1-1-2017

# Clinicopathology And Molecular Determinants Underlying Benign Breast And Breast Cancer Lesions

Andreana Holowatyj Holowatyj
*Wayne State University,*

# CLINICOPATHOLOGY AND MOLECULAR DETERMINANTS UNDERLYING BENIGN BREAST AND BREAST CANCER LESIONS

by

**ANDREANA NATALIE HOLOWATYJ**

**DISSERTATION**

Submitted to the Graduate School

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

**DOCTOR OF PHILOSOPHY**

2017

MAJOR: CANCER BIOLOGY

Approved By:

_____

Advisor Date

_____

_____

_____

_____

## DEDICATION

To my parents –

Tato, you always said, "*Try your hardest in everything you do. I love you.*"
Thank you for being my hero, and my guardian angel, these last fifteen years.

Mama, your resolute strength and unconditional love just inspire me every day.
"*I'll love you forever. I'll like you for always."*

## ACKNOWLEDGMENTS

My greatest thanks are owed to a plethora of individuals, without whom my thesis and career accomplishments to date would not have been at all possible. First and foremost to my mentor, Michele L. Cote, who has fostered my growth as a doctoral student, scientist, and instructor, has taught me valuable lessons both in and outside of the lab, has provided exemplary mentorship throughout my career at the Wayne State University School of Medicine and Barbara Ann Karmanos Cancer Institute, and has served as an outstanding role model for me to emulate in my career. I am ever grateful for *always* having you there for me, especially when I needed it most, and encouraging *every* single ambitious and wild idea I have had throughout my career, of which there were many. I am truly fortunate to be able to call you my mentor and a friend; my successes as a doctoral student are a direct testament to your constant support and the phenomenal, lasting impact you have made on my professional and personal growth.

I owe many thanks as well to Julie J. Ruterbusch, who spent countless hours teaching me every detail of “SAS-ing” and data analysis, letting me hang out in her office, helping brainstorm and supporting my endless and sometimes nearly impossible project ideas *always* with enthusiasm, and answering what probably seemed like never-ending technical questions. I cannot express enough gratitude to you both for all the unwavering support, encouragement, and friendship that together you have provided me over these last few years: for believing in me to push beyond limits and boundaries, for helping me not just achieve but crush my goals, for having constant faith in me even when I had my doubts (especially #bringonthetwenty), for always having my back, and

for providing a phenomenal learning environment and a *truly* remarkable and unforgettable doctoral career experience.

I am very appreciative to Zeng-Quan Yang for the two years of molecular biology doctoral training in his laboratory. The strong commitment you demonstrated towards my success as a doctoral student through various training and research opportunities has contributed significantly to my growth and development and continues into the future. From the day you invited me as an undergraduate student through the Cancer Biology Summer Fellowship Program all the way to helping me overcome immunoprecipitation experiment challenges for the binding project and serving on my committee, you have vastly expanded my knowledge, my passion for oncological research, and have helped shape my career and continued aspirations. The support provided to me by yourself, along with your lab members – particularly Jack Wu, Jinling Hou, Qin Ye, Hui Liu, Lanxin Liu, and Yuan-Yuan Jiang, who helped teach me new methods, helped expand my knowledge in biology, and who contributed various experiments and expertise – is forever acknowledged and appreciated, and has helped complete the story for this doctoral work.

To Ann G. Schwartz and Kendra L. Schwartz: my sincere thanks for your continuous encouragement and time from your busy schedules to mentor me and share your expertise. You have significantly contributed to my growth and career development. To Jennifer L. Beebe-Dimmer, thank you for introducing me to cancer epidemiology through your class. From that day forward, you have been a wonderful mentor and provided me unique opportunities to expand my training and knowledge, bringing it full circle with my teaching in your class this term. Your diligent efforts and

advice have helped immensely with my growth and career development. To Cathryn H. Bock and Kristen S. Purrington – thank you for being a constant encouragement to me, letting me expand my cancer knowledge and contribute to your studies, for your advice and mentorship. Thank you as well to Elena M. Stoffel and Laura S. Rozek for opportunities to learn from you both, and gain unique experience in global health and cancer genetics. I also extend my deepest gratitude to Lori A. Pile for being a constant with your open door for advice – especially during my teaching. The support and motivation you have each and all provided has been much appreciated.

Next, I owe immeasurable gratitude to Elisabeth I. Heath, David H. Gorski, and Jeffrey A. Triest – for letting me tag along around the clinic and operating room *every* single Tuesday or Thursday, contributing and taking interest in all of my research ideas and professional endeavors, providing your vast clinical expertise, having endless faith in my abilities and ambitions, and being constants with your unwavering support and encouragement over these years. Any time I had one reason to give up – you had this *magical* ability to find fifty-one reasons for me to keep fighting. It has been a privilege having the unique opportunity to learn from you all, and I truly could not be more fortunate to call you my mentors. Thank you to Joan Livingstone, Cathy Ciccotelli, and Donna Mysiewicz; you all have made the clinic an enjoyable experience and it has been a pleasure being able to learn from all of you.

I would also like to express gratitude to my dissertation committee members for their thoughtful questions, project suggestions, and constructive insights they put forward during this journey. In particular, my greatest thanks to Larry H. Matherly for his advice and constant encouragement – I can honestly say I would *not* have made it

through this journey without your support. Thank you to David H. Gorski for translational insight in the clinic, constant support and mentorship; Rouba Ali-Fehmi for clinical insight, teaching, and pathology expertise; Zeng-Quan Yang for the molecular biology training; and Katrina Studvent for her perspectives and providing translational input by acting as Komen liaison and patient advocate.

I would like to acknowledge my doctoral program - the Cancer Biology Program, in the Department of Oncology at the Wayne State University School of Medicine. Especially the Program Director, Larry H. Matherly; the Associate Director and Graduate Officer, George Brush; the Steering Committee; and the Program Manager, Nadia Daniel. Additional thanks to the Associate Dean of Graduate Scholars at the Wayne State University School of Medicine, Daniel Walz, and Deanna Dona in the Office of Graduate Scholars. Thank you all for the support, assistance, and encouragement.

Continuing on, thank you as well to the Department of Biological Sciences at Wayne State University for providing me with the part-time instructional opportunity and professional development during my doctoral career. Especially, thank you to David Njus for the teaching mentorship and valuable experience. I owe many thanks to Krystyn Purvis for her endless assistance and teaching advice when often called upon, and her steady support in and beyond the classroom.

To my students – I am privileged to have had the opportunity to teach you: for I have learned even more from what each of you brought to my classroom, from “boss status” to “give 100 – get 100” and “#letsdothis” over this time. You are all the reason my “professor magic” exists, and for that I am truly grateful.

Last but not least, thank you to additional members of the Cote lab, including Hyo Park and Asra Shaik, with whom I had the privilege of working with every day, always supported me, and have been great friends. My thanks to our collaborators for gene expression analyses at the Mayo Clinic Cancer Center in Jacksonville, Florida: in particular, Derek Radisky and Christine Mehner. Thank you as well to Greg Dyson in the Biostatistics Core at the Karmanos Cancer Institute for your fundamental help with gene expression analyses and always opening the door for all my questions and *endless* IPA runs. Thank you to our collaborators in the Department of Pathology at the Wayne State University School of Medicine, in particular Rouba Ali-Fehmi and Sudeshna Bandyopadhyay. Thank you to the BioBank and Correlative Sciences Core, in particular Julie Boerner, for help with the benign breast tissue work. My thanks to the Epidemiology Department and Population Sciences and Disparities Research Program, especially: Lea Staschke, Tara Baird, Dave Pandolfi, Sharon Moton, and Chandra Walton. Thank you to the Susan G. Komen for the Cure Graduate Training in Research Disparities Fellowship and the Wayne State University School of Medicine and Barbara Ann Karmanos Cancer Institute for career training support and funding.

They say it takes a village – and there are no words to express how fortunate I have been over these four years to have all of the immense support and mentorship that has defined my career, has contributed to my accomplishments and successes, and has shaped my future. Together, you all have gotten me through to the end of this journey – and helped establish a phenomenal start on my next adventure.

To *all* of you, and those I may have forgotten, my deepest thanks. I am much obliged.

# TABLE OF CONTENTS

## LIST OF FIGURES

# CHAPTER 1. GENERAL INTRODUCTION

The structure of the female breast is remarkable for its physiological function for lactation, and consists of muscular, lymphatic, connective, adipose, and epithelial tissues. On top of the pectoralis muscle and ribcage, the breast is situated between the edge of the sternum across to the center of the axilla. The breast is uniquely comprised of lobules, which collectively form the lobes of the breast – and produce milk for lactation in females. Milk ducts connect lobules and lobes of the breast to serve as a transport system for milk to the nipple (Figure 1). Fibrous connective and adipose tissues account for the majority of the breast mass, and are where ducts and lobules are spread throughout in the breast. Within adipose/fat tissues, a complex network of nerves, lymphatic vessels and nodes, blood vessels, ligaments and fibrous connective tissues serve to distribute elements throughout the body through the circulatory and lymphatic systems.

Breast cancer arises from uncontrolled cell proliferation in the breast, usually in the cells of lobules and terminal ducts, and is the most common cancer among women worldwide.[6] In the United States, breast cancer ranks as the most commonly diagnosed malignancy among women, accounting for 29.2% of all newly diagnosed cancers and 14.4% of female cancer deaths annually (Figure 2).[7,8] Breast cancer is a phenotypically diverse disease and consists of tumors with various molecular and pathologic characteristics, which are determinants for metastatic behavior and clinical outcome. Breast cancers are classified based on the histopathology of the tumor and extent of disease spread, into in situ and invasive types. Ductal carcinoma in situ (DCIS) is classified as a non-invasive cancer of the

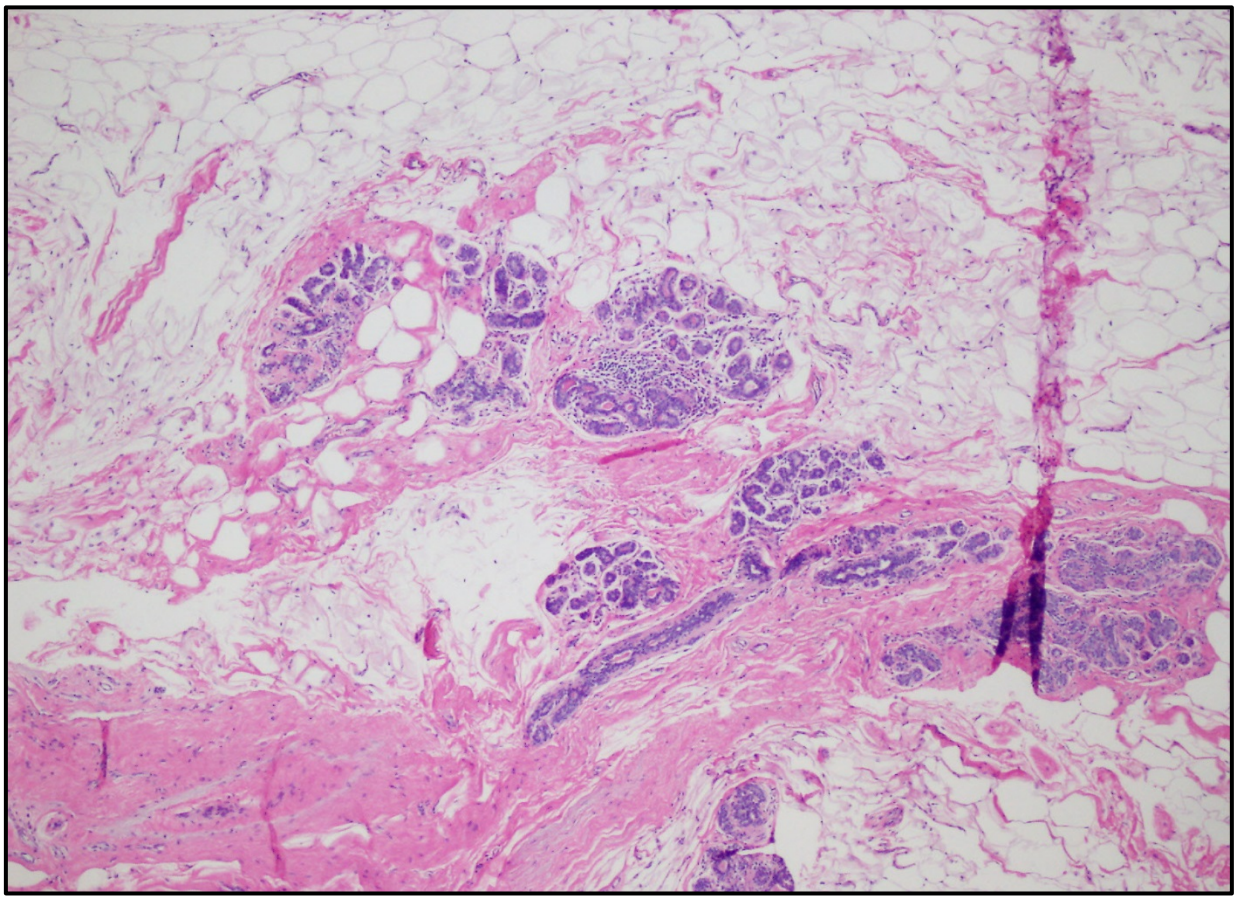

**Figure 1. Microscopic view of normal breast tissue.** Magnification: 100X. *Photograph courtesy of E. Abdulfatah, Wayne State University, Department of Pathology.*

**Estimated new cancer cases among females, 2016.**

| *Cancer site* | N | % |
|---|---|---|
| Breast | 246,660 | 29.2% |
| Lung and bronchus | 106,470 | 12.6% |
| Colon and rectum | 63,670 | 7.5% |
| Uterine corpus | 60,050 | 7.1% |
| Thyroid | 49,350 | 5.8% |
| Non-Hodgkin lymphoma | 32,410 | 3.8% |
| Melanoma of the skin | 29,510 | 3.5% |
| Leukemia | 26,050 | 3.1% |
| Pancreas | 25,400 | 3.0% |
| Kidney and renal pelvis | 23,050 | 2.7% |
| **All sites** | **843,820** | **100%** |

**Estimated cancer deaths among females, 2016.**

| *Cancer site* | N | % |
|---|---|---|
| Lung and bronchus | 72,160 | 25.6% |
| Breast | 40,450 | 14.4% |
| Colon and rectum | 23,170 | 8.2% |
| Pancreas | 20,330 | 7.2% |
| Ovary | 14,240 | 5.1% |
| Uterine corpus | 10,470 | 3.7% |
| Leukemia | 10,270 | 3.6% |
| Liver and intrahepatic bile duct | 8,890 | 3.2% |
| Non-Hodgkin lymphoma | 8,630 | 3.1% |
| Brain and other nervous system | 6,610 | 2.3% |
| **All sites** | **281,400** | **100%** |

**Figure 2. Leading cancer types for estimated new cancer cases and cancer deaths among females in the United States, 2016.** Cases are rounded to the nearest ten and excludes basal cell and squamous cell skin cancers and in situ carcinoma except urinary bladder.[7]

breast, where cells of the breast ducts have morphed abnormally into cancer cells. Unique to this type of breast cancer, these abnormal cells have not invaded through ductal walls into surrounding tissue of the breast and have low potential to spread (Figure 3A).[6] In contrast, the majority of diagnosed breast cancers are of the invasive type, where cancer cells have infiltrated through glandular or ductal walls into surrounding breast tissue (Figure 3B). Invasive breast cancers are clinically staged to determine the extent of disease spread throughout the primary site and regional lymph nodes, and the presence of cancer in distant organs. Clinical staging of breast cancer dictates clinical therapy regimens and patient prognosis.[6]

National Comprehensive Care Network (NCCN) guidelines for the treatment of invasive breast cancer outline systemic adjuvant therapies based on hormone receptor (HR) and human epidermal growth factor receptor 2 (HER2) statuses, where positive HR status is defined as expression of the estrogen (ER) and/or progesterone (PR) receptors.[9] Clinically, tumor cells are evaluated for these biological markers to approximate cancer molecular subtype based on expression profiling (Luminal A/B, HER2-enriched, basal). These molecular characteristics dictate course of treatment and therapies for patients. Thus, there is a critical need to understand the genetic and epigenetic abnormalities that are associated with the different types of breast cancer for the development of novel therapies.

Luminal breast cancers account for about 60% of all cases, are hormone receptor-positive (HR+), and can be classified based on HER2 status.[10] Luminal A breast cancers are HER2-negative (HR+/HER2-) and include ER+/PR+, ER+/PR-, and ER-/PR+ status. Their adjuvant treatment includes endocrine therapy with or without

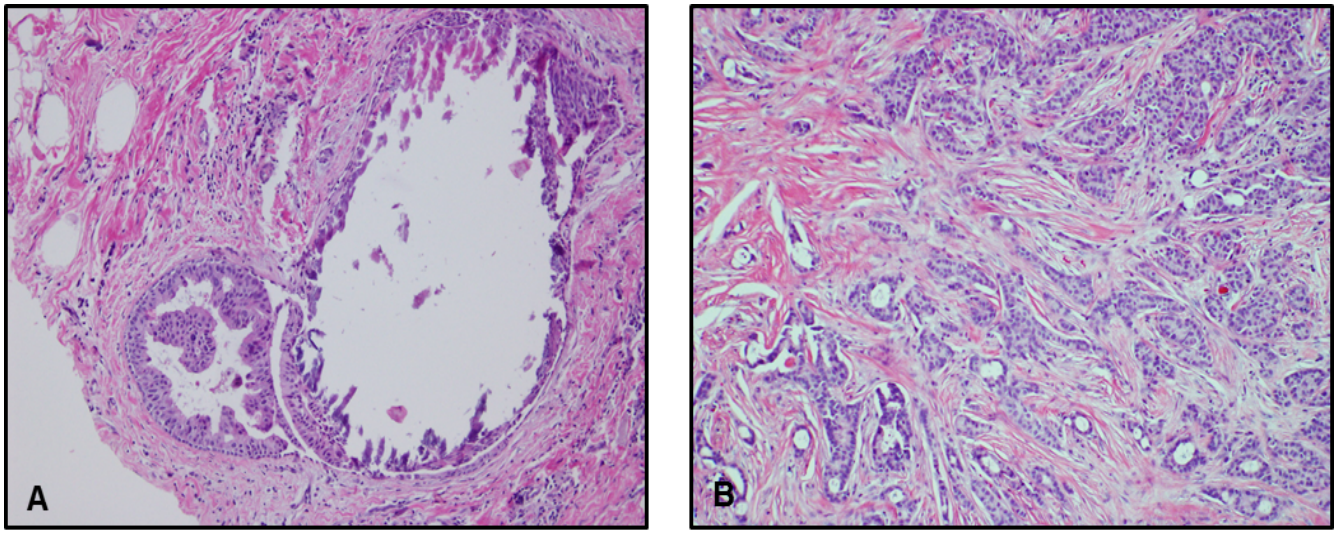


**Figure 3. Microscopic view of (A) ductal carcinoma in situ and (B) invasive ductal breast cancer tissues.** Magnification: 100X. *Photographs courtesy of E. Abdulfatah, Wayne State University, Department of Pathology.*

multimodality chemotherapy, based on tumor size, lymph node status, and, more recently, the 21-gene recurrence score.[9] Luminal A breast cancers are associated with the most favorable short-term prognosis due to favorable responses to endocrine therapy.[11,12] However, assessment of long-term prognosis demonstrates similar or worse overall survival for Luminal A cases as compared to other subtypes.[13]

Luminal B tumors tend to be more aggressive, demonstrate HER2-enrichment (HR+/HER2+), and encompass ER+/PR+, ER+/PR-, and ER-/PR+ cases. Recommended treatment for Luminal B tumors includes anthracycline-based multimodality chemotherapy containing trastuzumab, followed by a one-year course of trastuzumab and five years of endocrine therapy.[9] Together, Luminal breast cancer subtypes are associated with the best short-term prognoses for patients, attributable to favorable responses to hormonal therapy.[11,12]

Clinical differences among Luminal breast cancers can be attributed to the opposing effects of estrogen and progesterone on tumor progression. Estrogen supports tumor growth but suppresses progression, whereas progesterone supports tumor progression and is associated with more aggressive disease.[14] In the absence of estrogen signaling (ER- tumors), high progesterone levels in women have been shown to support tumor progression without opposition from estrogen.[15,16]

The most aggressive subtype of breast cancer is basal-like, where aberrant tumor cells are generally triple-negative, meaning estrogen receptor (ER)-negative, progesterone receptor (PR)-negative, and lacking human epidermal growth factor receptor 2 (HER2/ERBB2)-enrichment.[17,18] Basal-like tumors are associated with higher rates of metastasis and death, and the treatment for basal breast cancer consists of

standard chemotherapy regimens as no effective molecularly-targeted therapies have been developed.

Taken together, breast cancer is a disease consisting of over twenty histological subtypes and four molecular subtypes.[19] Unique presentations of these subtypes, in combination with variable risk factors of disease, make this heterogeneous disease more challenging to treat and profile for patients.[20] Thus, there is a critical need to develop a better understanding of the disease risks and disparities, as well as a better understanding of aggressive breast cancer subtypes, particularly among high risk populations, to ultimately yield novel discoveries that will impact clinical therapies and improve survival outcomes of patients.

# CHAPTER 2. DISPARITIES AMONG FEMALE PATIENTS DIAGNOSED WITH BREAST CANCER

## INTRODUCTION

*Racial/ethnic disparities in breast cancer incidence and survival*

Among breast cancer cases, racial/ethnic disparities persist, where higher incidence rates are observed among whites than blacks (128.1 versus 124.3 per 100,000, respectively) aged 45 years and older at diagnosis, although recent data suggest these rates are converging.[7] Some of this disparity can be attributed to differences in clinical characteristics, such that black women tend to be disproportionately diagnosed with early-onset disease (age < 35 years at diagnosis) and with more aggressive tumors.[7,21,22] Indeed, previous studies have demonstrated that among Luminal A tumors, race-associated biological factors contribute to poorer outcomes in blacks compared to white women.[23,24]

While overall age-adjusted breast cancer mortality has declined significantly since 1990, racial disparities have widened, as black women suffer from approximately 42% higher mortality rates compared to white or Hispanic women.[8,21,25] Differences in the uptake of cancer screening and access to high-quality treatment have also been suggested as contributory factors to racial disparities in breast cancer mortality, even in early-stage disease.[26-29] Population-based studies have reported that pre-menopausal black women under 50 years of age have approximately 1.5-fold higher incidence of triple negative breast cancers and poorer survival outcomes, partly attributed to limited systemic treatment options for this aggressive breast cancer subtype.[17,30-32]

Geographic variations also contribute to racial and ethnic disparities in breast cancer incidence and mortality. Among black women as represented by 18

Surveillance, Epidemiology, and End Results (SEER) program registries across the United States, age-adjusted incidence for in situ female breast cancer rates per 100,000 females are 31.54 and 31.02 for whites and blacks, respectively (Figure 4A). Rates in Detroit for female in situ breast cancers are higher, however, at 36.73 and 36.05 cases per 100,000 white and black females, respectively. This disparity persists among invasive breast cancers for females, as national rates of invasive breast cancer per 100,000 females are 127.96 and 125.19 for whites and blacks, respectively. In Detroit, incidence rates are 133.06 and 131.31 for white and black females (Figure 4B). This disparity in incidence rates for in situ and invasive breast cancers among white and black women in Detroit versus the United States demonstrate a unique population that needs to be further investigated to understand the etiology and development of tumors in this population. Further, the disparities observed between white and black women nationally are paralleled in Detroit and are consistent with prior studies that examine racial and ethnic disparities in breast cancer incidence and mortality.

*Age at breast cancer diagnosis and familial history of breast cancer*

Nearly 7% of all breast cancer cases in women are diagnosed before the age of 40 years,[33] and among women 20 to 39 years of age, breast cancer still ranks as the leading cause of cancer death. Women diagnosed with breast cancer at a young age tend to present with higher grade, hormone-insensitive tumors with more frequent spread to regional lymph nodes compared to older patients.[34] It is also established that patients age < 35 years whose surgical treatment consists of breast-conserving surgery and radiation have a greater risk of local recurrence at ten years compared to patients

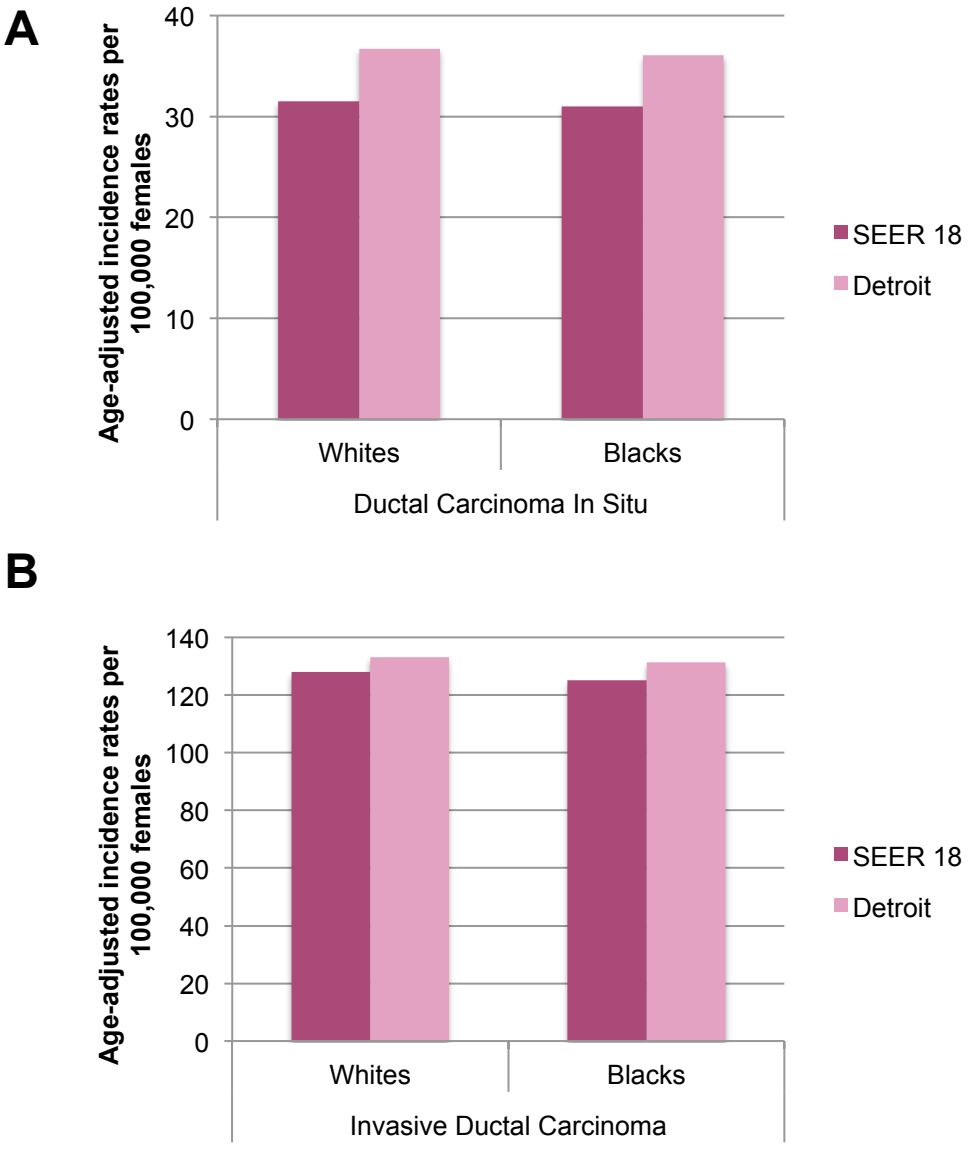


**Figure 4. Age-adjusted Surveillance, Epidemiology, and End Results (SEER) incidence rates per 100,000 females for (A) in situ and (B) invasive breast cancer by registry and race, 2009-2013.**

who opt for a mastectomy.[35] Further, women under 40 years of age at diagnosis are more likely to undergo mastectomy for breast cancer treatment compared to older women.[36,37] Independent of stage, histology, and extent of disease at breast cancer diagnosis, women under 40 years of age diagnosed with AJCC clinical stage I-II breast cancer have a lower rate of survival.[34,38-41] Young women with breast cancer also tend to consider consequences of treatment with regard to premature menopause, body image, fertility, and risk of secondary cancers.[42,43]

Young-onset breast cancer has also been shown to be associated with an increased familial risk of disease. Lalloo et al. reported that among thirty-seven women with strong family history of breast cancer diagnosed at age < 30 years, about half harbored a BRCA mutation.[44] By race, Churpek et al. observed that one in four women had a BRCA mutation among black patients with young-onset disease, a familial history of breast cancer, or triple-negative breast cancer.[45] BRCA mutation frequencies were also reported to be higher among black women, compared to whites, in a population-based Florida study of 396 women.[46] While heritable mutation frequencies may be higher among black women, reports have revealed that black women have a lower uptake of genetic testing services compared to whites.[47-50]

*Menopausal and socioeconomic status on breast cancer risk and outcomes*

Early menarche and late menopause are also health factors that contribute to breast cancer risk.[51] Progesterone levels are higher in pre-menopausal women, typically those diagnosed with breast cancer under the age of 50, compared to post-menopausal women over age 60.[52,53] Thus, post-menopausal women who naturally circulate high

levels of endogenous sex hormones have increased risk of breast cancer development compared to counterparts with low hormone levels.[54]

Access to screening, genetic testing, high-quality treatment, and delays in surgical care has also previously been attributed to disparities in breast cancer mortality.[27,28] Insurance status has also been shown to be a strong risk factor for poor outcomes among adolescents and young adults, particularly among tumor types that are responsive to early detection.[55] Differences in access to high-quality cancer care may contribute to disparities in breast cancer diagnosis and treatment.

