

https://dspace.org/

The germline genetic landscape of paediatric central nervous sysetem tumours

Chatergoon, Chetana

2025
University of the Witwatersrand, Johannesburg
https://hdl.handle.net/10539/48052

Chatergoon, Chetana. (2025). The germline genetic landscape of paediatric central nervous sysetem tumours [Master`s dissertation, University of the Witwatersrand, Johannesburg]. WIReDSpace. https://hdl.handle.net/10539/48052

Downloaded from a DSpace repository. https://dspace.org/

THE GERMLINE GENETIC LANDSCAPE OF PAEDIATRIC CENTRAL NERVOUS SYSTEM TUMOURS

Chetana Chatergoon

Student number: 2687027

Supervisor: Dr Lindie Lamola

Co-supervisor: Dr Khuthala Mnika



UNIVERSITY OF THE
WITWATERSRAND,
JOHANNESBURG

A research report (in the format of a "submissible" paper) submitted to the Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, in partial fulfilment of the requirements for the degree of Master of Science in Medicine (Genomic Medicine) by Research and Coursework

Johannesburg, 2025

NATIONAL HEALTH LABORATORY SERVICE
School of Pathology, University of the Witwatersrand
DIVISION OF HUMAN GENETICS
Hospital Street, Johannesburg, 2001 | PO Box 1038, Johannesburg, 2000
[T] +27 11 489 9223 | [M] +27 78 080 8841 | [F] +27 11 489 9226 | [E] human.genetics@nhls.ac.za

29 November 2024

Re: Research report submitted to the Faculty of Health Sciences, University of the Witwatersrand, by Ms. Chetana Chatergoon in partial fulfilment of the requirements for the degree Master of Science in Medicine (Genomic Medicine)

Dear examiner,

Thank you for agreeing to examine the research report submitted by Ms. Chetana Chatergoon (student number 2687027) entitled "The germline genetic landscape of paediatric central nervous system tumours." This research report is being submitted by the candidate in the form of a "submissible" paper as per the Style Guide for Thesis, Dissertation and Research Reports (updated March 2016) by the University of the Witwatersrand Faculty of Health Sciences. The research report constitutes 33% towards the final mark of the MSc (Med) Genomic Medicine degree for which the candidate is enrolled.

The candidate and her supervisors have chosen to submit the paper to the Journal of Oncology. This international journal is a peer-reviewed open-access medical journal published by Frontiers. The submissible paper attached is therefore written in accordance with the author guidelines of the stipulated journal. These guidelines have been attached for your reference in the appendix section of this document. The only deviation from these guidelines is that the manuscript has been presented as one concise document (including all tables and figures) for ease of marking, rather than being submitted separately.

The research protocol has also been attached as an appendix for the purpose of providing an extended literature review and further contextualise the study. The protocol has already been assessed and approved by the Faculty of Health Sciences Postgraduate Assessors Committee and is therefore not for examination.

Please refer to the Table of Contents for a complete list of the contents provided in this research report.

Yours sincerely,

Chetana Chatergoon (Candidate)

Lindie Lamola (Supervisor)

Khuthala Mnika (co- supervisor)

This report is intended solely to record the observations and/or opinion of the writer. It does not constitute a medico-legal report

Chairperson: Prof Jeffrey Mphahlele CEO: Prof Koleka Mlisana
Physical Address: 1 Modderfontein Road, Sandringham, Johannesburg, South Africa Postal Address: Private Bag X8, Sandringham, 2131, South Africa
Tel: +27 (0) 11 386 6000/ 0860 00 NHLS(6457) www.nhls.ac.za
Practice number: 5200296

Declaration

I, Chetana Chatergoon, declare that this Research Report (in the format of a “submissible” paper) is my own, unaided work. It is being submitted for the Degree of Master of Science in Medicine (Genomic Medicine) by Research and Coursework at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.



Chetana Chatergoon

Student number: 2687927

4th day of March 2024 in Johannesburg

Contribution of the Candidate

Student's contribution to the article(s) and agreement of co-author(s)

I, Chetana Chatergoon, student number 2687027, declare that this Research Report, entitled "The germline genetic landscape of paediatric central nervous system tumours", is my own work and that I contributed significantly towards the research findings presented in the paper intended for publication below.

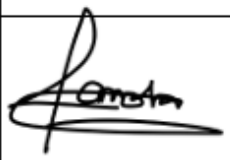


Signature of Student Date: 29/11/2024



Signature of Primary Supervisor Date: 30/11/2024

Agreement by co-authors: By signing this declaration, the co-authors listed below agree to the use of the research findings intended for a published article by the student as part of their Research Report.

Authors	Name	Signature	Date
1st	Ms. Chetana Chatergoon		30/11/2024
2nd	Dr. Khuthala Mnika		02/12/2024
3rd	Dr. Lindie Lamola		30/11/2024

Presentations arising from this research project

1. YEGG Symposium, 26th October, Johannesburg, Gauteng (rapid-fire poster presentation)
2. SASHG Biennial Congress 2024, 27th-29th October, Johannesburg, Gauteng (poster presentation)

Abstract

Central nervous system (CNS) cancers represent the most prevalent solid tumours in children and are among the leading causes of cancer-related deaths in this population. Inherited germline variants in cancer-predisposing genes (CPG), elevate the risk of cancer development during childhood. In South Africa, there is a scarcity of research on the prevalence of paediatric cancers and the influence of inherited genetic factors that heighten the risk of early cancer development in children. Next-generation sequencing technologies have played a pivotal role in identifying and understanding the contribution of germline variants to the development of paediatric cancers, advancing both research and diagnostic efforts globally. Through the design and analysis of a virtual panel including 56 CPGs, this study aimed to identify germline genetic variants in CPGs in a subset of paediatric patients with CNS tumours. Whole-exome sequencing (WES) data from 13 individuals diagnosed with CNS tumours, including medulloblastoma, astrocytoma, optic nerve glioma, craniopharyngioma, and pineoblastoma, and unspecified brain tumours were analysed. The workflow included annotation, filtering, prioritisation, and rigorous quality control of the candidate variants. Eleven candidate variants were subsequently classified in accordance with the guidelines established by the American College of Medical Genetics and Genomics and the Association for Molecular Pathology (ACMG-AMP). Pathogenic variants were detected in the *ATM* (c.2200dup; p.Val734GlyfsTer4; c.8911C>T; p.Gln2971Ter), and *NF1* (c.625C>T; p.Gln209Ter) genes. These variants were identified in individuals with medulloblastoma and optic nerve glioma, respectively. The positive pick-up rate of this study was 15,38%, which is in-keeping with international published oncology research. This underscores the role of *ATM* and *NF1* variants in contributing to cancer susceptibility in individuals with paediatric CNS tumours. This study contributes to our understanding of the germline genetic landscape underlying paediatric CNS tumours in South Africa.

Acknowledgements

I would like to express gratitude to my supervisor, Dr. Lindie Lamola, whose guidance, patience, and encouragement have been the cornerstone of this research. Her invaluable expertise and unwavering belief in my abilities have been pivotal in shaping the direction and success of this research. I also extend my heartfelt gratitude to my co-supervisor, Dr. Khuthala Mnika, whose profound knowledge and insightful perspectives have played a crucial role in refining this research.

I am deeply grateful to everyone at the Division of Human Genetics for fostering a supportive and collaborative environment. Their collective expertise and generosity in sharing knowledge have significantly contributed to both my academic and personal growth. This research would not have been possible without the advice and guidance of my supervisors, friends, and colleagues in the department, whose work has laid the foundation for my own.

I extend my deepest appreciation to my classmates for their unwavering support, and all of the laughs that we have shared throughout this journey. The shared experiences, camaraderie, and encouragement we've offered one another have been a constant source of joy and motivation, especially during the more challenging moments of this Master's degree. This truly has been a collective effort, and I am profoundly grateful for the sense of community we've built together.

Lastly, I extend my heartfelt thanks to my family and friends, whose love and steadfast support have been my greatest strength. Their unwavering belief in me has been a wellspring of inspiration and motivation throughout this journey. I could not have reached this point without them, and for that, I am forever grateful. You are my foundation, and I dedicate this work to you.

A special thanks to the individuals and families whose genetic data have informed this work. Your courage in contributing to research will pave the way to a deeper understanding of the germline genetic landscape of central nervous system tumours.

Table of Contents

Declaration iii

Presentations arising from this research project v

Abstract vi

Acknowledgements vii

List of figures x

List of abbreviations xi

Research report in the format of a "submissible" paper xiii

1. Introduction 1

2. Materials and methods 3

2.1 Ethical considerations 3

2.2 Study design 3

2.3 Study cohort 3

2.5 Variant annotation and filtering 4

2.6 Variant prioritisation 5

2.7 Variant quality assessment 5

2.8 Variant classification using ACMG-AMP guidelines 5

3. Results 6

3.1 Patient characteristics 6

3.2 Variant annotation and filtering 6

3.3 Shortlisted variants 8

3.4 Variant quality assessment 9

3.5 Variant classification according to ACMG-AMP guidelines 9

3.5.1 Pathogenic variants 10

3.5.1.1 PC97: Medulloblastoma 10

3.5.1.2 PC111: Optic nerve glioma 11

3.5.2 Variants of uncertain significance 11

4.2 PC111: Optic nerve glioma 16

4.3 PC96, 86, 139, 107, 137, and 89 17

4.4 Limitations of the study 18

4.5 Future work 18

5. Conclusion 18

6. Declarations 20

Statement of Ethics 20

Data Availability Statement 20

Acknowledgement 20
Author Contributions 20
References 21
Appendices 26
Appendix A: Supplementary Tables and Figures 26
Appendix B: Ethics Clearance Certificate (M240758) 42
Appendix C: Ethics Clearance Certificate (M240262) 44
Appendix D: Approved Research Protocol 46
Appendix E: Turnitin Report and Plagiarism Declaration 66
Appendix F: Author guidelines for Frontiers in Oncology 73

List of figures

Figure 1. Distribution of coding sequence variant types among the eight individuals after annotation with Ensembl VEP..... 7

List of tables

Table 1. Study cohort details including IonXpress codes, patient IDs, suspected CNS cancers, patient sex, and age of onset.....6

Table 2. Number of variants before and after filtering, associated genes and disorders or syndromes.....7

Table 3. Shortlisted moderate and high impact variants in cancer predisposition genes.....9

Table 4. Results following the classification of shortlisted variants according to the ACMG-AMP guidelines.....12

Supplementary Table 1. List of candidate paediatric CNS tumour predisposition genes (n=56).....28

Supplementary Table 2. Details, supporting evidence and ACMG-AMP classification of pathogenic variants (n=3).....34

Supplementary Table 3. Details, supporting evidence and ACMG-AMP classification of variants of uncertain significance (n=8).....35

List of abbreviations

ACMG-AMP	American College of Medical Genetics and Genomics and Association of Medical Pathology
AT	Ataxia-telangiectasia
BAI	Binary Alignment Index
BAM	Binary Alignment Map
BCGene	Brain Cancer Gene Database
BRCA1	Breast cancer 1 gene
CADD	Combined Annotation Dependent Depletion
CNS	Central nervous system
CNV	Copy number variant
CPG	Cancer predisposition gene
CPS	Cancer predisposition syndrome
CSRD	Cysteine and serine-rich domain
DNA	Deoxyribonucleic acid
FATHMM-MLK	Functional Analysis Through Hidden Markov Models – Multiple Kernel Learning
gnomAD	Genome Aggregation Database
HEAT	Huntingtin, Elongation factor 3, protein phosphatase 2A, and TOR1 repeat
HGVS	Human Genome Variation Society
IGV	Integrative Genomics Viewer
MAF	Minor allele frequency
MLH1	MutL protein homolog 1 gene
MMR	Mismatch repair
MSH6	MutS homolog 6 gene
NCBI	National Centre for Biotechnology Information
NF1	Neurofibromatosis Type 1
NF1	Neurofibromin 1
NBN	Nibrin
PMS2	PMS1 homolog 2, mismatch repair system component gene
POLE	DNA polymerase epsilon, catalytic subunit gene
RAS	Rat sarcoma gene

SIFT	Sorting Intolerant From Tolerant
TP53	Tumour protein p53 gene
VCF	Variant Call Format
VEP	Variant Effect Predictor
VUS	Variant of Uncertain Significance
WES	Whole-exome sequencing

Research report in the format of a "submissible" paper

To be submitted to Frontiers in Oncology.

Research Article

The germline genetic landscape of paediatric central nervous system tumours

Chetana Chatergoon[1], Khuthala Mnika[1,2], Lindie Lamola[1]

[1]Division of Human Genetics, National Health Laboratory Service, and School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

[2]Division of Human Genetics, Department of Pathology, Faculty of Health Sciences, & Institute of Infectious Disease and Molecular Medicine, University of Cape Town , South Africa

Corresponding Author: Chetana Chatergoon

Division of Human Genetics, Corner Hospital and De Korte Streets, JHB, South Africa, 2001

E-mail address: 2687027@students.wits.ac.za

Abstract for journal submission

Central nervous system (CNS) cancers are the most common type of paediatric solid tumours and have some of the highest cancer-related mortality rates. Paediatric CNS tumours occur in patients with a genetic predisposition in 8 to 12% of cases. Next-generation sequencing technologies have been used to identify putative disease-causing variants in cancer predisposition genes, with reported diagnostic yields ranging from 10% to 35% using Whole-exome sequencing (WES). This study aimed to identify germline genetic variants in cancer predisposition genes in a subset of paediatric patients with CNS tumours. Pre-generated WES data from 13 individuals with different types of CNS tumours including: medulloblastoma, astrocytoma, optic nerve glioma, craniopharyngioma, pineoblastoma, and unspecified brain tumours was analysed. A virtual panel consisting of 56 genes associated with germline cancer predisposition syndromes and childhood CNS tumours was investigated. Variants were annotated, filtered, and prioritised. Shortlisted variants were classified following the American College of Medical Genetics and Genomics and Association for Molecular Pathology (ACMG-AMP) guidelines. Notably, variants in the *ATM* (c.2200dup; p.Val734GlyfsTer4; c.8911C>T; p.Gln2971Ter), and *NF1* (c.625C>T; p.Gln209Ter) genes were classified as pathogenic, in cases of medulloblastoma and optic nerve glioma. This study forms part of a larger study which is investigating the genetics of paediatric cancers in South Africa, The PeCan (Paediatric Cancer) project. Knowledge gained in this study advances our understanding of the germline genetic landscape underlying paediatric CNS tumours in South Africa, and the positive pick-up rate (15,38%) is in-keeping with both global and South African oncology research.

Keywords: paediatric cancer, central nervous system tumours, cancer genetics, germline variants, whole exome sequencing

1. Introduction

Paediatric central nervous system (CNS) tumours are brain and spinal cord cancers that originate from the abnormal proliferation of nerve or supportive cells (Feh et al., 2021; Sait et al., 2021). Malignant brain tumours are the most common solid cancer in children (Capasso et al., 2020; Raja et al., 2021). These tumours have a global incidence of approximately 5.5 per 100,000 and a mortality rate of 0.72 per 100,000, making them the leading cause of cancer-related mortality in children (Capasso et al., 2020). Their incidence in Africa is 0.66 per 100,000, however, there is no data on the incidence of CNS tumours in South Africa, reflecting gaps in regional cancer reporting and research (Raja et al., 2021). Common paediatric CNS tumours include astrocytoma, medulloblastoma, craniopharyngioma, optic nerve glioma, pineoblastoma, and other unspecified gliomas (Sait et al., 2021). Feh et al. (2021) reported the following frequencies of CNS cancers in Africa between 1960 and 2017, astrocytoma represented 24.70% of cases, 4.26% were medulloblastoma, 4.07% were craniopharyngioma, glioblastoma and unspecified gliomas were 2.20% and pineoblastoma made up 0.19% (Feh et al., 2021).

Approximately 75% to 85% of cancers are attributed to sporadic somatic pathogenic variants, whereas the remaining 15% to 25% are associated with germline genetic variants in cancer predisposition genes (CPG) (Capasso et al., 2020; Sait et al., 2021). A subset of these inherited cancers arises from underlying cancer predisposition syndromes (CPS), which are characterised by germline mutations that significantly increase the risk of developing early-onset cancers compared to the general population (Gröbner et al., 2018). It is estimated that 8% to 12% of paediatric cancers occur in individuals with germline predisposing pathogenic variants, with 8% to 19% of these occurring in a classical CPG that is associated with a CPS (Sait et al., 2021). Although, research indicates that this percentage may be higher in children with central nervous system (CNS) tumours (Kline et al., 2017). Akhavanfard et al. (2020) reported that 7% to 8% of individuals diagnosed with solid tumours before the age of twenty years harboured pathogenic germline variants in classical CPGs, with high-grade gliomas accounting for 25% of these cases, which is the highest proportion among all tumours.

The rarity and complexity of paediatric CNS tumours pose significant challenges for research. According to the Brain Cancer Gene Database (BCGene), 1,421 genes are associated with brain cancer. However, only 27 genes have been referenced in more than 20 National Centre for Biotechnology Information (NCBI) PubMed abstracts, while approximately 881 genes are cited just once (Zhao et al., 2021). This highlights the need for further studies to elucidate the functions of these genes. The landscape of germline predisposing variants in children remains largely unexplored, and investigating these mutations could significantly advance our understanding of tumorigenesis.

Next-generation sequencing has proven effective for investigating the genetic landscape of CNS tumours, enabling the detection of single nucleotide variants and small insertions or deletions in coding regions (Akhavanfard et al., 2020; Gröbner et al., 2018). This approach has demonstrated a significant diagnostic yield, with Kline et al. (2017) reporting a diagnostic yield of 35% in a cohort of 32 patients with CNS tumours through targeted sequencing of 510 CPGs. Similarly, Gröbner et al. (2018) found that 7.6% of all paediatric tumours were likely associated with pathogenic germline variants. Next-generation sequencing is a high-throughput technique that facilitates parallel DNA sequencing, enabling the identification of genetic variants through comparisons with a reference genome (Parsons et al., 2016). NGS can be implemented within different applications, including whole genome sequencing, whole exome sequencing (WES), RNA sequencing, targeted sequencing, panel sequencing, methylation sequencing, and a set of array-based sequencing (Kline et al., 2017). Studies by Parsons et al. (2016) and Akhavanfard et al. (2020) have employed NGS, specifically WES, to explore germline variants in paediatric patients diagnosed with solid tumours. Their findings demonstrated the utility of NGS and bioinformatics in understanding the genetic underpinnings of solid tumours. For instance, Parsons et al. (2016) used WES to examine CNS and non-CNS malignancies, reporting that 10% of their cohort carried germline variants relevant to their disease phenotypes. Similarly, Akhavanfard et al. (2020) identified pathogenic or likely pathogenic germline variants in 10% of their cases.

While WES has been transformative in advancing precision oncology and identifying genetic variants in CPGs, its application has been largely limited to research in developed parts of the world, such as the United States and Europe (Gröbner et al., 2018; Kline et al., 2017; Feh et al., 2021). These international studies have elucidated the contribution of germline variants to paediatric cancer development, but have largely excluded populations from underrepresented regions, such as Africa (Rotimi, Rotimi and Salhia, 2021). To date, there has only been one study in South Africa utilising WES to investigate the genetic landscape of paediatric solid tumours (Manolas, 2023). This study focused on mutation profiling in a cohort of 32 paediatric solid tumour patients and reported a diagnostic yield of approximately 9.4%. The findings highlight the potential of WES as a valuable tool in understanding the genetic landscape of paediatric cancers within this population.

CNS tumours are underreported in Africa, contributing to the misconception that they are rare across the continent (Feh et al., 2021; Rotimi, Rotimi and Salhia, 2021). Systematic studies on their incidence and mortality have not been conducted in Africa (Feh et al., 2021). Consequently, current research may not adequately capture the diverse genetic landscapes and unique predispositions to CNS cancers prevalent in African populations. Rotimi et al. (2021) highlighted that despite Africa's substantial genetic diversity, only 0.329% of global cancer studies focus on African populations, with just 0.016%

addressing cancer genetics and genomics. Papers exclusively reporting on paediatric CNS cancers constituted 38.5% of existing literature. Expanding research to comprehensively investigate germline variants in African populations could provide insights into tumorigenesis, improve early detection, and enable personalised treatment strategies tailored to the continent's unique genetic diversity (Rotimi, Rotimi and Salhia, 2021). Therefore, the aim of this study is to identify germline genetic variants in cancer predisposition genes in a subset of paediatric patients with CNS tumours.

2. Materials and methods

2.1 Ethical considerations

Ethical clearance to conduct this study was obtained from the University of the Witwatersrand, Medical Research Ethics Committee (approval number: M240758). This is a sub-study that falls under the Paediatric Cancer (PeCan) project (M240262). Prior to the commencement of this study, consent to use the patient samples for research purposes was obtained during the recruitment process.

2.2 Study design

This study was retrospective in nature. Cases were obtained from the PeCan project and patients were recruited between 2019 and 2023. DNA was extracted from whole blood samples and WES was performed using the Ion GeneStudio S5™ Series Sequencer by ThermoFisher Scientific, using the Ion AmpliSeq™ Library Kit Plus (ThermoFisher Scientific, Waltham). For the current study, files containing pre-generated WES data in the Variant Call Format (VCF), Binary Alignment Map (BAM) and index file format (BAI) were obtained for data analysis.

2.3 Study cohort

The study cohort included thirteen paediatric patients with various CNS tumours recruited through the PeCan project. These included: medulloblastoma (one), astrocytoma (five), optic nerve glioma (three), craniopharyngioma (one), pineoblastoma (one), and other unspecified brain tumours (two). None of the individuals in the study cohort have a reported family history of CNS tumours or any form of cancer.

2.4 Candidate gene selection

A critical literature review was conducted to identify candidate CPGs that are associated with germline inheritance or CPS's. A list of 56 candidate genes associated with germline inheritance in medulloblastoma, astrocytoma, optic nerve glioma, craniopharyngioma, pineoblastoma, and other

unspecified brain tumours was curated. A literature search was performed using the NCBI PubMed database (https://pubmed.ncbi.nlm.nih.gov/), BCGene (https://bioinfo-minzhao.org/soft/bcgene/), and Google Scholar (https://scholar.google.com/), focusing on studies published within the last ten years. The search terms included “paediatric CNS tumours”, “paediatric CNS tumours genetics”, “molecular profiling paediatric brain tumours”, “germline inheritance” and “paediatric medulloblastoma/ astrocytoma/ optic nerve glioma/ craniopharyngioma/ pineoblastoma,”.

Sources included literature reviews and original research studies using curated CNS cancer gene panels to identify common genes underlying paediatric cancers. Genes that overlapped between commercial panels, for example, the VistaSeq Brain/CNS/PNS Cancer Panel and Invitae Hereditary Nervous System/Brain Cancer Panel were prioritised, as multiple panels included the same genes (Alonso-Luna et al., 2023; NIH genetic testing registry, 2024). A comprehensive list of the candidate genes and their details can be found in Supplementary Table 1 (Appendix A).

2.5 Variant annotation and filtering

The Ensembl Variant Effect Predictor (VEP) tool (https://www.ensembl.org/Tools/VEP) was utilised to annotate variants, by comparing the short-read sequences generated from WES to the GRCh37.p13 reference genome (McLaren et al., 2016). Genetic variants in candidate genes were identified by comparing the VCF file sequences with the reference genome. Annotations were made, such as the type, impact, location, and consequence of the variant, as well as its functional effect on the gene and its protein. Human Genome Variation Society (HGVS) nomenclature (HGVSc and HGVSp) was assigned to known variants (Dunnen et al., 2016).

Variant filtering was performed according to the evaluation of disease mechanism and clinical significance. This was performed manually, with the application of frequency filters on Microsoft Excel, database searches and literature searches to exclude common variants and predict the functional effect of moderate and high impact variants. A tiered approach was taken, and a different subset of candidate CPGs (Supplementary Table 1; Appendix A) was selected for the analysis of each sample, based on the type of CNS tumour. If no candidate variants were identified in these genes, the second tier consisted of general CNS cancer susceptibility genes, such as oncogenes, DNA mismatch repair (MMR) and tumour suppressor genes. Variants that were intronic, synonymous, or had a minor allele frequency (MAF) greater than 0.05 in the Genome Aggregation Database (gnomAD) were excluded (Karczewski et al., 2020). Annotations were saved as a TXT file and imported into Microsoft Excel for filtering. These filtering steps resulted in a refined list of candidate variants for further investigation.

2.6 Variant prioritisation

The candidate variants were prioritised using scores from in silico pathogenicity prediction tools. These included Combined Annotation-Dependent Depletion (CADD), which assigns deleteriousness scores to variants (https://cadd.gs.washington.edu/); as well as Sorting Intolerant from Tolerant (SIFT) (http://sift.bii.a-star.edu.sg/), AlphaMissense (https://alphamissense.hegelab.org/), and Functional Analysis Through Hidden Markov Models - Multiple Kernel Learning (FATHMM-MLK) (http://fathmm.biocompute.org.uk/fathmmMKL.htm), which predict the impact of amino acid substitutions on protein structure and function (Cheng et al., 2023; Kumar, Henikoff and Ng, 2009; Rentzsch et al., 2018; Shihab et al., 2015).

Additional resources, such as variant-specific databases, were consulted to aid in variant prioritisation. These resources included gnomAD (https://gnomad.broadinstitute.org/), UniProt (https://www.uniprot.org/), NCBI ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/), VarSome (https://varsome.com/), and Franklin (https://franklin.genoox.com). Published literature was reviewed using the NCBI PubMed database and Google Scholar. Thereafter, a shortlist including variants of uncertain significance (VUS), pathogenic and likely pathogenic variants in each patient was generated.

2.7 Variant quality assessment

A BAM file was used to evaluate the quality of variants in the shortlisted cases. The variant sites were visualised using Integrative Genomics Viewer (IGV) (https://software.broadinstitute.org/software/igv/), a robust genome browser for confirming true positive variants (Robinson et al., 2022). The coverage of the target sequences was assessed and variants with at least 30 reads were deemed reliable. Allele ratios were analysed to determine the zygosity of each variant.

2.8 Variant classification using ACMG-AMP guidelines

The shortlisted variants were classified using the American College of Medical Genetics and Genomics and Association of Medical Pathology (ACMG-AMP) guidelines (Richards et al., 2015). A combination of weighted codes was used to classify each shortlisted variant as: of uncertain significance, likely pathogenic and pathogenic. A literature search was conducted to find previous reports of the identified variants.

3. Results

3.1 Patient characteristics

The cohort included 13 paediatric patients with CNS tumours, consisting of 9 females and 4 males, aged between one and eleven years old. Twelve participants were of African ancestry, with the exception of PC111, whose ancestry is unknown. Table 1 provides a comprehensive overview of the study cohort, including Patient IDs and type CNS cancers, patient sex, age of onset and ancestral background.

