Understanding the Biology of Triple Negative Breast Cancer and Breast Cancer Stem Cells in Patients of Diverse Ethnicities.

by

Evelyn Mawunyo Jiagge

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Cancer Biology)
In The University of Michigan
2017

Doctoral Committee:

Professor Sofia D Merajver Co-Chair
Professor Max S Wicha, Co-Chair
Dr. Gary D Hammer
Dr. Elizabeth R. Lawlor
Dr. Jacques Eduardo Nor

Acknowledgement

My profound gratitude to my advisor, Dr. Sofia D Merajver for her guidance, time, ever-present support and mentorship.

My sincere gratitude to my thesis co-chair, Dr. Max S Wicha for his support and guidance.

I am very thankful to the members of my dissertation committee: Dr. Gary Hammer, Dr. Beth Lawlor and Dr. Jacques E Nor for their great support, time and for directing my work.

Thank you very much, Dr. Lisa A Newman and ICSBCS team for the enormous amount of resource and time you invested into this work.

Thank you to the many wonderful people in the Merajver and Wicha labs who contributed in diverse ways to my work. Special thanks to Dr. Sean McDermott for teaching me to be a scientist and Dr. Rabia Gilani for contributing significantly to my work.

Many thanks to our collaborators at the Komfo Anokye teaching hospital, Ghana who sacrificed their time and energy to prepare patients for this work.

Thank you to our collaborators at TGEN, especially Shukmei Wong for the many hours they invested into preparing and sequencing our sample.

Thank you to Dr. Jun Li and the Li lab for analyzing my data.

My sincere thanks to the University of Michigan flow Cytometry core for their time and patience in analyzing my many samples.

I am grateful to the ULAM and Cancer center pathology labs for the work you performed on my samples.

A special thanks to all the patients in Ghana and in the US who consented to their samples being used for this work

My sincere gratitude and love to my dear husband and the love of my life, Dr. Sylvester Antwi for sacrificing everything to make this work possible.

Thank you Nyameye and Judah for allowing me to do this work.

This work was supported in part by grants: Komen promise grant, the Breast Cancer Research Foundation (MW, SDM), the Metavivor Foundation (SDM), the UM CCC core grant and Rackham Barbour Scholarship.

Preface

A version of Chapter II, Breast Cancer and African Ancestry: Lessons Learned at the 10-Year Anniversary of the Ghana-Michigan Research Partnership and International Breast Registry was published is accepted in Journal of Global Oncology and online February 5, 2016. The list of authors is as follows: Evelyn Jiagge, Joseph Kwaku Oppong, Jessica Bensenhaver, Francis Aitpillah, Kofi Gyan, Ishmael Kyei, Ernest Osei-Bonsu, Ernest Adjei, Michael Ohene-Yeboah, Kathy Toy, Karen Eubanks Jackson, Marian Akpaloo, Dorcas Acheampong, Beatrice Antwi, Faustina Obeng Agyeman, Zainab Alhassan, Linda Ahenkorah Fondjo, Osei Owusu-Afriyie, Robert Newman Brewer, Amma Gyamfuah, Barbara Salem, Timothy Johnson, Max Wicha, Sofia Merajver, Celina Kleer, Judy Pang, Emmanuel Amankwaa-Frempong, Azadeh Stark, Francis Abantanga, Lisa Newman, Baffour Awuah

A portion of chapter III, Comparative Analysis of Breast Cancer Phenotypes in African American, White American, and West Versus East African patients: Correlation Between African Ancestry and Triple-Negative Breast Cancer, is accepted in Annals of Surgical Oncology and was published online November 23, 2016. The list of authors are as follows Evelyn Jiagge, Aisha Souleiman Jibril, Dhananjay Chitale, Jessica M. Bensenhaver, Baffour Awuah,

Mark Hoenerhoff , Ernest Adjei, Mahteme Bekele, Engida Abebe, S. David Nathanson, Kofi Gyan, Barbara Salem, Joseph Oppong, Francis Aitpillah, Ishmael Kyei, Ernest Osei Bonsu,

Erica Proctor , Sofia D. Merajver, Max Wicha, Azadeh Stark

A version of Chapter IV, ALDH expression characterizes tumor initiating stem cells in Triple Negative breast cancers across different ethnicities is being submitted to Cancer Research.

The author list is as follows: Evelyn Jiagge, Shukmei Wong , Rabia Gilani, Sean McDermott PhD, Qingxuan Song, Tahra Luther, Jungsoo Chang, Jessica Bensenhaver, Michele Dziubinski, Joseph Kwaku Oppong, Francis Aitpillah, Ishmael Kyei, Ernest Osei-Bonsu, Ernest Adjei, Awuah Baffour, Jessica Aldrich, Ahmet Kurdoglu, David Craig, Jeff Trent, Jun Li, Lisa Newman, Max Wicha, John Carpten, Sofia Merajver,

Table of Contents

List of Figures

List of Tables

Abstract

Triple negative breast cancer is one of the most aggressive breast cancer subtypes, for which there are no approved targeted therapies. In US alone, the incidence of TNBC is highest in women with African ancestry (AA); in western sub-Saharan Africa, single-institution studies show that TNBC constitutes 40-80% of all breast cancer cases. There is an urgent need to find actionable targets in TNBC of all ethnicities, but especially in patients with African ancestry, whose tumors are suspected to be more aggressive. Here, we sought to better understand the biology of TNBC by finding genes and pathways that are differentially expressed in the stem cell population of patient derived xenografts (PDX) created with TNBC tumors from Ghanaian (GH), AA and White American (WA) women and the effect of these differentially expressed genes on the stem cell phenotype in these primary tumors.

We first established an international, inter-institutional collaboration with a teaching hospital in Ghana (Komfo Anokye teaching hospital, KATH) for the acquisition of patients' samples. From these samples, we created patient derived xenograft models that serve as a renewable source of tumor tissue to study the cancer biology. Breast cancer stem cells, the small population of cells that have been shown to mediate breast tumor initiation, metastasis, and resistance to

conventional therapy have also been reported to mediate the heterogeneity of TNBC and are especially abundant in TNBC in AA women. We therefore sought to understand the biology of these stem cells in our primary patient samples and their associated PDX models.

We report here that, through our collaboration, we now have a better understanding of the diversity of the breast cancers that occur in Ghana and other parts of Africa, we have contributed to an improvement in the diagnosis and treatment of breast cancer patients and have contributed to enriching the human resource at the KATH in Ghana. We have successfully created a cohort of PDXs from TNBC patients in Ghana as well as African-American and White American patients. By sequencing the stem cells in these tumors, we have identified the aldehyde dehydrogenase (ALDH) expressing stem cell population to represent the stem cells in these aggressive tumors from this diverse population. We identified 14 genes that were simultaneously differentially expressed between the two-breast cancer stem cell sub-populations in (ALDH+ vs the CD44+/CD24-/EPCAM+) as well as ALDH+ vs bulk (p-value <0.001, FDR < 0.05 for both comparisons). The 3 most significant genes were matrix metalloproteinase 2 (MMP2) and Protocadherin 7 (PCDH7), both known to be involved in breast cancer metastasis and Probable carboxypeptidase X1 **(**CPXM1), a carboxylase. Inhibiting MMP2 expression in the PDX cells grown in suspension resulted in significant reduction in the ALDH+ cell population. Also, the ALDH+ and not the CD44+/CD24- cells formed spheres in serum free media. The WNT, MAPK and

TGF-beta pathways known to mediate metastasis were all significantly up-regulated in the ALDH+ population with down regulation of biosynthetic pathways which were up-regulated in the CD44+/CD24- population. Further studies are ongoing on pathway modulation in ALDH1+ cells based on these findings.

Together, these findings demonstrate the importance of international collaborations with mutual benefits. It also shows that TNBC, being a very heterogeneous disease, may be driven by a population of tumor cells that can be targeted to improve patient outcome.

Chapter 1 : Introduction

Breast Cancer Burden

Breast cancer (BC) is the most common cancer diagnosed in women worldwide with an estimated 1.7 million new cases diagnosed in women worldwide in 2012[3]. Globally, the burden of breast cancer exceeds all other cancers and the incidence is increasing. In 2015 there was an estimated 2.4 million incident cases making it the most common cancer. It is the fifth leading cause of cancer deaths for both men and women. For women, breast cancer was the leading cause of death in 2015.[4]

In the united states, approximately one in eight women (12%) will develop invasive breast cancer over the course of her lifetime. In 2016, it was estimated that, more than 3.5 million women living in the United States have a history of invasive breast cancer, an additional 246 660 new cases of invasive breast cancer are expected to be diagnosed and approximately 40450 would die as a result of it [5]. The median age at diagnosis of breast cancer is much younger than other common cancers. It is 61 years for breast cancer compared with 70 years for lung cancer and 68 years for colorectal cancer with About 19% of breast cancers diagnosed in women ages 30 to 49 years[5] [4]. There is therefore more

room for continuous research in breast cancer aiming at improving patient outcome.

Breast Cancer Incidence and Mortality

Globally, the incidence rates for breast cancer varies nearly four-fold across the world regions, with rates ranging from 27 per 100,000 in Middle Africa and Eastern Asia to 92 in Northern America. Higher breast cancer incidence in more developed countries is mostly attributed to the routine mammographic screening which helps in early detection of breast cancers [6]. Other non-hereditary factors more prevalent in western countries and are known to increase the incidence of breast cancer include weight gain after age 18 years, excess body weight (for postmenopausal breast cancer), use of menopausal hormone therapy, physical inactivity, alcohol consumption, and reproductive and hormonal factors, such as a long menstrual history, over use of oral contraceptives, and nulliparity or later age at first birth [7]. the incidence in the more developed parts of the world in now considered stable [7]. Breastfeeding may decrease the risk of some subtypes of breast cancer. In contrast, breast cancer incidence rates in many other countries, especially low to middle income countries (LMICs), continue to increase. The causes of these increases are not completely understood, but are thought to include changes in reproductive pattern as well as increased awareness and screening[6].

Although the incidence rate is higher in the western more developed countries (92 per 100000 in North America) compared to developing countries (27 per 100000 in middle Africa), the mortality rate is far higher in the developing world (20 in 100000 in western Africa) compared to the western world (6 per 100000 in Eastern Asia)[3], Figure 1.1. This global disparity in breast cancer outcome can mainly be attributed to factors such as lower life expectancy as well as low socioeconomic status in the developing world that leads to lack of screening modalities, delayed diagnosis and inadequate treatment among others. In a study by Ward et al, it was demonstrated that the poorest communities in any ethnic group had the worst prognosis in any cancer, stating the fact that poor socioeconomic status (SES) is an adverse prognostic factor in cancer outcome irrespective of Ethnicity[8].

Low SES however, is not necessarily the only determinant of BC ethnicity-related outcome disparity as it does not explain for biological features such as high prevalence of aggressive BC subtypes, younger age distribution and higher

incidence of male breast cancer in western Africa for example[1,9-11].

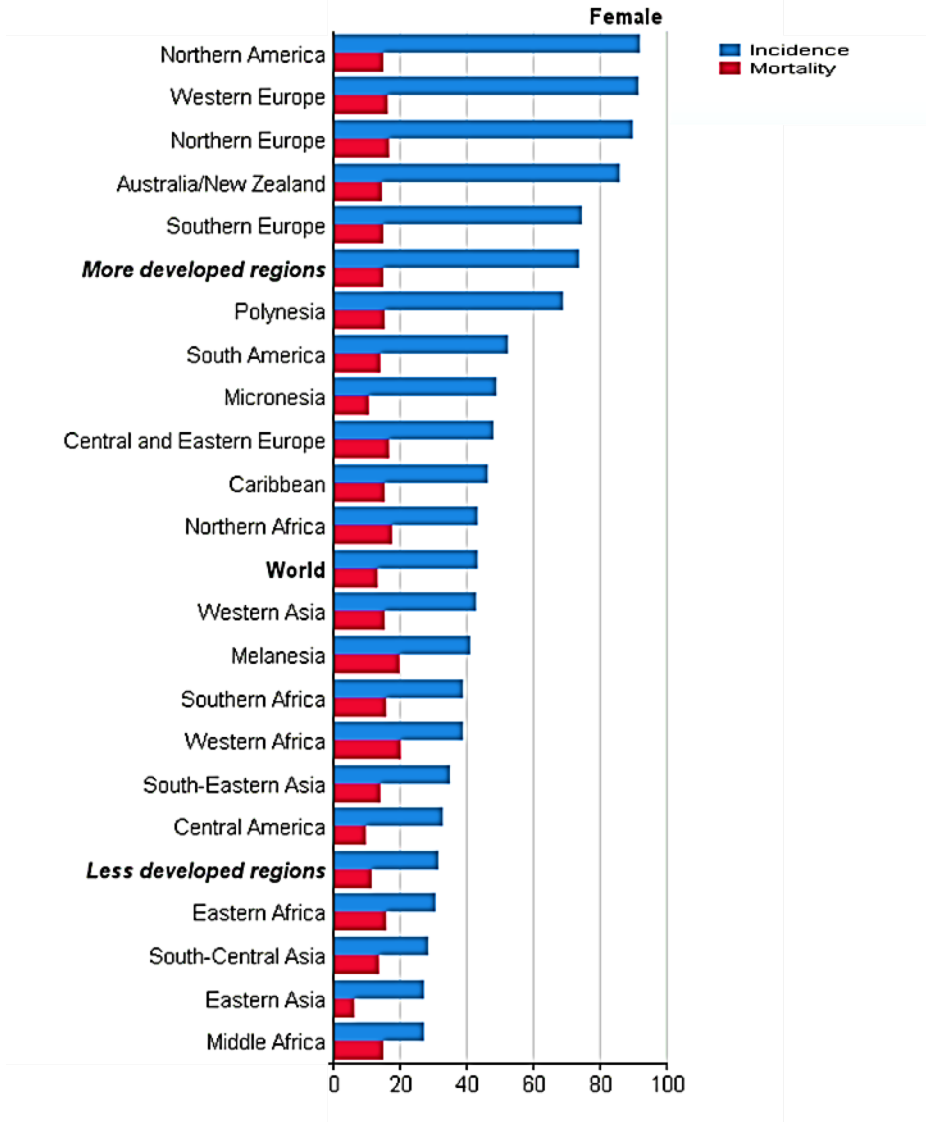


Figure 1-1: GLOBOAN 2012 global breast cancer incidence to mortality ratio

In a large (13,000 African American (AA) and 70,000 white Americans WA)) met-analysis conducted by Newman and coworkers in the USA reporting

on the survival of AA and WA breast cancer, they adjusted for SES, age and stage of diagnosis and showed that AA background was associated with a mortality hazard ratio of 1.27 (95% CI 1.18-1.38)[12,13]. This can be explained in part by the high incidence of an aggressive breast cancer subtype that lacks targeted therapy, in AA women as compared to WA women.

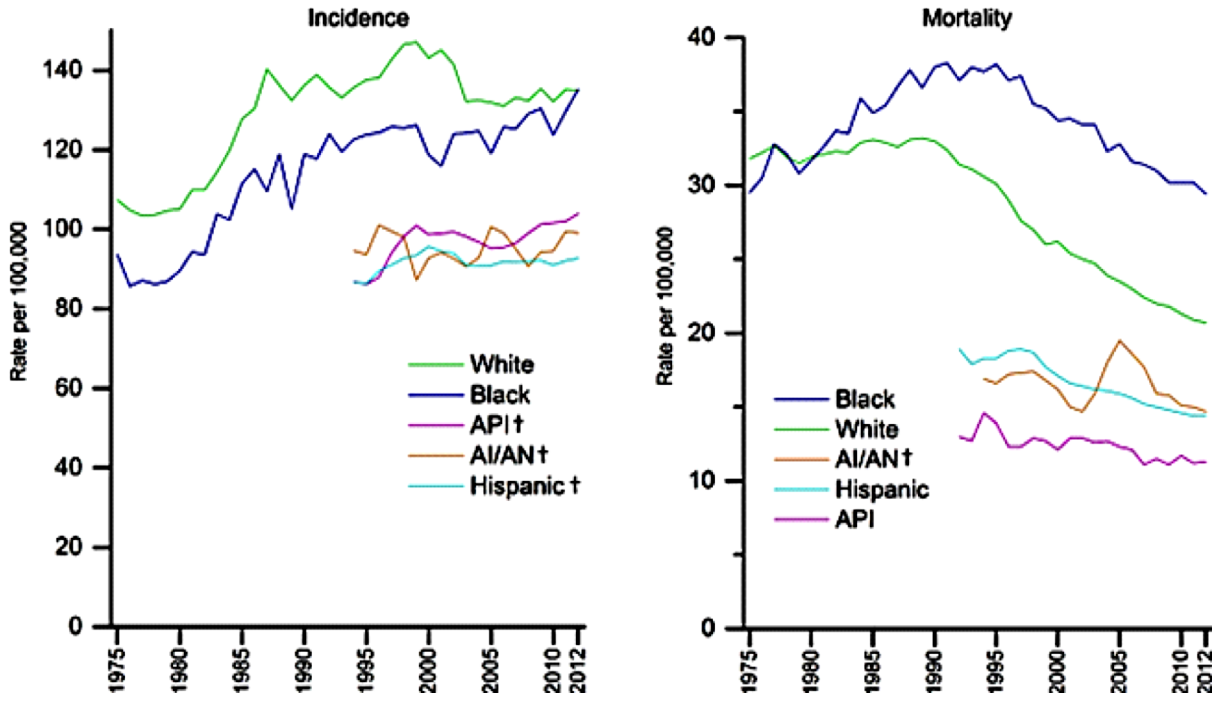


Figure 1-2 : Trends in Female Breast Cancer Incidence and Mortality Rates by Race/Ethnicity, United States, 1975 to 20121

Rates are per 100,000 females, age adjusted to the 2000 US standard population, and incidence rates are adjusted for reporting delay. †Rates are 3-year moving averages. NH indicates non-Hispanic; API, Asian/Pacific Islander; AI/AN, American Indian/Alaska Native. Sources: Incidence: Surveillance, Epidemiology, and End Results program, National Cancer Institute, 2015. Mortality: National Center for Health Statistics, Centers for Disease Control and Prevention, as provided by the Surveillance, Epidemiology, and End Results program, National Cancer Institute, 2015. For Hispanics, incidence data do not include cases from the Alaska Native Registry and mortality data exclude New Hampshire and Oklahoma. Data for American Indians/Alaska Natives are based on Contract Health Service Delivery Area (CHSDA) counties.[1]

In the United States, during 2008 through 2012 (the most recent 5 years of data available), overall breast cancer incidence rates increased among black women (0.4% per year) but did not change significantly among white women. Breast cancer incidence rates for whites and blacks converged in 2012, reflecting

the increased incidence in black women and relatively stable rates in white women, figure 1-2.

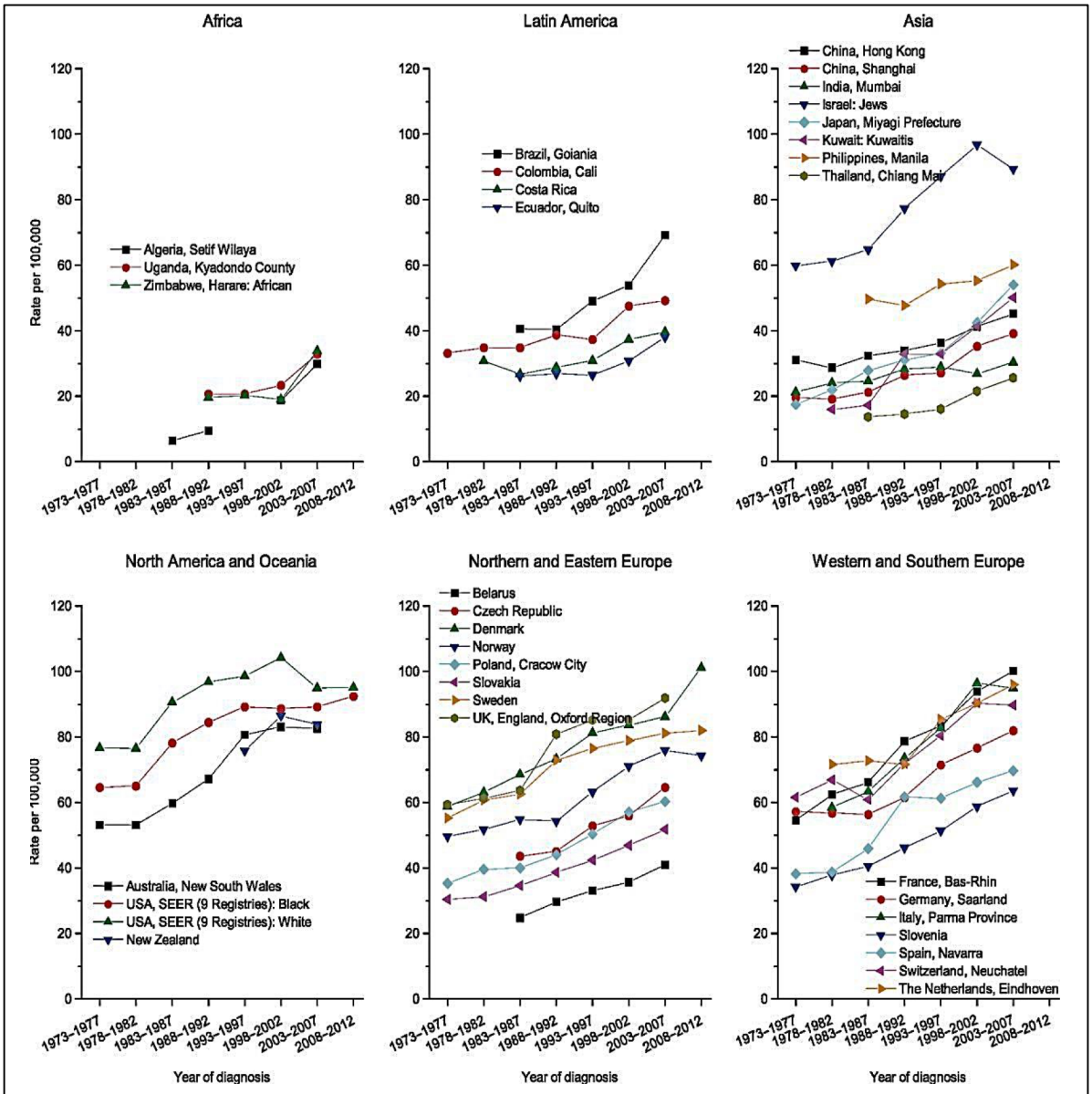


Figure 1-3: Female breast incidence trends, select countries, 1973–2012 (Compiled from: US, Surveillance, Epidemiology, and End Results [SEER] Program, 9 registries, 1973–2012; Denmark, Norway, Sweden, 2008–2012 NORDCAN; all others, Cancer Incidence in Five Continents, Volumes I–X.)[2]

A striking divergence in long-term breast cancer mortality trends between black and white women emerged in the early 1980s and has since continued to widen, (figure 1-2); and, by 2012, mortality rates were 42% higher in black women than in white women. Trends in breast cancer mortality rates by race/ethnicity are also shown in figure 1-3. From 2003 through 2012, breast cancer mortality rates declined annually by 1.8% in whites and only 1.4% in blacks[1]. Several studies indicate that the difference in breast cancer mortality are likely due to a combination of factors, including differences in incidence rates (at least during the recent time period), stage at diagnosis, obesity, comorbidities access, adherence, and response to state-of-the art treatments and a difference in tumor biology[1,14,15]. Treatment of breast cancer is stratified based on the sub-type of the tumor.

Breast Cancer Sub-types

Breast cancer is a heterogeneous disease with a high degree of diversity within a tumor. The heterogeneity of breast cancer makes it a challenging solid tumor to diagnose and treat. Historically, many factors have been investigated as a means to stratify patients, to prognosticate and to decide on treatment options. These include age, parity and family history among others. However, tumor receptor status has provided the most useful information in predicting prognosis and responsiveness to treatment. Using Immunohistochemistry, the expression of estrogen receptor (ER), progesterone receptor (PR), and overexpression of

human epidermal growth factor receptor 2 (HER2/neu) is measured[16-18]. Breast cancers are classified with respect to the presence or absence of these receptors. Patients whose breast tumors express estrogen receptor are classified as ER-positive (ER+) and breast tumor which do not express ER are classified as ER-negative disease with approximately 70% of breast cancer patients having ER+ disease. When ER is present, it drives cancer growth and is a therapeutic target and these patients are treated with endocrine therapy[19] which aims at either reducing the estrogen receptor activity or reduce receptor levels within breast cancer cells. Overexpression of the Human epidermal growth factor receptor 2 (HER2/neu) classifies breast cancer as HER2 positive breast cancer. Persistent activation of signaling pathways as a result of amplification of the human epidermal growth factor receptor 2 (HER2) in patients with HER2-positive breast cancer leads to a biologically aggressive malignancy with heightened sensitivity to cytotoxic chemotherapy[20]. In these patients, the addition of HER2 targeted therapy (trastuzumab and pertuzumab) both in the neoadjuvant and adjuvant setting, significantly improves the patient outcome[21]. TNBCs, classified based on the absence and not the presence of receptors, thus cannot be treated with any of these targed therapies. Amongst individual breast cancer subtypes, those classified as triple negative breast cancers (TNBCs) are especially lethal due to their high metastatic potential and high propensity to recur. As a group, TNBCs lack expression of hormone receptors (ER-✔ and PR) and lack overexpression of human epidermal growth factor receptor 2 (ErbB2/HER2).

Thus, therapies directed against these robust targets are not effective against TNBC. TNBC is the only subset of breast cancers for which there are no FDA-approved targeted therapies and clinicians are entirely reliant upon cytotoxic agents for tumor control. Clearly, there is a significant need for targeted therapeutics against TNBC.

Triple Negative Breast Cancer

Triple negative breast cancer (TNBC) is the subtype of breast cancer with the worst prognosis and is defined by having less than one per cent expression of the estrogen and progesterone receptor,[17] and no overexpression of the HER2 oncogene; either 0, 1+ by immunohistochemistry or 2+ with fluorescence in situ hybridization negative[18,22].

Triple-negative breast cancer accounts for approximately 20 percent of breast cancers diagnosed worldwide, which amounts to almost 200,000 cases each year[23]. Triple-negative breast cancer is more commonly diagnosed in women younger than 40 years compared with hormone-positive breast cancer. In one study, there was a twofold higher attributable risk of triple-negative breast cancer in women under 40 years compared with women over 50 years (odds ratio [OR] 2.13, 95% CI 1.34-3.39) [23]

Using gene expression profiling, about 80% of TNBC cases belong to the basal BC subtype which is biologically very aggressive[24,25]. By definition, patients with TNBC cannot be treated with endocrine therapy, nor with anti-HER2 proto-

oncogene targets since the tumor is defined by the absence and not the presence of these targets. TNBC has a 5-year overall survival of 64% vs. 81% in non-TNBC, it has the highest rate of visceral and distant relapse, and the shortest survival after the development of metastasis[26-28]. As there are no approved targeted therapies for TNBC, chemotherapy is the mainstay of systemic treatment in either the neoadjuvant setting or after loco-regional therapy with surgery and radiotherapy [22,29]. Patients with TNBC have a higher rate of complete pathological response after neoadjuvant chemotherapy (22% v 11%; P = .034) as compared to non-TNBC, which correlates with a better prognosis in TNBC[27,30]. Eight out of ten of TNBC patients have residual disease and they are at an increased risk of relapse and metastasis[26,28] with a higher rate of brain metastasis as compared to non-TNBC[25]. The unique molecular complexity and extremely heterogeneity of the TNBC has attracted a lot of research efforts to define molecular targets. Several attempts have been made to classify TNBC into subtypes, with the aim of describing prognostic categories[31-34]; however, these sub-classifications are yet to translate into direct clinical benefit.

