UNIVERSITÀ
DI SIENA
1240

Dipartimento di Biotecnologie Mediche

Dottorato in Genetica Oncologia e Medicina Clinica (GenOMeC)

XXXVIII° Ciclo

Coordinatrice: Prof.ssa Ilaria Meloni

Transcriptomic and in-silico characterization of ethnic differences in breast cancer

Settore scientifico disciplinare: MED/04

Candidata

Franciska Piccolo

Dipartimento di Medicina Sperimentale e Clinica,

Università degli studi di Firenze

Firma digitale del/della candidato/a

Firmato digitalmente da
FRANCISKA PICCOLO
02/12/2025 11:02:44
con la Firma Digitale cheFIRMA!

Supervisore

Elena Lastraioli

Università degli Studi di Firenze

Anno accademico di conseguimento del titolo di Dottore di ricerca

2024/25

Università degli Studi di Siena
Dottorato in Genetica, Oncologia e Medicina Clinica (GenOMeC)
XXXVIII° Ciclo

Data dell'esame finale

_____

Commissione giudicatrice

_____
_____
_____
_____
_____

Supplente

_____

Abstract

Breast cancer is a global healthcare burden, with growing rates all over the world and heterogeneous presentation, which affects prognosis and survival. It can be classified according to oestrogen, progesterone and HER2 receptor status into ER+/PR+ breast cancer, HER2+ breast cancer and triple negative breast cancer, which present the most aggressive phenotype. Ethnic differences are present in breast cancer incidence, mortality and biological signatures such as polymorphisms frequency, gene expression, DNA methylation and immune infiltrate.

The aim of my project was to characterize the genetic and molecular signatures in breast cancer from patients of different ethnicities, particularly regarding gene expression. We performed *in silico* analysis on public datasets and RNAseq was applied in a cohort of 12 TNBC European patients, and 2 HER2 + breast cancer, 7 luminal A and 4 luminal B breast tumours, belonging to different ethnicities.

On the TCGA cohort, we identified transcriptomic and mutational differences. Different distribution of breast cancer subtypes across ethnicities was confirmed, with HER2+ subtype being more common in Asian patients and basal subtype in African patients. Differential expression analysis between ethnicities was then performed, revealing lower levels of cell adhesion molecules in Black patients, while Asian patients presented higher expression of genes of the ErbB pathway in the basal subtype. In the HER2+ subtype and in the luminal A subtype, White patients present an enrichment in several metabolic pathways compared to other ethnicities. Among differentially expressed genes, some of them impacted on overall survival: these include *CXCL6* and *CXCL8*, *CLCA2*, *SCN1A KCNK3*, *KCNIP3* ion channels and *SLC30A10* transporter, and some lncRNAs. Caucasian patients' enrichment of genes involved in metabolic pathways compared to other ethnicities and *KCNIP3* higher expression in African patients were confirmed through RNAseq experiments.

Differential expression analysis between basal and other subtypes revealed enrichment in IL-17, cell cycle, Wnt signalling pathway, and stemness, and downregulation of tight junctions, drug metabolism and chemical carcinogenesis, while comparison with healthy tissue revealed enrichment in cell cycle and neutrophil trap formation, and downregulation of antigen processing and presentation, MHC II

protein complex and cell adhesion molecules compared to healthy mammary tissue.

RNA sequencing on TNBC samples identified several novel lncRNAs, which presented a correlation with genes involved in inflammation and various types of breast cancer and with distant metastasis and progression free survival.

In conclusion, we identified several differences in the transcriptomic profile of breast tumours from patients of different ethnicities, spanning from immune response, chemokine expression to ion channels to non-protein coding genes. Pathways enriched in Black patients, which have the most aggressive breast cancer, were often also enriched in TNBC compared to other subtypes or healthy tissue, highlighting their importance in cancer aggressivity and bad prognosis. Ion channels and transporters and lncRNAs demonstrated to be key genes for the characterization of breast cancer biology.

1. Introduction

1.1. Cancer

Cancer is a global healthcare burden, being responsible for 16.8% deaths worldwide. Almost half of cancer diagnosis (49.2%) and 56.1% of deaths occur in Asia, proportionally to the percentage of the global resident population. Europe, instead, has higher cancer incidence and mortality compared to the percentage of the population (22% vs 9%). Cancer mortality is greater in Asia and Africa, mostly due to late diagnosis. The most common type of cancer diagnosed worldwide is lung cancer, followed by female breast cancer, colorectal cancer, prostate cancer and stomach cancer. The most common cause of cancer death is lung cancer, followed by colorectal, liver, female breast and stomach cancer. In women, the most common type of cancer and first cause of cancer death is breast cancer, while in men it is lung cancer (Figure 1) (Bray et al., 2024).

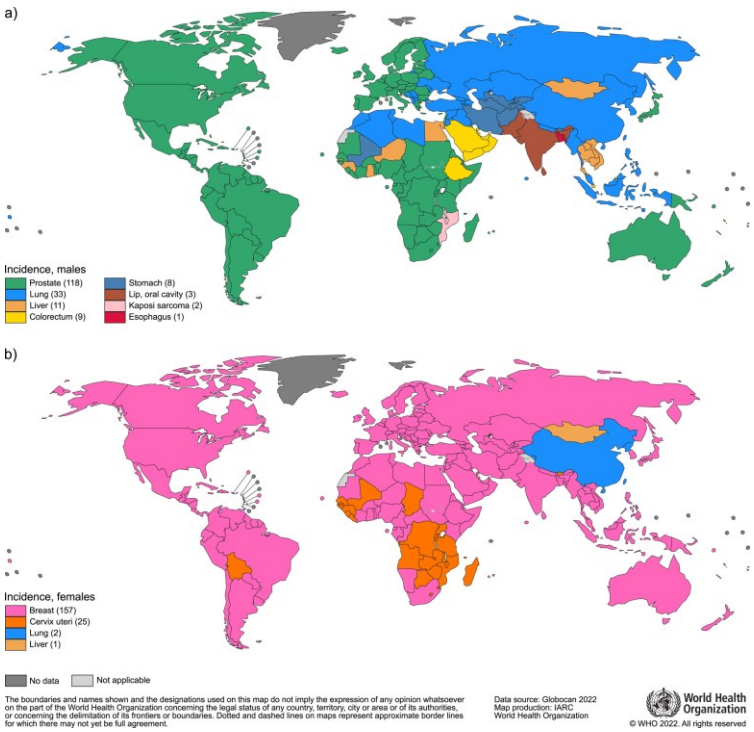

Figure 1. Global maps representing the incidence of the most common types of cancer in 2022 in each country among (A) men and (B) women (Bray et al., 2024)

In men, incidence rates are the highest in Australia and New Zealand and the lowest in western Africa, while mortality is the highest in central American countries and the

lowest in eastern Europe. Women also have the highest incidence in Australia and New Zealand, while the lowest incidence rates are detected in south-central Asia. Their mortality is the highest in central America and south-central Asia and the lowest in Melanesia. It is predicted that cancer incidence will grow in low and medium Human Development Index (HDI) countries (Bray et al., 2024). Cancer risk increases with HDI level, especially in males. Prostate cancer is the most frequent cancer in males in all countries, except in medium HDI countries, where the most common is oral cavity cancer. In high HDI countries the highest mortality is caused by lung and colorectal cancer, while in low HDI countries prostate and liver cancer are responsible for the highest mortality rates. Regarding females, low-to-medium HDI countries have higher breast and cervical cancer incidence and death, while in high HDI countries lung, breast and colorectal cancer incidence is higher, as well as death rates associated with lung and colorectal cancer (Cao et al., 2024). Arabian countries have higher incidence and mortality regarding non-Hodgkin and Hodgkin lymphomas, bladder, breast and liver cancer (Al-Muftah & Al-Ejeh, 2023). Singapore has the highest incidence among South-east Asian countries, while mortality is the highest in Laos (men) and Brunei (women). The most common type of cancer in women in this area is breast cancer as well; it is also the leading cause of cancer death in several of them, while for men the most common and responsible for the highest mortality is lung cancer (Dee et al., 2025).

It must be pointed out that there are inequalities in the way cancer data is estimated in each country. Only 30% of countries whose data are collected in GLOBOCAN have a cancer registry, while for 70% of them best estimates are used (Hemminki & Kaaks, 2024). GLOBOCAN estimation of cancer incidence and mortality for sub–Saharan Africa have poor quality for most criteria and are based on incomplete population, as the main sources of data are hospitals, which are mostly located in major cities, and there is a disproportion between urban and rural areas of registration of cancers (Ayubi et al., 2024).

The incidence of oncological diseases in the young population is on the rise not only in high HDI countries, but also in lower HDI ones, where the mortality is also higher (Hughes et al., 2024).

1.2. Breast cancer

Breast cancer generally develops from the ductal epithelium (ductal carcinoma) or lobule (lobular carcinoma). The breast is made of 15-20 lobes radiating from the nipple, which is reached by a variable number of ducts (from 5 to 20) (Ramsay et al., 2005). It is composed by glandular and adipose tissue, held together by Cooper's ligaments, which anchor the breast to the underlying muscle tissue. Each lobe is composed of lobules, clusters of alveoli containing mammary epithelial cells with secretive activity. Small ducts from the alveoli merge into a larger duct draining the lobules, which also merge in a unique duct for each lobe. Adipose tissue is present between lobes (Zucca-Matthes et al., 2016).

1.3. Epidemiology

Female breast cancer is the second cause of global cancer incidence, representing 11.6 % of all cancer cases and 6.9% of cancer deaths. The highest frequency has been recorded in France, Australia, New Zealand, north America and northern Europe. While in high HDI countries the incidence in higher, the mortality is lower compared to transitioning countries. The highest mortality is found in Melanesia, western Africa and Micronesia/Polynesia (Bray et al., 2024; Liao, 2024). Furthermore, breast cancer incidence is growing in developing nations, tending to reach a pattern similar to that of developed countries (N. Li et al., 2019).

Breast cancer is the most common type of cancer in females in several countries, such as China, India, countries in the eastern Mediterranean region and African countries such as Ethiopia (Dandena et al., 2024; Malvia et al., 2017; Qiu et al., 2021; Zahedi et al., 2020).

In 2022, 2.3 million new cases and 670000 deaths from breast cancer occurred, with annual rates increasing by 1-5%. In countries with very high HDI, mortality decreased but by 2050, it is predicted that new cases and deaths will increase by 38% and 68%, mostly impacting lower HDI countries (Kim et al., 2025). According to the American cancer society, breast cancer rates vary among different ethnicities, with White women having the highest rate, followed by African American women, Hispanic women, native American women and, lastly Asian Americans and Pacific Islanders, who have the lowest incidence (Menon et al., 2025).

1.4. Breast cancer classification

1.4.1. Histopathological Classification

Breast cancer is a heterogeneous disease, which has been classified in different subtypes according to receptor expression, gene expression, and histology. Histological type and grade are fundamental characteristics of breast cancer and divide tumours according to their growth pattern and differentiation. Histological grade is analysed through morphological analysis through haematoxylin-eosin staining, by a pathologist.

Histological tumour grade is based on the differentiation of the tumour tissue and influences survival (Rakha, Reis-Filho, Baehner, et al., 2010). Histopathological classification is based on tumour morphology and divides breast tumours into invasive ductal carcinoma (IDC) (50-75% of diagnosed breast cancers), lobular adenocarcinoma (10-15% of breast cancer), and other rare breast cancer categories. Ductal carcinoma in situ (DCIS) is the precursor lesion of invasive ductal carcinoma (Makki, 2015), while invasive lobular carcinomas (ILC) arise from a precursor lesion called atypical lobular hyperplasia (ALH) or from lobular carcinoma *in situ*. They are typically harder to detect, exhibit a luminal subtype and are oestrogen and progesterone positive (McCart Reed et al., 2021).

Mucinous carcinomas make up 2-5% of breast cancer are characterized by mucin production. They can be further subdivided into mixed mucinous and pure mucinous subtypes, with the former being associated to a better prognosis (Marrazzo et al., 2020). Tubular carcinomas arise from cells that form glands and tubules and constitute 1-2% of breast cancers. Both mucinous and tubular carcinomas are more common in postmenopausal women and have a better prognosis compared to other breast cancer histotypes (Roux et al., 2019). These tumours are for Estrogen (ER) and Progesterone (PR) receptor positives and negative for the type 2 receptor of the Epidermal Growth Factor (HER2) (Masood, 2016). Moreover, these two subtypes are influenced by strong genetic predispositions and familial clustering (Henry & Cannon-Albright, 2019).

Medullary carcinoma constitutes about 5% of all cases, is a highly aggressive poorly differentiated tumour with high numbers of tumour infiltrating lymphocytes (TILs).

These tumours are often present younger patients with BRCA mutations (Cserni, 2020). Apocrine breast cancer, which constitutes 1-4% of all cases, has apocrine differentiation, lacks a peripheral myoepithelial layer and shows high histological grade (Nascimento & Otoni, 2020; Seal et al., 2009). Apocrine carcinomas are often triple negative, expressing basal markers (Masood, 2016), AR overexpression and deregulation of PIK3CA and PTEN pathways (Vranic et al., 2017).

Metaplastic carcinoma represents 1% of breast cancer cases and is more common in post-menopausal women, has a poor differentiation, larger size at presentations and higher rates of metastasis (Nascimento & Otoni, 2020). These tumours are negative for oestrogen, progesterone and HER2 receptors (Schwartz et al., 2013). Cribriform carcinomas are associated with favourable prognosis, constitute 1-3% of cases and affect older patients. They have a survival rate of 90-100% at 10 years from diagnosis (Cong et al., 2015; Nascimento & Otoni, 2020). Neuroendocrine carcinomas constitute 0.5-5% of all cases and are more common in older patients, with attributes similar to neuroendocrine tumours of the gastrointestinal system and lung, chromogranin A , synaptophysin and neuron specific enolase expression (Y. Li et al., 2017; Nascimento & Otoni, 2020).

1.4.2. TNM staging

The staging of malignant tumours is carried out using TNM classification, which evaluates the size of the tumour, and the presence of lymph node metastases and distant metastases.

T (tumour): describes the size of the primary tumour and its invasion of adjacent tissue, its values go from Tis (carcinoma in situ), T0 (no presence of tumour) to T4. T1 indicates tumours with size up to 0.5-1.0 cm (T1b) or 1.0-2.0cm (T1c). T2 tumours are between 2-5 cm. T3 indicates large tumours, while T4 locally advanced ones.

N (nodes): describes regional lymph nodes involvement, varies from N1 to N3 in an increasing order. NX is used when lymph node invasion cannot be assessed. N0 indicates the absence of lymph node invasion. N1a indicates from 1 to 3 positive lymph nodes, N2a from 4 to 9 while N3a indicates more than 10 positive lymph nodes. It is a fundamental prognostic factor.

M (metastasis): identifies the presence of distant metastasis, categorized in M0 (distant metastasis not present) and M1 (a metastasis is detected in an area distant from the primary tumour). Metastasis is the primary cause of cancer-related mortality. The most common sites of metastasis for breast cancer are lungs, brain, bone and liver (Firatligil-Yildirir et al., 2023).

This system allows for the classification of tumours in stages, going from stage 0, indicating a carcinoma in situ, to stage I (localized cancer, either T1 or T2, with no lymph node involvement), stage II (localized cancer, presenting lymph node invasion, T1 or T2 and N1), stage III (locally advanced cancer of bigger dimensions, from T1 to T4, and N2-N3 node invasion) and stage IV (metastatic cancer) (Benson, 2003; Rosen & Sapra, 2025).

1.4.3. Tumour grade

Tumour grade evaluates the tumour cells differentiation. G1 tumours are well differentiated, with cells resembling healthy mammary cells, while high grade tumours (G3-G4) are poorly differentiated and are more aggressive. G2 tumours are moderately differentiated, with an intermediate behaviour (Telloni, 2017).

1.4.4. The Nottingham prognostic index

Nottingham prognostic Index (NPI) is a prognostic model based on lymph node stage, tumour size and pathological grade, dividing patients in three prognostic groups (Haybittle et al., 1982), which were then amplified to five (Todd et al., 1987) and later six: excellent prognostic group (EPF), good (GPG), moderate I (MPG), moderate II (MPG), poor (PPG) and very poor (VPG) (Blamey et al., 2007). This model retains its predicting ability in most independent populations, and it's useful for prognosis prediction (Phung et al., 2019). Nottingham grade I tumours form a homogeneous group, with survival being influenced by lymph node status. Grade 2 tumours, other than lymph node status, are also influenced by PR and HER2 expression. Grade 3 tumours show overexpression of genes participating in cell cycle regulatory pathways (*BIRC5, CDC6, FOXM1, TOP2A, MYBL2, UBE2C*), higher rates of disease recurrence and poor overall survival (OS) and disease free survival (DFS); they have a significant recurrence free survival (RFS) difference with grade 1 tumours (Freitas et al., 2025; Peiris et al., 2017). NPI is used for both prognosis and survival at

15 years from the diagnosis and suitability of treatments in primary breast cancer (Galea et al., 1992).

1.4.5. Breast cancer molecular classification

Breast cancer can be classified according to expression of hormonal receptors: oestrogen receptor (ER), progesterone receptor (PR/PgR), and human epidermal growth factor receptor 2 (HER2/neu). In the latest years, hormonal receptor expression, histological grade and multigene prognostic markers have been added in the prognostic staging system (Sawaki et al., 2019).

There are different oestrogen receptors: Erα, Erβ, and G-protein coupled ERs (GPR30 and GPER). ERα/ESR1 is the oestrogen receptor most involved in breast cancer. It contains six domains: the N-terminal domain which can activate gene transcription, the C-terminal domain, containing the DNA binding domain (DBD) which also enables homodimerization, the D domain, enabling nuclear translocation, the ligand binding domain and the F domain (Habara & Shimada, 2022). The oestrogen receptor's physiological signalling pathway provides that, after binding of oestrogen to ERα, the receptor translocates into the nucleus, activating the transcription of target genes. In breast cancer, it also has extranuclear functions, promoting the stimulation of Src, MAPK, PI3K and protein kinase C in the cytosol, stimulating cell migration. ERα signalling has a role in metastasis, regulation of cell migration and EMT (Y. Li et al., 2010; Saha Roy & Vadlamudi, 2012).

Progesterone receptor is a nuclear receptor composed of a DNA binding domain (DBD), a C-terminal ligand binding domain (LBD) and an N-terminal domain. It is one of the oestrogen receptors target genes, being upregulated by it (Z. Li et al., 2022). Following progesterone binding, it upregulates the expression of RANKL and NFkB pathway (Pecci et al., 2022)

HER2/neu, whose coding gene is *ERBB2* localized in the chromosome 17q21, is a tyrosine kinase receptor, able to heterodimerize with EGFR. In the nucleus, it stimulates breast cancer cell proliferation by acting as a coactivator of STAT3 (Teklemariam et al., 2024). Its structure includes an extracellular, ligand binding domain, a transmembrane domain and an intracellular domain with tyrosine kinase activity. It does not have a natural ligand and is constitutively active. In breast cancer,

it promotes proliferation due to its overexpression (Moasser, 2007).

Oestrogen receptor positive tumours comprise the majority of breast cancers (75%), while progesterone receptor is present in 55-65% of them. ER and PR are often co-expressed; the lack of PR in ER+ tumours correlates with resistance to endocrine therapy. Oestrogen receptor positive breast cancers generally have a better prognosis. Quantification of ki67, a proliferation index, is performed on tumour samples through immunohistochemistry, in order to stratify tumours, and as a prognostic tool (Rakha & Green, 2017). HER2 protein expression and/or gene amplification is present in 13-20% of breast tumours, with its expression being inversely correlated with the expression of ER and PR. Both the oestrogen receptor and HER2 can represent therapeutic targets (Rakha, Reis-Filho, & Ellis, 2010). Tumours lacking the expression of these biomarkers are called triple negative (TNBC). ER+/PR+ breast tumours have a better prognosis compared to both TNBC and HER2+ (Rakha, Reis-Filho, & Ellis, 2010).

Triple negative breast cancer constitutes around 15-20% of breast cancers and is the most aggressive form of this disease. TNBC patients are significantly younger than breast cancer patients diagnosed with other breast cancer subtypes, present larger tumours, a higher frequency of lymph node positivity, and higher probability of recurrence and distant metastasis occurrence, also due to the lack of therapeutic targets (Dent et al., 2007). TNBC is more frequent in women of African descent compared to Caucasian women in every age range (22% vs 11%), and in young and premenopausal women (Amirikia et al., 2011; Bauer et al., 2007; Morris et al., 2007; O'Brien et al., 2010). TNBC shows significant expression of vimentin, whose expression is inversely correlated with the expression of ER and NF-kB (Elzamly et al., 2018). In TNBC alterations in pathway of cell cycle, DNA repair, nucleotide synthesis, metabolic pathways, NF-kB signalling, inflammatory response and angiogenesis have been described (Ossovskaya et al., 2011). The genomic landscape of TNBC is heterogeneous, with *TP53* and *PIK3CA* mutations and *MYC* amplifications being the most common, present respectively in the 75-80%, 10% and 26% of cases (Weisman et al., 2016).

Breast cancer subtypes represent distinct biological entities, both clinically and biologically, with different gene expression patterns, genomic alterations and clinical

outcomes (Sørlie et al., 2001, 2003).

Employing DNA microarrays and gene expression patterns of 50 genes (PAM50), Perou and colleagues classified breast cancer into five intrinsic subtypes: luminal A, luminal B, HER2-enriched, basal and normal-like (Perou et al., 2000) (Figure 2).

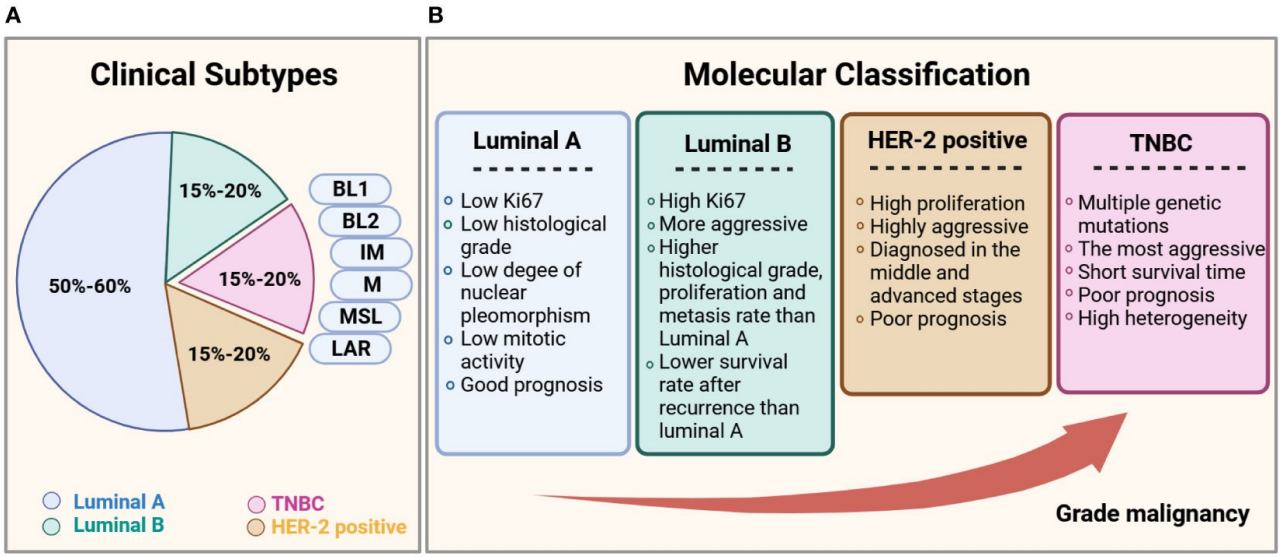

Figure 2. Molecular classification of breast cancer (Pan, Zhao, et al., 2023).

Luminal A: it is characterized by ER and PR expression, lack of HER2 expression, and low Ki67. It compromises half of new diagnosis and includes invasive ductal carcinoma, invasive lobular carcinoma and other less common histological subtypes like tubular, cribriform and mucinous. It has a better prognosis compared to other subtypes and it is not highly invasive. It shows high expression of genes involved in the luminal signature (*ESR1, GATA3, FOXA1, XBP1, MYB*). These genes are also mostly mutated in this subtype. *GATA3* and *FOXA1* are mutated in a mutually exclusive fashion. Luminal A also show a high mutation frequency of *PIK3CA* (Koboldt et al., 2012).

Luminal B: it is ER and PR positive, HER2 negative, with high Ki67 expression. It comprises 20-30% of new breast cancer diagnosis. Compared to luminal A subtype, it is more aggressive, more common in younger patients and has a higher frequency of *TP53* mutations (29% vs 12%). Compared to luminal A, luminal B breast cancers are associated with higher grade and higher frequency of nodal metastasis, as well as a higher expression of genes with a role in mitosis and cell proliferation (Hashmi et al.,

2018). Both luminal A and B subtypes have a pattern resembling the luminal epithelial component of the breast (Koboldt et al., 2012), and include tubular, cribriform, lobular and mucinous carcinomas (Provenzano et al., 2018).

HER2-enriched subtype: it constitutes 15-20% of new diagnosis, is ER and PR negative and defined by an overexpression of HER2/neu, often caused by *ERBB2* gene amplification, which can be assessed with Fluorescence *In Situ* Hybridization (FISH). Sometimes, it can show amplifications of genes in the same genomic region as *ERBB2*; it has low expression of genes involved in the luminal clusters (Makki, 2015).

Basal subtypes: overlaps with the triple negative subtype in around 75% of cases and has a worse prognosis compared to other subtypes (Kreike et al., 2007). It shows low expression of genes involved in luminal and HER2 gene clusters and high expression of basal cluster and myoepithelial cell phenotype, including epithelial cytokeratins 5/6, EGFR, c-KIT and vimentin. It has a high proliferation rate and a high frequency of mutations in *TP53* (Livasy et al., 2006; Nielsen et al., 2004). It is the most common subtype in carriers of the *BRCA1* germline mutation, that predisposes to this subtype (Koboldt et al., 2012; Sørlie et al., 2003). The BRCA1 and BRCA2 pathways are altered in sporadic basal cancers as well (Perou & Børresen-Dale, 2011). Basal tumours have higher mitotic index and higher grade than other subtypes (L. A. Carey et al., 2006). TNBC and tumours belonging to the basal subtype share many features, such as high grade, elevated mitotic count, necrosis, *TP53* mutations, increased CAV1, CAV2 expression and decreased AR expression (L. Carey et al., 2010).

Normal-like: this cancer's transcriptomic characteristics resemble the normal breast tissue. Up to 10% of breast cancer can be classified as normal-like, presenting as either ER positive or E negative, with an intermediate prognosis. However, some studies showed that this results could be caused by normal cell populations contamination (Parker et al., 2009).

Due to the diverse expression patterns, it has been proposed that basal and luminal tumours arise from different cells of origin.

Using PAM50 assay to classify breast cancer patients into intrinsic subtypes is prognostic for DFS and OS, differently from classification through

immunohistochemistry of six markers (ER, PR, HER2, ki67, EGFR and CK5/6): the classification in luminal subtypes, for example, predicts tamoxifen benefit (Chia et al., 2012). Furthermore, different subtypes tend to metastasise to different organ and tissues, with luminal A and B subtypes prioritizing bone metastasis, HER2 enriched more frequently gives rise to liver and lung metastasis and TNBC metastasizing mostly to the lung (Firatligil-Yildirir et al., 2023).

The different subtypes have also distinct mutation rates, with Luminal A subtype having *PIK3CA* as its most mutated gene, while in luminal B subtype, *PIK3CA* and *TP53* are similarly mutated. In the HER2 subtype there is high rate of *TP53* and *PIK3CA* mutations, and *ERBB2* gene amplification. The basal subtypes presents TP53 mutations in 80% of the cases, with truncating mutations being more common than in other subtypes (Vuong et al., 2014). The prognosis and overall survival of different subtypes varies as well, with luminal subtypes having the best prognosis and basal-like and HER2-enriched subtypes the worst one (Parker et al., 2009).

Another subtype was later identified: the claudin-low subtype, which presents low expression of genes involved in tight junctions and cell-to-cell adhesion. These tumours have reduced expression of luminal genes and are mostly ER negative, often belonging to the triple negative subtype (Herschkowitz et al., 2007). Conversely, they have high expression of EMT markers and genes involved in immune response and cancer stem cells, and are associated with poor prognosis (Prat et al., 2010).

TNBC can be further classified into additional molecular subtypes according to their gene expression profile (Kreike et al., 2007): initially the classification was composed of six subtypes, which were then refined to four: basal-like 1 (BL1), basal-like 2 (BL2), mesenchymal-like (M), mesenchymal stem-like (MSL), immunomodulatory (IM), luminal androgen receptor (LAR) (Figure 3). BL1 and BL2 express high expression of genes involved in cell proliferation and DNA damage response and often correlate with BRCA1 and BRCA2 mutations; the mesenchymal-like subtype is enriched in TFG-B, mTOR, Wnt-B-catenin and VEGF signalling pathways; the MSL subtype has similar enrichment to the mesenchymal one, but is also enriched in various growth factor signalling pathways, calcium and G-protein signalling and ABC transporters; the IM subtype is enriched in pathways playing roles in immunity processes, such as Th1/Th2 pathways, cytokine signalling and antigen processing and

presentation; lastly, the LAR subtype is characterized by an androgen receptor (AR) gene signature and expression of luminal cytokeratin. The MSL and IM subtypes were later excluded from the classification, as their transcriptomic characteristics were identified to stem from infiltrating lymphocytes and tumour-associated stromal cells (Lehmann et al., 2011, 2016).

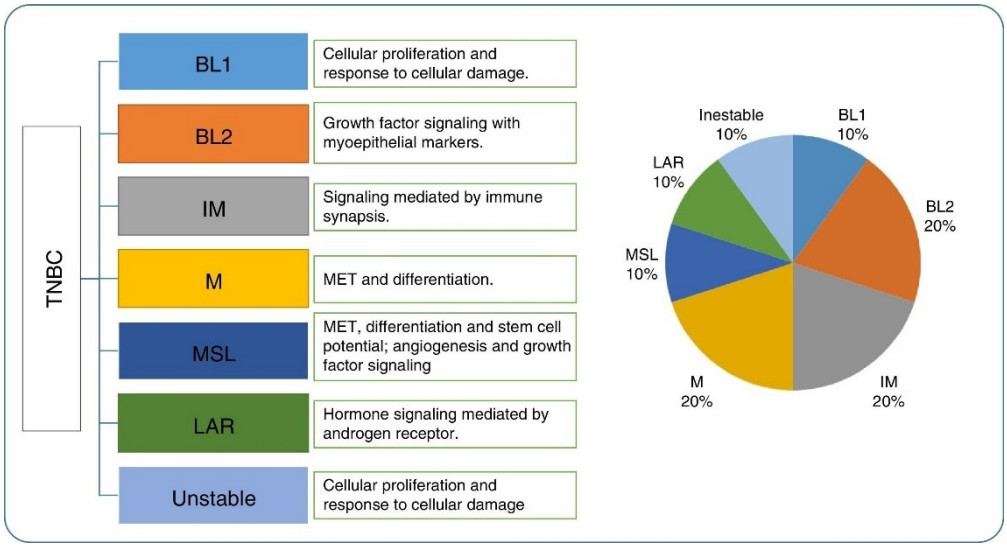

Figure 3. Molecular classification of triple negative breast cancer, (Uscanga-Perales et al., 2016)

Another classification, based on profiling of DNA and RNA, identified four TNBC subtypes, which partly overlapped Lehmann's classification: luminal androgen receptor (LAR), mesenchymal (MES), basal-like immunosuppressed (BLIS) and basal-like immune-activated (BLIA). BLIS tumours have the lowest prognosis, while BLIA are associated with the best one. Gene amplification can drive gene overexpression, with different amplified genes being specific to different subtypes. *CCND1* and *FGFR2* are amplified in LAR tumours, while *MAGOHB* amplification is present in MES, BLIS and BLIA tumours. In LAR and MES tumours, regulators of cell cycle and DNA repair genes are downregulated; in MES and BLIA tumours an upregulation of immune signalling and immune-related death pathways is present instead (Burstein et al., 2015). Mutations in *TP53* gene are present in all TNBC subtypes, while *PIK3CA* mutations are more frequent in the LAR subtype. *BRCA1* is

more commonly mutated in the mesenchymal subtype. *VHL* mutations are common in LAR. The BLIS subtype has the highest number of amplifications and deletions, suggesting higher levels of genomic instability. The mesenchymal subtype is enriched in the RTK-RAS pathways, which suggest a possible sensitivity to RTK inhibitors (R. Q. Li et al., 2024).

Subtyping of TNBC can also be assessed according to metabolomics, dividing it into a lipogenic, a glycolytic and a mixed subtype. In the second subtype, inhibiting the lactate dehydrogenases has been proven to enhance tumour response to anti-PD1 immunotherapy (Gong et al., 2021).

1.4.6. Prognostic assays

Several biomarker assays based on gene expression, with prognostic validity are commercially available and used to guide adjuvant systemic therapy decisions in women with early-stage breast cancer: these comprehend MammaPrint, Oncotype DX, EndoPredict and CanAssistBreast (CAB).

MAMMAPRINT is a 70-gene signature test for predicting clinical outcome and development of metastasis in women with breast cancer in the initial stage. The 70 genes in the assay are involved in regulating cell cycle, invasion, metastasis and angiogenesis, and include cyclin E2, MCM6, and MMP9, which are upregulated in tumours with poor prognosis (van ’t Veer et al., 2002). MAMMAPRINT was the first fully commercialized multivariate predictive test employing a microarray-based multigene assay, FDA approved, and is offered to women under 61 years of age, with lymph node negative breast cancer, either ER positive or negative (Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Working Group, 2016; Slodkowska & Ross, 2009).

Oncotype DX is a 21 gene recurrence score (RS) used as a diagnostic tool and to verify the benefits of chemotherapy in early-stage breast cancer. Higher oncotype score (>=25) is associated with grade III tumours, higher Ki67 and poor NPI (Durrani et al., 2021). Oncotype DX testing increases confidence in treatment decision for patients with ER+, node negative, early-stage breast tumours. Low RS predicts little positive effect of chemotherapy (Albain et al., 2010; Holt et al., 2013). Higher Oncotype DX risk score is associated with ER+, PR- breast tumours (Chaudhary et

al., 2016).

EndoPredict is a molecular predictor of recurrence in ER+ breast cancer, based on a multigene assay which can be employed even on formalin fixed paraffine embedded (FFPE) tumour tissues. It is a prognostic risk score independent of clinicopathological risk factors, but can be combined with data such as nodal status and tumour size (Filipits et al., 2011).

CanAssist Breast (CAB) is a proteomic test analysing the expression of five biomarkers (CD44, ABCC4, ABCC11, N-cadherin and panCadherin) plus clinical pathological parameters (tumour size, tumour grade and node status). It provides risk stratification in early stage, receptor positive breast cancer, which is useful for taking accurate therapeutic decisions, segregating patients into high risk and low risk groups for cancer recurrences and distant metastasis occurrence in 5 years, and it is particularly accurate in finding low-risk patients, which can be spared chemotherapy. Its validity was proved in different cohorts from various geographic regions, but is at the moment only approved in India (Gunda et al., 2022; Ramkumar et al., 2018). CanAssist Breast and Oncotype DX show high concordance between each other (Sengupta et al., 2020).

PREDICT is a tool based on clinicopathological parameters (size of the tumour, tumour grade, lymph node status and ER status) and details of adjuvant chemotherapy. It is able to predict the overall survival and disease outcome following surgery, with or without adjuvant chemotherapy, in all groups of patients except in patients older than 75 years (Wishart et al., 2010). It has been developed on UK breast cancer patients, and validated on cohorts from various different countries, like Canada, Malaysia, and the Netherlands (Wong et al., 2015). In later years, the model has been updated to include the benefits of radiotherapy and harms of both chemo and radiotherapy (Grootes et al., 2024). In a new version (PREDICT Plus) prognostic validity has been confirmed also in patients with HER2+ tumours, regarding both OS and breast cancer specific survival (BCSS) (Wishart et al., 2012).

Many studies have shown the significance of gene expression profiles in predicting prognosis, clinical outcomes and response to therapies.

1.5. Diagnosis, monitoring and screening of breast cancer

The diagnosis of breast cancer can happen following a symptom like pain, discharge, visible changes in the breast, self-physical examination performed by the patients themselves which results in a medical assessment or through screening. Most diagnoses are confirmed through imaging techniques, which include ultrasounds (US), mammography and magnetic resonance imaging (MRI). The most commonly used tools for the diagnosis of breast cancer are ultrasound and mammography.

Ultrasounds are a safe technique, as they do not employ ionizing radiations, and are mostly used in younger women presenting little risk breast cancer, pregnant and breastfeeding women (Jafari et al., 2018). Although safe and effective, ultrasound's addition to mammography for women at elevated risk of breast cancer increases the rate of false positives (Berg et al., 2008).

Mammography, instead, is an X-ray technique, capable to detect breast cancer lesions with a high resolution before they become palpable (Nicosia et al., 2023). However, mammography cannot differentiate between liquid lesions like cysts and solid lesions, at difference from the use of ultrasounds (A. Evans et al., 2018). To standardise mammography reporting, the breast imaging-reporting and data system (BI-RADS) scale was created, consisting of seven categories, from 0 to 6. Category 0 indicates cases which require additional evaluations and/or comparisons with other mammographic scans; category 1 is a negative mammography with no presence of tumours, while categories 2 and 3 show the presence of a benign or likely benign lesion. Category 4 is divided in 4a, 4b and 4c, and indicates a suspicious abnormality, which requires a biopsy, in a growing percentage of probability of being malignant. Category 5 shows a highly suspicious malignant lesion, while category 6 shows a biopsy-proven malignancy (Magny et al., 2025). In order to use a lower level of radiation and have an easier post processing of images, digital mammography has been developed (Alcantara et al., 2014).

MRI is used in patients after surgery and/or therapy, to identify recurrences, residual tumour, metastasis, occult, multicentric or contralateral disease, and to evaluate response to therapy (Weinstock et al., 2015). Contrast agents based on gadolinium or iodine are able to enhance the quality and sensitivity of images (Jafari et al., 2018). US and MRI imaging are both more sensitive to detect invasive cancer than mammography, however they can overestimate tumour size, so different imaging

techniques, combined to clinical examination are usually used to obtain higher sensitivity and specificity (Berg et al., 2004; Dorn et al., 2013).

Tomosynthesis allows three-dimensional breast studies, with superior accuracy compared to normal mammography, and its use combined with mammography increases cancer detection power (Friedewald et al., 2014). Imaging techniques employing radioactive isotopes which emit gamma photons or positrons include single photon emission computed tomography (SPECT) and positron emission tomography (PET), and can identify both the tumour and axillary metastasis, their response to chemotherapy and physiological information such as glucose metabolism, blood flow and oxygenation. The most used radiotracer is 18F-fluorodeoxyglucose. Other radiopharmaceuticals have been developed able to asses ER expression, HER2 status and other cell surface receptors (Bénard & Turcotte, 2005; Jafari et al., 2018).

The latest improvements in artificial intelligence are influencing diagnostic imaging, as deep learning-based AI systems can improve breast cancer detection through digital mammography or digital breast tomosynthesis (DBT). These systems can reduce false negatives and positives, enhancing screening outcomes and increasing accuracy. However, at the moment, no standardized systems or guidelines are available (Díaz et al., 2024)

Among novel biomarkers for the diagnosis and monitoring of breast cancer is free circulating tumour DNA (ctDNA), whose levels in the bloodstream can help to discriminate women with breast cancer from healthy individuals and monitor disease recurrence in patients after breast surgery and therapies (Catarino et al., 2008).

Effective cancer screening should be able to detect a cancerous mass during the preclinical phase, before the development of symptoms, which often leads to a more favourable prognosis. As probability of developing breast cancer increases with age, guidelines recommend mammographic screening every two years starting from 45 years. In women with family history of breast cancer or known BRCA1/2 or other high risk genes mutations, screening is performed every year, to detect aggressive cancers which could arise during intervals between screenings. Breast self-examination is encouraged by doctors so that women can identify changes in their breast tissue and prompt a clinical examination (Loibl et al., 2024; McDonald et al.,

2016; Overmoyer, 1999). Screening ultrasounds in women with negative mammography has been shown to be able to identify additional cancers which were not notable in mammography's, with a recall rate of 5.8% (Hwang et al., 2015).

1.6. Risk factors

Both genetic and environmental factors play a great role in risk of developing breast cancer.

Among environmental factors influencing breast cancer risk are diet, physical activity, tobacco smoking, alcohol intake, body mass index (BMI), breastfeeding and hormonal contraceptives. The 4th edition of the European code against cancer recommends healthy diets, to avoid smoking and avoid or limit alcohol intake and sun exposure and be physically active (Norat et al., 2015).

Healthy diets comprise high intakes of fruit, vegetables, whole grain foods and low intakes of red and processed meats and salt. It has been shown that dietary fibre intake lowers breast cancer risk (Aune et al., 2012). Similarly, it is important to limit foods with high levels of sugar and fat and high caloric content, and sugary drinks, both to prevent the insurgence of cancer and to improve overall survival after diagnosis. Vegetable consumption is associated with lower breast cancer recurrence in patients undergoing endocrine therapy (Thomson et al., 2011; Vrieling et al., 2013).

The intake of fatty acids in the diets is reflected by serum levels of trans-monounsaturated fatty acid. High levels of them, in particular palmitoleic and oleic acids, are connected with increased risk of breast cancer. Cis-monounsaturated fatty acid, instead, do not influence breast cancer risk (Chajès et al., 2008). Cooking with hydrogenated fats increases breast cancer risk compared to use of olive oil, as the former contains linoleic acid and the latter oleic acid (a monounsaturated fat). Interestingly, fatty acid intake was found to increase risk of breast cancer in White and Latin women, but not in African American women (J. Wang et al., 2008). High total and saturated fat intake correlate with the development of receptor positive breast cancers (Sieri et al., 2014). High levels of fat in the diet increase breast cancer risk also due to its capability of increasing oestrogen serum levels in post-menopausal women (Turner, 2011). Furthermore, some dietary compounds have been demonstrated to be associated with anti-cancer effects. These include polyunsaturated

fatty acids, carotenoids, curcumin and green tea (Banikazemi et al., 2018).

A healthy body weight reduces the risk of breast cancer. Body mass index (BMI) at diagnosis shows a correlation with prognosis in breast cancer: obesity correlates with poor prognosis in ER+ PR+ breast cancer, whereas being underweight is associated with poor prognosis in the HER2+ subtype (Jeon et al., 2015). Obesity and type 2 diabetes in postmenopausal women increase risk of developing oestrogen-dependent breast cancer due to adipose tissue inflammation and proinflammatory cytokines and prostaglandins, which are characteristic of obesity, while in premenopausal women they are associated with invasive, metastatic, triple negative breast cancer (Rose et al., 2015). Obesity was associated with short overall survival in breast cancer in both pre- and post-menopausal women (Chan et al., 2014). There are correlations between the consumption of food and beverages which promote weight gain and triple negative breast cancer in premenopausal women. Studies suggest that over 40% of post-menopausal breast cancer diagnoses could be prevented by reducing alcohol, excess body weight and inactivity (Castelló et al., 2015). Physical activity, low consumption of fat, processed foods, avoidance of tobacco smoking and alcohol are associated with lower breast cancer risk, in both menopausal and premenopausal women. This leads to adiposity reduction and reduction of psychological stress (Kruk, 2015). Physical activity has been demonstrated not only to reduce risk of breast cancer, but also to be able to reduce breast cancer related death (Lahart et al., 2015).

Alcohol consumption even in moderate quantities is associated with increased breast cancer probability, with risk being from 15% to 51% higher than in women that do not drink alcohol. It especially influences the development of hormone positive breast cancer, due to its susceptibility between menarche and first pregnancy, when alcohol consumption in young women and adolescent girls is higher (Y. Liu et al., 2015).

Steroid hormone levels in the bloodstream (oestrogen, testosterone and sex hormone-binding globulin), increase breast cancer risk, are correlated with mammographic density and increase significantly after alcohol consumption (Frydenberg et al., 2015). Mammographic density higher than 70%, in fact, puts women at bigger risk of developing breast cancer compared to women with lower density. Consuming alcohol increases breast cancer risk from 15% to 51%, according to the amount of alcohol assumed. Early first pregnancy reduces breast cancer risk, while late pregnancy

increases it (Howell et al., 2014). In particular, early full-term pregnancy taken to full term is protective towards receptor positive breast cancer. Breastfeeding for longer periods of time is protective against breast cancer as well, with a more pronounced effect on TNBC and ER+/PR+ breast cancer (Fu et al., 2024). This is particularly accentuated when comparing African American women to Caucasian ones; the former have lower rates of breastfeeding and higher rates of TNBC (Anstey et al., 2017). Pregnancies shorter than 33 weeks show no protective effect (Husby et al., 2018). The protective effect of pregnancy and breastfeeding could be caused by lower exposure to oestrogen and progesterone during this periods, which promote some subtypes of breast cancer growth (Surdacka et al., 2024).

Tobacco smoking is one of the most important risk factors in many types of cancer, including breast cancer (Parkin et al., 2011). Oral contraceptives (OCP) intake increase breast cancer risk, and the risk increases proportionally to the number of years of OCP consumption (Mørch et al., 2017; van Bommel et al., 2022). Both OCP oestrogen, progesterone or combined oestrogen/progesterone based) and hormone replacement therapy increase the risk of breast cancer (Poorolajal et al., 2021).

Women following health recommendations regarding diet, physical activity, smoking and alcohol intake have lower breast cancer risk compared to women who do not follow them (Castelló et al., 2015). Taking part in cancer screening programmes lowers breast cancer mortality.

Genetic factors, including both mutations in high-penetrance genes and polymorphisms in other genes impact the risk of developing breast cancer in combination (polygenic risk score PRS). *BRCA1, BRCA2* and *TP53* are considered high risk genes, while *PALB1, BRIP, ATM* and *CHEK2* moderate risk genes (Lalloo & Evans, 2012). In patients with family history of breast cancer, germline testing is recommended, also to select patients who could benefit from PARP inhibitor therapy. Currently, multigene panel testing (MGPT) allows comprehensive genetic testing, including not only *BRCA1/2* but also *PALB2, TP53, PTEN, STK11* and *CDH11*, whose pathogenic variants can influence decision making regarding surgery, adjuvant therapy and family risk assessment (Botty van den Bruele et al., 2024; Howell et al., 2014).

1.7. Therapy

1.7.1. Surgery

The primary treatment option for breast cancer is surgery, that always includes the sentinel lymph node biopsy (SNLB).

Oncoplastic resection is a therapeutic procedure for patients with biopsy-proven breast cancer, with the objective of completely removing the lesion with clean margins and trying to obtain a good cosmetic result. The procedures includes measuring skin-to-tumour distance and choosing the more appropriate kind of excision to perform, which changes according to the tumour's presence in the upper or lower pole, and the need to conserve the breast (Silverstein et al., 2014)

Mastectomy, a more invasive technique, and axillary dissection significantly reduces the risk of locoregional recurrence compared with breast conserving surgery, axillary dissection and breast radiotherapy (BCT), however a pooled analysis of different randomized trials showed that they have comparable effects on mortality and long-term survival. BCT correlates with risk of locoregional recurrence (Jatoi & Proschan, 2005). Sentinel lymph node biopsy's results are influenced by neoadjuvant chemotherapy administration, which renders the procedure less effective, with a lower detection rate and higher levels of false negatives. This also happens in patients subjected to the biopsy in a too early stage (Kuehn et al., 2013). When not breast conserving, mastectomy can be followed by postmastectomy breast reconstruction, according to the patient's wishes, to ensure quality of life. Breast reconstruction can be implant-based or autogenous tissue-based. Donor tissues include the abdomen, back, buttocks and tights (Cordeiro, 2008).

Patients presenting a small tumour (<2cm in diameter) in the early stage, may benefit from breast conserving surgery, or lumpectomy, or quadrantectomy, which are less invasive, and axillary dissection rather than the more impactful mastectomy, which affects quality of life and mental health while not showing better outcomes regarding disease free and overall survival for this kind of tumours (Veronesi et al., 1981). Conservative surgery shows good results, with less side effects than total mastectomy and a similar long-term overall survival (Luini et al., 2005; Veronesi et al., 2002).

Axillary lymph node dissection (ALND) is fundamental in staging cancer and taking further therapeutic decisions. To avoid invasive dissection of lymph nodes with no cancer cells invasion, only the sentinel node can be biopsied (Veronesi et al., 1997). It can be omitted in node negative invasive breast tumours smaller than 5cm and should not be recommended in early-stage breast cancer with no nodal metastasis, where SLNB should be preferred (Park et al., 2025). Sentinel lymph node biopsy (SLNB) can be recommended in patients cT3-cT4 or multicentric tumours with negative axillary lymph nodes at clinical analysis (Park et al., 2025). SLNB is associated with less shoulder stiffness, pain and swelling in the arm than complete ALND (Loibl et al., 2024).

1.7.2. Endocrine therapy

Endocrine therapies are used in ER+ breast cancer, as they target the oestrogen receptor-activated pathways, either by inhibiting the binding of oestradiol to the oestrogen receptor (Selective oestrogen receptor modulators, or SERMs, of which the most used is tamoxifen) or reduce the blood concentration of oestrogens by inhibiting aromatase (aromatase inhibitors, or AIs). Endocrine therapy is generally administered after surgery. AIs can bind to aromatase either reversibly or irreversibly. Aromatase inhibitors are mostly used in post-menopausal women, while in younger women, tamoxifen is preferred. Tamoxifen is usually administered for 5 years after surgery, even if studies suggest that 10 years of adjuvant therapy may offer better survival rates and lower recurrence (Davies et al., 2013; Tevaarwerk et al., 2016). Endocrine therapy is effective in both early and advanced hormone positive breast cancer (Reinbolt et al., 2015).

1.7.3. Chemotherapy

Chemotherapy can be administered both before surgery, to diminish the size of the tumour and render surgery more effective or even making an inoperable tumour into an operable one (neoadjuvant chemotherapy, NACT) or after (adjuvant chemotherapy, NAT), to eliminate residual tumour cells and decrease recurrence risk. Neoadjuvant chemotherapy is used before surgery, with better outcomes than surgery without NACT (Early Breast Cancer Trialists' Collaborative Group (EBCTCG), 2018). Preoperative chemotherapy is the preferred approach for locally advanced

disease, especially if following surgery is breast conserving (Redden & Fuhrman, 2013).

Adjuvant systemic therapy is indicated in high-risk patients, meaning patients with high grade tumours and according to their receptor status. Adjuvant chemotherapy is suggested for HER2+ tumours and TNBC, especially in young patients (Wöckel et al., 2018).

Anthracyclines (doxorubicine), taxanes (docetaxel, paclitaxel) and platinum based therapies (carboplatin) are among the most commonly used pharmaceuticals both in early and advanced stage breast cancer (Poggio et al., 2018; Qari et al., 2024). Chemotherapy is particularly utilized in TNBC, which lacks receptorial targets. Combined anthracyclines and taxanes reduce the risk of recurrence and mortality (Marra & Curigliano, 2021). Adjuvant chemotherapy is the standard post-surgery treatment for TNBC. Most luminal A tumours do not require chemotherapy, while some cases of luminal b tumours could require it. In HER2+ breast cancer, chemotherapy is usually administered combined with HER2 directed therapy (Loibl et al., 2024).

1.7.4. Radiotherapy

Radiotherapy on ductal carcinoma in situ (DCIS) has been shown to be able to reduce 10-year risk of recurrence by 15%, regardless of age, extent of surgery, endocrine therapy, margin status, grade and tumour size. It is especially effective on older women (Early Breast Cancer Trialists' Collaborative Group (EBCTCG), 2010). Following surgery, radiotherapy is recommended in highly aggressive tumours (T3 or T4 stage), with several positive lymph nodes or increased risk of recurrence such as triple negativity or HER2 positivity and shows a low rate of local recurrence at 5 years. It also has low chronic toxicity (Keller et al., 2012; McCormick et al., 2015). Radiotherapy is effective also in women receiving systemic therapy (cyclophosphamide, methotrexate, fluorouracil, tamoxifen) (EBCTCG (Early Breast Cancer Trialists' Collaborative Group), 2014).

1.7.5. Target therapy

Modern molecular studies on breast cancer have identified several molecular targets,

specific of breast cancer cell, or fundamental for its survival and potentially targeted by drugs.

HER2/neu positive tumours are susceptible to the targeting of this growth factor receptor with recombinant monoclonal antibodies. The first targeted therapy employing an antibody in HER2+ metastatic breast cancer was trastuzumab that induces receptor downregulation (Figure 4). The supplement of trastuzumab to chemotherapy (paclitaxel or doxorubicin) is associated with longer overall, up to 37%, and disease-free survival at ten years. Its most important side effect is cardiac toxicity when combined with anthracyclines. For this reason, regular cardiac assessment is necessary for patients receiving this combined therapy. The combination with non-anthracyclines, in fact, has better outcomes, due to less side effects (Perez et al., 2014; Slamon et al., 2001).

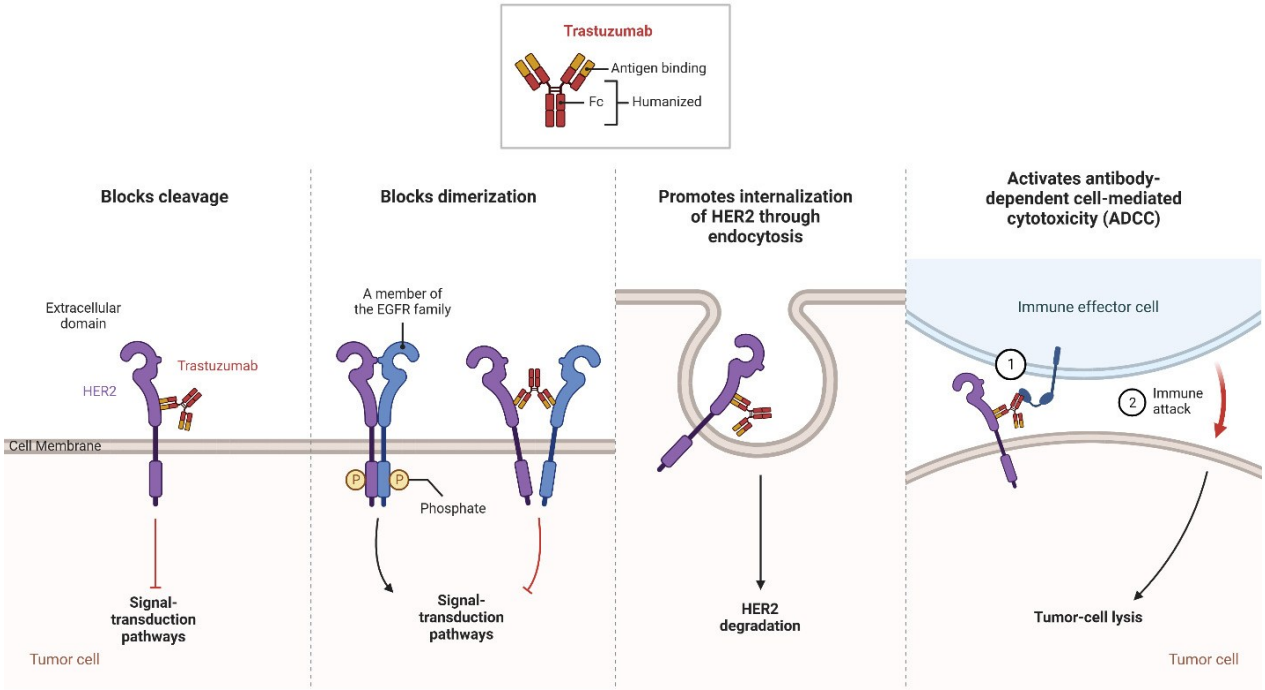

Figure 4. Mechanism of action of trastuzumab, (Fang, 2023)

Another anti-HER2 monoclonal antibody, pertuzumab, shows improved efficacy if administered together with trastuzumab. Pertuzumab targets another portion of the

receptor, blocking its heterodimerization with EGFR. Furthermore, adding docetaxel improves complete response, with more efficacy than chemotherapy with one single antibody. In a small number of patients, the administration of both antibodies brings complete response even without chemotherapy (Gianni et al., 2012; Swain et al., 2015). Antibodies can be bind to other pharmaceuticals, forming antibody-drug conjugates (ADCs), such as chemotherapeutics or radiopharmaceutics, to target the toxicity specifically to the tumour, with minimal side effects (Jacobs et al., 2022).

HER2 tyrosine kinase activity can also be targeted by tyrosine kinase inhibitors (TKI). Tyrosine kinase inhibitors are not specific for HER2 but can target other tyrosine kinase receptors as well. The first approved inhibitor targeted both HER2 and EGFR and is called lapatinib (Singh et al., 2022).

Other targeted therapies, which are not specific for HER2 positive breast cancer include CDK4/6 inhibitors, mTOR inhibitors like everolimus, and PI3K inhibitors for HR positive breast cancer with resistance to endocrine therapy (Jacobs et al., 2022).

Regarding TNBC, which do not express targetable receptors, immune checkpoint inhibitors (ICIs) have been tested, like antibodies inhibiting the PD1/PD-L1 binding (pembrolizumab, atezolizumab) rendering immune response against the tumour more effective (Kwapisz, 2020). In breast cancers with mutation in BRCA1/2, poly-ADP-ribose polymerase (PARP) inhibitors have been demonstrated to be effective, due to the chromosomal instability induced in cells lacking an effective DNA repair system (Jacobs et al., 2022). A potential target specific to TNBC, which could be targeted with ADCs, has been found in trophoblast cell surface antigen 2 (Trop2) (Hu et al., 2024).

1.8. Ethnic differences in breast cancer

While breast cancer is a burden in every nation of the world, affecting women of every race, ethnicity and nationality, the incidence rate, mortality, subtypes distribution and molecular characteristics have been demonstrated to vary according to ethnicity. Nevertheless, especially in the past, most oncology studies did not take into consideration the ethnicity of the patients. Breast cancer studies were carried

mostly on Caucasian women. Cell lines used for studies are derived from Caucasian patients in 80-90% of cases, while in clinical trials, 90% of recruited patients were Caucasian, even when racial minorities represent 40% of the US population (Clarke et al., 2022).

Among women under 45 years of age, Black women have higher incidence rates compared to White women. They also have the highest proportion of grade III-IV tumours, in all age groups, and the highest incidence of TNBC (23% versus 10-15% in other ethnicities). In all age groups, White women have higher incidence of breast cancer compared to Hispanic, Asian, Pacific Islander and Native American women. The most common subtype in all ethnicities is the HR+/HER2- one, however White women have higher rates of HR+/HER2- compared to other ethnicities, while the least common subtype across all ethnicities is the HR-/HER2+, with higher rates in Black and Asian women (Figure 5) (Shoemaker et al., 2018; Telli et al., 2011).

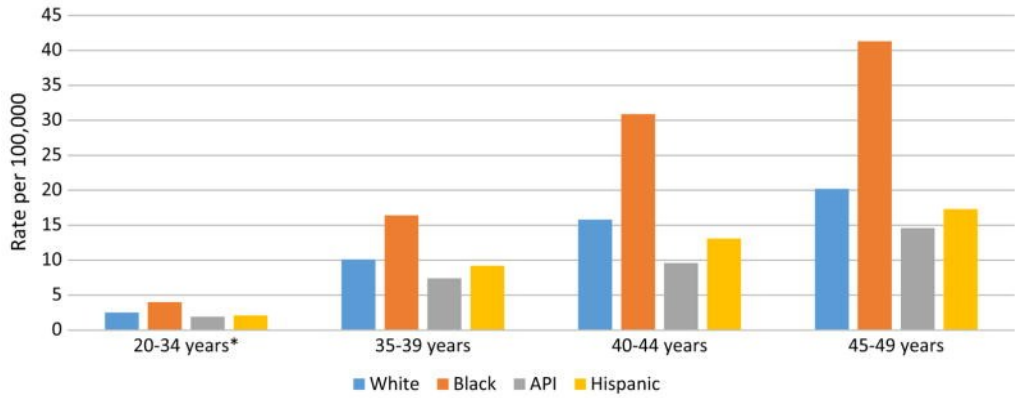

Figure 5. Incidence of triple negative breast cancer in women of various ethnicities, (Shoemaker et al., 2018)

In Norway, a study on immigrants and non-immigrant populations, showed that people from sub–Saharan Africa, South Asia and Eastern Europe have a lower incidence of luminal A breast cancer compared to Norwegian women. South Asian immigrants have a higher incidence of HER2+ tumours (Hjerkind et al., 2022). Furthermore, Black women are less likely to be diagnosed with lobular disease (Williams et al., 2019), Asian patients are less likely to have the BLIS subtype and more likely to have the LAR subtype, while the opposite is true for Hispanics (Ding et

al., 2019).

Many studies have confirmed the higher incidence of ER and PR negative tumours in African American women, who also show higher numbers of highly mitotic and necrotic tumours compared to White women (Porter et al., 2004). The higher frequency of TNBC in Black women also explains their higher mortality rate and worse overall survival as, when adjusting for subtype, the mortality for receptor negative breast cancer is not higher compared to White women (Sparano et al., 2012).

Compared to white, Asian and Pacific Islander women, Hispanic, Black and native American women are more likely to present with larger tumours, higher grades and numbers of lymph nodes and receptor negativity. Japanese women are less likely to have stage III or IV tumours while Filipinos, Hawaiians, Indians and Pakistanis are at higher risk (C. I. Li et al., 2003). African American patients tend to develop more AR-negative tumours compared to Caucasian Americans, and among this quadruple negative kind of breast cancer (QNBC) they have higher frequency of IM and BL1 subtype compared to White patients. Generally, they have lower AR gene expression both in TNBC and non-TNBC tumour types compared to White patients (Davis et al., 2018).

These statistics on Black women are not only relative to the United States, as in a study on Tanzanian and Italian patients it was found that tumours from Tanzanian patients were associated with higher histological grade and more advanced disease. Tanzanian patients also presented more ER negative tumours, higher proliferation (Ki67 expression) and higher numbers of TNBC (Amadori et al., 2014).

Furthermore, even among the same race there are differences: Black American women have higher rates of TNBC compared to Black Caribbean women, who, in fact, have a better overall survival, in particular Black Hispanic patients (Barreto-Coelho et al., 2019). Differences in the incidence and characteristics of breast cancer can also vary not only across wide racial groups but also among smaller ethnic groups, such as different populations in the same subcontinent. One example is a study conducted in Kenya, whose population can be divided into 3 main groups: Bantus, Nilotic and Cushitic. Cushites women are diagnosed at younger ages compared to the other two groups. This could be due to their different reproductive

profiles, with younger ages at first pregnancy and older age at menopause. Nilotes presents more often with ER negative, larger tumours and a higher incidence of lymphovascular invasion (LVI) compared to Bantus (Sayed et al., 2018). In Asia, Chinese and Japanese patients have higher proportions of good prognosis luminal A subtype compared to Filipino and Koreans. Filipinos, instead, have higher proportion of HER2+ cancer, and Koreans higher proportion of TNBC (Chuang et al., 2012).

In studies conducted in the United States, almost all ethnicities have higher probability of presenting with stage IV, receptor negative breast cancer compared to white women, with a strong impact on survival. This is particularly accentuated in Black and Puerto Rican women (Ooi et al., 2011). Ethnic minorities have higher risk of early recurrence compared to White women, which remain after adjusting for socioeconomic factors (Chua et al., 2024; Hill et al., 2018).

Among women with ER/PR+ breast cancer, mortality is similar for Hispanic and Asian American women, but higher for African American women compared to White women (John et al., 2021). In early stage breast cancer the survival is similar (Rizzo et al., 2015), however women belonging to ethnic minorities have double the risk of failing to receive adjuvant treatments and have higher levels of comorbidity (Bickell et al., 2006). The differences in death rates are attenuated when all patients are able to access equal high quality care (Leonard-Murali et al., 2023).

In fact, younger Black women have higher risk of non-initiating adjuvant endocrine therapy and to have delayed treatment initiations (more than 60 days from diagnosis). This is also influenced by financial access, comorbidities and level of communication (Reeder-Hayes et al., 2019; Sheppard et al., 2018). Black patients undergoing aromatase inhibitor treatment have higher median atherosclerotic cardiovascular disease (ASCVD) risk and are more likely to develop hypertension, diabetes and being obese compared to White patients, with an increase in side effects and poor prognosis (Gallicchio et al., 2017). Furthermore Black ethnicity, along with low-income, is associated with a lower probability of being tested using oncotypeDx to guide chemotherapy decisions (Alsten et al., 2024), while Caucasian ethnicity is associated with a higher probability of being tested compared to all other races (Orucevic et al., 2016).

Other than lacking regular care, Black women are less likely to participate in mammographic screening, explaining the advanced stages and larger tumour size observed at diagnosis in this subgroup (Dunn et al., 2024).

1.8.1. Molecular differences in breast cancer across ethnic groups

The differences between ethnic groups do not only include clinicopathological factors, but also mutations, gene expression, protein expression and epigenetics. Population studies in a physiological situation have demonstrated that different populations show genetic diversity. This is not only true across different continents, but even in populations belonging to the same nation, such as Kazaks, Uyghurs and Huis in China. Population specific patterns include genes involved in immunity and metabolism, with different gene expression across groups. Gene expression differences can be regulated by polymorphisms, which are known to be differentially distributed across populations. The identified differences are associated with physiological factors, such as facial morphology, skin pigmentation and hair thickness, however genetic differences between population have also been suggested to take a role in cancer development (Pan, Wen, et al., 2023).

Many studies, in various cohorts of patients have shown differentially expressed genes between ethnicities, including mRNAs, microRNAs (miRNAs) and long non-coding RNAs (lncRNAs) (Ping et al., 2020). In liver hepatocellular carcinoma, the expression of *CCNB1, CCNE2, CHEK1, FOXM1, HRAS, MSH2 RAD51* and *STMN1* is higher in Asian patients, while *TIGAR* is higher in White subjects. In prostate adenocarcinoma, *CTNNB1, FGF19* and *RB1* are more elevated in White individuals, while *FGFR4* and *RASSF7* are more expressed in Black patients. In stomach adenocarcinoma, white patients have higher PRDX1 protein expression while VHL is increased in Asian samples. In general, homology directed repair (HDR) is differentially regulated between Black and White patients, with *CCNB1, CCNE1, CCNE2* and *FOXM1* expression being dependent of the race of the patient in several types of cancer (Lei et al., 2023).

European ancestry is associated with higher expression of stemness markers in basal breast cancer, while in luminal A subtype it is associated with DNA repair, oxidative phosphorylation, MYC targets, pyrimidine metabolism and proteasome. TNF-A

signalling, KRAS target and inflammatory response, on the contrary is negatively associated with European ancestry in luminal A (Telonis et al., 2023). Cell cycle regulators are higher in tumours from African American women: among these are Cyclin E, p16 and p53, while cyclin D1 was lower (Porter et al., 2004). The VEGF-hypoxia signature in their basal breast tumours is upregulated as well, correlating with poor prognosis (Han et al., 2024).

Differentially expressed genes between white and black women include *CRYBB2, PSPHL* and *SOS1*, which play a role in cell growth, invasion, metastasis and immune response. *PSPHL* and *SOS1* are also differentially expressed in benign breast tissue (Field et al., 2012). Other differentially expressed genes are *MKK3*, genes involved in androgen response and interferon alfa signalling, which are also associated with poor prognosis and recurrence in Black patients and genes involved in fatty acid metabolism, IL-6/JAK/STAT3 signalling, that are associated with recurrence in White patients, (Strand et al., 2024; X. Yang et al., 2020). In a TNBC dataset differentially expressed genes between Black and White patients involved cell cycle, chromatin organization, DNA repair and cellular response to stress. DNA repair genes like *XRCC1, XRCC4, BRCA1* and *PARP1* were also upregulated in Black patients with ER+ breast cancer (Mazumder et al., 2022; McAndrew et al., 2024). Black patients have enriched immortalization signature (*TERT, DRAP1, PQBP1*), cell cycle (*RB1*, has-let-7a, *E2F1, c-MYC, TERT*), cell migration and invasion (miR-135b) (Andey et al., 2020).

MicroRNA expression is also variable. In a study on European Americans and Yoruba people from Nigeria 16% of miRNAs were differentially expressed, correlating with differentially expressed mRNAs, many of them correlating with sensitivity to chemotherapeutic agent, immune response, regulation of ion transport, regulation of cell death and cytokine binding. Their expression is influenced by SNPs, whose allele frequencies also differ from an ethnicity to the other. Among differentially expressed miRNAs is the miR-30 family, which is correlated with carboplatin sensitivity (R. S. Huang et al., 2011). Differential alternative splicing is present between African American and Caucasian American patients, due to differential expression of genes involved in splicing, such as *SNRNP70*, which is increased in breast cancer of Black patients compared to White patients (Al Abo et al.,

2021).

Allele frequency of SNPs varies by race, both in protein coding genes and in genes coding for microRNAs. The SNPs associated with breast cancer risk differ in African and Caucasian women, and those affecting both populations are rare (Yao et al., 2013). Among genetic polymorphisms influencing breast cancer risk is a gain of function mutation in the *EDAR* gene, present at high frequencies in Asian and native American populations, causing an increase in the risk of developing metabolic syndrome and correlates with lower breast density, both risk factors for breast cancer (Coletta et al., 2021). In Black women, a variant allele of *BRAF* is associated with higher breast cancer risk, while genetic variants in *AKT, AKT1, RAPTOR, TSC2* are also associated with breast cancer risk and also BMI, fat mass index, waist to hip ratio (WHR), which are also related to breast cancer risk (Ilozumba et al., 2023). A polymorphism in *CCND1* is associated with increased breast cancer risk in Caucasian patients, while the association is not significant in Asian patients (Lu et al., 2009). The rs10069690 SNP in *TERT* is more common in Black patients, and associated with higher odds of developing basal-like breast cancer (Huo et al., 2017).

The insulin like growth factor, growth hormone and leptin pathways have a role on cancer development and are subjected to germline genetic variations. In African American women, SNPs in *BAIAP2, CALM2, CSNK2A1, BRAF, BAD* and *MAPK3* were associated with breast cancer (Ruiz-Narváez et al., 2016). Black women also have variants in various hormone pathways are associated with breast cancer (Haddad et al., 2015).

In a cohort of northern Indian patients a polymorphism in the *CYP17* gene was associated with early onset breast cancer risk (Chakraborty et al., 2007). A study on Tunisian women, instead, highlighted the contribution of *TGFB1* polymorphisms to breast cancer susceptibility and pathological presentation, with associations with the TNBC subtype, tumour size, distant metastasis and response to therapy (Hadj-Ahmed et al., 2019).

