



Unique Apoptotic Gene Dysregulation in Kenyan Breast Cancer Patients Compared to African American and Caucasian Women

Marion N. Nyamari^{*1}, John Gitau², Kennedy Wanjau Mwangi³, Carol W. Hunja⁴, Sammy Wambua^{5,6}, and Francis Makokha²

¹Department of Biochemistry and Biotechnology, Pwani University, Kilifi, Kenya;

²Directorate of Research & Innovation, Mount Kenya University, Thika, Kenya;

³International Livestock Research Institute, Nairobi, Kenya; ⁴Department of Life Sciences, South-Eastern Kenya University, Kitui, Kenya; ⁵Department of Biological Sciences, Pwani University, Kilifi, Kenya, and ⁶Research and Conservation Support Sciences, Kilifi, Kenya

***Corresponding author:** Marion N. Nyamari. Email address: marionnyaboke01@gmail.com

ORCID: 0000-0003-0004-9982

DOI: <https://dx.doi.org/10.4314/ajhs.v38i4.11>

This work is distributed Open Access under the Creative Commons Attribution 4.0 (CC BY 4.0).
Copyright resides with the authors

Abstract

Background: Kenya records about 6,000 new diagnoses each year of breast cancer (BC), a major cause of cancer-related deaths in the country. The apoptotic pathway is well studied and proven to play a crucial role in cancer onset and progression, but studies have not been extensively done on Kenyan women with BC. The study's objective was to compare the gene expression patterns of apoptotic genes in Kenyan women with breast cancer to those of Caucasian women and African American women based on demography, subtype, and tumour stage.

Methods: Transcriptomic data from 46 African Americans, 66 Caucasians retrieved from The Cancer Genome Atlas (TCGA), and 60 Kenyan RNA-seq data (45 samples from AIC Kijabe/Agaa Khan hospitals and 15 Triple-Negative Breast Cancer (TNBC) from Gene Expression Atlas (GEO)) were studied. The Kenyan patients were part of a longitudinal BC genomics study in Kenya, and the TCGA data were retrospectively selected based on availability of RNA-seq data, documented ethnicity, and tumour stage and subtype classification. HISAT2 v2.2.0 aligned Kenyan transcripts to the reference genome, and HTSeq-count v2.0.2 determined gene counts. EdgeR v4.2 was used to analyse differential gene expression, and Enrichr v3.1 and ClusterProfiler v4.0 were used for enrichment analysis.

Results: The analysis revealed 23 population-specific apoptotic genes differentially expressed across the 60 Kenyan samples, with 9 upregulated and 14 downregulated. In comparison, 80 upregulated and 25 downregulated unique apoptotic genes were identified in the Caucasian dataset, while 6 upregulated and 1 downregulated unique apoptotic genes were recorded from the African-American dataset. Kenyan BC patients (without controlling for subtype or stage) had a unique elevation of the anti-apoptotic gene, Insulin-like Growth Factor 1 Receptor (IGF1R). Kenyan women with Luminal A subtype, compared to other subtypes studied and compared to Caucasian and African-American women, had a distinct upregulation of Interleukin-1 beta (IL1B). The IL1B gene prevents the progression and spread of tumour cells in prostate cancer, with over-exposure of tumour cells to IL1B in vitro leading to resistance to a variety of anticancer medicines in mice. On the other hand, IGF1R has been associated with poor prognosis in patients with elevated expression of the gene in Stage I, in non-small-cell lung cancer.

Conclusion: These results highlight population-specific molecular differences. IGF1R and IL1B are among genes with potential as Kenyan-specific diagnostic targets.

Keywords: Apoptosis, Kenya, Breast cancer

[*Afr. J. Health Sci.* 2025;38(4): Article 11. <https://doi.org/10.4314/ajhs.v38i4.11>]

Background

Breast cancer (BC) ranks as the cancer that affects most women in Kenya, with 6,700 new cases and 3,100 BC-related deaths recorded in 2024 (1). Kenyan patients often exhibit advanced stages and are much younger

at diagnosis, with 54% of cases diagnosed under the age of 50 (2,3). Between the years 2020 - 2023, 5.1% of all BC diagnoses were Stage I, 23.8% Stage II, 31.4% Stage III, and 37.2% Stage IV (4). Nationwide efforts are being made to enhance awareness, screening,

and treatment access with the aim of reducing this burden (5).

BC is classified into Luminal A, B, HER2+, and Triple-negative BC (TNBC) subtypes based on hormone receptor presence and HER2 gene expression (6–10). Apoptosis ensures cellular balance by eliminating precancerous and virus-infected cells, but cancer cells have evolved various ways to bypass it (11,12). Apoptosis-related genes include caspases, Tumour Necrosis Factor (TNF), p53 gene, and B-cell lymphoma 2 (Bcl-2) (13).

Downregulation of caspase-9 and caspase-3 and changes in p53 have been linked with breast cancer (14,15). Other anti-apoptotic mechanisms in BC include over-expression of growth factor receptors and downregulation of death receptors (15,16).

Targeting apoptosis may be a cancer-fighting strategy as TRAIL and the STAT3 (17) signalling pathway have previously been studied as a therapeutic target (17–19). Chemotherapy activates inhibitors of apoptosis proteins (IAP) expression, such as BCL-2 proteins, which prevent apoptosis through blocking cytochrome-c release from the mitochondria and increasing the resistance of a tumour to therapy (20).

We present a comparative analysis of apoptotic gene expression in Kenyan, African American, and Caucasian women with breast cancer. We examined differentially expressed apoptotic genes shared and unique to each population. We also analysed apoptotic gene expression patterns across breast cancer subtypes and tumour stages. This comparative approach aims to identify population-specific molecular signatures that may inform targeted diagnostic and therapeutic strategies for Kenyan BC patients.

Methodology

Study design and population

This Kenyan dataset is part of an ongoing longitudinal breast cancer genomics study conducted at AIC Kijabe and Aga Khan hospitals. Additionally, 15 Triple-Negative Breast Cancer (TNBC) samples from Gene Expression Atlas (GEO Accession:

GSM4223947) were included in the Kenyan study. Caucasian and African American datasets were retrospectively selected from the Cancer Genome Atlas data.

Study participants

The Kenyan cohort (n=60) represents all accessible RNA-seq data from Kenyan breast cancer patients during the study period. It comprised 45 samples from 23 patients at AIC Kijabe and Aga Khan hospitals and 15 TNBC samples from the GEO database (Accession: GSE142258). This sample size is larger than other published Kenyan studies (e.g., Saleh *et al.*, [9] (n = 20) and provided adequate statistical power for differential gene expression analysis. The Cancer Genome Atlas (TCGA) data included Caucasian (n=66) and African American (n=46) patients. Kenyan, African American, and Caucasian patients had a mean age of 48 years, 56 years, and 55 years, respectively. All samples were categorised by molecular subtype and Stage I - III (Table 1).

Inclusion and exclusion criteria

To minimise selection bias, all patients with a histologically confirmed diagnosis of breast cancer were eligible for inclusion in the longitudinal Kenyan study as long as they agreed to give a signed informed consent. Data with missing documentation of ethnicity or subtype classification of the tumour were excluded from the study.

Data collection of Kenyan samples

Tumour and adjacent non-tumour tissues were collected from each consenting patient and immediately frozen in liquid nitrogen and stored at -80 ° C at the Aga Khan University Hospital biorepository. The clinical data, that is the histology and immunohistochemistry/molecular subtyping, were entered into an Excel sheet soon after the collection of the samples. The data was collected in real time as soon as the results from the pathology lab were available.

Library processing

Fresh frozen tumour and adjacent normal samples from 45 participants were shipped on dry ice to the Laboratory of Human Carcinogenesis at the National Cancer Institute (USA) for total ribonucleic acid (RNA)

extraction and sequencing. As a quality control measure, samples from 23 participants whose biopsies had 50% tumour were selected for RNA extraction. The total RNA was isolated from frozen breast tumours and normal tissue using the TRIzol reagent (Thermo Fisher Scientific) with DNase treatment according to the manufacturer's instructions (21). The integrity of the isolated RNA was evaluated with the Agilent TapeStation High Sensitivity RNA ScreenTape (Agilent Technologies) following the manufacturer's guidelines, ensuring only high-quality RNA proceeded to sequencing (22).

RNA sequencing

Polymerase Chain Reaction (PCR) was used for RNA amplification before sequencing. The total RNA-seq sequencing was performed at the Sequencing Facility, Leidos Biomedical Research, Inc., Frederick National Laboratory for Cancer Research, USA. The samples were pooled and sequenced on NovaSeq S2 (Illumina) using NEBNext UltraII Directional RNA Library Prep and paired-end sequencing (23).

Data analysis

Kenyan breast cancer samples underwent quality control using FastQC v0.11.3 and MultiQC v1.13, with a Q30 cut-off (24,25). Low-quality bases and adapter sequences were removed using Cutadapt v3.1 (26). High-quality reads were aligned to the reference genome (Hg38) using HISAT2 v2.2.0, and alignment files were processed with

Samtools v1.3.1 (27,28). Gene-level counts were produced using HTSeq-count v2.0.2 (29).

TCGA data were pre-processed using STAR alignment and FPKM normalisation (30,31). Gene counts were normalised using FPKM and FPKM-UQ (30). Trimmed Mean of M-values normalisation was applied to all the samples. Differential gene expression analysis was conducted using edgeR (32) in R v4.2.0 (33), with visualisation in ggplot2 v3.3.6 (34). We set a threshold for differential expression at $FDR < 0.05$ and fold change ≥ 1 or ≤ -1 to ensure statistical significance while requiring at least a 2-fold expression change to capture biologically meaningful differences. Apoptotic genes were identified using GeneCards v5.7 (a consensus list of known apoptotic genes in Supplementary 1) (35). Shared and uniquely expressed apoptotic genes were compared between populations, subtypes, and stages for the Kenyan, Caucasian, and African American BC patients using custom Python scripts and Venn diagrams created with ggVennDiagram v1.2.2 (36). Gene Ontology (EnrichR v3.1) and gene set enrichment analysis (ClusterProfiler v4.0) identified key pathways (37,38). Protein-protein interaction networks of significantly expressed genes in each subtype and BC stage were generated at medium confidence (score ≥ 0.4) using the STRING database (39).

