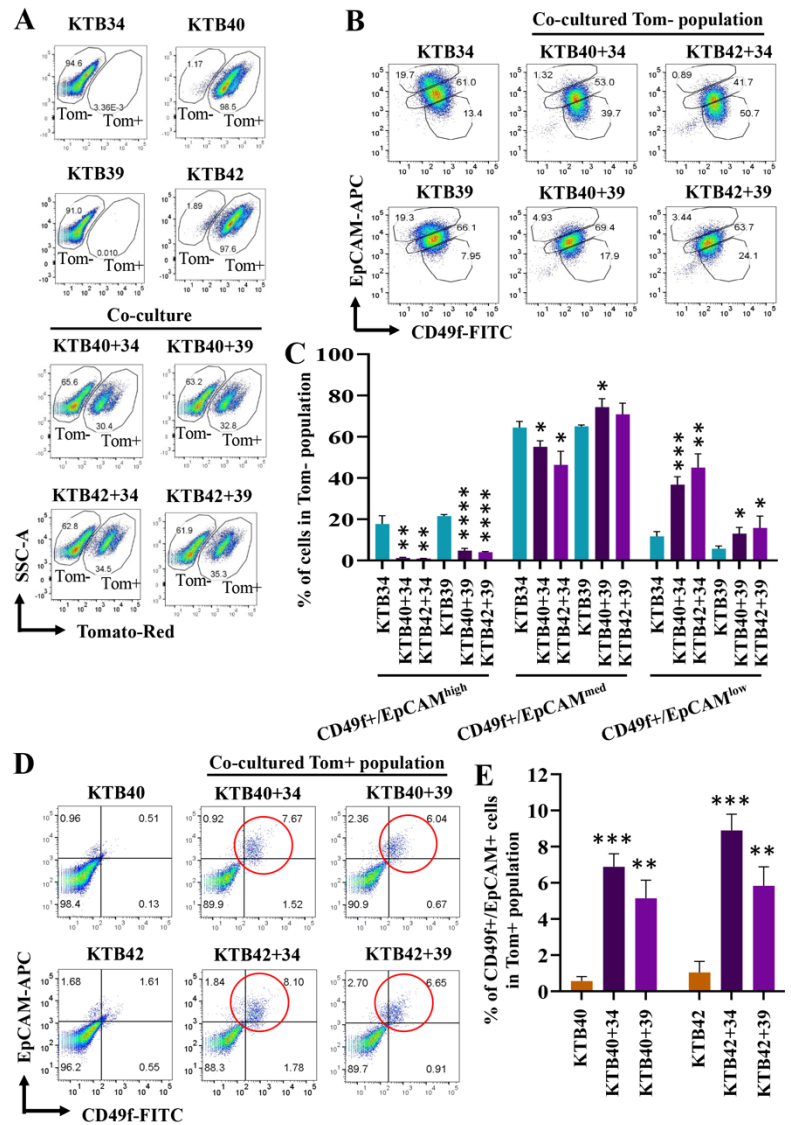


Figure 2: The effects of PZP cells on trans-differentiation of epithelial luminal progenitor cells. (A) Gating of tomato- (Tom-) and tomato+ (Tom+) populations of KTB34, KTB39, and co-cultured KTB40/KTB42 and KTB34/KTB39 cell lines. (B) CD49f and EpCAM staining patterns of KTB34, KTB39, and co-cultured KTB40/KTB42 and KTB34/KTB39 cell lines. (C) Quantification of CD49f⁺/EpCAM^{high}, CD49f⁺/EpCAM^{med}, and CD49f⁺/EpCAM^{low} populations in Tom- cell population. (D) CD49f and EpCAM staining patterns of KTB40, KTB42, and co-cultured KTB40/KTB42 and KTB34/KTB39 cell lines. Only Tom+ population was analyzed. (E) Quantification of CD49f⁺/EpCAM⁺ population among Tom+ population. *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001 by ANOVA.

adjacent to tumors (NATs) compared to KTB-normal in EA women (last report and the enclosed manuscript), phospho-STAT3 levels were higher in NATs compared to KTB-normal of EA women (**Fig. 3D**). Thus, genetic ancestry dependent variability in levels of stromal PZP cells has an impact on normal and cancerous breast biology.

PZP cells recapitulate stromal structures when co-implanted with tumors cells in NSG mice.

We implanted H-Ras^{G12V} plus SV40-T/t antigen transformed epithelial cell line KTB34 either alone or along with PZP cell lines KTB40 or KTB42. KTB40 and KTB42 on their own did not develop into tumors (**Fig. 4A**). Transformed KTB34 cells formed slow growing tumors and inclusion of PZP cells accelerated tumor growth, although modestly. With small number of animals used, results are not significant and we will be repeating these experiments. As per histology, KTB34-derived tumors were metaplastic cystic adenocarcinomas whereas tumors with PZP cells showed abnormal stroma and/or invaded muscles (**Fig. 4B**). Staining for ZEB1 revealed PZP cells encircling tumors. Tumors derived from transformed KTB34 cells alone contained ZEB1+ stroma, which suggests the infiltration of ZEB1+ mouse stromal cells. Tumor cells but not ZEB1+ stromal cells stained for phospho-STAT3, which clearly indicates tumor-specific activation of STAT3 signaling pathway. Together, these results suggest that PZP cells alter the phenotype of tumors *in vivo*.

Specific aim 2: To investigate the role of hyperactive c-MET signaling in breast tumorigenesis under duffy-null background

Progress: We had difficulty in immortalizing epithelial cell lines from duffy-homozygous carriers. However, we were able to generate two transformed cell lines from duffy-null carriers. An immortalized and transformed variant of duffy-heterozygous carrier have been created. KTB8 and KTB39 represent duffy-wild type, KTB41 represents duffy-heterozygous, KTB406 and KTB470 are duffy-homozygous. All are derived from breast tissues AA women. However, immortalized variants of KTB406 and KTB470 proved difficult to maintain. Therefore, all studies included cells overexpressing H-RAS^{G12V} and SV40-T/t antigens. We examined the expression levels of ACKR1 whose loss/reduced expression is associated with duffy phenotype in these cell lines to confirm that breast epithelial cells of duffy heterozygous and duffy-homozygous carriers express lower levels of ACKR1 compared to breast epithelial cells from duffy-wild type donors (**Fig. 5**).

Transformed duffy-heterozygous and duffy-homozygous cells display elevated phospho-STAT3 and phospho-p65 subunit of NF-κB: Duffy mutations are expected to increase signaling by chemokines/cytokines such as IL-8 and CCL2. Other chemokines/cytokines affected by duffy mutations include CXCL1, CXCL2, CXCL3, CXCL5, CXCL6, CXCL11, CCL5, CCL7, CCL11, CCL13, CCL14 and CCL17⁴. Using transformed cells we initiated studies to determine whether duffy phenotypes affect signaling downstream of IL-8 and CCL2. Although we were not able to see the effects of duffy phenotype on signaling by IL-8 and CCL2, basal activity of specific signaling pathways were higher in duffy-homozygous or duffy heterozygous cells. We maintained cells under serum- and growth factor-free conditions for four hours and then treated with IL-8 or CCL2 for one hour. Cell extracts were examined for the levels of phosphorylated-STAT3 (Y705), phosphorylated AKT (S473),

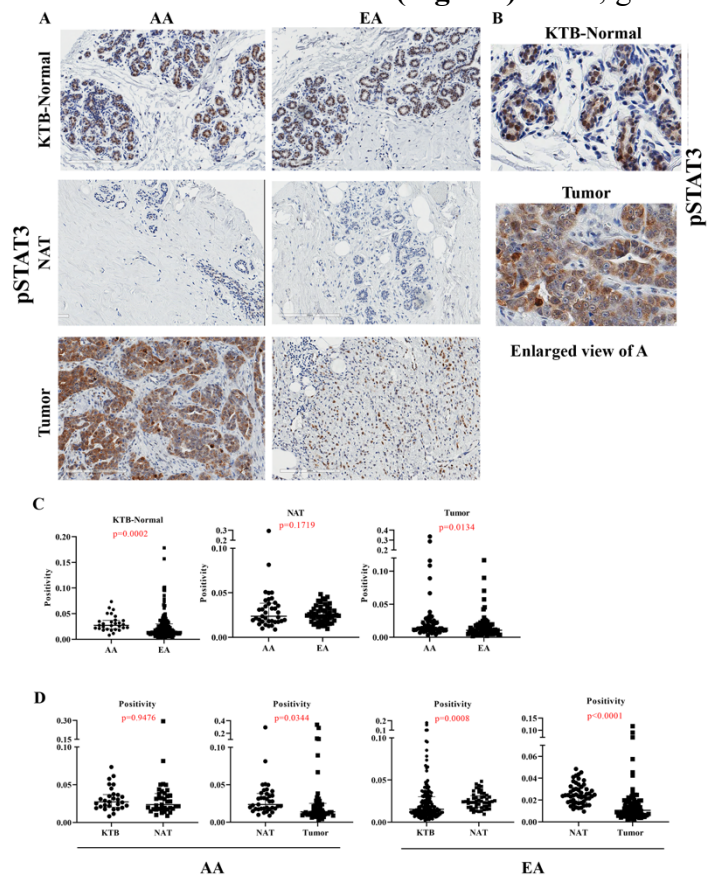


Figure 3: Phospho-STAT3 levels in KTB-normal, NATs, and breast tumors. (A) Representative IHC of pSTAT3 in KTB-normal, NATs, and/or tumors of women of AA, and EA. (B) Enlarged view of pSTAT3 expression in KTB-normal and tumor. (C) Differences in pSTAT3 positivity between KTB-normal, NATs and tumors of AA and EA women (D) Differences in pSTAT3 positivity between KTB-normal and NATs and between NATs and tumors of AA and EA women.

phosphorylated-ERK (T202/Y204), and phospho-p65 (S536) levels. We observed consistently higher levels of phospho-STAT3 and phospho-p65 in transformed Duffy-heterozygous and duffy-homozygous null cells compared to duffy wild type cells (**Fig. 6**). IL-8 and CCL2 treatment did not cause any changes in their levels. For unknown reasons, expression levels of total AKT, ERK1/2, STAT3 and p65 are also higher in duffy-homozygous cells compared to wild type cells. We note that duffy-heterozygous KTB41 and duffy wild type KTB8 expressed similar levels of STAT3 but heterozygous expressed higher levels of pSTAT3. These studies are being repeated under various conditions to ensure the above results are reproducible and meaningful to understand the effects of duffy phenotype on breast biology. These results suggest the following: 1) Epithelial cells of duffy carriers display constitutive NF- κ B activity; 2) While normal breast epithelial cells of AA women have higher levels of phospho-STAT3 compared to EA women, level is even higher in duffy carrier AA women.

To independently verify the above results, we implanted transformed KTB8 (wild type), KTB41 (heterozygous), KTB406 (null) and KTB470 (null) cells into NSG mice and generated tumors. Cell lines were established from primary tumors and/or lung metastasis. Due to bacterial contamination while culturing, we lost tumor-derived KTB41 cell line. However, lung metastasis-derived cell line was available (MKTB41 R+S). Under full growth media (with FBS and other growth factors including EGF and insulin), we observed elevated pERK levels in duffy null cells compared to wild type cells (**Fig. 7**). pAKT levels were similar across different cell types. As we reported in the original grant application, we also observed elevated phospho-cMET (Y1349) in transformed duffy-heterozygous and duffy-homozygous cells compared to duffy-wild type cells. Thus, specific basal signaling networks are altered in duffy-deficient cells compared to duffy-wild type cells.

Duffy mutation is associated with altered parainflammatory signals: Parainflammation is defined as elicitation of a low-grade inflammatory response by the epithelium through expression of cytokines/chemokines and these secreted factors causing chronic inflammation⁵. The presence of this chronic inflammation predisposes for cancer⁵. We used cytokine/chemokine array, which measures 104 cytokines/chemokines secreted by cells. Half-a-million cells were plated for 24 hours and then switched over to FBS and growth factor free media. Conditioned media was collected after 24 hours in serum and growth factor-free media and the expression levels of cytokines/chemokines were measured. We used four cell lines; tumor-derived KTB8 (TKTB8 R+S, duffy-

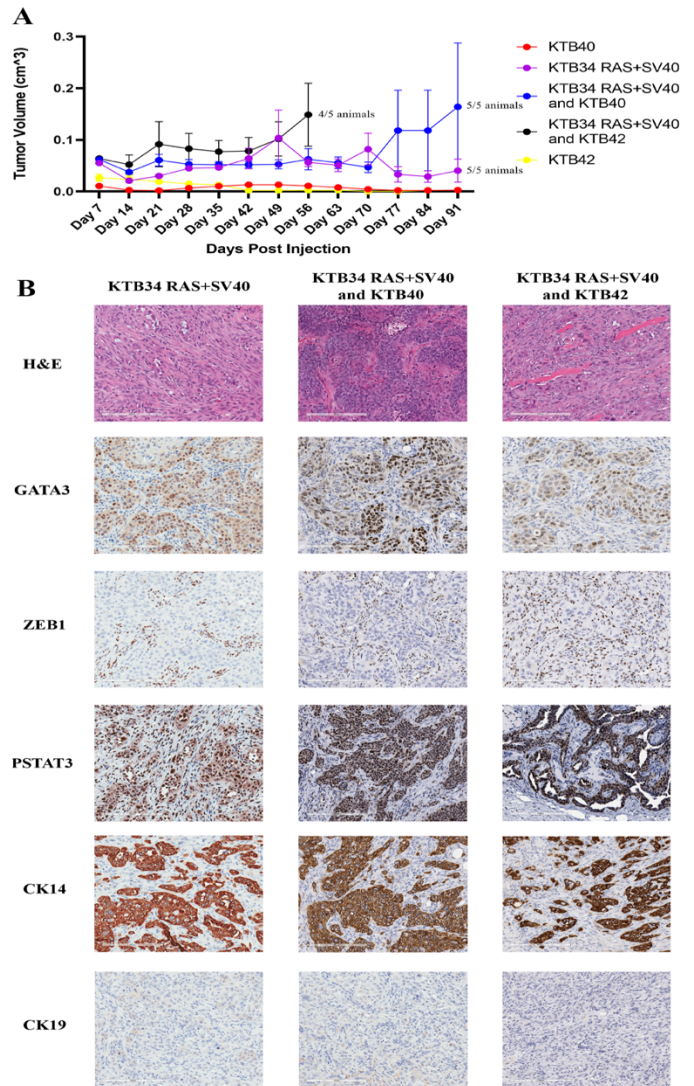


Figure 4: PZP cells alter the characteristics of epithelial cell-derived tumors: A) Growth patterns of tumors developed from transformed KTB34 cells with and without PZP cell co-injection. Tumors appear to grow faster with PZP cells although differences did not reach statistical significance. **B)** Injected PZP cells reconstitute the stroma of tumors based on ZEB1 staining pattern. Also note that pSTAT3 is present only in tumor cells but not stromal cells.

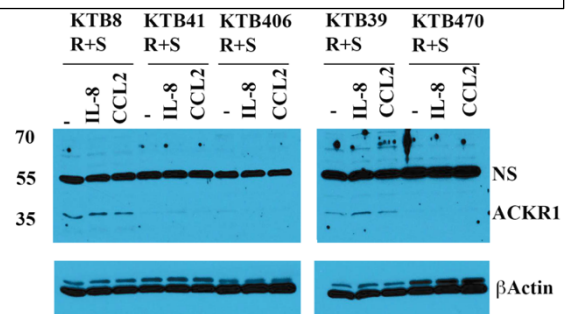


Figure 5: Duffy-heterozygous and duffy-homozygous transformed breast epithelial cells express very little ACKR1 protein compared to transformed breast epithelial cells from duffy wild type donor. NS= non-specific.

wild type from AA)), KTB34 (TKTB34 R+S, from EA), KTB41 (MKTB41 R+S, lung metastasis-derived, duffy-