Korean patients have a higher incidence of *BRCA1/2* mutations than the TCGA cohort, especially in younger patients. They also present more frequently *TP53* and *ERBB2* mutations or amplifications, despite a lower mutational burden (Kan et al.,

2018). TNBC shows a different mutational landscape in African American and Caucasian patients. *TP53* is mutated in 46% of Black patients, and only in 27% of White, while *MLL3* is mutated in 12% of Black patients and 6% in White (Miyashita et al., 2023; Strand et al., 2024). *PIK3CA* and *NCOR1* are more frequently mutated in White TNBC patients compared to Black ones (Omilian et al., 2020), while *KMT2C* mutations are more frequent (Miyashita et al., 2023). Black TNBC patients usually have higher numbers of truncating mutations than Whites (Q. Liu et al., 2019).

When studying hereditary breast cancer genes (*BRCA1, BRCA2, CHEK2, ATM*), White and Black patients have the highest frequency of pathogenic variants, while Asian and Hispanic patients show the highest frequency of variants of uncertain significance (VUS) (T. Jones et al., 2021).

DNA methylation levels can also vary by ethnicity. In a study comparing normal breast tissue of European Americans and African Americans, almost 500 CpG islands were found to be differentially methylated, with methylation levels correlating with differential gene expression across the two groups. The region where the methylation occurs also varies by ethnicity, being more frequent in promoter regions in European Americans and in gene body in African Americans (Song et al., 2015). ER-, PR- tumours in Black women under 50 years of age have higher percentage of methylation compared to Caucasian women of the same age group in *HIN-1, TWIST, Cyclin D2* and *RASSF1A* (Mehrotra et al., 2004). *FOXA1* is hypermethylated and lowly expressed in Black women in particular due to parity without breastfeeding, correlating to the higher frequency of ER- tumours in this group (T.-Y. D. Cheng et al., 2019; Sribenja et al., 2021).

DSC2, KCNK4, GSTM1, AXL, DNAJC15, HBII-52, TUSC3 and *TES* were found to be differentially methylated in the Carolina Breast cancer study, which compared Black patients with White patients, with *DSC2, KCNK4, AXL* and *TUSC3* having higher methylation in Black patients and *GSTM1, HBII-52, TES* and *DNAJC15* being methylated at higher levels in white patients. *AXL, DNAJC15, KCNK4*, and *GSTM1* were also confirmed to be differentially methylated through the analysis of The Cancer Genome Atlas (TCGA). Differential methylation in *AXL, GSTM1, DSC2* and *DNAJC15* was also observed in lymphoblastoid cell lines, highlighting that they are probably correlated to ethnicity and not specific of the disease (Conway et al., 2015).

FOXA1 is hypermethylated in tumours from Black patients, correlating with reduced RNA and protein expression. In Black women, higher methylation is associated with number of children and non-breast feeding (Espinal et al., 2017). *CCNE1* is more methylated and less expressed in White samples, while PTEN is more methylated in Black patients (Lei et al., 2023).

The tumour microenvironment (TME) and extracellular matrix (ECM) around the tumour show marked differences between different ethnicities, as well (Stone et al., 2023). A Korean cohort showed a more immune reactive TME, with higher numbers of tumour infiltrating lymphocytes (TILs), while TME in TCGA samples, which are mostly Caucasian, is more immune suppressed (Kan et al., 2018).

Black patients have higher density of CD163+ tumour associated macrophages (TAMs) and immunosuppressive macrophages (M2) compared to Hispanic and Caucasian patients, with highly proliferative activity, which correlates with their lower overall survival (Koru-Sengul et al., 2016), caused likely by higher levels of chemotaxis, neovascularization and tumour vessel density (Martin et al., 2009). Black patients also have a more exhausted CD8+ T cell profile, which is associated with poor survival (Yao et al., 2021). African Americans have higher expression of cytokines (*CD47*, *TGFB1*, *NFKB1*), associated with the transcriptional repressor *Kaiso*. This difference is present not only in the tumour, but also in the exosomes derived from the tumour. The expression of *Kaiso* is associated with breast cancer, copy number alterations and M2 macrophage polarization in women of African ancestry (Ahmed et al., 2023).

Regarding breast cancer stem cells, CD44+/CD24- cells and *ALDH1* have been associated with DFS and OS in Asian patients, but not in European ones, and with TNBC or basal subtype in all races (Gyan et al., 2021).

Variations in genetics, epigenetics, transcriptomic and immune infiltrate, influence not only the prognosis and characteristics of the tumours, but also the response to therapy. *ABCB1* is a gene involved in resistance to various chemotherapies, including paclitaxel, as it limits the absorption of xenobiotics and promotes their efflux into the bile and urine. It has four major SNPs, whose frequency varies according to ethnicity. Some of its polymorphisms are more common in Asian populations and confer a

certain resistance to paclitaxel, while experiments on non-Asian cells lines show them to be more sensitive to the drug (Kwon, 2009). *NOTCH4* is increased in Egyptian patients who received neoadjuvant chemotherapy, while in English patients who received the same treatment, *NOTCH1* is increased (Kamal et al., 2018). In TNBC cell lines derived from Black women, *RAD51* is overexpressed, epigenetically regulated by miR-214-5p, which is downregulated. As *RAD51* has a role in double strand break repairs, its high expression correlates with resistance to PARP inhibitors like Olaparib, playing a role in the worse prognosis of Black TNBC patients (Mani et al., 2023).

Despite these data, clinical trials are performed mostly on Caucasian patients, and cell lines, biobanks samples and genomic data from public datasets are in majority from patients of Caucasian ethnicity or do not data regarding the ethnicity of the patients (Guerrero et al., 2018). Given that differences in prognosis are ascribable not only to health and social inequalities but to biological differences as well, it is imperative to characterize tumours of patients belonging to less-studied ethnic groups, for a better understanding and management of the disease.

1.9. Ion channels in breast cancer

Among biomarkers which have gained importance in the latest years in oncology are ion channels and transporters (Figure 6). Expression profiling of ion channels in breast cancer can help in profiling the disease. A 30-gene signature has been shown to predict clinical outcome, with ion channel profiling varying according to p53 mutational status, ER status and histological grade (Ko et al., 2013).

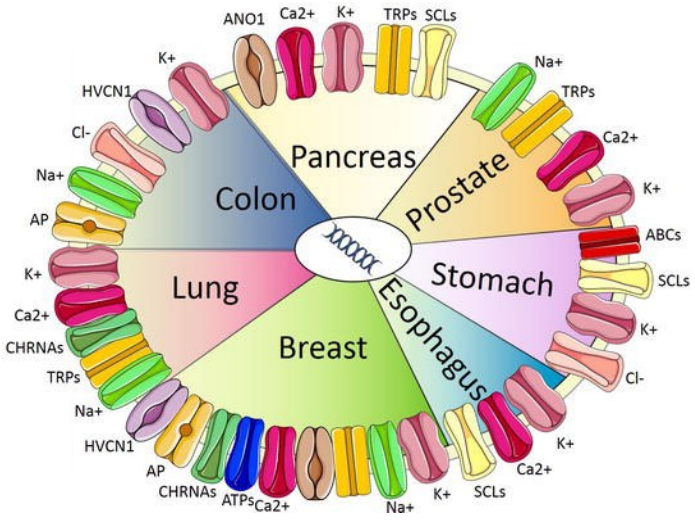

Figure 6. Ion channels and their role in different types of cancer (Iorio et al., 2019)

These molecules can be employed not only as biomarkers, but also as therapeutic targets, due to their role in various features of oncogenesis. Ion channels involved in cancer include potassium, sodium, calcium and chloride channels, aquaporins and transporters (Lastraioli et al., 2015). Sodium channel *SCN5A* and in particular its neonatal splice variant (nNav1.5) is highly expressed in breast cancer, promoting metastasis. In fact, sodium currents have been demonstrated to influence migration and invasion capability of breast cancer cells, and their blocking with tetrodotoxin (TTX) lowers cell invasiveness (Roger et al., 2003). *SCN5A* and its neonatal form are upregulated in TNBC, promoting cell proliferation, migration and invasion; the inhibition of the channels induces apoptosis and enhances the effect of paclitaxel (Erdogan et al., 2023). Downregulating nNav1.5 increases cell sensitivity to tetrodotoxin and suppresses cell migration. High expression of nNav1.5 correlates not only with higher migration ability but also with glutamate levels which, in turn, correlate with bone metastasis (Azahar et al., 2022; Brackenbury et al., 2007).

Potassium channels have been proved to control breast cancer metastasis through the β-catenin signalling pathway in TNBC (Breuer et al., 2019). *KCNH1*/Eag1/Kv10.1 was the first potassium channel which was associated with oncogenesis (Pardo & Stühmer, 2008). Physiologically, it is present only in the central nervous system, but it was found to be expressed in tumour cells, regulated by *p53* and *E2F1*. It has a role in

cell proliferation, survival, angiogenesis, migration and invasion (Pardo et al., 1999). *Eag1* is upregulated in both invasive and non-invasive cancer cells, regulating calcium influx, cytoskeleton rearrangement and therefore cell migration. Its expression also correlates with the expression of *HIF1α*, suggesting a role in hypoxia (Downie et al., 2008; Lai et al., 2014; Ouadid-Ahidouch et al., 2016). The expression of Eag1 channel is linked with malignancy, metastasis and poor prognosis in breast cancer. Its expression is higher in TNBC compared to other subtypes, and correlated with tumour size, advanced stage and positive lymph nodes (G.-X. Liu et al., 2015). Blocking Eag1 currents inhibits proliferation in all subtypes of breast cancer. Astemizole, a Eag1 blocker, has been tested in cell cultures in combination with a TKI, gefitinib, resulting in a reduction of cell proliferation and increase of apoptosis (García-Quiroz et al., 2019), and with calcitriol, resulting in a suppression of tumour growth (García-Quiroz et al., 2014). The expression of *Eag1* is, in fact, downregulated by calcitriol through a vitamin D receptor (VDR) dependent mechanism (García-Becerra et al., 2010).

Voltage gated potassium channel Kv1.3 plays a role in proliferation and apoptosis as well and it is expressed in leukemic T cells, prostate and breast cancer. In breast cancer, it shows upregulation in stage I, but it decreases with stage, increasing again in advanced stage and metastatic breast cancer. Its low expression in high stages and poorly differentiated breast cancer is due to promoter hypermethylation (Brevet et al., 2009). Its expression can be lowered by statins, potentiating doxorubicin's chemotherapeutic activity; its currents can be blocked with tetraethylammonium (TEA), and by tamoxifen blocking cell proliferation (Jang et al., 2009; Teisseyre et al., 2019; Thomas et al., 2003).

KCNMA1, a calcium activated potassium channel, shows genomic amplification in prostate, breast, ovarian and endometrial cancer of higher stages and grades. Its expression is upregulated in breast cancer, promoting proliferation (Oeggerli et al., 2012) and especially in breast cancer metastasizing to the brain compared to breast cancer metastasizing to other organs. It also promotes trans endothelial migration (Khaitan et al., 2009). *KCNK9* gene is also present in a genomic region frequently amplified in breast cancer, resulting in an upregulation of the channel tumorigenesis and resistance to hypoxia (Mu et al., 2003).

Kca3.1/*KCNN4* plays a role in tumorigenesis in various oncological diseases. In breast cancer, it promotes cell proliferation, calcium entry together with *TRPC1* and radioresistance (Mohr et al., 2019). High expression of Kca3.1 correlates with poor prognosis (Faouzi et al., 2016)

The expression hERG1 potassium channel correlates with breast cancer subtype, being the highest in luminal A and the lowest in basal tumours. Moreover, it correlates with ER expression and low proliferative index, and therefore with a better progression free survival (Iorio et al., 2018). Hyperstimulation of hERG1 through NS1643 inhibits breast cancer cell proliferation, causes cell cycle arrest and activates a senescence program through ROS production and DNA damage (Fukushiro-Lopes et al., 2017; Lansu & Gentile, 2013).

The expression of *KCNK5, KCNK9* and *KCNK12* in TNBC is upregulated, while the one of *KCNK6* and *KCNK15* is downregulated. Their expression patterns correlates with CpG methylation (hypomethylation for *KCNK5* and *KCNK9*) (Dookeran et al., 2017).

The expression of *AQP1* is associated with high tumour grade, TNBC subtype and poor prognosis (Otterbach et al., 2010). AQP5 is highly expressed in breast cancer, promoting cell proliferation and migration (Jung et al., 2011).

The *NHE1* exchanger dysregulates pH in the microenvironment, with a protumorigenic activity. It also promotes invasion by regulating metalloproteinases (Amith & Fliegel, 2013).

TRPC3 and *TRPC6* are upregulated in breast cancer cell, forming a complex. Silencing of the corresponding genes reduces breast cancer cell growth (Aydar et al., 2009).

CACNA2D3 expression is upregulated by chemotherapeutic azacytidine in cells showing CpG methylation. The CpG islands of the gene are frequently methylated in breast cancer metastasizing to the central nervous system, regardless of subtype (Palmieri et al., 2012).

Finally, chloride channel *CLCA2* is often downregulated and its promoter hypermethylated in breast cancer; it also represents a predictive factor for resistance

to doxorubicin. Furthermore, its genic locus, 1p31, is often deleted in sporadic breast cancer (X. Li et al., 2004; Walia et al., 2009). Its expression is induced by p53 in response to DNA damage, by binding to its promoter, and it inhibits cell migration and invasion. *CLCA2* inactivation also enhances the expression of focal adhesion kinase (FAK) and its signalling pathway (Sasaki et al., 2012). *CLCA2* and its interactor EVA1 are associated with mammary epithelial differentiation (Ramena et al., 2016). It colocalizes with integrin b4, playing an important role in cell adhesion: in fact, loss of *CLCA2* promotes epithelial-mesenchymal transition, promoting metastasis (Connon et al., 2004; Walia et al., 2012). Notably, the expression of *CLCA2* is associated with a better overall survival in Black women with TNBC (Purrington et al., 2020).

1.10. Long non-coding RNAs in breast cancer

Another class of molecules, playing a role as regulatory elements in a wide array of biological processes, is represented by long non-coding RNAs (lncRNAs). lncRNAs are defined as transcripts longer than 200 nucleotides, transcribed by RNA polymerase II and with a role in regulating gene expression through chromatin binding, interactions with chromatin-modifying complexes and through RNA interference (RNAi) (Guttman et al., 2009; Khalil et al., 2009; Martianov et al., 2007).

They have also been identified as critical components in breast cancer. They show different expression levels according to the subtype of breast cancer: NEAT1 and GAS5, for example are downregulated in TNBC compared to luminal subtypes, and XIST is downregulated in TNBC compared to HER2+ tumours (Marcu et al., 2021; Salama et al., 2020). Other lncRNAs are instead overexpressed in TNBC, such as THOR, AFAP1-AS1 and LINC01550 (García-Hernández et al., 2024; Y. Liu et al., 2024). Their expression can also be detected in plasma, and was confirmed for ANRIL, HIF1A-AS2 and UCA1 (M. Liu et al., 2017).

The expression of lncRNAs is variable even across different subtypes of breast cancer (Beltrán-Anaya et al., 2019), such as that it has even been employed to classify TNBC tumours in subtypes, similarly to how they were classified by Lehmann and collaborators. This subtypes are defined as LINC0511-enriched, overlapping with

BLIS and IM and characterized by pronounced immune infiltration, LINC00393-enriched, still overlapping with BLIS but with reduced immune signalling and poorer survival; finally, are the FIRRE-enriched and normal-like subtypes (Vishnubalaji et al., 2022)

Moreover, the expression of lncRNAs has been correlated with prognosis, with HOTAIR being linked to recurrence and AC105383 with poor prognosis (M. F. Evans et al., 2020; Q. Liu et al., 2020).

In breast cancer, lncRNAs take part in a variety of processes, such as DNA damage and DNA damage response. LINP1 enhances DNA double-strand break repair, by acting as a scaffold for the non-homologous end joining (NHEJ) pathway, in fact its inhibition sensitizes TNBC cells to radiotherapy and doxorubicin (Y. Zhang et al., 2016).

lncRNAs have an effect on metabolism as well: GHET1 enhances hypoxia induced glycolysis, promoting cell proliferation and invasion, while LINK-A activates HIF1A through phosphorylation even in normoxic conditions (A. Lin et al., 2016; Y. Wang & Liu, 2021).

Moreover, lncRNAs regulate cancer stem cells (CSCs), promoting stem-like features, the expression of stemness markers (Shima et al., 2018; Z. Xu et al., 2020) and, consequently, epithelial-mesenchymal transition (EMT). LINC00921, which would be a tumour and EMT suppressor, is downregulated in TNBC (J. Zhang et al., 2022), while MALAT1 is upregulated and enhances EMT features in hypoxia (Dragonetti et al., 2024). The effect of lncRNAs is exerted also on angiogenesis: AFAP1-AS1 induces hypoxia-induced angiogenesis and vasculogenic mimicry and its targeting could be part of antiangiogenic therapy. A similar role is performed by HOTAIR, which is strongly upregulated during hypoxia (García-Hernández et al., 2024; Lozano-Romero et al., 2020)

lncRNAs influences the tumour microenvironment. M2 macrophage polarization is induced by MALAT1, which is expressed in extracellular vesicles and activates the Hippo/YAP pathway, promoting an immunosuppressive TME (Adewunmi et al., 2023)

Moreover, lncRNAs have also been identified as therapeutic targets: siRNAs targeting DARS-AS1 inhibit its activity in enhancing migration, EMT and invasion, suppressing tumour growth and sensitizing tumour cells to doxorubicin (X. Liu et al., 2022, 2023).

2. Aim

The primary objective of this thesis was to identify and characterize the molecular and genetic features of breast cancer across different ethnic groups, with a particular emphasis on transcriptomic analyses.

We investigated gene expression profiles obtained from both publicly accessible databases and breast cancer samples analysed using RNA sequencing, followed by comprehensive bioinformatic evaluation.

In addition to examining protein-coding genes, this study explored the role of long non-coding RNAs (lncRNAs) and their contribution to breast cancer pathology.

Finally, special attention was given to ion channels and transporters, which are increasingly recognized as key players in cancer progression, influencing cellular proliferation, migration, and metastasis.

3. Materials and Methods

3.1. BRCA public datasets

Clinical features, gene expression and mutational data of 528 BRCA patients were obtained from The Cancer Genome Atlas (TCGA), with open access. Only patients whose available data included survival data (.tsv format), subtype, race, gene expression (tpm files) and somatic mutational (maf files) data were selected. Data for the survival analysis was downloaded through UCSC Xena Browser (University of California, Santa Cruz, http://xena.ucsc.edu/) (Goldman et al., 2020). The accession number for this dataset was phs000178.

Transcriptomic and mutational data were downloaded, prepared and normalized using the R (version 4.4.0) package 'TCGAbiolinks' version 2.32.0 (https://bioconductor.org/packages/TCGAbiolinks/) (Colaprico et al., 2016; Mounir et al., 2019), an R package for the retrieval and analysis of TCGA data.

The project "TCGA-BRCA", data category "Transcriptome Profiling" or "Simple Nucleotide Variation", data type "Gene Expression Quantification" or "Masked Somatic Mutation" and the barcodes of the patients with "Asian", "Black or African American" and "White" in the "Race Category" category were selected, as it is the classification used in TCGA data. Furthermore, patients were divided according to their subtype into "Basal", "HER2+", "Luminal A" and "Luminal B".

Healthy controls for the comparison with our TNBC samples were downloaded from the GTEX database (Lonsdale et al., 2013), selecting "Breast Mammary Tissue", and selected based on sex (female) and age similar to the patients', using the accession number phs000424.v10.p2. 16 samples were used, the same number of patients in the TNBC cohort.

3.2. Subtyping of TNBC

Classification of triple negative tumours into different TNBC subtypes was performed employing the online tool TNBCtype (https://cbc.app.vumc.org/tnbc/) (Chen et al., 2012), which is able to assign TNBC subtypes to samples according to gene expression.

3.3. Somatic mutations analysis

The presence of somatic mutations and mutational signatures in TCGA samples was evaluated with maftools, v2.24.0 (Mayakonda et al., 2018). Maftools is a package for the analysis and visualization of genomic data, employing files in the Mutation Annotation Format (MAF).

3.4. Patient selection

The RNAseq analysis was conducted on 16 samples of triple negative breast cancer, 2 HER2+ breast cancer (one African and one European), 7 luminal A breast tumour (one Arab, three Asian, three European), and 4 luminal B breast tumour (4 European and one African), selected from the section of Pathological Anatomy of the Department of Health Sciences, University of Florence. Clinical data were obtained from Archimed database of Oncological Radiotherapy, department of Oncology, Careggi University Hospital, Florence. Ethical approval for the use of these archival samples was obtained from the ethical committee.

3.5. RNA extraction from FFPE tumour samples

RNA was extracted from formalin fixed paraffin embedded tumour samples. Only samples with tumour cellularity higher than 80% were selected. For each sample 2 sections of 10mm each were deparaffinized with xylene and RNA was extracted using the AllPrep DNA/RNA FFPE (QIAGEN, cat. no. 80234) kit as per the manufacturer guidelines. Briefly, xylene was eliminated with 100% ethanol, pellet was resuspended, treated with proteinase K at 56°C for 15 minutes and cooled on ice for 3 minutes. Samples were then centrifuged at 20000 g for 15 minutes to separate pellet (containing DNA) from supernatant (containing RNA). The supernatant was then incubated at 80°C for 15 minutes to remove modification in the RNA caused by formaldehyde. Extraction was performed in order to extract not only messenger RNA but also microRNAs. Supernatant was then transferred to an RNAeasy MinElute spin column to isolate RNA, washed with wash buffers, treated with DNAse I to remove DNA contamination and washed again. The column membrane was then dried, and RNA was eluted in 14 ul of RNAse free water. RNA was quantified by measuring absorbance at 260nm with a spectrophotometer. Eluted RNA was stored at -80°C.

3.6. RNA sequencing, transcripts identification and quantification

RNAseq was performed on extracted RNA by Novogene, using the Illumina platform NovaSeq 6000 (Pipeline in Figure 7). rRNA was removed from total RNA by using specific probes, and mRNA was purified using poly-T oligo-attached magnetic beads. cDNA was synthetized using random hexamer primers and dNTPs, and adapters were then added. PCR products were then purified with AMPure XP beads to obtain strand specific library. The library was then quantified using Qubit and Bioanalyzer (Agilent Technologies, CA, USA) for size distribution detection.

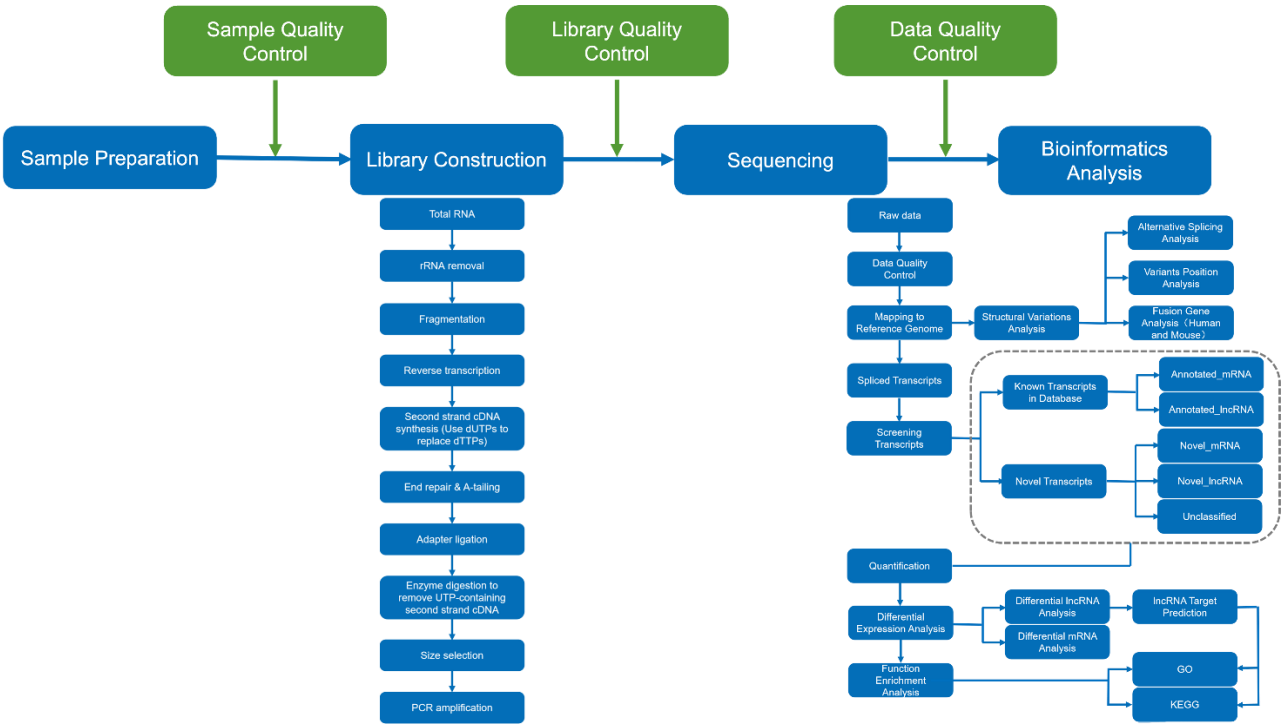

Figure 7. *Pipeline of Illumina sequencing and bioinformatic analysis*

Fluorescence obtained from Illumina platform were transformed into short reads, which were then recorded in FASTQ by base calling using CASAVA. The raw reads were filtered to obtain the clean reads and improve the quality of analysis. Reads containing adapters, containing more than 10% of bases that could not be determined and low-quality reads (reads with Qscore/Quality value <= 5 in 50% of the bases) were excluded. Clean reads were mapped to the reference genome using HISAT2 software (v2.2.1). A total of 12 million raw reads per sample were obtained.

Stringtie (v2.2.3) was then used to concatenate reads into transcripts and quantify

them (Pertea et al., 2015) and to merge them. The reference genome used was Homo Sapiens(GRCh38/hg38). The assembler transfrags were then compared to reference transcripts with Gffcompare (v0.12.6) to determine if they were sufficiently different to be considered novel in order to identify novel genes, exons, optimize the start and end information of known transcripts and predict the coding potential of the selected new transcripts. The protein coding potential of novel transcripts was evaluated through three different algorithms: Coding-Non-Coding Index (CNCI), Coding Potential Calculator (CPC), and Protein Function Prediction (PFP). CNCI (L. Sun et al., 2013) is a signature tool able to distinguish protein-coding and non protein-coding sequences. CPC (Kong et al., 2007) can access the protein-coding potential of a transcript based on the quality of the open reading frame (ORF) and other features. Finally, PFP (Hawkins et al., 2006) is able to assign functions and annotations to transcripts. Only novel sequences which were identified as non-coding by all three algorithms were categorized as novel lncRNAs.

Expression levels of mapped, spliced, screened transcripts and predicted transcripts were then quantified, by estimating the abundance of transcripts. FPKM (Fragments Per Kilobase of transcript sequence per Million base pairs sequenced) was used to estimate gene expression levels. To identify putative target genes and the potential function of novel lncRNAs, a Pearson correlation between transcripts and known mRNA was performed, with positive correlation set as $r > 0.7$.

Variant site analysis was performed as well, employing GATK (v4.6.0.0) (McKenna et al., 2010) and SnpEff (Cingolani et al., 2012).

3.7. Gene expression analysis

Principal component analysis of the TCGA cohort was performed with PCAtools, v2.20.0 (Blighe & Lun, 2025), employing gene expression. PCA plots, boxplots and barplots were designed with ggplot2, v4.4.0 (Wickham, 2016).

Employing the R package DESeq2 1.44.0 (https://www.bioconductor.org/packages/DESeq/) (Love et al., 2014), which identifies differentially expressed genes between different conditions, differential expression analysis to evaluate transcriptomic data differences across two groups of patients was performed: patients belonging to different ethnicities (Asian vs Black or

African American – Asian vs White – Black or African American vs White). The patients with basal subtypes were compared with each of the other subtypes (HER2+, Luminal A, Luminal B). For all comparisons, genes with a fold change higher than 1.5 or lower than -1.5 and with an adjusted p-value <0.05 were considered differentially expressed. Adjusted p-value was calculated using the Benjamini-Hochberg method control the false discovery rate (FDR).

Differential expression analysis was also performed on primary TNBC samples, comparing them to healthy controls from the GTEX database (Lonsdale et al., 2013) and between patients of different ethnicities in our cohort. edgeR (v4.0.16) (Robinson et al., 2010) was employed by Novogene in this case to analyse differential expression in sequenced samples, due to its better capability of handling samples with high RNA expression variability.

3.8. Enrichment analyses

Enrichment analyses were performed employing the R packages 'clusterProfiler' version 4.12.2 and 'enrichplot' version 1.24.2 (https://bioconductor.org/packages/enrichplot/; https://yulab-smu.top/biomedical-knowledge-mining-book/) (T. Wu et al., 2021; Yu et al., 2012, 2015), which provide gene functional annotation and the possibility to plot enrichment results. Differentially expressed genes (DEGs) were assigned to different pathways using the Gene Ontology (GO, https://geneontology.org/) (CRAN: Survival Citation Info, n.d.), Kyoto Encyclopedia of Genes and Genomes (KEGG, https://www.kegg.jp/) and Disgenet (DGN, https://disgenet.com/) databases. Enrichment analysis was performed on significant differential expressed genes, with threshold set as adjusted p-value < 0.05.

3.9. Survival analyses

Survival curves were plotted employing the R package 'survival' version 3.7.0 (https://cran.r-project.org/web/packages/survival) (Therneau, 2024). Patients were divided into groups according to the gene expression of the selected gene or novel lncRNA, using the median value of each gene as a cutoff, and their effect on survival was evaluated using the log-rank test statistical method.

4. Results

4.1. Breast cancer subtype, race and incidence in the TCGA cohort

Before conducting transcriptomic data analysis, we assessed the distribution of different subtypes of breast cancer across different ethnicities, and the proportional contribution of each ethnicity to the total case numbers for each subtype (Figure 8).

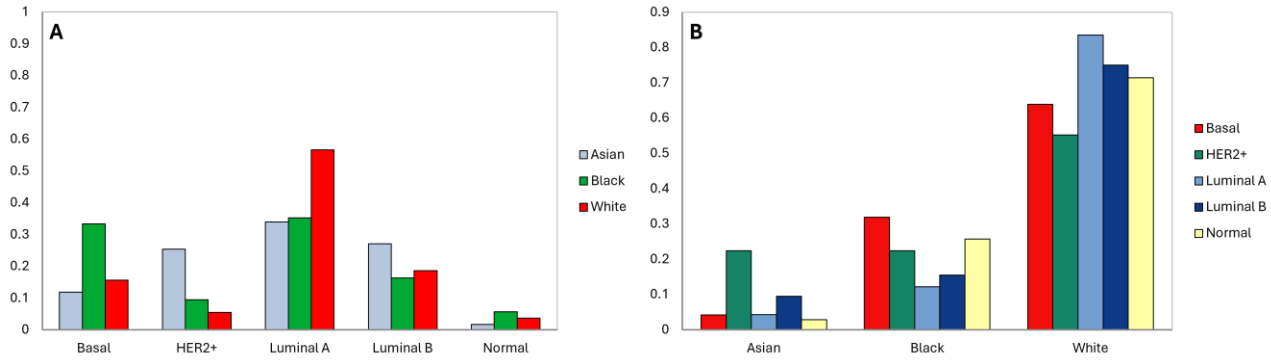

Figure 8. A) Breast cancer subtype distribution in Asian, Black and White patients B) Incidence of different subtypes of breast cancer in Asian, Black and White patients.

Characteristics of selected patients are summarized in Table 1.

	Category	**Number of patients**	**Percentage (%)**
Age (years)			
	<40	69	7.81
	40-49	182	20.61
	50-59	235	26.61
	60-69	225	25.48
	>= 70	172	19.48
Subtype			
	Basal	163	18.46
	HER2+	65	7.36

	Luminal A	453	51.30
	Luminal B	167	18.91
	Normal	35	3.96
Race/Ethnicity			
	Asian	*58*	*9.22*
	Black or African American	*150*	*23.85*
	White	*421*	*66.93*
Stage			
	Stage I	161	18.23
	Stage II	516	58.44
	Stage III	193	21.86
	Stage IV	13	1.47

Table 1. Demographic and clinical characteristics of selected TCGA breast cancer patients

In the basal subtype, Black patients constituted the biggest fraction, while Asian patients were the most numerous in the HER2+ subtype and White patients in the luminal A one, with results aligning to published literature. Similar observations were made in the subtype distribution in each ethnicity, with the basal subtype being the most frequent in Black patients, the HER2+ the most frequent in the Asian cohort and the luminal A the most frequent among White patients.

We then performed principal component analysis (PCA) on breast cancer samples based on transcriptomic data. Each subtype was well separated from the others, with the exception of the luminal subtypes, which show some overlap (Figure 9A). The basal subtype was clearly separated from the other subtypes, while the HER2+ one appeared as intermediate between luminal subtypes and the basal one (Figure 9A).

Different ethnicities, instead, did not show a clear-cut division between each other (Figure 9B).

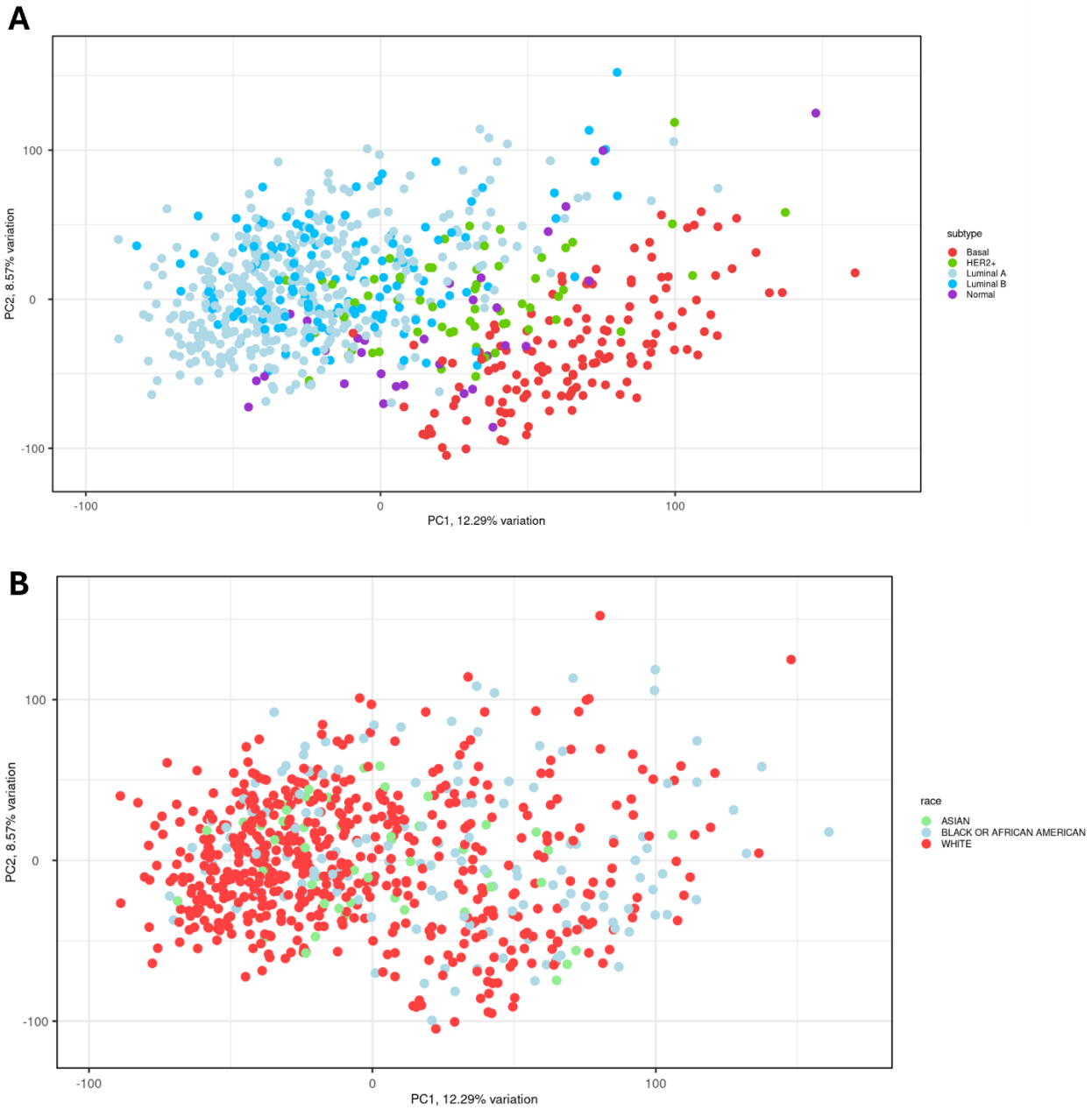

Figure 9. Principal component analysis based on gene expression of TCGA samples. Colours of the sample represent different A) breast cancer subtype and B) ethnicity

When examining differences within each subtype, no separation can be observed, probably due to the low number of non-White patients, at least in part (Figure 10). Conversely, separation by subtype is evident within each individual ethnic group, with basal and luminal subtypes being distinctly separated from one another (Figure

11).

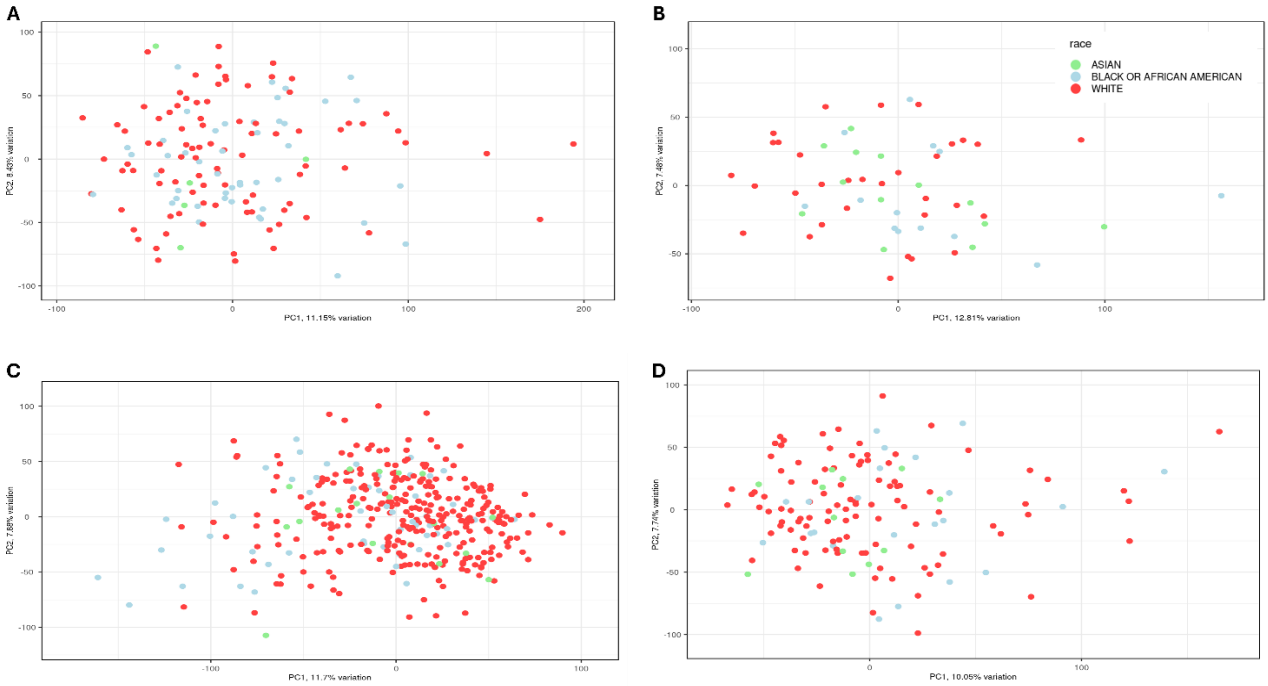

Figure 10. Principal component analysis in different breast cancer subtypes, with different colours representing different ethnicities. A) Basal subtype B) HER2+ subtype C) Luminal A subtype D) Luminal B subtype

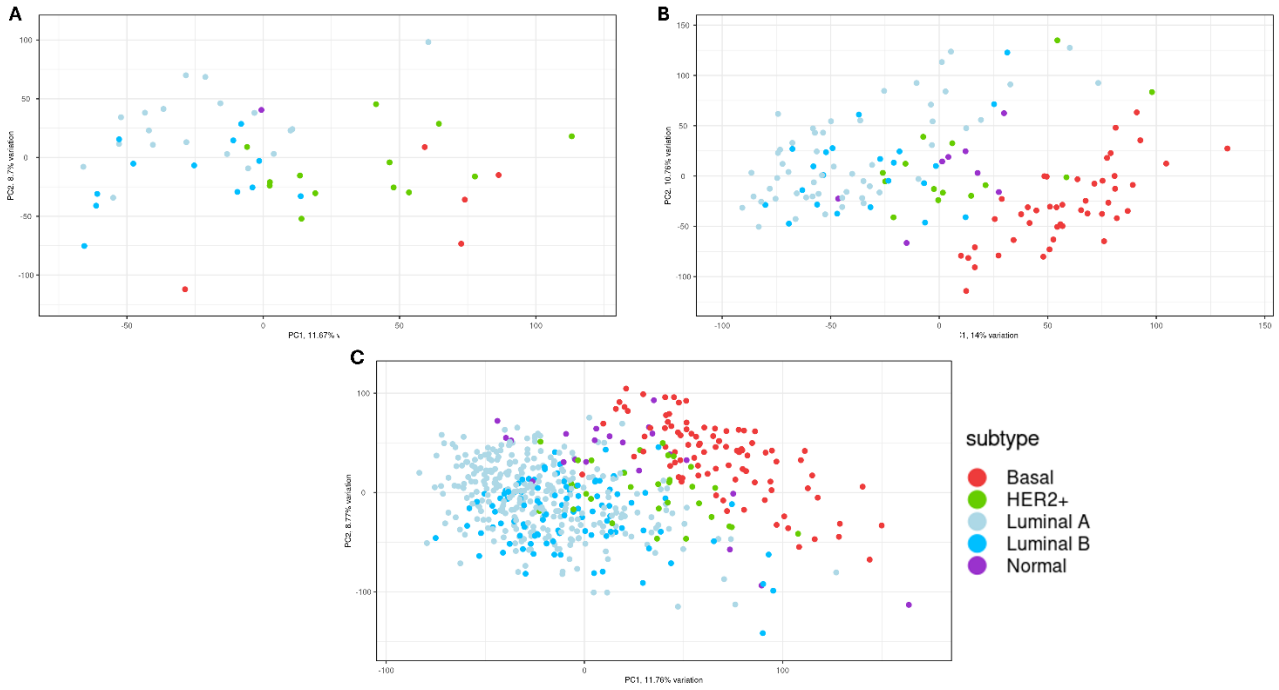

Figure 11. Principal component analysis in A) Asian patients B) Black patients C) White patients. Different colours indicate different subtypes

4.2. Differential gene expression and pathway enrichment across different ethnicities

We performed a first differential expression analysis between patients belonging to different ethnic groups (Asian vs Black or African American; Asian vs White; Black or African American vs White), identifying several differentially expressed genes (DEGs) (Figure 12). However, due to the intrinsic differences between breast cancer subtypes and the different distribution of subtypes across ethnicities, we decided to perform another differential expression analysis, between ethnicities in the same subtype. Differential expression analysis was normalized according to tumour stage.

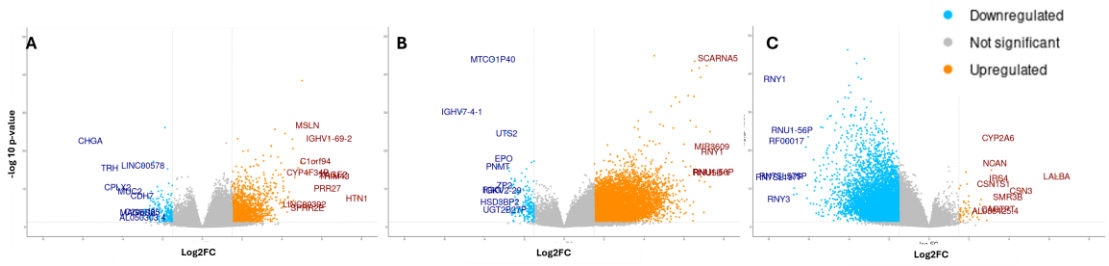

Figure 12. Volcano plot representing differentially expressed genes between A) Asian and Black patients, B) Asian and White patients and C) Black and White patients. Orange dots: upregulated genes. Blue dots: downregulated genes. Grey dots: non-significant genes

We therefore obtained differential gene expression between ethnicity by subtype and performed enrichment analysis on resulting DEGs.

In the basal subtype (Figure 6), Asian patients have higher expression of genes involved in ErbB signalling pathway compared to Black patients (Figure 13A), while White patients have higher hormone signalling. Black patients, instead, show an upregulation of chemical carcinogenesis through DNA adducts and a downregulation of cell adhesion molecules, compared to Asian patients. This pattern aligns with the higher aggressiveness and metastatic potential observed in breast tumours among Black patients.

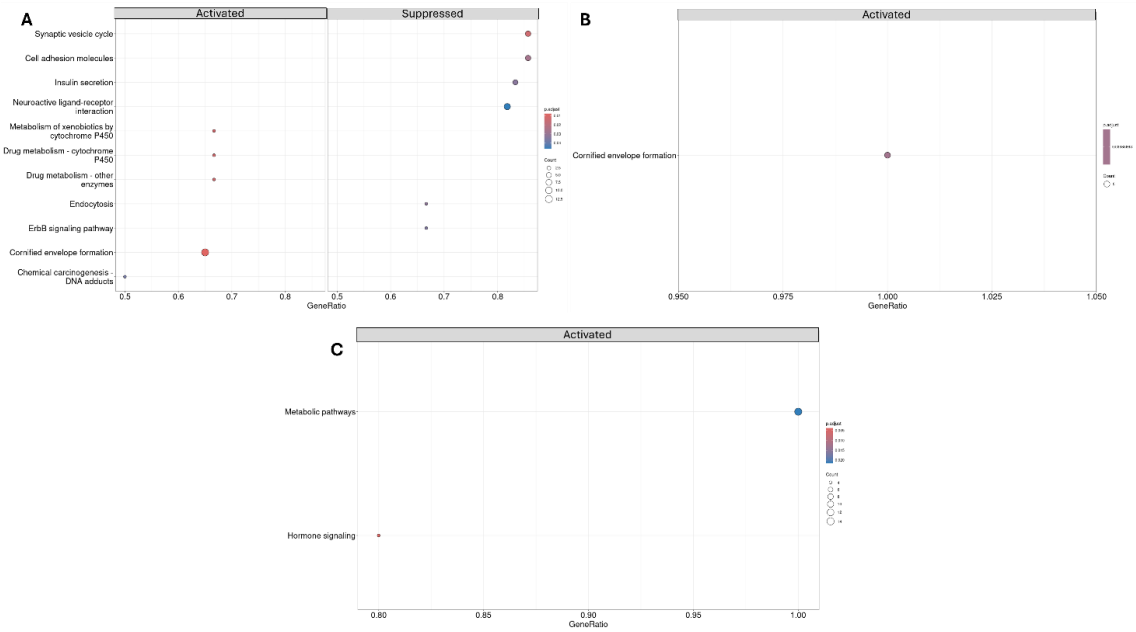

Figure 13. Differentially enriched pathways (Kegg, Kyoto Encyclopedia of Genes and Genomes) in the basal subtype. A) Pathways upregulated and downregulated in Black patients compared to Asian patients B) Pathways upregulated and downregulated in White patients compared to Asian patients C) Pathways upregulated and downregulated in White patients compared to Black patients

Detailed analysis further showed that Black patients present upregulation of glutathione S-transferase (GST), which is involved in metabolism of aromatic hydrocarbons (Figure 14).

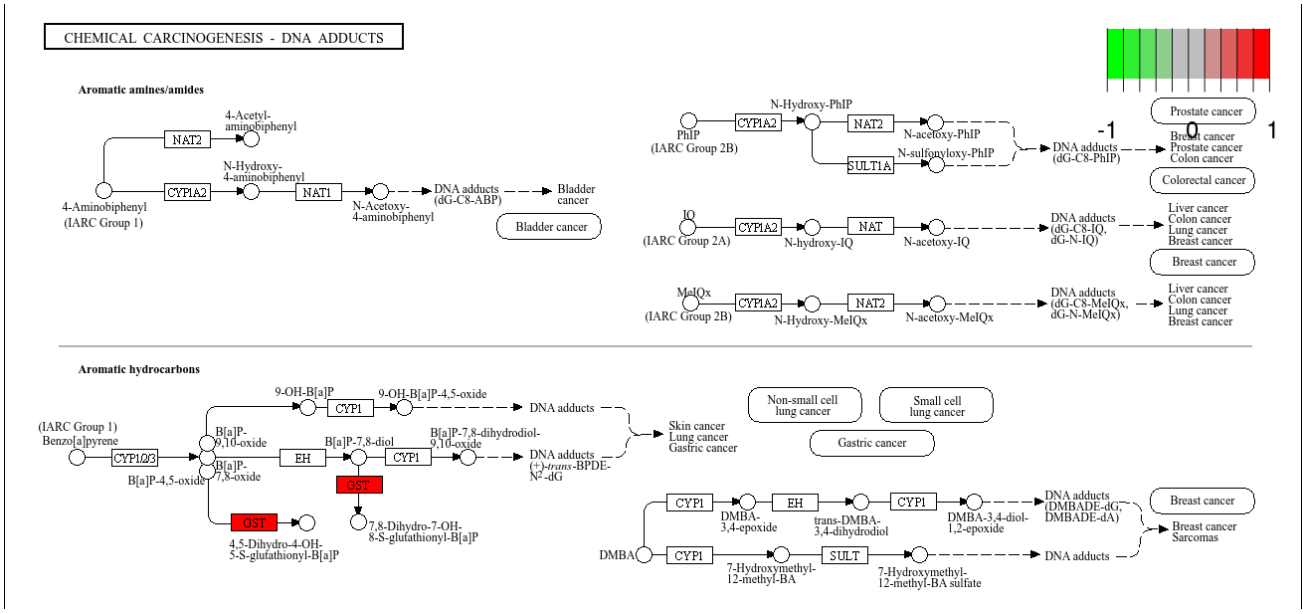

Figure 14. Pathview of the chemical carcinogenesis pathway (DNA adducts) with differential expressed genes. Green genes are downregulated, red genes are upregulated.

In the HER2+ subtype (Figure 15), enrichment analysis was not feasible between Asian and Black patients, due to the low number of differentially expressed genes identified. White patients showed an increased expression of t-he biosynthesis of cofactors compared to Asian and Black patients and of several metabolic pathways compared to Black patients only: drug and xenobiotics metabolism by cytochrome P450 and other enzymes, pentose and glucuronate interconversions, porphyrin metabolism. Chemical carcinogenesis was also enriched in this group, both involving DNA adducts and receptor activation. Consistent with the basal subtype findings, Black patients in this cohort also show a downregulation of cell adhesion molecules compared to White patients.

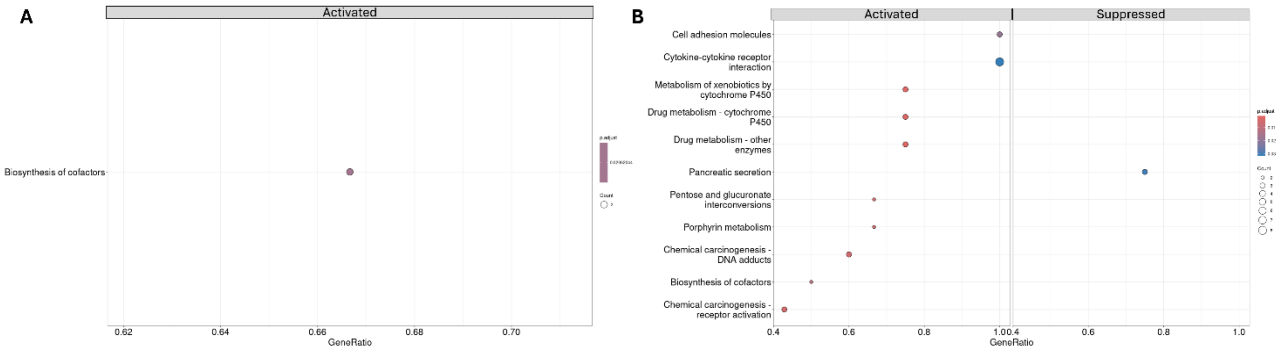

Figure 15. Differentially enriched pathways (Kegg, Kyoto Encyclopedia of Genes and Genomes) in the HER2+ subtype. A) Pathways upregulated and downregulated in White patients compared to Asian patients B) Pathways upregulated and downregulated in White patients compared to Black patients

The luminal A subtype (Figure 16) also lacked sufficient DEGs to perform pathway enrichment between Asian and Black patients. White patients exhibited significant enrichment of metabolic pathways, compared to other ethnicities. This includes retinol metabolism and drug and xenobiotics metabolism (compared to Asian patients), and glycolysis, gluconeogenesis, pyruvate metabolism, retinol metabolism and drug and xenobiotics metabolism (compared to Black patients). Conversely, cholesterol metabolism was more enriched in Black patients. White patients have also showed upregulation of MAPK and PI3K-Akt signalling pathway.

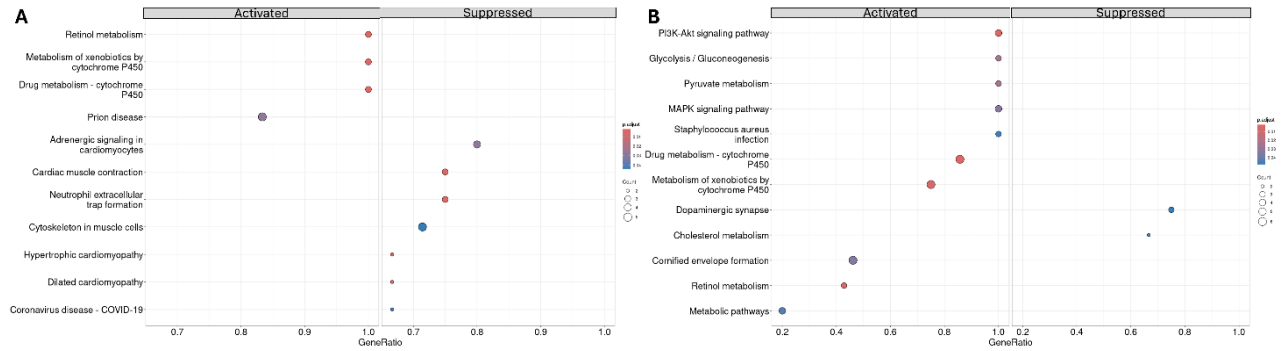

Figure 16. Differentially enriched pathways (Kegg, Kyoto Encyclopedia of Genes and Genomes) in the luminal A subtype. A) Pathways upregulated and downregulated in White patients compared to Asian patients B) Pathways upregulated and

downregulated in White patients compared to Black patients

Finally, in the luminal B subtype (Figure 17), White patients have an enrichment of IL-17 signalling pathway compared to Asian patients, and of chemical carcinogenesis compared to Black patients.

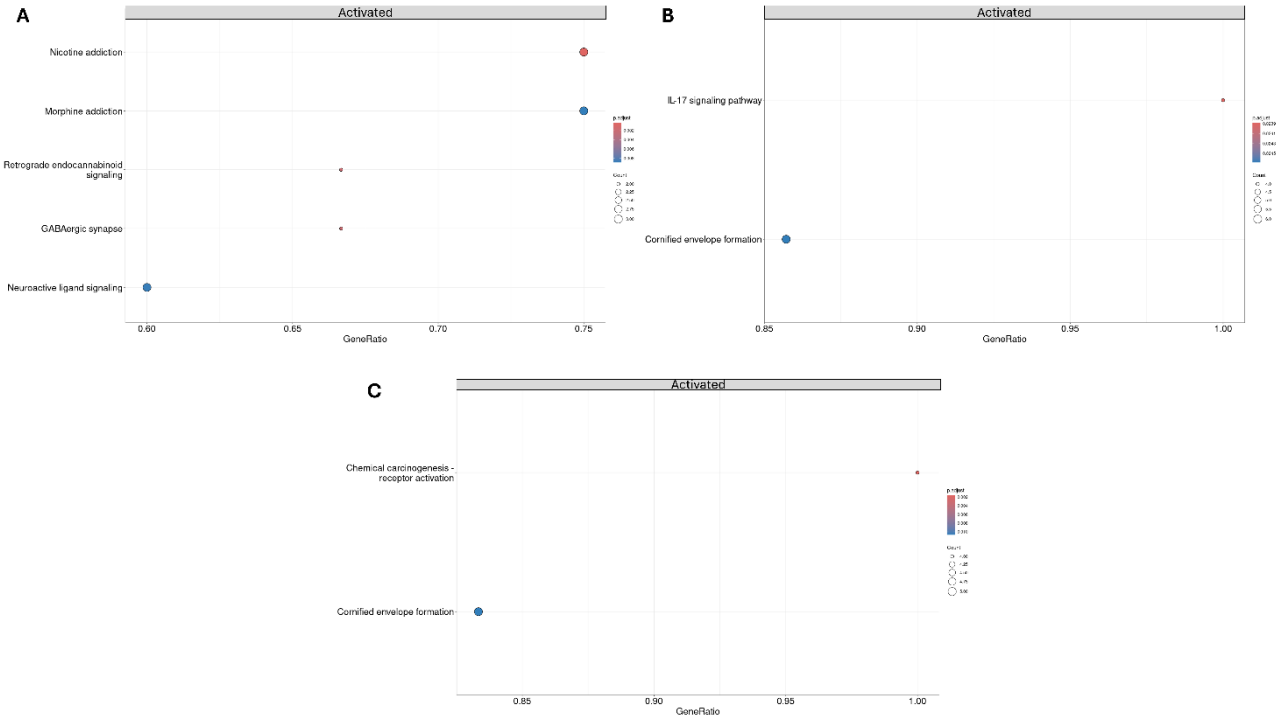

Figure 17. Differentially enriched pathways (Kegg, Kyoto Encyclopedia of Genes and Genomes) in the luminal B subtype. A) Pathways upregulated and downregulated in Black patients compared to Asian patients B) Pathways upregulated and downregulated in White patients compared to Asian patients C) Pathways upregulated and downregulated in White patients compared to Black patients

4.3. Differential gene expression and pathway enrichment across different ethnicities, by TNBC subtype

Moreover, we performed differential expression analysis and enrichment analysis on patients of different ethnicities within the same triple negative breast cancer subtype,

after performing classification through the TNBCtype online tool. However, the limited numbers of non-white patients restricted the comparison: no analysis with Asian patients was possible, and enrichment analysis between Black and White patients was feasible only in the basal-like 1 and mesenchymal subtypes (Figure 18). Out of all basal breast tumours from TCGA, only 10 were classified as BL-1, with 7 being from White patients and 3 from Black ones. Mesenchymal tumours were 17, 14 White and 3 Black. In the BL-1 subtype, White patients present an upregulation of Il-17 signalling pathway and cornified envelope, while in the mesenchymal subtype, they show a suppression of hormone signalling.

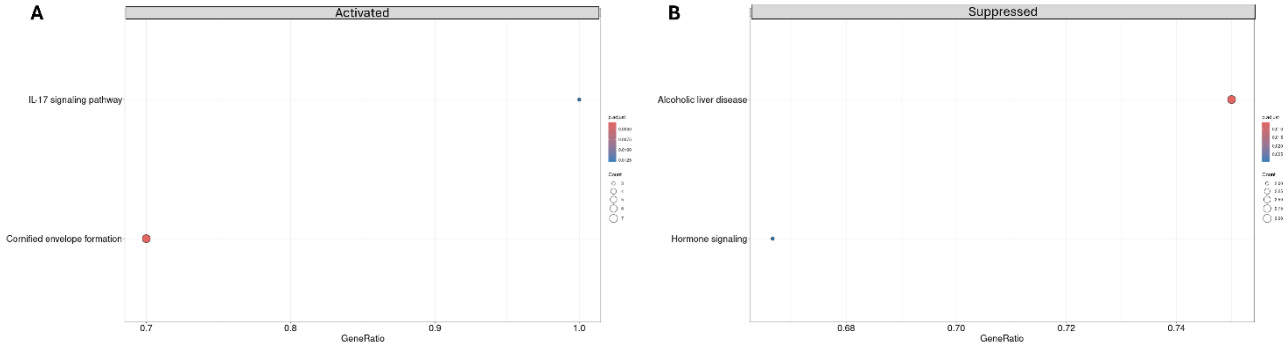

Figure 18. A) Upregulated pathways in White patients compared to Black patients in the basal-like 1 subtype B) Downregulated pathways in White patients compared to Black patients in the mesenchymal subtype

4.4. Differentially expressed genes and impact on survival

A screening on differentially expressed genes was performed, identifying several differentially expressed genes with an impact on survival, either in specific breast cancer subtypes, particular ethnicities or in breast cancer patients overall, highlighting ethnic and subtype specific prognostic markers.

The chemokine encoding genes *CXCL6* and *CXCL8* (Figures 19 and 20) are more expressed in tumours of White patients compared to Black ones, respectively in the basal and HER2+ subtypes. Although not significant, their expression is also higher in Asian patients. In both subtypes, a higher expression level of these genes correlated

with better prognosis.

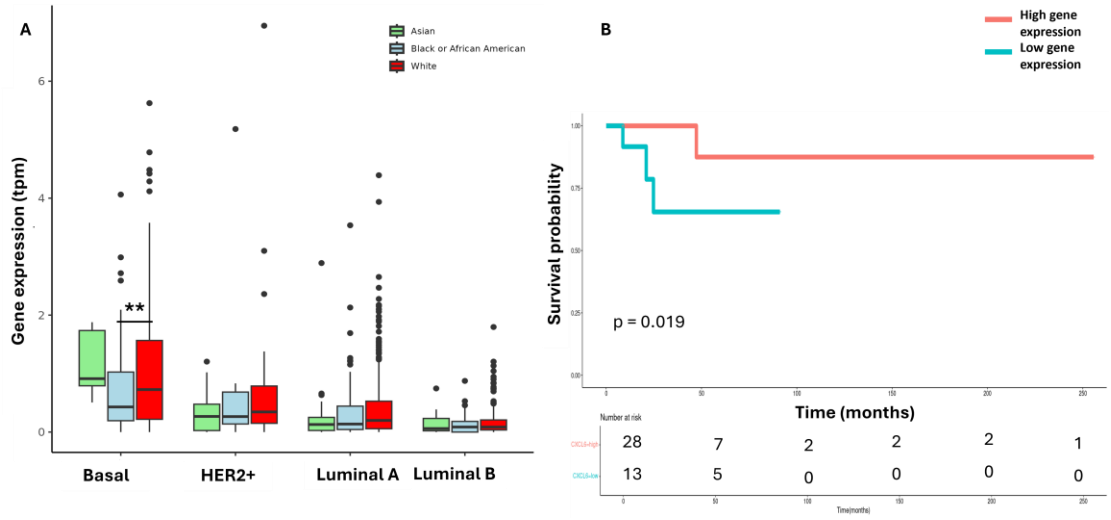

*Figure 19. CXCL6 gene expression (tpm) by race in different subtypes of breast cancer (A), and survival curves according to its expression levels in Black patients with basal breast cancer (B). Asterisks represent statistical significance: *: adjusted p-value ≤ 0.05 **: adjusted p-value ≤ 0.01 ***: adjusted p-value ≤ 0.001*

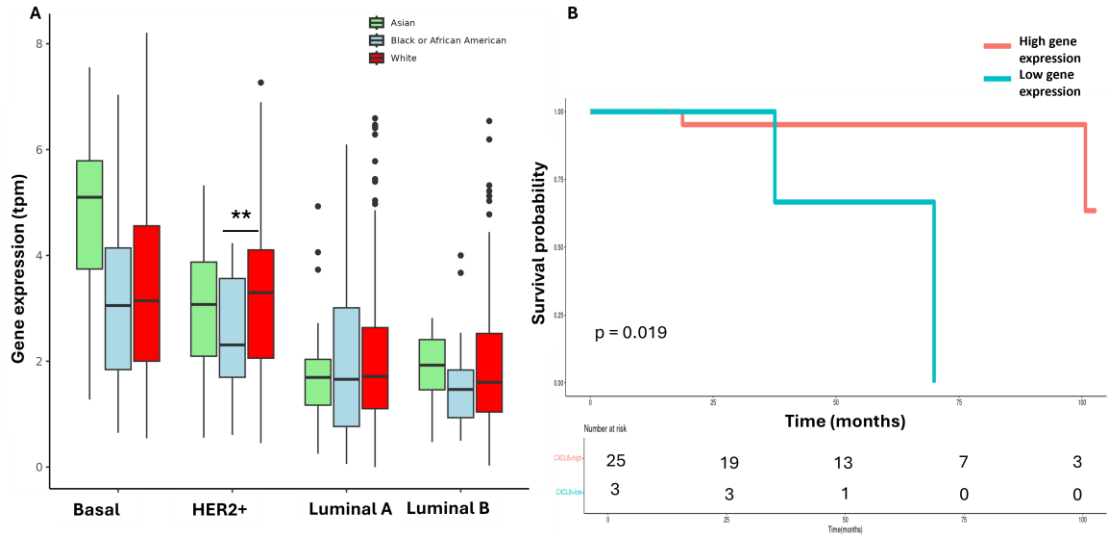

*Figure 20. CXCL8 gene expression (tpm) by race in different subtypes of breast cancer (A), and survival curves according to its expression levels in White patients with HER2+ breast cancer (B), regardless of ethnicity. Asterisks represent statistical significance: *: adjusted p-value ≤ 0.05 **: adjusted p-value ≤ 0.01 ***: adjusted p-value ≤ 0.001*

Glutathione S-transferase Mu1 (*GSTM1*) (Figure 21) is significantly downregulated in basal tumours of Asian patients compared to both Black and White patients. Its high expression, in fact, correlates with a worse overall survival in this subtype. Similarly, *THBS3* (Figure 22) was also downregulated in basal breast cancer patients of Asian ethnicity, with Black patients showing the highest expression and correlating with a worse overall survival.

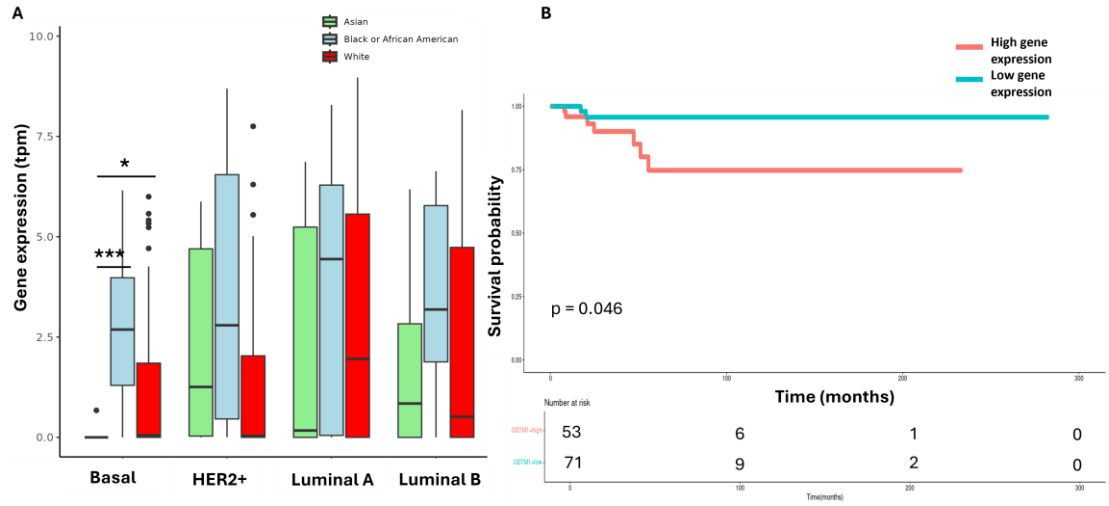

*Figure 21. GSTM1 gene expression (tpm) by race in different subtypes of breast cancer (A), and survival curves according to its expression levels in basal breast cancer (B) regardless of ethnicity. Asterisks represent statistical significance: *: adjusted p-value ≤ 0.05 **: adjusted p-value ≤ 0.01 ***: adjusted p-value ≤ 0.001*

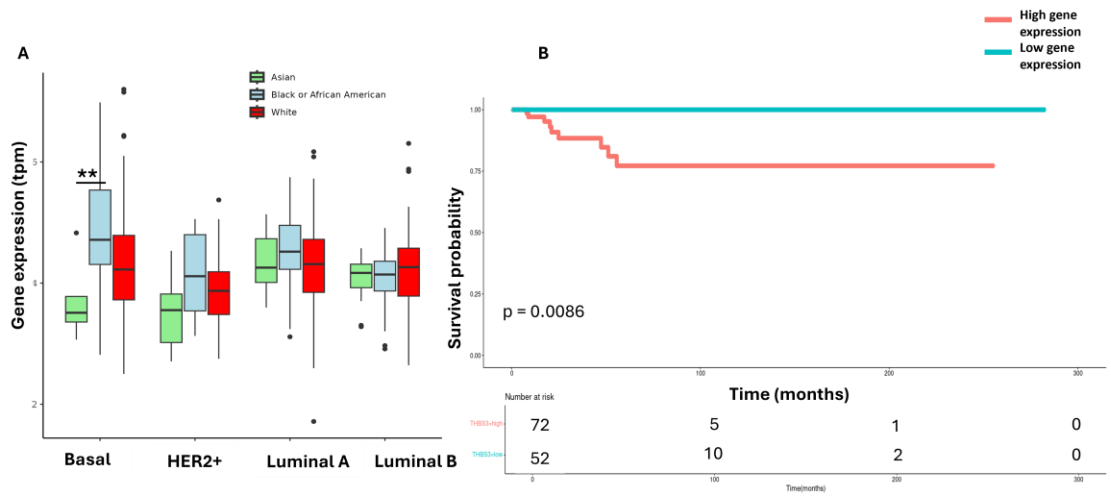

*Figure 22. THBS3 gene expression by race (tpm) in different subtypes of breast cancer (A), and survival curves according to its expression levels in basal breast cancer (B) regardless of ethnicity. Asterisks represent statistical significance: *: adjusted p-value ≤ 0.05 **: adjusted p-value ≤ 0.01 ***: adjusted p-value ≤ 0.001*

Focusing on ion channels and transporter, it emerged that a small set of genes influences survival: *CLCA2*, *SCN1A, ABCA4, KNCK3, and KCNIP3*.

