TRIPLE NEGATIVE BREAST CANCER AND AFRICAN ANCESTRY

IDENTIFYING EPIDEMIOLOGICAL AND GENETIC FACTORS UNDERLYING THE DISPARITY IN INCIDENCE AND OUTCOMES OF TRIPLE NEGATIVE BREAST CANCERS (TNBC) IN WOMEN OF AFRICAN ANCESTRY (WAA)

By: SHAWN MATTHEW HERCULES, B.Sc., MPH

A Thesis Submitted to the School of Graduate Studies

in Partial Fulfillment of the

Requirements for the Degree Doctor of Philosophy

DOCTOR OF PHILOSOPHY (2021) McMaster University

(Biology) Hamilton, Ontario

TITLE: **Identifying Epidemiological and Genetic Factors Underlying the Disparity in Incidence and Outcomes of Triple Negative Breast Cancers (TNBC) in Women of African Ancestry (WAA)**

AUTHOR: Shawn Matthew Hercules, B.Sc., MPH (University of the West Indies)

SUPERVISOR: Dr. Juliet M. Daniel

NUMBER OF PAGES: xviii, 215

LAY ABSTRACT

Breast cancer (BCa) is a leading cause of cancer-related death in women worldwide. Although Caucasian women are diagnosed with BCa more than women of African ancestry (WAA), more WAA unfortunately die from BCa. The reasons for this disparity are currently unknown, however, a higher proportion of WAA are diagnosed with an aggressive type of BCa called triple negative breast cancer (TNBC). This might partially explain the high cancer death rate in WAA. To understand this disparity in BCa incidence and outcomes, we investigated TNBC disease trends across the African continent and in Barbados (a Caribbean island with predominantly African ancestry) and found a high proportion of TNBC diagnoses in Barbados and West African countries. We also discovered a novel genetic profile within these groups that may be useful to develop new cancer therapies that would decrease TNBC aggressiveness and death in these populations.

ABSTRACT

Breast cancer (BCa) is a leading cause of cancer-related female deaths worldwide and is a complex disease consisting of many different subtypes with varying clinical course and outcomes. Triple negative breast cancer (TNBC), an aggressive and highly metastatic subtype, is most prevalent in women of African ancestry (WAA) but the causes of this disparity are not fully understood. The goal of this study was to investigate the epidemiological and genetic profiles in ancestrally-related WAA in Barbados and Nigeria to advance knowledge and lay the foundation for development of improved or novel BCa therapeutics.

To gain insight about TNBC across the African continent, a systematic review and meta-analysis was conducted. TNBC frequencies on average across Africa were estimated at 26.8% but were highest in West African countries (46.0%). We also sought to identify the epidemiological profile of BCa in Barbados—a Caribbean island with significant West African ancestry. We reviewed pathological reports for BCa from the sole public hospital in Barbados and compared those data with USA population-based data. We found a high prevalence of high prevalence of TNBC amongst women diagnosed with breast cancer in Barbados (25%), compared to 21% in non-Hispanic Black and 10% in non-Hispanic White women in the USA for the 2010-2016 period.

We also investigated the somatic mutational profile of WAA with TNBC in Barbados and Nigeria using whole exome sequencing (WES) of formalin-fixed paraffin-embedded TNBC tissues. This investigation revealed novel and pathogenic variants in well-known cancer-associated genes such as *TP53, BRCA1* and *MDC1*. The somatic

mutation signature in Nigerian tissues correlated with aflatoxin signature, implying a role for environmental factors influencing the genomics profile in this cohort. Copy number variants were revealed at high frequencies for *PIK3CA, FGFR2* and *HIF1AN* genes. Collectively, these findings uncovered novel epidemiological and genetic trends in WAA with high prevalence of the aggressive TNBC subtype.

ACKNOWLEDGEMENTS

I am overwhelmed with joy right now as I write this, knowing that I could not have reached this milestone in my personal journey so far (academic and otherwise), without the names mentioned here and so many more! I wish I could thank each and every person that helped me along this journey in the acknowledgements section and I could perhaps write an entire thesis doing so! Though some names may be missed, I am thankful and am elated to say thank you. There is SO MUCH I could say about each of you, but I will try to keep it short here!

To my parents, Mark and Leslyn Hercules and my aunt Verna Edwards—THANK YOU! From day one, you have all supported me and encouraged me to be my best. I witnessed each of you work extremely hard to ensure that I had all the tools needed for academic success, from my preschool days right up until this moment. Thank you for proof-reading my assignments in secondary school and for putting up with me arguing with you about how to correctly do math equations, lol! Thank you for always cheering me on and most of all believing in me, in times when I was doubtful of my capabilities. One of the scariest factors about leaving Barbados to pursue my Ph.D. at McMaster was the thought of leaving you and missing out on special moments with you all but I am so grateful to God for technology, your "good morning/good night" messages in our Whatsapp group and our weekly FaceTimes that made us even stronger than I would have imagined. And of course, my brother Jevon that always kept us laughing and in light spirits on our family FT calls. You are an inspiration Jevon, to leave Barbados TWICE in pursuit of your education. I am immensely proud of you and that I get to call you brother! Congrats on defending your Master's just a few days ago! Also, a huge thanks for you assisting me with collecting data for my project—you worked at the QEH for almost a year collecting tissues and clinical data and I am forever grateful to you for doing that.

To my (adopted) grand-parents and family, Elaine and Leroy McCarthy and the entire McCarthy family—you did not hesitate to accept me into your warm family as though I was one of the "grandkids". You all ensured that I had a family away from home and I thank you for that. Thank you for all the holiday dinners, summer BBQs and afternoon lunches after church and for being hospitable to my family whenever they came to visit.

Thank you to my first ever friend at McMaster, Dr. Sarah Marzec! I randomly ran into you outside of Barb's office and we connected since then. Thanks for being a great friend that I could be 100% myself around. Thanks for being my party-goer buddy for our first few grad school parties, lol. Also, thanks to Dr. Caroline Cauret, Dr. Andrew "Dede" Tupper and Chris Hughes for being supportive friends, especially at the beginning of grad school and for "family" Sunday dinners and "orphan" Thanksgiving dinners, lol. Thank you, Dr. Daniella Lato, for being one of the most supportive and sweetest friends ever! I sincerely miss our "mental health" walks in Coote's Paradise when we just decided to take a break

from the stress of grad school. Thanks for always accepting me just as I am and for making me feel "human" in the midst of grad school craziness.

To Dr. Shaiya Robinson—my SISTER! Sis, thank you, thank you thank you! This was one of the most unexpected lab friendships to have developed! Started off as lab spouses, lol, to now siblings. Our friendship is so tight knit now that I honestly feel as though you are a biological sister to me. Thank you for inviting me to your family's home for my first Christmas away from home and for always opening your home to me whenever I needed to stay over in Toronto for a conference or just for fun! Thank you sis! Especially for encouraging me to be comfortable with ALL aspects of myself and not just "Shawnie Poo the scientist". Thank you for also pushing me constantly, to be my BEST and not settle for less. You continue to be an inspiration to me, ASSISTANT PROFESSOR DOCTOR Shai Bear. Love you! <3

To Dr. Blessing Bassey-Archibong, THANK YOU! I remember the first day visiting the lab, you sat me down and talked about Kaiso and tons of bioinformatics tools for almost an hour—I had no idea what you were talking about at the time (LOL), but I could see that you were extremely passionate about what you were explaining. Thank you for establishing all the collaborations in Nigeria—my Ph.D. project as well as this incredible research that we do—would not have been possible without this, so thank you! Thank you for being the best travel companion to Nigeria. We did lots of "waka waka" by hopping on a different plane or long drive at least twice every week for the short time we were there and your drive and passion for this research further shone through in our time in Nigeria. Thank you for always being extremely honest with me about my work and always being there as an academic support, even after you've left the lab. Thank you for editing my thesis on such short notice!

To the "best friend that I live with" (lol), Lindyann Lessey… these past two years of knowing you turned out exceptionally better than we both thought! We are extreme polar opposites, yet we still get along really well and surprisingly don't hate each other lol. Thank you for persisting about living together even though I thought it was a bad idea—I cannot imagine better housemates (thanks also Shane!) to have gone through multiple lockdowns during a pandemic with. The first few weeks were pretty rough for me, as you know, but you ensured that I got through those initial rough stages and into a more positive mindset. Thank you for always showing initiative in the lab, helping keep the lab clean and for helping me with my experiments when needed. You have so much potential, and I am so excited about your future! Remember to pause, listen and reflect as this would greatly help you to get through grad school with less stress! You can do it! I believe in you! Love and appreciate you!

To Shane, thanks for always being so chill and for helping cheer me up with "you can do it/you got this" if you heard me deeply sigh or make a disgruntled noise while working… especially over the past year! Thanks for your positive attitude and for never telling me that

I “shouldn’t order that KFC/Quikchik”—you’re a real one lol. Looking forward to seeing your success in the Gupta lab, and beyond.

To all of the undergrads that I have crossed paths with in the Daniel lab—thank you! I will try to keep this section in order of when I met you. Will also keep it brief and only note the first few things that come to mind about each of you. But again, thank you all (and the ones not mentioned) for helping me on this journey in one way or another. To Sinthu for having deep intellectual/spiritual conversations about God and religion, spontaneously going with me to get my nose pierced and our random trips to Niagara. To Yulika for being an awesome help in the lab with my experiments and for pushing me to go to the gym more haha. The Naija girls and soon to be MDs, Noroh and Toyemi, thank you for all of your help as well and for being awesome undergrads. To the systematic review team… first up, to Stefan (!!!) my first real “official” undergrad mentee… I learned so much about the mentorship process through meeting you. Thank you for always being so inquisitive about our experiments and excited about science. Thank you for helping me realise my mentorship style. Thank you also to Chen for your insight with the systematic review and getting the forest plots started! Thank you to Meena also for all of your help throughout your various Daniel lab projects but especially the last few months of my Ph.D. and being extremely helpful with whatever “small task” I’ve had over the last few months. THANK YOU! Allysia… girl… sis… LOL. I don’t even know where to start. Thanks for helping me realise that I can design an experiment that I’ve never done (ayyy, that dot-blot EMSA) and have it turn out really well! Thanks for your assistance with the TMA scoring and with my IHC optimisations. Thanks for being a friend and sister during your time in the Daniel lab and even afterwards. You are one of my closest friends, and I thank you for your contributions along the journey.

Thanks to the quarandrag crew (Jevon, Sam, Lindyann, Chris, Topher, Naomi, Kali, Émile, Lily, Allysia, Dan, Evelyn and anyone else that came) for watching Drag Race virtually on Zoom during the lockdowns! Wow, those Friday nights really helped my mental health and helped me escape from the “doom and gloom” of the pandemic. Also, thanks for your constructive criticism about my lewks hahaha! Also thank EACH of you individually for your friendship which definitely grew stronger over the last year with our Zoom sessions. Special thanks also to my friends from Barbados, literal DAY ONES, namely Russell, Topher, Dominique, Katrina, Lorimer & Carla. Also, thanks to Catherine and Miriam my CDRC work buddies! You all were there from before the journey started and stayed throughout, if I needed to vent, laugh, process etc. Thank you for the gift of friendship and for always being there. Love you all.

To my love, Dr. Calum Neish—you’ve only been in my life for a short time but left an indelible impact so far. Thank you for motivating me to keep going, even when I at times had no energy left or simply wanted to throw in the towel. Thanks for helping to pull me out of my Ph.D. slump a few weeks ago and for holding me whenever I needed to be held

and for being a listening ear and a strong support throughout. Thank you for all of the delicious food and precious moments together while I was stressed about finishing the Ph.D. Thank you for helping me see the light at the end of the Ph.D. tunnel! I'll save the acknowledgement section from all the mushy stuff, but I just had to say thank you! I am so proud to call you my life partner. Love you.

My supervisory committee has been a huge part of my journey and I am extremely grateful to each of you. To Dr. Andrew McArthur—thank you for believing in me and helping me to also believe in myself! Your enthusiasm for this project and in my abilities to do it, greatly encouraged me to keep going, so thank you! Dr. André Bédard—thank you for always bringing me back to the molecular aspect of my project and for keeping me sharp with questions about signalling pathways and for your discussions about what my project was really testing for my hypotheses, thank you! To Dr. Greg Pond—thank you for helping me see the large amount of data I had! I did not even realise at the time, that the epidemiological data I had could be transformed into a manuscript, so thank you for helping me see that and for your continued statistical advice. To Dr. Zarko Manojlovic—you were not a part of my supervisory committee, but you greatly helped me with understanding the bioinformatics aspect of my project. Thanks for our many calls, emails, and Slack messages for helping that project come together. To Dr. Kim Blenman—our paths only connected over the past year, and I am extremely grateful for that. Thank you for always being critical and for making me see aspects about the manuscripts that I hadn't seen before. Thank you for always being encouraging and for helping me put my best foot forward with this very important work.

To my supervisor, Dr. Juliet Daniel… so much to thank you for! From the day we met in Barbados until now, you have been an inspiration to me. You see the "best" version of Shawn that I don't always see, and you keep pushing me to get there (as hesitatingly as I receive this sometimes, it is very much appreciated and needed). Thank you for taking me from Barbados and training me into the person I am today, not only academically but also as a well-rounded individual. Training in your lab has competitively equipped me for the future. Thank you for always being patient with me and encouraging me to think deeper about my work and for helping me to always see the bigger picture. Thank you for reading my thesis while you were on vacation! Thank you for playing such a critical role in my journey to self-discovery. There is obviously lots more that I can say but I will save that for the defence party! LOL!

I also cannot say thank you enough to all the patients and relatives of the patients that consented into this study. Without you, this project would not be possible. Thank you to all of the various pathology and surgical departments at our various locations for your participation in this project. The work that each of you have put into this has and will continue to pay off in the future for helping women of African ancestry beat triple negative breast cancer!

LIST OF FIGURES

CHAPTER 1

CHAPTER 2

N/A

CHAPTER 3

CHAPTER 4

CHAPTER 5

CHAPTER 6

CHAPTER 7

APPENDIX

LIST OF TABLES

CHAPTER 1

N/A

CHAPTER 2

N/A

CHAPTER 3

LIST OF ABBREVIATIONS

AA	**A**frican **A**merican
AFB1	**A**flatoxin **B1**
ATCC	**A**merican **T**ype **C**ulture **C**ollection
AJCC	**A**merican **J**oint **C**ommittee on **C**ancer
AR	**A**ndrogen **R**eceptor
BAM	**B**inary **A**lignment **M**ap
BCa	**B**reast **Ca**ncer
BL	**B**asal-**l**ike
BWA	**B**urrows' **W**heeler **A**lignment
CW	**C**aucasian **W**omen
CWAA	**C**aribbean **W**omen of **A**frican **A**ncestry
CA	**C**aucasian **A**merican
COSMIC	**C**atalogue **O**f **S**omatic **M**utations **I**n **C**ancer
DAB	**D**i**a**mino**b**enzidine
DNA	**D**eoxy-ribo**n**ucleic **a**cid
DMEM	**D**ulbecco's **M**odified **E**agle's **M**edium
DRFI	**D**istant **R**ecurrence-**F**ree **I**nterval
EMT	**E**pithelial-to-**m**esenchymal **t**ransition
ER	**E**strogen **R**eceptor
FDA	**F**ood & **D**rug **A**dministration
FISH	**F**luorescence **i**n **s**itu **H**ybridization
FBS	**F**etal **B**ovine **S**erum
FFPE	**F**ormalin-**F**ixed **P**araffin-**E**mbedded
FTB	**F**etch **T**arget **B**uffer
GATK	**G**enome **A**nalysis **T**ool **K**it
GEO	**G**ene **E**xpression **O**mnibus
gDNA	**G**enomic **D**eoxy-ribo**n**ucleic **a**cid

GLOBOCAN	**Glob**al **Can**cer Incidence, Mortality and Prevalence
GR	**G**lucocorticoid **R**eceptor
HCC	**H**epato**c**ellular **C**arcinoma
HDI	**H**uman **D**evelopment **I**ndex
HER2	**H**uman **E**pidermal Growth Factor **R**eceptor-**2**
HR	**H**ormone **R**eceptor
HRR	**H**omologous **R**ecombination **R**epair
IHC	**I**mmuno**h**isto**c**hemistry
IM	**Im**munomodulatory
LAR	**L**uminal **A**ndrogen **R**eceptor
LMIC	**L**ow- and **M**iddle-**I**ncome **C**ountry
LN	**L**ymph **N**ode
lncRNA	**L**ong **n**on-**c**oding **RNA**
miR	**mi**cro-**R**NA
MSL	**M**esenchymal **S**tem **L**ike
NHB	**N**on-**H**ispanic **B**lack
NHW	**N**on-**H**ispanic **W**hite
NW	**N**igerian **W**omen
OS	**O**verall **S**urvival
PARP	**P**oly **A**DP-**r**ibose **P**olymerase
PBS	**P**hosphate-**B**uffered **S**aline
PBS-T	**P**hosphate-**B**uffered **S**aline **T**ween
PD-L1	**P**rogrammed **D**eath-**L**ike **1**
POZ-ZF	**Po**xvirus and **Z**inc **F**inger
PR	**P**rogesterone **R**eceptor
RNA	**R**ibo**n**ucleic **a**cid
RPMI 1640	**R**oswell **P**ark **M**emorial **I**nstitute (**1640**)
SEP	**S**ocio**e**conomic **P**osition

SEER	**S**urveillance of **E**pidemiology and **E**nd **R**esults
SES	**S**socio**e**conomic **S**tatus
SSB	**S**ingle-**S**tranded **B**reaks
SNP	**S**ingle **N**ucleotide **P**olymorphism
TF	**T**ranscription **F**actor
TCGA	**T**he **C**ancer **G**enome **A**tlas
TMA	**T**issue **M**icro**a**rray
TNBC	**T**riple **N**egative **B**reast **C**ancer
UNG	**U**racil-**N**-**G**lycosylase
UNS	**Uns**table
WAA	**W**omen of **A**frican **a**ncestry
WEA	**W**omen of **E**uropean **a**ncestry
WES	**W**hole **e**xome **s**equencing

Table of Contents

CHAPTER 1: INTRODUCTION

1.1 OVERVIEW

Breast **c**ancer (BCa) is a leading cause of death in women worldwide, second only to lung cancer in developed countries (Bray *et al.*, 2018; Siegel *et al.*, 2017). Intriguingly, while BCa incidence for **w**omen of **E**uropean **a**ncestry (WEA) is higher than the incidence for **w**omen of **A**frican **a**ncestry (WAA), mortality rates show the opposite trend (DeSantis *et al.*, 2019; Moller *et al.*, 2016). Emerging evidence suggests that environmental factors such as diet, exercise, socioeconomic status (SES) and BCa screening (e.g., mammography) may play important roles in BCa aetiology and outcomes (Gordon, 2003; Halmin *et al.*, 2008; Linnenbringer *et al.*, 2017). However, other population-based studies that have adjusted for BCa screening rates, SES and other lifestyle factors between WAA and WEA, still found that WAA had a higher mortality rate than WEA (DeSantis *et al.*, 2016; Newman *et al.*, 2006; Wyatt *et al.*, 2017), implying that other factors (*e.g.,* biological factors) are driving this disparity. Notably, racial disparities are not unique to BCa – as other complex pathologies (e.g. prostate cancer, diabetes, cardiovascular disease, hypertension) also have racial disparities attributed to a combination of lifestyle and genetic risk factors (Nead *et al.*, 2015; Reddon *et al.*, 2016). This study sought to investigate the epidemiological and genetic profiles of WAA with TNBC in Nigeria and Barbados to elucidate potential risk factors responsible for TNBC aggressiveness in this group and to determine potential therapeutic and diagnostic measures for WAA with TNBC.

1.2 BREAST CANCER (BCA) EPIDEMIOLOGY AND HISTORY

Cancer is broadly characterised by site-specific uncontrolled cell proliferation and invasion of normal tissues (Hanahan & Weinberg, 2011). It is an increasing public health concern as incidence rates are projected to escalate exponentially over the next decade (Fidler *et al.*, 2018; Torre *et al.*, 2016). Growing global cancer incidence rates could be attributed to longer life expectancy as well as the rise in cancer-associated risk factors such as smoking, diet and other environmental factors (Bray *et al.*, 2018). Higher incidence could also be a result of more surveillance, the establishment of cancer registries, and other public health measures that promote cancer awareness and screening (Marcus & National Cancer Institute (U.S.), 2019). Globally, there has also been a surge in cancer-associated mortality with the top five leading causes of cancer-related deaths worldwide being lung, colorectal, liver, stomach and female breast cancers (Sung *et al.*, 2021). Recent estimates suggest that cancer is the first or second leading cause of premature deaths in 91 out of 172 countries globally with **l**ow- and **m**iddle- **i**ncome **c**ountries (LMICs) contributing a substantial proportion of those 91 countries (Fidler *et al.*, 2018). This however was not always the case — in 1990, cancer was ranked the 7th cause of death in LMICs compared to 2019 estimates suggesting that cancer is now the 2nd cause of death in these countries (IHME, 2020). This shift in cancer mortality in LMICs has surpassed cancer-specific mortality in higher income countries, *e.g.,* the estimated cumulative risk of dying from cancer among women in 2018 was higher in East Africa (11.4%) than the risks estimated in North America (8.6%) (Ferlay *et al.*, 2019; Fidler *et al.*, 2018). Specifically in women, BCa is the leading cause of cancer-related deaths worldwide even in countries within the low/middle **h**uman **d**evelopment

index (HDI) which is a measure of a country's levels of social and economic development (Fidler *et al.*, 2018). The highest mortality rates from female BCa were estimated to be from Melanesia, Western Africa, Micronesia/Polynesia and the Caribbean where Barbados had the highest estimated mortality worldwide (Sung *et al.*, 2021). While the cause(s) for increased mortality among these countries are not easily identifiable, it could be due to the complex interplay of epidemiological, environmental, and biological factors. It is no surprise that SES and environmental factors have been found to be associated with a higher mortality rate for BCa, as these factors play critical roles in cancer onset and treatment. It has recently been found that women in the USA with lower **s**ocio**e**conomic **s**tatus (SES) are less likely to have post-breast conservation surgery radiation, adjuvant chemotherapy and axillary surgery (all of which can significantly impact survival rates) compared to women with higher SES (Dreyer *et al.*, 2018). A similar trend was also seen across Europe, where SES showed a positive correlation with BCa incidence but had an inverse relationship with BCa mortality (Lundqvist *et al.*, 2016). These associations lay the foundation for explaining mortality rate but cannot fully explain the disparities observed in BCa aggressiveness and therefore biological factors implicated in the classification of BCa should also be considered and examined in greater detail.

1.3 CLINICAL AND MOLECULAR CLASSIFICATION OF BREAST CANCER AND TREATMENT OPTIONS

Breast cancer and its aetiology have been extensively studied and recorded for centuries in ancient civilizations. Dating back to ~ 2400 B.C., BCa was described in the Edwin Smith Papyrus ancient Egyptian text as a "bulging tumor on the breast with no treatment"

(Breasted, 1930). Until the 19th century, various treatment options involving salts and Egyptian ointments were used to manage BCa (Hajdu, 2011). Presently, BCa management techniques have evolved from these methods and the first documented introduction of mastectomy in the 1880s (Zurrida & Veronesi, 2015) to include more advanced surgical methods, chemotherapy, radiation therapy, and most recently targeted therapy.

At the onset of BCa, normal tissues accumulate genomic abnormalities such as mutations and progress to ductal hyperplasia then **d**uctal **c**arcinoma ***i**n **s**itu* (DCIS) which can then develop to **i**nvasive **d**uctal **c**arcinoma (IDC) (Shackney & Silverman, 2003). This progression can take place over an undefined period and women diagnosed with hyperplastic lesions or DCIS have an increased relative risk of up to 10-fold for an invasive carcinoma to develop (Allred *et al.*, 2001; Dupont & Page, 1989). Additionally, the relative risk for developing an invasive carcinoma increases with age at diagnosis of hyperplastic lesion. However, the progression is highly correlated to the proliferative capacity of the DCIS, largely understood as the grade of the tumour — i.e. high grade DCIS lesions are likely to progress to high grade IDC lesions (Simpson *et al.*, 2005). Dating back to as early as 1928, tumours were morphologically assessed for features such as tubule formation, nuclei size and hyperchromatism as prognostic factors (Patey & Scarff, 1928). Currently, the most commonly used system to assess for grade is the Elston/Nottingham modification of Bloom-Richardson system (Elston & Ellis, 1991) which allows pathologists to determine the proliferative ability of tumour cells and has been found to predict the rate of tumour growth and hence prognosis. Lower grade (grade 1) is associated with low proliferative ability, thus with a better prognosis, whereas high and intermediate tumour grades (grade

2 and 3) are associated with higher proliferative ability and thus worse prognosis (Rakha *et al.*, 2010). In addition to the assessment for grading tumour specimens, pathological and clinical staging is routinely conducted which further adds prognostic and anatomical value.

The most accepted and gold standard **A**merican **J**oint **C**ommittee on **C**ancer (AJCC) staging assessment is based on the size of the tumour (T), lymph node involvement (N) and secondary sites where the tumour could have migrated to (M), henceforth deemed as metastases (Amin *et al.*, 2017). The first edition was published in 1977 on the basis of TNM staging and has since evolved to the eighth and latest edition published in 2017 which also takes other prognostic markers (*e.g.,* histological grade) and molecular markers into consideration for staging in an effort to transition from a population-based approach to a more personalized approach . This level of staging takes into account the expression of three recognized and accepted biomarkers – **E**strogen **R**eceptor (ER), **P**rogesterone **R**eceptor (PR), and **H**uman **E**pidermal Growth Factor **R**eceptor **2** (HER2) – via **i**mmuno**h**isto**c**hemistry (IHC) and *in situ* hybridization where necessary. In addition to these three markers, Ki-67 expression is also assessed as well as the mutation frequency in multigene panel tests (Amin *et al.*, 2017). These panels assess genomic variations for common genetic BCa markers such as *BRCA1, BRCA2* and *TP53* (Sabatier *et al.*, 2014). These three markers have been associated with the aggressive nature of BCa and are thus of clinical importance for cancer staging and appropriate treatment options with a genomics approach. The development and widespread access to cutting-edge tools such as mammographic imaging, IHC, genomic sequencing, and open access genomics data has greatly improved the screening, diagnosis, and treatment of BCa.

1.4 GENOMICS AND PROGNOSTIC VALUE IN BCA

With the decrease in costs associated with high throughput sequencing of the human genome, there has been a sharp uptake in the number of sequencing projects with the goal to offer personalized/precision medicine for a range of diseases and disorders (Carrasco-Ramiro *et al.*, 2017). Though personalized medicine is still in the early stages for BCa, there have been considerable advances in this field. Color diagnostics was developed to sequence DNA obtained from saliva samples for hereditary/germline variants and has been used globally by physicians and genetic counsellors to conduct multi-gene testing on ~30 genes and assess risks for developing breast and ovarian cancer (Adedokun *et al.*, 2020; Crawford *et al.*, 2017). There are also more commercial genomics tests such as 23andme that can be used without physician recommendation to assess variants in genomic DNA (gDNA) obtained from saliva. Tools like these have contributed significantly to our understanding of the genetic profile of BCa and have aided in identifying mutations in other genes apart from the traditional *BRCA1* and *BRCA2* mutations. In addition to these tools that assess germline variants and pre-tumour onset genomic profiles, clinicians have embarked on incorporating results from multi-gene panels for somatic profile assessment using either **f**ormalin-**f**ixed **p**araffin-**e**mbedded (FFPE) or fresh frozen tumour samples. It is now mandated in North America by the AJCC's recently released guidelines for cancer staging, to include molecular profiling and the use of multi-gene panels such as Mammaprint, EndoPredict and Oncotype Dx to appropriately stage the progression of BCa (Amin *et al.*, 2017). For example, Mammaprint uses DNA microarray and qRT-PCR technology on fresh frozen tissue for interrogation of variants in up to 70 genes in BCa

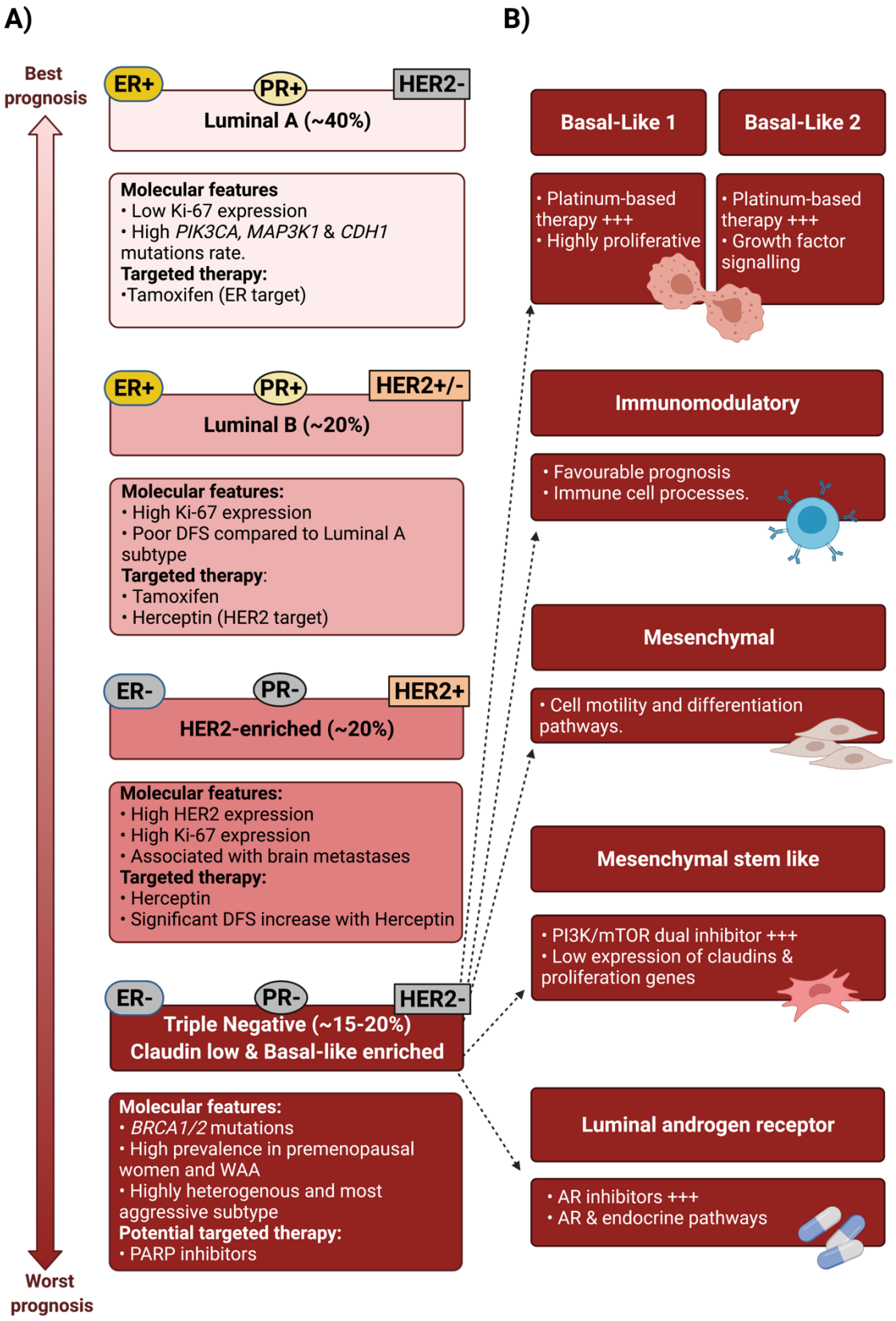
patients, regardless of **h**ormone **r**eceptor (HR) status and generates predictive results for risk of relapse, response and benefit of neoadjuvant and adjuvant chemotherapy respectively (Sabatier *et al.*, 2014). There is also now a stronger focus on patients with low frequency mutations (5-10% frequency), that when combined, may affect specific molecular pathways that have significant therapeutic potential (Lusito *et al.*, 2019; Sabatier *et al.*, 2014). For instance, though defects in *CDH1* are well documented in breast and other cancers, there is a lower frequency of these mutations in BCa when compared to the mutation frequency of *BRCA1*, *TP53* and *PIK3CA*. Recently, preclinical synthetic lethality—which is the combined lethal effect of defects or variations in two genes that would otherwise not be lethal if these variations occurred independently—was achieved *in vitro*, *in vivo* and *ex vivo* in tumours with *CDH1* defects when ROS1 inhibitors (such as crizotinib and foretinib) were used (Bajrami *et al.*, 2018). Additionally, using TNBC mice models, **w**hole **e**xome **s**equencing (WES) and **RNA seq**uencing (RNA-Seq), tumours with *Fgfr2* fusion accompanied by *Brca1* mutation showed better response to combination therapies that targeted both genes compared to tumours treated solely with only one (i.e. FGFR2 inhibitor and PARP inhibitor) (Liu *et al.*, 2018). Thus, identification of variants in novel genes and understudied pathways may be important keys to unlocking therapeutic agents. Also, combination of novel potential biomarkers or genomic variants in large repositories online or otherwise will greatly aid in development of therapeutic agents.

1.4.1 Breast Cancer Subtypes, Heterogeneity and Classification

Recent advancements within the field of BCa diagnostics have increased awareness about various BCa subtypes and hold promise to aid in more advanced therapeutic options for

patients. Using a combination of IHC and gene expression profiling, BCa can be classified into six subtypes (Prat & Perou, 2011). These include the less aggressive luminal A and B subtypes, the highly aggressive HER2-enriched, claudin-low and basal-like BCa subtypes (Prat & Perou, 2011) and the normal-like subtype (**Figure 1.1**). Luminal A and B BCa are typically referred to as HR-positive BCa due to the presence of ER and/or PR (Yanagawa *et al.*, 2012). Claudin-low and basal-like BCa subtypes are synonymously referred to as **t**riple-**n**egative **BC**a (TNBC) due to the likely absence of ER, PR and HER2 (Alluri & Newman, 2014) and their high-grade tumour status. In addition to gene-based classification, IHC-based classification of select markers (ER, PR, HER2 and Ki-67) has also been adopted by the St. Gallen International Expert Consensus for the aforementioned BCa subtypes and show high concordance with prognosis and potential for various therapeutic options for BCa patients (Goldhirsch *et al.*, 2013; Goldhirsch, Wood, Coates, Gelber, Thürlimann, *et al.*, 2011).

Figure 1.1: Molecular and therapeutic landscapes of (A) intrinsic breast cancer subtypes and (B) TNBC subtypes. (A) There is considerable molecular heterogeneity for BCa subtypes in the context of ER, PR and HER2 expression. Luminal A tumours are HR-positive (ER and PR positive), HER2-negative, and express low levels of the proliferation marker Ki-67. These tumours can be treated with Tamoxifen, an ER-targeted therapy that prevents ER-mediated BCa growth. Luminal B tumours are HR and HER2 positive or negative, and express low levels of Ki-67. Tamoxifen and Herceptin (HER2-targeted therapy for HER2-positive Luminal B cases) can be used with this Luminal B subtype; however, there is poorer **d**isease-**f**ree **s**urvival (DFS) than Luminal A tumours. HER2 enriched tumours are HR-negative but HER2 positive with high levels of Ki-67 expression. These tumours can be effectively treated with Herceptin which has led to a significant increase in DFS for HER2-enriched tumours after its introduction as a BCa therapy. **(B)** TNBC tumours are negative for HR and HER2, with high frequency of *BRCA1/2* mutations. TNBC can be further divided into at least six different subtypes (Basal-like 1 & 2, Immunomodulatory, Mesenchymal, Mesenchymal Stem Like, and Luminal Androgen Receptor) with enrichment for varying pathways and differential responses to available treatment options. Created with BioRender.com


A)
B)
Best prognosis
Worst prognosis
ER+ PR+ HER2-
Luminal A (~40%)
Molecular features
• Low Ki-67 expression
• High PIK3CA, MAP3K1 & CDH1 mutations rate.
Targeted therapy:
•Tamoxifen (ER target)
ER+ PR+ HER2+/-
Luminal B (~20%)
Molecular features:
• High Ki-67 expression
• Poor DFS compared to Luminal A subtype
Targeted therapy:
• Tamoxifen
• Herceptin (HER2 target)
ER- PR- HER2+
HER2-enriched (~20%)
Molecular features:
• High HER2 expression
• High Ki-67 expression
• Associated with brain metastases
Targeted therapy:
• Herceptin
• Significant DFS increase with Herceptin
ER- PR- HER2-
Triple Negative (~15-20%)
Claudin low & Basal-like enriched
Molecular features:
• BRCA1/2 mutations
• High prevalence in premenopausal women and WAA
• Highly heterogenous and most aggressive subtype
Potential targeted therapy:
• PARP inhibitors
Basal-Like 1
Basal-Like 2
• Platinum-based therapy +++
• Highly proliferative
• Platinum-based therapy +++
• Growth factor signalling
Immunomodulatory
• Favourable prognosis
• Immune cell processes.
Mesenchymal
• Cell motility and differentiation pathways.
Mesenchymal stem like
• PI3K/mTOR dual inhibitor +++
• Low expression of claudins & proliferation genes
Luminal androgen receptor
• AR inhibitors +++
• AR & endocrine pathways

1.2.1.1 Luminal A subtype

It is predicted that between 40-50% of diagnosed breast tumours are of the Luminal A subtype (Gao & Swain, 2018). This subtype is molecularly characterized by the overexpression of ER/PR, low or no expression of HER2 and low (<14%) Ki-67 expression (Goldhirsch, Wood, Coates, Gelber, Thurlimann, *et al.*, 2011). The proliferative marker Ki-67 has recently been found to significantly correlate with histological grade in a large cohort of patients regardless of BCa subtyping (Liang *et al.*, 2020) and low Ki-67 expression predicts a favourable prognosis. Luminal A BCa also has a favourable prognosis because of targeted endocrine therapies such as Tamoxifen, which are typically administered to treat ER-positive tumours. Tamoxifen was discovered in the 1960s and was clinically approved in 1985 as this was a promising therapy for ER-positive tumours (Tremont *et al.*, 2017). Tamoxifen works by competitively binding to ER and preventing estrogens from binding to ER thus inhibiting tumour proliferation (Krauss & Stickeler, 2020).

The Luminal A subtype has a 25-year **d**istant **r**ecurrence-**f**ree **i**nterval (DRFI) rate of up to 87% with Tamoxifen treatment (Yu *et al.*, 2019). Luminal A tumours also have the highest frequency of *PIK3CA* mutations (~45%), *MAP3K1* (~13%), *CDH1* (~9%) compared to other BCa subtypes and clinical trials focused on this genomics profile are currently being conducted (Santarpia *et al.*, 2016). The normal-like subtype is similar to the Luminal A subtype via IHC (ER/PR overexpression, low HER2 and low Ki-67) and was discovered due to its high similarity with normal breast tissue and low tumour cellularity (<50%) (Prat & Perou, 2011). Though data are limited since the normal-like

subtype was defined with a small number of tissues, the Luminal A subtype has a better prognosis than normal-like tumors (Prat & Perou, 2011).

1.2.1.2 Luminal B subtype

Luminal B BCa is also characterised by the overexpression of ER and can be either HER2 positive or negative; however, the 25-year DRFI is 70% for patients diagnosed with Luminal B BCa compared to 87% for Luminal A tumours treated with Tamoxifen (Yu *et al.*, 2019). Other studies have also found that Luminal B BCa has an overall poorer disease-free survival rate compared to Luminal A, has a predisposition to relapse in bone and pleural tissue and also has a tendency to be insensitive to endocrine therapy (compared with Luminal A subtype) and chemotherapy (compared to basal-like and HER2 enriched subtypes) (Tran & Bedard, 2011). This BCa subtype has lower expression of luminal-related proteins PR, FOXA1 but not ER (Prat *et al.*, 2015). Additionally, Luminal B tumours are typically more proliferative than Luminal A tumours since there is higher expression of proliferation markers Ki-67 when compared with Luminal A tumours. High Ki-67 expression, proliferative ability and poorer survival are thus hallmarks that distinguish Luminal B tumours from Luminal A tumours.

1.2.1.3 HER2 subtype

The HER2-enriched BCa subtype is characterised by overexpression of the *HER2/neu* oncogene. *HER2/neu* encodes a tyrosine kinase that when overexpressed, activates signaling pathways involved in cell proliferation and invasiveness (e.g. PI3K/AKT and Ras/Raf/MEK/MAPK), ultimately leading to an overall aggressive tumour (Hou *et al.*,

2020; Slamon *et al.*, 1987). This BCa subtype had one of the worst prognoses of the BCa subtypes before the introduction of HER2-targeted therapy. Trastuzumab (Herceptin) is a monoclonal antibody that binds to HER2 to prevent activation of downstream signalling pathways that promote tumour proliferation and differentiation (Baselga *et al.*, 1998). Since FDA approval of Herceptin in 1998 there has been a gradual decrease in mortality rates for patients diagnosed with this HER2-enriched BCa subtype. (Patel *et al.*, 2020). Recent data after 11-year follow-up from the landmark HERceptin Adjuvant (HERA) clinical trial (NCT00045032) show that one year administration of Herceptin to patients with HER2-positive early BCa, significantly reduced risk of death and increased **d**isease-**f**ree **s**urvival (DFS) by ~25% and ~24% respectively compared to patients that did not receive Herceptin (Cameron *et al.*, 2017). Establishment of this therapy and others currently being investigated will largely help reduce the mortality rate from this aggressive BCa subtype that affects ~ 20% of BCa cases and typically associates with metastases to the brain which is often fatal (Patel *et al.*, 2020).

1.2.1.4 Claudin-low subtype

The Claudin-low subtype is a recently discovered subtype added to the intrinsic BCa subtypes due to distinct molecular profiling. This subtype is mostly ER, PR and HER2 negative and is characterised by high metastatic activity and has low expression of tight junction and cell-cell adhesion proteins such as E-Cadherin and Claudin-3, -4, and -7 (Prat & Perou, 2011). Another hallmark of this BCa subtype is immune and stromal cell infiltration (Fougner *et al.*, 2020). Through hierarchal clustering, the Claudin-low subtype overlaps substantially with the basal-like subtype; however, there is a lower signature for

proliferation associated with this subtype (low Ki-67 expression) compared to the basal-like group (Fougner *et al.*, 2020; Prat & Perou, 2011). These tumours are highly heterogenous, have considerable overlap with the other intrinsic subtypes and have significantly less *TP53* mutations when compared with basal-like, Luminal A and normal-like subtypes (Fougner *et al.*, 2020). The claudin-low subtype may benefit from inhibition of **t**ransforming **g**rowth **f**actor **b**eta (TGFβ) signalling since this subtype has been associated with a stem-cell like phenotype that is induced by TGFβ in claudin-low cell lines (Garrido-Castro *et al.*, 2019). Breast tumours with high activation of TGFβ signalling are highly proliferative and have a high propensity to metastasize through EMT (Adorno *et al.*, 2009; Pang *et al.*, 2016; Xie *et al.*, 2018). The TGFβ signalling pathway has been established as one of the major pathways influencing the aggressive nature of BCa. Claudin-low tumours also have high expression of immune markers such as **p**rogrammed **d**eath **l**igand **1** (PD-L1), suggesting that these tumours may be eligible candidates for PD-L1 therapy. Approximately 30% of BCa samples express PD-L1 which decreases T-cell proliferation and increases T-cell apoptosis in BCa cells (Mittendorf *et al.*, 2014). Expression of this T-cell inhibitory molecule has been implicated in allowing tumour cells to evade the host immune system (Mittendorf *et al.*, 2014). Indeed, therapeutic agents for PD-L1 such as atezolizumab have recently proven to be beneficial for patients with metastatic breast cancer (Kagihara *et al.*, 2020; Kang & Syed, 2020) and thus are of clinical importance for this subset of patients and potentially for Claudin-low tumours that express PD-L1.

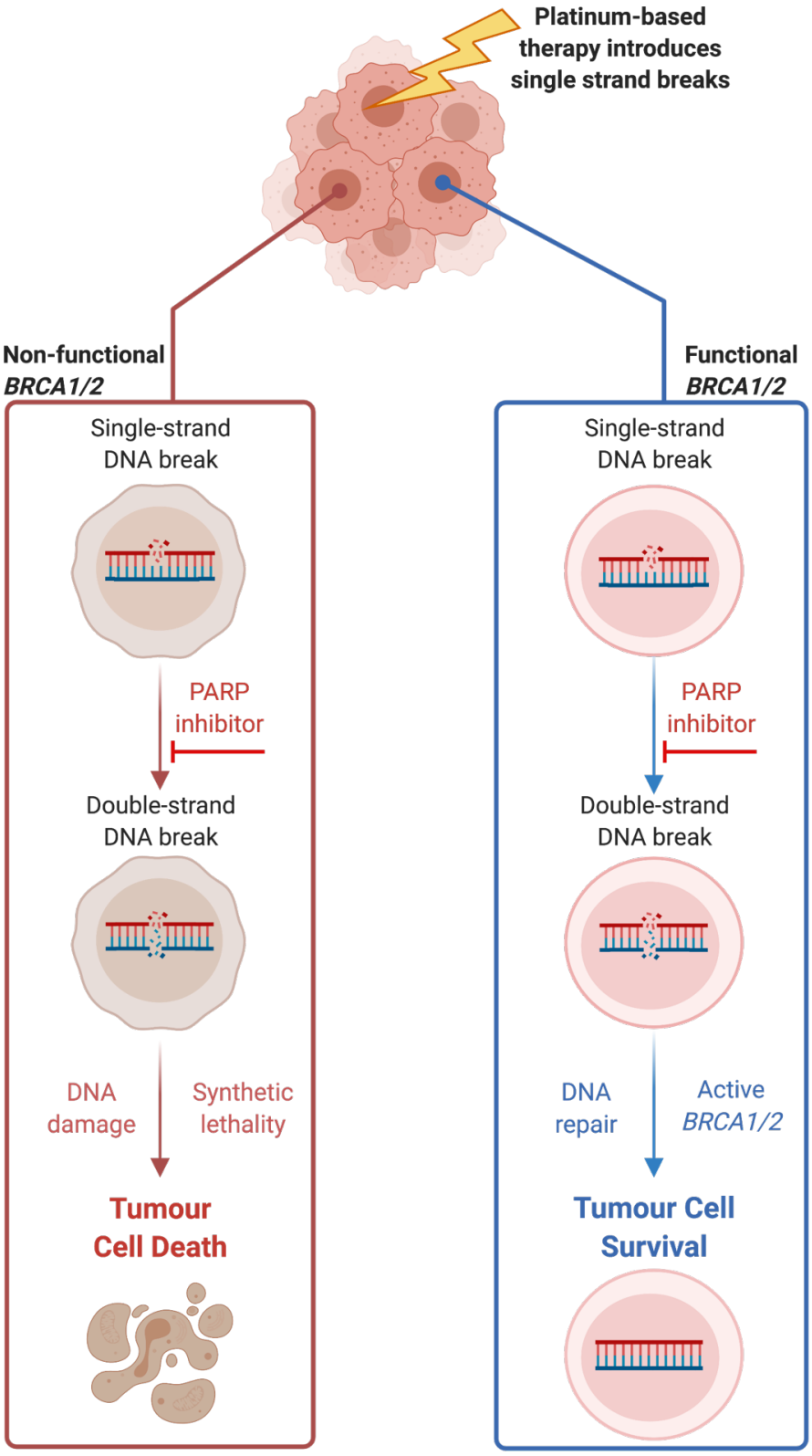
1.5 TRIPLE NEGATIVE BREAST CANCER (TNBC)

TNBC represents ~15% of all BCa cases and is an operational term that refers to a heterogeneous collection of breast cancers (largely overlapping with claudin-low and basal-like BCa) (Alluri & Newman, 2014; Gordon, 2003; Prat & Perou, 2011). The basal-like subtype, commonly referred to as triple-negative, is also characterised by low to no expression of ER, PR and HER2. However, it is highly proliferative with high Ki-67 expression and genomic instability. Though a majority of basal-like tumours are HR and HER2 negative, ~ 10% of basal-like tumours are HR positive (Prat & Perou, 2011). Unlike ER and HER2 positive tumors for which there are targeted therapies such as Tamoxifen and Herceptin respectively, there are no established targeted therapies for TNBC which lack expression of these targeted receptors (Hurvitz & Mead, 2016). Thus, patients with early stage TNBC are generally treated with conventional chemotherapies such as platinum-salts (e.g. Cisplatin), docetaxel, and paclitaxel in combination with more specific chemotherapeutics like Bevacizumab (Lehmann *et al.*, 2016; Miles *et al.*, 2010; Miller *et al.*, 2007). Compared to other subtypes, TNBC tumours intriguingly have a higher pathological complete response (pCR) rate – which is a measure of the absence of invasive/*in situ* cancer in the breast and/or axillary nodes – but lower 3-year progression free and overall survival (Biswas *et al.*, 2019; Liedtke *et al.*, 2008). Although some patients with TNBC are sensitive to chemotherapy treatment (O'Reilly *et al.*, 2015), most TNBC tumors are typically resistant to these treatments and often have high recurrence resulting in a poor prognosis (Clark *et al.*, 2014; Kassam *et al.*, 2009).

Since TNBC is often associated with *BRCA1* mutations and "BRCAness", researchers have investigated how these mutations can sensitize TNBC tumours to various forms of chemotherapeutic agents (Chen, Wu, *et al.*, 2018). It was recently found that a subset of TNBC patients with deleterious *BRCA1/2* germline mutations responded significantly better to carboplatin (platinum-based therapy) than docetaxel (taxane-based therapy) (Tutt *et al.*, 2018). Tumour cells deficient for *BRCA1/2* cells are particularly sensitive to platinum-based therapies due to the fact that inter-strand cross-links formed by these therapies can only be adequately repaired by **h**omologous **r**ecombination DNA **r**epair (HRR) (Turner & Tutt, 2012). Since *BRCA1* and *BRCA2* are required for HRR, cells with mutations in *BRCA1/2*-altering protein function will undergo cell death (Turner & Tutt, 2012). Taxane-based therapies were also found to be beneficial for TNBC patients with *BRCA1*/2 germline mutations (Bignon *et al.*, 2018). However, this type of therapy may lead to increased risk of chemotherapy-related gastrointestinal toxicities, specifically in *BRCA2* mutation carriers (Bayraktar *et al.*, 2020). Additionally, **p**oly (**A**DP-**r**ibose) **p**olymerase (PARP) inhibitors have been shown to be an effective treatment for TNBC, especially those with *BRCA1/2* mutations (**Figure 1.2**) (Papadimitriou *et al.*, 2018; Sharma, 2017). The combination of PARP inhibition and inactive *BRCA1/2*, creates an environment for synthetic lethality of tumour cells (Lord & Ashworth, 2017; O'Neil *et al.*, 2017). These polymerases, specifically *PARP1* and *PARP2* are critical in response to **s**ingle-**s**tranded DNA **b**reaks (SSB) as they recruit other proteins such as XRCC1 to the site of the SSB to assemble repair factors at the break (Lord & Ashworth, 2017). In cells with PARP inhibition and *BRCA1*/*2* mutations, HRR cannot occur to maintain cell viability and the

tumour cell undergoes cell death (**Figure 1.2**) (Beniey *et al.*, 2019). Thus, PARP inhibitors are a type of targeted therapy that has led to considerable improvements in TNBC patients with *BRCA1/2* defects (Geenen *et al.*, 2018). Ongoing studies combining PARP inhibitors with other inhibitors such as EGFR inhibitors and other chemotherapeutics to enhance their cytotoxic effects are currently underway (Stringer-Reasor *et al.*, 2021). Most recently, on July 26th 2021, the U.S. Food & Drug Administration (FDA) approved pembrolizumab in combination with chemotherapy for high-risk early stage TNBC patients as well as for patients with locally recurrent or metastatic TNBC with PD-L1 expression (NCT03036488). This recent therapy further highlights that consistent investigation into the molecular profiles of breast tumours is crucial for uncovering novel therapies, especially for the difficult to treat TNBC subtype.

Figure 1.2: Synthetic lethality in breast cancer through BRCA1 and PARP. Tumour cells treated with platinum-based therapies such as cisplatin undergo single-strand DNA breaks. In the presence of PARP inhibitors, PARPs are not recruited to the site of the single-strand breaks for DNA repair and thus double strand breaks (DSB) are created. If cells have functional *BRCA1/2*, DSBs can be repaired through homologous recombination DNA repair and tumour cells survive and proliferate. However, cell death is initiated for cells without functional *BRCA1/2*. Created with and adapted from BioRender.com


Platinum-based therapy introduces single strand breaks
Non-functional BRCA1/2
Functional BRCA1/2
Single-strand DNA break
PARP inhibitor
Double-strand DNA break
DNA damage
Synthetic lethality
Tumour Cell Death
Single-strand DNA break
PARP inhibitor
Double-strand DNA break
DNA repair
Active BRCA1/2
Tumour Cell Survival

1.5.1 TNBC Heterogeneity and Genetic Signatures

Although all TNBC tumours share receptor negative status for ER, PR and HER2, there is substantial genomic and pathological heterogeneity within this subtype. Original classification by Lehmann and colleagues suggested that there were at least six different molecular subtypes of TNBC tumors (**Figure 1.1**) (Lehmann *et al.*, 2011). These subtypes as defined by Lehmann *et al.* included two **b**asal-**l**ike (BL1 and BL2), an **i**mmuno**m**odulatory (IM), a **m**esenchymal (M), a **m**esenchymal **s**tem-**l**ike (MSL) and a **l**uminal **a**ndrogen **r**eceptor (LAR) subtype (**Figure 1.1**) (Lehmann *et al.*, 2011). However, after taking tumor microenvironment into consideration, further analyses by this group narrowed the subtypes to four main groups – M, LAR, BL1 and BL2 (Palma *et al.*, 2015). Transcript levels from immune and normal stromal cells showed that the IM and MSL subtypes represent tumours with substantial infiltrating lymphocytes and tumour-associated mesenchymal cells, respectively (Palma *et al.*, 2015), which suggests that these subtypes may be candidates for various forms of immunotherapeutic agents. Furthermore, additional characterization of TNBC tumours further grouped them into LAR, mesenchymal (MES), basal-like immunosuppressed (BLIS) and basal-like immune-activated (BLIA) groups that differ by biomarker expression and potential targets of treatment for each subtype (Burstein *et al.*, 2015). In general, TNBC classification is based on gene expression and DNA profiling, and is important as each subtype responds differently to various forms of chemotherapy/treatment options and have varying clinical courses (Lehmann *et al.*, 2016; Palma *et al.*, 2015). Though no significant differences were found by Lehmann and colleagues between these subtypes for overall and distant relapse-

free survival, pCR rates were significantly higher in patients with BL1 subtype (49%) compared to all other subtypes (31%) (Lehmann *et al.*, 2016). Indeed, *in vitro* studies have shown that cell lines of the basal-like subtype (BL1 & BL2) respond significantly better to cisplatin (platinum-based therapy) compared to bicalutamide (**A**ndrogen **r**eceptor (AR) inhibitor) whereas the LAR subtype responds significantly better to bicalutamide, and the mesenchymal-like responds better to cisplatin or PI3K/mTOR dual inhibitor NVP-BEZ235 (Lehmann *et al.*, 2011). The BL1 and BL2 response to platinum-based therapies could be due to the high mutation rate in *BRCA1/2* within this subtype. Conversely, the LAR subtype is not typically enriched for *BRCA1/2* mutations, and are characterized by the presence of AR, a luminal pattern of gene expression (e.g. high FOXA1, GATA3) despite ER-negativity and mutations in *PIK3CA* and *KMT2C* (Garrido-Castro *et al.*, 2019). In addition to the clinical use of AR antagonists like bicalutimide, a clinical trial is currently being conducted to determine the efficacy of combining enzalutamide (another AR antagonist) with alpelisib (a PI3K inhibitor) in patients that have no PTEN expression (NCT03207529). Therefore, having tumours appropriately subtyped and having a more comprehensive understanding of the molecular profile at the genomic, transcriptomic and proteomic levels will lead to more effective therapeutic choices and is now an emerging field of personalized medicine.

Biomarker discovery for TNBC is crucial as further discovery will help to reveal novel diagnostic, prognostic, and therapeutic agents for this aggressive subtype. Currently, TNBC biomarkers are separated into three main groups (i) *BRCA1/BRCA2* and DNA repair genes, (ii) immune checkpoint inhibitors and (iii) PI3KCA and PTEN as predictive markers

(Cocco *et al.*, 2020). The recent discovery of PARP inhibitors in TNBC patients with defective *BRCA1/2* has unveiled multiple types of PARP inhibitors from various pharmaceutical technology with varying degrees of efficacy in TNBC (Cocco *et al.*, 2020). Additionally, other recent TNBC biomarkers of interest include the transcription factor (TF) Kaiso that has been implicated in TNBC proliferation, invasion, metastasis and chemoresistance (Bassey-Archibong *et al.*, 2016; Bassey-Archibong, Rayner, *et al.*, 2017; Vermeulen *et al.*, 2012).

1.6 THE TRANSCRIPTION FACTOR KAISO AND TNBC

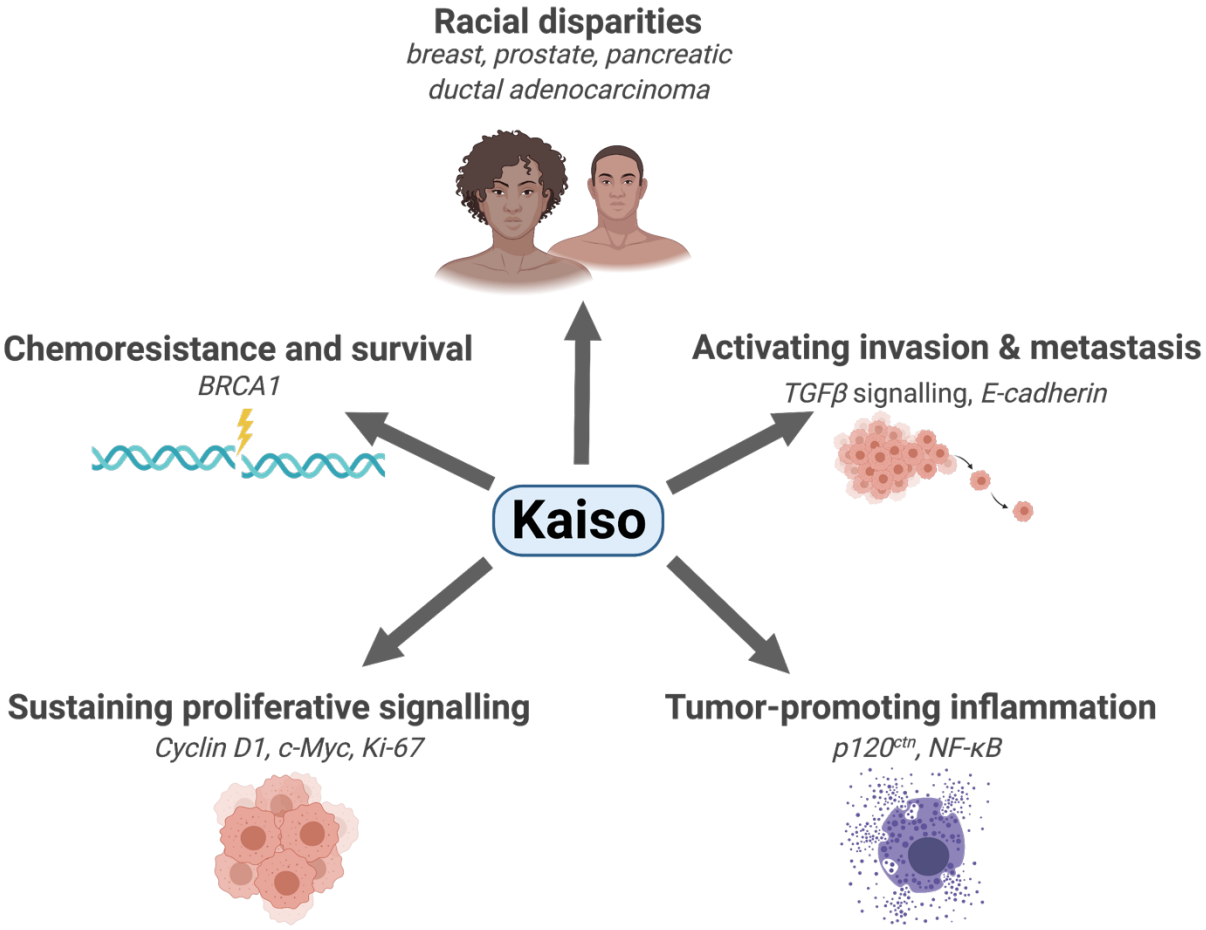
1.5.1 Kaiso and its role in biological processes

Recently, the poxvirus and zinc finger (POZ-ZF) transcription factor Kaiso was reported to be highly expressed in TNBC (Bassey-Archibong *et al.*, 2016; Jones *et al.*, 2014; Vermeulen *et al.*, 2012). Kaiso was discovered over two decades ago in association with the cell adhesion Armadillo protein **p120**-**c**a**t**eni**n** ($p120^{ctn}$) (Daniel & Reynolds, 1999). Unlike many other POZ-ZF transcription factors, Kaiso has the unique ability to recognise and bind to DNA via the sequence specific **K**aiso **b**inding **s**ite (KBS; TCCTGCNA), the palindromic TCTCGCGAGA sequence, and at **me**thylated **CpG** (meCpG) islands (Blattler *et al.*, 2013; Daniel & Reynolds, 1999; Daniel *et al.*, 2002; Prokhortchouk *et al.*, 2001; Raghav *et al.*, 2012). Remarkably, Kaiso has the ability to act as a repressor or activator in a context-dependent manner, unlike many other POZ-ZF transcription factors that mainly act as repressors (Kelly & Daniel, 2006).

In cancer, Kaiso has been implicated in several tumour-related processes such as **e**pithelial to **m**esenchymal **t**ransition (EMT), cell proliferation, apoptosis, inflammation,

and hypoxia as extensively reviewed (Pierre *et al.*, 2019, **Chapter 6**). Clinically, Kaiso expression and localisation has been linked to poor prognosis in breast, prostate, colorectal and lung cancers (Dai *et al.*, 2009; Jones *et al.*, 2012; Lopes *et al.*, 2008; Vermeulen *et al.*, 2012). For example, higher Kaiso expression was observed in primary & metastatic tumor biopsies compared to normal colon tissues in the context of colorectal cancer (Pierre *et al.*, 2015). Furthermore, nuclear Kaiso expression correlated significantly with higher histological grade in prostate and breast cancer, along with invasive ductal carcinoma (IDC) diagnoses, and ER negativity (Jones *et al.*, 2012). Notably, Kaiso interacts with the tumor suppressor p53 to activate transcription of pro-apoptotic genes, but when Kaiso interacts with mutant p53, it represses transcription of pro-apoptotic genes (Bassey-Archibong, Rayner, *et al.*, 2017; Koh *et al.*, 2015). While these data are promising for establishing Kaiso as a biomarker for BCa, especially in TNBC with high *TP53* mutation rate, further investigation is warranted to fully elucidate Kaiso's role in various cancers.

Figure 1.3. Kaiso roles in tumorigenesis. Kaiso activates tumour cell invasion and metastasis via regulation of TGFβ signalling, miR-200 and E-cadherin expression. In Kaiso depleted TNBC cells, TGFβ signalling pathway was reduced as well as decreased expression of EMT proteins such as E-Cadherin (Bassey-Archibong *et al.*, 2016). In prostate cancer cells, Kaiso regulates EMT through mimR-200 family (Abisoye-Ogunniyan *et al.*, 2018). In several mouse models, Kaiso promotes intestinal inflammation. Kaiso sustains proliferative signalling through Cyclin-D1, Myc, and Ki-67. Kaiso is implicated in chemoresistance and BCa overall survival through its interaction with BRCA1. Kaiso is also implicated in racial disparities in multiple cancer types. Created with BioRender.com and adapted from **Chapter 6, Figure 2.**


Racial disparities
breast, prostate, pancreatic ductal adenocarcinoma
Chemoresistance and survival
BRCA1
Activating invasion & metastasis
TGFβ signalling, E-cadherin
Kaiso
Sustaining proliferative signalling
Cyclin D1, c-Myc, Ki-67
Tumor-promoting inflammation
p120ctn, NF-κB

1.5.2 Kaiso and TNBC

Several studies from our group and others have implicated Kaiso in the progression of TNBC. High Kaiso expression correlates with the poor overall survival of African American BCa patients compared to their Caucasian counterparts (Jones *et al.*, 2014), which suggests a role for high Kaiso expression in the racial disparity affiliated with BCa mortality rates in WAA. In support of these previous findings, we found that Kaiso is highly expressed in WAA TNBC tissues compared to Caucasian TNBC tissues, with high Kaiso nuclear localisation and staining intensity associated with increasing percentage of African ancestry (Bassey-Archibong *et al.*, 2017; **Appendix Chapter 4**). Additionally, high Kaiso and BRCA1 expression together were found to correlate with poor overall survival in breast cancer patients (Bassey-Archibong, Rayner, *et al.*, 2017). Functionally, Kaiso expression was found to be essential for the proliferation, survival and metastasis of TNBC as Kaiso depletion increased the apoptosis and reduced the proliferation and metastases of TNBC cells to secondary organs in mouse models (Bassey-Archibong, Rayner, *et al.*, 2017). Concomitant with these changes in cell function, Kaiso-depleted TNBC cells exhibited decreased expression of key EMT markers such as Vimentin, Slug and Zeb1 (Bassey-Archibong *et al.*, 2016). *In vivo*, Kaiso-depleted TNBC xenografts had delayed tumor onset and reduced metastatic ability to secondary organs (Bassey-Archibong *et al.*, 2016; Bassey-Archibong, Rayner, *et al.*, 2017). Furthermore, Kaiso depletion resulted in increased sensitivity of TNBC cells to the chemotherapeutic agent Cisplatin (Bassey-Archibong, Rayner, *et al.*, 2017). These studies provided key insights about Kaiso's biological roles in TNBC and suggest that Kaiso may be a potential biomarker for TNBC aggressiveness.

Together, these findings raise the possibility that: 1) Kaiso plays a crucial role in TNBC aggressiveness; 2) WAA diagnosed with TNBC have a predisposition to increased Kaiso expression; and 3) Kaiso plays a role in the racial disparity in TNBC prevalence and outcome in WAA.

1.7 TNBC AND RACIAL DISPARITIES

In addition to genetic factors such as Kaiso possibly contributing to the racial disparity in BCa incidence and outcomes, multiple studies across the USA are pointing towards the "weathering hypothesis" proposed by Linnenbringer and colleagues to *"explain the roles of the dynamic interplay among psychosocial stressors, physiological & behavioral responses, and genomic pathways [and how they] contribute to the increased risk of hormone receptor (HR)- breast cancer among Black women"* (Linnenbringer *et al.*, 2017). Though this is a relatively new research field, more researchers are now examining the interplay of these factors and how they impact BCa onset. Notably, HR positive BCa is more frequently diagnosed in WEA, while TNBC is diagnosed at a higher rate in young WAA (Moller *et al.*, 2016; Stark *et al.*, 2010). This observation is corroborated by a recent population-based report in the USA that found non-Hispanic Black (NHB) women and Hispanic women had higher odds of TNBC diagnosis (over 2-fold for NHB) compared to non-Hispanic White (NHW) women (Scott *et al.*, 2019). This study and others have shown that in addition to higher TNBC prevalence rates, WAA experience a more aggressive clinical course of TNBC compared to women of other ethnicities (Brewster *et al.*, 2014; Dietze *et al.*, 2015; Stark *et al.*, 2010). Specifically, TNBC patients of African ancestry were diagnosed at a younger age and had larger and higher-grade tumors compared with

Caucasian women (Brewster *et al.*, 2014; Stark *et al.*, 2010). Moreover, beyond the TNBC diagnosis, it was found that, WAA with TNBC had higher mortality rates than Caucasian women with TNBC (Dietze *et al.*, 2015). To explain this, investigations into what drives this disparity, from an environmental perspective were done. One population-based study of Californian women found that Black women living in neighbourhoods with more Black residents had significantly lower odds of being diagnosed with TNBC relative to the HR+/HER2− subtype (Linnenbringer *et al.*, 2020). This perhaps is associated with the perceived psychosocial stressors of living in a more segregated neighbourhood compared to living in a neighbourhood where residents are more comfortable with their fellow neighbours. In this study, they also found that living in areas with higher neighborhood median household incomes was associated with lower odds of TNBC relative to the HR+/HER− subtype in both Black and White women. Another population-based study from the USA found a higher risk associated with ER- breast cancer in Black women born in Jim Crow states (e.g. Alabama, Mississippi) compared to Black women born in other states. Notably, this association was not found in White women across these two comparator groups (Krieger *et al.*, 2017). Jim Crow states are those 21 states and the District of Columbia in the USA that practiced legal racial discrimination until it was outlawed by the US Civil Rights Act in 1946 (Murray, 1950). Supporting these studies, it was recently found that Black women residing in low SES neighbourhoods, particularly those with lower density of Black residents, had a greater relative risk of TNBC (Qin *et al.*, 2020). While these findings are correlations and do not show causal associations, they provide evidence of how social isolation can impact tumorigenesis. In murine models, it

has been documented that social isolation leads to younger onset of mammary tumours as well as increased rates of mammary tumour growth and size (McClintock *et al.*, 2005; Williams *et al.*, 2009). TNBC onset can therefore not simply be defined by genomic aberrations and alterations alone. The interplay of SES, social stressors, behavioural patterns, and genetic variations should also be strongly taken into consideration when trying to determine why these disparities exist.

While it was initially postulated that SES and access to health care alone were key factors underlying the high BCa mortality rates in WAA, these factors do not fully explain the racial disparity associated with TNBC prevalence and outcome in young WAA and suggest other factors, including biological factors, as key causes of this disparity (Iqbal *et al.*, 2015). Recently, a high frequency of pathogenic and likely pathogenic variants for cancer-associated genes such as *BRCA1, BRCA2* and *PALB2* was highlighted in Caribbean women with breast and ovarian cancer (George *et al.*, 2021). Additionally, a higher *TP53* mutation rate was observed in Nigerian (62%) and African American women (46%) compared with Caucasian women with breast cancer (29%) (Pitt *et al.*, 2018). Notably, Pitt *et al.* used data collected from The Cancer Genome Atlas (TCGA), which was created to profile and analyze genomics data at the DNA, RNA, protein, and epigenetics levels for a variety of cancers. This rich dataset has enabled researchers to investigate genetic alterations that could form the basis for novel therapies and progressing to a more personalized approach to therapy. However, the TCGA database and others collecting such data lack representation of cancer cases from diverse ethnic populations.