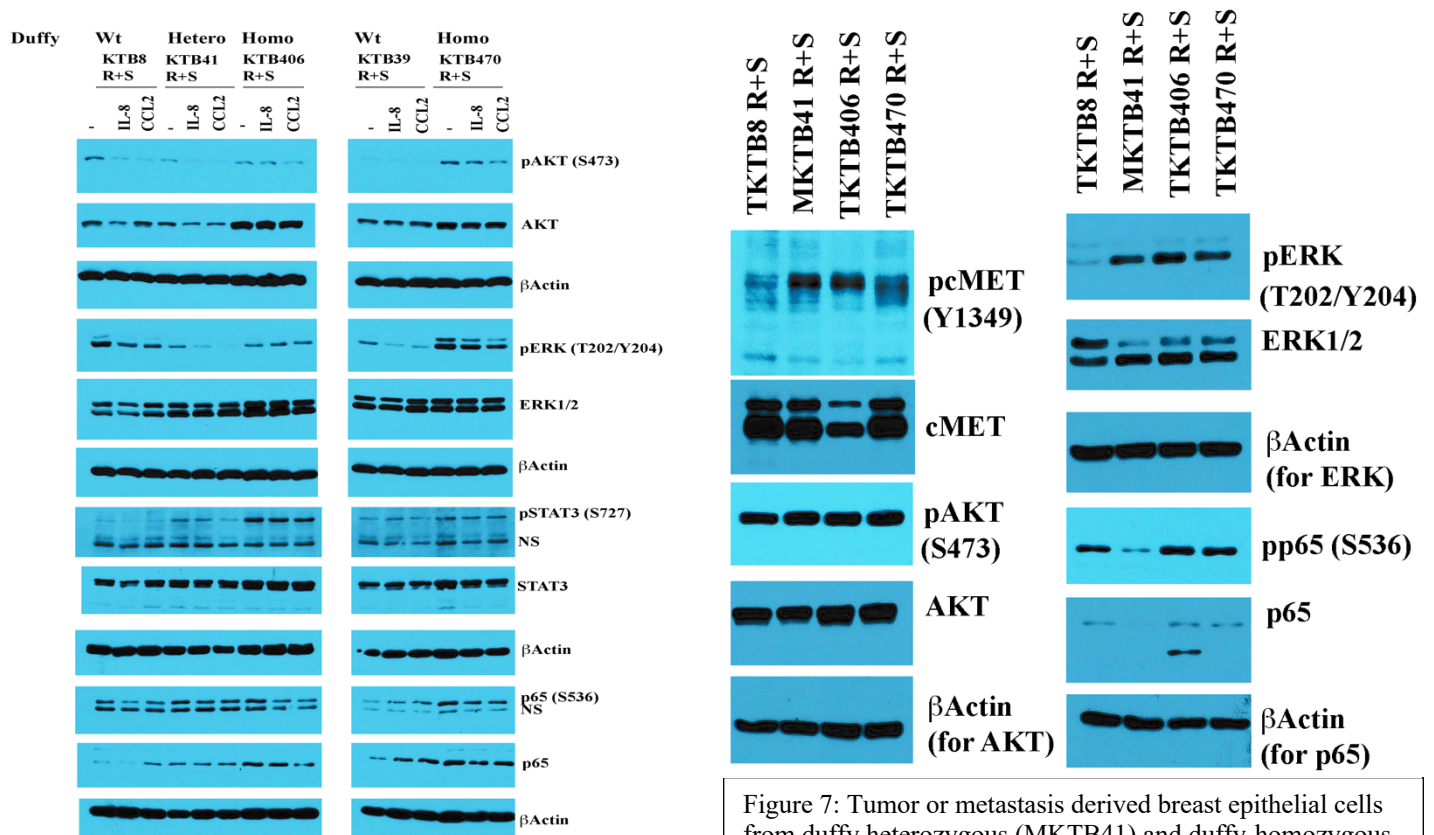


Figure 6: Distinct differences in signaling networks in transformed duffy-heterozygous/homozygous cells compared to duffy wild type cells. Indicated cells were grown in serum-free and growth factor-free media for four hours and then treated with either IL-8 (500 ng/ml) or CCL2 (200 ng/ml) for one hour and then examined for various signaling molecules. Basal expression of several of these signaling molecules and consequently their phospho-variants are higher in duffy-heterozygous/homozygous variants compared to wild type cells. We are currently working to determine whether these are artifacts or true differences by using different culture conditions.

Figure 7: Tumor or metastasis derived breast epithelial cells from duffy heterozygous (MKTB41) and duffy-homozygous (TKTB406 and TKTB470) display elevated cMET and ERK signaling compared to tumor-derived cells from duffy-wild type (TKTB8) donor.

cytokines. This observation needs to be reproduced using additional EA and AA-derived cell lines as GDF15 is associated with cancer-induced systemic effects including cachexia, whereas VEGF increases angiogenesis^{6, 7}. IGFBP2 could be responsible for elevated constitutive NF-κB in duffy-null and Duffy-heterozygous cells as IGFBP2 is a known inducer of NF-κB⁸. Overall, these results clearly suggest the influence of duffy phenotype on parainflammatory status of epithelial cells. Next year will focus on further evaluation of parainflammatory phenotype using *in vivo* models.

With respect to drug sensitivity, we have examined sensitivity of TKTB8 R+S, TKTB34 R+S, MKTB R+S, and TKTB406 R+S cells to paclitaxel. At 0.5, 1 and 2 μM concentration tested, we did not find any difference in sensitivity (**Fig. 9**). Sensitivity at lower concentrations is currently being tested.

4 Impact: Outcome of this proposal will form the basis of a project to be submitted to NCI as a SPORE in January 2023. The project will focus on further developing ZEB1 and phospho-STAT3 as biomarkers of tumor phenotype in African American women and to test whether inhibitor of IL-6R Tocilizumab with carboplatin will be effective against triple negative breast cancers in African American women compared to women of European ancestry. Genentech has been approached for drugs. The proposed clinical trial endpoints include: The primary efficacy endpoint be overall response rate based on RECIST 1.1 criteria. The secondary clinical endpoints will include progression free survival and toxicity in treated patients.

The tertiary endpoints will include change in circulating IL-6, sIL-6R, TNF α , and C-reactive protein levels over time.

Exploratory endpoints will include change in T-cell subsets (CD4+, CD8+, T-regs) and myeloid precursors (MDSC, DC, macrophages) overtime. Changes in tumor ZEB1, IL-6, IL-6R, and pSTAT3 will be measured. We will also explore changes in immune-related markers based on multiplexed digital spatial profiling⁹. ctDNA will be quantified serially. However, launching of this clinical trial will depend on willingness of Genentech to provide drug and raising sufficient funds.

We will also investigate whether genotyping of AA women for duffy phenotype is necessary to identify tumors that respond or do not respond to specific therapy.

5 Changes/problems: We had some difficulty in maintaining immortalized cell lines from duffy carriers. However, transformed variants are available to continue our studies. The laboratory is still working to identify oncogenes other than H-RAS^{G12V} + SV40-T/t antigens to transform breast epithelial cells. If we are successful, few of the experiments noted above will be repeated with different oncogenes. Phospho-STAT3 was identified as a potential biomarker of breast tumors in AA women and phospho-p65 and IGFBP2 as markers in AA women who are also duffy carriers (need to be reproduced with more samples and under different conditions). Therefore, all future preclinical studies of this project will characterize phospho-STAT3, phospho-p65. If our preliminary observation of duffy-heterozygous and duffy-null epithelial cells expressing higher level of IGFBP2 is reproduced using independent methods, few of mechanistic studies will focus on this molecule. Similarly, if AA epithelial specific expression of GDF15 and VEGF are reproduced, additional studies will focus on these molecules.

6 Products: The following manuscript is under review at Nature Communication and the preprint has been made available publicly through bioRxiv (<https://doi.org/10.1101/2022.07.14.500115>). The abstract has also been submitted to San Antonio Breast Cancer Symposium 2022.

Brijesh Kumar, Katie Batic, Poornima Bhat-Nakshatri, Maggie M Granatir, Rebekah Joann Addison, Megan Szymanski, Lee Ann Baldrige, Constance J. Temm, George Sandusky, Sandra K Althouse, Anna Maria Storniolo, and Harikrishna Nakshatri

Influence of genetic ancestry on breast stromal cells provides biologic basis for increased incidence of metaplastic breast cancer in women of African descent (manuscript under review).

We have created a TMA that contains breast tumors from 28 AA women with 8 among them being duffy carriers. This TMA has been stained for CD8 (for T cells) and CD68 (for macrophages). Although we did not observe differences between duffy-wild type and duffy-carrier groups, this TMA is continued to be analyzed for other markers including ZEB1 and phospho-STAT3. The TMA has also been shared with Dr. Peggi Angel of Medical University of South Carolina for post-translational modification of extracellular matrix proteins. As a part of our collaboration, we recently demonstrated distinct differences in N-glycosylation of matrix proteins in

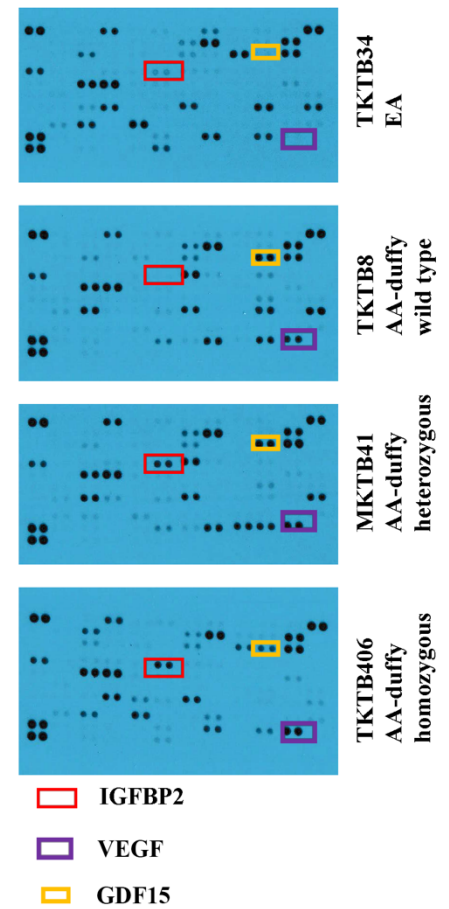


Figure 8: Cytokine/chemokine array identify IGFBP2 as a cytokine released by duffy-heterozygous/homozygous cells. VEGF and GDF15 are released by three AA-derived but not one EA-derived cell line.

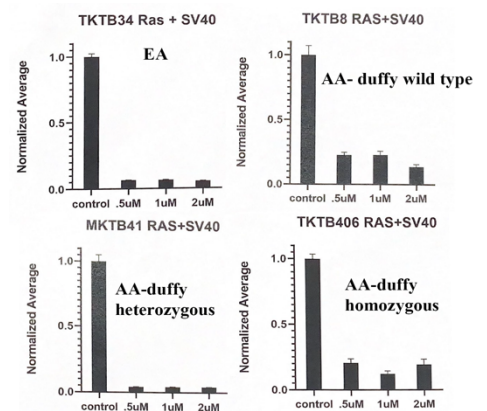


Figure 9: Transformed duffy-wild type, duffy-heterozygous and duffy-homozygous cells show similar sensitivity to paclitaxel at doses tested. Lower of concentration of drugs are currently being tested.

breast tissues of AA women compared to EA women¹⁰. Her group will be investigating whether there are further differences if tissues are from duffy carriers compared to tumor tissues from duffy-wild type AA women.

7 Participants & Other Collaborating Organizations:

Participants:

Brijesh Kumar
Katie Batic
Poornima Bhat-Nakshatri
Aditi Khatpe
Anna Maria Storniolo
George Sandusky
Sandra K Althouse
Komen Tissue Bank

8 Special reporting requirements: None

9 Appendices: Manuscript under review

References

1. Joshi PA, *et al.* PDGFRalpha(+) stromal adipocyte progenitors transition into epithelial cells during lobulo-alveologenesis in the murine mammary gland. *Nature communications* **10**, 1760 (2019).
2. Farup J, *et al.* Human skeletal muscle CD90(+) fibro-adipogenic progenitors are associated with muscle degeneration in type 2 diabetic patients. *Cell metabolism* **33**, 2201-2214 e2211 (2021).
3. Lin WH, *et al.* STAT3 phosphorylation at Ser727 and Tyr705 differentially regulates the EMT-MET switch and cancer metastasis. *Oncogene* **40**, 791-805 (2021).
4. Hughes CE, Nibbs RJB. A guide to chemokines and their receptors. *FEBS J* **285**, 2944-2971 (2018).
5. Aran D, *et al.* Widespread parainflammation in human cancer. *Genome Biol* **17**, 145 (2016).
6. Suriben R, *et al.* Antibody-mediated inhibition of GDF15-GFRAL activity reverses cancer cachexia in mice. *Nat Med* **26**, 1264-1270 (2020).
7. Apte RS, Chen DS, Ferrara N. VEGF in Signaling and Disease: Beyond Discovery and Development. *Cell* **176**, 1248-1264 (2019).
8. Gao S, *et al.* IGFBP2 Activates the NF-kappaB Pathway to Drive Epithelial-Mesenchymal Transition and Invasive Character in Pancreatic Ductal Adenocarcinoma. *Cancer Res* **76**, 6543-6554 (2016).
9. Omilian AR, *et al.* Multiplexed digital spatial profiling of invasive breast tumors from Black and White women. *Molecular oncology* **16**, 54-68 (2022).
10. Rujchanarong D, *et al.* Metabolic Links to Socioeconomic Stresses Uniquely Affecting Ancestry in Normal Breast Tissue at Risk for Breast Cancer. *Frontiers in oncology* **12**, 876651 (2022).

Influence of genetic ancestry on breast stromal cells provides biologic basis for increased incidence of metaplastic breast cancer in women of African descent

Brijesh Kumar¹, Katie Batic¹, Poornima Bhat-Nakshatri¹, Maggie M Granatir², Rebekah Joann Addison², Megan Szymanski², Lee Ann Baldrige², Constance J. Temm², George Sandusky², Sandra K Althouse³, Anna Maria Storniolo⁴, and Harikrishna Nakshatri^{1,5,6*}

¹Department of Surgery, Indiana University School of Medicine, Indianapolis, IN 46202, USA

²Department of Pathology and Laboratory Medicine, Indiana University School of Medicine, Indianapolis, IN 46202, USA

³Department of Biostatistics and Health Data Science, Indiana University School of Medicine, Indianapolis, IN 46202, USA

⁴Department of Medicine, Indiana University School of Medicine, Indianapolis, IN 46202, USA

⁵Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, IN 46202, USA.

⁶VA Roudebush Medical Center, Indianapolis, IN 46202, USA

*Corresponding author: Harikrishna Nakshatri, B.V.Sc., PhD.

C218C, 980 West Walnut Street

Indianapolis, IN 46202, USA

317 274 0396 (fax)

email: hnakshat@iupui.edu

Running title: Influence of genetic ancestry on breast stromal cells.

Keywords: breast cancer, genetic ancestry, normal breast, PROCR, ZEB1, PDGFR α ,

breast microenvironment

ABSTRACT

The biologic basis of genetic ancestry-dependent variability in disease incidence and outcome is just beginning to be explored. We recently reported enrichment of a population of ZEB1-expressing cells located adjacent to the ductal epithelial cells in the normal breast of women of African Ancestry (AA) compared to European Ancestry (EA). By establishing and characterizing cell lines corresponding to these cells and validating *in vitro* findings with tissue microarrays of healthy breast tissue from AA, EA and Latina Ancestry (LA) women, we demonstrate that these cells have the properties of fibroadipogenic/mesenchymal stromal cells that express PROCN and PDGFR α . PROCN+/ZEB1+/PDGFR α + cells, hence renamed as PZP cells, are enriched in the normal breast tissues of AA compared to EA or LA women. *In vitro*, PZP cells trans-differentiated into adipocytes or osteocytes. In co-culture conditions, PZP:epithelial cell communication resulted in luminal epithelial cells acquiring basal/stem cell characteristics and increased expression of IL-6 suggesting the impact of this communication on breast epithelial hierarchy and the microenvironment. Consistent with this possibility, the level of phospho-STAT3, which is a downstream target of IL-6, was higher in the normal and cancerous breast tissues of AA compared to EA women. PZP cells transformed with HRas^{G12V} \pm SV40-T/t antigens generated metaplastic carcinoma in NSG mice suggesting that these cells could be the cell-of-origin of metaplastic breast cancers. Collectively, these results identify a stromal cell component that could influence the biology of breast cancer in AA women.

INTRODUCTION

Mortality from triple negative breast cancer (TNBC) is much higher in women of African Ancestry (AA) compared to women of European Ancestry (EA) even when controlled for socioeconomic status ¹. Even in cases of estrogen receptor positive (ER+)/HER2-/node negative breast cancers with intermediate Oncotype recurrence score ², clinical outcome is worse in AA women compared to EA women ³. Furthermore, ductal carcinoma in situ (DCIS), which is typically considered to be of minimal risk of recurrence, tends to be more deleterious to AA women compared to EA women ⁴. With the recent discovery of the presence of an extra 296,485,284 base pairs DNA in populations of African descent affecting 315 distinct protein-coding genes ⁵, the disparity in breast cancer outcome could, therefore, partly be due to genetic ancestry-dependent variability in normal and cancerous breast biology.

To discover genetic ancestry-dependent differences in normal breast biology, we initiated studies using healthy breast tissues of women of different genetic ancestry donated to the Susan G. Komen Tissue Bank (KTB) at the IU Simon Comprehensive Cancer Center (IUSCCC). We had previously demonstrated that cultured cells from breast tissues of AA women are enriched for PROCR⁺/EpCAM⁻ and CD44⁺/C24⁻ cells compared to EA women ⁶. In the mouse mammary gland, the PROCR⁺/EpCAM⁻ cells have been demonstrated to function as multi-potent stem cells ⁷. Subsequently, we demonstrated that the normal breasts of AA women contain elevated number of ZEB1⁺ cells that surround the ductal epithelial cells compared to EA women ⁸. Several different functions have been attributed to ZEB1 in breast cancer. For example, ZEB1-mediated epithelial to mesenchymal transition (EMT) confers a metastasis-initiating cell state to cancer cells ⁹. Stromal ZEB1

expression in breast cancer inversely correlates with abundance of multiple immune cell types¹⁰. ZEB1 expression has been linked to claudin-low and metaplastic breast cancer subtypes, both of which are more common in AA compared to EA women^{11, 12, 13, 14}. These observations prompted us to further characterize stromally located ZEB1⁺ cells for their potential role in breast cancer initiation and progression within the context of genetic ancestry.

