

Burden of rare genetic variants in genes associated with cancer among Malawian cervical cancer patients



Authors:

Samuel D. Gwayi¹ 
Tamiwe Tomoka² 
Emile R. Chimusa³ 
George Fedoriw² 
Benjamin Kumwenda^{1,4} 

Affiliations:

¹Department of Biomedical Sciences, School of Life Science and Allied Health Professions, Kamuzu University of Health Sciences, Blantyre, Malawi

²Department of Medicine, University of North Carolina, Lilongwe Project, Lilongwe, Malawi

³Department of Human Genetics and Forensic Genetics, Northumbria University, Newcastle, United Kingdom

⁴Department of Food Biotechnology, University of Johannesburg, Johannesburg, South Africa

Corresponding author:

Samuel Gwayi,
sgwayi@kuhes.ac.mw

Dates:

Received: 21 Oct. 2025

Accepted: 19 Dec. 2025

Published: 15 May 2026

How to cite this article:

Gwayi SD, Tomoka T, Chimusa ER, Fedoriw G, Kumwenda B. Burden of rare genetic variants in genes associated with cancer among Malawian cervical cancer patients. *Afr J Lab Med*. 2026;15(1), a3073. <https://doi.org/10.4102/ajlm.v15i1.3073>

Read online:



Scan this QR code with your smart phone or mobile device to read online.

Background: Cervical cancer (CC) is one of the most common cancer types affecting women globally. Cervical cancer is largely associated with human papillomavirus infections; however, approximately 5% to 11% of CC cases are non-human papillomavirus virus-related. Malawi has the second highest CC prevalence and mortality rate worldwide.

Objective: This study investigated the burden of rare genetic variants in genes associated with CC among Malawian women.

Methods: Ethical approval was obtained from the National Health Science Research committee on 28 August 2023. Whole-genome sequencing was performed on 20 Malawian CC patients, followed by variant discovery and annotation using the Genome Analysis Toolkit and Ensembl's Variant Effector Predictor. Test for Rare Variants Against Public Database was performed on qualifying variants using 76156 genomes from the Genome Aggregation database as controls. Bonferroni correction was applied to account for multiple testing.

Results: We identified 372 genes with a significant burden of rare variants ($p < 0.05$), including *FAT1* ($p = 0.0003$), a known tumour suppressor gene associated with CC. After Bonferroni correction ($p < 6.7 \times 10^{-05}$), significance shifted to *HTR3B*, *EPHA8*, *ABCC2*, and *SEC23B* genes linked to various cancer types.

Conclusion: A significant burden of rare variants associated with CC was observed in known genes associated with lung, ovarian and osteosarcoma cancers, suggesting an increased population risk of developing CC and other cancers among Malawian women that needs further investigation.

What this study adds: The high burden of rare variants in the *FAT1* gene, and a significant rare-variant burden in other cancer associated genes, supports a broader CC genetic risk and possible novel roles requiring further exploration.

Keywords: cervical cancer; human papillomavirus; whole genome sequencing; qualifying variants; rare variants; Malawian women.

Introduction

Cervical cancer (CC) affects cells of the lower part of the uterus, called the cervix.¹ The cancer occurs because of the accelerated growth of cervical cells and resistance to programmed cell death, resulting in the formation and accumulation of abnormal cells.² These are cancerous cells that invade neighbouring tissues and spread through metastasis.

Cervical cancer is a global problem, as it is one of the leading causes of death among women globally.³ It is estimated that 500 000 cases and over 300 000 deaths occurred globally because of CC in 2018, with an incidence rate of 1.31 per 10 000 women.⁴ It is the second most common cause of death among women in developing countries.⁵ Cervical cancer (CC) patients experience local disease progression, causing fistulas, pain, urethral obstruction, and significant suffering, culminating in death.⁶ There are geographical variations in the distribution of CC globally, with the western Asia and Australia-New Zealand regions having as little as 1.0 case per 10 000 women. The African continent was found to have a very high CC burden, with an average age-standardised incidence rate (ASIR) of between 1.5 and 4.2 cases per 10 000 women, with the exception of North Africa. This was also observed in Eastern Europe, Southeastern Asia, Southern America, and the

Copyright: © 2026. The Authors. Licensee: AOSIS. This work is licensed under the Creative Commons Attribution 4.0 International (CC BY 4.0) license (<https://creativecommons.org/licenses/by/4.0/>).