Table 1. Study cohort details including patient IDs, CNS cancers, patient sex, and age of onset

Patient ID and type of CNS cancer	Patient sex	Age of onset (years (y))	Ancestral background
PC15: Brain tumour (unspecified)	Male	3y	African
PC75: Optic nerve glioma	Female	2y	African
PC78: Astrocytoma	Male	11y	African
PC85: Pineoblastoma	Female	7y	African
PC86: Craniopharyngioma	Female	7y	African
PC89: Astrocytoma	Female	6y	African
PC96: Astrocytoma	Male	6y	African
PC97: Medulloblastoma	Female	4y	African
PC98: Astrocytoma	Male	1y	African
PC107: Astrocytoma	Female	5y	African
PC111: Optic nerve glioma	Female	1y	Unknown
PC137: Optic nerve glioma	Female	3y	African
PC139: Brain tumour (unspecified)	Female	3y	African

3.2 Variant annotation and filtering

A total of 542,605 variants were identified among 13 patients, with an average of 41,739 variants detected per individual. Among the coding sequence variants identified, synonymous variants were the most prevalent, followed by missense variants. Frameshift and in-frame deletions were the least frequently observed, according to Ensembl VEP variant annotations (Figure 1).

Following variant filtering, one frameshift, two nonsense, and eight missense variants were identified in CPGs associated with medulloblastoma, optic nerve glioma, astrocytoma, craniopharyngioma, and an unspecified brain tumour. Eleven candidate variants in eight genes were shortlisted. The number of variants before filtering and following prioritisation, as well as the gene in which variants were identified, and the associated disorder is summarised in Table 2.

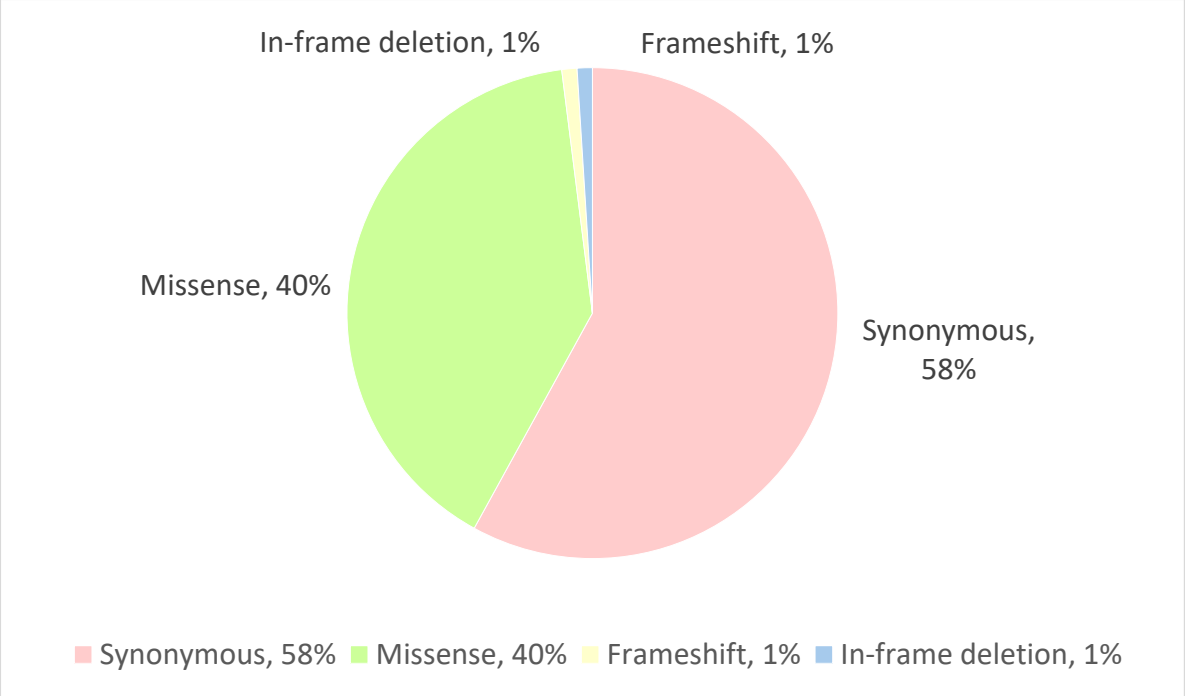


Figure 1. Distribution of coding sequence variant types among the eight individuals after annotation with Ensembl VEP. Approximately 58% of the observed variants were synonymous, 40% were missense, while frameshift and in-frame deletion variants each accounted for 1%.

Table 2. Number of variants before and after filtering, associated genes and disorders or syndromes

Patient ID and type of CNS cancer	Number of variants	Prioritised variants	Genes	Disorder or syndrome
PC97: Medulloblastoma	42,102	2	*ATM*	Ataxia-telangiectasia
PC111: Optic nerve glioma	35,902	1	*NF1*	Neurofibromatosis type 1
PC96: Astrocytoma	41,143	3	*PMS2*, *NF1*, *POLE*	Turcot syndrome Type 1/ Lynch syndrome; Neurofibromatosis type 1; Polymerase Proofreading-Associated Polyposis syndrome
PC86: Craniopharyngioma	40,684	1	*MSH6*	Constitutional mismatch repair deficiency syndrome/ Lynch syndrome/ Turcot syndrome Type 1
PC139: Brain tumour (unspecified)	36,181	1	*MLH1*	Constitutional mismatch repair deficiency syndrome/ Lynch syndrome/ Turcot syndrome Type 1
PC107: Astrocytoma	44,998	1	*PMS1*	Lynch syndrome
PC137: Optic nerve glioma	43,815	1	*IDH1*	-
PC89: Astrocytoma	42,465	1	*PMS1*	Lynch syndrome
PC75: Optic nerve glioma	39,429	0	-	-
PC78: Astrocytoma	40,197	0	-	-
PC85: Pineoblastoma	42,237	0	-	-
PC98: Astrocytoma	35,134	0	-	-
PC15: Brain tumour (unspecified)	58,318	0	-	-

Abbreviations: ATM- Ataxia telangiectasia mutated; NF1- Neurofibromin 1; PMS2- PMS1 homolog 2, mismatch repair system component; POLE- POLE DNA polymerase epsilon, catalytic subunit, MSH6- MutS homolog 6; MLH1- MutL protein homolog 1; PMS1- PMS1 homolog 1, mismatch repair system component; Isocitrate dehydrogenase 1.

3.3 Shortlisted variants

Following variant prioritisation, a shortlist of pathogenic, likely pathogenic and variants of uncertain significance was generated. Eleven candidate variants were identified in eight CPGs across eight individuals (Figure 2). No variants of interest were detected in CPGs in samples PC15, 75, 85, and 98, therefore they were excluded from the shortlist.

The majority of the prioritised variants were observed once in the cohort. PC97, a female with medulloblastoma, was identified as a compound heterozygote for two different pathogenic *Ataxia*

telangiectasia mutated (*ATM*) variants. A pathogenic variant in the *Neurofibromin 1* (*NF1*) gene was identified in PC111, a female with optic nerve glioma. A variant in the *PMS1 homolog 1, mismatch repair system component* (*PMS1*) gene was identified in two individuals with astrocytoma (PC89 and 107). The remainder of the cohort had missense variants in Tumour suppressor, DNA MMR genes, such as *PMS1 homolog 2, mismatch repair system component* (*PMS2*), *MutL protein homolog 1* (*MLH1*), *MutS homolog 6* (*MSH6), Isocitrate dehydrogenase 1* (*IDH1*), and *POLE DNA polymerase epsilon, catalytic subunit* (*POLE*) ,and *NF1*. Variants were prioritised based on their reported clinical significance on NCBI ClinVar, moderate/high predicted impact and their low minor allele frequency (MAF) in gnomAD control data. Details such as variant type and genomic position, and HGVSc and HGVSp nomenclature are summarised in Table 3.

3.4 Variant quality assessment

The variants were reviewed using IGV to determine their zygosity and evaluate their read depth for reliability. All of the candidate variants were confirmed to be heterozygous, with allele frequencies approximating a 50:50 ratio (Supplementary Figure 1, Appendix A). Coverage assessment showed read depths exceeding 30x for all samples, confirming that all identified variants were true variants.

3.5 Variant classification according to ACMG-AMP guidelines

A comprehensive search for supporting evidence was performed using resources such as gnomAD, AlphaMissense, CADD, SIFT, UniProt, FATHMM-MLK, NCBI ClinVar, VarSome, Franklin, and NCBI PubMed to aid in variant classification. Germline pathogenic variants were identified in two out of thirteen patients with CNS tumours, representing 15,38% of the cohort (Figure 2). Two *ATM* variants (NM_000051.4:c.2200dup; c.8911C>T) and one *NF1* variant (NM_000267.3:c.625C>T) were classified as pathogenic based on ACMG-AMP codes PSV1 (null variant in a gene where loss-of-function (LOF) is a known disease mechanism), PM2 (absence in population databases), and PP5 (pathogenicity cited in clinical literature) (Table 4; Supplementary Table 2, Appendix A). Eight VUSs were also identified, attributed to either conflicting evidence, or limited data. Five of the eight VUSs were considered to be "hot," indicating that they had considerable supportive evidence, and could potentially be upgraded to likely pathogenic or pathogenic with additional information, such as confirmation of de novo status (Table 5). Supporting evidence, including ClinVar IDs, gnomAD African/African American allele frequencies, and *in silico* pathogenicity prediction scores, is detailed in Supplementary Table 3 (Appendix A).

Table 3. Shortlisted moderate and high impact variants in cancer predisposition genes

Patient ID and type of CNS cancer	Gene and variant	Variant type	Genomic location	HGVSc and HGVSp
PC97: Medulloblastoma	*ATM*(c.2200dup; p.Val734GlyfsTer4)	Frameshift	11:108127016-108127016	NM_000051.4:c.2200dup NP_000042.3:p.Val734GlyfsTer4
	ATM(c.8911C>T; p.Gln2971Ter)	Nonsense	11:108235869-108235869	NM_000051.4:c.8911C>T NP_000042.3:p.Gln2971Ter
PC111: Optic nerve glioma	*NF1*(c.625C>T; p.Gln209Ter)	Nonsense	17:29508478-29508478	NM_000267.3:c.625C>T NP_000258.1:p.Gln209Ter
PC96: Astrocytoma	*PMS2*(c.863C>T; p.Ala288Val)	Missense	7:6027128-6027128	NM_001322003.2:c.863C>T NP_001308932.1:p.Ala288Val
	NF1(c.3169G>A; p.Ala1057Thr)	Missense	17:29557915-29557915	NM_001042492.3:c.3169G>A NP_001035957.1:p.Ala1057Thr
	POLE(c.3311C>T; p.Thr1104Met	Missense	12:133234521-133234521	NM_006231.4:c.3311C>T NP_006222.2:p.Thr1104Met
PC86: Craniopharyngioma	*MSH6*(c.3180A>C; p.Leu1060Phe)	Missense	2:48030566-48030566	NM_000179.3:c.3180A>C NP_000170.1:p.Leu1060Phe
PC139: Brain tumour (unspecified)	*MLH1*(c.1013A>G; p.Asn338Ser)	Missense	3:37061929-37061929	NM_000249.4:c.1013A>G NP_001341559.1:p.Asn338Ser
PC107: Astrocytoma	*PMS1*(c.79G>C; p.Glu27Gln)	Missense	2:190656614-190656614	NM_000534.5:c.79G>C NP_000525.1:p.Glu27Gln
PC137: Optic nerve glioma	*IDH1*(c.964G>A; p.Gly322Arg)	Missense	2:209104614-209104614	NM_001282387.1:c.964G>A NP_001269316.1:p.Gly322Arg
PC89: Astrocytoma	*PMS1*(c.79G>C; p.Glu27Gln)	Missense	2:190656614-190656614	NM_001128143.2:c.79G>C NP_001121615.1:p.Glu27Gln

3.5.1 Pathogenic variants

3.5.1.1 PC97: Medulloblastoma

Two variants in the *ATM* gene (OMIM 607585) were identified in PC97, a compound heterozygote female diagnosed with medulloblastoma (Age of onset: 4 years, 11 months) (Table 4). The first was a frameshift variant, NM_000051.4:c.2200dup (NP_000042.3:p.Val734GlyfsTer4), and the second was a nonsense variant, NM_000051.4:c.8911C>T (NP_000042.3:p.Gln2971Ter). Both variants were classified as pathogenic according to ACMG-AMP guidelines based on the following evidence: the c.2200dup variant, located in exon 14, and the c.8911C>T variant, located in exon 62, create premature translational stop signals which disrupt protein function through nonsense-mediated decay or truncation of the ATM protein (PSV1_Very strong); both variants were absent in gnomAD control data (PM2_Moderate); and reputable sources have reported the variants as pathogenic or likely

pathogenic, with one ClinVar submission for the c.2200dup variant (ClinVar ID: 2128482), and six ClinVar submissions for the c.8911C>T variant (ClinVar ID: 582761) (PP5_Supporting).

3.5.1.2 PC111: Optic nerve glioma

A heterozygous nonsense variant was identified in the *NF1* gene (OMIM 613113), NM_000267.3:c.625C>T (NP_000258.1:p.Gln209Ter) in PC111, a female diagnosed with optic nerve glioma (Age of onset: 1 year, 10 months) (Table 4). It was classified as pathogenic according to ACMG-AMP guidelines based on the following evidence: The c.625C>T variant , located in exon 6 of the *NF1* gene, is a substitution that changes the amino acid from a glutamine to a stop codon within a coding exon, resulting in an absent or disrupted protein product (PSV1_Very strong); the variant was absent in gnomAD control data (PM2_Moderate); and reputable sources have identified the variant as pathogenic, as shown by 6 ClinVar submissions (ClinVar ID: 187069) (PP5_Supporting).

3.5.2 Variants of uncertain significance

Eight VUSs were identified in genes which are integral to DNA repair pathways, tumour suppression, and genomic stability (Church et al., 2013; DeRycke et al., 2017; Weber et al., 2023). These variants were detected in six patients with different CNS tumour subtypes, including astrocytoma (PC96, 107, and 89), craniopharyngioma (PC86), optic nerve glioma (PC137), and an unspecified brain tumour (PC139) (Table 5). To assess their potential impact on protein structure and function, the variants underwent *in-silico* analysis using computational tools such as CADD, SIFT, AlphaMissense, and FATHMM-MLK. Details of the *in-silico* scores, and other supporting evidence, including ClinVar IDs, and gnomAD African/African American allele frequencies, is provided in Supplementary Table 3 (Appendix A).

Conflicting classification criteria or insufficient evidence to determine the role of these variants in CNS tumours led to their classification as VUSs. Among these, five VUSs were considered to be "hot", as they had substantial supporting evidence, suggesting that they could potentially be reclassified as likely pathogenic or pathogenic if further data becomes available. In contrast, three "cold" variants were identified, supported by benign evidence, as indicated by BP4 and BP6 criteria (Table 5). The currently available data is insufficient to fully elucidate the significance of these variants in CNS cancer, underscoring the need for further investigation and validation.

Table 4. Results following the classification of shortlisted variants according to the ACMG-AMP guidelines

Patient ID and type of CNS cancer	Gene and variant	Read depth	Classification criteria and supporting evidence	ACMG-AMP classification
PC97: Medulloblastoma	*ATM*(c.2200dup; p.Val734GlyfsTer4)	188	PSV1_Very strong: Frameshift variant in *ATM* gene (LOF is a known mechanism of disease); PM2_Moderate: Variant is absent on gnomAD; PP5_Supporting: Multiple ClinVar submissions suggesting pathogenicity	Pathogenic
	ATM(c.8911C>T; p.Gln2971Ter)	119	PSV1_Very strong: Nonsense variant in *ATM* gene (LOF is a known mechanism of disease); PM2_Moderate: Variant is absent on gnomAD; PP5_Supporting: Multiple ClinVar submissions suggesting pathogenicity	Pathogenic
PC111: Optic nerve glioma	*NF1*(c.625C>T; p.Gln209Ter)	38	PSV1_Very strong: Nonsense variant in *NF1* gene (LOF is a known mechanism of disease); PM2_Moderate: Variant is absent on gnomAD; PP5_Supporting: Multiple ClinVar submissions suggesting pathogenicity	Pathogenic
PC96: Astrocytoma	*PMS2*(c.863C>T; p.Ala288Val)	82	PM1_Moderate: Located in a mutational hotspot (Varsome); PM2_Moderate: gnomAD AFR/AA AF = 0.000006431; PP3_Strong: Multiple *in-silico* prediction tools suggest pathogenicity	Uncertain significance*
	NF1(c.3169G>A; p.Ala1057Thr)	79	PM1_Moderate: Located in a mutational hotspot (Varsome); PM2_Moderate: gnomAD AFR/AA AF = 0.0000619, PP2_Supporting: *NF1* gene has a low rate of benign missense variation; BP4_Supporting: Multiple *in-silico* prediction tools suggest a benign effect	Uncertain significance
	POLE(c.3311C>T; p.Thr1104Met)	69	PM1_Moderate: Located in a mutational hotspot (Varsome); PM2_Moderate: gnomAD AFR/AA AF = 0.000185; BP4_Supporting: Multiple *in-silico* prediction tools suggest a benign effect	Uncertain significance
PC86: Craniopharyngioma	*MSH6*(c.3180A>C; p.Leu1060Phe)	192	PM1_Moderate: Located in a mutational hotspot (Varsome); PM2_Moderate: Variant is absent on gnomAD; PP3_Strong: Multiple *in-silico* prediction tools suggest pathogenicity	Uncertain significance*
PC139: Brain tumour (unspecified)	*MLH1*(c.1013A>G; p.Asn338Ser)	71	PM1_Moderate: Located in a mutational hotspot (Varsome); PM2_Moderate: gnomAD AFR/AA AF = 0.0000615; PP3_Strong: Multiple *in-silico* prediction tools suggest pathogenicity; BP6_Supporting: Multiple ClinVar submissions suggesting variant is benign	Uncertain significance

PC107: Astrocytoma	*PMS1*(c.79G>C; p.Glu27Gln)	97	PM2_Moderate: gnomAD AFR/AA AF = 0.000984; PP3_Strong: Multiple *in-silico* prediction tools suggest pathogenicity	Uncertain significance*
PC137: Optic nerve glioma	*IDH1*(c.964G>A; p.Gly322Arg)	117	PM2_Moderate: Variant is absent on gnomAD; PP3_Strong: Multiple *in-silico* prediction tools suggest pathogenicity	Uncertain significance*
PC89: Astrocytoma	*PMS1*(c.79G>C; p.Glu27Gln)	108	PM2_Moderate: gnomAD AFR/AA AF = 0.000984; PP3_Strong: Multiple *in-silico* prediction tools suggest pathogenicity	Uncertain significance*

Abbreviations: ACMG-AMP- American College of Medical Genetics and Genomics, and Association for Molecular Pathology, LOF- loss-of-function, gnomAD- Genome Aggregation Database, AFR/AA AF- African/African American allele frequency. "Hot" variants of uncertain significance are indicated with stars (*).

4. Discussion

A thorough investigative approach was utilised to explore the germline genetic landscape of paediatric CNS tumours. The aim of this study was to identify germline pathogenic variants in 56 CPGs. Using pre-generated WES data from thirteen paediatric patients with CNS tumours, a candidate-gene-based strategy was employed to identify potential pathogenic and likely pathogenic variants. The analysis focused on candidate genes associated with both germline inheritance and CPS's, facilitating the investigation of a wide spectrum of genes associated with different types of CNS tumours. This approach initially yielded numerous variants, which were refined through rigorous filtering to a final set of eleven candidate variants in eight patients.

In this study, 3 pathogenic variants and eight VUSs were identified. Germline pathogenic variants were identified in 2 out of 13 patients with CNS tumours (15,38%), while six patients (46,15%) carried VUSs. Germline pathogenic variants were associated with medulloblastoma and optic nerve glioma, whereas astrocytoma cases exhibited the highest prevalence of VUSs. Individuals harboured variants in tumour suppressor genes such as *ATM*, *NF1*, and *POLE*, as well as DNA MMR genes including *MLH1*, *MSH6*, *PMS1*, and *PMS2*, and *IDH1.* Tumour suppressor genes regulate cell growth and prevent uncontrolled division, DNA MMR genes maintain genomic stability by correcting replication errors, and *IDH* genes play a role in cellular metabolism (Dahlin et al., 2020; Guerrini-Rousseau et al., 2024; Kim et al., 2021; Kline et al., 2017). Variants in genes that impair the regulation of cell proliferation are often central to the pathogenesis of CNS tumours and significantly increase the risk of cancer development (Parsons et al., 2016). All of the genes in which variants were identified have been previously linked to CNS tumours and other adult-onset cancers, including breast, prostate, endometrial, colorectal, gastric, thyroid, and leukaemia (Church et al., 2013; Dahlin et al., 2020; DeRycke et al., 2017; Gerges et al., 2013; Guerrini-Rousseau et al., 2024; Packer et al., 2020; Huang et al., 2013; Kim et al., 2021; Wang et al., 2021).

The comparison of positive pick-up rates among various studies highlights important insights into the genetic landscape of solid tumours. The positive pick-up rate of this study was 15,38%, in-keeping with

both South African, and global published literature. In another South African study by Manolas (2023), the diagnostic yield was approximately 9,4% in a cohort of 32 paediatric solid tumour patients. Internationally, research by Parsons et al. (2016) and Akhavanfard et al. (2020) identified pathogenic germline variants in 10% of their respective cohorts. These findings indicate that diagnostic yields differ slightly among populations, they reflect a consistent trend in identifying germline variants patients with solid tumours. Further research is needed to explore and compare these results with similar studies, locally and internationally, to enhance the understanding of germline pathogenic variants in paediatric CNS tumorigenesis.

4.1 PC97: Medulloblastoma

Two pathogenic variants, c.2200dup and c.8911C>T, were identified in the *ATM* gene in PC97, a female who is a compound heterozygote, and potentially affected with Ataxia-telangiectasia (AT). Biallelic mutations in the *ATM* gene are responsible for this disorder, which significantly increases the risk of tumorigenesis, including breast, ovarian, prostate, pancreatic, gastric, thyroid, and CNS tumours, in affected individuals (Hall et al., 2021; Mehrdad and Parvin, 2018; Stucci et al., 2021). This autosomal recessive neurodegenerative disorder is characterised by cerebellar ataxia, ocular telangiectasias, premature ageing, and immunodeficiency (Hall et al., 2021). Individuals with variants in *ATM* are at higher risk of developing cancer at an earlier stage in life compared to those without these mutations. These mutations confer a predisposition to cancer, with an estimated lifetime risk of 30–40% (Liberzon et al., 2003). Heterozygous carriers of *ATM* pathogenic variants have elevated cancer risks, including a 2- to 5-fold increase in adult-onset breast cancer risk compared to non-carriers. Additionally, moderate to high risks have been observed for pancreatic cancer (over 4-fold), gastric cancer (3-fold), and prostate cancer (2- to 3-fold). Modest yet significant elevations in risks for other cancers, such as colorectal cancer (1.5-fold), have also been reported (Hall et al., 2021; Mehrdad and Parvin, 2018; Stucci et al., 2021).

The *ATM* gene consists of 66 exons and is located on chromosome 11 (11q22-23) (Phan and Rezaeian, 2021). Upon its autophosphorylation, the ATM protein acts as an initiator and coordinator of the DNA damage response (Liberzon et al., 2003). ATM is a multifunctional protein kinase that is pivotal in regulating cell cycle progression, DNA damage response, double-strand break repair mechanisms, and cellular apoptosis by orchestrating the phosphorylation of various substrates (Stucci et al., 2021). To date, 739 disease-causing mutations in the *ATM* gene have been identified, distributed across all exons without evident hotspots. Approximately 70% of these mutations are frameshift or nonsense variants, which lead to the truncation of the ATM protein (Mehrdad and Parvin, 2018). This disorder highlights

the importance of the tumour suppressor gene, *ATM* in maintaining genomic stability and underscores the potential implications of *ATM* mutations in cancer susceptibility.

Genomic integrity and stability are critical in the brain, as neurons are terminally differentiated and lack the ability to divide (Sait et al., 2021). To protect against DNA damage, neurons have developed multiple overlapping mechanisms for repair. Disruptions in these processes can result in CNS defects, as evidenced by mutations in DNA repair genes that often lead to syndromes with pronounced neurological abnormalities (Mehrdad and Parvin, 2018). For example, AT, is characterised by progressive degeneration of the cerebellar cortex and is associated with CNS tumours, including medulloblastoma, astrocytoma, and gliomas. This underscores the essential role of DNA repair in the health and survival of neural cells (Liberzon et al., 2003; Stucci et al., 2021).

The *ATM*:c.2200dup variant is a duplication at nucleotide position 2200 in the coding sequence, affecting amino acid residue 734, which is located in the N- terminus of the ATM protein (UniProt ID: Q13315). This region spans from amino acid position one to position 1373, lies upstream of ATM's major functional domains, and does not belong to a specifically named functional domain. The duplication introduces a frameshift mutation, resulting in an altered reading frame and a premature stop codon downstream. This truncates the protein, likely preventing the translation of downstream critical domains essential for ATM's functionality in DNA damage signalling and repair.

The N-terminal region of ATM is composed of several α-helical Huntingtin, Elongation factor 3, protein phosphatase 2A, and TOR1 (HEAT) repeat motifs, that act as scaffolds, enabling interactions with both proteins and DNA. This region is critical for binding substrates and cofactors, including *Nibrin* (*NBN*), *Tumour protein P53* (*TP53*), and *Breast cancer 1* (*BRCA1*) (Stucci et al., 2021). These motifs are integral to ATM's ability to recruit and bind various proteins at DNA damage sites, a vital step in orchestrating effective DNA repair processes (Phan and Rezaeian, 2021). Additionally, this region includes a leucine zipper motif that is crucial for ATM dimerisation and interactions with other proteins (Phan and Rezaeian, 2021; Stucci et al., 2021).

Pathogenic truncating variants in the N-terminal region can disrupt these interactions, impairing ATM's ability to maintaining genomic stability. This could lead to an accumulation of DNA damage and mutations that contribute to cancer development. *ATM* is a tumour suppressor gene, therefore LOF variants, such as those causing truncation or impairing its protein interactions, can disrupt its protective role against tumorigenesis. In the absence of functional ATM activity, cells may inadequately respond to oncogenic stress, resulting in uncontrolled cell proliferation and tumorigenesis (Phan and Rezaeian, 2021).