In this study, we established several ZEB1⁺ cell lines, all derived from the breast tissues of AA women. These cells express PROCR and PDGFR α , similar to recently described stromal adipogenic progenitors cells of the mouse mammary gland that trans-differentiate into multiple cell types¹⁵. Based on cell surface expression of PROCR and PDGFR α combined with ZEB1 expression, we have named these cells as ‘PZP cells’ for simplicity. Under appropriate growth conditions, these cells differentiated into adipogenic or osteogenic lineages and expressed the adipogenic marker Peroxisome Proliferator Activated Receptor gamma (PPAR γ). Cell surface marker profiles revealed similarity of these cells to human prostate-derived mesenchymal stem cells (CD44⁺/CD90⁺/CD73⁺/CD105⁺ populations)¹⁶. PZP:breast epithelial cell communication resulted in elevated expression of IL-6. Furthermore, epithelial cells acquired CD49⁺/EpCAM⁻ basal cell characteristics, which has been shown to increase with aging and to increase susceptibility to breast cancer^{17, 18}. PZP cells, when transformed with oncogenes, generated metaplastic carcinomas in NSG mice suggesting that these cells are cell-of-origin of metaplastic carcinomas of the breast¹⁹.

MATERIALS AND METHODS

Primary cell culture, Immortalization and Lentiviral transduction

Primary breast epithelial cell lines were created from fresh or cryopreserved, de-identified normal breast tissues donated to KTB by healthy AA women; all subjects provided written informed consent prior to donation. All experiments were carried out in accordance with the approved guidelines of the Indiana University Institutional Review Board. International Ethical Guidelines for Biomedical Research involving human subjects were followed. Generation of primary cells from cryopreserved tissues, culturing method and media composition have been described previously²⁰. While primary breast tissues cultured from EA donors generated predominantly tightly attached cuboidal epithelial cell colonies, tissues from AA donors yielded both loosely attached elongated cell clusters and cuboidal epithelial colonies. Differential trypsinization was used to separate two population of cells for immortalization. Cells were immortalized by human telomerase gene (hTERT) using the retrovirus vector pLXSN-hTERT or the lentivirus vector pCLX-PGK-hTERT vector (#114315, Addgene). Retrovirus and lentivirus preparations, infection of primary cells, and selection for immortalized cells by 100 µg/ml G418 (only with pLXSN-hTERT, 61-234, Corning) have been described previously²¹. Cells were transformed with oncogenes *H-Ras*^{G12V}, and SV40-T/t antigens expressing lentiviruses using vectors pLenti CMV-RasV12-Neo (w108-1) (HRAS^{G12V}, #22259, Addgene), and pLenti-CMV/TO-SV40 small + Large T (w612-1) (#22298, Addgene), respectively, as before²². PZP KTB40 and KTB42 cells were modified to express Tomato Red or GFP markers using pCDH-EF1-Luc2-P2A-tdTomato (#72486, Addgene) or pCDH-EF1-Luc2-P2A-Cop GFP (#72485,

Addgene) lentivirus. Cell lines in the laboratory are routinely tested for *Mycoplasma* ~every 6 months (last testing was done on February 24, 2022).

Genetic ancestry mapping and genotype analysis

Leukocyte DNA from the same donors was obtained from KTB. Blood was collected using the BD-Vacutainer spray-coated K3EDTA tubes (Becton Dickinson). Tubes were centrifuged for 15 min at 2000 RPM. Once the upper phase (plasma) was removed, the tubes were stored in -80°C. Peripheral blood leukocyte DNA was extracted using AutoGenprep 965 (Autogen, Inc.). Genotyping was performed using the KASP technology (LGC Genomics) and a 41-SNP panel (labeled 41-AIM panel) selected from Nievergelt *et al*²³. Genotype analysis using the 41 SNPs panel along with a Bayesian clustering method (Structure Software V2.3.4) was able to discern continental origins including European (Caucasian)/Middle East, East Asia, Central/South Asia, Africa, Americas, and Oceania. A reference set was obtained from the Human Genome Diversity Project (HGDP).

Flow cytometry analysis and cell sorting

Adherent cells were collected by trypsinization, stained using antibodies PROCR (CD201)-PE (130-105-256), EpCAM-APC (130-091-254) EpCAM-PE (130-091-253), (Miltenyi Biotech Inc.), PDGFR α -PE (#323506, Biolegend), CD26-APC (#563670, BD Biosciences), CD105-PE (#12-1057-42, Invitrogen), CD73-PE (561014), CD90-APC (559869), CD44-APC (559942), CD24-PE (555428) (BD Pharmingen), CD10-APC (#340922, BD Biosciences), CD49f-APC (FAB13501A, R&D Systems), and were acquired using a BD LSR II flow cytometer. Data were analyzed using FlowJo software. Forward and side scatter were used to ensure that only live cells were considered in the

analysis. Gating was done using appropriate FITC (555573), PE (555749) and APC (555576) (BD Pharmingen) isotype control antibodies and only a representative isotype control for two fluorescent markers are shown. Samples were analyzed and sorted using BD FACSAria and SORPAria.

RNA extraction and quantitative real-time Reverse Transcription Polymerase Chain Reaction (qRT-PCR)

Total RNA was isolated using RNeasy Kit (74106, Qiagen) and 2 µg of RNA was used to synthesize cDNA with Bio-RAD iScript cDNA Synthesis Kit (170-8891, Bio-Rad). qRT-PCR was performed using Taqman universal PCR mix and predesigned gene expression assays with best coverage from Applied Biosystems. The following primer sets were used in our study: ACTB (Hs01060665_g1), ZEB1 (Hs01566408_m1), PPARG (Hs01115513_m1), RUNX1 (Hs01021971_m1), IL-6 (Hs00174131_m1), TAGLN (Hs06633192_s1), WISP1 (Hs05047584_s1), TNC (Hs01115665_m1), CSF1 (Hs00174164_m1), SPP1 (Hs00959010_m1), IL-33 (Hs04931857_m1), DLL1 (Hs00194509_m1), F3 (Hs01076029_m1), NRG1 (Hs01101538_m1), IL-8 (Hs00174103_m1), CMTM6 (Hs00215083_m1), MIF (Hs00236988_g1), MFGE8 (Hs00983890_m1), POSTN (Hs01566750_m1).

Generation of TMA

We created a TMA comprising healthy breast tissue donated to the KTB (KTB-normal) from 49 women of African (AA), 154 of European (EA) and 46 of Hispanic/Latinas (LA) ancestry. The TMA also contained matched normal adjacent to tumor (NAT) and tumor tissue of approximately 50 donors. Part of the EA TMA has been

described previously ²⁴. Genetic ancestry information of KTB-normal donors and histopathology of tumors are shown in supplementary Figure 1 (Fig. S1) and Table S1, respectively. TMA was analyzed by immunohistochemistry (IHC) for PROCR (MAB22451, R&D Systems), ZEB1 (3G6, #14-9741-82, eBioscience), PDGFR α (#3174, Cell Signaling Technology), phospho-STAT3 (#ab30647, Abcam), Estrogen Receptor alpha (ER α , Clone EP1, Dako IR 084), GATA3 (sc-268, Santa Cruz Biotechnology), and FOXA1 (sc-6553, Santa Cruz Biotechnology) expression. IHC was done in a Clinical Laboratory Improvement Amendments (CLIA)-certified Indiana University Histopathology Laboratory and evaluated by three pathologists in a blinded manner. Quantitative measurements were done using the automated Aperio imaging system (ScanScope CS) and analysis was conducted using an FDA-approved algorithm. Positivity and H-scores were recorded and statistically analyzed as described previously ⁸. Data were analyzed in three different ways: (i) expression differences between AA, EA and LA KTB-normal; (ii) expression differences between KTB-normal and NATs; and (iii) expression differences between NATs and tumors.

Adipogenic and osteogenic differentiation assay

For adipogenic differentiation assay, PZP cells were plated in tissue culture dishes pre-coated with laminin-5-rich conditioned media from the 804G cell line in DMEM supplemented with 10% FBS (#26140-079) and 10 ng/ml bFGF (#3718-FB-010), and maintained at 37°C in a 5% CO₂ incubator. Cells were allowed to grow to confluence and were then held at confluence for 2 days without changing of the media prior to exposure to the adipocyte differentiation cocktail (1 μ g/ml insulin, 0.25 μ g/ml dexamethasone, and 0.5 mM IBMX) in fresh media without addition of bFGF. The adipocytes differentiation media

was changed every 3 days until day 12 for harvest. Cells were fixed and stained with oil red O (#ECM950, Millipore) according to the manufacturer's instructions.

For osteogenic differentiation assay, PZP cells were plated onto tissue culture dishes in 50% DMEM/F-12 (#A4192001) supplemented with 10% FBS, 1% L-glutamine (#SCR028, Millipore Sigma). Cells were allowed to grow for 2-3 days to attain confluence, after which the normal media was removed and osteogenic differentiation medium was added. This media change corresponded to differentiation day 1. The osteogenic differentiation media contained 50% DMEM/F-12 supplemented with 10% FBS, 1% L-glutamine, 5 nM dexamethasone (#SCR028, Millipore Sigma), 250 μ M L-ascorbic acid 2-phosphate (#SCR028, Millipore Sigma) and 10 mM β - glycerophosphate (#SCR028, Millipore Sigma). The media was changed twice a week. After 3 weeks, differentiated cells were washed in PBS and fixed in 95% methanol for 10 mins. Cells were stained with 2% alizarin red solution (Catalog no.# ECM815, Millipore) for 5 mins, then rinsed with water and imaged under fluorescence microscope to visualize the Ca²⁺ deposits.

Co-culture and Cytokine/chemokine array

50% of each epithelial cell line (KTB34 and KTB39) and PZP cells (KTB40 and KTB42) were co-cultured in growth media for 2 days for RNA analysis. For cytokine/chemokine arrays, culturing of individual and co-culture was done as above and allowed to grow till the plates were confluent. Plates were washed with PBS and incubated with DMEM/F12 without other ingredients for 24 hours. Conditioned media (CM) was collected and subjected to cytokine/chemokine array. The cytokine/chemokine protein array was performed according to the manufacturer's instructions (#ARY022B, R&D Systems).

Antibodies and Western blotting

Primary antibodies used included rabbit anti-Ras (#3965, Cell Signaling Technologies), mouse anti-SV40-T antigens (MABF121, EMD Millipore), and mouse anti- β -actin (A5441, Sigma-Aldrich). Anti-mouse (#7076), and anti-rabbit (#7074) horseradish peroxidase (HRP) linked secondary antibodies were purchased from Cell Signaling Technologies. Cell lysates were prepared in radioimmunoassay buffer and analyzed by western blotting as described previously ²².

Xenograft study and IHC of tumors and lungs

The Indiana University Animal Care and Use Committee approved the use of animals in this study and all procedures were performed as per NIH guidelines. Transformed cells with 50% basement membrane matrix (BME) type 3 (3632-005-02, Trevigen) (total 100 μ L volume) were implanted into the mammary fat pad of 5 to 6 week old female NSG (NOD/SCID/IL2R^{gnull}) mice. A 60 day slow release 17 β -estradiol (0.72 mg) pellet (SE-121, Innovative Research of America) was implanted at the time of mammary fat pad injection. Tumor growth was measured weekly and tumor volume was calculated as described previously ²². After 2-4 months, tumors and lungs were collected and processed for hematoxylin and eosin (H&E), ER α , GATA3, FOXA1, CK5/6 (IR 780, Dako), CK8 (35bH11, N1560, Dako), CK14 (LL002, ab7800, Abcam), and CK19 (IR 615, Dako) staining as described previously ²².

Statistical analysis

Statistical analyses were conducted using Prism software program (version 6.0) and SAS software (version 9.4). Data were analyzed using one-way ANOVA. P values <0.05 were considered statistically significant. Nonparametric Wilcoxon rank-sum tests were used for unpaired analysis, as positivity and H-scores were not normally distributed, whereas nonparametric Wilcoxon signed-rank tests were used for paired analyses as described previously ⁸. For comparisons between three groups, the Kruskal-Wallis test was used.

RESULTS

Generation of **PROCR⁺/ZEB1⁺/PDGFR α ⁺** (PZP) cell lines from breast tissues of healthy AA women

We created six immortalized PZP cell lines from AA (KTB40, KTB42, KTB32, KTB53, KTB57, and KTB59) by overexpressing human telomerase gene (hTERT) using primary cells isolated and propagated from core breast biopsies of self-reported healthy black women. Enrichment of African ancestry informative markers in these donors was confirmed through genotyping (**Fig. 1A**). To phenotypically characterize and document heterogeneity, these cells were subjected to flow cytometry using various epithelial, stem, mesenchymal, and fibroblast markers. These cell lines were predominantly PROCR⁺/EpCAM⁻ (**Fig. 1B, C, and Fig. S2A**). To further characterize stem cell -related gene expression, we compared these immortalized variants with the immortalized PROCR[±]/EpCAM⁺ luminal/basal cell variants from EA and AA women we described previously ²¹. PROCR⁺/EpCAM⁻ cell lines expressed significantly higher levels of ZEB1 compared to PROCR[±]/EpCAM⁺ (KTB34 and KTB39) breast epithelial cell lines described previously ²¹ (**Fig. 1D**). Gene expression pattern in KTB34 cell line from EA overlaps Luminal-A intrinsic subtype of breast cancer, whereas gene expression in KTB39 from AA overlaps with Normal-like intrinsic subtype of breast cancer ²¹. Morphologically, PROCR⁺/EpCAM⁻ cell lines showed features of epithelial to mesenchymal transition (EMT), which is expected based on ZEB1 expression (**Fig. S2B**).

A recent study described PDGFR α ⁺ stromal cells as adipogenic progenitors of the mammary gland that transdifferentiate into epithelial cells and migrate into the duct when stimulated by PDGF-C ¹⁵. Interestingly, these cells also expressed PROCR ¹⁵. We

examined whether human breast PROCN⁺/ZEB1⁺ cells express PDGFR α . Indeed, >70% of cells were PDGFR α ⁺ (**Fig. 1E and F; Fig. S2C**), and these cells also expressed PDGFR β (**Fig. S2D**).

Trans-differentiating properties of PZP cells

In order to determine to what extent PZP cells show similarity to adipogenic progenitors described in the mouse mammary gland and fibroadipogenic progenitors (FAPs) described in the skeletal muscle²⁵ with respect to trans-differentiation into adipogenic and osteogenic lineages, PZP cells were subjected to adipogenic and osteogenic differentiation growth conditions. Indeed, PZP cells differentiated into adipocytes (**Fig. 2A**) and the differentiated cells expressed adipocyte differentiation marker PPAR γ (**Fig. 2B**). Adipogenic trans-differentiation capability of PZP cells was further confirmed through RUNX1 overexpression (**Fig. 2C**). PZP cells cultured in osteogenic media for 21 days showed osteogenic differentiation as evident from mineralization of matrix with Ca²⁺ and positive alizarin red staining (**Fig. 2D**). Thus, PZP cells enriched in the breasts of AA women could correspond to multipotent cells that can transdifferentiate into different cell types based on environmental cues.

PZP cells express lobular fibroblast markers

In breasts, loose connective tissue is unique to the terminal duct lobular units (TDLUs), which drain into the interlobular ducts, which in turn are embedded in a more dense connective tissue²⁶. CD105^{high} TDLU-resident lobular fibroblasts display properties different from interlobular fibroblasts²⁷. While the CD105^{high} lobular fibroblasts resemble mesenchymal stem cells (MSCs) both by phenotype and function, CD26^{high} interlobular cells remain fibroblast restricted²⁷. CD105^{high}/CD26^{low} and CD105^{low}/CD26^{high} lineages

are considered to represent lobular and interlobular human breast fibroblastic cells (HBFCs), respectively ²⁸. To further characterize the PZP cell lines, we examined the CD105 and CD26 staining pattern. PZP cells are enriched for CD105^{high}/CD26^{low} population with inter-individual variability in the ratio between CD105^{high}/CD26⁻ and CD105^{high}/CD26^{low} (**Fig. 3A - C**), which suggest that PZP cells reside predominantly in the lobules of breasts.

CD90⁻/CD73⁺ and CD73⁺/CD90⁺ are described as rare endogenous pluripotent somatic stem cells and potential mesenchymal stem cells, respectively ²⁹. PZP cells contained CD90⁻/CD73⁺ and CD73⁺/CD90⁺ subpopulations, with remarkable inter-individual variability in the ratio between CD90⁻/CD73⁺, CD90^{low}/CD73⁺ and CD90^{high}/CD73⁺ (**Fig. 3D – G**). Isotype controls to generate gates are shown in **Fig. S2E**. CD44 and CD24 are the “original” markers used to characterize cancer stem cells (CSCs) in breast cancer ³⁰. Interestingly, PZP cells displayed CD44⁺/CD24⁻ phenotype (**Fig. 3H and I**). However, lack of EpCAM expression in PZP cells suggest that not all CD44⁺/CD24⁻ cells of the breast including breast tumors have stem cell properties. CD10 marker is used to isolate myoepithelial cells, although a recent study showed CD10 positivity of cancer associated fibroblasts ^{31, 32}. PZP cells showed CD10⁺ phenotype (**Fig. 3J - K**). CD49f^{high}/EpCAM^{low}, CD49f^{high}/EpCAM^{medium}, and CD49f^{low}/EpCAM^{high} cells are described as breast stem, luminal progenitor, and mature/differentiated cells, respectively ³³. None of the PZP cell lines were positive for CD49f and EpCAM (**Fig. S2F**).