Ethical considerations

The study was approved by the Institutional Scientific Ethics Review Committee (ISREC) at The Aga Khan University in Nairobi, Kenya (Approval number 2018/REC-80 (v9)).

Table 1

Clinical and Demographic Characteristics of the Participants Included in the Study

Ethnicity	Subtypes				
	Luminal A	Luminal B	HER2-enriched	TNBC	Normal
Kenya - Mean age =48 yrs.	12 (20%)	4 (6.67%)	3 (5%)	18 (30%)	23 (38.33%)
African American - Mean =56 yrs.	12 (25.53%)	11 (23.40%)	6 (12.76%)	12 (25.53%)	6 (12.76%)
Caucasian - Mean = 55 yrs.	12 (20%)	12 (20%)	10 (16.67%)	12 (20%)	20 (33.33%)
Stages					
	Stage I	Stage II	Stage III	Normal	Undefined
Kenya	6	6	21	19	4
African American	15	20	5	6	0
Caucasian	8	24	9	20	5

* undefined = no staging information was provided in the original clinical data. These samples were excluded from stage-specific analyses but were included in overall population comparisons and subtype-specific analyses where staging was not required.

Results

Table 1 presents sample distribution by population, age, breast cancer subtype, and stage. Values are presented as n (%).

Shared and uniquely expressed apoptotic genes among Kenyan patients

The Caucasian BC patients exhibited the highest count of differentially expressed apoptotic genes (n=133 total), of which 105 genes were unique, while 28 genes were shared

with African American and/or Kenyan populations. Kenyan breast cancer patients had 29 total, while African American breast cancer patients had 30 total differentially expressed apoptotic genes, as seen in Figure 1 A- C in the appendix). The CDK1 gene was the only shared upregulated gene between Kenyan and African American (Figure 2 A- B). We found a total of nine unique upregulated apoptotic genes for the Kenyan dataset and fourteen unique downregulated apoptotic genes for the Kenyan patients (Table 2).

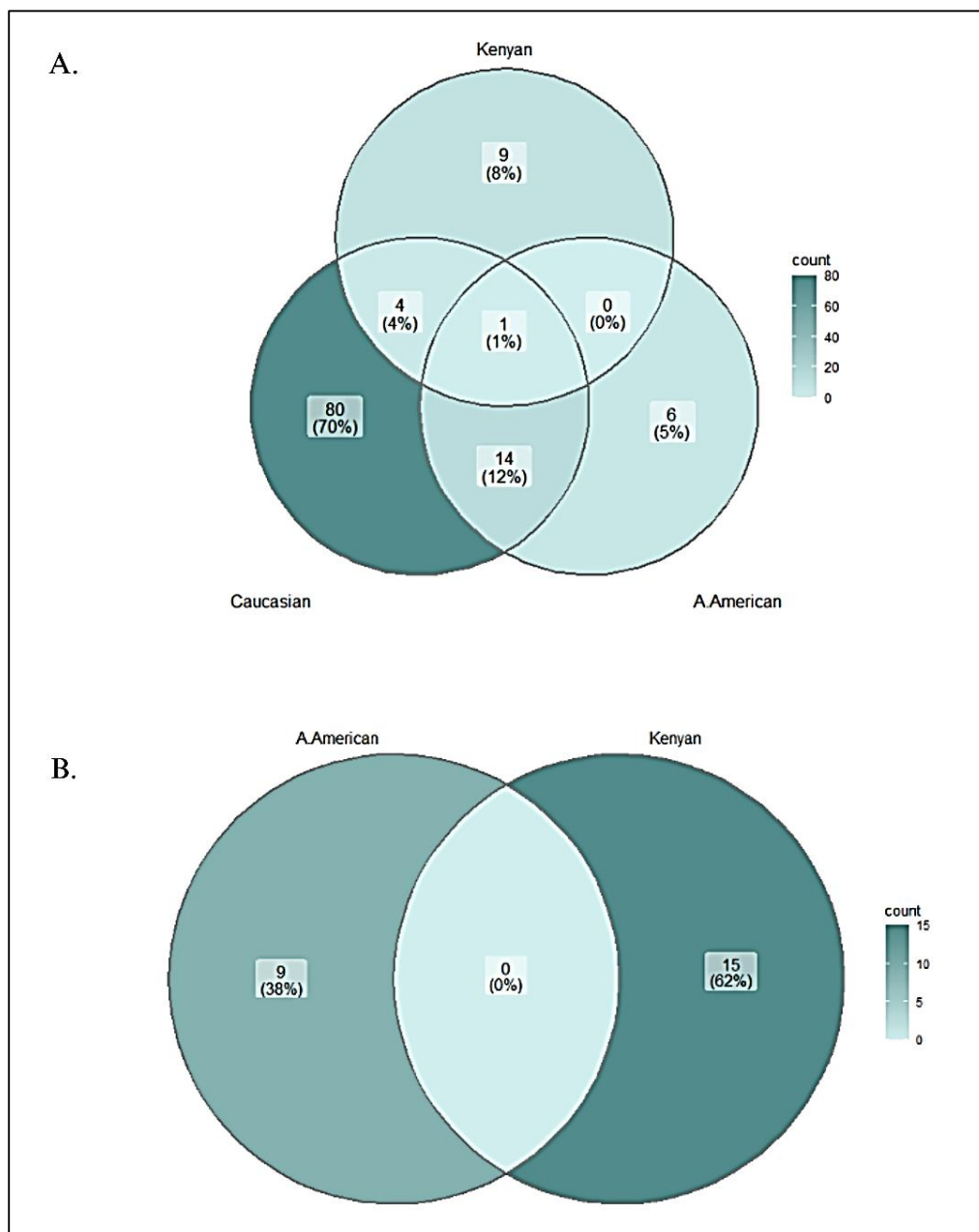


Figure 2

A. shared and unique upregulated genes per population, B. shared and unique downregulated genes, Kenyan vs African American women. For a detailed list of specific genes in each region, refer to Table 2.

Several biological processes were highly significantly enriched among Kenyan patients, including negative regulation of apoptosis (p-value= 2.0e-07), tumour necrosis factor-mediated pathways (p-value= 4.0e-07), and negative regulation of programmed cell death (p-value= 2.0e-07). The highly significant enrichment of these anti-apoptotic pathways suggests active suppression of cell death mechanisms in Kenyan breast tumours (Supplementary 2).

The cytokine-cytokine receptor interaction pathway, p53 signalling pathway, which is involved in intrinsic apoptosis, necroptosis involving RNF103-CHMP3, BIRC3, CHMP3, and IL1B genes, and endocytosis pathways were activated for the Kenyan patients. Adherens junction involving CREBBP and IGF1R genes, proteasome, spliceosome, and breast cancer pathway involving IGF1R and TNFSF11 were suppressed for the same patients (Supplementary 3).

Level of apoptotic gene expression based on subtype

Kenyan patients had 13 genes uniquely upregulated among patients with Luminal-A (including IGF1R, PPP2R2C, and PRKCZ), 15 genes among Luminal-B subtype patients (including PSMD3, CYCS, and CDK5R1), 11 genes in HER2-enriched subtype (including FASLG, HLA-DQA1, and GNG3), and 38 genes among patients with TNBC subtype (including RNF103-CHMP3, CHMP3, DSG1, and TNFRSF6B), (Supplementary 4). Moreover, the Luminal A subtype had 15 unique downregulated genes, the Luminal B subtype had 37 downregulated genes, the HER2-enriched subtype had 3 downregulated genes, and the triple-negative subtype had 68 downregulated genes, including PPP2R5A, YWHAG, MYC, and PRKAR1B. We visualised the top enriched pathways colored by statistical significance ($-\log_{10}$ FDR). The analysis of functional pathways revealed shared and unique significantly enriched pathways.

Table 2

Unique upregulated and downregulated genes by population

BC patients	Unique upregulated apoptotic genes	Unique down-regulated apoptotic genes
Kenyan	(n = 9) RNF103-CHMP3, DSG1, CHMP3, TNFRSF6B, DSG3, DCC, BIRC3, IL1B, GNB4	(n = 14) IGF1R, PTHLH, MAPT, PSMB8, KRT18, HSPA1A, PSMD3, HSPB1, SQSTM1, GNA14, PPP2R3B, ADCY1, DAPK3, HLA-A
Caucasian	(n = 80) GNG4, CDKN2A, CLSPN, LMNB1, CDC25A, BCL2A1, UNC5A, GZMB, TNFRSF18, PSMD3, TNFSF4, TNFSF18, PSMB3, BCL2L14, UNC5B, CDC25B, CD3D, HMGB2, RIPK2, CDK5R1, CD27, TNFRSF4, CD3E, TNFRSF13B, HLA-A, NFKBIE, MAPK15, PSME2, IL18, KRT18, HSPB1, LIMK1, FASLG, PARP1, BIK, PRDX1, BAK1, CD3G, HLA-C, HLA-DQA1, TNFSF13B, HLA-DOB, MAGED1, IL1A, BCL2L15, TNFSF15, PSMB9, SFN, CYCS, PSMB4, HLA-B, IKBKE, BMF, CD70, YWHAZ, PSMD11, TNF, HLA-F, PSMA5, PSMD14, GNAS, PTRH2, PSMC4, CASP6, MAPKAPK2, SMPD3, PIK3R3, IRF1, PPM1G, RELB, PSMB8, CASP3, PSMB2, DAPK1, PSME4, PRDX2, SLC25A5, HLA-DRB1, PTPA, HLA-DQA2	(n = 25) ELANE, TP63, TP53AIP1, FOS, GSN, Y_RNA, AIFM2, TNFRSF10D, PPP2R2B, CFL2, GAS2, MAPK10, JUN, DSG1, GNG7, ARHGAP10, BCL6, GCK, BOK, PPP2R1B, IRS1, AXIN2, EGF, VIM, BCL2L2
African-American	(n = 6) DAPK3, STUB1, PSMD4, GNB2, PSMB5, HRAS	(n = 1) ITGA6

Key: Population-specific differentially expressed genes are not shared across populations. Upregulated = higher in tumour vs. normal, Downregulated = lower in tumour vs. normal.

