Esophageal Cancer Genomics in Africa: Recommendations for Future Research

Hannah Simba1,2*, Gerard Tromp3,4,5,6,7, Vikash Sewram1, Christopher G Mathew6,9, Wenlong C. Chen8,10 and Helena Kuivaniemi3,4,5

1African Cancer Institute, Department of Global Health, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa, 2Environment and Lifestyle Epidemiology Branch, International Agency for Research on Cancer (IARC/WHO), Lyon, France, 3Division of Molecular Biology and Human Genetics, Department of Biomedical Sciences, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa, 4DSI-NRF Centre of Excellence for Biomedical Tuberculosis Research, Stellenbosch University, Cape Town, South Africa, 5South African Medical Research Council Centre for Tuberculosis Research, Stellenbosch University, Cape Town, South Africa, 6Bioinformatics Unit, South African Tuberculosis Bioinformatics Initiative, Stellenbosch University, Cape Town, South Africa, 7Centre for Bioinformatics and Computational Biology, Stellenbosch University, Stellenbosch, South Africa, 8Sydney Brenner Institute for Molecular Bioscience, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa, 9Division of Human Genetics, National Health Laboratory Service and School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa, 10National Cancer Registry, National Health Laboratory Service, Johannesburg, South Africa

Keywords: esophageal cancer, esophageal squamous cell carcinoma, genomics, genomic medicine, African esophageal cancer corridor

INTRODUCTION

Esophageal cancer (EC) is an aggressive malignancy and a major health burden documented as the sixth most common cause of cancer mortality worldwide (Bray et al., 2018). Over 80% of EC cases and deaths are reported in developing countries, where the esophageal squamous cell carcinoma (ESCC) subtype is more common, compared to the adenocarcinoma (EAC) subtype (Bray et al., 2018). EC has a peculiar geographical distribution, with high incidence rates reported in Asia and the African ESCC corridor (A bene et al., 2017; McCormack et al., 2017; Bray et al., 2018). Malawi has the highest ESCC incidence rate globally for both men and women, followed by Kenya and Zimbabwe in Africa (Bray et al., 2018). In South Africa, ESCC is the 10th most common cancer for men and the 11th most common cancer for women (National Cancer Registry, 2019), and has the 10th highest incidence of ESCC in Africa (Bray et al., 2018). Incidence rates are, however, disproportionately higher in the Eastern Cape Province, where it is the most common cancer for men, and the second most common cancer for women (Sondyala et al., 2015). Due to a lack of ESCC early detection markers, late diagnosis and poor prognosis are the norm. Additionally, distinct ESCC molecular subtypes have not been identified, which could provide opportunities for targeted and novel therapies.

The African ESCC corridor (Figure 1), which spans from the eastern to the southern part of Africa is characterized by high incidence rates, young age at presentation, delayed presentation, as well as poor outcomes and survival (Van Loon et al., 2018; Asombang et al., 2019). Risk factors associated with ESCC in high-risk areas include tobacco smoking, alcohol consumption, polycyclic aromatic hydrocarbon exposure, poor diet, hot beverages, poor oral hygiene, microbiome, and genetic factors (Asombang et al., 2019; Chetwood et al., 2019; Simba et al., 2019; Simba, 2021).

Limited Number of Genomic Studies on ESCC in Africa

There is an apparent lack of genomic studies on ESCC on African populations, therefore the genetic etiology is poorly understood and implementation of genomic medicine for ESCC remains elusive. The striking geographical distribution of ESCC suggests that ESCC etiology is multifactorial, with
shared and locally relevant environmental and genetic risk factors. There are several mutational signatures in ESCC that have been linked to environmental exposures, including tobacco smoking and alcohol consumption (Talukdar et al., 2013; Matejcic and Parker, 2015). It is, therefore, important to incorporate environmental and lifestyle risk factors when investigating genetic factors. Identification and quantification of these gene-environment (GxE) interactions provides a platform for targeted interventions and may explain the high incidence and mortality rates in Africa. In this opinion piece, we summarize the current state of ESCC genomic research in Africa, which include candidate gene studies, whole-exome studies and genome-wide association studies (GWAS). The first genome-wide methylation profiling of ESCC, with 327 tissues from nine African, Asian and South American populations, identified novel DNA methylation events in ESCC tumors (Talukdar et al., 2021). Importantly, there is a lack of multi-omics studies.