PZP cells are enriched in the normal breasts of women of AA compared to women of EA and Latina Ancestry (LA)

In our previous study, we had demonstrated that ZEB1⁺ cells in the normal breast are located in the stroma with close proximity to ductal epithelial cells ⁸. We generated a tissue microarray (TMA) comprising healthy breast tissues from 49 AA, 154 EA and 46 LA women. Ancestry marker distribution pattern of donors is shown in **Fig. S1**. TMA was analyzed for protein levels of PROCR, ZEB1 and PDGFR α by IHC. Because of our previous observation of elevated ZEB1⁺ cells in the normal tissues adjacent to tumors (NATs) compared to breast tissues from healthy donors in case of EA ⁸, TMA also contained paired breast tumors and NATs of AA and EA women. To distinguish tissues of clinically healthy donors from NATs, clinically healthy tissues are labelled as KTB-normal. Pathologic features of tumors are described in **Table S1**. Representative IHC staining patterns of PROCR in KTB-normal, NATs, and tumors are shown in **Fig. 4A** and **B** and statistical analyses are presented in **Fig. 4C-E** and in **Table 1**. PROCR expression was observed in both ductal and stromal cells (**Fig. 4A** and **B**). PROCR-expressing cells were enriched in the KTB-normal breast tissue of AA women compared with EA and LA women (**Fig. 4B, Table 1**). NATs of AA and EA women contained significantly higher levels of PROCR⁺ cells compared to KTB-normal suggesting the field effect of tumors on PROCR expression and NATs are not “normal” (**Fig. 4D**). Tumors of EA women displayed a modest increase in PROCR⁺ cells compared with those of NATs, while AA women did not show differences between NATs and tumors (**Fig. 4E**). Thus, PROCR⁺ cells are intrinsically higher in the normal breasts of AA women.

Consistent with our previous report regarding ZEB1 ⁸, KTB-normal breast tissues of AA women displayed higher levels of ZEB1⁺ cells compared to EA women. Interestingly, KTB-normal breast tissues of LA also demonstrated higher levels of ZEB1⁺

cells compared to EA women (**Fig. 5A-C; Table 1**). Note remarkable inter-individual heterogeneity in ZEB1⁺ cells among LA women, possibly reflecting heterogeneity in genetic ancestry of LA. NATs of AA and EA women contained significantly higher levels of ZEB1⁺ cells compared to corresponding KTB-normal (**Fig. 5D**). Tumors of EA women displayed higher levels of ZEB1⁺ cells compared to NATs (**Fig. 5E**). Therefore, the normal breasts of AA women exhibit intrinsically higher levels of ZEB1⁺ cells but ZEB1⁺ cell numbers increase in the breast with cancer in EA women, consistent with our previous report ⁸.

PDGFR α expression in the normal breast tissues was low and mostly in stromal cells (**Fig. 6A**). However, tumor epithelium showed some degree of positivity. PDGFR α -expressing cells were enriched specifically in the KTB-normal breast tissues of AA women compared with EA and LA women (**Fig. 6A-C; Table 1**). Interestingly, KTB-normal breast tissues contained significantly higher levels of PDGFR α ⁺ cells compared to NATs only in AA women, while no difference between KTB-normal breasts and NATs of EA women was observed (**Fig. 6D**). Furthermore, we did not observe a significant difference between NATs and tumors of AA women (**Fig. 6E**). Therefore, PDGFR α ⁺ cells are intrinsically higher in the normal breasts of AA women. Taken together, our data suggest that PZP cells are elevated in the normal breasts of AA women compared to EA women, although expression of PROCR and PDGFR α in epithelial cells makes data interpretation bit difficult.

We next examined the same TMA for luminal epithelial markers to rule out possible bias in our observations. Because FOXA1 along with another pioneer factor GATA3 and ER α form a lineage-restricted hormone-responsive signaling network in the normal breast

³⁴, we examined the expression levels of these markers. Representative staining patterns of ER α , GATA3 and FOXA1 are shown in **Fig. S3**, **Fig. S4** and **Fig. S5**, respectively. We did not observe genetic ancestry-dependent differences in ER α levels between the KTB-normal of AA, EA and LA women (**Fig. S3A-C**). NATs of AA women contained modestly higher level of ER α ⁺ cells compared with those of KTB-normal, while EA women did not show the difference (**Fig. S3D**). Tumors of EA women contained significantly higher level of ER α ⁺ cells compared with those of NATs, whereas AA women did not show the difference (**Fig. S3E**), which is consistent with pathologic features of tumors in EA as majority of them were ER⁺ tumors (**Table S1**). Similar to ER α , GATA3 expression in the normal breasts did not show genetic ancestry-dependent differences (**Fig. S4A-C**). NATs of AA women contained significantly higher level of GATA3⁺ cells compared with those of KTB-normal, while EA women did not show the difference (**Fig. S4D**). Tumors of EA women contained significantly higher level of GATA3⁺ cells compared with those of NATs, which is consistent with clinical features of tumors, whereas AA women did not show the difference (**Fig. S4E**). FOXA1 also did not show any genetic ancestry-dependent differences between KTB-normal of AA, EA, and LA women (**Fig. S5A-C**). NATs of AA and EA women contained significantly higher level of FOXA1⁺ cells compared with those of KTB-normal (**Fig. S5D**). As expected, tumors of EA women contained significantly higher level of FOXA1⁺ cells compared with those of NATs, whereas AA women did not show the difference (**Fig. S5E**). Importantly, these results suggest that the stromal but not the epithelial compartment of the normal breast shows genetic ancestry dependent variability in cell composition, at least with markers examined.

Modeling the effects of PZP cells on breast tumorigenesis

To obtain potential insight into signaling pathway alterations in epithelial and PZP cells as a consequence of their cross-talk, we performed cytokine/chemokine profiling of factors secreted by an immortalized luminal epithelial cell line, a PZP cell line and both co-cultured together (50% of each cell line) for ~24 hours. If the expression under the co-culture condition was much higher than expression in either cell types alone, we interpreted those results as cooperative effect on gene expression. If the expression under co-culture condition was the same or lower than expression in either cell type alone, we interpreted those results as no effect of cell-cell interaction. We selected the Luminal-A epithelial cell line KTB34 for this experiment because this cell line upon transformation generates adenocarcinoma, similar to human breast cancer ^{21, 22}. While luminal epithelial cell line expressed several ligands such as PDGF-AA and osteopontin, which can affect trans-differentiation of PZP cells, PZP cells expressed factors such as EGF, HGF and SDF-1 α , which can signal in luminal cells (**Fig. 7A; Table 2**). Interestingly, IL-6 is produced only under coculture condition (**Fig. 7A; Table 2**). We further confirmed the IL-6 production under co-culture condition at mRNA level by qRT-PCR (**Fig. 7B**). We suspect PZP cells produce IL-6 in response to interaction with luminal cells as basal expression of IL-6 was much higher in PZP cells compared to luminal cell line (**Fig. S6A**) and luminal cells secreted IL-1 α , which we have previously shown to induce IL-6 in stromal cells ³⁵. There appears to be specificity in cytokine production under co-culture conditions as we did not observe elevated production of IL-8 under co-culture condition of PZP and epithelial cells (**Fig. S6B**).

PZP-epithelial cells interaction alters the expression of specific genes

We performed an extensive literature search to identify potential genes whose expression could be altered due to stromal-epithelial cell interaction. For example, transgelin (TAGLN) is known to be a specific marker of smooth muscle differentiation and highly expressed in the myoepithelial cells and fibroblastic cells of benign breast tissue with limited expression in luminal cells ^{36,37}. A population of subepithelial cells that lines the entire villus-crypt axis of intestine express high levels of PDGFR α , DLL1, F3, and EGF-family ligand Neuregulin 1 (NRG1) ³⁸. In addition, NRG1 is also expressed in mesenchymal cells adjacent to the proliferative crypts ³⁸. Expression of these genes were examined in individual cell types and under co-culture. Three PZP cell lines (KTB32, KTB40, KTB42) and two epithelial cell lines (KTB34, KTB39), and co-culture of PZP and epithelial cell lines (50% of each cell line) were used. We observed abundant TAGLN expression in PZP cells, while epithelial cells expressed at low level. Interestingly, the expression of TAGLN was further increased under co-culture condition (**Fig. 7C**). We found a low level of DLL1, F3, and NRG1 expression in PZP cell lines except high level of NRG1 in KTB42. DLL1, F3 and NRG1 are expressed predominantly in epithelial cell lines. In co-cultured cells, expression of DLL1 and F3 was additive depending on the cell type (**Fig. S6C - E**). Taken together, these results indicate that PZP cells correspond to stromal cells that interact with epithelial cells to alter gene expression in a reciprocal manner. However, these cells are unlikely to function similar to subepithelial mesenchymal cells described in the intestine ³⁸.

Potential role of PZP cells in immune cell modulation in the microenvironment

Apart from IL-6, several other factors have been shown to alter the immune microenvironment. For example, WNT1-inducible-signaling pathway protein 1 (WISP1)

expression affects the clinical prognosis through associations with macrophage M2 polarization, and immune cell infiltration in pan-cancer and helps to maintain CSC properties in glioblastoma ³⁹. PZP cell lines displayed higher expression of WISP1 compared to luminal cell lines, while additive expression was observed under co-culture (**Fig. 7D**). Tenascin-C (TNC) promotes inflammatory response by inducing the expression of multiple proinflammatory factors in innate immune cells such as microglia and macrophages. TNC drives macrophage differentiation and polarization predominantly towards an M1-like phenotype ⁴⁰. TNC is expressed mostly by epithelial cells and this expression was unaffected under co-culture (**Fig. 7E**). CSF1 controls both the differentiation and immune regulatory function macrophages ⁴¹. PZP cell lines displayed higher expression of CSF1 compared to epithelial cell lines and the expression unaffected under co-culture (**Fig. 7F**). Osteopontin (SPP1 or OPN), secreted by myofibroblasts, promotes M2 macrophage polarization through STAT3/PPAR γ pathway ⁴². PZP cell lines displayed high expression of SPP1 compared to epithelial cell lines and the expression remained additive under co-culture condition (**Fig. 7G**). IL-33 is known to be upregulated in metastases-associated fibroblasts, and the upregulation of IL-33 activates type 2 inflammation in the metastatic microenvironment and facilitates eosinophils, neutrophils, and inflammatory monocytes recruitment to lung metastases ⁴³. Co-culturing of PZP and epithelial cells did not affect IL-33 expression (**Fig. 7H**). CMTM6 maintains the expression of PD-L1 in tumor cells to regulate anti-tumor immunity ⁴⁴. Epithelial cells but not PZP cells expressed higher levels of CMTM6, which was unaffected by co-culture condition (**Fig. S6F**). Macrophage migration inhibitory factor (MIF) is an essential cytokine that is involved in the regulation of macrophage function in host defense through the suppression

of anti-inflammatory effects of glucocorticoids ⁴⁵. Both PZP and epithelial cell lines expressed MIF, but co-culture condition did not further effect expression (**Fig. S6G**). Secretion of MFGE8 can reprogram macrophages from an M1 (proinflammatory) to an M2 (anti-inflammatory but pro-tumorigenic) phenotype ⁴⁶. MFGE8 also induces the production of basic fibroblast growth factor that is responsible for fibroblast migration and proliferation ⁴⁷. PZP cell lines displayed high expression of MFGE8 which was unaffected by co-culture (**Fig. S6H**). Periostin (POSTN) is predominantly secreted by stromal fibroblasts to promote the proliferation of tumor cells. POSTN is also an essential factor for macrophage recruitment in the tumor microenvironment and involved in the interactions between macrophages and cancer cells ⁴⁸. PZP cell lines displayed high expression of POSTN and the expression was unaffected by co-culture condition (**Fig. S6I**). We noted variation in the expression of select genes amongst PZP cell lines (WISP1, CSF1, NRG1, and POSTN, for example), suggesting inter-individual variability in gene expression in stromal cells, similar to epithelial cell types we described previously ²¹. Collectively, these results suggest that PZP-epithelial interaction in the breast could impact the levels of select chemokines/cytokines in the breast environment (IL-6 and TAGLN for example) with consequential effects on the normal and/or tumor immune environment.

The effects of breast cancer: PZP cell interaction on trans-differentiation of epithelial luminal progenitor cells

In order to further investigate the intercellular communication between PZP and epithelial cells, we generated stable tdTomato-labeled KTB40 and KTB42 cell lines using pCDH-EF1-Luc2-P2A-tdTomato lentivirus (**Fig. S7A**). To determine whether co-culturing alter the phenotype of epithelial cells, tdTomato-labeled PZP (KTB40, KTB42), epithelial

(KTB34, KTB39), and co-cultured PZP and epithelial cells were analyzed by flow cytometry using CD49f and EpCAM, which can differentiate luminal mature (CD49f⁻/EpCAM⁺), luminal progenitor (CD49f⁺/EpCAM⁺) and basal/stem (CD49f⁺/EpCAM⁻) cells. Co-cultured epithelial cells displayed an increase in CD49f⁺/EpCAM^{-/low} and reduced CD49f^{-/low}/EpCAM^{+/high} cell populations in co-culture condition (**Fig. 8A-C; Fig. S7B**). tdTomato-Red⁺ PZP cells on their own were CD49f⁻/EpCAM⁻), a small fraction of these cells displayed a unique CD49f⁺/EpCAM⁺ phenotype under co-culture condition (**Fig. 8D and E; Fig. S7C**). These results suggest that similar to mouse mammary fibroadipogenic cells¹⁵, PZP cells can potentially acquire epithelial characteristics under specific conditions. Our attempts of characterize flow-sorted CD49f⁺/EpCAM⁺ trans-differentiated PZP cells were not successful because of dominant growth of few contaminating PZP cells in the sorted population of cells.

Transformed PZP cells generate metaplastic carcinoma.

We recently reported that cell-of-origin but not the oncogenic mutations determine the histotypes of breast tumors using breast epithelial cell lines derived from multiple donors and a defined set of oncogenes²². We used the same strategy to determine whether PZP cells are cell-of-origin of specific malignancy of the breast. Mutant Ras is one of the potent oncogenes used to transform breast epithelial cells *in vitro*⁴⁹. Although initially considered not a relevant oncogene in breast cancer and breast cancer-related studies that utilized Ras oncogenes were often dismissed upon by reviewers, recent studies have clearly shown the role of Ras oncogene in endocrine resistance and metastasis of luminal breast cancer⁵⁰. Furthermore, the majority of breast cancer cell lines used as model systems have either mutation in Ras genes themselves or in downstream effectors or negative regulators

of Ras pathway and Ras pathway is the major signaling pathway activated in TNBCs and Claudin-low breast cancer subtype ⁵¹. We transformed PZP cell lines with HRas^{G12V}, SV40-T/t antigen and combination of both HRas^{G12V} and SV40-T/t antigen using lentivirus, since this combination was the most effective in breast epithelial cell transformation ²². Western blotting was used to detect the overexpression of mutant HRas^{G12V} (**Fig. S8A**), SV40-T/t antigen (**Fig. S8B**), and combination of both HRas^{G12V} and SV40-T/t (**Fig. S8C**). Phase contrast images of PZP transformed cell lines are shown in **Fig. S8D**. Transformation of PZP cells with activated HRas^{G12V} increased the fraction of cells that have acquired epithelial phenotype and express EpCAM, particularly in KTB42 (**Fig. 9A and B; Fig. S8D**). Transformed PZP cells expressed the stem/basal cell marker CD49f with inter-individual variability (**Fig. 9A; Fig. S9A**). Transformation also altered the cell surface profiles of mesenchymal stem cell marker CD90 (**Fig. 9C; Fig. S9A**). Transformed PZP cells were CD201⁺ and CD44⁺ (**Fig. 9B; Fig. S9A and S9B**). Thus, PZP cells transdifferentiate upon transformation by acquiring CD49f expression.

We next determined whether cells expressing oncogenes are tumorigenic in NSG mice. Indeed, five million transformed cells in 50% matrigel implanted into the mammary gland of 6-7 week old female NSG (NOD/SCID/IL2Rgnull) mice (n=5) progressed into tumors. All animals injected with HRas^{G12V} + SV40-T/t and four out of five animals injected with HRas^{G12V} transformed cells but not parental immortalized cell lines (n=5) generated tumors. Tumor was resected and analyzed by H&E staining and expression of luminal markers estrogen receptor alpha (ER α), GATA3, and FOXA1 and cytokeratins CK5/6, CK8, CK14 and CK19. The luminal cells express cytokeratin 19 (CK19), while basal cell types express cytokeratin 5/6 (CK5/6) and cytokeratin 14 (CK14), and cells

expressing both CK14 and CK19 show luminal progenitor phenotype or luminal/basal hybrid cells of the breast⁵². KTB42-HRas^{G12V} cell-derived tumor was ER α ⁻/GATA3⁻/FOXA1⁻ (**Fig. 9D**). Surprisingly, KTB42-HRas^{G12V}-derived tumor was also CK5/6⁻/CK8⁻/CK14⁻/CK19⁻ (**Fig. S9C**). KTB42 cell line transformed with both mutant HRas^{G12V} and SV40-T/t antigen also developed tumors in NSG mice. KTB42- HRas^{G12V}+SV40-T/t cell-derived tumor was ER α ⁻/GATA3⁻/FOXA1⁻ (**Fig. 9E**), and CK5/6⁻/CK14⁻/CK19⁻ (**Fig. S9D**). Unlike luminal breast epithelial cell derived tumors obtained after transformation with the same set of oncogenes²², which metastasized to lungs, these tumors did not show extensive lung metastasis (**Fig. 9F and G**). Histologically, these tumors are metaplastic carcinomas, which comprise 0.08-0.2% of all breast neoplasms (**Fig. 9G**)¹⁹.

Normal breast and/or breast tumors of AA women have higher levels of phospho-STAT3.