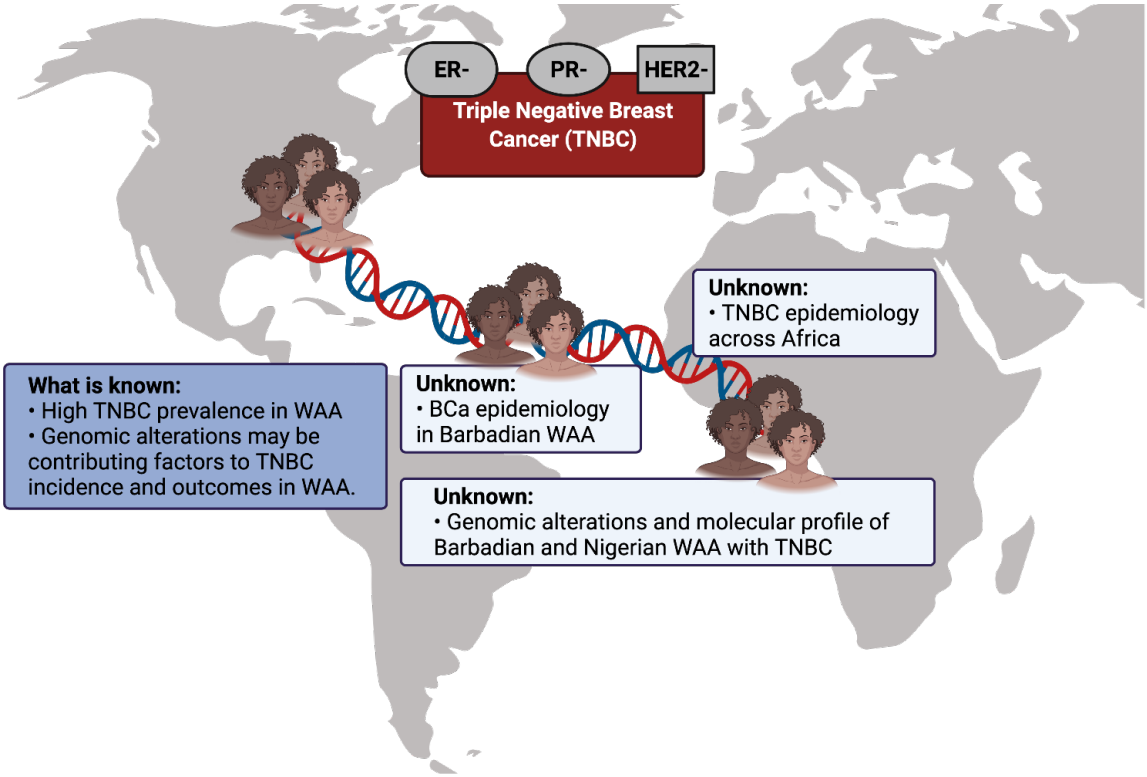
Currently, across all cancer types within the TCGA, there is an estimated 80% representation of European American cases whereas only 9% of African American samples were included (Yuan *et al.*, 2018). Thus, there has been an increased focus of genomics studies to include representation of minoritized groups from other ancestries, such as people of African ancestry. Genotyping studies have revealed that alleles such as the Duffy-null allele are associated with TNBC risk in a cohort enriched with African ancestry participants (Martini *et al.*, 2021). This study also revealed that the frequency of a protective variant for *ANKLE1* was lower in Ghanaian TNBC participants compared to European American TNBC participants which was intriguing as *ANKLE1* has been previously associated with DNA repair and has been used as a candidate for non-small cell lung cancer therapeutics (Whitehurst *et al.*, 2007). Further genotyping efforts have identified various loci associated with African ancestry and TNBC within the introns and/or near to previously known cancer-associated genes such as *KCNK2* and genes not previously associated with cancer such as *C5orf56* (Adedokun *et al.*, 2021). Though not much is understood about *KCNK2* in BCa, one study highlighted decreased mRNA expression in invasive breast tissue (Williams *et al.*, 2013). Additionally, *C5orf56* is **l**ong **n**on-**c**oding **RNA** (lncRNA) gene with little known about function and/or regulatory pathways. There is now an expansion of functional validation of these genetic signatures in WAA as well as how environmental factors also play a role in these genetic signatures found in WAA with TNBC.

RATIONALE & RESEARCH GOALS

The high prevalence and poor prognosis of TNBC in WAA emphasizes the importance of unravelling the genetic, molecular, and epidemiological factors that may be involved in

driving the early onset of TNBC in young women of African ancestry. Currently, the exact causes for high prevalence and poor outcomes have not been investigated widely and thus remain unknown. Recent studies have found that TNBC prevalence is highest among WAA across the diaspora with the highest TNBC prevalence reported in West-African populations (*e.g.,* Nigeria 50%, Ghana 82%), which are the founder population of most WAA in North America and the Caribbean (Mannix, 1962; Stark *et al.*, 2010). To date, there have been no comprehensive systematic reviews and meta-analyses investigating the frequency and prevalence of TNBC across Africa to fully understand the variation of BCa aggressiveness across the continent. Additionally, little is known about BCa subtypes and aggressiveness in the Caribbean where there is a higher percentage of African ancestry compared to North America. This study thus utilized ancestrally-related West African (Nigerian) and Caribbean (Barbadian) TNBC cohorts to investigate the epidemiological trends and molecular genetic profile in WAA. As most cancer genomics research to date have low inclusion of African and other diverse populations, potential therapeutic targets and diagnostic biomarkers in WAA might be overlooked. This work aims to investigate the genetic profile of DNA extracted from TNBC tissues obtained from both Nigerian and Barbadian women using WES to determine if there are any somatic mutations and/or genetic signatures that may be specific to TNBC in these WAA.

Figure 1.4: Conceptual framework for understanding TNBC in WAA. WAA have a high prevalence of TNBC and may have specific genomic and molecular patterns that contribute to tumour aggressiveness and TNBC prevalence in WAA. However, several unknown variables remain regarding the cause of TNBC prevalence in WAA namely (i) BCa epidemiology in Barbadian WAA—a population with a high percentage West African ancestry and the highest BCa incidence/mortality in the Caribbean, (ii) TNBC epidemiology across the African continent and (iii) specific genomic and molecular profile of ancestrally-related Barbadian and Nigerian WAA with TNBC. Model diagram created with BioRender.com


ER-
PR-
HER2-
Triple Negative Breast Cancer (TNBC)
What is known:
• High TNBC prevalence in WAA
• Genomic alterations may be contributing factors to TNBC incidence and outcomes in WAA.
Unknown:
• BCa epidemiology in Barbadian WAA
Unknown:
• TNBC epidemiology across Africa
Unknown:
• Genomic alterations and molecular profile of Barbadian and Nigerian WAA with TNBC

HYPOTHESIS AND OBJECTIVES

WAA diagnosed with TNBC share common genetic signatures and epidemiological trends that could lead to targeted interventions for effective treatments.

1. Compare the BCa epidemiological profile of women with a high African ancestry in Barbados with non-Hispanic white and non-Hispanic Black women in the USA.
2. Investigate TNBC frequency across African countries.
3. Investigate the genomic profile of WAA diagnosed with TNBC.
4. Investigate the expression profile of Kaiso and other BCa biomarkers in a **t**issue **m**icro**a**rray (TMA) comprised of Nigerian and Barbadian TNBC tissues.

CHAPTER 2: MATERIALS AND METHODS

Search Strategy and Selection Criteria

We used Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher *et al.*, 2009) as well as Meta-analysis of Observational Studies in Epidemiology (MOOSE) guidelines (Stroup *et al.*, 2000) as frameworks for our systematic review and meta-analysis (**Chapter 4**). On April 25th 2021, we searched PubMed, EMBASE, African Journals Online and Web of Science for relevant articles without date or language restrictions. Start date of the search was from inception of each database. A detailed version of our search strategy used in PubMed was modified for other databases. The search strategy and these modifications can be found in Table S1. All search terms were Medical Subject Heading terms, including TNBC terms ("TNBC", "triple-negative*", "triple negative") and terms for African countries ("Africa", "African" and names of all 54 African countries) and outcome variables ("rate*", "prevalence", "epidemiology"). We included all studies that met the following inclusion criteria (i) studies that use breast cancer tissue samples from indigenous African women of any age, in any care setting (clinical, community, *etc.*) and at any geographic location (ii) sample size of eligible participants ≥ 40 (as slightly more stringent criteria since normal distribution could be assumed at n= 30 (Krithikadatta, 2014)) (iii) studies that demonstrate at least one of the following: report on receptor status of breast tumours including ER, PR, and HER2 (iv) any primary study including but not limited to observational studies, cohort and case-control studies, cross-sectional studies, *etc.*

We excluded the follow types of publications: editorials, single case reports, case series, and commentaries, studies that assessed diagnostic measures and treatment options for women with TNBC in the absence of assessment of its prevalence, studies conducted in non-African nations without assessment of indigenous African TNBC rates or that of first-generation African immigrants. Study selection began by screening titles and abstracts of articles collected after employing the search strategy. The full text of these articles was then reviewed to assess inclusion. Two data abstractors independently reviewed articles at both the title/abstract and full text review stages. When there were discrepancies, a consensus was made in consultation with a third reviewer. The protocol for this review was not registered. Non-English studies (French) were included after translation through Google Translate followed by verification of translation by a French speaker.

Quality Assessment

Studies which passed full text review (Table S2) were evaluated for risk of bias using a tool developed by Hoy and colleagues specifically intended for prevalence-studies (Hoy *et al.*, 2012). Each study was assessed according to ten items assessing internal validity (Table S3) and assigned to have either low (score of 1) or high (score of 0) risk of bias for each question by two independent reviewers. A third reviewer mediated discrepancy and a final score per question was agreed upon. Studies were then classified based on the total score for all questions in the quality assessment tool as having a high (<=5), moderate (6-8), or low (>= 8) risk of bias.

Data Analysis

After reviewing full-text articles, a heat map was constructed with the number of studies, TNBC frequency and number of participants per country across African populations for unique studies using Google Sheets. All meta-analyses, meta-regressions and sensitivity analyses were completed using R (version 4.0.2) (R Core Team, 2020). Using the *metaprop* package in R, we conducted a meta-analysis of TNBC frequency among indigenous African women with breast cancer, stratified by country, region, risk of bias assessment, year of publication and the use of a validated tool for assessing receptor status. Freeman-Tukey double-arcsine transformation was used to stabilize the variances (Borges Migliavaca *et al.*, 2020) and a random effects model for our meta-analyses. Pooled TNBC frequency was estimated separately per country and per region as two studies included data from more than one region. When studies investigated African and non-African participants, only data from African participants were included in meta-analyses. Heterogeneity between studies was assessed with Cochran's Q, I^2, and H statistics. Meta-regression was done to explore heterogeneity using *metareg* package. We used Egger's test to investigate publication bias and small study effects using the *metabias* package. The full code can be found here: https://github.com/daniel-lab-mcmaster/Shawn_Rcodes/blob/main/full_penul.R.

Systematic review and meta-analysis findings are reported in **Chapter 4**

Study Patient Populations

Case reports from the Pathology Department at the Queen Elizabeth Hospital (QEH) in Barbados were examined between January 2007 and December 2017. Data were extracted

and cases were selected from all histologically confirmed BC cases during this time. A total of 2,136 breast cancer cases were identified. Data were also extracted from the Surveillance of Epidemiology and End Results (SEER) database for non-Hispanic Black (NHB) and non-Hispanic White (NHW) BCa cases in the USA with receptor status data (2010-2016). Archived formalin-fixed and paraffin-embedded (FFPE) tumor tissue blocks of Nigerian TNBC patients diagnosed at the Lagos University Teaching Hospital (LUTH), Abia State University Teaching Hospital (ABSUTH), University of Calabar Teaching Hospital (UCTH), University of Port Harcourt Teaching Hospital (UPTH), and FOMAS Hospital in Nigeria from 2011 – 2017, and Barbadian TNBC patients diagnosed at the Queen Elizabeth Hospital (QEH), Barbados from 2002 – 2017 respectively, were obtained from the Pathology Department at the respective hospitals after approval by each Ethics Committee. The FFPE tissue blocks for the TMA construction were then shipped to the Developmental Histology Lab at the Yale Pathological Tissue Services (YPTS), Yale University (Connecticut, New Haven, USA), where tissue sections of each block were hematoxylin & eosin (H&E) stained for histological confirmation. Representative tumor areas of each Nigerian and Barbadian FFPE specimens were then selected for the construction of a pilot tissue microarray (TMA). ER, PR and HER2 status of samples was confirmed at each institution by immunohistochemistry (IHC). Any tissue specimen with less than 1% staining for ER and PR was scored as negative. Similarly, 0 or +1 scoring for HER2 was considered negative. Available clinico-pathological information including age, tumor grade, tumor stage and lymph node involvement were retrieved from pathology reports and physician notes at these institutions. TNBC tissue samples of African American (AA) and

Caucasian American (CA) patients diagnosed at the Yale-New Haven Hospital, Connecticut, New Haven, USA from 1996 – 2004, were obtained by purchasing the Yale Tissue Microarray 347 (YTMA-347) that was generated at the Developmental Histology Lab, (YPTS, Yale University, USA). The ER, PR and HER2 status of these tissues was determined by IHC at the Developmental Histology Lab. The clinical-pathological features of all patients utilized in these studies are reported in **Chapters 3 and Appendix Chapters 2 and 3**.

Immunohistochemistry

Prior to staining with primary antibodies, FFPE TNBC tissues were de-waxed by warming on a slide warmer at 60°C for 20 minutes, followed by de-paraffinization in xylenes 3 times for 5 minutes each, and rehydrated in a decreasing ethanol gradient (100%, 95% and 70%). Antigen retrieval was achieved by boiling slides at 98°C for 40 minutes in a low pH buffer (pH 6.0) solution or high pH buffer (pH 9.0) solution (DAKO, Glostrup, Denmark) for the Nigerian, Barbadian, AA and CA TNBC tissues used in **Appendix Chapters 1, 2 and 3**. Endogenous peroxidase activity was quenched by treatment with 3% hydrogen peroxide in 1X **p**hosphate **b**uffer **s**aline (PBS), while non-specific staining was blocked by incubating tissues in a blocking solution comprised of 1X PBS and 5% normal donkey or goat serum (depending on the species the secondary antibody was raised in) for 1 hour at room temperature. Endogenous biotin, biotin receptors, and avidin binding sites on tissues were blocked using the Avidin/Biotin blocking kit (Vector Labs, CA, USA) preceding primary antibody incubations. Tissues were then incubated overnight at 4°C with the following

antibodies: mouse anti-Kaiso monoclonal 6F (1:1000; gift from Dr. Reynolds), mouse anti-p120 monoclonal 15D2 (1:500; BD Biosciences 610133), rabbit anti-WASF3 polyclonal (1:75; Sigma-Aldrich SAB4500164), and diluted in a solution containing 1X PBS and 1% normal donkey or goat serum depending on the antibody. Tissues were subsequently washed two times for 10 minutes each in 1X PBS solution containing 0.05% Tween-20 (PBS-T) and once for 10 minutes each in 1X PBS. Secondary antibody incubations were performed at room temperature for 2 hours with biotinylated goat anti-mouse, donkey anti-mouse or donkey anti-rabbit antibody at a dilution of 1:1000. Washes were performed as described above with PBS and PBS-T. To visualize immunostaining, tissues were incubated in Vectastain (Vector Labs) for 30 minutes and diaminobenzidine (DAB) (Vector Labs) for 2 - 10 minutes (depending on time taken to achieve adequate colour development). Tissues were subsequently counterstained with Harris hematoxylin (Sigma), differentiated in acid ethanol (0.3% HCl in 70% ethanol) and blued in Scott's tap water substitute. Dehydration of tissues was achieved by incubating slides in an ascending ethanol gradient (70%, 95% and 100%). Tissues were then dried in xylenes two times for 5 minutes each and mounted using Polymount (Polysciences Inc., Warrington, PA, USA). Negative controls were attained by excluding primary antibody. Positive controls were attained by staining mouse tissues of varying organs. IHC-stained images were acquired using the Aperio Slide scanner (Leica Biosystems, ON, Canada). The TMAs were scored blindly by two pathologists, and their scores averaged to give a final mean score value that was used for all subsequent analyses in **Appendix Chapters 1, 2 and 3**. The staining intensity scores for each stained tissue was then calculated using the formula for the modified histochemical

score (H-score) system which is 3x (percentage of cells with maximum/marked intensity staining) + 2x (percentage of cells with moderate intensity staining) + 1x (percentage of cells with mild intensity staining) to give a final score value ranging from 0 – 300.

DNA Isolation

Ten microns (μM) thick FFPE tissue sections on twenty slides were scraped and placed in NAVY RINO tubes (Next Advance, Troy, NY) with stainless steel beads and 160 μL of deparaffinization solution (Qiagen, Hilden, Germany) and homogenized using the Bullet Blender (Next Advance) for 5 minutes at speed 12. The tubes were then incubated for 3 minutes at 56˚C. The deparaffinization solution + DNA was then transferred to a tube with 55 μL of RNase-free water, 25 μL **F**etch **T**arget **B**uffer (FTB) Buffer, and 20 μL Proteinase K. The solution was then incubated for 1 hour at 56˚C, shaking at 900 rpm. The tubes were then transferred to a 90˚C block for 1 hour to de-crosslink the DNA before centrifugation. The lower clear phase was then transferred to a new microcentrifuge tube with 115 μL of RNase-free water. 35 μL of **U**racil-**N**-**G**lycosylase (UNG) was added to the solution and incubated for 1 hour at 50˚C. After an hour, the solution was centrifuged followed by addition of 2 μL of RNase A and incubated at room temperature for 2 minutes. The sample solution was purified utilizing QIAamp MinElute Column (Qiagen) and eluted into Tris-HCl (10 mM).

Library Preparation

The quality and quantity of isolated DNA was measured using the Genomic DNA Screen Tape Assay (Agilent Technologies, Santa Clara, CA). The concentration of genomic DNA

(gDNA) larger than 200 bp was then calculated, and at least 200 ng of DNA greater than 200 bp in size was sheared in 50 µL of nuclease-free water with the Covaris E220 using the 96 microTUBE Plate (Covaris, Woburn, MA) with the following conditions: Step one, peak power 175, duty factor 10, cycles per burst 200, average power 17.5, and time 360 s. Then step two, peak power 2.5, duty factor 0.1, cycles per burst 50, average power 0, and time 1.0 s. After the shearing protocol, the DNA was quantified using the High Sensitivity D1000 Screen Tape Assay (Agilent Technologies). 100-500 ng of sheared DNA was placed into 50 µL of nuclease-free water and the library was prepared using the KAPA Hyper Prep Kit (Roche, Basel, Switzerland). Dual-indexed adapters were ligated to end-prepped and A-tailed DNA for 1 hour and 10 cycles of PCR were performed on the libraries. The prepared libraries were then quantified using Agilent's D1000 Screen Tape Assay. 200-1000 ng of adapter-ligated libraries was placed into the exome capture reaction, using Agilent's SureSelect XT Low-Input Reagent Kit and the V6 + CUSTOM capture library with additional probe content that included: (1) 44,000 evenly spaced probes for more informative copy number analysis, (2) probes that tile completely across 20 known tumor suppressor genes to capture potential inactivating structural events, and (3) probes that tile across 20 specific non-coding regions known to be involved in cancer translocations. Integrated DNA Technologies (IDT) xGen Universal Blocking Oligos TS i5 and i7 (Integrated DNA Technologies, Coralville, IA) were used in addition to the blocking oligo mix included in the SureSelect XT-LI reagent kit. After exome capture, the libraries were quantified using Agilent's High Sensitivity D1000 Screen Tape Assay. Each library was

normalized to 5 nM pool and sequenced on Illumina's NovaSeq 6000, using the NovaSeq S2 flow cell V1 chemistry 300 cycles kit at 150x150 (Illumina, San Diego, CA).

General Pipeline Methods

Best practices were adhered to for each tool mentioned within this section and the next section of this chapter unless otherwise stated. Whole-exome sequences were aligned by **B**urrows **W**heeler **A**lignment (BWA) (v0.7.17) to GRCh38. Quality score errors were detected by **G**enome **A**nalysis **T**ool**k**it (GATK's) Base Recalibrator (v4.0.10.1). Picard Tools (v1.128) was used to merge aligned **b**inary **a**lignment **m**aps (BAMs) and mark duplicate reads. Germline **V**ariant **C**all **F**ormat (VCF) of BAM were obtained by GATK's Haplotype Caller using GATK best practices, Samtools MPileUp together with BCFtools (v1.2), and Freebayes (v1.1.0-6-gf069ec6). Somatic variant calling was performed by MuTect2 (Cibulskis *et al.*, 2013) to ensure compatible comparison with TCGA. MuTect2 somatic variant calling files (VCF) for each patient in this study were converted to MAF files using the vcf2maf v1.6.19 tool (Kandoth *et al.*). Data from TCGA were downloaded from: https://portal.gdc.cancer.gov/projects/TCGA-BRCA. The TNBC subtypes were extracted and divided into self-reported Caucasian/European American (EA) and African American (AA) race. Each variant was validated using IGV (v2.7.2) (Robinson *et al.*, 2011). Validation included viewing each respective variant for each participant using IGV to validate that the stated variant from our variant calling pipeline was visually observed in IGV for each type of mutation except splice site mutations.

Downstream Bioinformatics Methods

Gene frequencies in our WES data were performed by Unified Optimal Sequence Kernal Association testing in R. Visualization of somatic variants was performed using maftools (v2.6.05) (Mayakonda *et al.*, 2018) and R packages pheatmap (v1.0.12), ggplot2 (v3.3.3), VennDiagram (v1.6.20), and ggrepel (v0.9.1). Mutation signature from WES data was computed using Mutational Signature in Cancer (MuSiCa) (Díaz-Gay *et al.*, 2018) and **C**atalogue **O**f **S**omatic **M**utations **I**n **C**ancer (COSMIC) (v2) release for somatic mutational signatures. Copy number analysis was performed utilizing Nexus Copy Number v10 (Biodiscovery) and focal analysis was performed by GISTIC (v2.0) (Mermel *et al.*, 2011). CNV heatmap was plotted using the oncoprint function in the R package ComplexHeatmap (v2.6.2) and pheatmap. To deduce ancestry information from tumor DNA, 1000 Genomes Project phase 3 VCF release was used as our reference population (Auton *et al.*, 2015). Data were transformed to numeric genotypes by using PLINK (v1.90b6.7). Principle component analysis (PCA) was performed using the R v3.6.0 function prcomp to establish ancestry distributions mapped by the anchor population.

Statistical Analyses

Crude incidence rate per 100,000 years was calculated by dividing the number of incident cases by the number of women in the Barbados population and multiplying by 100,000. Age-specific incidence rates were then calculated in 18 age groups with respect to the population for each age group as recorded in the census data for that time period. Standard unpaired Student's t-test with Welch's correction was used for pairwise

comparison of means. Chi square analysis was used to assess the difference in clinico-pathological features between the Nigerian, Barbadian, AA, and CA cohorts. Data are presented as mean ± SEM where applicable. For all statistical tests, p values <0.05 denote statistical significance. For the sequencing project, Type I error (α) and type II error (β) were set at 0.05 and 0.1 respectively. Chi-square test and paired student t-test, as appropriate, were used to examine bivariate association of somatic differences between two cohorts. Benjamini-Hochberg was used for multiple tests. For statistical analysis and visualization, GraphPad Prism 8 was implemented (GraphPad Software, Inc.) and Rv3.6.0 packages: circlize (0.4.6), ComplexHeatmap (1.99.7), dplyr (0.8.0.1), ggplot2 (3.1.1), ggpubr (0.2), maftools (2.0.05), plyr (1.8.4), png (0.1-7), qvalue (2.16.0), reshape2 (1.4.3), stringr (1.4.0), TCGAbiolinks (2.12.6), tidyr (0.8.3), and tools (3.6.0) were utilized. All R codes used within this thesis and supplementary data for **Chapter 5** can be found here: https://github.com/daniel-lab-mcmaster/Shawn_Rcodes

CHAPTER 3: HIGH TRIPLE NEGATIVE BREAST CANCER PREVALENCE AND AGGRESSIVE PROGSOSTIC FACTORS IN BARBADIAN WOMEN WITH BREAST CANCER

Preface

This chapter consists of the published article entitled: ***"High triple negative breast cancer prevalence and aggressive prognostic factors in Barbadian women with breast cancer"*** by Hercules SM, Hercules JC, Ansari A, Date SAJ, Skeete DHA, Smith Connell SP, Pond GR and Daniel JM. (Cancer. 2020; 126(10), 2217-2224. doi: 10.1002/cncr.32771) in its original form. This is an open-access article.

In this manuscript, we investigated BCa prevalence over the 2007-2016 period in Barbados with an emphasis on receptor status. We collected these data from pathological reports during this time period and also collected data representing non-Hispanic Black (NHB) and non-Hispanic White (NHW) women in the USA using the SEER database. We found that there was an estimated 25% TNBC prevalence in Barbados for the 2010-2016 time period, compared to a 21% and 10% prevalence reported for NHB and NHW women respectively ($p < 0.001$) for which receptor status was readily available from the SEER database. We also found that age-specific incidence rates were higher among Barbadian women (15-59) in younger age groups compared to those in the USA.

Contributions: SM Hercules wrote the manuscript, collected and analysed data represented in all figures and tables. JC Hercules, A Ansari, Dr. SAJ Date, Dr. DHA Skeete and Dr. SP Smith Connell assisted with data collection from the Queen Elizabeth Hospital.

Dr. JM Daniel and SM Hercules conceived the study and co-wrote the manuscript. Dr. JM Daniel and Dr. GR Pond also provided significant intellectual guidance throughout the study. All other co-authors assisted with editing the manuscript text. Dr. JM Daniel provided funding for the project.

Original Article

High Triple-Negative Breast Cancer Prevalence and Aggressive Prognostic Factors in Barbadian Women with Breast Cancer

Shawn M. Hercules, BSc, MPH [1,2]; Jevon C. Hercules, BSc[3]; Amna Ansari, BSc[1]; Stephanie A. J. Date, MBBS[2,4]; Desiree H. A. Skeete, MBBS[2,4,5]; Suzanne P. Smith Connell, MD[4,6]; Gregory R. Pond, PhD[7,8]; and Juliet M. Daniel, PhD[1,2]

BACKGROUND: Women of African ancestry (WAA) are disproportionately affected by triple-negative breast cancer (TNBC), which remains one of the most clinically challenging breast cancer (BCa) subtypes. This study investigated the prevalence of TNBC and epidemiological trends for BCa in Barbados, a Caribbean island with a high percentage of African ancestry. **METHODS:** Pathology reports for all BCa cases between 2007 and 2016 were collected from the sole hospital in Barbados and reviewed. The clinicopathological data collected included age, tumor grade, lymph node status, and hormone receptor status as determined by immunohistochemistry. BCa data for non-Hispanic white (NHW) and non-Hispanic black (NHB) American populations were accessed from the Surveillance, Epidemiology, and End Results database. **RESULTS:** There were 1997 BCa cases in Barbados between 2007 and 2016 for an estimated incidence rate of 135.1 per 100,000 women in Barbados (standardized to the US population, where the standardized incidence rates for NHBs and NHWs were 141.4 and 152.6 per 100,000, respectively). Age-specific incidence rates in Barbados for this period were consistently higher in younger age groups (15-59 years) in comparison with NHWs and NHBs. Between 2010 and 2016 in Barbados, a TNBC prevalence of 25% was observed, whereas TNBC prevalences of 21% and 10% were observed in NHBs and NHWs, respectively. **CONCLUSIONS:** The BCa incidence was higher in younger Barbadian women than NHWs and NHBs, and the TNBC prevalence was ~2.5 times higher than the prevalence in NHWs. This hints at a possible genetic predisposition and other socioeconomic factors that could explain the high TNBC prevalence and aggressive clinical course in WAA globally. ***Cancer* 2020;126:2217-2224**. © *2020 American Cancer Society*.

KEYWORDS: Barbados, breast cancer (BCa), epidemiology, racial disparities, triple-negative breast cancer (TNBC).

INTRODUCTION

Breast cancer (BCa) is the most frequently diagnosed cancer in women and a leading cause of cancer-related deaths in women worldwide.[1] Intriguingly, in the United States, BCa incidence rates for non-Hispanic white (NHW) women have historically been higher than those for non-Hispanic black (NHB) women, but recently, the rates have started to converge.[2-4] Nonetheless, NHB women continue to have higher mortality rates than NHW women.[2-4] NHB women are also diagnosed with BCa at younger ages and exhibit a more aggressive clinical course in comparison with NHW women.[2,5,6] The causes of these racial disparities in BCa incidence and outcomes are currently unknown, but they could in part be due to the interplay of socioeconomic status and psychosocial, environmental, and genetic factors.[7]

NHBs have a high prevalence of the very aggressive and clinically challenging BCa subtype called triple-negative breast cancer (TNBC), which is characterized by the absence of 3 therapeutically relevant BCa biomarkers: estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2).[2,8] Currently, TNBC patients are treated with a combination of nontargeted chemotherapies, radiotherapy, and surgery.[9] Although a subset of TNBC tumors are sensitive to taxane- and anthracycline-based chemotherapies,[10] 5-year survival rates for TNBC patients are still significantly lower than those for non-TNBC patients across any stage at diagnosis even after adjustments for race, age, and other covariates.[11] Because of the lack of targeted treatment options for TNBC and the high propensity of triple-negative tumors to metastasize to vital organs such as the lungs and brain,[12] TNBC is associated with lower survival in comparison with the other BCa subtypes.[13,14] This aggressive subtype has been shown to have prevalence rates ranging between 27% and 82% across

Corresponding Author: Juliet M. Daniel, PhD, Department of Biology, McMaster University, 1280 Main St W, LSB-331, Hamilton, ON, Canada L8S4K1 (danielj@mcmaster.ca).

[1]Department of Biology, McMaster University, Hamilton, Ontario, Canada; [2]African Caribbean Cancer Consortium, Philadelphia, Pennsylvania; [3]Department of Mathematics, University of the West Indies at Mona, Kingston, Jamaica; [4]Faculty of Medical Sciences, University of the West Indies at Cave Hill, Bridgetown, Barbados; [5]Department of Pathology, Queen Elizabeth Hospital, Bridgetown, Barbados; [6]Department of Radiation Oncology, Queen Elizabeth Hospital, Bridgetown, Barbados; [7]Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada; [8]Department of Oncology, McMaster University, Hamilton, Ontario, Canada

Stephanie A. J. Date is an Eastern Caribbean Health Outcomes Research Network/Yale Transdisciplinary Collaborative Center Consortium Senior Fellow.

We thank the physicians and administrative staff in the Departments of Pathology and Radiotherapy at the Queen Elizabeth Hospital (Bridgetown, Barbados) for their assistance with retrieving pathological reports and patient notes when necessary.

Additional supporting information may be found in the online version of this article.

DOI: 10.1002/cncr.32771, **Received:** October 27, 2019; **Revised:** January 17, 2020; **Accepted:** January 22, 2020, **Published online** March 10, 2020 in Wiley Online Library (wileyonlinelibrary.com)

sub-Saharan Africa[15] and as high as 17% in Puerto Rico and 14% in Guadeloupe across the Caribbean,[16] although there is a vast dearth of knowledge on TNBC prevalence in that region. These observations compel us to understand and identify the key factors contributing to the increased prevalence of this often fatal BCa subtype across various populations of women of African ancestry (WAA).

In the Caribbean, Barbados has the second highest age-standardized incidence rate and third highest mortality rate of BCa behind the Bahamas, which has the highest BCa incidence and mortality rates in the Caribbean.[16] Barbados is a small Caribbean island with an estimated population of ~280,000; approximately 92% identify as black.[17] The island has a single public hospital, Queen Elizabeth Hospital (QEH), and a publicly funded health care system.[17] Although ongoing research seeks to elucidate the causes of the racial disparities of TNBC in the United States, little research has been done on other populations with high African ancestry, such as Caribbean populations. This study sought to determine the epidemiology of BCa and the prevalence of TNBC in Barbados in comparison with the United States over a 10-year period (2007-2016). We also sought to compare clinicopathological variables of Barbadian BCa patients with and without TNBC. We hypothesized that Barbadian women would have a higher TNBC prevalence, worse disease at diagnosis, and poorer prognostic factors in comparison with NHB or NHW women in the United States.

MATERIALS AND METHODS

Study Population

Case reports for the period between January 2007 and December 2016 were collected from the Pathology Department at QEH in Barbados, and clinicopathological data were extracted from the case reports of all histologically confirmed BCa cases during this time period. A total of 1997 BCa cases were identified. The study was approved by the institutional review boards at McMaster University, the University of the West Indies at Cave Hill, and the QEH. Data for the comparators (NHB and NHW American women) were extracted from the November 2018 release of the National Cancer Institute's 2018 Surveillance, Epidemiology, and End Results (SEER) database.

BCa Subtype Classification

BCa cases were subclassified on the basis of immunohistochemistry (IHC) staining to determine the expression levels of ER, PR, and HER2 in breast tumor tissues as per routine pathology testing at QEH. The primary antibodies used at QEH were anti-ER (1D5; Dako), anti-PR (A0098; Dako), and anti-Her2 (HercepTest; Dako). To determine the estrogen and progesterone hormone receptor (HR) status in invasive breast carcinomas, the threshold for positivity was immunoreactivity in at least 1% of the tumor cell nuclei, whereas carcinomas with <1% tumor cell immunoreactivity were considered HR–. HER2 positivity was scored according to the American Society of Clinical Oncology/College of American Pathologists guidelines.[18] IHC results were obtained from pathology reports, and in lieu of molecular testing, IHC was used to subtype histologically confirmed cases. Tumors that were HR+ (ER+ and/or PR+) and HER2– were classified as luminal A, HR+/HER2+ tumors were classified as luminal B, HR–/HER2+ tumors (or HER2-enriched) were classified as HER2+, and HR–/HER2– tumors were classified as TNBC. For the US SEER data, the receptor status was available only between 2010 and 2016.

Statistical Analyses

The crude incidence rate per 100,000 years was calculated by the division of the number of incident cases by the number of women in each population and multiplication by 100,000. Age-specific incidence rates were then calculated in 18 age groups (0-4, 5-9, 10-14, 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65-59, 70-74, 75-79, 80-84, and ≥85 years) with respect to the population for each age group as recorded in the census data for that time period. Age-standardized incidence rates were calculated via the direct method. The number of BCa cases for each age group was divided by the population within that age group to give the age-specific incidence rates, which were then multiplied by the age distribution for the US population and subsequently summed. Standard unpaired Student *t* tests with Welch's correction were used for pairwise comparisons of means. Chi-square tests were used to assess the differences in clinicopathological features between luminal A, luminal B, HER2+, and TNBC groups. Data are presented as means and standard deviations where applicable. For all statistical tests, *P* values <.05 denote statistical significance, and all tests were 2-sided. Missing data are common in this patient population and other developing nations. Because missing data may have been due to systematic issues (eg, the types of surgical specimens collected [biopsies vs mastectomies] and pathological reports not being fully updated) and not due to reasons related to disease severity, we assumed that data were missing at random. We then repeated our analyses with only cases that had recorded lymph node (LN) involvement (579

of 1997 cases) and cases that had the receptor status recorded for all 3 receptors (n = 842) versus those that did not (n = 1056) to validate the significance of our findings. Consequently, all descriptive outcomes are presented with the number of cases with information available used as the denominator.

RESULTS

Barbadian Epidemiological Overview

The total estimated resident population for Barbados in 2010 was 277,821 with 144,803 females (52%).[17] A total of 1997 BCa cases were identified at the QEH between January 2007 and December 2016, and the total number of annual BCa cases ranged between 160 and 251 cases over this period. The crude incidence rate for BCa over this 10-year period was 135.1 per 100,000. The lowest (110.5 per 100,000) and highest incidence rates (171.9 per 100,000) were reported in 2015 and 2008, respectively. For the patients with available data, the mean age at diagnosis was 57.9 ± 14.2 years, with 629 patients being diagnosed with grade 2 carcinomas and 608 being diagnosed with grade 3 carcinomas (47% and 45%, respectively, of the 1342 patients with known grade information; Table 1). Three hundred and forty-one patients (59% of the 579 patients with available data) had positive LN involvement, whereas 238 patients (41%) had no LN involvement (Table 1).

TABLE 1. Breast Cancer Epidemiological and Clinicopathological Profile in Barbados (2007-2016)

Characteristic	Breast Cancer Cases (n = 1997)
Age, mean (SD), y	57.9 (14.2)
Age, No. (%)	
<50 y	578 (29)
≥50 y	1378 (71)
Missing	41
Grade, No. (%)	
1	105 (8)
2	629 (47)
3	608 (45)
Not assessed	655
Lymph node involvement, No. (%)	
Yes	341 (59)
No	238 (41)
Not assessed	1418
ER status, No. (%)	
Positive	601 (64)
Negative	332 (36)
Not available	1064
PR status, No. (%)	
Positive	384 (41)
Negative	541 (59)
Not available	1072
HER2 status, No. (%)	
Positive	138 (16)
Negative	711 (84)
Not available	1144
Breast cancer subtype, No. (%)	
Luminal A	496 (59)
Luminal B	60 (7)
HER2+	78 (9)
TNBC	209 (25)

Abbreviations: ER, estrogen receptor; HER2, human epidermal growth factor receptor 2; PR, progesterone receptor; SD, standard deviation; TNBC, triple-negative breast cancer.
Age was missing for 41 cases. Grade, lymph node involvement, ER status, PR status, and HER2 status were not assessed for 655, 1418, 1064, 1072, and 1144 cases, respectively. For HER2 status, 34 of 1144 cases were equivocal.

Barbadian WAA Have a High TNBC Prevalence and an Aggressive Clinical Course

Overall, 209 patients (25% of the 842 patients with data available for all receptors) were identified as triple-negative (ER–/PR–/HER2–). The mean age at the time of diagnosis for TNBC patients was 55.5 ± 13.1 years, whereas it was 57.9 ± 14.3 years for non-TNBC patients ($P = .029$; Fig. 1). The mean age at the time of diagnosis for luminal A patients was significantly higher (58.3 ± 14.2 years) than the mean ages of TNBC patients (55.5 ± 13.1 years) and luminal B patients (53.6 ± 14.1 years; $P = .014$; Fig. 1 and Table 2). Among 176 TNBC patients, 126 of the tumors (72%) were recorded as grade 3 carcinomas, whereas for non-TNBC patients, only 166 of the cases (39%) were grade 3 carcinomas ($P < .0001$; Fig. 1). HER2-enriched and TNBC patients had the highest percentages of grade 3 carcinomas recorded (67% and 79%, respectively; $P < .0001$). Surprisingly, although there were no significant differences in LN status after stratification by BCa subtypes, LN positivity was observed in 58% of non-TNBC patients and in 65% of TNBC patients (Fig. 2C). Further stratification by subtype revealed that luminal A, luminal B, and HER2+ enriched groups had the highest percentages of cases with LN involvement within their respective groupings. To further validate these findings, we repeated all analyses with only cases that had recorded LN involvement (579 of 1997 cases). An analysis of this smaller data set validated our findings with the larger data set (Supporting Fig. 1 and Supporting Tables 1 and 2). We also investigated clinicopathological variables between cases that had the receptor status recorded for all 3 receptors (n = 842) and cases that did not have the receptor status recorded for these receptors (n = 1056), and we found that there was no significance between the 2 groups for any of the variables examined (Supporting Fig. 2 and Supporting Table 3).

Higher TNBC Prevalence and Poorer Prognosis in Barbadian WAA in Comparison With NHBs and NHWs

According to the available data, there was an estimated BCa incidence rate of 135.1 per 100,000 women in Barbados (standardized to the US population, where

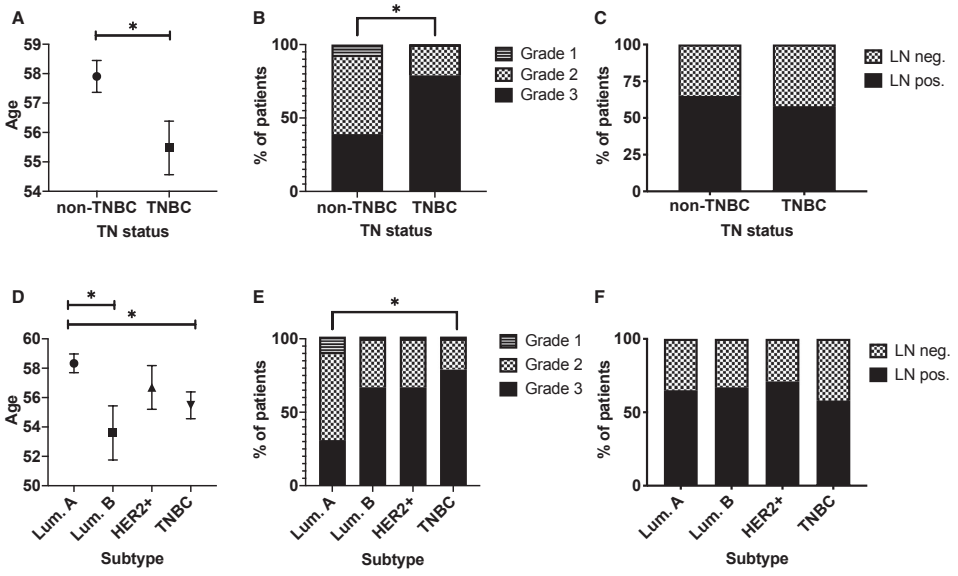


Figure 1. Biomarker status associated with clinicopathological variables. (A) TNBC patients were diagnosed at significantly younger ages than non-TNBC patients ($P < .05$). (B) TNBC patients were diagnosed with higher grade tumors than non-TNBC patients ($P < .0001$). (C) Non-TNBC patients exhibited a trend toward more LN-positive tumors than TNBC patients ($P = .24$). (D) TNBC and Lum. B patients were diagnosed at significantly younger ages than Lum. A patients ($P < .014$). (E) TNBC and HER2+ patients were diagnosed with higher grade tumors than Lum. A and Lum. B patients ($P < .0001$). (F) HER2+, Lum. A, and Lum. B patients exhibited a trend toward more LN-positive tumors than TNBC patients ($P = .567$). Error bars represent standard errors of the mean. HER2 indicates human epidermal growth factor receptor 2; LN, lymph node; Lum., luminal; TN, triple-negative; TNBC, triple-negative breast cancer. *indicates $P < .0001$ to $< .05$.

TABLE 2. Clinicopathological Differences Between Breast Cancer Subtypes in Barbadian Women

Characteristic	Missing	Cases (n = 1997)	Breast Cancer Subtypes				*P*
			Luminal A	Luminal B	HER2+	TNBC	
Age, mean (SD), y	41	57.9 (14.2)	58.3 (14.2)	53.6 (14.1)	56.7 (13.1)	55.5 (13.1)	.014
Grade, No. (%)	655						<.0001
1		105	41 (11)	2 (4)	1 (2)	3 (2)	
2		629	232 (63)	21 (40)	18 (33)	47 (27)	
3		608	93 (25)	29 (56)	36 (65)	126 (71)	
Lymph node involvement, No. (%)	1418						.788
Yes		97	60 (65)	7 (67)	11 (71)	19 (58)	
No		54	30 (35)	3 (33)	5 (29)	16 (42)	

Abbreviations: HER2, human epidermal growth factor receptor 2; SD, standard deviation; TNBC, triple-negative breast cancer.

the standardized incidence rates for NHBs and NHWs were 141.4 and 152.6 per 100,000, respectively) where Barbadian and NHB cases had higher percentages of grade 3 tumors (45% and 47%, respectively) than NHW BCa cases (30%; $P < .0001$), who had more low-grade (grade 1) tumors (24%) than NHB and Barbadian women (14% and 8%, respectively; Fig. 2B). We observed a higher percentage of Barbadian cases with LN positivity (59%) in comparison with NHB (33%) and NHW cases (24%; $P < .0001$). NHW women had a higher percentage of cases diagnosed locally to the breast with no regional LN involvement (76%) in comparison

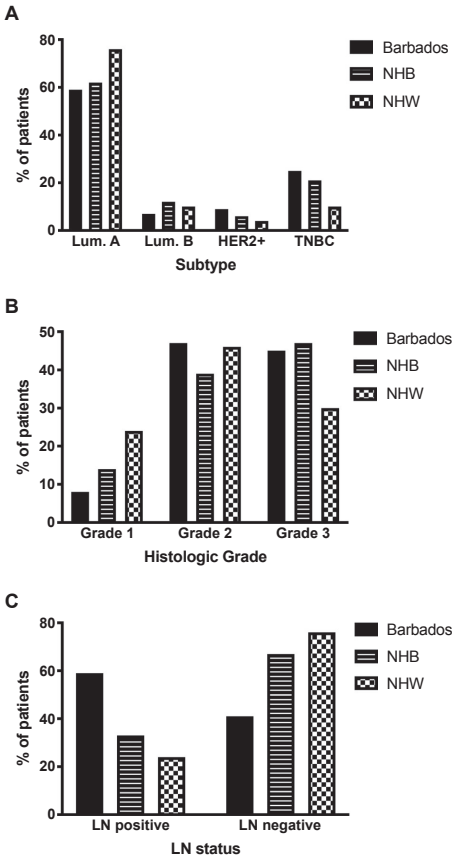


Figure 2. Barbadian and NHB women with breast cancer are diagnosed with a more aggressive clinical profile. (A) There was a higher percentage of TNBC in Barbadian and NHB women (25% and 21%, respectively) compared to NHW women (10%; $P < .0001$). (B) Higher percentages of Barbadian and NHB patients were diagnosed with grade 3 carcinoma (45% and 47%, respectively) compared to NHW patients (30%), who had a higher percentage of grade 1 carcinoma (24%) than Barbadian and NHB women (7% and 14%, respectively; $P < .0001$). (C) Barbadian and NHB patients had higher LN involvement (59% and 33%, respectively) compared to NHW women (25%), most of whom were LN-negative ($P < .0001$). HER2 indicates human epidermal growth factor receptor 2; LN, lymph node; Lum., luminal; NHB, non-Hispanic black; NHW, non-Hispanic white; TNBC, triple-negative breast cancer.

with NHB (67%) and Barbadian women (41%; Fig. 2B). Between 2010 and 2016, there was a higher prevalence of TNBC in Barbadian women (25%) in comparison with NHB (21%) and NHW American women (10%; $P < .0001$; Fig. 2C). Ranging from 1.06 to 318.8 per 100,000 women, age-specific incidence rates were higher for Barbadian women at younger ages (within the 15- to 59-year subgroups) in comparison with NHB and NHW American women for the same age group (0.29-279.1 and 0.11-281.3 per 100,000 women, respectively; Fig. 3). Similarly, the peak BCa incidence for Barbadian women occurred between the ages of 65 and 69 years, whereas the highest BCa incidence for NHW and NHB American women was recorded in the 75- to 79-year age group according to SEER data.

DISCUSSION

Although the 5-year BCa survival rates have increased globally with the advent of new therapies, diagnostics, and increased awareness, BCa is still a major public health problem globally, especially in low- to middle-income countries and small island developing states such as Barbados. The last published study on BCa incidence for this population reported a crude incidence rate of 78.6 per 100,000 between 2002 and 2006.[19] In our study, we found that there was a crude incidence rate of 135.1 per 100,000 between 2007 and 2016. This increased BCa incidence rate in Barbados is in agreement with findings from a global study estimating that age-standardized incidence rates for BCa globally increased by 12% between 2005 and 2015 with an even larger increase (~26%) in countries with a lower overall socioeconomic status.[20] Within the Barbadian context, this stark increase could also be due to increased BCa awareness and increased screening across the island.

In comparison with NHB and NHW American women, among whom incidence rates were highest at older ages (60-85+ years), we found that age-specific incidence rates were consistently higher among younger Barbadian women (across the 15 to 59-year age groups). In addition, BCa incidence in Barbadian women was highest in the 65 to 69-year age group, which is 10 years younger than that reported for both NHB and NHW American women, who exhibited the highest incidence in the 75 to 79-year age group according to data collected from the SEER data set (2007-2016 time period). Recent population-based studies from Trinidad and Tobago also found increased BCa incidence among Afro-Trinidadians within the 45 to 54-year age group,[21,22] whereas similar studies in Nigeria reported the highest BCa incidence rates within the 55 to 64-year age group.[23]

In addition to early-onset BCa, we found that Barbadian and NHB women presented with high percentages of grade 2 and 3 carcinomas and more LN

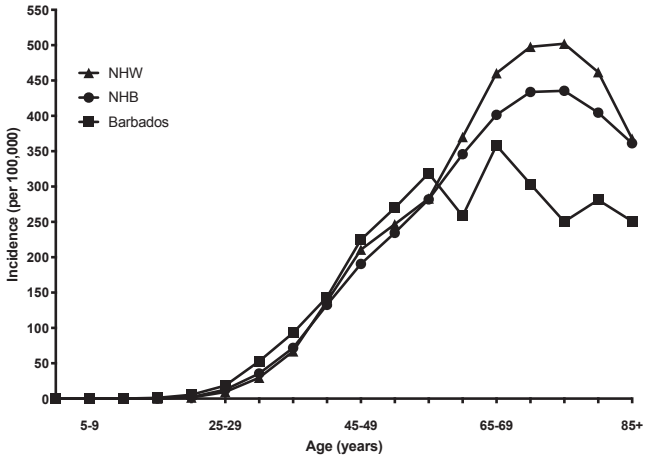


Figure 3. Age-specific breast cancer incidence rates between 2010 and 2016 for Barbados and for NHB and NHW American women. The highest breast cancer incidence rates were recorded among Barbadian women aged 54-59 and 65-69 years, whereas the highest breast cancer incidence rates in both NHB and NHW American women were recorded in the 75-79 years age group. NHB indicates non-Hispanic black; NHW, non-Hispanic white.

involvement at diagnosis in comparison with NHW women. Interestingly, although TNBC patients were diagnosed with more grade 3 tumors and a lower frequency of LN positivity than the non-TNBC patient group, further stratification of non-TNBC patients revealed that the luminal B, luminal A, and HER2+ groups had the highest percentages of LN involvement at diagnosis, and this suggested a greater likelihood of metastasis. This is consistent with the findings of Loibl and Gianni,[24] who found a correlation with HER2+ positivity and brain metastases in up to 50% of patients diagnosed. We also found that more WAA in our study cohort presented with the aggressive TNBC disease at diagnosis in comparison with NHW American women. The prevalence of TNBC in the United States has been reported previously as ~11% for NHWs and 22% for NHBs.[25] Our analysis of the SEER data set revealed similar results with rates of ~10% and 21%, respectively. The TNBC prevalence rate of ~25% for this period in Barbados is thus only slightly higher than the NHB rate of ~21%.

These younger BCa incidence rates in WAA (Barbadian and NHB women) across distinct but ancestrally-related populations are striking and hint at a possible genetic predisposition to early-onset BCa. Furthermore, an early age of BCa onset has also been associated with higher grades, high recurrence rates, contralateral BCa disease, and triple-negative and HER2-enriched subtypes.[26] This evidence strongly suggests genetic components contributing to the higher age-specific incidence rates and aggressive nature of BCa in younger Barbadian women because socioeconomic status alone is unlikely to explain clinical factors such as age at diagnosis, TNBC at diagnosis, and tumor grade. Indeed, it has also been previously shown that a family history of BCa and a history of benign breast disease increase BCa risk approximately 3 times more in cases versus controls in the Barbadian population, and this further hints at a genetic component for BCa in this population.[27]

One potentially relevant gene that has been identified as a BCa biomarker and is widely used for genetic testing is *BRCA1*. Across the African continent, researchers have identified *BRCA1* mutation rates as high as 18% in African women with BCa.[28] After analysis of a panel of 10 BCa-related genes in the genomes of 258 women, a recent study found that germline mutations for *BRCA1/2* genes were the most common in North American WAA.[29] This is in agreement with studies that identified founder mutations in *BRCA1* in ~23% and ~12.5% of Bahamian[30] and Jamaican BCa cases,[31] respectively. Another independent study using data from The Cancer Genome Atlas identified *TP53* as being more highly mutated in black women in comparison with white women with BCa.[32] They further found that CRYBB2

was more highly expressed in tumors of black women than tumors of white women, and this was further stratified by receptor status (ie, TNBC tumors expressed more CRYBB2).[32] This could be another potential candidate biomarker implicated in the aggressive nature of TNBC prevalence and outcomes in WAA. In addition to these candidates, another potentially relevant gene that has been found to be overexpressed in TNBC and aggressiveness is the transcription factor Kaiso, which has been associated with poorer overall survival and aggressive clinical outcomes in BCa and other cancers.[33] Studies from our laboratory and others have reported higher Kaiso expression in TNBC samples from WAA in comparison with white women.[34,35] Interestingly, high Kaiso expression also correlates with poor survival outcomes in African American men diagnosed with prostate cancer.[36]

Although our study provides a better understanding of the incidence of BCa and the distribution of BCa-associated clinical factors in Barbados, it was limited by the information reported in the pathology reports, many of which did not include data for LN involvement and receptor status. The lack of some data can be partially explained by the fact that LN involvement is typically reported for lumpectomy or mastectomy cases but not for core or needle biopsy cases, in which LNs are not surgically removed. For the years in which the surgery type was investigated (2013 and 2016), 47% of the reports were pending reports with no type of surgery noted, 32% noted mastectomies, and 21% noted core biopsies. The LN status was not noted in these pending reports or reports of core biopsies and was noted for only 65% of patients with larger surgical specimens such as mastectomy specimens (data not shown). Although health care is publicly funded in Barbados and breast tumor classification by IHC is standard of care in Barbados, some patients may have received BCa surgery and follow-up outside Barbados, and those clinical reports would not have been captured within this study. This would affect the number of BCa cases and the TNBC prevalence noted in the study. Notwithstanding these limitations, we have demonstrated for the first time the clinical profile of BCa cases in Barbados with a high percentage of African ancestry (estimated to be ~92%).[17] The high incidence of early-onset and premenopausal BCa in the Caribbean region warrants additional work to identify the genetic profile and relevant associations that can be leveraged for improved therapies, prognostic markers, and increased screening practices in this population.

FUNDING SUPPORT

This work was funded in part by the Canadian Breast Cancer Foundation and the Juravinski Hospital and Cancer Center Foundation.

CONFLICT OF INTEREST DISCLOSURES

Gregory R. Pond reports personal fees from Takeda and Astra-Zeneca outside the submitted work; he also has a close family member who is employed by Roche Canada and owns stock in Roche. The other authors made no disclosures.

AUTHOR CONTRIBUTIONS

Shawn M. Hercules: Study design, data collection, data analysis, data interpretation, writing of the manuscript, and revision of the manuscript. **Jevon C. Hercules:** Data collection and review and revision of the manuscript. **Amna Ansari:** Data collection and review and revision of the manuscript. **Stephanie A. J. Date:** Data collection and review and revision of the manuscript. **Desiree H. A. Skeete:** Data collection and review and revision of the manuscript. **Suzanne P. Smith Connell:** Data collection and review and revision of the manuscript. **Gregory R. Pond:** Data analysis and review and revision of the manuscript. **Juliet M. Daniel:** Study design, data collection, data interpretation, writing of the manuscript, and revision of the manuscript.

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68:394-424.
2. DeSantis CE, Fedewa SA, Goding Sauer A, Kramer JL, Smith RA, Jemal A. Breast cancer statistics, 2015: convergence of incidence rates between black and white women. *CA Cancer J Clin*. 2016;66:31-42.
3. Moller H, Henson K, Luchtenborg M, et al. Short-term breast cancer survival in relation to ethnicity, stage, grade and receptor status: national cohort study in England. *Br J Cancer*. 2016;115:1408-1415.
4. DeSantis CE, Ma J, Gaudet MM, et al. Breast cancer statistics, 2019. *CA Cancer J Clin*. 2019;69:438-451.
5. Dietze EC, Sistrunk C, Miranda-Carboni G, O'Regan R, Seewaldt VL. Triple-negative breast cancer in African-American women: disparities versus biology. *Nat Rev Cancer*. 2015;15:248-254.
6. Stark A, Kleer CG, Martin I, et al. African ancestry and higher prevalence of triple-negative breast cancer: findings from an international study. *Cancer*. 2010;116:4926-4932.
7. Williams DR, Mohammed SA, Shields AE. Understanding and effectively addressing breast cancer in African American women: unpacking the social context. *Cancer*. 2016;122:2138-2149.
8. Sung H, DeSantis CE, Fedewa SA, Kantelhardt EJ, Jemal A. Breast cancer subtypes among Eastern-African–born black women and other black women in the United States. *Cancer*. 2019;125:3401-3411.
9. Engebraaten O, Vollan HKM, Borresen-Dale AL. Triple-negative breast cancer and the need for new therapeutic targets. *Am J Pathol*. 2013;183:1064-1074.
10. Yao H, He G, Yan S, et al. Triple-negative breast cancer: is there a treatment on the horizon? *Oncotarget*. 2017;8:1913-1924.
11. Li X, Yang J, Peng L, et al. Triple-negative breast cancer has worse overall survival and cause-specific survival than non–triple-negative breast cancer. *Breast Cancer Res Treat*. 2017;161:279-287.
12. Foulkes WD, Smith IE, Reis-Filho JS. Triple-negative breast cancer. *N Engl J Med*. 2010;363:1938-1948.
13. Lin NU, Vanderplas A, Hughes ME, et al. Clinicopathologic features, patterns of recurrence, and survival among women with triple-negative breast cancer in the National Comprehensive Cancer Network. *Cancer*. 2012;118:5463-5472.
14. Kumar P, Aggarwal R. An overview of triple-negative breast cancer. *Arch Gynecol Obstet*. 2016;293:247-269.
15. Der EM, Gyasi RK, Tettey Y, et al. Triple-negative breast cancer in Ghanaian women: the Korle Bu Teaching Hospital experience. *Breast J*. 2015;21:627-633.

16. Ragin C, Banydeen R, Zhang C, et al. Breast cancer research in the Caribbean: analysis of reports from 1975 to 2017. *J Glob Oncol*. 2018;4:1-21.
17. Barbados Statistical Service. 2010 Population and Housing Census. Vol. I. Barbados Statistical Service; 2013.
18. Wolff AC, Hammond MEH, Allison KH, et al. Human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline focused update. *J Clin Oncol*. 2018;36:2105-2122.
19. Hennis AJ, Hambleton IR, Wu SY, Leske MC, Nemesure B; Barbados National Cancer Study Group. Breast cancer incidence and mortality in a Caribbean population: comparisons with African-Americans. *Int J Cancer*. 2009;124:429-433.
20. Fitzmaurice C, Allen C, Barber RM, et al. Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 32 cancer groups, 1990 to 2015: a systematic analysis for the Global Burden of Disease Study. *JAMA Oncol*. 2017;3:524-548.
21. Warner WA, Lee TY, Badal K, et al. Cancer incidence and mortality rates and trends in Trinidad and Tobago. *BMC Cancer*. 2018;18:712.
22. Warner WA, Morrison RL, Lee TY, et al. Associations among ancestry, geography and breast cancer incidence, mortality, and survival in Trinidad and Tobago. *Cancer Med*. 2015;4:1742-1753.
23. Jedy-Agba E, Curado MP, Ogunbiyi O, et al. Cancer incidence in Nigeria: a report from population-based cancer registries. *Cancer Epidemiol*. 2012;36:e271-e278.
24. Loibl S, Gianni L. HER2-positive breast cancer. *Lancet*. 2017;389: 2415-2429.
25. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2017. *CA Cancer J Clin*. 2017;67:7-30.
26. Gomez-Flores-Ramos L, Alvarez-Gomez RM, Villarreal-Garza C, Wegman-Ostrosky T, Mohar A. Breast cancer genetics in young women: what do we know? *Mutat Res*. 2017;774:33-45.
27. Nemesure B, Wu SY, Hambleton IR, Leske MC, Hennis AJ; Barbados National Cancer Study Group. Risk factors for breast cancer in a black population—the Barbados National Cancer Study. *Int J Cancer*. 2009;124:174-179.
28. Abbad A, Baba H, Dehbi H, et al. Genetics of breast cancer in African populations: a literature review. *Glob Health Epidemiol Genom*. 2018;3:e8.
29. Churpek JE, Walsh T, Zheng Y, et al. Inherited predisposition to breast cancer among African American women. *Breast Cancer Res Treat*. 2015;149:31-39.
30. Donenberg T, Lunn J, Curling D, et al. A high prevalence of BRCA1 mutations among breast cancer patients from the Bahamas. *Breast Cancer Res Treat*. 2011;125:591-596.
31. Donenberg T, Ahmed H, Royer R, et al. A survey of BRCA1, BRCA2, and PALB2 mutations in women with breast cancer in Trinidad and Tobago. *Breast Cancer Res Treat*. 2016;159:131-138.
32. Huo D, Hu H, Rhie SK, et al. Comparison of breast cancer molecular features and survival by African and European ancestry in The Cancer Genome Atlas. *JAMA Oncol*. 2017;3:1654-1662.
33. Pierre CC, Hercules SM, Yates C, Daniel JM. Dancing from bottoms up—roles of the POZ-ZF transcription factor Kaiso in cancer. *Biochim Biophys Acta Rev Cancer*. 2019;1871:64-74.
34. Bassey-Archibong BI, Kwiecien JM, Milosavljevic SB, et al. Kaiso depletion attenuates transforming growth factor-β signaling and metastatic activity of triple-negative breast cancer cells. *Oncogenesis*. 2016; 5:e208.
35. Bassey-Archibong BI, Hercules SM, Rayner LGA, et al. Kaiso is highly expressed in TNBC tissues of women of African ancestry compared to Caucasian women. *Cancer Causes Control*. 2017;28:1295-1304.
36. Jones J, Wang H, Zhou J, et al. Nuclear Kaiso indicates aggressive prostate cancers and promotes migration and invasiveness of prostate cancer cells. *Am J Pathol*. 2012;181:1836-1846.

Supplementary Figures:

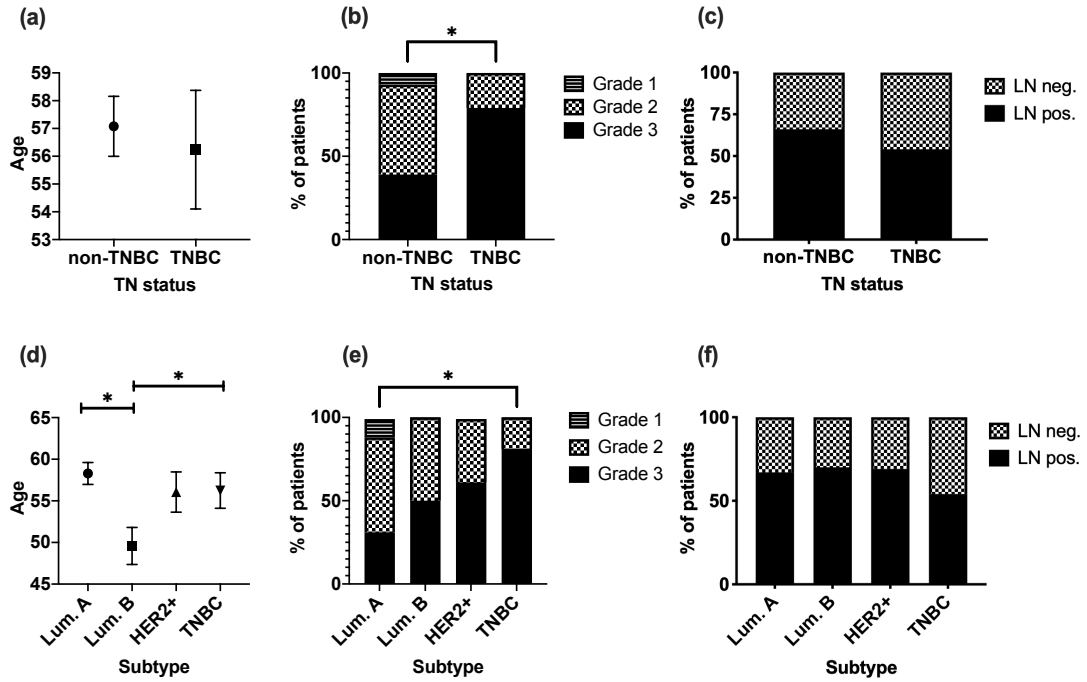


Figure S1: Biomarker status associated with clinicopathological variables. (a) No difference reported between age at diagnosis and TNBC status ($p = 0.720$). (b) TNBC patients were diagnosed with higher grade tumors than non-TNBC patients ($p < 0.0001$). (c) Non-TNBC patients exhibited a trend towards more LN-positive tumors than TNBC patients ($p = 0.453$). (d) Luminal B patients were diagnosed at significantly younger ages than Luminal A ($p < 0.05$) and TNBC patients ($p < 0.05$) (e) There was a higher trend of Lum. B, TNBC and HER2+ patients with higher grade tumors than Luminal A patients. (f) HER2+, Luminal A and Luminal B patients exhibited a trend towards more LN-positive tumors than TNBC patients ($p = 0.788$). Error bars represent standard error of the mean.

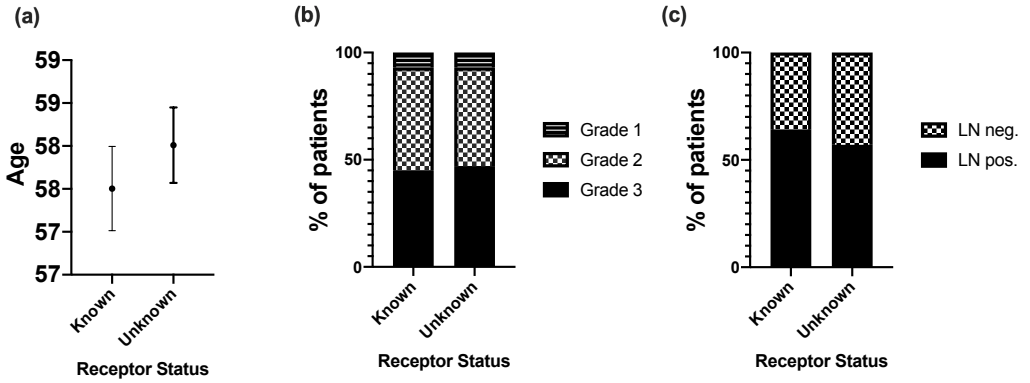


Figure S2: Known vs. unknown receptor status is not associated with clinicopathological variables. a) No difference reported between age at diagnosis and known receptor status ($p = 0.446$). (b) No difference reported between histological grade at diagnosis and known receptor status ($p = 0.320$). (c) No difference reported between lymph node status and known receptor status ($p = 0.134$). Error bars represent standard error of the mean.

Supplementary Tables:

Table S1: Breast Cancer Epidemiological and Clinicopathological Profile in Barbados (2007- 2016) with Lymph Node Status Recorded

Characteristic	**Breast cancer diagnoses** N=579, (n %)
Mean age, years (SD)	56.0 (13.1)
Age, years	
<50	184 (32)
≥50	386 (68)
Missing	9
Grade	
1	29 (7)
2	174 (44)
3	196 (49)
Not assessed	180
Lymph node involvement	
Yes	341 (59)
No	238 (41)
Not assessed	0
ER status	
Positive	117 (61)
Negative	73 (39)
Not available	389
PR status	
Positive	70 (37)
Negative	117 (63)
Not available	392
HER2 status	
Positive	29 (17)
Negative	144 (83)
Not available	406
BCa Subtypes	
Lum. A	104 (61)
Lum. B	12 (7)
HER2+	17 (10)
TNBC	38 (22)

Table S2: Clinicopathological Differences Between BCa Subtypes in Barbadian Women with Lymph Node Status Recorded

Characteristic	Breast Cancer Subtypes, n (%)						
	N Not assessed	Number of cases N= 550	Luminal A	Luminal B	HER2+	TNBC	*P*
Mean age, y (SD)	9	56.0 (12.9)	58.3 (13.4)	49.6 (7.7)	56.1 (9.9)	56.2 (13.2)	0.153
Grade	180						-
1		22	10 (11)	0 (0)	0 (0)	0 (0)	
2		165	50 (57)	4 (50)	5 (38)	7 (19)	
3		188	27 (31)	4 (50)	8 (61)	29 (81)	
Lymph node involvement	0						0.7880
Yes		97	60 (65)	7 (67)	11 (71)	19 (58)	
No		54	30 (35)	3 (33)	5 (29)	16 (42)	

Table S3: Clinicopathological Differences Between Barbadian Women with Known vs. Unknown Receptor Status

Characteristic	Breast Cancer Subtypes, n (%)		
	Known receptor status	Unknown receptor status	*P*
Mean age, y (SD)	57.5 (13.4)	58.0 (14.3)	0.446
Grade			0.320
1	47 (7)	47 (7)	
2	317 (49)	284 (45)	
3	283 (44)	302 (48)	
Lymph node involvement			0.134
Yes	109 (64)	221 (57)	
No	61 (36)	167 (43)	

CHAPTER 4: PREVALENCE OF TRIPLE NEGATIVE BREAST CANCER ACROSS AFRICA: A SYSTEMATIC REVIEW AND META-ANALYSIS

Preface

This chapter consists of the submitted manuscript entitled ***"Triple Negative Breast Cancer Prevalence in Africa: a Systematic Review and Meta-analysis"*** by Hercules SM, Alnajar M, Chen C, Mladjenovic SM, Shipeolu A, Perkovic O, Pond GR, Mbuagbaw L, Blenman KRM and Daniel JM which has been submitted to BMJ Open in its original form and is under review. In this manuscript, we investigate the distribution of, and research progress made related to TNBC across Africa. The paucity of data and information on TNBC in Africa highlighted the need of a systematic review and meta-analysis of TNBC frequency across the African continent. Our search strategy identified a total of 67 articles representing 20 countries across the continent. Our meta-analysis revealed that pooled TNBC estimates were highest across West African populations (46.0%) compared to North (24.3%), East (22.6%), Southern (16.1%) and Central African regions (14.9%). This study also revealed that there is a dearth of research on breast cancer receptor status across the continent and a lack of population-based data. Improved understanding of BCa in continental Africa can inform strategies for disease detection and management for WAA across the world.

