Department of Microbiology, Immunology and Parasitology, College of Health Sciences, School of Medicine, Addis Ababa University

Evaluating Microbiota Signatures, Intrinsic Subtypes, *PIK3CA* Gene Mutations and Alternative Diagnostic Modalities Among Ethiopian Breast Cancer Patients

By:

Zelalem Desalegn Woldesonbet (B.Sc., M.Sc)

A Dissertation Submitted to the Department of Microbiology, Immunology and Parasitology (DMIP), School of Medicine, College of Health Sciences, Addis Ababa University in Partial Fulfilment of the Requirements for the Degree of Doctor of Philosophy (PhD) in Medical Microbiology

June, 2024

Addis Ababa Ethiopia

Evaluating Microbiota Signatures, Intrinsic Subtypes, *PIK3CA* Gene Mutations and Alternative Diagnostic Modalities Among Ethiopian Breast Cancer Patients

Ph.D. candidate: Zelalem Desalegn Woldesonbet, Department of Microbiology, Immunology and Parasitology, School of Medicine, College of Health Sciences, Addis Ababa University, Ethiopia

E-mail: zelalem.desalegn@aau.edu.et/ tzollove@gmail.com

Main advisors:

1. **Tamrat Abebe (M.Sc., PhD, Assistant Professor):** Department of Microbiology, Immunology & Parasitology School of Medicine, College of Health Sciences, Addis Ababa University Tikur Anbessa Specialized Hospital Email: tamrat.abebe@aau.edu.et
2. **Eva Kantelhardt (M.Sc., MD, PhD, Professor):** Institute for Medical Epidemiology, Biometry and Computer Science, Medical faculty of the Martin Luther University Halle-Wittenberg; Email: eva.kantelhardt@gmx.de

Co-advisors:

3. **Athena Starlford-Davenport (PhD, Associate Professor):** Department of Genetics, Genomics and Informatics, The University of Tennessee Health Science Center, Memphis, TN, USA
4. **Martina Vetter (PhD):** Department of Gynecology, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany

Collaborators:

5. **Dr. Adamu Addissie (MD, PhD, Associate Professor):** Department of preventive Medicine, School of Public Health, Addis Ababa University
6. **Dr. Mathewos Assefa (MD, Associate Professor):** Department of Oncology, School of Medicine, Addis Ababa University
7. **Dr. Yonas Bekuretsion (MD, Associate Professor):** Department of Pathology, School of Medicine, Addis Ababa University

DISSERTATION DECLARATION

Declaration from PhD candidate:

I, Zelalem Desalegn Woldesonbet, declare that this dissertation is my original work. I have written and submitted to AAU-CHS, Department of Microbiology, Immunology and Parasitology never to any other institution in any form for evaluation. All the information here is dully acknowledged and I have never used any other source except those cited ones.

PhD candidate: **Zelalem Desalegn Woldesonbet, BSc, MSc**

Signature: _____

Date: _____

Contact address: Department of Microbiology, Immunology and Parasitology, College of Health Sciences and Addis Ababa University, P.O. Box 9086

Mobile #: +251912029567

Email: zelalem.desalegn@aau.edu.et

Declaration from Advisors:

I/ we, hereby declare that I/we have advised this Ph.D. work, the dissertation has been prepared under my/our supervision and has been submitted with my/our approval.

1. Dr. Tamrat Abebe (M.Sc, PhD, Assistant Professor)
   Signature: _____ Date: _____

2. Professor Eva Kantelhardt (MD, PhD)
   Signature: Eva Kantelhardt Date: _____

3. Athena Starlford-Davenport (MSc, PhD, Associate Professor)
   Signature: Athena Starlford-Davenport Date: _____
4. Martina Vetter (PhD)
   Signature: *Martina Vetter* Date: _____

Evaluating Microbiota Signatures, Intrinsic Subtypes, *PIK3CA* Gene Mutations and Alternative Diagnostic Modalities Among Ethiopian Breast Cancer Patients

ABSTRACT

Background: Female breast cancer incidence has risen to unprecedented levels. It is a complex disease presented with distinct morphology, biological behaviors, clinical outcome and prognosis. Consequently, classification of breast cancer, understanding the molecular events including *PIK3CA* mutations and exploration of environmental factors including microbiota have a central role in understanding tumor biology. However, there is limited evidences in sub-Saharan Africa (SSA) context including Ethiopia; therefore, this study was conducted to generate evidence-based data to improving breast cancer care in Ethiopia. Additionally, considering the limited infrastructure for breast cancer diagnosis, we did method evaluation between PCR methods and immunohistochemistry to suggest alternative methods.

Objectives: The study aims to evaluate microbiota signatures, determine intrinsic subtypes of breast cancer, determine the rate of *PIK3CA* gene mutations and investigate the diagnostic performance of PCR to determine ER, PR and HER2 markers against immunohistochemistry.

Methods: A pathologically confirmed breast cancer patients recruited to the study. Clinical and histopathological data were collected using a questionnaire coupled with a follow up data for survival analysis. Formalin-fixed paraffin-embedded (FFPE) tumor tissue and paired (normal adjacent and breast tumor tissues) samples were collected. IHC analysis, PAM50-based gene expression profiling, and 16S rRNA gene-based sequencing were carried out. PAM50 based assay was employed to classify breast cancer into distinct intrinsic subtypes. Hot spot mutations across exon 9 and 20 in *PIK3CA* gene was determined using TaqMan® Mutation detection assay applying a competitive Allele-Specific TaqMan® PCR platform. 16S rRNA gene amplicons were PCR-amplified using dual-barcoded primers targeting the V4 region followed by high through put sequencing in Illumina MiSeq. Concordance between IHC and PCR method was evaluated using Cohen kappa coefficient. SPSS, Python, and R-packages were employed for data analysis. Survival analysis was performed using Kaplan-Meier and Cox regression.

Result: The tumor samples were classified into PAM50 intrinsic subtypes as follows: 104 samples (31.1%) were luminal A, 91 samples (27.2%) were luminal B, 62 samples (18.6%) were HER2-enriched and 77 samples (23.1%) were basal-like. The intrinsic subtypes were found to be associated with clinical and histopathological parameters such as hormone receptor status (estrogen, progesterone), HER2 status, Ki-67 proliferation index, and tumor differentiation, but not with age, tumor size or histological type. IHC groups were found to be correlated with intrinsic subtypes.

In this study, the overall mutation across *PIK3CA* gene demonstrated that the prevalence was turned to be 20.4% (46/226). The rate of mutation was 24.4% (ER+), 27.3% (PR+) and 20% (HER2-). A marked difference in *PIK3CA* gene mutation was identified across IHC groups. High proportion of mutations noted among HR+/any HER2 tumors (HR+/HER2−: 24.1%; HR+/HER2+ tumors: 27.8%). The rate of *PIK3CA* mutations had a statistically significant association with specific prognostic features. In Cox's multivariate regression analysis, only advanced pT stage and HR−/HER2+ tumors were independent predictors of impaired survival.

Bacterial 16SrRNA sequencing revealed distinct differences in microbiota composition and abundance with respect to prognostic features. We observed no significant difference in alpha diversity between the tumor and normal adjacent tissues. *Sphingobium, Anaerococcus, Corynebacterium, Delftia*, and *Enhydrobacter* were most significantly decreased in breast tumors compared to NAT tissues. Several microbial genera significantly differed by clinicopathological factors in Ethiopian women with breast cancer. Significant differences in *Exiguobacterium* (p=0.0097) and several additional genera were observed among tumors according to intrinsic subtypes, IHC group, and advanced-stage of the disease. Specifically, the genus *Burkholderia* more strongly correlated with aggressive triple negative breast cancer and basal-like breast tumors. The genera *Alkanindiges, Anoxybacillus, Leifsonia*, and *Exiguobacterium* most strongly correlated with HER2-E tumors. Luminal A and luminal B tumors also correlated with *Anoxybacillus* but not as strongly as HER2-E tumors. A relatively higher abundance of the genus *Citrobacter* most significantly correlated with advanced-stage breast tumors compared to early-stage tumors.

Finally, we determined the performance of mRNA and protein-based techniques. Accordingly, while comparing IHC Vs end point PCR method, the expression of marker increased by 40.37% (ER), 30.33% (PR) and 32.83% (HER2) when using FFT samples than FFPE. Analyzing FFPE samples, the sensitivity was lower across the specific markers. A statistically significant agreement was noted between endpoint PCR and IHC for: ER (k = 0.332; P-value = 0.001, PgR (k=0.135 and P-value= 0.001; HER2 (k=0.134; P-value= 0.021) in FFPE tissues. Concordance of biomarker expression between the qPCR and IHC methods was 94.44% (kappa =0.758) for ER, 91.67% (kappa = 0.806) for PR, and 46.67% (kappa = 0.436) for HER2.

Conclusion: The distribution of intrinsic subtypes confirms previous immunohistochemistry-based studies from Ethiopia showing potentially endocrine-sensitive tumors in more than half of the patients. The findings indicated that PAM50-based classification offers a robust method for the molecular classification of tumor. While assessing mutations across *PIK3CA* gene, the proportion of mutation was high in favorable clinical and histopathological features including HR+ tumors. Considering endocrine therapy resistance challenge, these group of patients with *PIK3CA* mutation can benefit from combining endocrine therapy, chemotherapy, and targeted PI3K activity inhibitors. Though not yet explored in our setting and statistically significant, a specific microbial abundance noted between NAT and breast tumor. Again, a marked difference in the relative distribution of distinct microbiota across clinical and histopathological features was observed. Furthermore, a strong correlation of certain microbial communities with paired breast tissues and further prognostic features. The study highlighted that mammary microbiota composition differs significantly across tumor and cohort clinical and histopathological features which encourage to conduct further wide-scale study strengthen the current data and further utilize microbiota as a potential marker. Besides understanding the molecular biology and environmental factors including microbiota, working towards alternative diagnostic modalities is essential. Accordingly, applying PCR method, biomarker expression was higher while utilizing FFT than FFPE. Improvement of agreement noted between qPCR and IHC in FFT samples. Finally, the finding confirmed that mRNA-based breast cancer marker detection could be a complementary and/ or an alternative method in Ethiopia taking into account the limited IHC service availability.

Keywords: Breast cancer, Ethiopia, IHC, Intrinsic subtypes, Microbiota, Mutation, Overall survival, PCR, *PIK3CA*, Prognostic features,

DEDICATION

This particular work has been dedicated to almighty God, my family, advisors, my country-Ethiopia, friends, colleagues and collaborators.

God, your unconditional love and grace was massive and I shall not leave your house forever and dwelling in your presence is more than everything. I am who I am because of you; You made me beautifully and wonderfully to your glory. Thank you and praise to be your glorified name!

Such great work and accomplishment would not have been possible without the genuine love, support and encouragement from my wife Tigist Rahmet (Honey). Dear, I am grateful for your uninterrupted input and understanding the situation I have been through in all the ups and downs. You were, are and will be my greatest divine gift and blessings which I cannot compare with anything in this world. Thanks for taking me to the top of the mountain and helped a lot to the accomplishment I achieved as a researcher and/ or scientist in my area of field. Honey, I am also grateful for your irreplaceable gifts: Amen (Ami), Eliada (Eldu) and Aviya (Avu). Though it was difficult to leave these gifts and you, your outstanding home management and leadership gave me the confidence and enabled me to had protected time abroad to realize my dream. I could not forget the smile on your face on the day of my return from abroad that was awesome. You are always my permanent love and blessings!

Dear Dr. Tamrat, your support and contribution to my personal and profession development was great. You were always by my side and taught me the reward of hard working in life! Dear Professor Eva (Marthin Luther University, Halle, Germany) and Dr.Athena-Starlford (University of Tennessee Health Science Center, USA) your kind support and gracious follow up was unforgettable and impactful.

ACKNOWLEDGMENTS

First and most, I would like to thank the almighty God for his unlimited mercy and abundant provision. It is also a great pleasure to extend my heartfelt appreciation to my wife Tigist Rahmet and my beautiful angels: Amen, Eliada, and Aviya) for their unreserved support and patience for handling responsibilities in my absence.

I am highly excited to thank my advisors Dr. Tamrat Abebe Professor Eva Kantelhardt, Dr. Athena Starlford-Davenport and Dr. Martina Vetter for their unreserved guidance and scientific support during the conceptualization up until writing my dissertation. I highly acknowledge their utmost contribution in the preparation of draft manuscript and reviewing papers before submission to peer-reviewed reputable journals. It is a great pleasure to extend my appreciation to Dr. Endale Anberber, Dr. Yonas Bekuretsion and Dr. Mathewos Assefa for providing unlimited support in the identification/clinical evaluation of patients, development of data collection instrument, tracking clinical and/or histopathological parameters and collection of breast tissue samples. Without these key renowned specialists, it would have not been possible to reach where we are.

Also, I would like to thank residents at oncology department (Dr. Amanuel, Dr. Jilcha, Dr. Yasin, oncologic nurses and people working in Card room (Mulu and Nemi) for their continuous support during data collection and approaching breast cancer patients for follow up data collection. Taking this opportunity, I would like to extend my appreciation to laboratory technician in Germany including Kathrine and Sandy. Kathrine was there in all the process and her support was quite incredible and tremendous from the initial phase of sample preparation through actual laboratory activities.

I would like to thank Susan G. Komen foundation for the financial support via Marthin-Luther University. Our heartfelt thanks go to Marthin-Luther University for covering my travel and accommodation expenses from DAAD-PAGEL fund. Furthermore, my deepest gratitude goes to University of Tennessee Health center and Dr. Athena's research team for covering the expenses to carry out the experiment focusing on sequencing 16S rRNA gene for microbiota signature evaluation.

Last but not least, I would like to extend my appreciation to the study participants for being willing to be part of the study because without their voluntary participation realization of the project would have been impossible.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS AND/OR ACRONYMS

ASCO:	American Society of Clinical Oncology
ASVs:	Amplicon Sequence Variants
BC:	Breast Cancer
BCCLs:	Breast Cancer Cell Lines
BHD:	Breast Cancer Homology Domain
CAP:	College of American Pathologists
CISH:	Chromogen in-situ Hybridization
COSMIC:	Catalog Of Somatic Mutations in Cancer
CT:	Threshold Cycle
DCIS:	Ductal Carcinoma In-Situ
DFS:	Disease-Free Survival
DNA:	Deoxyribonucleic Acid
ER:	Estrogen Receptor
ESMO:	European Society for Medical Oncology
FDA:	Food and Drug Administration
FFPE:	Formalin-Fixed Paraffin-Embedded
FFT:	Fresh-frozen Tissue
GEO:	Gene Expression Omnibus
GEP:	Gene Expression Profiling
GOI:	Gene of Interest
HDI:	Human Development Index
HER2:	Human Epidermal Growth Factor Receptor 2
HIC:	High-income Countries

HMP:	Human Microbiome Project
IARC:	International Agency for Research on Cancer
IHC:	Immunohistochemistry
IS:	Intrinsic Subtypes
KM:	Kaplan-Meier
LCIS:	Lobular Carcinoma In-Situ
LMICs:	Low middle-income Countries
MFS:	Metastasis Free Survival
MiRNA:	Micro Ribonucleic Acid
mRNA:	Messenger RNA
NCDs:	Non-Communicable Diseases
NGS:	Next-Generation Sequencing
NST:	No Special Type
OS:	Overall Survival
OTUs:	Operational Taxonomic Unit
PAM50:	Prediction Analysis of Microarrays 50
PCoA:	Principal Coordinate Analysis
PCR:	Polymerase Chain Reaction
PgR:	Progesterone Receptor
PI3K:	Phosphatidylinositol 3-kinase (PI3K)
PTEN:	Phosphatase and Tension Homolog
RAS BD:	RAS-Binding Domain
RFS:	Recurrence-free Survival
ROR:	Risk-of Relapse
RT-qPCR:	Quantitative Reverse Transcriptase Polymerase Chain Reaction

SDI:	Shannon's Diversity Index
SOPs:	Standard Operating Procedures
SSA:	Sub-Saharan Africa
TCGA:	The Cancer Genome Atlas
TKR:	Tyrosine Kinase Receptor
TME:	Tumor Microenvironment
TNBC:	Triple-negative Breast Cancer
WHO:	World Health Organization
WLWHA:	Women Living with HIV/AIDS

1. INTRODUCTION

1.1. Background and Rationale

The epidemiologic trend of diverse cancers is escalating globally (Farmer et al., 2010). Consequently, the incidence and death rate associated with cancer is expected to nearly double by 2030 (Bray et al., 2012; Ferlay et al., 2015; Soerjomataram et al., 2012). The burden of cancer is increasing particularly in economically disadvantaged populations where the majority of cases are diagnosed in late stages (Gebrehiwot et al., 2014). Breast cancer (BC) is the most frequently diagnosed malignancy worldwide with the highest mortality in SSA (Getachew et al., 2020; Sung et al., 2021). In 2020, an estimated 2.3 million cases of female BC were diagnosed globally, and about 685,000 women died from the disease (Sung et al., 2021). Despite the burden of cancer, there is a global integrated and cooperative efforts on prevention and control of cancer lead by the WHO initiative focused on BC and beyond (Awedew et al., 2022).

Low- and middle-income countries account for approximately 60% of worldwide BC incidence, and women in low and middle-income countries (LMICs) are approximately 64% more likely to die as a result of barriers to getting quality treatments (WHO/IARC, 2018). Noncommunicable diseases (NCDs), including BC, are on the rise in Africa, owing to developments in health care, which have resulted in higher life expectancy and earlier identification of cancer. In Africa, BC accounts for one in every four cancer diagnoses and one in every five cancer deaths among women (Ferlay et al., 2010).

In LMICs, 50-70% of patients have advanced-stage disease at diagnosis, which contributes to high death rates (Shamseddine et al., 2010). Unfortunately, cancer mortality rates in African countries are not similar to those in high-income countries (HIC) (Vanderpuye et al., 2017), reaching intolerable levels. Despite the growing disease burden, SSA countries continue to battle the disease with suboptimal diagnostic capacity, access to life-saving targeted therapies for BC , low health literacy levels resulting in late stage at initial presentation, and comprehensive universal health coverage (Vanderpuye et al., 2021). The Breast Health Global Initiative has stressed the importance of downstaging strategies that promote early detection and prompt access to appropriate therapy in LMICs if cancer is to be managed optimally (Anderson et al., 2011). The challenge of the optimal management of BC in SSA is not only compromised in terms of access to care as outlined above but also actual data on the incidence that is critical for planning of cancer. Several factors could account for this but the lack of a well-established cancer registry is the main factor (WHO, 2018).

In Ethiopian context, new cases of BC is increasing and data shows that it is becoming one of the most common cancer types attributable to high rates of morbidity and mortality (Abate, 2016; Lukong et al., 2017; Timotewos et al., 2018). Supporting this evidence, a study from this setting demonstrated that it is the most common cancer accounting for 33% of cancers among women and 23% of all cancers in Ethiopia with an estimated national age standardized incidence rate of 43 per 100,000 population in women (Memirie et al., 2018). Despite the burden of the disease, research-evidence on tumor heterogeneity and tumor biology in Ethiopia is very limited. Therefore, evaluation of the BC biology is critical considering the genetic diversity of African/Ethiopian population.

With the advancement of molecular biology, systems biology and genome sciences in the last decades, our understanding about BC disease has been dramatically expanded at cellular, molecular and genomic levels. Again, it is important to distinguish between the various subtypes because they have different prognoses and treatment implications (Feng et al., 2018a). Evidences indicated that BC exhibits different histopathological and biological features which ultimately result in distinct tumor behaviors that lead to a diverse treatment responses demanding to devise and implement context driven cancer management and treatment strategies (Blows et al., 2010). Therefore, there is a high need to accurately group the tumor into clinically distinct subtypes and appropriately tailor the management.

Furthermore, a variety of cell signaling pathways including PI3K-AKT-mTOR was identified as a complex intracellular pathway involved in bringing change in a given cell's biological characteristics to cause cancer, mediate pathogenesis, and facilitate the progression of human cancers (Yuan & Cantley, 2008). Alteration in tyrosine kinase receptor (TKR)-phosphatidylinositol 3-kinase (PI3K) signaling pathways have been investigated as a frequent phenomenon among a wide range of human cancers. Of the many known aberrations, PI3K activation is the most important molecular event engaged in the regulation of several signaling pathways controlling cell survival, apoptosis, proliferation, motility, and adhesion (J. J. Zhao et al., 2005). A high frequency of *PIK3CA* mutations has been reported in various cancer models (Bader et al., 2006; J. W. Lee et al., 2005; J. J. Zhao et al., 2005). The prevalence of *PIK3CA* oncogene mutation is the most common among BC patients (mutation rate: 20% to 40%) (Saal et al., 2005; Stemke-hale et al., 2009) following TP53 suppressor gene.

A series of research evidence revealed that BC cells harbor high PI3 K pathway activity which is known as the most common aberrations besides the HER2 amplification and P53 mutation (Hernandez-Aya & Gonzalez-Angulo, 2011; Miron et al., 2010; Yamamoto et al., 2011). Intriguingly, multiple studies demonstrated that *PIK3CA* mutations is considered as more common in estrogen receptor positive (30% - 40%) than in estrogen receptor negative breast tumors (10% - 20%) (Kalinsky et al., 2009; Saal et al., 2005; Samuels et al., 2005). In studies which focused on prognostic significance of *PIK3CA* gene mutations in BC, contradicting results were reported (Pang et al., 2014; Zardavas et al., 2018). In a study conducted by Li and co-workers (G. Li et al., 2018), their finding confirmed that mutations across the gene somehow related to unfavorable clinical outcomes.

Cancer is caused by accumulation of spontaneous mutations in the course of DNA replication related to uncontrolled potent stem cells which undergo a relatively large number of cell divisions; however, with accumulated knowledge, environment critically affect cancer risk, progression and outcome (Ashford et al., 2015; C. C. Harris, 2016). Breast carcinogenesis and/or cancer growth and/or drug sensitivity has a multifactorial etiology—perhaps the least well-characterized aspect being that of the distant environmental influences, namely, the microbiota that inhabit humans. Microbes present in breast tissue given the known presence of bacteria in human milk (Urbaniak et al., 2012).

Breast tissues have been found to contain a distinct microbiome that differs from pathologically characterized normal, benign, and malignant breast tissue. Growing evidence suggests a link between an imbalance in the composition and number of bacteria in breast tissue (known as microbial dysbiosis) and BC in women (Smith et al., 2019). The microenvironment in and around the tumor contains a diverse range of cell types, including microbiota. Pathophysiological alterations in cells and microbial composition may have a substantial impact on tumor growth. While significant advances have been made in cancer diagnosis and treatment, microbial contributions to tumor formation remain unknown (Thompson et al., 2017). Considering the burden of BC in Ethiopia, due emphasis has to be given to understanding the molecular biology and the diverse tumor heterogeneity to get the full picture of the problem and comprehensively improving cancer care. Nowadays, researchers are in quest of finding additional diagnostic and prognostic marker which are helpful for early detection and predicting the prognosis and therapeutic benefits of patients with BC.

1.2. Statement of the problem

Breast cancer is the major cause of cancer morbidity among adult women in Ethiopia, accounting for one-third of all female cancers and one in every five cancer cases (Memirie et al., 2018; Timotewos et al., 2018). Every year, the country observes an estimated 16,133 new BC cases and 9061 fatalities (Bray et al., 2018). In Ethiopia, the majority of BC patients present with late-stage disease and are aged 35-45 years, which is related with suffering and premature death (Ayele et al., 2021, 2022). Alarmingly, more than 70% of women in Ethiopia were diagnosed at an advanced stage of the disease, and with limited access to patient-tailored treatment, their chances of survival are very low (Kantelhardt. et al., 2014). To avoid premature death, early case detection, diagnostic and treatment capacities for BC must be improved (Shita et al., 2023). To realize such objective, elucidating BC tumor biology, molecular aberrations of cancer driving genes, microbiota signatures and working towards alternative diagnostic modalities has importance.

BC is a heterogeneous group of tumors associated with a wide range of morphological features and molecular subtypes resulting unique biological behaviors, presentation, and prognosis. Therefore, understanding the tumor molecular pattern has fundamental importance to identify patients who will be benefited particularly from a given treatment based on expression of molecular markers. In the last couple of years, distinct intrinsic molecular subtypes of BC (Luminal A, Luminal B, HER2-enriched, Basal-like and Claudin-low) have been identified and intensively studied (Prat, Pineda, et al., 2015).

BC intrinsic subtypes (IS) were first described in 2000 by Perou and colleagues who utilized DNA microarrays representing more than 8000 genes in 65 breast tumor surgical specimens and 17 cultured cell lines (Perou et al., 2000). This new approach provided new insights into the biology of BC and affected the therapeutic approach of BC (Desalegn et al., 2022). In Ethiopia much is not known about the intrinsic subtypes of breast cancer and only few laboratories provide the IHC testing to classify BC and tailor the treatment. The advancement in molecular tools and data from gene expression profiling (GEP) has had a considerable impact on our understanding of BC biology, classification modalities and tailoring treatment approach and monitoring of treatment outcome.

Furthermore, taking into account BC heterogeneity and complexity in terms of molecular biology, classification and clinical outcome/ tumor behavior, further understanding of tumor biology with respect to mutation along certain gene has paramount importance in applying targeted therapy and predicting patient clinical outcome.

A wide range of genes which are susceptible to oncogenic mutations are related to the development of BC. In the last decades, the intrinsic subtyping and tailor of the treatment has improved the management of BC significantly. However, in patients with metastatic BC, the disease is becoming incurable and among the associated factors identified the abreactions affecting the (PI3K)/Akt/mammalian target of rapamycin (mTOR) pathway is one. More specifically, this is seen more among HR+ BC patients (Fusco et al., 2021). Somatic mutations in *PIK3CA* account for close to 30% of the processes that enhance PI3K/AKT/mTOR signaling, the most common oncogenic signaling pathway linked to BC (Shimoi et al., 2018). However, though HR+ BC is common in Ethiopia, the extent of this aberrations in Ethiopian BC patients is not documented.

Findings indicated that mutation in *PIK3CA* gene associated with favorable prognostic features including ER-positive/HER2-negative tumors; there is also a new evidence of *PIK3CA* mutation prevalence in HER2-positive tumors (Elwy et al., 2017). Along the *PIK3CA* gene, more than 80% of somatic mutations concentrated in exon 9 (E542K or E545K) and the exon 20 (H1047R or H1047L) (X. Zhao, 2013). Nowadays, a wide-range of research-based evidences are coming out about the relevance of targeting PI3K_AKT_mTOR signaling pathways taking into account its role in various biological activities; however, there are no studies in Ethiopia addressing on this topic. Therefore, elucidation of the features of BC with *PIK3CA* mutation is important for targeted treatment and predicting potential prognoses outcome.

Besides the genetic and non-genetic bases of BC, evidences demonstrate that a wide range of factors including environmental factor are attributable for BC development. With recent knowledge, environmental factors which is explained by a diverse microbial colonization of host cells and/or tissue becoming a center of research theme and there is a strong push to fully characterize the bacteria associated with different parts of the body under different health conditions. These studies have been made possible with the use of deep-sequencing technologies, and sites once thought of as sterile, such as the stomach, bladder, and lungs, have now been shown to harbor an indigenous microbiota (Bik et al., 2006; Wolfe et al., 2012). Since microbiota do have role in health and disease, exploring their abundance in patients with BC would have a relevance in targeting them as a diagnostic and prognostic marker to improve BC management in Ethiopia context.

In general, understanding BC classification, mutation in *PIK3CA* and exploiting the role of microbiota would have a role in improving BC care in Ethiopia. In general, determination of tumor markers is important strategies for the clinical management of cancer patients, guiding in diagnostic, staging, examination of response to treatment, monitoring of recurrence and metastasis, and development of new treatment modalities (Banin Hirata et al., 2014).

The challenges of limited laboratory infrastructure result in sub-optimal care. Therefore, exploration of alternative diagnostic modalities and evaluating their performance in comparison with a gold standard method is essential for better BC management in the context of economically disadvantaged setting like Ethiopia. IHC is considered as a gold standard in determining individual surface protein expression (ER, PR and HER2) and classifying BC into subgroups and tailoring treatment/management. However, IHC despite its being a game changer in BC management, it is associated with a number of problems.

Despite international efforts to standardize the cutoffs and the method (Goldhirsch et al., 2009; Kurosumi, 2007; Price & Goldhirsch, 2005), a complete standardization of IHC is difficult owing to the inherent semiquantitative nature of the technology. In clinical practice, substantial problems in biomarker diagnostics in BC have been reported (Allred et al., 2008; Layfield et al., 2000). Though immunohistochemistry (IHC) on formalin-fixed paraffin-embedded (FFPE) is the gold standard for the assessment of hormonal receptor status, HER2 and Ki67 (Allegra et al., 1980; Cuzick et al., 2011; Dowsett et al., 2011), gene expression-based method including polymerase chain reaction (PCR) also identified as alternative approach complementing immunohistochemical techniques to quantify the level of expression of gene of interest (GOI).

Nowadays, researchers are in quest of finding working diagnostic and prognostic marker which are helpful for early detection and predicting prognosis and therapeutic response to benefit patients with BC care in Ethiopia.

1.3. Significance of the project

Cancer remains a serious public health issue in Africa, owing primarily to the country's limited resources and healthcare infrastructure, which leads to inefficient preventive and control health measures. Though there are optimistic initiatives to battle breast cancer, the effectiveness of BC control remains a concern, particularly in resource-limited areas. BC is a diverse and complex disease. It is made up of various biological subgroups with different behaviors and responses to therapy. Gene expression investigations have found numerous unique BC intrinsic subtypes (IS) that differ dramatically in both prognosis and therapeutic targets in cancer cells. IS are becoming increasingly significant in providing a comprehensive picture of the molecular genetics of a patient's tumor, predicting prognosis, and, most importantly, selecting a therapeutic approach.

In addition to genetics, the tumor microenvironment as well as the environment with which the individual interacts are believed to contributes to the development of the disease, although the factors involved in the development of BC are not well known. Emerging evidence suggests the interaction of microbiome with host as important factor at some stage in the carcinogenesis process. Among these, the role of microbiota in the initiation and progression of cancer is getting the greatest attention and more specifically that of the mammary microbiota in BC. As a result, research related with microbiota is becoming a center of research theme and various works are ongoing to explore their role in promotion of carcinogenesis and on the other hand to look for their protective effect. The exploration of a such interactions as well as the role of signature of genes involved in BC initiation and progression is becoming rate limiting and may provide a more accurate diagnostic, prognostic and therapeutic tool.

With previous study, it has been confirmed that mutations in certain gene noted in certain group of BC patients. In patients with hormone receptor positive tumors, frequency of *PIK3CA* mutations was higher as a result targeted therapy is recommended to enhance BC care. On top of exploring BC biology, considering the limited laboratory infrastructure in Ethiopian context, looking for cost-effective, technically easy and rapid method for the determination of molecular markers is critical for better cancer diagnosis and selecting treatment to enhance BC care in Ethiopia.

Despite the public health importance of BC in Ethiopia, much is not known about the population level burden and biology of the cancer. There is an acute demand of research-based evidences regarding population-based distribution of IHC groups, intrinsic subtypes, rate and role of somatic mutation in *PIK3CA* gene, microbiota signatures and alternative diagnostic modalities among Ethiopian BC patients.

In this study,

- Microbiota signatures among pathologically confirmed BC patients were evaluated and correlated with prognostic features of the BC patients.
    - The finding provided crucial evidence on the correlation of microbiota signatures with the clinical and pathological information.
- The distribution of intrinsic subtypes in Ethiopian patients with BC explored and it has been correlated with clinical and histopathological parameters
    - The finding enabled to identify the circulating distinct BC molecular subtypes and their correlation with clinical and pathological parameters
    - The research output would pave a way for understanding the whole picture of tumor biology in Ethiopian BC patients. Hence, this would have relevance in understanding tumor heterogeneity in Ethiopian population and selection effective treatment;
- The frequency of molecular aberrations in *PIK3CA* gene and its relation with patient overall survival evaluated.
    - Such exploration could help to understanding the burden of mutation to benefit a group of BC patients.
- Finally, we explored the diagnostic performance of PCR methods against a reference method (IHC) using FFPE and FFT samples
    - It enabled to see alternative PCR based diagnostic approach to determine specific protein markers

Generating evidences on basic biological aspects, molecular and genetic bases of BC could assist in the development of novel and targeted therapies as a means of realizing the full potential of personalized medicine for BC care in Ethiopia.

2. LITERATURE REVIEW

2.1. Global epidemiology

BC incidence and fatality rates have risen over the last three decades. Between 1990 and 2016, the incidence of BC more than quadrupled in 60 of 102 nations, while fatalities doubled in 43 of 102 countries. Current forecasts indicate that by 2030, the global number of new cases diagnosed will be 2.7 million per year, with 0.87 million fatalities. BC incidence is expected to rise further in low- and middle-income countries due to westernization of lifestyles (e.g., delayed pregnancies, reduced breastfeeding, low age at menarche, lack of physical activity, and poor diet), as well as improved cancer registration and detection (Smolarz et al., 2022).

BC has supplanted lung cancer as the most commonly diagnosed cancer globally, accounting for one in every eight cancer diagnoses and 2.3 million new cases in both sexes combined (Sung et al., 2021). It accounted for a quarter of all cancer cases in females in 2020, making it the most often diagnosed cancer in women, and its burden has been increasing in many regions of the world, particularly in transitioning countries. In 2020, an expected 685,000 women died from BC accounting for 16% or one out of every six cancer deaths in women. Due to an insufficient public health response to this development, the World Health Organization (WHO) has launched the Global Breast Cancer Initiative (Anderson et al., 2021). In women, BC was the most diagnosed cancer in 157 (out of 185) countries, followed by cervical cancer in 23, mainly SSA, countries and the leading cause of cancer death in 110 countries, followed by cervical cancer in 36 and lung cancer in 25 countries. Yet, great geographic variation exists in incidence and mortality rates ranging from highest incidence (Figure 1) (Arnold et al., 2022).

.

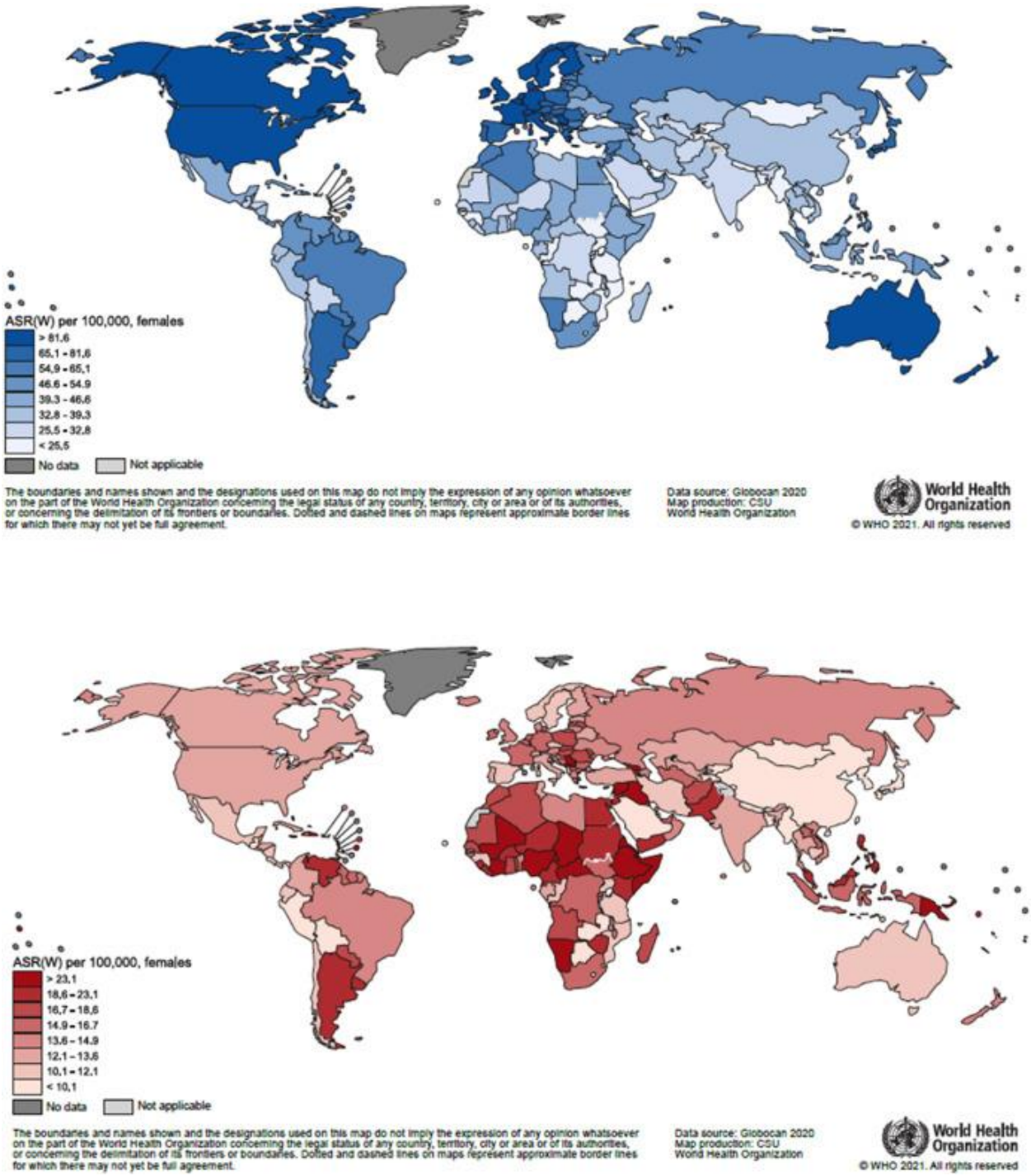


Figure 1: Breast cancer incidence (top, blue) and mortality (bottom, red) rates per 100,000 females, adjusted for age (Breast cancer incidences and fatalities by nation) (Arnold et al., 2022)

According to the 2020 GLOBOCAN report, the leading cancer surpassing lung cancer was female breast (11.7%) followed by lung cancer (11.4%). Referring to mortality, again BC was top ranked (Sung et al., 2021). Figure 2 shows age-standardized (world) top 10 cancers incidence and mortality rates distribution according to GLOBOCAN 2020 report.

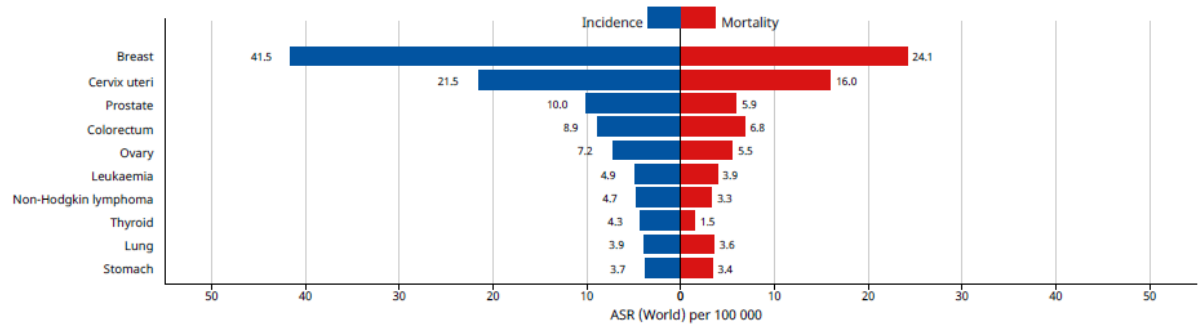


Figure 2: Age-standardized (World) incidence and death rates for the top ten malignancies (Sung et al., 2021)

Across the globe, BC occurs as the most common cancer and it is posing a formidable public health problem. Besides its high burden, BC is constituted of a group of biologically and molecularly heterogeneous diseases (Feng et al., 2018b). It is identified as a frequent malignancy globally and notably, this malignancy is the 1st or 2nd most common cancer in women in all but nine countries worldwide (Figure 3) (McCormack et al., 2018).

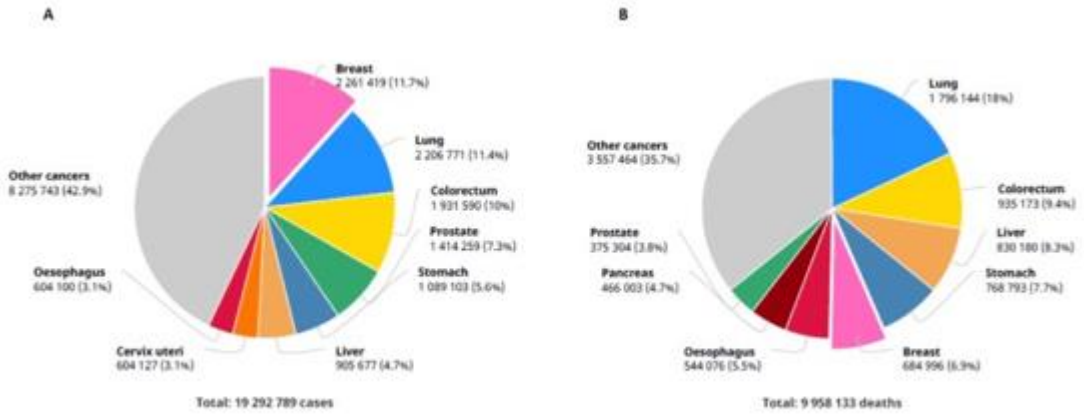


Figure 3: Number of new cases (A) and fatalities (B) in 2020 for both sexes and all ages (Sung et al., 2021)

2.2. Breast cancer in Africa

According to the GLOBOCAN reports, cancer burden disparities and heterogeneity are widely documented in transitioned and transitioning nations. According to reports, cancer incidence rates for men and women are three times greater in HIC than in LMICs (Soerjomataram & Bray, 2021). Emerging economies are dealing with a double burden of lifestyle and poverty-related malignancies. Though there are many factors that influence cancer incidence rates, urbanization, lifestyle, and increased life expectancy all have a substantial impact. When looking at the overall cancer burden in Africa, 1.1 million new cases and 711,429 deaths were anticipated for 2020. Egypt, Nigeria, South Africa, Ethiopia, and Morocco were the top five countries in terms of cancer incidence and deaths in 2020, accounting for 45% of new cases and 44% of cancer deaths (Table 1) (Sharma et al., 2022).

Overall, thorough statistics on the prevalence of BC in SSA are scarce (Pace & Shulman, 2016). In the majority of SSA nations where infrastructure has been created and is primarily dedicated to Maternal and Child Health (MCH) and infectious diseases, BC is emerging as a public health problem, with a large increase in incidence across the continent (Akarolo-Anthony, 2010). BC accounts for one out of every four cancer diagnoses and one out of every five cancer-related fatalities in Sub-Saharan Africa (Ferlay et al., 2010). In this context, BC is a rapidly expanding public health concern as the epidemiological transition shifts from mostly infectious causes to chronic, NCDs. As a result, several populations have experienced an increase in incidence of up to 5% each year (Coughlin & Ekwueme, 2009; Sabetghadam & Sabetghadam, 2013). Though the incidence of BC is significantly higher in industrialized countries, the developing world accounts for half of total cancer new cases and almost 60% of deaths (Jemal et al., 2011).

Table 1: Country-specific cancer incidence and mortality in Africa (Sharma et al., 2022)

Population	Incidence [95% UI]	ASIR	ASMR	Mortality [95% UI]	MIR
Algeria	58,418 [56,916–59,960]	135.3	76.1	32,802 [31,875–33,756]	0.56
Angola	20,327 [18,854–21,915]	130.6	86.5	12,599 [11,627–13,652]	0.62
Benin	6,747 [5,981–7,610]	95.8	69.3	4,662 [4,065–5,347]	0.69
Botswana	2,010 [1,874–2,156]	109.5	63.4	1,112 [1,036–1,193]	0.55
Burkina Faso	12,045 [10,722–13,531]	112	84.4	8,695 [7,649–9,884]	0.72
Burundi	7,929 [7,038–8,932]	135.5	103.8	5,701 [5,060–6,422]	0.72
Cabo Verde	825 [723–941]	179	107.6	488 [428–557]	0.59
Cameroon	20,745 [19,144–22,480]	127.6	86.6	13,199 [12,132–14,359]	0.64
Central African Republic	2,675 [2,270–3,152]	100.4	76.8	1,957 [1,676–2,285]	0.73
Chad	8,575 [7,277–10,104]	102.3	77.3	6,083 [5,210–7,103]	0.71
Comoros	609 [541–686]	111.8	79	411 [365–463]	0.67
Republic of the Congo	2,478 [2,242–2,739]	84.4	56.6	1,595 [1,428–1,781]	0.64
Côte d'Ivoire	17,300 [16,454–18,189]	123.3	88	11,760 [11,128–12,428]	0.68
Democratic Republic of Congo	48,839 [41,448–57,548]	102.9	74.9	34,412 [29,472–40,180]	0.7
Djibouti	765 [679–862]	91.0	65.3	527 [468–594]	0.69
Egypt	134,632 [131,480–137,860]	159.4	108.6	89,042 [86,773–91,371]	0.66
Equatorial Guinea	927 [787–1,092]	110.1	76.2	592 [507–691]	0.64
Eritrea	2,408 [2,137–2,713]	100.8	72.1	1,670 [1,482–1,881]	0.69
Eswatini	992 [917–1,074]	127.9	82.2	613 [560–671]	0.62
Ethiopia	77,352 [73,647–81,243]	106.7	75.3	51,865 [49,048–54,844]	0.67
France, La Réunion	2,965 [2,815–3,123]	212.5	97.8	1,523 [1,411–1,644]	0.51
Gabon	1,750 [1,516–2,020]	115.8	71.5	1,030 [882–1,203]	0.59
Ghana	24,009 [21,821–26,416]	115.9	80.6	15,802 [14254–17518]	0.66
Guinea	7,871 [7,297–8,490]	116.6	90.7	5,888 [5,465–6,344]	0.75
Guinea–Bissau	1,127 [902–1,408]	107.2	83.3	836 [679–1,029]	0.74
Kenya	42,116 [40,219–44,103]	149.2	103.2	27,092 [25,708–28,550]	0.64
Lesotho	1,876 [1,711–2,057]	109	74.8	1,256 [1,148–1,374]	0.67
Liberia	3,552 [2,842–4,439]	119.3	90.5	2,603 [2,114–3,205]	0.73
Libya	7,661 [7,063–8,309]	132.2	88.2	4,750 [4,381–5,150]	0.62
Madagascar	20,681 [18,358–23,298]	124.3	87.7	13,837 [12,282–15,588]	0.67
Malawi	17,936 [16,569–19,416]	154.2	113.4	12,454 [11,444–13,553]	0.69
Mali	14,185 [13,301–15,127]	140.3	107.5	10,234 [9,540.9–10,977]	0.72
Mauritania	3,079 [2,464–3,848]	108.6	78.3	2,121 [1,723–2,611]	0.69
Mauritius	3,050 [2,910–3,197]	155	72.4	1,504 [1,400–1,615]	0.49
Morocco	59,370 [57,415–61,392]	148.3	87.9	35,265 [34,015–36,561]	0.59
Mozambique	25,446 [23,874–27,122]	135.3	100.4	18,014 [16,766–19,355]	0.71
Namibia	3,345 [3,199–3,497]	198.3	116.1	1,876 [1,783–1,974]	0.56
Niger	9,787 [8,054–11,893]	78.4	62.2	7,382 [5,995–9,089]	0.75
Nigeria	124,815 [118,101–131,911]	110.4	74.8	78,899 [74,234–83,857]	0.63
Rwanda	8,835 [7,948–9,821]	113.9	81.4	6,044 [5,365–6,809]	0.68
Sao Tome and Principe	151 [69–328]	126.4	84.5	103 (47–224)	0.68
Senegal	11,317 [9,056–14,143]	119.5	87.4	7,893 [6,410–9,718]	0.7
Sierra Leone	4,708 [4,579–4,840]	102.1	76.9	3,389 [3,288–3,493]	0.72
Somalia	10,134 [8,995–11,417]	118.1	90.9	7,439 [6,603–8,380]	0.73
South Africa	108,168 [107,065–109,282]	209.5	111.7	56,802 [56,187–57,423]	0.53
South Sudan	6,312 [5,603–7,111]	94.7	72.2	4,633 [4,112–5,219]	0.73
Sudan	27,382 [26,675–28,108]	95.7	63.2	17,055 [16,600–17,523]	0.62
Tanzania	40,464 [38,682–42,328]	133.7	94.3	26,945 [25,586–28,376]	0.67
The Gambia	1,035 [904–1,184]	79.5	65.7	810 [706–930]	0.78
Togo	5,208 [4,765–5,693]	111.7	79.1	3,468 [3,176–3,787]	0.67
Tunisia	19,446 [18,737–20,182]	133.5	78.7	11,855 [11,404–12,323]	0.61
Uganda	34,008 [32,620–35,455]	153.8	112.4	22,992 [21,938–24,096]	0.68
Zambia	13,831 [13,007–14,708]	153.5	103.3	8,672 [8,086–9,300]	0.63
Zimbabwe	16,083 [15,394–16,802]	200.4	139.4	10,676 [10,163–11,215]	0.66
Africa	1,109,209 [965,409–1,274,430]	132.1	88.8	711,429 [611,604–827,547]	0.64

Incidence, all-age new cases; Mortality, all-age deaths; ASIR, age-standardized incidence rate (cases per 100,000); ASMR, age-standardized mortality rate (deaths per 100,000); MIR, mortality-to-incidence ratio. Data source: GLOBOCAN 2020 (International Agency for Research on Cancer). The figures inside square brackets depict 95% uncertainty intervals.

BC death rates in most Western countries have declined in recent years as a result of better treatment options and earlier identification (Autier et al., 2010; X. Wang et al., 2015). BC mortality is frequently greater in many low- and middle-income countries, including SSA (Winters et al., 2017) and rising in Asia, despite lower incidence, due to delayed presentation, late detection, and restricted treatment options. In 2020, transitioning nations accounted for 30% of all BC deaths globally; women in transitioning countries had 17% higher mortality rates than women in transitioned countries (Arnold et al., 2022). BC was the most common malignancy among African women in 2020, with 186,598 new cases and 85,787 deaths. BC , cervical cancer, and prostate cancer were the most common cancer types in the majority of African countries (Table 2) (Sharma et al., 2022).

Despite the load, Africa's cancer infrastructure is very weak. Only 16 of 43 countries have a cancer policy, strategy, or action plan in place, but 28 have a cancer registry. Despite the fact that breast and cervical cancer were the main malignancies in African females, only 15 countries provided BC screening via mammography at the public primary health care level, whereas 25 nations provided low-cost screening via clinical breast examination (Sharma et al., 2022).

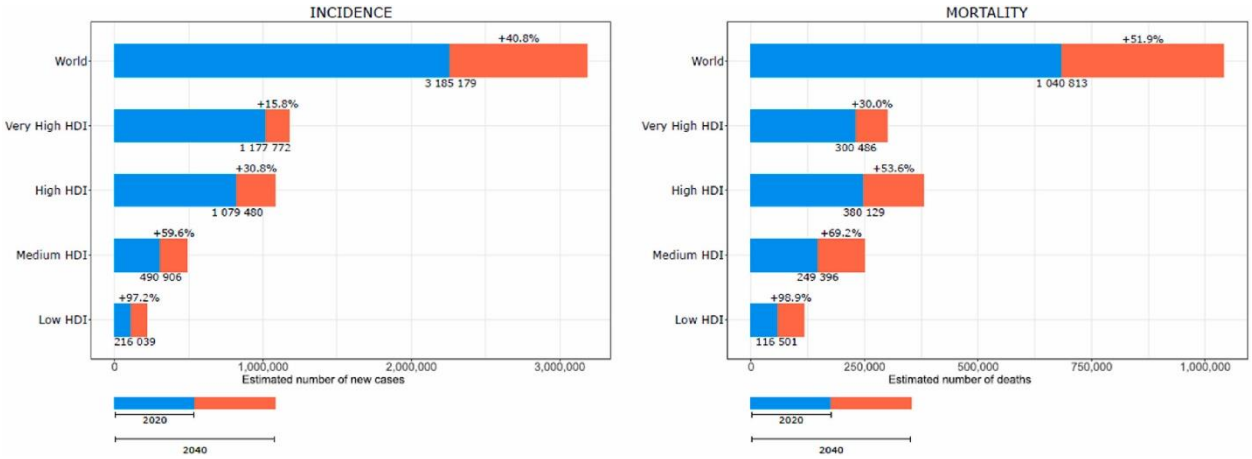
Table 2: Cancer-specific incidence and death rates across African nations (Sharma et al., 2022).

Cancer	Incidence [95% UI]	Mortality [95% UI]	ASIR	ASMR	MIR
Bladder	33,196 [27,881–39,525]	18,747 [15,372–22,862]	4.5	2.7	0.56
Brain and central nervous system	18,264 [14,370–23,213]	15,157 [11,539–19,909]	1.9	1.7	0.83
Breast	186,598 [173,041–201,217]	85,787 [77,648–94,779]	40.7	19.4	0.46
Cervix uteri	117,316 [105,999–129,842]	76,745 [68,380–86,133]	25.6	17.7	0.65
Colorectum	66,198	42,875	8.4	5.6	0.65
Corpus uteri	14,024 [10,920–18,010]	4,042 [3,041–5,372]	3.5	1	0.29
Gallbladder	5,454 [3,310–8,988]	4,249 [2,407–7,499]	0.75	0.59	0.78
Hodgkin lymphoma	10,815 [7,994–14,632]	4,315 [3,060–6,086]	0.96	0.43	0.4
Hypo pharynx	2,065 [1,051–4,058]	1,439 [667–3,103]	0.26	0.18	0.7
Kaposi sarcoma	25,010 [20,630–30,320]	13,066 [9,660–17,673]	2.2	1.2	0.52
Kidney	17,718 [13,606–23,073]	10,850 [8,035–14,651]	1.8	1.2	0.61
Larynx	9,908 [7,264–13,515]	6,636 [4,662–9,446]	1.3	0.91	0.67
Leukemia	32,138 [26,882–38,421]	23,891 [19,500–29,271]	3.2	2.6	0.74
Lip, oral cavity	14,286 [11,198–18,225]	8,088 [6,131–10,669]	1.8	1	0.57
Liver	70,542 [61,484–80,935]	66,944 [57,256–78,271]	8.8	8.5	0.95
Lung	45,988 [40,814–51,817.3]	41,171 [35,945–47,157]	6.2	5.6	0.9
Melanoma of skin	6,963 [5,360–9,046]	2,679 [1,989–3,608]	0.9	0.37	0.38
Mesothelioma	1,119 [593–2,110]	1,038 [505–2,135]	0.15	0.14	0.93
Multiple myeloma	8,491 [6,127–11,767]	7,069 [4,877–10,245]	1.1	0.97	0.83
Nasopharynx	10,041 [6,984–14,437]	6,600 [4,367–9,974]	1.1	0.77	0.66
Non–Hodgkin lymphoma	50,516 [43,758–58,317]	30,960 [26,294–36,453]	5.2	3.5	0.61
Esophagus	27,546 [23,474–32,324]	26,097 [21,756–31,305]	3.6	3.4	0.95
Oropharynx	2,913 [1,613–5,262]	1,782 [909–3,491]	0.36	0.23	0.61
Ovary	24,263 [19,547–30,117]	17,008 [13,301–21,748]	5.4	4	0.7
Pancreas	17,070 [13,706–21,260]	16,549 [12,893–21,242]	2.3	2.3	0.97
Penis	2,060 [1,013–4,191]	942 [420–2,113]	0.53	0.25	0.46
Prostate	93,173 [83,906–103,463]	47,249 [41,941–53,228]	29.7	16.3	0.51
Salivary glands	4,920 [2,987–8,104]	2,960 [1,678–5,222]	0.57	0.38	0.6
Stomach	32,402 [26,783–39,200]	27,945 [22,502–34,705]	4.2	3.7	0.86
Testis	3,302 [1,878–5,806]	1,084 [570–2,060]	0.61	0.24	0.33
Thyroid	18,457 [14,616–23,307]	4,443 [3,408–5,793]	2	0.62	0.24
Vagina	2,001 [1,031–3,883]	1,102 [518–2,343]	0.45	0.26	0.55
Vulva	5,144 [3,330–7,946]	2,858 [1,743–4,687]	1.1	0.66	0.56
All cancers	**1,109,209 [965,409–1,274,430]**	**711,429 [611,604–827,547]**	**132.1**	**88.8**	**0.64**

Incidence, all-age new cases; Mortality, all-age deaths; ASIR, age-standardized incidence rate (cases per 100,000); ASMR, age-standardized mortality rate (deaths per 100,000); MIR, mortality-to-incidence ratio. Data source: GLOBOCAN 2020 (International Agency for Research on Cancer). The figures inside square brackets depict 95% uncertainty intervals.

It is disturbingly projected that the incidence of cancer doubles in fewer resourced regions of Africa. Compared to other regions, Africa is predicted to witness the greatest growth in cancer incidence and mortality between 2020 and 2040. By 2040, there will be 1.4 million cancer-related deaths and 2.1 million cases in Africa, up from 1.1 million cases and 711, 429 deaths in 2020 (Sharma et al., 2022). It is anticipated that the annual number of newly diagnosed cases of BC will rise by more than 40% by 2040, to over 3 million cases.

In a similar vein, it is anticipated that the death rate from BC will rise by more than 50%, from 685,000 in 2020 to 1 million in 2040 (Figure 4). Transitioning nations will have a disproportionately high rise in new cases and deaths, which are predicted to double by 2040 (from 110,000 to 216,000 and from 59,000 to 116,000, respectively), particularly in low HDI nations. Transitioning countries will account for 30.1% of fatalities and 18.4% of instances of BC in 2020; by 2040, these numbers would climb to 35.2% and 22.2%, respectively. This estimate is purely the result of population aging and increase, and it could be further adjusted by shifts in incidence rates (Arnold et al., 2022).



.

Figure 4: Estimated number of BC cases and deaths from 2020 to 2040, based on HDI level (Arnold et al., 2022)

The incidence of five major cancers (breast, cervical, prostate, colorectal, and liver) is expected to double by 2040. By 2040, 12 neoplasms would have more than 50,000 cases per year, whereas nine neoplasms in Africa will have 50,000 or more cancer deaths (Figure 5).

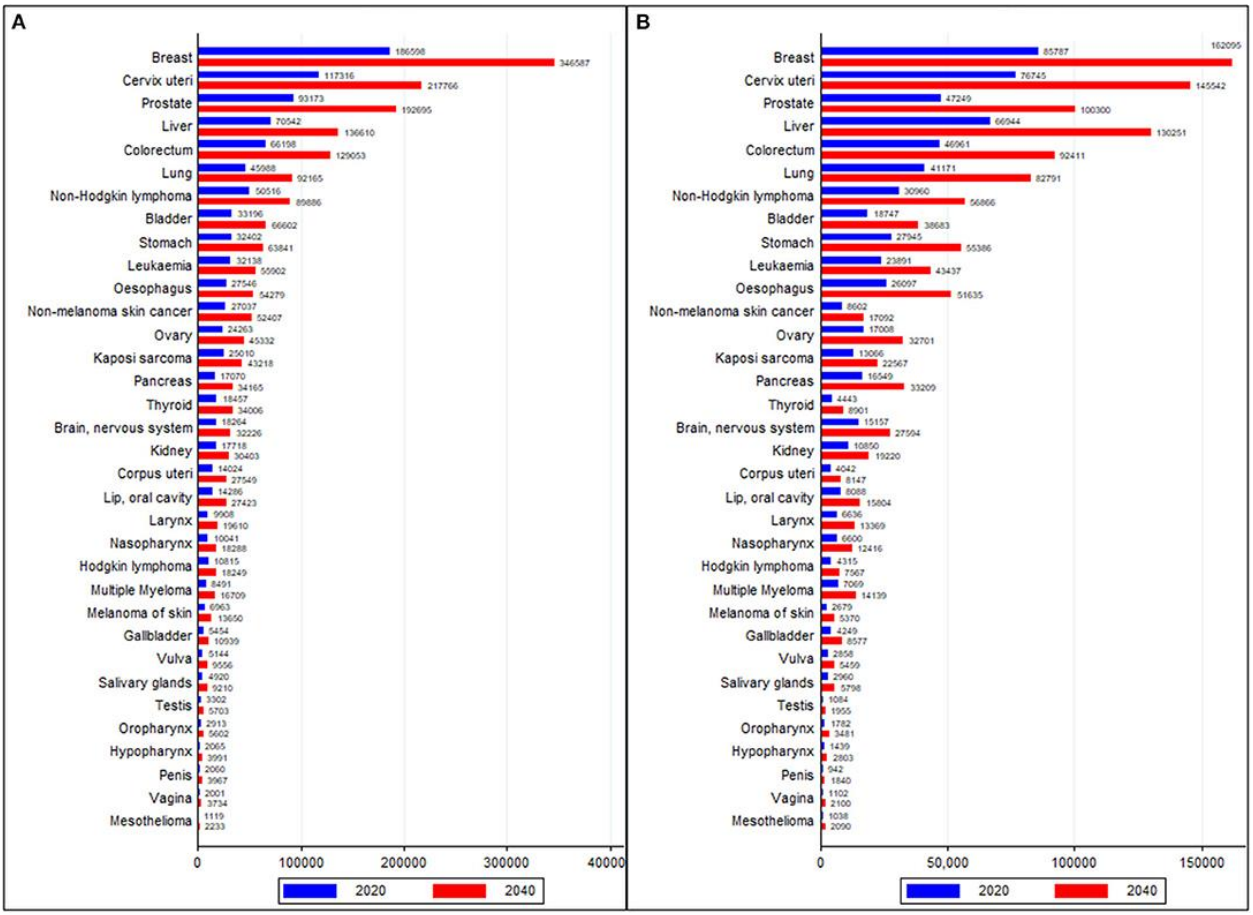


Figure 5: Cancer burden by cancer group in 2020 and projected values in 2040 (A) Incidence and (B) Deaths (Sharma et al., 2022)

2.3. The situation of breast cancer in Ethiopia

According to evidence-based data, the burden of BC is extremely substantial because it accounts for one-fourth of all cancer incidences, followed by cervical (Rick et al., 2020). In Ethiopia, it is becoming the leading cause of cancer-related death among females, accounting for one-third of all cancer cases among females and one-fifth of all cancer cases (Memirie et al., 2018; Timotewos et al., 2018). As a result, Ethiopia has an estimated 16,133 new BC cases and 9061 BC fatalities each year (Bray et al., 2018). Figure 6 depicts the distribution of BC in Ethiopia according to the GLOBOCAN 2020 data; accordingly, BC accounts for the lion's share of total new cancer cases (n=77,352), accounting for 16,133 (Sung et al., 2021). Another study from rural Ethiopia found that communicable diseases caused approximately 32% of 788 fatalities, neoplasms 12.1%, and pregnancy/maternal mortality 9.4%. BC was the second highest neoplasm, accounting for 21 (2.7%) of all fatalities and ranking among the top five causes of noncommunicable disease (Ayele et al., 2022).

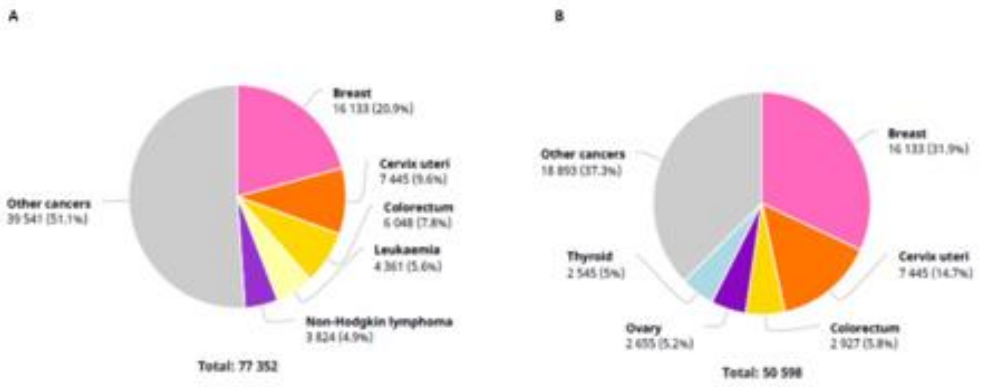


Figure 6: Number of new cases in 2020, both sexes (A) and females (B), of all ages in Ethiopia (Sung et al., 2021)

Despite the fact that NCDs such as cardiovascular disease, cancer, diabetes, and chronic respiratory diseases are on the rise in Ethiopia, the country's health institutions have historically focused on communicable disease control. As a result, the silent epidemic of NCDs now puts a 'double burden of sickness' on the country, which, unless handled, would overwhelm it in the near future (FMoH, 2015). Despite efforts in research and attention being paid to cancer, there is still a research shortage in this area in our environment. Previous hospital-based research has found that BC accounts for 18.3 to 33% of all detected cancer cases. Based on established facts, investigations revealed that Tikur Anbessa Specialized Hospital, the only hospital in the country where cancer patients can receive oncology care, has had consistent annual growth (Kantelhardt et al., 2014; Woldeamanuel et al., 2013).

Unfortunately, the majority of patients had advanced disease at the time of diagnosis (Wondemagegnhu, 2015). Alarmingly, over 80% of reported occurrences of cancer are detected at an advanced stage, when little can be done to treat the condition. This is mostly due to a lack of awareness of cancer signs and symptoms, insufficient screening, early detection, and treatment services, inadequate diagnostic facilities, and poorly structured referral (FMoH, 2015). Another TASH study, which analyzed the pattern and treatment result of 137 biopsy-proven BC cases from 1995 to 1999, found that the majority of BC patients had stage III disease. 45.9% of BC patients who were followed for a short amount of time experienced recurrences. Interestingly, the research result stated that the majority of the patients were young and in a late stage due to a high risk of recurrences (Ersumo, 2006).

Unlike African countries with a higher prevalence of TNBC, a study in Ethiopia found that the majority of BC patients were ER-positive. Again, the study found that the age of the patients at diagnosis was negatively associated to the fraction of ER-negative tumors, but was unaffected by histology or stage. Tamoxifen medication is therefore suggested for people with both known and uncertain ER status (Kantelhardt et al., 2014). Aside from the significant burden, a prior study found that 83.4% of patients presented with advanced-stage illness at diagnosis (Shita et al., 2023). According to the aforementioned studies, BC case-loads are steadily growing. Unfortunately, the findings revealed that two-thirds of the patients had advanced BC (Stages III and IV), which had occurred at a young age by the time the diagnosis was confirmed. The final statement emphasized the importance of establishing cancer registration centers, raising community awareness, and implementing an intensive early detection plan (Gebretsadik et al., 2021). In keeping with prior research, findings from Southern and South-Western Ethiopia revealed that the majority of BC patients were stage III with a relatively young age at diagnosis (Belachew et al., 2023).

According to a previous large-scale study aimed at determining the survival status of 1070 BC patients, women with this cancer in Ethiopia had a favorable 5-year outcome with 45% metastasis-free survival (MFS) when compared to reports from Sub-Saharan African countries, but the figure was still much lower than that of developed countries (Kantelhardt. et al., 2014). Despite the availability of treatment at tertiary health facilities, a survival rate of less than 60% has been observed, highlighting the need for additional improvements in early detection, diagnosis, and treatment accessibility to prevent premature mortality in these women (Shita et al., 2023).

Another study from Ethiopia, which focused on survival, found that the estimated 1- and 2-year overall survival probability rates were 78 and 53%, respectively, with 2-year survival being higher in patients with lymph node-negative disease, cT1-2 tumors, and hormone receptor-positive disease (Eber-Schulz et al., 2018).

2.4. Breast cancer risk factors

Table 3 shows a large number of risk factors for BC, including both modifiable and non-modifiable ones. However, for 70% of breast cancer patients, no risk factors are discovered (Smolarz et al., 2022).

2.4.1. Genetic predispositions

2.4.1.1. Mutations in the BRCA1 or BRCA2 gene

Advances in sequencing technology have allowed multigene testing, or "panel testing," a viable approach for identifying genetic abnormalities that may be related with a risk of breast cancer. Genetic testing for cancer risk is a standard aspect of medical practice. Until recently, testing was limited to patients with a significant family history of cancer and focused on a small number of genes linked to a high risk of cancer or specific cancer syndromes. Testing with broader gene panels is now possible because to inexpensive sequencing. However, for many genes on such panels, evidence of a link to cancer is often limited, and precise estimates of the cancer risks associated with variations are sometimes unavailable (Easton et al., 2015).

Breast cancer refers to a set of physiologically and molecularly diverse diseases that began in the breast. While the risk factors connected with this disease differ from other malignancies, genetic predisposition, most notably mutations in the *BRCA1* or *BRCA2* gene, is a key causative factor for this malignancy (Feng et al., 2018a). The BRCA1 gene, found on chromosome 17, is a suppressor gene that produces a nuclear protein that is responsible for genome stability. This protein, along with the products of other suppressor genes, signal transduction genes, and DNA damage detection, forms a protein complex that binds to RNA polymerase II and interacts with histone deacetylase, altering transcription, DNA repair, and recombination. The BRCA1 protein, along with the BRCA2 gene product, which is also a suppressor gene located on chromosome 13, is notably active in the repair of double DNA strand breaks by homologous recombination (Roy et al., 2016).

Overall, 5-10% of BC are caused by gene mutations inherited from a parent. The most prevalent cause of hereditary BC is an inherited mutation in the *BRCA1* or *BRCA2* genes. About 10% to 20% of BC patients have a first-degree relative with the disease. Up to 20% of women who

have a family history of breast cancer have a mutation in one of the BC susceptibility genes (*BRCA1* or *BRCA2*). Clinical data suggests that *BRCA1* mutation-related BC that do not overexpress ER or ERBB2 are related to the development of basal epithelial markers, which are already associated with ER/ERBB2-negative malignancies (Feng et al., 2018a).

According to statistics, women with a *BRCA1* mutation have a 55-65% lifetime risk of developing BC. Women with a BRCA2 mutation face a lifetime risk of 45%. A woman with a BRCA1 or BRCA2 gene mutation has around a 70% probability of developing BC by the age of 80. Women who carry one of these two mutations are also more likely to be diagnosed with BC at a younger age and to have disease in both breasts. In contrast, *BRCA1* mutations are less common in BC in men, and *BRCA2* mutations are associated with a lifetime BC risk of only around 6.8%. Testing of mutations in the *BRCA1* and *BRCA2* genes, as well as other less typically altered genes such as *PTEN* or *TP53*, in women at high risk can help in early detection and/or prevention of BC (Feng et al., 2018a).

2.4.1.2. Genetic testing of mutations other than BRCA1 and BRCA2 genes

Testing of mutations in the *BRCA1* and *BRCA2* genes, as well as other less typically altered genes such as *PTEN* or *TP53*, in women at high risk can help in early detection and/or prevention of BC (Feng et al., 2018a). Other suppressor genes with high-penetration mutations that predispose to breast cancer include TP53 (Li-Fraumeni syndrome) and PTEN (Cowden syndrome). Women with Li-Fraumeni syndrome had a 54% chance of developing breast cancer by the age of 70. Cowden's syndrome patients are 25-50% more likely to get breast cancer over their lifetime. However, both hereditary disorders are extremely rare (Angeli et al., 2020; Chamseddine et al., 2022).

Genetic testing for breast cancer susceptibility is widely utilized, although many genes have minimal evidence of a relationship with breast cancer, imprecise underlying risk estimates, and no credible subtype-specific risk estimates. In research finding, protein-truncating mutations in five genes (ATM, BRCA1, BRCA2, CHEK2, and PALB2) were linked to an increased risk of breast cancer overall, with a P value of less than 0.0001. Protein-truncating mutations in four additional genes (BARD1, RAD51C, RAD51D, and TP53) were related with a higher risk of breast cancer overall, with a P value of less than 0.05 and a Bayesian false-discovery probability of less than 0.05. For protein-truncating mutations in 19 of the remaining 25 genes, the odds ratio for breast cancer was less than 2.0 at the 95% confidence level (Leila Dorling, Sara Carvalho, Jamie Allen, Anna González-Neira, Craig Luccarini, Cecilia Wahlström, 2021).

For protein-truncating variants in *ATM* and *CHEK2*, odds ratios were higher for estrogen receptor (ER)–positive disease than for ER-negative disease; for protein-truncating variants in *BARD1*, *BRCA1*, *BRCA2*, *PALB2*, *RAD51C*, and *RAD51D*, odds ratios were higher for ER-negative disease than for ER-positive disease. Rare missense variants (in aggregate) in *ATM*, *CHEK2*, and *TP53* were associated with a risk of breast cancer overall with a P value of less than 0.001. For *BRCA1*, *BRCA2*, and *TP53*, missense variants (in aggregate) that would be classified as pathogenic according to standard criteria were associated with a risk of breast cancer overall, with the risk being similar to that of protein-truncating variants (Leila Dorling, Sara Carvalho, Jamie Allen, Anna González-Neira, Craig Luccarini, Cecilia Wahlström, 2021).

An exciting research finding was obtained from a population-based study of genes previously implicated in breast cancer. In this particular work, sequencing was performed using a custom multigene amplicon-based panel to identify germline pathogenic variants in 28 cancer-predisposition genes. Accordingly, pathogenic mutations in 12 recognized breast cancer risk genes were found in 5.03% of case patients and 1.63% of controls. Pathogenic mutations in BRCA1 and BRCA2 were linked to a high incidence of breast cancer. Pathogenic PALB2 variants were linked to a moderate risk. Pathogenic mutations in BARD1, RAD51C, and RAD51D were linked to an increased risk of estrogen receptor-negative and triple-negative breast cancer, while ATM, CDH1, and CHEK2 were linked to an increased risk of estrogen receptor-positive breast cancer. Pathogenic variants in 16 putative breast cancer-predisposition genes, including the c.657_661del5 founder pathogenic mutation in NBN, were not linked to an elevated risk of breast cancer (Yadav et al., 2021).

A genome-wide association study of breast cancer with 18,034 cases and 22,104 controls of African descent. Genetic variants at 12 loci were associated with breast cancer risk). These included a low-frequency missense variant rs61751053 in ARHGEF38 with overall breast cancer and a common variant rs76664032 at chromosome 2q14.2 with triple-negative breast cancer. Approximately 15.4% of TNBC cases carried six risk alleles in three genome-wide association study-identified TNBC risk variants, with an odds ratio of 4.21 (95% confidence range = 2.66-7.03) compared to those holding fewer than two risk alleles. A polygenic risk score (PRS) demonstrated an area under the receiver operating characteristic curve of 0.60 for the prediction of breast cancer risk, outperforming PRS. This study significantly increased the population diversity in genetic studies for breast cancer and shown the efficacy of PRS for risk prediction in females of African ancestry (Jia et al., 2024).

2.4.2. Non-genetic risk factors

In a broader context, variables such as widespread urbanization, changing patterns of reproductive and environmental risks, obesity, decreased physical activity, and rising life expectancy are thought to be predisposing factors for the occurrence of breast cancer (Akarolo-Anthony, 2010; Lingwood et al., 2008). In economically disadvantaged countries or low-income countries these factors are the cause for the steady rise in new BC cases. Similarly, higher incidence rates in transitioned countries reflect a longstanding higher prevalence of reproductive, hormonal and behavioral risk factors (Arnold et al., 2022).

Even while the risk factors for BC in Africa are similar to those in high-income countries, differences in risk factor occurrences could explain the disparity between African and high-income countries (Vanderpuye et al., 2017).

Table 3: Hereditary and non-genetic risk factors for BC (Smolarz et al., 2022)

Non-Modifiable Factors	Modifiable Factors
Female sex	Hormonal replacement therapy
Older age	Diethylstilbestrol
Family history (of breast or ovarian cancer)	Physical activity
Genetic mutations	Overweight/obesity
Race/ethnicity	Alcohol intake
Pregnancy and breastfeeding	Smoking
Menstrual period and menopause	Insufficient vitamin supplementation
Density of breast tissue	Excessive exposure to artificial light
Previous history of breast cancer	Intake of processed food
Non-cancerous breast diseases	Exposure to chemicals
Previous radiation therapy	Other drugs

2.5. Pathophysiology, heterogeneity and evolution of breast cancer

2.5.1. Pathophysiology of breast cancer

The exact mechanism by which BC originates is unknown; nonetheless, major efforts have been made to molecularly characterize and identify BC formation and evolution. At the cell of origin level, the clonal evolution model (in which mutations accumulate, epigenetic changes occur in tumor cells, and the 'fittest' cells survive) and the cancer stem cell model (in which only the precursor cancer cells initiate and sustain progression) are both implicated, which is complicated by the fact that cancer stem cells can also evolve clonally (Bombonati & Sgroi, 2011).

Morphologically, there are various lesions and genetic alterations, ranging from normal glands to malignancies. At the molecular level, findings reveal that BC proceeds via two separate molecular pathways, the majority of which are associated with ER expression, tumor grade, and proliferation. Furthermore, the identification of BC susceptibility genes has shed light on certain aspects of the pathophysiology of both spontaneous and inherited breast cancer.

2.5.2. Heterogeneity and evolution of breast cancer

Breast cancer tumor heterogeneity is one of the hallmarks of malignancy, which includes intertumoral heterogeneity seen in BC from different people and intratumor heterogeneity caused by the existence of various cell populations inside a single tumor. BC intratumor heterogeneity is the primary barrier to the development of effective treatments and customized medicine. For a long time, BC phenotypic heterogeneity was employed to classify breast tumors according to histological categories. The clinical implications of tumor heterogeneity were well understood early on, and BC was one of the first solid tumor types to establish the clinical and treatment implications of heterogeneity for cellular phenotypes by analyzing estrogen receptor expression (Feng et al., 2018a). Intratumor heterogeneity at the functional, genetic, and cellular levels has begun to be appreciated as molecular biology and genomics techniques have advanced, and the identification of intrinsic molecular subtypes based on global gene expression profiling studies in BC has been pioneered, with a relatively rapid translation of this knowledge into clinical BC management (Perou et al., 2000; Sørlie et al., 2001).

2.5.3. Major signaling pathways in breast cancer development and progression

On the molecular level, there are striking similarities between normal development and cancer progression (Huebner & Ewald, 2014; Macias & Hinck, 2012). Normal human development is carefully regulated by complex signaling pathways that allow cells to communicate with one another and their surroundings. Not unexpectedly, cancer cells and CSCs disrupt or hijack many of the same signaling pathways. In essence, cancer is caused by genetic and epigenetic changes that allow cells to bypass the mechanisms that typically limit their proliferation, survival, and migration. Many of these changes correspond to signaling pathways controlling cell proliferation and division, cell death, cell differentiation and fate, and cell motility. Thus, activating mutations in proto-oncogenes can promote hyperactivation of certain signaling pathways, whereas inactivating tumor suppressors removes key negative signaling regulators. The primary signaling pathways that control normal mammary gland growth and BC stem cell functions include estrogen receptor (ER) signaling, HER2 signaling, and canonical Wnt signaling (Feng et al., 2018a).

*Additional/Other signaling pathways in BC***:** A number of hormones and signaling mechanisms regulate normal mammary stem cells and growth (Stingl et al., 2006). In addition to the three pathways mentioned above, several other pathways and their interactions play key roles in controlling normal mammary development and, if dysregulated, breast cancer development. These include CDKs (Cyclin dependent kinases), Notch, SHH, PI3K/Akt/mTOR, and others (Nwabo Kamdje et al., 2014).

2.6. Somatic gene mutations in breast cancer

2.6.1. Oncogenic mutations of *PIK3CA* in breast cancer

Although less common and have a less severe effect on BC risk than BRCA mutations, hereditary abnormalities in several other genes can also cause BC (Feng et al., 2018a). According to a study of early breast malignancies (Nik-zainal et al., 2019), the most frequently mutant and/or amplified genes in tumor cells include *TP53* (41%), *PIK3CA* (30%), *MYC* (20%), *PTEN* (16%), *CCND1* (16%), *ERBB2* (13%), *FGFR1* (11%), and *GATA3* (10%). These genes encode cell-cycle modulators that are either repressed (for example, p53) or activated (for example, cyclin D1), thereby sustaining proliferation and/or inhibiting apoptosis, inhibiting activated oncogenic pathways (MYC, HER2, and FGFR1), or inhibiting no longer repressed elements (*PTEN*) (Figure 7). The bulk of mutations impacting 100 suspected breast cancer drivers are exceedingly rare (Yates & Desmedt, 2017), hence most BC are generated by several, low-penetrant mutations that act cumulatively.

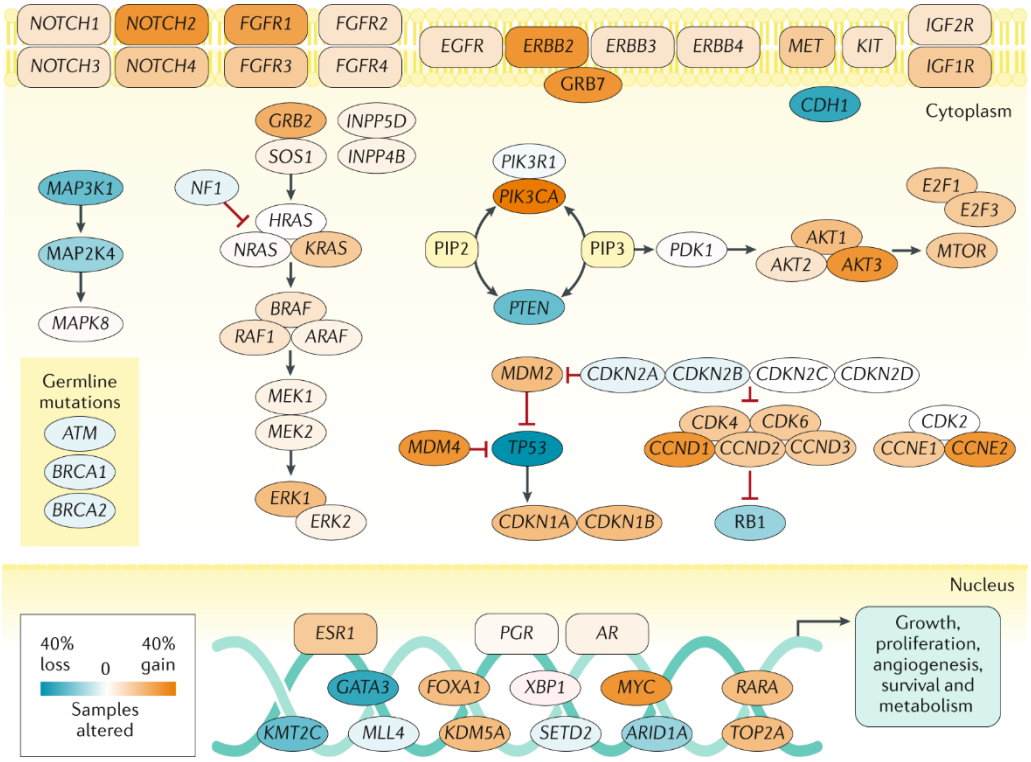


Figure 7: Molecular mutations in breast cancer. *The Cancer Genome Atlas data on breast tumor DNA copy number and somatic mutations were used to calculate the prevalence of each genetic variation in 792 breast cancer patients (all subtypes). Each gene is shaded according to the overall frequency of changes. Orange indicates a high level of amplification and/or possible gain-of-function mutations, whereas blue indicates homozygous deletions and/or potential loss-of-function mutations* (Harbeck et al., 2019)

2.6.1.1 Structure of PI3K

Of the major intracellular signaling pathways, PI3K/Akt/mTOR responds to the presence of various nutrients, metabolites including hormones and growth factor stimulation. This signaling pathway play a central role in the growth of tumor cell and proliferation (Cantley, 2002). Taking into account their structural composition and specificity to distinct substrate, PI3K is classified into 3 classes (I–III). In addition to this major classification, Class I PI3K is again classified into class IA and IB. Of the two classes under Class I, Class IA PI3K emerged as the most important and strongly associated with cancer (Figure 8). PI3K consists of a catalytic domain (p110) and a regulatory domain (p85). Three isoforms of p110 were identified: p110α (encoded by PIK3CA), p110β, and p110δ. p110δ is mostly expressed in white blood cells, while p110α and p110β are widely expressed in all types of cells (P. Liu et al., 2009).

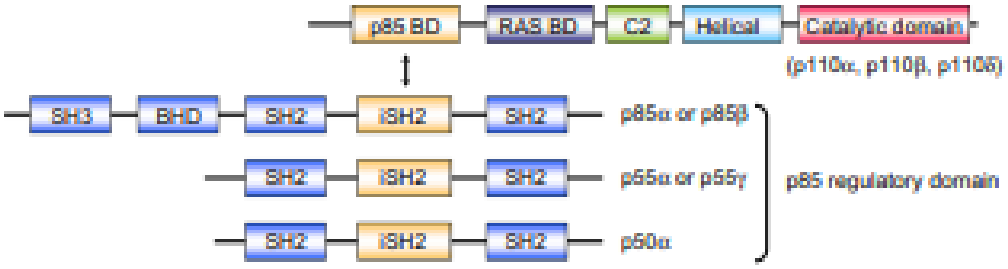


Figure 8: Structure of class IA PI3K. *Class IA PI3Ks are heterodimers made up of p110 and p85 subunits. The p110 catalytic isoforms include p110α, p110β, and p110δ. The p110 isoforms have five different domains: an amino-terminal p85-binding domain (p85 BD), a RAS-binding domain (RAS BD), a potential membrane-binding domain (C2), the helical domain, and the carboxy-terminal kinase catalytic domain. There are three p85 isoforms: p85α (with splice variants p55α and p50α), p85β, and p55γ. They share three core domains, which include a p110-binding domain known as the inter-Src homology 2 (iSH2) domain, as well as two SH2 domains. The expanded N-terminal portions of the longer isoforms, p85α and p85β, contain an SH3 domain and a breast cancer homology domain (BHD)* (Mukohara, 2015)

2.6.1.2. PI3K role in cancer

Before a signaling pathways, biological function is established, it goes through a succession of phosphorylations. In this process, PI3Ks play a critical role in mediating the phosphorylation of phosphatidylinositol 4,5 bisphosphate, or PIP2, to phosphatidylinositol 3,4,4-triphosphate, or PIP3, which then activates the phosphorylation of Akt, a serine/threonine kinase, which influences cancer cell cycling, survival, and growth (L. Zhao & Vogt, 2008a). This pathway is suppressed by the tumor suppressor protein phosphatase and tensin homolog (PTEN), which catalyzes PIP3 dephosphorylation to PIP2. Cellular PIP3 levels are thus determined by the conflict between PI3K and PTEN (Avan et al., 2016; Salmena et al., 2008).

In BC , exons 9 and 20 commonly involved and mutation in PTEN and *PIK3CA* is considered as the most frequent molecular aberrations which is recorded human malignancies (Samuels & Velculescu, 2004). As a result, the PI3K pathway can be activated in an Akt-independent manner, and Akt-independent *PIK3CA* mutations can cause cancer (Bruhn et al., 2013; W. Zhang et al., 2012). mTOR is a serine/threonine protein kinase located downstream of PI3K and Akt. The term mTOR refers to two distinct complexes, mTORC1 and mTORC2, each with a unique method of action. Rapamycin and its analogs target the mTORC1 protein. Even though mTORC1 is much better studied and described, it is currently believed that mTORC2 is blocked by these medicines at sufficient levels and impacts cell metabolism and cancer cell development (Figure 17) (Sarbassov et al., 2006; Wander et al., 2011).

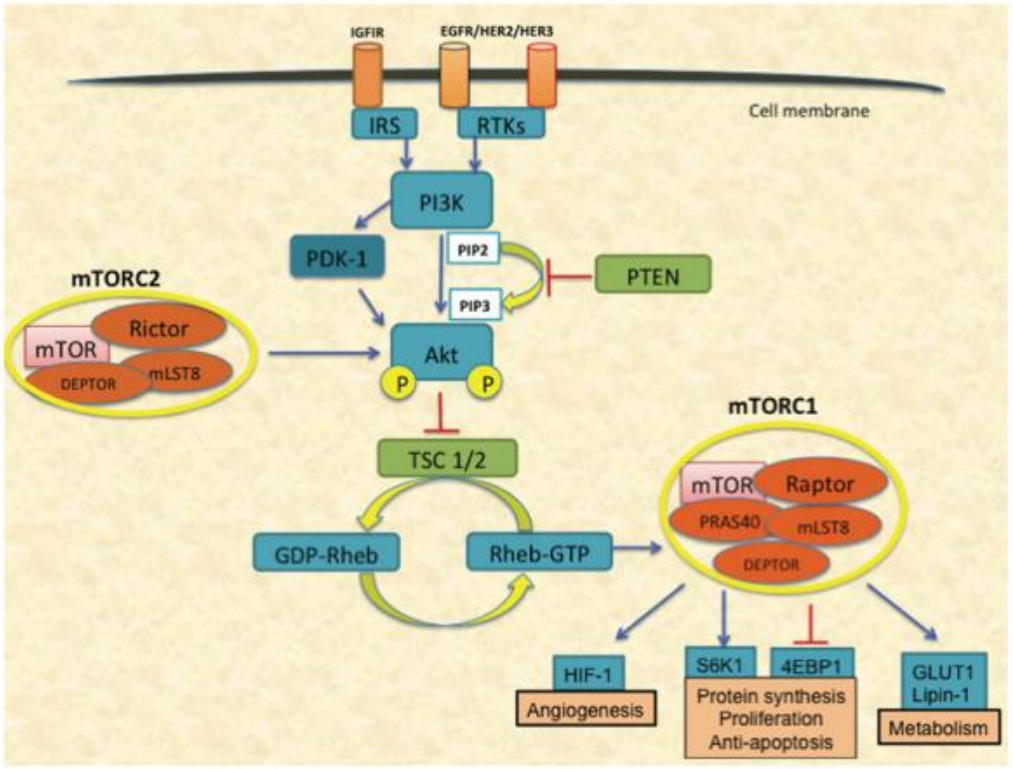


Figure 9: Illustration of the PI3K/Akt/mTOR pathway (Paplomata & Regan, 2014)

2.6.1.3. PIK3CA mutations and prognosis in breast cancer

Somatic mutations of *PIK3CA* coding p110α in solid tumors were first documented in 2004 (Samuels et al., 2004). The bulk of *PIK3CA* somatic mutations occur in two "hot spots": E542K or E545K in exon 9, and H1047R or H1047L in exon 20. Both types of mutations were found to be gain-of-function mutations with transforming capacity (Isakoff et al., 2005; J. J. Zhao et al., 2005).

Mutation/amplification, loss/mutation of the phosphatase and tensin homolog, Akt overexpression/overactivation, and modulation of tuberous sclerosis protein 1 and 2 tumor suppressors, can be often observed in HR+ BC. Different *PIK3CA* mutations affecting multiple domains of the protein have been shown in figure 10 (Fusco et al., 2021). In patients with BC , the most common alterations clustered in the helical (exon 9 p.E545K and E542K) or the kinase (exon 20 p.H1047R) domains (Samuels et al., 2004).

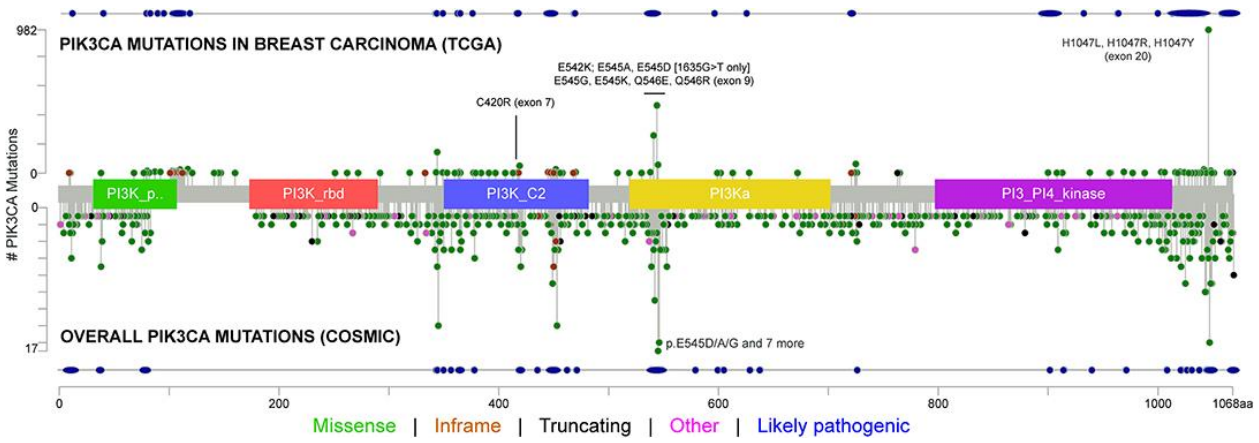


Figure 10: Type of mutations, frequency, and affected PIK3CA domains among breast tumors *from The Cancer Genome Atlas (TCGA) Network and selected solid tumors from the Catalog of Somatic Mutations in Cancer (COSMIC) datasets, which include the mutations and exons covered by the FDA-approved RT-PCR test. This analysis includes non-small cell lung cancer (squamous cell and adenocarcinoma), esophageal cancer, stomach cancer, colorectal cancer, cholangiocarcinoma, pancreatic cancer, liver cancer, bladder cancer, prostate adenocarcinoma, uterine cancer (endometrioid, serous, and carcinosarcoma), ovarian cancer, and invasive breast cancer. The types of mutations and their predicted pathogenicity are color-coded according to the legend at the bottom* (Fusco et al., 2021)

Among studies, the relationship between *PIK3CA* mutations and hormone receptor status has garnered the most attention. A meta-analysis of 26 studies revealed a substantial link between *PIK3CA* mutations and ER and PgR expressions (Pang et al., 2014). An early study found a positive connection between *PIK3CA* mutation and HER2 overexpression. However, following research found no connection between *PIK3CA* and HER2 status (Barbareschi et al., 2007; Maruyama et al., 2007). Furthermore, other studies discovered that *PIK3CA* mutations were associated with favorable prognosis, including a non-triple negative subtype (Gonzaiez-Angulo et al., 2009), luminal A subtype and lower Ki67 levels (López-Knowles et al., 2014) and smaller tumor size (Loi et al., 2013). Overall, it appears safe to claim that *PIK3CA* mutations are most likely detected in luminal-type (HR-positive/HER2-negative) cancers, especially those with markers indicating less aggressive tumor characteristics.