CLCA2 (Figure 23) had the lowest expression in Black patients across all subtypes, while its expression is higher in Asian patients. Furthermore, survival analysis showed that high expression of *CLCA2* correlates with a better prognosis. This is particularly accentuated in Black patients, the only group in which this correlation is statistically significant. In the other ethnicities, instead, the correlation is less striking (White), or absent (Asian). *CLCA2* was also more highly expressed in the HER2+ subtype, compared to the other subtypes.

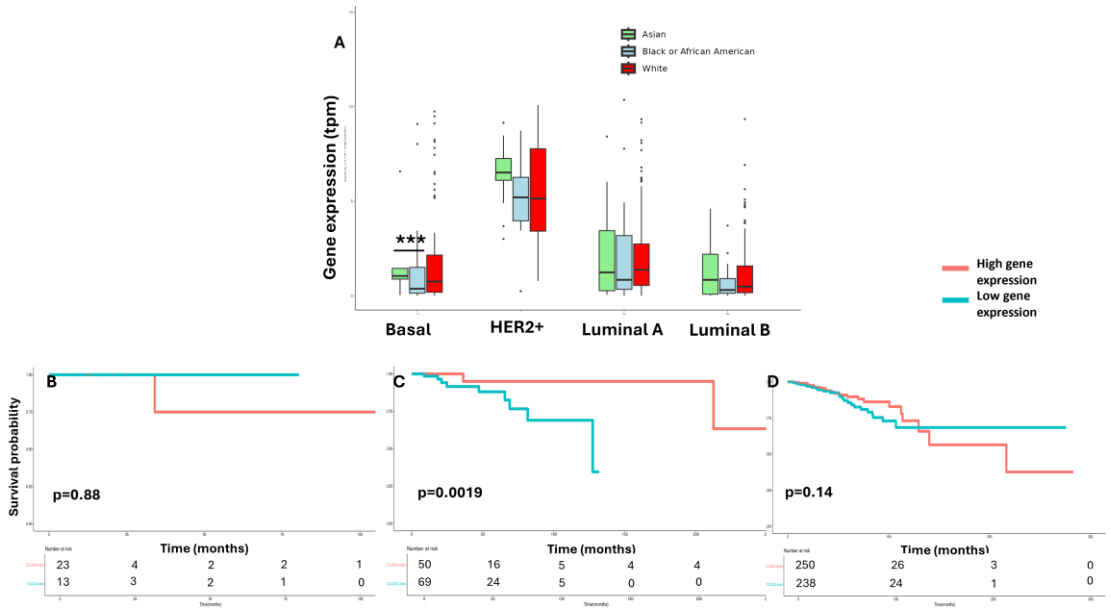

*Figure 23. CLCA2 gene expression (tpm) in different subtypes of breast cancer (A), and survival curves according to its expression levels, in Asian (B), Black (C) and White (D) patients. Asterisks represent statistical significance: *: adjusted p-value ≤ 0.05 **: adjusted p-value ≤ 0.01 ***: adjusted p-value ≤ 0.001*

SCN1A, instead, shows a quite higher expression in Black patients compared to other ethnicity (Figure 24), and the difference is statistically significantly in all subtypes except the HER2+ one. Higher expression of *SCN1A* correlates with a worse overall survival, aligned with Black patients' worse prognosis and more aggressive disease.

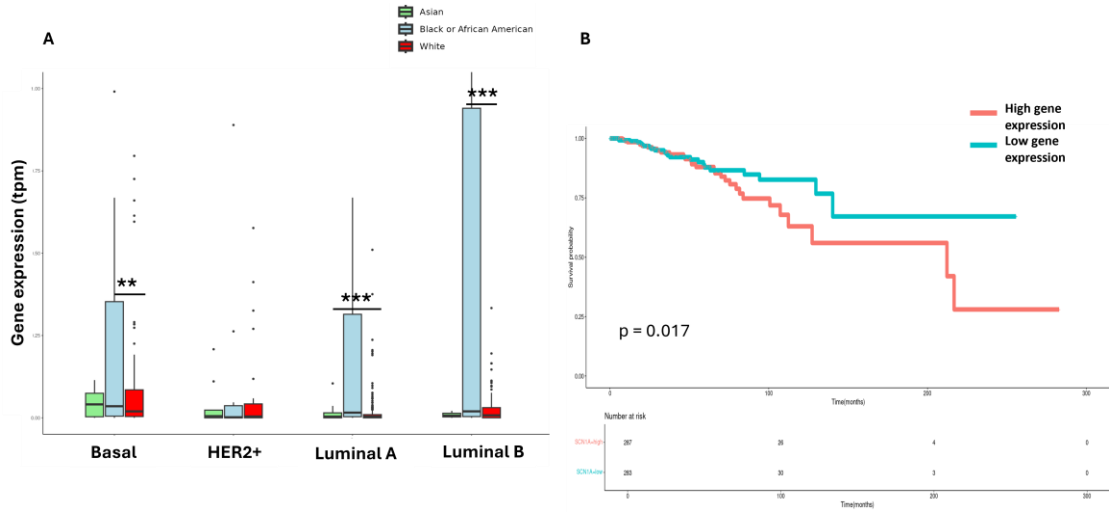

*Figure 24. A) SCN1A expression (tpm) in the different subtypes of breast cancer according to ethnicity. B) Survival curve of breast cancer patients according to the expression of SCN1A. Asterisks represent statistical significance: *: adjusted p-value ≤ 0.05 **: adjusted p-value ≤ 0.01 ***: adjusted p-value ≤ 0.001*

The ABC transporter *ABCA4*, whose high expression correlates with a better overall survival (OS) in luminal B breast cancer, was found to have a lower expression in Black luminal B breast cancer patients (Figure 25).

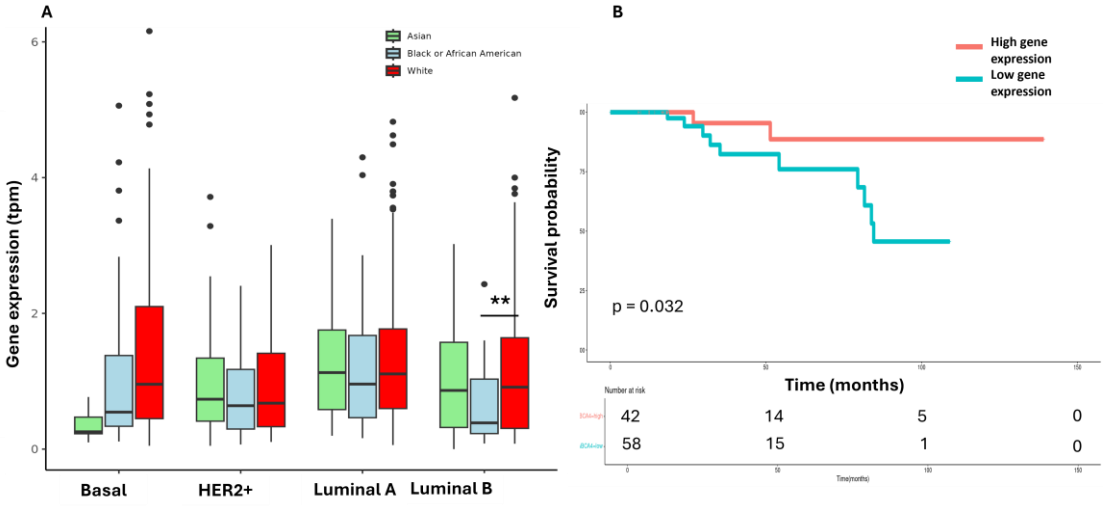

*Figure 25. A) ABCA4 expression (tpm) in the different subtypes of breast cancer according to ethnicity. B) Survival curve of luminal B breast cancer patients according to the expression of ABCA4, regardless of ethnicity. Asterisks represent statistical significance: *: adjusted p-value ≤ 0.05 **: adjusted p-value ≤ 0.01 ***: adjusted p-value ≤ 0.001*

The potassium channel gene *KCNK3* (Figure 26) was more expressed in Asian patients of luminal B subtype and correlates with a worse overall survival.

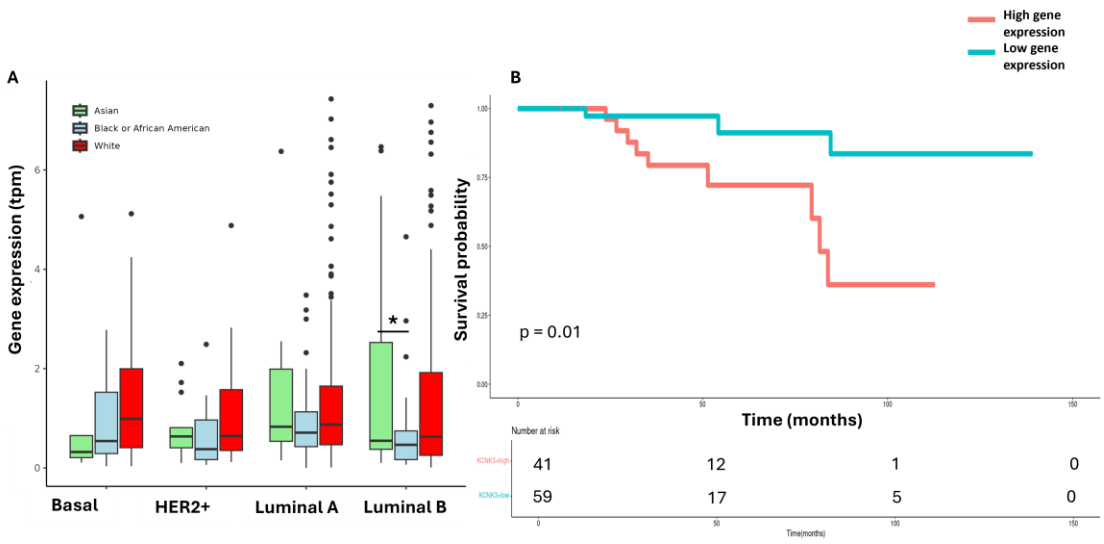

*Figure 26. A) KCNK3 expression (tpm) in the different subtypes of breast cancer according to ethnicity. B) Survival curve of luminal B breast cancer patients according to the expression of KCNK3 regardless of ethnicity. Asterisks represent statistical significance: *: adjusted p-value ≤ 0.05 **: adjusted p-value ≤ 0.01 ***: adjusted p-value ≤ 0.001*

Conversely, the potassium channel interacting protein *KCNIP3* showed a higher expression in Black basal breast cancer patients and its expression correlated with a worse overall survival in breast cancer without being subtype-specific (Figure 27).

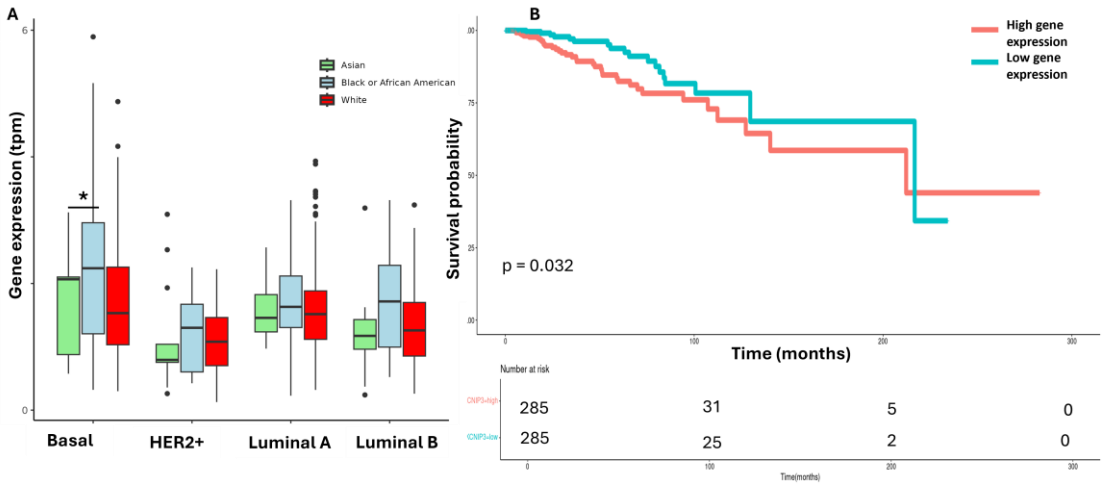

*Figure 27. A) KCNIP3 expression (tpm) in the different subtypes of breast cancer according to ethnicity. B) Survival curve of breast cancer patients (all subtypes and ethnicities) according to the expression of KCNIP3. Asterisks represent statistical significance: *: adjusted p-value ≤ 0.05 **: adjusted p-value ≤ 0.01 ***: adjusted p-value ≤ 0.001*

Non-protein coding genes were also found to be differentially expressed across ethnic groups. Among those are lncRNAs such as *LINC00624* and *LINC02579* and transcripts with no known role: *AC092979.1, AL031848.1* and *AL590666.3*.

LINC00624 was upregulated in HER2+ breast cancer of Black patients and its high expression correlates with worse overall survival in breast cancer, in general (Figure 28).

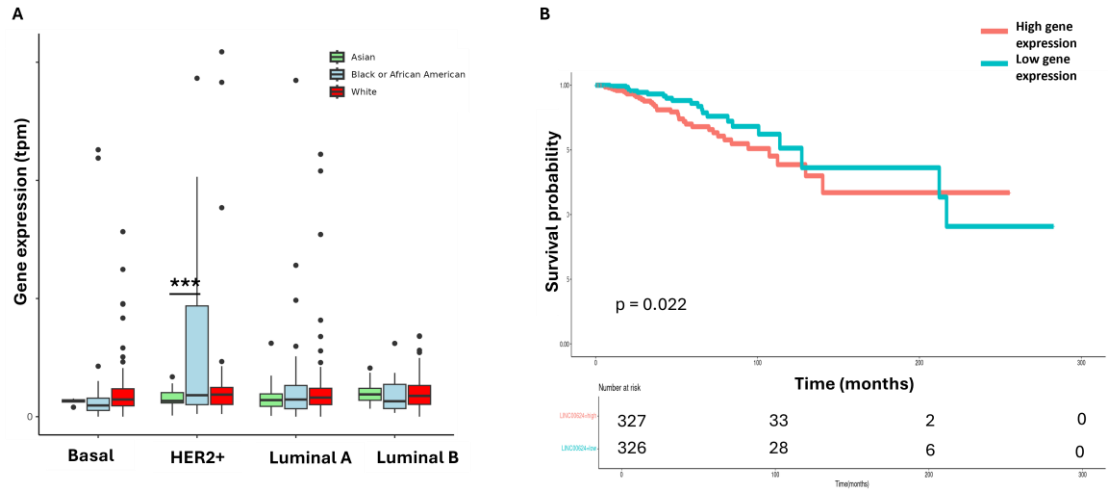

*Figure 28. LINC00624 gene expression (tpm) by race in different subtypes of breast cancer (A), and survival curves according to its expression levels in breast cancer (all subtypes and ethnicities) (B). Asterisks represent statistical significance: *: adjusted p-value ≤ 0.05 **: adjusted p-value ≤ 0.01 ***: adjusted p-value ≤ 0.001*

LINC02579 high expression, instead, correlates with a better prognosis in breast cancer with no subtype specificity, and is upregulated in Asian patients in the luminal B subtype (Figure 29).

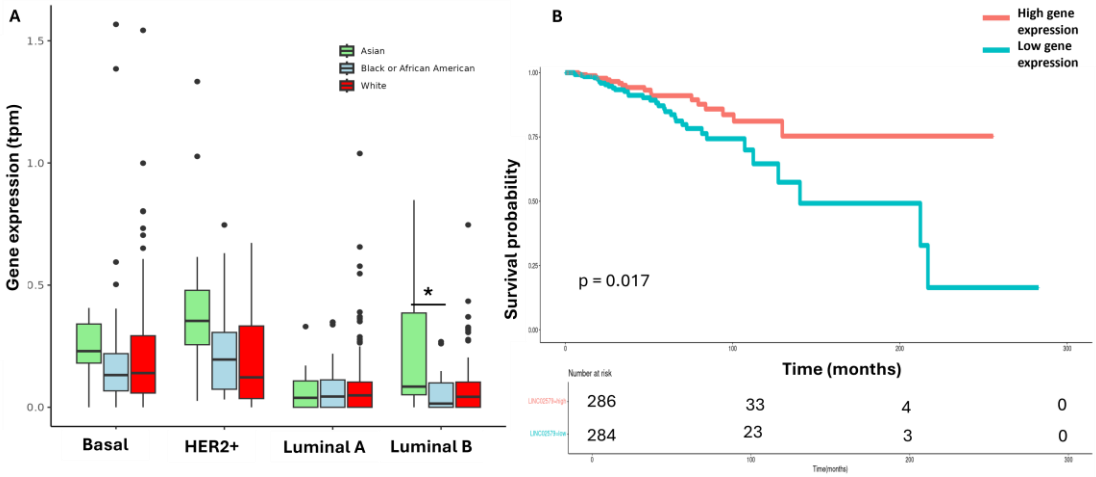

*Figure 29. LINC02579 gene expression (tpm) by race in different subtypes of breast cancer (A), and survival curves according to its expression levels in breast cancer (all subtypes and ethnicities) (B). Asterisks represent statistical significance: *: adjusted p-value ≤ 0.05 **: adjusted p-value ≤ 0.01 ***: adjusted p-value ≤ 0.001*

AC092979.1, which appears more characteristic of luminal subtypes compared to the others, is upregulated in White patients, and correlates with a worse prognosis in the luminal A subtype (Figure 30).

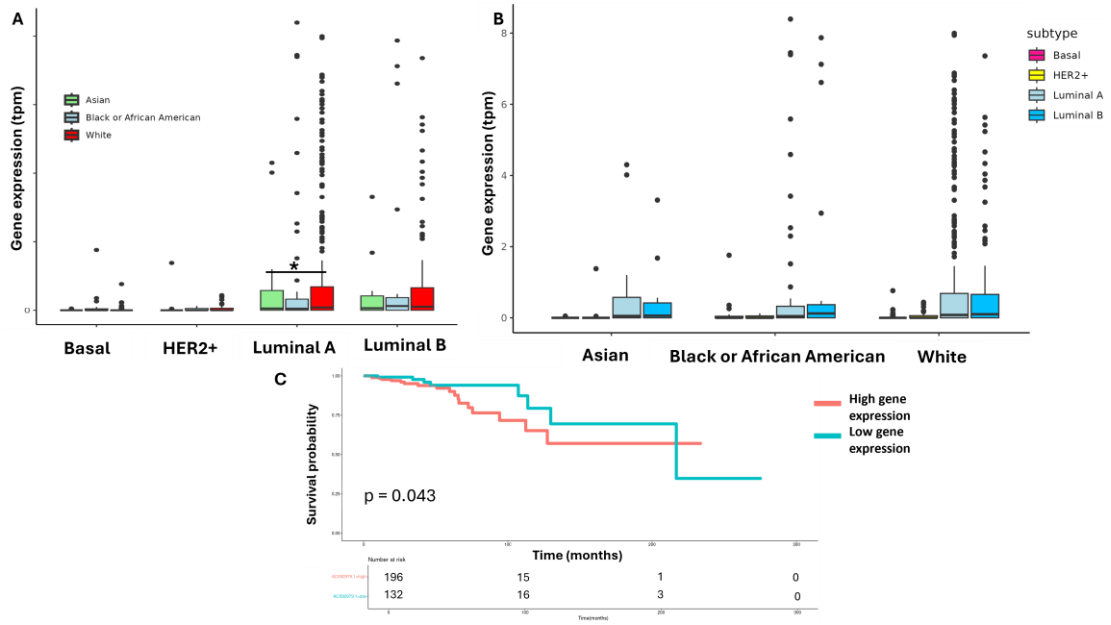

*Figure 30. AC092979.1 gene expression (tpm) by race different subtypes of breast cancer (A), gene expression by subtype in different ethnicities (B) and survival curves according to its expression levels in the luminal A subtype regardless of ethnicity (C). Asterisks represent statistical significance: *: adjusted p-value ≤ 0.05 **: adjusted p-value ≤ 0.01 ***: adjusted p-value ≤ 0.001*

AL031848.1 (Figure 31) correlates with a worse prognosis in the basal subtype and is upregulated in Black basal breast cancer patients.

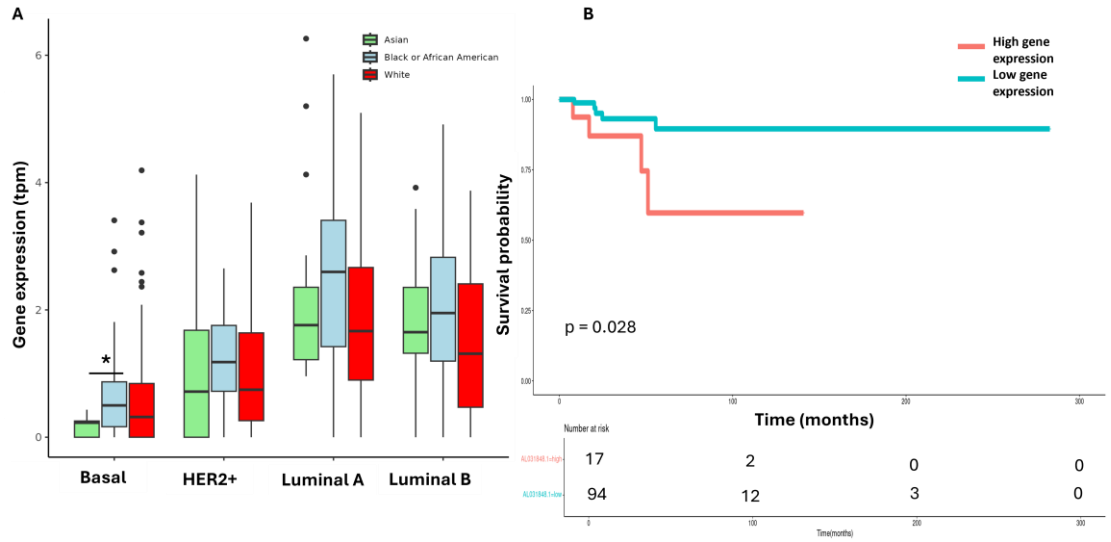

*Figure 31. AL031848.1 gene expression (tpm) by race in different subtypes of breast cancer (A), and survival curves according to its expression levels in the basal subtype (B, regardless of ethnicity. Asterisks represent statistical significance: *: adjusted p-value ≤ 0.05 **: adjusted p-value ≤ 0.01 ***: adjusted p-value ≤ 0.001*

AL590666.3 also correlates with poor prognosis, but in the luminal A subtype and is upregulated in Black patients in both basal and luminal A subtype (figure 32).

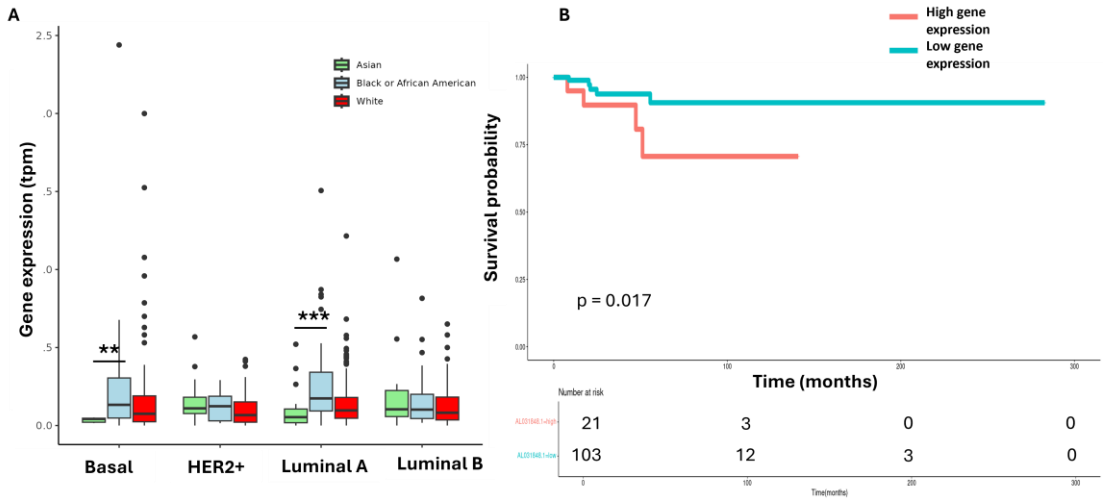

*Figure 32. AL590666.3 gene expression (tpm) by race in different subtypes of breast cancer (A), and survival curves according to its expression levels in the luminal A subtype, regardless of ethnicity (B). Asterisks represent statistical significance: *: adjusted p-value ≤ 0.05 **: adjusted p-value ≤ 0.01 ***: adjusted p-value ≤ 0.001*

4.5. Mutational profile across subtypes

Different subtypes demonstrated to have different incidence of mutations as well.

In the basal subtype, *FAT3* was mutated in 43% of Asian patients, but only in 4% and

8% of Black and White patients, respectively. *FLG2* was mutated with significantly higher frequency in Black patients (12% versus 0% in Asian patients and 2% in White patients), and a similar mutation rate was found for *CDH26* (10% in Black patients, 0% in Asian and White patients). On the other hand, *CREBBP* was found to be mutated only in White patients (10%). It is important to note that some genes, such as *MUC16, TTN, SYNE1* and *USH2A*, are reported by the software as FLAG genes, intended as large genes which often present mutations, but are not likely to harbour driver mutations relevant to cancer development or biology (Figure 33A).

In general, the WNT and Hippo pathways are more frequently mutated in Asian patients compared to other ethnicities, while the opposite is true for RTK-RAS. Notch, instead, has a similar mutation rate between Asian and Black patients, while it is more often mutated in White patients (Figure 33B). Both Asian and Black patients present APOBEC cytidine deaminase as their most common mutational signature, while this is not present in White patients, where defects in DNA repair are more likely to originate mutations.

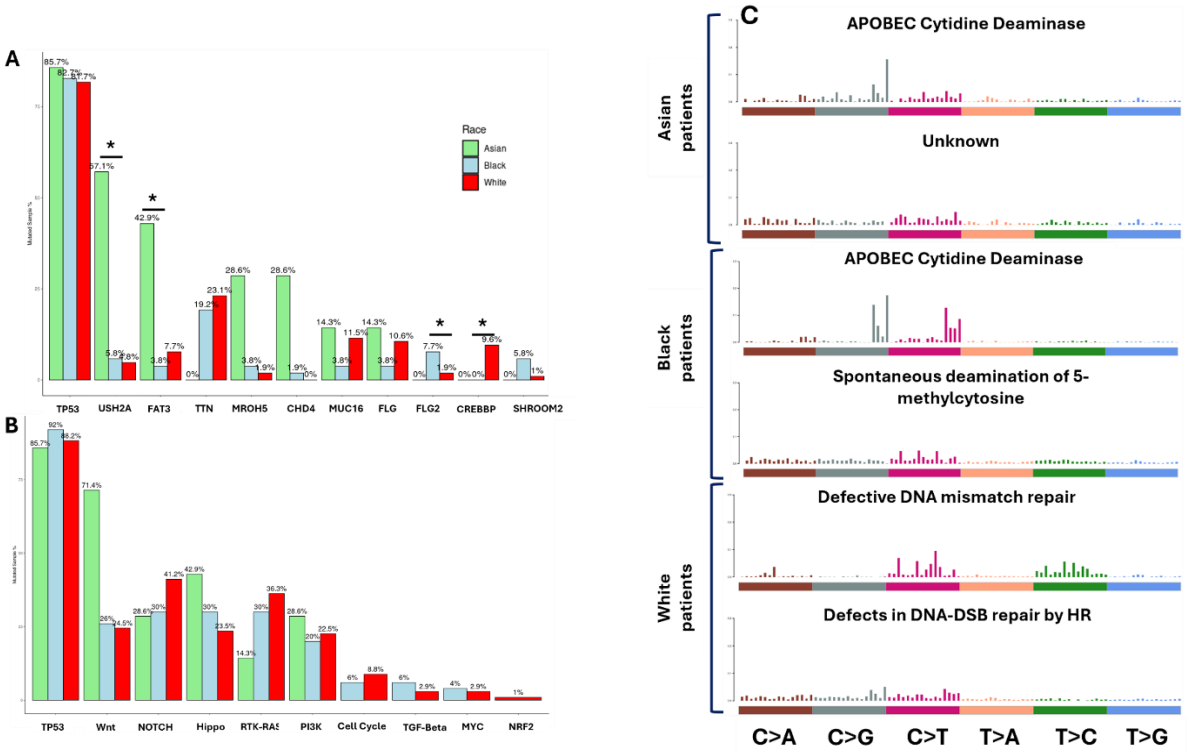

Figure 33. A) Top mutated genes in basal breast cancer in different ethnicities B) Top pathways affected by mutations in basal breast cancer in different ethnicities C) Top

mutational signatures in basal breast cancer in different ethnicities

In HER2+ subtype, there was no significant difference in mutated genes, however *KMT2D* mutations were more common in Asian patients (20%) versus White patients (9%) while no mutations were detected in Black patients.

Black patients had higher mutation rate of *BLM, COL12A1* and *KMT2C*, all present in 27% of patients. *BLM* mutations were found only in 3% of White patients and in no Asian patient, while *COL12A1* and *KMT2C* have a similar mutation rate in Asian and White patients (7% and 6% respectively) (Figure 34).

Hippo and TGF-beta pathways are less mutated in Asian patients (20% and 0%) compared to Black (45% and 9%) and White patients (31% and 11%). The Notch pathway is more frequently mutated compared to black patients, and the PI3KCA pathway is mutated at higher frequency compared to White patients. WNT is part of the pathway with the highest variability in mutation rate, with 13% in Asian patients, 36% in Black patients and 20% in White patients. The most common mutational signatures are similar between ethnicities.

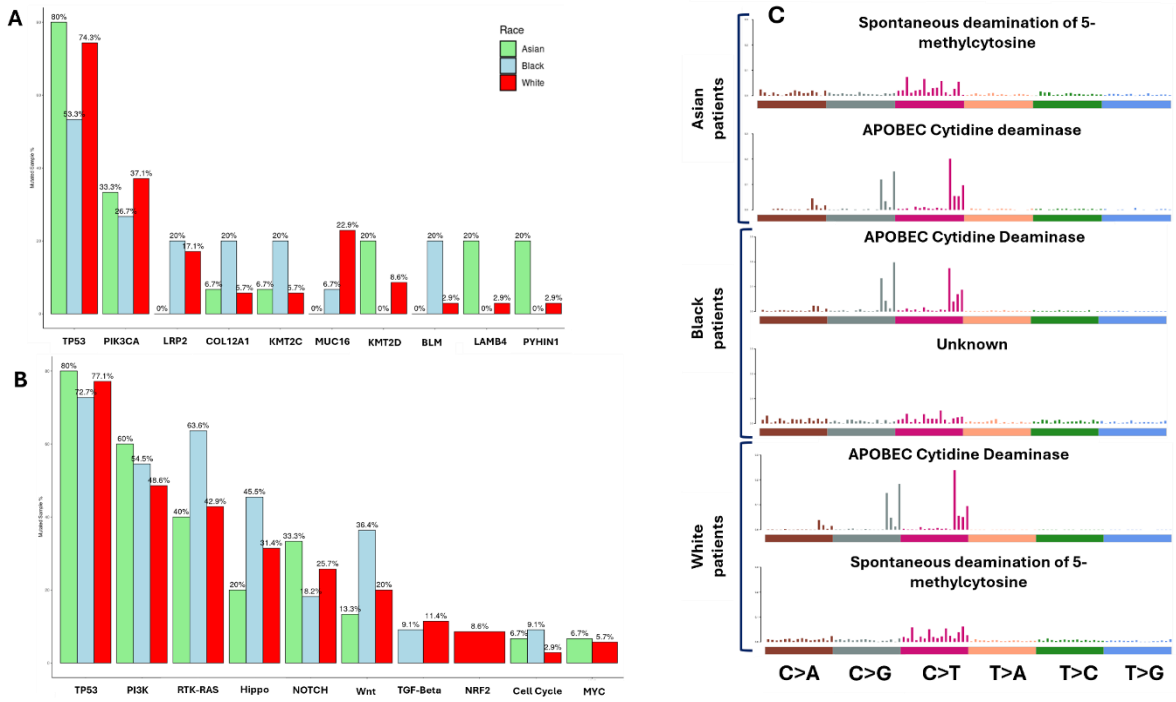

Figure 34. A) Top mutated genes in HER2+ breast cancer in different ethnicities B) Top pathways affected by mutations in HER2+ breast cancer in different ethnicities C) Top mutational signatures in HER2+ breast cancer in different ethnicities

In the luminal A subtype (Figure 35), Asian and White patients have a similar *PIK3CA* mutation rate (53%), while it is lower in Black patients (37%). *CDH1* and *GATA3* are more frequently mutated in Asian patients (24%) compared to Black (10% and 0%) and White patients (14%).

RTK-RAS pathway has the highest mutation rate in Asian patients (32% vs 17%), while Notch and Hippo have the lowest mutation rate (6% and 0% vs 23-25% and 11-16%). Defects in POLE polymerase are the most common cause of mutations in White patients but not in the other groups, where spontaneous deamination of 5-methylcytosine is more common (Figure 35C).

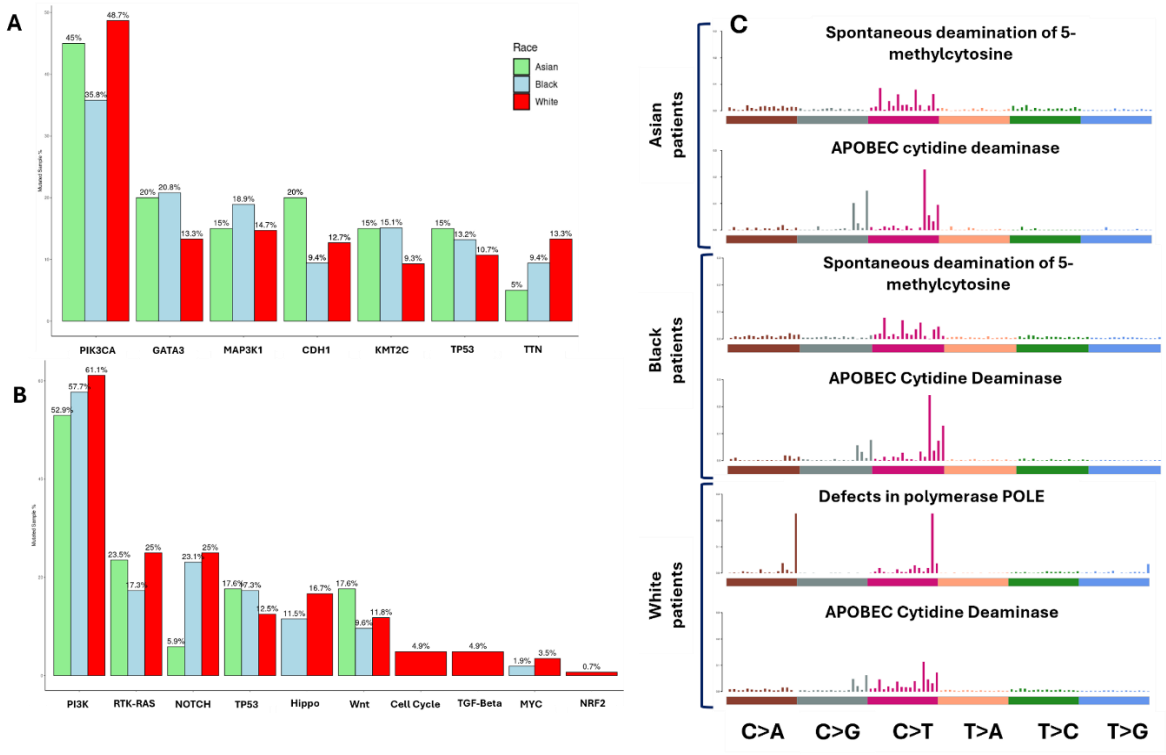

Figure 35. A) Top mutated genes in luminal A breast cancer in different ethnicities B) Top pathways affected by mutations in luminal A breast cancer in different ethnicities C) Top mutational signatures in luminal A breast cancer in different ethnicities

In the luminal B subtype (Figure 36), *GATA3* is significantly more frequently mutated in Asian patients compared to other ethnicities (47% versus 21% in Black patients and 18% in White patients); *ARID1A* is more commonly mutated in Asian patients as well (20% vs 0% vs 4%).

Genes involved in the PI3K and Wnt pathways have the highest mutation rate in Asian patients compared to other ethnicities. Black patients have the lowest mutation rate in both PI3KCA and TP53, and the highest in RTK-RAS. Notch, instead, has the lowest mutation rate in Asian patients. APOBEC cytidine deaminase is the most common mutational signature in all ethnic groups, with spontaneous deamination of 5-methylcytosine being second in all but White patients, who instead present mutations caused by defects in the DNA mismatch repair pathway (Figure 36).

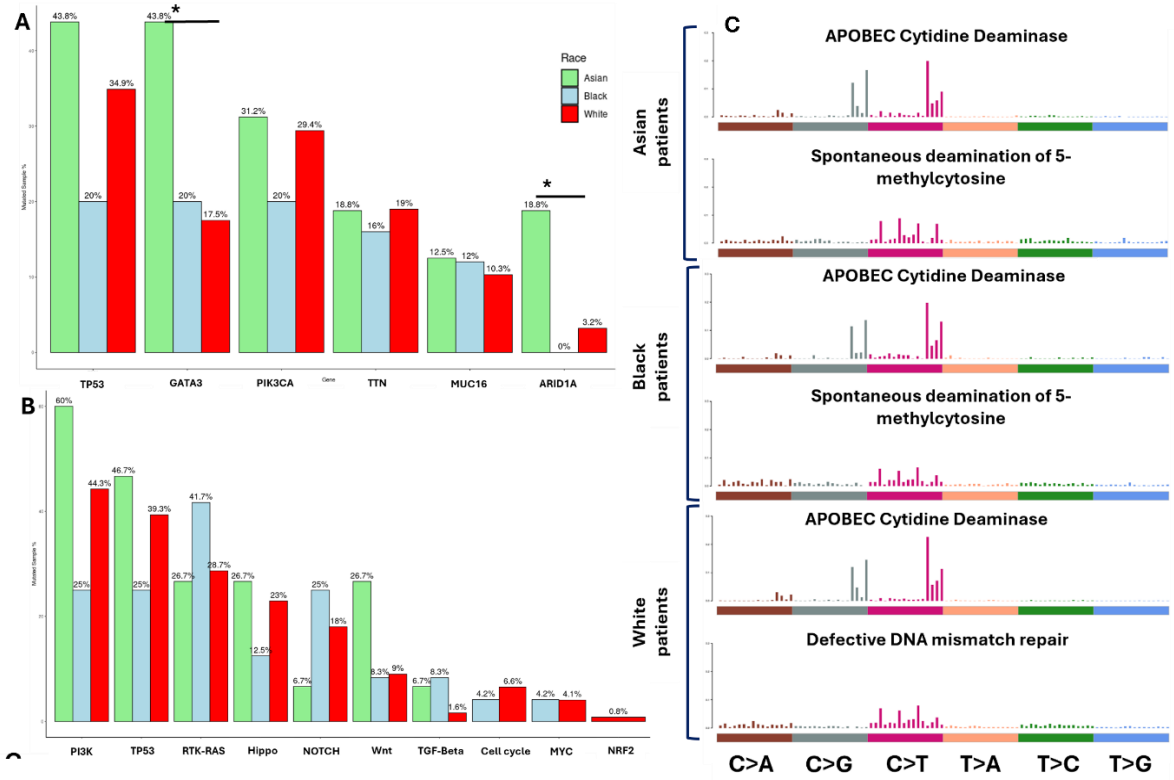

Figure 36. A) Top mutated genes in luminal B breast cancer in different ethnicities B) Top pathways affected by mutations in luminal B breast cancer in different ethnicities C) Top mutational signatures in luminal B breast cancer in different ethnicities

4.6. Differential gene expression across breast cancer subtypes

We performed an analysis on TCGA samples to determine differential gene expression between breast cancer subtypes, comparing patients with triple negative breast cancer with patients with breast cancer belonging to other subtypes: HER2 +, Luminal A and Luminal B.

Our comparisons included 150 TNBC patients versus 78 HER2+ breast cancer, 150 Luminal A breast cancer and 150 Luminal B breast cancer, selected from patients in Table 1. The reduced number of patients was chosen due to computational limitations.

MAplots showing the distribution of the differentially expressed genes (DEGs) is shown in Figure 37. The plots display gene expression levels against Log2 fold changes between different basal breast cancer and other subtypes. It can be observed that across all comparisons a substantial number of genes are differentially expressed (blue dots) while a smaller number of genes show no deregulation.

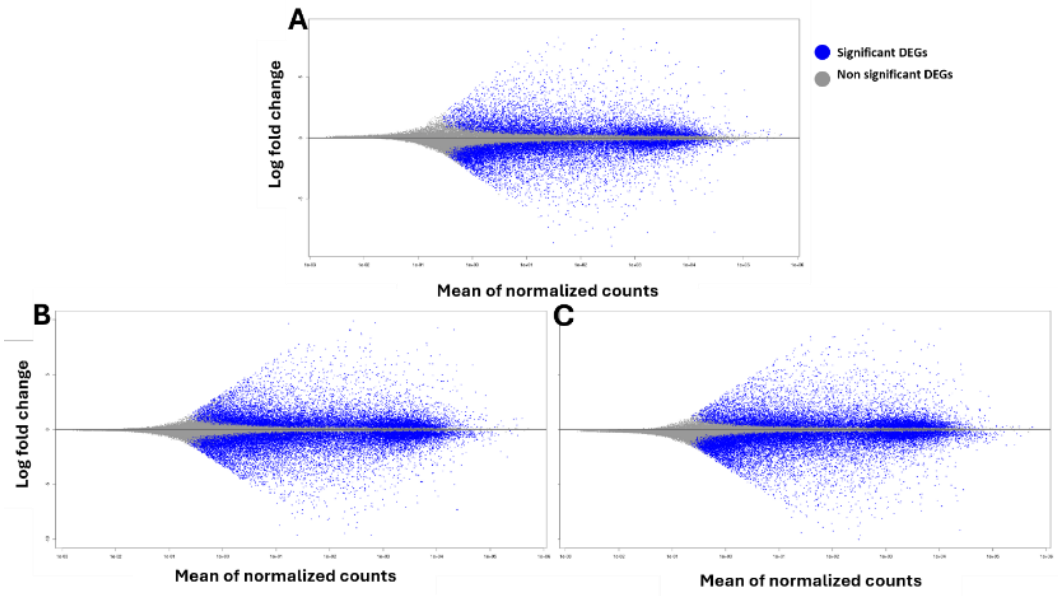

Figure 37. Distribution of differentially expressed genes in basal breast cancer compared to a) HER2+ breast cancer b) luminal A breast cancer c) luminal B breast cancer. Blue dots: significant DEGs; Grey dots: non-significant DEGs

In Figure 38 are presented Volcano plots displaying Log2FoldChange between different basal breast cancer and other subtypes and their significance, with orange dots representing upregulated genes and blue dots downregulated ones.

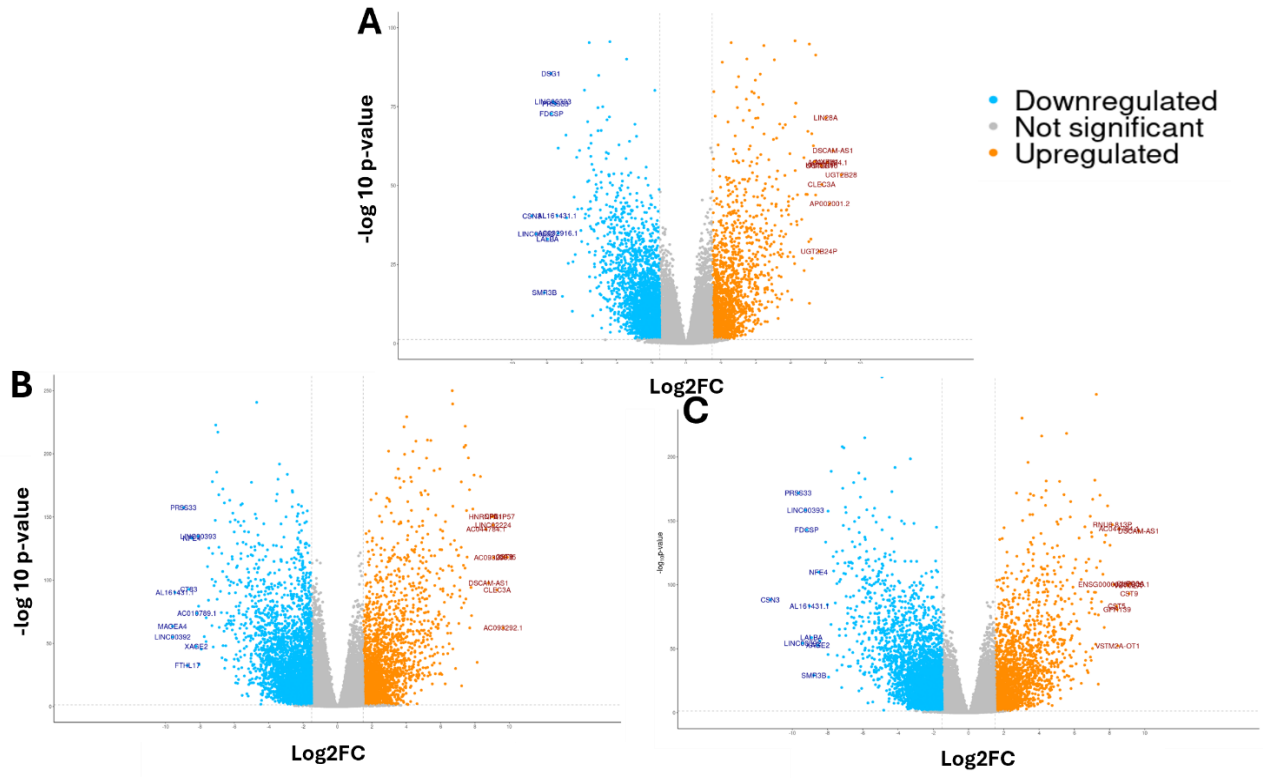

Figure 38. Volcano plot showing DEGs in patients with basal breast cancer compared to compared to a) HER2+ breast cancer b) luminal A breast cancer c) luminal B breast cancer. Labelled genes are the ones with the highest fold change, both upregulated (orange dots) and downregulated (blue dots)

From the volcano plots it can be observed that in every comparison a high number of genes turned out to be dysregulated, either up- (orange dots) and down-regulated (blue dots) supporting the hypothesis that differences in the genetic landscape are present in the different molecular subtypes of breast cancer.

In the following tables (Table 2-4) top 10 DEGs (downregulated and upregulated in TNBC compared respectively to HER2+ breast cancer, luminal A breast cancer and luminal B breast cancer) are listed.

GeneName	Log2FC	AdjPValue	GeneName	Log2FC	PValue	AdjPValue
CLEC3A	-7.345	2.40E-35	SERPINB4	3.971	6.72E-16	1.06E-13
RF00017	-7.187	3.20E-11	SERPINB3	3.714	8.47E-11	5.29E-09

ABCC12	-6.954	5.89E-42	CA6	3.501	6.15E-17	1.13E-14
RNU1-88P	-6.862	1.00E-12	AL035425.3	3.462	1.77E-15	2.57E-13
SULT1C3	-6.777	1.22E-09	CERS3	3.346	1.73E-23	8.46E-21
AP002001.2	-6.634	6.90E-15	SPRR2G	3.333	1.21E-07	3.56E-06
DSCAM-AS1	-6.595	1.05E-21	AL161431.1	3.293	8.14E-08	2.50E-06
AP002001.1	-6.324	3.00E-19	CMTM5	3.198	9.85E-13	9.00E-11
NXPH1	-6.319	1.35E-23	SERPINB7	3.194	5.76E-20	1.72E-17
PNMT	-6.167	4.04E-56	LALBA	3.186	1.11E-06	2.44E-05

Table 2. Top 10 downregulated and upregulated genes in basal breast cancer compared to HER2+ breast cancer.

GeneName	Log2FC	PValue	AdjPValue	GeneName	Log2FC	PValue	AdjPValue
CST5	-9.740	1.54E-115	3.26E-113	MAGEA4	9.582	9.01E-64	3.82E-62
AC093292.1	-9.661	1.44E-63	6.10E-62	LINC00392	9.578	3.61E-55	1.14E-53
CLEC3A	-9.295	1.66E-93	1.76E-91	AL161431.1	9.186	1.12E-85	9.70E-84
AC093297.1	-9.017	5.37E-117	1.18E-114	PRSS33	8.937	6.65E-156	4.66E-153
CST9	-8.981	3.09E-96	3.64E-94	FTHL17	8.713	6.35E-33	8.95E-32
LINC02224	-8.831	2.10E-136	8.34E-134	CT83	8.627	3.51E-92	3.57E-90
CPB1	-8.645	1.63E-132	6.09E-130	NFE4	8.459	1.18E-132	4.46E-130
CARTPT	-8.149	5.08E-36	8.13E-35	XAGE2	8.202	7.19E-48	1.77E-46

ENSG00000286208.1	-8.014	2.72E-161	2.35E-158	AC010789.1	8.155	1.19E-73	7.01E-72
AC008663.1	-7.957	1.33E-186	2.67E-183	LINC00393	8.075	2.12E-133	8.07E-131

Table 3. Top 10 downregulated and upregulated genes in basal breast cancer compared to luminal A breast cancer

GeneName	Log2FC	PValue	AdjPValue	GeneName	Log2FC	PValue	AdjPValue
DSCAM-AS1	-9.656	4.54E-143	2.60E-140	CSN3	11.28	3.02E-90	3.97E-88
CLEC3A	-9.185	6.30E-102	1.17E-99	PRSS33	9.799	5.11E-179	9.02E-176
CST9	-9.111	7.78E-94	1.11E-91	LINC00392	9.419	3.76E-55	1.66E-53
VSTM2A	-9.063	9.54E-101	1.68E-98	FDCSP	9.385	3.70E-152	2.59E-149
AC044784.1	-8.506	6.86E-143	3.88E-140	AL161431.1	9.285	2.91E-90	3.84E-88
VSTM2A-OT1	-8.469	5.21E-53	2.14E-51	LINC00393	9.134	9.12E-153	6.71E-150
GPR139	-8.420	5.06E-81	4.96E-79	LALBA	8.906	1.68E-59	8.67E-58

CST5	-8.398	6.14E-84	6.49E-82	SMR3B	8.780	5.27E-30	8.31E-29
ENSG00000287900.1	-8.247	3.72E-101	6.65E-99	PRR27	8.622	4.63E-60	2.42E-58
RNU6-813P	-8.153	1.19E-147	7.31E-145	XAGE2	8.565	4.56E-53	1.88E-51

Table 4. Top 10 downregulated and upregulated genes in basal breast cancer compared to luminal B breast cancer

To identify genes consistently dysregulated in TNBC, we intersected the sets of upregulated and downregulated genes from each comparison. The top 10 intersected DEGs are reported in Table 5, along with their mean Log2FoldChange values.

GeneName	Mean Log2FC	GeneName	Mean Log2FC
LINC00392	-9.410	CLEC3A	9.179
AL161431.1	-9.018	CST9	9.111
PRSS33	-8.956	DSCAM-AS1	8.791
CSN3	-8.842	AC044784.1	8.540
NFE4	-8.473	CST5	8.400
XAGE2	-8.221	ENSG00000286208.1	7.888
SMR3B	-8.119	VSTM2A	7.802
LINC00393	-8.094	TFF1	7.782
ZFP42	-8.010	CPB1	7.622
LALBA	-7.936	LINC02224	7.539

Table 5. Top 10 common downregulated and upregulated genes in basal breast cancer compared to other subtypes.

4.7. Pathway analysis

We then performed pathway enrichment analyses, using KEGG ontology, on the differentially expressed genes, whose results are reported in Figure 39.

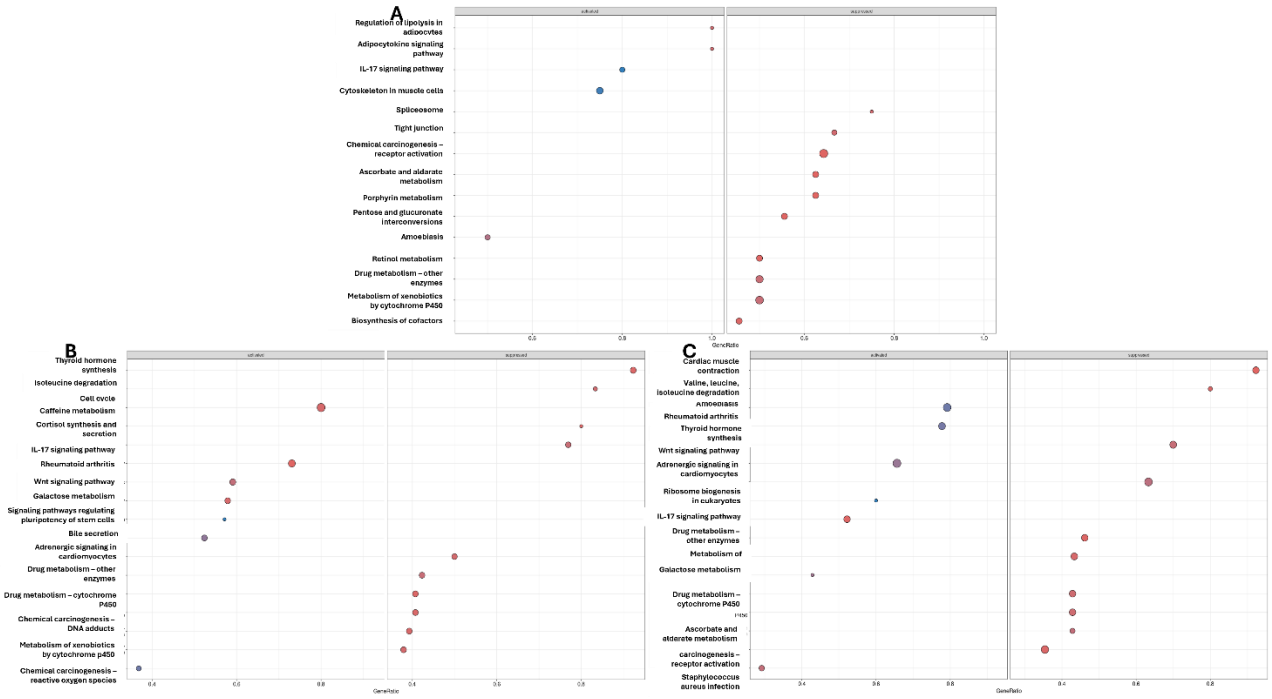

Figure 39. Dot plots of differentially enriched pathways (KEGG) in basal breast cancer compared to HER2+ (A), luminal A (B) and luminal B breast cancer (C)

Regarding KEGG pathways, the basal subtype is enriched in genes involved in IL-17 signalling pathway, while it has a downregulation in tight junctions, chemical carcinogenesis by receptor activation and drug metabolism compared to the HER2+ subtype (Figure 39A). Compared to the luminal A subtype, basal tumours exhibit upregulation of genes associated with cell cycle, IL17 signalling pathway, Wnt signalling pathway, signalling regulating the pluripotency of stem cells, and downregulation of drug metabolism and chemical carcinogenesis, both by DNA adducts and by reactive oxygen species (Figure 39B). Similarly, when compared to the luminal B subtype, the basal subtype shows upregulation of genes involved in Wnt signalling pathway and IL-17 signalling pathway, alongside downregulation of genes involved in chemical carcinogenesis and drug metabolism (Figure 39C).

4.8. Association with clinical data

Clinical data were available for 14 out of 16 patients, and are shown in Table 6. Among these, 3 patients (21.4%) were diagnosed with infiltrating apocrine breast cancer, while the remaining 11 (78.6%) had ductal infiltrating breast cancer. Tumour grades were either grade 2 (21.4%) or 3 (78.6%). No patients showed metastasis, and only one of them had a relapse (7.1%). Regarding TNM staging, 8 patients (57.1%) were classified as T1, 5 patients (35.7%) as T2 and 1 patient (7.1%) as T3. Lymph node metastasis was absent (N0) in 5 patients (35.7%), while other 5 (35.7%) were staged as N1, 1 as N2 and 2 (14.3%) as N3.

	Category	**Number of patients**	**Percentage (%)**
Histotype			
	Ductal infiltrating	11	78.57
	Infiltrating apocrine	3	21.43
Grade			
	G2	3	23.08
	G3	10	76.92
Tumour stage			
	T1	8	57.14
	T2	5	35.71
	T3	1	7.14
N stage			
	N0	5	38.46
	N1	5	38.46
	N2	1	7.69

	N3	2	15.38

Table 6. Clinical characteristics of TNBC patients

4.9. RNAseq data of TNBC samples

RNA sequencing was performed on a cohort of 16 TNBC patients. Principal component analysis (PCA) plots of mRNA and lncRNA expression showed overall similarity among samples, except for two distinct outliers visible as separated circles in both plots (Figure 40).

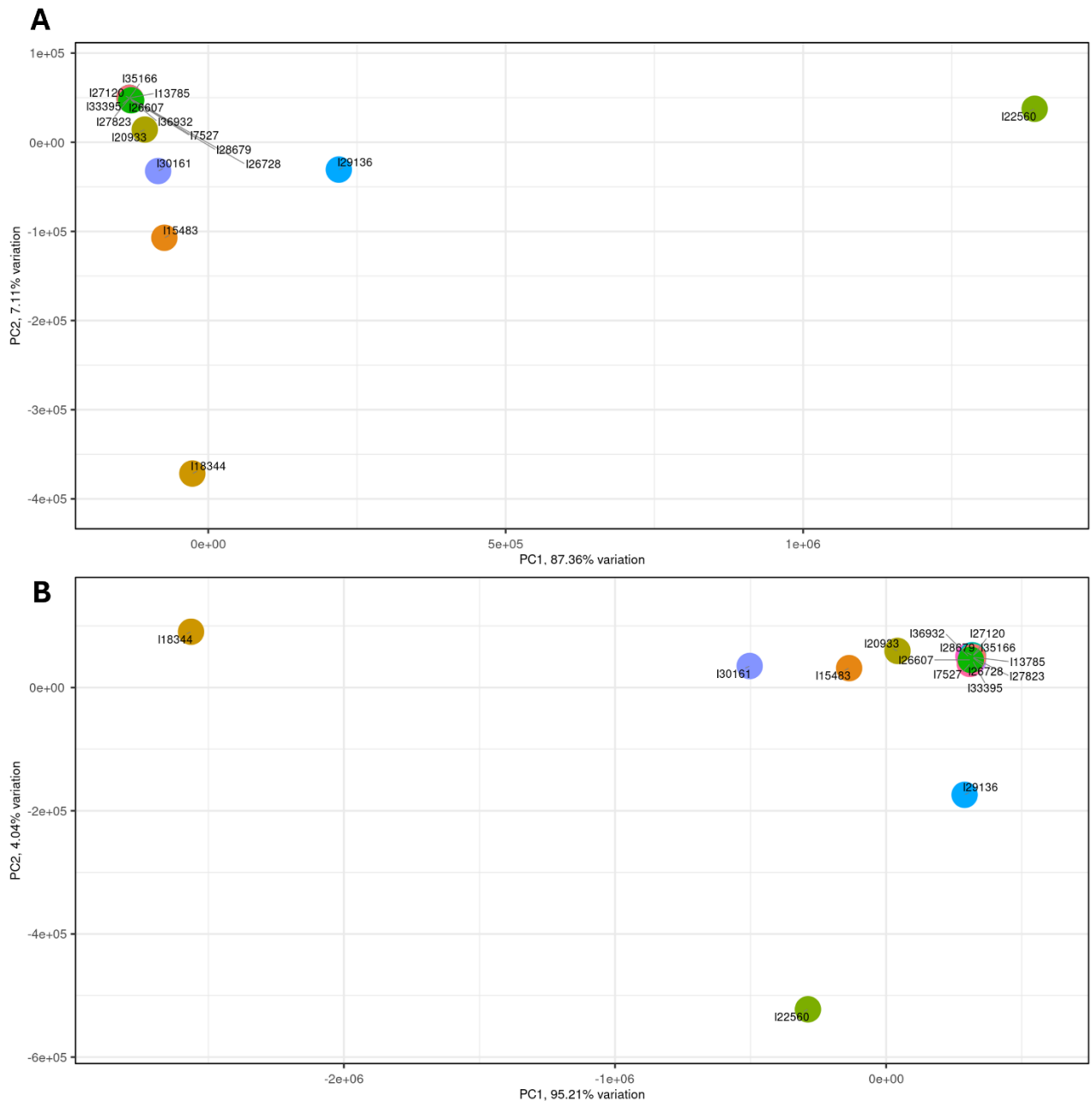

Figure 40. Principal component analysis of 16 TNBC based on A) mRNA expression and B) lncRNA expression (read counts)

Gene expression varied across samples, as depicted in the PCA plot, with only 5 samples showing notable gene co-expression. Correlation analysis between samples, based on both mRNA and lncRNA expression was conducted (Figure 41): in the correlation matrix it can be observed that some level of correlation was present

between the five samples with better RNA yield, but not with the other samples, and was less marked when based on lncRNA expression.

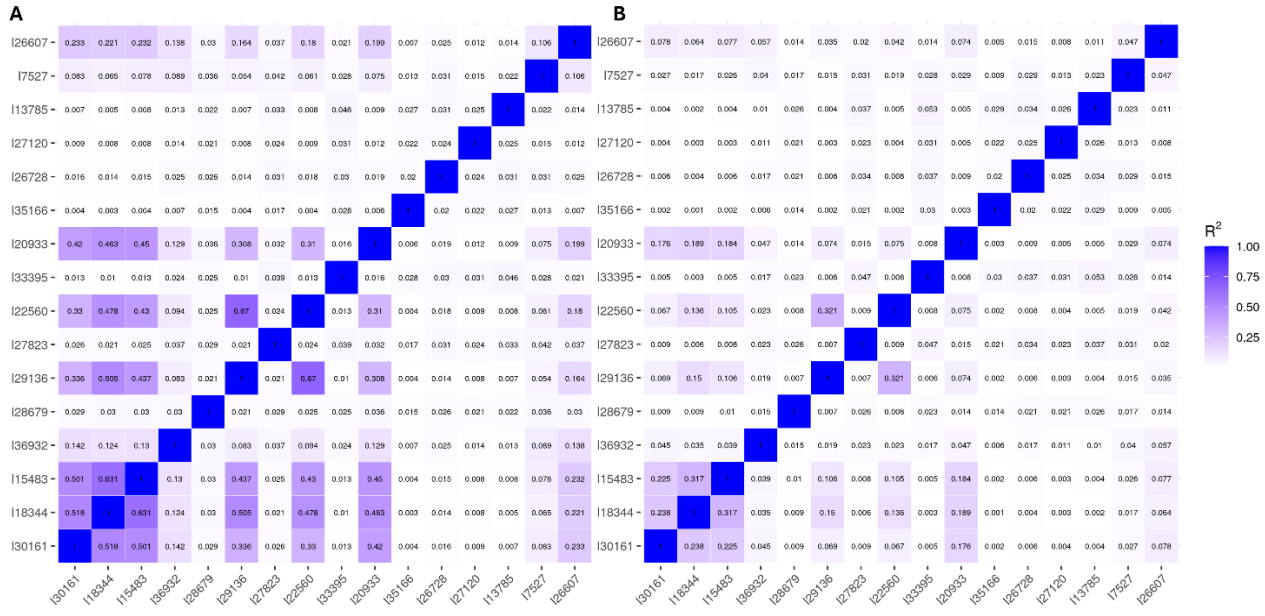

Figure 41. Correlation between samples based on A) mRNA expression and B) lncRNA expression

Possibly due to RNA quality and quantity, only 5 samples exhibited significant gene co-expression, as shown in Figure 42. The co-expression Venn diagram illustrates the number of genes uniquely expressed within each sample, while overlapping regions indicate the number of genes that are co-expressed in two or more samples. A total of 66 genes were found to be co-expressed among these samples.

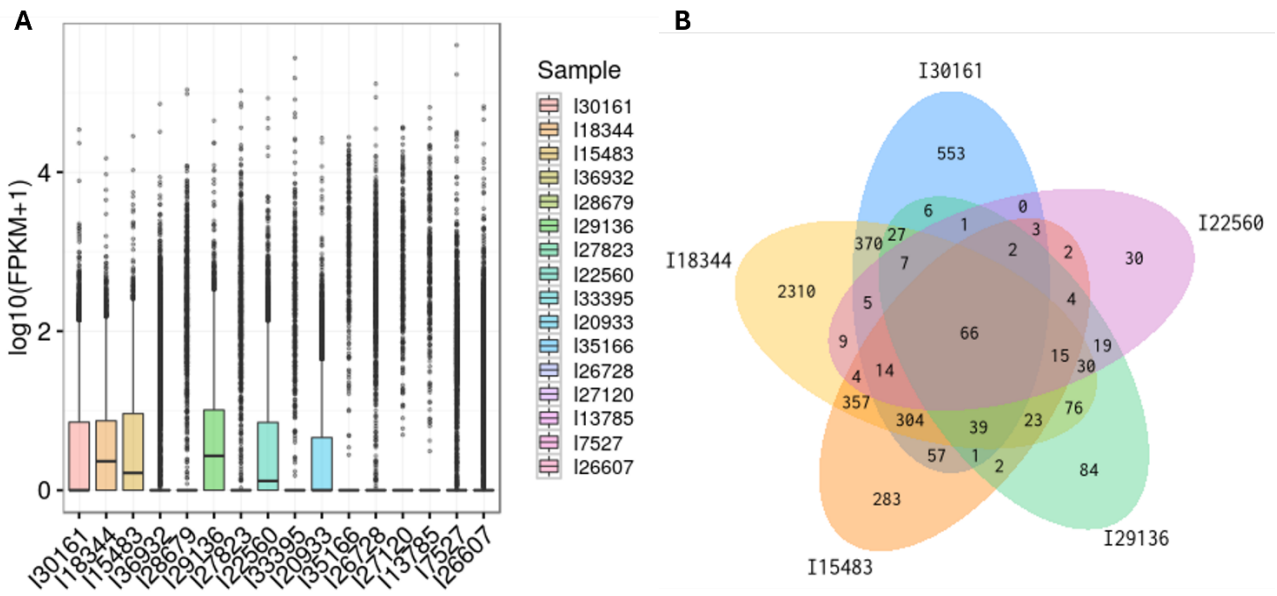

Figure 42. A) mean gene expression in each sample. B) Co-expressed genes in the five samples with better RNA quality

Differential expression analysis was performed between TNBC samples and healthy controls retrieved from the GTEX database. Figure 43A presents the principal component analysis (PCA) distinguishing TNBC from healthy samples: as shown, the two groups of samples are well separated with most of the samples clustered in definite areas, with one outlier in each group. Figure 43B illustrates gene counts for upregulation and downregulation across comparisons; although samples are highly variable, the majority of comparisons show substantially more upregulated genes than downregulated ones.

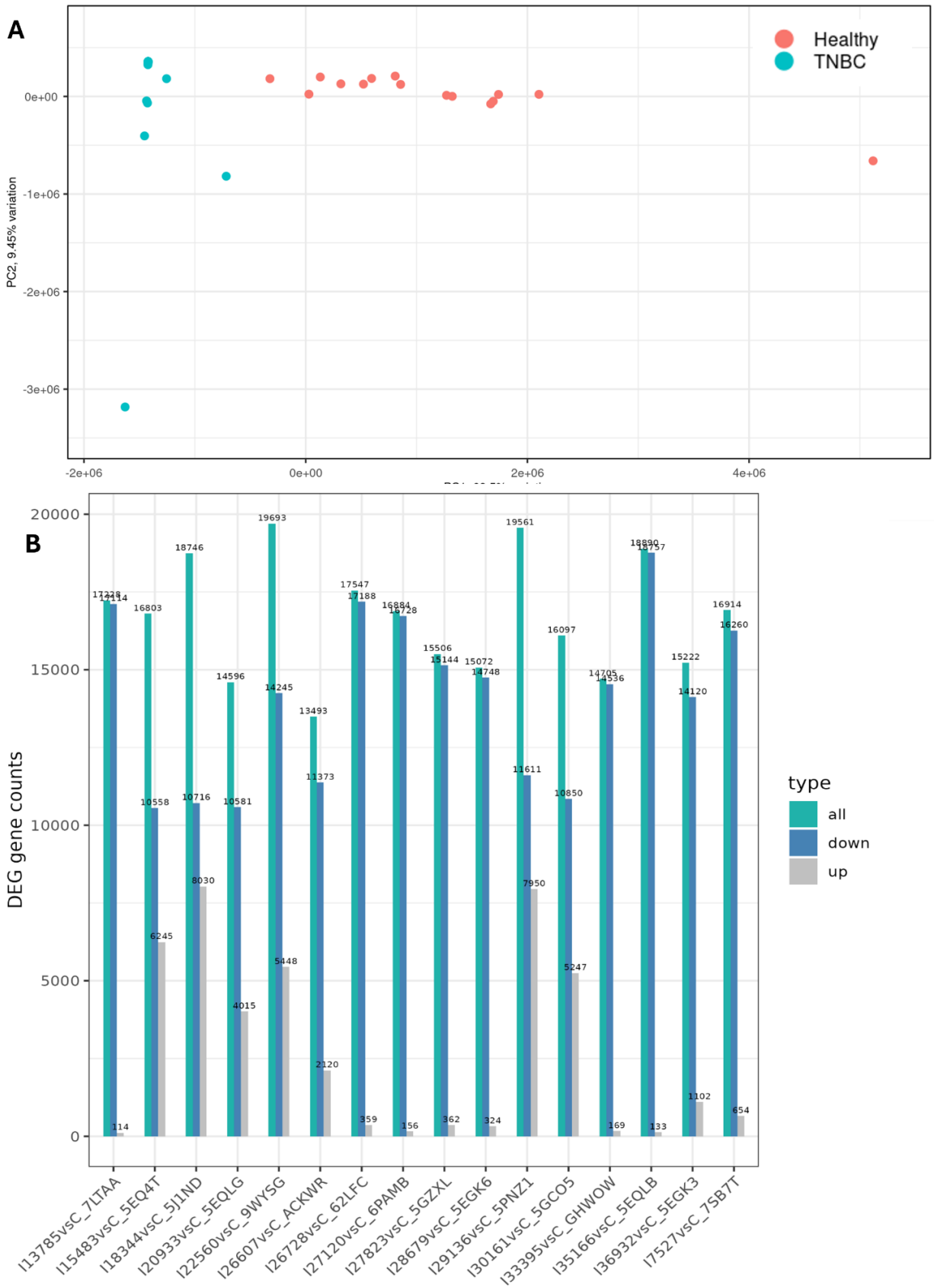

Figure 43. A) Principal component analysis of breast cancer samples and healthy

samples. B) Number of differentially expressed genes between triple negative breast cancer samples and healthy controls.