Contributions: SM Hercules conceptualized and designed the study, conducted the literature search, data collection, data analysis, data interpretation and writing the original manuscript draft, reviewing and editing. M Alnajar conducted the literature search, data

collection, writing of the original draft, review and editing. C Chen and SM Mladjenovic conducted the literature search, data collection, data analysis and writing of the original draft, review and editing. A Shipeolu assisted with the conceptualisation, study design and writing the original draft, reviewing, and editing. O Perkovic assisted with search strategy. Dr. G Pond assisted with data analyses, data interpretation reviewing and editing the manuscript. Dr. L. Mbuagbaw assisted with data interpretation, reviewing, and editing the manuscript. Dr. KRM Blenman assisted with data interpretation, reviewing, and editing the manuscript. Dr. JM Daniel also conceptualized and designed the study, assisted with data interpretation and writing the original draft, reviewing, and editing and funding acquisition.

Triple Negative Breast Cancer Prevalence in Africa: a Systematic Review and Meta-analysis

Authors: Shawn M. Hercules, Meena Alnajar, Chen Chen, Stefan M. Mladjenovic, Ajarat Shipeolu, Olga Perkovic, Greg R. Pond, Lawrence Mbuagbaw, Kim RM Blenman and Juliet M. Daniel*

*Corresponding Author: danielj@mcmaster.ca

Shawn M. Hercules, MPH: Department of Biology, McMaster University, Hamilton, Ontario, Canada; African Caribbean Cancer Consortium, Philadelphia, Pennsylvania, United States of America. hercules@mcmaster.ca

Meena Alnajar, B.Sc.: Department of Biology, McMaster University, Hamilton, Ontario, Canada[a]; Osgoode Hall Law School, York University, Toronto, Ontario, Canada[b]. meenaalnajar2020@osgoode.yorku.ca

Chen Chen, B.Sc.: Department of Biology, McMaster University, Hamilton, Ontario, Canada[a]; Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada[b]. ccl.chen@mail.utoronto.ca

Stefan M. Mladjenovic, B.Sc.: Department of Biology, McMaster University, Hamilton, Ontario, Canada[a]; Institute of Biomedical Engineering, University of Toronto, Toronto, Ontario, Canada[b]; Terrence Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, Ontario, Canada[b]. stefan.mladjenovic@mail.utoronto.ca

Ajarat Shipeolu, B.HSc.: Department of Biology, McMaster University, Hamilton, Ontario, Canada[a]; Temerty Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada[b]. bolade.shipeolu@mail.utoronto.ca

Olga Perkovic, MLS: McMaster University Library, McMaster University, Hamilton, Ontario, Canada. perkovi@mcmaster.ca

Greg R. Pond, Ph.D., PStat: Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada; Department of Oncology, McMaster University, Hamilton, Ontario, Canada. gpond@mcmaster.ca

Lawrence Mbuagbaw, MD, MPH, Ph.D.: Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada; Biostatistics Unit, Father Sean O'Sullivan Research Centre, St Joseph's Healthcare, Hamilton, Ontario, Canada. mbuagblc@mcmaster.ca

Kim R. M. Blenman, Ph.D., M.S.: Department of Internal Medicine, Section of Medical Oncology, Yale Cancer Center, Yale University, New

Haven, Connecticut, United States of America.
kim.blenman@yale.edu

Juliet M. Daniel, Ph.D.: Department of Biology, McMaster University, Hamilton, Ontario, Canada; African Caribbean Cancer Consortium, Philadelphia, Pennsylvania, United States of America.
danielj@mcmaster.ca

[a]Previous affiliation; [b]Current affiliation(s)

Word count: 3039 words

Abstract:

Objective: To investigate frequency of triple negative breast cancer (TNBC) across continental Africa.

Study Design and Setting: For this systematic review and meta-analysis, PubMed, EMBASE, African Journals Online and Web of Science were searched for studies on TNBC frequency. All studies reporting TNBC frequencies among African women were included. Editorials, single case reports, and studies conducted in non-African nations without assessment of TNBC frequency were excluded. Screening and data abstraction were completed independently by two reviewers and a random-effects meta-analysis was performed to obtain the pooled frequency of TNBC in African populations.

Results: 1808 potentially eligible studies were identified of which 67 were included in the systematic review and 60 were included in the meta-analysis. Pooled frequency of TNBC was highest across West Africa, 46.0% [n= 15, 95% CI 38.8-53.3, I^2= 92%] and lowest in Central Africa, 14.9% [n= 1, 95% CI 8.1-23.3]. TNBC estimates were higher for studies that used Allred guidelines for ER/PR status compared to ASCO/CAP guidelines, and for studies that used older versions of ASCO/CAP guidelines for assessing HER2 status.

Conclusion: TNBC frequency was variable with the highest frequency reported in West Africa. Greater emphasis should be placed on establishing protocols for assessing receptor status due to the variability among studies.

What is new:

- Estimates for the aggressive and difficult to treat triple negative breast cancer (TNBC) are variable across the African continent where highest pooled estimates were found in West African nations (~46% TNBC frequency).
- Methods used to assess receptor status are not consistent across the African continent and TNBC estimates were higher in studies that used less stringent cut-offs.
- Poor prognostic clinical factors (*e.g.,* young age at diagnosis, high grade, and lymph node involvement) were observed in most studies done across Africa and highlights the need for early detection across the continent.

Keywords

Triple negative breast cancer, TNBC, Africa, breast cancer, women of African ancestry

AUTHORS CONTRIBUTIONS:

Shawn M. Hercules, MPH: conceptualization, study design, literature search, data collection, data analysis, data interpretation, co-writing original draft, review & editing.

Meena Alnajar, B.Sc.: literature search, data collection, data analysis, co-writing original draft, review & editing.

Chen Chen, B.Sc.: literature search, data collection, data analysis, co-writing original draft, review & editing.

Stefan M. Mladjenovic, B.Sc.: literature search, data collection, data analysis, co-writing original draft, review & editing.

Ajarat Shipeolu, B.HSc.: conceptualization, study design, co-writing original draft, review & editing.

Olga Perkovic, MLS: study design, literature search, review & editing.

Gregory R. Pond, Ph.D., PStat: data analysis, data interpretation, review & editing.

Lawrence Mbuagbaw, MD, MPH, Ph.D.: data analysis, data interpretation, review & editing.

Kim R. M. Blenman, Ph.D., M.S.: data interpretation, review & editing.

Juliet M. Daniel, Ph.D.: conceptualization, study design, data interpretation, supervision, funding acquisition, review & editing.

Funding:

This study was partially funded by the Natural Sciences and Engineering Research Council of Canada (NSERC; grant number RGPIN-2020-06822), Canadian Breast Cancer Foundation/Canadian Cancer Society Research Institute (CBCF/CCSRI; grant number 316252) and Canadian Institute for Health Research (CIHR; grant number PJT-173223) for data collection storage and trainee stipends. We have not been paid to write this article by a pharmaceutical company or any other agency.

Competing interests Disclosure:

The authors declare that they have no competing interests.

Ethics approval:

Not necessary for this manuscript.

Availability of data and materials:

The data underlying this article are available in the article and in its online supplementary material.

INTRODUCTION

Breast cancer (BCa) mortality rates have markedly increased across Africa where estimated age-standardized rates in 2020 ranged from 20 deaths per 100,000 women across Northern Africa to 27 deaths per 100,000 women in Western Africa [1-3]. BCa has thus been dubbed an emerging epidemic in Africa [4]. Women of African ancestry (WAA) across the globe and indigenous African women are more likely to receive a poorer BCa prognosis compared to women of other ancestries [5]. Multiple studies often posit poorer prognosis is a result of healthcare systems across Africa, where there is limited capacity and health infrastructure, *e.g.* inadequate screening and diagnostic services [6]. However, BCa prevalence by subtype, as defined by receptor status, is strikingly different for indigenous African women compared to the BCa prevalence profile of Western countries and is not simply explained by limited access to healthcare [5].

Immunohistochemistry (IHC) is routinely used to classify BCa into molecular subtypes according to the presence or absence of the estrogen (ER) and progesterone (PR) hormone receptors and the human epidermal growth factor receptor 2 (HER2). Triple negative breast cancer (TNBC) is characterized by the lack of expression of all three biomarkers (ER, PR, HER2), which makes TNBC untreatable with targeted therapies such as Tamoxifen and Herceptin [7, 8]. TNBC is often associated with earlier disease-onset, advanced-stage tumors and aggressive disease progression when compared with other BCa subtypes [9, 10]. TNBC has also been shown to disproportionately affect African women, and younger WAA and Hispanic women in North America [10], where the prevalence of TNBC in WAA has been estimated to be more than twice the prevalence in non-Hispanic White women [11]. WAA also have a higher mortality rate from TNBC and more advanced stage at diagnosis [12]. Thus, it is important to investigate what seems to be an ancestral predisposition to TNBC since the reasons for this racial disparity in TNBC prevalence and outcomes are not fully understood. Studies to date have not compiled adequate information on TNBC frequency or considered reported frequency and routine practices associated with diagnosis and treatment across the African continent.

This systematic review and meta-analysis aims to increase understanding and knowledge regarding the frequency of TNBC across the African continent. The paucity of data and information on TNBC in Africa underscores the importance and urgency of such a review. A previous review reported higher ER-negativity among West African countries compared to East African countries [5] and a previous meta-analysis investigated ER-positivity across Africa [13] but no review or meta-analysis has been solely focused on the frequency of TNBC cases across the African continent. This review complements current biomedical research on TNBC and provides context for areas where TNBC research continues to expand. Improved understanding of TNBC frequency in continental Africa can further inform strategies for BCa detection and management for WAA

globally. Due to shared ancestry between North American WAA and indigenous West African women [14], we hypothesize that there will be higher TNBC prevalence rates in countries across West Africa compared to other regions (North, East, Central, Southern) across Africa.

METHODS

Search Strategy and Selection Criteria

We used Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines as a framework for our systematic review and meta-analysis [15] as well as Meta-analysis of Observational Studies in Epidemiology (MOOSE) guidelines [16]. On April 25th 2021, we searched PubMed, EMBASE, African Journals Online and Web of Science for relevant articles without date or language restrictions. Start date of the search was from inception of each database. A detailed version of our search strategy used in PubMed was modified for other databases. Search strategy and these modifications can be found in Table S1. Briefly, all search terms were Medical Subject Heading terms, including TNBC terms ("TNBC", "triple-negative*", "triple negative") and terms for African countries ("Africa", "African" and names of all 54 African countries) and outcome variables ("rate*", "prevalence", "epidemiology"). We included all studies that met the inclusion criteria. The inclusion criteria were as follows: studies that use breast cancer tissue samples from indigenous African women of any age, in any care setting (clinical, community, *etc.*) and at any geographic location; sample size of eligible participants ≥ 40 (as slightly more stringent criteria since normal distribution could be assumed at n= 30 [17]); studies that demonstrate at least one of the following: report on receptor status of breast tumours including ER, PR, and HER2; any primary study including but not limited to observational studies, cohort and case-control studies, cross-sectional studies, *etc.*

We excluded: editorials, single case reports, case series, and commentaries; studies that assessed diagnostic measures and treatment options for women with TNBC in the absence of assessment of its prevalence; studies conducted in non-African nations without assessment of indigenous African TNBC rates or that of first-generation African immigrants. Study selection began by screening titles and abstracts of articles collected after employing the search strategy. The full text of these articles was then reviewed to assess inclusion. Two data abstractors independently reviewed articles at both the title/abstract and full text review stages. When there were discrepancies, a consensus was made in consultation with a third reviewer. The protocol for this review was not registered. Non-English studies were included after translation through Google Translate followed by verification of translation by a French speaker.

Quality Assessment

Studies which passed full text review (Table S2) were evaluated for risk of bias using a tool developed by Hoy and colleagues specifically intended for prevalence-studies [18]. Each study was assessed according to ten items assessing internal validity (Table S3) and assigned to have either low (score of 1) or high (score of 0) risk of bias for each question by two independent reviewers. A third reviewer mediated discrepancy and a final score per question was agreed upon. Studies were then classified based on the total score for all questions in the quality assessment tool as having a high (<=5), moderate (6-8), or low (>= 8) risk of bias.

Data Analysis

After reviewing full-text articles, a heat map was constructed with the number of studies, TNBC frequency and number of participants per country across African populations for unique studies using Google Sheets. All meta-analyses, meta-regressions and sensitivity analyses were completed using R (version 4.0.2) [19]. Using the *metaprop* package in R, we conducted a meta-analysis of

TNBC frequency among indigenous African women with breast cancer, stratified by country, region, risk of bias assessment, year of publication and the use of a validated tool for assessing receptor status. Freeman-Tukey double-arcsine transformation was used to stabilize the variances [20] and a random effects model for our meta-analyses. Pooled TNBC frequency was estimated separately per country and per region as two studies included data from more than one region. When studies investigated African and non-African participants, only data from African participants were included in meta-analyses. Heterogeneity between studies was assessed with Cochran's Q, I^2, and H statistics. Meta-regression was done to explore heterogeneity using *metareg* package. We used Egger's test to investigate publication bias and small study effects using the *metabias* package.

RESULTS

Of the 1808 records identified, 1032 remained after removing duplicates from the various databases. After screening titles and abstracts, 932 records were excluded due to irrelevance. The full text for the remaining 109 records were screened and an additional 42 studies were excluded because they did not meet eligibility criteria, leaving 67 relevant studies for inclusion (Figure 1). Our search strategy identified eligible studies from 20 countries across continental Africa. Nigeria and Tunisia had the highest number of eligible studies (8 studies each), followed by Morocco (7 studies), Algeria, Egypt, and South Africa (6 studies each), Ghana and Kenya (4 studies each) and South Africa (3 studies); there were only two studies from Uganda and only one study each from Botswana, Democratic Republic of Congo, Ethiopia, Guinea, Malawi, Mali, Mozambique, Rwanda, Senegal, and Sudan (Table S2). Five studies included data from multiple countries and regions across the continent (East and West Africa, Algeria, Egypt, Ethiopia, Ghana, Morocco, Namibia, Senegal, South Africa, and Tunisia). Five studies were translated from French to English and were all based in North Africa. Summary of clinical data can be found in Table 1 and Table S2.

Figure 1: PRISMA flowchart of all articles included in the systematic review analysis

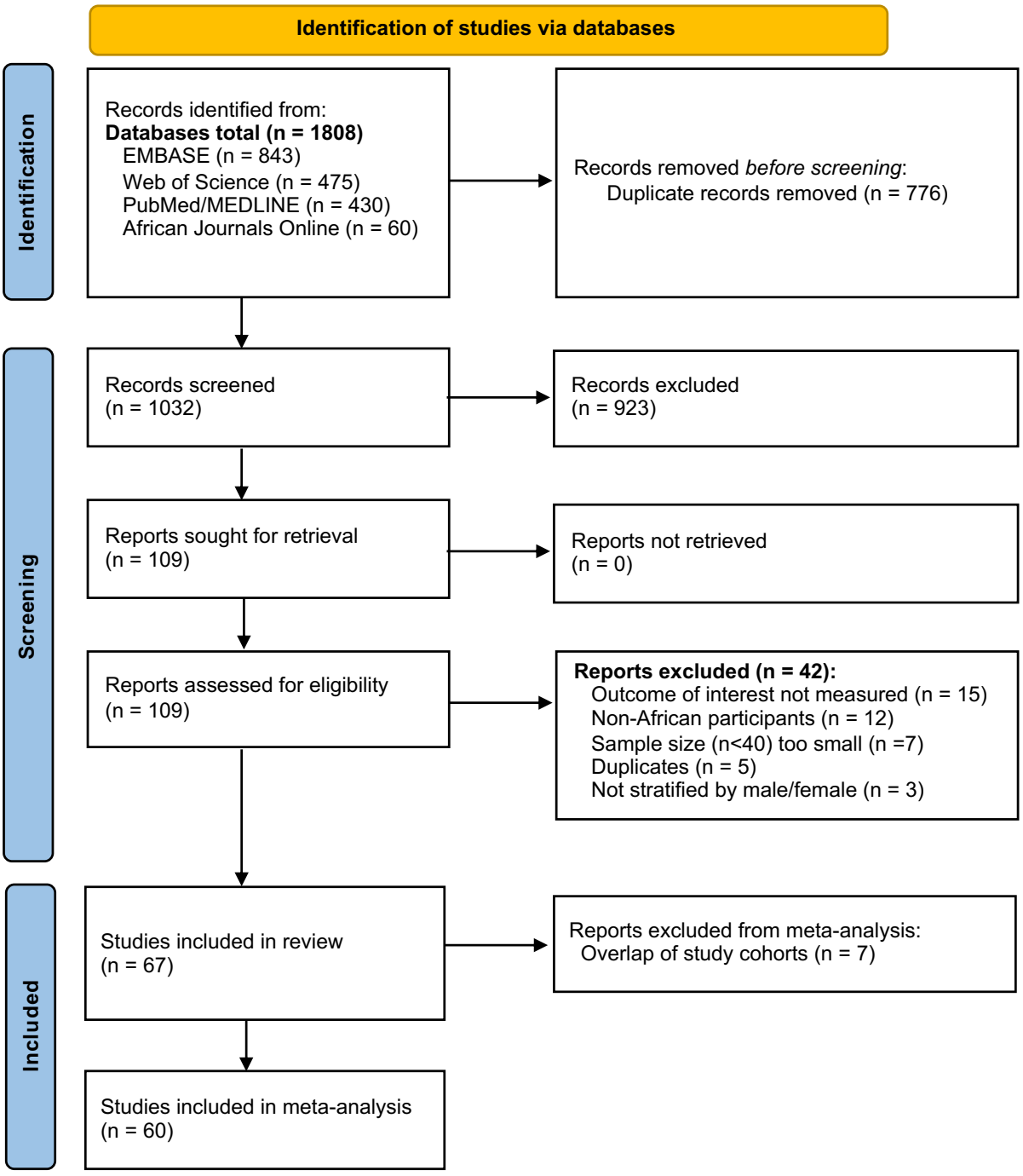


Table 1: Summary of study variables for included studies across Africa and by Region

Study variables	All studies (*n*= 67)	North African studies (*n*= 29)	East African studies (*n*= 10)	Southern African studies (*n*= 8)	Central African study (*n*= 1)	West African studies (*n*= 17)	West and East African studies (*n*= 2)
Mean age	30.62- 60.80 (*n*= 43)	30.62- 58.30 (*n*= 18)	43.00- 54.00 (*n*= 7)	53.60- 60.80 (*n*= 5)	50.00 (*n*= 1)	43.07- 52.05 (*n*= 12)	not investigated
Median age	36.00- 60.00 (*n*= 19)	36.00- 52.00 (*n*= 10)	48.00- 50.00 (*n*= 4)	54.00- 60.00 (*n*=3)	49.00 (*n*=1)	not investigated	56.00 (*n*= 1)
Min-max age recorded	13.00-105.00 (*n*= 43)	17.00-105.00 (*n*= 26)	13.00-87.00 (*n*= 7)	27.00- 96.00 (*n*= 2)	29.00- 82.00 (*n*= 1)	20.00- 96.00 (*n*= 8)	14.00- 95.00 (*n*= 1)
Grade 1 tumors	0.00%-60.64% (*n*= 53)	0.00%-22.00% (*n*= 24)	3.33%-25.00% (*n*= 8)	6.03%-60.64% (*n*= 6)	1.15% (*n*= 1)	1.50%- 27.87% (*n*= 13)	9.56% [W], 15.25% [E] (*n*= 1)
Grade 2 tumors	18.18%-82.29% (*n*= 54)	26.66%-82.29% (*n*= 24)	35.00%-51.11% (*n*= 8)	29.79%-53.22% (*n*= 7)	51.72% (*n*= 1)	18.18%-63.53% (*n*= 13)	33.52% [W], 30.81% [E] (*n*= 1)
Grade 3 tumors	9.57%-78.18% (*n*= 59)	13.94%-73.33% (*n*= 28)	31.25%-53.51% (*n*= 8)	8.40%-62.07% (*n*= 7)	47.13% (*n*= 1)	36.83%-78.18% (*n*= 14)	56.92% [W], 43.95% [E] (*n*= 1)
Positive lymph node status	12.10%-99.20% (*n*= 42)	21.40%-95.00% (*n*= 23)	49.00%-77.19% (*n*= 6)	12.10%-68.10% (*n*= 3)	not investigated	41.90%-99.20% (*n*= 10)	not investigated
Premenopausal	29.10%-73.0% (*n*= 30)	30.00%-67.50% (*n*= 13)	34.00%-64.56% (*n*= 5)	37.10% (n= 1)	60.92% (*n*= 1)	29.10%-73.00% (*n*= 7)	not investigated
TNBC frequency	5.60%-82.0% (*n*= 67)	6.00%- 66.25% (*n*= 29)	17.00%-37.70% (*n*=10)	5.60%-22.50% (*n*= 8)	14.94% (*n*= 1)	25.90%-82.20% (*n*= 17)	24.10%-32.80% (*n*= 2)

* two studies investigated both West and East African populations in their respective manuscripts, [W]- West African participants, [E]- East African participants

All 67 studies reported TNBC frequency from specific hospital/health facility settings, 34 of which were conducted via academic and academic/university teaching hospitals. Fourteen studies were prospective whereas the others (n= 53) were retrospective studies. Additionally, 20 studies included some form of biased sampling (e.g., all metastatic cases, tissue microarrays, age cut-offs) whereas the remaining studies included were either random sampling or population based. Of the included studies, eight (12%), thirty-seven (55%) and twenty-two (33%) were classified as low, moderate and high risk of bias, respectively, (Figure 2, Figure S1, Figure S2, Table S4) after using the risk of bias assessment tool for prevalence studies by Hoy *et al.* [18]. Most studies were scored as high risk of bias due to data acquisition (e.g., study population). However, according to criteria set by the Hoy *et al.* risk of bias tool, data were interpreted appropriately for most studies (e.g., having a clear definition of TNBC, appropriate numerator/denominator for frequencies). The eight studies that were low risk of bias, were based in Algeria (n= 1), Botswana (n= 1), Malawi (n= 1), Rwanda (n= 1), South Africa (n= 2) or multiple countries (n= 2).

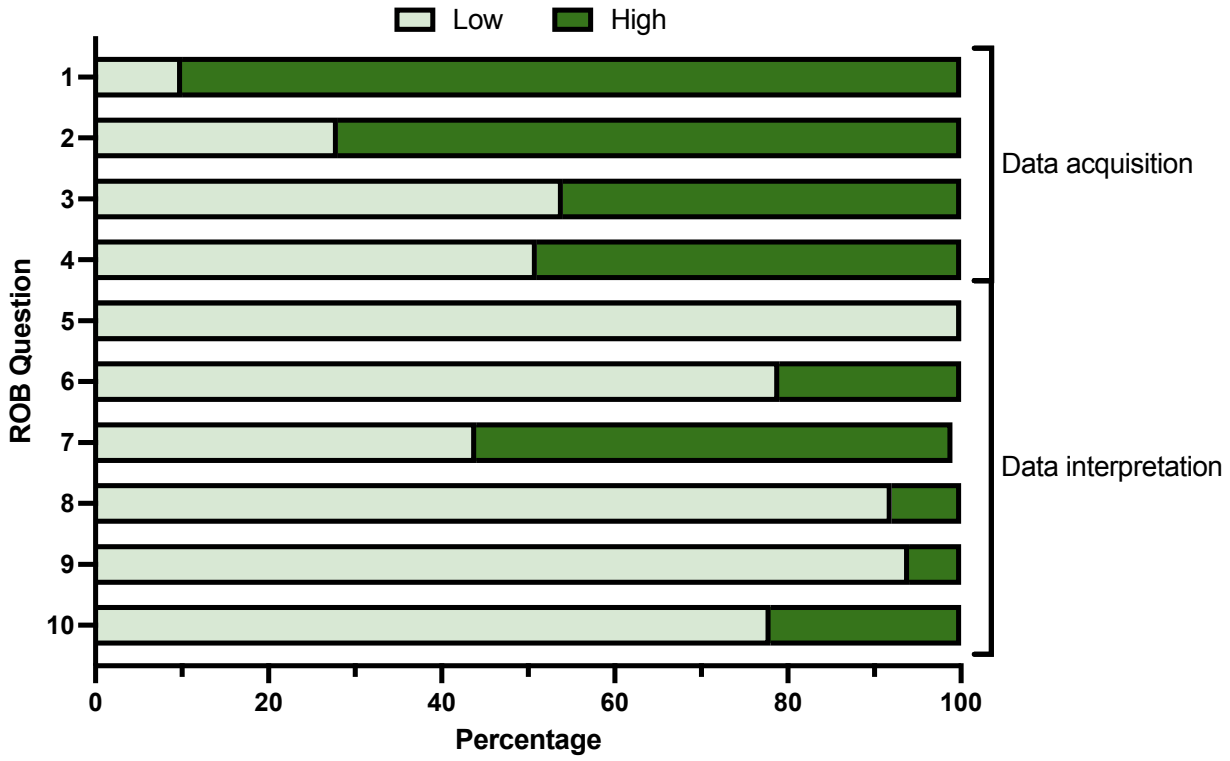


Figure 2: Risk of bias (ROB) assessment per question for all articles included in systematic review. High risk of bias was noted specifically for questions on data acquisition

After identifying unique study populations per country (n= 60, Figure 1), TNBC estimates from the meta-analysis, number of studies and participants per country were highlighted (Figure 3). Pooled TNBC frequency estimates per country (Table S5) ranged from 13.5% in South Africa (95% CI: 8.8%-19.1%, I2 = 78%) to 55.3% in Ghana (95% CI: 45.1%-65.2%, I2 = 83%). When investigating estimates per region (Figure 4), TNBC frequency was lowest in Central Africa (n=1, 14.9%; 95% CI 8.1%-23.3%) and highest in West Africa (n=15, 46.0%; 95%CI: 38.8%-29.3%, I2 = 92%). For these two analyses with pooled estimates per country and per region, heterogeneity (I2) was estimated at 94%, indicative of high between study variability. Pooled TNBC estimates were also stratified by use of validated tool for receptor status testing. Pooled TNBC frequency was higher (n= 25, 30.5%; 95%CI: 26.3%-34.9%, I2 = 95%) in studies that reported the use of a validated tool for assessing receptor expression, when compared to studies that did not (n= 35, 24.8%; 95%CI:

21.9%-27.9%, I2 = 93%) (Figure S3). Between study heterogeneity was high (I2= 94%) and meta-regression showed that these estimates were different (Figure S4A, P = 0.032, b coefficient= 0.061). When investigating TNBC estimates by the tool used for ER/PR and HER2 cut-offs, pooled TNBC frequency was higher in studies that used Allred 1998 for ER/PR cut-offs and in older versions of ASCO/CAP guidelines when compared to more recent versions for HER2 status (Figures 5, 6). There was also an association between publication year and TNBC estimates where meta-regression showed a decrease in effect size with increasing publication year (Figure S4B, $P < 0.001$, b coefficient= -0.017). When investigating the effect of risk of bias on study estimates, none was observed. Influence analyses showed that heterogeneity was not largely impacted by single studies. However, after conducting Egger's test, publication bias was identified (Figure S5, $P < 0.001$).

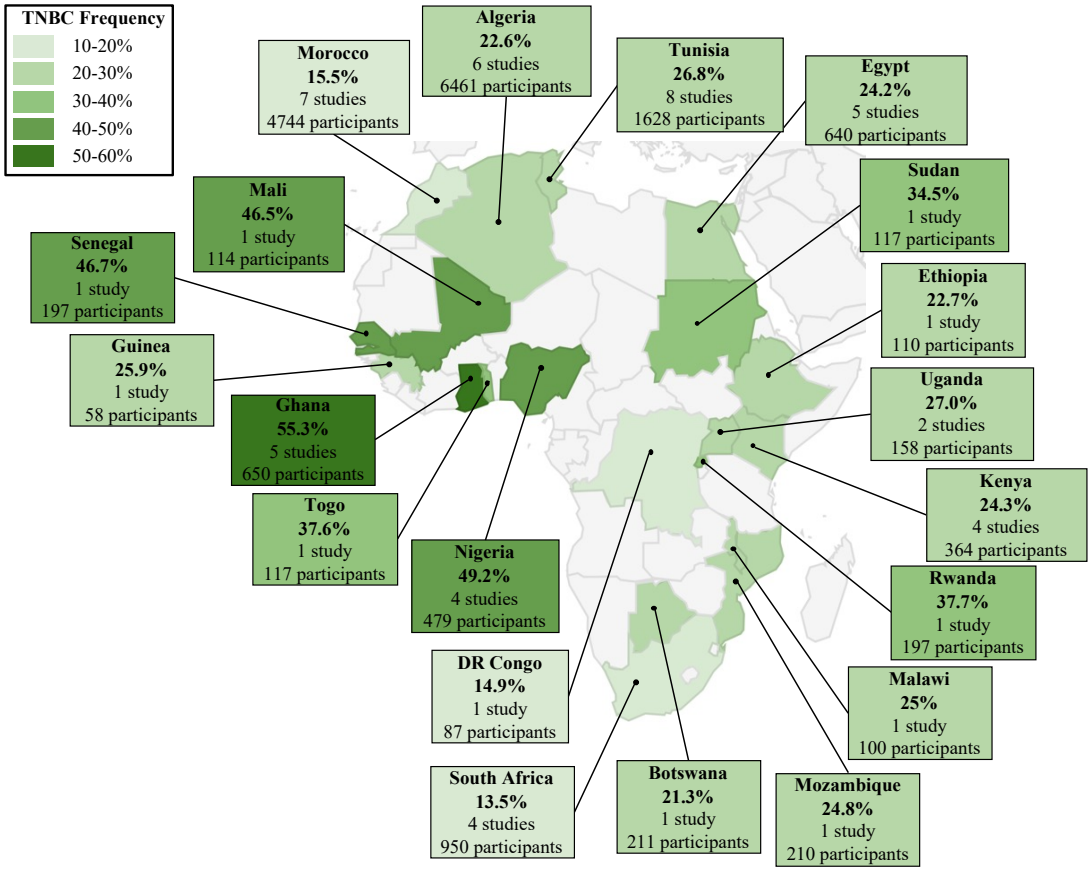


Figure 3: Pooled TNBC frequencies out of all BCa subtypes from studies done across African countries. Data represent pooled TNBC frequencies among all BCa subtypes reported in unique studies done in populations from stated countries. Pooled frequencies were calculated if the country had more than one study as stated within the meta-analysis. Estimates do not account for heterogeneity, IHC cut-offs and size of the respective populations

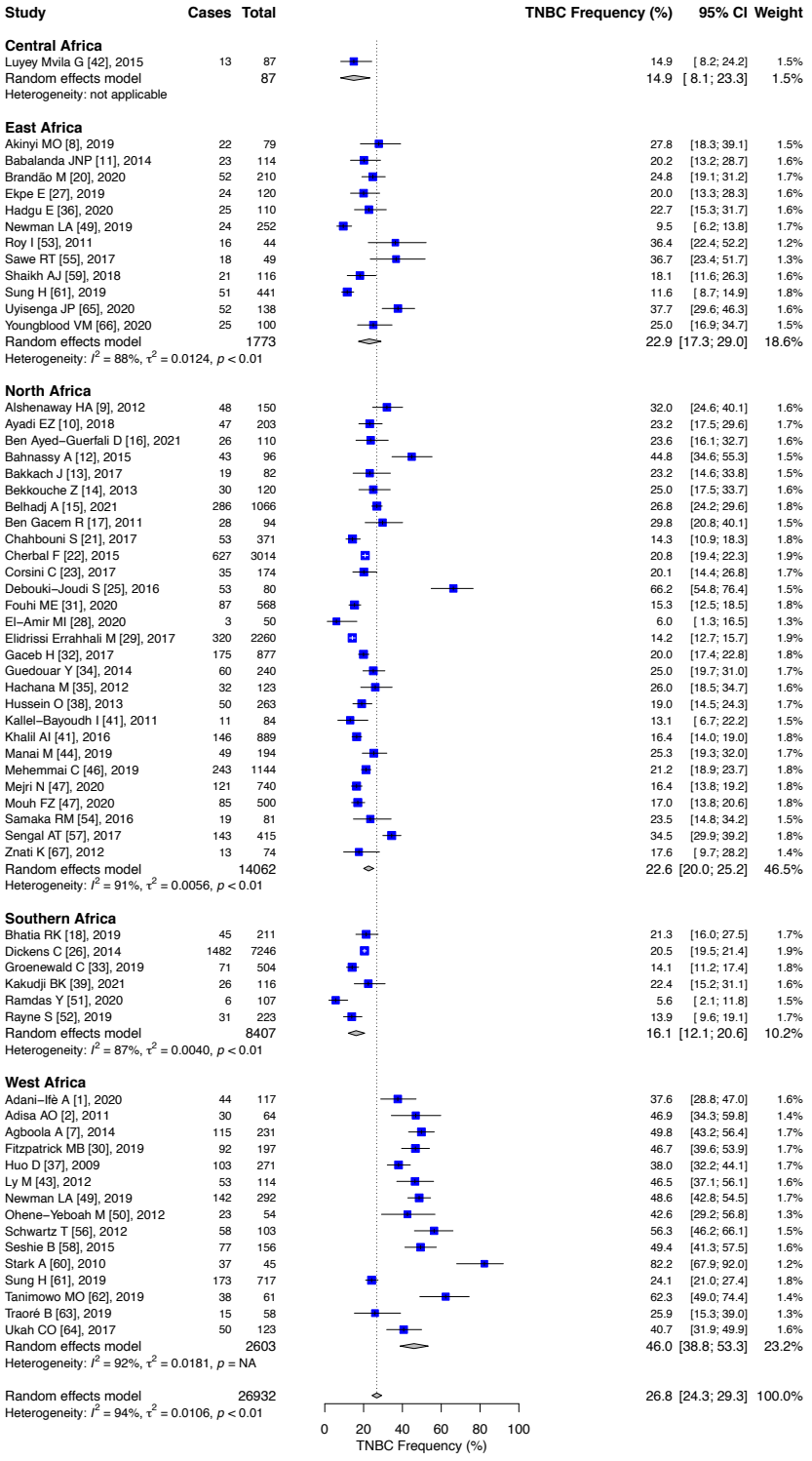


Figure 4: Pooled TNBC frequency in Africa by region. Cases are defined as participants in a study that were identified as triple negative, and total is the number of breast cancer participants with known receptor status in the study.

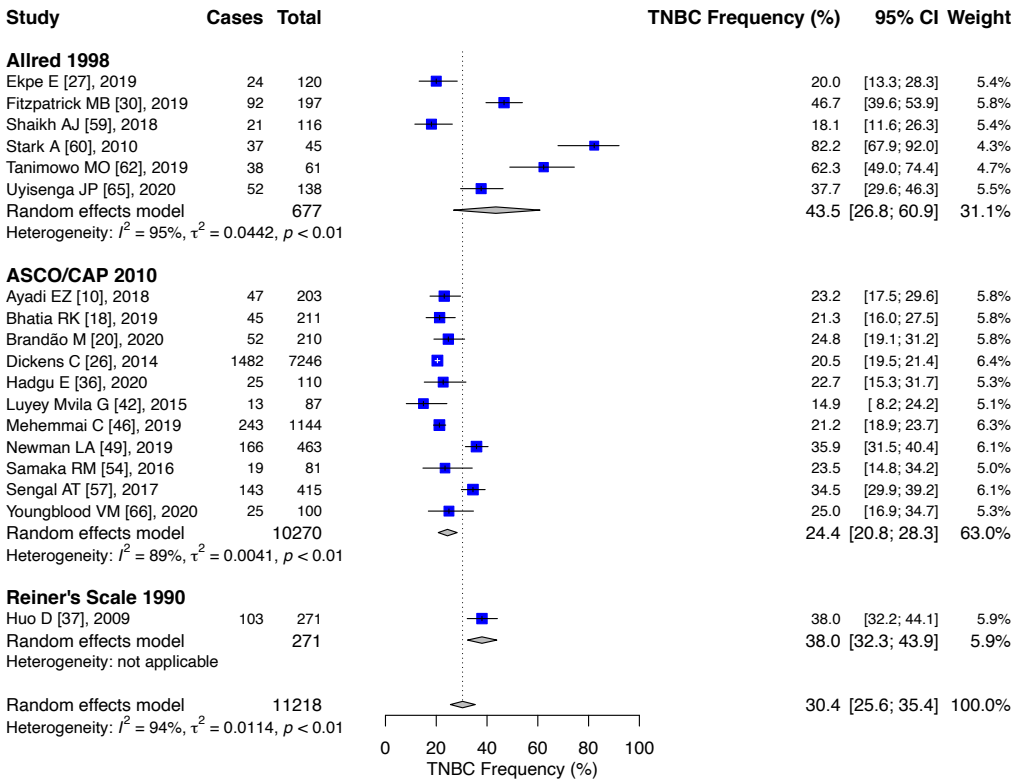


Figure 5: Pooled TNBC frequency in Africa by tool used for ER/PR status. Cases are defined as participants in a study that were identified as triple negative, and total is the number of breast cancer participants with known receptor status in the study.

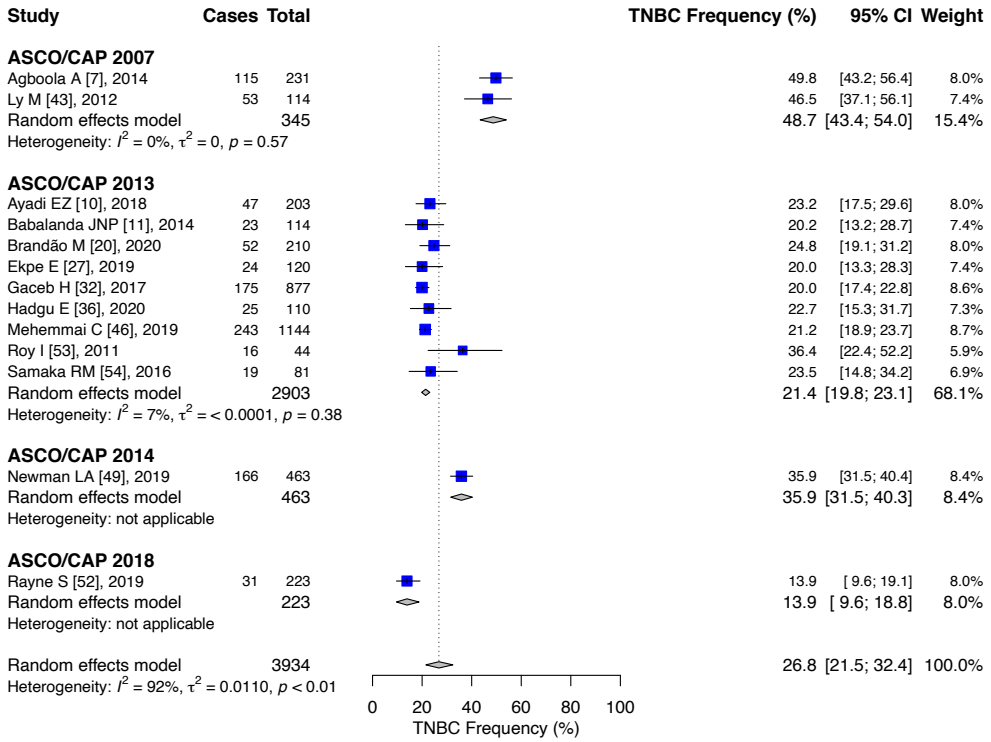


Figure 6: Pooled TNBC frequency in Africa by tool used for HER2 status. Cases are defined as participants in a study that were identified as triple negative, and total is the number of breast cancer participants with known receptor status in the study.

DISCUSSION

In this systematic review and meta-analysis of 67 studies on African women with BCa, we found that there was a high frequency of TNBC (27.1%) in cases reported across Africa although it varied depending on country and region. TNBC frequency was highest in West African populations (46.0%) compared to other regions across continental Africa (14.9%-22.9%). This is consistent with increased TNBC/ER-negative prevalence observed in populations with high West African ancestry [14, 21, 22] in the Caribbean [23, 24] and in North America [25]. In a USA population-based study (2010-2014), TNBC prevalence in non-Hispanic White women was estimated to be ~ 8% whereas it was ~15% in non-Hispanic Black women [22]. Additionally, TNBC prevalence was estimated to be ~8% and ~25% in White and Black women respectively in a UK cancer registry-based population in London [26]. These high frequencies of TNBC across Africa and the African diaspora are concerning as triple negative breast tumors have a greater propensity to metastasize to vital organs such as the brain [27] and are typically more aggressive due to lack of targeted therapies.

When investigating clinical factors, the reported mean and/or median age at diagnosis was under 50 in 35 out of the 47 studies reporting age. Young age at diagnosis (under age 40) has been previously reported to be associated with triple negative and HER2-positive cancers as well as more aggressive clinical outcome [28]. Indeed, this poor prognosis of TNBC patients was evident as most of the included studies reported a high percentage of grade 3 tumours, lymph node positivity and TNBC frequency. It must however be noted that a younger age at diagnosis is also routinely observed in lower- and middle-income countries as this is also reflective of the population structure [29]. Therefore, this observed lower age at diagnosis may be indicative of the population

distribution rather than the intrinsic aggressive biology of the tumors. To consider this possibility, we investigated associations with mean and median age at diagnosis and TNBC estimates and found no association (Figure S4C p= 0.230, β coefficient= -0.004, Figure S4D p= 0.320, β coefficient= -0.008). More advanced stage tumors and lymph node involvement at presentation may also be attributable to poor infrastructure and lack of BCa awareness and screening. A recent study of BCa across sub-Saharan Africa found that the majority of cases diagnosed were late-stage, emphasizing the need for early diagnosis [30]. Two separate studies from Nigeria and Ghana both found that most of the information obtained about BCa was from mass media and there was a general poor knowledge of BCa-associated risk factors [31, 32]. The Ghanaian-based study also found that the rate of breast self-examination (BSE), and clinical breast examination (CBE) were higher than that of obtaining mammograms [32] which emphasizes the need to promote screening programs in a culturally-relevant setting. It should be noted however that mammography has been associated with a two times higher chance of detecting ER-positive BCa compared to ER-negative BCa [33] which might be contributing to the relatively higher TNBC frequency observed across West African countries when compared to Southern African countries where mammography is more accessible.

Many studies were excluded on the account of not assessing ER, PR and HER2 status. With respect to receptor status, 30 out of the 67 studies used validated guidelines (ASCO/CAP [34-39], Allred [40] or Reiner's Scale [41]) for receptor status cut-offs. Such variability in classifying receptor status (i.e., the use of other guidelines with different cut-offs) affects the resulting treatment for patients with a diagnosis and how TNBC frequencies are calculated in each study. After stratifying studies by use of ASCO-CAP guidelines which account for specimen fixation and cut-offs for ER/PR expression at 1%, there was a decrease in pooled TNBC frequency (24.4%)

compared to those that used Allred or Reiner's Scale (cutoff at 10%) resulting in TNBC frequency of 43.5% and 38.0% respectively (Figure 5). A similar trend was observed for ASCO/CAP guidelines with respect to assessing HER2 status where more recent guidelines correlated with lower TNBC frequency compared to older guidelines (Figure 6). Thus, meta-regression with publication year was done and indeed there was an association ($P < 0.001$, β coefficient= -0.017, Figure S4). A similar trend was recently reported for East Africa- based studies conducted before and after 2013; ER/PR positivity was lower before 2013 compared to after 2013 [42]. The variability in how receptor status is assessed highlights a need for increased capacity to conduct immunohistochemical receptor status testing to further enhance BCa diagnoses and classification. It must be noted however, that IHC might not be feasible for many of the hospitals/health centers across Africa and international collaborations should be encouraged to assist with building such capacity. One Nigerian study noted cost to be a barrier – IHC was performed on only 31% of the reported cases [43]. Another Nigerian-based study stated that ASCO/CAP guidelines could not be adhered to for HER2 due to the high cost of fluorescence *in situ* hybridization in the case of an equivocal HER2 score [44]. In contrast, South Africa has an extensive healthcare system with a comprehensive standardized national public health system for routinely assessing breast cancer receptor status [45]. The disparity in access to diagnostic and therapeutic tools across the continent could also be contributing to the lack of receptor status data and higher BCa burden reported here.

To our knowledge, this is the first systematic review and meta-analysis with an in-depth analysis on TNBC frequency across continental Africa. However, there are some limitations to be considered. IHC and specimen collection and processing are not equally accessible, and neither are they uniformly done across the African continent. Specimen fixation, storage time of the samples

and other pre-analytical IHC variables have been previously shown to impact accuracy of IHC results [46]. It was estimated that up to 20% of IHC results globally are inaccurate based off of these pre-analytical variables [47]. The true frequency of TNBC is therefore not ascertainable due to lack of population-based data. Additionally, there are no validated search strategies for observational studies thus, some studies may be missed. As expected, there was considerable heterogeneity. High heterogeneity could be due to the large number of studies, or the varying ethnicities across Africa. However, given the available data, this is the best estimate of TNBC frequency.

This study provides the closest estimate of TNBC frequency across the different regions of continental Africa. Considerations should be made at the country-level to address IHC protocols and adherence to ASCO-CAP guidelines wherever possible. There is a clear disparity across the continent with respect to diagnostic and therapeutic tools that needs to be effectively addressed to prevent BCa burden. Priority should also be given to implementing culturally relevant BCa awareness programs as these have been proven to increase cancer awareness knowledge and thus could decrease preventable deaths from BCa [48]. There is also a dearth of knowledge across the continent about BCa subtype prevalence in general. This should be addressed as soon as possible by the establishment of cancer registries before the burden of BCa and other chronic diseases drastically increase with the epidemiological transition that has already started to take place across Africa.

Acknowledgements

The co-authors would like to thank Jessica Garden for reviewing and amending our French to English Google-translated documents.

REFERENCES

1. Pace, L.E. and L.N. Shulman, *Breast Cancer in Sub-Saharan Africa: Challenges and Opportunities to Reduce Mortality.* Oncologist, 2016. **21**(6): p. 739-44.
2. DeSantis, C.E., *et al.*, *International Variation in Female Breast Cancer Incidence and Mortality Rates.* Cancer Epidemiol Biomarkers Prev, 2015. **24**(10): p. 1495-506.
3. *International Agency for Research on Cancer. GLOBOCAN 2020.* April 14th, 2021]; Available from: http://globocan.iarc.fr/Default.aspx.
4. Huo, D., *et al.*, *Population differences in breast cancer: survey in indigenous African women reveals over-representation of triple-negative breast cancer.* J Clin Oncol, 2009. **27**(27): p. 4515-21.
5. Newman, L.A. and L.M. Kaljee, *Health Disparities and Triple-Negative Breast Cancer in African American Women: A Review.* JAMA Surg, 2017. **152**(5): p. 485-493.
6. Lukong, K.E., Y. Ogunbolude, and J.P. Kamdem, *Breast cancer in Africa: prevalence, treatment options, herbal medicines, and socioeconomic determinants.* Breast Cancer Res Treat, 2017. **166**(2): p. 351-365.
7. Prat, A., *et al.*, *Clinical implications of the intrinsic molecular subtypes of breast cancer.* Breast, 2015. **24 Suppl 2**: p. S26-35.
8. Bassey-Archibong, B.I., *et al.*, *Kaiso is highly expressed in TNBC tissues of women of African ancestry compared to Caucasian women.* Cancer Causes Control, 2017. **28**(11): p. 1295-1304.
9. Wright, N., *et al.*, *Panoptic Overview of Triple-Negative Breast Cancer in Nigeria: Current Challenges and Promising Global Initiatives.* J Glob Oncol, 2018. **4**: p. 1-20.
10. DeSantis, C.E., *et al.*, *Breast cancer statistics, 2019.* CA Cancer J Clin, 2019. **69**(6): p. 438-451.
11. Stark, A., *et al.*, *African ancestry and higher prevalence of triple-negative breast cancer: findings from an international study.* Cancer, 2010. **116**(21): p. 4926-32.
12. Brewster, A.M., M. Chavez-MacGregor, and P. Brown, *Epidemiology, biology, and treatment of triple-negative breast cancer in women of African ancestry.* Lancet Oncol, 2014. **15**(13): p. e625-e634.
13. Eng, A., V. McCormack, and I. dos-Santos-Silva, *Receptor-defined subtypes of breast cancer in indigenous populations in Africa: a systematic review and meta-analysis.* PLoS Med, 2014. **11**(9): p. e1001720.
14. Zakharia, F., *et al.*, *Characterizing the admixed African ancestry of African Americans.* Genome Biol, 2009. **10**(12): p. R141.

15. Moher, D., *et al.*, *Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement.* Ann Intern Med, 2009. **151**(4): p. 264-9, W64.

16. Stroup, D.F., *et al.*, *Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group.* JAMA, 2000. **283**(15): p. 2008-12.

17. Krithikadatta, J., *Normal distribution.* J Conserv Dent, 2014. **17**(1): p. 96-7.

18. Hoy, D., *et al.*, *Assessing risk of bias in prevalence studies: modification of an existing tool and evidence of interrater agreement.* J Clin Epidemiol, 2012. **65**(9): p. 934-9.

19. R Core Team, *R: A language and environment for statistical computing. R Foundation for Statistical Computing.* 2020.

20. Borges Migliavaca, C., *et al.*, *How are systematic reviews of prevalence conducted? A methodological study.* BMC Med Res Methodol, 2020. **20**(1): p. 96.

21. Moreno-Estrada, A., *et al.*, *Reconstructing the population genetic history of the Caribbean.* PLoS Genet, 2013. **9**(11): p. e1003925.

22. Scott, L.C., *et al.*, *Update on triple-negative breast cancer disparities for the United States: A population-based study from the United States Cancer Statistics database, 2010 through 2014.* Cancer, 2019. **125**(19): p. 3412-3417.

23. Hercules, S.M., *et al.*, *High triple-negative breast cancer prevalence and aggressive prognostic factors in Barbadian women with breast cancer.* Cancer, 2020. **126**(10): p. 2217-2224.

24. Ragin, C., *et al.*, *Breast Cancer Research in the Caribbean: Analysis of Reports From 1975 to 2017.* J Glob Oncol, 2018. **4**: p. 1-21.

25. DeSantis, C.E., *et al.*, *Breast cancer statistics, 2017, racial disparity in mortality by state.* CA Cancer J Clin, 2017. **67**(6): p. 439-448.

26. Jack, R.H., *et al.*, *Differences in breast cancer hormone receptor status in ethnic groups: a London population.* Eur J Cancer, 2013. **49**(3): p. 696-702.

27. Jin, J., *et al.*, *Incidence, pattern and prognosis of brain metastases in patients with metastatic triple negative breast cancer.* BMC Cancer, 2018. **18**(1): p. 446.

28. Gómez-Flores-Ramos, L., *et al.*, *Breast cancer genetics in young women: What do we know?* Mutat Res, 2017. **774**: p. 33-45.

29. Bidoli, E., *et al.*, *Worldwide Age at Onset of Female Breast Cancer: A 25-Year Population-Based Cancer Registry Study.* Sci Rep, 2019. **9**(1): p. 14111.

30. Jedy-Agba, E., *et al.*, *Stage at diagnosis of breast cancer in sub-Saharan Africa: a systematic review and meta-analysis.* Lancet Glob Health, 2016. **4**(12): p. e923-e935.

31. Azubuike, S., *Breast cancer risk factors and signs: How much do Nigerian women know?* International Journal of Advanced Medical and Health Research, 2017. **4**(1): p. 40-43.

32. Opoku, S.Y., M. Benwell, and J. Yarney, *Knowledge, attitudes, beliefs, behaviour and breast cancer screening practices in Ghana, West Africa.* Pan Afr Med J, 2012. **11**: p. 28.

33. Howlader, N., *et al.*, *US incidence of breast cancer subtypes defined by joint hormone receptor and HER2 status.* J Natl Cancer Inst, 2014. **106**(5).

34. Allison, K.H., *et al.*, *Estrogen and Progesterone Receptor Testing in Breast Cancer: ASCO/CAP Guideline Update.* J Clin Oncol, 2020. **38**(12): p. 1346-1366.

35. Wolff, A.C., *et al.*, *Human Epidermal Growth Factor Receptor 2 Testing in Breast Cancer: American Society of Clinical Oncology/College of American Pathologists Clinical Practice Guideline Focused Update.* J Clin Oncol, 2018. **36**(20): p. 2105-2122.

36. Wolff, A.C., *et al.*, *Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update.* J Clin Oncol, 2013. **31**(31): p. 3997-4013.

37. Wolff, A.C., *et al.*, *Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update.* Arch Pathol Lab Med, 2014. **138**(2): p. 241-56.

38. Hammond, M.E., *et al.*, *American society of clinical oncology/college of american pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer.* J Oncol Pract, 2010. **6**(4): p. 195-7.

39. Wolff, A.C., *et al.*, *American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer.* J Clin Oncol, 2007. **25**(1): p. 118-45.

40. Allred, D.C., *et al.*, *Prognostic and predictive factors in breast cancer by immunohistochemical analysis.* Mod Pathol, 1998. **11**(2): p. 155-68.

41. Reiner, A., *et al.*, *Immunocytochemical localization of estrogen and progesterone receptor and prognosis in human primary breast cancer.* Cancer Res, 1990. **50**(21): p. 7057-61.

42. Popli, P., *et al.*, *Receptor-Defined Breast Cancer in Five East African Countries and Its Implications for Treatment: Systematic Review and Meta-Analysis.* JCO Glob Oncol, 2021. **7**: p. 289-301.

43. Ukah, C.O., *et al.*, *The Immunohistochemical Profile of Breast Cancer in Indigenous Women of Southeast Nigeria.*

44. CO., U., *et al.*, *The Immunohistochemical Profile of Breast Cancer in Indigenous Women of Southeast Nigeria* Annals of Medical and Health Sciences Research, 2017. **7**(2): p. 83-87.

45. Cubasch, H., *et al.*, *Breast cancer survival in Soweto, Johannesburg, South Africa: A receptor-defined cohort of women diagnosed from 2009 to 11.* Cancer Epidemiol, 2018. **52**: p. 120-127.

46. Gown, A.M., *Diagnostic Immunohistochemistry: What Can Go Wrong and How to Prevent It.* Arch Pathol Lab Med, 2016. **140**(9): p. 893-8.

47. Agrawal, L., *et al.*, *Understanding preanalytical variables and their effects on clinical biomarkers of oncology and immunotherapy.* Semin Cancer Biol, 2018. **52**(Pt 2): p. 26-38.

48. Jemal, A. and O.W. Brawley, *Increasing cancer awareness and prevention in Africa.* Ecancermedicalscience, 2019. **13**: p. 939.

SUPPORTING DOCUMENTS

Table S1: Search strategy used for Databases

Database	Search terms
Pubmed/MEDLINE	(Africa OR African OR "West* Africa" OR "East* Africa" OR "South* Africa" OR "North* Africa" OR "Central Africa" OR Algeria OR Angola OR Benin OR Botswana OR "Burkina Faso" OR Burundi OR Cabo Verde OR Cameroon OR Central African Republic OR "CAR" OR Chad OR Comoros OR "DRC" OR Congo OR "Republic of Congo" OR "Ivory Coast" OR Cote d'Ivoire OR Djibouti OR Egypt OR "Equatorial Guinea" OR Eritrea OR Eswatini OR Swaziland OR Ethiopia OR Gabon OR Gambia OR Ghana OR Guinea OR "Guinea-Bissau" OR Kenya OR Lesotho OR Liberia OR Libya OR Madagascar [PL] OR Malawi OR Mali OR Mauritania OR Mauritanie OR Mauritius OR Morocco OR Mozambique OR Namibia OR Niger OR Nigeria OR Rwanda OR "Sao Tome and Principe" OR Senegal OR Seychelles OR "Sierra Leone" OR Somalia OR "South Africa" OR "South Sudan" OR Sudan OR Tanzania OR Togo OR Tunisia OR Uganda OR Zambia OR Zimbabwe) AND (TNBC OR "triple-negative*" OR "triple negative*" OR "triple negative") AND (rate* OR prevalence OR incidence OR epidemiology)
Web of Science	TS = (Africa or African or "West* Africa" or "East* Africa" or "South* Africa" or "North* Africa" or "Central Africa" or Algeria or Angola OR Benin OR Botswana OR "Burkina Faso" OR Burundi OR Cabo Verde OR Cameroon OR Central African Republic OR "CAR" OR Chad OR Comoros OR "DRC" OR Congo OR "Republic of Congo" OR "Ivory Coast" OR Cote d'Ivoire OR Djibouti OR Egypt OR "Equatorial Guinea" OR Eritrea OR Eswatini OR Swaziland OR Ethiopia OR Gabon OR Gambia OR Ghana OR Guinea OR "Guinea-Bissau" OR Kenya OR Lesotho OR Liberia OR Libya OR Madagascar [PL] OR Malawi OR Mali OR Mauritania OR Mauritanie OR Mauritius OR Morocco OR Mozambique OR Namibia OR Niger OR Nigeria OR Rwanda OR "Sao Tome and Principe" OR Senegal OR Seychelles OR "Sierra Leone" OR Somalia OR "South Africa" OR "South Sudan" OR Sudan OR Tanzania OR Togo OR Tunisia OR Uganda OR Zambia OR Zimbabwe) AND TS= (TNBC OR "triple-negative*" OR "triple negative*" OR "triple negative") AND TS= (rate* OR prevalence OR incidence OR epidemiology)
Embase	(Africa OR African OR "West* Africa" OR "East* Africa" OR "South* Africa" OR "North* Africa" OR "Central Africa" OR Algeria OR Angola OR Benin OR Botswana OR "Burkina Faso" OR Burundi OR Cabo Verde OR Cameroon OR Central African Republic OR "CAR" OR Chad OR Comoros OR "DRC" OR Congo OR "Republic of Congo" OR "Ivory Coast" OR Cote d'Ivoire OR Djibouti OR Egypt OR "Equatorial Guinea" OR Eritrea OR Eswatini OR Swaziland OR Ethiopia OR Gabon OR Gambia OR Ghana OR Guinea OR "Guinea-Bissau" OR Kenya OR Lesotho OR Liberia OR Libya OR Madagascar OR Malawi OR Mali OR Mauritania OR Mauritanie OR Mauritius OR Morocco OR Mozambique OR Namibia OR Niger OR Nigeria OR Rwanda OR "Sao Tome and Principe" OR Senegal OR Seychelles OR "Sierra Leone" OR Somalia OR "South Africa" OR "South Sudan" OR Sudan OR Tanzania OR Togo OR Tunisia OR Uganda OR Zambia OR Zimbabwe) AND (TNBC OR "triple-negative*" OR "triple negative*" OR "triple negative") AND (rate* OR prevalence OR incidence OR epidemiology)
African Journals Online (AJOL)	(TNBC OR "triple-negative*" OR "triple negative*" OR "triple negative") AND (rate* OR prevalence OR incidence OR epidemiology)

Table S2: Summary of data from included studies

Author	Population	Region	Mean age ± SD	Median age (min-max)	Grade 3 tumors	Positive Nodes	Number of TNBC cases	Participants studied with known receptor status (all participants)	Frequency of TNBC	Risk of Bias Assessment
Adani-Ifè A, 2020 [1]	Togo	West Africa	52.05 ±12.38	(30-85)	42.70%	41.90%	44	117 (312)	37.60%	Moderate
Adisa AO, 2011 [2]	Nigeria	West Africa	45.90	(20-81)	/	/	30	64 (385)	46.88%	High
Agboola A, 2014 [3]	Nigeria	West Africa	/	/	35.70%	99.20%	99	155 (199)	54.36%	High
Agboola A, 2014 [4]	Nigeria	West Africa	/	/	/	94.20%	55	100 (308)	55.00%	Moderate
Agboola A, 2013 [5]	Nigeria	West Africa	/	/	/	92.70%	81	149 (308)	63.87%	High
Agboola A, 2016 [6]	Nigeria	West Africa	/	/	34.12%	91.76%	114	187 (255)	60.96%	Moderate
Agboola A, 2014 [7]	Nigeria	West Africa	/	/	38.50%	92.20%	115	231 (231)	49.78%	High
Akinyi MO, 2019 [8]	Kenya	East Africa	48.00 ±14.5	/	/	71%	22	79 (79)	27.85%	High
Alshenawy HA, 2012 [9]	Egypt	North Africa	58.30 ±10.1	/	20.00%	/	48	150 (150)	26.67%	Moderate
Ayadi EZ, 2018 [10]	Tunisia	North Africa	35.00 ±4.2	(17-40)	46.85%	64.10%	47	203 (238)	23.15%	Moderate
Babalanda JNP, 2014 [11]	Uganda	East Africa	47	(13-87)	53.51%	77.19%	23	114 (114)	20.18%	Moderate

Bahnassy A, 2015 [12]	Egypt	North Africa	TNBC: 51.91± 12.34 non-TNBC: 52.77 ± 12.13 Controls: 35 ± 13.94	TNBC: (30-78) non-TNBC: (27-81) Controls: (22-64)	15.63%	72.91%	43	96 (96)	44.79%	Moderate
Bakkach J, 2017 [13]	Morocco	North Africa	/	36 (/)	40.20%	58.50%	19	82 (331)	23.00%	High
Bekkouche Z, 2013 [14]	Algeria	North Africa	49.72	(30-84)	73.33%	35.83%	30	120 (120)	25.00%	High
Belhadj A, 2021 [15]	Algeria	North Africa	/	50 (21-88)	38.86%	64.70%	286	1066 (1140)	27.00%	Moderate
Ben Ayed-Guerfali D, 2021 [16]	Tunisia	North Africa	/	42.78 (27-65)	26.78%	59.70%	26	110 (134)	23.60%	High
Ben Gacem R, 2011 [17]	Tunisia	North Africa	48.70	47 (31-75)	43.62%	55.26%	28	94 (94)	29.79%	Moderate
Bhatia RK, 2019 [18]	Botswana	Southern Africa	/	54 (27-96)	25.76%	/	45	211 (384)	21.30%	Low
Bogan D, 2017 [19]	Egypt	North Africa	52.38 ±12.17	(27-81)	13.94%	66.67%	43	96 (96)	44.79%	Moderate
Brandão M, 2020 [20]	Mozambique	East Africa	/	48 (/)	31.25%	61.43%	52	210 (262)	25.00%	Moderate
Chahbouni S, 2017 [21]	Morocco	North Africa	48.00	(22-99)	/	/	53	371 (390)	13.59%	High
Cherbal F, 2015 [22]	Algeria	North Africa	/	48.5 (22-84)	31.30%	/	627	3014 (3014)	20.80%	Moderate
Corsini C, 2017 [23]	Multiple (France, Algeria, Egypt, Morocco, Tunisia)	North Africa	/	/	44.74%	42%	35	174 (456)	20.10%	Moderate
Cubasch H, 2018 [24]	South Africa	Southern Africa	54.40 ±14.2	/	39.53%	/	126	581 (602)	21.70%	Low

Debouki-Joudi S, 2016 [25]	Tunisia	North Africa	49.30	(24-80)	26.22%	55.20%	53	80 (91)	66.25%	High
Dickens C, 2014 [26]	Multiple (South Africa & Namibia)	Southern Africa	/	/	33.87%	/	1482	7246 (12361)	20.45%	Low
Ekpe E, 2019 [27]	Kenya	East Africa	47.00 ±12	(21-79)	50.00%	/	24	120 (125)	20.00%	Moderate
El-Amir MI, 2020 [28]	Egypt	North Africa	54.20 ±10.6	(32-77)	18.00%	72%	3	50 (50)	6.00%	High
Elidrissi Errahhali M, 2017 [29]	Morocco	North Africa	48.70	(19-105)	18.00%	64.80%	320	2260 (2406)	14.20%	Moderate
Fitzpatrick MB, 2019 [30]	Senegal	West Africa	47.00 ±12.7	/	76.17%	/	92	197 (522)	46.70%	Moderate
Fouhi ME, 2020 [31]	Morocco	North Africa	51.60	(23-89)	32.60%	35.70%	87	568 (668)	15.30%	High
Gaceb H, 2017 [32]	Algeria	North Africa	48.44	(22-84)	28.50%	/	175	877 (877)	19.95%	Low
Groenewald C, 2019 [33]	South Africa	Southern Africa	53.60 ±13.4	/	/	68.10%	71	504 (505)	14.10%	Moderate
Guedouar Y, 2014 [34]	Algeria	North Africa	47.00	(23-78)	24.58%	38.33%	60	240 (240)	25.00%	High
Hachana M, 2012 [35]	Tunisia	North Africa	/	/	40.18%	21.40%	32	123 (123)	26.02%	Moderate
Hadgu E, 2020 [36]	Ethiopia	East Africa	43.00 ±14	50 (22-75)	50.00%	/	25	110 (112)	23.00%	Moderate
Huo D, 2009 [37]	Multiple (Nigeria & Senegal)	West Africa	44.80	/	39.15%	72%	103	271 (378)	27.00%	Low
Hussein O, 2013 [38]	Egypt	North Africa	53.00 ±11.6	52	17.39%	/	50	263 (263)	19.00%	High

Kakudji BK, 2021 [39]	South Africa	Southern Africa	56.35 ±14.09	/	62.07%	64.66%	26	116 (116)	22.50%	High
Kallel-Bayoudh I, 2011 [40]	Tunisia	North Africa	50.00 ±11	(27-75)	35.44%	65%	11	84 (84)	13.00%	High
Khalil AI, 2016 [41]	Morocco	North Africa	50.20 ±11.34	(17-93)	35.48%	49.40%	146	889 (1277)	16.40%	Moderate
Luyey Mvila G, 2015 [42]	Democratic Republic of Congo	Central Africa	50.00	49 (29-82)	47.13%	/	13	87 (87)	14.94%	Moderate
Ly M, 2012 [43]	Mali	West Africa	46.00	(25-82)	78.18%	91%	53	114 (114)	46.00%	Moderate
Manai M, 2019 [44]	Tunisia	North Africa	/	42 (24-62)	42.20%	95%	49	194 (210)	25.00%	Moderate
McCormack VA, 2013 [45]	South Africa	Southern Africa	55.30	/	42.34%	/	209	1027 (1247)	20.40%	Low
Mehemmai C, 2019 [46]	Algeria	North Africa	/	48.5 (21-84)	26.99%	/	243	1144 (1162)	21.24%	Moderate
Mejri N, 2020 [47]	Tunisia	North Africa	/	48.5 IBCa 49 non-IBCa (/)	46.20% IBCa 36.10% non-IBCa	50%	121	740 (740)	16.40%	High
Mouh FZ, 2020 [48]	Morocco	North Africa	/	47 (/)	34.92%	60.10%	85	500 (500)	17.00%	High
Newman LA, 2019 [49]	Multiple (Ghana, Ethiopia, United States)	West and East Africa	/	56 (14-95)	/	/	Ghanaian: 142 Ethiopian: 24	Ghana: 292 (618) Ethiopia: 144 (252)	Ghana: 49.00% Ethiopia: 17.00%	High
Ohene-Yeboah M, 2012 [50]	Ghana	West Africa	49.10 ±15.3	(20-96)	53.70%	/	23	54 (330)	42.70%	Moderate
Ramdas Y, 2020 [51]	South Africa	Southern Africa	60.80 ±9.3	60 (43-86)	9.57%	12.10%	6	107 (107)	5.60%	High
Rayne S, 2019 [52]	South Africa	Southern Africa	/	56	51.10%	/	31	223 (365)	14.00%	Moderate
Roy I, 2011 [53]	Uganda	East Africa	/	/	40.00%	/	16	44 (64)	36.00%	Moderate

Samaka RM, 2016 [54]	Egypt	North Africa	/	(30-81)	44.44%	77.80%	19	81 (81)	23.50%	Moderate
Sawe RT, 2017 [55]	Kenya	East Africa	51.90	48.5 (16-84)	42.22%	/	18	49 (49)	36.70%	Moderate
Schwartz T, 2012 [56]	Ghana	West Africa	49.00 ±13.4	/	52.52%	/	58	103 (104)	56.30%	Moderate
Sengal AT, 2017 [57]	Sudan	North Africa	48.80 ±13.5	(20-90)	41.79%	52.70%	143	415 (560)	34.50%	Moderate
Seshie B, 2015 [58]	Ghana	West Africa	52.50 ±12.1	(24-77)	31.58%	66.00%	77	156 (165)	49.40%	Moderate
Shaikh AJ, 2018 [59]	Kenya	East Africa	54.00	(27-83)	33.33%	49%	21	116 (123)	17.00%	Moderate
Stark A, 2010 [60]	Ghana	West Africa	48.00 ±13.7	/	76.00%	/	37	45 (75)	82.20%	Moderate
Sung H, 2019 [61]	Multiple (USA, Caribbean, West Africa, East Africa)	West and East Africa	/	/	West Africa: 56.92% East Africa: 43.95%	/	West Africa: 173 East Africa: 51	West Africa: 717 (811) East Africa: 441 (484)	West Africa: 24.10% East Africa: 11.60%	Moderate
Tanimowo MO, 2019 [62]	Nigeria	West Africa	43.07 ±11.19	(23-70)	29.51%	/	38	61 (61)	62.30%	High
Traoré B, 2019 [63]	Guinea	West Africa	48.20 ±10.9	(31-80)	36.21%	62.10%	15	58 (58)	25.90%	Moderate
Ukah CO, 2017 [64]	Nigeria	West Africa	46.30 ±12.3	(24-85)	26.83%	/	50	123 (397)	40.70%	Moderate
Uyisenga JP, 2020 [65]	Rwanda	East Africa	49.70 ±12.98	(17-86)	/	70.06%	52	138 (138)	37.70%	Low
Youngblood VM, 2020 [66]	Malawi	East Africa	/	49 (21-80)	37.00%	77%	25	100 (100)	25.00%	Low
Znati K, 2012 [67]	Morocco	North Africa	30.62	(18-35)	45.07%	51%	13	74 (74)	18.00%	High

/ data not stated

IBCa- inflammatory breast cancer

Figure S1: Risk of bias assessment of all articles by (A) region, (B) country and (C) overall risk of bias assessment for all articles included. Southern Africa had the highest percentage of low risk of bias. Multiple denotes more than one country and/or region or was investigated within that particular study.

