ISSN: 2579-079X

International Digital Organization for Scientific Research IDOSR JST/26/1220000
IDOSR JOURNAL OF SCIENCE AND TECHNOLOGY 12(2):113-117, 2026.
https://doi.org/10.59298/IDOSR/JST/26/122.113117

Genomic Studies and Biomarker Discovery for Cancer in East Africa Current Evidence, Sequencing Gaps and Priorities for Locally Relevant Diagnostics

Muhindo Anitah

Department of Pharmacy Kampala International University Uganda

Email: anitah.muhindo@studwc.kiu.ac.ug

ABSTRACT

Cancer genomics has improved cancer classification, risk assessment, and treatment selection, yet East African populations remain poorly represented in sequencing research and biomarker development. Studies across Kenya, Uganda, Tanzania, Rwanda, Ethiopia, and regional networks have focused mainly on breast, oesophageal, prostate, cervical, HPV-related cancers, and endemic Burkitt lymphoma. Researchers have used genotyping, genome-wide studies, exome and genome sequencing, RNA sequencing, viral analysis, immunohistochemistry, and circulating tumour DNA. Although these studies show feasibility and reveal possible ancestry-, infection-, and exposure-related differences, most remain exploratory because cohorts are small, clinically diverse, and concentrated in tertiary centres. Regional validation, tumour–normal pairing, and long-term outcome data are limited, while local pathology, biobanking, sequencing, bioinformatics, and interpretation capacity remain uneven. Progress requires representative multicountry cohorts, standardised methods, African-governed data resources, locally validated assays, ethical oversight, strong pathology services, treatment access, and implementation research to ensure genomic discoveries improve patient outcomes across diverse East African settings.

Keywords: cancer genomics; biomarkers; East Africa; sequencing; precision oncology; liquid biopsy; molecular diagnostics; African ancestry

INTRODUCTION

Cancer biomarkers include measurable DNA, RNA, protein, epigenetic, viral or cellular features that inform risk, detection, diagnosis, prognosis, treatment selection or monitoring. Some are inherited germline variants; others are acquired somatic alterations in a tumour [1]. A molecular difference is not automatically a useful biomarker. Clinical translation requires a defined intended use, reproducible assay, validated threshold, evidence of added value over existing care and a feasible clinical action [2]. East Africa has a distinctive research case. Its populations contain extensive genetic diversity and fine-scale structure that are poorly represented by broad labels such as "Black" or "African ancestry." Cancer patterns are also shaped by infection human papillomavirus (HPV), Epstein–Barr virus (EBV), hepatitis viruses, HIV and malaria-related immune effects alongside environmental exposures, diet, tobacco, alcohol and health-system delay [3, 4]. Biomarkers derived largely from European-ancestry cohorts may therefore transfer imperfectly. This review considers the East African Community and adjacent countries that participate in regional cancer research, with most available evidence arising from Kenya, Uganda and Tanzania. It distinguishes discovery research from clinically validated tools and asks three questions: what genomic work exists, where the sequencing and evidence gaps lie, and which diagnostic-development pathway is most likely to yield equitable value.

Why Population Representation Matters

African genomes contain greater nucleotide diversity and shorter linkage-disequilibrium blocks than non-African genomes. This can improve fine-mapping of causal variants, but it also means that arrays, imputation panels and polygenic scores trained elsewhere may perform poorly. Rare benign variants may be labelled "variants of

uncertain significance" when ancestry-matched controls are absent [5]. In tumours, ancestry can correlate with somatic drivers, copy-number changes, mutational signatures, immune context and outcome, although social exposures and care access must not be mistaken for genetic effects. The underrepresentation is extreme at the tumour-genome level. A 2025 review identified only five whole-cancer-genome databases containing patients from sub-Saharan Africa, spanning breast, oesophageal, prostate and Burkitt lymphoma [6]. A systematic review of next-generation breast-cancer sequencing found only ten eligible studies across seven African countries through September 2024 [7]. East African cohorts form only a fraction of these already small resources.