Previous study observed that patients with *PIK3CA*mut tumors represented 35.7%. in their finding *PIK3CA* mutation rates were lower in TNBC (16%) compared to HR+/HER2 (42%) and HER2+ (31%) BC (Martínez-Saéz et al., 2020). In another finding, *PIK3CA*-mutation rate was 26.7% and *PIK3CA*-mutations were significantly more frequent in steroid hormone-receptor (SHR)-positive HER2-negative (31.4%), and G1 and G2 tumors (32.8%) (Reinhardt et al., 2023). *PIK3CA* activating mutations, which encode the p110α catalytic subunit, are found in roughly 30% of breast cancer cells and are more common in ER+ breast cancer cells (P. Liu et al., 2009; Saal et al., 2005).

Eighty percent of *PIK3CA* mutations occur in two 'hot areas' inside exons 9 and 20 that encode the helical and kinase domains, respectively. The E542K and E545K (exon 9) mutations may lead to a gain-of-function by breaking the inhibitory connection between p110α and p85 (Miled et al., 2007). The H1047R (exon 20) mutation may cause an allosteric shift that mimics Ras-GTP binding, rendering this mutant independent of Ras-GTP interaction (L. Zhao & Vogt, 2008b).

In most previously conducted study, the clinical significance of mutations in *PIK3CA* gene among patients with newly diagnosed BC has been exhaustively determined. Such study revealed that *PIK3CA* mutations have been correlated with good prognostic features in patients with BC. These include positive expression of ER, smaller tumor size and low histological grade (Boyault et al., 2012; Cizkova et al., 2012; Kalinsky et al., 2009; Pérez-Tenorio et al., 2007). During the follow-up phase, it was observed that *PIK3CA* mutant patients had a better prognosis than wild-type patients in the first three years, but this vanished with longer follow-up. Consistent with these results, a single center retrospective cohort analysis of 590 patients revealed that mutations in the *PIK3CA* gene were strongly linked with improved clinical outcomes (Kalinsky et al., 2009). Supporting the aforementioned studies, an ample of research findings (Kalinsky et al., 2009; Maruyama et al., 2007; Pérez-Tenorio et al., 2007) documented that mutations in *PIK3CA* gene significantly and independently correlated with better recurrence-free survival (RFS). In those papers, *PIK3CA*-mutated status was associated with markers of good prognosis (Kalinsky et al., 2009). However, the studies highlighted that the clinical consequences differ according to the status of established molecular markers (ER, PgR, and HER2).

In the new era of personalized medicine, BC patients are routinely offered a "molecular diagnosis" in order to allow for tailored treatments that can potentially improve their survival outcomes. Predictors of response to targeted therapies are needed to select the patients that are more likely to respond and monitor the therapeutic benefit in real-time (Sobhani et al., 2018). Substantial investment and research have succeeded in producing a range of small molecule drugs targeting PI3K, which have undergone clinical trials in cancer patients. Such trials have recently resulted in the FDA approval of Alpelisib/BYL719, a PI3K p110α selective inhibitor for advanced BCs which are hormone receptor positive, HER2 negative and *PIK3CA* mutant (Berns et al., 2007).

2.6.1.4. PIK3CA mutations and therapy resistance in vitro

In a variety of preclinical cell line and xenograft models, mutation in *PIK3CA* gene found to be associated with resistance to HER2 and endocrine therapies. In those patient with HER2 - positive tumors, a wide range of preclinical studies demonstrated that mutations in this gene turned correlated with associated with resistance to HER2 blockade with trastuzumab (Berns et al., 2007; Kataoka et al., 2010). Another study revealed that these *PIK3CA* gene mutations might cause trastuzumab resistance, although the E545K and H1047R-HER2 overexpressing BC cell lines were responsive to GDC-0941, a pan-PI3K inhibitor (Junttila et al., 2009).

Furthermore, numerous lines of evidence suggest that PI3K signaling pathway activation has emerged as a biological modulator of endocrine resistance in patients with HR+ intrinsic subtypes, including luminal BC (Juncker-Jensen et al., 2006; Massarweh et al., 2008; Miller et al., 2010; Zardavas et al., 2012). The PI3K/Akt/mTOR pathway has also been linked to endocrine resistance. Preclinical investigations have revealed that Akt can activate the estrogen receptor (ER) pathway regardless of estrogen availability, and that combining mTOR inhibitors with endocrine therapy can overcome resistance (Paplomata et al., 2013).

Because *PIK3CA* mutations are most commonly detected in ER-positive cancers, the topic of whether these mutations cause resistance to endocrine treatments is quite interesting. In vitro investigations shown that PI3K and AKT can activate the ER in the absence of estrogen, and constitutively active AKT induces tamoxifen resistance (Campbell et al., 2001). One study found that tumors with *PIK3CA* mutations were less likely to respond to these medications than those with wild-type *PIK3CA* (Ellis et al., 2010). The aforementioned study, which included tamoxifen and observation groups, did not find a prognostic value for *PIK3CA* mutation assessed with RFS from adjuvant tamoxifen treatment (Beelen et al., 2014). Another study examining *PIK3CA* genotypes in tumors from 447 ER-positive metastatic patients who had received first-line tamoxifen treatment discovered no connection between genotypes and treatment outcome as defined by TTP (Ramirez-Ardila et al., 2013).

Less common gene mutations in BC: the less common gene mutations including A*TM*-telangiectasia, *TP5*-Li-Fraumeni syndrome, *PTEN*-Cowden syndrome, *STK11* or *LKB1*-Peutz-Jeghers syndrome, *PALB2* (Partner and Localizer of *BRCA2*), *CHEK2 and CDH1* well elaborated in previous study (Feng et al., 2018a).

2.7. Classification of breast cancer

2.7.1. Histological classification

Currently, clinical practice employs a surrogate classification of five subgroups based on histological and genetic features (Harbeck et al., 2019). Given the difficulty in obtaining gene expression data, there is a significant preference for routine histopathology testing. As a result, a group of specialists from the St. Gallen consensus conference proposed classifying (Goldhirsch et al., 2011) BCs using histopathological surrogate markers for genetic intrinsic subtypes. Over the last 10-15 years, therapy approaches have changed to account for this heterogeneity, with a focus on more physiologically guided medicines and treatment de-escalation to prevent treatment-related side effects. Despite the underlying biological variability that drives modern treatments, some characteristics, such as the impact of locoregional tumor burden or metastatic patterns, are common and influence therapy (Harbeck et al., 2019).

In addition to commonly utilized biomarkers such as the estrogen receptor (ER), progesterone receptor (PgR), and human epidermal growth factor receptor (HER2) discovered by immunohistochemistry (IHC), the application of molecular technologies may improve patient risk assessment and therapy prediction. A combination of biomarkers comprising ER, PgR, HER2, and Ki-67 has been used to determine BC treatment. These biomarkers have been used to classify BC into several molecular subtypes (Coates et al., 2015). Within the vast group of varied breast carcinomas, there are several varieties of BC characterized by their invasiveness in relation to the initial tumor locations. It is critical to distinguish between the various subtypes since they have varied prognoses and treatment options (Feng et al., 2018a).

To analyze the morphology of BC , we must first determine whether the tumor is confined to the epithelial component of the breast or has infiltrated the surrounding stroma, and then ensure that the tumor appears along mammary ducts or lobes (Vuong et al., 2014). The histological procedure, cell type, cell amount, type and location of secretion, IHC immunohistochemical profile, and architectural characteristics all contribute to determining whether the tumor is ductal or lobular (Makki, 2015; Nounou et al., 2015).

There are numerous types of BC ; however, the following are the most common: medullary carcinoma, metaplastic carcinoma, apocrine carcinoma, mucinous carcinoma, cribriform carcinoma, tubular carcinoma, neuroendocrine carcinoma, classic lobular carcinoma, and pleomorphic lobular carcinoma (Masood, 2016). Figure 11 depicts the major subtypes of invasive breast cancer.

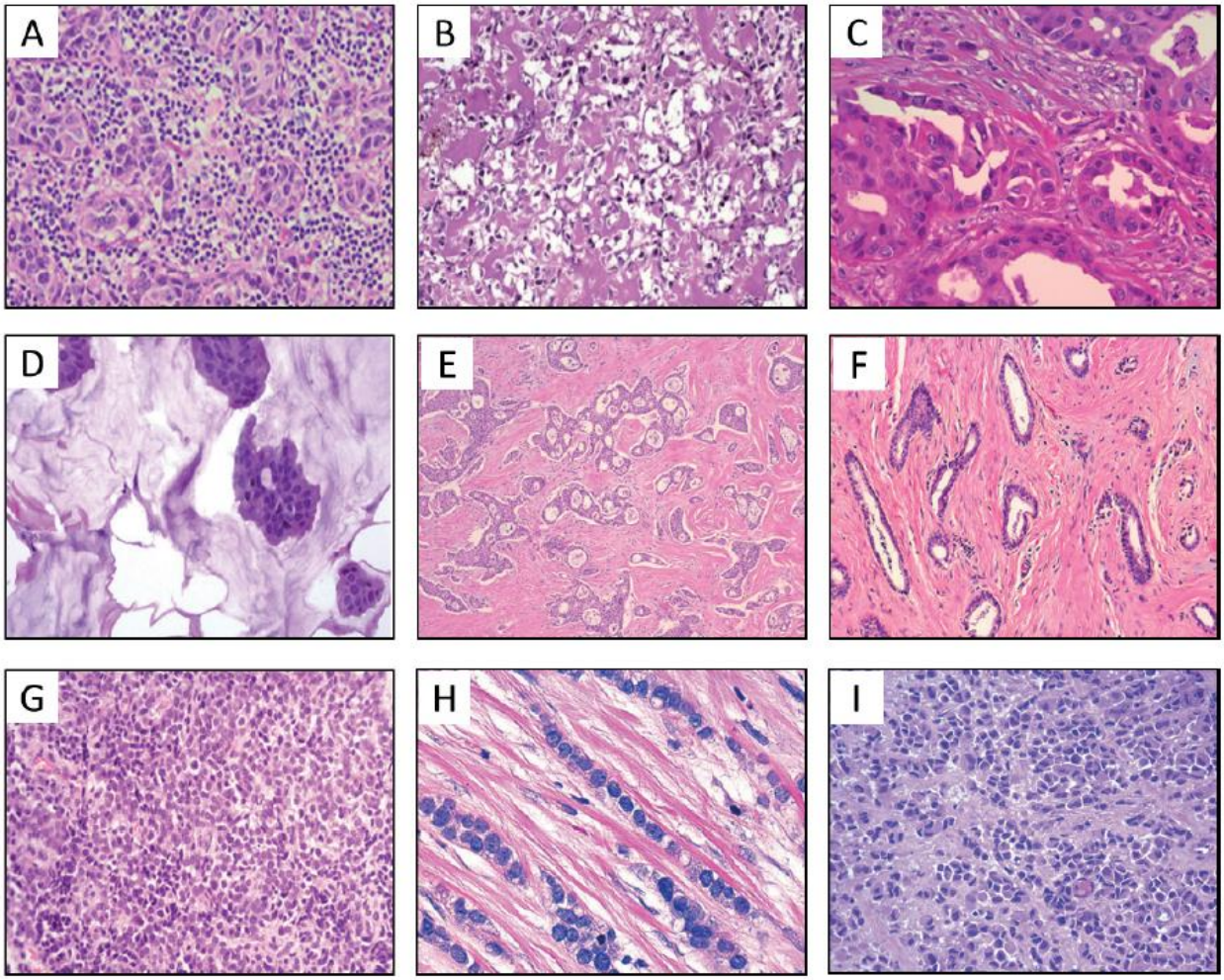


Figure 11: Morphological variants representing the primary subtypes of invasive breast carcinomas. *(A) medullary carcinoma, (B) metaplastic carcinoma, (C) apocrine carcinoma, (D) mucinous carcinoma, (E) cribriform carcinoma, (F) tubular carcinoma, (G) neuroendocrine carcinoma, (H) classic lobular carcinoma, and (I) pleomorphic lobular carcinoma* (Nascimento & Otoni, 2020)

In 2008, the American Society of Clinical Oncology (ASCO) and the College of American Pathologists (CAP) conducted an investigation with a panel of experts with the goal of developing a guideline to improve the accuracy of IHC-based ER and PR testing, as well as the utility of such receptors as predictive markers. Their research found that up to one-fifth of recent IHC studies of these markers worldwide were incorrect. Variations in preanalytical parameters, set cutoffs/thresholds to categorize receptor status as positive, and data interpretation criteria are only a few of the many elements that contribute to erroneous outcomes. Importantly, it was suggested that the assay be interpreted as positive ER when there are at least 1% positive tumor nuclei in the sample on testing, taking into account the projected reactivity of internal and external controls (Hammond et al., 2010). Given these constraints, alternative methods, like as molecular tools, for discovering the molecular biology of BC are critical.

Despite efforts to standardize these markers (especially Ki-67), lack of repeatability between laboratories remains a serious issue (Cuzick et al., 2011; Dowsett et al., 2011). There is now increasing evidence that gene expression-based subtyping is superior and more reliable than a standardized immunohistochemical classification algorithm (T. O. Nielsen et al., 2010; Wallden et al., 2015).

2.7.2. Molecular classification: gene expression based

Changes in gene expression reflect virtually any biological condition, including developmental states, cellular responses to external stimuli, and pathogenic states. The amount of related mRNA molecules in the cell is one way to quantify a gene's expression level. While a single gene's expression level cannot define biological states, the ability to monitor gene expression on a genome-wide scale using technologies like microarrays enables researchers to identify gene expression signatures linked with a certain biological phenomenon or response. Gene expression signatures may consist of tens or hundreds of genes (Kulkarni, 2011). Over the last decade, genomic and transcriptome data have been combined at unprecedented scales, revealing discrete cancer subtypes, essential molecular drivers, clonal evolutionary trajectories, and prognostic markers (Nolan et al., 2023).

The advent of RNA-based molecular profiling has profoundly influenced our understanding of BC heterogeneity and impacted patient stratification and treatment selection over the past two decades. Five primary intrinsic molecular subtypes have emerged as a result of pioneering microarray expression profiling studies (Nolan et al., 2023). Initially Perou et al. identified 476 genes from 65 BC s and matched normal tissues. Four intrinsic subtypes were recognized through hierarchical clustering, namely basal-like, Erb-B2^{+}, normal breast-like, and luminal epithelial/ER^{+} (Perou et al., 2000; Sørlie et al., 2001). Subsequent update by Sorlie et al. further subdivided the luminal subtype into luminal A and B with luminal A BCs displaying the most optimistic prognosis in terms of distant metastasis (Sørlie et al., 2003).

Global gene expression analysis is an extremely powerful molecular biology tool that is used to investigate basic biology, diagnose disease, facilitate drug discovery and development, customize therapies to specific pathologies, and produce databases of information about biological systems. As a result, expression analysis is one of the most widely used methodologies in modern biology, with over 750,000 expression datasets in the NCBI Gene Expression Omnibus (GEO) public database. Breast cancer is not a single disease; thus, it is defined by a variety of intrinsic molecular subtypes, risk factors, clinical characteristics, and treatment outcomes (Perou et al., 2000).

Despite its technical difficulties and cost, analyzing gene expression profile pattern has been shown to be a significant step in characterizing the distinct BC molecular subtypes (Perou et al., 2000). It has been observed that a thorough grasp of the human genome and gene expression analysis has allowed researchers to investigate the detailed biology of tumors over the last 15 years (Group et al., 2012; Network, 2012).

BC is a physiologically and phenotypically diverse group of diseases with varying clinical and therapeutic response patterns (Feng et al., 2018a). In this era of modern medicine, only morphological classification (nuclear grade, tubular grade, mitotic index, histological grade, and architectural characteristics) and clinical pathological parameters (tumor size, lymph node involvement, metastasis) are insufficient to predict the true behavior of breast tumor pathophysiology (Fragomeni et al., 2018; Masood, 2016). The cellular and molecular heterogeneity of breast cancers, as well as the enormous number of genes that could be involved in governing cell proliferation, death, and differentiation, highlight the need of examining several genetic alterations simultaneously. A systematic analysis of the expression patterns of hundreds of genes in tumors using cDNA microarrays, as well as their linkage to specific phenotypic variation aspects, may offer the foundation for an improved cancer taxonomy (Alizadeh et al., 2000; Hedenfalk et al., 2001; Perou et al., 2000).

BC comprises multiple biological entities characterized by heterogeneity in pathology, genomic alterations, gene expression, and the tumor microenvironment (TME), which collectively influence clinical behavior and treatment response. However, the classic parameters of histopathology, tumor size and grade, nodal involvement, and marker expression currently being used to guide treatment decisions are imperfect, particularly in the case of advanced cancers, which eventually develop resistance. Hence, there is a pressing need to better predict response to therapy and a need to improve selection of optimized therapy (Nolan et al., 2023).

Our understanding of this disease has grown considerably at the cellular, molecular, and genomic levels during the last few decades, thanks to tremendous advances in molecular biology, systems biology, and genome sciences (Feng et al., 2018b). In the last decade, numerous efforts have been made to supplement the morphological classification of breast carcinoma with molecular parameters that can provide a clearer appreciation for the heterogeneity of BC and better prediction of tumor behavior to improve therapeutic strategies (Makki, 2015).

Advances in technology and biological sciences have paved the path for genomic cancer research and the identification of possible molecular targets for cancer detection and treatment. Though it has advantages, there are numerous potential challenges associated with cancer diagnosis based on cell morphology, including invasive tissue retrieval, the inability to distinguish between clinically relevant subtypes, and the presence of avoidable and significant inter- and intra-observer variability. In response to these challenges, new genomic, transcriptomic, and proteomic methods have evolved with the goal of identifying cancer at the molecular level. This shift in analytical paradigms may lead to quicker diagnosis, better categorization, and more accurate prognoses than previous histological techniques (Elsheikh et al., 2008).

Currently, BC can be divided into multiple intrinsic subtypes based on gene expression profiles. Perou and coworkers were pioneers in defining intrinsic subtypes based on gene expression research in 2000 (Sørlie et al., 2001). BC intrinsic subtypes (IS) are biologically distinct entities, characterized by specific natural gene expression patterns (Perou et al., 2000; Sørlie et al., 2001). Gene expression profiling (GEP) has been used to predict the prognosis of BC, with the goal of identifying individuals with a favorable enough prognosis to avoid adjuvant chemotherapy (Sotiriou & Pusztai, 2009; Weigelt et al., 2010).

In recent decades, various genetic and molecular classification methods have been described with a prognostic objective. The most well-known subtypes of BC s include luminal A, luminal B, HER2-enriched, basal-like, and normal-like (Goldhirsch et al., 2011).

The first molecular hallmark of BC was discovered in 2000 by analyzing the expression of a collection of genes within the tumor that might predict the clinical outcome (Sørlie et al., 2001). Sørlie et al. created a molecular portrait' of BC using 456 cDNA clones. Tumors were classified into five intrinsic subtypes with different clinical outcomes: luminal A, luminal B, HER2 over-expression, basal, and normal-like tumors (Perou et al., 2000; Sørlie et al., 2001). The argument for such classification is that the changes in gene expression patterns between cancer subtypes reflect the malignancies' intrinsic distinctions at the molecular level (Sørlie et al., 2003).

Due to the low prognostic and predictive effectiveness of existing classifications, at the turn of the century, new approaches were sought to uncover the molecular basis for BC heterogeneity. Perou et al. used a hierarchical clustering analysis of gene expression profiling to identify molecularly defined groups of BC tissue (luminal A, luminal B, HER2-enriched, basal-like, and normal-like) with different biological and clinical characteristics (Perou et al., 2000).

This molecular categorization has been demonstrated to have prognostic value and to predict the response to chemotherapy in the context (Sørlie et al., 2003).

The intrinsic gene list included hundreds of genes thought to reflect distinct tumor characteristics. However, various problems hampered IS deployment in the clinic. First, most of the technologies used required fresh-frozen tissue; second, microarray technology is resource-intensive and time-consuming; third, the original classification based on hierarchical cluster analysis could only be applied retrospectively to sufficiently large cohorts of patients; and fourth, the entire sample-to-result process required a centralized laboratory and a controlled environment (Schettini, 2022). These factors prompted Perou and colleagues to create a therapeutically relevant assay, which was originally disclosed in 2009 (Bernard et al., 2009). First, they were able to minimize the number of intrinsic genes from over 2000 to 50 (note that fewer genes were related with lower subtype identification accuracy, particularly for non-Basal-like subtypes) (Figure 12).

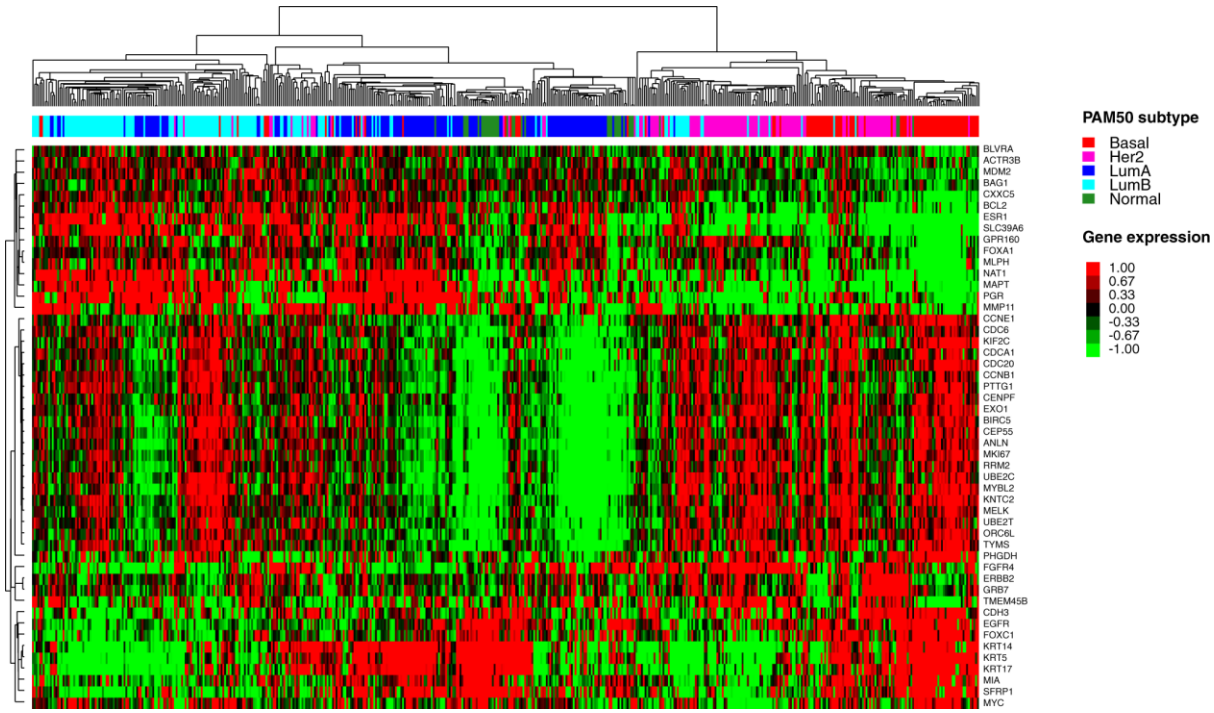


Figure 12: Heat diagram depicting the genes and expression patterns of the non-commercial PAM50 intrinsic subtypes**.** *In a set of 527 archived breast cancer fresh-frozen paraffin-embedded (FFPE) samples from Dr. Prat's laboratory, median-centered unsupervised hierarchical clustering representing the breast cancer intrinsic subtypes (Luminal A, Luminal B, HER2-Enriched, Basallike) and the Normal-like group was performed using the research-based PAM50 assay on the nCounter® platform. Each column reflects one patient's sample. The colors red, green, and black denote comparatively high, low, and median gene expression, respectively. The PAM50 gene list is shown on the right side of the heatmap. The unsupervised cluster and heatmap were generated using R 3.6.1, Cluster 3.0, and Javatreeview 1.1.6r4 for MacOSX. LumA luminal A, LumB luminal B, Her2 HER2-enhanced, Basal Basal-like, Normal Normal-like* (Prat & Perou, 2011)

Parker et al. published a study report in 2009 detailing the development of prediction analysis of microarrays (PAM) with a 50 gene set for standardizing subtype classification. When subtyping breast cancer, they used a panel of PAM50 genes and a quantitative reverse transcriptase polymerase chain reaction (qRT-PCR) technique that was validated on formalin-fixed paraffin-embedded tissues (Bernard et al., 2009; Chia et al., 2013). In 2000, Perou and Sorlie proposed the phrase "Molecular Classification" in breast cancer for the first time, based on a comprehensive assessment of gene expression differences (Perou et al., 2000). In this study, breast cancer was divided into different sub-groups according to various gene expression: "Luminal" (often differentiated in two or three subgroups; reflecting ER, ER regulatory genes, and the expression of genes expressed in normal luminal epithelial cells), "HER-2 positive" (reflecting ErbB2 / HER-2 amplification and overexpression), "basal" (reflecting ER, PR, and HER-2 negative and the expression of genes expressed in normal breast basal and / myoepitelial cells. Though gene expression overcomes the various shortcomings of molecular classification, it has several downsides.

The main constraints of gene signature profiling are problems in duplicating specific gene sets, testing costs, and reporting standards. However, gene signature cannot be replaced by IHC surrogates, which have considerable differences from genetic profiling (C. Kim et al., 2004). In summary, clinical practice now employs a surrogate classification of five subgroups based on histological and molecular features (Figure 13) (Harbeck et al., 2019).

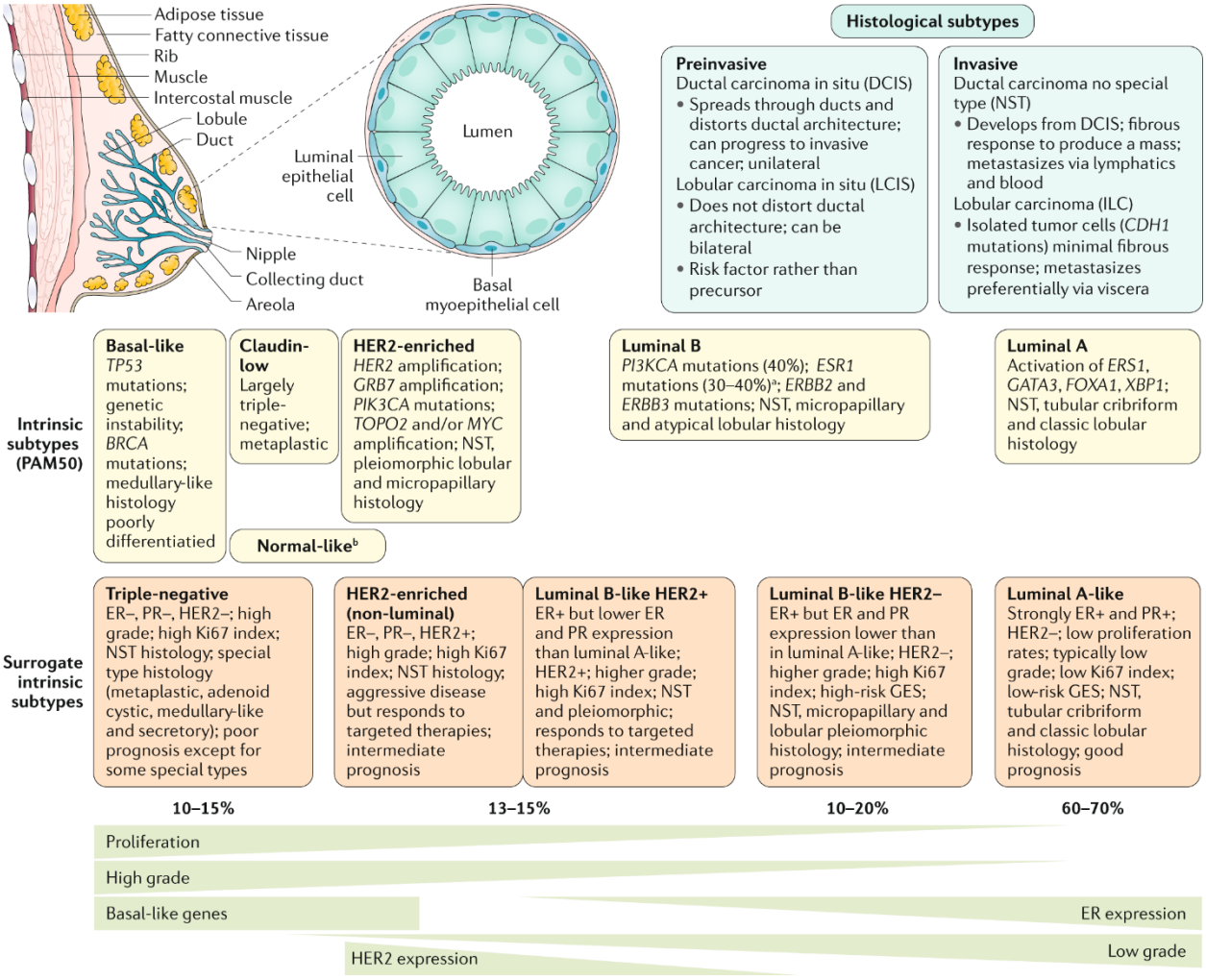


Figure 13: Surrogate classification of BC subtypes based on histology and molecular parameters. *All breast cancers develop in the terminal duct lobular units (the functional unit of the breast) of the collecting duct. Histological and molecular traits have crucial therapeutic implications, and numerous categories have been proposed based on these qualities. The histological subtypes described here (top right) are the most common subtypes of breast cancer; invasive lesions are ductal carcinoma (now referred to as 'no special type' (NST)) and lobular carcinoma; preinvasive counterparts are ductal carcinoma in situ and lobular carcinoma in situ (or lobular neoplasia). Perou and Sorlie1 have intrinsic subtypes based on a 50-gene expression profile (PAM50). Surrogate intrinsic subtypes are commonly utilized in clinical settings and are based on histological and immunohistochemical expression of important proteins such as the estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor receptor 2 (HER2), and the proliferative marker Ki67. Tumors that express ER and/or PR are referred to as 'hormone receptor-positive', while tumours that do not express ER, PR, or HER2 are referred to as 'triple-negative'. The relative positioning of the boxes corresponds to the qualities (such as proliferation and grade) in green. - indicates negative; + indicates positive. GES stands for Gene Expression Signature. A. ESR1 mutations caused by aromatase inhibitor targeted therapy. B.Artifact; expression of normal breast components as a result of reduced cancer cell density* (Harbeck et al., 2019)

2.7.2.1. Commercially available molecular assay:

The immunohistochemical (IHC) panel comprising these four biomarkers (ER/PR/HER2/Ki-67) has been deemed effective and relevant in the stratification of these molecular entities. To improve risk stratification and prognosis, multigene assays like Oncotype DX (Genomic Health, Redwood, CA, USA), Prosigna PAM50, and Mammaprint (Netherlands Cancer Institute™ and Agendia™, Netherland) have been developed and widely used in clinical practice (Nascimento & Otoni, 2020).

BC intrinsic subtypes have been identified based on the transcription of a predefined gene expression (GE) profiles and algorithm (prediction analysis of microarray 50 gene set, PAM50) (Y. J. Chen et al., 2021).The most widely accepted IS are the Luminal A, Luminal B, HER2-Enriched, and Basal-like (Prat & Perou, 2011).

Intrinsic subtypes are determined from a pre-filtered gene set which vary the most across different subjects and vary the least within the same subject such as pre-/post-chemotherapy or primary/metastatic pairs (Perou et al., 2000). The final version of intrinsic taxonomy, prediction analysis of microarray 50 gene set (PAM50) was trained from 189 BCs and 29 normal tissues and was validated from the combined dataset of 761 microarrays (Bernard et al., 2009). Constitutional 50 genes were the reduced gene set from the 1906 candidate genes of 189 BC, of which 122 were profiled by both microarray and quantitative reverse-transcriptase polymerase chain reaction (qRT-PCR). Intrinsic subtypes defined by PAM50 was advocated as being prognostic and predictive independent of clinical factors such as ER, HER2, and proliferation markers.

Using the PAM50 gene signature assay, measurement of the expression levels of 50 genes in surgically resected tissue to classify cancer into four intrinsic subtypes (Luminal A, Luminal B, HER2-enriched, and Basal-like). Such classification has a prognostic role in cancer patients who have not had adjuvant systemic therapy and have been treated with tamoxifen. In addition to identifying BC intrinsic subtype, the assay generates a personalized score estimating a patient's probability of disease recurrence by balancing the molecular subtype correlations, a subset of proliferation genes, and pathologic tumor size (Bernard et al., 2009). Perou, Sorlie, and others pioneered the use of the cDNA microarray approach to classify BC into separate subgroups based on similarities in gene expression profiles (Perou et al., 2000; Sørlie et al., 2001, 2003).

Thus, these investigations found that there are BC subgroups with distinct gene expression patterns that represent the individual phenotype, disease prognosis, and systemic therapy planning (Tsang & Tse, 2020). Genes related to the expression of estrogen receptors (ER), progesterone receptors (PR), HER2 (Human epidermal growth factor receptor 2), and cell proliferation regulator (Ki-67) are primarily responsible for the segregation of the molecular subtypes of BC (Lukong, 2017).

The PAM50 assay uses the nCounter Analysis System to simplify the workflow for local pathology labs (Prosigna™ BC Gene Signature Assay by NanoString Technologies, Seattle). For the purpose, either fresh frozen and/ or FFPE samples could be utilized. In specifically, the technique employs multiplexed gene-specific fluorescently labeled probe pairs (Geiss et al., 2008) to measure the GE profile with equal ease and efficiency. A recent clinical validation was performed using RNA isolated from over 1000 FFPE tumor specimens from the ATAC clinical study. This study found that the Prosigna risk of recurrence (ROR) score, based on the PAM50 GE signature, provided substantial prognostic information when evaluating the likelihood of distant recurrence in HR+, post-menopausal BC patients (Dowsett et al., 2013) treated with endocrine treatment alone.

NanoString Technologies represents the final significant participant in transcriptomic cancer diagnostics. The Prosigna BC Prognostic Gene Signature Assay assesses a patient's risk of BC recurrence by interrogating various pathways and analyzing the gene expression profile of 50 genes from a frozen or FFPE sample of tumor cells. This test surpasses the clinical testing service OncotypeDX because it can classify more women as high- or low-risk rather than the ambiguous "medium." Nanostring also specializes in miRNA analysis and its unique nCounter Gene Expression Code Sets, which evaluate expression levels of up to 800 genes at once with great precision, accuracy, and detection of fractional fold changes. Nanostring is the next generation of gene expression profiling firms, with technologies that can assess tissue, blood lysates, and FFPEs. The advantage of Nanostring's Prosigna is that the technology and equipment are simple enough to employ in a clinic, eliminating the need to send samples to specialized labs (Sager et al., 2015).

The platform's main benefit is that Nanostring is a relatively new gene expression profiling approach capable of producing precise genomic data from small volumes of formalin-fixed patient samples. According to several publications, the approach is used in prognostic and predictive research in chemotherapy trials (J. Lee et al., 2014; S. Liu et al., 2015) and randomized placebo-controlled investigations (Shike et al., 2014).

2.8. Intrinsic subtype as molecular biomarkers

Biomarkers are defined as "a biological molecule found in body fluids or tissues that indicates normal or abnormal circumstances, or a condition or disease, including cancer." Biomarkers come in a variety of forms, including proteins, nucleic acids, antibodies, and peptides. Interestingly, biomarkers offer a wide range of potential applications in oncology, including risk assessment, screening, differential diagnosis, prognosis prediction, therapy response prediction, and disease progression monitoring (Henry & Hayes, 2012). Molecular biomarkers have become increasingly important in detecting cancer patients with varying prognoses and predicting chemotherapy response. Molecular studies targeting these biomarkers are now routinely conducted in local pathology labs to assist with BC therapy decisions (T. Nielsen et al., 2014).

Accumulating evidence suggests that breast tumors with varied histological and biological characteristics exhibit distinct behaviors that result in different treatment responses and should be treated differently (Blows et al., 2010). Thus, precise classification of breast tumors into clinically relevant subgroups is critical for therapeutic decision making and is therefore urgently required. The intrinsic genes detected at least four distinct molecular subtypes: luminal A, luminal B, HER2-enriched, and basal-like. Luminal tumors are mostly ER-positive and are further divided into luminal-A, which are histologically low-grade, and luminal-B, which are mostly high-grade. HER2-positive cancers exhibit amplification and/or overexpression of the ERBB2 gene and are usually of high grade. Basal-like tumors, on the other hand, are ER-negative, PgR-negative, and HER2-negative (thus 'triple-negative'). These categories are quite similar with pathological classification based on ER, PgR, and HER2 status, as well as proliferation markers or histological grade (Perou et al., 2000; Sørlie et al., 2001).

BC comprises multiple biological entities characterized by heterogeneity in pathology, genomic alterations, gene expression, and the TME which collectively influence clinical behavior and treatment response. However, the classic parameters of histopathology, tumor size and grade, nodal involvement, and marker expression currently being used to guide treatment decisions are imperfect, particularly in the case of advanced cancers, which eventually develop resistance. Hence, there is a pressing need to better predict response to therapy and a need to improve selection of optimized therapy (Nolan et al., 2023). Several studies have shown that each molecular subtype has unique clinical consequences.

BC patients with the molecular subtype Luminal A had the highest overall and disease-free survival, while luminal B and HER2 enriched individuals have intermediate outcomes, and basal-like patients do the worst (Sørlie et al., 2001). In our everyday clinical pathology practice, the clinical utility of assigning invasive BC beyond standard histologic type, histologic grade, and ER/PR/HER2 status has yet to be established. The translation of molecular techniques into clinical practice to support morphological features and immunohistochemistry profiles is crucial, but it is also difficult. The subtypes of BC s identified by their gene signature are listed below (Makki, 2015).

2.8.1. Luminal breast cancer subtype

Luminal A BC's expression is similar to that of normal luminal breast epithelium. BC classified as luminal A have been shown to overexpress ER-regulated genes, under express a HER2 gene cluster, and express less proliferation-related genes. In both neoadjuvant and metastatic contexts, such tumors respond well to endocrine treatment but less well to cytotoxic drugs. Nearly 40% of all BC are categorized as luminal A tumors, which are associated with a good prognosis (Hu et al., 2006; Voduc et al., 2010).

Luminal B BC contain much lower expression of ER-related genes, variable expression of the HER2 gene cluster, and relatively high expression of proliferation-related genes. They account for around 20% of all BC s and are less prevalent than the luminal A-like subtype. These tumors have also been reported to contain genomic instability, TP53 mutations, which are associated with a higher risk of relapse and a poorer prognosis. Luminal tumors are known to be substantially more resistant to cytotoxic chemotherapy, as indicated by low pathological complete response rates following neoadjuvant chemotherapy (Carey et al., 2007; De Ronde et al., 2010; Glück et al., 2012).

2.8.2. HER2 enriched breast cancer subtype

This subtype accounts for 20–30% of all breast tumors. It is distinguished by elevated expression of HER2 proliferative genes and decreased expression of luminal genes (Perou et al., 2000). Luminal clusters contain luminal cytokeratins and other luminal-associated markers such as human endogenous retrovirus envelope PL1, X-box-binding protein 1, hepatocyte nuclear factor 3, GATA-binding protein 3, Annexin XXXI, and estrogen receptor 1, among others (Pauletti et al., 1996; Perou et al., 2000; Pollack et al., 1999). These tumours are commonly, but not always, HER2-positive and ER/PgR-negative. When compared to luminal A cancers, these tumors had a worse clinical outcome (Sørlie et al., 2003).

2.8.3. Basal-like breast cancer subtype

It accounts for roughly 15% of invasive ductal BCs. The tumor's name is derived from its common GE patterns with normal basal epithelial cells. Basal epithelial cells have a GE cluster composed of keratins 5, 6, and 17, integrin-β4, laminin, and fatty-acid binding protein 7 (Perou et al., 2000; Sørlie et al., 2001). Tumors are classified as ER-negative, PgR-negative, HER2-negative, CK5⁄6-positive, and/or EGFR (HER1)-positive by IHC (Carey et al., 2006). They are classified as ER/PgR and HER2 negative ("triple negative") because to low expression of the luminal and HER2 gene clusters. However, triple negative (TN) and basal B cells are not synonymous. Triple negative BCs are a more diverse group of disorders than basal-like BCs. Approximately 30% of TN tumors are non-basal-like (Bertucci et al., 2008). This subtype is distinguished by a high prevalence of BRCA1 (breast cancer type 1 susceptibility gene) mutations, enhanced genomic instability, strong expression of the proliferation cluster of genes, and a high histologic grade (Bayraktar & Glück, 2013).

Table 4 shows that histopathological subtypes are defined using IHC4 (ER, PgR, HER2 and ki-67) and tumor grading from G1 to G4. Again, it is obvious that molecular intrinsic subtyping is attainable by evaluating the expression of BC -related genes with gene expression signatures. Accordingly, Goldhirsch A et al described IHC and gene expression profile-based classification, including: Luminal A-like cancers (ER positive and/or PgR positive, HER2 negative, grade 1 or 2); Luminal B/HER2-negative-like tumors (ER positive and/or PgR positive, HER2 negative, grade 3), Luminal B/HER2-positive-like tumors (ER positive and/or PgR positive, HER2 positive, all grades), HER2-positive/non-luminal (ER negative and PgR negative, HER2 positive, all grades), HER2 positive(luminal) (ER positive and PgR positive, HER2 positive, all grades), and TNBC BC tumors (Goldhirsch et al., 2013).

Table 4: Breast cancer surrogate and molecular subtyping with four biomarkers (Goldhirsch et al., 2013)

Molecular subtypes	Luminal A	Luminal B		HER2 enriched	Basal-like
IHC	Luminal A-like	Luminal B-like HER2-	Luminal B –like HER2+	HER2+	Triple negative
	ER+	ER+	ER+	ER -	ER-
	PgR+/-	PgR+/-	PgR+/-	PgR- (<10%)	PgR- (<10%)
	HER2-	HER2-	HER2+	HER2+	HER2-
	G1, G2	G3	G1, G2, G3	G1, G2, G3	G1, G2, G3
	Ki-67low(<20%)	Ki-67 high(>=20%)	any Ki-67	any Ki-67	any Ki-67
Therapy	Endocrine	Endocrine and chemotherapy	Endocrine and chemotherapy and anti-HER2	Chemotherapy and anti-HER2	Chemotherapy

It has been noted that distinct molecular subtypes of BC have different clinical outcome. Accordingly, patients with luminal A and B type BC have better prognosis where as those with Basal type experience worse prognosis (Dai et al., 2015) as it is shown in figure 14.

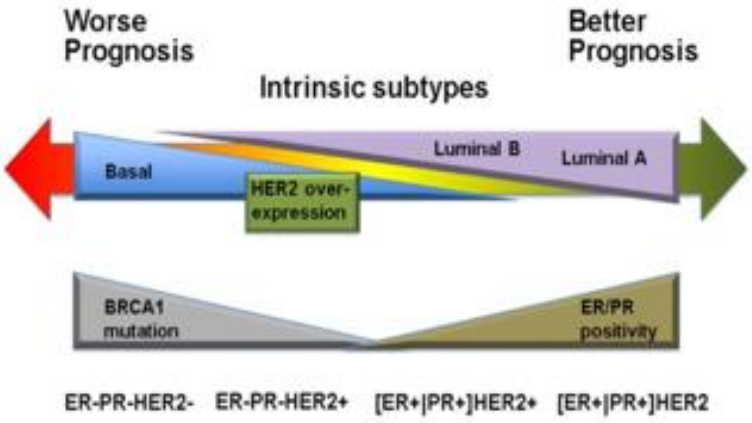


Figure 14: Comparison of molecular and pathological subtypes (Dai et al., 2015)

In Ethiopia, as in our earlier work (Desalegn et al., 2022) , the majority of the molecular subtypes were Luminal A, followed by Luminal B, with TNBC BC accounting for almost one-fourth (Hadgu et al., 2018). The median age at diagnosis differed significantly between molecular subtypes ($P < 0.05$). However, there were no statistically significant variations in tumor grade, histology, or stage among the molecular subtypes of BC (Hadgu et al., 2018).

2.9. Breast cancer diagnosis and management

Current BC diagnosis and management strategies include screening recommendations, diagnostic imaging and pathologic assessment to determine disease extent, surgery and radiation treatment, and a variety of systemic options such as chemotherapy, endocrine therapy, and targeted agents (Figure 15) (Mcdonald et al., 2016). BC is typically detected by screening or by a symptom (such as pain or a palpable tumor) that prompts a diagnostic evaluation. Screening healthy women is related with the discovery of cancers that are smaller, have a reduced risk of metastasis, are more amenable to breast-conserving and restricted axillary surgery, and require less chemotherapy (Sung et al., 2021). This situation leads to lower treatment-related morbidity and increased survival.

Screening for BC is an effective way to discover early-stage disease and increase cancer patients' survival rates. In recent decades, many developed countries have introduced population-based BC screening programs, which have helped to reduce mortality and advanced cancer rates (Ren et al., 2022).

In general, screening programs provide for earlier diagnosis, more effective treatment, and a higher chance of success (Khatib, 2006). Mammography is the only screening modality that has been proved to reduce BC -related death (Berry et al., 2005). Screening mammography reduces BC mortality by 19% on average (Pace & Keating, 2014), with women in their 40s benefiting less (15%) and women in their 60s benefiting more (32%).

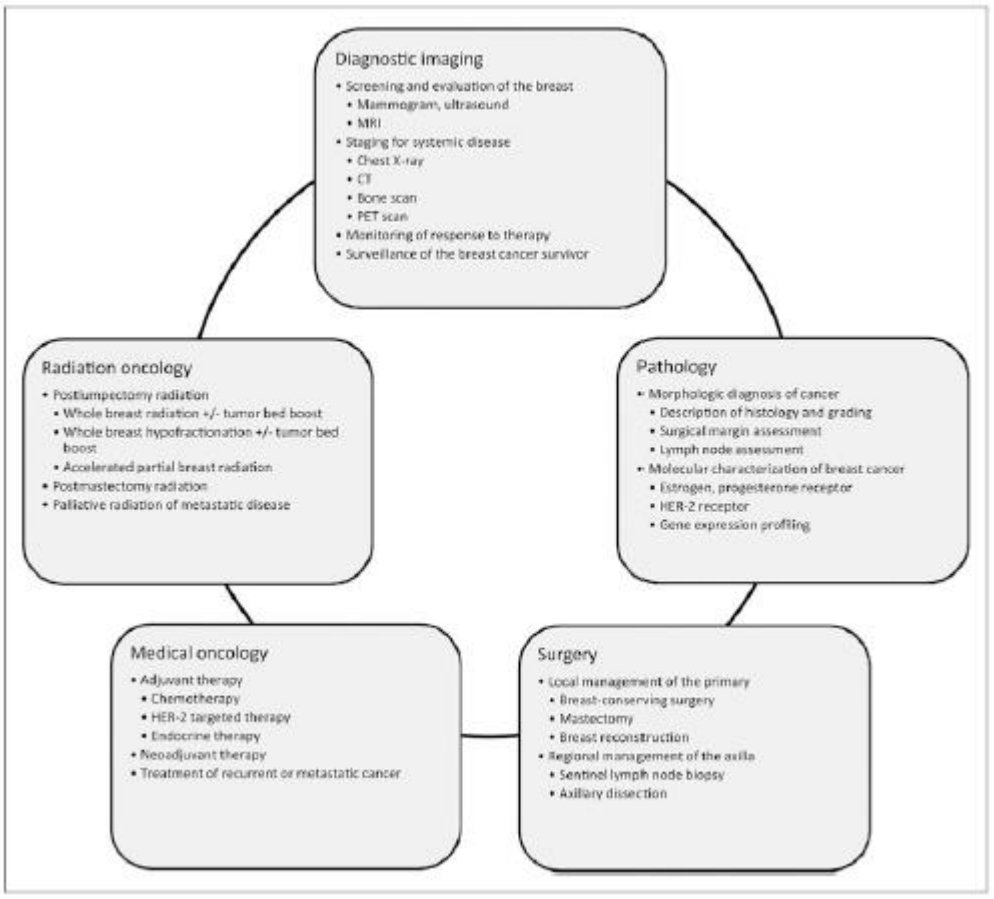


Figure 15: Breast cancer diagnostic and treatment strategy (Mcdonald et al., 2016)

The continued advancement of BC diagnosis and management has resulted in a paradigm shift away from uniform treatment regimens and toward "precision medicine" that targets the unique genetic makeup of patients and tumors. BC patients' treatment can be personalized by combining analysis of standard immunohistochemical markers and gene expression with data from anatomic imaging and targeted functional imaging studies to tailor both treatment planning and response assessment (Figure 16) (Mcdonald et al., 2016).

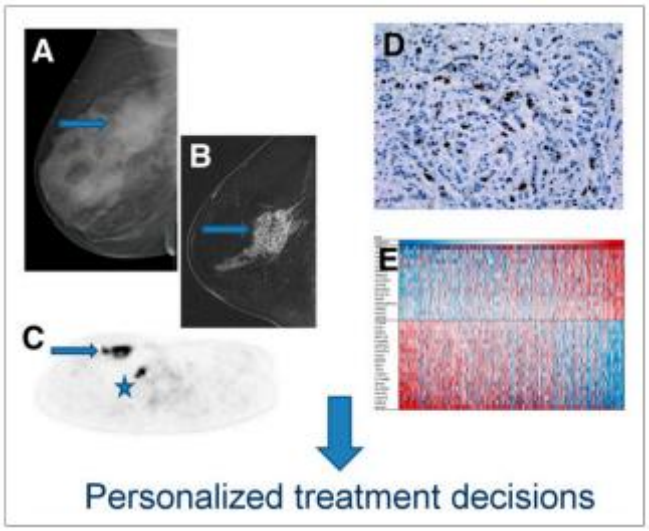


Figure 16: Precision medicine represents the future of cancer therapy. *Imaging-derived structural and functional information is linked with immunohistochemistry and genetic data to make tailored therapy options. (A) breast cancer (arrow) is poorly visible on mammography. (B) Variable gadolinium contrast agent uptake during breast MRI (arrow) suggests intratumoral blood flow heterogeneity. (C) Variable uptake of the 18F-FDG PET radiotracer (arrow) reveals variability in glucose metabolism in primary breast cancer. Extraaxillary nodal disease is also shown (*). (D) This data is paired with that from immunohistochemistry assays and mRNA Iression (E) to create a comprehensive tumor profile for therapy planning* (Mcdonald et al., 2016).

Because of low awareness of cancer signs and symptoms, insufficient screening and early detection and treatment services, insufficient diagnostic facilities, and poorly structured referral, it is difficult for the vast majority of Ethiopians to access cancer treatment services, resulting in long wait times and the progression of many potentially curable tumors to incurable stages (FMoH, 2015). Therefore, Ethiopia has developed a National Cancer Control Plan for 2016-2020, with the purpose of improving cancer prevention, early detection, improved diagnosis and treatment, palliative care, cancer monitoring, registration, and research (FMoH, 2015).

Therefore, considering cancer is not a single disease and is complicated enough to complicate care, a multifaceted approach must be implemented for better prevention and control. As a result, studying the pattern of gene expression, mutation profile, and microbiome signatures among women with BC is extremely beneficial in a variety of ways. In general, understanding the biological properties of the tumor, as well as the identification of indicators and the evaluation of microbial signatures, will be utilized as an essential parameter in estimating prognosis and selecting appropriate therapy for BC patients. Despite the prevalence of BC in Ethiopia, there is no and/or a severe lack of research-based data on the molecular biology of BC, the burden of mutation in BC -causing genes, and the involvement of microbiota in BC.

2.10. PCR as alternative diagnostic modality

Exploration of the expression of established biomarker including estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor receptor 2 (HER2) and a proliferation marker (Ki67) along with the clinical factor is essential considering their profound importance in prognosis and therapeutic implications. Again, with the advent of molecular assays, IHC based analysis has been complemented with multigene assays (Curigliano et al., 2017; L. N. Harris et al., 2016).

Considering the clinical utility of biomarker determination, a wide range of techniques are employed for the purpose. Of these, immunohistochemistry-based assessment of ER, PR, HER2, and Ki67 is an established standard of care in routine clinical practice (Allison et al., 2020; L. Harris et al., 2007; Wolff et al., 2018). Though this technique is quite important in predicting prognosis and selection of treatment which ultimately support optimal care of patients with BC, the observed quality and accuracy issue remain problematic due to intra- and inter-observer variation (Allison et al., 2020; T. O. Nielsen et al., 2021; Wolff et al., 2018).

Despite international efforts to standardize the cutoffs and the method (Goldhirsch et al., 2009; Kurosumi, 2007; Price & Goldhirsch, 2005), a complete standardization of IHC for assessment of estrogen and progesterone receptors is difficult owing to the inherent semiquantitative nature of the technology. In clinical practice, substantial problems in biomarker diagnostics in BC have been reported (Allred et al., 2008; Layfield et al., 2000). Though immunohistochemistry (IHC) on formalin-fixed paraffin-embedded (FFPE) is the gold standard for the assessment of hormonal receptor status, HER2 and Ki67 (Allegra et al., 1980; Cuzick et al., 2011; Dowsett et al., 2011), gene expression-based method applying PCR also identified as alternative approach to quantify the level of expression of gene of interest (GOI). However, its integration not yet possible in the European Society for Medical Oncology (ESMO) or ASCO/ CAP guidelines (Allison et al., 2020; L. Harris et al., 2007; T. O. Nielsen et al., 2021; Varga et al., 2017; Wolff et al., 2018; N. C. Wu et al., 2018).

2.11. Microbiota signature and its link with breast cancer

2.11.1. Definition and Overview

The terms microbiota and microbiome are often used as synonyms, but express two different meanings: microbiota refers to the totality of microbes (bacteria, archaea, fungi, viruses and protozoa) in a particular environment, in other words, it refers to taxonomy and abundance of community members, while the microbiome is the totality of genomes of a microbiota and it is often used to describe the entity of microbial traits (functions) encoded by a microbiota (Schlaeppi & Bulgarelli, 2015).

Independent organs display distinct microbiota composition; also, basic and functionally relevant inter-individual variability of microbiota was observed, indicating that they have the potential to be a determinant of disease, including cancer formation (Huttenhower et al., 2012). In support of the data, epidemiological studies have suggested that the human microbiome is a cause of several global malignancies (16-18%) (De Martel et al., 2012). According to established research, the relationship between gene and environment determines cancer susceptibility and progression. Despite their importance in health and disease, the various microbial communities that live in the gut and beyond have received little attention and have not been extensively studied until lately. However, a series of research discoveries identified microbiota as environmental variables to which humans are frequently exposed, and human microbiome investigations found a substantial difference in the relative abundance of particular microorganisms in tumor cancer patients compared to controls (Bultman, 2014).

Interestingly, body sites that have been previously considered as sterile such as the pancreas, the prostate, the lungs and the breast have been found to harvest unique microbial environments (Parida & Sharma, 2020). Given the fact that the breast has a naturally occurring nutrient rich and fatty environment and knowing that bacteria found in the skin have direct access to the mammary ducts through the nipple, it is not surprising that the breast contains a vast array of bacteria in its tissues (Urbaniak et al., 2014). In case of BC, though the association between women's microbiota and BC has been explored to a certain extent to date, it is still unclear whether the host's microbiota is causative or contributory to this disease (Fernández et al., 2018).

There are many theories regarding the origins of the breast microbiota, ranging from translocation of microbiota from the skin through the nipple, to the translocation of microbiota during lactation (Fernández et al., 2018). Whatever the origin, the breast tissue contains a high diversity of bacteria.

Using Shannon's diversity index (SDI), an index used to characterize species diversity, studies have concluded that the mammary microbiota has an average SDI of 3.6 (Chung et al., 2017). This is a relatively high diversity since the gut and oral cavity, which are considered to have a variety of bacterial communities, have an SDI of 3.9 and 6.5 respectively, while the vagina which is known to have a low diversity of bacteria has an SDI between 0.46 and 2.9 (Urbaniak et al., 2014). The knowledge that breasts tissue has its own unique microbiome has sparked interest as to its involvement in the development of BC (Fernández et al., 2018). Even though this area of research has yet to be fully explored, there have been various studies that have proved to be fruitful.

2.11.2. Microbial mechanism of oncogenesis /carcinogenesis

The role of microbes in cancer requires increased attention. About 13% of the global cancer burden is directly associated with a specific infection (de Martel et al., 2020). As a human body is composed of 10-100 trillion microorganisms and it is inevitable to assess their association with development, progression, and treat outcome of BC. More specifically, since the risk factor of BC is not fully elucidated it is high time to investigate the role of microbiota in BC. Both the human microbiome project (HMP) and integrated HMP (iHMP) projects in healthy human and non-infectious health conditions respectively had revealed unique set of microbiotas in the different locations in our body and their changes following the health conditions (Proctor et al., 2019).

It was early in 2014 that the breast is described not sterile as it was assumed (Urbaniak et al., 2014). Since then, several studies had revealed microbial community living in breast tissues are changed following the development of BC (Smith et al., 2019). Though it has been confirmed that microbiota have a role in health and disease, researchers have raised the question on how the mammary microbiome contributes in modulating the risk of BC development. It is a known fact that a difference in microbiota composition contributes to disease development and progress utilizing multiple pathways; however, there is still a dilemma and unclear whether such differences are an effect or a cause of the respective disease (Donnet-Hughes et al., 2010; Goedert et al., 2015; Urbaniak et al., 2014, 2016). On the other hand, it is up until now elusive whether there is a specific microbial signature causing a particular disease condition or mediating of breast tumorigenesis (Fernández et al., 2018).

The various ways by which microorganisms become oncogenic can be classified into four broad stages. These include direct impacts on signal transduction pathways in host cells, modulation/chronic inflammation, induction of DNA damage, alterations in host physiology, metabolite production, and effects on other microorganisms' contributions to oncogenesis/tumor suppression. As a result, these mechanisms were proposed to modify cancer susceptibility and progression; however, these processes are not exclusive because many key microbial agents are known to have several pathways pertaining to more than one process (Bhatt et al., 2017). Furthermore, microbiome may influence medication metabolism and host immune response (Louise Pouncey et al., 2018). Figure 17 demonstrates the significance of the microbiome in cancer initiation and development at various stages of cancer pathogenesis.

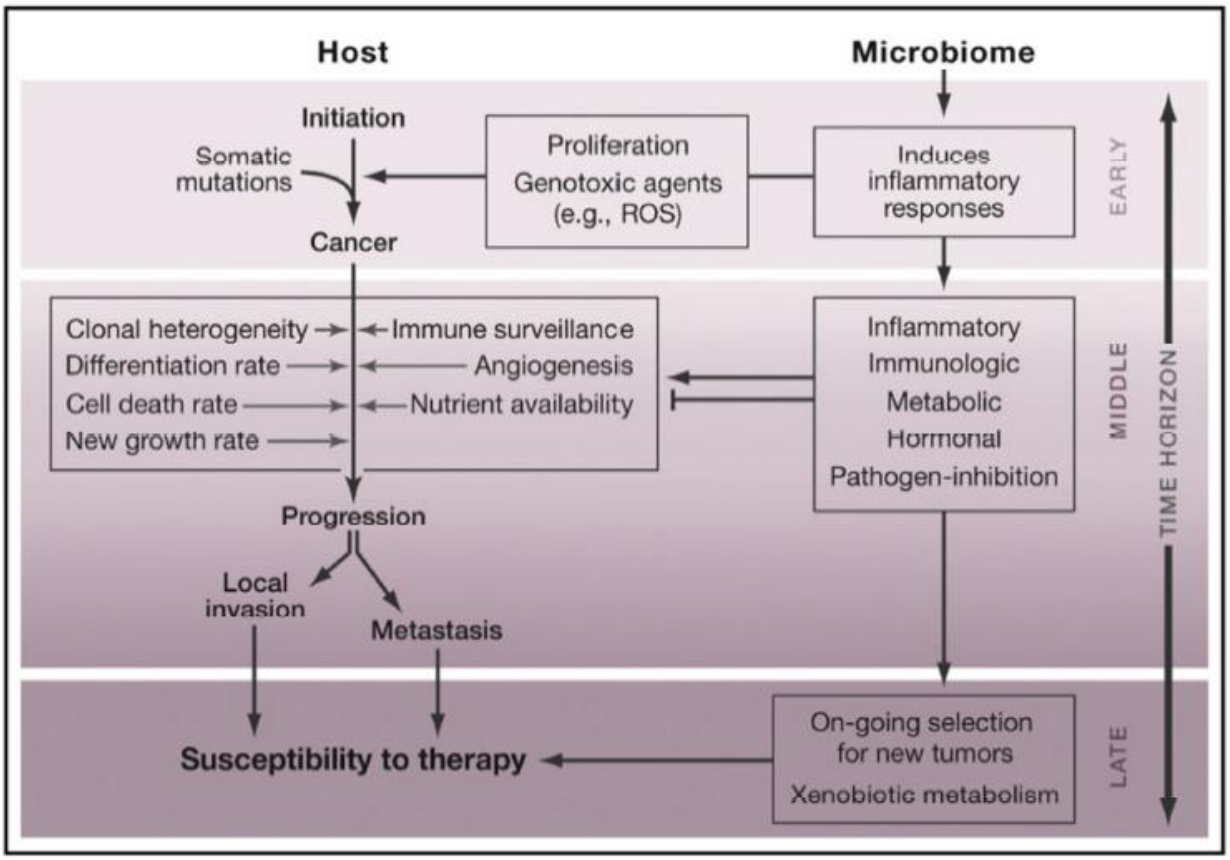


Figure 17: Mechanisms by which the microbiome can promote malignant transformation and cancer dissemination (Plottel & Blaser, 2011)

2.11.3. Gut microbiota and breast cancer

Breast cancer is associated with microbial dysbiosis in both the gut and breast tissue (J. Chen et al., 2019; Nejman et al., 2020). Supporting prior findings, dysbiosis has a role in the etiology of breast cancer. Such data revealed that women with BC have a gut microbiota with changed composition and decreased diversity when compared to healthy controls

Human gastrointestinal tract is the best investigated microbiota and is serving as a model for understanding host–microbiota interactions and disease. Main phyla of the gut microbiota includes *Firmicutes*, *Bacteriodetes*, *Actinobacteria*, *Proteobacteria*, *Fusobacteria*, *Verrucomicrobia*, *Tenericutes* and *Lentisphaerae*; and main genera are *Bacteroides*, *Clostridium*, *Faecalibacterium*, *Eubacterium*, *Ruminococcus*, *Peptococcus*, *Peptostreptococcus*, *Lactobacillus*, *Streptococcus*, *Streptomyces* and *Bifidobacterium* (Rea et al., 2018).

Because the bulk of gut bacteria are gram-negative anaerobes, more than half (50%) cannot be cultivated using standard microbiological procedures. As a result, a thorough analysis of the structural and functional complexity of the gut microbiota will be conducted using molecular techniques. Until recently, our understanding of the ecology of human-associated microorganisms was mostly based on bacterial culture. However, our ability to optimize growth conditions may restrict the effectiveness of culture-based approaches. As a result, fastidious and low-abundance organisms may not be detected utilizing culture methods. Up to 60% of organisms identified using molecular methods will not grow in normal bacterial culture media (Han et al., 2009).

Despite an explosion of knowledge into the human microbiome, these communities remain the body's dark matter. Microbial communities are difficult to research because they are so complicated. There could be hundreds of distinct species, yet enumerating what creatures are there using traditional microbiological procedures is impossible because many organisms have never been cultured and may require unusual, yet unknown, growing circumstances. Furthermore, the abundance of some bacteria might vary by orders of magnitude, necessitating extensive sampling to identify the less prevalent individuals. Next-generation sequencing (NGS) technologies were developed to enable more comprehensive investigations, including targeted 16S rRNA gene sequencing and whole-genome shotgun sequencing of bacteria in communities.

While molecular detection has increased our understanding of microbiota, many sequences remain unclassifiable, indicating a vast microbial world. Next-generation sequencing tools have revealed the microbiota's intricacy. Figure 18 depicts the paths for sample collection and microbiota analysis.

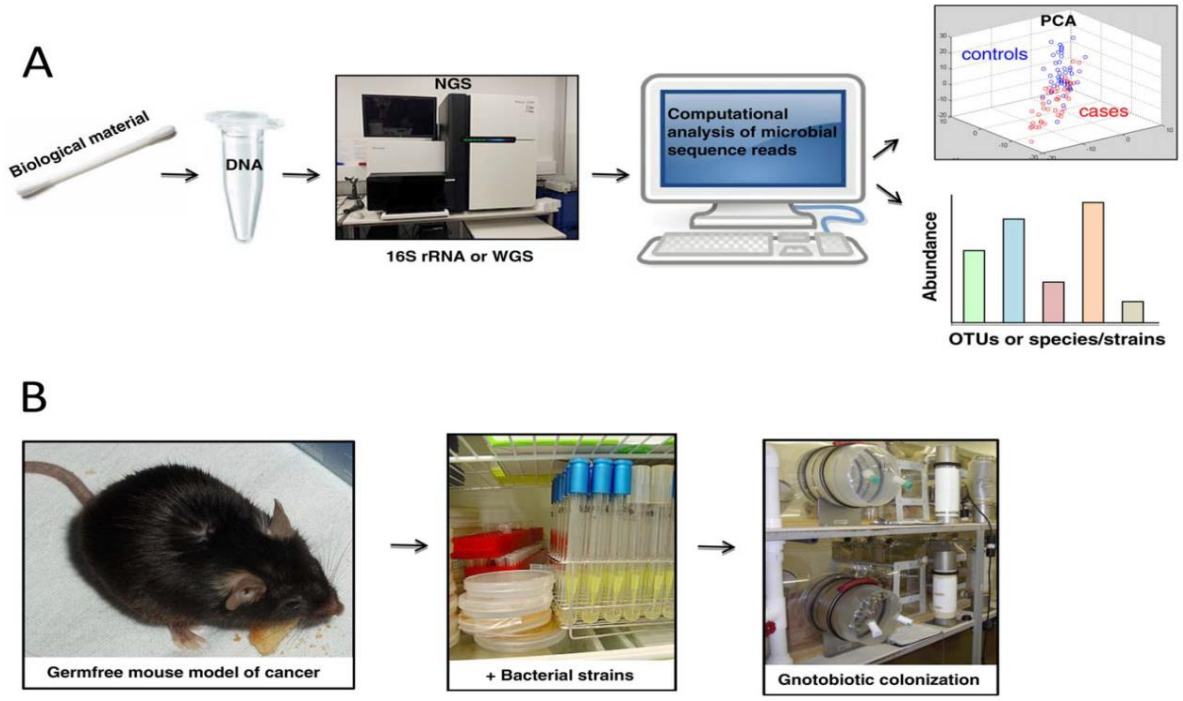


Figure 18: Sample collection process and a brief microbiome analysis strategy (Bhatt et al., 2017).

Several factors are involved in determining the composition and abundance gastrointestinal intestinal microbiota factors, including age (i.e. during diagnosis), racial disparity, diet, maternal colonization, and hygiene, as well as host genetics and environmental exposures (Keku, 2015; Turnbaugh et al., 2009). In normal circumstances, a symbiotic interaction (described as normobiosis) in between host and microbiota is essential to control a balance (homeostasis) in GIT. In many ways, such relationship helps the host o mediate various biological activities. However, any alteration of homeostasis content that interrupt this connection, the so-called “dysbiosis”, may cause bad consequences including promote the development of diseases (Schwabe & Jobin, 2013).

A series of research works were carried out and they showed the composition of microbiota and its link with BC development and progression (Fuhrman et al., 2014; Goedert et al., 2015; Hieken et al., 2016; Luu et al., 2017; Urbaniak et al., 2014, 2016; Xuan et al., 2014).

A pilot study concentrating on profiling gut microbiota in a group of postmenopausal patients, evaluated BC case patients (n=48), pretreatment vs control patients (n=48). To investigate microbial patterns in fecal DNA, an Illumina sequencing technology was used. A pilot study found that postmenopausal women with BC have altered gut microbiota composition (β-diversity) and lower α-diversity compared to control patients. However, the impact of this diversity on cancer risk and prognosis is unknown (Goedert et al., 2015).

Though quite a large number and diversity of microbiotas have been identified by various researchers, some works indicated that certain bacterial orders overrepresented in BC intestinal microbiome. And largely *Clostridiales* and *Bacteriodes* were the predominant organisms which do have an assumed estrogen metabolism potential and have been shown to co-associate with metabolite to parent estrogen ratios (Fuhrman et al., 2014). A study among 60 healthy postmenopausal women revealed that the composition and diversity of the gut microbiota were associated with patterns of estrogen metabolism. Accordingly, the study demonstrated that the relative abundances of the order *Clostridiales* and the genus *Bacteroides* were directly and inversely related with the ratio estrogen metabolites to estrogen parents, respectively; however, associations were independent of age, body mass index, and other study design factors (Fuhrman et al., 2014). Studies that focus on gut microbiota and breast cancer partially confirm correlations between specific intestine bacterial orders and breast cancer. However, the research published to yet are tiny and cannot draw any meaningful conclusions (J. Yang et al., 2017).

2.11.4. Mammary microbiota and breast cancer

2.11.4.1. Microbiota in normal, NAT and breast tumor

In women, cancer in the breast is the leading malignancy (Sung et al., 2021). Cancer in general is a complex disease, where a multitude of genomic and physiological alterations occur incessantly in the tumor tissue, contributing to the complexity of disease treatment and management. The precise etiology for BC is still unknown, but the combination of genetic, epigenetic, and environmental factors has been identified. Nevertheless, those genetic-epigenetic determinants and well-established risk factors only could explain a limited amount of the global burden of this disease (Barnes et al., 2011). As a result, the origin of a vast majority of BC cases (estimated to be as high as 70%) still remains unknown (Lakritz et al., 2015). However, since BC is linked to many modifiable risk factors and affects many women worldwide, it has become a growing priority to identify new biomarkers to facilitate the screening and management of BC (Fernández et al., 2018).

The environmental factors that influence microbiota have not been thoroughly investigated (Plaza-DÍaz et al., 2019; N. Wang et al., 2021). There is a remarkable symbiotic connection between people and various microbiota (Paul et al., 2015); yet, a change in microbiota that results in "disequilibrium" may facilitate the development of multiple disorders, including cancer (Xue et al., 2018). A variety of factors contribute significantly to the microbiota's composition. Nutritional habits, environmental factors, hereditary characteristics, drugs, and infections are identified as contributing factors to their composition (El-Sayed et al., 2021).

The human breast is not sterile but contains a diverse and unique community of bacteria, distinct from that found at other body sites, regardless of the location sampled within the breast, age, country of origin, history of pregnancy, presence/absence of breast malignancy, and method of DNA preparation and sequencing technologies (Urbaniak et al., 2014; Xuan et al., 2014). BC tissue is rich in microbiota (Hieken et al., 2016; Urbaniak et al., 2014, 2016). This microbiome, which is enriched in the breast environment and the tumor microenvironment, may alter the consequences of carcinogenesis and therapeutic interventions in breast tissue (N. Wang et al., 2021). Previous research revealed that different bacteria had covered mammary glands. One of the first published research on the microbiota of BC tissues used next-generation sequencing (NGS) on DNA isolated from breast tumor tissue and paired normal neighboring tissue from the same patient (Lakritz et al., 2015; Xuan et al., 2014) to identify a connection between microbial dysbiosis and BC .

Breast tissue contains a distinct microbiome, and emerging data suggests a relationship between an imbalance in the composition and number of bacteria (Smith et al., 2019), as well as the possibility that particular bacteria may be associated with different therapeutic responses (Fernández et al., 2018). However, just a few research have looked into the composition of microbiota in breast tumor tissue and NAT in women, with disputed or conflicting findings in both groups. Though efforts have been made in recent decades, such studies are still in their early stages, and the evidence on whether a difference exists between tumor and nearby histologically normal tissue from cancer patients is unclear, necessitating further inquiry (Fernández et al., 2018). Relatively, the identified Shannon diversity index (SDI) averages tended 3.6 which is more or less comparable to GIT and oral cavity; however, the diversity is by far higher than observed in the vagina. Such findings confirmed that mammary glands are inhabited by a diverse population of microbiota (Lakritz et al., 2015; Xuan et al., 2014) although their engagement in BC remains at its infancy. Therefore, microbial dysbiosis seems to be a hallmark of human BC (Shapira et al., 2013; Xuan et al., 2014). Therefore, it is crucial to understand how these sporadic BC s arise in order to develop preventative strategies against this devastating disease.

A study was conducted to compare the microbiota composition of tumors and normal adjacent to tumors in estrogen receptor (ER)-positive BC patients (n=20) with tissue from healthy donors. The study found a 10-fold drop in the absolute numbers of bacteria in cancer versus normal tissues. The researchers evaluated the compositional abundance of microbiota species in tumors to control tissues. This study found that *M. radiotolerans* appeared more frequently in breast tumor tissue, but *Sphingomonas yanoikuyae* numbers decreased in the tumor. However, in normal tissue, the opposite was true; the species *Sphingomonas* was shown to be more common. In summary, their findings suggest that breast cancers have a different microbial composition than normal tissue (Reddy, 1978).

Interestingly, while *S. yanoikuyae* was detected in almost half of normal adjacent tissue samples, it was not detected at all in the paired corresponding tumor tissues, whereas *M. radiotolerans* was detected in both the normal adjacent and the tumor tissue samples. This observation suggests that *S. yanoikuyae* may have a protective function in the breast, potentially due to its ability to express certain ligands that activate important cancer immuno-surveillance mediators (Xuan et al., 2014).

A finding from a BC stage study found that patients with grade III cancer had a borderline significantly higher absolute count of *Blautia sp*. than patients with grade I cancer (P = 0.048). Furthermore, absolute numbers of *Bifdobacterium* and *Blautia*, as well as *Faecalibacterium prausnitzii* and *Blautia*, varied according to clinical stage (Bard et al., 2015; Luu et al., 2017).

Another study found that *Proteobacteria* and *Firmicutes* were the most common phyla found in breast tissue where there is a change in the composition and abundance of microbiota between healthy and breast tumor (Smith et al., 2019). There is also a good relationship between BC and taxa of lesser abundance, such as genera *Fusobacterium*, *Atopobium*, *Gluconacetobacter*, *Hydrogenophaga*, and *Lactobacillus* (Hieken et al., 2016). Another finding indicated that a higher relative abundance of *Bacillus*, *Enterobacteriaceae*, and *Staphylococcus* appreciated. Again, the bacteria *Escherichia coli* and *Staphylococcus epidermidis*, all identified among patients with BC where they initiate DNA double- stranded breaks in HeLa cells (Urbaniak et al., 2016).

A study in Canada and Ireland examined the presence of a microbiome in the mammary glands of 81 cancer-free women. Bacteria were found in tissue from several breast regions in women aged 18 to 90, regardless of breastfeeding history. The main phylum was *Proteobacteria*. *Bacillus* (11.4%), *Acinetobacter* (10.0%), *Enterobacteriaceae* (8.3%), *Pseudomonas* (6.5%), *Staphylococcus* (6.5%), *Propionibacterium* (5.8%), *Comamonadaceae* (5.7%), *Gammaproteobacteria* (5.0%), and *Prevotella* (5.0%) were the most common species among the Canadian samples. The most common taxa in Irish samples were *Enterobacteriaceae* (30.8%), *Staphylococcus* (12.7%), *Listeriawelshimeri* (12.1%), *Propionibacterium* (10.1%), and *Pseudomonas* (5.3%) (Urbaniak et al., 2014).

Smith et al. and colleagues discovered that the microbial composition of normal, normal pair, and tumor tissues differed by phylum, class, and family. At the phylum level, *Proteobacteria* (classes *Betaproteobacteria*, *Alphaproteobacteria*, and *Gammaproteobacteria*) predominated in the breast, followed by *Firmicutes*, *Bacteroidetes*, and *Actinobacteria*. *Oxalobacteraceae*, a family in the *Proteobacteria* phylum, dominated normal, normal pair, and malignant tissues. Furthermore, the relative abundance of the *Pseudomonadaceae* family (phylum *Proteobacteria*), a bacterium associated with antibiotic resistance, was higher in normal pair and tumor breast tissues than in normal tissues. In keeping with earlier investigations, Class *Clostridia*, *Bacteroidia*, WPS_2, and family *Ruminococcaceae* were most numerous in tumor samples (Figure 19) (Smith et al., 2019).

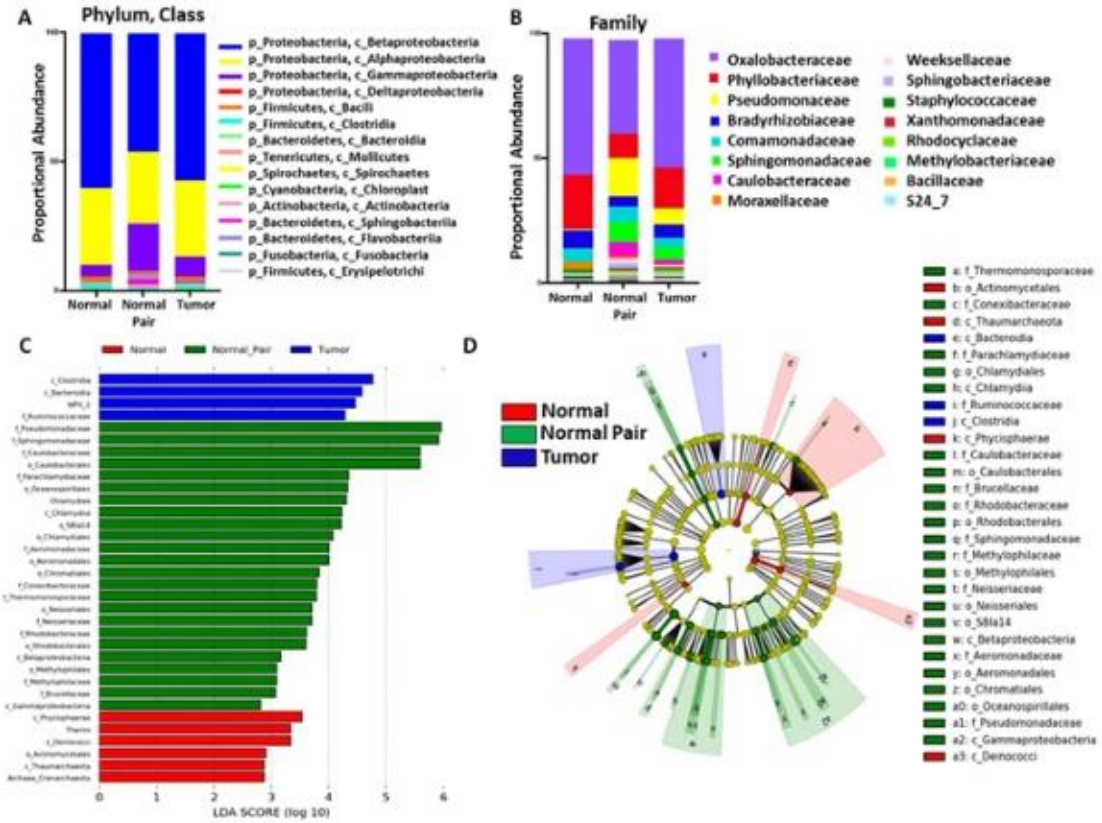


Figure 19: Breast microbiome distribution in normal, normal-paired, and tumor breast tissues (Smith et al., 2019)

The Shannon index between normal, normal pair, and tumor disease states was significantly different. Normal and normal pair breast tissues had significantly higher alpha diversity as assessed by Richness ($p = 0.017$), Chao1 ($p = 0.0021$), and Fisher's alpha ($p = 0.00087$) metrics, compared with breast tumor tissue. Differences in beta diversity across the group indicated that normal samples clustered significantly different than tumor samples (Adonis: $R^2 = 0.039$, $p = 0.002$, 999 permutations). Normal and adjacent normal tissue samples displayed greater dissimilarity along PC2 (5.54%) compared with breast tumor tissues (Smith et al., 2019).

In general, the TME is made up of a variety of cell types, including microbiota, and data suggest that pathologic changes in these could have a substantial impact on cancer formation (Thompson et al., 2017). With current knowledge, alterations in the microbiota of the mammary and gut tissue are associated with the development of BC (Urbaniak et al., 2014).

It is therefore important for more research to be carried out to identify the bacteria and molecular mechanisms involved in breast carcinogenesis. With today ground breaking scientific evidences, the microbiota is increasingly getting significance and becoming a factor correlated to the development of cancer and determine response to anti-cancer therapies (Zitvogel et al., 2016).

2.11.4.2. Microbiota versus patient and/or tumor features relation

Multiple bacterial genera showed significant differences in relative abundance when stratified by breast tissue type (tumor, tumor adjacent normal, high-risk, healthy control), cancer stage, grade, histologic subtype, receptor status, lymphovascular invasion, or node-positive status, even after controlling for confounding variables (Tzeng et al., 2021). The microbiota of ER+ malignant tumors compared to healthy tissues. Xuan et al (Xuan et al., 2014) investigated the potential role of breast microbiota on ER+ BC by comparing the DNA of the micro- biota found in the breast tumor tissue compared to the normal adjacent tissue from the same patients and healthy women that underwent reduction mammoplasty. *Sphingomonas yanoikuyae* was the most abundant bacterium found in normal adjacent breast tissue, while the bacterium *Methylobacterium radiotolerans* was abundant in ER+ breast tumor tissues (Xuan et al., 2014).

The microbiota in ER+ and PR+ malignant tumors compared to ER- PR- tumors. Wang et al (Parida & Sharma, 2020) observed that the microbial diversity and abundance was higher in hormone-positive BC than in hormone-negative BC. This could be due to the ability of some microbes to metabolize and regulate the bioavailability of steroid hormones in the former (Parida & Sharma, 2020; H. Wang et al., 2017). Additionally, they found *Methylbacterium* to be significantly decreased in hormone-positive BC when compared to hormone-negative BC (Parida & Sharma, 2020).

A study by Banerjee et al (21) was carried out to investigate the specific microbial signatures in 4 different BC subtypes i.e., HER-2-positive (HER-2+), hormone receptor-positive (PR+/ER+), triple-negative (PR-, ER-, HER-2-) and triple- positive (PR+, ER+, HER-2+) BC. When microbiota of HER-2-positive (HER-2+), hormone receptor-positive (PR+ /ER+), triple-negative (PR-, ER-, HER-2-) and triple-positive (PR+, ER+, HER-2+) BC compared, *Sphyngomonas and Mycobacterium* microbiota found in all four BC subtypes whereas higher levels of *Brevundiomana* noted in PR+ /ER+ and triple-positive when compared BC to PR- /ER- and triple-negative subtypes (Banerjee et al., 2018)..

Increased levels of *Mycobacterium* bacterium identified in PR- and ER- BC identified. Additionally, they found that each subtype had a unique microbial signature; Triple-negative had the least complex microbial signature, while PR+ /ER+ was associated with the most complex microbial signature (Banerjee et al., 2018).

In another study, microbial differences examination using LEfSe analysis aimed at identifying significant breast tumor subtype-associated biomarkers revealed that TNBC was more abundant in genus *Streptococcaceae*, and *Ruminococcus* (both phylum *Firmicutes*) (LDA > 3.5). Luminal B tumors were most abundant in genus *Clostridium* (phylum Firmicutes) whereas Luminal A tumors were most abundant in order Xanthomonadales (phylum *Proteobacteria*) (LDA > 5). HER2 tumors were abundant in genus *Akkermasia* (phylum *Verrucomicrobia*) (LDA = 4). A Spearman heatmap demonstrated phyla level changes across HER2, luminal A, luminal B, and TNBC tumor subtypes where luminal subtypes demonstrated greater *Tenericutes*, *Proteobacteria*, and *Planctomycetes* phyla. HER2 breast tumors demonstrated greatest abundance of phyla Thermi and *Verrucomicrobia* while TNBC tumors demonstrated the highest total abundance of phyla *Euryarchaeota*, *Cyanobacteria*, and *Firmicutes* (Smith et al., 2019).

In general, with all the limitation of the previous studies including the low number of participants and recruiting women who underwent cosmetic surgeries as healthy controls, identifying BC subtype signatures through microbiota composition could pave way for potential BC screening techniques. Nevertheless, it is evident that investigations regarding microbiota residing in normal breast tissue and its role in BC development is still in its infancy, but can potentially have a huge role in developing screening and therapy techniques for BC (Toumazi et al., 2021).

2.11.5. Microbiota role in breast cancer

2.11.5.1. Prognostic and predictive role

The microbiome's modulatory properties could be utilized to a variety of healthcare interventions. According to reports, microbiota could be considered a complementary prognostic or predictive biomarker for treatment results in BC by delving deeply into compositional abundance (Gopalakrishnan et al., 2018). Previous research has demonstrated that the microbiome can assist improve therapeutic efficacy, reduce drug toxicity, and prevent cancer (Picardo et al., 2019). Other research have found that microbiota can be used to diagnose, predict risk and illness progression, and prevent disease (Ehrlich, 2016).

Prospective clinical investigations to validate relationships between microbiota compositions and their involvement in predicting treatment response would be critical in determining the efficacy of these techniques (Fessler et al., 2019).

Therefore, in the near future, through investigation of the composition microbiome, one parameter may be incorporated with other known correlates of outcome such as T cell infiltration and tumor mutational burden to 1) predict potential efficacy with a given immunotherapy and 2) inform additional interventions via the microbiota to improve immunotherapy potency or alternatively reduce treatment-related toxicity (Behrouzi et al., 2019).

In the future, along with other BC molecular markers, it will be vital to consider the microbiome as one of many parameters to be added into considerations for personalized cancer therapy (Behrouzi et al., 2019).

2.11.5.2. Microbiota as a next generation medicine

Due to the newly established role that the microbiota potentially plays in the development of BC, it is important to explore how interventions that affect the composition of the microbiome (e.g., via the use of probiotics) may affect the process of breast carcinogenesis. The most common micro- organisms used in probiotics are lactic acid bacteria (LAB) which have been shown to have multiple health benefits for the host. Various other studies have been conducted both in vitro and in vivo demonstrating the effects of probiotics on BC (Ali et al., 2015).

A collection of invitro studies using cell lines were conducted to demonstrate role of probiotic strain on BC. The findings showed that the microbiota found to increase apoptosis, decrease in cell proliferation, increased cytotoxic effects on TNBC cells, a greater downregulation of hypoxia-associated genes, the downregulation of the NF-κB pathway lead to apoptosis of ER-negative BC cell lines, inhibition of growth of BC cell lines (Bharti et al., 2015; Esfandiary et al., 2016; Hassan et al., 2016; Kadirareddy et al., 2016).

Similar to those invitro studies, studies on animal model revealed that some bacteria are key players in BC. The specific phenotypic strain identified as increasing or decreasing distinct cytokines, induction of cellular apoptosis, delayed BC development, delay in onset of pre-neoplastic features, activates CD4+ and CD25+ cells, ↓growth rate of tumors, ↑survival compared with the control, improved disease prognosis in mice with highly metastatic breast tumors, ↓cell mitosis, inhibition of inflammation in tumor environment, modulation of immune system and regulation of angiogenesis proteins and genes (Toumazi et al., 2021).

Though efforts have been made so far to understanding probiotic role of some microbiota in BC, it is important to note that research on the effects of probiotics on BC is still at its infancy and larger and more elaborate studies should be conducted to conclude if they have a protective long-term anticancer effects and to find the optimum dosage, strain and regimen for a treatment with probiotics to be effective.

Taken together, research explores new dimensions of the impact of the microbiome and evaluates the differences in patients and healthy individuals for finding biomarker patterns, but insufficient information is available, and further research in this subject appears to be very valuable. On the other hand, new therapeutic approaches known as personalized medicine have opened a new window in medical science, and the link between microbiome and personalized medicine appears to be one of the most intriguing aspects of subsequent research, with a pivotal impact on the treatment of diseases such as cancer. As a result, considering the novelty of the interaction between these two axes, there are few studies in this area. Specific strains may be able to regulate cancer progression. Therefore, timely and wide-scale study related to this topic is relevant to explore the role of microbiota in breast cancer development in the context of Ethiopian patients with breast cancer.

3. RESEARCH QUESTIONS AND OBJECTIVES

3.1. Research questions

1. Does the breast microbiota abundance and composition among breast cancer patients in Ethiopia differ from other geographical location?
2. Is there a difference in microbial abundance and diversity among paired normal adjacent and tumor tissue from breast cancer patients of Ethiopia?
3. Does the microbiota abundance and diversity level vary per clinical and histopathological parameters among breast cancer patients of Ethiopia?
4. What is the distribution of intrinsic breast cancer subtypes among breast cancer patients in Ethiopia? And how it is in line with IHC and clinical and histopathological parameters?
5. Is there a difference in the prevalence of *PIK3CA* gene mutation among breast cancer patients in Ethiopia when compared with other countries? What is its relation with survival?
6. Does the performance of mRNA-based technique (End point PCR, qPCR) comparable with IHC and intrinsic subtypes for the determination of surface receptors including ER, PR and HER2?

3.2. Research Objectives

3.2.1. General objectives

The objective of this study was to determine the intrinsic subtypes, microbiome profile, *PIK3CA* gene mutations and to evaluate the performance of PCR based determination of breast cancer subtypes using IHC as gold standard among breast cancer patients in Ethiopia.

3.2.2. Specific objectives

- To determine PAM50 based intrinsic subtypes and correlation with clinical and histopathological features among patients with breast cancer
- To profile microbiota signatures from paired breast tumor and normal adjacent tissue among patients with breast cancer
- To correlate the microbiota signatures with clinical and histopathological characteristics among patients with breast cancer
- To evaluate the prevalence of hotspot mutations in *PIK3CA* gene among patients with breast cancer
- To determine the prognostic role of *PIK3CA* mutations
- To compare the performance of mRNA-based technique (Endpoint PCR, qPCR) in relation to IHC technique to determine ER, PR and HER2 among patients with breast cancer

4. MATERIALS AND METHODS

4.1. Study setting

This study was conducted in selected public and private hospital in Addis Ababa where large number of BC surgery is performed and one rural hospital that performed a large number of BC surgery. The hospitals include Tikur Anbessa Specialized Hospital, Zewditu Memorial Hospital, Yekatit-12 Hospital Medical College, Bethzatha hospital from Addis Ababa, and Aira from rural hospitals in Western Ethiopia. Area map of the study setting (Figure 20).

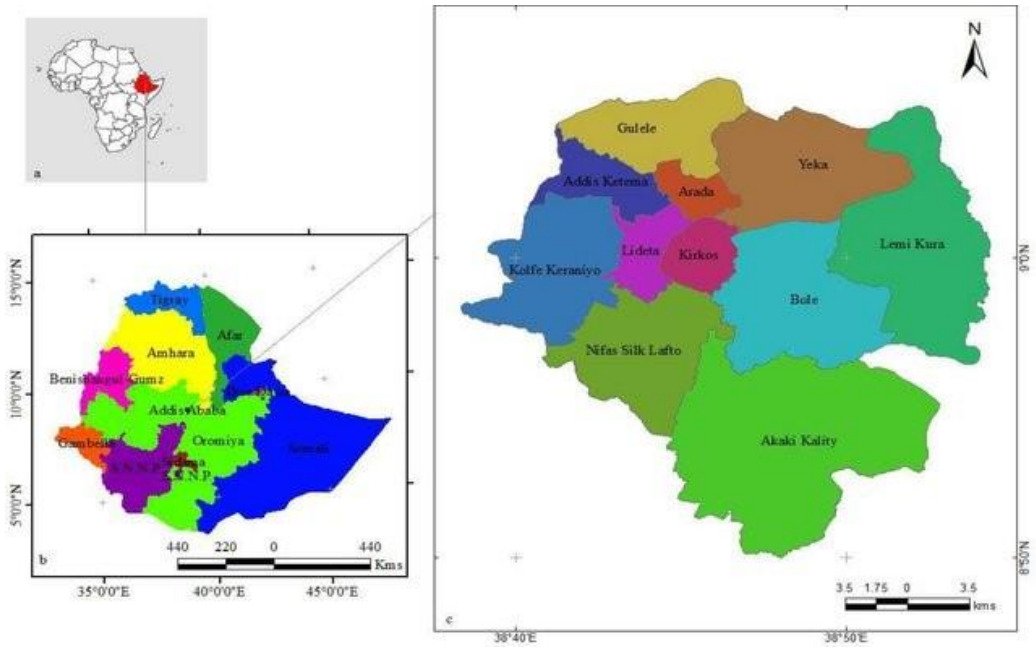


Figure 20: Location map of the study area in Addis Ababa, with Sub-cities

4.2. Study design and period

A retrospective and prospective cohort study design was employed to evaluate BC intrinsic subtypes, microbiota signatures, molecular aberrations in *PIK3CA* gene, and perform performance evaluation of PCR versus IHC among patients from Ethiopia from 1) 2018 to 2022 G.C (prospective cohort); 2) 2013 to 2017 G.C (retrospective cohort).

4.3. Study population

The study population were those pathologically confirmed BC patients diagnosed at selected hospitals in Ethiopia. The retrospective cohort comprised of patients who were diagnosed from 2013 to 2017. G.C whereas the prospective cohort was composed of pathologically confirmed BC patients and recruited to the study prospectively (2018 onwards).

4.4. Sample size calculation and sampling technique

Retrospective cohort

In case of the retrospective cohort, we included all pathologically conformed BC patients who were diagnosed from the year 2013 to 2017 G.C and had complete clinical and histopathological parameters along with a follow up phone calling to collect data on history of death status.

Prospective cohort:

With respect to prospective cohort, the single population proportion formula was employed. The study participants were included to the study conveniently.

$$n = \frac{(Z\ a/2)^2\ P\ Q}{d^2}$$

Where: - n = Sample size required

Za/2: Standard normal distribution abscissa corresponding to 95%confidence interval

d= desired level of precision/Marginal error (5%)

P= Proportion in the target population to BC; P=20.8%

Power of test: 83%.

4.5. Inclusion and exclusion criteria

4.5.1. Inclusion criteria

- Adults who were diagnosed with pathologically confirmed BC with complete clinical, histopathological and follow up data
- Willingness to be part of the study

4.5.2. Exclusion criteria

- Severely ill patients during the study period
- Patients who have not received antibiotic treatment 3 months prior to recruitment, or any neoadjuvant therapy

4.6. Study variables

4.6.1. Independent variables

Clinical and histopathological parameters: age (i.e. at or during diagnosis), simple/clinical stage (T1/T2: early; T3/T4: advanced; Note: nodal status not included for the staging purpose, only based on tumor size), tumor grade, IHC (ER, PR, HER2 and Ki-67status)

4.6.2. Dependent variables

Microbiota composition/signatures, intrinsic subtypes, mutations in *PIK3CA* gene, overall survival (OS) and performance of the test method [sensitivity, specificity, diagnostic accuracy, positive predictive value (PPV) and negative predictive value (NPV)]

4.7. Participant recruitment procedures

Different participant recruitment approach was employed depending on the nature of the cohort.