Since PZP-epithelial cell interaction resulted in elevated IL-6 expression *in vitro* and PZP cells are present at higher levels in the normal breast of AA women, it is expected that signals downstream of IL-6 should be higher in the breast tissues of AA women compared to EA women. Phosphorylation of STAT3 at residue S727 is one the major downstream signal upon binding of IL-6 to its receptor and S727 phosphorylated STAT3 promotes mesenchymal to epithelial transition of early disseminated cancer cells⁵³. We used phospho-STAT3 as a surrogate marker to determine IL-6 activity in the normal breast tissues and breast tumors of AA and EA women. Although cells in the KTB-normal contained lower levels of phospho-STAT3 as we could estimate on positivity, positivity was still higher with healthy breast tissues of AA women compared to EA women (**Fig. 10A-C**). A modestly higher levels of phospho-STAT3 were also observed in tumors of AA

compared to EA women (**Fig. 10C**). Consistent with elevated levels of PZP cells in NATs compared to KTB-normal in EA women, phospho-STAT3 levels were higher in NATs compared to KTB-normal of EA women (**Fig. 10D**). Thus, genetic ancestry dependent variability in levels of stromal PZP cells has an impact on normal and cancerous breast biology.

DISCUSSION

At the most recent American Cancer Society estimate, breast cancer mortality rate in AA women is significantly higher compared to EA women, while breast cancer incidence in AA women is lower⁵⁴. This difference cannot be explained solely by differences in socioeconomic status or access to health care. This suggests that the biology of normal breast epithelial cells between these two groups differs, which may contribute to altered susceptibility to tumor initiation, progression and/or metastasis. This study focused on one particular stromal cell type, which can contribute to genetic ancestry dependent variability in breast biology.

Cell surface marker profiling and transdifferentiation studies indicate similarities between PZP cells and tissue resident mesenchymal stem cells, which have transdifferentiating and regenerative capacity⁵⁵. PZP cells contain three subpopulations, CD90⁻/CD73⁺, CD90^{low}/CD73⁺ and CD90^{high}/CD73⁺. These cells also differentiate into adipocyte and osteogenic lineages. Similar cell type in the mouse mammary gland functions as adipogenic progenitor and differentiates into luminal and basal epithelia during mammary morphogenesis¹⁵. We observed epithelial differentiation of PZP cells only upon transformation by HRas^{G12V}. In this respect, PZP cells are the breast counterpart of PDGFR α ⁺ intestinal stromal progenitors or fibroadipogenic progenitors (FAPs) of the skeletal muscle. FAPs cells are multipotent, and are capable of adipogenesis, fibrogenesis, osteogenesis and chondrogenesis^{38, 55}. While intestine and skeletal muscle undergo extensive regeneration upon injury, normal breast undergoes regeneration or modification during the pregnancy-lactation-involution cycle. Therefore, it is not surprising that such stromal cells exist in the breast.

PZP cells also displayed stromal fibroblast-like properties. Morsing M. et al. reported that a distinct fibroblast lineage in the human breast, which supports luminal epithelial progenitor growth and possessing adipogenic capacity, gathers around terminal ductal lobular units (TDLUs)²⁸. These cells are likely derived from lobular origin, which is evident by the enrichment of CD105^{high}/CD26^{low} population. We found that PZP cells stably maintain the CD105^{high}/CD26^{low} phenotype with extended culture. Lobular-like fibroblasts regulate epithelial morphogenesis and differentiation typical of the TDLU through TGFβ signaling pathway^{27, 28}. Early myoepithelial progenitors are susceptible to cues from lobular-like and interlobular-like fibroblasts in terms of luminal differentiation repertoire^{27, 28}. Since PZP cells expressed several secretory factors, more than luminal cells, few of these factors could provide specific cues to luminal cells.

Co-culture experiments provided a clue to how alterations in number of PZP cells in the breast could influence the breast microenvironment. For example, we observed a higher level of TAGLN expression under co-culture conditions compared to either cell types cultured separately. Recent studies have shown that TAGLN localizes in the cytoplasm of benign and malignant tumor cells in the breast tissue³⁶. It is strongly expressed in the cytoplasm of myoepithelial and fibroblastic cells of the benign breast tissue; however normal luminal cells are predominately negative or have a very weak cytoplasmic expression. TAGLN differentially expressed among molecular breast cancer subtypes and predominantly upregulated in higher grade triple negative breast cancers³⁶. Expression of TAGLN is positively correlated with high Ki-67 and low ER and PR status. TAGLN expression is found in 33% of the invasive carcinoma³⁶. Other studies showed that TAGLN is directly associated with migratory and invasive ability of tumorigenic cells

and TAGLN overexpression increases the invasiveness of both tumorigenic and nontumorigenic cells ⁵⁶.

PZP cell-epithelial cell interaction leading to elevated local production of IL-6 could have a profound effect on the tissue microenvironment. In this respect, genetic ancestry-dependent variability in breast tumor immune microenvironment has already been demonstrated, particularly in relation to CD8⁺ T-cell exhaustion ⁵⁷. IL-6 is pleiotropic cytokine that contributes to EMT and drives metastasis in ER⁺ breast cancer ⁵⁸. The IL-6/JAK/STAT3 pathway plays a major role in tumor growth and development of several human cancers. Previous studies report a higher levels of IL-6 in chronic inflammatory conditions, such as inflammatory bowel disease, rheumatoid arthritis, and in various cancers including breast cancer ⁵⁹. IL-6 activates the JAK/STAT3 signaling, which drives the tumor proliferation, survival, invasiveness, angiogenesis and immunosuppression ⁶⁰. Consistent with the possibility of genetic ancestry-dependent variability in PZP cells causing elevated local IL-6 levels and downstream signaling, we observed elevated phospho-STAT3 levels in normal and tumors of AA women compared to EA women.

Co-culturing of epithelial cells with PZP cells caused changes in the phenotype of epithelial cells with a fraction of luminal progenitor cells acquiring basal like properties. Recent studies have suggested that luminal cells with basal cell properties are enriched for gene signatures of basal-like breast cancers and number of cells with this property increase with aging ¹⁷. Whether locally produced IL-6 is responsible for luminal cells acquiring basal phenotype remains to be determined although there is evidence in the literature for IL-6 conferring plasticity to epithelial cells ⁶¹. Currently, at least four IL-6 targeting drugs

are either approved or in clinical trial⁶². It remains to be determined whether IL-6 targeting drugs are more effective against breast tumors of AA women compared to EA women.

Metaplastic carcinoma of the breast, although rare, is more common in AA women compared to EA women¹³. Our results indicate that PZP cells are the cell-of-origin of this cancer type. Identifying cell-of-origin allows further evaluation of signaling pathways activated in this cancer type and eventually development of targeted therapies. Metaplastic carcinomas of the breast are CD49f⁺/EpCAM⁻⁶³ and can be subdivided into metaplastic, claudin- (META-CLOW) and metaplastic-squamous, and metaplastic-non-CLOW, non-Squamous. PZP cells acquired CD49f⁺/EpCAM⁻ phenotype upon transformation and tumors derived from these cells displayed squamous features, indicating PZP cells are the cell-of-origin of metaplastic-squamous subtype.

In summary, this study identified and established PZP cell lines from healthy breasts of AA women and these cells phenotypically resemble tissue resident mesenchymal stromal/stem cells with transdifferentiation capabilities and can regulate signaling in neighboring epithelial cells. Although we identified only two factors, TAGLN and IL-6, whose levels in the tissue microenvironment could be influenced due to PZP-epithelial cell interaction, there may be additional signaling pathway activation as a consequence of this interaction and potentially influence the microenvironment. Because of the ability of PZP cells to influence signaling in epithelial cells and differences in their number based on genetic ancestry, signaling pathway in epithelial cells of the breast may show genetic ancestry-dependent variability in signaling under both normal and cancerous condition and drugs that target crosstalk between these two cell types can be new therapeutic agents. These drugs may show genetic ancestry-dependent variability in their potency. Therefore,

in addition to including patients of different genetic ancestry in clinical trials, evaluation of outcomes of the clinical trial needs to take genetic ancestry into consideration. Some therapeutics, either as monotherapy or in combination, may be effective only in patients of specific genetic ancestry. Mechanistically, how PZP cell numbers are elevated in healthy breasts of AA women is unknown. FAPs, which are similar to PZP cells, increase in numbers in skeletal muscle of individuals with type II diabetes²⁵. Whether type II diabetes, which is more common in AA than EA⁶⁴, is responsible for increased PZPs in the breasts of AA women remains to be determined.

ACKNOWLEDGMENTS

We thank members of the Indiana University Simon Comprehensive Cancer Center (IUSCCC) flow cytometry core, animal facility, and tissue procurement cores and the Susan G Komen Tissue Bank for various tissues and reagents. We also thank members of the Komen Tissue Bank and IUSCCC tissue bank including Jill Henry, Mary Cox, Rana German, Rockey Pam, and Julia von Arx for their work in generating TMA and associated data. We thank Dr. Natascia Marino for contributing part of EA TMA. We thank countless number of women for donating their breast tissue for research purpose as well as volunteers who facilitated tissue collection. This work is supported by DOD- W81XWH-15-1-0707, DOD-W81XWH-20-1-0577 and Susan G. Komen for the Cure (DRS20645418) to HN. Susan G. Komen for the Cure, Breast Cancer Research Foundation and Vera Bradley Foundation for Breast Cancer Research provide funding support to Komen Normal Tissue Bank.

AUTHOR CONTRIBUTIONS

Conception and design: BK, HN

Development of methodology: BK, PBN, LAB, CJT, AMS, HN

Acquisition of data: BK, KB, PBN, MMG, RJA, MS, HN

Analysis and interpretation of data: BK, KB, PBN, GS, SKA, HN

Writing, review and/or revision of the manuscript: BK, HN

Administrative, technical, or material support: HN, AMS, PBN

Study supervision: HN

DECLARATION OF INTERESTS

Authors have no conflict of interest to declare.

DATA AVAILABILITY: This manuscript does not report any high throughput data. All reagents created as part of the manuscript will be distributed upon meeting requirements of the Indiana University.

Table 1. Compare H-score and Positivity between race in all patients within Normal tissue for PROCR, ZEB1 and PDGFR α

Marker	variable label	Race				P-value
		COLUMN_OVERALL	African American	European	Latino	
PROCR	Positivity	0.17 (0.04, 0.55)	0.30 (0.06, 0.55) (N=31)	0.15 (0.04, 0.41) (N=129)	0.17 (0.04, 0.39) (N=33)	<.0001
	H-Score	27.97 (5.45, 127.65)	56.41 (10.46, 127.65)	24.37 (6.68, 83.50)	26.06 (5.45, 72.30)	<.0001
ZEB1	Positivity	0.01 (0.00, 0.24)	0.01 (0.00, 0.10) (N=33)	0.01 (0.00, 0.12) (N=144)	0.02 (0.00, 0.24) (N=28)	0.0380
	H-Score	1.62 (0.15, 30.72)	2.21 (0.24, 16.47)	1.46 (0.15, 19.76)	2.75 (0.40, 30.72)	0.0076
PDGFR α	Positivity	0.10 (0.01, 0.75)	0.28 (0.03, 0.75) (N=35)	0.09 (0.01, 0.73) (N=154)	0.09 (0.02, 0.56) (N=18)	<.0001
	H-Score	16.30 (1.43, 110.23)	36.96 (6.25, 110.23)	14.27 (1.43, 88.61)	12.61 (2.69, 73.08)	<.0001

Note: Values expressed as median (min, max)

Note: P-value comparisons across race categories are based on Kruskal-Wallis

Table 2. Cytokine/chemokine array identified factors secreted by luminal and PZP cells either alone or together.

Shared	Angiogenin, Angiopoietin 2, BDNF, DKK1, Emmprin, IFN γ , IGFBP2, IGFBP3, IL-8, kallikrein 3, MIF, Osteopontin, Pentraxin 3, Sperine E1, ST2 and VEGF
Luminal and in co-culture	GDF15, GRO α , IL-1 α , Lipocalcin 2, MMP-9, PDGF-AA
PZP and in co-culture	Angiopoietin 1, Complement factor D, Cystatin C, EGF, ENA-78 (CXCL5), Endoglin, FGF19, HGF, SDF-1 α , Thrombospondin-1, uPAR
Coculture-specific	IL-6

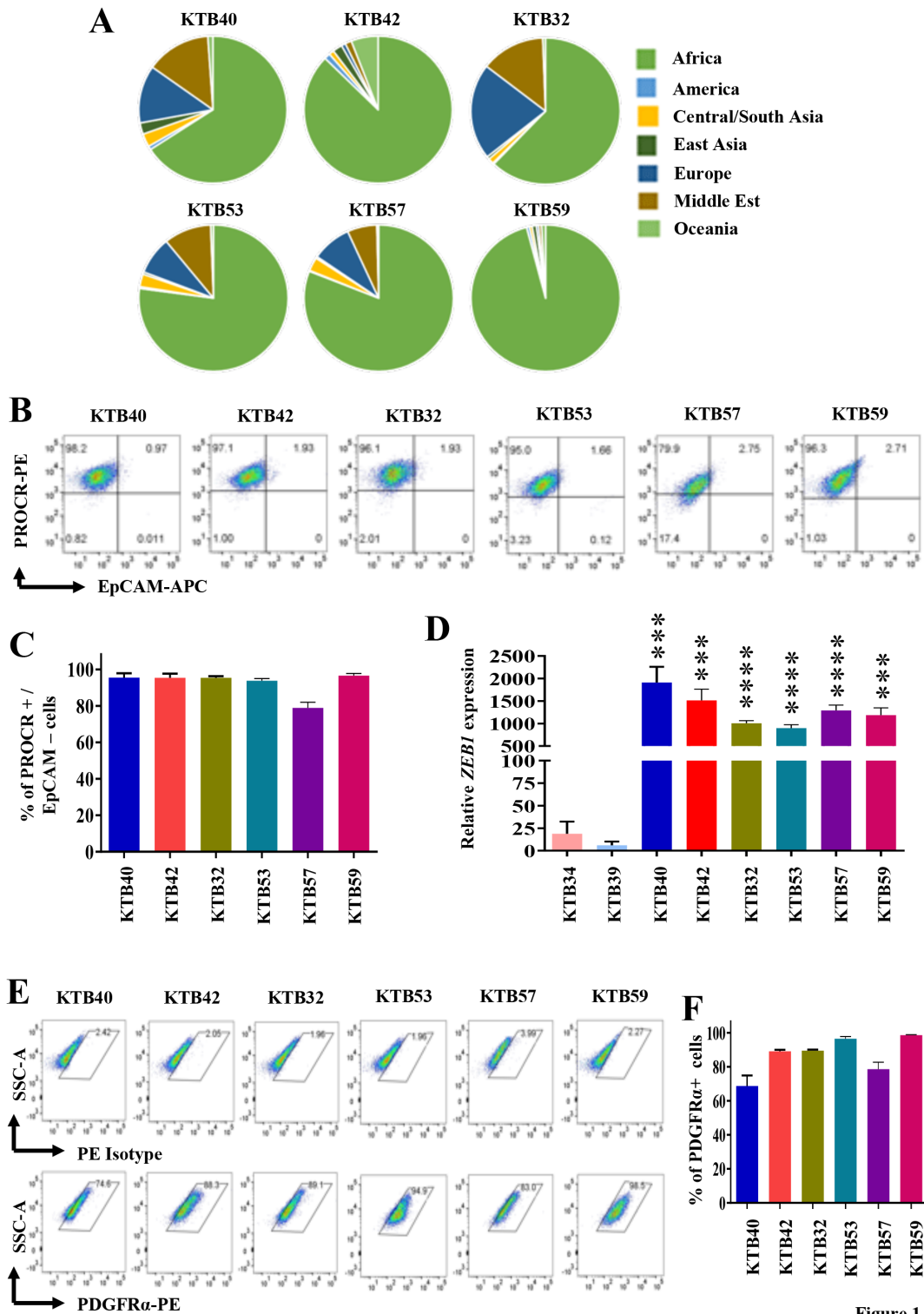


Figure 1

Figure 1: Establishment of PROCR⁺/ZEB1⁺/PDGFR α ⁺ (PZP) cell lines from the

healthy breast tissues of AA women. (A) Genetic ancestry mapping of breast tissue donors (KTB40, KTB42, KTB32, KTB53, KTB57, and KTB59) using a panel of 41-SNP. (B) PROCR⁺/EpCAM⁻ cells are enriched in established cell lines from AA women. (C) Quantitation of PROCR⁺/EpCAM⁻ cells. (D) ZEB1 expression levels in various PROCR⁺/EpCAM⁻ cell lines compared to EpCAM⁺ (KTB34 and KTB39) luminal cell lines. (E) PROCR⁺/ZEB1⁺ cell lines express PDGFR α as determined by flow cytometry. (F) Quantitation of PDGFR α ⁺ cells. ***p<0.001, ****p<0.0001 by ANOVA.