In the United states, 15% of all breast cancers are TNBC, in contrast, the Carolina BC study reported that 39% [35]of all invasive breast tumors diagnosed in premenopausal AA women were TNBC, highlighting the potential the potential role of TNBC in BC cancer outcome disparity. Several other studies have reported that the population based incidence of TNBC in AA in twice that of WA in all age categories. This includes the 2015 “Annual Report to the Nation on the

Status of Cancer" which cited a 27.2 vs 14.4 per 100 000 women respectively [36-38]. The TNBC phenotype is associated with an increased in hereditary susceptibility to BC irrespective of family history[39]. The AA population of the USA constitutes a genetic admixture of several ancestral populations with their African ancestry originating largely from Western Africa through forced migration from the trans-Atlantic slave trade[40]. It is therefore worthy of note that the incidence of TNBC in west African women has been reported to be as high as 80% of all breast cancer diagnoses in several institutional based studies. While African American women have a higher incidence of TNBC than European American women in the USA, West African women with pure African ancestry have a much higher incidence. The average incidence is 10 - 21%, 21 - 46%, and up to 82% for WA, AA and west Africans respectively[35,41-45]. The question begging to be asked therefore is, is African ancestry associated with a hereditary susceptibility to TNBC and can we better understand the biology of TNBC if we study the genetic makeup of these tumors in women with African ancestry. Considering the fact that the African continent is the oldest continent and being the origin of all other ethnicities, we might find gene and pathways signaling in the African tumors.

Hereditary Breast Cancer and Breast Cancer in Blacks

Approximately 10% of all breast cancers are hereditary[46]. Hereditary breast cancers are autosomal dominant in transmission and occur at an earlier age.

Defects in the *BRCA1* and *BRCA2* genes accounts for most cases of hereditary breast cancer with several additional mutations now being associated with an increased risk of developing breast cancer[46]. The BRCA1 and BRCA2 genes, are tumor suppressor genes that were discovered in the mid-1990s[47,48]. A woman who carries a BRCA1 mutation has a 57% lifetime risk of developing breast cancer, whereas a woman with a BRCA2 mutation has a 49% risk[49,50]. Once a woman is diagnosed with breast cancer, if she is a carrier of a BRCA1 or BRCA2 mutation, she has an increased risk of developing a second breast cancer in the contralateral breast.

Up to 20 percent of patients with triple-negative breast cancer harbor a breast cancer gene (BRCA1/2) mutation (Even among patients unselected for family history of breast cancer), particularly in *BRCA1*[51]. By contrast, less than 6 percent of all breast cancers are associated with a BRCA mutation. Population-based studies have reported that Black women are approximately two to three times more likely to have triple-negative breast cancer (TNBC) than White women[37,52]. There is limited data on the prevalence of deleterious *BRCA1* and *BRCA2 (BRCA1/2)* in blacks with most of the available data arising from population-based studies or from cohorts comprising highly selected families[53]. In individuals referred for genetic testing, the prevalence of *BRCA1/2* mutation in women of African ancestry ranged from 12.1 to 15.6 % and were significantly higher than in women of Western European ancestry (ranged from 7.5 to 12.1 %). In western sub-Saharan Africa, where there is the highest reported

prevalence of TNBC, almost, nothing is known of the prevalence of BRCA gene mutations in TNBC patients. Given the high rate of early onset breast cancer and TNBC in Black women, it is important to determine the prevalence of *BRCA1/2* mutations in this population[54]

Breast Cancer Treatment

When breast cancer is detected early without detectable distant metastases is a considered a potentially curable disease. After diagnosis, therapy concepts are usually decided by a multidisciplinary team. The therapeutic option for each patient is determined based on the results of clinical examination and breast imaging (mammography, breast ultrasound), after a tissue diagnosis of malignancy has been established. In patients with early-stage-breast cancer, surgical options include breast-conserving therapy (breast-conserving surgery plus radiation therapy [RT]) or mastectomy (with or without RT). In patients with locally advanced inoperable breast cancer, neoadjuvant systemic therapy (chemotherapy usually) is instituted. These patients are usually not candidates for breast conservation at their initial presentation. Neoadjuvant treatment improves the rate of breast conservation without compromising survival outcomes. Following surgery, patients may receive adjuvant radiation therapy to maximize loco regional control[55]. Patients with HER2-positive breast cancer receive one year of trastuzumab following completion of surgery[55-59]. Patients with hormone receptor-positive breast cancer receive adjuvant endocrine

therapy. Patients with hormone receptor-negative breast cancer receive no further treatment provided they completed the planned neoadjuvant chemotherapy regimen. These patients then begin post-treatment surveillance. By exploring new approaches and finding vulnerabilities in TNBC, it might be possible to find targeted therapies for these patients in the adjuvant setting. One such approach is studying the cancer stem cells that have been shown to mediate tumor metastasis.

Cancer Stem Cells

Recent data suggest that cancer cells are heterogeneous and that only a small and discrete subset possesses self-renewal capacity. This particular subpopulation is thought to be responsible for sustenance of tumor growth and recurrence, termed "cancer stem cells" (CSCs)[60-63]. The cancer stem cell model proposes that the cells in a tumor are organized in a hierarchical manner with the stem cell sitting at the top of the hierarchy. The stem cell has the potential of self-renewal as well as the ability to differentiate into all the other cells that makes up the bulk tumor[64,65]. For successful control of tumor growth, it has been proposed that these cancer stem cells must be eradicated. Dick and colleagues pioneered this field in 1997 and demonstrated that only cells with a CD34+/CD38- phenotype were able to initiate human acute myelogenous leukemia (AML) in non-obese diabetic/severe combined immunodeficient (NOD/SCID) mice[60]. These cells were capable of self-renewal and recapitulation of the leukemia

similar to the parental tumor phenotype in vivo. Furthermore, leukemic stem cells are not functionally homogeneous but consist of a distinct hierarchically arranged class similar to normal hematopoietic stem cells[60]. In breast cancer, Al-Hajj and colleagues were the first who identified breast CSC phenotype. They demonstrated that as few as 100 cells with a CD44+CD24-/low lineage- phenotype isolated from human breast tumors and malignant pleural effusion can initiate breast cancer in NOD/SCID mice reflecting the self-renewal capability of this phenotype. In contrast, injection of tens of thousands of cells with different phenotypes failed to form any tumors[61]. Further study showed that although the CD44+CD24-/low lineage- population accounts for only less than 10% in the primary tumors, the majority (71%) of the disseminated breast cancer cells detected in bone marrow exhibit this breast CSC phenotype, highlighting its significant role in future relapse and distant metastasis[61]. Moreover, there is a direct correlation between the prevalence of this phenotype in the primary tumor and the risk for distant metastasis[66-68]. Several groups demonstrated that self-renewal pathways as well as multiple “stemness” genes are highly expressed in CD44+CD24-/low breast cancer cells. These include hedgehog signaling components such as PTCH1, Gli1, Gli2, and the polycomb gene Bmi-1[69,70].

Aldehyde dehydrogenase (ALDH) is a family of enzymes important in the oxidation of alcohol and vitamin A as well as in resistance to cyclophosphamide[71]. The human aldehyde dehydrogenase (ALDH) superfamily currently consists of 19 known putatively functional genes in 11 families and 4

subfamilies with distinct chromosomal locations[72-74]. Of the vast ALDH families and subfamilies, it has been shown that the ALDH enzymes that are involved in normal stem cells as well as cancer stem cells (CSCs) include the ALDH1, ALDH3A1, ALDH4A1, and ALDH7A1[72]. Cytosolic ALDH1 is highly expressed in human and murine hematopoietic stem and progenitor cells[75].[76] More recently, Ginestier and colleagues[77] have demonstrated that both normal and breast cancer cells with higher ALDH1 expression have stem/progenitor properties. These cells with a higher ALDH expression level have the highest tumorigenicity in xenotransplantation and the most extensive lineage differentiation potential. High ALDH activity in breast cancer also distinguishes the tumorigenic cell fraction that is capable of self-renew and recapitulating the heterogeneity of the original tumor. Furthermore, expression of ALDH1 detecting by immunohistochemically (IHC) staining is found to be an independent predictive marker of poor prognosis[78].

Breast Cancer Stem Cells

There is now an abundance of evidence supporting the existence of a small sub-population of breast tumor cells with self-renewing potential, termed breast cancer stem cells (BCSC) or tumor initiating cells, that plays a major role in tumor recurrence and metastasis[79,80]. The existence of stem cells has been demonstrated in different tumors including leukemia[60], breast[61], colon[81] and prostate[82] among others. In these various studies, the stem cells were enriched

for by the expression of protein markers or the presence of the enzyme aldehyde dehydrogenase (ALDH) in the cell. In breast cancer, these stem cells can be isolated from the tumor mass using fluorescent activated cell sorting (FACS) when they are CD44+/CD24-/EPCAM+[61,83]. Furthermore, using the ALDEFLUOR assay, stem-like cells can also be isolated by measuring the activity of the intracellular aldehyde dehydrogenase enzyme in intact breast cancer cells[77] These cells have also been shown to be especially abundant in TNBC[77 84,85], are responsible for resistance to chemotherapy and radiotherapy,[86-88] and contribute to the heterogeneity observed in TNBC [89]. ALDH+ BCSC have been shown to mediate metastasis and poor clinical outcome in breast cancer[90]. While the existence of phenotypic BCSC is supported by a large body of preclinical and clinical evidence, the pathways that contribute to the maintenance of this cell population are not clearly defined[91]

Breast Cancer Ctem cells and Breast Cancer in Blacks

The expression of ALDH1, the protein associated with the ALDEFLUOR activity, has been associated with TNBC and aggressive breast cancers in African women[85]. In our previous work, we showed that, a significantly high expression of ALDH1 in breast lesions (benign and malignant) of Ghanaian women. We also showed a higher expression of ALDH1 in TNBC as compared to other breast cancer subtypes[84]in these women. We have therefore hypothesized that the key to unraveling the complexity of TNBC and thus finding new targets

for its treatment lies in understanding signaling integrations and gene expression in the BCSCs of TNBC in women with African ancestry in comparison with other ethnicities.

Implications of CSCs for Cancer Therapy

Although most cancer patients may respond to therapy, an initial response is rarely translated into long-term survival. This is comparable to cutting a dandelion off at ground level. Although this will remove the visible portion of the weed, the hidden root remains underground and will eventually regrow. In order to entirely eradicate the weed, the unseen root also needs to be eliminated to prevent regrowth ("the dandelion phenomenon")[92]. Most conventional chemotherapies mainly affect differentiated cancer cells that make up the bulk of a tumor, but are often ineffective against CSCs[92]. There are several mechanisms for chemotherapy resistance in CSCs. Similar to normal stem cells, CSCs are usually in quiescent G0 phase and enter cell cycling only when they receive proper signals from their niches. Therefore, CSCs are insensitive to S-phase specific chemotherapies.

CSCs have a particularly efficient drug efflux pump systems compared to non-CSCs. Hirschmann-Jax et al[93]. Identified ABC transporters as the means by which side population (SP) have an increased capability to efflux chemotoxic drugs in various tumor cells. SP have been shown to overlap in breast cancer with $CD44^{+}/CD24^{-}$CSCs[94] in breast cancer. CSCs are hypothesized to have

higher radio resistance, by a superior DNA damage repair than non-CSCs akin to normal stem cells[95]. In a recent study by Bao et al[96]. Radiotherapy was shown to enrich $CD133^+$ glioma CSCs two- to four-fold, and these enriched cells had lower rates of apoptosis and higher activation of DNA damage repair (i.e. ATM, Rad17, Chk1, and Chk2). Cancer stem cells have also been shown to have lower levels of reactive oxygen species (ROS) than the non-cancer stem cells[96].

Moreover, CSCs express high level of anti-apoptotic proteins, particularly members of the Bcl-2 family[97], and drug transporters, such as ABCG2 (BCRP) and P-glycoprotein. [98,99] The current challenge in anti-cancer drug development is therefore to find strategies to selectively target and eradicate CSCs. Perhaps, CSC directed therapies should be administered in conjunction with conventional debulking methods like surgery, chemotherapy, and radiation. It is also important to find the right model for studying breast cancer stem cells as it represents a sub-population of the cancer cells.

Models for Breast Cancer Research

The most widely used model for pre-clinical breast cancer research *in vitro* culture of established breast cancer cell lines. The use of these cell lines has many technical advantages; the complete of environmental conditions and standardized culture conditions ensuring the reproducibility of results between experiments and laboratories. It is cheaper than animal experiments and the use of these cell lines are almost infinite. It is however very limited. Although the cell

is cryopreserved, they undergo dedifferentiation resulting from multiple subcultures, leading to loss of special characteristics. The simplicity of cell cultures implies difficulties in extrapolating results to human tumors. They turn be less heterogeneous than human tumors and they lack stroma[100]. By contrast, Patient derived xenograft (PDX) tumors that are created when cancerous tissue from a patient's primary tumor is implanted directly into an immunodeficient mouse have been shown to conserve the gene expression, histology, and architecture and heterogeneity of the original primary tumor from which they are developed. They turn to be more expensive and labor intensive but are more similar to the human tumor[101-103].

Most of the available data on the biology of TNBC was generated using cell lines developed mainly from white patient tumors, which do not serve as a true representation of the heterogeneous nature of TNBC. They also utilize only the bulk tumor cells instead of the BCSCs, which are responsible for the increased rate of metastasis. In this study, we used a unique approach to finding novel targets in the BCSC of TNBC using PDX developed from women with different ancestries[104] [105].

References

1. DeSantis CE, Fedewa SA, Goding Sauer A, Kramer JL, Smith RA, Jemal A. Breast cancer statistics, 2015: Convergence of incidence rates between black and white women. *CA Cancer J Clin.* Jan-Feb 2016;66(1):31-42.

2. Torre LA, Siegel RL, Ward EM, Jemal A. Global Cancer Incidence and Mortality Rates and Trends--An Update. *Cancer Epidemiol Biomarkers Prev.* Jan 2016;25(1):16-27.

3. GLOBOCAN. International Agency for Research on Cancer (IARC) and World Health Organization (WHO) GLOBOCAN 2012: Estimated cancer incidence, mortality and prevalence worldwide in 2012. 2012.

4. Global Burden of Disease Cancer C, Fitzmaurice C, Allen C, et al. Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability-Adjusted Life-years for 32 Cancer Groups, 1990 to 2015: A Systematic Analysis for the Global Burden of Disease Study. *JAMA Oncol.* Dec 03 2016.

5. Miller KD, Siegel RL, Lin CC, et al. Cancer treatment and survivorship statistics, 2016. *CA Cancer J Clin.* Jul 2016;66(4):271-289.

6. Althuis MD, Dozier JM, Anderson WF, Devesa SS, Brinton LA. Global trends in breast cancer incidence and mortality 1973-1997. *Int J Epidemiol.* Apr 2005;34(2):405-412.

7. Chlebowski RT, Manson JE, Anderson GL, et al. Estrogen plus progestin and breast cancer incidence and mortality in the Women's Health Initiative Observational Study. *J Natl Cancer Inst.* Apr 17 2013;105(8):526-535.

8. Ward E, Jemal A, Cokkinides V, et al. Cancer disparities by race/ethnicity and socioeconomic status. *CA Cancer J Clin.* Mar-Apr 2004;54(2):78-93.

9. Anderson WF, Rosenberg PS, Menashe I, Mitani A, Pfeiffer RM. Age-related crossover in breast cancer incidence rates between black and white ethnic groups. *J Natl Cancer Inst.* Dec 17 2008;100(24):1804-1814.

10. Nahleh ZA, Srikantiah R, Safa M, Jazieh AR, Muhleman A, Komrokji R. Male breast cancer in the veterans affairs population: a comparative analysis. *Cancer.* Apr 15 2007;109(8):1471-1477.

11. Anderson WF, Althuis MD, Brinton LA, Devesa SS. Is male breast cancer

similar or different than female breast cancer? *Breast cancer research and treatment.* Jan 2004;83(1):77-86.

12. Newman LA, Martin IK. Disparities in breast cancer. *Curr Probl Cancer.* May-Jun 2007;31(3):134-156.

13. 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.* Mar 20 2006;24(9):1342-1349.

14. Ooi SL, Martinez ME, Li CI. Disparities in breast cancer characteristics and outcomes by race/ethnicity. *Breast cancer research and treatment.* Jun 2011;127(3):729-738.

15. Curtis E, Quale C, Haggstrom D, Smith-Bindman R. Racial and ethnic differences in breast cancer survival: how much is explained by screening, tumor severity, biology, treatment, comorbidities, and demographics? *Cancer.* Jan 01 2008;112(1):171-180.

16. Polyak K. Heterogeneity in breast cancer. *J Clin Invest.* Oct 2011;121(10):3786-3788.

17. Hammond ME, Hayes DF, Dowsett M, et al. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer (unabridged version). *Arch Pathol Lab Med.* Jul 2010;134(7):e48-72.

18. Wolff AC, Hammond ME, Hicks DG, et al. Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. *J Clin Oncol.* Nov 1 2013;31(31):3997-4013.

19. Colleoni M, Montagna E. Neoadjuvant therapy for ER-positive breast cancers. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO.* Sep 2012;23 Suppl 10:x243-248.

20. Cortazar P, Zhang L, Untch M, et al. Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis. *Lancet.* Jul 12 2014;384(9938):164-172.

21. Gianni L, Eiermann W, Semiglazov V, et al. Neoadjuvant chemotherapy

with trastuzumab followed by adjuvant trastuzumab versus neoadjuvant chemotherapy alone, in patients with HER2-positive locally advanced breast cancer (the NOAH trial): a randomised controlled superiority trial with a parallel HER2-negative cohort. *Lancet.* Jan 30 2010;375(9712):377-384.

22. Malorni L, Shetty PB, De Angelis C, et al. Clinical and biologic features of triple-negative breast cancers in a large cohort of patients with long-term follow-up. *Breast cancer research and treatment.* Dec 2012;136(3):795-804.

23. Trivers KF, Lund MJ, Porter PL, et al. The epidemiology of triple-negative breast cancer, including race. *Cancer Causes Control.* Sep 2009;20(7):1071-1082.

24. Anders CK, Carey LA. Biology, metastatic patterns, and treatment of patients with triple-negative breast cancer. *Clin Breast Cancer.* Jun 2009;9 Suppl 2:S73-81.

25. Carey L, Winer E, Viale G, Cameron D, Gianni L. Triple-negative breast cancer: disease entity or title of convenience? *Nat Rev Clin Oncol.* Dec 2010;7(12):683-692.

26. Dent R, Trudeau M, Pritchard KI, et al. Triple-negative breast cancer: clinical features and patterns of recurrence. *Clin Cancer Res.* Aug 1 2007;13(15 Pt 1):4429-4434.

27. Liedtke C, Mazouni C, Hess KR, et al. Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. *J Clin Oncol.* Mar 10 2008;26(8):1275-1281.

28. Dent R, Hanna WM, Trudeau M, Rawlinson E, Sun P, Narod SA. Pattern of metastatic spread in triple-negative breast cancer. *Breast cancer research and treatment.* May 2009;115(2):423-428.

29. National Comprehensive Cancer Network. National Comprehensive Cancer Network Breast Cancer. 2015.

30. Carey LA, Dees EC, Sawyer L, et al. The triple negative paradox: primary tumor chemosensitivity of breast cancer subtypes. *Clin Cancer Res.* Apr 15 2007;13(8):2329-2334.

31. Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast

tumours. *Nature.* Aug 17 2000;406(6797):747-752.

32. Lehmann BD, Bauer JA, Chen X, et al. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J Clin Invest.* Jul 2011;121(7):2750-2767.

33. Prat A, Perou CM. Deconstructing the molecular portraits of breast cancer. *Mol Oncol.* Feb 2011;5(1):5-23.

34. Shah SP, Roth A, Goya R, et al. The clonal and mutational evolution spectrum of primary triple-negative breast cancers. *Nature.* Jun 21 2012;486(7403):395-399.

35. Carey LA, Perou CM, Livasy CA, et al. Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. *JAMA.* Jun 7 2006;295(21):2492-2502.

36. Amirikia KC, Mills P, Bush J, Newman LA. Higher population-based incidence rates of triple-negative breast cancer among young African-American women : Implications for breast cancer screening recommendations. *Cancer.* Jun 15 2011;117(12):2747-2753.

37. Bauer KR, Brown M, Cress RD, Parise CA, Caggiano V. Descriptive analysis of estrogen receptor (ER)-negative, progesterone receptor (PR)-negative, and HER2-negative invasive breast cancer, the so-called triple-negative phenotype: a population-based study from the California cancer Registry. *Cancer.* May 01 2007;109(9):1721-1728.

38. Kohler BA, Sherman RL, Howlader N, et al. Annual Report to the Nation on the Status of Cancer, 1975-2011, Featuring Incidence of Breast Cancer Subtypes by Race/Ethnicity, Poverty, and State. *J Natl Cancer Inst.* Jun 2015;107(6):djv048.

39. Newman LA, Reis-Filho JS, Morrow M, Carey LA, King TA. The 2014 Society of Surgical Oncology Susan G. Komen for the Cure Symposium: triple-negative breast cancer. *Annals of surgical oncology.* Mar 2015;22(3):874-882.

40. Newman LA. Health Disparities, African Americans, and Triple Negative Breast Cancer:
Oncologic Anthropology representing the intersection between the epidemiology of cancer, population genetics and race as a sociopolitical construct. *JAMA.* 2017.

41. Stark A, Kleer CG, Martin I, et al. African ancestry and higher prevalence of triple-negative breast cancer: findings from an international study. *Cancer.* Nov 1 2010;116(21):4926-4932.

42. Morris GJ, Naidu S, Topham AK, et al. Differences in breast carcinoma characteristics in newly diagnosed African-American and Caucasian patients: a single-institution compilation compared with the National Cancer Institute's Surveillance, Epidemiology, and End Results database. *Cancer.* Aug 15 2007;110(4):876-884.

43. Stead LA, Lash TL, Sobieraj JE, et al. Triple-negative breast cancers are increased in black women regardless of age or body mass index. *Breast cancer research : BCR.* 2009;11(2):R18.

44. Lund MJ, Trivers KF, Porter PL, et al. Race and triple negative threats to breast cancer survival: a population-based study in Atlanta, GA. *Breast cancer research and treatment.* Jan 2009;113(2):357-370.

45. McCormack VA, Joffe M, van den Berg E, et al. Breast cancer receptor status and stage at diagnosis in over 1,200 consecutive public hospital patients in Soweto, South Africa: a case series. *Breast cancer research : BCR.* 2013;15(5):R84.

46. Peng S, Lu B, Ruan W, Zhu Y, Sheng H, Lai M. Genetic polymorphisms and breast cancer risk: evidence from meta-analyses, pooled analyses, and genome-wide association studies. *Breast cancer research and treatment.* Jun 2011;127(2):309-324.

47. Hall JM, Lee MK, Newman B, et al. Linkage of early-onset familial breast cancer to chromosome 17q21. *Science.* Dec 21 1990;250(4988):1684-1689.

48. Wooster R, Neuhausen SL, Mangion J, et al. Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13. *Science.* Sep 30 1994;265(5181):2088-2090.

49. Chen S, Parmigiani G. Meta-analysis of BRCA1 and BRCA2 penetrance. *J Clin Oncol.* Apr 10 2007;25(11):1329-1333.

50. Stuckey AR, Onstad MA. Hereditary breast cancer: an update on risk assessment and genetic testing in 2015. *American journal of obstetrics and gynecology.* Aug 2015;213(2):161-165.

51. Gonzalez-Angulo AM, Timms KM, Liu S, et al. Incidence and outcome of BRCA mutations in unselected patients with triple receptor-negative breast cancer. *Clin Cancer Res.* Mar 01 2011;17(5):1082-1089.

52. Kwan ML, Kushi LH, Weltzien E, et al. Epidemiology of breast cancer subtypes in two prospective cohort studies of breast cancer survivors. *Breast cancer research : BCR.* 2009;11(3):R31.

53. Pal T, Bonner D, Kim J, et al. Early onset breast cancer in a registry-based sample of African-american women: BRCA mutation prevalence, and other personal and system-level clinical characteristics. *The breast journal.* Mar-Apr 2013;19(2):189-192.

54. Lynce F, Smith KL, Stein J, et al. Deleterious BRCA1/2 mutations in an urban population of Black women. *Breast cancer research and treatment.* Aug 2015;153(1):201-209.

55. Ebctcg, McGale P, Taylor C, et al. Effect of radiotherapy after mastectomy and axillary surgery on 10-year recurrence and 20-year breast cancer mortality: meta-analysis of individual patient data for 8135 women in 22 randomised trials. *Lancet.* Jun 21 2014;383(9935):2127-2135.

56. Veronesi U, Cascinelli N, Mariani L, et al. Twenty-year follow-up of a randomized study comparing breast-conserving surgery with radical mastectomy for early breast cancer. *N Engl J Med.* Oct 17 2002;347(16):1227-1232.

57. Colleoni M, Sun Z, Price KN, et al. Annual Hazard Rates of Recurrence for Breast Cancer During 24 Years of Follow-Up: Results From the International Breast Cancer Study Group Trials I to V. *J Clin Oncol.* Mar 20 2016;34(9):927-935.

58. Brewster AM, Hortobagyi GN, Broglio KR, et al. Residual risk of breast cancer recurrence 5 years after adjuvant therapy. *J Natl Cancer Inst.* Aug 20 2008;100(16):1179-1183.

59. van Maaren MC, de Munck L, de Bock GH, et al. 10 year survival after breast-conserving surgery plus radiotherapy compared with mastectomy in early breast cancer in the Netherlands: a population-based study. *Lancet Oncol.* Aug 2016;17(8):1158-1170.

60. Bonnet D, Dick JE. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat Med.* Jul

1997;3(7):730-737.

61. 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.* Apr 1 2003;100(7):3983-3988.

62. Piccirillo SG, Reynolds BA, Zanetti N, et al. Bone morphogenetic proteins inhibit the tumorigenic potential of human brain tumour-initiating cells. *Nature.* Dec 07 2006;444(7120):761-765.

63. Kim CF, Jackson EL, Woolfenden AE, et al. Identification of bronchioalveolar stem cells in normal lung and lung cancer. *Cell.* Jun 17 2005;121(6):823-835.

64. Wicha MS, Liu S, Dontu G. Cancer stem cells: an old idea--a paradigm shift. *Cancer research.* Feb 15 2006;66(4):1883-1890; discussion 1895-1886.

65. Durante F. Nesso fisio-pathologico tra la struttura dei nei matemi e la genesi di alcuni tumori maligni. *Arch Memor Observ Chir Pract, 11, 217–226.* 1874.

66. Buess M, Rajski M, Vogel-Durrer BM, Herrmann R, Rochlitz C. Tumor-endothelial interaction links the CD44(+)/CD24(-) phenotype with poor prognosis in early-stage breast cancer. *Neoplasia.* Oct 2009;11(10):987-1002.

67. Ricardo S, Vieira AF, Gerhard R, et al. Breast cancer stem cell markers CD44, CD24 and ALDH1: expression distribution within intrinsic molecular subtype. *J Clin Pathol.* Nov 2011;64(11):937-946.

68. Mylona E, Giannopoulou I, Fasomytakis E, et al. The clinicopathologic and prognostic significance of CD44+/CD24(-/low) and CD44-/CD24+ tumor cells in invasive breast carcinomas. *Hum Pathol.* Jul 2008;39(7):1096-1102.

69. Peitzsch C, Kurth I, Kunz-Schughart L, Baumann M, Dubrovska A. Discovery of the cancer stem cell related determinants of radioresistance. *Radiother Oncol.* Sep 2013;108(3):378-387.

70. Karamboulas C, Ailles L. Developmental signaling pathways in cancer stem cells of solid tumors. *Biochimica et biophysica acta.* Feb 2013;1830(2):2481-2495.