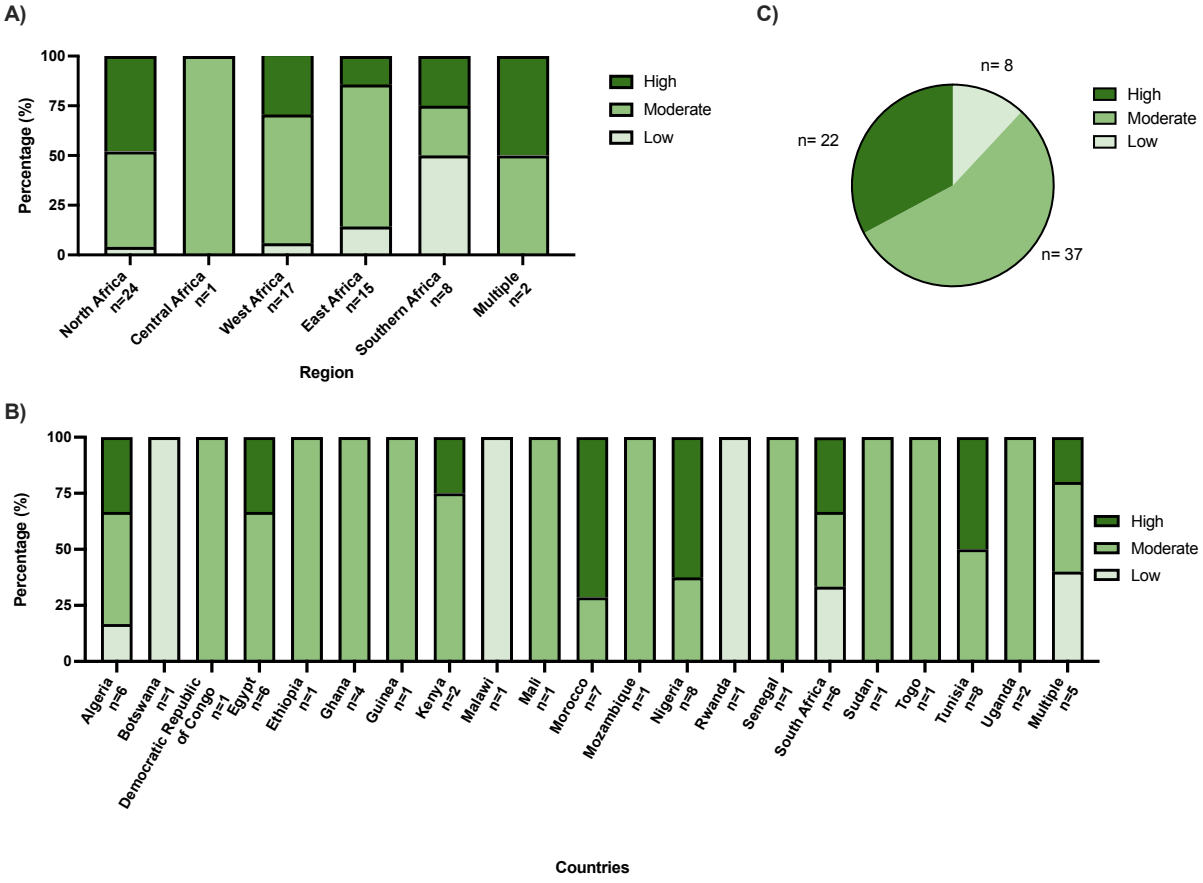


Table S3: Questions for Risk of Bias Assessment, modified from Hoy *et al.*, 2012

1. Was the study's target population a close representation of the national population in relation to relevant variables, e.g. age, sex, occupation?
2. Was the sampling frame a true or close representation of the target population?
3. Was some form of random selection used to select the sample, or was a census undertaken?
4. Was the likelihood of non-response bias minimal?
5. Were data collected directly from the subjects (as opposed to a proxy)?
6. Was an acceptable case definition used in the study? Was TNBC defined as ER, PR, HER2 negative? (Or cited) I.e. were cutoffs used?
7. Was the study instrument that measured the parameter of interest (e.g. appropriate guidelines for cut-offs ASCO-CAP) shown to have reliability and validity?
8. Was the same mode of data collection used for all subjects? Was the information from each patient collected in the same way (e.g clinical referral appointment)?
9. Was there a shortest prevalence or incidence period identified?
10.Were the numerator(s) and denominator(s) for the parameter of interest appropriate?

Table S4: Raw Values for Risk of Bias Assessment for Included Studies

Author	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	TOTAL	Overall ROB
Adani-Ifè A, 2020 [1]	0	0	0	0	1	1	1	1	1	1	6	**Moderate**
Adisa AO, 2011 [2]	0	0	1	0	1	0	0	0	1	0	3	**High**
Agboola A, 2014 [3]	0	0	0	0	1	1	1	1	1	0	5	**High**
Agboola A, 2014 [4]	0	0	1	0	1	1	1	1	1	1	7	**Moderate**
Agboola A, 2013 [5]	0	0	0	0	1	1	1	1	0	1	5	**High**
Agboola A, 2016 [6]	0	0	0	1	1	0	1	1	1	1	6	**Moderate**
Agboola A, 2014 [7]	0	0	0	0	1	0	1	1	1	1	5	**High**
Akinyi MO, 2019 [8]	0	0	0	1	1	1	0	1	0	0	4	**High**
Alshenawy HA, 2012 [9]	0	0	1	1	1	1	0	1	1	1	7	**Moderate**
Ayadi EZ, 2018 [10]	0	0	0	1	1	1	1	1	1	1	7	**Moderate**
Babalanda JNP, 2014 [11]	0	0	1	1	1	1	1	1	1	1	8	**Moderate**
Bahnassy A, 2015 [12]	0	0	1	1	1	1	0	1	1	1	7	**Moderate**
Bakkach J, 2017 [13]	0	0	0	1	1	0	0	1	1	0	4	**High**
Bekkouche Z, 2013 [14]	0	0	1	0	1	1	0	1	1	0	5	**High**
Belhadj A, 2021 [15]	1	0	1	1	1	1	0	1	1	1	8	**Moderate**
Ben Ayed-Guerfali D, 2021 [16]	0	0	0	0	1	0	0	1	1	0	3	**High**
Ben Gacem R, 2011 [17]	0	0	1	1	1	1	0	1	0	1	6	**Moderate**
Bhatia RK, 2019 [18]	1	1	1	1	1	1	1	1	1	1	10	**Low**
Bogan D, 2017 [19]	0	0	0	1	1	1	1	1	1	1	7	**Moderate**
Brandão M, 2020 [20]	0	0	1	1	1	1	1	0	1	1	7	**Moderate**
Chahbouni S, 2017 [21]	0	0	0	0	1	1	0	1	1	0	4	**High**
Cherbal F, 2015 [22]	1	1	1	0	1	1	0	1	1	1	8	**Moderate**
Corsini C, 2017 [23]	0	0	1	1	1	0	0	1	1	1	6	**Moderate**
Cubasch H, 2018 [24]	1	1	1	1	1	1	0	1	1	1	9	**Low**
Debouki-Joudi S, 2016 [25]	0	0	0	0	1	1	0	1	0	1	4	**High**
Dickens C, 2014 [26]	1	1	1	1	1	1	1	1	1	1	10	**Low**
Ekpe E, 2019 [27]	0	0	0	0	1	1	1	1	1	1	6	**Moderate**

El-Amir MI, 2020 [28]	0	0	0	0	1	0	0	1	1	1	4	**High**
Elidrissi Errahhali M, 2017 [29]	0	1	1	0	1	1	0	1	1	1	7	**Moderate**
Fitzpatrick MB, 2019 [30]	0	0	0	0	1	1	1	1	1	1	6	**Moderate**
Fouhi ME, 2020 [31]	0	1	0	0	1	0	0	1	1	0	4	**High**
Gaceb H, 2017 [32]	1	1	1	0	1	1	1	1	1	1	9	**Low**
Groenewald C, 2019 [33]	0	0	0	1	1	1	0	1	1	1	6	**Moderate**
Guedouar Y, 2014 [34]	0	1	0	0	1	1	0	1	1	0	5	**High**
Hachana M, 2012 [35]	0	0	1	1	1	1	0	1	1	1	7	**Moderate**
Hadgu E, 2020 [36]	0	0	0	1	1	1	1	1	1	1	7	**Moderate**
Huo D, 2009 [37]	0	1	1	1	1	1	1	1	1	1	9	**Low**
Hussein O, 2013 [38]	0	0	1	0	1	0	0	1	1	1	5	**High**
Kakudji BK, 2021 [39]	0	0	0	1	1	1	0	0	1	1	5	**High**
Kallel-Bayoudh I, 2011 [40]	0	0	0	0	1	1	0	1	1	1	5	**High**
Khalil AI, 2016 [41]	0	1	1	0	1	1	0	1	1	1	7	**Moderate**
Luyey Mvila G, 2015 [42]	0	0	1	0	1	1	1	1	1	1	7	**Moderate**
Ly M, 2012 [43]	0	0	1	1	1	1	1	1	1	1	8	**Moderate**
Manai M, 2019 [44]	0	1	1	0	1	1	0	1	1	1	7	**Moderate**
McCormack VA, 2013 [45]	1	1	1	1	1	1	0	1	1	1	9	**Low**
Mehemmai C, 2019 [46]	0	0	1	1	1	1	1	1	1	1	8	**Moderate**
Mejri N, 2020 [47]	0	0	0	0	1	0	0	1	1	0	3	**High**
Mouh FZ, 2020 [48]	0	0	0	0	1	1	0	1	1	1	5	**High**
Newman LA, 2019 [49]	0	0	0	0	1	0	1	1	1	1	5	**High**
Ohene-Yeboah M, 2012 [50]	0	0	1	0	1	1	0	1	1	1	6	**Moderate**
Ramdas Y, 2020 [51]	0	0	0	1	1	0	0	1	1	1	5	**High**
Rayne S, 2019 [52]	0	0	1	0	1	1	1	0	1	1	6	**Moderate**
Roy I, 2011 [53]	0	0	0	0	1	1	1	1	1	1	6	**Moderate**
Samaka RM, 2016 [54]	0	0	0	1	1	1	1	1	1	1	7	**Moderate**
Sawe RT, 2017 [55]	0	1	1	0	1	1	0	1	1	0	6	**Moderate**

Schwartz T, 2012 [56]	0	0	0	1	1	1	0	1	1	1	6	**Moderate**
Sengal AT, 2017 [57]	0	0	1	1	1	1	1	1	1	1	8	**Moderate**
Seshie B, 2015 [58]	0	1	1	1	1	1	0	1	1	0	7	**Moderate**
Shaikh AJ, 2018 [59]	0	0	1	1	1	1	1	1	1	1	8	**Moderate**
Stark A, 2010 [60]	0	0	1	0	1	1	1	1	1	0	6	**Moderate**
Sung H, 2019 [61]	0	1	1	1	1	1	0	1	1	1	8	**Moderate**
Tanimowo MO, 2019 [62]	0	0	0	1	1	0	1	0	1	1	5	**High**
Traoré B, 2019 [63]	0	1	1	0	1	1	0	1	1	0	6	**Moderate**
Ukah CO, 2017 [64]	0	1	0	1	1	1	0	1	1	1	7	**Moderate**
Uyisenga JP, 2020 [65]	0	1	1	1	1	1	1	1	1	1	9	**Low**
Youngblood VM, 2020 [66]	0	1	1	1	1	1	1	1	1	1	9	**Low**
Znati K, 2012 [67]	0	0	0	0	1	0	0	1	1	0	3	**High**

Table S5: Pooled TNBC Frequcny in Africa by Country

Country (no. of studies)	Total participants	Pooled TNBC Frequency, % [95% CI]	Heterogeneity (I^2), %	Weight, %
Algeria (*n*= 6)	6461	22.6 [20.3- 25.1]	76	12.0
Botswana (*n*= 1)	211	21.3 [16.0- 27.1]	not applicable	1.9
Democratic Republic of Congo (*n*= 1)	87	14.9 [8.1- 23.3]	not applicable	1.7
Egypt (*n*= 5)	640	24.2 [14.0- 36.1]	90	8.8
Ethiopia (*n*= 1)	110	22.7 [15.3- 31.1]	not applicable	1.8
Ghana (*n*= 4)	358	57.6 [43.1- 71.4]	86	6.6
Guinea (*n*= 1)	58	25.9 [15.3- 38.0]	not applicable	1.6
Kenya (*n*= 4)	364	24.3 [17.3- 32.0]	62	6.8
Malawi (*n*= 1)	100	25.0 [17.0- 34.0]	not applicable	1.8
Mali (*n*= 1)	114	46.5 [37.4- 55.7]	not applicable	1.8
Morocco (*n*= 7)	4744	15.5 [14.1- 16.9]	30	13.5
Mozambique (*n*= 1)	210	24.8 [19.1- 30.8]	not applicable	1.9
Nigeria (*n* = 4)	479	49.2 [41.4- 57.1]	62	6.9
Rwanda (*n*= 1)	138	37.7 [29.8- 45.9]	not applicable	1.8
Senegal (*n*= 1)	197	46.7 [39.8- 53.7]	not applicable	1.9
South Africa (*n*= 4)	950	13.5 [8.8- 19.1]	78	7.5
Sudan (*n*= 1)	415	34.5 [30.0- 39.1]	not applicable	2.0
Togo (*n*= 1)	117	37.6 [29.0- 46.6]	not applicable	1.8
Tunisia (*n*= 8)	1628	26.8 [18.6- 35.9]	93	14.6
Uganda (*n*= 2)	158	27.0 [12.7- 44.1]	76	3.2
Total	17539	27.4 [24.5- 30.3]	94	100

Figure S2: Pooled TNBC frequency in Africa by risk of bias assessment. Cases are defined as participants in a study that were identified as triple negative, and total is the number of breast cancer participants with known receptor status in the study.

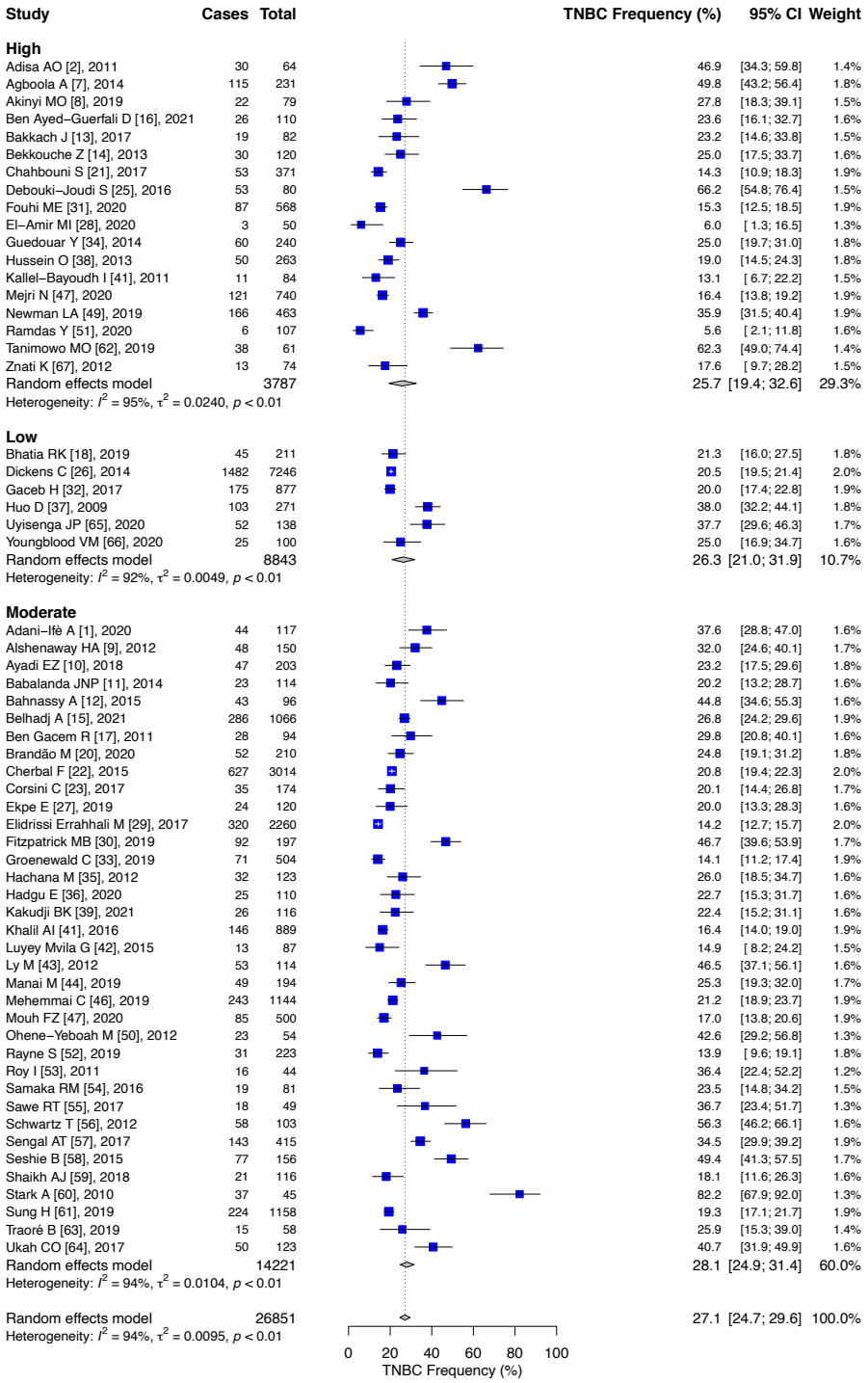


Figure S3: Pooled TNBC frequency in Africa by use of validated tool for receptor status. Cases are defined as participants in a study that were identified as triple negative, and total is the number of breast cancer participants with known receptor status in the study.

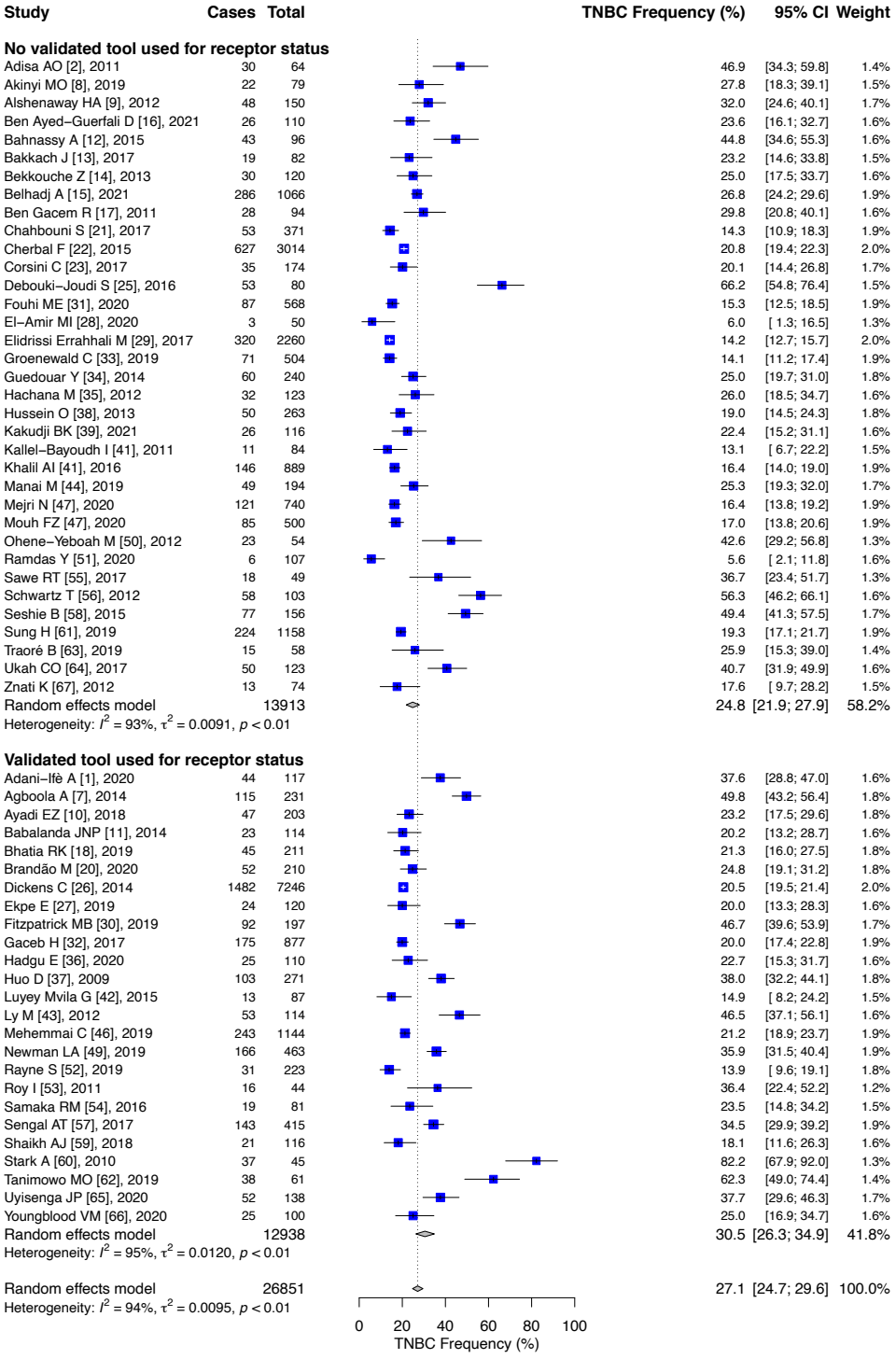


Figure S4: Meta-regression shows association with year of publication but not with age or use of validated tool for receptor status. The association of use of validated tool to assess receptor status (A), year of publication (B), mean age at diagnosis (C) and median age at diagnosis (D) on heterogeneity of treatment effects between studies was investigated by a meta-regression. P <0.05 indicates significant association.

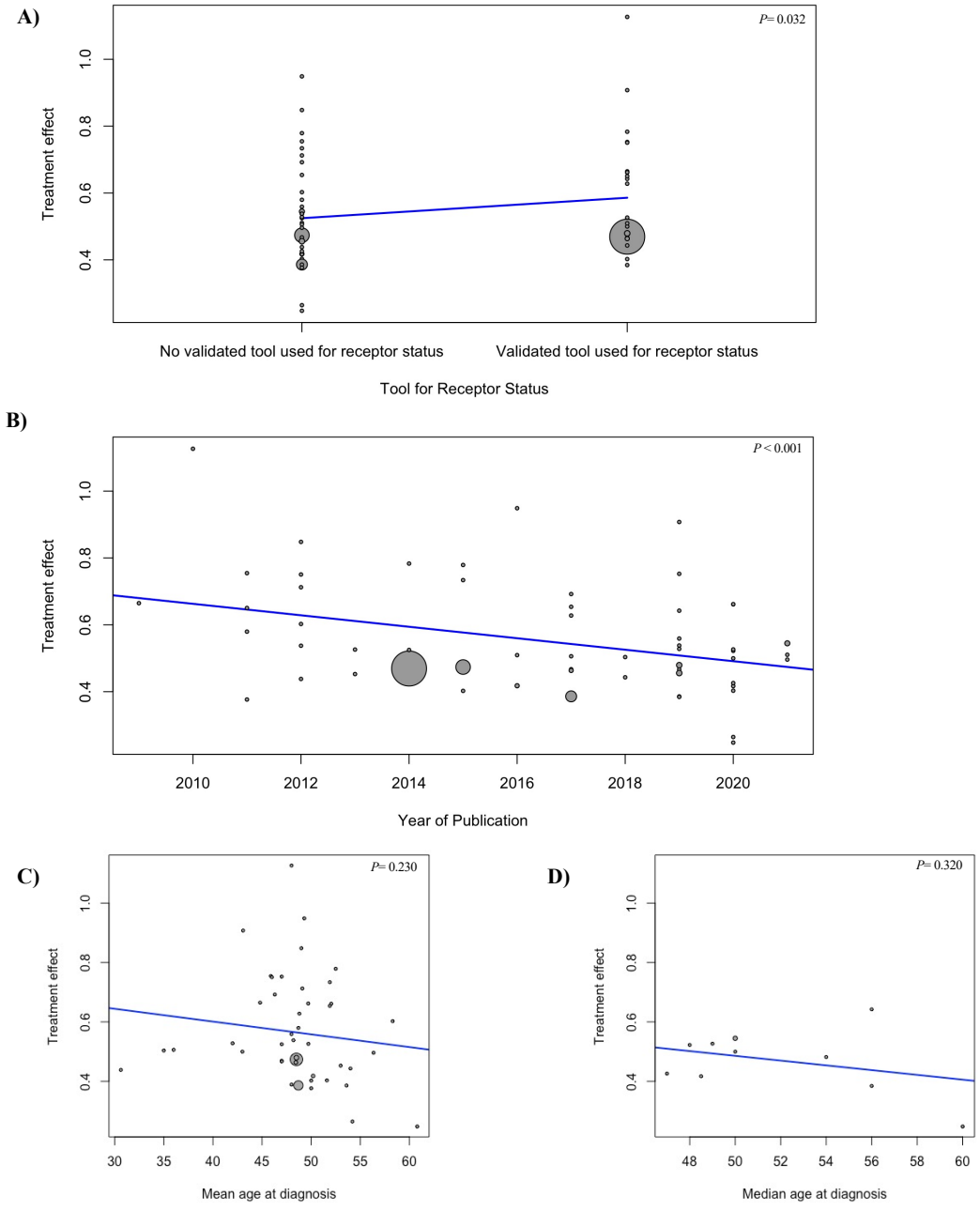
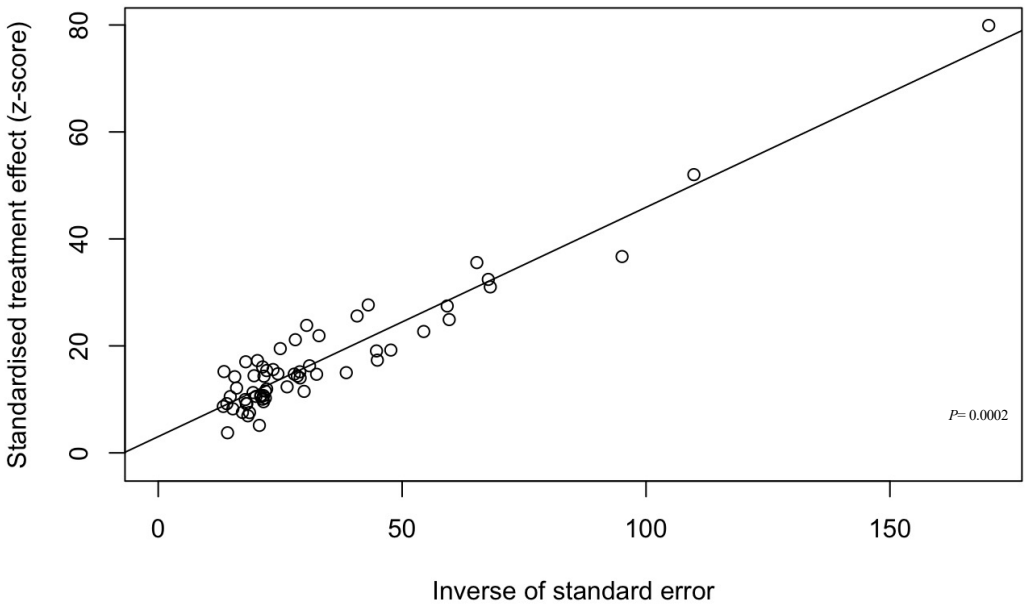


Figure S5: Egger's test shows high publication bias. Publication bias between studies for TNBC frequency investigated using Egger's test and a significant association was found, inferring significant publication bias. P <0.05 indicates significant association.



References for studies included in Systematic Review:

1. Adani-Ife, A., *et al.*, *Breast cancer in togolese women: immunohistochemistry subtypes.* Bmc Womens Health, 2020. **20**(1).
2. Adisa, A.O., *et al.*, *Metastatic breast cancer in a Nigerian tertiary hospital.* Afr Health Sci, 2011. **11**(2): p. 279-84.
3. Agboola, A.O.J., *et al.*, *Clinicopathological and molecular significance of Sumolyation marker (ubiquitin conjugating enzyme 9 (UBC9)) expression in breast cancer of black women.* Pathology - Research and Practice, 2014. **210**(1): p. 10-17.
4. Agboola, A., *et al.*, *Molecular profiling of breast cancer in Nigerian women identifies an altered p53 pathway as a major mechanism underlying its poor prognosis compared with British counterpart.* The Malaysian journal of pathology, 2014. **36 1**: p. 3-17.
5. Agboola, A.O.J., *et al.*, *Cell Proliferation (KI-67) Expression Is Associated with Poorer Prognosis in Nigerian Compared to British Breast Cancer Women.* ISRN Oncology, 2013. **2013**: p. 675051.
6. Agboola, A., *et al.*, *Tumour cell membrane laminin expression is associated with basal-like phenotype and poor survival in Nigerian breast cancer.* The Malaysian journal of pathology, 2016. **38**: p. 83-92.
7. Agboola, A., *et al.*, *PIASγ expression in relation to clinicopathological, tumour factors and survival in indigenous black breast cancer women.* Journal of Clinical Pathology, 2014. **67**(4): p. 301.
8. Akinyi, M.O., *et al.*, *Correlation between receptor status and presence of axillary lymph node metastasis in breast cancer in Kenya.* Annals of African Surgery, 2019. **16**(2).
9. Alshenawy, H.A., *Prevalence of androgen receptors in invasive breast carcinoma and its relation with estrogen receptor, progesterone receptor and Her2/neu expression.* Journal of the Egyptian National Cancer Institute, 2012. **24**(2): p. 77-83.
10. Ayadi, E.Z., *et al.*, *Prognostic Value of BCL2 in Women Patients with Invasive Breast Cancer.* Asian Pac J Cancer Prev, 2018. **19**(12): p. 3557-3564.
11. Babalanda, J., *et al.*, *Receptor Status and Associated Clinico- Histopathological Characteristics among Women with Breast Cancer in a Ugandan Tertiary Hospital.* The Annals of African Surgery, 2014. **11**.
12. Bahhnassy, A., *et al.*, *Transforming growth factor-β, insulin-like growth factor I/insulin-like growth factor I receptor and vascular endothelial growth factor-A: Prognostic and predictive markers in triple-negative and non-triple-negative breast cancer.* Mol Med Rep, 2015. **12**(1): p. 851-864.
13. Bakkach, J., *et al.*, *Clinicopathologic and prognostic features of breast cancer in young women: a series from North of Morocco.* BMC Women's Health, 2017. **17**(1): p. 106.

14. Bekkouche, Z., *et al.*, *Caractéristiques des carcinomes mammaires triple-négatifs dans l'Ouest-algérien.* Journal Africain du Cancer / African Journal of Cancer, 2013. **5**(3): p. 155-161.

15. Belhadj, A., *et al.*, *Prevalence and prognosis of molecular phenotypes in breast cancer patients by age: a population-based retrospective cohort study in western Algeria.* Pan Afr Med J, 2021. **38**: p. 88.

16. Ben Ayed-Guerfali, D., *et al.*, *Novel and recurrent BRCA1/BRCA2 germline mutations in patients with breast/ovarian cancer: a series from the south of Tunisia.* J Transl Med, 2021. **19**(1): p. 108.

17. Ben Gacem, R., *et al.*, *Clinicopathologic significance of DNA methyltransferase 1, 3a, and 3b overexpression in Tunisian breast cancers.* Human Pathology, 2012. **43**(10): p. 1731-1738.

18. Bhatia, R.K., *et al.*, *Report of clinico-pathological features of breast cancer in HIV-infected and uninfected women in Botswana.* Infect Agent Cancer, 2019. **14**: p. 28.

19. Bogan, D., *et al.*, *The role of BRCA1-IRIS in the development and progression of triple negative breast cancers in Egypt: possible link to disease early lesion.* BMC Cancer, 2017. **17**(1): p. 329.

20. Brandão, M., *et al.*, *Breast cancer subtypes: implications for the treatment and survival of patients in Africa-a prospective cohort study from Mozambique.* ESMO Open, 2020. **5**(5): p. e000829.

21. Chahbouni, S., *et al. CLINICOPATHOLOGIC FEATURES AND MOLECULAR SUBTYPES OF BREAST CANCER IN FEZ-MEKNES REGION (MOROCCO): A STUDY OF 390 PATIENTS*. 2017.

22. Cherbal, F., *et al.*, *Distribution of molecular breast cancer subtypes among Algerian women and correlation with clinical and tumor characteristics: A population-based study.* Breast disease, 2015. **35**: p. 95-102.

23. Corsini, C., *et al.*, *Early onset breast cancer: differences in risk factors, tumor phenotype, and genotype between North African and South European women.* Breast Cancer Research and Treatment, 2017. **166**(2): p. 631-639.

24. Cubasch, H., *et al.*, *Breast cancer survival in Soweto, Johannesburg, South Africa: A receptor-defined cohort of women diagnosed from 2009 to 11.* Cancer Epidemiology, 2018. **52**: p. 120-127.

25. Debouki-Joudi, S., *et al.*, *CpG methylation of APC promoter 1A in sporadic and familial breast cancer patients.* Cancer Biomarkers, 2016. **16**: p. 1-9.

26. Dickens, C., *et al.*, *Racial Comparison of Receptor-Defined Breast Cancer in Southern African Women: Subtype Prevalence and Age–Incidence Analysis of Nationwide Cancer Registry Data.* Cancer Epidemiology Biomarkers & Prevention, 2014. **23**(11): p. 2311.

27. Ekpe, E., *et al.*, *Metastatic Breast Cancer in Kenya: Presentation, Pathologic Characteristics, and Patterns-Findings From a Tertiary Cancer Center.* Journal of Global Oncology, 2019. **5**.

28. El-Amir, M.I., *et al.*, *Role of ATG16LI (rs2241880) and Interleukin 10 (rs1800872) Polymorphisms in Breast Cancer Among Egyptian Patients.* Egypt J Immunol, 2020. **27**(1): p. 65-76.

29. Elidrissi Errahhali, M., *et al.*, *First report on molecular breast cancer subtypes and their clinico-pathological characteristics in Eastern Morocco: series of 2260 cases.* BMC Women's Health, 2017. **17**(1): p. 3.

30. Fitzpatrick, M.B., *et al.*, *Pathology of Senegalese breast cancers.* Pan Afr Med J, 2019. **34**: p. 67.

31. Fouhi, M.E., *et al.*, *[Epidemiological and anatomopathological profile of breast cancer at the Ibn Rochd University Hospital, Casablanca].* Pan Afr Med J, 2020. **37**: p. 41.

32. Gaceb, H., *et al.*, *Clinicopathological and Molecular Study of Triple-Negative Breast Cancer in Algerian Patients.* Pathology & Oncology Research, 2017. **24**(2): p. 297-308.

33. Groenewald, C., *et al.*, *Axillary lymph node dissection for patients with invasive breast cancer at Charlotte Maxeke and Chris Hani Baragwanath Academic Hospitals.* S Afr J Surg, 2019. **57**(4): p. 18-24.

34. Guedouar, Y., *et al.*, *Évaluation phénotypique des sous-types moléculaires en carcinologie mammaire dans une population de l'Ouest algérien.* Journal Africain du Cancer / African Journal of Cancer, 2014. **6**(3): p. 150-158.

35. Hachana, M., *et al.*, *Investigation of human JC and BK polyomaviruses in breast carcinomas.* Breast Cancer Research and Treatment, 2012. **133**(3): p. 969-977.

36. Hadgu, E., *et al.*, *Distribution and characteristics of androgen receptor (AR) in breast cancer among women in Addis Ababa, Ethiopia: A cross sectional study.* Plos One, 2020. **15**(5).

37. Huo, D., *et al.*, *Population Differences in Breast Cancer: Survey in Indigenous African Women Reveals Over-Representation of Triple-Negative Breast Cancer.* Journal of Clinical Oncology, 2009. **27**(27): p. 4515-4521.

38. Hussein, O., *et al.*, *Hormone Receptors and Age Distribution in Breast Cancer Patients at a University Hospital in Northern Egypt.* Breast Cancer: Basic and Clinical Research, 2013. **7**: p. BCBCR.S12214.

39. Kakudji, B.K., *et al.*, *Breast cancer molecular subtypes and receptor status among women at Potchefstroom Hospital: a cross-sectional study.* Pan Afr Med J, 2021. **38**: p. 85.

40. Kallel-Bayoudh, I., *et al.*, *Bcl-2 expression and triple negative profile in breast carcinoma.* Medical Oncology, 2011. **28**(1): p. 55-61.

41. Khalil, A., *et al.*, *Cancer du sein au Maroc: profil phénotypique des tumeurs.* The Pan African Medical Journal, 2016. **25**.

42. Luyeye Mvila, G., *et al.*, *Prognostic features of breast cancer differ between women in the Democratic Republic of Congo and Belgium.* The Breast, 2015. **24**(5): p. 642-648.

43. Ly, M., *et al.*, *High Incidence of Triple-Negative Tumors in Sub-Saharan Africa: A Prospective Study of Breast Cancer Characteristics and Risk Factors in Malian Women Seen in a Bamako University Hospital.* Oncology, 2012. **83**(5): p. 257-263.

44. Manai, M., *et al.*, *Inflammatory breast cancer in 210 patients: A retrospective study on epidemiological, anatomo-clinical features and therapeutic results.* Mol Clin Oncol, 2019. **10**(2): p. 223-230.

45. McCormack, V.A., *et al.*, *Breast cancer receptor status and stage at diagnosis in over 1,200 consecutive public hospital patients in Soweto, South Africa: a case series.* Breast cancer research : BCR, 2013. **15**(5): p. R84-R84.

46. Mehemmai, C., *et al.*, *BRCA1 and BRCA2 Germline Mutation Analysis in Hereditary Breast/Ovarian Cancer Families from the Aures Region (Eastern Algeria): First Report.* Pathology & Oncology Research, 2019. **26**(2): p. 715-726.

47. Mejri, N., *et al.*, *Reproductive Risk Factors of Inflammatory Breast Cancer according to Luminal, HER2-Overexpressing, and Triple-Negative Subtypes: A Case Comparison Study.* Oncol Res Treat, 2020. **43**(5): p. 204-210.

48. Mouh, F.Z., *et al.*, *Clinicopathological, Treatment and Event-Free Survival Characteristics in a Moroccan Population of Triple-Negative Breast Cancer.* Breast Cancer (Auckl), 2020. **14**: p. 1178223420906428.

49. Newman, L.A., *et al.*, *Hereditary Susceptibility for Triple Negative Breast Cancer Associated With Western Sub-Saharan African Ancestry: Results From an International Surgical Breast Cancer Collaborative.* Ann Surg, 2019. **270**(3): p. 484-492.

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

51. Ramdas, Y., C.A. Benn, and M. van Heerden, *First Intraoperative Radiation Therapy Center in Africa: First 2 Years in Operation, Including COVID-19 Experiences.* JCO Glob Oncol, 2020. **6**: p. 1696-1703.

52. Rayne, S., *et al.*, *Unraveling the South African Breast Cancer Story: The Relationship of Patients, Delay to Diagnosis, and Tumor Biology With Stage at Presentation in an Urban Setting.* Journal of Surgical Research, 2019. **235**: p. 181-189.

53. Roy, I. and E. Othieno, *Breast carcinoma in Uganda: microscopic study and receptor profile of 45 cases.* Arch Pathol Lab Med, 2011. **135**(2): p. 194-9.

54. Samaka, R.M. and S.F. Younes, *Androgen Receptor Expression in Breast Carcinoma of Egyptian Patients.* J Clin Diagn Res, 2016. **10**(11): p. Ec17-ec21.

55. Sawe, R.T., *et al.*, *Tumor infiltrating leukocyte density is independent of tumor grade and molecular subtype in aggressive breast cancer of Western Kenya.* Tropical Medicine and Health, 2017. **45**(1): p. 19.

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

57. Sengal, A.T., *et al.*, *Comparison of Receptor-Defined Breast Cancer Subtypes Between German and Sudanese Women: A Facility-Based Cohort Study.* Journal of Global Oncology, 2017(4): p. 1-12.

58. Seshie, B., *et al.*, *A retrospective analysis of breast cancer subtype based on ER/PR and HER2 status in Ghanaian patients at the Korle Bu Teaching Hospital, Ghana.* BMC Clinical Pathology, 2015. **15**(1): p. 14.

59. Shaikh, A.J., *et al.*, *Mammographic Breast Density and Breast Cancer Molecular Subtypes: The Kenyan-African Aspect.* BioMed Research International, 2018. **2018**: p. 6026315.

60. Stark, A., *et al.*, *African ancestry and higher prevalence of triple-negative breast cancer.* Cancer, 2010. **116**(21): p. 4926-4932.

61. Sung, H., *et al.*, *Breast cancer subtypes among Eastern-African–born black women and other black women in the United States.* Cancer, 2019. **125**(19): p. 3401-3411.

62. Tanimowo, M.O., *et al.*, *Histopathological and immunohistochemical characteristics of breast carcinomas in Uyo, subtropical region of Africa.* Medical Journal of Zambia, 2019. **46**(2): p. 100-108.

63. Traore, B., *et al.*, *Molecular profile of breast cancers in Guinean oncological settings.* Pan African Medical Journal, 2019. **33**.

64. CO., U., *et al.*, *The Immunohistochemical Profile of Breast Cancer in Indigenous Women of Southeast Nigeria* Annals of Medical and Health Sciences Research, 2017. **7**(2): p. 83-87.

65. Uyisenga, J.P., *et al.*, *Prevalence of Histological Characteristics of Breast Cancer in Rwanda in Relation to Age and Tumor Stages.* Horm Cancer, 2020. **11**(5-6): p. 240-249.

66. Youngblood, V.M., *et al.*, *Outcomes and prognostic factors for women with breast cancer in Malawi.* Cancer Causes Control, 2020. **31**(4): p. 393-402.

67. Znati, K., *et al.*, *Cancer du sein chez la femme jeune dans le Nord-Est du Maroc.* Gynécologie Obstétrique & Fertilité, 2012. **42**(3): p. 149-154.

CHAPTER 5: ANALYSIS OF THE GENOMIC LANDSCAPES OF BARBADIAN AND NIGERIAN WOMEN WITH TRIPLE NEGATIVE BREAST CANCER (TNBC)

Preface This chapter consists of the submitted manuscript entitled "***Analysis of the genomic landscapes of Barbadian and Nigerian women with High prevalence of triple negative breast cancer (TNBC)"*** by Hercules SM., Liu X., Bassey-Archibong BI., Skeete DHA., Smith Connell S., Daramola A, Banjo AA., Ekanem IA., Ebughe G, Agan T, Udosen J., Obiorah CC., Ojule AC., Misauno MA., Dauda AM., Egbujo EC., Xu Y, Jin Y., Chang S., Hercules JC., Ansari A., Brain I., MacColl C., Carpten JD., Bédard A., Pond GR., Blenman KRM., Manojlovic Z., Daniel JM submitted to Cancer Causes and Control and under review in its original form. In this manuscript, we investigated the genomic profile of Barbadian and Nigerian WAA with triple negative breast cancer.

Contributions: SM Hercules and Dr. JM Daniel conceptualized and designed the study, SM Hercules collected TNBC tissues and associated clinical data from Barbados, clinical data from Nigeria, analysed data, interpreted data, wrote original draft, reviewed & edited the manuscript. X Liu analysed data and provided bioinformatics support with data interpretation assistance. Dr. BI Bassey-Archibong collected TNBC tissues from Nigeria and established collaborations with Nigerian Physicians and researchers. She also contributed intellectual input and feedback. Drs. DHA Skeete and S Smith Connell assisted with collection of TNBC tissues and associated clinical data in Barbados and interpreted pathological reports. Drs. A Daramola, AA Banjo, IA Ekanem, G Ebughe, T Agan, J Udosen, CC Obiorah, AC Ojule, MA Misauno, AM Dauda and EC Egbujo assisted with collection of Nigerian BCa tissues. Y Xu, Y Jin and S Chang assisted with bioinformatics support. Dr. JM Daniel, J Hercules and A Ansari assisted with collection of TNBC

tissues and associated clinical data from Barbados. Drs. I Brain and C MacColl provided pathological support. Dr. JD Carpten provided bioinformatics support. Dr. A Bédard provided supervision throughout the course of this project and edited the manuscript. Dr. GR Pond provided statistical support and edited the manuscript. Dr. KRM Blenman provided data interpretation support. Dr. Z Manojlovic provided bioinformatics support and supervision. Dr. JM Daniel provided significant guidance and intellectual input and funded the study. All authors edited the manuscript text.

Full Title: Analysis of the genomic landscapes of Barbadian and Nigerian women with triple negative breast cancer

Short Title: Genomics, women of African ancestry and triple negative breast cancer

Authors: Shawn M. Hercules, Xiyu Liu, Blessing B. I. Bassey-Archibong, Desiree H. A. Skeete, Suzanne Smith Connell, Adetola Daramola, Adekunbiola A. Banjo, Ima-Obongg Ekanem, Godwin Ebughe, Thomas Agan, Joe Udosen, Christopher Obiorah, Aaron C. Ojule, Michawl A. Misauno, Ayuba M. Dauda, Ejike C. Egbujo, Yuxin Xu, Yili Jin, Sharon Chang, Jevon C. Hercules, Amna Ansari, Ian Brain, Christina MacColl, John D. Carpten, André Bédard, Greg R. Pond, Kim R. M. Blenman, Zarko Manojlovic, Juliet M. Daniel.*

*Corresponding author: danielj@mcmaster.ca

Author Affiliations

Shawn M. Hercules, MPH: Department of Biology, McMaster University, Hamilton, Ontario, Canada; African Caribbean Cancer Consortium, Philadelphia, Pennsylvania, United States of America. hercules@mcmaster.ca

Xiyu Liu, M.S..: Department of Translational Genomics, Keck School of Medicine, University of Southern California, Los Angeles, California, United States of America. xiyuliu@usc.edu

Blessing Bassey-Archibong, Ph.D.: Stem Cell and Cancer Research Institute (SCC-RI), McMaster University, Hamilton, Ontario, Canada. basseybi@mcmaster.ca

Desiree H. A. Skeete, MBBS: African Caribbean Cancer Consortium, Philadelphia, Pennsylvania, United States of America; Faculty of Medical Sciences, University of the West Indies at Cave Hill, Bridgetown, Barbados; Department of Pathology, Queen Elizabeth Hospital, Bridgetown, Barbados. desiree.skeete@cavehill.uwi.edu

Suzanne Smith Connell, M.D.: [a]Faculty of Medical Sciences, University of the West Indies at Cave Hill, Bridgetown, Barbados; [a]Department of Radiation Oncology, Queen Elizabeth Hospital, Bridgetown, Barbados. [b]Cancer Specialists Inc, Barbados spsmithmd@gmail.com

Adetola Daramola, MBBS: Department of Anatomic and Molecular Pathology, Lagos University Teaching Hospital, Lagos, Nigeria. aodaramola@cmul.edu.ng

Adekunbiola A F. Banjo, MBBS: Department of Anatomic and Molecular Pathology, Lagos University Teaching Hospital, Lagos, Nigeria. aafbanjo@cmul.edu.ng

Godwin Ebughe, MBBS: Department of Pathology, University of Calabar Teaching Hospital, Calabar, Nigeria. gobeng02@gmail.com

Thomas Agan, MBBS: Department of Obstetrics & Gynaecology, College of Medical Sciences, University of Calabar Teaching Hospital. evitomtu@yahoo.com

Ima-Obong Ekanem, MBBS: Department of Pathology, College of Medical Sciences, University of Calabar Teaching Hospital. imaekanem2013@gmail.com

Joe E. Udosen, MBBS: Division of General and Breast Surgery, University of Calabar Teaching Hospital josephessien@unical.edu.ng

Christopher Obiorah, MBBS: Department of Anatomical Pathology, University of Port Harcourt Teaching Hospital, Port Harcourt, Nigeria. christopher.obiorah@uniport.edu.ng

Aaron C. Ojule, MBBS: Department of Chemical Pathology, University of Port Harcourt Teaching Hospital, Port Harcourt, Nigeria. ojuleac@gmail.com

Michael A Misauno, MBBS: Department of Surgery, Jos University Teaching Hospital, Jos, Nigeria. micoyedim@yahoo.com

Ayuba M. Dauda, MBBS: Department of Pathology, Jos University Teaching Hospital, Jos, Nigeria. mdayuba10th@gmail.com

Ejike C. Egbujo, MBBS: Meena Histopathology and Cytology Laboratory, Jos, Nigeria. anatomyejike@gmail.com

Jevon C. Hercules, B.Sc.: [a]Department of Mathematics, University of the West Indies at Mona, Kingston, Jamaica; [b]Wang Yanan Institute for Studies in Economics, Xiamen University, Xiamen, China. jevon.c.hercules@gmail.com

Amna Ansari, B.HSc.(PA): Department of Biology, McMaster University, Hamilton, Ontario, Canada. ansara4@mcmaster.ca

Ian Brain, MD: Department of Pathology and Molecular Medicine, McMaster University, Hamilton, Ontario, Canada.

ian.brain@medportal.ca

Christina MacColl, MD: Department of Pathology and Molecular Medicine McMaster University, Hamilton, Ontario, Canada. christine.maccoll@medportal.ca

Yili Xu, M.S.: Department of Translational Genomics, Keck School of Medicine, University of Southern California, Los Angeles, California, United States of America. yilixu@usc.edu

Yuxin Jin, M.S.: Department of Translational Genomics, Keck School of Medicine, University of Southern California, Los Angeles, California, United States of America. yuxinjin@usc.edu

Sharon Chang, MPH: Department of Translational Genomics, Keck School of Medicine, University of Southern California, Los Angeles, California, United States of America. changsha@usc.edu

John D. Carpten, Ph.D.: Department of Translational Genomics, Keck School of Medicine, University of Southern California, Los Angeles, California, United States of America. carpten@usc.edu

André Bédard, Ph.D.: Department of Biology, McMaster University, Hamilton, Ontario, Canada. abedard@mcmaster.ca

Greg R. Pond, Ph.D., P. Stat: Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada; Department of Oncology, McMaster University, Hamilton, Ontario, Canada. gpond@mcmaster.ca

Kim R. M. Blenman, Ph.D., M.S.: Department of Internal Medicine, Section of Medical Oncology, Yale Cancer Center, Yale University, New Haven, Connecticut, United States of America. kim.blenman@yale.edu

Zarko Manojlovic, Ph.D.: Department of Translational Genomics, Keck School of Medicine, University of Southern California, Los Angeles, California, United States of America. zmanojlo@usc.edu

Juliet M. Daniel, Ph.D.: Department of Biology, McMaster University, Hamilton, Ontario, Canada; African Caribbean Cancer Consortium,

Philadelphia, Pennsylvania, United States of America.
danielj@mcmaster.ca

[a]Previous affiliation; [b]Current affiliation(s)

AUTHORS CONTRIBUTIONS:

Shawn M. Hercules, MPH: conceptualization, study design, data collection, data analysis, data interpretation, co-writing original draft, review & editing.

Xiyu Liu, M.S: data analysis, review & editing.

Blessing Bassey-Archibong, Ph.D.: data collection, co-writing original draft, review & editing.

Desiree H. A. Skeete, MBBS: data collection, review & editing.

Suzanne Smith Connell, MD: data collection, review & editing.

Adetola Daramola, MBBS: data collection, review & editing.

Adekunbiola A F. Banjo, MBBS: data collection, review & editing.

Godwin Ebughe, MBBS: data collection, review & editing.

Thomas Agan, MBBS: data collection, review & editing.

Ima-Obong Ekanem, MBBS: data collection, review & editing.

Joe E. Udosen, MBBS: data collection, review & editing.

Christopher C. Obiorah, MBBS: data collection, review & editing.

Aaron C. Ojule, MBBS: data collection, review & editing.

Michael A. Misauno, MBBS: data collection, review & editing

Dauda A. Madachi, MBBS: data collection, review & editing

Ejike C. Egbujo, MBBS: data collection, review & editing

Jevon C. Hercules, B.Sc.: data collection, review & editing

Amna Ansari, B. HSc.(PA): data collection, review & editing

Ian Brain, MD: pathological assessment

Christina MacColl, MD: pathological asssessment

Yili Xu, M.S.: data analysis

Yuxin Jin, M.S.: data analysis

Sharon Chang, MPH: data analysis

John D. Carpten, Ph.D.: data analysis, data interpretation, review

André Bédard, Ph.D.: data interpretation, review & editing

Gregory R. Pond, Ph.D., P.Stat.: data interpretation, review & editing

Kim R. M. Blenman, Ph.D., M.S.: data analysis, data interpretation, review & editing.

Zarko Manojlovic, Ph.D.: study design, data analysis, data interpretation, co-writing original draft, supervision, review & editing.

Juliet M. Daniel, Ph.D.: conceptualization, study design, data interpretation, supervision, funding acquisition, review & editing,

Funding:

This study was partially funded by the Natural Sciences and Engineering Research Council of Canada (NSERC; grant number RGPIN-2020-06822), Canadian Breast Cancer Foundation/Canadian Cancer Society Research Institute (CBCF/CCSRI; grant number 316252) and Canadian Institutes of Health Research (CIHR; grant number PJT-173223) for data collection storage and trainee stipends. We have not been paid to write this article by a pharmaceutical company or any other agency.

Data availability statement:

Raw sequencing data will be submitted to NCBI SRA prior to the acceptance of the manuscript.

Abstract

Triple negative breast cancer (TNBC) is an aggressive breast cancer subtype that disproportionately affects women of African ancestry (WAA) and is often associated with high recurrence rates and metastasis. Although there is a high prevalence of TNBC across West Africa and in women of the African diaspora, there has been no comprehensive genomics study to investigate the mutational profile of ancestrally related women across the Caribbean and West Africa. This multisite cross-sectional study used 31 formalin-fixed paraffin-embedded (FFPE) samples obtained from Barbadian and Nigerian TNBC participants. High-resolution whole exome sequencing (WES) was performed on the Barbadian and Nigerian TNBC samples to identify their mutational profiles and comparisons were made to African American, European American and Asian American sequencing data obtained from The Cancer Genome Atlas (TCGA). Whole exome sequencing was conducted on tumors with an average of 382x coverage and 4335x coverage for pooled germline (n= 22) non-tumor samples. The most frequently mutated genes in our WAA cohorts were *NBPF12*, *PLIN4*, *TP53* and *BRCA1*. In the TCGA TNBC cases, these genes had a lower mutation rate, except for *TP53* (32% in our cohort; 63% in TCGA-African American; 67% in TCGA-European American; 63% in TCGA-Asian). For all altered genes, there were no differences in frequency of mutations between WAA TNBC groups including the TCGA-African American cohort. For copy number variants, high frequency alterations were observed in *PIK3CA, TP53, FGFR2* and *HIF1AN* genes. This study provides in-depth insights into the underlying genomic alterations in WAA-TNBC samples and shines light on the importance of inclusion of under-represented populations in cancer genomics and biomarker studies.

Author Summary

Breast cancer (BCa) is a leading cause of cancer-related death in women worldwide. Surprisingly although less Black women in the USA are diagnosed with BCa on a yearly basis compared to White women, more Black women die from BCa. This may be due to the fact that a higher proportion of Black women in the USA are diagnosed with a very aggressive and difficult to treat BCa subtype called triple negative breast cancer (TNBC). Currently, little is known about mutations in Black women that could lead to TNBC so we collected DNA from 31 ancestrally-related Barbadian and Nigerian BCa samples and sequenced their DNA. In comparison to mutations from Black and White women in the USA, our results revealed mutations in well-known cancer-driving genes such as *BRCA1* and *TP53* as well as in other genes not previously associated with breast cancer. We also found mutations that could be of therapeutic benefit. This study reports for the first time novel and common mutations between these ancestrally-related groups and shines light on the importance of inclusion of under-represented populations in cancer genomics studies.

Word count: 3077

INTRODUCTION

Breast cancer (BCa) is currently the second leading cause of cancer-related deaths in women worldwide [1] and is routinely categorized into different subtypes based on the amplification of human epidermal growth factor receptor 2 (HER2) and expression of estrogen receptor (ER) and progesterone receptor (PR) [2]. Tumors that lack expression for these three receptors are classified as **t**riple **n**egative **b**reast **c**ancer (TNBC). These tumors are typically more aggressive with advanced grade and stage at diagnosis and limited targeted therapies due to the absence of HER2, ER and PR [3]. Mounting evidence indicates a prevalence of TNBC in West-African women and **w**omen of **A**frican **a**ncestry (WAA) in the Caribbean (~25% in Barbados), the USA (~22%) and the UK (~22%) compared to non-Hispanic White women (11%) [4-6]. The reasons for this disparity are currently unknown; however, recent studies allude to an intricate interplay of environmental and genetic risk factors [7-9].

Due to the aggressive nature of TNBC, there has been an increased interest in investigating molecular biomarkers that could be relevant for therapeutics, diagnostics and prognostics. Recently, it was found that a subset of TNBC patients with deleterious *BRCA1/2* germline mutations responded significantly better to carboplatin (platinum-based therapy) than docetaxel (taxane-based therapy) [10]. In addition to the clinical utility of germline variants, the identification of novel somatic “driver” mutations has been shown to play critical roles in the development of targeted therapies in breast and other solid tumors [11, 12]. Therefore, identification of molecular targets and subsequent development of targeted therapies will be of great importance in improving the overall survival rates of TNBC patients.

Large cancer genomics databases, such as The Cancer Genome Atlas (TCGA), have been useful in understanding the genomics landscape of a variety of tumors. However, to date, most of the TCGA breast cancer biospecimens are from women of European ancestry (~80%) despite higher TNBC prevalence in women of African and Hispanic ancestry [13, 14]. Due to the low percentage of African ancestry cases within the TCGA and other similar repositories, researchers have embarked on conducting independent sequencing studies on these populations to explore and understand their unique genomics landscapes [15-20]. Herein, we have conducted **w**hole **e**xome **s**equencing (WES) of ancestrally-related WAA with TNBC in Nigeria and Barbados to further understand the genomics landscape of these groups. Previous genetic association studies have estimated a high percentage of West African (specifically Nigerian) ancestry in Barbadian cohorts, thus the rationale for these two groups [21-23]. Our findings were compared with the genomic signature of TNBC cases of African American (TCGA-AA) and European American (TCGA-EA) within the TCGA database.

METHODS

Patient population

Formalin-fixed paraffin-embedded (FFPE) specimens with corresponding clinical data were collected from the Pathology Department at the Queen Elizabeth Hospital (QEH) in Barbados and Lagos University Teaching Hospital (LUTH), University of Calabar Teaching Hospital (UCTH), University of Port Harcourt Teaching Hospital (UPTH), and Jos University Teaching Hospital (JUTH) in Nigeria. Fifty-nine FFPE TNBC samples were selected at random for DNA extraction and WES, however only 31 samples passed quality control. Protocols for specimen collection as outlined within the respective institutional review boards were adhered to

and patients consented to give their samples. The study was approved by the institutional review boards at McMaster University, the University of the West Indies – Cave Hill, the QEH, LUTH, UCTH, UPTH and JUTH. Tissues were assessed for ER, PR and HER2 via IHC at their respective institutions and BCa subtype status was further confirmed in Canada. The Allred algorithm [47] was used to calculate the scores.

Pathologic assessment and DNA isolation

Hematoxylin and eosin (H&E) stained slides were made from FFPE samples for pathological interrogation of tumor enrichment. 10 μm FFPE tissue sections on slides were scraped and placed in NAVY RINO tubes (Next Advance, Troy, NY) with stainless steel beads and 160 μL of deparaffinization solution (Qiagen, Hilden, Germany). Samples were homogenized using the Bullet Blender (Next Advance) for 5 minutes at speed 12. Samples were incubated for 3 minutes at 56˚C and processed according to the manufacturer's instructions. DNA was isolated from pathologically assessed tumor enriched regions and uninvolved "normal" sections.

Whole-Exome Sequencing

Quality and quantity of DNA was measured using the Genomic DNA Screen Tape Assay (Agilent Technologies, Santa Clara, CA) and Qubit. The concentration of genomic DNA (gDNA) larger than 200 bp was then calculated, and at least 200 ng of DNA greater than 200 bp was sheared in 50 μL of nuclease-free water with the Covaris E220 using the 96 microTUBE Plate (Covaris, Woburn, MA). Library was prepared using KAPA Hyper Prep Kit (Roche, Basel,

Switzerland) using 100-500 ng of sheared DNA according to the manufacturer's instructions. Individual tumor adapter-ligated libraries were enriched into the exome capture reaction, and for germline each adapter-ligated library was pooled before proceeding to capture using Agilent's SureSelect Human All Exon V6 + custom probes capture library kit [48]. Each library was pooled and sequenced on Illumina's NovaSeq 6000 (Illumina, San Diego, CA) using 300 cycle kit. Raw FASTQs were generated using the industry standard BCL2FASTQ v1.8.4

Primary Informatics Methods

Whole-exome sequences were aligned by BWA (v0.7.17) to GRCh38. Quality score errors were detected by GATK's Base Recalibrator (v4.0.10.1). Picard Tools (v1.128) was used to merge aligned BAMs and mark duplicate reads. Germline Variant Call Format (VCF) of BAM were obtained by GATK's Haplotype Caller using GATK best practices, Samtools MPileUp together with BCFtools (v1.2), and Freebayes (v1.1.0-6-gf069ec6). Somatic variant calling was performed by MuTect2[49] to ensure compatible comparison between with TCGA. MuTect2 somatic variant calling files (VCF) for each patient in this study were converted to MAF files using the vcf2maf v1.6.19 tool [50]. TCGA data were downloaded from: https://portal.gdc.cancer.gov/projects/TCGA-BRCA. The TNBC subtypes were extracted and divided into self-reported Caucasian/European American (EA) and African American (AA) race. Each variant was validated using IGV_(v2.7.2) [46].

Downstream Bioinformatics Methods

Gene frequencies in our WES data were performed by Unified Optimal Sequence Kernal Association testing in R. Visualization of somatic variants was performed using maftools (v2.6.05) [51] and R packages pheatmap (v1.0.12), ggplot2 (v3.3.3), VennDiagram (v1.6.20), and ggrepel (v0.9.1). Mutation signature from WES data was computed using Mutational Signature in Cancer (MuSiCa) [52]. Copy number analysis was performed utilizing Nexus Copy Number v10 (Biodiscovery) and focal analysis was performed by GISTIC (v2.0) [53]. CNV heatmap was plotted using the oncoprint function in the R package ComplexHeatmap (v2.6.2) and pheatmap. To deduce ancestry information from tumor DNA, 1000 Genomes Project phase 3 VCF release was used as our reference population [54]. Data were transformed to numeric genotypes by using PLINK (v1.90b6.7). Principle component analysis (PCA) was performed using the R v3.6.0 function prcomp to establish ancestry distributions mapped by the anchor population.

Type I error (α) and type II error (β) were set at 0.05 and 0.1, respectively. Chi-square test and paired student t-test, as appropriate, were used to examine bivariate association of somatic differences between two cohorts. Benjamini-Hochberg was used for multiple tests. For statistical analysis and visualization, GraphPad Prism 8 was implemented (GraphPad Software, Inc.) and Rv3.6.0 packages: circlize (0.4.6), ComplexHeatmap (1.99.7), dplyr (0.8.0.1), ggplot2 (3.1.1), ggpubr (0.2), maftools (2.0.05), plyr (1.8.4), png (0.1-7), qvalue (2.16.0), reshape2 (1.4.3), stringr (1.4.0), TCGAbiolinks (2.12.6), tidyr (0.8.3), and tools (3.6.0).

RESULTS

Participant characteristics

Mean age at diagnosis for all participants was 49.9 years (Table S1). Specifically, for Nigerian women, mean age at diagnosis was 43.2 years old which was significantly younger than the mean age at diagnosis for Barbadian women (53.9, $P < 0.05$). For all participants with grade data (n= 25), 92% (n= 23) were diagnosed with intermediate and high grade (grade 2 or grade 3) carcinoma whereas only 8% (n= 2) where classified as grade 1 (Table S1). WES data were collected from The Cancer Genome Atlas (TCGA) breast cancer project, stratified by TNBC subtype and participant recorded race (African American, n=24 [TCGA_TNBC_AA]; European American, n=63 [TCGA_TNBC_EA]) for comparative analyses.

Mutation contributions and distributions

A summary of the sequencing pipeline is depicted in Figure 1. To assess genomic alterations in TNBC in WAA we performed WES on tumors yielding a mean output of 30,625 Mbases per sample and an average of 382x coverage (Table S2). An internal pool approach of germline DNA samples was derived from each patient to ensure better somatic estimation instead of using available Euro-centric references [24]. Germline DNA for each sample was individually indexed before being pooled into the final capture using the same probe sets as tumor samples (Table S2). The germline pool yielded 281,792 Mbase of data and an average of 4335x coverage showing relative germline contribution to each sample (Table S2). WES identified an average of 707 non-silent somatic mutations per tumor that was higher compared to an exome-sequenced cohort of TNBC in TCGA with a mean of 87 non-silent mutations per sample. This difference is most likely due to the residual increase in private germline variants that is contributed by the diverse African genome, as well as the larger exome capture set (this study= ~80 Mbp compared to TCGA= ~34 Mbp). Using ancestry informative markers [25] and principal component analysis, Barbadian samples independently clustered among themselves, the African Caribbean in Barbados

(ACB) group and among the Americans of African Ancestry in Southwest USA (ASW) groups (Figure S1). Nigerian samples clustered among the Yoruba in Ibadan, Nigeria (YRI) clusters and the Esan in Nigeria (ESN) clusters (Figure S1).

Somatic mutation analysis

There are 94 unique non-silent somatic variants that are enriched in Barbadian TNBC tissues, and 72 unique non-silent mutations enriched in Nigerian TNBC tissues (Figure 2A, Table S3). There were also 56 commonly mutated genes (Figure 2A) shared between the Barbadian and Nigerian study samples and 78 commonly mutated genes (Figure 2A) shared between the four cohorts included (Barbadian, Nigerian, TCGA-TNBC-EA/AA samples). Notably, the TCGA-TNBC-EA group had 2,401 genes in isolation that were not shared with other cohorts in our study. Global comparison of somatic variants with TCGA-TNBC-EA group (Figure 2B) identified 2 pseudogenes that exhibited an increase in mutation frequency in the Nigerian and Barbados cohort compared to TCGA-EA TNBC (*TNRC18P2* and *DDX12P*, $P < 0.05$; $q < 0.1$). However, there were no significantly mutated genes between our study samples and the TCGA-AA group (Figure S2A). Commonly mutated genes (at least 15% of samples harbouring a mutation) in our WAA cohorts were cancer-associated genes – *NBPF12*, *PLIN4, TP53, ZNF717, TAP1, KMT2D, PIEZO1* and *BRCA1* (Figure 2C). *HMCN2* and *MBD3L3* were also commonly mutated but not previously associated with cancer. In comparison with TCGA data for TNBC cases, these most frequently mutated genes in our WAA cohort were not as frequently mutated except for *TP53* (32% in our combined cohort; 63% in TCGA-TNBC-AA; 67% in TCGA-TNBC-EA; 63% in TCGA-Asian). For *TP53* specifically, there were five novel variants identified (HGVSc annotation: c.994-1_1023del, c.838_863del, c.368_374del,

c.844_845insT, c.382_383del) that were all predicted to have high impact as defined by PolyPhen2 [26].

COSMIC Mutational Signatures

According to somatic mutation signatures as defined by COSMIC [27], most of our samples had a moderate to strong correlation to Signatures 1, 3 and 6 (Figure 3). These signatures correspond to age, *BRCA1/BRCA2* and defective DNA mismatch repair/microsatellite instability (small INDELs) respectively. Among the Barbadian samples, there was a weak correlation for Signature 10, POLE (ultra-hypermutation) which was not seen in the Nigerian samples or the TCGA groups. There were also six Nigerian samples (54%) that showed a correlation to Signature 24 (Aflatoxin) which was not observed in the Barbadian samples or the TCGA dataset.

Comparison of TNBC copy number variation

Copy number analysis identified several regions of the genome associated with common copy number gain and loss (Figure S3, Table S7-8). Twenty-eight out of the 31 samples (90%) harbored copy number gains in *PIK3CA,* and a copy number loss for *TP53* was seen in 23 of the 31 samples (74%). Interestingly, copy number loss was seen in 30 samples (97%) for *FGFR2* and in 24 samples (77%) for *HIF1AN*. Overall, CNVs were observed for BCa-related genes (e.g., *PIK3CA, ERBB2, TP53, FOXA1*), other cancer-related genes (e.g., *ROBO2, ELN, CELF4*) and other notable genes (e.g., *DPP7, CYP26A1*). To further expand upon our understanding of bi-allelic events across genes, an integrated analysis was performed across frequently altered genes. This analysis revealed a predominance of bi-allelic (copy number loss and non-silent mutation) in *TP53* in our study (Figure 4).

DISCUSSION

Herein, we report data from WES of ancestrally-related WAA (Barbadian and Nigerian) with TNBC, which revealed pathogenic and novel variants for *TP53* and *BRCA1* as well as other BCa implicated genes such as *MDC1* (Figure 2C). This is in concordance with the high mutation rate (SNPs and CNV) for *TP53* that is typically seen in TNBC [28], and more importantly that is observed in Nigerian and American Black women with TNBC [19]. We also observed a high mutation rate for *NBPF12* and *PLIN4*. Interestingly, *in silico* analyses of BCa genomics data from TCGA and the International Cancer Genome Consortium have identified *NBPF12* as a BCa-driver gene with an estimated 0.3% substitution rate [29]. *NBPF12* belongs to the neuroblastoma breakpoint family (NBPF) of genes that are located on chromosome 1, are highly conserved across primates and are highly expressed in breast tissue [30]. *PLIN4* is located on chromosome 19 and is a member of the perilipin family that is implicated in adipocyte stability and obesity [31, 32]. Notably, high PLIN4 expression has recently been implicated in TNBC chemoresistance [33]. Although beyond the scope of this study, more functional studies need to be performed to evaluate how these genes and others identified in our study are implicated in TNBC tumorigenesis and disparity in WAA. Further studies are also needed to determine the function of the 78 genes that were commonly enriched between all study groups (Barbados, Nigeria, TCGA-AA and TCGA-EA), as they might be particularly useful for TNBC drug development regardless of ancestry.

In addition to investigating variation in individual genes, we used the COSMIC database of somatic mutations and investigated individual signatures [27], which combines base substitutions with signatures such as DNA mismatch repair. Of note, there was an enrichment of signature 24 (Aflatoxin) in 6 Nigerian samples. This signature is typically observed in a subset of hepatocellular carcinoma (HCC) liver cancers with known exposures to aflatoxin, a mycotoxin

that grows on grains across West Africa and is commonly consumed among these populations [34]. It was recently documented in a 10-year study that aflatoxin contamination in crops such as maize and groundnut are common across sub-Saharan African countries such as Nigeria [35], which may be contributing to breast and liver cancers across these sub-Saharan African nations. Two independent studies reported a considerable number of liver metastases from breast cancers in two Nigerian populations [36, 37], raising the possibility that the correlation with the aflatoxin somatic signature observed in our Nigerian TNBC cases may play a role in this phenomenon. Follow-up studies on this signature in breast tumor samples should be investigated to further delineate this relationship. Nonetheless, these findings highlight the interplay of environmental risk factors with genetics, in this case somatic mutations, and how they could lead to tumorigenic outcomes.