Candidate Gene Studies in African ESCC
We recently performed a systematic review of candidate gene studies carried out for ESCC in African populations (Simba et al., 2019). Studies were done on South African, Malawian and Kenyan populations only (Simba et al., 2019). Only one exome sequencing study from Malawi on ESCC tumors was found (Liu et al., 2016). In the germline, 25 SNPs were reported to be associated with ESCC in 20 different genes, and 22 genes had somatic mutations (Simba et al., 2019). These studies had several limitations: 1) small sample sizes leading to poor statistical power; 2) no cohort studies; and 3) no standardized methods of data collection, reporting, or analysis. In addition, the studies did not correct for population structure in cases and controls, which is particularly relevant in African populations known for high levels of genomic diversity (Choudhury et al., 2020).

Genome-Wide Association Studies in African ESCC are Needed
All GWAS for ESCC published to date were conducted in Asian or European populations (Cui et al., 2009; Wang et al., 2010; McKay et al., 2011; Wu et al., 2012; Wu et al., 2013; Wu et al., 2014; Abnet et al., 2017). These studies identified several associated loci, including PLC1, RUNX1 and CHEK2. However, the transferability and replication of these loci in African populations were limited, even in a well-powered study (Chen et al., 2019). Thus there is a need for ESCC GWAS studies in populations from high-risk regions of sub-Saharan Africa. GWAS are currently in progress in the ERICA-SA study (https://www.samrc.ac.za/intramural-research-units/evolving-risk-factors-cancers-african-populations-ERICA-SA), Johannesburg Cancer study (Chen et al., 2020), and the African Esophageal Cancer Consortium (AfrECC) (Van Loon et al.,...
These studies are likely to provide a much clearer understanding of the genetic etiology of African ESCC. GWAS data can also be used to develop polygenic risk scores for stratifying disease risk (Fritsche et al., 2021).

**Somatic Mutations in Tumor Tissue From African ESCC Patients**

Genomic analysis of tumor tissues by DNA and RNA sequencing in large-scale global projects such as the Cancer Genome Atlas (TCGA) (The Cancer Genome Atlas Research Network, 2017) has led to a huge increase in knowledge of the genes and somatic mutations which drive tumor development. This in turn has fueled major progress in the development of cancer therapies that target the molecular pathways important for tumor development. No samples from African ESCC cases were included in the TCGA project. Information on the genomic profiles of African cancers is, however, very limited. In ESCC, whole-exome and RNA sequencing of tumors from 59 Malawian patients (Liu et al., 2016) observed similar genetic aberrations to those reported in Asian and North American patients and included mutations of well-established driver genes such as TP53, CDKN2A, NFE2L2, CHEK2, NOTCH1, FAT1, and FBXW7. Analyses also detected signatures associated with aging and cytidine deaminase (APOBEC) activity but, surprisingly, not of tobacco smoke.

A recent study (Moody et al., 2021) investigated mutational signatures in 552 ESCC patients from high-incidence regions (Iran, Kenya, Tanzania, China, and Malawi) including three African countries, and low-incidence regions (Brazil, Japan and the UK) using whole-genome sequencing. Similar mutational profiles across all countries were found. Specific mutational signatures and ESCC risk factors were detected for tobacco, alcohol, opium and germline variants, and also highlighted APOBEC activation as an important step in tumor development. No evidence of an unknown exogenous mutagen associated with a mutational signature, which could explain ESCC variation in incidence, was found.

**DISCUSSION**

**Benefits of Genomic Data in ESCC Research**

Genomic studies provide critical information about the pathobiology of diseases, which improves our understanding of the risk and heritability of ESCC, risk prediction for populations and individuals, and contributes to cancer prevention (Figure 1). ESCC genomic research must follow rigorous guidelines to ensure reproducibility and reliability of results. Prioritization of ESCC genomic medicine in African populations will help elucidate the genetic etiology of ESCC, giving insights on variants, biological mechanisms, and the GxE interactions associated with ESCC development. This information will facilitate developing algorithms for predicting ESCC prognosis and survival of the patients.