71. Sladek NE. Leukemic cell insensitivity to cyclophosphamide and other oxazaphosphorines mediated by aldehyde dehydrogenase(s). *Cancer Treat Res.* 2002;112:161-175.

72. Marchitti SA, Brocker C, Stagos D, Vasiliou V. Non-P450 aldehyde oxidizing enzymes: the aldehyde dehydrogenase superfamily. *Expert opinion on drug metabolism & toxicology.* Jun 2008;4(6):697-720.

73. Black WJ, Stagos D, Marchitti SA, et al. Human aldehyde dehydrogenase genes: alternatively spliced transcriptional variants and their suggested nomenclature. *Pharmacogenet Genomics.* Nov 2009;19(11):893-902.

74. Vasiliou V, Nebert DW. Analysis and update of the human aldehyde dehydrogenase (ALDH) gene family. *Hum Genomics.* Jun 2005;2(2):138-143.

75. Corti S, Locatelli F, Papadimitriou D, et al. Identification of a primitive brain-derived neural stem cell population based on aldehyde dehydrogenase activity. *Stem Cells.* Apr 2006;24(4):975-985.

76. Kastan MB, Schlaffer E, Russo JE, Colvin OM, Civin CI, Hilton J. Direct demonstration of elevated aldehyde dehydrogenase in human hematopoietic progenitor cells. *Blood.* May 15 1990;75(10):1947-1950.

77. Ginestier C, Hur MH, Charafe-Jauffret E, et al. ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. *Cell Stem Cell.* Nov 2007;1(5):555-567.

78. Woodward WA, Krishnamurthy S, Lodhi A, et al. Aldehyde dehydrogenase1 immunohistochemical staining in primary breast cancer cells independently predicted overall survival but did not correlate with the presence of circulating or disseminated tumors cells. *J Cancer.* 2014;5(5):360-367.

79. Charafe-Jauffret E, Ginestier C, Iovino F, et al. Breast cancer cell lines contain functional cancer stem cells with metastatic capacity and a distinct molecular signature. *Cancer research.* Feb 15 2009;69(4):1302-1313.

80. Dontu G, Abdallah WM, Foley JM, et al. In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells. *Genes Dev.* May 15 2003;17(10):1253-1270.

81. O'Brien CA, Pollett A, Gallinger S, Dick JE. A human colon cancer cell capable of initiating tumour growth in immunodeficient mice. *Nature.* Jan 4 2007;445(7123):106-110.

82. Collins AT, Berry PA, Hyde C, Stower MJ, Maitland NJ. Prospective identification of tumorigenic prostate cancer stem cells. *Cancer research.* Dec 1 2005;65(23):10946-10951.

83. Perrone G, Gaeta LM, Zagami M, et al. In situ identification of CD44+/CD24- cancer cells in primary human breast carcinomas. *PLoS One.* 2012;7(9):e43110.

84. Schwartz T, Stark A, Pang J, et al. Expression of aldehyde dehydrogenase 1 as a marker of mammary stem cells in benign and malignant breast lesions of Ghanaian women. *Cancer.* Feb 1 2013;119(3):488-494.

85. Nalwoga H, Arnes JB, Wabinga H, Akslen LA. Expression of aldehyde dehydrogenase 1 (ALDH1) is associated with basal-like markers and features of aggressive tumours in African breast cancer. *Br J Cancer.* Jan 19 2010;102(2):369-375.

86. Moitra K, Lou H, Dean M. Multidrug efflux pumps and cancer stem cells: insights into multidrug resistance and therapeutic development. *Clin Pharmacol Ther.* Apr 2011;89(4):491-502.

87. Dave B, Chang J. Treatment resistance in stem cells and breast cancer. *J Mammary Gland Biol Neoplasia.* Mar 2009;14(1):79-82.

88. Li X, Lewis MT, Huang J, et al. Intrinsic resistance of tumorigenic breast cancer cells to chemotherapy. *J Natl Cancer Inst.* May 7 2008;100(9):672-679.

89. Campbell LL, Polyak K. Breast tumor heterogeneity: cancer stem cells or clonal evolution? *Cell cycle.* Oct 1 2007;6(19):2332-2338.

90. Charafe-Jauffret E, Ginestier C, Iovino F, et al. Aldehyde dehydrogenase 1-positive cancer stem cells mediate metastasis and poor clinical outcome in inflammatory breast cancer. *Clin Cancer Res.* Jan 1 2010;16(1):45-55.

91. Bianchini G, Balko JM, Mayer IA, Sanders ME, Gianni L. Triple-negative breast cancer: challenges and opportunities of a heterogeneous disease. *Nat Rev Clin Oncol.* May 17 2016.

92. Jones RJ, Matsui W. Cancer Stem Cells: From Bench to Bedside. *Biol Blood Marrow Transplant.* Jan 2007;13(Suppl 1):47-52.

93. Hirschmann-Jax C, Foster AE, Wulf GG, et al. A distinct "side population" of cells with high drug efflux capacity in human tumor cells. *Proc Natl Acad Sci U S A.* Sep 28 2004;101(39):14228-14233.

94. Tanaka H, Nakamura M, Kameda C, et al. The Hedgehog signaling pathway plays an essential role in maintaining the CD44+CD24-/low subpopulation and the side population of breast cancer cells. *Anticancer Res.* Jun 2009;29(6):2147-2157.

95. Kim Y, Joo KM, Jin J, Nam DH. Cancer stem cells and their mechanism of chemo-radiation resistance. *Int J Stem Cells.* May 2009;2(2):109-114.

96. Bao S, Wu Q, McLendon RE, et al. Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. *Nature.* Dec 07 2006;444(7120):756-760.

97. Wang S, Yang D, Lippman ME. Targeting Bcl-2 and Bcl-XL with nonpeptidic small-molecule antagonists. *Seminars in oncology.* Oct 2003;30(5 Suppl 16):133-142.

98. Pardal R, Clarke MF, Morrison SJ. Applying the principles of stem-cell biology to cancer. *Nature reviews. Cancer.* Dec 2003;3(12):895-902.

99. Scharenberg CW, Harkey MA, Torok-Storb B. The ABCG2 transporter is an efficient Hoechst 33342 efflux pump and is preferentially expressed by immature human hematopoietic progenitors. *Blood.* Jan 15 2002;99(2):507-512.

100. Zhao W, Jaganathan S, Turkson J. A cell-permeable Stat3 SH2 domain mimetic inhibits Stat3 activation and induces antitumor cell effects in vitro. *The Journal of biological chemistry.* Nov 12 2010;285(46):35855-35865.

101. Zhang H, Cohen AL, Krishnakumar S, et al. Patient-derived xenografts of triple-negative breast cancer reproduce molecular features of patient tumors and respond to mTOR inhibition. *Breast cancer research : BCR.* 2014;16(2):R36.

102. Zhao X, Liu Z, Yu L, et al. Global gene expression profiling confirms the molecular fidelity of primary tumor-based orthotopic xenograft mouse models of medulloblastoma. *Neuro-oncology.* May 2012;14(5):574-583.

103. de Plater L, Lauge A, Guyader C, et al. Establishment and characterisation of a new breast cancer xenograft obtained from a woman carrying a germline BRCA2 mutation. *Br J Cancer.* Oct 12 2010;103(8):1192-1200.

104. Chen Q, Boire A, Jin X, et al. Carcinoma-astrocyte gap junctions promote brain metastasis by cGAMP transfer. *Nature.* May 26 2016;533(7604):493-498.

105. Fang D, Nguyen TK, Leishear K, et al. A tumorigenic subpopulation with stem cell properties in melanomas. *Cancer research.* Oct 15 2005;65(20):9328-9337.

Chapter 2 : Breast Cancer and African Ancestry: Importance of International Collaborations

Abstract

Women with African ancestry in western, sub-Saharan Africa and in the United States represent a population subset facing an increased risk of being diagnosed with biologically-aggressive phenotypes of breast cancer that are negative for the estrogen receptor, the progesterone receptor, and the HER2/*neu* marker. These tumors are commonly referred to as triple negative breast cancer. Disparities in breast cancer incidence and outcome related to racial/ethnic identity motivated the establishment of a breast cancer research partnership between the Komfo Anokye Teaching Hospital in Kumasi, Ghana and the University of Michigan Comprehensive Cancer Center in Ann Arbor, Michigan, and the Henry Ford Health System (HFHS) in Detroit, Michigan. This research collaborative has featured educational training programs as well as scientific investigations related to the comparative biology of breast cancer in Ghanaian African, African American, and White American patients. Currently, this International Breast Registry has expanded to include African American patients throughout the United States by partnering with the Sisters Network, Inc. (the only network for black breast cancer survivors in the united states); additional

sites in Ghana (representing west Africa) as well as Ethiopia (representing east Africa). Herein we review the history and results of this international program.

Background

Financial constraints are a constant reality in low-and middle-income countries, and pose enormous barriers to both quantifying and addressing the cancer burden in sub-Saharan Africa. The limited data available have been generated by the Globocan 2008 database of the International Agency for Research on Cancer (IARC) and from recent attempts to report population-based cancer incidence rates from selected countries such as Uganda and Ghana [1-6]. Taken together, these resources suggest that breast cancer is an increasing problem, and the likely explanations for this expanding health threat include more prolonged longevity in many African communities (as healthcare access improves, and with breast cancer risk increasing as women age); adoption of westernized, higher-fat dietary patterns (which can increase breast cancer risk in both pre- and postmenopausal women); and increased utilization of reproductive patterns that are more prevalent in western populations (such as delayed childbearing and reduced overall parity, which increased risk of hormone-sensitive breast cancer). Rising rates of breast cancer cases are particularly alarming within sub-Saharan African countries, where the already-overburdened healthcare system is ill-prepared to afford early detection and multidisciplinary

treatment programs. The disturbing and excessively high mortality-to-incidence ratio of breast cancer in sub-Saharan Africa compared to other parts of the world are depicted in Figure 1.

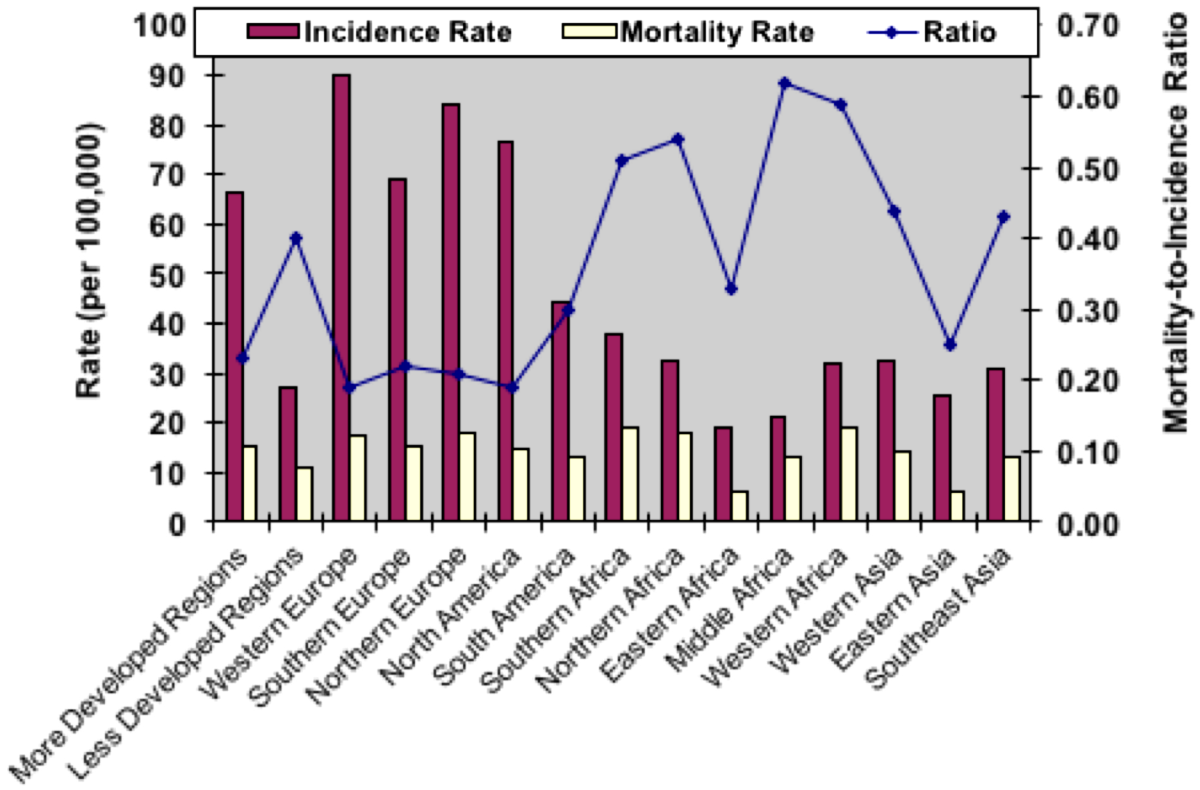


Figure 2-1: International variation in breast cancer incidence, mortality, and mortality-incidence ratios. Data from Jamal et al7, Globocan 20088, Newman et al 9 and Jiagge et al10.

In the United States, race/ethnicity-associated disparities in breast cancer incidence and outcome have been documented for many decades. Breast cancer mortality rates are disproportionately higher for African American compared to White/European women, and African American women tend to be diagnosed with breast cancer at younger ages.

The mortality differences have historically been ascribed to socioeconomic differences, since poverty rates and inadequate health care are more prevalent among African Americans. The younger age distribution for breast cancer, studies documenting higher mortality risk after adjusting for socioeconomic status, and advancing insights regarding the higher frequency of biologically aggressive tumors (such as triple negative breast cancer, TNBC) within the African American community have fueled speculation that African ancestry itself might be associated with hereditary susceptibility for specific patterns of breast cancer[11-13]. These issues provided the foundation for establishing the International Breast Registry (IBR), a breast cancer research partnership between the Komfo Anokye Teaching Hospital (KATH), the University of Michigan (UM), and the Henry Ford Health System (HFHS). This program is now being overseen by Dr. Sofia Merajver at the University of Michigan with our collaborators at the HFHS. Exploratory conversations and bi-continental introductory visits occurred in 2004-2005, and the institutional review boards of each institution provided their initial human research ethics approvals in 2006. The early goals of this collaborative were therefore related to studying the biology of breast cancer in women with African ancestry, and indeed the first joint publication from this team was a manuscript providing “how-to” guidelines for other investigators regarding the conduct of cancer research in developing/low- and middle-income countries.[14]

The IBR has grown enormously since its inception. While this collaborative continues to feature a robust breast tumor repository that has provided exciting data regarding breast cancer in Ghanaian African as well as African American women, it has also expanded its portfolio of educational and training exchange programs. In this review, we summarize the various outcomes of this international effort over the past ten years, as presented in peer-reviewed publications and academic meeting abstracts. We will also review the non-research related productivity of this partnership, featuring investment in the oncology services infrastructure of Ghana.

Breast Cancer and African Ancestry- Patterns identified through the IBR

The concept of subtyping breast cancer has assumed increasing importance as our knowledge of targeted therapy has advanced. Invasive breast cancers that are positive for the estrogen receptor and/or the progesterone receptor can be managed systemically with a variety of endocrine agents- tamoxifen for premenopausal patients; and tamoxifen or one of the aromatase inhibitors for postmenopausal patients. Tumors that overexpress HER2/*neu* benefit greatly from targeted, anti-HER2 therapy with trastuzamab and/or pertuzamab. Equally important and relevant to treatment planning is the fact that these targeted agents are contraindicated in cases that are negative for these markers, and utilizing them in cases that are negative for all three markers- tumors known as triple negative breast cancer (TNBC)- would result in exposing

patients unnecessarily to the toxicity of an ineffective regimen. Studies of breast cancer in the United States have revealed that frequency and population-based incidence rates of TNBC are significantly higher in African American women compared to women of other racial/ethnic identities [15,16]. Speculation that African ancestry could be associated with inherited susceptibility for TNBC prompted comparisons of breast tumor phenotypes in African American and western, sub-Saharan African women, as these two population subsets have shared ancestry resulting from the colonial-era slave trade[13].

One of the initial publications of the KATH-Michigan IBR collaborative therefore focused on comparisons of TNBC prevalence in White/Caucasian Americans compared to African American and Ghanaian women. The Henry Ford Health System (HFHS), which provides care to the robustly-diverse metropolitan Detroit Michigan community, served as the source for the comparison patient population. This study confirmed that TNBC accounted for the majority of KATH breast cancer patients[17]. Subsequent reports from this partnership, based upon larger sample sizes, have documented similar results. The most recent analyses, based upon immunohistochemistry studies of 234 Ghanaian breast cancer patients (with complete marker profiling on 173 invasive tumors), reveal that 92 (53.2%) are triple-negative[18]. In contrast, the lowest frequency of TNBC was seen in WA patients (15.5%); and TNBC frequency for AA patients was intermediate between these two extremes at 30%. This pattern persisted in subset analysis of patients younger than 50 years. These findings

are consistent with the theory that extent of African ancestry correlates with likelihood of being diagnosed with TNBC, since AA patients represent a genetically admixed population.

Table 2-1 summarizes the results of studies characterizing the breast cancer burden of Ghanaian women based upon data from the IBR research collaborative. While studies from the IBR have been enlightening with regard to understanding the breast cancer burden of western sub-Saharan Africa as exemplified by Ghanaian women, expansion of the registry has provided opportunities to study breast cancer in East Africa as well. Contributions from partners at the St. Paul's Millennium Hospital in Addis Ababa have yielded interesting preliminary findings regarding breast tumor phenotypes from Ethiopian women. Immunohistochemistry performed at UM on 95 invasive breast cancers from Addis Ababa have revealed a low frequency of TNBC, at 9%[19]. In another study at the Merajver lab with colleagues from Ethiopia through the UMAPS program, we performed immunohistochemistry on 100 invasive breast tumors from Ethiopia and found the incidence of TNBC to be about 4%(waiting for results of FISH analysis, data not yet published). Patterns of the African diaspora and shared ancestry may explain these disparate results regarding the prevalence of TNBC in AA and Ghanaian women compared to Ethiopians. The colonial-era slave trade that was controlled by Europeans for the purpose of generating a labor supply in North America focused on capture and enforced trans-Atlantic transport of west Africans[20]. In contrast, enslavement of east

Africans from the region of Ethiopia was predominately perpetrated by Arab traders and enforced migration to northern Africa and Asia[25,26]. Fig 2-2 is shows the population migration patterns from different parts of Africa[20].

Table 2-1: Features of breast cancer in Ghana based upon studies from the Ghana-Michigan Breast Cancer Research Partnership

Study	**Sample size and Source of Breast Cancers in Africa**	**Frequency of Triple Negative Tumors**	**Others features Identified**
Stark, 2010[17]	75	82%	Frequency of TNBC among African Americans= 33%; Frequency of TNBC among White/Caucasian Americans=10% P<0.01
Schwartz, 2013[21]	104	56%	Frequency mammary stem cell marker ALDH1 expression in TNBC cases 53% versus 33% in non-TNBC cases P<0.05
Proctor, 2015[22]	147	61%	Frequency of mammary stem cell marker ALDH1 expression in TNBC cases= 45%; Frequency Androgen Receptor expression in TNBC cases= 24%
Pang, 2012[23]	100	60%	Oncogene EZH2 more prevalent in both nucleus and cytoplasm of TNBC
Der, 2015[24]	219	58%	N/A
Jiagge and Newman, 2015[18]	234	53%	N/A
Newman, 2016[19]	234 (KATH, Ghana) 95 (St. Paul's Millenium Hospital, Addis Ababa, Ethiopia)	Ghana: 53% Ethiopia: 9%	Increased frequency (36%) of HER2/neu-overexpressing breast cancers in Ethiopia

AA women are therefore more likely to have shared ancestry with women from western Africa, and this has been confirmed by genotyping studies of markers *are associated with geographic ancestry, commonly referred-to as Ancestry Informative Markers (AIMs)*[27-29].

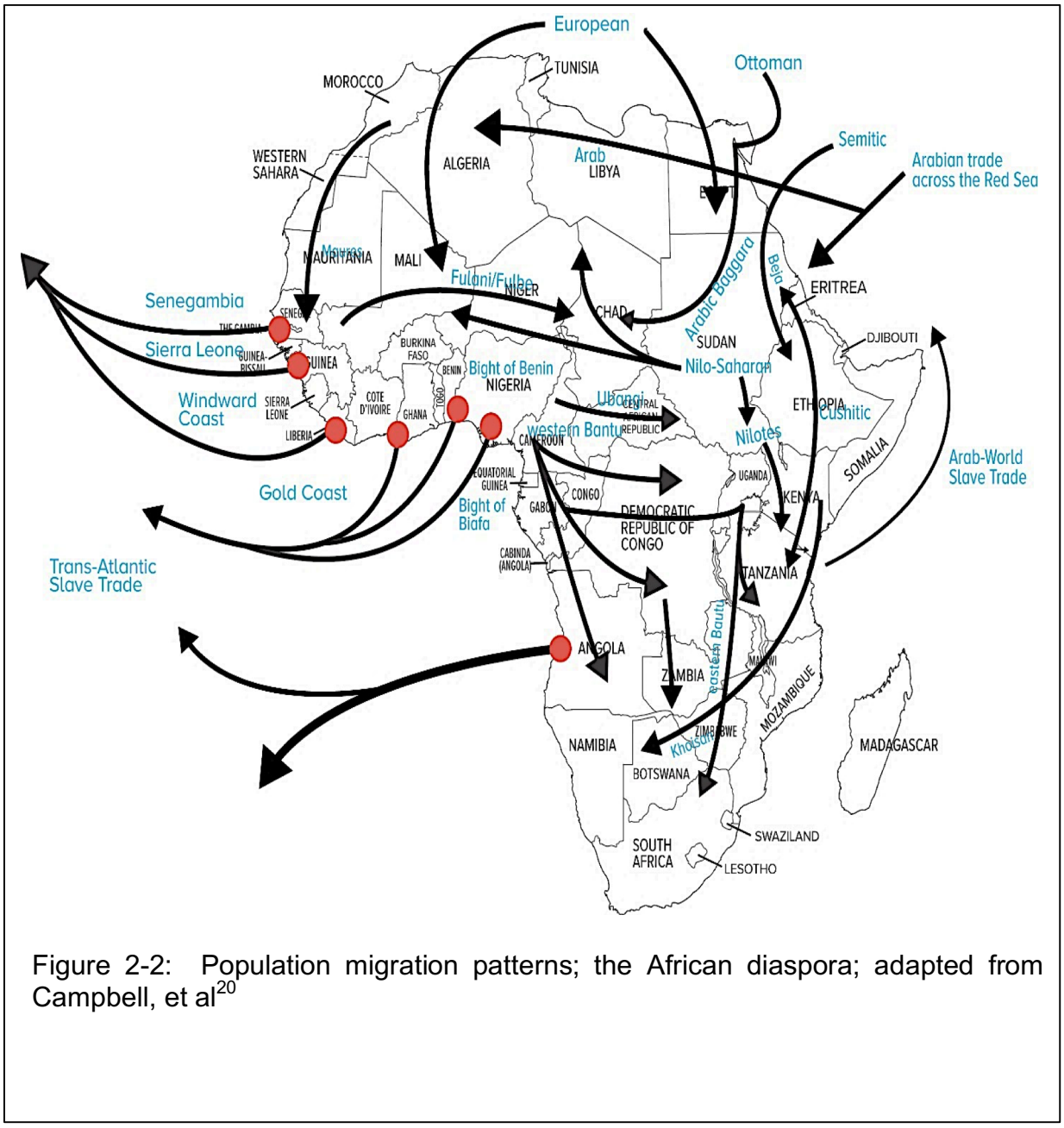


Figure 2-2: Population migration patterns; the African diaspora; adapted from Campbell, et al[20]

Studies of breast tumor biology have generated the “stem cell” theory, which hypothesizes that mammary tumor virulence and metastatic risk is driven

by a small subset of cells within the cancer known as the stem cells. Efforts to identify and characterize the mammary stem cells therefore represent an exciting body of research. UM researchers have pioneered studies of the mammary stem cell hypothesis, and have reported on aldehyde dehydrogenase-1 (ALDH1) expression as a reliable, immunohistochemically-detectable marker of the mammary stem cell as well as being associated with more virulent tumors[30]. Studies of KATH tumors at UM demonstrated elevated expression of this marker in both benign and malignant breast tissue from Ghanaian women[21].

The polycomb group protein EZH2 is another molecule that has been implicated in mammary stem cell and TNBC progression, as demonstrated by UM researchers involved with the IBR[31,32]. We therefore incorporated studies of EZH2 into the IBR. This marker was found to have unique patterns of expression in Ghanaian breast specimens studied through the KATH-UM partnership[23].

Lastly, TNBC subtyping is a promising avenue for gaining insights regarding more refined, personalized treatment of TNBC[33-37]. Existing research suggests that there are at least six different TNBC subtypes, and some of the distinguishing features are related to luminal-like characteristics seen in the Androgen Receptor subtype versus the stem cell-like characteristics seen in the mesenchymal subtypes. Despite the disproportionately high frequency of TNBC among women with African ancestry, the TNBC subtyping research has been based almost exclusively on datasets representing White American, European, and Asian populations. No African datasets of gene expression profiles have

been available for inclusion in this body of research thus far. These issues motivated exploratory analyses of tissue from the IBR looking at expression of both androgen receptor and ALDH1 as immunohistochemistry surrogates for TNBC subtypes. Among the Ghanaian tumors, an intriguing finding appeared suggesting a novel TNBC subtype, featuring co-dominance of both androgen receptor and ALDH1 pathways[38].

As noted above, the partnerships within the IBR oncology teams has generated valuable insights regarding breast cancer phenotypes that are more prevalent among women with West African ancestry, such as triple negative tumors. Patient-derived xenografts (PDXs) represent an exciting research strategy for studying breast tumor biology and novel therapies. Implantation of breast cancer fragments into mice mammary fat pads yields a renewable supply of human tumors that can be used for a variety of in vivo rodent model experiments. Through the KATH-UM research partnership, a series of PDXs have been created based upon the tumors from Ghanaian; African American; and White American breast cancer patients.

Breast Cancer in Ghana- Diagnostic and Treatment Advances Promoted by the Ghana-Michigan Partnership

The IBR collaborative has also served as an investment in improving the clinical services available to Ghanaian women with breast problems. Historically, the KATH breast clinics relied predominantly upon open surgical diagnostic

biopsies in order to confirm or rule out the presence of cancer in any woman presenting with a breast abnormality or mass. The scheduling and implementation of a surgical diagnostic procedure is time-consuming and utilizes valuable, costly operating room resources. Furthermore, this sequence increases the risk that the affected patient (who has likely already travelled a distance and expended personal finances in order to seek medical attention) may be lost to follow-up. The early years of this partnership featured a training program in utilizing percutaneous core needle biopsies to diagnose breast cancer on-site during the outpatient clinic visit. The success of this program in terms of accurately and efficiently establishing a diagnosis of breast cancer has been reviewed and documented[39].

Multidisciplinary/multimodality treatment of breast cancer requires immunohistochemistry resources that can efficiently and accurately assess for expression of the estrogen receptor, the progesterone receptor, and HER2/*neu*. These markers are critical in being able to appropriately determine a patient's response to endocrine therapy and/or targeted anti-HER2/*neu* therapy. Lacking this molecular marker information, hormone receptor-negative breast cancer patients may face the toxicity of ineffective endocrine therapy, and hormone receptor-positive breast cancer patients may miss the opportunity to receive life-saving endocrine therapy. Similarly, HER2/*neu* expression can identify tumors that are quite sensitive to chemotherapy and anti-HER2/*neu* agents. The Ghana-Michigan partnership featured a training program in immunohistochemistry for

the KATH pathology team, as well as the development of a resource supply and allocation system that has enabled the KATH Breast Oncology Program to routinely generate their own molecular marker reports for each patient diagnosed with invasive breast cancer (Figure 2-2).