Caribbean, which also had a world-age standardised rate of about 1.5 cases per 10000 women.⁷

Within Africa, sub-Saharan countries are severely affected by CC, as indicated by the average ASIR.⁸ The region has an overall average ASIR of 3.49 per 10000 women. However, the incidence rate varies greatly within this region, with reported rates of 4.34 per 10000 women in Southern Africa, 4.01 per 10000 in Eastern Africa, 2.96 per 10000 in Western Africa, and 2.68 per 10000 in Central Africa.⁹

Malawi, a sub-Saharan country, has the second highest CC mortality rate and ASIR, after Eswatini, which has the highest global CC mortality rate. Globally, the mortality rate and ASIR are 7.3 per 100 000 per year and 13.3 per 100000 per year. Malawi has a CC mortality rate and ASIR of 51.5 per 100000 per year and 67.9 per 100000 per year. The high prevalence of HIV and human papillomavirus infections have been attributed to the high incidence of CC in the country.¹⁰ Thus, there have been efforts by the Malawi government through the Ministry of Health (MoH) to combat CC in the country. Programmes promoting CC screening as well as vaccination against human papillomavirus as part of disease control have been implemented in various health care centres.¹¹ However, despite these efforts, the incidence of CC continues to increase, and the treatment of invasive cancer is associated with poor prognosis. While high-risk HPVs are largely associated with CC, host genetic susceptibility plays another major role in disease development.¹² Many genetic-based studies conducted in different populations to elucidate the genetic markers associated with CC have identified single nucleotide polymorphisms and variants.^{13,14,15,16} These studies revealed that, in addition to human papillomavirus infections, genetic variants contribute to the disease. Genetic variants are broadly classified as common or rare based on their minor allele frequency (MAF) in a given population. Rare genetic variants, defined as variants with a MAF of at most 10% ($MAF \leq 0.1$), which are usually not detected during genome-wide association studies, and mutations in genes such as *BRCA1*, *BRCA2*, *TP53*, *ATM*, *CHEK2*, *BRIP1*, and *PALB2*, have also been reported to play a significant role in cancer predisposition.^{17,18,19,20} The rare variants are evaluated via a burden association test, which assesses their collective impact on diseases that are complex in nature.²¹

Despite the high CC incidence and mortality rate in Malawi, rare genetic markers that may predispose women to CC have not been investigated. Thus, this study investigated the burden of rare genetic variants in genes associated with CC among Malawian women.

Methods

Ethical considerations

Prior to participation, consent was obtained from each participant via a written informed consent form. The participants were assigned study identification numbers,

and no participant name was used during sample processing and data analysis. Data analysis was performed on a secured bioinformatics server with access granted only to the main investigator. Prior to conducting the study, the study protocol was submitted to Malawi National Health Science Research Committee for ethical review. The National Health Science Research Committee approved the study on 28 August 2023, with the approval number 4048. Sample collection, processing, and whole-genome sequencing were carried out from 01 October 2023 to 19 July 2024.

Sample collection and sequencing

This study used a case-control association test design utilising whole-genome sequences from 20 Malawian women diagnosed with CC as cases, and 76156 genomes from the Genome Aggregation Database as controls.²² This was a pilot study, and given the limited sample size, we used a rare variant burden test, which is appropriate because it aggregates multiple rare variants within a gene or genomic region into a single test statistic rather than testing each variant individually, thereby enhancing the ability to detect collective effects that single-variant tests would be underpowered to find.²³ The patients' average age was 49.05 years, with a minimum age of 31 years and maximum age of 64 years. The clinical characteristics of the study cohort included HIV status, with four out of 20 participants (20%) testing HIV positive and the remaining 16 (80%) being HIV negative at the time of diagnosis. All participants were newly diagnosed with CC, and aside from HIV infection, there were no documented comorbidities in the clinical records. The study site was the Malawi National Cancer Center, in Lilongwe, Malawi. Participants were enrolled in the study on the basis of their confirmed diagnosis of CC.

After obtaining consent to conduct the study, we collected whole blood samples from the patients between 01 October and 30 November 2023. DNA extraction was performed using the Qiagen Max Kit (QIAGEN GmbH, QIAGEN Str. 1, D-40724 Hilden, Germany).²⁴ The whole-genome sequencing with an average reads depth of 30X, was done at the Centre for Genomics Research (CGR) in the United Kingdom using the NovaSeq X Plus Illumina platform (Illumina, Inc., San Diego, California, USA) in June 2024.

Data management and bioinformatics analysis

Whole-genome sequence reads for the 20 samples were obtained from the Centre for Genomics Research server in FASTQ format and stored on a secure local server running Ubuntu 18.04.5 LTS (Canonical Ltd, London, UK) for analysis in July 2024. Data trimming was performed using cutadapt version 4.5Q (Marcel Martin, National Bioinformatics Infrastructure Sweden) with the option -O 3 to ensure that trimming of adapter sequences was done at the 3' end, matching 3 base pairs (bp) or more. Low-quality bases were removed using a minimum window quality score of 20. Reads shorter than 15 bp were also removed, followed by quality assessment using FastQC and MultiQC.^{25,26} All of the