The differentially expressed genes exhibited substantial variation across all comparisons. Figure 44 displays Venn diagrams highlighting the overlap of DEGs shared between comparisons, identifying the genes most consistently upregulated or downregulated across groups.

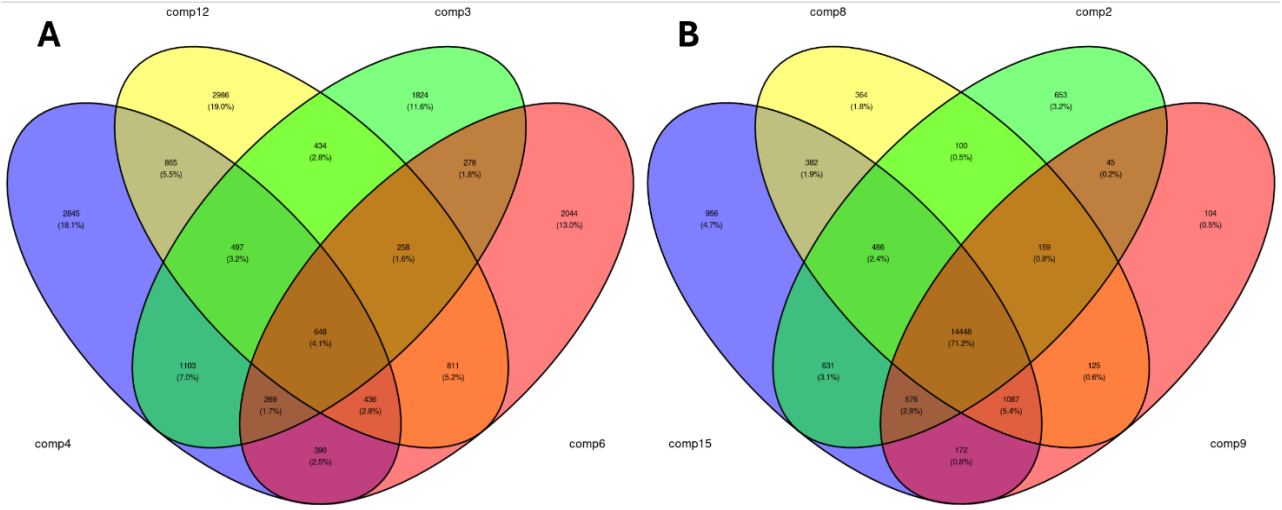

Figure 44. Venn diagrams showing the most common DEGs across different comparisons

4.10. Enrichment

Enrichment analysis was performed on DEGs, for each tumour versus healthy sample. However, to better characterise the differences between TNBC samples and healthy controls, we intersected dysregulated pathways to find the most common altered Gene Ontology (GO) (Figure 45A and 45B) and KEGG (Figure 45C and 45D) pathways, both up and downregulated, across the five samples exhibiting gene co-expression.

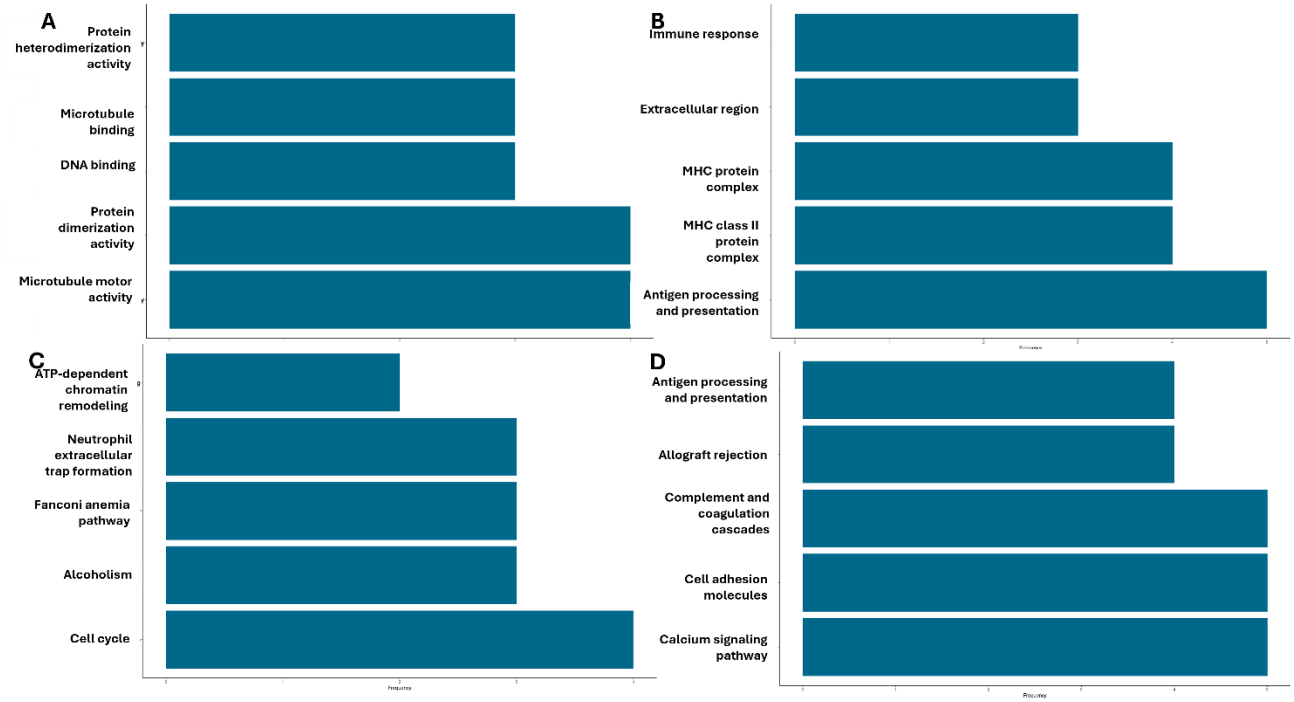

Figure 45. A) and B) Common GO dysregulated pathways. C) and D) common KEGG dysregulated pathways

Among the dysregulated GO pathways, upregulated terms included microtubule motor activity, protein dimerization, microtubule binding, DNA binding. Downregulated GO pathways included antigen processing and presentation and MHC II protein complex. Dysregulated KEGG pathways included, among upregulated ones, Fanconi anaemia, cell cycle, chromatin remodelling and neutrophil trap formation. Among downregulated KEGG pathways there were antigen processing and presentation, cell adhesion molecules and calcium signalling pathway.

Furthermore, we performed enrichment analysis on common differentially expressed genes. Overall, TNBC samples demonstrated upregulation in protein targeting to the membrane and to the endoplasmic reticulum, and downregulation in antigen processing and presentation, MHC II protein complex assembly, collagen containing extracellular matrix, components of ribosomes and protein binding (Figure 46).

Figure 46. Enriched pathways according to common DEGs. A) Biological process, upregulated B) Biological process, downregulated, C) Cellular component, upregulated D) Cellular component, downregulated, E) Molecular function, downregulated

4.11. Identification of single nucleotide polymorphisms in the transcripts

RNA sequencing also identified insertions/deletions (INDELs) and single nucleotide polymorphisms (SNPs) within the RNA sequence. Figure 47 presents the impact of INDELs (A) and SNPs (B) on genes in each sample. Several variants act as modifiers in most of the cases, and are present in all samples, whereas the ones with high, moderate or low impact are less frequent. The number of INDELs and SNPs is highly

variable across samples, ranging from as few as six or seven in some samples to several thousands in others, thus the variant counts are reported on a logarithmic scale for better visualization.

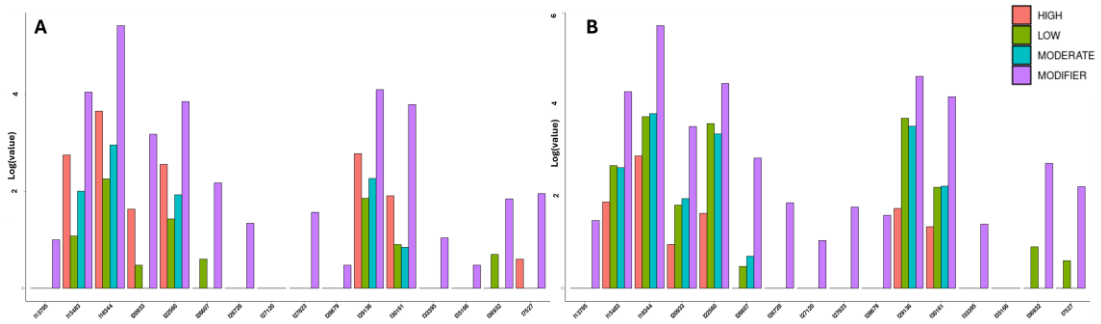

Figure 47. Number and effect of INDELs (A) and SNPs (B) in each sample

4.12. lncRNAs identification and correlation with known messenger RNA

RNA sequencing identified both known and novel transcripts, belonging to both mRNA and lncRNA categories. The characteristics of these transcripts are summarized in Figure 48, showing that novel mRNAs and lncRNAs share similar features with their known counterpart.

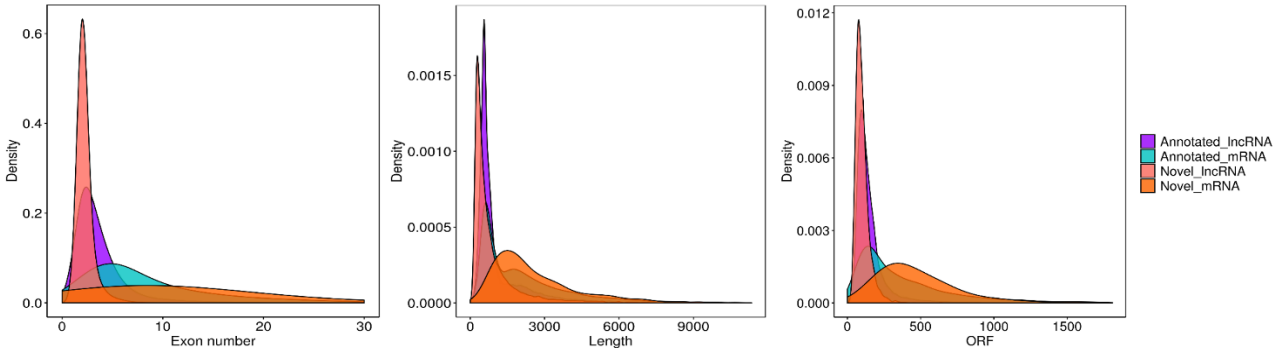

Figure 48. Exon number (A), length of transcript (B) and ORFs (C) of novel and annotated mRNAs and lncRNAs

Novel transcript underwent protein coding potential analysis through three different algorithms: CPC, CNCI, PFAM (Figure 49A). The intersection of these three methods identified 11999 novel lncRNAs, which were further classified into intronic lncRNAs

(49%), overlapping lncRNAs (21.6%), lincRNAs (20.6%) and antisense lncRNAs (8.6%). The genomic distribution of novel lncRNAs is presented in Figure 49C.

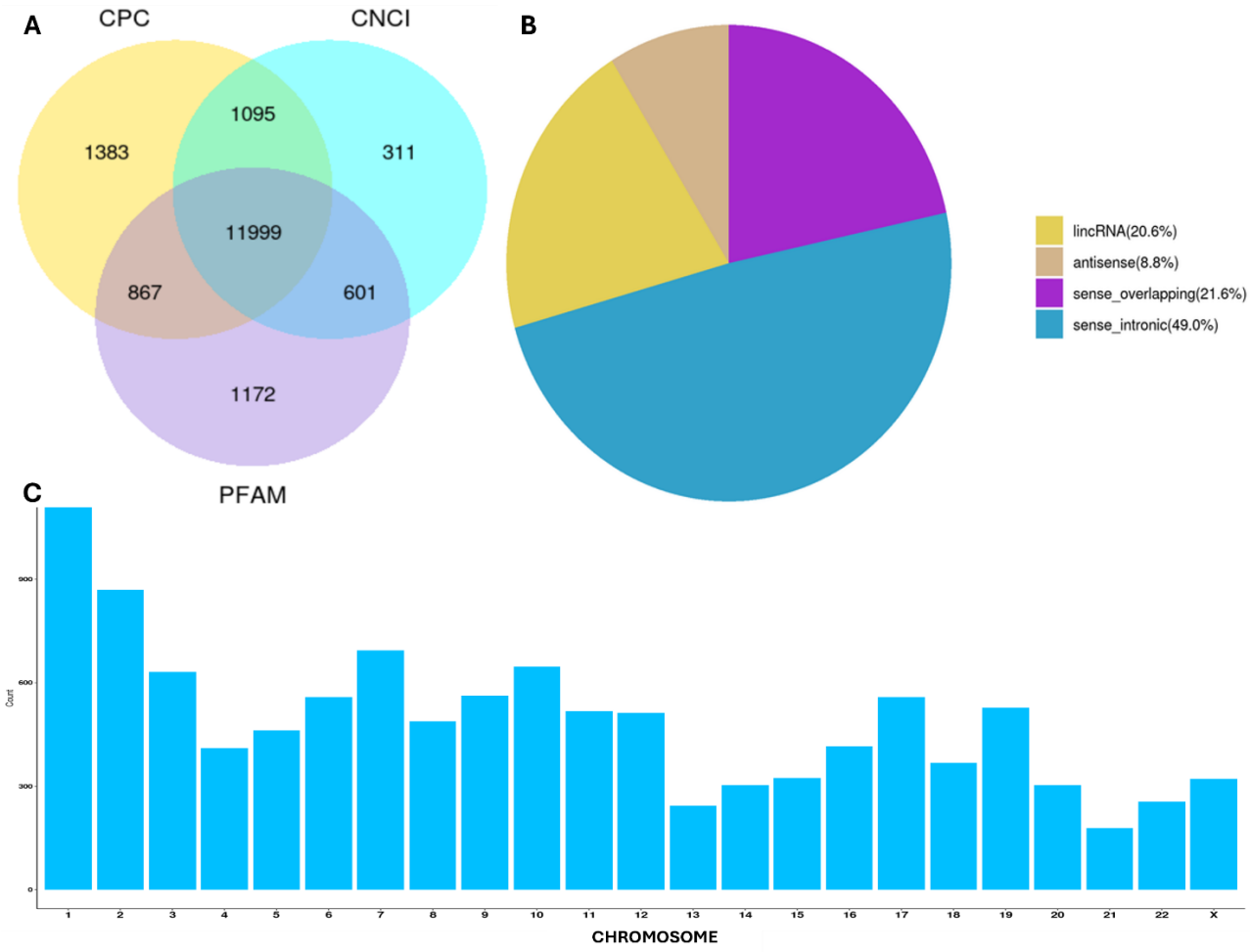

Figure 49. A) Protein coding potential analysis of novel transcripts, B) classification of novel lncRNAs and C) distribution across chromosomes of novel lncRNAs

Figure 50 shows the most common across samples annotated (A) and novel (B) lncRNAs.

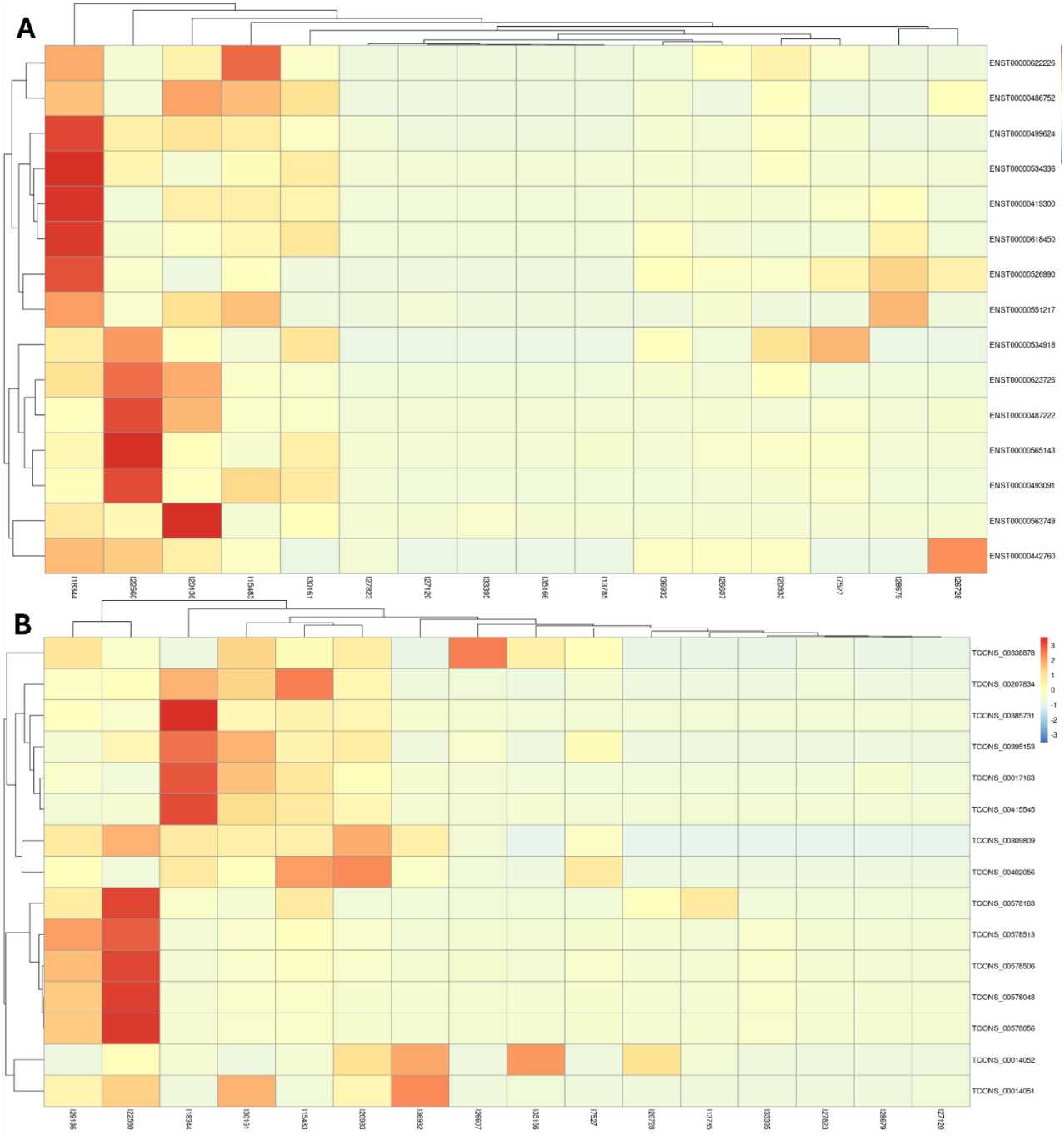

Figure 50. Top 15 most common annotated (A) and novel (B) lncRNAs

We then evaluated the potential co-expression between lncRNAs and mRNAs using Pearson correlation analysis. A positive correlation between a lncRNA and a mRNA was defined as a Pearson correlation greater than 0.7 and a p-value less than 0.05. Furthermore, we performed enrichment analysis on co-expressed genes, to elucidate the possible biological role of the identified lncRNAs.

Figure 51 shows GO categories (Biological Process/BP, Cellular Component/CC, Molecular Function/MF) and Disease Gene Network (DGN) enriched pathways. DGN enrichment revealed involvement in inflammation, various types of cancer (meningioma, sarcoma), invasive breast carcinoma and intraductal carcinoma.

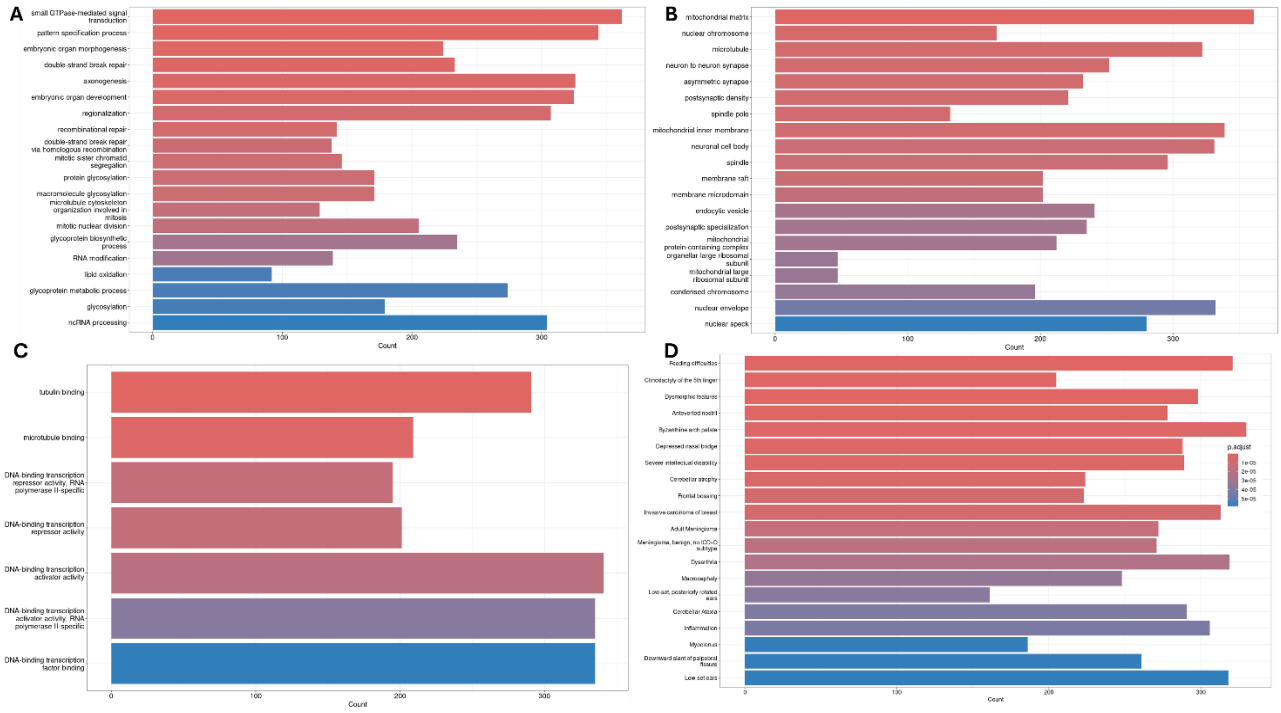

Figure 51. Pathway enrichment in genes co-expressed with novel lncRNAs A) BP enriched pathways, B) CC enriched pathways C) MF enriched pathways D) DGN enriched pathways

Genes associated with various breast cancer subtypes (basal-like breast carcinoma, familial breast cancer, HER2-negative breast cancer and invasive ductal breast carcinoma), were all tightly interconnected (Figure 52).

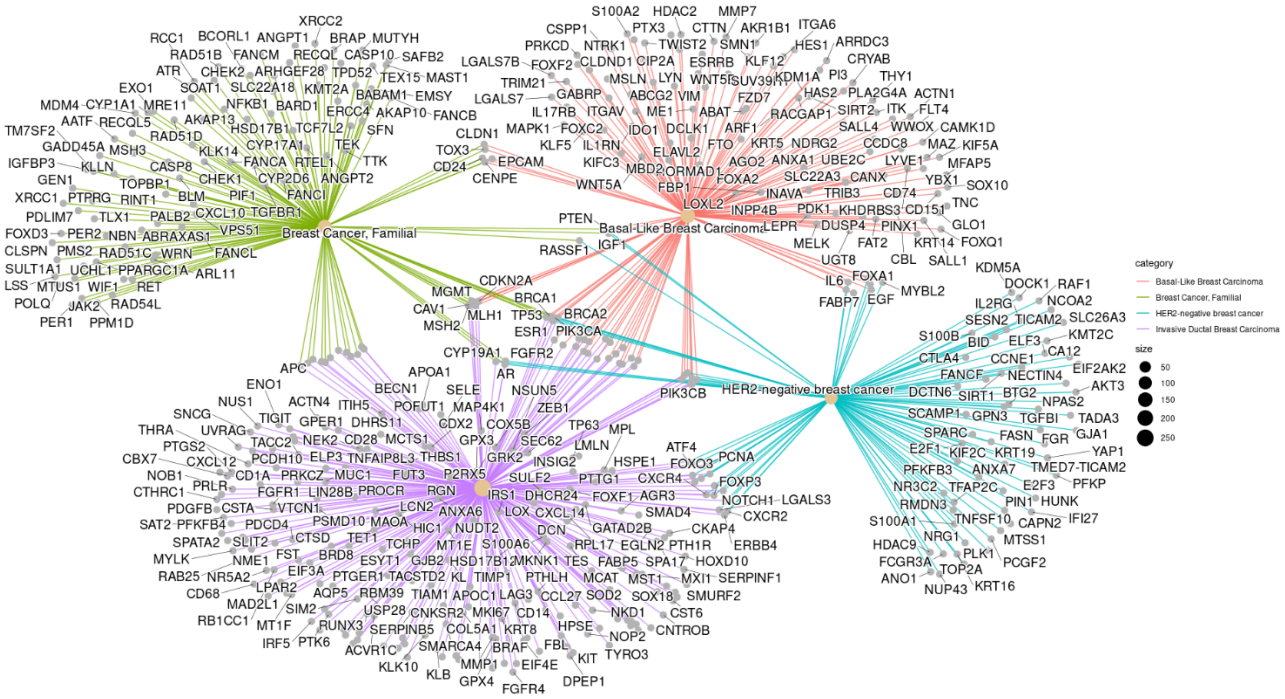

Figure 52. Network of enriched genes in DGN pathways involved in breast carcinoma

KEGG pathway analysis showed enrichment in cytokine-cytokine receptor interactions, various types of chemical carcinogenesis (receptor activation, DNA adducts), basal cell carcinoma and microRNAs in cancer (Figure 53).

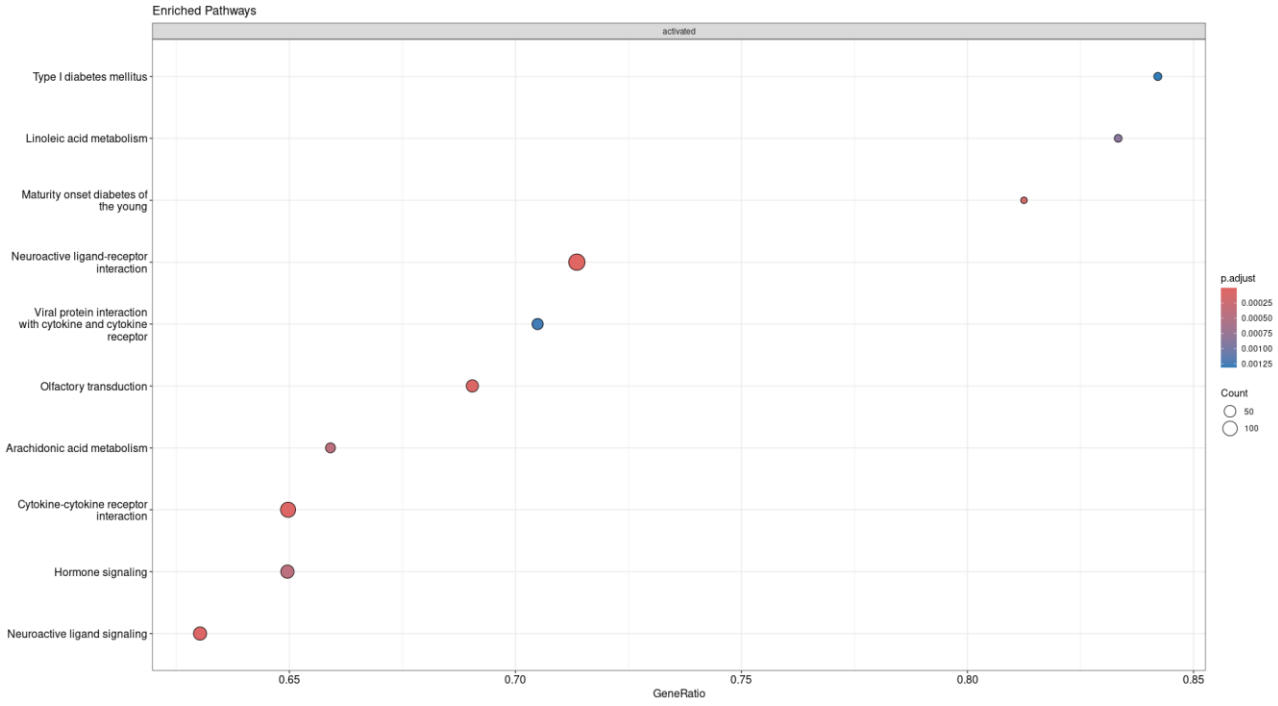

Figure 53. Dot plot of KEGG enriched pathways co-expressed with novel lncRNAs

We performed differential expression analysis based on the occurrence of relapse. No significant genes were found; however, we found one novel lncRNA (XLOC_407427) to have a higher expression in the patient who experienced relapse (Table 7).

GeneName	Symbol	Log2FoldChange	PValue	AdjustedPValue
XLOC_407427		12.674	0.024249	0.989
ENSG00000156219	ART3	12.108	0.031355	0.989
ENSG00000205221	VIT	11.668	0.038048	0.989
ENSG00000189283	FHIT	11.362	0.043026	0.989
ENSG00000121904	CSMD2	11.341	0.043787	0.989
ENSG00000166535	A2ML1	11.102	0.048418	0.989
ENSG00000173698	ADGRG2	9.8605	0.041535	0.989

Table 7. Top upregulated genes in the patient who presented relapse compared to other patients

We also performed DEG analysis according to the histotype, with the infiltrating apocrine histotype having a substantial number of downregulated genes compared to the ductal infiltrating histotype (Figure 54).

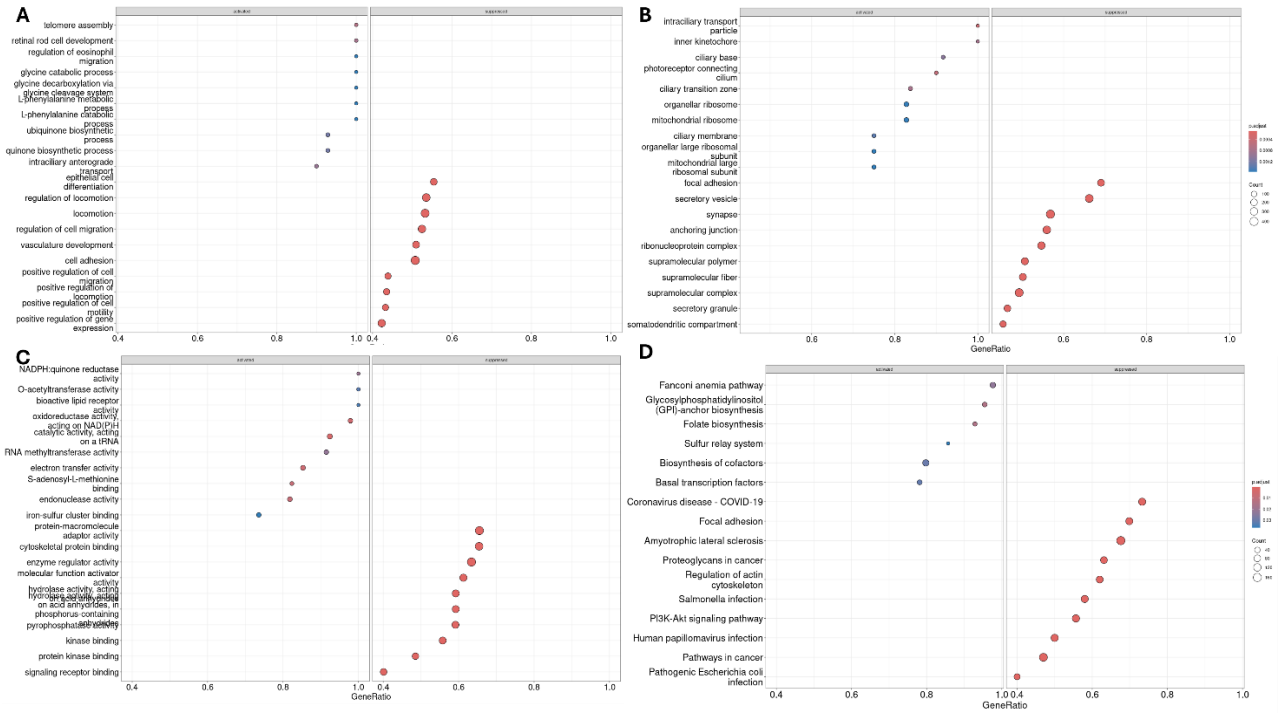

Figure 54. Differentially enriched pathways between histotypes A) BP, B) CC C) MF, D) KEGG

Notably, the infiltrating apocrine histotype showed a lower expression of genes involved in cell adhesion and cell motility.

4.13. Novel lncRNAs effect on survival

We performed survival analysis according to the expression (categorised as high or low) of novel identified lncRNAs and four of them (TCONS_00137122, TCONS_00332234, TCONS_00578113, TCONS_00255783) turned out to have significant effect on distant metastasis free survival (DMFS) (TCONS_00137122, TCONS_00332234 and TCONS_00578113) and on progression free survival (PFS) (TCONS_00578113 and TCONS_00255783). The high expression of TCONS_00332234 correlated with a better DMFS (p=0.032), while the higher expression of the other three lncRNAs correlated with a worse DMFS or PFS. TCONS_00578113 showed an effect on survival in both DMFS (p=0.036) and PFS (p=0.021) (Figure 55).

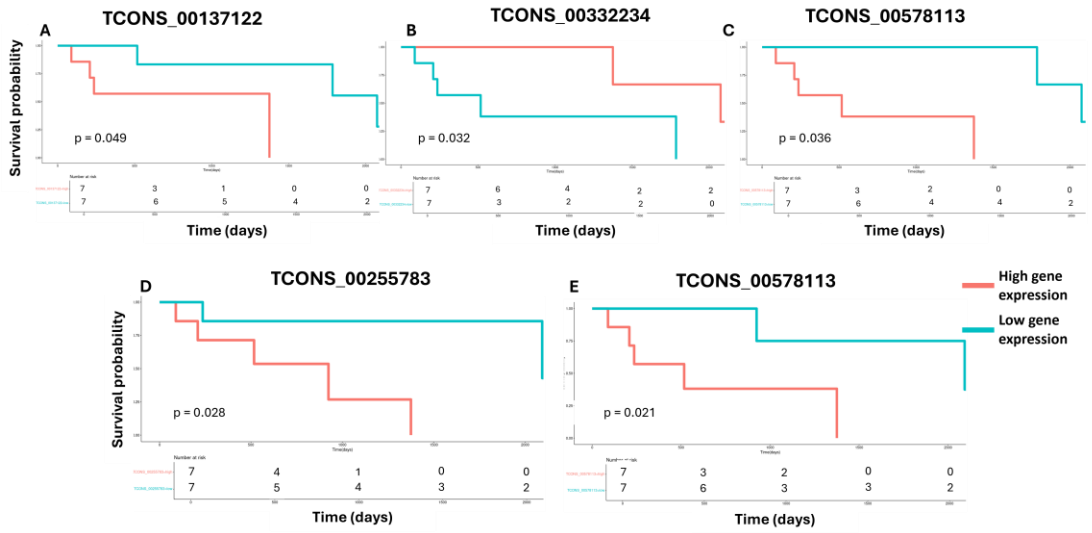

Figure 55. Survival curves according to the expression of novel lncRNA. A) Distant metastasis free survival (DMFS) according to the expression of TCONS_00137122 B) DMFS according to the expression of TCONS_00332234 C) DMFS according to the expression of TCONS_00578113 D) Progression Free Survival (PFS) according to the expression of TCONS_00255783 E) PFS according to the expression of TCONS_00578113

4.14. RNAseq on FFPE samples from patients belonging to different ethnic groups

We also performed RNAseq on FFPE samples of patients from different ethnic groups, to analyse their transcriptomic profile. Their demographic and clinical characteristics are shown in Table 8.

	Category	**Number of patients**	**Percentage (%)**
Age (years)			
	<40	3	30.00
	40-49	4	40.00

	50-59	2	20.00
	60-69	1	10.00
Subtype			
	HER2+	2	20.00
	Luminal A	5	50.00
	Luminal B	3	30.00
Race/Ethnicity			
	African	2	16.67
	Arab	2	16.67
	Asian	3	25.00
	European	5	41.67
Grade			
	G1	4	40.00
	G2	6	60.00
Tumour stage			
	T1	5	50.00
	T2	4	40.00
	T3	1	10.00
N **stage**			
	N0	4	40.00
	N1	4	40.00

	N2	1	10.00
	N3	1	10.00

Table 8. Demographic and clinical characteristics of breast cancer patients on who RNAseq was performed

In Figure 56, mean gene expression of each sample, and principal component analysis (PCA) of the samples according to lncRNA and mRNA expression is shown. Mean gene expression (Figure 56A) shows that most samples have similar RNA quality and comparable gene expression. In Figure 56B and 56C, PCA of the samples is displayed, showing closeness between six samples based on both lncRNA and mRNA expression.

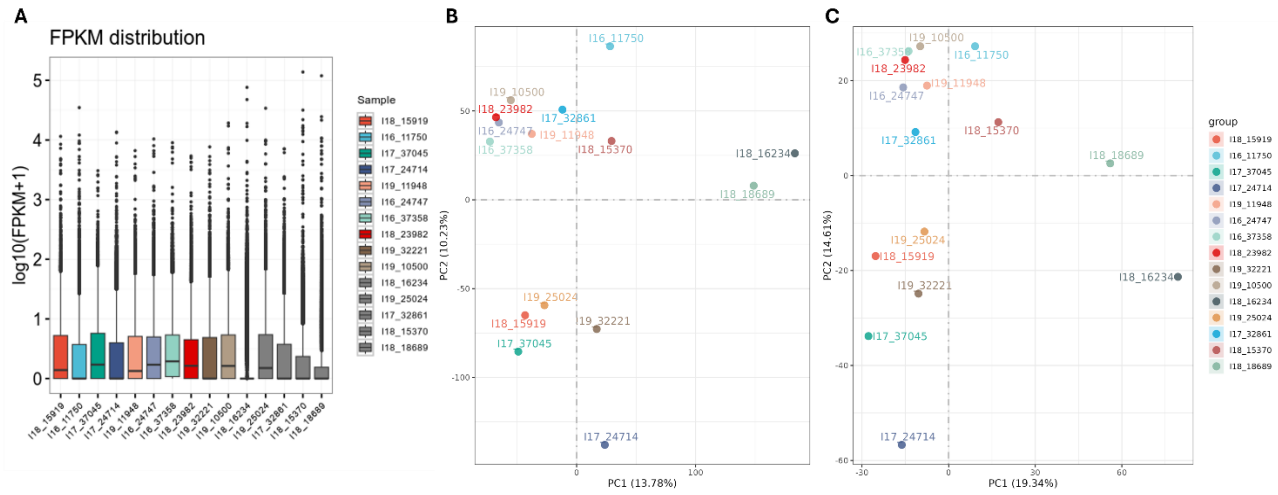

Figure 56. A) Mean gene expression in each sample B) Principal component analysis based on lncRNA expression C) Principal component analysis based on mRNA expression

Pearson correlation analysis based on mRNA expression (Figure 57B) showed a correlation between samples I19-10500, I19-11948, I16-24747 and I16-37358, which were also close to each other in the PCA plot. These similarities, however, were not related to ethnicity, as the samples are respectively European, Asian, African and European, but they all belong to luminal subtypes (either A or B).

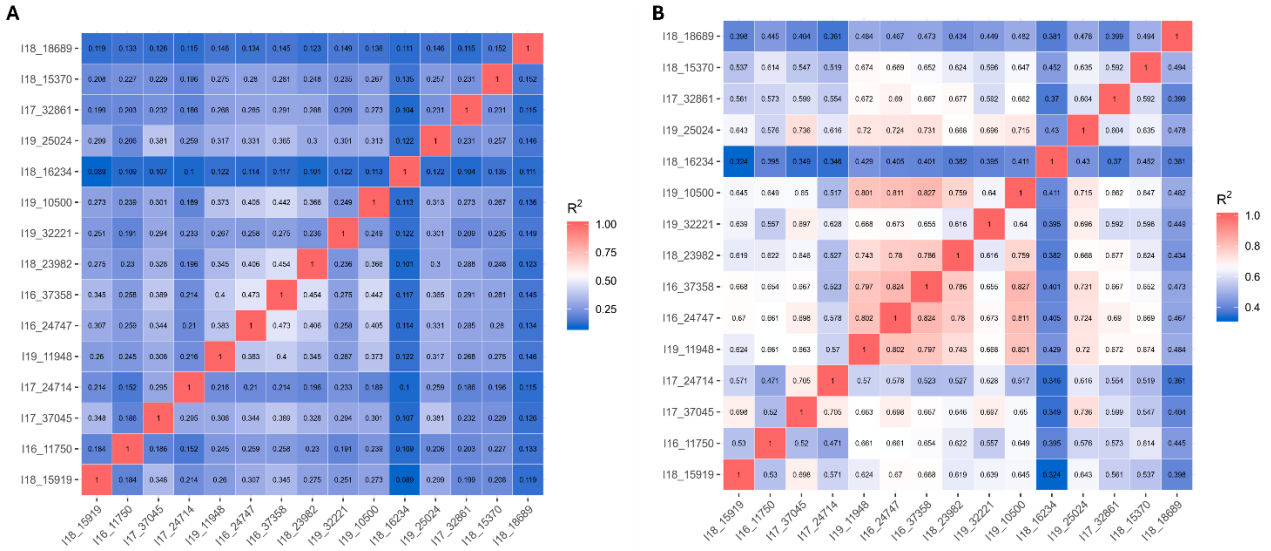

Figure 57. Pearson correlation between samples according to (A) lncRNA and (B) mRNA expression

Venn diagrams were then plotted to identify genes co-expressed in the same ethnicity (Figure 58), selecting two African patients (Figure 58A, E), two Arab patients (Figure 58B, F), three Asian patients (Figure 58C, G) and five European patients (Figure 58 D, H). African patients shared 6023 mRNAs and 6485 lncRNAs; Arab patients showed co-expression of 3953 mRNAs and 3776 lncRNAs; Asian patients co-expressed 4082 mRNAs and 3216 lncRNAs; finally, European patients co-expressed 6568 mRNAs and 6772 lncRNAs.

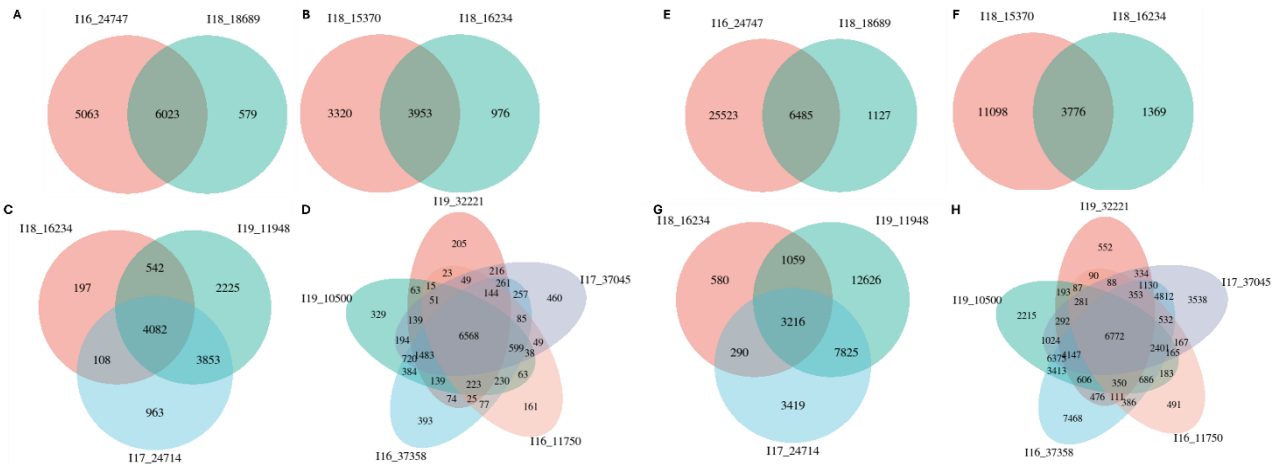

Figure 58. Venn diagrams of co-expressed genes in African (A, E), Arab (B, F), Asian (C, G), and European (D, H) patients, based on mRNA (A, B, C, D) and lncRNA (E, F, G, H) expression

An enrichment analysis was then performed on co-expressed genes (Figure 59). Co-expressed genes specific in African patients were involved in drug, selenocompound and glucose metabolism, growth hormones and regulation of TRP channels by inflammation. Arab patients, instead, presented enrichment in adipocytokine signalling pathway and metabolism of steroids and nucleotides.

Asian patients presented enrichment in glycosylphosphatidylinositol anchor biosynthesis, motor proteins and stem cells.

Finally, European patients were enriched in ECM-receptor interaction and folate metabolism.

PI3K-Akt signalling pathway enrichment was present in all ethnicities, suggesting a key role of this pathway in breast carcinogenesis. Allograft rejection and other autoimmune pathways were also present in every ethnicity.

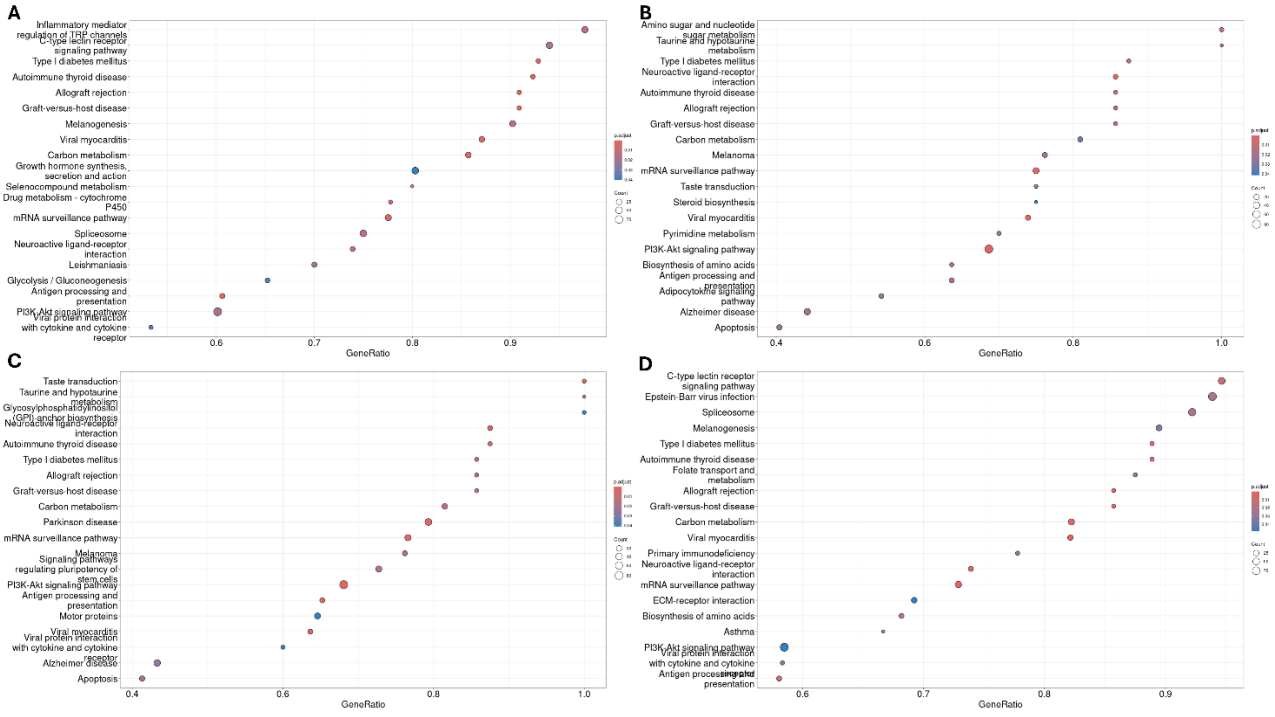

Figure 59. Pathway enrichment (KEGG) of co-expressed genes in African (A), Arab (B), Asian (C) and European (D) patients

Differential expression analysis was then performed on pairs of patients from different ethnicities (Figure 60). Differential expression of mRNA was more evident than differential expression of lncRNAs.

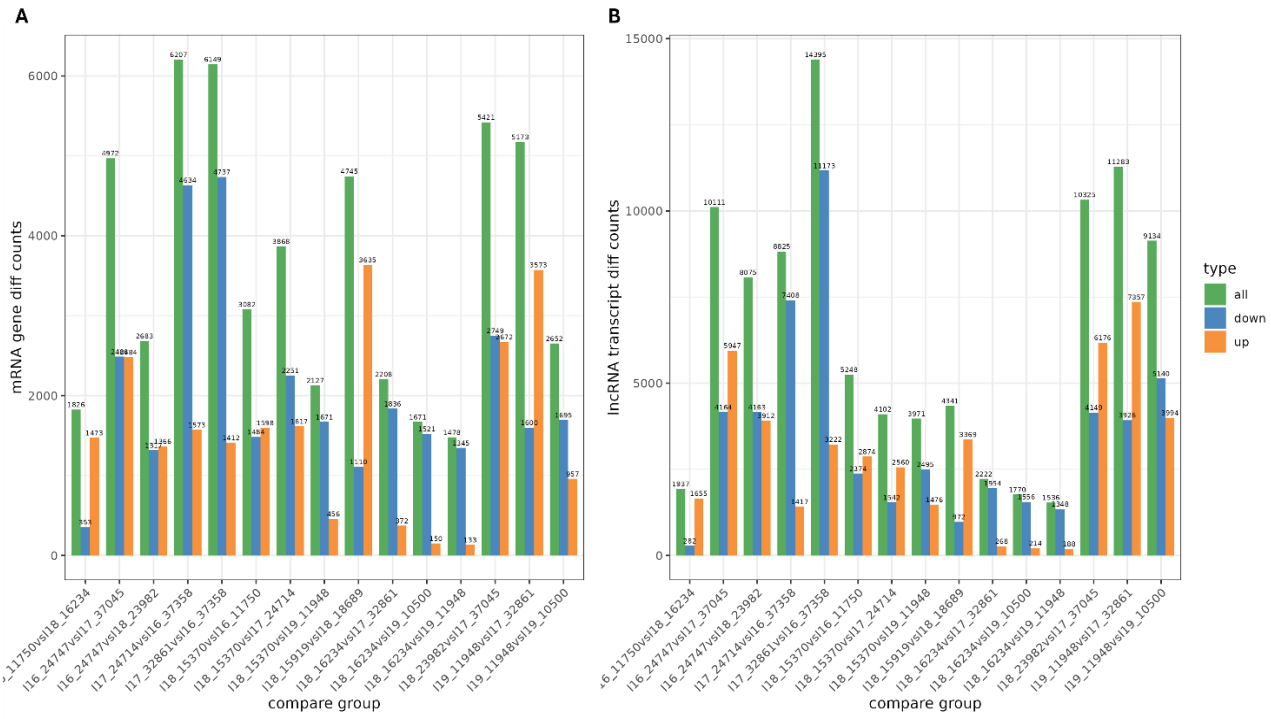

Figure 60. Differentially expressed mRNA (A) and lncRNA (B) between patients of different ethnicities

In the HER2+ subtype, a comparison was made between one African and one European patient, revealing the upregulation of metabolism pathways (sulfur metabolism, cysteine and methionine metabolism, arginine biosynthesis) and the downregulation of immune response in the European patient compared to the African patient (Figure 61).

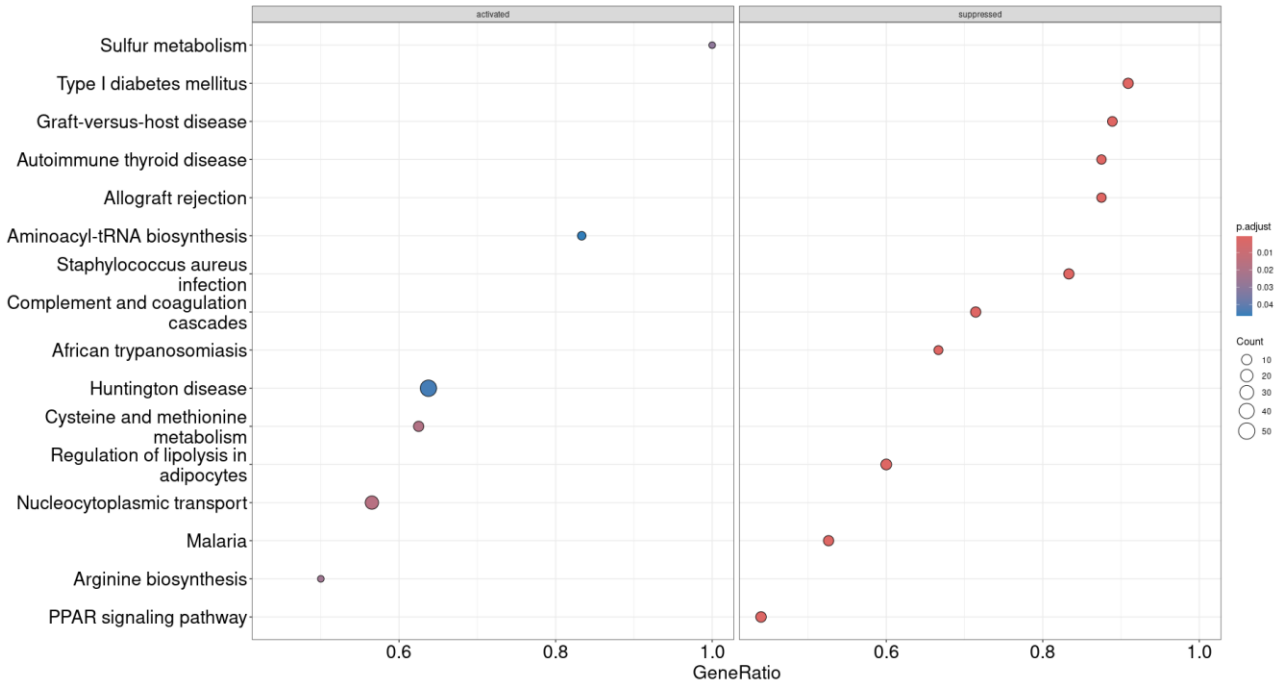

Figure 61. Pathway enrichment of upregulated and downregulated genes in European patients compared to African patients in HER2+ subtype

In the luminal A subtype, Arab patients present downregulation of proteoglycans in cancer and colorectal cancer (Figure 62) compared to European ones.

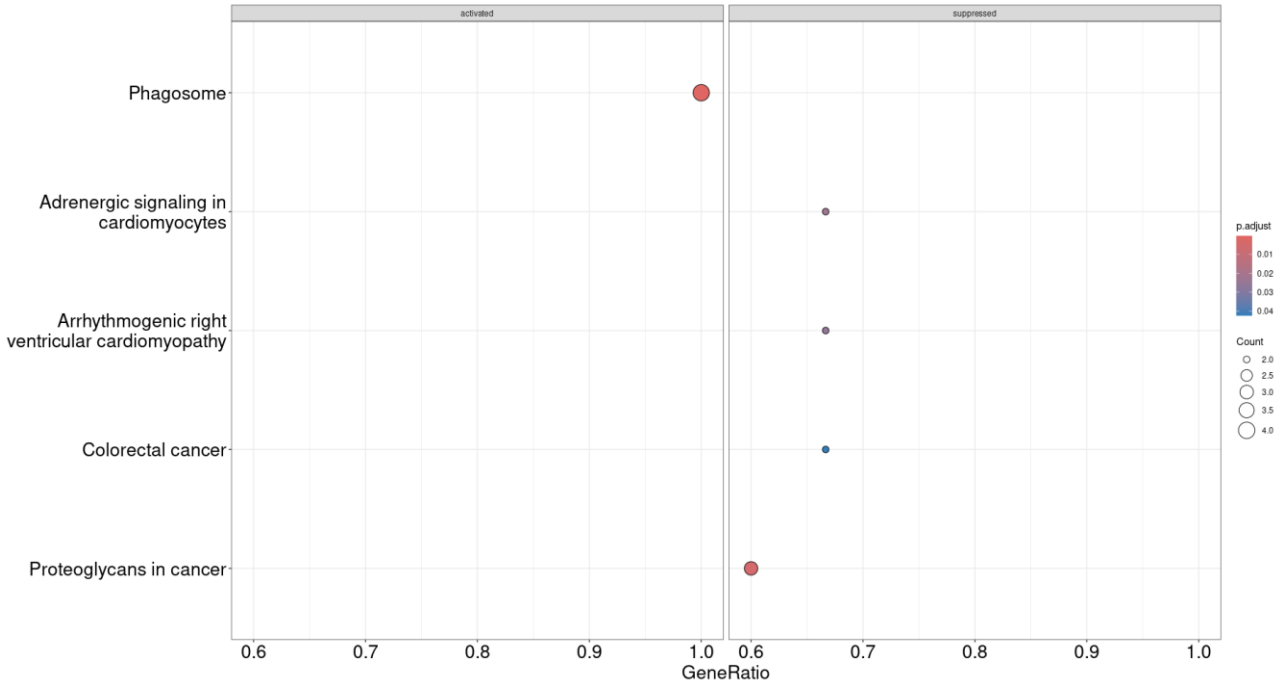

Figure 62. Pathway enrichment of upregulated and downregulated genes in Arab patients compared to European patients in luminal A subtype

Compared to European (Figure 63), Asian patients present upregulation in necroptosis, choline metabolism in cancer, TGF-B signalling pathway, and Hippo signalling pathway, and downregulation of cellular senescence and cell cycle.

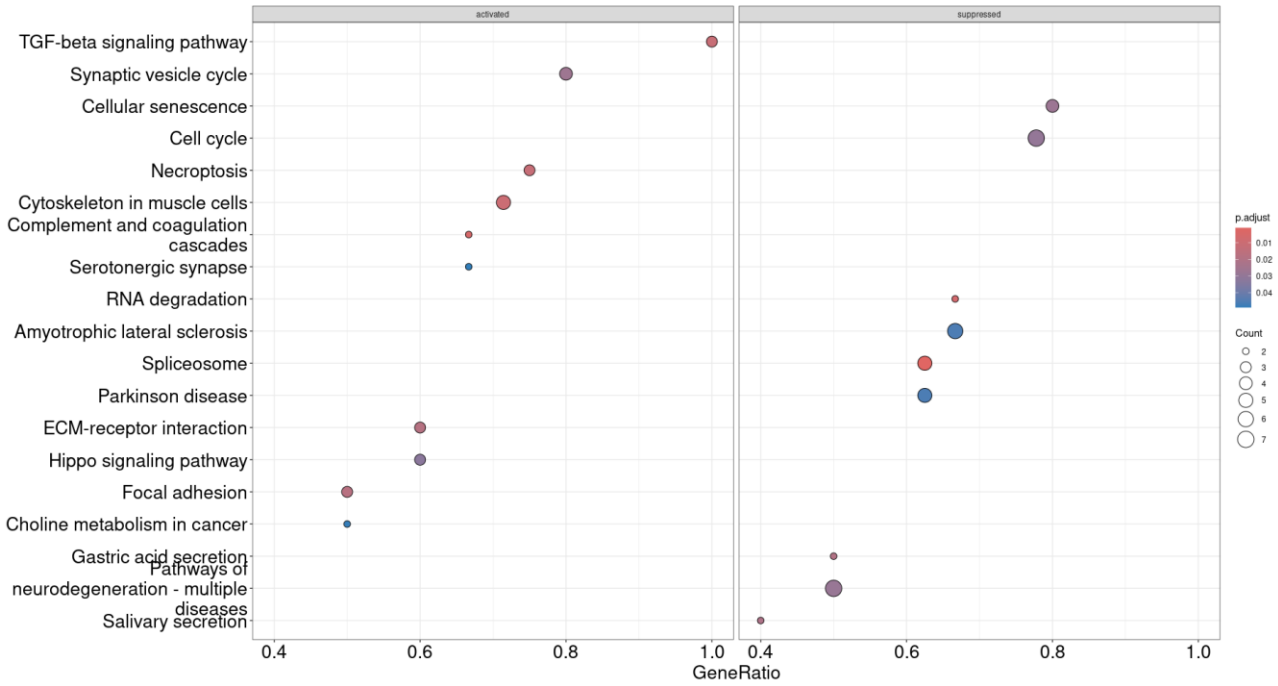

Figure 63. Pathway enrichment of upregulated and downregulated genes in Asian patients compared to European patients in the luminal A subtype

Compared to Asian patients, Arab ones have upregulation of PPAR signalling pathway and apoptosis (Figure 64).

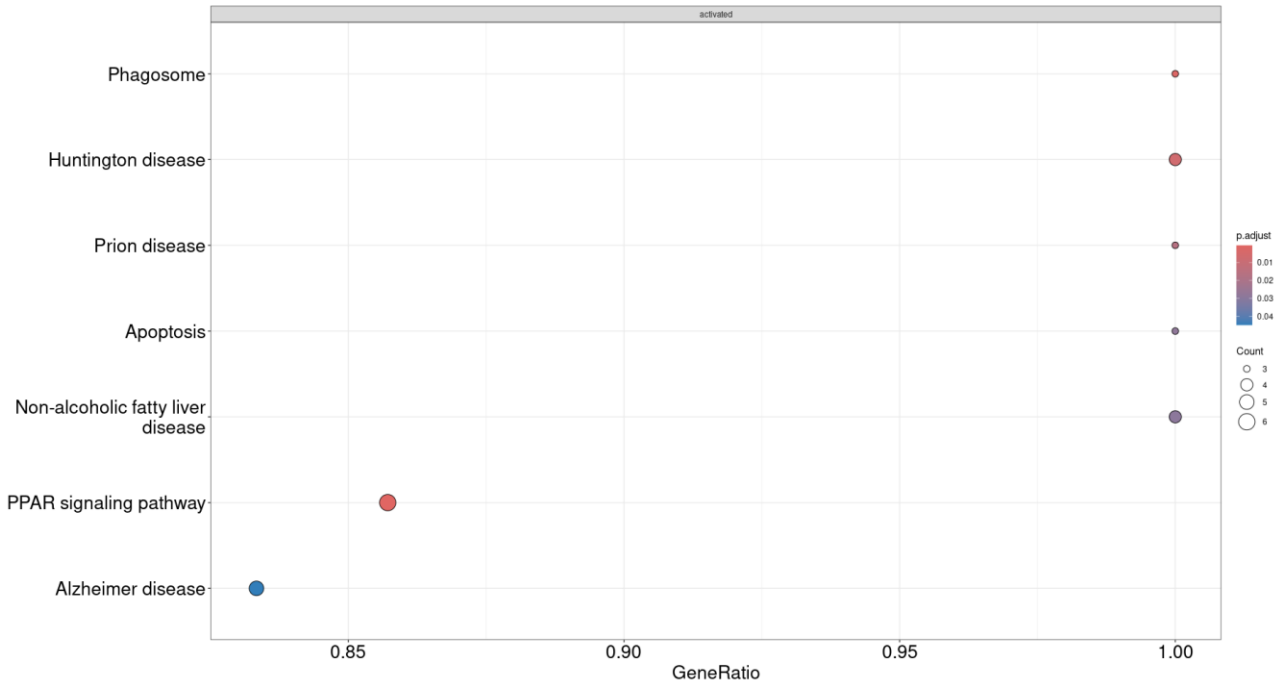

Figure 64. Pathway enrichment of upregulated genes in Arab patients compared to Asian patients in the luminal A subtype

In the luminal B subtype, African patients have enrichment of signalling pathways (cAMP, adrenergic signalling, cortisol) and downregulation of cell cycle (Figure 65).

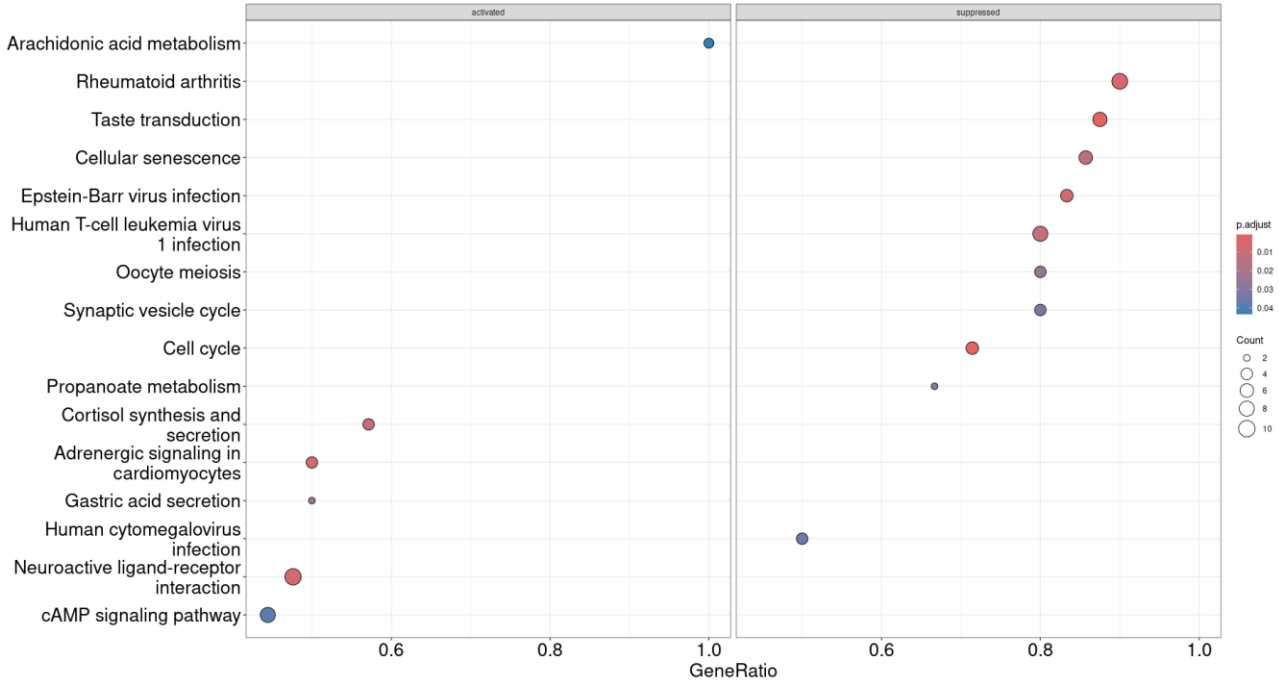

Figure 65. Pathway enrichment of upregulated and downregulated genes in African patients compared to European patients in the luminal B subtype

Regarding specific genes, we could not confirm the differential expression found in the TCGA cohort, however this could be caused by the low number of patients in our cohort.

Among ion channel genes only *KCNIP3* was confirmed as upregulated in our cohort of African patients compared to other ethnic groups (Figure 66).

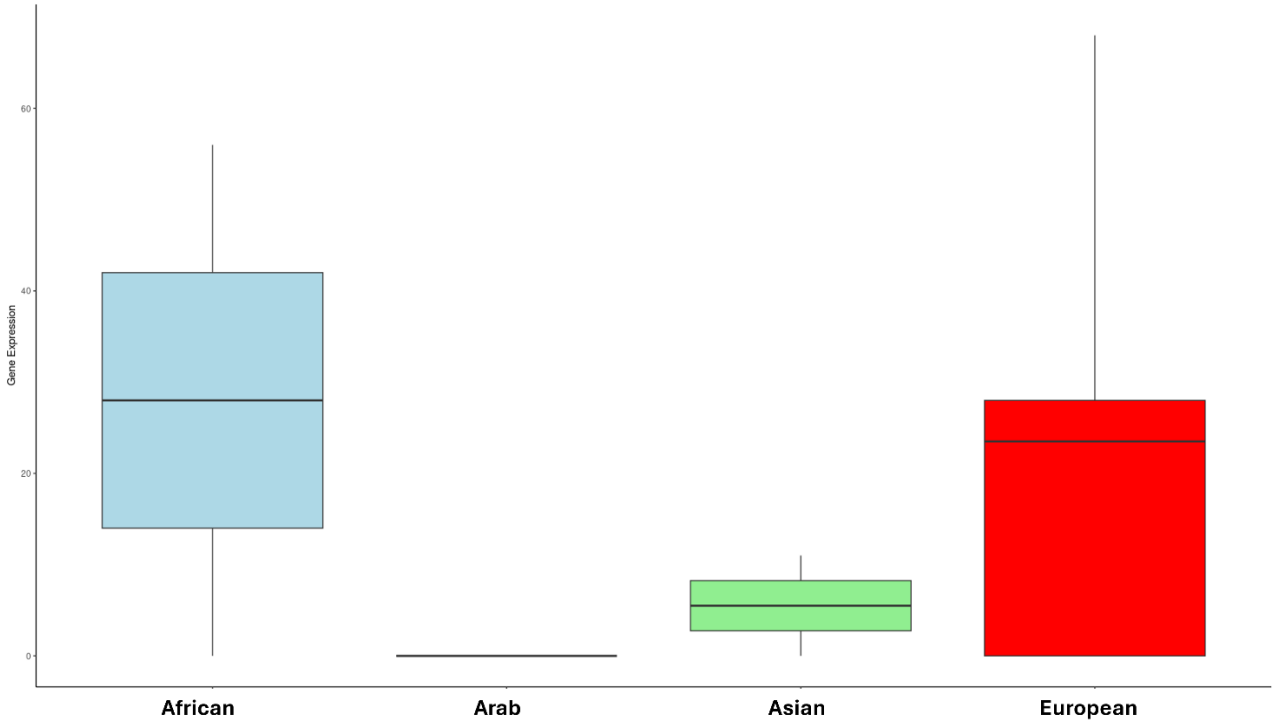

Figure 66. KCNIP3 RNA expression in different ethnic groups

5. Discussion

Breast cancer is the most commonly diagnosed type of cancer in women in almost all countries, with few exceptions, with its incidence growing rapidly in both developed and developing countries, and mortality increasing in the latter (N. Li et al., 2019). Both incidence, mortality and biology of breast cancer however have shown to be ethnicity or race-dependent, with women of African descent presenting more aggressive disease and worse outcomes compared to other ethnic groups and women of Asian descent having lower incidence rates and mortality of breast cancer (Menon et al., 2025; Telli et al., 2011). Although the higher mortality of breast cancer in ethnic minorities has been attributed to a worse or lacking medical care, less participation in regular screening and a worse stage at prognosis (Dunn et al., 2024; Ooi et al., 2011), some variability remains after adjusting for socioeconomical factors (Hill et al., 2018).The intrinsic biology of breast cancer remains variable across ethnic groups, including breast cancer subtype frequency, mutation rate, gene expression, DNA methylation and tumour microenvironment composition. In this project, the different distribution of breast cancer subtypes was confirmed from the TCGA cohort, with Asian patients presenting more often with the HER2+ subtype (Hjerkind et al., 2022); furthermore, they present higher expression of genes involved in the ERBB pathway compared to Black patients even in the basal subtype, suggesting a higher involvement of this pathway in general breast cancer development in this ethnicity, a characteristic they share with Native American women, whose ERBB2 expression is proportional to the percentage of Native American ancestry (Marker et al., 2020; Rey-Vargas et al., 2025). This similarity could be due to the genetic link between Asian and Amerindian populations (Tokunaga et al., 2001)

Consistently with literature-confirmed higher aggressivity of breast cancer occurring in Black patients, the tumours present lower expression of adhesion molecules in basal and HER2+ subtype, which correlate with higher metastatic capability and invasion. Involved molecules include claudins (CLDN), molecules involved in tight junctions of endothelial and epithelial cells, and *VCAM1* and *SELE*, which are involved in leukocyte trans endothelial migration. Claudins have been shown to

inhibit breast cancer growth and metastasis (Dong et al., 2023; Jin et al., 2024) and a claudin-low subtype of TNBC has been identified, exhibiting a more aggressive phenotype compared to claudin-rich ones (Yadav et al., 2022). Tight junctions and cell adhesion molecules were also found to be downregulated in the basal subtype compared to the HER2+ subtype and in our cohort of TNBC patients compared to healthy tissue.