Table 1: Current Landscape by Cancer Type

Cancer	East African activity	Candidate biomarker value	Main limitation
Breast	Kenyan tumour WES/RNA-seq; molecular pathology and tumour-microenvironment cohorts	Somatic drivers, expression classes, neoantigens, germline risk	Small cohorts; subtype and treatment heterogeneity; little prospective validation
Oesophageal	Consortia in Kenya and Tanzania; aetiologic, germline and tumour studies	Mutational signatures, drivers, exposure-linked markers	Few paired genomes and outcome-linked cohorts
Prostate	Ugandan participation in African GWAS and genomic consortia	Ancestry-specific susceptibility and molecular taxonomy	East African numbers remain modest; PRS transferability
Cervical/HPV	Kenya–Uganda HPV/HIV consortium; genotype and persistence studies	HPV genotype, viral load/integration and methylation candidates	Many epidemiologic studies, fewer tumour multi-omics and triage validations
Burkitt lymphoma	Northern Uganda germline studies; Uganda–Tanzania ctDNA validation	EBV/tumour sequences, rapid diagnosis and monitoring	Cost, low DNA fractions, infrastructure and clinical-utility evidence

Cancer genomics in East Africa is advancing, but evidence remains uneven and clinical translation requires caution. In breast cancer, Kenyan studies using whole-exome and RNA sequencing have shown that local tumour profiling is feasible and may reveal mutations, expression patterns and neoantigens, yet small samples mean these findings are exploratory rather than ready for diagnosis or vaccines [8]. Strengthening reliable oestrogen receptor, progesterone receptor and HER2 testing should remain a priority before genomic panels are introduced. For oesophageal squamous-cell carcinoma, regional collaborations are investigating recurrent drivers, mutational signatures, inherited susceptibility and exposure-related molecular patterns, but larger standardised tumour–normal cohorts with treatment and survival data are still needed. Prostate-cancer consortia across Africa have identified ancestry-specific risk architectures and molecular subtypes, although East African participation is too limited for country-specific prediction, so polygenic scores must be locally calibrated and compared with simpler clinical tools [9]. Cervical cancer offers the clearest near-term pathway because high-risk HPV testing is causal, actionable and suitable for self-sampling; genotype data, methylation, viral integration and other biomarkers may improve triage after prospective validation. Endemic Burkitt lymphoma research in Uganda and Tanzania is assessing host genetics and circulating tumour DNA, which could speed diagnosis and reduce invasive procedures, but limited samples, costs, supply chains and uncertain clinical benefit remain major barriers to implementation across diverse regional health systems and patient populations [10].

Molecular Platforms and Biomarker Classes

Molecular platforms offer complementary ways to identify and validate cancer biomarkers. Germline studies, including GWAS, rare-variant analyses, and family-based research, can reveal inherited susceptibility, but they require ancestry-matched controls and careful consent because findings may affect relatives. Somatic DNA testing through targeted panels, whole-exome sequencing, or whole-genome sequencing can detect tumour drivers and mutational signatures, although paired normal tissue is needed to separate inherited from acquired variants [11]. Transcriptomic and epigenomic approaches, such as RNA expression and methylation profiling, can classify tumours and support diagnostic triage, but results are highly sensitive to sample handling and mixed cell populations. Viral genomics, particularly HPV and EBV genotyping, integration, and viral-load assessment, is especially relevant in regions with infection-associated cancers [12]. Liquid biopsy using circulating tumour DNA may assist diagnosis and monitor treatment response or recurrence, but performance varies by tumour type, stage, assay depth, and pre-analytical quality. Proteomics and immunophenotyping may provide lower-cost clinical tests after discovery, while multi-omics can uncover broader pathways, though small datasets risk overfitting and

require independent validation. Across all platforms, robust pathology, standardized workflows, and reproducible validation remain essential [13].

The Sequencing-Data Gap

The sequencing-data gap in East African cancer genomics reflects weaknesses across cohort design, reference resources, specimen quality, and local infrastructure. Many studies rely on small, highly selected referral-centre cohorts that overrepresent advanced disease, exclude patients unable to access biopsy or surgery, and lack complete staging, treatment, and follow-up data. These limitations inflate effect estimates, exaggerate rare variants, and prevent meaningful analysis by ancestry, sex, HIV status, or tumour subtype [14]. Interpretation is further hindered by limited East African representation in population databases, insufficient matched normal tissue, and dependence on European or TCGA controls that may introduce ancestry, environmental, clinical, and healthcare confounding. Reliable findings therefore require local normal genomes, fine-scale population controls, and regionally governed data aggregation. Sample integrity is equally critical because poor fixation, cold-chain failures, low tumour cellularity, necrosis, and degraded nucleic acids can undermine sequencing, making pathology review, standard procedures, annotation, and biobanking essential [15]. Finally, limited sequencing platforms, reagent shortages, maintenance, power, connectivity, storage, computing, and specialist expertise often force sample export, delaying analysis, restricting training, and shifting control of data, authorship, and infrastructure to external partners, and reinforcing dependency created by short grants and externally controlled systems [16].