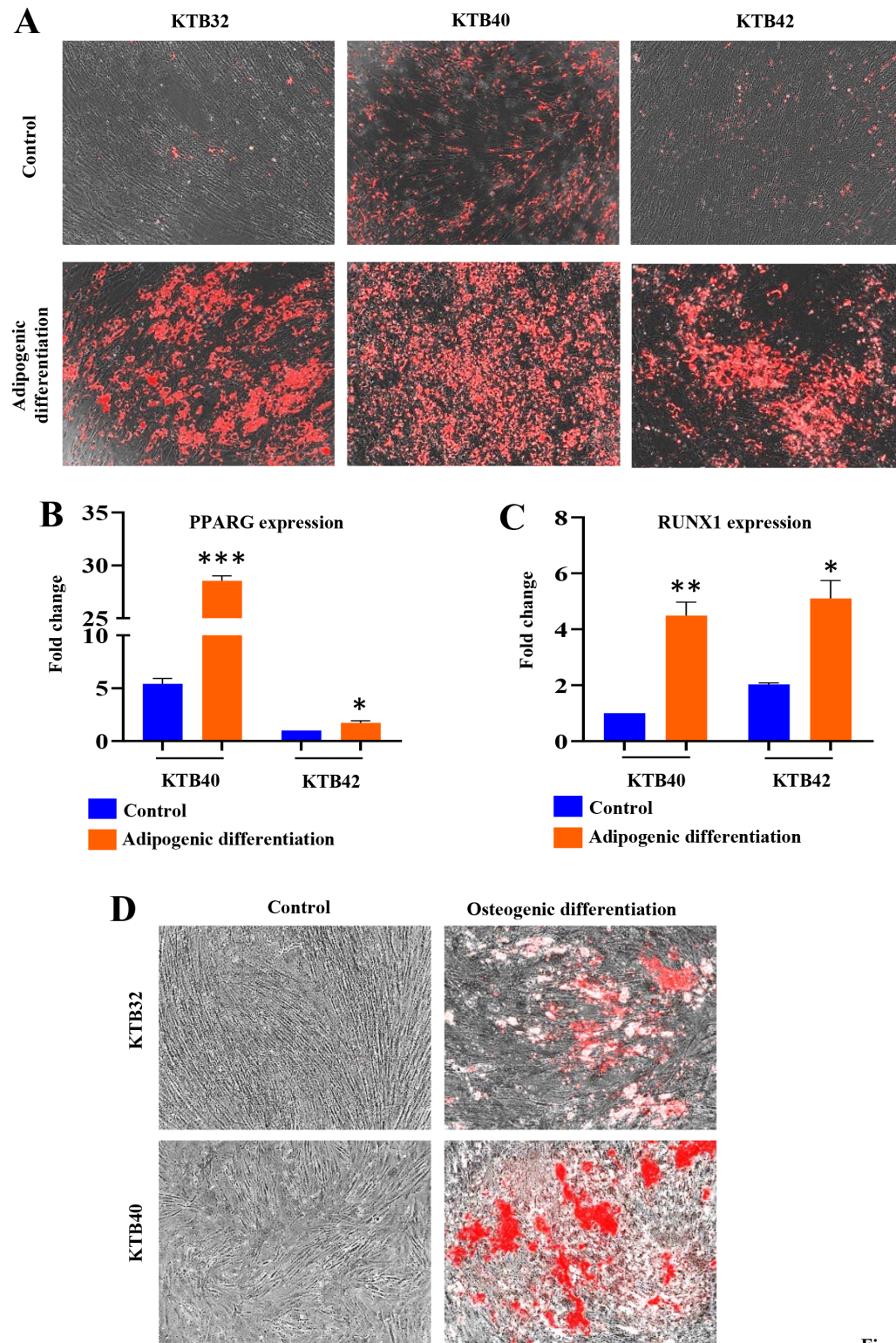


Figure 2

Figure 2: Trans-differentiating properties of PZP cells. (A) PZP cells undergo

adipogenic differentiation under appropriate growth condition. Neural lipids stain red upon Oil Red-O staining. (B) Adipogenic differentiated PZP cells show elevated level of adipocytes differentiation marker PPAR γ . (C) Adipogenic differentiation of PZP cells was confirmed through RUNX1 overexpression. (D) PZP cells undergo osteogenic differentiation under appropriate growth condition. Mineralization of matrix with Ca²⁺ was visualized by alizarin red staining. *p<0.05, **p<0.01, ***p<0.001 by ANOVA.

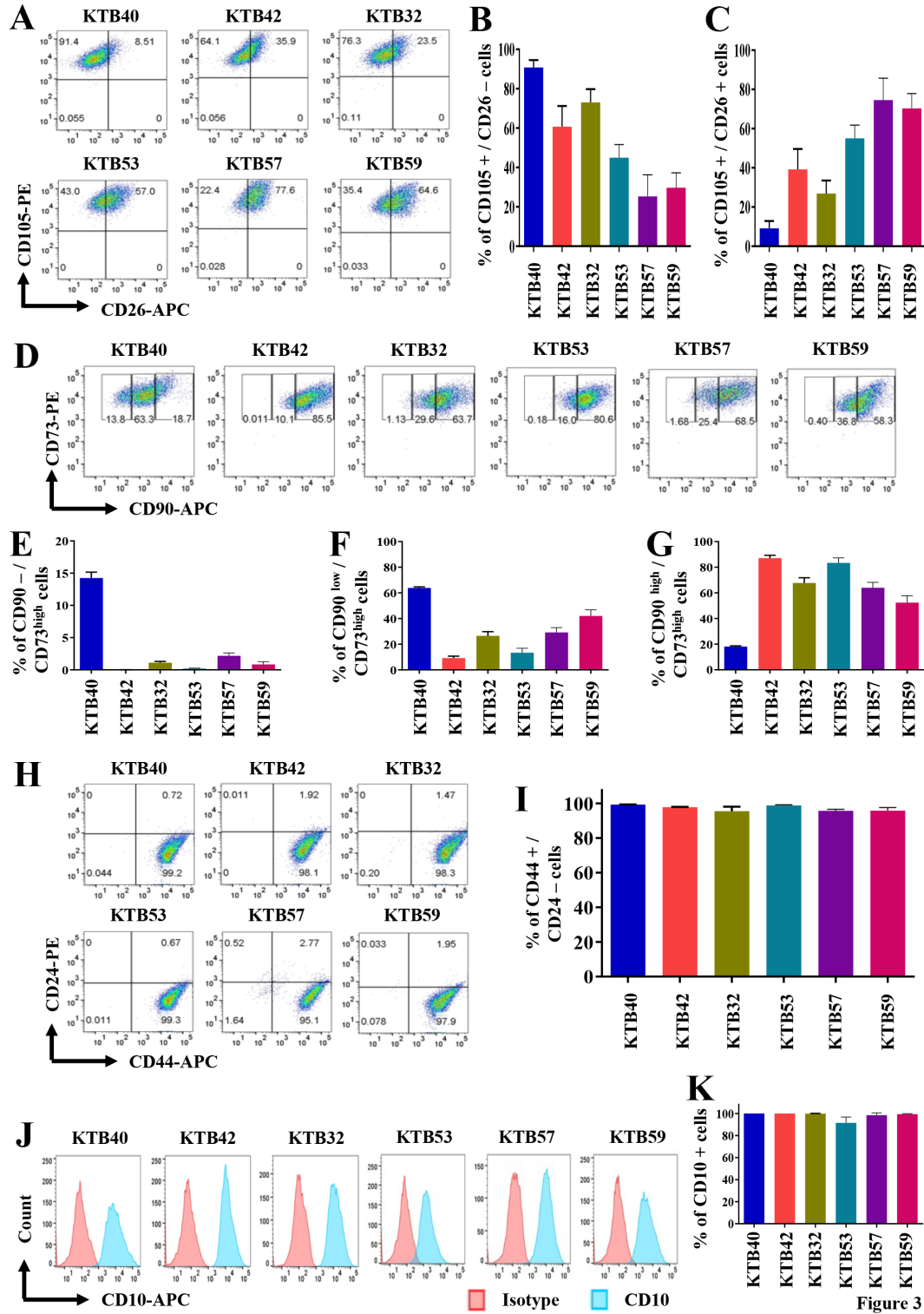


Figure 3

Figure 3: Phenotypic characterization of PZP cells. (A) PZP cell lines were stained with

CD105 and CD26 antibodies to identify the lobular and interlobular origin of human breast fibroblastic cells. (B and C) Quantification of CD105^{high}/CD26⁻ and CD105^{high}/CD26^{low} population of cells. (D) PZP cell lines were stained with CD90 and CD73 antibodies to identify rare endogenous pluripotent somatic stem cells and potential mesenchymal stem cells. (E, F and G) Quantification of CD90⁻/CD73^{high}, CD90^{low}/CD73^{high}, and CD90^{high}/CD73^{high} population of cells. (H) PZP cell lines were stained with CD44 and CD24 antibodies to determine whether their phenotype overlaps with cancer stem cells. (I) Quantification of CD44⁺/CD24⁻ population. (J) PZP cell lines were stained with CD10 antibody to determine overlap with myoepithelial cell marker expression. (K) Quantification of CD10⁺ population.

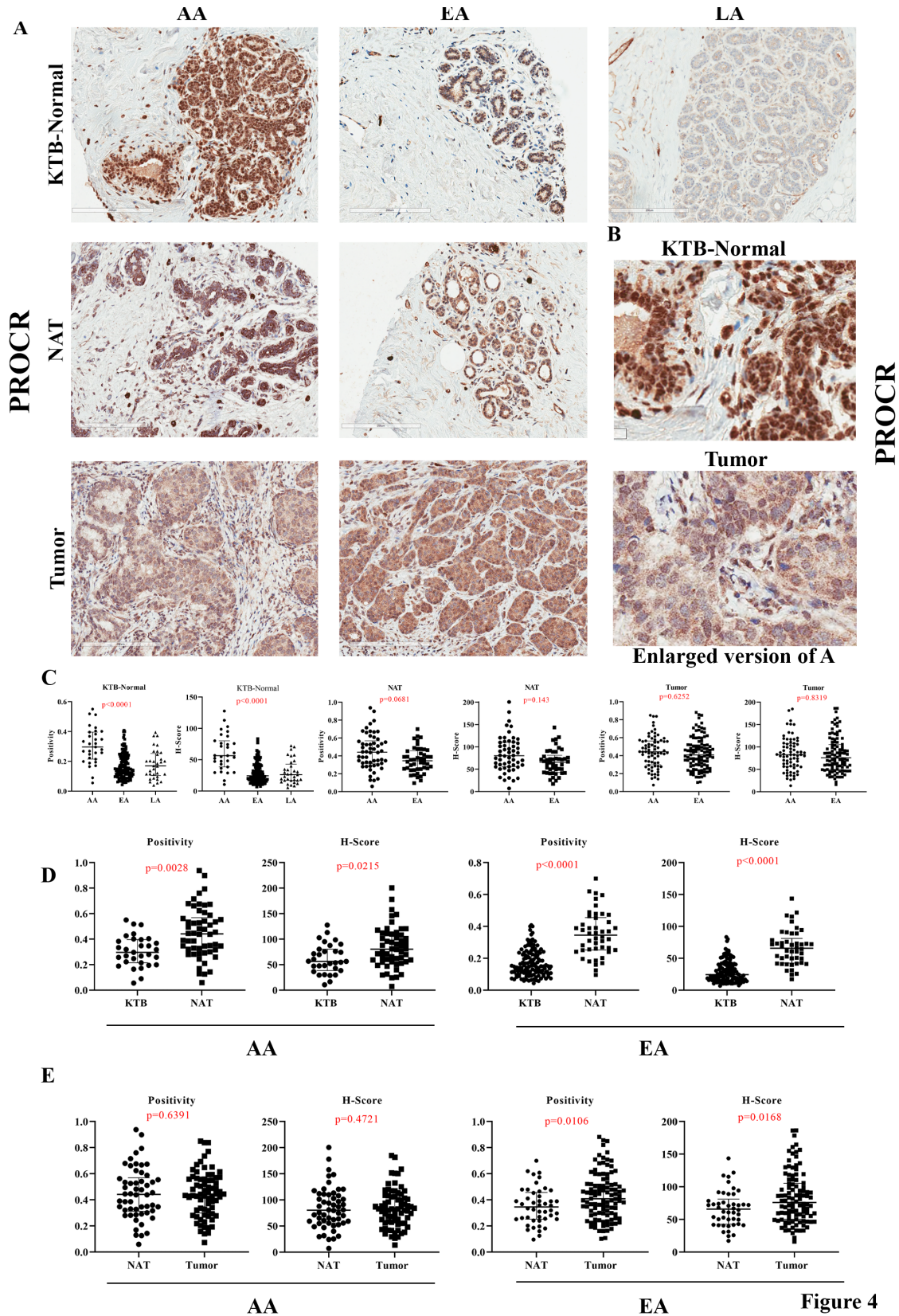


Figure 4: PROCR expression pattern in KTB-normal, NATs, and breast tumors. (A)

Representative IHC of PROCR in KTB-normal, NATs, and/or tumors of AA, EA, and LA women. (B) Enlarged view of PROCR expression in KTB-normal and tumor. (C) Differences in PROCR expression (positivity and H-score) between KTB-normal of AA, EA and LA women. Differences in PROCR expression (positivity and H-score) between KTB-normal, NATs and/or tumors of AA, EA, and LA women. (D) Differences in PROCR expression (positivity and H-score) between KTB-normal and NATs in AA and EA women. (E) Differences in PROCR expression (positivity and H-score) between NATs and tumors in AA and EA women.

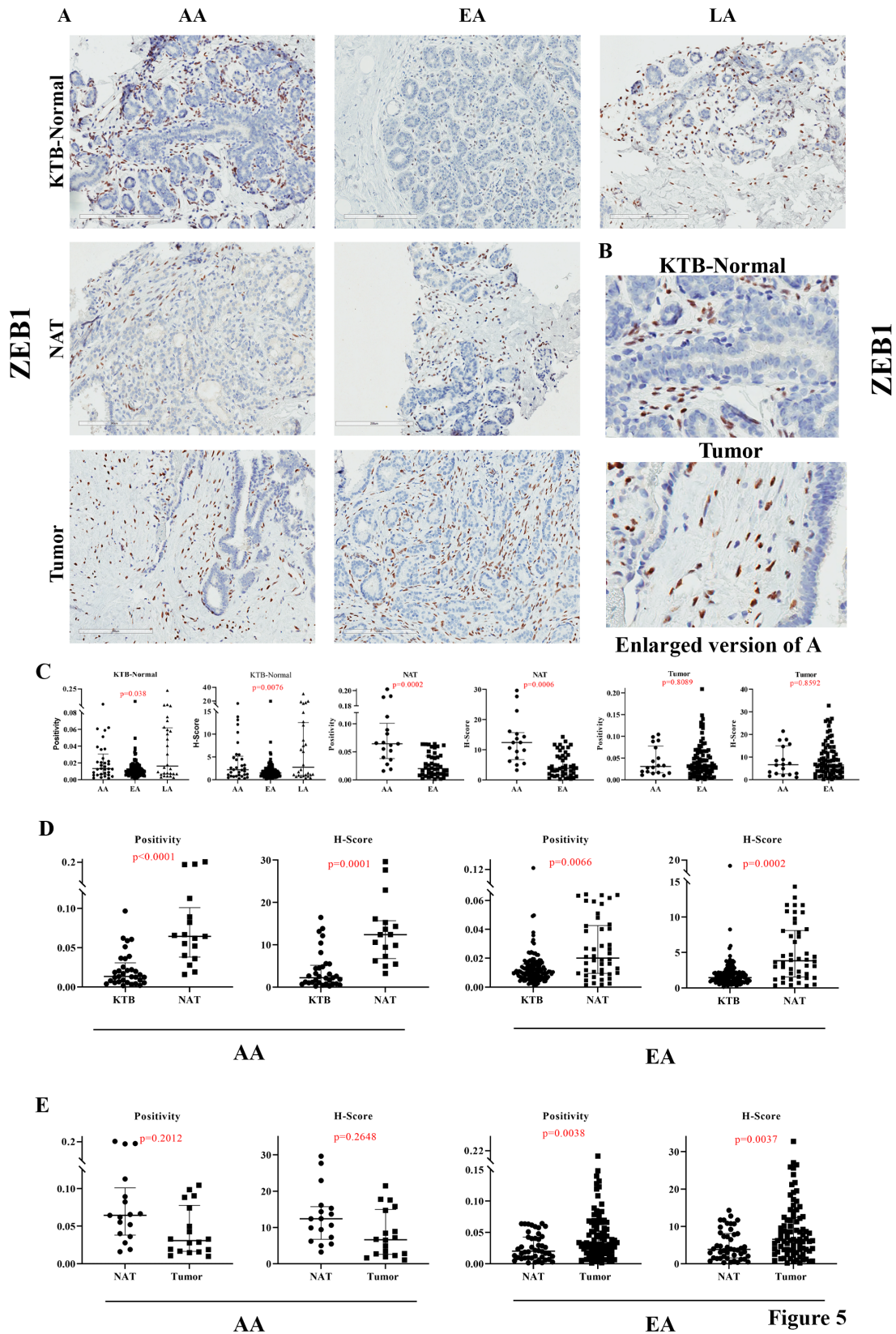


Figure 5: ZEB1 expression pattern in KTB-normal, NATs, and breast tumors. (A)

Representative IHC of ZEB1 in KTB-normal, NATs, and/or tumors of AA, EA and LA women. (B) Enlarged view of ZEB1 expression in KTB-normal and tumor. Note ZEB1⁺ cells surround epithelial cell clusters. (C) Differences in ZEB1 expression (positivity and H-score) between KTB-normal of AA, EA and LA women. Differences in ZEB1 expression (positivity and H-score) between KTB-normal, NATs and/or tumors of AA, EA, and LA women. (D) Differences in ZEB1 expression (positivity and H-score) between KTB-normal and NATs in AA and EA women. (E) Differences in ZEB1 expression (positivity and H-score) between NATs and tumors in AA and EA women.