The *ATM*:c.8911C>T variant is a is a nucleotide change at position 8911 in the coding sequence of the *ATM* gene, corresponding to amino acid position 2971. This position falls within the phosphatidylinositol 3-/4-kinases catalytic domain of the ATM protein, which spans residues 2686 to 2998 (UniProt ID: Q13315). This variant results in a premature stop codon (p.Gln2971Ter), which likely leads to a truncated protein that cannot perform its critical role in maintaining genomic stability. Genetic variants in this domain are understudied, resulting in a lack of research into the disease mechanism of this specific variant.

4.2 PC111: Optic nerve glioma

A pathogenic variant (c.625C>T) was identified in the *NF1* gene in a female patient with optic nerve glioma, PC111, which is associated with Neurofibromatosis type 1. Neurofibromatosis type 1 is an autosomal dominant condition, characterised by café-au-lait spots, Lisch nodules, neurofibromas, and an increased risk of developing optic pathway gliomas, astrocytoma, and other CNS tumours, breast, ovarian, and colorectal cancer, and leukaemia (Packer et al., 2020; Wang et al., 2021). Approximately 15% to 20% of individuals with NF1 develop optic pathway gliomas and astrocytoma (Packer et al., 2020).

The *NF1* gene consists of 57 exons and is located on chromosome 17 (17q11.2) (Mo et al., 2022). *NF1* functions as a tumour suppressor gene, and encodes neurofibromin, which negatively regulates the *Rat sarcoma* (*RAS*) proto-oncogene through its RAS GTP-ase activating activity. Loss-of-function mutations in *NF1* lead to overactivation of RAS signalling, promoting cell growth and increasing the risk of both benign and malignant tumours (Xu et al., 2018). The *NF1* gene exhibits significant variability, with over 3,000 unique mutations reported (Wang et al., 2021). This high variability is compounded by the gene's lack of mutational hotspots, the presence of pseudogenes, and its broad mutation spectrum. The overall pathogenic mutation rate in *NF1* is approximately 92% (Wang et al., 2021). Roughly half (55%) of these pathogenic variants result in a truncated neurofibromin protein. Of these, 13% are nonsense variants, 42% are frameshift variants, and 10% are missense variants, further highlighting the gene's mutational diversity (Okumura, Ozaki and Niida, 2014).

The *NF1*:c.625C>T variant is a is a single nucleotide substitution at position 625 in the coding sequence of the *NF1* gene, corresponding to amino acid position 209. This position lies within the N-terminal region of the neurofibromin protein, which spans residues one to 1198 (UniProt ID: P21359). This region plays a role in stabilising the protein structure and may contribute to interactions with other cellular components. Variants in this part of the protein could disrupt neurofibromin's ability to regulate RAS signalling indirectly, potentially leading to aberrant cell proliferation and tumorigenesis associated with NF1 (Mo et al., 2022). Amino acid position 209 is not part of a specific functional

domain, but it lies next to the cysteine- and serine-rich domain (CSRD), located between amino acids 543 and 909, and encoded by exons 14 to 21. This domain enhances the association of neurofibromin with actin following phosphorylation (Mo et al., 2022). Individuals carrying a variant in this domain are at an increased risk of developing optic glioma (Xu et al., 2018). One publication that documents the *NF1*:c.625C>T variant, is a case report which described a patient with the nonsense variant (Okumura, Ozaki and Niida, 2014). Her mother was also a confirmed carrier of the variant, without evidence of somatic mosaicism (Okumura, Ozaki and Niida, 2014). Further functional assays are required to validate the specific impact of this variant.

4.3 PC96, 86, 139, 107, 137, and 89

Notably, 5 of the 11 candidate variants identified (45,45%) were in DNA MMR genes. Additionally, VUSs in MMR genes were observed in 5 out of 13 patients, accounting for 38,46% of the cohort. The MMR genes *MLH1, MSH6, PMS1,* and *PMS2* are associated with hereditary cancer predisposition syndromes that elevate the risk of colorectal, endometrial, and CNS tumours, including Constitutional mismatch repair deficiency, Lynch, and Turcot (Type 1) syndrome (DeRycke et al., 2017; Gerges et al., 2013; Guerrini-Rousseau et al., 2024; Kim et al., 2022). Variants in *IDH1* drive gliomas and secondary glioblastomas, while *POLE* is associated with Polymerase proofreading-associated polyposis syndrome and malignancies, including colorectal, endometrial, ovarian and CNS cancers such as astrocytoma, glioblastoma, and medulloblastoma (Church et al., 2013; DeRycke et al., 2017; Weber et al., 2023). This underscores the role of these genes in both inherited and sporadic cancer predisposition.

Variants in genes with no reported association to the diagnosis of the patient were identified, such as *MSH6* and *IDH1* variants in patients with craniopharyngioma and optic nerve glioma, respectively. Although *MSH6* and *IDH1* variants have not been reported in the literature as being specifically associated with craniopharyngioma and optic nerve glioma, these genes have been previously implicated in the pathogenesis of CNS tumours (DeRycke et al., 2017; Gerges et al., 2013; Guerrini-Rousseau et al., 2024). These findings highlight the need for further investigation to determine the potential roles of *MSH6* and *IDH1* in the pathogenesis of these tumour types.

Two of the five patients with astrocytoma (PC89 and 107) carried VUSs in *PMS1*(c.79G>C), and one patient (PC96) carried a VUS in PMS2(c.863C>T), associated with hereditary colorectal cancer. There are no previous reports of *PMS2*(c.863C>T; p.Ala288Val), *NF1* (c.3169G>A; p.Ala1057Thr), *IDH1*(c.964G>A; p.Gly322Arg), *MLH1*(c.1013A>G; p.Asn338Ser) and *MSH6*(c.3180A>C; p.Leu1060Phe). These variants have not been previously associated with the individuals' cancer phenotypes and therefore require further investigation. The c.79G>C variant in *PMS1*, a colorectal CPG has been reported as a VUS in literature once and is associated with Lynch syndrome (DeRycke et al.,

2017). The *POLE* variant c.3311C>T (p.Thr1104Met) has been documented in the context of endometrial cancer by Church et al. (2013).

4.4 Limitations of the study

This study was limited to a pre-defined set of candidate genes, which, while comprehensive, may have excluded other relevant genes implicated in the pathogenesis of paediatric CNS tumours. Therefore, it is possible that there are germline variants in other genes not analysed in this study, as well as epigenetic alterations, or copy number variants (CNV) underlying the different CNS tumours. The lack of comprehensive knowledge regarding the genetic mechanisms underlying paediatric CNS tumours limits the scope of variant interpretation, particularly when interpreting VUSs. Without functional studies or additional evidence, the classification of VUSs remains challenging, potentially leaving clinically relevant variants unrecognised. Another challenge in classifying VUSs is the underrepresentation of African populations in genomic databases like gnomAD, leading to limited allele frequency data for accurate variant interpretation. Since this cohort comprised African patients, this underrepresentation likely contributed to the high number of VUSs observed. These patients were singletons, therefore ACMG-AMP codes regarding *de novo* occurrences (PM6 or PM2) and segregation (PP1) were not applied, as there was no information about inheritance patterns and variant co-segregation.

4.5 Future work

Future studies are essential to validate the findings of this research, and to extend the scope of investigation. All candidate variants identified through WES data analysis should be confirmed using Sanger sequencing. Re-analysis of the WES data may be warranted as new genomic data becomes available and VUSs may be reclassified. Alternative candidate genes may be analysed for patients where no putative pathogenic variants were identified. Incorporating complementary approaches, such as CNV analysis, epigenetic alterations could uncover structural variants that might contribute to the germline genetic landscape of paediatric CNS tumours. Future studies should prioritise comprehensive genomic analyses, experimental validation, and the inclusion of diverse populations to address the limitations of this study.

5. Conclusion

This study investigated the spectrum of germline variants in CPGs, encompassing both genes associated with germline inheritance and those related to CPS's. The positive pick-up rate of this study was 15,38%, which is in-keeping with international published oncology research. Pathogenic variants

were identified in two individuals, in 2 of the 56 genes examined. Germline pathogenic variants were identified in medulloblastoma and optic nerve glioma patients, while astrocytoma cases exhibited the highest number of VUSs. Innovative genetic technologies such as WES and comprehensive bioinformatics analyses facilitate the identification of genetic variants involved in tumour development. This study highlights the challenges associated with interpreting VUSs in underrepresented populations, including those of South African descent. This study contributes to a broader understanding of the germline genetic landscape of paediatric CNS tumours in South Africa and addresses a gap in global oncology research.

6. Declarations

Statement of Ethics

This sub-study protocol was reviewed and approved by the Medical Research Ethics Committee of the University of the Witwatersrand (approval number: M240758), under the parent study (approval number: M240262). The ethical clearance certificate is provided in Appendix B.

Data Availability Statement

Data available upon request.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Funding Sources

This study was supported by the National Research Foundation, the South African Medical Research Council, and the Female Academic Leaders Fellowship. The funders had no role in the design, data collection, data analysis, and reporting of this study.

Acknowledgement

Drs K Mnika and L Lamola assisted C Chatergoon with data analysis. Drs. J Geel , G Naidu , and N Mabaso aided in patient recruitment.

Author Contributions

L Lamola designed the study. C Chatergoon, K Mnika, and L Lamola contributed to the acquisition and interpretation of the data. C Chatergoon analysed the data and drafted the manuscript. C Chatergoon, M Mnika, and L Lamola revised the manuscript for important intellectual content. M Mnika and L Lamola supervised the research. All authors approved the final version of the manuscript.

References

Alonso-Luna, O. Mercado-Celis, G.E., Melendez-Zajgla, J., Zapata-Tarres, M. and Mendoza, C.E. (2023) 'The genetic era of childhood cancer: Identification of high-risk patients and germline sequencing approaches,' *Annals of Human Genetics*, 87(3), p. 81–90. https://doi.org/10.1111/ahg.12502.

Akhavanfard, S. Padmanabhan, R., Yehia, L., Cheng, F., and Eng, C. (2020) 'Comprehensive germline genomic profiles of children, adolescents and young adults with solid tumors,' *Nature Communications*, 11(1). https://doi.org/10.1038/s41467-020-16067-1.

Capasso, M., Montella, A., Tirelli, M., Maiorino, T., Cantalupo, S. and Iolascon, A. (2020) 'Genetic predisposition to solid pediatric cancers,' *Frontiers in Oncology*, *10*, p.590033. https://doi.org/10.3389/fonc.2020.590033.

Cheng, J., Novati, G., Pan, J., Bycroft, C., Žemgulytė, A., Applebaum, T., et al. (2023) 'Accurate proteome-wide missense variant effect prediction with AlphaMissense,' *Science*, 381(6664). https://doi.org/10.1126/science.adg7492.

Church, D.N., Briggs, S.E., Palles, C., Domingo, E., Kearsey, S.J., Grimes, et al. (2013) 'DNA polymerase ε and δ exonuclease domain mutations in endometrial cancer,' *Human Molecular Genetics*, 22(14), p. 2820–2828. https://doi.org/10.1093/hmg/ddt131.

Dahlin, A.M., Wibom, C., Andersson, U., Bybjerg-Grauholm, J., Deltour, I., Hougaard, D., *et al.* (2020) 'A genome-wide association study on medulloblastoma,' *Journal of Neuro-Oncology*, 147(2), p. 309–315. https://doi.org/10.1007/s11060-020-03424-9.

DeRycke, M.S. Gunawardena, S., Balcom, J.R., Pickart, A.M., Waltman, L.A., French, A.J., et al. (2017) 'Targeted sequencing of 36 known or putative colorectal cancer susceptibility genes,' *Molecular Genetics & Genomic Medicine*, 5(5), p. 553–569. https://doi.org/10.1002/mgg3.317.

Dunnen, J.T.D., Dalgleish, R., Donna, R., Maglott, K., Hart, R.K., Greenblatt, M.S., McGowan, J., et al. (2016) 'HGVS Recommendations for the description of sequence Variants: 2016 update,' *Human Mutation*, 37(6), p. 564–569. https://doi.org/10.1002/humu.22981.

Feh, M.K.-N.M., Lyon, K.A., Brahmaroutu, A.V., Tadipatri, R., and Fonkem, E. (2021) 'The need for a central brain tumor registry in Africa: A review of central nervous system tumors in Africa from 1960 to 2017,' *Neuro-Oncology Practice*, 8(3), p. 337–344. https://doi.org/10.1093/nop/npaa086.

Gerges, N., Fontebasso, A.M., Albrecht, S., Faury, D., and Jabado, N. (2013) 'Pediatric high-grade astrocytomas: a distinct neuro-oncological paradigm,' *Genome Medicine*, 5(7), p. 66. https://doi.org/10.1186/gm470.

Gröbner, S.N., Worst, B.C., Weischenfeldt, J., Buchhalter, I., Kleinheinz, K., Rudneva K., et al. (2018). 'The landscape of genomic alterations across childhood cancers', Nature, 555(7696), p.321-327. doi: https://doi.org/10.1038/nature25480.

Guerrini-Rousseau, L., Merlevede, J., Denizeau, P., Andreiuolo, F., Varlet, P., Puget, S., et al. (2024) 'Glioma oncogenesis in the constitutional mismatch repair deficiency (CMMRD) syndrome,' *Neuro-Oncology Advances*, 6(1). https://doi.org/10.1093/noajnl/vdae120.

Gump, J.M., Donson, A., Nirks, D., Amani, M., Rao, K., Ringer, M., et al. (2015) 'Identification of targets for rational pharmacological therapy in childhood craniopharyngioma,' *Acta Neuropathologica Communications*, 3(1). https://doi.org/10.1186/s40478-015-0211-5.

Hall, M.J. Bernhisel, R., Hughes, E., Larson, K., Rosenthal, E.T., Singh, N.A., et al. (2021) 'Germline Pathogenic Variants in the Ataxia Telangiectasia Mutated (ATM) Gene are Associated with High and Moderate Risks for Multiple Cancers,' *Cancer Prevention Research*, 14(4), p. 433–440. https://doi.org/10.1158/1940-6207.capr-20-0448.

Huang, L.M. (2014) 'Association Between ERCC2 Polymorphisms and Glioma Risk: a Meta-analysis,' *Asian Pacific Journal of Cancer Prevention*, 15(11), p. 4417–4422. https://doi.org/10.7314/apjcp.2014.15.11.4417.

Karczewski, K.J., Francioli, L.C., Tiao, G., Cummings, B.B., Alföldi, J., Wang, Q., et al., 2020. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*, 581(7809), p.434-443. Available at: https://doi.org/10.1038/s41586-020-2308-7.

Kim, H. Lim, K.Y., Park, J.W., Kang, J., Won, J.K., Lee, K., et al. (2021) 'Sporadic and Lynch syndrome-associated mismatch repair-deficient brain tumors,' *Laboratory Investigation*, 102(2), p. 160–171. https://doi.org/10.1038/s41374-021-00694-3.

Kline C.N., Joseph N.M., Grenert J.P., van Ziffle J., Talevich E., Onodera C., et al. (2017) 'Targeted next-generation sequencing of pediatric neuro-oncology patients improves diagnosis, identifies pathogenic germline mutations, and directs targeted therapy,' *Neuro-Oncology*, 19(5), p.699-709. https://doi.org/10.1093/neuonc/now254.

Kumar, P., Henikoff, S. and Ng, P.C. (2009) 'Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm,' *Nature Protocols*, 4(7), p. 1073–1081. https://doi.org/10.1038/nprot.2009.86.

Li, B.K. Vasiljevic, A., Dufour, C., and Yao, F., (2019) 'Pineoblastoma segregates into molecular sub-groups with distinct clinico-pathologic features: a Rare Brain Tumor Consortium registry study,' *Acta Neuropathologica*, 139(2), p. 223–241. https://doi.org/10.1007/s00401-019-02111-y.

Liberzon, E. Avigad, S., Cohen, I.J., Yaniv, I., Michovitz, S., and Zaizov, R. (2003) 'ATM gene mutations are not involved in medulloblastoma in children,' *Cancer Genetics and Cytogenetics*, 146(2), p. 167–169. https://doi.org/10.1016/s0165-4608(03)00130-4.

Manolas, E. (2023). Mutation Profiling of Paediatric Solid Tumours in a Cohort of South African Patients. [Dissertation]. [Johannesburg (South Africa)]: the University of the Witwatersrand.

McLaren, W., Gil, L., Hunt, S.E., Riat, H.S., Ritchie, G.R.S., Thormann, A., Flicek, P. and Cunningham, F. (2016). 'The Ensembl Variant Effect Predictor', *Genome Biology*, 17(1), p.1-14. https://doi.org/10.1186/s13059-016-0974-4

Mehrdad, A.E. and Parvin, M. (2018) 'ATM in breast and brain tumors: a comprehensive review,' *Cancer Biology and Medicine*, 15(3), p. 210. https://doi.org/10.20892/j.issn.2095-3941.2018.0022.

Mo, J. Moye, S.L., McKay , R.M., Le, L.Q. (2022) 'Neurofibromin and suppression of tumorigenesis: beyond the GAP,' *Oncogene*, 41(9), p. 1235–1251. https://doi.org/10.1038/s41388-021-02156-y.

Okumura, A., Ozaki, M. and Niida, Y. (2014) 'Development of a practical NF1 genetic testing method through the pilot analysis of five Japanese families with neurofibromatosis type 1,' *Brain and Development*, 37(7), p. 677–689. https://doi.org/10.1016/j.braindev.2014.11.002.

Packer, R.J Iavarone, A., Jones, D.T., Blakeley, J.O., Bouffet, E., Fisher, M.J. et al. (2020) 'Implications of new understandings of gliomas in children and adults with NF1: report of a consensus conference,' *Neuro-Oncology*, 22(6), p. 773–784. https://doi.org/10.1093/neuonc/noaa036.

Parsons, D.W. A., Yang, Y., Wang, T., Scollon, S., Bergstrom, K., Kerstein, R.A., et al. (2016) 'Diagnostic yield of clinical tumor and germline Whole-Exome sequencing for children with solid tumors,' *JAMA Oncology*, 2(5), p. 616-624. https://doi.org/10.1001/jamaoncol.2015.5699.

Phan, L.M. and Rezaeian, A.-H. (2021) 'ATM: main features, signaling pathways, and its diverse roles in DNA damage response, tumor suppression, and cancer development,' *Genes*, 12(6), p. 845. https://doi.org/10.3390/genes12060845.

Raja, N., Hayes, L., Basta, N., and McNally, R.J.Q. (2021) 'International trends in the incidence of brain tumours in children and young-adults and their association with indicators of economic development,' *Cancer Epidemiology*, 74, p. 102006. https://doi.org/10.1016/j.canep.2021.102006.

Rentzsch, P., Witten, D., Cooper, G. M., Shendure, J., and Kircher, M. (2019) 'CADD: predicting the deleteriousness of variants throughout the human genome,' *Nucleic Acids Research*, 47(D1), p. D886–D894. https://doi.org/10.1093/nar/gky1016.

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., et al. (2015) 'Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology,' *Genetics in Medicine*, 17(5), p. 405–424. https://doi.org/10.1038/gim.2015.30.

Robinson, J.T. Thorvaldsdottir, H., Turner, D. and Mesirov, J.P (2022) 'igv.js: an embeddable JavaScript implementation of the Integrative Genomics Viewer (IGV),' *Bioinformatics*, 39(1). https://doi.org/10.1093/bioinformatics/btac830.

Rotimi, S.O., Rotimi, O.A. and Salhia, B. (2021) 'A review of Cancer Genetics and Genomics studies in Africa,' *Frontiers in Oncology*, 10. https://doi.org/10.3389/fonc.2020.606400.

Sait, S.F., Walsh, M.F., and Karajannis, M.A. (2021) 'Genetic syndromes predisposing to pediatric brain tumors,' *Neuro-Oncology Practice*, 8(4), p. 375–390. https://doi.org/10.1093/nop/npab012.

Shihab, H.A., Rogers, M.F., Gough, J., Mort, M., Cooper, D.N., Day, I.N., et al. (2015) 'An integrative approach to predicting the functional effects of non-coding and coding sequence variation,' *Bioinformatics*, 31(10), p. 1536–1543. https://doi.org/10.1093/bioinformatics/btv009.

Stucci, L.S., Internò V., Tucci, M., Perrone, M., Mannavola, F., Palmirotta, R., et al. (2021) 'The ATM gene in Breast Cancer: Its relevance in clinical practice,' *Genes*, 12(5), p. 727. https://doi.org/10.3390/genes12050727.

VistaSeq Brain/CNS/PNS Cancer Panel - Clinical test - NIH Genetic Testing Registry (GTR) - NCBI (2024). https://www.ncbi.nlm.nih.gov/gtr/tests/GTR000556512/.

Wang, W. Wei, C.J, Cui, X.W., Li, Y.H., Gu, Y.H., Gu, B., et al. (2021) 'Impacts of NF1 gene mutations and genetic modifiers in neurofibromatosis type 1,' *Frontiers in Neurology*, 12. https://doi.org/10.3389/fneur.2021.704639.

Weber, C.A., Krönke, N., Volk, V., Auber, B., Förster, A., Trost, D., et al. (2023) 'Rare germline variants in POLE and POLD1 encoding the catalytic subunits of DNA polymerases ε and δ in glioma families,' *Acta Neuropathologica Communications*, 11(1). P.184. https://doi.org/10.1186/s40478-023-01689-5.

Xu, M. Xiong, H., Han, Y., Li, C., Mai, S., Huang, Z., et al. (2018) 'Identification of mutation regions on NF1 responsible for High- and Low-Risk development of optic pathway glioma in neurofibromatosis type I,' *Frontiers in Genetics*, 9. https://doi.org/10.3389/fgene.2018.00270.

Zhao, M. Liu, Y., Ding, G., Qu, D., and Qu, H. (2021) 'Online database for brain cancer-implicated genes: exploring the subtype-specific mechanisms of brain cancer,' *BMC Genomics*, 22(1). p.458 https://doi.org/10.1186/s12864-021-07793-x.

Appendices

Appendix A: Supplementary Tables and Figures

Supplementary Table 1. List of candidate paediatric CNS tumour predisposition genes (n=56)

Supplementary Table 2. Details, supporting evidence and ACMG-AMP classification of pathogenic variants (n=3)

Supplementary Table 3. Details, supporting evidence and ACMG-AMP classification of variants of uncertain significance (n=8)

Supplementary Figure 1. Candidate variant visualisation using IGV

Appendix B: Ethics Clearance Certificate (M240758)

Appendix C: Ethics Clearance Certificate (M180885)

Appendix D: Ethics Clearance Certificate (M240262)

Appendix E: Approved Research Protocol

Appendix F: Turnitin Report and Plagiarism Declaration

Appendix G: Author Guidelines for Frontiers in Oncology

Appendix A: Supplementary Tables and Figures

Supplementary Table 1. List of candidate paediatric CNS tumour predisposition genes (n=56)

Type of CNS tumour	Gene	Cancer predisposition disorders and syndromes	Pathway/function	Reference
Medulloblastoma				
	PTCH1	Gorlin syndrome	SHH	Kline et al., 2017
	SUFU	Gorlin syndrome	SHH	Kline et al., 2017
	PTCH2	Gorlin syndrome	SHH	Kline et al., 2017
	SMO	Gorlin syndrome	SHH	Kline et al., 2017
	GLI1	Greig syndrome	SHH/WNT	Kline et al., 2017
	APC	Turcot syndrome Type 2/ Familial adenomatous polyposis	Tumour suppressor	Kline et al., 2017
	MLH1	Constitutional mismatch repair deficiency syndrome/ Lynch syndrome/ Turcot syndrome Type 1	DNA repair pathway	Kim et al., 2022
	MSH2	Constitutional mismatch repair deficiency syndrome/ Lynch syndrome/ Turcot syndrome Type 1	DNA repair pathway	Kim et al., 2022
	MSH6	Constitutional mismatch repair deficiency syndrome/ Lynch syndrome/ Turcot syndrome Type 1	DNA repair pathway	Kim et al., 2022
	PMS1	Lynch syndrome	DNA repair pathway	Kim et al., 2022
	PMS2	Turcot syndrome Type 1/ Lynch Syndrome	DNA repair pathway	Kline et al., 2017
	BRCA1	Hereditary breast and ovarian cancer/ Fanconi anaemia	SHH	Kline et al., 2017

	BRCA2	Hereditary breast and ovarian cancer/ Fanconi anaemia	SHH	Kline et al., 2017
	PALB2	Fanconi anaemia	Tumour suppressor	Kline et al., 2017
	TP53	Li-Fraumeni syndrome	Tumour suppressor	Kline et al., 2017
	ATM	Ataxia-telangiectasia	DNA repair pathway	Dahlin et al., 2020
	FMR1	Fragile X syndrome	SHH/WNT	Dahlin et al., 2020
	NBN	Nijmegen breakage syndrome	Damaged DNA repair	Dahlin et al., 2020
	NF2	Neurofibromatosis Type 2	Tumour suppressor	Kline et al., 2017
	BLM	Bloom's syndrome	Damaged DNA repair	Dahlin et al., 2020
	VHL	von Hippel–Lindau disease	Tumour suppressor	Parsons et al., 2016
	CREBBP	Rubinstein-Taybi syndrome	Cellular differentiation	Parsons et al., 2016
	EP300	Rubinstein-Taybi syndrome	Stem call differentiation	Parsons et al., 2016
Astrocytoma				
	CDKN2A	Familial melanoma-astrocytoma syndrome	Tumour suppressor	Kline et al., 2017
	CDKN2B	-	Tumour suppressor	Kline et al., 2017
	PTEN	Cowden syndrome	Tumour suppressor	Kline et al., 2017
	NF1	Neurofibromatosis Type 1	Tumour suppressor	Gerges et al., 2013
	TP53	Li-Fraumeni syndrome	Tumour suppressor	Kline et al., 2017
	TSC2	Tuberous sclerosis complex	Tumour suppressor	Parsons et al., 2016
	PTPN114	Noonan syndrome	Ras-MAPK signalling	Gerges et al., 2013
	NBN	Nijmegen breakage syndrome	Damaged DNA repair	Dahlin et al., 2020