Differential gene expression analysis identified several genes whose expression not only varies between different ethnic groups but also influences breast cancer overall survival (OS).

CXCL6 and CXCL8 are proinflammatory chemokines induced by IL-17, whose role in cancer has been associated with angiogenesis, invasion and metastasis and tumour immune suppression (Kolls & Lindén, 2004; Lee et al., 2008). IL-17 is associated with shorter disease free and overall survival and often upregulated in TNBC, contributing to neutrophil infiltration and tumour progression (Dawod et al., 2020; Khalid et al., 2023). This was confirmed in the TCGA cohort, where IL-17 signalling pathway is upregulated in the basal subtype compared to the other subtypes, and in our TNBC cohort, which showed an upregulation of genes involved in neutrophil extracellular trap formation compared to healthy tissue. *CXCL8* has higher expression in TNBC compared to healthy tissue, and is associated with a worse progression free survival, while *CXCL6* has been less studied in the context of breast cancer (R. Huang et al., 2023; Tang et al., 2025). In the TCGA basal breast cancer cohort, *CXCL6* was significantly more expressed in White patients compared to Black patients. Interestingly, a correlation with survival was only found in Black patients, where a high expression of this chemokine correlated with a better OS. In contraposition with literature data regarding *CXCL8*, its expression correlated with better prognosis in the HER2 + subtype, where its expression is lower in Black patients compared to the other groups. Generally, this proinflammatory pathway was associated with features of a more aggressive phenotype (basal subtype, Black ethnicity), however the expression of single chemokines was not.

The expression of *GSTM1* is higher in Black patients and correlates with worse overall survival in basal breast cancer. *GSTM1* and other glutathione transferases polymorphisms have been associated with cancer risk, with its null genotype

correlating with worse prognosis (Ambrosone et al., 2001; B. A. Jones et al., 2009). THBS3 is a glycoprotein involved in cell-cell interaction that has not been studied in the context of breast cancer, however it was found that elevated expression correlates with poor prognosis in renal clear cell carcinoma and colorectal cancer (J. Huang et al., 2025; G. Wang et al., 2023). Its expression has a similar effect on survival in basal breast cancer and is higher in Black patients compared to Asian patients.

Differentially expressed ion channels included chloride, sodium and potassium channels, ABC transporters and solute carriers. This class of proteins has been highlighted in recent years as players in various types of cancer, impacting not only prognosis, but also cell proliferation and metastasis.

Calcium sensitive chloride channel accessory 2 *CLCA2* has the highest expression in the HER2+ subtype. Its expression is induced by replicative senescence and oxidative stress through p53, and also promotes senescence, playing a tumour suppressor role. Its expression is reduced in prostate cancer, breast cancer and cervical squamous cell carcinoma (Tanikawa et al., 2012; X. Yang et al., 2021). In breast cancer, its downregulation can be caused by promoter hypermethylation, and correlates with a worse survival in TNBC patients of Black ethnicity (X. Li et al., 2004; Purrington et al., 2020). Across subtypes, it is consistently lower in Black patients and higher in Asian patients. Its effect on survival reported by the literature was confirmed, affecting significantly Black patients, at difference from patients belonging to other ethnicities.

While voltage gated sodium channels (VGSCs) have been demonstrated to play important roles in cancer, *SCN1A* has not been studied extensively, especially regarding its expression (H. Liu et al., 2024). Gender specific *SCN1A* polymorphisms were found in colorectal cancer patients treated with 5-fluorouracil and associated with recurrence (Benhaim et al., 2014). Its mutations are associated with familial migraine, epilepsy and intellectual disability (Ademuwagun et al., 2024; Fry et al., 2016). However, in studies related to these diseases, knockdown of *SCN1A* was associated with cancer pathways, PI3K-Akt signalling, MAPK signalling, and other ion channels (Shi et al., 2019). In oesophageal squamous carcinoma it is related to chemotherapy resistance, being a key gene in the protein network (Xie et al., 2022). Its expression is higher in ovarian cancer compared to normal tissue (R. Gao et al.,

2010). In the TCGA cohort, its high expression correlates with a worse OS. Among different ethnicities it shows marked different expression in all subtypes except HER2+, with Black patients consistently showing a higher expression, suggesting a role as an ethnic specific biomarker.

ABCA4 is upregulated in breast cancer cells resistant to chemotherapy, contributing to multidrug resistance (MDR) (Y. Liu et al., 2005). Polymorphisms in *ABCA4* have been associated with response to neoadjuvant chemotherapy (Hlaváč et al., 2020). In the TCGA cohort, it correlates with better prognosis in luminal B breast cancer and is downregulated in Black patients compared to White ones.

KCNK3 correlates with worse overall survival and it is upregulated in Asian patients with luminal B subtype tumours. Its expression has not been studied in breast cancer, however in lung adenocarcinoma it is downregulated and correlated to proliferation and glucose metabolism (G. Lin et al., 2022). In breast cancer, instead, SNPs have been associated with lymphedema following breast cancer surgery (Smoot et al., 2017).

KCNIP3 has been identified in the TCGA dataset as poor prognostic factor for breast cancer (Zhou et al., 2020). Its expression is the highest in basal breast cancer patients of Black ethnicity and correlates with a worse OS in breast cancer overall.

lncRNAs have been identified as novel and major players in breast cancer, and we identified several of them differentially expressed between ethnicities.

LINC00624 has been found to enhance liver cancer progression by disrupting an epigenetic repression complex (Z. Li et al., 2021). In HER2+ breast cancer, it inhibits the antitumour effect of HER2 targeted therapy by inhibiting Interferon I pathway activation and MHC I antigen presentation and CD8+ T cell infiltration (Q. Zhang et al., 2022). We confirmed its correlation with a worse overall survival in the TCGA cohort, where it is upregulated in HER2+ breast cancer Black patients. LIN*C02579, AC092979.1, AL031848.1,* and *AL590666.3* are not reported in literature regarding oncology, however their impact on survival and their differential expression across ethnicities highlights the effect of lncRNAs in breast cancer development and characteristics.

Triple negative breast tumours with high mutational burden (TMB) generally have a better prognosis. *FAT3* mutations correlate with lower expression, better prognosis and higher immune infiltrate (C. Gao et al., 2021). High levels of *FAT3* mutations have been associated with high TMB and better prognosis in oesophageal cancer as well (Guo et al., 2021). This correlates with the higher mutation rate of *FAT3* in Asian patients in the TCGA cohort, who have better survival.

In another study, it was reported that Black breast cancer patients were more likely to have *TP53* and *FAT3* mutations and less likely to have *PIK3CA, CDH1, DDR2* and *GATA3* mutations compared to other ethnicities (Van Alsten et al., 2025; Yao et al., 2025). Through mutational profiling of TCGA samples, we confirmed this lower PIK3CA mutation rate, although statistical significance was not reached.

ARID1A and *PIK3CA* co-mutations are associated with immune related pathways in luminal breast cancer, predicting response to immunotherapy (Anabel Sinberger et al., 2023). *ARID1A* mutations are also common in endocrine resistant ER+ tumours (X. Cheng et al., 2021; G. Xu et al., 2020). The higher mutation rate in Asian patients could therefore suggest a better response of this ethnic group to immunotherapy rather than endocrine therapy.

KMT2C and *KMT2D* are genes encoding histone methyltransferases frequently mutated in breast cancer, and associated with poor prognosis (Tinsley et al., 2024). In TNBC, their mutations drive brain metastasis (Seehawer et al., 2024). *KMT2C* is more frequently mutated in luminal A breast cancer, while *KMT2D* mutations more frequently occur in HER2+ breast cancer (L. Liu et al., 2014). In HER2+ breast cancer, *KMT2C* mutations are more common in Black patients, while *KMT2D* is more frequently mutated in Asian patients. *KMT2D* mutations were not found in Black patients, while White patients have a similar mutation rate of *KMT2D* and *KMT2C*.

These aspects of breast tumours can be attributed to both germline mutations and polymorphisms, intrinsic differences in healthy breast tissue and in different lifestyles, including use of hormonal contraceptives, age at first pregnancy, breast feeding and use of alcohol and tobacco, which vary in the different ethnic groups (Anstey et al., 2017; Thomson et al., 2011).

TNBC have a higher enrichment of stem cells compared to other breast cancer

subtypes (Djamgoz, 2025). Compared to luminal A subtype, TNBC was shown to have higher expression of genes regulating stem cells pluripotency. Wnt pathway is upregulated in TNBC as well, contributing to cancer development, interaction with TME and drug resistance (King et al., 2012; Merikhian et al., 2021; M. X. Sun et al., 2025). Our TNBC patients also show downregulation of antigen processing and presentation compared to healthy samples. TNBC can lose tumour antigen presentation and HLA1 is often downregulated and correlated with poorer prognosis (Pedersen et al., 2017; Y. Zheng et al., 2023). In fact, MHC II antigen processing pathways are upregulated in patients with no relapse (Forero-Torres et al., 2015). We did not perform a direct comparison between TCGA and GTEx datasets, as several articles have already addressed this matter, to prevent redundancy (Bou-Dargham et al., 2023; G. Zheng et al., 2021).

Apocrine breast carcinoma is a rarer subtype, with higher levels of cadherin binding, extracellular matrix signalling and a worse outcome compared to ductal carcinoma (Dellapasqua et al., 2013; N. Zhang et al., 2017; Zhu et al., 2024, p. 2024). In our cohort, infiltrating apocrine histotype correlates with higher tumour stage, and lower expression of genes involved in cell adhesion and motility.

Among key genes to analyse to better understand tumour biology and ethnical differences are ion channels and lncRNAs. Biological differences should be integrated with lifestyle and socioeconomic aspects, to identify specific factors impacting the different survival and response to treatment between ethnic groups even when presenting the same tumour subtype.

While we could not confirm specific differential expression of singular genes due to low number of patients and high transcriptomic variability, European patients showed enrichment in metabolic pathways compared to other ethnicities in both our FFPE cohort and TCGA cohort. Asian patients presented enrichment in stem cell pluripotency and GPI-anchor biosynthesis, which affect the tumour microenvironment and T-cell response (H. Wu et al., 2024). These two pathways are also linked as GPI-anchored proteins are more highly expressed in malignant and more dedifferentiated breast tumours (P. Zhao et al., 2012). *KCNIP3* was upregulated in Black or African patients in both our cohort and TCGA cohort.

PI3K-Akt signalling pathway, which promotes cell proliferation, survival and motility (H. Li et al., 2021), appears to be a key pathway in breast tumorigenesis in all ethnic groups.

We were able to identify several novel transcripts in triple negative breast cancer, however the use of FFPE samples, which yields lower RNA quality with possible fragmentation and chemical modifications (Pignatta et al., 2025; Y. Zhao et al., 2019), will require further examination and experimental validation of transcripts, with a focus on the novel lncRNA which were found to impact patients' prognosis. In the recent years, it has been established that lncRNAs play an important role in breast cancer pathogenesis, prognosis, and response to therapy, impacting a variety of pathways (Aranza-Martínez et al., 2021; Q. Yang et al., 2023), therefore their analysis and identification of novel transcripts hold significant potential for the improvement of breast cancer diagnosis and treatment.

This pilot study aimed to explore ethnic differences in breast cancer samples from patients belonging to diverse racial and ethnic backgrounds. Comparisons between single samples, while not statistically significant, were performed to explore potential trends and provide insights for studies on larger numbers of patients. Moreover, although mRNA expression analysis can be useful for patient stratification and prognosis, future research will also need to analyse protein expression and functionality.

6. Conclusion

In this thesis, several differentially expressed genes and pathways between patients of African, Asian, Middle Eastern and Caucasian ancestry were identified, both through *in silico* analysis on the TCGA cohort and RNAseq on FFPE samples.

Asian patients were characterized by not only higher rates of HER2+ breast cancer, but also by higher levels of ErbB pathway in other subtypes. They also presented higher levels of mutations in *FAT3* and *ARID1A*, which are both correlated with better response to immunotherapy.

Black patients were confirmed as the ethnic group most affected by the basal subtype, with a pathway enrichment characteristic of higher aggressiveness, like low expression of adhesion molecules and claudins which define the claudin-low subtype of TNBC, a more aggressive subtype. Black patients also presented a higher rate of *PIK3CA* mutation.

White patients presented a dysregulation of various metabolic pathways, both in the TCGA and FFPE cohorts.

Several ion channels and transporters were differentially expressed and influenced survival: *CLCA2*, whose expression is lower in Black patients as confirmed by literature and correlates with worse survival. *SCN1A*, instead has not been studied in breast cancer, however it presented a significant higher expression in Black patients, correlating with worse survival, similarly to ovarian and oesophageal carcinomas. *KCNIP3* was found to be upregulated in Black breast cancer patients in both TCGA and our FFPE cohort, correlating with a worse overall survival.

lncRNAs were found to be important genes in both ethnic differences and in TNBC.

TNBC presented a claudin-low phenotype compared to other subtypes and healthy tissue, and low expression of other adhesion molecules and tight junctions as well. Other pathways differentially expressed between TNBC and other subtypes or healthy tissue include stem cells pluripotency, antigen processing and presentation.

While the numbers of patients belonging to ethnic minorities were limited, we were able to identify several variations characterizing different ethnicities, which can

impact their survival and response to therapy.

The data reported in this thesis are extremely relevant although preliminary and additional analyses in a bigger cohort are warranted to confirm these observations.

7. References

Ademuwagun, I. A., Adam, Y., Rotimi, S. O., Syrbe, S., Radtke, M., Hentschel, J., Lemke, J. R., & Adebiyi, E. (2024). Exome sequencing in Nigerian children with early-onset epilepsy syndromes. *Epilepsia Open*, *10*(1), 222–232. https://doi.org/10.1002/epi4.13106

Adewunmi, O., Shen, Y., Zhang, X. H.-F., & Rosen, J. M. (2023). Targeted Inhibition of lncRNA Malat1 Alters the Tumor Immune Microenvironment in Preclinical Syngeneic Mouse Models of Triple-Negative Breast Cancer. *Cancer Immunology Research*, *11*(11), 1462–1479. https://doi.org/10.1158/2326-6066.CIR-23-0045

Ahmed, M. S. U., Lord, B. D., Adu Addai, B., Singhal, S. K., Gardner, K., Salam, A. B., Ghebremedhin, A., White, J., Mahmud, I., Martini, R., Bedi, D., Lin, H., Jones, J. D., Karanam, B., Dean-Colomb, W., Grizzle, W., Wang, H., Davis, M., & Yates, C. C. (2023). Immune Profile of Exosomes in African American Breast Cancer Patients Is Mediated by Kaiso/THBS1/CD47 Signaling. *Cancers*, *15*(8), 2282. https://doi.org/10.3390/cancers15082282

Al Abo, M., Hyslop, T., Qin, X., Owzar, K., George, D. J., Patierno, S. R., & Freedman, J. A. (2021). Differential alternative RNA splicing and transcription events between tumors from African American and White patients in The Cancer Genome Atlas. *Genomics*, *113*(3), 1234–1246. https://doi.org/10.1016/j.ygeno.2021.02.020

Albain, K. S., Barlow, W. E., Shak, S., Hortobagyi, G. N., Livingston, R. B., Yeh, I.-T., Ravdin, P., Bugarini, R., Baehner, F. L., Davidson, N. E., Sledge, G. W., Winer, E. P., Hudis, C., Ingle, J. N., Perez, E. A., Pritchard, K. I., Shepherd, L., Gralow, J. R., Yoshizawa, C., … Hayes, D. F. (2010). Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: A retrospective analysis of a randomised trial. *The Lancet Oncology*, *11*(1), 55–65. https://doi.org/10.1016/S1470-2045(09)70314-6

Alcantara, D., Leal, M. P., García-Bocanegra, I., & García-Martín, M. L. (2014). Molecular imaging of breast cancer: Present and future directions. *Frontiers in Chemistry*, *2*, 112. https://doi.org/10.3389/fchem.2014.00112

Al-Muftah, M., & Al-Ejeh, F. (2023). Cancer Incidence and Mortality Estimates in Arab Countries in 2018: A GLOBOCAN Data Analysis. *Cancer Epidemiology, Biomarkers & Prevention*, *32*(12), 1738. https://doi.org/10.1158/1055-9965.EPI-23-0520

Alsten, S. C. V., Dunn, M. R., Hamilton, A. M., Ivory, J. M., Gao, X., Kirk, E. L., Nsonwu-Farley, J. S., Carey, L. A., Abdou, Y., Reeder-Hayes, K. E., Roberson, M. L., Wheeler, S. B., Emerson, M. A., Hyslop, T., & Troester, M. A. (2024). Disparities in OncotypeDx Testing and Subsequent Chemotherapy Receipt by Geography and Socioeconomic Status. *Cancer Epidemiology, Biomarkers & Prevention : A Publication of the American Association for Cancer Research,*

Cosponsored by the American Society of Preventive Oncology, *33*(5), 654. https://doi.org/10.1158/1055-9965.EPI-23-1201

Amadori, D., Serra, P., Bravaccini, S., Farolfi, A., Puccetti, M., Carretta, E., Medri, L., Nanni, O., Tumedei, M. M., Kahima, J., & Masalu, N. (2014). Differences in biological features of breast cancer between Caucasian (Italian) and African (Tanzanian) populations. *Breast Cancer Research and Treatment*, *145*(1), 177–183. https://doi.org/10.1007/s10549-014-2903-0

Ambrosone, C. B., Sweeney, C., Coles, B. F., Thompson, P. A., McClure, G. Y., Korourian, S., Fares, M. Y., Stone, A., Kadlubar, F. F., & Hutchins, L. F. (2001). Polymorphisms in Glutathione S-Transferases (GSTM1 and GSTT1) and Survival after Treatment for Breast Cancer1. *Cancer Research*, *61*(19), 7130–7135.

Amirikia, K. C., Mills, P., Bush, J., & Newman, L. A. (2011). Higher Population-Based Incidence Rates of Triple-Negative Breast Cancer Among Young African American Women: Implications for Breast Cancer Screening Recommendations. *Cancer*, *117*(12), 2747–2753. https://doi.org/10.1002/cncr.25862

Amith, S. R., & Fliegel, L. (2013). Regulation of the Na+/H+ Exchanger (NHE1) in Breast Cancer Metastasis. *Cancer Research*, *73*(4), 1259–1264. https://doi.org/10.1158/0008-5472.CAN-12-4031

Anabel Sinberger, L., Zahavi, T., Sonnenblick, A., & Salmon-Divon, M. (2023). Coexistent ARID1A-PIK3CA mutations are associated with immune-related pathways in luminal breast cancer. *Scientific Reports*, *13*(1), 20911. https://doi.org/10.1038/s41598-023-48002-x

Andey, T., Attah, M. M., Akwaaba-Reynolds, N. A., Cheema, S., Parvin-Nejad, S., & Acquaah-Mensah, G. K. (2020). Enhanced immortalization, HUWE1 mutations and other biological drivers of breast invasive carcinoma in Black/African American patients. *Gene*, *763*, 100030. https://doi.org/10.1016/j.gene.2020.100030

Anstey, E. H., Shoemaker, M. L., Barrera, C. M., O'Neil, M. E., Verma, A. B., & Holman, D. M. (2017). Breastfeeding and Breast Cancer Risk Reduction: Implications for Black Mothers. *American Journal of Preventive Medicine*, *53*(3 Suppl 1), S40–S46. https://doi.org/10.1016/j.amepre.2017.04.024

Aranza-Martínez, A., Sánchez-Pérez, J., Brito-Elias, L., López-Camarillo, C., Cantú de León, D., Pérez-Plasencia, C., & López-Urrutia, E. (2021). Non-Coding RNAs Associated With Radioresistance in Triple-Negative Breast Cancer. *Frontiers in Oncology*, *11*, 752270. https://doi.org/10.3389/fonc.2021.752270

Aune, D., Chan, D. S. M., Greenwood, D. C., Vieira, A. R., Rosenblatt, D. A. N., Vieira, R., & Norat, T. (2012). Dietary fiber and breast cancer risk: A systematic review and meta-analysis of prospective studies. *Annals of Oncology*, *23*(6), 1394–1402. https://doi.org/10.1093/annonc/mdr589

Aydar, E., Yeo, S., Djamgoz, M., & Palmer, C. (2009). Abnormal expression, localization and interaction of canonical transient receptor potential ion

channels in human breast cancer cell lines and tissues: A potential target for breast cancer diagnosis and therapy. *Cancer Cell International*, *9*(1), 23. https://doi.org/10.1186/1475-2867-9-23

Ayubi, E., Lyus, R., Brhlikova, P., & Pollock, A. M. (2024). A critical appraisal of the quality of data submitted by sub-Saharan African cancer registries to GLOBOCAN 2020. *JRSM Open*, *15*(1), 20542704231217888. https://doi.org/10.1177/20542704231217888

Azahar, I. I., Sharudin, N. A., Noor Din, A. H. M., Che Has, A. T., Mohd Nafi, S. N., Jaafar, H., & Mokhtar, N. F. (2022). nNav1.5 expression is associated with glutamate level in breast cancer cells. *Biological Research*, *55*(1), 18. https://doi.org/10.1186/s40659-022-00387-1

Banikazemi, Z., Haji, H. A., Mohammadi, M., Taheripak, G., Iranifar, E., Poursadeghiyan, M., Moridikia, A., Rashidi, B., Taghizadeh, M., & Mirzaei, H. (2018). Diet and cancer prevention: Dietary compounds, dietary MicroRNAs, and dietary exosomes. *Journal of Cellular Biochemistry*, *119*(1), 185–196. https://doi.org/10.1002/jcb.26244

Barreto-Coelho, P., Cerbon, D., Schlumbrecht, M., Parra, C. M., Hurley, J., & George, S. H. L. (2019). Differences in breast cancer outcomes amongst Black US-born and Caribbean-born immigrants. *Breast Cancer Research and Treatment*, *178*(2), 433–440. https://doi.org/10.1007/s10549-019-05403-9

Bauer, K. R., Brown, M., Cress, R. D., Parise, C. A., & Caggiano, V. (2007). Descriptive analysis of estrogen receptor (ER)-negative, progesterone receptor (PR)-negative, and HER2-negative invasive breast cancer, the so-called triple-negative phenotype: A population-based study from the California cancer Registry. *Cancer*, *109*(9), 1721–1728. https://doi.org/10.1002/cncr.22618

Beltrán-Anaya, F. O., Romero-Córdoba, S., Rebollar-Vega, R., Arrieta, O., Bautista-Piña, V., Dominguez-Reyes, C., Villegas-Carlos, F., Tenorio-Torres, A., Alfaro-Riuz, L., Jiménez-Morales, S., Cedro-Tanda, A., Ríos-Romero, M., Reyes-Grajeda, J. P., Tagliabue, E., Iorio, M. V., & Hidalgo-Miranda, A. (2019). Expression of long non-coding RNA ENSG00000226738 (LncKLHDC7B) is enriched in the immunomodulatory triple-negative breast cancer subtype and its alteration promotes cell migration, invasion, and resistance to cell death. *Molecular Oncology*, *13*(4), 909–927. https://doi.org/10.1002/1878-0261.12446

Bénard, F., & Turcotte, É. (2005). Imaging in breast cancer: Single-photon computed tomography and positron-emission tomography. *Breast Cancer Research*, *7*(4), 153–162. https://doi.org/10.1186/bcr1201

Benhaim, L., Gerger, A., Bohanes, P., Paez, D., Wakatsuki, T., Yang, D., LaBonte, M. J., Ning, Y., El-Khoueiry, R., Loupakis, F., Zhang, W., Laurent-Puig, P., & Lenz, H. J. (2014). Gender-specific profiling in SCN1A polymorphisms and time-to-recurrence in patients with stage II/III colorectal cancer treated with adjuvant 5-fluoruracil chemotherapy. *The Pharmacogenomics Journal*, *14*(2), 135–141. https://doi.org/10.1038/tpj.2013.21

Benson, J. R. (2003). The TNM staging system and breast cancer. *The Lancet Oncology*, *4*(1), 56–60. https://doi.org/10.1016/S1470-2045(03)00961-6

Berg, W. A., Blume, J. D., Cormack, J. B., Mendelson, E. B., Lehrer, D., Böhm-Vélez, M., Pisano, E. D., Jong, R. A., Evans, W. P., Morton, M. J., Mahoney, M. C., Larsen, L. H., Barr, R. G., Farria, D. M., Marques, H. S., & Boparai, K. (2008). Combined Screening with Ultrasound and Mammography Compared to Mammography Alone in Women at Elevated Risk of Breast Cancer: Results of the First-Year Screen in ACRIN 6666. *JAMA : The Journal of the American Medical Association*, *299*(18), 2151–2163. https://doi.org/10.1001/jama.299.18.2151

Berg, W. A., Gutierrez, L., NessAiver, M. S., Carter, W. B., Bhargavan, M., Lewis, R. S., & Ioffe, O. B. (2004). Diagnostic Accuracy of Mammography, Clinical Examination, US, and MR Imaging in Preoperative Assessment of Breast Cancer. *Radiology*, *233*(3), 830–849. https://doi.org/10.1148/radiol.2333031484

Bickell, N. A., Wang, J. J., Oluwole, S., Schrag, D., Godfrey, H., Hiotis, K., Mendez, J., & Guth, A. A. (2006). Missed opportunities: Racial disparities in adjuvant breast cancer treatment. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, *24*(9), 1357–1362. https://doi.org/10.1200/JCO.2005.04.5799

Blamey, R. W., Ellis, I. O., Pinder, S. E., Lee, A. H. S., Macmillan, R. D., Morgan, D. a. L., Robertson, J. F. R., Mitchell, M. J., Ball, G. R., Haybittle, J. L., & Elston, C. W. (2007). Survival of invasive breast cancer according to the Nottingham Prognostic Index in cases diagnosed in 1990–1999. *European Journal of Cancer*, *43*(10), 1548–1555. https://doi.org/10.1016/j.ejca.2007.01.016

Blighe, K., & Lun, A. (2025). *PCAtools: PCAtools: Everything Principal Components Analysis*. Bioconductor. http://bioconductor.org/packages/PCAtools/

Botty van den Bruele, A., Weiss, A., Laronga, C., Boughey, J. C., Agnese, D., Amin, M., Barrio, A. V., Botty van den Bruele, A., Burke, E., Dirbas, F., Fayanju, O., Kantor, O., Kumar, S., Nguyen, T. T., Park, K. U., Plichta, J., Showalter, S. L., Singh, P., Tranakas, N., … Society of Surgical Oncology Breast Disease Site Work Group. (2024). Updated Guidelines on When to Consider Germline Testing for Patients with Breast Cancer. *Annals of Surgical Oncology*, *31*(9), 5479–5482. https://doi.org/10.1245/s10434-024-15638-6

Bou-Dargham, M. J., Sha, L., Sarker, D. B., Krakora-Compagno, M. Z., Chen, Z., Zhang, J., & Sang, Q.-X. A. (2023). TCGA RNA-Seq and Tumor-Infiltrating Lymphocyte Imaging Data Reveal Cold Tumor Signatures of Invasive Ductal Carcinomas and Estrogen Receptor-Positive Human Breast Tumors. *International Journal of Molecular Sciences*, *24*(11), 9355. https://doi.org/10.3390/ijms24119355

Brackenbury, W. J., Chioni, A.-M., Diss, J. K. J., & Djamgoz, M. B. A. (2007). The neonatal splice variant of Nav1.5 potentiates in vitro invasive behaviour of

MDA-MB-231 human breast cancer cells. *Breast Cancer Research and Treatment*, *101*(2), 149–160. https://doi.org/10.1007/s10549-006-9281-1

Bray, F., Laversanne, M., Sung, H., Ferlay, J., Siegel, R. L., Soerjomataram, I., & Jemal, A. (2024). Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, *74*(3), 229–263. https://doi.org/10.3322/caac.21834

Breuer, E.-K., Fukushiro-Lopes, D., Dalheim, A., Burnette, M., Zartman, J., Kaja, S., Wells, C., Campo, L., Curtis, K. J., Romero-Moreno, R., Littlepage, L. E., Niebur, G. L., Hoskins, K., Nishimura, M. I., & Gentile, S. (2019). Potassium channel activity controls breast cancer metastasis by affecting β-catenin signaling. *Cell Death & Disease*, *10*(3), 180. https://doi.org/10.1038/s41419-019-1429-0

Brevet, M., Haren, N., Sevestre, H., Merviel, P., & Ouadid-Ahidouch, H. (2009). DNA Methylation of Kv1.3 Potassium Channel Gene Promoter is Associated with Poorly Differentiated Breast Adenocarcinoma. *Cellular Physiology and Biochemistry*, *24*(1–2), 25–32. https://doi.org/10.1159/000227810

Burstein, M. D., Tsimelzon, A., Poage, G. M., Covington, K. R., Contreras, A., Fuqua, S. A. W., Savage, M. I., Osborne, C. K., Hilsenbeck, S. G., Chang, J. C., Mills, G. B., Lau, C. C., & Brown, P. H. (2015). Comprehensive Genomic Analysis Identifies Novel Subtypes and Targets of Triple-negative Breast Cancer. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research*, *21*(7), 1688–1698. https://doi.org/10.1158/1078-0432.CCR-14-0432

Cao, W., Qin, K., Li, F., & Chen, W. (2024). Socioeconomic inequalities in cancer incidence and mortality: An analysis of GLOBOCAN 2022. *Chinese Medical Journal*, *137*(12), 1407–1413. https://doi.org/10.1097/CM9.0000000000003140

Carey, L. A., Perou, C. M., Livasy, C. A., Dressler, L. G., Cowan, D., Conway, K., Karaca, G., Troester, M. A., Tse, C. K., 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*.

Carey, L., Winer, E., Viale, G., Cameron, D., & Gianni, L. (2010). Triple-negative breast cancer: Disease entity or title of convenience? *Nature Reviews Clinical Oncology*, *7*(12), 683–692. https://doi.org/10.1038/nrclinonc.2010.154

Castelló, A., Martín, M., Ruiz, A., Casas, A. M., Baena-Cañada, J. M., Lope, V., Antolín, S., Sánchez, P., Ramos, M., Antón, A., Muñoz, M., Bermejo, B., Juan-Ferré, A. D., Jara, C., Chacón, J. I., Jimeno, M. A., Rosado, P., Díaz, E., Guillem, V., … Researchers, E. (2015). Lower Breast Cancer Risk among Women following the World Cancer Research Fund and American Institute for Cancer Research Lifestyle Recommendations: EpiGEICAM Case-Control Study. *PLOS ONE*, *10*(5), e0126096.

https://doi.org/10.1371/journal.pone.0126096
Catarino, R., Ferreira, M. M., Rodrigues, H., Coelho, A., Nogal, A., Sousa, A., & Medeiros, R. (2008). Quantification of Free Circulating Tumor DNA as a Diagnostic Marker for Breast Cancer. *DNA and Cell Biology*, *27*(8), 415–421. https://doi.org/10.1089/dna.2008.0744
Chajès, V., Thiébaut, A. C. M., Rotival, M., Gauthier, E., Maillard, V., Boutron-Ruault, M.-C., Joulin, V., Lenoir, G. M., & Clavel-Chapelon, F. (2008). Association between Serum trans-Monounsaturated Fatty Acids and Breast Cancer Risk in the E3N-EPIC Study. *American Journal of Epidemiology*, *167*(11), 1312–1320. https://doi.org/10.1093/aje/kwn069
Chakraborty, A., Murthy, N. S., Chintamani, C., Bhatnagar, D., Mohil, R. S., Sharma, P. C., & Saxena, S. (2007). CYP17 gene polymorphism and its association with high-risk north Indian breast cancer patients. *Journal of Human Genetics*, *52*(2), Article 2. https://doi.org/10.1007/s10038-006-0095-0
Chan, D. S. M., Vieira, A. R., Aune, D., Bandera, E. V., Greenwood, D. C., McTiernan, A., Rosenblatt, D. N., Thune, I., Vieira, R., & Norat, T. (2014). Body mass index and survival in women with breast cancer—Systematic literature review and meta-analysis of 82 follow-up studies. *Annals of Oncology*, *25*(10), 1901–1914. https://doi.org/10.1093/annonc/mdu042
Chaudhary, L. N., Jawa, Z., Szabo, A., Visotcky, A., & Chitambar, C. R. (2016). Relevance of progesterone receptor immunohistochemical staining to Oncotype DX recurrence score. *Hematology/Oncology and Stem Cell Therapy*, *9*(2), 48. https://doi.org/10.1016/j.hemonc.2015.12.001
Chen, X., Li, J., Gray, W. H., Lehmann, B. D., Bauer, J. A., Shyr, Y., & Pietenpol, J. A. (2012). TNBCtype: A Subtyping Tool for Triple-Negative Breast Cancer. *Cancer Informatics*, *11*, 147–156. https://doi.org/10.4137/CIN.S9983
Cheng, T.-Y. D., Yao, S., Omilian, A. R., Khoury, T., Buas, M. F., Ondracek, R. P., Sribenja, S., Bshara, W., Hong, C.-C., Bandera, E. V., Davis, W., Higgins, M. J., & Ambrosone, C. B. (2019). FOXA1 protein expression in ER+ and ER− breast cancer in relation to parity and breastfeeding in Black and White women. *Cancer Epidemiology, Biomarkers & Prevention : A Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology*, *29*(2), 379. https://doi.org/10.1158/1055-9965.EPI-19-0787
Cheng, X., Zhao, J.-X., Dong, F., & Cao, X.-C. (2021). ARID1A Mutation in Metastatic Breast Cancer: A Potential Therapeutic Target. *Frontiers in Oncology*, *11*, 759577. https://doi.org/10.3389/fonc.2021.759577
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, J. S., Bernard, P. S., Perou, C. M., Ellis, M. J., & Nielsen, T. O. (2012). A 50-Gene Intrinsic Subtype Classifier for Prognosis and Prediction of Benefit from Adjuvant Tamoxifen. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research*, *18*(16), 4465–4472. https://doi.org/10.1158/1078-

0432.CCR-12-0286

Chua, A. V., Sheng, H., Liang, E., Gandhi, S., Kwan, M. L., Ergas, I. J., Roh, J. M., Laurent, C. A., Yan, L., Khoury, T., Ambrosone, C. B., Kushi, L. H., & Yao, S. (2024). Epidemiology of early vs late recurrence among women with early stage estrogen receptor–positive breast cancer in the Pathways Study. *JNCI: Journal of the National Cancer Institute*, *116*(10), 1621–1631. https://doi.org/10.1093/jnci/djae128

Chuang, E., Christos, P., Flam, A., McCarville, K., Forst, M., Shin, S., Vahdat, L., Swistel, A., Simmons, R., Osborne, M., Moore, A., Mazumdar, M., & Klein, P. (2012). Breast Cancer Subtypes in Asian-Americans Differ According to Asian Ethnic Group. *Journal of Immigrant and Minority Health*, *14*(5), 754–758. https://doi.org/10.1007/s10903-012-9577-7

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., Land, S. J., Lu, X., & Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly*, *6*(2), 80–92. https://doi.org/10.4161/fly.19695

Clarke, S., Chin, S. N., Dodds, L., George, S. H. L., & Badal, S. (2022). Racial disparities in breast cancer preclinical and clinical models. *Breast Cancer Research*, *24*(1), 56. https://doi.org/10.1186/s13058-022-01551-x

Colaprico, A., Silva, T. C., Olsen, C., Garofano, L., Cava, C., Garolini, D., Sabedot, T. S., Malta, T. M., Pagnotta, S. M., Castiglioni, I., Ceccarelli, M., Bontempi, G., & Noushmehr, H. (2016). TCGAbiolinks: An R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Research*, *44*(8), e71–e71. https://doi.org/10.1093/nar/gkv1507

Coletta, D. K., Hlusko, L. J., Scott, G. R., Garcia, L. A., Vachon, C. M., Norman, A. D., Funk, J. L., Shaibi, G. Q., Hernandez, V., De Filippis, E., & Mandarino, L. J. (2021). Association of EDARV370A with breast density and metabolic syndrome in Latinos. *PloS One*, *16*(10), e0258212. https://doi.org/10.1371/journal.pone.0258212

Cong, Y., Qiao, G., Zou, H., Lin, J., Wang, X., Li, X., Li, Y., & Zhu, S. (2015). Invasive cribriform carcinoma of the breast: A report of nine cases and a review of the literature. *Oncology Letters*, *9*(4), 1753–1758. https://doi.org/10.3892/ol.2015.2972

Connon, C. J., Kawasaki, S., Yamasaki, K., Quantock, A. J., & Kinoshita, S. (2004). *The quantification of hCLCA2 and colocalisation with integrin b4 in stratified human epithelia*.

Conway, K., Edmiston, S. N., Tse, C.-K., Bryant, C., Kuan, P. F., Hair, B. Y., Parrish, E. A., May, R., & Swift-Scanlan, T. (2015). Racial variation in breast tumor promoter methylation in the Carolina Breast Cancer Study. *Cancer Epidemiology, Biomarkers & Prevention: A Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology*, *24*(6), 921–930. https://doi.org/10.1158/1055-9965.EPI-14-1228

Cordeiro, P. G. (2008). Breast Reconstruction after Surgery for Breast Cancer. *New England Journal of Medicine*, *359*(15), 1590–1601. https://doi.org/10.1056/NEJMct0802899

CRAN: survival citation info. (n.d.). Retrieved 28 May 2025, from https://cran.r-project.org/web/packages/survival/citation.html

Cserni, G. (2020). Histological type and typing of breast carcinomas and the WHO classification changes over time. *Pathologica*, *112*(1), 25–41. https://doi.org/10.32074/1591-951X-1-20

Dandena, F. G., Teklewold, B. T., Darebo, T. D., & Suga, Y. D. (2024). Epidemiology and clinical characteristics of breast cancer in Ethiopia: A systematic review. *BMC Cancer*, *24*, 1102. https://doi.org/10.1186/s12885-024-12822-5

Davies, C., Pan, H., Godwin, J., Gray, R., Arriagada, R., Raina, V., Abraham, M., Alencar, V. H. M., Badran, A., Bonfill, X., Bradbury, J., Clarke, M., Collins, R., Davis, S. R., Delmestri, A., Forbes, J. F., Haddad, P., Hou, M.-F., Inbar, M., … Peto, R. (2013). Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years after diagnosis of oestrogen receptor-positive breast cancer: ATLAS, a randomised trial. *Lancet*, *381*(9869), 805–816. https://doi.org/10.1016/S0140-6736(12)61963-1

Davis, M., Tripathi, S., Hughley, R., He, Q., Bae, S., Karanam, B., Martini, R., Newman, L., Colomb, W., Grizzle, W., & Yates, C. (2018). AR negative triple negative or 'quadruple negative' breast cancers in African American women have an enriched basal and immune signature. *PloS One*, *13*(6), e0196909. https://doi.org/10.1371/journal.pone.0196909

Dawod, B., Liu, J., Gebremeskel, S., Yan, C., Sappong, A., Johnston, B., Hoskin, D. W., Marshall, J. S., & Wang, J. (2020). Myeloid-derived suppressor cell depletion therapy targets IL-17A-expressing mammary carcinomas. *Scientific Reports*, *10*(1), 13343. https://doi.org/10.1038/s41598-020-70231-7

Dee, E. C., Laversanne, M., Bhoo-Pathy, N., Ho, F. D. V., Feliciano, E. J. G., Eala, M. A. B., Ting, F. I. L., Ginsburg, O., Moraes, F. Y., Gyawali, B., Gomez, S. L., Ng, K., Wu, J. F., Jain, U., Jain, B., Columbres, R. C., Matsuda, T., Sangrajrang, S., Sinuraya, E. S., … Bray, F. (2025). Cancer incidence and mortality estimates in 2022 in southeast Asia: A comparative analysis. *The Lancet Oncology*, *26*(4), 516–528. https://doi.org/10.1016/S1470-2045(25)00017-8

Dellapasqua, S., Maisonneuve, P., Viale, G., Pruneri, G., Mazzarol, G., Ghisini, R., Mazza, M., Iorfida, M., Rotmensz, N., Veronesi, P., Luini, A., Goldhirsch, A., & Colleoni, M. (2013). Immunohistochemically Defined Subtypes and Outcome of Apocrine Breast Cancer. *Clinical Breast Cancer*, *13*(2), 95–102. https://doi.org/10.1016/j.clbc.2012.11.004

Dent, R., Trudeau, M., Pritchard, K. I., Hanna, W. M., Kahn, H. K., Sawka, C. A., Lickley, L. A., Rawlinson, E., Sun, P., & Narod, S. A. (2007). Triple-Negative Breast Cancer: Clinical Features and Patterns of Recurrence. *Clinical Cancer Research*, *13*(15), 4429–4434. https://doi.org/10.1158/1078-0432.CCR-06-

3045
Díaz, O., Rodríguez-Ruíz, A., & Sechopoulos, I. (2024). Artificial Intelligence for breast cancer detection: Technology, challenges, and prospects. *European Journal of Radiology*, *175*. https://doi.org/10.1016/j.ejrad.2024.111457
Ding, Y. C., Steele, L., Warden, C., Wilczynski, S., Mortimer, J., Yuan, Y., & Neuhausen, S. L. (2019). Molecular subtypes of triple-negative breast cancer in women of different race and ethnicity. *Oncotarget*, *10*(2), 198–208. https://doi.org/10.18632/oncotarget.26559
Djamgoz, M. B. A. (2025). Stemness of Cancer: A Study of Triple-negative Breast Cancer From a Neuroscience Perspective. *Stem Cell Reviews and Reports*, *21*(2), 337–350. https://doi.org/10.1007/s12015-024-10809-0
Dong, Y., Jin, Q., Sun, M., Qi, D., Qu, H., Wang, X., & Quan, C. (2023). CLDN6 inhibits breast cancer metastasis through WIP-dependent actin cytoskeleton-mediated autophagy. *Journal of Experimental & Clinical Cancer Research : CR*, *42*, 68. https://doi.org/10.1186/s13046-023-02644-x
Dookeran, K. A., Zhang, W., Stayner, L., & Argos, M. (2017). Associations of two-pore domain potassium channels and triple negative breast cancer subtype in The Cancer Genome Atlas: Systematic evaluation of gene expression and methylation. *BMC Research Notes*, *10*(1), 475. https://doi.org/10.1186/s13104-017-2777-4
Dorn, P. L., Al-Hallaq, H. A., Haq, F., Goldberg, M., Abe, H., Hasan, Y., & Chmura, S. J. (2013). A Prospective Study of the Utility of Magnetic Resonance Imaging in Determining Candidacy for Partial Breast Irradiation. *International Journal of Radiation Oncology, Biology, Physics*, *85*(3), 615–622. https://doi.org/10.1016/j.ijrobp.2012.06.014
Downie, B. R., Sánchez, A., Knötgen, H., Contreras-Jurado, C., Gymnopoulos, M., Weber, C., Stühmer, W., & Pardo, L. A. (2008). Eag1 Expression Interferes with Hypoxia Homeostasis and Induces Angiogenesis in Tumors. *The Journal of Biological Chemistry*, *283*(52), 36234–36240. https://doi.org/10.1074/jbc.M801830200
Dragonetti, M., Turco, C., Benedetti, A., Goeman, F., Forcato, M., Scalera, S., Allegretti, M., Esposito, G., Fazi, F., Blandino, G., Donzelli, S., & Fontemaggi, G. (2024). The lncRNAMALAT1-WTAP axis: A novel layer of EMT regulation in hypoxic triple-negative breast cancer. *Cell Death Discovery*, *10*, 276. https://doi.org/10.1038/s41420-024-02058-4
Dunn, M. R., Metwally, E. M., Vohra, S., Hyslop, T., Henderson, L. M., Reeder-Hayes, K., Thompson, C. A., Lafata, J. E., Troester, M. A., & Butler, E. N. (2024). Understanding mechanisms of racial disparities in breast cancer: An assessment of screening and regular care in the Carolina Breast Cancer Study. *Cancer Causes & Control*, *35*(5), 825–837. https://doi.org/10.1007/s10552-023-01833-5
Durrani, S., Al-Mushawa, F., Heena, H., Wani, T., & Al-Qahtani, A. (2021). Relationship of Oncotype Dx score with tumor grade, size, nodal status,

proliferative marker Ki67 and Nottingham Prognostic Index in early breast cancer tumors in Saudi Population. *Annals of Diagnostic Pathology*, *51*, 151674. https://doi.org/10.1016/j.anndiagpath.2020.151674

Early Breast Cancer Trialists' Collaborative Group (EBCTCG). (2010). Overview of the Randomized Trials of Radiotherapy in Ductal Carcinoma In Situ of the Breast. *Journal of the National Cancer Institute. Monographs*, *2010*(41), 162–177. https://doi.org/10.1093/jncimonographs/lgq039

Early Breast Cancer Trialists' Collaborative Group (EBCTCG). (2018). Long-term outcomes for neoadjuvant versus adjuvant chemotherapy in early breast cancer: Meta-analysis of individual patient data from ten randomised trials. *The Lancet. Oncology*, *19*(1), 27–39. https://doi.org/10.1016/S1470-2045(17)30777-5

EBCTCG (Early Breast Cancer Trialists' Collaborative Group). (2014). Effect of radiotherapy after mastectomy and axillary surgery on 10-year recurrence and 20-year breast cancer mortality: Meta-analysis of individual patient data for 8135 women in 22 randomised trials. *Lancet (London, England)*, *383*(9935), 2127–2135. https://doi.org/10.1016/S0140-6736(14)60488-8

Elzamly, S., Badri, N., Padilla, O., Dwivedi, A. K., Alvarado, L. A., Hamilton, M., Diab, N., Rock, C., Elfar, A., Teleb, M., Sanchez, L., & Nahleh, Z. (2018). Epithelial-Mesenchymal Transition Markers in Breast Cancer and Pathological Responseafter Neoadjuvant Chemotherapy. *Breast Cancer: Basic and Clinical Research*, *12*, 1178223418788074. https://doi.org/10.1177/1178223418788074

Erdogan, M. A., Yuca, E., Ashour, A., Gurbuz, N., Sencan, S., & Ozpolat, B. (2023). *SCN5A* promotes the growth and lung metastasis of triple-negative breast cancer through EF2-kinase signaling. *Life Sciences*, *313*, 121282. https://doi.org/10.1016/j.lfs.2022.121282

Espinal, A. C., Buas, M. F., Wang, D., Cheng, D. T.-Y., Sucheston-Campbell, L., Hu, Q., Yan, L., Payne-Ondracek, R., Cortes, E., Tang, L., Gong, Z., Zirpoli, G., Khoury, T., Yao, S., Omilian, A., Demissie, K., Bandera, E. V., Liu, S., Ambrosone, C. B., & Higgins, M. J. (2017). FOXA1 hypermethylation: Link between parity and ER-negative breast cancer in African American women? *Breast Cancer Research and Treatment*, *166*(2), 559–568. https://doi.org/10.1007/s10549-017-4418-y

Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Working Group. (2016). Recommendations from the EGAPP Working Group: Does the use of Oncotype DX tumor gene expression profiling to guide treatment decisions improve outcomes in patients with breast cancer? *Genetics in Medicine*, *18*(8), 770–779. https://doi.org/10.1038/gim.2015.173

Evans, A., Trimboli, R. M., Athanasiou, A., Balleyguier, C., Baltzer, P. A., Bick, U., Camps Herrero, J., Clauser, P., Colin, C., Cornford, E., Fallenberg, E. M., Fuchsjaeger, M. H., Gilbert, F. J., Helbich, T. H., Kinkel, K., Heywang-Köbrunner, S. H., Kuhl, C. K., Mann, R. M., Martincich, L., … Sardanelli, F. (2018). Breast ultrasound: Recommendations for information to women and

referring physicians by the European Society of Breast Imaging. *Insights into Imaging*, *9*(4), 449–461. https://doi.org/10.1007/s13244-018-0636-z

Evans, M. F., Vacek, P. M., Sprague, B. L., Stein, G. S., Stein, J. L., & Weaver, D. L. (2020). Microarray and RNA in situ hybridization assay for recurrence risk markers of breast carcinoma and ductal carcinoma in situ: Evidence supporting the use of diverse pathways panels. *Journal of Cellular Biochemistry*, *121*(2), 1736–1746. https://doi.org/10.1002/jcb.29409

Fang, M. (2023). Drug Treatments for HER2 Positive Breast Cancer. *Critical Debates in Humanities, Science and Global Justice*, *1*(1). https://criticaldebateshsgj.scholasticahq.com/article/89660-drug-treatments-for-her2-positive-breast-cancer

Faouzi, M., Hague, F., Geerts, D., Ay, A.-S., Potier-Cartereau, M., Ahidouch, A., & Ouadid-Ahidouch, H. (2016). Functional cooperation between KCa3.1 and TRPC1 channels in human breast cancer: Role in cell proliferation and patient prognosis. *Oncotarget*, *7*(24), 36419–36435. https://doi.org/10.18632/oncotarget.9261

Field, L. A., Love, B., Deyarmin, B., Hooke, J. A., Shriver, C. D., & Ellsworth, R. E. (2012). Identification of differentially expressed genes in breast tumors from African American compared with Caucasian women. *Cancer*, *118*(5), 1334–1344. https://doi.org/10.1002/cncr.26405

Filipits, M., Rudas, M., Jakesz, R., Dubsky, P., Fitzal, F., Singer, C. F., Dietze, O., Greil, R., Jelen, A., Sevelda, P., Freibauer, C., Müller, V., Jänicke, F., Schmidt, M., Kölbl, H., Rody, A., Kaufmann, M., Schroth, W., Brauch, H., … for the EP Investigators. (2011). A New Molecular Predictor of Distant Recurrence in ER-Positive, HER2-Negative Breast Cancer Adds Independent Information to Conventional Clinical Risk Factors. *Clinical Cancer Research*, *17*(18), 6012–6020. https://doi.org/10.1158/1078-0432.CCR-11-0926

Firatligil-Yildirir, B., Yalcin-Ozuysal, O., & Nonappa. (2023). Recent advances in lab-on-a-chip systems for breast cancer metastasis research. *Nanoscale Advances*, *5*(9), 2375–2393. https://doi.org/10.1039/D2NA00823H

Forero-Torres, A., Varley, K. E., Li, Y., Chen, D., Grizzle, W. E., Downs-Kelly, E., Burwell, T. C., Buchsbaum, D. J., Myers, R. M., & LoBuglio, A. F. (2015). MHC II antigen presentation pathway expression in triple-negative breast cancer. *Journal of Clinical Oncology*, *33*(15_suppl), 1066–1066. https://doi.org/10.1200/jco.2015.33.15_suppl.1066

Freitas, A. J. A. D., Causin, R. L., Calfa, S., Santana, I. V. V., Balie, M., Hirai, W., Laus, A. C., Araujo, H. S., Hidalgo Filho, C. M., De Oliveira Bombarda, F., Da Silva, V. D., Evangelista, A. F., De Pádua Souza, C., Reis, R. M., & Marques, M. M. C. (2025). Exploring the Nottingham classification: Assessing gene expression profiles in breast cancer patients and their association with outcomes. *Breast Cancer Research and Treatment*, *212*(2), 237–250. https://doi.org/10.1007/s10549-025-07718-2

Friedewald, S. M., Rafferty, E. A., Rose, S. L., Durand, M. A., Plecha, D. M.,

Greenberg, J. S., Hayes, M. K., Copit, D. S., Carlson, K. L., Cink, T. M., Barke, L. D., Greer, L. N., Miller, D. P., & Conant, E. F. (2014). Breast Cancer Screening Using Tomosynthesis in Combination With Digital Mammography. *JAMA*, *311*(24), 2499–2507. https://doi.org/10.1001/jama.2014.6095

Fry, A. E., Rees, E., Thompson, R., Mantripragada, K., Blake, P., Jones, G., Morgan, S., Jose, S., Mugalaasi, H., Archer, H., McCann, E., Clarke, A., Taylor, C., Davies, S., Gibbon, F., Te Water Naude, J., Hartley, L., Thomas, G., White, C., … Kerr, M. P. (2016). Pathogenic copy number variants and SCN1A mutations in patients with intellectual disability and childhood-onset epilepsy. *BMC Medical Genetics*, *17*, 34. https://doi.org/10.1186/s12881-016-0294-2

Frydenberg, H., Flote, V. G., Larsson, I. M., Barrett, E. S., Furberg, A.-S., Ursin, G., Wilsgaard, T., Ellison, P. T., McTiernan, A., Hjartåker, A., Jasienska, G., & Thune, I. (2015). Alcohol consumption, endogenous estrogen and mammographic density among premenopausal women. *Breast Cancer Research*, *17*(1), 103. https://doi.org/10.1186/s13058-015-0620-1

Fu, S., Ke, H., Yuan, H., Xu, H., Chen, W., & Zhao, L. (2024). Dual role of pregnancy in breast cancer risk. *General and Comparative Endocrinology*, *352*, 114501. https://doi.org/10.1016/j.ygcen.2024.114501

Fukushiro-Lopes, D. F., Hegel, A. D., Rao, V., Wyatt, D., Baker, A., Breuer, E.-K., Osipo, C., Zartman, J. J., Burnette, M., Kaja, S., Kouzoukas, D., Burris, S., Jones, W. K., & Gentile, S. (2017). Preclinical study of a Kv11.1 potassium channel activator as antineoplastic approach for breast cancer. *Oncotarget*, *9*(3), 3321–3337. https://doi.org/10.18632/oncotarget.22925

Galea, M. H., Blamey, R. W., Elston, C. E., & Ellis, I. O. (1992). The Nottingham prognostic index in primary breast cancer. *Breast Cancer Research and Treatment*, *22*(3), 207–219. https://doi.org/10.1007/BF01840834

Gallicchio, L., Calhoun, C., Riseberg, D., & Helzlsouer, K. (2017). Cardiovascular Health among Black and White Breast Cancer Patients Initiating Aromatase Inhibitor Therapy. *The Breast Journal*, *23*(2), 206–209. https://doi.org/10.1111/tbj.12709

Gao, C., Li, H., Liu, C., Xu, X., Zhuang, J., Zhou, C., Liu, L., Feng, F., & Sun, C. (2021). Tumor Mutation Burden and Immune Invasion Characteristics in Triple Negative Breast Cancer: Genome High-Throughput Data Analysis. *Frontiers in Immunology*, *12*, 650491. https://doi.org/10.3389/fimmu.2021.650491

Gao, R., Shen, Y., Cai, J., Lei, M., & Wang, Z. (2010). Expression of voltage-gated sodium channel α subunit in human ovarian cancer. *Oncology Reports*, *23*(5), 1293–1299. https://doi.org/10.3892/or_00000763

García-Becerra, R., Díaz, L., Camacho, J., Barrera, D., Ordaz-Rosado, D., Morales, A., Ortiz, C. S., Avila, E., Bargallo, E., Arrecillas, M., Halhali, A., & Larrea, F. (2010). Calcitriol inhibits *Ether-à go-go* potassium channel expression and cell proliferation in human breast cancer cells. *Experimental Cell Research*, *316*(3), 433–442. https://doi.org/10.1016/j.yexcr.2009.11.008

García-Hernández, A. P., Corona, D. N., Carlos-Reyes, Á., Sierra-Martínez, M., Acosta-Altamirano, G., Cisneros-Villanueva, M., Pérez-Navarro, Y., Ibarra-Sierra, E., Marchat, L. A., & López-Camarillo, C. (2024). The lncRNA AFAP1-AS1 is upregulated in metastatic triple-negative breast tumors and controls hypoxia-activated vasculogenic mimicry and angiogenesis. *BMC Cancer*, *24*, 1332. https://doi.org/10.1186/s12885-024-13019-6

García-Quiroz, J., García-Becerra, R., Santos-Martínez, N., Barrera, D., Ordaz-Rosado, D., Avila, E., Halhali, A., Villanueva, O., Ibarra-Sánchez, M. J., Esparza-López, J., Gamboa-Domínguez, A., Camacho, J., Larrea, F., & Díaz, L. (2014). In vivo dual targeting of the oncogenic Ether-à-go-go-1 potassium channel by calcitriol and astemizole results in enhanced antineoplastic effects in breast tumors. *BMC Cancer*, *14*(1), 745. https://doi.org/10.1186/1471-2407-14-745

García-Quiroz, J., González-González, M. E., Díaz, L., Ordaz-Rosado, D., Segovia-Mendoza, M., Prado-García, H., Larrea, F., & García-Becerra, R. (2019). Astemizole, an Inhibitor of Ether-à-Go-Go-1 Potassium Channel, Increases the Activity of the Tyrosine Kinase Inhibitor Gefitinib in Breast Cancer Cells. *Revista de Investigación Clínica*, *71*(3), 186–194. https://doi.org/10.24875/RIC.18002840

Gerardo Daniel Castro, J., & Castro, A. (2014). Alcohol drinking and mammary cancer: Pathogenesis and potential dietary preventive alternatives. *World Journal of Clinical Oncology*, *5*(4), 713–729. https://doi.org/10.5306/wjco.v5.i4.713

Gianni, L., Pienkowski, T., Im, Y.-H., Roman, L., Tseng, L.-M., Liu, M.-C., Lluch, A., Staroslawska, E., Haba-Rodriguez, J. de la, Im, S.-A., Pedrini, J. L., Poirier, B., Morandi, P., Semiglazov, V., Srimuninnimit, V., Bianchi, G., Szado, T., Ratnayake, J., Ross, G., & Valagussa, P. (2012). Efficacy and safety of neoadjuvant pertuzumab and trastuzumab in women with locally advanced, inflammatory, or early HER2-positive breast cancer (NeoSphere): A randomised multicentre, open-label, phase 2 trial. *The Lancet Oncology*, *13*(1), 25–32. https://doi.org/10.1016/S1470-2045(11)70336-9

Goldman, M. J., Craft, B., Hastie, M., Repečka, K., McDade, F., Kamath, A., Banerjee, A., Luo, Y., Rogers, D., Brooks, A. N., Zhu, J., & Haussler, D. (2020). Visualizing and interpreting cancer genomics data via the Xena platform. *Nature Biotechnology*, *38*(6), 675–678. https://doi.org/10.1038/s41587-020-0546-8

Gong, Y., Ji, P., Yang, Y.-S., Xie, S., Yu, T.-J., Xiao, Y., Jin, M.-L., Ma, D., Guo, L.-W., Pei, Y.-C., Chai, W.-J., Li, D.-Q., Bai, F., Bertucci, F., Hu, X., Jiang, Y.-Z., & Shao, Z.-M. (2021). Metabolic-Pathway-Based Subtyping of Triple-Negative Breast Cancer Reveals Potential Therapeutic Targets. *Cell Metabolism*, *33*(1), 51-64.e9. https://doi.org/10.1016/j.cmet.2020.10.012

Grootes, I., Wishart, G. C., & Pharoah, P. D. P. (2024). An updated PREDICT breast cancer prognostic model including the benefits and harms of radiotherapy. *Npj*

Breast Cancer, *10*(1), 6. https://doi.org/10.1038/s41523-024-00612-y

Guerrero, S., López-Cortés, A., Indacochea, A., García-Cárdenas, J. M., Zambrano, A. K., Cabrera-Andrade, A., Guevara-Ramírez, P., González, D. A., Leone, P. E., & Paz-y-Miño, C. (2018). Analysis of Racial/Ethnic Representation in Select Basic and Applied Cancer Research Studies. *Scientific Reports*, *8*(1), 13978. https://doi.org/10.1038/s41598-018-32264-x

Gunda, A., Eshwaraiah, M. S., Gangappa, K., Kaur, T., & Bakre, M. M. (2022). A comparative analysis of recurrence risk predictions in ER+/HER2− early breast cancer using NHS Nottingham Prognostic Index, PREDICT, and CanAssist Breast. *Breast Cancer Research and Treatment*, *196*(2), 299–310. https://doi.org/10.1007/s10549-022-06729-7

Guo, Z., Yan, X., Song, C., Wang, Q., Wang, Y., Liu, X.-P., Huang, J., Li, S., & Hu, W. (2021). FAT3 Mutation Is Associated With Tumor Mutation Burden and Poor Prognosis in Esophageal Cancer. *Frontiers in Oncology*, *11*, 603660. https://doi.org/10.3389/fonc.2021.603660

Guttman, M., Amit, I., Garber, M., French, C., Lin, M. F., Feldser, D., Huarte, M., Zuk, O., Carey, B. W., Cassady, J. P., Cabili, M. N., Jaenisch, R., Mikkelsen, T. S., Jacks, T., Hacohen, N., Bernstein, B. E., Kellis, M., Regev, A., Rinn, J. L., & Lander, E. S. (2009). Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature*, *458*(7235), 223–227. https://doi.org/10.1038/nature07672

Gyan, E., Owiredu, W. K. B. A., Fondjo, L. A., Jackson, A. M., Green, A. R., & Rahman, G. A. (2021). A review of the racial heterogeneity of breast cancer stem cells. *Gene*, *796–797*, 145805. https://doi.org/10.1016/j.gene.2021.145805

Habara, M., & Shimada, M. (2022). Estrogen receptor α revised: Expression, structure, function, and stability. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology*, *44*(12), e2200148. https://doi.org/10.1002/bies.202200148

Haddad, S. A., Lunetta, K. L., Ruiz-Narváez, E. A., Bensen, J. T., Hong, C.-C., Sucheston-Campbell, L. E., Yao, S., Bandera, E. V., Rosenberg, L., Haiman, C. A., Troester, M. A., Ambrosone, C. B., & Palmer, J. R. (2015). Hormone-Related Pathways and Risk of Breast Cancer Subtypes in African American Women. *Breast Cancer Research and Treatment*, *154*(1), 145. https://doi.org/10.1007/s10549-015-3594-x

Hadj-Ahmed, M., Ghali, R. M., Bouaziz, H., Habel, A., Stayoussef, M., Ayedi, M., Hachiche, M., Rahal, K., Yacoubi-Loueslati, B., & Almawi, W. Y. (2019). Transforming growth factor beta 1 polymorphisms and haplotypes associated with breast cancer susceptibility: A case-control study in Tunisian women. *Tumor Biology*, *41*(8), 1010428319869096. https://doi.org/10.1177/1010428319869096

Han, Y. J., Liu, S., Hardeman, A., Rajagopal, P. S., Mueller, J., Khramtsova, G., Sanni, A., Ajani, M., Clayton, W., Hurley, I. W., Yoshimatsu, T. F., Zheng, Y., Parker,

J., Perou, C. M., & Olopade, O. I. (2024). The VEGF-Hypoxia Signature Is Upregulated in Basal-like Breast Tumors from Women of African Ancestry and Associated with Poor Outcomes in Breast Cancer. *Clinical Cancer Research*, OF1–OF10. https://doi.org/10.1158/1078-0432.CCR-23-1526

Hashmi, A. A., Aijaz, S., Khan, S. M., Mahboob, R., Irfan, M., Zafar, N. I., Nisar, M., Siddiqui, M., Edhi, M. M., Faridi, N., & Khan, A. (2018). Prognostic parameters of luminal A and luminal B intrinsic breast cancer subtypes of Pakistani patients. *World Journal of Surgical Oncology*, *16*(1), 1. https://doi.org/10.1186/s12957-017-1299-9

Hawkins, T., Luban, S., & Kihara, D. (2006). Enhanced automated function prediction using distantly related sequences and contextual association by PFP. *Protein Science : A Publication of the Protein Society*, *15*(6), 1550–1556. https://doi.org/10.1110/ps.062153506

Haybittle, J. L., Blamey, R. W., Elston, C. W., Johnson, J., Doyle, P. J., Campbell, F. C., Nicholson, R. I., & Griffiths, K. (1982). A prognostic index in primary breast cancer. *British Journal of Cancer*, *45*(3), 361–366. https://doi.org/10.1038/bjc.1982.62

Hemminki, K., & Kaaks, R. (2024). The new GLOBOCAN methods paper demonstrates global inequalities in data sources. *International Journal of Cancer*, ijc.35300. https://doi.org/10.1002/ijc.35300

Henry, N. L., & Cannon-Albright, L. A. (2019). Breast cancer histologic subtypes show excess familial clustering. *Cancer*, *125*(18), 3131–3138. https://doi.org/10.1002/cncr.32198

Herschkowitz, J. I., Simin, K., Weigman, V. J., Mikaelian, I., Usary, J., Hu, Z., Rasmussen, K. E., Jones, L. P., Assefnia, S., Chandrasekharan, S., Backlund, M. G., Yin, Y., Khramtsov, A. I., Bastein, R., Quackenbush, J., Glazer, R. I., Brown, P. H., Green, J. E., Kopelovich, L., … Perou, C. M. (2007). Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. *Genome Biology*, *8*(5), R76. https://doi.org/10.1186/gb-2007-8-5-r76

Hill, D. A., Friend, S., Lomo, L., Wiggins, C., Barry, M., Prossnitz, E., & Royce, M. (2018). Breast cancer survival, survival disparities, and guideline-based treatment. *Breast Cancer Research and Treatment*, *170*(2), 405–414. https://doi.org/10.1007/s10549-018-4761-7

Hjerkind, K. V., Johansson, A. L. V., Trewin, C. B., Russnes, H. G., & Ursin, G. (2022). Incidence of breast cancer subtypes in immigrant and non-immigrant women in Norway. *Breast Cancer Research : BCR*, *24*, 4. https://doi.org/10.1186/s13058-021-01498-5

Hlaváč, V., Václavíková, R., Brynychová, V., Koževnikovová, R., Kopečková, K., Vrána, D., Gatěk, J., & Souček, P. (2020). Role of Genetic Variation in ABC Transporters in Breast Cancer Prognosis and Therapy Response. *International Journal of Molecular Sciences*, *21*(24), 9556. https://doi.org/10.3390/ijms21249556

Holt, S., Bertelli, G., Humphreys, I., Valentine, W., Durrani, S., Pudney, D., Rolles, M., Moe, M., Khawaja, S., Sharaiha, Y., Brinkworth, E., Whelan, S., Jones, S., Bennett, H., & Phillips, C. J. (2013). A decision impact, decision conflict and economic assessment of routine Oncotype DX testing of 146 women with node-negative or pNImi, ER-positive breast cancer in the UK. *British Journal of Cancer*, *108*(11), 2250–2258. https://doi.org/10.1038/bjc.2013.207

Howell, A., Anderson, A. S., Clarke, R. B., Duffy, S. W., Evans, D. G., Garcia-Closas, M., Gescher, A. J., Key, T. J., Saxton, J. M., & Harvie, M. N. (2014). Risk determination and prevention of breast cancer. *Breast Cancer Research*, *16*(5), 446. https://doi.org/10.1186/s13058-014-0446-2

Hu, Y., Wang, C., Liang, H., Li, J., & Yang, Q. (2024). The treatment landscape of triple-negative breast cancer. *Medical Oncology*, *41*(10), 236. https://doi.org/10.1007/s12032-024-02456-9

Huang, J., Zhang, T., Li, H., Li, Z., Yin, S., Liu, Y., Zhang, C., Qiu, Y., & Yu, H. (2025). Ubiquitination-Dependent LLGL2 Degradation Drives Colorectal Cancer Progression via THBS3 mRNA Stabilization. *Advanced Science*, *n/a*(n/a), e01656. https://doi.org/10.1002/advs.202501656

Huang, R. S., Gamazon, E. R., Ziliak, D., Wen, Y., Im, H. K., Zhang, W., Wing, C., Duan, S., Bleibel, W. K., Cox, N. J., & Dolan, M. E. (2011). Population differences in microRNA expression and biological implications. *RNA Biology*, *8*(4), 692–701. https://doi.org/10.4161/rna.8.4.16029

Huang, R., Wang, Z., Hong, J., Wu, J., Huang, O., He, J., Chen, W., Li, Y., Chen, X., & Shen, K. (2023). Targeting cancer-associated adipocyte-derived CXCL8 inhibits triple-negative breast cancer progression and enhances the efficacy of anti-PD-1 immunotherapy. *Cell Death & Disease*, *14*(10), 703. https://doi.org/10.1038/s41419-023-06230-z

Hughes, T., Harper, A., Gupta, S., Frazier, A. L., Van Der Graaf, W. T. A., Moreno, F., Joseph, A., & Fidler-Benaoudia, M. M. (2024). The current and future global burden of cancer among adolescents and young adults: A population-based study. *The Lancet Oncology*, *25*(12), 1614–1624. https://doi.org/10.1016/S1470-2045(24)00523-0