From Discovery to a Locally Relevant Diagnostic

Developing a locally relevant diagnostic requires moving systematically from discovery to implementation. First, define the intended use, since risk prediction, early detection, differential diagnosis, prognosis, treatment selection, and monitoring demand different study designs and decision thresholds. Biomarkers should be discovered in representative, adequately powered, multicountry cohorts using paired tumour–normal samples and harmonised clinical data. They must then be replicated independently in East African populations and evaluated across ancestry, HIV status, sex, and tumour site [17]. Complex signatures should be converted into robust, affordable assays, such as PCR, targeted sequencing, methylation testing, or immunohistochemistry, that local laboratories can operate reliably. Analytical validity requires evidence on accuracy, precision, detection limits, failure rates, specimen stability, contamination, and inter-laboratory reproducibility. Clinical validity should be demonstrated through pre-specified sensitivity, specificity, discrimination, calibration, and added value over existing methods. Clinical utility and cost must also be assessed to confirm that testing improves decisions and outcomes without worsening inequity [18]. Finally, implementation needs regulation, procurement, quality assurance, training, interpretation, treatment access, and ongoing monitoring. The final clinical product may differ greatly from the discovery platform, and a biomarker is useful only when linked care is accessible, affordable, and safe.

Ethical, Legal and Social Considerations

Genomic research raises questions about broad consent, future use, family implications, re-identification, export, benefit sharing and return of results. These concerns are amplified when samples originate in East Africa but data, patents and commercial value accrue elsewhere. Governance should specify custodianship, access committees, publication expectations, community engagement, controlled data sharing and routes for returning clinically actionable findings [19]. Consent materials must distinguish research from guaranteed clinical benefit. Indigenous and minority communities require particular safeguards against group stigma and exploitative ancestry claims. Equity also concerns who receives the eventual test. Urban private laboratories may adopt panels long before public systems. Research should include rural referral pathways, turnaround time, sample transport and household cost from the outset. A marker that is analytically elegant but accessible only to a small affluent group may widen cancer disparities [20].

Strategic Priorities for East Africa

East Africa should establish a multicountry cancer-genome initiative initially targeting breast, cervical, oesophageal, prostate, paediatric, and infection-related cancers. The programme should build representative tumour–normal cohorts linked to standardised pathology, staging, treatment, survival, and exposure data, with deliberate inclusion of rural and underserved populations. A meaningful proportion of samples should be sequenced locally, supported by long-term investment in equipment maintenance, reagents, computing, and data stewardship. Strong pathology, biobanking, and pre-analytic systems must underpin genomics, using shared quality standards and external proficiency testing [21]. The initiative should develop ancestry-aware germline reference resources that reflect fine-scale population structure rather than treating African ancestry as uniform. Priority applications include prospective validation of HPV triage markers, childhood lymphoma ctDNA, and limited actionable tumour panels. Assays should be designed around public-sector realities, including specimen type, turnaround time, workforce, equipment, quality assurance, and treatment access. Regional bioinformatics training, career pathways, molecular tumour boards, and secure shared computing should replace short-term workshops [22]. Governance must be African-led, with equitable authorship, transparent material-transfer agreements, benefit sharing, independent replication, health-economic assessment, and proof of clinical utility

before biomarkers are promoted as East African specific. This approach improves relevance, sustainability, equity, and impact.

Evidence Gaps and Research Cautions

Country coverage is highly uneven; Rwanda, Burundi, South Sudan and many parts of Ethiopia and Tanzania contribute little tumour-sequencing data. Paediatric cancers other than Burkitt lymphoma, colorectal cancer, liver cancer, haematological malignancies and rare tumours are poorly characterised. Germline and somatic findings are often mixed conceptually, and few studies address tumour evolution, recurrence or treatment response [23]. Multi-omics studies face severe overfitting when the number of features greatly exceeds the number of patients. Apparent ancestry differences may reflect stage, exposure, pathology or treatment. Finally, "population-specific" should not imply biological isolation: the goal is representative evidence and calibrated tools, not racial essentialism.