	BRCA2	Fanconi anaemia	Tumour suppressor	Kline et al., 2017
	IDH1	-	Lipid metabolism	Gerges et al., 2013
	IDH2	-	Lipid metabolism	Gerges et al., 2013
	H3F3A	-	Histone modification	Gerges et al., 2013
	SETD2	-	Histone modification	Gerges et al., 2013
	PIK3CA	-	PI3K/AKT/mTOR	Gerges et al., 2013
	MYBL1	-	Regulator of male meiosis	Gerges et al., 2013
	FGFR1	Kallmann syndrome	Cell growth and division	Gerges et al., 2013
	POLE	Polymerase Proofreading-Associated Polyposis syndrome	DNA proofreading	Kline et al., 2017
	ERCC2	Xeroderma pigmentosum	Damaged DNA repair	Huang, 2014
Optic nerve glioma				
	NF1	Neurofibromatosis Type 1	Tumour suppressor	Kline et al., 2017
	BRAF	-	Cell growth signalling	Gerges et al., 2013
	FGFR1	Kallmann syndrome	Cell growth and division	Gerges et al., 2013
	CIC	-	Tumour suppressor	Kline et al., 2017
	KIAA1549	-	Ras-MAPK	Gerges et al., 2013
	KRAS	Noonan syndrome	Cell growth and division	Gerges et al., 2013
	PTPN11	Noonan syndrome	Cell growth and division	Kline et al., 2017

	RAF1	Noonan syndrome	Cell growth and division	Gerges et al., 2013
Pineoblastoma				
	DICER1	Dicer 1 syndrome	MicroRNA biogenesis	Li et al., 2020
	DROSHA	Hereditary breast and ovarian cancer	DNA repair pathway	Li et al., 2020
	TP53	Li-Fraumeni syndrome	Tumour suppressor	Li et al., 2020
	RB1	Hereditary retinoblastoma	Tumour suppressor	Li et al., 2020
	PNET	-	Tumour suppressor	Li et al., 2020
Craniopharyngioma				
	CTNNB1	-	WNT	Gump et al., 2015
	APC	FAP	Tumour suppressor	Kline et al., 2017
	EGFR	-	Cell growth and division	Kline et al., 2017
	KMT2D	Kabuki syndrome	Tumour suppressor	Gump et al., 2015
	PMS1	Lynch syndrome	DNA repair pathway	Kim et al., 2022
	PMS2	Turcot syndrome Type 1/ Lynch syndrome	DNA repair pathway	Kim et al., 2022
	MLH1	Constitutional mismatch repair deficiency syndrome/ Lynch syndrome/ Turcot syndrome Type 1	DNA repair pathway	Kim et al., 2022
	MSH2	Constitutional mismatch repair deficiency syndrome/ Lynch syndrome/ Turcot syndrome Type 1	DNA repair pathway	Kim et al., 2022
	MSH6	Constitutional mismatch repair deficiency syndrome/ Lynch syndrome/ Turcot syndrome Type 1	DNA repair pathway	Kim et al., 2022

Brain tumours (not specified)				
MMRD brain tumors	*MLH1*	Constitutional mismatch repair deficiency syndrome/ Lynch syndrome/ Turcot syndrome Type 1	DNA repair pathway	Kim et al., 2022
MMRD brain tumors	*MSH2*	Constitutional mismatch repair deficiency syndrome/ Lynch syndrome/ Turcot syndrome Type 1	DNA repair pathway	Kim et al., 2022
	MSH3	Lynch syndrome	DNA repair pathway	Kim et al., 2022
High grade gliomas (glioblastoma)	*MSH6*	Constitutional mismatch repair deficiency syndrome/ Lynch syndrome	DNA repair pathway	Kim et al., 2022
	PMS1	Lynch syndrome	DNA repair pathway	Kim et al., 2022
	PMS2	Turcot syndrome Type 1/ Lynch Syndrome	DNA repair pathway	Kim et al., 2022
	BRCA1	Hereditary breast and ovarian cancer/ Fanconi anaemia	Tumour suppressor	Kline et al., 2017
	BRCA2	Hereditary breast and ovarian cancer/ Fanconi anaemia	Tumour suppressor	Kline et al., 2017
	FANCF	Fanconi anaemia	DNA repair pathway	Kline et al., 2017
	CHEK2	Von Hippel-Lindau syndrome	Tumour suppressor	Dahlin et al., 2020
	PTEN	-	Tumour suppressor	Gerges et al., 2013
	EGFR1	-	Cell growth and division	Gerges et al., 2013
	IDH1	-	Lipid metabolism	Gerges et al., 2013
	IDH2	-	Lipid metabolism	Gerges et al., 2013
	VHL	von Hippel-Lindau syndrome	Cell growth and division	Dahlin et al., 2020

	TP53	Li-Fraumeni syndrome	Tumour suppressor	Kline et al., 2017
	TSC1	Tuberous sclerosis complex	Tumour suppressor	Parsons et al., 2016
	TSC2	Tuberous sclerosis complex	Tumour suppressor	Parsons et al., 2016
	NF1	Neurofibromatosis Type 1	Tumour suppressor	Kline et al., 2017
	NF2	Neurofibromatosis Type 2	Tumour suppressor	Kline et al., 2017
	ERCC2	-	Nucleotide excision repair	Huang, 2014

Abbreviations: SHH- Sonic hedgehog, WNT- WNT/β-catenin signalling pathway, Ras-MAPK- Ras/RAF/MEK/ERK (mitogen-activated protein kinase), PI3K/AKT/mTOR- phosphatidylinositol-3 kinase (PI3K)/protein kinase B (Akt)/mammalian target of rapamycin (mTOR).

Supplementary Table 2. Details, supporting evidence and ACMG-AMP classification of pathogenic variants (n=3)

Patient ID	PC97		PC111
Type of CNS cancer	Medulloblastoma		Optic nerve glioma
Gene	*ATM*	*ATM*	*NF1*
Disorder associated with gene	Ataxia-telangiectasia	Ataxia-telangiectasia	Neurofibromatosis type 1
RefSeq	NM_000051.4	NM_000051.4	NM_000267.3:c.625C>T
Coding mutation	c.2200dup	c.8911C>T	c.625C>T
Protein change	p.Val734GlyfsTer4	p.Gln2971Ter	p.Gln209Ter
Genomic location	11:108127016-108127016	11:108235869-108235869	17:29508478-29508478
ClinVar ID	2128482	582761	187069
Read depth	188	119	38
Zygosity	Het	Het	Het
ACMG-AMP classification	Pathogenic	Pathogenic	Pathogenic
Criteria Met for Classification and Evidence	PSV1, PM2, PP5	PSV1, PM2, PP5	PSV1, PM2, PP5

Abbreviations: ATM- Ataxia telangiectasia mutated, NF1- Neurofibromin 1, ACMG-AMP- American College of Medical Genetics and Genomics, and Association for Molecular Pathology, Het- heterozygous.

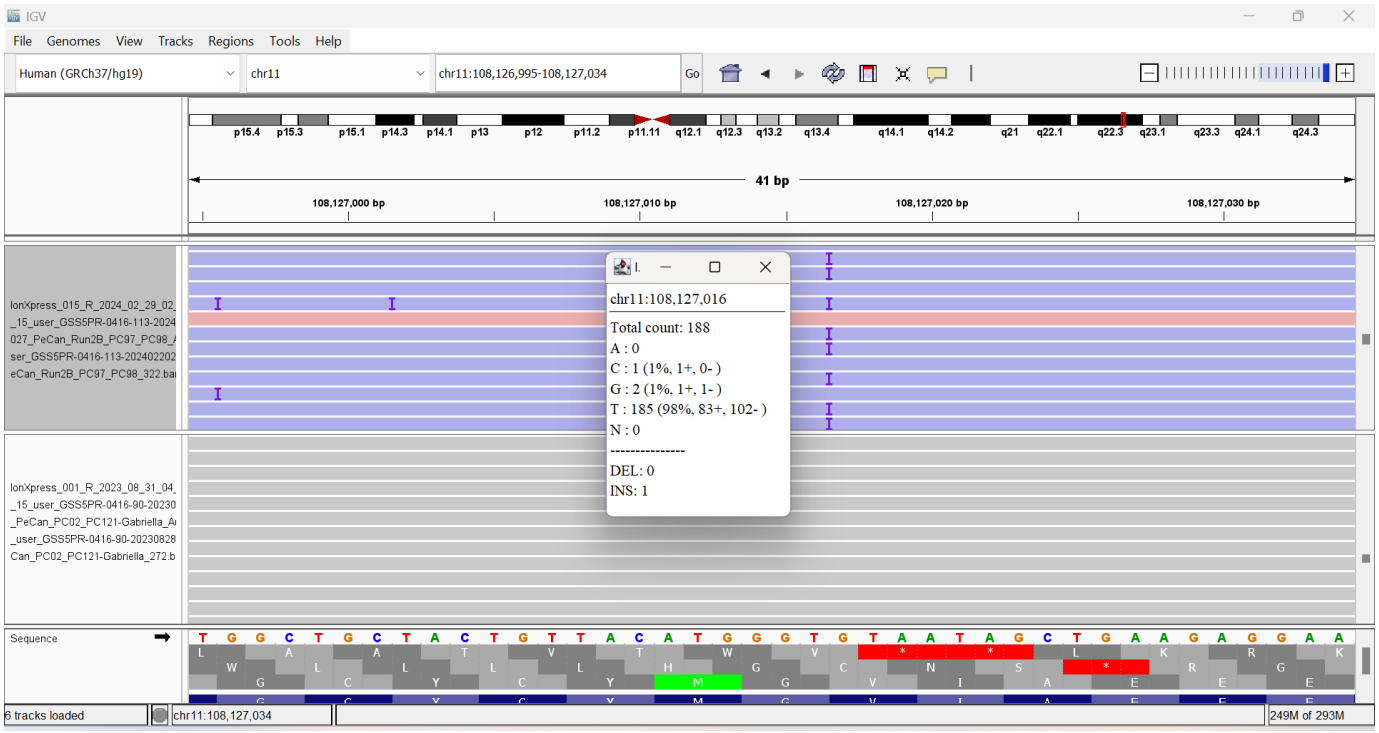
Supplementary Table 3. Details, supporting evidence and ACMG-AMP classification of variants of uncertain significance (n=8)

Patient ID	PC96			PC86	PC139	PC107	PC137	PC89
Type of CNS cancer	Astrocytoma			Craniopharyngioma	Brain tumour (unspecified)	Astrocytoma	Optic nerve glioma	Astrocytoma
Gene	*PMS2*	*NF1*	*POLE*	*MSH6*	*MLH1*	*PMS1*	*IDH1*	*PMS1*
Disorder associated with gene	TS Type 1/ LS	NF1	PPAP syndrome	CMMRD syndrome/ LS/ TS Type 1	CMMRD syndrome/ LS/ TS Type 1	LS	-	LS
RefSeq	NM_001322003.2	NM_001042492.3	NM_006231.4	NM_000179.3	NM_000249.4	NM_001128143.2	NM_001282387.1	NM_001128143.2
Coding mutation	c.863C>T	c.3169G>A	c.3311C>T	c.3180A>C	c.1013A>G	c.79G>C	c.964G>A	c.79G>C
Protein change	p.Ala288Val	p.Ala1057Thr	p.Thr1104Met	p.Leu1060Phe	p.Asn338Ser	p.Glu27Gln	Gly322Arg	p.Glu27Gln
Genomic location	7:6027128-6027128	17:29557915-29557915	12:133234521-133234521	2:48030566-48030566	3:37061929-37061929	2:190656614-190656614	2:209104614-209104614	2:190656614-190656614
ClinVar accession	186233	656225	405803	486908	89605	599334	Not Reported	599334
gnomAD AFR/AA AF	0.000006431	0.0000619	0.000185	-	0.0000615	0.000984	-	0.000984
CADD/SIFT /AlphaMissense/ FATHMM-MKL	22,70/ 0,04/ 0.71/ 0.98	20,40/ 0,27/ 0,06/ 0,92	18,82/ 0,06/ 0,08/ 0,79	24.3/ 0.06/ 0.468/ 0.79188	4.041/ 0.167/ 0.1036/ 0.9922	27.7/ 0/ 0.935/ 0.9956	8.755/ 0/ 0.7053/ 0.9866	27.7/ 0/ 0.935/ 0.9956
Read depth	82	79	69	192	71	97	117	108
Zygosity	Het	Het	Het	Het	Het	Het	Het	Het
ACMG-AMP classification	Uncertain significance*	Uncertain significance	Uncertain significance	Uncertain significance*	Uncertain significance	Uncertain significance*	Uncertain significance*	Uncertain significance*
Criteria Met for Classification and Evidence	PM1, PM2, PP3	PM1, PM2, PP2, BP4	PM1, PM2, BP4	PM1, PM2, PP3	PM1, PM2, PP3, BP6	PM2, PP3	PM2, PP3	PM2, PP3

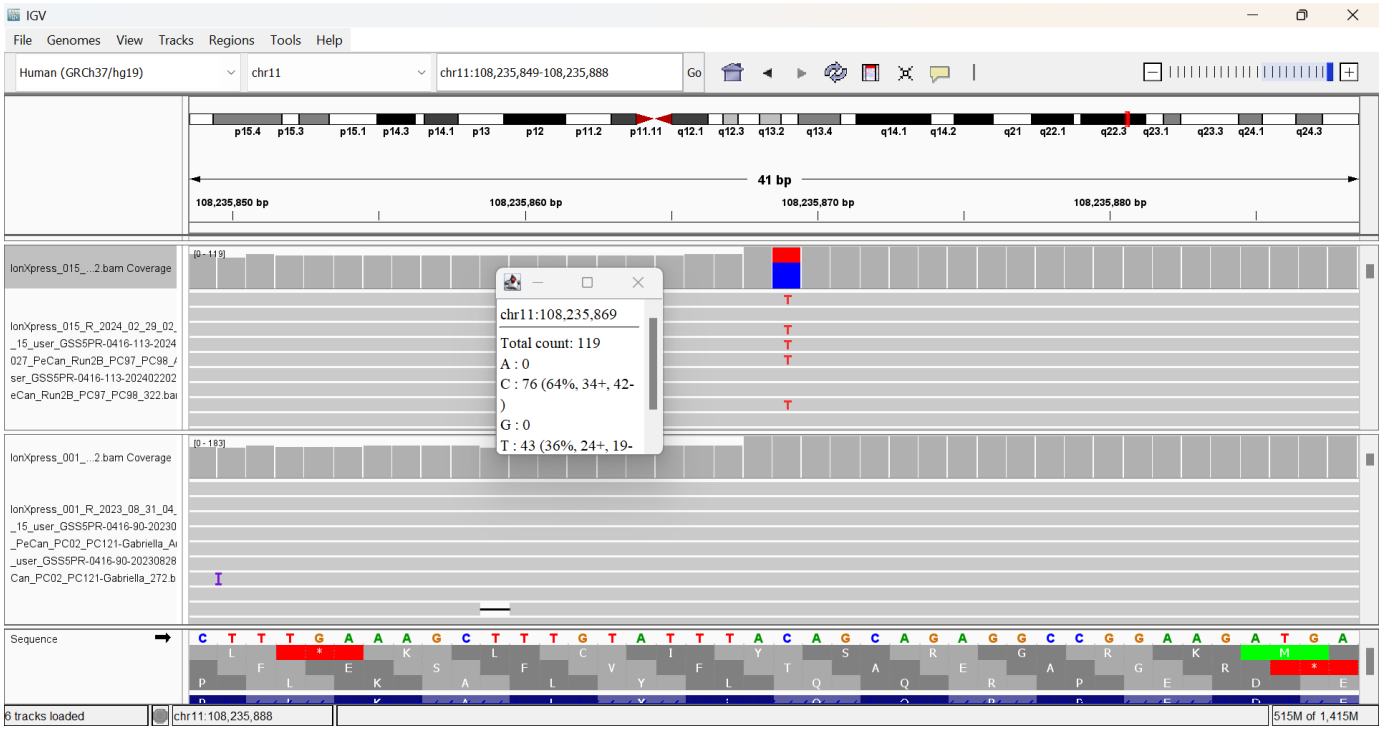
Abbreviations: PMS2- PMS1 homolog 2, mismatch repair system component; NF1- Neurofibromin 1, POLE- POLE DNA polymerase epsilon, catalytic subunit, MSH6- MutS homolog 6; MLH1- MutL protein homolog 1; PMS1- PMS1 homolog 1, mismatch repair system component; Isocitrate dehydrogenase 1. TS- Turcot syndrome, LS- Lynch syndrome, NF1- Neurofibromatosis Type 1, PAAP- Polymerase proofreading-associated polyposis, CMMRD- Constitutional mismatch repair deficiency, CADD- Combined Annotation-Dependent Depletion (CADD), SIFT- Sorting tolerant

from intolerant, FATHMM-MLK- Functional Analysis Through Hidden Markov Models - Multiple Kernel Learning, Het- heterozygous ACMG-AMP- American College of Medical Genetics and Genomics, and Association for Molecular Pathology, gnomAD- Genome Aggregation Database, AFR/AA AF- African/African American allele frequency. "Hot" variants of uncertain significance are indicated with stars (*).

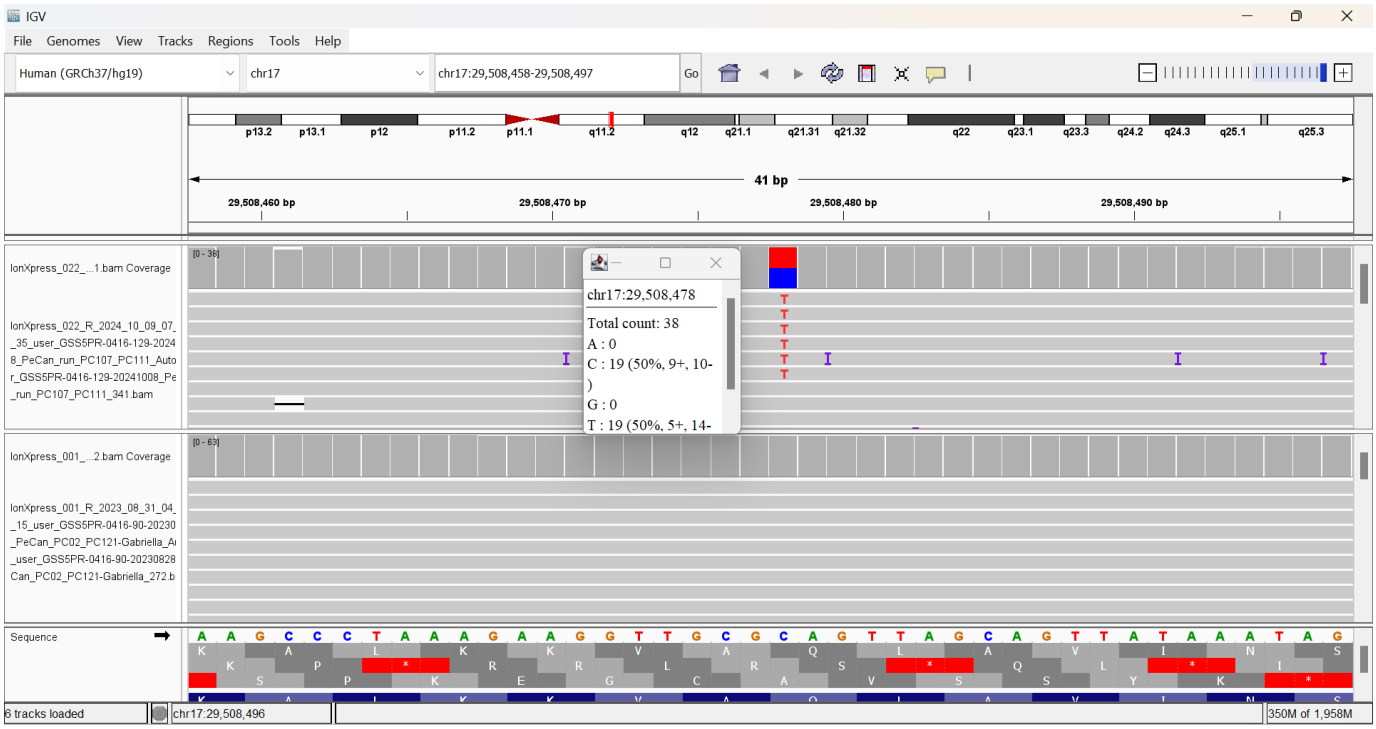
(A)


IGV
File Genomes View Tracks Regions Tools Help
Human (GRCh37/hg19)
chr11
chr11:108,126,995-108,127,034
Go
41 bp
108,127,000 bp
108,127,010 bp
108,127,020 bp
108,127,030 bp
chr11:108,127,016
Total count: 188
A : 0
C : 1 (1%, 1+, 0-)
G : 2 (1%, 1+, 1-)
T : 185 (98%, 83+, 102-)
N : 0
DEL: 0
INS: 1
Sequence
6 tracks loaded
chr11:108,127,034
249M of 293M

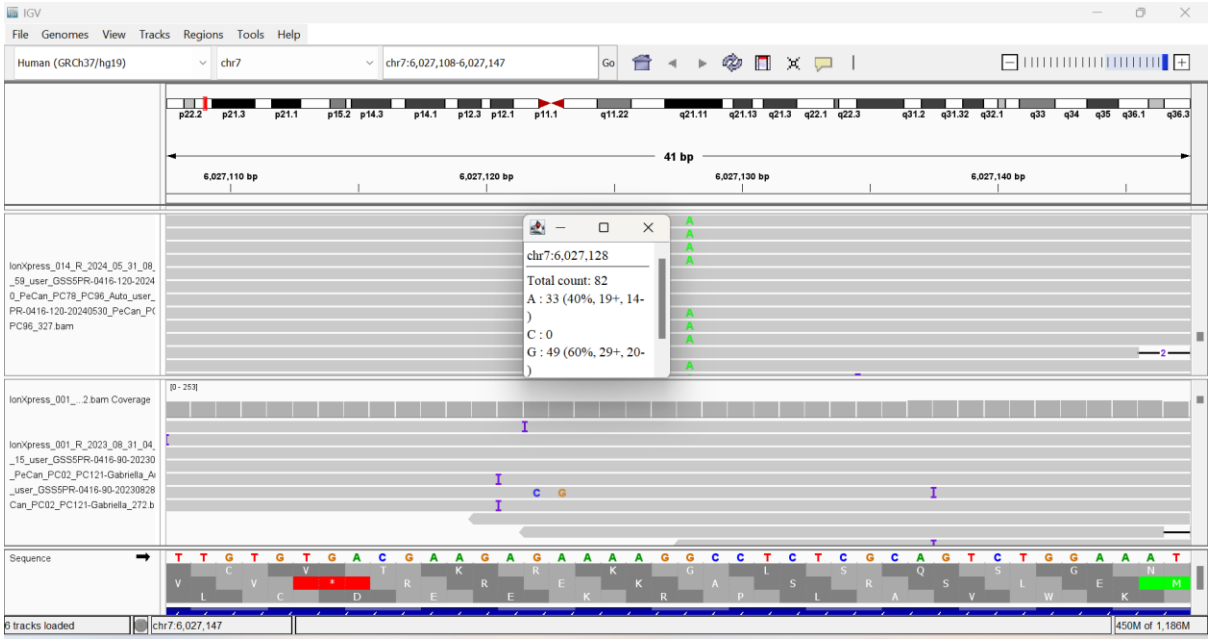
(B)


IGV
File Genomes View Tracks Regions Tools Help
Human (GRCh37/hg19)
chr11
chr11:108,235,849-108,235,888
Go
41 bp
108,235,850 bp
108,235,860 bp
108,235,870 bp
108,235,880 bp
chr11:108,235,869
Total count: 119
A : 0
C : 76 (64%, 34+, 42-)
G : 0
T : 43 (36%, 24+, 19-
Sequence
6 tracks loaded
chr11:108,235,888
515M of 1,415M

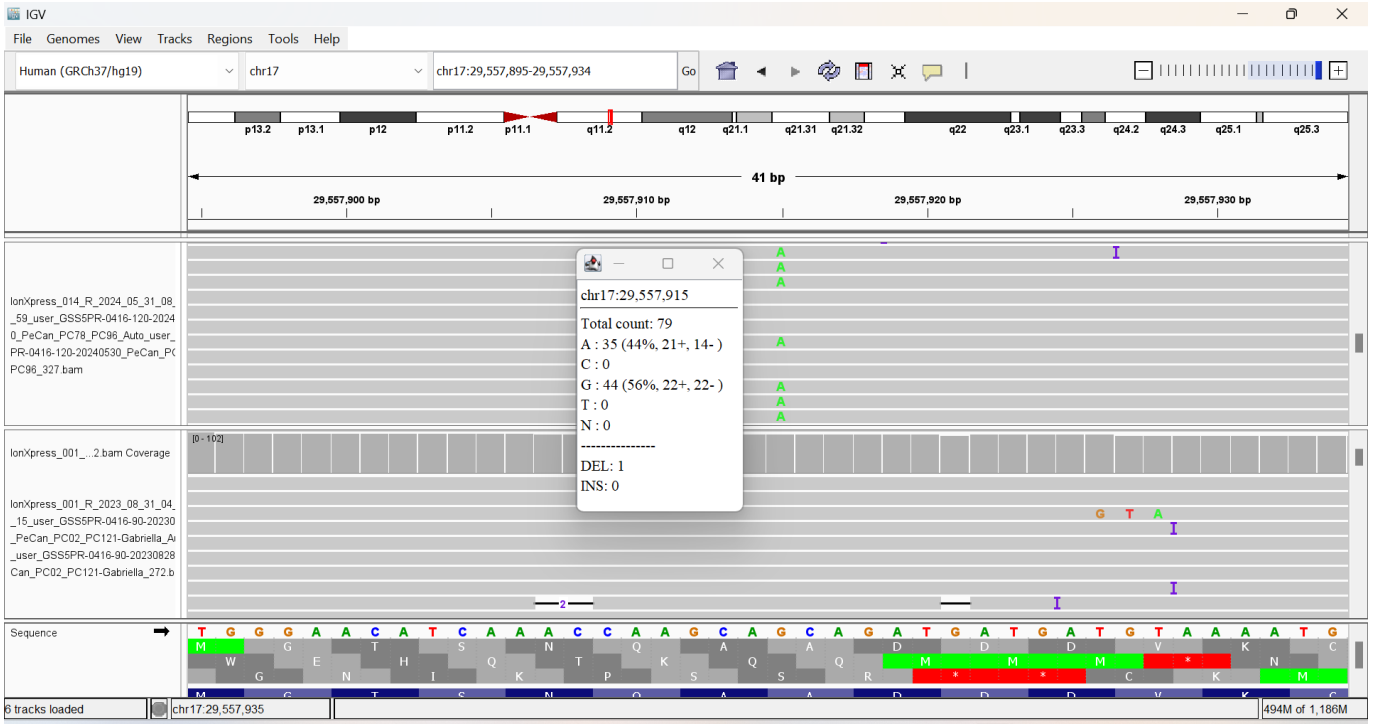
(C)