Programmed cell death, intercellular signal transduction, and apoptosis were enriched in multiple subtypes, suggesting common regulatory responses. Condition-specific pathways such as the Wnt signalling pathway and pluripotency (Luminal A), miRNA regulation and DNA damage response (Luminal B), positive regulation of cell death (TNBC), and positive regulation of response to stimulus (HER2+) revealed distinct biological processes activated or suppressed in a subtype-dependent manner (Figure 3).

The Kenyan patients with the TNBC subtype exhibited a higher number of downregulated apoptotic genes across all

subtypes (Supplementary 4). To explore functional associations among significantly expressed apoptotic genes in TNBC patients, we constructed protein–protein interaction (PPI) networks (confidence score threshold ≥ 0.4). Separate networks were generated for up- and down-regulated genes to highlight condition and direction-specific interaction patterns (Figure 4). The downregulated network was notably denser relative to the upregulated network, suggesting a higher degree of coordinated suppression of immune and cell death pathways in TNBC. Both networks showed significantly more protein interactions than would be expected by chance.

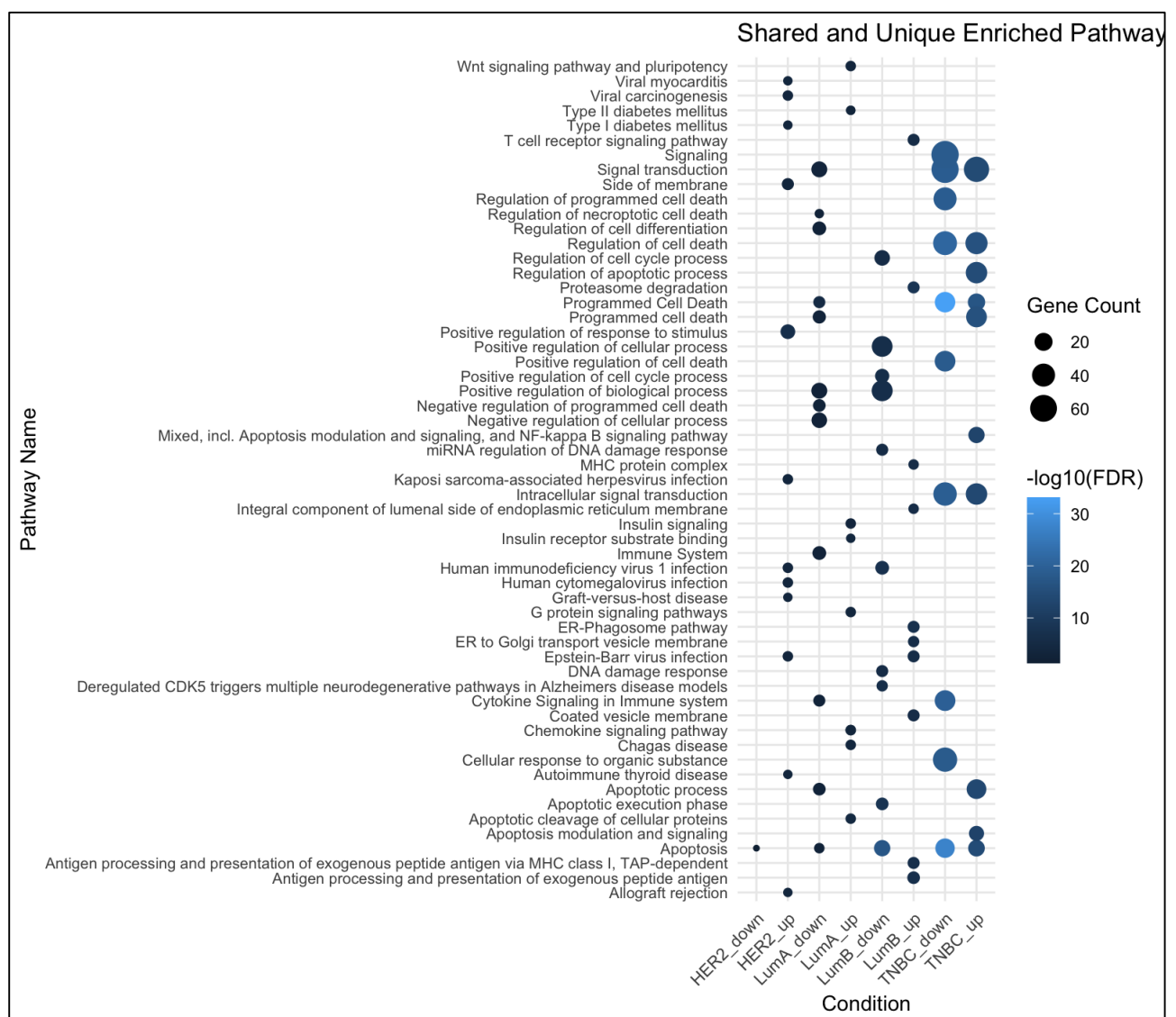


Figure 3
Bubble plot of the top significantly enriched pathways per subtype, based on up- or down-regulated gene sets. The size of the bubbles corresponds to the number of genes contributing to the enrichment, and the colour indicates the statistical significance, with lighter blue representing higher $-\log_{10}(\text{FDR})$ values.

This indicates that these differentially expressed proteins are not randomly co-expressed but rather form biologically meaningful, coordinated interaction networks, which results in combined functional consequences rather than independent effects. Up-regulated gene networks were enriched in apoptosis modulation and signalling, signal transduction, and regulation of cell death. Down-regulated networks were involved in cytokine signalling in immune systems and programmed death (Figure 3).

Level of apoptotic gene expression based on tumour stage

Comparing the expression levels of apoptotic genes between Stage I and Stage II, Kenyan patients had 152 total differentially expressed apoptotic genes. Among these, 8 genes were upregulated, while 144 genes were significantly underexpressed (Supplementary 5). In Stage I and Stage III, 200 apoptotic genes were significantly expressed. Seven of these genes were upregulated (Supplementary 4). However, for Stage II patients vs Stage III patients, two apoptotic genes were downregulated (PRKCZ and GSDMD). Shared and unique biological pathways revealed regulation of signalling and signal transduction, and regulation of cellular processes were uniquely enriched in Stage I (Figure 5).

Discussion

Due to the limited availability of African transcriptomic data, African American cohorts have often served as reference populations for African ancestry groups (40,41). This study addressed this gap by analysing Kenyan-specific transcriptomic data.

The study of BC tumour and control tissues from Kenyan BC patients (age, subtype, and tumour stage not adjusted) revealed differences in apoptotic gene expression. A comparison of the up- and down-regulated genes between African American BC patients and Kenyan patients revealed a single shared differentially expressed apoptotic gene. Our results showed more similarity in profiles of gene expression between African American and Caucasian BC patients (42).

Kenyan patients had a unique upregulation of DSG1, DCC, and IL1B; these genes have been previously reported to be pro-apoptotic genes (43). The DCC gene is a CASP3 substrate that works by activating the apoptotic pathway; it has also been reported to suppress apoptosis when the gene product binds to the netrin-1 ligand (19). The gene IL1B has also been widely studied in prostate cancer and found to suppress the proliferation of tumour cells; however, tumour cells over-exposed to the gene gained resistance to a variety of anticancer chemotherapy options in mice (43).

Our results suggest that the unique upregulated pro-apoptotic genes act as a protective mechanism against the proliferation of cancer cells for these patients. Well-studied genes like DSG3, BIRC3, TNFRSF6B, and GNB4 were among the upregulated anti-apoptotic genes uniquely found among Kenyan patients (44–47). According to Wang *et al.* (47), the growth and spread of breast cancer cells is linked to the overexpression of the gene GNB4. The BIRC3 gene is also reported to fast-track the progression of brain tumours in glioma (44). In another study, TNFRSF6B overexpression has been reported in patients with pancreatic cancer and linked to a poorer prognosis among these patients (46).

Our results suggest the growth of BC tumour cells and their resistance to apoptosis may be linked to the overexpression of the unique anti-apoptotic genes we identified. In our study, IGF1R, an inhibitor of the apoptotic pathway, was identified as being upregulated uniquely among Luminal A Kenyan BC patients (48).

A previous study on breast cancer patients in Brazil reported similar overexpression of IGF1R in their patients with Luminal-A and Luminal-B subtypes (49). This is further corroborated by Farabaugh *et al.* (50), who observed that the IGF1R was upregulated among samples with ER+ subtypes. Previous studies on IGF1R treatments have been unsuccessful and reported poor prognosis in high-expression, non-small-cell lung cancer patients (49,51).