CONCLUSION

Cancer-genomic research in East Africa has moved beyond proof of concept, but it has not yet produced a broad portfolio of regionally validated clinical biomarkers. Strongest opportunities lie where burden, biological rationale and feasible action align: HPV-based cervical triage, molecular refinement of common breast and oesophageal cancers, ancestry-aware prostate risk research, and ctDNA diagnosis for endemic childhood lymphoma. Progress now depends on larger tumour–normal cohorts, complete clinical annotation, regional replication, high-quality specimens, African-controlled data infrastructure and affordable assay conversion. Locally relevant diagnostics will not emerge from sequencing alone; they require a continuous pathway from representative discovery to analytical validity, clinical utility, treatment access and equitable implementation.

REFERENCES

1. Uti, D. E., Omang, W. A., Idowu, A. O., Egbung, G. E. (2026). Theranostic Nanoparticles in Prostate Cancer: Disrupting Hypoxia-Induced Glycolysis by Targeting Hypoxia-Inducible Factor-1 Alpha and Downstream Metabolites. Cancer Medicine, 15(1). https://doi.org/10.1002/cam4.71519
2. Li, H., Meng, L., Wang, H., Cui, L., Sheng, H., Zhao, P., et al: Precise identification of somatic and germline variants in the absence of matched normal samples. Brief. Bioinform. 26, bbae677 (2024). https://doi.org/10.1093/bib/bbae677
3. Omang, W.A., Bawa, I., Ugwu, O.P.C., Egba, S.I. (2026). Next-generation strategies against prostate cancer: natural products and nanomedicine targeting prostate cancer stem cells. Journal of Men S Health, 22(1), 27. https://doi.org/10.22514/jomh.2026.003
4. Mozzi, A., Pontremoli, C., Sironi, M.: Genetic susceptibility to infectious diseases: Current status and future perspectives from genome-wide approaches. Infect. Genet. Evol. 66, 286–307 (2018). https://doi.org/10.1016/j.meegid.2017.09.028
5. Uti, D. E., Alum, E. U., Ugwu, O. P. (2026). Nanovaccines in Cancer Immunoprevention and Immunotherapy: Current Insights and Future Directions. Cancer Treatment and Research Communications, 101166. https://doi.org/10.1016/j.ctarc.2026.101166
6. Fergus, K.B., Newberg, J., Greenstein, R., Fejerman, L., Carvajal-Carmona, L., Collisson, E., et al Association between ancestry and tumor somatic mutations in a large national cohort of women with breast cancer. NPJ Breast Cancer. 11, 117 (2025). https://doi.org/10.1038/s41523-025-00815-x
7. Egwu CK, Manjula VS, et al. Overcoming the Black Box Challenge: Building Trust in Artificial Intelligence Algorithms in Oncology. *Technology in Cancer Research & Treatment.* 2026;25. doi:10.1177/15330338261434649
8. Wagutu, G., Gitau, J., Mwangi, K., Murithi, M., Melly, E., Harris, A.R., Sayed, S., Ambs, S., Makokha, F.: Whole exome-seq and RNA-seq data reveal unique neoantigen profiles in Kenyan breast cancer patients. Front. Oncol. 14, 1444327 (2024). https://doi.org/10.3389/fonc.2024.1444327
9. Akwari, A.A., Okoroh, P.N., Aniokete, U.C., Abba, J.N., Uti, D.E. Phytochemicals as modulators of ferroptosis: a novel therapeutic avenue in cancer and neurodegeneration. *Mol Biol Rep* **52**, 636 (2025). https://doi.org/10.1007/s11033-025-10752-4
10. Burdier, F.R., Waheed, D.-N., Nedjai, B., Steenbergen, R.D.M., Poljak, M., Baay, M., Vorsters, A., Van Keer, S.: DNA methylation as a triage tool for cervical cancer screening – A meeting report. Prev. Med. Rep. 41, 102678 (2024). https://doi.org/10.1016/j.pmedr.2024.102678
11. Nwuruku, O.A., Edwin, N.: Harnessing nature: plant-derived nanocarriers for targeted drug delivery in cancer therapy. Phytomedicine Plus. 5, 100828 (2025). https://doi.org/10.1016/j.phyplu.2025.100828
12. De Paolis, E., Perrucci, A., Albertini Petroni, G., Conca, A., Corsi, M., Urbani, A., Minucci, A.: Genomics and Epigenomics Approaches for the Quantification of Circulating Tumor DNA in Liquid Biopsy: Relevance of a Multimodal Strategy. Int. J. Mol. Sci. 26, 10982 (2025). https://doi.org/10.3390/ijms262210982