When investigating copy number variation, we observed a high enrichment of *PIK3CA* amplifications, which was also observed in previous studies [28]. Our analysis also revealed a predominance of bi-allelic (copy number loss and non-silent) mutations in *TP53* (Figure 4) which has been previously associated with poor outcome in multiple myeloma [38]. *In silico* analysis of *TP53* copy number loss has also highlighted its prognostic value in breast cancer [39].

Notably, almost every sample (30/31) harboured a copy number loss at 10q26.12 - q26.13 that includes the *FGFR2* gene. Multiple studies highlight an over-expression of FGFR2 and FGFR1 in TNBC [40-43] so this was an unexpected finding. Indeed, there is currently a clinical trial for inhibition of *FGFR2* (ClinicalTrials.gov identifier: NCT04526106) in solid tumors. This copy number loss of *FGFR2* highlights a novel genetic alteration in WAA not previously observed that might be protective in these populations. Further studies to investigate FGFR2

mutations and/or expression in WAA with TNBC (perhaps using different specimens - e.g. fresh frozen samples) should shed light on this phenomenon CNV data from FFPE samples can be "noisy" and CNV analyses are rapidly evolving with new tools being developed over time for better interpretation [44].

Twenty-four samples harboured a copy loss for *HIF1AN* which is the inhibitor of *HIF-1α*. In TNBC, *HIF-1α* is highly expressed and implicated in the renewal of cancer stem cells and epithelial-to-mesenchymal transition that is highly associated with metastasis [45]. The frequent copy loss of *HIF1AN* might thus be associated with the aggressive nature of TNBC observed in WAA since it prevents *HIF-1α* inhibition. Follow-up RNA sequencing, proteomic profiling and/or targeted sequencing experiments investigating the transcriptome and proteome of these WAA-TNBC cohorts will elucidate genes and pathways of interest in WAA-TNBC with therapeutic implications.

Limitations

To differentiate between somatic and germline mutations, genomic DNA is routinely extracted from peripheral blood, saliva, and adjacent healthy tissue representing germline spectrum of genomics data. This however was not possible due to the limited resource settings in Barbados and Nigeria at the time of data collection. We acknowledge this limitation and created a pooled non-tumour sample derived from the low quantity patient's germline adjacent normal breast tissue. This was the best approach instead of only relying on the current Euro-centric databases for germline subtraction with low representation of African ancestry genomics data [24]. We also performed a manual IGV [46] review of each highlighted mutation event to remove any false positives due to the higher potential of inflation in false positive somatic

calling. To address other issues with extracting high quality nucleic acids derived from FFPE tissues, we sequenced our tumour and non-tumour samples at high-resolution sequencing depth (382x and 4,335x respectively) to increase confidence in our mutation calls. Our pooled germline sample approach might be a useful application for studies of solid tumors with limited germline availability in other resource-limited populations/healthcare facilities.

Conclusions

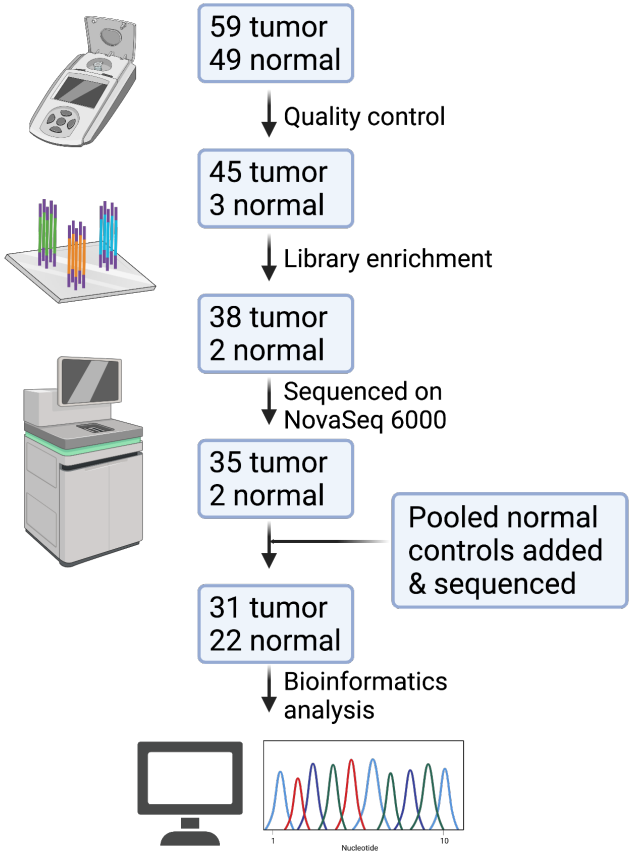
To our knowledge this is the first study to investigate the somatic mutational landscape of TNBC patients from populations with related African ancestry – West Africa (Nigeria) and Caribbean (Barbados). We identified pathogenic and likely pathogenic variants and novel variants in cancer-associated genes (e.g., *TP53*, *BRCA1*, *MDC1*) and novel mutations in other potential genes of interest (e.g., *NBPF12*, *PLIN4*, *FGFR2*). These variants may be useful for development of future therapeutic options, both unique to our WAA-TNBC cohorts and universally for all women diagnosed with TNBC. Furthermore, to better reflect our global population, more collaborative studies need to be done to increase genomic data from diverse populations within genomics databases. This would allow researchers to identify genetic risks and therapeutic options for diverse populations.

ACKNOWLEDGEMENTS

We would like to thank the staff and personnel of the Pathology and Oncology Departments at all of the sites in Barbados (QEH) and Nigeria (LUTH, UCTH, UPTH, JUTH) for their assistance with collecting samples used for this project. We would also like to thank Dr. Andrew McArthur for Bioinformatics guidance and support throughout this project.

MAIN FIGURES:

Figure 1: Overview of sequencing pipeline. DNA was extracted from 59 tumour-enriched samples and 49 adjacent uninvolved samples. After quality control, library enrichment and sequencing on NovaSeq6000, 31 tumour samples and 22 pooled normal adjacent controls were successfully sequenced. Model Diagram created with BioRender.com.



<u>Figure 2:</u> Barbadian and Nigerian women with TNBC harbour different genetic alterations than European American (EA)and African American (AA) women with TNBC. **(A)** Global analysis of all altered genes revealed that only 78 genes are shared among the four groups (Barbadian, Nigerian, TCGA-TNBC-EA, TCGA-TNBC-AA) and 2,401 genes are unique to the EA group in comparison to the other cohorts. **(B)** Global comparison of somatic variants with TCGA-TNBC-EA group identified 2 pseudogenes (*TNRC18P2* and *DDX12P*) with an increase in mutation frequency in the Nigerian and Barbadian cohorts. **(C)** The most frequently mutated genes in Barbadian and Nigerian samples were not enriched in the TCGA dataset except for *TP53* (third gene from the top of the gene list).

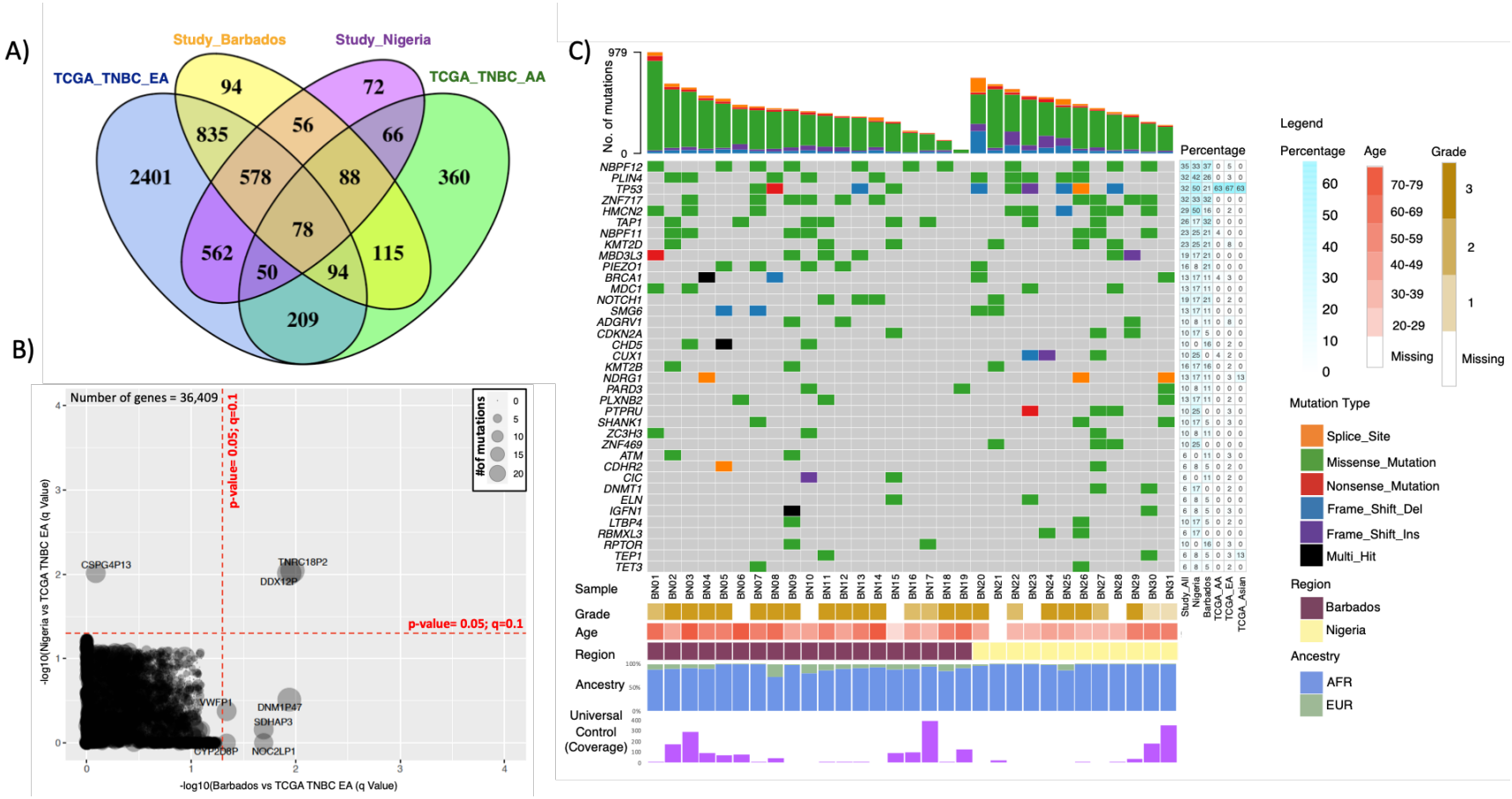


Figure 3: Mutation signature contributions for Barbadian and Nigerian TNBC samples show high correlation to COSMIC Signatures 1, 3 and 6. Using COSMIC somatic mutation signatures, Age, BRCA1/BRCA2 and Defective DNA MMR/MSI (small INDELS) were enriched signatures for Barbadian and Nigerian samples. These signatures were also enriched in TCGA-AA, TCGA-EA women with TNBC and overall, all breast cancer cases within TCGA.

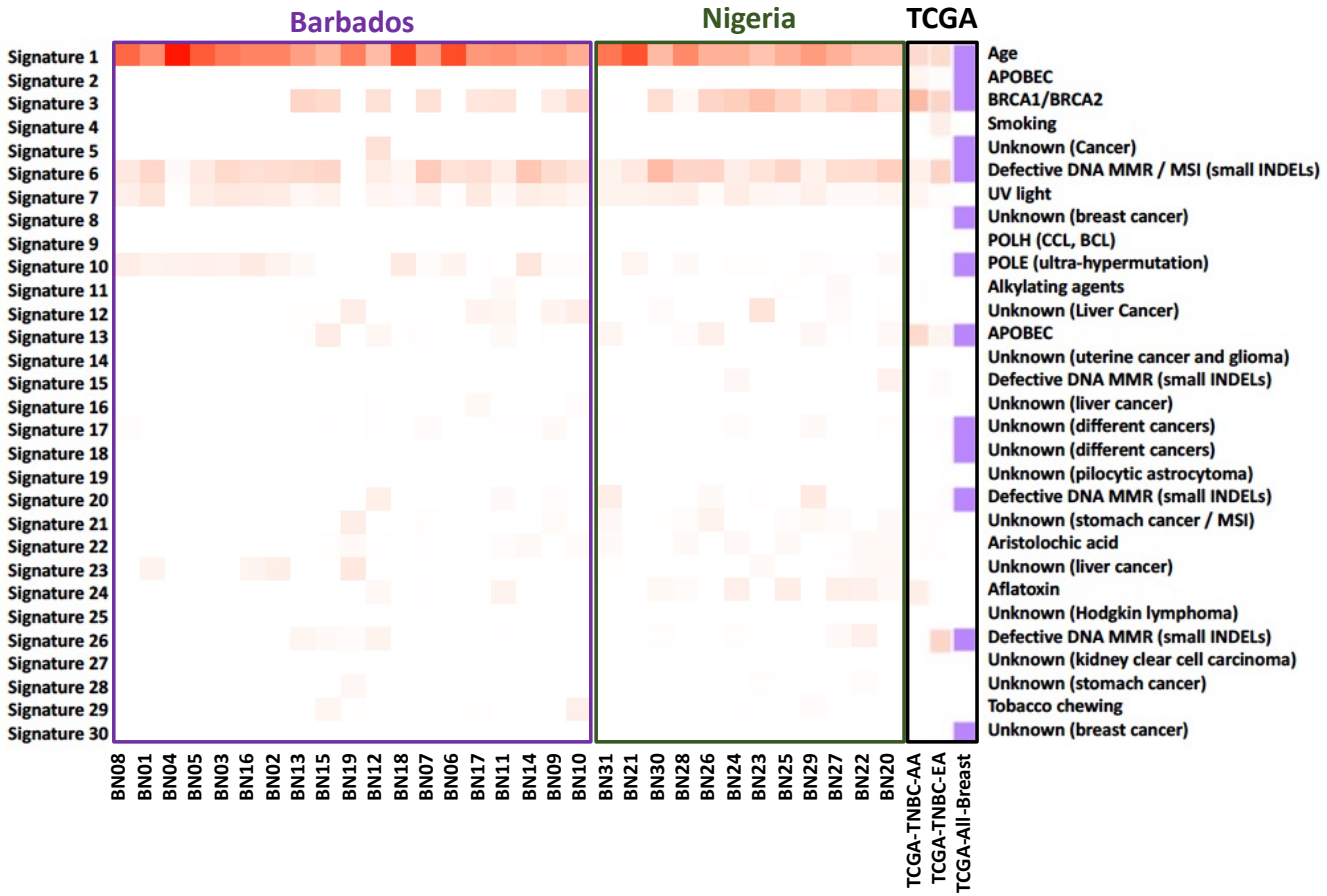
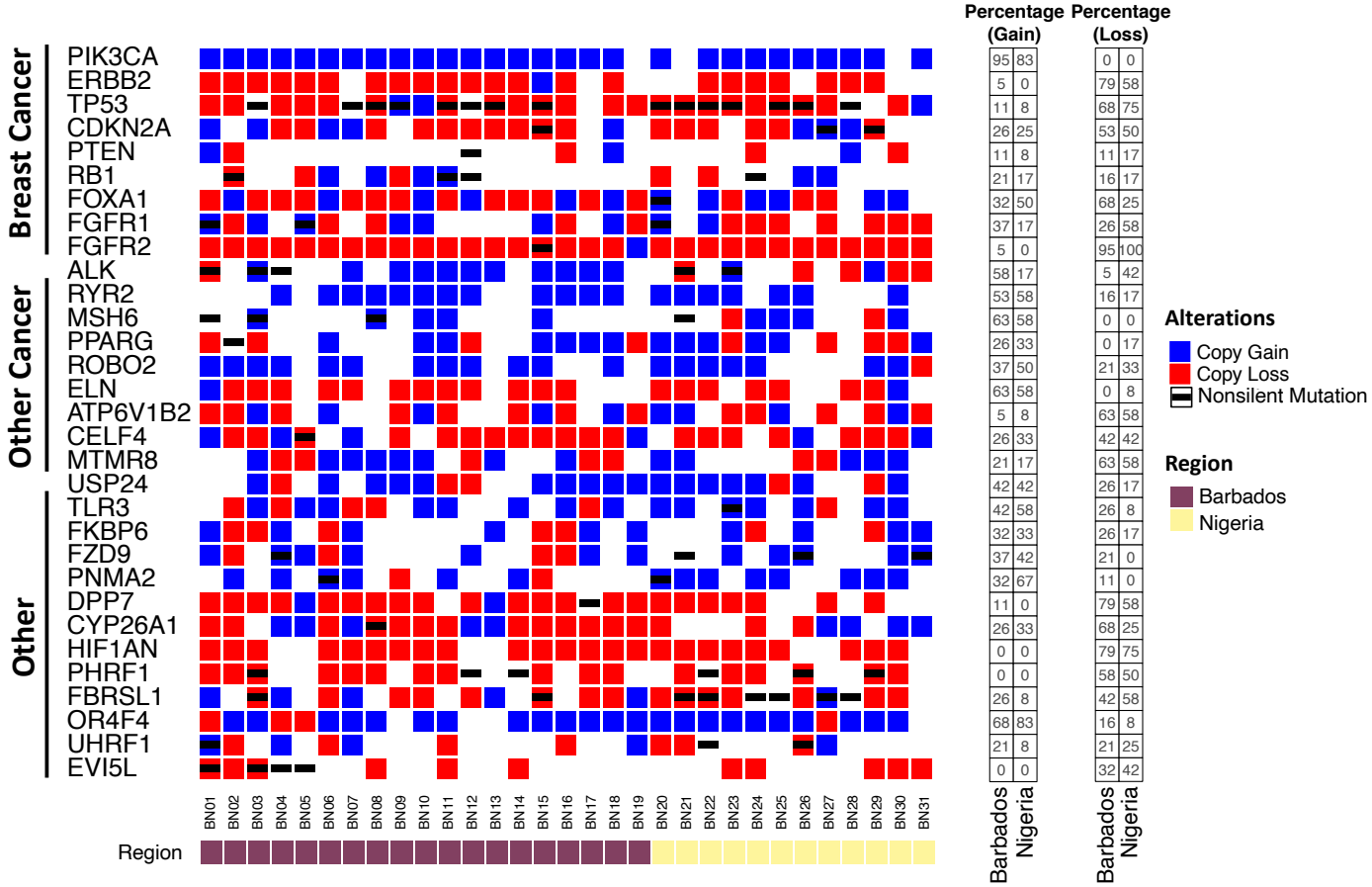


Figure 4: Copy number variation (CNV) analysis revealed no differences between Barbadian and Nigerian samples. Using NEXUS Copy Number (v10) analysis toolkits, most common CNV were investigated. The copy number gain across each sample is presented in blue and copy number loss in red. The black line indicates a non-silent somatic mutation. The bottom panel indicates region of sample origin. Left panels show percent alterations across each copy number event across two different populations. Bi-allelic mutation illustrated by black bar and either blue or red square indicating non-silent somatic mutation and either copy gain or less respectively. No significance ($p < 0.05$) was detected in this cohort of samples.



REFERENCES

1. Ferlay, J., *et al.*, *Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods.* Int J Cancer, 2019. **144**(8): p. 1941-1953.
2. Prat, A., *et al.*, *Clinical implications of the intrinsic molecular subtypes of breast cancer.* Breast, 2015. **24 Suppl 2**: p. S26-35.
3. Li, X., *et al.*, *Triple-negative breast cancer has worse overall survival and cause-specific survival than non-triple-negative breast cancer.* Breast Cancer Res Treat, 2017. **161**(2): p. 279-287.
4. Hercules, S.M., *et al.*, *High triple-negative breast cancer prevalence and aggressive prognostic factors in Barbadian women with breast cancer.* Cancer, 2020. **126**(10): p. 2217-2224.
5. Bowen, R.L., *et al.*, *Early onset of breast cancer in a group of British black women.* Br J Cancer, 2008. **98**(2): p. 277-81.
6. DeSantis, C.E., *et al.*, *Breast cancer statistics, 2015: Convergence of incidence rates between black and white women.* CA Cancer J Clin, 2016. **66**(1): p. 31-42.
7. Linnenbringer, E., *et al.*, *Associations between breast cancer subtype and neighborhood socioeconomic and racial composition among Black and White women.* Breast Cancer Res Treat, 2020. **180**(2): p. 437-447.
8. Qin, B., *et al.*, *Neighborhood social environmental factors and breast cancer subtypes among Black women.* Cancer Epidemiol Biomarkers Prev, 2020.
9. Williams, D.R., S.A. Mohammed, and A.E. Shields, *Understanding and effectively addressing breast cancer in African American women: Unpacking the social context.* Cancer, 2016. **122**(14): p. 2138-49.
10. Tutt, A., *et al.*, *Carboplatin in BRCA1/2-mutated and triple-negative breast cancer BRCAness subgroups: the TNT Trial.* Nat Med, 2018. **24**(5): p. 628-637.
11. Pereira, B., *et al.*, *The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes.* Nat Commun, 2016. **7**: p. 11479.
12. Yu, B., S.A. O'Toole, and R.J. Trent, *Somatic DNA mutation analysis in targeted therapy of solid tumours.* Transl Pediatr, 2015. **4**(2): p. 125-38.
13. Yin, L., *et al.*, *Triple-negative breast cancer molecular subtyping and treatment progress.* Breast Cancer Res, 2020. **22**(1): p. 61.
14. Yuan, J., *et al.*, *Integrated Analysis of Genetic Ancestry and Genomic Alterations across Cancers.* Cancer Cell, 2018. **34**(4): p. 549-560.e9.

15. Brewster, A.M., M. Chavez-MacGregor, and P. Brown, *Epidemiology, biology, and treatment of triple-negative breast cancer in women of African ancestry.* Lancet Oncol, 2014. **15**(13): p. e625-e634.

16. Davis, M., *et al.*, *Identification of Distinct Heterogenic Subtypes and Molecular Signatures Associated with African Ancestry in Triple Negative Breast Cancer Using Quantified Genetic Ancestry Models in Admixed Race Populations.* Cancers, 2020. **12**(5).

17. Martini, R., *et al.*, *Investigation of triple-negative breast cancer risk alleles in an International African-enriched cohort.* Sci Rep, 2021. **11**(1): p. 9247.

18. Newman, L.A., *et al.*, *Hereditary Susceptibility for Triple Negative Breast Cancer Associated With Western Sub-Saharan African Ancestry Results From an International Surgical Breast Cancer Collaborative.* Annals of Surgery, 2019. **270**(3): p. 484-492.

19. Pitt, J.J., *et al.*, *Characterization of Nigerian breast cancer reveals prevalent homologous recombination deficiency and aggressive molecular features.* Nat Commun, 2018. **9**(1): p. 4181.

20. George, S.H.L., *et al.*, *Gene Sequencing for Pathogenic Variants Among Adults With Breast and Ovarian Cancer in the Caribbean.* JAMA Netw Open, 2021. **4**(3): p. e210307.

21. Murray, T., *et al.*, *African and non-African admixture components in African Americans and an African Caribbean population.* Genet Epidemiol, 2010. **34**(6): p. 561-8.

22. Benn-Torres, J., *et al.*, *Admixture and population stratification in African Caribbean populations.* Ann Hum Genet, 2008. **72**(Pt 1): p. 90-8.

23. Cropp, C.D., *et al.*, *8q24 risk alleles and prostate cancer in African-Barbadian men.* Prostate, 2014. **74**(16): p. 1579-88.

24. Halperin, R.F., *et al.*, *A method to reduce ancestry related germline false positives in tumor only somatic variant calling.* BMC Med Genomics, 2017. **10**(1): p. 61.

25. Manojlovic, Z., *et al.*, *Comprehensive molecular profiling of 718 Multiple Myelomas reveals significant differences in mutation frequencies between African and European descent cases.* PLoS Genet, 2017. **13**(11): p. e1007087.

26. Adzhubei, I., D.M. Jordan, and S.R. Sunyaev, *Predicting functional effect of human missense mutations using PolyPhen-2.* Curr Protoc Hum Genet, 2013. **Chapter 7**: p. Unit7.20.

27. Forbes, S.A., *et al.*, *COSMIC: High-Resolution Cancer Genetics Using the Catalogue of Somatic Mutations in Cancer.* Curr Protoc Hum Genet, 2016. **91**: p. 10.11.1-10.11.37.

28. Shi, Y., *et al.*, *Therapeutic landscape in mutational triple negative breast cancer.* Mol Cancer, 2018. **17**(1): p. 99.

29. Rajendran, B.K. and C.X. Deng, *Characterization of potential driver mutations involved in human breast cancer by computational approaches.* Oncotarget, 2017. **8**(30): p. 50252-50272.

30. Vandepoele, K., *et al.*, *A novel gene family NBPF: intricate structure generated by gene duplications during primate evolution.* Mol Biol Evol, 2005. **22**(11): p. 2265-74.

31. Wolins, N.E., *et al.*, *Adipocyte protein S3-12 coats nascent lipid droplets.* J Biol Chem, 2003. **278**(39): p. 37713-21.

32. Richardson, K., *et al.*, *The PLIN4 variant rs8887 modulates obesity related phenotypes in humans through creation of a novel miR-522 seed site.* PLoS One, 2011. **6**(4): p. e17944.

33. Sirois, I., *et al.*, *A Unique Morphological Phenotype in Chemoresistant Triple-Negative Breast Cancer Reveals Metabolic Reprogramming and PLIN4 Expression as a Molecular Vulnerability.* Mol Cancer Res, 2019. **17**(12): p. 2492-2507.

34. Ladep, N.G., *et al.*, *Problem of hepatocellular carcinoma in West Africa.* World J Hepatol, 2014. **6**(11): p. 783-92.

35. Bandyopadhyay, R., *et al.*, *"Ground-Truthing" Efficacy of Biological Control for Aflatoxin Mitigation in Farmers' Fields in Nigeria: From Field Trials to Commercial Usage, a 10-Year Study.* Front Microbiol, 2019. **10**: p. 2528.

36. Adisa, A.O., *et al.*, *Metastatic breast cancer in a Nigerian tertiary hospital.* Afr Health Sci, 2011. **11**(2): p. 279-84.

37. Abur, P.P., L.M. Yusufu, and V.I. Odigie, *Pattern of visceral metastasis from breast cancer patients in Ahmadu Bello University Teaching Hospital, Zaria, North Western Nigeria.* Archives of Medicine and Surgery, 2019. **4**(2): p. 31.

38. Thanendrarajan, S., *et al.*, *The level of deletion 17p and bi-allelic inactivation of.* Haematologica, 2017. **102**(9): p. e364-e367.

39. Smith, J.C. and J.M. Sheltzer, *Systematic identification of mutations and copy number alterations associated with cancer patient prognosis.* Elife, 2018. **7**.

40. Jafarian, A.H., *et al.*, *The Relationship Between Fibroblastic Growth Factor Receptor-1 (FGFR1) Gene Amplification in Triple Negative Breast Carcinomas and Clinicopathological Prognostic Factors.* Iran J Pathol, 2019. **14**(4): p. 299-304.

41. Kim, S., *et al.*, *FGFR2 promotes breast tumorigenicity through maintenance of breast tumor-initiating cells.* PLoS One, 2013. **8**(1): p. e51671.

42. Cunningham, D.L., *et al.*, *Differential responses to kinase inhibition in FGFR2-addicted triple negative breast cancer cells: a quantitative phosphoproteomics study.* Sci Rep, 2020. **10**(1): p. 7950.

43. Fletcher, M.N., *et al.*, *Master regulators of FGFR2 signalling and breast cancer risk.* Nat Commun, 2013. **4**: p. 2464.

44. Filia, A., *et al.*, *High-Resolution Copy Number Patterns From Clinically Relevant FFPE Material.* Sci Rep, 2019. **9**(1): p. 8908.

45. Cecil, D.L., *et al.*, *Immunization against HIF-1α Inhibits the Growth of Basal Mammary Tumors and Targets Mammary Stem Cells.* Clin Cancer Res, 2017. **23**(13): p. 3396-3404.

46. Robinson, J.T., *et al.*, *Integrative genomics viewer.* Nat Biotechnol, 2011. **29**(1): p. 24-6.

47. Allred, D.C., *et al.*, *Prognostic and predictive factors in breast cancer by immunohistochemical analysis.* Mod Pathol, 1998. **11**(2): p. 155-68.

48. Zhrebker, L., *et al.*, *Case report: whole exome sequencing of primary cardiac angiosarcoma highlights potential for targeted therapies.* BMC Cancer, 2017. **17**(1): p. 17.

49. Cibulskis, K., *et al.*, *Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples.* Nat Biotechnol, 2013. **31**(3): p. 213-9.

50. Kandoth, C., *et al.*, *mskcc/vcf2maf: vcf2maf v1.6.16*. 2018.

51. Mayakonda, A., *et al.*, *Maftools: efficient and comprehensive analysis of somatic variants in cancer.* Genome Res, 2018. **28**(11): p. 1747-1756.

52. Diaz-Gay, M., *et al.*, *Mutational Signatures in Cancer (MuSiCa): a web application to implement mutational signatures analysis in cancer samples.* BMC Bioinformatics, 2018. **19**(1): p. 224.

53. Mermel, C.H., *et al.*, *GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers.* Genome Biol, 2011. **12**(4): p. R41.

54. Genomes Project, C., *et al.*, *A global reference for human genetic variation.* Nature, 2015. **526**(7571): p. 68-74.

Supporting information captions:

Figure S1: Ancestry of Barbadian and Nigerian samples confirmed by principal component analysis. Principal component plot across all samples by location of where samples were collected Barbadian (black squares) and Nigerian (black triangles) using ancestry informative markers derived from the whole-exome sequencing. African and European sub-populations from 1000 Genomes Project phase 3 populations were used as anchor populations.

Figure S2: Barbadian and Nigerian women with TNBC harbour similar frequencies of genetic alterations with African American women with TNBC. Global comparison of somatic variants with TCGA-TNBC-AA group identified did not identify any changes in mutation frequency in the Nigerian and Barbados cohorts.

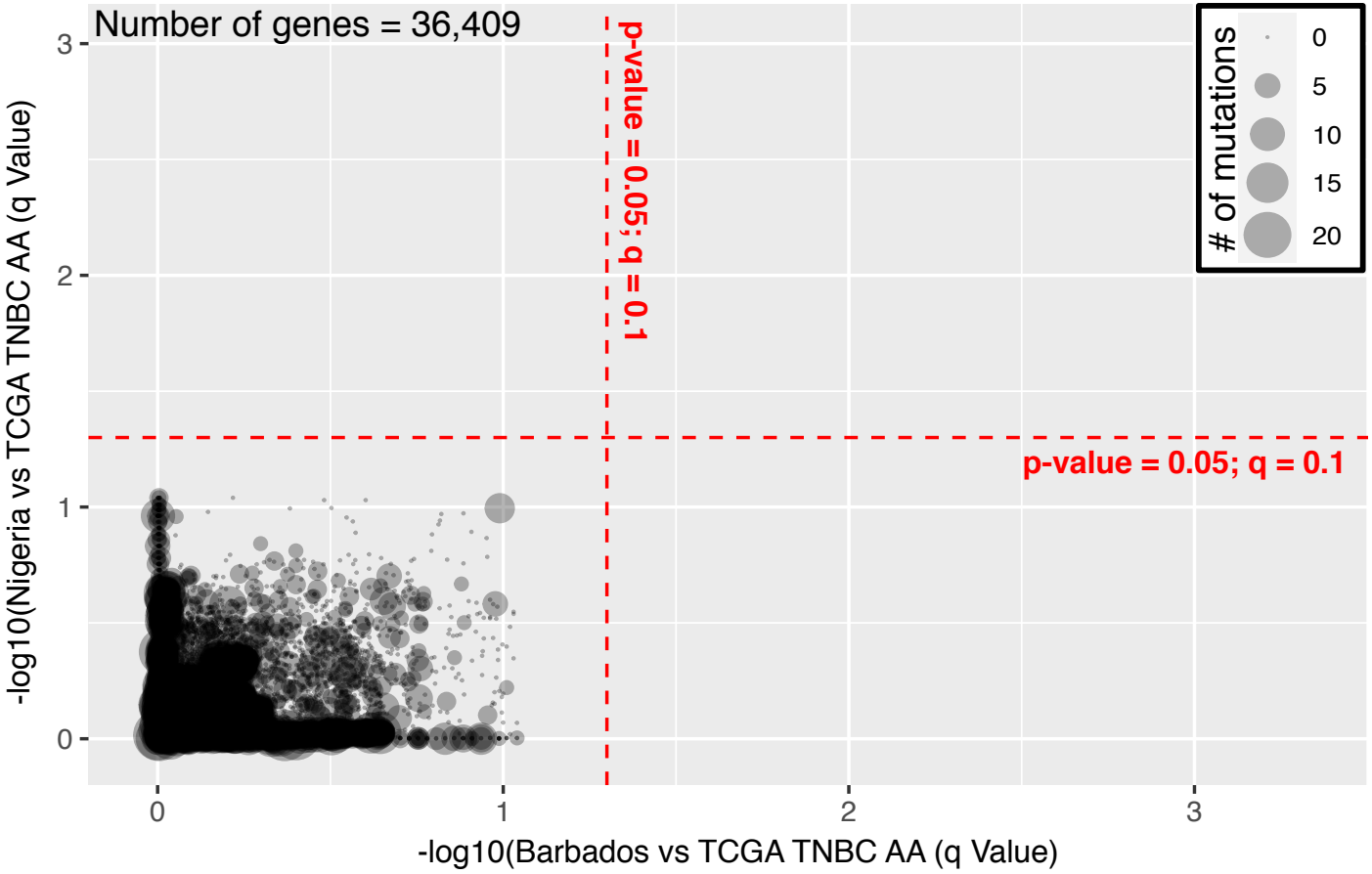
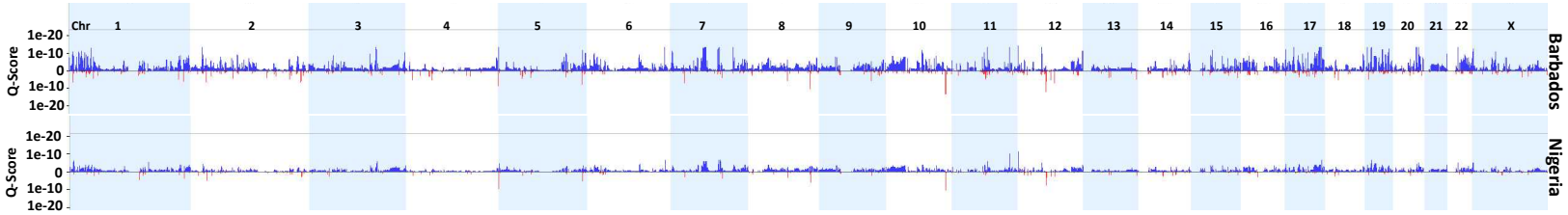


Figure S3: Copy number analysis across each chromosome between Barbadian and Nigerian participants reveal no major differences across the exome. Focal copy number analysis was performed using GISTIC (v2.0) and revealed no major differences. The plot shows the q-score, with the copy number gains indicated in blue and copy number losses in red. Chromosome positions are indicated along the x-axis.



CHAPTER 6: DANCING FROM BOTTOMS UP—ROLES OF THE POZ-ZF TRANSCRIPTION FACTOR KAISO IN CANCER.

Preface

This chapter consists of the published article entitled: ***"Dancing from bottoms up—roles of the POZ-ZF transcription factor Kaiso in Cancer"*** (Pierre *et al.*, 2019) by Pierre CC*, Hercules SM*, Yates C and Daniel JM. (Biochim Biophys Acta Rev Cancer, 1871(1), 64-74. doi:10.1016/j.bbcan.2018.10.005) in its original form. This is not an open-access article.

In this article we summarise and discuss all that was known up to the time of publication about Kaiso's function and role as a transcription factor, and its postulated roles in various cancer sites (colorectal, lung, prostate, breast, glioma and chronic myeloid leukemia) and the racial disparities in breast and prostate cancer incidence and outcomes in people of African ancestry.

Contributions:

Dr. JM Daniel and Dr. CC Pierre conceived the review and co-wrote the manuscript with SM Hercules. Dr. C Yates co-wrote one section of the manuscript (Kaiso and prostate cancer). SM Hercules created all figures and tables within the manuscript. Dr. JM Daniel also provided significant intellectual guidance throughout the review process. All co-authors edited the manuscript.

BBA - Reviews on Cancer 1871 (2019) 64–74

Contents lists available at ScienceDirect

BBA - Reviews on Cancer

journal homepage: www.elsevier.com/locate/bbacan

Review

Dancing from bottoms up – Roles of the POZ-ZF transcription factor Kaiso in Cancer

Christina C. Pierre[a,1,2], Shawn M. Hercules[a,2], Clayton Yates[b], Juliet M. Daniel[a,*]

[a] Department of Biology, McMaster University, Hamilton, Ontario L8S 4K1, Canada
[b] Department of Biology, Center for Cancer Research, Tuskegee University, Tuskegee, AL, USA

ARTICLE INFO

Keywords:
Kaiso
POZ-ZF transcription factors
Racial disparities in Cancer
EMT
Inflammation

ABSTRACT

The POZ-ZF transcription factor Kaiso was discovered two decades ago as a binding partner for p120^{ctn}. Since its discovery, roles for Kaiso in diverse biological processes (epithelial-to-mesenchymal transition, apoptosis, inflammation) and several signalling pathways (Wnt/β-catenin, TGFβ, EGFR, Notch) have emerged. While Kaiso's biological role in normal tissues has yet to be fully elucidated, Kaiso has been increasingly implicated in multiple human cancers including colon, prostate, ovarian, lung, breast and chronic myeloid leukemia. In the majority of human cancers investigated to date, high Kaiso expression correlates with aggressive tumor characteristics including proliferation and metastasis, and/or poor prognosis. More recently, interest in Kaiso stems from its apparent correlation with racial disparities in breast and prostate cancer incidence and survival outcomes in people of African Ancestry. This review discusses Kaiso's role in various cancers, and Kaiso's potential for driving racial disparities in incidence and/or outcomes in people of African ancestry.

1. Introduction

The POZ-ZF transcription factor Kaiso (whose unique name is derived from the "Kaiso" genre of Caribbean music that has its origins in West Africa and is associated with dancing) was discovered almost two decades ago in association with p120 catenin (p120^{ctn}), an Armadillo catenin responsible for regulating E-cadherin stability and turnover in adherens junctions [1]. At the time, several members of the POZ-ZF transcription factor family had been characterized as mediators of vertebrate development and cancer, hinting that Kaiso may also function in these processes [2]. As more evidence emerged surrounding the tumor suppressing role of E-cadherin, and the critical role of p120^{ctn} in regulating E-cadherin function, significant efforts were made and are still ongoing to determine how Kaiso might function in tumorigenesis, perhaps as a regulator of E-cadherin– or p120^{ctn}– related functions. Interestingly, early studies of Kaiso's function in *Xenopus* development identified a role for Kaiso in antagonizing canonical Wnt signalling [3–5] and this set the stage for studies aimed at investigating Kaiso's modulation of this pathway in the context of cancer. To date, multiple studies have implicated Kaiso in several different human cancers, however characterizing Kaiso as a *bona fide* tumor suppressor or oncogene has been challenging, as Kaiso's function appears to be highly context-dependent. Nonetheless, diverse roles for Kaiso in key cancer-related processes and signalling pathways have been identified, revealing exciting avenues for further research. In this review, we will discuss some of the insight that has been gleaned into Kaiso's structure and function and explore its identified roles in tumorigenesis as well as potential areas for continued study.

1.1. An overview of Kaiso structure, DNA binding and transcriptional activity

As is characteristic of the POZ-ZF family of proteins, Kaiso possesses an amino-terminal BTB/POZ (Broad complex, Tramtrack and Bric à brac/Poxvirus and Zinc finger, hereafter POZ) protein-protein interaction domain and a carboxy-terminal zinc finger (ZF) domain (Fig. 1). The highly conserved POZ domain mediates homo- and heterodimeric interactions with other POZ and non-POZ proteins, whereas the ZF domain mediates DNA-binding, reviewed in [2]. Kaiso also contains several putative Serine/Threonine phosphorylation sites that have remained relatively understudied, although we and others have routinely observed a doublet consistent with a phosphorylated form of Kaiso in immunoblot analysis. Indeed, preliminary studies from the Daniel Lab have determined that Kaiso is predominantly phosphorylated on

* Corresponding author at: Dept. of Biology, LSB-331, McMaster University, 1280 Main St. West., Hamilton, ON L8S 4K1, Canada.
E-mail address: danielj@mcmaster.ca (J.M. Daniel).
[1] Current address: Department of Pathology, University of Virginia, Charlottesville, VA, USA.
[2] Co-first authors.

https://doi.org/10.1016/j.bbcan.2018.10.005
Received 15 August 2018; Received in revised form 5 October 2018; Accepted 7 October 2018
Available online 09 November 2018
0304-419X/

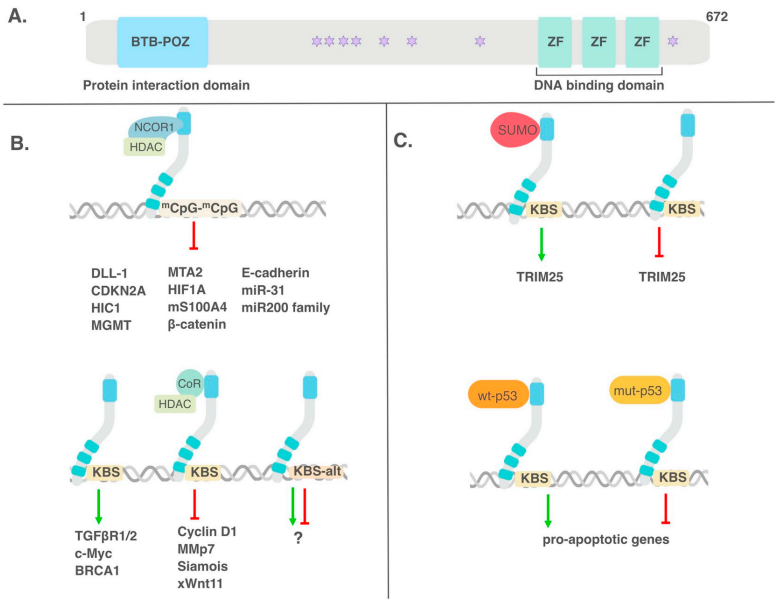


Fig. 1. Kaiso as a transcriptional activator and repressor. (A) Kaiso possesses a BTB/POZ protein-protein interaction domain and three DNA-binding ZFs. Serine/Threonine phosphroylation sites indicated by stars. (B) Kaiso negatively regulates expression of many genes in a methylation-dependent manner and via the core KBS. Kaiso also positively regulates expression of other genes via the core KBS. Kaiso's r (C) Kaiso also switches between an activator or repressor dependent on its SUMOylation status. When SUMOylated, Kaiso acts as an activator but when de-SUMOylated, Kaiso acts as a repressor. Moreover, when Kaiso interacts with wild-type p53, it was found to activate pro-apoptotic genes whereas when associated with mutant p53, Kaiso repressed pro-apoptotic genes.

Table 1
Correlation of Kaiso expression with histopathological/clinical features of cancers.

Cancer Site	Methodology	Histopathological/ Clinical Correlation	References
Colorectal	Colon cancer TMA, GEO database	Higher Kaiso expression observed in primary & metastatic tumor biopsies compared to normal colon tissues	Pierre *et al.* [48]
Prostate	PCa TMA	Higher Kaiso expression malignant prostate tumors compared to BPH; high Kaiso expression correlated with high tumor grade, Gleason score & AAs	Jones *et al.* [14]
	TCGA database	High Kaiso/low mir-31 expression correlated with lower overall survival relative to patients with either high Kaiso only or miR-31 low only tumors	Wang *et al.* [16]
Lung	NSCLC TMA	Higher cytoplasmic expression correlated with advanced stage lung cancer & poorer overall patient survival; higher cytoplasmic expression in tissues of NSCLC patients compared to NBE patient tissues	Dai *et al.* [28]
Pancreas	TMA; ONCOMINE database	High Kaiso expression in PDAC tissues compared with low Kaiso expression in normal/adjacent tissues; high Kaiso expression in PDAC tissues correlated with higher grade & tumor size	Jones *et al.* [98]
Breast	Breast cancer TMA	Higher nuclear Kaiso expression associated with higher nuclear tumor grade	Vermeulen *et al.* [29]
	Breast cancer TMA	High nuclear Kaiso expression positively correlated with invasion, lymph node metastases & poorer /reduced overall survival of patients with IDC	Jones *et al.* [13]
	TCGA database	High Kaiso/BRCA1 expression correlated with lower overall survival of patients diagnosed with TNBC & other BC subtypes	Bassey-Archibong *et al.* [23]
	TNBC TMA	Higher nuclear Kaiso expression observed in TNBC tissues obtained from WAA compared to Caucasian women	Bassey-Archibong *et al.* [97]

Serine/Threonine residues (our unpublished data).

In our effort to fully describe the complexity of Kaiso's biological function, it would be remiss not to mention the ongoing debates surrounding Kaiso's DNA binding site(s) and its function as a transcriptional activator or repressor. Three different DNA sites have been reported as *bona fide* Kaiso binding sites, although to date, there is a lack of agreement as to which is the most favored or preferred Kaiso binding site. During our initial characterization of Kaiso, we identified a consensus DNA binding site, TC**TGCNA** (core sequence bolded; N is any nucleotide) that was termed the Kaiso Binding Site (KBS) and we demonstrated an interaction between recombinant Kaiso proteins and KBS's in the *matrilysin* promoter by EMSA [6]. In the same study, we also confirmed Kaiso's binding to methylated CpG dinucleotides in oligonucleotides derived from the *S100A4/Metastasin* gene promoter [6], which was confirmed by Prokhortchouk *et al.* in an independent study [7]. These findings led to Kaiso being classified as a dual-specificity DNA-binding transcription factor [6]. Further analyses by our lab to determine the relative affinity of Kaiso for the KBS versus methyl-CpG (me-CpG) dinucleotides suggested that Kaiso possessed higher affinity for the KBS than for me-CpG dinucleotides [6]. However, a 2012

study by Raghav *et al.* – which found that promotor-proximal tethering of SMRT in terminal adipogenesis is mediated by Kaiso through a conserved, methylated TCTCGCGAGA motif – demonstrated a stronger interaction between Kaiso and the methylated TCTCGCGAGA motif than between Kaiso and the originally identified KBS [8]. In this study, the nucleotides flanking the core methylated CGCG motif were also shown to strengthen Kaiso's binding to the me-CpG motif, as mutation of these flanking sequences reduced the affinity of Kaiso for this site [8]. Following these early studies however, the relative affinity of Kaiso for the KBS versus me-CpG dinucleotides continues to be a topic of debate.

Through the analysis of Kaiso ChIP-seq datasets downloaded from the UCSC (University of California Santa Cruz) browser (from the ENCODE consortium), Blattler *et al.* demonstrated three key attributes of Kaiso's endogenous interaction with DNA [9]. First, ~ 77% of Kaiso peaks were shown to overlap with RNA polymerase II peaks and enriched for active histone modifications, which suggested that Kaiso mostly binds to DNA regions that are actively transcribed or primed for transcription. Second, 36–43% of Kaiso peaks contained the TCTCGC GAGA motif identified by Raghav *et al.* [8], while the original KBS was not identified in any of the Kaiso peak sets. This finding suggests that additional studies should be conducted to clarify Kaiso's DNA-binding properties and the importance/contributions of these different motifs to Kaiso's respective downstream biological functions. Notably, most sites bound by Kaiso, including those containing the TCTCGCGAGA motif, were found to be unmethylated, indicating that Kaiso may prefer to bind to unmethylated DNA sites for transcriptional activity [9].

Irrespective of which site is the preferred or "correct" site for Kaiso binding, further research to clarify and identify Kaiso's binding sites is warranted to enable a complete understanding of Kaiso's transcriptional properties and function. This will also require an in-depth analysis of the role of the Kaiso-p120^{ctn} interaction, since the interaction of p120^{ctn} with Kaiso's DNA-binding domain results in inhibition of Kaiso's transcriptional activity and repression of target genes in the context of KBS-containing gene promoters [1,10]. The cellular contexts under which p120^{ctn} interacts with Kaiso are still not fully understood, and thus determining the upstream signalling pathways that regulate the Kaiso-p120^{ctn} interaction will be essential for fully understanding Kaiso's role in normal and tumorigenic processes.

Another layer of complexity with regards to Kaiso's transcriptional role in normal and tumorigenic processes was heralded in 2006 upon the identification of two Kaiso-like proteins, ZBTB4 and ZBTB38 [11]. Intriguingly, ZBTB4, like Kaiso, was found to exhibit dual-specificity DNA binding to the KBS consensus sequence and me-CpG dinucleotides, while ZBTB38 was found to only bind me-CpG dinucleotides [11,12]. Despite their structural similarity, Kaiso did not heterodimerize with ZBTB38 or ZBTB4, although an interaction between ZBTB38 and ZBTB4 was observed. Furthermore, no interaction between p120^{ctn} and either of these two Kaiso-like proteins was observed [11]. Nonetheless, Kaiso, ZBTB4 and ZBTB38 were classified as a new family of methylated DNA-binding transcription factors that regulate gene expression and may function redundantly depending on cell or tissue context [11].

1.2. Kaiso - transcriptional repressor, activator or both?

Kaiso's binding to methylated CpG dinucleotides, coupled with studies demonstrating interactions between Kaiso and SMRT-containing corepressor complexes (Fig. 1), led to its initial characterization as a transcriptional repressor [7,8]. Indeed, most studies to date have reported on Kaiso's transcriptional repression of target genes including *E-cadherin* [13,14], *Wnt 11* [4], *matrilysin* [15], *HIF1A* [16], *CDKN2A* [17], *miR-31* [18] and the *miR-200* family [19]. While fewer studies have reported on Kaiso-mediated transcriptional activation of target genes [20,21], recent studies suggest that transcriptional activation may be Kaiso's preferred mode of transcriptional regulation. In support of this notion, we have found that twice as many genes are downregulated in response to Kaiso depletion in HCT-116 colon carcinoma cells than those that are upregulated (our unpublished data).

The notion that Kaiso may function as both an activator and repressor of transcription is not unprecedented since another POZ-ZF protein, Miz-1, has been reported to demonstrate both transcriptional repression and activation activities depending on the cofactors with which it interacts [22]. We have observed a similar phenomenon with Kaiso; its interaction with wild-type p53 activates transcription of pro-apoptotic genes, while its interaction with mutant p53 potentially represses transcription of pro-apoptotic genes (Fig. 1) [23,24]. Additionally, a recent study showed that while Kaiso's transcriptional activities were mostly repressive in HeLa cells, many pathways were also activated in the presence of Kaiso [25]. This study also found opposing transcriptional activities of Kaiso between two cell lines, confirming that Kaiso DNA-binding and transcriptional activity is context and cell-type specific [25].

A recent study by Zhenilo *et al.* investigated the effect of post translational modifications on Kaiso's transcriptional activity [26] and found that covalent linkage of small ubiquitin-like modifier (SUMO) polypeptides functioned as a molecular switch to regulate Kaiso's function as an activator or repressor [26]. When SUMOylated, Kaiso acted as an activator but when deSUMOylated, Kaiso acted as a repressor (Fig. 1). Using HEK293 cells, the authors demonstrated that Kaiso is SUMOylated under normal conditions at lysine 93 (K93); however, under hyperosmotic stress, Kaiso was found to be deSUMOylated [26]. This was the first study to elucidate a role for post-translational modifications on Kaiso function and indicates that further work is needed to examine other mechanisms (e.g. phosphorylation) through which Kaiso could switch between its activating and repressive functions.

It is noteworthy that to date, studies reporting on Kaiso's transcriptional activation all focus on Kaiso's role when bound to the KBS. Thus, future studies are needed to elucidate whether the TCTCGCGAGA motif is also functionally linked to Kaiso-mediated transcriptional activation. Regardless of Kaiso's specific transcriptional properties or preferred mode of DNA binding, it is clear that Kaiso and many of its putative target genes are implicated in or play key roles in tumorigenesis.

2. A Jack of all trades: the multifaceted functions of Kaiso in tumor development and progression

In a seminal review, Hanahan and Weinberg summarized the key acquired capabilities that are necessary for tumor growth and progression [27]. As we continue to unravel the details of Kaiso's function, it has become evident that Kaiso is able to mediate several of these vital tumor-acquired capabilities including invasion, metastasis, apoptosis, cell proliferation, and inflammation (Fig. 2; key references listed in Table 2). Thus, not surprisingly, Kaiso has been increasingly implicated in various human cancers (colon, prostate, lung, breast, ovarian, chronic myeloid leukemia) [17, 28–31]. Here we review in depth, studies investigating Kaiso's role in various cancers and tumor-related pathways.

2.1. Kaiso in Colorectal Cancer - Wnt pathway regulation or inflammation?

During the initial characterization of Kaiso as a transcription factor, several independent studies reported that Kaiso repressed a subset of Wnt/β-catenin target genes and inhibited Wnt signalling in several model systems [15,32]. Park and colleagues were the first to report a role for Kaiso in the Wnt signalling cascade, a well-established cascade whose malfunction plays a key role in colon cancer [33]. Using a *Xenopus* model, they demonstrated that morpholino-mediated Kaiso depletion resulted in increased Wnt reporter activity and increased expression of several Wnt signalling targets including *siamois*, *c-Myc*, *c-Fos* and *cyclinD1* [32]. Notably, Kaiso-mediated inhibition of Wnt target gene expression was attenuated by p120^{ctn} overexpression [32]. Importantly, Kaiso was shown to rescue the duplicate-axis phenotype

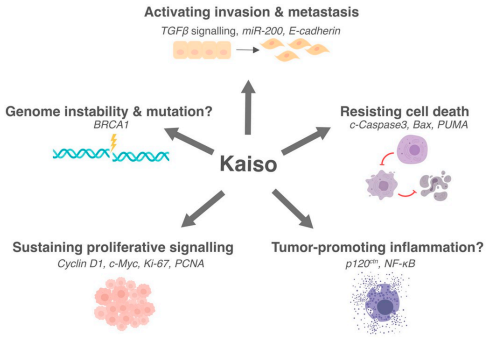


Fig. 2. Kaiso regulates several tumor-related processes as defined by Hanahan & Weinberg [27]. Kaiso activates tumor cell invasion and metastasis via regulation of TGFβ signalling, miR-200 family and E-cadherin expression. Kaiso sustains proliferative signalling through Cyclin-D1,c-Myc, Ki-67 and PCNA. As observed in several mouse models, Kaiso promotes inflammation. Kaiso is also implicated in resisting cell death through the regulation of c-Caspase3, Bax and PUMA. Finally, Kaiso is implicated in genome instability through its regulation of BRCA1 (Table 1.

induced by constitutive Wnt signalling [32]. Based on these studies it was postulated that cross-talk existed between the p120^{ctn}-Kaiso signalling trajectory and the canonical Wnt signalling pathway [34]. However, the interplay of Kaiso and p120^{ctn} with canonical Wnt signalling in *Xenopus* development was further complicated by a later study that implicated Kaiso as a bimodal regulator of Wnt signalling [35]. In this study, both Kaiso depletion and overexpression inhibited Wnt signalling, while mild ectopic Kaiso expression resulted in increased Wnt signalling activity [35]. Unfortunately, none of these studies examined the expression levels or relevance of the Kaiso-like protein ZBTB4 on Wnt signalling in these models, and thus their apparent paradoxical findings may be explained by functional redundancy or competition with ZBTB4. Notwithstanding this paradox, Kaiso's role as a regulator of Wnt signalling in *Xenopus* was solidified by a 2009 study that investigated a molecular mechanism for the interaction of the *Xenopus* homolog of Kaiso (xKaiso) with xTcf3 [5]. Ruzov *et al.* demonstrated that xKaiso interacted directly with xTcf3, leading to xTcf3 dissociation from the promoters of Wnt target genes [5]. These seminal studies examining Kaiso's function in *Xenopus* development and Wnt signalling provided the rationale for further exploring Kaiso's role in Wnt signalling in the context of mammalian model systems.

Studies in mammalian cell culture models further supported a *bona fide* role for Kaiso as a modulator of canonical Wnt signalling. Our lab and others reported that Kaiso represses a subset of Wnt target genes (*cyclinD1, MMP7*) in mammalian cultured cells [15,36,37]. Cell culture studies also hinted at a negative feedback loop whereby activation of the Wnt signalling pathway would inhibit Kaiso's repressive effects on Wnt target genes [38]. Specifically, CK1ε phosphorylates p120^{ctn} in response to Wnt stimulation, which promotes p120^{ctn}'s association with Kaiso, and inhibits Kaiso's transcriptional functions (Fig. 3). In an extension of the studies performed by Ruzov *et al.*, Kaiso was also found to interact with TCF4 and β-catenin in a manner that was mutually exclusive from its association with p120^{ctn} [39]. Consequently, the Kaiso-p120^{ctn} interaction hinders Kaiso's ability to bind TCF4 and β-catenin to modulate Wnt signalling [38].

Given the evidence linking Kaiso to Wnt signalling and the role of Wnt signalling in intestinal homeostasis and disease, the mammalian intestine represented an obvious choice for further studies into the physiological relevance of the Kaiso-Wnt relationship. Most colorectal cancers (CRCs), ~70%, are sporadic and driven by point mutations that follow a specific succession, with the first mutation occurring in the *adenomatous polyposis coli (Apc)* tumor suppressor gene [40]. APC is a component of a large multi-protein complex that regulates the stability of β-catenin, the downstream effector of the canonical Wnt signalling cascade [41]. Most *Apc* mutations that have been characterized in CRC result in a truncated APC protein, such that even in the absence of ligand-induced activation of the Wnt pathway, β-catenin is stabilized and translocates to the nucleus where it activates a plethora of tumorigenesis-promoting Wnt target genes [41–43].

The first study to examine the role of Kaiso in intestinal tumorigenesis was conducted in 2005, when Prokhortchouk *et al.* examined the effect of Kaiso depletion on polyp formation in the *Apc$^{Min/+}$* mouse model of intestinal neoplasia [31]. Surprisingly, Kaiso-deficient *Apc$^{Min/+}$* mice exhibited longer lifespans, delayed tumor onset and smaller tumors compared to their *Apc$^{Min/+}$* counterparts, suggesting that Kaiso functions in a pro-tumorigenic capacity in the intestinal setting [31]. Notably, the phenotype of Kaiso-deficient mice mimicked that of mice depleted for Mbd2, a methyl−CpG binding repressor [44]. Hypermethylation of tumor suppressor genes in various cancers is well documented [20], and has been proposed as a mechanism by which Kaiso may exert its effects in CRC [17]. Indeed, Kaiso depletion in CRC cell lines resulted in increased expression of three hypermethylated tumor suppressor genes, *CDKN2A, HIC1*, and *MGMT* and sensitized cells to the cytotoxic drug etoposide [17]. The expression of Kaiso in human intestinal tissue and in the *Muc2$^{-/-}$* mouse model, which develops invasive colorectal tumors akin to those of patients with inflammatory bowel disease, was also investigated. Remarkably, Kaiso expression was found to be increased in the tumors of *Muc2$^{-/-}$* mice compared to controls, although no difference was observed in Kaiso expression between matched normal and tumor human intestinal tissues in this study by Prokhortchouk [31].

In a quest to further understand the role of Kaiso in intestinal tumorigenesis and canonical Wnt signalling, we generated an intestinal-specific Kaiso overexpressing mouse model (*Kaiso$^{Tg/+}$*) [45]. While Kaiso overexpression alone was not sufficient to drive the development of tumors in the murine intestine, *Kaiso$^{Tg/+}$* mice developed crypt hyperplasia and chronic intestinal inflammation [45], which is an established risk factor for CRC. In fact, in patients with inflammatory bowel disease, the cumulative risk of developing colitis-associated cancer (CAC) ranges between 18 and 20% [16]. The development of chronic intestinal inflammation in Kaiso overexpressing mice is intriguing on two fronts. First, in contrast to sporadic CRC, *Apc* mutation occurs only in advanced stages of IBD- associated carcinogenesis if at all [16], which suggests a Wnt-independent mechanism by which Kaiso may promote intestinal tumorigenesis. Secondly, it suggests opposing roles for Kaiso (pro-) and p120^{ctn} (anti-) in intestinal inflammation/tumorigenesis since limited ablation of p120^{ctn} in murine intestines induced an inflammatory response characterized by an epithelial barrier defect with infiltrating neutrophils [46,47], and promoted intestinal adenomas [47]. Most importantly, these findings hint that p120^{ctn} and Kaiso may play a role in IBD progression to CAC.

We expanded our studies into the effect of Kaiso overexpression on intestinal homeostasis by crossing our *Kaiso$^{Tg/+}$* mice with *Apc$^{Min/+}$* mice [16,48]. Consistent with earlier studies by Prokhortchouk *et al.*, *Kaiso$^{Tg/+}$:Apc$^{Min/+}$* mice exhibited significantly reduced lifespans and an approximately 3-fold increase in polyp number relative to control mice [48]. *Kaiso$^{Tg/+}$:Apc$^{Min/+}$* mice also developed chronic intestinal inflammation, albeit at a younger age than *Kaiso$^{Tg/+}$* mice [48]. We also investigated the expression of Wnt target genes in the intestines of *Kaiso$^{Tg/+}$* and *Kaiso$^{Tg/+}$:Apc$^{Min/+}$* mice [48]. Notably, Wnt target gene expression was increased rather than decreased in both *Kaiso$^{Tg/+}$* and *Kaiso$^{Tg/+}$:Apc$^{Min/+}$* mice; however the increases noted in *Kaiso$^{Tg/+}$* mice compared to control mice were not significant [16,48]. We also observed that increased Kaiso expression positively correlates with tumor stage in a human CRC tissue microarray (TMA) [48] and analysis of publicly available CRC microarray datasets revealed a statistically

Table 2
Summary of cancer-related processes and genes regulated by Kaiso.

Cancer site	Process affected as highlighted by research study	Gene(s)/Protein(s) Implicated	Findings	References
Colorectal	Survival, tumor burden	–	Delayed polyp onset, extended lifespan & smaller polyps upon Kaiso depletion in $Apc^{Min/+}$ mice	Prokhortchouk *et al.* [31]
	Chemotherapeutic response	CDKN2A, HIC1 and &MGMT	Increased expression of tumor suppressor genes CDKN2A, HIC-1 & MGMT in Kaiso depleted CRC cell lines; Kaiso-depleted cells were sensitized to the cytotoxic drug etoposide	Lopes *et al.* [17]
	Inflammation	–	Chronic intestinal inflammation & crypt hyperplasia on intestinal-specific overexpression of Kaiso in C57BL/6 J mice	Chaudhary *et al.* [45]
	Survival, tumor burden, inflammation	Wnt target genes, Pro-inflammatory markers	Shortened lifespan & increased polyp burden upon intestinal-specific overexpression of Kaiso in $Apc^{Min/+}$ mice; increased expression of Wnt target genes & pro-inflammatory proteins	Pierre *et al.* [48]
	–	JAG1	Positive regulation of JAG1 observed in both mice & CRC cell lines; JAG1 expression associated with Notch-mediated colon cancer	Robinson *et al.* [52]
Prostate	Migration, invasion	E-cadherin	Reduced migration & invasion of Kaiso-depleted prostate cancer (PCa) cells with increased E-cadherin expression	Jones *et al.* [14]
	Migration, invasion, proliferation	miR-31	Reduced migration & invasion of Kaiso-depleted PCa cells with increased miR-31 expression	Wang *et al.* [16]
	Metastasis, tumor burden	miR-200 family; EMT regulators	Increased expression of miR-200 family (associated with suppression of EMT) and decreased expression of EMT regulators in Kaiso-depleted PCa cells; decreased tumor growth & fewer metastases in Kaiso-depleted PC3 xenografts	Abisoye-Ogunniyan *et al.* [19]
Lung	Proliferation, invasion	–	Increased proliferation & invasion in Kaiso-depleted lung cancer cell lines	Dai *et al.* [28]
	–	Cyclin-D1	Kaiso overexpression decreased Cyclin-D1 & Cyclin-E1 expression	Jiang *et al.* [37]
	–	β-catenin	Kaiso overexpression reduced β-catenin expression	Liu *et al.* [69]
CML	Proliferation and differentiation	SCF, Wnt11, C/EBPα, c-Myb, GATA-2, PU.1, CD15, CD11b, CD33 and CD117	Kaiso depletion increased proliferation and decreased expression of differentiation markers	Cofre*et al.* [30]
Breast	EMT, metastasis	EMT regulators; Cell adhesion factors	Induction of MET phenotype with decreased expression of key EMT markers & enhanced expression of cell-adhesion proteins in Kaiso-depleted TNBC cell lines; reduced metastasis of xenografted Kaiso-depleted TNBC cell lines.	Jone *et al.* [13]; Bassey-Archibong et al. [81]
	EMT, metastasis	TGFβR1/2	Decreased expression of TGFβR1/2 with attenuation of TGFβ signalling in Kaiso-depleted TNBC cells; Constitutively active TGFβR1/2 rescued TGFβ signalling but not metastatic potential of xenografted Kaiso-depleted TNBC cells.	Bassey-Archibong *et al.* [81]
	Proliferation	CyclinD1, c-Myc. Ki67,PCNA	Reduced proliferation & restored anchorage dependence of Kaiso-depleted TNBC cell lines; reduced tumor volume of xenografted Kaiso-depleted TNBC cells; reduced expression of proliferation markers	Bassey-Archibong *et al.* [23]
	Apoptosis	C-Caspase3, Bax, PUMA	Induction of apoptosis with increased expression of pro-apoptotic proteins in mutant p53 Kaiso-depleted TNBC cells; enhanced sensitivity to cisplatain	
	DNA Repair	BRCA1	Increased expression of BRCA1 in Kaiso-depleted TNBC cell lines	
Glioma	Proliferation	PCNA, CDK2, CDK4, cyclin D1, cyclin E1, p21 and p27	Kaiso depletion inhibited proliferation of a glioma cell line accompanied by reduced expression of cell cycle regulators PCNA, CDK2, CDK4, Cyclin D1 and Cyclin E1 and increased expression of p21 and p27.	Wang *et al.* [90]
	Metastasis	E-cadherin, Vimentin, N-cadherin	Kaiso depletion resulted in reduced expression of mesenchymal markers (Vimentin, N-cadherin) and increased expression of epithelial markers (E-cadherin)	

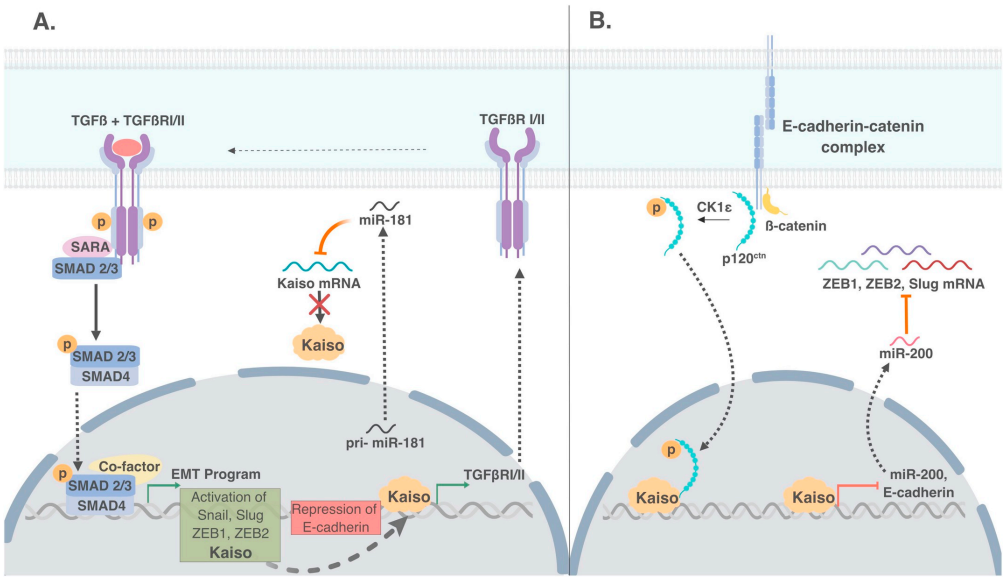


Fig. 3. Possible mechanisms of action of Kaiso in EMT. (A) Kaiso regulates TGF-β-mediated EMT and is itself regulated by the TGFβ signalling cascade which activates the expression of other pro-metastatic genes in triple negative breast tumors. The increase in Kaiso expression by TGFβ could be via the Smad transcription complex. Kaiso also appears to function in a positive feedback loop to increase TGFβ signalling (via increased TGFβRI & II expression) through as yet unidentified mechanism(s). (B) Kaiso represses E-cadherin expression through repression of miR-200 and possibly via sequestration of phosphorylated $p120^{ctn}$ in the nucleus. Phosphorylated $p120^{ctn}$ binds Kaiso, leading to its sequestration in the nucleus and E-cadherin degradation.

significant positive correlation between increased Kaiso expression and colon tumor tissues compared to normal tissues [16,48], in contrast to the Prokhortchouk study [31]. Collectively, these studies indicate that in mammalian models, Kaiso promotes rather than suppresses tumorigenesis as was initially anticipated based on our findings regarding Kaiso and Wnt signalling in *Xenopus*.

Another striking phenotype that we observed in $Kaiso^{Tg/+}$ mice was the expansion of the secretory cell population (i.e. goblet, Paneth and enteroendocrine cells) in the small and large intestines, which was accompanied by an overall decrease in cell proliferation [45]. This phenotype was reminiscent of that observed upon inhibition of the Notch signalling pathway in the intestine [49–51]. Indeed, expression of the Notch pathway target gene *Hes1* was decreased in $Kaiso^{Tg/+}$ mice [45], hinting at a novel role for Kaiso in regulating cell fate decisions through Notch signalling in the murine intestine. Notably, Kaiso represses the expression of both the Notch1 receptor and the Notch ligand Dll-1 in CRC cell lines [52] but promoted Jagged-1 expression, which is implicated in colon cancer progression. This finding provides an additional mechanism whereby Kaiso may promote CRC and adds another layer of complexity to Kaiso's role in intestinal homeostasis and colon cancer.