Retrospective cohort:

In case of the retrospective cohort patients' card/chart review was employed as a strategy to identify BC patients. In the oncology department, we identified patients' cards and verified that the respective breast cases were accompanied with biopsy report. For those cases with biopsy report, we have collected the FFPE block in the pathology department. In all the cases, we reviewed the completeness of clinical and pathological data in the medical chart. Accordingly, only patients with the required information and qualified with the inclusion criteria were included into the study conveniently. Then, the corresponding FFPE tissue were collected from each pathology department using patient biopsy number as unique identifier.

Prospective cohort:

On the other hand, the prospective cohort was established by conveniently recruiting newly diagnosed and pathologically confirmed BC patients prospectively. The prospective cohort were followed up with a regular phone calling to check if they are alive or dead to qualify for overall survival analysis.

4.8. Data collection instrument

A structured questionnaire was employed to collect data about patient demographic and disease characteristic at the time of the diagnosis including clinical/ histopathological parameters and follow up data (history of treatment, endpoint measures including recurrence, metastasis and event [alive, dead]).

Data including age (i.e. at or during diagnosis), residence, cancer type, pathological type, histologic type, tumor size, node status, history of metastasis and histologic grade was abstracted from patients' medical chart. The data collection was done by residents from Oncology department following briefing about the aim of the project and the detail procedures as per the protocol. Additionally, five years follow up data from the cancer registry, medical record review and phone call was made to know whether BC patients alive or dead.

4.9. Sample collection, transportation and storage

4.9.1. FFPE sample

The FFPE tumor blocks were collected from the respective health facilities pathology department using a pre-registered biopsy number. The collected specimens were kept in a dry, moisture and dust free environment and then shipped to Germany without directly exposing the paraffin embedded tissue to direct sun light. Pathological examination was done on the samples centrally at pathology department in Germany. FFPE Samples with verified quality status and confirmed primary tumor were subjected to downstream application (IHC analysis, PAM50 based BC subtyping, microbiota abundance evaluation, *PIK3CA* mutation analysis and PCR based marker detection (ER, PR and HER2).

4.9.2. Fresh breast tissue/ breast biopsy samples

Following the surgical resection, the surgeon identified and collected a paired cancerous and cancer free- normal adjacent (NAT, collected from 5 cm away from tumor enriched area) breast tissue specimens from the surgically resected breast tissue. The collected samples were chopped into pieces and transferred to cryovials. Finally, the samples were immediately snap frozen in liquid nitrogen. The snap frozen tissue later kept in deep freezer for downstream application (IHC analysis, PAM50 based BC subtyping, microbiota abundance evaluation, *PIK3CA* mutation analysis and PCR based marker detection (ER, PR and HER2).

4.10. Sample preparation and nucleic acid extraction

4.10.1. RNA extraction from FFPE

Two to four 10µm FFPE-sections were taken from the FFPE tissue block using Microtomy and transferred to a 1.5ml cryotube. Before the RNA extraction, tissue de-paraffinization was carried out using D-Limonene/Xylene followed by consecutive washing and incubation with ethanol and absolute alcohol. Tissues with > 80% of tumor considered for RNA extraction and the following experiments. RNA isolation was performed using miRNeasy FFPE Mini Kit® (QIAGEN®, Germany) with standard protocol and following manufacturer instruction.

RNA concentration and quality was measured using Nanophotometer. The extracted RNA was stored under -80^{o}C for further analysis.

4.10.2. RNA extraction fresh frozen breast tissue

Tissues was excised from the fresh frozen tumor material. Independent excised tissues with 2 to 4 10um FFT-sections were collected for: H and E staining, IHC analysis and RNA extraction. MiRNeasy mini kit (QIAGEN®, Germany) was employed for RNA extraction. First the tumor tissue was soaked 700 ul of QIAzol lysis reagent. Briefly vortexed and incubated at room temperature. Then, 140 µl chloroform added to the homogenate and then centrifuged at 4^{o}C for 15 minutes.

The upper aqueous phase was transferred to a new collection tube. Into the collected material, 100% ethanol added and mixed thoroughly by pipetting up and down several times. Following that, 700µl of the sample mix transferred into RNeasy Mini spin column holding a 2 ml collection tube. The lid closed and vortexed gently and centrifuge at ≥8000 x g (≥10,000 rpm) for 15second at room temperature. After adding 350µl of RWt buffer, DNase I and RDD buffer mix was added directly into the spin column and incubated at room temperature for 15 minutes. After the second wash with 350µl of RWT buffer, 500µl RPE buffer added and centrifugation was carried out for 2 minutes at 10,000PM. Finally, the RNA was eluted using 30ul of RNase free water. The extracted RNA was stored at -80^{o}C for downstream analysis.

4.10.3. DNA extraction from FFPE

DNA was extracted from the FFPE tumor block using QIAamp® DNA FFPE Tumor Tissue kit (Qiagen, Hilden, Germany) according to the manufacturer instructions. Genomic DNA was isolated from frozen specimens using a NucleoSpin Tissue kit (Clontech, Mountain View, CA, USA). The quality of genomic DNA was measured using the UV absorbance (A260/A280) and high-quality DNA with A260/A280 ratio >1.7 was used for further analysis. The extracted DNA was stored at -80^{o}C to perform *PIK3CA* mutational analysis.

4.10.4. DNA extraction from fresh frozen tissue

DNA was isolated from breast tissues under a sterile lamina flow hood using the Qiagen DNA Isolation kit (Qiagen, Germany) and quantified by Nanodrop and Quant-iT PicoGreen dsDNA Assay Kit (Invitrogen). Blank controls were used for quality control. Specimens were placed into a MoBio PowerMag Soil DNA Isolation Bead Plate. DNA was extracted following MoBio's instructions on a KingFisher robot. The extracted DNA was stored at -80 ^{o}C for Microbiota profile evaluation.

4.11. Experimental design and procedures

4.11.1. Immunohistochemistry (IHC) analysis

IHC analysis was carried out using specific antibodies according to the manufacturer instruction described as follows:

- ER: (Clone: Ab-11; Company: Thermo scientific; order nr. MS-354-P1; host: Mouse; dilution: 1:150)
- PR (Clone: PR 636, DAKO; order nr. M3569; host: mouse; dilution: 1:100), and
- HER2 (Clone: PR 636; DAKO; order nr. SK001; host: rabbit; dilution: rtu)
- Ki67 (Clone: SP6; Company: Thermo scientific; order nr. RM-9106-S; host: mouse; dilution: 1:250).

Expression of hormone receptors (ER and PR) was considered positive when they exceeded 1% of nuclear staining in tumor cells. The immune reactive score (IRS) of the respective receptor reading was observed and interpreted as follows:

- Positivity of ER and PR status was declared when the IRS was >0;
- HER2 positivity was defined where IRS score was ≥3; however, when IRS value = 2, it was interpreted as equivocal and subjected to chromogen in-situ hybridization (CISH) for confirmation.
- The Ki67 proliferation index was graded as 'low' if Ki-67 stained in < 20% of tumor cells or 'high' when at least 20% of tumor cells stained positive.
- The histological grading was performed using the Elston-Ellis grading system (Nottingham Grading System).

4.11.2. Chromogenic In Situ Hybridization (CISH) for HER2

Clinical relevance: The ERBB2 gene (erb-b2 receptor tyrosine kinase 2, a.k.a. HER2 and NEU) is located in the chromosomal region 17q12 and encodes the cellular growth factor receptor p185. Amplification of the proto-oncogene ERBB2, observed in approximately 20% of all BC samples, and has been correlated with a poor prognosis of the disease. To verify equivocal results by IHC technique, we also performed CISH test for HER2.

Test principle: The chromogenic in situ hybridization (CISH) technique allows the detection and visualization of specific nucleic acid sequences in cell preparations. Hapten-labeled nucleotide fragments, so called CISH probes, and their complementary target sequences in the preparations are co-denatured and subsequently allowed to anneal during hybridization.

Afterwards, unspecific and unbound probe fragments are removed by stringency washing steps. Duplex formation of the labeled probe can be visualized using primary (unmarked) antibodies, which are detected by secondary polymerized enzyme-conjugated antibodies. The enzymatic reaction with chromogenic substrates leads to the formation of colored precipitates. After counterstaining the nucleus with a nuclear dye, hybridized probe fragments are visualized by light microscopy.

Brief procedure: Qualitative detection of ERBB2 was performed using ZytoDot 2C SPEC ERBB2/CEN 17 Probe Kit ZytoDot 2C SPEC ERBB2/CEN 17 Probe kit (Zytovision GmbH, Bremerhaven/ Germany). The required consumables including alcohol were prepared and slides were pretreated with different concentration of alcohol and washed. After this step, denaturation and hybridization followed. Finally, post-hybridization and detection carried out following data interpretation protocol as per the manufacturer instruction.

Interpretation of results: The hybridization signals of Digoxigenin-labeled polynucleotides appear dark green colored (ERBB2 gene region), and Dinitrophenyl-labeled polynucleotides appear bright red (CEN 17). Normal situation: In interphases of normal cells or cells without an amplification involving the ERBB2 gene region, two distinct dot-shaped green signals and two distinct dot-shaped red signals appear. Aberrant situation: In cells with an amplification of the ERBB2 gene region, an increased number of green signal or green signal clusters will be observed (Figure 15).

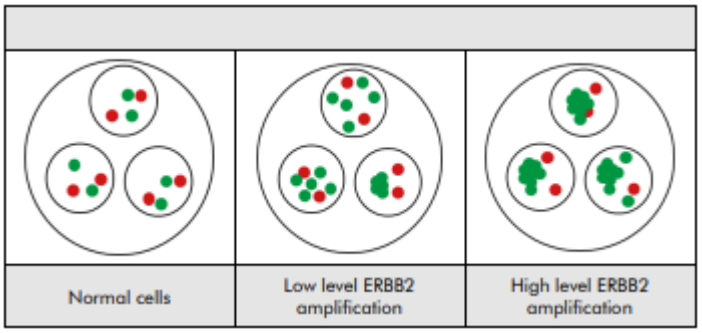


Figure 21: Expected results in normal and aberrant nuclei

4.11.3. PAM50 based gene expression profile

PAM50 based gene expression was measured using the NanoString nCounter® Analysis System (NanoString Technologies, Seattle, WA, USA). The system measures the relative abundance of each gene mRNA of interest using a multiplexed hybridization assay, and digital readouts of fluorescent barcoded probes that are hybridized to each transcript are created. The data collection/ acquisition was carried out in the nCounter® Digital Analyzer. Data import, quality control, and normalization of expression values were conducted with the nSolver version 4.0 (NanoString Technologies Inc.). Background subtraction from raw transcript counts was performed through negative input controls. Following reference-normalization by dividing the geometric mean of six references-control genes (ACTB, G6PD, RPLP0, TBP, TFRC, and UBB), normalized counts were log2 transformed before further analyses. Intrinsic subtype classification was calculated in R using the PAM50 predictor and was applied to the nearest PAM50 centroid algorithm Bioclassifier (Bernard et al., 2009).

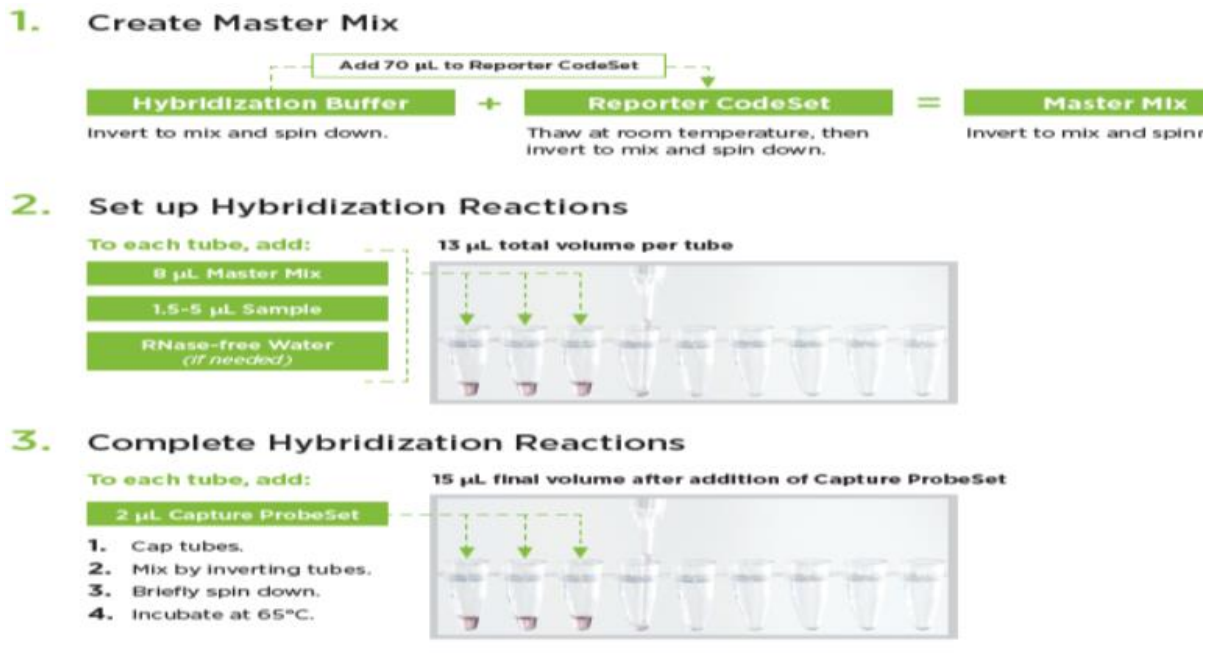


Figure 22: Quick-guide work flow for the Nanostring experiment (Nanostring Technologies, 2017)

4.11.4. Microbiome profiling: PCR, sequencing and sequence processing

Quantitative PCR: Bacterial 16S rRNA genes were PCR-amplified with dual-barcoded primers targeting the V4 region (515F 5'-GTGCCAGCMGCCGCGGTAA-3', and 806R 5'GGACTACHVGGGTWTCTAAT-3'), as per the protocol (Kozich et al., 2013). 25 µl reaction mix using iQ SYBR Green Supermix (Bio-Rad) were run on Applied Biosystems StepOne Plus instrument in triplicate. Linear regression was used to determine copy numbers of samples, based on CT of standards. Melt Curve Reaction specificity was assessed using a melt curve from 55 oC to 95 oC, held at 0.5 oC increment for 1s.

Sequencing: Amplicons were sequenced with an Illumina MiSeq using the 300-bp paired-end kit (v.3). MiSeq-generated Fastq files were quality-filtered, sequences were denoised, taxonomically classified using Silva (v. 138) as the reference database, and clustered into 97%-similarity operational taxonomic units (OTUs) with the mothur software package http://www.mothur.org; v. 1.44.1) (Schloss et al., 2009), following the recommended procedure (https://www.mothur.org/wiki/MiSeq_SOP; accessed Nov 2020). High quality reads were classified using Silva v. 138 as the reference database. We aggregated OTUs into each taxonomic rank, and plotted the relative abundance of the most abundant ones. Some samples have been dropped from the analysis due to low read counts (Figure 23).

Analytical flowchart

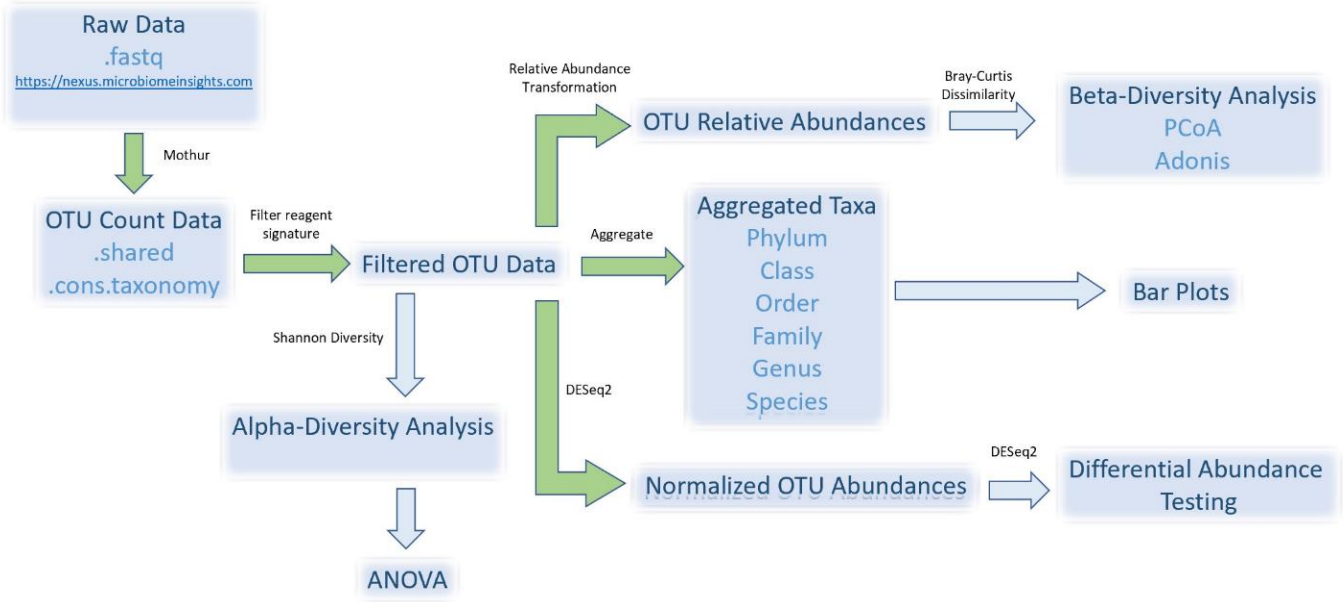


Figure 23: Step by step analytical flow of microbiome analysis

4.11.5. *PIK3CA* mutation detection

Tumor tissue of pathologically confirmed BC patients was examined for *PIK3CA* gene mutation analysis. TaqMan® Mutation Detection (Competitive Allele-Specific TaqMan® PCR Assay Applied Biosystems by life technologies, Carlsbad, CA 92008 USA) was employed to detect mutations. Accordingly, mutations in *PIK3CA* gene at the following three hotspot codons: 1) **p.E542K**:c1624G>A 2) **p.E545K**:c1633G>A 3) **p.H1047R**:c3140A>G were examined. These hotspots account for approximately 80% of all *PIK3CA* mutations (Baker et al., 2012; Ligresti et al., 2009) . The PCR thermal condition was set: 10 minutes at 95ºC; 15 second at 92 ºC and 1 minute at 58 ºC for 5 cycles; 15 second for 92 ºC and 1 minute for 60 ºC for 40 cycles. After amplification, the CT values are determined by the Applied Biosystems®' real-time PCR instrument software (ABI StepOnePlus v2.1). In mutation analysis calculations, the difference between the Ct value of each mutant allele assay and the CT value of the gene reference assay is calculated. This ΔCT value represents the quantity of the specific mutation allele detected within the sample; it was used to determine sample mutation status by comparison to a predetermined detection ΔCT cutoff value.

In each run, DNA derived from BC cell lines (BCCLs) including MCF-7, BT-20 and M296 were utilized as a positive control. On the other hand, no template control (NTC) was used as a negative control.

In this study, we analyzed the frequency of individual hotspot mutations and we combined all together to calculate the overall frequency of combined mutations in each hotspot site. Accordingly, the mutation status was assigned as *PIK3CA* mutated when there was mutation in at least one or more hotspots whereas tumor was defined as wild type when there was no mutation in any of the assessed hotspot's site. Figure 18 shows participants recruitment, sample collection and laboratory analysis procedures of *PIK3CA* mutation.

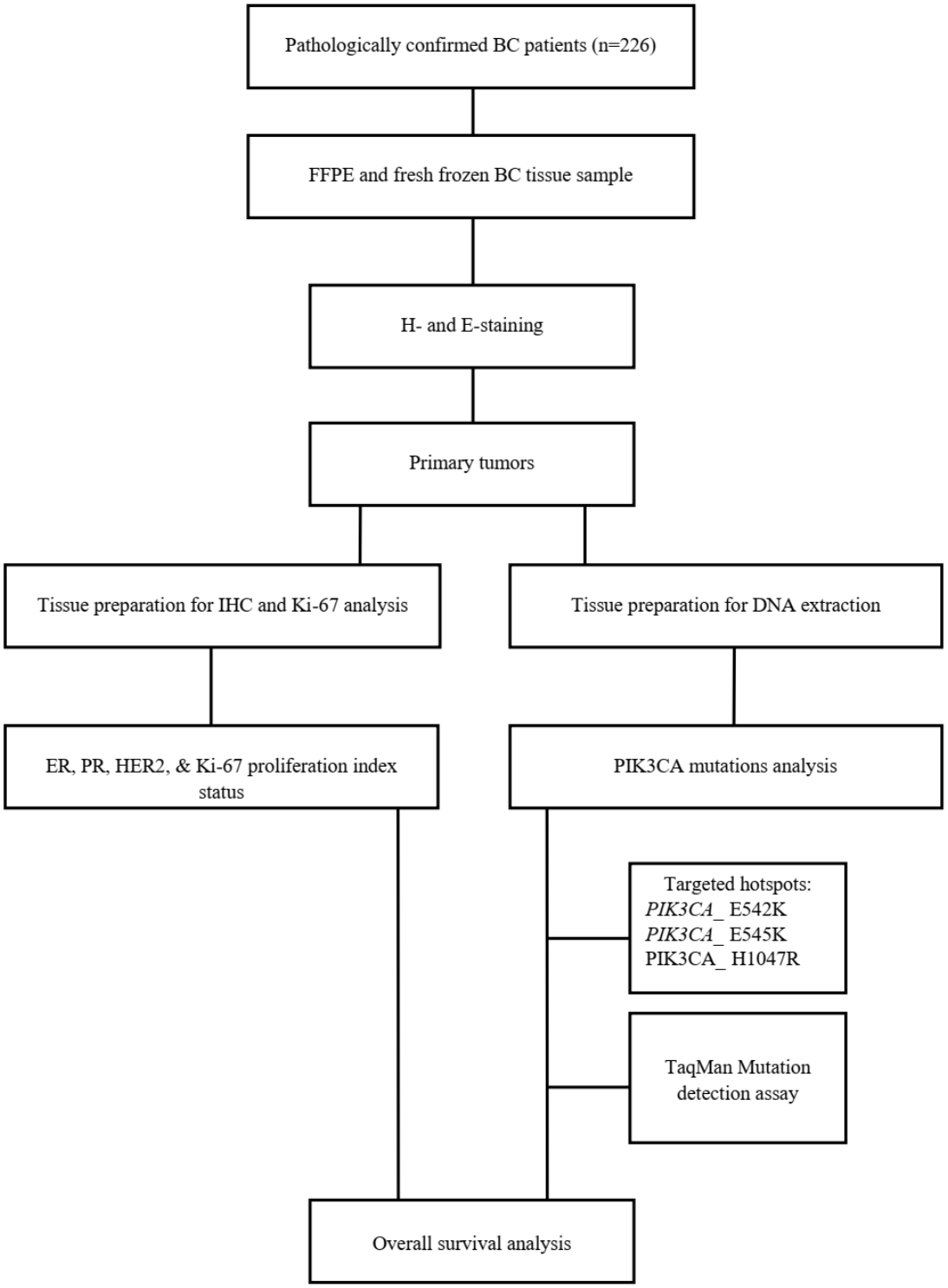


Figure 24: Flow of participants recruitment and laboratory experiment procedures

4.11.6. End Point PCR and qPCR experiment

4.11.6.1. cDNA synthesis procedures

cDNA synthesis was carried out using Biozym Kit (Biozym, Germany) following manufacturer instruction. Primary, the entire reagents thawed up at room temperature. Following that master mix constituted of dNTP Mix (10 mM each), RNase Inhibitor, Oligo (dT)12–18 (10 µM) – or, Hexamer Primer (25 µM) – or, 5× cDNA Synthesis Buffer, Reverse Transcriptase and PCR Grade Water Variable was prepared. And then, 11 µl of the master mix mix transferred to 0.2ul PCT tubes followed by adding 2 µl (500ng) of the starting material/RNA. Finally, the whole mix briefly placed into the PCR machine at a given PCR condition. A variety of cDNA concentration applied to select optimum cDNA concentration to run the PCR.

4.11.6.2. End point PCR experiment design

Cell line control: A Positive (BT474) and negative control (BT20) BCCLs used for the optimization and keeping quality control of the experiment (Table 5).

Table 5: Cell culture control

Cell line	ESR1	PGR	ERBB2	Starting material concentration
BT-474	Positive	Positive	Positive	25ng/ µl
BT-20	negative	Negative	negative	25ng/ µl

Reference gene and primer set: *RPL37A* was used as a reference gene. A specific primer sequence for each of GOI was derived from a published paper (Müller et al., 2011).

Table 6: Reference genes and primer nucleotide sequence

Gene name	Primer name	Nucleotide sequence	Amplicon length
RPL37A	RPL37A FWD	5´-TGTGGTTCCTGCATGAAGACA-3´	65
	RPL37A REV	5´-GTGACAGCGGAAGTGGTATTGTAC-3´	
ESR1	ESR1 FWD	5´-GCCAAATTGTGTTTGATGGATTAA-3´	73
	ESR1 REV	5´-GACAAAACCGAGTCACATCAGTAATAG-3´	
PGR	PGR FWD	5´-AGCTCATCAAGGCAATTGGTTT-3´	100
	PGR REV	5´-ACAAGATCATGCAAGTTATCAAGAAGTT-3	
ERBB2	ERBB2 FWD	5´-CCAGCCTTCGACAACCTCTATT-3´	87
	ERBB2 REV	5´-TGCCGTAGGTGTCCCTTTG-3´	

PCR conditions: The experiment started with variable concentration of cDNA (100ng, 50ng and 25 ng); however, finally 50ng of cDNA used as a starting material for the purpose. Moreover, variable PCR conditions including adjusting the cycle number (25 cycle for ERBB2 gene and 30 cycle for ESR1 and PGR genes) performed to set optimum conditions for the establishment of the method. Herewith below presented the PCR conditions for the respective GOI.

For RPLA37, ESR1, PGR:

Hot Start	95°C	1 min	
Denaturation	95°C	20 sec	**30 cycles**
Annealing	60°C	20 sec	
Elongation	72°C	30 sec	

For ERBB2:

Hot Start	95°C	1 min	
Denaturation	95°C	20 sec	**25 cycles**
Annealing	60°C	20 sec	
Elongation	72°C	30 sec	

Gel documentation: After running the PCR, 6ul of the PCR product and 3ul of 50bp DNA ladder loaded into the agarose gel. Following that, the voltage was set at 60V for five minutes and at 160 V for 25 minutes. Finally, the gel documentation performed using Image Quant LAS 4000 luminescent analyzer.

4.11.6.3. TaqMan® Gene Expression Assays

The extracted RNA was converted into cDNA to perform real time PCR based gene expression analysis among the GOI. Using the cDNA as a staring material, expression of the three genes determined using TaqMan® Gene Expression Assays. Briefly speaking, 50ng/ul of cDNA used as a starting material with two technical duplicates of each sample. RPLPO and HPRT used as a reference gene and endogenous control used in each run. Reverse transcription as well as Real-Time PCR for gene of interest expression analysis carried out using specific primer sequence for GOI as described before (Müller et al., 2011). A master of master mix composed of Master Mix (2X), TaqMan® Gene Expression Assay (20X) and Nuclease free water was prepared in a 1.5ml Eppendorf tube. And then, 9ul of the master mix transferred in to a 96 well plate. Finally, 1ul of cDNA (50ng) cDNA material added to each well and covered with adhesive tape. The 96 well plate placed in to Thermal cycler and the PCR condition adjusted accordingly. In all the experiment/run, each sample measured in duplicate.

Relative gene expression and cutoff

The Relative expression levels of genes of interest for each breast tumor sample in relation to endogenous control calculated based on the threshold cycle (CT) value using the following formula: $\Delta Ct = 20 - (CtGOI - Ct\ Endogenous\ control)$. ΔCt values positively correlate with relative gene expression. The optimal cut-off points for endogenous gene-normalized data were defined for classification of *ESR1*, *PGR*, and *ERBB2* status based on the column scatter plots. Positivity for ESR1 mRNA was defined as a ΔCt of more than 14.5, negativity for *ESR1* was determined as a ΔCt less than 13.5 Those few samples that were between 13.5 and 14.5 were designated as equivocal for ESR1 mRNA expression. For *PGR* mRNA the upper and lower cutoff were 12 and 11, respectively, the samples between both cutoffs were regarded as equivocal. For ERBB2 mRNA the upper and lower cutoff were 19 and 18, respectively, the samples between both cutoffs were regarded as equivocal (Müller et al., 2011).

4.12. Quality Assurance

To ensure the quality of the generated data and the quality of the analysis internal quality assurance systems was in place for all laboratory procedures. All the methods used has been validated as fit for the purpose before use in the study and if an amendment and/or modification are made it will be documented in the log books and reported to all members of the research group. SOPs developed were strictly followed for all laboratory procedures.

4.13. Data Processing and Analysis

Data entry and analysis was done using SPSS version 22. Descriptive statistics were caried out and data was presented using tables and figures. Frequency distributions of intrinsic subtypes was estimated for the overall cohort and within subgroups defined by clinical and histopathological categories. For PAM50 analysis, data was normalized and analyzed using R-software. Logistic regression employed to explore association between independent variable and outcome variables. Strength of association was measured using odds ratio. Statistically significant association was declared at a p value of <0.05. Survival analysis was performed using Kaplan-Meier and Cox regression model.

Two-by-two tables compare IHC versus end point PCR and IHC versus qPCR for ER, PR and HER2. Cross tabulation was performed to record the overlapping count of measures between the reference method and the training cohort. We employed MedCalc® statistical software, to estimate the sensitivity, specificity, PPV, NPV and diagnostic accuracy of the methods. Level of concordance between IHC and the respective method was carried out using Kappa measures of agreement (Table 7) (Landis & Koch, 1977). The approximate significance was declared at P-value $\leq$0.05.

Table 7: Kappa agreement cutoff and interpretation

Kappa	**Strength of agreement**
<0	No agreement
0.01 to 0.20	None to slight
> 0.2 $\leq$ 0.4	Fair
> 0.4 $\leq$ 0.6	Moderate
>0.6 $\leq$ 0.8	Substantial / good
> 0.8 $\leq$ 1	Perfect / very good

Alpha diversity was estimated with the Shannon index on raw OTU abundance tables after filtering out contaminants. The significance of diversity differences was tested with ANOVA or linear mixed model depending on the study design. To estimate beta diversity across samples, we excluded OTUs occurring with a count of less than 3 in at least 10 % of the samples and then computed Bray-Curti's indices.

We visualized beta diversity, emphasizing differences across samples, using Principal Coordinate Analysis (PCoA) ordination. Variation in community structure was assessed with permutational multivariate analyses of variance (PERMANOVA) with treatment group as the main fixed factor and using 999 permutations for significance testing. A dot-dash circle in ordination plot represent for the central of each cluster (Nguyen et al., 2020). All analyses were conducted in the R environment.

Relative abundance of OTUs in the heatmap was Min-Max normalized to be comparable in the scale from 0 to 1. Spearman's correlation analysis was performed between normalized-OTUs to construct a correlation matrix using Spearman's correlation coefficients, which was then converted into a distance matrix by (1—correlation coefficients). Next, hierarchical clustering was performed on the distance matrix to build a tree using the complete-linkage clustering algorithm where branch lengths reflect the degree of association between OTUs (i.e., shorter branches indicate that relative abundance of OTUs were more similar among individuals) (Nguyen et al., 2020).

4.14. Ethical consideration

Ethical approval was obtained from Addis Ababa University College of Health Sciences-Institutional Review Board (IRB) (IRB protocol #092/17/17) and National Research Ethics Review Committee (NRERC) (NRERC protocol # MOSHE/RD/2018). Permission letter was obtained from the selected health facilities prior to undertaking the study. The study participants were informed and their written informed consent was collected before collection of any clinical sample and information. Confidentiality was maintained using unique code and the document was kept locked. Participation in the study was on voluntary basis and participants were informed about the possibility of withdrawal from the study at any phase of the project without losing any health-related service. Permission to conduct the study was obtained from all institutions involved. We have communicated findings to attending physician based on their request to facilitate the decision-making process/ treatment selection.

5. RESULTS

5.1. Intrinsic subtype distribution among Ethiopian breast cancer patients

5.1.1. Clinical and histopathological characteristics of patients

The median age of the cancer patients was 40 years. Most of the tumors were G3 (58.1%). Majority of the tumors were identified as “no special type” (NST) or invasive ductal carcinoma (IDC) of the breast. Furthermore, 69.5% and 21.9% of the tumors were hormone receptor positive (ER+ and/ or PR+) and HER2 positive, respectively. Two-third of the tumors turned out with high Ki-67 proliferation index. The IHC analysis revealed that higher proportion of the tumors were HR+ /HER2- (55.9%) followed by HR- /HER2- (22.2%) phenotype (Table 8).

Table 8: Patient and tumor characteristics

Parameters	All n=334 (100%)
Age Group (years)	
< 50	201 (60.2)
≥ 50	95 (28.4)
Unknown	38 (11.4)
Tumor Size	
T1 or T2	168 (50.3)
T3 or T4	126 (37.7)
unknown	40 (12.0)
Histological Types	
NST	303 (90.7)
Non-NST	31 (9.3)
Tumor Grade	
G1 or G2	140 (41.9)
G3	194 (58.1)
Estrogen Receptor Status	
ER-positive (≥1%)	184 (55.1)
ER-negative (<1%)	150 (44.9)
Progesterone Receptor Status	
PgR-positive (≥1%)	157 (47.0)
PgR-negative (<1%)	177 (53.0)
Hormone Receptor Status	
HR-positive	232 (69.5)
HR-negative	102 (30.5)
HER2 Status	
Negative	261 (78.1)
Positive	73 (21.9)
Ki-67 Proliferation Index	
Low (<20%)	132 (39.5)
High (≥20%)	202 (60.5)
IHC Group	
HR+/HER2−	187 (55.9)
HR+/HER2+	45 (13.5)
HR−/HER2+	28 (8.4)
HR−/HER2−	74 (22.2)

5.1.2. Distribution of intrinsic subtypes

5.1.2.1. Gene expression across intrinsic subtypes

The expression levels of the 50 gene included in the assay (Figure 25) revealed a gene expression signature which was unique for each subtype.

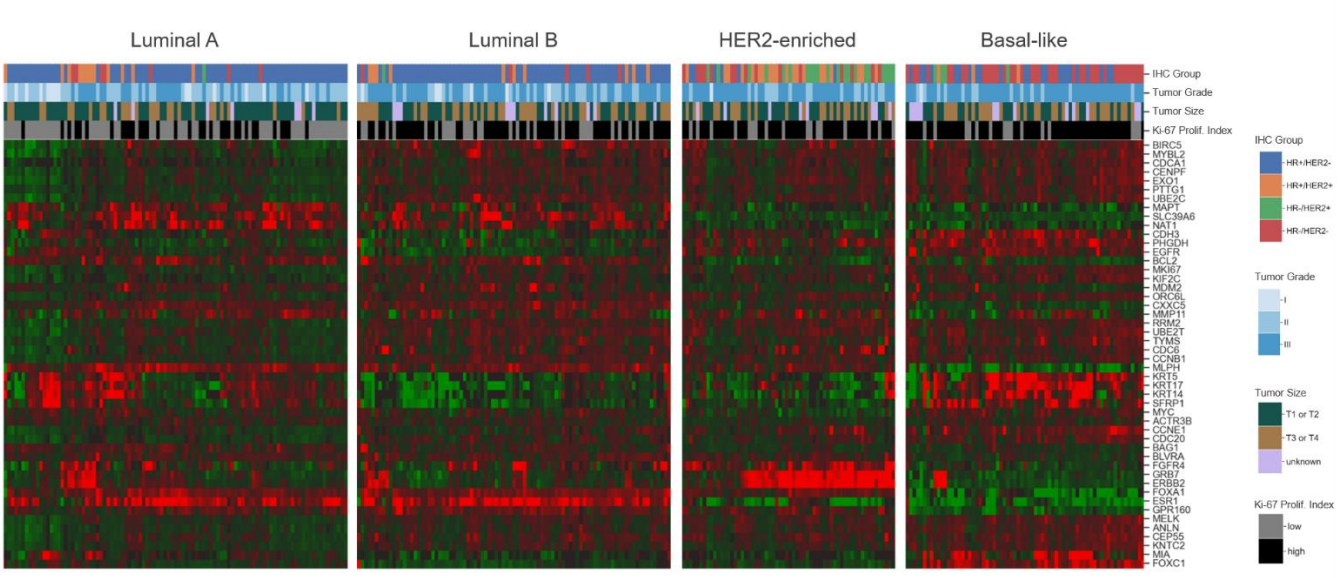


Figure 25: Gene Expression heatmap of the 50 loci used for the PAM50 classification of 334 BC samples *[The 334 samples are grouped horizontally according to their intrinsic subtype, which are indicated above each block. Red tiles denote overexpression, while green tiles correspond to under expression. The four horizontal bars above the heatmaps indicate the classification of samples according to IHC groups, tumor grade, tumor size and Ki-67 proliferation index (top-down, respectively, with color codes for each bar given at the right side of the figure). HR: Hormone Receptor; HER2: Human Epidermal Growth Factor Receptor 2]*

5.1.2.2. Intrinsic subtypes by clinical and histopathological parameters

An overview of the prevalence of intrinsic subtypes and the distribution of clinical and histopathological parameters among the PAM50-based intrinsic subtypes is given in Table 9. Accordingly, the classification of intrinsic subtypes based on the PAM50 assay yielded 104 (31.1%) luminal A, 91(27.2%) luminal B, 62 (18.6%) HER2-enriched and 77 (23.1%) basal-like (Table 9). Majority of luminal tumors identified with in favorable clinical and histopathological features. Out of luminal B, HER2-enriched and basal-like tumors, most tended grade 3 tumors and expressed high level of Ki-67 as expected.

Similar to luminal A tumors, most luminal B tumors tended ER+ and PR+. Among HER2-enriched tumors, most of it identified as HER2+ whereas high proportion basal-like tumors observed HER2-.

Table 9: The distribution of clinical and histopathological parameters within intrinsic subtypes

Parameters	All	Luminal A	Luminal B	HER2-enriched	Basal-like
	n=334 (100%)	n=104 (31.1%)	n=91 (27.2%)	n=62 (18.6%)	n=77 (23.1%)
Age Group (years)					
< 50	201 (60.2%)	66 (63.5%)	58 (63.7%)	33 (53.2%)	44 (57.1%)
≥ 50	95 (28.4%)	29 (27.9%)	24 (26.4%)	21 (33.9%)	21 (27.3%)
unknown	38 (11.4%)	9 (8.7%)	9 (9.9%)	8 (12.9%)	12 (15.6%)
Tumor Size					
T1 or T2	168 (50.3%)	68 (65.4%)	39 (42.9%)	31 (50.0%)	30 (39.0%)
T3 or T4	126 (37.7%)	29 (27.9%)	41 (45.1%)	23 (37.1%)	33 (42.9%)
Unknown	40 (12.0%)	7 (6.7%)	11 (12.1%)	8 (12.9%)	14 (18.2%)
Histological Types					
NST	303 (90.7%)	94 (90.4%)	86 (94.5%)	54 (87.1%)	69 (89.6%)
Non-NST	31 (9.3%)	10 (9.6%)	5 (5.5%)	8 (12.9%)	8 (10.4%)
Tumor Grade**					
G1 or G2	140 (41.9%)	71 (68.3%)	37 (40.7%)	18 (29.0%)	14 (18.2%)
G3	194 (58.1%)	33 (31.7%)	54 (59.3%)	44 (71.0%)	63 (81.8%)
Estrogen Receptor Status**					
ER-positive (≥1%)	184 (55.1%)	77 (74.0%)	69 (75.8%)	18 (29.0%)	20 (26.0%)
ER-negative (<1%)	150 (44.9%)	27 (26.0%)	22 (24.2%)	44 (71.0%)	57 (74.0%)
Progesterone Receptor Status**					
PgR-positive (≥1%)	157 (47.0%)	69 (66.3%)	60 (65.9%)	13 (21.0%)	15 (19.5%)
PgR-negative (<1%)	177 (53.0%)	35 (33.7%)	31 (34.1%)	49 (79.0%)	62 (80.5%)
Hormone Receptor Status**					
HR-positive	232 (69.5%)	95 (91.3%)	85 (93.4%)	26 (41.9%)	26 (33.8%)
HR-negative	102 (30.5%)	9 (8.7%)	6 (6.6%)	36 (58.1%)	51 (66.2%)
HER2 Status**					
Negative	261 (78.1%)	91 (87.5%)	80 (87.9%)	21 (33.9%)	69 (89.6%)
Positive	73 (21.9%)	13 (12.5%)	11 (12.1%)	41 (66.1%)	8 (10.4%)
Ki-67 Proliferation Index**					
Low (<20%)	132 (39.5%)	66 (63.5%)	31 (34.1%)	17 (27.4%)	18 (23.4%)
High (≥20%)	202 (60.5%)	38 (36.5%)	60 (65.9%)	45 (72.6%)	59 (76.6%)

HER2: Human Epidermal Growth Factor Receptor 2; NST: No special type; ER: Estrogen Receptor; PgR: Progesterone Receptor; HR: Hormone Receptor

When we look a closer look at specific subtypes with in clinical and histopathological features (Table 10), majority of T1/ T2 tumors (63.7%) identified as luminal subtypes whereas within T3/T4 tumors, 26.2% were basal like. Of G1/G2 tumors, most (76.9%) turned luminal tumors; however, out of G3 tumors, the unfavorable subtypes were predominant (55.2%). Majority of the luminal tumors were linked positive ER, PR and HR status unlike basal like tumors.

With respect to age, there was only slight difference in the distribution of intrinsic subtypes across age as shown in Figure 26.

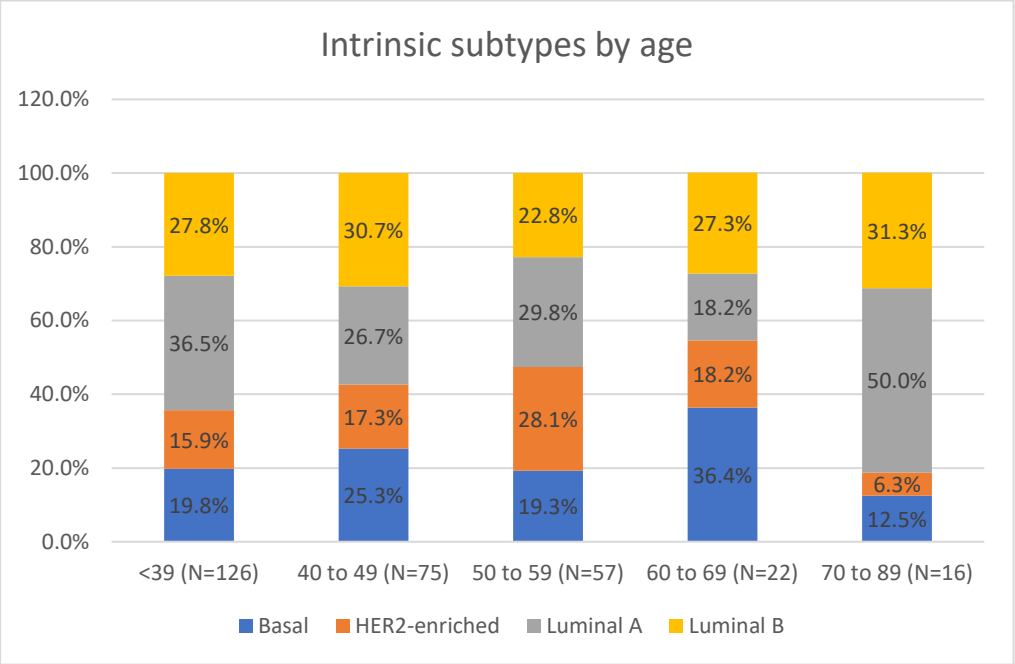


Figure 26: Intrinsic subtypes distribution across age

A detailed analysis focused on the whole measured range of the prognostic parameters ER, PgR, HER2 and Ki-67 proliferation index is visualized in Figure 27, Table 10. While the IRS scores of ER and PgR spanned evenly between 0 and 12 for both luminal subtypes, these accumulated below IRS 2 for HER-enriched and Basal-like samples (Figure 27A, B). This underlines the associations described between intrinsic subtypes and the expression of the steroid hormone receptors detected by IHC (see Table 10 for binary assessments).

The violin plot of DAKO scores made the aforementioned association between HER2-enriched samples and positive HER2 status evident. The accumulation of scores > 2 was almost exclusively seen in HER2-enriched samples (Figure 27C and Table 10).

Similarly, specimens of low proliferation rates (< 20%) were more likely to be of the Luminal A subtype (Figure 27D and Table 10). However, low proliferation rates were also present in the other intrinsic subtypes.

The highest proliferation rates were observed among Basal-like samples. The univariate analysis from X^2-test shows tumor grade, estrogen receptor, progesterone receptor, hormone receptor, HER2 status, Ki-67 and IHC group had a statistically significant association with intrinsic subtypes (P-value <0.01; 99% CI).

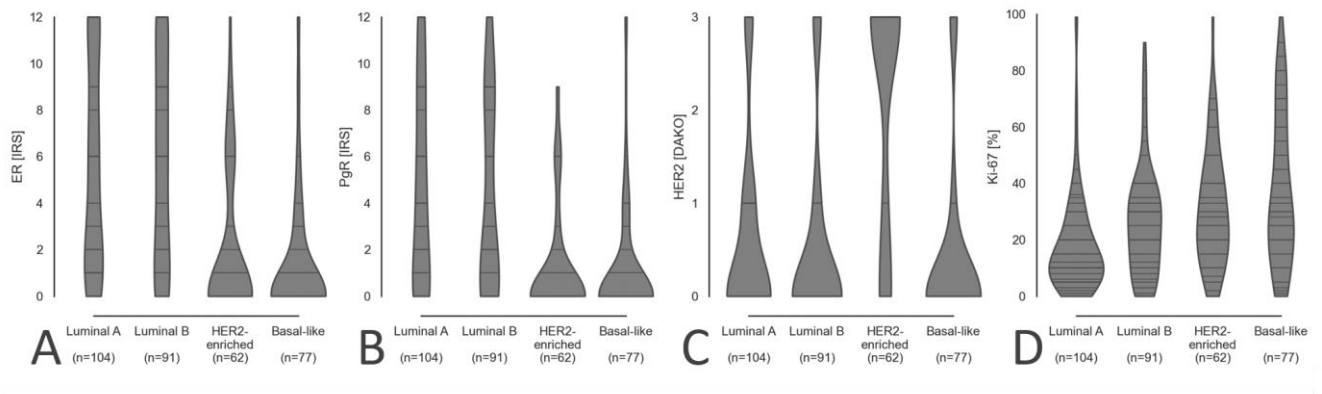


Figure 27: Scores of Estrogen Receptor, Progesterone Receptor, HER2 and Ki-67 Proliferation Index among intrinsic subtypes. *The horizontal widths of the violin plots correspond to the frequency distribution of data points. Horizontal bars shown in the plots denote data available for those values. ER: Estrogen Receptor; PgR: Progesterone Receptor; IRS: Immune-reactive score. ER: Estrogen Receptor; PgR: Progesterone Receptor; HER2: Human Epidermal Growth Factor Receptor 2*

Table 10: Relationship between clinical or histopathological parameters and intrinsic subtypes, with proportions given within each row

Parameters	All n = 334 100%	Luminal A n = 104 31.10%	Luminal B n = 91 27.20%	HER2-enriched n = 62 18.60%	Basal-like n = 77 23.10%	P-value
Age group (years)						0.67
< 50	201 (100%)	66 (32.8%)	58 (28.9%)	33 (16.4%)	44 (21.9%)	
≥ 50	95 (100%)	29 (30.5%)	24 (25.3%)	21 (22.1%)	21 (22.1%)	
Unknown	38 (100%)	9 (23.7%)	9 (23.7%)	8 (21.1%)	12 (31.6%)	
Tumor size						0.01
T1 or T2	168 (100%)	68 (40.5%)	39 (23.2%)	31 (18.5%)	30 (17.9%)	
T3 or T4	126 (100%)	29 (23%)	41 (32.5%)	23 (18.3%)	33 (26.2%)	
unknown	40 (100%)	7 (17.5%)	11 (27.5%)	8 (20%)	14 (35%)	
Histological type						0.45
NST	303 (100%)	94 (31%)	86 (28.4%)	54 (17.8%)	69 (22.8%)	
Non-NST	31 (100%)	10 (32.3%)	5 (16.1%)	8 (25.8%)	8 (25.8%)	
Tumor grade*						$3.35 \times 10\text{-}11$
G1 or G2	140 (100%)	71 (50.7%)	37 (26.4%)	18 (12.9%)	14 (10%)	
G3	194 (100%)	33 (17%)	54 (27.8%)	44 (22.7%)	63 (32.5%)	
Estrogen receptor						$5.11 \times 10\text{-}16$
ER-positive (≥ 1%)	184 (100%)	77 (41.8%)	69 (37.5%)	18 (9.8%)	20 (10.9%)	
ER-negative (< 1%)	150 (100%)	27 (18%)	22 (14.7%)	44 (29.3%)	57 (38%)	
Progesterone receptor						$6.99 \times 10\text{-}15$
PgR-positive (≥ 1%)	157 (100%)	69 (43.9%)	60 (38.2%)	13 (8.3%)	15 (9.6%)	
PgR-negative (< 1%)	177 (100%)	35 (19.8%)	31 (17.5%)	49 (27.7%)	62 (35%)	
Hormone receptor						$4.43 \times 10\text{-}25$
HR-positive	232 (100%)	95 (40.9%)	85 (36.6%)	26 (11.2%)	26 (11.2%)	
HR-negative	102 (100%)	9 (8.8%)	6 (5.9%)	36 (35.3%)	51 (50%)	
HER2 status*						$7.56 \times 10\text{-}19$
Negative	261 (100%)	91 (34.9%)	80 (30.7%)	21 (8%)	69 (26.4%)	
Positive	73 (100%)	13 (17.8%)	11 (15.1%)	41 (56.2%)	8 (11%)	
Ki-67 proliferation index*						$2.48 \times 10\text{-}8$
Low (< 20%)	132 (100%)	66 (50%)	31 (23.5%)	17 (12.9%)	18 (13.6%)	
High (≥ 20%)	202 (100%)	38 (18.8%)	60 (29.7%)	45 (22.3%)	59 (29.2%)	
IHC Group*						$5.185 \times 10\text{-}40$
HR+/HER−	187 (100%)	83 (44.4%)	75 (40.1%)	6 (3.2%)	23 (12.3%)	
HR+/HER+	45 (100%)	12 (26.7%)	10 (22.2%)	20 (44.4%)	3 (6.7%)	
HR−/HER+	28 (100%)	1 (3.6%)	1 (3.6%)	21 (75%)	5 (17.9%)	
HR−/HER2−	74 (100%)	8 (10.8%)	5 (6.8%)	15 (20.3%)	46 (62.2%)	

**Parameters for which a $p < 0.01$ (from a χ^2 test for independence) was observed, denoting a lack of independence between histopathological parameters and intrinsic subtypes. ER: estrogen receptor; HER2: human epidermal growth factor receptor 2; HR: hormone receptor; NST: no special type; PgR: progesterone receptor.*

5.1.3. Intrinsic subtypes by IHC groups

We investigated the distribution of the combined IHC markers with respect to PAM50 based intrinsic subtypes. Of the tumors identified as luminal A and luminal B by PAM50 assay, 79.8% and 82.4 of them were HR+ /HER2-, respectively. Of the basal-like subtypes, 59.7% were HR- /HER- using IHC techniques. Within HER2 enriched subtype, 33.9% were HR-/HER2+. On the other hand, correlation of IHC group across PAM50 based intrinsic subtypes observed. Among samples classified as HR+ /HER2-, 44.4% and 40.1% were classified as luminal A and luminal B subtype, respectively. Amongst HR- /HER2- tumors by IHC, 62% and 20% were basal-like and HER2-enriched, respectively. Of those with HR- / HER2+ IHC-type, 75% and 18% of tumors were HER2-enriched and basal-like by PAM50 gene classifier (Table 10, Figure 28).

A strong correlation was observed between intrinsic subtypes and IHC groups (p-value from χ^2 test for independence < 0.001, Figure 28). This association was driven mainly by the luminal subtypes, from which 84.5% (n=158) were grouped as HR-positive and HER2-negative (Figure 28, upper left). Out of the 62 samples classified as HER2-enriched, 66.1% (n=41) were defined as HER2-positive irrespective of HR status (Figure 28).

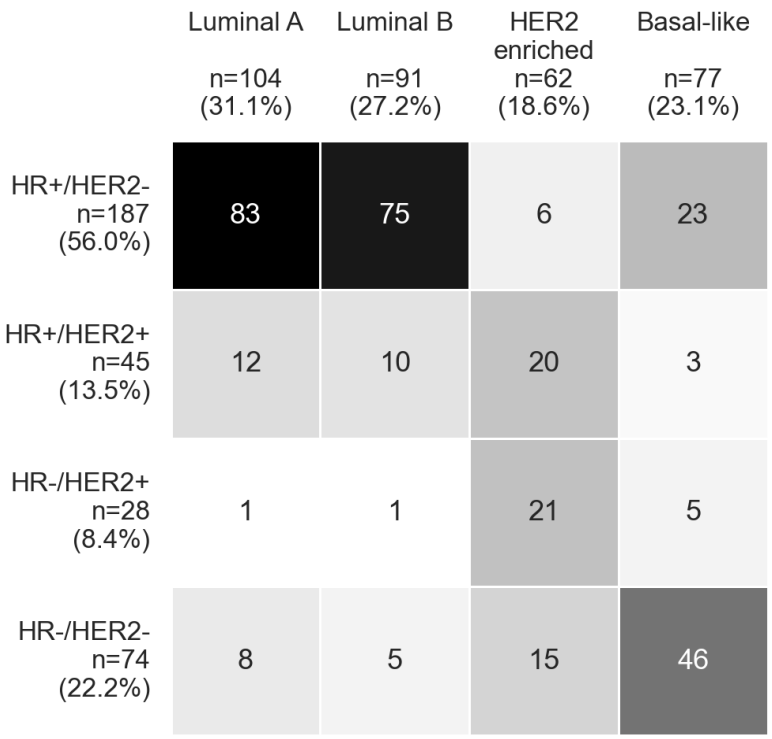


Figure 28: Color-coded cross table of 334 BC tissue samples grouped according to IHC groups (rows) and intrinsic subtypes (columns). *The cell color gradient indicates the relationship in terms of a strong discordance (white) to a strong concordance (black) between IHC and intrinsic subtype classifications. HR: Hormone Receptor; HER2: Human Epidermal Growth Factor Receptor 2*

5.1.4. Association between intrinsic subtypes and histopathological parameters

A multivariate analysis was carried out to observe the relationship between intrinsic subtypes with clinical and histopathological parameters. Luminal A samples had a more than three-fold increased probability for favorable characteristics such as higher tumor differentiation (G1 or G2) and low Ki-67 proliferation index (OR 3.19, 99% CI 1.47-6.92 and OR 3.65, 99% CI 1.72-7.72, respectively, Table 11).

Basal-like samples were associated with lower differentiation (G3; OR 3.43, 95% CI 1.26-9.34, Table 11). A multivariate regression allowed us to confirm the strong correlation (ORs > 20) between the luminal intrinsic subtypes and the HR+/HER2– IHC group, after adjusting for age, tumor size, histological type, tumor grade and Ki-67 proliferation index (Table 11). No further significant associations between intrinsic subtypes and histopathological parameters were found. In Table 11, the multivariate analysis clearly indicated that luminal A subtypes had a statistically significant association with tumor grade 1 to 2, low Ki-67 expression, HR+/HER2- tumors whereas luminal B subtypes had association only with HR+/HER2-. With respect to HER2-enriched tumors, they had a statistically significant association with HR+/HER2+, HR-/HER2+, HR-/HER2- tumors; however, basal like tumors had association with tumor grade 3 and IHC group HR-/HER2-.

The results of a multivariate regression confirmed this, as HER2-enriched samples were significantly associated with a very high probability of being HER2 positive (independent of HR status, ORs > 20, Table 11). HER2-enriched samples were also associated with TNBC, although with a lower odds ratio (OR 6.7, 99% CI 1.77-25.39, Table 11).

The samples classified as Basal-like had an increased probability of being TNBC (46 of 77 samples, 60%), as compared with triple-positive (HR+/HER2+) samples (Table 11).

Table 11: Results of multivariate logistic regression of clinical and histopathological parameters, taken as predictive variables for intrinsic subtypes

Parameters	**Luminal A (n=104)**		**Luminal B (n=91)**		**HER2-enriched (n=62)**		**Basal-like (n=77)**	
	OR (99% CI)	P-value	OR (99% CI)	P-value	OR (99% CI)	P-value	OR (99% CI)	P-value
Age Group (years)								
< 50 (n=201)	1.47 (0.68-3.16)	0.2	0.95 (0.46-1.97)	0.87	0.76 (0.30-1.92)	0.45	1.00 (0.43-2.31)	0.99
≥ 50 (n=95)	Ref		Ref		Ref		Ref	
Tumor Size								
T1 or T2 (n=168)	Ref		Ref		Ref		Ref	
T3 or T4 (n=126)	0.66 (0.30-1.44)	0.17	1.74 (0.84-3.62)	0.05	0.72 (0.28-1.88)	0.38	1.04 (0.45-2.42)	0.89
Histological Type								
NST (n=303)	1.36 (0.39-4.78)	0.53	1.88 (0.46-7.76)	0.25	0.50 (0.11-2.31)	0.24	0.65 (0.16-2.67)	0.43
Non-NST (n=31)	Ref		Ref		Ref		Ref	
Tumor Grade								
G1 or G2 (n=140)	**3.19 (1.47-6.92)**	**1.21×10^{-4}**	Ref		Ref		Ref	
G3 (n=194)	Ref		1.16 (0.52-2.57)	0.63	1.45 (0.50-4.14)	0.37	**3.43 (1.26-9.34)**	**1.54×10^{-3}**
Ki-67 Proliferation index								
Low (n=132)	**3.65 (1.72-7.72)**	**9×10^{-6}**	Ref		Ref		Ref	
High (n=202)	Ref		1.64 (0.76-3.56)	0.1	1.50 (0.54-4.13)	0.31	2.40 (0.95-6.09)	0.02
IHC Group								

HR+/HER2− (n=187)	**20.52 (1.36-309.24)**	**4.12×10^{-3}**	**20.62 (1.43-297.79)**	**3.51×10^{-3}**	Ref		2.88 (0.54-15.32)	0.1
HR+/HER2+ (n=45)	16.30 (0.94-281.70)	0.01	6.80 (0.41-111.73)	0.08	**22.62 (5.86-87.36)**	**2.74×10^{-9}**	Ref	
HR−/HER2+ (n=28)	Ref		Ref		**88.30 (18.12-430.32)**	**3.16×10^{-13}**	3.52 (0.45-27.62)	0.12
HR−/HER2− (n=74)	3.89 (0.22-68.16)	0.22	1.86 (0.10-33.46)	0.58	**6.70 (1.77-25.39)**	**2.33×10^{-4}**	**30.58 (5.49-170.50)**	**2.94×10^{-7}**

**Bold letters: cases with p-values < 0.01. OR: Odds Ratio. CI: Confidence Interval. Ref: Category taken as reference within each parameter for each test. HER2: Human Epidermal Growth Factor Receptor 2; NST: No special type; HR: Hormone Receptor*

5.1.5. IHC types across clinical and histopathological features

Table 12 illustrates the distribution of IHC types within the clinical and histopathological parameters. Two-third of T1 to T2 tumors were tended HR+ /HER2- (61.9%). Of G3 tumors, 45.9% were HR+ /HER2-. Of ER+ and PgR+ tumors, 83.2% and 82.8% were HR+ /HER2-. Among HER2- tumors, 71.6% were HR+ /HER2- whereas within HER2+ tumors, 61.6% were HR+ /HER2+ followed by HR- /HER2+ (38.4%). Within tumors with low Ki-67 proliferation, 65.2% were HR+ /HER2- followed basal-like tumors (19.7%).

Table 12: Distribution of clinical and histopathological parameters among IHC groups

Parameters	**All**	**HR+/HER2− (%)**	**HR+/HER+ (%)**	**HR−/HER+ (%)**	**HR−/HER2− (%)**
	334	187 (55.9)	45 (13.5)	28 (8.4)	74 (22.2)
Age Group (years)					
< 50	201	118 (58.7)	30 (14.9)	14 (7.0)	39 (19.4)
≥ 50	95	53 (55.8)	10 (10.5)	12 (12.6)	20 (21.1)
Unknown	38	16 (42.1)	5 (13.2)	2 (5.3)	15 (39.5)
Tumor Size					
T1 or T2	168	104 (61.9)	21 (12.5)	14 (8.3)	29 (17.3)
T3 or T4	126	65 (51.6)	19 (15.1)	12 (9.5)	30 (23.8)
Unknown	40	18 (45.0)	5 (12.5)	2 (5.0)	15 (37.5)
Histological Type					
NST	303	170 (56.1)	43 (14.2)	24 (7.9)	66 (21.8)
Non-NST	31	17 (54.8)	2 (6.5)	4 (12.9)	8 (25.8)
Tumor Grade					
G1 or G2	140	98 (70.0)	13 (9.3)	9 (6.4)	20 (14.3)
G3	194	89 (45.9)	32 (16.5)	19 (9.8)	54 (27.8)
Estrogen Receptor status					
ER-Positive (≥1)	184	153 (83.2)	31 (16.8)	0 (0.0)	0 (0.0)
ER-negative (<1)	150	34 (22.7)	14 (9.3)	28 (18.7)	74 (49.3)
Progesterone Receptor Status					
PgR-positive (≥1)	157	130 (82.8)	26 (16.6)	0 (0.0)	1 (0.6)
PgR-negative (<1)	177	57 (32.2)	19 (10.7)	28 (15.8)	73 (41.2)
Hormone Receptor Status					
Positive	232	187 (80.6)	45 (19.4)	0 (0.0)	0 (0.0)
Negative	102	0 (0.0)	0 (0.0)	28 (27.5)	74 (72.5)
HER2 Status					
Negative	261	187 (71.6)	0 (0.0)	0 (0.0)	74 (28.4)
Positive	73	0 (0.0)	45 (61.6)	28 (38.4)	0 (0.0)
Ki-67 Proliferation index					
Low (<20)	132	86 (65.2)	9 (6.8)	11 (8.3)	26 (19.7)
High (≥20)	202	101 (50.0)	36 (17.8)	17 (8.4)	48 (23.8)

5.2. Frequency of mutation in *PIK3CA* gene and its prognostic role in Ethiopian breast cancer patients

In this study, beyond determining intrinsic subtypes of Ethiopian BC patients, we profiled *PIK3CA* gene mutation by targeting three hot spot sites using qPCR. *PIK3CA* mutation analysis was carried out in a total of 226 breast tumor tissues.

Table 13 shows the clinical and histopathological features for which *PIK3CA* analysis results are available. The median age of the participants was 40 years and more than half of the patients were less than 50 years. Close to 50% of the patients were presented with early-stage disease (pT1, pT2). Majority of the tumors identified G3 (54.4%). With IHC analysis, 74.3%, 56.6% and 22.6% of the tumors were ER+, PR+ and HER2+ in order. With respect to Ki-67, greater than half of the tumors had high Ki-67 proliferation index. IHC type was defined using HR and HER2 status. Accordingly, significant number of the tumors were HR+/HER2− (62.4%) followed by HR+/HER2+ (15.9%).

Table 13: Patient and tumor histopathological features among Ethiopian Patients

Parameters	**All, N= 226**
Age (years)	
< 50	125 (55.3%)
>= 50	65 (28.8%)
Unknown	36 (15.9%)
Tumor size	
pT1, pT2	103 (45.6%)
>pT2	78 (34.5%)
Unknown	45 (19.9%)
Tumor Grade	
G1 or G2	103 (45.6%)
G3	123 (54.4)
ER Status	
positive (≥ 1%)	168 (74.3%)
negative (<1)	58 (25.7)
PR Status	
positive (≥ 1%)	128 (56.6%)
negative (< 1)	98 (43.4%)
HR Status	
positive	177 (78.3%)
negative	49 (21.7%)
HER2 Status	
negative	175 (77.4%)
positive	51 (22.6%)
Ki-67 proliferation Index	
Low (< 20%)	103 (45.6%)
High (≥ 20%)	123 (54.4%)
IHC group	
HR+/HER2-	141 (62.4%)
HR+/HER2+	36 (15.9%)
HR-/HER2+	15 (6.6%)
HR-/HER2-	34 (15%)
Intrinsic subtypes	
Luminal A	71 (31.4%)
Luminal B	45 (19.9%)
HER2-Enriched	36 (15.9%)
Basal-like	41 (18.1%)
Unknown	33 (14.6%)

5.2.1. Distribution of overall *PIK3CA* gene mutation and its association with clinical and histopathological features

5.2.1.1. Overall mutation

The overall prevalence of *PIK3CA* mutation was 20.4%. Table 14 described the distribution of *PIK3CA* within each clinical and histopathological feature and showed its observed association using binary logistic regression. Higher proportion of *PIK3CA* mutation mainly observed among young (24%), early disease stage (19.4%), and in moderate or poorly differentiated tumors-G1 or G2 (22.3%). High frequency of *PIK3CA* mutation was observed within ER+ (24.4%), PR+ (27.3%), HER2+ (21.6%) and high Ki-67 proliferation index (19.5%). Of the mutated tumors, majority tended HR+ status (> 95%).

The finding explored the frequency of *PIK3CA* mutation across distinct IHC groups. A marked difference was observed in the *PIK3CA* gene mutation frequency across the four major tumor subgroups. The distribution of mutation across IHC group tended 24.1% (HR+/HER2−), 27.8% (HR+/HER2) and a lower proportion accounting 2.9% identified in HR−/HER2− tumors.

5.2.1.2. overall mutation rate association with clinical and histopathological features

In this study, the association of *PIK3CA* with the corresponding clinical and histopathological feature explored (Table 14). Accordingly, favorable clinical and histopathological parameters including ER+ (P-value =0.014), PR+ (P-value= 0.004), HR+ (P-value= 0.006), HR+/HER2- (P-value = 0.023) and HR+/HER2+ (P-value =0.019) had a statistically significant association with overall *PIK3CA* mutation. We didn't observe a statistically significant association other than these parameters and overall mutations including intrinsic subtypes.

Table 15, shows the clinical and histopathological features of *PIK3CA* mutated tumors. A table analogous to Table 14, but with percentages given vertically (for each category in each variable) is given in Table 15. In the finding, the majority of the mutated tumors were ER+ (89.1%%) and PR+ (76.1%) whereas most of the mutated tumors were turned HER2− (76.1%). Out of the mutated tumors, 73.9% and 21.7% were HR+/HER2− and HR+/HER2+, respectively whereas 2.2% of the mutated tumors were either HR−/HER2+ or HR−/HER2− . In this study, most of the mutated tumors had a high Ki-67 proliferation index (52.2%).

Table 14: Overall *PIK3CA* mutation and its association with clinical and histopathological parameters among Ethiopian patients with BC

Parameters		***PIK3CA* mutation status**		**Logistic regression**	
	All N=226 (100%)	wild type N=180 (79.6%)	mutated N=46 (20.4%)	univariate OR (95% CI)	P-value
Age (years)					
< 50	125 (100%)	95 (76%)	30 (24%)	2.250 (0.965-5.244)	0.06
≥ 50	65 (100%)	57 (87.7%)	8 (12.3%)	Ref.	
Unknown	36 (100%)	28 (77.8%)	8 (22.2%)	2.036 (0.692-5.991)	0.197
Tumor Size					
pT1, pT2	103 (100%)	83 (80.6%)	20 (19.4%)	1.012 (0.480-2.132)	0.975
> pT2	78 (100%)	63 (80.8%)	15 (19.2%)	Ref.	
Unknown	45 (100%)	34 (75.6%)	11 (24.4%)	1.359 (0.562-3.285)	0.496
Tumor Grade					
G1 or G2	103 (100%)	80 (77.7%)	23 (22.3%)	1.250 (0.654-2.391)	0.5
G3	123 (100%)	100 (81.3)	23 (18.7%)	Ref.	
ER Status					
positive (≥ 1%)	168 (100%)	127 (75.6%)	41 (24.4)	3.422 (1.282-9.138)	**0.014***
negative (< 1)	58 (100%)	53 (91.4%)	5 (8.6%)	Ref.	
PR Status					
positive (≥ 1%)	128 (100%)	93 (72.7%)	35 (27.3%)	2.977 (1.423-6.225)	**0.004***
negative (< 1)	98 (100%)	87 (88.8%)	11 (11.2%)	Ref.	
HR Status					
positive	177 (100%)	133 (75.1%)	44 (24.9%)	7.774 (1.814-33.328)	**0.006***
negative	49 (100%)	47 (95.9%)	2 (4.1%)	Ref.	
HER2 Status					
negative	175 (100%)	140 (80%)	35 (20%)	1.100 (0.513-2.360)	0.807
positive	51 (100%)	40 (78.4%)	11 (21.6%)	Ref.	
Ki-67 proliferation index					
Low (< 20%)	103 (100%)	81 (78.6%)	22 (21.4%)	1.120 (0.586-2.143)	0.731
High (≥ 20%)	123 (100%)	99 (80.5%)	24 (19.5%)	Ref.	
IHC group					
HR+/HER2-	141 (100%)	107 (75.9%)	34 (24.1%)	10.486 (1.382-79.563)	**0.023**
HR+/HER2+	36 (100%)	26 (72.2%)	10 (27.8%)	12.692 (1.525-105.626)	**0.019**
HR-/HER2+	15 (100%)	14 (93.3%)	1 (6.7%)	2.357 (0.138-40.402)	0.554
HR-/HER2-	34 (100%)	33 (97.1%)	1 (2.9%)	Ref.	
PAM50 intrinsic subtypes					
Luminal A	71 (100%)	55 (77.5%)	16 (22.5%)	2.691 (0.833-8.689)	0.098
Luminal B	45 (100%)	34 (75.6%)	11 (24.4%)	2.993 (0.870-10.295)	0.082
HER2-Enriched	36 (100%)	27 (75%)	9 (25%)	3.083 (0.859-11.068)	0.084
Basal-like	41 (100%)	37 (90.2%)	4 (9.8%)	Ref.	
Unknown	33 (100%)	27(81.8%)	6 (18.2%)	2.056 (0.528-8.000)	0.299

** And bold: Parameters for which a P-value (from a $\chi 2$ test for independence) below 5% was observed,showing dependency between PIK3CA mutation status and histopathological parameters; CI: Confidence interval; HER2: Human Epidermal Growth Factor Receptor 2; ER: Estrogen Receptor; PgR: Progesterone Receptor; HR: Hormone Receptor; OR: Odds ratio; TNBC: triple-negative breast cancer*

Table 15: Overall *PIK3CA* mutation and its association with clinical and histopathological parameters

Parameters		***PIK3CA* mutation status**		**Logistic regression**	
	All	wild type	mutated		
	N=226 (100%)	N=180 (79.6%)	N=46 (20.4%)	univariate OR (95% CI)	P-value
Age (years)					
< 50	125 (55.3%)	95 (52.8%)	30 (65.2%)	2.250 (0.965-5.244)	0.06
≥ 50	65 (28.8%)	57 (31.7%)	8 (17.4%)	Ref.	
unknown	36 (15.9%)	28 (15.6%)	8 (17.4%)	2.036 (0.692-5.991)	0.197
Tumor Size					
pT1, pT2	103 (45.6%)	83 (46.1%)	20 (43.5%)	1.012 (0.480-2.132)	0.975
> pT2	78 (34.5%)	63 (35%)	15 (32.6%)	Ref.	
unknown	45 (19.9%)	34 (18.9%)	11 (23.9%)	1.359 (0.562-3.285)	0.496
Tumor Grade					
G1 or G2	103 (45.6%)	80 (44.4%)	23 (50%)	1.250 (0.654-2.391)	0.5
G3	123 (54.4%)	100 (55.6%)	23 (50%)	Ref.	
ER Status					
positive (≥ 1%)	168 (74.3%)	127 (70.6%)	41 (89.1%)	3.422 (1.282-9.138)	**0.014***
negative (< 1)	58 (25.7%)	53 (29.4%)	5 (10.9%)	Ref.	
PR Status					
positive (≥ 1%)	128 (56.6%)	93 (51.7%)	35 (76.1%)	2.977 (1.423-6.225)	**0.004***
negative (< 1)	98 (43.4%)	87 (48.3%)	11 (23.9%)	Ref.	
HR Status					
positive	177 (78.3%)	133 (73.9%)	44 (95.7%)	7.774 (1.814-33.328)	**0.006***
negative	49 (21.7%)	47 (26.1%)	2 (4.3%)	Ref.	
HER2 Status					
negative	175 (77.4%)	140 (77.8%)	35 (76.1%)	1.100 (0.513-2.360)	0.807
positive	51 (22.6%)	40 (22.2%)	11 (23.9%)	Ref.	
Ki-67 proliferation index					
Low (< 20%)	103 (45.6%)	81 (45%)	22 (47.8%)	1.120 (0.586-2.143)	0.731
High (≥ 20%)	123 (54.4%)	99 (55%)	24 (52.2%)	Ref.	
IHC group					
HR+/HER2-	141 (62.4%)	107 (59.4%)	34 (73.9%)	10.486 (1.382-79.563)	**0.023**
HR+/HER2+	36 (15.9%)	26 (14.4%)	10 (21.7%)	12.692 (1.525-105.626)	**0.019**
HR-/HER2+	15 (6.6%)	14 (7.8%)	1 (2.2%)	2.357 (0.138-40.402)	0.554
HR-/HER2-	34 (15.0%)	33 (18.3%)	1 (2.2%)	Ref.	
PAM50 intrinsic subtypes					
Luminal A	71 (31.4%)	55 (30.6%)	16 (34.8%)	2.691 (0.833-8.689)	0.098
Luminal B	45 (19.9%)	34 (18.9%)	11 (23.9%)	2.993 (0.870-10.295)	0.082
HER2-Enriched	36 (15.9%)	27 (15%)	9 (19.6%)	3.083 (0.859-11.068)	0.084
Basal-like	41 (18.1%)	37 (20.6%)	4 (8.7%)	Ref.	
unknown	33 (14.6%)	27 (15%)	6 (13%)	2.056 (0.528-8.000)	0.299

CI: Confidence interval; HER2: Human Epidermal Growth Factor Receptor 2; ER: Estrogen Receptor; PgR: Progesterone Receptor; HR: Hormone Receptor; OR: Odds ratio; TNBREAST CANCER triple-negative breast cancer

5.2.2. Exon specific mutation and its association with clinical and histopathological parameters

5.2.2.1. Exon specific PIK3CA mutation

In this study, we assessed exon specific distribution of *PIK3CA* mutation across clinical and histopathological parameters (Table 16). With a Chi-square analysis, ER (P-value= 0.01), PR (P-value= 0.003) and HR (P-value= 0.001) positive tumors had a statistically significant association with overall *PIK3CA* mutations. Furthermore, HR+/HER2- tumors (P-value=0.014) had a statistically significant association with overall *PIK3CA* mutations. The mutation rates for the three hot spot sites were 0.4% for COSMIC C760 (E542K), 2.2% for COSMIC 763 (E545K) and 17.7% for COSMIC C775 (H1047R). In this study, we didn't find co-occurrence of hot spot mutations along *PIK3CA* gene region across in each sample. Tumor samples harboring only mutation across a single hot spot region accounts 20.4%.

5.2.2.2. Exon specific mutation association with clinical and histopathological features

We did a Chi-square analysis to show the association of clinical and histopathological features with specific mutation across *PIK3CA* gene. Accordingly, we observed that ER+ (P-value = 0.015), PR+ (P-value = 0.01), HR+ (P-value = 0.006) and HR+/HER2- (P-value= 0.027) tumors had a statistically significant association with exon 20 mutation (Table 16). However, a statistically significant association between exon 20 mutation and the corresponding features not observed.

Table 16: Exon specific *PIK3CA* mutations across clinical and histopathological parameters among primary BC patients in Ethiopia

Parameters		**Exon 9 mutation status**			**Exon 20 mutation status**		
	All N=226 (100%)	wild type N=220 (97.3%)	mutated N=6 (2.7%)	P-value	wild type N=187 (82.7%)	mutated N=39 (17.3%)	P-value
Age (years)							
< 50	125 (100%)	120 (96%)	5 (4%)	0.338	101 (80.8%)	24 (19.2%)	0.238
≥ 50	65 (100%)	64 (98.5%)	1 (1.5%)		58 (89.2%)	7 (10.8%)	
unknown	36 (100%)	36 (100%)	0 (0%)		28 (77.8%)	8 (22.2%)	
Tumor Size							
pT1, pT2	103 (100%)	100 (97.1%)	3 (2.9%)	0.431	87 (84.5%)	16 (15.5%)	0.362
> pT2	78 (100%)	75 (96.2%)	3 (3.8%)		66 (84.6%)	12 (15.4%)	
unknown	45 (100%)	45 (100%)	0(0%)		34 (75.6%)	11 (24.4%)	
Tumor Grade							
G1 or G2	103 (100%)	99 (96.1%)	4 (3.9%)	0.293	84 (81.6%)	19 (18.4%)	0.665
G3	123 (100%)	121 (98.4%)	2 (1.6%)		103 (83.7%)	20 (16.3%)	
ER Status							
positive (≥ 1%)	168 (100%)	163 (97%)	5 (3%)	0.609	133 (79.2%)	35 (20.8%)	**0.015***
negative (<1)	58 (100%)	57 (98.3%)	1 (1.7%)		54 (93.1%)	4 (6.9%)	
PR Status							
positive (≥ 1%)	128 (100%)	124 (96.9%)	4 (3.1%)	0.615	98 (76.6%)	30 (23.4%)	**0.01***
negative (< 1)	98 (100%)	96 (98)	2 (2%)		89 (90.8%)	9 (9.2%)	
HR Status							
positive	177 (100%)	171 (96.6%)	6 (3.4%)	0.191	140 (79.1%)	37 (20.9%)	**0.006***
negative	49 (100%)	49 (100%)	0 (0%)		47 (95.9%)	2 (4.1%)	
HER2 Status							
negative	175 (100%)	169 (96.6%)	6 (3.4%)	0.18	147 (84%)	28 (16%)	0.354
positive	51 (100%)	51 (100%)	0 (0%)		40 (78.4%)	11 (21.6%)	
Ki-67 proliferation index							
Low (< 20%)	103 (100%)	99 (96.1%)	4 (3.9%)	0.293	85 (82.5%)	18 (17.5%)	0.936
High (≥ 20%)	123 (100%)	121 (98.4%)	2 (1.6%)		102 (82.9%)	21 (17.1%)	
IHC group							
HR+/HER2-	141 (100%)	135 (95.7%)	6 (4.3%)	0.294	114 (80.9%)	27 (19.1%)	0.027
HR+/HER2+	36 (100%)	36 (100%)	0 (0%)		26 (72.2%)	10 (27.8%)	
HR-/HER2+	15 (100%)	15 (100%)	0 (0%)		14 (93.3%)	1 (6.7%)	
HR-/HER2-	34 (100%)	34(100%)	0 (0%)		33 (97.1%)	1 (2.9%)	
PAM50 intrinsic subtypes							
Luminal A	71 (100%)	68 (95.8%)	3 (4.2%)	0.764	58 (81.7%)	13 (18.3%)	0.451
Luminal B	45 (100%)	44 (97.8%)	1 (2.2%)		35 (77.8%)	10 (22.2%)	
HER2-Enriched	36 (100%)	35 (97.2%)	1 (2.8%)		28 (77.8%)	8 (22.2%)	
Basal-like	41 (100%)	41 (100%)	0 (0%)		37 (90.2%)	4 (9.8%)	
unknown	33 (100%)	32 (97%)	1(3%)		29 (87.9%)	4 (12.1%)	

In table 17, exon specific mutated tumors turned within in favourable clinical and histopathological features.

Table 17: Overall *PIK3CA* mutation status and exon specific *PIK3CA* mutations across clinical and histopathological parameters

Parameters		**Exon 9 mutation status**			**Exon 20 mutation status**		
	All N=226 (100%)	wild type N=220 (97.3%)	mutated N=6 (2.7%)	P-value	wild type N=187 (82.7%)	mutated N=39 (17.3%)	P-value
Age (years)							
< 50	125 (55.3%)	120 (54.5%)	5 (83.3%)	0.338	101(54%)	24(61.5%)	0.238
≥ 50	65 (28.8%)	64 (29.1%)	1 (16.7%)		58(31%)	7(17.9%)	
unknown	36 (15.9%)	36 (16.4%)	0 (0%)		28(15%)	8(20.5%)	
Tumor Size							
pT1, pT2	103 (45.6%)	100 (45.5%)	3 (50%)	0.431	87(46.5%)	16(41%)	0.362
> pT2	78 (34.5%)	75 (34.1%)	3 (50%)		66 (35.3%)	12 (30.8%)	
unknown	45 (19.9%)	45 (20.5%)	0 (0%)		34 (18.2%)	11 (27.2%)	
Tumor Grade							
G1 or G2	103 (45.6%)	99 (45%)	4 (66.7%)	0.293	84 (44.9%)	19 (48.7%)	0.665
G3	123 (54.4%)	121 (55%)	2 (33.3%)		103 (55.1%)	20 (51.3%)	
ER Status							
positive (≥ 1%)	168 (74.3%)	163 (74.1%)	5 (83.3%)	0.609	133 (71.1%)	35 (89.7%)	**0.015***
negative (<1%)	58 (25.7%)	57 (25.9%)	1 (16.7%)		54 (28.9%)	4 (10.3%)	
PR Status							
positive (≥1%)	128 (56.6%)	124 (56.4%)	4 (66.7%)	0.615	98 (52.4%)	30 (76.9%)	**0.005***
negative (< 1%)	98 (43.4%)	96 (43.6%)	2 (33.3%)		89 (47.6%)	9 (23.1%)	
HR Status							
positive	177 (78.3%)	171 (77.7%)	6 (100%)	0.191	140 (74.9%)	37 (94.9%)	**0.006***
negative	49 (21.7%)	49 (22.3%)	0 (0%)		47 (25.1%)	2 (5.1%)	
HER2 Status							
negative	175 (77.4%)	16. 9(76.8%)	6 (100%)	0.18	147 (78.6%)	28 (71.8%)	0.354
positive	51 (22.6%)	51 (23.2%)	0 (0%)		40 (21.4%)	11 (28.2%)	
Ki-67 proliferation index							
Low (< 20%)	103 (45.6%)	99 (45%)	4 (66.7%)	0.293	85 (45.5%)	18 (46.2%)	0.936
High (≥ 20%)	123 (54.4%)	121 (55%)	2 (33.3%)		102 (54.5%)	21 (53.8%)	
IHC group							
HR+/HER2-	141 (62.4%)	135 (61.4%)	6(100%)	0.294	114 (61%)	27 (69.2%)	**0.027**
HR+/HER2+	36 (15.9%)	36 (16.4%)	0 (0%)		26 (13.9%)	10 (25.6%)	
HR-/HER2+	15 (6.6%)	15 (6.8%)	0 (0%)		14 (7.5%)	1 (2.6%)	
HR-/HER2-	34 (15.0%)	34 (15.5%)	0 (0%)		33 (17.6%)	1 (2.6%)	
PAM50 intrinsic subtypes							
Luminal A	71 (31.4%)	68 (30.9%)	3 (4.2%)	0.764	58 (31%)	13 (33.3%)	0.451
Luminal B	45 (19.9%)	44 (20%)	1 (2.2%)		35 (18.7%)	10 (25.6%)	
HER2-Enriched	36 (15.9%)	35 (15.9%)	1 (2.8%)		28 (15%)	8 (20.5%)	
Basal-like	41 (18.1%)	41 (18.6%)	0 (0%)		37 (19.8%)	4 (10.3%)	
unknown	33 (14.6%)	32 (14.5%)	1 (16.7%)		32 (14.5%)	1 (16.7%)	

5.2.2. *PIK3CA* gene mutation and prognosis

5.2.2.1. Overall survival for the cohort

We performed survival analysis among 150 pathologically confirmed BC patients considering death status as endpoint measure. Kaplan-Meier estimates and Cox (multivariate) hazard risk models were used to evaluate the survival of patients with BC as well as the influence of certain variables upon survival. Close to 70% of the patients managed to survive in the first 36 months of the follow up period (Figure 29). Though we did not notice statistical significance, the KM curve indicated that the number of death of patients with BC increased as time goes by.

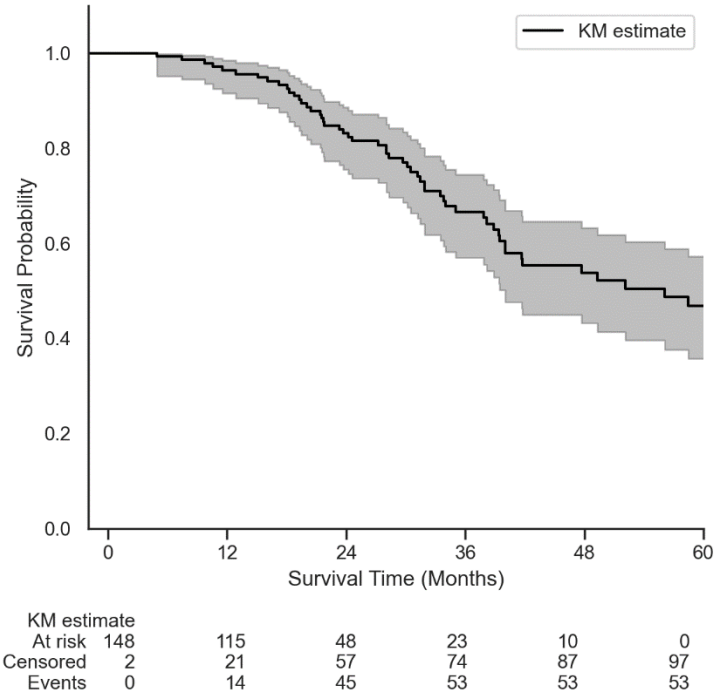


Figure 29: Kaplan Meier curve illustrating OS status for the overall cohort (N=150)

5.2.2.2. Overall survival and PIK3CA mutation

We also tested whether *PIK3CA* mutations were associated with patient OS. While observing the impact of *PIK3CA* mutation on the probability OS, the KM plot indicated that patients with mutated *PIK3CA* gene had a better survival overtime though we found no evidence of an impact of *PIK3CA* mutation on OS (P-value from the log rank test=0.98, visual representation given in Figure 30).

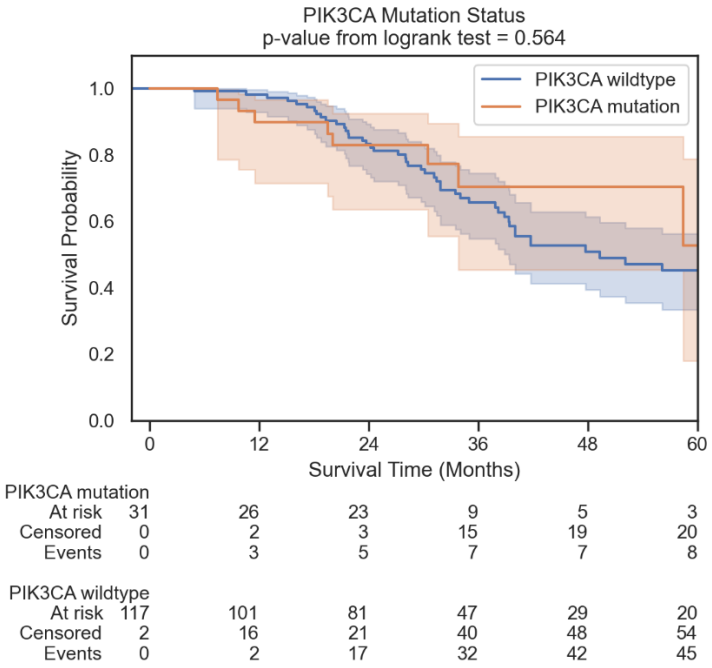


Figure 30: OS of BC patients with respect to mutation status of *PIK3CA* gene

5.2.2.3. PIK3CA mutation influence on OS with respect to IHC and Ki-67 status

In Figure 31 A to H, the influence of *PIK3CA* mutation on patient OS in relation to specific clinical and histopathological features observed. Accordingly, the effect of *PIK3CA* mutation profile on OS of the patient in ER-Vs ER+ (P-value from log rank test= 0.112 Vs 0.375), HR-Vs HR+ (P-value from log rank test= 0.124 Vs 0.51), HER2-Vs HER2+ (P-value from log rank test= 0.439 Vs 0.778) and low Ki-67 Vs high Ki-67 (P-value from log rank test= 0.755 Vs 0.22) tumors demonstrated that there was no significant difference. Though the association did not reach a statistical significance, *PIK3CA* mutated tumors coupled with favorable clinical and histopathological parameters seemingly have a better OS survival.

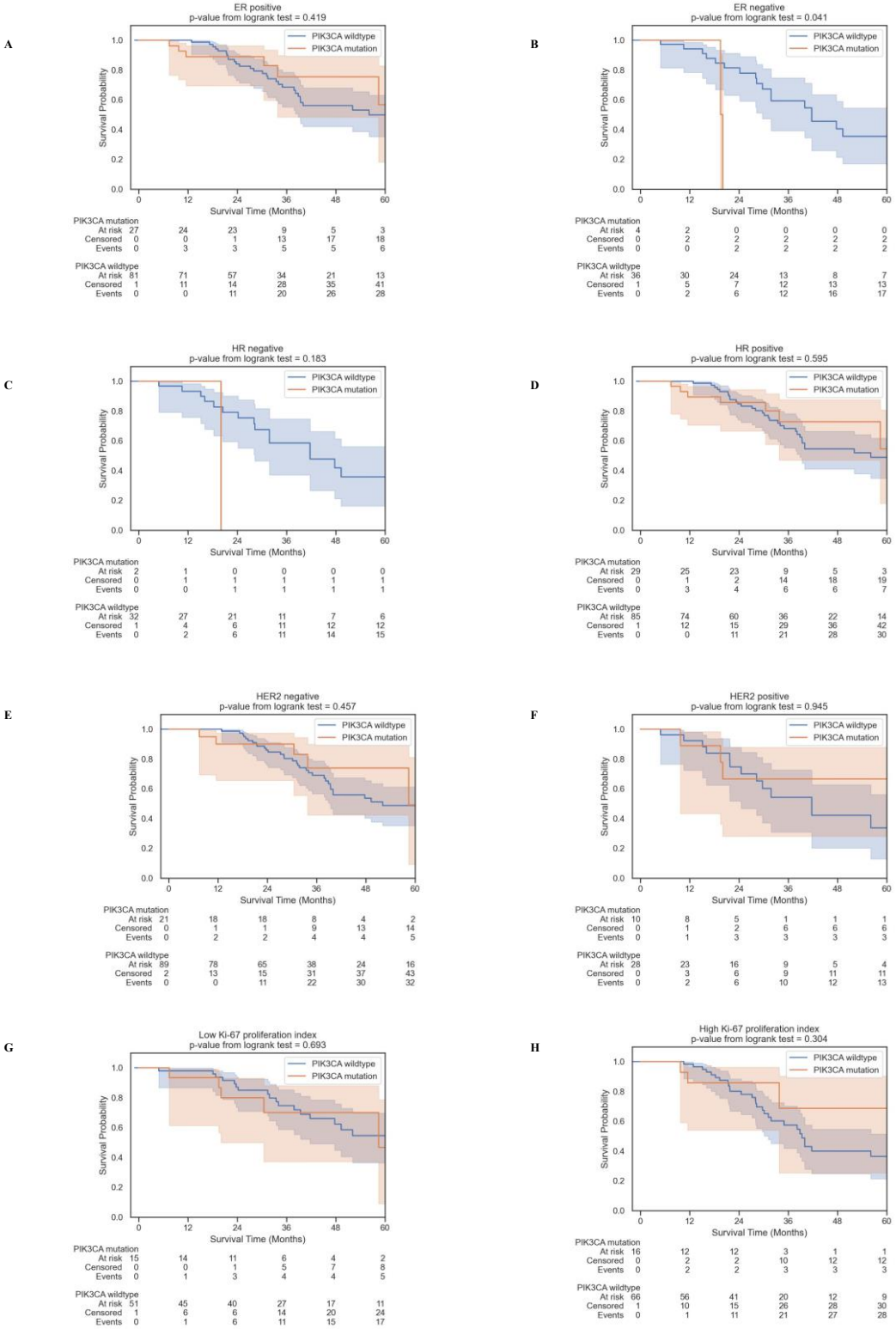


Figure 31: *PIK3CA* gene mutation impact on OS stratified by distinct clinical and histopathological parameters *(Figure 25 A to H shows PIK3CA mutation impact on OS among patients with the specific clinical and histopathological parameters in order (ER, HR, HER2 and Ki-67)*

However, in Cox's multivariate regression analysis, advanced pT stage (HR= 1.81; 95% CI = 0.99 to 3.29; P-value= 0.05) and IHC group HR−/HER2+ (HR= 3.01; 95% CI = 1.38 to 6.55; P-value= 0.02) tumors were independent predictors of impaired survival as compared to other clinical and histopathological parameters (Figure 32).