sample sequences that passed the per sequence quality score, mini quality score and per base N content, and GC content, and also passed the duplication and adapter contamination level quality checks, were included for further analysis. Assembly and variant discovery were performed using Burrows–Wheeler Aligner and Genome Analysis Toolkit following best practice guidelines.²⁷ The reads were aligned to the Genome Reference Consortium Human Genome Building 38 patch 14 via the Burrows–Wheeler Aligner.²⁸ Variant calling was performed on recalibrated bases via Genome Analysis Toolkit Haplotype Caller, followed by joint genotype calling on the individual sample generating a genotype Variant Call Format file. The combined genotype Variant Call Format file contained raw variants for all 20 samples. Single nucleotide polymorphisms were separated from Insertion and Deletions, creating two separate files. These were then filtered using Genome Analysis Toolkit hard-filter before being combined again into one file containing an analysis-ready Variant Call Format file.

Variant annotation and filtering

A server-based Ensembl Variant Effector Predictor version 108 (EMBL-EBI [European Bioinformatics Institute]). was downloaded onto the server and the variants were annotated based on their genomic location and impact using Variant Effector Predictor.²⁹ Variants in coding regions were selected and reannotated for their predicted impact using Sorting Intolerant from Tolerant and Polymorphism Phenotyping function predictors, and their frequency in the 1000 Genomes Project database and Genome Aggregation Database. In this study, all variants with a MAF of less than or equal to 0.1 were classified as rare variants.³⁰

On the basis of this process, all deleterious, probably or possibly damaging variants with MAF of ≤ 0.1 in the Genome Aggregation Database and the 1000 Genomes Project database were classified as qualifying variants and selected for burden association testing.

Burden association test and gene annotation

Our analysis leveraged public control genome data rather than matched individual-level genotype data; as such, we employed Test for Rare variants Against Public Database for rare variant burden testing. This association test is based on comparing the burden of rare variants in genes between cases and public controls using summary statistics, which is not feasible with methods such as Sequence Kernel Association Test that require individual-level genotypes and covariate information.³¹ Thus, a single nucleotide polymorphism file was created based on the qualifying variants, and genotypes in cases and controls were obtained using the single nucleotide polymorphism file. A one-sided Fisher's exact test was performed using the default Test for Rare variants Against Public Database script called *Burden R*.³¹ A p -value threshold of less than 0.05 was initially used, followed by Bonferroni correction to account for multiple testing.³² This generated a corrected p -value of 6.7×10^{-05} . Genes with significant burden of rare variants were

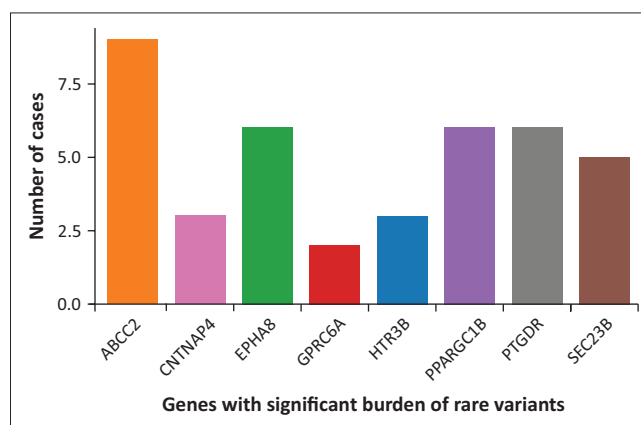
annotated via the Database for Annotation, Visualization and Integrated Discovery tool for disease association and pathway analysis.³³

Results

There were 17443881 variants called from the 20 Malawian CC cases using Genome Analysis Toolkit. Among these, 9206 variants were classified as high or moderate impact variants present in coding regions of the genomes. Following further filtering of the variants based on Sorting Intolerant from Tolerant, Polymorphism Phenotype v2 (SIFT is developed by researchers at the Fred Hutchinson Cancer Research Center (FHCRC) and the J. Craig Venter Institute while Polyphen-v2 is developed and maintained by the Sunyaev Lab at Brigham and Women's Hospital, Harvard Medical School) and MAF criteria, 1173 variants were classified as qualifying variants for the association test. These variants were detected in 740 genes among the CC patients. Following the burden association test, 372 genes had an initial significant burden of rare variants ($p \leq 0.05$). Bonferroni correction for multiple testing resulted in 51 genes with a significant burden of rare variants on the basis of the adjusted p -value ($p < 6.7 \times 10^{-05}$).

Among the genes with initial significance burden was the *FAT1* gene ($p = 0.0003$), whereas after applying the Bonferroni correction and annotation of genes with significant burden using the Database for Annotation, Visualization and Integrated Discovery tool, the significant burden of rare variants was observed in 35 genes associated with human diseases, including eight genes, *HTR3B*, *EPHA8*, *ABCC2* and *SEC23B*, which were associated with different cancer types on the basis of Database for Annotation, Visualization and Integrated Discovery annotation (Figure 1 and Table 1).