This partnership has also generated opportunities to utilize telemedicine technology for real-time international collaboration and multidisciplinary discussion. The UM partners invested in dedicated internet and teleconference equipment on-site at KATH. This teleconference unit allows the UM and KATH teams to discuss patient care on a weekly basis and to share conference proceedings through live interactions.

Our Ghanaian partners have also been able to promote general breast health awareness programs as a consequence of the expanded breast cancer attention generated by this research collaborative. KATH PI and manuscript co-senior author B.A. has worked with regional herbalists, primary care physicians, and nurses to conduct educational seminars promoting breast cancer early detection strategies. Lacking adequate financial resources as well as accurate population-based statistics on breast cancer incidence, it is not feasible to conduct community-wide mammographic screening programs. These early detection programs therefore largely focus on dissemination of information regarding clinical signs and symptoms of breast cancer (dominant lump, bloody nipple discharge, etc.) and the importance of prompt biopsy with initiation of treatment. While we cannot quantify the effectiveness of our efforts with regard to

breast cancer stage distribution in Ghana, our KATH colleagues have established a tumor registry office (Kumasi Cancer Registry, Dr. Baffour Awuah, Director; unpublished data; personal communication) and have seen increasing volumes of breast cancer annually as well as anecdotal observations of more women presenting with operable, earlier-stage disease. Strengthening this Tumor Registry (which has struggled with consistent personnel and completeness of data collection) remains a high priority for the Michigan-Ghana Collaborative and we have also invested in advanced training of tumor registry personnel Although lumpectomy and breast radiation is an option that is available to breast cancer patients in Ghana, few women present with tumors that are amenable to the breast conservation approach despite anecdotal observations of an earlier stage distribution for KATH patients. Furthermore, concerns regarding inadequate pre- and postoperative mammographic imaging availability has generated suspicion that breast-conserving surgery cannot be planned with optimal information regarding extent of disease and adequacy of resection."

Breast Cancer and African Ancestry: The Ghana-Michigan Partnership as a Model for Expansion

The multifaceted success of the Michigan partnership with the Komfo Anokye Teaching Hospital has established the foundation for expansion and collaboration with other healthcare facilities in Africa. As a consequence, this program has grown, with exchange programs that have included St. Paul's

Millennium Hospital in Ethiopia as well as three additional sites in Ghana: the Korle Bu Teaching Hospital in Accra; the Tamale Teaching Hospital in Tamale; and the Sunyani Teaching Hospital in Sunyani. Table 2 summarizes the studies that have been published and presented through these partnerships.

International expansion efforts beyond Africa are also in development, as partnerships evolve with the All India Institute of Medical Sciences in New Delhi and the Third Xiangya Hospital, Central South University.

Table 2-2: Published studies and reports related to the Ghana-Michigan Breast Cancer Collaborative

Report	Year
Guide for Investigators Conducting International Research Involving Developing Nations[14]	2010
African Ancestry and Higher Prevalence of Triple-Negative Breast Cancer[17]	2010
Implementation of a Percutaneous Core Needle Biopsy Training Program: Results from the University of Michigan-Komfo Anokye Teaching Hospital breast cancer research partnership[39]	2011
Invasive Breast Carcinomas in Ghana: High frequency of high-grade, basal-like histology and high EZH2 expression[23]	2012
Expression of Aldehyde Dehydrogenase 1 as a Marker of Mammary Stem Cells in Benign and Malignant Breast Lesions of Ghanaian Women[40]	2013
Androgen receptor expression in Ghanaian breast cancer cases: novel correlation with ALDH1 in triple-negative tumors[22,38]	2014
Distinct Pathways Differentiate the CD44+ Mesenchymal-like from the ALDH+ Epithelial-Like Phenotype of Triple negative Breast Cancer Stem Cells[41]	2014
Creating a Comprehensive Patient Derived Xenograft Bank to Represent Racial Disparities in Triple Negative Breast Cancer[42]	2014
Global surgical oncology disease burden: addressing disparities via global surgery initiatives: the University of Michigan International Breast Cancer Registry [10]	2015
Comparative analysis of breast cancer phenotypes in African American, White American, and West versus East African patients- Correlation between African ancestry and triple negative breast cancer[19]	2016

The KATH-UM relationship has served as a platform for training the physician-scientists from Ghana. More than a dozen Ghanaians have spent time at UM for observorships and research programs in the Merajver lab as well. E.J. is a surgeon from KATH, currently completing her PhD work at UM in Cancer Biology, after which she will return to Ghana to assume leadership of a Ghana-based translational research program.

Within the United States, expansion of this registry has involved partnership with the Sisters Network, Inc. The Sisters Network is a national organization of African American breast cancer survivors[43], and currently has a membership of nearly three thousand women in thirty chapters across twenty-two states. The Sisters Network membership contributes to the International Registry by providing recruitment opportunities at their national as well as local meetings. Recruitment involves participants agreeing to provide access to medical records and saliva specimens suitable for DNA extraction and genotyping studies. In their 2016 national conference held at Detroit Michigan, the UM comprehensive cancer center featured with educational talks, counseling opportunities and a chance to get involved in the lives of the very important group of breast cancer survivors.

Long-term goals of this collaboration include studies of germline breast cancer risk in women with diverse racial-ethnic backgrounds in the United States compared to international populations.

We are in the process of establishing a cancer research institution in Ghana, that will serve as a model research institution for serve the western African sub-region. It is our hope that the institution, that is registered under the name, Research Institute of Health, Ghana (RIHG), will study chronic diseases in the African population with the hope of beret understanding the biology of these diseases and developing new interventions. It will partner with all the major Universities and teaching hospitals in Ghana, the west African sub-region and the United States.

We enthusiastically look forward to this expanded research as a valuable contribution to precision medicine initiatives. We are also committed to ongoing investment in the breast oncology services available to patients in all of our partnering sites as well as throughout other low and middle income countries. Establishment of core biopsy programs, immunohistochemistry programs, and multidisciplinary tumor board conferences are examples of the services that are promoted through our research collaborative.

Other investigators with an interest in this type of international cancer research should be mindful of evolving policies that regulate transport of human tissue via commercial airlines. Updates on regulatory requirements and restrictions can be obtained at the Centers for Disease Control website (http://www.cdc.gov/importation/index.html). Our group typically transported breast tissues for immunohistochemistry studies as formalin-fixed specimens embedded in paraffin blocks. The fresh specimens for PDX work were

transported in dry ice. Commercial carriers have regulations for labeling and packaging of dry ice, including maximum limits for quantity of dry ice that can be transported. Updates on these regulations can be accessed through the particular carrier's "Dangerous Goods" office and website. Investigators should also work with their institution's liaison to the Occupational Safety and Health Administration (OSHA) for training in handling dry ice, and further information can be obtained from OSHA directly (www.osha.gov).

International partnerships represent a powerful and unique opportunity to advance insights regarding the etiology of domestic disparities in breast cancer burden related to racial-ethnic identity. These efforts also provide valuable cultural, academic and educational exchange programs as well as opportunities to strengthen the oncology services in under-resourced countries. Our program is certainly not unique; "medical missionary"- type work has existed for many decades. Hopefully this summary of our ten-year experience with international breast cancer outreach and research will provide motivation for others to add to this growing field.

References

1. Forouzanfar MH, Foreman KJ, Delossantos AM, et al. Breast and cervical cancer in 187 countries between 1980 and 2010: a systematic analysis. *Lancet.* 2011;378(9801):1461-1484.

2. Ferlay J, Bray F, Pisani P, Parkin DM. Globocan 2002: Cancer Incidence, Mprtality and Prevalence Worldwide. *IARC CancerBase No 5 version 20, IARCPress, Lyon.* 2004.

3. Globocan Fact Sheet, http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx. Accessed April 11, 2015.

4. Jemal A, Bray F, Forman D, et al. Cancer burden in Africa and opportunities for prevention. *Cancer.* 2012.

5. Parkin DM, Nambooze S, Wabwire-Mangen F, Wabinga HR. Changing cancer incidence in Kampala, Uganda, 1991-2006. *Int J Cancer.* 2010;126(5):1187-1195.

6. Laryea DO, Awuah B, Amoako YA, et al. Cancer incidence in Ghana, 2012: evidence from a population-based cancer registry. *BMC Cancer.* 2014;14:362.

7. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. *CA Cancer J Clin.*61(2):69-90.

8. Jaeger BA, Jueckstock J, Andergassen U, et al. Evaluation of two different analytical methods for circulating tumor cell detection in peripheral blood of patients with primary breast cancer. *BioMed research international.* 2014;2014:491459.

9. Newman LA. Breast Cancer Disparities: High-Risk Breast Cancer and African Ancestry. *Surg Oncol Clin N Am.* 2014;23(3):579-592.

10. Jiagge E, Bensenhaver JM, Oppong JK, Awuah B, Newman LA. Global surgical oncology disease burden: addressing disparities via global surgery initiatives: the university of michigan international breast cancer registry. *Ann Surg Oncol.* 2015;22(3):734-740.

11. Jatoi I, Becher H, Leake CR. Widening disparity in survival between white and African-American patients with breast carcinoma treated in the U. S. Department of Defense Healthcare system. *Cancer.* 2003;98(5):894-899.

12. 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.* 2006;24(9):1342-1349.

13. Fregene A, Newman LA. Breast cancer in sub-Saharan Africa: how does it relate to breast cancer in African-American women? *Cancer.* 2005;103(8):1540-1550.

14. Martin IK, Awuah B, Newman LA. Guide for investigators conducting international cancer research involving developing nations. *Cancer.* 2010;116(6):1396-1399.

15. Kohler BA, Sherman RL, Howlader N, et al. Annual report to the nation on the status of cancer, 1975-2011, featuring incidence of breast cancer subtypes by race/ethnicity, poverty, and state. *J Natl Cancer Inst.* 2015;107(6).

16. Newman LA, Reis-Filho JS, Morrow M, Carey LA, King TA. The 2014 Society of Surgical Oncology Susan G. Komen for the Cure Symposium: triple-negative breast cancer. *Annals of surgical oncology.* 2015;22(3):874-882.

17. Stark A, Kleer CG, Martin I, et al. African ancestry and higher prevalence of triple-negative breast cancer: findings from an international study. *Cancer.* 2010;116(21):4926-4932.

18. Jiagge E, Awuah B, Bensenhaver J, et al. Comparative analysis of breast cancer phenotypes in African American, White American, and African patients- Correlation between African ancestry and triple negative breast cancer. Paper presented at: San Antonio Breast Cancer Symposium; December 12, 2015, 2015; San Antonio, TX.

19. Newman L, Jiagge E, Bensenhaver J, Stark A, Awuah B, Jibril A. Comparative analysis of breast cancer phenotypes in African American, White American, and West versus East African patients- Correlation between African ancestry and triple negative breast cancer. Paper presented at: Society of Surgical Oncology Annual Cancer Symposium; March 4, 2016, 2016; Boston, MA.

20. Campbell MC, Hirbo JB, Townsend JP, Tishkoff SA. The peopling of the African continent and the diaspora into the new world. *Curr Opin Genet Dev.* 2014;29:120-132.

21. Schwartz T, Stark A, Pang J, et al. Expression of ALDH1 as a marker of mammary stem cells in benign and malignant breast lesions of Ghanaian women. *Cancer, in press.* 2012.

22. Proctor E, Kidwell KM, Jiagge E, et al. Characterizing Breast Cancer in a Population with Increased Prevalence of Triple-Negative Breast Cancer: Androgen Receptor and ALDH1 Expression in Ghanaian Women. *Ann Surg Oncol.* 2015.

23. Pang J, Toy KA, Griffith KA, et al. Invasive breast carcinomas in Ghana: high frequency of high grade, basal-like histology and high EZH2 expression. *Breast Cancer Res Treat.* 2012;135(1):59-66.

24. Der EM, Gyasi RK, Tettey Y, et al. Triple-Negative Breast Cancer in Ghanaian Women: The Korle Bu Teaching Hospital Experience. *Breast J.* 2015;21(6):627-633.

25. Harich N, Costa MD, Fernandes V, et al. The trans-Saharan slave trade - clues from interpolation analyses and high-resolution characterization of mitochondrial DNA lineages. *BMC evolutionary biology.* 2010;10:138.

26. Richards M, Rengo C, Cruciani F, et al. Extensive female-mediated gene flow from sub-Saharan Africa into near eastern Arab populations. *Am J Hum Genet.* 2003;72(4):1058-1064.

27. Stefflova K, Dulik MC, Barnholtz-Sloan JS, Pai AA, Walker AH, Rebbeck TR. Dissecting the within-Africa ancestry of populations of African descent in the Americas. *PLoS One.* 2011;6(1):e14495.

28. Salas A, Carracedo A, Richards M, Macaulay V. Charting the ancestry of African Americans. *Am J Hum Genet.* 2005;77(4):676-680.

29. Bryc K, Auton A, Nelson MR, et al. Genome-wide patterns of population structure and admixture in West Africans and African Americans. *Proc Natl Acad Sci U S A.* 2010;107(2):786-791.

30. Ginestier C, Hur MH, Charafe-Jauffret E, et al. ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. *Cell Stem Cell.* 2007;1(5):555-567.

31. Kunju LP, Cookingham C, Toy KA, Chen W, Sabel MS, Kleer CG. EZH2 and ALDH-1 mark breast epithelium at risk for breast cancer development. *Mod Pathol.* 2011;24(6):786-793.

32. Moore HM, Gonzalez ME, Toy KA, Cimino-Mathews A, Argani P, Kleer CG. EZH2 inhibition decreases p38 signaling and suppresses breast cancer motility and metastasis. *Breast Cancer Res Treat.* 2013;138(3):741-752.

33. Christgen M, Geffers R, Kreipe H, Lehmann U. IPH-926 lobular breast cancer cells are triple-negative but their microarray profile uncovers a luminal subtype. *Cancer Sci.* 2013;104(12):1726-1730.

34. Lehmann BD, Bauer JA, Chen X, et al. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J Clin Invest.* 2011.

35. Lehmann BD, Bauer JA, Chen X, et al. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J Clin Invest.* 2011;121(7):2750-2767.

36. Lehmann BD, Pietenpol JA. Identification and use of biomarkers in treatment strategies for triple-negative breast cancer subtypes. *J Pathol.* 2014;232(2):142-150.

37. Masuda H, Baggerly KA, Wang Y, et al. Differential response to neoadjuvant chemotherapy among 7 triple-negative breast cancer molecular subtypes. *Clin Cancer Res.* 2013;19(19):5533-5540.

38. Proctor E, Jiagge E, Kleer C, et al. Androgen receptor expression in Ghanaian breast cancer cases: novel correlation with ALDH1 in triple-negative tumors. Paper presented at: Society of Surgical Oncology Annual Cancer Symposium2014; Phoenix, AZ.

39. Awuah B, Martin IK, Takyi V, et al. Implementation of a percutaneous core needle biopsy training program: results from the University of Michigan-Komfo Anokye Teaching Hospital breast cancer research partnership. *Ann Surg Oncol.* 2011;18(4):957-960.

40. Schwartz T, Stark A, Pang J, et al. Expression of aldehyde dehydrogenase 1 as a marker of mammary stem cells in benign and malignant breast lesions of Ghanaian women. *Cancer.* 2013;119(3):488-494.

41. Jiagge EM, Wong S, Luther T, et al. Distinct pathways differentiate the CD44+ mesenchymal-like from the ALDH+ epithelial-like phenotype of triple negative breast cancer stem cells. Late-breaking abstract #59, April 7, 2014, San Diego, CA. Paper presented at: American Association for Cancer research (AACR) Annual Meeting 20142014; San Diego, CA.

42. Luther T, Jiagge E, Lewis MT, et al. Creating a comprehensive Patient Derived Xenograft (PDX) to represent racial disparities in triple negative breast cancer. Paper presented at: Seventh AACR Conference on The Science of Cancer Health Disparities in Racial/Ethnic Minorities and the Medically Underserved, November 9–12, 2014, San Antonio, Texas. 2014.

43. Newman LA, Jackson KE. Advocacy Corner: Sisters Network, Inc. *Journal of Oncology Practice.* 2009;November 1, 2009:313-314.

Chapter 3 : Correlation Between African Ancestry, Triple Negative Breast Cancer and BRCA Gene Expression

Abstract

Triple negative breast cancer (TNBC) is the most common breast cancer subtype diagnosed in women with BRCA1 mutation. It is more common among African American (AA) and western sub-Saharan African breast cancer (BC) patients compared to White Americans (WA) and Europeans. Though women in western sub-Saharan Africa have the highest single institutional incidence of BC, there is very little known about the prevalence of BRCA mutation among these patients. Also, very little is known about TNBC in east Africa. Here we compared the prevalence of TNBC between Western Sub-Saharan African, Eastern African, AA and WA women. Considering the shared phenotype between TNBC patients and patients with BRCA mutation, we evaluated the expression of the BRCA and BRCA2 protein in FFPE samples from Ghanaian BC patients and also began studies evaluating the occurrence of BRCA1and BRCA2 germline and somatic mutations in Ghana TNBC patients.

We first evaluated Invasive BC diagnosed 1998-2014 by immunohistochemistry: WA and AA patients from the Henry Ford Health System in Detroit, Michigan; Ghanaian/west Africans from the Komfo Anokye Teaching

Hospital in Kumasi, Ghana; and Ethiopian/east Africans from the St. Paul's Hospital Millennium Medical College in Addis Ababa, Ethiopia. Histopathology and immunohistochemistry for estrogen receptor (ER), progesterone receptor (PR) and HER2/*neu* expression was performed at the University of Michigan on formalin-fixed, paraffin-embedded samples from all cases.

We then performed IHC staining of TNBC tumors from Ghanaian patients for the expression of the BRCA1 and BRCA2 protein. We also sequenced six TNBC tumors from Ghanaian patients with match normal DNA obtained from the saliva.

234 Ghanaian (mean age 49 yrs.); 94 Ethiopian (mean age 43); 271 AA (mean age 60); and 321 WA (mean age 62; P=0.001 for the four comparisons) patients were compared. ER-negative and TNBC were more common among Ghanaian and AA compared to WA and Ethiopian cases (frequency ER-negativity 67.5% and 37.1% versus 19.8% and 28.7% respectively, p<0.0001; frequency TNBC 53.2% and 29.8% versus 15.5% and 15%, respectively, p<0.0001). Among patients younger than 50, prevalence of TNBC remained highest among Ghanaians (50.8%) and AA (34.3%) compared to WA and Ethiopians (15.9% in each); (P<0.0001). more than 50% of Ghanaian TNBC patients had no expression of both BRCA1 and BRCA2 proteins and we identified novel BRCA1/2 gene mutations in Ghanaian tumors, some that can lead to truncation to of the protein.

Conclusions:

This study confirms an association between TNBC and West African ancestry; TNBC frequency is highest in west African women, intermediate among AA patients consistent with genetic admixture following the west Africa-based trans-Atlantic slave trade and lowest in Eastern African and WA women. TNBC frequency was low among Ethiopians/East Africans; this may reflect less shared ancestry between AA and Ethiopians. The very high occurrence of absent expression of BRCA1 and BRCA2 protein in Ghanaian patients and the relatively high frequency of germline mutations indicate the need to study BRCA and its associated genes in this understudied population.

Introduction

African American women bear a disproportionately high share of the breast cancer mortality burden in the United States, and this outcome disparity has been increasing. Socioeconomic disadvantages that are more prevalent in the African American community, such as poverty rates and being under-insured, undeniably contribute to barriers in accessing the health care system and result in poorer cancer control overall.

However, there are features characterizing the epidemiology of breast cancer in African Americans that cannot be easily ascribed to socioeconomic inequities. Until recently population-based lifetime incidence rates of breast cancer have been lower for African American compared to White American women; the average age at breast cancer diagnosis is younger for African American patients; and African American women have higher rates of hormone receptor negative and the biologically more-virulent triple negative tumors. Lastly, population-based incidence rates of male breast cancer are higher in the African American community.

A similar pattern characterizes BC in patients with BRCA1 mutation: they occur at a younger age, are aggressive, have a higher incidence of TNBC and have a higher incidence of male breast cancer. I western Sub-Saharan Africa, available data indicates a similar pattern of BC phenotype with the incidence of TNBC being even higher; as high as 80%. There is very little reported however on the incidence of BRCA mutations in these patients. This is largely due to the

cost involved in BRCA testing, and that lack of availability of the test to those who may be able to afford. It is however very important to know the prevalence of BRCA mutations in this population.

We therefore sought to evaluate the question of whether African ancestry is associated with some heritable marker of risk for particular patterns of breast cancer pathogenesis by studying breast tumor phenotypes in four different population subsets: White Americans; African Americans; Ghanaians (representing west Africa); and Ethiopians (representing east Africa).

In the Ghanaian population, we first performed a preliminary analysis by staining for BRCA1 and BRCA2 protein using IHC staining of FFPE samples from the patients. We also investigated the incidence of BRCA1 and BRCA2 and associated genes germline and somatic mutations in Ghanaian TNBC.

Methods

Ethics

This research effort represents clinic pathology studies that are components of an international breast cancer registry, approved by the Institutional Review Boards of the University of Michigan; the Henry Ford Health System; and the human research ethics approval and/or institutional departmental approval equivalents for the Komfo Anoyke Teaching Hospital in Kumasi Ghana (Committee on Human Research Publication and Ethics, Kwame

Nkrumah University of Science and Technology) and the St. Paul's Hospital Millennium Medical College in Addis Ababa, Ethiopia. Samples and results were de-identified/anonymized prior to analyses.

Pathology and Immunohistochemistry

Histopathology to confirm the diagnosis of breast cancer and immunohistochemistry for molecular marker studies were performed on the specimens from African American and White American cases by the Department of Pathology at the Henry Ford Health System. Evaluation of the Ghanaian and Ethiopian tumor specimens were performed at the University of Michigan North Campus Research Complex. Nuclear expression of hormone receptor (ER and PR) proteins was detected with specific monoclonal antibodies using a labeled streptavidin-biotin immunoperoxidase method. The immunohistochemical assay was performed on deparaffinized formalin-fixed tissue sections of the specimens. Monoclonal mouse antibodies to human ER (DAKO clone ID5) and to human PR (DAKO clone PgR636) were used with a DAKO automated immunostainer following the manufacturer's protocol. Immunohistochemistry for HER2/*neu* staining was performed using the HerceptTest (DAKO, Glostrup, Denmark), an FDA-approved clinical test that qualitatively identifies by light microscopy p185 HER2 overexpression in breast cancer cells. Molecular marker staining was interpreted by in compliance with ASCO/CAP guidelines and as per Fitzgibbons et al[1-3]. Tumors were scored as ER/PR-negative if they had less than

1% nuclear staining. Confirmed ER and PR positive tumors served as positive controls, and normal adjacent mammary gland ductules present in the sections of tumor served as internal positive controls for the hormone receptors. The expression of HER2 was scored based on recommendations from Fitzgibbons et al[1]. Grading was based on the degree and intensity of membrane labeling of tumor cells, on a scale from 0-3+, as follows: grade 0 (no observable labeling or faint, incomplete, or barely detectable membrane labeling in <10% of tumor cells), 1^+ (faint, incomplete, or barely detectable membrane labeling in >10% of tumor cells), 2^+ (incomplete and/or weak to moderate complete membrane staining in >10% of tumor cells, or complete, intense membrane labeling in <10% of tumor cells) or 3^+ (intense, complete membrane labeling in >10% of tumor cells). A specimen scored as 0 or 1+ was classified as HER2/*neu* negative, and specimens scored as 3+ were considered positive. Those specimens with a grade of 2+ were considered equivocal, and follow-up fluorescent in situ hybridization (FISH) was used to assess amplification of the HER2/*neu* gene in the cases with 2+.

Exome Sequencing and Data Analysis

Exome library construction: 150-200ng of patient normal and tumor were used for exome library construction. The genomic DNA was sheared to sizes between 150-200bp using the Covaris S2 system as per manufacturer protocol. Libraries were constructed using an in-house protocol using a combination of

Kapa LTP kit and Agilent' SureSelect Human All ExonV5 capture probes. Exome libraries were quantified with Qubit high sensitivity DNA assay and size was determined with the BioAnalyzer using the High Sensitivity DNA Assay. Paired end NGS: All NGS was carried out using Illumina HiSeq 2000 with a read length of 83 cycles x 7 cycles x 83 cycles. Pooled libraries were first clustered onto flowcells using Illumina's cBot and HiSeq Paired End Cluster Generation kits at a concentration of 14pM. Analysis of the resulting data was done with an in-house developed pipeline.

Ascertainment of Cases

African American and White American: At Henry Ford Health System, women diagnosed with their first primary invasive breast cancer between January 1st, 2001 and December 31st, 2006 were identified from the Pathology Information System (Co-Path). The initial date of entry reflects the year that HFHS implemented an institutional wide policy mandating routine assessment of the HER2 biomarker for all cases of invasive breast cancer. The other eligibility criteria were: 1) the initial diagnosis of breast cancer at HFHS; 2) AJCC TNM stages I-IV; and 3) insurance through the HMO plan. The first two eligibility requirements were imposed to minimize variation in the standards of the pathologic diagnostic criteria. The last criterion was imposed to reduce the potential confounding effect of socioeconomic status, i.e. limited or no access to healthcare, on the pathologic prognostic indicators of breast cancer. Data on

date of birth/age at diagnosis, and racial-ethnic identity were collected from electronic medical records.

Ghanaian: Samples were obtained through the archived resources of the Department of Pathology at the Komfo Anokye Teaching Hospital in Kumasi, Ghana. They reflected a convenience-based collection of invasive breast cancer specimens in women diagnosed 1998-2014. Patient age at diagnosis and gender was the only uniformly-available clinical feature on these cases, as provided by the Department of Pathology records.

Ethiopian: Samples were obtained through the archived resources of the Department of Pathology at the St. Paul's Hospital Millennium Medical College in Addis Ababa, Ethiopia. They reflected a convenience-based collection of invasive breast cancer specimens in women diagnosed 2001-2014. Patient age at diagnosis and gender was the only uniformly-available clinical feature on these cases, as provided by the Department of Pathology records.

Statistical Analysis

Parametric and non-parametric statistical techniques, as appropriate, were performed to discern the distribution of clinic pathologic variables between African-Americans and White-Americans populations. We then dichotomized women into two age groups, <50 years versus ≥50 years and conducted a sub-set analysis to discern the distribution of clinic pathologic variables between African-American and White-American women younger than age 50 years.

To discern the potential association between African heritage and the risk of triple negative breast cancer, we developed multivariable cumulative logistic regression models. We had data for three variables, age at diagnosis, breast cancer subtype and tumor grades. We first estimated the individual effect of each variable and their interactions with the outcome of interest, breast cancer subtype. Variables with a P-value of ≤0.10 were considered as the candidate variables. Interactions between variables also were tested at P-value ≤0.05. The final model included three candidate variables, age at diagnosis, tumor grade and ethnic heritage. All statistical tests were two-sided and analyses were performed using SAS vs. 9.1 (SAS Institute, Cary, NC).

Results

Tables 1 and 2 demonstrate the clinic pathologic features for the 272 African American compared to 321 White American patients from the Henry Ford Health System, revealing statistically significant higher frequencies of high-grade, estrogen receptor-negative and triple negative tumors among the African American cases; frequency of HER2/neu-overexpressing tumors was similar between these two groups. The African American patients had a numerically but statistically non-significant younger mean age at diagnosis compared to the White American patients (60 versus 62; $p<0.09$). The AA patients had a similar

stage distribution compared to the WA patients overall and in the subset analysis of patients younger than age 50 years.