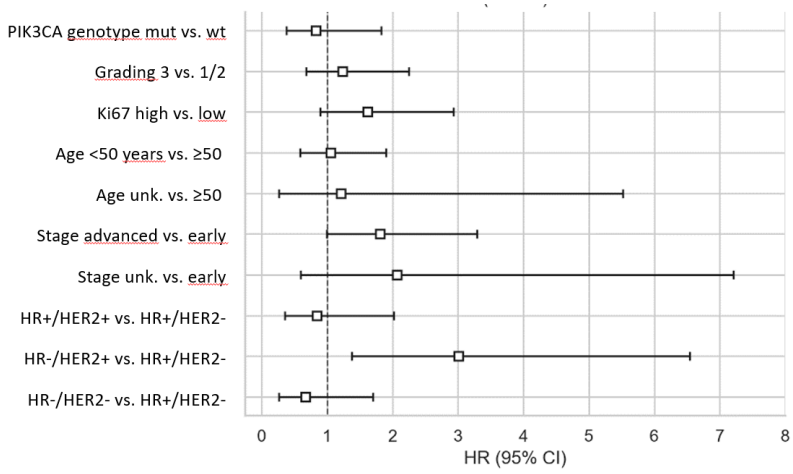


Figure 32: Multivariate Cox's regression analysis including clinical, histopathological parameters and *PIK3CA* genotype (N=150) (**mut**: mutated; **wt**: wildtype; **unk**: unknown)

5.3. Evaluation of microbiota signatures among Ethiopian breast cancer patients

In this study, after characterizing the tumor by using the PAM50 assay and analyzing the *PIK3CA* mutations rates, we profiled the microbiome at the tumor and normal adjust breast tissues. In this project, we compared the difference in microbiota abundance in paired breast tissue samples (tumor versus normal adjacent tissue). Furthermore, we have seen the association of microbiota composition with respect to clinical and histopathological parameters. This study aimed to enhance our understanding of mammary microbiota abundance, composition and its role in patients with BC. A total of 100 breast tissue samples constituted of fresh frozen NAT and tumor breast tissue specimens from 50 Ethiopian women processed to determine relative abundance by sequencing targeting the V4 hypervariable region of 16S rRNA amplicon.

5.3.1. Microbiome study cohort characteristics

Table 18 depicts the clinical and histopathological features of patients with BC which are included for microbiome study. The mean age of patients included was 46 ± 1.92 years. Of the total (n= 50), two-third were premenopausal and half of the patients were with early-stage BC (T1, T2) at the time of diagnosis. Thirty four percent of patients with BC had an advanced disease stage (T3, T4). Majority of the tumors were G3 (52%) and HR+/HER2- (60%).

5.3.2. Taxonomic agglomeration and/or transformation

In the present study, we did taxonomic agglomeration at genus level or the height of the rank for illustration/grouping purpose. Figure 33 shows displayed the tree structure before and after agglomeration at genus level.

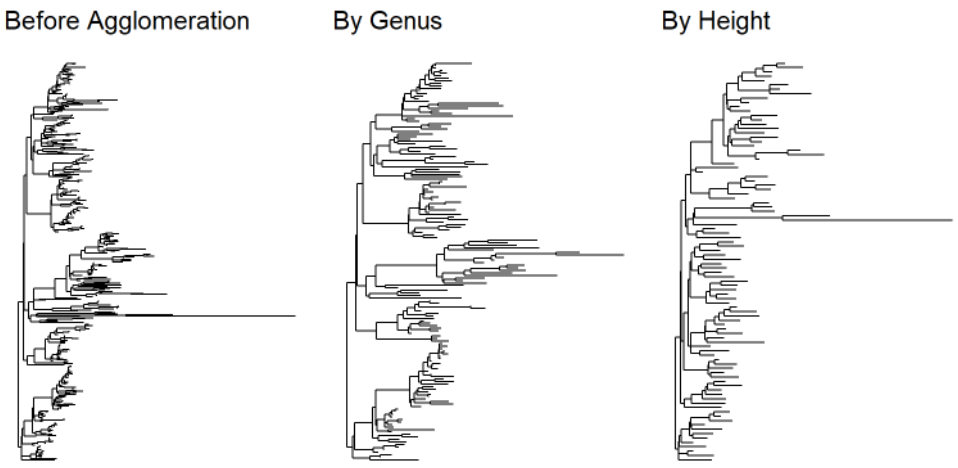


Figure 33: Tree structures for taxonomic agglomeration

Table 18: Histopathological parameters and patients with BC features included for microbiome analyses

Variable	Frequency (%)
Mean age (years ± SE)	46 ± 1.92
Minimum	27
Maximum	83
UICC stage*	
Early (pT1, pT2)	25 (50)
Advanced (>pT2)	17 (34)
Unknown	8 (16)
Grade	
Grade 1 and 2	24 (48)
Grade 3	26 (52)
ER status	
Positive (≥ 1%)	44 (88)
Negative (< 1%)	6 (12)
PR status	
Positive (≥ 1%)	29 (58)
Negative (< 1%)	21 (42)
HER2 status	
Positive	17 (34)
Negative	33 (66)
Ki-67 proliferation index	
Low (≤20%)	16 (32)
High (> 20%)	34 (68)
IHC type (IHC)	
HR+/HER2 -	30 (60)
HR+/HER2+	14 (28)
HR-/HER2+	3 (6)
HR-/HER2-	3 (6)

** Denotes simplified tumor stage taking into account only pathological tumor size; ; ER: estrogen receptor; PR: progesterone receptor; HER2: human epidermal growth factor 2 receptor; HR: hormone receptor; IHC: immunohistochemistry.*

5.3.3. Breast tumor tissue exhibits distinct microbiome composition from NAT of the same women

5.3.3.1. NAT versus breast tumor read counts

We sequenced 16Sv4 amplicons generated from human breast tissue samples on a MiSeq. MiSeq-generated Fastq files were quality-filtered and clustered into 97% similarity operational taxonomic unit (OTUs) using the mothur software package [http://www.mothur.org]. After quality-filtering and inferring amplicon sequence variants (ASVs), quality-filtered reads accounting for an average of 1569 were generated per sample and used for further analysis.

Rarefaction curves relating number of sequencing reads compared to the number of ASVs or genera are shown in Figure 34 A and B.

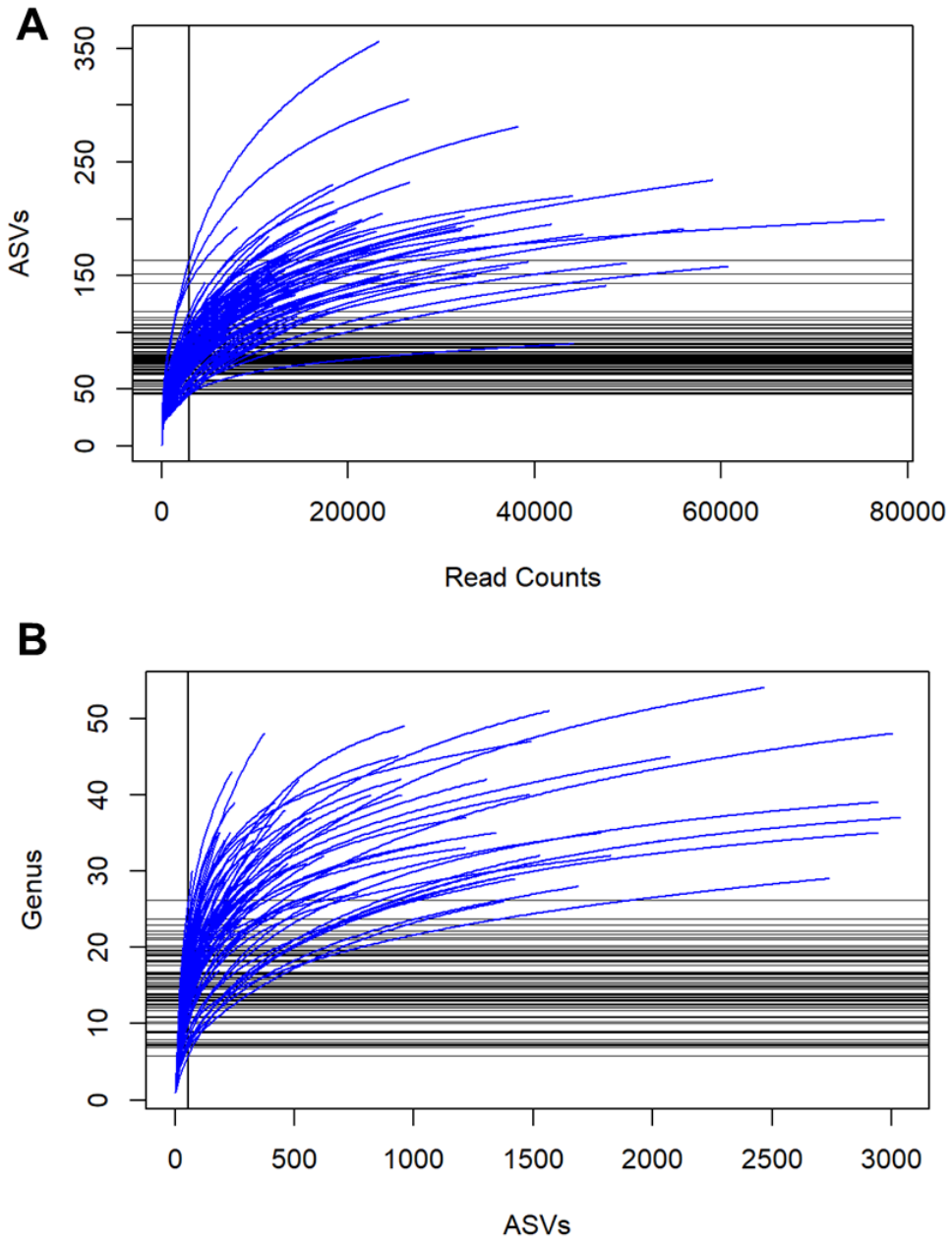


Figure 34: Graphs show rarefaction curves from human breast tissue samples. ***(A)*** *Rarefaction curve relating the number of sequencing reads compared to the number of amplicon sequence variants (ASVs).* ***(B)*** *Rarefaction curve relating the number of ASVs compared to the number of microbial genera*

In this study 17 breast tissue samples have been dropped from the analysis due to low read counts. Taxonomical annotation was carried out and high-quality reads were classified using Silva v. 138 as the reference database. Sequencing quality for R1 and R2 was determined using FastQC 0.11.5. These reads reflect the total number of high-quality sequences that align with 16Sv4, clustered into OTUs and were assigned taxonomic classification.

Any ambiguous or low-quality data were discarded from the subsequent analyses. As shown in Figure 35 higher read count has been recorded in NAT than breast tumor tissues.

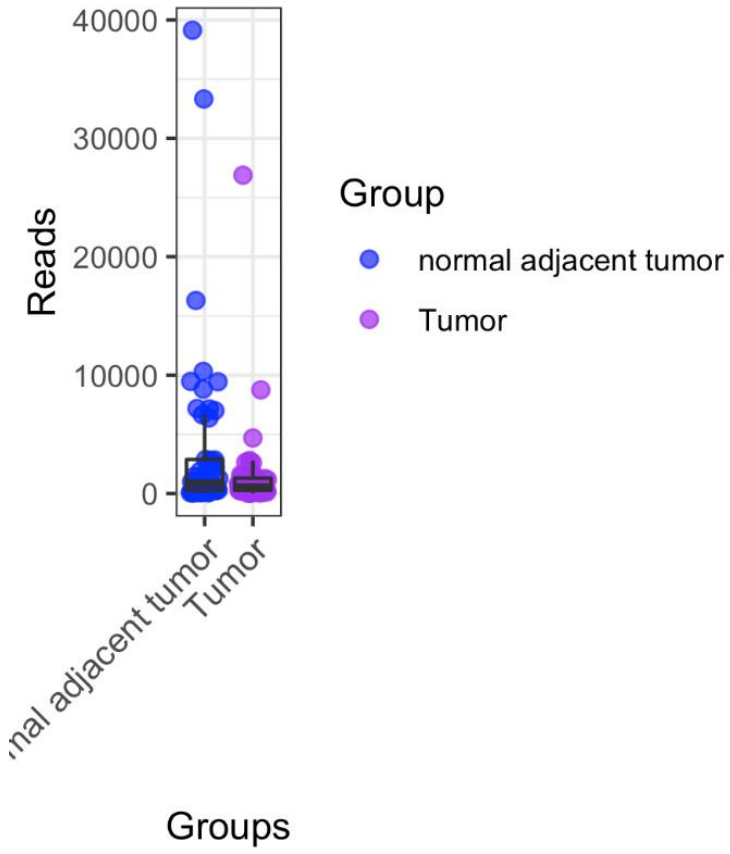


Figure 35: Box-and-whisker plot illustrating the total number of qualities filtered per sample

5.3.4. Taxonomic distribution

5.3.4.1. Overall microbiota distribution across breast tissues

OTUs aggregated into each taxonomic rank, and plotted the relative abundance of the most abundant ones. The 16S rRNA based sequencing identified 5 phyla, 7 classes, 16 orders, 16 families, and 16 genera across all the breast tissue samples. With a closer view, the predominant phylum in the overall breast tissues tended to be *Proteobacteria* (48.38% on average), *Firmicutes* (22.10%), *Actinobacteriota* (14.95% on average) in their order. Figure 36 shows Phylum level microbiota total abundance across all BC tissue samples.

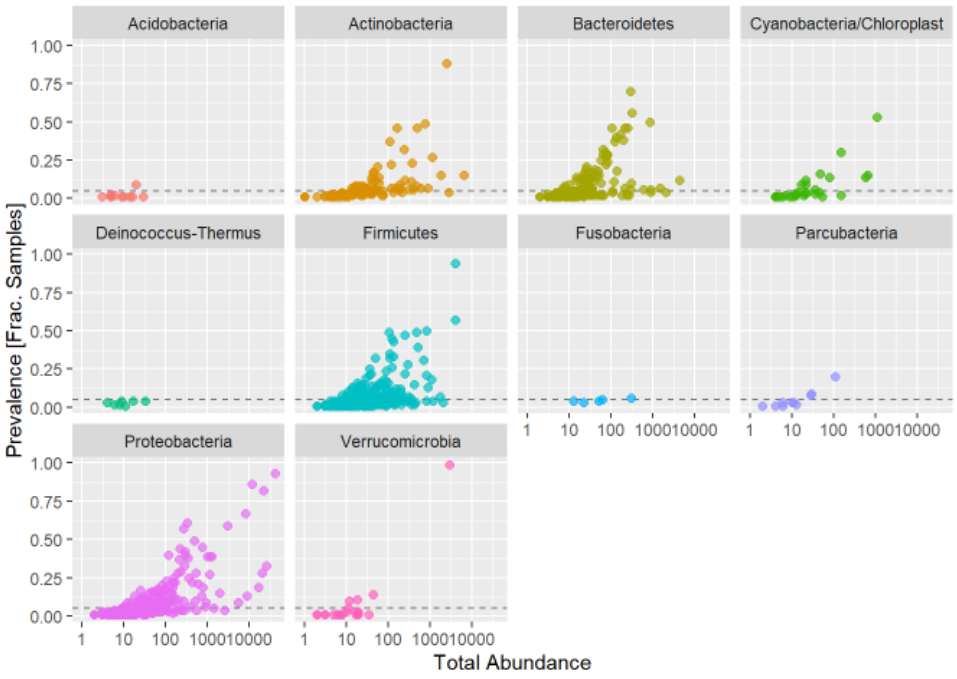


Figure 36: Phylum level microbiota total abundance across all BC tissue samples

5.3.4.2. Microbiota differ between adjacent normal and breast tumor tissue of the same individual

The aggregated taxa were visualized at each taxonomic rank using taxonomic bar plots. The unfilled portion of the bar plots represent lower-abundance taxa. We then determined the taxonomic composition and abundance of breast microbial profiles of tumor and NAT tissues. Phylum abundances was correlated between tumor and normal tissue. In this study, the relative abundances of the most abundant phyla and genera appear to be quite different between NAT and breast tumor tissues.

At phylum level, as shown in Figure 37, the predominant bacteria were Proteobacteria followed by *Firmicutes* in tumor sample. Similarly, NAT samples had a higher proportion of the *Proteobacteria* phylum followed by *Firmicutes*. In both breast tissue types, the three predominant phyla were *Proteobacteria* (48.4%), followed by *Firmicutes* (22.1%) and *Actinobacteria* (15.0%) as shown in Figure 37.

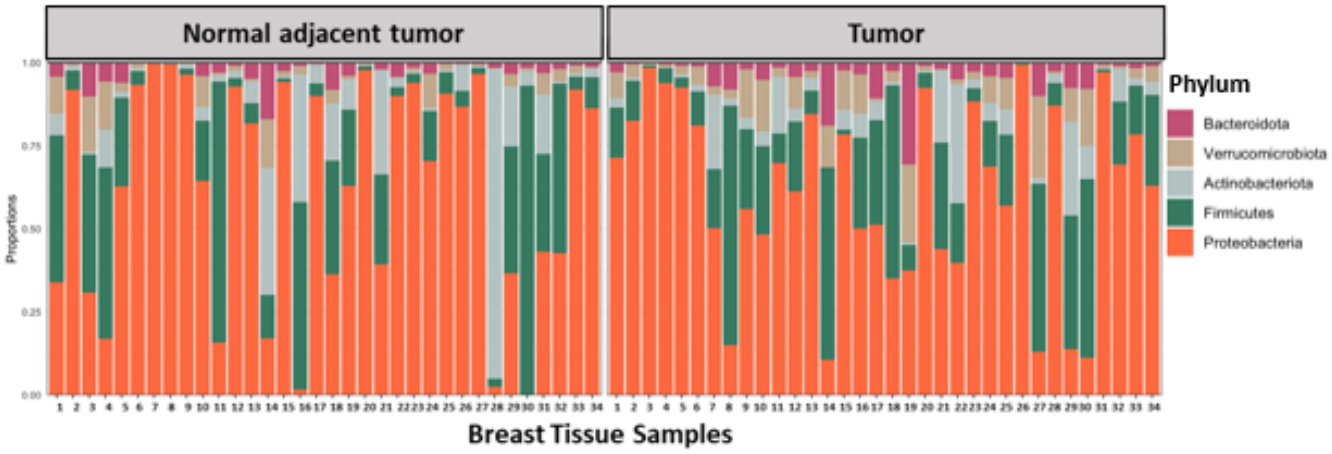


Figure 37: Taxonomic composition of the breast microbiome, *depicted as average relative abundances at the phylum for normal and tumor tissue are shown per each enrolled subject as a stacked bar-plot*

Referring class, *Alphaproteobacteria* (41.5%), *Bacilli* (15.2%), *Actinobacteria* (15.0%), and *Bacteroidia* (7.33%) were the most abundant among all tumor and NAT tissue samples (Figure 38A). Considering all breast tissue samples, the most prevalent families were *Sphingomonadaceae* (20.78%), *Staphylococcaceae* (13.2%), *Beijerinckiaceae* (13.1%), and *Corynebacteriaceae* (9.44%) (Figure 38B). *Sphingomonas* (20.54%), *Staphylococcus* (13.18%), *Methylobacterium-Methylorubrum* (13.09%), and *Corynebacterium* (9.44%) were identified as the most predominant genera (Figure 38C). We also observed that 64.1% of microbial genera were shared between NAT and tumor groups. However, 26.1% of genera were unique to NAT tissues, whereas approximately 10% were unique to tumors.

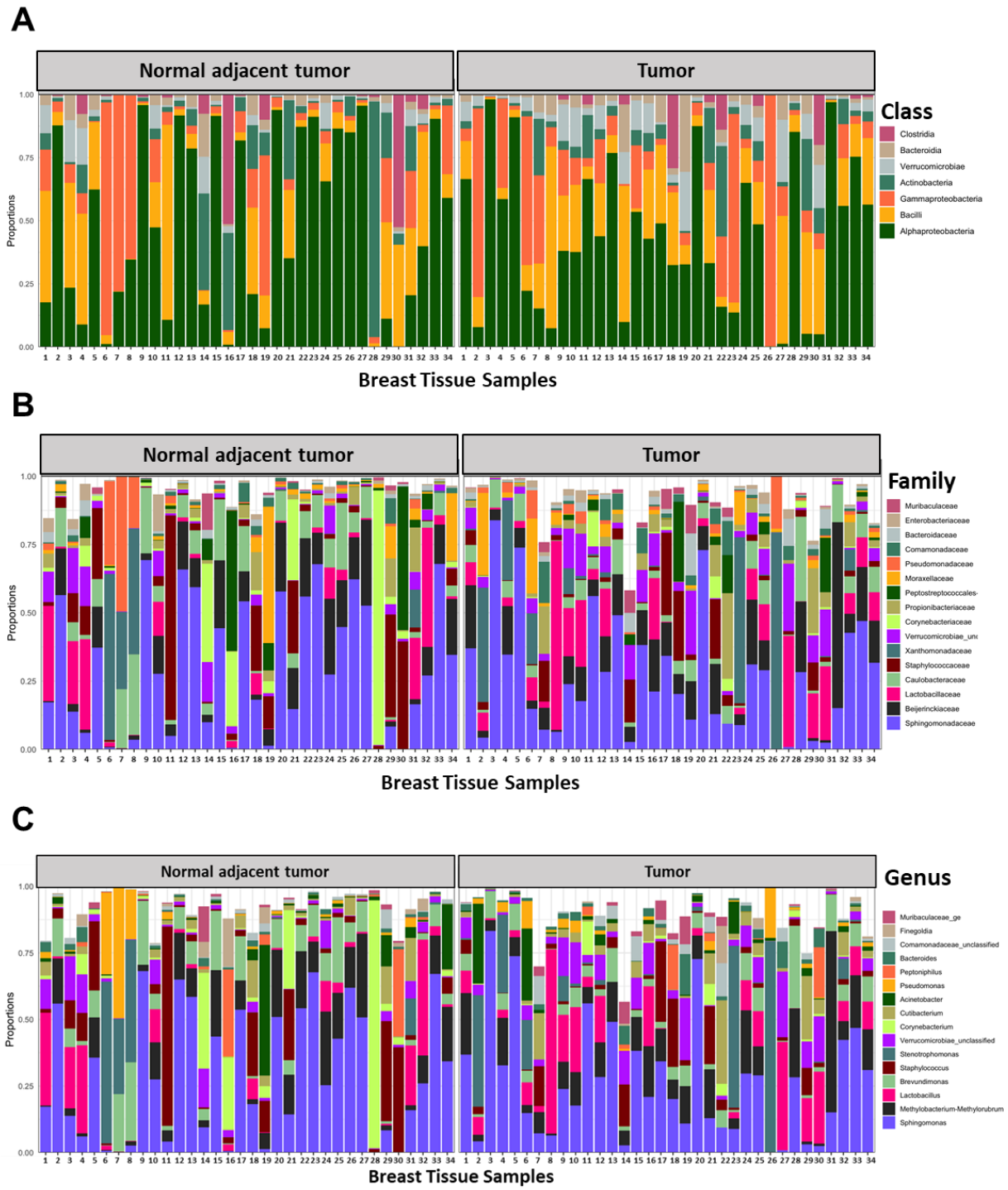


Figure 38: Bar plots illustrating taxonomic composition of the breast microbiome, *the relative abundances of microbiota that differ between NAT and breast tumor tissues by (A) class, (B) family, and (C) genus taxonomic levels. The unfilled portion of the bar plots represent lower-abundance taxa*

5.3.5. Differential genus abundance in tumor relative to NAT

We observed the differential abundance of microbiota at genus level in tumor relative to NAT (Table 19). Paired Wilcox signed rank test was done to see whether microbiota at genus level differ in tumor relative to NAT. Accordingly, a total of 18 distinct taxa had a statistically significant difference in tumor relative to normal tissue at genus level (P-value < 0.05) (Table 19).

Significantly apparent different taxonomic microbial composition between NAT and tumor tissue were noted at genus level. The most significantly different genera in tumor compared to NAT tissues included *Sphingobium* (p<0.0001), *Anaerococcus* (p<0.0001), *Corynebacterium* (p=0.0012), and *Delftia* (p=0.0031).

Table 19: Significant genera by paired Wilcoxon signed-rank in paired tumor relative to normal adjacent tumor tissues

Genus	Median IQR	P-value	BH95
Sphingobium	-0.347 (-2.134, 0)	0.0001	0.0105
Anaerococcus	-0.159 (-1.475, 0)	0.0001	0.0105
Corynebacterium	-0.693 (-2.359, 0.484)	0.0012	0.0632
Delftia	0 (-1.746, 0)	0.0031	0.1225
Enhydrobacter	0 (-1.314, 0)	0.0065	0.187
Cloacibacterium	-0.784 (-1.708, 0)	0.0071	0.187
Polaromonas	0 (0, 0)	0.0103	0.214
Peptoniphilus	0 (-1.314, 0)	0.0134	0.214
Varibaculum	0 (0, 0)	0.0135	0.214
Finegoldia	0 (-1.386, 0)	0.0142	0.214
Staphylococcus	-0.656 (-2.223, 0.444)	0.0149	0.214
Micrococcus	0 (-0.693, 0)	0.0175	0.2261
Prevotella	-0.347 (-1.365, 0)	0.0186	0.2261
Variovorax	0 (0, 0)	0.0239	0.2697
Coprobacter	0 (0, 0)	0.0316	0.3329
Brevibacterium	0 (0, 0)	0.0343	0.3387
Rhizobacter	0 (0, 0)	0.0411	0.3786
Bergeyella	0 (0, 0)	0.0477	0.3786

Abbreviations: BH: q value of false discovery rate using Benjamin Hochberg 1995 method; IQR: Interquartile Range; P<0.05 is statistically significant

5.3.6. Microbiota abundance and association with prognostic features

5.3.6.1. Abundance of microbiota across staging and IHC group

The relative distribution of microbiota was observed across various features. In clinical stage I and II tumors, at the phylum level, Proteobacteria dominated followed by *Firmicutes* and *Actinobacteria* whereas *Verrucomicrobia* and *Cyanobacteria* were lower in abundance in each category (Figure 39A). Similar to HR status and intrinsic subtype, relative abundance of family *Sphingomonadaceae* and genus *Sphingomonas* dominated in early staged (T1, T2) compared to more advanced stage tumors (T3, T4). In advanced stage, the abundance of *Pseudomonas* was higher followed by *Sphingomonas*.

The relative abundances of microbiota differ across IHC groups. The predominant bacterial phylum, family, and genus across the four IHC breast tumor groups were *Proteobacteria*, *Sphingomonadaceae*, and *Sphingomonas*, respectively, as shown in Figure 39. Microbial composition at the lower taxonomic level (family and genus) were more abundant and/or diverse in HR+ HER- and HR + HER2+ tumors compared to HR- and TNBC (Figure 39B). On the other hand, TNBC tumors having the lowest abundance of *Proteobacteria* compared to the other tumor types. Among HR+/HER+ and TNBC tumors, Actinobacteria was the second leading microbiota unlike the other two groups.

At family level, *Shingomonadacae* and *Pseudomonadacae* identified as dominant taxa among patients with early disease state (Figure 39C). We evaluated that microbiota abundance at a family level indicated that *Sphingomonadacae* abundance was higher in HR-/HER2+ and TNBC tumors Figure 39D). While exploring the relative distribution of microbiota at genus level, *Sphingomonas* was identified as the prevalent microbiota in HR-/HER+ and TNBC tumors (Figure 39E).

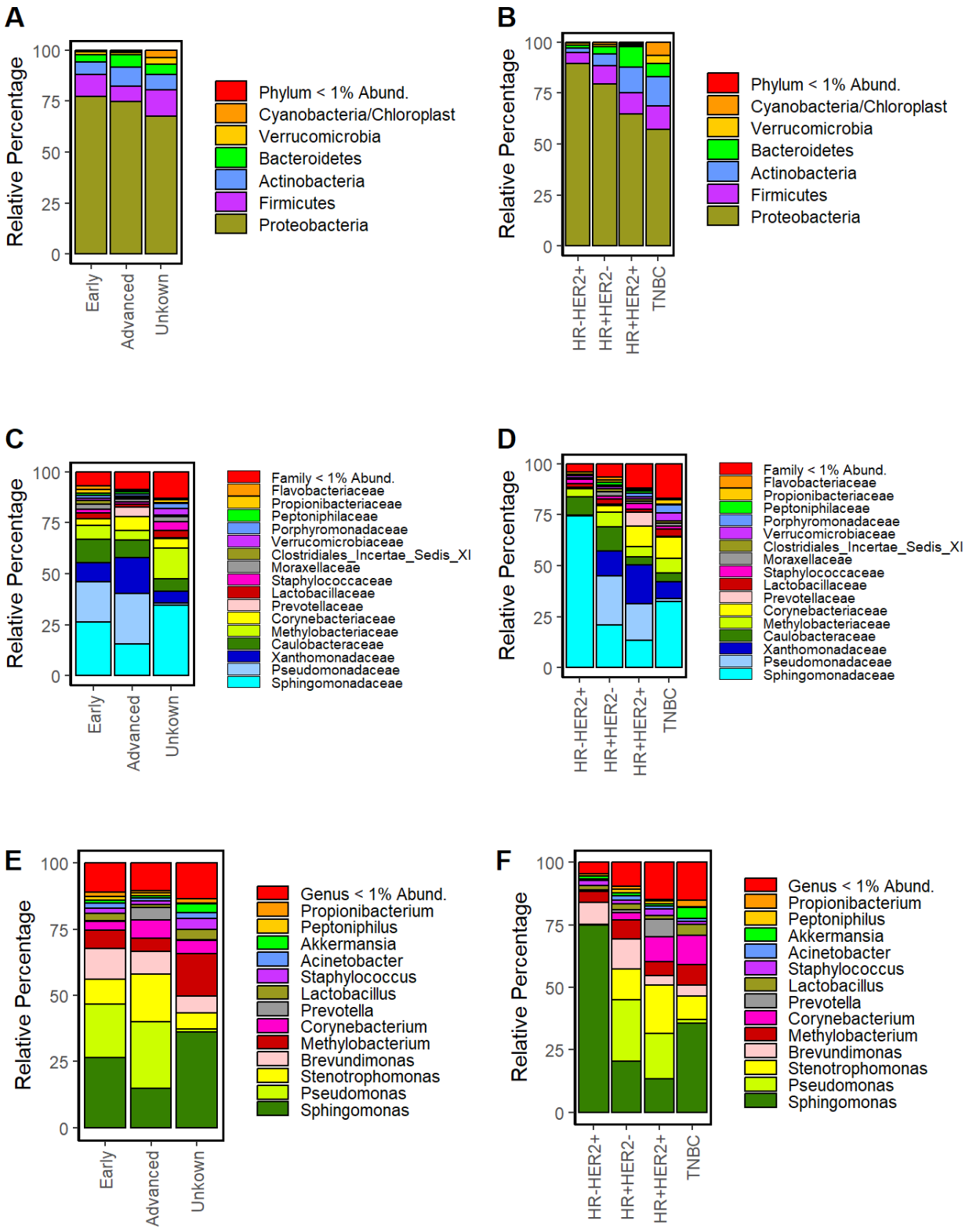


Figure 39: Breast bacterial community composition in relation to histopathological features*: a. Phylum, b. Family, c.Genus distribution stratified by clinical stage and IHC type. All the rare taxa are collapsed in "other" (relative abundance < 1% in all samples)*

5.3.6.2. Microbiota abundance with respect to intrinsic subtypes

Microbiota composition was determined across intrinsic subtypes (Figure 40). Accordingly, the predominant bacterial phylum across the four intrinsic breast tumor subtypes was *Proteobacteria* followed by *Firmicutes* (Figure 40A), whereas at the family level, *Sphingomonadaceae* were predominant in luminal A, HER2-E and basal-like breast tumor intrinsic subtypes (Figure 40B). In the luminal B subtype, the relative abundance of such microbiota was lower, including *Sphingomonadaceae*; however, *Pseudomonadaceae* and *Xanthomonadaceae* were relatively higher. Clearly, the genus *Sphingomonas* was the leading microbiota across intrinsic subtypes except in Luminal B where *Stentotrophomonas* was predominant followed by Pseudomonas (Figure 40C).

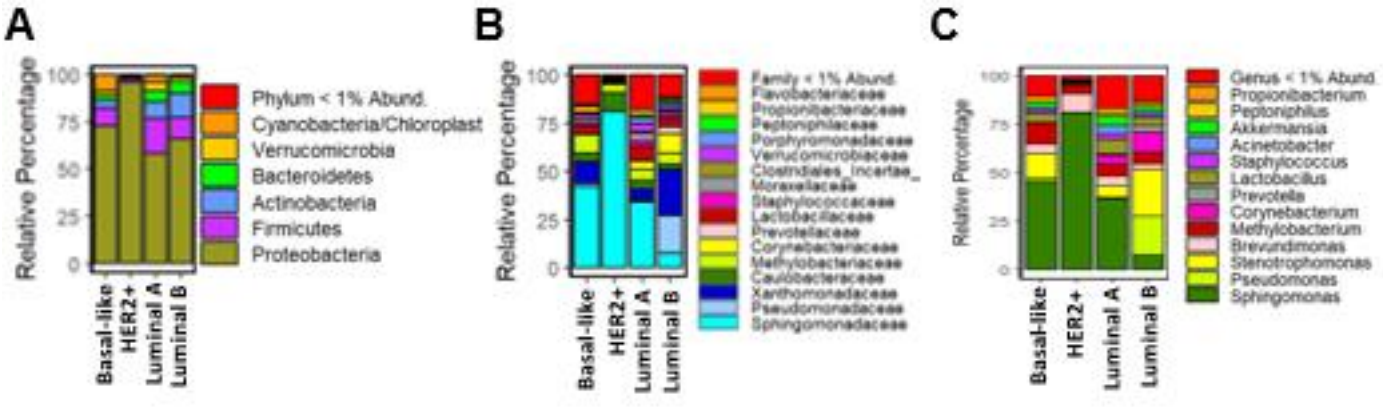


Figure 40: Microbiota abundance across intrinsic subtypes, *relative percentage of breast microbiota at, A) Phylum; B) Family; and C) Genus taxonomic level. All the rare taxa are collapsed in "other" (relative abundance < 1% in all samples)*

5.3.7. Microbiota are significantly associated with prognostic breast tumor features

5.3.7.1. Differential genus abundance in relation to PAM50

In Table 20, the differential genus abundance in tumor in relation to intrinsic subtypes explored. At the lower taxonomic level (genus), our linear model analysis revealed differences in abundance of microbial genera by intrinsic subtype. Most genera abundance had a statistically significant association with in distinct intrinsic subtypes relative to luminal B ($P< 0.05$). As shown in Table 20, at the lower taxonomic level (genus), our linear regression analysis revealed differences in abundance of microbial genera by intrinsic subtype. A statistically significant abundance of *Polaromonas* ($p = 0.0005$), *Varibaculum* ($p = 0.0087$), *Exiguobacterium* ($p = 0.0097$), *Bifidobacterium* ($p = 0.0382$), and several other genera between luminal A, HER2-E, and basal-like tumors were observed overall when compared to luminal B tumors.

Table 20: Significant genera in tumors according to PAM50 (linear model)

Genus	Level	Estimate	SE	P
	luminal A	-0.022	0.055	
Exiguobacterium	HER2-E	0.344	0.1	0.0097
	basal-like	-0.022	0.1	
	luminal A	0.208	0.245	
Bifidobacterium	HER2-E	-0.741	0.448	0.0382
	basal-like	1.007	0.448	
	luminal A	-0.035	0.089	
Varibaculum	HER2-E	0.562	0.162	0.0087
	basal-like	-0.035	0.162	
	luminal A	0	0.044	
Polaromonas	HER2-E	0	0.08	5.00E-04
	basal-like	0.366	0.08	
	luminal A	-0.067	0.083	
Aeromonas	HER2-E	-0.067	0.151	0.0052
	basal-like	0.53	0.151	
	luminal A	0.093	0.119	
Romboutsia	HER2-E	-0.067	0.217	0.0198
	basal-like	0.7	0.217	
	luminal A	0.222	0.145	
Pedobacter	HER2-E	0	0.266	0.0082
	basal-like	0.924	0.266	
	luminal A	-0.022	0.089	
Leifsonia	HER2-E	0.671	0.163	0.0016
	basal-like	-0.022	0.163	
	luminal A	0.031	0.079	
Cupriavidus	HER2-E	-0.022	0.145	0.0095
	basal-like	0.514	0.145	

5.3.7.2. Differential genus abundance in relation to IHC type

A linear model was employed to examine the difference of microbiota abundance at genus level. In table 21, while observing the differential genus abundance in tumor according to IHC type, a significant difference across IHC type was observed. *Exiguobacterium, Rhodopseudomonas, Varibaculum, Polaromonas, Aeromonas, Romboutsia, Pedobacter, Leifsonia,* and *Cupriavidus* had a statistically significant differential genus abundance in relation to IHC type (P-value < 0.05). Significant differences in the relative percentage of microbiota according to IHC group was documented. We did further analysis to understand IHC group difference in the relative abundances of the bacteria at genus level among the different IHC types using a linear model.

We found that *Exiguobacterium* (p=0.007), *Rhodopseudomonas* (p=0.0376), *Varibaculum* (p=0.0087), *Leifsonia* (*p* = 0.0016) and several other genera were significantly less abundant in HR+HER2-, HR+HER2+ and TNBC tissues compared to HR-HER2+ tumors.

Table 21: Significant genera in tumors according to IHC status (linear model)

Genus	**Level**	**Estimate**	**SE**	**P**
Exiguobacterium	HR+HER2-	-0.366	0.1	0.007
	HR+HER2+	-0.317	0.105	
	TNBC	-0.366	0.135	
Rhodopseudomonas	HR+HER2-	-0.647	0.588	0.0376
	HR+HER2+	0.282	0.618	
	TNBC	-0.157	0.793	
Varibaculum	HR+HER2-	-0.561	0.162	0.0086
	HR+HER2+	-0.597	0.171	
	TNBC	-0.597	0.219	
Polaromonas	HR+HER2-	0	0.08	5.00E-04
	HR+HER2+	0	0.084	
	TNBC	0.366	0.108	
Aeromonas	HR+HER2-	0.023	0.151	0.0047
	HR+HER2+	0.099	0.159	
	TNBC	0.597	0.204	
Romboutsia	HR+HER2-	0.116	0.218	0.0226
	HR+HER2+	0.05	0.229	
	TNBC	0.768	0.294	
Pedobacter	HR+HER2-	0.037	0.272	0.0198
	HR+HER2+	0.128	0.286	
	TNBC	0.924	0.367	
Leifsonia	HR+HER2-	-0.67	0.163	0.0016
	HR+HER2+	-0.693	0.172	
	TNBC	-0.693	0.22	
Cupriavidus	HR+HER2-	0.046	0.145	0.0086
	HR+HER2+	0	0.152	
	TNBC	0.536	0.195	

NOTE: HR-HER2- tumors serve as the reference group: Abbreviations: SE: Standard Error; P<0.05 is statistically significant

5.3.7.3. Differential genus abundance in relation to clinical/ simple stage

Table 22 depicts the differential genus abundance in tumor according to clinical/ simple stage stage. Specifically, genus *Stenotrophomonas* (p=0.015), *Corynebacterium* (p=0.049), *Prevotella* (p=0.024), *Actinomyces* (p=0.021) and several additional genera were significantly more abundant in advanced staged tumors compared to early-stage tumors.

Table 22: Significant genera that differ between advanced vs early-stage tumors

Genus	Estimate*	SE	P-value
Stenotrophomonas	1.51	0.597	0.0155
Corynebacterium	1.122	0.554	0.0494
Prevotella	1.11	0.475	0.0244
Actinomyces	0.814	0.338	0.0207
Streptococcus	0.668	0.317	0.0416
Anaerococcus	0.618	0.303	0.0484
Citrobacter	0.571	0.196	0.0059
Dermacoccus	0.252	0.103	0.0192
Blastococcus	0.194	0.086	0.0294

*Estimate: Early staged tumors serve as the reference group; Abbreviations: SE: Standard Error; P<0.05 is statistically significant

5.3.8. Specific breast microbial communities correlate with clinicopathological features in breast cancer

Lastly, we performed a multivariate regression analysis to further identify microbial genera with significantly different abundances when tumor was stratified by IHC group or HR/HER2 status, intrinsic subtype, and stage. Specific taxa were significantly associated with IHC type, intrinsic subtype and stage of disease (Figure 41). As shown in Fig 35A, HR-HER2+, HR+ HER 2 +ve or HER2 -ve tumors and TNBC had distinct statistically significant predominant bacterial genera. In HR-HER2+ tumors, *Bradyrhizobium* and *Exiguobacterium*, in HR+ HER 2+/- tumors *Thermicanus*, *Romboutsia*, *Pedobacter*, *Burkholderia* and *Aeromonas*, and in TNBC tumors, *Rhodopseudomonas* and *Polaromonas* were among the predominant genera of bacteria communities. Interestingly, TNBC tumors correlated positively with several taxa including genus *Burkholderia*, *Aeromonas*, *Mogibacterium*, *Paracoccus*, *Thermicanus*, *Pedobacter*, and *Romboutsia* (Figure 41A).

In particular, HR-HER2+ tumors strongly correlated with genus *Alkanindiges*, *Anoxybacillus*, *Rhodopseudomonas*, and *Leifsonia* whereas genus *Anoxybacillus* slightly correlated positively with HR+HER2-. Similarly, HR+HER+ tumors correlated positively with only genus *Rhodopseudomonas* (Figure 41A).

Similar to the IHC types, the intrinsic subtypes, Luminal A and Luminal B tumors had similarity in type as well as abundance of predominant genera, and this includes taxa, including *Thermicanus*, *Romboutsia*, *Paracoccus*, *Cupriavidus*, *Burkholderia*, *Anoxybacillus*, and *Alkanindiges*. HER2+ tumors strongly correlated with *Anoxybacillus*, *Alkanindiges*, *Leifsonia*, and *Exiguobacterium*. Basal-like tumors strongly correlated with same genera observed in TNBC tumors except there was a weak association with *Cuprivadus* (Figure 41B) that was not observed in TNBC tumors (Figure 41A).

Lastly, we observed a stronger correlation with genera *Jeogalicoccus* and *Citrobacter* in advance staged tumors, whereas early staged tumors more strongly correlated with *Jeogalicoccus* tumors (Figure 41C).

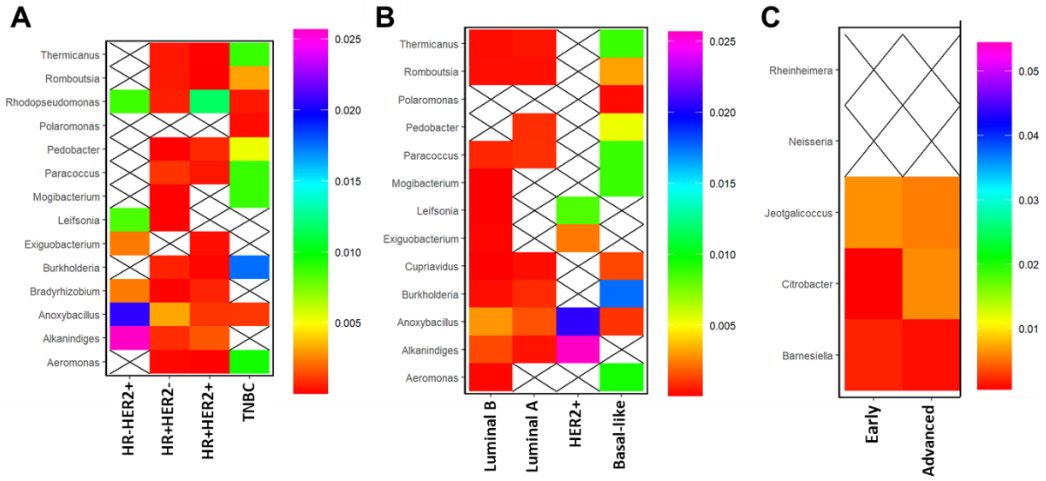


Figure 41: Specific bacterial genera correlate with clinicopathologic features**.** *Mean relative abundances (proportions) of bacterial genera that were differentially present in breast tumors stratified by (A) IHC type (b) PAM50 breast tumor subtype, and (C) early or advanced stage. Crossed-out boxes indicate samples for which specific genera were not detected. Color bars vary on a logarithmic scale. All genera shown had FDR-corrected p-value < 0.05 by Kruskal–Wallis H-test after adjustments for age*

5.3.9. Microbiota Diversity: Alpha and Beta diversity

5.3.9.1. Alpha and Beta diversity between NAT and breast tissue

The alpha diversity was measured using both Shannon and Simpson and plotted as a box-plot. To demonstrate the relative richness and show how evenly OTUs are distributed in a sample, are the describing evenness (how evenly distributed these OTUS are) in a given sample, alpha diversity (Shannon index) was computed and illustrated for each sample. In general, (alpha) diversity (measured with the Shannon Index, which weighs the numbers of species and their abundances in each sample) is very similar among groups.

In the following Violin plot (Figure 42), the relative alpha diversity level between tumor relative to normal tissue was illustrated. ANOVA was used to measure the statistical significance of the distribution of OUT across sample.

To compare differences in breast microbial diversity between the paired NAT and tumor tissues, we measured α-diversity, defined as within-group taxonomic richness and evenness. Interestingly, the Shannon index was not significantly different between NAT (mean: 3.05 ± 0.49) and tumor tissues (mean: 3.11 ± 0.50) (p= 0.07); however, alpha diversity measured by the Simpson index revealed slightly higher alpha diversity in tumors compared to NAT (p = 0.043). Additionally, Simpson index did not reveal significant differences in α-diversity between NAT (0.90 ± 0.50) and tumor (0.95 ± 0.50) tissues.

On the other hand, to obtain a graphical representation of microbiome composition similarity among samples, we summarized OTU abundances into Bray-Curtis dissimilarities and performed a PCoA ordination of unweighted UniFrac distances. With this approach, the microbiome composition for all specimens were analyzed. To determine differences in beta diversity, we visualized the overall differences between the breast microbiome profile of tumor and NAT tissues using principal coordinate analysis (PCoA) of Bray–Curtis dissimilarity. Permutational analysis of variance (*adonis* R function, or Permanova) determined no significant differences in beta-diversity. The tumor tissue and NAT clustered significantly differently between the two groups (R^2 = 0.664; p = 0.049) with greater dissimilarity along PC1 (23.1% variation); however, we also did observe overlap between the two groups.

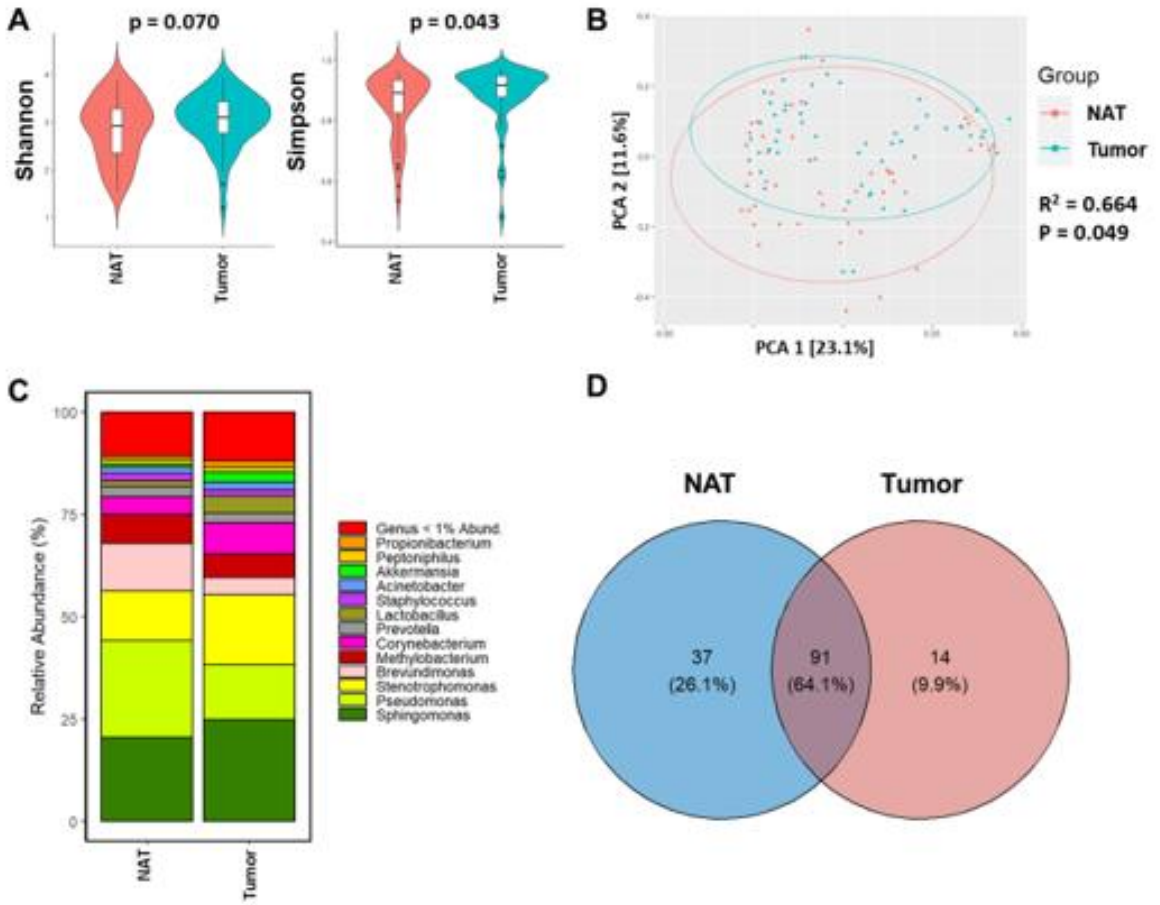


Figure 42: Breast bacterial community composition varies by patient BC status and normal adjacent tumor (NAT) and tumor tissue types. *(**A**) Violin plots show median and interquartile range in bacterial alpha diversity as measured by Shannon and Simpson diversity indices within (tumor and NAT) breast tissue from Ethiopian* BC *patients. p-values were obtained from Kruskal–Wallis tests. (**B**) Principal coordinates (PC) plots show the clustering pattern of tumor and NAT based on unweighted UniFrac distance and are colored by sample types (red circles—NAT, teal—Tumor samples); p = 0.049 and R^2 = 0.664. (**C**) Taxonomic composition of the breast microbiome, depicted as proportional abundances at the genus level for NAT and corresponding tumor tissue of the same donor. (**D**) Venn diagram showing the unique and shared bacterial genera between the tumor and normal groups*

5.3.9.2. Alpha and Beta diversity with respect to tumor stage, intrinsic subtypes and IHC

By tumor sage

We next investigated the relationship between tumor stage and the relative abundances among microbial taxa. Shannon (p = 0.475) and Simpson (p = 0.358) indexes were not significantly different between early and advanced stage of disease, respectively. PCA showed no significant separation between early and advanced stages of BC (R^2 = 0.533, p = 0.16). However, visually there is a clear separation between the two groups. There were eight patients with unknown stage information which did not significantly change the results (Figure 43A, B).

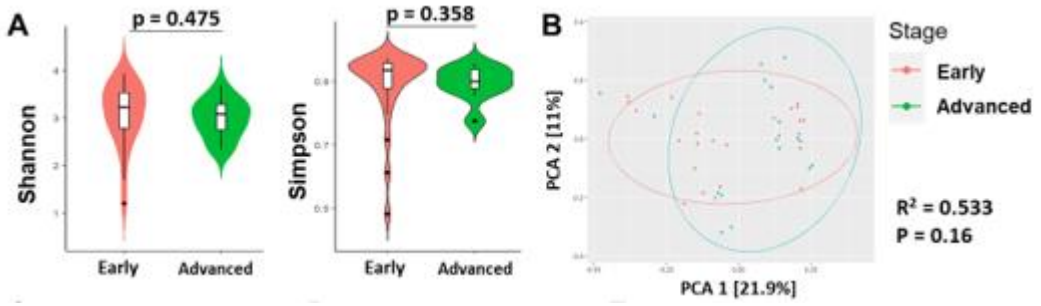


Figure 43: Breast bacterial community composition varies between early and advanced stages of BC. *(**A**) Violin plots show median and interquartile range as measured by Shannon and Simpson diversity indices based upon breast tumor stage. p-value results were obtained from Kruskal–Wallis tests. (**B**) PC plots show the clustering pattern among early (red circles), advanced (green circles), and unknown (blue circles) tumor stage based on unweighted UniFrac distance; p = 0.16 and R^2 = 0.533. (**B**) Taxonomic composition of the breast microbiome, depicted as relative percentage at the (**B**) phylum, (**C**) family, and (**D**) genus level*

By intrinsic subtypes

Since we observed significant differences in the relative percentage of microbiota according to IHC group/HR/HER2 status, we hypothesized that microbial differences could exist across intrinsic subtypes. Similar to IHC type, Shannon and Simpson indexes were not significantly different between the four intrinsic subtypes (*p*-value = 0.329 vs. *p*-value = 0.310, respectively) (Figure 44A), which was also observed among the PCA analysis (R^2 = 0.779, *p* = 0.378) (Figure 44B).

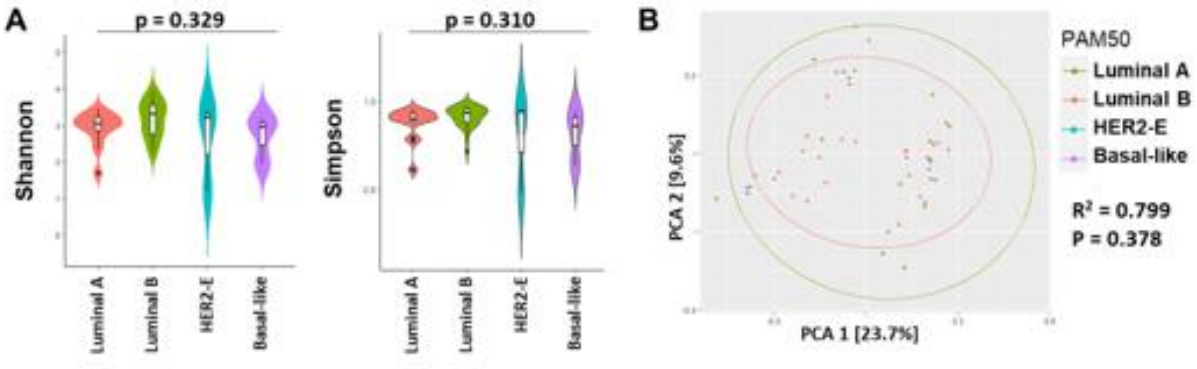


Figure 44: Breast bacterial community composition varies by PAM50 intrinsic subtypes. *(**A**) Violin plots show median and interquartile range as measured by Shannon and Simpson diversity indices within breast tissue from Ethiopian BC patients. p-value results were obtained from Kruskal–Wallis tests. (**B**) PC plot shows the clustering pattern of the intrinsic subtypes based on unweighted UniFrac distance and is colored by sample types (green—luminal A, red circles—luminal B, teal—HER2E, and purple circles—basal-like tumor samples); p = 0.378 and R^2 = 0.799*

By IHC group

We next determined whether microbial composition and abundance were associated with breast tumor prognostic features, including by IHC group. The Shannon index (p = 0.717) and Simpson index (p = 0.748) did not differ significantly by IHC group/ HR/HER2 status, respectively (Figure 45).

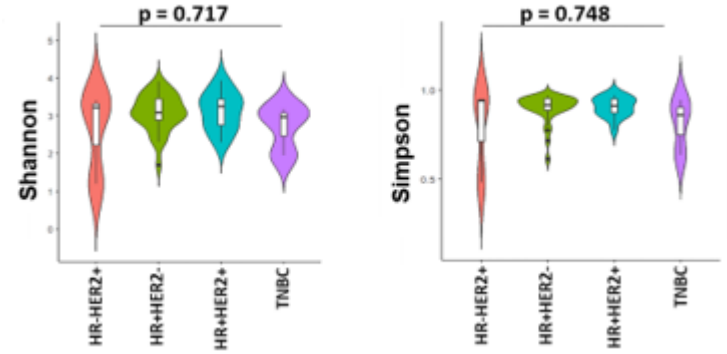


Figure 45: Breast bacterial community composition varies by immunohistochemical (IHC) status. *(**A**) Violin plots show median and interquartile range as measured by Shannon and Simpson diversity indices with regard to hormone receptor (HR) and HER2 status of the tumor tissue. p-value results were from Kruskal–Wallis tests*

5.4. mRNA and Protein Level Biomarker Expression Concordance for *ESR1*/ER, *PGR*/PgR and *ERBB2*/HER2 in Ethiopian Patients with BC: Implications for Diagnostic Alternatives in Ethiopia

5.4.1. Concordance between End point PCR Vs IHC using FFPE

The concordance between mRNA and protein level expression of established BC molecular markers was investigated in independent FFPE tissue (n=160 for Endpoint PCR Vs IHC), FFT (n=91 for Endpoint PCR versus IHC, 56 for qPCR Vs IHC) and paired (FFPE and FFTT from same women) tissues (n=50 for endpoint PCR Vs IHC) from same women.

The endpoint PCR analysis was performed on three GOI as follows: ESR1, PGR and ERBB2 applying specific primer set. During the analysis, the scoring was classified in to three "0" (negative), "1" (weak positive) and "3" (strongly positive) based on amplicon band intensity. In this study using independent FFPE tissues, approximately 40%, 8.2% and 5% of the tumors were either weakly and/or strongly positive for ER, PR and HER2, respectively.

Majority (73.12%) and 38.75% of the tumors identified positive for ER using IHC and Endpoint PCR, respectively. In both methods, 37.5% and 25.62% of the samples turned out positive and negative for ER, respectively. A difference in the classification of biomarker status noted applying mRNA and protein level expression method. Of the total (n=117) samples identified as ER positive with IHC, almost half of the samples negative by Endpoint PCR.

When examining the alignment in detecting negative cases, of the IHC confirmed ER negative samples (n=43), almost all (n=41) were detected as negative by Endpoint PCR technique. The analysis revealed that the sensitivity and specificity of the PCR accounts 51.28% (95% CI: 41.87% to 60.63%) and 95.35% (95% CI: 84.19% to 99.43%), respectively. Considering this, the PPV and NPP value found to be 96.77% (95% CI: 88.46% to 99.16%) and 41.84% (95% CI: 37.13% to 46.70%), respectively with 63.12% (95% CI: 55.15% to 70.61%) diagnostic accuracy. Finally, the finding revealed that there was a statistically significant fair agreement between the reference and Endpoint PCR method (Kappa value =0.332; P-value = 0.001) in detecting ER.

Utilizing IHC, 55% of the tumors were PgR positive; however, the end point PCR method detected 8.125% of the samples positive for this marker. Both methods capable of detecting a total of 8.125% and 45% tumors positive and negative for PgR, respectively.

Out of the 88 samples identified PgR positive with the reference method, 85.277% of the samples were negative by Endpoint PCR. On the other hand, out of the samples identified as negative by IHC (n=72), all were turned negative by End point PCR.

Furthermore, the finding demonstrated that the sensitivity and specificity of the end point PCR method was 14.77% (95% CI: 8.11% to 23.94%) and 100% (95% CI: 95.01% to 100%), respectively. Again, the PPV and NPP value was 100% (95% CI: 75.29% to 100%) and 48.98% (95% CI: 46.81% to 51.15%), respectively with diagnostic accuracy of 53.12% (95% CI: 45.08% to 61.05%). Speaking of concordance, a statistically significant slight agreement noted between IHC and Endpoint PCR method considering this specific marker (Kappa value =0.135; P-value = 0.001).

With the IHC and Endpoint PCR method, 23.75% and 5.625% of the tumors were HER2 positive and negative, respectively. Both methods identified that 3.125% and 20.625% of the tumors were HER2 positive and negative, respectively. Of the 33 samples identified HER2 positive by IHC, five (15.15%) samples were positive by Endpoint PCR method.

Out of the samples identified as HER2 negative by IHC (76.25%, n=122)), majority (96.72%, n=118) were negative by Endpoint PCR. In this study, the analysis demonstrated that the sensitivity and specificity accounts 13.16% (95% CI: 4.41% to 28.09%), 96.72% (95% CI: 91.82% to 99.10%), respectively. The PPV and NPP value was documented 55.56% (95% CI: 26.11% to 81.56%) and 78.15% (95% CI: 75.88% to 80.25%), respectively with a diagnostic accuracy of 76.88% (95% CI: 69.56% to 83.17%). The expression of the marker confirmed having a statistically significant slight agreement between the reference method and Endpoint PCR method (Kappa value =0.134; P-value = 0.021).

Table 23: Performance evaluation between Endpoint PCR and IHC FFPE samples from patients with BC

		IHC status [ER-FFPE]	
		negative	positive
Endpoint PCR [*ESR1*]	Negative	41	57
	positive	2	60
	Sensitivity, n (95% CI)	51.28% (41.87% to 60.63%)	
	Specificity, n (95% CI)	95.35% (84.19% to 99.43%)	
	PPV, n (95% CI)	96.77% (88.46% to 99.16%)	
	NPV, n (95% CI)	41.84% (37.13% to 46.7%)	
	Accuracy, n (95% CI)	63.12% (55.15% to 70.61%)	
	Kappa value	0.332	
	P-value	0.001	
		IHC status [PgR-FFPE]	
		negative	positive
Endpoint PCR [*PGR*]	negative	72	75
	positive	0	13
	Sensitivity, n (95% CI)	14.77% (8.11% to 23.94%)	
	Specificity, n (95% CI)	100% (95.01% to 100%)	
	PPV, n (95% CI)	100% (75.29% to 100%)	
	NPV, n (95% CI)	48.98% (46.81% to 51.15%)	
	Accuracy, n (95% CI)	53.12% (45.08% to 61.05%)	
	kappa value	0.135	
	P-value	0.001	
		IHC status [HER2-FFPE]	
		negative	positive
Endpoint PCR [*ERBB2*]	negative	118	33
	positive	4	5
	Sensitivity, n (95% CI)	13.16% (4.41% to 28.09%)	
	Specificity, n (95% CI)	96.72% (91.82% to 99.10%)	
	PPV, n (95% CI)	55.56% (26.11% to 81.56%)	
	NPV, n (95% CI)	78.15% (75.88% to 80.25%)	
	Accuracy, n (95% CI)	76.88% (69.56% to 83.17%)	
	Kappa value	0.134	
	P-value	0.021	

5.4.2. Concordance between End point PCR Vs IHC using fresh frozen tissue

The GOI and the scoring of the band intensity considered to this specific samples were similar to the FFPE samples. Accordingly, a total of 79.2%, 38.5% and 30.8% of the samples were positive for ER, PgR and HER2, respectively using End point PCR method.

The detail of the diagnostic performance of the reference and end point PCR method is shown in table 24. Using these samples, IHC and End point PCR method identified a total of 89.01% and 79.12% tumors positive for ER, respectively.

Both methods detected higher proportion of ER positive tumors (72.52%); however, only small proportion of overlap documented in detecting ER negative cases (4.39%). Of the total, 81 samples identified ER positive with the reference method, 66(75%) samples) were found positive by Endpoint PCR method. Of the total 11 samples tended negative by IHC method, 4 (40%) were negative by Endpoint PCR technique. Additionally, the analysis revealed that the sensitivity and specificity the investigational device accounts 81.48% (95% CI: 71.3% to 89.25%), 40% (95% CI: 12.16% to 73.76%), respectively. Specific to ER, the PPV and NPP value was 91.67% (95% CI: 86.78% to 94.86%) and 21.05% (95% CI: 9.91% to 39.27%), respectively with diagnostic accuracy of 76.92% (95% CI: 66.92% to 85.11%). Evaluating the level of agreement, this study showed that there was no statistically significant agreement between the reference and the Endpoint PCR method (Kappa value =0.154; P-value =0.115).

Using the independent FFT samples, 63.73% and 38.46% of the tumors were identified PgR positive using IHC and End point PCR method, respectively. Both methods detected a total of 28.57% and 26.37% tumors positive and negative for PgR, respectively. Out of the 58 samples positive for PgR by reference method, close to half (44.82%, n=26)) samples were positive by Endpoint PCR method; however, amongst 33 samples negative for PgR with IHC, majority (72.72%, n=24) were negative by Endpoint PCR technique.

Furthermore, the sensitivity and specificity accounts for 44.83% (95% CI: 31.74% to 58.46%), 72.73% (95% CI: 54.48% to 86.70%), respectively. While looking other parameters, the PPV and NPP value was 74.29% (95% CI: 60.70% to 84.38%) and 42.86% (95% CI: 35.44% to 50.61%), respectively with a diagnostic accuracy of 54.95% (95% CI:44.16 % to 65.40 %). But there was no statistically significant concordance between IHC and Endpoint PCR method (Kappa value =0.153; P-value =0.098).

Employing protein and mRNA-based methods, IHC and the PCR identified a total of 29.67% and 38.46% samples positive for HER2, respectively. A total of 12.08% and 43.95% of the tumors were detected HER2 positive and negative, respectively by both methods. Of the total samples (27/91) tended positive by IHC, considerable samples (40.74%, 11) were positive with Endpoint PCR method; however, of those samples identified as HER2 negative by IHC (64/91), more than two-third (62.5%, n=40) were negative by Endpoint PCR technique.

The analysis revealed that the sensitivity and specificity accounts 40.74% (95% CI: 22.39% to 61.20%), and 62.50% (95% CI: 49.51% to 74.30%), respectively. The finding presented that the PPV and NPP value was 31.43% (95% CI: 20.85% to 44.37%) and 71.43% (95% CI: 63.42% to 78.28%), respectively with 56.04% (95% CI: 45.25 % to 66.44%) diagnostic accuracy. The finding indicated that there was no a statistically significant concordance between IHC and Endpoint PCR method (Kappa value =0.030; P-value =0.772).

Table 24: Performance evaluation between Endpoint PCR and IHC in FFT samples from patients with BC

		IHC status [ER-FFT]	
		negative	positive
Endpoint PCR [*ESR1*]	negative	4	15
	positive	6	66
	Sensitivity, n (95% CI)	81.48% (71.30% to 89.25%)	
	Specificity, n (95% CI)	40% (12.16% to 73.76%)	
	PPV, n (95% CI)	91.67% (86.78% to 94.86%)	
	NPV, n (95% CI)	21.05% (9.91% to 39.27%)	
	Accuracy, n (95% CI)	76.92% (66.91% to 85.11%)	
	Kappa value	0.154	
	P-value	0.115	
		IHC status [PgR-FFT]	
		negative	positive
Endpoint PCR [*PGR*]	negative	24	32
	positive	9	26
	Sensitivity, n (95% CI)	44.83% (31.74% to 58.46%)	
	Specificity, n (95% CI)	72.73% (54.48% to 86.7%)	
	PPV, n (95% CI)	74.29% (60.7% to 84.38%)	
	NPV, n (95% CI)	42.86% (35.44% to 50.61%)	
	Accuracy, n (95% CI)	54.95% (44.16% to 65.4%)	
	kappa value	0.153	
	P-value	0.098	
		IHC status [HER2-FFT]	
		negative	positive
Endpoint PCR [*ERBB2*]	negative	40	16
	positive	24	11
	Sensitivity, n (95% CI)	40.74% (22.39% to 61.2%)	
	Specificity, n (95% CI)	62.5% (49.51% to 74.3%)	
	PPV, n (95% CI)	31.43% (20.85% to 44.37%)	
	NPV, n (95% CI)	71.43% (63.42% to 78.28%)	
	Accuracy, n (95% CI)	56.04% (45.25% to 66.44%)	
	Kappa value	0.03	
	P-value	0.772	

5.4.3. Concordance between End point PCR Vs IHC using paired (FFT Vs FFPE) tissues

5.4.3.1. Concordance between endpoint PCR Vs IHC using paired tissue of same women

Analysis was done on a total of 50 paired breast tissue samples to observe the trend in ER status in same women. Accordingly, 29(58%) of FFPE and 40(80%) of FFT samples were positive for ER using end point PCR from same patient. The finding confirmed that ER positivity increased by 22% when FFT samples utilized from same women. On the other hand, 11(22%) of FFPE and 25 (50%) of FFT samples were positive for PgR using Endpoint PCR from same patient with a 28% increment while using FFT samples to determine PgR status from same women. On the other hand, 3(6%) of FFPE and 15(34%) of FFT samples were positive for HER2 using end point PCR from same patient. The positivity of HER2 status increased by 28% while using FFT instead of FFPE samples to determine the protein among patients with BC.

Finally, we did analysis on 50 paired breast tissue where ER, PgR and HER positivity increased by 22%, 28% and 28%, respectively when using FFT samples. Additionally, the sensitivity increased for: ER from 65.85 [FFPE] to 82.61% [FFT] samples; PgR: from 31.43% [FFPE] to 57.58% [FFT]; HER2: from 37.50% [FFPE] to 43.75% [FFT]. On the other hand, the percentage of specificity declined in FFT samples unlike FFPE samples across all markers. In the paired tissue, a statistically significant agreement between the method was identified while using FFPE instead of FFT samples for ER [k=0.287: fair agreement; P-value=0.016], PgR [k=0.216: fair agreement; P-value=0.014] and HER2 [k=0.502: moderate agreement; P-value=0.001].

Table 25: Performance evaluation between Endpoint PCR and IHC using paired breast tissue among patients with BC

		IHC status [ER-FFPE]		**IHC status [ER-FFT]**	
		negative	positive	negative	positive
Endpoint PCR [*ESR1*]	negative	7	14	2	8
	positive	2	27	2	38
	Sensitivity, n (95% CI)	65.85% (49.41% to 79.92%)		82.61% (68.58% to 92.18%)	
	Specificity, n (95% CI)	77.78% (39.99% to 97.19%)		50% (6.76% to 93.24%)	
	PPV, n (95% CI)	93.10% (79.59% to 97.91%)		95% (87.60% to 98.08%)	
	NPV, n (95% CI)	33.33% (22.39% to46.43%)		20% (7.23% to 44.49%	
	Accuracy, n (95% CI)	68% (53.30% to 80.48%)		80% (66.28% to 89.97%)	
	Kappa value	0.287		0.194	
	P-value	0.016		0.118	
		IHC status [PgR-FFPE]		**IHC status [PgR-FFT]**	
		negative	positive	negative	positive
Endpoint PCR [*PGR*]	negative	15	24	11	14
	positive	0	11	6	19
	Sensitivity, n (95% CI)	31.43% (16.85% to 49.29%)		57.58% (39.22% to 74.52%)	
	Specificity, n (95% CI)	100% (78.20 to 100%)		64.71% (38.33% to 85.79%)	
	PPV, n (95% CI)	100% (71.51% to 100%)		76% (60.96% to 86.53%)	
	NPV, n (95% CI)	38.46% (33.31% to 43.89%)		44% (31.62% to 57.18%)	
	Accuracy, n (95% CI)	52% (37.42% to 66.34%)		60% (45.18% to 73.59%)	
	kappa value	0.216		0.2	
	P-value	0.014		0.136	
		IHC status [HER2-FFPE]		**IHC status [HER2-FFT]**	
		negative	positive	negative	positive
Endpoint PCR [*ERBB2*]	negative	42	5	24	9
	positive	0	3	10	7
	Sensitivity, n (95% CI)	37.50% (8.52% to 75.51%)		43.75% (19.75% to 70.12%)	
	Specificity, n (95% CI)	100% (91.59% to 100%)		70.59% (52.52% to 84.9%)	
	PPV, n (95% CI)	100% (29.24% to 100%)		41.18% (24.64% to 59.98%)	
	NPV, n (95% CI)	89.36% (83.08% to 93.49%)		72.73% (62.18% to 81.22%)	
	Accuracy, n (95% CI)	90% (78.19% to 96.67%)		62% (47.17% to 75.35%)	
	Kappa value	0.502		0.141	
	P-value	0.001		0.318	

5.4.4. Concordance between qPCR analysis Vs IHC using fresh frozen tissue

The qPCR analysis was performed for three GOI as follows: ESR1, PGR and ERBB2 applying specific primer set. Accordingly, a total of 12(21.4%) and 44 (78.6%) samples were negative and positive for ESR1, respectively. A total of 25 (44.6%) and 30(53.6%) samples were negative and positive for PGR. A total of 46 (82.1%) and 10(17.9%) samples were negative and positive for ERBB2, respectively. The detail of the diagnostic performance of the qPCR method is shown in table 26.

Out of the total sample, IHC and qPCR method identified 83.2% and 81.4% of the samples positive for ER, respectively. In this study, variation in the assignment of ER status observed applying the protein expression and gene expression platform. Out of the samples identified ER positive with the reference method, majority (94.44%, n=34) of the samples were positive by qPCR method. On the other hand, of those samples identified as ER negative by IHC, 6(85.71%) were negative by qPCR technique.

In this study, the analysis revealed that the sensitivity and specificity accounts 94.44% (95% CI: 81.34% to 99.32 %), 85.71% (95% CI: 42.13% to 99.64%), respectively. The finding presented that the PPV and NPP value was 97.14% (95% CI: 84.68% to 99.52(%) and 75% (95% CI: 43% to 92.27%), respectively with diagnostic accuracy of 93.02% (95% CI: 80.94% to 98.54%). Accordingly, the finding confirmed that there was a statistically significant substantial agreement between the reference method and qPCR method (Kappa value =0.758; P-value =0.001).

Amongst the total, a total 57.14% and 57.14% of the tumors identified PgR positive with IHC and qPCR, respectively. By both methods, 52.38% and 66.66% of the tumors were positive and negative respectively. In this study, variation in the assignment of PgR status observed applying the protein and mRNA expression method. Out of out of the samples identified PgR positive with the reference method, most of the tumors (91.66%, n=22) were positive by qPCR. However, out of the samples identified as PgR negative by IHC (18/42), majority (88.88, n=16) were negative by qPCR. Moreover, the analysis revealed that the sensitivity and specificity accounts 91.67% (95% CI: 73% to 98.97%), 88.89% (95% CI: 65.29% to 98.62%), respectively. The finding demonstrated that the PPV and NPP value was 91.67% (95% CI: 74.76% to 97.61%) and 88.89% (95% CI: 67.75% to 96.82%), respectively with a diagnostic accuracy of 90.48% (95% CI: 77.38% to 97.34%). The finding indicated that there was a statistically significant perfect concordance between IHC and qPCR method (Kappa value =0.806.; P-value =0.001).

Amongst the total, IHC technique identified a total 15(34.88%) HER2 positive tumors whereas 9(20.93%) of the sample picked positive with qPCR. Accordingly, out of the total samples (15/43/) identified HER2 positive with the reference method, 7(46.66%) samples were found positive by qPCR method. On the other hand, of those samples identified as HER2 negative by IHC (28/43), 26(92.86%) were negative by qPCR technique.

Additionally, the analysis revealed that the sensitivity and specificity of accounts 46.67% (95% CI: 21.27% to 73.41%), and 92.86% (95% CI: 76.5% to 99.12%), respectively. The finding presented that the PPV and NPP value was 77.78% (95% CI: 45.31% to 93.67%) and 76.47% (95% CI: 66.69% to 84.06%), respectively with a diagnostic accuracy of 76.74% (95% CI: 61.37% to 88.24%). The finding indicated that there was a statistically significant moderate agreement between IHC and qPCR method to determine HER2 (Kappa value =0.436; P-value =0.002).

Table 26: Performance evaluation between qPCR and IHC in FFT samples from patients with BC

		IHC status [ER-FFT]	
		negative	positive
qPCR [*ESR1*]	negative	6	2
	positive	1	34
	Sensitivity, n (95% CI)	94.44% (81.34% to 99.32%)	
	Specificity, n (95% CI)	85.71% (42.13% to 99.64%)	
	PPV, n (95% CI)	97.14% (84.68% to 99.52%)	
	NPV, n (95% CI)	75% (43% to 92.27)	
	Accuracy, n (95% CI)	93.02% (80.94% to 98.54%)	
	Kappa value	0.758	
	P-value	0.001	
		IHC status [PgR-FFT]	
		negative	positive
qPCR [*PGR*]	negative	16	2
	positive	2	22
	Sensitivity, n (95% CI)	91.6% (73% to 98.97%)	
	Specificity, n (95% CI)	88.89% (65.29% to 98.62%)	
	PPV, n (95% CI)	91.67% (74.76% to 97.61%)	
	NPV, n (95% CI)	88.89% (67.75% to96.82%)	
	Accuracy, n (95% CI)	90.48% (77.38% to 97.34%	
	kappa value	0.806	
	P-value	0.001	
		IHC status [HER2-FFT]	
		negative	positive
qPCR [*ERBB2*]	negative	26	8
	positive	2	7
	Sensitivity, n (95% CI)	46.67% (21.27% to 73.41%)	
	Specificity, n (95% CI)	92.86% (76.50% to 99.12%)	
	PPV, n (95% CI)	77.78% (45.31% to 93.67%)	
	NPV, n (95% CI)	76.47% (66.69% to 84.06%)	
	Accuracy, n (95% CI)	76.74% (61.37% to 88.24%)	
	Kappa value	0.436	
	P-value	0.002	

6. DISCUSSION

Breast cancer is common cancer in Ethiopia. There are many studies done on the epidemiology, screening and treatment outcome. However, few data are available on the biology of BC in Ethiopia. This study is the first study that addressed the biology of BC in Ethiopia using state-of-the art technology. The project includes studying the multi-gene expression pattern among 334 BC patients, *PIK3CA* mutation analysis, comparing microbiota profiling of normal adjust and tumor tissue of BC and PCR based (conventional and qPCR) based analysis of important biomarkers used in the management of BC. The multigene expression analysis using the Nanostring technology allowed us to learn the molecular epidemiology of BC in Ethiopia.

Patients and histopathological features

In this study, two-third patients recruited were younger in agreement with previous study (Bennis et al., 2012; Kuzhan et al., 2013; Sayed et al., 2014; Traoré et al., 2019). The median age was slightly lower from African countries including Eastern Morocco (Elidrissi Errahhali et al., 2017) and Kenya (Sayed et al., 2014). But, it was in contrast with developed countries in Western, Arab and Asian countries where BC is prevalent in older age groups (Aiad et al., 2014; Blamey et al., 2010; Kallel et al., 2012; Khabaz, 2014; Parise & Caggiano, 2014; Runnak et al., 2012). Moreover, relatively younger age of women with BC is already reported in several studies in Africa (Effi et al., 2017; Fregene & Newman, 2005). The difference in age during cancer diagnosis among patients with BC could be explained by the heterogeneity in the structure of the population where we have more young population compared to the west. Considering the aggressiveness of tumor at younger age group, possible strategy has to be devised for screening, early detection/diagnostic modalities and application of effective treatment. Further wide scale research works could elaborate age-specific tumor molecular biology and the multiple factor's role in the development of BC at younger age.

In consistent with previous study (Traoré et al., 2019), most tumors were G3 followed by G2. Again, anther PAM50 based signature and overall survival study revealed tumors with higher grade (Pu et al., 2020). Previous study from Ethiopia indicated that tumors grade 2 were predominant followed by advanced histological grade 3(Hadgu et al., 2018). Similar to study from Ethiopia, finding from South Africa indicated that higher percentage of tumors were grade 2 (Dix-Peek et al., 2023).

Supporting, the aforementioned findings, in report from Sudanese and German cohort, Sudanese women had a higher grade than German women confirmed higher proportion of tumors were higher grade (Sengal et al., 2018).

The increased rate of high-grade tumors observed in this study is in part possibly due to late diagnosis of BC unlike developed countries where early diagnosis is in place. Since majority of the tumors from our finding tended to be with moderate or fast-growing nature and had a higher potential of spreading, timely diagnosis and initiation of treatment is critical. However, we have to be cautious while utilizing findings from health facility for various reasons. In agreement with our finding, previous study shown that most patients had early stage BC based on simplified staging (H. K. Kim et al., 2019) and invasive ductal carcinoma was the predominant histological subtype (Adani-Ifè et al., 2020; Bennis et al., 2012; Cherbal et al., 2015) followed by less common histological subtypes.

In this study, higher percentage of tumor with T1/T2 size would provide relevant information to a surgeon to undergo excision or mastectomy which ultimately prevent the spread of the tumor laterally or beyond. Furthermore, the early disease state during diagnosis suggests that there are improvements in health promotion of NCD, diagnosis and/or early detection and health-seeking behavior of the public. But much has to be done to improving early detection through employing cost-effective but reliable diagnostic modalities to track cases as early as possible. Moreover, the conventional methods have to be supported with advanced molecular tools for better management of patients considering tumor BC complexity and tumor heterogeneity.

In support of our findings, previous study including from Ethiopia and others indicated that higher percentage of tumors were ER+ and PR+ markers (Adani-Ifè et al., 2020; Pu et al., 2020; Sayed et al., 2014; Traoré et al., 2019). In countries where the prevalence of hormone receptor-positive tumors is very high, it is possible to apply hormonal therapy even among patients in whom HR status has not been conducted (Adjei et al., 2014; Kantelhardt et al., 2014). In our finding, one-fifth of the tumors were HER2 positive which is by far lower than previous reports from Eastern Morocco but close to other SSA countries (Adebamowo et al., 2008; Kantelhardt et al., 2014; Ly et al., 2012).

Based on ER, PR and HER2 status, we determined the IHC molecular subtyping. With this respect, previous study revealed that higher percentages of tumors were HR+/HER2– but it was in contrast with our finding because the least proportion in their study was TNBC which was the second ranked IHC group in the current stud (H. K. Kim et al., 2019; Luangxay et al., 2019; Tamaki et al., 2013; L. Wang et al., 2021). Study from Ethiopia also confirmed that HR+/HER2- tumors were predominantly identified (Hadgu et al., 2018). The dominance of HR+/HER2- tumor could be related to the high expression of both ER and PR proteins across breast tissues. Unlike our finding where majority of IHC-based group was categorized as HR+ /HER2-, previous study showed that the predominant was Triple negative , followed by HR+ / HER2- followed by HR- /HER2+ groups (Adani-Ifè et al., 2020).

Pattern of subtypes

In this study, applying RNA-expression based subtyping, more than half of the tumors were endocrine responsive. We observed 31% Luminal A, 27% Luminal B, 19% HER2 enriched and 23% Basal-like tumors where close to two-third were clearly the luminal subtypes. Though luminal subtypes somehow still dominant, considerably shifted to the more aggressive subtypes compared to e.g. data from the United States (Nurses' Health Study) reporting 46% Luminal A, 18% Luminal B, 14% HER2-enriched and 15% Basal-like (Kensler et al., 2020). Such direct comparison needs to be interpreted with care. The luminal A is most common and less aggressive whereas the luminal B is associated with higher grade and poor clinical outcome. The luminal type was followed by basal like which is also aggressive type of BC.

Despite late presentation of BC patients to health facility and lack of optimal care, the high mortality rate of BC in Ethiopia and LMIC may be associated with the predominance of aggressive BC types (luminal B and basal like, accounting for 50%) but it needs population-based study to make generalization. In Ethiopia, like women in other lower middle income countries (LMIC), have poorer overall BCsurvival rates (Kantelhardt. et al., 2014; Unger-Saldaña, 2014). Stage at diagnosis as well as availability of optimal care were among the factors associated with poor BCsurvival outcomes. Since, we didn't address the correlation of intrinsic subtypes with overall survival, further large-scale work is required. In previous study from Ethiopia (2014), it has been reported that two thirds of tumors from a patient cohorts from Addis Ababa had endocrine responsive disease (Kantelhardt et al., 2014).

Another cohort from rural Ethiopia showed more than half of the patients with positive hormone receptor status (Eber-Schulz et al., 2018). The third study published in 2018 again from Addis Ababa also reported 65% of receptor positive disease (Hadgu et al., 2018). In Sub-Saharan Africa, luminal A is more common in some countries (Adebamowo et al., 2008; Effi et al., 2017). Strengthening previous evidences, African-American women were more likely to have basal-like tumors compared with non-Hispanic whites (Sweeney et al., 2014). In another study from South Africa, typing using the PAM50 gave 19.3% luminal-A, 32.5% luminal-B, 23.5% HER2-enriched and 24.6% basal-like (Dix-Peek et al., 2023).

Again, in consistent with the current PAM50 based BC classifier results, the IHC based analysis from Ethiopia and beyond demonstrated that large majority of the tumor were luminal A and luminal B (Bennis et al., 2012; Hadgu et al., 2018; H. K. Kim et al., 2019; Serrano-Gomez et al., 2016; L. Wang et al., 2021). The finding is in agreement with previous studies which applied molecular subtyping using the PAM50 classifier and we found that the luminal A intrinsic subtype was predominant, followed by luminal B, basal-like and HER2-enriched (Kensler et al., 2019; Malvia et al., 2019; T. O. Nielsen et al., 2010; Ohnstad et al., 2017; L. Wang et al., 2021).

Another research-based PAM50 signature and long-term BCsurvival study clearly indicated that luminal subtypes were the leading followed by basal like and HER2-enriched (Pu et al., 2020). However, Luminal B which is the second most common molecular subtype in our study (26%) is present at higher proportion among our study participants when compared to the rate seen in the Western high resource regions, Asia and middle east (10–20%) (Alnegheimish et al., 2016). This was also in agreement with an African study where the typing using the PAM50 showed that luminal-B(32.5%) was dominant followed by basal-like (24.6%) (Dix-Peek et al., 2023). In another study from Africa, it has been shown that East and central SSA has similar rates of luminal, South Africa with high and west Africa with lower proportion of luminal type BC and TNBC more pronounced in West Africa but comparable in the other parts of Africa (Bauer et al., 2023).

Therefore, the considerable number of cases in this study harboring aggressive tumor phenotype which could be the possible reason for poor clinical outcome and related consequences among patients living with BC. Therefore, further understanding of population-based distribution of intrinsic subtypes is essential to identify such group of patients and implement effective BC intervention and management strategy.

In contrast with the present study, though luminal subtypes were dominant, the PAM50 based investigation showed that luminal B was prevalent followed by luminal A, HER2-enriched and basal like chronologically (Jorgensen et al., 2014). Unlike our study, another PAM50-based study demonstrated that basal-like ranked first followed by luminal A, luminal B, and HER2-enriched subtypes in order (Prat, Fan, et al., 2015). Additional study suggested that luminal A was dominant followed by basal-like subtypes (Kumar et al., 2015) with equal proportion of the other subtypes. Another PAM50 study indicated that luminal A, basal-like, HER2-enriched and luminal B as the prevalent subtypes in order (H. K. Kim et al., 2019). Based on the IHC analysis, unlike our findings where luminal B type ranked second, findings from previous study recorded luminal B as the predominant type followed by luminal A (Cherbal et al., 2015; Traoré et al., 2019). When using markers beyond HR and HER2, in previous work reported high prevalence of luminal B BC subtype (37.2%) followed by luminal A (26.3%) (Serrano-Gomez et al., 2016).

Such discrepancy in the proportion of intrinsic subtypes could be emanated from difference in genetic basis, tumor heterogeneity, inter-laboratory/classification criteria variation, methodological difference in evaluation of expression (Maisonneuve et al., 2014). Confirming this evidence, Cancer subtypes luminal A, luminal B, basal-like, and $HER2^{+}/ER^{-}$ were distributed as 64%, 11%, 11%, and 5% for whites, and 48%, 8%, 22%, and 7% for African Americans, respectively in the Carolina BC study (O'Brien et al., 2010). However, since the PAM50 analysis is more or less more coherent compared to the IHC data across Africa, our reliable results may help our oncologists and policy makers in what way to get prepared in the management of BC in Ethiopia.

As per documented evidences, country specific studies demonstrated that there are ethnic differences in the relative prevalence of BC intrinsic subtypes and found different distributions according to ethnicity. Though unclear, discrepancy in such intrinsic subtypes explained by difference in environmental exposures including the way people lead their life and nutritional related factors (Serrano-Gomez et al., 2016; Sweeney et al., 2014). But this is a gray area where it requires large scale population-based exploration of the subtype distribution to define context driven preventive and therapeutic measures. Since the current finding indicated that majority of the intrinsic subtypes are luminal type, hormonal therapy could be the possible alternative for treatment regardless of hormone receptor status.

But a nation-wide study investigating distribution intrinsic subtypes and further exploration of tumor biology is essential to capture the large number of different ethnic groups in Ethiopia which could possibly help to have a broaden understanding of BC horizon.

Comparing with other geographic regions

High quality data on tumor subtypes e.g. from the cancer genome atlas about tumors from African patients is lacking, but patients with African compared to European ancestry had a higher likelihood of basal-like (odds ratio OR 1.67) and HER2-enriched (OR 2.22) tumors(Sweeney et al., 2014) that has to be mentioned, that African-American patients genetic background is dominated from West Africa (Newman, 2014). Geographic variations have been observed across Africa concerning the composition of different hormone-receptor-based types no longer considering all tumors in Africa aggressive. A large meta-analysis found that pattern of tumor subtypes in East Africa appear to be more favorable compared to other geographic regions such as West Africa (Eng et al., 2014).

Clinical and histopathological factors influencing intrinsic subtypes

Our convenient hospital cohort has specific features. The age composition as well as proportions of early and later stage disease may influence the overall proportions of subtypes. In this study, the high proportion of patients below the age of 50 (approx. 60%) and advanced grade differed from high-income settings. This probably resulted in the lower proportions of luminal A subtypes compared to Western cohorts. In general, cohorts from sub–Saharan Africa seem to similarly represent a young population and commonly comprise of late-stage tumors compared to Western settings.

Our finding declared that the prevalence of luminal subtype was slightly higher among a younger age group. Moreover, the study finding demonstrated that luminal subtypes were dominant across all age groups. This contrasts with high-income countries data, in which older age groups usually present higher proportions of luminal subtype BC. These endocrine sensitive tumors are associated with obesity (Miyagawa et al., 2015) lack of physical activity, fewer children and less breast-feeding. In contrast with the current finding, black women of all ages had a statistically significantly lower frequency of luminal A type in younger than older patients whereas all other subtype frequencies were higher in black women (Troester et al., 2018).

Since these risk factors are not very common yet in Ethiopia, we noted that postulate that therefore such tumors are lacking in our cohort. Interestingly, previous studies showed that HER2- luminal tumors are more common among older, postmenopausal women and non-luminal subtypes are more common in younger women (Jenkins et al., 2014; Millikan et al., 2008).

Pooled data from publicly available clinical and gene expression analysis and use of PAM50 algorithm to classify BC intrinsic subtypes had shown incidence of luminal A, B, and A+B increase with age while basal like decrease with age (Jenkins et al., 2014). In our cohort, age was not associated with intrinsic subtype, whereas tumor grade was associated with intrinsic subtypes and this is consistent with other findings (T. O. Nielsen et al., 2010).

In a cohort from the United States that patients with luminal A tumors are on average 5-6 years older than patients with the other subtypes (Allott et al., 2020). A comparison of receptor status between Sudanese and German patients also showed considerably more low-grade, receptor-positive tumors in older German women (Sengal et al., 2018). Therefore, cohorts with fewer older patients lack these luminal tumors which is the case in our study. Additionally, body composition has been reported to influence subtype. Obesity has been associated with higher risk of Luminal A type BC (Miyagawa et al., 2015). Since obesity is still rare in Ethiopia, less Luminal A BCs are expected (Kwan et al., 2015). These endocrine sensitive tumors are associated with obesity (Miyagawa et al., 2015) lack of physical activity, fewer children and less breast-feeding. Unfortunately, we didn't explore the relationship between intrinsic subtypes and obesity, reproductive characteristics and behavior related parameters in our study; however, such exploration is essential to identify certain group of patients to be benefit from existing and/ or new intervention approach.

With respect to subtypes distribution across tumor grade, our study was in consistent with previous study (T. O. Nielsen et al., 2010) where higher proportion of luminal A followed by luminal B subtypes were G1/G2 where as in grade 3 tumor, higher percentage (18.9%) of the tumor were basal-like followed by luminal B (16.2%). Finding from Africa demonstrated that tumors with a luminal subtype were more likely to be of lower grade (grade 1 or 2) than basal-like subtypes (75.0% grade 3) (Dix-Peek et al., 2023). In agreement with the current finding , higher percentage of histological G3 tumors found identified in TNBC, basal-like, and HER2-enriched tumors in order (Serrano-Gomez et al., 2016) whereas in another study two third of basal-like subtypes were G3 (Kumar et al., 2015).

In support of generated evidences, (Kensler et al., 2019), the dominance of luminal subtypes across almost all favorable clinical and histopathological parameters convey that almost none are excluded from probable benefit of endocrine treatment.

In agreement with accumulated evidences, our study findings revealed that majority of the tumors expressed ER and PR in which luminal subtypes were dominant. Majority of the of HER2 positive tumors were non-luminal subtype with higher proportion among sub-group HER2-enriched subtypes. Consequently, targeted therapy for HER2 positive cases has to be in place for better care. In summary, in parallel with the higher proportion PAM50 based luminal type tumor, the IHC finding strengthen the direction towards to offering endocrine therapy by trained health professionals across Ethiopian health facility regardless of IHC marker status.

Concordance between IHC groups and PAM50-based intrinsic subtypes

A line of research findings demonstrated that the discovery of molecular subtypes has created a new tool for clinicians and researchers to further understand BC biology (Kwan et al., 2015). Though IHC is an important tool or gold standard to classify BC and guide therapy; however, even basic IHC test is not widely and readily available in Ethiopia. Even in capital cities, 15 out of 20 centers experienced frequent power cuts and four out of twenty had no immunohistochemistry in country (Ziegenhorn et al., 2020). But, the relevance of IHC-surrogates determination including ER, PR, HER2 and Ki67 well recommended according to 2011 St. Gallen guideline in the treatment plan of BC (Goldhirsch et al., 2013).

Despite its benefit, IHC is associated with false positive and false negative rate of up to 25% (Di Leo et al., 2012). Although it is expensive, the use of molecular subtyping to confirm the epidemiology of BC types in Ethiopia is one the major contribution of this work. This helps the policy makers to take into account what type of BC is predominant and plan what is required in the management of the disease. In our study, both methods had considerable similarities, especially for basal-like or triple-negative and HER2-enriched or HER2-positive types. In our finding we have observed significant linearity between intrinsic subtypes and IHC type somehow reassures the dominance of endocrine responsive disease and both methods are valuable in the classification of BC in the setting despite the known drawbacks. IHC results appear reliable as a basis for decision to advise endocrine treatment to the patient despite lack of standardized external quality assurance measures for tissue processing and immunohistochemistry performance (Jorns, 2019).

Due to the fact that also for Ethiopian patients' routine immunohistochemistry is not always available, seeing that more than half of patients have endocrine responsive disease confirmed by RNA-expression as well as immunohistochemistry gives confidence to offer endocrine treatment for patients with unknown receptor status.

Assessing tumor biology with RNA-expression analysis is considered gold standard, at the same time immunohistochemistry has been used as basis to prove effectiveness of endocrine treatment. When comparing both methods, certain discrepancies can be seen. In line with previous reports (Sørlie et al., 2001), we have observed a substantial mismatch between the classification based on gene expression (PAM50 intrinsic subtypes) and on immunohistochemistry (IHC groups using estrogen, progesterone and HER2).

Individualized therapy – luminal types

In this study, of the total tumors identified as luminal A and luminal B by PAM50-based classifier, more than three-fourth of the tumors were classified as HR positive/HER2 negative by IHC techniques whereas with in HR+/HER2– sub-cohort, the finding indicated that greater than 80 % of tumors were grouped as luminal followed by non-luminal phenotype. In previous works with in HR+/HER2– sub-cohort 58.3% were Luminal A and 80.0% of HR– /HER2+ sub-cohorts were HER2-enriched type (Bernard et al., 2009; Paquet & Hallett, 2015).