Genomics can be used as a tool to address health disparities in cancer (Balogun and Olopade, 2021). There are differences in ESCC incidence and mortality between European and African populations. Whilst inadequate health care systems, lack of access, and poor quality of care have an obvious role, genomics can also be used to address these gaps in incidence and mortality. Genomic research can identify variants, pathways and biomarkers associated with increased risk and mortality, to be used in precision medicine (Figure 1). Ultimately, ESCC genomic research in African populations should not only contribute to understanding the etiology but also generate evidence that can be translated to prevention and therapeutics.

A recent study (Choudhury et al., 2020) explored human migration and the breadth of genomic diversity in 426 African individuals from 50 ethnolinguistic groups. Over three million previously undescribed single nucleotide variants were identified. The authors recommended “broader characterization of the genomic diversity of African individuals to understand human ancestry and improve health” (Choudhury et al., 2020). It remains to be determined if ancestral events, migration, the admixture of populations, and adaptation to exposures in the African ESCC corridor play a role in the demographic and geographic aspects of ESCC.

The Pan-Cancer Analysis of Whole Genomes Consortium performed the most comprehensive meta-analysis of cancer genomes to date, using 2,658 tumors and 38 tissues (Gieslak and Chinnaiyan, 2020). Of these, 98 were EAC tumors, but none were ESCC samples. The study assessed key aspects of cancer genomics, i.e., cancer drivers (Campbell et al., 2020), non-coding changes (Rheinbay et al., 2020), mutational signatures (Li et al., 2020), structural variants (Alexandrov et al., 2020), cancer evolution (Gerstung et al., 2020), and RNA alterations (Calabrese et al., 2020). However, none of the investigators were from Africa. Genomic analysis of substantial numbers of African ESCC patients together with detailed epidemiological data is needed to further explore the origins of African ESCC.

**Barriers to ESCC Genomics Research in Africa**

The lack of investment in genomic medicine in Africa has led to most genomic medicine knowledge being founded on genomes of European ancestry, despite African populations displaying higher levels of genetic diversity. A major barrier to implementing genomic research in Africa is inadequate infrastructure including poorly equipped facilities, erratic power supply, inadequate biotechnology and information technology infrastructure as well as the high cost of genomics tools and the costs associated with implementing genomic medicine (Adebamowo et al., 2018; Munung et al., 2018). These aspects further impact biospecimen collection, transportation and storage, which is pivotal in planning and conducting genomic studies. In Africa, the ethical aspects surrounding genomics research are also challenging and obtaining informed consent is compounded by language barriers and low literacy levels.
Recommendations for Future ESCC Studies and Furthering Genomics Research in Africa

African genomes harbor the most genetic diversity and variation, and yet are the least genetically characterized. This means that genetic variants of medical relevance remain unknown (Choudhury et al., 2020). For ESCC this impedes progress on applying genomics in understanding etiology, tailored screening and therapeutic interventions, promoting health equity and ultimately reducing the burden of ESCC in the African ESCC corridor. To fully understand the etiology of ESCC and provide tools for genomic medicine, ESCC research in Africa should: 1) follow a multidisciplinary approach to study interactions between genomic, environmental and lifestyle factors; 2) foster collaborations and data sharing to accelerate progress; 3) use standardized methods for analysis and reporting; 4) use larger study samples to adequately detect GxE interactions; and 5) control for population stratification in admixed populations. Additionally, openness and precision in reporting of methods are needed to improve reproducibility. In a cancer biology study aimed at replicating 193 experiments from 53 high-impact papers, none of the published studies had sufficient information for repeating the experiments (Errington et al., 2021). For genetic association studies, the STRectifying the REPoting of Genetic Association studies (STREGA) statement (Little et al., 2009) should be used as a checklist to assess quality of reporting and methods.