# RESULTS:

## DISPARITIES IN SURGICAL THERAPY AMONG FEMALE PATIENTS WITH YOUNG ONSET EARLY-STAGE BREAST CANCER

*Description of study cohort*

The purpose of the study was to compare uptake of surgical therapy, specifically breast-conserving surgery (BCS), versus mastectomy among non-Hispanic white (NHW), non-Hispanic black (NHB), and Hispanic women under 40 years of age diagnosed with AJCC clinical stage I-II breast cancer, the majority of whom are treated first with surgery. A total of 6,449 incident early-stage breast cancer cases diagnosed in women age <40 years were identified from the Surveillance, Epidemiology, and End Results (SEER) database with race categorized as NHW, NHB, and Hispanic (4,013 NHW; 1,059 NHB; 1,377 Hispanic) who underwent surgical treatment from 2010 through 2013 (Figure 5). 10.7% of these young-onset breast cancer cases were diagnosed before the age of 30 (690 of 6,449 cases), and NHB women were diagnosed younger than NHWs and Hispanics (p-value = 0.002 and 0.0324, respectively) (Figure 5). NHW patients were more likely to be insured as compared to NHB or Hispanics (86.2% vs 69.4% and 66.2%, respectively; p-value = <0.0001 and <0.0001, respectively).

Consistent with previous reports,[17,31,32] triple negative breast cancers were significantly more frequent among NHBs compared to NHWs and Hispanics (26.6% vs 17.8% and 22.3%, respectively; p-value = <0.0001 and 0.002, respectively) (Figure 5). Also consistent with previous reports,[56-59] NHW and Hispanic patients tended to present with smaller primary tumors (p-value = 0.0004 and 0.0315, respectively) as well

| | NH White | | NH Black | | Hispanic | | p-value* | | |
|---|---|---|---|---|---|---|---|---|---|
| | N | % | N | % | N | % | NHW:NHB | NHW:H | NHB:H |
| Total | 4013 | | 1059 | | 1377 | | | | |
| **Age at Diagnosis** | | | | | | | 0.002 | 0.7237 | 0.0324 |
| <25 | 56 | 1.40% | 18 | 1.70% | 22 | 1.60% | | | |
| 25-29 | 352 | 8.80% | 109 | 10.30% | 133 | 9.70% | | | |
| 30-34 | 1087 | 27.10% | 336 | 31.70% | 368 | 26.70% | | | |
| 35-39 | 2518 | 62.70% | 596 | 56.30% | 854 | 62.00% | | | |
| Mean (std) | 34.9 (3.8) | | 34.5 (3.9) | | 34.7 (3.8) | | | | |
| **Subtype** | | | | | | | <0.0001 | 0.002 | 0.0844 |
| HR+/HER2- | 2163 | 53.90% | 511 | 48.30% | 720 | 52.30% | | | |
| HR+/HER2+ | 696 | 17.30% | 159 | 15.00% | 207 | 15.00% | | | |
| HR-/HER2+ | 231 | 5.80% | 56 | 5.30% | 76 | 5.50% | | | |
| HR-/HER2- | 714 | 17.80% | 282 | 26.60% | 307 | 22.30% | | | |
| Unknown | 209 | 5.20% | 51 | 4.80% | 67 | 4.90% | | | |
| **AJCC Clinical Stage** | | | | | | | <0.0001 | <0.0001 | 0.4038 |
| IA | 1587 | 39.50% | 317 | 29.90% | 426 | 30.90% | | | |
| IB | 145 | 3.60% | 35 | 3.30% | 33 | 2.40% | | | |
| IIA | 1341 | 33.40% | 429 | 40.50% | 535 | 38.90% | | | |
| IIB | 940 | 23.40% | 278 | 26.30% | 383 | 27.80% | | | |
| **Tumor Grade** | | | | | | | <0.0001 | <0.0001 | 0.5007 |
| I (Well differentiated) | 412 | 10.30% | 67 | 6.30% | 102 | 7.40% | | | |
| II (Moderately differentiated) | 1398 | 34.80% | 308 | 29.10% | 429 | 31.20% | | | |
| III (Poorly differentiated) | 2028 | 50.50% | 619 | 58.50% | 774 | 56.20% | | | |
| IV (Undifferentiated) | 15 | 0.40% | 7 | 0.70% | 10 | 0.70% | | | |
| Unknown | 160 | 4.00% | 58 | 5.50% | 62 | 4.50% | | | |
| **Tumor Size**[a] | | | | | | | 0.0004 | 0.3839 | 0.0315 |
| ≤5.0 cm | 3900 | 97.20% | 1006 | 95.00% | 1329 | 96.50% | | | |
| >5.0 cm | 113 | 2.80% | 53 | 5.00% | 45 | 3.30% | | | |
| **Nodal Involvement** | | | | | | | 0.4047 | 0.0559 | 0.452 |
| No | 2598 | 64.70% | 671 | 63.40% | 852 | 61.90% | | | |
| Yes | 1415 | 35.30% | 388 | 36.60% | 525 | 38.10% | | | |
| **Insurance** | | | | | | | <0.0001 | <0.0001 | 0.0186 |
| Insured | 3459 | 86.20% | 735 | 69.40% | 912 | 66.20% | | | |
| Uninsured | 90 | 2.20% | 43 | 4.10% | 90 | 6.50% | | | |
| Medicaid | 407 | 10.10% | 262 | 24.70% | 358 | 26.00% | | | |
| Unknown | 57 | 1.40% | 19 | 1.80% | 17 | 1.20% | | | |

*p-value calculations do not include unknown values.

NHW:NHB = Non-Hispanic White:Non-Hispanic Black

NHW:H = Non-Hispanic White: Hispanic

NHB:H = Non-Hispanic Black: Hispanic

[a]3 patients had unknown tumor size.

**Figure 5. Summary of clinicodemographic characteristics by race/ethnicity among women diagnosed with young-onset early-stage breast cancer; Surveillance, Epidemiology, and End Results (SEER) 18, 2010-2013.**

| | NH White | | NH Black | | Hispanic | | p-value* | | |
|---|---|---|---|---|---|---|---|---|---|
| | N | % | N | % | N | % | NHW:NHB | NHW:H | NHB:H |
| Total | 4013 | | 1059 | | 1377 | | | | |
| **Surgical Therapy**[b] | | | | | | | <0.0001 | 0.0042 | 0.2244 |
| BCS | 1251 | 31.20% | 399 | 37.70% | 487 | 35.40% | | | |
| Mastectomy | 2757 | 68.70% | 657 | 62.00% | 889 | 64.60% | | | |
| **Radiation Therapy** | | | | | | | 0.0374 | 0.9256 | 0.0671 |
| No | 2194 | 54.70% | 531 | 50.10% | 755 | 54.80% | | | |
| Yes | 1608 | 40.10% | 452 | 42.70% | 550 | 39.90% | | | |
| Unknown | 211 | 5.30% | 76 | 7.20% | 72 | 5.20% | | | |
| **BCS and Radiation°** | | | | | | | 0.0076 | <0.0001 | 0.165 |
| BCS only | 277 | 22.10% | 110 | 27.60% | 160 | 32.90% | | | |
| BCS and Radiation | 898 | 71.80% | 250 | 62.70% | 295 | 60.60% | | | |
| Unknown | 76 | 6.10% | 39 | 9.80% | 32 | 6.60% | | | |
| **Mastectomy and Radiation°** | | | | | | | 0.0059 | 0.0923 | 0.298 |
| Mastectomy only | 1914 | 69.40% | 418 | 63.60% | 594 | 66.80% | | | |
| Mastectomy and Radiation | 710 | 25.80% | 202 | 30.70% | 255 | 28.70% | | | |
| Unknown | 133 | 4.80% | 37 | 5.60% | 40 | 4.50% | | | |

*p-value calculations do not include unknown values.

NHW:NHB = Non-Hispanic White:Non-Hispanic Black

NHW:H = Non-Hispanic White: Hispanic

NHB:H = Non-Hispanic Black: Hispanic

BCS: Breast-conserving surgery/lumpectomy.

[b]9 patients had unknown information on type of surgical therapy.

°Among individuals who had BCS or mastectomy, respectively.

**Figure 6. Summary of treatment characteristics by race/ethnicity among women diagnosed with young-onset early-stage breast cancer; Surveillance, Epidemiology, and End Results (SEER) 18, 2010-2013.**

as lower grade tumors (p-value < 0.0001 and 0.0001, respectively) compared to NHB women (Figure 5). No significant differences were observed by race/ethnicity for lymph node involvement.

In terms of surgical therapy, both NHBs and Hispanics were significantly less likely to undergo mastectomy than NHW women (p-value = <0.0001 and 0.0042, respectively) (Figure 6). NHB patients were significantly more likely to undergo radiation treatment compared to NHWs (p-value = 0.0374). To further assess racial differences in therapeutic uptake, receipt of radiation therapy was evaluated among each surgical type. As presented in Figure 6, among women who received BCS, NHWs were significantly more likely to receive radiation as compared to NHBs and Hispanics (p-value = 0.0076 and <0.0001, respectively). In contrast, among women who underwent mastectomy, NHWs were significantly less likely than NHB to undergo radiation (p-value = 0.0059). Further stratification by stage at diagnosis demonstrated that among NHBs who underwent a mastectomy, 69.1% were diagnosed at stage II versus only 58.8% of NHW patients with stage II disease (p = <0.0001, data not shown). No significant differences were observed with Hispanic women.

*Racial/ethnic disparities in receipt of surgical therapy*

Odds ratios (OR) for receipt of mastectomy, adjusted for age, race/ethnicity, insurance status, subtype, stage, tumor size, lymph node involvement, and radiation therapy, are presented in Figure 7. NHB and Hispanic women were 24% and 21% less likely to undergo a mastectomy compared to NHW women (OR 0.76, 95% CI 0.63-0.90; OR 0.79, 95% CI 0.67-0.92; respectively). No significant differences in surgical therapy

were observed by breast cancer subtype. Compared to women aged 35 to 39 years, those diagnosed at younger ages were 1.33 to 2.23-fold more likely to have a mastectomy. Uninsured women were 23% less likely to undergo a surgical mastectomy (OR 0.77, 95% CI 0.65-0.92).

Stratification of cases by tumor subtype (using NHW cases as the referent) revealed that NHB and Hispanic individuals were less likely to undergo a mastectomy among all subtypes (Figure 8). NHB and Hispanic women were 65% and 50% less likely to undergo a mastectomy among women with HER2-enriched tumors, compared to NHW (OR 0.35, 95% CI 0.17-0.71; and OR 0.50, 95% CI 0.26-0.94, respectively). Significant differences were also observed between Hispanic and NHW women diagnosed with Luminal A (HR+/HER2-) tumors (OR 0.77, 95% CI 0.62-0.96). Among patients diagnosed with Luminal A (HR+/HER2-) and triple negative (HR-/HER2-) tumors, younger age was significantly associated with an increased likelihood to undergo a surgical mastectomy compared to BCS (Figure 8). By stage, women with AJCC stage II tumors were significantly more likely to undergo a mastectomy compared to stage I cases among Luminal A tumors (OR 1.63, 95% CI 1.31-2.01).

| | Mastectomy[§*] | |
|---|---|---|
| | OR (95% CI) | p-value |
| **Race/Ethnicity** | | |
| Non-Hispanic White | Ref | |
| Non-Hispanic Black | 0.76 (0.63-0.90) | 0.0019 |
| Hispanic | 0.79 (0.67-0.92) | 0.0037 |
| **Age at Diagnosis** | | |
| <25 | 2.23 (1.27-3.90) | 0.0051 |
| 25-29 | 1.44 (1.15-1.82) | 0.0019 |
| 30-34 | 1.33 (1.15-1.55) | 0.0001 |
| 35-39 | Ref | |
| **Subtype** | | |
| HR+/HER2- | Ref | |
| HR+/HER2+ | 0.99 (0.84-1.18) | 0.9471 |
| HR-/HER2+ | 1.21 (0.91-1.60) | 0.1825 |
| HR-/HER2- | 0.89 (0.75-1.04) | 0.1416 |
| **AJCC Clinical Stage** | | |
| I | Ref | |
| II | 1.33 (1.14-1.55) | 0.0003 |
| **Insurance** | | |
| Insured | Ref | |
| Medicaid | 1.24 (0.87-1.78) | 0.2372 |
| Uninsured | 0.77 (0.65-0.92) | 0.0037 |

*Model adjusted for race/ethnicity (NH White, NH Black, Hispanic), age at diagnosis (5 year groups), clinical subtype (Luminal A/B, HER2, Triple negative), stage (I/II), tumor size (≤5.0 cm versus >5.0 cm), nodal involvement (yes/no), radiation therapy (yes/no), insurance (insured, Medicaid, uninsured).
[§]Referent group is patients who received breast-conserving surgery.

**Figure 7. Adjusted odds ratios for early-stage breast cancer patient clinicodemographic characteristics, SEER 18, 2010-2013.**

| | Mastectomy[§]* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | HR+/HER2- | | HR+/HER2+ | | HR-/HER2+ | | HR-/HER2- | |
| | OR (95% CI) | p-value | OR (95% CI) | p-value | OR (95% CI) | p-value | OR (95% CI) | p-value |
| **Age at Diagnosis** | | | | | | | | |
| <25 | 4.17 (1.68-10.34) | 0.002 | 1.71 (0.61-4.78) | 0.304 | 5.28 (0.52-53.73) | 0.16 | 0.63 (0.19-2.12) | 0.4545 |
| 25-29 | 1.57 (1.13-2.19) | 0.0078 | 1.39 (0.84-2.32) | 0.2035 | 0.80 (0.32-2.01) | 0.6374 | 1.52 (0.94-2.46) | 0.086 |
| 30-34 | 1.30 (1.06-1.59) | 0.0115 | 1.23 (0.86-1.76) | 0.2505 | 0.95 (0.52-1.73) | 0.8729 | 1.64 (1.21-2.23) | 0.0015 |
| 35-39 | Ref | | Ref | | Ref | | Ref | |
| **Race/Ethnicity** | | | | | | | | |
| Non-Hispanic White | Ref | | Ref | | Ref | | Ref | |
| Non-Hispanic Black | 0.88 (0.68-1.14) | 0.3407 | 0.65 (0.43-1.00) | 0.051 | 0.35 (0.17-0.71) | 0.0033 | 0.77 (0.54-1.08) | 0.1261 |
| Hispanic | 0.77 (0.62-0.96) | 0.022 | 0.70 (0.47-1.04) | 0.0766 | 0.50 (0.26-0.94) | 0.0323 | 0.96 (0.68-1.35) | 0.797 |
| **AJCC Clinical Stage** | | | | | | | | |
| I | Ref | | Ref | | Ref | | Ref | |
| II | 1.63 (1.31-2.01) | <0.0001 | 1.01 (0.69-1.47) | 0.9698 | 1.07 (0.57-2.00) | 0.8303 | 1.01 (0.73-1.41) | 0.9362 |

*Model adjusted for race/ethnicity (NH White, NH Black, Hispanic), age at diagnosis (5 year groups), stage (I/II), tumor size (≤5.0 cm versus >5.0 cm), nodal involvement (yes/no), radiation therapy (yes/no), and insurance (insured, Medicaid, uninsured).
[§]Referent group is patients who received breast-conserving surgery.

**Figure 8. Adjusted odds ratios for early-stage breast cancer patient clinicodemographic characteristics by subtype, SEER 18, 2010-2013.**

# RESULTS:

## HER2 STATUS AND DISPARITIES IN LUMINAL BREAST CANCERS

*Differences among ER+/PR+, ER+/PR-, and ER-/PR+ HR-positive breast cancers*

Expression of the estrogen (ER) and/or progesterone (PR) receptors was defined as hormone-receptor positive (HR+) disease.[9] To study clinical, demographic and socioeconomic differences within Luminal A (HR+/HER2-) and Luminal B (HR+/HER2+) clinical breast cancer subtypes, 134,639 patients with HR+ and known HER2 receptor status were gathered from the Surveillance, Epidemiology, and End Results (SEER) database.[60] Among these patients, 118,285 (87.8%) cases were HR+/HER2- (Luminal A) (Figure 9), and 16,354 (12.2%) were HR+/HER2+ (Luminal B) tumors (Figure 10).

For Luminal A (HR+/HER2-) cases, 102,087 (86.3%) were ER+/PR+; 14,994 (12.7%) ER+/PR-; and 1,204 (1.0%) ER-/PR+ (Figure 9). Luminal A subtype patients demonstrated significantly different distributions by age, race/ethnicity, tumor size, AJCC clinical stage, and SES measures of poverty ($p<0.0001$). Specifically, Luminal A cases with ER-/PR+ status were more likely to be diagnosed at a younger age, to be non-Hispanic (NH) black or Hispanic, to live in counties with higher poverty, to have larger tumors, and to present with later stage disease (Figure 9).

Of the Luminal B (HR+/HER2+) tumors, 11,391 (69.7%) were ER+/PR+; 4,491 (27.4%) ER+/PR-; and 472 (2.9%) ER-/PR+ (Figure 10). Patients with Luminal B breast cancers had distributions that varied by age, race/ethnicity, and AJCC clinical stage. Compared with ER+/PR- and ER+/PR+ patients, Luminal B cases with ER-/PR+ status were more likely to be diagnosed at a younger age, to be NH Asian, Pacific Islander, or

| | | Luminal A (HER2-) | | | | | | p-value* | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | ER+ / PR+ | | ER+ / PR- | | ER- / PR+ | | ER+/PR+: | ER+/PR+: | ER+/PR-: |
| | N | N | % | N | % | N | % | ER+/PR- | ER-/PR+ | ER-/PR+ |
| **Total** | 118,285 | 102,087 | | 14,994 | | 1,204 | | | | |
| **Age at Diagnosis** | | | | | | | | <0.0001 | <0.0001 | <0.0001 |
| <50 | 21,777 | 19,299 | 18.9% | 2,081 | 13.9% | 397 | 33.0% | | | |
| 50-64 | 43,550 | 37,261 | 36.5% | 5,849 | 39.0% | 440 | 36.5% | | | |
| 65-74 | 28,705 | 24,836 | 24.3% | 3,670 | 24.5% | 199 | 16.5% | | | |
| ≥75 | 24,253 | 20,691 | 20.3% | 3,394 | 22.6% | 168 | 14.0% | | | |
| **Race/Ethnicity** | | | | | | | | <0.0001 | <0.0001 | <0.0001 |
| NH White | 85,717 | 74,701 | 73.2% | 10,297 | 68.7% | 719 | 59.7% | | | |
| NH Black | 10,540 | 8,371 | 8.2% | 1,930 | 12.9% | 239 | 19.9% | | | |
| Asian/Pacific Islander | 9,117 | 7,956 | 7.8% | 1,094 | 7.3% | 67 | 5.6% | | | |
| Am.Indian/AlaskaNative | 633 | 555 | 0.5% | 71 | 0.5% | 7 | 0.6% | | | |
| Hispanic | 11,429 | 9,766 | 9.6% | 1,496 | 10.0% | 167 | 13.9% | | | |
| Unknown | 849 | 738 | 0.7% | 106 | 0.7% | 5 | 0.4% | | | |
| **AJCC Stage** | | | | | | | | <0.0001 | <0.0001 | <0.0001 |
| 0-I | 64,491 | 57,029 | 55.9% | 7,018 | 46.8% | 444 | 36.9% | | | |
| II | 34,905 | 29,601 | 29.0% | 4,809 | 32.1% | 495 | 41.1% | | | |
| III | 11,352 | 9,329 | 9.1% | 1,868 | 12.5% | 155 | 12.9% | | | |
| IV | 4,979 | 3,997 | 3.9% | 900 | 6.0% | 82 | 6.8% | | | |
| Unknown | 2,558 | 2,131 | 2.1% | 399 | 2.7% | 28 | 2.3% | | | |
| **Tumor Size** | | | | | | | | 0.1503 | <0.0001 | <0.0001 |
| ≤0.5 cm | 10,091 | 8,724 | 8.5% | 1,316 | 8.8% | 51 | 4.2% | | | |
| >0.5 cm | 104,379 | 90,257 | 88.4% | 13,021 | 86.8% | 1,101 | 91.4% | | | |
| Unknown | 3,815 | 3,106 | 3.0% | 657 | 4.4% | 52 | 4.3% | | | |
| **Poverty Index** | | | | | | | | 0.0833 | <0.0001 | <0.0001 |
| Q1 | 28,333 | 24,551 | 24.0% | 3,516 | 23.4% | 266 | 22.1% | | | |
| Q2 | 27,526 | 23,778 | 23.3% | 3,513 | 23.4% | 235 | 19.5% | | | |
| Q3 | 28,566 | 24,569 | 24.1% | 3,736 | 24.9% | 261 | 21.7% | | | |
| Q4 | 33,837 | 29,170 | 28.6% | 4,226 | 28.2% | 441 | 36.6% | | | |
| Unknown | 23 | 19 | 0.0% | 3 | 0.0% | 1 | 0.1% | | | |

**p-value calculations do not include unknown values.*

**Figure 9. Patient and tumor characteristics of Luminal A invasive breast cancer cases.**[4] Summary of clinical and demographic characteristics of hormone receptor-positive (HR+), HER2- Luminal A breast cancers in women with invasive breast cancer: Surveillance, Epidemiology and End Results 18, 2010-2012. *Open access copyright permissions for this article permitted re-use of this figure from Cancer Medicine.*

| | | Luminal B (HER2+) | | | | | | p-value* | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | ER+ / PR+ | | ER+ / PR- | | ER- / PR+ | | ER+/PR+: | ER+/PR+: | ER+/PR- : |
| | N | N | % | N | % | N | % | ER+/PR- | ER-/PR+ | ER-/PR+ |
| **Total** | 16,354 | 11,391 | | 4,491 | | 472 | | | | |
| **Age at Diagnosis** | | | | | | | | <0.0001 | 0.7842 | <0.0001 |
| <50 | 4,621 | 3,518 | 30.9% | 950 | 21.2% | 153 | 32.4% | | | |
| 50-64 | 6,631 | 4,431 | 38.9% | 2,016 | 44.9% | 184 | 39.0% | | | |
| 65-74 | 2,886 | 1,951 | 17.1% | 855 | 19.0% | 80 | 16.9% | | | |
| ≥75 | 2,216 | 1,491 | 13.1% | 670 | 14.9% | 55 | 11.7% | | | |
| **Race/Ethnicity** | | | | | | | | 0.3661 | <0.0001 | <0.0001 |
| NH White | 10,793 | 7,545 | 66.2% | 2,970 | 66.1% | 278 | 58.9% | | | |
| NH Black | 1,918 | 1,306 | 11.5% | 555 | 12.4% | 57 | 12.1% | | | |
| Asian/Pacific Islander | 1,491 | 1,038 | 9.1% | 400 | 8.9% | 53 | 11.2% | | | |
| Am.Indian/AlaskaNative | 105 | 63 | 0.6% | 30 | 0.7% | 12 | 2.5% | | | |
| Hispanic | 1,921 | 1,352 | 11.9% | 503 | 11.2% | 66 | 14.0% | | | |
| Unknown | 126 | 87 | 0.8% | 33 | 0.7% | 6 | 1.3% | | | |
| **AJCC Stage** | | | | | | | | <0.0001 | <0.0001 | 0.003 |
| 0-I | 6,383 | 4,534 | 39.8% | 1,710 | 37.9% | 139 | 29.2% | | | |
| II | 5,792 | 4,073 | 35.8% | 1,536 | 34.2% | 183 | 38.8% | | | |
| III | 2,463 | 1,697 | 14.9% | 676 | 15.1% | 90 | 19.1% | | | |
| IV | 1,260 | 781 | 6.9% | 437 | 9.7% | 42 | 8.9% | | | |
| Unknown | 456 | 306 | 2.7% | 132 | 2.9% | 18 | 3.8% | | | |
| **Tumor Size** | | | | | | | | <0.0001 | 0.1152 | 0.7848 |
| ≤0.5 cm | 1,348 | 861 | 7.6% | 443 | 9.9% | 44 | 9.3% | | | |
| >0.5 cm | 14,178 | 9,983 | 87.6% | 3,800 | 84.6% | 395 | 83.7% | | | |
| Unknown | 828 | 547 | 4.8% | 248 | 5.5% | 33 | 7.0% | | | |
| **Poverty Index** | | | | | | | | 0.1446 | 0.6807 | 0.3295 |
| Q1 | 3,648 | 2,514 | 22.1% | 1,025 | 22.8% | 109 | 23.1% | | | |
| Q2 | 3,717 | 2,585 | 22.7% | 1,030 | 22.9% | 102 | 21.6% | | | |
| Q3 | 3,814 | 2,635 | 23.1% | 1,078 | 24.0% | 101 | 21.4% | | | |
| Q4 | 5,172 | 3,654 | 32.1% | 1,358 | 30.2% | 160 | 33.9% | | | |
| Unknown | 3 | 3 | 0.0% | 0 | 0.0% | 0 | 0.0% | | | |

**p-value calculations do not include unknown values.*

**Figure 9. Patient and tumor characteristics of Luminal B invasive breast cancer cases.**[4] Summary of clinical and demographic characteristics of hormone receptor-positive (HR+), HER2+ Luminal B breast cancers in women with invasive breast cancer: Surveillance, Epidemiology and End Results 18, 2010-2012. *Open access copyright permissions for this article permitted re-use of this figure from Cancer Medicine.*

Hispanic, and to be diagnosed at later stages of disease (Figure 10). Among Luminal B tumors, no significant differences in SES were observed, measured with census tract county-level poverty (p=0.1446, p=0.6807, and p=0.3295, respectively).

An additional 19,020 breast cancer patients had borderline or unknown status for ER, PR, or HER2, and were not included in this analysis. Among these women, 67.7% were NH white (12,876 cases), 11.7% were NH black (2,227 cases), 10.9% were Hispanic (2,068 cases), 7.4% were Asian or Pacific Islander (1,404 cases), 0.5% were American Indian/Alaska Native (91 cases), and 1.9% (354 cases) had unknown race/ethnicity (data not shown). 83.0% of unknown cases were diagnosed with invasive breast cancer at age 50 years and older (15,793 cases). 61.5% of cases were diagnosed with early stage (AJCC stage 0-II) disease (11,701 cases), 17.8% were diagnosed with later stage (AJCC stage III-IV) cancer, and 20.1% (3,929 cases) did not have information on AJCC stage. 59.1% of cases (11,242 cases) resided in areas where at least 31% of residents were living in poverty (Q3 or Q4). Together, cases with unknown receptor status tended to be NH white, older, and diagnosed with early stage disease (data not shown).

Overall, it was observed that, regardless of HER2 status, ER-/PR+ cases were more likely to be diagnosed in young patients (age < 50 years) and to present with later stage (stage III-IV) disease, but were less likely to be non-Hispanic white, as compared to ER+/PR- or ER+/PR+ patients.

*Socioeconomic disparities in Luminal A breast cancers*

Significant differences in area-based poverty were also noted among patients with Luminal A (HR+/HER2-) tumors. To further explore the relationship of socioeconomic status and HR+ status (ER+/PR+, ER+/PR-, or ER-/PR+) within each breast cancer subtype, multinomial logistic regression models were used. Because race and age are associated with socioeconomic status, analyses were adjusted for these variables to determine whether socioeconomic disparities persisted in this population.