As a future perspective, nanostring-based subtype determination can also support individualized therapy in Ethiopia. Since a large proportion of tumors are larger than 2 cm or lymph node-positive, without additional prognostic markers, according to the National Cancer Control Network (NCCN) harmonized guidelines, nearly all patients would need chemotherapy. As an example, utilizing PAM50-based subtypes, we were able to split the large group of patients with HR-positive and HER2-negative tumors into clinically relevant, more homogenous intrinsic subtypes. Patients with HR+/HER2 tumors have been regarded as an especially challenging BC group due to high variance concerning clinical outcome (Balic et al., 2019). These patients include cases of low, intermediate and high risk of recurrence. Differentiating the HR+/HER2- group into Luminal A and non-luminal A intrinsic subtypes increases the ability to assess recurrence risk in a more individualized manner, since chemotherapy is not recommended for patients with Luminal A tumors (Chia et al., 2013; Jensen et al., 2018).

Individualized therapy – HER2 and basal-like types

The subtype determination can also personalize treatment for patients with aggressive tumors. The PAM50 assay allowed the identification of 62 tumors as HER2-enriched, regardless of HER2 status. Within HER2-enriched intrinsic subtype, 21(33.9%) were HR-/HER2+; however, within HR-/HER + tumors, 75% were classified as HER2-enriched subtype.

In recent study, of patients with HER2+ by IHC, majority had HER2-enriched intrinsic subtype (Carey et al., 2016). Among those tumors with basal-like intrinsic subtypes, two-third were HR-/HER- by IHC grouping. Of samples identified as HR- /HER2- by IHC, two-third and one-fifth were basal-like and HER2-enriched by PAM50 assay. Using the IHC method, only 41 patients would get an anti-HER2 therapy. This means that the PAM50 assay leads to an approximately 50% incremented number of patients who would benefit from drugs suppressing the HER2 signaling pathway (anti-HER2 therapy). Our results are concordant with the work of Perou and colleagues in the sense that two thirds of the samples classified as HER2-positive were found to be HER2-enriched (Perou et al., 2000).

Finally, 77 samples were classified as basal-like through the PAM50 assay. Out of these, only 46 (60%) were concordant with the triple-negative class via the IHC method. These samples are arguably the ones that would benefit most from chemotherapy, 40% of the basal-like samples are non-TNBC (Perou et al., 2000). In a study from Africa, the basal-like and TNC had the highest concordance, while the luminal-A and IHC-A group had the lowest concordance (Dix-Peek et al., 2023) Conversely, approximately 38% of the TNBC tumors were non-basal-like. Thus, both PAM50 and IHC classification methods yield heterogeneous groups where personalized recommendations would improve by additional molecular subtyping as proposed before (Hartung et al., 2021; Lehmann et al., 2011).

Association of intrinsic subtypes with clinical and histopathological parameters

In univariate analysis, the finding revealed that where tumor grade, ER, PgR, HR, HER2 status, Ki-67 and IHC group had a statistically significant association with intrinsic subtypes. While running the multivariate analysis, the study depicted that luminal A subtypes found to be associated with tumor grade 1 to 2, low Ki-67 expression, HR+/HER2- tumors whereas luminal B subtypes had association only with HR+/HER2-.

Regarding HER2-enriched tumors, they had a statistically significant association with HR+/HER2+, HR-/HER2+, HR-/HER2- tumors; however, basal like tumors found associated with tumor grade 3 and IHC group HR-/HER2-.

In previous IHC based typing, age at or during diagnosis, grade and recurrence had a statistically significant association with intrinsic subtype (Serrano-Gomez et al., 2016). Moreover, tumor grade, age at diagnosis and menopausal status had association with PAM50-based intrinsic subtypes (Pu et al., 2020). In previous research work, intrinsic subtypes were significantly different by premenopausal and postmenopausal status (Cherbal et al., 2015).

Another finding demonstrated that there was statistically significant difference in median age at or during diagnosis between the molecular subtypes where there was a bimodal distribution of molecular subtypes in different age ranges with luminal B subtype being more common at younger ages and luminal A subtype more prevalent at older ages (Hadgu et al., 2018).

PIK3CA gene mutations

The second major objective of this particular study was to investigate molecular aberrations including *PIK3CA* gene mutations and its prognostic role among BC cohorts. With previous evidences, molecular testing for genetic and genomic variations is becoming integral part of cancer care. A wide range of systematic investigations on cancer genomes had shown various somatic mutation profiles in each type of cancer, allowing better understanding of tumorigenesis, disease etiology, prognosis, and possibly aid implementation of personalized treatment (Van Hoeck et al., 2019).

Generally speaking, the availability of methods to detect genetic basis of the disease as well as the identification of targetable mutations are the driving forces behind the development in BC care. PI3K/Akt/mTOR pathway is shown to play a critical role in cellular proliferation, differentiation, migration, and angiogenesis. To this end, in the last couple of years the PI3K/Akt/mTOR pathway is under scrutiny for targeted therapy and identification of prognostic biomarkers. Currently, there are FDA approved drugs including Alpelisib to treat patients with HR+/HER2- tumors and carrying mutation in *PIK3CA* gene (André et al., 2019). To benefit patients with this treatment, stratifying patients based on IHC group, mRNA based intrinsic subtyping and determination of hotspot and/or sequence based whole exon profiling in *PIK3CA* gene is essential.

Therefore, we performed this study to determine *PIK3CA* mutation rate as well its prognostic role in Ethiopian BC patients where the large majority of them are HR+. Consequently, elucidation of mutations in *PIK3CA* gene would be insightful enough to apply targeted treatment and predicting clinical outcome.

PIK3CA mutation prevalence

In this study we included 226 newly diagnosed BC patients with, our results showed 20.4% of the tumors had mutations which is in consistent with previous study from France and India (Ahmad et al., 2016; Cizkova et al., 2013). Though we could not find much studies in Africa, few studies showed regional diversity. A study from Nigeria showed that the proportion of *PIK3CA* mutations was 35.8% (22/53) among HER2+ BC tumors (Ogenyi et al., 2021).

The rate of *PIK3CA* mutation turned 15.7% in Egypt (Elwy et al., 2017) among HER2+ disease and this is consistent with our findings in HER2-enriched groups the *PIK3CA* mutation rate was 15.9%.

Since the aforementioned studies from Africa focused on HER2+ disease, interpretation should be performed with caution; however, similarly, our finding confirmed that the proportion of *PIK3CA* mutation was higher in HER2+ tumors. A study that compared the role of genetic background in *PIK3CA* gene mutation had shown the rate of mutation was higher among Caucasians (34%) than in African America (23%) (Ademuyiwa et al., 2017) (335) and the rate in African American is comparable to our data.

In contrast to our study, higher proportion of *PIK3CA* gene mutation was observed in studies from China (38.05% to 46.5%) (Lian et al., 2020; Maruyama et al., 2007), Japan (29%) (Deng et al., 2019), India (34%) (Sudhakar et al., 2016), Brazil (27%) (Mangone et al., 2012), Iran (45%) (Tabesh et al., 2017) and Germany (26.7%) (Reinhardt et al., 2023). In another pooled data analysis from 10,319 and 6338 patients, *PIK3CA* mutations occurred in 32% and 35.7% of the patients respectively (Martínez-Saéz et al., 2020; Zardavas et al., 2018). In general, the aforementioned difference in the frequency and distribution of *PIK3CA* mutation is likely related to variations in geographical region and the genetic makeup of the population. Additionally, hotspot regions analyzed, sample source preservations, methods used for DNA extraction, and sensitivity of the assay employed. Since our qPCR method only targets 3 hotspots, a large-scale population-based study scanning mutational variants in whole set of exons 9 and 20 along with exploring downstream PI3K/AKT/mTOR signaling pathway would be needed.

PIK3CA mutation and clinical and histopathological features

The correlation of *PIK3CA* mutations with hormone-receptor status has received extensive investigation. Several studies evaluating the prognosis of BC patients suggest that *PIK3CA* mutations are ‘‘good mutations and *PIK3CA* gene mutations are significantly associated with both ER and PR expression”, which are believed to be favorable clinicopathologic features of BC (H. Wu et al., 2019). In various previous studies, multivariate analysis revealed none of the clinical and histopathological parameters had significant association with *PIK3CA* mutation in agreement with several studies where observed higher frequency of mutation in good prognostic features did not reach a statistically significant level (Michelucci et al., 2009; Tabesh et al., 2022).

In general, the relationship between *PIK3CA* mutations and clinical-pathological features of BC is inconclusive in different studies. The observed *PIK3CA* mutations in our study was prominent in the young age (<50) and G1 or G2 tumor which is comparable with previous study (Cizkova et al., 2013; Lian et al., 2020). In contrast, another study from India showed that *PIK3CA* mutations were observed frequently in older patients, smaller size tumors, and grade II tumors; however, none of the differences were significant (Ahmad et al., 2016). In our study, most *PIK3CA* mutations were observed in HR+/any HER2 tumors. Similar to our findings, accumulated evidences also showed that mutated tumors were more frequent in HR+ tumors (Buttitta et al., 2006; Saal et al., 2005; Stemke-hale et al., 2009). Accordingly, high frequency of mutated tumors tended ER+ positive in various studies (Lian et al., 2020; Maruyama et al., 2007; Palimaru et al., 2013).

However, despite the high *PIK3CA* mutation among favorable markers (ER+), exploration of the significance of the *PIK3CA* mutations among ER+ cases that received tamoxifen or not demonstrated that ER+ patients benefit from tamoxifen independently of the PIK3CA status (Pérez-Tenorio et al., 2007). While reviewing association of HER2 with *PIK3CA* mutation, previous studies showed contradicting findings; in one hand, *PIK3CA* mutations were associated with HER2− clinicopathological subtype (Buttitta et al., 2006) and in contrast, no association was determined (Barbareschi et al., 2007).

Studies have shown that acquired resistance to endocrine-based therapy is a challenge (C. Y. Liu et al., 2017). *PIK3CA* mutations have been associated with resistance to endocrine therapy, HER2-directed therapy, and cytotoxic therapy. Previous study proved that *PIK3CA*-mutated HR+/HER2- metastatic BC was less sensitive to chemotherapy (Mosele et al., 2020). Though

ET is the foremost medical treatment (Lei et al., 2020) , a significant percentage of patients, however, eventually have resistance due to several mechanisms, including the dysregulation of phosphoinositide 3 kinase (PI3K/Akt/mammalian target of rapamycin (mTOR) signaling pathway (Vasan et al., 2019).

Therefore, mutational analysis and combining ET and targeted therapy plays a role particularly in advanced disease. The *PIK3CA* inhibitor Alpelisib can specifically be utilized in postmenopausal women and men with advanced BC whose tumors have the *PIK3CA* mutation and are hormone receptor (HR) positive and HER2 negative. Such therapy is not available in Ethiopia, but could eventually be considered given the prevalence of *PIK3CA* mutations. Further studies targeting patients with HR+ tumors are warranted to understanding the benefit of the FDA approved PI3K inhibitor together with the established regimen in treating group of BC patients.

Recently, findings are unraveling the relation of HER2 over expression and *PIK3CA* mutation. Though established targeted therapy, Trastuzumab, benefits in treating BC that overexpress HER2, evidences are showing that not all women derive benefit, and resistant tumor phenotype is inevitable. In conclusion, activation of the PI3 kinase pathway, either via loss of the tumor suppressor *PTEN* or through oncogenic stimulation of *PIK3CA*, can mediate trastuzumab resistance (Esteva et al., 2010; Park & Davidson, 2007; L. Wang et al., 2011). Consequently, evaluation of *PIK3CA* mutation and PI3K pathway activation pathways could provide a biomarker to identify patients unlikely to respond to a trastuzumab-based therapy (Berns et al., 2007). In such circumstances, combining inhibitors of the PI3 kinase pathway in conjunction with trastuzumab therapy is critical to derive benefit to these patients. In a recent study on cell lines, it has been shown that cell lines with *PIK3CA* mutations (E545K and H1047R) were more resistant to trastuzumab than those without mutations (Kataoka et al., 2010). Therefore, planning PI3K inhibitors as a treatment modality could be an option to overcome resistance towards trastuzumab treatment.

Published data showed that the frequency of somatic mutations of the *PIK3CA* gene differed markedly according to hormone receptor and HER2 status; mutations occurred in 30–50% of ER +/HER2− early BC (Zardavas et al., 2018). In agreement with previous study, the present study confirmed that HR+/HER2− tumors had the highest proportion of mutations compared to HR−/any HER2 tumors which had the least. It has widely been proposed that *PIK3CA* mutations are associated with luminal subtype (Cizkova et al., 2012). Less *PIK3CA* mutations are observed in TNBC compared to HR+/any HER2 among in 10 publicly available studies

(Martínez-Saéz et al., 2020). In the present study, this results in the largest majority of tumors with mutations being HR+/any HER2. But, a comparatively low frequency of *PIK3CA* mutation (13.6 %) observed with TNBC tumors (Ahmad et al., 2016). In summary, our results of *PIK3CA* mutations within distinct IHC-types are comparable with published results from other geographic regions.

Association of PIK3CA gene mutation with clinicopathological features

Our data from Ethiopia showed that higher proportion of *PIK3CA* mutations were observed in younger patients, G1 or G2 tumor, early disease state (T1, T2), HR+, HR+/any HER2 and tumor with high Ki-67 proliferation index. Additionally, ER+, PR+, HR+ and HR+/HER2- tumors had a statistically significant association with overall *PIK3CA* and hot specific mutation. In previous studies, associations of *PIK3CA* mutations and clinical and pathological features were controversial. In agreement with our study, some studies showed association with favorable factors such as low histopathological grade, small macroscopic tumor size, ER+, PR+ , and HER2− tumors, low Ki67 labeling index, increased age, lymph node negativity, luminal/human epidermal growth factor receptor 2-enriched subtypes (Cizkova et al., 2012; Kalinsky et al., 2009; Maruyama et al., 2007; Pang et al., 2014; H. Wu et al., 2019; Zardavas et al., 2018). Controversially, some study showed that mutations were associated with larger tumor size (S. Y. Li et al., 2006), and older age (Mosele et al., 2020). In consistent with our study, another study showed a positive correlation between *PIK3CA* mutation and HER2 overexpression (Saal et al., 2005). Meanwhile, in some study, mutations were negatively correlated with triple-negative BC subtype but were not associated with age at diagnosis, tumor stage, lymph node status, tumor size, or HER2 status (H. Wu et al., 2019). Though not statistically significant, lower proportion of mutation noted in our study in agreement with the above study.

Though evidence demonstrated that *PIK3CA* mutations were associated with favorable prognostic factors such as ER and PR expression, it has been declared that mutation by itself is unlikely to be the single pivotal determinant of favorable responses to endocrine treatment (Pang et al., 2014). However, there is strong evidence that the signaling pathways associated with *PIK3CA* mutations was indicative of better clinical outcomes in ER/HER2 BC patients (Loi et al., 2013). Therefore, investigation of PIK3/Akt/mTOR signaling pathway activity is more important than the *PIK3CA* mutation itself with respect to prognosis.

Studies determining whether *PIK3CA* mutations are beneficial in tamoxifen-treated BC patients with other molecular features such as *PTEN* loss or *AKT1* mutations are warranted. But, despite the observed discrepancies, evidences indicated that the high frequency of *PIK3CA* mutation and the discovery of hotspot mutations have important clinical implications for diagnosis, prognosis and therapy (Karakas et al., 2006).

PIK3CA mutation and its prognostic role

In BC, *PIK3CA* mutations are considered as driver mutations; however, their prognostic role is yet not well illustrated. A previous study declared that mutations in the *PIK3CA* gene were associated with better recurrence-free survival (Kalinsky et al., 2009). In contrast, other previous findings showed that patients with *PIK3CA* mutations showed significantly worse survival, particularly those with positive estrogen receptor expression status or non-amplified HER2 (S. Y. Li et al., 2006). Among HR + /HER2- mBC pooling evidence across multiple studies from 33 arms across 11 trials indicated that *PIK3CA* mutation was associated with shorter median OS suggesting a negative prognostic value of *PIK3CA* mutations in metastatic BC (Fillbrunn et al., 2022). Survival analysis in a total of 150 patients showed that patients with wild type or mutations of *PIK3CA* had no differences in outcome.

In this study, multivariable analysis showed that advanced pT stage and HR-/HER2+ tumor types had the most unfavorable outcome in comparison with respective favorable clinical and histopathological features. However, *PIK3CA* mutational status did not have prognostic value neither in the multivariable nor among the stratified clinical and histopathological parameters. Prognostic and predictive role of *PIK3CA* mutation in different types of BC has shown differing results (Lux et al., 2016). Study from Italy, demonstrated that there was no association between the survival and *PIK3CA* mutational status (Michelucci et al., 2009).

In HER2+ metastatic or early BC having a wild type or mutated *PIK3CA* did not show any difference (Ibrahim et al., 2022). Another meta-analysis that assessed the prognostic role of PI3KCA mutation among 1,929 BC patients) identified the *PIK3CA* mutation was associated with poor outcomes (Sobhani et al., 2018). Besides, treatment of HR+ and HER2- BC patients who had *PIK3CA* mutations with α-specific PI3K inhibitor (Alpelisib (BYL719), had prolonged progression-free survival (André et al., 2019), indicating the predictive role of *PIK3CA* mutation for the PI3K inhibitor.

Findings of Studies on Breast Tissue Microbiome

After characterizing the Ethiopian BC cohort, we did microbial profiling of the paired tumor and normal adjust breast tissue. The goal was to see the difference in microbial profile and the association with clinicopathological features. In contrary to large and rigorous cross-sectional studies (Peters et al., 2023), we couldn't show a statistically significant altered microbiome composition between normal adjacent and tumor tissues. However, our finding indicated that specific microbiota abundance noted across clinical and histopathological features. Unfortunately, we couldn't compare it with healthy control tissues. Besides, our sample size was not big enough to show differences. In this study we proved, it is possible to get good enough samples for microbiota analysis in our settings, and also showed some kinds of trends in the microbiota difference between normal adjacent and tumor tissue. In the large rigor studies, the pattern of microbiota perturbation was not consistent (Peters et al., 2023). We next plan to conduct large study with prospective follow up on progression and treatment outcome.

A recent, narrative review on BC microbiota role in BC risk and progression indicates, only few evidences are available on the relationship between breast tumor microbiota and prognosis (Peters et al., 2023). However, several lines of evidence indicate the potential role of mammary microbiota in mediating BC development (J. Zhang et al., 2021). Interestingly, previous works indicated that there is a geographic variation in the type of microbiota among BC and healthy patients (Fernández et al., 2018). Such evidence points to distinct microbial signatures that might be related to cancer development and response to certain treatment (Fernández et al., 2018). However, the extent to which microbiota alterations (dysbiosis) contributes to BC development is unclear. Although the discovery of microbiota differences in breast milk was previously explored, recent works have begun to investigate the association between breast tissue microbial dysbiosis with cancer. Accordingly, these efforts indicate that in healthy breast and BC tissues, the phylum *Proteobacteria* predominated followed by *Firmicutes* (Hieken et al., 2016; Urbaniak et al., 2014).

In agreement with our study, previous studies demonstrated the breast is characterized by a high predominance of *Proteobacteria*, followed by the phyla *Firmicutes* and *Actinobacteria* (Hieken et al., 2016; Urbaniak et al., 2014; J. Zhang et al., 2021). In another study, *Proteobacteria* were prominently identified in the tumor tissues and conversely in the non-cancerous adjacent tissues Actinobacteria were with increased abundance (Thompson et al., 2017).

Furthermore, previous studies highlighted the difference in breast microbial composition and abundance between normal adjacent and tumor tissues and/or healthy controls (Smith et al., 2019; H. Wang et al., 2017). All the above mentioned microbes thrive in a fatty-acid-rich environment in the breast and are positively associated with adiposity (Murphy et al., 2015). Thus, it is not surprising that these bacteria are abundant in the breast. At the family level, our findings are in agreement with Klann et al., who also observed a higher relative abundance the family *Ruminococcaceae* and the genus *Akkermansia* in breast tumor tissue as compared to normal tissue and a lower relative abundance of *Bacteroides* and *Sutterella* in breast tumor compared to normal tissue (Klann et al., 2020).

More recently, exciting research on the breast microbiome to identify distinct microbial signatures considering different parameters, including race of patients with BC , tumor stage, and breast tumor subtype, was carried out (Smith et al., 2022). In agreement with our findings, it was observed that the phylum *Proteobacteria* was most abundant across all tissues, followed by *Firmicutes*, *Bacteroidetes*, and *Actinobacteria* regardless of race, and a higher abundance of the genus *Ralstonia*, which could explain in part a portion of the BC racial disparities (Smith et al., 2022). Considering the diverse population in the Ethiopian context, conducting a study to profile microbiota and understanding its association with certain prognostic features would be of paramount importance diagnosis, prognosis and treatment plan in certain group of patients with BC.

In efforts towards understanding the role of microbiota in health and disease, until now, it has remained elusive and/or unclear whether there is a difference in microbial composition in breast tumors and paired NAT of the same individual with BC (Urbaniak et al., 2014; Xuan et al., 2014). While comparing tumor tissue with NAT, the microbiota composition showed a distinct bacterial profile, suggesting the oncogenic effect of specific bacterial taxa (Hieken et al., 2016). In terms of relative distribution of microbiota, our sequencing data of NAT and tumor groups are supported by another study where out of 11 differentially abundant operational taxonomic units (OTUs), the majority of OTUs were abundant in paired normal tissue and close to one-third of the OTUs were abundant in tumor tissue. The bacterium *Sphingomonas yanoikuyae* was enriched in paired normal tissue, similar to our findings, whereas the bacterium *Methylobacterium radiotolerans* was most significantly enriched and prevalent in paired in tumor tissue (Xuan et al., 2014).

Microbial differences in paired breast tissues (tumor versus NAT) have also been identified in women from different geographical locations, which suggests that environment or lifestyle factors might impact the relative abundance of certain microbes in the breast microenvironment (Lehouritis et al., 2015; Urbaniak et al., 2014).

Generally, the breast TME is composed of a variety of cell types and microbiota. Studies suggest that pathophysiological changes occurring within the breast cells could have a significant impact on tumor growth (Thompson et al., 2017). Additionally, it has been discovered that certain microbial species identified in the human breast are potent agents that have the potential to trigger DNA damage, genomic instability, and alterations of molecular events in the form of mutations and epigenetic modifications (I. O. Lee et al., 2010). For instance, functional *in vitro* studies revealed that *Escherichia coli* and *Staphylococcus epidermidis* isolated from BC patients trigger DNA double-stranded breaks in HeLa cells (Urbaniak et al., 2016). Therefore, microbial communities within a host could be considered as an additional environmental factor that may contribute to or be influence carcinogenesis (German et al., 2023). However, still it is difficult to associate the significance of microbiota in the initiation, progression, and prediction of therapeutic responses among patients with BC.

Findings Associated with Clinical and Histopathological Features

In agreement with our findings, other studies reported a correlation between specific breast microbial taxa and certain prognostic features such as stage (Smith et al., 2019), receptor status (Banerjee et al., 2018; Nejman et al., 2020; H. Wang et al., 2017), and lymphovascular invasion (H. Wang et al., 2017). In our study, we found specific genera that were strongly associated with different intrinsic subtypes and by stage of disease. Specifically, we observed that the genus *Burkholderia* most strongly correlated with aggressive triple negative (TNBC) and basal-like breast tumors compared to less aggressive luminal A, luminal B, and HER2-E tumors. In support of our findings, a recent insightful study by Hoskinson et al. noted that a higher abundance of several microbial families, including *Burkholderiaceae*, was associated with early BC development when analyzing normal breast tissue compared to tissues donated prior to and after BC diagnosis (Hoskinson et al., 2022). These findings suggest that microbes in the genus *Burkholderia* may play an oncogenic role in the development of aggressive BC, such as TNBC and basal-like tumors.

In addition, we observed that several genera, including *Alkanindiges*, *Anoxybacillus*, *Leifsonia*, and *Exiguobacterium* most strongly correlated with HER2-E tumors. Luminal A and luminal B tumors weakly correlated with *Alkanindiges* and *Anoxybacillus*. Interestingly, Modica and colleagues reported that lymphovascular invasion correlated negatively with *Alkanindiges* in HER2-E tumors, which could explain the weak association between less aggressive luminal A and luminal B tumors in our study (Di Modica et al., 2022). On the other hand, *Alkanindiges* was correlated with invasive ductal carcinomas (Tzeng et al., 2021). Furthermore, we observed a relative higher abundance of the genus *Citrobacter* most significantly correlated with advanced-stage breast tumors compared to early-stage tumors. A supporting study by Yang et al. reported that women with malignant BC exhibited enriched levels of *Citrobacter* in their gut compared to those with benign tumors where *Citrobacter* was further associated with elevated glycan and lipopolysaccharide biosynthesis (P. Yang et al., 2021).

With a comparatively large sample size aimed at identifying bacterial genera with statistically different abundances, another study identified *Porphyromonas*, *Lacibacter*, *Ezakiella*, and *Fusobacterium* as being more abundant in higher-stage compared with lower-stage tumors (Tzeng et al., 2021). These findings highlight the complex nature of the microbiome and tumor interactions in the breast. Although we did not observe the distribution across tumor grade, in a separate study, multiple genera were significantly associated with histologic grade, with a number of genera present only in grade 1 tumors (Tzeng et al., 2021). However, additional studies using a larger sample size and comparable numbers of different groups of BCs are needed to confirm these associations. Collectively, most findings consistent with our study revealed that the breast houses a unique microbiota that can be distinguished based upon hormone receptor status, molecular subtype, and stage of disease (Hieken et al., 2016; Thompson et al., 2017; Urbaniak et al., 2014; Xuan et al., 2014).

Problems with Breast Microbiome Studies

Normally, a wide range of research findings employ various platforms to investigate breast tissue microbial profiles. Consequently, comparing and interpretating findings from different studies can be challenging due to differences in sample retrieval, processing, and methods employed to prevent bacterial contamination. According to recent evidence, microbiota of BC tissues are quite different from microbiota in breast tissues of women without BC (H. Wang et al., 2017). Though various studies have indicated that breast microbiota might play a role in mediating breast carcinogenesis (Donnet-Hughes et al., 2010; Goedert et al., 2015; Urbaniak et al., 2014), scientists are posing questions about how the microbiome might play a role in

modulating the risk of BC development. The hypothesis “changes in the composition of breast microbiota may also contribute to disease development and progression through several pathways”is still unclear whether the host´s microbial differences are a consequence or a cause of this human disease. This distinct phenomenon could be explained and/or interpreted in two ways: (1) alteration in microbiota profiles and/or dysbiosis comes first in the course of the carcinogenic process and may establish a microenvironment that predispose to cancer or (2) there is no correlation between the two events. However, knowledge of the microbiota of BC patients remains in its infancy. Given this context, additional studies have to be carried out to broaden knowledge on this topic and to understand which of the two possibilities occurs in patients with BC.

Considering the observed methodological differences in breast microbiome studies, including sample size, sample type, employed amplicon amplification technique, DNA extraction method utilized, sequencing approach, and potentially other factors, appreciating the precise difference in the composition and abundance of microbiota within a group of variables could be challenging in interpreting conclusions. However, in support of our findings, the literature indicates that differences in community composition across data sets can be attributed to ethnicity, dietary habits, geography, lactation status, method of sample collection, and platform of sequencing and data analysis (Parida & Sharma, 2020).

In addition to the contradicting scenarios referring to the role of microbiota as either a cause or effect of cancer, the differences in sample collection and processing, methodological differences in sequencing, and variations in recruited patient groups, along with the distinct clinical and histopathological parameters, underrepresentation of some geographical regions, and patient-oriented studies only with few inconsistent and/or contradicting in vitro results, could ultimately make the generalization of the role of microbiota in health and disease difficult in a broader context. Therefore, further studies addressing all of the noted pitfalls are necessary, particularly in developing countries including Ethiopia to improve BC care.

Why Shannon Diversity Was Not Different between NAT and Tumor Tissues

In our study, Shannon diversity or richness of microbial taxa was not found to be significantly different between NAT and tumor tissues; however, microbial evenness defined by Simpson’s index was slightly significant between the two groups. In support of our study, a previous study showed that adjacent normal paired breast tissue had a higher microbial diversity and richness than normal and tumor tissues (Smith et al., 2019).

Additionally, a previous study revealed that alpha diversity was significantly higher in normal compared to tumor samples where in unweighted UniFrac measures, breast tumor samples clustered distinctly from normal samples (R2 = 0.130; $p = 0.01$) (Klann et al., 2020). One study from Ghana revealed that alpha diversity was strongly and inversely associated with BC by tumor stage and molecular subtype (Byrd et al., 2021). In other studies, it was shown that BC patients had statistically significantly altered microbiota composition (beta diversity) and lower alpha diversity compared with healthy patients (Hoskinson et al., 2022) or breast tissues (Meng et al., 2018; Smith et al., 2019), whereas another study reported that there was no difference in bacterial communities between tumor tissue and NAT (Urbaniak et al., 2016), which may be explained by sample retrieval approaches (Hieken et al., 2016; H. Wang et al., 2017). Since we did not include breast tissue from women without BC, such a discrepancy in the relative alpha diversity between normal tissue and tumor samples is anticipated. Such a scenario will strengthen a well-articulated future study that includes diverse samples and specific patient groups.

Strengths and limitations

Our study has important strength referring the primary objective. Firstly, the study for the first time included a large number of samples from multiple health facilities in Ethiopia. Secondly, the cohort included a considerable proportion of young patients and advanced tumors thus representing a typical setting in Africa. Thirdly, we employed a multigene assay for gene expression profiling which is the gold standard for molecular classification as well as hormone-receptor status by immunohistochemistry. Both methods similarly showed a large share of endocrine responsive tumors. Certain limitations have to be taken into account. Firstly, the cohort is a convenient hospital-based sample and does not reflect the true pattern within the population. Given the resource-limited setting, population-based sample collection was not feasible. Moreover, follow up data was only available for some of the cohort because of the difficulty in obtaining the data in this setting. Again, complete therapy record was not available and was not taken into account while carrying out survival analysis.

Referring the microbiota investigation, though the finding we found is exciting and paved a way to understanding the composition of microbiota in Ethiopian patients, we believe that it would have been better if we included more samples to draw the full picture in the context. In this topic, we did not to include environmental and diet related parameters which are in most of the case a key variable in relation microbiota profile.

In previous studies, we have seen that they consider healthy controls, group of subjects who underwent breast cosmetic surgery (where we do not have such surgery in our context), to make comparison with breast tumor tissue which is lacked in study. Our findings have paved a way to understanding the composition of breast microbiota in Ethiopian BC Patients. Despite the plethora of microbiome studies in developed countries, it has been shown that African populations are under-represented, and there is an acute demand for studies related to the microbiome among African populations (Allali et al., 2022).

Therefore, taking into account context-specific factors and considerable limited evidence, microbiome studies must be replicated in Africa to extend our understanding of BC biology and to identify potential biomarkers used for prognosis and predicting therapeutic response. We believe that having a larger sample size consisting of fresh frozen breast tissues from healthy controls (without BC) and BC patients with benign and malignant tumors might assist in drawing conclusions as to whether with microbial differences are a consequence or a cause of BC. Additional patient information on lifestyle factors, such as diet and environmental exposures, are also critical components for understanding the role of microbial dysbiosis on the breast microenvironment.

mRNA versus IHC based ER, PR and HER 2determination

In this work, we used extracted RNA from FFPE and/or FFT tissue to examine the expression of the 3-BCrelated biomarkers ESR1, PGR, and HER2 utilizing performance PCR against IHC technique. A previous study demonstrated the viability of routinely assessing ESR1, PGR, and MKI67 using RT-qPCR to help choose BC patients for neoadjuvant treatment (Sinn et al., 2017). In our study, expression of marker was higher in FFT compared to FFPE samples. Utilizing the FFT samples, positivity increased for ER, for PgR and for HER2. Despite the increment in sensitivity, specificity was declined in FFT when compared with independent FFPE samples (n=160). Though getting FFT sample is not simple, considering this sample is critical to enhance BC care in parallel with using FFPE samples.

On the other hand, while analyzing FFPE samples, the sensitivity was lowered across the three markers whereas the specificity was higher unlike FFT samples. In case of FFPE samples, the sample quality is compromised because of formalin so that it diminishes positivity. Interestingly, a statistically significant agreement was noted between endpoint PCR and IHC for: ER (k = 0.332), PgR (k=0.135) and HER2 (k=0.134) in FFPE tissues.

With the 50 paired samples, positivity increased by 22% in ER, 28% in PgR and 28% in HER2 in FFT relative to FFPE samples. In matched tissues, a statistically significant moderate agreement was observed when using FFPE unlike FFT samples. Furthermore, concordance of biomarker expression between the qPCR and IHC methods was 94.44% (kappa =0.758) for ER, 91.67% (kappa = 0.806) for PR, and 46.67% (kappa = 0.436) for HER2.

In a recent study, the standard method IHC and quantitative RT-kPCR were compared for the assessment of ESR1, PGR, and HER2 expression levels. The results showed an excellent concordance between the two approaches for ESR1 status, with an agreement of 98%. PGR and HER2 expression agreements were somewhat lower, at 83% and 90%, respectively (Müller et al., 2011). The overall agreement between qIHC of ER and PgR and RT-qPCR of ESR1 and PGR was 96.6 and 91.4%, respectively, in a prior study. However, the overall agreement between visual IHC and RT-qPCR was somewhat lower for ER/ESR1 and PR/PGR (91.2 and 92.9%, respectively (Sinn et al., 2017).

In conclusion, a prior feasibility study demonstrates that the novel method's wide dynamic range, repeatability, and established agreement between IHC and RT-PCR bring up new possibilities in the realm of personalized medicine. This method is a significant technical advancement for the implementation of repeatable and economical testing of such biomarkers in clinical routine and research studies using archived FFPE material in molecular pathology diagnostic testing, since many upcoming biomarkers from gene expression studies are measured on the mRNA level (Müller et al., 2011). Due to the relatively small sample size, these data should be considered preliminary and worth validating in larger datasets.

In this recent study, the detection of protein expression with PCR based method possible when compared with the reference method. However, qPCR method has the upper hand in detecting the expression of surrogate markers than Endpoint PCR. Therefore, though qPCR methods are more expensive, they could serve as an alternative technique for breast cancer molecular markers determination so that a better treatment selection would be realized. On the other hand, we have observed that there is a difference in the level of detection of the molecular marker while using FFT and FFPE breast tissues. With this respect we noted that a significant positivity in FFT samples rather than FFPE samples. The possible explanation could be formalin cause cleavage of the nucleic acid so that it diminishes the level of detection of a particular mRNA transcript. Consequently, we recommend the use of PCR method and quality material to improving the detection rate and use the method as an alternative diagnostic modality to utilize appropriate treatment and enhance breast care in Ethiopia context.

7. CONCLUSION

In summary, we confirmed previous immunohistochemistry results showing a considerable proportion (>50%) of patients with BC were eligible for endocrine treatment. This will allow the decentralization of BC treatment with limited side-effects and could be administered even at lower-level health facilities. A study from Ethiopia proved that a strategy involving "cancer nurses" supporting patients during their 5-year treatment can improve adherence (Getachew et al., 2022). Efforts are needed to provide sufficient access to immunohistochemistry service but in the meantime can encourage to utilization of endocrine treatment for patients with even unknown receptor status until we get into states where IHC service is widely available. While administering the hormonal therapy, due attention is critical because of the emergence of endocrine therapy resistance associated with accumulation of molecular aberrations and/ or mutations related to various BC driving genes including *PIK3CA*.

With respect to molecular aberrations, the finding confirmed that *PIK3CA* gene mutation was higher in Ethiopian BC patients. Moreover, high proportion of mutated tumors were within favorable clinical and histopathological parameters in support of established research evidences. The proportion of *PIK3CA* mutations differed by BC subtype with HR+/HER2− disease having the highest proportion followed by HER2+ disease. We noted a statistically significant association between favorable clinical and histopathological parameters and the frequency of *PIK3CA* mutation. The survival analysis indicated that advanced stage and HR negative tumors tend to have worst overall survival. However, to understanding the whole picture of molecular pathways, predicting clinical outcomes and selecting effective treatment, coupling *PIK3CA* mutation spectrum study with further exploration of signaling pathways activation downstream of PI3K is essential.

The current study, which utilized fresh frozen breast tumor tissues and NAT from the same women, reported for the first time a unique microbial signature that correlates with prognostic features, including stage, IHC status, and PAM50 intrinsic subtypes, among Ethiopian women with BC. We identified 14 microbiota genera in breast tumor tissues that were distinct from NAT tissues. *Burkholderia* most strongly correlated with aggressive triple negative (TNBC) and basal-like breast tumors; whereas *Alkanindiges*, *Anoxybacillus, Leifsonia,* and *Exiguobacterium* most strongly correlated with HER2-E tumors. Luminal A and luminal B tumors also correlated with *Anoxybacillus* but not as strongly as HER2-E tumors. A relatively higher abundance of the genus *Citrobacter* most significantly correlated with advanced-stage breast tumors compared to early-stage tumors.

This is the first study to report an association between breast microbial dysbiosis and clinicopathological factors in Ethiopian women. Our findings encourage further precise characterization of local microbes in BC patients in Ethiopia for drug discovery and targeted microbial-based therapeutics, thus improving the prognosis and quality of life of BC patients. A future epidemiological study taking into account sample size, more controlled environmental conditions, high throughput metagenomics, and follow-up data is essential to strengthen conclusions related to clinical outcomes.

Furthermore, performance evaluation between PCR methods and IHC technique evaluated using FFPE and FFT samples. Accordingly, biomarker expression was higher while utilizing FFT than FFPE samples. In paired tissue, a statistically significant agreement was observed between the two methods in FFPE samples though sensitivity was higher in FFT samples. Improvement of agreement noted between qPCR and IHC in FFT samples. Finally, the finding confirmed that mRNA-based BC marker detection could be a complementary and/ or an alternative method in Ethiopia taking into account limited IHC service availability.

8. REFERENCES

Abate. (2016). Trends of Breast Cancer in Ethiopia. International Journal of Cancer Research and Molecular Mechanisms (ISSN 2381-3318), 2(1), 2–6.

Adani-Ifè, A., Amégbor, K., Doh, K., & Darré, T. (2020). Breast cancer in togolese women: immunohistochemistry subtypes. BMC Women's Health, 20(1), 1–7.

Adebamowo, C. A., Famooto, A., Ogundiran, T. O., & Aniagwu, T. (2008). Immunohistochemical and molecular subtypes of breast cancer in Nigeria. 17674190.

Ademuyiwa, F. O., Tao, Y., Luo, J., Weilbaecher, K., & Ma, C. X. (2017). Differences in the mutational landscape of triple-negative breast cancer in African Americans and Caucasians. Breast Cancer Research and Treatment, 161(3), 491–499. https://doi.org/10.1007/s10549-016-4062-y

Adjei, E. K., Owusu-afriyie, O., Awuah, B., & Stalsberg, H. (2014). Hormone receptors and Her2 expression in breast cancer in sub-Saharan Africa . A comparative study of biopsies from Ghana and Norway. 2–3. https://doi.org/10.1111/tbj.12261

Ahmad, F., Badwe, A., Verma, G., Bhatia, S., & Das, B. R. (2016). Molecular evaluation of PIK3CA gene mutation in breast cancer : determination of frequency , distribution pattern and its association with clinicopathological findings in Indian patients. 33(7), 10–11. https://doi.org/10.1007/s12032-016-0788-y

Aiad, H. A., Wahed, M. M. A., Asaad, N. Y., El-tahmody, M., & Elhosary, E. (2014). Immunohistochemical expression of GPR30 in breast carcinoma of Egyptian patients : an association with immunohistochemical subtypes. 2–3. ttps://doi.org/10.1111/apm.12241

Akarolo-Anthony, S. (2010). Introduction Global trends in cancer epidemiology. 12(Suppl 4). http://breast-cancer-research.com/supplements/12/S4/S8

Ali, A., Fooladi, I., Yazdi, M. H., & Pourmand, M. R. (2015). Th1 Cytokine Production Induced by Lactobacillus acidophilus in BALB / c Mice Bearing Transplanted Breast Tumor. 8(4), 1–5. https://doi.org/10.5812/jjm.8(4)2015.17354

Alizadeh, A. A., Eisen, M. B., Davis, R. E., Ma, C., Lossos, I. S., Rosenwald, A., Boldrick, J. C., Sabet, H., Tran, T., Chan, W. C., Greiner, T. C., Weisenburger, D. D., Armitage, J. O., Warnke, R., Levy, R., Wilson, W., & Grever, M. R. (2000). Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. 2–3. https://doi.org/10.1038/35000501

Allali, I., Abotsi, R. E., Tow, L. A., Thabane, L., & Zar, H. J. (2022). Human microbiota research in Africa : a systematic review reveals gaps and priorities for future research. 2021.

Allegra, J. C., Lippman, M. E., Thompson, E. B., Simon, R., Barlock, A., Green, L., Huff, K. K., Do, H. M. Y. T., Aitken, S. C., & Warren, R. (1980). Estrogen receptor status: an important variable in predicting response to endocrine therapy in metastatic breast cancer. European Journal of Cancer (1965), 16(3), 323–331. https://doi.org/10.1016/0014-2964(80)90348-5

Allison, K. H., Hammond, M. E. H., Dowsett, M., McKernin, S. E., Carey, L. A., Fitzgibbons, P. L., Hayes, D. F., Lakhani, S. R., Chavez-MacGregor, M., Perlmutter, J., Perou, C. M., Regan, M. M., Rimm, D. L., Symmans, W. F., Torlakovic, E. E., Varella, L., Viale, G., Weisberg, T. F., McShane, L. M., & Wolff, A. C. (2020). Estrogen and progesterone receptor testing in breast cancer: ASCO/CAP guideline update. Journal of Clinical

Oncology, 38(12), 1346–1366. https://doi.org/10.1200/JCO.19.02309

Allott, E. H., Shan, Y., Chen, M., Sun, X., Garcia-Recio, S., Kirk, E. L., Olshan, A. F., Geradts, J., Earp, H. S., Carey, L. A., Perou, C. M., Pfeiffer, R. M., Anderson, W. F., & Troester, M. A. (2020). Bimodal age distribution at diagnosis in breast cancer persists across molecular and genomic classifications. Breast Cancer Research and Treatment, 179(1), 185–195. https://doi.org/10.1007/s10549-019-05442-2

Allred, D. C., Clinic, G. H., Craig, D., & Allred, D. C. (2008). O ncologist Commentary : Hormone Receptor Testing in Breast Cancer : 1134–1136. https://doi.org/10.1634/theoncologist.2008-0184

Alnegheimish, N. A., Alshatwi, R. A., Alhefdhi, R. M., Arafah, M. M., Alrikabi, A. C., & Husain, S. (2016). Molecular subtypes of breast carcinoma in Saudi Arabia: A retrospective study. Saudi Medical Journal, 37(5), 506–512. https://doi.org/10.15537/smj.2016.5.15000

Anderson, B. O., Cazap, E., El Saghir, N. S., Yip, C. H., Khaled, H. M., Otero, I. V., Adebamowo, C. A., Badwe, R. A., & Harford, J. B. (2011). Optimisation of breast cancer management in low-resource and middle-resource countries: Executive summary of the Breast Health Global Initiative consensus, 2010. The Lancet Oncology, 12(4), 387–398. https://doi.org/10.1016/S1470-2045(11)70031-6

Anderson, B. O., Ilbawi, A. M., Fidarova, E., Weiderpass, E., & Stevens, L. (2021). The Global Breast Cancer Initiative : a strategic collaboration to strengthen health care for non-communicable diseases. 33691141. https://doi.org/10.1016/S1470-2045(21)00071-1

André, F., Ciruelos, E., Rubovszky, G., Campone, M., Loibl, S., Rugo, H. S., Iwata, H., Conte, P., Mayer, I. A., Kaufman, B., Yamashita, T., Lu, Y.-S., Inoue, K., Takahashi, M., Pápai, Z., Longin, A.-S., Mills, D., Wilke, C., Hirawat, S., & Juric, D. (2019). Alpelisib for PIK3CA -Mutated, Hormone Receptor–Positive Advanced Breast Cancer . New England Journal of Medicine, 380(20), 1929–1940. https://doi.org/10.1056/nejmoa1813904

Angeli, D., Salvi, S., & Tedaldi, G. (2020). Genetic predisposition to breast and ovarian cancers: How many and which genes to test? International Journal of Molecular Sciences, 21(3). https://doi.org/10.3390/ijms21031128

Arnold, M., Morgan, E., Rumgay, H., Mafra, A., Singh, D., Laversanne, M., Vignat, J., Gralow, J. R., Cardoso, F., Siesling, S., & Soerjomataram, I. (2022). Current and future burden of breast cancer: Global statistics for 2020 and 2040. Breast, 66(September), 15–23. https://doi.org/10.1016/j.breast.2022.08.010

Ashford, N. A., Bauman, P., Brown, H. S., Clapp, R. W., Finkel, A. M., Gee, D., Hattis, D. B., Martuzzi, M., Sasco, A. J., & Sass, J. B. (2015). Cancer risk: Role of environment. Science, 347(6223), 727. https://doi.org/10.1126/science.aaa6246

Autier, P., Boniol, M., Vecchia, C. La, Vatten, L., Gavin, A., & Héry, C. (2010). Disparities in breast cancer mortality trends between 30 European countries : retrospective trend analysis of WHO mortality database. 23–24. https://doi.org/10.1136/bmj.c3620

Avan, A., Narayan, R., Giovannetti, E., Peters, G. J., Avan, A., Giovannetti, E., & Peters, G. J. (2016). World Journal of Clinical Oncology © 2016. 7(5), 352–370. https://doi.org/10.5306/wjco.v7.i5.352

Awedew, A. F., Asefa, Z., & Belay, W. B. (2022). National Burden and Trend of Cancer in Ethiopia, 2010-2019: a systemic analysis for Global burden of disease study. Scientific Reports, 12(1), 12736. https://doi.org/10.1038/s41598-022-17128-9

Ayele, W., Addissie, A., Wienke, A., Unverzagt, S., & Jemal, A. (2021). Breast Awareness , Self-Reported Abnormalities , and Breast Cancer in Rural Ethiopia : A Survey of 7 , 573 Women and Predictions of the National Burden. 2–3. https://doi.org/10.1002/onco.13737

Ayele, W., Führer, A., Braun, G. A., Formazin, F., Wienke, A., Taylor, L., Unverzagt, S., Addissie, A., & Kantelhardt, E. J. (2022). Breast cancer morbidity and mortality in rural Ethiopia: data from 788 verbal autopsies. BMC Women's Health, 22(1), 1–9. https://doi.org/10.1186/s12905-022-01672-7

Bader, A. G., Kang, S., & Vogt, P. K. (2006). Cancer-specific mutations in PIK3CA are oncogenic in vivo. Proceedings of the National Academy of Sciences of the United States of America, 103(5), 1475–1479. https://doi.org/10.1073/pnas.0510857103

Baker, C. L., Vaughn, C. P., & Samowitz, W. S. (2012). A PIK3CA pyrosequencing-based assay that excludes pseudogene interference. Journal of Molecular Diagnostics, 14(1), 56–60. https://doi.org/10.1016/j.jmoldx.2011.08.004

Balic, M., Thomssen, C., Würstlein, R., Gnant, M., & Harbeck, N. (2019). St. Gallen/Vienna 2019: A brief summary of the consensus discussion on the optimal primary breast cancer treatment. Breast Care, 14(2), 103–110. https://doi.org/10.1159/000499931

Banerjee, S., Tian, T., Wei, Z., Shih, N., Feldman, M. D., Peck, K. N., DeMichele, A. M., Alwine, J. C., & Robertson, E. S. (2018). Distinct microbial signatures associated with different breast cancer types. Frontiers in Microbiology, 9(MAY). https://doi.org/10.3389/fmicb.2018.00951

Banin Hirata, B. K., Oda, J. M. M., Losi Guembarovski, R., Ariza, C. B., Oliveira, C. E. C. De, & Watanabe, M. A. E. (2014). Molecular markers for breast cancer: Prediction on tumor behavior. Disease Markers, 2014, 1–28. https://doi.org/10.1155/2014/513158

Barbareschi, M., Buttitta, F., Felicioni, L., Cotrupi, S., Barassi, F., Del Grammastro, M., Ferro, A., Palma, P. D., Galligioni, E., & Marchetti, A. (2007). Different prognostic roles of mutations in the helical and kinase domains of the PIK3CA gene in breast carcinomas. Clinical Cancer Research, 13(20), 6064–6069. https://doi.org/10.1158/1078-0432.CCR-07-0266

Bard, J., Luu, H. T., Dravet, F., Michel, C., Moyon, T., Pagniez, A., Nazih, H., & Bobin-dubigeon, C. (2015). Relationship Between Intestinal Microbiota and Clinical Characteristics of Patients with Early Stage Breast Cancer. The FASEB Journal, 29, 12–13.

Barnes, B. B. E., Steindorf, K., Hein, R., Flesch-Janys, D., & Chang-Claude, J. (2011). Population attributable risk of invasive postmenopausal breast cancer and breast cancer subtypes for modifiable and non-modifiable risk factors. Cancer Epidemiology, 35(4), 345–352. https://doi.org/10.1016/j.canep.2010.11.003

Bauer, M., Vetter, M., Stückrath, K., Yohannes, M., Desalegn, Z., Yalew, T., Bekuretsion, Y., Kenea, T. W., Joffe, M., van den Berg, E. J., Nikulu, J. I., Bakarou, K., Manraj, S. S., Ogunbiyi, O. J., Ekanem, I. O., Igbinoba, F., Diomande, M., Adebamowo, C., Dzamalala, C. P., … Kantelhardt, E. J. (2023). Regional Variation in the Tumor Microenvironment, Immune Escape and Prognostic Factors in Breast Cancer in Sub-Saharan Africa. Cancer Immunology Research, 11(6), 720–731. https://doi.org/10.1158/2326-6066.CIR-22-0795

Bayraktar, S., & Glück, S. (2013). Molecularly targeted therapies for metastatic triple-negative breast cancer. Breast Cancer Research and Treatment, 138(1), 21–35. https://doi.org/10.1007/s10549-013-2421-5

Beelen, K., Opdam, M., Severson, T. M., Koornstra, R. H. T., Vincent, A. D., Wesseling, J., Muris, J. J., Berns, E. M. J. J., Vermorken, J. B., van Diest, P. J., & Linn, S. C. (2014). PIK3CA mutations, phosphatase and tensin homolog, human epidermal growth factor receptor 2, and insulin-like growth factor 1 receptor and adjuvant tamoxifen resistance in postmenopausal breast cancer patients. Breast Cancer Research, 16(1), 1–10. https://doi.org/10.1186/bcr3606

Behrouzi, A., Nafari, A. H., & Siadat, S. D. (2019). The significance of microbiome in personalized medicine. Clinical and Translational Medicine, 8(1). https://doi.org/10.1186/s40169-019-0232-y

Belachew, E. B., Desta, A. F., Deneke, D. B., & Fenta, B. D. (2023). Clinicopathological Features of Invasive Breast Cancer : A Five-Year Retrospective Study in Southern and South-Western Ethiopia. 1–15. https://doi.org/10.3390/medicines10050030

Bennis, S., Abbass, F., Akasbi, Y., Znati, K., Joutei, K. A., El Mesbahi, O., & Amarti, A. (2012). Prevalence of molecular subtypes and prognosis of invasive breast cancer in north-east of Morocco: Retrospective study. BMC Research Notes, 5. https://doi.org/10.1186/1756-0500-5-436

Bernard, P. S., Parker, J. S., Mullins, M., Cheung, M. C. U., Leung, S., Voduc, D., Vickery, T., Davies, S., Fauron, C., He, X., Hu, Z., Quackenhush, J. F., Stijleman, I. J., Palazzo, J., Matron, J. S., Nobel, A. B., Mardis, E., Nielsen, T. O., Ellis, M. J., & Perou, C. M. (2009). Supervised risk predictor of breast cancer based on intrinsic subtypes. Journal of Clinical Oncology, 27(8), 1160–1167. https://doi.org/10.1200/JCO.2008.18.1370

Berns, K., Horlings, H. M., Hennessy, B. T., Madiredjo, M., Hijmans, E. M., Beelen, K., Linn, S. C., Gonzalez-Angulo, A. M., Stemke-Hale, K., Hauptmann, M., Beijersbergen, R. L., Mills, G. B., van de Vijver, M. J., & Bernards, R. (2007). A Functional Genetic Approach Identifies the PI3K Pathway as a Major Determinant of Trastuzumab Resistance in Breast Cancer. Cancer Cell, 12(4), 395–402. https://doi.org/10.1016/j.ccr.2007.08.030

Berry, D. A., Cronin, K. A., Plevritis, S. K., Fryback, D. G., Clarke, L., Zelen, M., Mandelblatt, J. S., Yakovlev, A. Y., Habbema, J. D. F., & Feuer, E. J. (2005). Effect of Screening and Adjuvant Therapy on Mortality from Breast Cancer. New England Journal of Medicine, 353(17), 1784–1792. https://doi.org/10.1056/nejmoa050518

Bertucci, F., Finetti, P., Cervera, N., Esterni, B., Hermitte, F., Viens, P., & Birnbaum, D. (2008). How basal are triple-negative breast cancers? International Journal of Cancer, 123(1), 236–240. https://doi.org/10.1002/ijc.23518

Bharti, V., Singh, S., Jain, N., & Ahirwal, L. (2015). Cytotoxicity of live whole cell , heat killed cell and cell free extract of lactobacillus strain in u-87 human glioblastoma cell line and mcf-7 breast cancer cell line LACTOBACILLUS STRAIN IN U-87 HUMAN GLIOBLASTOMA. August 2020.

Bhatt, A. P., Redinbo, M. R., & Bultman, S. J. (2017). The role of the microbiome in cancer development and therapy. CA: A Cancer Journal for Clinicians, 67(4), 326–344. https://doi.org/10.3322/caac.21398

Bik, E. M., Eckburg, P. B., Gill, S. R., Nelson, K. E., Purdom, E. A., Francois, F., Perez-Perez, G., Blaser, M. J., & Relman, D. A. (2006). Molecular analysis of the bacterial microbiota in the human stomach. Proceedings of the National Academy of Sciences of the United States of America, 103(3), 732–737. https://doi.org/10.1073/pnas.0506655103

Blamey, R. W., Ball, G., Cataliotti, L., Fourquet, A., Gee, J., & Holli, K. (2010). ONCOPOOL

- a European database for 16 , 944 cases of breast cancer. 2–3. https://doi.org/10.1016/j.ejca.2009.09.009

Blows, F. M., Driver, K. E., Schmidt, M. K., Broeks, A., Leeuwen, F. E. Van, Wesseling, J., Cheang, M. C., Gelmon, K., Nielsen, T. O., Blomqvist, C., Foulkes, W. D., Heikkinen, T., Nevanlinna, H., Akslen, L. A., Be, L. R., Couch, F. J., Wang, X., Cafourek, V., Olson, J. E., … Paul, D. (2010). Subtyping of Breast Cancer by Immunohistochemistry to Investigate a Relationship between Subtype and Short and Long Term Survival : A Collaborative Analysis of Data for 10 , 159 Cases from 12 Studies. 7(5). https://doi.org/10.1371/journal.pmed.1000279

Bombonati, A., & Sgroi, D. C. (2011). The molecular pathology of breast cancer progression. 23–24. https://doi.org/10.1002/path.2808

Boyault, S., Drouet, Y., Navarro, C., Bachelot, T., Lasset, C., Treilleux, I., Tabone, E., Puisieux, A., & Wang, Q. (2012). Mutational characterization of individual breast tumors: TP53 and PI3K pathway genes are frequently and distinctively mutated in different subtypes. Breast Cancer Research and Treatment, 132(1), 29–39. https://doi.org/10.1007/s10549-011-1518-y

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: A Cancer Journal for Clinicians, 68(6), 394–424. https://doi.org/10.3322/caac.21492

Bray, F., Jemal, A., Grey, N., Ferlay, J., & Forman, D. (2012). Global cancer transitions according to the Human Development Index (2008-2030): A population-based study. The Lancet Oncology, 13(8), 790–801. https://doi.org/10.1016/S1470-2045(12)70211-5

Bruhn, M. A., Pearson, R. B., Hannan, R. D., & Sheppard, K. E. (2013). AKT-independent PI3-K signaling in cancer – emerging role for SGK3. 281–292.

Bultman, S. J. (2014). Emerging roles of the microbiome in cancer. Carcinogenesis, 35(2), 249–255. https://doi.org/10.1093/carcin/bgt392

Buttitta, F., Felicioni, L., Barassi, F., Martella, C., Paolizzi, D., Fresu, G., Salvatore, S., Cuccurullo, F., Mezzetti, A., Campani, D., & Marchetti, A. (2006). PIK3CA mutation and histological type in breast carcinoma: High frequency of mutations in lobular carcinoma. Journal of Pathology, 208(3), 350–355. https://doi.org/10.1002/path.1908

Byrd, D. A., Vogtmann, E., Wu, Z., Han, Y., Wan, Y., Clegg-Lamptey, J. N., Yarney, J., Wiafe-Addai, B., Wiafe, S., Awuah, B., Ansong, D., Nyarko, K., Hullings, A. G., Hua, X., Ahearn, T., Goedert, J. J., Shi, J., Knight, R., Figueroa, J. D., … Sinha, R. (2021). Associations of fecal microbial profiles with breast cancer and nonmalignant breast disease in the Ghana Breast Health Study. International Journal of Cancer, 148(11), 2712–2723. https://doi.org/10.1002/ijc.33473

Campbell, R. A., Bhat-Nakshatri, P., Patel, N. M., Constantinidou, D., Ali, S., & Nakshatri, H. (2001). Phosphatidylinositol 3-kinase/AKT-mediated activation of estrogen receptor α: A new model for anti-estrogen resistance. Journal of Biological Chemistry, 276(13), 9817–9824. https://doi.org/10.1074/jbc.M010840200

Cantley, L. C. (2002). The phosphoinositide 3-kinase pathway. Science, 296(5573), 1655–1657. https://doi.org/10.1126/science.296.5573.1655

Carey, L. A., Berry, D. A., Cirrincione, C. T., Barry, W. T., Pitcher, B. N., Harris, L. N., Ollila, D. W., Krop, I. E., Henry, N. L., Weckstein, D. J., Anders, C. K., Singh, B., Hoadley, K.

A., Iglesia, M., Cheang, M. C. U., Perou, C. M., Winer, E. P., & Hudis, C. A. (2016). Molecular Heterogeneity and Response to Neoadjuvant Human Epidermal Growth Factor Receptor 2 Targeting in CALGB 40601 , a Randomized Phase III Trial of Paclitaxel Plus Trastuzumab With or Without Lapatinib. 2–4. https://doi.org/10.1200/JCO.2015.62.1268

Carey, L. A., Dees, E. C., Sawyer, L., Gatti, L., Moore, D. T., Collichio, F., Ollila, D. W., Sartor, C. I., Graham, M. L., & Perou, C. M. (2007). The triple negative paradox: Primary tumor chemosensitivity of breast cancer subtypes. Clinical Cancer Research, 13(8), 2329–2334. https://doi.org/10.1158/1078-0432.CCR-06-1109

Carey, L. A., Perou, C. M., Livasy, C. A., Dressler, L. G., Cowan, D., Conway, K., Karaca, G., Troester, M. A., Chiu, K. T., Edmiston, S., Deming, S. L., Geradts, J., Cheang, M. C. U., Nielsen, T. O., Moorman, P. G., Earp, H. S., & Millikan, R. C. (2006). Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. Journal of the American Medical Association, 295(21), 2492–2502. https://doi.org/10.1001/jama.295.21.2492

Chamseddine, R. S., Wang, C., Yin, K., Wang, J., Singh, P., Zhou, J., Robson, M. E., Braun, D., & Hughes, K. S. (2022). Penetrance of male breast cancer susceptibility genes: a systematic review. Breast Cancer Research and Treatment, 191(1), 31–38. https://doi.org/10.1007/s10549-021-06413-2

Chen, J., Douglass, J., Prasath, V., Neace, M., Atrchian, S., Manjili, M. H., Shokouhi, S., & Habibi, M. (2019). The microbiome and breast cancer: a review. Breast Cancer Research and Treatment, 178(3), 493–496. https://doi.org/10.1007/s10549-019-05407-5

Chen, Y. J., Huang, C. S., Phan, N. N., Lu, T. P., Liu, C. Y., Huang, C. J., Chiu, J. H., Tseng, L. M., & Huang, C. C. (2021). Molecular subtyping of breast cancer intrinsic taxonomy with oligonucleotide microarray and NanoString nCounter. Bioscience Reports, 41(8), 1–10. https://doi.org/10.1042/BSR20211428

Cherbal, F., Gaceb, H., Mehemmai, C., Saiah, I., Bakour, R., Rouis, A. O., Boualga, K., Benbrahim, W., & Mahfouf, H. (2015). Distribution of molecular breast cancer subtypes among Algerian women and correlation with clinical and tumor characteristics: A population-based study. Breast Disease, 35(2), 95–102. https://doi.org/10.3233/BD-150398

Chia, S. K., Bramwell, V. H., Tu, D., Shepherd, L. E., Jiang, S., Vickery, T., Mardis, E., Leung, S., Ung, K., Pritchard, K. I., Parker, S., Bernard, P. S., Perou, C. M., Ellis, M. J., Torsten, O., & Carolina, N. (2013). Prediction of Benefit from Adjuvant Tamoxifen. 18(16), 4465–4472. https://doi.org/10.1158/1078-0432.CCR-12-0286.A

Chung, Y. R., Kim, H. J., Kim, Y. A., Chang, M. S., Hwang, K., & Park, S. Y. (2017). Diversity index as a novel prognostic factor in breast cancer. 8(57), 97114–97126.

Cizkova, M., Dujaric, M. E., Lehmann-Che, J., Scott, V., Tembo, O., Asselain, B., Pierga, J. Y., Marty, M., De Cremoux, P., Spyratos, F., & Bieche, I. (2013). Outcome impact of PIK3CA mutations in HER2-positive breast cancer patients treated with trastuzumab. British Journal of Cancer, 108(9), 1807–1809. https://doi.org/10.1038/bjc.2013.164

Cizkova, M., Susini, A., Vacher, S., Cizeron-Clairac, G., Andrieu, C., Driouch, K., Fourme, E., Lidereau, R., & Bièche, I. (2012). PIK3CA mutation impact on survival in breast cancer patients and in ERα, PR and ERBB2-based subgroups. Breast Cancer Research, 14(1), 1–9. https://doi.org/10.1186/bcr3113

Coates, A. S., Winer, E. P., Goldhirsch, A., Gelber, R. D., Gnant, M., Thürlimann, B., Senn, H., & Members, P. (2015). Tailoring therapies — improving the management of early

breast cancer : St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2015. May, 1533–1546. https://doi.org/10.1093/annonc/mdv221

Coughlin, S. S., & Ekwueme, D. U. (2009). Breast cancer as a global health concern. Cancer Epidemiology, 33(5), 315–318. https://doi.org/10.1016/j.canep.2009.10.003

Curigliano, G., Burstein, H. J., Winer, E. P., Gnant, M., Dubsky, P., Loibl, S., Colleoni, M., M. Regan, M., Piccart-Gebhart, M., Senn, H. J., Thürlimann, B., André, F., Baselga, J., Bergh, J., Bonnefoi, H., Brucker, S. Y., Cardoso, F., Carey, L., Ciruelos, E., … Xu, B. (2017). De-escalating and escalating treatments for early-stage breast cancer: The St. Gallen International Expert Consensus Conference on the Primary Therapy of Early Breast Cancer 2017. Annals of Oncology, 28(8), 1700–1712. https://doi.org/10.1093/annonc/mdx308

Cuzick, J., Dowsett, M., Pineda, S., Wale, C., Salter, J., Quinn, E., Zabaglo, L., Mallon, E., Green, A. R., Ellis, I. O., Howell, A., & Buzdar, A. U. (2011). Prognostic value of a combined estrogen receptor , progesterone receptor , Ki-67 , and human epidermal growth factor receptor 2 immunohistochemical score and comparison with the Genomic Health recurrence score in early breast cancer. 10–11. https://doi.org/10.1200/JCO.2010.31.2835

Dai, X., Li, T., Bai, Z., Yang, Y., Liu, X., Zhan, J., & Shi, B. (2015). Breast cancer intrinsic subtype classification , clinical use and future trends. 5(10), 2929–2943.

De Martel, C., Ferlay, J., Franceschi, S., Vignat, J., Bray, F., Forman, D., & Plummer, M. (2012). Global burden of cancers attributable to infections in 2008: A review and synthetic analysis. The Lancet Oncology, 13(6), 607–615. https://doi.org/10.1016/S1470-2045(12)70137-7

de Martel, C., Georges, D., Bray, F., Ferlay, J., & Clifford, G. M. (2020). Global burden of cancer attributable to infections in 2018: a worldwide incidence analysis. The Lancet Global Health, 8(2), e180–e190. https://doi.org/10.1016/S2214-109X(19)30488-7

De Ronde, J. J., Hannemann, J., Halfwerk, H., Mulder, L., Straver, M. E., Vrancken Peeters, M. J. T. F. D., Wesseling, J., Van De Vijver, M., Wessels, L. F. A., & Rodenhuis, S. (2010). Concordance of clinical and molecular breast cancer subtyping in the context of preoperative chemotherapy response. Breast Cancer Research and Treatment, 119(1), 119–126. https://doi.org/10.1007/s10549-009-0499-6

Deng, L., Zhu, X., Sun, Y., Wang, J., Zhong, X., Li, J., Hu, M., & Zheng, H. (2019). Prevalence and prognostic role of PIK3CA/AKT1 mutations in Chinese breast cancer patients. Cancer Research and Treatment, 51(1), 128–140. https://doi.org/10.4143/crt.2017.598

Desalegn, Z., Yohannes, M., Porsch, M., Stückrath, K., Anberber, E., Santos, P., Bauer, M., Addissie, A., Bekuretsion, Y., Assefa, M., Worku, Y., Taylor, L., Abebe, T., Kantelhardt, E. J., & Vetter, M. (2022). Intrinsic subtypes in Ethiopian breast cancer patient. Breast Cancer Research and Treatment, 196(3), 495–504. https://doi.org/10.1007/s10549-022-06769-z

Di Leo, A., Turner, N. H., Malorni, L., & Guarducci, C. (2012). Introduction: Luminal A and B: How Curable are they? Annals of Oncology, 23(Supplement 9), ix27. https://doi.org/10.1093/annonc/mds371

Di Modica, M., Arlotta, V., Sfondrini, L., Tagliabue, E., & Triulzi, T. (2022). The Link Between the Microbiota and HER2+ Breast Cancer: The New Challenge of Precision Medicine. Frontiers in Oncology, 12, 1–2. https://doi.org/10.3389/fonc.2022.947188

Dix-Peek, T., Phakathi, B. P., van den Berg, E. J., Dickens, C., Augustine, T. N., Cubasch, H.,

Neugut, A. I., Jacobson, J. S., Joffe, M., Ruff, P., & Duarte, R. A. B. (2023). Discordance between PAM50 intrinsic subtyping and immunohistochemistry in South African women with breast cancer. Breast Cancer Research and Treatment, 199(1), 1–12. https://doi.org/10.1007/s10549-023-06886-3

Donnet-Hughes, A., Perez, P. F., Doré, J., Leclerc, M., Levenez, F., Benyacoub, J., Serrant, P., Segura-Roggero, I., & Schiffrin, E. J. (2010). Potential role of the intestinal microbiota of the mother in neonatal immune education. Proceedings of the Nutrition Society, 69(3), 407–415. https://doi.org/10.1017/S0029665110001898

Dowsett, M., Nielsen, T. O., A'Hern, R., Bartlett, J., Coombes, R. C., Cuzick, J., Ellis, M., Henry, N. L., Hugh, J. C., Lively, T., McShane, L., Paik, S., Penault-Llorca, F., Prudkin, L., Regan, M., Salter, J., Sotiriou, C., Smith, I. E., Viale, G., … Hayes, D. F. (2011). Assessment of Ki67 in Breast Cancer: Recommendations from the international Ki67 in breast cancer working Group. Journal of the National Cancer Institute, 103(22), 1656–1664. https://doi.org/10.1093/jnci/djr393

Dowsett, M., Sestak, I., Lopez-Knowles, E., Sidhu, K., Dunbier, A. K., Cowens, J. W., Ferree, S., Storhoff, J., Schaper, C., & Cuzick, J. (2013). Comparison of PAM50 risk of recurrence score with oncotype DX and IHC4 for predicting risk of distant recurrence after endocrine therapy. Journal of Clinical Oncology, 31(22), 2783–2790. https://doi.org/10.1200/JCO.2012.46.1558

Easton, D. F., Pharoah, P. D. P., Antoniou, A. C., Tischkowitz, M., Tavtigian, S. V., Nathanson, K. L., Devilee, P., Meindl, A., Couch, F. J., Southey, M., Goldgar, D. E., Evans, D. G. R., Chenevix-Trench, G., Rahman, N., Robson, M., Domchek, S. M., & Foulkes, W. D. (2015). Gene-Panel Sequencing and the Prediction of Breast-Cancer Risk. New England Journal of Medicine, 372(23), 2243–2257. https://doi.org/10.1056/nejmsr1501341

Eber-Schulz, P., Tariku, W., Reibold, C., Addissie, A., Wickenhauser, C., Fathke, C., Hauptmann, S., Jemal, A., Thomssen, C., & Kantelhardt, E. J. (2018). Survival of breast cancer patients in rural Ethiopia. Breast Cancer Research and Treatment, 170(1), 111–118. https://doi.org/10.1007/s10549-018-4724-z

Effi, A. B., Aman, N. A., Koui, B. S., Koffi, K. D., Traoré, Z. C., & Kouyate, M. (2017). Immunohistochemical determination of estrogen and progesterone receptors in breast cancer: Relationship with clinicopathologic factors in 302 patients in Ivory Coast. BMC Cancer, 17(1), 1–6. https://doi.org/10.1186/s12885-017-3105-z

Ehrlich, S. D. (2016). Le microbiote intestinal humain influe sur la santé et la maladie. Comptes Rendus - Biologies, 339(7–8), 319–323. https://doi.org/10.1016/j.crvi.2016.04.008

El-Sayed, A., Aleya, L., & Kamel, M. (2021). Microbiota's role in health and diseases. Environmental Science and Pollution Research, 28(28), 36967–36983. https://doi.org/10.1007/s11356-021-14593-z

Elidrissi Errahhali, M., Elidrissi Errahhali, M., Ouarzane, M., El Harroudi, T., Afqir, S., & Bellaoui, M. (2017). First report on molecular breast cancer subtypes and their clinico-pathological characteristics in Eastern Morocco: Series of 2260 cases. BMC Women's Health, 17(1), 1–11. https://doi.org/10.1186/s12905-016-0361-z

Ellis, M. J., Lin, L., Crowder, R., Tao, Y., Hoog, J., Snider, J., Davies, S., Deschryver, K., Evans, D. B., Steinseifer, J., Bandaru, R., Liu, W., Gardner, H., Semiglazov, V., Watson, M., Hunt, K., Olson, J., & Baselga, J. (2010). Phosphatidyl-inositol-3-kinase alpha catalytic subunit mutation and response to neoadjuvant endocrine therapy for estrogen receptor positive breast cancer. Breast Cancer Research and Treatment, 119(2), 379–390.

https://doi.org/10.1007/s10549-009-0575-y

Elsheikh, T. M., Asa, S. L., Chan, J. K. C., DeLellis, R. A., Heffess, C. S., LiVolsi, V. A., & Wenig, B. M. (2008). Interobserver and intraobserver variation among experts in the diagnosis of thyroid follicular lesions with borderline nuclear features of papillary carcinoma. American Journal of Clinical Pathology, 130(5), 736–744. https://doi.org/10.1309/AJCPKP2QUVN4RCCP

Elwy, F., Helwa, R., Leithy, A. A. El, Shehab, Z., Assem, M. M., & Hassan, N. H. A. (2017). PIK3CA mutations in HER2-positive Breast Cancer Patients ; Frequency and Clinicopathological Perspective in Egyptian Patients. 18, 57–64. https://doi.org/10.22034/APJCP.2017.18.1.57

Eng, A., McCormack, V., & dos-Santos-Silva, I. (2014). Receptor-Defined Subtypes of Breast Cancer in Indigenous Populations in Africa: A Systematic Review and Meta-Analysis. PLoS Medicine, 11(9). https://doi.org/10.1371/journal.pmed.1001720

Ersumo, T. (2006). Breast Cancer in an Ethiopian Population, Addis Ababa. East and Central African Journal of Surgery, 11(1), 81–86.

Esfandiary, A., Sc, M., Taherian-esfahani, Z., Sc, M., Abedin-do, A., & Sc, M. (2016). Lactobacilli Modulate Hypoxia-Inducible Factor (HIF) -1 Regulatory Pathway in Triple Negative Breast Cancer Cell Line. 18(2), 237–244.

Esteva, F. J., Guo, H., Zhang, S., Santa-Maria, C., Stone, S., Lanchbury, J. S., Sahin, A. A., Hortobagyi, G. N., & Yu, D. (2010). PTEN, PIK3CA, p-AKT, and p-p70S6K status: Association with trastuzumab response and survival in patients with HER2-positive metastatic breast cancer. American Journal of Pathology, 177(4), 1647–1656. https://doi.org/10.2353/ajpath.2010.090885

Farmer, P., Frenk, J., Knaul, F. M., Shulman, L. N., Alleyne, G., Armstrong, L., Atun, R., Blayney, D., Chen, L., Feachem, R., Gospodarowicz, M., Gralow, J., Gupta, S., Langer, A., Lob-levyt, J., Neal, C., Mbewu, A., Mired, D., Piot, P., … Seffrin, J. R. (2010). Expansion of cancer care and control in countries of low and middle income : a call to action. 376(9758), 1–2. https://doi.org/10.1016/S0140-6736(10)61152-X

Feng, Y., Spezia, M., Huang, S., Yuan, C., Zeng, Z., Zhang, L., Ji, X., Liu, W., Huang, B., Luo, W., Liu, B., Lei, Y., Du, S., Vuppalapati, A., Luu, H. H., Haydon, R. C., He, T. C., & Ren, G. (2018a). Breast cancer development and progression: Risk factors, cancer stem cells, signaling pathways, genomics, and molecular pathogenesis. Genes and Diseases, 5(2), 77–106. https://doi.org/10.1016/j.gendis.2018.05.001

Feng, Y., Spezia, M., Huang, S., Yuan, C., Zeng, Z., Zhang, L., Ji, X., Liu, W., Huang, B., Luo, W., Liu, B., Lei, Y., Du, S., Vuppalapati, A., Luu, H. H., Haydon, R. C., He, T. C., & Ren, G. (2018b). Breast cancer development and progression: Risk factors, cancer stem cells, signaling pathways, genomics, and molecular pathogenesis. Genes and Diseases, 5(2), 77–106. https://doi.org/10.1016/j.gendis.2018.05.001

Ferlay, J., Shin, H. R., Bray, F., Forman, D., Mathers, C., & Parkin, D. M. (2010). Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. International Journal of Cancer, 127(12), 2893–2917. https://doi.org/10.1002/ijc.25516

Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., Parkin, D. M., Forman, D., & Bray, F. (2015). Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. International Journal of Cancer, 136(5), E359–E386. https://doi.org/10.1002/ijc.29210

Fernández, M. F., Reina-Pérez, I., Astorga, J. M., Rodríguez-Carrillo, A., Plaza-Díaz, J., & Fontana, L. (2018). Breast cancer and its relationship with the microbiota. International Journal of Environmental Research and Public Health, 15(8), 1–7. https://doi.org/10.3390/ijerph15081747

Fessler, J., Matson, V., & Gajewski, T. F. (2019). Exploring the emerging role of the microbiome in cancer immunotherapy. Journal for ImmunoTherapy of Cancer, 7(1), 1–15. https://doi.org/10.1186/s40425-019-0574-4

Fillbrunn, M., Signorovitch, J., André, F., Wang, I., Lorenzo, I., Ridolfi, A., Park, J., Dua, A., & Rugo, H. S. (2022). PIK3CA mutation status , progression and survival in advanced HR + / HER2 - breast cancer : a meta - analysis of published clinical trials. 1–11. https://doi.org/10.1186/s12885-022-10078-5

FMoH. (2015). National Cancer Control Plan 2016 - 2020 of Ethiopia. Disease Prevention and Control Directoriate, October 2015, 83. http://bmg.bund.de/en.html

Fragomeni, S. M., Sciallis, A., & Jeruss, J. S. (2018). Molecular Subtypes and Local-Regional Control of Breast Cancer. Surgical Oncology Clinics of North America, 27(1), 95–120. https://doi.org/10.1016/j.soc.2017.08.005

Fregene, A., & Newman, L. A. (2005). Breast cancer in sub-Saharan Africa : how does it relate to breast cancer in African-American women ? 2–3. https://doi.org/10.1002/cncr.20978

Fuhrman, B. J., Feigelson, H. S., Flores, R., Gail, M. H., Xu, X., Ravel, J., & Goedert, J. J. (2014). Associations of the fecal microbiome with urinary estrogens and estrogen metabolites in postmenopausal women. Journal of Clinical Endocrinology and Metabolism, 99(12), 4632–4640. https://doi.org/10.1210/jc.2014-2222

Fusco, N., Malapelle, U., Fassan, M., Marchiò, C., Buglioni, S., Zupo, S., Criscitiello, C., Vigneri, P., Dei Tos, A. P., Maiorano, E., & Viale, G. (2021). PIK3CA Mutations as a Molecular Target for Hormone Receptor-Positive, HER2-Negative Metastatic Breast Cancer. Frontiers in Oncology, 11(March), 1–9. https://doi.org/10.3389/fonc.2021.644737

Gebrehiwot, H., Hailu, T., & Gidey, G. (2014). Knowledge and Attitude towards Breast Cancer among Mekelle University Female Regular Undergraduate Students, Tigray Region, Ethiopia,. Scholars Journal of Applied Medical Sciences (SJAMS, 2(2D), 766–772. www.saspublisher.com

Gebretsadik, A., Bogale, N., & Negera, D. G. (2021). Epidemiological Trends of Breast Cancer in Southern Ethiopia : A Seven-Year Retrospective Review. 28, 1–7. https://doi.org/10.1177/10732748211055262

Geiss, G. K., Bumgarner, R. E., Birditt, B., Dahl, T., Dowidar, N., Dunaway, D. L., Fell, H. P., Ferree, S., George, R. D., Grogan, T., James, J. J., Maysuria, M., Mitton, J. D., Oliveri, P., Osborn, J. L., Peng, T., Ratcliffe, A. L., Webster, P. J., Davidson, E. H., & Hood, L. (2008). Direct multiplexed measurement of gene expression with color-coded probe pairs. Nature Biotechnology, 26(3), 317–325. https://doi.org/10.1038/nbt1385

German, R., Marino, N., Hemmerich, C., Podicheti, R., Rusch, D. B., Stiemsma, L. T., Gao, H., Xuei, X., Rockey, P., & Storniolo, A. M. (2023). Exploring breast tissue microbial composition and the association with breast cancer risk factors. 1–21.

Getachew, S., Addissie, A., Seife, E., Wakuma, T., Unverzagt, S., Jemal, A., Taylor, L., Wienke, A., & Kantelhardt, E. J. (2022). Breast Nurse Intervention to Improve Adherence to Endocrine Therapy Among Breast Cancer Patients in South Ethiopia. The Oncologist,

27(8), e650–e660. https://doi.org/10.1093/oncolo/oyac081

Getachew, S., Tesfaw, A., Kaba, M., Wienke, A., Taylor, L., Kantelhardt, E. J., & Addissie, A. (2020). Perceived barriers to early diagnosis of breast Cancer in south and southwestern Ethiopia: a qualitative study. BMC Women's Health, 20(1), 2–3. https://doi.org/10.1186/s12905-020-00909-7

Glück, S., Ross, J. S., Royce, M., McKenna, E. F., Perou, C. M., Avisar, E., & Wu, L. (2012). TP53 genomics predict higher clinical and pathologic tumor response in operable early-stage breast cancer treated with docetaxel-capecitabine ± Trastuzumab. Breast Cancer Research and Treatment, 132(3), 781–791. https://doi.org/10.1007/s10549-011-1412-7

Goedert, J. J., Jones, G., Hua, X., Xu, X., Yu, G., Flores, R., Falk, R. T., Gail, M. H., Shi, J., Ravel, J., & Feigelson, H. S. (2015). Investigation of the Association Between the Fecal Microbiota and Breast Cancer in Postmenopausal Women: A Population-Based Case-Control Pilot Study. Journal of the National Cancer Institute, 107(8), 1–5. https://doi.org/10.1093/jnci/djv147

Goldhirsch, A., Ingle, J. N., Gelber, R. D., Coates, A. S., Thürlimann, B., & Senn, H. (2009). Thresholds for therapies : highlights of the St Gallen International Expert Consensus on the primary therapy of early breast cancer 2009. 22–23. https://doi.org/10.1093/annonc/mdp322

Goldhirsch, A., Winer, E. P., Coates, A. S., Gelber, R. D., Piccart-Gebhart, M., Thürlimann, B., Senn, H. J., Albain, K. S., André, F., Bergh, J., Bonnefoi, H., Bretel-Morales, D., Burstein, H., Cardoso, F., Castiglione-Gertsch, M., Coates, A. S., Colleoni, M., Costa, A., Curigliano, G., … Wood, W. C. (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. Annals of Oncology, 24(9), 2206–2223. https://doi.org/10.1093/annonc/mdt303

Goldhirsch, A., Wood, W. C., Coates, A. S., Gelber, R. D., Thürlimann, B., & Senn, H. J. (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. Annals of Oncology, 22(8), 1736–1747. https://doi.org/10.1093/annonc/mdr304

Gonzaiez-Angulo, A. M., Stemke-Hale, K., Palla, S. L., Carey, M., Agarwal, R., Meric-Berstam, F., Traina, T., Hudis, C., Hortobagyi, G. N., Gerald, W. L., Mills, G. B., & Hennessy, B. T. (2009). Androgen receptor levels and association with PIK3CA mutations and prognosis in breast cancer. Clinical Cancer Research, 15(7), 2472–2478. https://doi.org/10.1158/1078-0432.CCR-08-1763

Gopalakrishnan, V., Spencer, C. N., Nezi, L., Reuben, A., Andrews, M. C., Karpinets, T. V., Prieto, P. A., Vicente, D., Hoffman, K., Wei, S. C., Cogdill, A. P., Zhao, L., Hudgens, C. W., Hutchinson, D. S., Manzo, T., Petaccia De Macedo, M., Cotechini, T., Kumar, T., Chen, W. S., … Wargo, J. A. (2018). Gut microbiome modulates response to anti-PD-1 immunotherapy in melanoma patients. Science, 359(6371), 97–103. https://doi.org/10.1126/science.aan4236

Group, M., Co-chairs, committee, W., committee, S., Institute, T. and clinical data source sites: U. of C. R. U. K. C. R., Agency, B. C. C., Nottingham, U. of, London, K. C., Biology, M. I. of C., Institute, C. genome/transcriptome characterization centres: U. of C. R. U. K. C. R., & Institute, D. analysis subgroup: U. of C. R. U. K. C. R. (2012). The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature, 1–7.

http://dx.doi.org/10.1038/nature10983%5Cnpapers2://publication/doi/10.1038/nature10983

Hadgu, E., Seifu, D., Tigneh, W., Bokretsion, Y., Bekele, A., Abebe, M., Sollie, T., Merajver, S. D., Karlsson, C., & Karlsson, M. G. (2018). Breast cancer in Ethiopia: Evidence for geographic difference in the distribution of molecular subtypes in Africa. BMC Women's Health, 18(1), 1–8. https://doi.org/10.1186/s12905-018-0531-2

Hammond, M. E. H., Hayes, D. F., Dowsett, M., Allred, D. C., Hagerty, K. L., Badve, S., Fitzgibbons, P. L., Francis, G., Goldstein, N. S., Hayes, M., Hicks, D. G., Lester, S., Love, R., Mangu, P. B., McShane, L., Miller, K., Osborne, C. K., Paik, S., Perlmutter, J., … Wolff, A. C. (2010). American society of clinical oncology/college of american pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. Journal of Clinical Oncology, 28(16), 2784–2795. https://doi.org/10.1200/JCO.2009.25.6529

Han, Y. W., Shen, T., Chung, P., Buhimschi, I. A., & Buhimschi, C. S. (2009). Uncultivated bacteria as etiologic agents of intra-amniotic inflammation leading to preterm birth. Journal of Clinical Microbiology, 47(1), 38–47. https://doi.org/10.1128/JCM.01206-08

Harbeck, N., Penault-Llorca, F., Cortes, J., Gnant, M., Houssami, N., Poortmans, P., Ruddy, K., Tsang, J., & Cardoso, F. (2019). Breast cancer. In Nature Reviews Disease Primers (Vol. 5, Issue 1). https://doi.org/10.1038/s41572-019-0111-2

Harris, C. C. (2016). Editorial. Carcinogenesis, 37(1), 1. https://doi.org/10.1093/carcin/bgv246

Harris, L., Fritsche, H., Mennel, R., Norton, L., Ravdin, P., Taube, S., Somerfield, M. R., Hayes, D. F., Jr, R. C. B., & Society, A. (2007). American Society of Clinical Oncology 2007 update of recommendations for the use of tumor markers in breast cancer. 26(12), 10–11. https://doi.org/10.1200/JCO.2007.14.2364

Harris, L. N., Ismaila, N., McShane, L. M., Andre, F., Collyar, D. E., Gonzalez-Angulo, A. M., Hammond, E. H., Kuderer, N. M., Liu, M. C., Mennel, R. G., Van Poznak, C., Bast, R. C., & Hayes, D. F. (2016). Use of biomarkers to guide decisions on adjuvant systemic therapy forwomenwith early-stage invasive breast cancer: American Society of clinical Oncology clinical practice guideline. Journal of Clinical Oncology, 34(10), 1134–1150. https://doi.org/10.1200/JCO.2015.65.2289

Hartung, C., Porsch, M., Stückrath, K., Kaufhold, S., Staege, M. S., Hanf, V., Lantzsch, T., Uleer, C., Peschel, S., John, J., Pöhler, M., Weigert, E., Buchmann, J., Bürrig, K. F., Schüler, K., Bethmann, D., Große, I., Kantelhardt, E. J., Thomssen, C., & Vetter, M. (2021). Identifying High-Risk Triple-Negative Breast Cancer Patients by Molecular Subtyping. Breast Care, 16(6), 637–647. https://doi.org/10.1159/000519255

Hassan, Z., Mustafa, S., Rahim, R. A., & Isa, N. M. (2016). Anti-breast cancer effects of live , heat-killed and cytoplasmic fractions of Enterococcus faecalis and Staphylococcus hominis isolated from human breast milk. https://doi.org/10.1007/s11626-015-9978-8

Hedenfalk, I., Duggan, D., Chen, Y., Radmacher, M., Bittner, M., Simon, R., Meltzer, P., & Gusterson, B. (2001). Gene-expression profiles in hereditary breast cancer. 344(8), 22–23. https://doi.org/10.1056/NEJM200102223440801

Henry, N. L., & Hayes, D. F. (2012). Cancer biomarkers 5. 6, 0–6. https://doi.org/10.1016/j.molonc.2012.01.010

Hernandez-Aya, L. F., & Gonzalez-Angulo, A. M. (2011). Targeting the Phosphatidylinositol 3-Kinase Signaling Pathway in Breast Cancer. The Oncologist, 16(4), 404–414.

https://doi.org/10.1634/theoncologist.2010-0402

Hieken, T. J., Chen, J., Hoskin, T. L., Walther-Antonio, M., Johnson, S., Ramaker, S., Xiao, J., Radisky, D. C., Knutson, K. L., Kalari, K. R., Yao, J. Z., Baddour, L. M., Chia, N., & Degnim, A. C. (2016). The microbiome of aseptically collected human breast tissue in benign and malignant disease. Scientific Reports, 6(July), 1–10. https://doi.org/10.1038/srep30751

Hoskinson, C., Zheng, K., Gabel, J., Kump, A., German, R., Podicheti, R., Marino, N., & Stiemsma, L. T. (2022). Composition and Functional Potential of the Human Mammary Microbiota Prior to and Following Breast Tumor Diagnosis. MSystems, 7(3), 1–18. https://doi.org/10.1128/msystems.01489-21

Hu, Z., Fan, C., Oh, D. S., Marron, J. S., He, X., Qaqish, B. F., Livasy, C., Carey, L. A., Reynolds, E., Dressler, L., Nobel, A., Parker, J., Ewend, M. G., Sawyer, L. R., Wu, J., Liu, Y., Nanda, R., Tretiakova, M., Orrico, A. R., … Perou, C. M. (2006). The molecular portraits of breast tumors are conserved across microarray platforms. BMC Genomics, 7, 1–12. https://doi.org/10.1186/1471-2164-7-96

Huebner, R. J., & Ewald, A. J. (2014). Cellular foundations of mammary tubulogenesis. Seminars in Cell and Developmental Biology, 31, 124–131. https://doi.org/10.1016/j.semcdb.2014.04.019

Huttenhower, C., Gevers, D., Knight, R., Abubucker, S., Badger, J. H., Chinwalla, A. T., Creasy, H. H., Earl, A. M., Fitzgerald, M. G., Fulton, R. S., Giglio, M. G., Hallsworth-Pepin, K., Lobos, E. A., Madupu, R., Magrini, V., Martin, J. C., Mitreva, M., Muzny, D. M., Sodergren, E. J., … White, O. (2012). Structure, function and diversity of the healthy human microbiome. Nature, 486(7402), 207–214. https://doi.org/10.1038/nature11234

Ibrahim, I. H., Abd El-Aziz, H. G., Amer, N. N. L., & Abd El-Sameea, H. S. (2022). Mutational pattern of PIK3CA exon 20 in circulating DNA in breast cancer. Saudi Journal of Biological Sciences, 29(4), 2828–2835. https://doi.org/10.1016/j.sjbs.2022.01.002

Isakoff, S. J., Engelman, J. A., Irie, H. Y., Luo, J., Brachmann, S. M., Pearline, R. V., Cantley, L. C., & Brugge, J. S. (2005). Breast cancer-associated PIK3CA mutations are oncogenic in mammary epithelial cells. Cancer Research, 65(23), 10992–11000. https://doi.org/10.1158/0008-5472.CAN-05-2612

Jemal, A., Bray, F., Center, M. M., Ferlay, J., Ward, E., & Forman, D. (2011). Global cancer statistics. CA: A Cancer Journal for Clinicians, 61(2), 69–90. https://doi.org/10.3322/caac.20107

Jenkins, E. O., Deal, A. M., Anders, C. K., Prat, A., Perou, C. M., Carey, L. A., & Muss, H. B. (2014). Age-specific changes in intrinsic breast cancer subtypes : a focus on older women. 2–3. https://doi.org/10.1634/theoncologist.2014-0184

Jensen, M. B., Lænkholm, A. V., Nielsen, T. O., Eriksen, J. O., Wehn, P., Hood, T., Ram, N., Buckingham, W., Ferree, S., & Ejlertsen, B. (2018). The Prosigna gene expression assay and responsiveness to adjuvant cyclophosphamide-based chemotherapy in premenopausal high-risk patients with breast cancer. Breast Cancer Research, 20(1), 1–10. https://doi.org/10.1186/s13058-018-1012-0

Jia, G., Ping, J., Guo, X., Yang, Y., Tao, R., Li, B., Ambs, S., Mollie, E., Chen, Y., Gu, J., Hu, J. J., Huo, D., John, E. M., Li, C. I., Li, J. L., Nathanson, K. L., Nemesure, B., Olopade, O. I., Pal, T., … Shu, X. (2024). Genome-wide association analyses of breast cancer in women of African ancestry identify new susceptibility loci and improve risk prediction.

Nature Genetics, 56(5), 819–826. https://doi.org/10.1038/s41588-024-01736-4

Jorgensen, C. L. T., Nielsen, T. O., Bjerre, K. D., Liu, S., Wallden, B., Balslev, E., Nielsen, D. L., & Ejlertsen, B. (2014). PAM50 breast cancer intrinsic subtypes and effect of gemcitabine in advanced breast cancer patients. Acta Oncologica, 53(6), 776–787. https://doi.org/10.3109/0284186X.2013.865076

Jorns, J. M. (2019). Challenges in routine estrogen receptor, progesterone receptor, and HER2/neu evaluation. Archives of Pathology and Laboratory Medicine, 143(12), 1444–1449. https://doi.org/10.5858/arpa.2019-0205-RA

Juncker-Jensen, A., Lykkesfeldt, A. E., Worm, J., Ralfkiær, U., Espelund, U., & Jepsen, J. S. (2006). Insulin-like growth factor binding protein 2 is a marker for antiestrogen resistant human breast cancer cell lines but is not a major growth regulator. Growth Hormone and IGF Research, 16(4), 224–239. https://doi.org/10.1016/j.ghir.2006.06.005

Junttila, T. T., Akita, R. W., Parsons, K., Fields, C., Lewis Phillips, G. D., Friedman, L. S., Sampath, D., & Sliwkowski, M. X. (2009). Ligand-Independent HER2/HER3/PI3K Complex Is Disrupted by Trastuzumab and Is Effectively Inhibited by the PI3K Inhibitor GDC-0941. Cancer Cell, 15(5), 429–440. https://doi.org/10.1016/j.ccr.2009.03.020

Kadirareddy, R. H., Ghantavemuri, S., & Devi, U. M. (2016). Probiotic Conjugated Linoleic Acid Mediated Apoptosis in Breast Cancer Cells by Downregulation of NF-κB. 17, 3395–3403.