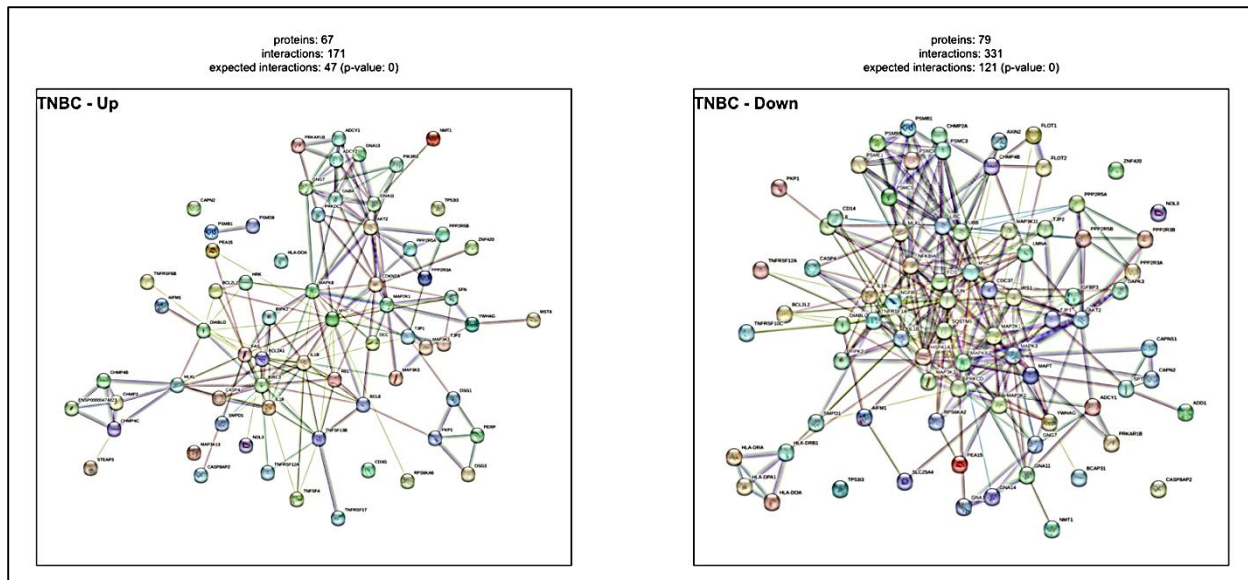


Figure 4

Protein-protein interaction networks of upregulated and downregulated genes in TNBC. Nodes represent proteins encoded by the uniquely expressed apoptotic genes, while edges represent known and predicted interactions.

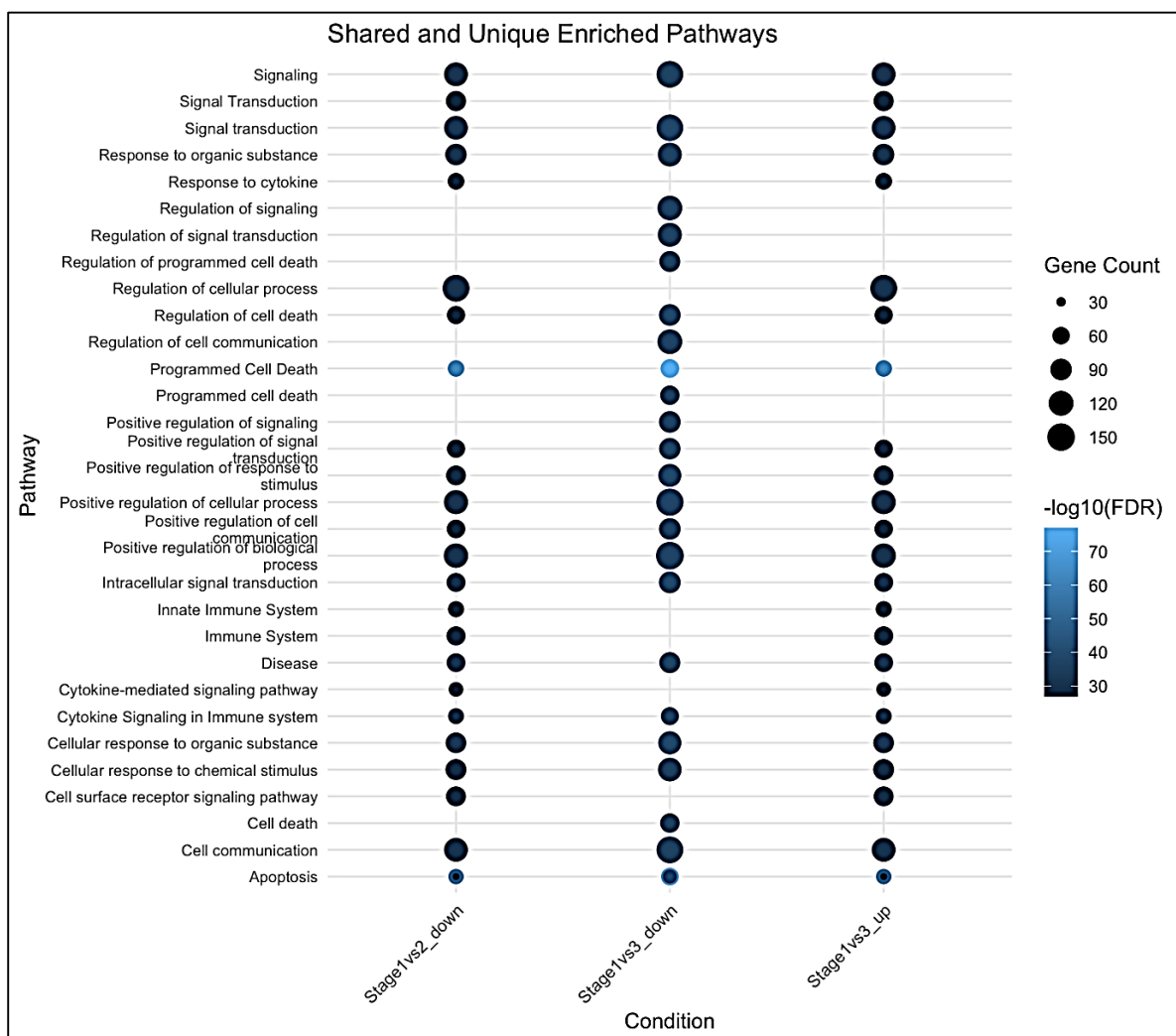


Figure 5

Bubble plot of the top significantly enriched pathways by stage, based on up- or down-regulated gene sets. The size of the bubbles corresponds to the number of genes contributing to the enrichment, and the colour indicates the statistical significance, with lighter blue representing higher $-\log_{10}(\text{FDR})$ values.

The presence of shared pathways across subtypes among Kenyan women with BC reflects a core response to cancer proliferation, while the distinct pathway enrichments underlie functional differences specific to the subtype. Our study analysed gene expression in Kenyan Stage I compared to Stage II BC patients, finding a higher proportion of downregulated genes in Stage I. This suggests that at early stages, there is suppressed pro-apoptotic signalling. When comparing Stage II to Stage III, both GSDMD (pro-apoptotic) and PRKCZ (anti-apoptotic) were downregulated in Stage III. Although PRKCZ downregulation would typically promote cell death, the simultaneous suppression of GSDMD creates a net survival advantage for tumour cells. This suggests that in advanced-stage tumours, suppression of pro-apoptosis outweighs any death-promoting effects from reduced anti-apoptotic signalling, ultimately favouring tumour progression.

Limitations

Some of the assumptions we made included homogeneity of the study populations and that subjects did not have other comorbidities or chronic illnesses that would affect gene expression. However, we could not adjust for additional potential confounders such as age, body mass index, comorbidities, treatment history, or environmental factors due to limited sample size and incomplete clinical data availability, particularly for the publicly available datasets. We also assumed relative homogeneity within each subtype and stage, though unmeasured confounding from patient characteristics or technical batch effects may have influenced results. Future studies with larger sample sizes and comprehensive clinical annotation would enable multivariable adjustment for these confounders. Stage-subtype interactions could not be examined simultaneously due to insufficient sample sizes. Our study population size was limited by the lack of previous studies among Kenyan patients.

This study was limited to comparing three populations. While Asian, Hispanic, and other ethnic populations also experience breast cancer with potentially unique molecular

profiles, their inclusion was beyond the scope and resources of this comparative analysis. Future studies should expand to include these populations for a more comprehensive understanding of global breast cancer heterogeneity.

Conclusion

In comparing apoptotic gene expression across all three populations, we identified distinct molecular signatures. Kenyan patients showed unique dysregulation of 23 genes (including IL1B and IGF1R), African American patients had 7 unique genes (including DAPK3 and STUB1), while Caucasian patients exhibited the largest set of unique genes, 105 genes, including an extensive HLA gene family. Despite some shared genes like CDK1 between Kenyan and African American populations, the predominance of population-specific dysregulation underscores substantial genetic heterogeneity in breast cancer across populations and ethnicities.

These findings have important policy implications. We recommend that health policymakers invest in genomic research infrastructure in Kenya and Africa, develop population-specific diagnostic and treatment guidelines, and prioritise funding for African-focused research. This would improve early detection, treatment efficacy, and survival outcomes for African breast cancer patients.