Huo, D., Hu, H., Rhie, S. K., Gamazon, E. R., Cherniack, A. D., Liu, J., Yoshimatsu, T. F., Pitt, J. J., Hoadley, K. A., Troester, M., Ru, Y., Lichtenberg, T., Sturtz, L. A., Shelley, C. S., Benz, C. C., Mills, G. B., Laird, P. W., Shriver, C. D., Perou, C. M., & Olopade, O. I. (2017). Comparison of Breast Cancer Molecular Features and Survival by African and European Ancestry in The Cancer Genome Atlas. *JAMA Oncology*, *3*(12), 1654–1662. https://doi.org/10.1001/jamaoncol.2017.0595

Husby, A., Wohlfahrt, J., Øyen, N., & Melbye, M. (2018). Pregnancy duration and breast cancer risk. *Nature Communications*, *9*, 4255. https://doi.org/10.1038/s41467-018-06748-3

Hwang, J.-Y., Han, B.-K., Ko, E. Y., Shin, J. H., Hahn, S. Y., & Nam, M. Y. (2015). Screening Ultrasound in Women with Negative Mammography: Outcome

Analysis. *Yonsei Medical Journal*, *56*(5), 1352–1358. https://doi.org/10.3349/ymj.2015.56.5.1352

Ilozumba, M. N., Yaghjyan, L., Datta, S., Zhao, J., Hong, C.-C., Lunetta, K. L., Zirpoli, G., Bandera, E. V., Palmer, J. R., Yao, S., Ambrosone, C. B., & Cheng, T.-Y. D. (2023). mTOR pathway candidate genes and obesity interaction on breast cancer risk in Black women from the Women's Circle of Health Study. *Cancer Causes & Control : CCC*, *34*(5), 431. https://doi.org/10.1007/s10552-022-01657-9

Iorio, J., Duranti, C., Lastraioli, E., Iorio, J., Duranti, C., & Lastraioli, E. (2019). Ion Channels and Transporters as Cancer Biomarkers and Targets for Diagnostics with Antibodies. In *Translational Research in Cancer*. IntechOpen. https://doi.org/10.5772/intechopen.90401

Iorio, J., Meattini, I., Bianchi, S., Bernini, M., Maragna, V., Dominici, L., Casella, D., Vezzosi, V., Orzalesi, L., Nori, J., Livi, L., Arcangeli, A., & Lastraioli, E. (2018). hERG1 channel expression associates with molecular subtypes and prognosis in breast cancer. *Cancer Cell International*, *18*, 93. https://doi.org/10.1186/s12935-018-0592-1

Jacobs, A. T., Martinez Castaneda-Cruz, D., Rose, M. M., & Connelly, L. (2022). Targeted therapy for breast cancer: An overview of drug classes and outcomes. *Biochemical Pharmacology*, *204*, 115209. https://doi.org/10.1016/j.bcp.2022.115209

Jafari, S. H., Saadatpour, Z., Salmaninejad, A., Momeni, F., Mokhtari, M., Nahand, J. S., Rahmati, M., Mirzaei, H., & Kianmehr, M. (2018). Breast cancer diagnosis: Imaging techniques and biochemical markers. *Journal of Cellular Physiology*, *233*(7), 5200–5213. https://doi.org/10.1002/jcp.26379

Jang, S.-H., Kang, K.-S., Ryu, P.-D., & Lee, S.-Y. (2009). Kv1.3 voltage-gated K^+ channel subunit as a potential diagnostic marker and therapeutic target for breast cancer. *BMB Reports*, *42*(8), 535–539. https://doi.org/10.5483/BMBRep.2009.42.8.535

Jatoi, I., & Proschan, M. A. (2005). Randomized Trials of Breast-Conserving Therapy Versus Mastectomy for Primary Breast Cancer: A Pooled Analysis of Updated Results. *American Journal of Clinical Oncology*, *28*(3), 289. https://doi.org/10.1097/01.coc.0000156922.58631.d7

Jeon, Y. W., Kang, S. H., Park, M. H., Lim, W., Cho, S. H., & Suh, Y. J. (2015). Relationship between body mass index and the expression of hormone receptors or human epidermal growth factor receptor 2 with respect to breast cancer survival. *BMC Cancer*, *15*(1), 865. https://doi.org/10.1186/s12885-015-1879-4

Jin, Q., Qi, D., Zhang, M., Qu, H., Dong, Y., Sun, M., & Quan, C. (2024). CLDN6 inhibits breast cancer growth and metastasis through SREBP1-mediated RAS palmitoylation. *Cellular & Molecular Biology Letters*, *29*, 112. https://doi.org/10.1186/s11658-024-00629-y

John, E. M., McGuire, V., Kurian, A. W., Koo, J., Shariff-Marco, S., Gomez, S. L.,

Cheng, I., Keegan, T. H. M., Kwan, M. L., Bernstein, L., Vigen, C., & Wu, A. H. (2021). Racial/Ethnic Disparities in Survival after Breast Cancer Diagnosis by Estrogen and Progesterone Receptor Status: A Pooled Analysis. *Cancer Epidemiology, Biomarkers & Prevention*, *30*(2), 351–363. https://doi.org/10.1158/1055-9965.EPI-20-1291

Jones, B. A., Christensen, A. R., Wise, J. P., & Yu, H. (2009). Glutathione *S*-transferase polymorphisms and survival in African-American and White colorectal cancer patients. *Cancer Epidemiology*, *33*(3), 249–256. https://doi.org/10.1016/j.canep.2009.08.004

Jones, T., Trivedi, M., Jiang, X., Silverman, T., Underhill, M., Chung, W., Kukafka, R., & Crew, K. (2021). Racial and Ethnic Differences in BRCA1/2 and Multigene Panel Testing Among Young Breast Cancer Patients. *Journal of Cancer Education*, *36*(3), 463–469. https://doi.org/10.1007/s13187-019-01646-8

Jung, H. J., Park, J.-Y., Jeon, H.-S., & Kwon, T.-H. (2011). Aquaporin-5: A Marker Protein for Proliferation and Migration of Human Breast Cancer Cells. *PLOS ONE*, *6*(12), e28492. https://doi.org/10.1371/journal.pone.0028492

Kamal, M., Nafie, E. H. O., Elsers, S., Alanwar, S., Ibrahim, R., Farag, F., Mlees, M., Simões, B. M., Spence, K., Santiago-Gómez, A., Salem, M. L., & Clarke, R. B. (2018). Ethnicity influences breast cancer stem cells' drug resistance. *The Breast Journal*, *24*(4), 701–703. https://doi.org/10.1111/tbj.13021

Kan, Z., Ding, Y., Kim, J., Jung, H. H., Chung, W., Lal, S., Cho, S., Fernandez-Banet, J., Lee, S. K., Kim, S. W., Lee, J. E., Choi, Y.-L., Deng, S., Kim, J.-Y., Ahn, J. S., Sha, Y., Mu, X. J., Nam, J.-Y., Im, Y.-H., … Park, Y. H. (2018). Multi-omics profiling of younger Asian breast cancers reveals distinctive molecular signatures. *Nature Communications*, *9*(1), 1725. https://doi.org/10.1038/s41467-018-04129-4

Keller, L. M. M., Sopka, D. M., Li, T., Klayton, T., Li, J., Anderson, P. R., Bleicher, R. J., Sigurdson, E. R., & Freedman, G. M. (2012). Five-year Results of Whole Breast Intensity Modulated Radiation Therapy for the Treatment of Early Stage Breast Cancer: The Fox Chase Cancer Center Experience. *International Journal of Radiation Oncology, Biology, Physics*, *84*(4), 881–887. https://doi.org/10.1016/j.ijrobp.2012.01.069

Khaitan, D., Sankpal, U. T., Weksler, B., Meister, E. A., Romero, I. A., Couraud, P.-O., & Ningaraj, N. S. (2009). Role of KCNMA1gene in breast cancer invasion and metastasis to brain. *BMC Cancer*, *9*(1), 258. https://doi.org/10.1186/1471-2407-9-258

Khalid, F., Takagi, K., Sato, A., Yamaguchi, M., Guestini, F., Miki, Y., Miyashita, M., Hirakawa, H., Ohi, Y., Rai, Y., Sagara, Y., Sasano, H., & Suzuki, T. (2023). Interleukin (IL)-17A in triple-negative breast cancer: A potent prognostic factor associated with intratumoral neutrophil infiltration. *Breast Cancer*, *30*(5), 748–757. https://doi.org/10.1007/s12282-023-01467-0

Khalil, A. M., Guttman, M., Huarte, M., Garber, M., Raj, A., Rivea Morales, D.,

Thomas, K., Presser, A., Bernstein, B. E., van Oudenaarden, A., Regev, A., Lander, E. S., & Rinn, J. L. (2009). Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(28), 11667–11672. https://doi.org/10.1073/pnas.0904715106

Kim, J., Harper, A., McCormack, V., Sung, H., Houssami, N., Morgan, E., Mutebi, M., Garvey, G., Soerjomataram, I., & Fidler-Benaoudia, M. M. (2025). Global patterns and trends in breast cancer incidence and mortality across 185 countries. *Nature Medicine*, *31*(4), 1154–1162. https://doi.org/10.1038/s41591-025-03502-3

King, T. D., Suto, M. J., & Li, Y. (2012). The Wnt/β-catenin signaling pathway: A potential therapeutic target in the treatment of triple negative breast cancer. *Journal of Cellular Biochemistry*, *113*(1), 13–18. https://doi.org/10.1002/jcb.23350

Ko, J.-H., Ko, E. A., Gu, W., Lim, I., Bang, H., & Zhou, T. (2013). Expression profiling of ion channel genes predicts clinical outcome in breast cancer. *Molecular Cancer*, *12*(1), 106. https://doi.org/10.1186/1476-4598-12-106

Koboldt, D. C., Fulton, R. S., McLellan, M. D., Schmidt, H., Kalicki-Veizer, J., McMichael, J. F., Fulton, L. L., Dooling, D. J., Ding, L., Mardis, E. R., Wilson, R. K., Ally, A., Balasundaram, M., Butterfield, Y. S. N., Carlsen, R., Carter, C., Chu, A., Chuah, E., Chun, H.-J. E., … MD Anderson Cancer Center. (2012). Comprehensive molecular portraits of human breast tumours. *Nature*, *490*(7418), 61–70. https://doi.org/10.1038/nature11412

Kolls, J. K., & Lindén, A. (2004). Interleukin-17 Family Members and Inflammation. *Immunity*, *21*(4), 467–476. https://doi.org/10.1016/j.immuni.2004.08.018

Kong, L., Zhang, Y., Ye, Z.-Q., Liu, X.-Q., Zhao, S.-Q., Wei, L., & Gao, G. (2007). CPC: Assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Research*, *35*(Web Server issue), W345–W349. https://doi.org/10.1093/nar/gkm391

Koru-Sengul, T., Santander, A. M., Miao, F., Sanchez, L. G., Jorda, M., Glück, S., Ince, T. A., Nadji, M., Chen, Z., Penichet, M. L., Cleary, M. P., & Torroella-Kouri, M. (2016). Breast cancers from black women exhibit higher numbers of immunosuppressive macrophages with proliferative activity and of crown-like structures associated with lower survival compared to non-black Latinas and Caucasians. *Breast Cancer Research and Treatment*, *158*(1), 113–126. https://doi.org/10.1007/s10549-016-3847-3

Kreike, B., van Kouwenhove, M., Horlings, H., Weigelt, B., Peterse, H., Bartelink, H., & van de Vijver, M. J. (2007). Gene expression profiling and histopathological characterization of triple-negative/basal-like breast carcinomas. *Breast Cancer Research : BCR*, *9*(5), R65. https://doi.org/10.1186/bcr1771

Kruk, J. (2015). Lifestyle Components and Primary Breast Cancer Prevention. *Asian Pacific Journal of Cancer Prevention*, *15*(24), 10543–10555.

https://doi.org/10.7314/APJCP.2014.15.24.10543
Kuehn, T., Bauerfeind, I., Fehm, T., Fleige, B., Hausschild, M., Helms, G., Lebeau, A., Liedtke, C., Minckwitz, G. von, Nekljudova, V., Schmatloch, S., Schrenk, P., Staebler, A., & Untch, M. (2013). Sentinel-lymph-node biopsy in patients with breast cancer before and after neoadjuvant chemotherapy (SENTINA): A prospective, multicentre cohort study. *The Lancet Oncology*, *14*(7), 609–618. https://doi.org/10.1016/S1470-2045(13)70166-9
Kwapisz, D. (2020). Pembrolizumab and atezolizumab in triple-negative breast cancer. *Cancer Immunology, Immunotherapy : CII*, *70*(3), 607–617. https://doi.org/10.1007/s00262-020-02736-z
Kwon, W. S. (2009). G-T haplotype (2677G>T/A and 3435C>T) of ABCB1 gene polymorphisms is associated with ethnic differences to paclitaxel sensitivity in cancer cells with different gene expression pattern. *Cancer Letters*.
Lahart, I. M., Metsios, G. S., Nevill, A. M., & Carmichael, A. R. (2015). Physical activity, risk of death and recurrence in breast cancer survivors: A systematic review and meta-analysis of epidemiological studies. *Acta Oncologica*, *54*(5), 635–654. https://doi.org/10.3109/0284186X.2014.998275
Lai, Q., Wang, T., Guo, Q., Zhang, Y., Wang, Y., Yuan, L., Ling, R., He, Y., & Wang, W. (2014). Positive correlation between the expression of hEag1 and HIF-1α in breast cancers: An observational study. *BMJ Open*, *4*(5), e005049. https://doi.org/10.1136/bmjopen-2014-005049
Lalloo, F., & Evans, D. G. (2012). Familial Breast Cancer. *Clinical Genetics*, *82*(2), 105–114. https://doi.org/10.1111/j.1399-0004.2012.01859.x
Lansu, K., & Gentile, S. (2013). Potassium channel activation inhibits proliferation of breast cancer cells by activating a senescence program. *Cell Death & Disease*, *4*(6), e652–e652. https://doi.org/10.1038/cddis.2013.174
Lastraioli, E., Iorio, J., & Arcangeli, A. (2015). Ion channel expression as promising cancer biomarker. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, *1848*(10), 2685–2702. https://doi.org/10.1016/j.bbamem.2014.12.016
Lee, J. W., Wang, P., Kattah, M. G., Youssef, S., Steinman, L., DeFea, K., & Straus, D. S. (2008). Differential Regulation of Chemokines by IL-17 in Colonic Epithelial Cells1. *The Journal of Immunology*, *181*(9), 6536–6545. https://doi.org/10.4049/jimmunol.181.9.6536
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. *The Journal of Clinical Investigation*, *121*(7), 2750–2767. https://doi.org/10.1172/JCI45014
Lehmann, B. D., Jovanović, B., Chen, X., Estrada, M. V., Johnson, K. N., Shyr, Y., Moses, H. L., Sanders, M. E., & Pietenpol, J. A. (2016). Refinement of Triple-Negative Breast Cancer Molecular Subtypes: Implications for Neoadjuvant Chemotherapy Selection. *PLOS ONE*, *11*(6), e0157368. https://doi.org/10.1371/journal.pone.0157368

Lei, B., Jiang, X., & Saxena, A. (2023). TCGA Expression Analyses of 10 Carcinoma Types Reveal Clinically Significant Racial Differences. *Cancers*, *15*(10), 2695. https://doi.org/10.3390/cancers15102695

Leonard-Murali, S., Nathanson, S. D., Springer, K., Baker, P., & Susick, L. (2023). Early Breast Cancer Survival of Black and White American Women with Equal Diagnostic and Therapeutic Management. *European Journal of Surgical Oncology : The Journal of the European Society of Surgical Oncology and the British Association of Surgical Oncology*, *49*(3), 583–588. https://doi.org/10.1016/j.ejso.2022.11.101

Li, C. I., Malone, K. E., & Daling, J. R. (2003). Differences in Breast Cancer Stage, Treatment, and Survival by Race and Ethnicity. *Archives of Internal Medicine*, *163*(1), 49. https://doi.org/10.1001/archinte.163.1.49

Li, H., Prever, L., Hirsch, E., & Gulluni, F. (2021). Targeting PI3K/AKT/mTOR Signaling Pathway in Breast Cancer. *Cancers*, *13*(14), 3517. https://doi.org/10.3390/cancers13143517

Li, N., Deng, Y., Zhou, L., Tian, T., Yang, S., Wu, Y., Zheng, Y., Zhai, Z., Hao, Q., Song, D., Zhang, D., Kang, H., & Dai, Z. (2019). Global burden of breast cancer and attributable risk factors in 195 countries and territories, from 1990 to 2017: Results from the Global Burden of Disease Study 2017. *Journal of Hematology & Oncology*, *12*, 140. https://doi.org/10.1186/s13045-019-0828-0

Li, R. Q., Yan, L., Zhang, L., Ma, H. X., Wang, H. W., Bu, P., Xi, Y. F., & Lian, J. (2024). Genomic characterization reveals distinct mutational landscapes and therapeutic implications between different molecular subtypes of triple-negative breast cancer. *Scientific Reports*, *14*(1), 12386. https://doi.org/10.1038/s41598-024-62991-3

Li, X., Cowell, J. K., & Sossey-Alaoui, K. (2004). CLCA2 tumour suppressor gene in 1p31 is epigenetically regulated in breast cancer. *Oncogene*, *23*(7), 1474–1480. https://doi.org/10.1038/sj.onc.1207249

Li, Y., Du, F., Zhu, W., & Xu, B. (2017). Neuroendocrine carcinoma of the breast: A review of 126 cases in China. *Chinese Journal of Cancer*, *36*(1), 45. https://doi.org/10.1186/s40880-017-0211-x

Li, Y., Wang, J.-P., Santen, R. J., Kim, T.-H., Park, H., Fan, P., & Yue, W. (2010). Estrogen Stimulation of Cell Migration Involves Multiple Signaling Pathway Interactions. *Endocrinology*, *151*(11), 5146–5156. https://doi.org/10.1210/en.2009-1506

Li, Z., Lu, X., Liu, Y., Zhao, J., Ma, S., Yin, H., Huang, S., Zhao, Y., & He, X. (2021). Gain of LINC00624 Enhances Liver Cancer Progression by Disrupting the Histone Deacetylase 6/Tripartite Motif Containing 28/Zinc Finger Protein 354C Corepressor Complex. *Hepatology*, *73*(5), 1764. https://doi.org/10.1002/hep.31530

Li, Z., Wei, H., Li, S., Wu, P., & Mao, X. (2022). The Role of Progesterone Receptors in Breast Cancer . *Drug Design, Development and Therapy*, *16*, 305–314. https://doi.org/10.2147/DDDT.S336643

Liao, L. (2024). Inequality in breast cancer: Global statistics from 2022 to 2050. *The Breast : Official Journal of the European Society of Mastology*, *79*, 103851. https://doi.org/10.1016/j.breast.2024.103851

Lin, A., Li, C., Xing, Z., Hu, Q., Liang, K., Han, L., Wang, C., Hawke, D. H., Wang, S., Zhang, Y., Wei, Y., Ma, G., Park, P. K., Zhou, J., Zhou, Y., Hu, Z., Zhou, Y., Marks, J. R., Liang, H., … Yang, L. (2016). The LINK-A lncRNA Activates Normoxic HIF1α Signaling in Triple-negative Breast Cancer. *Nature Cell Biology*, *18*(2), 213–224. https://doi.org/10.1038/ncb3295

Lin, G., Lin, L., Lin, H., Chen, W., Chen, L., Chen, X., Chen, S., Lin, Q., Xu, Y., & Zeng, Y. (2022). KCNK3 inhibits proliferation and glucose metabolism of lung adenocarcinoma via activation of AMPK-TXNIP pathway. *Cell Death Discovery*, *8*, 360. https://doi.org/10.1038/s41420-022-01152-9

Liu, G.-X., Yu, Y.-C., He, X.-P., Ren, S.-N., Fang, X.-D., Liu, F., & He, Y. (2015). *Expression of eag1 channel associated with the aggressive clinicopathological features and subtype of breast cancer*.

Liu, H., Weng, J., Huang, C. L.-H., & Jackson, A. P. (2024). Voltage-gated sodium channels in cancers. *Biomarker Research*, *12*(1), 70. https://doi.org/10.1186/s40364-024-00620-x

Liu, L., Kimball, S., Liu, H., Holowatyj, A., & Yang, Z.-Q. (2014). Genetic alterations of histone lysine methyltransferases and their significance in breast cancer. *Oncotarget*, *6*(4), 2466–2482. https://doi.org/10.18632/oncotarget.2967

Liu, M., Xing, L.-Q., & Liu, Y.-J. (2017). A three-long noncoding RNA signature as a diagnostic biomarker for differentiating between triple-negative and non-triple-negative breast cancers. *Medicine*, *96*(9), e6222. https://doi.org/10.1097/MD.0000000000006222

Liu, Q., Wang, Z., Kong, X., Wang, X., Qi, Y., Gao, R., Fang, Y., & Wang, J. (2020). A Novel Prognostic Signature of mRNA-lncRNA in Breast Cancer. *DNA and Cell Biology*, *39*(4), 671–682. https://doi.org/10.1089/dna.2019.5223

Liu, Q., Yao, S., Zhao, H., Hu, Q., Kwan, M. L., Roh, J. M., Ambrosone, C. B., Kushi, L. H., Liu, S., & Zhu, Q. (2019). Early-onset triple-negative breast cancer in multiracial/ethnic populations: Distinct trends of prevalence of truncation mutations. *Cancer Medicine*, *8*(4), 1845. https://doi.org/10.1002/cam4.2047

Liu, X., Zhang, G., Yu, T., He, J., Liu, J., Chai, X., Zhao, G., Yin, D., & Zhang, C. (2022). Exosomes deliver lncRNA DARS-AS1 siRNA to inhibit chronic unpredictable mild stress-induced TNBC metastasis. *Cancer Letters*, *543*, 215781. https://doi.org/10.1016/j.canlet.2022.215781

Liu, X., Zhang, G., Yu, T., Liu, J., Chai, X., Yin, D., & Zhang, C. (2023). CL4-modified exosomes deliver lncRNA DARS-AS1 siRNA to suppress triple-negative breast cancer progression and attenuate doxorubicin resistance by inhibiting autophagy. *International Journal of Biological Macromolecules*, *250*, 126147. https://doi.org/10.1016/j.ijbiomac.2023.126147

Liu, Y., Nguyen, N., & Colditz, G. A. (2015). Links between Alcohol Consumption and Breast Cancer: A Look at the Evidence. *Women's Health*, *11*(1), 65–77.

https://doi.org/10.2217/WHE.14.62
Liu, Y., Peng, H., & Zhang, J.-T. (2005). Expression Profiling of ABC Transporters in a Drug-Resistant Breast Cancer Cell Line Using AmpArray. *Molecular Pharmacology*, *68*(2), 430–438. https://doi.org/10.1124/mol.105.011015
Liu, Y., Zhang, M., Sun, J., Zhang, J., Gu, B., Li, J., Pan, B., & Zhao, H. (2024). Construction and validation of immune-associated lncRNA model for predicting immune status and therapeutic reactions of triple-negative breast cancer. *American Journal of Translational Research*, *16*(9), 4355. https://doi.org/10.62347/VIXN9362
Livasy, C. A., Karaca, G., Nanda, R., Tretiakova, M. S., Olopade, O. I., Moore, D. T., & Perou, C. M. (2006). Phenotypic evaluation of the basal-like subtype of invasive breast carcinoma. *Modern Pathology*, *19*(2), 264–271. https://doi.org/10.1038/modpathol.3800528
Loibl, S., André, F., Bachelot, T., Barrios, C. H., Bergh, J., Burstein, H. J., Cardoso, M. J., Carey, L. A., Dawood, S., Mastro, L. D., Denkert, C., Fallenberg, E. M., Francis, P. A., Gamal-Eldin, H., Gelmon, K., Geyer, C. E., Gnant, M., Guarneri, V., Gupta, S., … Harbeck, N. (2024). Early breast cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up☆. *Annals of Oncology*, *35*(2), 159–182. https://doi.org/10.1016/j.annonc.2023.11.016
Lonsdale, J., Thomas, J., Salvatore, M., Phillips, R., Lo, E., Shad, S., Hasz, R., Walters, G., Garcia, F., Young, N., Foster, B., Moser, M., Karasik, E., Gillard, B., Ramsey, K., Sullivan, S., Bridge, J., Magazine, H., Syron, J., … Moore, H. F. (2013). The Genotype-Tissue Expression (GTEx) project. *Nature Genetics*, *45*(6), 580–585. https://doi.org/10.1038/ng.2653
Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550. https://doi.org/10.1186/s13059-014-0550-8
Lozano-Romero, A., Astudillo-de la Vega, H., Terrones-Gurrola, M. C. del R., Marchat, L. A., Hernández-Sotelo, D., Salinas-Vera, Y. M., Ramos-Payan, R., Silva-Cázares, M. B., Nuñez-Olvera, S. I., Hernández-de la Cruz, O. N., & López-Camarillo, C. (2020). HOX Transcript Antisense RNA HOTAIR Abrogates Vasculogenic Mimicry by Targeting the AngiomiR-204/FAK Axis in Triple Negative Breast Cancer Cells. *Non-Coding RNA*, *6*(2), 19. https://doi.org/10.3390/ncrna6020019
Lu, C., Dong, J., Ma, H., Jin, G., Hu, Z., Peng, Y., Guo, X., Wang, X., & Shen, H. (2009). CCND1 G870A polymorphism contributes to breast cancer susceptibility: A meta-analysis. *Breast Cancer Research and Treatment*, *116*(3), 571–575. https://doi.org/10.1007/s10549-008-0195-y
Luini, A., Gatti, G., Galimberti, V., Zurrida, S., Intra, M., Gentilini, O., Paganelli, G., Viale, G., Orecchia, R., Veronesi, P., & Veronesi, U. (2005). Conservative Treatment of Breast Cancer: Its Evolution. *Breast Cancer Research and Treatment*, *94*(3), 195–198. https://doi.org/10.1007/s10549-004-7376-0
Magny, S. J., Shikhman, R., & Keppke, A. L. (2025). Breast Imaging Reporting and

Data System. In *StatPearls*. StatPearls Publishing. http://www.ncbi.nlm.nih.gov/books/NBK459169/

Makki, J. (2015). Diversity of Breast Carcinoma: Histological Subtypes and Clinical Relevance. *Clinical Medicine Insights: Pathology*, *8*, CPath.S31563. https://doi.org/10.4137/CPath.S31563

Malvia, S., Bagadi, S. A., Dubey, U. S., & Saxena, S. (2017). Epidemiology of breast cancer in Indian women. *Asia-Pacific Journal of Clinical Oncology*, *13*(4), 289–295. https://doi.org/10.1111/ajco.12661

Mani, C., Acharya, G., Saamarthy, K., Ochola, D., Mereddy, S., Pruitt, K., Manne, U., & Palle, K. (2023). Racial differences in RAD51 expression are regulated by miRNA-214-5P and its inhibition synergizes with olaparib in triple-negative breast cancer. *Breast Cancer Research: BCR*, *25*(1), 44. https://doi.org/10.1186/s13058-023-01615-6

Marcu, A., Nitusca, D., Vaduva, A., Baderca, F., Cireap, N., Coricovac, D., Dehelean, C. A., Seclaman, E., Ilina, R., & Marian, C. (2021). Long Non-Coding RNA Expression in Laser Micro-Dissected Luminal A and Triple Negative Breast Cancer Tissue Samples—A Pilot Study. *Medicina*, *57*(4), 371. https://doi.org/10.3390/medicina57040371

Marker, K. M., Zavala, V. A., Vidaurre, T., Lott, P. C., Vásquez, J. N., Casavilca-Zambrano, S., Calderón, M., Abugattas, J. E., Gómez, H. L., Fuentes, H. A., Picoaga, R. L., Cotrina, J. M., Neciosup, S. P., Castañeda, C. A., Morante, Z., Valencia, F., Torres, J., Echeverry, M., Bohórquez, M. E., … COLUMBUS Consortium. (2020). Human Epidermal Growth Factor Receptor 2–Positive Breast Cancer Is Associated with Indigenous American Ancestry in Latin American Women. *Cancer Research*, *80*(9), 1893–1901. https://doi.org/10.1158/0008-5472.CAN-19-3659

Marra, A., & Curigliano, G. (2021). Adjuvant and Neoadjuvant Treatment of Triple-Negative Breast Cancer With Chemotherapy. *The Cancer Journal*, *27*(1), 41–49. https://doi.org/10.1097/PPO.0000000000000498

Marrazzo, E., Frusone, F., Milana, F., Sagona, A., Gatzemeier, W., Barbieri, E., Bottini, A., Canavese, G., Rubino, A. O., Eboli, M. G., Rossetti, C. M., Testori, A., Errico, V., Luca, A. D., & Tinterri, C. (2020). Mucinous breast cancer: A narrative review of the literature and a retrospective tertiary single-centre analysis. *The Breast*, *49*, 87–92. https://doi.org/10.1016/j.breast.2019.11.002

Martianov, I., Ramadass, A., Serra Barros, A., Chow, N., & Akoulitchev, A. (2007). Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript. *Nature*, *445*(7128), 666–670. https://doi.org/10.1038/nature05519

Martin, D. N., Boersma, B. J., Yi, M., Reimers, M., Howe, T. M., Yfantis, H. G., Tsai, Y. C., Williams, E. H., Lee, D. H., Stephens, R. M., Weissman, A. M., & Ambs, S. (2009). Differences in the Tumor Microenvironment between African-American and European-American Breast Cancer Patients. *PLOS*

ONE, *4*(2), e4531. https://doi.org/10.1371/journal.pone.0004531
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
Mayakonda, A., Lin, D.-C., Assenov, Y., Plass, C., & Koeffler, H. P. (2018). Maftools: Efficient and comprehensive analysis of somatic variants in cancer. *Genome Research*, *28*(11), 1747–1756. https://doi.org/10.1101/gr.239244.118
Mazumder, A., Jimenez, A., Ellsworth, R. E., Freedland, S. J., George, S., Bainbridge, M. N., & Haricharan, S. (2022). The DNA damage repair landscape in Black women with breast cancer. *Therapeutic Advances in Medical Oncology*, *14*, 17588359221075458. https://doi.org/10.1177/17588359221075458
McAndrew, H., Rigler, J., Yeguvapalli, S., & Chitrala, K. N. (2024). *Analysis of gene expression profiles to elucidate racial differences in African American and White patients with Triple-negative breast cancer* (p. 2024.05.29.596478). bioRxiv. https://doi.org/10.1101/2024.05.29.596478
McCart Reed, A. E., Kalinowski, L., Simpson, P. T., & Lakhani, S. R. (2021). Invasive lobular carcinoma of the breast: The increasing importance of this special subtype. *Breast Cancer Research : BCR*, *23*, 6. https://doi.org/10.1186/s13058-020-01384-6
McCormick, B., Winter, K., Hudis, C., Kuerer, H. M., Rakovitch, E., Smith, B. L., Sneige, N., Moughan, J., Shah, A., Germain, I., Hartford, A. C., Rashtian, A., Walker, E. M., Yuen, A., Strom, E. A., Wilcox, J. L., Vallow, L. A., Small, W., Pu, A. T., … White, J. (2015). RTOG 9804: A Prospective Randomized Trial for Good-Risk Ductal Carcinoma In Situ Comparing Radiotherapy With Observation. *Journal of Clinical Oncology*, *33*(7), 709–715. https://doi.org/10.1200/JCO.2014.57.9029
McDonald, E. S., Clark, A. S., Tchou, J., Zhang, P., & Freedman, G. M. (2016). Clinical Diagnosis and Management of Breast Cancer. *Journal of Nuclear Medicine*, *57*(Supplement 1), 9S-16S. https://doi.org/10.2967/jnumed.115.157834
McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., & DePristo, M. A. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, *20*(9), 1297–1303. https://doi.org/10.1101/gr.107524.110
Mehrotra, J., Ganpat, M. M., Kanaan, Y., Fackler, M. J., McVeigh, M., Lahti-Domenici, J., Polyak, K., Argani, P., Naab, T., Garrett, E., Parmigiani, G., Broome, C., & Sukumar, S. (2004). Estrogen Receptor/Progesterone Receptor-Negative Breast Cancers of Young African-American Women Have a Higher Frequency of Methylation of Multiple Genes than Those of Caucasian Women1. *Clinical Cancer Research*, *10*(6), 2052–2057. https://doi.org/10.1158/1078-0432.CCR-03-0514
Menon, G., Alkabban, F. M., & Ferguson, T. (2025). Breast Cancer. In *StatPearls*.

StatPearls Publishing. http://www.ncbi.nlm.nih.gov/books/NBK482286/
Merikhian, P., Eisavand, M. R., & Farahmand, L. (2021). Triple-negative breast cancer: Understanding Wnt signaling in drug resistance. *Cancer Cell International*, *21*, 419. https://doi.org/10.1186/s12935-021-02107-3
Miyashita, M., Bell, J. S. K., Wenric, S., Karaesmen, E., Rhead, B., Kase, M., Kaneva, K., De La Vega, F. M., Zheng, Y., Yoshimatsu, T. F., Khramtsova, G., Liu, F., Zhao, F., Howard, F. M., Nanda, R., Beaubier, N., White, K. P., Huo, D., & Olopade, O. I. (2023). Molecular profiling of a real-world breast cancer cohort with genetically inferred ancestries reveals actionable tumor biology differences between European ancestry and African ancestry patient populations. *Breast Cancer Research*, *25*(1), 58. https://doi.org/10.1186/s13058-023-01627-2
Moasser, M. M. (2007). The oncogene HER2; Its signaling and transforming functions and its role in human cancer pathogenesis. *Oncogene*, *26*(45), 6469–6487. https://doi.org/10.1038/sj.onc.1210477
Mohr, C. J., Gross, D., Sezgin, E. C., Steudel, F. A., Ruth, P., Huber, S. M., & Lukowski, R. (2019). KCa3.1 Channels Confer Radioresistance to Breast Cancer Cells. *Cancers*, *11*(9), 1285. https://doi.org/10.3390/cancers11091285
Mørch, L. S., Skovlund, C. W., Hannaford, P. C., Iversen, L., Fielding, S., & Lidegaard, Ø. (2017). Contemporary Hormonal Contraception and the Risk of Breast Cancer. *New England Journal of Medicine*, *377*(23), 2228–2239. https://doi.org/10.1056/NEJMoa1700732
Morris, G. J., Naidu, S., Topham, A. K., Guiles, F., Xu, Y., McCue, P., Schwartz, G. F., Park, P. K., Rosenberg, A. L., Brill, K., & Mitchell, E. P. (2007). Differences in breast carcinoma characteristics in newly diagnosed African–American and Caucasian patients: A single-institution compilation compared with the National Cancer Institute's Surveillance, Epidemiology, and end results database. *Cancer*, *110*(4), 876–884. https://doi.org/10.1002/cncr.22836
Mounir, M., Lucchetta, M., Silva, T. C., Olsen, C., Bontempi, G., Chen, X., Noushmehr, H., Colaprico, A., & Papaleo, E. (2019). New functionalities in the TCGAbiolinks package for the study and integration of cancer data from GDC and GTEx. *PLoS Computational Biology*, *15*(3), e1006701. https://doi.org/10.1371/journal.pcbi.1006701
Mu, D., Chen, L., Zhang, X., See, L.-H., Koch, C. M., Yen, C., Tong, J. J., Spiegel, L., Nguyen, K. C. Q., Servoss, A., Peng, Y., Pei, L., Marks, J. R., Lowe, S., Hoey, T., Jan, L. Y., McCombie, W. R., Wigler, M. H., & Powers, S. (2003). Genomic amplification and oncogenic properties of the KCNK9 potassium channel gene. *Cancer Cell*, *3*(3), 297–302. https://doi.org/10.1016/S1535-6108(03)00054-0
Nascimento, R. G. D., & Otoni, K. M. (2020). Histological and molecular classification of breast cancer: What do we know? *Mastology*, *30*, e20200024. https://doi.org/10.29289/25945394202020200024
Nicosia, L., Gnocchi, G., Gorini, I., Venturini, M., Fontana, F., Pesapane, F., Abiuso,

I., Bozzini, A. C., Pizzamiglio, M., Latronico, A., Abbate, F., Meneghetti, L., Battaglia, O., Pellegrino, G., & Cassano, E. (2023). History of Mammography: Analysis of Breast Imaging Diagnostic Achievements over the Last Century. *Healthcare*, *11*(11), 1596. https://doi.org/10.3390/healthcare11111596

Nielsen, T. O., Hsu, F. D., Jensen, K., Cheang, M., Karaca, G., Hu, Z., Hernandez-Boussard, T., Livasy, C., Cowan, D., Dressler, L., Akslen, L. A., Ragaz, J., Gown, A. M., Gilks, C. B., Van De Rijn, M., & Perou, C. M. (2004). Immunohistochemical and Clinical Characterization of the Basal-Like Subtype of Invasive Breast Carcinoma. *Clinical Cancer Research*, *10*(16), 5367–5374. https://doi.org/10.1158/1078-0432.CCR-04-0220

Norat, T., Scoccianti, C., Boutron-Ruault, M.-C., Anderson, A., Berrino, F., Cecchini, M., Espina, C., Key, T., Leitzmann, M., Powers, H., Wiseman, M., & Romieu, I. (2015). European Code against Cancer 4th Edition: Diet and cancer. *Cancer Epidemiology*, *39*, S56–S66. https://doi.org/10.1016/j.canep.2014.12.016

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 : An Official Journal of the American Association for Cancer Research*, *16*(24), 6100–6110. https://doi.org/10.1158/1078-0432.CCR-10-1533

Oeggerli, M., Tian, Y., Ruiz, C., Wijker, B., Sauter, G., Obermann, E., Güth, U., Zlobec, I., Sausbier, M., Kunzelmann, K., & Bubendorf, L. (2012). Role of KCNMA1 in Breast Cancer. *PLOS ONE*, *7*(8), e41664. https://doi.org/10.1371/journal.pone.0041664

Omilian, A. R., Wei, L., Hong, C.-C., Bandera, E. V., Liu, S., Khoury, T., Ambrosone, C. B., & Yao, S. (2020). Somatic Mutations of Triple-Negative Breast Cancer: A Comparison between Black and White Women. *Breast Cancer Research and Treatment*, *182*(2), 503. https://doi.org/10.1007/s10549-020-05693-4

Ooi, S. L., Martinez, M. E., & Li, C. I. (2011). Disparities in Breast Cancer Characteristics and Outcomes by Race/Ethnicity. *Breast Cancer Research and Treatment*, *127*(3), 729–738. https://doi.org/10.1007/s10549-010-1191-6

Orucevic, A., Heidel, R. E., & Bell, J. L. (2016). Utilization and impact of 21-gene recurrence score assay for breast cancer in clinical practice across the United States: Lessons learned from the 2010 to 2012 National Cancer Data Base analysis. *Breast Cancer Research and Treatment*, *157*(3), 427–435. https://doi.org/10.1007/s10549-016-3833-9

Ossovskaya, V., Wang, Y., Budoff, A., Xu, Q., Lituev, A., Potapova, O., Vansant, G., Monforte, J., & Daraselia, N. (2011). Exploring Molecular Pathways of Triple-Negative Breast Cancer. *Genes & Cancer*, *2*(9), 870–879. https://doi.org/10.1177/1947601911432496

Otterbach, F., Callies, R., Adamzik, M., Kimmig, R., Siffert, W., Schmid, K. W., & Bankfalvi, A. (2010). Aquaporin 1 (AQP1) expression is a novel characteristic feature of a particularly aggressive subgroup of basal-like breast carcinomas.

Breast Cancer Research and Treatment, *120*(1), 67–76. https://doi.org/10.1007/s10549-009-0370-9

Ouadid-Ahidouch, H., Ahidouch, A., & Pardo, L. A. (2016). Kv10.1 K+ channel: From physiology to cancer. *Pflügers Archiv - European Journal of Physiology*, *468*(5), 751–762. https://doi.org/10.1007/s00424-015-1784-3

Overmoyer, B. (1999). *BREAST CANCER SCREENING*.

Palmieri, C., Rudraraju, B., Monteverde, M., Lattanzio, L., Gojis, O., Brizio, R., Garrone, O., Merlano, M., Syed, N., Lo Nigro, C., & Crook, T. (2012). Methylation of the calcium channel regulatory subunit α2δ-3 (CACNA2D3) predicts site-specific relapse in oestrogen receptor-positive primary breast carcinomas. *British Journal of Cancer*, *107*(2), 375–381. https://doi.org/10.1038/bjc.2012.231

Pan, Y., Wen, J., Ning, Z., Yuan, Y., Liu, X., Yang, Y., Guan, Y., Lu, Y., Mamatyusupu, D., & Xu, S. (2023). Comparative Genomic and Transcriptomic Analyses Reveal the Impacts of Genetic Admixture in Kazaks, Uyghurs, and Huis. *Molecular Biology and Evolution*, *40*(3), msad054. https://doi.org/10.1093/molbev/msad054

Pan, Y., Zhao, Q., He, H., Qi, Y., Bai, Y., Zhao, J., & Yang, Y. (2023). TRPML1 as a potential therapeutic target for triple-negative breast cancer: A review. *Frontiers in Oncology*, *13*. https://doi.org/10.3389/fonc.2023.1326023

Pardo, L. A., del Camino, D., Sánchez, A., Alves, F., Brüggemann, A., Beckh, S., & Stühmer, W. (1999). Oncogenic potential of EAG K(+) channels. *The EMBO Journal*, *18*(20), 5540–5547. https://doi.org/10.1093/emboj/18.20.5540

Pardo, L. A., & Stühmer, W. (2008). Eag1 as a cancer target. *Expert Opinion on Therapeutic Targets*, *12*(7), 837–843. https://doi.org/10.1517/14728222.12.7.837

Park, K. U., Somerfield, M. R., Anne, N., Brackstone, M., Conlin, A. K., Couto, H. L., Dengel, L. T., Eisen, A., Harvey, B. E., Hawley, J., Kim, J. N., Lasebikan, N., McDonald, E. S., Pradhan, D., Shams, S., Vega, R. M., Thompson, A. M., & Torres, M. A. (2025). Sentinel Lymph Node Biopsy in Early-Stage Breast Cancer: ASCO Guideline Update. *Journal of Clinical Oncology*, *43*(14), 1720–1741. https://doi.org/10.1200/JCO-25-00099

Parker, J. S., Mullins, M., Cheang, M. C. U., Leung, S., Voduc, D., Vickery, T., Davies, S., Fauron, C., He, X., Hu, Z., Quackenbush, J. F., Stijleman, I. J., Palazzo, J., Marron, J. S., Nobel, A. B., Mardis, E., Nielsen, T. O., Ellis, M. J., Perou, C. M., & Bernard, P. S. (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

Parkin, D. M., Boyd, L., & Walker, L. C. (2011). 16. The fraction of cancer attributable to lifestyle and environmental factors in the UK in 2010. *British Journal of Cancer*, *105*(Suppl 2), S77–S81. https://doi.org/10.1038/bjc.2011.489

Pecci, A., Ogara, M. F., Sanz, R. T., & Vicent, G. P. (2022). Choosing the right partner

in hormone-dependent gene regulation: Glucocorticoid and progesterone receptors crosstalk in breast cancer cells. *Frontiers in Endocrinology*, *13*. https://doi.org/10.3389/fendo.2022.1037177

Pedersen, M. H., Hood, B. L., Beck, H. C., Conrads, T. P., Ditzel, H. J., & Leth-Larsen, R. (2017). Downregulation of antigen presentation-associated pathway proteins is linked to poor outcome in triple-negative breast cancer patient tumors. *Oncoimmunology*, *6*(5), e1305531. https://doi.org/10.1080/2162402X.2017.1305531

Peiris, H., Mudduwa, L., & Thalagala, N. (2017). *The value of Nottingham grade in breast cancer re-visited in the Sri Lankan setting*.

Perez, E. A., Romond, E. H., Suman, V. J., Jeong, J.-H., Sledge, G., Geyer, C. E., Martino, S., Rastogi, P., Gralow, J., Swain, S. M., Winer, E. P., Colon-Otero, G., Davidson, N. E., Mamounas, E., Zujewski, J. A., & Wolmark, N. (2014). Trastuzumab Plus Adjuvant Chemotherapy for Human Epidermal Growth Factor Receptor 2–Positive Breast Cancer: Planned Joint Analysis of Overall Survival From NSABP B-31 and NCCTG N9831. *Journal of Clinical Oncology*, *32*(33), 3744–3752. https://doi.org/10.1200/JCO.2014.55.5730

Perou, C. M., & Børresen-Dale, A.-L. (2011). Systems Biology and Genomics of Breast Cancer. *Cold Spring Harbor Perspectives in Biology*, *3*(2), a003293. https://doi.org/10.1101/cshperspect.a003293

Perou, C. M., Sørlie, T., Eisen, M. B., Van De Rijn, M., Jeffrey, S. S., Rees, C. A., Pollack, J. R., Ross, D. T., Johnsen, H., Akslen, L. A., Fluge, Ø., Pergamenschikov, 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

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T.-C., Mendell, J. T., & Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature Biotechnology*, *33*(3), 290–295. https://doi.org/10.1038/nbt.3122

Phung, M. T., Tin Tin, S., & Elwood, J. M. (2019). Prognostic models for breast cancer: A systematic review. *BMC Cancer*, *19*, 230. https://doi.org/10.1186/s12885-019-5442-6

Pignatta, S., Tazzari, M., Indio, V., Tumedei, M. M., Limarzi, F., Tauceri, F., Tebaldi, M., Fanelli, D., & Bulgarelli, J. (2025). Comparative analysis of library preparation approaches for FFPE gene expression profiling and related recommendations. *Scientific Reports*, *15*(1), 29592. https://doi.org/10.1038/s41598-025-12992-7

Ping, J., Guo, X., Ye, F., Long, J., Lipworth, L., Cai, Q., Blot, W., Shu, X.-O., & Zheng, W. (2020). Differences in gene-expression profiles in breast cancer between African and European-ancestry women. *Carcinogenesis*, *41*(7), 887–893. https://doi.org/10.1093/carcin/bgaa035

Poggio, F., Bruzzone, M., Ceppi, M., Pondé, N. F., Valle, G. L., Mastro, L. D.,

Azambuja, E. de, & Lambertini, M. (2018). Platinum-based neoadjuvant chemotherapy in triple-negative breast cancer: A systematic review and meta-analysis. *Annals of Oncology*, *29*(7), 1497–1508. https://doi.org/10.1093/annonc/mdy127

Poorolajal, J., Heidarimoghis, F., Karami, M., Cheraghi, Z., Gohari-Ensaf, F., Shahbazi, F., Zareie, B., Ameri, P., & Sahraei, F. (2021). Factors for the Primary Prevention of Breast Cancer: A Meta-Analysis of Prospective Cohort Studies. *Journal of Research in Health Sciences*, *21*(3), e00520. https://doi.org/10.34172/jrhs.2021.57

Porter, P. L., Lund, M. J., Lin, M. G., Yuan, X., Liff, J. M., Flagg, E. W., Coates, R. J., & Eley, J. W. (2004). Racial differences in the expression of cell cycle-regulatory proteins in breast carcinoma: Study of young African American and white women in Atlanta, Georgia. *Cancer*, *100*(12), 2533–2542. https://doi.org/10.1002/cncr.20279

Prat, A., Parker, J. S., Karginova, O., Fan, C., Livasy, C., Herschkowitz, J. I., He, X., & Perou, C. M. (2010). Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Research*, *12*(5), R68. https://doi.org/10.1186/bcr2635

Provenzano, E., Ulaner, G. A., & Chin, S.-F. (2018). Molecular Classification of Breast Cancer. *PET Clinics*, *13*(3), 325–338. https://doi.org/10.1016/j.cpet.2018.02.004

Purrington, K. S., Knight, J., Dyson, G., Ali-Fehmi, R., Schwartz, A. G., Boerner, J. L., & Bandyopadhyay, S. (2020). CLCA2 expression is associated with survival among African American women with triple negative breast cancer. *PloS One*, *15*(4), e0231712. https://doi.org/10.1371/journal.pone.0231712

Qari, A. S., Mowais, A. H., Alharbi, S. M., Almuayrifi, M. J., Asiri, A. A. A., Alwatid, S. A., Aljohani, A. A., Alanazi, R. M., Thoubaity, F. A., Qari, A. S., Mowais, A. H., Alharbi, S. M., Almuayrifi, M. J., Asiri, A. A. A., Alwatid, S. A., Aljohani, A. A., Alanazi, R. M., & Thoubaity, F. A. (2024). Adjuvant and Neoadjuvant Therapy for Breast Cancer: A Systematic Review. *European Journal of Breast Health*. https://doi.org/10.4274/ejbh.galenos.2024.2023-12-16

Qiu, H., Cao, S., & Xu, R. (2021). Cancer incidence, mortality, and burden in China: A time-trend analysis and comparison with the United States and United Kingdom based on the global epidemiological data released in 2020. *Cancer Communications*, *41*(10), 1037–1048. https://doi.org/10.1002/cac2.12197

Rakha, E. A., & Green, A. R. (2017). Molecular classification of breast cancer: What the pathologist needs to know. *Pathology*, *49*(2), 111–119. https://doi.org/10.1016/j.pathol.2016.10.012

Rakha, E. A., Reis-Filho, J. S., Baehner, F., Dabbs, D. J., Decker, T., Eusebi, V., Fox, S. B., Ichihara, S., Jacquemier, J., Lakhani, S. R., Palacios, J., Richardson, A. L., Schnitt, S. J., Schmitt, F. C., Tan, P.-H., Tse, G. M., Badve, S., & Ellis, I. O. (2010). Breast cancer prognostic classification in the molecular era: The role

of histological grade. *Breast Cancer Research : BCR*, *12*(4), 207. https://doi.org/10.1186/bcr2607

Rakha, E. A., Reis-Filho, J. S., & Ellis, I. O. (2010). Combinatorial biomarker expression in breast cancer. *Breast Cancer Research and Treatment*, *120*(2), 293–308. https://doi.org/10.1007/s10549-010-0746-x

Ramena, G., Yin, Y., Yu, Y., Walia, V., & Elble, R. C. (2016). CLCA2 Interactor EVA1 Is Required for Mammary Epithelial Cell Differentiation. *PloS One*, *11*(3), e0147489. https://doi.org/10.1371/journal.pone.0147489

Ramkumar, C., Buturovic, L., Malpani, S., Kumar Attuluri, A., Basavaraj, C., Prakash, C., Madhav, L., Doval, D. C., Mehta, A., & Bakre, M. M. (2018). Development of a Novel Proteomic Risk-Classifier for Prognostication of Patients With Early-Stage Hormone Receptor–Positive Breast Cancer. *Biomarker Insights*, *13*, 1177271918789100. https://doi.org/10.1177/1177271918789100

Ramsay, D., Kent, J., Hartmann, R., & Hartman, P. (2005). Anatomy of the lactating human breast redefined with ultrasound imaging. *Journal of Anatomy*, *206*(6), 525–534. https://doi.org/10.1111/j.1469-7580.2005.00417.x

Redden, M. H., & Fuhrman, G. M. (2013). Neoadjuvant Chemotherapy in the Treatment of Breast Cancer. *Surgical Clinics of North America*, *93*(2), 493–499. https://doi.org/10.1016/j.suc.2013.01.006

Reeder-Hayes, K. E., Mayer, S. E., Olshan, A. F., Wheeler, S. B., Carey, L. A., Tse, C.-K., Bell, M. E., & Troester, M. A. (2019). Race and delays in breast cancer treatment across the care continuum in the Carolina Breast Cancer Study. *Cancer*, *125*(22), 3985–3992. https://doi.org/10.1002/cncr.32378

Reinbolt, R. E., Mangini, N., Hill, J. L., Levine, L. B., Dempsey, J. L., Singaravelu, J., Koehler, K. A., Talley, A., & Lustberg, M. B. (2015). Endocrine Therapy in Breast Cancer: The Neoadjuvant, Adjuvant, and Metastatic Approach. *Seminars in Oncology Nursing*, *31*(2), 146–155. https://doi.org/10.1016/j.soncn.2015.02.002

Rey-Vargas, L., Bejarano-Rivera, L. M., López-Correa, P., Ballen-Lozano, D. F., & Serrano-Gómez, S. J. (2025). Differences in the distribution of HER2-positive breast tumors according to ethnicity and genetic variants in ERBB2: A special focus on Asian and Latina women. *Frontiers in Oncology*, *15*, 1635681. https://doi.org/10.3389/fonc.2025.1635681

Rizzo, J. A., Sherman, W. E., & Arciero, C. A. (2015). Racial disparity in survival from early breast cancer in the department of defense healthcare system. *Journal of Surgical Oncology*, *111*(7), 819–823. https://doi.org/10.1002/jso.23884

Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, *26*(1), 139–140. https://doi.org/10.1093/bioinformatics/btp616

Roger, S., Besson, P., & Le Guennec, J.-Y. (2003). Involvement of a novel fast inward sodium current in the invasion capacity of a breast cancer cell line. *Biochimica*

et Biophysica Acta (BBA) - Biomembranes, *1616*(2), 107–111. https://doi.org/10.1016/j.bbamem.2003.07.001

Rose, D. P., Gracheck, P. J., & Vona-Davis, L. (2015). The Interactions of Obesity, Inflammation and Insulin Resistance in Breast Cancer. *Cancers*, *7*(4), 2147–2168. https://doi.org/10.3390/cancers7040883

Rosen, R. D., & Sapra, A. (2025). TNM Classification. In *StatPearls*. StatPearls Publishing. http://www.ncbi.nlm.nih.gov/books/NBK553187/

Roux, P., Knight, S., Cohen, M., Classe, J. M., Mazouni, C., Chauvet, M.-P., Reyal, F., Colombo, P.-E., Jouve, E., Chopin, N., Daraï, E., Coutant, C., Lambaudie, E., & Houvenaeghel, G. (2019). Tubular and mucinous breast cancer: Results of a cohort of 917 patients. *Tumori Journal*, *105*(1), 55–62. https://doi.org/10.1177/0300891618811282

Ruiz-Narváez, E. A., Lunetta, K. L., Hong, C.-C., Haddad, S., Yao, S., Cheng, T.-Y. D., Bensen, J. T., Bandera, E. V., Haiman, C. A., Troester, M. A., Ambrosone, C. B., Rosenberg, L., & Palmer, J. R. (2016). Genetic variation in the insulin, insulin-like growth factor, growth hormone, and leptin pathways in relation to breast cancer in African-American women: The AMBER consortium. *NPJ Breast Cancer*, *2*, 16034. https://doi.org/10.1038/npjbcancer.2016.34

Saha Roy, S., & Vadlamudi, R. K. (2012). Role of Estrogen Receptor Signaling in Breast Cancer Metastasis. *International Journal of Breast Cancer*, *2012*, 654698. https://doi.org/10.1155/2012/654698

Salama, E. A., Adbeltawab, R. E., & El Tayebi, H. M. (2020). XIST and TSIX: Novel Cancer Immune Biomarkers in PD-L1-Overexpressing Breast Cancer Patients. *Frontiers in Oncology*, *9*, 1459. https://doi.org/10.3389/fonc.2019.01459

Sasaki, Y., Koyama, R., Maruyama, R., Hirano, T., Tamura, M., Sugisaka, J., Suzuki, H., Idogawa, M., Shinomura, Y., & Tokino, T. (2012). CLCA2, a target of the p53 family, negatively regulates cancer cell migration and invasion. *Cancer Biology & Therapy*, *13*(14), 1512–1521. https://doi.org/10.4161/cbt.22280

Sawaki, M., Shien, T., & Iwata, H. (2019). TNM classification of malignant tumors (Breast Cancer Study Group). *Japanese Journal of Clinical Oncology*, *49*(3), 228–231. https://doi.org/10.1093/jjco/hyy182

Sayed, S., Moloo, Z., Wasike, R., Bird, P., Oigara, R., Njoroge, F. W., Shaikh, A. J., Prasad, S. V., Vinayak, S., Gierach, G. L., Dawsey, S. M., Palakal, M., Fan, S., Mullooly, M., Chauhan, R., Okiro, P., Gakinya, S., Nzioka, A., Kyobutungi, C., … Saleh, M. (2018). Ethnicity and breast cancer characteristics in Kenya. *Breast Cancer Research and Treatment*, *167*(2), 425–437. https://doi.org/10.1007/s10549-017-4511-2

Schwartz, T. L., Mogal, H., Papageorgiou, C., Veerapong, J., & Hsueh, E. C. (2013). Metaplastic breast cancer: Histologic characteristics, prognostic factors and systemic treatment strategies. *Experimental Hematology & Oncology*, *2*(1), 31. https://doi.org/10.1186/2162-3619-2-31

Seal, M., Wilson, C., Naus, G. J., Chia, S., Bainbridge, T. C., & Hayes, M. M. (2009). Encapsulated apocrine papillary carcinoma of the breast—a tumour of

uncertain malignant potential: Report of five cases. *Virchows Archiv*, *455*(6), 477–483. https://doi.org/10.1007/s00428-009-0834-7

Seehawer, M., Li, Z., Nishida, J., Foidart, P., Reiter, A. H., Rojas-Jimenez, E., Goyette, M.-A., Yan, P., Raval, S., Munoz Gomez, M., Cejas, P., Long, H. W., Papanastasiou, M., & Polyak, K. (2024). Loss of Kmt2c or Kmt2d drives brain metastasis via KDM6A-dependent upregulation of MMP3. *Nature Cell Biology*, *26*(7), 1165–1175. https://doi.org/10.1038/s41556-024-01446-3

Sengupta, A. K., Gunda, A., Malpani, S., Serkad, C. P. V., Basavaraj, C., Bapat, A., & Bakre, M. M. (2020). Comparison of breast cancer prognostic tests CanAssist Breast and Oncotype DX. *Cancer Medicine*, *9*(21), 7810–7818. https://doi.org/10.1002/cam4.3495

Sheppard, V. B., De Mendoza, A. H., He, J., Jennings, Y., Edmonds, M. C., Oppong, B. A., & Tadesse, M. G. (2018). Initiation of Adjuvant Endocrine Therapy in Black and White Women With Breast Cancer. *Clinical Breast Cancer*, *18*(5), 337-346.e1. https://doi.org/10.1016/j.clbc.2017.12.002

Shi, X., He, W., Guo, S., Zhang, B., Ren, S., Liu, K., Sun, T., & Cui, J. (2019). RNA-seq Analysis of the SCN1A-KO Model based on CRISPR/Cas9 Genome Editing Technology. *Neuroscience*, *398*, 1–11. https://doi.org/10.1016/j.neuroscience.2018.11.052

Shima, H., Kida, K., Adachi, S., Yamada, A., Sugae, S., Narui, K., Miyagi, Y., Nishi, M., Ryo, A., Murata, S., Taniguchi, H., Ichikawa, Y., Ishikawa, T., & Endo, I. (2018). Lnc RNA H19 is associated with poor prognosis in breast cancer patients and promotes cancer stemness. *Breast Cancer Research and Treatment*, *170*(3), 507–516. https://doi.org/10.1007/s10549-018-4793-z

Shoemaker, M. L., White, M. C., Wu, M., Weir, H. K., & Romieu, I. (2018). Differences in breast cancer incidence among young women aged 20–49 years by stage and tumor characteristics, age, race, and ethnicity, 2004–2013. *Breast Cancer Research and Treatment*, *169*(3), 595–606. https://doi.org/10.1007/s10549-018-4699-9

Sieri, S., Chiodini, P., Agnoli, C., Pala, V., Berrino, F., Trichopoulou, A., Benetou, V., Vasilopoulou, E., Sánchez, M.-J., Chirlaque, M.-D., Amiano, P., Quirós, J. R., Ardanaz, E., Buckland, G., Masala, G., Panico, S., Grioni, S., Sacerdote, C., Tumino, R., … Krogh, V. (2014). Dietary Fat Intake and Development of Specific Breast Cancer Subtypes. *JNCI: Journal of the National Cancer Institute*, *106*(5), dju068. https://doi.org/10.1093/jnci/dju068

Silverstein, M. J., Mai, T., Savalia, N., Vaince, F., & Guerra, L. (2014). Oncoplastic breast conservation surgery: The new paradigm. *Journal of Surgical Oncology*, *110*(1), 82–89. https://doi.org/10.1002/jso.23641

Singh, D. D., Lee, H.-J., & Yadav, D. K. (2022). Clinical updates on tyrosine kinase inhibitors in HER2-positive breast cancer. *Frontiers in Pharmacology*, *13*. https://doi.org/10.3389/fphar.2022.1089066

Slamon, D. J., Leyland-Jones, B., Shak, S., Fuchs, H., Paton, V., Bajamonde, A., Fleming, T., Eiermann, W., Wolter, J., Pegram, M., Baselga, J., & Norton, L.

(2001, March 15). *Use of Chemotherapy plus a Monoclonal Antibody against HER2 for Metastatic Breast Cancer That Overexpresses HER2* (world) [Research-article]. The New England Journal of Medicine; Massachusetts Medical Society. https://doi.org/10.1056/NEJM200103153441101

Slodkowska, E. A., & Ross, J. S. (2009). MammaPrint™ 70-gene signature: Another milestone in personalized medical care for breast cancer patients. *Expert Review of Molecular Diagnostics*, *9*(5), 417–422. https://doi.org/10.1586/erm.09.32

Smoot, B., Kober, K. M., Paul, S. M., Levine, J. D., Abrams, G., Mastick, J., Topp, K., Conley, Y., & Miaskowski, C. (2017). Potassium Channel Candidate Genes Predict the Development of Secondary Lymphedema Following Breast Cancer Surgery. *Nursing Research*, *66*(2), 85–94. https://doi.org/10.1097/NNR.0000000000000203

Song, M.-A., Brasky, T. M., Marian, C., Weng, D. Y., Taslim, C., Dumitrescu, R. G., Llanos, A. A., Freudenheim, J. L., & Shields, P. G. (2015). Racial differences in genome-wide methylation profiling and gene expression in breast tissues from healthy women. *Epigenetics*, *10*(12), 1177–1187. https://doi.org/10.1080/15592294.2015.1121362

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*, *100*(14), 8418–8423. https://doi.org/10.1073/pnas.0932692100

Sparano, J. A., Wang, M., Zhao, F., Stearns, V., Martino, S., Ligibel, J. A., Perez, E. A., Saphner, T., Wolff, A. C., Sledge, G. W., Wood, W. C., & Davidson, N. E. (2012). Race and hormone receptor-positive breast cancer outcomes in a randomized chemotherapy trial. *Journal of the National Cancer Institute*, *104*(5), 406–414. https://doi.org/10.1093/jnci/djr543

Sribenja, S., Maguire, O., Attwood, K., Buas, M. F., Palmer, J. R., Ambrosone, C. B., & Higgins, M. J. (2021). Deletion of Foxa1 in the mouse mammary gland results in abnormal accumulation of luminal progenitor cells: A link between reproductive factors and ER-/TNBC breast cancer? *American Journal of Cancer Research*, *11*(6), 3263.

Stone, A., Kalahiki, C., Li, L., Hubig, N., Iuricich, F., & Dunn, H. (2023). Evaluation of breast tumor morphologies from African American and Caucasian patients.

Computational and Structural Biotechnology Journal, *21*, 3459–3465. https://doi.org/10.1016/j.csbj.2023.06.019

Strand, S. H., Houlahan, K. E., Branch, V., Lynch, T., Rivero-Guitiérrez, B., Harmon, B., Couch, F., Gallagher, K., Kilgore, M., Wei, S., DeMichele, A., King, T., McAuliffe, P., Curtis, C., Owzar, K., Marks, J. R., Colditz, G. A., Hwang, E. S., & West, R. B. (2024). Analysis of ductal carcinoma in situ by self-reported race reveals molecular differences related to outcome. *Breast Cancer Research*, *26*(1), 127. https://doi.org/10.1186/s13058-024-01885-8

Sun, L., Luo, H., Bu, D., Zhao, G., Yu, K., Zhang, C., Liu, Y., Chen, R., & Zhao, Y. (2013). Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. *Nucleic Acids Research*, *41*(17), e166. https://doi.org/10.1093/nar/gkt646

Sun, M. X., Zhu, H. C., Yu, Y., Yao, Y., Li, H. Y., Feng, F. B., Wang, Q. Y., Liu, R. J., & Sun, C. G. (2025). Role of the Wnt signaling pathway in the complex microenvironment of breast cancer and prospects for therapeutic potential (Review). *International Journal of Oncology*, *66*(5), 1–22. https://doi.org/10.3892/ijo.2025.5742

Surdacka, L. M., Jakubas, A., Jagiełło, J., Daniłowska, K., Picheta, N., & Gil-Kulik, P. (2024). Epigenetic and Immune Mechanisms Linking Breastfeeding to Lower Breast Cancer Rates. *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research*, *30*, e945451. https://doi.org/10.12659/MSM.945451

Swain, S. M., Baselga, J., Kim, S.-B., Ro, J., Semiglazov, V., Campone, M., Ciruelos, E., Ferrero, J.-M., Schneeweiss, A., Heeson, S., Clark, E., Ross, G., Benyunes, M. C., & Cortés, J. (2015). Pertuzumab, Trastuzumab, and Docetaxel in HER2-Positive Metastatic Breast Cancer. *The New England Journal of Medicine*, *372*(8), 724–734. https://doi.org/10.1056/NEJMoa1413513

Tang, S., Zhang, Y., Song, L., Hui, K., & Jiang, X. (2025). High CXCL8 expression predicting poor prognosis in triple-negative breast cancer. *Anti-Cancer Drugs*, *36*(3), 246–252. https://doi.org/10.1097/CAD.0000000000001678

Tanikawa, C., Nakagawa, H., Furukawa, Y., Nakamura, Y., & Matsuda, K. (2012). CLCA2 as a p53-inducible senescence mediator. *Neoplasia (New York, N.Y.)*, *14*(2), 141–149. https://doi.org/10.1593/neo.111700

Teisseyre, A., Palko-Labuz, A., Sroda-Pomianek, K., & Michalak, K. (2019). Voltage-Gated Potassium Channel Kv1.3 as a Target in Therapy of Cancer. *Frontiers in Oncology*, *9*, 933. https://doi.org/10.3389/fonc.2019.00933

Teklemariam, A. B., Muche, Z. T., Agidew, M. M., Mulu, A. T., Zewde, E. A., Baye, N. D., Adugna, D. G., Maru, L., & Ayele, T. M. (2024). Receptor tyrosine kinases and steroid hormone receptors in breast cancer: *Review of recent evidences*. *Metabolism Open*, *24*, 100324. https://doi.org/10.1016/j.metop.2024.100324

Telli, M. L., Chang, E. T., Kurian, A. W., Keegan, T. H. M., McClure, L. A., Lichtensztajn, D., Ford, J. M., & Gomez, S. L. (2011). Asian ethnicity and

breast cancer subtypes: A study from the California Cancer Registry. *Breast Cancer Research and Treatment*, *127*(2), 471–478. https://doi.org/10.1007/s10549-010-1173-8

Telloni, S. M. (2017). Tumor Staging and Grading: A Primer. In V. Espina (Ed.), *Molecular Profiling: Methods and Protocols* (pp. 1–17). Springer. https://doi.org/10.1007/978-1-4939-6990-6_1

Telonis, A. G., Rodriguez, D. A., Spanheimer, P. M., Figueroa, M. E., & Goel, N. (2023). Genetic Ancestry-specific Molecular and Survival Differences in Admixed Patients With Breast Cancer. *Annals of Surgery*, *279*(5).