Kalinsky, K., Jacks, L. M., Heguy, A., Patil, S., Drobnjak, M., Bhanot, U. K., Hedvat, C. V., Traina, T. A., Solit, D., Gerald, W., & Moynahan, M. E. (2009). PIK3CA mutation associates with improved outcome in breast cancer. Clinical Cancer Research, 15(16), 5049–5059. https://doi.org/10.1158/1078-0432.CCR-09-0632

Kallel, I., Khabir, A., Boujelbene, N., Abdennadher, R., & Daoud, J. (2012). EGFR overexpression relates to triple negative profile and poor prognosis in breast cancer patients in Tunisia. 2–3. https://doi.org/10.3109/10799893.2012.664552

Kantelhardt., Zerche, P., Mathewos, A., Trocchi, P., Addissie, A., Aynalem, A., Wondemagegnehu, T., Ersumo, T., Reeler, A., Yonas, B., Tinsae, M., Gemechu, T., Jemal, A., Thomssen, C., Stang, A., & Bogale, S. (2014). Breast cancer survival in Ethiopia: A cohort study of 1,070 women. International Journal of Cancer, 135(3), 702–709. https://doi.org/10.1002/ijc.28691

Kantelhardt, J. J., Mathewos, A., Aynalem, A., Wondemagegnehu, T., Jemal, A., Vetter, M., Knauf, E., Reeler, A., Bogale, S., Thomssen, C., Stang, A., Gemechu, T., Trocchi, P., & Yonas, B. (2014). The prevalence of estrogen receptor-negative breast cancer in Ethiopia. BMC Cancer, 14(1), 1–6. https://doi.org/10.1186/1471-2407-14-895

Karakas, B., Bachman, K. E., & Park, B. H. (2006). Mutation of the PIK3CA oncogene in human cancers. British Journal of Cancer, 94(4), 455–459. https://doi.org/10.1038/sj.bjc.6602970

Kataoka, Y., Mukohara, T., Shimada, H., Saijo, N., Hirai, M., & Minami, H. (2010). Association between gain-of-function mutations in PIK3CA and resistance to HER2-targeted agents in HER2-amplified breast cancer cell lines. Annals of Oncology, 21(2), 255–262. https://doi.org/10.1093/annonc/mdp304

Keku, T. O. (2015). The gastrointestinal microbiota and colorectal cancer. 2014(45), 1–28. https://doi.org/10.1152/ajpgi.00360.2012

Kensler, K. H., Sankar, V. N., Wang, J., Zhang, X., Christopher, A., Baker, G. M., Parker, J. S., Hoadley, K. A., Stancu, A. L., Pyle, M. E., Collins, L. C., Hunter, D. J., Eliassen, A. H., Susan, E., Tamimi, R. M., & Heng, Y. J. (2020). HHS Public Access. 1–19. https://doi.org/10.1158/1055-9965.EPI-18-0863.PAM50

Kensler, K. H., Sankar, V. N., Wang, J., Zhang, X., Rubadue, C. A., Baker, G. M., Parker, J. S., Hoadley, K. A., Stancu, A. L., Pyle, M. E., Collins, L. C., Hunter, D. J., Eliassen, A. H., Hankinson, S. E., Tamimi, R. M., & Heng, Y. J. (2019). PAM50 molecular intrinsic subtypes in the nurses' health Study cohorts. Cancer Epidemiology Biomarkers and Prevention, 28(4), 798–806. https://doi.org/10.1158/1055-9965.EPI-18-0863

Khabaz, M. N. (2014). Immunohistochemistry subtypes (ER / PR / HER) of breast cancer : where do we stand in the West of Saudi Arabia ? 2–3. https://doi.org/10.7314/apjcp.2014.15.19.8395

Khatib, O. (2006). Guidelines for the Early Detection and Screening of Breast Cancer. EMRO Technical Publications Series, 30. In WHO World Health Organization.

Kim, C., Baker, J., Ph, D., Cronin, M., Ph, D., Baehner, F. L., Walker, M. G., Ph, D., Watson, D., Ph, D., Park, T., Ph, D., Hiller, W., Fisher, E. R., Wickerham, D. L., Bryant, J., Ph, D., & Wolmark, N. (2004). A Multigene Assay to Predict Recurrence of Tamoxifen-Treated, Node-Negative Breast Cancer. 2817–2826.

Kim, H. K., Park, K. H., Kim, Y., Park, S. E., Lee, H. S., Lim, S. W., Cho, J. H., Kim, J. Y., Lee, J. E., Ahn, J. S., Im, Y. H., Yu, J. H., & Park, Y. H. (2019). Discordance of the PAM50 intrinsic subtypes compared with immunohistochemistry-based surrogate in breast cancer patients: Potential implication of genomic alterations of discordance. Cancer Research and Treatment, 51(2), 737–747. https://doi.org/10.4143/crt.2018.342

Klann, E., Williamson, J. M., Tagliamonte, M. S., & Ukhanova, M. (2020). Microbiota composition in bilateral healthy breast tissue and breast tumors. Cancer Causes & Control, 31(11), 1027–1038. https://doi.org/10.1007/s10552-020-01338-5

Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., & Schloss, P. D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the miseq illumina sequencing platform. Applied and Environmental Microbiology, 79(17), 5112–5120. https://doi.org/10.1128/AEM.01043-13

Kulkarni, M. M. (2011). Digital multiplexed gene expression analysis using the nanostring ncounter system. Current Protocols in Molecular Biology, 94(SUPPL.94), 25–27. https://doi.org/10.1002/0471142727.mb25b10s94

Kumar, N., Patni, P., Agarwal, A., Khan, M. A., & Parashar, N. (2015). Prevalence of molecular subtypes of invasive breast cancer: A retrospective study. Medical Journal Armed Forces India, 71(3), 254–258. https://doi.org/10.1016/j.mjafi.2015.04.006

Kurosumi, M. (2007). R e v i e w Article Immunohistochemical Assessment of Hormone Receptor Status Using a New Scoring System (J-Score) in Breast Cancer. 14(2), 189–193.

Kuzhan, A., Adli, M., Eryigit Alkis, H., & Caglayan, D. (2013). Hormone receptor and HER2 status in patients with breast cancer by races in southeastern Turkey. Journal of B.U.ON., 18(3), 619–622.

Kwan, M. L., Kroenke, C. H., Sweeney, C., Bernard, P. S., Weltzien, E. K., Castillo, A., Factor, R. E., Maxfield, K. S., Stijleman, I. J., Kushi, L. H., Quesenberry, C. P., Habel, L. A., &

Caan, B. J. (2015). Association of high obesity with PAM50 breast cancer intrinsic subtypes and gene expression. BMC Cancer, 15(1), 1–10. https://doi.org/10.1186/s12885-015-1263-4

Lakritz, J. R., Poutahidis, T., Mirabal, S., Varian, B. J., Levkovich, T., Ibrahim, Y. M., Ward, J. M., Teng, E. C., Fisher, B., Parry, N., Lesage, S., Alberg, N., Gourishetti, S., Fox, J. G., Ge, Z., & Erdman, S. E. (2015). Gut bacteria require neutrophils to promote mammary tumorigenesis. Oncotarget, 6(11), 9387–9396. https://doi.org/10.18632/oncotarget.3328

Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. Biometrics, 33(1), 159. https://doi.org/10.2307/2529310

Layfield, L. J., Gupta, D., & Mooney, E. E. (2000). Assessment of Tissue Estrogen and Progesterone Receptor Levels : A Survey of Current Practice , Techniques , and Quantitation Methods. 2–3. https://doi.org/10.1046/j.1524-4741.2000.99097.x

Lee, I. O., Kim, J. H., Choi, Y. J., Pillinger, M. H., Kim, S., Blaser, M. J., & Lee, Y. C. (2010). Helicobacter pylori CagA Phosphorylation Status Determines the gp130-activated SHP2 / ERK and JAK / STAT Signal Transduction Pathways in Gastric Epithelial Cells * □. 285(21), 16042–16050. https://doi.org/10.1074/jbc.M110.111054

Lee, J., Sohn, I., Do, I. G., Kim, K. M., Park, S. H., Park, J. O., Park, Y. S., Lim, H. Y., Sohn, T. S., Bae, J. M., Choi, M. G., Lim, D. H., Min, B. H., Lee, J. H., Rhee, P. L., Kim, J. J., Choi, D. Il, Tan, I. B., Das, K., … Kim, S. (2014). Nanostring-based multigene assay to predict recurrence for gastric cancer patients after surgery. PLoS ONE, 9(3), 1–6. https://doi.org/10.1371/journal.pone.0090133

Lee, J. W., Soung, Y. H., Kim, S. Y., Lee, H. W., Park, W. S., Nam, S. W., Kim, S. H., Lee, J. Y., Yoo, N. J., & Lee, S. H. (2005). PIK3CA gene is frequently mutated in breast carcinomas and hepatocellular carcinomas. Oncogene, 24(8), 1477–1480. https://doi.org/10.1038/sj.onc.1208304

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. Journal of Clinical Investigation, 121(7), 2750–2767. https://doi.org/10.1172/JCI45014

Lehouritis, P., Cummins, J., Stanton, M., Murphy, C. T., McCarthy, F. O., Reid, G., Urbaniak, C., Byrne, W. L., & Tangney, M. (2015). Local bacteria affect the efficacy of chemotherapeutic drugs. Scientific Reports, 5, 1–2. https://doi.org/10.1038/srep14554

Lei, J. T., Anurag, M., Haricharan, S., Gou, X., & Ellis, M. J. (2020). Endocrine therapy resistance : new insights. 48(Suppl 1), 15–20. https://doi.org/10.1016/S0960-9776(19)31118-X

Leila Dorling, Sara Carvalho, Jamie Allen, Anna González-Neira, Craig Luccarini, Cecilia Wahlström, K. A. P. (2021). Breast Cancer Risk Genes — Association Analysis in More than 113,000 Women. N Engl J Med, 384(5), 428–439. https://doi.org/10.1056/NEJMoa1913948

Li, G., Guo, X., Chen, M., Tang, L., Jiang, H., Day, J. X., Xie, Y., Peng, L., Xu, X., Li, J., Wang, S., Xiao, Z., Dai, L., & Wang, J. (2018). Prevalence and spectrum of AKT1, PIK3CA, PTEN and TP53 somatic mutations in Chinese breast cancer patients. PLoS ONE, 13(9), 1–18. https://doi.org/10.1371/journal.pone.0203495

Li, S. Y., Rong, M., Grieu, F., & Iacopetta, B. (2006). PIK3CA mutations in breast cancer are associated with poor outcome. 2, 3–9.

Lian, J., Xu, E. W., Xi, Y. F., Wang, H. W., Bu, P., Wang, J. F., & Wang, L. X. (2020). Clinical-Pathologic Analysis of Breast Cancer With PIK3CA Mutations in Chinese Women. Technology in Cancer Research and Treatment, 19, 1–7. https://doi.org/10.1177/1533033820950832

Ligresti, G., Militello, L., Steelman, L. S., Cavallaro, A., Basile, F., Nicoletti, F., Stivala, F., Mccubrey, J. A., & Libra, M. (2009). PIK3CA mutations in human solid tumors Role in sensitivity to various therapeutic approaches. 8(9), 1352–1358.

Lingwood, R. J., Boyle, P., Milburn, A., Ngoma, T., Arbuthnott, J., McCaffrey, R., Kerr, S. H., & Kerr, D. J. (2008). The challenge of cancer control in Africa. Nature Reviews Cancer, 8(5), 398–403. https://doi.org/10.1038/nrc2372

Liu, C. Y., Wu, C. Y., Petrossian, K., Huang, T. T., Tseng, L. M., & Chen, S. (2017). Treatment for the endocrine resistant breast cancer: Current options and future perspectives. Journal of Steroid Biochemistry and Molecular Biology, 172, 166–175. https://doi.org/10.1016/j.jsbmb.2017.07.001

Liu, P., Cheng, H., Roberts, T. M., & Zhao, J. J. (2009). Targeting the phosphoinositide 3-kinase pathway in cancer. Nature Reviews Drug Discovery, 8(8), 627–644. https://doi.org/10.1038/nrd2926

Liu, S., Chapman, J. A. W., Burnell, M. J., Levine, M. N., Pritchard, K. I., Whelan, T. J., Rugo, H. S., Albain, K. S., Perez, E. A., Virk, S., Barry, G., Gao, D., O'Brien, P., Shepherd, L. E., Nielsen, T. O., & Gelmon, K. A. (2015). Prognostic and predictive investigation of PAM50 intrinsic subtypes in the NCIC CTG MA.21 phase III chemotherapy trial. Breast Cancer Research and Treatment, 149(2), 439–448. https://doi.org/10.1007/s10549-014-3259-1

Loi, S., Michiels, S., Lambrechts, D., Fumagalli, D., Claes, B., Kellokumpu-Lehtinen, P. L., Bono, P., Kataja, V., Piccart, M. J., Joensuu, H., & Sotiriou, C. (2013). Somatic mutation profiling and associations with prognosis and trastuzumab benefit in early breast cancer. Journal of the National Cancer Institute, 105(13), 960–967. https://doi.org/10.1093/jnci/djt121

López-Knowles, E., Segal, C. V., Gao, Q., Garcia-Murillas, I., Turner, N. C., Smith, I., Martin, L. A., & Dowsett, M. (2014). Relationship of PIK3CA mutation and pathway activity with antiproliferative response to aromatase inhibition. Breast Cancer Research, 16(3). https://doi.org/10.1186/bcr3683

Louise Pouncey, A., James Scott, A., Leslie Alexander, J., Marchesi, J., & Kinross, J. (2018). Gut microbiota, chemotherapy and the host: The influence of the gut microbiota on cancer treatment. Ecancermedicalscience, 12, 1–9. https://doi.org/10.3332/ecancer.2018.868

Luangxay, T., Virachith, S., Hando, K., Vilayvong, S., Xaysomphet, P., Arounlangsy, P., Phongsavan, K., Mieno, M. N., Honma, N., Kitagawa, M., & Sawabe, M. (2019). Subtypes of breast cancer in Lao P.D.R.: A study in a limited-resource setting. Asian Pacific Journal of Cancer Prevention, 20(2), 589–594. https://doi.org/10.31557/APJCP.2019.20.2.589

Lukong, K. E. (2017). Understanding breast cancer – The long and winding road. BBA Clinical, 7, 64–77. https://doi.org/10.1016/j.bbacli.2017.01.001

Lukong, K. E., Ogunbolude, Y., & Kamdem, J. P. (2017). Breast cancer in Africa: prevalence, treatment options, herbal medicines, and socioeconomic determinants. Breast Cancer Research and Treatment, 166(2), 351–365. https://doi.org/10.1007/s10549-017-4408-0

Luu, T. H., Michel, C., Bard, J. M., Dravet, F., Nazih, H., & Bobin-Dubigeon, C. (2017). Intestinal Proportion of Blautia sp. is Associated with Clinical Stage and Histoprognostic Grade in Patients with Early-Stage Breast Cancer. Nutrition and Cancer, 69(2), 267–275. https://doi.org/10.1080/01635581.2017.1263750

Lux, M. P., Fasching, P. A., Schrauder, M. G., Hein, A., Jud, S. M., Rauh, C., & Beckmann, M. W. (2016). The PI3K Pathway: Background and Treatment Approaches. Breast Care, 11(6), 398–404. https://doi.org/10.1159/000453133

Ly, M., Antoine, M., Dembélé, A. K., Levy, P., Rodenas, A., Touré, B. A., Badiaga, Y., Dembélé, B. K., & Bagayogo, D. C. (2012). 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. 2–3. https://doi.org/10.1159/000341541

Macias, H., & Hinck, L. (2012). Mammary gland development. 2–3. https://doi.org/10.1002/wdev.35

Maisonneuve, P., Disalvatore, D., Rotmensz, N., Curigliano, G., Colleoni, M., Dellapasqua, S., Pruneri, G., Mastropasqua, M. G., Luini, A., Bassi, F., Pagani, G., Viale, G., & Goldhirsch, A. (2014). Proposed new clinicopathological surrogate definitions of luminal A and luminal B (HER2-negative) intrinsic breast cancer subtypes. 16(3), 1–9. https://doi.org/10.1186/bcr3679

Makki, J. (2015). Diversity of breast carcinoma: Histological subtypes and clinical relevance. Clinical Medicine Insights: Pathology, 8(1), 23–31. https://doi.org/10.4137/CPath.s31563

Malvia, S., Bagadi, S. A. R., Pradhan, D., Chintamani, C., Bhatnagar, A., Arora, D., Sarin, R., & Saxena, S. (2019). Study of Gene Expression Profiles of Breast Cancers in Indian Women. Scientific Reports, 9(1), 1–15. https://doi.org/10.1038/s41598-019-46261-1

Mangone, F. R., Bobrovnitchaia, I. G., Salaorni, S., Manuli, E., & Nagai, M. A. (2012). PIK3CA exon 20 mutations are associated with poor prognosis in breast cancer patients. Clinics, 67(11), 1285–1290. https://doi.org/10.6061/clinics/2012(11)11

Martínez-Saéz, O., Chic, N., Pascual, T., Adamo, B., Vidal, M., González-Farré, B., Sanfeliu, E., Schettini, F., Conte, B., Brasó-Maristany, F., Rodríguez, A., Martínez, D., Galván, P., Rodríguez, A. B., Martinez, A., Munõz, M., & Prat, A. (2020). Frequency and spectrum of PIK3CA somatic mutations in breast cancer. Breast Cancer Research, 22(1), 1–9. https://doi.org/10.1186/s13058-020-01284-9

Maruyama, N., Miyoshi, Y., Taguchi, T., Tamaki, Y., Monden, M., & Noguchi, S. (2007). Clinicopathologic analysis of breast cancers with PIK3CA mutations in Japanese women. Clinical Cancer Research, 13(2 I), 408–414. https://doi.org/10.1158/1078-0432.CCR-06-0267

Masood, S. (2016). Breast cancer subtypes: Morphologic and biologic characterization. Women's Health, 12(1), 103–119. https://doi.org/10.2217/whe.15.99

Massarweh, S., Osborne, C. K., Creighton, C. J., Qin, L., Tsimelzon, A., Huang, S., Weiss, H., Rimawi, M., & Schiff, R. (2008). Tamoxifen resistance in breast tumors is driven by growth factor receptor signaling with repression of classic estrogen receptor genomic function. Cancer Research, 68(3), 826–833. https://doi.org/10.1158/0008-5472.CAN-07-2707

McCormack, V. A., Febvey-Combes, O., Ginsburg, O., & dos-Santos-Silva, I. (2018). Breast cancer in women living with HIV: A first global estimate. International Journal of Cancer,

143(11), 2732–2740. https://doi.org/10.1002/ijc.31722

Mcdonald, E. S., Clark, A. S., Tchou, J., Zhang, P., & Freedman, G. M. (2016). Clinical Diagnosis and Management of Breast Cancer. 9–16. https://doi.org/10.2967/jnumed.115.157834

Memirie, S. T., Habtemariam, M. K., Asefa, M., Deressa, B. T., Abayneh, G., Tsegaye, B., Abraha, M. W., Ababi, G., Jemal, A., Rebbeck, T. R., & Verguet, S. (2018). Estimates of cancer incidence in Ethiopia in 2015 using population-based registry data. Journal of Global Oncology, 2018(4), 1–11. https://doi.org/10.1200/JGO.17.00175

Meng, S., Chen, B., Yang, J., Wang, J., Zhu, D., Meng, Q., & Zhang, L. (2018). Study of microbiomes in aseptically collected samples of human breast tissue using needle biopsy and the potential role of in situ tissue microbiomes for promoting malignancy. Frontiers in Oncology, 8(AUG), 1–14. https://doi.org/10.3389/fonc.2018.00318

Michelucci, A., Di Cristofano, C., Lami, A., Collecchi, P., Caligo, A., Decarli, N., Leopizzi, M., Aretini, P., Bertacca, G., Porta, R. P., Ricci, S., Rocca, C. Della, Stanta, G., Bevilacqua, G., & Cavazzana, A. (2009). PIK3CA in breast carcinoma: A mutational analysis of sporadic and hereditary cases. Diagnostic Molecular Pathology, 18(4), 200–205. https://doi.org/10.1097/PDM.0b013e31818e5fa4

Miled, N., Yan, Y., Hon, W. C., Perisic, O., Zvelebil, M., Inbar, Y., Schneidman-Duhovny, D., Wolfson, H. J., Backer, J. M., & Williams, R. L. (2007). Mechanism of two classes of cancer mutations in the phosphoinositide 3-kinase catalytic subunit. Science, 317(5835), 239–242. https://doi.org/10.1126/science.1135394

Miller, T. W., Hennessy, B. T., González-Angulo, A. M., Fox, E. M., Mills, G. B., Chen, H., Higham, C., García-Echeverría, C., Shyr, Y., & Arteaga, C. L. (2010). Hyperactivation of phosphatidylinositol-3 kinase promotes escape from hormone dependence in estrogen receptor-positive human breast cancer. Journal of Clinical Investigation, 120(7), 2406–2413. https://doi.org/10.1172/JCI41680

Millikan, R. C., Newman, B., Tse, C., Moorman, P. G., Smith, L. V, Labbok, M. H., Geradts, J., Jeannette, T., Jackson, S., Nyante, S., Livasy, C., Carey, L., Shelton, H., & Perou, C. M. (2008). Epidemiology of basal-like breast cancer. 109(1), 123–139.

Miron, A., Varadi, M., Carrasco, D., Li, H., Luongo, L., Kim, H. J., Park, S. Y., Cho, E. Y., Lewis, G., Kehoe, S., Iglehart, J. D., Dillon, D., Allred, D. C., Macconaill, L., Gelman, R., & Polyak, K. (2010). PIK3CA mutations in in situ and invasive breast carcinomas. Cancer Research, 70(14), 5674–5678. https://doi.org/10.1158/0008-5472.CAN-08-2660

Miyagawa, Y., Miyake, T., Yanai, A., Murase, K., Imamura, M., Ichii, S., Takatsuka, Y., Ito, T., Hirota, S., Saito, M., Kotoura, Y., Miyauchi, K., Fujimoto, Y., Hatada, T., Sasa, M., & Miyoshi, Y. (2015). Association of body mass index with risk of luminal A but not luminal B estrogen receptor-positive and HER2-negative breast cancer for postmenopausal Japanese women. Breast Cancer, 22(4), 399–405. https://doi.org/10.1007/s12282-013-0493-z

Mosele, F., Stefanovska, B., Lusque, A., Tran Dien, A., Garberis, I., Droin, N., Le Tourneau, C., Sablin, M. P., Lacroix, L., Enrico, D., Miran, I., Jovelet, C., Bièche, I., Soria, J. C., Bertucci, F., Bonnefoi, H., Campone, M., Dalenc, F., Bachelot, T., … André, F. (2020). Outcome and molecular landscape of patients with PIK3CA-mutated metastatic breast cancer. Annals of Oncology, 31(3), 377–386. https://doi.org/10.1016/j.annonc.2019.11.006

Mukohara, T. (2015). PI3K mutations in breast cancer: prognostic and therapeutic implications. Breast Cancer: Targets and Therapy, 2015:7, 111–123. https://doi.org/https://doi.org/10.2147/BCTT.S60696

Müller, B. M., Kronenwett, R., Hennig, G., Euting, H., Weber, K., Bohmann, K., Weichert, W., Altmann, G., Roth, C., Winzer, K. J., Kristiansen, G., Petry, C., Dietel, M., & Denkert, C. (2011). Quantitative determination of estrogen receptor, progesterone receptor, and HER2 mRNA in formalin-fixed paraffin-embedded tissue - A new option for predictive biomarker assessment in breast cancer. Diagnostic Molecular Pathology, 20(1), 1–10. https://doi.org/10.1097/PDM.0b013e3181e3630c

Murphy, E. A., Velazquez, K. T., & Herbert, K. M. (2015). Influence of high-fat diet on gut microbiota : a driving force for chronic disease risk. 18(5), 2015. https://doi.org/10.1097/MCO.0000000000000209

Nanostring Technologies. (2017). Gene Expression Data Analysis Guidelines. 1–23. www.nanostring.com

Nascimento, R. G. do, & Otoni, K. M. (2020). Histological and molecular classification of breast cancer: what do we know? Mastology, 30, 1–8. https://doi.org/10.29289/25945394202020200024

Nejman, D., Livyatan, I., Fuks, G., Gavert, N., Zwang, Y., Geller, L. T., Rotter-Maskowitz, A., Weiser, R., Mallel, G., Gigi, E., Meltser, A., Douglas, G. M., Kamer, I., Gopalakrishnan, V., Dadosh, T., Levin-Zaidman, S., Avnet, S., Atlan, T., Cooper, Z. A., … Straussman, R. (2020). The human tumor microbiome is composed of tumor type-specific intracellular bacteria. Science, 368(6494), 973–980. https://doi.org/10.1126/science.aay9189

Network, T. C. G. A. (2012). Comprehensive molecular portraits of human breast tumors The Cancer Genome Atlas Network. Nature, 490(7418), 61–70. https://doi.org/10.1038/nature11412.Comprehensive

Newman, L. A. (2014). Breast Cancer Disparities. High-Risk Breast Cancer and African Ancestry. Surgical Oncology Clinics of North America, 23(3), 579–592. https://doi.org/10.1016/j.soc.2014.03.014

Nguyen, N. K., Deehan, E. C., Zhang, Z., Jin, M., Baskota, N., Perez-Muñoz, M. E., Cole, J., Tuncil, Y. E., Seethaler, B., Wang, T., Laville, M., Delzenne, N. M., Bischoff, S. C., Hamaker, B. R., Martínez, I., Knights, D., Bakal, J. A., Prado, C. M., & Walter, J. (2020). Gut microbiota modulation with long-chain corn bran arabinoxylan in adults with overweight and obesity is linked to an individualized temporal increase in fecal propionate. Microbiome, 8(1), 1–21. https://doi.org/10.1186/s40168-020-00887-w

Nielsen, T. O., Leung, S. C. Y., Rimm, D. L., Dodson, A., Acs, B., Badve, S., Denkert, C., Ellis, M. J., Fineberg, S., Flowers, M., Kreipe, H. H., Laenkholm, A. V., Pan, H., Penault-Llorca, F. M., Polley, M. Y., Salgado, R., Smith, I. E., Sugie, T., Bartlett, J. M. S., … Hayes, D. F. (2021). Assessment of Ki67 in Breast Cancer: Updated Recommendations From the International Ki67 in Breast Cancer Working Group. Journal of the National Cancer Institute, 113(7), 808–819. https://doi.org/10.1093/jnci/djaa201

Nielsen, T. O., Parker, J. S., Leung, S., Voduc, D., Ebbert, M., Vickery, T., Davies, S. R., Snider, J., Stijleman, I. J., Reed, J., Cheang, M. C. U., Mardis, E. R., Perou, C. M., Bernard, P. S., & Ellis, M. J. (2010). A Comparison of PAM50 Intrinsic Subtyping with Immunohistochemistry and Clinical Prognostic Factors in Tamoxifen-Treated Estrogen Receptor – Positive Breast Cancer. 16(21), 5222–5233. https://doi.org/10.1158/1078-0432.CCR-10-1282

Nielsen, T., Wallden, B., Schaper, C., Ferree, S., Liu, S., Gao, D., Barry, G., Dowidar, N., Maysuria, M., & Storhoff, J. (2014). Analytical validation of the PAM50-based Prosigna Breast Cancer Prognostic Gene Signature Assay and nCounter Analysis System using formalin-fixed paraffin-embedded breast tumor specimens. BMC Cancer, 14(1), 1–14. https://doi.org/10.1186/1471-2407-14-177

Nik-zainal, S., Davies, H., Staaf, J., Ramakrishna, M., Glodzik, D., Zou, X., Martincorena, I., Alexandrov, L. B., & Martin, S. (2019). Landscape of somatic mutations in 560 breast cancer whole-genome sequences. 566(7742), 23–25. https://doi.org/10.1038/nature17676

Nolan, E., Lindeman, G. J., & Visvader, J. E. (2023). Deciphering breast cancer : from biology to the clinic. Cell, 186(8), 1708–1728. https://doi.org/10.1016/j.cell.2023.01.040

Nounou, M. I., Elamrawy, F., Ahmed, N., Abdelraouf, K., Goda, S., & Syed-Sha-Qhattal, H. (2015). Breast cancer: Conventional diagnosis and treatment modalities and recent patents and technologies supplementary issue: Targeted therapies in breast cancer treatment. Breast Cancer: Basic and Clinical Research, 9, 17–34. https://doi.org/10.4137/BCBCR.S29420

Nwabo Kamdje, A. H., Seke Etet, P. F., Vecchio, L., Muller, J. M., Krampera, M., & Lukong, K. E. (2014). Signaling pathways in breast cancer: Therapeutic targeting of the microenvironment. Cellular Signalling, 26(12), 2843–2856. https://doi.org/10.1016/j.cellsig.2014.07.034

O'Brien, K. M., Cole, S. R., Tse, C. K., Perou, C. M., Carey, L. A., Foulkes, W. D., Dressler, L. G., Geradts, J., & Millikan, R. C. (2010). Intrinsic breast tumor subtypes, race, and long-term survival in the carolina breast cancer study. Clinical Cancer Research, 16(24), 6100–6110. https://doi.org/10.1158/1078-0432.CCR-10-1533

Ogenyi, S. I., Onu, J. A., Ibeh, N. C., Madukwe, J. U., Onu, O. A., & Menkiti, F. E. (2021). PIK3CA , KI67 , Estrogen and Progesterone Receptors Expression Pattern in HER2 Positive Breast Cancers. Asian Pac J Cancer Biol, 6(4), 281–287. https://doi.org/10.31557/APJCB.2021.6.4.281

Ohnstad, H. O., Borgen, E., Falk, R. S., Lien, T. G., Aaserud, M., Sveli, M. A. T., Kyte, J. A., Kristensen, V. N., Geitvik, G. A., Schlichting, E., Wist, E. A., Sørlie, T., Russnes, H. G., & Naume, B. (2017). Prognostic value of PAM50 and risk of recurrence score in patients with early-stage breast cancer with long-term follow-up. Breast Cancer Research, 19(1), 1–12. https://doi.org/10.1186/s13058-017-0911-9

Pace, L. E., & Keating, N. L. (2014). A systematic assessment of benefits and risks to guide breast cancer screening decisions. Jama, 311(13), 1327–1335. https://doi.org/10.1001/jama.2014.1398

Pace, L. E., & Shulman, L. N. (2016). Breast Cancer in Sub-Saharan Africa: Challenges and Opportunities to Reduce Mortality. The Oncologist, 21(6), 739–744. https://doi.org/10.1634/theoncologist.2015-0429

Palimaru, I., Brügmann, A., Wium-Andersen, M. K., Nexo, E., & Sorensen, B. S. (2013). Expression of PIK3CA, PTEN mRNA and PIK3CA mutations in primary breast cancer: Association with lymph node metastases. SpringerPlus, 2(1), 1–7. https://doi.org/10.1186/2193-1801-2-464

Pang, B., Cheng, S., Sun, S. P., An, C., Liu, Z. Y., Feng, X., & Liu, G. J. (2014). Prognostic role of PIK3CA mutations and their association with hormone receptor expression in breast cancer: A meta-analysis. Scientific Reports, 4, 1–9.

https://doi.org/10.1038/srep06255

Paplomata, E., & Regan, R. O. (2014). The PI3K / AKT / mTOR pathway in breast cancer : targets , trials and biomarkers. Therapeutic Advances in Medical Oncology, 154–166. https://doi.org/10.1177/1758834014530023

Paplomata, E., Zelnak, A., & O'Regan, R. (2013). Everolimus: side effect profile and management of toxicities in breast cancer. Breast Cancer Research and Treatment, 140(3), 453–462. https://doi.org/10.1007/s10549-013-2630-y

Paquet, E. R., & Hallett, M. T. (2015). Absolute assignment of breast cancer intrinsic molecular subtype. Journal of the National Cancer Institute, 107(1), 1–2. https://doi.org/10.1093/jnci/dju357

Parida, S., & Sharma, D. (2020). Microbial Alterations and Risk Factors of Breast Cancer: Connections and Mechanistic Insights. Cells, 9(5), 1–29. https://doi.org/10.3390/cells9051091

Parise, C., & Caggiano, V. (2014). Disparities in the risk of the ER / PR / HER2 breast cancer subtypes among Asian Americans in California. 3–4. https://doi.org/10.1016/j.canep.2014.08.001

Park, B. H., & Davidson, N. E. (2007). PI3 kinase activation and response to Trastuzumab Therapy : what ' s neu with herceptin resistance ? 17936554. https://doi.org/10.1016/j.ccr.2007.10.004

Paul, B., Barnes, S., Demark-Wahnefried, W., Morrow, C., Salvador, C., Skibola, C., & Tollefsbol, T. O. (2015). Influences of diet and the gut microbiome on epigenetic modulation in cancer and other diseases. Clinical Epigenetics, 7(1), 1–11. https://doi.org/10.1186/s13148-015-0144-7

Pauletti, G., Godolphin, W., Press, M. F., & Slamon, D. J. (1996). Detection and quantitation of HER-2/neu gene amplification in human breast cancer archival material using fluorescence in situ hybridization. Oncogene, 13(1), 63–72.

Pérez-Tenorio, G., Alkhori, L., Olsson, B., Waltersson, M. A., Nordenskjöld, B., Rutqvist, L. E., Skoog, L., & Stål, O. (2007). PIK3CA mutations and PTEN loss correlate with similar prognostic factors and are not mutually exclusive in breast cancer. Clinical Cancer Research, 13(12), 3577–3584. https://doi.org/10.1158/1078-0432.CCR-06-1609

Perou, C. M., Sørile, T., Eisen, M. B., Van De Rijn, M., Jeffrey, S. S., Ress, C. A., Pollack, J. R., Ross, D. T., Johnsen, H., Akslen, L. A., Fluge, Ø., Pergammenschlkov, A., Williams, C., Zhu, S. X., Lønning, P. E., Børresen-Dale, A. L., Brown, P. O., & Botstein, D. (2000). Molecular portraits of human breast tumours. Nature, 406(6797), 747–752. https://doi.org/10.1038/35021093

Peters, B. A., Kelly, L., Wang, T., Loudig, O., & Rohan, T. E. (2023). The Breast Microbiome in Breast Cancer Risk and Progression: A Narrative Review. Cancer Epidemiology, Biomarkers & Prevention, OF1–OF11. https://doi.org/10.1158/1055-9965.epi-23-0965

Picardo, S. L., Coburn, B., & Hansen, A. R. (2019). The microbiome and cancer for clinicians. Critical Reviews in Oncology/Hematology, 141, 1–12. https://doi.org/10.1016/j.critrevonc.2019.06.004

Plaza-DÍaz, J., Álvarez-Mercado, A. I., Ruiz-Marín, C. M., Reina-Pérez, I., Pérez-Alonso, A. J., Sánchez-Andujar, M. B., Torné, P., Gallart-Aragón, T., Sánchez-Barrón, M. T., Reyes Lartategui, S., García, F., Chueca, N., Moreno-Delgado, A., Torres-Martínez, K., Sáez-

Lara, M. J., Robles-Sánchez, C., Fernández, M. F., & Fontana, L. (2019). Association of breast and gut microbiota dysbiosis and the risk of breast cancer: A case-control clinical study. BMC Cancer, 19(1), 1–9. https://doi.org/10.1186/s12885-019-5660-y

Plottel, C. S., & Blaser, M. J. (2011). Microbiome and malignancy. 11–12. https://doi.org/10.1016/j.chom.2011.10.003

Pollack, J. R., Perou, C. M., Alizadeh, A. A., Eisen, M. B., Pergamenschikov, A., Williams, C. F., Jeffrey, S. S., Botstein, D., & Brown, P. O. (1999). Genome-wide analysis of DNA copy-number changes using cDNA microarrays. Nature Genetics, 23(1), 41–46. https://doi.org/10.1038/12640

Prat, A., Fan, C., Fernández, A., Hoadley, K. A., Martinello, R., Vidal, M., Viladot, M., Pineda, E., Arance, A., Muñoz, M., Paré, L., Cheang, M. C. U., Adamo, B., & Perou, C. M. (2015). Response and survival of breast cancer intrinsic subtypes following multi-agent neoadjuvant chemotherapy. BMC Medicine, 13(1), 1–11. https://doi.org/10.1186/s12916-015-0540-z

Prat, A., & Perou, C. M. (2011). Deconstructing the molecular portraits of breast cancer. Molecular Oncology, 5(1), 5–23. https://doi.org/10.1016/j.molonc.2010.11.003

Prat, A., Pineda, E., Adamo, B., Galván, P., & Fernández, A. (2015). Clinical implications of the intrinsic molecular subtypes of breast cancer. 2–3. https://doi.org/10.1016/j.breast.2015.07.008

Price, K. N., & Goldhirsch, A. (2005). Clinical trial update : International Breast Cancer Study Group. 1–6. https://doi.org/10.1186/bcr1334

Proctor, L. M., Creasy, H. H., Fettweis, J. M., Lloyd-Price, J., Mahurkar, A., Zhou, W., Buck, G. A., Snyder, M. P., Strauss, J. F., Weinstock, G. M., White, O., & Huttenhower, C. (2019). The Integrative Human Microbiome Project. Nature, 569(7758), 641–648. https://doi.org/10.1038/s41586-019-1238-8

Pu, M., Messer, K., Davies, S. R., Vickery, T. L., Pittman, E., Parker, B. A., Ellis, M. J., Flatt, S. W., Marinac, C. R., Nelson, S. H., Mardis, E. R., Pierce, J. P., & Natarajan, L. (2020). Research-based PAM50 signature and long-term breast cancer survival. Breast Cancer Research and Treatment, 179(1), 197–206. https://doi.org/10.1007/s10549-019-05446-y

Ramirez-Ardila, D. E., Helmijr, J. C., Look, M. P., Lurkin, I., Ruigrok-Ritstier, K., Van Laere, S., Dirix, L., Sweep, F. C., Span, P. N., Linn, S. C., Foekens, J. A., Sleijfer, S., Berns, E. M. J. J., & Jansen, M. P. H. M. (2013). Hotspot mutations in PIK3CA associate with first-line treatment outcome for aromatase inhibitors but not for tamoxifen. Breast Cancer Research and Treatment, 139(1), 39–49. https://doi.org/10.1007/s10549-013-2529-7

Rea, D., Coppola, G., Palma, G., Barbieri, A., Prete, P. Del, Rossetti, S., Berretta, M., Perdonà, S., Turco, M. C., & Arra, C. (2018). Microbiota effects on cancer : from risks to therapies. 9(25), 17915–17927.

Reddy, T. O. (1978). Effect of intestinal microflora and dietary fat on 3,2'-dimethyl-4-aminobiphenyl-induced colon carcinogenesis in F344 ratsv. 41(4), 1363–1367.

Reinhardt, K., Stückrath, K., Hartung, C., Kaufhold, S., Uleer, C., Hanf, V., Lantzsch, T., Peschel, S., John, J., Pöhler, M., Bauer, M., & Karl, F. (2023). PIK3CA - mutations in breast cancer. Breast Cancer Research and Treatment, 2022, 483–493. https://doi.org/10.1007/s10549-022-06637-w

Ren, W., Chen, M., Qiao, Y., & Zhao, F. (2022). Global guidelines for breast cancer screening:

A systematic review. Breast, 64(March), 85–99. https://doi.org/10.1016/j.breast.2022.04.003

Rick, T. J., Habtamu, B., Abreha, A., Solomon, B., Tigeneh, W., Assefa, M., & Incrocci, L. (2020). Ethiopia: How the Care of 100 Million People Pivots on a Single Cobalt Teletherapy Machine. International Journal of Radiation Oncology Biology Physics, 106(2), 230–235. https://doi.org/10.1016/j.ijrobp.2019.06.015

Roy, R., Chun, J., & Powell, S. N. (2016). BRCA1 and BRCA2: different roles in a common pathway of genome protection. Nat Rev Cancer, 12(1), 68–78. https://doi.org/10.1038/nrc3181.BRCA1

Runnak, M. A., Hazha, M. A., Hemin, H. A., Wasan, A. A., Rekawt, R. M., & Michael, H. D. (2012). A population-based study of Kurdish breast cancer in northern Iraq : Hormone receptor and HER2 status . A comparison with Arabic women and United States SEER data. ???, 12(1), 1. https://doi.org/10.1186/1472-6874-12-16

Saal, L. H., Holm, K., Maurer, M., Memeo, L., Su, T., Wang, X., Yu, J. S., Malmström, P. O., Mansukhani, M., Enoksson, J., Hibshoosh, H., Borg, Å., & Parsons, R. (2005). PIK3CA mutations correlate with hormone receptors, node metastasis, and ERBB2, and are mutually exclusive with PTEN loss in human breast carcinoma. Cancer Research, 65(7), 2554–2559. https://doi.org/10.1158/0008-5472-CAN-04-3913

Sabetghadam, M., & Sabetghadam, S. (2013). Investigating Epidemiology and cancer prevalence (breast cancer) in GhotbAbad city. Advances in Environmental Biology, 7(6), 1123–1126.

Sager, M., Yeat, N. C., Pajaro-Van Der Stadt, S., Lin, C., Ren, Q., & Lin, J. (2015). Transcriptomics in cancer diagnostics: Developments in technology, clinical research and commercialization. Expert Review of Molecular Diagnostics, 15(12), 1589–1603. https://doi.org/10.1586/14737159.2015.1105133

Salmena, L., Carracedo, A., & Pandolfi, P. P. (2008). Tenets of PTEN tumor suppression. 18455982. https://doi.org/10.1016/j.cell.2008.04.013

Samuels, Y., Diaz, L. A., Schmidt-Kittler, O., Cummins, J. M., DeLong, L., Cheong, I., Rago, C., Huso, D. L., Lengauer, C., Kinzler, K. W., Vogelstein, B., & Velculescu, V. E. (2005). Mutant PIK3CA promotes cell growth and invasion of human cancer cells. Cancer Cell, 7(6), 561–573. https://doi.org/10.1016/j.ccr.2005.05.014

Samuels, Y., & Velculescu, V. E. (2004). Oncogenic Mutations of PIK3CA in Human Cancers nd es Bio sci No t D ist r. 4101. https://doi.org/10.4161/cc.3.10.1164

Samuels, Y., Wang, Z., Bardelli, A., Silliman, N., Ptak, J., Szabo, S., Yan, H., Gazdar, A., Powell, S. M., Riggins, G. J., Willson, J. K. V., Markowitz, S., Kinzler, K. W., Vogelstein, B., & Velculescu, V. E. (2004). High Frequency of Mutations of the PIK3CA Gene in Human Cancers. Science, 304(5670), 554. https://doi.org/10.1126/science.1096502

Sarbassov, D. D., Ali, S. M., Sengupta, S., Sheen, J., Hsu, P. P., Bagley, A. F., Markhard, A. L., & Sabatini, D. M. (2006). Prolonged Rapamycin Treatment Inhibits mTORC2 Assembly and Akt / PKB. 159–168. https://doi.org/10.1016/j.molcel.2006.03.029

Sayed, S., Moloo, Z., Wasike, R., Bird, P., Oigara, R., Govender, D., Kibera, J., Carrara, H., & Saleh, M. (2014). Is breast cancer from Sub Saharan Africa truly receptor poor? Prevalence of ER/PR/HER2 in breast cancer from Kenya. Breast, 23(5), 591–596. https://doi.org/10.1016/j.breast.2014.06.006

Schettini, F. (2022). OPEN A perspective on the development and lack of interchangeability of the breast cancer intrinsic subtypes. 1–4. https://doi.org/10.1038/s41523-022-00451-9

Schlaeppi, K., & Bulgarelli, D. (2015). The plant microbiome at work. 25514681. https://doi.org/10.1094/MPMI-10-14-0334-FI

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., Lesniewski, R. A., Oakley, B. B., Parks, D. H., Robinson, C. J., Sahl, J. W., Stres, B., Thallinger, G. G., Van Horn, D. J., & Weber, C. F. (2009). Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. Applied and Environmental Microbiology, 75(23), 7537–7541. https://doi.org/10.1128/AEM.01541-09

Schwabe, R. F., & Jobin, C. (2013). The microbiome and cancer. 11–12. https://doi.org/10.1038/nrc3610

Sengal, A. T., Mukhtar, N. S. H., Vetter, M., Elhaj, A. M., Bedri, S., Hauptmann, S., Thomssen, C., Mohamedani, A. A., Wickenhauser, C., & Kantelhardt, E. J. (2018). Comparison of receptor-defined Breast Cancer Subtypes between German and sudanese women: A facility-based cohort study. Journal of Global Oncology, 2018(4), 1–12. https://doi.org/10.1200/JGO.2017.010082

Serrano-Gomez, S. J., Sanabria-Salas, M. C., Hernández-Suarez, G., García, O., Silva, C., Romero, A., Mejía, J. C., Miele, L., Fejerman, L., & Zabaleta, J. (2016). High prevalence of luminal B breast cancer intrinsic subtype in Colombian women. Carcinogenesis, 37(7), 669–676. https://doi.org/10.1093/carcin/bgw043

Shamseddine, A., Tfayli, A., Temraz, S., & Abou Mrad, R. (2010). Breast cancer in low- and middle-income countries: An emerging and challenging epidemic. Journal of Oncology, 2010. https://doi.org/10.1155/2010/490631

Shapira, I., Sultan, K., Lee, A., & Taioli, E. (2013). Evolving Concepts: How Diet and the Intestinal Microbiome Act as Modulators of Breast Malignancy. ISRN Oncology, 2013, 1–10. https://doi.org/10.1155/2013/693920

Sharma, R., Nanda, M., Fronterre, C., Sewagudde, P., Amponsah-manu, F., & Ssentongo, P. (2022). Mapping Cancer in Africa : A Comprehensive and Comparable Characterization of 34 Cancer Types Using Estimates From GLOBOCAN. 10(April), 1–14. https://doi.org/10.3389/fpubh.2022.839835

Shike, M., Doane, A. S., Russo, L., Cabal, R., Reis-Filho, J. S., Gerald, W., Cody, H., Khanin, R., Bromberg, J., & Norton, L. (2014). The effects of soy supplementation on gene expression in breast cancer: A randomized placebo-controlled study. Journal of the National Cancer Institute, 106(9). https://doi.org/10.1093/jnci/dju189

Shimoi, T., Noguchi, E., Hamada, A., Yamagishi, M., Hirai, M., Sudo, K., Shimomura, A., Yonemori, K., Shimizu, C., Yoshida, M., Nishikawa, T., Yunokawa, M., Kinoshita, T., Fukuda, T., Fujiwara, Y., & Tamura, K. (2018). PIK3CA mutation profiling in patients with breast cancer , using a highly sensitive detection system. November 2017, 2558–2566. https://doi.org/10.1111/cas.13696

Shita, A., Yalew, A. W., Id, E. S., Afework, T., Tesfaw, A., Hagos, Z., Id, G., Rabe, F., Taylor, L., Kantelhardt, E. J., & Id, S. G. (2023). PLOS ONE Survival and predictors of breast cancer mortality in South Ethiopia : A retrospective cohort study. 1–16. https://doi.org/10.1371/journal.pone.0282746

Sinn, H., Schneeweiss, A., Keller, M., Schlombs, K., Laible, M., Seitz, J., Lakis, S., Veltrup,

E., Altevogt, P., Eidt, S., Wirtz, R. M., & Marmé, F. (2017). Comparison of immunohistochemistry with PCR for assessment of ER , PR , and Ki-67 and prediction of pathological complete response in breast cancer. 1–10. https://doi.org/10.1186/s12885-017-3111-1

Smith, A., Cao, X., Gu, Q., Kubi Amos-Abanyie, E., Tolley, E. A., Vidal, G., Lyn-Cook, B., & Starlard-Davenport, A. (2022). Characterization of the Metabolome of Breast Tissues from Non-Hispanic Black and Non-Hispanic White Women Reveals Correlations between Microbial Dysbiosis and Enhanced Lipid Metabolism Pathways in Triple-Negative Breast Tumors. Cancers, 14(17), 2–3. https://doi.org/10.3390/cancers14174075

Smith, A., Pierre, J. F., Makowski, L., Tolley, E., Lyn-Cook, B., Lu, L., Vidal, G., & Starlard-Davenport, A. (2019). Distinct microbial communities that differ by race, stage, or breast-tumor subtype in breast tissues of non-Hispanic Black and non-Hispanic White women. Scientific Reports, 9(1), 1–2. https://doi.org/10.1038/s41598-019-48348-1

Smolarz, B., Zadrożna Nowak, A., & Romanowicz, H. (2022). Breast Cancer—Epidemiology, Classification, Pathogenesis and Treatment (Review of Literature). Cancers, 14(10), 1–27. https://doi.org/10.3390/cancers14102569

Sobhani, N., Roviello, G., Corona, S. P., Scaltriti, M., Ianza, A., Bortul, M., Zanconati, F., & Generali, D. (2018). The prognostic value of PI3K mutational status in breast cancer : a meta-analysis. 1–11. https://doi.org/10.1002/jcb.26687

Soerjomataram, I., & Bray, F. (2021). Planning for tomorrow: global cancer incidence and the role of prevention 2020–2070. Nature Reviews Clinical Oncology, 18(10), 663–672. https://doi.org/10.1038/s41571-021-00514-z

Soerjomataram, I., Lortet-Tieulent, J., Parkin, D. M., Ferlay, J., Mathers, C., Forman, D., & Bray, F. (2012). Global burden of cancer in 2008: A systematic analysis of disability-adjusted life-years in 12 world regions. The Lancet, 380(9856), 1840–1850. https://doi.org/10.1016/S0140-6736(12)60919-2

Sørlie, T., Perou, C. M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., Hastie, T., Eisen, M. B., Van De Rijn, M., Jeffrey, S. S., Thorsen, T., Quist, H., Matese, J. C., Brown, P. O., Botstein, D., Lønning, P. E., & Børresen-Dale, A. L. (2001). Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proceedings of the National Academy of Sciences of the United States of America, 98(19), 10869–10874. https://doi.org/10.1073/pnas.191367098

Sørlie, T., Tibshirani, R., Parker, J., Hastie, T., Marron, J. S., Nobel, A., Deng, S., Johnsen, H., Pesich, R., Geisler, S., Demeter, J., Perou, C. M., Lønning, P. E., Brown, P. O., Børresen-Dale, A. L., & Botstein, D. (2003). Repeated observation of breast tumor subtypes in independent gene expression data sets. Proceedings of the National Academy of Sciences of the United States of America, 100(14), 8418–8423. https://doi.org/10.1073/pnas.0932692100

Sotiriou, C., & Pusztai, L. (2009). Gene-Expression Signatures in Breast Cancer. New England Journal of Medicine, 360(8), 790–800. https://doi.org/10.1056/nejmra0801289

Stemke-hale, K., Gonzalez-angulo, A. M., Lluch, A., Neve, R. M., Kuo, W., Davies, M., Carey, M., Hu, Z., Guan, Y., Symmans, W. F., Pusztai, L., Nolden, L. K., Horlings, H., Hung, M., Vijver, M. J. Van De, Valero, V., & Gray, J. W. (2009). NIH Public Access. 68(15), 6084–6091. https://doi.org/10.1158/0008-5472.CAN-07-6854.An

Stingl, J., Eirew, P., Ricketson, I., Shackleton, M., Vaillant, F., Choi, D., & Li, H. I. (2006).

Purification and unique properties of mammary epithelial stem cells. 16395311. https://doi.org/10.1038/nature04496

Sudhakar, N., George Priya Doss, C., Kumar, D. T., Chakraborty, C., Anand, K., & Suresh, M. (2016). Deciphering the impact of somatic mutations in exon 20 and exon 9 of PIK3CA gene in breast tumors among Indian women through molecular dynamics approach. Journal of Biomolecular Structure and Dynamics, 34(1), 29–41. https://doi.org/10.1080/07391102.2015.1007483

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: A Cancer Journal for Clinicians, 71(3), 209–249. https://doi.org/10.3322/caac.21660

Sweeney, C., Bernard, P. S., Factor, R. E., Kwan, M. L., Habel, L. A., Quesenberry, C. P., Shakespear, K., Weltzien, E. K., Stijleman, I. J., Davis, C. A., Ebbert, M. T. W., Castillo, A., Kushi, L. H., & Caan, B. J. (2014). Intrinsic subtypes from PAM50 gene expression assay in a population-based breast cancer cohort: Differences by age, race, and tumor characteristics. Cancer Epidemiology Biomarkers and Prevention, 23(5), 714–724. https://doi.org/10.1158/1055-9965.EPI-13-1023

Tabesh, G. A., Izadi, P., & Fereidooni, F. (2022). The High Frequency of PIK3CA Mutations in Iranian Breast Cancer Patients. 11–14.

Tabesh, G. A., Izadi, P., Fereidooni, F., Nader, A., & Razavi, E. (2017). The High Frequency of PIK3CA Mutations in Iranian Breast Cancer Patients. 35(1), 10–11. https://doi.org/10.1080/07357907.2016.1247455

Tamaki, M., Kamio, T., Kameoka, S., Kojimahara, N., & Nishikawa, T. (2013). The relevance of the intrinsic subtype to the clinicopathological features and biomarkers in Japanese breast cancer patients. World Journal of Surgical Oncology, 11, 1–13. https://doi.org/10.1186/1477-7819-11-293

Thompson, K. J., Ingle, J. N., Tang, X., Chia, N., Jeraldo, P. R., Walther-Antonio, M. R., Kandimalla, K. K., Johnson, S., Yao, J. Z., Harrington, S. C., Suman, V. J., Wang, L., Weinshilboum, R. L., Boughey, J. C., Kocher, J. P., Nelson, H., Goetz, M. P., & Kalari, K. R. (2017). A comprehensive analysis of breast cancer microbiota and host gene expression. PLoS ONE, 12(11), 1–17. https://doi.org/10.1371/journal.pone.0188873

Timotewos, G., Solomon, A., Mathewos, A., Addissie, A., Bogale, S., Wondemagegnehu, T., Aynalem, A., Ayalnesh, B., Dagnechew, H., Bireda, W., Kroeber, E. S., Mikolajczyk, R., Bray, F., Jemal, A., & Kantelhardt, E. J. (2018). First data from a population based cancer registry in Ethiopia. Cancer Epidemiology, 53, 93–98. https://doi.org/10.1016/j.canep.2018.01.008

Toumazi, D., Daccache, S. E. L., & Constantinou, C. (2021). An unexpected link : The role of mammary and gut microbiota on breast cancer development and management (Review). https://doi.org/10.3892/or.2021.8031

Traoré, B., Koulibaly, M., Diallo, A., & Bah, M. (2019). Molecular profile of breast cancers in Guinean oncological settings. Pan African Medical Journal, 33, 1–7. https://doi.org/10.11604/pamj.2019.33.22.18189

Troester, M. A., Sun, X., Allott, E. H., Geradts, J., Cohen, S. M., Tse, C. K., Kirk, E. L., Thorne, L. B., Mathews, M., Li, Y., Hu, Z., Robinson, W. R., Hoadley, K. A., Olopade, O. I., Reeder-Hayes, K. E., Earp, H. S., Olshan, A. F., Carey, L. A., & Perou, C. M. (2018).

Racial Differences in PAM50 Subtypes in the Carolina Breast Cancer Study. Journal of the National Cancer Institute, 110(2), 176–182. https://doi.org/10.1093/jnci/djx135

Tsang, J. Y. S., & Tse, G. M. (2020). Molecular Classification of Breast Cancer. Advances in Anatomic Pathology, 27(1), 27–35. https://doi.org/10.1097/PAP.0000000000000232

Turnbaugh, P. J., Hamady, M., Yatsunenko, T., Cantarel, B. L., Duncan, A., & Ley, R. E. (2009). A core gut microbiome in obese and lean twins. 1–14. https://doi.org/10.1038/nature07540

Tzeng, A., Sangwan, N., Jia, M., Liu, C. C., Keslar, K. S., Downs-Kelly, E., Fairchild, R. L., Al-Hilli, Z., Grobmyer, S. R., & Eng, C. (2021). Human breast microbiome correlates with prognostic features and immunological signatures in breast cancer. Genome Medicine, 13(1), 22–23. https://doi.org/10.1186/s13073-021-00874-2

Unger-Saldaña, K. (2014). Challenges to the early diagnosis and treatment of breast cancer in developing countries. World Journal of Clinical Oncology, 5(3), 465–477. https://doi.org/10.5306/wjco.v5.i3.465

Urbaniak, C., Burton, J. P., & Reid, G. (2012). Breast, milk and microbes: A complex relationship that does not end with lactation. Women's Health, 8(4), 385–398. https://doi.org/10.2217/whe.12.23

Urbaniak, C., Cummins, J., Brackstone, M., Macklaim, J. M., Gloor, G. B., Baban, C. K., Scott, L., O'Hanlon, D. M., Burton, J. P., Francis, K. P., Tangney, M., & Reida, G. (2014). Microbiota of human breast tissue. Applied and Environmental Microbiology, 80(10), 3007–3014. https://doi.org/10.1128/AEM.00242-14

Urbaniak, C., Gloor, G. B., Brackstone, M., Scott, L., Tangney, M., & Reida, G. (2016). The microbiota of breast tissue and its association with breast cancer. Applied and Environmental Microbiology, 82(16), 5039–5048. https://doi.org/10.1128/AEM.01235-16

Van Hoeck, A., Tjoonk, N. H., Van Boxtel, R., & Cuppen, E. (2019). Portrait of a cancer: Mutational signature analyses for cancer diagnostics. BMC Cancer, 19(1), 1–14. https://doi.org/10.1186/s12885-019-5677-2

Vanderpuye, V., Dadzie, M.-A., Huo, D., & Olopade, O. I. (2021). Assessment of Breast Cancer Management in Sub-Saharan Africa. JCO Global Oncology, 7, 1593–1601. https://doi.org/10.1200/go.21.00282

Vanderpuye, V., Grover, S., Hammad, N., Prabhakar, P., Simonds, H., Olopade, F., & Stefan, D. C. (2017). An update on the management of breast cancer in Africa. Infectious Agents and Cancer, 12(1), 1–12. https://doi.org/10.1186/s13027-017-0124-y

Varga, Z., Lebeau, A., Bu, H., Hartmann, A., & Penault-llorca, F. (2017). An international reproducibility study validating quantitative determination of ERBB2 , ESR1 , PGR , and MKI67 mRNA in breast cancer using MammaTyper ®. 1–2. https://doi.org/10.1186/s13058-017-0848-z

Vasan, N., Toska, E., & Scaltriti, M. (2019). Overview of the relevance of PI3K pathway in HR-positive breast cancer. 30(Supplement 10), 3–11. https://doi.org/10.1093/annonc/mdz281

Voduc, K. D., Cheang, M. C. U., Tyldesley, S., Gelmon, K., Nielsen, T. O., & Kennecke, H. (2010). Breast cancer subtypes and the risk of local and regional relapse. Journal of Clinical Oncology, 28(10), 1684–1691. https://doi.org/10.1200/JCO.2009.24.9284

Vuong, D., Simpson, P. T., Green, B., Cummings, M. C., & Lakhani, S. R. (2014). Molecular classification of breast cancer. Virchows Archiv, 465(1), 1–14. https://doi.org/10.1007/s00428-014-1593-7

Wallden, B., Storhoff, J., Nielsen, T., Dowidar, N., Schaper, C., Ferree, S., Liu, S., Leung, S., Geiss, G., Snider, J., Vickery, T., Davies, S. R., Mardis, E. R., Gnant, M., Sestak, I., Ellis, M. J., Perou, C. M., Bernard, P. S., & Parker, J. S. (2015). Development and verification of the PAM50-based Prosigna breast cancer gene signature assay. 1–14. https://doi.org/10.1186/s12920-015-0129-6

Wander, S. A., Hennessy, B. T., & Slingerland, J. M. (2011). Science in medicine Next-generation mTOR inhibitors in clinical oncology : how pathway complexity informs therapeutic strategy. 121(4), 1231–1241. https://doi.org/10.1172/JCI44145.tion

Wang, H., Altemus, J., Niazi, F., Green, H., Calhoun, B. C., Sturgis, C., Grobmyer, S. R., & Eng, C. (2017). Breast tissue, oral and urinary microbiomes in breast cancer. Oncotarget, 8(50), 88122–88138. https://doi.org/10.18632/oncotarget.21490

Wang, L., Li, Q., Aushev, V. N., Neugut, A. I., Santella, R. M., Teitelbaum, S., & Chen, J. (2021). PAM50- and immunohistochemistry-based subtypes of breast cancer and their relationship with breast cancer mortality in a population-based study. Breast Cancer, 28(6), 1235–1242. https://doi.org/10.1007/s12282-021-01261-w

Wang, L., Zhang, Q., Zhang, J., Sun, S., Guo, H., Jia, Z., Wang, B., Shao, Z., Wang, Z., & Hu, X. (2011). PI3K pathway activation results in low efficacy of both trastuzumab and lapatinib. BMC Cancer, 11(1), 248. https://doi.org/10.1186/1471-2407-11-248

Wang, N., Sun, T., & Xu, J. (2021). Tumor-related Microbiome in the Breast Microenvironment and Breast Cancer. Journal of Cancer, 12(16), 4841–4848. https://doi.org/10.7150/JCA.58986

Wang, X., Bannon, F., Ahn, J. V, Johnson, C. J., & Bonaventure, A. (2015). Global surveillance of cancer survival 1995-2009 : analysis of individual data for 25 , 676 , 887 patients from 279 population-based registries in 67 countries (CONCORD-2). 35–37. https://doi.org/10.1016/S0140-6736(14)62038-9

Weigelt, B., Mackay, A., A'hern, R., Natrajan, R., Tan, D. S. P., Dowsett, M., Ashworth, A., & Reis-Filho, J. S. (2010). Breast cancer molecular profiling with single sample predictors: A retrospective analysis. The Lancet Oncology, 11(4), 339–349. https://doi.org/10.1016/S1470-2045(10)70008-5

WHO/IARC. (2018). Cancer Today 2018 data.

WHO. (2018). Breast Cancer Diagnosis, prevention and Screening. World Health Organization.

Winters, S., Martin, C., Murphy, D., & Shokar, N. K. (2017). Breast Cancer Epidemiology , Prevention , and Screening. 23–24. https://doi.org/10.1016/bs.pmbts.2017.07.002

Woldeamanuel, Y. W., Girma, B., & Teklu, A. M. (2013). Cancer in Ethiopia. The Lancet Oncology, 14(4), 289–290. https://doi.org/10.1016/S1470-2045(12)70399-6

Wolfe, A. J., Toh, E., Shibata, N., Rong, R., Kenton, K., FitzGerald, M. P., Mueller, E. R., Schreckenberger, P., Dong, Q., Nelson, D. E., & Brubaker, L. (2012). Evidence of uncultivated bacteria in the adult female bladder. Journal of Clinical Microbiology, 50(4), 1376–1383. https://doi.org/10.1128/JCM.05852-11

Wolff, A. C., Hammond, M. E. H., Allison, K. H., Harvey, B. E., Mangu, P. B., Bartlett, J. M. S., Bilous, M., Ellis, I. O., Fitzgibbons, P., Hanna, W., Jenkins, R. B., Press, M. F., Spears, P. A., Vance, G. H., Viale, G., Mcshane, L. M., & Dowsett, M. (2018). Human Epidermal Growth Factor Receptor 2 Testing in Breast Cancer : American Society of Clinical Oncology / College of American Pathologists Clinical Practice Guideline Focused Update. 1–2. https://doi.org/10.5858/arpa.2018-0902-SA

Wondemagegnhu, T. (2015). Pattern of Cancer in Tikur Anbessa Specialized Hospital Oncology Center in Ethiopia from 1998 to 2010. International Journal of Cancer Research and Molecular Mechanisms (ISSN 2381-3318), 1(1), 1–5. https://doi.org/10.16966/2381-3318.103

Wu, H., Wang, W., Du, J., Li, H., Wang, H., Huang, L., Xiang, H., Xie, J., Liu, X., Li, H., & Lin, W. (2019). The distinct clinicopathological and prognostic implications of PIK3CA mutations in breast cancer patients from central China. Cancer Management and Research, 11, 1473–1492. https://doi.org/10.2147/CMAR.S195351

Wu, N. C., Wong, W., Ho, K. E., Chu, V. C., Rizo, A., Davenport, S., Kelly, D., Makar, R., Jassem, J., Duchnowska, R., Biernat, W., Radecka, B., Fujita, T., Klein, J. L., Stonecypher, M., Ohta, S., Juhl, H., Weidler, J. M., Bates, M., & Press, M. F. (2018). Comparison of central laboratory assessments of ER, PR, HER2, and Ki67 by IHC/FISH and the corresponding mRNAs (ESR1, PGR, ERBB2, and MKi67) by RT-qPCR on an automated, broadly deployed diagnostic platform. Breast Cancer Research and Treatment, 172(2), 327–338. https://doi.org/10.1007/s10549-018-4889-5

Xuan, C., Shamonki, J. M., Chung, A., DiNome, M. L., Chung, M., Sieling, P. A., & Lee, D. J. (2014). Microbial dysbiosis is associated with human breast cancer. PLoS ONE, 9(1), 1–7. https://doi.org/10.1371/journal.pone.0083744

Xue, M., Ji, X., Liang, H., Liu, Y., Wang, B., Sun, L., & Li, W. (2018). The effect of fucoidan on intestinal flora and intestinal barrier function in rats with breast cancer. Food and Function, 9(2), 1214–1223. https://doi.org/10.1039/c7fo01677h

Yadav, S., Boddicker, N. J., Samara, R., Klebba, J., Ambrosone, C. B., Culver, H. A., Auer, P., Bandera, E. V, Bernstein, L., Bertrand, K. A., Burnside, E. S., Carter, B. D., Eliassen, H., Gapstur, S. M., Gaudet, M., Haiman, C., Hodge, J. M., Hunter, D. J., Jacobs, E. J., … Goldgar, D. E. (2021). A Population-Based Study of Genes Previously Implicated in Breast Cancer. https://doi.org/10.1056/NEJMoa2005936

Yamamoto, S., Tsuda, H., Takano, M., Iwaya, K., Tamai, S., & Matsubara, O. (2011). PIK3CA mutation is an early event in the development of endometriosis-associated ovarian clear cell adenocarcinoma. Journal of Pathology, 225(2), 189–194. https://doi.org/10.1002/path.2940

Yang, J., Tan, Q., Fu, Q., Zhou, Y., Hu, Y., Tang, S., Zhou, Y., Zhang, J., Qiu, J., & Lv, Q. (2017). Gastrointestinal microbiome and breast cancer: correlations, mechanisms and potential clinical implications. Breast Cancer, 24(2), 220–228. https://doi.org/10.1007/s12282-016-0734-z

Yang, P., Wang, Z., Peng, Q., Lian, W., & Chen, D. (2021). Comparison of the Gut Microbiota in Patients with Benign and Malignant Breast Tumors: A Pilot Study. Evolutionary Bioinformatics, 17, 23–24. https://doi.org/10.1177/11769343211057573

Yates, L. R., & Desmedt, C. (2017). Translational Genomics : Practical Applications of the Genomic Revolution in Breast Cancer. 28572257. https://doi.org/10.1158/1078-0432.CCR-16-2548

Yuan, & Cantley. (2008). PI3K pathway alterations in cancer. Oncogene, 27(41), 5497–5510. https://doi.org/10.1038/onc.2008.245.PI3K

Zardavas, D., Fumagalli, D., & Loi, S. (2012). Phosphatidylinositol 3-kinase/AKT/mammalian target of rapamycin pathway inhibition: A breakthrough in the management of luminal (ER +HER2-) breast cancers? Current Opinion in Oncology, 24(6), 623–634. https://doi.org/10.1097/CCO.0b013e328358a2b5

Zardavas, D., Te Marvelde, L., Milne, R. L., Fumagalli, D., Fountzilas, G., Kotoula, V., Razis, E., Papaxoinis, G., Joensuu, H., Moynahan, M. E., Hennessy, B. T., Bieche, I., Saal, L. H., Stal, O., Iacopetta, B., Jensen, J. D., O'Toole, S., Lopez-Knowles, E., Barbaraeschi, M., … Loi, S. (2018). Tumor PIK3CA genotype and prognosis in early-stage breast cancer: A pooled analysis of individual patient data. Journal of Clinical Oncology, 36(10), 981–990. https://doi.org/10.1200/JCO.2017.74.8301

Zhang, J., Xia, Y., & Sun, J. (2021). Breast and gut microbiome in health and cancer. Genes and Diseases, 8(5), 581–589. https://doi.org/10.1016/j.gendis.2020.08.002

Zhang, W., Haines, B. B., Efferson, C., Zhu, J., Ware, C., Kunii, K., Tammam, J., Angagaw, M., Hinton, M. C., Keilhack, H., Paweletz, C. P., Zhang, T., Winter, C., Sathyanarayanan, S., Cheng, J., Zawel, L., Fawell, S., Gilliland, G., & Majumder, P. K. (2012). Tr a n s l a t i o n a l O n c o l o g y Evidence of mTOR Activation by an AKT-Independent Mechanism Provides Support for the Combined Treatment of PTEN-Deficient Prostate Tumors with mTOR and. 5(6), 422–429. https://doi.org/10.1593/tlo.12241

Zhao, J. J., Liu, Z., Wang, L., Shin, E., Loda, M. F., & Roberts, T. M. (2005). The oncogenic properties of mutant p110α and p110β phosphatidylinositol 3-kinases in human mammary epithelial cells. Proceedings of the National Academy of Sciences of the United States of America, 102(51), 18443–18448. https://doi.org/10.1073/pnas.0508988102

Zhao, L., & Vogt, P. K. (2008a). Class I PI3K in oncogenic cellular transformation. Oncogene, 27(41), 5486–5496. https://doi.org/10.1038/onc.2008.244

Zhao, L., & Vogt, P. K. (2008b). Helical domain and kinase domain mutations in p110α of phosphatidylinositol 3-kinase induce gain of function by different mechanisms. Proceedings of the National Academy of Sciences of the United States of America, 105(7), 2652–2657. https://doi.org/10.1073/pnas.0712169105

Zhao, X. (2013). PIK3CA gene mutation associated with poor prognosis of lung adenocarcinoma. 497–502.

Ziegenhorn, H. V., Frie, K. G., Ekanem, I. O., Ebughe, G., Kamate, B., Traore, C., Dzamalala, C., Ogunbiyi, O., Igbinoba, F., Liu, B., Bauer, M., Thomssen, C., Parkin, D. M., Wickenhauser, C., & Kantelhardt, E. J. (2020). Breast cancer pathology services in sub-Saharan Africa: A survey within population-based cancer registries. BMC Health Services Research, 20(1), 1–9. https://doi.org/10.1186/s12913-020-05752-y

Zitvogel, L., Ayyoub, M., Routy, B., & Kroemer, G. (2016). Microbiome and Anticancer Immunosurveillance. Cell, 165(2), 276–287. https://doi.org/10.1016/j.cell.2016.03.001

9. LIST OF PUBLICATIONS, POTENTIAL MANUSCRIPT AND PRESENTED PAPERS

11.1. Published papers included in this thesis:

1. **Zelalem Desalegn** , Meron Yohannes , Martin Porsch, Kathrin Stückrath, Endale Anberber, Pablo Santos, Marcus Bauer, Adamu Addissie, Yonas Bekuretsion, Mathewos Assefa, Yasin Worku, Lesley Taylor, Tamrat Abebe, Eva Johanna Kantelhardt, Martina Vetter. Intrinsic subtypes in Ethiopian breast cancer patient. Breast cancer Res Treat. 2022 Dec;196(3):495-504. doi: 10.1007/s10549-022-06769-z. https://pubmed.ncbi.nlm.nih.gov/36282363/
2. **Zelalem Desalegn**, Alana Smith, Meron Yohannes , Xueyuan Cao, Endale Anberber, Yonas Bekuretsion, Mathewos Assefa , Marcus Bauer, Martina Vetter , Eva Johanna Kantelhardt , Tamrat Abebe , Athena Starlard-Davenport . Human Breast Tissue Microbiota Reveals Unique Microbial Signatures that Correlate with Prognostic Features in Adult Ethiopian Women with breast cancer. Cancers (Basel). 2023 Oct 9;15(19):4893. doi: 10.3390/cancers15194893.https://pubmed.ncbi.nlm.nih.gov/37835588/

11.2. Co-authored publications (not included in this thesis)

1. Dessiet Oma [1], Maria Teklemariam , Daniel Seifu, **Zelalem Desalegn** , Endale Anberbir , Tamrat Abebe , Solomon Mequannent, Solomon Tebeje, Wajana Lako Labisso. Immunohistochemistry versus PCR Technology for Molecular Subtyping of breast cancer: Multicentered Experiences from Addis Ababa, Ethiopia. J Cancer Prev. 2023 Jun 30;28(2):64-74. doi: Paper published with collaborator in Germany

11.3. **Potential Manuscript for publication**

1. Frequency of mutations in *PIK3CA* gene and its prognostic role among Ethiopian patients with breast cancer
2. mRNA and Protein Level Biomarker Expression Concordance for ESR1/ER, PGR/PgR and ERBB2/HER2 in Ethiopian Patients with breast cancer: Implications for Diagnostic Alternatives in Ethiopia
3. Immune landscape of the tumor microenvironment in Ethiopian breast cancer patients.
4. Immune cell populations profile among Ethiopian patients with breast cancer

11.4. Attended conferences to present thesis derived abstract (oral/ poster)

1. Tumor immunology meets oncology annual conference, Halle, Germany, 2018
2. Tumor immunology meets oncology annual conference, Halle, Germany, 2019
3. American Association for cancer research (AACR) November 1 to 5, 2018
4. AORTIC 2019 conference in Maputo from November 5 – 8 2019 conference, Mozambique
5. 1st International Conference of Hospital Partnerships 2022, Berlin German
6. AORTIC 2023 conference in Senegal from November 2-6 2023

10. APPENDIXES

Appendix I: Immunohistochemical staining protocol

name	clone	host	order nr.	company	dilution
Estrogen receptor (ER)	Ab-11	mouse	MS-354-P1	Thermo Scientific	1:150
Progesterone receptor (PgR)	PgR 636	mouse	M3569	Dako	1:100
Ki-67	SP6	rabbit	RM-9106-S	Thermo Scientific	1:250

1. Principle of the method

Paraffin-embedded tissue sections are first deparaffinized and rehydrated. Endogenous peroxidase activity in the tissue may cause non-specific staining. This enzyme activity can be blocked by incubation with 3% H_2O_2-solution (peroxide block). Epitopes of antigens are being masked and often are no longer accessible for primary antibodies, because of the strong cross-links between proteins induced by formalin fixation. Heat induced epitope retrieval (HIER) in buffer solutions of different compositions and pH-values restore structures of the epitopes making them more accessible to specific antibodies. The next step is incubation with the specific primary antibody. After washing, the enhancement reagent ("PostBlock") is applied and incubated. A second washing is followed by the application of the HRP-polymer. Any excess of unbound HRP-polymer is thoroughly washed away after incubation. The addition of the chromogenic substrate starts the enzymatic reaction of the peroxidase which leads to color precipitation where the primary antibody is bound. The color can be observed with a light microscope. The chromogen used determines the color. The chromogen DAB forms a dark brown precipitate.

2. Material and reagents

2.1 reagents provided

- primary antibody (ER, PgR, Ki67)
- retrieval buffer: HercepTest™ Epitope Retrieval Solution (Containing Detergent) (10x)
- antibody diluent: Zytomed Antibody diluent (ready-to-use) (ZUC025)
- Zytomed ZytoChem Plus HRP Polymer Kit (POLHRP-100):

- blocking solution (not needed, because antibody diluent includes blocking solution)
- Post Block (ready-to-use)
- HRP-Polymer (ready-to-use)

- wash buffer: Zytomed Wash Buffer (20x) (ZUC020)
- DAB Substrate Kit: Zytomed DAB Chromogen + Zytomed DAB Substrate buffer (DAB057)

2.2 Materials required but not supplied

- Positive control tissue:
  - Estrogen receptor - MCF-7 cells, Breast CA
  - Progesterone receptor – normal breast, cervix uteri, prostate
  - Ki67 - SKBR3 cells, Spleen, tonsil, or breast CA
- Slides: poly-L-lysine coated slides or silanized slides (e.g. SuperFrost Plus)
- Xylene or suitable substitutes
- Ethanol
- distilled water
- 30% H_2O_2 (hydrogen peroxide) solution
- Counter stain solution: Hematoxylin Mayer´s
- Mounting medium, coverslip

2.3 Reagent preparation

- Reagents should be at room temperature when used
- Tissue sections have to be completely covered with the different reagents in order to avoiddrying out
- **Endogenous peroxidase blocking reagent**/ Hydrogen peroxide (H_2O_2):
  - Dilute 30% H_2O_2 into 1:10 with distilled water, to get an 3% H_2O_2-dilution
  - E.g. 20ml of 30% H_2O_2 + 180ml distilled water
  - H_2O_2 is light-sensitive >> protect from sun and prepare every day fresh

Epitope retrieval solution:

- Dilute a sufficient quantity of HercepTest™ Epitope Retrieval Solution 1:10 using distilled ordeionized water
- Diluted retrieval solution can be reused tree times
- The diluted working strength solution is stable for about 1 week at room temperature or onemonth at 2-8°C depending on the ambient temperature
- Discard diluted solution if cloudy in appearance

antibody dilution:

- 100µl diluted antibody per slide
- antibody diluent includes blocking solution, that's why blocking solution of ZytoChem-PlusHRP Polymer-Kit is not needed
- Estrogen receptor (**ER**): **1:150** (e.g.: 6µl antibody + 900µl antibody diluent >> remains for 9samples)
- Progesteron receptor (**PgR**): **1:100**

Ki-67: 1:250

- **wash buffer** preparation:
  - Dilute Wash Buffer concentrate 1:20 with deionized or distilled water and mix thoroughly
  - The diluted working strength solution is stable for about 1 week depending on the ambienttemperature
- **DAB** working solution preparation:
  - Add 4 drops of DAB Chromogen (DAB concentrate) to one bottle of DAB Substrate Buffer andmix thoroughly.
  - The working solution should be prepared freshly at the day of use. Once the two reagents arecombined, the resulting solution can be used for up to six hours. Excess working solution needs to be disposed as hazardous substance.
  - The DAB chromogen is hazardous to your health. Wear protective clothing to avoid contactof reagents or specimen with eye, skin or mucous membrane

3. Procedure

Important points before starting

- preheating drying oven to 60°C
- vegetable steamer:
  - add distilled water in the lower reservoir up to the maximum mark
  - preheating steamer, with plastic beaker with lid, including HIER antigen retrieval buffer for25minutes
  - caution: add additional distilled water in the lower reservoir of the steamer at latest timeafter 50minutes of using

Prepare formalin-fixed, paraffin-embedded tissue sections

- Section the paraffin-embedded tissue block at 1-2 µm thickness on a microtome and float in a40°C water bath containing distilled water.
- Transfer the sections onto glass slides suitable for immunohistochemistry (*e.g*. Superfrost Plus)
- Allow the slides to drying oven at 60°C for one hour

Deparaffinization

- 3x xylene >> 3 minutes each
- 2x 96% ethanol >> 3 minutes each
- 1x 50% ethanol >> 3 minutes
- 1x distilled water >> 3 minutes
- **important notes**: change xylene and ethanol every week and/or after 100 slides

Endogenous peroxidase blocking

- incubating sections in 3% H_2O_2 solution at room temperature for 7 min
- rinse in 2 changes distilled water, 1 min each
- **important notes**: 3% H_2O_2 light-sensitive >> protect from sun and prepare every day fresh 3%H_2O_2-dilution from 30% H_2O_2

Antigen retrieval (pH6) -Sodium citrate pH6.0

- buffer is given from the DAKO HERCEP-Kit and has to be diluted in 1:10 (see 2.3)

- **important notes**: Slides should be placed in a plastic rack and vessel for this procedure. Standard glass histology staining racks and vessels will crack when heated.

vegetable steamer:

- additional information and pictures, how to use steamer you will find on *appendix A*
- add distilled water in the lower reservoir of the steamer up to the maximum mark
- place plastic rack (with lid) with antigen retrieval buffer into steamer and allow to preheat buffer for 25 minutes
- check temperature before adding slides: should be reached 89°C (*)
- add slides into preheated retrieval buffer and incubate for:
    - 30 minutes for PgR and Ki67
    - 40 minutes for ER
- **Important notes**: add additional distilled water in the lower reservoir of the steamer at latest time after 50minutes of using
- Take off container with retrieval buffer including slides out from the steamer and allow to cool down at room temperature for 20 min
- Wash slides with distilled water
- Bring slides into diluted wash buffer

(*) temperature is set because of the high altitude of Addis Abeba. In Germany temperature should reach 95°C before adding slides into retrieval buffer

Antigen-Detection

- Bring slides into wet chamber or Coverplates (see *appendix B*)
- 100µl of diluted primary antibody >> 40 minutes
- Wash 3x with diluted wash buffer
- 100µl of post block (yellow) >> 20 minutes
- Wash 3x with diluted wash buffer
- 100µl of HRP-Polymer (red) >> 30 minutes
- Wash 3x with diluted wash buffer

DAB staining

- Prepare DAB working solution always fresh:
  - Add 4 drops of DAB Chromogen into one bottle of DAB Substrate buffer
  - Mix by shaking carefully
- Take out slide from coverplate to a wet chamber or use extra coverplate, which are only used forDAB
- Add 100µl DAB working solution to the section on the slide
- Incubation add room temperature for 10minutes
- Wash with distilled water
- **Important notes**: DAB working solution is stable for at least 6 hours. DAB is a suspectedcarcinogen. Handle with care. Wear gloves, lab coat and eye protection.

Counterstaining

- Hematoxylin solution, Mayer´s for 30seconds
- Rinse the slides in running tap water for 5 minutes
- Important notes: change Hematoxylin every week

Dehydration and Mounting

- 1x distilled water >> 1 minutes
- 1x 50% ethanol >> 2 minutes
- 2x 96% ethanol >> 2 minutes each
- 3x xylene >> 2 minutes each
- coverslip using mounting solution (*e.g.* Permount)
- mounted slides can be stored at room temperature permanently

Appendix II: Immunohistochemical staining (IHC) for HER2

name	clone	host	order nr.	company	dilution
human epidermal growth factor receptor 2 (HER2)	NA	rabbit	SK001	Dako	rtu

1. Principle of the method

HercepTest™ is a semi-quantitative immunocytochemical assay to determine HER2 protein overexpression in breast cancer tissues. Paraffin-embedded tissue sections are first de-paraffinised and rehydrated. Endogenous peroxidase activity in the tissue may cause non-specific staining. This enzyme activity can be blocked by incubation with 3% H_2O_2-solution (peroxide block). Epitopes of antigens are being masked and often are no longer accessible for primary antibodies, because of the strong cross-links between proteins induced by formalin fixation. Heat induced epitope retrieval (HIER) in buffer solutions of different compositions and pH-values restore structures of the epitopes making them more accessible to specific antibodies. The next step is incubation with the specific primary antibody. After washing, Visualization reagent is applied and incubated. This reagent consists of both secondary goat anti- rabbit immunoglobulin molecules and horseradish peroxidase molecules linked to a common dextranpolymer backbone, thus eliminating the need for sequential application of link antibody and peroxidase-conjugated antibody. Any excess of unbound Visualization reagent is thoroughly washed away after incubation. The addition of the chromogenic substrate starts the enzymatic reaction of the peroxidase which leads to color precipitation where the primary antibody is bound. The color canbe observed with a light microscope. The chromogen used determines the color. The chromogen DAB forms a dark brown precipitate.

2. Material and reagents

<u>reagents provided</u>

- Positive control tissue: HercepTest^{TM} control slides = Each slide contains sections of three formalin-fixed, paraffin-embedded breast carcinoma cell lines representing different levels of HER2 protein expression: MDA-231 (0), MDA-175 (1+), and SK-BR-3 (3+). The Control Slides havebeen heat treated for better adherence of sections to glass slides. Any additional heat treatmentof Control

Slides performed to improve the adherence of sections to glass slides may compromise staining results.

- retrieval buffer: HercepTest™ Epitope Retrieval Solution (Containing Detergent) (10x)
- wash buffer: Zytomed Wash Buffer (20x) (ZUC020)
- primary antibody: HercepTest™ Rabbit Anti-Human HER2 (ready-to-use)
- HRP Polymer-system: HercepTest™ Visualization Reagent = Dextran polymer conjugated withhorseradish peroxidase and affinity-isolated goat anti-rabbit immunoglobulins (ready-to-use)
- DAB Substrate system: HercepTest™ DAB Substrate Buffer + HercepTest™ DAB Chromogen
- peroxidase blocking: HercepTest™ Peroxidase-Blocking Reagent = NOT NEEDED
- HercepTest™ negative control reagent = NOT NEEDED

Materials required but not supplied

- Slides: poly-L-lysine coated slides or silanized slides (e.g. SuperFrost Plus)
- Xylene or suitable substitutes
- Ethanol
- 30% H_2O_2 (hydrogen peroxide) solution
- distilled water
- Counter stain solution: Hematoxylin Mayer´s
- Mounting medium, coverslip

Reagent preparation

- Reagents should be at room temperature when used
- Tissue sections have to be completely covered with the different reagents in order to avoiddrying out
- **Endogenous peroxidase blocking reagent**/ Hydrogen peroxide (H_2O_2):
  - Dilute 30% H_2O_2 into 1:10 with distilled water, to get an 3% H_2O_2-dilution
  - E.g. 20ml of 30% H_2O_2 + 180ml distilled water

- H_2O_2 is light-sensitive >> protect from sun and prepare every day fresh

Epitope retrieval solution:

- Dilute a sufficient quantity of HercepTest™ Epitope Retrieval Solution 1:10 using distilled ordeionized water
- Diluted retrieval solution can be reused tree times
- The diluted working strength solution is stable for about 1 week at room temperature or onemonth at 2-8°C depending on the ambient temperature
- Discard diluted solution if cloudy in appearance

- **wash buffer** preparation:
  - Dilute Wash Buffer concentrate 1:20 with deionized or distilled water and mix thoroughly
  - The diluted working strength solution is stable for about 1 week depending on the ambienttemperature
- **DAB** working solution preparation:
  - Add 25µl of DAB Chromogen per 1ml of DAB Substrate Buffer and mix thoroughly
  - Once the two reagents are combined, the resulting solution can be used for up to 5 dayswhen stored at 2-8°C
  - Excess working solution needs to be disposed as hazardous substance
  - The DAB chromogen is hazardous to your health. Wear protective clothing to avoid contactof reagents or specimen with eye, skin or mucous membrane

3. Procedure

Important points before starting

- preheating drying oven to 60°C
- vegetable steamer:
  - add distilled water in the lower reservoir up to the maximum mark
  - preheating steamer, with plastic beaker with lid, including HIER antigen retrieval buffer for25minutes

- caution: add additional distilled water in the lower reservoir of the steamer at latest time after 50 minutes of using

Prepare formalin-fixed, paraffin-embedded tissue sections

- Section the paraffin-embedded tissue block at 1-2 µm thickness on a microtome and float in a 40°C water bath containing distilled water.
- Transfer the sections onto glass slides suitable for immunohistochemistry (*e.g.* Superfrost Plus)
- Allow the slides to dry in drying oven at 60°C for one hour

Deparaffinization

- 3x xylene >> 3 minutes each
- 2x 96% ethanol >> 3 minutes each
- 1x 50% ethanol >> 3 minutes
- 1x distilled water >> 3 minutes
- **important notes**: change xylene and ethanol every week and/or after 100 slides

Endogenous peroxidase blocking

- incubating sections in 3% H_2O_2 solution at room temperature for 7 min
- rinse in 2 changes distilled water, 1 min each
- **important notes**: 3% H_2O_2 light-sensitive >> protect from sun and prepare every day fresh 3% H_2O_2-dilution from 30% H_2O_2

Antigen retrieval (pH6) -Sodium citrate pH6.0

- buffer has to be diluted in 1:10 (see 2.3)
- **important notes**: Slides should be placed in a plastic rack and vessel for this procedure. Standard glass histology staining racks and vessels will crack when heated.

vegetable steamer:

- additional information and pictures, how to use steamer you will find on *appendix A*
- add distilled water in the lower reservoir of the steamer up to the maximum mark

- place plastic rack (with lid) with antigen retrieval buffer into steamer and allow to preheat bufferfor 25 minutes
- check temperature before adding slides: should be reached 89°C (*)
- add slides into preheated retrieval buffer an incubate for 40 minutes
- **Important notes**: add additional distilled water in the lower reservoir of the steamer at latesttime after 50minutes of using
- Take off container with retrieval buffer including slides out from the steamer and allow to cooldown at room temperature for 20 min
- Wash slides with distilled water
- Bring slides into diluted wash buffer

(*) temperature is set because of the high altitude of Addis Abeba. In Germany temperature should reach 95°C before adding slides into retrieval buffer

Antigen-Detection

- Bring slides into wet chamber or coverplates (see *appendix B*)
- 100µl of primary antibody: HercepTest™ Rabbit Anti-Human HER2
- Incubate at room temperature for 30 minutes
- Wash 3x with diluted wash buffer
- 100µl of HRP Polymer-system: HercepTest™ Visualization Reagent
- Incubate at room temperature for 30 minutes
- Wash 3x with diluted wash buffer

DAB staining

- Prepare DAB working solution:
  - 25µl of DAB Chromogen per 1ml of DAB Substrate Buffer
  - Mix by shaking carefully
- Take out slide from coverplate to a wet chamber or use extra coverplate, which are only used forDAB
- Add 100µl DAB working solution to the section on the slide

- Incubation add room temperature for 10minutes
- Wash with distilled water
- **Important notes**: DAB working solution is stable for at least 5 days when stored at 2-8°C. DAB is asuspected carcinogen. Handle with care. Wear gloves, lab coat and eye protection.

Counterstaining

- Hematoxylin solution, Mayer´s for 30 seconds
- Rinse the slides in running tap water for 5 minutes
- Important notes: change Hematoxylin every week

Dehydration and Mounting

- 1x distilled water >> 1 minutes
- 1x 50% ethanol >> 2 minutes
- 2x 96% ethanol >> 2 minutes each
- 3x xylene >> 2 minutes each
- coverslip using mounting solution (*e.g*. Permount)
- mounted slides can be stored at room temperature permanently

Important notes:

- Start to adding reagents after ALL the wash buffer is run through the coverplates (no remaining buffer in the upper reservoir), otherwise you diluted your reagent in the wrongway
- reuse the coverplates after cleaning it with brush and dishwashing detergent, following byrinsing with distilled water

BENCHTOP PROTOCOL for IHC of HER2

- dry sections in oven at 65°C >> 1 hour

- deparaffinization:
  - 3x xylene >> 3 minutes each
  - 2x 96% ethanol >> 3 minutes each

- 1x 50% ethanol >> 3 minutes
- 1x distilled water >> 3 minutes

- endogenous peroxidase blocking:
  - 3% H_2O_2 >> 7 minutes
  - 2x distilled water >> 1 minutes each

- antigen retrieval (pH6) -Sodium citrate pH6.0
  - diluted retrieval buffer into vegetable steamer
  - preheating >> 25 minutes
  - slides into retrieval buffer >> 40 minutes
  - Take out retrieval buffer container (including slides) and let cool at room temperature

    >> 20 min utes
  - 2x distilled water >> 1 minutes each
  - 1x wash buffer >> 1 minute
- Bring slides into wet chamber or Coverplates

- Antigen detection
  - 100µl of primary antibody >> 30 minutes
  - Wash 3x with diluted wash buffer
  - 100µl of Visualization Reagent >> 30 minutes
  - Wash 3x with diluted wash buffer
- Take out slides from Coverplates

- DAB staining
    - Add 25µl of DAB Chromogen into 1ml of DAB Substrate buffer to produce DAB workingsolution
    - 100µl DAB working solution to slide >> 10 minutes
    - 1x wash buffer >> 1 minute
- Counterstaining
    - Hematoxylin solution, Mayer´s >> 30 seconds
    - Rinse in tap water >> 5 minutes

- Dehydrating and mounting
    - 1x distilled water >> 1 minutes
    - 1x 50% ethanol >> 2 minutes
    - 2x 96% ethanol >> 2 minutes each
    - 3x xylene >> 2 minutes each
    - coverslip using mounting solution

Appendix III: Chromogenic in situ hybridization (CISH) for ERBB2 gene amplifications

Purpose: Qualitative detection of ERBB2

Assay procedure

I. Preparatory steps

(1) Prepare an ethanol series (70%, 90%, and 100% ethanol solutions): Dilute 100% ethanol with deionized or distilled water. These solutions can be stored in suitable containers and can be re-used.

(2) Heat Pretreatment Solution EDTA (PT2): Heat to 98°C in a covered staining jar.

(3) Preparation of 3% H2O2: Dilute 1 part 30% H2O2 in 9 parts 100% methanol.

(4) ZytoDot 2C CISH Probe: Bring to room temperature before use.

II. Pretreatment (dewax/proteolysis)

(1) Incubate slides for 10 min at 70°C (e.g., on a hot plate).

(2) Incubate slides for 2x 5 min in xylene.

(3) Incubate slides for 3x 3 min in 100% ethanol.

(4) Incubate slides for 5 min in 3% H2O2.

(5) Wash slides 2x 1 min in deionized or distilled water.

(6) Incubate for 15 min in pre-warmed Heat Pretreatment Solution EDTA (PT2) at 98°C. Use eight slides per staining jar (add dummy slides if needed).

(7) Transfer slides immediately to deionized or distilled water and wash for 2x 2 min.

(8) Apply (dropwise) Pepsin Solution (ES1) to the specimen and incubate for 5-15 min at 37°C in a humidity chamber. As a general rule, we recommend to ascertain the optimum time for proteolysis in pre-tests.

(9) Immerse slides in deionized or distilled water.

(10) Dehydration in: 70%, 90%, and 100% ethanol, each for 1 min.

(11) Air dry sections. Note: Make sure to completely dry sections prior to probe application.

III. Denaturation and hybridization

(1) Pipette 10 µl of the ZytoDot 2C CISH Probe onto each pretreated specimen.

(2) Cover specimens with a 22 mm x 22 mm coverslip (avoid trapped bubbles) and seal the coverslip. We recommend using rubber cement (e.g., Fixogum) for sealing.

(3) Place slides on a hot plate or hybridizer and denature specimens for 5 min at 79°C.

(4) Transfer the slides to a humidity chamber and hybridize overnight at 37°C (e.g., in a hybridization oven).

IV. Day 2: Preparatory steps

(1) Wash Buffer SSC (WB1): For stringency wash, heat to 80°C in a covered staining jar.

(2) Preparation of 1x Wash Buffer TBS: Dilute 1 part of 20x Wash Buffer TBS (WB5) in 19 parts deionized or distilled water. Diluted 1x Wash Buffer TBS is stable for one week when stored at 2-8°C.

(3) Anti-DIG/DNP-Mix (AB14), HRP/AP-Polymer-Mix (AB13), AP-Red Solution A (SB6a), AP-Red Solution B (SB6b), HRP-Green Solution A (SB7a), HRP-Green Solution B (SB7b) Nuclear Blue Solution (CS2), Mounting Solution (alcoholic) (MT4): Bring to room temperature before use.

V. Post-hybridization and detection

(1) Carefully remove the rubber cement or glue.