Figure 11 summarizes results from models adjusted for area-based poverty, age, and race/ethnicity. Using ER+/PR+ tumors as the referent outcome in each subtype and poverty quartile Q1 as the referent covariable, it was found that women with Luminal A breast cancer who live in counties with higher poverty were more likely to be diagnosed with ER-/PR+ disease (Q4: OR = 1.20, 95% CI = 1.03-1.40) (Figure 11). Women diagnosed with ER-/PR+ Luminal A disease were 1.7-fold more likely to under 50 years of age compared to women with ER+/PR+ tumors in these area-based poverty-adjusted models. In addition, NH blacks were at an increased risk of being diagnosed with ER-/PR+ Luminal A breast cancers (OR = 2.62, 95% CI = 2.25-3.05) (Figure 11). Notably, age, race and poverty were not associated with ER-/PR+ disease in Luminal B breast cancers (Figure 11).

| | Area-Based Poverty[§] | | | |
|---|---|---|---|---|
| | Luminal A (HER2-)[◊] | | Luminal B (HER2+)[₡] | |
| | ER+ / PR- | ER- / PR+ | ER+ / PR- | ER- / PR+ |
| | OR (95% CI) | OR (95% CI) | OR (95% CI) | OR (95% CI) |
| **Age at Diagnosis, years** | | | | |
| <50 | 0.67 (0.64-0.71) | 1.68 (1.47-1.93) | 0.59 (0.54-0.65) | 1.01 (0.81-1.26) |
| 50-64* | 1 | 1 | 1 | 1 |
| 65-74 | 0.96 (0.91-1.00) | 0.70 (0.59-0.82) | 0.97 (0.88-1.07) | 0.99 (0.75-1.29) |
| ≥75 | 1.08 (1.03-1.13) | 0.72 (0.60-0.86) | 1.01 (0.90-1.12) | 0.94 (0.69-1.28) |
| **Race/Ethnicity** | | | | |
| NH White* | 1 | 1 | 1 | 1 |
| NH Black | 1.75 (1.66-1.85) | 2.62 (2.25-3.05) | 1.14 (1.02-1.28) | 1.18 (0.88-1.58) |
| Asian/Pacific Islander | 1.05 (0.98-1.12) | 0.77 (0.59-0.99) | 1.03 (0.91-1.17) | 1.36 (1.00-1.84) |
| Am. Indian/Alaska Native | 0.96 (0.75-1.23) | 1.22 (0.57-2.57) | 1.27 (0.82-1.97) | 5.15 (2.74-9.67) |
| Hispanic | 1.19 (1.12-1.26) | 1.48 (1.25-1.76) | 1.04 (0.93-1.16) | 1.31 (0.99-1.73) |
| **Poverty Index** | | | | |
| Q1* | 1 | 1 | 1 | 1 |
| Q2 | 1.02 (0.97-1.07) | 0.89 (0.75-1.07) | 0.96 (0.87-1.07) | 0.90 (0.68-1.19) |
| Q3 | 1.01 (0.96-1.06) | 0.88 (0.74-1.05) | 0.97 (0.88-1.08) | 0.88 (0.66-1.16) |
| Q4 | 0.93 (0.89-0.98) | 1.20 (1.03-1.40) | 0.88 (0.79-0.97) | 0.98 (0.76-1.27) |

[§]Model adjusted for age, race, and poverty (quartiles).

[◊]Referent group is HER2-/ER+/PR+.

[₡]Referent group is HER2+/ER+/PR+.

*Referent covariable.

**Figure 11. Adjusted odds ratios for patient demographics and socioeconomic status by hormone receptor-positive (HR+) breast cancers.**[4] *Open access copyright permissions for this article permitted re-use of this figure from Cancer Medicine.*

## DISCUSSION

Together, these studies have identified disparities in breast cancer incidence that could contribute to differences in survival outcomes for patients at higher-risk of disease development. In particular, these studies focused on: (1) racial/ethnic differences in surgical treatment among early-stage young onset breast cancers; and (2) associations among Luminal breast cancers and socioeconomic status as assessed by age, race/ethnicity, and a measurement of county-level poverty and found that within Luminal A (HR+/HER2-) and Luminal B (HR+/HER2+) cancers, clinical and demographic characteristics varied.

*Disparities in surgical therapy among patients with young onset early-stage breast cancer*

To minimize known disparities in screening utilization by race/ethnicity,[61,62] the first study population was restricted to women diagnosed with breast cancer who do not undergo routine mammographic screening. This population-based study of early-stage breast cancers diagnosed in women aged < 40 years identified age-related, racial, and ethnic disparities in uptake of surgical therapy. Both NHBs and Hispanics were less likely to undergo a mastectomy among all breast cancer subtypes compared to NHW women, with the largest disparities observed among HER2-enriched (HR-/HER2+) tumors.

Regardless of race, younger age at diagnosis was associated with increased uptake of mastectomy. Specifically, compared to women aged 35 to 39 years, it was observed that women diagnosed at younger ages (20-34 years of age) were

significantly more likely to undergo mastectomy regardless of race/ethnicity, consistent with surgical trends by age at diagnosis.[36,37] Women diagnosed with breast cancer at a young age tend to present with higher grade, hormone-insensitive tumors with more frequent spread to regional lymph nodes compared to older patients.[35] It is also established that patients age < 35 whose surgical treatment consists of BCS and radiation have a greater risk of local recurrence at ten years compared to patients who opt for a mastectomy.[35]

Fear of recurrence, avoidance of adjuvant side effects, clinical indicators, and perceived favorable survival outcomes have been reported as patient reasons for electing surgical mastectomy.[63-66] Indeed, the perception that more radical surgery results in improved chances of survival persists even though several large multi-institutional randomized trials have not shown a difference in overall survival between patients who undergo breast-conserving surgery and those who undergo mastectomy[67-74] and there is some evidence suggesting that women who undergo breast-conserving surgery plus radiation might have a slightly higher breast cancer-specific survival rate.[75,76]

In the first study, it was observed that NHB patients were significantly more likely to undergo radiation therapy, which would be expected in a population significantly more likely to undergo BCS than mastectomy as the primary treatment for breast cancer. However, the increase in radiation treatment in the NHB is not accounted for by differences in mastectomy rate. The differences between NHW and NHB for surgical treatment differ by over 5 percent of the population, with larger disparities observed among women who underwent breast-conserving surgery. Indeed, these findings are

consistent with other studies that indicate NHB women are less likely to receive radiation after BCS.[77,78]

The focus on early-stage breast cancer cases in women diagnosed < 40 years, and examination of subtype-specific differences in uptake of surgery reduced the impact that clinical differences may have on patient treatment decisions. Neoadjuvant chemotherapy is commonly used prior to surgery for TNBC and HER2+ breast cancer in order to facilitate BCS or, in the case of HER2+ tumors, because HER2-targeted neoadjuvant chemotherapy regimens containing trastuzumab ± pertuzumab can result in high pathologic complete response rates.[79,80] Pathologic complete response after neoadjuvant chemotherapy is associated with improvements in overall survival and with long-term patient outcomes.[81,82] While SEER does not capture information on systemic therapy, no significant differences were observed by race/ethnicity in the uptake of surgical mastectomy among TNBC cases when stratified by subtype.

The racial/ethnic disparities in surgical therapy that were observed persist even after adjustment, as uninsured women were significantly less likely to undergo a surgical mastectomy compared to insured patients. Consistent with these observations, studies evaluating the cost comparison of mastectomy versus BCS for early-stage breast cancer found higher short-term costs but lower long-term costs for BCS than mastectomy.[83,84] Cost and access to high-quality care could potentially contribute to disparities in surgery treatment decisions for breast cancer. While differences in exposure profiles (e.g., smoking history), medical history (e.g., menarche, parity, obesity), or genetic susceptibility were unavailable for study in SEER, limiting this study

population to breast cancers diagnosed in women aged < 40 years allowed for reduction of menopausal status as an associated breast cancer risk in this study.

Patient preferences play a critical role in electing a surgical mastectomy for breast cancer, with increased patient involvement in surgery treatment decisions being associated with a greater likelihood of mastectomy as the surgical treatment for breast cancer.[63,85,86] A recent study by Thomas, et al. among invasive breast cancer cases across all ages and cancer stage from the National Cancer Data Base (NCDB) noted similar findings, with increased rates in utilization of BCS among NHBs.[87] This study is unique in that the cohort was limited to young-onset (diagnosed at age<40 years) cases of early-stage breast cancer where surgery is the primary treatment, and observed that NHB and Hispanics were significantly less likely to undergo mastectomy compared to NHW women. These results appear inconsistent with those of Katz, et al.,[85] who observed that NHB and Hispanic women were more likely to receive a recommendation for mastectomy and received less information about BCS. However, they also observed racial differences in the surgical decision making process in which NHB and Hispanics were more likely to undergo mastectomy when they perceived the primary decision-maker to be the surgeon compared to NHW women, who were much less likely to undergo mastectomy in that situation.[85] Reasons for this unexpected propensity of NHWs compared to NHBs and Hispanics to choose more radical surgery, even though it does not improve overall survival, are unclear, but could involve fear of recurrence and better access to reconstructive surgery.[63,65-67] Patient preferences among racial groups can also impact surgical decision making, including importance of having breasts and overall interest in breast preservation. In a recent study by Jagsi, et al.,[88]

NHB patients were less likely to undergo contralateral prophylactic mastectomy compared to NHWs, although uptake of this surgical procedure was low, particularly when a patient reported that their surgeon recommended against it. Unfortunately, with SEER data alone these discrepancies cannot be addressed. This is an area for future investigation to understand how patient involvement in shared decision-making regarding the treatment of their breast cancer can be maintained without leading to overtreatment.

*HER2 status and disparities in Luminal breast cancers*

In this second study, associations were examined among Luminal breast cancers and socioeconomic status as assessed by age, race/ethnicity, and a census level measurement of county-level poverty and found that within Luminal A (HR+/HER2-) and Luminal B (HR+/HER2+) cancers, clinical and demographic characteristics varied. Consistent with the findings that clinical differences among luminal breast cancers can be attributed to the opposing effects of estrogen and progesterone on tumor progression,[12-14] it was observed in this study that regardless of HER2 status, women with ER-/PR+ tumors were more likely to present with later stage (stage III-IV) disease compared to ER+/PR+ or ER+/PR- cases. These results suggest that differences in HR+ (ER-/PR+ versus ER+/PR- or ER+/PR+) tumor biology are likely to be clinically significant and play a role in breast cancer disease, regardless of HER2 status.

Demographic characteristics of patients, including age, also varied within each Luminal breast cancer subtype. Progesterone levels are higher in pre-menopausal women, typically those diagnosed with breast cancer under the age of 50 years,

compared to post-menopausal women over the age of 60.[52,53] This study showed that women under the age of 50 were at an increased risk of developing ER-/PR+ Luminal A disease, while women over the age of 60 were at a decreased risk compared to ER+/PR+ disease. This observation is consistent with reports that high progesterone levels (occurring only in the luteal phase and in pregnancy) will induce breast cancer cell invasiveness and metastasis in the absence of estrogen or the estrogen receptor. In contrast, age was not associated with increased risk of ER-/PR+ disease among Luminal B cases. This observation may be explained by findings using experimental models, that overexpression of HER2 supports aggressive tumor growth in luminal breast cancers.[89]

Racial and ethnic differences were also noted among Luminal A cancers, as non-Hispanic black women were most likely to develop ER-/PR+ disease. Indeed, previous studies have demonstrated that among Luminal A tumors, race-associated biological factors contribute to poorer outcomes in black women compared to non-Hispanic white women.[23,24] Among women with Luminal B tumors, non-Hispanic Asian or Pacific Islander or American Indian/Alaska Native individuals were at an increased risk of developing ER-/PR+ tumors. However, caution should be taken in interpreting results in these racial/ethnic categories due to small sample size. Earlier findings have observed that Hispanic white women are more likely to present with more aggressive tumors and be diagnosed with ER+/PR- disease.[90] In this study, Hispanic individuals were at an increased risk of developing ER+/PR- or ER-/PR+ disease compared to ER+/PR+ status in both Luminal A and B breast cancer subtypes after adjusting for age, race, and area-based poverty.

While clinical and patient characteristics differed within each luminal breast cancer subtype, disparities in socioeconomic status were found to persist only among Luminal A tumors. Luminal A breast cancers are associated with the most favorable short-term prognosis due to favorable responses to endocrine therapy.[11,12] However, assessment of long-term prognosis demonstrates similar or worse overall survival for Luminal A cases as compared to other subtypes.[13] In this study, among Luminal A cases, it was observed that ER-/PR+ disease was associated with residing in areas of higher poverty even after adjusting for age and race/ethnicity. Caution should be exercised in the interpretation for results of ER-/PR+ cases due to small sample size. The relationship between socioeconomic status and development of Luminal A breast cancers demonstrates that while race and SES are correlated, each plays an independent role in contributing to disease among Luminal A tumors. These disparities were not observed among Luminal B tumors, suggesting that HER2 status may be associated with risk factors that affect the socioeconomic status of patients.

This study is the first to assess disparities among hormone receptor-positive (HR+) breast cancer subtypes in the context of HER2 status using SEER patient data. Previous studies have investigated the role of HR+ status without HER2 information, or have analyzed the role by breast cancer subtype. For example, a recent study by Parise, et al. that used the California Cancer Registry found that socioeconomic status moderately altered racial disparities and risk of mortality in particular breast cancer subtypes.[91] Similar results have been observed among other tumor sites.[92] While these findings suggest differences within Luminal A or B breast cancer subtypes, available data for HER2 are limited to three years of diagnosis, which does not allow further

analysis for SES disparities in mortality at this time. The use of data from the population-based SEER program is a strength of this study, as it allows for the inclusion of a considerable number of pathologically verified cases making these results more generalizable to the larger United States population. Another strength is that these data are of high quality and database entries are standardized and continuously monitored for accuracy. However, use of SEER data does have limitations. Family history, lifestyle-related factors (eg. obesity, reproductive factors, and environmental exposures), modality of diagnosis, and chemotherapy data, are not available for study. Patients whose ER, PR, or HER2 status was unknown were also unable to be examined.

Scientific, clinical, and public health implications can be inferred from this study. First, the findings are consistent with preclinical and clinical data regarding the opposing effects of the estrogen and progesterone receptors in breast cancer growth and progression. Further research is needed to analyze the opposing effects of hormone receptors in the context of HER2 status. Hormone receptor-positive (HR+) breast cancers are currently defined as ER and/or PR-positive tumors. In addition, these observations suggest that additional work is required to assess clinical differences observed, particularly between ER-/PR+ and ER+/PR- breast cancers, as this information could potentially be used to improve systemic adjuvant treatments. Disparities in socioeconomic status among Luminal A (HR+/HER2-) breast cancer patients may be associated with risks of recurrence and mortality, and identifying barriers in patient access to medical care can seek to improve patient outcomes in underserved, high-risk populations.

*Conclusions*

Taken together, these findings conclude that: (1) for women age<40 years diagnosed with early-stage breast cancer, NHB and Hispanic women are considerably less likely to undergo mastectomy compared to NHWs, a disparity that was most striking for HER2-enriched tumors; (2) while race/ethnicity and socioeconomic status are correlated, each plays an independent role in contributing to disease among Luminal A tumors.

Identification of age-related, racial, and ethnic disparities in uptake of surgical therapy for breast cancer contribute to improved understanding of the reasons for NHWs to choose more radical surgery, even though it does not improve overall survival. Further, detailed investigations of differences in tumor biology and association of race/ethnicity and socioeconomic status among hormone receptor-positive breast cancers, particularly those with HER2-negative status (HR+/HER2-), may lead to the identification of additional prognostic markers, direct resources to underserved populations for screening, and improve adjuvant treatments to better long-term patient outcomes.

## PATIENTS AND METHODS

*Data Sources and Case Selection*

Data were obtained from the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) program.[93] The SEER program collects cancer incidence and mortality data from 18 population-based cancer registries covering approximately 30 percent of the US population.[7] SEER*Stat is a free program provided by SEER to access and analyze information in the publically available database. The SEER 18 incidence dataset includes information from the following registries: Alaska Native Tumor Registry, Arizona Indians, Cherokee Nation, Connecticut, Detroit, Atlanta, Greater Georgia, Rural Georgia, San Francisco-Oakland, San Jose-Monterey, Greater California, Hawaii, Iowa, Kentucky, Los Angeles, Louisiana, New Jersey, New Mexico, Seattle-Puget Sound, and Utah. The dataset used is publicly available and was exempt from human subjects review. Therefore, an Institutional Review Board (IRB) protocol was not required for these studies.

*Disparities in surgical therapy among female patients with young onset early-stage breast cancers analysis*

The analysis was restricted to women diagnosed during the years 2010 to 2013 in this study (n=244,819), as diagnosis year 2010 is the first year for which HER2 (human epidermal growth factor receptor 2) status data are available. Study data, including breast cancer subtype, demographic characteristics, tumor stage, and surgical therapy were identified across SEER registries using standardized coding protocols based on pathology reports and hospital medical records. Case patients diagnosed with

adenocarcinomas, including ductal, lobular, and mixed neoplasm histology, were included (n=799 excluded). Patients aged ≥40 years at diagnosis were excluded (n=232,761), and analysis was limited to women diagnosed with American Joint Commission of Cancer (AJCC) clinical stage I-II tumors (n=3,309 excluded), with intent concentrate on patients whose primary therapy is most commonly surgery. Patients with self-reported race/ethnicity classified as non-Hispanic white (NHW), non-Hispanic African American or black (NHB), and Hispanic were included; subjects whose race was coded as Asian or Pacific Islander, American Indian or Alaskan Native, other, or unknown were not included in this analysis due to small sample size (n=1,137). Patients who did not undergo surgical therapy were also excluded (n=362). The final cohort thus consisted of 6,449 female patients with primary young-onset early-stage breast cancer.

Clinical and demographic variables examined included: age at diagnosis (5-year groups), AJCC stage (IA-B, IIA-B), surgical therapy (BCS or mastectomy), tumor size (≤5.0 cm versus >5.0 cm), nodal involvement (yes/no), radiation therapy (yes/no), and insurance status (insured, uninsured, Medicaid). Breast cancer subtype was derived from SEER based on joint hormone receptor (HR) and HER2 status, and tumors were classified into four mutually-exclusive categories: HR+/HER2- (Luminal A), HR+/HER2+ (Luminal B), HR-/HER2+ (HER2-enriched), and HR-/HER2- (triple negative). Surgical therapy (BCS versus mastectomy) was categorized based on SEER program surgical summary codes. BCS was defined as removal of the gross primary tumor and some of the breast tissue and included SEER surgery codes 20-24: partial mastectomy, lumpectomy, re-excision of the biopsy site and segmental mastectomy. Mastectomy was defined as removal of all breast tissue with or without removal of the nipple-areolar

complex. Mastectomy procedures included SEER surgery codes 30, 40-76, 80: subcutaneous (nipple sparing) mastectomy; total (simple) mastectomy with or without removal of uninvolved contralateral breast; modified radical mastectomy; bilateral mastectomy; radical mastectomy; extended radical mastectomy; and mastectomy, NOS.

Differences in demographic and tumor characteristics by race were examined by chi-square tests for categorical variables and t-tests for continuous variables. Logistic regression was used to estimate odds ratios (ORs) and 95% confidence intervals (CI) to quantify associations between surgical therapy (mastectomy as the outcome with BCS as referent) and various demographic/clinical factors. Demographic variables assessed included: age (5 year groups), race/ethnicity (NHW, NHB, Hispanic), and insurance status (insured, uninsured, Medicaid). Clinical variables assessed included: breast cancer subtype and AJCC stage. Models were adjusted for age, race/ethnicity, stage, tumor size, nodal involvement, radiation therapy, and insurance status; based on patients having complete information for each of these covariables. Analyses were then stratified by subtype.

*HER2 and socioeconomic disparities in luminal breast cancers analysis*

A case listing session in SEER*Stat was run on the SEER 18 incidence dataset to obtain demographic, tumor characteristics, and socioeconomic information on breast cancers. Women with invasive breast cancer diagnosed from 2010 to 2012 were included in this study. Diagnosis year 2010 is the first year for which HER2 status data were available. Study data, including ER, PR, and HER2 status, demographic

characteristics, and tumor stage, were identified across SEER registries using standardized coding rules based on pathology reports and hospital medical records. Case patients diagnosed with nevi and melanomas; soft tissue tumors and sarcomas, NOS; fibromatous neoplasms; lipomatous neoplasms; myomatous neoplasms; fibroepithelial neoplasms; synovial-like neoplasms; blood vessel tumors; osseous and chondromatous neoplasms; miscellaneous bone tumors; gliomas; nerve sheath tumors; and granular cell tumors & alveolar soft part sarcoma histologies were excluded. To assess breast cancer subtype, tumors were classified into six mutually exclusive categories: ER+/PR+/HER2-; ER+/PR-/HER2-; ER-/PR+/HER2-; ER+/PR+/HER2+; ER+/PR-/HER2+; ER-/PR+/HER2+. Analysis was restricted to exclude patients without positive/negative ER, PR, or HER2 statuses (n=19,020). Case patients with ER-/PR-/HER2+ (HER2-enriched), and ER-/PR-/HER2- (basal-like) receptor status were also excluded (n=26,736). The final analytic data set consisted of 134,639 breast cancer patients.

The variables of interest included age at diagnosis (<50, 50-64, 65-74, 75+ years), race/ethnicity (non-Hispanic (NH) white, non-Hispanic black, non-Hispanic Asian or Pacific Islander, non-Hispanic American Indian, Hispanic), American Joint Commission of Cancer (AJCC) clinical stage, and tumor size (≤0.5 cm, >0.5 cm). An approximation of socioeconomic status (SES) was evaluated using contextual measures of area-based poverty. The percent of persons and families whose incomes were below 200% of the poverty level were calculated in SEER using county attribute data from the US Census Bureau's 2008-2012 American Community Survey. The cohort of breast cancer cases was categorized into quartiles based on the percentage

of persons with incomes below 200% of the poverty level distribution within the data set. Quartile 1 (Q1) (<24%) represented the fourth of the cohort that reside in areas where less than 24% of residents are low income. Quartile 2 (Q2) (24%≤31% low income), quartile 3 (Q3) (31%≤39%), and quartile 4 (Q4) (>39%) were defined such that Q4 contained the fourth of the cohort that reside in areas with the highest proportion of low income residents. Follow-up for each case was current within 22 months of the annual submission date (November 1, 2014).

Pair-wise comparisons of hormone receptor categories by demographic and tumor characteristics were examined using chi-square tests. Multinomial logistic regression was used to estimate odds ratios (OR) and 95% confidence intervals (CI) to quantify associations between breast cancer hormone receptor status and various demographic factors. Analyses were stratified by HER2 status and the reference outcome category was ER+/PR+ disease. Demographics variables assessed, included: age at diagnosis (<50, 50-64, 65-74, 75+), race/ethnicity (NH white, NH black, NH Asian or Pacific Islander, NH American Indian, Hispanic), and area-based poverty level (quartiles). These factors were adjusted for age, race/ethnicity and poverty level, based on patients having complete information for each of these co-variables.

*Statistical analysis*

All data were analyzed using SAS version 9.4 statistical software (SAS Institute; Cary, NC). All statistical tests were two-sided. A p-value of <0.05 was considered to be statistically significant.

# CHAPTER 3. CLINICOPATHOLOGY AND MOLECULAR PRECURSORS FOR TIME TO BREAST CANCER AMONG WOMEN WITH BENIGN BREAST LESIONS

## INTRODUCTION

Each year in the United States, over 1.5 million breast biopsies are performed.[94,95] The majority of these lesions are classified as benign;[96] however, benign breast disease (BBD) is an established breast cancer risk factor, among which distinct pathological features are associated with higher risk of malignancy.[2,97-100] Benign breast disease pathological features are loosely categorized into three groups using Dupont and Page criteria, where proliferative disease with atypia, or atypical hyperplasia, is associated with the greatest risk of subsequent breast cancer diagnosis and nonproliferative disease without atypia is associated with lowest risk.[101-103]

Other pathological findings, including columnar alterations (Figure 12D), can also contribute to an increased risk of subsequent breast cancer development (Figure 13).[104-106] Columnar cell lesions of the breast represent the morphological spectrum of alterations that includes: columnar cell change and columnar cell hyperplasia.[107] Increasing studies have shown that columnar alterations may be a marker of breast cancer risk in women with benign breast disease.[104] Yet the clinical relevance of these alterations is still undetermined.[105] Additional benign breast disease pathological features such as cysts, ductal ectasia, ductal hyperplasia, intra-ductal papilloma, lobular hyperplasia, and schlerosing adenosis also trended towards an increased relative risk of subsequent breast cancer development in black women (Figures 12 & 13).

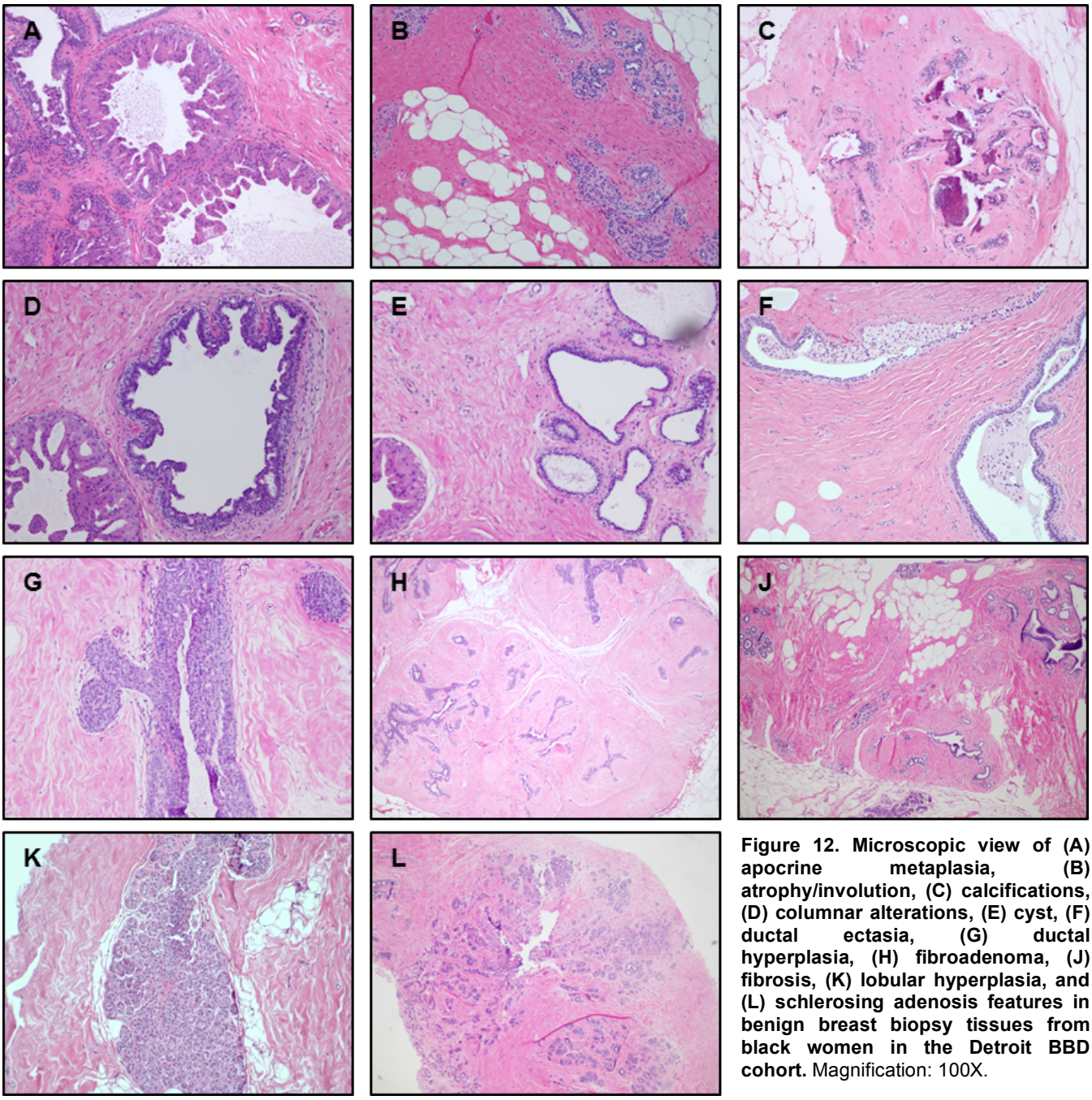


**Figure 12. Microscopic view of (A) apocrine metaplasia, (B) atrophy/involution, (C) calcifications, (D) columnar alterations, (E) cyst, (F) ductal ectasia, (G) ductal hyperplasia, (H) fibroadenoma, (J) fibrosis, (K) lobular hyperplasia, and (L) schlerosing adenosis features in benign breast biopsy tissues from black women in the Detroit BBD cohort.** Magnification: 100X.