Tevaarwerk, A. J., Wisinski, K. B., & O'Regan, R. M. (2016). Endocrine Therapy in Premenopausal Hormone Receptor–Positive Breast Cancer. *Journal of Oncology Practice*, *12*(11), 1148–1156. https://doi.org/10.1200/JOP.2016.016865

Thomas, D., Gut, B., Karsai, S., Wimmer, A.-B., Wu, K., Wendt-Nordahl, G., Zhang, W., Kathöfer, S., Schoels, W., Katus, H. A., Kiehn, J., & Karle, C. A. (2003). Inhibition of cloned HERG potassium channels by the antiestrogen tamoxifen. *Naunyn-Schmiedeberg's Archives of Pharmacology*, *368*(1), 41–48. https://doi.org/10.1007/s00210-003-0766-8

Thomson, C. A., Rock, C. L., Thompson, P. A., Caan, B. J., Cussler, E., Flatt, S. W., & Pierce, J. P. (2011). Vegetable intake is associated with reduced breast cancer recurrence in tamoxifen users: A secondary analysis from the Women's Healthy Eating and Living Study. *Breast Cancer Research and Treatment*, *125*(2), 519–527. https://doi.org/10.1007/s10549-010-1014-9

Tinsley, E., Bredin, P., Toomey, S., Hennessy, B. T., & Furney, S. J. (2024). *KMT2C* and *KMT2D* aberrations in breast cancer. *Trends in Cancer*, *10*(6), 519–530. https://doi.org/10.1016/j.trecan.2024.02.003

Todd, J. H., Dowle, C., Williams, M. R., Elston, C. W., Ellis, I. O., Hinton, C. P., Blamey, R. W., & Haybittle, J. L. (1987). Confirmation of a prognostic index in primary breast cancer. *British Journal of Cancer*, *56*(4), 489–492. https://doi.org/10.1038/bjc.1987.230

Tokunaga, K., Ohashi, J., Bannai, M., & Juji, T. (2001). Genetic link between Asians and native Americans: Evidence from HLA genes and haplotypes. *Human Immunology*, *62*(9), 1001–1008. https://doi.org/10.1016/S0198-8859(01)00301-9

Turner, L. B. (2011). A meta-analysis of fat intake, reproduction, and breast cancer risk: An evolutionary perspective. *American Journal of Human Biology*, *23*(5), 601–608. https://doi.org/10.1002/ajhb.21176

Uscanga-Perales, G. I., Santuario-Facio, S. K., & Ortiz-López, R. (2016). Triple negative breast cancer: Deciphering the biology and heterogeneity. *Medicina Universitaria*, *18*(71), 105–114. https://doi.org/10.1016/j.rmu.2016.05.007

Van Alsten, S. C., Love, M. I., Calhoun, B. C., Butler, E. N., Perou, C. M., Hoadley, K. A., & Troester, M. A. (2025). Genomic Analysis Reveals Racial and Age-Related Differences in the Somatic Landscape of Breast Cancer and the

Association with Socioeconomic Factors. *Cancer Research*, *85*(7), 1327–1340. https://doi.org/10.1158/0008-5472.CAN-24-1788

van Bommel, M. H. D., IntHout, J., Veldmate, G., Kets, C. M., de Hullu, J. A., van Altena, A. M., & Harmsen, M. G. (2022). Contraceptives and cancer risks in BRCA1/2 pathogenic variant carriers: A systematic review and meta-analysis. *Human Reproduction Update*, *29*(2), 197–217. https://doi.org/10.1093/humupd/dmac038

van 't Veer, L. J., Dai, H., van de Vijver, M. J., He, Y. D., Hart, A. A. M., Mao, M., Peterse, H. L., van der Kooy, K., Marton, M. J., Witteveen, A. T., Schreiber, G. J., Kerkhoven, R. M., Roberts, C., Linsley, P. S., Bernards, R., & Friend, S. H. (2002). Gene expression profiling predicts clinical outcome of breast cancer. *Nature*, *415*(6871), 530–536. https://doi.org/10.1038/415530a

Veronesi, U., Cascinelli, N., Mariani, L., Greco, M., Saccozzi, R., Luini, A., Aguilar, M., & Marubini, E. (2002). Twenty-Year Follow-up of a Randomized Study Comparing Breast-Conserving Surgery with Radical Mastectomy for Early Breast Cancer. *New England Journal of Medicine*, *347*(16), 1227–1232. https://doi.org/10.1056/NEJMoa020989

Veronesi, U., Paganelli, G., Galimberti, V., Viale, G., Zurrida, S., Bedoni, M., Costa, A., Cicco, C. de, Geraghty, J. G., Luini, A., Sacchini, V., & Veronesi, P. (1997). Sentinel-node biopsy to avoid axillary dissection in breast cancer with clinically negative lymph-nodes. *The Lancet*, *349*(9069), 1864–1867. https://doi.org/10.1016/S0140-6736(97)01004-0

Veronesi, U., Saccozzi, R., Vecchio, M. D., Banfi, A., Clemente, C., Lena, M. D., Gallus, G., Greco, M., Luini, A., Marubini, E., Muscolino, G., Rilke, F., Salvadori, B., Zecchini, A., & Zucali, R. (1981). Comparing Radical Mastectomy with Quadrantectomy, Axillary Dissection, and Radiotherapy in Patients with Small Cancers of the Breast. *New England Journal of Medicine*, *305*(1), 6–11. https://doi.org/10.1056/NEJM198107023050102

Vishnubalaji, R., Elango, R., & Alajez, N. M. (2022). LncRNA-Based Classification of Triple Negative Breast Cancer Revealed Inherent Tumor Heterogeneity and Vulnerabilities. *Non-Coding RNA*, *8*(4), 44. https://doi.org/10.3390/ncrna8040044

Vranic, S., Feldman, R., & Gatalica, Z. (2017). Apocrine carcinoma of the breast: A brief update on the molecular features and targetable biomarkers. *Biomolecules and Biomedicine*, *17*(1), 9–11. https://doi.org/10.17305/bjbms.2016.1811

Vrieling, A., Buck, K., Seibold, P., Heinz, J., Obi, N., Flesch-Janys, D., & Chang-Claude, J. (2013). Dietary patterns and survival in German postmenopausal breast cancer survivors. *British Journal of Cancer*, *108*(1), 188–192. https://doi.org/10.1038/bjc.2012.521

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

Walia, V., Ding, M., Kumar, S., Nie, D., Premkumar, L. S., & Elble, R. C. (2009). hCLCA2 Is a p53-Inducible Inhibitor of Breast Cancer Cell Proliferation. *Cancer Research*, *69*(16), 6624–6632. https://doi.org/10.1158/0008-5472.CAN-08-4101

Walia, V., Yu, Y., Cao, D., Sun, M., McLean, J. R., Hollier, B. G., Cheng, J., Mani, S. A., Rao, K., Premkumar, L., & Elble, R. C. (2012). Loss of breast epithelial marker hCLCA2 promotes epithelial-to-mesenchymal transition and indicates higher risk of metastasis. *Oncogene*, *31*(17), 2237–2246. https://doi.org/10.1038/onc.2011.392

Wang, G., Qu, F., Zhou, J., Zhu, B., & Gao, Y. (2023). Elevated THBS3 predicts poor overall survival for clear cell renal cell carcinoma and identifies LncRNA/RBP/THBS3 mRNA networks. *Cell Cycle*, *22*(3), 316–330. https://doi.org/10.1080/15384101.2022.2117910

Wang, J., John, E. M., Horn-Ross, P. L., & Ingles, S. A. (2008). Dietary Fat, Cooking Fat, and Breast Cancer Risk in a Multiethnic Population. *Nutrition and Cancer*, *60*(4), 492–504. https://doi.org/10.1080/01635580801956485

Wang, Y., & Liu, S. (2021). LncRNA GHET1 Promotes Hypoxia-Induced Glycolysis, Proliferation, and Invasion in Triple-Negative Breast Cancer Through the Hippo/YAP Signaling Pathway. *Frontiers in Cell and Developmental Biology*, *9*, 643515. https://doi.org/10.3389/fcell.2021.643515

Weinstock, C., Campassi, C., Goloubeva, O., Wooten, K., Kesmodel, S., Bellevance, E., Feigenberg, S., Ioffe, O., & Tkaczuk, K. H. R. (2015). Breast magnetic resonance imaging (MRI) surveillance in breast cancer survivors. *SpringerPlus*, *4*(1), 459. https://doi.org/10.1186/s40064-015-1158-5

Weisman, P. S., Ng, C. K. Y., Brogi, E., Eisenberg, R. E., Won, H. H., Piscuoglio, S., De Filippo, M. R., Ioris, R., Akram, M., Norton, L., Weigelt, B., Berger, M. F., Reis-Filho, J. S., & Wen, H. Y. (2016). Genetic Alterations of Triple Negative Breast Cancer By Targeted Next Generation Sequencing And Correlation With Tumor Morphology. *Modern Pathology : An Official Journal of the United States and Canadian Academy of Pathology, Inc*, *29*(5), 476–488. https://doi.org/10.1038/modpathol.2016.39

Wickham, H. (2016). Data Analysis. In H. Wickham (Ed.), *Ggplot2: Elegant Graphics for Data Analysis* (pp. 189–201). Springer International Publishing. https://doi.org/10.1007/978-3-319-24277-4_9

Williams, L. A., Hoadley, K. A., Nichols, H. B., Geradts, J., Perou, C. M., Love, M. I., Olshan, A. F., & Troester, M. A. (2019). Differences in race, molecular and tumor characteristics among women diagnosed with invasive ductal and lobular breast carcinomas. *Cancer Causes & Control: CCC*, *30*(1), 31–39. https://doi.org/10.1007/s10552-018-1121-1

Wishart, G. C., Azzato, E. M., Greenberg, D. C., Rashbass, J., Kearins, O., Lawrence, G., Caldas, C., & Pharoah, P. D. (2010). PREDICT: A new UK prognostic model that predicts survival following surgery for invasive breast cancer. *Breast Cancer Research : BCR*, *12*(1), R1. https://doi.org/10.1186/bcr2464

Wishart, G. C., Bajdik, C. D., Dicks, E., Provenzano, E., Schmidt, M. K., Sherman, M., Greenberg, D. C., Green, A. R., Gelmon, K. A., Kosma, V.-M., Olson, J. E., Beckmann, M. W., Winqvist, R., Cross, S. S., Severi, G., Huntsman, D., Pylkäs, K., Ellis, I., Nielsen, T. O., … Pharoah, P. D. P. (2012). PREDICT Plus: Development and validation of a prognostic model for early breast cancer that includes HER2. *British Journal of Cancer*, *107*(5), 800–807. https://doi.org/10.1038/bjc.2012.338

Wöckel, A., Albert, U.-S., Janni, W., Scharl, A., Kreienberg, R., & Stüber, T. (2018). The Screening, Diagnosis, Treatment, and Follow-Up of Breast Cancer. *Deutsches Ärzteblatt International*, *115*(18), 316–323. https://doi.org/10.3238/arztebl.2018.0316

Wong, H.-S., Subramaniam, S., Alias, Z., Taib, N. A., Ho, G.-F., Ng, C.-H., Yip, C.-H., Verkooijen, H. M., Hartman, M., & Bhoo-Pathy, N. (2015). The Predictive Accuracy of PREDICT: A Personalized Decision-Making Tool for Southeast Asian Women With Breast Cancer. *Medicine*, *94*(8), e593. https://doi.org/10.1097/MD.0000000000000593

Wu, H., Wu, Z., Li, H., Wang, Z., Chen, Y., Bao, J., Chen, B., Xu, S., Xia, E., Ye, D., & Dai, X. (2024). Glycosylphosphatidylinositol anchor biosynthesis pathway-based biomarker identification with machine learning for prognosis and T cell exhaustion status prediction in breast cancer. *Frontiers in Immunology*, *15*. https://doi.org/10.3389/fimmu.2024.1392940

Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., Fu, X., Liu, S., Bo, X., & Yu, G. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *The Innovation*, *2*(3), 100141. https://doi.org/10.1016/j.xinn.2021.100141

Xie, F., Zhang, D., Qian, X., Wei, H., Zhou, L., Ding, C., Pan, W., & Ye, Q. (2022). Analysis of cancer-promoting genes related to chemotherapy resistance in esophageal squamous cell carcinoma. *Annals of Translational Medicine*, *10*(2), 92. https://doi.org/10.21037/atm-21-7032

Xu, G., Chhangawala, S., Cocco, E., Razavi, P., Cai, Y., Otto, J. E., Ferrando, L., Selenica, P., Ladewig, E., Chan, C., Da Cruz Paula, A., Witkin, M., Cheng, Y., Park, J., Serna-Tamayo, C., Zhao, H., Wu, F., Sallaku, M., Qu, X., … Toska, E. (2020). ARID1A determines luminal identity and therapeutic response in estrogen-receptor-positive breast cancer. *Nature Genetics*, *52*(2), 198–207. https://doi.org/10.1038/s41588-019-0554-0

Xu, Z., Liu, C., Zhao, Q., Lü, J., Ding, X., Luo, A., He, J., Wang, G., Li, Y., Cai, Z., Wang, Z., Liu, J., Liu, S., Li, W., & Yu, Z. (2020). Long non-coding RNA CCAT2 promotes oncogenesis in triple-negative breast cancer by regulating stemness of cancer cells. *Pharmacological Research*, *152*, 104628. https://doi.org/10.1016/j.phrs.2020.104628

Yadav, R., Kumar, Y., Dahiya, D., & Bhatia, A. (2022). Claudins: The Newly Emerging Targets in Breast Cancer. *Clinical Breast Cancer*, *22*(8), 737–752. https://doi.org/10.1016/j.clbc.2022.09.001

Yang, Q., Wang, J., Yang, H., Fu, Y., & Zhang, X. (2023). Roles of lncRNA in the diagnosis and prognosis of triple-negative breast cancer. *Journal of Zhejiang University. Science. B*, *24*(12), 1123–1140. https://doi.org/10.1631/jzus.B2300067

Yang, X., Amgad, M., Cooper, L. A. D., Du, Y., Fu, H., & Ivanov, A. A. (2020). High expression of MKK3 is associated with worse clinical outcomes in African American breast cancer patients. *Journal of Translational Medicine*, *18*(1), 334. https://doi.org/10.1186/s12967-020-02502-w

Yang, X., Cao, J.-L., Yang, F.-N., Li, X.-F., Tao, L.-M., & Wang, F. (2021). Decreased expression of CLCA2 and the correlating with immune infiltrates in patients with cervical squamous cell carcinoma: A bioinformatics analysis. *Taiwanese Journal of Obstetrics and Gynecology*, *60*(3), 480–486. https://doi.org/10.1016/j.tjog.2021.03.016

Yao, S., Cheng, T.-Y. D., Elkhanany, A., Yan, L., Omilian, A., Abrams, S. I., Evans, S., Hong, C.-C., Qi, Q., Davis, W., Liu, S., Bandera, E. V., Odunsi, K., Takabe, K., Khoury, T., & Ambrosone, C. B. (2021). Breast Tumor Microenvironment in Black Women: A Distinct Signature of CD8+ T-Cell Exhaustion. *JNCI Journal of the National Cancer Institute*, *113*(8), 1036. https://doi.org/10.1093/jnci/djaa215

Yao, S., Graham, K., Shen, J., Campbell, L. E. S., Singh, P., Zirpoli, G., Roberts, M., Ciupak, G., Davis, W., Hwang, H., Khoury, T., Bovbjerg, D. H., Jandorf, L., Pawlish, K. S., Bandera, E. V., Liu, S., Ambrosone, C. B., & Zhao, H. (2013). Genetic variants in microRNAs and breast cancer risk in African American and European American women. *Breast Cancer Research and Treatment*, *141*(3), 447–459. https://doi.org/10.1007/s10549-013-2698-4

Yao, S., Wei, L., Hu, Q., Liu, S., Manojlovic, Z., Fiorica, P. N., Long, M., Zirpoli, G. R., Cai, Q., Long, J., Ping, J., Barnard, M. E., Jin, Y., Murakami, M., Wang, J., Zhu, Q., Davis, W., Chen, J., Ondracek, R. P., … Carpten, J. D. (2025). Mutational landscape of triple-negative breast cancer in African American women. *Nature Genetics*, *57*(9), 2166–2176. https://doi.org/10.1038/s41588-025-02322-y

Yu, G., Wang, L.-G., Han, Y., & He, Q.-Y. (2012). clusterProfiler: An R Package for Comparing Biological Themes Among Gene Clusters. *OMICS : A Journal of Integrative Biology*, *16*(5), 284–287. https://doi.org/10.1089/omi.2011.0118

Yu, G., Wang, L.-G., Yan, G.-R., & He, Q.-Y. (2015). DOSE: An R/Bioconductor package for disease ontology semantic and enrichment analysis. *Bioinformatics*, *31*(4), 608–609. https://doi.org/10.1093/bioinformatics/btu684

Zahedi, R., Molavi Vardanjani, H., Baneshi, M. R., Haghdoost, A. A., Malekpour Afshar, R., Ershad Sarabi, R., Tavakoli, F., & Zolala, F. (2020). Incidence trend of breast Cancer in women of eastern Mediterranean region countries from 1998 to 2019: A systematic review and meta-analysis. *BMC Women's Health*, *20*, 53. https://doi.org/10.1186/s12905-020-00903-z

Zhang, J., Zhang, L., Wang, J., Zhao, J., Zhao, X., Zhang, C., Han, P., & Geng, C.

(2022). Long non-coding RNA linc00921 suppresses tumorigenesis and epithelial-to-mesenchymal transition of triple-negative breast cancer via targeting miR-9-5p/LZTS2 axis. *Human Cell*, *35*(3), 909–923. https://doi.org/10.1007/s13577-022-00685-6

Zhang, N., Zhang, H., Chen, T., & Yang, Q. (2017). Dose invasive apocrine adenocarcinoma has worse prognosis than invasive ductal carcinoma of breast: Evidence from SEER database. *Oncotarget*, *8*(15), 24579–24592. https://doi.org/10.18632/oncotarget.15597

Zhang, Q., Xiu, B., Zhang, L., Chen, M., Chi, W., Li, L., Guo, R., Xue, J., Yang, B., Huang, X., Shao, Z.-M., Huang, S., Chi, Y., & Wu, J. (2022). Immunosuppressive lncRNA LINC00624 promotes tumor progression and therapy resistance through ADAR1 stabilization. *Journal for Immunotherapy of Cancer*, *10*(10), e004666. https://doi.org/10.1136/jitc-2022-004666

Zhang, Y., He, Q., Hu, Z., Feng, Y., Fan, L., Tang, Z., Yuan, J., Shan, W., Li, C., Hu, X., Tanyi, J. L., Fan, Y., Huang, Q., Montone, K., Dang, C. V., & Zhang, L. (2016). Long noncoding RNA LINP1 regulates double strand DNA break repair in triple negative breast cancer. *Nature Structural & Molecular Biology*, *23*(6), 522–530. https://doi.org/10.1038/nsmb.3211

Zhao, P., Nairn, A. V., Hester, S., Moremen, K. W., O'Regan, R. M., Oprea, G., Wells, L., Pierce, M., & Abbott, K. L. (2012). Proteomic Identification of Glycosylphosphatidylinositol Anchor-dependent Membrane Proteins Elevated in Breast Carcinoma *. *Journal of Biological Chemistry*, *287*(30), 25230–25240. https://doi.org/10.1074/jbc.M112.339465

Zhao, Y., Mehta, M., Walton, A., Talsania, K., Levin, Y., Shetty, J., Gillanders, E. M., Tran, B., & Carrick, D. M. (2019). Robustness of RNA sequencing on older formalin-fixed paraffin-embedded tissue from high-grade ovarian serous adenocarcinomas. *PLoS ONE*, *14*(5), e0216050. https://doi.org/10.1371/journal.pone.0216050

Zheng, G., Zhang, C., & Zhong, C. (2021). Identification of potential prognostic biomarkers for breast cancer using WGCNA and PPI integrated techniques. *Annals of Diagnostic Pathology*, *50*, 151675. https://doi.org/10.1016/j.anndiagpath.2020.151675

Zheng, Y., Li, S., Tang, H., Meng, X., & Zheng, Q. (2023). Molecular mechanisms of immunotherapy resistance in triple-negative breast cancer. *Frontiers in Immunology*, *14*. https://doi.org/10.3389/fimmu.2023.1153990

Zhou, X., Xiao, C., Han, T., Qiu, S., Wang, M., Chu, J., Sun, W., Li, L., & Lin, L. (2020). Prognostic biomarkers related to breast cancer recurrence identified based on Logit model analysis. *World Journal of Surgical Oncology*, *18*, 254. https://doi.org/10.1186/s12957-020-02026-z

Zhu, Y., Long, M., Shi, W., He, T., Xie, F., Qian, A., Liu, Y., Hu, T., & Tang, S. (2024). *Proteomics Landscape of Triple-Negative Apocrine Breast Carcinoma Reveals Molecular Mechanisms of Tumorigenesis and Characteristics of Chemo-insensitivity*.

Zucca-Matthes, G., Urban, C., & Vallejo, A. (2016). Anatomy of the nipple and breast ducts. *Gland Surgery*, *5*(1), 32–36. https://doi.org/10.3978/j.issn.2227-684X.2015.05.10

8. Appendices

Appendix A

This appendix contains R scripts used for data analysis.

Script for PCA plotting

```
library(tidyverse)

library(ggplot2)

library(dplyr)

library(ggfortify)

library(reshape2)

library(factoextra)

library(ggpubr)

library(ggcorrplot)

library(PCAtools)


#load gene expression and clinical data

df <- read.csv("Geneexp.csv", header=TRUE, row.names = 1)

samples <- read.csv("clinical.csv", header=TRUE, row.names = 1)

df_clean <- df_numeric[, apply(df_numeric, 2, function(x) var(x, na.rm = TRUE) !=
0)]

pca <- prcomp(t(df_clean), scale. = TRUE)

#plots by race or subtype

biplot(p, x="PC1", y="PC2", showLoadings = FALSE, colby = "race", colkey =
c("lightgreen", "lightblue", "brown1"), lab = "", pointSize = 5, labSize = 7,
max.overlaps = 100, legendPosition = "right")

biplot(p, x="PC1", y="PC2", showLoadings = FALSE, colby = "subtype", colkey =
c("brown2", "chartreuse3","lightblue", "deepskyblue", "darkorchid3"), lab = "",
pointSize = 5, labSize = 7, max.overlaps = 100, legendPosition = "right")


Download, processing and differential expression analysis

library(TCGAbiolinks)
```

```
library(dplyr)
library(stringr)
library(DESeq2)
library(ggplot2)
library(ggrepel)
library(tidyr)
library(PCAtools)
library(ggalt)

#list of samples from TCGA (txt1, txt2, txt3 are respectively list of Asian, Black or African American and White patients), data processing
txt1 <-“barcordes of TCGA samples, separated by space”
txt2 <-<-“barcordes of TCGA samples, separated by space”
txt3 <-<-“barcordes of TCGA samples, separated by space”
split_txt1 <- strsplit(txt1, split = "\\s")[[1]]
listSamples1 <- split_txt1
query1 <- GDCquery(
  project = "TCGA-BRCA",
  data.category = "Transcriptome Profiling",
  data.type = "Gene Expression Quantification",
  barcode = listSamples1)
GDCdownload(query1)
BRCA1 <- GDCprepare(query1)
BRCAoutliers1 <- TCGAanalyze_Preprocessing(BRCA1)

split_txt2 <- strsplit(txt2, split = "\\s")[[1]]
listSamples2 <- split_txt2
query2 <- GDCquery(
  project = "TCGA-BRCA",
```

```
  data.category = "Transcriptome Profiling",
  data.type = "Gene Expression Quantification",
  barcode = listSamples2)
GDCdownload(query2)
BRCA2 <- GDCprepare(query2)
BRCAoutliers2 <- TCGAanalyze_Preprocessing(BRCA2)

split_txt3 <- strsplit(txt3, split = "\\s")[[1]]
listSamples3 <- split_txt3
query3 <- GDCquery(
  project = "TCGA-BRCA",
  data.category = "Transcriptome Profiling",
  data.type = "Gene Expression Quantification",
  barcode = listSamples3)
GDCdownload(query3)
BRCA3 <- GDCprepare(query3)
BRCAoutliers3 <- TCGAanalyze_Preprocessing(BRCA3)
#union of two files to compared (AB = Asian vs Black or African American)
AB <- cbind(BRCAoutliers1, BRCAoutliers2)

# load clinical data, check the presence of samples barcodes in both dataframes
coldata <- read.csv("coldataAB.csv", row.names = 1)
dim(AB)
all(colnames(AB) %in% rownames(coldata))  # Should be TRUE
coldata <- coldata[colnames(AB), ]
common_samples <- intersect(colnames(AB), rownames(coldata))
colnames(AB)
colnames(AB) <- substr(colnames(AB), 1, 16)
```

```
AB <-as.data.frame(AB)
common_samples <- intersect(colnames(MBW), rownames(coldata))
MBW <- MBW[, common_samples]
coldata <- coldata[common_samples, ]

#Differential expression analysis based on race, normalized by tumor stage
dds <- DESeqDataSetFromMatrix (countData = AB, colData=coldata,
                        design = ~(race + stage))
dds <- DESeq(dds)
res <- results(dds)
plotMA(res, ylim=c(-7,30))

#creation of a table of significant DEGs
gene_table <- data.frame(
 GeneName = rownames(res),
 Log2FoldChange = res$log2FoldChange,
 PValue = res$pvalue,
 AdjustedPValue = res$padj
) %>%
 mutate(
  diffexpressed = case_when(
   Log2FoldChange >= 1.5 & AdjustedPValue < 0.05 ~ "UPREGULATED",
   Log2FoldChange <= -1.5 & AdjustedPValue < 0.05 ~ "DOWNREGULATED",
   TRUE ~ "NOT SIGNIFICANT"
  )
 ) %>%
 filter(diffexpressed != "NOT SIGNIFICANT") %>%
 arrange(Log2FoldChange)
```

```
write.csv(gene_table1, file = "DEGsAB.csv")



#selection of top up and downregulated genes to highlight, plotting volcanoplot

DEG <- gene_table1

filtered_df <- head(DEG, 10)

filtered_df2 <- tail(DEG, 10)

ggplot(data = DEG, aes(x = Log2FoldChange, y = -log10(PValue), col =
diffexpressed, label = rownames(DEG))) +

 geom_vline(xintercept = c(-1.5, 1.5 ), col = "gray", linetype = 'dashed') +

 geom_hline(yintercept = -log10(0.05), col = "gray", linetype = 'dashed') +

 geom_point(size = 2) +

 scale_color_manual(values = c("deepskyblue", "grey", "darkorange"),

                    labels = c("Downregulated", "Not significant", "Upregulated")) +

 scale_x_continuous(breaks = seq(-10, 10, 2), limits = c(-12, 12)) +

 scale_y_continuous(limits = c(0, 305)) +

 labs(color = 'Severe',

    x = expression("log"[2]*"FC"),

    y = expression("-log"[10]*"p-value")) +

 theme_classic()+

 theme(axis.text=element_text(size=10), legend.text = element_text(size=20),
axis.title=element_text(size=15, face="bold.italic"))+

 geom_text(data = filtered_df2, aes(Log2FoldChange, y=-log10(PValue), label =
GeneName), size = 5, color= "darkred")+

 geom_text(data = filtered_df, aes(Log2FoldChange, y=-log10(PValue), label =
GeneName), size = 5, color= "darkblue")
```

Script for enrichment analysis

```
library(DOSE)

library(enrichplot)
```

```
library(ggupset)
library(pheatmap)
library(clusterProfiler)
library(org.Hs.eg.db)
library(RColorBrewer) # for a colourful plot
library(ggplot2)
library(europepmc)
library(scales)
library(pathview)

# load DEGs file
df <- read.csv("DegluminalBBW.csv", header=TRUE)
original_gene_list <- df$Log2FoldChange
names(original_gene_list) <- df$GeneName
gene_list<-na.omit(original_gene_list)

# order genes
gene_list = sort(gene_list, decreasing = TRUE)
#selezione dell'organismo di riferimento
organism <- org.Hs.eg.db
#enrichment Gene Ontology (ont can be “BP”, “CC”, “MF”, “ALL”)
gse <- gseGO(geneList=gene_list,
        ont ="BP",
        keyType = "SYMBOL",
        nPerm = 10000,
        minGSSize = 3,
        maxGSSize = 800,
        pvalueCutoff = 0.05,
```

```
        verbose = TRUE,
        OrgDb = organism,
        pAdjustMethod = "none")
#plots
dotplot(gse, showCategory=5, split=".sign", font.size=20) + facet_grid(.~.sign)
enrich <- pairwise_termsim(gse)
emapplot(enrich, showCategory = 15)
# categorySize can be either 'pvalue' or 'geneNum'
cnetplot(gse, categorySize="pvalue", foldChange=gene_list, showCategory = 5,
color_category="bisque1",
      color.params=list(foldChange = df$Log2FoldChange), circular = FALSE,
colorEdge = TRUE)
ridgeplot(gse, showCategory = 22) + labs(x = "enrichment distribution")



# convert gene names in ENTREZID
ids<-bitr(df$GeneName, fromType = "SYMBOL", toType = "ENTREZID",
OrgDb=organism)
dedup_ids = ids[!duplicated(ids[c("SYMBOL")]),]
df2 = df[df$GeneName %in% dedup_ids$SYMBOL,]


duplicated_rows <- duplicated(df2$GeneName)
sum(duplicated_rows)  # Count duplicates
df2 <- df2[!duplicated(df2$GeneName), ]


df2$GeneName = dedup_ids$ENTREZID


# Creation of list of genes with ENTREZID nomenclature
kegg_gene_list <- df2$Log2FoldChange
names(kegg_gene_list) <- df2$GeneName
```

```
kegg_gene_list = sort(kegg_gene_list, decreasing = TRUE)

kegg_organism = "hsa"

#Enrichment

kk2 <- gseKEGG(geneList     = kegg_gene_list,

           organism     = kegg_organism,

           nPerm        = 10000,

           minGSSize    = 3,

           maxGSSize    = 800,

           pvalueCutoff = 0.05,

           pAdjustMethod = "none",

           keyType      = "ncbi-geneid")

#plots

dotplot(kk2, showCategory = 10 , split=".sign", font.size=20) + facet_grid(.~.sign)

enrichk<-pairwise_termsim(kk2)

emapplot(enrichk)

# categorySize can be either 'pvalue' or 'geneNum'

cnetplot(kk2, categorySize="pvalue", foldChange=gene_list)

ridgeplot(kk2) + labs(x = "enrichment distribution")

entrez_ids <- unique(unlist(strsplit(kk2@result$core_enrichment, "/")))

mapping <- bitr(entrez_ids, fromType = "ENTREZID", toType = "SYMBOL", OrgDb
= org.Hs.eg.db)

kk2@result$core_enrichment <- sapply(strsplit(kk2@result$core_enrichment,
"/"), function(x) {

  paste(mapping$SYMBOL[mapping$ENTREZID %in% x], collapse = "/")

})

scores <- setNames(df$Log2FoldChange, df$GeneName)

color.params = list(foldChange = scores)

p <- cnetplot(kk2, showCategory = 5, color.params=color.params, colorEdge=
TRUE)

min.value <- floor( min(p$data$color, na.rm = TRUE) )
```

```
max.value <- ceiling( max(p$data$color, na.rm = TRUE) )

p2 <- p + scale_color_gradientn(name = "fold change", colours = c("red", "blue"),
values = rescale(c(min.value, 0, max.value)),

                    limits= c(min.value, max.value), breaks=c(min.value , 0,
max.value) )

 p3 <- p + scale_color_gradientn(name = "Log2FoldChange", colours = c("red",
"blue"),

                    values = rescale(c(min.value, 0, max.value)),

                    limits= c(min.value, max.value),

                    breaks=c(min.value , 0, max.value) )

 print(p3)


 pv.out <- pathview(gene.data = kegg_gene_list, pathway.id = "00983",

             species = kegg_organism, kegg.native = T,

             sign.pos = kegg_gene_list)
```

Script for Boxplots

```
library(ggpubr)

library(dplyr)

library(ggplot2)

library(reshape2)

#loading expression and clinical data and join them

geneexpBRCA <- read.csv("geneexpBRCA.csv")

clin <- read.csv("clinical.csv")

df_with_colnames <- rbind(colnames(geneexpBRCA), geneexpBRCA)

colnames(df_with_colnames)[1] <- "Original_Column_Names"

transposed_data <- t(df_with_colnames)

transposed_df <- as.data.frame(transposed_data, stringsAsFactors = FALSE)

colnames(transposed_df) <- transposed_df[1, ]
```

```
transposed_df <- transposed_df[-1, ]
rownames(transposed_df) <- NULL
genest<- transposed_df
geneclin<-left_join(clin, genest, by=c("PATIENT_ID"="gene"))

#delete rows with empy cells
df_clean <- na.omit(geneclin)
rows_with_empty <- apply(df_clean, 1, function(row) any(row == ""))
df_cleaned <- df_clean[!rows_with_empty, ]
osdata <- df_cleaned
original_colnames <- colnames(df_clean)
write.csv(osdata, "osdata.csv")
#load list of genes to plot
gene <- read.csv("filenames.csv")
gene_list <- gene[,1]
#filter, when necessary, by subtype or ethnicity
osdatatnbc <- osdata %>%
 filter(subtype == "Basal")
osdatatnbc <- osdata %>%
 filter(race == "Asian")

#filter genes of interest according to the list of genes
filtered_df <- subset(geneexpBRCA, gene %in% gene_list)
genes <- setdiff(colnames(gene), "race")
#crea e salva i plot per ogni gene, dividendoli per etnia e sottotipo
for (gene in gene_list) {
 df_gene <- merged_df %>%
  filter(gene == !!gene) %>%
```

```
  filter(!is.na(subtype), !is.na(race))
 if (nrow(df_gene) == 0) {
  message(paste("No data available for gene:", gene))
  next
 }
 #plot and save boxplots, with different colors according to ethnicity/race
 plot <- ggplot(df_gene, aes(x = subtype, y = count, fill = race)) +
  geom_boxplot() +
  theme_classic() +
  scale_x_discrete(labels = function(x) gsub("\\.", " ", x)) +
  scale_fill_manual(values = c("lightgreen", "lightblue", "red")) +
  labs(
   x = "",
   y = "Gene Expression",
   title = paste("Boxplot of", gene, "Expression by Subtype and Race")
  ) +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))

 ggsave(filename = paste0("boxplot_", gene, ".png"), plot = plot, width = 8, height
= 6)
}

 #plot and save boxplots, with different colors according to subtype
for (gene in gene_list) {
 df_gene <- merged_df %>%
  filter(gene == !!gene) %>%
  filter(!is.na(subtype), !is.na(race))
 if (nrow(df_gene) == 0) {
  message(paste("No data available for gene:", gene))
```

```
  next
 }
  plot <- ggplot(df_gene, aes(x = race, y = count, fill = subtype)) +
  geom_boxplot() +
  theme_classic() +
  scale_x_discrete(labels = function(x) gsub("\\.", " ", x)) +
  scale_fill_manual(values = c("deeppink", "yellow","lightblue", "deepskyblue"))
+
  labs(
    x = "",
    y = "Gene Expression",
    title = paste("Boxplot of", gene, "Expression by Subtype and Race")
  ) +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))


 ggsave(filename = paste0("2boxplot_", gene, ".png"), plot = plot, width = 8,
height = 6)
}


#boxplot of a specific gene
ggplot(df, aes(x = race, y = SCN1A, fill = race)) +
 geom_boxplot() +
 theme_classic() +
 scale_x_discrete(labels = function(x) gsub("\\.", " ", x)) +
 scale_fill_manual(values = c("lightgreen", "lightblue", "red")) +
 labs(
  x = "Subtype and Race",
  y = "Gene Expression",
  title = paste("Boxplot of Gene Expression for", gene, "by Subtype and Race")
```

```
 ) +
 theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
 facet_wrap(~subtype)
```

Script mutational differences

```
library(TCGAbiolinks)
library(dplyr)
library(stringr)
library(ggplot2)
library(ggrepel)
library(BSgenome.Hsapiens.UCSC.hg38)
library(NMF)
library(tidyr)
library(maftools)

clinical <- read.csv("clinical.csv", stringsAsFactors = FALSE)
clinical_data <- clinical %>%
 filter(subtype %in% c("Basal"))
#list of samples from TCGA (txt1, txt2, txt3 are respectively list of Asian, Black or African American and White patients), data processing
txt1 <-“barcordes of TCGA samples, separated by space”
txt2 <-<-“barcordes of TCGA samples, separated by space”
txt3 <-<-“barcordes of TCGA samples, separated by space”

split_txt1 <- strsplit(txt1, split = "\\s")[[1]]
listSamples1 <- split_txt1
query1 <- GDCquery(
 project = "TCGA-BRCA",
 data.category = "Simple Nucleotide Variation",
```

```
  data.type = "Masked Somatic Mutation",
  barcode = listSamples1)
GDCdownload(query1)
maf1 <- GDCprepare(query1)
maf1 <- read.maf(maf1)
sum1 <- getSampleSummary(maf1)
#plots
plotmafSummary(maf = maf1)
rainfallPlot(maf1, detectChangePoints = TRUE)
math1<-math.score(maf1)
plotVaf(maf1)

split_txt2 <- strsplit(txt2, split = "\\s")[[1]]
listSamples2 <- split_txt2
query2 <- GDCquery(
  project = "TCGA-BRCA",
  data.category = "Simple Nucleotide Variation",
  data.type = "Masked Somatic Mutation",
  barcode = listSamples2)
GDCdownload(query2)
maf2 <- GDCprepare(query2)
maf2 <-read.maf(maf2)
getSampleSummary(maf2)

plotmafSummary(maf = maf2)
rainfallPlot(maf2)
math2<-math.score(maf2)
plotVaf(maf2)
```

```
split_txt3 <- strsplit(txt3, split = "\\s")[[1]]
listSamples3 <- split_txt3
query3 <- GDCquery(
  project = "TCGA-BRCA",
  data.category = "Simple Nucleotide Variation",
  data.type = "Masked Somatic Mutation",
  barcode = listSamples3)
GDCdownload(query3)
maf3 <- GDCprepare(query3)
maf3 <-read.maf(maf3)
getSampleSummary(maf3)
plotmafSummary(maf = maf3)
rainfallPlot(maf3, pointSize = 1)
math1<-math.score(maf1)
plotVaf(maf3)


#plots
oncoplot(maf1)
oncoplot(maf2)
oncoplot(maf3)
coBarplot(maf1, maf2)
coBarplot(maf1, maf3)
coBarplot(maf2, maf3)

#mutation type
titv1 = titv(maf1, useSyn = TRUE)
```

```
titv2 = titv(maf2, useSyn = TRUE)

titv3 = titv(maf3, useSyn = TRUE)


#mutational signatures

tnm1 <- trinucleotideMatrix(maf1, ref_genome =
"BSgenome.Hsapiens.UCSC.hg38")

sig1 <- extractSignatures(tnm1, n = 2, plotBestFitRes = FALSE, parallel = 4,
pConstant = 0.05)

plotSignatures(sig1)

plotApobecDiff(tnm1, maf1, pVal = 0.05, title_size = 1, axis_lwd = 1, font_size =
1.2)

enrich1<-signatureEnrichment(maf1, sig1, minMut = 5, useCNV = FALSE, fn =
NULL)

plotEnrichmentResults(enrich1)


#pathway mutations

drive1 <- oncodrive(maf1)

plotOncodrive(drive1)

drive2 <- oncodrive(maf2)

plotOncodrive(drive2)

drive3 <- oncodrive(maf3)

plotOncodrive(drive3)

pA <- pathways(maf1)

plotPathways(maf1, pathlist = pA)

pB <- pathways(maf2)

plotPathways(maf2, pathlist = pB)

pW <- pathways(maf3)

plotPathways(maf3, pathlist = pW)


#differentially mutated genes
```

```
mafAW <- mafCompare(maf1, maf3)

forestPlot(mafAW)

mafBW <- mafCompare(maf2, maf3)

forestPlot(mafBW)

mafAB <- mafCompare(maf1, maf2)

forestPlot(mafAB)

#barplots

genelist <- read.csv("genelist.csv", header = TRUE, stringsAsFactors = FALSE)

genes_of_interest <- genelist[[1]]

mut1 <- maf1@data %>% select(Tumor_Sample_Barcode, Hugo_Symbol) %>%
mutate(Cohort = "Asian")

mut2 <- maf2@data %>% select(Tumor_Sample_Barcode, Hugo_Symbol) %>%
mutate(Cohort = "Black")

mut3 <- maf3@data %>% select(Tumor_Sample_Barcode, Hugo_Symbol) %>%
mutate(Cohort = "White")

mut_all <- bind_rows(mut1, mut2, mut3)

mut_all$Sample_ID <- substr(mut_all$Tumor_Sample_Barcode, 1, 16)

merged <- merge(mut_all, clinical_data, by.x = "Sample_ID", by.y = "SAMPLE_ID")


# Filter for genes in genelist

filtered <- merged %>% filter(Hugo_Symbol %in% genes_of_interest)

total_samples <- clinical_data %>%

 group_by(race) %>%

 summarise(Total = n())

mutation_counts <- merged %>%

 filter(Hugo_Symbol %in% genes_of_interest) %>%

 group_by(Hugo_Symbol, race) %>%

 summarise(Mutated = n_distinct(Sample_ID)) %>%

 left_join(total_samples, by = "race") %>%

 mutate(Percent = round(100 * Mutated / Total, 1))
```

```
races <- unique(clinical_data$race)

genes <- genes_of_interest


# Create full grid

full_grid <-expand.grid(Hugo_Symbol = genes, race = races, stringsAsFactors =
FALSE)

mutation_counts <- merged %>%

 filter(Hugo_Symbol %in% genes) %>%

 group_by(Hugo_Symbol, race) %>%

 summarise(Mutated = n_distinct(Sample_ID), .groups = "drop")

total_samples <- clinical_data %>%

 group_by(race) %>%

 summarise(Total = n(), .groups = "drop")

plot_data <- full_grid %>%

 left_join(mutation_counts, by = c("Hugo_Symbol", "race")) %>%

 left_join(total_samples, by = "race") %>%

 mutate(Mutated = replace_na(Mutated, 0),

      Percent = round(100 * Mutated / Total, 1))

ggplot(plot_data, aes(x = reorder(Hugo_Symbol, -Percent), y = Percent, fill =
race)) +

 geom_bar(stat = "identity", position = "dodge", color="black") +

 geom_text(aes(label = paste0(Percent, "%")),

        position = position_dodge(width = 0.9),

        vjust = -0.5, size = 6) +

 scale_fill_manual(values = c("lightgreen", "lightblue", "red")) +

 theme_minimal() +

 labs(x = "Gene", y = "Mutated Sample %") +

 theme(axis.text.x = element_text(angle = 45, hjust = 1))+

 theme_classic()
```

```
#barplots of mutated pathways
A <- pA %>% mutate(Race = "Asian")
B <- pB %>% mutate(Race = "Black")
W <- pW %>% mutate(Race = "White")

colnames(A)
pathway_data <- bind_rows(A, B, W) %>%
rename(Percent = Fraction_mutated_samples)

ggplot(pathway_data, aes(x = reorder(Pathway, -Percent), y = Percent * 100, fill =
Race)) +
 geom_bar(stat = "identity", position = "dodge", color = "black") +
 geom_text(aes(label = paste0(round(Percent * 100, 1), "%")),
        position = position_dodge(width = 0.9),
        vjust = -0.5, size = 6) +
 scale_fill_manual(values = c("Asian" = "lightgreen", "Black" = "lightblue",
"White" = "red")) +
 theme_minimal() +
 labs(x = "Pathway", y = "Mutated Sample %") +
 theme(axis.text.x = element_text(angle = 45, hjust = 1))+
 theme_classic()
```

Script Kaplan Meier

```
library("survival")
library("survminer")
library("gProfileR")
library("genefilter")
library(ggplot2)
```

```
library(dplyr)


#Combine expression data with clinical data

clinical<-left_join(data_clinical_patient, data_clinical_sample,
by=c("PATIENT_ID"="PATIENT_ID"))

df_with_colnames <- rbind(colnames(geneexpBRCA), geneexpBRCA)

colnames(df_with_colnames)[1] <- "Original_Column_Names"

transposed_data <- t(df_with_colnames)

transposed_df <- as.data.frame(transposed_data, stringsAsFactors = FALSE)

colnames(transposed_df) <- transposed_df[1, ]

transposed_df <- transposed_df[-1, ]

rownames(transposed_df) <- NULL

genest<- transposed_df

geneclin<-left_join(clinical150, genest, by=c("PATIENT_ID"="gene"))

df_clean <- na.omit(geneclin)

rows_with_empty <- apply(df_clean, 1, function(row) any(row == ""))

df_cleaned <- df_clean[!rows_with_empty, ]

osdata <- df_cleaned

original_colnames <- colnames(df_clean)

write.csv(osdata, "osdata.csv")


osdata[, 19:ncol(osdata)] <- lapply(osdata[, 19:ncol(osdata)], as.numeric)

is_numeric <- as.data.frame(sapply(osdata, is.numeric))

print(is_numeric)


original_colnames <- colnames(osdata)

#divides samples in “high” and “low” expression of each gene, based on median
value

osdata <- as.data.frame(lapply(seq_along(osdata), function(i) {
```

```
 if (i <= 18) {
  return(osdata[[i]])
 } else {
  return(ifelse(osdata[[i]] >= median(osdata[[i]], na.rm = TRUE), "high", "low"))
 }
}))
colnames(osdata) <- original_colnames
#list of DEGS to plot
DEGs <- read.csv("degs.csv")
genes <- DEGs[, 1]
gene_list <- genes

#Plots survival curves
for (gene in gene_list) {
 if (gene %in% colnames(osdata)) {
  # Prepare survival object
  surv_object <- Surv(time = osdata$os_time, event = osdata$os_status
  # Check unique values in the gene expression data
  unique_values <- unique(osdata[[gene]])
  # Proceed only if there are at least two groups
  if (length(unique_values) > 1) {
   # Fit survival model
   fit <- survfit(surv_object ~ osdata[[gene]], data = osdata)

   # Perform log-rank test to get the p-value
   surv_diff <- survdiff(surv_object ~ osdata[[gene]], data = osdata)
   pval <- 1 - pchisq(surv_diff$chisq, length(surv_diff$n) - 1)
```

```
  # Proceed if the p-value is less than or equal to 0.05
  if (pval <= 0.05) {
    legend_labels <- if (length(unique_values) == 2) {
      c("High Expression", "Low Expression")
    } else {
      paste("Expression of", gene)
    }

    # Plot and save Kaplan-Meier curve with customized legend
    plot <- ggsurvplot(
      fit,
      data = osdata,
      pval = TRUE,
      xlab = "Time (months)",
      ylab = "Survival probability",
      title = paste("Kaplan-Meier Curve for", gene),
      legend.title = "",
      legend.labs = legend_labels,
      lwd = 4
    )
    # Save the plot
    ggsave(filename = paste0("KM_curve_", gene, ".png"), plot = plot$plot)
  } else {
    print(paste("P-value for gene", gene, "is greater than 0.05. Plot not saved."))
  }
} else {
  print(paste("Gene", gene, "has only one group. Plot not saved."))
}
```

```
  } else {
    print(paste("Gene", gene, "not found in dataframe"))
  }
}
```

Appendix B

This appendix contains tables of differentially expressed genes

Table B1: Top 50 differentially expressed genes between Asian patients and Black patients in the basal subtype

GeneName	Log2FoldChange	PValue	AdjustedPValue	Expression
SCRT2	-7.32052	0.000139	0.008429	DOWNREGULATED
LHFPL3	-6.93264	1.45E-05	0.002089	DOWNREGULATED
AC005550.2	-6.58974	2.22E-08	2.15E-05	DOWNREGULATED
MC3R	-6.49625	0.000125	0.007937	DOWNREGULATED
SLC17A8	-6.35684	2.89E-06	0.000687	DOWNREGULATED
AC004158.1	-5.97977	1.60E-05	0.002235	DOWNREGULATED
AL162384.1	-5.81191	1.32E-06	0.0004	DOWNREGULATED
CHGA	-5.78174	1.74E-11	5.43E-08	DOWNREGULATED
AC093001.1	-5.7136	1.19E-08	1.23E-05	DOWNREGULATED
KCNA6	-5.64793	2.67E-10	4.98E-07	DOWNREGULATED
LRTM2	-5.50505	1.07E-07	6.02E-05	DOWNREGULATED
CLDN10	-5.45087	4.49E-08	3.39E-05	DOWNREGULATED
OTOGL	-5.41498	1.01E-14	5.66E-11	DOWNREGULATED
ST18	-5.40952	1.06E-21	1.48E-17	DOWNREGULATED
LINC01612	-5.40416	5.28E-05	0.004797	DOWNREGULATED
GFI1B	-5.30396	3.11E-13	1.25E-09	DOWNREGULATED
SCG3	-5.23642	3.80E-09	4.84E-06	DOWNREGULATED
RANBP20P	-5.22964	3.43E-06	0.000763	DOWNREGULATED
TEX101	-5.17398	2.48E-08	2.20E-05	DOWNREGULATED
MUC6	-5.05557	5.65E-09	6.60E-06	DOWNREGULATED
TBXT	-4.95424	4.46E-05	0.004416	DOWNREGULATED
SYNPR	-4.86698	0.00011	0.007448	DOWNREGULATED
ENSG00000286742.1	-4.70188	2.95E-11	8.19E-08	DOWNREGULATED
LINC01231	-4.63242	4.47E-08	3.39E-05	DOWNREGULATED
PIANP	-4.61731	3.68E-09	4.84E-06	DOWNREGULATED

KCTD4	-4.60747	4.59E-07	0.000178	DOWNREGULATED
C5orf58	-4.54873	3.73E-14	1.74E-10	DOWNREGULATED
ENSG00000288638.1	-4.54245	0.000157	0.009071	DOWNREGULATED
PCSK2	-4.53308	3.37E-05	0.003716	DOWNREGULATED
AC089998.1	-4.52746	7.37E-06	0.001329	DOWNREGULATED
LINC00052	-4.51628	0.000288	0.012669	DOWNREGULATED
ZP2	-4.49483	0.000161	0.009227	DOWNREGULATED
AC005833.3	-4.42876	2.51E-08	2.20E-05	DOWNREGULATED
ENSG00000286191.1	-4.41988	0.000479	0.016935	DOWNREGULATED
AL021395.1	-4.35439	0.00013	0.008105	DOWNREGULATED
PLP1	-4.27456	0.000173	0.009522	DOWNREGULATED
PCSK1	-4.22524	1.98E-08	1.98E-05	DOWNREGULATED
CACNG5	-4.20534	8.84E-09	9.91E-06	DOWNREGULATED
TDRD1	-4.19759	3.86E-08	3.10E-05	DOWNREGULATED
LINC01158	-4.09298	0.000503	0.017446	DOWNREGULATED
DCX	-4.09265	0.000142	0.008531	DOWNREGULATED
AC116565.1	-4.09103	0.000188	0.010044	DOWNREGULATED
AP000802.1	-4.07982	2.65E-05	0.003173	DOWNREGULATED
SLC7A3	-4.04569	2.06E-05	0.002682	DOWNREGULATED
SPIC	-4.03896	3.16E-06	0.000716	DOWNREGULATED
POU3F3	-4.01742	0.000648	0.020474	DOWNREGULATED
AL591686.2	-3.98182	0.001407	0.032499	DOWNREGULATED
AL354824.2	-3.96904	0.000337	0.013617	DOWNREGULATED
COL19A1	-3.95433	1.17E-06	0.00037	DOWNREGULATED
LINC02037	-3.9532	0.000209	0.010641	DOWNREGULATED
MTCO1P53	4.60854	1.20E-05	0.001849	UPREGULATED
LINC02041	4.665955	0.00064	0.020376	UPREGULATED
SNORC	4.680679	4.38E-07	0.000177	UPREGULATED
LINC00491	4.71134	0.000681	0.020978	UPREGULATED
AC239800.2	4.71149	8.47E-07	0.000293	UPREGULATED
AC112493.1	4.72476	0.001706	0.036767	UPREGULATED
LINC02437	4.738779	1.37E-05	0.002035	UPREGULATED

AMY2A	4.759245	0.001079	0.027993	UPREGULATED
HSPB3	4.764976	0.000568	0.018821	UPREGULATED
ALPP	4.827389	8.36E-05	0.00637	UPREGULATED
RF00019	4.843127	1.33E-10	2.88E-07	UPREGULATED
GABRA5	4.860373	0.00032	0.013302	UPREGULATED
PAUPAR	4.862043	0.002146	0.042071	UPREGULATED
AC107419.1	4.870355	0.000304	0.012934	UPREGULATED
AJ003147.2	4.917827	1.09E-05	0.00178	UPREGULATED
AL365434.1	4.991223	0.002394	0.044067	UPREGULATED
U2AF1L5	5.016203	3.65E-13	1.28E-09	UPREGULATED
ASPG	5.103913	2.12E-07	0.000103	UPREGULATED
KLK11	5.111148	2.36E-06	0.000588	UPREGULATED
SPRR3	5.208378	0.001168	0.029507	UPREGULATED
KRT17P3	5.223098	2.96E-07	0.000134	UPREGULATED
ENSG00000285644.1	5.295261	2.39E-07	0.000114	UPREGULATED
FAM133A	5.321075	0.000235	0.011276	UPREGULATED
IGKV1D-27	5.390962	8.66E-06	0.001498	UPREGULATED
TBC1D3D	5.395626	0.000492	0.017151	UPREGULATED
CHP2	5.480967	0.000629	0.02017	UPREGULATED
RTP3	5.490948	0.00099	0.026487	UPREGULATED
KRT14	5.711638	4.02E-09	4.90E-06	UPREGULATED
RPS26P34	5.719358	8.85E-05	0.006671	UPREGULATED
AC021504.1	5.742757	0.001839	0.038443	UPREGULATED
AC073264.3	5.774914	4.30E-07	0.000177	UPREGULATED
RPS26P4	5.877234	8.34E-05	0.00637	UPREGULATED
PRR27	5.908456	0.001205	0.029951	UPREGULATED
TPSP2	5.926503	6.98E-06	0.001287	UPREGULATED
UGT2B17	5.943867	3.17E-06	0.000716	UPREGULATED
IGLVI-70	6.053686	4.42E-07	0.000177	UPREGULATED
LINC01297	6.058951	3.69E-08	3.04E-05	UPREGULATED
AC004870.2	6.115593	7.20E-05	0.005866	UPREGULATED
LINC00668	6.190804	2.79E-06	0.000668	UPREGULATED

GNGT1	6.44227	0.000353	0.014022	UPREGULATED
AL034346.1	6.578895	1.07E-07	6.02E-05	UPREGULATED
LINC00392	6.832694	0.000423	0.015669	UPREGULATED
AC010789.1	6.860946	7.89E-07	0.00028	UPREGULATED
OR56A3	7.373789	6.90E-05	0.005671	UPREGULATED
GSTM1	7.483748	5.89E-11	1.38E-07	UPREGULATED
MINOS1P3	7.876877	6.11E-16	4.28E-12	UPREGULATED
BPIFB1	8.360349	6.20E-10	1.09E-06	UPREGULATED
AL157778.1	9.035519	5.48E-06	0.001066	UPREGULATED
IGHV1-69-2	10.67907	2.76E-17	2.58E-13	UPREGULATED
XAGE2	24.30209	1.79E-40	5.01E-36	UPREGULATED

Table B2: Top 50 differentially expressed genes between Asian patients and White patients in the basal subtype

GeneName	Log2FoldChange	PValue	AdjustedPValue	Expression
MTND1P23	-7.68221	7.84E-35	2.26E-30	DOWNREGULATED
CRISP3	-6.94778	6.81E-08	8.31E-05	DOWNREGULATED
UGT2B11	-5.54151	0.000137	0.024771	DOWNREGULATED
TEX101	-5.5402	5.84E-09	9.90E-06	DOWNREGULATED
SCG3	-5.47985	5.44E-14	2.24E-10	DOWNREGULATED
LHFPL3	-5.28285	0.000134	0.024771	DOWNREGULATED
LINC01231	-5.28137	6.38E-10	1.31E-06	DOWNREGULATED
SHISA3	-5.12103	6.50E-11	1.56E-07	DOWNREGULATED
GFI1B	-4.96793	9.32E-12	2.69E-08	DOWNREGULATED
AC011298.1	-4.92063	8.92E-06	0.004016	DOWNREGULATED
TBXT	-4.78214	4.22E-05	0.012548	DOWNREGULATED
PIANP	-4.71749	1.02E-12	3.66E-09	DOWNREGULATED
KCNH6	-4.64294	5.54E-10	1.23E-06	DOWNREGULATED
SPINK8	-4.54137	3.64E-05	0.011774	DOWNREGULATED
AC004158.1	-4.52515	0.000159	0.026383	DOWNREGULATED
PGAM1P4	-4.24859	0.000162	0.026593	DOWNREGULATED
AC022733.2	-4.22113	0.000105	0.022292	DOWNREGULATED

ST18	-4.20563	7.21E-12	2.31E-08	DOWNREGULATED
TRIM55	-4.16406	2.03E-06	0.001426	DOWNREGULATED
CARM1P1	-4.10234	1.77E-05	0.006548	DOWNREGULATED
PLP1	-3.98173	0.00042	0.047838	DOWNREGULATED
LINC02037	-3.86012	0.000102	0.022292	DOWNREGULATED
CHGA	-3.76042	0.000214	0.031318	DOWNREGULATED
SYT6	-3.74212	1.01E-07	0.000112	DOWNREGULATED
MEGF11	-3.73848	8.87E-08	0.000102	DOWNREGULATED
CCDC92B	-3.72522	4.50E-05	0.013221	DOWNREGULATED
KCNC1	-3.63836	2.27E-09	4.10E-06	DOWNREGULATED
SPIC	-3.61849	0.000139	0.024859	DOWNREGULATED
CACNG5	-3.58295	3.28E-06	0.002054	DOWNREGULATED
PCAT18	-3.57934	0.000437	0.048758	DOWNREGULATED
VNN1	-3.57392	1.48E-08	2.25E-05	DOWNREGULATED
ENSG00000286742.1	-3.53804	6.80E-06	0.003379	DOWNREGULATED
C9orf129	-3.45907	0.000157	0.026383	DOWNREGULATED
ENSG00000285791.1	-3.45401	1.76E-06	0.001266	DOWNREGULATED
C5orf58	-3.44866	1.04E-06	0.000836	DOWNREGULATED
SPOCK3	-3.43701	0.000118	0.023742	DOWNREGULATED
ADCYAP1	-3.42601	0.000411	0.047136	DOWNREGULATED
HCN4	-3.40173	7.24E-05	0.01768	DOWNREGULATED
HMGB1P1	-3.39698	7.47E-09	1.20E-05	DOWNREGULATED
HIST1H1T	-3.3878	9.34E-06	0.004095	DOWNREGULATED
CNGA3	-3.3846	9.38E-06	0.004095	DOWNREGULATED
AC073150.1	-3.31225	1.65E-06	0.001238	DOWNREGULATED
MTCO1P40	-3.30015	1.42E-09	2.73E-06	DOWNREGULATED
SLC35F4	-3.24043	3.61E-05	0.011774	DOWNREGULATED
OTOGL	-3.23173	3.12E-06	0.001996	DOWNREGULATED
FOXD3	-3.17218	0.00015	0.026042	DOWNREGULATED
AC010931.2	-3.09171	4.12E-06	0.002519	DOWNREGULATED
POU3F1	-3.06017	5.46E-06	0.002969	DOWNREGULATED
UTS2	-3.03874	0.000238	0.034147	DOWNREGULATED

ANK1	-3.03393	1.74E-07	0.000179	DOWNREGULATED
AL034346.1	4.111533	0.000388	0.045413	UPREGULATED
AJ003147.2	4.117897	0.0003	0.038168	UPREGULATED
AL358394.1	4.152617	1.23E-05	0.004894	UPREGULATED
ENSG00000287292.1	4.161434	0.000159	0.026383	UPREGULATED
UGT2B17	4.197201	0.000246	0.034809	UPREGULATED
ENSG00000287342.1	4.413936	2.45E-05	0.008593	UPREGULATED
CCDC144A	4.413983	1.38E-06	0.001075	UPREGULATED
AL136018.1	4.480286	0.000256	0.035309	UPREGULATED
BPIFB1	4.55391	6.84E-05	0.01717	UPREGULATED
U2AF1L5	4.601284	1.72E-08	2.48E-05	UPREGULATED
PSG4	4.717033	0.000361	0.04322	UPREGULATED
KLK11	4.745725	8.53E-06	0.003962	UPREGULATED
GABRA3	4.788666	1.22E-05	0.004894	UPREGULATED
AC005829.2	4.792623	0.000159	0.026383	UPREGULATED
RF00019	4.79989	1.29E-14	6.20E-11	UPREGULATED
AC098850.3	4.820377	1.88E-05	0.006845	UPREGULATED
ENSG00000285644.1	4.935442	1.58E-05	0.006078	UPREGULATED
LINC01297	4.94253	2.85E-06	0.001865	UPREGULATED
DISC1FP1	5.029574	0.000258	0.035388	UPREGULATED
MTRNR2L1	5.096595	6.30E-06	0.003244	UPREGULATED
ERVMER61-1	5.11386	8.41E-07	0.000693	UPREGULATED
OR52N2	5.208179	0.000389	0.045413	UPREGULATED
ENSG00000286868.1	5.26022	0.000126	0.024316	UPREGULATED
MAGEA11	5.268782	0.000396	0.045624	UPREGULATED
AC107419.1	5.425095	0.000175	0.02769	UPREGULATED
AC011487.2	5.5607	2.19E-05	0.007788	UPREGULATED
DSG1	5.65797	2.29E-06	0.001569	UPREGULATED
TDRD12	5.674993	3.26E-07	0.000293	UPREGULATED
AC073264.3	5.841284	1.19E-05	0.004894	UPREGULATED
LINC00668	5.976454	5.97E-05	0.016076	UPREGULATED
ENSG00000286358.1	6.007135	5.09E-05	0.014377	UPREGULATED

IGHV1-69-2	6.125069	8.78E-06	0.004015	UPREGULATED
CASC9	6.159277	2.84E-05	0.009404	UPREGULATED
MAPK8IP1P2	6.236529	0.000105	0.022292	UPREGULATED
MINOS1P3	6.251476	6.65E-05	0.01717	UPREGULATED
GNGT1	6.265137	0.000153	0.026252	UPREGULATED
GSTM1	6.282519	3.98E-05	0.012073	UPREGULATED
AC021504.1	6.321309	0.000207	0.031002	UPREGULATED
PRR27	6.414254	0.00017	0.027548	UPREGULATED
SPRR3	6.560668	0.000103	0.022292	UPREGULATED
AC010789.1	6.77601	3.74E-05	0.011961	UPREGULATED
FAM230C	7.061061	0.000155	0.026365	UPREGULATED
MAGEA12	8.645171	9.28E-05	0.020726	UPREGULATED
NAA11	8.873471	6.81E-05	0.01717	UPREGULATED
FTHL17	9.393427	9.50E-05	0.021054	UPREGULATED
CGA	10.29185	3.33E-08	4.57E-05	UPREGULATED
FP236383.3	15.51405	2.17E-19	1.57E-15	UPREGULATED
XAGE2	22.46504	3.47E-33	5.01E-29	UPREGULATED
LINC02377	25.49218	1.28E-16	7.35E-13	UPREGULATED
ENSG00000288657.1	27.53741	2.64E-26	2.53E-22	UPREGULATED

Table B3: Top 50 differentially expressed genes between Black patients and White patients in the basal subtype

GeneName	Log2FoldChange	PValue	AdjustedPValue	Expression
IGLVI-70	-6.05767	7.91E-32	5.00E-28	DOWNREGULATED
MTND1P23	-6.05672	1.19E-45	2.75E-41	DOWNREGULATED
AC132825.1	-5.81777	7.76E-26	2.45E-22	DOWNREGULATED
RPS26P34	-5.38856	2.03E-22	5.50E-19	DOWNREGULATED
RPS26P4	-5.25169	9.57E-19	1.81E-15	DOWNREGULATED
MTCO1P40	-5.0227	1.45E-45	2.75E-41	DOWNREGULATED
MTCO1P53	-4.76478	1.15E-35	9.78E-32	DOWNREGULATED
AC100757.2	-4.56027	3.98E-18	6.56E-15	DOWNREGULATED
MTATP8P2	-4.5166	3.78E-36	4.78E-32	DOWNREGULATED

AC013457.1	-4.30493	1.37E-15	1.79E-12	DOWNREGULATED
AL445685.2	-4.20967	5.32E-13	4.69E-10	DOWNREGULATED
MTND4P24	-3.80737	1.29E-35	9.78E-32	DOWNREGULATED
ENSG00000286266.1	-3.74247	5.25E-11	2.41E-08	DOWNREGULATED
IGLL1	-3.72077	4.07E-14	4.67E-11	DOWNREGULATED
RPL10P6	-3.66444	2.02E-27	7.68E-24	DOWNREGULATED
RTP3	-3.61858	5.09E-06	0.000182	DOWNREGULATED
KRT4	-3.53223	1.54E-10	5.26E-08	DOWNREGULATED
IGHV1OR15-6	-3.46881	4.35E-08	5.31E-06	DOWNREGULATED
IGHV1-69-2	-3.38728	1.09E-07	1.07E-05	DOWNREGULATED
RPS26P38	-3.31974	1.13E-16	1.65E-13	DOWNREGULATED
IGKV1D-27	-3.27866	4.85E-11	2.32E-08	DOWNREGULATED
IGLV1-41	-3.27538	5.26E-12	3.69E-09	DOWNREGULATED
MAGEA9	-3.26869	0.005762	0.031015	DOWNREGULATED
RPL10P9	-3.25145	2.90E-31	1.57E-27	DOWNREGULATED
AC107373.2	-3.16934	0.005627	0.030477	DOWNREGULATED
LINC01749	-3.14399	2.40E-07	1.98E-05	DOWNREGULATED
SNORD3B-1	-3.12166	3.02E-19	6.73E-16	DOWNREGULATED
AL451165.1	-3.08149	0.000331	0.003895	DOWNREGULATED
AC113367.2	-3.04714	7.61E-05	0.001348	DOWNREGULATED
AL445189.2	-3.03165	1.99E-18	3.59E-15	DOWNREGULATED
ACAN	-3.01268	8.72E-14	9.19E-11	DOWNREGULATED
IGKV2-40	-2.99386	1.33E-06	7.00E-05	DOWNREGULATED
IGHV3-64	-2.99084	9.64E-12	6.20E-09	DOWNREGULATED
AL110505.1	-2.97134	4.40E-07	3.08E-05	DOWNREGULATED
MTCO3P22	-2.96653	3.20E-24	9.34E-21	DOWNREGULATED
IGKV1D-43	-2.90779	2.19E-10	7.05E-08	DOWNREGULATED
AC239800.2	-2.86391	1.81E-11	9.93E-09	DOWNREGULATED
AL050403.1	-2.84373	1.19E-05	0.000347	DOWNREGULATED
KRTAP3-1	-2.82754	4.49E-05	0.00091	DOWNREGULATED
IGLV3-19	-2.81684	2.81E-11	1.46E-08	DOWNREGULATED
KRT83	-2.78177	9.69E-14	9.67E-11	DOWNREGULATED

IGLCOR22-2	-2.77821	0.001581	0.01225	DOWNREGULATED
TBC1D3E	-2.74941	6.23E-06	0.000212	DOWNREGULATED
BPIFB1	-2.74385	1.51E-07	1.38E-05	DOWNREGULATED
IGHJ6	-2.73984	2.47E-10	7.75E-08	DOWNREGULATED
AC023157.1	-2.70925	9.54E-16	1.29E-12	DOWNREGULATED
FGFBP2	-2.69722	1.40E-11	8.18E-09	DOWNREGULATED
CYP4F34P	-2.69603	0.000116	0.001811	DOWNREGULATED
ZBTB8OSP2	-2.64984	4.34E-13	4.01E-10	DOWNREGULATED
AC093809.1	-2.60808	2.32E-12	1.73E-09	DOWNREGULATED
LINC01179	2.477181	0.000958	0.008474	UPREGULATED
SLC25A15P2	2.481408	0.000653	0.006433	UPREGULATED
PNPLA5	2.485838	6.98E-05	0.001263	UPREGULATED
UNC5D	2.490741	1.44E-05	0.000398	UPREGULATED
ADGRG7	2.49264	7.98E-06	0.000256	UPREGULATED
ECEL1	2.514708	1.33E-12	1.09E-09	UPREGULATED
C14orf39	2.522598	1.17E-06	6.33E-05	UPREGULATED
PRR27	2.5469	0.001933	0.014109	UPREGULATED
UCN3	2.572802	0.000774	0.007269	UPREGULATED
MUC6	2.619998	6.68E-10	1.69E-07	UPREGULATED
VCX	2.625434	0.00017	0.002405	UPREGULATED
AC074389.2	2.628988	8.73E-05	0.001485	UPREGULATED
TRIML1	2.644932	0.000202	0.00274	UPREGULATED
LINC02197	2.659504	0.000553	0.005691	UPREGULATED
AC073174.1	2.692154	0.000186	0.002581	UPREGULATED
CSN1S1	2.770588	7.97E-05	0.001392	UPREGULATED
WASIR1	2.780971	9.27E-07	5.36E-05	UPREGULATED
CLDN10-AS1	2.810652	0.00032	0.003793	UPREGULATED
GPR151	2.81893	0.000179	0.002505	UPREGULATED
AC093496.1	2.839711	5.28E-05	0.001026	UPREGULATED
SLC6A10P	2.870554	5.62E-05	0.001077	UPREGULATED
AC011294.1	2.952918	1.41E-06	7.32E-05	UPREGULATED
FAM83C	2.986474	1.13E-05	0.000334	UPREGULATED

TDRD1	3.036331	8.45E-09	1.43E-06	UPREGULATED
NPY2R	3.046246	2.99E-06	0.000123	UPREGULATED
AL159987.1	3.069021	0.00175	0.013125	UPREGULATED
MAPK8IP1P2	3.148082	5.26E-06	0.000186	UPREGULATED
PPP4R3C	3.188776	0.00748	0.03753	UPREGULATED
AC005336.3	3.194611	0.001946	0.014173	UPREGULATED
AC111149.2	3.2591	0.006527	0.033923	UPREGULATED
SERPINB13	3.295168	1.78E-07	1.56E-05	UPREGULATED
AP000844.2	3.309297	1.19E-05	0.000347	UPREGULATED
DCX	3.326361	1.05E-11	6.40E-09	UPREGULATED
DCAF4L2	3.339317	0.002125	0.015114	UPREGULATED
CTAG2	3.340764	0.000756	0.007142	UPREGULATED
SMR3B	3.34226	0.006806	0.034941	UPREGULATED
CLCA2	3.462219	1.76E-12	1.42E-09	UPREGULATED
TDRD9	3.513952	2.40E-18	4.14E-15	UPREGULATED
ENSG00000286619.1	3.594953	6.64E-31	3.15E-27	UPREGULATED
PAGE1	3.97087	0.000209	0.002811	UPREGULATED
MTRNR2L1	3.994734	2.80E-17	4.24E-14	UPREGULATED
AL035425.4	4.436849	0.000897	0.008092	UPREGULATED
CGA	4.866055	2.30E-13	2.18E-10	UPREGULATED
RNU1-88P	5.160175	2.93E-06	0.000122	UPREGULATED
RF00017	5.228428	5.93E-05	0.001116	UPREGULATED
RNU1-20P	6.413318	0.00041	0.004565	UPREGULATED
RN7SL801P	7.067791	0.004501	0.025938	UPREGULATED
RNU1-117P	8.292627	0.007988	0.039345	UPREGULATED
RN7SL681P	8.927748	0.004294	0.025058	UPREGULATED
RN7SL504P	10.14151	0.001179	0.00987	UPREGULATED

Table B4: Top 50 differentially expressed genes between Asian patients and Black patients in the HER2+ subtype

GeneName	Log2FoldChange	PValue	AdjustedPValue	Expression
ANKRD18B	-5.23285	0.000104	0.01346	DOWNREGULATED

CEL	-4.42125	2.49E-06	0.001567	DOWNREGULATED
KLK7	-4.29428	5.18E-06	0.002517	DOWNREGULATED
MUC21	-4.22828	3.90E-05	0.008025	DOWNREGULATED
AL353705.4	-4.17635	5.99E-05	0.010118	DOWNREGULATED
EREG	-4.05069	1.21E-09	4.17E-06	DOWNREGULATED
LINC01748	-4.01262	0.000107	0.013671	DOWNREGULATED
SBSN	-3.83141	0.000142	0.015718	DOWNREGULATED
LINC01179	-3.81591	0.000881	0.039652	DOWNREGULATED
CEACAM5	-3.70469	8.52E-05	0.011891	DOWNREGULATED
A2ML1	-3.70168	2.56E-07	0.000314	DOWNREGULATED
ERVV-2	-3.65341	0.001162	0.044604	DOWNREGULATED
EYA1	-3.64951	1.73E-05	0.004912	DOWNREGULATED
DSG3	-3.5874	4.88E-05	0.00894	DOWNREGULATED
AL137845.2	-3.52885	0.000101	0.013256	DOWNREGULATED
ENSG00000287519.1	-3.52672	0.000291	0.02185	DOWNREGULATED
IGLV2-18	-3.44256	9.18E-05	0.01241	DOWNREGULATED
IGLV3-27	-3.4372	0.000191	0.018315	DOWNREGULATED
AP001574.1	-3.4224	0.000289	0.02185	DOWNREGULATED
IGHV3-13	-3.41792	9.86E-06	0.003266	DOWNREGULATED
LINC00704	-3.40733	0.000373	0.025051	DOWNREGULATED
FAM133A	-3.33419	0.000673	0.033935	DOWNREGULATED
PCP4L1	-3.29865	9.41E-05	0.012534	DOWNREGULATED
AC245517.5	-3.29203	0.000141	0.015675	DOWNREGULATED
HS3ST6	-3.24813	0.000905	0.040143	DOWNREGULATED
IGHV1-14	-3.21933	0.000901	0.040128	DOWNREGULATED
SLC12A1	-3.1201	9.24E-05	0.012434	DOWNREGULATED
ETNPPL	-3.07628	0.00025	0.020697	DOWNREGULATED
DCDC2	-3.05241	0.000107	0.013671	DOWNREGULATED
ADGRF2	-3.00728	0.000149	0.016274	DOWNREGULATED
LY6G6C	-2.97854	0.000255	0.020697	DOWNREGULATED
LINC01517	-2.93185	0.000303	0.022273	DOWNREGULATED
GABRP	-2.93173	1.25E-06	0.001004	DOWNREGULATED