IGV
File Genomes View Tracks Regions Tools Help
Human (GRCh37/hg19)
chr17
chr17:29,508,458-29,508,497
Go
41 bp
29,508,460 bp
29,508,470 bp
29,508,480 bp
29,508,490 bp
chr17:29,508,478
Total count: 38
A : 0
C : 19 (50%, 9+, 10-)
G : 0
T : 19 (50%, 5+, 14-
Sequence
6 tracks loaded
chr17:29,508,496
350M of 1,958M

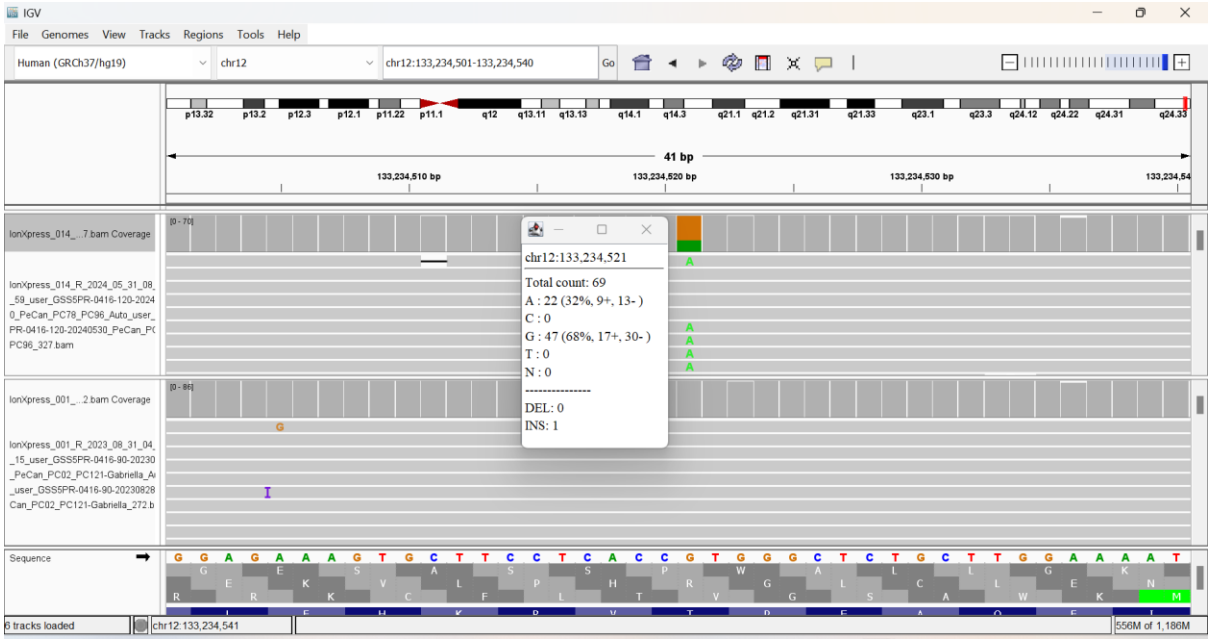
(D)


IGV
File Genomes View Tracks Regions Tools Help
Human (GRCh37/hg19)
chr7
chr7:6,027,108-6,027,147
Go
41 bp
6,027,110 bp
6,027,120 bp
6,027,130 bp
6,027,140 bp
chr7:6,027,128
Total count: 82
A : 33 (40%, 19+, 14-)
C : 0
G : 49 (60%, 29+, 20-)
Sequence
6 tracks loaded
chr7:6,027,147
450M of 1,186M

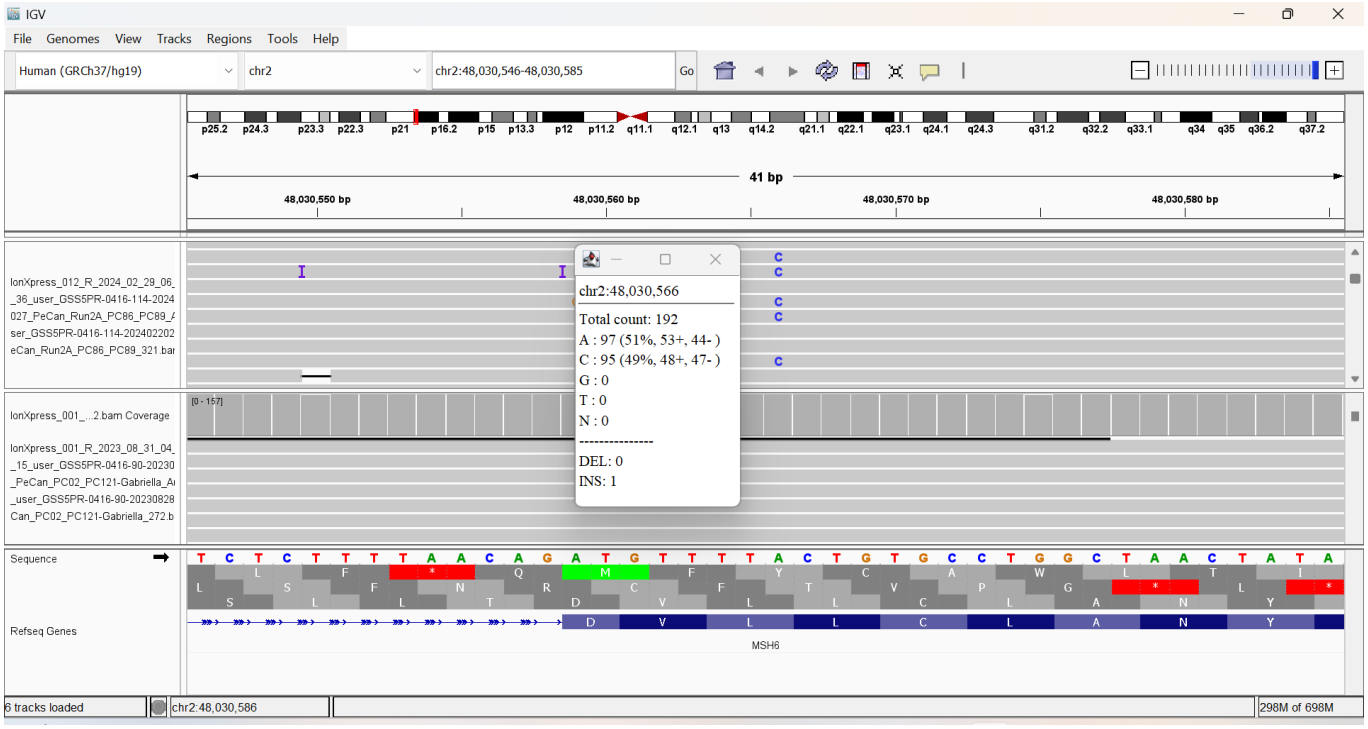
(E)


IGV
File Genomes View Tracks Regions Tools Help
Human (GRCh37/hg19)
chr17
chr17:29,557,895-29,557,934
Go
41 bp
29,557,900 bp
29,557,910 bp
29,557,920 bp
29,557,930 bp
chr17:29,557,915
Total count: 79
A : 35 (44%, 21+, 14-)
C : 0
G : 44 (56%, 22+, 22-)
T : 0
N : 0
DEL: 1
INS: 0
Sequence
6 tracks loaded
chr17:29,557,935
494M of 1,186M

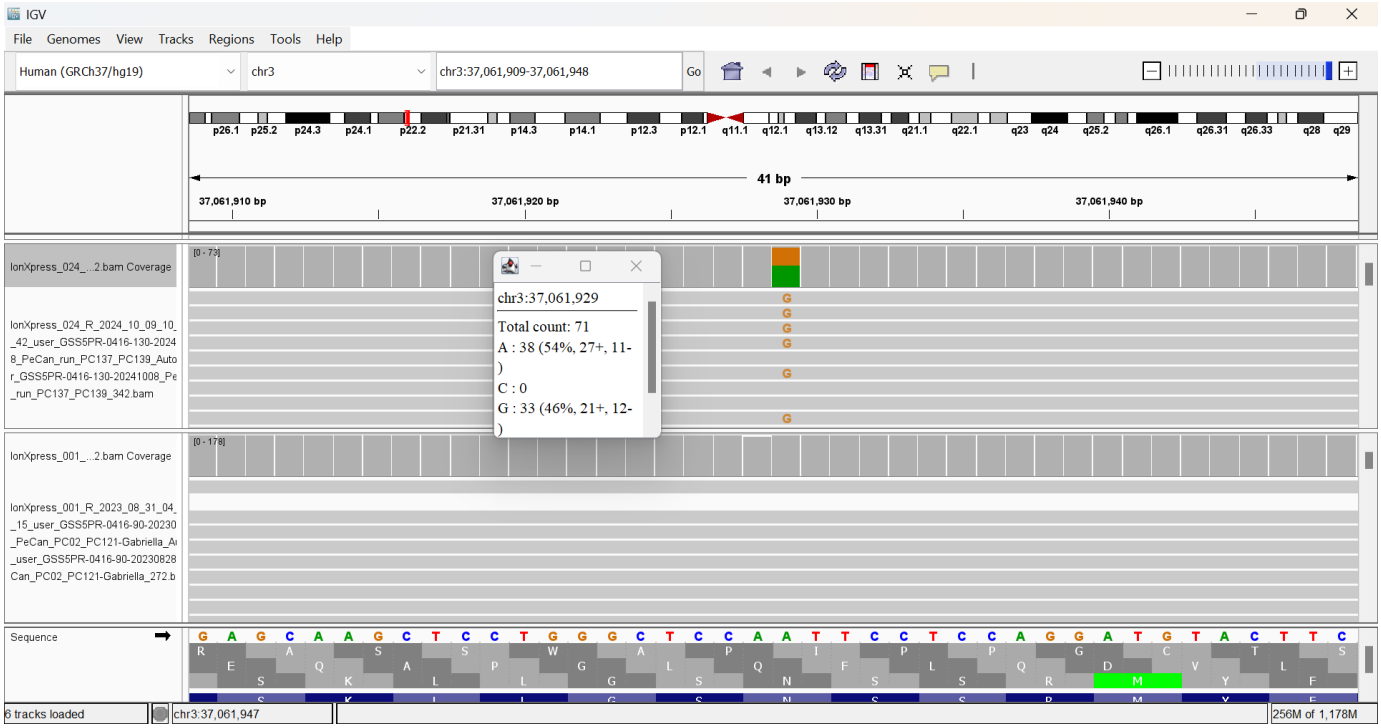
(F)


IGV
File Genomes View Tracks Regions Tools Help
Human (GRCh37/hg19)
chr12
chr12:133,234,501-133,234,540
Go
41 bp
133,234,510 bp
133,234,520 bp
133,234,530 bp
chr12:133,234,521
Total count: 69
A : 22 (32%, 9+, 13-)
C : 0
G : 47 (68%, 17+, 30-)
T : 0
N : 0
DEL: 0
INS: 1
Sequence
6 tracks loaded
chr12:133,234,541
556M of 1,186M

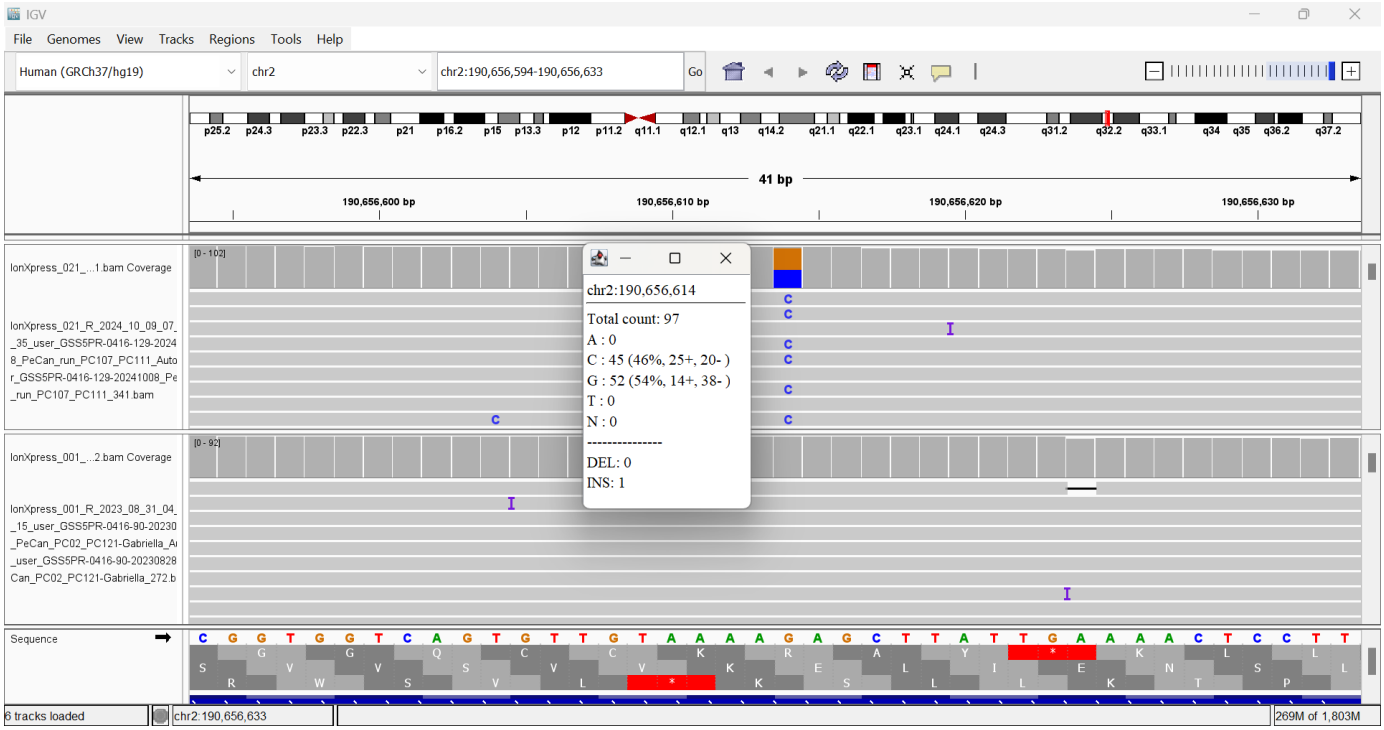
(G)


IGV
File Genomes View Tracks Regions Tools Help
Human (GRCh37/hg19)
chr2
chr2:48,030,546-48,030,585
Go
41 bp
48,030,550 bp
48,030,560 bp
48,030,570 bp
48,030,580 bp
chr2:48,030,566
Total count: 192
A : 97 (51%, 53+, 44-)
C : 95 (49%, 48+, 47-)
G : 0
T : 0
N : 0
DEL: 0
INS: 1
Sequence
Refseq Genes
MSH6
6 tracks loaded
chr2:48,030,586
298M of 698M

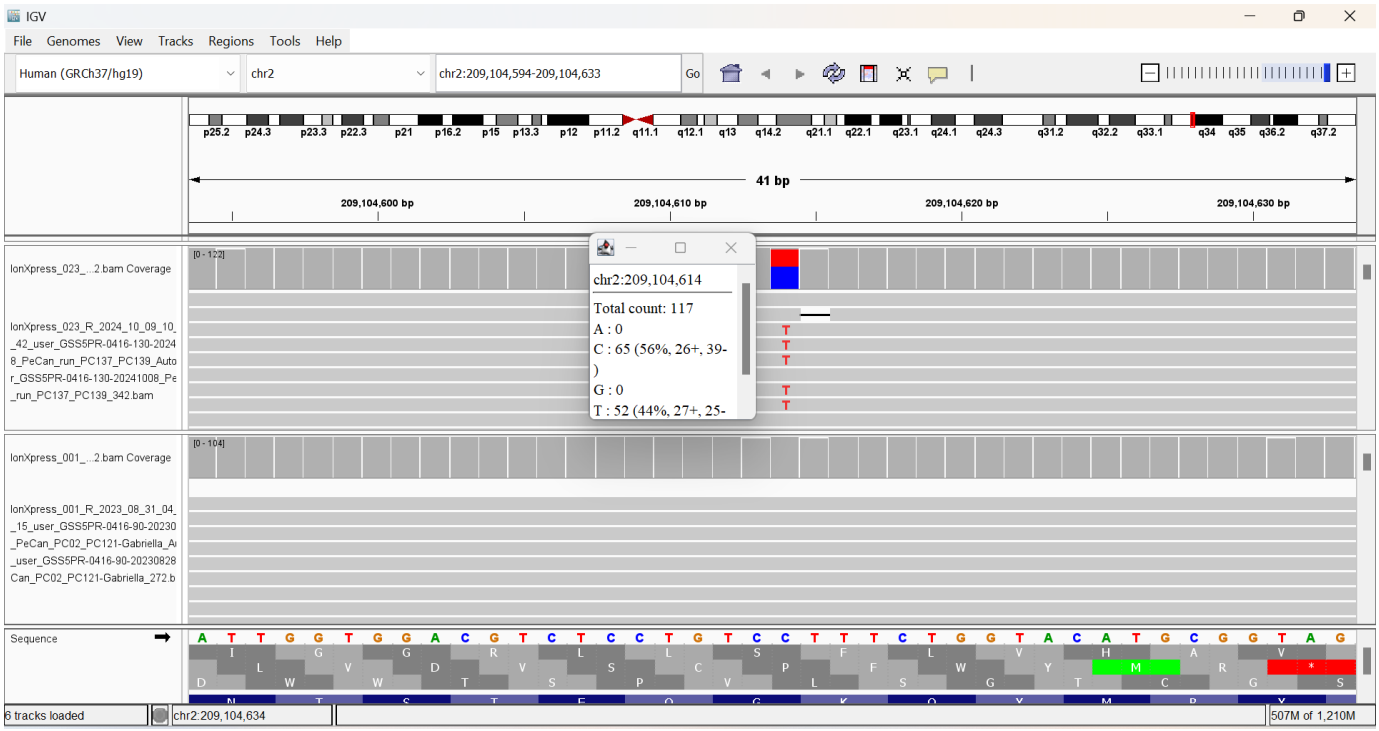
(H)


IGV
File Genomes View Tracks Regions Tools Help
Human (GRCh37/hg19)
chr3
chr3:37,061,909-37,061,948
Go
41 bp
37,061,910 bp
37,061,920 bp
37,061,930 bp
37,061,940 bp
chr3:37,061,929
Total count: 71
A : 38 (54%, 27+, 11-)
C : 0
G : 33 (46%, 21+, 12-)
Sequence
6 tracks loaded
chr3:37,061,947
256M of 1,178M

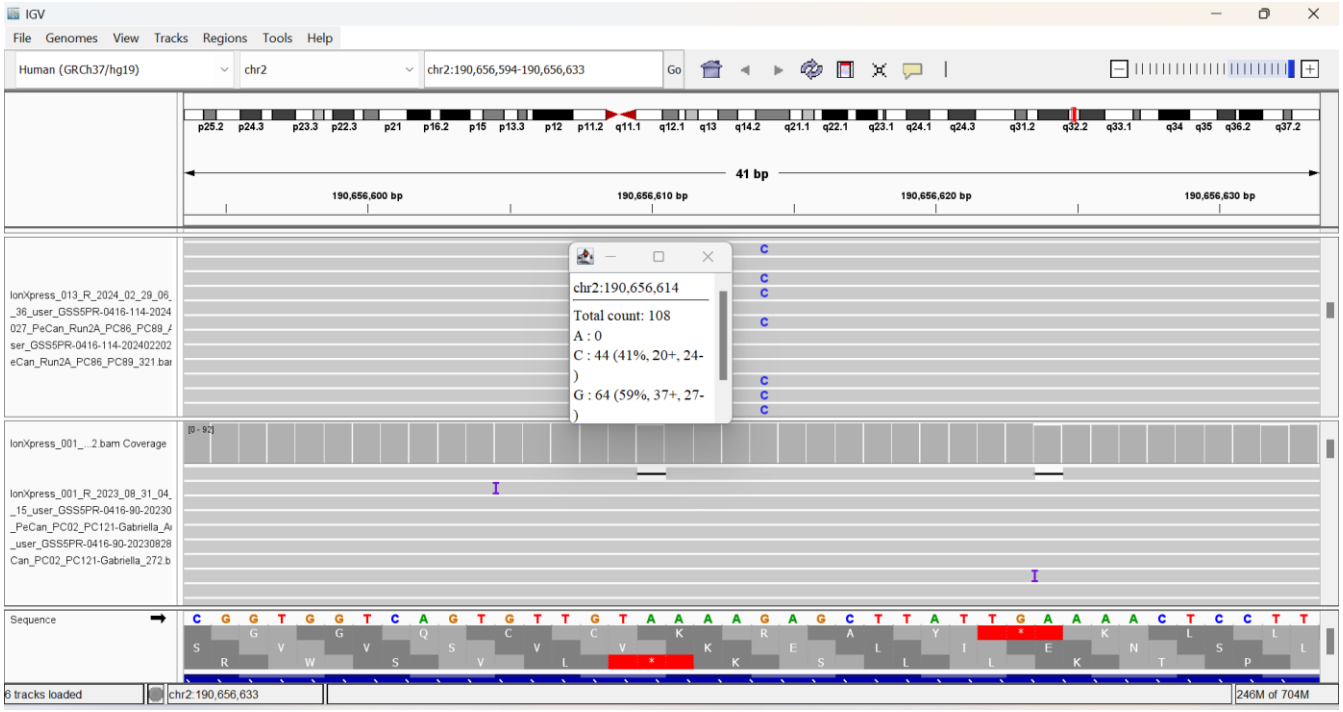
(I)


IGV
File Genomes View Tracks Regions Tools Help
Human (GRCh37/hg19)
chr2
chr2:190,656,594-190,656,633
Go
41 bp
190,656,600 bp
190,656,610 bp
190,656,620 bp
190,656,630 bp
IonXpress_021_...1.bam Coverage
IonXpress_021_R_2024_10_09_07_..._35_user_GSS5PR-0416-129-2024 8_PeCan_run_PC107_PC111_Auto r_GSS5PR-0416-129-20241008_Pe _run_PC107_PC111_341.bam
chr2:190,656,614
Total count: 97
A : 0
C : 45 (46%, 25+, 20-)
G : 52 (54%, 14+, 38-)
T : 0
N : 0
DEL: 0
INS: 1
IonXpress_001_...2.bam Coverage
IonXpress_001_R_2023_08_31_04_..._15_user_GSS5PR-0416-90-20230 _PeCan_PC02_PC121-Gabriella_Ai _user_GSS5PR-0416-90-20230828 Can_PC02_PC121-Gabriella_272.b
Sequence
6 tracks loaded
chr2:190,656,633
269M of 1,803M

(J)


IGV
File Genomes View Tracks Regions Tools Help
Human (GRCh37/hg19)
chr2
chr2:209,104,594-209,104,633
Go
41 bp
209,104,600 bp
209,104,610 bp
209,104,620 bp
209,104,630 bp
IonXpress_023_...2.bam Coverage
IonXpress_023_R_2024_10_09_10_..._42_user_GSS5PR-0416-130-2024 8_PeCan_run_PC137_PC139_Auto r_GSS5PR-0416-130-20241008_Pe _run_PC137_PC139_342.bam
chr2:209,104,614
Total count: 117
A : 0
C : 65 (56%, 26+, 39-)
G : 0
T : 52 (44%, 27+, 25-
IonXpress_001_...2.bam Coverage
IonXpress_001_R_2023_08_31_04_..._15_user_GSS5PR-0416-90-20230 _PeCan_PC02_PC121-Gabriella_Ai _user_GSS5PR-0416-90-20230828 Can_PC02_PC121-Gabriella_272.b
Sequence
6 tracks loaded
chr2:209,104,634
507M of 1,210M

(K)



Supplementary Figure 1. Candidate variant visualisation using IGV. Visualisation of each candidate variant for the determination of zygosity and coverage assessment using Integrative Genomics Viewer (IGV). (A) PC97: *ATM*(c.2200dup; p.Val734GlyfsTer4), (B) *ATM*(c.8911C>T; p.Gln2971Ter), (C) PC111: *NF1*(c.625C>T; p.Gln209Ter), (D) PC96: *PMS2*(c.863C>T; p.Ala288Val), (E) PC96: *NF1*(c.3169G>A; p.Ala1057Thr), (F) PC96: *POLE*(c.3311C>T; p.Thr1104Met), (G) PC86: *MSH6*(c.3180A>C; p.Leu1060Phe), (H) PC139: *MLH1*(c.1013A>G; p.Asn338Ser), (I) PC107: PMS1(c.79G>C; p.Glu27Gln), (J) PC137: IDH1(c.964G>A; p.Gly322Arg), (K) PC89: PMS1(c.79G>C; p.Glu27Gln)

Appendix B: Ethics Clearance Certificate (M240758)

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG	HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

Office of the Deputy Vice-Chancellor (Research and Innovation)

TO: Ms C Chatergoon
School of Pathology
Division of Human Genetics
Faculty of Health Sciences

E-mail: 2687027@students.wits.ac.za

CC: Supervisor: Drs L Lamola and M Kuthala
Lindie.Lamola@wits.ac.za
and <HREC-Medical Research Office@wits.ac.za>

FROM: Mr Iain Burns
Human Research Ethics Committee (Medical)
Tel: 011 717 1252

E-mail: Iain.Burns@wits.ac.za

DATE: 19 August 2024

REF: R14/49

PROTOCOL NO: **M240758** (This is your ethics application reference number. Please quote it in all enquiries, oral or written, relating to this study.)

PROJECT TITLE: *The germline genetic landscape of paediatric central nervous system tumours*

Please find attached the Clearance Certificate for the above project. I hope it goes well and that an article in a recognized publication comes out of it. This will reflect well on your professional standing and contribute to Government funding of the University.

MSWorks2000/Iain0007/Clearscan.wps

UNIVERSITY OF THE
WITWATERSRAND,
JOHANNESBURG

R49 Ms C Chatergoon

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)
CLEARANCE CERTIFICATE NO. M240758

NAME: (Principal Investigator)	Ms C Chatergoon	**DEGREE:** MSc
DEPARTMENT:	School of Pathology Division of Human Genetics Faculty of Health Sciences	
TITLE:	*The germline genetic landscape of paediatric central nervous system tumours*	
DATE CONSIDERED:	Ad hoc	
DECISION:	Approved unconditionally	
CONDITIONS:	Sub-study under M24/02/62	
NOTE:	If contact information regarding student study participants is required, please contact the Registrar's office <Nicoleen.Potgieter@wits.ac.za>	
SUPERVISORS:	Drs L Lamola and M Kuthala	
APPROVED BY:	Ms M van Niekerk, Co-Chairperson, HREC (Medical)	

DATE OF APPROVAL: 19 August 2024 **EXPIRY DATE:** 18 August 2029

This Clearance Certificate is valid for 5 years from the date of approval. An extension may be applied for.