Furthering genomic medicine in Africa requires leveraging existing infrastructure and learning from the extensive experience of current genomic medicine implementations in other countries (AESA, 2020). This requires significant infrastructure, including access to clinical facilities and high-throughput genotyping and sequencing facilities. Creating a unified ESCC genomics research hub in Africa will require standardized sample collection and well-managed biorepositories with the capacity to store and manage biospecimens. Strong information technology infrastructure is needed that is capable of managing, storing and analyzing big data (Wonkam and Mayosi, 2014). In addition, capacity building is needed to create a critical mass of bioinformaticians, and provide genomic medicine training programs for healthcare professionals.

CONCLUSION

Genomics is an invaluable approach in providing unbiased information about the pathogenesis of ESCC. The information could be used to predict risk, screen asymptomatic individuals, diagnose more accurately and develop targeted treatments. We are still far from being able to implement genomic medicine for ESCC in Africa since genomic information on African ESCC patients is very limited.

AUTHOR CONTRIBUTIONS

HS, HK, and GT conceptualized the idea for the “Opinion” manuscript. HS, HK, GT, VS, CGM, and WCC wrote sections of the manuscript. All authors reviewed and approved the final manuscript.

FUNDING

HS acknowledges the Beit Trust fund for providing a Doctoral Scholarship (https://beittrust.org.uk/beit-trust-scholarships), the Collaboration for Evidence-based Healthcare and Public Health in Africa (CEBHA+; https://www.cebha-plus.org/), as part of the Research Networks for Health Innovation in Sub-Saharan Africa, Funding Initiative of the German Federal Ministry of Education and Research, Margaret McNamara Education Grant (https://www.mmeg.org/), and the L’Oreal UNESCO for Women in Science Award (https://en.unesco.org/news/unesco-and-foundation-loreal-recognize-20-young-women-scientists-sub-saharan-africa). HK was supported by the Faculty of Medicine and Health Sciences, Stellenbosch University, South Africa. VS was supported by the African Cancer Institute and National Research Foundation (UID:132783). GT was supported by the South African Tuberculosis Bioinformatics Initiative (SATBII), a Strategic Health Innovation Partnership grant from the South African Medical Research Council (https://www.samrc.ac.za/) and South African Department of Science and Innovation (https://www.dst.gov.za/). WCC and CGM were supported by the South African Medical Research Council (with funds received from the South African National Department of Health) and the UK Medical Research Council (with funds from the UK Government’s Newton Fund) (MRC-RFA-SHIP 01-2015).

REFERENCES

Abnet, C. C., Arnold, M., and Wei, W.-Q. (2018). Epidemiology of Esophageal Squamous Cell Carcinoma. Gastroenterology 154 (2), 360–373. doi:10.1053/j.gastro.2017.08.023
Ademshumo, S. N., Francis, V., Tambo, E., Diallo, S. H., Landouré, G., Nembeware, V., et al. (2018). Implementation of Genomics Research in Africa: Challenges and Recommendations. Glob. Health Action. 11 (1), 1419033. doi:10.1080/16549716.2017.1419033
AESA (2020). “A Framework for the Implementation of Genomic Medicine for Public Health in Africa,” in Alliance for Accelerating Excellence in Science in Africa (Nairobi).

Alexandrov, L. B., Kim, J., Kim, J., Haradhvala, N. J., Huang, M. N., Tian Ng, A. W., et al. (2020). The Repertoire of Mutational Signatures in Human Cancer. Nature 578 (7793), 94–101. doi:10.1038/s41586-020-1943-3
Asombang, A. W., Chishinga, N., Nkhoma, A., Chipaila, J., Nsokolo, B., Manda-Mapalo, M., et al. (2019). Systematic Review of Esophageal Cancer in Africa: Epidemiology, Risk Factors, Management and Outcomes. Wjg 25 (31), 4512–4533. doi:10.3748/wjg.v25.i31.4512
Balogun, O. D., and Olopade, O. I. (2021). Addressing Health Disparities in Cancer with Genomics. Nat. Rev. Genet. 22 (10), 621–622. doi:10.1038/s41576-021-00390-4
Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA: a Cancer J. Clinicians 68 (6), 394–424. doi:10.3322/caac.21492