Author information

- Marion N. Nyamari- 0000-0003-0004-9982
- John Gitau- 0000-0001-8079-1102
- Kennedy Wanjau Mwangi- 0000-0003-2561-1521
- Carol W. Hunja- 0000-0001-8122-4927
- Sammy Wambua- 0000-0002-3041-2714
- Francis Makokha- 0000-0003-3324-4426

References

1. Kenya NCI. Cancer Awareness | National Cancer Institute Kenya [Internet]. [cited 2025 Dec 18]. Available from: <https://www.ncikenya.go.ke/awareness>
2. Kenya Launches National Cancer Control Strategy 2017-2022 | WHO | Regional Office for Africa [Internet]. 2024 [cited

- 2024 Nov 17]. Available from: <https://www.afro.who.int/news/kenya-launches-national-cancer-control-strategy-2017-2022>
3. Sayed S, Fan S, Moloo Z, Wasike R, Bird P, Saleh M, *et al.* Breast cancer risk factors in relation to molecular subtypes in breast cancer patients from Kenya. *Breast Cancer Research*. 2021 Jun 26;23(1):68. doi:10.1186/s13058-021-01446-3
4. Breast Cancer Status in Kenya – National Cancer Registry of Kenya (NaCaRe-KE) – National Cancer Institute Kenya [Internet]. 2025 Dec 4 [cited 2025 Dec 18]. Available from: <https://wp.ncikenya.go.ke/2024/11/15/breast-cancer-status-in-kenya-national-cancer-registry-of-kenya-nacare-ke/>
5. Antabe R, Kansanga M, Sano Y, Kyeremeh E, Galaa Y. Utilization of breast cancer screening in Kenya: what are the determinants? *BMC Health Services Research*. 2020 Mar 18;20(1):228. doi:10.1186/s12913-020-5073-2
6. Howlader N, Cronin KA, Kurian AW, Andridge R. Differences in Breast Cancer Survival by Molecular Subtypes in the United States. *Cancer Epidemiology, Biomarkers & Prevention*. 2018 May 31;27(6):619–26. doi:10.1158/1055-9965.EPI-17-0627
7. Hashmi AA, Aijaz S, Khan SM, Mahboob R, Irfan M, Zafar NI, *et al.* Prognostic parameters of luminal A and luminal B intrinsic breast cancer subtypes of Pakistani patients. *World Journal of Surgical Oncology*. 2018 Jan 2;16(1):1. doi:10.1186/s12957-017-1299-9
8. Saleh M, Chandrashekar DS, Shahin S, Agarwal S, Kim HG, Behring M, *et al.* Comparative analysis of triple-negative breast cancer transcriptomics of Kenyan, African American and Caucasian Women. *Translational Oncology*. 2021 Jul 1;14(7):101086. doi:10.1016/j.tranon.2021.101086
9. Pommier RM, Sanlaville A, Tonon L, Kielbassa J, Thomas E, Ferrari A, *et al.* Comprehensive characterization of claudin-low breast tumors reflects the impact of the cell-of-origin on cancer evolution. *Nat Commun*. 2020 Jul 9;11(1):3431. doi:10.1038/s41467-020-17249-7
10. Ramos-Solis N, Yeh ES. 6.04 - Triple Negative Breast Cancer. In: Kenakin T, editor. *Comprehensive Pharmacology* [Internet]. Oxford: Elsevier; 2022 [cited 2024 Nov 17]. p. 35–48. Available from: <https://www.sciencedirect.com/science/article/pii/B9780128204726000566> doi:10.1016/B978-0-12-820472-6.00056-6
11. Pfeffer CM, Singh ATK. Apoptosis: A Target for Anticancer Therapy. *International Journal of Molecular Sciences*. 2018 Feb;19(2):2. doi:10.3390/ijms19020448
12. Carneiro BA, El-Deiry WS. Targeting apoptosis in cancer therapy. *Nat Rev Clin Oncol*. 2020 Jul;17(7):395–417. doi:10.1038/s41571-020-0341-y
13. Jan R, Chaudhry G e S. Understanding Apoptosis and Apoptotic Pathways Targeted Cancer Therapeutics. *Adv Pharm Bull*. 2019 Jun 1;9(2):2. doi:10.15171/apb.2019.024
14. Zou R, Zhao W, Xiao S, Lu Y. A Signature of Three Apoptosis-Related Genes Predicts Overall Survival in Breast Cancer. *Front Surg*. 2022 May 17;9. doi:10.3389/fsurg.2022.863035
15. Thapa S, Rather RA, Singh SK, Bhagat M, Thapa S, Rather RA, *et al.* Insights into the Role of Defective Apoptosis in Cancer Pathogenesis and Therapy. In: *Regulation and Dysfunction of Apoptosis* [Internet]. IntechOpen; 2021 [cited 2024 Nov 17]. Available from: <https://www.intechopen.com/chapters/76842> doi:10.5772/intechopen.97536
16. Zhang J, Lu CY, Chen CH, Chen HM, Wu SY. Effect of pathologic stages on postmastectomy radiation therapy in breast cancer receiving neoadjuvant chemotherapy and total mastectomy: A Cancer Database Analysis. *Breast*. 2020 Dec;54:70–8. doi:10.1016/j.breast.2020.08.017 PubMed PMID: 32947148; PubMed Central PMCID: PMC7501458.
17. Kim C, Kim B. Anti-Cancer Natural Products and Their Bioactive Compounds Inducing ER Stress-Mediated Apoptosis: A Review. *Nutrients*. 2018 Aug;10(8):8. doi:10.3390/nu10081021
18. Sant DW, Mustafi S, Gustafson CB, Chen J, Slingerland JM, Wang G. Vitamin C promotes apoptosis in breast cancer cells by increasing TRAIL expression. *Sci Rep*. 2018 Mar 28;8(1):5306. doi:10.1038/s41598-018-23714-7
19. Chen J, Du H, Zhang Y, Chen H, Zheng M, Lin P, *et al.* Netrin-1 Prevents Rat Primary Cortical Neurons from Apoptosis via the DCC/ERK Pathway. *Front Cell Neurosci*.

- 2017 Dec 13;11.
doi:10.3389/fncel.2017.00387
20. Neophytou CM, Trougakos IP, Erin N, Papageorgis P. Apoptosis Deregulation and the Development of Cancer Multi-Drug Resistance. *Cancers*. 2021 Jan;13(17):17. doi:10.3390/cancers13174363
21. Rio DC, Ares M, Hannon GJ, Nilsen TW. Purification of RNA using TRIzol (TRI reagent). *Cold Spring Harb Protoc*. 2010 Jun;2010(6):pdb.prot5439. doi:10.1101/pdb.prot5439 PubMed PMID: 20516177.
22. Padmanaban A, Salowsky R, Cher C. RNA Quality Control using the Agilent 2200 TapeStation System.
23. Biolabs NE. Protocol for use with Purified mRNA or rRNA Depleted RNA and NEBNext Ultra II Directional RNA Library Prep ..., published on Dec 19, 2018 [Internet]. 2018 Dec 19 [cited 2026 Mar 8]. Available from: <https://www.protocols.io/view/protocol-for-use-with-purified-mrna-or-rrna-deplet-14egnx7w6l5d/v1>
24. Babraham Bioinformatics - FastQC: A Quality Control tool for High Throughput Sequence Data [Internet]. [cited 2024 Nov 17]. Available from: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
25. MultiQC: summarize analysis results for multiple tools and samples in a single report | Bioinformatics | Oxford Academic [Internet]. [cited 2024 Nov 17]. Available from: <https://academic.oup.com/bioinformatics/article/32/19/3047/2196507>
26. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*. 2011 May 2;17(1):1. doi:10.14806/ej.17.1.200
27. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-Based Genome Alignment and Genotyping with HISAT2 and HISAT-genotype. *Nature biotechnology*. 2019 Aug 2;37(8):907. doi:10.1038/s41587-019-0201-4 PubMed PMID: 31375807.
28. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, *et al*. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009 Aug 15;25(16):2078–9. doi:10.1093/bioinformatics/btp352 PubMed PMID: 19505943; PubMed Central PMCID: PMC2723002.
29. Putri GH, Anders S, Pyl PT, Pimanda JE, Zanini F. Analysing high-throughput sequencing data in Python with HTSeq 2.0. *Bioinformatics*. 2022 May 13;38(10):2943–5. doi:10.1093/bioinformatics/btac166
30. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, *et al*. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2012 Oct 25;29(1):15. doi:10.1093/bioinformatics/bts635 PubMed PMID: 23104886.
31. Picard Tools - By Broad Institute [Internet]. [cited 2025 Jan 6]. Available from: <https://broadinstitute.github.io/picard/>
32. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data | Bioinformatics | Oxford Academic [Internet]. [cited 2025 Jan 6]. Available from: <https://academic.oup.com/bioinformatics/article/26/1/139/182458>
33. R: The R Project for Statistical Computing [Internet]. [cited 2025 Jan 6]. Available from: <https://www.r-project.org/>
34. Wickham H. ggplot2: Elegant Graphics for Data Analysis [Internet]. New York, NY: Springer; 2009 [cited 2025 Jan 6]. Available from: <https://link.springer.com/10.1007/978-0-387-98141-3> doi:10.1007/978-0-387-98141-3
35. Safran M, Rosen N, Twik M, BarShir R, Stein TI, Dahary D, *et al*. The GeneCards Suite. In: Abugessaisa I, Kasukawa T, editors. *Practical Guide to Life Science Databases* [Internet]. Singapore: Springer Nature; 2021 [cited 2025 Jan 6]. p. 27–56. Available from: https://doi.org/10.1007/978-981-16-5812-9_2 doi:10.1007/978-981-16-5812-9_2
36. Gao CH, Yu G, Cai P. ggVennDiagram: An Intuitive, Easy-to-Use, and Highly Customizable R Package to Generate Venn Diagram. *Front Genet*. 2021 Sep 7;12. doi:10.3389/fgene.2021.706907
37. Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, *et al*. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Research*. 2016 Jul 8;44(W1):W90–7. doi:10.1093/nar/gkw377
38. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS*. 2012 May;16(5):284–7. doi:10.1089/omi.2011.0118 PubMed PMID:

- 22455463; PubMed Central PMCID: PMC3339379.
39. Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, *et al.* The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res.* 2023 Jan 6;51(D1):D638–46. doi:10.1093/nar/gkac1000 PubMed PMID: 36370105; PubMed Central PMCID: PMC9825434.
40. Pinder LF, Nelson BD, Eckardt M, Goodman A. Article Commentary: A Public Health Priority: Disparities in Gynecologic Cancer Research for African-Born Women in the United States. *Clin Med Insights Womens Health.* 2016 Jan 1;9:CMWH.S39867. doi:10.4137/CMWH.S39867
41. Full article: African American and Black Caribbean feelings of closeness to Africans [Internet]. [cited 2025 Jan 27]. Available from: <https://www.tandfonline.com/doi/full/10.1080/1070289X.2016.1208096>
42. Huo D, Hu H, Rhie SK, Gamazon ER, Cherniack AD, Liu J, *et al.* Comparison of Breast Cancer Molecular Features and Survival by African and European Ancestry in The Cancer Genome Atlas. *JAMA Oncology.* 2017 Dec 1;3(12):1654–62. doi:10.1001/jamaoncol.2017.0595
43. Rébé C, Ghiringhelli F. Interleukin-1 β and Cancer. *Cancers.* 2020 Jul;12(7):7. doi:10.3390/cancers12071791
44. Gressot LV, Doucette T, Yang Y, Fuller GN, Manyam G, Rao A, *et al.* Analysis of the inhibitors of apoptosis identifies BIRC3 as a facilitator of malignant progression in glioma. *Oncotarget.* 2016 Apr 8;8(8):12695–704. doi:10.18632/oncotarget.8657
45. Li X, Ahmad US, Huang Y, Uttagomol J, Rehman A, Zhou K, *et al.* Desmoglein-3 acts as a pro-survival protein by suppressing reactive oxygen species and doming whilst augmenting the tight junctions in MDCK cells. *Mechanisms of Ageing and Development.* 2019 Dec 1;184:111174. doi:10.1016/j.mad.2019.111174
46. Zhang C, Li H, Huang Y, Tang Y, Wang J, Cheng Y, *et al.* Integrative analysis of TNFRSF6B as a potential therapeutic target for pancreatic cancer. *Journal of Gastrointestinal Oncology.* 2021 Aug;12(4). doi:10.21037/jgo-21-303
47. Wang B, Li D, Rodriguez-Juarez R, Farfus A, Storzynsky Q, Malach M, *et al.* A suppressive role of guanine nucleotide-binding protein subunit beta-4 inhibited by DNA methylation in the growth of anti-estrogen resistant breast cancer cells. *BMC Cancer.* 2018 Aug 13;18(1):817. doi:10.1186/s12885-018-4711-0
48. Zhang JX, Yan XJ, Wu S, Liu Q, Ma JM. KLF16 overexpression deleteriously affects the proliferation and migration of retinoblastoma by transcriptionally repressing BCL2L15. *Biochem Biophys Res Commun.* 2020 Sep 3;529(4):977–83. doi:10.1016/j.bbrc.2020.06.027 PubMed PMID: 32819608.
49. Rocha FS, Maués JH da S, Pereira CMBL, Moreira-Nunes CA, Burbano RMR. Analysis of Increased EGFR and IGF-1R Signaling and Its Correlation with Socio-Epidemiological Features and Biological Profile in Breast Cancer Patients: A Study in Northern Brazil. *BCTT.* 2021 May 21;13:325–39. doi:10.2147/BCTT.S308554
50. Farabaugh SM, Litzenger BC, Elangovan A, Pecar G, Walheim L, Atkinson JM, *et al.* IGF1R constitutive activation expands luminal progenitors and influences lineage differentiation during breast tumorigenesis. *Developmental Biology.* 2020 Jul 1;463(1):77–87. doi:10.1016/j.ydbio.2020.04.007
51. Hua H, Kong Q, Yin J, Zhang J, Jiang Y. Insulin-like growth factor receptor signaling in tumorigenesis and drug resistance: a challenge for cancer therapy. *Journal of Hematology & Oncology.* 2020 Jun 3;13(1):64. doi:10.1186/s13045-020-00904-3



Supplementary data

Supplementary 1

Laboratory processing and sequencing of RNA

The tissue samples were frozen in liquid nitrogen immediately after surgical excision and stored at -80 °C at the Aga Khan University Hospital biorepository. Pathological information on the clinical sample was abstracted from the medical records and included patient's age, tumor grade, tumor stage, and hormone receptor status, among others (Table 1). Fresh frozen tumor and adjacent normal samples from 45 participants were shipped on dry ice to the Laboratory of Human Carcinogenesis at the National Cancer Institute (USA) for RNA extraction and sequencing. After examination, samples from 23 participants that at least 50% tumor were selected for RNA extraction.

The total ribonucleic acid (RNA) was isolated from frozen breast tumors and normal tissue using the TRIzol reagent (Thermo Fisher Scientific) with DNase treatment. The integrity of the isolated RNA was evaluated with the Agilent TapeStation High Sensitivity RNA ScreenTape (Agilent Technologies). Polymerase Chain Reaction (PCR) was used for RNA amplification before sequencing. The total RNAseq sequencing was performed at the Sequencing Facility, Leidos Biomedical Research, Inc., Frederick National Laboratory for Cancer Research, USA. The samples were pooled and sequenced on NovaSeq S2 using NEBNext UltraII Directional RNA Library Prep and paired-end sequencing.



Supplementary 2

Apoptotic genes

CONSENSUS LIST OF APOPTOTIC GENES							
SLC25A4	SLC25A5	SLC25A6	APAF1	FAS	BAD	BAK1	BAX
BID	BIK	CASP8	CASP9	IGF1R	MCL1	MT-RNR2	PMAIP1
HRK	AIFM1	BCL2L10	BCL2L11	BBC3	CYCS	DIABLO	MOAP1
LMNA	LMNB1	CASP6	PARP1	CASP2	DFFB	FASLG	CASP3
DFFA	FADD	MAP3K1	ARHGDIB	DAXX	PAK2	CRADD	LMNB2
RIPK1	ACTA1	BIRC3	SPTAN1	AKT2	BIRC2	MAPK8	MAP2K4
TRAF2	PRKDC	RB1	RIPK2	TNFRSF10 B	CASP4	HSPB1	MAP3K5
PAK1	PTK2	APP	TNFRSF10 A	PIDD1	TOP1	IGF1	LIMK1
GSN	SREBF1	SATB1	MADD	IKBKE	CFL2	TNFSF10	MAPK10
PRF1	CASP1	CHUK	IKBKB	ACTG1	IKBKG	MAPK9	MAP3K14
BTK	SMPD3	TRAF1	UBE2I	E2F1	ACTB	MAP2K1	MAPKAPK2
SMPD2	TFDP1	TFDP2	TP53	SFN	MAP3K7	MAPKAPK 3	MAPK1
GSDME	NAIP	C1QBP	CARD8	UACA	PPP3CC	PPP3R1	GSDMD
AVEN	DYNLL1	NMT1	PPP1R13B	YWHAG	YWHAH	YWHAZ	PRKD1
TP63	MIR29B2	TP53BP2	APIP	TNFRSF10 C	BIRC5	NFKBIE	TNFRSF4
MIR29B1	PTRH2	TP73	YWHAE	TNFRSF10 D	NFKBIB	TNFSF11	BCL2L13
PEA15	TNFRSF1 B	YWHAH	AIFM2	TNFRSF25	TNFSF15	PPP2CB	GNB4
TNFRSF17	IRAK2	BIRC7	NFKBIA	TRAF6	GNB1	GNAS	PIK3CG
TNFRSF12A	NGF	NGFR	TRAF5	TNFSF4	EGFR	PPP2CA	GNG2
IRAK1	BCL2L12	NFKB2	TNFSF8	PRKCZ	PIK3CD	BOK	PIK3CA
TNFSF13B	PTHLH	ITGB4	ITGA6	RPS6KA2	GSK3B	GNG3	GNG8
BCL2L15	TRAF3	PIK3R3	RPS6KA1	BCL2L14	GNB3	PDP1	PPP3CA
GSK3A	TNFRSF8	GNG11	GNB5	PDPK1	GNB2	PPP3CB	GNA13
ADCY1	HRAS	PRKACA	GNGT2	MAP2K2	GNA11	GNG13	PPP2R5B
GRB2	GNG5	GNGT1	PRKAR1A	GNG10	PTPA	GNA12	PPP2R5C
PRKAR2B	CDK1	PRKACB	GNG7	PPP1CA	PPP2R5A	BIRC8	MAP3K11
PPM1A	PPM1B	PPM1G	PPM1F	PPP2R3A	OMA1	TNFRSF21	CHMP2A
SHC1	PPM1D	BECN1	EGF	BAG4	PSME4	CHMP6	TICAM2
PIK3R2	ADCY3	PPP2R2B	PPP2R2C	CD70	H1-0	IGFBP3	FLOT2
RPS6KB2	PPP2R2A	RPS6KA5	RPS6KA6	STK26	PKP1	PERP	PSMC6
ADCY2	RPS6KA4	CDKN1A	CD27	MAP3K3	PSMB4	HLA-DQA1	RPS27A
PPP2R1B	HSPA1A	FOS	ENDOG	PSMB3	PSME3	DBNL	PSMD11
PPP2R3B	GCK	MAP3K2	LTBR	PSMC5	UNC5B	APC	PRKCQ
RELA	MAP2K3	PSMD14	BCAP31	DCC	OPA1	PSMA7	PSMB2
MAPK7	NOL3	HSP90AB 1	MAP3K13	DNM1L	PRKCD	PSMB1	CHMP2B
MAPK6	HSP90AA 1	BEX3	MAPK14	CTNNB1	PSMA6	PSMC3	PRDX2
SIVA1	ELK1	MAP2K5	OCLN	PSMC1	PSMC2	MLKL	HLA-DPA1
TNFSF18	ATF2	TMED7-TICAM2	TICAM1	PSMD7	PSMD8	IRF2	CD3E
HIPK2	PSMA8	HMGB2	PSMD9	PSMC4	ARHGAP10	SDCBP	HLA-DOA
TNFRSF18	DSP	MAPT	PSMD10	ROCK1	PSMD6	PSMF1	HLA-F
MAPK15	APPL1	HMGB1	TJP1	CD14	TLR4	ELANE	SOD2
H1-1	PSMA5	PSMB10	STK24	CHMP4B	IRF1	CHMP3	CDC25B
DSG3	PSMA4	PSMD5	CDC37	TJP2	BNIP3L	TP53AIP1	CD3D
LY96	H1-5	PSMA3	PSMB9	IL18	NLRC4	CHM	TNFRSF11 A