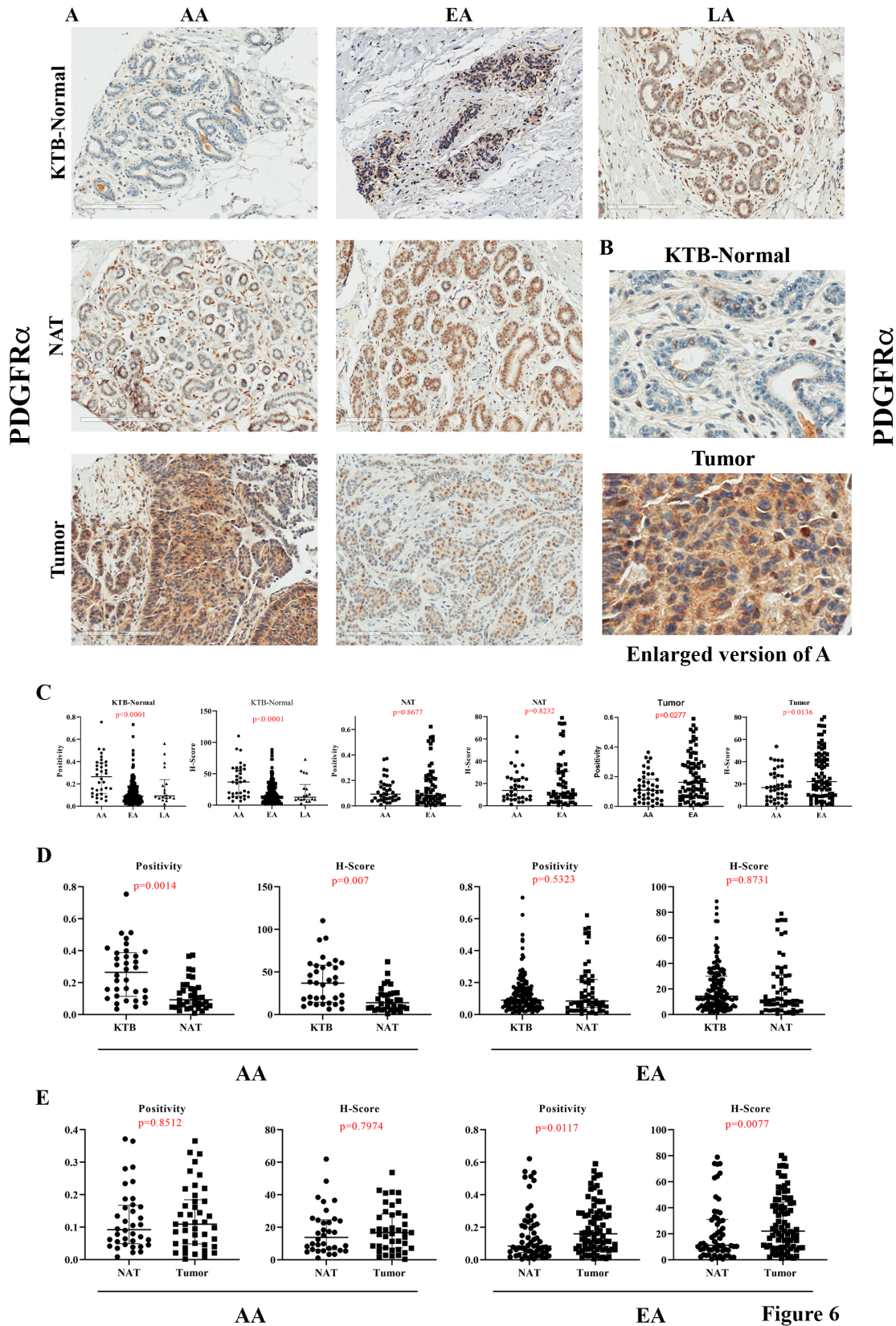


Figure 6: PDGFR α expression pattern in KTB-normal, NATs, and breast tumors. (A)

Representative IHC of PDGFR α in KTB-normal, NATs, and/or tumors of AA, EA, and LA women. (B) Enlarged view of PDGFR α expression in KTB-normal and tumor. (C) Differences in PDGFR α expression (positivity and H-score) between KTB-normal of AA, EA, and LA women. Differences in PDGFR α expression (positivity and H-score) between KTB-normal, NATs and/or tumors of AA, EA, and LA women. (D) Differences in PDGFR α expression (positivity and H-score) between KTB-normal and NATs in AA and EA women. (E) Differences in PDGFR α expression (positivity and H-score) between NATs and tumors in AA and EA women.

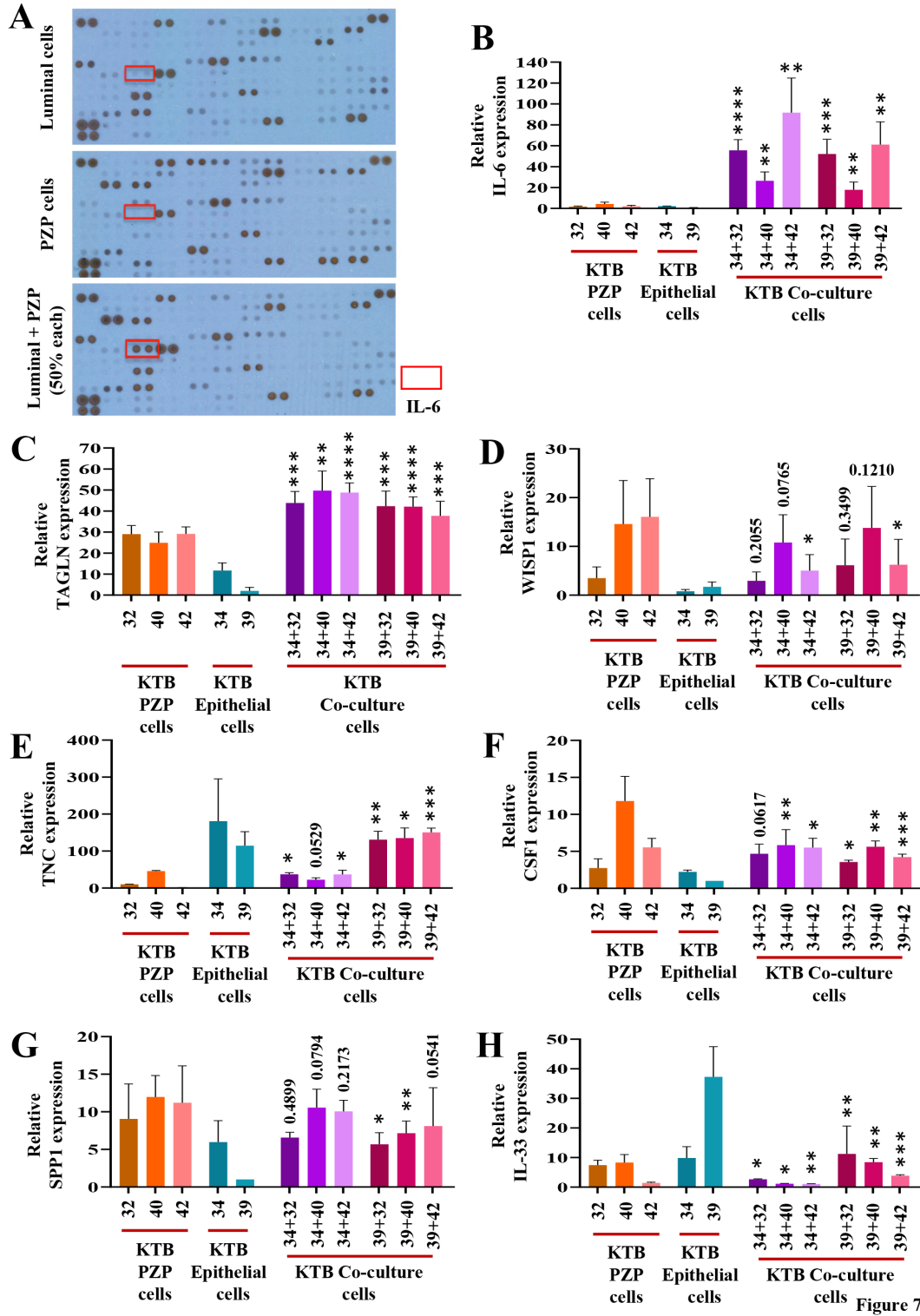


Figure 7: PZP-Epithelial cells interaction alters the gene expression. (A) IL-6

expression was detected only when luminal and PZP cells were co-cultured. R&D systems cytokine/chemokine array was used to identify secreted factors by luminal and PZP cells either alone or together (50% each). Expression of IL-6 (B), TAGLN (C), WISP1 (D), TNC (E), CSF1 (F), SPP1 (G), and IL-33 (H) in PZP cell lines (KTB32, KTB40, KTB42), luminal progenitor (epithelial cells; KTB34, KTB39), and co-culture of PZP and luminal progenitor cell lines. Statistical significance (p values) was determined by comparing KTB32/40/42 (PZP cells), KTB34/KTB39 (epithelial cells) and respective co-cultured PZP + epithelial cells as indicated in the figure. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$ by ANOVA.

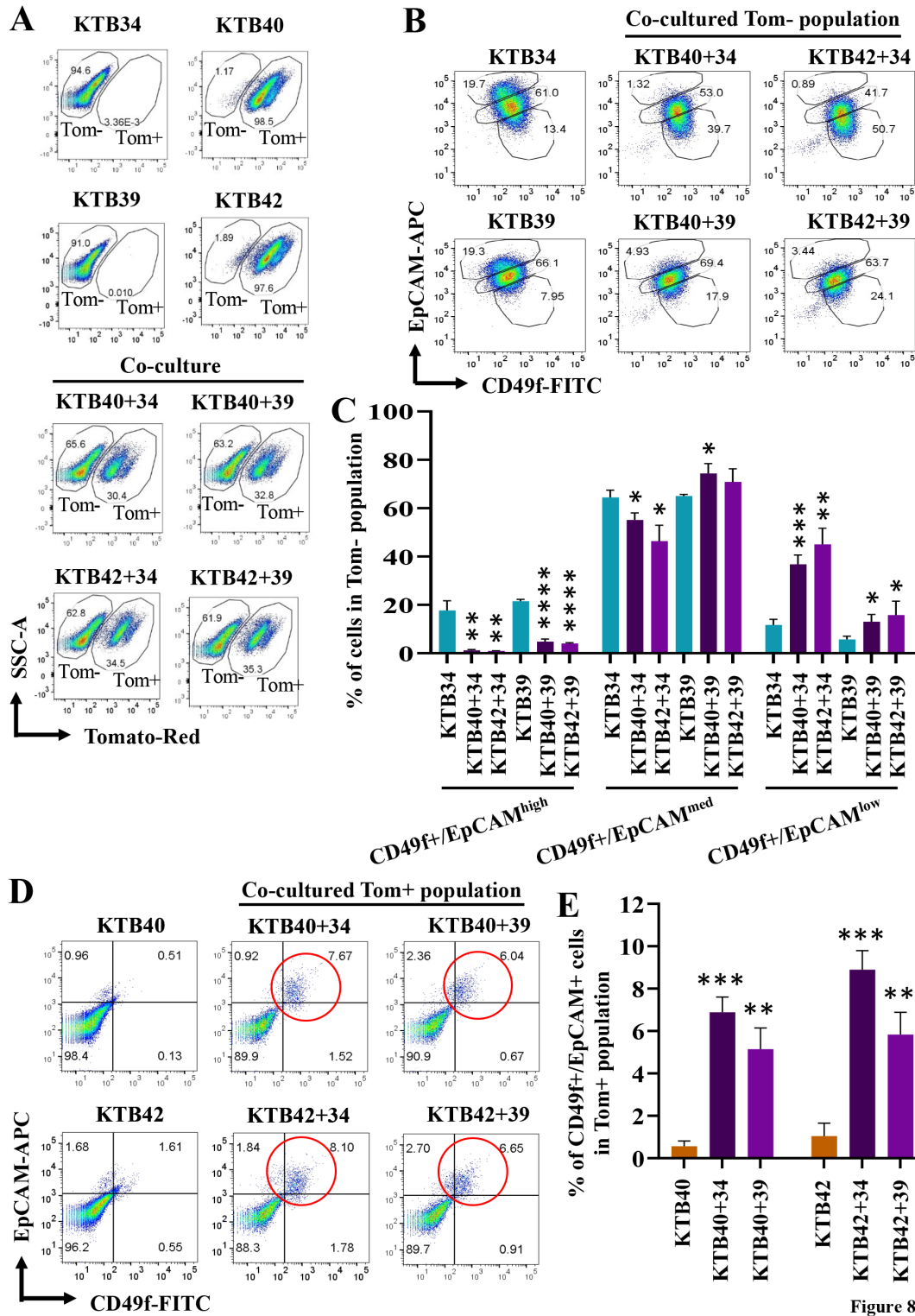


Figure 8

Figure 8: The effects of PZP cells on trans-differentiation of epithelial luminal

progenitor cells. (A) Gating of tomato- (Tom-) and tomato+ (Tom+) populations of KTB34, KTB39, and co-cultured KTB40/KTB42 and KTB34/KTB39 cell lines. (B) CD49f and EpCAM staining patterns of KTB34, KTB39, and co-cultured KTB40/KTB42 and KTB34/KTB39 cell lines. (C) Quantification of CD49f⁺/EpCAM^{high}, CD49f⁺/EpCAM^{med}, and CD49f⁺/EpCAM^{low} populations in Tom- cell population. (D) CD49f and EpCAM staining patterns of KTB40, KTB42, and co-cultured KTB40/KTB42 and KTB34/KTB39 cell lines. Only Tom+ population was analyzed. (E) Quantification of CD49f⁺/EpCAM⁺ population among Tom+ population. *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001 by ANOVA.

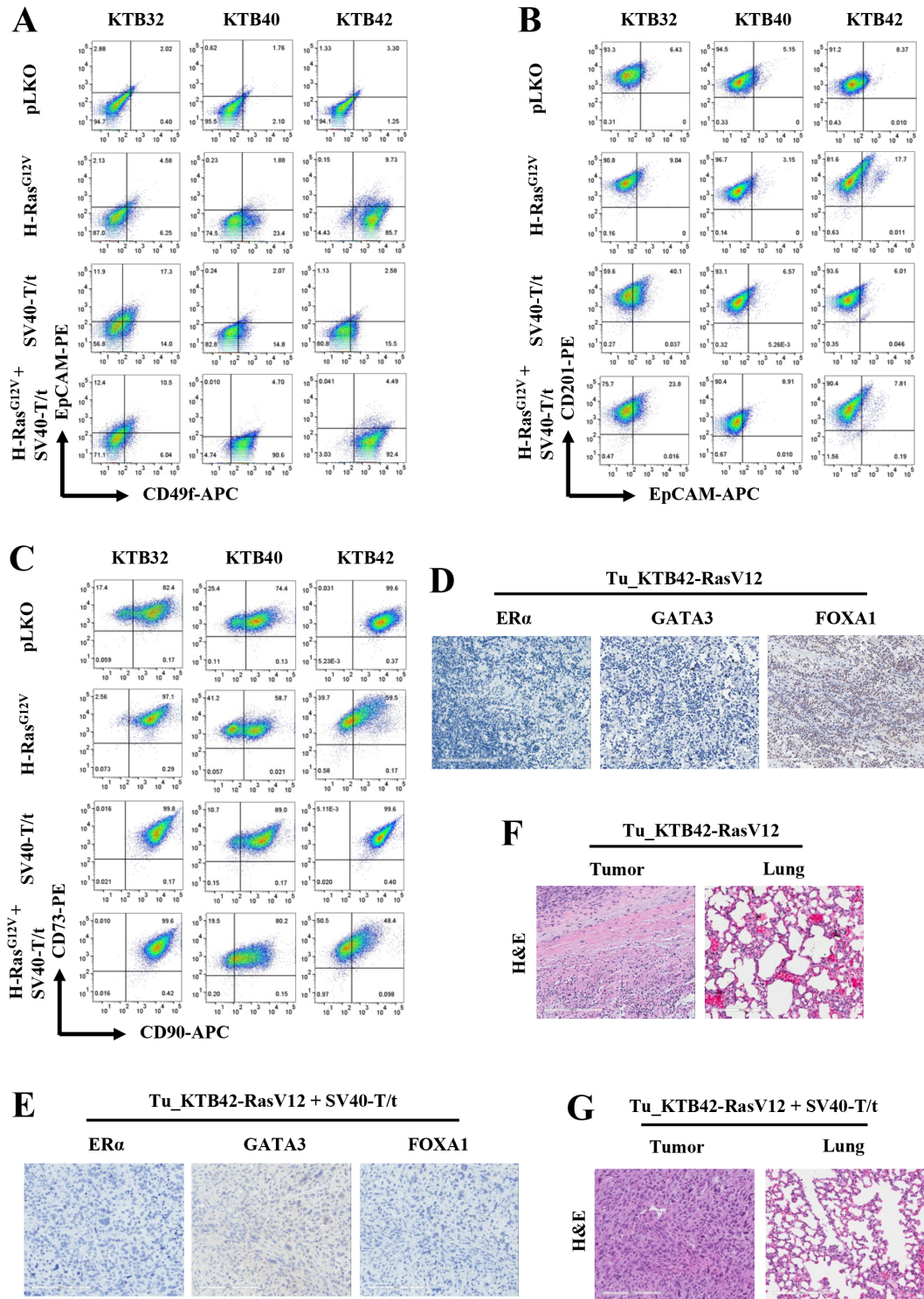


Figure 9

Figure 9: PZP cells transformed with Ras and SV40-T/t antigen are tumorigenic. (A)

CD49f and EpCAM staining patterns of immortalized and transformed PZP (KTB32, KTB40 and KTB42) cell lines. CD49f⁺/EpCAM⁻, CD49f⁺/EpCAM⁺ and CD49f⁻/EpCAM⁺ cells correspond to stem/basal, progenitor and differentiated cells, respectively. (B) CD201 and EpCAM staining patterns of immortalized and transformed PZP (KTB32, KTB40, and KTB42) cell lines. (C) CD90 and CD73 staining patterns of immortalized and transformed PZP (KTB32, KTB40, and KTB42) cell lines. (D) IHC analyses of luminal markers ER α , GATA3 and FOXA1 in tumors developed from KTB42-HRas^{G12V} transformed cells. (E) IHC analyses of luminal markers ER α , GATA3 and FOXA1 in tumors developed from the KTB42-HRas^{G12V}+SV40-T/t antigen transformed cells. (F) Tumor developed from KTB42-HRas^{G12V} transformed cells did not show extensive lung metastasis. (G) Tumor developed from KTB42-HRas^{G12V}+SV40-T/t antigen transformed cells did not show lung metastasis. H&E staining shows the type of tumors.