(2) Remove the coverslip by submerging the slides in Wash Buffer SSC (WB1) at room temperature for 5 min. WB1 can be reused once. Store at 2-8°C for a maximum of one week.

(3) Wash slides for 5 min in Wash Buffer SSC (WB1) at 80°C. Use eight slides per staining jar (add dummy slides if needed).

(4) Wash slides 2x 1 min in deionized or distilled water.

(5) Immerse slides in 1x Wash Buffer TBS. (6) Apply Anti-DIG/DNP-Mix (AB14) (1-2 drops per slide) to the slides and incubate for 15 min at 37°C in a humidity chamber.

(7) Wash slides 3x 1 min in 1x Wash Buffer TBS.

(8) Apply HRP/AP-Polymer-Mix (AB13) (1-2 drops per slide) to the slides and incubate for 15 min at 37°C in a humidity chamber.

(9) Wash slides 3x 1 min in 1x Wash Buffer TBS.

(10) Prepare AP-Red Solution (working solution): fill 1 ml AP-Red Solution B (SB6b) in a graduated cup and add one drop (30 µl) of AP-Red Solution A (SB6a). Mix well.

(11) Apply AP-Red Solution (1-2 drops per slide) to the slides and incubate for 10 min at RT.

(12) During the incubation, prepare HRP-Green Solution (working solution): fill 1 ml HRP-Green Solution B (SB7b) in a graduated cup and add two drops (2x 20 µl) HRP-Green Solution A (SB7a). Mix well.

(13) Wash slides for 2 min in deionized or distilled water.

(14) Apply HRP-Green Solution dropwise (1-2 drops per slide) to the slides and incubate for 10 min at room temperature.

(15) Wash slides for 2 min in deionized or distilled water.

(16) Counterstain specimens for 2

min with Nuclear Blue Solution (CS2).

(17) Transfer slides into a staining jar and wash 2 min under cold running tap water.

(18) Dehydrate 3x 30 s in 100% ethanol (use very pure ethanol).

(19) Incubate slides for 2x 30 s in xylene (use very pure xylene). Do not prolong or shorten the incubation time as this might result in loss of signals!

(20) Avoiding trapped bubbles, cover the samples with a coverslip (22 mm x 22 mm; 24 mm x 32 mm) by using Mounting Solution (alcoholic) (MT4). Allow 20-30 min for the coverslip to become immobilized.

(21) Evaluate stained specimens by using light microscopy.

VI. Interpretation of results

The hybridization signals of Digoxigenin-labeled polynucleotides appear dark green colored (ERBB2 gene region), and Dinitrophenyl-labeled polynucleotides appear bright red (CEN 17). **Normal situation:** In interphases of normal cells or cells without an amplification involving the ERBB2 gene region, two distinct dot-shaped green signals and two distinct dot-shaped red signals appear.

Aberrant situation: In cells with an amplification of the ERBB2 gene region, an increased number of green signal or green signal clusters will be observed. Overlapping signals may appear as brown signals

Appendix IV: Procedure for purification of RNA from FFPE sections (miRNeasy FFPE Kit – Qiagen)

1. Principle of the method

Xylene dissolve and remove Paraffin. Denaturation buffer and proteinase K lyses tissue. Incubation at 80°C reverses formalin crosslinking and destroy proteinase K. DNase treatment eliminate all genomic DNA, including very small DNA fragments that are often present in FFPE samples after prolonged formalin fixation and/or long storage times. RNA binds under specific buffer condition to silica-based membrane, and contaminants are wash away. RNA is eluted from the membrane by RNase-free water. Purified RNA from FFPE samples has a lower molecular weight than RNA from fresh frozen tissue, because there is a higher fragmentation, depending on type and age of the sample and fixation conditions, improving RNA yield and quality as well as RNA performance in downstream enzymatic assays.

2. Material and reagents

Starting material

- 2- up to 4 sections, each with a thickness of up to 10µm and a surface area of up to 250mm²(use more sections, if surface area is lower than 250mm²) in a 1,5ml-tube

Important points before starting

- Prepare **DNase I stock** solution by dissolving the lyophilized DNase I (1500 K units) in 550 µl RNase-free water, mix by gently inventing the vial, aliquot and store at -20°C for up to9months. Do not refreeze aliquot after thawing.
- If using Buffer **RPE** for the first time, add 44ml of 100% ethanol (tick the check box on thebottle label to indicate that ethanol has been added) and mix by shaking
- Set a thermal mixer, heating block or water bath to 56°C
- If available, set a second thermal mixer, heating block or water bath to 80°C

3. Procedure

deparaffinization

- Add **800µl D-Limonene** OR Xylene
    - Vortex vigorously for 10 s
    - Incubation for 5 minutes at room temperature
    - Centrifuge 2 minutes at full speed
    - Carefully remove the supernatant by pipetting without disturbing the pellet
- Repeat previous step

Add 800µl 96% Ethanol

- Vortex vigorously for 10 s
- Centrifuge 2 minutes at full speed
- Carefully remove the supernatant by pipetting without disturbing the pellet

- Repeat previous step
- Incubate for 5-15 minutes at 55°C (thermal mixer, heating block or water bath) until pellet isdried

lyse

- **150µl buffer PKD + 15µl Proteinase K** and mix by vortexing
- Incubation 1hour at 56°C (Heating-block) and shaking at 1.000rpm
- After one hour:
    - Check if there is still some tissue inside the tube
    - Briefly spin down
    - Add another 7µl Proteinase K to each sample
    - Incubation 90minutes at 56°C (Heating-block) and shaking at 1000rpm

Heat

- Incubation 15minutes at 80°C (Heating-block)

  Caution! If you use one heating-block, take of samples at room temperature untiltemperature of heating-block reaches 80°C, otherwise RNA will be destroyed

- Incubation 3minutes at crushed ice

- Centrifugation 15minutes at full speed
- Pipet supernatant in a new 2ml tube

Caution! don´t destroy pellet at the bottom of tube; pellet can be discarded

proceed miRNeasy FFPE Handbook 01/2020 – step 13

DNase I-digestion

Mastermix per sample: **16µl DNase Booster Buffer + 10µl DNase I stock**

- 26µl of mastermix to each sample
- Mix gently by inverting tube
- Incubation 15minutes at room temperature

Bind

Add **320µl RBC**-buffer

- Mix thoroughly
- Briefly spin down

Add **1120µl 100% Ethanol** (2x 560µl)

- Mix by pipetting up and down

Transfer 700µl of the sample to a RNeasy MinElute spin column

- Centrifuge 30seconds at 8000g (~10000rpm)
- Discard the flow-through
- Reuse the collection tube in the next step

Caution! Storage of spin-column at 4°C

- Repeat last step until the entire sample has passed through the column

wash

Add 500µl RPE-buffer

- Centrifuge 30seconds at 8000g (~10000rpm)
- Discard the flow-through
- Reuse the collection tube in the next step

Add 500µl RPE-buffer

- Centrifuge 2 minutes at 8000g (~10000rpm)
    - Discard collection-tube with flow-through and place column into a new collectiontube

Centrifuge with open lid for 5 minutes at full speed

- Discard the collection tube with the flow-through
- Place column in a new 1,5ml-tube (elution tube, labeled)

Elution

- Add **17µl RNase-free water** directly to the spin column filter

Caution! Don't touch filter directly with the pipet tip

- Incubation 1 minute at room temperature
- Centrifuge 1 minute at full speed
- Replace flow-through to the spin column filter
- Incubation 1 minute at room temperature
- Centrifuge 1 minute at full speed
- Discard column and store RNA at -80°C

Appendix V: Isolation of RNA from fresh frozen tissue

(miRNeasy Mini Kit – Qiagen)

1. **Principle of the method**

Tissue should be sectioned in 5-10µm slices or prepared as powder by laboratory ball mills. Samples are homogenized in QIAzol Lysis reagent by using pestle. This lysis reagent is a monophasic solution ofphenol and guanidine thiocyanate, which lysis the cells, inhibit RNases and remove most of DNA and proteins by organic extraction. After adding chloroform and a centrifugation step, the sample is separated into 3 phases: the upper, aqueous phase contains RNA; the white precipitate/ interphase contains DNA and some proteins; the lower, pinkish, organic phase contains most of the proteins. The upper, aqueous phase is separated into a new tube and ethanol is added, to provide optimal binding conditions for RNeasy Mini spin column. RNA binds to silica-based membrane of the spin columns andcontaminants are washed away by using specific wash buffers. RNA is elute from the membrane by RNase-free water.

2. Material and reagents

Starting material

- Prepare 6-10 sections, each with a thickness of up to 10µm in a 1,5ml-tube; avoid to thaw tissueduring cryogenic sectioning, handling and storing at -80°C

Important points before starting

- It is recommended to do on-column DNase digestion with RNase-free DNase-Set (Cat.No.

  79254) which is not supplied

- Prepare **DNase I stock** solution by dissolving the lyophilized DNase I (1500 K units) in

  550 µl RNase-free water, mix by gently inventing the vial, aliquot and store at -20°C for up to9 months. Do not refreeze aliquot after thawing.

- If using Buffer **RPE** for the first time, add 44ml of 96%-100% ethanol (tick the check box onthe bottle label to indicate that ethanol has been added) and mix by shaking

- If using Buffer **RWT** for the first time, add 30ml of 96%-100% ethanol (tick the check box onthe bottle label to indicate that ethanol has been added) and mix by shaking

3. Procedure

Tissue preparation

- Avoid to thaw tissue during the whole tissue preparation; tissue should be stored at -80°C
- Cleaning cryostat with 70% ethanol after every sample and dry surface with paper
- Determine amount of tissue, for using up to 100mg of it
- Bring tissue into cryostat specimen block by using cryostat glue
    - 10µm sections
    - 60-100 sections into a 1,5ml tube by using pipette tip to avoid to touch sections

lyse

- Add **700µl QIAzol Lysis Reagent** to the frozen tissue and mix by vortexing or homogenize byusing pestle at room temperature

  proceed miRNeasy MiniHandbook 11/2020 -Protocol: Purification of Total RNA, Including Small RNAs, from Animal Tissues – step 4
- Incubation 5min at room temperature
- **140µl chloroform** (chloroform is stored at 4°C)
    - Mix vigorously untill samle become milky and light-pink
    - Incubation 2-3min at room temperature56°C

separation

- Centrifugation **15min** at full speed at **4°C**
- Transfer upper aqueous phase to a new 1,5ml tube and avoid to touching interphase withpipette tip (amount of upper phase should be round about 180 µl)

Add 525µl 100% ethanol

- Mix thoroughly by petting up and down

- Proceed immediately to the next step

bind

- **700µl of sample** into a RNeasy Mini **spin column**
    - Centrifuge 30seconds at 8000g (~10000rpm)
    - Discard the flow-through
    - Reuse the collection tube in the next step
- Repeat last step with the remaining sample

Add 350µl RWT-buffer

- Centrifuge 30seconds at 8000g (~10000rpm)
- Discard the flow-through
- Reuse the collection tube in the next step

DNase I-digestion

- Prepare Mastermix per sample: **70µl RDD Buffer + 10µl DNase I stock**
    - Add 80µl DNase mastermix directly at the filter of the spin column without touchingthe filter
    - Incubation 15minutes at room temperature

Add 350µl RWT-buffer

- Centrifuge 30seconds at 8000g (~10000rpm)
- Discard the flow-through
- Reuse the collection tube in the next step

wash

Add 500µl RPE-buffer

- Centrifuge 30seconds at 8000g (~10000rpm)
- Discard the flow-through
- Reuse the collection tube in the next step

Add 500µl RPE-buffer

- Centrifuge 2 minutes at 8000g (~10000rpm)

- Discard collection-tube with flow-through and place column into a new collectiontube

- Centrifuge with open lid for 1 minutes at full speed
    - Discard the collection tube with the flow-through
    - Place column in a new 1,5ml-tube (elution tube, labeled)

Elution

- Add **20µl RNase-free water** directly to the spin column filter

Caution! Don't touch filter directly with the pipet tip

- Centrifuge 1 minute at 8000g (~10000rpm)

- Add another **20µl RNase-free water** directly to the spin column filter

Caution! Don't touch filter directly with the pipet tip

- Centrifuge 1 minute at 8000g (~10000rpm)

- Discard column and store RNA at -80°C

Appendix VI: Protocol for RNA quality ad concentration checking

Machine: Infinite M200 with computer accessory

- Can measure 16 samples at a time
- First clean the cool plate with distilled water to which the blank and the samples are loaded independently

➢ Micropipette with Tips
➢ distilled –water/RNase free water
➢ water log book

Step-wise procedure

1. Set up the wavelength for measuring the concentration and RNA quality
   - 260/280 just to check the presence of good RNA concentration with the corresponding absorbance ration (should be greater than 1.7 and ration of 2 is best)
   - 260/230 to check the quality of the RNA: should be greater than 1.3(you can see in graph)
   - If there is poor concentration there will be also poor quality with graph
2. Set the blanking by loading 2µl of nuclease free water in to each cool plate loading port, close the plate and put back then run by pressing start
   - ➢ avoid bubble formation which compromise RNA measurement
   - ➢ Write sample ID while the machine blanking
3. Take out the cool plate and add 2µl of the sample and then run the machine: before loading arrange your samples accordingly
4. Automatically you can have the RNA concentration on the excel sheet for each sample
5. The machine will calculate the ratio of the absorbance at 260/280 for each sample and the corresponding concentration in ng/µl
   - The quality can be seen by looking into the graph produced
   - Click standard from the software put the wave length range in between 230-330 in every 5 nm
6. Finally save the file and clean the cool plate with distilled water
7. Store the RNA at -80^{0}C using specific box number with the corresponding number and letter orientation

Appendix VII: Gene Expression CodeSet RNA Hybridization Protocol

GENERAL PROBE HANDLING WARNING: During setup of the reaction, do not vortex or pipette vigorously to mix as it may shear the CodeSet. Mixing should be done by flicking or inverting the tubes. If using a microfuge to spin down tubes, do not spin any faster than 1,000 rpm for more than 30 seconds. Do not "pulse" to spin because that will cause the centrifuge to go to maximum speed and may spin the CodeSet out of solution.

IMPORTANT: Check the reagent labels before you begin to ensure the correct reagents are being utilized. NOTE: A thermal cycler with a heated lid is required for this protocol. NanoString recommends a thermal cycler with a programmable heated lid. Models without programmable lids may reach a high temperature that causes tubes to melt or deform during extended or overnight hybridization times, and if used, should be set to ensure that the heated lid does not exceed 110°C.

The final hybridization reaction will contain the following components: 10µL Reporter CodeSet, 10µL hybridization buffer, a total volume of 5µL of sample RNA (100ng), and 5µL Capture ProbeSet.

1. Pre-heat the thermal cycler to 65°C with a heated lid at 70°C.

NOTE: If using cell lysates, see MAN-10051, Preparing RNA and Lysates from Fresh Frozen Samples. NOTE: If using Panel Plus or CodeSet Plus, refer to the CodeSet RNA Hybridization Protocol with Panel Plus or CodeSet Plus.

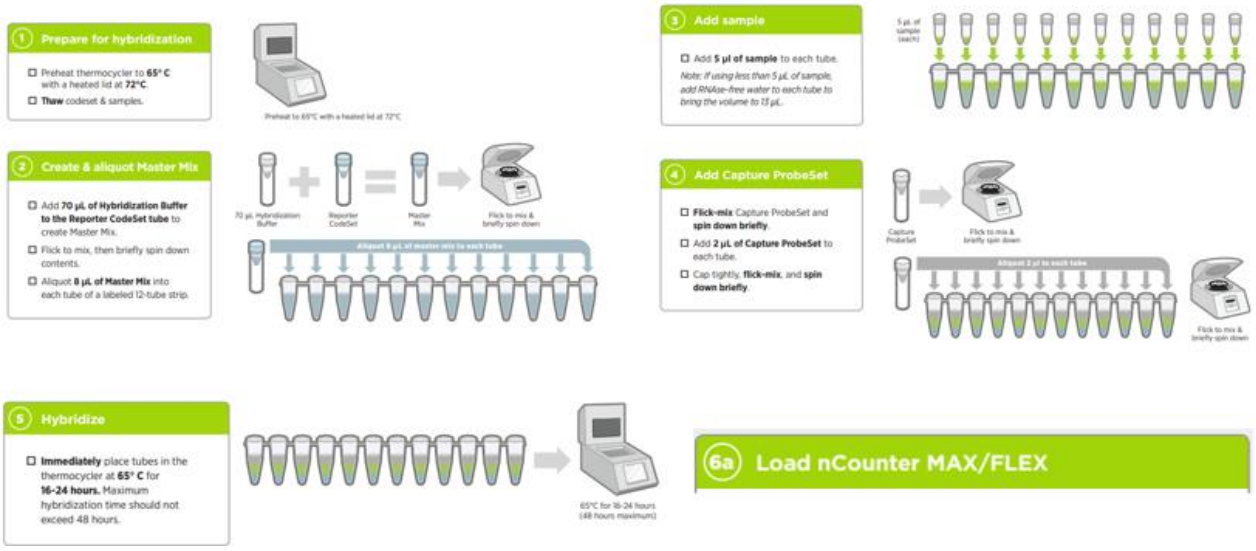
2. Remove Reporter CodeSet and Capture ProbeSet tubes from the freezer and thaw at room temperature. Invert or flick the tube several times to mix well and briefly spin down reagents. IMPORTANT: After it has thawed, inspect the tube of Reporter CodeSet to make sure no colored precipitate is present. If you see a colored precipitate, heat the entire tube to 75°C for 10 minutes and cool at room temperature before using.

3. Create a hybridization Master Mix by adding the following reagents to the tube containing the Reporter CodeSet.

Do not remove the Reporter CodeSet from the tube, add components directly into the CodeSet tube. Do not add the Capture ProbeSet to the master mix.

Component	Master Mix (µL)	Per Reaction (µL)
Reporter CodeSet	In tube (42)	3
Hybridization Buffer	70	5
Total Volume	**112**	**8**

4. Flick or invert the tube repeatedly to mix then briefly spin down the Master Mix.
5. Label a 12-tube PCR hybridization strip. If necessary, ensure the strip will fit in a microfuge or picofuge by cutting both the strip tube and its lid in half prior to setting up the reactions, taking care not to crack the tubes.
6. Prepare hybridization reactions using a fresh tip for pipetting into each well:
   - Add 8 µL of Master Mix to each well of a strip tube.
   - Add 5 µL of sample (or diluted Panel Standard) to each tube containing Master Mix. If you are using a Panel Standard, see the section Panel Standard Usage for details.
   - Mix the Capture ProbeSet tube by inverting or flicking, and briefly spin down the contents.
   - Add 2 µL of Capture ProbeSet to each tube.
   - Cap the strip tubes tightly and mix by inverting the tubes several times and flicking to ensure complete mixing. f. Spin briefly and immediately place the tubes in a pre-heated 65°C thermal cycler.
7. Incubate hybridization reactions for at least 16 hours. Maximum hybridization time should not exceed 48 hours.
8. Incubate at 4°C once desired hybridization time is reached and process the following day. Do not leave the reactions at 4°C for more than 24 hours or increased background may result. NOTE: Counts continue to accumulate with time at 65°C, with total counts typically increasing 5% per hour between 16 and 24 hours. Although a 16-hour incubation is adequate for most purposes, a longer incubation increases sensitivity by increasing counts without significantly increasing background.
9. Once the hybridization reactions have been removed from the thermal cycler, proceed immediately to an nCounter Prep Station or SPRINT as described in the nCounter Analysis System User Manual. Do not store hybridizations at 4°C.



Appendix VIII: DNA purification from FFPE section

1. Principle of the method

Xylene dissolve and remove Paraffin. Denaturation buffer and proteinase K lyses tissue. Incubation at 90°C reverses formalin crosslinking and destroy proteinase K. DNA binds under specific buffer condition to silica-based membrane, and contaminants are wash away. DNA is eluted from the membrane by buffer ATE or water. Purified DNA from FFPE samples has a lower molecular weight than DNA from fresh frozen tissue, because there is a higher fragmentation, depending on type and age of the sample and fixation conditions.

2. Material and reagents

Starting material –

- up to 4 sections, each with a thickness of up to 10µm and a surface area of up to 250mm² (use more sections, if surface area is lower than 250mm²) in a 1,5ml-tube

Important points before starting

- If using Buffer AW1 for the first time, add 25ml of 100% ethanol (tick the check box on the bottle label to indicate that ethanol has been added) and mix by shaking
- If using Buffer AW2 for the first time, add 30ml of 100% ethanol (tick the check box on the bottle label to indicate that ethanol has been added) and mix by shaking
- In case of any precipitation in buffer ATL or AL, dissolve by heating to 70°C with gentle agitation - Set a thermal mixer, heating block or water bath to 56°C
- If available, set a second thermal mixer, heating block or water bath to 90°C

3. Procedure

deparaffinization

- Add 800µl D-Limonene OR Xylene
    - Vortex vigorously for 10 s
    - Incubation for 5 minutes at room temperature
    - Centrifuge 2 minutes at full speed
    - Carefully remove the supernatant by pipetting without disturbing the pellet
- Repeat previous step
- Add 800µl absolute/ 100% Ethanol
    - Vortex vigorously for 10 s
    - Centrifuge 2 minutes at full speed
    - Carefully remove the supernatant by pipetting without disturbing the pellet
- Repeat previous step
- Incubate for 5-15 minutes at 55°C (thermal mixer, heating block or water bath) until pellet is dried

<u>lyse</u>

- 180µl buffer ATL + 20µl Proteinase K and mix by vortexing
- Incubation 1hour at 56°C (Heating-block) and shaking at 800rpm
- After one hour: - Check if there is still some tissue inside the tube
    - Briefly spin down - Add another 7µl Proteinase K to each sample
    - Incubation for one more hour at 56°C (Heating-block) and shaking at 800rpm

<u>Heat</u>

- Incubation 1 hour at 90°C (Heating-block)

Caution! If you use one heating-block, take of samples at room temperature until temperature of heating-block reaches 90°C

- Incubation 5minutes at room temperature
- Briefly centrifuge the tube to remove drops from the inside of the lid
- Transfer 200µl of supernatant in a new 1,5ml-tube, without disturbing pellet; pellet can be discarded

3.1. Bind

- Add 200µl buffer AL to the 200µl supernatant and mix by vortexing

- Add 200µl 100% ethanol and mix by vortexing
- Spin down briefly to remove drops from inside the lid
- Transfer entire lysate to the QIAamp MinElute column
    - 1min at 800rpm centrifugation
    - Place QIAamp MinElute column in a clean 2ml collection tube
    - Discard collection tube containing flow-through

3.2. wash

- Add 500µl AW1-buffer - 1min at 800rpm centrifugation
    - Place QIAamp MinElute column in a clean 2ml collection tube
    - Discard collection tube containing flow-through
- Add 500µl AW2-buffer
    - 1min at 800rpm centrifugation
    - Place QIAamp MinElute column in a clean 2ml collection tube
    - Discard collection tube containing flow-through
- 3min at full speed centrifugation to dry membrane completely
    - Place column in a new 1,5ml-tube (elution tube, labeled)
    - Discard collection tube containing flow-through

3.3. Elution

- Add 30µl ATE-buffer directly to the spin column filter
    - Caution! Don’t touch filter directly with the pipet tip - Incubation 5 minute at room temperature
    - Centrifuge 1 minute at full speed

Appendix IX: DNA extraction from fresh frozen tissue (QIAamp DNA Mini Kit)

1. Cut up to 25 mg of tissue into small pieces, place in a 1.5 ml microcentrifuge tube, and add 180 µl of Buffer ATL
2. Add 20 µl Proteinase K, mix by vertexing, and incubate at 56°C until the tissue is completely lysed.
    - Vortex occasionally during incubation to disperse the sample, or place in a shaking water bath or on a rocking platform
3. Briefly centrifuge the 1.5 ml microcentrifuge tube to remove drops from the inside of the lid.
    - Continue with step 3a, or if RNA-free genomic DNA is required, continue with step 3b.

3.1. Add 200 µl Buffer AL to the sample, mix by pulse-vortexing for 15 s, and incubate at 70°C for 10 min.
    - Briefly centrifuge the 1.5 ml microcentrifuge tube to remove drops from inside the lid

3.2. First add 4 µl RNaseA (100 mg/ml), mix by pulse-vortexing for 15 s, and incubate for 2 min at room temperature.

3.3. Briefly centrifuge the 1.5 ml microcentrifuge tube to remove drops from inside the lid before adding 200 µl Buffer AL to the sample.
    - Mix again by pulse-vortexing for 15 s, and incubate at 70°C for 10 min.
    - Briefly centrifuge the 1.5 ml microcentrifuge tube to remove drops from inside the lid

4. Add 200 µl ethanol (96–100%) to the sample, and mix by pulse-vortexing for 15 s.
    - After mixing, briefly centrifuge the 1.5 ml microcentrifuge tube to remove drops from inside the lid.
5. Carefully apply the mixture from step 4 (including the precipitate) to the QIAamp spin column (in a 2 ml collection tube) without wetting the rim.
    - Close the cap, and centrifuge at 6000 x g (8000 rpm) for 1 min.
    - Place the QIAamp spin column in a clean 2 ml collection tube (provided), and discard the tube containing the filtrate.
6. Carefully open the QIAamp spin column and add 500 µl Buffer AW1 without wetting the rim.
    - Close the cap, and centrifuge at 6000 x g (8000 rpm) for 1 min.

- Place the QIAamp spin column in a clean 2 ml collection tube (provided), and discard the collection tube containing the filtrate

7. Carefully open the QIAamp spin column and add 500 µl Buffer AW2 without wetting the rim.
    - Close the cap and centrifuge at full speed (20,000 x g; 14,000 rpm) for 3 min.
    - Continue directly with step 8, or to eliminate any chance of possible Buffer AW2 carryover, perform step 7a, and then continue with step 8.

    (Optional): Place the QIAamp spin column in a new 2 ml collection tube (not provided) and discard the collection tube containing the filtrate.
    - Centrifuge at full speed for 1 min
8. Place the QIAamp spin column in a clean 1.5 ml microcentrifuge tube (not provided), and discard the collection tube containing the filtrate.
    - Carefully open the QIAamp spin column and add 200 µl Buffer AE or distilled water.
    - Incubate at room temperature for 1 min, and then centrifuge at 6000 x g (8000 rpm) for 1 min.
9. Repeat step 8.

Appendix X: Mutation analysis protocol

(TaqMan® Mutation Detection Assay)

Experiment overview

Two experiment types: Two types of experiments are required for mutation detection analysis:

1. Detection ΔCT cutoff determination. Run the mutant allele assay and the corresponding gene reference assay on wild type gDNA samples that are from the same sample type as the test sample (e.g. gDNA from FFPE tissue samples). Run at least three different wild type samples and three technical replicates of each sample. ΔCT values are calculated for each sample run with a mutant allele assay - gene reference assay pair. The average ΔCT value for all samples is then calculated and is used to derive the detection ΔCT cutoff value for the mutant allele assay.

2. Mutation detection. Run the test sample with a mutant allele assay(s) and corresponding gene reference assay. The ΔCT value for the mutant allele assay - gene reference assay pair is calculated and this value is compared to the predetermined detection ΔCT cutoff value to determine the sample mutation status.

1. Run purified gDNA with one or more mutant allele assays and a corresponding gene reference assay. For each real-time PCR, combine the gDNA with:

• A TaqMan® Mutation Detection Assay. The assay contains two primers and a FAM™ dye-labeled MGB probe to detect a mutant allele or gene reference target. Mutant allele assays also contain an MGB oligonucleotide blocker.

• TaqMan® Genotyping Master Mix. The master mix contains AmpliTaq Gold® DNA Polymerase, UP (Ultra Pure), dNTPs, and buffer required for the qPCR reactions

• (Optional) TaqMan® Mutation Detection IPC Reagent Kit. The kit contains an internal positive control (IPC) template, two primers, and a VIC® dye-labeled TAMRA™ probe. You can add the IPC reagents to any mutation detection assay reaction to distinguish true target negatives from PCR failure or inhibition.

2. Run the reactions on a real-time PCR system, using a universal mutation detection thermal-cycling protocol. After the run, the real-time PCR system analysis software determines the CT values for the mutation detection assays and (optional) IPC reagent reactions.

3. Export the Results or Results Table from the real-time PCR system software as *.csv or *.txt files, then import the files into the Mutation Detector™ Software for downstream analysis.

WORK FLOW

Prepare the samples:

1. Extract, purify and quantitate gDNA: Apply QIAamp® DNA FFPE Tissue Kit; Quantitate the gDNA by measuring the UV absorbance (A260/A280). Be sure that the human gDNA that you use has an A260/A280 ratio >1.7.

Perform the PCR

Sample input range

Sample	Input range	Recommended amount
gDNA from FFPE tissues	10–50 ng	20 ng
gDNA from fresh-frozen tissues	1–50 ng	20 ng
gDNA from cell lines	1–100 ng	20 ng

* The recommended lower limit amount of gDNA input for FFPE tissue samples is higher than that for other types of samples because FFPE tissue samples contain crosslinked, fragmented, and damaged DNA template. The amount of functional template in FFPE tissue gDNA samples is significantly lower than in other types of gDNA samples. The recommended upper limit amount of gDNA input for tissue samples is less than that for cell line samples because PCR inhibitors are typically found in tissue gDNA sample preparations.

*The mutant allele assays can detect as low as 0.1% mutated DNA in a background of wild type DNA.

To detect:

• 0.1% mutation, use >= 20 ng of functional gDNA template in the PCR

• 1% mutation, use >= 2 ng of functional gDNA template in the PCR

Prepare the PCR mix and the PCR plate

Detection ΔCT cutoff determination experiments

In a detection ΔCT cutoff determination experiment, run three or more wild type gDNA samples, and three technical replicates of each sample, with a mutant allele assay(s) and paired

gene reference assay. The same amount of gDNA must be used for each sample. Prepare the PCR mix and the PCR plate as follows:

1. For each sample, calculate the total number of reactions required.

2. Calculate the total volume required each component: volume for 1 reaction × total number of reactions

3. Include 10% excess volume to account for pipetting errors.

Component	Volume for 1 reaction	
	20 µL reaction (96-well plate)	**10 µL reaction (384-well plate)**
TaqMan® Genotyping Master Mix, 2X	10.0 µL	5.0 µL
Prepared gDNA sample†	2.0 – 4.0 µL	1.0 – 2.0 µL‡
(Optional) 50X Exogenous IPC Template DNA	0.4 µL	0.2 µL
(Optional) 10X Exogenous IPC Mix	2.0 µL	1.0 µL
Nuclease-free water	Add water to 18.0 µL	Add water to 9.0 µL
Total volume of super mix	**18.0 µL**	**9.0 µL**
Example: Total volume for 3 technical replicates (includes 10% extra volume for pipetting errors)	**59.4 µL**	**29.7 µL**

† We recommend that you input 20 ng of gDNA; the volume of gDNA sample should not be greater than 20% of the total reaction volume.

‡ For 10-µL reactions, we recommend that you use the same amount of gDNA as for 20-µL reactions to obtain the same sensitivity of mutation detection.

4. Label a 1.5-mL microcentrifuge tube, add all components to the labeled tube, cap the tube, then vortex the tube briefly to mix the components.

5. Centrifuge the tube briefly to spin down the contents and eliminate air bubbles.

6. For each set of technical replicates, transfer aliquots of super mix to a microcentrifuge tube, then add the TaqMan® Mutation Detection Assay (mutant allele or gene reference assay) to each tube.

If the total reaction volume for one reaction is...	Add the following volumes to the microcentrifuge tube...†	
	Super mix	**TaqMan® Mutation Detection Assay (10X)**
20 µL	59.4 µL	6.6 µL
10 µL	29.7 µL	3.3 µL

† The super mix and assay volumes listed are for three technical replicates.

7. Add the appropriate volume of PCR mix to each reaction well of a PCR plate:

PCR plate	Volume of PCR mix per reaction
96-well	20 µL
384-well	10 µL

8. Cover the plate with an optical adhesive film.

9. Centrifuge the plate briefly to spin down the contents and eliminate air bubbles.

Note: PCR, and the mutation detection application in particular, is a highly sensitive experimental method. In order to avoid false positive amplifications due to contamination from amplifications previously performed in your lab, please follow the recommended "PCR Good Laboratory Practices,"

Mutation detection experiments

In a mutation detection experiment, run the test sample with a mutant allele assay(s) and corresponding gene reference assay. Technical replicates are not required. The amount of test sample gDNA used should be similar to the amount of wild type gDNA sample that was used for your detection ΔCT cutoff determination experiments. Prepare the PCR mix and the PCR plate as follows:

1. For each sample, calculate the total number of reactions required.

2. Calculate the total volume required each component: volume for 1 reaction × total number of reactions

3. Include 10% extra volume to compensate for pipetting errors

Component	Volume for 1 reaction	
	20 µL reaction (96-well plate)	10 µL reaction (384-well plate)
TaqMan® Genotyping Master Mix, 2X	10.0 µL	5.0 µL
Prepared gDNA sample†	2.0 – 4.0 µL	1.0 – 2.0 µL‡
(Optional) 50X Exogenous IPC Template DNA	0.4 µL	0.2 µL
(Optional) 10X Exogenous IPC Mix	2.0 µL	1.0 µL
Nuclease-free water	Add water to 18.0 µL	Add water to 9.0 µL
Total volume of super mix	**18.0 µL**	**9.0 µL**

† We recommend that you input 20 ng of gDNA; the volume of gDNA sample should not be greater than 20% of the total reaction volume.
‡ For 10-µL reactions, we recommend that you use the same amount of gDNA as for 20-µL reactions to obtain the same sensitivity of mutation detection.

4. Label a 1.5-mL microcentrifuge tube, add all components to the labeled tube, cap the tube, then vortex the tube briefly to mix the components.

5. Centrifuge the tube briefly to spin down the contents and eliminate air bubbles

6. Add the appropriate volume of super mix to each reaction well of a PCR plate

PCR plate	Volume of super mix per reaction
96-well	18.0 µL
384-well	9.0 µL

Add the appropriate volume of TaqMan® Mutation Detection Assay (mutant allele or gene reference assay) to each reaction well:

PCR plate	Volume of TaqMan® Mutation Detection Assay (10×) per reaction
96-well	2.0 µL
384-well	1.0 µL

8. Cover the plate with an optical adhesive film. 9. Centrifuge the plate briefly to spin down the contents and eliminate air bubble

Set up the plate document or experiment and start the run

In the real-time PCR system software:

1. Select the experiment type: Absolute Quantitation or Quantitation - Standard Curve.

2. For each well that contains a reaction, apply a sample name, assay name, and target or detector name. For downstream analysis with the Mutation Detector™ Software, note the following:

3. Enter sample quantity values in the real-time PCR system software:

a. Select Standard as the task for each well of interest.

b. Enter a numeric value. We recommend that the numeric values you enter are relevant to the ng amount of gDNA or copies of DNA input

4. Set the following thermal-cycling conditions:

• Run mode – Standard

• Sample volume – 10 µL (384-well plates) or 20 µL (96-well plates)

• Thermal-cycling profile – See the table below

Stage	Temp.	Time (mm:ss)	Cycles	Data collection
1	95°C	10:00	1	No
2	92°C	00:15	5	No
	58°C	01:00		No
3	92°C	00:15	40	No
	60°C	01:00		FAM™ or VIC®dye†

† FAM dye is the reporter for TaqMan Mutation Detection Assays; VIC dye is the reporter for the IPC reagents (from the TaqMan Mutation Detection IPC Reagent Kit).

5. Load the reaction plate into the real-time PCR instrument, then start the run.

Analyze the data

Analysis steps

Briefly, the analysis steps are:

1. Analyze the data in the real-time PCR system software, using the following analysis settings:

• Manual CT (threshold cycle): 0.2

• Automatic Baseline: On The real-time PCR system software determines the CT values for the mutation detection assays and (optional) IPC reagent reactions.

2. View the amplification plots and/or CT values for all reaction

3. Export the Results or Results Table from the real-time PCR system software as a *.csv or *.txt file. 4. Import the *.csv or *.txt file(s) into the Mutation Detector™ Software.

The Mutation Detector Software can:

• Calculate detection ΔCT cutoff values

• Determine the presence or absence of a mutation in a sample and

• Quantitate the percent mutation in a sample (when assay calibration values are available

How the Mutation Detector™ Software performs the mutation analysis

Determine the detection ΔCT cutoff value

In a typical detection ΔCT cutoff determination experiment, three or more wild type gDNA samples, and three technical replicates of each sample, are run with a mutant allele assay and paired gene reference assay.

To perform the real-time calculations for the detection ΔCT cutoff values, the Mutation Detector™ Software: 1. Calculates the ΔCT value between amplification reactions for the mutant allele assay (negative control) and the gene reference assay, depending on whether or not a calibration ΔCT value is available for the assay pair, as follows:

ΔCT = CT (mutant allele assay negative control) – CT (gene reference assay)

or

Normalized ΔCT = [CT (mutant allele assay negative control) – CT (gene reference assay)] – Calibration ΔCT

Where, The calibration ΔCT value is the inherent CT difference between the mutant allele assay and the corresponding gene reference assay. These values are currently available for a subset of the mutation detection assays

2. Averages the ΔCT value or normalized ΔCT value from each sample.

3. Calculates the detection ΔCT cutoff value, as follows: Detection ΔCT cutoff = Average ΔCT – (3 × the standard deviation or 2CT, use whichever one is greater)

Note: We recommend adjusting any detection ΔCT cutoff values that are greater than 9.96 to 9.96 ΔCT, which is equivalent to a 0.1% detection cutoff

Determine the mutation status of a sample

The Mutation Detector™ Software determines the mutation status of a sample by first calculating the ΔCT value between amplification reactions for a mutant allele assay and a corresponding gene reference assay, as follows, depending on whether or not a calibration ΔCT value is available for the assay pair. ΔCT = CT(mutant allele assay) – CT(gene reference assay) or Normalized ΔCT = [CT(mutant allele assay) – CT(gene reference assay)] – Calibration ΔCT where: The calibration ΔCT value is the inherent CT difference between the mutant allele assay and the corresponding reference assay. These values are currently available for a subset of the mutation detection assays.

The software then compares the ΔCT value to a predetermined detection ΔCT cutoff value to determine if a mutation is detected within the sample:

If the ΔC_T value is...	Then...
Greater than the detection ΔC_T cutoff value	The software classifies the gDNA sample as **mutation not detected**. The sample is either mutation negative, or below the limit of detection for the TaqMan® Mutation Detection Assays.
Less than the detection ΔC_T cutoff value	The software classifies the gDNA sample as **mutation detected**, and calculates the percent mutation in the sample, if the calibration ΔC_T value for the assay pair is available.

Determine the percent mutation For assay pairs having calibration ΔCT values, quantitative mutation analysis is possible. The Mutation Detector™ Software determines the percent mutation present in a mutation-positive sample based on the normalized ΔCT value between amplification reactions for a mutant allele assay and a corresponding gene reference assay using the following calculation:

%mutation = [1/2 ÷ (1/2 + 1)] × 100%

Note: The software does not calculate a percent mutation for assays that do not have calibration ΔCT values or that have ΔCT results greater than the detection ΔCT cutoff value.

Appendix XI: DNA extraction from fresh frozen tissue (for Microbiota analysis)

(TissueLyser and the DNeasy® Blood & Tissue Kit)

Important points before starting

• If using the DNeasy Blood & Tissue Kit for the first time, read "Important Notes" in the DNeasy Blood & Tissue Handbook. Ensure that you are familiar with operating the TissueLyser.

• All centrifugation steps are carried out at room temperature (15–25°C) in a microcentrifuge.

• Vortexing should be performed by pulse-vortexing for 5–10 s.

• Use 10 mg tissue as starting material, increasing the amount if the protocol works satisfactorily.

Things to do before starting

• Buffer ATL and Buffer AL may form precipitates upon storage. If necessary, warm to 56°C until the precipitates have fully dissolved.

• Buffer AW1 and Buffer AW2 are supplied as concentrates. Before using for the first time, add the appropriate amount of ethanol (96–100%) as indicated on the bottle to obtain a working solution.

• Preheat a thermomixer, shaking water bath, or rocking platform to 56°C for use in step 3.

Procedure

1. Pipet 180 µl Buffer ATL into a 2 ml Safe-Lock microtube.
2. Add one stainless steel bead to each tube. For best results, use 5 mm (mean diameter) stainless steel beads.
3. Add 10 mg tissue to the tube, and assemble the TissueLyser. Use 10 mg tissue as starting material, increasing the amount if the protocol works satisfactorily.
4. Homogenize on the TissueLyser for 20 s at 15 Hz. Do not exceed this time as it may result in DNA shearing.
5. Centrifuge the sample briefly to ensure that all the tissue debris is on the bottom of the tube.
6. Add 20 µl proteinase K to the tube. Note: Add 40 µl proteinase K if using RNAlater stabilized tissues.
7. Incubate for 56°C for 1 h in a shaker incubator. Lysis time varies depending on the sample

processed. In general, samples are lysed in 1 h. If it is more convenient, samples can be lysed overnight; this will not affect them adversely.

8. Centrifuge the 2 ml microtube briefly to remove drops from inside the lid. Optional: If RNA-free genomic DNA is required, add 4 µl RNase (100 mg/ml) and incubate for 5 min at room temperature before continuing with step 9.
9. Add 200 µl Buffer AL to the sample (220 µl Buffer AL if using RNAlater stabilized tissues), and mix thoroughly by vortexing.
10. Add 200 µl ethanol (96–100%), and mix again thoroughly by vortexing. It is essential that the sample, Buffer AL, and ethanol are mixed immediately and thoroughly by vortexing or pipetting to yield a homogeneous solution. Buffer AL and ethanol can be premixed and added together in one step to save time when processing multiple samples. A white precipitate may form on addition of Buffer AL and ethanol. This precipitate does not interfere with the DNeasy procedure. Some sample types may form a gelatinous lysate after addition of Buffer AL and ethanol. In this case, vigorously shaking or vortexing the preparation is recommended.
11. Pipet the mixture from step 9 (including any precipitate) into the DNeasy Mini spin column placed in a 2 ml collection tube (provided). Centrifuge at ≥6000 x g (8000 rpm) for 1 min. Discard flow-through and collection tube.
12. *Place the DNeasy Mini spin column in a new 2 ml collection tube (provided), add 500 µl Buffer AW1, and centrifuge for 1 min at ≥6000 x g (8000 rpm). Discard flowthrough and collection tube.
13. *Place the DNeasy Mini spin column in a new 2 ml collection tube (provided), add 500 µl Buffer AW2, and centrifuge for 3 min at 20,000 x g (14,000 rpm) to dry the DNeasy membrane. Discard flow-through and collection tube. It is important to dry the membrane of the DNeasy Mini spin column, since residual ethanol may interfere with subsequent reactions. This centrifugation step ensures that no residual ethanol will be carried over during the following elution. Following the centrifugation step, remove the DNeasy Mini spin column carefully so that the column does not come into contact with the flow-through, since this will result in carryover of ethanol. If carryover of ethanol occurs, empty the collection tube, then reuse it in another centrifugation for 1 min at 20,000 x g (14,000 rpm).
14. Place the DNeasy Mini spin column in a clean 1.5 ml or 2 ml microcentrifuge tube (not provided), and pipet 200 µl Buffer AE directly onto the DNeasy membrane.

15. Incubate at room temperature for 1 min, and then centrifuge for 1 min at ≥6000 x g (8000 rpm) to elute. Elution with 100 µl (instead of 200 µl) increases the final DNA concentration in the eluate, but also decreases the overall DNA yield (see DNeasy Blood & Tissue Handbook).

16. Recommended: For maximum DNA yield, repeat elution once as described in step 13. This step leads to increased overall DNA yield. A new microcentrifuge tube can be used for the second elution step to prevent dilution of the first eluate. Alternatively, to combine the eluates, the microcentrifuge tube from step 13 can be reused for the second elution step. Note: Do not elute more than 200 µl into a 1.5 ml microcentrifuge tube because the DNeasy Mini spin column will come into contact with the eluate.

Appendix XII: 16S rRNA PCR amplification and sequencing (Illumina MiSeq)

***16S Prep Workflow*:** Initial Set up

1. Reconstitute indexed primers to 100 µM.
2. Prepare 100µl 10 µM aliquots of indexed primers.
3. Array aliquots into four 96 well plates.
   Use the following scheme:
   a. A701 – A712 with A501 – A508
   b. A701 – A712 with B501 – B508
   c. B701 – B712 with B501 – B508
   d. B701 – B712 with A501 – A508

4. Extract template DNA and array in 96 well format leaving two wells open. (One for a negative water control and another for the positive Mock Community control)

5. Using Illumina Experiment Manager, create a sample plate for each 96 well plate of template. Choose indexes that correspond to one of the four index pair plates above. See Appendix C for instruction on creating a custom assay in IEM.

6. Using Illumina Experiment Manager, create a sample sheet for the run. Ensure that index choices are compatible with one another and there is sufficient diversity in the index reads so as to activate both light channels every cycle.

7. The MiSeq requires base diversity on every cycle. 16S is a low diversity library. With MiSeq software v2.2, 16S libraries can be loaded with 5% PhiX. Additionally, other high diversity samples such as metagenomes can be run simultaneously. This requires manually editing the sample sheet. Older software versions required "hardcoding" the matrix and phasing/pre-phasing values. See Appendix B.

PCR

Note: These steps may be performed using an epMotion or similar automated pipetting system.
1. Dispense 17 µl of Accuprime Pfx Supermix into each well of a new 96 well plate. 2. Using a multichannel pipette, transfer 1 µl of template DNA per well to the corresponding well on the PCR plate.
3. Using a multichannel pipette, transfer 2 µl of each paired set of index primers to the corresponding well on the PCR plate. Be sure to follow the layout chosen in the sample sheet.

4. Add 1 µl of PCR grade dH2O to the negative control well, and 1 µl of Mock Community at a 1:3 dilution to the positive control well.

5. Repeat for up to four 96 well plates. Seal plates, vortex briefly and spin down contents.

6. Place in thermocycler.

PCR Program

Use the following program:

95C 2:00

-----30 cycles--------

95C 00:20

55C 00:15

72C 5:00

72C 10:00

4C end

Gel Electrophoresis

1. A random row of 12 should be selected from each PCR plate and run on a gel to confirm success of the PCR.
2. 2. Use 2 µl of sample, 4 µl of loading dye in a 1% agarose gel.
3. Run at 100v for 30 minutes alongside a standard ladder.
4. Photograph gel under UV. Check to be sure there is a band for every well.

Cleanup, Normalization, and Pooling

Use the SequalPrep Normalization Plate Kit

1. Transfer 18 µl of PCR product from PCR plate to corresponding well on the normalization plate.

2. Add 18 µl of Binding Buffer. Mix by pipetting, sealing, vortexing, and spinning briefly.

3. Incubate at room temperature for 60 minutes. Note: can incubate overnight if needed. Extra time does not improve results.

4. Aspirate the liquid from the wells. Do not scrape the sides.

5. Add 50 µl of Wash Buffer and pipette up and down twice, then aspirate immediately. Ensure there is no residual wash buffer in any wells.

6. Add 20 µl of Elution Buffer. Mix by pipetting up and down 5 times. Seal, vortex, and spin briefly.

7. Incubate at room temperature for 5 minutes.

8. Create a pool from each plate. Take 5 µl of each well to pool. The use of an empty 96 well plate may facilitate the use of multichannel pipettes.

9. Freeze the remaining sample for later use.

Library QC

1. Prepare the following dilutions of each pooled library in PCR grade H2O:

a. 1:1

b. 1:10

c. 1:1000 (dilute in several steps for better results)

d. 1:2000

e. 1:4000

2. Agilent Bioanalyzer Trace

a. Prepare Gel-Dye mix if not already prepared.

b. Let reagents equilibrate to room temperature.

c. Turn Bioanalyzer on and load 350 µl of dH2O onto electrode cleanser and place in analyzer for 5 minutes.

d. Open a high sensitivity chip and place on the priming station. Base plate should be a position "C" and syringe clip should be at lowest position.

e. Load 9.0 µl of gel-dye mix to position 12 market with a large "G". Ensure the syringe plunger is at 1.0 ml and close the station. Press plunger until it is held by clip. f. Wait for exactly 60 seconds then release the plunger clip. Wait an additional 5 seconds, then slowly pull the plunger back to the 1.0 ml position.

g. Open the priming station. Pipette 9.0 µl of gel-dye mix into the other wells marked "G" in positions 4,8,and 16.

h. Pipette 5.0 µl of marker to all wells excluding the right column. (No marker positions 4,8,12, and 16)

i. Load 1 µl of ladder into position 15 marked by the ladder symbol.

j. Pipette 1 µl of each of dilutions a – b above.

Top row Plate 1 Pool 1:1 x 1, 1:10 x 2. Second row Plate 2 pool 1:1 x 1, 1:10 x 2. Third row Plate 3 Pool 1:1 x 1, 1:10 x 2. Bottom row Plate 4 Pool 1:1 x 1, 1:10 x 2.

k. Place chip in the designated vortex for 1 minute, then transfer chip to the Bioanalyzer.

l. Open the 2100 Expert software and select the HS DNA Assay. Enter sample names/dilutions for each of the test wells. Click Start.

m. Print .pdf when run finishes.

3. Kapa Q-PCR Library Quantification

a. Before Q-PCR reaction setup, add 1 ml Primer Premix (10X) to the 5 ml bottle of KAPA SYBR® FAST Q-PCR Master Mix (2X) and mix by vortexing for 10 sec. Record the date of Primer Premix addition on the KAPA SYBR® FAST Q-PCR Master Mix bottle.

b. Reaction can be either 10 µl or 20 µl. A 10 µl reaction volume is recommended.

c. Prepare a 96 well Q-PCR plate compatible with the real time thermocycler. There are six standards. Each should be run in triplicate. Each pool at each dilution should be run in triplicate.

d. For 10 µl reaction volume dispense 6 µl of master mix into each well needed.

e. Pipette 4 µl of standards and library dilutions into appropriate wells. Mix by pipetting. Vortex and spin optional.

f. Place plate in thermocycler. Start control software

g. Program the following cycle

i. Initial Activation 95 °C 5 minutes

ii. 35 cycles

1. Denaturation 95 °C 30 seconds

2. Annealing 60 °C 45 seconds

3. If library fragment size exceeds 700bp, extend annealing step to 90 seconds.

iii. Perform melt curve to check for primer/adaptor dimer

h. Assign wells and group replicates.

i. Enter values for standards

i. Std. 1 20pM

ii. Std. 2 2pM

iii. Std. 3 0.2pM

iv. Std. 4 0.02pM

v. Std. 5 0.002pM

vi. Std. 6 0.0002pM

vii. Note: The concentrations provided here are for the DNA Standards as supplied in the kit, and are NOT the concentrations in the reactions. Provided that the volume of template added to each reaction is the same for Standards and for library samples (i.e. 4 µl in each case), there is no need to account for these volumes when calculating the concentrations of library samples, nor should one need to calculate the concentration of template in the reaction.

j. Run program

k. To calculate library concentration use the following formula:

i. Average x (452/Avg fragment length from bioanalyzer) x dilution factor

ii. Use the average of the triplicate data points corresponding to the most concentrated library DNA dilution that falls within the dynamic range of the DNA Standards to calculate the concentration of the undiluted library.

iii. Do not include outliers in calculation. If there is more than one outlier in a group, the assay must be repeated.

4. Create normalized pools from each plate by diluting to the concentration of the least concentrated plate. Create a single final pool by adding equal amounts of each post qpcr normalized pool. Final pool must be >10µl in total volume. 40-80µl is ideal.

Sequencing

1. Remove a 500 cycle reagent cartridge from the -20 °C freezer. Place in room temperature water bath for one hour. Place HT1 buffer tube in 4 °C fridge.

2. Copy sample sheet to sample sheet folder on MiSeq. Rename sample sheet to match barcode of reagent cartridge.

3. When the reagent cartridge has thawed, dry bottom with paper towel. Invert the cartridge repeatedly to check each well is thawed. This also serves to mix the reagents. Place in hood.

4. Thaw library, PhiX, and sequencing primers. Check to make sure HT1 is thawed.

5. Place 3 µl of the Read 1 Sequencing Primer(s) into a clean PCR tube. Repeat in separate tubes for the Index Primer(s) and Read 2 Sequencing Primer(s).

6. Using a 1000 µl pipette tip, break the foil over wells 12, 13, 14, and 17.

7. Use an extra-long 100 µl tip with the pipettor set on 75 µl to transfer the 3 µl of Read 1 Sequencing Primer to the bottom of well 12 and pipette 10X to mix. Repeat this process spiking the Index Primer into well 13 and the Read 2 Sequencing Primer into well 14.

8. Prepare a fresh dilution of 0.2N NaOH.

9. To a 1.5ml tube add 10 µl of library, and 10 µl of 0.2N NaOH. To a separate tube add 2 µl PhiX, 3 µl PCR grade water, and 5 µl of 0.2N NaOH. Vortex both tubes to mix and spin for 1 minute at 400rcf. Note: NaOH concentration on the flow cell must remain under 0.001N. Adjusting the concentration of the NaOH used to denature the DNA to 0.1N may be necessary.

10. Allow the tubes to incubate at room temperature for 5 minutes. Immediately add 980 µl of HT1 to the library tube, and 990 µl HT1 to the PhiX tube.

11. Use HT1 to dilute both the library and PhiX to 10pM. For a 5% PhiX run, combine 950 µl of 3.5pM Library and 50 µl PhiX in a final tube. Vortex. Load 600 µl of this solution into well 17 on the reagent cartridge. See example below:

a. (1.45 nM library x 10 µl) + (0.2N NaOH x 10 µl) + 980 µl HT1 = 14.5pM Lib, 0.002N NaOH

b. (14.5pM lib x 241.38 µl) + 758.62 µl HT1 = 3.5pM lib, 0.00048N NaOH

c. [(10nM PhiX x 2 µl) + 3 µl H2O] + (0.2N NaOH x 5 µl) + 990 µl HT1 = 20pM PhiX, 0.001N NaOH

d. (20pM PhiX x 175 µl) + 825 µl HT1 = 3.5pM PhiX, 0.0000175N NaOH

e. (3.5pM Lib x 950 µl) + (3.5M PhiX x 50 µl) = solution loaded f. Solution loaded is 3.5pM overall with a 3.325pM Library concentration, 0.175pM PhiX concentration, and 0.0000457N NaOH

13. Set reagent cartridge aside. Unbox flow cell and PR2 bottle.

14. Thoroughly rinse the flow cell with Milli-Q water. Carefully dry by blotting with lint free wipes (Kimwipes). Give special attention to the edges and points of intersection between the glass and plastic.

15. Wet a new wipe with 100% alcohol and wipe the glass on both sides avoiding the rubber intake ports.

16. Visually inspect the flow cell to ensure there are no blemishes, particles, or fibers on the glass.

17. Follow on screen instructions and load the flow cell, reagent cartridge, and PR2 bottle. Empty and replace the waste bottle.

18. Ensure the machine recognizes the correct sample sheet and the run parameters are correct.

19. Wait for the MiSeq to perform its pre-run checks, and press start.

Run Monitoring

1. The run should be monitored periodically using Illumina Sequence Analysis Viewer.

2. Ideal parameters for a 95% 16S run:

a. Cluster density 700-800k/mm2

b. >85% clusters passing filter

c. 5% aligned (amount of PhiX)

d. No spikes in corrected intensity plot

e. All indices identified following index reads

f. Final >Q30 score of >70% Final Steps 1. Perform post run wash.

2. Dispose of liquid waste in appropriate hazardous jug and reagent cartridge in hazardous bucket.

3. When MiSeq Reporter finishes, copy the fastq files from the analysis folder to the run folder on the NAS drive.

4. Perform maintenance wash if r**equired**

Appendix XIII: Ethical approval and MTA

Department of Microbiology, Immunology andParasitology (DMIP)

Department Research Ethics Review Committee (DRERC)

Meeting No: DRERC/03/2020 Date: 29 May 2020

Protocol Title: Evaluation of Distinct Microbiota Signatures and Exploration of GeneExpression Profile among Breast Cancer Patients in Ethiopia

Principal Investigator	Zelalem Desalegn	
Institute/Department	CHS-AAU/DMIP	
Type of review	[X] Initial Review:	[ ] Amendment [ ] Other (specify):_____
Elements Reviewed	[ ] Attached	[ ] Not attached
Decision of the meeting	[ ] Approved	[ ] Approved with Recommendation
	[ ] Revision requested	[ ] Disapproved
Action Required	[X] Send to IRB	[ ] Authorize Implementaion

Obligations of the PI:

i. Should comply with the standard international and national scientific and ethical guidelines
ii. All amendments and changes made in protocol and consent form needs DREC approval
iii. The PI should report Serious Adverse Events (SAE) within 10 days of the event
iv. End of the study, including thesis work and manuscript should be reported to the DREC

Follow up report expected in:

3 Months_____ 6 Months __X___ 9 Months _____ one year _____

Asrat Hailu (Prof)

Chair, DRERC

Signature

Date: 29/05/2020

Department Stamp

ADDIS ABABA UNIVERSITY, COLLEGE OF HEALTH SCIENCES (IRB)
አዲስ አበባ ዩኒቨርስቲ፡ ጤናሳይንስ ኮሌጅ
Institutional Review Board

ANNEX 3
Form AAUMF 03-008

IRB's Decision

Meeting No: 08/2020 Meeting Date: August 26,2020
Protocol number: 078/20/DMIP

Protocol Title: Evaluation of Distinct Microbiota Signatures and Exploration of Gene Expression Profile among Breast Cancer Patients in Ethiopia		
Principal Investigator:	Zelalem Desalegn	
Institute:	College of Health Sciences, AAU	
Elements Reviewed (AAUMF 01-008)	☑ Attached ☐ Not attached	
Review of Revised Application ☐ Yes ☐ No	Date of Previous review:	
Decision of the meeting:	☑Approved ☐Approved with Recommendation ☐Resubmission ☐ Disapproved	

I. Elements approved-
1. Protocol Version No: 2
2. Protocol Version Date:
3. Informed consent Version No. 2
4. Informed Consent Version Date:

II. Obligations of the PI-
1. Should comply with the standard international & national scientific and ethical guidelines
2. All amendments and changes made in protocol and consent form needs IRB approval
3. The PI should report SAE within 10 days of the event
4. End of the study, including manuscripts and thesis works should be reported to the IRB
5. The PI should report non-compliance and unanticipated events

III. TO NERC ☐

Institution Review Board (IRB) Approval: Period from: November 25, 2020 to November 24, 2021
Follow up report expected in
3 Months ___ 6 months ___ 9 months ✔ one year ____

Chairperson, IRB
Dr. Adamu Addissie
Signature
Date: 25/11/2020

Ref.No. MoSHE//RD/

Date: 13 JAN 2020

Addis Ababa University, College of Health Science

Ethiopia,

Subject: Letter of renewal

We are writing this letter in reference to your renewal request letter wth Ref.No CHS/RTTD/221/2019 dated on December 17/2019.

After having in depth review of your request on ""**The Genetic Land Scope of Breast Cancer and its Interaction with Host Immunity in Ethiopian Breast Cancer Patients**"" by Dr.Tamrat Abebe. The National Research Ethics Review Committee has accepted your renewal request for **one year** from (**January 13, 2020 to January 12,2021).** This is, therefore, to notify that the ethical approval is renewed and your group can proceed in accordance to the latest approved document. Please ensure that you submit a biannual report and an annual renewal application 30 days prior to expire date. We are confident that you as PI of the project and your esteemed organization will monitor the ethical implication of the project as it is stipulated in the latest approved document.

Sincerely

Hiwot Kassu Gizaw (Professor)
State Minister

Cc.

- Office of HE the State Minister
- DG for Science & Research Affairs
- NRERC Secretareiate
  MoSHE
- Dr.Tamrat Abebe (PI)
  AAUCHS

ADDIS ABABA UNIVERSITY, COLLEGE OF HEALTH SCIENCES (IRB)

አዲስ አበባ ዩኒቨርስቲ ጤና ሳይንስ ኮሌጅ

Institutional Review Board

ANNEX 3

Form AAUMF 03-008

IRB's Decision

Meeting No: 01/2018 Date: February 13,2018

Protocol number: 092/17/DMIP

Protocol Title: The genetic land scope of Breast cancer and its interaction with host immunity in Ethiopian Breast cancer patients		
Principal Investigator:	Dr. Tamrat Abebe et al	
Institute:	College of Health Sciences, AAU	
Elements Reviewed (AAUMF 01-008)	☑ Attached	☐ Not attached
Review of Revised Application ☐ Yes ☐ No	Date of Previous review:	
Decision of the meeting:	☑ Approved ☐ Approved with Recommendation ☐ Resubmission ☐ Disapproved	

I. Elements approved-
1. Protocol Version No: 02
2. Protocol Version Date: 15/01/2018
3. Informed consent Version No: 02
4. Informed Consent Version Date: 15/01/2018

II. Obligations of the PI-
1. Should comply with the standard international & national scientific and ethical guidelines
2. All amendments and changes made in protocol and consent form needs IRB approval
3. The PI should report SAE within 10 days of the event
4. End of the study, including manuscripts and thesis works should be reported to the IRB
5. The PI should report non-compliance and unanticipated events

III. TO NERC ☒

Institution Review Board (IRB) Approval: Period from: December 16, 2019 to December 15, 2020
Follow up report expected in: 3 Months____ 6 Months__X__ 9 Months _ __One year____

Chairperson, IRB
Dr. Adamu Addissie

Signature
Date: 16/12/2019

Institutional Review Board Office (IRB) Faculty of Medicine ADDIS ABABA UNIVERSITY

Director of Research & Technology Transfer, CHS
Dr. Wondwossen Amogne

Signature
Date 25 Dec 19

Material Transfer Agreement

This Material Transfer Agreement (MTA) has been prepared for use by Addis Ababa University, Department of Microbiology, Immunology and Parasitology, Ethiopia and Klinik und Poliklinik für Gynäkologie und Institut für Med. Epidemiologie, Biometrie u. Informatik Medizinische Fakultät Martin-Luther-Universität Halle-Wittenberg Universitätsklinikum Halle (Saale), Germany in all transfer of research material (samples, derivatives, and specimens) related to the protocol **"The genetic landscape of breast cancer and its interaction with host immunity in Ethiopian breast cancer patients"**

Provider: Dr. Tamrat Abebe, Addis Ababa University, Department of Microbiology Immunology and Parasitology, Ethiopia

Recipient: Klinik und Poliklinik für Gynäkologie und Institut für Med. Epidemiologie, Biometrie u. Informatik Medizinische Fakultät Martin-Luther-Universität Halle-Wittenberg Universitätsklinikum Halle (Saale)

1. Provider agrees to transfer to recipient's designated (provider) the following research materials /specimen. Formalin-fixed paraffine embedded tumor block, tumor biopsy DNA, RNA, Plasma, serum and PBMC.

The research material will only be used for research purposes as described in the protocol by recipient's investigator in designated laboratory for the research project described below, under suitable containment conditions. This research material will not be used for commercial purposes such as screening, production or sale for which a commercialization license may be required. Recipient agrees to comply with all National and International guidelines rules and regulations applicable to the Research Project and the handling of the Research Material.

a) Are the research materials of human origin?

Yes No [√] [ ]

b) If yes, are they collected according to the details in the protocol and in adherence to National Research Ethics Review Committee (NRERC) and Addis Ababa University College of Health Sciences (AAU-CHS)Ethics Review Committee recommendations and their approval.

Yes No [√] [ ]

1

2. This research material and its derivatives will be used by recipient's investigator solely in connection with the following research project ("Research Project") described with specificity as follows '' **"The genetic landscape of breast cancer and its interaction with host immunity in Ethiopian breast cancer patients"**

3. In all presentations or written publications concerning the research project, recipient will acknowledge provider's contribution of this research material unless requested otherwise.
4. This research material represents a significant contribution on the part of provider and is considered proprietary to provider. Recipient therefore agrees to retain control over this research material and further agrees not to transfer the research material to other people not under her/his direct supervision without advance written approval of provider. The research material will be disposed of as agreed upon per protocol at the end of completion of the project on 2018.
5. The provider does not take any responsibility for loss, damage, wastage or spoilage of the research material during or after shipment to the address provided by the recipient under conditions agreed to in the protocol on shipment of the samples. This research material is provided as a service to the research community. IT IS BEING SUPPLIED TO RECIPIENT WITH NO WARRANTIES, EXPRESS OR IMPLIED, INCLUDING ANY WARRANTY OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Provider makes no representations that the use of the research material will not infringe any patent or proprietary right of third parties.
6. The recipient shall notify the provider in witting of any intention, improvement, modification discovery or development to the material or the information made by recipient or parties, collaborating with recipient, here in after referred to as "invention". Nothing in this agreement shall, however, be construed as conveying to the provider any rights under any patents or other intellectual property to such invention, other than as explicitly provided herein. At its option the provider shall be entitled to receive sample of any materials derived from the Materials for its own research and evaluation purposes only.
7. The under- signed provider and recipient expressly certify and affirm that the contents of any statements made herein are truthful and accurate.
8. Any additional terms (use an attached page if necessary):
9. The provider maintains, ownership right of the research material and its unmodified derivatives unless stated otherwise.

The provider will retain a copy (aliquot) of every sample sent abroad as much as possible for local research needs.

2

Material Transfer Agreement

Signature page

For Recipient:

Read and Acknowledged by:

Recipient's Investigator Duly Authorized Signature

Ove Kandelhardt

Signature/ Stamp

Martin-Luther-Universität
Halle-Wittenberg
Institut für Medizinische Epidemiologie,
Biometrie und Informatik
06097 Halle (Saale)

Date: 15.1.18

Mailing Address for Material: Mailing Address for Notices:

For Provider:

Provider's Investigator Duly Authorized Signature

Dr. Tamrat Abebe

Signature/ Stamp

Date: 24-1-18

Mailing Address: Mailing Address for Notices:

Addis Ababa University Addis Ababa University
Department of Microbiology, Immunology and Parasitology
School of Medicine, College of Health
TikurAnbessa Specialized Hospital
Second floor room number 76
P.o.Box 9086, Addis Ababa, Ethiopia

Tel: +251911447227

Fax: +251- 1- 551 30 99Fax: +251- 1- 551 30 99

3

Published Paper One

Breast Cancer Research and Treatment (2022) 196:495–504
https://doi.org/10.1007/s10549-022-06769-z

PRECLINICAL STUDY

Intrinsic subtypes in Ethiopian breast cancer patient

Zelalem Desalegn[1,2] · Meron Yohannes[3] · Martin Porsch[4] · Kathrin Stückrath[5] · Endale Anberber[6] · Pablo Santos[7] · Marcus Bauer[2,8] · Adamu Addissie[9] · Yonas Bekuretsion[10] · Mathewos Assefa[11] · Yasin Worku[12] · Lesley Taylor[13] · Tamrat Abebe[1] · Eva Johanna Kantelhardt[5,7] · Martina Vetter[5]

Received: 9 August 2022 / Accepted: 6 October 2022 / Published online: 25 October 2022

Abstract

Purpose The recent development of multi-gene assays for gene expression profiling has contributed significantly to the understanding of the clinically and biologically heterogeneous breast cancer (BC) disease. PAM50 is one of these assays used to stratify BC patients and individualize treatment. The present study was conducted to characterize PAM50-based intrinsic subtypes among Ethiopian BC patients.

Patients and methods Formalin-fixed paraffin-embedded tissues were collected from 334 BC patients who attended five different Ethiopian health facilities. All samples were assessed using the PAM50 algorithm for intrinsic subtyping.

Results The tumor samples were classified into PAM50 intrinsic subtypes as follows: 104 samples (31.1%) were luminal A, 91 samples (27.2%) were luminal B, 62 samples (18.6%) were HER2-enriched and 77 samples (23.1%) were basal-like. The intrinsic subtypes were found to be associated with clinical and histopathological parameters such as steroid hormone receptor status, HER2 status, Ki-67 proliferation index and tumor differentiation, but not with age, tumor size or histological type. An immunohistochemistry-based classification of tumors (IHC groups) was found to correlate with intrinsic subtypes.

Conclusion The distribution of the intrinsic subtypes confirms previous immunohistochemistry-based studies from Ethiopia showing potentially endocrine-sensitive tumors in more than half of the patients. Health workers in primary or secondary level health care facilities can be trained to offer endocrine therapy to improve breast cancer care. Additionally, the findings indicate that PAM50-based classification offers a robust method for the molecular classification of tumors in the Ethiopian context.

Keywords Ethiopia · Breast cancer · PAM50 · Intrinsic subtyping

✉ Eva Johanna Kantelhardt
eva.kantelhardt@uk-halle.de

1 Department of Microbiology, Immunology, and Parasitology, School of Medicine, Addis Ababa University, Addis Ababa, Ethiopia

2 Global Health Working Group, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany

3 School of Medical Laboratory Sciences, Addis Ababa University, Addis Ababa, Ethiopia

4 Institute of Computer Science, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany

5 Department of Gynecology, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany

6 Department of Surgery, School of Medicine, Addis Ababa University, Addis Ababa, Ethiopia

7 Institute of Medical Epidemiology, Biostatistics and Informatics, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany

8 Institute of Pathology, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany

9 School of Public Health, College of Health Sciences, Addis Ababa University, Addis Ababa, Ethiopia

10 Department of Pathology, School of Medicine, Addis Ababa University, Addis Ababa, Ethiopia

11 Department of Oncology, School of Medicine, Addis Ababa University, Addis Ababa, Ethiopia

12 College of Medicine and Health Science, Wollo University, Dessie, Wollo, Ethiopia

13 City of Hope National Medical Center, Duarte, CA, USA

Introduction

Breast cancer (BC) is one of the most commonly diagnosed cancer worldwide with an estimated 2.3 million new cases in 2020. Improved screening, systemic drug treatment such as endocrine therapy, chemotherapy, targeted therapy, immunotherapy, and surgery, radiotherapy have contributed significantly in reducing mortality [1]. In the resource-limited settings, BC incidence and mortality tend to increase [2]. In the context of Sub-Saharan Africa, access to diagnosis and cancer care have been identified as important determinants for survival [3]. In Ethiopia, BC is one of the most common cancers [4]. Previous survival studies demonstrated that women in Ethiopia had a favorable 5-year outcome with 45% metastasis free survival [5]. In another study in rural Ethiopia, the 2-year overall survival was only 53% [6].

Prognostic and predictive factors such as age, tumor size, nodal status, grading, Ki-67 proliferation index, angioinvasion, hormone receptor (HR) status, human epidermal growth factor receptor 2 (HER2) status are used in the clinical routine for prognosis of the patient and predicting benefit from therapy [7]. In addition, molecular biomarkers have an increasingly important role tailoring the individualized treatment recommendation for BC patients in the selection of chemotherapy, endocrine or immunotherapy regimen [8].

Gene expression analysis has helped identify distinct molecular signatures in breast cancer that have different prognostic outcomes in addition to clinical and histopathological features. Multi-gene assays in early breast cancer are now routinely used in clinical practice and are integrated into national and international guidelines [9]. Parker and colleagues simplified the profiling algorithms using 50 genes for classification in one of the intrinsic subtypes: luminal A, luminal B, HER2-enriched, and basal-like [10, 11]. The PAM50 classifier provides in addition to traditional clinical and histopathological biomarkers prognostic and predictive value for BC patients [11].

Studies from different regions in Africa have reported proportions of estrogen receptor-positive disease varying between 20 and 70% supposedly due to differences in genetic background, tumor size, but also quality issues have been discussed [12]. Our own data from Ethiopia showed 65% estrogen receptor-positive tumors assessed by immunohistochemistry [13]. The present study was conducted to assess the performance of the PAM50 classifier in Ethiopian breast cancer specimens and to assess associations of the intrinsic subtypes with protein expression of estrogen receptor, progesterone receptor and HER2.

Patients and methods

Patients and samples

All female patients with invasive carcinoma of the breast were included in this study in accordance with the REMARK criteria [14]. Formalin-fixed paraffin-embedded (FFPE) tissues from pathologically confirmed BC samples ($n = 334$) were collected by ZD and MY from different hospitals and pathology laboratories in Ethiopia (Tikur Anbessa Specialized Hospital, Zewditu Memorial Hospital, Yekatit 12 Hospital, Bethzatha General Hospital, St. Paul's Hospital Millennium Medical College, Aira General Hospital). Clinical and histopathological data including age at diagnosis, tumor size, histologic type and tumor differentiation were retrieved from patient's medical records. An overview of all parameters considered here is given in the Supplementary Table S1.

IHC and gene expression analysis

FFPE tissue was employed for the analysis of the expression of ER, PgR, HER2, and Ki-67 using immunohistochemistry (IHC). IHC analyses were carried out using specific antibodies according to the manufacturers' instructions as follows:

- ER: Clone Ab-11; Thermo Scientific, MA; USA, Catalog Number MS-354-P1, mouse host, dilution 1:150
- PgR: Clone PgR 636; DAKO, CA, USA, Catalog Number M3569, mouse host, dilution 1:100
- HER2: Clone DG44; DAKO; CA, USA, Catalog Number SK001, rabbit host, ready to use
- Ki-67: Clone SP6; Thermo Scientific; MA, USA, Catalog Number RM-9106-S; mouse host; dilution 1:250

Expression of HR and HER2 status was performed according to the current guidelines [15], positivity of ER and PgR status was declared when the IRS was > 0. If at least one of the markers were positive, the HR status was defined as positive. HER2 status was assessed according to ASC-CAP guidelines [16], HER2 DAKO 3 was considered positive, HER2 DAKO 0 and DAKO 1 was considered HER2-negative. Tissue with HER2 DAKO 2 was interpreted as equivocal and cromogen in situ hybridization (CISH) was performed for confirmation. Based on previous reports, Ki-67 proliferation index was graded as 'low'

if Ki-67 staining was positive in < 20% of tumor cells or 'high' when at least 20% of tumor cells stained positive [17, 18]. The histological grading was performed using Elston–Ellis grading system [19].

After identification of the tumor areas on HE-stained slides by the pathologist, 3–5 adjacent unstained tumor slides (10 µm) were processed using miRNeasy FFPE Mini Kit® (Qiagen) according to the manufacturer's protocol. The extracted RNA concentration and quality was measured using Nanophotometer.

Relative gene expression was measured using the NanoString nCounter® Analysis System (NanoString Technologies, Seattle, WA, USA) using a multiplexed hybridization assay, digital readouts of fluorescent barcoded probes which hybridize with each mRNA sequence of interest. The data collection was carried out in the nCounter® Digital Analyzer. Data import, quality control, and normalization of expression levels were conducted with the nSolver software version 4 (NanoString Technologies, Inc.). Background subtraction from raw transcript counts was performed through negative input controls. Following reference-normalization by dividing the geometric mean of six references-control genes (*ACTB*, *G6PD*, *RPLP0*, *TBP*, *TFRC* and *UBB*), normalized counts were log2-transformed prior to all downstream analyses. Intrinsic subtype classification was calculated using the nearest PAM50 centroid algorithm Bioclassifier and NanoStringNorm implemented in R [11, 20].

Endpoints and statistical analysis

As a first objective, the distribution of the intrinsic subtypes within an Ethiopian cohort was defined. The secondary objectives were the associations of the intrinsic subtypes to clinical and histopathological parameters including the IHC groups, applying logistic regression. In order to tackle multiple testing, we reduced the subgroup analyses to pre-defined, well-accepted and clinical relevant groups (e.g., age, tumor size, tumor grade/differentiation, ER status, PgR status, HER2 status, Ki-67 proliferation index and IHC groups). Statistical significance was declared for p-values < 1% (two-sided Pearson's chi-square tests for independence, with Yates' correction for continuity when relevant). Statistical analyses were carried out using SPSS 25 (IBM, Armonk NY, USA).

Results

Distribution of intrinsic subtypes

The classification of intrinsic subtypes based on the PAM50 assay yielded 104 luminal A (31.1%), 91 luminal B (27.2%), 62 HER2-enriched (18.6%) and 77 basal-like samples (23.1%). The expression levels of the 50 loci included in the assay (Fig. 1) revealed a gene expression signature which was unique for each subtype. An overview of the distribution of selected clinical and histopathological parameters among the PAM50-based intrinsic subtypes is given in Table 1.

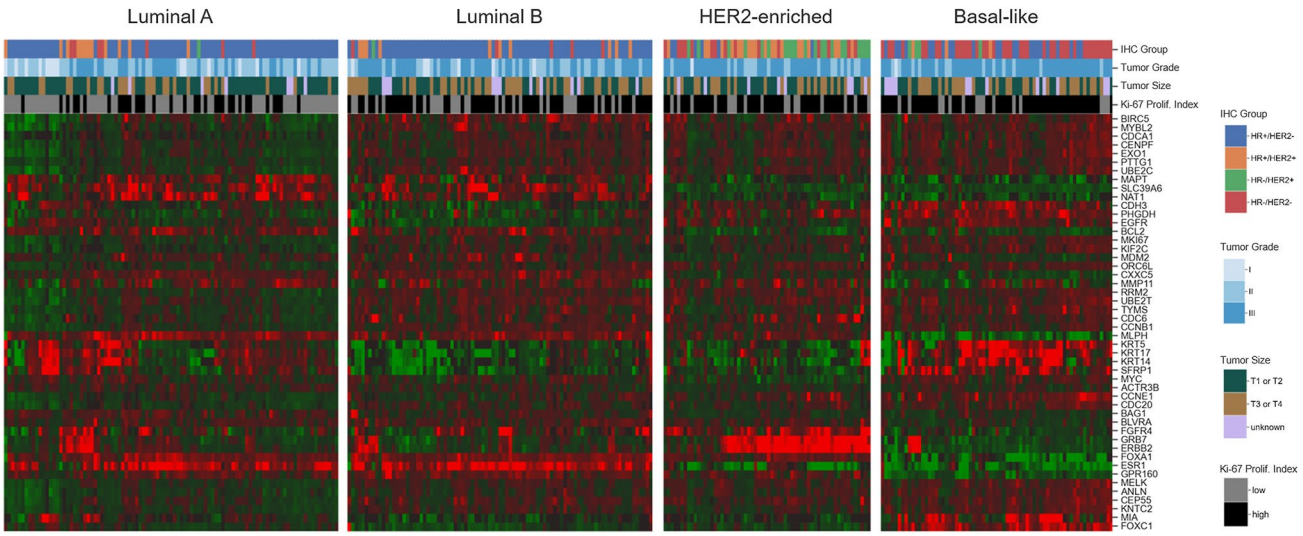


Fig. 1 Gene Expression heatmap of the 50 loci used for the PAM50 classification of 334 BC samples. The 334 samples are grouped horizontally according to their intrinsic subtype, which are indicated above each block. Red tiles denote overexpression, while green tiles correspond to underexpression. The four horizontal bars above the heatmaps indicate the classification of samples according to IHC groups, tumor grade, tumor size and Ki-67 proliferation index (top-down, respectively, with color codes for each bar given at the right side of the figure). *HR* hormone receptor, *HER2* human epidermal growth factor receptor 2

Table 1 Distribution of clinical and histopathological parameters among intrinsic subtypes

Parameters	All $n=334$	Luminal A $n=104$	Luminal B $n=91$	HER2−enriched $n=62$	Basal-like $n=77$	p-value
Age group (years)						0.67
< 50	201 (60.2%)	66 (63.5%)	58 (63.7%)	33 (53.2%)	44 (57.1%)	
≥ 50	95 (28.4%)	29 (27.9%)	24 (26.4%)	21 (33.9%)	21 (27.3%)	
Unknown	38 (11.4%)	9 (8.7%)	9 (9.9%)	8 (12.9%)	12 (15.6%)	
Tumor size						0.01
T1 or T2	168 (50.3%)	68 (65.4%)	39 (42.9%)	31 (50.0%)	30 (39.0%)	
T3 or T4	126 (37.7%)	29 (27.9%)	41 (45.1%)	23 (37.1%)	33 (42.9%)	
Unknown	40 (12.0%)	7 (6.7%)	11 (12.1%)	8 (12.9%)	14 (18.2%)	
Histological types						0.45
NST	303 (90.7%)	94 (90.4%)	86 (94.5%)	54 (87.1%)	69 (89.6%)	
Non-NST	31 (9.3%)	10 (9.6%)	5 (5.5%)	8 (12.9%)	8 (10.4%)	
Tumor grade*						$\mathbf{3.35 \times 10^{-11}}$
G1 or G2	140 (41.9%)	71 (68.3%)	37 (40.7%)	18 (29.0%)	14 (18.2%)	
G3	194 (58.1%)	33 (31.7%)	54 (59.3%)	44 (71.0%)	63 (81.8%)	
Estrogen receptor status*						$\mathbf{5.11 \times 10^{-16}}$
Positive (≥ 1%)	184 (55.1%)	77 (74.0%)	69 (75.8%)	18 (29.0%)	20 (26.0%)	
Negative (< 1%)	150 (44.9%)	27 (26.0%)	22 (24.2%)	44 (71.0%)	57 (74.0%)	
Progesterone receptor status*						$\mathbf{6.99 \times 10^{-15}}$
Positive (≥ 1%)	157 (47.0%)	69 (66.3%)	60 (65.9%)	13 (21.0%)	15 (19.5%)	
Negative (< 1%)	177 (53.0%)	35 (33.7%)	31 (34.1%)	49 (79.0%)	62 (80.5%)	
Hormone receptor status*						$\mathbf{4.43 \times 10^{-25}}$
Positive	232 (69.5%)	95 (91.3%)	85 (93.4%)	26 (41.9%)	26 (33.8%)	
Negative	102 (30.5%)	9 (8.7%)	6 (6.6%)	36 (58.1%)	51 (66.2%)	
HER2 status*						$\mathbf{7.56 \times 10^{-19}}$
Negative	261 (78.1%)	91 (87.5%)	80 (87.9%)	21 (33.9%)	69 (89.6%)	
Positive	73 (21.9%)	13 (12.5%)	11 (12.1%)	41 (66.1%)	8 (10.4%)	
Ki-67 proliferation index*						$\mathbf{2.48 \times 10^{-8}}$
Low (< 20%)	132 (39.5%)	66 (63.5%)	31 (34.1%)	17 (27.4%)	18 (23.4%)	
High (≥ 20%)	202 (60.5%)	38 (36.5%)	60 (65.9%)	45 (72.6%)	59 (76.6%)	

HER2 human epidermal growth factor receptor 2, *NST* no special type, *ER* estrogen receptor, *PgR* progesterone receptor, *HR* hormone receptor
*Parameters for which a p-value (from a χ^2 test for independence) below 1% was observed, denoting a lack of independence between histopathological parameters and intrinsic subtypes

Intrinsic subtypes and IHC groups

A strong correlation was observed between intrinsic subtypes and IHC groups (p-value from χ^2 test for independence < 0.001, Fig. 2). This association was driven mainly by the luminal subtypes, from which 81.6% ($n=158$) were grouped as HR-positive and HER2-negative (Fig. 2, upper left). A multivariate regression allowed us to confirm the strong correlation (ORs > 20) between the luminal intrinsic subtypes and the HR+/HER2–IHC group, after adjusting for age, tumor size, histological type, tumor grade and Ki-67 proliferation index (Table 2).

Out of the 62 samples classified as HER2-enriched, 66.1% ($n=41$) were defined as HER2–positive irrespective of HR status (Fig. 2). The results of a multivariate regression confirmed this, as HER2-enriched samples were significantly associated with a very high probability of being HER2-positive (independent of HR status, ORs > 20, Table 2). HER2-enriched samples were also associated with TNBC, although with a lower odds ratio (OR 6.7, 99% CI 1.77–25.39, Table 2).

The samples classified as basal-like had an increased probability of being TNBC (46 of 77 samples, 60%), as compared with triple-positive (HR+/HER2+) samples (Table 2).

Intrinsic subtypes and histopathological parameters

Luminal A samples had a more than three-fold increased probability for favorable characteristics such as higher tumor differentiation (G1 or G2) and low Ki-67 proliferation index (OR 3.19, 99% CI 1.47–6.92 and OR 3.65, 99% CI 1.72–7.72, respectively, Table 2). Basal-like samples were

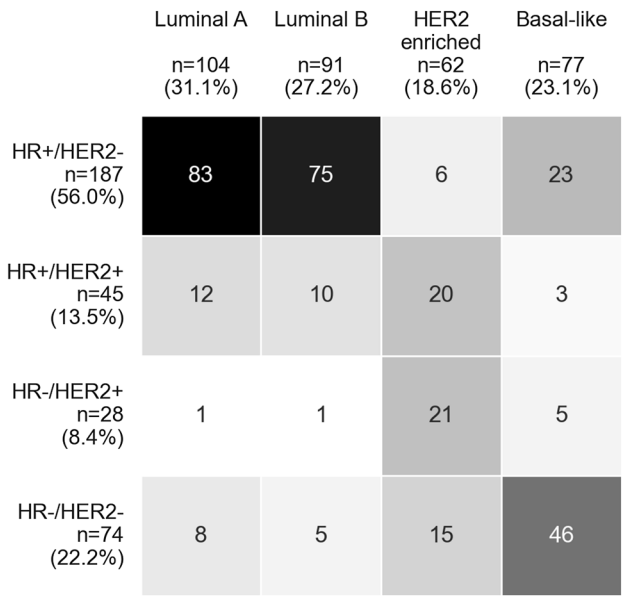


Fig. 2 Color-coded crosstable of 334 BC tissue samples grouped according to IHC groups (rows) and intrinsic subtypes (columns). The cell color gradient indicates the relationship in terms of a strong discordance (white) to a strong concordance (black) between IHC and intrinsic subtype classifications. *HR* hormone receptor, *HER2* human epidermal growth factor receptor 2

associated with lower differentiation (G3; OR 3.43, 95% CI 1.26–9.34, Table 2). No further significant associations between intrinsic subtypes and histopathological parameters were found. A detailed analysis focused on the whole measured range of the prognostic parameters ER, PgR, HER2 and Ki-67 proliferation index is visualized in Fig. 3. While the IRS scores of ER and PgR spanned evenly between 0 and 12 for both luminal subtypes, these accumulated below IRS 2 for HER2-enriched and basal-like samples (Fig. 3A, B). This underlines the associations described between intrinsic subtypes and the expression of the steroid hormone receptors detected by IHC (see Table S2 for binary assessments). The violin plot of DAKO scores made the aforementioned association between HER2-enriched samples and positive HER2 status evident. The accumulation of scores > 2 was almost exclusively seen in HER2-enriched samples (Fig. 3C; Table S2). Similarly, specimens of low proliferation rates (< 20%) were more likely to be of the luminal A subtype (Fig. 3D; Table S2). However, low proliferation rates were also present in the other intrinsic subtypes. The highest proliferation rates were observed among basal-like samples.

Discussion

In this prospective cohort of Ethiopian patients with BC, intrinsic PAM50-based subtypes of tumors were determined in addition to receptor status by immunohistochemistry ($n = 334$). Both methods had considerable similarities, especially for basal-like or triple-negative and HER2-enriched or HER2-positive types. The high proportion of patients below the age of 50 (approx. 60%) and large tumor sizes (T3 and T4, approx. 40%) differed from high-income settings. This probably resulted in the lower proportions of luminal A subtypes compared to Western cohorts.

Pattern of subtypes

We reported in 2014 that two thirds of tumors from a patient cohorts from Addis Ababa had endocrine responsive disease [13]. Another cohort from rural Ethiopia showed more than half of the patients with positive hormone receptor status [6]. The third study published in 2018 again from Addis Ababa also reported 65% of receptor-positive disease [21]. Within the current study, immunohistochemistry as well as RNA-expression analysis were used to assess endocrine responsiveness. The results also revealed that even using RNA-expression based subtyping, more than half of the tumors were endocrine responsive. In detail, we found 31.1% luminal A, 27.2% luminal B, 18.6% HER2-enriched and 23.1% basal-like tumors. This is considerably shifted to the more aggressive subtypes compared, for example, to data from the United States (Nurses Health Study) that reported 46% luminal A, 18% luminal B, 14% HER2-enriched, 15% basal-like, and 8% normal-like subtypes [22]. Such direct comparison needs to be interpreted with care.

Comparing with other ethnic groups

High quality data on tumor subtypes from the cancer genome atlas about tumors from African patients is lacking. Patients with African compared to European ancestry had a higher likelihood of basal-like (odds ratio OR 1.67) and HER2-enriched (OR 2.22) tumors [23]. African–American patients genetic background has been reported as predominantly from West Africa [24].

Geographic variations have been observed across Africa concerning the composition of different hormone receptor-based types no longer characterizing all tumors in Africa as having aggressive phenotypes. A large meta-analysis found that patterns of tumor subtypes in East Africa appear to be more favorable compared to other geographic regions such as West Africa [12].

Table 2 Results of multivariate logistic regression of clinical and histopathological parameters, taken as predictive variables for intrinsic subtypes

Parameters	Luminal A ($n=104$)		Luminal B ($n=91$)		HER2−enriched ($n=62$)		Basal-like ($n=77$)	
	OR (99% CI)	*p*-value	OR (99% CI)	*p*-value	OR (99% CI)	*p*-value	OR (99% CI)	*p*-value
Age group (years)								
<50 ($n=201$)	1.47 (0.68–3.16)	0.20	0.95 (0.46–1.97)	0.87	0.76 (0.30–1.92)	0.45	1.00 (0.43–2.31)	0.99
≥50 ($n=95$)	Ref		Ref		Ref		Ref	
Tumor size								
T1 or T2 ($n=168$)	Ref		Ref		Ref		Ref	
T3 or T4 ($n=126$)	0.66 (0.30–1.44)	0.17	1.74 (0.84–3.62)	0.05	0.72 (0.28–1.88)	0.38	1.04 (0.45–2.42)	0.89
Histological type								
NST ($n=303$)	1.36 (0.39–4.78)	0.53	1.88 (0.46–7.76)	0.25	0.50 (0.11–2.31)	0.24	0.65 (0.16–2.67)	0.43
Non-NST ($n=31$)	Ref		Ref		Ref		Ref	
Tumor grade								
G1 or G2 ($n=140$)	**3.19 (1.47–6.92)**	**1.21×10^{-4}**	Ref		Ref		Ref	
G3 ($n=194$)	Ref		1.16 (0.52–2.57)	0.63	1.45 (0.50–4.14)	0.37	**3.43 (1.26–9.34)**	**1.54×10^{-3}**
Ki-67 proliferation index								
Low (<20%) ($n=132$)	**3.65 (1.72–7.72)**	**9×10^{-6}**	Ref		Ref		Ref	
High (≥20%) ($n=202$)	Ref		1.64 (0.76–3.56)	0.10	1.50 (0.54–4.13)	0.31	2.40 (0.95–6.09)	0.02
IHC group								
HR+/HER2− ($n=187$)	**20.52 (1.36–309.24)**	**4.12×10^{-3}**	**20.62 (1.43–297.79)**	**3.51×10^{-3}**	Ref		2.88 (0.54–15.32)	0.10
HR+/HER2+($n=45$)	16.30 (0.94–281.70)	0.01	6.80 (0.41–111.73)	0.08	**22.62 (5.86–87.36)**	**2.74×10^{-9}**	Ref	
HR−/HER2+($n=28$)	Ref		Ref		**88.30 (18.12–430.32)**	**3.16×10^{-13}**	3.52 (0.45–27.62)	0.12
HR−/HER2− ($n=74$)	3.89 (0.22–68.16)	0.22	1.86 (0.10–33.46)	0.58	**6.70 (1.77–25.39)**	**2.33×10^{-4}**	**30.58 (5.49–170.50)**	**2.94×10^{-7}**

OR odds ratio, *CI* confidence interval, *Ref* category taken as reference within each parameter for each test, *HER2* human epidermal growth factor receptor 2, *NST* no special type, *HR* hormone receptor

*Bold letters: cases with *p*-values < 0.01

Other factors influencing subtype composition

Our convenient hospital cohort has specific features. The age composition as well as proportions of early and later stage disease may influence the overall proportions of subtypes. In general, cohorts from sub Saharan Africa seem to similarly represent a young population and commonly comprise of late stage tumors compared to Western settings. It was shown in a cohort from the United States that patients with luminal A tumors are on average 5–6 years older than patients with the other subtypes [25]. A comparison of receptor status between Sudanese and German patients also showed considerably more low-grade, receptor-positive tumors in older German women [26]. Therefore cohorts with fewer older patients lack these luminal tumors. Additionally, body composition has been reported to influence subtype. Obesity has been associated with higher risk of luminal A type breast cancer [27]. Since obesity is still rare in Ethiopia, less luminal A breast cancers are expected [28].

Comparing subtypes and IHC

Assessing tumor biology with RNA-expression analysis is considered gold standard, at the same time immunohistochemistry has been used as basis to prove effectiveness of endocrine treatment. When comparing both methods,

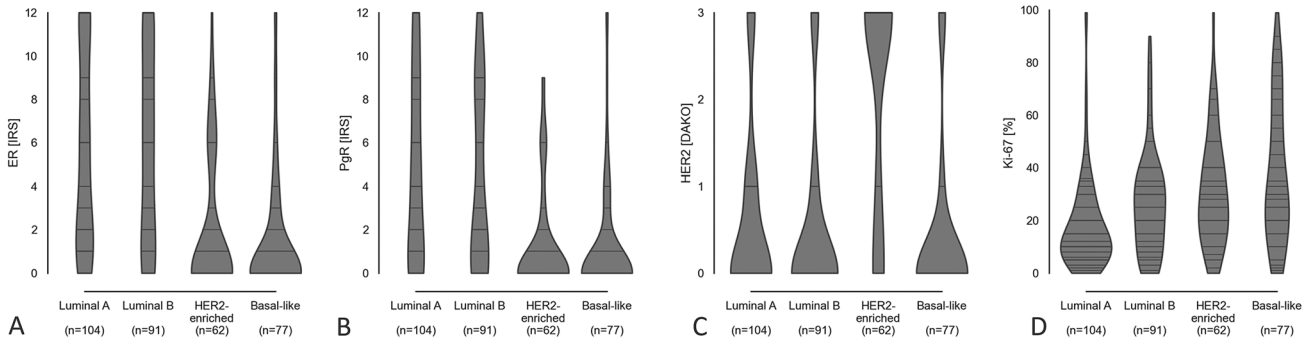


Fig. 3 Scores of Estrogen Receptor, Progesterone Receptor, HER2 and Ki-67 Proliferation Index among intrinsic subtypes. The horizontal widths of the violin plots correspond to the frequency distribution of data points. Horizontal bars shown in the plots denote data available for those values. *ER* estrogen receptor, *PgR* progesterone receptor, *IRS* immune-reactive score, *ER* estrogen receptor, *PgR* progesterone receptor, *HER2* human epidermal growth factor receptor 2

certain discrepancies can be seen. Consistent with previous reports [29], we have observed a substantial mismatch between the classification based on gene expression (PAM50 intrinsic subtypes) and on immunohistochemistry (IHC groups using ER, PgR and HER2). Still, hormone receptor-positive tumors were mainly classified as luminal subtypes. Therefore, immunohistochemistry results appear reliable as a basis for decision to advise endocrine treatment to the patient despite lack of standardized external quality assurance measures for tissue processing and immunohistochemistry performance [30].