DECLARATION OF INVESTIGATORS

To be completed in duplicate and **ONE COPY** returned to the Research Office secretariat on the 3rd floor, Phillip Tobias Building, Parktown, University of the Witwatersrand, Johannesburg.

I/we fully understand the conditions under which I am/we are authorized to carry out the above-mentioned research and I/we undertake to ensure compliance with these conditions. Should any departure be contemplated from the research protocol as approved, I/we undertake to submit details to the Committee. **I agree to submit a yearly progress report** in the format available at https://wits.ac.za/research/researcher-support/research-ethics/ethics-committees/. This report is due annually, for the duration of the Clearance Certificate, beginning on the first anniversary of Date of Approval above. Unreported changes to the study may invalidate the clearance given by the HREC (Medical).

Signature of Principal Investigator Date

Appendix C: Ethics Clearance Certificate (M240262)

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG	HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

Office of the Deputy Vice-Chancellor (Research and Innovation)

TO:	Dr L Lamola School of Pathology Division of Human Genetics National Health Laboratory Service E-mail: Lindie.Lamola@wits.ac.za
CC:	Supervisor: Not applicable <> and <HREC-Medical Research Office@wits.ac.za>
FROM:	Mr Iain Burns Human Research Ethics Committee (Medical) Tel: 011 717 1252 E-mail: Iain.Burns@wits.ac.za
DATE:	11 March 2024
REF:	R14/49
PROTOCOL NO:	**M240262** (This is your ethics application reference number. Please quote it in all enquiries, oral or written, relating to this study.)
PROJECT TITLE:	*Inherited cancer predisposing mutations in a cohort of childhood cancer patients*

Please find attached the Clearance Certificate for the above project. I hope it goes well and that an article in a recognized publication comes out of it. This will reflect well on your professional standing and contribute to Government funding of the University.



MSWorks2000/Iain0007/Clearscan.wps

UNIVERSITY OF THE WITWATERSRAND JOHANNESBURG

R49 Dr L Lamola

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)
CLEARANCE CERTIFICATE NO. M240262

NAME: Dr L Lamola **DEGREE:** N/A
(Principal and Co-Investigators)

DEPARTMENT: School of Pathology
Division of Human Genetics
National Health Laboratory Service

PROJECT TITLE: *Inherited cancer predisposing mutations in a cohort of childhood cancer patients*

DATE CONSIDERED: Ad hoc

DECISION: Approved unconditionally

CONDITIONS: Renewal of M18/08/55 - expired on 25 February 2024

NOTE: If contact information regarding student study participants is required, please contact the Registrar's office <Nicoleen.Potgieter@wits.ac.za>

SUPERVISOR: Not applicable

APPROVED BY:
Professor P Ruff, Chairperson, HREC (Medical)

DATE OF APPROVAL: 11 March 2024 **EXPIRY DATE:** 10 March 2029

This Clearance Certificate is valid for 5 years from the date of approval. An extension may be applied for.

DECLARATION OF INVESTIGATORS

To be completed in duplicate and **ONE COPY** returned to the Research Office secretariat on the 3rd floor, Phillip Tobias Building, Parktown, University of the Witwatersrand, Johannesburg.

I/we fully understand the conditions under which I am/we are authorized to carry out the above-mentioned research and I/we undertake to ensure compliance with these conditions. Should any departure be contemplated from the research protocol as approved, I/we undertake to submit details to the Committee. **I agree to submit a yearly progress report** in the format available at https://wits.ac.za/research/researcher-support/research-ethics/ethics-committees/. This report is due annually, for the duration of the Clearance Certificate, beginning on the first anniversary of Date of Approval above. Unreported changes to the study may invalidate the clearance given by the HREC (Medical).

Signature of Principal Investigator

13/03/2024
Date

Appendix D: Approved Research Protocol

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

FACULTY OF HEALTH SCIENCES

CANDIDATE'S SURNAME: [Please print]	Chatergoon	FIRST NAME/S:	Chetana	STUDENT NUMBER:	2687027
CURRENT QUALIFICATIONS: BSc Human Physiology, Genetics and Psychology BSocSci (Hons) Psychology BHSc (Hons) Human Genetics					
TEL: N/A	CELL: 060 533 4091	E-MAIL: 2687027@students.wits.ac.za		FAX: N/A	
DEGREE FOR WHICH PROTOCOL IS BEING SUBMITTED: Msc(Med) Genomic Medicine					
PART-TIME OR FULL-TIME: Full time					
FIRST REGISTERED FOR THIS DEGREE:	TERM : 1			YEAR: 2024	
DEPARTMENT: Human Genetics					
TITLE OF PROPOSED RESEARCH: An investigation into the germline genetic landscape of paediatric central nervous system tumours using whole exome sequencing data					
CANDIDATE'S SIGNATURE: Chatergoon				DATE: 10/04/2024	
SUPERVISOR 1 (NAME & SURNAME): Lindie Lamola				% Supervision	
SUPERVISOR'S QUALIFICATIONS: BSc, BSc (Med) Hons, MSc, PhD					
SUPERVISOR'S DEPARTMENT: Human Genetics					
SUPERVISOR'S ADDRESS / TEL / E-MAIL: lindie.lamola@wits.ac.za					
SUPERVISOR 2 (NAME & SURNAME): Khuthala Mnika				% Supervision	
SUPERVISOR'S QUALIFICATIONS: BSc, BSc (Med) Hons, MSc, PhD					
SUPERVISOR'S ADDRESS / TEL / E-MAIL: k.mnika@uct.ac.za					
SUPERVISOR 3 (NAME & SURNAME):				% Supervision	
SUPERVISOR'S QUALIFICATIONS					
SUPERVISOR'S ADDRESS / TEL / E-MAIL:					
SYNOPSIS OF RESEARCH:(Brief summary of proposed research project; between 200-300 words only; with sub-headings: an introduction and justification for study, aim/s, proposed methodology and expected outcome/s) [Use reverse side of this page if more space is required]					

WITS ETHICS NOT REQUIRED: Yes No WITS ETHICS PENDING: (Yes) No WITS ETHICS APPROVED: Yes No (circle appropriate symbol)* ***Please note the final human ethics clearance certificate or animal ethics certificate must be available prior to starting research**		IF Y SUPPLY ETHICS CLEARANCE CERTIFICATE AS ATTACHMENT AND INCLUDE ETHICS NUMBER HERE:
As supervisor/s, I/we confirm that I have read the protocol which has been submitted for assessment.		
SIGNATURE OF SUPERVISOR/S:	.....	.....
SIGNATURE PG OFFICE STAFF	REGISTERED YES..... NO.....	STAMP

SYNOPSIS OF RESEARCH CONTINUED

Introduction

Central nervous system (CNS) tumours are the second most common type of paediatric cancer and have some of the highest cancer-related mortality rates. Paediatric CNS tumours occur in patients with a genetic predisposition in 8-12% of cases, indicative of cancer predisposition syndromes. Next generation sequencing technology can be used to identify putative disease-causing variants in cancer predisposition genes, this has potential translational relevance This study forms part of a larger study which is investigating the genetics of paediatric cancers in South Africa. Knowledge gained in this study will advance our understanding of the germline genetic landscape underlying paediatric CNS tumours.

Aim

The aim of this project is to identify germline genetic variants in cancer predisposition genes in a subset of paediatric patients with CNS tumours.

Methodology

Candidate genes associated with the following cancers of the CNS: medulloblastoma, astrocytoma, optic nerve glioma, craniopharyngioma, pineoblastoma and other unspecified brain tumours will be investigated using a set of pre-generated data. A bioinformatic analysis will take place using Whole exome sequencing data from 15 patients with CNS tumours. A bioinformatics pipeline will be created to allow for genetic variants to be annotated, filtered, and prioritised. Thereafter a shortlist of variants will undergo classification using the American College of Medical Genetics and Genomics and Association of Medical Pathology (ACMG-AMP) guidelines.

Expected outcome

Variants of uncertain significance, likely pathogenic and pathogenic variants will be identified in the specified genes in each of the 15 files. Knowledge gained in this study is intended to advance our understanding of the landscape of germline genetic changes that underlie CNS tumours.



UNIVERSITY OF THE
WITWATERSRAND,
JOHANNESBURG



NATIONAL HEALTH
LABORATORY SERVICE

The germline genetic landscape of paediatric central nervous system tumours

Research proposal

Chetana Chatergoon (2687027)

MSc (Med) Genomic Medicine

Division of Human Genetics, School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, National Health Laboratory Service (NHLS)

Supervisors: Dr. Lindie Lamola

BSc, BSc (Med) Hons, MSc, PhD

Dr. Khuthala Mnika

BSc, BSc (Med) Hons, MSc, PhD

Contents

List of tables and figures 4
1 Introduction 6
1.1 Common paediatric cancers of the central nervous system 8
1.2 Central nervous system cancer research in Africa 8
1.3 The genetic landscape of paediatric central nervous system tumours 9
1.4 Molecular investigation of the genetic landscape of tumours using whole-exome sequencing 10
1.5 Study rationale 11
2 Aims and objectives 11
3 Methods 12
3.1 Study cohort 12
3.2 Performing a thorough literature review for candidate gene selection 12
3.3 Bioinformatics analysis 13
3.3.1 Variant annotation and prioritisation 13
3.3.2 Variant quality analysis using Integrative Genomics Viewer 14
3.3.3 Variant classification using ACMG-AMP guidelines and published literature 15
4 Ethics 15
5 Timing 15
6 Funding 16
7 Limitations 16
8 References 17

List of tables and figures

Figure 1. Distribution of paediatric cancers in patients younger than 19 years old 8

Table 1. Frequency of paediatric cancers of the central nervous system in children younger than 15 years of age versus children and adults of all ages 9

List of abbreviations

1. Cancer predisposition syndrome (CPS)
2. Germline predisposing pathogenic variant (GPV)
3. Central nervous system (CNS)
4. Whole-exome sequencing (WES)
5. Medulloblastoma (MB)
6. Astrocytoma (AS)
7. Optic nerve glioma (ONG)
8. Craniopharyngioma (CP)
9. Pineoblastoma (PB)
10. Constitutional mismatch repair deficiency syndrome (CMMRD)
11. Pilocytic astrocytoma (PA)
12. American College of Medical Genetics and Genomics and Association of Medical Pathology (ACMG-AMP)
13. Integrative Genomics Viewer (IGV)
14. Variant Call Format (VCF)
15. Binary Map Alignment (BAM)
16. Ensembl Variant Effect Predictor (VEP)
17. genome Aggregation Database (gnomAD)
18. Combined Annotation Dependent Depletion (CADD)

1 Introduction

Cancer is a term used to describe a large group of diseases that can occur in any organ or tissue when uncontrolled proliferation may result in metastasis of abnormal cells to other parts of the body (Zhang et al., 2015). Between 75-85% of cancers are caused by sporadic somatic pathogenic variants, whilst the remaining 15-25% of cases arise from germline genetic variants in cancer susceptibility genes (Sait et al., 2021). A portion of these inherited cancers are known to result from an underlying cancer predisposition syndrome (CPS) (Gröbner et al., 2018). Cancer predisposition syndromes are associated with characteristic germline mutations and increase the likelihood of developing early-onset cancers in comparison to the general population. It is estimated that 8-12% of paediatric cancers arise in patients with germline predisposing pathogenic variants (GPV's), with 8-19% of these occurring in a classical tumour predisposition gene that underlies a known CPS (Sait et al., 2021). Although, research indicates that this percentage may be higher in children with central nervous system (CNS) tumours (Kline et al., 2017). The most commonly occurring paediatric cancers are represented in Figure 1, the orange portion corresponds to CNS cancers, which are the second most common paediatric cancer, after leukaemia, and the most common solid tumour occurring in paediatric patients.

Paediatric tumours of the brain and spinal cord arise from the abnormal proliferation of nerve or supportive cells and can be either malignant (58%) or non-malignant (42%), they are referred to as CNS tumours (Sait et al., 2021). The global incidence of CNS tumours is approximately 5.5 per 100 000 and the mortality rate is 0.72 per 100,000 population, making CNS tumours the leading cause of cancer-related mortality in children (Capasso et al., 2020).

The intracranial location of these tumours poses challenges to surgery, and survivors face life-long morbidity. Chemotherapy and radiotherapy have cytotoxic effects and patients may be at risk of developing cognitive disabilities or secondary cancers (Parsons et al., 2016).

Investigating the genetic landscape of paediatric CNS cancers is an essential first step towards identifying targetable driver mutations that may inform individualised treatment based on genetic factors as opposed to routine standard-of-care treatments.

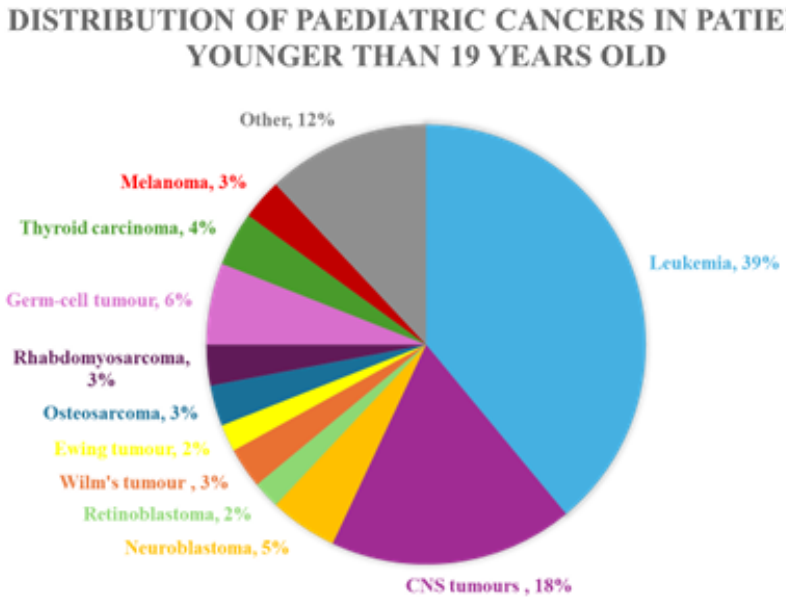


Figure 1. Distribution of paediatric cancers in patients younger than 19 years old. The frequencies were calculated from Centres for Disease Control and Prevention data and based on incidence in United States for the years 1999–2016 (adapted from Capasso et al., 2020).

Precision oncology has emerged as advances in next-generation sequencing technologies, including whole-exome sequencing (WES) have contributed to the characterisation of an increasing number of cancer predisposition genes, thus providing new avenues for investigating CNS tumour risk at the genomic level (Kline et al., 2017). Studies and molecular tumour profiling that inform treatment programmes are predominantly carried out in first world countries. Consequently, paediatric CNS cancer remains understudied in Africa, highlighting a critical gap in global oncology research that can be addressed through the application of cutting-edge genetic technologies, such as WES. The comprehensive investigation of the

genetic landscape of germline variants and their implications in tumorigenesis in Africa may enhance risk assessment, early detection, and personalised treatment strategies tailored to the unique genetic profiles of individuals within the African context (Rotimi et al., 2021).

1.1 Common paediatric cancers of the central nervous system

Paediatric CNS tumours are classified based on their intracranial location, histopathological factors, and more recently, genomic factors, and are distinctive from their adult counterparts in terms of their histology and genetic landscape (Gröbner et al., 2018). Table 1 summarises the frequency of some of the most common paediatric cancers, including astrocytoma (AS), medulloblastoma (MB), craniopharyngioma (CP), optic nerve glioma (ONG), pineoblastoma (PB), and other unspecified gliomas.

Table 1. Frequency of paediatric cancers of the central nervous system in children younger than 15 years of age versus children and adults of all ages.

Tumour type	Children younger than 15(%)	Children and adults (%)
Gliomas	70	45
Astrocytoma	30	15
Medulloblastoma	20	4
Craniopharyngioma	10	3
Glioblastoma	5	15
Optic nerve glioma	5	1
Pineoblastoma	1	0.1

Table 1 indicates the frequency of paediatric cancers of the central nervous system in children younger than 15 years of age versus children and adults of all ages (adapted from Miabi et al., 2007).

1.2 Central nervous system cancer research in Africa

While numerous global studies in high- and middle-income countries have investigated the genetic architecture of CNS tumours, there remains a notable dearth of research focusing on African populations. Cases of CNS tumours are sparsely reported in Africa, resulting in the misconception that CNS cancers are a rarity across the entire continent (Rotimi et al., 2021). However, systematic studies on CNS tumour incidence and mortality have never been performed in Africa (Mbi Feh et al., 2021). Consequently, current research may not adequately

capture the diverse genetic landscapes and unique predispositions to CNS cancers prevalent in African populations. The high genetic diversity observed in African populations necessitates the need for further research. Rotimi et al. (2021) conducted a literature search including publications on CNS cancer in Africa in the period spanning from 1990 to December 2019. It was concluded that only 0.329% of cancer studies globally were on African populations, and a further 0.016% were on cancer genetics and genomics from Africa. Papers exclusively reporting on paediatric CNS cancers constituted 38.5% of existing literature. Mbi Feh et al. (2021) reported the following frequencies of CNS cancers in Africa between 1960 and 2017, AS represented 24.70% of cases, 4.26% were MB, 4.07% were CP, glioblastoma and unspecified gliomas were 2.20% and PB made up 0.19%.

1.3 The genetic landscape of paediatric central nervous system tumours

Paediatric CNS tumours can occur in patients with a genetic predisposition, potentially revealing underlying CPS's driven by GPV's. The complete spectrum of GPV's that contribute to the development of paediatric CNS tumours remains partially understood (Dahlin et al., 2020). Germline GPV's can be *de novo*, arising spontaneously in the gametes, or they can be inherited across multiple generations (Sait et al., 2021). The majority of cancer predisposition genes are involved in fundamental biological pathways linked to tumorigenesis, such as chromatin remodelling, cell-cycle control, transcriptional regulation, DNA repair, and apoptosis (Kline et al., 2017). Consequently, alterations in these genes impair the regulation of cell proliferation and significantly increase the risk of cancer development (Parsons et al., 2016).

Several CPS's are inherited in an autosomal dominant manner, wherein one allele of a tumour suppressor gene is inactivated in the germline and the somatic inactivation of the remaining wildtype allele causes the development of CNS neoplasms. In the case of paediatric cancers, this loss of heterozygosity often occurs during childhood and is a consequence of

nondisjunction during cell division (Sait et al., 2021). In other circumstances, activating mutations occur in oncogenes in *trans* with the germline variant (Waszak et al., 2018). Examples of autosomal dominant CPS's are Gorlin syndrome, Li-Fraumeni syndrome, Lynch syndrome, and Neurofibromatosis type 1. Autosomal recessive CPS's are the consequence of homozygous or compound heterozygous GPV's that result in biallelic inactivation of a tumour suppressor gene (Helfferich et al., 2016). Examples of autosomal recessive CPS's are Fanconi anaemia and constitutional mismatch repair deficiency syndrome (CMMRD).

1.4 Molecular investigation of the genetic landscape of tumours using whole-exome sequencing

Modern molecular technologies, such as WES are applicable in paediatric oncology by serving as a tool to study the complex genetic landscape underlying tumour etiology (Capasso et al., 2020). This technique is particularly useful, as by analysing the coding regions of the genome, germline variants may be characterised and the interplay between hereditary predisposition and oncogenic drivers may be studied (Kline et al., 2017). Additionally, there may be immediate treatment implications if it is determined that a cancer has arisen due to a CPS. For example, considering the significant risk of developing secondary cancers owing to DNA double-strand breaks caused by gamma-radiation in cells haploinsufficient for functional tumour suppressor proteins, current recommendations urge minimal use or avoidance of radiotherapy and chemotherapy when possible in patients with Li-Fraumeni syndrome, Fanconi anaemia, Neurofibromatosis type 1and Nijmegen breakage syndrome (Kline et al., 2017).

Identifying GPV's provides insight into the genes that drive tumorigenesis. This has potential translational relevance in prognostication and individualised therapy design (Parsons et al., 2016). Between 40-50% of primary paediatric tumours harbour a potentially targetable genetic germline or somatic variant (Gröbner et al., 2018). This highlights an opportunity for clinicians to leverage targeted therapies and precision medicine approaches to mitigate disease

progression and formulate more effective and less harmful treatment programmes. An example is the use of Temozolomide, a drug commonly used for treating glioblastomas, which has been determined to be less effective in MMR-deficient tumours, providing a growth advantage to tumour cells. Therefore, this drug should be avoided in patients with CMMRD and immunotherapies, such as immune checkpoint inhibitors should be considered (Guerrini-Rousseau et al., 2019).

1.5 Study rationale

Paediatric CNS cancer remains understudied in Africa, highlighting a critical gap in global oncology research that can be addressed through the application of cutting-edge genetic technologies, such as WES. There is a need for a comprehensive investigation of the genetic landscape of germline genetic variants and their implications in tumorigenesis in South Africa. By leveraging WES, the knowledge gap may be bridged through the investigation of the spectrum of germline genetic variants underlying paediatric CNS cancer in Africa. This study forms part of a larger study, the PeCan project, which is investigating the genetics of paediatric cancers in South Africa. Knowledge gained in this study is intended to advance our understanding of the germline genetic changes that underlie CNS tumours.

2 Aims and objectives

The aim of this project is to identify germline genetic variants in cancer predisposition genes in a subset of paediatric patients with CNS tumours. The objectives are as follows:

1. To identify candidate genes associated with the following cancers of the CNS: MB, AS, ONG, CP, PB, and other unspecified brain tumours.
2. To perform a bioinformatics analysis on pre-generated WES data on up to 100 cancer predisposition genes associated with the aforementioned cancers.

3. To analyse the quality of the shortlisted variants using Integrative Genomics Viewer (IGV) (Robinson et al., 2022).
4. To identify variants in the genes of interest and classify them according to the American College of Medical Genetics and Genomics and Association of Medical Pathology (ACMG-AMP) guidelines (Richards et al., 2015).

3 Methods

3.1 Study cohort

This study includes a sample of fifteen paediatric patients diagnosed with different types of brain tumours/cancers including MB, AS, ONG, CP, PB, and other unspecified brain tumours, that were recruited as part of the PeCan project (Ethics clearance pre-2024: M180855; and 2024 onwards: M240262). Prior to the commencement of this study, consent to use the patient samples was obtained during the recruitment process. DNA was extracted from blood and saliva samples and WES was performed using the Ion GeneStudio S5™ Series Sequencer by (ThermoFisher Scientific, Waltham). Files containing pre-generated WES data in the Variant Call Format (VCF), Binary Map Alignment (BAM) and index file format (BAI) will be obtained for data analysis in this study. Currently, the data is being stored on a local server at the University of the Witwatersrand. My supervisor, Dr. Lindie Lamola, has access, and the data will be made available to me following the approval of my study protocol and ethics application.

3.2 Performing a thorough literature review for candidate gene selection

A critical literature review will take place in order to identify candidate cancer susceptibility genes that are known to be associated with germline inheritance or CPS's. Sources will include literature reviews and original research studies using curated CNS cancer gene panels to

identify common genes underlying paediatric cancers. Specifically, genes that overlap between commercial panels such as the VistaSeq Brain/CNS/PNS Cancer Panel and Invitae Hereditary Nervous System/Brain Cancer Panel will be prioritised, as multiple panels include the same genes (Alonso-Luna et al., 2023; NIH genetic testing registry, 2024). Notably, variants occurring in African Americans will be prioritised, due to the similarity in populations and lack of CNS tumour data in South Africa. A list of up to 100 candidate genes associated with germline inheritance in medulloblastoma, astrocytoma, optic nerve glioma, craniopharyngioma, pineoblastoma, and other unspecified brain tumours will be drafted.

3.3 Bioinformatics analysis

3.3.1 Variant annotation and prioritisation

Next-generation sequencing generates thousands of sequence variants that must be annotated, filtered, and prioritized before interpretation. This large amount of data is stored in VCF, BAM, and BAI files that require bioinformatics analysis. Ensembl Variant Effect Predictor (VEP) (https://www.ensembl.org/info/docs/tools/vep/index.html) will be used for variant annotation of the candidate genes (McLaren et al., 2016). The impact, location, and consequence of the variants, as well as their functional effect on the gene and its protein will be recorded. During annotation, identifiers such as Human Genome Variation Society nomenclature (HGVSc and HGVSp) will be assigned to known genes, variants, and proteins. Variant and frequency data from databases such as the genome Aggregation Database (gnomAD), and others, will be accessed to obtain the minor allele frequencies of the variants from diverse ancestries.

The identified genetic variants will be filtered based on their clinical significance using the process of variant prioritization. Firstly, genes that have not been identified as candidate genes will be excluded. Once the candidate genes have been isolated, variants with an allele frequency above 1% will be filtered out because the prevalence of CPS's in the population is

expected to be low (Richards et al., 2015). This range is to accommodate genes associated with rare genetic disorders, such as CMMRD, as well more common ones, such as NS.

Prioritisation parameters will be decided on considering the evaluation of disease mechanism, the predicted effect on the regulatory region, and their clinical significance. The consequences of any variants in regulatory regions will be determined. Variants that are not expected to have an effect on functional genes, or protein structure and function will be removed. The remaining variants will be analysed using SIFT and Polyphen among other *in silico* prediction tools to determine the effects of the variant on protein function, and the potential consequences of amino acid substitutions (Flanagan, Patch and Ellard, 2010). Combined Annotation Dependent Depletion (CADD) will be used to measure the deleteriousness of a particular variant (Rentzsch et al., 2021). Clinvar will be used as a tool to determine clinical significance and benign and likely benign variants are excluded (Landrum et al., 2020). A shortlist will then be compiled, including variants of uncertain significance, likely pathogenic and pathogenic variants.

3.3.2 Variant quality analysis using Integrative Genomics Viewer

The quality and coverage of each variant will be assessed, and the BAM and BAI files will be visualised using Integrative Genomics Viewer (IGV). This is a powerful genome browser used to determine true positive variants by ruling out sequencing errors, such as artefacts, and it allows for increased confidence in variant calls (Robinson et al., 2022). The quality of the variants will be assessed by evaluating the read alignment, read depth, and variant allele frequency. The coverage of the target sequences will be determined by examining the pileup of sequencing reads at the variant positions and variants with over thirty reads will be considered true variants. The variant allele frequency will be measured and zygosity of the variants will be determined.

3.3.3 Variant classification using ACMG-AMP guidelines and published literature

The shortlisted variants will be interpreted using published literature and classified using the American College of Medical Genetics and Genomics and Association of Medical Pathology (ACMG-AMP) guidelines. The ACMG-AMP guidelines describe a set of codes that are assigned to variants ranging from benign, likely benign, variant of uncertain significance, likely pathogenic and pathogenic. Each code is weighted, ranging from very strong, strong, moderate, and supporting (Richards et al., 2015). A combination of these codes will be used to classify each shortlisted variant. A literature search will be conducted to find previous reports of the identified variants.