13. Nwuruku, A.O. Alum, E.U., Edwin, N. Targeting Oxidative Stress in Cancer Management: The Role of Antioxidant Phytochemicals. KIU J. Health Sci., 4(2): 1-10. https://doi.org/10.59568/KJHS-2024-4-2-01
14. Corpas, M., Pius, M., Poburennaya, M., Guio, H., Dwek, M., Nagaraj, S., Lopez-Correa, C., Popejoy, A., Fatumo, S.: Bridging genomics' greatest challenge: The diversity gap. Cell Genomics. 5, 100724 (2025). https://doi.org/10.1016/j.xgen.2024.100724
15. Alum, E.U. AI-driven biomarker discovery: enhancing precision in cancer diagnosis and prognosis. *Discov Onc* **16**, 313 (2025). https://doi.org/10.1007/s12672-025-02064-7
16. Uti, D.E., Atangwho, I.J., Alum, E.U. *et al.* Antioxidants in cancer therapy mitigating lipid peroxidation without compromising treatment through nanotechnology. *Discover Nano* **20**, 70 (2025). https://doi.org/10.1186/s11671-025-04248-0
17. Lafon, T., Weingart, M., Vaidie, J., Calfee, C.S., Jacob, S.T., Freund, Y., et al: Challenges in early detection and prognostication of sepsis: new approaches from the emergency department and intensive care unit. eClinicalMedicine. 94, 103864 (2026). https://doi.org/10.1016/j.eclinm.2026.103864
18. Edeh, O.F., Ainbeyoona, C., Unveiling the microbial orchestra: exploring the role of microbiota in cancer development and treatment. *Discov Onc* **16**, 646 (2025). https://doi.org/10.1007/s12672-025-02352-2
19. Li, Y., Bowen, D.J., Yu, J.-H.: Informed consent in genetic and genomic studies in Sub-Saharan Africa: a systematic review of bioethical issues. BMC Med. Ethics. 26, 99 (2025). https://doi.org/10.1186/s12910-025-01170-z
20. Fegade, B., Chaudhari, S.Y., Likhar, R.V. *et al.* Design, synthesis, molecular docking and molecular dynamics studies of some 3-methoxy flavone derivatives as an anti-breast cancer agent. *Discov Onc* **16**, 773 (2025). https://doi.org/10.1007/s12672-025-02491-6.
21. Magadani, R., Ndinteh, D.T., Roux, S., Nangah, L.P., Atangwho, I.J., et al. Cytotoxic Effects of *Lecaniodiscus Cupanioides* (Planch.) Extract and Triterpenoids-derived Gold Nanoparticles On MCF-7 Breast Cancer Cell Lines. Anticancer Agents Med Chem. 2025 Jan 27. doi: 10.2174/0118715206325529241004064307. Epub ahead of print. PMID: 39871566.
22. Armin R, Sami G T, Mark W C, Benjamin M G (2012). *Fusobacterium nucleatum* infection mimicking metastatic cancer. *Indian Journal of Gastroenterology, 31, 198-200.*
23. Manar M, Elnaem M. H, Ejaz C., Ismail I., Jamalludin Ab R, Ahlam N K, How S H (2021). Cancer-associated thrombosis: a clinical scoping review of the risk assessment models across solid tumours and haematological malignancies. International Journal of General Medicine, 3881-3897.

CITE AS: Muhindo Anitah. (2026). Genomic Studies and Biomarker Discovery for Cancer in East Africa Current Evidence, Sequencing Gaps and Priorities for Locally Relevant Diagnostics. IDOSR JOURNAL OF SCIENCE AND TECHNOLOGY 12(2):113-117, 2026.
https://doi.org/10.59298/IDOSR/JST/26/122.113117