PSMB11	KPNB1	PSMB8	PSMD4	CHMP4C	ATM	BCL2L1	GNG4
DSG2	PSMA2	PSME2	STAT3	UNC5A	NDRG1	VDAC1	PRKACG
H1-4	PSMB7	UBB	UBC	RIPK3	HLA-DMB	NUMA1	SOS2
KPNA1	PSMD3	FLOT1	STUB1	IL1B	HLA-E	CASP8AP2	TNFSF14
FOXO3	PSME1	CASP5	OGT	CHMP7	HLA-DMA	GZMB	DAPK3
CAPN2	BMX	HLA-C	HLA-DRB5	CDC25A	TNFRSF13B	KRT18	ADD1
CDK5	UBA52	HLA-DRB4	CAPNS2	CDK5R1	CD40	HTRA2	DSG1
CAPN1	STEAP3	HLA-A	HLA-B	CAST	PLEC	FAF1	PSMD1
DYNLL2	PSMD13	BABAM2	PTPN13	GNA15	FNTA	PELI1	HLA-DOB
BCL2	BCL2A1	MAPK3	RELB	RPS6KA3	AXIN1	HLA-G	PRDX1
RAF1	TNFRSF1A	IL1A	MYC	BIRC6	SEM1	ACIN1	ZNF420
BMF	AKT1	SEPTIN4	SOS1	PRKAR1B	DAPK2	CDC25C	TMEM219
CASP7	TNFRSF6B	CDKN2A	PIK3CB	GNG12	DAPK1	PRELID3A	BCL6
XIAP	AKT3	REL	PPM1L	GNA14	MAP3K8	RABGGTA	BCL2L2
CFLAR	TFAP2A	TNFSF13	MAPK12	CDH1	PSMD12	CD247	VDAC2
SUMO1	NFKB1	PIK3R1	PTPN11	H1-2	CREBBP	LIPF	CASP10
IRS1	NFX1	RPS6KB1	PPP2R5E	MAGED1	TP53INP1	PSMA1	TNF
GAS2	YWHAQ	PRKAR2A	TNFSF12	RABGGTB	CD3G	PSMB6	TRADD
HLA-DQA2	H1-3	TP53I3	PPP2R1A	PDCD6IP	CD40LG	PSMD2	JUN
GOLGA2	HLA-DRA	PRELID1	PPP3R2	CHMP4A	SQSTM1	CLSPN	VIM
HLA-DRB1	CAPNS1	HLA-DQB1	AXIN2	TRIAP1	PPTC7	RNF103-HMP3	SMPD1
HLA-DRB3	PSMB5	ACTC1	ERCC6	AIF1L	BAG1	BCL10	BAK1P1
BLK	YWHAEP5						

Supplementary 3

Unique upregulated and downregulated genes by population

BC patients	Unique upregulated apoptotic genes	Unique down regulated apoptotic genes
Kenyan	(n = 9) RNF103-CHMP3, DSG1, CHMP3, TNFRSF6B, DSG3, DCC, BIRC3, IL1B, GNB4	(n = 14) IGF1R, PTHLH, MAPT, PSMB8, KRT18, HSPA1A, PSMD3, HSPB1, SQSTM1, GNA14, PPP2R3B, ADCY1, DAPK3, HLA-A
Caucasian	(n = 80) GNG4, CDKN2A, CLSPN, LMNB1, CDC25A, BCL2A1, UNC5A, GZMB, TNFRSF18, PSMD3, TNFSF4, TNFSF18, PSMB3, BCL2L14, UNC5B, CDC25B, CD3D, HMGB2, RIPK2, CDK5R1, CD27, TNFRSF4, CD3E, TNFRSF13B, HLA-A, NFKBIE, MAPK15, PSME2, IL18, KRT18, HSPB1, LIMK1, FASLG, PARP1, BIK, PRDX1, BAK1, CD3G, HLA-C, HLA-DQA1, TNFSF13B, HLA-DOB, MAGED1, IL1A, BCL2L15, TNFSF15, PSMB9, SFN, CYCS, PSMB4, HLA-B, IKBKE, BMF, CD70, YWHAZ, PSMD11, TNF, HLA-F, PSMA5, PSMD14, GNAS, PTRH2, PSMC4, CASP6, MAPKAPK2, SMPD3, PIK3R3, IRF1, PPM1G, RELB, PSMB8, CASP3, PSMB2, DAPK1, PSME4, PRDX2, SLC25A5, HLA-DRB1, PTPA, HLA-DQA2	(n = 25) ELANE, TP63, TP53AIP1, FOS, GSN, Y_RNA, AIFM2, TNFRSF10D, PPP2R2B, CFL2, GAS2, MAPK10, JUN, DSG1, GNG7, ARHGAP10, BCL6, GCK, BOK, PPP2R1B, IRS1, AXIN2, EGF, VIM, BCL2L2
African-American	(n = 6) DAPK3, STUB1, PSMD4, GNB2, PSMB5, HRAS	(n = 1) ITGA6

Supplementary 4

Unique upregulated and downregulated genes based on subtypes for the Kenyan population

Subtypes	No. of upregulated apoptotic genes	No. of downregulated apoptotic genes	Unique upregulated apoptotic genes	Unique downregulated apoptotic genes
Luminal A	13	15	IGF1R, PPP2R2C, PRKCZ, MAPT, GNA14, OCLN, ADCY1, PIDD1, PRELID3A, IKBKB, PPP2R3B, ACIN1, PRKCD	CFLAR, ITGA6, PDP1, SDCBP, PRKACA, BNIP3L, PELI1, RPS6KA3, TLR4, AKT3, PPM1L, TNFSF14, SOD2, TNFRSF21, PKP1
Luminal B	15	37	PSMD3, CYCS, CDK5R1, HLA-F, HLA-A, CD3D, PTRH2, LMNB2, HLA-DQB1, SREBF1, PSMA6, PSMB2, CDC25B, HLA-C, MAGED1	CDH1, H1-5, MAPK15, CLSPN, CYCS, H1-3, GNG13, CDK1, CDC25A, UNC5A, CDK5R1, HLA-F, BIRC5, E2F1, SMPD3, LMNB1, TNFSF15, TFAP2A, GNB3, CDC25C, DSP, CD3D, LIMK1, PTRH2, UNC5B, TP73, BIK, LMNB2, HLA-DQB1, IKBKE, PSMA6, PSMB2, H1-4, CDC25B, PIK3R2, BBC3, MAGED1
HER2+	11	3	FASLG, HLA-DQA1, GNG3, PSMD8, TRAF5, BMF, PIK3CG, HLA-B, ACTB, MAPKAPK2, TRAF1	PTPN13, NGF, HLA-DRB5
TNBC	38	68	RNF103-CHMP3, CHMP3, DSG1, TNFRSF6B, DSG3, STK26, CDKN2A, CD3G, HRK, BCL2A1, ADCY2, BIRC3, GNB4, PERP, MAP3K13, PIK3R3, DCC, CHMP4C, RPS6KA6, FAS, TNFSF4, TNFRSF17, STEAP3, REL, PRKDC, TNFSF13B, SFN, PKP1, RIPK2, IL1B, GNA13, CASP8AP2, DIABLO, PPP2R3A, MAPK8, IL18, MAP3K2, ZNF420	PPP2R5A, YWHAG, MYC, PRKAR1B, MAP2K1, MLKL, CHMP4B, GNA11, PEA15, BCL6, ADCY1, MAP3K3, CASP4, PSMD8, BCL2L2, NMT1, CAPN2, AKT2, TP53I3, AIFM1, NOL3, HLA-DOA, TJP2, PPP2R5B, TNFRSF12A, GNG7, SMPD1, PSMB1, TJP1, TNFRSF1A, PSMC5, FLOT2, UBC, IGFBP3, MAP3K11, SLC25A4, TNFRSF10C, PRKCD, CAPNS1, AXIN2, DAPK3, PPP2R3B, SPTAN1, PSMC3, ADD1, PSMB6, CDC37, IRS1, HLA-DRA, NFKBIA, CD14, HLA-DRB1, RPS6KA2, HLA-DPA1, NGFR, HSPA1A, MAPK3, BCAP31, SQSTM1, LMNA, GNA14, UBB, CHMP2A, PSME1, JUN, MAPT, FLOT1, FOS

Supplementary 5

Upregulated and downregulated apoptotic genes based on the stage of BC at time of diagnosis