While no studies to date have pinpointed the precise molecular mechanisms by which Kaiso potentiates intestinal tumorigenesis, and a definitive link between Kaiso and Wnt signalling in intestines has not been established, it is clear that in the context of the intestine, Kaiso functions in a pro-tumorigenic manner. Thus, ongoing efforts in our lab and others seek to characterize Kaiso's role in intestinal inflammation and investigate its contribution to CAC and CRC using various mouse models.

2.2. Kaiso as a regulator of EMT and metastasis in prostate cancer

The first report examining Kaiso's role in prostate cancer (PCa) surveyed Kaiso expression and subcellular localization in various tumor types including a small cohort of PCa patients [53]. In a subsequent study with a larger patient cohort that included normal, adjacent normal, and benign prostatic hyperplasia (BPH) samples, Kaiso was found to be overexpressed in malignant prostate tumors and its high expression correlated significantly with tumor grade, Gleason score and race [14]. Specifically, low-Gleason grade tumors exhibited predominantly cytoplasmic Kaiso expression, while the high-Gleason grade and metastatic tumors displayed increased nuclear Kaiso expression. Two additional noteworthy observations from this study were that, i) African-American (AA) men expressed more nuclear Kaiso compared to Caucasian men, and ii) nuclear Kaiso was present in cells within the basal layer of adjacent normal tissue. Interestingly, multiple reports have indicated that adult prostate stem cells are present in the basal layer of both human and mouse prostate tissues and serve as developmental stimulants in PCa progression [39]. Indeed, a recent study by Yates *et al.* utilized Keratin-14-CreER/ROSA-LacZ and Keratin-5-tTA-TRE-H2BGFP lineage tracing mouse models and reported that Keratin-5 and Keratin-14 positive basal cells survive multiple rounds of castration and hormone manipulation, thus highlighting their stem cell characteristics [39]. It is unclear whether Kaiso expression in the basal layer is playing a role in castration resistance and/or PCa progression, but further investigation should be done to fully elucidate the role of Kaiso in the basal layer.

Previous studies suggest that DNA methylation is responsible for modulating gene expression that ultimately affects morphological changes in PCa cells and epithelial-to-mesenchymal transition (EMT) [54–56]. This is especially important in PCa, as PCa cells frequently

adopt an EMT phenotype [57–59]. EGFR has been well characterized as a promoter of EMT in multiple tumor types, including PCa [60]. To gain insight into Kaiso's mechanism of action in metastatic PCa, several well-characterized PCa cell culture models of EMT (LNCaP, DU-145, DU-145WT - which genetically overexpress EGFR, and PC-3, which exhibit a similar pattern of gene expression with human prostate tumor tissues), were examined and it was found that the more invasive and metastatic cell lines demonstrated increased nuclear Kaiso expression [14]. Interestingly, Kaiso expression was increased and Kaiso localized to the nucleus when EGFR was overexpressed or cells were treated with EGF. Furthermore, Kaiso bound directly to CpG islands in the E-cadherin promoter of PC-3 cells and Kaiso depletion resulted in robust re-expression of E-cadherin, similar to that observed on treatment with the demethylation agent 5-aza-CdR [14]. Intriguingly, a previous report by Yates *et al.* found that inhibition of the autocrine EGFR loop (and likely the EGFR-induced hepatocyte growth factor/c-met autocrine loop), either by direct disruption of the signalling loop or by secondary site signalling trans-attenuation, resulted in decreased cell motility and invasiveness, concomitant with E-cadherin re-expression in both DU-145 and PC-3 cells. Inhibition of the autocrine EGFR loop was observed in Kaiso-depleted cells hinting at a role for Kaiso in the EGFR-driven EMT phenotype of aggressive PCa.

To gain further insight into the mechanisms underlying Kaiso-mediated PCa development and progression, Wang *et al.* examined miRNA arrays and found that several miRNAs were upregulated in PCa [18]. Thirteen miRNAs, including miR-31, were significantly differentially expressed in Kaiso-depleted PC3 cells and a direct, methylation-dependent association between Kaiso and the endogenous miR-31 promoter was reported. PCa patients with Kaiso high and miR-31 low tumors had worse overall survival relative to patients with only Kaiso high tumors or miR-31 low tumors, suggesting that Kaiso promotes poor PCa survival via regulation of miR-31 expression. Most recently, Kaiso was also implicated in regulating the miR-200 family that suppresses EMT in PCa cells [19]. Upon Kaiso depletion in selected PCa cells, there was a significant increase of these miRNAs—similar to that observed upon demethylation—and a decrease in downstream EMT targets such as ZEB1/2, Twist and Snail. Furthermore, using an *in vivo* mouse xenograft model, Kaiso-depleted PC3 cells subcutaneously injected into mice exhibited decreased tumor growth and less metastases compared to mice injected with control PC3 cells [19]. Collectively, these data further highlight Kaiso's role in EMT and PCa tumor metastasis. Since acquisition of the mesenchymal phenotype by carcinoma cells is not permanent and tumor cells that have undergone EMT may later revert to an epithelial state at the secondary metastatic site through a mesenchymal-epithelial-transition (MET) [61–65], Kaiso could play a major role in mediating these transitional states of the cell throughout tumor progression and metastasis.

2.3. Kaiso and lung cancer

Two research groups recently investigated the role of Kaiso in lung cancer development and patient outcomes. Analysis of representative tissue cores from 294 non-small cell lung cancer (NSCLC) patients by Dai and colleagues, showed significantly increased cytoplasmic expression of Kaiso in NSCLC tissues compared to normal bronchial epithelium (NBE) tissues [28]. Dai *et al.* also found that high Kaiso expression correlated significantly with advanced stage and lymph node metastasis in a cohort of lung cancer patients [28]. Furthermore, patients with high levels of cytoplasmic Kaiso expression had significantly lower survival rates compared to those with no cytoplasmic Kaiso expression. Notably, nuclear Kaiso was only detected in a few NSCLC cases (~5%) but it was not associated with any clinicopathological features. In contrast to what was observed in NSCLC and NBE tissues, Dai *et al.* found that Kaiso was primarily localized in the nucleus of three lung cancer cell lines used (BE1, LETP-A-2 and SPC-A-1). Using proliferation and invasion assays, they found that down-regulation of nuclear Kaiso expression led to increased proliferation and invasion of the three lung cancer cell lines used [28]. In a subsequent study investigating Kaiso function in NSCLC, cytoplasmic Kaiso expression was positively correlated with high TNM stage, lymph node metastases and poor prognosis of NSCLC patients [66]. Notably, high cytoplasmic co-expression of Kaiso and δ-catenin was not detected in normal lung tissue [66].

The dynamic subcellular localization of Kaiso in NSCLC as described above is regulated by its interaction with $p120^{ctn}$. In a study investigating the role of the Kaiso-$p120^{ctn}$ interaction in various lung cancer cell lines, it was found that Kaiso bound primarily to $p120^{ctn}$ isoform 3 (p120-3A) and not isoform 1 (p120-1A) in these cell lines [67]. Transient transfection studies using A549 lung cancer cells revealed that both cytoplasmic and nuclear Kaiso expression was increased by p120-3A overexpression but was unchanged by p120-1A overexpression [67]. Moreover p120-3A-bound Kaiso was shuttled out of the nucleus through the chromosomal region maintenance export pathway [67]. Zhang *et al.* also found that Kaiso's binding to $p120^{ctn}$ was dependent on $p120^{ctn}$ serine-288 phosphorylation (Ps288) in lung cancer cells and tissues. They further showed that higher levels of Kaiso co-precipitated with phosphorylated $p120^{ctn}$ compared to unphosphorylated $p120^{ctn}$, suggesting that serine-288 phosphorylation may act as the molecular switch that mediates $p120^{ctn}$-Kaiso binding in lung cancer cells [67]. Most recently, Iderzorig *et al.* [68] found that $p120^{ctn}$ co-localizes with Kaiso in the nucleus in tyrosine kinase inhibitor (TKI)-resistant NSCLC cells, while in TKI-sensitive NSCLC cells reduced co-localization was observed. Interestingly, TKI-resistant NSCLC cells also exhibited low levels of E-cadherin expression and high expression of EMT markers such as Twist, Slug and Snail. This is in contrast to TKI-sensitive NSCLC cells, where E-cadherin was expressed and low levels of the aforementioned EMT markers were observed [68]. Given the dependence of the Kaiso-$p120^{ctn}$ interaction on the status of $p120^{ctn}$ phosphorylation [67], it is possible that $p120^{ctn}$ is hyper-phosphorylated in TKI-resistant NSCLC cells, resulting in its sequestration by Kaiso and consequently, a reduction in E-cadherin protein stability concomitant with increased cell motility or metastasis. This phenomenon, though not explored, is likely in TKI-resistant NSCLC tumors, since Dai *et al.* have reported overexpression of Kaiso in NSCLC [28].

Additional studies to elucidate Kaiso's role in lung cancer have proposed a role for Kaiso in regulating β-catenin expression in a methylation-dependent manner in lung cancer [69]. Notably, increased β-catenin expression in the nucleus and cytoplasm correlates with more aggressive clinical course and lower survival, and is extensively reviewed in [70]. Liu *et al.* found that treatment of lung cancer cell lines with the demethylation agent 5-Aza-CdR, resulted in increased β-catenin expression. However, high Kaiso expression significantly down-regulated β-catenin mRNA expression and suggested that Kaiso may be regulating β-catenin expression. Indeed, ChIP analysis revealed that Kaiso associated with the endogenous β-catenin promoter via-methylated CpG dinucleotides [69].

Recently, studies by Zhang *et al.* investigating Kaiso's role in lung cancer found that cigarette smoke induced nuclear-to-cytoplasmic shuttling of Kaiso, thus reducing Kaiso's ability to transcriptionally silence the expression of its pro-tumorigenic target genes (e.g. c-Fos) [71]. Cigarette smoke promoted complex formation between $p120^{ctn}$ and the cytoplasmic tail of Mucin-1 (MUC1-CT) and enhanced Kaiso binding to $p120^{ctn}$, an association that resulted in Kaiso shuttling to the cytoplasm. The nuclear-to-cytoplasmic shuttling of Kaiso was dependent on $p120^{ctn}$ as lung epithelial cells lacking $p120^{ctn}$ did not exhibit any shuttling of Kaiso from the nucleus [71].

Perhaps the most crucial findings that have been gleaned from the study of Kaiso in lung cancer cells is that Kaiso's subcellular localization and consequently, its function, can be altered through its interaction with phosphorylated $p120^{ctn}$. These findings lend insight into the upstream signalling events that can regulate Kaiso function, which have remained relatively understudied.

2.4. Kaiso as a regulator of cell differentiation and proliferation in chronic myeloid leukemia (CML)

The role of Kaiso's cytoplasmic localization was further explored in CML by Cofre *et al.* [30]. Using immunofluorescence and immunoblot analyses, Kaiso was observed to be highly expressed in the cytoplasm of these cells rather than the nucleus. Similar to what was found in lung cancer cells, knock down of Kaiso (and/or $p120^{ctn}$) led to increased proliferation in these cells, possibly through increased expression of the proliferation marker SCF, and hinted at a possible role for Kaiso by itself or in association with $p120^{ctn}$ for CML proliferation [30]. Examination of the expression of hematopoietic differentiation genes Wnt11, C/EBPα, c-Myb, GATA-2 and PU.1 in Kaiso and/or $p120^{ctn}$ knock-down CML cells revealed that c-MyB expression was increased while Wnt 11, PU-1, C/EBPα and Gata-2 expression were decreased when compared to scrambled knock-down cells [30]. Knock down of Kaiso and/or $p120^{ctn}$ also led to decreased expression of the global cell differentiators CD15, CD11b, CD33 and CD117. Notably in the clinical context, high expression of cytoplasmic Kaiso was observed in the more aggressive form of CML [30]. These findings were the first to implicate Kaiso in CML and further investigation is warranted to fully assess Kaiso's role and mechanism of action in CML.

2.5. Kaiso and breast cancer

Mounting evidence suggests that Kaiso functions as an important driver of aggressive breast cancers like the triple negative breast cancer (TNBC) subtype. Interestingly, TNBCs are characterized by high proliferation indices, high rates of recurrence and a propensity to metastasize [72–74]. Currently, TNBCs lack targeted therapies due to their lack of estrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor-2 (HER2) expression. [75]. Although some therapeutic response is achieved through chemotherapy, patients with TNBC often relapse, thus associating this BC subtype with poor outcomes and shortened overall survival [76–79]. Notably, TNBC, is more frequently diagnosed in women of African ancestry, and Kaiso's role in TNBC racial disparities in incidence and outcomes will be discussed in a later section.

Vermeulen *et al.* examined Kaiso expression and localization in a large cohort of normal and tumor breast tissues and found that high-grade breast tumors and tumors with a high mitotic activity index exhibited more nuclear Kaiso expression compared to low-grade tumors [29]. They further found that Kaiso expression negatively correlated with ERα positivity, and that nuclear Kaiso was significantly higher in aggressive BC, particularly the basal/triple negative and HER2-amplified breast cancers compared to luminal-type BC [29]. Notably, ~70% of hereditary (*BRCA1*-associated) breast cancers exhibited nuclear Kaiso expression compared to ~30% of sporadic carcinomas [29].

In a subsequent study by Jones *et al.*, it was found that high nuclear Kaiso expression positively correlated with local invasion, lymph node metastases and poorer overall survival in a large cohort of invasive ductal carcinoma (IDC) tumor tissue samples [13]. Jones *et al.* also found TNBC cell lines (MDA-MB-231 & MDA-MB-468, hereafter MDA-231 & MDA-468) exhibited higher nuclear expression of Kaiso compared to the non-metastatic MCF-7 breast cancer cell line where Kaiso mainly localized to the cytoplasm [13]. Kaiso depletion in the highly metastatic MDA-231 and MDA-468 cell lines resulted in reduced cell migration in both wound healing and Boyden chamber assays, and reduced invasion through matrigel assays, while Kaiso overexpression in MCF-7 cells resulted in their increased migration and invasion [13]. These findings hinted that Kaiso may be involved in the regulation of the EMT process that is characterized by loss of E-cadherin expression and is a key contributor to metastases and poor outcomes of cancer patients [13]. Indeed, Kaiso depletion in MDA-231 and MDA-468 cells that typically lack E-cadherin expression, resulted in re-expression of E-cadherin with an associated reduction in expression of mesenchymal-related factors including N-cadherin, Cadherin 11 and Vimentin [13]. Chromatin immunoprecipitation assays also revealed that Kaiso binds the E-cadherin promoter, but treatment of cells with the demethylating agent 5-Aza-CdR abolished Kaiso binding, suggesting that this interaction is methylation-dependent [13].

Multiple signal transduction pathways regulate EMT, but one of the most studied is the TGFβ signalling pathway, which directly activates the expression of key transcriptional mediators of EMT such as Snail, Slug, ZEB1, ZEB2, and Twist [80]. Notably, we found that Kaiso-depleted MDA-231 breast tumor cells exhibited decreased TGFβ signalling, proliferated slowly in vivo and unlike parental MDA-231 cells, did not metastasize to lungs or liver in mouse xenograft models [81]. Kaiso depletion also resulted in increased expression of the TGFβ receptors, TGFβR1 and TGFβR2, and Kaiso was found to bind to KBS sites in the promoter regions of both genes [81]. Remarkably, TGFβ signalling induced Kaiso expression in TNBC cells at both the transcript and protein levels [81], which hinted at a positive feedback loop between Kaiso and TGFβ signalling in TNBC cells. Notably, TGFβ signalling plays a paradoxical role in breast cancer; in early stage breast cancer, TGFβ acts as a tumor suppressor by inhibiting cell proliferation and inducing apoptosis, but in advanced and aggressive breast cancers, TGFβ promotes progression and metastasis partly through induction of EMT [82]. Thus, our findings suggested that Kaiso may be a key driver of metastasis in TNBC (Fig. 3).

Beyond its role in TNBC metastasis, we more recently demonstrated that Kaiso also plays a critical role in TNBC cell proliferation and survival [23]. Kaiso-depleted MDA-231 cells exhibited delayed tumor onset and reduced expression of the cell proliferation markers Ki67 and CyclinD1 in a murine xenograft model [23]. Kaiso depletion also resulted in increased apoptosis of TNBC cells expressing mutant p53 and up-regulation of the pro-apoptotic genes PUMA and Bax [23]. Intriguingly, BRCA1 expression was reduced in Kaiso-depleted Hs587T and MDA-231 cells, and high co-expression of Kaiso and BRCA1 correlated with poor overall survival of TNBC patients. Notably, some TNBC subtypes demonstrate decreased sensitivity to chemotherapeutic agents [83]. Thus our finding that Kaiso-depleted Hs578T and MDA-231 cells exhibit increased Cisplatin sensitivity suggest that Kaiso may promote BRCA-1-mediated chemotherapy resistance in TNBC [23].

It is noteworthy that in breast tumor cells expressing wild-type p53, Kaiso functions in a pro-apoptotic capacity [23]. This supports independent studies showing that Kaiso activates the pro-apoptotic gene *APAF1* by forming a protein complex with wild-type p53, which in turn binds and activates the *APAF1* promoter [24]. Another mechanism by which Kaiso may promote apoptosis of breast tumor cells is through methylation-dependent repression of the glucocorticoid receptor (GR) [84]. Indeed, Kaiso overexpression reduced the anti-apoptotic effect of GR activation, thus highlighting a possible adjuvant role for Kaiso in breast cancer patients treated with glucocorticoid combined chemotherapy [84]. Collectively, these studies highlight paradoxical roles for Kaiso in apoptosis and indicates that even within the same tissue type, Kaiso's role is highly complex and context-dependent.

2.6. Kaiso, miR-181 and EMT in glioma

Further support of Kaiso's role as a regulator of EMT and proliferation in cancer cells was recently demonstrated in glioma cell lines. The miR-181 family inhibits proliferation, invasion, migration and EMT of gliomas by targeting several genes [85–89]. High Kaiso expression was found in glioma tissue compared to adjacent normal tissues and importantly, a significant negative correlation was found between Kaiso and miR-181 expression in high grade gliomas [90]. miR-181a bound to Kaiso's 3′ UTR and ectopic expression of a miR-181a mimic resulted in reduced expression of both Kaiso mRNA and protein suggesting that Kaiso is a miR-181a target. Furthermore, Kaiso overexpression rescued the anti-proliferative and anti-invasive effects of a miR-181a mimic suggesting that miR-181 elicits its inhibitory effects on proliferation,

invasion, migration and EMT through Kaiso suppression [90]. Similar to findings in breast tumor cell lines, Kaiso depletion inhibited the proliferation of a glioma cell line accompanied by reduced expression of cell cycle regulators PCNA, CDK2, CDK4, Cyclin D1 and Cyclin E1 and increased expression of p21 and p27. Reduced expression of mesenchymal markers (Vimentin, N-cadherin) and increased expression of epithelial markers (E-cadherin) was also observed upon depletion of Kaiso in glioma cells [90], similar to what has been reported in both breast and prostate cancer cells [19,81]. In addition to reinforcing Kaiso's role as a regulator of EMT, these studies have also identified the miR-181 family as novel regulators of Kaiso expression.

3. Kaiso and racial disparities in cancer

In the past decade, increasing evidence indicates that genetically based racial disparities exist in various human cancers including breast and prostate. The high BC mortality rates in women of African ancestry (WAA) despite having lower incidence rates and lifetime risks of BC than Caucasian women, initially led many to think that this disparity was linked to socio-economic status [91–95]. However, several studies now suggest that genetic risk factors/predisposition contribute to the racial disparity in the prevalence and mortality of patients diagnosed with the TNBC subtype [93,97,96]. Hence, the recent discovery that high Kaiso expression correlates with aggressive BC subtypes like TNBC and shorter metastasis-free survival in WAA [13,29], as well as with aggressive PCa in African American (AA) men [14] suggests that high Kaiso expression may be linked to the racial disparity in prevalence and/or outcomes of aggressive cancers in people of African ancestry.

As a first step in testing this hypothesis, the Daniel and Yates labs independently examined different cohorts of breast cancer patients and observed that nuclear Kaiso is significantly overexpressed in breast tumor tissues of patients with an African ancestry [13,97]. Jones *et al.* further observed that nuclear Kaiso was overexpressed in primary tumor and paired lymph node metastases of AA women compared to Caucasian women (CaW) [13], and AA women with nuclear Kaiso had a decreased overall survival compared to CaW with nuclear Kaiso expression. We extended these findings in a pilot tissue microarray (TMA) study where we examined Kaiso expression and subcellular localization in a TMA comprised of TNBC tissues from West African (Nigeria) and Caribbean (Barbadian) women. These patient cohorts were selected based on the premise that they represent more homogeneous populations of African ancestry than African Americans. Bassey-Archibong *et al.* found that Nigerian and Barbadian women were diagnosed with TNBC at much younger ages than AA and CaW [97], and that nuclear Kaiso levels were significantly higher in Nigerian, Barbadian and AA women compared with CaW. The high levels of nuclear Kaiso expression in women of African Ancestry compared to their Caucasian counterparts, suggests a role for Kaiso in TNBC racial disparity, and forms the premise for ongoing studies in the Daniel lab to further decipher this phenomenon. Collectively, these studies raise the exciting possibility that Kaiso may be a useful prognostic marker for TNBC, especially in WAA.

Notably, high Kaiso expression also correlates significantly with aggressive prostate tumors in AA men who have more aggressive tumors compared to Caucasian men [14]. Kaiso expression, particularly nuclear Kaiso, is increased in AA PCa patients, and is associated with a worse overall survival in AA men compared to Caucasian men [14]. Finally, high Kaiso expression has also been linked to racial disparities in pancreatic ductal adenocarcinoma (PDAC) [98]. Jones *et al.* observed higher Kaiso expression in PDAC tissues of AA patients when compared with normal/adjacent tissues [98]. They also found a significant relationship between high Kaiso expression, higher grade and tumor size in AA PDAC patients. Collectively these studies highlight an unexpected and unique role for Kaiso in racial disparities in cancer.

4. Concluding remarks

In the past decade, significant strides have been made in understanding how Kaiso contributes to various human cancers. Generally, high Kaiso expression correlates with poor prognosis and worse clinical outcomes in colorectal, prostate, lung and breast cancers. We also have more insight into the molecular mechanisms via which Kaiso may drive these outcomes, since it is now appreciated that Kaiso regulates several tumor-associated processes including EMT, proliferation, apoptosis and inflammation.

Contrary to early reports that characterized Kaiso as a transcriptional repressor, it is evident that the manner in which Kaiso functions to modulate transcription may depend on several different factors including its interaction partners (e.g. wild-type vs. mutant p53) as well as the site at which it binds DNA. Most molecular studies to date have investigated Kaiso binding to the consensus KBS or in the context of methylation; however, few studies have examined the Kaiso binding motif identified by Raghav *et al.* [8]. Novel regulatory pathways may be identified using this motif and further enhance our understanding of Kaiso's roles in tumorigenic processes. Other factors that regulate the activity of Kaiso include (i) changes in subcellular localization, which can occur upon binding to phosphorylated $p120^{ctn}$, (ii) phosphorylation, which may alter Kaiso's protein interactions, subcellular localization and transcriptional activity and (iii) SUMOylation, which alters Kaiso's transcriptional activity. However, it is still unclear the precise cascade of events that lead to $p120^{ctn}$-, phosphorylation- or SUMO-mediated regulation of Kaiso and hence these factors represent exciting avenues for further investigation. Furthermore, Kaiso's regulation by miR181a and other as yet unknown miRNAs demand further investigation in various cell lines and contexts.

Another exciting avenue of study revolves around the findings that Kaiso can promote inflammation in the intestine. The mechanisms and molecular pathways via which Kaiso promotes inflammation, Kaiso's interplay with $p120^{ctn}$ in the inflammatory process and whether Kaiso elicits its pro-inflammatory effects in other tissues are all questions that warrant further investigation.

With increasing data demonstrating that high Kaiso expression correlates with aggressive tumor characteristics, most likely through Kaiso's regulation of EMT and cell proliferation in multiple cancers, Kaiso represents a candidate for further investigation as a diagnostic or prognostic marker and potentially a therapeutic target. One important question to be addressed is what causes high Kaiso expression in aggressive tumors. Finally, the link between high Kaiso expression and the racial disparities in cancer outcomes was an unexpected finding and ongoing studies seek to determine the molecular mechanism, if any, underlying this phenomenon. Two decades worth of progress into understanding Kaiso's relevance in various cancers, has led to many intriguing findings and has indeed primed us for an exciting new era of scientific discovery around Kaiso function.

Acknowledgements

All images created with BioRender.

Funding

JMD was funded by Canadian Institutes of Health Research (CIHR) MOP-84320, Natural Sciences and Engineering Research Council (NSERC) RGPIN6442-15, Juravinski Hospital and Cancer Center Foundation (JHCCF) R-153, Canadian Breast Cancer Foundation/Canadian Cancer Society (CBCF/CCS) 316252; CCP was supported by The Richard Fuller Memorial Scholarship for Excellence in Science; CY was funded by the National Institutes of Health/Research Centers at Minority Institutions (NIH/RCMI) U54-MD007585-26, National Institutes of Health/National Cancer Institute (NIH/NCI)U54 CA118623 , (NIH/NCI) 1 R21 CA188799-01; and the US

Army Department of Defense, PC120913, W81XWH-10-1-0543.

References

[1] J.M. Daniel, A.B. Reynolds, The catenin p120(ctn) interacts with Kaiso, a novel BTB/POZ domain zinc finger transcription factor, Mol. Cell. Biol. 19 (5) (1999) 3614–3623.
[2] K.F. Kelly, J.M. Daniel, POZ for effect - POZ-ZF transcription factors in cancer and development, Trends Cell Biol. 16 (11) (2006) 578–587.
[3] S.W. Kim, et al., Isolation and characterization of XKaiso, a transcriptional repressor that associates with the catenin Xp120(ctn) in Xenopus laevis, J. Biol. Chem. 277 (10) (2002) 8202–8208.
[4] S.W. Kim, et al., Non-canonical Wnt signals are modulated by the Kaiso transcriptional repressor and p120-catenin, Nat. Cell Biol. 6 (12) (2004) 1212–1220.
[5] A. Ruzov, et al., The interaction of xKaiso with xTcf3: a revised model for integration of epigenetic and Wnt signalling pathways, Development 136 (5) (2009) 723–727.
[6] J.M. Daniel, et al., The p120(ctn)-binding partner Kaiso is a bi-modal DNA-binding protein that recognizes both a sequence-specific consensus and methylated CpG dinucleotides, Nucleic Acids Res. 30 (13) (2002) 2911–2919.
[7] A. Prokhortchouk, et al., The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor, Genes Dev. 15 (13) (2001) 1613–1618.
[8] S.K. Raghav, et al., Integrative genomics identifies the corepressor SMRT as a gatekeeper of adipogenesis through the transcription factors C/EBPbeta and KAISO, Mol. Cell 46 (3) (2012) 335–350.
[9] A. Blattler, et al., ZBTB33 binds unmethylated regions of the genome associated with actively expressed genes, Epigenetics Chromatin 6 (1) (2013) 13.
[10] K.F. Kelly, et al., NLS-dependent nuclear localization of $p120^{ctn}$ is necessary to relieve Kaiso-mediated transcriptional repression, J. Cell Sci. 117 (2004) 2675–2686.
[11] G.J. Filion, et al., A family of human zinc finger proteins that bind methylated DNA and repress transcription, Mol. Cell. Biol. 26 (1) (2006) 169–181.
[12] N. Sasai, M. Nakao, P.A. Defossez, Sequence-specific recognition of methylated DNA by human zinc-finger proteins, Nucleic Acids Res. 38 (15) (2010) 5015–5022.
[13] J. Jones, et al., Nuclear localization of Kaiso promotes the poorly differentiated phenotype and EMT in infiltrating ductal carcinomas, Clin. Exp. Metastasis 31 (5) (2014) 497–510.
[14] J. Jones, et al., Nuclear Kaiso indicates aggressive prostate cancers and promotes migration and invasiveness of prostate cancer cells, Am. J. Pathol. 181 (5) (2012) 1836–1846.
[15] C.M. Spring, et al., The catenin p120ctn inhibits Kaiso-mediated transcriptional repression of the beta-catenin/TCF target gene matrilysin, Exp. Cell Res. 305 (2) (2005) 253–265.
[16] D. Saraggi, et al., The molecular landscape of colitis-associated carcinogenesis, Dig. Liver Dis. 49 (4) (2017) 326–330.
[17] E.C. Lopes, et al., Kaiso contributes to DNA methylation-dependent silencing of tumor suppressor genes in colon cancer cell lines, Cancer Res. 68 (18) (2008) 7258–7263.
[18] H. Wang, et al., Kaiso, a transcriptional repressor, promotes cell migration and invasion of prostate cancer cells through regulation of miR-31 expression, Oncotarget 7 (5) (2016) 5677–5689.
[19] A. Abisoye-Ogunniyan, et al., Transcriptional repressor Kaiso promotes epithelial to mesenchymal transition and metastasis in prostate cancer through direct regulation of miR-200c, Cancer Lett. 431 (2018) 1–10.
[20] B.P. Madakashira, K.C. Sadler, DNA methylation, nuclear organization, and cancer, Front. Genet. 8 (2017) 76.
[21] M. Rodova, et al., Regulation of the rapsyn promoter by kaiso and delta-catenin, Mol. Cell. Biol. 24 (16) (2004) 7188–7196.
[22] K.E. Wiese, et al., The role of MIZ-1 in MYC-dependent tumorigenesis, Cold Spring Harb. Perspect. Med. 3 (12) (2013) a014290.
[23] B.I. Bassey-Archibong, et al., Kaiso depletion attenuates the growth and survival of triple negative breast cancer cells, Cell Death Dis. 8 (3) (2017) e2689.
[24] D.I. Koh, et al., Transcriptional activation of APAF1 by KAISO (ZBTB33) and p53 is attenuated by RelA/p65, Biochim. Biophys. Acta 1849 (9) (2015) 1170–1178.
[25] A. Pozner, T.W. Terooatea, B.A. Buck-Koehntop, Cell-specific Kaiso (ZBTB33) regulation of cell cycle through Cyclin D1 and Cyclin E1, J. Biol. Chem. 291 (47) (2016) 24538–24550.
[26] S. Zhenilo, et al., DeSUMOylation switches Kaiso from activator to repressor upon hyperosmotic stress, Cell Death Differ. 25 (2018) 1938–1951.
[27] D. Hanahan, R.A. Weinberg, Hallmarks of cancer: the next generation, Cell 144 (5) (2011) 646–674.
[28] S.D. Dai, et al., Cytoplasmic Kaiso is associated with poor prognosis in non-small cell lung cancer, BMC Cancer 9 (2009) 178.
[29] J.F. Vermeulen, et al., Nuclear kaiso expression is associated with high grade and triple-negative invasive breast cancer, PLoS One 7 (5) (2012) e37864.
[30] J. Cofre, et al., Knock-down of Kaiso induces proliferation and blocks granulocytic differentiation in blast crisis of chronic myeloid leukemia, Cancer Cell Int. 12 (1) (2012) 28.
[31] A. Prokhortchouk, et al., Kaiso-deficient mice show resistance to intestinal cancer, Mol. Cell. Biol. 26 (1) (2006) 199–208.
[32] J.I. Park, et al., Kaiso/p120-catenin and TCF/beta-catenin complexes coordinately regulate canonical Wnt gene targets, Dev. Cell 8 (6) (2005) 843–854.
[33] E.M. Schatoff, B.I. Leach, L.E. Dow, Wnt signaling and colorectal cancer, Curr. Colorectal Cancer Rep. 13 (2) (2017) 101–110.
[34] J.M. Daniel, Dancing in and out of the nucleus: p120(ctn) and the transcription factor Kaiso, Biochim. Biophys. Acta 1773 (1) (2007) 59–68.
[35] H. Iioka, S.K. Doerner, K. Tamai, Kaiso is a bimodal modulator for Wnt/beta-catenin signaling, FEBS Lett. 583 (4) (2009) 627–632.
[36] N.S. Donaldson, et al., Kaiso represses the cell cycle gene cyclin D1 via sequence-specific and methyl-CpG-dependent mechanisms, PLoS One 7 (11) (2012) e50398.
[37] G. Jiang, et al., P120-catenin isoforms 1 and 3 regulate proliferation and cell cycle of lung cancer cells via beta-catenin and Kaiso respectively, PLoS One 7 (1) (2012) e30303.
[38] B. Del Valle-Perez, et al., Wnt controls the transcriptional activity of Kaiso through CK1epsilon-dependent phosphorylation of p120-catenin, J. Cell Sci. 124 (Pt 13) (2011) 2298–2309.
[39] C. Zhao, et al., Dual regulatory switch through interactions of Tcf7l2/Tcf4 with stage-specific partners propels oligodendroglial maturation, Nat. Commun. 7 (2016) 10883.
[40] I. Marmol, et al., Colorectal carcinoma: a general overview and future perspectives in colorectal cancer, Int. J. Mol. Sci. 18 (1) (2017) .
[41] E.R. Fearon, Molecular genetics of colorectal cancer, Annu. Rev. Pathol. 6 (2011) 479–507.
[42] R. Gryfe, Inherited colorectal cancer syndromes, Clin. Colon Rectal. Surg. 22 (4) (2009) 198–208.
[43] J. Behrens, B. Lustig, The Wnt connection to tumorigenesis, Int. J. Dev. Biol. 48 (5–6) (2004) 477–487.
[44] O.J. Sansom, et al., Deficiency of Mbd2 suppresses intestinal tumorigenesis, Nat. Genet. 34 (2) (2003) 145–147.
[45] R. Chaudhary, et al., The POZ-ZF transcription factor Kaiso (ZBTB33) induces inflammation and progenitor cell differentiation in the murine intestine, PLoS One 8 (9) (2013) e74160.
[46] W.G. Smalley-Freed, et al., p120-catenin is essential for maintenance of barrier function and intestinal homeostasis in mice, J. Clin. Invest. 120 (6) (2010) 1824–1835.
[47] W.G. Smalley-Freed, et al., Adenoma formation following limited ablation of p120-catenin in the mouse intestine, PLoS One 6 (5) (2011) e19880.
[48] C.C. Pierre, et al., Kaiso overexpression promotes intestinal inflammation and potentiates intestinal tumorigenesis in Apc(Min/+) mice, Biochim. Biophys. Acta 1852 (9) (2015) 1846–1855.
[49] S. Fre, et al., Notch signals control the fate of immature progenitor cells in the intestine, Nature 435 (7044) (2005) 964–968.
[50] J. Jensen, et al., Control of endodermal endocrine development by Hes-1, Nat. Genet. 24 (1) (2000) 36–44.
[51] T. Ueo, et al., The role of Hes genes in intestinal development, homeostasis and tumor formation, Development 139 (6) (2012) 1071–1082.
[52] S.C. Robinson, et al., Kaiso differentially regulates components of the Notch signaling pathway in intestinal cells, Cell Commun. Signal. 15 (1) (2017) 24.
[53] A. Soubry, et al., Expression and nuclear location of the transcriptional repressor Kaiso is regulated by the tumor microenvironment, Cancer Res. 65 (6) (2005) 2224–2233.
[54] H. Ross-Adams, et al., HNF1B variants associate with promoter methylation and regulate gene networks activated in prostate and ovarian cancer, Oncotarget 7 (2016) 74734–74746.
[55] G. Palma, et al., Triple negative breast cancer: looking for the missing link between biology and treatments, Oncotarget 6 (29) (2015) 26560–26574.
[56] S. Majid, et al., miR-23b represses proto-oncogene Src kinase and functions as methylation-silenced tumor suppressor with diagnostic and prognostic significance in prostate cancer, Cancer Res. 72 (24) (2012) 6435–6446.
[57] S. Josson, et al., Tumor-stromal interactions influence radiation sensitivity in epithelial - versus mesenchymal-like prostate cancer cells, J. Oncol. 2010 (2010).
[58] C. Yates, et al., Novel three-dimensional organotypic liver bioreactor to directly visualize early events in metastatic progression, Adv. Cancer Res. 97 (2007) 225–246.
[59] C.C. Yates, et al., Co-culturing human prostate carcinoma cells with hepatocytes leads to increased expression of E-cadherin, Br. J. Cancer 96 (8) (2007) 1246–1252.
[60] M.I. Khan, et al., Role of epithelial mesenchymal transition in prostate tumorigenesis, Curr. Pharm. Des. 21 (10) (2015) 1240–1248.
[61] E.A. Ruiz-Narvaez, et al., Genetic variation in the insulin, insulin-like growth factor, growth hormone, and leptin pathways in relation to breast cancer in African-American women: the AMBER consortium, NPJ Breast Cancer 2 (2016).
[62] B. Senguven, et al., Comparison of methods for the extraction of DNA from formalin-fixed, paraffin-embedded archival tissues, Int. J. Med. Sci. 11 (5) (2014) 494–499.
[63] R.L. Siegel, K.D. Miller, A. Jemal, Cancer Statistics, 2017, CA Cancer J. Clin. 67 (1) (2017) 7–30.
[64] S.G. Ahn, et al., Molecular classification of triple-negative breast cancer, J. Breast Cancer 19 (3) (2016) 223–230.
[65] H.M. Jeong, et al., Targeted exome sequencing of Korean triple-negative breast cancer reveals homozygous deletions associated with poor prognosis of adjuvant chemotherapy-treated patients, Oncotarget 8 (37) (2017) 61538–61550.
[66] S.D. Dai, et al., Upregulation of δ-catenin is associated with poor prognosis and enhances transcriptional activity through Kaiso in non-small-cell lung cancer, Cancer Sci. 102 (1) (2011) 95–103.
[67] P.X. Zhang, et al., p120-catenin isoform 3 regulates subcellular localization of Kaiso and promotes invasion in lung cancer cells via a phosphorylation-dependent mechanism, Int. J. Oncol. 38 (6) (2011) 1625–1635.
[68] T. Iderzorig, et al., Comparison of EMT mediated tyrosine kinase inhibitor resistance in NSCLC, Biochem. Biophys. Res. Commun. 496 (2) (2018) 770–777.
[69] Y. Liu, et al., Kaiso interacts with p120-catenin to regulate beta-catenin expression at the transcriptional level, PLoS One 9 (2) (2014) e87537.

[70] J. Jin, et al., Prognostic significance of β-catenin expression in patients with non-small cell lung cancer: a meta-analysis, Transl. Lung Cancer Res. 6 (1) (2017) 97–108.
[71] L. Zhang, et al., Cigarette smoke mediates nuclear to cytoplasmic trafficking of transcriptional inhibitor kaiso through MUC1 and P120-catenin, Am. J. Pathol. 186 (12) (2016) 3146–3159.
[72] R. Lindner, et al., Molecular phenotypes in triple negative breast cancer from African American patients suggest targets for therapy, PLoS One 8 (11) (2013) e71915.
[73] J. Iqbal, et al., Differences in breast cancer stage at diagnosis and cancer-specific survival by race and ethnicity in the United States, JAMA 313 (2) (2015) 165–173.
[74] T. Keenan, et al., Comparison of the genomic landscape between primary breast cancer in African American versus white women and the association of racial differences with tumor recurrence, J. Clin. Oncol. 33 (31) (2015) 3621–3627.
[75] O. Clark, et al., Targeted therapy in triple-negative metastatic breast cancer: a systematic review and meta-analysis, Core Evid. 9 (2014) 1–11.
[76] F. Kassam, et al., Survival outcomes for patients with metastatic triple-negative breast cancer: implications for clinical practice and trial design, Clin. Breast Cancer 9 (1) (2009) 29–33.
[77] G. Houvenaeghel, et al., Axillary lymph node micrometastases decrease triple-negative early breast cancer survival, Br. J. Cancer 115 (9) (2016) 1024–1031.
[78] A. Wushou, et al., Development of triple-negative breast cancer radiosensitive gene signature and validation based on transcriptome analysis, Breast Cancer Res. Treat. 154 (1) (2015) 57–62.
[79] W. Sun, et al., Prognostic analysis of triple-negative breast cancer patients treated with adjuvant chemotherapy of fluorouracil, epirubicin and cyclophosphamide, Oncol. Lett. 11 (3) (2016) 2320–2326.
[80] Y. Katsuno, S. Lamouille, R. Derynck, TGF-beta signaling and epithelial-mesenchymal transition in cancer progression, Curr. Opin. Oncol. 25 (1) (2013) 76–84.
[81] B.I. Bassey-Archibong, et al., Kaiso depletion attenuates transforming growth factor-beta signaling and metastatic activity of triple-negative breast cancer cells, Oncogene 5 (2016) e208.
[82] J.M. Zarzynska, Two faces of TGF-Beta1 in breast cancer %, J. Mediat. Inflammation 2014 (2014) 16.
[83] E.A. O'Reilly, et al., The fate of chemoresistance in triple negative breast cancer (TNBC), BBA Clin. 3 (2015) 257–275.
[84] L. Zhou, et al., Kaiso represses the expression of glucocorticoid receptor via a methylation-dependent mechanism and attenuates the anti-apoptotic activity of glucocorticoids in breast cancer cells, BMB Rep. 49 (3) (2016) 167–172.
[85] S.X. Huang, et al., The correlation of microRNA-181a and target genes with poor prognosis of glioblastoma patients, Int. J. Oncol. 49 (1) (2016) 217–224.
[86] Z.M. Shi, et al., MiRNA-181b suppresses IGF-1R and functions as a tumor suppressor gene in gliomas, RNA 19 (4) (2013) 552–560.
[87] E. Ayala-Ortega, et al., Epigenetic silencing of miR-181c by DNA methylation in glioblastoma cell lines, BMC Cancer 16 (2016) 226.
[88] X.F. Wang, et al., MiR-181d acts as a tumor suppressor in glioma by targeting K-ras and Bcl-2, J. Cancer Res. Clin. Oncol. 138 (4) (2012) 573–584.
[89] L. Shi, et al., hsa-mir-181a and hsa-mir-181b function as tumor suppressors in human glioma cells, Brain Res. 1236 (2008) 185–193.
[90] L. Wang, et al., Kaiso (ZBTB33) downregulation by mirna-181a inhibits cell proliferation, invasion, and the epithelial-mesenchymal transition in glioma cells, Cell. Physiol. Biochem. 48 (3) (2018) 947–958.
[91] C. Desantis, et al., Breast cancer statistics, 2011, CA Cancer J. Clin. 61 (6) (2011) 409–418.
[92] Society, A.C, http://www.cancer.org/Cancer/BreastCancer/, (2011).
[93] K.C. Amirikia, 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 (12) (2011) 2747–2753.
[94] D. Huo, et al., Population differences in breast cancer: survey in indigenous African women reveals over-representation of triple-negative breast cancer, J. Clin. Oncol. 27 (27) (2009) 4515–4521.
[95] A. Stark, et al., African ancestry and higher prevalence of triple-negative breast cancer: findings from an international study, Cancer 116 (21) (2010) 4926–4932.
[96] P. Khosravi-Shahi, L. Cabezon-Gutierrez, S. Custodio-Cabello, Metastatic triple negative breast cancer: Optimizing treatment options, new and emerging targeted therapies, Asia Pac. J. Clin. Oncol. 14 (2017) 32–39.
[97] B.I. Bassey-Archibong, et al., Kaiso is highly expressed in TNBC tissues of women of African ancestry compared to Caucasian women, Cancer Causes Control 28 (11) (2017) 1295–1304.
[98] J. Jones, et al., African Americans with pancreatic ductal adenocarcinoma exhibit gender differences in Kaiso expression, Cancer Lett. 380 (2) (2016) 513–522.

CHAPTER 7: DISCUSSION

The inherent difficulty to appropriately assess and treat TNBC has resulted in a growing interest by many researchers. This BCa subtype disproportionately affects WAA who exhibit the highest TNBC prevalence across West Africa and the West African diaspora. Furthermore, WAA diagnosed with TNBC have higher associated mortality rates compared to women of other ancestries (Alluri & Newman, 2014; DeSantis *et al.*, 2019; DeSantis *et al.*, 2017; Martini *et al.*, 2021; Newman *et al.*, 2006; Stark *et al.*, 2010) with increasing studies indicating an intricate interplay of environmental and biological/genomic factors as the causal factors behind the racial disparity in TNBC incidence and outcomes. Recent studies have focused on investigating the exome/genome of WAA with TNBC in order to unveil pathogenic and novel mutations in genes associated with carcinogenesis and which may have therapeutic and clinical relevance (Adedokun *et al.*, 2021; Churpek *et al.*, 2015; Melissa Davis *et al.*, 2020; Martini *et al.*, 2021). Of note, novel, and pathogenic germline *BRCA1*, *BRCA2* and *PALB2* mutations as well as high somatic mutation frequency of *TP53* have been reported in various cohorts of WAA with BCa (Adedokun *et al.*, 2020; Donenberg *et al.*, 2011; Friebel *et al.*, 2019; George *et al.*, 2021; Huo *et al.*, 2017). Additionally, identifying TNBC biomarkers in WAA has seen an expansion in research as high throughput sequencing and molecular techniques have become more financially feasible to achieve. One such potential biomarker is Kaiso which has been implicated in metastasis, apoptosis and proliferation in TNBC through *in vivo* and *in vitro* experiments (Bassey-Archibong *et al.*, 2016; Bassey-Archibong, Rayner, *et al.*, 2017; Kwiecien *et al.*, 2017). However, despite the increasing investigation on TNBC epidemiology and genetics

in WAA, little is known about the (i) BCa epidemiological profile of women with a high African ancestry in the Caribbean compared to non-Hispanic white and non-Hispanic Black women in the USA, (ii) TNBC prevalence across African countries, (iii) TNBC-associated mutations and alterations in WAA, or (iv) the relative expression of Kaiso and other BCa biomarkers in TNBC tissues from West African (e.g. Nigerian) and Caribbean (e.g. Barbadian) women. This thesis presents findings on these research gaps and adds to our understanding about TNBC epidemiological trends, the tumour biology of WAA, and potential TNBC therapeutic targets.

7.1 BCA EPIDEMIOLOGY IN BARBADOS

Barbados is a small developing island state situated as the most eastern island in the island chain in the Caribbean Sea. Biomedical research—specifically BCa research—has not been extensively pursued in this Caribbean island due to under-resourced healthcare systems. The most recent analyses from GLOBOCAN estimated that Barbados has the highest BCa mortality globally (Sung *et al.*, 2021) which is concerning for this limited resource small island. Ragin and colleagues reviewed BCa estimates from studies done across the Caribbean and found that Barbados had the second highest age standardized BCa incidence (94.7 per 100,000), second to the Bahamas (98.9 per 100,000) (Ragin *et al.*, 2018). However, the limited number of studies investigating BCa incidence and/or mortality in Barbados have not investigated clinical-pathological variables and BCa subtype data. For our study, we conducted a comprehensive pathological investigation of all BCa cases recorded in Barbados from the sole public hospital, QEH for the 2007-2016 ten-year period. With a population of approximately 277,821 of which 144,803 are females according to

Barbados 2010 census data (Service, 2013), we estimated a crude BCa incidence of 135.1 per 100,000 females in Barbados for the 2007-2016 period (**Chapter 3**). Our study also showed a disproportionately higher age-specific incidence of BCa in young (15-59) Barbadian females, when compared with age-specific incidence of NHB and NHW in the USA (**Chapter 3, Figure 3**). Young age at diagnosis has previously been associated with aggressive clinical course and genetic factors influencing the BCa diagnosis (Anders *et al.*, 2008; Gómez-Flores-Ramos *et al.*, 2017). Greater consideration across the island should therefore be placed on genomic testing as well as on adjusting the BCa screening program to diagnose BCa at earlier ages and an earlier stage instead of at a more advanced stage. Indeed, Ragin and colleagues additionally highlighted that emphasis should be placed across the Caribbean region on hereditary vs. sporadic BCa as there were multiple reports of young age at diagnosis (Ragin *et al.*, 2018).

Barbadian females had higher frequency of lymph node (LN) positivity and Grade 3 carcinoma diagnoses when compared with NHB and NHW females with BCa in the USA (**Chapter 3, Figure 2**). As LN positivity and grade 3 tumours indicate a high propensity to metastasize to other organs and high proliferative ability of tumour cells respectively, this clinical profile could be indicative of poor prognosis and thus may be contributing factors to the high BCa mortality in Barbados as estimated by GLOBOCAN. Early diagnosis through consistent screening programs should be a priority for Barbados as perhaps, some of these more advanced tumours might have been caught in earlier stages. However, a recent study investigating approaches to cancer control across the Caribbean region highlighted that there was no organised BCa national screening programme present in

Barbados which might be an indicator of lack of resources and that mammography services may be primarily used for diagnostic measures rather than screening purposes (Spence *et al.*, 2019). However, establishing a national BCa screening programme in Barbados as a priority would be beneficial to screening and diagnosing BCa at earlier stages to decrease preventable BCa related morbidity and mortality.

We also investigated BCa subtype prevalence over the 2010-2016 period for comparison with the USA as receptor status data were fully collected from 2010 in the SEER database. We estimated a 25% TNBC prevalence over this period in comparison with 21% and 10% in NHB and NHW females in the USA respectively (**Chapter 3, Figure 2**). This reported prevalence is higher than what was previously reported for TNBC across the Caribbean where TNBC frequency among of all BCa cases ranged from approximately 14%-17.3% (Ragin *et al.*, 2018) in Caribbean-based BCa studies. High TNBC prevalence in Barbados further indicates a poor BCa prognosis among the Barbadian population which is of concern. A previous review investigating the epidemiology of TNBC in WAA, highlighted body weight and distribution (measured either by overweight/obesity status, or waist to hip ratio), parity, short duration of breastfeeding and age younger than 40 years as factors that positively associated with TNBC in WAA populations in the USA (Brewster *et al.*, 2014). Population-based data in Barbados have estimated that the prevalence of overweight and obese adults among Barbadian females is ~60% (Spence *et al.*, 2019) which may account for or contribute to the high TNBC prevalence observed among this population. For our study, anthropometric measures such as body weight were not collected; however, further investigation into these risk factors and their associations with

BCa subtypes in Barbados would shed significant light on such risk factors and further public health measures to adequately address them. Indeed, a previous Barbadian based study found an association between BMI and BCa risk that indicates a stronger relationship in older females for this association (Nemesure, Wu, Hennis, *et al.*, 2009). Furthermore, another population-based Barbadian study also showed associations between age at first full-term pregnancy, parity (as a protective factor), family history of BCa and history of benign breast disease with BCa among African Barbadian women (Nemesure, Wu, Hambleton, *et al.*, 2009). Interrogation into the BCa epidemiology in a small island such as Barbados would greatly assist policymakers with ensuring appropriate treatment options are available for the unique distribution of patients diagnosed with the various BCa subtypes, i.e., since there is a relatively high prevalence of TNBC among this population, more emphasis could be placed on utilizing other chemotherapeutic agents such as PARP inhibitors, which have proven to be beneficial for TNBC patients. Additionally, IHC could be routinely done to assess PD-L1 expression as targeted therapies such as atezolizumab in combination with nab-paclitaxel have been beneficial for metastatic TNBC patients that are positive for PD-L1 (Kagihara *et al.*, 2020; Kang & Syed, 2020; Narayan *et al.*, 2020). Thus, comprehensive knowledge about the BCa epidemiology of any population, especially frequency of the BCa subtypes, is highly beneficial for mitigating poor BCa outcomes.

7.2 EPIDEMIOLOGY OF TNBC ACROSS AFRICA

Some countries across Africa have already started to undergo an epidemiological transition and have started to be affected by chronic non communicable diseases (NCDs) such as hypertension, cancer and diabetes with a higher burden when compared with

communicable diseases such as tuberculosis, malaria and acquired immunodeficiency syndrome (Sung *et al.*, 2021). To understand the burden of BCa subtypes, namely TNBC, across Africa, we conducted a systematic review and meta-analysis of all studies reporting data on BCa subtypes with African women. Previous reports investigating BCa subtypes across Africa only reported ER negativity across the continent, BCa survival rates or BCa subtypes among Eastern Africa, but there were no comprehensive reviews or meta-analyses done to show the differences in TNBC frequency across the continent (Newman & Kaljee, 2017; Popli *et al.*, 2021; Ssentongo *et al.*, 2019). Elucidating the distribution of TNBC frequency across the continent is necessary to better understand the high TNBC prevalence reported in the African diaspora worldwide.

In our study, we found that TNBC frequency was highest in studies investigating BCa in West African women compared with studies investigating BCa in women from other regions across Africa (**Chapter 3, Table 1, Figure 3, Figure 4**). Specifically, studies conducted on Ghanaian, Nigerian and Senegalese women with BCa reported the highest TNBC prevalence across Africa (**Chapter 3, Figure 3, Table S2, Table S5**), which was consistent with our hypothesis that TNBC frequency would be highest in West African countries. This hypothesis was derived from studies showing higher TNBC prevalence in WAA in North America, the UK, and Caribbean islands (where women have high African ancestry and are ancestrally related to West Africa). Due to the transatlantic slave trade, it is postulated that the African diaspora across the Caribbean and North America were largely from West African countries such as Nigeria and Ghana. This has been confirmed by multiple genotyping efforts and investigating ancestry informative markers within these

populations (Benn-Torres *et al.*, 2008; Geronimus *et al.*, 2006; Moreno-Estrada *et al.*, 2013; Murray *et al.*, 2010; Zakharia *et al.*, 2009). Notably, pooled TNBC frequency across African countries with higher admixture of Arabian ancestry (North and East Africa) was lower (**Chapter 3, Figure 3, Figure 4**) indicating that West African ancestry may be associated with TNBC prevalence. A recent population-based study conducted in the USA found that TNBC prevalence was highest in West-African born women, followed by Caribbean born women then East-African born women, who had the lowest TNBC prevalence of the three groups (Sung *et al.*, 2019). This indicates that Black women across the USA, and arguably the rest of the diaspora, are not a monolith and ancestry should be strongly considered when investigating "racial disparities", not only in BCa but also in other diseases. The search strategy for our systematic review however only identified studies from 20 out of the 54 countries across Africa which indicates that there is a significant knowledge gap about BCa and BCa subtypes across the continent. However, we are confident that our study provides the best current estimate of TNBC frequency across the continent given the available studies.

Our research showed that receptor status assessment tools and techniques varied considerably across studies which also led to variable estimates of TNBC frequency. We found that TNBC estimates were higher in studies that used the Allred algorithm instead of ASCO/CAP guidelines for ER/PR status, while studies that assessed HER2 using older ASCO/CAP guidelines reported higher TNBC (**Chapter 3, Figure 5, Figure 6**). We also found an association between year of publication and TNBC estimates (**Chapter 3, Figure S4B**). This association indicates that either (i) the tools used to assess receptor status are

more stringent, and that tissue handling, processing and analysis has improved over time allowing for decreased false positives, or (ii) with increasing BCa awareness programs, more women are being screened and there is less inflation of TNBC prevalence. Indeed, a recent systematic review and meta-analysis on BCa subtypes across East Africa, found a similar association with year of publication where ER-positivity was noted to be higher in studies after 2013 compared with studies done before 2013 (Popli *et al.*, 2021). One key difference between the Allred algorithm and ASCO/CAP guidelines for assessing ER/PR status is that for the Allred algorithm, the cut-off for receptor status positivity could be as high as 10% (Allred *et al.*, 1998), whereas for ASCO/CAP guidelines, 1% is the more stringent cut-off (Allison *et al.*, 2020); therefore, it would be expected to observe higher receptor positivity for studies that used the Allred algorithm. Additionally, more recent versions of ASCO/CAP guidelines for HER2 testing include the interrogation of HER2 via IHC as well as through fluorescent *in situ* hybridization (FISH) where cut-off for positivity moved from 30% to 10% positive tumour cells via IHC between the 2007 and current guidelines (Ahn *et al.*, 2020). Indeed, our results reflect differences in TNBC estimates when taking the tool used for receptor status into account.

Additionally, across African nations there is a clear disparity in the laboratory techniques used for tumor assessment and the associated financial costs to the patient. One Nigerian study indicated that cost was a barrier for IHC tests and that FISH was not routinely done for HER2 equivocal scores as outlined in ASCO/CAP guidelines because of costs, whereas, in South Africa, the comprehensive public health system eliminates financial barriers to accessing care and having breast tumours appropriately subtyped (CO.

et al., 2017; Cubasch *et al.*, 2018). This inconsistency in access to healthcare and facilities for adequate tumour assessment could further explain the high variability in pooled TNBC estimates across the continent and emphasizes the need for international collaborations to assist countries across Africa (and other LMICs) with improved access to these tools to properly diagnose and treat BCa patients. A recent small study with physicians and nurses across Africa, noted that medical/hospital supplies and direct monetary support were their highest priority to improve efforts at decreasing BCa burden at their respective institutions (Gyan *et al.*, 2019). Another study interrogating inequities in BCa treatment across sub-Saharan Africa highlighted a vast disparity of treatment rates in Nigeria compared to Namibia: in Nigeria ~ 40% of women were not treated within 1-year of diagnosis compared to the Namibian population where almost all patients (~99%) received treatment within 1-year of diagnosis (Foerster *et al.*, 2019). Foerster and colleagues report SES and belief in traditional medicine to be barriers to treatment in the untreated populations. Increasing cancer screening and access to care in these nations should thus be a priority for various funding entities that have the resources to do so.

When assessing knowledge about BCa, two separate studies from Nigeria and Ghana both found that most of the information obtained about BCa was from mass media and there was a general poor knowledge of BCa-associated risk factors (Azubuike, 2017; Opoku *et al.*, 2012). The Ghanaian-based study also found that the rate of breast self-examination (BSE), and clinical breast examination (CBE) were higher than that of obtaining mammograms (Opoku *et al.*, 2012) which emphasizes the need to promote screening programs in a culturally-relevant setting. It should be noted however that

mammography has been associated with a two times higher chance of detecting ER-positive BCa compared to ER-negative BCa (Howlader *et al.*, 2014) which might be contributing to the relatively higher TNBC frequency observed across West African countries when compared to Southern African countries where mammography is more accessible. Though somewhat controversial, mammography has been an effective screening tool in high-resource settings; however, it has been postulated that this may not be an effective screening tool for resource-limited settings or for detecting advanced staged tumors (Black & Richmond, 2019). In comparison, as reviewed by Corbex *et al.*, multiple studies across resource-limited settings have recommended CBE which requires less resources and has been demonstrated to decrease mortality (Corbex *et al.*, 2012). This further emphasizes the need for culturally relevant approaches to reduce BCa burden within resource-limited populations, such as many African countries. A recent study by Jemal *et al.* (Jemal & Brawley, 2019) outlined the effectiveness of cultural relevance to increase cancer awareness; cancer awareness programs where politicians are actively involved in the campaign were found to be effective in increasing knowledge about cancer across various African countries. This type of approach could be vital in decreasing cancer burden and mortality across Africa with the epidemiological transition that will lead to increasing cancer incidence and mortality, specifically across sub-Saharan Africa.

7.3 SOMATIC MUTATION ANALYSIS REVEALS PATHOGENIC AND NOVEL VARIANTS IN WAA

Due to the high prevalence of TNBC in WAA, multiple cancer biologists and clinician-scientists have launched research projects to investigate the mutational landscape of WAA

with TNBC to identify "driver genes" that may serve as drug targets and/or as diagnostic tools and lead to better treatment options. A recent review of BCa genetics in populations across Africa however found that the majority of studies focused on *BRCA1* and *BRCA2* mutations and not wide interrogation of the genome (Abbad *et al.*, 2018). Indeed, this was also true for the Caribbean where studies have focused on *BRCA1*, *BRCA2* and *PALB2* across the region (Donenberg *et al.*, 2016; Donenberg *et al.*, 2011; George *et al.*, 2021; Lerner-Ellis *et al.*, 2017). However, in the USA where resources are not limited, more unbiased genomics studies have been done using large genotyping datasets and whole exome/genome sequencing efforts. Previous studies have reported high *TP53* mutation rates in African American women with TNBC as well as novel mutations such as rs10069690 for *TERT* associated with ER/PR- and TNBC risk specifically in African American women (Brewster *et al.*, 2014; Huo *et al.*, 2017). To date, there have been no comprehensive studies investigating the mutational profile of ancestrally-related West African (Nigeria) and Caribbean (Barbados) women with TNBC.

We conducted WES on DNA collected from 31 Barbadian and Nigerian FFPE TNBC samples to investigate shared and unique somatic mutation signatures between these groups and subsequently compared this mutational profile with that of African American and European American TNBC samples from the TCGA database. After investigating the overlap of genes within these groups, we found 78 commonly mutated genes across all four cohorts (Barbadian, Nigerian, TCGA-AA and TCGA-EA) that might have potential as actionable targets (**Chapter 5, Figure 2A**). Intriguingly, we found that there were 2,401 unique variants in the TCGA- EA cohort that are potentially associated with the

overwhelming large sample size of that group within the TCGA database. To increase genomics representation within these databases to reflect our global population, more collaborative studies need to be done to further our understanding of genetic risks and therapeutic opportunities for diverse populations. The low inclusion (~9%) of African ancestry data within the TCGA database (Yuan *et al.*, 2018) is problematic and more studies need to be done to include diverse populations within these large genomics databases. It must be noted however, that there is a high degree of hesitancy to participate in health-based data collection studies among **A**frican **A**mericans (AAs) and arguably, the African diaspora worldwide. In the USA, there is a centuries long history of segregation, racism and unfair treatment of AA. In current times, these historical and ongoing atrocities towards AA are postulated to fuel hesitancy for COVID-19 vaccine uptake in AAs (Bajaj & Stanford, 2021).

After global analysis of the mutation frequency of all 36,409 genes in each of the cohorts (Barbadian, Nigerian, TCGA-AA and TCGA-EA) no significantly mutated genes were identified apart from two pseudogenes (*TNRC18P2, DDX12P*) with no currently known biological or clinical relevance (**Chapter 5, Figure 2B**). Although pseudogene function in general is not widely understood, it is postulated that since pseudogenes share high sequence similarity with their associated parental genes, they may interact with long-non-coding RNAs (lncRNAs) for regulatory purposes (Milligan & Lipovich, 2014). While there is currently no known function for these two pseudogenes, future investigation of their function and roles in signaling pathways and regulatory networks will shed light on their relevance in our WAA cohorts and TCGA-EA with TNBC. In addition to this

discovery, we also found novel, pathogenic and likely pathogenic variants in well-known cancer-associated genes (e.g. *BRCA1, MDC1, TP53*) which may be potential driver mutations within these cohorts (**Chapter 5, Figure 2C**). As previously stated, these therapies might be useful for WAA cohorts with high *BRCA1* mutations (George *et al.*, 2021; Papadimitriou *et al.*, 2018; Turner & Tutt, 2012; Tutt *et al.*, 2018). Notably, mutations in *MDC1* are also implicated in chemosensitivity to platinum-based therapies (Ruff *et al.*, 2020), suggesting that platinum-based therapies should be considered as a possible treatment of TNBC in WAA.

In addition to non-silent mutations in *BRCA1*, somatic signature analysis showed strong correlations for age, *BRCA1/BRCA2* and defective DNA mismatch repair/microsatellite instability (small INDELs) respectively, where there was a weak correlation for Signature 10, POLE (ultra-hypermutation) in Barbadian samples and a correlation to Signature 24 (Aflatoxin) in Nigerian samples (**Chapter 5, Figure 3**). Somatic mutation signatures for this study were derived from the **C**atalogue **o**f **S**omatic **M**utations **i**n **C**ancer (COSMIC) database which includes somatic mutation data from over 1 million tumour samples and more than 25,000 publications (Tate *et al.*, 2019). These signatures are defined by six substitutions (C>A, C>G, C>T, T>A, T>C, and T>G) and the corresponding bases immediately 5' and 3' to each mutated base. For example, signature 10 that is highly correlated to POLE hypermutation is defined by C>A and T>G mutations at TCT and TTT sites respectively where the underlined nucleotide is mutated (Forbes *et al.*, 2016). Interestingly, when investigating this signature with the TCGA database, this is not a signature that was observed in AA or EA TNBC samples but was observed in BCa

samples including all BCa subtypes and self-reported race (**Chapter 5, Figure 3**). This signature has been mostly associated with colorectal and uterine cancers, and altered activity of DNA polymerase POLE is proposed to be the underlying mutational process driving this signature (Haradhvala *et al.*, 2018). Additionally, this signature is associated with hypermutated samples which was seen in our data where the highest mutated samples were among Barbadian samples in comparison with Nigerian samples (**Chapter 5, Figure 2C**). These cohort specific correlations to somatic signatures 10 and 24 in Barbadian and Nigerian women respectively were interesting, especially signature 24 in Nigerian women.

It is well documented that there is a large consumption of unprocessed maize and groundnut across Nigeria which contributes to high levels of aflatoxin being consumed in these populations (Bandyopadhyay *et al.*, 2019; Ladep *et al.*, 2014). This signature is associated with **h**epato**c**ellular **c**arcinoma (HCC) liver cancer, specifically in cancer samples with known aflatoxin exposure (likely from **af**latoxin **B1** (AFB1)) and C>A mutations at GCC are common for this signature (Forbes *et al.*, 2016). Interestingly, this signature was discovered through WES of 11 HCC samples with known aflatoxin exposure (likely from African or Asian origins) with presence of *TP53* p.R249S mutations in some of the samples (Huang *et al.*, 2017). Though we did not find this mutation in any of our samples, we found an ~50% *TP53* mutation rate in our Nigerian samples (**Chapter 5, Figure 2C**). A recent report also found other *TP53* mutations in tumour samples that had aflatoxin exposure and somatic mutational signature (Chawanthayatham *et al.*, 2017). This report also noted that after aflatoxin exposure in mice, the somatic mutational signature was seen as early as 10 weeks post-aflatoxin exposure, even though tumours were seen at

72 weeks post-aflatoxin exposure. Though liver tumours were not assessed in our Nigerian participants, this *in vivo* mouse study shone light on additional mutational events as a result of inflammation and oxidative stress that are acquired over time before liver tumours were established in their mouse models (Chawanthayatham *et al.*, 2017). Hence, screening efforts for liver cancer may be of importance for this population with high aflatoxin exposure.

Indeed, apart from HCC samples, previous evidence has found AFB1-DNA adducts present in tumour and normal breast, colorectal and cervical tissues where higher concentration of AFB1-DNA adducts were observed in tumour samples (Harrison *et al.*, 1993). This study was crucial in early studies about aflatoxins and their relevance as carcinogens. More recent studies have shown AFB1 to be involved in regulating micro-RNA (miR) 25-5p through *in silico* analyses (Marchese *et al.*, 2018). This miR intriguingly has been implicated in multiple cancers but specifically was shown to be involved in promoting proliferation in TNBC cells potentially through AKT and ERK-MAPK signalling pathways (Chen, Pan, *et al.*, 2018). The finding of aflatoxin signature in our Nigerian TNBC tissues adds context to the potential of environmental factors (*e.g.,* groundnut consumption) contributing to somatic mutations. Though this finding is based on association and not necessarily causation, it is interesting nevertheless and should be followed up to better understand how environmental factors such as diet can have an impact on BCa tumorigenesis and outcomes.

We also identified bi-allelic loss of *TP53* which is consistent with other reports in WAA with TNBC (Pitt *et al.*, 2018). Additionally, copy number analysis revealed high

amplification for *PIK3CA*, and high copy loss for *FGFR2* and *HIF1AN* in our Barbadian and Nigerian cohorts (**Chapter 5, Figure 4**). Surprisingly we did not find amplification of *ZBTB33* (the gene encoding Kaiso) in our tissues as this was expected from previous reports indicating high Kaiso expression in WAA TNBC tissues (Bassey-Archibong *et al.*, 2016; Singhal *et al.*, 2021). However, the X chromosome, where *ZBTB33* is located, did not have high sequencing coverage due to short, fragmented reads from our FFPE samples. We aim to conduct targeted sequencing of this region to investigate this further.

Of note, clinical trials are currently underway (NCT04216472, NCT04251533 and NCT03207529) with *PIK3CA* inhibitors which could be therapeutically beneficial to TNBC patients with *PIK3CA* amplifications (Chang *et al.*, 2021). Additionally, there are current clinical trials for FGFR2 inhibitors for TNBC patients as *FGFR2* is likely to be amplified and/or increased FGFR signalling has been previously reported (Cunningham *et al.*, 2020; Jafarian *et al.*, 2019; Kim *et al.*, 2013). Interestingly, our Barbadian and Nigerian samples show a stark copy number loss at 10q26.12 - q26.13 which includes the *FGFR2* gene which was an unexpected finding. Further studies should investigate this phenomenon to understand if (i) this is indeed a *bona fide* mutation signature and (ii) *FGFR2* is lost in other TNBC WAA cohorts. Overall, this novel genomics profile could provide benefits for women with TNBC regardless of ancestry.

7.4 GENE EXPRESSION PROFILING IN WAA TNBC TISSUES

Studies have largely focused on *in silico* analyses using repositories such as TCGA and GEO for interrogation of gene mutations, RNA expression, predicted protein changes and protein expression in various subsets of patients and tumour types. As an alternate

approach, we constructed a WAA-TNBC TMA with 46 Barbadian and 28 Nigerian TNBC tissues. This TMA is a unique and useful resource as it was generated directly from patient samples consented into the study from these under-studied and resource-constrained locations with high TNBC prevalence (**Chapters 3 & 4**). The TMA allows us to interrogate protein subcellular localisation and semi-quantitative expression of TNBC biomarkers on multiple samples (n=74) during one IHC experiment. Using this TMA, we investigated sub-cellular localisation and semi-quantitative expression of Kaiso protein. Kaiso has been previously implicated in TNBC aggressiveness, metastasis, proliferation and chemoresistance (Bassey-Archibong *et al.*, 2016; Bassey-Archibong, Rayner, *et al.*, 2017; Singhal *et al.*, 2021; Vermeulen *et al.*, 2012). We found higher relative nuclear Kaiso H-scores in Nigerian and Barbadian TNBC tissues and lower H-scores in African American and Caucasian American TNBC tissues in a separate commercially available TMA from Yale University (Bassey-Archibong, Hercules, *et al.*, 2017a; **Appendix Chapter 3, Figure 3, Figure 4, Table 1**). This is consistent with previous findings from our lab using *in silico* analyses where we reported that Kaiso expression was highest in African American TNBC tissues compared to Caucasian American TNBC tissues (Bassey-Archibong *et al.*, 2016). The cause for this high Kaiso expression is unknown since the regulatory pathways that regulate Kaiso expression and function are also unknown. Nonetheless, the correlation of high Kaiso expression with aggressive cancers (TNBC, prostate, pancreatic) and disparities in outcomes in TNBC and prostate cancer (Pierre *et al.*, 2019, **Chapter 6)** strongly suggests an important role for Kaiso in the racial disparities observed in TNBC and possibly other cancers.