As shown in Table 3, both of the American population subsets had older mean ages at diagnosis compared to the 234 Ghanaian patients (mean age 49 years) and the 94 Ethiopian patients (mean age 43 years; $p<0.001$). High-grade pathology was significantly more common among the tumors of the African American, Ghanaian, and Ethiopian patients compared to those of the White American patients (50.4%; 53.8%; and 53.6% compared to 33.7%, respectively; $p<0.001$).

Figure 1 demonstrates results of the molecular marker and phenotype comparisons. Frequency of ER-negative and triple negative tumors was highest in the Ghanaians (67.5% and 53.2% respectively), lowest in the White Americans (19.8% and 15.5%, respectively), and intermediate in the African Americans (37.1% and 29.8%, respectively). Frequency of these phenotypes in the Ethiopian cases was similar to the White American cases (28.7% and 15.0%, respectively). Differences in the distribution of these phenotypes across all four population subsets were statistically significant

Table 3-1: Clinic pathologic features of African American and White American patients from the Henry Ford Health System

Variable	African Americans N=272 (%)	White Americans N=321 (%)	P-Value
Estrogen Receptor Status			
Positive	171 (62.9)	256 (80.2)	<.0001
Negative	101 (37.1)	63 (19.8)	
Missing	0	2	
Progesterone Receptor Status			
Positive	158 (58.1)	236 (74.2)	<.001
Negative	114 (41.9)	82 (25.8)	
Missing	0	3	
HER2 Biomarker Status			
Positive	51 (18.7)	53 (16.7)	.5088
Negative	221 (81.2)	265 (83.3)	
Missing	0	3	
Histologic Grade			
1	33 (12.3)	79 (24.9)	<.0001
2	100 (37.3)	131 (41.3)	
3	135 (50.4)	107 (33.7)	
Missing	4	4	
TNM STAGE			0.4986
I	168	189	
II	73	92	

BC= breast cancer; Dx= diagnosis

(p=0.0001 for ER; p<0.0001 for TNBC). Frequency of HER2/neu-overexpressing tumors was low in the White American, African American, and Ghanaian cases (16.7%, 18.7% and 20.1%, respectively) but significantly higher in the Ethiopian cases (33.3%; p=0.0048). As shown by Figure 1b, subset analysis for the patients younger than age 50 years revealed patterns similar to those seen for patients of all ages. Frequency of triple negative tumors was highest in the Ghanaian patients (50.8%), lowest in the White American and Ethiopian patients (15.9% in each), and intermediate in the African Americans (34.3%; p=<0.0001). Among patients younger than 50 years, the frequency of HER2/*neu*-overexpressing cancers was similar for all four subsets.

.

Table 3-2: Clinic pathologic features of African American and White American patients from the Henry Ford Health System, patients younger than 50 years.

Variable	African-Americans N=67 (%)	White-Americans N= 67 (%)	P-Value
Estrogen Receptor Status			
Positive	38 (56.7)	50 (74.9)	0.0142
Negative	29 (43.3)	15 (23.1)	
Missing	0	2	
Progesterone Receptor Status			
Positive	36 (53.7)	50 (74.9)	0.0034
Negative	31 (46.3)	14 (21.5)	
Missing	0	3	
HER2 Biomarker Status			
Positive	15 (22.4)	13 (19.7)	0.7046
Negative	52 (77.6)	53 (80.3)	
Missing	0	1	
Histologic Grade			
1	4 (6.0)	12 (17.9)	0.0008
2	17 (25.4)	28 (41.8)	
3	46 (68.6)	27 (40.3)	
Missing	0	0	
TNM STAGE			

Table 3-3: Clinic pathologic features of African American, White American, Ghanaian and Ethiopian breast cancer cases

	Median Age (range)	**# High-Grade Pathology (%)**
White American N= 321	62 yrs (31-91)	107 (33.7%)
African American N= 272	60 yrs (27-87)	135 (50.4%)
Ghanaian N= 234	49 yrs (24-92)	84 (53.8%)
Ethiopian N= 94	43 yrs (23-76)	44 (53.6%)
p Value	**<0.001**	**<0.001**

yrs.= years

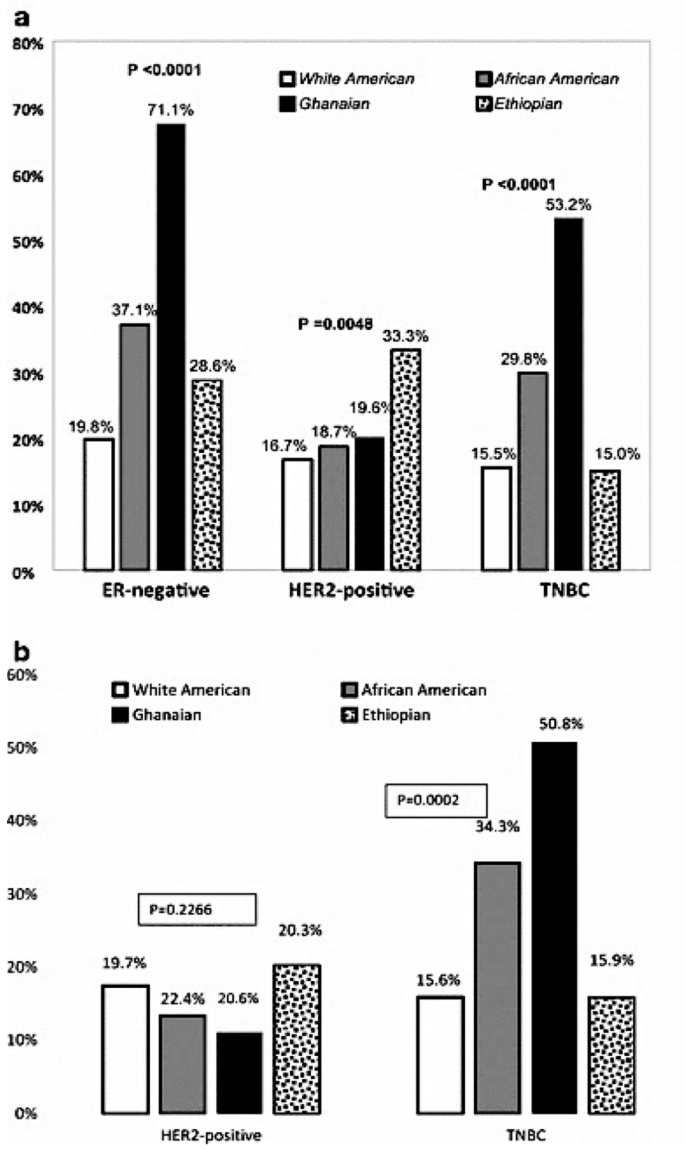


Figure 3-1: A comparison of breast cancer phenotypes across different ethnicities: **a** Comparison of phenotypes between African American, White American, Ghanaian, and Ethiopian breast cancer patients. **b** Comparison of phenotypes between African American, White American, Ghanaian, and Ethiopian breast cancer patients younger than 50 years

Table 3-4: Comparisons of phenotypes between couplets of population subsets.

	HR+ HER2-	**HR+ HER2+**	**HR- HER2+**	**TNBC**	**P-Value**
Ethiopians vs. White Americans	53.6% vs. 63.5%	20.3% vs. 17.5%	10.1% vs. 3.2%	15.9% vs. 15.9%	**0.77**
Ethiopians vs. African Americans	53.6% vs. 43.3%	20.3% vs. 13.4%	10.1% vs. 9.0%	15.9% vs. 34.3%	**0.02**
Ethiopians vs. Ghanaians	53.6% vs. 28.6%	20.3% vs. 11.1%	10.1% vs. 9.5%	15.9% vs. 50.8%	**<0.0001**
African Americans vs. White Americans	43.3% vs. 63.5%	13.4% vs. 17.5%	9.0% vs. 3.2%	34.3% vs. 15.9%	**0.015**
African Americans vs. Ghanaians	43.3% vs. 28.6%	13.4% vs. 11.1%	9.0% vs. 9.5%	34.3% vs. 50.8%	**0.012**
White Americans vs. Ghanaians	63.5% vs. 28.6%	17.5% vs. 11.1%	3.2% vs. 9.5%	15.9% vs. 50.8%	**<0.0001**

Table 4 demonstrates phenotype distributions for the four different patient populations as couplet comparisons, revealing that the Ethiopians and White Americans were the most similar (p=0.77). Statistically significant differences were observed in comparisons of all other couplets, largely driven by the highest frequencies of triple negative tumors in the Ghanaian patients and moderately high frequencies in the African Americans.

Expression of BRCA1 and BRCA2 protein in Ghanaian TNBC patients.

With the confirmation that the Ghanaian patients had the highest proportion of TNBC, we sought to determine if this observation is related in any way to BRCA gene expression. We performed an initial screening on FFPE samples from 72 invasive Ghanaian TNBC for the expression of the BRCA1 and BRCA2 protein by immunohistochemistry. The purpose of this evaluation was to score positive or negative BRCA1 and 2 status in the slides provided (Figure 2). Since BRCA1 and 2 are nuclear proteins, and the current literature scores based on nuclear localization, samples were scored as positive for immunolabeling for either marker when nuclear localization was observed. Cytoplasmic staining was considered background, and not taken into consideration of positive status. BRCA1 staining was generally very specific, with little cytoplasmic staining or other background. However, BRCA2 staining was often cytoplasmic and granular, and thus background staining was observed in many sections. However, several specimens were completely negative for each marker, with no background staining.

Fig.2 shows the expression of the BRCA1 and BRCA2 proteins. In 28 of the tumors (39%) there was no expression of either BRCA1 or BRCA2. 20 tumors (28%) expressed both proteins. 13 tumors (18%) expressed the BRCA1 and not BRCA2

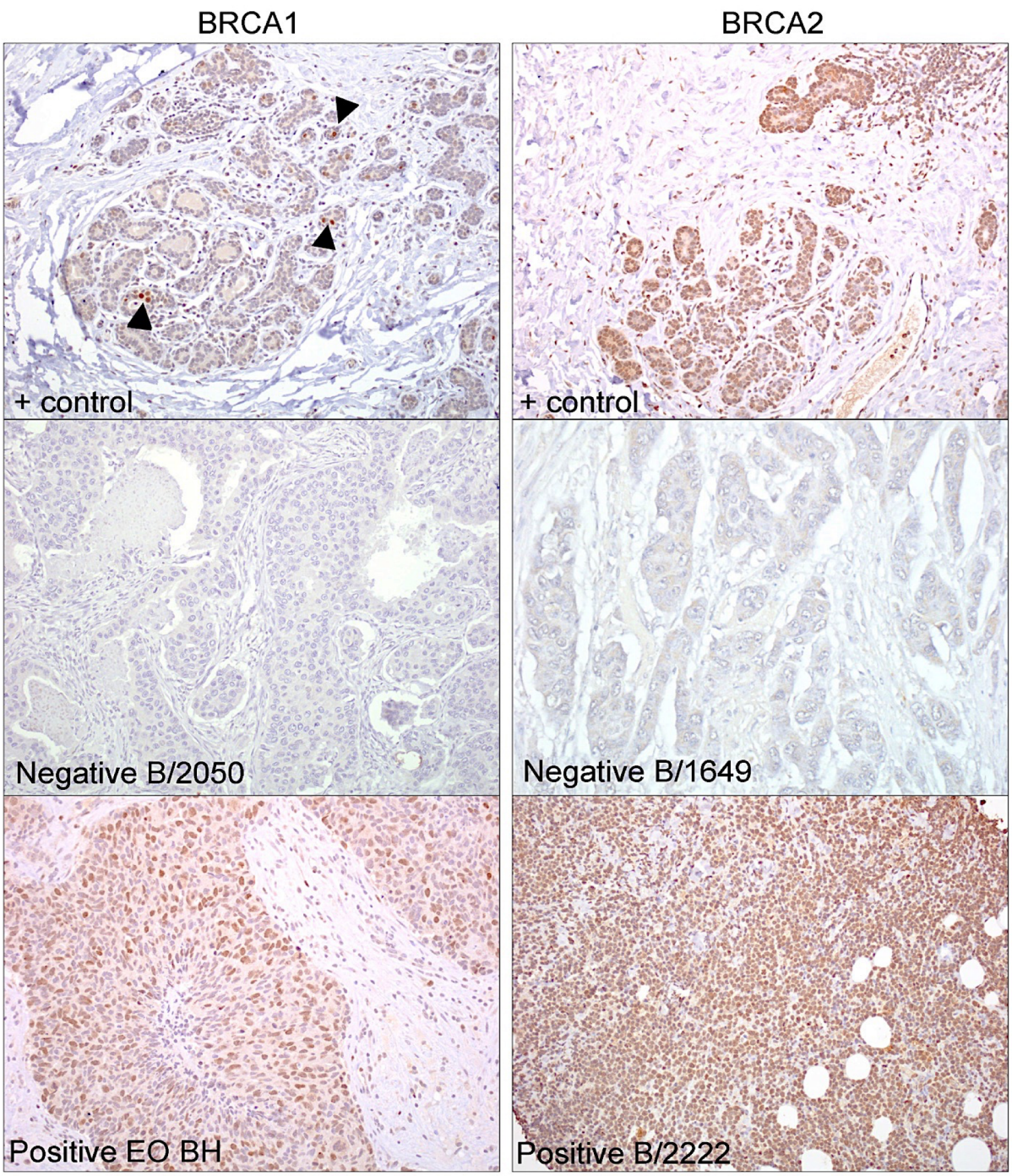


Figure 3-2: BRCA immunostaining

BRCA1 (left column) and BRCA2 (right column) immunolabeling in human breast cancer samples. Control tissues (top) show focal nuclear immunoreactivity to BRCA1 (left, arrowheads) and diffuse nuclear immunoreactivity to BRCA2 (right). Samples B/2050 and B/1649 show diffuse negative immunoreactivity to antibodies for BRCA1 (left) and BRCA2 (right), respectively. Sample EO BH showed multifocal positive nuclear immunoreactivity for BRCA1 (left), and sample B/2222 showed diffuse strong positive nuclear immunoreactivity for BRCA2 (right).

whiles 5 tumors (7%) expressed the BRCA2 and not BRCA1. 46% of the tumors therefore did not express the BRCA1 protein with 57% not expressing the BRCA2 protein.

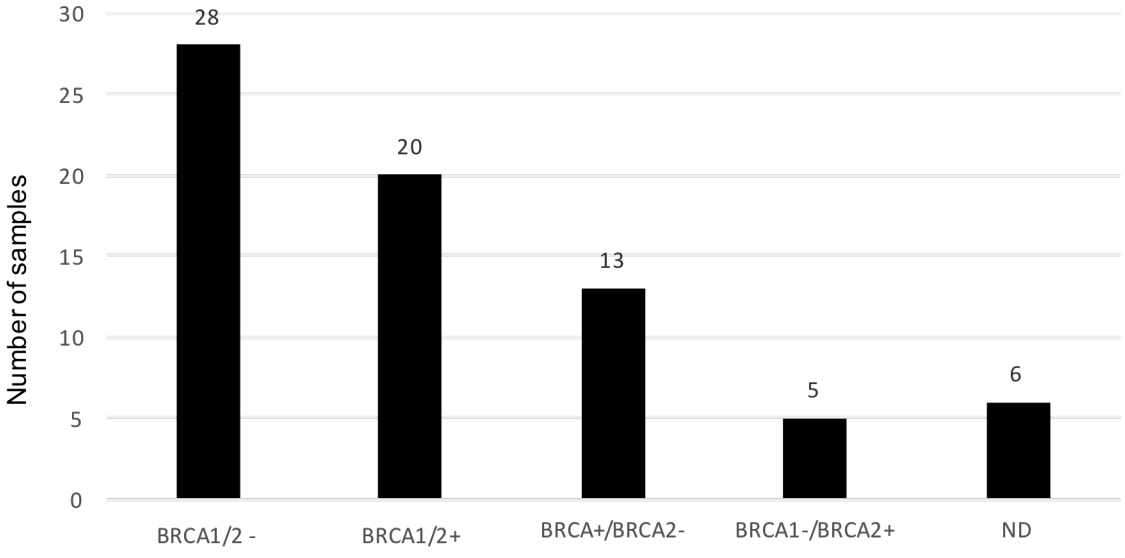


Figure 3-3: Results of BRCA1 and BRCA2 IHC staining in Ghanaian BC patients BRCA1/2-; tumor had no expression of both BRCA1 and BRCA2 proteins. BRCA1/2+; tumor expresses both BRCA1 and BRCA2 proteins. BRCA1+/BRCA2-; tumors express the BRCA1 and not BRCA2 protein. BRCA1-/BRCA2+: only BRCA2 protein is expressed. ND; non-diagnostic.

Several different genetic and epigenetic alterations including mutations, promoter methylation and post-translational alteration (silencing) may account for the loss of protein expression observed here. To have some understanding of the cause of the protein loss, we performed whole exome sequencing on 6 Ghanaian TNBC tumors and matched normal DNA extracted from the respective patient's saliva as described above. We used extensive bioinformatics analysis to

identify somatic variants and germline mutations in these tumors. Bioinformatics was performed by the Jun Li lab, UMICH.

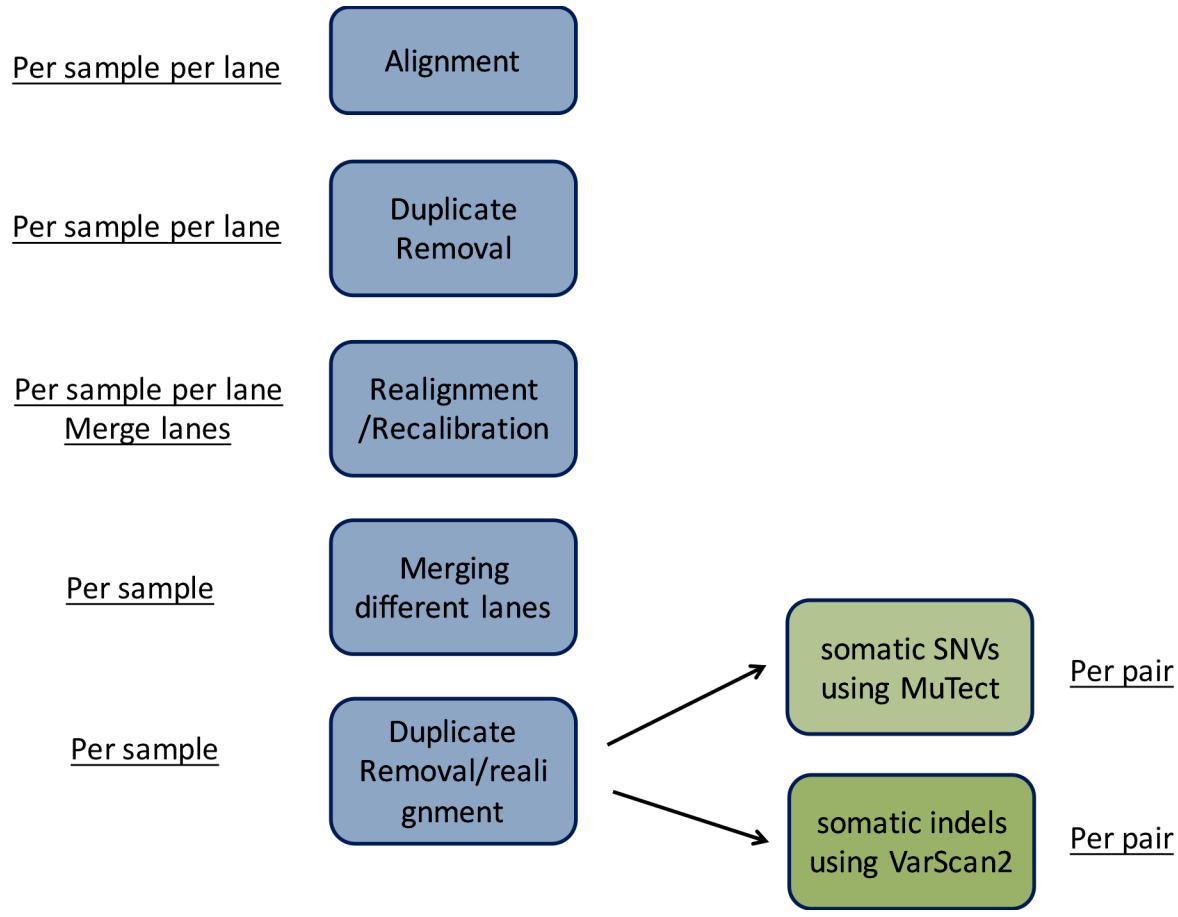


Figure 3-4: Pipeline for somatic variant calling.

There are 7 tumor/normal sample pairs from TNBC patients from Ghana, Africa. Normal DNA extracted from saliva specimens, and tumor DNA extracted from frozen tumor samples. 14 samples were exome sequenced using Illumina Hiseq2000 in 2 batches. First batch has sample.ID P1, P5, P6, and P9 and second batch has sample.ID P4410, P4430 and P4440. Each sample was

spread across 3 lanes. We used the following pipeline for DNA processing (Figure 1). The coverage varies between 27X to 69X and the average is ~50X. First batch have relatively lower coverage than the 2nd batch. After data processing, we used MuTect and VarScan2 to call the somatic variants. Table 5 summarizes the number of indels and SNVs identified in each patient tumor. We observed as few as 4 Indels in P4110 tumor and 8 in patient P4130.

Table 3-5: Somatic variants for 6 Ghanaian TNBC tumor/normal DNA pair

sample	P1	P5	P6	P9	P4110	P4130
SNV	803	224	10	90	121	57
Indel	59	27	58	45	4	8

Somatic SNVs : predicted by MuTect and filter by Allele Frequency in normal samples less than 0.05. somatic Indels : predicted by VarScan2 (p-value<0.01). Filter: Allele Frequency in normal samples less than 0.1 and alternative allele has less than 3 copies.

Recurrently Mutated Genes

After cross comparing the exonic variants in each sample, we have identified 35 recurrent mutated genes present in more than one sample, within 30 recurrent genes, like NOTCH1, TNN and TNF655 occurred in 2 samples and GMPS, EP400, TP53, PPP2R2B, and PKHD1L1 occurred 3 or 4 samples. Table 6 shows the list of genes that were recurrently mutated in the 6 samples.

.

Table 3-6: list of genes that are recurrently mutated in the 6 Ghanaian TNBC tumors

KIAA1211	ICA1	PKHD1L1
ABCD1	ITGAL	PLCH2
ARID1B	JADE1	PLXNA3
ATG2A	KIAA0922	PPP2R2B
CSPG4	MUC4	SCN11A
EP400	NLRC3	SLC35B2
EYS	NOTCH1	TP53
GMPS	NPAS4	TTN
GREB1	NRD1	ZNF655
HLA-DRB5	OTOF	

Germline mutation analysis was also performed on these samples with emphasis of the BRCA1 and BRCA2 and its pathway-associated genes on four of the patient sample pairs. (P1, P5, P6 and P9). In sample P1, germline analysis of exome data revealed a novel BRCA2 nonsense variant that would likely truncate the vast majority of the protein. Although this mutation has never been reported, it could be significant given the likely functional impact of the mutation. In sample P6, we identified 4 rare previously reported BRCA2 germline variants, however none were deemed pathogenic by ClinVar. In patient P9, there was one BRCA1 and two BRCA2 variants, however these were rare previously reported variants that were not deemed functionally or clinically relevant.

Discussion

A meta-analysis from ten years ago evaluated survival rates of more than thirteen thousand African American and seventy-five thousand White American breast cancer patients, and found that African American identity was associated with a statistically significant outcome disadvantage, even after adjusting for socioeconomic status (mortality hazard 1.28; 95% confidence interval 1.18-1.38)[4]. The importance of tumor biology in the complex disparities picture was highlighted in the 2015 Annual Report to the Nation on the Status of Cancer[5], demonstrating that population-based incidence rates of triple negative breast cancer are approximately two-fold higher for African Americans compared to all other population subsets, and this pattern is seen in all age categories. Population-based incidence rates of male breast cancer are also higher for African Americans compared to White Americans, and a study from the California Cancer Registry found a threefold higher frequency of triple negative tumors among African American men with breast cancer compared to White American men[6].

The American Cancer Society and the Surveillance Epidemiology and End Results Program have recently shown that breast cancer incidence rates have been increasing for African American women, to the point where they are now equal to the incidence rates in White American women[7]. Many of the advances made in systemic therapy for breast cancer have been in targeted therapies for endocrine-sensitive and HER2/neu-overexpressing tumors. Since the frequency

of triple negative tumors are higher in African American breast cancer patients, these treatment advances are less effective in this population subset. The disproportionate effectiveness of targeted therapies in African American breast cancer patients, coupled with the increasing incidence of breast cancer in African Americans has resulted a widening of the mortality disparity between African Americans and White Americans, which is now a 42% difference.

Against this background landscape of worsening breast cancer survival disparities related to racial-ethnic identity, this study evaluated the association between African ancestry and triple negative breast cancer by studying four different patient populations: African Americans; White Americans; Ghanaians (representing west Africa) and Ethiopians (representing east Africa). The two populations of African patients were significantly younger than the American patients, but this pattern is likely heavily influenced by the shorter overall life expectancy of individuals born in low and middle-income countries. We also found highest frequencies of ER-negative and triple negative tumors in African American and west African/Ghanaian breast cancer patients. White Americans and Ethiopians had similarly low frequencies for these phenotypes. Interestingly, frequency of HER2/*neu*-overexpressing tumors was elevated among Ethiopian breast cancer patients compared to the three other population subsets. These patterns persisted in subset analyses looking at patients younger than age 50 years.

Our study suggests that breast cancer patterns of African Americans are similar to those of west Africans, but differ from those of east Africans. Patterns of the Africa diaspora and forced population migration through the slave trade from several centuries ago may explain these observations[8-10]. The colonial-era trans-Atlantic slave trade, which was largely controlled by the Europeans, resulted in west Africans (including the ancestors of present-day Ghanaians) being brought to the Americas. In contrast, the slave trade from east Africa was controlled by Arab groups, and brought ancestors of contemporary Ethiopians to the mid-East and Asia. African Americans therefore are likely to have more shared ancestry with west Africans/Ghanaians compared to Ethiopians.

Confirming that our study cases are broadly representative of national and international populations was challenging, owing to the paucity of published data on breast cancer phenotypes in Africa. Table 5 summarizes comparisons of our study African American and White American cases to population-based data from the SEER Program[7], demonstrating that patterns of disease in the HFHS cases were indeed comparable to national statistics. The study Ghanaian cases were generally comparable to other reports of breast cancer in Ghana with regard to the young median age at diagnosis and the high frequency of ER-negative as well as triple negative tumors[11,12]. A note-worthy exception was the study by Adjei et al[13], comparing a relatively small number of Ghanaian breast tumors (n=51) to the tumors of Norway, with all immunohistochemistry performed at the Norwegian collaborating cancer facility. This study found a frequency of

triple negative tumors of only 22% among the Ghanaian cases, however of note this study also reported a notably low frequency of triple negative breast cancers in the comparison Norwegian population, at only 7%. The frequency of TNBC among Ghanaian cases therefore was threefold higher than the comparison Norwegian cases, a statistically significant difference (p=0.018). Very little data are available on molecular marker expression of breast cancers in Ethiopia, but work conducted by a German group has reported similarly low prevalence of ER-negative breast cancer in Ethiopian cases from Addis Ababa University (34.7%) compared to our St. Paul's Millennium Hospital study (28.7%)[14]. Interestingly, SEER-based data evaluating breast cancers among women in the United States diagnosed 1996-2008 with diverse African ancestral backgrounds confirms differences in frequency of ER-negative breast cancer related to heritage[15]. East African-born patients (186 patients, predominantly from Ethiopia or Eritrea) had ER-negative tumors in 22% of cases, compared to 33% ER-negative tumors among 143 west African-born patients (mostly from Nigeria). Frequency of ER-negative tumors among White Americans (18%) was similarly low as that observed among east African-born patients. In contrast, US-born African Americans had higher prevalence of ER-negative tumors (31.2%).