*Photographs courtesy of E. Abdulfatah, Wayne State University, Department of Pathology.*

In contrast, lobular involution or atrophy has also been associated with breast cancer risk, with higher proportions of involution being protective (Figures 12B & 13). Previous studies have shown that higher proportions of atrophy are indeed associated with a decreased risk of breast cancer development, although given the small sample size of patients this association was not found to be significant (Figure 13). Thus, there is a critical need to understand risk factors unique to the black population to identify high-risk women who can benefit from closer surveillance, earlier diagnosis and better patient prognosis.[2]

Although it has been shown that benign breast lesions confer an increased breast cancer risk among both blacks and whites,[2,99,100,103] little is known about the molecular profiles of benign breast lesions and subsequent breast tumorigenesis, particularly in the black population. Gene expression profiling on breast biopsy tissue from black and white women was recently performed by Field et al. and demonstrated differences in molecular profiles between races, but these nonmalignant tissues demonstrated no evidence of benign conditions.[108] No studies to date have examined the transcriptome of benign breast lesions from black women. Thus, this study sought to identify novel molecular precursors that can predict shorter time to breast cancer diagnosis among benign breast lesions of black women. Profiling the transcriptome of these premalignant lesions served to help identify black women diagnosed with benign breast disease at highest risk who could largely benefit from increased surveillance, chemopreventive strategies, and earlier detection, leading to improved survival.

| | Breast Cancer Status | | Relative Risk[a] | p value |
|---|---|---|---|---|
| | Negative | Positive | (95% C.I.) | |
| **Apocrine Metaplasia** | | | | |
| No | 927 | 38 | 1.0 (Ref) | |
| Yes | 422 | 17 | 0.93 (0.52-1.67) | 0.80 |
| **Atrophy/Involution** | | | | |
| None | 237 | 11 | 1.0 (Ref) | |
| Partial | 700 | 27 | 0.75 (0.37-1.55) | 0.44 |
| Complete | 218 | 12 | 0.84 (0.34-2.07) | 0.71 |
| **Columnar Alterations** | | | | |
| No | 1127 | 40 | 1.0 (Ref) | |
| Yes | 222 | 15 | 1.84 (0.99-3.39) | ***0.05*** |
| **Cyst** | | | | |
| No | 809 | 28 | 1.0 (Ref) | |
| Yes | 540 | 27 | 1.41 (0.82-2.42) | 0.21 |
| **Ductal Ectasia** | | | | |
| No | 1294 | 52 | 1.0 (Ref) | |
| Yes | 57 | 3 | 1.33 (0.40-4.41) | 0.64 |
| **Ductal Hyperplasia** | | | | |
| None | 928 | 36 | 1.0 (Ref) | |
| Yes | 415 | 19 | 1.14 (0.64-2.01) | 0.66 |
| **Fibrosis** | | | | |
| No | 505 | 21 | 1.0 (Ref) | |
| Yes | 749 | 31 | 0.97 (0.55-1.71) | 0.91 |
| **Intra-Ductal Papilloma** | | | | |
| None | 1220 | 48 | 1.0 (Ref) | |
| 1 or more | 130 | 7 | 1.26 (0.55-2.84) | 0.59 |
| **Lobular Hyperplasia** | | | | |
| No | 1330 | 54 | 1.0 (Ref) | |
| Yes | 21 | 1 | 1.08 (0.14-8.21) | 0.94 |
| **Sclerosing Adenosis** | | | | |
| No | 1266 | 48 | 1.0 (Ref) | |
| Yes | 84 | 7 | 2.20 (0.96-5.01) | 0.06 |
| **Overall Impression** | | | | |
| Non-Proliferative Disease | 912 | 33 | 1.0 (Ref) | |
| Proliferative Disease without Atypia | 388 | 17 | 1.16 (0.64-2.12) | 0.62 |
| Proliferative Disease with Atypia | 39 | 5 | 3.29 (1.21-8.93) | ***0.02*** |

[a] *Adjusted for age and year at biopsy.*
*CI: confidence interval*

**Figure 13. Pathological characteristics of benign breast tissue from biopsies and association with risk of subsequent breast cancer among black women in metropolitan Detroit.** Adapted from Cote, et al.[2] *Copyright permissions for re-use of this figure were obtained from the American Association for Cancer Research (AACR).*

# RESULTS:

## CLINICOPATHOLOGICAL AND MOLECULAR CHARACTERISTICS OF BENIGN BREAST LESIONS FROM BLACK WOMEN SUBSEQUENTLY DIAGNOSED WITH BREAST CANCER

*Clinicopathological characteristics of study cohort*

A total of 3,759 black women from metropolitan Detroit with benign breast lesions comprise the Detroit benign breast disease cohort. Among these individuals, thirty-six women who were subsequently diagnosed with breast cancer were selected based on tissue availability (16.6%, 36 of 217 cases) to comprise this study sample (Figure 14). Mean age at first breast biopsy among these women was 55.3 years, ranging from 29 to 76 years. Similar to the entire benign breast disease cohort,[2] half of the benign samples were from excisional biopsies (n=18, 50%) and half were from core needle biopsies performed between 1997 and 2003.

Evaluation of other breast cancer risk factors, such as patient body mass index (BMI), revealed that the 72.2% of women (26 of 36 cases) were classified as overweight (BMI: 25- ≤30 kg/m$^2$) or obese (BMI: >30 kg/m$^2$) (Figure 14). BMI ranged from 20.5 to 52.9 kg/m$^2$, with a mean BMI in this cohort of 30.7 (sd 8.8). A self-reported history of breast cancer in first-degree relatives was observed in 36.1% of cases, whereas over 61.1% (22 of 36 cases) reported a familial history of cancer. Further, 13.9% of cases (5 of 36) underwent menarche at an early age, between 9 to 11 years, and 13.9% of women were nulliparous in this cohort. In total, 63.9% of women were clinically assessed as post-menopausal (Figure 14). A known history of cigarette smoking was reported for 47.2% of cases in this cohort (17 of 36 cases).

| *Characteristic* | *N* | *%* |
|---|---|---|
| | **36** | |
| **Age at Breast Biopsy, years** | | |
| <50 | 11 | 30.6% |
| 50-59 | 12 | 33.3% |
| 60-64 | 8 | 22.2% |
| 70+ | 5 | 13.9% |
| Mean (sd, years) | | 55.2 (11.4) |
| **Year of Breast Biopsy** | | |
| 1997-1999 | 14 | 38.9% |
| 2000-2001 | 12 | 33.3% |
| 2001-2003 | 10 | 27.8% |
| **Type of Breast Biopsy** | | |
| Excisional | 18 | 50.0% |
| Core/Needle | 18 | 50.0% |
| **Body Mass Index (kg/m$^2$)** | | |
| Normal (18.5 - ≤25) | 10 | 27.8% |
| Overweight (25- ≤30) | 12 | 33.3% |
| Obese, Class I (30- ≤35) | 5 | 13.9% |
| Obese, Class II (35+) | 9 | 25.0% |
| Mean (sd, kg/m$^2$) | | 30.7 (8.8) |
| **Breast Cancer History in First-Degree Relatives** | | |
| No | 18 | 50.0% |
| Yes | 13 | 36.1% |
| Unknown | 5 | 13.9% |
| **Age at Menarche, years** | | |
| 9-11 | 5 | 13.9% |
| 12-13 | 13 | 36.1% |
| 14+ | 9 | 25.0% |
| Unknown | 9 | 25.0% |
| **Menopausal Status** | | |
| Pre-menopausal | 2 | 5.6% |
| Peri-menopausal | 4 | 11.1% |
| Post-menopausal | 23 | 63.9% |
| Unknown | 7 | 19.4% |
| **Number of Pregnancies** | | |
| 0 | 5 | 13.9% |
| 1-2 | 8 | 22.2% |
| 3-4 | 10 | 27.8% |
| 5+ | 6 | 16.7% |
| Unknown | 7 | 19.4% |
| **Cigarette Smoking History** | | |
| No | 17 | 47.2% |
| Yes | 17 | 47.2% |
| Unknown | 2 | 5.6% |

**Figure 14. Clinical and demographic characteristics of thirty-six black women with benign breast lesions and a subsequent breast cancer diagnosis from the metropolitan Detroit cohort.**

| *Feature* | *N* | *%* |
|---|---|---|
| | **36** | |
| **Apocrine Metaplasia˚** | | |
| No | 22 | 62.90% |
| Yes | 13 | 37.10% |
| **Atrophy/Involution** | | |
| No | 11 | 30.60% |
| 1-24% TDLU | 14 | 38.90% |
| 25-74% TDLU | 9 | 25.00% |
| Unknown | 2 | 5.60% |
| **Calcifications˚** | | |
| No | 16 | 45.70% |
| Yes | 19 | 54.30% |
| **Columnar Alterations˚** | | |
| No | 20 | 57.10% |
| Yes | 14 | 40.00% |
| Atypia | 1 | 2.90% |
| **Cyst˚** | | |
| No | 13 | 37.10% |
| Yes | 22 | 62.90% |
| **Ductal Ectasia˚** | | |
| No | 31 | 88.60% |
| Atypical | 4 | 11.40% |
| **Ductal Hyperplasia˚** | | |
| No | 20 | 57.10% |
| Atypical | 3 | 8.60% |
| Mild | 6 | 17.10% |
| Moderate/Florid | 6 | 17.10% |
| **Fibroadenoma** | | |
| No | 25 | 69.40% |
| Yes | 11 | 30.60% |

| *Feature* | *N* | *%* |
|---|---|---|
| | **36** | |
| **Fibrosis˚** | | |
| No | 16 | 45.70% |
| Yes | 18 | 51.40% |
| Marked | 1 | 2.90% |
| **Intraductal Papilloma˚** | | |
| No | 30 | 85.70% |
| Single | 4 | 11.40% |
| Multiple | 1 | 2.90% |
| **Mucocele-Like Lesions˚** | | |
| No | 35 | 100.00% |
| Yes | 0 | 0.00% |
| **Lobular Hyperplasia˚** | | |
| No | 35 | 100.00% |
| Yes | 0 | 0.00% |
| **Radial Scar** | | |
| No | 32 | 88.90% |
| Yes (<5 mm) | 4 | 11.10% |
| **Sclerosing Adenosis˚** | | |
| No | 29 | 82.90% |
| Yes | 6 | 17.10% |
| **Overall Impression** | | |
| Non-proliferative disease without atypia | 17 | 47.20% |
| Proliferative disease without atypia | 15 | 41.70% |
| Atypical hyperplasia | 4 | 11.10% |
| **Breast Cancer Behavior** | | |
| Ductal carcinoma in situ | 10 | 27.80% |
| Invasive ductal carcinoma | 26 | 72.20% |

˚Pathology review for 1 case noted that this benign breast feature was not applicable in the lesion.

**Figure 15. Pathological features and overall impression of benign breast lesions from thirty-six black women who were subsequently diagnosed with breast cancer from the metropolitan Detroit cohort.** TDLU, terminal duct lobular unit.

Figure 15 presents the pathological features of benign breast biopsies. The majority of the sample presented with cysts (62.9%), calcifications (54.3%), and fibrosis (51.4%). Proliferative disease with atypia, atypical hyperplasia, was present in 11.1% of cases. Of these patients, 27.8% of patients (10 of 36 women) went on to develop ductal carcinoma in situ (DCIS), while 62.2% (26 of 36 women) were diagnosed with invasive ductal carcinoma (Figures 2 & 15). At fifteen years, the cumulative incidence for breast cancer was 2.24% for DCIS and 5.34 % for invasive ductal carcinoma, with a median time of 6.4 years to breast cancer diagnosis (Figure 16). Mean age at subsequent breast cancer diagnosis was 62.1 years (sd 11.4 years; data not shown).

Figure 17 highlights clinical and demographic characteristics of subsequent breast cancer cases diagnosed between 1999 and 2012 among the thirty-six women in the study sample. Mean age at diagnosis among the 10 cases of DCIS was 58.3 years (sd 10.4 years), while mean age at diagnosis for invasive breast cancer cases was older, with 63.5 years (sd 11.7 years). ER status was positive among all DCIS cases with known receptor information, while 69.2% of all invasive breast cancer cases were classified as ER-positive (Figure 17). Further assessment of progesterone receptor status revealed that 80.0% of all DCIS cases were positive, in comparison to 73.1% of invasive cancer cases. Evaluation of HER2 status revealed that no cases of DCIS were positive among those that could be evaluated, and 11.5% of all invasive cases were HER2-positive. Notably, 23.1% of invasive breast cancers (6 of 26 cases) were found to be of the triple negative subtype (ER-/PR-/HER2-negative) (data not shown).

Examination of primary tumor size revealed that 15.4% of all invasive breast cancers were larger than 5 centimeters in size, indicating these cases would be

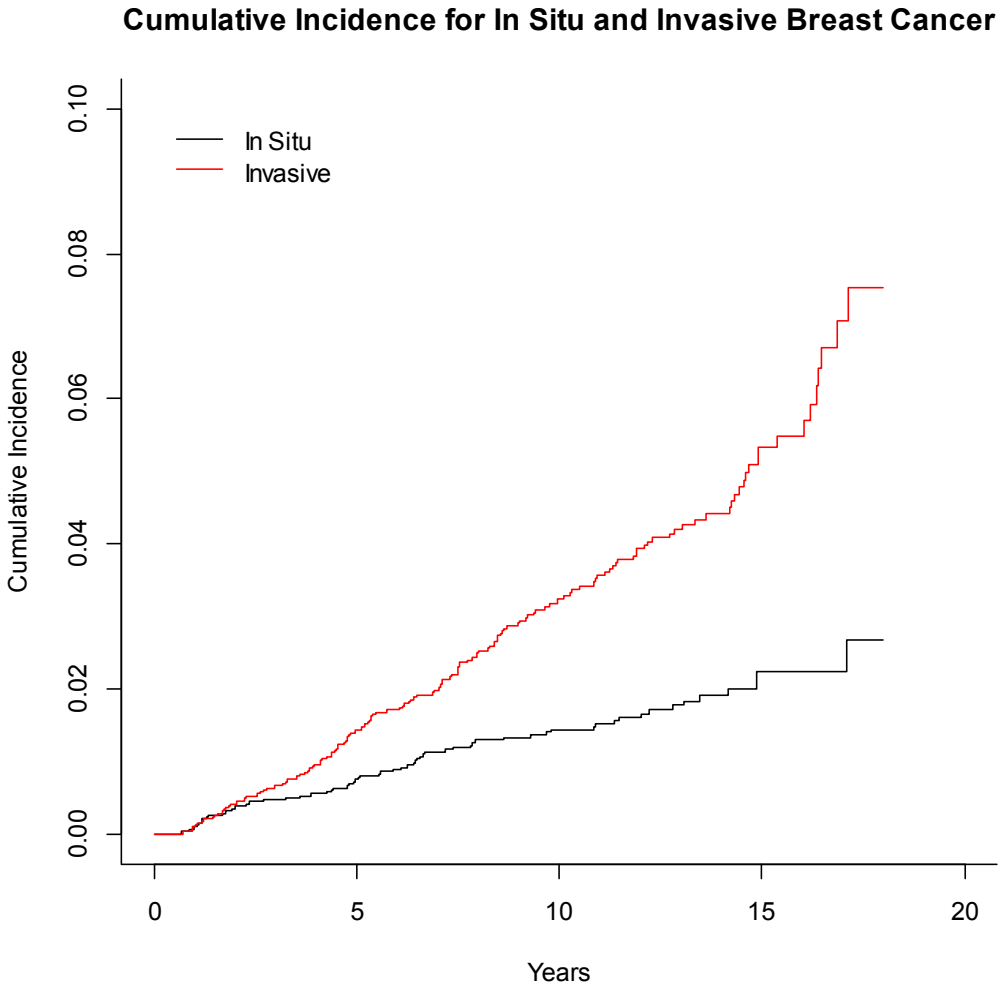


**Figure 16. Cumulative incidence for in situ and invasive breast cancer among 217 black women with benign breast lesions in the Detroit BBD cohort.** (63 in-situ cancers and 154 invasive cancers). *Figure generated using R software.*

| Characteristic | *In situ* N | *In situ* % | *Invasive* N | *Invasive* % |
|---|---|---|---|---|
| | **10** | | **26** | |
| **Age at Diagnosis, years** | | | | |
| <50 | 1 | 10.0% | 2 | 7.7% |
| 50-59 | 6 | 60.0% | 8 | 30.8% |
| 60-69 | 1 | 10.0% | 7 | 26.9% |
| 70+ | 2 | 20.0% | 9 | 34.6% |
| Mean (sd, years) | 58.3 (10.4) | | 63.5 (11.7) | |
| **Year of Diagnosis** | | | | |
| 1999-2003 | 3 | 30.0% | 5 | 19.2% |
| 2004-2008 | 6 | 60.0% | 11 | 42.3% |
| 2009-2012 | 1 | 10.0% | 10 | 38.5% |
| **Estrogen Receptor (ER) Status** | | | | |
| Negative | 0 | 0.0% | 8 | 30.8% |
| Positive | 9 | 90.0% | 18 | 69.2% |
| Unknown | 1 | 10.0% | 0 | 0.0% |
| **Progesterone Receptor (PR) Status** | | | | |
| Negative | 1 | 10.0% | 7 | 26.9% |
| Positive | 8 | 80.0% | 19 | 73.1% |
| Unknown | 1 | 10.0% | 0 | 0.0% |
| **HER2 Status** | | | | |
| Negative | 9 | 90.0% | 23 | 88.5% |
| Positive | 0 | 0.0% | 3 | 11.5% |
| Unknown | 1 | 10.0% | 0 | 0.0% |

**Figure 17. Demographic and pathological characteristics of breast cancers from thirty-six black women with benign breast disease in metropolitan Detroit.**

diagnosed with AJCC stage IIB or higher clinical stage (Figure 18). Over half of all invasive cases demonstrated regional lymph node involvement (53.8%), as compared to 30.0% of DCIS cases. While 20.0% of women diagnosed with DCIS received chemotherapy, 61.5% of the individuals diagnosed with invasive breast cancer received chemotherapy in a neoadjuvant or adjuvant setting (16 of 26 cases) (Figure 18). For both DCIS and invasive cancers, the majority of these black women underwent breast-conserving surgery (60.0% and 57.7% of DCIS and invasive cases, respectively). 76.9% of invasive cases underwent radiation in first course of treatment, while in contrast the majority of women diagnosed with DCIS did not receive radiation therapy. At last follow-up, 77.8% of all black women in the study cohort were alive (Figure 18).

*Molecular precursors for time to cancer among black women with benign breast lesions*

To investigate molecular precursors of benign breast lesions associated with shorter time to breast cancer diagnosis, LHRs for 67,528 transcripts were calculated, adjusted for age at breast biopsy and overall impression of the benign breast lesion. 18,749 transcripts (absolute LHR>1.0 and p-value<0.05) were entered into Ingenuity Pathway Analysis (IPA) software (Figure 19). The top up-regulated molecule associated with time to breast cancer diagnosis was histone lysine demethylase 4C (KDM4C) (LHR, 12.949; p-value=0.00177), and the top down-regulated molecule was deformed epidermal autoregulatory factor 1 (DEAF1) (LHR, -9.865; p-value=0.00227) (Figure 20).

| Characteristic | *In situ* N | *In situ* % | *Invasive* N | *Invasive* % |
|---|---|---|---|---|
| | **10** | | **26** | |
| **Tumor Size** | | | | |
| Microscopic foci | 2 | 20.0% | 0 | 0.0% |
| ≤ 5 cm | 5 | 50.0% | 22 | 84.6% |
| > 5 cm | 1 | 10.0% | 4 | 15.4% |
| Unknown | 2 | 20.0% | 0 | 0.0% |
| **Regional Lymph Nodes** | | | | |
| Not examined | 7 | 70.0% | 2 | 7.7% |
| Negative | 3 | 30.0% | 14 | 53.8% |
| Positive | 0 | 0.0% | 10 | 38.5% |
| **Chemotherapy** | | | | |
| No | 8 | 80.0% | 7 | 26.9% |
| Yes | 2 | 20.0% | 16 | 61.5% |
| Unknown | 0 | 0.0% | 3 | 11.5% |
| **Surgery** | | | | |
| None | 0 | 0.0% | 2 | 7.7% |
| Breast-conserving surgery | 6 | 60.0% | 15 | 57.7% |
| Mastectomy | 4 | 40.0% | 9 | 34.6% |
| **Radiation Therapy** | | | | |
| No | 8 | 80.0% | 6 | 23.1% |
| Yes | 2 | 20.0% | 20 | 76.9% |
| **Vital Status** | | | | |
| Alive | 9 | 90.0% | 19 | 73.1% |
| Deceased | 1 | 10.0% | 7 | 26.9% |

**Figure 18. Clinical characteristics of breast cancer from thirty-six black women with benign breast disease in metropolitan Detroit.**

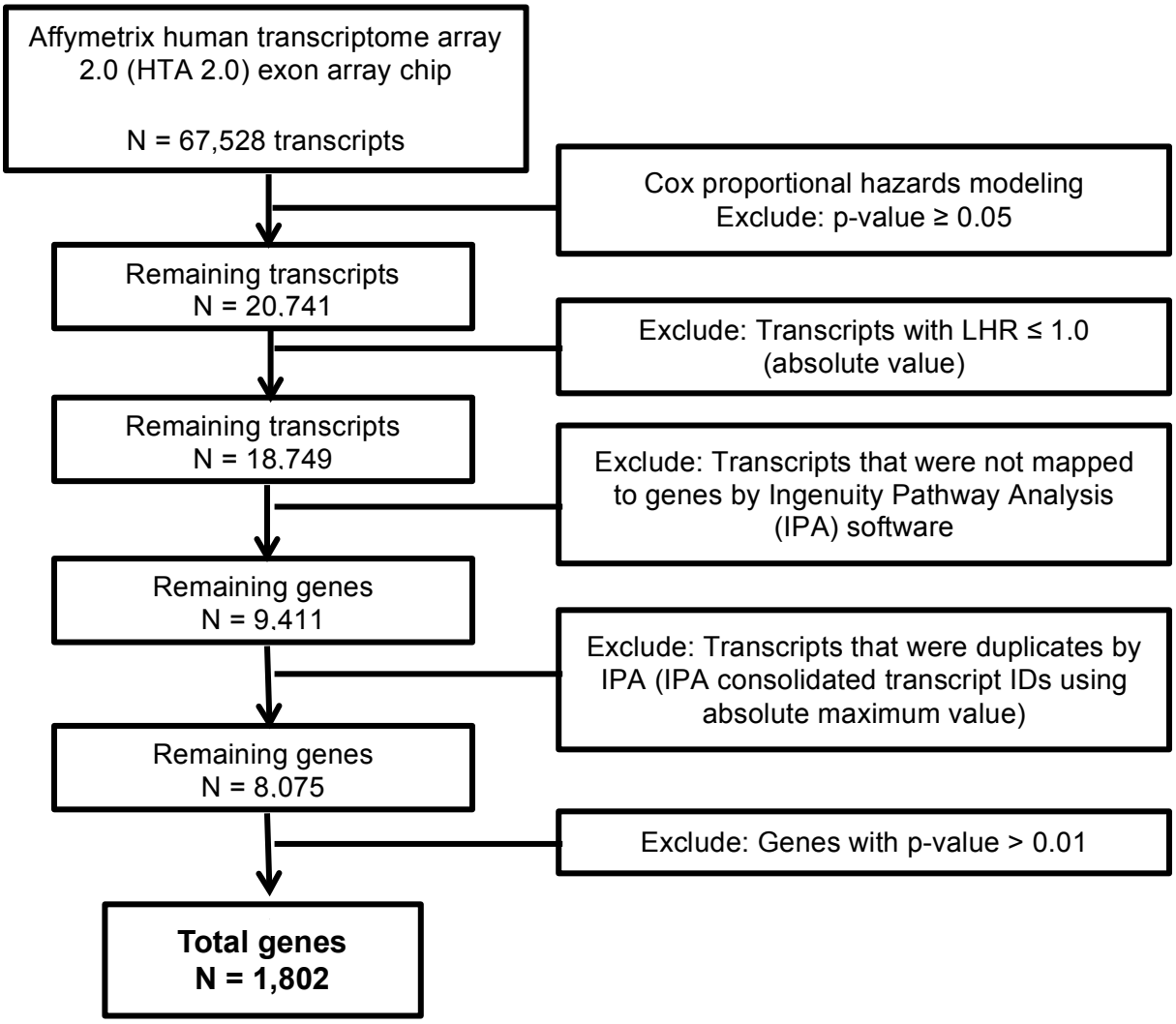


**Figure 19. Analysis flowchart for the study.**

| *Log Hazard Ratio* | *p-value* | *Transcript Cluster ID* | *Symbol* | *Entrez Gene Name* |
|---|---|---|---|---|
| **12.949** | 0.00177 | TC09002929.hg.1 | KDM4C | lysine demethylase 4C |
| **11.682** | 0.00008 | TC11002221.hg.1 | TRPC6 | transient receptor potential cation channel subfamily C member 6 |
| **10.3** | 0.00393 | TC21000441.hg.1 | DSCR3 | DSCR3 arrestin fold containing |
| **9.523** | 0.00003 | TC0X000084.hg.1 | NHS | NHS actin remodeling regulator |
| **8.923** | 0.00178 | TC07001050.hg.1 | NOM1 | nucleolar protein with MIF4G domain 1 |
| **-8.072** | 0.00068 | TC02000440.hg.1 | CCT7 | chaperonin containing TCP1 subunit 7 |
| **-8.227** | 0.00406 | TC13000217.hg.1 | WDFY2 | WD repeat and FYVE domain containing 2 |
| **-8.39** | 0.00214 | TC05001005.hg.1 | TSPAN17 | tetraspanin 17 |
| **-9.308** | 0.00235 | TC19001497.hg.1 | YIF1B | Yip1 interacting factor homolog B, membrane trafficking protein |
| **-9.865** | 0.00227 | TC11001244.hg.1 | DEAF1 | DEAF1, transcription factor |

**Figure 20. Gene expression profiling data of top up- and down-regulated molecules associated with shorter time to breast cancer among benign breast biopsy lesions from thirty-six black women.**

# RESULTS:

## HISTONE LYSINE DEMETHYLASE 4C, KDM4C: THE TOP UP-REGULATED MOLECULE ASSOCIATED WITH TIME TO BREAST CANCER DIAGNOSIS

Of the 67,528 transcripts that were evaluated through gene expression profiling, analyses revealed that KDM4C was the top up-regulated molecule to be associated with a shorter time to breast cancer diagnosis among black women with benign breast disease (Figure 20). Strikingly, my early doctoral work under the mentorship of Zeng-Quan Yang, was primarily focused on the dysregulation of histone lysine demethylases in cancers, and in particular the expression patterns and molecular mechanism of KDM4C in breast cancer.

Histone demethylases play essential roles in dynamically regulating gene expression and chromatin architecture through histone lysine methylation, and are thus implicated in developmental processes and tumorigenesis.[109-112] Many lines of evidence suggest that genetic alteration and dysregulation of histone lysine demethylases are associated with breast cancer initiation and progression, where the effect is to activate expression of oncogenes, repress expression of tumor suppressors, alter DNA mismatch repair, disrupt chromosomal stability, or interact with key hormonal receptors which control cellular proliferation.[3,113-115]

Histones can be post-translationally modified through covalent addition of a number of distinct chemical moieties through dynamic processes, including methylation.[116-120] Methylation at various lysine residues, degree of methylation, and location of the methylated histone within a specific gene locus all yield different transcriptional and biological outcomes. Histone lysine methylation is the principal

chromatin-regulatory mechanism that influences fundamental nuclear processes and gene expression. Histone methylation is known to occur on the lysine residues of histones 3 and 4 (H3, H4), and the linker histone H1, isotype 4 (H1.4), and modifications of these residues are associated with distinct functional outcomes. Methylation of H3K4, H3K36 and H3K79 is generally associated with gene activation, while methylation of H3K9, H3K27, H3K56, H4K20 and H1.4K26 is linked to transcriptional repression.[121,122]

Structurally, the histone lysine demethylases are a diverse group of proteins that can be subdivided into two distinct groups based on their mechanism of demethylation. Amine oxidase demethylases, the first group, can oxidize their substrate by FAD to generate an imine that gets hydrolyzed (Figure 21A).[5,123] Histone lysine demethylases regulate chromatin architecture and transcription, and play critical roles in epigenetic signaling. The first functional enzymatic family of histone lysine demethylases in humans includes the lysine specific demethylase (LSD1, also known as KDM1A), which, along with the structurally similar KDM1B (LSD2), consist of the flavin adenine dinucleotide (FAD)-dependent amine oxidases, which can remove mono- and dimethyl histone lysine marks.[124-126] These amine oxidases, however, are unable to demethylate tri-methyl lysine residues since they require a lone pair of electrons only present on mono- and dimethyl lysine histone residues.