4 Ethics

The ethics certificate numbers of the PeCan project are M180855 (pre-2024) and M240262 (2024- onwards). An application for this study was submitted to the University of Witwatersrand Human Research Ethics Committee (HREC) on the 15th of March 2024. It is currently pending, as the HREC will sit on the 19th of April 2024 to decide. This study will commence as soon as ethics approval is obtained, ideally by June.

5 Timing

	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb
Literature review	■	■											
Ethics application	■	■											
Preparing proposal		■	■										
Research protocol submission			■										
Ethics clearance					■								
Data analysis					■	■	■	■	■				
Final write-up and submission of research project									■	■			
Final presentation-research report defence													■

6 Funding

This research project is a bioinformatics analysis of pre-generated sequencing data from the PeCanproject; therefore no additional costs associated with this project. The funders for the PeCan project include the National Research Foundation (NRF), Female Academic Leaders Fellowship (FALF) and South African Medical Research Council (SAMRC).

7 Limitations

The use of WES limits the detection of complex and structural variants, which can only be detected through whole-genome sequencing (WGS). Next-generation sequencing technologies, including WES, can produce false positives due to sequencing errors, especially in regions with high GC content, repetitive sequences, or homopolymer runs. Variants of uncertain significance are complicated to interpret due to the lack of representation of African genomes on public databases and variants in genes that are outside the scope of the candidate gene list will be missed. Additionally, Sanger sequencing is a possible validation step that could be taken to ensure the reliability and accuracy of any findings. This validation is beyond the scope of the current study, due to time constraints, but could be explored in potential future studies under the PeCan project.

Pathogenic variants identified in this study will be fed back to the parent study and feedback will be given to the patients, this may inform genetic testing for at risk relatives. Patient feedback is beyond the scope of this study and will be performed by genetic counselors. The impact of this study is in its contribution to the knowledge of the germline genetic landscape of CNS tumours.

8 References

Alonso-Luna, O., Mercado-Celis, G.E., Melendez-Zajgla, J., Zapata-Tarres, M. and Mendoza-Caamal, E. (2023). 'The genetic era of childhood cancer: Identification of high-risk patients and germline sequencing approaches', *Annals of Human Genetics*, *87*(3), p.81-90. doi: https://doi.org/10.1111/ahg.12502

Capasso, M., Montella, A., Tirelli, M., Maiorino, T., Cantalupo, S. and Iolascon, A. (2020). 'Genetic predisposition to solid pediatric cancers', *Frontiers in Oncology*, 10. doi:10.3389/fonc.2020.590033.

Dahlin, A.M., Wibom, C., Andersson, U., Bybjerg-Grauholm, J., Deltour, I., Hougaard, D.M., Scheurer, M.E., Lau, C.C., McKean-Cowdin, R., Kennedy, R.J. and Hung, L.T. (2020). 'A genome-wide association study on medulloblastoma', *Journal of neuro-oncology*, *147*, p.309-315. doi: https://doi.org/10.1007/s11060-020-03424-9

Gröbner, S.N., Worst, B.C., Weischenfeldt, J., Buchhalter, I., Kleinheinz, K., Rudneva, V.A., Johann, P.D., Balasubramanian, G.P., Segura-Wang, M., Brabetz, S., Bender, S., Hutter, B., Sturm, D., Pfaff, E., Hübschmann, D., Zipprich, G., Heinold, M., Eils, J., Lawerenz, C., Erkek, S., Lambo, S., Waszak, S., Blattmann, C., Borkhardt, A., Kuhlen, M., Eggert, A., Fulda, S., Gessler, M., Wegert, J., Kappler, R., Baumhoer, D., Burdach, S., Kirschner-Schwabe, R., Kontny, U., Kulozik, A.E., Lohmann, D., Hettmer, S., Eckert, C., Bielack, S., Nathrath, M., Niemeyer, C., Richter, G.H., Schulte, J., Siebert, R., Westermann, F., Molenaar, J.J., Vassal, G., Witt, H., Zapatka, M., Burkhardt, B., Kratz, C.P., Witt, O., Tilburg, C.M.V., Kramm, C.M., Fleischhack, G., Dirksen, U., Rutkowski, S., Frühwald, M., Hoff, K. Von, Wolf, S., Klingebiel, T., Koscielniak, E., Landgraf, P., Koster, J., Resnick, A.C., Zhang, J., Liu, Y., Zhou, X., Waanders, A.J., Zwijnenburg, D.A., Raman, P., Brors, B., Weber, U.D., Northcott, P.A., Pajtler, K.W., Kool, M., Piro, R.M., Korbel, J.O., Schlesner, M., Eils, R., Jones, 8 D.T.W., Lichter, P.,

Chavez, L., Pfister, S.M. (2018). 'The landscape of genomic alterations across childhood cancers', *Nature*, *555*(7696), p.321-327. doi: https://doi.org/10.1038/nature25480.

Guerrini-Rousseau, L., Varlet, P., Colas, C., Andreiuolo, F., Bourdeaut, F., Dahan, K., Devalck, C., Faure-Conter, C., Genuardi, M., Goldberg, Y., Kuhlen, M., Moalla, S., Opocher, E., Perez-Alonso, V., Sehested, A., Slavc, S., Unger, S., Wimmer, K., Grill, J., Brugières, L. (2019). 'Constitutional mismatch repair deficiency–associated brain tumors: report from the European C4CMMRD consortium', *Neuro-oncology advances*, *1*(1), p.vdz033. doi: https://doi.org/10.1093/noajnl/vdz033

Flanagan, S.E., Patch, A.-M. and Ellard, S. (2010). 'Using SIFT and PolyPhen to Predict Loss-of-Function and Gain-of-Function Mutations', Genetic Testing and Molecular *Biomarkers*, 14(4), p.533–537. doi:https://doi.org/10.1089/gtmb.2010.0036.

Helfferich, J., Nijmeijer, R., Brouwer, O.F., Boon, M., Fock, A., Hoving, E.W., Meijer, L., den Dunnen, W.F. and de Bont, E.S. (2016). 'Neurofibromatosis type 1 associated low grade gliomas: a comparison with sporadic low grade gliomas', *Critical reviews in oncology/hematology*, *104*, p.30-41. doi: https://doi.org/10.1016/j.critrevonc.2016.05.008

Kline, C.N., Joseph, N.M., Grenert, J.P., van Ziffle, J., Talevich, E., Onodera, C., Aboian, M., Cha, S., Raleigh, D.R., Braunstein, S. and Torkildson, J. (2017). 'Targeted next-generation sequencing of pediatric neuro oncology patients improves diagnosis, identifies pathogenic germline mutations, and directs targeted therapy', *Neuro-oncology*, *19*(5), p.699-709. doi:https://doi.org/10.1093/neuonc/now254.

Landrum, M.J., Chitipiralla, S., Brown, G.R., Chen, C., Gu, B., Hart, J., Hoffman, D., Jang, W., Kaur, K., Liu, C. and Lyoshin, V. (2020). 'ClinVar: improvements to accessing data', *Nucleic acids research*, *48*(1), p.835-844. doi: https://doi.org/10.1093/nar/gkz972

Mbi Feh, M.K.N., Lyon, K.A., Brahmaroutu, A.V., Tadipatri, R. and Fonkem, E. (2021). 'The need for a central brain tumor registry in Africa: a review of central nervous system tumors in Africa from 1960 to 2017', *Neuro-oncology practice*, *8*(3), p.337-344. doi: https://doi.org/10.1093/nop/npaa086

Miabi, Z., Midia, M. and Midia, R. (2007). 'Evaluation of pediatric intracranial neoplasm in Tabriz Iran', *Acta Medica Iranica*, 45(6), p.443-446.

McLaren, W., Gil, L., Hunt, S.E., Riat, H.S., Ritchie, G.R.S., Thormann, A., Flicek, P. and Cunningham, F. (2016). 'The Ensembl Variant Effect Predictor', *Genome Biology, 17,* 0.1-14. doi:https://doi.org/10.1186/s13059-016-0974-4

Parsons, D.W., Roy, A., Yang, Y., Wang, T., Scollon, S., Bergstrom, K., Kerstein, R.A., Gutierrez, S., Petersen, A.K., Bavle, A. and Lin, F.Y. (2016). 'Diagnostic yield of clinical tumor and germline whole-exome sequencing for children with solid tumors', *JAMA oncology*, *2*(5), p.616-624. doi: https://doi.org/10.1001/jamaoncol.2015.5699.

Rentzsch, P., Schubach, M., Shendure, J. and Kircher, M. (2021). 'CADD-Splice—improving genome-wide variant effect prediction using deep learning-derived splice scores', *Genome medicine*, *13*, p.1-12. doi: https://doi.org/10.1186/s13073-021-00835-9

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W.W., Hegde, M., Lyon, E., Spector, E., Voelkerding, K., Rehm, H.L. and ACMG Laboratory Quality Assurance Committee (2015). 'Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology', *Genetics in medicine : official journal of the American College of Medical Genetics,* *17*(5), p.405-423. doi:https://doi.org/10.1038/gim.2015.30.

Robinson, J.T., Thorvaldsdottir, H., Turner, D. and Mesirov, J.P. (2022). 'igv.js: an embeddable JavaScript implementation of the Integrative Genomics Viewer (IGV)', *Bioinformatics*, 39(1), p.1076-1082. doi: https://doi.org/10.1093/bioinformatics/btac830.

Rotimi, S.O., Rotimi, O.A. and Salhia, B. (2021). 'A review of cancer genetics and genomics studies in Africa', *Frontiers in Oncology*, *10*, p.606400. doi: https://doi.org/10.3389/fonc.2020.606400.

Sait, S., Walsh, M.F. and Karajannis, M.A. (2021). 'Genetic syndromes predisposing to pediatric brain tumors', *Neuro-Oncology Practice*, *8*(4), p.375-390. doi: https://doi.org/10.1093/nop/npab012.

Waszak, S.M., Northcott, P.A., Buchhalter, I., Robinson, G.W., Sutter, C., Groebner, S., Grund, K.B., Brugières, L., Jones, D.T., Pajtler, K.W. and Morrissy, A.S. (2018). 'Spectrum and prevalence of genetic predisposition in medulloblastoma: a retrospective genetic study and prospective validation in a clinical trial cohort', *The Lancet Oncology*, *19*(6), p.785-798. doi: http://dx.doi.org/10.1016/S1470-2045(18)30242-0

Zhang, J., Walsh, M.F., Wu, G., Edmonson, M.N., Gruber, T.A., Easton, J., Hedges, D., Ma, X., Zhou, X., Yergeau, D.A., Wilkinson, M.R., Vadodaria, B., Chen, X., McGee, R.B., Hines-Dowell, S., Nuccio, R., Quinn, E., Shurtleff, S.A., Rusch, M., Patel, A., Becksfort, J.B., Wang, S., Weaver, M.S., Ding, L., Mardis, E.R., Wilson, R.K., Gajjar, A., Ellison, D.W., Pappo, A.S., Pui, C.-H., Nichols, K.E., Downing, J.R. (2015). 'Germline Mutations in Predisposition Genes in Pediatric Cancer', *New England Journal of Medicine*, *373*(24), p.2336-2346. doi: https://doi.org/10.1056/nejmoa1508054.

VISTASEQ brain/CNS/PNS Cancer Panel - clinical test - NIH genetic testing registry (GTR) - NCBI (2024) *National Center for Biotechnology Information*. https://www.ncbi.nlm.nih.gov/gtr/tests/GTR000556512/

Appendix E: Turnitin Report and Plagiarism Declaration

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

FACULTY OF HEALTH SCIENCES

PLAGIARISM DECLARATION TO BE SIGNED BY ALL HIGHER DEGREE STUDENTS

SENATE PLAGIARISM POLICY: APPENDIX ONE

I Chetana Chatergoon (Student number: 2687027) am a student registered for the degree of MSc (Med) Genomic Medicine in the academic year 2024.

I hereby declare the following:

- I am aware that plagiarism (the use of someone else's work without their permission and/or without acknowledging the original source) is wrong.
- I confirm that the work submitted for assessment for the above degree is my own unaided work except where I have explicitly indicated otherwise.
- I have followed the required conventions in referencing the thoughts and ideas of others.
- I understand that the University of the Witwatersrand may take disciplinary action against me if there is a belief that this is not my own unaided work or that I have failed to acknowledge the source of the ideas or words in my writing.
- I have included as an appendix a report from "Turnitin" (or other approved plagiarism detection) software indicating the level of plagiarism in my research document.

Signature: C Chatergoon Date: 29/11/24

MSc Final draft- Research in submissible format.docx

ORIGINALITY REPORT

20%	10%	10%	11%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to University of Witwatersrand Student Paper	8%
2	www.frontiersin.org Internet Source	1%
3	www.mdpi.com Internet Source	1%
4	www.ncbi.xyz Internet Source	1%
5	core.ac.uk Internet Source	1%
6	Lele Li, Xia Tian, Vaughan Woodzell, Richard A. Gibbs, Bo Yuan, Eric Venner. "Tracking updates in clinical databases increases efficiency for variant reanalysis", Genetics in Medicine Open, 2024 Publication	1%
7	Hunt, Callum. "Characterisation of Chromium Coated Zirconium Alloys for Accident Tolerant Fuel Claddings", The University of Manchester (United Kingdom), 2024	<1%

	Publication	
8	www.karger.com Internet Source	<1 %
9	hdl.handle.net Internet Source	<1 %
10	rcastoragev2.blob.core.windows.net Internet Source	<1 %
11	Pamela Salman, Sergio Panay, René Fernández, Mauricio Mahave, Cristian Soza-Ried. "Evidence of response to pembrolizumab in a patient with Lynch syndrome-related metastatic colon cancer", OncoTargets and Therapy, 2018 Publication	<1 %
12	Guerra, Joana Sofia Gonçalves. "POLE Somatic Mutations in Advanced Colorectal Cancer", Universidade do Porto (Portugal), 2024 Publication	<1 %
13	Michael J. Hall, Ryan Bernhisel, Elisha Hughes, Katie Larson et al. " Germline Pathogenic Variants in the Ataxia Telangiectasia Mutated ( ) Gene are Associated with High and Moderate Risks for Multiple Cancers ", Cancer Prevention Research, 2021 Publication	<1 %
14	herkules.oulu.fi Internet Source	<1 %

15	digitalassets.lib.berkeley.edu Internet Source	<1 %
16	Juan Mo, Stefanie L. Moye, Renee M. McKay, Lu Q. Le. "Neurofibromin and suppression of tumorigenesis: beyond the GAP", Oncogene, 2022 Publication	<1 %
17	bmcmedgenomics.biomedcentral.com Internet Source	<1 %
18	www2.mdpi.com Internet Source	<1 %
19	discovery.ucl.ac.uk Internet Source	<1 %
20	pesquisa.bvsalud.org Internet Source	<1 %
21	umu.diva-portal.org Internet Source	<1 %
22	unsworks.unsw.edu.au Internet Source	<1 %
23	Guffanti, Federica. "DNA Repair and Response to Therapy: Exploring Their Relation in Epithelial Ovarian Cancer Models", Open University (United Kingdom), 2021 Publication	<1 %

24	Jiang, Susiyan. "DNA Double Strand Breaks: To Repair, or Not to Repair.", New York Medical College Publication	<1 %
25	Anirban Das, Liana Nobre. "Genomics in pediatric high-grade gliomas: Hope or hype: Practical implications for resource-limited settings", Pediatric Hematology Oncology Journal, 2023 Publication	<1 %
26	F. Macdonald, C.H.J. Ford, A.G. Casson. "Molecular Biology of Cancer", CRC Press, 2004 Publication	<1 %
27	Submitted to Queen Mary and Westfield College Student Paper	<1 %
28	kipdf.com Internet Source	<1 %
29	Couto, Bárbara Incio Costa. "Reclassificação de Variantes de Significado Clínico Desconhecido no Estudo do Exoma Solicitado na Prática Clínica", Universidade do Porto (Portugal), 2024 Publication	<1 %
30	aiab.weebly.com Internet Source	<1 %

	Complementary Therapies in Clinical Practice, 2020 Publication	
39	wiredspace.wits.ac.za Internet Source	<1 %
40	www.uniprot.org Internet Source	<1 %
41	Mashu Futagawa, Tetsuya Okazaki, Eiji Nakata, Chika Fukano et al. "Genotypes and phenotypes of neurofibromatosis type 1 patients in Japan: A Hereditary Tumor Cohort Study", Human Genome Variation, 2024 Publication	<1 %
42	Melissa J. Reeves, Kerry E. Goetz, Bin Guan, Ehsan Ullah, Delphine Blain, Wadih M. Zein, Santa J. Tumminia, Robert B. Hufnagel. " Genotype-phenotype associations in a large -related retinopathy cohort ", Human Mutation, 2020 Publication	<1 %
43	Nguyen, Ngoc Minh Phuong. "New Genes and Mechanisms of Recurrent Hydatidiform Moles.", McGill University (Canada), 2020 Publication	<1 %
44	Starinsky-Elbaz, S.. "The pre-GAP-related domain of neurofibromin regulates cell migration through the LIM kinase/cofilin	<1 %

pathway", Molecular and Cellular Neuroscience, 200911
Publication

45	academic.oup.com Internet Source	<1 %
46	par.nsf.gov Internet Source	<1 %
47	www.biorxiv.org Internet Source	<1 %
48	www.omicsdi.org Internet Source	<1 %

Exclude quotes	On	Exclude matches	< 8 words
Exclude bibliography	On		

Appendix F: Author guidelines for Frontiers in Oncology

Author guidelines

Writing and formatting

Title

The title should be concise, omitting terms that are implicit and, where possible, be a statement of the main result or conclusion presented in the manuscript. Abbreviations should be avoided within the title.

Witty and creative titles are welcome, but only if relevant and within measure. Consider if a title meant to be thought-provoking might be misinterpreted as offensive or alarming. In extreme cases, the editorial office may veto a title and propose an alternative.

Authors should avoid:

- titles that are a mere question without giving the answer
- unambitious titles, for example starting with 'Towards,' 'A description of,' 'A characterization of' or 'Preliminary study on'
- vague titles, for example starting with 'Role of', 'Link between', or 'Effect of' that do not specify the role, link, or effect
- including terms that are out of place, for example the taxonomic affiliation apart from species name.

Abstract

As a primary goal, the abstract should make the general significance and conceptual advance of the work clearly accessible to a broad readership. The abstract should be no longer than a single paragraph and should be structured, for example, according to the **IMRAD format**. For the specific structure of the abstract, authors should follow the requirements of the article type or journal to which they're submitting. Minimize the use of abbreviations and do not cite references, figures or tables.

For clinical trial articles, please include the unique identifier and the URL of the publicly accessible website on which the trial is registered.

Manuscript length

We encourage you to closely follow the article word count lengths given in the 'Article types' page of the journals. The manuscript length includes only the main body of the text, footnotes, and all citations within it, and excludes the abstract, section titles, figure and table captions, funding statement, acknowledgments, and references in the bibliography. Please indicate the number of words and the number of figures and tables included in your manuscript on the first page.

Sections

The manuscript is organized by headings and subheadings. The section headings should be those appropriate for your field and the research itself. You may insert up to 5 heading levels into your manuscript (i.e.,: 3.2.2.1.2 Heading Title).

For Original Research articles, it is recommended to organize your manuscript in the following sections or their equivalents for your field.

Introduction
Succinct, with no subheadings.

Materials and methods
This section may be divided by subheadings and should contain sufficient detail so that when read in conjunction with cited references, all procedures can be repeated. For experiments reporting results on animal or human subject research, an ethics approval statement should be included in this section (for further information, see the 'Bioethics' section of our **policies and publication ethics**.)

Results
This section may be divided by subheadings. Footnotes should not be used and must be transferred to the main text.

Discussion
This section may be divided by subheadings. Discussions should cover the key findings of the study: discuss any prior research related to the subject to place the novelty of the discovery in the appropriate context, discuss the potential shortcomings and limitations on their interpretations, discuss their integration into the current understanding of the problem and how this advances the current views, speculate on the future direction of the research, and freely postulate theories that could be tested in the future.

For further information, please check the descriptions defined in the journal's 'Article types' page, in the 'For authors' menu on every journal page.

Language editing

Frontiers requires manuscripts submitted to meet international English language standards to be considered for publication.

For authors who would like their manuscript to receive language editing or proofreading to improve the clarity of the manuscript and help highlight their research, we recommend the language-editing services provided by the following external partners.

Note that sending your manuscript for language editing does not imply or guarantee that it will be accepted for publication by a Frontiers journal. Editorial decisions on the scientific content of a manuscript are independent of whether it has received language editing or proofreading by these partner services or other services.

Editage
We recommend the language-editing service provided by our external partner Editage. These services may be particularly useful for researchers for whom English is not the primary language. They can help to improve the grammar, syntax, and flow of your manuscript prior to submission. Frontiers' authors will receive a 10% discount by visiting the following link: **editage.com/frontiers**.

The Charlesworth Group
We recommend the Charlesworth Group's author services, who has a long-standing track record in language editing and proofreading. This is a third-party service for which Frontiers' authors will receive a 10% discount by visiting the following link: **www.cwauthors.com/frontiers**.

Language style

The default language style at Frontiers is American English. If you prefer your article to be formatted in British English, please specify this on the first page of your manuscript. For any questions regarding style, we recommend authors to consult the **Chicago Manual of Style**.

Inclusive language guidelines

Frontiers is an inclusive publisher and we ask that all submissions are in line with our inclusive language policy. When preparing your manuscript for submission, take a mindful approach towards personal biases and a concerted effort to limit their influence. Authors should remove any suggestion or implication of superiority or inferiority of one person over another based on age, gender, race, ethnicity, culture, sexual orientation, disability, religion, or socio-economic class. We ask authors to use inclusive language practices and awareness of diversity, equity, and inclusion into their research and keep it at the forefront during the composition of their findings.

External guidance that may be useful is available at C4DISC's **Guidelines on Inclusive Language and Images in Scholarly Communication**.

Furthermore, when drafting your work, please take into account the following considerations

In general, seek to avoid

- language that could be deemed insulting, profane, or derogatory.
- descriptors that identify personal attributes such as age, gender, race, ethnicity, culture, sexual orientation, disability, or health conditions, where they are not critically relevant to the discussion.
- any form of language that suggests a particular culture or group as the default or standard.

And where feasible:

- proactively ask individuals or groups how they would prefer to be referenced.
- adhere to the SAGER guidelines for reference to sex and gender in research.

Remember, the language we use can influence perceptions, evoke emotions, and shape perspectives. Let's work together to nurture an inclusive, respectful, and empowering discourse in science.

Guidelines for artificial intelligence and related technologies

These guidelines cover acceptable uses of generative AI technologies such as Large Language Models (ChatGPT, Jasper) and text-to-image generators (DALL-E 2, Midjourney, Stable Diffusion) in the writing or editing of manuscripts submitted to Frontiers.

Use of AI to create written or visual content

Authors should not list a generative AI technology as a co-author or author of any submitted manuscript. Generative AI technologies cannot be held accountable for all aspects of a manuscript and consequently do not meet the criteria required for authorship.

If the author of a submitted manuscript has used written or visual content produced by or edited using a generative AI technology, this use must follow all Frontiers guidelines and policies. Specifically, the author is responsible for checking the factual accuracy of any content created by the generative AI technology. This includes, but is not limited to, any quotes, citations or references. Figures produced by or edited using a generative AI technology must be checked to ensure they

accurately reflect the data presented in the manuscript. Authors must also check that any written or visual content produced by or edited using a generative AI technology is free from plagiarism.

If the author of a submitted manuscript has used written or visual content produced by or edited using a generative AI technology, such use must be acknowledged in the acknowledgements section of the manuscript and the methods section if applicable. This explanation must list the name, version, model, and source of the generative AI technology.
We encourage authors to upload all input prompts provided to a generative AI technology and outputs received from a generative AI technology in the supplementary files for the manuscript.

The entire document should be single-spaced and must contain page and line numbers in order to facilitate the review process. The manuscript should be written using either Word or LaTeX. See above for templates.

Abbreviations and nomenclatures

The use of abbreviations should be kept to a minimum. Non-standard abbreviations should be avoided unless they appear at least four times, and must be defined upon first use in the main text. Consider also giving a list of non-standard abbreviations at the end, immediately before the acknowledgments.

Equations should be inserted in editable format from the equation editor.

Italicize gene symbols and use the approved gene nomenclature where it is available. For human genes, please refer to the HUGO Gene Nomenclature Committee (**HGNC**). New symbols for human genes should be submitted to the HGNC **here**. Common alternative gene aliases may also be reported, but should not be used alone in place of the HGNC symbol. Nomenclature committees for other species are listed **here**. Protein products are not italicized.

We encourage the use of Standard International Units in all manuscripts.

Chemical compounds and biomolecules should be referred to using systematic nomenclature, preferably using the recommendations by the International Union of Pure and Applied Chemistry (IUPAC).

Astronomical objects should be referred to using the nomenclature given by the International Astronomical Union (IAU) provided **here**.

Life Science Identifiers (LSIDs) for ZOOBANK registered names or nomenclatural acts should be listed in the manuscript before the keywords. An LSID is represented as a uniform resource name (URN) with the following format: urn:lsid:<Authority>:<Namespace>:<ObjectID>[:<Version>]

For more information on LSIDs please see the 'Code' section of our **policies and publication ethics**.

Enhancing search engine optimization (SEO)

There are a few simple ways to maximize your article's discoverability and search results.

- Include a few of your article's keywords in the title of the article
- Do not use long article titles
- Pick 5-8 keywords using a mix of generic and more specific terms on the article subject(s)
- Use the maximum amount of keywords in the first two sentences of the abstract

- Use some of the keywords in level 1 headings

References

Preparing and formatting references

Submissions to Frontiers must be grounded in relevant and up to date peer-reviewed, academic research, and this should be reflected in the accompanying reference lists.

Authors are welcome to use online referencing tools in preparation of their manuscript. Some useful resources include **RefMe**, **Zotero**, and **Mendeley**.

- The citation of non-academic and non-peer-reviewed sources (e.g. blog posts, website content), as well as anonymous sources or commercial websites should be avoided or kept to a minimum
- Authors should avoid citing content that is not directly relevant to the scope of the article and the journal
- Reference lists should reflect the current status of knowledge in the field, avoid bias, and not include a high proportion of citations to the same authors or sources, school of thought, etc.
- The length of the reference list should be appropriate depending on the article type, covering the relevant literature through sufficient referencing
- Authors should ensure that references are accurate, that all links are accessible, and that the citations/references adhere to the reference styles outlined below

Frontiers' journals use one of two reference styles, either Harvard (author-date) or Vancouver (numbered). These formats should be adhered to for the in-text citations and the reference lists. Please check our **help center** to find the correct style for the journal to which you're submitting.