Table 3-7: Comparison of study patient populations with other reported studies of breast cancer in African American, White American, Ghanaian, and Ethiopian women

U.S. Breast Cancer Patients		**Average Age at Dx**	**ER-Negative**	**TNBC**
African American	Present Study, HFHS;	60 yrs	37.1%	29.8%
	ACS/SEER[7, 15]; n=19,734	59 yrs[15]; 58 yrs[7]	38.9%[15]	22%[7]
White American	Present Study, HFHS; n=321	62 yrs	19.8%	15.5%
	ACS/SEER[7, 15]; n=154,222	64 yrs[15]; 62 yrs[7]	21.4%[15]	11%[7]

Ghanaian Breast Cancer Patients	**Average Age at Dx**	**ER-Negative**	**TNBC**
Present Study, *KATH; n= 234*	49 yrs	68%	53%
KATH, Kumasi *Ohene-Yeboah[11], 2012; n=*	49 yrs	53%	43%
KATH, Kumasi *Adjei[13], 2014; n= 51*	51 yrs	24%	22%
KBTH, Accra *Der[12], 2015; n= 223*	52 yrs	82%	58%

Ethiopian Breast Cancer Patients	**Average Age at Dx**	**ER-Negative**
Present Study, SPMH; n= 94	43 yrs	28.7%
Addis Ababa, *Kantelhardt[14], 2014; n= 352*	43 yrs (ER-positive) 40 yrs (ER-negative)	34.7%

ACS/SEER= American Cancer Society/Surveillance Epidemiology and End Results Program; Dx= diagnosis; ER= estrogen receptor; TNBC= triple negative breast cancer; yrs= years.

Approximately 7% of breast cancer cases in the general population are estimated to harbor mutations in the BRCA gene; these women are found primarily in families characterized by multiple cases of the early onset of breast cancer[16]. The incidence of BRCA mutation is however very high in TNBC patients (30.8% needs reference).

BRCA mutation prevalence differs as well by ethnicity and race: AA (20.4%), Ashkenazi Jewish (50%), Asian (28.5%), Caucasian (33.3%), and Hispanic (20%).[17]

In patients with BRCA1 mutation, up to 70% of the BC are triple negative. There is a very high overlap in the morphologic and phenotypic characteristics of sporadic TNBC and BRCA1 associated BC and BC occurring in women with African ancestry; younger age distribution, high grade lack of expression of the estrogen receptor. These strong resemblances have been said to possibly suggest a commonality in one or more defects in the functions of the BRCA1 pathway for both *BRCA1*-associated and sporadic TNBC[18]. Identification of TNBC patients with BRCA1 mutation has helped in identifying family members who are at an increased risk for developing breast cancer and also for planning treatment options. In western Sub-Saharan Africa, where there is the reported highest incidence of TNBC, very little is known about the prevalence of BRCA mutations. Here we report an unusually high incidence of loss of expression of BRCA proteins; 57% of the tumors had loss of expression of both BRCA1 and BRCA2 with 39% having loss of expression of BRCA1, though the exact cause of

loss of protein expression is not fully understood, our sequencing data shows a high incidence of BRCA variants in the Ghanaian TBNC tumors with the identification of novel mutations that would likely lead to truncation of the protein. This preliminary data suggests the possibility that women in Western sub-Saharan Africa may harbor a very high incidence of BC susceptibility genes.

Clearly, more research is warranted in the study of breast cancer related to African ancestry. Our study suggests that west African ancestry is associated with inherited susceptibility for hormone receptor negative and triple negative breast cancer. It also suggests the possibility of a very high incidence of BRCA gene alterations in this population. Future work should strive to document population-based data on breast cancer incidence and mortality in different areas of Africa and also study the incidence of BRCA mutations in African TNBC patients. This research and its relevance to breast cancer disparities in the United States can be refined by utilizing germline genotyping studies and Ancestry Informative Markers to quantify extent of east versus west African ancestry in African American breast cancer patients[19]. Furthermore, gene expression studies of tumors in breast cancer patients with African ancestry are needed so that we can define the extent to which TNBC subtypes are similar or different across diverse populations[20].

This type of research holds great promise with regard to the critically important and relevant goals of understanding disparities in breast cancer burden

related to racial-ethnic identity domestically; investing in the oncology care systems of LMIC; and identifying additional markers of hereditary susceptibility for triple negative breast cancer.

References

1. Wang J, Sang D, Xu B, et al. Value of Breast Cancer Molecular Subtypes and Ki67 Expression for the Prediction of Efficacy and Prognosis of Neoadjuvant Chemotherapy in a Chinese Population. *Medicine.* 2016;95(18):e3518.

2. Hammond ME, Hayes DF, Wolff AC, Mangu PB, Temin S. American society of clinical oncology/college of American pathologist's guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. *J Oncol Pract.* 2010;6(4):195-197.

3. Wolff AC, Hammond ME, Hicks DG, et al. Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. *J Clin Oncol.* 2013;31(31):3997-4013.

4. 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.* 2006;24(9):1342-1349.

5. Kohler BA, Sherman RL, Howlader N, et al. Annual Report to the Nation on the Status of Cancer, 1975-2011, Featuring Incidence of Breast Cancer Subtypes by Race/Ethnicity, Poverty, and State. *J Natl Cancer Inst.* 2015;107(6):djv048.

6. Chavez-Macgregor M, Clarke CA, Lichtensztajn D, Hortobagyi GN, Giordano SH. Male breast cancer according to tumor subtype and race: a population-based study. *Cancer.* 2013;119(9):1611-1617.

7. DeSantis CE, Fedewa SA, Goding Sauer A, Kramer JL, Smith RA, Jemal A. Breast cancer statistics, 2015: Convergence of incidence rates between black and white women. *CA Cancer J Clin.* 2015.

8. Campbell MC, Hirbo JB, Townsend JP, Tishkoff SA. The peopling of the African continent and the diaspora into the new world. *Curr Opin Genet Dev.* 2014;29:120-132.

9. Salas A, Carracedo A, Richards M, Macaulay V. Charting the ancestry of African Americans. *Am J Hum Genet.* 2005;77(4):676-680.

10. Bryc K, Auton A, Nelson MR, et al. Genome-wide patterns of population structure and admixture in West Africans and African Americans. *Proc Natl Acad Sci U S A.* 2010;107(2):786-791.

11. Ohene-Yeboah M, Adjei E. Breast cancer in Kumasi, Ghana. *Ghana Med J.* 2012;46(1):8-13.

12. Der EM, Gyasi RK, Tettey Y, et al. Triple-Negative Breast Cancer in Ghanaian Women: The Korle Bu Teaching Hospital Experience. *Breast J.* 2015;21(6):627-633.

13. Adjei EK, Owusu-Afriyie O, Awuah B, Stalsberg H. Hormone receptors and Her2 expression in breast cancer in sub-Saharan Africa. A comparative study of biopsies from Ghana and Norway. *Breast J.* 2014;20(3):308-311.

14. Kantelhardt EJ, Mathewos A, Aynalem A, et al. The prevalence of estrogen receptor-negative breast cancer in Ethiopia. *BMC Cancer.* 2014;14:895.

15. Jemal A, Fedewa SA. Is the prevalence of ER-negative breast cancer in the US higher among Africa-born than US-born black women? *Breast Cancer Res Treat.* 2012;135(3):867-873.

16. Claus EB, Schildkraut JM, Thompson WD, Risch NJ. The genetic attributable risk of breast and ovarian cancer. *Cancer.* 1996;77(11):2318-2324.

17. Greenup R, Buchanan A, Lorizio W, et al. Prevalence of BRCA mutations among women with triple-negative breast cancer (TNBC) in a genetic counseling cohort. *Annals of surgical oncology.* 2013;20(10):3254-3258.

18. Moran MS. Should triple-negative breast cancer (TNBC) subtype affect local-regional therapy decision making? *Am Soc Clin Oncol Educ Book.* 2014:e32-36.

19. Palmer JR, Ruiz-Narvaez EA, Rotimi CN, et al. Genetic susceptibility loci for subtypes of breast cancer in an African American population. *Cancer Epidemiol Biomarkers Prev.* 2013;22(1):127-134.

20. Lehmann BD, Bauer JA, Chen X, et al. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J Clin Invest.* 2011;121(7):2750-2767.

Chapter 4 : ALDH Expression Characterizes Tumor Initiating Stem Cells in Triple Negative Breast Cancers across Different Ethnicities.

Abstract

Triple negative breast cancer (TNBC), one of the most aggressive breast cancer subtypes, carries a guarded prognosis and is the only subtype for which there are no approved targeted therapies. Chemotherapy is the mainstay of treatment both in the neo- and adjuvant settings and for metastatic disease. Improving patient outcomes requires a better understanding of the genotype and phenotypes of cells that mediate aggressiveness and metastases.

In the US, the incidence of TNBC is highest in women with African ancestry (AA); in western sub-Saharan Africa, single-institution studies show that TNBC constitutes 40- 80% of all breast cancers. Given the Caucasian/AA survival disparity in breast cancer, there is an urgent need to find actionable targets in TNBC of all ethnicities, but especially in TNBC in AA, which are suspected to be more aggressive. There is the urgent need to better understand the biology of TNBC in order to develop effective therapies targeting genes and pathways responsible for metastatic behaviors.

Breast cancer stem cells, the small population of cells in breast tumors that have been shown to mediate breast tumor initiation, metastasis and

resistance to conventional therapy have also been reported to mediate the heterogeneity of TNBC and are especially abundant in TNBC in women with African ancestry.

In this study, we sought to find differentially expressed and activated genes and pathways in the stem cell populations of patient derived TNBC xenografts from Ghanaian, African American, and Caucasian patients. By these analyses we identified that the ALDH1 expressing cancer stem cells of these tumors segregate apart from bulk populations and CD24/44 stem cells, and were enriched with genes and pathways that are involved in tumor metastasis including the WNT, MAPK and TGF-beta pathways. We also identified novel genes that are significantly up-regulated in this population that may serve as potential targets in further studies.

Introduction

Triple negative breast cancer (TNBC), defined by having less than one per cent expression of the estrogen and progesterone receptor,[1] and no overexpression of the HER2 oncogene;[2],[3] carries a guarded prognosis. TNBC has a 5-year overall survival of 64% compared to 81% in non-TNBC; TNBC has the highest rate of visceral and distant relapse, and, importantly, the shortest survival after the onset of clinical metastasis[4-6]. It appears that the significant gains in survival from breast cancer achieved in the last 2 decades have largely bypassed TNBC patients. As there are no approved targeted therapies for TNBC, chemotherapy is the mainstay of treatment in either the neoadjuvant setting or after loco-regional therapy with surgery and radiotherapy [3,7] and for metastatic disease outside of clinical trials. Patients with TNBC have a higher rate of complete pathological response after neoadjuvant chemotherapy (22% v 11%; P = .034) as compared to non-TNBC[5,8]. However, for those 80% of TNBC patients with partial pathologic response, their survival is much worse than for patients with non TNBC with similar response. This apparent paradox is due to TNBC patients with residual disease having a much rapid relapse and much lower survival once metastases are detected[4,6]. The molecular complexity and extreme heterogeneity of TNBC makes it difficult to define and credential molecular targets. Several attempts have been made to classify TNBC into subtypes, with the aim of describing prognostic categories[9-12]; however, these sub-classifications have not yet influenced clinical management.

An abundance of evidence supports the existence in breast cancer of tumor initiating cells with self-renewing potential, termed breast cancer stem cells (BCSC), which play a major role in tumor recurrence and metastasis[13,14]. These cells have been shown to be especially abundant in TNBC[15],[16,17], are key mediators of resistance to conventional chemotherapy and radiotherapy,[18-20] and contribute to the heterogeneity of TNBC [21]. The BCSC populations are themselves diverse and can be isolated from the tumor mass using fluorescent activated cell sorting (FACS) when they are CD44+/CD24-/EPCAM+[22,23]. Furthermore, using the ALDEFLUOR assay, stem-like cells can also be isolated by measuring the activity of the intracellular aldehyde dehydrogenase enzyme in intact breast cancer cells[15]. ALDH+ BCSC have been shown to mediate metastasis and poor clinical outcomes in breast cancer[24]. While the existence of phenotypic BCSC is supported by a large body of preclinical and clinical evidence, the pathways that contribute to the maintenance of this cell population have not yet been clearly defined in patient tumors.[25]

Another striking characteristic of the epidemiology of TNBC is its high prevalence in women with African ancestry[26] compared to Caucasian women. While African American women have a lower incidence of breast cancer than European American women in the USA, multiple studies have reported a much higher proportion of TNBC as a component of all breast cancers in sub-Saharan African women with pure African ancestry. The proportions of TNBC in all breast

cancers by race are 10-21%, 21-46% and up to 82% for Caucasians, African Americans (AA), and sub-Saharan Africans respectively[26-31].

The high incidence of TNBC in women with African ancestry has been suspected as one of the potential determinants for the higher mortality observed in African American women with breast cancer, which has been maintained over the last 3 decades. Importantly, previous reports specifically have associated expression of ALDH1, the main protein associated with the ALDEFLUOR assay activity, with TNBC and aggressive breast cancers in African women[17]. Indeed, in our previous work, we showed a significantly high expression of ALDH1 in breast lesions (both benign and malignant) of Ghanaian women. We also showed a higher expression of ALDH1 in TNBC as compared to other breast cancer subtypes[16]in this native African population. It is important to note that AA in the United States represent varying degrees of genetic admixture from local as well as several ancestral populations, including “true African” ancestry originating largely from western Sub-Saharan African through forced migration from the trans-Atlantic slave trade. Thus, it is possible that African ancestry might be associated with a heritable risk for developing TNBC and studies of pathway alterations in ancestrally pure populations compared to admixed populations may offer a new understanding of the biology of this disease.

We have therefore hypothesized that the key to unraveling the complexity of TNBC and thus finding new targets for its treatment lies in understanding signaling integration and gene expression in the BCSCs of TNBC in women with

African ancestry in comparison with other ethnicities, looking both for common changes and unusual new genes or pathways that may be altered to render this subtype quite different from the other breast cancer subtypes.

So far, most of the data on the biology of TNBC was generated using cell lines developed mainly from Caucasian patient tumors, which, by virtue of their prolonged passage in 2D cultures, may not serve as a true representation of the heterogeneous nature of TNBC. This study utilizes sorted BCSC populations from unique patient materials to address the gap in knowledge about genetic and pathway determinants of the phenotype that is TNBC. In this article, we will also discuss our unique approach to finding novel targets in the BCSC of TNBC using PDX developed from women with different ancestries including patients from Ghana.

Methods

Patient Derived Xenograft Generation

The PDXs that were used for this study were generated at three locations: The University of Michigan (UM), Van Andel Institute (Grand Rapids, MI) and Baylor College of Medicine (Houston, TX). For the PDXs generated at UM, the samples were obtained from patients receiving treatment either at the Komfo Anokye Teaching Hospital (KATH) in Kumasi Ghana or at the UM Hospital. Institutional Review Board approval was obtained at both institutions and patient

informed consent was also obtained. At KATH, excess patient tumor tissue was taken from the operating room, cut into small pieces and slow frozen at -80oC in 10% DMSO/ 90%FBS. The samples were then transported overnight on dry ice in a commercial flight to UM. For samples obtained from patients, undergoing breast cancer treatment at the UM, excess patient biopsy sample was obtained. For tumor implantation, the tumor pieces were rapidly thawed and rinsed off with HBSS or PBS, cut into smaller fragments (approximately 2 mm in diameter), mixed with matrigel (BD), and injected into the mammary fat pad of female NOD/scid/IL2Rγ^{null} (NSG) mice with an 18-gauge needle, 25 µl of matrigel per injection site. Xenograft tumors were taken out when they reached a size of 1 to 1.5 cm. PDXs generated at Van Andel Institute and Baylor College of Medicine were re-implanted in NSG mice.

Tissue Dissociation and Flow Cytometry

Xenograft tumors were digested using the Tumor Dissociation Kit (Miltenyi # 130-095-929) and the gentleMACS Octo Dissociator with the Human Tissue dissociation protocol, running the “gentleMACS Octo Dissociator program 37C_h_TDK_1 (Miltenyi). Mouse cells were removed using the Mouse Cell Depletion Kit (Miltenyi # 130-104-694) and MultiMACS Cell24 Separator Plus (Miltenyi). The cells were then stained with: LIVE/DEAD Near-IR Fixable Dead Cell Stain (Invitrogen, Catalog #L10119) anti-mouse MHC Class I (H-2Kd) eFluor 450 (eBiosciences, Catalog #48-5957-80), Aldefluor (Stem Cell Technology,

Catalog #01700), CD44-APC (Clone G44-26; BD), CD24-PECy7 (Clone ML5; Biolegend), and EpCAM-PE (Clone EBA-1; BD). Half of the stained cells were incubated with DEAB (Stem Cell Technology, Catalog #01700), an inhibitor of ALDH. An aliquot of cells was also stained with Live-Dead Near-IR, anti-mouse H2KD eFluor 450, and mouse IgG1 isotype antibodies individually conjugated to APC, PECy7, or PE (BD). These isotype-stained cells were used to set gates for CD44 APC CD24-PECy7 and EPCAM PE. Single color controls were used for compensation. The DEAB control was used to set the gate for Aldeflour+ cells. Cell sorting was done on a MoFlo Astrios flow cytometer.

Breast Cancer Stem Cells Whole Transcriptome Library Construction and Paired End NGS:

BCSC populations from PDX samples from 5 women each of various ethnicities (African American, Caucasian, Ghanaian) were FACS sorted to sub-select and collect ALDH+, CD24-CD44+EPCAM+ cells, and bulk cells. Total RNA was extracted from all 3 cell fractions using Single Cell RNA Purification kit (Norgen Biotek Corp., Ontario, Canada) following the manufacturer's protocol. Quality and quantity of total RNA was assessed using Agilent 2200 TapeStation Instrument (Agilent Technologies, Santa Clara, CA) and NanoDrop 1000 Spectrophotometer (Thermo Fisher Scientific, Waltham, MA), respectively. Whole Transcriptome libraries with an initial input of 10ng RNA were constructed using Nugen Ovation RNASeq V2 (NuGen, San Carlos, CA) for cDNA generation

and Kapa LTP Library preparation kit (Kapa Biosystems, Wilmington, MA). The libraries were then sequenced on the Illumina HiSeq platform (Ilumina, San Diego, CA) with 83 cycles x 7 cycles x 83 cycles read length.

Data Analysis

Raw sequencing data in the form of .bcl files were converted to .fastq files using bcl2fastq 1.8.4 (Ilumina, San Diego, CA) and aligned against the human reference genome GRCh 37.74 using STAR 2.3.1 Aligner with default parameters. (https://github.com/alexdobin/STAR). Aligned reads were assigned to genomic features (GRCh37 genes) and quantified by HTSeq-count 0.6.1. (http://www-huber.embl.de/users/anders/HTSeq/doc/overview.html). HtSeq-count generated the CPM values (counts per million) for each gene of each samples. EdgeR was used to identify differential gene expression with both fisher-exact test and generalized linear models (FDR<=0.05), with an input of CPM value with log2 transformation. Prior to the PCA analysis, six samples (CA:3, AA: 3) were removed due to a poor data quality. Log2 values of each patient's bulk CPM were used for TNBC subtyping (http://cbc.mc.vanderbilt.edu/tnbc/). Pathway analysis was performed with iPathwayGuide™ (Advaita Bioinformatics, Plymouth, MI) and LRpath (lrpath.ncibi.org).

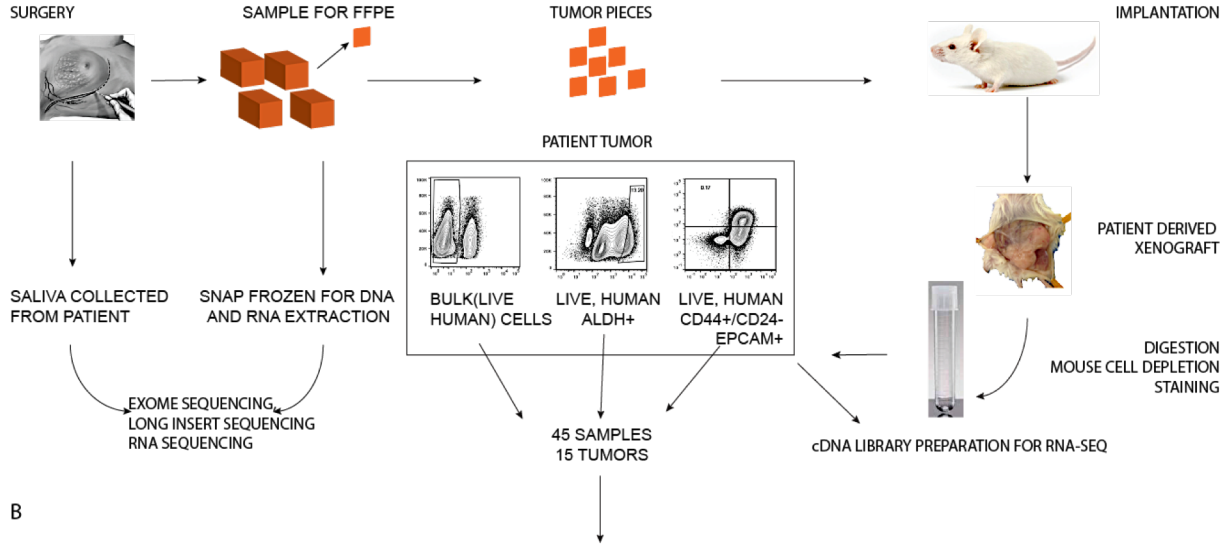
Tumorsphere Formation Assay

Cells obtained from FACS sorting of PDX samples were plated in a 96-ultralow attachment flat bottom plates (Corning) at 5,000 cells/well in 100ul of serum free media Serum-free media (SFM) contains: DMEM, 1X B27 (Gibco), 50 ug/mL insulin, 1X non-essential amino acids, 10ng/mL EGF, 10 ng/mL bFGF, and antibiotics. Tumorspheres formed after 3-5 days and each experiment was performed in triplicate

Results

Creating PDX based TNBC Models from Tumors of Women of Different Ancestries.

We succeeded in implanting cancerous tissues directly from patients into immune deficient mice, thus providing a renewable source of tissue for experiments. Patient derived xenograft (PDX) tumors have been shown to conserve the gene expression, histology, and architecture of the original primary tumor from which they are developed[32-34]. To study the gene expression in BCSC, which typically constitute less than 1% of the total tumor mass, we needed a large renewable source of breast tumors. We first explored the possibility of creating PDX from TNBC patients in sub-Saharan Africa (Ghana). To do this, our team of surgeons and oncologists traveled to Komfo



B

ETHNICITY	GHANAIAN					AFRICAN AMERICAN					CAUCASIAN				
SAMPLE	GUM-07	GUM-13	GUM-17	GUM-28	MUM-12	2147	VARI-006	3887	4664	4913	VARI-037	VARI-068	VARI-004	MUM-02	MC1
SUB-TYPE	TNBC	TNBC	TNBC	TNBC	TNBC	TNBC	TNBC	TNBC	TNBC	TNBC	TNBC	TNBC	TNBC	TNBC	TNBC
PATIENTAGE	74	47	50	46	54		50				58	53	55	27	
HISTOLOGY	IDC G2	IDC G3	IDC G3	IDC G3		IDC	PD IDC LN	CR IDC	IDC	IDC	IDC G3	IDC G3	IDC G3	IDC G3	

Figure 4-1: A. processing and stem cell characterization from the patient derived xenograft

A. Tumor biopsy collected directly from patients were used for the experiments; biopsy samples were divided for different experimental procedures. For each biopsy sample, a part was slow frozen and later implanted into NSG mice as described in the methods section. Tumors were harvested once they reached to 1-1.5cm in diameter. Freshly harvested PDX tumors cut into small pieces were enzymatically digested and processed into single cells. Tumor cells isolated after the removal of mouse cells were stained and flow sorted. Three cell populations; BULK (Live, Human) ALDH+(Live, Human, ALDH+) and CD44+/CD24-(Live, Human, CD44+/CD24-/EPCAM+) were sorted for. Each sort was done for three biological replicates of the tumor all at the same passage. A total of 15 tumors including 5 African American, 5 Ghanaian and 5 Caucasian were processed and analyzed as described above

B: Clinical and pathological data of the patient and the primary tumor from which the PDX were developed respectively.
GUM refers to tumors that were obtained from Ghana, MUM is for tumors obtained from UM patients, VARI for PDXs from Van Andel Institute and the one with numbers only are from Baylor College of Medicine

Anokye Teaching Hospital in Ghana, Africa. After obtaining institutional review board approval and patient informed consent, we operated on BC patients and obtained samples of their tumors. Figure 4-1A. These tumors were frozen and transported to the UM as described. A piece of the tumor was at the same time fixed in formalin to create FFPE from which slides could be cut and stained for H&E as well as immunohistochemistry studies to confirm the marker composition of the tumor, as shown in Figure 4-1A. A piece was also snap frozen for RNA and DNA extraction at a later time. Saliva was also collected to have a matched normal genomic DNA for tumor/normal genomic comparisons. In the process of acquiring these tumors from Ghana, we provided infrastructural, educational, and technical support to the faculty, staff, and patients of the hospital[35].

At the UM, frozen primary tumor samples were orthotopically implanted in mice to create PDXs as described. We made 8 trips to Ghana over a period of three years and obtained tumor from 36 breast cancer patients, 16 of whom had TNBC. We implanted ten of the TNBC tumors in immunosuppressed mice and developed PDXs from five of them. We also established two PDXs from TNBC tumors obtained from patients undergoing surgery at the UM after obtaining informed consent. In addition, we collected early passage established PDXs from our collaborators in Van Andel and Baylor hospital. Figure 4-1B shows the patient ancestry, age, and histology of the primary tumors from which the PDXs were developed.