The other known 32 histone lysine demethylases belong to the JmjC (JumonjiC)-domain containing group.[3,127] These demethylases are characterized by a JmjC domain with defined enzymatic activity. Residues within predicted cofactor binding sites are conserved.[123] The mechanism of demethylation requires five residues within the JmjC domain that bind to both the α-ketoglutarate (α-KG) and iron cofactors to undergo a

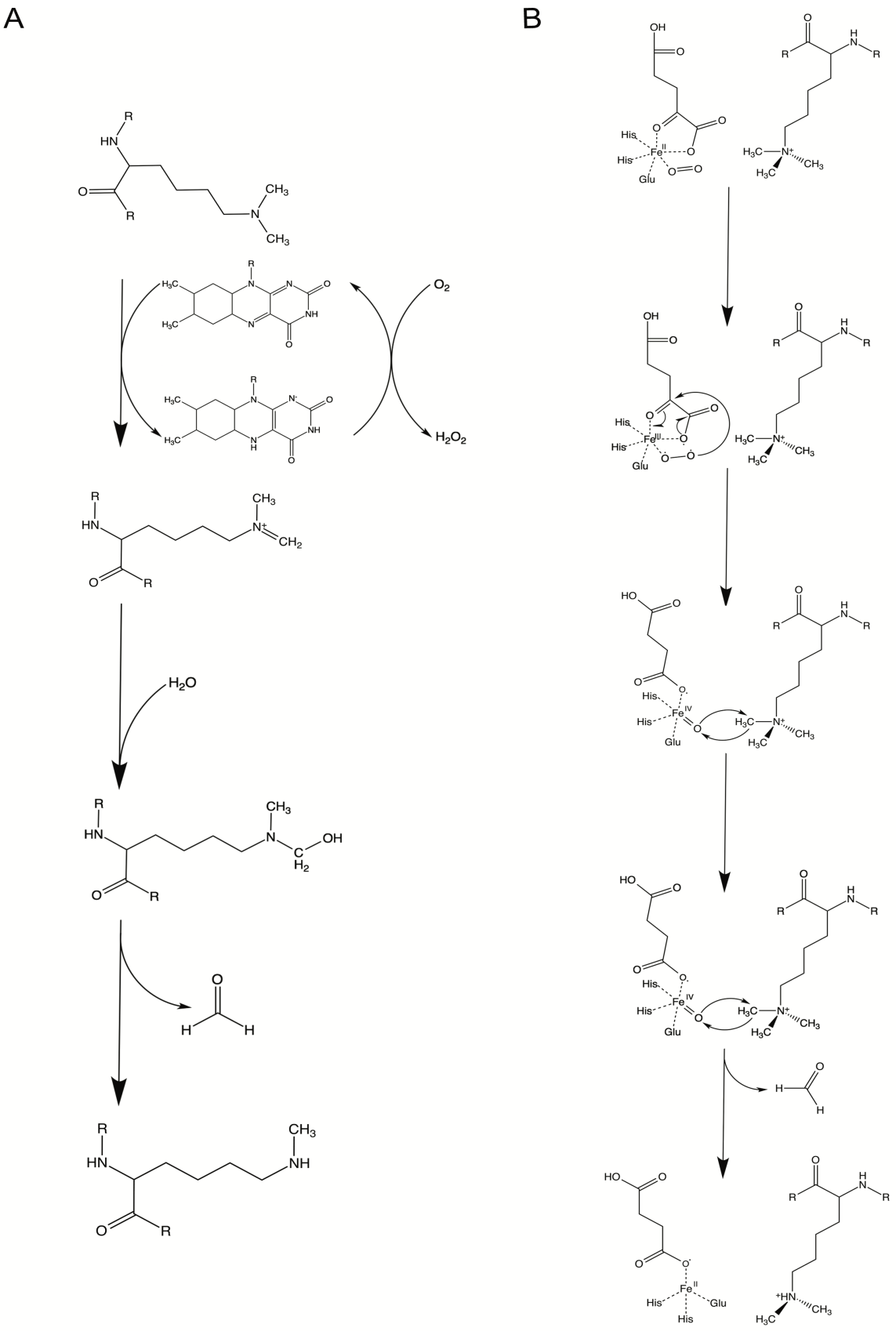


**Figure 21. Mechanisms of histone lysine demethylation.**[5] (A) Mechanism of LSD1 demethylation via an amine oxidation reaction. (B) Mechanism of JmjC protein demethylation via a hydroxylation reaction. *Figure designed using ChemDraw software. Open access copyright permissions for this article permitted*

hydroxylation reaction, which allow for removal of a methyl mark (Figure 21B).[5] Often, the JmjC domain can be found in combination with other protein domains, which function in substrate specificity. Given the correlation between particular methyl marks and the transcriptional state of genes, it has been proposed that the activity of specific histone lysine demethylases contributes to different transcriptional and biological outcomes, depending on the histone lysine demethylase substrate.[128-130]

*Genomic and protein structures of the KDM4 demethylase subfamily*

KDM4C, also referred to as JMJD2C or GASC1 (Gene Amplified in Squamous Cell Carcinoma 1), is overexpressed in numerous cancers including esophageal squamous cell carcinoma, breast and prostate cancers, medulloblastoma, metastatic lung sarcomatoid carcinoma, in primary mediastinal B-cell lymphoma and Hodgkin's lymphoma, and in acute myeloid leukemia.[131-138] Recurring evidence supports that *KDM4C* overexpression results from aberrant amplification of chromosome 9 at the 9p23-24 foci.[131] It is also aberrantly expressed as a fusion partner to the immunoglobulin heavy chain gene (IGH) in mucosa-associated lymphoma, following 9p translocation.[139]

Previous studies have demonstrated that *KDM4* genes are amplified and overexpressed in various tumor types, including lung, breast, esophageal, prostate cancers and lymphoma.[131,134,138-140] KDM4C is part of the large KDM4 subfamily that consists of five functional KDM4 member genes (*KDM4A*-*E*) in humans (Figure 22). Genes encoding *KDM4A*, *B* and *C* localize to human chromosomes 1p34.1, 19p13.3, and 9p24.1, respectively. *KDM4D* localizes to human chromosome 11q21, and forms a

cluster with two additional intronless *KDM4* genes, *KDM4E* and *KDM4F*.[141] Previously, *KDM4E* and *F* were considered pseudogenes, however *KDM4E* expression has recently been observed, suggesting its role as a functional gene.[121,142,143]

The KDM4A, B and C proteins, which share more than 50% sequence identity, each contain JmjN, JmjC, two plant homeodomains (PHD) and two Tudor domains. KDM4D and KDM4E, in contrast, are considerably shorter proteins that lack the C-terminal region, including the PHD and Tudor domains (Figure 22). As with all JmjC-domain containing demethylases, the KDM4 JmjC domain bears catalytic function while the JmjN domain interacts extensively with JmjC and provides structural integrity.[133,144] Beyond the catalytic core of KDM4A-C, the C-terminal PHD and Tudor domains bear important histone reader functions, likely contributing to efficient nuclear localization. Recent studies found that the Tudor domains of KDM4C are responsible for binding H3K4me3 activation marks at gene promoters.[145] In addition, the PHD domains in other histone regulatory proteins have been demonstrated to bind unmodified, methylated, and/or acetylated histone residues on one or more histone tails, offering flexibility in directing epigenetic modifications.

Previous studies have indicated that KDM4A and C are broadly expressed in mouse and/or human tissues, while KDM4D and E are predominantly expressed in the mouse testes.[121,138,146,147] To establish a comprehensive profile of genomic alterations for *KDM4A-E* in human cancer, a large-scale meta-analysis was conducted of the genetic amplifications, deletions and mutations reported across 52 databases in the Cancer Genomics cBioPortal.[148,149] An overview of these data reveal that *KDM4A-E* are altered across many tumor types (Figure 23). These data are complemented by a

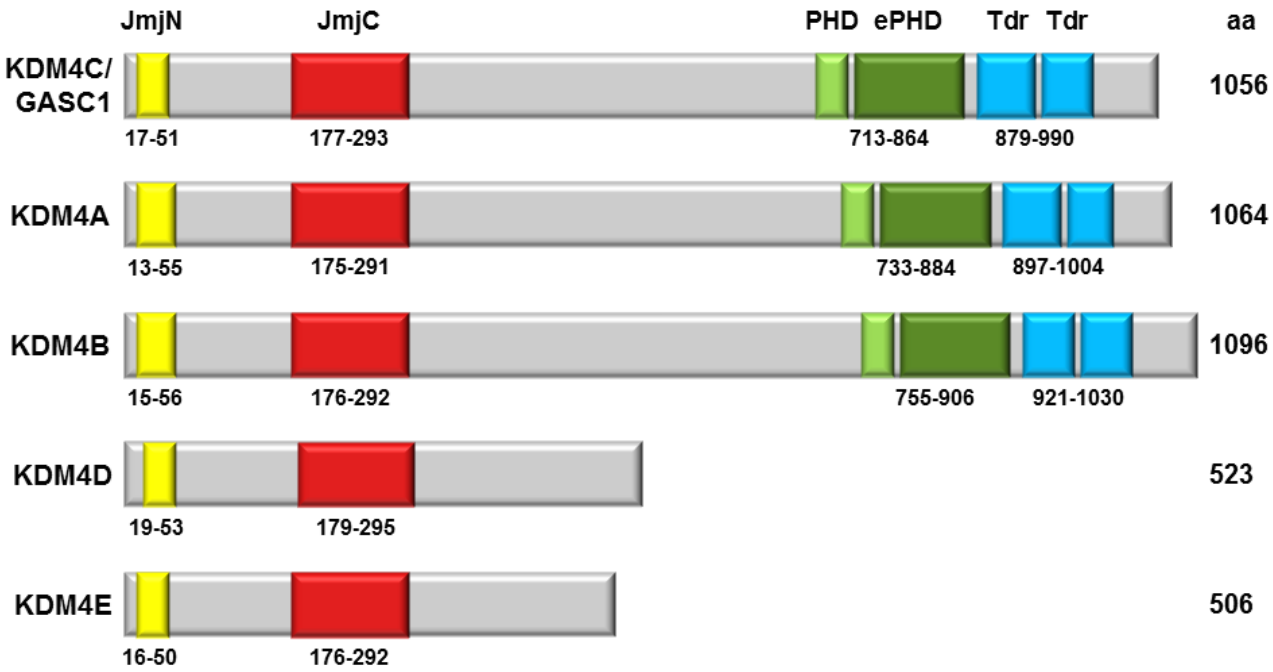


**Figure 22. Schematic structure and functional domains of the human KDM4 family**: the N-terminal Jumonji (Jmj) domains form the histone demethylase catalytic core, whereas the C-terminal plant homeo domain (PHD) and Tudor domains mediate specific lysine recognition.

recent analysis of 4,934 cancer copy number profiles from The Cancer Genome Atlas (TCGA) Pan-Cancer data set, which has revealed significant amplifications of the *KDM4C* genomic region in human cancer cells.[150]

*Genetic alterations of KDM4 histone demethylases across breast cancer subtypes*

To systematically investigate genetic alterations of the KDM4 demethylases in breast cancer, genome sequencing data of 976 primary breast cancer samples were first analyzed from The Cancer Genome Atlas (TCGA) database *via* cBioPortal.[148,149] In cBioPortal, copy numbers are computed using the GISTIC algorithm, which identifies the putative copy number level as high-level amplification, low-level gain, diploid, heterozygous deletion, or homozygous deletion.[148,149] For mRNA expression analysis, the Z-score is used to determine whether a gene is upregulated or downregulated relative to normal adjacent tissue or relative to all other tumor samples that are diploid for the gene. *KDM4C* was found to have the highest frequency (13.4%) of genetic alteration, including high-level amplification, homozygous deletion, mutation, upregulation, and downregulation, whereas *KDM4B* had the lowest frequency (6.1%) of genetic alteration among the 976 TCGA breast cancer specimens (data not shown). For *KDM4C*, 2.5% of samples had high-level amplification and 0.7% had homozygous deletion, whereas 1.9% of samples had high-level amplification or homozygous deletion of *KDM4A*, *B* and *D* (Figure 24).[1]

Correlations between gene expression and copy number have been used widely to prioritize driver oncogenes in human cancer, because mRNA expression can successfully translate the effect of elevated copy number to cancer initiation and

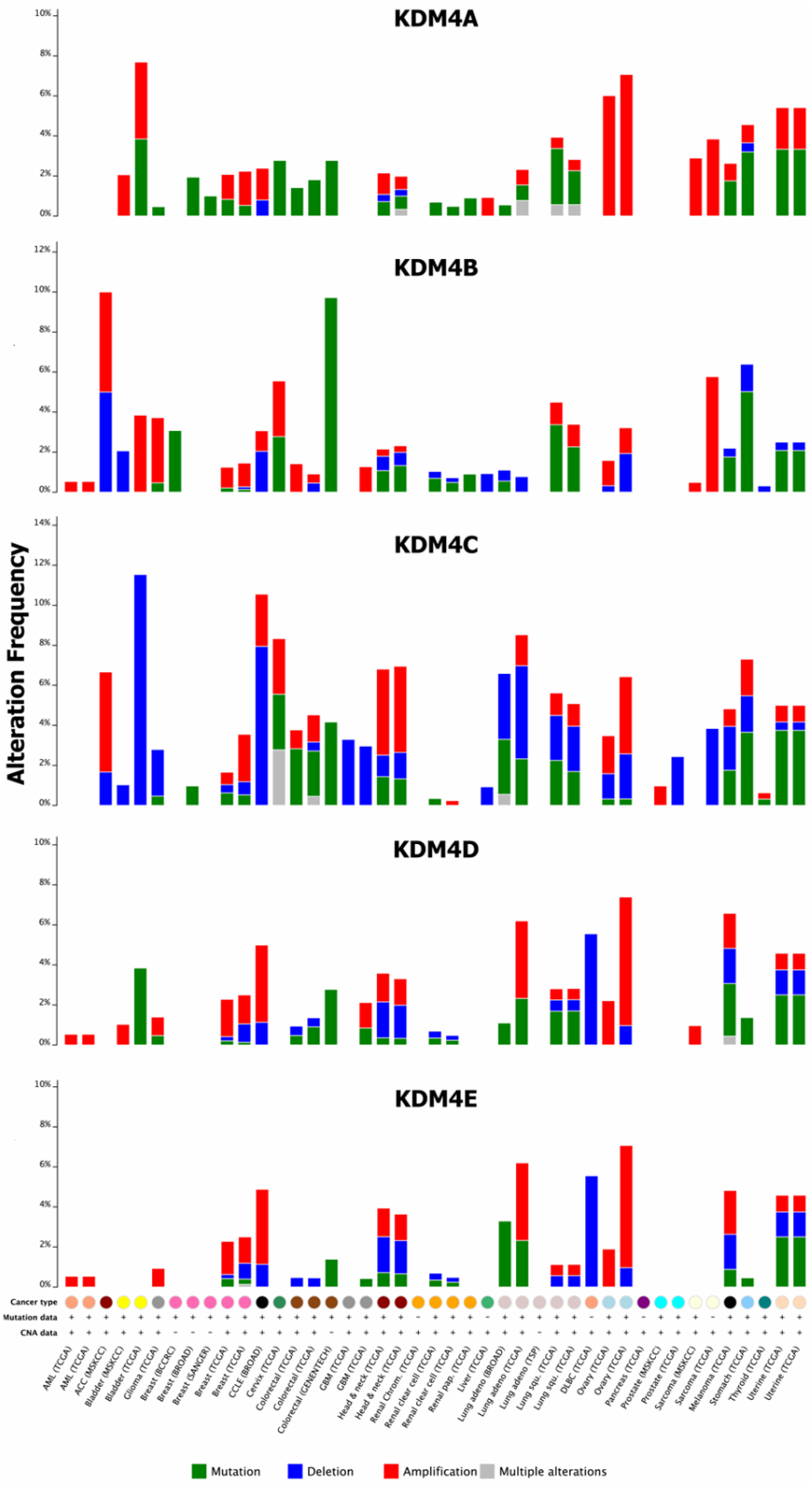


**Figure 23. Alteration frequencies of *KDM4* subfamily genes identified in human tumors of multiple origins reported across 52 databases held in the Cancer Genomics cBioPortal.**[3] Alteration frequencies are displayed for each of four categories, including: genetic amplifications (red), deletions (blue), mutations (green) or multiple alterations (grey). *Open access copyright permissions for this article permitted re-use of this figure from the American Journal of Translational Research.*

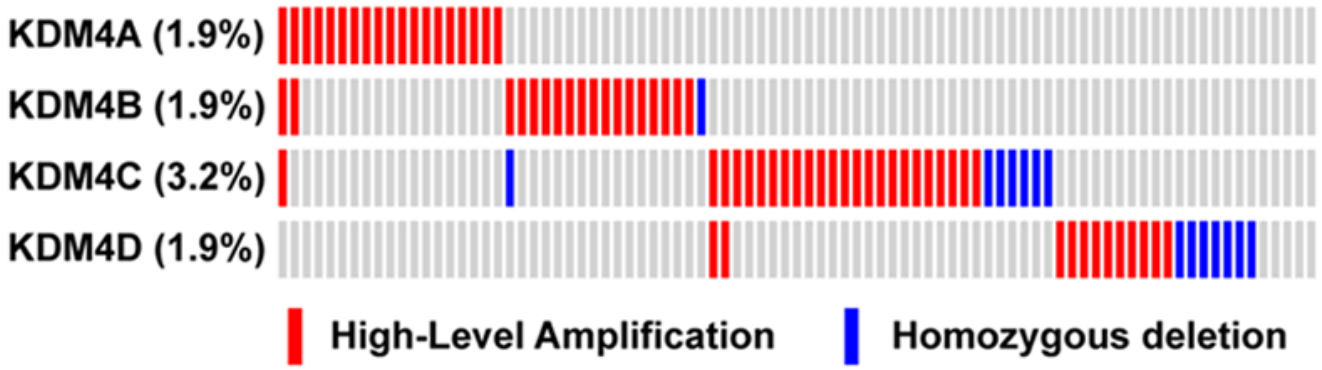


**Figure 24. KDM4A-D copy number alterations in breast cancer.**[1] High-level amplification and homozygous deletion of KDM4 subfamily members in TCGA breast cancer dataset (n=976). Data are displayed with the Oncoprint tool from cBioPortal. *Open access copyright permissions for this article permitted re-use of this figure from the American Journal of Cancer Research.*

progression. Thus, examination of the association between copy number and mRNA expression of KDM4A, B, C and D was performed by using Spearman's rank correlation in TCGA breast cancer specimens. Correlation between copy number and mRNA expression was strongest for KDM4C (R=0.61), followed by KDM4A (R=0.52); remaining KDM4 members had correlation coefficients less than 0.50 (R=0.44 for KDM4D, and R=0.28 for KDM4B) (Figure 25).

Next, it was sought to determine whether the high-level amplification or expression level of each KDM4 member was related to breast cancer subtype. Of the TCGA breast cancer samples, 493 had subtype data available, including 8 normal-like, 220 Luminal A, 121 Luminal B, 55 HER2+, and 89 basal-like breast cancers. Due to the small sample size (n=8) of the normal-like subtype, those samples were excluded from this analysis. *KDM4C* and *KDM4D* amplification was found to have the highest frequency (12.4% and 3.6%, respectively) in basal-like breast cancer. The highest frequency of *KDM4A* amplification (3.6%) was in HER2-subtype samples, whereas *KDM4B* amplification (3.3%) was highest in Luminal B breast cancer (Figure 26). To determine whether mRNA expression of each KDM4 subfamily member is associated with a specific subtype of breast cancer, the mRNA expression levels of KDM4A, B, C and D were compared across different subtypes of breast cancer specimens. Indeed, mRNA expression levels of KDM4A, C, and D were significantly higher in basal-like breast cancer (*P*=0.005, 6.01E-05, and 3.04E-07, respectively), whereas KDM4B mRNA was highly expressed in luminal breast cancer and less expressed in basal-like breast cancer (Figure 27).

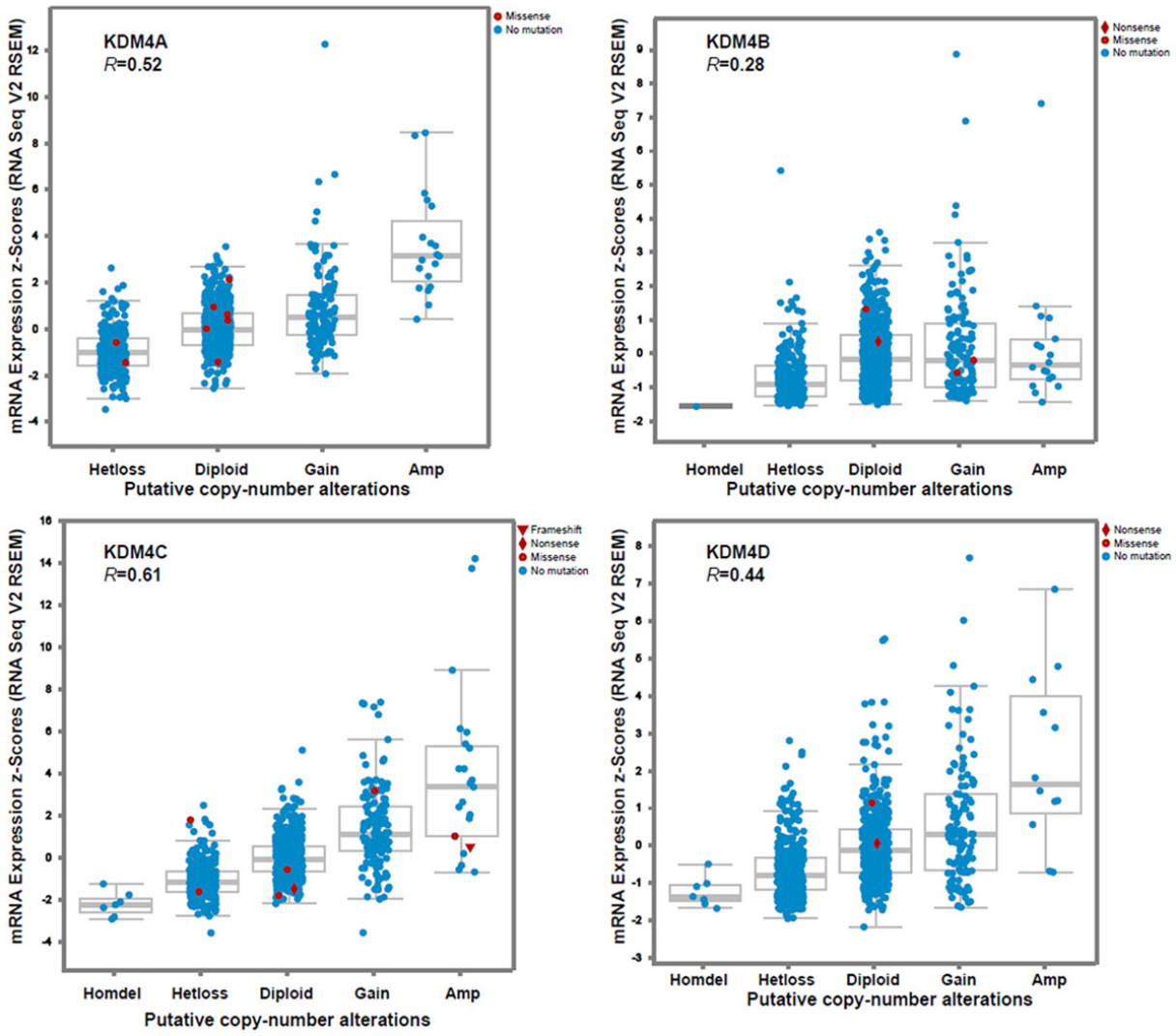


**Figure 25. KDM4A, B, C, and D mRNA expression (y-axis) versus copy number (x-axis) in 976 TCGA primary breast cancer samples.**[1] Amp: high-level amplification; Gain: low level gain; Hetloss: heterozygous deletion; and Homdel: homozygous deletion. *Open access copyright permissions for this article permitted re-use of this figure from the American Journal of Cancer Research.*

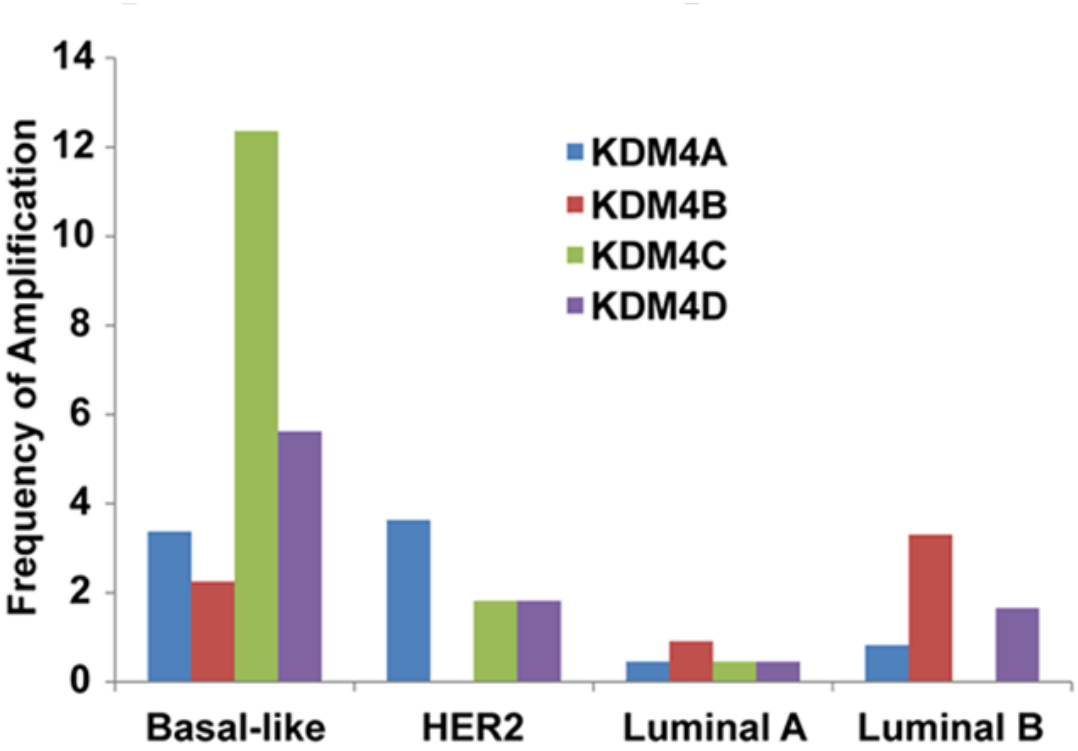


**Figure 26. KDM4A-D high-level amplification frequencies in different subtypes of breast cancer.**[1] Frequencies (%) of high-level amplification of KDM4A, B, C, and D in different subtypes of breast cancer. *Open access copyright permissions for this article permitted re-use of this figure from the American Journal of Cancer Research.*

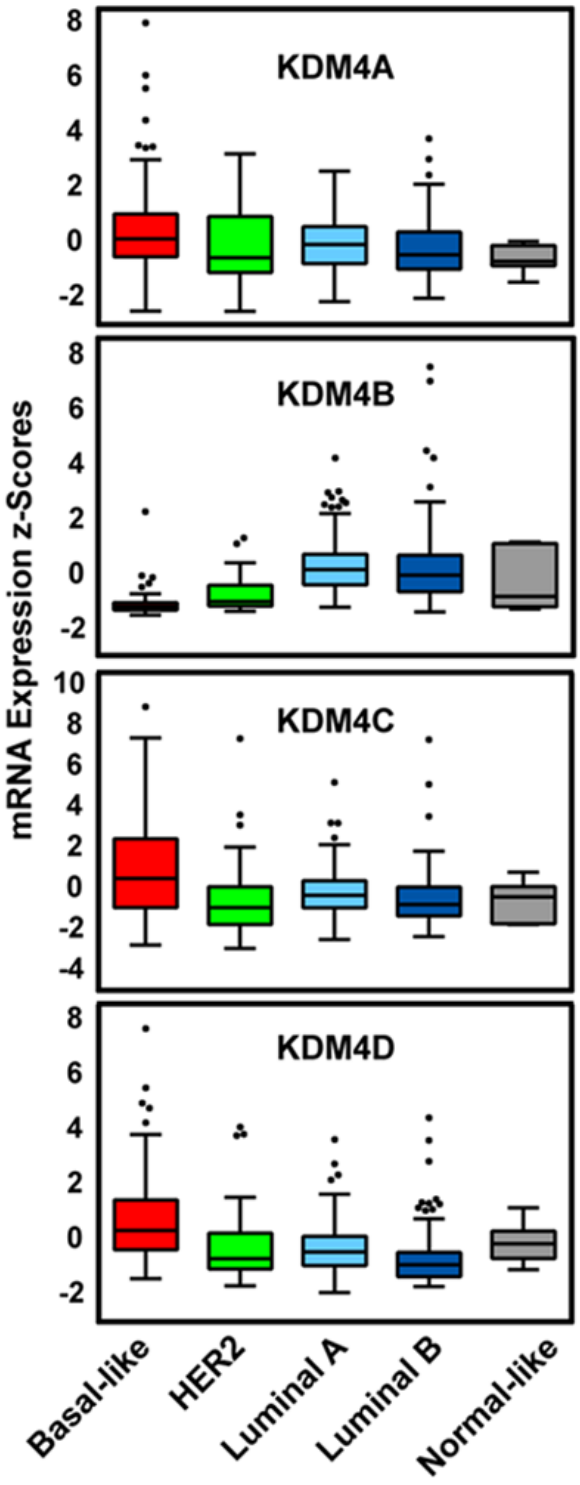


**Figure 27. KDM4A-D expression levels in different subtypes of breast cancer.**[1] Expression levels of KDM4A, B, C, and D across five subtypes of breast cancer based on TCGA database. The differences in KDM4A, B, C, and D mRNA levels among breast cancer subtypes are statistically significant ($P < 0.001$). *Open access copyright permissions for this article permitted re-use of this figure from the American Journal of Cancer Research.*

*KDM4C* was validated to be amplified at high levels in 12.4% of basal-like tumors. Incorporating low level-gain (39 of 89 cases, 43.8%) leads to more than half of basal-like tumors with an increased *KDM4C* copy number. Furthermore, KDM4A, C, and D had high mRNA expression levels in basal-like breast cancer. High-level amplification and expression of KDM4A, C, and D tended toward mutual exclusivity (Figure 24). Thus, KDM4A, C, and D might individually contribute to the aggressive phenotypes of basal-like breast cancer. Taken together, these data indicate that amplification and overexpression of KDM4A, C, and D are more prevalent in aggressive, basal-like breast cancers, while KDM4B overexpression is more prevalent in ER+, luminal breast cancers.

*Expression of KDM4 demethylases and histone methylation marks in breast cancer cell lines*

Next, the relative mRNA and protein expression levels of KDM4A, B, C and D in a panel of breast cancer cell lines were examined (Figures 28 & 29). MCF10A, an immortalized but non-tumorigenic breast epithelial cell line, was used as the control. Quantitative RT-PCR (qRT-PCR) assays were used to measure the mRNA expression level of KDM4A, B, C and D in 17 breast cancer cell lines, including nine basal-like, two HER2+, and six Luminal lines (Figure 28). To determine KDM4A, B, C and D protein abundance more precisely relative to their histone demethylase function, nuclear extracts from eight breast cancer cell lines and MCF10A were isolated and probed with antibodies that recognize KDM4A, B, C and D (Figure 29). KDM4A mRNA was found to be overexpressed in several ER+ luminal cell lines (such as HCC1428) and basal-like

cell lines (such as MDA-MB-468). KDM4B was strikingly overexpressed at the mRNA level in luminal cell lines, yet nuclear protein was also abundant in basal-like lines, such as Colo824 (Figure 29).