- All citations of published works in the text, figures, or tables must be in the reference list and vice-versa.
- The names of the first six authors followed by et al. and the DOI (when available) should be provided.
- Given names of authors should be abbreviated to initials (e.g. Smith, J., Lewis, C.S., etc.).
- The reference list should only include articles that are published or accepted.
- Unpublished data, submitted manuscripts, or personal communications should be cited within the text only, for article types that allow such inclusions. Where additional details are available, these will be included as footnotes.
- For accepted but unpublished works use 'in press' instead of page numbers.
- Data sets that have been deposited to an online repository should be included in the reference list. Include the version and unique identifier when available.
- Personal communications should be documented by a letter of permission.
- Website URLs should be included as footnotes.

- Any inclusion of verbatim text must be contained in quotation marks and should clearly reference the original source.
- Preprints can be cited provided that a DOI or archive URL is available, and the citation clearly mentions that the contribution is a preprint. If a peer-reviewed journal publication for the same preprint exists, the official journal publication is the preferred source. See the preprints section for each reference style below for more information.

Harvard reference style (author-date)

Reference examples for Frontiers' journals using the Harvard referencing system can be found below. For examples of other sources, and for general questions regarding the Harvard reference style, please refer to the **Chicago Manual of Style**.

- References should include the full last name and first name initials of the first six authors, followed by et al. and the year of publication in brackets.
- Alphabetical order is followed for the reference list.

Reference guidelines - Harvard

Source	Reference list entry	In-text citation
Article in a print journal	Sondheimer, N., and Lindquist, S. (2000). Rnq1: an epigenetic modifier of protein function in yeast. Mol. Cell. 5, 163-172.	(Sondheimer and Lindquist, 2000)
Article in an online journal	Tahimic, C.G.T., Wang, Y., Bikle, D.D. (2013). Anabolic effects of IGF-1 signaling on the skeleton. Front. Endocrinol. 4:6. doi: 10.3389/fendo.2013.00006	(Tahimic et al., 2013)
Article or chapter in a book	Sorenson, P. W., and Caprio, J. C. (1998). "Chemoreception," in The Physiology of Fishes, ed. D. H. Evans (Boca Raton, FL: CRC Press), 375-405.	(Sorenson and Caprio, 1998)
Book	Cowan, W. M., Jessell, T. M., and Zipursky, S. L. (1997). Molecular and Cellular Approaches to Neural Development. New York: Oxford University Press.	(Cowan et al., 1997)

Source	Reference list entry	In-text citation
Abstract	Hendricks, J., Applebaum, R., and Kunkel, S. (2010). A world apart? Bridging the gap between theory and applied social gerontology. Gerontologist 50, 284-293. Abstract retrieved from Abstracts in Social Gerontology database. (Accession No. 50360869)	(Hendricks et al., 2010)
Website	World Health Organization. (2018). E. coli. https://www.who.int/news-room/fact-sheets/detail/e-coli [Accessed March 15, 2018].	(World Health Organisation, 2018)
Patent	Marshall, S. P. (2000). Method and apparatus for eye tracking and monitoring pupil dilation to evaluate cognitive activity. U.S. Patent No 6,090,051. Washington, DC: U.S. Patent and Trademark Office.	(Marshall, 2000)
Data	Perdiguero P, Venturas M, Cervera MT, Gil L, Collada C. Data from: Massive sequencing of Ulms minor's transcriptome provides new molecular tools for a genus under the constant threat of Dutch elm disease. Dryad Digital Repository. (2015) http://dx.doi.org/10.5061/dryad.ps837	(Perdiguero et al., 2015)
Theses and dissertations	Smith, J. (2008) Post-structuralist discourse relative to phenomological pursuits in the deconstructivist arena. [dissertation/master's thesis]. [Chicago (IL)]: University of Chicago	(Smith, 2008)
Preprint	Smith, J. (2008). Title of the document. Preprint repository name [Preprint]. Available at: https://persistent-url (Accessed March 15, 2018).	(Smith, 2008)

Source	Reference list entry	In-text citation
One author	Hesse-Biber, S. (2010). Qualitative Approaches to Mixed Methods Practice. Qualitative Inquiry, 16(6), 455-468. https://doi.org/10.1177/1077800410364611	(Hesse-Biber, 2010)
Two authors	Duvail, S., Hamerlynck, O. (2007) The Rufiji River flood: plague or blessing?. Int. J. Biometeorol. 52, 33–42. https://doi.org/10.1007/s00484-007-0105-8	(Duvail and Hamerlynck, 2007)
More than two authors	Leemhuis C, Thonfeld F, Näschen K, Steinbach S, Muro J, Strauch A, López A, Daconto G, Games I, Diekkrüger B. (2017) Sustainability in the Food-Water-Ecosystem Nexus: The Role of Land Use and Land Cover Change for Water Resources and Ecosystems in the Kilombero Wetland, Tanzania. Sustainability. 9(9):1513. https://doi.org/10.3390/su9091513	(Leemhuis et al., 2017)
Same author(s); same year	Huang, X-G. (2016a). Electromagnetic fields and anomalous transports in heavy-ion collisions—a pedagogical review. Rep. Prog. Phys. 79 076302. doi: 10.1088/0034-4885/79/7/076302 Huang, X-G. (2016b). Simulating Chiral Magnetic and Separation Effects with Spin-Orbit Coupled Atomic Gases. Scientific Reports. 6:20601. doi: 10.1038/srep20601	(Huang, 2016a, 2016b)
Same author(s); different years	Sedrakian, A. (2007). The physics of dense hadronic matter and compact stars. Progress in Particle and Nuclear Physics. 58(1):168-246. doi: 10.1016/j.ppnp.2006.02.002 Sedrakian, A. (2016). Axion cooling of neutron stars. Phys. Rev. D 93:065044. doi. 10.1103/PhysRevD.93.065044	(Sedrakian, 2007, 2016)

Source	Reference list entry	In-text citation
Same first author; different author list	Quimque, M. T., Notarte, K. I., Letada, A., Fernandez, R. A., and Pilapil, D. Y. 4th., Pueblos, K.R., Agbay, J.C., Dahse, H.M., Wenzel-Storjohann, A., Tasdemir, D., Khan, A., Wei, D.Q., Gose Macabeo, A.P. (2021a). Potential Cancer- and Alzheimer's Disease-Targeting Phosphodiesterase Inhibitors fromUvaria alba: Insights from In Vitro and Consensus Virtual Screening. ACS Omega. 6, 8403–8417. doi: 10.1021/acsomega.1c00137 Quimque, M. T. J., Notarte, K. I. R., Fernandez, R. A. T.,Mendoza,M. A. O., Liman, R. A. D., Lim, J. A. K., et al. (2021b). Virtual screening-driven drug discovery of SARSCoV2 enzyme inhibitors targeting viral attachment, replication, post-translational modification and host immunity evasion infection mechanisms. J Biomol Struct Dyn. 39, 4316–4333. doi: 10.1080/07391102.2020.1776639	(Quimque et al., 2021a; Quimque et al., 2021b)
Different authors; same surname	Khan, S. M., Khan, M., Alouffi, A., Almutairi, M. M., Numan, M., Ullah, S., et al. (2023). Phylogenetic position of Haemaphysalis kashmirensis and Haemaphysalis cornupunctata, with Notes on Rickettsia spp. Genes. 14, 360. doi: 10.3390/genes14020360 Khan, Z., Shehla, S., Alouffi, A., Kashif Obaid, M., Zeb Khan, A., Almutairi, M. M., et al. (2022). Molecular survey and genetic characterization of Anaplasma marginale in ticks collected from livestock hosts in Pakistan. Animals. 12, 1708. doi: 10.3390/ani12131708	(Khan Z. et al., 2022; Khan S. M. et al., 2023)
Publishing in a Humanities and Social Sciences journal	Farrell, H. (2012). The Consequences of the Internet for Politics. Annual Review Political Science. 15, 35-52. http://dx.doi.org/10.1146/annurev-polisci-030810-110815	(Farrell, 2012, p. 40)
Personal communications		"We thank L. Li (personal communication,

Source	Reference list entry	In-text citation
		December, 2018) for noting this ambiguity.”

Vancouver reference style (numbered)

Reference examples for Frontiers’ journals using the Vancouver referencing system can be found below. For more examples of citing other documents and general questions regarding the Vancouver reference style, please refer to **Citing Medicine**.

- In-text citations in the Vancouver reference style should be numbered consecutively in order of appearance in the text and identified by Arabic numerals in parenthesis.
- Use square brackets for physics and mathematics articles.
- The abbreviation ‘Ref’ should not be used, e.g.: [e.g., (1)] should NOT read [e.g. Ref. (1)].
- Style for comparing a citation should follow the number format, e.g. [cf. (1)]. The same applies when using ‘see’, e.g. [see (46)].
- References should be numbered and listed chronologically, according to the order they appear in the text.

Reference guidelines - Vancouver

Source	Reference list entry	In-text citation
Article in a print journal	1. Sondheimer N, Lindquist S. Rnq1: an epigenetic modifier of protein function in yeast. Mol Cell. (2000) 5:163-172.	(1)
Article in an online journal	2. Tahimic CGT, Wang Y, Bikle DD. Anabolic effects of IGF-1 signaling on the skeleton. Front Endocrinol. (2013) 4:6. doi: 10.3389/fendo.2013.00006	(2)

Source	Reference list entry	In-text citation
Article or chapter in a book	3. Sorenson PW, Caprio JC. "Chemoreception". In: Evans DH, editor. The Physiology of Fishes. Boca Raton, FL: CRC Press (1998). p. 375-405.	(3)
Book	4. Cowan WM, Jessell TM, Zipursky SL. Molecular and Cellular Approaches to Neural Development. New York: Oxford University Press. (1997). p. 345.	(4)
Abstract	5. Christensen S, Oppacher F. An analysis of Koza's computational effort statistic for genetic programming. In: Foster JA, editor. Genetic Programming. EuroGP 2002: Proceedings of the 5th European Conference on Genetic Programming; 2002 Apr 3–5; Kinsdale, Ireland. Berlin: Springer. (2002). p. 182–91.	(5)
Website	6. World Health Organization. E. coli. (2018).https://www.who.int/news-room/fact-sheets/detail/e-coli[Accessed March 15, 2018].	(6)
Patent	7. Pagedas AC, inventor; Ancel Surgical R&D Inc., assignee. Flexible Endoscopic Grasping and Cutting Device and Positioning Tool Assembly. United States patent US 20020103498. (2002).	(7)
Data	8. Perdiguero P, Venturas M, Cervera MT, Gil L, Collada C. Data from: Massive sequencing of Ulms minor's transcriptome provides new molecular tools for a genus under the constant threat of Dutch elm disease. Dryad Digital Repository. (2015). http://dx.doi.org/10.5061/dryad.ps837	(8)

Source	Reference list entry	In-text citation
Theses and dissertations	9. Smith J. (2008) Post-structuralist discourse relative to phenomenological pursuits in the deconstructivist arena [dissertation/master's thesis]. Chicago (IL): University of Chicago. (2008).	(9)
Preprint	10. Kingma DP, Ba J. Adam: A method for stochastic optimization. arXiv [preprint]. (2014). Available at: https://arxiv.org/abs/1412.6980 (Accessed June 20, 2014).	(10)
Unpublished reference		"We thank L. Li (personal communication, December, 2018) for noting this ambiguity."
Named citation – One author	20. Hesse-Biber S. Qualitative Approaches to Mixed Methods Practice. Qualitative Inquiry. (2010) 16(6): 455-468. doi: 10.1177/1077800410364611	"Hesse-Biber (20) found…"
Named citation – Two authors	43. Duvail S, Hamerlynck O. The Rufiji River flood: plague or blessing?. Int J Biometeorol. (2007) 52:33–42. doi: 10.1007/s00484-007-0105-8	"… as stated by Duvail and Hamerlynck (43)."
Named citation – Multiple authors	44. Leemhuis C, Thonfeld F, Näschen K, Steinbach S, Muro J, Strauch A, et al. Sustainability in the Food-Water-Ecosystem Nexus: The Role of Land Use and Land Cover Change for Water Resources and Ecosystems in the Kilombero Wetland, Tanzania. Sustainability. (2017) 9(9):1513. doi: 10.3390/su9091513	"In research conducted by Leemhuis et al. (44)..."

Source	Reference list entry	In-text citation
Multiple citations	8. Yuan S, Yao X, Yang H, Zhang Y, Liu H, Sun J, et al. Research note: genetic diversity of duck circoviruses circulating in partial areas of Guangdong province, southern China. Poult Sci. (2022) 101:102032. doi: 10.1016/j.psj.2022.102032 … 29. Liu H, Li LX, Sun WC, Shi N, Sun XT, Jin NY, et al. Molecular survey of duck circovirus infection in poultry in southern and southwestern China during 2018 and 2019. BMC Vet Res. (2020) 16:80. doi: 10.1186/s12917-020-02301-x 30. Zhang X, Jiang S, Wu J, Zhao Q, Sun Y, Kong Y, et al. An investigation of duck circovirus and co-infection in Cherry Valley ducks in Shandong Province, China. Vet Microbiol. (2009) 133:252–6. doi: 10.1016/j.vetmic.2008.07.005	"… reported in these regions (8, 29, 30), are…"
Publishing in physics or mathematics	1. Lee YT, Vardi A, Tordjman M. A hybrid self-aligned MIS-MESFET architecture for improved diamond-based transistors. Appl Phys Lett (2020) 117:202101. doi:10.1063/5.0023662	"…including transistors [1]…"

Citation of retracted articles

Authors are responsible for ensuring the integrity of the references they cite, including verifying whether any article they intend to cite has been retracted. Citing retracted articles can undermine the credibility of your manuscript and should only occur where it is essential for the context of your discussion.

Guidelines for citing retracted articles:

1. **Verification:** Authors are responsible for checking the status of all cited works before submission. Lists of retracted articles can often be found in databases such as PubMed and the Retraction Watch database. Authors can check the version of the record which is expected to include a visible notice of retraction.
2. **Disclosure:** If it is necessary to cite a retracted article, authors must clearly indicate the retraction status in the reference list.
3. **Justification:** Authors are encouraged to discuss the reasons for citing the retracted article in the manuscript, explaining its relevance and the implications of its retracted status on the discussion.

Adhering to these guidelines is considered best practice as it helps maintain the scientific integrity of the published literature and ensures that all references within the manuscript contribute validly to the scholarly discourse. Ensuring the accuracy and transparency of citations protects authors from

inadvertently perpetuating misinformation and supports the overall trustworthiness of research publications.

Acknowledgements

This is a short text to acknowledge the contributions of specific colleagues, institutions, or agencies that aided the efforts of the authors. Should the content of the manuscript have previously appeared online, such as in a thesis or preprint, this should be mentioned here, in addition to listing the source within the reference list.

Acknowledgements are designed to recognize individuals directly involved in the research. We request that authors avoid personal expressions of gratitude toward figures or entities unrelated or indirectly related to the production of the manuscript. While such acknowledgements are understood to be significant to the authors, they are considered indirectly involved in the research unless they pertain to specific research activities. It remains the responsibility of the authors to ensure that these guidelines are adhered to in the final publication.

Templates

If working with Word please use our **Word templates**. If you wish to submit your article as LaTeX, we recommend our **LaTeX templates**.

For LaTeX files, please ensure all relevant manuscript files are uploaded: .tex file, PDF, and .bib file (if the bibliography is not already included in the .tex file).

During the **interactive review**, authors are encouraged to upload versions using track changes. Editors and reviewers can only download the PDF file of the submitted manuscript.

Figures, tables, and images

Figures, tables, and images: rights and permissions

All figures, tables, and images will be published under a Creative Commons **CC-BY license**, and permission must be obtained for use of copyrighted material from other sources (including re-published/adapted/modified/partial figures and images from the internet). It is the responsibility of the authors to acquire the licenses, follow any citation instructions requested by third-party rights holders, and cover any supplementary charges.

For additional information, please see the 'Image manipulation' section of our **policies and publication ethics**.

Figures and images: style guidelines

We require figures to be submitted individually, in the same order as they are referred to in the manuscript; the figures will then be automatically embedded at the end of the submitted manuscript. Ensure that each figure is mentioned in the text and in numerical order.

For figures with more than one panel, panels should be clearly indicated using labels (A), (B), (C), (D), etc. However, do not embed the part labels over any part of the image. These labels will be replaced during typesetting according to Frontiers' journal style. For graphs, there must be a self-explanatory label (including units) along each axis.

For LaTeX files, figures should be included in the provided PDF. In case of acceptance, our production office might require high-resolution files of the figures included in the manuscript in EPS, JPEG or TIF/TIFF format.

To upload more than one figure at a time, save the figures (labeled in order of appearance in the manuscript) in a zip file and upload them as 'Supplementary material presentation.'

Please note that figures not in accordance with the guidelines will cause substantial delay during the production process.

Captions

Captions should be preceded by the appropriate label, for example 'Figure 1.' Figure captions should be placed at the end of the manuscript. Figure panels are referred to by bold capital letters in brackets: (A), (B), (C), (D), etc.

Image size and resolution requirements

Figures should be prepared with the PDF layout in mind. Individual figures should not be longer than one page and with a width that corresponds to one column (85 mm) or two columns (180 mm).

All images must have a resolution of 300 dpi at final size. Check the resolution of your figure by enlarging it to 150%. If the image appears blurry, jagged, or has a stair-stepped effect, the resolution is too low.

The text should be legible and of high quality. The smallest visible text should be no less than eight points in height when viewed at actual size.

Solid lines should not be broken up. Any lines in the graphic should be no smaller than two points wide.

Please note that saving a figure directly as an image file (JPEG, TIF) can greatly affect the resolution of your image. To avoid this, one option is to export the file as PDF, then convert into TIFF or EPS using a graphics software.

Format and color image mode

The following formats are accepted: TIF/TIFF (.tif/.tiff), JPEG (.jpg), and EPS (.eps) (upon acceptance). Images must be submitted in the color mode RGB.

Images of chemical structures

Chemical structures should be prepared using ChemDraw or a similar program. If working with ChemDraw please use our **ChemDraw template**. If working with another program please follow the guidelines below.

- Drawing settings: chain angle, 120° bond spacing, 18% width; fixed length, 14.4 pt; bold width, 2.0 pt; line width, 0.6 pt; margin width, 1.6 pt; hash spacing, 2.5 pt. Scale 100% Atom Label settings: font, Arial; size, 8 pt
- Assign all chemical compounds a bold, Arabic numeral in the order in which the compounds are presented in the manuscript text.

Table requirements and style guidelines

Tables should be inserted at the end of the manuscript in an editable format. If you use a word processor, build your table in Word. If you use a LaTeX processor, build your table in LaTeX. An empty line should be left before and after the table.

Table captions must be placed immediately before the table. Captions should be preceded by the appropriate label, for example 'Table 1.' Please use only a single paragraph for the caption.

Ensure that each table is mentioned in the text and in numerical order.

Large tables covering several pages cannot be included in the final PDF for formatting reasons. These tables will be published as supplementary material.

Tables which are not according to the above guidelines will cause substantial delay during the production process.

During production, tables will be formatted according to Frontiers' house style. Here is an example of a formatted table.

This is a heading	This is a heading
This is a subheading	
This is an example text	This is an example text
This is an example text	This is an example text
This is an example text	This is an example text

Accessibility

We encourage authors to make the figures and visual elements of their articles accessible for the visually impaired. Effective use of color can help people with low visual acuity, or color blindness, understand all the content of an article.

These guidelines are easy to implement and are in accordance with the W3C Web Content Accessibility Guidelines (**WCAG 2.1**), the standard for web accessibility best practices.

Ensure sufficient contrast between text and its background
People who have low visual acuity or color blindness could find it difficult to read text with low contrast background color. Try using colors that provide maximum contrast.

WC3 recommends the following contrast ratio levels:

- Level AA, contrast ratio of at least 4.5:1
- Level AAA, contrast ratio of at least 7:1

Supplementary material

Selecting supplementary material

Data that are not of primary importance to the text, or which cannot be included in the article because they are too large or the current format does not permit it (such as videos, raw data traces, and PowerPoint presentations), can be uploaded as supplementary material during the submission procedure and will be displayed along with the published article. All supplementary files are deposited to figshare for permanent storage and receive a DOI.

Supplementary material is not typeset, so please ensure that all information is clearly presented without tracked changes/highlighted text/line numbers, and the appropriate caption is included in the file. To avoid discrepancies between the published article and the supplementary material, please do not add the title, author list, affiliations or correspondence in the supplementary files.

File type requirements

The supplementary material can be uploaded as:

- data sheet (Word, Excel, CSV, CDX, FASTA, PDF or Zip files)
- presentation (PowerPoint, PDF or Zip files)
- image (CDX, EPS, JPEG, PDF, PNG or TIF/TIFF),
- table (Word, Excel, CSV or PDF)
- audio (MP3, WAV or WMA)
- video (AVI, DIVX, FLV, MOV, MP4, MPEG, MPG or WMV).

Technical requirements for supplementary images:

- 300 DPIs
- RGB color mode.

For supplementary material templates (LaTeX and Word), see our **supplementary material templates**.

Submitting information

Choosing where to submit

Open access and copyright

All Frontiers' articles from July 2012 onwards are published with open access under the Creative Commons **CC-BY license** (the current version is CC-BY, version 4.0). This means that the author(s) retains copyright, but the content is free to download, distribute, and adapt for commercial or non-commercial purposes, given appropriate attribution to the original article.

Upon submission, the author(s) grants Frontiers a license to publish, including to display, store, copy, and reuse the content. The CC-BY Creative Commons attribution license enables anyone to use the publication freely, given appropriate attribution to the author(s) and citing Frontiers as the original publisher. The CC-BY Creative Commons attribution license does not apply to third-party materials that display a copyright notice to prohibit copying. Unless the third-party content is also subject to a CC-BY Creative Commons attribution license, or an equally permissive license, the author(s) must comply with any third-party copyright notices.

Registration with Frontiers

Corresponding and all submitting authors must **register** with Frontiers before submitting an article. You must be logged in to your personal Frontiers account to submit an article.

For any co-author who would like their name on the article abstract page and PDF to be linked to a Frontiers profile on the **Loop network**, please **register** before the final publication of the paper.

Article type

We require authors to select the appropriate article type for their manuscript and to comply with the article type descriptions defined in the journal's 'Article types' page, which can be found under the 'About journal' menu in 'For authors' on every Frontiers journal page. Please pay close attention to the word count limits.

Keywords

All article types require a minimum of five and a maximum of eight keywords.

CrossMark policy

CrossMark is a multi-publisher initiative to provide a standard way for readers to locate the current version of a piece of content. By applying the CrossMark logo Frontiers is committed to maintaining the content it publishes and to alerting readers to changes if and when they occur.

Clicking on the CrossMark logo will tell you the current status of a document and may also give you additional publication record information about the document.

For Corrigenda, General Commentaries, and Editorials, the title of your manuscript should have the following format.

- 'Corrigendum: [Title of original article]'
- General Commentaries:
  'Commentary: [Title of original article]'
  'Response: Commentary: [Title of original article]'
- 'Editorial: [Title of Research Topic]'

Authors and affiliations

All names are listed together and separated by commas. Provide exact and correct author names as these will be indexed in official archives. Affiliations should be keyed to the author's name with superscript numbers and be listed as follows:

- Laboratory, Institute, Department, Organization, City, State abbreviation (only for United States, Canada, and Australia), and Country (without detailed address information such as city zip codes or street names).

Example: Max Maximus1
1 Department of Excellence, International University of Science, New York, NY, United States.

Correspondence

The corresponding author(s) should be marked with an asterisk in the author list. Provide the exact contact email address of the corresponding author(s) in a separate section.

Example: Max Maximus*
maximus@iuscience.edu

If any authors wish to include a change of address, list the present address(es) below the correspondence details using a unique superscript symbol keyed to the author(s) in the author list.

Equal contributions

The authors who have contributed equally should be marked with a symbol (†) in the author list of the doc/latex and pdf files of the manuscript uploaded at submission.

Please use the appropriate standard statement(s) to indicate equal contributions:

- **Equal contribution:** These authors contributed equally to this work
- **First authorship:** These authors share first authorship
- **Senior authorship:** These authors share senior authorship
- **Last authorship:** These authors share last authorship
- **Equal contribution and first authorship:** These authors contributed equally to this work and share first authorship
- **Equal contribution and senior authorship:** These authors contributed equally to this work and share senior authorship
- **Equal contribution and last authorship:** These authors contributed equally to this work and share last authorship

Example: Max Maximus 1†, John Smith2† and Barbara Smith1
†These authors contributed equally to this work and share first authorship

Consortium/group and collaborative authors

Consortium/group authorship should be listed in the manuscript with the other author(s).

In cases where authorship is retained by the consortium/group, the consortium/group should be listed as an author separated by a comma or 'and'. The consortium/group name will appear in the author list, in the citation, and in the copyright. If provided, the consortium/group members will be listed in a separate section at the end of the article.

For the collaborators of the consortium/group to be indexed in PubMed, they do not have to be inserted in the Frontiers submission system individually. However, in the manuscript itself, provide a section with the name of the consortium/group as the heading followed by the list of collaborators, so they can be tagged accordingly and indexed properly.

Example: John Smith, Barbara Smith and The Collaborative Working Group.

In cases where work is presented by the author(s) on behalf of a consortium/group, it should be included in the author list separated with the wording 'for' or 'on behalf of.' The consortium/group will not retain authorship and will only appear in the author list.

Example: John Smith and Barbara Smith on behalf of The Collaborative Working Group.

These guidelines cover acceptable uses of generative AI technologies such as Large Language Models (ChatGPT, Jasper) and text-to-image generators (DALL-E 2, Midjourney, Stable Diffusion) in the writing or editing of manuscripts submitted to Frontiers.

Scope statement

When you submit your manuscript, you will be required to summarize in 200 words your manuscript's scope and its relevance to the journal and/or specialty section you're submitting to. The aim is to convey to editors and reviewers how the contents of your manuscript fit within the selected journal's scope.

This statement will not be published with your article if it is accepted for publication. The information will be used during the initial validation and review processes to assess whether the manuscript is a suitable fit for the chosen journal and specialty.

We encourage you to consider carefully where to submit your manuscript, as submissions to an unsuitable journal or specialty will result in delays and increase the likelihood of manuscript rejection.

If you are submitting to a Research Topic, please also clarify how your submission is suited to the specific topic.