One signaling pathway implicated in regulating Kaiso expression is the TGFβ signaling pathway. Findings from our lab elucidated a positive feedback loop in TNBC cells between Kaiso and TGFβ signalling (Bassey-Archibong *et al.*, 2016) which suggests that TGFβ signalling may be regulating Kaiso expression in TNBC. To assess this, investigation of *ZBTB33* promoter region through ChIP-PCR for TGFβ components such as SMAD and ZEB1 transcription factors (TFs) would shed light on their association at the promoter region, implicating regulatory effects. Furthermore, luciferase promoter-reporter assays investigating these TFs would demonstrate whether their association with the promoter would activate or repress Kaiso expression. Additionally, it was previously demonstrated that Kaiso expression increased incrementally over a 24-hour period upon treatment in PCa cells (Jones *et al.*, 2012). This suggests that EGF and associated signalling pathways may also regulated Kaiso overexpression. Both TGFβ and EGF signalling pathways have been previously associated with TNBC in various studies (Changavi *et al.*, 2015; Costa *et al.*, 2017; Hashmi *et al.*, 2019; Katayama *et al.*, 2019; Xu *et al.*, 2018; Zhang *et al.*, 2017) and the complex relationships between these pathways and Kaiso expression should be examined more to fully understand their respective roles in Kaiso regulation, expression as well as TNBC aggressiveness in WAA.

Kaiso has multiple roles in cancer, as discussed in our recent review (Pierre *et al.*, 2019; **Chapter 6**) and its high expression has been associated with higher tumour grade, larger tumour size, apoptosis and metastasis in TNBC (Bassey-Archibong *et al.*, 2016; Bassey-Archibong, Rayner, *et al.*, 2017; Jones *et al.*, 2014; Kwiecien *et al.*, 2017). Herein, we also found that Nigerian and Barbadian women were diagnosed at significantly younger

ages, at higher grade and had larger tumour size at diagnosis (Bassey-Archibong, Hercules, *et al.*, 2017a; **Appendix Chapter 3, Figure 1, Table 1**). To replicate these findings, a larger TMA was constructed with 103 Barbadian and 123 Nigerian BCa tissues. Intriguingly, we found that relative nuclear Kaiso expression was higher in Barbadian tissues compared with Nigerian tissues (**Appendix Chapter 2, Figure A2.1**). Due to this contradictory finding between the pilot TMA and the large TMA, ongoing experiments and analyses seek to understand and determine the cause of this discrepancy/difference. When clinical variables were investigated among patients in our larger WAA-TMA, a similar clinical profile as outlined in our cohorts for the pilot WAA-TMA were replicated in the larger WAA-TMA (**Appendix Chapter 2, Table A2.1**). This is consistent with other studies that have shown that regardless of BCa staging, NHB women still had a more aggressive clinical profile (high grade, lymph node involvement *etc.*) compared with NHW (Yedjou *et al.*, 2017). Recent characterisation of Kaiso found that both high nuclear and cytoplasmic expression of Kaiso was associated with poor survival, regardless of lymph node status with the strongest predictor being cytoplasmic Kaiso expression (Singhal *et al.*, 2021). Singhal and colleagues however did not find evidence of a racial disparity in their cohort when nuclear and cytoplasmic Kaiso H-scores were compared. The finding of stronger association with cytoplasmic Kaiso was certainly intriguing as most studies to date investigating Kaiso's role in cancer have focused on nuclear Kaiso expression. Evidence suggests that Kaiso shuttles in and out of the nucleus with its binding partner $p120^{ctn}$ (Pierre *et al.*, 2019; **Chapter 6**) and that Kaiso's transcriptional activity is regulated when bound to $p120^{ctn}$. We investigated $p120^{ctn}$ subcellular localisation and semi-quantitative expression in our

pilot WAA-TMA and found lower relative cytoplasmic $p120^{ctn}$ expression in Nigerian and Barbadian TNBC tissues compared to African American and Caucasian American TNBC tissues (**Appendix Chapter 1, Figure A1.1**). Since the decreased expression $p120^{ctn}$ is associated with cell motility, invasion and metastasis, our findings suggest that their its expression within our WAA-TMA cohorts compared to the AA and CA TNBC samples imply a higher metastatic potential in our cohorts. Indeed, epidemiological analysis of the Barbadian cohort did reveal a higher proportion of samples with positive LN status compared to NHB and NHW participants in the USA (**Chapter 3**).

Further to this finding, we investigated expression of actin remodeling protein, WAVE-3 in our pilot WAA-TMA. High expression of WAVE-3 has been previously associated with increased metastasis and low chemosensitivity (Davuluri *et al.*, 2014; Kulkarni *et al.*, 2012). Additionally, WAVE-3 and TGFβ interact via a positive feedback mechanism to promote EMT in TNBC (Taylor *et al.*, 2013). In our pilot WAA-TNBC-TMA, we found higher relative WAVE-3 expression in Nigerian and Barbadian TNBC tissues compared with African American and Caucasian American tissues. This is the first study to note a differential expression of WAVE-3 between different population groups. It was recently highlighted that not only does WAVE-3 play important roles with metastasis and cell motility, (i) it functions within these roles in its phosphorylated form and (ii) WAVE-3 phosphorylation is necessary for PI3K, TGFβ and EGF signalling and downstream effectors both *in vitro* and *in vivo* (Wang *et al.*, 2020). Further investigation into the phosphorylation status of WAVE3 within our WAA-TMA pilot should be pursued which might shed more light on WAVE3 activity as well as the interplay of TGFβ and EGF

signalling on Kaiso expression. Indeed, high Kaiso and WAVE3 expression within our WAA-TMA cohorts imply a correlational relationship between them. High relative nuclear and cytoplasmic WAVE-3 expression in addition to high relative nuclear Kaiso expression in WAA-TNBC tissues warrants further investigation into this relationship. Indeed, further studies need to be done to investigate Kaiso's various roles and subcellular localisation in TNBC carcinogenesis and within the racial disparity observed in BCa.

7.5 OUTSTANDING QUESTIONS AND FUTURE STUDIES

This thesis adds to current knowledge about the epidemiological and molecular profiles of WAA with TNBC, specifically Barbadian and Nigerian women. We were the first to report TNBC frequency across the African content, epidemiology of BCa in Barbados, expression profiling of Kaiso in WAA with TNBC as well as identifying a novel genomics profile in Barbadian and Nigerian women. However, the following unanswered questions remain which were not investigated within this thesis but were highlighted as potential future avenues for research: (i) what roles do environmental and genetic factors play in tumorigenesis in WAA with TNBC; (ii) how is Kaiso regulated and why have multiple studies shown high Kaiso expression in WAA; and (iii) what are the downstream effects, if any, of newly identified mutational profile in WAA with TNBC?

7.5.1 Interplay of environmental and genetic factors contributing to BCa tumorigenesis in WAA

While we found potential evidence of environmental factors (aflatoxin exposure) correlating with genomic signature in six Nigerian TNBC samples, there is no causative

association for unprocessed maize and groundnut consumption with this signature. Aflatoxin exposure has been shown to lead to liver cancer through potential genetic variation in *CYP2E1* (Elsamanoudy *et al.*, 2016). Additionally, its by-product AFB1, was recently shown to be associated with proliferation in BCa cells (Marchese *et al.*, 2018). It is postulated that environmental factors such as diet, exercise and stress could have epigenetic effects impacting genes involved in carcinogenesis. Perhaps, in this small subset of Nigerian patients, there are epigenetic effects underlying this signature. It would be prudent to conduct epigenetic sequencing of genes involved in the aflatoxin exposure pathway and the eventual carcinogenesis in the liver. Additionally, the high frequency of TNBC metastasis to the liver in WAA may be attributed to the aflatoxin signature and thus warrants further investigation.

Finally, collection of broader environmental risk factors in Nigeria and Barbados was not feasible due to their limited resource settings. If environmental risk factors were collected (*e.g.* behavioural risk factors such as smoking, diet, exercise; SES risk factors), we would have been able to associate them with our epidemiological and genetic findings. If possible, collection of these data would lead to a much richer data set that can further delineate contributions of environmental and genetic factors to tumorigenesis and perhaps how they relate to clinical parameters. Furthermore, at time of data collection it was not possible to obtain mortality data for these cohorts and thus the prognostic value of gene hits were not attainable. Future studies should focus on establishing infrastructures that can allow collection of such data to enhance the datasets within such genomics and gene expression profiling studies.

7.5.2 How is Kaiso regulated and why is there high expression in WAA with TNBC?

To date, multiple studies have highlighted high Kaiso expression to carcinogenesis and tumour aggressiveness. However, little is known about how Kaiso is regulated and what leads to high Kaiso mRNA and protein expression. It is unknown whether high Kaiso expression is due to high copy number or other mutations at the DNA level or through other pathways such as EGFR further activating Kaiso expression. We postulate targeted sequencing of Kaiso in WAA-TNBC tissues to better understand if this high expression is modulated at the DNA level in WAA with TNBC. Other signalling pathways such as TGFβ has been extensively studied in relation to Kaiso and TNBC (Bassey-Archibong *et al.*, 2016) however, the relationship between EGFR and Kaiso in TNBC should be investigated since little is known about how these pathways interact in TNBC. This could be another potential avenue for therapeutics in WAA as EGFR inhibitors are currently undergoing clinical trials for various solid tumours (NCT01137162). Additionally, we found relative lower $p120^{ctn}$ expression in Nigerian and Barbadian TNBC tissues (**Appendix chapter 1, Figure A1.1**). It could be postulated that low p120ctn expression in these WAA-TNBC cohorts is associated with high nuclear Kaiso expression where Kaiso can carry out its transcriptional regulation of target genes. Further studies should investigate this relationship to assess co-localisation of $p120^{ctn}$ and Kaiso in WAA-TNBC samples as well as the related effects of Kaiso's transcriptional activities when bound to $p120^{ctn}$ in the context of TNBC.

7.5.3 Relevance and validation of mutational profile identified

Multiple novel, pathogenic and likely pathogenic mutations were identified in our WAA-TNBC cohort consisting of Barbadian and Nigerian women. It is however currently unknown exactly what effects these variations would have on RNA expression, protein structure and function, as well as how these mutations may be suitable for various therapeutic targets. The first step should be to verify mutations via Sanger sequencing to before any further validation studies. Future follow-up studies require RNA-sequencing on these samples to further understand the effect of these non-silent mutations on RNA expression. Recent studies investigating RNA expression in WAA with TNBC have highlighted unique immune pathways that may be critical in identifying immunotherapeutic targets for these populations (M. Davis *et al.*, 2020). Additionally, using our unique WAA-TNBC-TMA, genes of interest identified should be interrogated to investigate if these alterations influence relative protein expression. As a first step, genes identified with copy number alterations should be investigated initially as the relationship between protein expression and copy number is more causal and direct in comparison with non-silent and nonsense mutations. To further understand the relevance of specific mutations on tumorigenesis, experiments should be designed to test the consequences of these mutations on protein structure and function using *in silico* analyses such as PANTHER and SIFT, and screened *in vitro* for response to various chemotherapeutic agents for cell death and viability. Thereafter, *in vivo* studies could be done to investigate the effects of those mutations with tumour growth, burden and progression using various mouse models. Additionally, due to the low quality and quantity of DNA extracted from our FFPE samples, additional WES or whole genome sequencing (WGS) experiments should be conducted

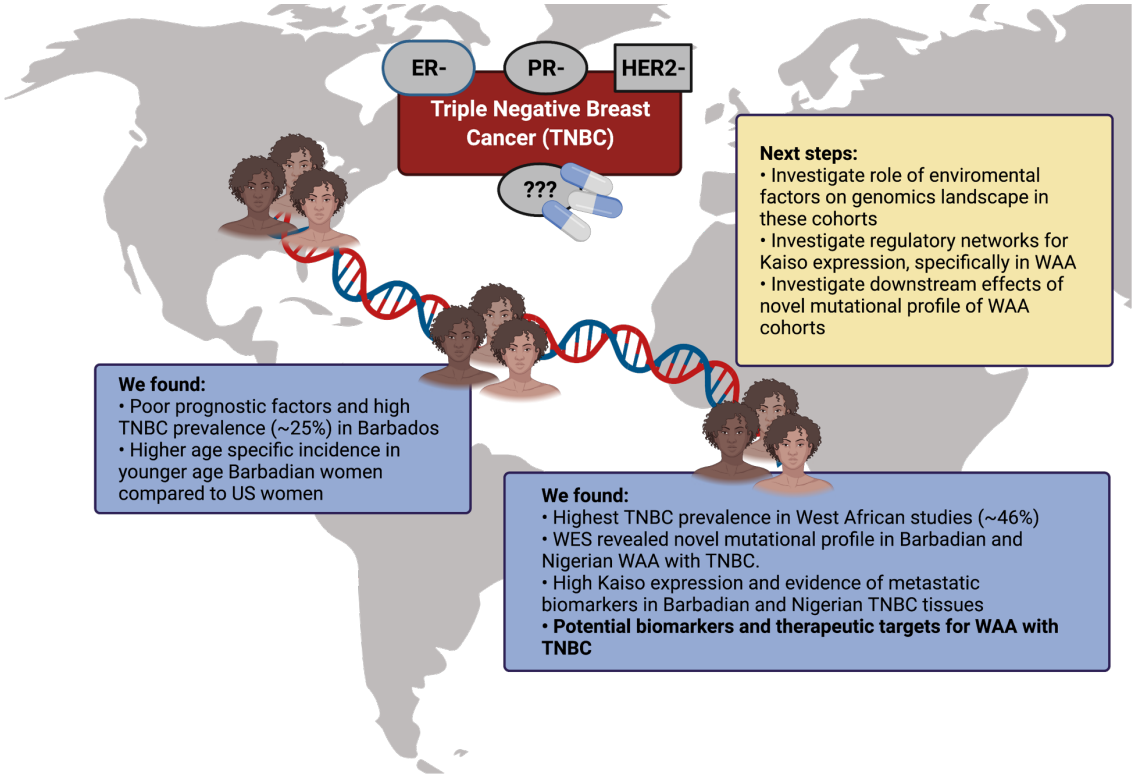
either with fresh frozen tissues which have higher DNA and RNA quality and yield, or with FFPE samples using more advanced techniques for nucleic acid extraction that have been developed after nucleic acid extraction for our study. The use of fresh frozen tissue could also be beneficial for creating patient-derived cell lines that have direct benefit for screening various chemotherapeutic agents for tailored treatment and precision medicine. Additionally, investigating other ancestrally-related WAA TNBC groups would be beneficial to determine if these signals are conserved and shared across multiple groups of WAA with TNBC further solidifying the shared genetic landscape between these populations and highlighting the inclusion of more diverse populations in TNBC research. We currently collected 171 Jamaican tissues for the inclusion on our large WAA-TNBC-TMA; however, due to COVID-19 restrictions and other logistical issues, we were unable to complete the construction of this TMA. However, expanding our experimental designs to include these samples would be of great value.

7.6 CONCLUDING REMARKS

This work has advanced our knowledge about TNBC epidemiology and tumour biology in ancestrally-related WAA. To our knowledge, this project is the first of its kind investigating TNBC in such cohorts (Nigeria and Barbados). We have shed significant light on the epidemiological and genetic factors that contribute to the high prevalence and associated mortality rates of TNBC in WAA in our study. This work has laid a foundation to develop better strategies to detect and treat the often-fatal TNBC especially in the WAA. This study will be effective for gaining genetic insight into TNBC, which currently lacks effective targeted therapies unlike ER/PR/HER2 positive breast cancers. Additionally, the WAA-

TMA cohorts will be a useful resource for BCa researchers worldwide investigating TNBC in WAA. We have generated a framework highlighting what we found through this thesis as well as next steps (**Figure 7.1**) since there are multiple exciting avenues for future research that will bring us closer to better therapeutic options for WAA with TNBC.

Figure 7.1: Conceptual framework for understanding TNBC in WAA and next steps. We found high frequency of TNBC across West African-based studies and a high prevalence of TNBC in Barbadian women. Shared genomic alterations between Nigerian and Barbadian TNBC cohorts may be beneficial for developing therapeutic agents for these groups. Next steps include assessing (i) the roles of environmental factors on the genomics landscape in these cohorts, (ii) the regulatory network(s) that increase Kaiso expression and (iii) the biological roles of identified somatic mutational profile. Figure created with BioRender.com


ER-
PR-
HER2-
Triple Negative Breast Cancer (TNBC)
???
Next steps:
• Investigate role of enviromental factors on genomics landscape in these cohorts
• Investigate regulatory networks for Kaiso expression, specifically in WAA
• Investigate downstream effects of novel mutational profile of WAA cohorts
We found:
• Poor prognostic factors and high TNBC prevalence (~25%) in Barbados
• Higher age specific incidence in younger age Barbadian women compared to US women
We found:
• Highest TNBC prevalence in West African studies (~46%)
• WES revealed novel mutational profile in Barbadian and Nigerian WAA with TNBC.
• High Kaiso expression and evidence of metastatic biomarkers in Barbadian and Nigerian TNBC tissues
• Potential biomarkers and therapeutic targets for WAA with TNBC

APPENDIX 1:

EXPRESSION OF TNBC BIOMARKERS IN WAA TNBC TISSUES

Rationale:

Kaiso was first identified as a binding partner with $p120^{ctn}$ (Daniel & Reynolds, 1999). When present, $p120^{ctn}$ binds and regulates the stability and turnover of the cell adhesion protein E-Cadherin that mediates epithelial cell-to-cell adhesion. Loss of the E-cadherin-p120ctn interaction is thus associated with enhanced motility, invasion and breast cancer metastasis. Increased expression of WAVE3 is also associated with metastasis and has been shown to be indirectly regulated by Kaiso through miRNAs. Women of African Ancestry (WAA) are disproportionately affected by more aggressive breast cancers, especially the TNBC subtype. However, the expression patterns of $p120^{ctn}$ and WAVE3 in TNBC tissues of WAA is currently unknown. Immunohistochemistry (IHC) was used to determine subcellular localisation and expression of these proteins within our WAA-TMA and a commercial TMA with a mixed cohort of African American and Caucasian American TNBC tissues. Pathologist's scoring was obtained for $p120^{ctn}$ and WAVE-3 IHCs. Detailed methods can be found in **Chapter 2**.

Results:

Previously, we reported that Kaiso is highly expressed in WAA TNBC tissues compared to CA TNBC tissues. In this study, IHC was used to evaluate the expression $p120^{ctn}$ and WAVE-3 in TNBC tissues from Nigerian, Barbadian, AA, and CA patients. As shown in Fig. A1.1 (representative images shown), we observed significantly higher cytoplasmic

p120^{ctn} expression in CA TNBC tissues compared to Nigerian and Barbadian TNBC tissues ($p < 0.0001$). Remarkably, we found a significantly higher WAVE3 cytoplasmic expression in Nigerian (n= 18) and Barbadian (n= 29) tissues compared to African American (n= 26) and Caucasian American (n= 57) tissues ($p < 0.0001$) (Figure A1.2) similar to Kaiso expression patterns in these tissues as reported in Chapter 5. We also found significantly higher nuclear WAVE3 expression in the Nigerian TNBC tissues compared to African American and Caucasian American TNBC tissues ($p < 0.0001$). These findings thus indicate a similar increase in WAVE3 expression in WAA tissues with higher African ancestry, which may suggest a role for both Kaiso and WAVE3 expression in the TNBC prevalence and mortality observed in WAA.

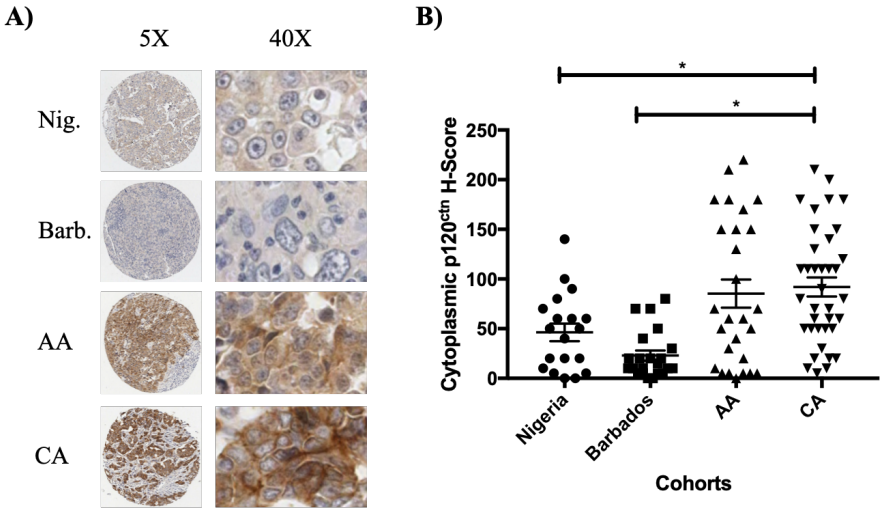


Figure A1.1: Comparative analysis of p120^{ctn} cytoplasmic expression. A) IHC images showing p120^{ctn} expression in Nigerian, Barbadian, AA, and CA TNBC tissues B) Significantly higher cytoplasmic expression in CA and AA tissues compared to Nigerian and Barbadian tissues. *$p<0.05$.

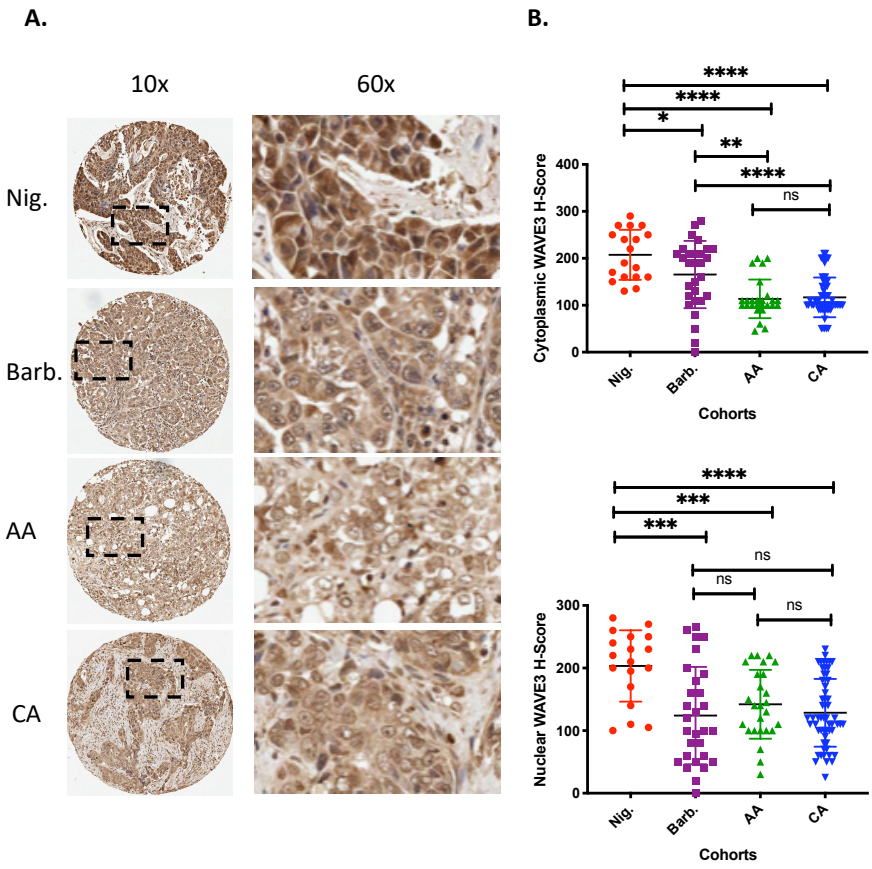


Figure A1.2: High cytoplasmic and nuclear WAVE3 expression in women of African ancestry TNBC tissues. (A) IHC images showing WAVE3 localization to both the nucleus and cytoplasm of Nigerian (Nig.), Barbadian (Barb.), AA, and CA TNBC tissues. (b) Graphical representation of nuclear and cytoplasmic WAVE3 expression in Nigerian (n = 18), Barbadian (n = 29), AA (n = 26), and CA (n = 57) TNBC tissues showing significantly higher cytoplasmic WAVE3 expression in Nigerian, and Barbadian TNBC tissues compared to African American and Caucasian American tissues. Nuclear WAVE3 expression was significantly higher in Nigerian TNBC tissues compared to African American and Caucasian American TNBC tissues. ns not significant, ****$p < 0.0001$, *** $p < 0.001$, ** $p < 0.005$, * $p < 0.05$

Discussion:

Herein, we found reduced cytoplasmic $p120^{ctn}$ expression in Nigerian and Barbadian TNBC tissues compared to AA and CA TNBC tissues. Since $p120^{ctn}$ also stabilizes E-Cadherin it can be postulated that tissues with reduced $p120^{ctn}$ expression will have less anchorage of E-Cadherin and higher EMT. Increased EMT can then lead to increased likelihood of metastasis (Mittal, 2018) and was evidence in findings from our pilot study where WAA had higher lymph node involvement than CA (Bassey-Archibong, Hercules, *et al.*, 2017). Furthermore, increased WAVE3 expression is also linked to metastasis and reduced overall survival and distant metastasis free survival (Sossey-Alaoui, 2013). Unpublished data from our lab suggests that Kaiso regulated WAVE3 expression through interaction with miR31 and miR200c. In TNBC cells, high Kaiso expression represses these miRNAs and leads to increased WAVE3 expression. Based on Kaiso and WAVE3 expression on the pilot TMA, there is a weak positive trend for nuclear Kaiso and WAVE3 expression (data not shown). Further work should be done to examine this relationship more closely as Kaiso may also be regulating EMT and metastasis through its interaction with miRNAs and WAVE3 in addition to what is already known with its regulation of EMT through TGFβ signalling pathways.

APPENDIX 2:

ANNOTATION AND CHARACTERISATION OF LARGE WAA-TMA.

Rationale:

Previous studies have shown that there is a correlation between Kaiso expression and poor prognosis and overall survival in breast cancer (Bassey-Archibong, Rayner, *et al.*, 2017; Vermeulen *et al.*, 2012). To this end, formalin fixed paraffin embedded (FFPE) specimens were obtained from WAA with TNBC in Nigeria and Barbados and a pilot WAA-TMA was constructed which resulted in a publication (Bassey-Archibong, Hercules, *et al.*, 2017). Characterization of this TMA revealed that Nigerian and Barbadian women in our study were diagnosed with TNBC at a younger age than African American (AA) and Caucasian American (CA) women. Analysis of Kaiso expression further revealed that nuclear Kaiso expression scores were significantly higher in Nigerian, Barbadian and AA women compared with CA women. As this was the pilot study, a larger WAA-TMA has been constructed for further validation analyses. In addition to this larger WAA-TMA, we have been given access to 171 TNBC patient FFPE specimens at a local hospital in Jamaica to construct a Jamaican WAA-TMA. It is expected that higher nuclear Kaiso expression is correlated with higher grade and younger age at diagnosis. Furthermore, it is expected that Kaiso expression is significantly correlated with other cancer biomarkers which play a role in cancer aggressiveness and outcomes. In our pilot TMA, we found that there was a trend for a weak negative correlation between nuclear Kaiso expression and BRCA1. With a larger cohort, we will be able to demonstrate this relationship more conclusively.

Methods:

Methods were performed as described in **Chapter 2**. In addition, FFPE BCa tissue blocks were collected from Barbados and Nigeria and shipped to Dr. Anita Bane at Juravinski/McMaster where the larger WAA-TMA was constructed. Using hematoxylin and eosin (H&E) stained slides for histopathological confirmation, tumor areas from each FFPE tissue block were selected for TMA construction. Any sample with less than 1% staining for ER and PR was scored negative; likewise, a score of 0 or +1 for HER2 was considered negative.

Results:

This TMA was constructed with a total of 123 Nigerian, 103 Barbadian. An additional TMA obtained from Yale included 32 AA, and 63 CA BC patients. The mean age at time of diagnosis for Nigerian women was 46.9 years compared to 50.3 years for Barbadian women, 49.5 years for AA women, and 56.2 years for CA women ($p = 0.0006$; Table A2.1). The percentage of younger women who presented with TNBC at time of diagnosis was significantly higher for the Barbadian cohort (66.7%; $n = 39$) compared with the Nigerian (64.7%; $n = 82$), AA (56.5%; $n = 23$), and CA (30.2%; $n = 53$) cohort ($p = 0.0014$) (Table A2.1). Low-grade tumors were seldom observed in the AA (0%; $n = 0$), Barbadian (2.1%; $n = 48$), and CA cohorts (4.1%; $n = 49$) compared to the Nigerian cohort (7.3%; $n = 68$) ($p = 0.0457$). Low-grade was defined as grade 1, medium-grade as grade 2, and high-grade as grade 3, respectively. Stage was not available for the Nigerian patients; however, approximately 25.0% ($n = 20$) of Barbadian women presented with higher stage (T3–T4)

tumours compared with 18.1% (n = 11) for AA and 0% (n = 32) for CA women with TNBC (p = 0.0145; Table A2.1). With respect to lymph node status, there were no observable differences between the groups. Intriguingly, there was statistically higher nuclear Kaiso expression in Barbadian BC samples compared to Nigerian BC samples ($p < 0.05$) (Figure A2.1). However, there was no differential expression for cytoplasmic scores.

Table A2.1: Clinico-pathological characteristics and analysis of study participants

Characteristic	Cohort				
	Nigerian n= 123	Barbadian n- 103	AA n= 32	CA n= 63	p
Mean age, y (SD)*	46.9 (13.5)	50.3 (13.5)	49.5 (10.1)	56.2 (12.0)	0.0006
Age, y*					
≤50	53	26	13	16	0.0014
50	29	13	10	37	
Grade*					
1	5	1	0	2	0.0457
2	29	14	4	23	
3	34	33	19	24	
Stage T*					
T1-T2		15	9	32	0.0145
T3- T4		5	2	0	
Stage N					
N1- N2		13	4	21	0.0627
N3- N4		11	7	7	

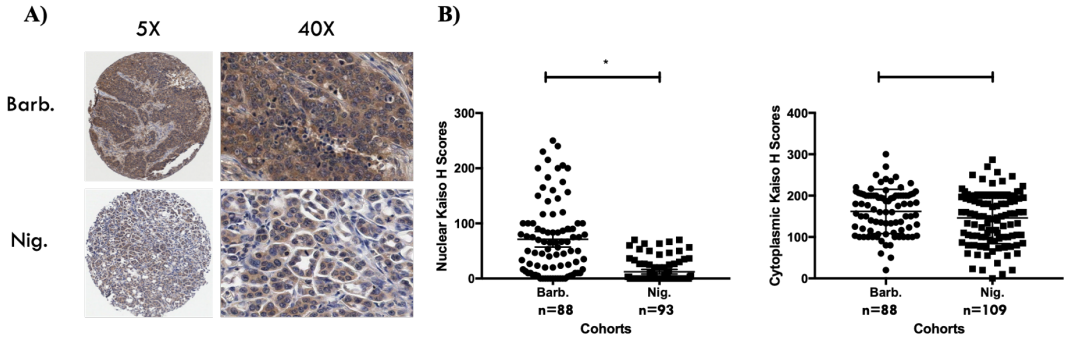


Figure A2.1: Comparative analysis of Kaiso nuclear and cytoplasmic expression. A) IHC images showing Kaiso expression in Nigerian and Barbadian TNBC tissues. B) Significantly higher Nuclear Kaiso expression in Barbadian TNBC tissues compared to Nigerian TNBC tissues. No statistically significant differential expression observed between Barbadian and Nigerian groups for cytoplasmic Kaiso expression. *$p<0.05$.

Discussion:

We found WAA in our cohort were diagnosed at younger ages, with higher grade tumours and larger tumours compared to CA (Table A2.1) as was also found in our previous work (Bassey-Archibong, Hercules, *et al.*, 2017). Interestingly, we found that there was higher nuclear Kaiso expression for Barbadian samples compared to Nigerian samples on this larger TMA which is contrary to what was previously published. Pathologist assessment of subsequently stained TMA including the AA and CA tissues are underway to fully investigate differential sub-cellular localisation and relative expression.

APPENDIX 3:

KAISO IS HIGHLY EXPRESSED IN TNBC TISSUES OF WOMEN OF AFRICAN ANCESTRY COMPARED TO CAUCASIAN WOMEN

Preface

This article reports the findings from the pilot TMA study conducted by Dr. Blessing Bassey-Archibong entitled: ***"Kaiso is highly expressed in TNBC tissues of women of African ancestry compared to Caucasian women"*** by Bassey-Archibong BI, Hercules SM, Rayner LGA, Skeete DHA, Connell SPS, Brain I, Daramola A, Banjo AAF, Byun JS, Gardner K, Dushoff J and Daniel JM, which has been reproduced in its original form (Cancer Causes and Control, 2017; 8(3): e2689). This is an open-access article distributed under the Creative Commons CC-BY License, which permits unrestricted reproduction and dissemination in any medium, provided the authors, attribution parties and sources are acknowledged.

This retrospective study examined the clinical parameters associated with TNBC in women of African ancestry (WAA) across Barbados, indigenous African women in a Nigerian cohort, African American and Caucasian American women, and Kaiso expression patterns in TNBC tissues obtained from WAA and Caucasian women. The study was conceived and designed in an effort to characterize Kaiso expression in TNBC tumors from WAA, as Kaiso was reported to be more highly expressed in TNBC, which is most prevalent in premenopausal WAA compared to Caucasian women. We found that WAA are diagnosed with TNBC at younger ages than Caucasian women and have more high-

grade and lymph node positive tumors. Importantly, we found that nuclear Kaiso is more highly expressed in TNBC tissues of WAA compared to Caucasian TNBC tissues. The highest nuclear Kaiso expression was observed in TNBC tissues from WAA that have a higher degree of African ancestry (Nigerian and Barbadian), suggesting a possible role for Kaiso in the racial disparity associated with TNBC prevalence.

Contributions:

Dr. JM Daniel and BI Bassey-Archibong conceived the study and co-wrote the manuscript. Dr. BI Bassey-Archibong executed the research trip to collect TNBC tissues from Nigerian women while SM Hercules undertook the research trip to collect TNBC tissues and clinical data from Barbadian women as well as data management of TMA maps, WAA-TMA and American TMA, H-scores and subsequent statistical analyses. Dr. BI Bassey-Archibong stained the TMAs and performed the statistical analysis that generated the Figures in 1A-C, and Figures 4A-D, and generated the data for Figures 3A and B. LG Rayner generated the data for Figure 2A and B. Dr. Desiree Skeete and Dr. Ian Brain scored the Kaiso immunostain represented in Figures 3B and Figures 4A-D. Dr. JM Daniel also provided significant guidance and intellectual input throughout the course of the study. All other authors assisted with the recruitment of patient populations (Dr. A Daramola, Dr. A Banjo, Dr. DHA Skeete and Dr. Smith Connell) and interpretation of the data (Dr. J Byun, Dr. K Gardner and Dr. J Dushoff). All authors edited the manuscript text.

Cancer Causes Control (2017) 28:1295–1304
DOI 10.1007/s10552-017-0955-2

ORIGINAL PAPER

Kaiso is highly expressed in TNBC tissues of women of African ancestry compared to Caucasian women

Blessing I. Bassey-Archibong[1] · Shawn M. Hercules[1] · Lyndsay G. A. Rayner[1] · Desiree H. A. Skeete[2,3] · Suzanne P. Smith Connell[3,4] · Ian Brain[5] · Adetola Daramola[6] · Adekunbiola A. F. Banjo[6] · Jung S. Byun[7] · Kevin Gardner[7] · Jonathan Dushoff[1] · Juliet M. Daniel[1]

Received: 25 February 2017 / Accepted: 31 August 2017 / Published online: 8 September 2017

Abstract
Purpose Triple-negative breast cancer (TNBC) is most prevalent in young women of African ancestry (WAA) compared to women of other ethnicities. Recent studies found a correlation between high expression of the transcription factor Kaiso, TNBC aggressiveness, and ethnicity. However, little is known about Kaiso expression and localization patterns in TNBC tissues of WAA. Herein, we analyze Kaiso expression patterns in TNBC tissues of African (Nigerian), Caribbean (Barbados), African American (AA), and Caucasian American (CA) women.
Methods Formalin-fixed and paraffin embedded (FFPE) TNBC tissue blocks from Nigeria and Barbados were utilized to construct a Nigerian/Barbadian tissue microarray (NB-TMA). This NB-TMA and a commercially available TMA comprising AA and CA TNBC tissues (AA-CA-YTMA) were subjected to immunohistochemistry to assess Kaiso expression and subcellular localization patterns, and correlate Kaiso expression with TNBC clinical features.
Results Nigerian and Barbadian women in our study were diagnosed with TNBC at a younger age than AA and CA women. Nuclear and cytoplasmic Kaiso expression was observed in all tissues analyzed. Analysis of Kaiso expression in the NB-TMA and AA-CA-YTMA revealed that nuclear Kaiso H scores were significantly higher in Nigerian, Barbadian, and AA women compared with CA women. However, there was no statistically significant difference in nuclear Kaiso expression between Nigerian versus Barbadian women, or Barbadian versus AA women.
Conclusions High levels of nuclear Kaiso expression were detected in patients with a higher degree of African heritage compared to their Caucasian counterparts, suggesting a role for Kaiso in TNBC racial disparity.

Keywords Kaiso · TNBC · Women of African ancestry · Breast cancer racial disparity

Electronic supplementary material The online version of this article (doi:10.1007/s10552-017-0955-2) contains supplementary material, which is available to authorized users.

✉ Juliet M. Daniel
danielj@mcmaster.ca

[1] Department of Biology, McMaster University, Hamilton, ON, Canada

[2] Department of Pathology, Queen Elizabeth Hospital (QEH), Bridgetown, Barbados

[3] Faculty of Medical Sciences, The University of the West Indies, Cave Hill Campus, Bridgetown, Barbados

[4] Department of Radiation Oncology, Queen Elizabeth Hospital (QEH), Bridgetown, Barbados

[5] Department of Pathology and Molecular Medicine, McMaster University, Hamilton, ON, Canada

[6] Department of Anatomic and Molecular Pathology, Lagos University Teaching Hospital (LUTH), Lagos, Nigeria

[7] Genetics Branch, National Institute of Health, Bethesda, MD, USA

Introduction

Breast cancer (BCa) is a complex disease that occurs mostly in females and is a leading cause of female deaths worldwide [1–3]. The triple-negative breast cancer (TNBC) subtype accounts for a disproportionate number of BCa deaths due to its highly aggressive nature and metastatic tendencies [4–6]. As the name implies, triple-negative tumors represent a subset of breast tumors that are negative for the estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor

receptor-2 (HER2) [7]. Most TNBC are classified as basal-like cancers and are generally characterized by high histologic/nuclear grade, increased rate of recurrence, and a greater frequency of epidermal growth factor receptor (EGFR) amplification, p53 mutations, and breast cancer type 1 (BRCA1) mutations [7, 8]. Due to their triple-negative status for ER, PR, and HER2, TNBCs lack targeted-treatment options, and cannot be treated with hormonal (Tamoxifen) or anti-HER2 therapies [7].

There is increasing evidence that TNBC occurs more frequently in young premenopausal African and AA women compared to Caucasian women [7, 9–14]. For example, Stark and colleagues reported that among Ghanaian BCa cases, there was a TNBC prevalence of ~82% compared to the USA where TNBC prevalence was ~33% and ~10% among AA and CA cases, respectively [11]. Similarly, Agboola et al. reported a high incidence of TNBC among BCa cases in Nigerian women (~48%) compared with British women (~14%) [14]. The trend of high TNBC prevalence in AA and African females strongly suggests an ancestral genetic predisposition to TNBC in women of African ancestry (WAA) [15–17]. More disturbing, however, is the poor survival rate of AA TNBC patients compared with Caucasian TNBC patients [10, 18], which underscores the urgency to identify potential prognostic or diagnostic TNBC biomarkers in WAA.

Recent studies have found a correlation between increased nuclear expression of the transcription factor Kaiso and poor overall survival of AA breast cancer and prostate cancer patients compared to their Caucasian counterparts [19, 20]. These data hint at a role for Kaiso in the racial disparity in outcomes associated with breast and prostate cancer. Kaiso was first identified as a binding partner of the E-cadherin catenin cofactor—p120-catenin [21]. Kaiso is a dual-specificity transcription factor and member of the POZ-ZF family of transcription factors [21–25] that are implicated in vertebrate development and tumorigenesis. Kaiso has been most often characterized as a transcriptional repressor [26], but some studies indicate that Kaiso can also function as a transcriptional activator [27, 28]. Notably, several Kaiso target genes identified to date (*cyclinD1, matrilysin, E-cadherin*) have been linked to tumor onset, invasion, and metastasis [29–31].

Since its discovery, Kaiso has been implicated in the poor prognostic outcomes of several cancers including colorectal, non-small cell lung cancer, prostate, pancreatic ductal adenocarcinoma, and TNBC [20, 32–35]. Studies from our lab and others indicates that Kaiso plays both pro-oncogenic and tumor suppressive roles in several human cancers [19, 20, 33, 34, 36–38]. Notably, in addition to being implicated in racial disparities in breast cancer outcomes, high Kaiso expression correlates significantly with ER-α negativity, and the aggressiveness of basal/TNBCs [35, 38]. To date however, no studies have specifically examined and compared Kaiso expression and subcellular localization in TNBC tissues from WAA, who have the highest prevalence and worst outcomes from TNBC compared to Caucasian women. In this retrospective study, we evaluated Kaiso expression in TNBC specimens from Nigerian, Barbadian, AA, and CA patients. We found that nuclear Kaiso expression was significantly increased in TNBC tissues of Nigerian, Barbadian, and AA patients compared with their Caucasian counterparts. While there was no significant difference in nuclear Kaiso expression in TNBC tissues of Nigerian versus Barbadian patients (who have a higher percentage of African ancestry compared to AA), we found significantly more nuclear Kaiso expression in Nigerian versus AA patients, and a trend towards higher nuclear Kaiso expression in Barbadian versus AA patients. Collectively, these findings suggest that Kaiso may play a role in the racial disparity associated with TNBC in WAA.

Methods

Study population and characteristics of tumor samples

FFPE TNBC tissue blocks of 28 Nigerian TNBC patients diagnosed between 2011 and 2013 at the Lagos University Teaching Hospital (LUTH), Nigeria, and 46 Barbadian TNBC patients diagnosed between 2002 and 2011 at the Queen Elizabeth Hospital (QEH), Barbados were obtained from the archives of the Department of Anatomic and Molecular Pathology at LUTH and the Department of Pathology at QEH after approval by LUTH and QEH Ethics committees, respectively. The FFPE specimens were then shipped to the Developmental Histology Lab at the Yale Pathological Tissue Services (YPTS), Yale University (Connecticut, New Haven, USA), where they were hematoxylin and eosin (H&E) stained for histopathological confirmation, before tumor areas from each FFPE tissue block were selected for the construction of a Nigerian and Barbadian TNBC tissue microarray (NB-TMA). ER, PR, and HER2 status of the Nigerian tissues were confirmed by immunohistochemistry (IHC) conducted at LUTH, while ER, PR, and HER2 status of the Barbadian tissues were confirmed by IHC conducted at QEH, Barbados, the Human Tissue Resource Center (Chicago, IL, USA) or the Immunohistochemistry Lab at the University of Miami, Miller School of Medicine (Clinical Research Building, Miami, FL, USA). Any sample with less than 1% staining for ER and PR was scored negative; likewise, 0 or +1 for HER2 was considered negative. Available clinico-pathological data (age, tumor pathology, lymph node involvement, and grade)

were retrieved from the hardcopy pathology reports at LUTH and QEH, and are summarized in Table 1.

For the AA and CA patient population, we utilized the Yale tissue microarray 347 (YTMA-347), which was generated at the Yale Developmental Histology Lab, and comprised of 20 AA and 43 CA usable TNBC specimens that were diagnosed at the Yale-New Haven Hospital, Connecticut, USA between 1996 and 2004. ER, PR, and HER2 status were determined by IHC at the Yale Developmental Histology Lab. The clinico-pathological features of the YTMA-347 cohort are summarized in Table 1.

Immunohistochemistry

5-μm tissue sections prepared from the NB-TMA tissue block and the purchased YTMA-347 tissue slides were deparaffinized by warming at 60 °C for 20 min, followed by immersion in xylenes for 10 min. Tissue sections were then rehydrated in descending ethanol dilutions before they were subjected to heat antigen retrieval in a low pH buffer (pH 6.0) solution (DAKO, Glostrup, Denmark). Endogenous biotin, biotin receptors, and avidin binding sites on tissues were subsequently blocked using the Avidin/Biotin blocking kit (Vector Laboratories, Inc., Burlingame, CA, USA), while endogenous peroxidase activity was quenched by treatment with 3% hydrogen peroxide. Tissue slides were stained with mouse anti-Kaiso 6F monoclonal (1:10,000; [39]) or mouse anti-human cytokeratin clones AE1/AE3 monoclonal (1:500; Dako North America, Inc., Carpinteria, CA, USA) primary antibodies overnight at 4 °C, followed by secondary antibody incubations at room temperature for 2 h with biotinylated donkey anti-mouse secondary antibody (Vector Labs; 1:1000). Tissues were subsequently incubated in Vectastain (Vector Labs) for 30 min, rinsed in 1X PBS, and then incubated in diaminobenzidine (DAB) (Vector Labs) for 10 min. Counterstaining was achieved by incubating tissues in Harris hematoxylin (Sigma) for 10–60 s, followed by rinsing in tap water or as described in [40]. Slides were then dehydrated in ascending alcohol dilutions, and cleared with two rounds of xylenes before being mounted using Polymount (Polysciences Inc., Warrington, PA, USA). Negative control staining data were achieved by slide incubation with secondary antibodies only. Images of stained slides were captured using the Aperio Slide scanner (Leica Biosystems, ON, Canada). Stained tissues were scored blindly by two Pathologists, and the scores averaged to give a final score value. The intensity of staining was scored as 0, 1, 2, or 3 representing no, mild, moderate, or high staining intensity. The modified histochemical score (H-score) system was then used to generate the total score for each tissue with values spanning 0–300 using the formula: 3 × (percentage of cells with high intensity staining (3+) + 2 × (percentage of cells with moderate intensity staining (2+) + 1 × (percentage of cells with mild intensity staining (1+) for each slide.

Table 1 Clinico-pathological characteristics and analysis of study participants

	Nigerian (%) $n = 28$	Barbadian (%) $n = 46$	African American (%) $n = 20$	Caucasian American (%) $n = 43$	χ^2 value	p value
Age (years)						
≤50	20 (71.4%)	21 (45.7%)	6 (30.0%)	10 (23.3%)	16.89	0.0007
>50	5 (17.9%)	25 (54.3%)	7 (35.0%)	27 (62.8%)		
Unknown[a]	3 (10.7%)	0 (0.0%)	7 (35.0%)	6 (13.9%)		
Grade						
1	5 (17.9%)	0 (0.0%)	2 (10.0%)	13 (30.2%)	63.59	<0.0001
2	8 (28.6%)	9 (19.6%)	11 (55.0%)	21 (48.9%)		
3	10 (35.7%)	35 (76.1%)	0 (0.0%)	1 (2.3%)		
Unknown[a]	5 (17.8%)	2 (4.3%)	7 (35.0%)	8 (18.6%)		
Stage T						
T1–T2	6 (21.4%)	18 (39.1%)	7 (35.0%)	23 (53.5%)	30.52	<0.0001
T3–T4	11 (39.3%)	1 (2.2%)	1 (5.0%)	0 (0.0%)		
Unknown[a]	11 (39.3%)	27 (58.7%)	12 (60.0%)	20 (46.5%)		
Stage N						
N0	4 (14.3%)	11 (23.9%)	7 (35.0%)	26 (60.5%)	10.23	0.02
N1-N3	11 (39.3%)	10 (21.7%)	13 (65.0%)	12 (27.9%)		
Unknown[a]	13 (46.4%)	25 (54.4%)	0 (0.0%)	5 (11.6%)		

[a]Unknown cases were exempted from analysis

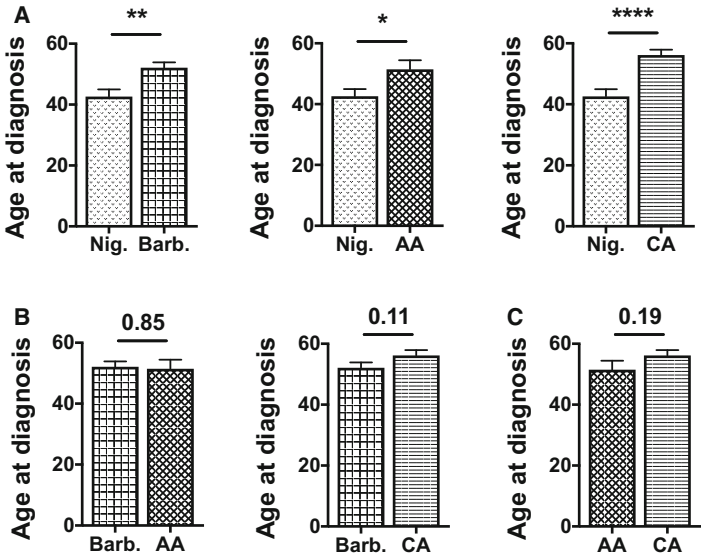


Fig. 1 Nigerian women are diagnosed with TNBC at younger ages than Barbadian, AA, and CA women. **a** The mean age at diagnosis for Nigerian TNBC patients was 42.6 years ($n = 25$) compared with 52.1 years for Barbadian women ($n = 46$), 51.5 years for AA women ($n = 13$), and 56.2 years for CA women ($n = 37$). No significant differences were observed between the mean age at diagnosis for Barbadian versus AA and CA TNBC patients (**b**) and for AA versus CA patients (**c**). $^{*}p < 0.05$, $^{**}p < 0.005$, $^{****}p < 0.0001$

Statistical analysis

GraphPad Prism statistical software (GraphPad Software Inc., La Jolla, CA, USA) was used for all statistical analyses. Standard unpaired Student's t test with Welch's correction was used for pairwise comparison of means. Chi square analysis was used to assess the difference in clinico-pathological features between the Nigerian, Barbadian, AA, and CA cohorts. Data are presented as mean ± SEM where applicable. For all statistical tests, p values <0.05 denote statistical significance.

Results

Clinico-pathological characteristics of study participants

This retrospective study involved a total of 28 Nigerian, 46 Barbadian, 20 African American (AA), and 43 Caucasian American (CA) TNBC patients. The mean age at time of diagnosis for Nigerian women was 42.6 years compared to 52.1 years for Barbadian women ($p = 0.002$), 51.5 years for AA women ($p = 0.03$), and 56.2 years for CA women ($p < 0.0001$; Fig. 1a). Comparison of the mean age at diagnosis between Barbadian, AA, and CA patients yielded no statistical significance (Fig. 1b, c). The percentage of younger women who presented with TNBC at time of diagnosis was significantly higher for the Nigerian cohort (71.4%; $n = 20$) compared with the Barbadian (45.7%; $n = 21$), AA (30.0%; $n = 6$), and CA (23.3%; $n = 10$) cohort ($p < 0.001$) (Table 1). Low-grade tumors were seldom observed in the Nigerian (17.9%; $n = 5$), Barbadian (0%; $n = 0$), and AA (10.0%; $n = 2$) cohorts compared to the CA (30.2%; $n = 13$) cohort ($p < 0.0001$; Table 1). Low-grade was defined as grade 1, medium-grade as grade 2, and high-grade as grade 3, respectively. Approximately 39.3% ($n = 11$) of Nigerian women presented with higher stage (T3–T4) tumors compared with 2.2% ($n = 1$) for Barbadian, 5.0% ($n = 1$) for AA, and 0% ($n = 0$) for CA women ($p < 0.0001$; Table 1). Finally, CA TNBC patients displayed a higher frequency of lymph node-negative tumors (60.5%; $n = 26$) compared with that observed in Nigerian (14.3%; $n = 4$), Barbadian (23.9%; $n = 11$), and AA (35.0%; $n = 7$) TNBC patients ($p = 0.02$; Table 1).

Kaiso is highly expressed in TNBC tissues of WAA compared to Caucasian women

Previously, we reported that Kaiso is highly expressed at the mRNA level in triple-negative tumors compared with

Cancer Causes Control (2017) 28:1295–1304

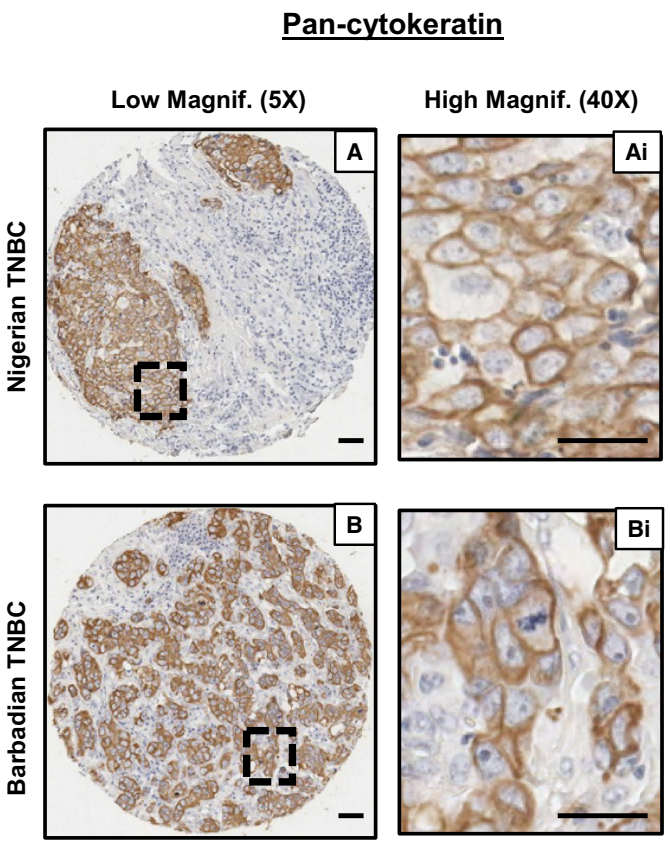


Fig. 2 Cytokeratin immunostaining of Nigerian and Barbadian TNBC tissues verifies tissue integrity. IHC images at low (5×) and high magnification (40×) show intact tissue cores (**a**, **b**) and membrane localization (**ai**, **bi**) of cytokeratin, which portrays good integrity of the Nigerian and Barbadian tissues. *Scale bar* 50 μm

hormone receptor-positive breast tumors in publicly available datasets downloaded from The Cancer Genome Atlas—TCGA website or the Gene Expression Omnibus—GEO website [35]. Thus, in this study, we utilized immunohistochemistry to specifically evaluate the expression and subcellular localization of Kaiso in TNBC tissues from Nigerian, Barbadian, AA, and CA patients. Tissue integrity of the Nigerian and Barbadian TNBC tissues was determined by immunostaining for pan-cytokeratin as described in the methods; Fig. 2a, b shows representative images of the tissue quality of the Nigerian and Barbadian TNBC tissues. As shown in Fig. 3a (representative images shown), Kaiso exhibited both nuclear and cytoplasmic localization in all TNBC tissues analyzed, with varying degrees of heterogeneity. Nuclear and cytoplasmic Kaiso staining intensity was scored as described in the methods, and Kaiso's relative expression in each TNBC cohort analyzed. As seen in Fig. 3b, we observed significantly higher cytoplasmic than nuclear Kaiso expression in the AA and CA TNBC cohorts ($p < 0.0001$), but did not find significant differences between nuclear and cytoplasmic Kaiso expression in the Nigerian and Barbadian TNBC cohorts.

Since nuclear but not cytoplasmic Kaiso expression is known to be associated with TNBC aggressiveness, and decreased survival of AA BCa patients [19, 38], we next performed comparative analysis of nuclear Kaiso expression between the Nigerian, Barbadian, AA, and CA cohorts. Interestingly, we observed a significantly higher level of nuclear Kaiso expression in TNBC tissues of patients of African ancestry (Nigerian, Barbadian, and AA) compared to their Caucasian counterparts (Fig. 4a). However, there was no significant difference between nuclear Kaiso expression in TNBC tissues of Nigerian and

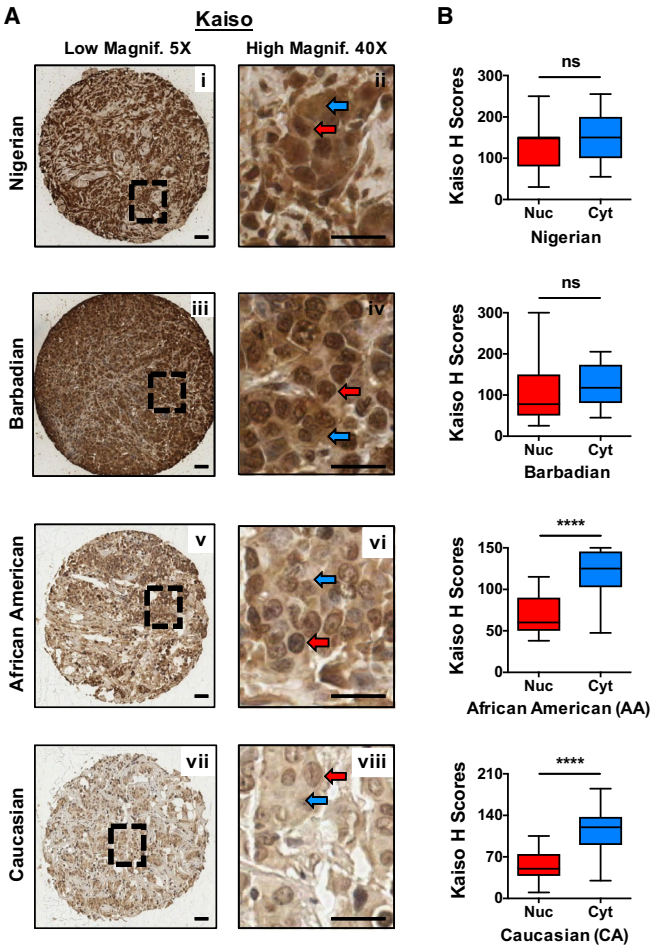


Fig. 3 Kaiso subcellular localization and expression in Nigerian, Barbadian, AA, and CA TNBC tissues. (**a**i–viii) IHC images showing Kaiso localization to both the nucleus and cytoplasm of Nigerian, Barbadian, AA, and CA TNBC tissues. (**b**) Graphical representation of nuclear and cytoplasmic Kaiso expression in Nigerian ($n = 19$), Barbadian ($n = 20$), AA ($n = 20$), and CA ($n = 39$) TNBC tissues. Cytoplasmic Kaiso expression was significantly higher than nuclear Kaiso expression in the AA and CA TNBC cohorts but not in the Nigerian and Barbadian TNBC cohorts. Red arrows indicate nuclear Kaiso staining, while blue arrows indicate cytoplasmic Kaiso staining. *Scale bar* 50 µm. *ns* not significant, $^{****}p < 0.0001$

Barbadian patients, who have ~99.8 and ~77.4% degree of African heritage, respectively [41, 42], or between TNBC tissues of Barbadian and AA patients, who have ~77.4 and ~72.5% degree of African heritage, respectively [42] (Fig. 4b). Remarkably however, there was significantly more nuclear Kaiso expression in TNBC tissues of Nigerian compared to AA patients (Fig. 4c), probably due to the higher degree of African heritage in Nigerian patients (~99.8%) compared to AA patients (~72.5%). Since TNBC is more prevalent in WAA compared to Caucasian women, these findings suggest a role for nuclear Kaiso expression levels in the racial disparity in TNBC prevalence.

Correlation between nuclear Kaiso expression and clinico-pathological features of study participants

Breast tumors of WAA are often associated with a higher histological grade and positive lymph node involvement compared to breast tumors of Caucasian women [11, 14]. Since previous studies from our lab and others have

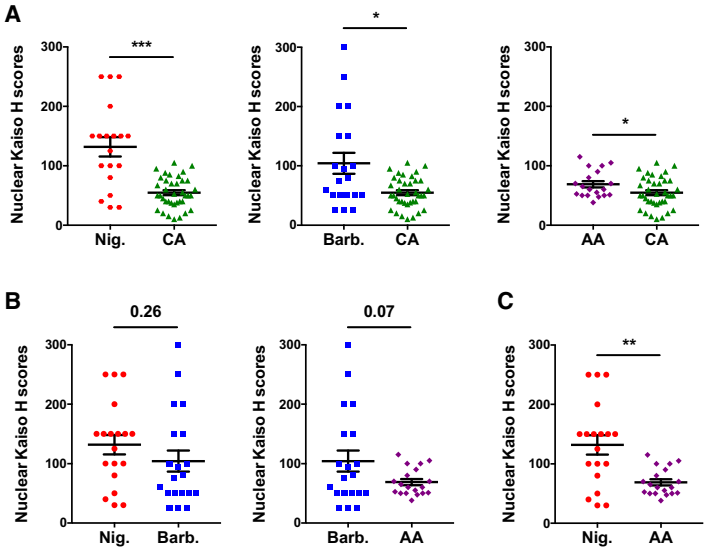


Fig. 4 Comparative analysis of nuclear Kaiso expression in Nigerian, Barbadian, AA, and CA TNBC tissues. Higher levels of nuclear Kaiso expression were detected in TNBC tissues of Nigerian, Barbadian, and AA compared with their Caucasian counterparts (**a**). Although no significant difference in nuclear Kaiso expression was observed between Nigerian versus Barbadian tissues, or between Barbadian versus AA tissues (**b**), there was a significant difference in nuclear Kaiso expression between Nigerian and AA TNBC tissues (**c**). $^{*}p < 0.05$, $^{**}p < 0.005$, $^{***}p < 0.001$

correlated increased Kaiso expression with advanced grade and metastasis of TNBC [35, 38], and lymph node involvement is an established prognostic marker for the metastatic potential of breast tumors [43], we next assessed the association of Kaiso expression with high-grade and lymph node involvement in Nigerian, Barbadian, AA, and CA patients. High-grade tumors were defined as grade 3 for Nigerian and Barbadian patients and grade 2 for AA and CA patients due to no analyzed grade 3 tumors in the AA and CA TNBC cohort (the only observed grade 3 CA patient could not be scored as a result of tissue loss). Low-grade tumors were thus defined as grades 1 and 2 for Nigerian and Barbadian patients, and grade 1 for AA and CA patients. Lymph node metastasis was considered positive if one or more lymph nodes were noted to contain cancer cells (n1–n3), and negative if there were no observed cancer cells in the lymph nodes (n0). Due to the small sample size used in the analysis, no significant correlation was found between high nuclear Kaiso expression and high-grade or lymph node-positive triple-negative tumors in any of the patient cohorts analyzed (Suppl. Figure 1).

Discussion

TNBC is most prevalent in WAA compared to Caucasian American/European females, but the reason for this disparity is currently unknown [11, 14, 16, 44]. Although poor socio-economic status has been linked to TNBC mortality in African and AA women, it does not fully explain the disproportionate prevalence and aggressiveness of TNBC in WAA compared to their Caucasian counterparts [17]. Thus, we and others have postulated that there may be an ancestral genetic predisposition to TNBC in WAA [17, 45].

Notably, a higher prevalence of TNBC has been reported in West-African women (Nigerians—65%, and Ghanaians—82.2%) compared with that reported in AA—$\sim$33% [9, 11, 46], thus supporting the idea of a relationship between percentage of African ancestry and TNBC prevalence. Since West-African countries such as Ghana and Nigeria are the founding ancestors of most WAA worldwide [41, 42, 47–49], we posit that there is a higher probability of identifying a founder mutation, if one exists, in Nigerian and Ghanaian populations, and also in more

homogeneous populations of the African Diaspora such as the Caribbean (e.g., Barbados).

Recent studies have linked high nuclear expression of the transcription factor Kaiso with increased TNBC aggressiveness [20, 38], and decreased survival of AA breast cancer patients compared with their Caucasian counterparts [19]. These reports suggest a link between increased nuclear Kaiso, TNBC aggressiveness/metastasis, and the racial disparity in prevalence/outcomes associated with breast cancer. Remarkably, our findings lend some credence to this hypothesis as we observed elevated expression of nuclear Kaiso in TNBC tissues from patients of African ancestry (Nigerians, Barbadians, and African Americans) compared to their Caucasian/European ancestry counterparts (CA) (see Fig. 4a). Thus, our previous findings in Kaiso-depleted mouse xenograft models [35, 51], where we demonstrated roles for Kaiso in TNBC cell growth, survival, and metastasis, may explain why high Kaiso-expressing triple-negative tumors in WAA are associated with a more aggressive phenotype and fatal outcomes than TNBC in Caucasian women.

Importantly, our findings highlight an interesting correlation between high nuclear Kaiso expression and percent African ancestry, which may be linked to the predisposition of young WAA to TNBC. However, this study is limited by the small sample size, the semi-quantitative method of analysis used, and lack of complete clinico-pathological information, which did not allow proper assessment of the correlation between Kaiso expression and the high tumor grade observed in African/Caribbean women compared to African American or Caucasian women. Additional studies using larger cohort sizes of West-African (Nigeria and others), Caribbean (Barbados and others), AA, and CA TNBC cases, coupled with quantitative methods of immunostain analysis such as the automated quantitative analysis (AQUA) system established by Rimm and colleagues [50], will undoubtedly provide more insight into the clinical relevance of nuclear Kaiso expression in the etiology of TNBC in WAA.

In conclusion, this is the first study to suggest a potential link between increased Kaiso expression and the predisposition of young WAA to TNBC. This observation, in addition to the previous identified roles for Kaiso in TNBC aggressiveness, metastasis, and poor overall survival in affected patients [35, 38, 51], raises two exciting possibilities: i) Kaiso expression could be utilized as a biomarker for the diagnosis and prognosis of TNBC in WAA and ii) Kaiso could be a molecular target for the development of treatment options against TNBC not only in WAA but also TNBC patients worldwide.

Acknowledgments We would like to thank the staff and personnel of the Pathology and Oncology departments at QEH and LUTH, for their assistance with the samples used for this project. We would also like to thank Lori Charette and Dr. David Rimm (Department of Pathology, Yale University School of Medicine, New Haven, Connecticut, USA) for their assistance with the construction of the Nigerian–Barbadian TMA. This work was funded in part by the Canadian Breast Cancer Foundation (CBCF), the Juravinski Hospital and Cancer Center Foundation (JHCCF), and the Natural Sciences and Engineering Research Council of Canada (NSERC). BIB-A was partly supported by a Schlumberger Faculty for the Future Fellowship.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All procedures performed in this retrospective study were in accordance with the ethical standards of LUTH and QEH, respectively. For this type of study formal consent is not required.