For this study, we used five Ghanaian, five African American, and five Caucasian PDXs. All primary patient samples were TNBC (TNBC status was

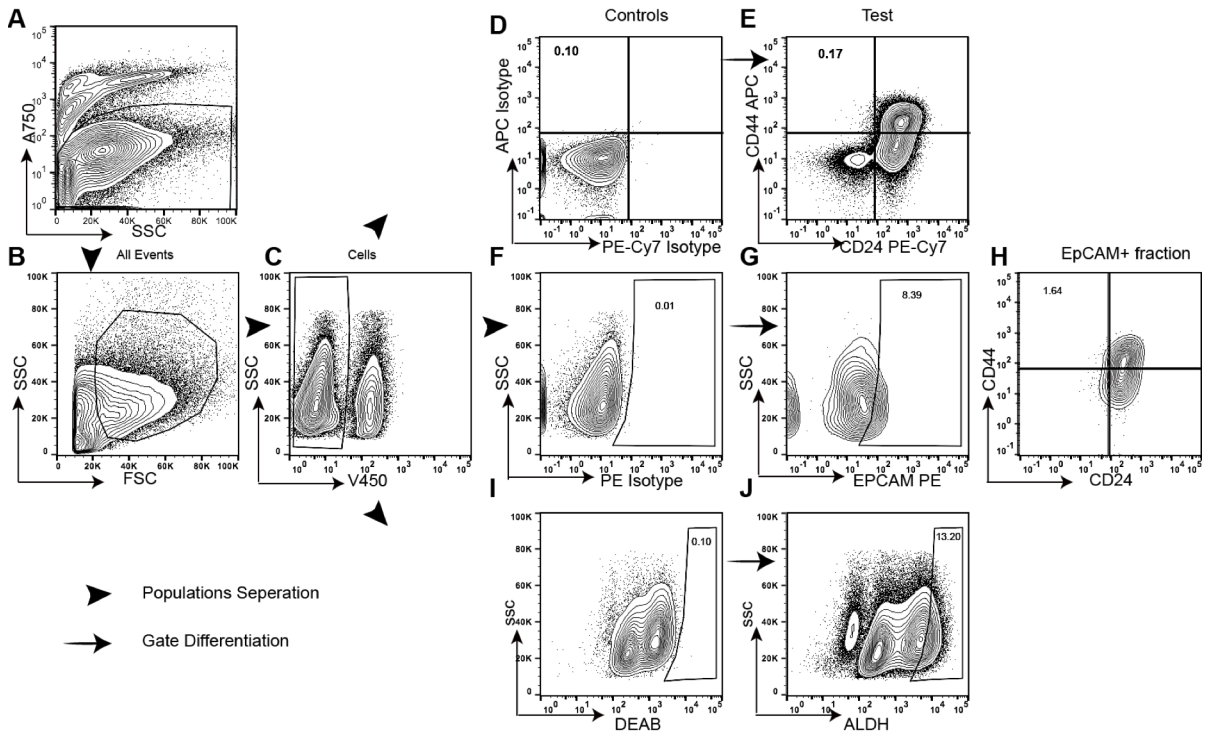


Figure 4-2: schematic for flow sorting for bulk and stem cell sub-populations
Single cells obtained from PDX was stained as described in the methods section. Sorting was first done to isolate live, Fig2 A and single cells, Fig B. Human cells were then gated of mice cells using Mouse H2KD in V450 channel, Fig C. Using cells stained with isotype controls, gates were set for APC (CD44) positive and PECy7 (CD24) negative at 0.1. Fig D and PE (EPCAM) at 0.1, Fig F. Using cells incubated in DEAB, the gate was set for ALDH+ cells at 0.10, Fig I. Of the total stained cells, Fig E shows the percentage that is CD44+/CD24-, 0.17. In the CD44+/ CD24- population, 0.39% were EPCAM positive, Fig G. In the total population of EPCAM+ cells, 1.64 percent are CD44+/CD24-, Fig H. In the total population of stained cells, 13.20 are ALDH+, J.

confirmed at

the UM, Van Andel or Baylor). All patients had high-grade invasive ductal carcinoma (IDC). Thirteen of the patient samples were from primary tumors and

two from metastatic locations in the axillary lymph nodes (LN) and chest wall metastasis.

Isolating Breast Cancer Stem Cells and Bulk Cell Sub-populations from Established PDX

The expression of the cell surface markers CD44+/CD24-(low) as well as the intra cellular enzyme ALDH have both been used as markers to identify BCSC and also demonstrate the classic BCSC phenotype of self-renewal and mammosphere formation.[15,23,36]. It was however noted that some tumors may not express either or both of these markers[22]. Using florescent activated cell sorting (FACS), we describe here a systematic procedure for simultaneously isolating the ALDH+, CD44+/CD24-/EPCAM+ and the bulk cells (which is the general fraction of live human tumor cells, not depleted of stem cells), in the established PDXs. The tumor was first processed into single cells and stained as described above. Single color controls for EPCAM PE, CD44 APC, CD24 PECy7, V45 H2KD, and A750 LIVE/DEAD were first used to set compensations for the FACS separation (Fig. 2). For each sample, two tubes of cell were stained as controls and one tube was stained for sorting. All cells were stained with the LIVE/DEAD reagent, figure 4.2A, and V450/ANTI MOUSE H2KD which is used to separate the human from mouse cells, figure 4.2C. During FACS sorting, a gate is set for fore scatter (FSC) vs. side Scatter (SSC) figure 4.2C. This is an arbitrary gate used to separate single viable cells from doublets and debris. The first tube of

control cells is stained with isotype controls for APC, PE, and PECy7 which are used to set the gates for CD44 (in APC channel), CD24 (in PECy7 channel) and EPCAM (in PE channel) each to 0.1 figure 4.2D and F, respectively.

In figure 4.2D, the upper left quadrant represents the CD44+/CD24- population. The second control cells, containing cells stained with ALDH, CD44APC, CD24 PECy7 in the presence of DEAD which blocks the action of the enzyme Aldehyde Dehydrogenase is used to set the gate for ALDH+ cells at 0.1%, figure 4.2I. With all gates set, cell sorting is done from the third tube with cells stained with LIVE/DEAD, ANTI-MOUSE H2KD, ALDH, CD44, CD24 and EPCAM. To obtain a CD44+/CD24-/EPCAM+ cell population, live, single, human cells were sorted in the APC vs PECy7 gate, E, (0.17% in this sample (GUM-13). The EPCAM+ cells are then sorted off the CD44+/CD24- fraction. As figure 4.2G denotes, in GUM13, 8.39% of all CD44+/CD24- cells were EPCAM+. The ALDH+ population was sorted of live, single human cells in the gate, pre-set with DEAB control, (13.20% in GUM13), figure 4.2J. A portion of the stained cells was sorted for the BULK population (these are LIVE, SINGLE and HUMAN). In all 15 PDXs, we were able to isolate both populations of stem cells.

Table 4-1: percentage of the different stem cell sub-populations in the 15 PDXs.

Sample name	CD44+/CD24- %	CD44+/CD24-/EPCAM+ %	ALDH+%
GUM-07	34.14	31.30	1.71
GUM-13	8.39	0.04	13.20
GUM-17	1.42	0.52	1.37
GUM-28	17.99	6.57	3.22
MUM-12	10.71	0.04	2.32
2147	0.28	0.11	0.36
VARI-006	0.21	0.00	0.36
4664	3.05	0.00	0.04
3887	0.42	0.08	0.18
4913	29.64	14.24	0.63
MUM-02	17.99	2.69	3.22
VARI-004	1.51	1.30	0.82
VARI-037	0.07	0.02	0.13
MC1	---	----	----
VARI-068	13.03	1.23	0.94

Table 4-1 shows the proportion of tumor cells represented by the different stem cell sub populations, in the 15 samples. Importantly, it was observed that in most of these tumors, only a small percentage of the CD44+/CD24- cells expressed EPCAM. These may represent the claudin-low tumor subtype which has a low expression of luminal markers and a high expression of mesenchymal markers[37].

ALDH+ Stem Cell Population has a Distinct Gene Expression Profile from CD44+/CD24- and BULK Population, in TNBC PDXs from all Ethnicities.

In order to better understand the biology of BCSCs in these ethnically diverse tumors, we performed RNA-seq on the isolated stem cell populations and compared the results to those obtained for the bulk cells. Total RNA was extracted from the isolated cells, (Table 4-2 shows the number of stem cells that was collected for each sample and the corresponding RNA integrity). Figure 4.3A summarizes the process we used for RNA sequencing and comprehensive bioinformatics analyses. We performed principal component analysis for 39 samples, excluding only six samples with poor data quality. Interestingly, the results for the subpopulations of each tumor clustered based on the tumor of origin, that is, the three samples from each tumor clustered together, as seen in Figure 4.3B. They did not initially cluster by cell population or by ethnicity.

Table 4-2: number of cells sorted and the respective RIN

Ethnicity	Sample	ALDH+ cell count	RIN	Sample	CD44+ cell count	RIN	Sample	Bulk cell count	RIN
Caucasian	MUM 02	330,000	9.2	MUM 02	972,000	5.7	MUM 02	400,000	9.6
	VAR 004	7,000	7	VAR 004	18,000	5.1	VAR 004	150,000	6.5
	037	266,000	9	037	16,400	7.3	037	4,000	8.4
	MCI	21,000	9.3	MCI	31,000	8.8	MCI	15,000	8.5
	VAR 068	1,070	8.5	VAR 068	2,030	8	VAR 068	1,019	7.8
African American	2147	50,000	9	2147	1,890	7.9	2147	20,000	10
	VARI 006	750	Conc too low	VARI 006	1,140	Conc too low	VARI 006	20,000	7.7
	4664	309	8.4	4664	124	7	4664	7,000	8.9
	3887	5,400	7.9	3887	5,000	7.6	3887	194,000	9
	4913	1,890	7.4	4913	95,000	6.4	4913	102,000	6.7
Ghanian	GUM 17	5,990	8.7	GUM 17	12,500	8.9	GUM 17	148,000	9.2
	MUM 12	31,000	5	MUM 12	94,000	7.7	MUM 12	100,000	7
	P3-2770	175	Conc too low	P3-2770	10,200	8.6	P3-2770	228	8.4
	GH-VB	>2000	8.4	GH-VB	>2000	8.6	GH-VB	>2000	8
	GUM 28	729,000	9	GUM 28	17,500	9.5	GUM 28	142,000	9.9

To better observe the effect of ethnicity or cell population, we performed mean centering of the triplet of samples from each tumor. Briefly, the average spectrum for each triplet was subtracted from the spectrum for each sample in the triplet. After mean centering,

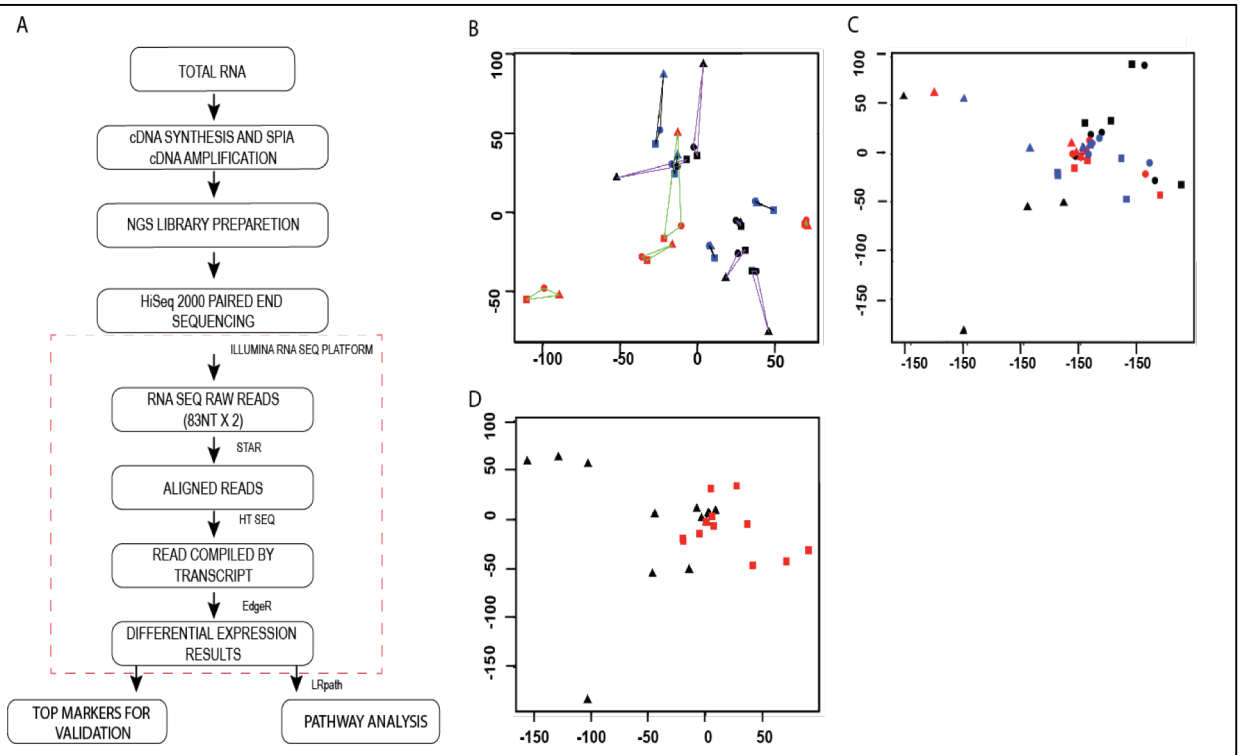


Figure 4-3: PCA plots of the global gene expression of 45 samples from 15 PDXs. Fig3A shows the pipeline for sample processing, sequencing and data analysis for the sorted samples. Prior to the PCA analysis, six samples (CA:3, AA: 3) were removed due to a poor data quality. Fig. 4B shows PC1 vs PC2 plot of gene expression for 39 sorted samples from 15 PDXs developed from women from three different ancestries. Fig. 4C shows the PC1 vs PC2 plot of the samples after mean-centering the three samples from each tumor. Fig 4D, PCA plot showing PCA1 vs PCA2 of only the ALDH+ and the CD44+/CD24- samples.

the ALDH+ samples separated out from the bulk and the CD44+/CD24-, regardless of the ethnicity. Surprisingly, there was no separation between the CD44+/CD24- and the bulk samples, figure 4-3C. To answer the question of whether the ALDH+ and the CD44+/CD24- cells are distinct sub-populations in these tumors, we performed PCA analysis of the ALDH+ and the CD44+/CD24- samples alone. Once again, ALDH+ samples separated out distinctly from the CD44+/CD24- samples figure 4-3D.

ALDH+ not CD44+/C24- Cells Exhibit Stem Cell Characteristics in Primary TNBCs.

Previous reports have indicated that ALDH expressing stem cell are important for the stem cell phenotype observed in TNBC especially in women with African ancestry[17]. Thus, with the ALDH+ population separating out, we compared their ability to form tumor spheres with that of the other two cell populations: bulk and CD44+/CD24-. The sphere formation assay, as described in the methods section, was performed with sorted stem and bulk cell sub-populations from three of the PDXs: VARI-068, VARI-004, and GUM-17, as seen in Fig 4. We observed that the ALDH+ cells formed a significantly higher number of spheres followed by the bulk cells, with the CD44+/CD24- cells forming the least number of spheres (figure 4-4) In the secondary passage, the ALDH+ cells maintained a significantly higher sphere formation over the bulk cells, with the CD44+/CD24- forming no spheres at the higher passage.

In Primary TNBC ALDH+ Stem Cells are Characterized by a Distinct Gene Expression Profile.

With the observation that the ALDH+ populations separate out from the bulk and CD44+/CD24- (which clustered together), and with the ALDH+ population forming significantly more spheres in 3D culture, we explored the genes and pathways that

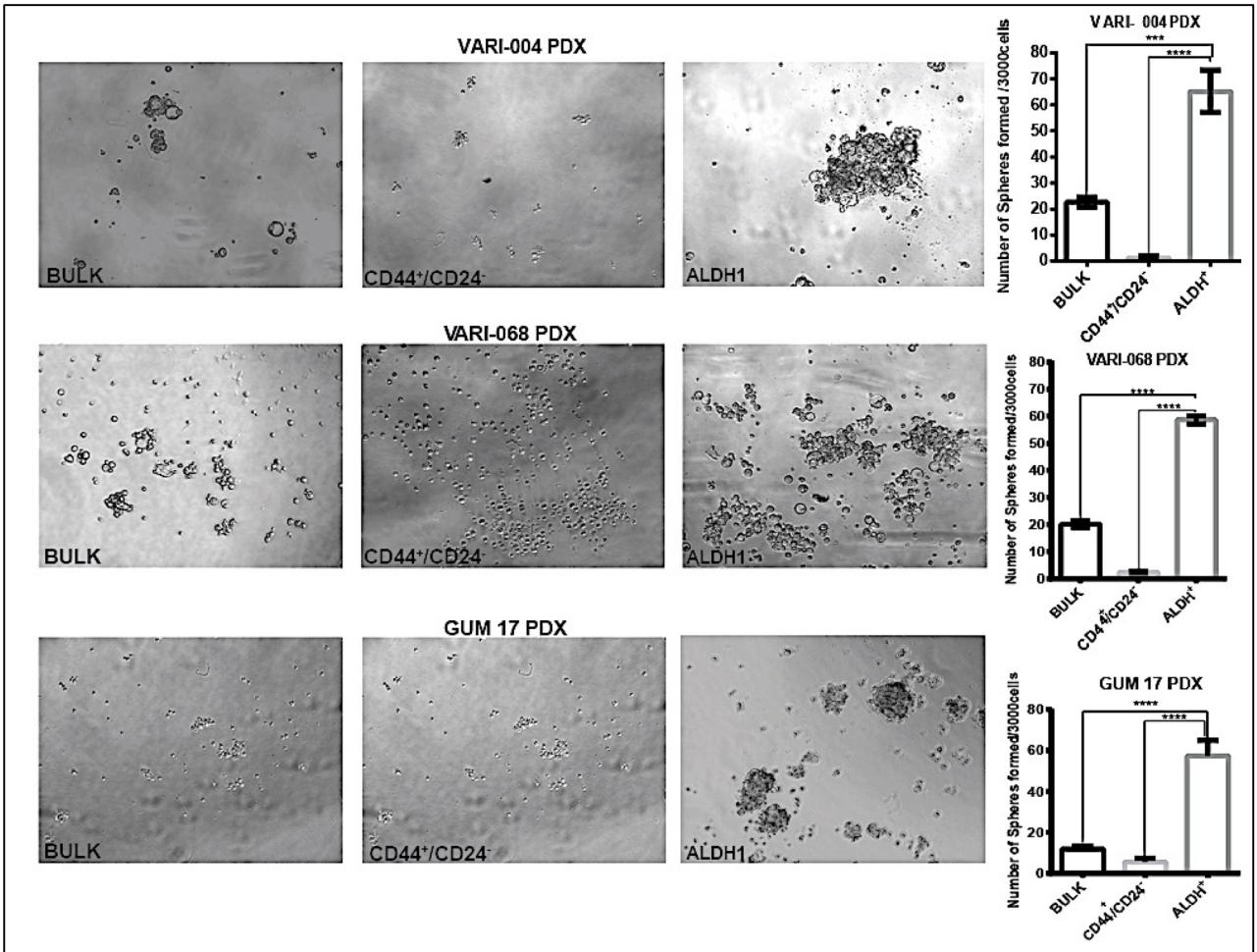


Figure 4-4: Tumorsphere formation assay for ALDH+, CD44+/CD24- and BULK cells from three different PDXs.

A shows tumor spheres formed from the BULK, CD44+/CD24- and ALDH+ sorted cells from VARI-068 and VARI-004 respectively, both TNBC PDXs. Cells were plated as described in the methods. Graph showing the average number of spheres formed per well of the 96-flat bottom ultra-low attachment plate for VARI-068 C, VARI-004, D.

separate the ALDH+ sub-population from the other two sub-populations. We identified 272 genes that were differentially expressed between the ALDH+ and the CD44+/CD24- sub-populations, (p-value 0.0008, FDR 0.05).

Table 4-3: ALDH enriched pathways

Pathway	Count	PValue
Wnt signaling pathway	13	6.88E-04
MAPK signaling pathway	17	0.002
Axon guidance	11	0.002
GnRH signaling pathway	9	0.005
Neurotrophin signaling pathway	10	0.006
TGF-beta signaling pathway	8	0.008
Endocytosis	12	0.010
Melanogenesis	8	0.017
Pathways in cancer	16	0.029
O-Glycan biosynthesis	4	0.046

Of these, 168 genes were down regulated while 104 were up regulated in the ALDH+ sub-population. Using LRpath, we identified ten pathways that are enriched in the ALDH+, (Table 4-3) and five pathways that are down regulated with respect to the CD44+/CD24- sub-population (Table 4-4).

The WNT and MAPK pathways were the most enriched in the ALDH+ population. TGF-beta signaling pathway was also significantly enriched in the ALDH+ cells. The pathways that were down regulated were mostly biosynthetic pathways, such as Aminoacyl-tRNA biosynthesis, N-Glycan biosynthesis and one carbon by folate.

. Comparing the ALDH+ to the BULK populations, we identified 373 genes that were significantly differentially expressed (p-value <0.001, FDR < 0.05). Of these, 372 were up regulated with the gene XIST being the only gene that was significantly down regulated in the ALDH+ vs BULK comparison. Among the pathways that were significantly enriched in the ALDH+ vs BULK comparison were WNT, MAPK and TGF-beta. We identified 14 genes that were simultaneously differentially expressed between the ALDH+ vs the CD44+/CD24- as well as ALDH+ vs bulk (p-value <0.001, FDR < 0.05). Thus, we have identified a set of 14 genes that separate the ALDH+ from CD44+/CD24- as well as BULK sub-populations, as seen in figure 4-5C. The three most significantly changed genes were CPXM1, MMP2, and PCDH7. CPXM1 is a gene that likely encodes a member of the carboxypeptidase family of proteins [provided by RefSeq, May 2010]. MMP2, (Matrix metalloproteinase 2) belongs to a broad family of zinc-dependent proteases that are important in extracellular matrix degradation, and MMP2 is important in breast tumor invasion and metastases and has been shown to mediate breast cancer brain metastasis[38].

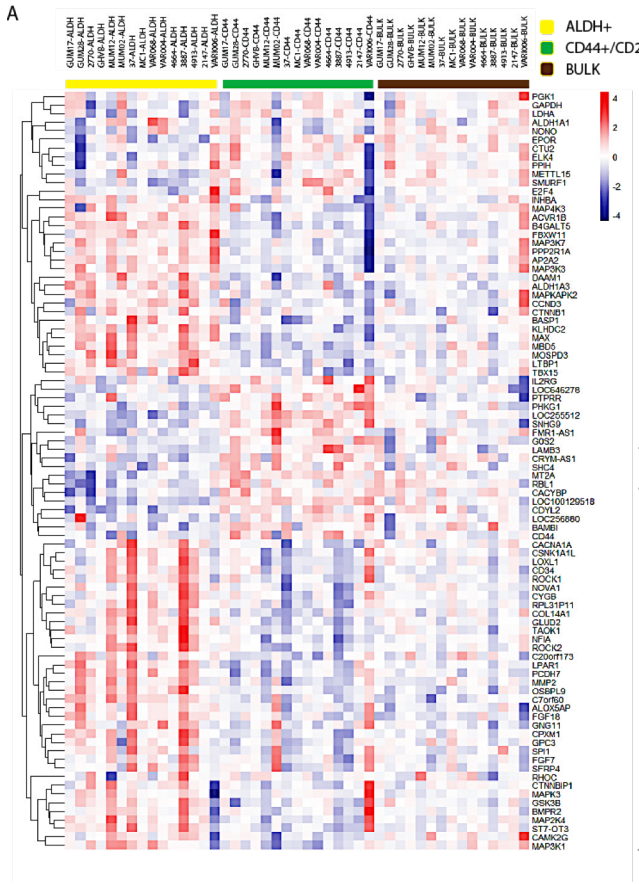
PROTOCADHERIN 7; PCDH7 or *BRAIN-HEART PROTOCADHERIN*; BHPCDH promotes the assembly of carcinoma-astrocyte gap junctions that promotes brain metastasis in breast and lung carcinomas. PCDH7 also activate pathways in brain metastatic cells (through activation of STAT1 and NFKB), by which they support tumor growth and chemoresistance[39]. PCDH7 has also been shown to induce bone metastases in breast cancer[40]. Thus two of the three most significantly altered genes with known functions, are involved in tumor metastases.

Table 4-4: CD44+/CD24- enriched pathways

Term	Count	PValue
Aminoacyl-tRNA biosynthesis	9	2.09E-05
N-Glycan biosynthesis	8	3.58E-04
RNA degradation	7	0.006
One carbon pool by folate	4	0.011
Ribosome	7	0.044

Significant signaling interactions exist between the genes that are differentially altered in ALDH+ stem cells in primary TNBC.

To better understand the signaling interactions between the genes that are differentially expressed between the ALDH+ and other cell populations, we first created a heat map with the genes that had the highest fold change difference



Gene	ALDH+/CD44+CD24- FC	ALDH+/CD44+CD24- FDR	BULK/ALDH+ FC	Bulk/ALDH+ FDR
CPXM1	6.189	0.029	-3.076	1.73E-13
MMP2	5.588	0.043	-2.336	5.80E-07
PCHD7	5.271	0.034	-3.345	1.08E-17
NFIA	5.024	0.041	-2.520	1.26E-06
RPL10L	4.413	0.049	-2.753	2.38E09
GDA	3.868	0.043	-3.073	7.05E-11
LPAR1	3.373	0.047	-2.215	6.23E-06
PRRT4	2.814	0.034	-1.407	0.031
SLC16A2	2.788	0.049	-1.442	0.025
MARCKS	2.537	0.049	0.724	0.606
FOXP1	2.434	0.038	-1.929	0.001
CPEB4	2.066	0.039	-1.965	9.10E-05
MAST4	2.065	0.0399	-1.607	0.021
ARHGAP40	1.923	0.049	-2.004	0.004

Figure 4-5: Heat map of significant genes and their known gene interactions
Heatmap showing the normalized log2CPM values 96 genes. The heat map was generated with dataset from 96 gene with the most significant differential expression between the ALDH+, the CD44+/CD24- and the bulk populations. It was generated using the "pheatmap" function on the statistical software package R. Data was not clustered according to cell lines, but rather prearranged according to cell line conditions. Using gene ontology, we generated a map showing the signal relationship between all the genes that were up regulated in the ALDH+ samples. **C.** All the genes that were significantly differentially expressed between the ALDH+ and the CD44+/CD24-

between the ALDH+ and the CD44+/CD24- (FC >2). We also added the genes that belonged to the pathways that were most significant between these two sub-populations using the IPATHWAY guide software, figure 4-5A. We identified two

separate clusters of genes that are upregulated in the ALDH+ stem cells, separated by a cluster of genes that were downregulated in the ALDH+ population. Using gene ontology and David statistical software[41] we identified known signaling interactions involving the genes that were upregulated in the ALDH+ stem cells figure 4-5B.

Discussion

Using PDXs, we are the first to create a renewable tumor resource from women with native African ancestry (Ghanaian) and establish a comparative gene profiling system between the BCSC of ethnically diverse tumors. In this manner, we definitively demonstrate the existence of two distinct BCSC sub-populations in primary TNBC. The ALDH expressing stem cells exhibited a gene expression profile that was significantly different from the bulk and the CD44+/CD24- stem cell population with the up-regulation of genes involved in metastasis in the ALDH-expressing stem cells. The importance of using new approaches to understand the biology TNBCs has become increasingly paramount as it becomes more evident that we are encountering a lack of success in translating initially potentially promising drugs for TNBC into the clinic. There is a dire need for credentialed new targets for rational signaling driven therapeutics.

Utilizing a series of novel sorting schemas, the relevant BCSC populations could be isolated to discern their gene expression profiles in comparison to each other and to the bulk cells. As BCSC have been robustly implicated in certain aspects of tumor aggressiveness such as metastatic growth and resistance to conventional therapy we hypothesized that there would be unique signatures of these stem-like cell subpopulations that may enable targeting them specifically, as part of multimodality approaches to attack TNBC and effect the highest potential benefit for patients[13,14,18-20,42]. In particular, we and others have shown that the prevalence of TNBC is higher in women of African extraction than in Caucasian women and thus we also surmised that the differences in stem cell populations may lead to therapies that will in turn help ameliorate the mortality disparities between Caucasians and African Americans.

Work by our group and others has demonstrated the existence of two distinct populations of breast cancer stem cells in breast tumors[15,22]. These are the CD44+/CD24- stem cells that are said to be mesenchymal-like (epithelial-to-mesenchymal transition, EMT) and the ALDH+ stem cells showed to be epithelial-like (mesenchymal-epithelial-transition, MET)[43] in primary breast tumors across all tumor subtypes. In this study focused on TNBC only, we have clearly demonstrated the existence on these two distinct populations of stem cells in TNBC. Interestingly however, in this study, we demonstrate that the ALDH+ stem cells are the only stem cell population that has a gene expression profile that is distinct and separate from the bulk population, as the CD44+/CD24-

did not separate out from the bulk cells. The ALDH+ stem cells displayed a gene profile pattern that would be considered mesenchymal-like with the up regulation of MMP2 and down regulation of KI67 (data not shown). Relative to the CD44+/CD24- stem cells, ALDH+ stem cells also had an up-regulation of pathways WNT, TGF-B, and MAPK, all of which are important for tumor metastasis and general therapeutic resistance. This is not surprising because the expression of ALDH1 has been shown to be important in TNBC as a whole and in TNBC in women with African ancestry in particular[15,17].