Consistent with the data from primary breast cancer specimens and previous findings, KDM4C is highly overexpressed in a set of basal-like breast cancer cell lines, including HCC1954 and Colo824, which both contain high-level KDM4C gene amplification (*Note:* Expression analysis indicate that the HCC1954 line belongs to the basal-like subtype even though it contains HER2 amplification).[131] Strikingly, immunoblot analysis did detected high levels of KDM4C protein expression in nuclear extracts from HCC1954 and Colo824 breast cancer cell lines (Figure 29). Similar to KDM4A, KDM4D is also overexpressed at the mRNA level in various ER+ luminal and basal-like cell lines. Interestingly, HCC1954 and Colo824 cells also showed KDM4D (11q21) gene amplification. However, KDM4D protein in these nuclear extracts was likely ubiquitously expressed in ER+ and basal breast cancers as well as MCF10A cells (Figure 29).

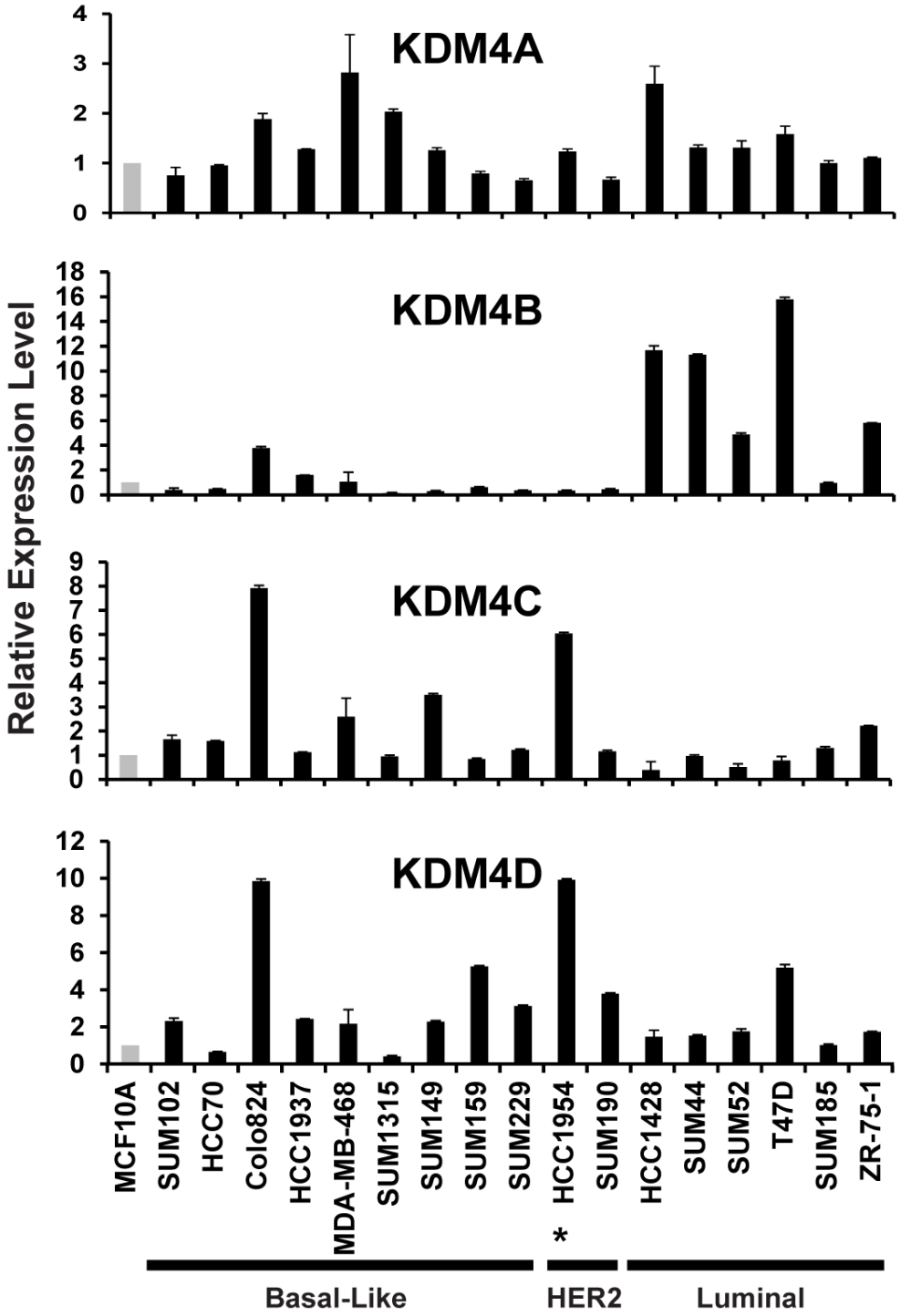


**Figure 28. KDM4 subfamily mRNA expression in breast cancer cell lines.**[1] mRNA expression levels of KDM4 demethylases in a panel of breast cancer cell lines were determined by qRT-PCR. mRNA expression levels in the MCF10A cells, an immortalized but nontumorigenic breast epithelial cell line, were arbitrarily set as 1. Relative expression levels in breast cancer cell lines are shown as fold changes compared with that of MCF10A cells. *The HCC1954 line belongs to the basal-like subtype even though it contains HER2 amplification. Thanks to Qin Ye for help with parts of this figure.

*Open access copyright permissions for this article permitted re-use of this figure from the American Journal of Cancer Research.*

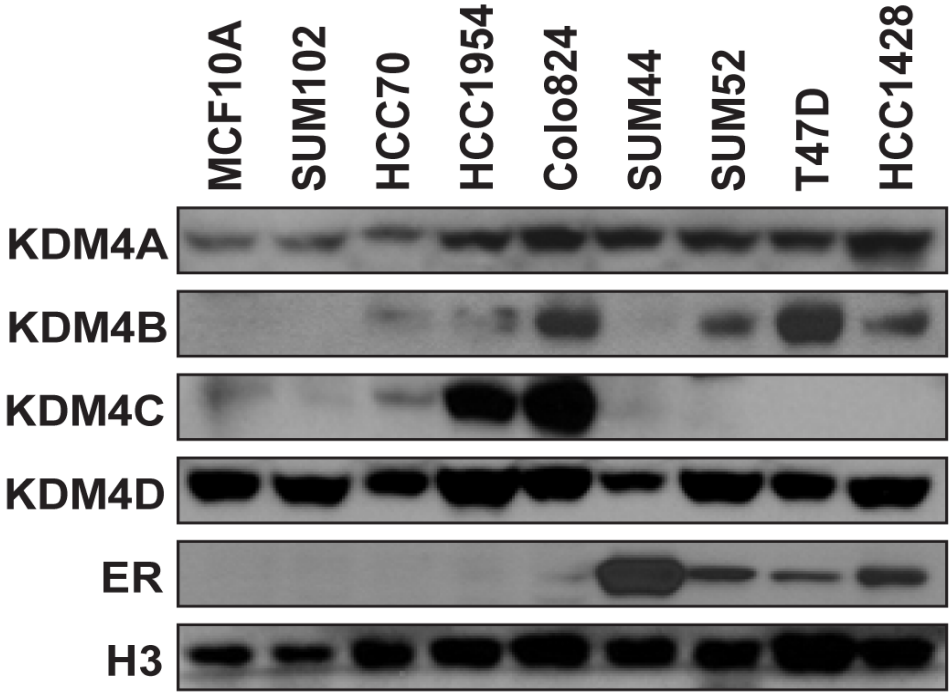


**Figure 29. KDM4 subfamily protein expression in breast cancer cell lines.**[1] KDM4A, B, C, D and Estrogen Receptor (ER) protein levels were analyzed by western blot in eight breast cancer cell lines and the MCF10A non-tumorigenic mammary epithelial cell line. Total H3 was used as the loading control. Thanks to Qin Ye for help with parts of this figure.
*Open access copyright permissions for this article permitted re-use of this figure from the American Journal of Cancer Research.*

# RESULTS:

## CHARACTERIZATION OF THE MOLECULAR PROFILE OF GENE ASSOCIATED WITH TIME TO BREAST CANCER IN BENIGN BREAST LESIONS

To identify networks and diseases associated with genes significantly associated with time to breast cancer, with help from the Biostatistics Core at Karmanos Cancer Institute, IPA mapped the 18,749 transcripts to 8,075 unique genes (Figure 19). Next, 1,805 of these genes (p-value<0.01) were analyzed to identify genetic networks and functional classifications of genes associated with time to breast cancer diagnosis. Overall, 80.7% of these genes (1,454 of 1,802 genes) were found to be up-regulated in association with shorter time to breast cancer diagnosis. Select significant genes were subsequently measured on Nanostring technology for all 36 samples in the cohort by collaborators at Mayo Clinic Cancer Center in Jacksonville, Florida, to validate the expression level differences observed by microarray (Figure 30).

Inflammatory disease ranked as the top disorder (p-value = 0.0483-0.00001) of molecules significantly associated with shorter time to breast cancer diagnosis (data not shown). Gene expression and organ morphology, cellular development, and cell morphology, were the top networks and associated network functions of genes associated with shorter time to breast malignancy (Figure 31).

To gain further insight into the top network of gene expression and organ morphology, molecules were mapped using IPA software and are shown in Figure 32. Of the eleven genes regulated by the estrogen receptor in this network, ten (90.9%) were found to be significantly associated with shorter time to breast cancer diagnosis.

Additional analysis revealed that of these ten molecules, the majority (70%; 7 of 10 molecules) were found to be up-regulated in association with shorter time to breast cancer diagnosis (CDH9, NCAM2, CLDN23, CNNM1, CDH12, PCDH20, FGFR3) (Figure 32).

Examination of the top canonical pathways and overlap of pathway molecules revealed significant overlap with 25.0% of molecules in the lipid antigen presentation by CD1 pathway (p-value = 0.0028) (Figure 33). Notably, this immune strategy uses CD1, a conserved family of major histocompatibility complex-like glycoproteins that capture lipid antigens for presentation to T-lymphocytes and contribute to innate and adaptive immune responses. Significant findings also included overlap with 18.4% of estrogen biosynthesis pathway molecules (p-value=0.0078), consistent with the synthesis of steroid hormones estrone and estradiol in breast tissue that is critical for development and function, as well as induction of tumor growth.

| *Gene Symbol* | *Entrez Gene Name* | *Accession Number* | *Target Region* |
|---|---|---|---|
| **ARF1** | **ADP ribosylation factor 1** | **NM_001024227.1** | **1371-1470** |
| CASP14 | caspase 14 | NM_012114.2 | 627-726 |
| **CD59** | **CD59 molecule** | **NM_000611.4** | **731-830** |
| CEBPA | CCAAT/enhancer binding protein alpha | NM_004364.2 | 1321-1420 |
| **DEAF1** | **DEAF1, transcription factor** | **NM_021008.2** | **1025-1124** |
| EGR2 | early growth response 2 | NM_000399.3 | 1892-1991 |
| **ESR1** | **estrogen receptor 1** | **NM_000125.2** | **1596-1695** |
| **ESR2** | **estrogen receptor 2** | **NM_001214903.1** | **167-266** |
| FGFR3 | fibroblast growth factor receptor 3 | NM_022965.2 | 3171-3270 |
| **GNB1** | **G protein subunit beta 1** | **NM_002074.2** | **2166-2265** |
| GRB2 | growth factor receptor bound protein 2 | NM_002086.4 | 413-512 |
| **HSD17B1** | **hydroxysteroid (17-beta) dehydrogenase 1** | **NM_000413.2** | **1575-1674** |
| **HUNK** | **hormonally up-regulated Neu-associated kinase** | **NM_014586.1** | **1281-1380** |
| **ITCH** | **itchy E3 ubiquitin protein ligase** | **NM_001257138.1** | **439-538** |
| **KAT2B** | **lysine acetyltransferase 2B** | **NM_003884.3** | **1221-1320** |
| **KDM4C** | **lysine demethylase 4C** | **NM_001146694.1** | **2033-2132** |
| **LCK** | **LCK proto-oncogene, Src family tyrosine kinase** | **NM_005356.2** | **1261-1360** |
| **LY6E** | **lymphocyte antigen 6 complex, locus E** | **NM_002346.2** | **381-480** |
| **PAX6** | **paired box 6** | **NM_000280.3** | **1174-1273** |
| POU5F1 | POU class 5 homeobox 1 | NM_002701.4 | 1226-1325 |
| **RHOA** | **ras homolog family member A** | **NM_001664.2** | **1231-1330** |
| **RIPK3** | **receptor interacting serine/threonine kinase 3** | **NM_006871.3** | **1541-1640** |
| **RNF11** | **ring finger protein 11** | **NM_014372.4** | **725-824** |
| **RUNX3** | **runt related transcription factor 3** | **NM_004350.1** | **2086-2185** |
| SALL4 | spalt-like transcription factor 4 | NM_020436.3 | 3225-3324 |
| SOX11 | SRY-box 11 | NM_003108.3 | 5651-5750 |
| **TBP** | **TATA-box binding protein** | **NM_001172085.1** | **588-687** |
| **XRCC5** | **X-ray repair complementing defective repair in Chinese hamster cells 5** | **NM_021141.3** | **833-932** |

**Figure 30. Differentially expressed gene custom code set (gene symbol, gene name, accession number, and target region) for Nanostring analysis.** Samples were loaded onto the Nanostring PrepStation for processing, and placed into the nCounter cartridge. The cartridge was transferred to the nCounter digital analyzer for image capture and data acquisition of fluorescent reporters. Measurements were taken at high sensitivity with 555 FOV. The positive controls measured on each array were used to normalize the observed nanostring reads, using linear regression to estimate the adjustment factor for each sample. The agreement between normalized Affymetrix probe and Nanostring gene expression values was analyzed by Spearman correlation using the software R. Nanostring validation was performed by collaborators at Mayo Clinic, Jacksonville, Florida.
Boldface text indicates a correlation p-value of < 0.05.

| *Networks and Associated Network Functions* | *Molecules in Network* | *Number of Focus Molecules* |
|---|---|---|
| Gene Expression, Organ Morphology | **ABCA4, ACTB,** ADCY, **ASB9, BCL11B, BRF1, C9, CASP14, CD58, CD59, CDH9, CDH12, CLDN23, CNNM1,** estrogen receptor, **FGFR3, GNAI2, GNB1, GNGT1, GTF2E1, HIST1H1A,** Hsp27, Hsp70, **LCK, LTB, LY6E, MAP1B, NCAM2, PCDH20, SOX11, TAF2, TAF5, TAF9, TAF1A, TBP.** | 31 |
| Cell Death and Survival, Cellular Growth and Proliferation, Cellular Development | **ACTR2, AP2B1, CCL4L1/CCL4L2, CEBPA, EOMES, EPHX1, FCER1G,** FSH, **GAPDH, GRB2, HELZ, HSD17B1,** Lh, **LST1, MIXL1, PAF1, PAX6, POU5F1, PTP4A1, RAB14, RAB1A, RALGAPA1, RBBP4, RIPK3,** RNA polymerase II, **SALL4, SERPINI1, SLC25A3,** Smad2/3, **SOD2, STS, TAC3, TNFRSF9, VHL, XRCC5.** | 31 |
| Cell Morphology, Organismal Abnormalities | **ANO3, ARF1, BCR, CAPNS1, CD63,** Creb, **DYNLL1, EDARADD, GPR34, HLA-DMB, HPF1, IL1RL2, KSR2, LITAF, mir-506, MIR17HG,** Mlc, **MYLK,** NFkB (complex), **NLRP2, OTUB2, PTPRS, RAB31, RAP1GAP, RFTN1, RGS5, RHOA, RNF11,** Rock, **RTF1, SLC8A1, SPIB, SUMO4, TRPC6, UBE2V1.** | 31 |

**Figure 31. Top three networks, associated network functions, molecules in network, and number of focus molecules significantly associated with shorter time to breast cancer diagnosis among thirty-six black women with benign breast lesions.** Boldface text indicates genes that were available in the dataset and significantly associated with shorter time to breast cancer diagnosis.

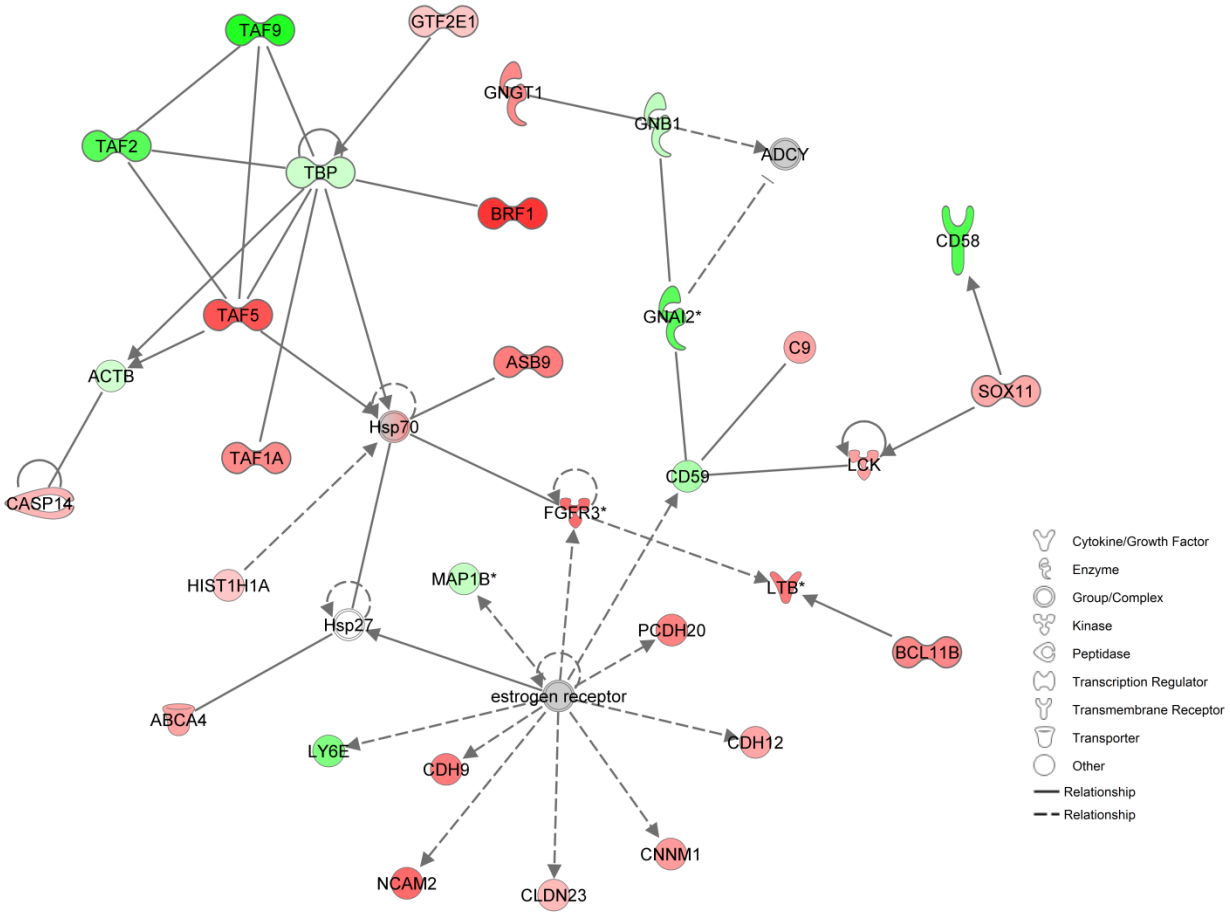


**Figure 32. Network of gene expression and organ morphology.** Red indicates a log hazard ratio (LHR) of > 1; green color indicates a LHR of < -1. Grey indicates the gene was not differentially expressed and white indicates the gene was not found to be statistically significant in the dataset. The legend indicates shapes for genes with known functions. Solid lines indicate a direct relationship and dashed lines represent an indirect relationship. *Figure generated using Ingenuity Pathway Analysis software.*

| *Pathway* | *p-value* | *Overlap* | *Overlap Molecules* |
|---|---|---|---|
| Lipid Antigen Presentation by CD1 | 0.0028 | 25.0% | CD3E, PDIA3, AP2B1, FCER1G, PSAP, CANX. |
| Retinoate Biosynthesis I | 0.0106 | 19.4% | SDR9C7, ALDH1A3, ALDH1A2, RDH16, AKR1C4, AKR1B10. |
| Estrogen Biosynthesis | 0.0078 | 18.4% | CYP4F8, CYP1A1, CYP4B1, HSD17B12, AKR1C4, HSD17B1, CYP2C8. |

**Figure 33. Top canonical pathways and overlap between pathway molecules and genes in benign breast lesions significantly associated with shorter time to breast cancer diagnosis.**

## DISCUSSION

Using the Detroit benign breast disease cohort, this analysis of molecular precursors among thirty-six black women with benign breast lesions who subsequently developed breast cancer identified genes involved in inflammatory disease and estrogen biosynthesis associated with shorter time to breast cancer diagnosis. These findings are novel in that this study is the first to identify potential molecular precursors associated with shorter time to breast cancer diagnosis among black women with benign breast lesions.

The inflammatory state of premalignant and malignant lesions is considered an enabling characteristic to promote tumor progression.[151] The top disorder associated with the significant gene set was inflammatory disease, a process where tissues of the body elicit a biological response to pathogens and damaged cells. Epidemiological studies suggest that non-steroidal anti-inflammatory drugs (NSAIDs) contribute to the reduction of breast cancer risk,[75] and Griffith et al. recently published that black women diagnosed with breast cancer demonstrated increased inflammation compared to black controls without a history of disease.[152] Chronic inflammation has been shown to drive breast cancer pathologies, with recent studies beginning to investigate epigenetic regulation of inflammation, particularly the role of histone lysine demethylases, to target tumorigenic mechanisms.[153-155]

Strikingly, histone lysine demethylase 4C (KDM4C) was the top-upregulated molecule in this analysis, with increased expression of this gene associated with shorter time to breast cancer diagnosis. Previous studies have demonstrated that KDM4C serves as a transforming oncogene and undergoes genetic alterations at a high

frequency in breast cancers.[1,138,156] KDM4C is amplified and overexpressed in aggressive basal-like breast cancers, and associated with poor patient prognosis. KDM4 demethylases function extensively in multiple cellular events throughout organismal development and homeostasis.[113-115,140,142,144,146,157-161] Despite the recent discovery that the KDM4 subfamily plays an essential role in regulating gene expression and chromatin architecture via H3K9 and H3K36 demethylation,[157] there is still much to learn about how KDM4 proteins are recruited to genomic loci, how they modulate histone demethylation and subsequently activate specific downstream targets in different cell types. Moreover, it is clear that KDM4 proteins cooperate in similar macromolecular complexes and processes,[158,160-162] yet the redundancies and interactions between them are still not well understood.

Considering the enormous potential of these epigenetic master regulators in modulating gene transcriptional programs, it is not surprising that their alterations are implicated in human diseases, particularly in cancer.[1,113,114,144] Utilizing The Cancer Genome Atlas breast cancer database and cell line models, KDM4A, B, C and D gene amplification and expression were analyzed relative to different breast cancer subtypes. Consistent with previous findings, *KDM4C* amplification had the highest frequency (12.4%) in basal-like breast cancer compared with other subtypes.[131] Furthermore, high expression of KDM4A, C, and D was found in the basal type and KDM4B was found in ER+ luminal-type breast cancers.

Within the KDM4 subfamily, KDM4A, B, and C show a high degree of homology in sequence and domain organization.[3] Although KDM4 demethylases all catalyze via the same demethylation reaction, recent evidence indicates that their normal cellular

functions are not completely redundant. Possible reasons for these unique functions are: (1) a distinct pattern of intracellular location, (2) cell-type-specific expression, (3) selective recruitment of the different KDM4 demethylases to their target genes, and (4) demethylase specific non-histone proteins.[113-115,140,157-159] Indeed, recent studies revealed that KDM4A is equally present in the cytoplasm and nucleus, KDM4B is more prevalent in the nucleus, and KDM4C strongly associates with chromatin.[159,163] It was also reported that protein levels of KDM4A and B, but not KDM4C, are highly regulated by ubiquitination and the proteasome since KDM4A and B are direct substrates of the E3 ubiquitins RNF8 and RNF168.[160] In contrast, inositol pyrophosphates regulate KDM4C-dependent histone demethylation.[164] More recent studies showed that KDM4B and C have distinct and combinatorial functions in mouse embryonic stem cell identity.[157] Most likely, the four KDM4 subfamily members have substantially overlapping but not completely redundant functions, and their functions are cell-type specific and context-dependent. Because epigenetic changes are reversible and histone demethylases are druggable,[128,144,165-167] despite challenges of overlapping targets they remain promising therapeutic targets for chemoprevention and intervention strategies.

Estrogen biosynthesis is a critical mechanism to produce estrogens involved in regulation of the female reproductive system. The synthesis of estrogen is widely regulated by the aromatase enzyme, a member of the cytochrome P450 superfamily.[168] Aromatase is comprised of a regulatory region of ten tissue-specific promoters for local estrogen biosynthesis in both normal physiological and breast cancer conditions.[169] Estrogen biosynthesis was identified as one of the top canonical pathways of differentially expressed genes. Higher expression levels of four cytochrome P450 family

member genes (CYP4F8, CYP1A1, CYP4B1, CYP2C8) in the estrogen biosynthesis pathway were significantly associated with shorter time to breast cancer diagnosis. Further, most estrogen-mediated signaling pathways are estrogen receptor-dependent and change during the ageing process.[170,171] This analysis also revealed that ten genes regulated by the estrogen receptor, including the increased expression of fibroblast growth factor receptor 3 (FGFR3), were associated with shorter time to breast cancer diagnosis. Notably, previous studies have demonstrated that FGFR3 plays a critical role in breast cancer development and response to endocrine therapy.[172,173] FGFR3 activation stimulates activation of the mitogen-activated protein kinase (MAPK) and phosphoinositide 3-kinase (PI3K) signaling pathways *in vitro*, and reduces sensitivity of breast cancer cells to tamoxifen.[173,174]

Activation of estrogen signaling pathways through inflammation-mediated upregulation of aromatase is critical to breast carcinogenesis. The severity of breast inflammation correlates with aromatase activity, as increased aromatase levels lead to enhanced estrogen biosynthesis and progesterone receptor upregulation.[175] Risk of developing estrogen and progesterone receptor-positive breast cancers are significantly increased among overweight and obese women,[176] as obesity is also associated with inflammation via elevated levels of circulating pro-inflammatory mediators that promote growth and tumorigenesis.[177,178] Dandona et al. demonstrated that increased levels of pro-inflammatory mediators correlate with poor prognosis among obese patients with breast cancer.[179] Regardless of breast cancer subtype, overweight and obese women suffer from greater tumor burden and worse disease-related outcomes versus leaner counterparts.[180,181] Overweight and obesity rates are nearly 1.5 times higher among

black adults compared to whites,[182] and black women have been shown to have the highest BMI among gender and race/ethnicity subgroups.[183] Aligned with this evidence, genes that are involved in inflammatory disease and estrogen biosynthesis were found to be associated with shorter time to breast cancer diagnosis among the cohort of black women, of which the majority was found to be overweight and obese.

The indirect association of the estrogen receptor and DNA influences the activity of transcription factors, and has been found to promote breast cancer cell proliferation through co-regulation of key cell cycle regulators.[184,185] It was noted that cellular growth and proliferation, death and survival, and development, also ranked as the top canonical pathway for genes associated with a shorter time to breast cancer diagnosis in benign lesions. Profiling the molecular characteristics of benign breast lesions from black women who subsequently developed breast cancer identified DEAF1, a transcription factor, as the top-downregulated molecule in this analysis. Prior reports have discovered that DEAF1 contributes to the proliferation of mammary epithelial cells and ductal side-branching to regulate normal development and breast cancer.[186,187] Breast cancer most commonly arises from the mammary ductal epithelium, and consists of tumors with various pathologic and molecular characteristics—which are determinants for metastatic behavior and clinical outcome. Previous studies demonstrated that black women with a history of benign breast disease have a higher risk of subsequent breast cancer.[2,188] Thus, further work is warranted to analyze the molecular and pathological characteristics of benign breast lesions and subsequent breast cancers developed among these black women to evaluate potential precursor molecules and biomarkers for increased surveillance and detection of cancer at early stages of disease.

The review of medical records to collect data on body mass index and additional information on important breast cancer risk factors, including history of tobacco use, age at menarche, and menopausal status, is a strength of this study. Complete information on breast density was unable to be collected, which would have provided a more detailed assessment of the biopsy tissue microenvironment. Yet linkage of benign breast disease cases to the Metropolitan Detroit Cancer Surveillance System (MDCSS) for identification of subsequent breast cancer development is another strength, as it allowed for detailed patient follow-up and identification of pathologically-verified cases. Use of the Affymetrix platform for gene expression profiling of benign FFPE samples and data processing also provided accurate interpretation of results.[189] Limitations of sample validation include using Nanostring technology to confirm these results, given the limited target regions of probes for genes. A larger sample size for transcriptome profiling of benign breast lesions is needed to further examine the predictive value of these molecular precursors to predict time to cancer, and additional studies are called for to see if these molecular markers are associated with subsequent breast cancer risk.