In Ethiopia, as well as many other countries in Sub-Saharan Africa, basic immunohistochemistry is not always accessible to all patients. Even in capital cities, 15 out of 20 centers experienced frequent power cuts and four out of twenty had no immunohistochemistry in the country [31]. Due to the fact that also for Ethiopian patients' routine immunohistochemistry is not always available, several confirmatory studies reported that half of patients have endocrine responsive disease (confirmed by RNA-expression as well as immunohistochemistry studies). This may provide clinicians more confidence in their decision process to offer endocrine treatment for patients with unknown receptor status.

Individualized therapy

As a future perspective, multiplexed hybridization assay-based subtype determination via nCounter® can also support individualized therapy in Ethiopia. Since a large proportion of tumors are larger than 2 cm or lymph node-positive, without additional prognostic markers, according to the National Cancer Control Network (NCCN) harmonized guidelines, nearly all patients would need chemotherapy. As an example, utilizing PAM50-based subtypes, we were able to split the large group of patients with HR-positive and HER2-negative tumors into clinically relevant, more homogenous intrinsic subtypes. Patients with HR+/HER2 tumors have been regarded as an especially challenging BC group due to high variance concerning clinical outcome [32]. These patients include cases of low, intermediate and high risk of recurrence. Differentiating the HR+/HER2 group into luminal A and non-luminal A intrinsic subtypes increases the ability to assess recurrence risk in a more individualized manner, since chemotherapy is not recommended for patients with luminal A tumors [33, 34]. This information also enables clinicians to counsel patients on the importance of adherence to adjuvant endocrine therapy long-term and surveillance for local or distant recurrence.

Individualized therapy: HER2-*enriched and basal-like types*

The subtype determination can also personalize treatment for patients with aggressive tumors. The PAM50 assay allowed the identification of 62 tumors as HER2-enriched, regardless of HER2 status. Using the IHC method, only 41 patients would get an anti-HER2 therapy. This means that the PAM50 assay leads to an approximately 50% incremented number of patients who would benefit from drugs suppressing the HER2 signaling pathway (anti-HER2 therapy). Our results are concordant with the seminal work of Perou and colleagues in the sense that two thirds of our samples classified as HER2-positive were found to be HER2-enriched (Table 1) [10].

Finally, 77 samples were classified as basal-like through the PAM50 assay. Out of these, only 46 (60%) were concordant with the triple-negative class via the IHC method. These samples are arguably the ones that would benefit most from chemotherapy, 40% of the basal-like samples are non-TNBC [10]. Conversely, approximately 38% of the TNBC tumors were non-basal-like. Thus, both PAM50 and IHC classification methods yield heterogeneous groups where personalized recommendations would improve by additional molecular subtyping as proposed before [35, 36].

Strengths and limitations

The strength of the study includes multi-center involvement in Ethiopia and the application of multigene assays for gene expression profiling which is the gold standard for molecular classification as well as hormone receptor status by immunohistochemistry involving relatively large sample size. Both methods similarly showed a large share of endocrine responsive tumors. Certain limitations have to be taken into account. Firstly, the cohort is a convenient hospital-based sample and does not reflect the true pattern within the population. Given the resource-limited setting, population-based sample collection is not feasible. Secondly, Ethiopian patients vary considerably linguistically, ethnically and culturally within the country. Therefore results cannot be generalized within the country and a nation-wide study investigating distribution intrinsic subtypes and further exploration of tumor biology is essential to capture the large number of different ethnic groups in Ethiopia which could possibly help to have a broaden understanding of BC biology.

Conclusion

To our knowledge, this is the first time that a large African BC patient group is characterized by molecular expression profiling. In summary, we confirmed previous immunohistochemistry results showing a considerable proportion of more than half the patients eligible for endocrine treatment. This allows utilization of a cost-effective treatment with very little side-effects that can be administered even at primary or secondary level health facilities. A study from Ethiopia proved that a strategy involving the specialized training of "cancer nurses" to support patients during their 5 year treatment can improve adherence [37].

Efforts are needed to provide sufficient access to immunohistochemistry service but in the meantime can encourage to utilization of endocrine treatment for patients with unknown receptor status. Within this project, colleagues from Ethiopia (ZD, MY, TA) received thorough training in tumor-banking, RNA-extraction and PAM50 subtyping. Additionally, during the time of the study and still ongoing is a larger training program to implement routine immunohistochemistry at Addis Ababa University involving weekly virtual meetings for case-discussions, several training courses in Ethiopia as well as in person mentorship in Germany. In 2022, consumables were provided and routine IHC for ER, PgR and HER2 determinations are performed in Ethiopia. Endpoint-PCR for research only was also implemented at Addis Ababa University as a step to further develop capacity in laboratory techniques. Since molecular methods are becoming available at lower prices, utilization of RNA-expression-based subtype assessment could become an option to optimize personalized treatment. One-step PCR technologies to assess basic receptor status are a possible compromise given the robustness of the methodology and at the same time relatively low cost availability.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10549-022-06769-z.

Acknowledgements We would like to express our gratitude to the study participants and data collectors without whom it would have not been possible to realize the study. The authors are grateful for Susan. G. Komem for granting the research project. Last but not least, our heartfelt thanks go to Addis Ababa University for covering the tuition fee expense during my study.

Author contributions ZD contributed to study design, sample and data acquisition, analysis, interpretation and writing of the original and final draft. MY contributed to sample acquisition. MY, KS, EJK, MV, LT and TA contributed to study design, data acquisition, data analysis, data interpretation and editing of the manuscript. ZD, KS and MV carried out the experimental work. EA, YB, MA and YA contributed to data acquisition and editing of the manuscript. MP and PS contributed to data analysis, plotting and statistics. All authors reviewed and approved the final version of the manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL. This study took advantage of grant by the Susan G. Komen organization assigned to the Medical Faculty of the Martin Luther University Halle-Wittenberg, Halle (Saale), Germany (Grant No. GTDR16378013). This work was supported by Else-Kroener-Foundation through Martin-Luther-University, Halle-Wittenberg, Germany, Grant No. 2018_HA31SP. This study was also support by a grant from the German Academic Exchange Service to Martin-Luther-University, Halle-Wittenberg, Germany (ID 57216764). The study was also supported by a grant from Hospital partnership through Deutsche Gesellschaft für Internationale Zusammenarbeit funded by the Ministry for Economic Cooperation and Development and the Else-Kroener-Fresenius Foundation (ID 81256434). EJK also received an unrestricted grant from F. Hoffmann-La Roche Ltd (27.5.2014)—the funder was not involved in the design, analysis or publication of the study.

Data availability The data generated in this study are available within the article and its supplementary data files. Raw data were generated and processed from the authors and are available on request to the corresponding authors.

Declarations

Competing interest The authors have no relevant financial or non-financial interests to disclose.

Ethical approval This study was performed in line with the principles of the Declaration of Helsinki. Ethical approval was waived by the local ethics committee of all institutions. This work was carried out under IRB approved protocols at Addis Ababa University (Protocol Number 092/17/17) and Ethiopian National Research Ethical Review Committee (Protocol MOSHE/RD).

Informed consent Informed written consent was obtained from all participants.

References

1. Feng Y, Spezia M, Huang S et al (2018) Breast cancer development and progression: risk factors, cancer stem cells, signaling pathways, genomics, and molecular pathogenesis. Genes Dis 5(2):77–106. https://doi.org/10.1016/j.gendis.2018.05.001
2. Sung H, Ferlay J, Siegel RL et al (2021) Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. Cancer J Clin 71(3):209–249. https://doi.org/10.3322/caac.21660
3. Joko-Fru WY, Griesel M, Mezger NCS et al (2021) Breast cancer diagnostics, therapy, and outcomes in Sub-Saharan Africa: a population-based registry study. J Natl Compr Cancer Netw. https://doi.org/10.6004/jnccn.2021.7011
4. Timotewos G, Solomon A, Mathewos A et al (2018) First data from a population based cancer registry in Ethiopia. Cancer Epidemiol 53:93–98. https://doi.org/10.1016/j.canep.2018.01.008
5. Kantelhardt EJ, Zerche P, Mathewos A et al (2014) Breast cancer survival in Ethiopia: a cohort study of 1070 women. Int J Cancer 135(3):702–709. https://doi.org/10.1002/ijc.28691
6. Eber-Schulz P, Tariku W, Reibold C et al (2018) Survival of breast cancer patients in rural Ethiopia. Breast Cancer Res Treat 170(1):111–118. https://doi.org/10.1007/s10549-018-4724-z
7. Gradishar WJ, Moran MS, Abraham J et al (2022) Breast cancer, version 3.2022, NCCN clinical practice guidelines in oncology. J Natl Compr Cancer Netw 20(6):691–722. https://doi.org/10.6004/jnccn.2022.0030
8. Andre F, Ismaila N, Allison KH et al (2022) Biomarkers for adjuvant endocrine and chemotherapy in early-stage breast cancer: ASCO guideline update. J Clin Oncol 40(16):1816–1837. https://doi.org/10.1200/JCO.22.00069
9. Giorgi Rossi P, Lebeau A, Canelo-Aybar C et al (2021) Recommendations from the European Commission initiative on breast cancer for multigene testing to guide the use of adjuvant chemotherapy in patients with early breast cancer, hormone receptor positive, HER-2 negative. Br J Cancer 124(9):1503–1512. https://doi.org/10.1038/s41416-020-01247-z
10. Perou CM, Sorlie T, Eisen M (2000) Molecular portraits of human breast tumours. Nature 406:747–752
11. Parker JS, Mullins M, Cheang MCU et al (2009) Supervised risk predictor of breast cancer based on intrinsic subtypes. J Clin Oncol 27(8):1160–1167. https://doi.org/10.1200/JCO.2008.18.1370
12. Eng A, McCormack V, dos-Santos-Silva I (2014) Receptor-defined subtypes of breast cancer in indigenous populations in Africa: a systematic review and meta-analysis. PLoS Med. https://doi.org/10.1371/journal.pmed.1001720
13. Kantelhardt EJ, Mathewos A, Aynalem A et al (2014) The prevalence of estrogen receptor-negative breast cancer in Ethiopia. BMC Cancer 14:895. https://doi.org/10.1186/1471-2407-14-895
14. McShane LM, Altman DG, Sauerbrei W et al (2005) REporting recommendations for tumour MARKer prognostic studies (REMARK). Br J Cancer 93(4):387–391. https://doi.org/10.1038/sj.bjc.6602678
15. Allison KH, Hammond MEH, Dowsett M et al (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
16. Wolff AC, Hammond MEH, Allison KH et al (2018) 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 36(20):2105–2122. https://doi.org/10.1200/JCO.2018.77.8738
17. Cheang MCU, Chia SK, Voduc D et al (2009) Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. J Natl Cancer Inst 101(10):736–750. https://doi.org/10.1093/jnci/djp082
18. Dowsett M, Nielsen TO, A'Hern R et al (2011) Assessment of Ki67 in breast cancer: recommendations from the International Ki67 in breast cancer working group. J Natl Cancer Inst 103(22):1656–1664. https://doi.org/10.1093/jnci/djr393
19. Elston CW, Ellis IO (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
20. Berchtold E, Vetter M, Gündert M et al (2019) Comparison of six breast cancer classifiers using qPCR. Bioinformatics 35(18):3412–3420. https://doi.org/10.1093/bioinformatics/btz103
21. Hadgu E, Seifu D, Tigneh W et al (2018) Breast cancer in Ethiopia: evidence for geographic difference in the distribution of molecular subtypes in Africa. BMC Womens Health 18(1):40. https://doi.org/10.1186/s12905-018-0531-2
22. Kensler KH, Sankar VN, Wang J et al (2018) PAM50 molecular intrinsic subtypes in the nurses' health study cohorts. Cancer Epidemiol Biomark Prev 28(4):798–806. https://doi.org/10.1158/1055-9965.EPI-18-0863
23. Sweeney C, Bernard PS, Factor RE et al (2014) Intrinsic subtypes from PAM50 gene expression assay in a population-based breast cancer cohort: differences by age, race, and tumor characteristics. Cancer Epidemiol Biomark Prev 23(5):714–724. https://doi.org/10.1158/1055-9965.EPI-13-1023
24. Newman LA (2014) Breast cancer disparities: high-risk breast cancer and African ancestry. Surg Oncol Clin N Am 23(3):579–592. https://doi.org/10.1016/j.soc.2014.03.014
25. Allott EH, Shan Y, Chen M et al (2020) Bimodal age distribution at diagnosis in breast cancer persists across molecular and genomic classifications. Breast Cancer Res Treat 179(1):185–195. https://doi.org/10.1007/s10549-019-05442-2
26. Sengal AT, Haj Mukhtar NS, Vetter M et al (2018) Comparison of receptor-defined breast cancer subtypes between German and Sudanese women: a facility-based cohort study. J Glob Oncol 4:1–12. https://doi.org/10.1200/JGO.2017.010082
27. Miyagawa Y, Miyake T, Yanai A et al (2015) Association of body mass index with risk of luminal A but not luminal B estrogen receptor-positive and HER2-negative breast cancer for postmenopausal Japanese women. Breast Cancer 22(4):399–405. https://doi.org/10.1007/s12282-013-0493-z
28. Kwan ML, Kroenke CH, Sweeney C et al (2015) Association of high obesity with PAM50 breast cancer intrinsic subtypes and gene expression. BMC Cancer 15:278. https://doi.org/10.1186/s12885-015-1263-4

29. Sørlie T, Perou CM, Tibshirani R et al (2001) Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci USA 98(19):10869–10874. https://doi.org/10.1073/pnas.191367098
30. Jorns JM (2019) Breast cancer biomarkers: challenges in routine estrogen receptor, progesterone receptor, and HER2/neu evaluation. Arch Pathol Lab Med 143(12):1444–1449. https://doi.org/10.5858/arpa.2019-0205-RA
31. Ziegenhorn H-V, Frie KG, Ekanem I-O et al (2020) Breast cancer pathology services in Sub-Saharan Africa: a survey within population-based cancer registries. BMC Health Serv Res. https://doi.org/10.1186/s12913-020-05752-y
32. Balic M, Thomssen C, Würstlein R et al (2019) St. Gallen/Vienna 2019: a brief summary of the consensus discussion on the optimal primary breast cancer treatment. Breast Care 14(2):103–110. https://doi.org/10.1159/000499931
33. Jensen M-B, Lænkholm A-V, Nielsen TO et al (2018) The Prosigna gene expression assay and responsiveness to adjuvant cyclophosphamide-based chemotherapy in premenopausal high-risk patients with breast cancer. Breast Cancer Res. https://doi.org/10.1186/s13058-018-1012-0
34. Chia SK, Bramwell VH, Tu D et al (2012) A 50-gene intrinsic subtype classifier for prognosis and prediction of benefit from adjuvant tamoxifen. Clin Cancer Res 18(16):4465–4472. https://doi.org/10.1158/1078-0432.CCR-12-0286
35. Lehmann BD, Bauer JA, Chen X et al (2011) Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. J Clin Investig 121(7):2750–2767. https://doi.org/10.1172/JCI45014
36. Hartung C, Porsch M, Stückrath K et al (2021) Identifying high-risk triple-negative breast cancer patients by molecular subtyping. Breast Care 16(6):637–647. https://doi.org/10.1159/000519255
37. Getachew S, Addissie A, Seife E et al (2022) Breast nurse intervention to improve adherence to endocrine therapy among breast cancer patients in South Ethiopia. Oncologist 27(8):e650–e660. https://doi.org/10.1093/oncolo/oyac081

Published Paper Two

cancers

MDPI

Article

Human Breast Tissue Microbiota Reveals Unique Microbial Signatures that Correlate with Prognostic Features in Adult Ethiopian Women with Breast Cancer

Zelalem Desalegn [1,2], Alana Smith [3], Meron Yohannes [1,2,4], Xueyuan Cao [5], Endale Anberber [6], Yonas Bekuretsion [7], Mathewos Assefa [8], Marcus Bauer [9], Martina Vetter [10], Eva Johanna Kantelhardt [2,10,11], Tamrat Abebe [1,2,†] and Athena Starlard-Davenport [3,*,†]

1 Department of Microbiology, Immunology, and Parasitology, School of Medicine, College of Health Sciences Addis Ababa University, Addis Ababa 9086, Ethiopia; zelalem.desalegn@aau.edu.et (Z.D.); meron.yohannes@aau.edu.et (M.Y.); tamrat.abebe@aau.edu.et (T.A.)
2 Global Health Working Group, Martin Luther University Halle-Wittenberg, 06097 Halle (Saale), Germany; eva.kantelhardt@uk-halle.de
3 Department of Genetics, Genomics and Informatics, College of Medicine, University of Tennessee Health Science Center, Memphis, TN 38163, USA; aantoin1@uthsc.edu
4 School of Medical Laboratory Sciences, Addis Ababa University, Addis Ababa 9086, Ethiopia
5 Department of Health Promotion and Disease Prevention, College of Nursing, University of Tennessee Health Science Center, Memphis, TN 38163, USA; xcao12@uthsc.edu
6 Department of Surgery, School of Medicine, Addis Ababa University, Addis Ababa 9086, Ethiopia; endale.anberber@aau.edu.et
7 Department of Pathology, School of Medicine, Addis Ababa University, Addis Ababa 9086, Ethiopia; yonas.bekuretsion@aau.edu.et
8 Department of Oncology, School of Medicine, Addis Ababa University, Addis Ababa 9086, Ethiopia; mathewos.assefa@aau.edu.et
9 Institute of Pathology, Martin Luther University Halle-Wittenberg, 06097 Halle (Saale), Germany; marcus.bauer@uk-halle.de
10 Department of Gynecology, Martin Luther University Halle-Wittenberg, 06097 Halle (Saale), Germany; martina.vetter@uk-halle.de
11 Institute of Medical Epidemiology, Biostatistics, and Informatics, Martin Luther University Halle-Wittenberg, 06097 Halle (Saale), Germany
* Correspondence: astarlar@uthsc.edu; Tel.: +1-901-448-3085
† These authors contributed equally to this work.

check for updates

Citation: Desalegn, Z.; Smith, A.; Yohannes, M.; Cao, X.; Anberber, E.; Bekuretsion, Y.; Assefa, M.; Bauer, M.; Vetter, M.; Kantelhardt, E.J.; et al. Human Breast Tissue Microbiota Reveals Unique Microbial Signatures that Correlate with Prognostic Features in Adult Ethiopian Women with Breast Cancer. *Cancers* **2023**, *15*, 4893. https://doi.org/10.3390/cancers15194893

Academic Editor: Robert-Alain Toillon

Received: 12 September 2023
Revised: 2 October 2023
Accepted: 5 October 2023
Published: 9 October 2023

Simple Summary: Breast cancer (BC) is the leading cause of cancer deaths among adult women in Ethiopia. The death toll associated with breast cancer is high among women of African ancestry. The cause of the disparity in mortality is unclear. Recently, studies conducted in the United States and other high-income countries highlighted the role of microbial dysbiosis in breast cancer initiation, growth and treatment outcome. However, whether differences in abundance and composition of microbiota are associated with clinical and histopathological parameters in Ethiopian women has not been studied. The aim of our study was to conduct microbial profiling on breast tumor and normal adjacent tissues of the same donor. Further, the study aimed to identify the association of the differences in microbial composition and abundance with clinicopathological factors in Ethiopian women with breast cancer. This is the first study to report an association between breast microbial dysbiosis and clinicopathological factors in Ethiopian women.

Abstract: Breast cancer (BC) is the leading cause of cancer mortality among women in Ethiopia. Overall, women of African ancestry have the highest death toll due to BC compared to other racial/ethnic groups. The cause of the disparity in mortality is unclear. Recently, studies conducted in the United States and other high-income countries highlighted the role of microbial dysbiosis in BC initiation, tumor growth, and treatment outcome. However, the extent to which inter-individual differences in the makeup of microbiota are associated with clinical and histopathological outcomes in Ethiopian women has not been studied. The goal of our study was to profile the microbiome in breast tumor and normal adjacent to tumor (NAT) tissues of the same donor and to identify associations between

microbial composition and abundance and clinicopathological factors in Ethiopian women with BC. We identified 14 microbiota genera in breast tumor tissues that were distinct from NAT tissues, of which *Sphingobium*, *Anaerococcus*, *Corynebacterium*, *Delftia*, and *Enhydrobacter* were most significantly decreased in breast tumors compared to NAT tissues. Several microbial genera significantly differed by clinicopathological factors in Ethiopian women with BC. Specifically, the genus *Burkholderia* more strongly correlated with aggressive triple negative (TNBC) and basal-like breast tumors. The genera *Alkanindiges*, *Anoxybacillus*, *Leifsonia*, and *Exiguobacterium* most strongly correlated with HER2-E tumors. Luminal A and luminal B tumors also correlated with *Anoxybacillus* but not as strongly as HER2−E tumors. A relatively higher abundance of the genus *Citrobacter* most significantly correlated with advanced-stage breast tumors compared to early-stage tumors. This is the first study to report an association between breast microbial dysbiosis and clinicopathological factors in Ethiopian women.

Keywords: breast microbiome; microbiota; Ethiopia; African ancestry; BC; PAM50 molecular intrinsic subtypes

1. Introduction

Globally, female breast cancer (BC) is becoming the most common malignancy [1]. In Ethiopia, it is the leading cause of cancer-morbidity, attributable to one-third of the total cancers occurring in women [2,3]. Annually, the incidence of BC accounts for a total of 16,133 cases and approximately 9000 deaths in Ethiopia [4]. It has been shown that more than 80% of cancer cases in Ethiopia are identified at an advanced stage, which has been attributed to a lack of resources necessary for early BC detection and prevention strategies [5]. Though there are additional unforeseen factors contributing to BC risk, the increasing incidence of BC in sub-Saharan Africa may be associated with a dynamic change in lifestyle behaviors, reproductive factors, and population aging [6–8].

Recently, scientific evidence demonstrated that microbial dysbiosis in the breast microenvironment may contribute to BC development [9–13]. Consequently, microbiota inhabiting breast tissue and/or the tumor microenvironment (TME) have a potential biological function in mediating carcinogenesis in breast tissue [10,14]. Variations in composition and functionality of the microbiota among BC cases in relation to healthy controls encourage the potential development of microbiome-derived biomarkers and future targeted interventions which ultimately have a central role in reducing the burden of BC [15].

Indeed, a combination of genetic, environmental, and lifestyle factors have been linked to BC [14], in which bacterial communities within the host could be one additional environmental factor associated with BC [16]. In general, supporting evidence from epidemiological studies confirms that various microbial species and/or their metabolites contribute to at least 16% of malignancies that occur across the globe [17]. However, it is unclear as to how microbial dysbiosis contributes to BC development. It is hypothesized that microbiota may utilize different mechanisms to promote cancer development, such as modulating inflammation via mediators and inflammatory response signaling pathways [18], triggering DNA damage [19], and/or releasing harmful gut microbiota-derived metabolites that mediate tumorigenesis or tumor suppression [20,21].

Regardless of factors including sample collection site, age of patients with BC, geographical variation, history of pregnancy, presence/absence of breast malignancy, method of DNA preparation, and sequencing technologies, the composition of mammary microbiota appeared diverse and different when compared with other body sites [13,22]. The unique breast microbiota pattern of healthy women demonstrated that the predominant phyla were Proteobacteria, Firmicutes, Actinobacteria, and Bacteriodetes in decreasing order, respectively. Though evaluation of microbiota in breast tumor tissue has recently received much attention, previous research works highlighted distinct microbiota in human milk which has already been established for several years. Similarity in composition

between human milk and breast tissue microbiota were noted, and Proteobacteria was identified as the most abundant phylum [18,19].

Furthermore, tumor microenvironment (TME) has become an important focus in understanding BC development and progression, and in determining responses to treatment over the last couple of decades [23]. The breast tumor microbiome has been shown to play an important role in modifying the TME, thereby potentially impacting treatment outcome [24]. In fact, it is recognized that the TME is an integral component when investigating the tumor microbiome [25]. This is further highlighted in recent studies that observed differences in breast microbial composition and density between NAT and breast tumor tissues and/or healthy controls [12,26].

Although the microbiomes of populations from developed countries have been extensively characterized, evidence-based data using African populations to address this topic remain limited [12,27–29], yet collecting such data has to be a high priority. Furthermore, to our knowledge, no microbiome-based BC studies have been conducted in Ethiopian women. It is critical for breast microbiome studies to be replicated in Africa due to inter-individual and population variability in cultural behaviors, environmental exposures, including infectious diseases, and genomic heterogeneity [30–34]. Therefore, this study was carried out to investigate microbiota abundance among patients with BC and explore its correlation with clinical and histopathological features in Ethiopian women with BC.

2. Materials and Methods

2.1. Patient Enrollment and Tissue Collection and Processing

Pathologically confirmed BC patients (age 16 to 83) were enrolled into the study from select hospitals in Addis Ababa, Ethiopia, according to the set inclusion and exclusion criteria. All study participants who were taking antibiotics were excluded from this study. BC patients were approached by physicians while attending for routine services and informed about the objective and procedures of data collection. After securing the written informed consent, we applied aseptic procedures and employed a standard protocol to obtain a fresh frozen surgically resected breast tissue specimen from patients with BC who underwent mastectomy. The specimens constituted a total of 100 surgically obtained tissues composed of breast tumor (T) tissue and normal adjacent tissue (NAT) pairs from the same donor (n = 50) and were included in this study for comparison from women of Ethiopian descent. Normal breast tissue immediately adjacent (up to 5 cm) to the collected breast tissue sample was evaluated and confirmed by a pathologist to be histologically free of any tumor cells or lesions. Upon receipt, breast tissue samples (n = 100) were immediately transferred to a sterile container containing liquid nitrogen and maintained at −80 °C until further processing. Clinical and histopathological data about the donor breast tissue specimen, including hormone receptor (HR) status, Ki-67 proliferation index, tumor grade, and stage of BC was obtained via medical chart review and pathological evaluation. This study was carried out according to the Declaration of Helsinki, and the research protocol was ethically approved by the Institutional Review Board of the College of Health Sciences, Addis Ababa University (IRB protocol #092/17/17), and at the University of Tennessee Health Science Center (IRB #19-06743).

2.2. Immunohistochemistry (IHC) and Gene Expression Analysis

Samples of the fresh frozen tissues were embedded and analyzed for the expression of estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor receptor 2 (HER2), and Ki-67 proliferation index using immunohistochemistry (IHC). IHC analyses was carried out using specific antibodies directed against the respective proteins as presented in Table 1.

Table 1. IHC analysis with specific antibodies directed against ER, PR, HER2 and Ki-67.

Protein	Clone	Manufacturer	Host	Dilution
ER	Clone Ab-11	Thermo Scientific; cat. no MS-354-P1 (Waltham, MA, USA)	mouse	1:150
PR	Clone PgR 636	DAKO; cat. no M3569 (Santa Clara, CA, USA)	mouse	1:100
HER2	Clone DG44	DAKO; cat. no SK001	rabbit	RTU
Ki-67	Clone SP6	Thermo Scientific; cat. no RM-9106-S	mouse	1:250

Following the IHC analysis, the immune reactive score reading was recorded and the status of ER and PR expression interpreted according to established guidelines [35].

In general, HR positivity was declared with IRS > 0 whereas negativity was defined as IRS < 1. HR status was declared positive when either ER and/or PR tended to be positive. ASC-CAP guidelines were employed to assess the status of HER2 [36]. On the other hand, the Ki-67 proliferation index was scored as follows:

- 'Low proliferation index:' when Ki-67 staining was positive in <20% of tumor cells; or
- 'High proliferation index': when at least 20% of tumor cells stained positive [37,38].

Histological grading was performed using the Elston–Ellis grading system [39].

2.3. RNA Extraction, Gene Expression Profiling, Normalization, and Intrinsic Subtyping

2.3.1. RNA Extraction

First the pathologist identified the tumor-enriched area on HE-stained slides. Then, multiple sections consisting of 10 µm breast tissue were utilized for RNA extraction. RNA was extracted using the miRNeasy FFPE Mini Kit®(Qiagen) following the manufacturer's recommendations. Before and after extraction, we decontaminated the workbench with disinfectant and chemicals to ensure a nuclease-free environment and to maintain the quality of quality of RNA extraction and further downstream applications. After measuring the concentration and the quality using a nanophotometer, the extracted RNA was stored at −80 °C for further downstream applications.

2.3.2. NanoString and PAM50 Assay

To determine intrinsic subtypes, a NanoString nCounter® Analysis System (NanoString Technologies, Seattle, WA, USA) was employed. We measured the relative gene expression using the PAM50 method, a multiplexed hybridization assay, where digital readouts of fluorescent barcoded probes hybridize with each mRNA sequence of interest. nCounter® Digital Analyzer was used for the acquisition of data. Importing data, verifying quality control, and normalization of expression levels were performed using nSolver software version 4. Furthermore, negative input controls were utilized for background subtraction from raw transcript counts. A total of six reference-control genes were used to perform reference-normalization by dividing the geometric mean of these genes. Thereafter, the normalized mRNA counts were log2-transformed. Finally, the distinct intrinsic subtype classification was estimated using the nearest PAM50 centroid algorithm in Bioclassifier and NanoStringNorm implemented in R [40] and according to the classification method as described previously [41].

2.4. DNA Extraction, PCR, and 16S rRNA Gene Sequencing and Processing

Using aseptic techniques, we isolated DNA from breast tissues under a sterile laminar flow hood using the Qiagen DNA Isolation kit (Qiagen, Germantown, MD, USA). We decontaminated the workbench with disinfectant and chemicals to ensure a nuclease-free environment. DNA quality and quantity was measured using the Nanodrop and a Quant-iT PicoGreen dsDNA Assay Kit (Invitrogen, Carlsbad, CA, USA). DNA samples were then shipped off on dry ice to Microbiome Insights (Vancouver, BC, Canada) for further

processing. Specimens were gently placed into a MoBio PowerMag Soil DNA Isolation Bead Plate, which is needed for further extraction of microbial DNA and for the isolation of nucleic acids without the binding of residual contaminants and purified on a KingFisher robot following MoBio's instructions. Bacterial 16S rRNA genes were PCR-amplified using dual-barcoded primers that target the V4 region of the bacterial genome, as per the protocol of Kozich et al. [42]. Amplicons were further sequenced using the 300-bp paired-end kit (v.2) on an Illumina MiSeq instrument. The potential for contamination was addressed by co-sequencing DNA amplified from specimens and from template-free negative controls and DNA extraction kit reagents processed the same way as the specimens. A cloned SUP05 DNA sample served as a positive control. Resulting sequences were filtered, then trimmed and denoised, and used to infer amplicon sequence variants (ASVs) exactly using the DADA2 sequencing data tool [43]. Subsequent ASVs were merged to create a sequence table, then chimeras were removed and taxonomically classified using the Ribosomal Database Project's Training Set 16 (11.5 release) as the reference database by the naive Bayesian classifier method implemented in DATA2. The Phyloseq program [44] was further used to import, store, analyze, and graphically display complex phylogenetic sequencing data.

2.5. Statistical Analysis

We estimated alpha diversity, defined as the observed richness or number of taxa of each sample, using the Shannon index on raw ASV abundance tables after filtering out all contaminants. The Simpson index is another indicator we used to further estimate microbial diversity and to evaluate species evenness. We conducted a Kruskal–Wallis test or Wilcoxon signed rank test to determine significant differences in diversity. To estimate beta diversity across samples, we excluded ASVs occurring with a count of less than 5 (5% of number of samples). We then computed Bray–Curtis indices. Beta diversity, defined as the variability in community composition or identity of taxa observed among the samples, was visualized using principal coordinate analysis (PCoA) ordination and emphasized differences across samples. Variation in community structure was assessed with permutational multivariate analysis of variance (PERMANOVA) with groups/subtypes as the main fixed factor and using 999 permutations for significance testing [45]. All analyses were conducted in the R environment.

3. Results

3.1. Clinicopathological Characteristics of the Study Participants

This study aimed to enhance our understanding of breast microbial abundance and composition in Ethiopian women with BC. Our study included fresh frozen tumor tissue and normal adjacent to tumor (NAT) specimens from 50 Ethiopian women diagnosed with BC, for a total of 100 breast tissue samples analyzed. Table 2 presents the clinicopathological characteristics of the study participants. The mean age of patients in this study was 46 ± 1.92 years. Of these women, 60% were premenopausal, 50% were with confirmed early stage at the time of diagnosis, and almost equal proportions of the tumors were intermediate to poorly differentiated (G2, G3). With regards to their IHC groups, 44 (88%) were HR+. Of the breast tumor tissues, 26% were luminal A, and 17% were luminal B.

3.2. Breast Tumor Tissue Exhibits Distinct Microbiome Composition from NAT Tissue of the Same Woman

In this study, we were able to sequence the 16s rRNA amplicon in 50 paired samples from tumor tissue and in NAT tissue. We sequenced 16Sv4 amplicons generated from human breast tissue samples and minimized the potential of bacterial contaminants by using multiple controls alongside the tissue samples. After quality-filtering and inferring amplicon sequence variants (ASVs), quality-filtered reads accounting for an average of 1569 were generated per sample and used for further analysis. Rarefaction curves relating

number of sequencing reads compared to the number of ASVs or genera are shown in Figure S1A,B.

Table 2. Patient and tumor characteristics.

Variable	Frequency (%)
Mean age (yrs ± SE)	46 ± 1.92
Minimum	27
Maximum	83
Menopausal status	
Pre	30 (60)
Post	17 (34)
Unknown	3 (6)
UICC stage	
Early (0–2A)	25 (50)
Advanced (2B–4)	17 (34)
Unknown	8 (16)
Grade	
1	0 (0)
2	24 (48)
3	26 (52)
IHC type	
HR+/HER2−	30 (60)
HR+/HER2+	14 (28)
HR−/HER2+	3 (6)
HR−/HER2−	3 (6)
ER status (IHC)	
Positive (≥**1%**)	44 (88)
Negative (<1%)	6 (12)
PR status (IHC)	
Positive (≥**1%**)	29 (58)
Negative (<1%)	21 (42)
HER2 status (IHC)	
Positive	17 (34)
Negative	33 (66)
Ki-67 proliferation index (IHC)	
Low (≥20%)	16 (32)
High (<20%)	34 (68)
Intrinsic subtype	
Luminal A	13 (26)
Luminal B	7 (17)
HER2-E	7 (14)
Basal-like	5 (10)
Unknown	18 (36)

In this study, alpha diversity was quantified to correlate the differences in breast microbial diversity between the tumor and NAT breast tissues. Interestingly, the Shannon index showed alpha diversity was not significantly different between tumor and NAT tissues ($p = 0.07$) (Figure 1A); however, alpha diversity measured by the Simpson index revealed slightly higher alpha diversity in tumors compared to NAT ($p = 0.043$). To determine differences in beta diversity, we visualized the overall differences between the breast microbiome profile of tumor and NAT tissues using principal coordinate analysis (PCoA) of Bray–Curtis dissimilarity (Figure 1B). The tumor tissue and NAT clustered significantly differently between the two groups ($R^2 = 0.664$; $p = 0.049$) with greater dissimilarity along PC1 (23.1% variation). We also observed overlap between the two groups.

We then determined the taxonomic composition and abundance of breast microbial profiles of tumor and NAT tissues. The 16S rRNA based sequencing identified 5 phyla, 7 classes, 16 orders, 16 families, and 16 genera across all the breast tissue samples. Across all tissue types, the three predominant phyla were Proteobacteria (48.4%), followed by Firmicutes (22.1%) and Actinobacteria (15.0%) as shown in Figure S2A. Alphaproteobacteria

(41.5%), Bacilli (15.2%), Actinobacteria (15.0%), and Bacteroidia (7.33%) were the most abundant classes among all tumor and NAT tissue samples (Figure S2B). Considering all breast tissue samples, the most prevalent families were *Sphingomonadaceae* (20.78%), *Staphylococcaceae* (13.2%), *Beijerinckiaceae* (13.1%), and *Corynebacteriaceae* (9.44%) (Figure S2C). We also observed that 64.1% of microbial genera were shared between NAT and tumor groups. However, 26.1% of genera were unique to NAT tissues, whereas approximately 10% were unique to tumors (Figure 1D). Genera most predominant in tumor compared to NAT tissues included *Sphingobium* ($p < 0.0001$), *Anaerococcus* ($p < 0.0001$), *Corynebacterium* ($p = 0.0012$), and *Delftia* ($p = 0.0031$) (Table S1).

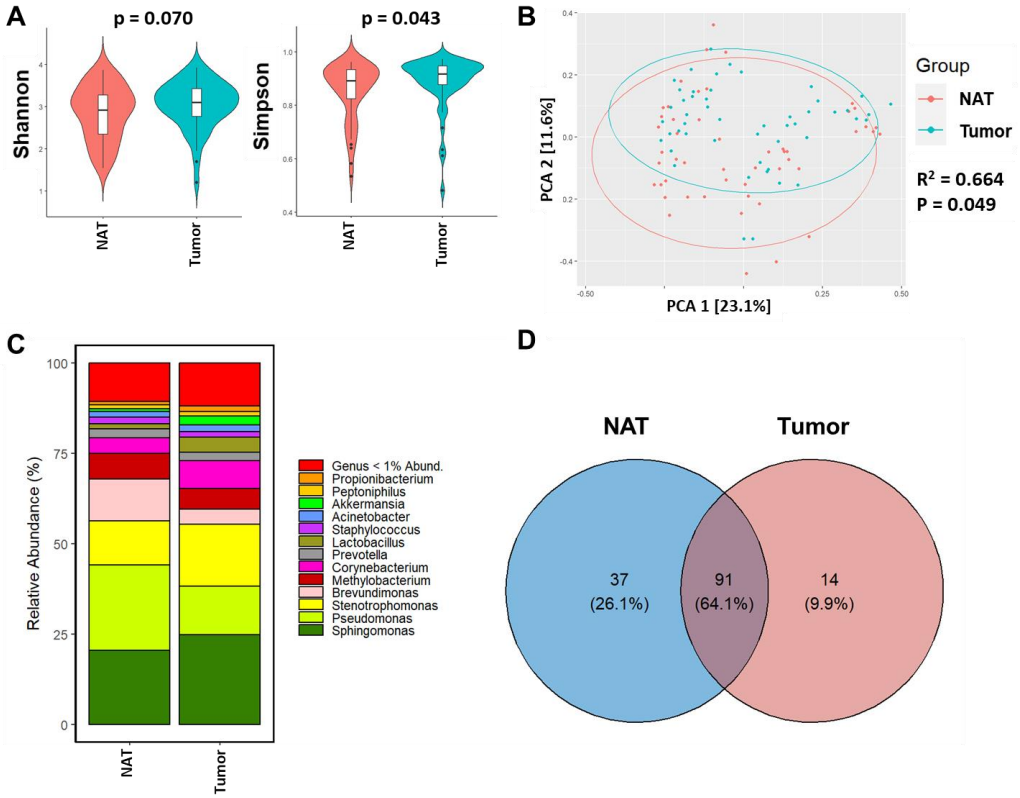


Figure 1. Breast bacterial community composition varies by patient BC status and normal adjacent tumor (NAT) and tumor tissue types. (**A**) Violin plots show median and interquartile range in bacterial alpha diversity as measured by Shannon and Simpson diversity indices within (tumor and NAT) breast tissue from Ethiopian BC patients. *p*-values were obtained from Kruskal–Wallis tests. (**B**) Principal coordinates (PC) plots show the clustering pattern of tumor and NAT based on unweighted UniFrac distance and are colored by sample types (red circles—NAT, teal—Tumor samples); $p = 0.049$ and $R^2 = 0.664$. (**C**) Taxonomic composition of the breast microbiome depicted as proportional abundances at the genus level for NAT and corresponding tumor tissue of the same donor. (**D**) Venn diagram showing the unique and shared bacterial genera between the tumor and normal groups.

3.3. Breast Microbial Communities Differ by Breast Tumor IHC Types, PAM50 Intrinsic Subtypes, and Stage of Disease

We next determined whether microbial composition and abundance were associated with breast tumor prognostic features, including by IHC group and PAM50-based intrinsic subtypes. The Shannon index ($p = 0.717$) and Simpson index ($p = 0.748$) did not differ significantly by HR status, respectively (Figure 2A). Proteobacteria was the predominant microbiota observed followed by Firmicutes (Figure 2B). The predominant bacterial phylum, family, and genus across the four IHC breast tumor groups were Proteobacteria, *Sphingomonadaceae*, and *Sphingomonas*, respectively, as shown in Figure 2B–D. The relative abundances of these phyla differed by HR status, with TNBC tumors having the lowest

abundance of Proteobacteria compared to the other tumor types. Microbial composition at the lower taxonomic level (family and genus) were more abundant and/or diverse in HR+ HER- and HR + HER2+ tumors compared to HR− and TNBC. We did further analysis to understand differences in relative abundances of the bacteria at the genus level among the different IHC types using a linear model. We found that *Exiguobacterium* ($p = 0.007$), *Varibaculum* ($p = 0.0087$), and *Leifsonia* ($p = 0.0016$) were significantly less abundant in HR+/HER2−, HR+/HER2+ and TNBC tissues compared to HR − HER2+ tumors (Table S2).

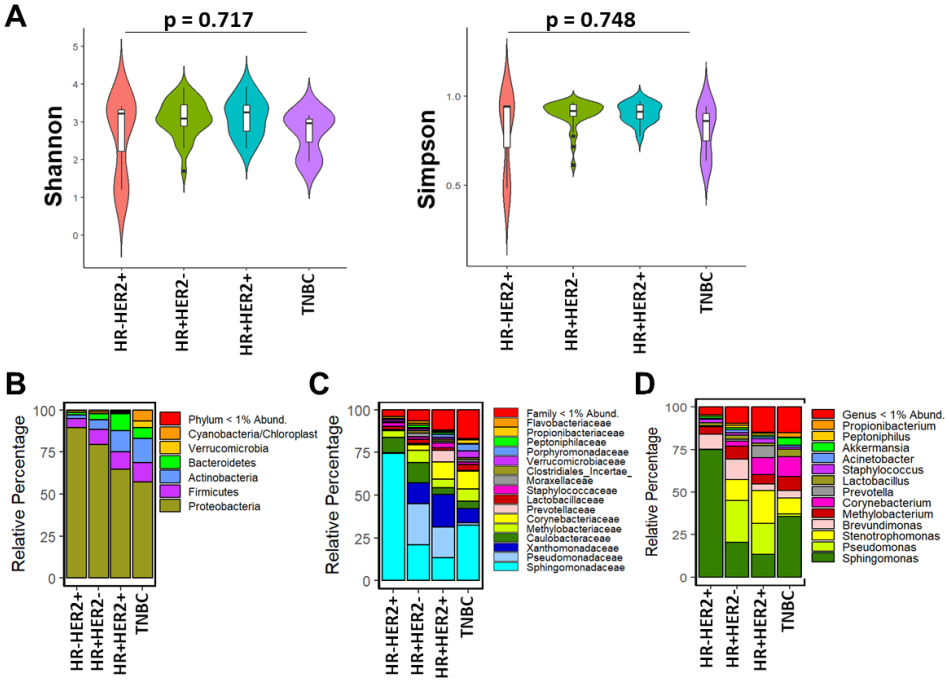


Figure 2. Breast bacterial community composition varies by immunohistochemical (IHC) status. (**A**) Violin plots show median and interquartile range as measured by Shannon and Simpson diversity indices with regard to hormone receptor (HR) and HER2 status of the tumor tissue. *p*-value results were from Kruskal–Wallis tests. Taxonomic composition of the breast microbiome is depicted as relative percentages at the (**B**) phylum, (**C**) family, and (**D**) genus level.

Since we observed significant differences in the relative percentage of microbiota according to HR status, we hypothesized that microbial differences exist across intrinsic subtypes. Similar to IHC type, Shannon and Simpson indexes were not significantly different between the four intrinsic subtypes (p-value = 0.329 vs. p-value = 0.310, respectively) (Figure 3A), which was also observed among the PCA analysis ($R^2 = 0.779$, $p = 0.378$) (Figure 3B). Furthermore, Proteobacteria were the predominant bacterial phylum across the four intrinsic breast tumor subtypes followed by Firmicutes (Figure 3C), whereas at the family level, *Sphingomonadaceae* were predominant in luminal A, HER2-E and basal-like breast tumor intrinsic subtypes (Figure 3D). In the luminal B subtype, the relative abundance of such microbiota was lower, including *Sphingomonadaceae*; however, *Pseudomonadaceae* and *Xanthomonadaceae* were relatively higher (Figure 3D). At the lower taxonomic level (genus), our linear regression analysis revealed differences in abundance of microbial genera by intrinsic subtype. As shown in Figure 3E, statistically significant abundance of *Polaromonas* ($p = 0.0005$), *Varibaculum* ($p = 0.0087$), *Exiguobacterium* ($p = 0.0097$), *Bifidobacterium* ($p = 0.0382$), and several other genera between luminal A, HER2-E, and basal-like tumors were observed overall when compared to luminal B tumors (Table S3).

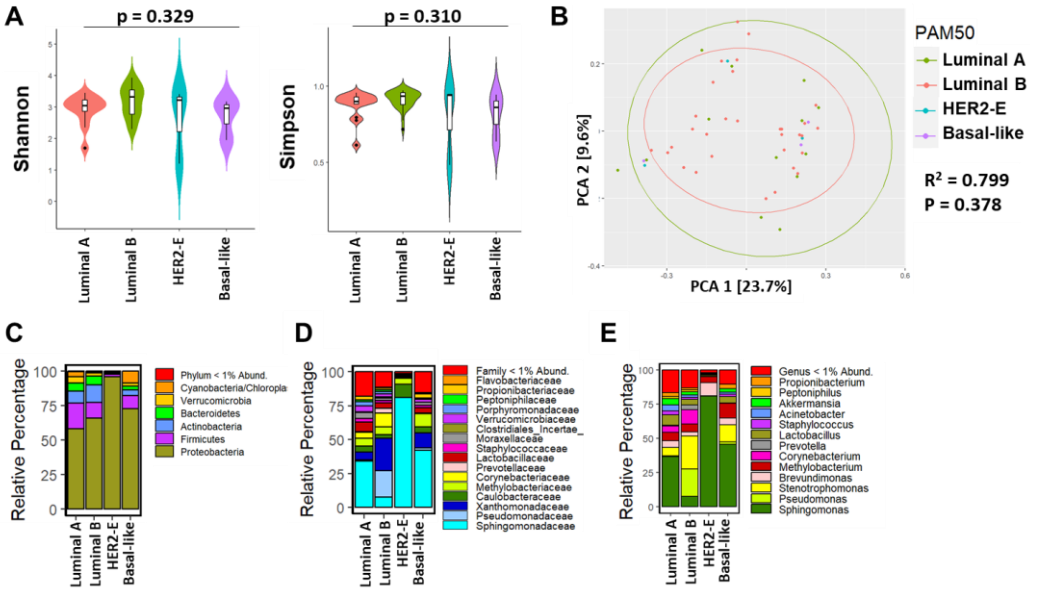


Figure 3. Breast bacterial community composition varies by PAM50 intrinsic subtypes. (**A**) Violin plots show median and interquartile range as measured by Shannon and Simpson diversity indices within breast tissue from Ethiopian BC patients. *p*-value results were obtained from Kruskal–Wallis tests. (**B**) PC plot shows the clustering pattern of the intrinsic subtypes based on unweighted UniFrac distance and is colored by sample types (green—luminal A, red circles—luminal B, teal—HER2E, and purple circles—basal-like tumor samples); $p = 0.378$ and $R^2 = 0.799$. Taxonomic composition of the breast microbiome, depicted as relative percentage at the (**C**) phylum, (**D**) family, and (**E**) genus level.

We next investigated the relationship between tumor stage and the relative abundances among microbial taxa. Shannon ($p = 0.475$) and Simpson ($p = 0.358$) indexes were not significantly different between early and advanced stage of disease, respectively (Figure 4A). PCA showed no significant separation between early and advanced stages of BC ($R^2 = 0.533$, $p = 0.16$). However, visually there is a clear separation between the two groups. There were eight patients with unknown stage information which did not significantly change the results (Figure 4B). At the phylum level, Proteobacteria dominated (Figure 4B). Similar to HR status and intrinsic subtype, relative abundance of family *Sphingomonadaceae* and genus *Sphingomonas* dominated in early-stage (T0-2A) compared to more advanced-stage tumors (T2B-4) (Figure 4C,D). Specifically, genus *Stenotrophomonas* ($p = 0.015$), *Corynebacterium* ($p = 0.049$), *Prevotella* ($p = 0.024$), *Actinomyces* ($p = 0.021$), and several additional genera were significantly more abundant in advanced-staged tumors compared to early-stage tumors (Table S4).

3.4. Breast Microbial Communities Correlate with Clinicopathological Features in BC

Lastly, we determined whether microbial communities correlated with certain clinicopathological features, including IHC group, intrinsic subtype, and stage, as observed in our study. In HR−/HER2+ BC, *Alkanindiges* and *Anoxybacillus* strongly correlated with this IHC group, whereas *Rhodopseudomonas* correlated strongly with HR+ HER2+ BC. Only the genus *Anoxybacillus* slightly correlated positively with HR + HER2- tumors. Interestingly, TNBC tumors correlated strongly with the genus *Burkholderia*, followed by *Thermicanus*, *Paracoccus*, *Mogibacterium*, and *Aeromonas* (Figure 5A).

Many of the same microbial patterns observed by IHC group were also found among the different intrinsic molecular subtypes. For instance, similar to HER2+ tumors, HER2-E tumors correlated strongly with *Alkanindiges* and *Anoxybacillus* (Figure 5B). Likewise, basal-like tumors correlated with the same genera as TNBC, with *Burkholderia* being the most strongly correlated genus. There was a weak association with *Cuprivadus* in basal-like tumors (Figure 5B) that was not observed in TNBC tumors using IHC determination

(Figure 5A). Luminal A and luminal B tumors correlated weakly with *Anoxybacillus*. Furthermore, there was not much of a difference in genera among the early- and advanced-stage BC, although a stronger correlation with the genus *Citrobacter* was observed in advanced breast tumors (Figure 5C).

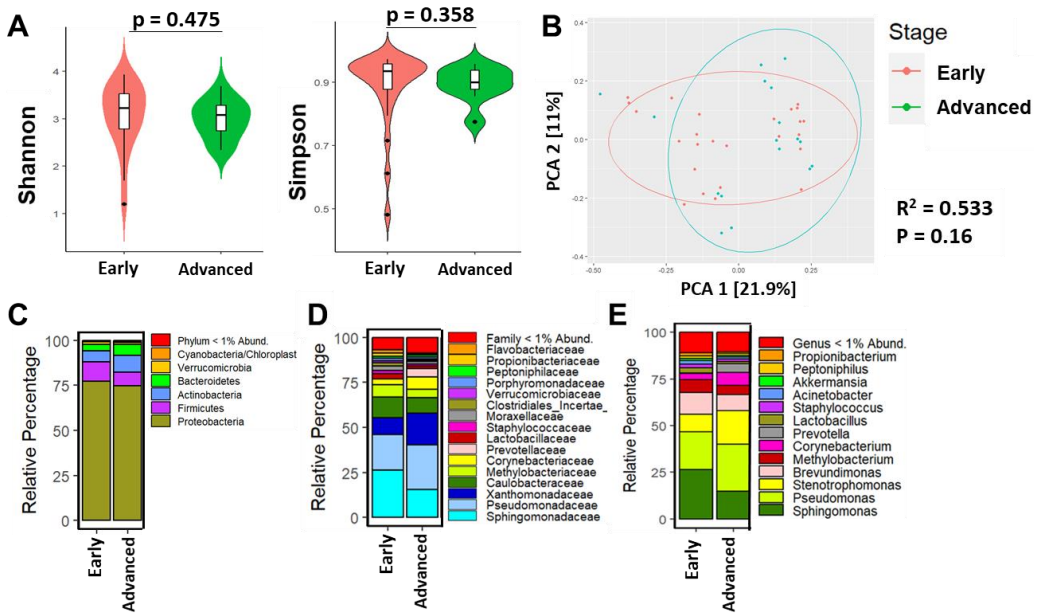


Figure 4. Breast bacterial community composition varies between early and advanced stages of BC. (**A**) Violin plots show median and interquartile range as measured by Shannon and Simpson diversity indices based upon breast tumor stage. *p*-value results were obtained from Kruskal–Wallis tests. (**B**) PC plots show the clustering pattern among early (red circles), advanced (green circles), and unknown (blue circles) tumor stage based on unweighted UniFrac distance; $p = 0.16$ and $R^2 = 0.533$. (**B**) Taxonomic composition of the breast microbiome, depicted as relative percentage at the (**C**) phylum, (**D**) family, and (**E**) genus level.

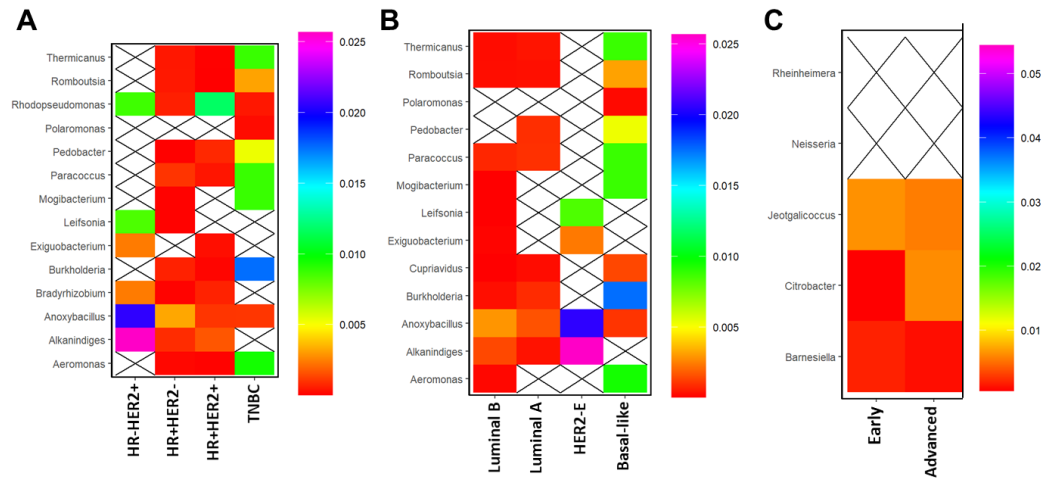


Figure 5. Specific bacterial genera correlate with clinicopathologic features. Mean relative abundances (proportions) of bacterial genera that were differentially present in breast tumors stratified by (**A**) IHC type (**B**) PAM50 breast tumor subtype, and (**C**) early or advanced stage. Crossed-out boxes indicate samples for which specific genera were not detected. Color bars vary on a logarithmic scale. All genera shown had FDR-corrected *p*-value < 0.05 by Kruskal–Wallis H-test.

In summary, we identified 14 microbiota genera in breast tumor tissues that were distinct from NAT tissues, of which *Sphingobium*, *Anaerococcus*, *Corynebacterium*, *Delftia*, and *Enhydrobacter* were most significantly decreased in breast tumors compared to NAT

tissues. Several microbial genera also significantly differed by clinicopathological factors in Ethiopian women with BC. Specifically, genera *Burkholderia*, followed by *Thermicanus*, *Paracoccus*, *Mogibacterium*, and *Aeromonas*, more strongly correlated with aggressive triple negative (TNBC) and basal-like breast tumors compared to less aggressive luminal A, luminal B, and HER2-E tumors. The genera *Alkanindiges*, *Anoxybacillus*, *Leifsonia*, and *Exiguobacterium* most strongly correlated with HER2-E tumors. Luminal A and luminal B tumors also correlated with *Anoxybacillus* but not as strongly as HER2-E tumors. A relative higher abundance of the genus *Citrobacter* most significantly correlated with advanced-stage breast tumors compared to early-stage tumors. This is the first study to report an association between breast microbial dysbiosis and clinicopathological factors in Ethiopian women.

4. Discussion

Women in Ethiopia, like women in other lower middle income countries (LMIC), have poorer overall BC survival rates [46,47]. Stage at diagnosis as well as availability of optimal care were among the factors associated with poor BC survival outcomes. Additionally, the role of biological factors including the tumor microenvironment (TME) that includes the microbiota have been shown to be associated with poor clinical outcomes [48].

BC is a leading cause of cancer morbidity among women of African ancestry. This is further compounded by underrepresentation of women of African ancestry in BC research studies. Therefore, we intended this study to generate evidence-based preliminary data to profile the microbiome in breast tumor and normal adjacent tissues of the same donor and to identify the association between differences in microbial composition and abundance and clinicopathological factors in Ethiopian women with BC. To this end, we used fresh frozen breast tumor and NAT tissues to profile microbial composition and abundance and to determine their association with clinicopathological features observed in BC.

4.1. Findings of Studies on Breast Tissue Microbiome

In recent years, research findings have begun to reveal the potential role of mammary microbiota in mediating BC development [49]. Current research findings indicate that the microbiota among BC patients differs from that of healthy women [14]. Supporting evidence reported a distinct mammary microbiota composition that also differed between breast tumor and normal breast tissues [11]. Such evidence points to distinct microbial signatures that might be related to cancer development and response to certain treatment [14]. However, the extent to which microbiota alterations (dysbiosis) contributes to BC development is unclear. Although the discovery of microbiota differences in breast milk was previously explored, recent works have begun to investigate the association between breast tissue microbial dysbiosis with cancer. Accordingly, these efforts indicate that healthy breast and BC tissues are composed of a unique microbiota in which the phylum Proteobacteria predominated in healthy breast tissues, followed by Firmicutes [11,50].

In agreement with our study, previous studies demonstrated that the breast is characterized by a high predominance of Proteobacteria, followed by the phyla Firmicutes and Actinobacteria [11,49,50]. In another study, Proteobacteria were prominently identified in the tumor tissues and conversely, in non-cancerous adjacent tissues, Actinobacteria abundance was increased [51]. Furthermore, previous studies highlighted the difference in breast microbial composition and abundance between normal adjacent and tumor tissues and/or healthy controls [12,26]. These microbes thrive in a fatty-acid-rich environment in the breast and are positively associated with adiposity [52]. Thus, it is not surprising that these bacteria are abundant in the breast. At the family level, our findings are in agreement with Klann et al., who also observed a higher relative abundance the family *Ruminococcaceae* and the genus *Akkermansia* in breast tumor tissue as compared to normal tissue and a lower relative abundance of *Bacteroides* and *Sutterella* in breast tumor compared to normal tissue [53]. More recently, exciting research on the breast microbiome to identify distinct microbial signatures considering different parameters, including race of patients

with BC, tumor stage, and breast tumor subtype, was carried out [28]. In agreement with our findings, it was observed that the phylum Proteobacteria was most abundant across all tissues, followed by Firmicutes, Bacteroidetes, and Actinobacteria regardless of race, and a higher abundance of the genus *Ralstonia*, which could explain in part a portion of the BC racial disparities [28]. Considering the diverse population in the Ethiopian context, conducting a study to profile microbiota and understanding its association with certain prognostic features would be of paramount importance in cancer management.

In efforts towards understanding the role of microbiota in health and disease, until now, it has remained elusive and/or unclear whether there is a difference in microbial composition in breast tumors and paired NAT of the same individual with BC [11,13]. While comparing tumor tissue with NAT, the microbiota composition showed a distinct bacterial profile, suggesting the oncogenic effect of specific bacterial taxa [50]. In terms of relative distribution of microbiota, our sequencing data of NAT and tumor groups are supported by another study where out of 11 differentially abundant operational taxonomic units (OTUs), the majority of OTUs were abundant in paired normal tissue and close to one-third of the OTUs were abundant in tumor tissue. The bacterium *Sphingomonas yanoikuyae* was enriched in paired normal tissue, similar to our findings, whereas the bacterium *Methylobacterium radiotolerans* was most significantly enriched and prevalent in paired in tumor tissue [13]. Microbial differences in paired breast tissues (tumor versus NAT) have also been identified in women from different geographical locations, which suggests that environment or lifestyle factors might impact the relative abundance of certain microbes in the breast microenvironment [11,54].

Generally, the breast TME is composed of a variety of cell types and microbiota. Studies suggest that pathophysiological changes occurring within the breast cells could have a significant impact on tumor growth [51]. Additionally, it has been discovered that certain microbial species identified in the human breast are potent agents that have the potential to trigger DNA damage, genomic instability, and alterations of molecular events in the form of mutations and epigenetic modifications [55]. For instance, functional in vitro studies revealed that *Escherichia coli* and *Staphylococcus epidermidis* isolated from BC patients trigger DNA double-stranded breaks in HeLa cells [22]. Therefore, microbial communities within a host could be considered an additional environmental factor that may contribute to or be influenced by carcinogenesis [56]. Considering that the data in this regard are in their infancy, precautions are required while inferring the significance of microbiota in the initiation, progression, and prediction of therapeutic responses among patients with BC.

4.2. Findings Associated with Clinical and Histopathological Features

In agreement with our findings, other studies reported a correlation between specific breast microbial taxa and certain prognostic features such as stage [12], receptor status [26,57,58], and lymphovascular invasion [26]. In our study, we found specific genera that were strongly associated with different intrinsic subtypes and by stage of disease. Specifically, we observed that the genus *Burkholderia* most strongly correlated with aggressive triple negative (TNBC) and basal-like breast tumors compared to less aggressive luminal A, luminal B, and HER2-E tumors. In support of our findings, a recent insightful study by Hoskinson et al. noted that a higher abundance of several microbial families, including *Burkholderiaceae*, was associated with early BC development when analyzing normal breast tissue compared to tissues donated prior to and after BC diagnosis [59]. These findings suggest that microbes in the genus *Burkholderia* may play an oncogenic role in the development of aggressive BCs, such as TNBC and basal-like tumors. In addition, we observed that several genera, including *Alkanindiges*, *Anoxybacillus*, *Leifsonia*, and *Exiguobacterium* most strongly correlated with HER2-E tumors. Luminal A and luminal B tumors weakly correlated with *Alkanindiges* and *Anoxybacillus*. Interestingly, Modica and colleagues reported that lymphovascular invasion correlated negatively with *Alkanindiges* in HER2-E tumors, which could explain the weak association between less aggressive luminal A and luminal B tumors in our study [60]. On the other hand, *Alkanindiges* was correlated with

invasive ductal carcinomas [61]. Furthermore, we observed a relative higher abundance of the genus *Citrobacter* most significantly correlated with advanced-stage breast tumors compared to early-stage tumors. A supporting study by Yang et al. reported that women with malignant BC exhibited enriched levels of *Citrobacter* in their gut compared to those with benign tumors where *Citrobacter* was further associated with elevated glycan and lipopolysaccharide biosynthesis [62].

With a comparatively large sample size aimed at identifying bacterial genera with statistically different abundances, another study identified *Porphyromonas*, *Lacibacter*, *Ezakiella*, and *Fusobacterium* as being more abundant in higher-stage compared with lower-stage tumors [61]. These findings highlight the complex nature of the microbiome and tumor interactions in the breast. Although we did not observe the distribution across tumor grade, in a separate study, multiple genera were significantly associated with histologic grade, with a number of genera present only in grade 1 tumors [61]. However, additional studies using a larger sample size and comparable numbers of different groups of BCs are needed to confirm these associations. Collectively, most findings consistent with our study revealed that the breast houses a unique microbiota that can be distinguished based upon hormone receptor status, molecular subtype, and stage of disease [11,13,50,51].

4.3. Problems with Breast Microbiome Studies

Normally, a wide range of research findings employ various platforms to investigate breast tissue microbial profiles. Consequently, comparing and interpretating findings from different studies can be challenging due to differences in sample retrieval, processing, and methods employed to prevent bacterial contamination. According to recent evidence, microbiota of BC tissues are quite different from microbiota in breast tissues of women without BC [26]. Though various studies have indicated that breast microbiota might play a role in mediating breast carcinogenesis [11,22,63,64], scientists are posing questions about how the microbiome might play a role in modulating the risk of BC development. The hypothesis is that changes in the composition of breast microbiota may also contribute to disease development and progression through several pathways, but it is still unclear whether the host´s microbial differences are a consequence or a cause of this human disease. This distinct phenomenon could be explained and/or interpreted in two ways: (1) alteration in microbiota profiles and/or dysbiosis comes first in the course of the carcinogenic process and can establish a microenvironment predisposed to cancer or (2) there is no correlation between the two events. However, knowledge of the microbiota of BC patients remains in its infancy. Given this context, additional studies have to be carried out to broaden knowledge on this topic and to understand which of the two possibilities occurs in patients with BC.

Considering the observed methodological differences in breast microbiome studies, including sample size, sample type, employed amplicon amplification technique, DNA extraction method utilized, sequencing approach, and potentially other factors, appreciating the precise difference in the composition and abundance of microbiota within a group of variables could be challenging in interpreting conclusions. However, in support of our findings, the literature indicates that differences in community composition across data sets can be attributed to ethnicity, dietary habits, geography, lactation status, method of sample collection, and platform of sequencing and data analysis [65].

In addition to the contradicting scenarios referring to the role of microbiota as either a cause or effect of cancer, the differences in sample collection and processing, methodological differences in sequencing, and variations in recruited patient groups, along with the distinct clinical and histopathological parameters, underrepresentation of some geographical regions, and patient-oriented studies only with few inconsistent and/or contradicting in vitro results, could ultimately make the generalization of the role of microbiota in health and disease difficult in a broader context. Therefore, further studies addressing all of the noted pitfalls are necessary, particularly in developing countries including Ethiopia to improve BC care.

4.4. Why Shannon Diversity Was Not Different between NAT and Tumor Tissues

In this study, Shannon diversity or richness of microbial taxa was not found to be significantly different between NAT and tumor tissues; however, microbial evenness defined by Simpson's index was slightly significant between the two groups. In support of our study, a previous study showed that adjacent normal paired breast tissue had a higher microbial diversity and richness than normal and tumor tissues [12]. Additionally, a previous study revealed that alpha diversity was significantly higher in normal compared to tumor samples where in unweighted UniFrac measures, breast tumor samples clustered distinctly from normal samples (R^2 = 0.130; p = 0.01) [53]. One study from Ghana revealed that alpha diversity was strongly and inversely associated with BC by tumor stage and molecular subtype [66]. In other studies, it was shown that BC patients had statistically significantly altered microbiota composition (beta diversity) and lower alpha diversity compared with healthy patients [59] or breast tissues [12,67], whereas another study reported that there was no difference in bacterial communities between tumor tissue and NAT [22], which may be explained by sample retrieval approaches [26,50]. Since we did not include breast tissue from women without BC, such a discrepancy in the relative alpha diversity between normal tissue and tumor samples is anticipated. Such a scenario will strengthen a well-articulated future study that includes diverse samples and specific patient groups.

4.5. Limitation of Our Study

Our findings have paved a way to understanding the composition of breast microbiota in Ethiopian BC patients. Despite the plethora of microbiome studies in developed countries, it has been shown that African populations are under-represented, and there is an acute demand for studies related to the microbiome among African populations [68]. Therefore, taking into account context-specific factors and considerable limited evidence, microbiome studies must be replicated in Africa to extend our understanding of BC and to identify potential biomarkers used for prognosis and predicting therapeutic response. We believe that having a larger sample size consisting of fresh frozen breast tissues from healthy controls (without BC) and BC patients with benign and malignant tumors might assist in drawing conclusions as to whether with microbial differences are a consequence or a cause of BC. Additional patient information on lifestyle factors, such as diet and environmental exposures, are also critical components for understanding the role of microbial dysbiosis on the breast microenvironment.

5. Conclusions

In summary, the current study, which utilized fresh frozen breast tumor tissues and NAT from the same women, reported for the first time a unique microbial signature that correlates with prognostic features, including stage, IHC status, and PAM50 intrinsic subtypes, among Ethiopian women with BC. We identified 14 microbiota genera in breast tumor tissues that were distinct from NAT tissues. *Burkholderia* most strongly correlated with aggressive triple negative (TNBC) and basal-like breast tumors, whereas *Alkanindiges*, *Anoxybacillus*, *Leifsonia*, and *Exiguobacterium* most strongly correlated with HER2-E tumors. Luminal A and luminal B tumors also correlated with *Anoxybacillus* but not as strongly as HER2-E tumors. A relatively higher abundance of the genus *Citrobacter* most significantly correlated with advanced-stage breast tumors compared to early-stage tumors. This is the first study to report an association between breast microbial dysbiosis and clinicopathological factors in Ethiopian women. Our findings encourage further precise characterization of local microbes in BC patients in Ethiopia for drug discovery and targeted microbial-based therapeutics, thus improving the prognosis and quality of life of BC patients. A future epidemiological study taking into account sample size, more controlled environmental conditions, high throughput metagenomics, and follow-up data is essential to strengthen conclusions related to clinical outcomes.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/cancers15194893/s1, Figure S1: Graphs show rarefaction curves from human breast tissue samples. (A) Rarefaction curve relating the number of sequencing reads compared to the number of amplicon sequence variants (ASVs). (B) Rarefaction curve relating the number of ASVs compared to the number of microbial genera. Figure S2: Bar plots illustrating the relative abundances of microbiota that differ between NAT and breast tumor tissues by (A) phylum, (B) class, and (C) family taxonomic levels. The unfilled portions of the bar plots represent lower-abundance taxa; Table S1: Significant genera by paired Wilcoxon signed-rank test in paired tumor relative to normal adjacent tissues; Table S2: Significant genera in tumors according to IHC status (linear model); Table S3: Significant genera in tumors according to PAM50 (linear model); Table S4: Significant genera that differ between advanced vs. early-stage tumors.

Author Contributions: Conceptualization, Z.D., A.S.-D. and T.A.; methodology, Z.D., X.C., T.A., A.S. and E.J.K.; software, X.C. and A.S.-D.; validation, Z.D., X.C., A.S.-D., X.C. and A.S.; formal analysis, X.C., Z.D., T.A., A.S. and A.S.-D.; investigation, Z.D., M.Y., A.S.-D., T.A., E.A., Y.B., M.A., A.S. and E.J.K.; resources, Z.D., T.A., A.S.-D., E.J.K. and M.V.; data curation, Z.D., X.C., A.S.-D., T.A., E.J.K. and M.V.; writing—original draft preparation, Z.D. and A.S.-D.; writing—review and editing, Z.D., A.S.-D., X.C., T.A., E.J.K., M.Y., E.A., M.A., M.B., Y.B. and M.V.; visualization, X.C., Z.D. and A.S.-D.; supervision, Z.D., A.S.-D., T.A. and E.J.K.; project administration, Z.D., A.S.-D., T.A. and E.J.K.; funding acquisition, Z.D., A.S.-D., T.A. and E.J.K. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by a grant from the Susan G. Komen organization assigned to the Medical Faculty of the Martin Luther University Halle-Wittenberg, Halle (Saale), Germany (Grant number GTDR16378013). EJK, MB, and AA received funding from the Else-Kroener-Foundation through Martin Luther University, Halle-Wittenberg, Germany (grant No. 2018_HA31SP). ZD was supported by Addis Ababa University. The APC was funded by Martin Luther University. AS-D was supported by funding from Methodist Mission at the University of Tennessee Health Science Center.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board (or Ethics Committee) of Addis Ababa University College of Health Sciences (protocol code number: 078/20/DMIP; Year of approval: 2020) and the University of Tennessee Health Science Center (IRB # 19-06743).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: All relevant datasets generated and/or analyzed during the current study are included in this article, and we may provide additional de-identifiable data upon request.

Acknowledgments: We would like to express our gratitude to the study participants and data collectors without whom it would have not been possible to realize the study. The authors are grateful to University of Tennessee Health Science Center, Athena Starlard-Davenport, and her research team for their valuable support throughout the process. Last but not least, our heartfelt thanks go to Addis Ababa University and Martin Luther University for providing partial financial support.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA. Cancer J. Clin.* **2021**, *71*, 209–249. [CrossRef] [PubMed]
2. Memirie, S.T.; Habtemariam, M.K.; Asefa, M.; Deressa, B.T.; Abayneh, G.; Tsegaye, B.; Abraha, M.W.; Ababi, G.; Jemal, A.; Rebbeck, T.R.; et al. Estimates of Cancer Incidence in Ethiopia in 2015 Using Population-Based Registry Data. *J. Glob. Oncol.* **2018**, *2018*, 1–11. [CrossRef] [PubMed]
3. Timotewos, G.; Solomon, A.; Mathewos, A.; Addissie, A.; Bogale, S.; Wondemagegnehu, T.; Aynalem, A.; Ayalnesh, B.; Dagnechew, H.; Bireda, W.; et al. First Data from a Population Based Cancer Registry in Ethiopia. *Cancer Epidemiol.* **2018**, *53*, 93–98. [CrossRef] [PubMed]
4. Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R.L.; Torre, L.A.; 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. [CrossRef]

5. MOH Federal Ministry of Health, Ethiopia. *Health Sector Transformation Plan, 2015/16–2019/20*; Ministry of Health: Addis Ababa, Ethiopia, 2015.
6. Akarolo-Anthony, S.N.; Ogundiran, T.O.; Adebamowo, C.A. Emerging Breast Cancer Epidemic: Evidence from Africa. *Breast Cancer Res.* **2010**, *12*, S8. [CrossRef]
7. Brinton, L.A.; Figueroa, J.D.; Awuah, B.; Yarney, J.; Wiafe, S.; Wood, S.N.; Ansong, D.; Nyarko, K.; Wiafe-Addai, B.; Clegg-Lamptey, J.N. Breast Cancer in Sub-Saharan Africa: Opportunities for Prevention. *Breast Cancer Res. Treat.* **2014**, *144*, 467–478. [CrossRef] [PubMed]
8. Figueroa, J.D.; Davis Lynn, B.C.; Edusei, L.; Titiloye, N.; Adjei, E.; Clegg-Lamptey, J.N.; Yarney, J.; Wiafe-Addai, B.; Awuah, B.; Duggan, M.A.; et al. Reproductive Factors and Risk of Breast Cancer by Tumor Subtypes among Ghanaian Women: A Population-Based Case–Control Study. *Int. J. Cancer* **2020**, *147*, 1535–1547. [CrossRef]
9. Plaza-DÍaz, J.; Álvarez-Mercado, A.I.; Ruiz-Marín, C.M.; Reina-Pérez, I.; Pérez-Alonso, A.J.; Sánchez-Andujar, M.B.; Torné, P.; Gallart-Aragón, T.; Sánchez-Barrón, M.T.; Reyes Lartategui, S.; et al. Association of Breast and Gut Microbiota Dysbiosis and the Risk of Breast Cancer: A Case-Control Clinical Study. *BMC Cancer* **2019**, *19*, 495. [CrossRef]
10. Wang, N.; Sun, T.; Xu, J. Tumor-Related Microbiome in the Breast Microenvironment and Breast Cancer. *J. Cancer* **2021**, *12*, 4841–4848. [CrossRef]
11. Urbaniak, C.; Cummins, J.; Brackstone, M.; Macklaim, J.M.; Gloor, G.B.; Baban, C.K.; Scott, L.; O'Hanlon, D.M.; Burton, J.P.; Francis, K.P.; et al. Microbiota of Human Breast Tissue. *Appl. Environ. Microbiol.* **2014**, *80*, 3007–3014. [CrossRef]
12. Smith, A.; Pierre, J.F.; Makowski, L.; Tolley, E.; Lyn-Cook, B.; Lu, L.; Vidal, G.; Starlard-Davenport, A. Distinct Microbial Communities That Differ by Race, Stage, or Breast-Tumor Subtype in Breast Tissues of Non-Hispanic Black and Non-Hispanic White Women. *Sci. Rep.* **2019**, *9*, 11940. [CrossRef] [PubMed]
13. Xuan, C.; Shamonki, J.M.; Chung, A.; DiNome, M.L.; Chung, M.; Sieling, P.A.; Lee, D.J. Microbial Dysbiosis Is Associated with Human Breast Cancer. *PLoS ONE* **2014**, *9*, e0083744. [CrossRef] [PubMed]
14. Fernández, M.F.; Reina-Pérez, I.; Astorga, J.M.; Rodríguez-Carrillo, A.; Plaza-Díaz, J.; Fontana, L. Breast Cancer and Its Relationship with the Microbiota. *Int. J. Environ. Res. Public Health* **2018**, *15*, 1747. [CrossRef]
15. Newman, T.M.; Vitolins, M.Z.; Cook, K.L. From the Table to the Tumor: The Role of Mediterranean and Western Dietary Patterns in Shifting Microbial-Mediated Signaling to Impact Breast Cancer Risk. *Nutrients* **2019**, *11*, 2565. [CrossRef] [PubMed]
16. Bhatt, A.P.; Redinbo, M.R.; Bultman, S.J. The Role of the Microbiome in Cancer Development and Therapy. *CA. Cancer J. Clin.* **2017**, *67*, 326–344. [CrossRef]
17. De Martel, C.; Ferlay, J.; Franceschi, S.; Vignat, J.; Bray, F.; Forman, D.; Plummer, M. Global Burden of Cancers Attributable to Infections in 2008: A Review and Synthetic Analysis. *Lancet Oncol.* **2012**, *13*, 607–615. [CrossRef]
18. Francescone, R.; Hou, V.; Grivennikov, S.I. Microbiome, Inflammation, and Cancer. *Cancer J.* **2014**, *20*, 181–189. [CrossRef]
19. Hsiao, Y.C.; Liu, C.W.; Chi, L.; Yang, Y.; Lu, K. Effects of Gut Microbiome on Carcinogenic DNA Damage. *Chem. Res. Toxicol.* **2020**, *33*, 2130–2138. [CrossRef]
20. Yang, Q.; Wang, B.; Zheng, Q.; Li, H.; Meng, X.; Zhou, F.; Zhang, L. A Review of Gut Microbiota-Derived Metabolites in Tumor Progression and Cancer Therapy. *Adv. Sci.* **2023**, *10*, 202207366. [CrossRef]
21. Dalal, N.; Jalandra, R.; Bayal, N.; Yadav, A.K.; Harshulika; Sharma, M.; Makharia, G.K.; Kumar, P.; Singh, R.; Solanki, P.R.; et al. Gut Microbiota-Derived Metabolites in CRC Progression and Causation. *J. Cancer Res. Clin. Oncol.* **2021**, *147*, 3141–3155. [CrossRef]
22. Urbaniak, C.; Gloor, G.B.; Brackstone, M.; Scott, L.; Tangney, M.; Reida, G. The Microbiota of Breast Tissue and Its Association with Breast Cancer. *Appl. Environ. Microbiol.* **2016**, *82*, 5039–5048. [CrossRef]
23. Mittal, S.; Brown, N.J.; Holen, I. The Breast Tumor Microenvironment: Role in Cancer Development, Progression and Response to Therapy. *Expert Rev. Mol. Diagn.* **2018**, *18*, 227–243. [CrossRef]
24. Chen, Y.; Wu, F.H.; Wu, P.Q.; Xing, H.Y.; Ma, T. The Role of The Tumor Microbiome in Tumor Development and Its Treatment. *Front. Immunol.* **2022**, *13*, 935846. [CrossRef] [PubMed]
25. Ciernikova, S.; Sevcikova, A.; Stevurkova, V.; Mego, M. Tumor Microbiome—An Integral Part of the Tumor Microenvironment. *Front. Oncol.* **2022**, *12*, 1063100. [CrossRef] [PubMed]
26. Wang, H.; Altemus, J.; Niazi, F.; Green, H.; Calhoun, B.C.; Sturgis, C.; Grobmyer, S.R.; Eng, C. Breast Tissue, Oral and Urinary Microbiomes in Breast Cancer. *Oncotarget* **2017**, *8*, 88122–88138. [CrossRef] [PubMed]
27. Parida, S.; Siddharth, S.; Xia, Y.; Sharma, D. Concomitant Analyses of Intratumoral Microbiota and Genomic Features Reveal Distinct Racial Differences in Breast Cancer. *npj Breast Cancer* **2023**, *9*, 4. [CrossRef]
28. Smith, A.; Cao, X.; Gu, Q.; Kubi Amos-Abanyie, E.; Tolley, E.A.; Vidal, G.; Lyn-Cook, B.; Starlard-Davenport, A. Characterization of the Metabolome of Breast Tissues from Non-Hispanic Black and Non-Hispanic White Women Reveals Correlations between Microbial Dysbiosis and Enhanced Lipid Metabolism Pathways in Triple-Negative Breast Tumors. *Cancers* **2022**, *14*, 4075. [CrossRef]
29. Thyagarajan, S.; Zhang, Y.; Thapa, S.; Allen, M.S.; Phillips, N.; Chaudhary, P.; Kashyap, M.V.; Vishwanatha, J.K. Comparative Analysis of Racial Differences in Breast Tumor Microbiome. *Sci. Rep.* **2020**, *10*, 14116. [CrossRef]
30. Reid, G.; Nduti, N.; Sybesma, W.; Kort, R.; Kollmann, T.R.; Adam, R.; Boga, H.; Brown, E.M.; Einerhand, A.; El-nezami, H. Harnessing Microbiome and Probiotic Research in Sub-Saharan Africa: Recommendations from an African Workshop. *BMC Microbiome* **2014**, *2*, 12. [CrossRef]

31. Gupta, V.K.; Paul, S.; Dutta, C. Geography, Ethnicity or Subsistence-Specific Variations in Human Microbiome Composition and Diversity. *Front. Microbiol.* **2017**, *7–8*, 01162. [CrossRef]
32. De Filippo, C.; Di Paola, M.; Ramazzotti, M.; Albanese, D.; Pieraccini, G.; Banci, E.; Miglietta, F.; Cavalieri, D.; Lionetti, P. Diet, Environments, and Gut Microbiota. A Preliminary Investigation in Children Living in Rural and Urban Burkina Faso and Italy. *Front. Microbiol.* **2017**, *8*, 01979. [CrossRef]
33. Brewster, R.; Tamburini, F.B.; Asiimwe, E.; Oduaran, O.; Hazelhurst, S.; Bhatt, A.S. Surveying Gut Microbiome Research in Africans: Toward Improved Diversity and Representation. *Trends Microbiol.* **2019**, *27*, 824–835. [CrossRef] [PubMed]
34. Choudhury, A.; Aron, S.; Sengupta, D.; Hazelhurst, S. African Genetic Diversity Provides Novel Insights into Evolutionary History and Local Adaptations. *Hum. Mol. Genet.* **2018**, *27*, 209–218. [CrossRef]
35. Allison, K.H.; Hammond, M.E.H.; Dowsett, M.; McKernin, S.E.; Carey, L.A.; Fitzgibbons, P.L.; Hayes, D.F.; Lakhani, S.R.; Chavez-MacGregor, M.; Perlmutter, J.; et al. Estrogen and Progesterone Receptor Testing in Breast Cancer: ASCO/CAP Guideline Update. *J. Clin. Oncol.* **2020**, *38*, 1346–1366. [CrossRef] [PubMed]
36. Wolff, A.C.; Elizabeth Hale Hammond, M.; Allison, K.H.; Harvey, B.E.; Mangu, P.B.; Bartlett, J.M.S.; Bilous, M.; Ellis, I.O.; Fitzgibbons, P.; Hanna, W.; 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. [CrossRef] [PubMed]
37. Cheang, M.C.U.; Chia, S.K.; Voduc, D.; Gao, D.; Leung, S.; Snider, J.; Watson, M.; Davies, S.; Bernard, P.S.; Parker, J.S.; et al. Ki67 Index, HER2 Status, and Prognosis of Patients with Luminal B Breast Cancer. *J. Natl. Cancer Inst.* **2009**, *101*, 736–750. [CrossRef]
38. Dowsett, M.; Nielsen, T.O.; A'Hern, R.; Bartlett, J.; Coombes, R.C.; Cuzick, J.; Ellis, M.; Henry, N.L.; Hugh, J.C.; Lively, T.; et al. Assessment of Ki67 in Breast Cancer: Recommendations from the International Ki67 in Breast Cancer Working Group. *J. Natl. Cancer Inst.* **2011**, *103*, 1656–1664. [CrossRef]
39. Elston, C.W.; Ellis, I.O. Pathological Prognostic Factors in Breast Cancer. I. The Value of Histological Grade in Breast Cancer: Experience from a Large Study with Long-Term. *Histopathology* **1991**, *19*, 9–10. [CrossRef]
40. Bernard, P.S.; Parker, J.S.; Mullins, M.; Cheung, M.C.U.; Leung, S.; Voduc, D.; Vickery, T.; Davies, S.; Fauron, C.; He, X.; et al. Supervised Risk Predictor of Breast Cancer Based on Intrinsic Subtypes. *J. Clin. Oncol.* **2009**, *27*, 1160–1167. [CrossRef]
41. Desalegn, Z.; Yohannes, M.; Porsch, M.; Stückrath, K.; Anberber, E.; Santos, P.; Bauer, M.; Addissie, A.; Bekuretsion, Y.; Assefa, M.; et al. Intrinsic Subtypes in Ethiopian Breast Cancer Patient. *Breast Cancer Res. Treat.* **2022**, *196*, 495–504. [CrossRef]
42. Kozich, J.J.; Westcott, S.L.; Baxter, N.T.; Highlander, S.K.; Schloss, P.D. Development of a Dual-Index Sequencing Strategy and Curation Pipeline for Analyzing Amplicon Sequence Data on the Miseq Illumina Sequencing Platform. *Appl. Environ. Microbiol.* **2013**, *79*, 5112–5120. [CrossRef] [PubMed]
43. Callahan, B.J.; Mcmurdie, P.J.; Rosen, M.J.; Han, A.W. DADA2: High-Resolution Sample Inference from Illumina Amplicon Data. *Nat. Methods* **2016**, *13*, 581–583. [CrossRef] [PubMed]
44. Mcmurdie, P.J.; Holmes, S. Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE* **2013**, *8*, e61217. [CrossRef] [PubMed]
45. Nguyen, N.K.; Deehan, E.C.; Zhang, Z.; Jin, M.; Baskota, N.; Perez-Muñoz, M.E.; Cole, J.; Tuncil, Y.E.; Seethaler, B.; Wang, T.; et al. Gut Microbiota Modulation with Long-Chain Corn Bran Arabinoxylan in Adults with Overweight and Obesity Is Linked to an Individualized Temporal Increase in Fecal Propionate. *Microbiome* **2020**, *8*, 118. [CrossRef] [PubMed]
46. Unger-Saldaña, K. Challenges to the Early Diagnosis and Treatment of Breast Cancer in Developing Countries. *World J. Clin. Oncol.* **2014**, *5*, 465–477. [CrossRef]
47. Kantelhardt, E.J.; Zerche, P.; Mathewos, A.; Trocchi, P.; Addissie, A.; Aynalem, A.; Wondemagegnehu, T.; Ersumo, T.; Reeler, A.; Yonas, B.; et al. Breast Cancer Survival in Ethiopia: A Cohort Study of 1,070 Women. *Int. J. Cancer* **2014**, *135*, 702–709. [CrossRef]
48. Bauer, M.; Vetter, M.; Stückrath, K.; Yohannes, M.; Desalegn, Z.; Yalew, T.; Bekuretsion, Y.; Kenea, T.W.; Joffe, M.; van den Berg, E.J.; et al. Regional Variation in the Tumor Microenvironment, Immune Escape and Prognostic Factors in Breast Cancer in Sub-Saharan Africa. *Cancer Immunol. Res.* **2023**, *11*, 720–731. [CrossRef]
49. Zhang, J.; Xia, Y.; Sun, J. Breast and Gut Microbiome in Health and Cancer. *Genes Dis.* **2021**, *8*, 581–589. [CrossRef]
50. Hieken, T.J.; Chen, J.; Hoskin, T.L.; Walther-Antonio, M.; Johnson, S.; Ramaker, S.; Xiao, J.; Radisky, D.C.; Knutson, K.L.; Kalari, K.R.; et al. The Microbiome of Aseptically Collected Human Breast Tissue in Benign and Malignant Disease. *Sci. Rep.* **2016**, *6*, 30751. [CrossRef]
51. Thompson, K.J.; Ingle, J.N.; Tang, X.; Chia, N.; Jeraldo, P.R.; Walther-Antonio, M.R.; Kandimalla, K.K.; Johnson, S.; Yao, J.Z.; Harrington, S.C.; et al. A Comprehensive Analysis of Breast Cancer Microbiota and Host Gene Expression. *PLoS ONE* **2017**, *12*, e188873. [CrossRef]
52. Murphy, E.A.; Velazquez, K.T.; Herbert, K.M. Influence of High-Fat Diet on Gut Microbiota: A Driving Force for Chronic Disease Risk. *Curr. Opin. Clin. Nutr. Metab. Care* **2015**, *18*, 515–520. [CrossRef] [PubMed]
53. Klann, E.; Williamson, J.M.; Tagliamonte, M.S.; Ukhanova, M. Microbiota Composition in Bilateral Healthy Breast Tissue and Breast Tumors. *Cancer Causes Control* **2020**, *31*, 1027–1038. [CrossRef]
54. Lehouritis, P.; Cummins, J.; Stanton, M.; Murphy, C.T.; McCarthy, F.O.; Reid, G.; Urbaniak, C.; Byrne, W.L.; Tangney, M. Local Bacteria Affect the Efficacy of Chemotherapeutic Drugs. *Sci. Rep.* **2015**, *5*, 14554. [CrossRef]

55. Lee, I.O.; Kim, J.H.; Choi, Y.J.; Pillinger, M.H.; Kim, S.; Blaser, M.J.; Lee, Y.C. Helicobacter Pylori CagA Phosphorylation Status Determines the Gp130-Activated SHP2 / ERK and JAK / STAT Signal Transduction Pathways in Gastric Epithelial Cells. *J. Biol. Chem.* **2010**, *285*, 16042–16050. [CrossRef] [PubMed]
56. German, R.; Marino, N.; Hemmerich, C.; Podicheti, R.; Rusch, D.B.; Stiemsma, L.T.; Gao, H.; Xuei, X.; Rockey, P.; Storniolo, A.M. Exploring Breast Tissue Microbial Composition and the Association with Breast Cancer Risk Factors. *Breast Cancer Res.* **2023**, *25*, 82. [CrossRef]
57. Nejman, D.; Livyatan, I.; Fuks, G.; Gavert, N.; Zwang, Y.; Geller, L.T.; Rotter-Maskowitz, A.; Weiser, R.; Mallel, G.; Gigi, E.; et al. The Human Tumor Microbiome Is Composed of Tumor Type-Specific Intracellular Bacteria. *Science* **2020**, *368*, 973–980. [CrossRef]
58. Banerjee, S.; Tian, T.; Wei, Z.; Shih, N.; Feldman, M.D.; Peck, K.N.; DeMichele, A.M.; Alwine, J.C.; Robertson, E.S. Distinct Microbial Signatures Associated with Different Breast Cancer Types. *Front. Microbiol.* **2018**, *9*, 951. [CrossRef]
59. Hoskinson, C.; Zheng, K.; Gabel, J.; Kump, A.; German, R.; Podicheti, R.; Marino, N.; Stiemsma, L.T. Composition and Functional Potential of the Human Mammary Microbiota Prior to and Following Breast Tumor Diagnosis. *mSystems* **2022**, *7*, 1–18. [CrossRef]
60. Di Modica, M.; Arlotta, V.; Sfondrini, L.; Tagliabue, E.; Triulzi, T. The Link Between the Microbiota and HER2+ Breast Cancer: The New Challenge of Precision Medicine. *Front. Oncol.* **2022**, *12*, 947188. [CrossRef]
61. Tzeng, A.; Sangwan, N.; Jia, M.; Liu, C.C.; Keslar, K.S.; Downs-Kelly, E.; Fairchild, R.L.; Al-Hilli, Z.; Grobmyer, S.R.; Eng, C. Human Breast Microbiome Correlates with Prognostic Features and Immunological Signatures in Breast Cancer. *Genome Med.* **2021**, *13*, 22–23. [CrossRef]
62. Yang, P.; Wang, Z.; Peng, Q.; Lian, W.; Chen, D. Comparison of the Gut Microbiota in Patients with Benign and Malignant Breast Tumors: A Pilot Study. *Evol. Bioinform.* **2021**, *17*, 23–24. [CrossRef] [PubMed]
63. Goedert, J.J.; Jones, G.; Hua, X.; Xu, X.; Yu, G.; Flores, R.; Falk, R.T.; Gail, M.H.; Shi, J.; Ravel, J.; et al. Investigation of the Association Between the Fecal Microbiota and Breast Cancer in Postmenopausal Women: A Population-Based Case-Control Pilot Study. *J. Natl. Cancer Inst.* **2015**, *107*, djv147. [CrossRef] [PubMed]
64. Donnet-Hughes, A.; Perez, P.F.; Doré, J.; Leclerc, M.; Levenez, F.; Benyacoub, J.; Serrant, P.; Segura-Roggero, I.; Schiffrin, E.J. Potential Role of the Intestinal Microbiota of the Mother in Neonatal Immune Education. *Proc. Nutr. Soc.* **2010**, *69*, 407–415. [CrossRef] [PubMed]
65. Parida, S.; Sharma, D. Microbial Alterations and Risk Factors of Breast Cancer: Connections and Mechanistic Insights. *Cells* **2020**, *9*, 1091. [CrossRef] [PubMed]
66. Byrd, D.A.; Vogtmann, E.; Wu, Z.; Han, Y.; Wan, Y.; Clegg-Lamptey, J.N.; Yarney, J.; Wiafe-Addai, B.; Wiafe, S.; Awuah, B.; et al. Associations of Fecal Microbial Profiles with Breast Cancer and Nonmalignant Breast Disease in the Ghana Breast Health Study. *Int. J. Cancer* **2021**, *148*, 2712–2723. [CrossRef]
67. Meng, S.; Chen, B.; Yang, J.; Wang, J.; Zhu, D.; Meng, Q.; Zhang, L. Study of Microbiomes in Aseptically Collected Samples of Human Breast Tissue Using Needle Biopsy and the Potential Role of in Situ Tissue Microbiomes for Promoting Malignancy. *Front. Oncol.* **2018**, *8*, 318. [CrossRef]
68. Allali, I.; Abotsi, R.E.; Tow, L.A.; Thabane, L.; Zar, H.J. Human Microbiota Research in Africa: A Systematic Review Reveals Gaps and Priorities for Future Research. *BMC Microbiome* **2022**, *9*, 241. [CrossRef] [PubMed]