A significant burden of rare variants was also observed in 27 genes associated with various human conditions. Most of these genes were associated with cardiovascular disease, as well as chemodependency and metabolism related abnormalities, based on Database for Annotation, Visualization and Integrated Discovery annotation (Figure 2).



ABCC2 (ATP binding cassette subfamily C member 2); CNTNAP4 (contactin associated protein-like 4); EPHA8 (EPH receptor A8); GPRC6A (G protein-coupled receptor class C group 6 member A); HTR3B (5-hydroxytryptamine receptor 3B); PPARGC1B (PPARG coactivator 1 beta); PTGDR (prostaglandin D2 receptor); SEC23B (SEC23 homolog B).

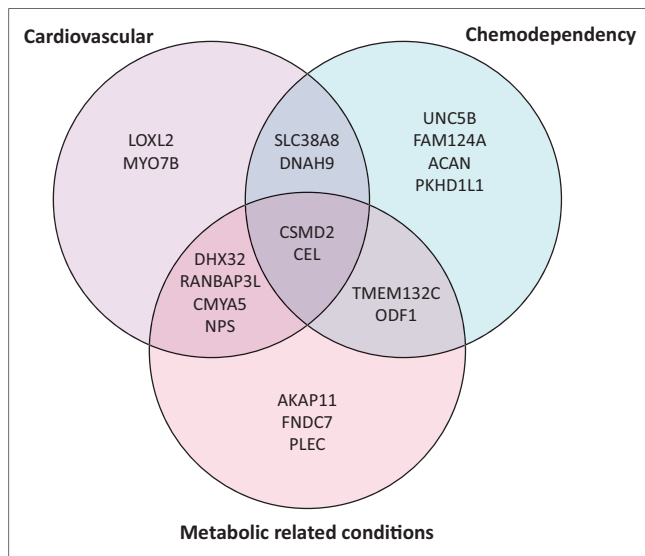
FIGURE 1: Known genes associated with cancer carrying a significant burden of rare variants among Malawian cervical cancer cases.

TABLE 1: List of genes with a high burden of rare variants associated with cancer.

Gene name	Gene description	Previously associated cancer type	TRAPD dominant model <i>p</i> -values	Association with cervical cancer among Malawian women
<i>ABCC2</i>	ATP binding cassette subfamily C member 2	Ovarian	3.52×10^{-07}	Significant
<i>CNTNAP4</i>	contactin associated protein family member 4	Osteosarcoma	2.21×10^{-05}	Significant
<i>EPHA8</i>	EPH receptor A8	Epithelial ovarian	2.42×10^{-05}	Significant
<i>FAT1</i>	FAT atypical cadherin 1	Cervical cancer	3.67×10^{-04}	Non-significant
<i>GPRC6A</i>	G protein-coupled receptor class C group 6 member A	Prostate	3.28×10^{-06}	Significant
<i>HTR3B</i>	5-hydroxytryptamine receptor 3B	Lung	3.28×10^{-06}	Significant
<i>PPARGC1B</i>	PPARG coactivator 1 beta	Pan-cancer	4.97×10^{-05}	Significant
<i>PTGDR</i>	prostaglandin D2 receptor	Prostanoid signalling in cancer	3.32×10^{-06}	Significant
<i>SEC23B</i>	SEC23 homolog B, COPII coat complex component	Noncanonical role in cancer	5.14×10^{-05}	Significant

TRAPD, Test for Rare variants Against Public Database; DAVID, Database for Annotation, Visualization and Integrated Discovery.

TRAPD dominant model *p*-values, associated conditions, cancer type based on DAVID annotation and observed association with cervical cancer are presented.



LOXL2 (lysyl oxidase like 2); MYO7B (myosin VIIb); DHX32 (DEAH-box helicase 32); RANBP3L (RAN binding protein 3 like); CMYA5 (cardiomyopathy associated 5); NPS (neuropeptide S); CSMD2 (CUB and Sushi multiple domains 2); CEL (carboxyl ester lipase); TMEM132C (transmembrane protein 132C); ODF1 (outer dense fiber of sperm tails 1); UNC5B (unc-5 netrin receptor B); FAM124A (family with sequence similarity 124 member A); ACAN (aggregran); PKHD1L1 (polycystic kidney and hepatic disease 1 like 1); AKAP11 (A-kinase anchoring protein 11); FNDC7 (fibronectin type III domain containing 7); PLEC (plectin); SLC38A8 (solute carrier family 38 member 8); DNAH9 (dynein axonemal heavy chain 9).

FIGURE 2: Genes with a significant burden of rare variants associated with major conditions in Malawian cervical cancer cases, including cardiovascular, metabolic and chemodependency-related conditions, as well as shared genes across these conditions, based on DAVID annotation.