Although the tumors in our study did not exhibit a gene expression signature that segregated by ethnicity, a firm conclusion cannot be made on this point at this time due to the limited sample size and the high degree of heterogeneity exhibited by these TNBCs. We note that this study represents the only study of the biology of TNBC and stem cells in which two thirds of the sample population is women with African ancestry with half of them being from women with pure African ancestry.

Recent studies suggest that ALDH1 expressing stem cells might be the cells of origin of basal breast cancers (TNBC), the subtype of breast cancer that occurs most commonly in patients with BRCA1 mutation and that the loss of the BRCA1 expression leads to the expansion of ALDH expressing cells[44-46]. In our previous work, we reported that benign breast tumors in Ghanaian women, (a population with the highest reported institutional based incidence of TNBC), have a very high expression of ALDH1, even higher than seen in the malignant

tumors[16]. The findings in this recent study that in aggressive TNBCs from women with African ancestry and Caucasian women, it is the ALDH expressing cells that represent the stem cell population is therefore very relevant to our findings. It will be important therefore to investigate the BRCA1 mutation frequency in African women with TNBC as well as downregulation of BRCA1 expression by other non-genetic means, both of which are speculated to be very high. Indeed, in our previous study (chapter 3), we identified that two-thirds of Ghanaian tumors did not express the BRCA protein.

The existence of molecular differences between or within breast tumors (tumor heterogeneity) has been well documented in breast cancers in general and more specifically in TNBC. These molecular differences occur either between different patients with the same tumor type (intertumor heterogeneity) or within the same patient (intratumor heterogeneity), which can occur in different parts of the same tumor or between a primary tumor and its metastasis[47-51]. Heterogeneity contributes to development of therapy resistance and raises concern for researchers, clinicians and patients as to the ultimate effectiveness of targeted drugs or drug combinations. The 15 tumors that were used for this study exhibited profound intertumor heterogeneity with the samples from each tumor clustering tightly together. Despite this large degree of molecular differences between the tumors, there were a number of genes and pathways that were very significantly differentially expressed in the ALDH expressing stem cells compared to the other populations. Therefore, focusing on targeting the

genes and pathways we discovered in this study has the potential to lead to therapies that may be effective across ethnicities by eliminating the cells with the highest potential for mesenchymal motion and metastatic growth. Taken together, we propose that our data supports the targeting of the ALDH1+ stem cell subpopulation, which may be an important adjunct strategy to conventional chemo radiation in TNBC.

Acknowledgements: Work supported in part by grants Komen promise grant, the Breast Cancer Research Foundation (MW, SDM), the Metavivor Foundation (SDM), and the UM CCC core, grant Rackham Barbour Scholarship

References

1. Hammond ME, Hayes DF, Dowsett M, et al. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer (unabridged version). *Arch Pathol Lab Med.* Jul 2010;134(7):e48-72.

2. Wolff AC, Hammond ME, Hicks DG, et al. Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. *J Clin Oncol.* Nov 1 2013;31(31):3997-4013.

3. Malorni L, Shetty PB, De Angelis C, et al. Clinical and biologic features of triple-negative breast cancers in a large cohort of patients with long-term follow-up. *Breast cancer research and treatment.* Dec 2012;136(3):795-804.

4. Dent R, Trudeau M, Pritchard KI, et al. Triple-negative breast cancer: clinical features and patterns of recurrence. *Clin Cancer Res.* Aug 1 2007;13(15 Pt 1):4429-4434.

5. Liedtke C, Mazouni C, Hess KR, et al. Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. *J Clin Oncol.* Mar 10 2008;26(8):1275-1281.

6. Dent R, Hanna WM, Trudeau M, Rawlinson E, Sun P, Narod SA. Pattern of metastatic spread in triple-negative breast cancer. *Breast cancer research and treatment.* May 2009;115(2):423-428.

7. National Comprehensive Cancer Network. National Comprehensive Cancer Network Breast Cancer. 2015.

8. Carey LA, Dees EC, Sawyer L, et al. The triple negative paradox: primary tumor chemosensitivity of breast cancer subtypes. *Clin Cancer Res.* Apr 15 2007;13(8):2329-2334.

9. Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumours. *Nature.* Aug 17 2000;406(6797):747-752.

10. Lehmann BD, Bauer JA, Chen X, et al. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J Clin Invest.* Jul 2011;121(7):2750-2767.

11. Prat A, Perou CM. Deconstructing the molecular portraits of breast cancer. *Mol Oncol.* Feb 2011;5(1):5-23.

12. Shah SP, Roth A, Goya R, et al. The clonal and mutational evolution spectrum of primary triple-negative breast cancers. *Nature.* Jun 21 2012;486(7403):395-399.

13. Charafe-Jauffret E, Ginestier C, Iovino F, et al. Breast cancer cell lines contain functional cancer stem cells with metastatic capacity and a distinct molecular signature. *Cancer research.* Feb 15 2009;69(4):1302-1313.

14. Dontu G, Abdallah WM, Foley JM, et al. In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells. *Genes Dev.* May 15 2003;17(10):1253-1270.

15. Ginestier C, Hur MH, Charafe-Jauffret E, et al. ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. *Cell Stem Cell.* Nov 2007;1(5):555-567.

16. Schwartz T, Stark A, Pang J, et al. Expression of aldehyde dehydrogenase 1 as a marker of mammary stem cells in benign and malignant breast lesions of Ghanaian women. *Cancer.* Feb 1 2013;119(3):488-494.

17. Nalwoga H, Arnes JB, Wabinga H, Akslen LA. Expression of aldehyde dehydrogenase 1 (ALDH1) is associated with basal-like markers and features of aggressive tumours in African breast cancer. *Br J Cancer.* Jan 19 2010;102(2):369-375.

18. Moitra K, Lou H, Dean M. Multidrug efflux pumps and cancer stem cells: insights into multidrug resistance and therapeutic development. *Clin Pharmacol Ther.* Apr 2011;89(4):491-502.

19. Dave B, Chang J. Treatment resistance in stem cells and breast cancer. *J Mammary Gland Biol Neoplasia.* Mar 2009;14(1):79-82.

20. Li X, Lewis MT, Huang J, et al. Intrinsic resistance of tumorigenic breast cancer cells to chemotherapy. *J Natl Cancer Inst.* May 7 2008;100(9):672-679.

21. Campbell LL, Polyak K. Breast tumor heterogeneity: cancer stem cells or clonal evolution? *Cell cycle.* Oct 1 2007;6(19):2332-2338.

22. 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.* Apr 1 2003;100(7):3983-3988.

23. Perrone G, Gaeta LM, Zagami M, et al. In situ identification of CD44+/CD24- cancer cells in primary human breast carcinomas. *PLoS One.* 2012;7(9):e43110.

24. Charafe-Jauffret E, Ginestier C, Iovino F, et al. Aldehyde dehydrogenase 1-positive cancer stem cells mediate metastasis and poor clinical outcome in inflammatory breast cancer. *Clin Cancer Res.* Jan 1 2010;16(1):45-55.

25. Bianchini G, Balko JM, Mayer IA, Sanders ME, Gianni L. Triple-negative breast cancer: challenges and opportunities of a heterogeneous disease. *Nat Rev Clin Oncol.* May 17 2016.

26. Stark A, Kleer CG, Martin I, et al. African ancestry and higher prevalence of triple-negative breast cancer: findings from an international study. *Cancer.* Nov 1 2010;116(21):4926-4932.

27. Carey LA, Perou CM, Livasy CA, et al. Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. *JAMA.* Jun 7 2006;295(21):2492-2502.

28. Morris GJ, Naidu S, Topham AK, et al. Differences in breast carcinoma characteristics in newly diagnosed African-American and Caucasian patients: a single-institution compilation compared with the National Cancer Institute's Surveillance, Epidemiology, and End Results database. *Cancer.* Aug 15 2007;110(4):876-884.

29. Stead LA, Lash TL, Sobieraj JE, et al. Triple-negative breast cancers are increased in black women regardless of age or body mass index. *Breast cancer research : BCR.* 2009;11(2):R18.

30. Lund MJ, Trivers KF, Porter PL, et al. Race and triple negative threats to breast cancer survival: a population-based study in Atlanta, GA. *Breast cancer research and treatment.* Jan 2009;113(2):357-370.

31. McCormack VA, Joffe M, van den Berg E, et al. Breast cancer receptor status and stage at diagnosis in over 1,200 consecutive public hospital patients in Soweto, South Africa: a case series. *Breast cancer research : BCR.* 2013;15(5):R84.

32. Zhang H, Cohen AL, Krishnakumar S, et al. Patient-derived xenografts of triple-negative breast cancer reproduce molecular features of patient tumors and respond to mTOR inhibition. *Breast cancer research : BCR.* 2014;16(2):R36.

33. Zhao X, Liu Z, Yu L, et al. Global gene expression profiling confirms the molecular fidelity of primary tumor-based orthotopic xenograft mouse models of medulloblastoma. *Neuro-oncology.* May 2012;14(5):574-583.

34. de Plater L, Lauge A, Guyader C, et al. Establishment and characterisation of a new breast cancer xenograft obtained from a woman carrying a germline BRCA2 mutation. *Br J Cancer.* Oct 12 2010;103(8):1192-1200.

35. Evelyn Jiagge JKO, Jessica Bensenhaver, Francis Aitpillah, Kofi Gyan, Ishmael Kyei, Ernest Osei-Bonsu, Ernest Adjei, Michael Ohene-Yeboah, Kathy Toy, Karen Eubanks Jackson, Marian Akpaloo, Dorcas Acheampong, Beatrice Antwi, Faustina Obeng Agyeman, Zainab Alhassan, Linda Ahenkorah Fondjo, Osei Owusu-Afriyie, Robert Newman Brewer, Amma Gyamfuah, Barbara Salem, Timothy Johnson, Max Wicha, Sofia Merajver, Celina Kleer, Judy Pang, Emmanuel Amankwaa-Frempong, Azadeh Stark, Francis Abantanga, Lisa Newman, Baffour Awuah. Breast Cancer and African Ancestry: Lessons Learned at the 10-Year Anniversary of the Ghana-Michigan Research Partnership and International Breast Registry. *JOURNAL OF GLOBAL ONCOLOGY.* 2016.

36. Ponti D, Costa A, Zaffaroni N, et al. Isolation and in vitro propagation of tumorigenic breast cancer cells with stem/progenitor cell properties. *Cancer research.* Jul 1 2005;65(13):5506-5511.

37. Prat A, Parker JS, Karginova O, et al. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast cancer research : BCR.* 2010;12(5):R68.

38. Mendes O, Kim HT, Lungu G, Stoica G. MMP2 role in breast cancer brain metastasis development and its regulation by TIMP2 and ERK1/2. *Clin Exp Metastasis.* 2007;24(5):341-351.

39. Chen Q, Boire A, Jin X, et al. Carcinoma-astrocyte gap junctions promote brain metastasis by cGAMP transfer. *Nature.* May 26 2016;533(7604):493-498.

40. Li AM, Tian AX, Zhang RX, Ge J, Sun X, Cao XC. Protocadherin-7 induces bone metastasis of breast cancer. *Biochem Biophys Res Commun.* Jul 5 2013;436(3):486-490.

41. Szklarczyk D, Franceschini A, Wyder S, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic acids research.* Jan 2015;43(Database issue):D447-452.

42. Bauer KR, Brown M, Cress RD, Parise CA, Caggiano V. Descriptive analysis of estrogen receptor (ER)-negative, progesterone receptor (PR)-negative, and HER2-negative invasive breast cancer, the so-called triple-negative phenotype: a population-based study from the California cancer Registry. *Cancer.* May 1 2007;109(9):1721-1728.

43. Liu S, Cong Y, Wang D, et al. Breast cancer stem cells transition between epithelial and mesenchymal states reflective of their normal counterparts. *Stem Cell Reports.* Jan 14 2014;2(1):78-91.

44. Molyneux G, Geyer FC, Magnay FA, et al. BRCA1 basal-like breast cancers originate from luminal epithelial progenitors and not from basal stem cells. *Cell Stem Cell.* Sep 3 2010;7(3):403-417.

45. Lim E, Vaillant F, Wu D, et al. Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. *Nat Med.* Aug 2009;15(8):907-913

46. Brooks MD, Burness ML, Wicha MS. Therapeutic Implications of Cellular Heterogeneity and Plasticity in Breast Cancer. *Cell Stem Cell.* Sep 3 2015;17(3):260-271.

47. Bedard PL, Hansen AR, Ratain MJ, Siu LL. Tumour heterogeneity in the clinic. *Nature.* Sep 19 2013;501(7467):355-364.

48. Burrell RA, McGranahan N, Bartek J, Swanton C. The causes and consequences of genetic heterogeneity in cancer evolution. *Nature.* Sep 19 2013;501(7467):338-345.

49. Yap TA, Gerlinger M, Futreal PA, Pusztai L, Swanton C. Intratumor heterogeneity: seeing the wood for the trees. *Sci Transl Med.* Mar 28 2012;4(127):127ps110.

50. Gerlinger M, Rowan AJ, Horswell S, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med.* Mar 8 2012;366(10):883-892.

51. Zardavas D, Irrthum A, Swanton C, Piccart M. Clinical management of breast cancer heterogeneity. *Nat Rev Clin Oncol.* Jul 2015;12(7):381-394.

Chapter 5 : Summary and Discussion

Introduction

Overall breast cancer mortality rates decreased by 36% from 1989 to 2012 in the united states after slowly increasing for many years (0.4% annually since 1975)[1]. As a result of this decline, 249,000 breast cancer deaths have been averted in US women from 1990 to 2012[2]. Declines in breast cancer mortality rates have been attributed to both improvements in treatment (e.g., adjuvant chemotherapy and hormonal therapy in the 1980s and targeted therapies in the 1990s) and early detection. This means that, if we better understand the biology of breast cancer, and thus find the right treatment, stratified to the right patients, there is the possibility of saving many lives, not only in the united states but globally.

The gain from recent advancement in treatment however is not evenly distributed for all breast cancer patients. Patients with triple negative breast cancer still have chemotherapy as the mainstay of treatment since no targeted therapy has been successful for this disease. The high proportion of triple negative breast cancer in women with African ancestry in the United States likely accounts in part for the striking divergence in long-term breast cancer mortality trends between black and white women which emerged in the early 1980s and

has since continued to widen such that by 2012, mortality rates were 42% higher in black women than in white women[1,3,4].

Triple negative breast cancer occur in younger patients, are high grade and are the most common breast cancer subtype in patients with BRCA1 mutation (70 % of breast cancers in patients with BRCA1 mutation are triple negative).

Globally the highest proportion is in western sub-Saharan Africa where single-institution studies reports proportions as high as 80% of all breast cancers. Considering the shared ancestry between AA women and women in western sub-Saharan African, there is a possibility that African ancestry may be associated with a higher chance of developing TNBC. This is further buttressed by the similar phenotype of breast cancer in women with African ancestry and that in women with hereditary breast cancer. Both groups have early onset breast cancer, high grade tumors and higher incidence if male breast cancer.

We therefore hypothesized that, the key to understanding the complexity of TNBC treatment lies in understanding its biology in women with African ancestry in comparison with other ethnicities.

To validate the hypothesis, I had three aims. The first aim was to create patients derived xenografts based on triple negative breast cancer from patients in western sub-Saharan Africa, African-American and white patients that will be used for experiments.

Considering the similarity between breast cancers black women and women with hereditary breast cancer (BRCA mutation), our second aim was to understand the correlation between African ancestry, triple negative breast cancer and BRCAness. BRCAness can be defined as a defect in double-strand break repair by homologous recombination[5].

Considering the fact that triple negative breast cancer patients have much rapid relapse after chemotherapy with lower survival in the metastatic setting, the next aim was to look at the gene expression in the cells that have been shown to mediate tumor recurrence, metastasis and resistance to conventional therapy[6-11]. These are the breast cancer stem cells that have also been shown to be particularly abundant in triple negative breast cancers, especially that in women with African ancestry[12,13].

Below is a summary of the chapters of this thesis, detailing our novel approach to understanding the biology of TNBC and thus finding possible new drug targets. It reports on our multidisciplinary, international outreach to better understand TNBC epidemiology and its impact on patients and care givers in chapter II, the possible role of BRCA mutation in chapter III and the importance of breast cancer stem cells to this disease in chapter IV.

Summary of Chapters

Chapter 2: The need to create patient derived xenografts from triple negative breast cancer patients in western sub-Saharan Africa motivated the establishment of a breast cancer research partnership between the Komfo Anokye Teaching Hospital in Kumasi, Ghana and the University of Michigan Comprehensive Cancer Center in Ann Arbor, Michigan. The aim of this collaboration is to better understand the contribution of African ancestry to the breast cancer phenotype observed in women with African ancestry, both in Africa and in the United States. This aim was pursued by a multidisciplinary collaborative partnership between these institution, its staff workers and patients.

After IRB approval and with informed consent, tissue samples were collected from patients receiving treatment for breast cancer. These samples were transported to the University of Michigan where they were implanted in immunosuppressed mice to create patient derived xenografts models for the study of the biology of these tumors. These patient-derived xenografts have been used for studying breast cancer stem cells and aggressive triple negative breast cancer biology. Tissue samples that were obtained through this collaboration has been studied extensively and has led to a better understanding of ethnic/racial influences of breast cancer. The benefits of the findings of this work is not only to women with African ancestry but to all TNBC patients who desperately need new treatments for their disease.

The collaborative work led to educational training programs for doctors in both institutions as well as scientific investigations related to the comparative biology of breast cancer in Ghanaian African, African American, and White American patients. In also led to a very significant improvement in the patient management protocols in KATH and direct benefits to patients as well.

Chapter 3: women with mutations in BRCA1 and BRCA2 gene, have hereditary susceptibility to breast cancer. Those women with BRCA1 mutation, have up to 70% chance of developing TNBC when they develop breast cancer. In patients with triple negative breast cancer, there is a 30% chance of harboring a mutation in the BRCA compared to 6% in other sporadic breast cancers. Women in western sub-Saharan Africa have the highest reported proportion of triple negative breast cancer but the prevalence of BRCA mutations or alterations in the BRCA associated genes and pathways are not known in these women. We first determined the expression of the BRCA1 and BRCA2 protein in formalin fixed paraffin embedded samples developed from triple negative breast cancers from these women. We discovered that, two thirds of the patient's sample did not express either or both of these tumors. This is the highest report of absence expression of BRCA protein indicating that this population might have a high prevalence of BRCAness.

To identify possible causes of the loss expression of the BRCA protein, we performed exome sequencing for four of these triple negative breast tumors for which we had fresh tumor samples. We identified novel BRCA1 and BRCA2

variants in three of the tumors, some of which will likely lead to truncation of the protein. This further solidifies the point that this might be a population with a very high prevalence of BRCA gene variants or variants in BRCA associated genes that leads to impairment in double-strand break repair by homologous recombination and so the need for further studies.

We also compared the characteristics of breast cancer in African-American women, white women, and women in western sub-Saharan Africa and in east Africa. We confirmed that the frequency of triple negative tumors was highest in the Ghanaian patients (50.8%), lowest in the White American and Ethiopian patients (15.9% in each), and intermediate in the African Americans (34.3%; p=<0.0001). Ghanaian and African-American tumors were more likely to be high grade while the Ethiopian tumors had the highest frequency of HER2 over expression. It was thus established in this study that significant racial /ethnic diversity exists in the breast cancer phenotype. This diversity might contribute to the disparity in breast cancer outcomes observed in women of different ethnicities.

Chapter 4: In previous work by our group, we showed a very high prevalence of expression of ALDH1, a marker of breast cancer stem cells in benign breast tumors in Ghanaian women, even higher expression than was observed for malignant tumors in this population. there is ample evidence

demonstrating that these breast cancer stem cells are responsible for tumor metastasis, resistance to chemotherapy and radiation and for tumor recurrence. We therefore sought to identify genes and pathways that are differentially expressed between the breast cancer stem cells and the bulk tumor cells in primary patient derived xenografts developed from tumors from women with divers ethnicities. Using flow cytometry, we sorted for the two populations of breast cancer stem cells (CD44+CD24-EPCAM+ and ALDH+) and bulk (live, human cells, not excluding stem cells) from 15 (5 Ghanaian, 5 African-American and 5 White) patient derived xenografts. We performed RNA sequencing (Illumina HiSeq platform) on the isolated populations and bulk cells (45 samples). Comprehensive bioinformatics analyses led to the identification of highly significantly differentially expressed genes and pathways between the cell populations. By principal component analysis, the tumors were very heterogeneous. However, the ALDH+ cells separated out from the CD44+/CD24- and the bulk cells. The genes that were significantly differentially expressed in the ALDH+ population compared with the other two populations had significant signaling interactions using gene ontology. We identified 14 genes that were simultaneously differentially expressed between the ALDH+ vs the CD44+/CD24- as well as ALDH+ VS bulk (p-value <0.001, FDR < 0.05). The 3 most significant genes were MMP2 and PCDH7, both known to be involved in breast cancer metastasis and CPXM1, a carboxylase. The WNT, MAPK and TGF-beta pathways known to mediate metastasis were all significantly up-regulated in the

ALDH+ population with down regulation of biosynthetic pathways which were up-regulated in the CD44+/CD24- population. In this study, we identified the ALDH+ stem cells to be the only stem cell population in these aggressive breast tumors predominantly from women with African ancestry with a significantly higher propensity to for mammospheres compared to the bulk and the CD44+/CD24-/EPCAM+ population.

Future Directions.

While the research here elucidates some of the biology of triple negative breast cancer as it pertains to women with African ancestry, there are so many unanswered questions. We demonstrated that, a high percentage of breast tumors from Ghanaian patients had loss of expression of BRCA proteins and three out of four of these patients whose tumors were sequenced had variants in the BRCA gene. However, we have not been able to validate that the loss of BRCA protein as observed in the immunohistochemistry staining is due to mutation in the BRCA gene. It will be important therefore to investigate the BRCA1 mutation frequency in African women with TNBC as well as down-regulation of BRCA1 expression by other non-genetic means, both of which are speculated to be very high.

Using PDXs, we are the first to create a renewable tumor resource from women with native African ancestry (Ghanaian) and establish a comparative gene profiling system between the BCSC of ethnically diverse tumors. This study represents about the only study of the biology of TNBC and stem cells in which two thirds of the sample population is of women with African ancestry with half of them being from women with pure African ancestry. The sample size used for this experiment was not large enough to determine an ethnic relationship with the gene and pathway expression. There is therefore the need to increase the sample size to determine if there are genes and pathways that are differentially expressed in triple negative breast cancer in women with African ancestry that can be targeted for therapy.

Our data infers that ALHD as a marker of breast cancer stem cells may represent an interesting link into the biology of the TNBC tumors. We have, shown that matrix metalloproteases are highly upregulated in ALDH expressing breast cancer stem cells. Several studies have shown that MMPs play an important role in tumor metastasis by degrading extracellular matrix. But the therapies directed against MMPs have not been effective in the clinic. In our studies, we have shown that MMP-2 is highly expressed and maintained in in vivo system in the breast cancer stem cells. Our data suggests that the disconnect in the preclinical studies and the clinical data may partly be due to inadequate model system used to study the potential role of MMPs as a target. Our studies impress upon the role of MMPs in connection to breast cancer stem

cell. It will be interesting to further study the role of MMP and ALDH-1 in relation to breast cancer stem cells. Apart from MMP we have identified several other genes that were significantly up-regulated in the ALDH expressing stem cells. These include CPXM1, a carboxylase whose function is yet to be determined[14] and PCDH7 which has been shown to induce metastasis in breast cancer[15]. Focusing on targeting the genes and pathways we discovered in this study has the potential to lead to therapies that may be effective across ethnicities, thereby eliminating the cells with the highest potential for mesenchymal motion and metastatic growth.

Final Remarks

Triple negative breast cancer is a very heterogeneous disease making it extremely difficult to target. Previous attempts to find targeted therapy for this disease has failed thus the need to find different approaches to understand its biology and find new drug targets is urgently needed. In this study, we chose a novel approach by studying the population with the highest proportion of the disease (African women) and the cells that most likely mediate the tumors aggressive phenotype (breast cancer stem cells). Here we have Studied the biology of breast cancer stem cells in actual patient tumors which has provided further insight into novel genes and pathways that may be potential drug targets in this disease. Targeting these genes and pathways may have the potential of

benefiting black women with triple negative breast cancer who have been underrepresented in previous research and clinical trials. It will also benefit all patient with triple negative breast cancer irrespective of their racial/ethnic background as the genes we identified did not cluster on the bases of ethnicity.

References

1. Howlader N NA, Krapcho M, et al. . SEER Cancer Statistics Review, 1975-2012, based on Nov 2014 SEER data submission. Bethesda, MD: National Cancer Institute;. 2015.

2. DeSantis CE, Fedewa SA, Goding Sauer A, Kramer JL, Smith RA, Jemal A. Breast cancer statistics, 2015: Convergence of incidence rates between black and white women. *CA Cancer J Clin.* 2016;66(1):31-42.

3. Berry DA, Cronin KA, Plevritis SK, et al. Effect of screening and adjuvant therapy on mortality from breast cancer. *N Engl J Med.* 2005;353(17):1784-1792.

4. Munoz D, Near AM, van Ravesteyn NT, et al. Effects of screening and systemic adjuvant therapy on ER-specific US breast cancer mortality. *J Natl Cancer Inst.* 2014;106(11).

5. Lord CJ, Ashworth A. BRCAness revisited. *Nature reviews Cancer.* 2016;16(2):110-120.

6. Dent R, Hanna WM, Trudeau M, Rawlinson E, Sun P, Narod SA. Pattern of metastatic spread in triple-negative breast cancer. *Breast cancer research and treatment.* 2009;115(2):423-428.

7. Liedtke C, Mazouni C, Hess KR, et al. Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. *J Clin Oncol.* 2008;26(8):1275-1281.

8. Dent R, Trudeau M, Pritchard KI, et al. Triple-negative breast cancer: clinical features and patterns of recurrence. *Clin Cancer Res.* 2007;13(15 Pt 1):4429-4434.

9. Carey LA, Dees EC, Sawyer L, et al. The triple negative paradox: primary tumor chemosensitivity of breast cancer subtypes. *Clin Cancer Res.* 2007;13(8):2329-2334.

10. Charafe-Jauffret E, Ginestier C, Iovino F, et al. Aldehyde dehydrogenase 1-positive cancer stem cells mediate metastasis and poor clinical outcome in inflammatory breast cancer. *Clin Cancer Res.* 2010;16(1):45-55.

11. Dontu G, Abdallah WM, Foley JM, et al. In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells. *Genes Dev.* 2003;17(10):1253-1270.

12. Nalwoga H, Arnes JB, Wabinga H, Akslen LA. Expression of aldehyde dehydrogenase 1 (ALDH1) is associated with basal-like markers and features of aggressive tumours in African breast cancer. *Br J Cancer.* 2010;102(2):369-375.

13. Schwartz T, Stark A, Pang J, et al. Expression of aldehyde dehydrogenase 1 as a marker of mammary stem cells in benign and malignant breast lesions of Ghanaian women. *Cancer.* 2013;119(3):488-494.

14. Ota T, Suzuki Y, Nishikawa T, et al. Complete sequencing and characterization of 21,243 full-length human cDNAs. *Nat Genet.* 2004;36(1):40-45.

15. Li AM, Tian AX, Zhang RX, Ge J, Sun X, Cao XC. Protocadherin-7 induces bone metastasis of breast cancer. *Biochem Biophys Res Commun.* 2013;436(3):486-490.