C5orf58	-2.91785	0.00024	0.020378	DOWNREGULATED
CLLU1OS	-2.87896	0.000834	0.038648	DOWNREGULATED
IGHV1-17	-2.87419	0.000986	0.041366	DOWNREGULATED
IGHV4-61	-2.85634	0.000775	0.036959	DOWNREGULATED
GCNT3	-2.85571	0.00061	0.032335	DOWNREGULATED
SPX	-2.83662	0.00011	0.013812	DOWNREGULATED
SLC6A14	-2.7361	0.001054	0.04306	DOWNREGULATED
IGLV10-54	-2.70889	5.53E-06	0.002544	DOWNREGULATED
SOSTDC1	-2.67208	7.81E-06	0.002884	DOWNREGULATED
KLRC2	-2.66389	0.000295	0.021901	DOWNREGULATED
AL022324.3	-2.62665	0.000984	0.04132	DOWNREGULATED
CCDC187	-2.61511	0.000635	0.032889	DOWNREGULATED
IGKV1-6	-2.60637	0.00047	0.02841	DOWNREGULATED
ADGRF4	-2.58341	0.000523	0.029894	DOWNREGULATED
ACKR2	-2.53414	0.001319	0.047467	DOWNREGULATED
IGHV3-72	-2.49923	0.000144	0.015775	DOWNREGULATED
IGLV2-8	-2.45116	0.000992	0.041524	DOWNREGULATED
CDC42EP5	3.551004	1.29E-09	4.17E-06	UPREGULATED
RPH3A	3.555289	1.28E-06	0.001004	UPREGULATED
APOD	3.55988	4.24E-05	0.008196	UPREGULATED
AC084759.3	3.573327	0.000485	0.028928	UPREGULATED
AP003716.1	3.610469	0.000247	0.020659	UPREGULATED
GNMT	3.68792	8.10E-07	0.000775	UPREGULATED
RNF126P1	3.691107	3.12E-06	0.001751	UPREGULATED
OSTN-AS1	3.728616	0.001453	0.049237	UPREGULATED
LINC00668	3.788718	0.000546	0.030614	UPREGULATED
AC073172.1	3.800489	0.0008	0.03779	UPREGULATED
AL513304.1	3.808526	0.000501	0.029232	UPREGULATED
PCDH10	3.814619	2.86E-06	0.001677	UPREGULATED
ENSG00000287826.1	3.891486	0.000515	0.02969	UPREGULATED
FAR2P4	3.902055	0.000317	0.022634	UPREGULATED
MKRN4P	3.965069	0.000291	0.02185	UPREGULATED

AC024610.2	3.973655	0.000653	0.033489	UPREGULATED
AL109615.3	3.987505	1.97E-05	0.005234	UPREGULATED
SLC15A5	4.042561	3.05E-05	0.007107	UPREGULATED
RPL10P9	4.057004	4.70E-12	4.04E-08	UPREGULATED
AC079949.1	4.095632	0.000196	0.018597	UPREGULATED
AC112243.1	4.228846	6.88E-05	0.010643	UPREGULATED
CD177	4.236682	3.83E-08	7.61E-05	UPREGULATED
AC131097.2	4.299423	6.43E-08	0.000111	UPREGULATED
ENSG00000287628.1	4.335057	0.00057	0.031218	UPREGULATED
SORD2P	4.358846	2.97E-09	8.52E-06	UPREGULATED
AKR1C2	4.414134	5.29E-06	0.002517	UPREGULATED
FCN2	4.433199	3.99E-07	0.000438	UPREGULATED
ENSG00000287064.1	4.478241	1.28E-08	3.31E-05	UPREGULATED
RPL10P6	4.509859	3.79E-11	1.63E-07	UPREGULATED
MS4A18	4.52281	0.000208	0.019045	UPREGULATED
COX20P1	4.5402	2.76E-11	1.56E-07	UPREGULATED
EDDM3B	4.554126	0.001158	0.044573	UPREGULATED
CHRNA2	4.584348	0.000655	0.033489	UPREGULATED
OFCC1	4.635874	7.11E-05	0.010734	UPREGULATED
SCGB3A1	4.69359	1.30E-07	0.000183	UPREGULATED
KRT13	4.714294	9.08E-05	0.012344	UPREGULATED
CBSL	4.715505	2.66E-05	0.00644	UPREGULATED
LINC00624	4.912374	3.02E-11	1.56E-07	UPREGULATED
SNHG27	4.943382	2.81E-06	0.001677	UPREGULATED
RPS26P34	5.092282	0.000335	0.023539	UPREGULATED
HBQ1	5.121255	2.76E-05	0.006531	UPREGULATED
AC113367.2	5.282371	6.76E-05	0.010643	UPREGULATED
HNRNPA1P57	5.458915	6.64E-05	0.010589	UPREGULATED
EDDM3A	5.498565	2.51E-05	0.006247	UPREGULATED
PCSK2	5.521164	3.82E-05	0.008025	UPREGULATED
ENSG00000287971.1	5.64877	2.14E-06	0.001415	UPREGULATED
CYP4F34P	6.033689	5.35E-05	0.009535	UPREGULATED

ENSG00000287609.1	6.359306	8.28E-08	0.000126	UPREGULATED
IGHV1-69-2	7.573878	2.67E-13	3.44E-09	UPREGULATED
MTND1P23	8.255771	4.03E-19	1.04E-14	UPREGULATED

Table B5: Top 50 differentially expressed genes between Asian patients and White patients in the HER2+ subtype

GeneName	Log2FoldChange	PValue	AdjustedPValue	Expression
AC021504.1	-5.48963	8.54E-07	0.001101	DOWNREGULATED
IGHV7-4-1	-5.25313	2.43E-10	2.40E-06	DOWNREGULATED
ENSG00000286944.1	-4.69059	0.000148	0.037912	DOWNREGULATED
IGHV7-40	-4.62315	0.000136	0.037277	DOWNREGULATED
PRSS1	-4.17553	5.77E-06	0.0045	DOWNREGULATED
AL137845.2	-3.98793	3.48E-10	2.58E-06	DOWNREGULATED
PPBP	-3.95689	1.08E-05	0.006822	DOWNREGULATED
POLD2P1	-3.68272	1.29E-05	0.007628	DOWNREGULATED
ATP2B3	-3.62681	4.21E-06	0.003671	DOWNREGULATED
KLK2	-3.48894	1.09E-08	4.62E-05	DOWNREGULATED
EYA1	-3.28757	3.51E-08	8.79E-05	DOWNREGULATED
AC055854.1	-3.25251	2.05E-06	0.002252	DOWNREGULATED
HMX3	-3.23966	0.000116	0.034129	DOWNREGULATED
ARSF	-3.21717	2.30E-05	0.011015	DOWNREGULATED
LINC00704	-3.21624	2.78E-06	0.002742	DOWNREGULATED
SBK2	-3.16347	4.14E-08	9.44E-05	DOWNREGULATED
IGHV1-14	-3.10883	2.45E-05	0.011015	DOWNREGULATED
ANGPTL7	-3.09742	0.000206	0.046271	DOWNREGULATED
IGKV2-29	-3.09714	0.000228	0.049452	DOWNREGULATED
UTS2	-3.09567	1.62E-07	0.000283	DOWNREGULATED
IGLV2-18	-3.04599	4.56E-06	0.003758	DOWNREGULATED
IGHV3-73	-3.0261	2.13E-05	0.010865	DOWNREGULATED
TGM5	-2.97475	2.28E-06	0.002328	DOWNREGULATED
HS3ST5	-2.95122	3.83E-05	0.015971	DOWNREGULATED
CU638689.4	-2.94174	2.40E-05	0.011015	DOWNREGULATED

AC091812.3	-2.93559	2.46E-05	0.011015	DOWNREGULATED
ADAMTS8	-2.9193	6.88E-08	0.000128	DOWNREGULATED
ERVMER61-1	-2.83374	0.000126	0.035476	DOWNREGULATED
EGR4	-2.81613	0.000119	0.034606	DOWNREGULATED
AC023157.1	-2.63056	0.000114	0.033699	DOWNREGULATED
TRPA1	-2.54988	1.32E-06	0.001632	DOWNREGULATED
ADGRF2	-2.53901	4.54E-05	0.018175	DOWNREGULATED
CHI3L2	-2.45886	9.45E-06	0.006363	DOWNREGULATED
TRIM55	-2.40276	0.000164	0.040166	DOWNREGULATED
ADAMTS7P1	-2.39276	4.62E-09	2.28E-05	DOWNREGULATED
IGHV3-38	-2.30724	0.000102	0.031296	DOWNREGULATED
LINC01468	-2.30186	7.46E-05	0.025695	DOWNREGULATED
PADI4	-2.27631	1.47E-05	0.008052	DOWNREGULATED
PCYT1B	-2.2696	6.24E-05	0.022292	DOWNREGULATED
RNF157-AS1	-2.18324	1.08E-05	0.006822	DOWNREGULATED
AC022784.1	-2.17172	9.49E-05	0.029743	DOWNREGULATED
LINC01010	-2.13086	8.04E-05	0.027084	DOWNREGULATED
AC108676.1	-2.09271	0.00015	0.037912	DOWNREGULATED
TRIM67	-2.08667	1.72E-05	0.009079	DOWNREGULATED
GAPDHP1	-2.06181	9.03E-05	0.029104	DOWNREGULATED
CDH2	-2.01429	9.33E-05	0.029719	DOWNREGULATED
TDO2	-1.90163	1.88E-05	0.009767	DOWNREGULATED
ACBD3-AS1	-1.79335	1.66E-05	0.008947	DOWNREGULATED
GAPDHP38	-1.7664	5.30E-07	0.000748	DOWNREGULATED
AC009509.1	-1.63208	7.75E-05	0.026401	DOWNREGULATED
SCUBE2	1.982768	4.05E-05	0.016662	UPREGULATED
AC094019.1	2.136939	7.40E-06	0.005349	UPREGULATED
GTF2H2B	2.256362	6.03E-05	0.021794	UPREGULATED
ITIH2	2.273385	2.45E-05	0.011015	UPREGULATED
RIMBP2	2.389923	0.000145	0.037912	UPREGULATED
ENSG00000286342.1	2.444523	1.23E-05	0.007489	UPREGULATED
FAM3D	2.511737	0.000193	0.044586	UPREGULATED

GAD1	2.69895	8.40E-06	0.005929	UPREGULATED
MYO18B	2.830913	0.00018	0.043313	UPREGULATED
RF00019	2.841342	1.97E-07	0.000307	UPREGULATED
TPTEP1	2.949295	2.07E-08	6.80E-05	UPREGULATED
ZNF99	2.955134	4.35E-06	0.003682	UPREGULATED
GLYAT	2.989306	8.88E-05	0.028929	UPREGULATED
HRK	3.046884	1.72E-06	0.001958	UPREGULATED
AC012640.3	3.054559	8.88E-05	0.028929	UPREGULATED
ENSG00000287647.1	3.075771	4.72E-05	0.018312	UPREGULATED
FAM95C	3.167616	0.000124	0.035476	UPREGULATED
AC131097.2	3.168308	1.36E-05	0.007772	UPREGULATED
KLK13	3.178369	7.25E-05	0.025274	UPREGULATED
ENSG00000286042.1	3.202737	4.76E-05	0.018312	UPREGULATED
SLC28A2	3.220987	3.80E-06	0.003522	UPREGULATED
AP000547.3	3.300492	3.56E-08	8.79E-05	UPREGULATED
COX20P1	3.32522	5.78E-08	0.000118	UPREGULATED
RPS4XP22	3.383214	1.72E-22	5.09E-18	UPREGULATED
RNF126P1	3.441494	4.15E-06	0.003671	UPREGULATED
MTRNR2L1	3.443195	1.41E-06	0.001671	UPREGULATED
SLC26A9	3.449308	4.51E-05	0.018175	UPREGULATED
TRDN	3.493295	6.39E-06	0.004855	UPREGULATED
LEP	3.611593	1.35E-05	0.007772	UPREGULATED
DMBT1	3.655522	4.92E-06	0.003939	UPREGULATED
FO538757.2	3.690975	9.76E-05	0.030138	UPREGULATED
SYT9	3.710408	2.13E-06	0.002252	UPREGULATED
TREHP1	3.753258	0.000167	0.04051	UPREGULATED
AC005829.2	3.845807	5.51E-05	0.020154	UPREGULATED
BEX1	3.847578	9.33E-06	0.006363	UPREGULATED
KBTBD12	4.00167	6.59E-05	0.02326	UPREGULATED
TDRD12	4.04473	1.31E-08	4.85E-05	UPREGULATED
FMO6P	4.050525	7.39E-06	0.005349	UPREGULATED
KLK11	4.05132	2.75E-05	0.011819	UPREGULATED

AC093525.6	4.055635	4.55E-07	0.000674	UPREGULATED
ENSG00000286619.1	4.718286	1.80E-12	2.67E-08	UPREGULATED
TF	4.777441	1.15E-09	6.84E-06	UPREGULATED
SLC15A5	4.796679	1.08E-05	0.006822	UPREGULATED
UGT1A6	5.012668	0.000147	0.037912	UPREGULATED
CBSL	5.097192	2.82E-08	8.36E-05	UPREGULATED
UGT1A1	5.933691	1.45E-05	0.008052	UPREGULATED
IGHV1-69-2	5.993299	5.95E-08	0.000118	UPREGULATED
NKX2-5	6.524644	7.87E-07	0.00106	UPREGULATED
CYP4F62P	7.556752	0.000138	0.037535	UPREGULATED
AL359458.1	17.10038	1.86E-07	0.000306	UPREGULATED

Table B6: Top 50 differentially expressed genes between Black patients and White patients in the HER2+ subtype

GeneName	Log2FoldChange	PValue	AdjustedPValue	Expression
MTND1P23	-8.14054	4.35E-39	1.30E-34	DOWNREGULATED
IGHV3-73	-6.23112	6.16E-14	2.31E-10	DOWNREGULATED
LY6L	-6.15627	1.64E-05	0.000897	DOWNREGULATED
IGLVI-70	-5.80737	7.09E-12	1.93E-08	DOWNREGULATED
AC114786.2	-5.65519	0.000496	0.008736	DOWNREGULATED
ADAM7	-5.61504	8.62E-08	2.35E-05	DOWNREGULATED
MTCO1P40	-5.54174	2.52E-25	3.78E-21	DOWNREGULATED
CYP4F34P	-5.2102	3.39E-07	6.13E-05	DOWNREGULATED
RPS26P34	-5.15463	1.62E-07	3.66E-05	DOWNREGULATED
CES1P1	-5.05141	4.92E-10	5.26E-07	DOWNREGULATED
CHRNA2	-5.03047	4.53E-05	0.00179	DOWNREGULATED
SLCO1B1	-5.02731	0.000811	0.011928	DOWNREGULATED
AC113367.2	-5.01972	1.58E-06	0.000185	DOWNREGULATED
RPS26P4	-4.97425	9.68E-05	0.00295	DOWNREGULATED
MTCO1P53	-4.57404	2.33E-14	1.16E-10	DOWNREGULATED
PCSK2	-4.46299	5.50E-07	8.50E-05	DOWNREGULATED
SLC26A3	-4.44996	7.97E-09	4.26E-06	DOWNREGULATED

TMIGD1	-4.33137	0.006749	0.049973	DOWNREGULATED
SORD2P	-4.28774	1.78E-19	1.77E-15	DOWNREGULATED
HPR	-4.27072	1.05E-07	2.76E-05	DOWNREGULATED
RPL10P6	-4.25771	3.87E-09	2.58E-06	DOWNREGULATED
CD177	-4.25159	1.43E-09	1.22E-06	DOWNREGULATED
RPL15P18	-4.18049	8.97E-11	1.34E-07	DOWNREGULATED
SULT1C2P2	-4.15522	0.002051	0.022573	DOWNREGULATED
AC087379.2	-4.10526	0.000999	0.013688	DOWNREGULATED
AC016885.3	-4.06885	9.73E-06	0.000641	DOWNREGULATED
ENSG00000286944.1	-4.04729	0.002188	0.023578	DOWNREGULATED
CLEC4GP1	-3.95935	5.40E-06	0.000429	DOWNREGULATED
RPS26P38	-3.92499	3.38E-09	2.38E-06	DOWNREGULATED
LCN8	-3.91539	6.08E-05	0.002144	DOWNREGULATED
RPL10P9	-3.89851	3.60E-10	4.15E-07	DOWNREGULATED
SERHL	-3.81568	1.40E-09	1.22E-06	DOWNREGULATED
PHGR1	-3.7461	0.000121	0.003386	DOWNREGULATED
SULT1C2P1	-3.74178	0.000622	0.010203	DOWNREGULATED
FGA	-3.69655	0.003019	0.029054	DOWNREGULATED
SORCS3	-3.68154	1.25E-05	0.000756	DOWNREGULATED
ENSG00000287852.1	-3.64047	0.001182	0.015374	DOWNREGULATED
AC010624.4	-3.61118	9.59E-06	0.000636	DOWNREGULATED
ANKRD34C-AS1	-3.59315	0.000129	0.003551	DOWNREGULATED
AL033397.2	-3.46987	1.17E-08	5.57E-06	DOWNREGULATED
AL136987.1	-3.45328	0.000395	0.007481	DOWNREGULATED
ARSF	-3.44899	3.00E-05	0.001353	DOWNREGULATED
MAFA-AS1	-3.43358	0.000118	0.003336	DOWNREGULATED
LINC01518	-3.4003	9.95E-05	0.003003	DOWNREGULATED
AP001284.1	-3.38783	0.000165	0.004134	DOWNREGULATED
AC010255.3	-3.3303	2.39E-07	4.90E-05	DOWNREGULATED
CDC42EP5	-3.31995	9.66E-15	7.24E-11	DOWNREGULATED
RPL31P28	-3.31919	0.00035	0.006851	DOWNREGULATED
ENSG00000287628.1	-3.2917	0.000208	0.004812	DOWNREGULATED

EDDM3B	-3.27282	0.00414	0.036046	DOWNREGULATED
PSORS1C3	2.650551	0.005367	0.042914	UPREGULATED
AL353705.4	2.652876	0.006656	0.049499	UPREGULATED
MTRNR2L1	2.670028	0.000742	0.011297	UPREGULATED
LINC01559	2.67745	0.003279	0.030742	UPREGULATED
DEFB1	2.69904	7.47E-05	0.002489	UPREGULATED
ERVV-2	2.70101	0.005783	0.045251	UPREGULATED
A2ML1	2.7133	6.45E-05	0.002238	UPREGULATED
RNA5SP18	2.746407	0.000252	0.005489	UPREGULATED
ENSG00000286619.1	2.767686	3.40E-05	0.001469	UPREGULATED
AC084859.1	2.775367	0.000498	0.008758	UPREGULATED
DUSP5P1	2.805899	6.68E-05	0.002291	UPREGULATED
ENSG00000287519.1	2.811225	0.001388	0.017122	UPREGULATED
GABRP	2.818523	2.14E-05	0.001083	UPREGULATED
SLC5A5	2.830515	0.000584	0.009773	UPREGULATED
SLC30A10	2.831199	0.002255	0.024085	UPREGULATED
SMYD1	2.832093	0.002934	0.028535	UPREGULATED
KLK13	2.837694	0.000683	0.010785	UPREGULATED
AC007563.3	2.837763	0.00229	0.024325	UPREGULATED
FMO6P	2.839622	0.000776	0.011594	UPREGULATED
EREG	2.873498	3.45E-06	0.00032	UPREGULATED
LINC01133	2.913958	1.08E-05	0.000685	UPREGULATED
AF064860.1	2.973283	0.000388	0.007366	UPREGULATED
AP000851.1	2.990601	0.000896	0.012762	UPREGULATED
LINC00898	2.998798	9.01E-05	0.002827	UPREGULATED
CHST9	3.016705	0.000104	0.003094	UPREGULATED
AC087257.2	3.037667	0.000957	0.013297	UPREGULATED
IGLVIVOR22-1	3.063369	0.001336	0.016691	UPREGULATED
GSTA3	3.070329	0.004522	0.038299	UPREGULATED
AC116614.1	3.091045	0.001709	0.019865	UPREGULATED
AC079062.1	3.198629	0.006613	0.049262	UPREGULATED
KBTBD12	3.236572	0.000415	0.00773	UPREGULATED

AC117386.2	3.255779	0.00507	0.0414	UPREGULATED
PROKR1	3.26056	0.001122	0.014858	UPREGULATED
AP001783.1	3.272936	0.00127	0.016125	UPREGULATED
TBX10	3.320034	0.000145	0.003798	UPREGULATED
LINC00958	3.346044	2.49E-05	0.001192	UPREGULATED
SLC26A9	3.370092	3.96E-05	0.00163	UPREGULATED
DCDC2	3.373644	8.39E-06	0.000582	UPREGULATED
SLC13A2	3.689204	7.19E-06	0.000524	UPREGULATED
CCNG2P1	3.847608	0.00346	0.031844	UPREGULATED
ERICH5	4.010616	1.42E-05	0.000815	UPREGULATED
BMPR1B-AS1	4.042234	5.82E-05	0.002083	UPREGULATED
MIR663AHG	4.055542	0.002226	0.023884	UPREGULATED
TF	4.10513	5.65E-08	1.66E-05	UPREGULATED
PRKAG3	4.153476	0.000104	0.003089	UPREGULATED
LINC01224	4.303037	9.71E-06	0.000641	UPREGULATED
CRISP3	4.370039	9.80E-05	0.002968	UPREGULATED
TEX13C	4.442638	0.001742	0.020126	UPREGULATED
UGT1A6	5.509567	0.00014	0.003733	UPREGULATED
UGT1A1	6.168225	1.52E-05	0.000854	UPREGULATED

Table B7: Top 50 differentially expressed genes between Asian patients and Black patients in the luminal A subtype

GeneName	Log2FoldChange	PValue	AdjustedPValue	Expression
SMR3B	-4.23412	7.20E-06	0.001526	DOWNREGULATED
OPRPN	-4.06763	9.73E-05	0.008023	DOWNREGULATED
MYBPC1	-3.9071	2.55E-08	2.79E-05	DOWNREGULATED
FGB	-3.69867	0.000329	0.015311	DOWNREGULATED
TCN1	-3.31353	3.55E-05	0.0043	DOWNREGULATED
FGG	-3.31224	0.000493	0.019444	DOWNREGULATED
SPOCK3	-3.24199	1.28E-07	9.01E-05	DOWNREGULATED
SUCNR1	-3.21001	6.08E-10	1.61E-06	DOWNREGULATED
SERPINA6	-3.08141	7.12E-05	0.006506	DOWNREGULATED

IGKV6D-41	-2.96754	0.000555	0.020962	DOWNREGULATED
AC093627.7	-2.95129	4.87E-13	3.09E-09	DOWNREGULATED
KLK12	-2.91262	0.002457	0.047953	DOWNREGULATED
KLK13	-2.70373	3.52E-05	0.0043	DOWNREGULATED
ACADL	-2.68309	5.87E-06	0.001329	DOWNREGULATED
CDH19	-2.67667	0.000154	0.010206	DOWNREGULATED
RERGL	-2.64046	1.74E-06	0.000593	DOWNREGULATED
CCDC144NL	-2.61066	0.000763	0.025347	DOWNREGULATED
ENSG00000287558.1	-2.53991	0.001303	0.034718	DOWNREGULATED
TPH1	-2.47888	3.81E-06	0.001005	DOWNREGULATED
FCGR3B	-2.46956	4.61E-07	0.00024	DOWNREGULATED
AC022431.1	-2.44521	0.000214	0.012305	DOWNREGULATED
AL132780.4	-2.43869	4.78E-07	0.000245	DOWNREGULATED
LIPK	-2.4202	0.000451	0.018491	DOWNREGULATED
C8orf34-AS1	-2.41327	0.000181	0.011391	DOWNREGULATED
ANGPTL7	-2.37273	8.49E-05	0.007352	DOWNREGULATED
AC022498.2	-2.33292	4.88E-05	0.005072	DOWNREGULATED
EN2	-2.31381	0.000808	0.026362	DOWNREGULATED
ENSG00000286150.1	-2.29835	4.24E-06	0.001058	DOWNREGULATED
MYOCD	-2.2897	2.47E-09	4.13E-06	DOWNREGULATED
MLC1	-2.28619	7.46E-08	6.12E-05	DOWNREGULATED
SCGB2A1	-2.28484	0.000596	0.021824	DOWNREGULATED
SLC34A2	-2.28374	0.000696	0.024129	DOWNREGULATED
CHRM1	-2.25722	2.57E-05	0.003563	DOWNREGULATED
BMP3	-2.23073	0.000183	0.011438	DOWNREGULATED
CP	-2.18613	0.000118	0.008797	DOWNREGULATED
SFRP5	-2.1763	0.000333	0.015418	DOWNREGULATED
RGS5	-2.17383	1.53E-09	3.24E-06	DOWNREGULATED
SCGB1D2	-2.17187	0.002164	0.04474	DOWNREGULATED
AL450332.1	-2.16538	2.43E-06	0.000778	DOWNREGULATED
PIK3C2G	-2.13687	1.96E-05	0.003051	DOWNREGULATED
AL359633.2	-2.13571	0.001455	0.036295	DOWNREGULATED

KCNG1	-2.13126	3.12E-05	0.004018	DOWNREGULATED
FDCSP	-2.12337	0.001339	0.035138	DOWNREGULATED
DES	-2.12307	1.25E-05	0.002223	DOWNREGULATED
LINC02015	-2.02375	0.000998	0.029742	DOWNREGULATED
AC005537.1	-2.01282	0.002	0.042587	DOWNREGULATED
CLDN10	-2.0033	1.60E-05	0.002677	DOWNREGULATED
UTS2	-1.97175	7.34E-06	0.001526	DOWNREGULATED
SLC8A2	-1.95826	0.000601	0.021939	DOWNREGULATED
ACTC1	-1.94623	0.000487	0.019301	DOWNREGULATED
MKX	3.113082	4.09E-06	0.001045	UPREGULATED
DLK1	3.114402	0.000922	0.028436	UPREGULATED
CYP2B6	3.128554	3.32E-07	0.000192	UPREGULATED
APOA5	3.137322	3.97E-05	0.00463	UPREGULATED
ASS1P11	3.142794	0.000326	0.015233	UPREGULATED
TUBA3E	3.145395	5.59E-09	8.05E-06	UPREGULATED
DDX11L8	3.149422	0.000266	0.013628	UPREGULATED
OR7E161P	3.154175	0.000149	0.01005	UPREGULATED
CYP2A6	3.160278	1.80E-06	0.000609	UPREGULATED
AP005121.1	3.18962	2.04E-06	0.000665	UPREGULATED
AC105399.1	3.198481	0.000113	0.008727	UPREGULATED
CSAG3	3.261432	0.001601	0.038346	UPREGULATED
KDM5D	3.268025	0.000586	0.021565	UPREGULATED
TCL1A	3.319445	8.50E-07	0.000369	UPREGULATED
SPRR1B	3.417275	0.000428	0.017969	UPREGULATED
RPS26P38	3.425951	5.91E-07	0.00028	UPREGULATED
AC015818.3	3.443018	0.000495	0.019469	UPREGULATED
AL445685.2	3.444785	3.61E-05	0.004341	UPREGULATED
IGKV1OR-3	3.55074	0.000371	0.016509	UPREGULATED
TUBA3C	3.552331	3.97E-08	3.78E-05	UPREGULATED
FAM210CP	3.554455	2.96E-08	3.13E-05	UPREGULATED
ENSG00000286358.1	3.643581	4.41E-05	0.004778	UPREGULATED
SLC5A12	3.689127	3.03E-07	0.000185	UPREGULATED

ANKRD20A9P	3.71607	3.19E-06	0.000902	UPREGULATED
ACTBP12	3.749358	8.84E-06	0.001742	UPREGULATED
EDDM3B	3.760549	5.10E-07	0.000249	UPREGULATED
GFY	3.762033	0.000118	0.008797	UPREGULATED
VSTM2A	3.782245	6.65E-05	0.006225	UPREGULATED
APOC3	3.786286	0.000566	0.021146	UPREGULATED
LINC01502	3.914534	2.57E-05	0.003563	UPREGULATED
ENSG00000287020.1	3.984132	0.001022	0.030186	UPREGULATED
COX6CP1	4.075604	2.14E-11	8.49E-08	UPREGULATED
AL592486.1	4.097243	3.44E-07	0.000195	UPREGULATED
SPANXC	4.144299	0.002073	0.043537	UPREGULATED
RPS26P47	4.330647	2.46E-09	4.13E-06	UPREGULATED
OR56A3	4.535317	0.00045	0.01847	UPREGULATED
AC132825.1	4.547798	1.24E-06	0.000488	UPREGULATED
MTCO1P53	4.685403	1.02E-14	1.08E-10	UPREGULATED
COX20P1	4.710614	7.23E-20	2.29E-15	UPREGULATED
TBC1D3D	4.716718	2.11E-07	0.000134	UPREGULATED
MAGEA12	5.080364	5.49E-05	0.005542	UPREGULATED
RPS26P4	5.354735	1.76E-09	3.29E-06	UPREGULATED
MTATP8P2	5.397509	6.37E-15	1.01E-10	UPREGULATED
PSG4	5.60152	1.21E-09	2.75E-06	UPREGULATED
UGT2B17	5.623199	2.07E-10	5.95E-07	UPREGULATED
CYP4F34P	5.944045	2.43E-11	8.57E-08	UPREGULATED
RPS26P34	6.01111	6.42E-11	2.04E-07	UPREGULATED
SCN1A	6.261565	6.51E-13	3.44E-09	UPREGULATED
MAGEA3	6.351383	0.000126	0.009094	UPREGULATED
SPANXB1	6.511293	1.64E-11	7.41E-08	UPREGULATED

Table B8: Top 50 differentially expressed genes between Asian patients and White patients in the luminal A subtype

GeneName	Log2FoldChange	PValue	AdjustedPValue	Expression
CASP14	-3.94755	6.56E-06	0.002986	DOWNREGULATED

IGKV2-29	-3.57157	5.09E-05	0.01299	DOWNREGULATED
AC023157.1	-3.40523	4.96E-09	1.34E-05	DOWNREGULATED
SUCNR1	-2.91381	3.28E-13	1.62E-09	DOWNREGULATED
SPINK1	-2.81198	2.53E-06	0.001362	DOWNREGULATED
EPO	-2.73726	2.32E-06	0.001297	DOWNREGULATED
CU638689.4	-2.60931	0.000473	0.045206	DOWNREGULATED
AC004832.6	-2.5382	8.78E-07	0.000675	DOWNREGULATED
MTCO1P40	-2.47562	5.57E-08	7.85E-05	DOWNREGULATED
CA9	-2.38585	6.85E-05	0.015978	DOWNREGULATED
IGLV10-54	-2.31447	0.000425	0.043819	DOWNREGULATED
UTS2	-2.31202	2.02E-06	0.001223	DOWNREGULATED
PPIAP58	-2.25803	3.13E-08	4.97E-05	DOWNREGULATED
IGKV2OR22-3	-2.15188	0.000539	0.04791	DOWNREGULATED
IGHV3-64	-2.03321	0.000294	0.03658	DOWNREGULATED
IGLV3-19	-1.98905	0.000209	0.030115	DOWNREGULATED
ACP7	-1.91321	2.41E-05	0.007673	DOWNREGULATED
NKX2-8	-1.88824	0.000275	0.035254	DOWNREGULATED
AC093627.7	-1.88409	1.48E-05	0.005223	DOWNREGULATED
C5orf67	-1.85009	0.000418	0.043763	DOWNREGULATED
PPIAP9	-1.84863	2.38E-08	4.15E-05	DOWNREGULATED
AC243829.1	-1.82111	0.000139	0.023485	DOWNREGULATED
NPIPB2	-1.78481	2.64E-05	0.008054	DOWNREGULATED
RETN	-1.76198	0.00023	0.031703	DOWNREGULATED
ADAMTS7P1	-1.72791	1.47E-07	0.000174	DOWNREGULATED
PPIAP21	-1.70848	2.96E-07	0.000324	DOWNREGULATED
IGHG2	-1.7056	0.000283	0.036015	DOWNREGULATED
AL158151.4	-1.64964	1.07E-05	0.004226	DOWNREGULATED
FAM163B	-1.64671	0.000373	0.041622	DOWNREGULATED
CCL3L1	-1.62056	6.89E-06	0.003089	DOWNREGULATED
B3GAT1	-1.59437	0.00046	0.045077	DOWNREGULATED
AC012441.2	-1.56196	0.000307	0.037066	DOWNREGULATED
RAPGEFL1	-1.55338	3.41E-05	0.009339	DOWNREGULATED

ISG15	-1.5351	0.000592	0.04969	DOWNREGULATED
PFN1P1	-1.53382	4.19E-08	6.20E-05	DOWNREGULATED
LINC02043	-1.52719	0.000253	0.033254	DOWNREGULATED
GNG8	-1.52357	8.19E-05	0.017962	DOWNREGULATED
HIST1H4C	-1.51108	9.10E-06	0.003742	DOWNREGULATED
EPPIN	2.93721	0.000435	0.044236	UPREGULATED
PRL	2.94619	0.000113	0.021689	UPREGULATED
AL160408.1	3.008454	0.000116	0.02173	UPREGULATED
AGXT	3.085101	5.68E-05	0.013991	UPREGULATED
DAPL1	3.114852	1.84E-06	0.001134	UPREGULATED
PPP1R1C	3.121622	8.54E-06	0.003662	UPREGULATED
LRP1B	3.145711	6.45E-06	0.002983	UPREGULATED
FOXE1	3.304475	1.06E-06	0.000748	UPREGULATED
UPK1B	3.334541	0.000107	0.021246	UPREGULATED
HOXB-AS3	3.401595	9.16E-07	0.000675	UPREGULATED
KRT1	3.409652	9.73E-09	2.06E-05	UPREGULATED
FUT6	3.419058	2.17E-05	0.007129	UPREGULATED
LINC02315	3.469455	0.000488	0.045301	UPREGULATED
TMPRSS11D	3.479892	7.56E-07	0.000675	UPREGULATED
COX20P1	3.522512	1.22E-08	2.41E-05	UPREGULATED
ENSG00000286329.1	3.631444	0.000284	0.036015	UPREGULATED
ALDH3A1	3.641692	1.08E-10	4.01E-07	UPREGULATED
CDH18	3.659195	0.000235	0.03195	UPREGULATED
ENSG00000286358.1	3.672481	1.31E-05	0.004847	UPREGULATED
GABRA1	3.704817	0.000531	0.047825	UPREGULATED
AC104035.1	3.752735	2.25E-06	0.001281	UPREGULATED
RPS4XP22	3.809311	2.57E-30	7.60E-26	UPREGULATED
AC005829.2	3.824037	9.23E-07	0.000675	UPREGULATED
COL2A1	3.883521	3.19E-08	4.97E-05	UPREGULATED
MTRNR2L1	4.020149	5.87E-09	1.45E-05	UPREGULATED
AP000785.1	4.022054	0.000447	0.044354	UPREGULATED
KRT4	4.044151	8.13E-07	0.000675	UPREGULATED

AC011632.1	4.065868	1.50E-06	0.000984	UPREGULATED
ENSG00000288235.1	4.076617	1.39E-07	0.000171	UPREGULATED
GABRG2	4.080405	0.000484	0.045301	UPREGULATED
AC074389.2	4.091587	0.000287	0.036015	UPREGULATED
CCDC144A	4.131197	2.44E-12	1.03E-08	UPREGULATED
AC092979.1	4.148553	0.000256	0.033456	UPREGULATED
RALYL	4.157981	0.000465	0.045206	UPREGULATED
FAM83C	4.221098	0.000261	0.033718	UPREGULATED
AKR1B10	4.284982	9.35E-09	2.06E-05	UPREGULATED
AC146944.4	4.285572	9.02E-06	0.003742	UPREGULATED
GATA5	4.357112	8.44E-07	0.000675	UPREGULATED
MAPK8IP1P2	4.370521	7.75E-05	0.017372	UPREGULATED
CYP2G2P	4.388746	2.15E-06	0.001246	UPREGULATED
AC098850.3	4.439892	1.43E-09	4.69E-06	UPREGULATED
USP32P1	4.711285	8.19E-15	4.85E-11	UPREGULATED
KRT13	4.843613	8.95E-07	0.000675	UPREGULATED
ENSG00000286619.1	5.134191	3.38E-17	3.33E-13	UPREGULATED
AC078882.1	5.238218	2.70E-07	0.000307	UPREGULATED
SPRR1A	5.470721	0.000206	0.029962	UPREGULATED
AC146944.3	5.685635	3.18E-06	0.001683	UPREGULATED
BRINP3	6.863824	3.73E-09	1.10E-05	UPREGULATED
PSG4	7.239712	1.01E-15	7.47E-12	UPREGULATED
TRH	7.962347	1.57E-20	2.33E-16	UPREGULATED

Table B9: Top 50 differentially expressed genes between Black patients and White patients in the luminal A subtype

GeneName	Log2FoldChange	PValue	AdjustedPValue	diffexpressed
SULT1C3	-5.63101	5.82E-07	1.05E-05	DOWNREGULATED
AC243829.5	-5.57687	1.63E-05	0.000168	DOWNREGULATED
MTATP8P2	-5.54296	2.76E-52	5.35E-48	DOWNREGULATED
MTCO1P40	-5.21261	2.06E-64	8.01E-60	DOWNREGULATED
RPS26P34	-5.01965	7.46E-18	6.43E-15	DOWNREGULATED

TBC1D3D	-4.74608	4.75E-09	1.86E-07	DOWNREGULATED
ENSG00000287448.1	-4.53222	1.00E-08	3.44E-07	DOWNREGULATED
KLK3	-4.45311	1.43E-20	2.21E-17	DOWNREGULATED
LINC01499	-4.30478	2.28E-06	3.25E-05	DOWNREGULATED
MAGEA6	-4.29716	0.000282	0.001827	DOWNREGULATED
RPS26P4	-4.2697	3.27E-13	6.13E-11	DOWNREGULATED
MSMB	-4.15325	6.25E-24	1.73E-20	DOWNREGULATED
MTCO1P53	-4.10048	1.18E-26	5.14E-23	DOWNREGULATED
RHOXF1P2	-4.07496	1.27E-05	0.000136	DOWNREGULATED
AL445685.2	-3.8786	2.21E-13	4.36E-11	DOWNREGULATED
SCN1A	-3.87019	6.97E-14	1.70E-11	DOWNREGULATED
LINC00668	-3.86063	6.70E-10	3.59E-08	DOWNREGULATED
AC113367.2	-3.7831	0.000897	0.004775	DOWNREGULATED
MTND4P24	-3.72276	8.77E-39	1.13E-34	DOWNREGULATED
UCN3	-3.71536	4.71E-05	0.000408	DOWNREGULATED
KLK2	-3.62871	1.08E-23	2.79E-20	DOWNREGULATED
IGLVI-70	-3.61976	1.34E-13	2.89E-11	DOWNREGULATED
RHOXF1P1	-3.61712	2.95E-08	8.65E-07	DOWNREGULATED
MTND1P23	-3.57248	1.05E-10	7.61E-09	DOWNREGULATED
AC104454.2	-3.53553	3.09E-05	0.000286	DOWNREGULATED
DPP10-AS1	-3.47049	7.05E-07	1.23E-05	DOWNREGULATED
CYP4F34P	-3.41952	3.63E-06	4.77E-05	DOWNREGULATED
TCAP	-3.26818	1.35E-34	1.31E-30	DOWNREGULATED
RPL10P6	-3.22087	6.96E-17	4.74E-14	DOWNREGULATED
FEZF1-AS1	-3.20074	6.26E-12	7.08E-10	DOWNREGULATED
ACTBP12	-3.19262	1.46E-10	1.01E-08	DOWNREGULATED
MAGEA3	-3.12037	0.001687	0.008053	DOWNREGULATED
SPANXB1	-3.1172	6.53E-06	7.83E-05	DOWNREGULATED
ZG16	-3.10392	2.08E-09	9.31E-08	DOWNREGULATED
APOC3	-3.1012	5.44E-08	1.45E-06	DOWNREGULATED
MYPN	-3.08143	2.49E-11	2.25E-09	DOWNREGULATED
AC084149.1	-3.06512	4.94E-06	6.20E-05	DOWNREGULATED

RPL10P9	-3.05604	3.28E-22	6.71E-19	DOWNREGULATED
ZBTB8OSP2	-3.05485	1.20E-16	7.14E-14	DOWNREGULATED
OR7E22P	-3.03993	1.31E-16	7.72E-14	DOWNREGULATED
WFDC10A	-3.03434	3.15E-05	0.00029	DOWNREGULATED
SORD2P	-3.02739	1.21E-32	9.37E-29	DOWNREGULATED
TBC1D3E	-3.02612	4.58E-10	2.59E-08	DOWNREGULATED
IGLL1	-2.97816	8.29E-12	8.72E-10	DOWNREGULATED
AC243829.1	-2.96989	1.91E-21	3.22E-18	DOWNREGULATED
AC137695.2	-2.96098	1.38E-12	2.02E-10	DOWNREGULATED
AC103769.1	-2.92415	8.89E-11	6.55E-09	DOWNREGULATED
AC004080.1	-2.88118	1.95E-06	2.84E-05	DOWNREGULATED
HSD3BP2	-2.85163	1.40E-05	0.000149	DOWNREGULATED
ENSG00000287443.1	-2.81038	0.001026	0.00534	DOWNREGULATED
CR392039.3	2.37005	0.001985	0.009211	UPREGULATED
CXADRP1	2.37058	5.71E-07	1.03E-05	UPREGULATED
AADACL2	2.379953	0.000225	0.001521	UPREGULATED
AC146944.3	2.390483	0.000567	0.003262	UPREGULATED
ZNF716	2.408957	0.003051	0.013132	UPREGULATED
GRAMD4P1	2.409569	7.31E-05	0.000593	UPREGULATED
MMP8	2.423261	2.53E-07	5.20E-06	UPREGULATED
MED15P5	2.434486	0.002937	0.01273	UPREGULATED
AC111149.2	2.435091	0.010778	0.036563	UPREGULATED
IGHV5-10-1	2.440619	1.71E-07	3.78E-06	UPREGULATED
GTF3AP6	2.443435	0.009837	0.033999	UPREGULATED
OPRPN	2.447774	4.44E-05	0.000388	UPREGULATED
ENSG00000285972.1	2.453974	0.00049	0.002889	UPREGULATED
FOXE1	2.478715	1.85E-07	4.01E-06	UPREGULATED
ENSG00000287069.1	2.569161	1.44E-06	2.22E-05	UPREGULATED
ENSG00000287609.1	2.588247	1.61E-07	3.59E-06	UPREGULATED
SYT10	2.589888	3.38E-05	0.000309	UPREGULATED
GJB6	2.608705	2.86E-08	8.44E-07	UPREGULATED
NTS	2.627926	1.87E-05	0.000189	UPREGULATED

AC087379.1	2.638814	0.00024	0.001602	UPREGULATED
AF127577.3	2.647837	0.000162	0.001161	UPREGULATED
AL033397.1	2.652564	0.013455	0.043612	UPREGULATED
LINC00491	2.662734	0.000337	0.002122	UPREGULATED
FGB	2.715594	1.55E-05	0.000161	UPREGULATED
AC023421.1	2.751324	1.23E-13	2.69E-11	UPREGULATED
CPB1	2.799442	5.49E-06	6.76E-05	UPREGULATED
SERPINB3	2.828871	0.007644	0.027712	UPREGULATED
KRT13	2.832461	1.56E-06	2.38E-05	UPREGULATED
CST5	2.883383	3.20E-07	6.35E-06	UPREGULATED
CTAG2	3.007032	0.004239	0.017196	UPREGULATED
RN7SKP256	3.0167	0.004567	0.018244	UPREGULATED
UGT2B4	3.119576	1.79E-08	5.63E-07	UPREGULATED
NANOGP3	3.139809	0.000173	0.001223	UPREGULATED
SPRR3	3.143499	0.000405	0.002469	UPREGULATED
SPRR1A	3.196608	0.000216	0.001471	UPREGULATED
SERPINB13	3.201775	0.000229	0.001542	UPREGULATED
SMR3B	3.268627	1.50E-06	2.30E-05	UPREGULATED
AL354984.2	3.375383	2.65E-07	5.41E-06	UPREGULATED
CYP2A7	3.398158	1.34E-12	1.99E-10	UPREGULATED
BPIFA1	3.464862	6.47E-05	0.000534	UPREGULATED
AC146944.4	3.640597	1.53E-09	7.11E-08	UPREGULATED
AKR1B10	3.797371	1.33E-15	5.93E-13	UPREGULATED
AL390061.1	3.827278	2.09E-05	0.000207	UPREGULATED
AC074389.2	3.855506	1.16E-07	2.72E-06	UPREGULATED
AC114786.2	3.856909	0.009041	0.031754	UPREGULATED
SERPINA6	3.897097	1.10E-11	1.10E-09	UPREGULATED
MKNK2P1	3.938738	1.49E-06	2.29E-05	UPREGULATED
LINC02302	4.248807	6.51E-09	2.40E-07	UPREGULATED
DCD	4.734812	3.54E-11	3.00E-09	UPREGULATED
ADH7	6.432831	4.31E-06	5.53E-05	UPREGULATED

Table B10: Top 50 differentially expressed genes between Asian patients and Black patients in the luminal B subtype

GeneName	Log2FoldChange	PValue	AdjustedPValue	Expression
ENSG00000287448.1	-6.49939	1.03E-06	0.001263	DOWNREGULATED
MAGEA3	-6.46239	6.48E-05	0.01583	DOWNREGULATED
TSPAN8	-5.75197	7.79E-09	2.19E-05	DOWNREGULATED
HMX1	-5.3764	7.42E-08	0.000123	DOWNREGULATED
LINC01639	-5.28927	2.16E-05	0.008186	DOWNREGULATED
DPYSL5	-5.25106	1.56E-10	7.31E-07	DOWNREGULATED
GABRQ	-4.65537	1.35E-08	3.16E-05	DOWNREGULATED
CPLX2	-4.35596	5.60E-05	0.014428	DOWNREGULATED
ZIC4	-4.16842	1.62E-05	0.006862	DOWNREGULATED
GABRA3	-4.091	4.53E-05	0.012833	DOWNREGULATED
KIF1A	-3.94951	1.05E-05	0.006176	DOWNREGULATED
AL772337.3	-3.89938	0.000242	0.032533	DOWNREGULATED
CHRNA4	-3.79933	2.03E-05	0.00781	DOWNREGULATED
C6orf223	-3.66412	5.77E-06	0.004155	DOWNREGULATED
CR382285.1	-3.59296	5.44E-05	0.014155	DOWNREGULATED
AC005696.4	-3.48905	1.48E-05	0.006862	DOWNREGULATED
MED15P9	-3.3273	0.000282	0.036148	DOWNREGULATED
LINC00461	-3.323	0.000235	0.032263	DOWNREGULATED
MC5R	-3.28792	1.58E-05	0.006862	DOWNREGULATED
LINC00578	-3.2342	1.78E-07	0.000277	DOWNREGULATED
RIMS2	-3.19698	5.07E-05	0.013579	DOWNREGULATED
CACNA1B	-3.12077	3.13E-07	0.000462	DOWNREGULATED
HNRNPA1P57	-3.11606	0.000122	0.022626	DOWNREGULATED
CPEB1-AS1	-3.02469	1.05E-05	0.006176	DOWNREGULATED
KIR2DS4	-2.95931	8.57E-05	0.018709	DOWNREGULATED
PON1	-2.93583	0.000327	0.03892	DOWNREGULATED
MOGAT2	-2.93502	1.62E-05	0.006862	DOWNREGULATED
SRRM4	-2.91233	0.00049	0.049165	DOWNREGULATED
C2CD4A	-2.88248	0.000136	0.023428	DOWNREGULATED

C15orf59-AS1	-2.85354	0.000324	0.03892	DOWNREGULATED
HIST1H3G	-2.8235	0.000333	0.039095	DOWNREGULATED
SEZ6	-2.8096	0.000125	0.022626	DOWNREGULATED
KCNK3	-2.77143	0.000443	0.046087	DOWNREGULATED
KIR2DL1	-2.77118	0.000484	0.048951	DOWNREGULATED
ATP1A2	-2.77103	1.82E-05	0.007297	DOWNREGULATED
MTND6P3	-2.73173	0.000507	0.049851	DOWNREGULATED
CNTN2	-2.7222	1.54E-05	0.006862	DOWNREGULATED
AC022498.2	-2.69556	5.77E-05	0.014608	DOWNREGULATED
UNC80	-2.66129	0.000192	0.028108	DOWNREGULATED
TMEM178B	-2.63172	1.86E-05	0.007297	DOWNREGULATED
KLK13	-2.61605	0.000303	0.037483	DOWNREGULATED
FER1L6	-2.61113	5.17E-05	0.013707	DOWNREGULATED
ENSG00000286590.1	-2.59674	0.000379	0.041513	DOWNREGULATED
VIPR2	-2.59262	7.68E-05	0.017701	DOWNREGULATED
AL132780.4	-2.58991	2.21E-06	0.002024	DOWNREGULATED
AL356234.2	-2.52201	0.000383	0.041715	DOWNREGULATED
AC124657.1	-2.47949	0.000139	0.02374	DOWNREGULATED
FAM19A4	-2.47695	0.000307	0.037821	DOWNREGULATED
HCN4	-2.45082	4.57E-05	0.012833	DOWNREGULATED
ADGRF4	-2.44225	2.53E-05	0.009126	DOWNREGULATED
C1orf167	2.404374	0.000498	0.04946	UPREGULATED
TPSP2	2.408294	6.67E-05	0.016169	UPREGULATED
CHIT1	2.427302	0.000129	0.023	UPREGULATED
ENSG00000285644.1	2.430518	4.52E-05	0.012833	UPREGULATED
FAM103A2P	2.457282	1.50E-06	0.001566	UPREGULATED
C2CD4C	2.473776	4.31E-05	0.012605	UPREGULATED
RGS6	2.479399	0.000233	0.032149	UPREGULATED
EPHA6	2.491572	0.000328	0.038947	UPREGULATED
KRT8P42	2.501635	0.00037	0.041159	UPREGULATED
DSCAML1	2.505689	0.0002	0.028886	UPREGULATED
AL445189.2	2.514724	0.000481	0.048951	UPREGULATED

AC005225.2	2.52054	4.87E-05	0.013317	UPREGULATED
AC051619.5	2.528452	1.33E-05	0.006862	UPREGULATED
ZBTB8OSP2	2.595168	0.000267	0.034813	UPREGULATED
KHDRBS2	2.603127	0.000121	0.022626	UPREGULATED
AC010624.1	2.616995	9.11E-06	0.005816	UPREGULATED
CA4	2.633872	0.000484	0.048951	UPREGULATED
SLC25A48	2.644126	3.72E-05	0.011732	UPREGULATED
PYY	2.664262	0.000143	0.023921	UPREGULATED
ENSG00000287826.1	2.788491	0.000179	0.027102	UPREGULATED
RPS26P38	2.789309	0.000182	0.02742	UPREGULATED
AC010624.4	2.8019	0.00039	0.042271	UPREGULATED
AC073264.3	2.857019	0.000102	0.020605	UPREGULATED
PNLIPRP2	2.916654	0.000145	0.023921	UPREGULATED
AP001056.1	2.91757	0.000218	0.030653	UPREGULATED
FAM196A	3.003742	0.000144	0.023921	UPREGULATED
LINC01956	3.047845	3.92E-05	0.011852	UPREGULATED
RPS26P4	3.147094	0.000503	0.049642	UPREGULATED
COX20P1	3.185067	4.57E-06	0.003472	UPREGULATED
LEP	3.22725	4.05E-05	0.012116	UPREGULATED
LCNL1	3.322574	0.000159	0.025343	UPREGULATED
MTND1P23	3.416393	0.000291	0.036661	UPREGULATED
SYNPO2L	3.490679	0.00024	0.032373	UPREGULATED
WT1-AS	3.545174	0.000127	0.022913	UPREGULATED
SEZ6L	3.547852	0.000178	0.027102	UPREGULATED
IGLV1-41	3.71404	3.80E-05	0.011852	UPREGULATED
U2AF1L5	3.803348	1.17E-15	6.60E-12	UPREGULATED
RPL10P9	3.860268	9.38E-17	6.59E-13	UPREGULATED
CBLN2	3.917282	2.74E-05	0.009519	UPREGULATED
ENSG00000285972.1	4.011675	0.000227	0.031566	UPREGULATED
RPS26P34	4.096269	8.46E-06	0.005529	UPREGULATED
FAM210CP	4.186403	4.35E-07	0.000612	UPREGULATED
GCGR	4.376815	8.13E-07	0.001088	UPREGULATED

AP000763.2	4.813088	7.43E-08	0.000123	UPREGULATED
RPL10P6	4.828384	1.12E-19	3.14E-15	UPREGULATED
HS3ST5	5.070117	3.13E-10	1.26E-06	UPREGULATED
UGT2B17	5.432456	9.79E-06	0.005981	UPREGULATED
TBC1D3D	6.020275	3.14E-08	6.31E-05	UPREGULATED
MTCO1P53	6.257422	9.19E-18	1.29E-13	UPREGULATED
IGHV1-69-2	7.847883	7.28E-17	6.59E-13	UPREGULATED

Table B11: Top 50 differentially expressed genes between Asian patients and White patients in the luminal B subtype

GeneName	Log2FoldChange	PValue	AdjustedPValue	Expression
RPL37P4	-6.14786	5.79E-05	0.011206	DOWNREGULATED
CHGA	-6.10217	2.47E-17	4.01E-13	DOWNREGULATED
TRH	-5.19482	2.76E-09	4.48E-06	DOWNREGULATED
SEC14L3	-5.13103	3.28E-11	1.18E-07	DOWNREGULATED
LHFPL4	-4.85874	4.49E-11	1.33E-07	DOWNREGULATED
AL353726.2	-4.7406	3.54E-06	0.001536	DOWNREGULATED
TSPAN8	-4.64947	5.68E-08	5.77E-05	DOWNREGULATED
CNTN6	-4.63755	1.89E-14	1.41E-10	DOWNREGULATED
IGHV7-4-1	-4.5123	5.79E-07	0.000355	DOWNREGULATED
LINC02468	-4.4936	1.81E-06	0.000937	DOWNREGULATED
CR382285.1	-4.42646	2.16E-14	1.41E-10	DOWNREGULATED
AC073869.5	-4.4102	4.02E-05	0.009013	DOWNREGULATED
MED15P9	-4.3786	7.44E-09	1.05E-05	DOWNREGULATED
C6orf223	-4.24284	1.37E-13	6.37E-10	DOWNREGULATED
AC011195.2	-4.21941	3.86E-11	1.26E-07	DOWNREGULATED
IGKV2-29	-4.18291	9.58E-07	0.000546	DOWNREGULATED
CEACAMP3	-3.98813	1.94E-05	0.005637	DOWNREGULATED
NR5A1	-3.70135	9.57E-07	0.000546	DOWNREGULATED
GLRA3	-3.65134	2.12E-05	0.006005	DOWNREGULATED
IGKV1D-17	-3.56231	2.78E-06	0.001275	DOWNREGULATED
IGKV2D-28	-3.53159	1.63E-08	2.12E-05	DOWNREGULATED

IGLVI-70	-3.5146	4.06E-07	0.000254	DOWNREGULATED
EPS8L3	-3.50805	4.69E-10	1.09E-06	DOWNREGULATED
AL139039.2	-3.50006	2.41E-05	0.006593	DOWNREGULATED
AL772337.1	-3.48528	5.15E-06	0.00212	DOWNREGULATED
MTCO1P42	-3.41111	6.60E-05	0.012123	DOWNREGULATED
CST8	-3.39259	5.75E-05	0.011197	DOWNREGULATED
AC025423.3	-3.37785	0.000485	0.046983	DOWNREGULATED
BCRP1	-3.34846	8.43E-08	7.83E-05	DOWNREGULATED
AC133681.1	-3.3265	0.000234	0.029174	DOWNREGULATED
CD5L	-3.27751	0.000209	0.027441	DOWNREGULATED
AC026369.1	-3.27586	1.79E-09	3.24E-06	DOWNREGULATED
SPACA3	-3.27017	1.88E-07	0.000145	DOWNREGULATED
BCAR4	-3.21576	9.89E-05	0.016071	DOWNREGULATED
MIR4740	-3.19077	5.80E-09	8.57E-06	DOWNREGULATED
AC114316.2	-3.1865	9.72E-05	0.015942	DOWNREGULATED
AQP5	-3.17174	3.98E-05	0.008993	DOWNREGULATED
UTS2	-3.14168	6.89E-15	7.46E-11	DOWNREGULATED
MYEOV	-3.07938	2.21E-08	2.66E-05	DOWNREGULATED
CU638689.4	-3.06771	0.000113	0.017176	DOWNREGULATED
LINC01748	-3.03454	4.40E-05	0.009539	DOWNREGULATED
IGKV2-28	-2.95323	3.26E-05	0.007656	DOWNREGULATED
AL109897.1	-2.93457	1.93E-05	0.005637	DOWNREGULATED
ENSG00000286062.1	-2.91117	9.80E-08	8.85E-05	DOWNREGULATED
AL139039.3	-2.86609	2.37E-06	0.001149	DOWNREGULATED
TCF24	-2.84344	4.69E-05	0.010057	DOWNREGULATED
SYNPO2L	-2.80696	0.000232	0.029098	DOWNREGULATED
LRP1B	-2.78474	0.000226	0.028851	DOWNREGULATED
AC096734.1	-2.71007	0.000124	0.018529	DOWNREGULATED
KIR2DS4	-2.62725	5.65E-06	0.002298	DOWNREGULATED
WNT3A	2.35603	0.000463	0.045793	UPREGULATED
DLX2	2.371902	0.000448	0.044909	UPREGULATED
PI16	2.412202	5.62E-05	0.011097	UPREGULATED

CYP2A7	2.428635	0.000381	0.039824	UPREGULATED
DND1P1	2.456578	3.21E-05	0.007655	UPREGULATED
CD36	2.471776	6.78E-08	6.48E-05	UPREGULATED
OR3A2	2.500442	8.24E-06	0.003116	UPREGULATED
SERPINB2	2.561375	0.000118	0.017853	UPREGULATED
LINC01956	2.633021	0.000107	0.016782	UPREGULATED
SLC6A4	2.650373	6.00E-05	0.011476	UPREGULATED
GTF2H2B	2.656313	2.99E-07	0.000203	UPREGULATED
S100A9	2.665329	2.78E-06	0.001275	UPREGULATED
AFP	2.721091	0.000232	0.029098	UPREGULATED
SORCS1	2.753877	0.000152	0.021713	UPREGULATED
HMGCS2	2.754063	6.95E-05	0.012487	UPREGULATED
PURPL	2.769917	0.000243	0.029973	UPREGULATED
GNAQP1	2.813038	2.25E-07	0.000161	UPREGULATED
CA4	2.850186	3.06E-05	0.007496	UPREGULATED
MTRNR2L1	2.87377	7.57E-06	0.002896	UPREGULATED
ENSG00000285668.1	2.880253	2.54E-05	0.006779	UPREGULATED
HS3ST5	2.927493	8.14E-05	0.013709	UPREGULATED
AC005829.2	3.010979	2.33E-05	0.006464	UPREGULATED
OR3A1	3.0729	5.75E-05	0.011197	UPREGULATED
SSTR5-AS1	3.096045	0.000311	0.034878	UPREGULATED
CCDC175	3.163883	0.000362	0.038522	UPREGULATED
AP001999.1	3.171337	0.000109	0.016904	UPREGULATED
U2AF1L5	3.18422	1.15E-10	3.11E-07	UPREGULATED
DDX11L10	3.258728	6.93E-07	0.00041	UPREGULATED
MKX	3.26284	2.27E-07	0.000161	UPREGULATED
SRP68P1	3.327187	0.00018	0.025008	UPREGULATED
CHRNB2	3.348958	3.39E-07	0.00022	UPREGULATED
USP32P1	3.515296	2.20E-09	3.76E-06	UPREGULATED
IGHV1-69-2	3.515805	0.00011	0.017021	UPREGULATED
AC073850.1	3.653188	1.49E-05	0.004697	UPREGULATED
AC105999.2	3.657357	0.0005	0.047527	UPREGULATED

CCDC144A	3.685476	4.37E-10	1.09E-06	UPREGULATED
CHGB	3.748572	2.67E-06	0.001258	UPREGULATED
CXADRP1	3.982732	0.000244	0.029987	UPREGULATED
S100A7	4.115606	2.48E-05	0.006728	UPREGULATED
ENSG00000286619.1	4.119332	7.92E-13	3.22E-09	UPREGULATED
RPS4XP22	4.160974	2.54E-26	8.27E-22	UPREGULATED
DCD	4.398783	0.000248	0.030304	UPREGULATED
AC002383.1	4.419508	8.63E-06	0.003224	UPREGULATED
MAPK8IP1P2	4.447549	5.09E-05	0.010436	UPREGULATED
ENSG00000288235.1	4.53927	1.02E-07	8.92E-05	UPREGULATED
AC098850.3	4.543853	1.73E-09	3.24E-06	UPREGULATED
AL050303.4	4.614358	1.60E-05	0.004866	UPREGULATED
PSG4	4.853786	5.60E-08	5.77E-05	UPREGULATED
S100A7A	5.089309	8.89E-06	0.003248	UPREGULATED
CARTPT	5.301501	2.08E-05	0.005921	UPREGULATED

Table B12: Top 50 differentially expressed genes between Black patients and White patients in the luminal B subtype

GeneName	Log2FoldChange	PValue	AdjustedPValue	Expression
SCN1A	-5.76814	5.68E-18	2.99E-14	DOWNREGULATED
AP000763.2	-5.62665	8.38E-42	2.65E-37	DOWNREGULATED
MTND1P23	-5.49279	3.22E-19	2.04E-15	DOWNREGULATED
MTCO1P53	-5.29341	1.85E-22	2.92E-18	DOWNREGULATED
RPS26P34	-4.71968	2.04E-17	8.07E-14	DOWNREGULATED
IGLVI-70	-4.57303	1.08E-13	2.28E-10	DOWNREGULATED
IGKV2-29	-4.52731	9.71E-10	6.40E-07	DOWNREGULATED
AC243829.5	-4.41134	0.000707	0.015753	DOWNREGULATED
ENSG00000285822.1	-4.31953	0.000868	0.017875	DOWNREGULATED
ENSG00000286308.1	-4.12	2.33E-05	0.001676	DOWNREGULATED
MTCO1P40	-4.02989	2.13E-14	5.18E-11	DOWNREGULATED
FCAMR	-3.85024	1.14E-08	5.77E-06	DOWNREGULATED

IGKV2D-29	-3.80246	6.04E-13	8.69E-10	DOWNREGULATED
TBC1D3D	-3.70786	4.55E-05	0.002591	DOWNREGULATED
VIL1	-3.69638	1.22E-14	3.22E-11	DOWNREGULATED
IGHV1-17	-3.66256	5.95E-08	1.98E-05	DOWNREGULATED
LINC02046	-3.64382	8.45E-05	0.003954	DOWNREGULATED
AL355478.1	-3.61601	0.000196	0.006973	DOWNREGULATED
IGLV3-29	-3.59319	1.98E-06	0.000298	DOWNREGULATED
IGLV3-30	-3.52039	0.00163	0.026019	DOWNREGULATED
IGLV3-16	-3.36869	7.84E-08	2.53E-05	DOWNREGULATED
HPR	-3.33936	4.55E-10	3.20E-07	DOWNREGULATED
IGHV1-24	-3.32129	2.27E-09	1.36E-06	DOWNREGULATED
PI15	-3.31394	7.46E-12	9.07E-09	DOWNREGULATED
RNF17	-3.29548	6.68E-08	2.20E-05	DOWNREGULATED
IGKV2-24	-3.29446	3.19E-10	2.40E-07	DOWNREGULATED
RPL10P6	-3.27699	4.19E-11	4.42E-08	DOWNREGULATED
MTATP8P2	-3.27324	8.99E-18	4.06E-14	DOWNREGULATED
CYP4F34P	-3.27088	0.000878	0.018021	DOWNREGULATED
AC093809.1	-3.24584	3.16E-17	1.11E-13	DOWNREGULATED
AL512631.1	-3.24582	2.86E-08	1.14E-05	DOWNREGULATED
IGHV1-69-2	-3.24148	0.00013	0.005253	DOWNREGULATED
TBC1D3E	-3.219	5.05E-09	2.75E-06	DOWNREGULATED
LINC02526	-3.2159	2.24E-05	0.00163	DOWNREGULATED
IGHV7-27	-3.10281	3.94E-05	0.002342	DOWNREGULATED
DLK1	-3.08815	2.17E-05	0.001614	DOWNREGULATED
IGLV3-21	-3.07416	3.00E-08	1.17E-05	DOWNREGULATED
IGKV2D-30	-3.04906	1.32E-06	0.000226	DOWNREGULATED
AL035252.2	-3.04073	4.78E-06	0.000575	DOWNREGULATED
IGHV3-79	-3.03529	1.42E-05	0.001205	DOWNREGULATED
AL445189.2	-3.02563	9.65E-13	1.33E-09	DOWNREGULATED
IGKV1D-17	-3.0254	4.81E-07	0.000114	DOWNREGULATED
IGKV1D-27	-3.0197	5.42E-07	0.000121	DOWNREGULATED
AC010127.1	-2.95224	1.40E-05	0.001196	DOWNREGULATED

LGALS4	-2.92736	2.91E-19	2.04E-15	DOWNREGULATED
RPL10P9	-2.89576	1.21E-13	2.40E-10	DOWNREGULATED
ZBTB8OSP2	-2.88819	1.23E-10	1.10E-07	DOWNREGULATED
LINC01748	-2.88459	4.06E-06	0.00051	DOWNREGULATED
MTND4P24	-2.87117	5.64E-12	7.14E-09	DOWNREGULATED
IGHV2-70D	-2.84904	1.04E-06	0.000195	DOWNREGULATED
FOXI1	2.731101	9.82E-06	0.000971	UPREGULATED
CYP2A7P2	2.735554	1.68E-06	0.000269	UPREGULATED
NATP	2.753687	0.000877	0.018009	UPREGULATED
AC023421.1	2.754552	4.03E-08	1.49E-05	UPREGULATED
RF00012	2.757552	0.000209	0.007302	UPREGULATED
CYP2A7	2.771496	1.80E-06	0.000283	UPREGULATED
SRGAP3-AS2	2.796348	0.000112	0.004779	UPREGULATED
AC096582.3	2.812038	0.001107	0.02061	UPREGULATED
DPYSL5	2.861691	0.00039	0.010746	UPREGULATED
INSM1	2.889872	1.97E-05	0.001501	UPREGULATED
NRXN1	2.909444	3.51E-07	8.76E-05	UPREGULATED
OTOP3	2.973049	0.001417	0.023807	UPREGULATED
KRT75	2.981151	8.12E-06	0.000858	UPREGULATED
LIPK	2.983473	9.67E-05	0.004386	UPREGULATED
KLK10	2.997548	4.57E-08	1.62E-05	UPREGULATED
ANKRD20A17P	3.007061	0.004152	0.046876	UPREGULATED
ENSG00000287448.1	3.018823	0.000294	0.009064	UPREGULATED
NGB	3.024376	5.76E-06	0.000652	UPREGULATED
ADGRG7	3.028707	0.000428	0.01137	UPREGULATED
ACTG1P22	3.119353	3.40E-05	0.002105	UPREGULATED
MKNK2P1	3.130226	0.004507	0.049287	UPREGULATED
ZIC1	3.223418	1.91E-07	5.30E-05	UPREGULATED
AL445183.3	3.315246	4.94E-07	0.000116	UPREGULATED
CHRNB2	3.318022	7.63E-10	5.13E-07	UPREGULATED
AL050303.4	3.351864	0.000126	0.005156	UPREGULATED
AC073325.1	3.360437	1.77E-05	0.001399	UPREGULATED

AC079310.1	3.375453	0.0018	0.027797	UPREGULATED
AL160408.1	3.409044	3.29E-09	1.86E-06	UPREGULATED
ASNSP1	3.422197	0.00133	0.022947	UPREGULATED
CPLX2	3.430038	7.79E-05	0.003755	UPREGULATED
ENSG00000288235.1	3.501061	4.98E-07	0.000116	UPREGULATED
AC087379.2	3.514074	0.000251	0.008244	UPREGULATED
AMER3	3.514977	4.09E-05	0.002405	UPREGULATED
GLYATL2	3.532866	1.03E-11	1.21E-08	UPREGULATED
AC005829.2	3.552357	8.86E-09	4.75E-06	UPREGULATED
AL354984.2	3.755542	7.03E-05	0.003507	UPREGULATED
AC026355.1	3.861197	2.50E-09	1.47E-06	UPREGULATED
HNRNPA1P57	4.040939	1.26E-10	1.10E-07	UPREGULATED
BAGE2	4.049493	0.000249	0.008199	UPREGULATED
KLK13	4.195407	6.72E-11	6.64E-08	UPREGULATED
LINC01446	4.224715	1.10E-05	0.001037	UPREGULATED
CXADRP1	4.259256	1.30E-06	0.000223	UPREGULATED
GRAMD4P1	4.328686	6.48E-05	0.003328	UPREGULATED
MAPK8IP1P2	4.375505	1.19E-06	0.000213	UPREGULATED
AC004949.1	4.585162	0.00089	0.01814	UPREGULATED
ENSG00000285711.1	4.600739	0.000718	0.015844	UPREGULATED
NELL1	5.320693	2.24E-09	1.36E-06	UPREGULATED
OLFM3	5.739524	7.17E-07	0.000145	UPREGULATED
SYT4	5.97158	2.04E-06	0.000302	UPREGULATED
LINC02077	6.544787	0.000582	0.013851	UPREGULATED