Discussion

This study investigated the burden of rare genetic variants (MAF < 0.1 in the 1000 Genome Project and Genome Aggregation Database), in genes associated with CC among Malawian women. To our knowledge, this is the first study conducted among Malawian women to investigate CC genetic markers, and the results provide significant insights into the genetic nature of CC among Malawian women.

The findings of this work highlight the notable burden of rare variants with a significant *p*-value of 0.0003 in the tumour suppressor *FAT1* gene. The gene controls cell respiration and signals for cell–cell contact when the physiological conditions are normal within the body. Our findings suggest that the *FAT1* gene may have a potential role in CC causation among Malawian women, as mutations in the gene have been linked to cervical, colorectal and breast, as well as head and neck cancers.³⁴ Suppressive mutations in the *FAT1* gene are

associated with type changes in epithelial cells and cancer cell initiation.³⁵ However, in our study, the association between *FAT1* gene and CC was not statistically significant.

After applying the Bonferroni correction (adjusted *p*-value, 6.7×10^{-05}), we identified a significant burden of rare variants in several genes, including the *HTR3B*, *EPHA8* and *ABCC2* genes. These are known genes associated with different cancer types, such as ovarian, lung, and breast cancers. The *HTR3B* gene, for example, because of its role in neurotransmitter signalling, has been associated with worse clinical manifestations among patients during cancer therapy, including nausea and vomiting.³⁶ Alterations in genes have been described as risk signatures for breast cancer and are associated with shorter overall survival in patients with breast and lung cancers.^{37,38} The high burden of rare variants in this gene suggests that the identified gene plays a significant role in CC causation among Malawian women. On the other hand, the *EPHA8* gene is known to play an important role in the formation of new vessels and cancer metastasis during tumour formation. The gene has been classified as a cancer-causing agent because of its role in ovarian cancer causation and progression.³⁹ Furthermore, the *ABCC2* gene plays a critical role in lung cancer treatment outcomes, as it has been associated with poor response to chemotherapy and overall survival in patients with lung cancer.⁴⁰ This gene has been identified as a possible candidate for lung cancer therapy.⁴¹ Thus, our study suggests that both *EPHA8* and *ABCC2* genes also play an important role CC development among Malawian women.

Other cancer-related genes with a significant burden of rare variants identified in this study included *SEC23B*, *GPRC6A*, *PPARGC1B*, *CNTNAP4*, and *PTGDR* genes. While higher expression of the *SEC23B* gene has been associated with highly metastatic breast cancer, variations in the gene have been classified as having a yet to be agreed upon role in cancer progression.^{42,43} On the other hand, *GPRC6A* gene overexpression is associated mostly with male prostate cancer.^{44,45} Apart from these genes, a recent study among the Chinese population revealed that variations in the *PPARGC1B* gene are associated with an increased risk of gastric cancer in both sexes.⁴⁶ However, as early as 2006, inherited single nucleotide polymorphisms in the *PPARGC1B* gene were associated with breast cancers within families.⁴⁷ A correlation between variations in the gene and breast cancer was also

reported in 2011, when it was discovered that individuals with some polymorphisms in the *PPARGC1B* gene were at increased risk of breast cancer.⁴⁸ Furthermore, regarding variations in the *CNTNAP4* and *PTGDR* genes, studies have shown that mutations in *CNTNAP4* and *PTGDR* genes are linked to osteosarcoma and colorectal cancer.^{49,50} Therefore, while the four genes have not been previously associated with CC, our findings suggest their role in CC disease among Malawian women.

Given that cancer is classified as a complex disease, the discovery of 27 additional genes identified as having a significant burden of rare variants and associated with conditions not directly related to cancer emphasises the diversity of biological pathways involved and the complexity of the interactions among these genetic factors. These interactions may indicate a role of the identified genetic factors in CC oncogenesis among Malawian women, and require further investigation.

Limitations

In this study, we utilised public genomes as controls rather than locally collected healthy Malawian women controls. This could introduce bias resulting from differences in genetic background or population structure. Future studies should include matched local controls to confirm and extend our findings. We were also unable to apply individual-level rare variant tests such as Sequence Kernel Association Test because of smaller sample size and lack of matched control data. Sequence Kernel Association Test requires individual-level genotype data, which is not available when using summary-level external control datasets. This study is based on genetic association data without experimental or functional validation of the rare variants identified in genes. Thus, the biological effects and mechanisms of the rare variants remain unclear. Future work should involve laboratory-based (in vitro or in vivo) studies to confirm the roles of the identified rare variants.

Conclusion

This study revealed that while a high burden of rare variants was present in a known gene associated with CC (*FAT1*) among Malawian women, this was not statistically supported. However, the study revealed that *HTR3B*, *EPHA8*, *ABCC2*, *SEC23B*, *GPRC6A*, *PPARGC1B*, *CNTNAP4*, and *PTGDR* genes, previously associated with other cancer types, had a significantly high burden of rare variants in CC cases compared with controls among Malawian women. The findings suggest a potential role of the identified genes in CC causation and shared cancer susceptibility among Malawian CC patients. However, future studies should focus on the functional roles and clinical implications of these genes in CC disease.