References

1. Hortobagyi GN, de la Garza Salazar J, Pritchard K, Amadori D, Haidinger R, Hudis CA et al (2005) The global breast cancer burden: variations in epidemiology and survival. Clin Breast Cancer. 6(5):391–401
2. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D (2011) Global cancer statistics. CA Cancer J Clin 61(2):69–90
3. Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C et al (2013) GLOBOCAN 2012 v1.0, Cancer incidence and mortality worldwide: IARC CancerBase No. 11 [Internet]. International Agency for Research on Cancer, Lyon
4. Oakman C, Viale G, Di Leo A (2010) Management of triple negative breast cancer. Breast 19(5):312–321
5. Carey LA (2011) Directed therapy of subtypes of triple-negative breast cancer. Oncologist 16(Suppl 1):71–78
6. Andre F, Zielinski CC (2012) Optimal strategies for the treatment of metastatic triple-negative breast cancer with currently approved agents. Ann Oncol 23(Suppl 6):vi46–vi51
7. Foulkes WD, Smith IE, Reis-Filho JS (2010) Triple-negative breast cancer. N Engl J Med 363(20):1938–1948
8. Irshad S, Ellis P, Tutt A (2011) Molecular heterogeneity of triple-negative breast cancer and its clinical implications. Curr Opin Oncol 23(6):566–577
9. Carey LA, Perou CM, Livasy CA, Dressler LG, Cowan D, Conway K et al (2006) Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. JAMA 295(21):2492–2502
10. Lund MJ, Trivers KF, Porter PL, Coates RJ, Leyland-Jones B, Brawley OW et al (2009) Race and triple negative threats to breast cancer survival: a population-based study in Atlanta, GA. Breast Cancer Res Treat 113(2):357–370
11. Stark A, Kleer CG, Martin I, Awuah B, Nsiah-Asare A, Takyi V et al (2010) African ancestry and higher prevalence of triple-

negative breast cancer: findings from an international study. Cancer 116(21):4926–4932
12. Bhikoo R, Srinivasa S, Yu TC, Moss D, Hill AG (2011) Systematic review of breast cancer biology in developing countries (part 1): Africa, the Middle East, Eastern Europe, Mexico, the Caribbean and South America. Cancers 3:2358–2381
13. Amirikia KC, Mills P, Bush J, Newman LA (2011) Higher population-based incidence rates of triple-negative breast cancer among young African-American women: implications for breast cancer screening recommendations. Cancer 117(12):2747–2753
14. Agboola AJ, Musa AA, Wanangwa N, Abdel-Fatah T, Nolan CC, Ayoade BA et al (2012) Molecular characteristics and prognostic features of breast cancer in Nigerian compared with UK women. Breast Cancer Res Treat 135(2):555–569
15. Amend K, Hicks D, Ambrosone CB (2006) Breast cancer in African-American women: differences in tumor biology from European-American women. Can Res 66(17):8327–8330
16. Boyle P (2012) Triple-negative breast cancer: epidemiological considerations and recommendations. Ann Oncol 23(Suppl 6):vi7–vi12
17. Dietze EC, Sistrunk C, Miranda-Carboni G, O'Regan R, Seewaldt VL (2015) Triple-negative breast cancer in African-American women: disparities versus biology. Nat Rev Cancer 15(4):248–254
18. Bauer KR, Brown M, Cress RD, Parise CA, Caggiano V (2007) Descriptive analysis of estrogen receptor (ER)-negative, progesterone receptor (PR)-negative, and HER2-negative invasive breast cancer, the so-called triple-negative phenotype: a population-based study from the California cancer Registry. Cancer 109(9):1721–1728
19. Jones J, Wang H, Karanam B, Theodore S, Dean-Colomb W, Welch DR et al (2014) Nuclear localization of Kaiso promotes the poorly differentiated phenotype and EMT in infiltrating ductal carcinomas. Clin Exp Metasis 31(5):497–510
20. Jones J, Wang H, Zhou J, Hardy S, Turner T, Austin D et al (2012) Nuclear kaiso indicates aggressive prostate cancers and promotes migration and invasiveness of prostate cancer cells. Am J Pathol 181(5):1836–1846
21. Daniel JM, Reynolds AB (1999) The catenin p120(ctn) interacts with Kaiso, a novel BTB/POZ domain zinc finger transcription factor. Mol Cell Biol 19(5):3614–3623
22. Kelly KF, Daniel JM (2006) POZ for effect—POZ-ZF transcription factors in cancer and development. Trends Cell Biol 16(11):578–587
23. Daniel JM, Spring CM, Crawford HC, Reynolds AB, Baig A (2002) The p120(ctn)-binding partner Kaiso is a bi-modal DNA-binding protein that recognizes both a sequence-specific consensus and methylated CpG dinucleotides. Nucleic Acids Res 30(13):2911–2919
24. Prokhortchouk A, Hendrich B, Jorgensen H, Ruzov A, Wilm M, Georgiev G et al (2001) The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor. Genes Dev 15(13):1613–1618
25. Donaldson NS, Pierre CC, Anstey MI, Robinson SC, Weerawardane SM, Daniel JM (2012) Kaiso represses the cell cycle gene cyclin D1 via sequence-specific and methyl-CpG-dependent mechanisms. PLoS ONE 7(11):e50398
26. Daniel JM (2007) Dancing in and out of the nucleus: p120(ctn) and the transcription factor Kaiso. Biochim Biophys Acta 1773(1):59–68
27. Rodova M, Kelly KF, VanSaun M, Daniel JM, Werle MJ (2004) Regulation of the rapsyn promoter by kaiso and delta-catenin. Mol Cell Biol 24(16):7188–7196
28. Blattler A, Yao L, Wang Y, Ye Z, Jin VX, Farnham PJ (2013) ZBTB33 binds unmethylated regions of the genome associated with actively expressed genes. Epigenetics Chromatin 6(1):13
29. Musgrove EA, Caldon CE, Barraclough J, Stone A, Sutherland RL (2011) Cyclin D as a therapeutic target in cancer. Nat Rev Cancer 11(8):558–572
30. Adachi Y, Yamamoto H, Itoh F, Hinoda Y, Okada Y, Imai K (1999) Contribution of matrilysin (MMP-7) to the metastatic pathway of human colorectal cancers. Gut 45(2):252–258
31. Onder TT, Gupta PB, Mani SA, Yang J, Lander ES, Weinberg RA (2008) Loss of E-cadherin promotes metastasis via multiple downstream transcriptional pathways. Can Res 68(10):3645–3654
32. Pierre CC, Longo J, Mavor M, Milosavljevic SB, Chaudhary R, Gilbreath E et al (2015) Kaiso overexpression promotes intestinal inflammation and potentiates intestinal tumorigenesis in Apc(-Min/+) mice. Biochim Biophys Acta 1852(9):1846–1855
33. Dai SD, Wang Y, Miao Y, Zhao Y, Zhang Y, Jiang GY et al (2009) Cytoplasmic Kaiso is associated with poor prognosis in non-small cell lung cancer. BMC Cancer 9:178
34. Jones J, Mukherjee A, Karanam B, Davis M, Jaynes J, Reams RR et al (2016) African Americans with pancreatic ductal adenocarcinoma exhibit gender differences in Kaiso expression. Cancer Lett 380(2):513–522
35. Bassey-Archibong BI, Kwiecien JM, Milosavljevic SB, Hallett RM, Rayner LG, Erb MJ et al (2016) Kaiso depletion attenuates transforming growth factor-β signaling and metastatic activity of triple-negative breast cancer cells. Oncogenesis 5:e208
36. Prokhortchouk A, Sansom O, Selfridge J, Caballero IM, Salozhin S, Aithozhina D et al (2006) Kaiso-deficient mice show resistance to intestinal cancer. Mol Cell Biol 26(1):199–208
37. Lopes EC, Valls E, Figueroa ME, Mazur A, Meng FG, Chiosis G et al (2008) Kaiso contributes to DNA methylation-dependent silencing of tumor suppressor genes in colon cancer cell lines. Can Res 68(18):7258–7263
38. Vermeulen JF, van de Ven RA, Ercan C, van der Groep P, van der Wall E, Bult P et al (2012) Nuclear Kaiso expression is associated with high grade and triple-negative invasive breast cancer. PLoS ONE 7(5):e37864
39. Daniel JM, Ireton RC, Reynolds AB (2001) Monoclonal antibodies to Kaiso: a novel transcription factor and p120ctn-binding protein. Hybridoma 20(3):159–166
40. Chaudhary R, Pierre CC, Nanan K, Wojtal D, Morone S, Pinelli C et al (2013) The POZ-ZF transcription factor Kaiso (ZBTB33) induces inflammation and progenitor cell differentiation in the murine intestine. PLoS ONE 8(9):e74160
41. Tracing African roots; exploring the ethnic origins of the Afro-Diaspora
42. Murray T, Beaty TH, Mathias RA, Rafaels N, Grant AV, Faruque MU et al (2010) African and non-African admixture components in African Americans and an African Caribbean population. Genet Epidemiol 34(6):561–568
43. Weigelt B, Peterse JL, van't Veer LJ (2005) Breast cancer metastasis: markers and models. Nat Rev Cancer 5(8):591–602
44. Huo D, Ikpatt F, Khramtsov A, Dangou JM, Nanda R, Dignam J et al (2009) Population differences in breast cancer: survey in indigenous African women reveals over-representation of triple-negative breast cancer. J Clin Oncol 27(27):4515–4521
45. Sawe RT, Kerper M, Badve S, Li J, Sandoval-Cooper M, Xie J et al (2016) Aggressive breast cancer in western Kenya has early onset, high proliferation, and immune cell infiltration. BMC Cancer 16:204
46. Adisa CA, Eleweke N, Alfred AA, Campbell MJ, Sharma R, Nseyo O et al (2012) Biology of breast cancer in Nigerian women: a pilot study. Ann Afr Med 11(3):169–175
47. Fraizer M (2005) Continuity and change in Caribbean immigration. People's World
48. Jackson FL (2008) Ancestral links of Chesapeake Bay region African Americans to specific Bight of Bonny (West Africa)

microethnic groups and increased frequency of aggressive breast cancer in both regions. Am J Hum Biol 20(2):165–173
49. Bryc K, Auton A, Nelson MR, Oksenberg JR, Hauser SL, Williams S et al (2010) Genome-wide patterns of population structure and admixture in West Africans and African Americans. Proc Natl Acad Sci USA 107(2):786–791
50. McCabe A, Dolled-Filhart M, Camp RL, Rimm DL (2005) Automated quantitative analysis (AQUA) of in situ protein expression, antibody concentration, and prognosis. J Natl Cancer Inst 97(24):1808–1815
51. Bassey-Archibong BI, Rayner LG, Hercules SM, Aarts CW, Dvorkin-Gheva A, Bramson JL et al (2017) Kaiso depletion attenuates the growth and survival of triple negative breast cancer cells. Cell Death Dis 8(3):e2689

REFERENCES

Abbad, A., Baba, H., Dehbi, H., Elmessaoudi-Idrissi, M., Elyazghi, Z., Abidi, O., & Radouani, F. (2018). Genetics of breast cancer in African populations: a literature review. *Glob Health Epidemiol Genom*, *3*, e8. https://doi.org/10.1017/gheg.2018.8

Abisoye-Ogunniyan, A., Lin, H., Ghebremedhin, A., Salam, A. B., Karanam, B., Theodore, S., . . . Yates, C. (2018). Transcriptional repressor Kaiso promotes epithelial to mesenchymal transition and metastasis in prostate cancer through direct regulation of miR-200c. *Cancer Lett*, *431*, 1-10. https://doi.org/10.1016/j.canlet.2018.04.044

Adedokun, B., Du, Z., Gao, G., Ahearn, T. U., Lunetta, K. L., Zirpoli, G., . . . Team, G. S. (2021). Cross-ancestry GWAS meta-analysis identifies six breast cancer loci in African and European ancestry women. *Nat Commun*, *12*(1), 4198. https://doi.org/10.1038/s41467-021-24327-x

Adedokun, B., Zheng, Y., Ndom, P., Gakwaya, A., Makumbi, T., Zhou, A. Y., . . . Huo, D. (2020). Prevalence of Inherited Mutations in Breast Cancer Predisposition Genes among Women in Uganda and Cameroon. *Cancer Epidemiol Biomarkers Prev*, *29*(2), 359-367. https://doi.org/10.1158/1055-9965.EPI-19-0506

Adorno, M., Cordenonsi, M., Montagner, M., Dupont, S., Wong, C., Hann, B., . . . Piccolo, S. (2009). A Mutant-p53/Smad complex opposes p63 to empower TGFbeta-induced metastasis. *Cell*, *137*(1), 87-98. https://doi.org/10.1016/j.cell.2009.01.039

Ahn, S., Woo, J. W., Lee, K., & Park, S. Y. (2020). HER2 status in breast cancer: changes in guidelines and complicating factors for interpretation. *J Pathol Transl Med*, *54*(1), 34-44. https://doi.org/10.4132/jptm.2019.11.03

Allison, K. H., Hammond, M. E. H., Dowsett, M., McKernin, S. E., Carey, L. A., Fitzgibbons, P. L., . . . Wolff, A. C. (2020). Estrogen and Progesterone Receptor Testing in Breast Cancer: ASCO/CAP Guideline Update. *J Clin Oncol*, *38*(12), 1346-1366. https://doi.org/10.1200/JCO.19.02309

Allred, D. C., Harvey, J. M., Berardo, M., & Clark, G. M. (1998). Prognostic and predictive factors in breast cancer by immunohistochemical analysis. *Mod Pathol*, *11*(2), 155-168.

Allred, D. C., Mohsin, S. K., & Fuqua, S. A. (2001). Histological and biological evolution of human premalignant breast disease. *Endocr Relat Cancer*, *8*(1), 47-61. https://doi.org/10.1677/erc.0.0080047

Alluri, P., & Newman, L. A. (2014). Basal-like and triple-negative breast cancers: searching for positives among many negatives. *Surg Oncol Clin N Am*, *23*(3), 567-577. https://doi.org/10.1016/j.soc.2014.03.003

Amin, M. B., Edge, S. B., Greene, F. L., & American Joint Committee on, C. (2017). *AJCC cancer staging manual* (Eighth edition ed.). Springer.

Anders, C. K., Hsu, D. S., Broadwater, G., Acharya, C. R., Foekens, J. A., Zhang, Y., . . . Blackwell, K. L. (2008). 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*(20), 3324-3330. https://doi.org/10.1200/JCO.2007.14.2471

Auton, A., Brooks, L. D., Durbin, R. M., Garrison, E. P., Kang, H. M., Korbel, J. O., . . . Consortium, G. P. (2015). A global reference for human genetic variation. *Nature*, *526*(7571), 68-74. https://doi.org/10.1038/nature15393

Azubuike, S. (2017). Breast cancer risk factors and signs: How much do Nigerian women know? [Rapid Communication]. *International Journal of Advanced Medical and Health Research*, *4*(1), 40-43. https://doi.org/10.4103/ijamr.ijamr_45_16

Bajaj, S. S., & Stanford, F. C. (2021). Beyond Tuskegee — Vaccine Distrust and Everyday Racism. *New England Journal of Medicine*, *384*(5), e12. https://doi.org/10.1056/NEJMpv2035827

Bajrami, I., Marlow, R., van de Ven, M., Brough, R., Pemberton, H. N., Frankum, J., . . . Lord, C. J. (2018). E-Cadherin/ROS1 Inhibitor Synthetic Lethality in Breast Cancer. *Cancer Discov*, *8*(4), 498-515. https://doi.org/10.1158/2159-8290.CD-17-0603

Bandyopadhyay, R., Atehnkeng, J., Ortega-Beltran, A., Akande, A., Falade, T. D. O., & Cotty, P. J. (2019). "Ground-Truthing" Efficacy of Biological Control for Aflatoxin Mitigation in Farmers' Fields in Nigeria: From Field Trials to Commercial Usage, a 10-Year Study. *Front Microbiol*, *10*, 2528. https://doi.org/10.3389/fmicb.2019.02528

Bankhead, P., Loughrey, M. B., Fernández, J. A., Dombrowski, Y., McArt, D. G., Dunne, P. D., . . . Hamilton, P. W. (2017). QuPath: Open source software for digital pathology image analysis. *Sci Rep*, *7*(1), 16878. https://doi.org/10.1038/s41598-017-17204-5

Baselga, J., Norton, L., Albanell, J., Kim, Y. M., & Mendelsohn, J. (1998). Recombinant humanized anti-HER2 antibody (Herceptin) enhances the antitumor activity of paclitaxel and doxorubicin against HER2/neu overexpressing human breast cancer xenografts. *Cancer Res*, *58*(13), 2825-2831.

Bassey-Archibong, B. I., Hercules, S. M., Rayner, L. G. A., Skeete, D. H. A., Smith Connell, S. P., Brain, I., . . . Daniel, J. M. (2017). Kaiso is highly expressed in TNBC tissues of women of African ancestry compared to Caucasian women. *Cancer Causes Control*. https://doi.org/10.1007/s10552-017-0955-2

Bassey-Archibong, B. I., Kwiecien, J. M., Milosavljevic, S. B., Hallett, R. M., Rayner, L. G., Erb, M. J., . . . Daniel, J. M. (2016). Kaiso depletion attenuates transforming growth factor-beta signaling and metastatic activity of triple-negative breast cancer cells. *Oncogenesis*, *5*, e208. https://doi.org/10.1038/oncsis.2016.17

Bassey-Archibong, B. I., Rayner, L. G., Hercules, S. M., Aarts, C. W., Dvorkin-Gheva, A., Bramson, J. L., . . . Daniel, J. M. (2017). Kaiso depletion attenuates the growth and survival of triple negative breast cancer cells. *Cell Death Dis*, *8*(3), e2689. https://doi.org/10.1038/cddis.2017.92

Bayraktar, S., Zhou, J. Z., Bassett, R., Gutierrez Barrera, A. M., Layman, R. M., Valero, V., & Arun, B. (2020). Clinical outcome and toxicity from taxanes in breast cancer patients with BRCA1 and BRCA2 pathogenic germline mutations. *Breast J*, *26*(8), 1572-1582. https://doi.org/10.1111/tbj.13922

Beniey, M., Haque, T., & Hassan, S. (2019). Translating the role of PARP inhibitors in triple-negative breast cancer. *Oncoscience*, *6*(1-2), 287-288. https://doi.org/10.18632/oncoscience.474

Benn-Torres, J., Bonilla, C., Robbins, C. M., Waterman, L., Moses, T. Y., Hernandez, W., . . . Kittles, R. A. (2008). Admixture and population stratification in African Caribbean populations. *Ann Hum Genet*, *72*(Pt 1), 90-98. https://doi.org/10.1111/j.1469-1809.2007.00398.x

Bignon, L., Fricker, J. P., Nogues, C., Mouret-Fourme, E., Stoppa-Lyonnet, D., Caron, O., . . . Velten, M. (2018). Efficacy of anthracycline/taxane-based neo-adjuvant chemotherapy on triple-

negative breast cancer in BRCA1/BRCA2 mutation carriers. *Breast J*, *24*(3), 269-277. https://doi.org/10.1111/tbj.12887

Biswas, T., Jindal, C., Fitzgerald, T. L., & Efird, J. T. (2019). Pathologic Complete Response (pCR) and Survival of Women with Inflammatory Breast Cancer (IBC): An Analysis Based on Biologic Subtypes and Demographic Characteristics. *Int J Environ Res Public Health*, *16*(1). https://doi.org/10.3390/ijerph16010124

Black, E., & Richmond, R. (2019). Improving early detection of breast cancer in sub-Saharan Africa: why mammography may not be the way forward. *Global Health*, *15*(1), 3. https://doi.org/10.1186/s12992-018-0446-6

Blattler, A., Yao, L., Wang, Y., Ye, Z., Jin, V. X., & Farnham, P. J. (2013). ZBTB33 binds unmethylated regions of the genome associated with actively expressed genes. *Epigenetics Chromatin*, *6*(1), 13. https://doi.org/1756-8935-6-13 [pii] 10.1186/1756-8935-6-13

Borges Migliavaca, C., Stein, C., Colpani, V., Barker, T. H., Munn, Z., Falavigna, M., & (PERSyst), P. E. R. S. R. M. G. (2020). How are systematic reviews of prevalence conducted? A methodological study. *BMC Med Res Methodol*, *20*(1), 96. https://doi.org/10.1186/s12874-020-00975-3

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., & Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*, *68*(6), 394-424. https://doi.org/10.3322/caac.21492

Breasted, J. H. (1930). The Edwin Smith Surgical Papyrus: published in facsimile and hieroglyphic transliteration with translation and commentary in two volumes.

Brewster, A. M., Chavez-MacGregor, M., & Brown, P. (2014). Epidemiology, biology, and treatment of triple-negative breast cancer in women of African ancestry. *Lancet Oncol*, *15*(13), e625-e634. https://doi.org/10.1016/S1470-2045(14)70364-X

Burstein, M. D., Tsimelzon, A., Poage, G. M., Covington, K. R., Contreras, A., Fuqua, S. A., . . . Brown, P. H. (2015). Comprehensive genomic analysis identifies novel subtypes and targets of triple-negative breast cancer. *Clin Cancer Res*, *21*(7), 1688-1698. https://doi.org/10.1158/1078-0432.CCR-14-0432

Cameron, D., Piccart-Gebhart, M. J., Gelber, R. D., Procter, M., Goldhirsch, A., de Azambuja, E., . . . Team, H. A. H. T. S. (2017). 11 years' follow-up of trastuzumab after adjuvant chemotherapy in HER2-positive early breast cancer: final analysis of the HERceptin Adjuvant (HERA) trial. *Lancet*, *389*(10075), 1195-1205. https://doi.org/10.1016/S0140-6736(16)32616-2

Carrasco-Ramiro, F., Peiró-Pastor, R., & Aguado, B. (2017). Human genomics projects and precision medicine. *Gene Ther*, *24*(9), 551-561. https://doi.org/10.1038/gt.2017.77

Chang, D. Y., Ma, W. L., & Lu, Y. S. (2021). Role of Alpelisib in the Treatment of PIK3CA-Mutated Breast Cancer: Patient Selection and Clinical Perspectives. *Ther Clin Risk Manag*, *17*, 193-207. https://doi.org/10.2147/TCRM.S251668

Changavi, A. A., Shashikala, A., & Ramji, A. S. (2015). Epidermal Growth Factor Receptor Expression in Triple Negative and Nontriple Negative Breast Carcinomas. *J Lab Physicians*, *7*(2), 79-83. https://doi.org/10.4103/0974-2727.163129

Chawanthayatham, S., Valentine, C. C., Fedeles, B. I., Fox, E. J., Loeb, L. A., Levine, S. S., . . . Essigmann, J. M. (2017). Mutational spectra of aflatoxin B. *Proc Natl Acad Sci U S A*, *114*(15), E3101-E3109. https://doi.org/10.1073/pnas.1700759114

Chen, H., Pan, H., Qian, Y., Zhou, W., & Liu, X. (2018). MiR-25-3p promotes the proliferation of triple negative breast cancer by targeting BTG2. *Mol Cancer*, *17*(1), 4. https://doi.org/10.1186/s12943-017-0754-0

Chen, H., Wu, J., Zhang, Z., Tang, Y., Li, X., Liu, S., & Cao, S. (2018). Association Between. *Front Pharmacol*, *9*, 909. https://doi.org/10.3389/fphar.2018.00909

Churpek, J. E., Walsh, T., Zheng, Y., Moton, Z., Thornton, A. M., Lee, M. K., . . . Olopade, O. I. (2015). Inherited predisposition to breast cancer among African American women. *Breast Cancer Res Treat*, *149*(1), 31-39. https://doi.org/10.1007/s10549-014-3195-0

Cibulskis, K., Lawrence, M. S., Carter, S. L., Sivachenko, A., Jaffe, D., Sougnez, C., . . . Getz, G. (2013). Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol*, *31*(3), 213-219. https://doi.org/10.1038/nbt.2514

Clark, O., Botrel, T. E., Paladini, L., & Ferreira, M. B. (2014). Targeted therapy in triple-negative metastatic breast cancer: a systematic review and meta-analysis. *Core Evid*, *9*, 1-11. https://doi.org/10.2147/ce.s52197

CO., U., C, E., DCD, A., IV, O., ME, O., OA, E., . . . SNC, A. (2017). The Immunohistochemical Profile of Breast Cancer in Indigenous Women of Southeast Nigeria *Annals of Medical and Health Sciences Research*, *7*(2), 83-87.

Cocco, S., Piezzo, M., Calabrese, A., Cianniello, D., Caputo, R., Lauro, V. D., . . . De Laurentiis, M. (2020). Biomarkers in Triple-Negative Breast Cancer: State-of-the-Art and Future Perspectives. *Int J Mol Sci*, *21*(13). https://doi.org/10.3390/ijms21134579

Corbex, M., Burton, R., & Sancho-Garnier, H. (2012). Breast cancer early detection methods for low and middle income countries, a review of the evidence. *Breast*, *21*(4), 428-434. https://doi.org/10.1016/j.breast.2012.01.002

Costa, R., Shah, A. N., Santa-Maria, C. A., Cruz, M. R., Mahalingam, D., Carneiro, B. A., . . . Giles, F. J. (2017). Targeting Epidermal Growth Factor Receptor in triple negative breast cancer: New discoveries and practical insights for drug development. *Cancer Treat Rev*, *53*, 111-119. https://doi.org/10.1016/j.ctrv.2016.12.010

Crawford, B., Adams, S. B., Sittler, T., van den Akker, J., Chan, S., Leitner, O., . . . van 't Veer, L. (2017). Multi-gene panel testing for hereditary cancer predisposition in unsolved high-risk breast and ovarian cancer patients. *Breast Cancer Res Treat*, *163*(2), 383-390. https://doi.org/10.1007/s10549-017-4181-0

Cubasch, H., Dickens, C., Joffe, M., Duarte, R., Murugan, N., Tsai Chih, M., . . . Ruff, P. (2018). Breast cancer survival in Soweto, Johannesburg, South Africa: A receptor-defined cohort of women diagnosed from 2009 to 11. *Cancer Epidemiol*, *52*, 120-127. https://doi.org/10.1016/j.canep.2017.12.007

Cunningham, D. L., Sarhan, A. R., Creese, A. J., Larkins, K. P. B., Zhao, H., Ferguson, H. R., . . . Heath, J. K. (2020). Differential responses to kinase inhibition in FGFR2-addicted triple negative breast cancer cells: a quantitative phosphoproteomics study. *Sci Rep*, *10*(1), 7950. https://doi.org/10.1038/s41598-020-64534-y

Dai, S. D., Wang, Y., Miao, Y., Zhao, Y., Zhang, Y., Jiang, G. Y., . . . Wang, E. H. (2009). Cytoplasmic Kaiso is associated with poor prognosis in non-small cell lung cancer. *BMC Cancer*, *9*, 178. https://doi.org/1471-2407-9-178 [pii]

10.1186/1471-2407-9-178

Daniel, J. M., & Reynolds, A. B. (1999). The catenin p120(ctn) interacts with Kaiso, a novel BTB/POZ domain zinc finger transcription factor. *Mol Cell Biol*, *19*(5), 3614-3623.

Daniel, J. M., Spring, C. M., Crawford, H. C., Reynolds, A. B., & Baig, A. (2002). The p120(ctn)-binding partner Kaiso is a bi-modal DNA-binding protein that recognizes both a sequence-specific consensus and methylated CpG dinucleotides. *Nucleic Acids Res*, *30*(13), 2911-2919.

Davis, M., Martini, R., Newman, L., Elemento, O., White, J., Verma, A., . . . Yates, C. (2020). Identification of Distinct Heterogenic Subtypes and Molecular Signatures Associated with African Ancestry in Triple Negative Breast Cancer Using Quantified Genetic Ancestry Models in Admixed Race Populations. *Cancers*, *12*(5), Article 1220. https://doi.org/10.3390/cancers12051220

Davis, M., Martini, R., Newman, L., Elemento, O., White, J., Verma, A., . . . Yates, C. (2020). Identification of Distinct Heterogenic Subtypes and Molecular Signatures Associated with African Ancestry in Triple Negative Breast Cancer Using Quantified Genetic Ancestry Models in Admixed Race Populations. *Cancers (Basel)*, *12*(5). https://doi.org/10.3390/cancers12051220

Davuluri, G., Schiemann, W. P., Plow, E. F., & Sossey-Alaoui, K. (2014). Loss of WAVE3 sensitizes triple-negative breast cancers to chemotherapeutics by inhibiting the STAT-HIF-1α-mediated angiogenesis. *JAKSTAT*, *3*(4), e1009276. https://doi.org/10.1080/21623996.2015.1009276

DeSantis, C. E., Fedewa, S. A., Goding Sauer, A., Kramer, J. L., Smith, R. A., & Jemal, A. (2016). Breast cancer statistics, 2015: Convergence of incidence rates between black and white women. *CA Cancer J Clin*, *66*(1), 31-42. https://doi.org/10.3322/caac.21320

DeSantis, C. E., Ma, J., Gaudet, M. M., Newman, L. A., Miller, K. D., Goding Sauer, A., . . . Siegel, R. L. (2019). Breast cancer statistics, 2019. *CA Cancer J Clin*, *69*(6), 438-451. https://doi.org/10.3322/caac.21583

DeSantis, C. E., Ma, J., Goding Sauer, A., Newman, L. A., & Jemal, A. (2017). Breast cancer statistics, 2017, racial disparity in mortality by state. *CA Cancer J Clin*, *67*(6), 439-448. https://doi.org/10.3322/caac.21412

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

Donenberg, T., Ahmed, H., Royer, R., Zhang, S., Narod, S. A., George, S., . . . Hurley, J. (2016). A Survey of BRCA1, BRCA2, and PALB2 mutations in women with breast cancer in Trinidad and Tobago. *Breast Cancer Res Treat*, *159*(1), 131-138. https://doi.org/10.1007/s10549-016-3870-4

Donenberg, T., Lunn, J., Curling, D., Turnquest, T., Krill-Jackson, E., Royer, R., . . . Hurley, J. (2011). A high prevalence of BRCA1 mutations among breast cancer patients from the Bahamas. *Breast Cancer Res Treat*, *125*(2), 591-596. https://doi.org/10.1007/s10549-010-1156-9

Dreyer, M. S., Nattinger, A. B., McGinley, E. L., & Pezzin, L. E. (2018). Socioeconomic status and breast cancer treatment. *Breast Cancer Res Treat*, *167*(1), 1-8. https://doi.org/10.1007/s10549-017-4490-3

Dupont, W. D., & Page, D. L. (1989). Relative risk of breast cancer varies with time since diagnosis of atypical hyperplasia. *Hum Pathol*, *20*(8), 723-725. https://doi.org/10.1016/0046-8177(89)90063-4

Díaz-Gay, M., Vila-Casadesús, M., Franch-Expósito, S., Hernández-Illán, E., Lozano, J. J., & Castellví-Bel, S. (2018). Mutational Signatures in Cancer (MuSiCa): a web application to implement

mutational signatures analysis in cancer samples. *BMC Bioinformatics*, *19*(1), 224. https://doi.org/10.1186/s12859-018-2234-y

Elsamanoudy, A. Z., Mohamed Neamat-Allah, M. A., Hisham Mohammad, F. A., Hassanien, M., & Nada, H. A. (2016). The role of nutrition related genes and nutrigenetics in understanding the pathogenesis of cancer. *J Microsc Ultrastruct*, *4*(3), 115-122. https://doi.org/10.1016/j.jmau.2016.02.002

Elston, C. W., & Ellis, I. O. (1991). Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology*, *19*(5), 403-410. https://doi.org/10.1111/j.1365-2559.1991.tb00229.x

Ferlay, J., Colombet, M., Soerjomataram, I., Mathers, C., Parkin, D. M., Piñeros, M., . . . Bray, F. (2019). Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *Int J Cancer*, *144*(8), 1941-1953. https://doi.org/10.1002/ijc.31937

Fidler, M. M., Bray, F., & Soerjomataram, I. (2018). The global cancer burden and human development: A review. *Scand J Public Health*, *46*(1), 27-36. https://doi.org/10.1177/1403494817715400

Foerster, M., Anderson, B. O., McKenzie, F., Galukande, M., Anele, A., Adisa, C., . . . McCormack, V. (2019). Inequities in breast cancer treatment in sub-Saharan Africa: findings from a prospective multi-country observational study. *Breast Cancer Res*, *21*(1), 93. https://doi.org/10.1186/s13058-019-1174-4

Forbes, S. A., Beare, D., Bindal, N., Bamford, S., Ward, S., Cole, C. G., . . . Campbell, P. J. (2016). COSMIC: High-Resolution Cancer Genetics Using the Catalogue of Somatic Mutations in Cancer. *Curr Protoc Hum Genet*, *91*, 10.11.11-10.11.37. https://doi.org/10.1002/cphg.21

Fougner, C., Bergholtz, H., Norum, J. H., & Sørlie, T. (2020). Re-definition of claudin-low as a breast cancer phenotype. *Nat Commun*, *11*(1), 1787. https://doi.org/10.1038/s41467-020-15574-5

Friebel, T. M., Andrulis, I. L., Balmaña, J., Blanco, A. M., Couch, F. J., Daly, M. B., . . . Rebbeck, T. R. (2019). BRCA1 and BRCA2 pathogenic sequence variants in women of African origin or ancestry. *Hum Mutat*. https://doi.org/10.1002/humu.23804

Gao, J. J., & Swain, S. M. (2018). Luminal A Breast Cancer and Molecular Assays: A Review. *Oncologist*, *23*(5), 556-565. https://doi.org/10.1634/theoncologist.2017-0535

Garrido-Castro, A. C., Lin, N. U., & Polyak, K. (2019). Insights into Molecular Classifications of Triple-Negative Breast Cancer: Improving Patient Selection for Treatment. *Cancer Discov*, *9*(2), 176-198. https://doi.org/10.1158/2159-8290.CD-18-1177

Geenen, J. J. J., Linn, S. C., Beijnen, J. H., & Schellens, J. H. M. (2018). PARP Inhibitors in the Treatment of Triple-Negative Breast Cancer. *Clin Pharmacokinet*, *57*(4), 427-437. https://doi.org/10.1007/s40262-017-0587-4

George, S. H. L., Donenberg, T., Alexis, C., DeGennaro, V., Dyer, H., Yin, S., . . . Hurley, J. E. (2021). Gene Sequencing for Pathogenic Variants Among Adults With Breast and Ovarian Cancer in the Caribbean. *JAMA Netw Open*, *4*(3), e210307. https://doi.org/10.1001/jamanetworkopen.2021.0307

Geronimus, A. T., Hicken, M., Keene, D., & Bound, J. (2006). "Weathering" and age patterns of allostatic load scores among blacks and whites in the United States. *Am J Public Health*, *96*(5), 826-833. https://doi.org/10.2105/AJPH.2004.060749

Goldhirsch, A., Winer, E. P., Coates, A. S., Gelber, R. D., Piccart-Gebhart, M., Thürlimann, B., . . . members, P. (2013). Personalizing the treatment of women with early breast cancer:

highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. *Ann Oncol*, *24*(9), 2206-2223. https://doi.org/10.1093/annonc/mdt303

Goldhirsch, A., Wood, W. C., Coates, A. S., Gelber, R. D., Thurlimann, B., Senn, H. J., & Panel, m. (2011). Strategies for subtypes--dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. *Ann Oncol*, *22*(8), 1736-1747. https://doi.org/10.1093/annonc/mdr304

Goldhirsch, A., Wood, W. C., Coates, A. S., Gelber, R. D., Thürlimann, B., Senn, H. J., & members, P. (2011). Strategies for subtypes--dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. *Ann Oncol*, *22*(8), 1736-1747. https://doi.org/10.1093/annonc/mdr304

Gordon, N. H. (2003). Socioeconomic factors and breast cancer in black and white Americans. *Cancer Metastasis Rev*, *22*(1), 55-65.

Gyan, K., Kyei, I., Susick, L. L., Chen, Y., Davis, M. B., Salem, B., & Newman, L. A. (2019). African Clinicians' Prioritization of Needs in International Breast Cancer Partnerships. *JAMA Surg*, *154*(2), 182-184. https://doi.org/10.1001/jamasurg.2018.4208

Gómez-Flores-Ramos, L., Álvarez-Gómez, R. M., Villarreal-Garza, C., Wegman-Ostrosky, T., & Mohar, A. (2017). Breast cancer genetics in young women: What do we know? *Mutat Res*, *774*, 33-45. https://doi.org/10.1016/j.mrrev.2017.08.001

Hajdu, S. I. (2011). A note from history: landmarks in history of cancer, part 1. *Cancer*, *117*(5), 1097-1102. https://doi.org/10.1002/cncr.25553

Halmin, M., Bellocco, R., Lagerlund, M., Karlsson, P., Tejler, G., & Lambe, M. (2008). Long-term inequalities in breast cancer survival--a ten year follow-up study of patients managed within a National Health Care System (Sweden). *Acta Oncol*, *47*(2), 216-224. https://doi.org/10.1080/02841860701769768

Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *Cell*, *144*(5), 646-674. https://doi.org/10.1016/j.cell.2011.02.013

Haradhvala, N. J., Kim, J., Maruvka, Y. E., Polak, P., Rosebrock, D., Livitz, D., . . . Getz, G. (2018). Distinct mutational signatures characterize concurrent loss of polymerase proofreading and mismatch repair. *Nat Commun*, *9*(1), 1746. https://doi.org/10.1038/s41467-018-04002-4

Harrison, J. C., Carvajal, M., & Garner, R. C. (1993). Does aflatoxin exposure in the United Kingdom constitute a cancer risk? *Environ Health Perspect*, *99*, 99-105. https://doi.org/10.1289/ehp.939999

Hashmi, A. A., Naz, S., Hashmi, S. K., Irfan, M., Hussain, Z. F., Khan, E. Y., . . . Faridi, N. (2019). Epidermal growth factor receptor (EGFR) overexpression in triple-negative breast cancer: association with clinicopathologic features and prognostic parameters. *Surgical and Experimental Pathology*, *2*(1), 1-7.

Hou, Y., Nitta, H., Parwani, A. V., & Li, Z. (2020). The assessment of HER2 status and its clinical implication in breast cancer. *Mini-Symposium: Breast Pathology*.

Howlader, N., Altekruse, S. F., Li, C. I., Chen, V. W., Clarke, C. A., Ries, L. A., & Cronin, K. A. (2014). US incidence of breast cancer subtypes defined by joint hormone receptor and HER2 status. *J Natl Cancer Inst*, *106*(5). https://doi.org/10.1093/jnci/dju055

Hoy, D., Brooks, P., Woolf, A., Blyth, F., March, L., Bain, C., . . . Buchbinder, R. (2012). Assessing risk of bias in prevalence studies: modification of an existing tool and evidence of interrater

agreement. *J Clin Epidemiol*, *65*(9), 934-939. https://doi.org/10.1016/j.jclinepi.2011.11.014

Huang, M. N., Yu, W., Teoh, W. W., Ardin, M., Jusakul, A., Ng, A. W. T., . . . Rozen, S. G. (2017). Genome-scale mutational signatures of aflatoxin in cells, mice, and human tumors. *Genome Res*, *27*(9), 1475-1486. https://doi.org/10.1101/gr.220038.116

Huo, D., Hu, H., Rhie, S. K., Gamazon, E. R., Cherniack, A. D., Liu, J., . . . Olopade, O. I. (2017). Comparison of Breast Cancer Molecular Features and Survival by African and European Ancestry in The Cancer Genome Atlas. *JAMA Oncol*, *3*(12), 1654-1662. https://doi.org/10.1001/jamaoncol.2017.0595

Hurvitz, S., & Mead, M. (2016). Triple-negative breast cancer: advancements in characterization and treatment approach. *Curr Opin Obstet Gynecol*, *28*(1), 59-69. https://doi.org/10.1097/gco.0000000000000239

IHME. (2020). *GBD Compare Data Visualization*. Retrieved June 9th from http://vizhub.healthdata.org/gbd-compare

Iqbal, J., Ginsburg, O., Rochon, P. A., Sun, P., & Narod, S. A. (2015). Differences in breast cancer stage at diagnosis and cancer-specific survival by race and ethnicity in the United States. *JAMA*, *313*(2), 165-173. https://doi.org/10.1001/jama.2014.17322

Jafarian, A. H., Kooshkiforooshani, M., Farzad, F., & Mohamadian Roshan, N. (2019). The Relationship Between Fibroblastic Growth Factor Receptor-1 (FGFR1) Gene Amplification in Triple Negative Breast Carcinomas and Clinicopathological Prognostic Factors. *Iran J Pathol*, *14*(4), 299-304. https://doi.org/10.30699/ijp.2019.96713.1952

Jemal, A., & Brawley, O. W. (2019). Increasing cancer awareness and prevention in Africa. *Ecancermedicalscience*, *13*, 939. https://doi.org/10.3332/ecancer.2019.939

Jones, J., Wang, H., Karanam, B., Theodore, S., Dean-Colomb, W., Welch, D. R., . . . Yates, C. (2014). Nuclear localization of Kaiso promotes the poorly differentiated phenotype and EMT in infiltrating ductal carcinomas. *Clin Exp Metastasis*, *31*(5), 497-510. https://doi.org/10.1007/s10585-014-9644-7

Jones, J., Wang, H., Zhou, J., Hardy, S., Turner, T., Austin, D., . . . Yates, C. (2012). Nuclear Kaiso indicates aggressive prostate cancers and promotes migration and invasiveness of prostate cancer cells. *Am J Pathol*, *181*(5), 1836-1846. https://doi.org/S0002-9440(12)00605-0 [pii] 10.1016/j.ajpath.2012.08.008

Kagihara, J. A., Andress, M., & Diamond, J. R. (2020). Nab-paclitaxel and atezolizumab for the treatment of PD-L1-positive, metastatic triple-negative breast cancer: review and future directions. *Expert Rev Precis Med Drug Dev*, *5*(2), 59-65. https://doi.org/10.1080/23808993.2020.1730694

Kandoth, C., Gao, J., qwangmsk, Mattioni, M., Struck, A., Boursin, Y., . . . Chavan, S. mskcc/vcf2maf: vcf2maf v1.6.16.

Kang, C., & Syed, Y. Y. (2020). Atezolizumab (in Combination with Nab-Paclitaxel): A Review in Advanced Triple-Negative Breast Cancer. *Drugs*, *80*(6), 601-607. https://doi.org/10.1007/s40265-020-01295-y

Kassam, F., Enright, K., Dent, R., Dranitsaris, G., Myers, J., Flynn, C., . . . Clemons, M. (2009). Survival outcomes for patients with metastatic triple-negative breast cancer: implications for clinical practice and trial design. *Clin Breast Cancer*, *9*(1), 29-33. https://doi.org/10.3816/CBC.2009.n.005

Katayama, H., Tsou, P., Kobayashi, M., Capello, M., Wang, H., Esteva, F., . . . Hanash, S. (2019). A plasma protein derived TGFβ signature is a prognostic indicator in triple negative breast cancer. *NPJ Precis Oncol*, *3*, 10. https://doi.org/10.1038/s41698-019-0082-5

Kelly, K. F., & Daniel, J. M. (2006). POZ for effect--POZ-ZF transcription factors in cancer and development. *Trends Cell Biol*, *16*(11), 578-587. https://doi.org/S0962-8924(06)00239-X [pii] 10.1016/j.tcb.2006.09.003

Kim, S., Dubrovska, A., Salamone, R. J., Walker, J. R., Grandinetti, K. B., Bonamy, G. M., . . . Reddy, V. A. (2013). FGFR2 promotes breast tumorigenicity through maintenance of breast tumor-initiating cells. *PLoS One*, *8*(1), e51671. https://doi.org/10.1371/journal.pone.0051671

Koh, D. I., An, H., Kim, M. Y., Jeon, B. N., Choi, S. H., Hur, S. S., & Hur, M. W. (2015). Transcriptional activation of APAF1 by KAISO (ZBTB33) and p53 is attenuated by RelA/p65. *Biochim Biophys Acta*, *1849*(9), 1170-1178. https://doi.org/10.1016/j.bbagrm.2015.07.008

Krauss, K., & Stickeler, E. (2020). Endocrine Therapy in Early Breast Cancer. *Breast Care (Basel)*, *15*(4), 337-346. https://doi.org/10.1159/000509362

Krieger, N., Jahn, J. L., & Waterman, P. D. (2017). Jim Crow and estrogen-receptor-negative breast cancer: US-born black and white non-Hispanic women, 1992-2012. *Cancer Causes Control*, *28*(1), 49-59. https://doi.org/10.1007/s10552-016-0834-2

Krithikadatta, J. (2014). Normal distribution. *J Conserv Dent*, *17*(1), 96-97. https://doi.org/10.4103/0972-0707.124171

Kulkarni, S., Augoff, K., Rivera, L., McCue, B., Khoury, T., Groman, A., . . . Sossey-Alaoui, K. (2012). Increased expression levels of WAVE3 are associated with the progression and metastasis of triple negative breast cancer. *PLoS One*, *7*(8), e42895. https://doi.org/10.1371/journal.pone.0042895

Kwiecien, J. M., Bassey-Archibong, B. I., Dabrowski, W., Rayner, L. G., Lucas, A. R., & Daniel, J. M. (2017). Loss of Kaiso expression in breast cancer cells prevents intra-vascular invasion in the lung and secondary metastasis. *PLoS One*, *12*(9), e0183883. https://doi.org/10.1371/journal.pone.0183883

Ladep, N. G., Lesi, O. A., Mark, P., Lemoine, M., Onyekwere, C., Afihene, M., . . . Taylor-Robinson, S. D. (2014). Problem of hepatocellular carcinoma in West Africa. *World J Hepatol*, *6*(11), 783-792. https://doi.org/10.4254/wjh.v6.i11.783

Lehmann, B. D., Bauer, J. A., Chen, X., Sanders, M. E., Chakravarthy, A. B., Shyr, Y., & Pietenpol, J. A. (2011). Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J Clin Invest*, *121*(7), 2750-2767. https://doi.org/10.1172/jci45014

Lehmann, B. D., Jovanović, B., Chen, X., Estrada, M. V., Johnson, K. N., Shyr, Y., . . . Pietenpol, J. A. (2016). Refinement of Triple-Negative Breast Cancer Molecular Subtypes: Implications for Neoadjuvant Chemotherapy Selection. *PLoS One*, *11*(6), e0157368. https://doi.org/10.1371/journal.pone.0157368

Lerner-Ellis, J., Donenberg, T., Ahmed, H., George, S., Wharfe, G., Chin, S., . . . Akbari, M. R. (2017). A high frequency of PALB2 mutations in Jamaican patients with breast cancer. *Breast Cancer Res Treat*, *162*(3), 591-596. https://doi.org/10.1007/s10549-017-4148-1

Liang, Q., Ma, D., Gao, R. F., & Yu, K. D. (2020). Effect of Ki-67 Expression Levels and Histological Grade on Breast Cancer Early Relapse in Patients with Different Immunohistochemical-based Subtypes. *Sci Rep*, *10*(1), 7648. https://doi.org/10.1038/s41598-020-64523-1

Liedtke, C., Mazouni, C., Hess, K. R., André, F., Tordai, A., Mejia, J. A., . . . Pusztai, L. (2008). Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. *J Clin Oncol*, *26*(8), 1275-1281. https://doi.org/10.1200/JCO.2007.14.4147

Linnenbringer, E., Gehlert, S., & Geronimus, A. T. (2017). Black-White Disparities in Breast Cancer Subtype: The Intersection of Socially Patterned Stress and Genetic Expression. *AIMS Public Health*, *4*(5), 526-556. https://doi.org/10.3934/publichealth.2017.5.526

Linnenbringer, E., Geronimus, A. T., Davis, K. L., Bound, J., Ellis, L., & Gomez, S. L. (2020). Associations between breast cancer subtype and neighborhood socioeconomic and racial composition among Black and White women. *Breast Cancer Res Treat*, *180*(2), 437-447. https://doi.org/10.1007/s10549-020-05545-1

Liu, H., Murphy, C. J., Karreth, F. A., Emdal, K. B., White, F. M., Elemento, O., . . . Cantley, L. C. (2018). Identifying and Targeting Sporadic Oncogenic Genetic Aberrations in Mouse Models of Triple-Negative Breast Cancer. *Cancer Discov*, *8*(3), 354-369. https://doi.org/10.1158/2159-8290.CD-17-0679

Lopes, E. C., Valls, E., Figueroa, M. E., Mazur, A., Meng, F. G., Chiosis, G., . . . Melnick, A. (2008). Kaiso contributes to DNA methylation-dependent silencing of tumor suppressor genes in colon cancer cell lines. *Cancer Res*, *68*(18), 7258-7263. https://doi.org/68/18/7258 [pii]

10.1158/0008-5472.CAN-08-0344

Lord, C. J., & Ashworth, A. (2017). PARP inhibitors: Synthetic lethality in the clinic. *Science*, *355*(6330), 1152-1158. https://doi.org/10.1126/science.aam7344

Lundqvist, A., Andersson, E., Ahlberg, I., Nilbert, M., & Gerdtham, U. (2016). Socioeconomic inequalities in breast cancer incidence and mortality in Europe-a systematic review and meta-analysis. *Eur J Public Health*, *26*(5), 804-813. https://doi.org/10.1093/eurpub/ckw070

Lusito, E., Felice, B., D'Ario, G., Ogier, A., Montani, F., Di Fiore, P. P., & Bianchi, F. (2019). Unraveling the role of low-frequency mutated genes in breast cancer. *Bioinformatics*, *35*(1), 36-46. https://doi.org/10.1093/bioinformatics/bty520

Mannix, D. P. (1962). *Black cargoes; a history of the Atlantic slave trade, 1518-1865*. Viking Press.

Marchese, S., Polo, A., Ariano, A., Velotto, S., Costantini, S., & Severino, L. (2018). Aflatoxin B1 and M1: Biological Properties and Their Involvement in Cancer Development. *Toxins (Basel)*, *10*(6). https://doi.org/10.3390/toxins10060214

Marcus, P. M., & National Cancer Institute (U.S.). (2019). *Assessment of cancer screening : a primer*. National Cancer Institute (US),.

Martini, R., Chen, Y., Jenkins, B. D., Elhussin, I. A., Cheng, E., Hoda, S. A., . . . Davis, M. B. (2021). Investigation of triple-negative breast cancer risk alleles in an International African-enriched cohort. *Sci Rep*, *11*(1), 9247. https://doi.org/10.1038/s41598-021-88613-w

Mayakonda, A., Lin, D. C., Assenov, Y., Plass, C., & Koeffler, H. P. (2018). Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res*, *28*(11), 1747-1756. https://doi.org/10.1101/gr.239244.118

McClintock, M. K., Conzen, S. D., Gehlert, S., Masi, C., & Olopade, F. (2005). Mammary cancer and social interactions: identifying multiple environments that regulate gene expression throughout the life span. *J Gerontol B Psychol Sci Soc Sci*, *60 Spec No 1*, 32-41. https://doi.org/10.1093/geronb/60.special_issue_1.32

Mermel, C. H., Schumacher, S. E., Hill, B., Meyerson, M. L., Beroukhim, R., & Getz, G. (2011). GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic

copy-number alteration in human cancers. *Genome Biol*, *12*(4), R41. https://doi.org/10.1186/gb-2011-12-4-r41

Miles, D. W., Chan, A., Dirix, L. Y., Cortes, J., Pivot, X., Tomczak, P., . . . Romieu, G. (2010). Phase III study of bevacizumab plus docetaxel compared with placebo plus docetaxel for the first-line treatment of human epidermal growth factor receptor 2-negative metastatic breast cancer. *J Clin Oncol*, *28*(20), 3239-3247. https://doi.org/10.1200/jco.2008.21.6457

Miller, K., Wang, M., Gralow, J., Dickler, M., Cobleigh, M., Perez, E. A., . . . Davidson, N. E. (2007). Paclitaxel plus bevacizumab versus paclitaxel alone for metastatic breast cancer. *N Engl J Med*, *357*(26), 2666-2676. https://doi.org/10.1056/NEJMoa072113

Milligan, M. J., & Lipovich, L. (2014). Pseudogene-derived lncRNAs: emerging regulators of gene expression. *Front Genet*, *5*, 476. https://doi.org/10.3389/fgene.2014.00476

Mittal, V. (2018). Epithelial Mesenchymal Transition in Tumor Metastasis. *Annu Rev Pathol*, *13*, 395-412. https://doi.org/10.1146/annurev-pathol-020117-043854

Mittendorf, E. A., Philips, A. V., Meric-Bernstam, F., Qiao, N., Wu, Y., Harrington, S., . . . Alatrash, G. (2014). PD-L1 expression in triple-negative breast cancer. *Cancer Immunol Res*, *2*(4), 361-370. https://doi.org/10.1158/2326-6066.CIR-13-0127

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med*, *151*(4), 264-269, W264. https://doi.org/10.7326/0003-4819-151-4-200908180-00135

Moller, H., Henson, K., Luchtenborg, M., Broggio, J., Charman, J., Coupland, V. H., . . . Purushotham, A. (2016). Short-term breast cancer survival in relation to ethnicity, stage, grade and receptor status: national cohort study in England. *Br J Cancer*. https://doi.org/10.1038/bjc.2016.335

Moreno-Estrada, A., Gravel, S., Zakharia, F., McCauley, J. L., Byrnes, J. K., Gignoux, C. R., . . . Bustamante, C. D. (2013). Reconstructing the population genetic history of the Caribbean. *PLoS Genet*, *9*(11), e1003925. https://doi.org/10.1371/journal.pgen.1003925

Murray, P. (1950). *States' Laws on Race and Color, and Appendices: Containing International Documents, Federal Laws and Regulations, Local Ordinances and Charts*. Women's Division of Christian Service, Board of Missions and Church ….

Murray, T., Beaty, T. H., Mathias, R. A., Rafaels, N., Grant, A. V., Faruque, M. U., . . . Barnes, K. C. (2010). African and non-African admixture components in African Americans and an African Caribbean population. *Genet Epidemiol*, *34*(6), 561-568. https://doi.org/10.1002/gepi.20512

Narayan, P., Wahby, S., Gao, J. J., Amiri-Kordestani, L., Ibrahim, A., Bloomquist, E., . . . Beaver, J. A. (2020). FDA Approval Summary: Atezolizumab Plus Paclitaxel Protein-bound for the Treatment of Patients with Advanced or Metastatic TNBC Whose Tumors Express PD-L1. *Clinical Cancer Research*, *26*(10), 2284-2289. https://doi.org/10.1158/1078-0432.ccr-19-3545

Nead, K. T., Li, A., Wehner, M. R., Neupane, B., Gustafsson, S., Butterworth, A., . . . Meyre, D. (2015). Contribution of common non-synonymous variants in PCSK1 to body mass index variation and risk of obesity: a systematic review and meta-analysis with evidence from up to 331 175 individuals. *Hum Mol Genet*, *24*(12), 3582-3594. https://doi.org/10.1093/hmg/ddv097

Nemesure, B., Wu, S. Y., Hambleton, I. R., Leske, M. C., Hennis, A. J., & Group, B. N. C. S. (2009). Risk factors for breast cancer in a black population--the Barbados National Cancer Study. *Int J Cancer*, *124*(1), 174-179. https://doi.org/10.1002/ijc.23827

Nemesure, B., Wu, S. Y., Hennis, A., Leske, M. C., & Group, B. N. C. S. (2009). Body size and breast cancer in a black population--the Barbados National Cancer Study. *Cancer Causes Control*, *20*(3), 387-394. https://doi.org/10.1007/s10552-008-9253-3

Newman, L. A., Griffith, K. A., Jatoi, I., Simon, M. S., Crowe, J. P., & Colditz, G. A. (2006). Meta-analysis of survival in African American and white American patients with breast cancer: ethnicity compared with socioeconomic status. *J Clin Oncol*, *24*(9), 1342-1349. https://doi.org/10.1200/jco.2005.03.3472

Newman, L. A., & Kaljee, L. M. (2017). Health Disparities and Triple-Negative Breast Cancer in African American Women: A Review. *JAMA Surg*, *152*(5), 485-493. https://doi.org/10.1001/jamasurg.2017.0005

O'Neil, N. J., Bailey, M. L., & Hieter, P. (2017). Synthetic lethality and cancer. *Nat Rev Genet*, *18*(10), 613-623. https://doi.org/10.1038/nrg.2017.47

O'Reilly, E. A., Gubbins, L., Sharma, S., Tully, R., Guang, M. H., Weiner-Gorzel, K., . . . McCann, A. (2015). The fate of chemoresistance in triple negative breast cancer (TNBC). *BBA Clin*, *3*, 257-275. https://doi.org/10.1016/j.bbacli.2015.03.003

Opoku, S. Y., Benwell, M., & Yarney, J. (2012). Knowledge, attitudes, beliefs, behaviour and breast cancer screening practices in Ghana, West Africa. *Pan Afr Med J*, *11*, 28.

Palma, G., Frasci, G., Chirico, A., Esposito, E., Siani, C., Saturnino, C., . . . D'Aiuto, M. (2015). Triple negative breast cancer: looking for the missing link between biology and treatments. *Oncotarget*, *6*(29), 26560-26574. https://doi.org/10.18632/oncotarget.5306

Pang, M. F., Georgoudaki, A. M., Lambut, L., Johansson, J., Tabor, V., Hagikura, K., . . . Fuxe, J. (2016). TGF-β1-induced EMT promotes targeted migration of breast cancer cells through the lymphatic system by the activation of CCR7/CCL21-mediated chemotaxis. *Oncogene*, *35*(6), 748-760. https://doi.org/10.1038/onc.2015.133

Papadimitriou, M., Mountzios, G., & Papadimitriou, C. A. (2018). The role of PARP inhibition in triple-negative breast cancer: Unraveling the wide spectrum of synthetic lethality. *Cancer Treat Rev*, *67*, 34-44. https://doi.org/10.1016/j.ctrv.2018.04.010

Patel, A., Unni, N., & Peng, Y. (2020). The Changing Paradigm for the Treatment of HER2-Positive Breast Cancer. *Cancers (Basel)*, *12*(8). https://doi.org/10.3390/cancers12082081

Patey, D., & Scarff, R. (1928). The position of histology in the prognosis of carcinoma of the breast. *Lancet*, 801-804.

Pierre, C. C., Hercules, S. M., Yates, C., & Daniel, J. M. (2019). Dancing from bottoms up - Roles of the POZ-ZF transcription factor Kaiso in Cancer. *Biochim Biophys Acta Rev Cancer*, *1871*(1), 64-74. https://doi.org/10.1016/j.bbcan.2018.10.005

Pierre, C. C., Longo, J., Mavor, M., Milosavljevic, S. B., Chaudhary, R., Gilbreath, E., . . . Daniel, J. M. (2015). Kaiso overexpression promotes intestinal inflammation and potentiates intestinal tumorigenesis in Apc(Min/+) mice. *Biochim Biophys Acta*, *1852*(9), 1846-1855. https://doi.org/10.1016/j.bbadis.2015.06.011

Pitt, J. J., Riester, M., Zheng, Y., Yoshimatsu, T. F., Sanni, A., Oluwasola, O., . . . Barretina, J. (2018). Characterization of Nigerian breast cancer reveals prevalent homologous recombination deficiency and aggressive molecular features. *Nat Commun*, *9*(1), 4181. https://doi.org/10.1038/s41467-018-06616-0

Popli, P., Gutterman, E. M., Omene, C., Ganesan, S., Mills, D., & Marlink, R. (2021). Receptor-Defined Breast Cancer in Five East African Countries and Its Implications for Treatment: Systematic Review and Meta-Analysis. *JCO Glob Oncol*, *7*, 289-301. https://doi.org/10.1200/GO.20.00398

Prat, A., & Perou, C. M. (2011). Deconstructing the molecular portraits of breast cancer. *Mol Oncol*, *5*(1), 5-23. https://doi.org/10.1016/j.molonc.2010.11.003

Prat, A., Pineda, E., Adamo, B., Galván, P., Fernández, A., Gaba, L., . . . Muñoz, M. (2015). Clinical implications of the intrinsic molecular subtypes of breast cancer. *Breast*, *24 Suppl 2*, S26-35. https://doi.org/10.1016/j.breast.2015.07.008

Prokhortchouk, A., Hendrich, B., Jorgensen, H., Ruzov, A., Wilm, M., Georgiev, G., . . . Prokhortchouk, E. (2001). The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor. *Genes Dev*, *15*(13), 1613-1618. https://doi.org/10.1101/gad.198501

Qin, B., Babel, R. A., Plascak, J. J., Lin, Y., Stroup, A. M., Goldman, N., . . . Llanos, A. A. M. (2020). Neighborhood social environmental factors and breast cancer subtypes among Black women. *Cancer Epidemiol Biomarkers Prev*. https://doi.org/10.1158/1055-9965.EPI-20-1055

R Core Team. (2020). R: A language and environment for statistical computing. R Foundation for Statistical

Computing. In.

Raghav, S. K., Waszak, S. M., Krier, I., Gubelmann, C., Isakova, A., Mikkelsen, T. S., & Deplancke, B. (2012). Integrative genomics identifies the corepressor SMRT as a gatekeeper of adipogenesis through the transcription factors C/EBPbeta and KAISO. *Mol Cell*, *46*(3), 335-350. https://doi.org/10.1016/j.molcel.2012.03.017

Ragin, C., Banydeen, R., Zhang, C., Ben, A., Calabrese, V., Villa, N. N., . . . Louden, D. (2018). Breast Cancer Research in the Caribbean: Analysis of Reports From 1975 to 2017. *J Glob Oncol*, *4*, 1-21. https://doi.org/10.1200/JGO.18.00044

Rakha, E. A., Reis-Filho, J. S., Baehner, F., Dabbs, D. J., Decker, T., Eusebi, V., . . . Ellis, I. O. (2010). Breast cancer prognostic classification in the molecular era: the role of histological grade. *Breast Cancer Res*, *12*(4), 207. https://doi.org/10.1186/bcr2607

Reddon, H., Gueant, J. L., & Meyre, D. (2016). The importance of gene-environment interactions in human obesity. *Clin Sci (Lond)*, *130*(18), 1571-1597. https://doi.org/10.1042/cs20160221

Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., & Mesirov, J. P. (2011). Integrative genomics viewer. *Nat Biotechnol*, *29*(1), 24-26. https://doi.org/10.1038/nbt.1754

Ruff, S. E., Logan, S. K., Garabedian, M. J., & Huang, T. T. (2020). Roles for MDC1 in cancer development and treatment. *DNA Repair (Amst)*, *95*, 102948. https://doi.org/10.1016/j.dnarep.2020.102948

Sabatier, R., Gonçalves, A., & Bertucci, F. (2014). Personalized medicine: present and future of breast cancer management. *Crit Rev Oncol Hematol*, *91*(3), 223-233. https://doi.org/10.1016/j.critrevonc.2014.03.002

Santarpia, L., Bottai, G., Kelly, C. M., Győrffy, B., Székely, B., & Pusztai, L. (2016). Deciphering and Targeting Oncogenic Mutations and Pathways in Breast Cancer. *Oncologist*, *21*(9), 1063-1078. https://doi.org/10.1634/theoncologist.2015-0369

Scott, L. C., Mobley, L. R., Kuo, T. M., & Il'yasova, D. (2019). Update on triple-negative breast cancer disparities for the United States: A population-based study from the United States Cancer Statistics database, 2010 through 2014. *Cancer*, *125*(19), 3412-3417. https://doi.org/10.1002/cncr.32207

Service, B. S. (2013). *2010 Population and Housing Census: Volume 1*. Barbados Statistical Services.

Shackney, S. E., & Silverman, J. F. (2003). Molecular evolutionary patterns in breast cancer. *Adv Anat Pathol*, *10*(5), 278-290. https://doi.org/10.1097/00125480-200309000-00003

Sharma, P. (2017). PARP inhibitor and platinum agent in triple negative breast cancer: utilizing innovative trial design to bring together something "new" and something "old". *Chin Clin Oncol*, *6*(1), 3. https://doi.org/10.21037/cco.2016.12.01

Siegel, R. L., Miller, K. D., & Jemal, A. (2017). Cancer Statistics, 2017. *CA Cancer J Clin*, *67*(1), 7-30. https://doi.org/10.3322/caac.21387

Simpson, P. T., Reis-Filho, J. S., Gale, T., & Lakhani, S. R. (2005). Molecular evolution of breast cancer. *J Pathol*, *205*(2), 248-254. https://doi.org/10.1002/path.1691

Singhal, S. K., Byun, J. S., Park, S., Yan, T., Yancey, R., Caban, A., . . . Gardner, K. (2021). Kaiso (ZBTB33) subcellular partitioning functionally links LC3A/B, the tumor microenvironment, and breast cancer survival. *Commun Biol*, *4*(1), 150. https://doi.org/10.1038/s42003-021-01651-y

Slamon, D. J., Clark, G. M., Wong, S. G., Levin, W. J., Ullrich, A., & McGuire, W. L. (1987). Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science*, *235*(4785), 177-182. https://doi.org/10.1126/science.3798106

Sossey-Alaoui, K. (2013). Surfing the big WAVE: Insights into the role of WAVE3 as a driving force in cancer progression and metastasis. *Semin Cell Dev Biol*, *24*(4), 287-297. https://doi.org/10.1016/j.semcdb.2012.10.006

Spence, D., Dyer, R., Andall-Brereton, G., Barton, M., Stanway, S., Argentieri, M. A., . . . Sarfati, D. (2019). Cancer control in the Caribbean island countries and territories: some progress but the journey continues. *Lancet Oncol*, *20*(9), e503-e521. https://doi.org/10.1016/S1470-2045(19)30512-1

Ssentongo, P., Lewcun, J. A., Candela, X., Ssentongo, A. E., Kwon, E. G., Ba, D. M., . . . Dodge, D. G. (2019). Pooled 5-year survival rate in breast cancer population in Africa stratified by sub-regions (East Africa, West Africa, South Africa and North Africa) from which the papers used in the meta-analysis were published. *Figshare*. https://doi.org/10.1371/journal.pone.0225039.g007

Stark, A., Kleer, C. G., Martin, I., Awuah, B., Nsiah-Asare, A., Takyi, V., . . . Newman, L. (2010). African ancestry and higher prevalence of triple-negative breast cancer: findings from an international study. *Cancer*, *116*(21), 4926-4932. https://doi.org/10.1002/cncr.25276

Stringer-Reasor, E. M., May, J. E., Olariu, E., Caterinicchia, V., Li, Y., Chen, D., . . . Yang, E. S. (2021). An open-label, pilot study of veliparib and lapatinib in patients with metastatic, triple-negative breast cancer. *Breast Cancer Research*, *23*(1), Article 30. https://doi.org/10.1186/s13058-021-01408-9

Stroup, D. F., Berlin, J. A., Morton, S. C., Olkin, I., Williamson, G. D., Rennie, D., . . . Thacker, S. B. (2000). Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA*, *283*(15), 2008-2012. https://doi.org/10.1001/jama.283.15.2008

Sung, H., DeSantis, C. E., Fedewa, S. A., Kantelhardt, E. J., & Jemal, A. (2019). Breast cancer subtypes among Eastern-African-born black women and other black women in the United States. *Cancer*, *125*(19), 3401-3411. https://doi.org/10.1002/cncr.32293

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin*, *71*(3), 209-249. https://doi.org/10.3322/caac.21660

Tate, J. G., Bamford, S., Jubb, H. C., Sondka, Z., Beare, D. M., Bindal, N., . . . Forbes, S. A. (2019). COSMIC: the Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Res*, *47*(D1), D941-D947. https://doi.org/10.1093/nar/gky1015

Taylor, M. A., Davuluri, G., Parvani, J. G., Schiemann, B. J., Wendt, M. K., Plow, E. F., . . . Sossey-Alaoui, K. (2013). Upregulated WAVE3 expression is essential for TGF-β-mediated EMT and metastasis of triple-negative breast cancer cells. *Breast Cancer Res Treat*, *142*(2), 341-353. https://doi.org/10.1007/s10549-013-2753-1

Torre, L. A., Siegel, R. L., Ward, E. M., & Jemal, A. (2016). Global Cancer Incidence and Mortality Rates and Trends--An Update. *Cancer Epidemiol Biomarkers Prev*, *25*(1), 16-27. https://doi.org/10.1158/1055-9965.EPI-15-0578

Tran, B., & Bedard, P. L. (2011). Luminal-B breast cancer and novel therapeutic targets. *Breast Cancer Res*, *13*(6), 221. https://doi.org/10.1186/bcr2904

Tremont, A., Lu, J., & Cole, J. T. (2017). Endocrine Therapy for Early Breast Cancer: Updated Review. *Ochsner J*, *17*(4), 405-411.

Turner, N. C., & Tutt, A. N. (2012). Platinum chemotherapy for BRCA1-related breast cancer: do we need more evidence? *Breast Cancer Res*, *14*(6), 115. https://doi.org/10.1186/bcr3332

Tutt, A., Tovey, H., Cheang, M. C. U., Kernaghan, S., Kilburn, L., Gazinska, P., . . . Bliss, J. M. (2018). Carboplatin in BRCA1/2-mutated and triple-negative breast cancer BRCAness subgroups: the TNT Trial. *Nat Med*, *24*(5), 628-637. https://doi.org/10.1038/s41591-018-0009-7

Vermeulen, J. F., van de Ven, R. A., Ercan, C., van der Groep, P., van der Wall, E., Bult, P., . . . Derksen, P. W. (2012). Nuclear Kaiso expression is associated with high grade and triple-negative invasive breast cancer. *PLoS One*, *7*(5), e37864. https://doi.org/10.1371/journal.pone.0037864 PONE-D-12-05852 [pii]

Wang, W., Kansakar, U., Markovic, V., Wang, B., & Sossey-Alaoui, K. (2020). WAVE3 phosphorylation regulates the interplay between PI3K, TGF-β, and EGF signaling pathways in breast cancer. *Oncogenesis*, *9*(10), 87. https://doi.org/10.1038/s41389-020-00272-0

Whitehurst, A. W., Bodemann, B. O., Cardenas, J., Ferguson, D., Girard, L., Peyton, M., . . . White, M. A. (2007). Synthetic lethal screen identification of chemosensitizer loci in cancer cells. *Nature*, *446*(7137), 815-819. https://doi.org/10.1038/nature05697

Williams, J. B., Pang, D., Delgado, B., Kocherginsky, M., Tretiakova, M., Krausz, T., . . . Conzen, S. D. (2009). A model of gene-environment interaction reveals altered mammary gland gene expression and increased tumor growth following social isolation. *Cancer Prev Res (Phila)*, *2*(10), 850-861. https://doi.org/10.1158/1940-6207.CAPR-08-0238

Williams, S., Bateman, A., & O'Kelly, I. (2013). Altered expression of two-pore domain potassium (K2P) channels in cancer. *PLoS One*, *8*(10), e74589. https://doi.org/10.1371/journal.pone.0074589

Wyatt, T. E., Pernenkil, V., & Akinyemiju, T. F. (2017). Trends in breast and colorectal cancer screening among U.S. adults by race, healthcare coverage, and SES before, during, and

after the great recession. *Prev Med Rep*, *7*, 239-245. https://doi.org/10.1016/j.pmedr.2017.04.001

Xie, F., Ling, L., van Dam, H., Zhou, F., & Zhang, L. (2018). TGF-β signaling in cancer metastasis. *Acta Biochim Biophys Sin (Shanghai)*, *50*(1), 121-132. https://doi.org/10.1093/abbs/gmx123

Xu, X., Zhang, L., He, X., Zhang, P., Sun, C., Lu, Y., & Li, F. (2018). TGF-β plays a vital role in triple-negative breast cancer (TNBC) drug-resistance through regulating stemness, EMT and apoptosis. *Biochem Biophys Res Commun*, *502*(1), 160-165. https://doi.org/10.1016/j.bbrc.2018.05.139

Yanagawa, M., Ikemot, K., Kawauchi, S., Furuya, T., Yamamoto, S., Oka, M., . . . Sasaki, K. (2012). Luminal A and luminal B (HER2 negative) subtypes of breast cancer consist of a mixture of tumors with different genotype. *BMC Res Notes*, *5*, 376. https://doi.org/10.1186/1756-0500-5-376

Yedjou, C. G., Tchounwou, P. B., Payton, M., Miele, L., Fonseca, D. D., Lowe, L., & Alo, R. A. (2017). Assessing the Racial and Ethnic Disparities in Breast Cancer Mortality in the United States. *Int J Environ Res Public Health*, *14*(5). https://doi.org/10.3390/ijerph14050486

Yu, N. Y., Iftimi, A., Yau, C., Tobin, N. P., van 't Veer, L., Hoadley, K. A., . . . Lindström, L. S. (2019). Assessment of Long-term Distant Recurrence-Free Survival Associated With Tamoxifen Therapy in Postmenopausal Patients With Luminal A or Luminal B Breast Cancer. *JAMA Oncol*, *5*(9), 1304-1309. https://doi.org/10.1001/jamaoncol.2019.1856

Yuan, J., Hu, Z., Mahal, B. A., Zhao, S. D., Kensler, K. H., Pi, J., . . . Zhang, L. (2018). Integrated Analysis of Genetic Ancestry and Genomic Alterations across Cancers. *Cancer Cell*, *34*(4), 549-560.e549. https://doi.org/10.1016/j.ccell.2018.08.019

Zakharia, F., Basu, A., Absher, D., Assimes, T. L., Go, A. S., Hlatky, M. A., . . . Tang, H. (2009). Characterizing the admixed African ancestry of African Americans. *Genome Biol*, *10*(12), R141. https://doi.org/10.1186/gb-2009-10-12-r141

Zhang, M., Wu, J., Mao, K., Deng, H., Yang, Y., Zhou, E., & Liu, J. (2017). Role of transforming growth factor-β1 in triple negative breast cancer patients. *Int J Surg*, *45*, 72-76. https://doi.org/10.1016/j.ijsu.2017.07.080

Zurrida, S., & Veronesi, U. (2015). Milestones in breast cancer treatment. *Breast J*, *21*(1), 3-12. https://doi.org/10.1111/tbj.12361