Taken together, these findings elucidate molecular precursors of benign breast lesions in black women that are associated with shorter time to breast cancer. While breast cancer survival rates have improved over the last two decades, breast cancer still accounts for 14% of all cancer-related deaths among women annually.[8] Identification of top canonical pathways and networks of genes associated with shorter time to breast cancer diagnosis can provide further insight into critical mechanisms of tumorigenesis and progression to a malignant state. Additional studies analyzing the molecular profiles of benign breast and breast cancer lesions among women,

particularly the disparate black population, can contribute to improvements in surveillance and detection, and ultimately improved patient survival.

## PATIENTS AND METHODS

*Study design and patient cohort*

This study population comprises patients selected from the metropolitan Detroit benign breast disease (BBD) cohort, which has been described previously by Cote et al.[2] The BBD cohort is composed of 3,759 women aged 21 to 88 years who self-reported African American/black race from metropolitan Detroit, Michigan, and who had been diagnosed with BBD between 1997 and 2010 at hospitals and clinics associated with the Detroit Medical Center and Wayne State University Department of Pathology. Exclusion criteria included: a previous breast biopsy, a history of invasive or in situ breast carcinoma prior to, or within 6 months, of the BBD biopsy, unilateral or bilateral mastectomy prior to or at diagnosis, prior breast reduction surgery, lipoma, fat necrosis, epidermal cysts, hematoma, accessory structure, phyllodes tumor, or a lymph node biopsy with no breast tissue.[188] Using the Metropolitan Detroit Cancer Surveillance System (MDCSS), a founding member of the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) program which has been continuously collecting population-based cancer data since 1973,[7] information on subsequent development of breast cancer was obtained.[2]

Thirty-six women selected based on tissue availability from the benign breast disease cohort who subsequently developed breast cancer with available benign breast lesion tissue blocks comprised this study population. The distribution of clinical and demographic characteristics of these patients was described using percentages. All study procedures were approved by the Wayne State University Institutional Review Board.

*Histology review*

Biopsy review has been previously described by Cote et al.[2] Lesions were classified on Dupont and Page criteria[101] into three categories: non-proliferative disease, proliferative disease without atypia, or proliferative disease with atypia (atypical hyperplasia). Pathological features assessed for each biopsy included: apocrine metaplasia, atrophy, calcifications, columnar alterations, cyst, ductal ectasia, ductal hyperplasia, fibroadenoma, fibrosis, intraductal papilloma, mucocele-like lesions, lobular hyperplasia, radial scar, sclerosing adenosis (Figure 12). The distribution of pathological characteristics of breast lesions/biopsies was described using percentages.

*Cumulative incidence*

The cumulative incidence for developing breast cancer was calculated using R statistical software, package 'cmprsk'.[190,191] Time to breast cancer diagnosis was evaluated considering behavior as a competing risk (in-situ versus invasive disease). All women in the metropolitan Detroit BBD cohort were evaluated in this analysis. Of the 3,759 black women with benign breast disease, 217 women developed breast cancer (63 in-situ cancers and 154 invasive cancers).

*Extraction of RNA and gene expression profiling*

RNA was extracted from 97 formalin-fixed, paraffin-embedded (FFPE) benign breast samples from the thirty-six black women using the High Pure RNA Paraffin Kit (Roche Diagnostics, Germany) by collaborators at Mayo Clinic Cancer Center in Jacksonville, Florida. RNA quantity and quality were assessed using the ND-1000

Spectrophotometer (Nanodrop, Wilmington, Delaware). One sample per patient was utilized for further analysis if the optical density 260/280 ratio was ≥1.8, the total RNA yield was ≥500 ng, and based on the highest DV200 value as a measure of RNA quality. Extracted RNA was labeled and hybridized to the Affymetrix human transcriptome array 2.0 (HTA 2.0) exon array chip containing > 6.0 million distinct probes covering 44,699 protein-coding genes (transcript clusters) and 22,829 non-protein-coding genes according to manufacturer's protocol at the Mayo Clinic Cancer Center.

*Transcriptomic and ingenuity pathway analyses*

Sample probe gene expression values were exported from Affymetrix and imported into the software R[190] for normalization, additional quality control, and analysis using the Bioconductor 'oligo' software package.[192] Working with the Biostatistics Core at Karmanos Cancer Institute, data were background corrected using the robust multi-array average algorithm, normalized using quantile normalization, and summarized using median polish. Cox proportional hazards models were used to determine which transcripts were associated with time-to-breast cancer.

Log hazard ratios (LHR) and standard error (SE) were estimated for each transcript after adjusting for age at breast biopsy (years), and benign breast biopsy overall impression (non-proliferative disease, proliferative disease without atypia, proliferative disease with atypia). LHRs > 0 indicate that a gene expression increase is associated with an earlier time-to-cancer diagnosis. LHRs < 0 indicate that a gene expression decrease is associated with an earlier time-to-cancer diagnosis. The

microarray data discussed in this publication will be deposited into the NCBI's Gene Expression Omnibus.

Transcripts with a p-value < 0.05 and a LHR of at least 1 (absolute value) were considered to be significant (Figure 19). Genes that were statistically significant at the 0.01 level were subsequently entered into Ingenuity Pathway Analysis (IPA) software (Ingenuity Systems, Mountain View, California) with the help of the Karmanos Cancer Institute Biostatistics Core to determine the most significantly enriched functions, networks, diseases, and pathways.

*RNA validation and correlation analysis*

Twenty-eight statistically significant genes (Figure 30) and ten housekeeping genes were submitted to create a custom code set for Nanostring technology analysis to confirm the results on alternative technology. The assay was performed at Mayo Clinic Cancer Center in Jacksonville, Florida, according to manufacturer's protocol (nCounter XT CodeSet Gene Expression Assay). Briefly, 100ng of extracted RNA was hybridized with the Reporter CodeSet and Capture ProbeSet for 18 hours at 65°C. Samples were loaded onto the Nanostring PrepStation for processing, and placed into the nCounter cartridge. The cartridge was transferred to the nCounter digital analyzer for image capture and data acquisition of fluorescent reporters. Measurements were taken at high sensitivity with 555 FOV. The positive controls measured on each array were used to normalize the observed nanostring reads, using linear regression to estimate the adjustment factor for each sample with help from the Biostatistics Core at the Karmanos Cancer Institute in Detroit, Michigan. The agreement between normalized

Affymetrix probe and Nanostring gene expression values was analyzed by Spearman correlation using the software R.

*Cell culture*

The cultures for the SUM series of breast cancer cell lines and nontransformed human mammary epithelial cell MCF10A line have been described in detail previously.[193,194] The Colo824 cell line was obtained from DSMZ, the SUM cell lines were obtained from Dr. Stephen P. Ethier, and all other cell lines in this study were obtained from ATCC (Manassas, VA). These lines were maintained in RPMI with 10% FBS (Atlanta Biologicals, Flowery Branch, GA, USA) according to DSMZ and ATCC protocols.

*The Cancer Genome Atlas (TCGA) data for breast cancer*

The DNA copy number, mutation, and RNA sequencing datasets of 976 breast cancer samples used in this research were obtained from the cBio Cancer Genomics Portal.[148,149] The copy number of each histone lysine demethylase was generated from the copy number analysis algorithms GISTIC (Genomic Identification of Significant Targets in Cancer) and categorized as copy number level per gene: “-2” is a deep loss (possibly a homozygous deletion), “-1” is a heterozygous deletion, “0” is diploid, “1” indicates a low-level gain, and “2” is a high-level amplification.

For mRNA expression data, the relative expression of an individual gene and the gene expression distribution in a reference population were analyzed. The reference population was either all tumors that are diploid for the gene in question, or, when available, normal adjacent tissue. The returned value indicates the number of standard

deviations away from the mean of expression in the reference population (Z-score). Somatic mutation data were obtained from exome sequencing.[148,149] The breast cancer subtype information was from a previous publication and the cBio Cancer Genomics Portal.[18,148,149]

*Semiquantitative RT-PCR reactions*

mRNA was prepared from human breast cancer cell lines and the MCF10A cell line by using an RNeasy Plus Mini Kit (QIAGEN). mRNA was mixed with qScript cDNA SuperMix (Quanta Biosciences, Gaithersburg, MD, USA), then converted into cDNA through a reverse-transcription reaction for real-time PCR reactions. Primer sets for genes of interest were ordered from Life Technologies (Carlsbad, CA, USA). A PUM1 primer set was used as a control. Semiquantitative RT-PCR was done using the FastStart Universal SYBR Green Master (Rox) (Roche Diagnostics Indianapolis, IN, USA).

*Immunoblotting and antibodies*

Whole cell lysates were prepared by scraping cells from the dishes into cold RIPA lysis buffer. Nuclear protein extracts from breast cancer cells and the MCF10A cells were prepared with an NE-PER Nuclear Protein Extraction Kit (Thermo Scientific, Rockford, IL, USA). Histone proteins from cells were isolated with the EpiQuik Total Histone Extraction Kit (Epigentek, Farmingdale, NY, USA). Protein content was estimated with the Bradford method.

A total of 20-100μg of total cell lysate was resolved by SDS-PAGE and transferred onto PVDF membrane. Antibodies used in the study included anti-KDM4D (Abcam ab93694, Cambridge, MA, USA), anti-KDM4B and C (Bethyl Laboratories A301-478A, Montgomery, TX, USA), anti-KDM4A, anti-H3, anti-H3K4me2, anti-H3K4me1, anti-H3K9me1, anti-H3K36me3, anti-H3K36me2 (Cell Signaling, Danvers, MA, USA), anti-H3K4me3, anti-H3K9me3, and anti-H3K36me1 (Active Motif, Carlsbad, CA, USA).

*Statistical analysis for KDM4 studies*

Statistical analyses were performed using R software (http://www.r-project.org) and Stata.[190,195] The correlations between copy numbers and mRNA levels of each histone lysine demethylase from 976 sequenced breast cancer specimens were analyzed using Spearman, Kendall, and Pearson correlation tests. The Spearman and Kendall tests are rank correlations: the Spearman coefficient relates the two variables while conserving the order of data points, and the Kendall coefficient measures the number of ranks that match in the data set. Although the Pearson correlation coefficient is the most widely used, it was deemed the least relevant to this study, as it measures only the strength of linear relationships and ignores all others. The “cor” function in R statistical software was used for computation, specifying in the code which type of test was desired (Spearman, Kendall, or Pearson). The difference in mRNA expression level for each histone lysine demethylase between the basal-like and the other cancer subtypes was calculated using Student’s *t*-test.

# CHAPTER 4. GENERAL DISCUSSION

Each year, over 1.5 million women undergo breast biopsies.[94,95] Among these women, black individuals are more likely to have been referred for biopsy as follow-up from recommended mammograms, even despite lower rates of screening.[196] While the majority of these biopsies performed result in benign findings,[96] the pathological characteristics of some of these benign breast lesions still confer an increased risk of subsequent breast cancer diagnosis in white and black populations.[2,99] These pathological features, including columnar alterations, fibroadenoma, sclerosing adenosis, cysts, and calcifications, contribute to the categorization of the benign breast lesion into loose groups for an overall impression according to Dupont and Page criteria,[101] and are highly prognostic for breast cancer risk in populations of black women.[2,188] This work is especially important given the racial disparities that exist in breast cancer incidence and survival, where black women are more likely to develop young-onset breast cancer with the disease diagnosed at more aggressive clinical stages.[6,8]

Identification of women at the highest risk of disease has clinical implications, including: increased surveillance, diagnosis at early stage of malignancy, chemoprevention and intervention, and ultimately, improved survival outcomes. The development of risk prediction models to assess cancer risk on a personalized patient basis began in the 1990s. Perhaps the most appropriate model to predict invasive breast cancer risk for black women is the recently established CARE model that is tailored to this high-risk racial/ethnic population.[197] Specific variables used to provide breast cancer risk estimates in the CARE model include: number of first-degree

relatives with a history of breast cancer, menarche age, age at first live birth, current age, and number of previous benign breast biopsies.[197] While this model takes into account the number of previous benign breast biopsies, there is no incorporation of detailed pathological or molecular findings from breast biopsies. Further, few models consider benign breast pathological findings in risk prediction, and at best construct analyses based on the overall impression of the benign breast lesion. It is important to note there are limitations associated with pathological impressions, including differences in histological spectra and clinical manifestations of these lesions. Therefore, molecular findings among benign breast lesions that predict shorter time to breast cancer diagnosis could be instrumental to the improvement of breast cancer risk characterization among black women with a history of benign breast disease.

Profiling the clinicopathology and molecular characteristics of benign breast lesions among black women who subsequently developed breast cancer in the Detroit cohort provides insight into both the histological and genomic scopes of benign breast disease, particularly in an understudied population with poor survival outcomes. These analyses identified the top up-regulated genes, including histone lysine demethylase 4C (KDM4C), pathways, diseases, and networks, including inflammation and estrogen biosynthesis, that are associated with a shorter time to breast cancer. These findings warrant further investigations to elucidate the dysregulation and molecular mechanisms of these modulators that contribute to breast carcinogenesis.

The clinical implications of these molecular findings are also a paradigm for clarifying future risk of breast cancer development and shorter time to breast cancer diagnosis among black women with benign breast disease. Stratifying the risk for 1.5

million women who undergo breast biopsies annually is critical to identify those at highest risk of subsequent malignancy who could be best served with allocation of additional interventions and clinical management, as breast cancer ranks as the second leading cause of cancer related deaths annually in the United States.[8] Elucidation of the biological basis of disparity in this high-risk population can contribute to ultimately understanding critical biological differences in patient prognosis.

On a molecular level, breast cancer progression is a complex, multifactorial process traditionally viewed as the stepwise accumulation of genetic alterations. Furthermore, through the use of next-generation sequencing, it has become apparent that vital epigenetic regulators, such as histone methyltransferases and demethylases, undergo genetic alterations at a high frequency in aggressive breast cancer. Together, these chromatin modifications play critical roles in controlling transcription, chromatin architecture, and cellular differentiation. However, the genomic landscape and clinical significance of histone lysine demethylases in breast cancer remain poorly characterized.

Dysregulation of the KDM4 demethylases has been documented in a variety of cancers, including lymphoma, medulloblastoma, and breast, prostate, colorectal, lung, gastric, esophageal, and renal cancers.[1,131,132,134,139,166,198,199] The KDM4 demethylases, A, B, C, and D, were the first identified demethylases to act on tri-methylated lysines.[133,143] Given these data and the finding that KDM4C was the top up-regulated molecule associated with shorter time to breast cancer diagnosis in benign breast lesions of black women, a meta-analysis of KDM4A, B, C and D in breast cancer was conducted and identified associations among recurrent copy number alterations, gene

expression and breast cancer subtypes. KDM4A, C and D are also significantly overexpressed in basal-like breast cancer, whereas KDM4B overexpression is more dominant in estrogen-receptor-positive, luminal breast cancer. These findings demonstrate genetic amplification and overexpression of the KDM4 demethylases in different subtypes of breast cancer, and provide information regarding the genomic and transcriptomic alterations of the KDM4 subfamily in different subtypes of breast cancer for a better understanding into the mechanisms of breast carcinogenesis.

Using a large-scale cancer genomics data set, four KDM genes (KDM1B, KDM4C, KDM5A, and JARID2) were highly amplified (>10%) and overexpressed in basal-like breast cancer.[156] Basal-like breast cancers are generally triple-negative, defined by lack of estrogen receptor, progesterone receptor, and HER2 oncoprotein expression.[17] Basal-like tumors are associated with higher rates of metastasis and death, and given the lack of effective molecularly targeted therapies developed, the treatment for basal breast cancer consists of standard chemotherapy regimens.[17,18] Thus, the establishment and validation of novel therapeutic strategies that inhibit progression of breast cancer remains a primary challenge. Since histone methylation is reversible and demethylases are druggable targets, histone lysine demethylase inhibitors may serve as a novel approach to target a subset of aggressive, basal-like breast cancers.

Understanding the genetic and epigenetic abnormalities that are associated with breast cancer subtypes will help identify novel subtype-specific targets for therapy. These findings add significant information to the genomic and transcriptomic profiles of histone lysine methyltransferases and demethylases in different breast cancer

subtypes. Indeed, breast cancer subtypes show distinct copy number alteration patterns and differential expression for each demethylase gene. These findings add layers of information to the genomic and transcriptomic profiles of the KDM4 subfamily in different subtypes of breast cancer. These results also lay the foundation for future studies to pre-clinically validate histone lysine demethylase inhibition as a therapeutic strategy for different subtypes of breast cancer.

As the most commonly diagnosed neoplasm among women in the United States, there exists a critical need for novel insights into the molecular underpinnings of breast cancer.[8] Together, these findings translate molecular and pathological findings into potential clinical benefits for breast cancer patients by understanding the clinicopathology and molecular precursors underlying benign breast and breast cancer lesions to provide insight into breast tumorigenesis.

# LITERATURE CITED

1. Ye Q, Holowatyj A, Wu J, et al: Genetic alterations of KDM4 subfamily and therapeutic effect of novel demethylase inhibitor in breast cancer. Am J Cancer Res 5:1519-1530, 2015.

2. Cote ML, Ruterbusch JJ, Alosh B, et al: Benign breast disease and the risk of subsequent breast cancer in African American women. Cancer Prev Res (Phila) 5:1375-1380, 2012.

3. Labbe RM, Holowatyj A, Yang ZQ: Histone lysine demethylase (KDM) subfamily 4: structures, functions and therapeutic potential. Am J Transl Res 6:1-15, 2013.

4. Holowatyj AN, Ruterbusch JJ, Ratnam M, et al: HER2 status and disparities in luminal breast cancers. Cancer Med 5:2109-2116, 2016.

5. Holowatyj A, Yang ZQ, Pile LA: Histone lysine demethylases in Drosophila melanogaster. Fly (Austin) 9:36-44, 2015.

6. American Cancer Society. Breast Cancer Facts & Figures 2015-2016. Atlanta: American Cancer Society, Inc. 2015.

7. Howlader N, Krapcho M, Miller D, Bishop K, Altekruse SF, Kosary CL, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ, Cronin KA (eds). SEER Cancer Statistics Review, 1975-2013, National Cancer Institute. Bethesda, MD, http://seer.cancer.gov/csr/1975_2013/, based on November 2015 SEER data submission, posted to the SEER web site, April 2016.

8. Siegel RL, Miller KD, Jemal A: Cancer statistics, 2016. CA Cancer J Clin 66:7-30, 2016.

9. National Comprehensive Care Network: Breast Cancer. NCCN Clinical Practice Guidelines in Oncology, 2015.

10. Perou CM, Borresen-Dale AL: Systems biology and genomics of breast cancer. Cold Spring Harb Perspect Biol 2011.

11. Early Breast Cancer Trialists' Collaborative Group: Tamoxifen for early breast cancer. Cochrane Database Syst Rev 2001.

12. Rastelli F, Crispino S: Factors predictive of response to hormone therapy in breast cancer. Tumori 94:370-383, 2008.

13. Blows FM, Driver KE, Schmidt MK, et al: Subtyping of breast cancer by immunohistochemistry to investigate a relationship between subtype and short and long term survival: a collaborative analysis of data for 10,159 cases from 12 studies. PLoS Med 7, 2010.

14. Wang X, Belguise K, Kersual N, et al: Oestrogen signalling inhibits invasive phenotype by repressing RelB and its target BCL2. Nat Cell Biol 9:470-478, 2007.

15. Kariagina A, Xie J, Langohr IM, et al: Progesterone decreases levels of the adhesion protein E-cadherin and promotes invasiveness of steroid receptor positive breast cancers. Horm Cancer 4:371-380, 2013.

16. Lin VC, Eng AS, Hen NE, et al: Effect of progesterone on the invasive properties and tumor growth of progesterone receptor-transfected breast cancer cells MDA-MB-231. Clin Cancer Res 7:2880-2886, 2001.

17. Carey LA, Perou CM, Livasy CA, et al: Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. JAMA 295:2492-2502, 2006.

18. The Cancer Genome Atlas Network: Comprehensive molecular portraits of human breast tumours. Nature 490:61-70, 2012.

19. Perou CM, Sorlie T, Eisen MB, et al: Molecular portraits of human breast tumours. Nature 406:747-752, 2000.

20. Yang XR, Chang-Claude J, Goode EL, et al: Associations of breast cancer risk factors with tumor subtypes: a pooled analysis from the Breast Cancer Association Consortium studies. J Natl Cancer Inst 103:250-263, 2011.

21. DeSantis CE, Siegel RL, Sauer AG, et al: Cancer statistics for African Americans, 2016: Progress and opportunities in reducing racial disparities. CA Cancer J Clin 66:290-308, 2016.

22. Morris GJ, Mitchell EP: Higher incidence of aggressive breast cancers in African-American women: a review. J Natl Med Assoc 100:698-702, 2008.

23. D'Arcy M, Fleming J, Robinson WR, et al: Race-associated biological differences among Luminal A breast tumors. Breast Cancer Res Treat 152:437-448, 2015.

24. O'Brien KM, Cole SR, Tse CK, et al: Intrinsic breast tumor subtypes, race, and long-term survival in the Carolina Breast Cancer Study. Clin Cancer Res 16:6100-6110, 2010.

25. Roseland ME, Pressler ME, Lamerato LE, et al: Racial differences in breast cancer survival in a large urban integrated health system. Cancer 121:3668-3675, 2015.

26. Chowdhury R, David N, Bogale A, et al: Assessing the Key Attributes of Low Utilization of Mammography Screening and Breast-self Exam among African-American Women. J Cancer 7:532-537, 2016.

27. George P, Chandwani S, Gabel M, et al: Diagnosis and surgical delays in African American and white women with early-stage breast cancer. J Womens Health 24:209-217, 2015.

28. Hill J, Watanabe-Galloway S, Shostrom V, et al: Breast Cancer Survival among African-Americans Living in the Midwest: Disparities and Recommendations to Decrease Mortality. J Natl Black Nurses Assoc 26:8-14, 2015.

29. Mortel M, Rauscher GH, Murphy AM, et al: Racial and Ethnic Disparity in Symptomatic Breast Cancer Awareness despite a Recent Screen: The Role of Tumor Biology and Mammography Facility Characteristics. Cancer Epidemiol Biomarkers Prev 24:1599-1606, 2015.

30. Alluri P, Newman LA: Basal-like and triple-negative breast cancers: searching for positives among many negatives. Surg Oncol Clin N Am 23:567-577, 2014.

31. Althuis MD, Brogan DD, Coates RJ, et al: Breast cancers among very young premenopausal women (United States). Cancer Causes Control 14:151-160, 2003

32. Amirikia KC, Mills P, Bush J, et al: Higher population-based incidence rates of triple-negative breast cancer among young African-American women: Implications for breast cancer screening recommendations. Cancer 117:2747-2753, 2011.

33. Anders CK, Johnson R, Litton J, et al: Breast cancer before age 40 years. Semin Oncol 36:237-249, 2009.

34. Anders CK, Hsu DS, Broadwater G, et al: Young age at diagnosis correlates with worse prognosis and defines a subset of breast cancers with shared patterns of gene expression. J Clin Oncol 26:3324-3330, 2008.

35. Chan A, Pintilie M, Vallis K, et al: Breast cancer in women < or = 35 years: review of 1002 cases from a single institution. Ann Oncol 11:1255-1262, 2000.

36. Freedman RA, Virgo KS, Labadie J, et al: Receipt of locoregional therapy among young women with breast cancer. Breast Cancer Res Treat 135:893-906, 2012.

37. McGuire KP, Santillan AA, Kaur P, et al: Are mastectomies on the rise? A 13-year trend analysis of the selection of mastectomy versus breast conservation therapy in 5865 patients. Ann Surg Oncol 16:2682-2690, 2009.

38. Bleyer A, Barr R, Hayes-Lattin B, et al: The distinctive biology of cancer in adolescents and young adults. Nat Rev Cancer 8:288-298, 2008.

39. Colleoni M, Rotmensz N, Robertson C, et al: Very young women (<35 years) with operable breast cancer: features of disease at presentation. Ann Oncol 13:273-279, 2002.

40. Nixon AJ, Neuberg D, Hayes DF, et al: Relationship of patient age to pathologic features of the tumor and prognosis for patients with stage I or II breast cancer. J Clin Oncol 12:888-894, 1994.

41. Winchester DP, Osteen RT, Menck HR: The National Cancer Data Base report on breast carcinoma characteristics and outcome in relation to age. Cancer 78:1838-1843, 1996.

42. Freedman RA, Partridge AH: Management of breast cancer in very young women. Breast 2:176-179, 2013.

43. Partridge A.H. GA, Gelber S., Gelber R. Breast Cancer in Young Women. L. Harris, Osborne Morrow (Eds.), Diseases of the Breast (4th ed), Lippincott Williams & Wilkins, Philadelphia (2010), pp. 1073–1082.

44. Lalloo F, Varley J, Moran A, et al: BRCA1, BRCA2 and TP53 mutations in very early-onset breast cancer with associated risks to relatives. Eur J Cancer 42:1143-1150, 2006.

45. Churpek JE, Walsh T, Zheng Y, et al: Inherited predisposition to breast cancer among African American women. Breast Cancer Res Treat 149:31-39, 2015.

46. Pal T, Bonner D, Cragun D, et al: A high frequency of BRCA mutations in young black women with breast cancer residing in Florida. Cancer 121:4173-4180, 2015.

47. Armstrong K, Micco E, Carney A, et al: Racial differences in the use of BRCA1/2 testing among women with a family history of breast or ovarian cancer. JAMA 293:1729-1736, 2005.

48. Jagsi R, Griffith KA, Kurian AW, et al: Concerns about cancer risk and experiences with genetic testing in a diverse population of patients with breast cancer. J Clin Oncol 33:1584-1591, 2015.

49. Sheppard VB, Graves KD, Christopher J, et al: African American women's limited knowledge and experiences with genetic counseling for hereditary breast cancer. J Genet Couns 23:311-322, 2014.

50. Sherman KA, Miller SM, Shaw LK, et al: Psychosocial approaches to participation in BRCA1/2 genetic risk assessment among African American women: a systematic review. J Community Genet 5:89-98, 2014.

51. Collaborative Group on Hormonal Factors in Breast Cancer: Menarche, menopause, and breast cancer risk: individual participant meta-analysis, including 118 964 women with breast cancer from 117 epidemiological studies. Lancet Oncol 13:1141-1151, 2012.

52. Geisler J: Breast cancer tissue estrogens and their manipulation with aromatase inhibitors and inactivators. J Steroid Biochem Mol Biol 86:245-253, 2003.

53. Pasqualini JR, Chetrite G, Blacker C, et al: Concentrations of estrone, estradiol, and estrone sulfate and evaluation of sulfatase and aromatase activities in pre- and postmenopausal breast cancer patients. J Clin Endocrinol Metab 81:1460-1464, 1996.

54. Key TJ: Endogenous oestrogens and breast cancer risk in premenopausal and postmenopausal women. Steroids 76:812-815, 2011.

55. Robbins AS, Lerro CC, Barr RD: Insurance status and distant-stage disease at diagnosis among adolescent and young adult patients with cancer aged 15 to 39 years: National Cancer Data Base, 2004 through 2010. Cancer 120:1212-1219, 2014.

56. Chagpar AB, Crutcher CR, Cornwell LB, et al: Primary tumor size, not race, determines outcomes in women with hormone-responsive breast cancer. Surgery 150:796-801, 2011.

57. Heimann R, Ferguson D, Powers C, et al: Race and clinical outcome in breast cancer in a series with long-term follow-up evaluation. J Clin Oncol 15:2329-2337, 1997.

58. Henson DE, Chu KC, Levine PH: Histologic grade, stage, and survival in breast carcinoma: comparison of African American and Caucasian women. Cancer 98:908-917, 2003.

59. McBride R, Hershman D, Tsai WY, et al: Within-stage racial differences in tumor size and number of positive lymph nodes in women with breast cancer. Cancer 110:1201-1208, 2007.

60. Howlader N, Altekruse SF, Li CI, et al: US incidence of breast cancer subtypes defined by joint hormone receptor and HER2 status. J Natl Cancer Inst 106, 2014.

61. Rauscher GH, Allgood KL, Whitman S, et al: Disparities in screening mammography services by race/ethnicity and health insurance. J Womens Health 21:154-160, 2012.

62. van Ravesteyn NT, Schechter CB, Near AM, et al: Race-specific impact of natural history, mammography screening, and adjuvant treatment on breast cancer mortality rates in the United States. Cancer Epidemiol Biomarkers Prev 20:112-122, 2011.

63. Campesino M, Koithan M, Ruiz E, et al: Surgical treatment differences among Latina and African American breast cancer survivors. Oncol Nurs Forum 39:324-331, 2012.

64. Fancher TT, Palesty JA, Thomas R, et al: A woman's influence to choose mastectomy as treatment for breast cancer. J Surg Res 153:128-131, 2009.

65. Morrow M, Jagsi R, Alderman AK, et al: Surgeon recommendations and receipt of mastectomy for treatment of breast cancer. JAMA 302:1551-1556, 2009.

66. Tate PS, McGee EM, Hopkins SF, et al: Breast conservation versus mastectomy: patient preferences in a community practice in Kentucky. J Surg Oncol 52:213-216, 1993.

67. Blichert-Toft M, Rose C, Andersen JA, et al: Danish randomized trial comparing breast conservation therapy with mastectomy: six years of life-table analysis. Danish Breast Cancer Cooperative Group. J Natl Cancer Inst Monogr:19-25, 1992.