Acknowledgements

We are grateful to the National Health Institute for their financial support through the Malawi Cancer Research Outcome Project, the Kamuzu University of Health Sciences

for their academic support during the project, and the staff members of the Malawi National Cancer Center, Kamuzu Central Hospital, the University of North Carolina Lilongwe project, Kamuzu University of Health Sciences Molecular Laboratory and University of North Carolina Chapel Hill for their support during the project.

This article is partially based on Samuel D. Gwayi thesis entitled 'Investigation of Genetic Variants Associated with Cervical Cancer among Malawian Women' towards the degree of Doctor of Philosophy in Biomedical Sciences, School of life Science and Allied Health Professions, Kamuzu University, Malawi, with supervisors Benjamin Kumwenda, Emile R Chimusa, George Fedoriw, Tamiwe Tomoka. Portions of the data, analysis, and discussion have been revised, updated, and adapted for publication as a journal article. The thesis is currently unpublished and not publicly available.

Competing interests

The authors reported that they received funding from National Institutes of Health, National Cancer Institute which may be affected by the research reported in the enclosed publication. The authors have disclosed those interests fully and have implemented an approved plan for managing any potential conflicts arising from their involvement. The terms of these funding arrangements have been reviewed and approved by the affiliated university in accordance with its policy on objectivity in research.

CRediT authorship contribution

Samuel D Gwayi: Conceptualisation, methodology, formal analysis, investigation, writing – original draft, visualisation, software, validation, data curation. Tamiwe Tomoka: Resources, writing – review and editing, supervision, funding acquisition. Emile R Chimusa: Methodology, writing – review and editing, supervision. George Fedoriw: Resources, writing – review and editing, supervision, funding acquisition. Benjamin Kumwenda: Methodology, project administration, resources, writing – review and editing, supervision, funding acquisition. All authors reviewed the article, contributed to the discussion of results, approved the final version for submission and publication, and take responsibility for the integrity of its findings.

Sources of support

This project was supported by the National Institutes of Health, National Cancer Institute (grant number D43CA260641).

Data availability

The datasets generated or analysed during the current study are available in the Genome Variation Map in National Genomics Data Center, Beijing Institute of Genomics, Chinese Academy of Sciences and China National Center for Bioinformation Repository, under accession number GVM001125.

Disclaimer

The views and opinions expressed in this article are those of the authors and are the product of professional research. It does not necessarily reflect the official policy or position of any affiliated institution, funder, agency, or that of the publisher. The authors are responsible for this article's results, findings, and content.