BC patients	Stages	Upregulated	Downregulated
Kenyan	stage I vs stage II	(n = 8) DSG1, DSG3, RNF103-CHMP3, CHMP3, TNFRSF6B, EGFR, PRKACB, PPM1L	(n = 144) DBNL, IRF2, RIPK1, STUB1, DYNLL1, RELA, SHC1, ACTB, NFKB1, NMT1, AVEN, PPM1G, MAP3K14, TJP2, PSMF1, MADD, PSMD13, ACTG1, PDPK1, RPS6KA2, PLEC, HSP90AA1, SMPD2, MLKL, GNG11, TP73, ARHGAP10, PPP2R5B, SEM1, BEX3, PPP1R13B, LMNA, LIMK1, SPTAN1, UBC, RPS6KA4, PSMB5, AKT2, TNFRSF4, MAP2K2, PTRH2, BCL2L12, PIK3R2, GNB2, AIFM2, GNA11, GSK3A, LTBR, PSMD9, TNFRSF1A, PSMB6, PSMB1, CAPN2, PTPA, RPS6KB2, PPP2R1A, PRKAR1B, CREBBP, TRADD, AIFM1, UBE2I, MAPKAPK3, BAX, MAP2K3, PSMC5, ACIN1, MAPK7, AXIN1, MAPKAPK2, TJP1, PSMC3, H1-3, IKBKB, H1-4, HSP90AB1, MAP3K11, NFKB2, CDKN1A, PARP1, ITGB4, ADD1, GNA15, TP53I3, CHMP4B, AKT1, PRKAR2B, MAPK3, NFKBIA, H1-2, UNC5B, TICAM1, H1-0, HLA-E, CAPN1, NOL3, IRS1, JUN, PSME2, SMPD1, IKBKG, CDC37, CD14, SLC25A4, FADD, IGFBP3, DAPK3, PSMD8, BBC3, TNFRSF13B, CAPNS1, BCAP31, IGF1R, UBB, CDH1, TNFRSF12A, FLOT2, SREBF1, NAIP, ADCY1, DAXX, PSMD3, CHMP2A, UNC5A, HLA-A, PRKCD, KRT18, HLA-B, GNA14, HLA-C, HLA-DRB3, SQSTM1, HLA-DRA, PPP2R2C, HLA-DOA, HSPB1, PSMB8, RPS6KA1, HLA-DPA1, SLC25A6, MAPT, PSMB3, FLOT1, PPP2R3B, HSPA1A
	Stage I vs stage III	(n = 7) RNF103-CHMP3, DSG1, CHMP3, TNFRSF6B, EGFR, PPM1L, CFL2	(n = 193) TMEM219, MAPK9, GNA12, HIPK2, BMF, PTPN13, CASP6, SHC1, TOP1, ARHGAP10, IRAK1, RIPK1, CASP8, CASP4, PSMD14, SLC25A5, ACTB, TNFRSF1A, C1QBP, GNAS, RPS6KA2, HMGB2, TJP2, MAGED1, PTRH2, BAD, PDPK1, NFKB1, PAK1, CASP7, HSP90AA1, PSMF1, GNG7, RELA, PSMB5, DBNL, GNG11, PIK3CD, PSMB1, STUB1, MAPK7, IRF2, PSMD4, ACTG1, PPM1G, DYNLL1, TRAF2, TFAP2A, NMT1, LMNB1, PSMD13, LMNB2, TNFRSF4, AVEN, PPP2CB, PPP2R5B, MAPKAPK2, AKT2, LIMK1, CD40, BIK, GSDMD, AIFM1, CHMP6, LTBR, MADD, TRADD, AIFM2, GNB2, SMPD2, GOLGA2, PPP2R1A, BEX3, CD14, MAPK12, UBC, NFKBIE, PSMB6, UBE2I, GNA11, GNA15, LMNA, MAP3K14, PPP1R13B, TP53I3, NFKBIA, APP, CAPN2, PPP1CA, PTPA, MLKL, GSK3A, PIDD1, SPTAN1, MAPKAPK3, CHMP4C, IKBKE, H1-2, SREBF1, JUN,

BC patients	Stages	Upregulated	Downregulated
			UNC5B, OCLN, MAP2K2, PIK3R2, TJP1, SMPD1, MAP3K11, PSMD3, TP73, SEM1, PRELID3A, RPS6KB2, RPS6KA4, H1-3, PARP1, HRAS, HSP90AB1, H1-4, GNG13, NUMA1, AKT1, NOL3, FOS, CREBBP, SLC25A4, CAPN1, PSMD9, PSMD8, PLEC, PSMC3, AXIN1, IGFBP3, PRKCZ, PRKAR2B, MAPK3, AXIN2, BAX, RELB, ADCY1, TICAM1, ADD1, TNFRSF12A, PTHLH, BCL2L12, IKBKG, CDH1, ITGB4, PSMC5, BCAP31, NFKB2, CAPNS1, H1-0, FLOT2, PRKAR1B, IKBKB, NGFR, FADD, PSME2, CHMP4B, ACIN1, UBB, TNFRSF13B, IRS1, HLA-F, DAPK3, CDC37, BBC3, HLA-A, NAIP, CHMP2A, PRKCD, DAXX, UNC5A, IGF1R, HLA-B, KRT18, PPP2R2C, GNA14, HLA-DRB3, HLA-C, HLA-DRA, HSPB1, HLA-DOA, SQSTM1, PSMB8, HLA-DPA1, PSMB3, RPS6KA1, SLC25A6, MAPT, FLOT1, PPP2R3B, HSPA1A
	stage II vs stage III	(n = 0)	(n = 2) PRKCZ and GSDMD
Caucasian	stage I vs stage II	(n = 3) EGFR, HRK (unique), and TNFSF18 (unique)	(n = 0)
	Stage I vs stage III	(n = 0)	(n = 0)
	stage II vs stage III	(n = 0)	(n = 1) BCL2A1 (unique)
African-American	stage I vs stage II	(n = 1) TNFSF11 (unique)	(n = 0)
	Stage I vs stage III	(n = 0)	(n = 0)
	stage II vs stage III	(n = 0)	(n = 2) ACTA1, and PRKACB

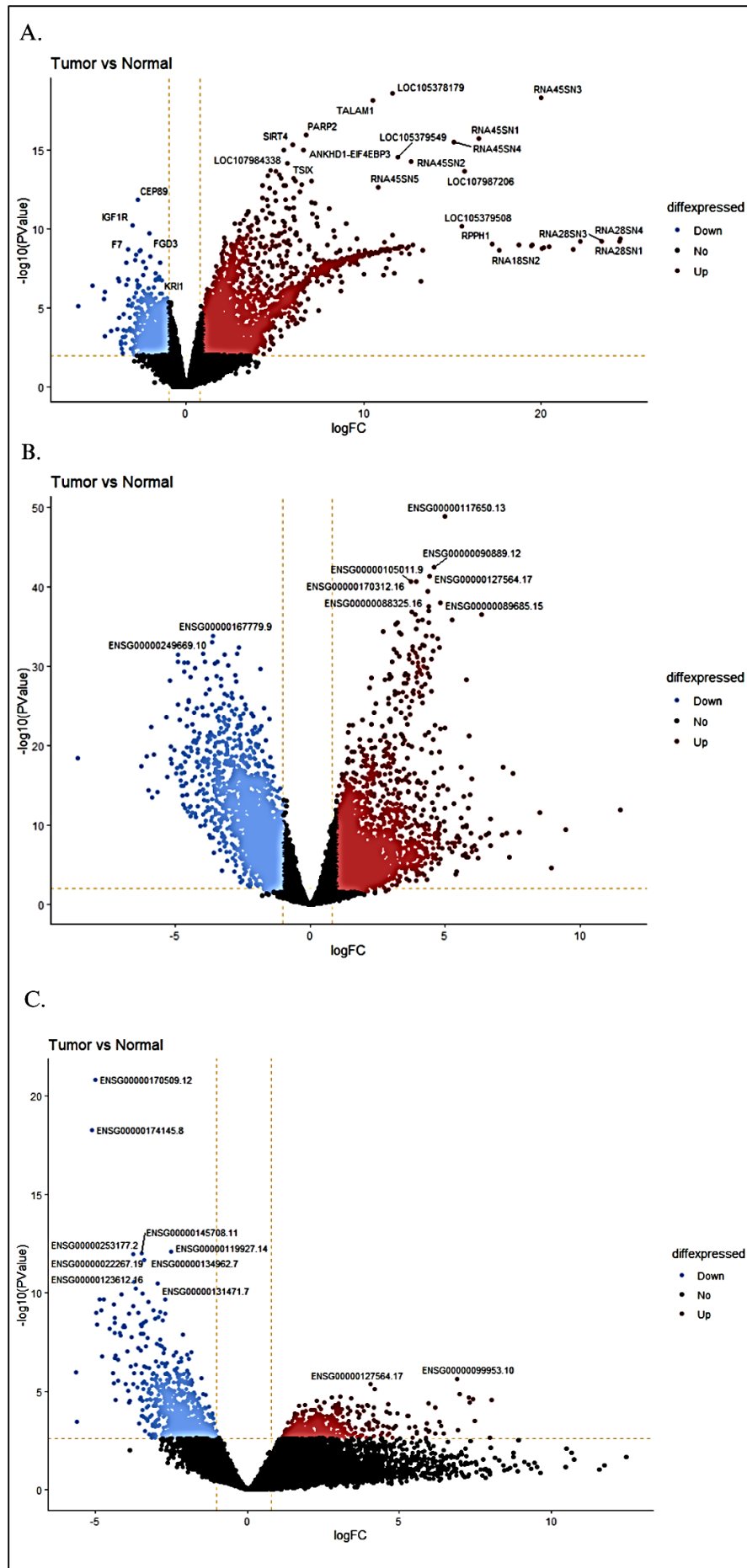


Figure 1.
A population-based volcano plot of the differentially expressed genes (DEG) in tumours vs control samples: A. African American, B. Caucasian, C. Kenyan. Red points represent significantly upregulated genes in tumour tissue, blue points represent significantly downregulated genes in tumour tissue, and grey points represent non-significantly DEG.