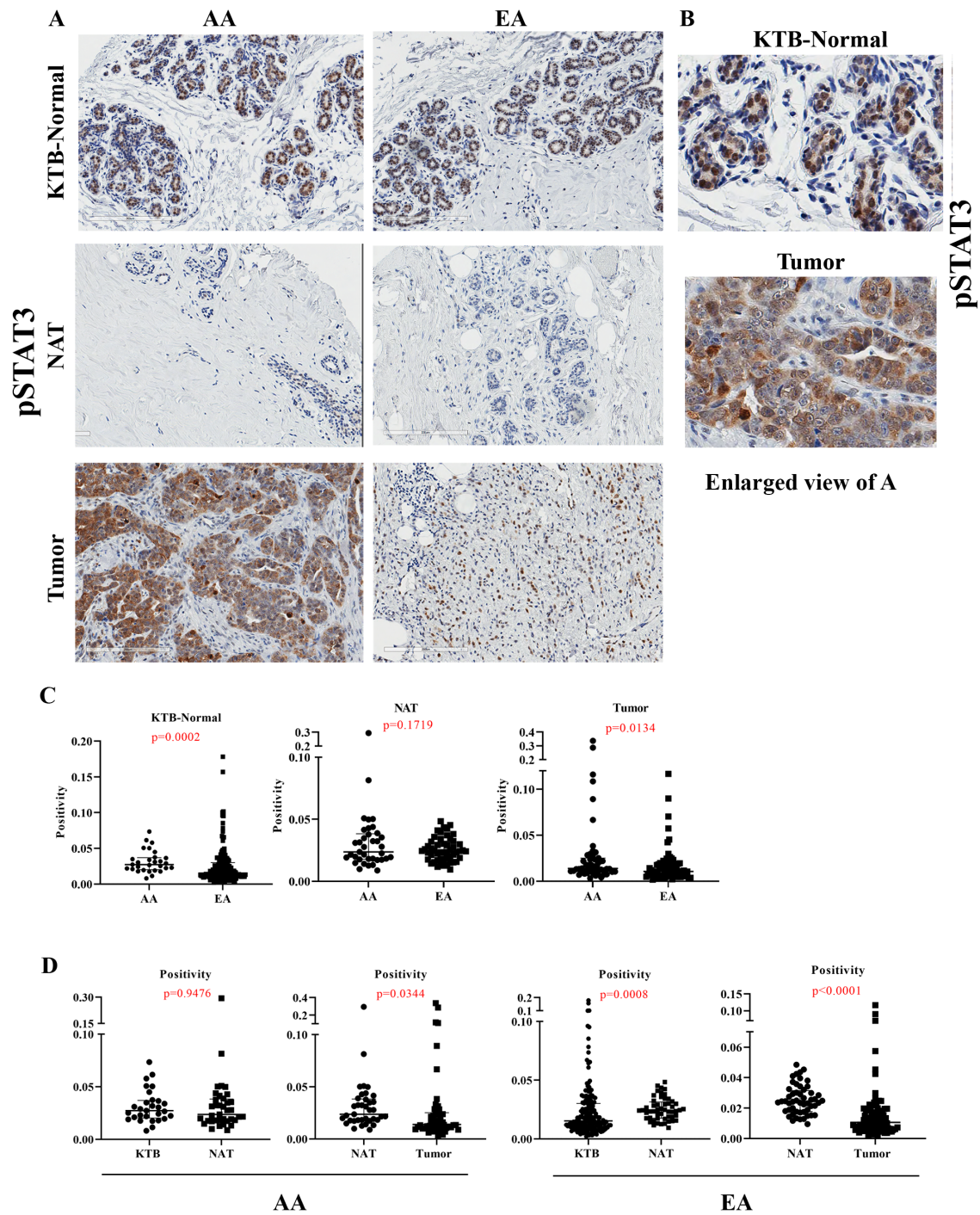


Figure 10

Figure 10: Phospho-STAT3 levels in KTB-normal, NATs, and breast tumors. (A)

Representative IHC of pSTAT3 in KTB-normal, NATs, and/or tumors of women of AA, and EA. (B) Enlarged view of pSTAT3 expression in KTB-normal and tumor. (C) Differences in pSTAT3 positivity between KTB-normal, NATs and tumors of AA and EA women (D) Differences in pSTAT3 positivity between KTB-normal and NATs and between NATs and tumors of AA and EA women.

REFERENCES

1. Newman LA, Griffith KA, Jatoi I, Simon MS, Crowe JP, Colditz GA. Meta-analysis of survival in African American and white American patients with breast cancer: ethnicity compared with socioeconomic status. *J Clin Oncol* **24**, 1342-1349 (2006).
2. Paik S, *et al.* A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med* **351**, 2817-2826 (2004).
3. Albain KS, *et al.* Race, Ethnicity, and Clinical Outcomes in Hormone Receptor-Positive, HER2-Negative, Node-Negative Breast Cancer in the Randomized TAILORx Trial. *J Natl Cancer Inst* **113**, 390-399 (2021).
4. Narod SA, Iqbal J, Giannakeas V, Sopik V, Sun P. Breast Cancer Mortality After a Diagnosis of Ductal Carcinoma In Situ. *JAMA Oncol* **1**, 888-896 (2015).
5. Sherman RM, *et al.* Assembly of a pan-genome from deep sequencing of 910 humans of African descent. *Nat Genet* **51**, 30-35 (2019).
6. Nakshatri H, Anjanappa M, Bhat-Nakshatri P. Ethnicity-Dependent and -Independent Heterogeneity in Healthy Normal Breast Hierarchy Impacts Tumor Characterization. *Scientific reports* **5**, 13526 (2015).
7. Wang D, *et al.* Identification of multipotent mammary stem cells by protein C receptor expression. *Nature* **517**, 81-84 (2014).
8. Nakshatri H, *et al.* Genetic ancestry-dependent differences in breast cancer-induced field defects in the tumor-adjacent normal breast. *Clin Cancer Res* **25**, 2848-2859 (2019).
9. De Cock JM, Shibue T, Dongre A, Keckesova Z, Reinhardt F, Weinberg RA. Inflammation Triggers Zeb1-Dependent Escape from Tumor Latency. *Cancer Res* **76**, 6778-6784 (2016).
10. Block CJ, Dyson G, Campeanu IJ, Watza D, Ratnam M, Wu G. A stroma-corrected ZEB1 transcriptional signature is inversely associated with antitumor immune activity in breast cancer. *Scientific reports* **9**, 17807 (2019).

11. Morel AP, *et al.* EMT inducers catalyze malignant transformation of mammary epithelial cells and drive tumorigenesis towards claudin-low tumors in transgenic mice. *PLoS genetics* **8**, e1002723 (2012).
12. Pezzi CM, Patel-Parekh L, Cole K, Franko J, Klimberg VS, Bland K. Characteristics and treatment of metaplastic breast cancer: analysis of 892 cases from the National Cancer Data Base. *Ann Surg Oncol* **14**, 166-173 (2007).
13. Zawati I, *et al.* Association of ZEB1 and Vimentin with poor prognosis in metaplastic breast cancer. *Ann Diagn Pathol* **59**, 151954 (2022).
14. 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).
15. Joshi PA, *et al.* PDGFRalpha(+) stromal adipocyte progenitors transition into epithelial cells during lobulo-alveologenesis in the murine mammary gland. *Nature communications* **10**, 1760 (2019).
16. Kim W, *et al.* RUNX1 is essential for mesenchymal stem cell proliferation and myofibroblast differentiation. *Proc Natl Acad Sci U S A* **111**, 16389-16394 (2014).
17. Gray GK, *et al.* A human breast atlas integrating single-cell proteomics and transcriptomics. *Dev Cell* **57**, 1400-1420 e1407 (2022).
18. Shalabi SF, *et al.* Evidence for accelerated aging in mammary epithelia of women carrying germline BRCA1 or BRCA2 mutations. *Nat Aging* **1**, 838-849 (2021).
19. Salemis NS. Metaplastic carcinoma of the breast with mesenchymal differentiation (carcinosarcoma). A unique presentation of an aggressive malignancy and literature review. *Breast Dis* **37**, 169-175 (2018).
20. Prasad M, *et al.* Dual TGFbeta/BMP Pathway Inhibition Enables Expansion and Characterization of Multiple Epithelial Cell Types of the Normal and Cancerous Breast. *Mol Cancer Res* **17**, 1556-1570 (2019).
21. Kumar B, *et al.* Normal breast-derived epithelial cells with luminal and intrinsic subtype-enriched gene expression document inter-individual differences in their differentiation cascade. *Cancer Res* **78**, 5107-5123 (2018).

22. Kumar B, *et al.* Bidirectional Regulatory Cross-Talk between Cell Context and Genomic Aberrations Shapes Breast Tumorigenesis. *Mol Cancer Res*, (2021).
23. Nievergelt CM, *et al.* Inference of human continental origin and admixture proportions using a highly discriminative ancestry informative 41-SNP panel. *Investig Genet* **4**, 13 (2013).
24. Marino N, *et al.* FAM83A is a potential biomarker for breast cancer initiation. *Biomark Res* **10**, 8 (2022).
25. Farup J, *et al.* Human skeletal muscle CD90(+) fibro-adipogenic progenitors are associated with muscle degeneration in type 2 diabetic patients. *Cell metabolism* **33**, 2201-2214 e2211 (2021).
26. Cardiff RD, Wellings SR. The comparative pathology of human and mouse mammary glands. *J Mammary Gland Biol Neoplasia* **4**, 105-122 (1999).
27. Morsing M, *et al.* Evidence of two distinct functionally specialized fibroblast lineages in breast stroma. *Breast Cancer Res* **18**, 108 (2016).
28. Morsing M, *et al.* Fibroblasts direct differentiation of human breast epithelial progenitors. *Breast Cancer Res* **22**, 102 (2020).
29. Roy S, *et al.* Rare somatic cells from human breast tissue exhibit extensive lineage plasticity. *Proc Natl Acad Sci U S A* **110**, 4598-4603 (2013).
30. Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF. Prospective identification of tumorigenic breast cancer cells. *Proc Natl Acad Sci U S A* **100**, 3983-3988 (2003).
31. Morel AP, *et al.* A stemness-related ZEB1-MSRB3 axis governs cellular pliancy and breast cancer genome stability. *Nat Med* **23**, 568-578 (2017).
32. Su S, *et al.* CD10(+)GPR77(+) Cancer-Associated Fibroblasts Promote Cancer Formation and Chemoresistance by Sustaining Cancer Stemness. *Cell* **172**, 841-856 e816 (2018).
33. Visvader JE, Stingl J. Mammary stem cells and the differentiation hierarchy: current status and perspectives. *Genes Dev* **28**, 1143-1158 (2014).

34. Eeckhoutte J, Keeton EK, Lupien M, Krum SA, Carroll JS, Brown M. Positive cross-regulatory loop ties GATA-3 to estrogen receptor alpha expression in breast cancer. *Cancer Res* **67**, 6477-6483 (2007).
35. Bhat-Nakshatri P, Newton TR, Goulet R, Jr., Nakshatri H. NF-kappaB activation and interleukin 6 production in fibroblasts by estrogen receptor-negative breast cancer cell-derived interleukin 1alpha. *Proc Natl Acad Sci U S A* **95**, 6971-6976 (1998).
36. Rao D, Kimler BF, Nothnick WB, Davis MK, Fan F, Tawfik O. Transgelin: a potentially useful diagnostic marker differentially expressed in triple-negative and non-triple-negative breast cancers. *Hum Pathol* **46**, 876-883 (2015).
37. Tawfik O, Rao D, Nothnick WB, Graham A, Mau B, Fan F. Transgelin, a Novel Marker of Smooth Muscle Differentiation, Effectively Distinguishes Endometrial Stromal Tumors from Uterine Smooth Muscle Tumors. *Int J Gynecol Obstet Reprod Med Res* **1**, 26-31 (2014).
38. Holloway EM, *et al.* Mapping Development of the Human Intestinal Niche at Single-Cell Resolution. *Cell Stem Cell* **28**, 568-580 e564 (2021).
39. Tao W, *et al.* Dual Role of WISP1 in maintaining glioma stem cells and tumor-supportive macrophages in glioblastoma. *Nature communications* **11**, 3015 (2020).
40. Yalcin F, Dzaye O, Xia S. Tenascin-C Function in Glioma: Immunomodulation and Beyond. *Advances in experimental medicine and biology* **1272**, 149-172 (2020).
41. Braza MS, *et al.* Neutrophil derived CSF1 induces macrophage polarization and promotes transplantation tolerance. *Am J Transplant* **18**, 1247-1255 (2018).
42. Yuan Q, *et al.* MyD88 in myofibroblasts enhances colitis-associated tumorigenesis via promoting macrophage M2 polarization. *Cell reports* **34**, 108724 (2021).
43. Shani O, *et al.* Fibroblast-Derived IL33 Facilitates Breast Cancer Metastasis by Modifying the Immune Microenvironment and Driving Type 2 Immunity. *Cancer Res* **80**, 5317-5329 (2020).
44. Burr ML, *et al.* CMTM6 maintains the expression of PD-L1 and regulates anti-tumour immunity. *Nature* **549**, 101-105 (2017).

45. Calandra T, Bucala R. Macrophage Migration Inhibitory Factor (MIF): A Glucocorticoid Counter-Regulator within the Immune System. *Crit Rev Immunol* **37**, 359-370 (2017).
46. Soki FN, *et al.* Polarization of prostate cancer-associated macrophages is induced by milk fat globule-EGF factor 8 (MFG-E8)-mediated efferocytosis. *J Biol Chem* **289**, 24560-24572 (2014).
47. Laplante P, *et al.* MFG-E8 Reprogramming of Macrophages Promotes Wound Healing by Increased bFGF Production and Fibroblast Functions. *J Invest Dermatol* **137**, 2005-2013 (2017).
48. Ma H, *et al.* Periostin Promotes Colorectal Tumorigenesis through Integrin-FAK-Src Pathway-Mediated YAP/TAZ Activation. *Cell reports* **30**, 793-806 e796 (2020).
49. Elenbaas B, *et al.* Human breast cancer cells generated by oncogenic transformation of primary mammary epithelial cells. *Genes Dev* **15**, 50-65. (2001).
50. Wright KL, *et al.* Ras Signaling Is a Key Determinant for Metastatic Dissemination and Poor Survival of Luminal Breast Cancer Patients. *Cancer Res* **75**, 4960-4972 (2015).
51. Wagner KU. Know thy cells: commonly used triple-negative human breast cancer cell lines carry mutations in RAS and effectors. *Breast Cancer Res* **24**, 44 (2022).
52. Villadsen R, *et al.* Evidence for a stem cell hierarchy in the adult human breast. *J Cell Biol* **177**, 87-101 (2007).
53. Lin WH, *et al.* STAT3 phosphorylation at Ser727 and Tyr705 differentially regulates the EMT-MET switch and cancer metastasis. *Oncogene* **40**, 791-805 (2021).
54. Dietze EC, Sistrunk C, Miranda-Carboni G, O'Regan R, Seewaldt VL. Triple-negative breast cancer in African-American women: disparities versus biology. *Nat Rev Cancer* **15**, 248-254 (2015).
55. Lemos DR, Duffield JS. Tissue-resident mesenchymal stromal cells: Implications for tissue-specific antifibrotic therapies. *Science translational medicine* **10**, (2018).

56. Lee EK, Han GY, Park HW, Song YJ, Kim CW. Transgelin promotes migration and invasion of cancer stem cells. *J Proteome Res* **9**, 5108-5117 (2010).
57. 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).
58. Siersbaek R, *et al.* IL6/STAT3 Signaling Hijacks Estrogen Receptor alpha Enhancers to Drive Breast Cancer Metastasis. *Cancer Cell* **38**, 412-423 e419 (2020).
59. Knupfer H, Preiss R. Significance of interleukin-6 (IL-6) in breast cancer (review). *Breast Cancer Res Treat* **102**, 129-135 (2007).
60. Johnson DE, O'Keefe RA, Grandis JR. Targeting the IL-6/JAK/STAT3 signalling axis in cancer. *Nature reviews Clinical oncology* **15**, 234-248 (2018).
61. Ebbing EA, *et al.* Stromal-derived interleukin 6 drives epithelial-to-mesenchymal transition and therapy resistance in esophageal adenocarcinoma. *Proc Natl Acad Sci U S A* **116**, 2237-2242 (2019).
62. Choy EH, De Benedetti F, Takeuchi T, Hashizume M, John MR, Kishimoto T. Translating IL-6 biology into effective treatments. *Nat Rev Rheumatol* **16**, 335-345 (2020).
63. Thennavan A, *et al.* Molecular analysis of TCGA breast cancer histologic types. *Cell Genom* **1**, (2021).
64. Ajuwon AM, Love R. Type 2 diabetes and depression in the African American population. *J Am Assoc Nurse Pract* **32**, 120-127 (2020).