References

- Denny L. Cervical cancer: Prevention and treatment. *Discov Med*. 2012;14(75):125–131.
- Janicek MF, Averette HE. Cervical cancer: Prevention, diagnosis, and therapeutics. *CA Cancer J Clin*. 2001;51(2):92–114; quiz 115–118. <https://doi.org/10.3322/canjclin.51.2.92>
- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68(6):394–424. <https://doi.org/10.3322/caac.21492>
- Tanaka LF, Schrieffer D, Radde K, Schaubberger G, Klug SJ. Impact of opportunistic screening on squamous cell and adenocarcinoma of the cervix in Germany: A population-based case-control study. *PLoS One*. 2021;16(7):e0253801. <https://doi.org/10.1371/journal.pone.0253801>
- Martínez-Rodríguez F, Limones-González JE, Mendoza-Almanza B, et al. Understanding cervical cancer through proteomics. *Cells*. 2021;10(8):1854. <https://doi.org/10.3390/cells10081854>
- Small W, Bacon MA, Bajaj A, et al. Cervical cancer: A global health crisis. *Cancer*. 2017;123(13):2404–2412. <https://doi.org/10.1002/cncr.30667>
- Arbyn M, Weiderpass E, Bruni L, et al. Estimates of incidence and mortality of cervical cancer in 2018: A worldwide analysis. *Lancet Glob Health*. 2020;8(2):e191–e203. [https://doi.org/10.1016/S2214-109X\(19\)30482-6](https://doi.org/10.1016/S2214-109X(19)30482-6)
- Stang A, Gianicolo E. Age standardization of epidemiological frequency measures. *Dtsch Arztebl Int*. 2025;122(14):387–392.
- Kamiza AB. Incidence and risk factors associated with cervical cancer in sub-Saharan Africa: A systematic review. *ARCH Women Health Care*. 2020;3(1):1–7.
- Gerstl S, Lee L, Nesbitt RC, et al. Cervical cancer screening coverage and its related knowledge in southern Malawi. *BMC Public Health*. 2022;22(1):295. <https://doi.org/10.1186/s12889-022-12547-9>
- Msyamboza KP, Mwagomba BM, Valle M, Chiumia H, Phiri T. Implementation of a human papillomavirus vaccination demonstration project in Malawi: Successes and challenges. *BMC Public Health*. 2017;17(1):599. <https://doi.org/10.1186/s12889-017-4526-y>
- De Freitas AC, Gurgel APAD, Chagas BS, Coimbra EC, Do Amaral CMM. Susceptibility to cervical cancer: An overview. *Gynecol Oncol*. 2012;126(2):304–311. <https://doi.org/10.1016/j.ygyno.2012.03.047>
- Shi Y, Li L, Hu Z, et al. A genome-wide association study identifies two new cervical cancer susceptibility loci at 4q12 and 17q12. *Nat Genet*. 2013;45(8):918–922. <https://doi.org/10.1038/ng.2687>
- Hu Z, Ma D. The precision prevention and therapy of HPV-related cervical cancer: New concepts and clinical implications. *Cancer Med*. 2018;7(10):5217–5236. <https://doi.org/10.1002/cam4.1501>
- Jia M, Han J, Hang D, et al. HLA-DP is the cervical cancer susceptibility loci among women infected by high-risk human papillomavirus: Potential implication for triage of human papillomavirus-positive women. *Tumour Biol*. 2016;37(6):8019–8025. <https://doi.org/10.1007/s13277-015-4673-7>
- Burk RD, Chen Z, Saller C, et al. Integrated genomic and molecular characterization of cervical cancer. *Nature*. 2017;543(7645):378–384. <https://doi.org/10.1038/nature21386>
- Masuda T, Low SK, Akiyama M, et al. GWAS of five gynecologic diseases and cross-trait analysis in Japanese. *Eur J Hum Genet*. 2020;28(1):95–107. <https://doi.org/10.1038/s41431-019-0495-1>
- Rasnic R, Linial N, Linial M. Expanding cancer predisposition genes with ultra-rare cancer-exclusive human variations. *Sci Rep*. 2020;10:13462. <https://doi.org/10.1038/s41598-020-70494-0>
- Bodmer W, Tomlinson I. Rare genetic variants and the risk of cancer. *Curr Opin Genet Dev*. 2010;20(3):262–267. <https://doi.org/10.1016/j.gde.2010.04.016>
- Lenkova K, Khusainova R, Mynyazeva R, et al. Germline variants in proto-oncogenes and tumor suppressor genes in women with cervical cancer. *Biomedicines*. 2024;12(11):2454. <https://doi.org/10.3390/biomedicines12112454>
- Auer PL, Lettre G. Rare variant association studies: Considerations, challenges and opportunities. *Genome Med*. 2015;7(1):16. <https://doi.org/10.1186/s13073-015-0138-2>
- Gudmundsson S, Singer-Berk M, Watts NA, et al. Variant interpretation using population databases: Lessons from gnomAD. *Hum Mutat*. 2022;43(8):1012–1030. <https://doi.org/10.1002/humu.24309>
- Riccio C, Jansen ML, Thalén F, Koliopoulos G, Link V, Ziegler A. Assessment of the functionality and usability of open-source rare variant analysis pipelines. *Brief Bioinform*. 2025;26(1):bbaf044. <https://doi.org/10.1093/bib/bbaf044>
- Tagliaferro SS, Zejnelagic A, Farrugia R, Wettinger SB. Comparison of DNA extraction methods for samples from old blood collections. *BioTechniques*. 2021;70(5):243–250. <https://doi.org/10.2144/btn-2020-0113>
- Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*. 2016;32(19):3047–3048. <https://doi.org/10.1093/bioinformatics/btw354>
- Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet-journal*. 2011;17(1):10–12. <https://doi.org/10.14806/ej.17.1.200>
- McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010;20(9):1297–1303. <https://doi.org/10.1101/gr.107524.110>
- Li H, Durbin R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics*. 2010;26(5):589–595. <https://doi.org/10.1093/bioinformatics/btp698>
- McLaren W, Gil L, Hunt SE, et al. The Ensembl Variant Effect Predictor. *Genome Biol*. 2016;17(1):122. <https://doi.org/10.1186/s13059-016-0974-4>
- Saint Pierre A, Génin E. How important are rare variants in common disease? *Brief Funct Genomics*. 2014;13(5):353–361. <https://doi.org/10.1093/bfgp/elu025>
- Guo MH, Plummer L, Chan YM, Hirschhorn JN, Lippincott MF. Burden testing of rare variants identified through exome sequencing via publicly available control data. *Am J Hum Genet*. 2018;103(4):522–534. <https://doi.org/10.1016/j.ajhg.2018.08.016>
- Etymologia: Bonferroni correction. *Emerg Infect Dis*. 2015;21(2):289. <https://doi.org/10.3201/eid2102.ET2102>
- Sherman BT, Hao M, Qiu J, et al. DAVID: A web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res*. 2022;50(W1):W216–W221. <https://doi.org/10.1093/nar/gkac194>
- Chen M, Sun X, Wang Y, et al. FAT1 inhibits the proliferation and metastasis of cervical cancer cells by binding β -catenin. *Int J Clin Exp Pathol*. 2019;12(10):3807–3818.
- Chen ZG, Saba NF, Teng Y. The diverse functions of FAT1 in cancer progression: Good, bad, or ugly? *J Exp Clin Cancer Res*. 2022;41(1):248. <https://doi.org/10.1186/s13046-022-02461-8>
- Tzvetkov MV, Meineke C, Oetjen E, Hirsch-Ernst K, Brockmöller J. Tissue-specific alternative promoters of the serotonin receptor gene *HTR3B* in human brain and intestine. *Gene*. 2007;386(1):52–62. <https://doi.org/10.1016/j.gene.2006.08.002>
- Chen JR, Huang MS, Lee YC, Lin MH, Yang YF. Potential clinical value of 5-hydroxytryptamine receptor 3C as a prognostic biomarker for lung cancer. *J Oncol*. 2021;2021(1):1901191. <https://doi.org/10.1155/2021/1901191>
- Zhan D, Wang X, Zheng Y, et al. Integrative dissection of 5-hydroxytryptamine receptors-related signature in the prognosis and immune microenvironment of breast cancer. *Front Oncol*. 2023;13:1147189. <https://doi.org/10.3389/fonc.2023.1147189>
- Liu X, Xu Y, Jin Q, et al. EphA8 is a prognostic marker for epithelial ovarian cancer. *Oncotarget*. 2016;7(15):20801. <https://doi.org/10.18632/oncotarget.8018>
- Campa D, Müller P, Edler L, et al. A comprehensive study of polymorphisms in ABCB1, ABCC2 and ABCG2 and lung cancer chemotherapy response and prognosis. *Int J Cancer*. 2012;131(12):2920–2928. <https://doi.org/10.1002/ijc.27567>
- Chen Y, Zhou H, Yang S, Su D. Increased ABCC2 expression predicts cisplatin resistance in non-small cell lung cancer. *Cell Biochem Funct*. 2021;39(2):277–286. <https://doi.org/10.1002/cbf.3577>
- LENG Jie QG. Mechanism of breast cancer centrosome regulatory protein SEC23B on tumor invasion and metastasis. *China Oncol*. 2023;33(2):152–161.
- Yehia L, Jindal S, Komar AA, Eng C. Non-canonical role of cancer-associated mutant SEC23B in the ribosome biogenesis pathway. *Hum Mol Genet*. 2018;27(18):3154–3164. <https://doi.org/10.1093/hmg/ddy226>
- Kalyvianaki K, Panagiotopoulos AA, Malamos P, et al. Membrane androgen receptors (OXER1, GPRC6A AND ZIP9) in prostate and breast cancer: A comparative study of their expression. *Steroids*. 2019;142:100–108. <https://doi.org/10.1016/j.steroids.2019.01.006>
- Pi M, Quarles LD. GPRC6A regulates prostate cancer progression. *Prostate*. 2012;72(4):399–409. <https://doi.org/10.1002/pros.21442>
- Chen B, Wang Y, Tang W, et al. Association between PPAR γ , PPARGC1A, and PPARGC1B genetic variants and susceptibility of gastric cancer in an Eastern Chinese population. *BMC Med Genomics*. 2022;15(1):274. <https://doi.org/10.1186/s12920-022-01428-0>
- Wirtenberger M, Tchatchou S, Hemminki K, et al. Associations of genetic variants in the estrogen receptor coactivators PPARGC1A, PPARGC1B and EP300 with familial breast cancer. *Carcinogenesis*. 2006;27(11):2201–2208. <https://doi.org/10.1093/carcin/bgl067>
- Li Y, Li Y, Wedrén S, et al. Genetic variation of ESR1 and its co-activator PPARGC1B is synergistic in augmenting the risk of estrogen receptor-positive breast cancer. *Breast Cancer Res*. 2011;13(1):R10. <https://doi.org/10.1186/bcr2817>
- Kalmár A, Péterfia B, Hollósi P, et al. DNA hypermethylation and decreased mRNA expression of MAL, PRIMA1, PTGDR and SFRP1 in colorectal adenoma and cancer. *BMC Cancer*. 2015;15(1):736. <https://doi.org/10.1186/s12885-015-1687-x>
- Qin Q, Ramesh S, Gomez-Salazar MJ, et al. CNTNAP4 signaling regulates osteosarcoma disease progression. *NPJ Precis Oncol*. 2023;7(1):1–13. <https://doi.org/10.1038/s41698-022-00344-x>