68. Fisher B, Anderson S, Bryant J, et al: Twenty-year follow-up of a randomized trial comparing total mastectomy, lumpectomy, and lumpectomy plus irradiation for the treatment of invasive breast cancer. N Engl J Med 347:1233-1241, 2002.

69. Fisher B, Redmond C, Poisson R, et al: Eight-year results of a randomized clinical trial comparing total mastectomy and lumpectomy with or without irradiation in the treatment of breast cancer. N Engl J Med 320:822-828, 1989.

70. Jatoi I, Proschan MA: Randomized trials of breast-conserving therapy versus mastectomy for primary breast cancer: a pooled analysis of updated results. Am J Clin Oncol 28:289-294, 2005.

71. Lichter AS, Lippman ME, Danforth DN, Jr., et al: Mastectomy versus breast-conserving therapy in the treatment of stage I and II carcinoma of the breast: a randomized trial at the National Cancer Institute. J Clin Oncol 10:976-983, 1992.

72. Sarrazin D, Le MG, Arriagada R, et al: Ten-year results of a randomized trial comparing a conservative treatment to mastectomy in early breast cancer. Radiother Oncol 14:177-184, 1989.

73. van Dongen JA, Bartelink H, Fentiman IS, et al: Randomized clinical trial to assess the value of breast-conserving therapy in stage I and II breast cancer, EORTC 10801 trial. J Natl Cancer Inst Monogr:15-18, 1992.

74. Veronesi U, Saccozzi R, Del Vecchio M, et al: Comparing radical mastectomy with quadrantectomy, axillary dissection, and radiotherapy in patients with small cancers of the breast. N Engl J Med 305:6-11, 1981.

75. Agrawal A, Fentiman IS: NSAIDs and breast cancer: a possible prevention and treatment strategy. Int J Clin Pract 62:444-449, 2008.

76. McCormick B: The mastectomy myth. Lancet Oncol, 2016.

77. Baquet CR, Mishra SI, Commiskey P, et al: Breast cancer epidemiology in blacks and whites: disparities in incidence, mortality, survival rates and histology. J Natl Med Assoc 100:480-488, 2008.

78. Li CI, Malone KE, Daling JR: Differences in breast cancer stage, treatment, and survival by race and ethnicity. Arch Intern Med 163:49-56, 2003.

79. Gianni L, Pienkowski T, Im YH, et al: Efficacy and safety of neoadjuvant pertuzumab and trastuzumab in women with locally advanced, inflammatory, or early HER2-positive breast cancer (NeoSphere): a randomised multicentre, open-label, phase 2 trial. Lancet Oncol 13:25-32, 2012.

80. Cortes J, Fumoleau P, Bianchi GV, et al: Pertuzumab monotherapy after trastuzumab-based treatment and subsequent reintroduction of trastuzumab: activity

and tolerability in patients with advanced human epidermal growth factor receptor 2-positive breast cancer. J Clin Oncol 30:1594-1600, 2012.

81. Guarneri V, Broglio K, Kau SW, et al: Prognostic value of pathologic complete response after primary chemotherapy in relation to hormone receptor status and other factors. J Clin Oncol 24:1037-1044, 2006.

82. 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 26:1275-1281, 2008.

83. Barlow WE, Taplin SH, Yoshida CK, et al: Cost comparison of mastectomy versus breast-conserving therapy for early-stage breast cancer. J Natl Cancer Inst 93:447-455, 2001.

84. Polsky D, Mandelblatt JS, Weeks JC, et al: Economic evaluation of breast cancer treatment: considering the value of patient choice. J Clin Oncol 21:1139-1146, 2003.

85. Katz SJ, Lantz PM, Janz NK, et al: Patient involvement in surgery treatment decisions for breast cancer. J Clin Oncol 23:5526-5533, 2005.

86. Hawley ST, Griggs JJ, Hamilton AS, et al: Decision involvement and receipt of mastectomy among racially and ethnically diverse breast cancer patients. J Natl Cancer Inst 101:1337-1347, 2009.

87. Thomas P, Killelea BK, Horowitz N, et al: Racial Differences in Utilization of Breast Conservation Surgery: Results from the National Cancer Data Base (NCDB). Ann Surg Oncol 23:3272-3283, 2016.

88. Jagsi R, Hawley ST, Griffith KA, et al: Contralateral Prophylactic Mastectomy Decisions in a Population-Based Sample of Patients With Early-Stage Breast Cancer. JAMA Surg, 2016.

89. Ithimakin S, Day KC, Malik F, et al: HER2 drives luminal breast cancer stem cells in the absence of HER2 amplification: implications for efficacy of adjuvant trastuzumab. Cancer Res 73:1635-1646, 2013.

90. Banegas MP, Li CI: Breast cancer characteristics and outcomes among Hispanic Black and Hispanic White women. Breast Cancer Res Treat 134:1297-1304, 2012.

91. Parise CA, Caggiano V: The Influence of Socioeconomic Status on Racial/Ethnic Disparities among the ER/PR/HER2 Breast Cancer Subtypes. J Cancer Epidemiol 8:134-156, 2015.

92. Kish JK, Yu M, Percy-Laurry A, et al: Racial and ethnic disparities in cancer survival by neighborhood socioeconomic status in Surveillance, Epidemiology, and End Results (SEER) Registries. J Natl Cancer Inst Monogr 2014:236-243, 2014.

93. Surveillance Epidemiology and End Results (SEER) Program (www.seer.cancer.gov) SEER*Stat Database: Incidence - SEER 18 Regs Research Data + Hurricane Katrina Impacted Louisiana Cases, Nov 2015 Sub (1973-2013 varying) - Linked To County Attributes - Total U.S., 1969-2013 Counties, National Cancer Institute, DCCPS, Surveillance Research Program, Surveillance Systems Branch. released April 2016, based on the November 2015 submission.

94. Liberman L: Clinical management issues in percutaneous core breast biopsy. Radiol Clin North Am 38:791-807, 2000.

95. Liberman L: Percutaneous image-guided core breast biopsy. Radiol Clin North Am 40:483-500, 2002.

96. Sickles EA, Ominsky SH, Sollitto RA, et al: Medical audit of a rapid-throughput mammography screening practice: methodology and results of 27,114 examinations. Radiology 175:323-327, 1990.

97. Connolly JL, Schnitt SJ: Benign breast disease. Resolved and unresolved issues. Cancer 71:1187-1189, 1993.

98. Goldacre MJ, Abisgold JD, Yeates DG, et al: Benign breast disease and subsequent breast cancer: English record linkage studies. J Public Health 32:565-571, 2010.

99. Hartmann LC, Sellers TA, Frost MH, et al: Benign breast disease and the risk of breast cancer. N Engl J Med 353:229-237, 2005.

100. Kabat GC, Jones JG, Olson N, et al: A multi-center prospective cohort study of benign breast disease and risk of subsequent breast cancer. Cancer Causes Control 21:821-828, 2010.

101. Dupont WD, Page DL: Risk factors for breast cancer in women with proliferative breast disease. N Engl J Med 312:146-151, 1985.

102. Dupont WD, Parl FF, Hartmann WH, et al: Breast cancer risk associated with proliferative breast disease and atypical hyperplasia. Cancer 71:1258-1265, 1993.

103. Marshall LM, Hunter DJ, Connolly JL, et al: Risk of breast cancer associated with atypical hyperplasia of lobular and ductal types. Cancer Epidemiol Biomarkers Prev 6:297-301, 1997.

104. Aroner SA, Collins LC, Schnitt SJ, et al: Columnar cell lesions and subsequent breast cancer risk: a nested case-control study. Breast Cancer Res, 2010.

105. Carley AM, Chivukula M, Carter GJ, et al: Frequency and clinical significance of simultaneous association of lobular neoplasia and columnar cell alterations in breast tissue specimens. Am J Clin Pathol 130:254-258, 2008.

106. Guray M, Sahin AA: Benign breast diseases: classification, diagnosis, and management. Oncologist 11:435-449, 2006.

107. Dabbs DJ, Carter G, Fudge M, et al: Molecular alterations in columnar cell lesions of the breast. Mod Pathol 19:344-349, 2006.

108. Field LA, Love B, Deyarmin B, et al: Identification of differentially expressed genes in breast tumors from African American compared with Caucasian women. Cancer 118:1334-1344, 2012.

109. Greer EL, Shi Y: Histone methylation: a dynamic mark in health, disease and inheritance. Nat Rev Genet 13:343-357, 2012.

110. Varier RA, Timmers HT: Histone lysine methylation and demethylation pathways in cancer. Biochim Biophys Acta 1815:75-89, 2011.

111. Kampranis SC, Tsichlis PN: Histone demethylases and cancer. Adv Cancer Res 102:103-169, 2009.

112. Kooistra SM, Helin K: Molecular mechanisms and potential functions of histone demethylases. Nat Rev Mol Cell Biol 13:297-311, 2012.

113. Black JC, Manning AL, Van Rechem C, et al: KDM4A lysine demethylase induces site-specific copy gain and rereplication of regions amplified in tumors. Cell 154:541-555, 2013.

114. Young LC, Hendzel MJ: The oncogenic potential of Jumonji D2 (JMJD2/KDM4) histone demethylase overexpression. Biochem Cell Biol 91:369-377, 2013.

115. Young LC, McDonald DW, Hendzel MJ: Kdm4b histone demethylase is a DNA damage response protein and confers a survival advantage following gamma-irradiation. J Biol Chem 21:376-388, 2013.

116. Chen T, Dent SY: Chromatin modifiers and remodellers: regulators of cellular differentiation. Nat Rev Genet 15:93-106, 2014.

117. Jenuwein T, Allis CD: Translating the histone code. Science 293:1074-1080, 2001.

118. Kouzarides T: Chromatin modifications and their function. Cell 128:693-705, 2007.

119. Li B, Carey M, Workman JL: The role of chromatin during transcription. Cell 128:707-719, 2007.

120. Rando OJ: Combinatorial complexity in chromatin structure and function: revisiting the histone code. Curr Opin Genet Dev 22:148-155, 2012.

121. Jack AP, Bussemer S, Hahn M, et al: H3K56me3 is a novel, conserved heterochromatic mark that largely but not completely overlaps with H3K9me3 in both regulation and localization. PLoS One, 2013.

122. Tu S, Bulloch EM, Yang L, et al: Identification of histone demethylases in Saccharomyces cerevisiae. J Biol Chem 282:62-71, 2007.

123. Klose RJ, Kallin EM, Zhang Y: JmjC-domain-containing proteins and histone demethylation. Nat Rev Genet 7:715-727, 2006.

124. Lee MG, Wynder C, Cooch N, et al: An essential role for CoREST in nucleosomal histone 3 lysine 4 demethylation. Nature 437:432-435, 2005.

125. Metzger E, Wissmann M, Yin N, et al: LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. Nature 437:436-439, 2005.

126. Shi Y, Lan F, Matson C, et al: Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell 119:941-953, 2004.

127. Johansson C, Tumber A, Che K, et al: The roles of Jumonji-type oxygenases in human disease. Epigenomics 6:89-120, 2014.

128. Ruthenburg AJ, Allis CD, Wysocka J: Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark. Mol Cell 25:15-30, 2007.

129. Martin C, Zhang Y: The diverse functions of histone lysine methylation. Nat Rev Mol Cell Biol 6:838-849, 2005.

130. Black JC, Van Rechem C, Whetstine JR: Histone lysine methylation dynamics: establishment, regulation, and biological impact. Mol Cell 48:491-507, 2012.

131. Liu G, Bollig-Fischer A, Kreike B, et al: Genomic amplification and oncogenic properties of the GASC1 histone demethylase gene in breast cancer. Oncogene 28:491-500, 2009.

132. Northcott PA, Nakahara Y, Wu X, et al: Multiple recurrent genetic events converge on control of histone lysine methylation in medulloblastoma. Nat Genet 41:465-472, 2009.

133. Cloos PA, Christensen J, Agger K, et al: The putative oncogene GASC1 demethylates tri- and dimethylated lysine 9 on histone H3. Nature 442:307-311, 2006.

134. Italiano A, Attias R, Aurias A, et al: Molecular cytogenetic characterization of a metastatic lung sarcomatoid carcinoma: 9p23 neocentromere and 9p23-p24 amplification including JAK2 and JMJD2C. Cancer Genet Cytogenet 167:122-130, 2006.

135. Sun LL, Holowatyj A, Xu XE, et al: Histone demethylase GASC1, a potential prognostic and predictive marker in esophageal squamous cell carcinoma. Am J Cancer Res 3:509-517, 2013.

136. Helias C, Struski S, Gervais C, et al: Polycythemia vera transforming to acute myeloid leukemia and complex abnormalities including 9p homogeneously staining region with amplification of MLLT3, JMJD2C, JAK2, and SMARCA2. Cancer Genet Cytogenet 180:51-55, 2008.

137. Ehrbrecht A, Muller U, Wolter M, et al: Comprehensive genomic analysis of desmoplastic medulloblastomas: identification of novel amplified genes and separate evaluation of the different histological components. J Pathol 208:554-563, 2006.

138. Yang ZQ, Imoto I, Fukuda Y, et al: Identification of a novel gene, GASC1, within an amplicon at 9p23-24 frequently detected in esophageal cancer cell lines. Cancer Res 60:735-739, 2000.

139. Vinatzer U, Gollinger M, Mullauer L, et al: Mucosa-associated lymphoid tissue lymphoma: novel translocations including rearrangements of ODZ2, JMJD2C, and CNN3. Clin Cancer Res 14:426-431, 2008.

140. Coffey K, Rogerson L, Ryan-Munden C, et al: The lysine demethylase, KDM4B, is a key molecule in androgen receptor signalling and turnover. Nucleic Acids Res 41:433-446, 2013.

141. Katoh Y, Katoh M: Comparative integromics on JMJD2A, JMJD2B and JMJD2C: preferential expression of JMJD2C in undifferentiated ES cells. Int J Mol Med 20:269-273, 2007.

142. Hillringhaus L, Yue WW, Rose NR, et al: Structural and evolutionary basis for the dual substrate selectivity of human KDM4 histone demethylase family. J Biol Chem 41:616-625, 2011.

143. Whetstine JR, Nottke A, Lan F, et al: Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. Cell 125:467-481, 2006.

144. Berry WL, Janknecht R: KDM4/JMJD2 histone demethylases: epigenetic regulators in cancer cells. Cancer Res 73:2936-2942, 2013.

145. Pedersen MT, Agger K, Laugesen A, et al: The demethylase JMJD2C localizes to H3K4me3-positive transcription start sites and is dispensable for embryonic development. Mol Cell Biol 34:1031-1045, 2014.

146. Iwamori N, Zhao M, Meistrich ML, et al: The testis-enriched histone demethylase, KDM4D, regulates methylation of histone H3 lysine 9 during spermatogenesis in the mouse but is dispensable for fertility. Biol Reprod 84:1225-1234, 2011.

147. Zhang QJ, Chen HZ, Wang L, et al: The histone trimethyllysine demethylase JMJD2A promotes cardiac hypertrophy in response to hypertrophic stimuli in mice. J Clin Invest 121:2447-2456, 2011.

148. Cerami E, Gao J, Dogrusoz U, et al: The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov 2:401-404, 2012.

149. Gao J, Aksoy BA, Dogrusoz U, et al: Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal 2013.

150. Zack TI, Schumacher SE, Carter SL, et al: Pan-cancer patterns of somatic copy number alteration. Nat Genet 45:1134-1140, 2013.

151. Hanahan D, Weinberg RA: Hallmarks of cancer: the next generation. Cell 144:646-674, 2011.

152. Griffith KA, Chung SY, Zhu S, et al: Insulin Resistance and Inflammation in Black Women with and without Breast Cancer: Cause for Concern. Ethn Dis 26:513-520, 2016.

153. Belkina AC, Denis GV: BET domain co-regulators in obesity, inflammation and cancer. Nat Rev Cancer 12:465-477, 2012.

154. Nicholas DA, Andrieu G, Strissel KJ, et al: BET bromodomain proteins and epigenetic regulation of inflammation: implications for type 2 diabetes and breast cancer. Cell Mol Life Sci, 2016.

155. Perrigue PM, Silva ME, Warden CD, et al: The histone demethylase jumonji coordinates cellular senescence including secretion of neural stem cell-attracting cytokines. Mol Cancer Res 13:636-650, 2015.

156. Liu H, Liu L, Holowatyj A, et al: Integrated genomic and functional analyses of histone demethylases identify oncogenic KDM2A isoform in breast cancer. Mol Carcinog 55:977-990, 2016.

157. Das PP, Shao Z, Beyaz S, et al: Distinct and combinatorial functions of Jmjd2b/Kdm4b and Jmjd2c/Kdm4c in mouse embryonic stem cell identity. Mol Cell 53:32-48, 2014.

158. Ipenberg I, Guttmann-Raviv N, Khoury HP, et al: Heat shock protein 90 (Hsp90) selectively regulates the stability of KDM4B/JMJD2B histone demethylase. J Biol Chem 14:681-687, 2013.

159. Kupershmit I, Khoury-Haddad H, Awwad SW, et al: KDM4C (GASC1) lysine demethylase is associated with mitotic chromatin and regulates chromosome segregation during mitosis. Nucleic Acids Res 6:168-182, 2014.

160. Mallette FA, Mattiroli F, Cui G, et al: RNF8- and RNF168-dependent degradation of KDM4A/JMJD2A triggers 53BP1 recruitment to DNA damage sites. EMBO J 31:1865-1878, 2012.

161. Tan MK, Lim HJ, Harper JW: SCF(FBXO22) regulates histone H3 lysine 9 and 36 methylation levels by targeting histone demethylase KDM4A for ubiquitin-mediated proteasomal degradation. Mol Cell Biol 36:87-99, 2011.

162. Van Rechem C, Black JC, Abbas T, et al: The SKP1-Cul1-F-box and leucine-rich repeat protein 4 (SCF-FbxL4) ubiquitin ligase regulates lysine demethylase 4A (KDM4A)/Jumonji domain-containing 2A (JMJD2A) protein. J Biol Chem 30:462-470, 2011.

163. Kim TD, Fuchs JR, Schwartz E, et al: Pro-growth role of the JMJD2C histone demethylase in HCT-116 colon cancer cells and identification of curcuminoids as JMJD2 inhibitors. Am J Transl Res 6:236-247, 2014.

164. Burton A, Azevedo C, Andreassi C, et al: Inositol pyrophosphates regulate JMJD2C-dependent histone demethylation. Proc Natl Acad Sci U S A 189:70-75, 2013.

165. Dawson MA, Kouzarides T: Cancer epigenetics: from mechanism to therapy. Cell 150:12-27, 2012.

166. Rui L, Emre NC, Kruhlak MJ, et al: Cooperative epigenetic modulation by cancer amplicon genes. Cancer Cell 18:590-605, 2010.

167. You JS, Jones PA: Cancer genetics and epigenetics: two sides of the same coin? Cancer Cell 22:9-20, 2012.

168. Santen RJ, Brodie H, Simpson ER, et al: History of aromatase: saga of an important biological mediator and therapeutic target. Endocr Rev 30:343-375, 2009.

169. Simpson ER, Zhao Y, Agarwal VR, et al: Aromatase expression in health and disease. Recent Prog Horm Res 52:185-213; discussion 213-214, 1997.

170. Endoh H, Maruyama K, Masuhiro Y, et al: Purification and identification of p68 RNA helicase acting as a transcriptional coactivator specific for the activation function 1 of human estrogen receptor alpha. Mol Cell Biol 19:5363-5372, 1999.

171. Ghosh S, Thakur MK: Interaction of estrogen receptor-alpha transactivation domain with nuclear proteins of mouse brain: p68 RNA helicase shows age- and sex-specific change. J Neurosci Res 87:1323-1328, 2009.

172. Penault-Llorca F, Bertucci F, Adelaide J, et al: Expression of FGF and FGF receptor genes in human breast cancer. Int J Cancer 61:170-176, 1995.

173. Tomlinson DC, Knowles MA, Speirs V: Mechanisms of FGFR3 actions in endocrine resistant breast cancer. Int J Cancer 130:2857-2866, 2012.

174. Herrera-Abreu MT, Pearson A, Campbell J, et al: Parallel RNA interference screens identify EGFR activation as an escape mechanism in FGFR3-mutant cancer. Cancer Discov 3:1058-1071, 2013.

175. Morris PG, Hudis CA, Giri D, et al: Inflammation and increased aromatase expression occur in the breast tissue of obese women with breast cancer. Cancer Prev Res (Phila) 4:1021-1029, 2011.

176. Cleary MP, Grossmann ME: Minireview: Obesity and breast cancer: the estrogen connection. Endocrinology 150:2537-2542, 2009.

177. Pierce BL, Ballard-Barbash R, Bernstein L, et al: Elevated biomarkers of inflammation are associated with reduced survival among breast cancer patients. J Clin Oncol 27:3437-3444, 2009.

178. van Kruijsdijk RC, van der Wall E, Visseren FL: Obesity and cancer: the role of dysfunctional adipose tissue. Cancer Epidemiol Biomarkers Prev 18:2569-2578, 2009.

179. Dandona P, Weinstock R, Thusu K, et al: Tumor necrosis factor-alpha in sera of obese patients: fall with weight loss. J Clin Endocrinol Metab 83:2907-2910, 1998.

180. Calle EE, Rodriguez C, Walker-Thurmond K, et al: Overweight, obesity, and mortality from cancer in a prospectively studied cohort of U.S. adults. N Engl J Med 348:1625-1638, 2003.

181. Maehle BO, Tretli S, Skjaerven R, et al: Premorbid body weight and its relations to primary tumour diameter in breast cancer patients; its dependence on estrogen and progesteron receptor status. Breast Cancer Res Treat 68:159-169, 2001.

182. Ogden CL, Carroll MD, Kit BK, et al: Prevalence of obesity among adults: United States, 2011-2012. NCHS Data Brief:1-8, 2013.

183. An R: Prevalence and Trends of Adult Obesity in the US, 1999-2012. ISRN Obes 2014.

184. Dahlman-Wright K, Qiao Y, Jonsson P, et al: Interplay between AP-1 and estrogen receptor alpha in regulating gene expression and proliferation networks in breast cancer cells. Carcinogenesis 33:1684-1691, 2012.

185. Stender JD, Frasor J, Komm B, et al: Estrogen-regulated gene networks in human breast cancer cells: involvement of E2F1 in the regulation of cell proliferation. Mol Endocrinol 21:2112-2123, 2007.

186. Barker HE, Smyth GK, Wettenhall J, et al: Deaf-1 regulates epithelial cell proliferation and side-branching in the mammary gland. BMC Dev Biol, 2008.

187. Cubeddu L, Joseph S, Richard DJ, et al: Contribution of DEAF1 structural domains to the interaction with the breast cancer oncogene LMO4. PLoS One, 2012.

188. Mitro SD, Ali-Fehmi R, Bandyopadhyay S, et al: Clinical characteristics of breast cancers in African-American women with benign breast disease: a comparison to the surveillance, epidemiology, and end results program. Breast J 20:571-577, 2014.

189. Callari M, Lembo A, Bianchini G, et al: Accurate data processing improves the reliability of Affymetrix gene expression profiles from FFPE samples. PLoS One, 2014.

190. R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

191. https://cran.r-project.org/package=cmprsk.

192. Carvalho BS, Irizarry RA: A framework for oligonucleotide microarray preprocessing. Bioinformatics 26:2363-2367, 2010.

193. Forozan F, Veldman R, Ammerman CA, et al: Molecular cytogenetic analysis of 11 new breast cancer cell lines. Br J Cancer 81:1328-1334, 1999.

194. Yang ZQ, Streicher KL, Ray ME, et al: Multiple interacting oncogenes on the 8p11-p12 amplicon in human breast cancer. Cancer Res 66:11632-11643, 2006.

195. Liu L, Kimball S, Liu H, et al: Genetic alterations of histone lysine methyltransferases and their significance in breast cancer. Oncotarget 6:2466-2482, 2015.

196. Kapp JM, Walker R, Haneuse S, et al: Are there racial/ethnic disparities among women younger than 40 undergoing mammography? Breast Cancer Res Treat 124:213-222, 2010.

197. Gail MH, Costantino JP, Pee D, et al: Projecting individualized absolute invasive breast cancer risk in African American women. J Natl Cancer Inst 99:1782-1792, 2007.

198. Suikki HE, Kujala PM, Tammela TL, et al: Genetic alterations and changes in expression of histone demethylases in prostate cancer. Prostate 70:889-898, 2010.

199. Luo W, Chang R, Zhong J, et al: Histone demethylase JMJD2C is a coactivator for hypoxia-inducible factor 1 that is required for breast cancer progression. Proc Natl Acad Sci U S A 109:3367-3376, 2012.

# ABSTRACT

## CLINICOPATHOLOGY AND MOLECULAR DETERMINANTS UNDERLYING BENIGN BREAST AND BREAST CANCER LESIONS

by

**ANDREANA NATALIE HOLOWATYJ**

**May 2017**

**Advisor:** Dr. Michele L. Cote, Ph.D., MPH

**Major:** Cancer Biology

**Degree:** Doctor of Philosophy

Despite converging incidence rates for breast cancers by race, disparities in mortality persist where black women suffer from poorer prognosis compared to white counterparts. To understand the clinical, demographic, and molecular characteristics underlying these disparities, we examined differences among patients with breast cancer to understand the role of human epidermal growth factor receptor 2 (HER2) status, age, and race/ethnicity among women diagnosed with hormone receptor-positive breast cancer, and disparities in surgical therapy among female patients with early stage young-onset breast cancer. Benign breast disease, another known risk factor for breast cancer, includes a histological spectrum of lesions, could contribute to disparities in survival among both black and white women is benign breast disease as little is known about benign breast tissue from black women. To better characterize the risk of breast cancer among black women with benign breast disease, we profiled the clinicopathology and molecular characteristics of benign breast lesions among black

women who subsequently developed breast cancer. Using the metropolitan Detroit benign breast disease cohort, we identified black women with benign breast lesions subsequently diagnosed with breast cancer. Gene expression profiling of benign breast disease tissue and Cox proportional hazards modeling were used to evaluate transcriptional variations associated with time to breast cancer. 1,802 genes were significantly associated with a shorter time to breast cancer diagnosis. The greatest transcriptional variation associated with time to diagnosis was histone lysine demethylase 4C (KDM4C), a histone-modifying enzyme. Given that information regarding the genomic and transcriptomic alterations of KDM4C and the KDM4 subfamily in different subtypes of breast cancer remains largely incomplete, we conducted a meta-analysis of KDM4A-D in breast cancer and identified associations among recurrent copy number alterations, gene expression, and breast cancer subtypes. We demonstrated that KDM4C amplification and overexpression was observed in aggressive, basal-like breast cancer tissues. Inflammatory disease ranked as the top disorder, and nearly a fifth of the genes classified in the estrogen biosynthesis pathway were significantly associated with time to breast cancer. Taken together, our findings identify molecular precursors for time to cancer among black women with benign breast disease. These results identify increased expression of critical molecular determinants, including KDM4C, and pathways that are significantly associated with shorter time to subsequent breast cancer development among black women. Further studies to better characterize the molecular profiles of benign breast and breast cancer lesions can lead to development of molecular classifiers for breast cancer risk and diagnosis among black women with benign breast lesions.

# AUTOBIOGRAPHICAL STATEMENT

Andreana Natalie Holowatyj, daughter of the late Stephen and Irene Holowatyj, was born and raised in the suburbs of Chicago, Illinois. Holowatyj remained in her home state of Illinois to pursue her undergraduate studies at Benedictine University in Lisle, where she earned a Bachelor of Arts degree with Honors in Medical Humanities and a minor in Mathematics in May 2013. During her undergraduate tenure, she not only completed certification as an Emergency Medical Technician-Basic (EMT-B), but also spent semesters overseas in the Czech Republic, New Zealand, and China.

Upon completion of her undergraduate studies, Holowatyj matriculated into the Cancer Biology Ph.D. program in the Department of Oncology at the Wayne State University School of Medicine and Karmanos Cancer Institute in August 2013. During her doctoral career, Holowatyj has continued her passions for global health and oncology by earning a total of thirty-two national and international awards to present her scientific work and findings to date, and has traveled to twenty-two countries around the world in this time. Notably, in February 2016 and 2017 Holowatyj was invited onto the AACR Congressional Briefing panel to discuss her oncology research with Congressmen on Capitol Hill. She has authored several research manuscripts and review articles to date, with additional manuscripts in submission, of which her recent first-author work was published in the *Journal of Clinical Oncology.* Holowatyj pursued her desire for teaching by also becoming a part-time instructor in the Department of Biological Sciences, beginning in January 2016 to date. After earning the Komen for the Cure Graduate Training in Disparities Research Fellowship, Holowatyj added a Graduate Certificate in Public Health Practice (GCPHP) degree from the Department of Family Medicine & Public Health Sciences to better understand the public health significance of her translational and molecular oncology research, and further her investigative research studies to assess the cancer burden among homeless persons; a degree which conferred in December 2016. She has also served on various institutional and national committees, including as a member on the American Association for Cancer Research AMC-led Fundraising Committee and as an ambassador for the European Association for Cancer Research. Upon completion of her Ph.D. in Cancer Biology, Holowatyj hopes to continue teaching and researching in the field of oncology.